Modeling in a Real and Complex World

United States Military Academy

DEPARTMENT OF THE ARMY

UNITED STATES MILITARY ACADEMY

West Point, New York 10996-1786

MADN-MATH July 2, 2008

MEMORANDUM FOR MA103 Students, Department of Mathematics, USMA

SUBJECT: MA 103 Program Director's Memorandum

1. MA103 is a 4.0 credit hour course and therefore, does not follow the Day 1 / Day 2 schedule of a 3.0 credit hour course. The course calendar can be found by navigating through the Course Admin link on the MA103 website (http://www.dean.usma.edu/math/courses/MA103/index.htm). In this course, you have an opportunity to learn what mathematical modeling is and how powerful the mathematical modeling you learn can be. You may or may not become a mathematician or an engineer. But your studies in this course will enhance your ability to think critically, a skill that can and will help you in any field or profession. The material that you will study in MA103 is organized into five blocks:
    a. Problems solving and modeling with continuous functions (Block I)
    b. Modeling with Discrete Dynamical Systems (Block II)
    c. Matrix operations and solutions to systems of equations (Block III)
    d. Modeling with systems of Discrete Dynamical Systems (Block IV)
    e. Continuous change (Block V)

2. Through hard work in each of the blocks above, you will become more successful students and learners in the following five areas:

a. Base of knowledge. You will learn several problem solving techniques in order to formulate and structure powerful mathematical models that can help you do many things. In particular, the experience gained in developing mathematical models in this course will help you answer many interesting questions like what savings plan do I need to begin in order to become a millionaire by a desired age, and how can I manipulate an image on my computer like those in many high tech movies.

b. Technology. You will have numerous opportunities to use powerful software programs to enhance your capability to investigate possible solutions of the mathematical models that you develop in the course. Specifically, you will become competent in the basic commands of a computer algebra system (Mathematica) and of a computer spreadsheet (Excel) in order to make important predictions about things in every day life that you are concerned about.

c. Communication. Being a good communicator is one of the most important characteristics of being a great leader. The fundamentals in successfully conveying how you want the troops under your command to perform an essential task and in describing your thought process in solving a mathematics problem are the same. All leaders must be able to clearly articulate their thoughts. You will have many opportunities to improve your communication skills both verbally and in writing. These opportunities include board presentations, various writing assignments, and the preparation of a technical report.

d. Confident and competent problems solvers. You will develop modeling and problem solving abilities through in-class experiences, homework exercises, and a group project. These events will require you to analyze real world problems, make critical assumptions, model the problem, solve the model, and then interpret your results. Being able to do these things will help you in becoming more confident and competent solvers of all types of problems.

e. Develop habits of mind. Some key components of habits of mind that you will become better at are: creativity, work ethic, thinking interdependently, critical thinking, lifelong learning, and curiosity. You can certainly reach a higher potential if all of the elements are incorporated and pursued simultaneously. Strategies will be implemented in this course that promote and develop each of these for you. Ultimately, you will be introduced to the importance of life long learning and will be encouraged to learn how you best learn and to develop good study habits.

3. You must take responsibility for your own learning and participate as an active learner. To realize the goals above, you must do several things:

- Success in this course depends heavily on your daily preparation. Dedicate the time required for success – we have designed this course so that the average student can succeed with between 1-2 hours of daily preparation. If you habitually prepare less than this, your understanding and performance may measurably suffer.
- Come prepared for class with worked or attempted problems, understanding, and questions. Come to class knowing what you don't know so that you can ask questions. Unless otherwise directed by your instructor, you are responsible for the assigned readings and problems prior to coming to class.
- Participate in the instruction and discussion – this is *your* education; take charge!
- Seek assistance when needed – from the text, your classmates, or your instructor. We have some of the most professional and caring instructors teaching our course. Although they want you to succeed, they cannot learn for you.

4. Course Evaluation Plan: Your performance in this course will be evaluated both in and out of class. Out of class efforts consist of homework assignments and projects. In-class assessments consist of written exercises, presentations, Written Partial Reviews (WPRs), and a comprehensive final examination. To evaluate your progress in reaching the goals in paragraph 3 (and to provide you with feedback on your learning), we will have the following assessments:

Event	**Points**	**Percentage**
2 Written Partial Reviews (WPR)	500	25.0 %
2 Course Wide Quizzes	250	12.5%
1 Term End Examination (TEE)*	500	25.0 %
1 Fundamental Concepts Exam	100	5.0 %
1 Project	200	10.0 %
1 Course wide homework	100	5.0 %
Instructor Points	350	17.5 %
Total Points	**2000**	**100 %**

*A score of less than 50% on the TEE (regardless of final course average) could result in course failure.

8. I think you will find this course interesting, very applicable, and enjoyable. On behalf of all the instructors that teach this course, welcome to MA103, your first course in the Department of Mathematical Sciences. Good luck this semester!

//original signed//
GERALD C. KOBYLSKI
LTC, EN
Program Director

THIS PAGE INTENTIONALLY LEFT BLANK

Course Introduction

The purpose of the USMA Department of Mathematical Sciences core program is to provide each cadet a broad mathematical education emphasizing intellectual discipline, mastery of reasoning, practical applications, and the role of mathematics in society.
Our goal is a mathematics program that develops habits of mind for effective problem solving by applying mathematical knowledge to formulate and validate while leveraging the power of technology to calculate and investigate. The course you are enrolled in presents a variety of mathematical tools to help you critically evaluate a problem to come to a logical conclusion. Specifically, the goals of the course are focused in five areas; you will revisit each of these five areas in all four math course you take at USMA.

- Base of knowledge. You will learn several problem solving techniques in order to formulate and structure powerful mathematical models that can help you do many things. In particular, the experience gained in developing mathematical models in this course will help you answer many interesting questions like what savings plan do I need to begin in order to become a millionaire by a desired age, and how can I manipulate an image on my computer like those in many high tech movies.

- Technology. You will have numerous opportunities to use powerful software programs to enhance your capability to investigate possible solutions of the mathematical models that you develop in the course. Specifically, you will become competent in the basic commands of a computer algebra system (Mathematica) and of a computer spreadsheet (Excel) in order to make important predictions about things in every day life that you are concerned about.

- Communication. Being a good communicator is one of the most important characteristics of being a great leader. The fundamentals in successfully conveying how you want the troops under your command to perform an essential task and in describing your thought process in solving a mathematics problem are the same. All leaders must be able to clearly articulate their thoughts. You will have many opportunities to improve your communication skills both verbally and in writing. These opportunities include board presentations, various writing assignments, and the preparation of a technical report.

- Confident and competent problems solvers. You will develop modeling and problem solving abilities through in-class experiences, homework exercises, and a group project. These events will require you to analyze real world problems, make critical assumptions, model the problem, solve the model, and then interpret your results. Being able to do these things will help you in becoming more confident and competent solvers of all types of problems.

- Develop habits of mind. Some key components of habits of mind that you will become better at are: creativity, work ethic, thinking interdependently, critical thinking, lifelong learning, and curiosity. You can certainly reach a higher potential if all of the elements are incorporated and pursued simultaneously. Strategies will be implemented in this course that promote and develop each of these for you. Ultimately, you will be introduced to the importance of life long learning and will be encouraged to learn how you best learn and to develop good study habits.

You must take responsibility for your own learning and participate as an active learner. To realize the goals above, you must do several things:

• Success in this course depends heavily on your daily preparation. Dedicate the time required for success – we have designed this course so that the average student can succeed with between 1-2 hours of daily preparation. If you habitually prepare less than this, your understanding and performance may measurably suffer.

• Come prepared for class with worked or attempted problems, understanding, and questions. Come to class knowing what you don't know so that you can ask questions. Unless otherwise directed by your instructor, you are responsible for the assigned readings and problems prior to coming to class.

• Participate in the instruction and discussion – this is *your* education; take charge!

• Seek assistance when needed – from the text, your classmates, or your instructor. We have some of the most professional and caring instructors teaching our course. Although they want you to succeed, they cannot learn for you.

Since the material covered in this course spans many topics from data fitting to the beginnings of Calculus, no single text is currently available to comprehensively cover the material. Professors in the USMA Department of Mathematical Sciences have written this reference designed specifically for this course. This text is also posted by lesson on the MA103 website at http://www.dean.usma.edu/math/courses/MA103/index.htm. You will want to use the online text frequently in order to access the many interactive websites we have developed for you. One of our goals in writing this text is to motivate you to read and understand the material in the textbook. With this in mind, this text is written in a very casual tone that we hope you will enjoy and appreciate.

The following professors have contributed to the development of this text:

Amanda Beecher	Peter Charbonneau
Amy H. Erickson	Andrew Glen
Alex Heidenberg	Michelle Isenhour
Heather Jackson	Gerald Kobylski
Joseph Lindquist	Shawn McMurran
Kerry Moores	Jack Picciuto
Jonathan Roginski	Don Small
Frank Wattenberg	

You have also received the text Calculus Early Transcendentals by James Stewart. You will use this text during the course both as a primary and secondary reference in order to make connections to Calculus throughout the course. This text will be the primary reference in MA104 Differential Calculus and MA205 Integral Calculus, your next two courses in the math sequence at USMA.

Chapter 1

Mathematical Modeling

1.1 Introduction

1.1.1 A Mathematical Model

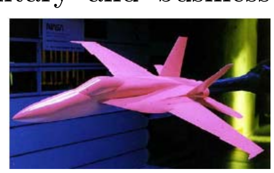
This course and book focus on the use of the power of mathematics to solve real world problems that are important and often urgent. The central concept of this book is *mathematical modeling*. A discussion of mathematical modeling would be incomplete without first defining what a mathematical model is. A **mathematical model** is a construct (e.g., a function or equation) that is designed to predict the behavior of a system. Because assumptions usually must be made, a model usually is an idealization of the actual system it represents. Certainly, all models have limitations. In fact, Dr. George Box, one of the most prominent statisticians of the 20th century, is credited with having said "all models are wrong – but some are useful."[1]

> A **mathematical model** is a construct (e.g., a function or equation) that is designed to predict the behavior of a system.

Mathematical models can be used to solve problems in many different kinds of situations. For example, they can be used to predict how a population may grow, estimate life expectancy, or determine how much a manufacturer should charge for a product.

It is possible to use mathematical models to represent physical phenomena. For example, many people enjoy playing video games, such as Super Swing Golf, available for the Nintendo Wii™ gaming system. The game's software detects your movements and uses a mathematical model that may consider such factors as the speed of your swing, your hand-eye coordination, the composition of the ball and club you selected, the density of the air, wind speed, the Coriolis Effect, and many other factors to represent the movement in your living room on the game's console.

Models are not only good for games; they are used in the military and business worlds to increase safety and save both time and money. Aircraft designers use mathematical models to predict how design modifications may affect handling and aerodynamic properties. The use of these models limits the amount of money the manufacturer has to spend on damaged aircraft and pilot flight hours. More importantly, pilot and crew exposure to dangerous situations is decreased.



[1]Box, George. E. P. Robustness in the Strategy of Scientific Model Building. In R. L. Launer, and G. N. Wilkinson, (eds.) Robustness in Statistics. New York: Academic Press

1.1.2 Problem Solving Processes

As problem solvers, we all have our own different levels of expertise and experience. Many problems appear complex or unsolvable at first glance – and perhaps some are. We can, however, always gain greater insight about a problem. But how? What are the steps? Where is the template that you can apply? There are many well-known methods or processes to solve problems. Perhaps you recall some type of process that you used in high school to solve problems. Maybe you have been using a method to tackle a problem, but didn't call it a process. Our minds often break up a problem into manageable steps to help us arrive at a logical conclusion. A problem solving process can help structure our thoughts as we solve the problem. Using a process can help ensure that we do not leave out crucial information and that we do consider all alternatives.

A problem solving process often used in the military is known as the Military Decision Making Process (MDMP). You will see more of it in your USMA military science courses as well as many future military courses during your career. The MDMP (Figure 1.1) illustrates the way military officers go from receiving a mission from a higher authority to execution of the mission at their level.

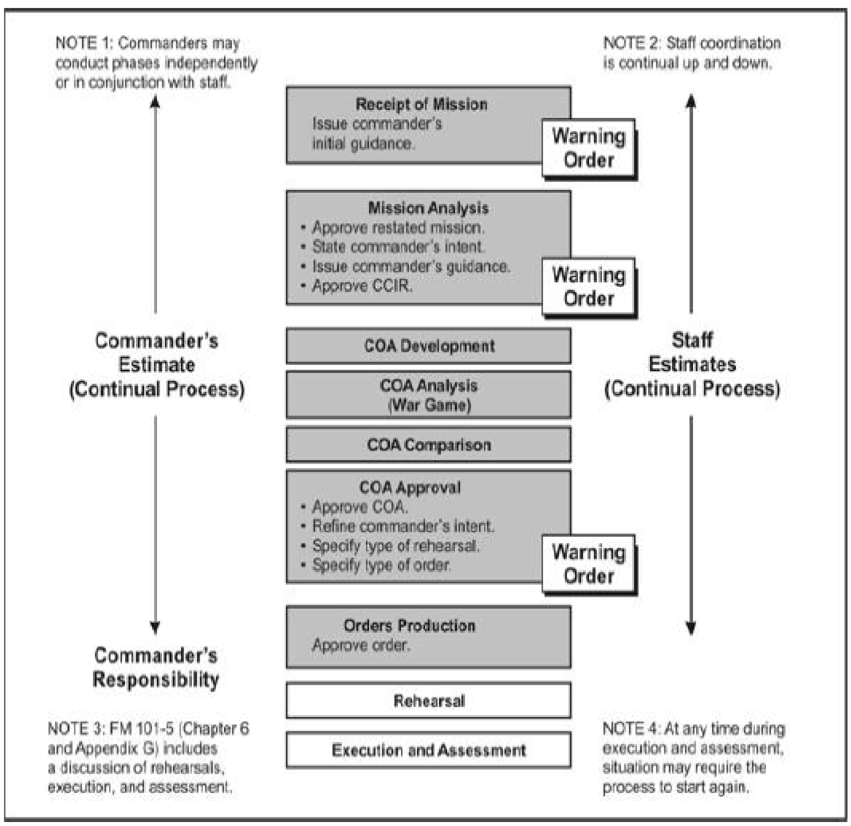


Figure 1.1: Military Decision Making Process

You will encounter another problem solving process next year in physics, shown in Table 1.1.

STEPS	PROBLEM-SOLVING FORMAT
READ	**1. FOCUS THE PROBLEM.** Gain a qualitative understanding of the problem, i.e., find out what is known and what is to be determined. a. Read the problem carefully, noting key words and clues.
DRAW	b. Visualize the situation and draw a sketch
GIVEN	c. Write down the given information.
ASSUMPTION(S)	d. Write down the assumptions that are stated, implied, or necessary to solve the problem.
TASK(S)	e. Write down what you have been tasked to do.
PHYSICS	**2. PLAN A SOLUTION.** Identify the fundamental physics concepts and associated diagrams required to solve this problem. Break the problem into parts (as necessary). Associate each part with a fundamental physics concept. Select the mathematical model that you will use to describe the concept.
SOLVE	**3. EXECUTE THE SOLUTION.** Develop a solution using the physics equations from the Physics Reference Card (PRC) and substitute the known quantities as necessary to arrive at a final numerical solution. a. Draw diagrams for the parts as appropriate. b. Use the physics equations associated with the fundamental physics concept(s) and logically and progressively develop a symbolic solution. For multiple part problems, this may require substituting quantities from one part into another part. c. Substitute known numerical quantities to arrive at a final numerical solution.
CHECK	**4. EVALUATE THE RESULT.** Ensure that you have answered the original task. a. Check for a sensible answer. b. Check for correct units and number of significant figures. c. Check for the correct direction for vector quantities.

Table 1.1: Physics Problem Solving Process

The problem solving process that we will highlight combines the process developed by Dr. George Polya in his book, How to Solve it, and the process developed by the USMA Chemistry Department, known as Given, Find, Plan, Solve (GFPS). This process is described below.

1. **Given.** Understanding what is given in the problem is the first step in solving the problem.
   - What is known?
   - What is unknown?
   - Define the variables of interest.
   - What assumptions must be made about the unknowns to solve the problem? Remember: all assumptions must be **valid** and **necessary**.
     - Valid: The assumption is accurate (e.g., assuming that a parabola would be an appropriate model to describe the path of a punted football).
     - Necessary: The assumption is required for the problem (e.g., not "the golf ball is orange" if we are determining how far Tiger Woods can drive a ball).
   - Visualize the situation and draw a picture, if possible.
   - Identify the units to be used throughout the problem.

2. **Find.**
   - What do we need to *find* to solve the problem?

3. **Plan.** Develop an idea of what solution techniques will be most appropriate in the problem and how you will apply them.
   - Have you worked this problem or a similar problem before? Your previous solution process may help this time, too.
   - If necessary, break the problem into smaller parts.
   - Ensure you account for all the important information in the problem – do not disregard important data or an important condition.
   - Identify the fundamental concepts and tools needed to solve the problem.
   - Identify the steps you will follow to solve the problem.
   - Determine the role of technology in your solution process.
   - Write down an estimated answer.

4. **Solve.**
   - Carry out the solution plan. *Check each step!*

5. **Reflect.**
   - Examine the solution obtained. *Does it make sense?* Is it close to your estimate? If not, why?
   - Ensure the units are correct.
   - Ensure the answer has an appropriate number of significant digits.

1.1.3 Mathematical Modeling

In section 1.1.1 we defined what a mathematical model is. Mathematical modeling then is the art of creating mathematical models. It is more than just creating a function or a model. Mathematical modeling incorporates all of the key elements of problems solving and is iterative in nature. Figure 1.2 shows the incorporation of the problem solving elements above into the iterative modeling process that we will use throughout this course.

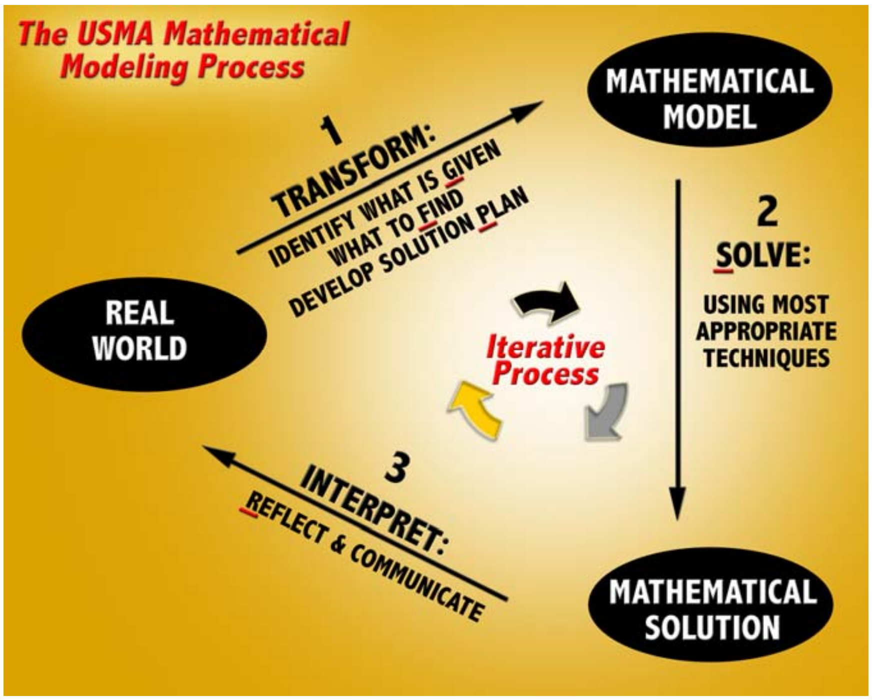


Figure 1.2: The Mathematical Modeling Triangle

Step 1 (Transforming the Real World Problem). Before developing the model, a thorough understanding of the problem at hand must be gained. Understand the root of the actual problem, not the symptoms of the problem. Know what is *given* and what you need to *find.* When this understanding is achieved, *transform* the problem into a mathematical model that can be solved using quantitative techniques. It is in the transformation of the problem into a mathematical model that you begin to develop the "plan of attack" to solve the problem. A model developed for one situation may not work in another. Beginning to think about the plan during the transformation step may save time later.

Step 2 (Solve). When the model is complete, finish the solution plan. After formulating the model, apply the most appropriate solution techniques to provide the desired answer. As with model selection, you will see in this course that choice of solution technique can be the difference between attaining an answer to a problem and not finishing.

Step 3 (Interpret the Solution). After the solution is attained, an important step in the modeling process is *interpreting* the answer, determining what the answer means in the context of the problem and putting it back into the simple, non-mathematical language for the decision maker. After communicating the solution, reflect on the results. Did the solution that you attained solve the problem posed? Did it reveal another problem that must be solved? If another problem is revealed, proceed around the modeling triangle as often as needed to develop an adequate solution or solutions.

We can demonstrate this process with a simple problem shown below.

Example 1 *John was born three years before his sister Joan and is now twice as old as she is. How old are John and Joan?*

- Step 1: Transforming the Real World Problem. What is given? We see the relationship between the ages of John and Joan. What do we need to find? John's and Joan's ages. Variable definitions: the letters x and y denote the age of John and Joan, respectively. Next, we must *transform* the information that we know into a mathematical model that we can solve quantitatively.
  - John was born three years before his sister, therefore his age is three greater than his sister, $x = y + 3$.
  - John is twice as old as his sister, therefore his age is twice that of his sister, $x = 2y$.
  - Our plan of attack looks to be the solution of a system of equations.
- Step 2: Solve the problem using the most appropriate techniques. We have two equations and two variables (unknowns) for which to solve. Our plan is to use substitution, let's implement the plan.
  - If $x = 2y$, we can use that relationship to put the equation $x = y + 3$ in terms of one variable and solve: $2y = y + 3$. Subtracting y from both sides results in $y = 3$.
  - Since we know that $x = 2y$, we can calculate x by substituting in $y = 3$: $x = 2(3) = 6$.
- Step 3: Interpret the solution. We have now solved for our variables of interest, it is now time to *interpret* the results in a manner that it easily understood. Since x represented John's age and y represented Joan's age, we answer the original question posed by stating "John is 6 years old; Joan is 3 years old." As we reflect on the answer, we see that it satisfies all of the conditions stated in the problem, so we are done. No more trips around the modeling triangle are necessary.

The following set of word problems is intended to provide practice using the modeling triangle to solve familiar kinds of problems. The modeling triangle is powerful; our goal in this course is to use it for solving actual, more complicated problems than those posed here. However, it is useful to begin slowly and increase the level difficulty of the problems, rather than starting with more complicated problems and floundering.

Question 1 *You are working with an aerial photograph taken by an unmanned aerial vehicle (UAV) of a meeting between an unknown enemy operative and a known enemy operative. The meeting was outside in bright sunlight. The known operative's height is 5 feet, 8 inches and the length of his shadow on the ground is three feet. The length of the unknown operative's shadow is 3 feet, 6 inches. You would like to know his height to help identify him at a later time. Solve this problem using the three steps of the mathematical modeling triangle.*

- *Understand the problem and transform it into a mathematical model*
- *Use appropriate solution techniques to attain a solution to the problem*
- *Communicate the solution in easily understandable terms*

For this problem, it is helpful to draw a sketch showing the situation. Note the two triangles in Figure 1.3 are similar because the shadows caused by the light rays from the sun are parallel. The lengths in this figure are expressed in inches.

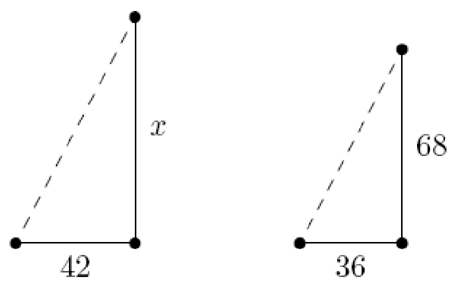


Figure 1.3: A Meeting Between Two Enemy Operatives

Question 2 *You are building a recreation center whose water supply will come from a well that is 150 feet east and 200 feet north of the recreation center. You will need to order pipe to connect the well to the recreation center. The pipe comes in 25 foot lengths. How many lengths do you need to order. Assume that there are no obstacles to burying the pipes and that they may be laid in a straight line from the well to the recreation center.*

Question 3 *Suppose that you want to make your squad leader's birthday special. Your squad mates and you have decided that to make it REALLY special, you will need to fill his room with balloons. Now, we aren't just talking about a couple of balloons that say Happy Birthday in bright pink letters – you want the balloons to fill all available space in his room. Would a couple of hundred balloons do it? Maybe a thousand balloons? How long will it take to do this and how much will it cost?*

Question 4 *How could you make your estimate of the total number of balloons required better? Describe your method and figure out the total number of balloons required using your improved method.*

Question 5 *How many doctors are there in the city of Boston?*

Notice these problems, like most of the problems you will encounter in this book, are located in the middle of the section, rather than at the end. Learning mathematical modeling is not a spectator sport. You will get the most out of this text by doing the problems as they are posed in the text. These problems are strategically placed to reinforce the concepts that immediately precede them. You will read about a concept to gain familiarity, and then do the problems to reinforce your conceptual understanding before moving on to the next concept.

1.1.4 An "Average" Model

Suppose you are playing a game in which you are guessing the heights of students that walk into the classroom. The good news is that you get to collect data to help you guess. The data you gather is below.

CDT 1	CDT 2	CDT 3	CDT 4	CDT 5	CDT 6	CDT 7	CDT 8	CDT 9	CDT 10
60"	72"	68"	62"	66"	78"	70"	72"	67"	61"

Recall from the first page of the text that a model "is a construct (e.g., a function or equation) that is designed to predict the behavior of a system." Based upon that definition and the data above, how tall would you guess the next cadet entering the room would be? There are several approaches to the question, but suppose that you were to earn 10 bonus points for guessing the answer exactly, nine for being off by an inch, 8 for two inches, etc. In this case, picking an extreme height could be disastrous for you – causing you to earn *no* bonus points.

How can we maximize the possible number of bonus points? The answer lies in your first mathematical model for the year – the *average,* or *mean* of the data. In our case, the model will be too high half the time and too low half the time, but we're as safe as we can be against either an extremely tall or short cadet entering the room.

Is it possible to make a model that is better than the average? Of course it is! In fact, in the next section we will begin learning about how to develop models that can be much more precise than simply taking the average. We will also learn how to compare models to determine which is better – stay tuned!

1.1.5 Why Mathematical Modeling is Important

A goal of this book and the United States Military Academy core mathematics program is to leverage the power of mathematics to help gain insight into the solutions of important and often urgent problems. In this course, students have developed mathematical models that have solved the following problems:

- Will a bridge collapse when a truck with a certain weight drives over it?
- What is the investment portfolio that will provide the most money in retirement, considering risk tolerance?
- How soon will a lake poisoned by contaminant be clean enough to provide a village with water?

- Given a diagram of an oceanic shipping port, where should I put the most protection to keep the port safest against attack?
- How does mathematics create image transformations seen in animated movies?
- How much money can I expect to earn if I play roulette at a casino?
- What is a viable schedule for a diabetic to follow to guide eating and insulin usage to maintain blood sugar in a healthy range?

Question 6 *How much peanut butter does the Corps of Cadets eat in a year?*

Question 7 *What does every problem solving process discussed in this reading have in common?*

Question 8 *Compare and contrast the modeling process that we have explained here with the process described by Stewart in your Calculus text on page 24.*

Question 9 Given the following real-valued functions, identify the independent and dependent variables, domain, and range.

a. $f(x) = 3x + 4$

b. $g(x) = 4x^{\frac{1}{2}}$

c. $h(x) = 4e^{x}$

d. $j(t) = 5\sin(t)$

1.2 Properties of Functions

In the previous section we discussed problem solving and mathematical modeling. As mentioned, these are the central concepts for this course. The models that are used in this course will most often take the form of a *function.* What *is* a function? Here is one acceptable definition (there are many others):

> A **function** is a rule that relates each input (domain value) to exactly one output (range value).

The fact that a function relates each input to exactly one output is important. How useful would a function be if for every input there was more than one result? How would you know which to use? Certainly having more than one output would complicate things as there would no longer be one solution at the end of your analysis. Each function, therefore, *yields one unique output for each input*.

> The **domain** is the set of all input values for a function, its *independent variable*.
> The **range** is the set of all output values for a function, its *dependent variable.*

Example 1 *Given the real-valued equation* $y = \sqrt{x}$, *what are the independent and dependent variables of the equation? What is the domain and range of the equation?*

From the preceding discussion we see that y depends on whatever value we choose for x. This makes y our dependent variable and x our independent variable. Since the domain is the set of all possible input values, we look at the possible values that x can take on. As long as x is greater than or equal to zero, the equation produces results that are real numbers (not imaginary numbers, like $\sqrt{-2}$); therefore, the domain of this equation is all real numbers greater than or equal to zero.

To determine the range, we substitute the domain values into the equation and see what we obtain as outputs. If x was zero (the smallest value in the domain), then y would be zero. If x was 4, then y would be 2. If x was 100, then y would be 10. We notice that as x gets large, so does y. In fact, as long as x is growing, so is y. This would make the range all real numbers greater than or equal to zero. Written in proper mathematical notation, the domain and range are:

$$Domain: \{x \mid 0 \le x \le \infty\} = [0, \infty)$$
$$Range: \{y \mid 0 \le y \le \infty\} = [0, \infty)$$

This notation should be read, the domain is the set of all x such that x is greater than or equal to zero and less than infinity. The range is defined similarly.

Is the equation in **Example 1** a function? Go back and reread the definition of a function. For every x value we get precisely one y value. So, we have a function. How can we test any equation to determine if it is indeed a function? This question is answered by the vertical line test.

The Vertical Line Test: A curve, when plotted, is the graph of a *function* if and only if no vertical line intersects the curve more than once.

Let's look at a couple of examples. Figure 1.4 represents two equations. The question of interest is, "Are these functions?" We notice that on the graph to the left we can put a vertical line anywhere on the graph and never intersect the curve more than once. That makes this a function! The graph on the right however is not. We notice that when we draw a vertical line on a portion of the graph, we can intersect our curve more than once.

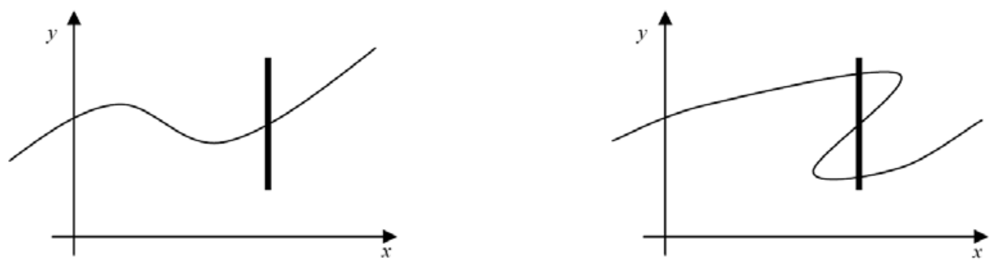


Figure 1.4: Example of the Vertical Line Test

Example 2 *Is the equation* $y^2 = x$ *a function of x?*

We first realize that this equation is written a bit differently than we normally see *x*-*y* equations. Usually we think of y as being the dependent variable, so we might want to start by solving this function for *y* in terms of *x*.

$$y^2 = x$$
$$\sqrt{y^2} = \pm\sqrt{x}$$
$$y = \pm\sqrt{x}$$

What does this say? Well, it says that if we input any value for *x* (in the domain) – except for zero – we will get two values for *y*. We recall that for our equation to be a function, we must have one and only one output for every input. Therefore, this equation is NOT a function. What would this equation look like if we wanted to graph it? Let's see in Figure 1.5.

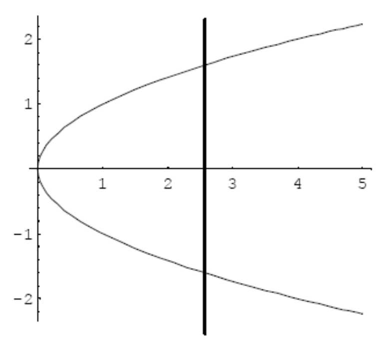


Figure 1.5: Plot of $y^2 = x$ with Vertical Line Test

As demonstrated by the vertical line, for at least some inputs, there are multiple outputs.

What else is there to say about functions? First, there is more than one way to express a function. A function can be represented in equation form, for example the **equation** $f(t)=3t-5$, which may model a vehicle's velocity with respect to time. We can illustrate our functions **graphically,** as in Figures 1.4 and 1.5. Suppose we were interested in values of the equation $f(t)=3t-5$ at certain key domain values – we may use a **table** such as Table 1.2. In relating the functions to our client, we would discuss the functions in **words**. For example, "we have determined that the velocity of the vehicle is proportional to time."

t	$f(t)$
0	-5
1	-2
2	1
3	4
4	7

Table 1.2: Tabular Representation of a Function

Question 1 *Which of the below relationships below are functions? Explain why or why not.*

a. The number of miles driven in your car versus the number of gallons of gasoline used.

b. The number of touchdowns a National Football League football player has at the end of a season.

c. The amount of snowfall that falls in Buffalo, NY (real snowy place) on any particular day of a given year.

Let's consider a situation in which we work in the West Point tailor shop and are preparing to put New Cadets in uniforms after R-Day. The tailor shop *could* provide uniforms that fit perfectly right now. Or, the tailor shop could provide uniforms that fit perfectly after the Beast Barracks summer. During R-Day, each cadet's height and weight are measured; however, if we wanted to know what the cadet's weight would be after the summer, we would need to use a *function* that related the cadet's height to his weight. **Note, in this scenario, we will assume the cadets measured are all male; we expect female measurements to be different. Table 1.3 shows heights and weights of Cadets just received into the Corps, after their first summer.

Weight (pounds)	Height (inches)
150	68
155	70
140	67
138	66
170	71
185	73
195	74
200	75
175	72
165	70

Table 1.3: Post-Beast Heights and Weights

As mentioned before…functions can be represented in four ways: in words, as graphs, in tables, and as equations. Here, we will focus on the last three, graphical, tabular and equation forms. Figure 1.6 shows the graphical representation of a function that best fits the data presented in Table 1.3.

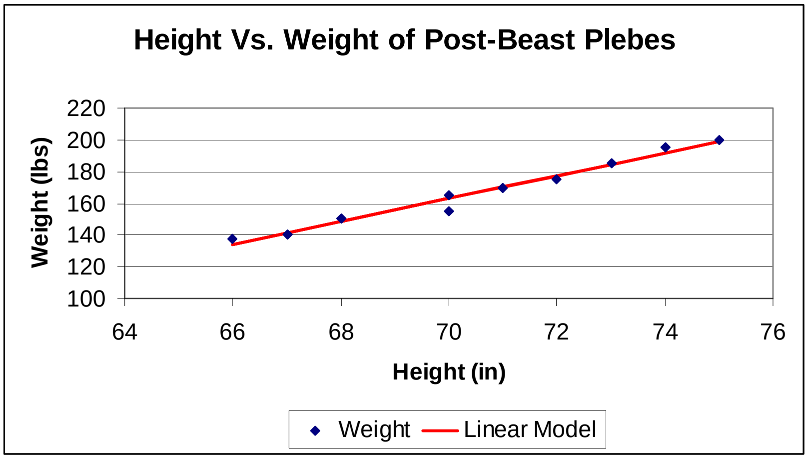


Figure 1.6: Heights and Weights of Plebes After Beast Barracks

The equation that represents the linear model shown above is:

$$Weight(Height) = 7.2\,(Height) - 341.3$$

Given the equation, $Weight(Height) = 7.2\,(Height) - 341.3$, we can determine the weight of a post-Beast Barracks cadet, based upon his height. In fact, using this model, if a cadet was 70 inches tall, we would expect him to weigh $Weight(70) = 7.2\,(70) - 341.3 \approx 163$ pounds.

What if we knew the cadet's weight, and wanted to determine the height? We would need to change the entire perspective of the problem and "flip-flop" the dependent and independent variables! We could certainly analyze the problem from the opposite or *inverse* perspective:

$$Weight = 7.2(Height) - 341.3$$
$$Weight + 341.3 = 7.2(Height)$$
$$\frac{Weight + 341.3}{7.2} = Height$$

Let's substitute in some know numbers to verify. We found that if a cadet was 70 inches tall, we would expect him to weigh about 163 pounds. What if we knew the cadet weighed 163 pounds? Could we find his height?

$$\frac{Weight + 341.3}{7.2} = Height$$
$$\frac{163 + 341.3}{7.2} = Height$$
$$70 \approx Height$$

We have just found that we can look at a relationship between variables in different ways: the way established in the equation, the table, the graph, or the sentences…or, we could look at it the opposite way. We can analyze height as a function of weight or weight as a function of height. In this case, we see a function, and its *inverse.* The *inverse* has the effect of "undoing" the original function, or looking at it in the opposite way.

We have seen that the function relating a cadet's height to his weight has an inverse function that relates his weight to his height. Let's consider a scenario. Suppose we are using a function that represents height (in feet) of a ball at time (t): $f(t) = -t^2 + 8t$.

Question 2 *Is the relation described in the equation above a function? Why or why not? Does it have an inverse* ***function****? Why or why not?*

Question 3 Find the domain and range of the following functions:

a. $f(x) = \dfrac{1}{3x-1}$

b. $f(x) = \sqrt{9 - x^2}$

c. $f(x) = \ln(x+3)$

What if we are trying to model something that has *ONE behavior for a certain part of the domain and ANOTHER behavior for the rest of the domain*? In this case, one function (rule) will not suffice. We need two functions (rules) to make it happen. In mathematics, we call these functions **piecewise functions**.

> **Piecewise Functions** are functions that are defined by different formulas in different parts of their domains.

Example 3 *A function f is defined by:*

$$f(x) = \begin{cases} x & \text{if } x \le 2 \\ x^2 & \text{if } x > 2 \end{cases}$$

Evaluate $f(0)$; $f(2)$; $f(4)$ *and sketch the graph.*

First, remember that a *function* is nothing more than a *rule*. For our given function, we have two rules that are defined on different portions of the domain. When we are working with inputs that are less than or equal to 2, we use the first rule. When we are working with inputs that are greater than 2, we use the second rule. Therefore:

$$f(0) = 0; \qquad f(2) = 2; \qquad f(4) = 16 .$$

If we want to graph this, it would look like the graph in Figure 1.7.

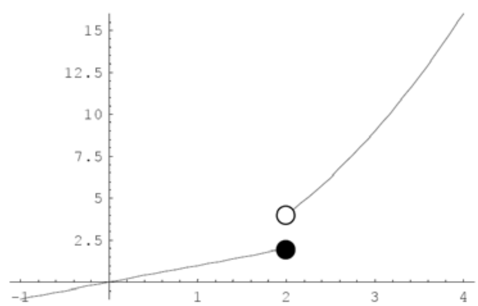


Figure 1.7: Plot of a Piecewise Function

Notice the open circle on the graph. This means the function is defined by the upper curve for every number greater than two, but not at *x=2*. Piecewise functions are handy when you have different behaviors occurring at different points of the domain…something you will see in the next section, as we discuss supply and demand. An example might be if you were to model your speed as you walk from your barracks room to math class. Perhaps you walk at a brisk pace for a while, but then you stop to open up a door. Two or more behaviors are happening that cannot be modeled by a single function.

We will be investigating many of the more common types of functions in this course. You need to get a feel for what these functions look like graphically. I am sure you realize what a line looks like and could pick one out of a line-up (pun intended) nine times out of ten. But, do you know what a logarithmic or exponential function looks like? As we review these functions, you will also need to become familiar with the effects of changing the parameters within functions. We will discuss parameters more in Lesson 4; you will see that as we change the parameters of a function, shape, orientation, and location of the function changes. The ability to visualize the shape of these functions will pay great dividends throughout the core math program as well as other coursework.

***Question 4** Each day you make your way to Thayer Hall from your barracks. Sketch a graph of your distance from your math classroom as a function of time. Pay close attention to the labels of your axes.

a. What type of function would best represent your graph?

b. What is the domain and range of your graph?

Question 5 Graphically depict an increasing function; then, a decreasing function. Explain why the function is increasing or decreasing.

THIS PAGE INTENTIONALLY LEFT BLANK

1.3 Properties of Linear Functions

We have discussed problem solving and the use of functions as tools to solve problems. We will now begin developing our skills in using models to solve problems. We typically begin the modeling process with simple models and build progressively better and more complex models. In fact, we may make several trips around the modeling triangle, building complexity and realism with each trip. Before we can build complexity, we must gain familiarity with the simplest of models, the linear model.

To illustrate the use of linear functions in modeling, we will use a story that is constantly in the news: the supply and demand that occurs as a result of the market price of oil per barrel. In this subsection we will look at the suppliers and consumers of a particular product whose production requires oil. Thus, when the price of oil rises the production cost also rises. We will focus on four related quantities:

- The **unit cost** of producing, transporting, and selling the product. As we build our model, we will use the letter c to denote this cost. We will analyze a situation in which the unit cost for the product in question is currently \$10.00 but that due to an anticipated rise in the price of oil the unit cost will rise to \$12.00.

- The **selling price** of the product. We will use the letter p to denote the selling price of the product in dollars.

- The **demand** for the product. This is the number of thousands of units of the product sold each week. It depends on the selling price. Consumers usually buy more when prices are low and buy less when prices are high. The relationship between the price and demand is called the *demand function.* We will denote this function $D(p)$. Companies often do a great deal of market research to determine the demand function because it helps them to set their price and to determine their production. In this subsection we will assume that prior analysis has helped determine the demand function to be:

$$D(p) = \begin{cases} 1000 - 25p, & 0 \le p \le 40 \\ 0, & p > 40 \end{cases} . \qquad (1)$$

Note that the price p is expressed in dollars and the demand $D(p)$ is expressed in thousands of units per week. Figure 1.7 shows a graph of this function. Notice that as the price rises, as expected the demand goes down.

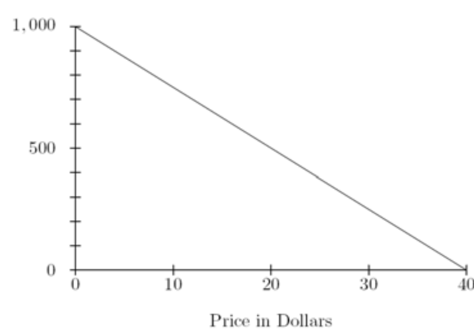


Figure 1.7: The Demand Function $D(p)$ (thousands of units per week)

- The **supply** of the product. This is the number (in thousands of units) of the product that are made each week. It is described by a function *S(p)* because the supply depends on the price. Producers usually produce more when the price is high and less when the price is low. Products are produced and sold in many different kinds of marketplaces. For some products there are a small number of large producers. For other products there are many smaller producers. Some products can be made easily and new producers can enter (or leave) the business easily. Other products are more difficult to make. For this example, we will assume that the product is easily made and is made by a large number of small producers who can enter or leave the market easily and who can easily increase or decrease their production. In this subsection we will assume that market analysis has helped determine the supply function to be a piecewise function, as defined in the previous section:

$$S(p) = \begin{cases} 0, & x < 10 \\ 40(p-10), & x \geq 10 \end{cases}. \qquad (2)$$

Figure 1.8 shows the supply function on the same set of axes as the demand function.

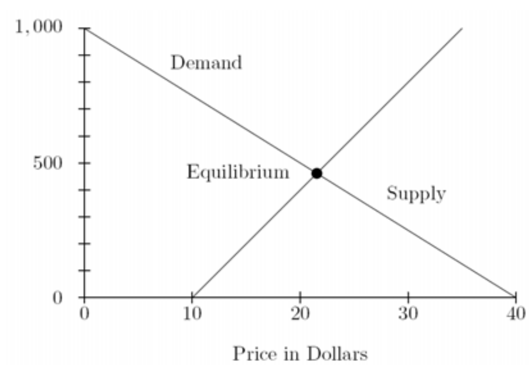


Figure 1.8: The Supply and Demand Functions (thousands of units per week)

Before going on, notice that we have just completed the first leg (transform) of the modeling triangle. We have been **given** our supply and demand functions, based on previous analysis. These functions provide us with more information about the system. In fact, \$10.00 in the supply function, $S(p) = 40(p-10)$, is the same number as the cost of producing, transporting, and selling our product. When the selling price is equal to this cost, the supply is zero because there is no point in producing a product if you can't make a profit. Notice also that as the price increases, the supply increases as well.

Now we are ready to do some mathematical analysis. Notice that when the price is low, the demand is above the supply, and when the price is high, the demand is below the supply. There is a point, marked by a dot in Figure 1.8 at which the supply and demand are equal. This point is called an equilibrium value because the producers will produce exactly enough units each week to fulfill the demand. Customers will be able to purchase exactly the number they desire. This is naturally the point we would like to **find.** We will discuss more about this application of the modeling triangle after we go into more detail about linear functions.

We say that the demand function depicted in Figure 1.7 and the supply function depicted in Figure 1.8 are "piecewise linear functions"; that is, the function is made up of more than one straight line, depending on the domain value.

Functions which have the same **average rate of change** on every interval are defined as *linear functions*.

> The **average rate of change** of a function is the **rate of change** of a function, *f(x)*, over an interval, I, where $I = b - a$, in which b is the right limit of the domain and a is the left limit.
>
> $$\text{Average Rate of Change} = \frac{\text{change in } f}{I} = \frac{\Delta\,\text{dependent variable}}{\Delta\,\text{independent variable}} = \frac{\Delta\, rise}{\Delta\, run} = \frac{f(b) - f(a)}{b - a}$$

The average rate of change of a function over an interval allows us to make a further classification of the function with which we're working. A function is said to be an **increasing function** if the average rate of change of *f(x)* is positive on every interval. Conversely, a function is classified as a **decreasing function** if the average rate of change of *f(x)* is negative on every interval. If the function is not increasing (or decreasing) on every interval, then it cannot be classified as such.

A *linear function* is a function that is in the form

$$f(x) = y = ax + d\ , \qquad (3)$$

where x is the independent variable and *f(x)* or y is the dependent variable; a is the average rate of change, or the **slope,** and d is the **y-intercept**. In previous courses, you may have used the equation $y = mx + b$ to describe a line. We will use equation (3) for the general form of the line for consistency with other functions throughout the course.

A variable represents an unknown quantity, or a quantity that *varies*. In equation (3) our x and y variables may take on many values because there is no restriction on the domain. However, for a given model the a and d quantities will always remain the same (constant). A fixed value within a function, like a or d in equation (3) is called a **parameter**.

Take some time right now to navigate to the interactive website located below to answer the question that follows.

http://www.dean.usma.edu/departments/math/MRCW/MA103/linear/live_graph.html

Question 1 *Select values for the slope of a line so that it is positive, negative, and zero. For which value is the function increasing? Decreasing? Neither?*

You should have noted from the interactive website that the larger the value of d is, the higher the intercept will be on the y-axis. The smaller the value of d is, the lower the intercept will be on the y-axis. We can conclude that the a and d parameters for the linear function determine its graph's slope (or shape) and location above (or below) the y-axis, respectively. In general, the parameters of any function (linear, trigonometric, exponential, etc.) determine its shape and location.

Parameters determine the *shape* and *location* of a function.

Since the parameter a in Equation 1 is constant, the graph of this function will always be a straight line. The sign of this parameter will determine if the line is increasing or decreasing. The magnitude of the parameter will determine how steep the line is. To demonstrate, let's consider examples.

Example 1 Suppose you have the three tables of information shown in Figure 1.9. Determine a function that describes the information in Table A. Plot and label the function.

Table A

x	$f(x)$
-5	-7.5
-4	-6
-3	-4.5
-2	-3
-1	-1.5
0	0
1	1.5
2	3
3	4.5
4	6
5	7.5

Table B

h	$g(h)$
-5	-2.5
-4	-2
-3	-1.5
-2	-1
-1	-0.5
0	0
1	0.5
2	1
3	1.5
4	2
5	2.5

Table C

t	$s(t)$
-5	10
-4	8
-3	6
-2	4
-1	2
0	0
1	-2
2	-4
3	-6
4	-8
5	-10

Figure 1.9: Data for Example 1

Step 1: *Transform the problem.* Let's consider Table A. We are **given** a table of data; the first step is to plot the data so we can see if it exhibits a pattern.

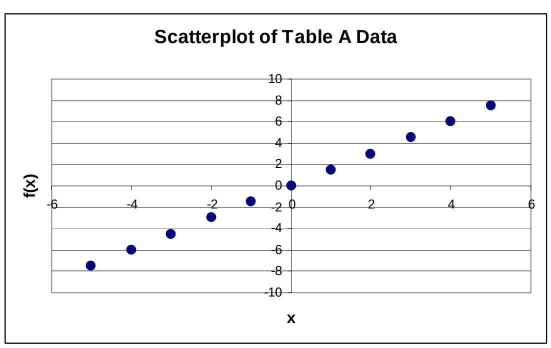


Figure 1.10: Scatterplot of Table A Data.

The data in Table A certainly exhibits a linear pattern, so we will continue the process of **finding** a line that fits through the data. We see that our data contains the point (0, 0), the origin. Because the data goes through the origin, the *y*-intercept, or 'd' value in the linear equation, will be equal to zero. Since the *y*-intercept is zero, there is one parameter left to find: our **plan** is to find the average rate of change, the slope of the line.

Step 2. *Solve the problem using appropriate solution techniques.* To **solve**, calculate the rate of change of the data at each point, using the formula on page 18. $Slope = \frac{\Delta rise}{\Delta run} = \frac{\Delta f(x)}{\Delta x}$. So, we see that the slope between the first two data points is $\frac{-6-(-7.5)}{-4-(-5)} = \frac{1.5}{1} = 1.5$. The slope between the remaining points is calculated in the same manner and works out to be the same value, 1.5.

Step 3. *Interpret the Solution.* The linear function we have developed to model the data in Table A is $f(x) = 1.5x + 0$, *or* $f(x) = 1.5x$. To **reflect** on the solution we have attained, it will be most helpful to draw a picture to make sure that our linear function goes through our data points. See Figure 1.11.

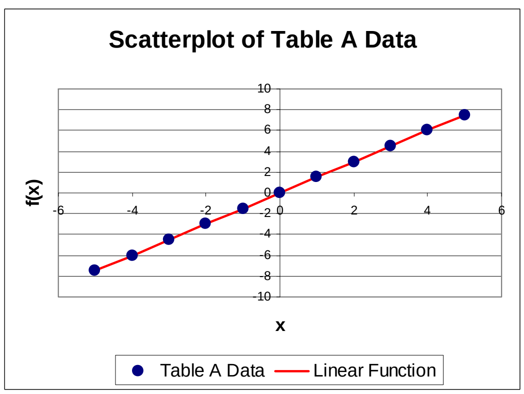


Figure 1.11: Linear Model Plotted Against Table A Data

The linear function developed to model the data we were given seems to fit the data points very well; therefore, we are satisfied that $f(x) = 1.5x$ is indeed a reasonable function to model this data set. The definition of a function is incomplete with knowing the domain and range that it is good for. In the case of this specific function, the domain and range are as follows:

$$Domain: \{x \mid -5 \le x \le 5\}$$

$$Range: \{y \mid -7.5 \le y \le 7.5, where\ y = f(x)$$

Question 2 *What are the functions that would best model the data sets given in Tables B and C in Figure 1.9?*

Now that we understand the properties of linear functions, let's use linear functions we have seen already, the supply and demand relationships established in equations (1) and (2), along with the modeling triangle and problem solving process described in Section 1.1, to solve a problem of interest.

Example 2 A supplier of goods is interested in the maximum price possible to sell a commodity at which there will be no wasted product. The price in a supply-demand system at which the supplier sells all possible product and consumers are able to buy as much as they want is the equilibrium price. What is the equilibrium point of the system described at the beginning of section 1.3?

REAL WORLD PROBLEM: Find an equilibrium point for a given supply and demand system.

Step 1. Transform problem into a mathematical model.

a. **Given:** Relationship of demand to price: $D(p) = 1000 - 25p$

Relationship of supply to selling price: $S(p) = \begin{cases} 0, & x < 10 \\ 40(p - 10), & x \geq 10 \end{cases}$

Variable declaration: $D(p)$ = Demand (thousands of units) as a function of selling price.

$S(p)$ = Supply (thousands of units) as a function of selling price

Definition of equilibrium price: $D(p) = S(p)$

b. **Find:** Equilibrium price as defined above.

c. Assume: Selling price is greater than \$10, so supply is greater than zero.

d. Solution **Plan**: Use algebraic manipulation to solve supply and demand

equations: $$D(p) = 1000 - 25p = S(p) = 40(p - 10)$$ $$1000 - 25p = 40p - 400$$

Step 2. **Solve** using appropriate solution techniques (algebraic manipulation).

$$1000 - 25p = 40p - 400$$
$$1400 = 65p$$
$$p = 21.54$$

MATHEMATICAL SOLUTION: Equilibrium price = \$21.54

Step 3. Communicate and **reflect** upon results.

The equilibrium selling price of \$21.54 is a reasonable price; it seems to match closely with the intersection of the supply and demand equations plotted in Figure 1.8.

In this lesson, we have focused on the *slope-intercept* form of a line. There are two other forms, the *point-slope* form and the *general form* of a line.

> The *slope-intercept form* is: $f(x) = ax + d$, where a is the slope (rate of change) and d is the *y*-intercept.
>
> The *point-slope* form of a line is: $y - y_0 = a(x - x_0)$, where a is the slope (rate of change) and (x_0, y_0) is a point on the line.
>
> The *general form* of a line is : $Ax+By+C=0$, where A, B, and C are constants.

Question 3 *Using the definition for average rate of change, your understanding of y-intercept, and your knowledge of the slope-intercept form of a line,*

a. Determine an equation describing each of the following tables of data.

Table A

x	*f(x)*
-5	-2
-4	-1.6
-3	-1.2
-2	-0.8
-1	-0.4
0	0
1	0.4
2	0.8
3	1.2
4	1.6
5	2

Table B

h	*g(h)*
-5	5.5
-4	5
-3	4.5
-2	4
-1	3.5
0	3
1	2.5
2	2
3	1.5
4	1
5	0.5

Table C

t	*s(t)*
-5	-13
-4	-11.4
-3	-9.8
-2	-8.2
-1	-6.6
0	-5
1	-3.4
2	-1.8
3	-0.2
4	1.4
5	3

b. Verify your equations for each table of data. Then, graph and label each function.

***Question 4** *Thayer Hall has four floors and is approximately 70 feet tall. Taylor Hall has nine floors and is approximately 180 feet tall. Develop a model that predicts the height of a building based on the number of floors it has. (Don't forget to include your model's domain and range.)*

Question 5 *Plot two lines that are parallel. What is relationship of their slopes? Ensure you draw the lines to scale as much as possible (graph paper helps).*

Question 6 *Plot two lines that are perpendicular. What is relationship of their slopes? Ensure you draw the lines to scale as much as possible (graph paper helps).*

THIS PAGE LEFT INTENTIONALLY BLANK

1.4 Modeling With Technology

In the previous section, we discussed the use of a linear model to solve a problem that we may see in economics: finding the equilibrium point in a system governed by traditional rules of supply versus demand. You were also given the functions that showed the supply and demand behaviors; these functions were both "determined by previous analysis." In this section, we will *do the analysis* to develop the supply and demand curves so we can find the equilibrium point of an economic system.

Suppose we are analysts for a large oil company and we want to analyze the behavior of the American oil market to make more money. The company has access to data collected through years of business. We will conduct an analysis of the price of oil (U.S. dollars per barrel), the demand of oil (millions of barrels per day in the United States), and the supply of oil (same units as demand).

First, we will develop a linear model that enables us to represent oil demand, given the price of oil. See Table 1.4 for the data (data is fictional, also linked on the course website).

Price ($ per Barrel)	Demand (Millions of Barrels per Day)
75	17.22
80	16.58
85	15.94
90	15.05
95	14.90
100	13.72
105	13.43

Table 1.4: Oil Demand vs. Price

We now have a problem that we must solve: develop a linear model for the data above, so it is possible to determine the equilibrium price of the system. What must we do to solve the problem? A process for solving the problem would certainly be in order!

Step 1. *Transform the Problem.* Given in Table 1.4 is the data we will analyze. The independent variable is the price of oil because that determines the demand of oil (the dependent variable). We must find a model that fits through the data so we can make predictions. Our first step in finding the most appropriate model is to use the skills learned in our first problem solving lab to plot the data. See Figure 1.12 for a plot of the data found in Table 1.4.

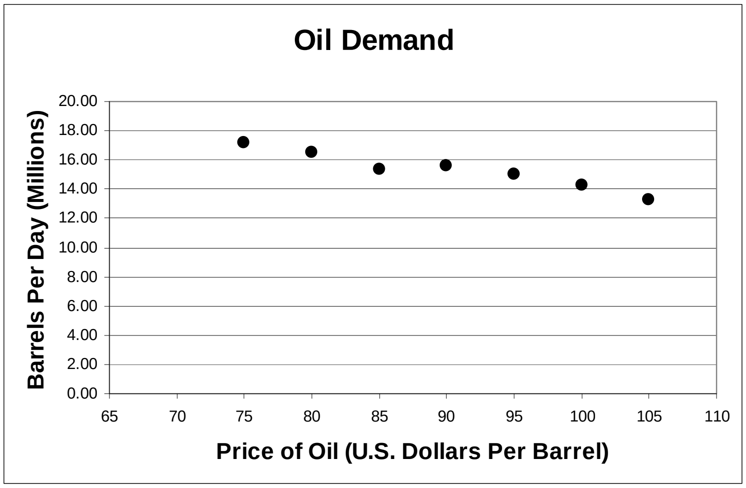


Figure 1.12: Plot of Oil Price vs. Demand

It is easy to see that a straight line will not perfectly fit through the data. That's OK! A model that you fit to actual data *very* rarely will fit perfectly through every data point – but we must do the best we can. It seems that a line will be able to give us a model that is "good enough" to solve the problem at hand: find the equilibrium point of our system. The plan we will use will be to find values for the parameters associated with a line (the slope and y-intercept).

Step 2. *Solve the Problem Using the Most Appropriate Techniques*. To finalize our solution plan, we should remind ourselves of the general form of the model we selected. The general form of a line is: $y = ax + d$. Notice that we have two parameters, 'a' and 'd.' We have two variables, 'x' and 'y.' We can solve for our parameters if we select values for the variables, forming a system of equations. Because we have two parameters we need to solve for, we will need two equations, selecting two data points. If the first and last data points are representative of the general trend of the data, it is common to select the first and last data points to estimate an initial model. Let's see what happens when we use these points: $(x, y) = (75, 17.22)$ and $(105, 13.43)$. The model we will have after solving for our two parameters should go through the two data points we used to develop the parameters – a useful fact when reflecting on our solution. Now that we have our two data points, let's form our equations:

$$17.22 = 75a + d$$
$$13.43 = 105a + d$$

We could solve these equations using substitution. Let's use one of the technology tools we have at our disposal – Mathematica, which would be much faster than substitution! In the problem solving lab, we learned how to solve two equations in two unknowns. Let's apply that knowledge here.

```
Solve[{17.22 == 75 a + d, 13.43 == 105 a + d}, {a, d}]

{{a → -0.126333, d → 26.695}}
```

Figure 1.13: Solving Two Equations in Two Unknowns (Mathematica)

The Mathematica output indicates that a line with a slope of about -0.1263 and *y*-intercept of approximately 26.695 will run through the two data points we selected. The final equation is: $y = -0.1263x + 26.695$.

Step 3. *Interpret the Solution.* Perhaps the most important step in the modeling and problem solving process is interpreting the solution – communicating it in non-mathematical terms and reflecting upon whether or not the solution we attain solves the problem posed. To help interpret the solution, let's plot a graph of the linear model we developed on the same axes as the data.

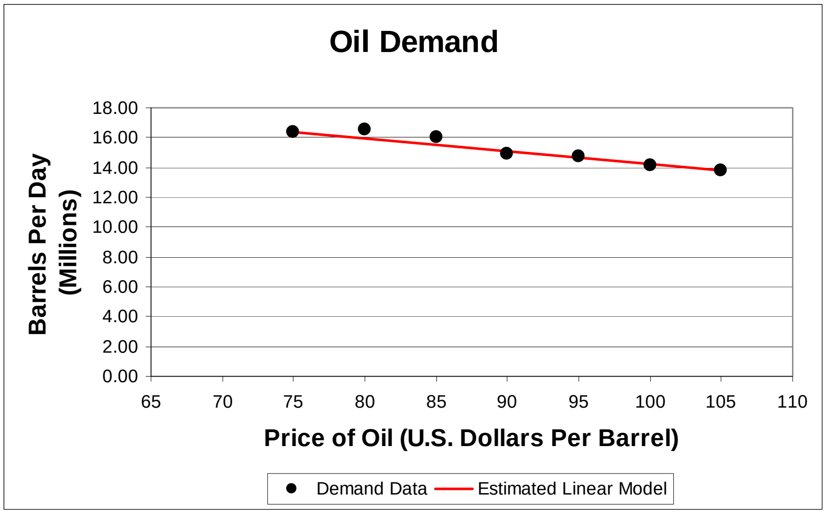


Figure 1.14: Illustration of Estimated Linear Model

The model we have developed, relating demand to price with the equation $\text{Oil\_Demand} = -0.1263\text{Oil\_Price} + 26.695$ seems do a nice job of describing the data provided in Table 1.4. The model does seem to fit directly through the first and last data points, as expected. The domain and range of the function follow:

$$\begin{aligned} &Domain(x): && [75,105] \\ &Range(y): && [13.43,17.22] \end{aligned}$$

People that do modeling for a living are curious people. Our initial assumption was that the first and last data points were "good enough" from which to create an initial model. What if we changed the assumption we used about which points to choose? Let's try the third and sixth data points, (85, 15.94) and (100, 13.72).

Again, using Mathematica to solve for our parameters, our slope and *y*-intercept are -0.148 and 28.52, respectively. The slopes are different by about .02 (about a 15% difference) and the y-intercepts by less than 2.0 (7%). It appears that the model's y-intercept is less *sensitive to change* than the model's slope. We complete a **sensitivity analysis** by testing how much change in an assumption will impact the final model.

> The Government Accountability Office (GAO) defines **sensitivity analysis** as the determination of how sensitive outcomes are to changes in the assumptions. The assumptions that deserve the most attention should be those with the greatest amount of uncertainty and effect on the outcome.

In general, we prefer a model that is *robust* against change. That means, when an assumption changes, the model remains essentially the same. It is undesirable to have a model that fluctuates wildly with changes in the assumptions.

We compared the two models, algebraically. We can also compare them graphically by overlaying the second model over the first model and data. As expected, the two developed models do not look much different (see Figure 1.15). In the next section, we will discuss methods to determine which of the algebraic models is truly the best model. For now, we will choose the first model we developed, $y = -0.1263x + 26.695$.

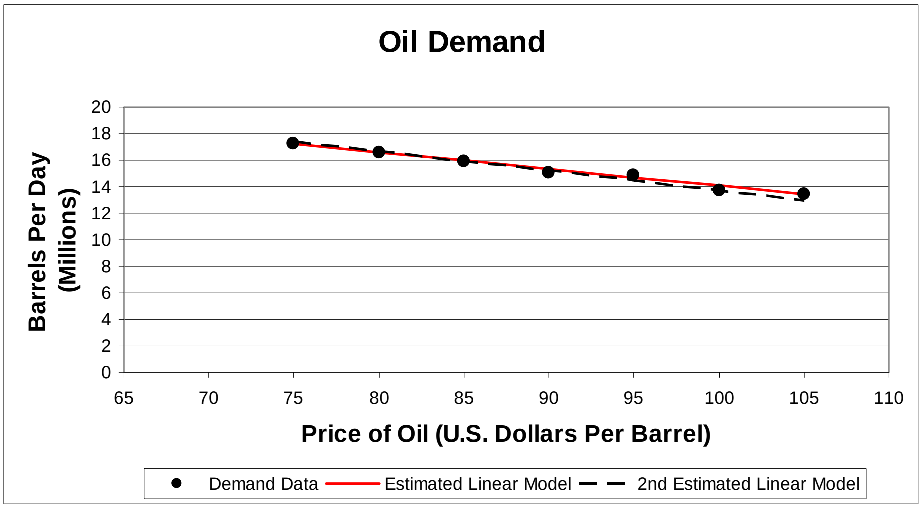


Figure 1.15: A Comparison of Linear Oil Demand Models

This linear model will enable us to find the equilibrium point of the oil supply and demand system, after the supply curve is plotted.

Question 1 *Given the data in Table 1.5 (and on course website), develop an equation for the supply of oil, in millions of barrels per day, given the price of a barrel of oil.*

Price ($ per Barrel)	Supply (Millions of Barrels per Day)
75	2.56
80	3.42
85	3.50
90	3.85
95	4.62
100	4.90
105	6.13

Table 1.5: Oil Supply vs. Price

After determining the model for calculating oil supply as a function of selling price, we are now prepared to determine the equilibrium price of the system we are analyzing. There are two possibilities for the equilibrium price: it could occur inside the current domain values of our functions, [75, 105]. It could also occur outside those values, meaning we must extend our models.

If we predict a value *inside* the current domain of our function, it is said that we are **interpolating**. For example, if the equilibrium point occurred at a selling price of $87, it would be a price *inside* our domain; therefore, we would *interpolate*.

If we were to predict a value outside the current domain, meaning we must *extend* our model, it is said that we are **extrapolating**. For example, if the equilibrium point occurred at a selling price of $115, it would be a price outside our domain. We would have to *extend* our model (it may still be a good model); therefore, we would *extrapolate.*

> We define **interpolation** as the act of making predictions *within* the domain of known values or data. **Extrapolation** is the act of predicting values *outside* of the domain of the data.

Question 2 *Determine the equilibrium point of the oil supply and demand system described in this section. Is the answer an* interpolation *or* extrapolation *of our model?*

Example 1 *In 1986, the Space Shuttle Challenger experienced a catastrophic failure in its solid rocket booster. The explosion, 73 seconds after liftoff, claimed both the crew and shuttle. The cause of explosion was later determined to be an o-ring failure in the right solid rocket booster. The final investigative report concluded that cold weather was a contributing factor.*

The o-rings in the solid rocket boosters on the space shuttle are designed to expand when heated to seal different chambers of the rocket so that solid rocket fuel is not ignited. According to engineering specifications, the o-rings must expand by at least 5% in order to ensure a safe launch. The temperature on the day of launch was 29 degrees F. O-ring expansion data was collected on the previous nine launches and is shown in Table 1.6.

Temp (degF)	93	88	87	81	73	72	68	64	55
% Expansion	22.3	21.0	20.6	19.7	18.7	19.0	17.3	16.2	15.5

Table 1.6: O-Ring Expansion Data for Space Shuttle Challenger Launches

If you were given this data prior to launch, what would you have recommended?

Step 1. *Transform the problem.* We must answer the question posed above: should the shuttle launch. We must identify what is <u>given</u> in the problem.

- We have data from the previous nine launches.
- We know that the temperature at time of launch was 29 degrees F.
- We know that o-rings must expand by 5% to ensure a safe launch.
- Can we draw a picture (graph)? YES! (see Figure 1.16 on the next page)

What must we <u>find</u> in the problem? We need a model to determine what the percent expansion of an o-ring will be if we have a launch temperature of 29^{O}F.

Because we need to find the expansion, that will be our *dependent* variable. The variable is the "cause" of the o-ring expansion is the temperature, the *independent* variable.

In addition to identifying what is given and what we need to find, there is an important assumption to be made: that o-rings react to temperature in a predictable manner at temperatures outside the domain of current launch temperatures (e.g., there is not a temperature at which the o-rings stop expanding).

Let's use Mathematica to plot this graph, instead of Excel. The following command will
result in the plot of our data.

```
ListPlot[{{93, 22.3}, {88, 21}, {87, 20.6}, {81, 19.7}, {73, 18.7}, {72, 19}, {68, 17.3}, {64, 16.2}, {55, 15.5}},
 PlotRange → {14, 23}, AxesLabel → {Temperature, O - ring_Expand}]
```

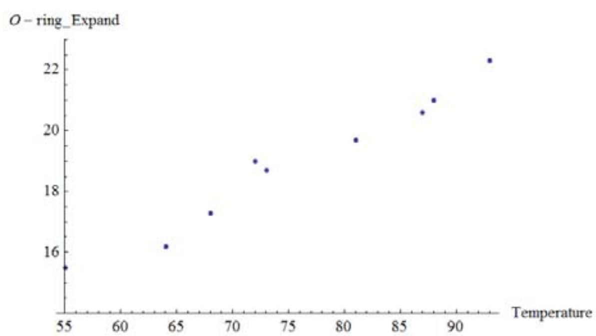


Figure 1.16: Mathematica Graph of O-Ring Data

The ListPlot command plots the data points, while the PlotRange command tells Mathematica the "y" values to plot, rather than giving the standard plot starting at zero. The AxesLabel command enable you to label the x and y axes. From this point on, it will be your choice whether to use Excel or Mathematica to plot data.

After graphing the data point, we can see that the data seem to follow a linear pattern; therefore, we should use our knowledge of linear models to develop a model that we can use to **extrapolate** whether or not the shuttle should launch at a temperature of 29^{O}F. Because we will use a linear model, the general form of the model is: $Expansion(Temp) = a(Temp) + d$.

Our plan will be to estimate values for the parameters (a and d) of our linear model. To estimate two parameters, we need two data points that represent the data well. We see the data in Figure 1.16. For this model we will assume that the first and last data points will lead us to reasonable estimates of our parameters. Keep in mind that when choosing data points to estimate the parameters, the goal should be to minimize the deviation from the general trend of the data.

Step 2. *Solve Using the Most Appropriate Technique.* Our plan will be to use the first and the last point to formulate two equations with two unknowns (a and d). We will then solve these two equations simultaneously and use the solutions to estimate the two parameters of the linear function. Finally, we'll need to determine our model's domain and range. Using our first and last data points, we create two equations with two unknowns. Our two equations are:

$$22.3 = a(93) + d$$
$$15.5 = a(55) + d$$

Figure 1.17 highlights the use of our *Mathematica* skills to solve both equations simultaneously for *a* and *d.* We find our estimated slope is 0.18 and our *y*-intercept is 5.66. Our final model is: $Expansion(Temp) = .18(Temp) + 5.66$.

```
Solve[{22.3 == a (93) + d, 15.5 == a (55) + d}, {a, d}]

{{a → 0.178947, d → 5.65789}}
```

Figure 1.17: Using Mathematica to Solve Two Simultaneous Equations

Thus, our model is: $Expansion(Temp) = 0.18 \cdot Temp + 5.66$. The data show our domain and range to be:

$$Domain: \{x \mid 55 \le x \le 93\}$$
$$Range: \{y \mid 15.5 \le y \le 22.3\}$$

Step 3. *Interpret the Solution.* The final step of the modeling and problem solving process is to communicate and reflect upon the solution that we have derived. Through interpreting the model developed, we see that on the day of launch, the temperature was 29 degrees F; therefore, our model predicts: $Expansion(29) = 0.18(29) + 5.66 = 10.88$

According to the model we created, the o-rings would have expanded nearly 11%. Our results indicate the o-ring expansion adheres to the engineering standards for a safe launch. It is upon reflection that we realize that our model may very well be flawed. After all, the Space Shuttle exploded. Was there something in the data that could have led us to that predict a dangerous launch?

Let's test an assumption to determine the sensitivity of our model. We assumed that the first and last data points were representative of the trend in our data. So, let's make a minor change to our selection and see what happens. Instead, we'll choose the second and last data points and see how much of an affect this has on our solution.

```
Solve[{21 == a (88) + d, 15.5 == a (55) + d}, {a, d}]

{{a → 0.166667, d → 6.33333}}
```

Figure 1.18: Using Technology to Test the Sensitivity of our Assumptions

Making this change slightly decreased our slope and increased our y-intercept (Figure 1.18). This also causes our new prediction on the day of launch to be: $Expansion(29) = 0.17(29) + 6.33 = 11.26$

Our new model predicts an even greater expansion than the last (former model is solid line, latter, dashed). Figure 1.19 shows the data versus the two possible models described in the text, and how the models are plotted.

```
Plot1 = ListPlot[{{93, 22.3}, {88, 21}, {87, 20.6}, {81, 19.7}, {73, 18.7}, {72, 19}, {68, 17.3}, {64, 16.2}, {55, 15.5}, {29, 11.26}},
  PlotRange → {5, 23}, AxesLabel → {Temperature, O - ring_Expand}]

f[x_] = .178947 x + 5.65789

5.65789 + 0.178947 x

Plot2 = Plot[f[x], {x, 29, 93}, PlotRange → {5, 23}]

g[x_] = .166667 x + 6.33333

6.33333 + 0.166667 x

Plot3 = Plot[g[x], {x, 29, 93}, PlotRange → {5, 23}, PlotStyle → {Dashed}]

Show[Plot1, Plot2, Plot3]
```

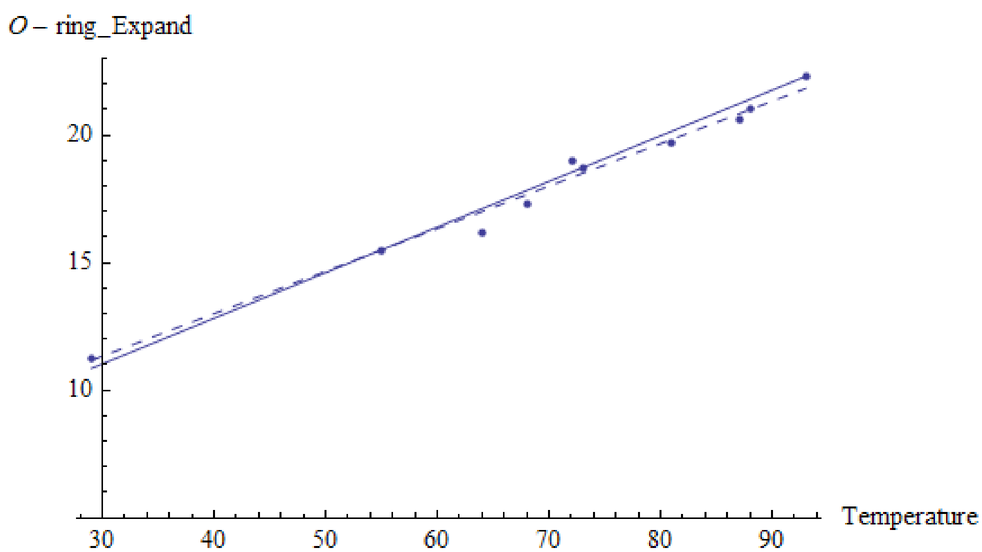


Figure 1.19: O-Ring Sensitivity Analysis

Our analysis concludes that the launch should have been safe; however, since we know this wasn't the case, perhaps our assumptions were not valid. Specifically, our assumption to select a linear model and our assumption that the data accurately represent what happens to o-rings at other temperatures may be invalid. We will revisit this idea later in the course.

Question 3 The data represented in Table 1.6 are posted on the course website. Using this data and a linear model, would it have been *possible to predict the Challenger's explosion?*

Question 4 *Given the following height and weight data from R-Day of 10 male new cadets,*

Weight (pounds)	Height (inches)
150	68
155	70
140	67
138	66
170	71
185	73
195	74
200	75
175	72
165	70

answer the following questions.

a. *What type of model do you think would be appropriate to predict a male new cadet's height given his weight?*
b. *Predict the height of a male cadet that weighs 100, 160 and 300 pounds.*
c. *Does your model have any limitations? What impact do these limitations have on the domain and range of your model?*

***Question 5** *The National Collegiate Athletic Association (NCAA) exists to promote a commitment to excellence in both the classroom and the "fields of friendly strife." Data is provided below that describes NCAA's membership since 1950*[1].

Year	Active Members
1950	362
1955	449
1960	524
1965	579
1970	645
1975	704
1980	738
1985	793
1990	828
1995	903
2000	977
2001	977
2002	1005
2003	1024
2004	1028
2005	1027

a. *Determine a model for NCAA membership as a function of time. Be sure to give the domain and range of your model.*

In its 2004 Annual Report, the NCAA noted that it anticipated that its membership growth rate will slow in coming years.

b. *Based on the prediction in the Annual Report, what characteristics would best describe a model of NCAA membership in the future?*

[1] 2004 NCAA Membership Report, Retrieved 26 January, 2006, from
http://www.ncaa.org/library/membership/membership˙report/2004/2004˙ncaa˙membership˙report.pdf

1.5 Model Evaluation I (Fit of Functions)

We have used our knowledge of problem solving and properties of linear functions to *develop* models to help communicate solutions to several problems, like identifying the height of an enemy operative, determining the size of cadets for uniform issue, and calculating the equilibrium price of a supply and demand system. Developing models is useful, but our models are limited if we do not know how good they are. This section will focus, not on the development of models, but on their *evaluation.* We will discuss how to determine a model's "goodness," thereby putting more credibility behind the predictions we make.

1.5.1 Model Evaluation (Subjective)

So far, we have informally used two tests to determine our model's goodness: the *nature of the data* test and the "eyeball" test. We will now formally define these tests.

- Nature of the Data: The underlying structure of the data; i.e., how the data would appear if collected in a "perfect world." For example, we would expect projectile motion to follow a parabolic trend as gravity acts on the projectile.
- Eyeball Test: A qualitative measure that determines how closely a model appears to fit a given data set.

We may consider a linear model to be "good" if we knew about the circumstances under which the data were collected and expected the modeled quantities to be proportional – a measure that the model matches the nature of the data. If we graphed the model and the prediction looked like it closely represented the data (passes the eyeball test), then we may call it a "good" model, like the models presented in Figure 1.20. Today, we will develop a means to quantitatively evaluate just how good a model is. How can we tell if one model is really better than another? We will explore this question and more in this lesson.

Let's consider the average demand, the model we discussed in Section 1.1, and the two demand models we developed in Section 1.4.

$$\begin{aligned} &Average\_Demand: && y = 15.21 \\ &Estimated\ Linear\ Model 1: && y = -0.1263x + 26.695 \\ &Estimated\ Linear\ Model 2: && y = -0.148x + 28.52 \end{aligned}$$

It appears in Figure 1.20 that the average does not "fit" the data well. In fact, the model underpredicts for the first half of the data and overpredicts for the second half of the data. Both estimated linear models seem to be a reasonable fit of our data set. Which line is the better linear model most accurately predicting oil demand given the price of oil per barrel? Both the solid and dashed lines seem to provide a good approximation, but they clearly are different lines (models). So which model best fits the data? We need to define a measure for best fit so that we can compare the models we developed.

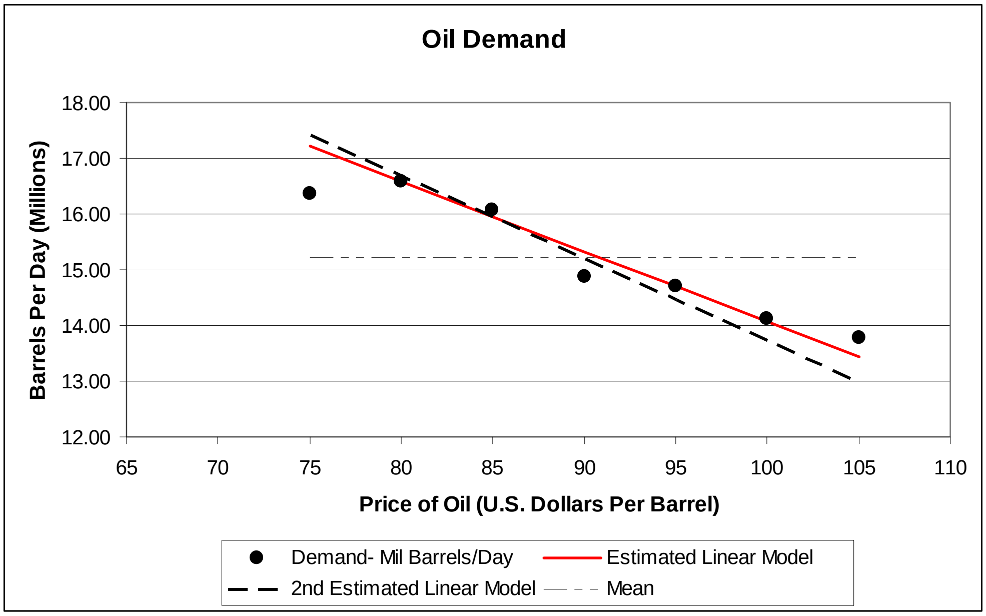


Figure 1.20: A Comparison of Linear Oil Demand Models

1.5.2 Model Evaluation (Sum of Squared Error)

One way to determine which model is "*best*" quantitatively is to measure the distance between the values predicted by the model, $\hat{y}$, and the actual data points, y. This distance is our error, e. (This symbol, *e, is not to be confused with the number, e ≈ 2.72).*

Let's look at the situation a bit closer. If we look at a single data point from Figure 1.20, say, (95 dollars, 14.72 million barrels) we see that both estimated linear models under-predict the actual amount of oil in demand. This is because both models lie slightly below the point. Does this mean that our models are wrong? Absolutely not! The models still capture the general trend of the data; they just predict that consumers will demand less oil, given the oil's selling price.

So, how good or bad are our models? To *quantify,* or put a numerical value on how good or bad they are, we measure the distance between our predictions and the actual data points (see Figure 1.21). In this case, our demand for oil is 14.72 million barrels our solid model predicts a demand of about 14.70 million barrels.

What is the error of the solid model? Well, it's $\textit{actual demand} - \textit{predicted demand}$, or in this case, it is $14.72 - 14.70 \approx 0.02$ million barrels.

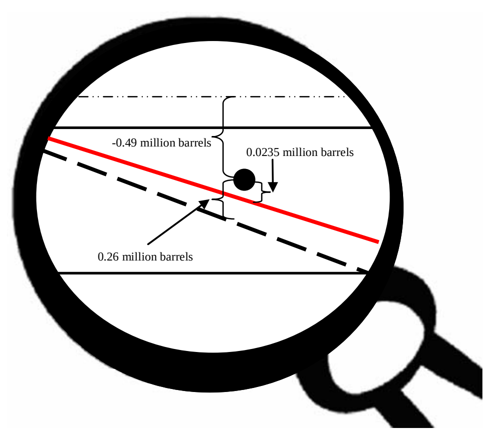


Figure 1.21: Closer Look at Our Errors

This calculation shows that the model under-predicts the actual demand by 0.02 million barrels of oil. What would have happened if we looked at the point just to the left, (90, 14.88). In this case, our model over-predicts; the error is $14.88 - 15.328 = -0.448$ million barrels.

So, what do we do with this information? How can we come up with a single number that *quantifies* how good or bad the entire model is? Let's look at a couple of possibilities.

A-Not-So-Good-Idea: If we sum the errors for each model, perhaps the model with the lowest sum would be the better of the two models. Notationally, with n points in a given model, this is written:

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

The above notation may be a little intimidating at first. If we read it from the inside out, it may become a little easier to understand. The quantity $(y_i - \hat{y}_i)$ means "the actual data point minus the predicted value." The subscript i helps us to keep track of which data point we are calculating. For example, $(y_1 - \hat{y}_1)$ is the calculation of the error for the first data point, where $(y_2 - \hat{y}_2)$ is the second error calculation. The symbol $\sum_{i=1}^{n}$ means "the sum from the first value to the n^{th} (n is the final value). So, putting it all together, $\sum_{i=1}^{n}(y_i - \hat{y}_i)$ means "the sum of errors, from 1 to n."

There is one major problem with this technique. If we assume a positive distance when the data point is above the line and a negative distance below the line, a value of zero may only mean that we overestimated and underestimated exactly the same amount. A value of zero would occur because the positive and negative errors would cancel each other out. We really want to measure the total deviation of the model from the data which this method does not capture.

A-Bit-Better-Idea: Perhaps we could take the absolute values of the distances. Notationally, with n points, this is written:

$$\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| = \sum_{i=1}^{n}\left|e_i\right|$$

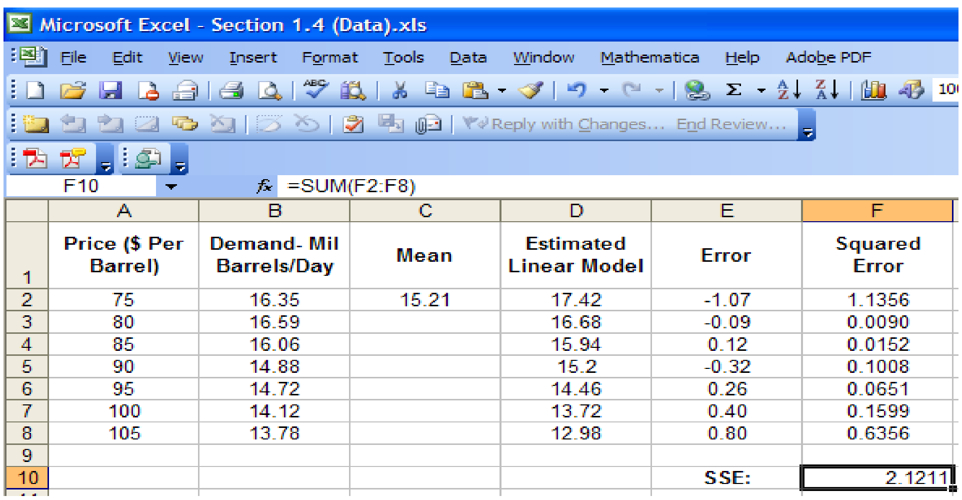
Using absolute values may seem reasonable, but when absolute values are involved, trying to minimize the error between the model and the actual data using calculus techniques becomes quite difficult.

<u>A-More-Acceptable-Idea:</u> A more common and acceptable method is to square the distances before we sum them. Notationally, with n points, this is written:

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

This accomplishes two purposes. First, it eliminates any negative errors (that might cancel out positive errors), and second, it emphasizes those data points that are furthest from our model. That is, we are severely penalized for data that are far from our model when the distance is squared. The *value of summing the square of the errors* is called the **Sum of Squared Errors** (***SSE***). If we compare two models using this method, the one with the smaller *SSE* indicates that the model "fits" the data better than the one with the larger *SSE*. We will consider other model evaluation methods in future lessons.

How can you calculate the *SSE* using MS *Excel*? First, create two columns next to your data and model to track the error and squared error terms. Figure 1.22 demonstrates one way to set up an MS *Excel* spreadsheet.



	A	B	C	D	E	F
1	Price ($ Per Barrel)	Demand- Mil Barrels/Day	Mean	Estimated Linear Model	Error	Squared Error
2	75	16.35	15.21	17.42	-1.07	1.1356
3	80	16.59		16.68	-0.09	0.0090
4	85	16.06		15.94	0.12	0.0152
5	90	14.88		15.2	-0.32	0.1008
6	95	14.72		14.46	0.26	0.0651
7	100	14.12		13.72	0.40	0.1599
8	105	13.78		12.98	0.80	0.6356
9						
10					SSE:	2.1211

Figure 1.22: Sample MS *Excel* Spreadsheet

Given a linear model with a slope of -0.148 and an intercept of 28.52, the model portrayed in Figure 1.22 displays an SSE value of about 2.1211. Is this good? What can we do with this value? SSE is a ***relative*** number. This means that it is only useful when we can compare it with the SSE values from other models.

Think about it this way – suppose that we had a data set that had one million points. By some stroke of genius, let's say you found a model that was only 0.1 units off from each of the one million data points. To the naked eye, the data and your model appear to be exactly the same. The error for each point is only 0.1 and the squared error for each of these points 0.01. Yet the *SSE* for this model is:

$$1{,}000{,}000\ \textit{data points} \cdot (0.1\ \textit{units})^2 = 10{,}000\ \textit{units}^2$$

What does this tell us? Is this a large or small number? Unfortunately, an SSE by itself does not tell us much. Without comparing this SSE to the SSEs of other models, we cannot tell if the SSE above is high or low.

Question 1 *Figure 1.22 shows the calculation of the SSE for the second estimated linear model, with parameters 'a'=-0.148 and 'd'=28.52. Is it better than the first estimated linear model developed in Section 1.4, with parameters 'a'=-.1263 and 'd'=26.695?*

Question 2 U*sing the Space Shuttle Challenger data set from Section 1.4, create a spreadsheet like the one shown in Figure 1.22. What is the SSE for the model estimated in Section 1.4? If you change the values of the slope and intercept for this model, can you get a smaller SSE?*

1.5.3 Model Evaluation (Coefficient of Determination)

Another quantitative measure commonly used in statistics is the coefficient of determination, or r^2. This value is bounded by zero and one ($0 \le r^2 \le 1$), and is ***a measure of the amount of variation of the data for which your model can account.*** So, an $r^2 = 0.83$ would mean that 83% of the variation of the data can be accounted for by your model, the remaining 17% of the variation must be the result of some other factors not covered in your model.

> The **coefficient of determination** or r^2 is a measure of the amount of variation of the data accounted for by the model and is bounded by zero and one ($0 \le r^2 \le 1$). $$r^2 = 1 - \frac{SSE}{SST} \qquad (1)$$

Before we calculate r^2, we must define a key term, the Sum of Squares Total, SST. Recall that the Sum of Squared Error was a measure of how the model you developed compared to the data for which it was developed. SST is similar, but instead of calculating the error between the data and a model you estimated, SST calculates the error between the data and the mean – the simplest model in our toolkit.

The mean is a simple model with just one parameter, so we expect it to be our least accurate model. Because SST quantifies the error of the least accurate model, we can say that it is the measure of how the worst model, the mean, compares to the data for which it was developed. SST is calculated in a similar manner to SSE,

$$SST = \sum_{i=1}^{n} (y_i - \bar{y}), \qquad (2)$$

where SST is the sum of squared deviations of the y-values of the data from the sample mean of the data. The sample mean is the average of the given outputs denoted $\bar{y}$. SST is the relative measure of how bad the mean is as a model, that's why it's called Sum of Squares Total – it's the total amount of squared error using the worst model.

If we divide SSE (the error between the data and the model you developed) by the SST we get the percentage of error in the data that our model does not explain.

The best possible model would have zero error ($SSE = 0$), accounting for 100% of the variability in the data. If we subtract the percentage of error that our model does *not* account for from 100% (or 1), we get a measure of error that the model *does* account for. Therefore, $r^2 = 1 - \frac{SSE}{SST}$.

Example 1 *In Section 1.4, you created a linear model for homework that predicted the height of an R-day cadet based on his or her weight. Our best fit model is:*

$$height(weight) = 0.135451(weight) + 47.94905$$
$$\text{Domain} : \{weight \mid 90 < weight < 300\} = (90, 300)$$
$$\text{Range} : \{height \mid 60.13 < height < 88.57\} = (60.13, 88.57)$$

It has an SSE of approximately 1.95. Now, we will turn our analysis to r^2 *to determine the accuracy of our model. Recall, we will need to calculate SST using Equation (2) on the previous page where SST is calculated by:*

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})$$

Since we determined the SSE of this model previously, we focus here on how to calculate the SST.

- First, we must determine the sample mean which is the average of the data values; here it is the average height. The average height is 70.6 inches.
- Next, we determine how far each of the data values is from this mean - that is, determine the deviations of the y-values of the data from the sample mean of the data.
- Then, we square each of these values.
- Finally, we sum them up.

The SST for this example is 80.4.

The final step in determining the r^2 is to use Equation (1) inputting our SSE and SST where appropriate.

$$r^2 = 1 - \frac{1.94705456}{80.4} = 0.975783$$

Here, our model accounts for 97.6% of the variation in our y-values. This indicates that our model is very accurate for this data set.

One attribute of r^2 is that, unlike SSE, it is a standardized or normalized measure. In other words, r^2 values do not need to be compared to substantiate the use of one model over another. If your model has a high r^2 it indicates it is accurate for the set of data you are modeling.

Caution! A high r^2 value *does indicate the model is accurate* for the data *but does not necessarily indicate that it is appropriate* for the situation being modeled. Similarly, a low r^2 (say in the low 40s) does not necessarily mean the model is not useful. For example, we can find polynomial models to fit every data set but not every data set should be modeled with a polynomial function. To ensure a model is both accurate and appropriate, employ more than one evaluation tool.

1.5.4 Model Evaluation (Number of Parameters)

An important consideration when modeling is the relationship between the number of parameters in a model and the measure of 'goodness' (often SSE or r^2). Every time we add a new parameter to a model, the SSE will get smaller and r^2 will always get larger, even though the added parameter may end up being of insignificant value. The question for modelers is, "when are enough parameters enough?" It's often best to use this notion of Albert Einstein, "Keep things as simple as possible, but no simpler."

Let's explore an example that will show use the effect of adding additional terms and parameters to a model.

Example 2 *In the following graph, Figure 1.23, we once again see the seven points of data that define the oil demand problem, this time with three models overlaid on the data.*

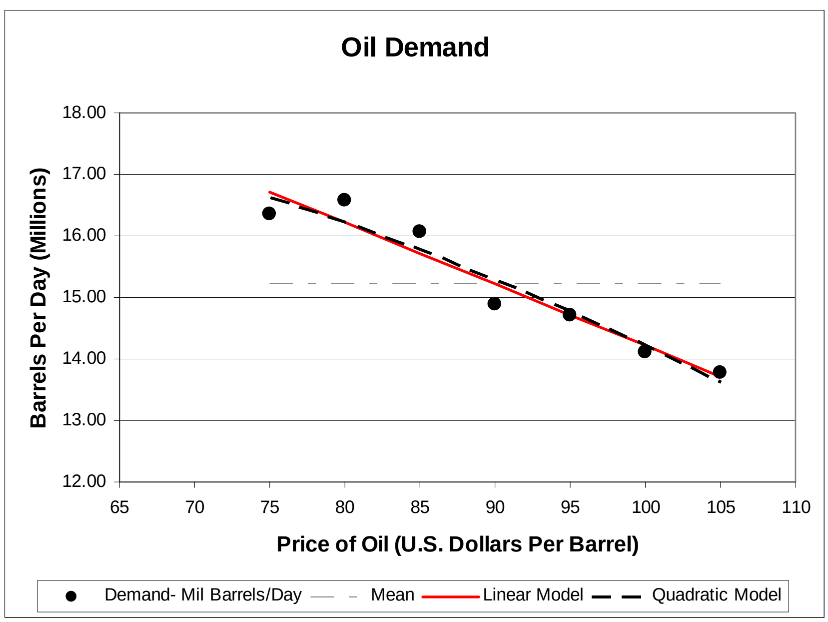


Figure 1.23: Oil Demand Modeled with Three Models

Clearly, the relationship between the independent variable and the dependent variable is decreasing. Let's model the data with polynomials having one, two, and three parameters (mean, linear with non zero y-intercept, and quadratic). NOTE: To estimate the 3-parameter model, we chose three representative data points and used a polynomial of the form $f(x) = a(x+c)^2 + d$ *, which you'll see again in section 1.8. Table 1.7 lists the model, minimum SSE, and corresponding parameter values for the data.*

parameters	*model*	*Min SSE*	*a*	*c*	*d*
1	*d*	*7.52*			*15.21*
2	$f(x) = ax + d$	*0.49*	*-0.1263*		*26.695*
3	$f(x) = a(x+c)^2 + d$	*0.51*	*-.0053*	*-81.99*	*16.611*

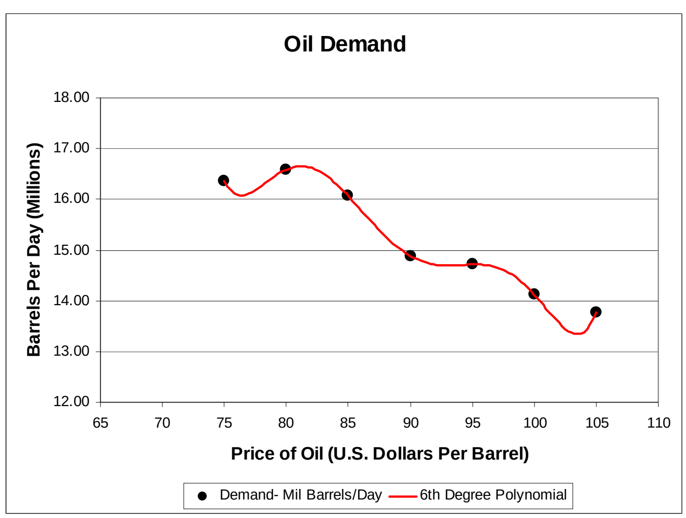
Table 1.7: Polynomial Models

From the graph and the table, we can note a few important things. First, the minimum SSE for the one-parameter model is an order of magnitude higher – has a change of more than 10 times – than the two-parameter model (a linear model). But, the improvement of going to a three-parameter model (a quadratic model) is only about 4%. Also, graphically the one-parameter model is clearly poor, while the other two are virtually indistinguishable from each other. Finally, then the third parameter, a, is introduced, its value (a = -0.0053) is so close to zero when compared to the scale of the other parameters, that it is not a useful addition to the modeling process.

Because the addition of the third parameter is not useful in this case, we would decide to use the simpler linear model. However, we noticed that the SSE decreased when we added a parameter to model, a fact that will always hold true. A larger number of terms will *always* result in lower SSE (and larger r^2). **In fact, a polynomial of degree *n* – 1 will always fit *n* data points.** For example, a data set of two points can be fit perfectly by a degree 1 polynomial (a line). Also, three points can be fit by a 2nd degree polynomial (a quadratic). Each would have a SSE = 0. A polynomial of degree n - 1 will always fit n data points. But if a polynomial of degree n – 2 can closely fit the data points, why use a more complicated model?

In Example 2, there were 7 data points. A 6-degree polynomial can model these points "perfectly," with an SSE of zero and r^2 of 1. A 6-degree polynomial is a very "wiggly" function (see picture at right). As you would expect, the function is also very complicated:

$$f(x) = .0000009x^6 - .0005x^5 + 0.1066x^4$$
$$-12.692x^3 + 847.71x^2 - 30102x + 444020$$



Notice that the model's trend is opposite of the data, seeming to indicate that for prices higher than \$105, demand will increase. It doesn't make sense that higher price would yield higher demand. This model, though it fits every data point, is not going to be useful for prediction.

As mentioned previously, in every case, as you add a parameter, SSE will decrease. But the important question is, "does the improvement in SSE justify the disadvantage of adding a parameter and increasing the complexity of the model?" The nature of the disadvantage of having too many parameters lies in the strength of the estimates for those parameters. Every parameter that needs to estimated essentially 'takes away' some of the information in the sample. Estimating two parameters from seven data points will be fairly easy and will result in relatively accurate estimates for those two parameters. On the other hand, estimating 6 parameters from 7 data points is far more complicated and will likely result in parameters of little consequence, such as a from Table 1.7. Furthermore, once you have more parameters than data points, it is not possible to estimate any of the parameters.

What's a good rule of thumb in determining how many parameters to add to your model? When adding an extra parameter doesn't give noticeable graphical improvement or a noticeable improvement on SSE, then stop adding parameters. What is noticeable? This is a tough question and often depends on the situation…another example of why modeling has a "science" component and an "art" component. Remember: for the model you develop to have meaning, you must be able to interpret its results to a decision maker. If you are unable to communicate the effect of an additional parameter, it's a good bet that you may not need that parameter!

Question 3 *You are working with a group of student interns for the National Center for Atmospheric Research (NCAR) in Boulder, Colorado, monitoring the global warming situation. NCAR has access to data about* CO_2 *concentrations in the atmosphere (below and on course website). Let x represent the number of years since 1995.*

years	0	1	2	3	4	5	6	7	8	9
CO_2 ppm	361.6	364.0	364.6	367.3	369.6	370.5	372.1	373.5	376.1	379.0

a. *Develop a linear model that fits these data. Show all work in how you obtained your model and put a graph of your model on the same axes as the scatterplot.*

b. *Select two different points to estimate a linear model.*

c. *Thoroughly evaluate your model's, meaning: use the nature of the data test, the eyeball test, SSE,* r^2*, and number of parameters. Is your model a good fit for the data?*

d. *These data were all observed in March of each year. Based on your model what were the* CO_2 *ppms in September 2001?*

e. *Based on your model what is your prediction for the* CO_2 *ppm in March 2100?*

f. *The equation used at NCAR to predict CO_2 ppm was $y_2 = 1.82x + 361.6$. Plot this equation with your model and the data on your scatter plot. Which equation seems to be the "better" model? Explain your reasoning.*

Question 4 *Using the models developed in Example 1 of Section 1.4 (O-Ring Data), determine the r^2 value for each model. Which appears to be the better model? Based on your understanding of the problem, interpret the meaning of your r^2 values and explain what it indicates about the best model. Discuss your findings.*

***Question 5** *You are working with a group of student interns for the park services at Yellowstone National Park in Wyoming. Old Faithful is one of 400 geysers within the park. It is named Old Faithful, because the time of next eruption can be predicted by the duration of Old Faithful's eruption. Since the original model was established, the intervals between eruptions have tended to increase. The park service would like you to set up a model with current data and answer a few of their questions with the model. Below are some current data on Old Faithful's eruptions.[1] The data is also linked on the course website.*

x duration in minutes	1.8	1.98	2.37	3.78	4.3	4.53	1.82	2.03	2.82	3.83	4.30	4.55	1.88	2.05
y interval in minutes	56	59	61	79	84	89	58	60	73	85	89	86	60	57

[1] *These data were obtained from Yellowstone National Park.*

a. *Based on the graph you develop, does it appear that interval is approximately a linear function of duration?*

b. *What is the slope of the line that models this data? Explain in practical terms (duration and interval) the meaning of this slope?*

c. *What is the y-intercept for your model? Does this intercept have any practical meaning? If so, what is it?*

d. *Thoroughly evaluate the model you developed, meaning: apply the nature of the data test, the eyeball test, and r^2.*

e. *Suppose that you observe an eruption that lasts 2 minutes and 40 seconds. Based on your model, predict when to expect the next eruption.*

f. *The equation used by the park service to predict the intervals between Old Faithful's eruptions is $y = 14x + 30$. Compare this equation with your model. Which equation seems to be the "better" model? Explain your reasoning.*

1.6 The Generalized Exponential Function

Take a moment to conduct an internet search on the word "exponential." When this chapter of the text was being written, Google returned over 7 million hits on sites or articles that somehow referenced "exponential." Below are just a few of the quotes you may find.

"The death toll from the Indian Ocean tsunami is set to rise ***exponentially*** *above current estimates of 150,000 as relief workers reach remote villages and survivors succumb to disease, UN officials warned today."*[1]

"Scientists are finding that electromagnetic flux of unknown spectrum in action may be the root cause of the ***exponentially*** *increasing number of earthquakes in the last eighteen months."*[2]

Exponential growth or decay…you hear the term used by newscasters, political pundits, and people you talk with, but what precisely does it mean?

You may have personal experience with exponential functions in a place that may be important to you in the future - an interest bearing savings account. Perhaps you have taken a class at some point in your life where you learned a function that allowed you to compute the value of an account that accrued interest continually. You may have heard it referred to as PERT, or seen it presented as Equation (1); this equation is an exponential function.

$$V = Pe^{rt} \tag{1}$$

where:

V is the value of the account at the time that you are interested (dependent variable)

P is the principal, or initial value invested (parameter)

e is the **base** (parameter)

r is the interest rate (parameter)

t is the time that has passed (in the same units as the period referred to in *rate*) since the initial investment. (independent variable)

What makes this function exponential is that there is a constant base raised to a variable power. In this case, the base is e which is also approximately 2.71828. The most basic form of the exponential function is:

$$f(x) = b^x \tag{2}$$

where b is the constant base (parameter) raised to a variable power x, the independent variable. **Do not confuse this with the power function (learned later) where the independent variable is the base. Remember: the exponential function has the independent variable in the exponent.**

[1] http://www.guardian.co.uk/tsunami/story/0,15671,1382973,00.html

[2] qd.typepad.com/19/2005/01/its_been_a_whol.html

1.6.1 Properties of Functions

Recall that a function is said to be an increasing function if the average rate of change of *f(x)* is positive on every interval. Conversely, a function is classified as a decreasing function if the average rate of change of *f(x)* is negative on every interval. Similarly, we say a function is increasing on an interval if the average rate of change of *f(x)* is positive on the interval. Also, a function is decreasing on an interval if the average rate of change of *f(x)* is negative on every interval. More precisely,

A function is **increasing** on an interval if $f(x_1) < f(x_2)$ whenever $x_1 < x_2$ and x_1, x_2 are in the interval.
A function is **decreasing** on an interval if $f(x_1) > f(x_2)$ whenever $x_1 < x_2$ and x_1, x_2 are in the interval.

If we trace a graph from left to right, we notice our hand will move up when the function is increasing and down when decreasing. In Section 1.3, we saw that a linear function increases when it has a positive slope and decreases with a negative slope. This extends to other functions – a positive rate of change indicates the function is increasing and negative, decreasing.

In the linear case, we saw that the rate of change was constant and its sign determined if the line was increasing or decreasing. We will see many examples of exponential functions that increase or decrease and their rates of change that also increase or decrease. The next subsection catalogs the three cases of exponential functions and their corresponding changing rates.

1.6.2 Three Cases of the Basic Exponential Function, b^x

There are three cases of this basic exponential function; each case has a different value for *b*.

Case 1: $b > 1$. This function increases as the domain values increase. This can be seen in the first table of Figure 1.24a. This case of the exponential function also increases at an increasing rate. This can be seen by investigating the average rates of change which are also shown in Figure 1.24a. Notice how these rates are increasing. Thus the function is not only increasing but increasing at an increasing rate. A graph of this type of function is shown in Figure 1.24b. These functions are normally referred to as growth functions.

Case 2: $0 > b > 1$. This function decreases as the domain values increase. This can be seen in the second table of Figure 1.24a. This case of the exponential function also decreases at an increasing rate. This can be seen by investigating the average rates of change. Notice in the second table of Figure 1.24a how these rates are increasing. Thus the function is not only decreasing but decreasing at an increasing rate. These functions are normally referred to as decay functions.

Case 3: $b = 1$. This function equals 1 everywhere. The average rate of change of this function will always be zero (Figure 1.24a) regardless of whether or not the domain values increase or decrease. Since this case results in a constant function (Figure 1.24), we will not investigate this case any further in this section.

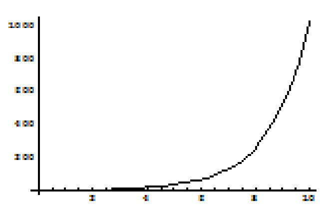
****NOTE**** We only consider values for b that are greater than or equal to zero. Why? Well, what happens when you raise a negative real number to a power? If the exponent is an even integer, then the answer is a positive real number. If the exponent is an odd integer, then the answer is a negative real number. If the exponent is any non-integer real number, then the answer is a complex number. In this course, we will focus on real-valued functions. That is, we will concentrate on functions whose domains and ranges are both subsets of the real numbers.

b = 2		
x	*f(x)=b^x*	*AVG RoC*
-2	0.250	
-1.5	0.354	0.207
-1	0.500	0.293
-0.5	0.707	0.414
0	1.000	0.586
0.5	1.414	0.828
1	2.000	1.172
1.5	2.828	1.657
2	4.000	2.343

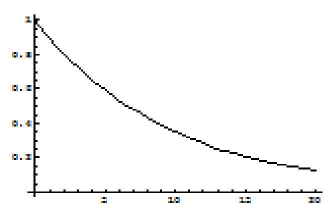
b = .5		
x	*f(x)=b^x*	*AVG RoC*
-2	4.000	
-1.5	2.828	-2.343
-1	2.000	-1.657
-0.5	1.414	-1.172
0	1.000	-0.828
0.5	0.707	-0.586
1	0.500	-0.414
1.5	0.354	-0.293
2	0.250	-0.207

b = 1		
x	*f(x)=b^x*	*AVG RoC*
-2	1	
-1.5	1	0
-1	1	0
-0.5	1	0
0	1	0
0.5	1	0
1	1	0
1.5	1	0
2	1	0

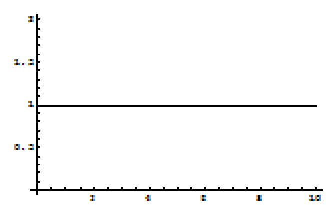
Figure 1.24a: Average Rates of Change for Exponential Functions (Cases 1-3) on Domain $-2 < x < 2$



Case 1: Increasing at an Increasing Rate



Case 2: Decreasing at an Increasing Rate



Case 3: Constant Function (Rate of Change = 0)

Figure 1.24b: Graphs Depicting Basic Exponential Functions

Question 1 *Plot graphs of each of the three cases for exponential functions, different from those shown in Figure 1.24b. Evaluate the functions over the domain [-3, 3] for each case. For Case 1, assume b = 1.5, for Case 2, assume b = .4, and for Case 3, assume b = 1. Identify the domain and range of each of these functions.*

1.6.3 The Generalized Exponential Function

The graphical and associated word descriptions of the behavior of exponential functions provided above will prove very useful in your modeling efforts. However, it is important to recognize other forms of the exponential function and their patterns of behavior as well. In this course, we will investigate the more generalized exponential function listed below.

$$f(x) = ab^x + d \qquad (3)$$

The three cases of the basic exponential function shown in Figure 1.24 share the property that the parameter $a = 1$ and that the parameter $d = 0$. The parameters for a generalized exponential function (in a similar manner to that for the linear function) control the shape and location of its graph.

Generalization 1. We can shift any graph of a function upward or downward by adding a parameter to the function:

$$y = f(x) + d$$

If d is positive, we will shift or translate the function's graph upward. If d is negative, we will shift or translate the function's graph downward. Unique to exponential functions is that the parameter d also indicates the location of a horizontal asymptote.

> A **horizontal asymptote** of a curve is a line $y = L$ that the function *f(x)* approaches as x approaches infinity.

> A **vertical asymptote** of a curve is a line $x = a$ where the values of the function *f(x)* approaches infinity as x approaches the undefined point a.

Generalization 2. By changing the a parameter, we can flip the function about the horizontal asymptote, or line $y = d$. We can also stretch and shrink the function vertically. Table 1.8 shows the vertical stretch and shrink of an exponential function, $f(x) = e^x$.

x	$f(x)=e^x$ "Base" Function	$f(x)=10e^x$ Vertical *Stretch*	$f(x)=0.10e^x$ Vertical *Shrink*
-2	.1353	1.353	.01353
-1	.3679	3.679	.03679
0	1	10	.1
1	2.7182	27.182	.27182
2	7.389	73.89	.7389

Table 1.8: Tabular Representation of Vertical Stretching and Shrinking

Note that the first column is the base function, the second includes a vertical *stretch*. The a parameter is increased, so for each value of the independent variable, the value of the vertically stretched function is larger than that of the base function. In the third column, we implement a vertical *shrink* by decreasing the value of the a parameter. For each value of the independent variable, the function value is smaller.

We can see the effect of changes made on the a parameter graphically, as well as in a table. The upper left picture in Figure 1.25 is $f(x)=e^x$. The upper right function shows the result of negating the function. In addition, we can vertically stretch or shrink the graph of any function by increasing or decreasing the a parameter. The bottom left function illustrates the effect of decreasing the value of the a parameter; the bottom right shows the effect of increasing the a parameter.

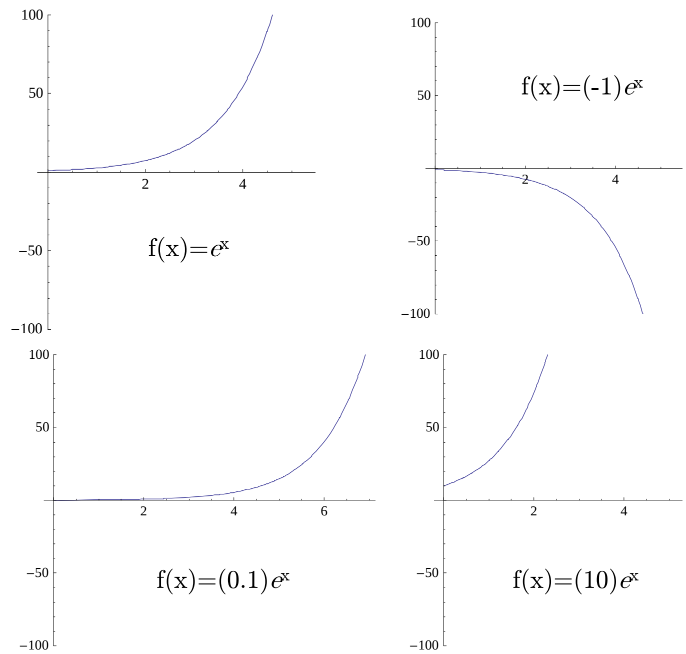


Figure 1.25: Illustrating the Effect of Changing the a Parameter in the Exponential Function

Generalization 3. The parameter b determines the shape of the function. Recall the three cases of the exponential function illustrated in Figure 1.24a and b. Another illustration of the impact of the b parameter is shown in Figure 1.26.

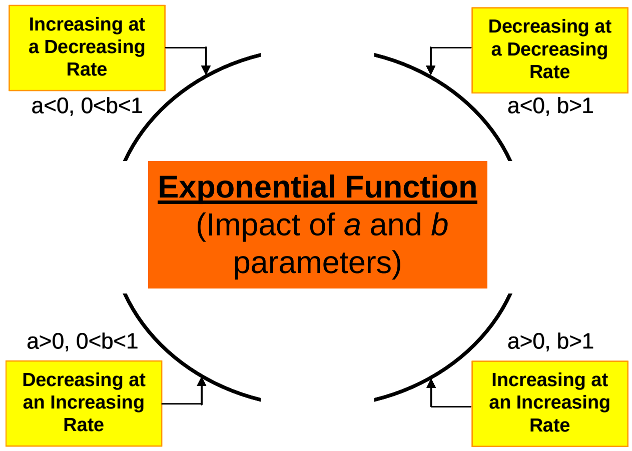


Figure 1.26: An Illustration of Generalizations 2 & 3.

The first case, increasing at a decreasing rate, means that our finger will move up as we trace left to right on the upper left curve in Figure 1.26 (increasing), but not as fast at the end as at the beginning. Next, decreasing at a decreasing rate means when traced, our finger moves down, but we actually decrease faster at the end. Let's explain in terms of the average rate of change, a concept we discussed in the context of linear models.

We see that the function in the upper right quadrant is decreasing; that is, the dependent variable is getting smaller as the independent variable increases. The next area to address is the rate at which the function is decreasing, an analysis of the rate of change. In doing this, always move from left to right on the path of the function.

First consider the ruler labeled "Step 1" in Figure 1.27, measuring the average rate of change between the points $(1, 7.5)$ and $(4, 7.0)$.

$$\textit{average rate of change} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{7.0 - 7.5}{4 - 1} = \frac{.5}{3} = -.1\overline{66}$$

Next, consider the ruler labeled "Step 2" in Figure 1.27, measuring the average rate of change between the points (4, 7.0) and (7, 4.5).

$$\textit{average rate of change} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{4.5 - 7.0}{7 - 4} = \frac{-2.5}{3} = -.8\overline{33}$$

Notice that -.833 < -.166, therefore the average rate of change (or the slope) is *decreasing* and the function is decreasing at a decreasing rate.

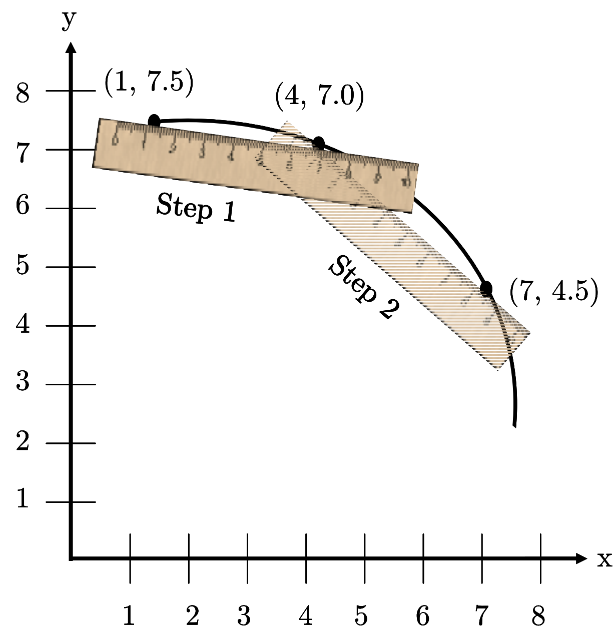


Figure 1.27: Illustration of Decreasing at a Decreasing Rate.

We will use Figures 1.26 and 1.27 to describe many standard functions in the upcoming lessons, including exponential functions in this section.

The box below summarizes the exponential properties we have discussed so far. You should have a good understanding of each of these.

Properties of Exponential Functions of the form $f(x) = ab^x + d$:

- If $a > 1$, and
  - $b > 1$, the function will increase at an increasing rate as the independent variable increases.
  - $0 < b < 1$, the function will decrease at an increasing rate as the independent variable increases.
  - $b = 1$, the function will remain constant an equal to the sum of the parameters a and d as the independent variable increases.

- If $a < 1$, and
  - $b > 1$, the function will decrease at a decreasing rate as the independent variable increases.
  - $0 < b < 1$, the function will increase at a decreasing rate as the independent variable increases.
  - $b = 1$, the function will remain constant an equal to the sum of the parameters a and d as the independent variable increases.

You should also have a good understanding of the Law of Exponents which is summarized in the following box.

Law of Exponents for Exponential Functions:

1) $a^{x+y} = a^x a^y$

2) $a^{x-y} = \dfrac{a^x}{a^y}$

3) $(a^x)^y = a^{xy}$

4) $(ab)^x = a^x b^x$

You may want to take some time right now to navigate to the interactive website located at:

http://www.dean.usma.edu/departments/math/MRCW/MA103/exponential/live graph.html

(You can also link to this website through the MA-103 course webpage). Once there, adjust the parameters to see how the graph changes. Intuitively, we see that if the equation had the values a=1, d=0, then we are back to our simplest type of exponential function described in Equation (2).

Question 2 *In your own words, describe how changes in the parameters of the generalized exponential function affect its shape.*

Question 3 *In Table 1.9 there are three sets of data that represent three different functions. Which data comes from an exponential function? Can you identify the other types of functions from the data? Do you need more information for the first two data sets?*

x	y
0	20.0000
1	21.0000
2	22.1000
3	23.2775
4	24.6425
5	26.2650

x	y
0	20.0000
1	21.0000
2	22.0500
3	23.1525
4	24.3101
5	25.5256

x	y
0	20.0000
1	21.0000
2	22.0000
3	23.0000
4	24.0000
5	25.0000

Table 1.9: Values for Question 3[3]

[3] This problem is from Functioning in the Real World, A Precalculus Experience by Gordon, Gordon, Tucker, and Siegel, pg. 83

1.6.4 Applications of the Generalized Exponential Function

Example 1 Modeling the Growth of an Investment.

Let's discuss the growth of money in a traditional savings account. If money in the account grew in a linear manner, the changes in the account from year to year, as we learned in Lesson 4, would all be the same (there would be a constant average rate of change). The second column of Table 1.11 illustrates how a \$1000 investment might actually grow in a savings account. You can see in the third column, the average rate of change between each successive year is NOT constant and thus, the growth is NOT linear. In fact, the average rates of change are increasing because the more money that is in the account, the more money there is to make additional interest. The balance of the account can be modeled with a Case 1 exponential function where parameter $a > 0$ and parameter $b > 1$.

Suppose your parents put money into a bank account for you after you graduated the sixth grade. You are not sure what the interest rate is, but based upon old annual statements, you can see the growth that has occurred. You would like to know how much will be in that account when you retire from the military after 20 years of service (23 years from now, year 7).

Year	Amount	AVG RoC
0	\$1,000.00	
1	\$1,051.27	51.27
2	\$1,105.17	53.90
3	\$1,161.83	56.66
4	\$1,221.40	59.57
5	\$1,284.03	62.62
6	\$1,349.86	65.83
7	\$1,419.07	69.21

Table 1.10: Average Rate of Change of an Exponential Function

Step 1: *Transform the problem.* We are **given** a table of data, we must examine the table.

- Define the variables:
  - Independent variable (input): year
  - Dependent variable (output): amount of money in account

- Nature of the data: Apply the quantitative measures of the nature of the data to gain an idea of which type of function may be most appropriate. See Table 1.10.

Notice that the data are increasing, as is the rate; therefore, we need a function that is increasing at an increasing rate: a Case 1 exponential function. Let's graph the data to verify our conjecture. See Figure 1.28.

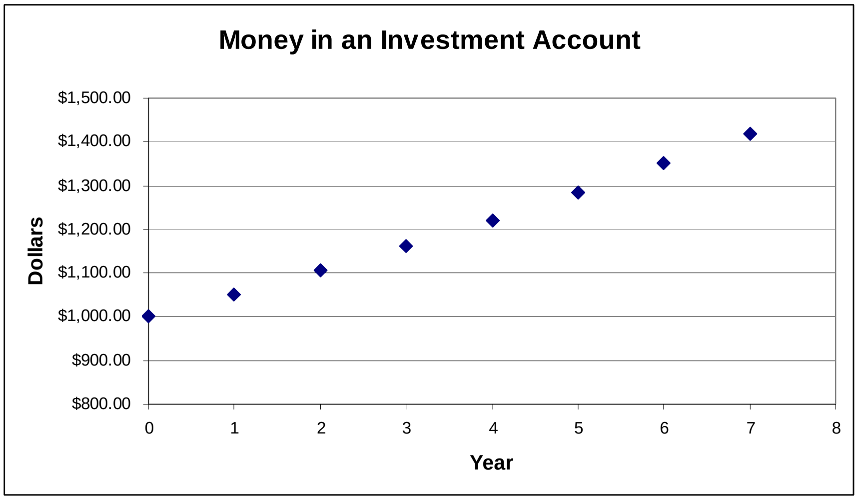


Figure 1.28: Scatterplot of Table 1.10 Data.

The data in Figure 1.28 seem to exhibit a *linear* pattern. Our quantitative measure of the nature of the data and our visual measure seem to conflict. We will explore the two models that explain what appears to be going on in the data set: the linear model (visual inspection) and exponential model (nature of the data). We know that we will have to **find** the model that fits through the data best.

- Assumptions
    - No money is withdrawn from the account
    - No additional deposits are made into the account
    - The data accurately represent what will happen to the population in the coming years.

Our plan is to use the model development and evaluation techniques that we have used to this point in the course to find the best model to determine how much money we will have after any given year in the future.

Step 2. *Solve the problem using appropriate solution techniques.* To **solve**, we will develop a linear model and an exponential model, then compare the two to decide which is best.

- Linear model development. First, we'll try the linear model. In general, our linear model will be:

$$Amount(year) = a(year) + d$$

Linear Model Parameter Estimation: Using modeling skills developed to date, we need to estimate the parameters (a and d) for a linear model. Let's use the first and last data point to form two equations we can solve for the two unknown parameters. The two equations follow:

$$1000 = a(0) + d$$
$$1419.07 = a(7) + d$$

We see by inspection that the d parameter (y-intercept) is 1000. We can use that information to solve for a.

$$
\begin{aligned}
1419.07 &= a(7) + 1000 \\
419.07 &= 7a \\
a &= \frac{419.07}{7} \approx 59.87
\end{aligned}
$$

The estimated model is:

$$
\begin{aligned}
&Amount\,(year) = 59.87(year) + 1000 \\
&Domain: \{year \mid 0 \le year \le 7\} \\
&Range: \{Amount \mid 1000 \le Amount \le 1419.07
\end{aligned}
$$

Overlaid with our data, the model is graphed in Figure 1.29.

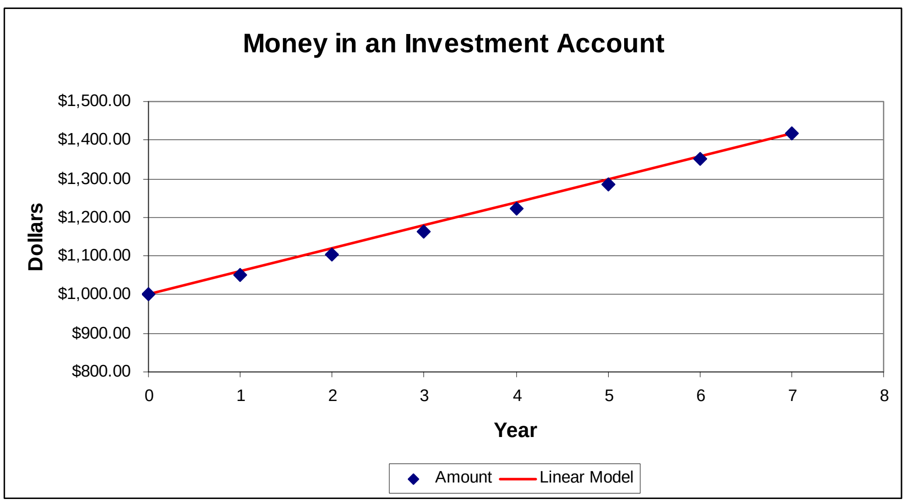


Figure 1.29: Investment Data Overlaid with Linear Model.

Using the eyeball test to evaluate Figure 1.29, we see that the estimated linear model seems to be a pretty good fit to the data.

Question 4 *Calculate the Sum of Squared Error and the coefficient of determination for the linear model shown in Figure 1.29. Use the two tests to determine the model's "goodness of fit."*

- Exponential model development. Now that we've developed a power model, it is time to develop our generalized exponential model, of the form:

$$Amount\,(year) = ab^{x} + d$$

As we saw earlier in the example, each of the points seems representative. It make sense to solve for all three parameters if we can, so we'll use the first, last, and middle points to form three equations in three unknowns and let Mathematica do the work for us. The three equations are:

$$1051.27 = ab^{1} + d$$
$$1221.40 = ab^{4} + d$$
$$1419.07 = ab^{7} + d$$

```
In[1]:= Solve[{1051.27 == a * b^1 + d, 1221.40 == a * b^4 + d, 1419.07 == a * b^7 + d},
          {a, b, d}]

Out[1]= {{d → 0.281732, a → -499.859 - 865.782 i, b → -0.525642 + 0.910438 i},
         {d → 0.281732, a → -499.859 + 865.782 i, b → -0.525642 - 0.910438 i},
         {d → 0.281732, a → 999.719, b → 1.05128}}
```

Figure 1.30: Using Mathematica to Solve for Three Parameters

Using the solution to the systems of equations in Figure 1.30, yields the final model:

$$Amount\,(year) = 999.719(1.05128)^{x} + 0.281732$$
$$Domain: \{year \mid 0 \le year \le 7\}$$
$$Range: \{Amount \mid 1000 \le Amount \le 1419.07$$

What if we encounter an example (there are many) where Mathematica cannot solve for three parameters simultaneously?

Let's assume a value for the b parameter to make the function *intrinsically linear* (in the form $ax+d$) and then solve for a and d. For this example, we see that the rate at which the function is increasing is not significant, so we could assume a value for b that is close to 1. Let's assume $b = 1.01$.

> A function is said to be ***intrinsically linear*** when it can be written in the form $f(x) = ax+d$. (Notice the form of the exponential function, after a 'b' parameter is estimated and values substituted in for representative points)
>
> $$1051.27 = a(1.01)^{1} + d \rightarrow 1051.27 = a(1.01) + d \rightarrow f(x) = ax + d$$

Use the second and last data points to form two equations to solve for the two unknown parameters:

$$1051.27 = a(1.01)^1 + d$$
$$1419.07 = a(1.01)^7 + d$$

```
Solve[{1051.27 == a (1.01)^1 + d,  1419.07 == a (1.01)^7 + d}, {a, d}]

{{a → 5919.34, d → -4927.26}}
```

Figure 1.31: Solving for Two Parameters in Mathematica

Using the solution to the systems of equations in Figure 1.31, yields the final model:

$$Amount(year) = 5919.34(1.01)^x - 4927.26$$
$$Domain: \{year \mid 0 \le year \le 7\}$$
$$Range: \{Amount \mid 1000 \le Amount \le 1419.07\}$$

Let's plot our exponential models, Figure 1.32, to see how well it appears to fit the data and how well it compares to the linear model we already plotted.

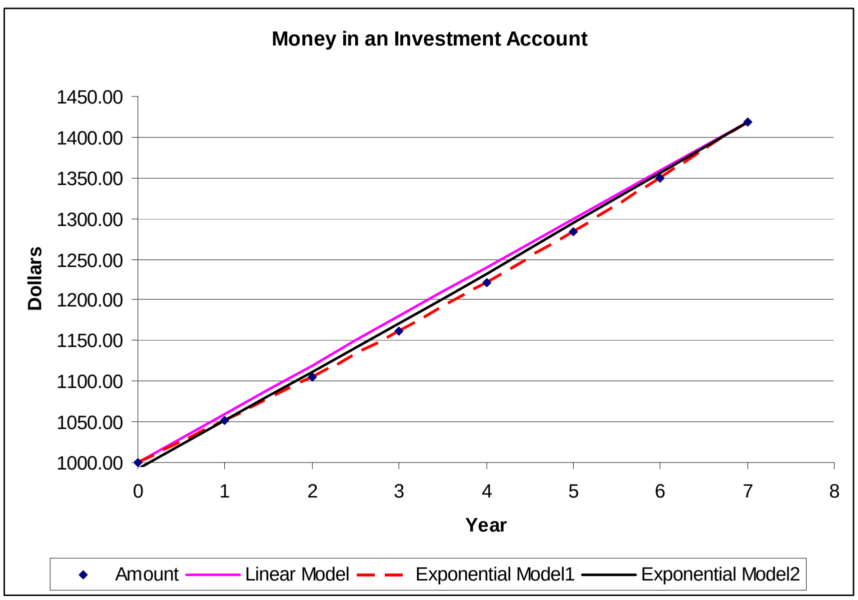


Figure 1.32: Investment Data Overlaid with Linear and Exponential Models

Step 3. *Interpret the results of the solution.* All models appear to fit the data relatively well, but which is best? Using the eyeball test, the exponential model1 seems to fit the data better than the estimated exponential model2 or the linear model. Because of the scale of the graph, we cannot be precisely sure that the eyeball test, a qualitative (subjective) measure is accurate enough in this case. We will verify our conjecture using each model's sum of squared error and coefficient of determination. See the results in Figure 1.33.

Year	Amount	Linear Model	Mean Error	Squared Mean Error	Model Error	Squared Model Error		
0	1000.00	1000.00	-209.55	43909.11	0.00	0.00		Linear
1	1051.27	1059.87	-158.28	25050.98	-8.60	73.96	a	59.87
2	1105.17	1119.74	-104.38	10894.14	-14.57	212.28	b	
3	1161.83	1179.61	-47.71	2276.72	-17.78	316.13	c	
4	1221.40	1239.48	11.86	140.54	-18.08	326.89	d	1000
5	1284.03	1299.35	74.49	5548.02	-15.32	234.70		
6	1349.96	1359.22	140.42	19716.37	-9.26	85.75		
7	1419.07	1419.09	209.53	43900.73	-0.02	0.00		
	Mean:	1209.55	**SST:**	151436.60	**SSE:**	1249.71		
					r^2:	0.99174763		
		Exponential Model1						
		1000.00	-199.06	39625.61	0.00	0.00		Expo1
		1051.27	-147.79	21842.43	0.00	0.00	a	999.719
		1105.16	-93.89	8815.68	0.01	0.00	b	1.05128
		1161.82	-37.23	1386.21	0.01	0.00	c	
		1221.38	22.34	498.99	0.02	0.00	d	0.281732
		1284.00	84.97	7219.59	0.03	0.00		
		1349.83	150.90	22770.26	0.13	0.02		
		1419.03	220.01	48403.59	0.04	0.00		
	Mean:	1199.06	**SST:**	150562.36	**SSE:**	0.02		
					r^2:	0.99999987		
		Exponential Model2						
		992.08	-203.45	41392.98	7.92	62.73		Expo2
		1051.27	-152.18	23159.56	0.00	0.00	a	5919.34
		1111.06	-98.28	9659.48	-5.89	34.68	b	1.01
		1171.44	-41.62	1732.44	-9.61	92.39	c	
		1232.43	17.95	322.11	-11.03	121.64	d	-4927.26
		1294.03	80.58	6492.71	-10.00	99.92		
		1356.24	146.51	21464.40	-6.28	39.42		
		1419.07	215.62	46490.84	0.00	0.00		
	Mean:	1203.45	**SST:**	150714.53	**SSE:**	450.77		
					r^2:	0.99700912		

Figure 1.33: Computation of Linear and Exponential SSE and r^2

As expected, the exponential models have a lower SSE and a higher r^2, and are therefore, a better fit to the data than the linear model. Exponential model1 is better than the exponential model2 which is also to be expected since we estimated the parameter b in the set-up of exponential model2.

Question 5 *Complete the* ***reflect*** *step of the modeling process for Example 1. Does the model you chose to have the best fit make sense based upon the number of parameters and your knowledge of the nature of investment data?*

Question 6 *Complete a sensitivity analysis of the two estimated models. Do your findings change significantly?*

Example 2 Modeling Prozac in the Bloodstream

Let's now analyze decay functions by considering one of the most widely taken drugs to treat depression, Prozac. If a person takes a certain single dose, it will eventually be eliminated from the bloodstream by the kidneys. We can assume that during a given fixed time period, the kidneys will remove a certain percentage of the drug from the bloodstream. In fact, it has been found that the kidneys remove one-fourth of the drug during any 24-hour period so that 75% of the drug will still remain.

It is unhealthy for two different antidepressants to work in the body at the same time. In fact, a person can have no more than 10mg of Prozac in their blood to safely begin another drug regimen. Let's assume that a person must change the prescription they are on from Prozac to another drug. The pharmacist tested the patient to determine how much Prozac is currently in the blood. The test revealed an amount of 60mg. Given this initial amount of Prozac in the blood, it is your job to advise the pharmacist when to safely prescribe the new drug.

Step 1: *Transform the problem.* We are **given** the rate at which Prozac is removed from the bloodstream (75% eliminated per day). Therefore, we know that our b parameter is $b = .75$...a decaying exponential function. We are also given the start point; at day zero, there is 60mg of Prozac in the blood.

- Define the variables:
  - Independent variable (input): time (days)
  - Dependent variable (output): amount of Prozac in the blood

- Assume: For this example, we will assume that the parameter $d = 0$ because after an infinite time period, the level of the drug in the blood will tend toward zero (i.e., there is a horizontal asymptote at $y = 0$ – recall earlier we mentioned that the d parameter equates to the horizontal asymptote). Therefore, our new exponential function becomes:

$$druglevel(t) = ab^t + 0 \qquad (4)$$

The goal of this problem is to **find** how many days it will take the Prozac to reach a safe level to administer the new drug, a level of 10mg. Our **plan** will be to develop a model for Prozac being eliminated from the blood stream. We can then iterate the function to see when the function drops below 10mg of Prozac in the bloodstream.

Step 2. Solve the problem using appropriate techniques. Given that the initial dose is 60mg, and our horizontal asymptote is at $t=0$, we are able to solve for the a parameter as follows:

$$\begin{aligned} druglevel(t) &= ab^t \\ druglevel(0) = 60 &= ab^0 \\ 60 &= a(1) \\ 60 &= a. \end{aligned}$$

And in general, the following function models the amount of Prozac in the bloodstream after t days,

$$druglevel(t) = 60(.75)^t \qquad (5)$$

$$\begin{aligned} &\text{Domain}: \{t \mid 0 \le t \le 50\} = [0, 50] \text{ (after 50 days amount of drug is negligible)} \\ &\text{Range}: \{druglevel \mid 0 \le druglevel \le 60\} = [0, 60]. \end{aligned}$$

Therefore, the drug after each 24-hour time period can be determined by

$$\begin{aligned} &druglevel(1) = 60(.75)^1 \\ &druglevel(2) = 60(.75)^2 = 60(.75)(.75) \\ &etc... \end{aligned}$$

***Question 7** *Using Equation (5), modeling the drug level in the bloodstream, determine how much remains at time=3, 4, and 5 days. Determine how many milligrams of Prozac your body metabolizes between each successive day for days 0 to 5. What do you notice about the differences between the previous and successive drug levels?*

***Question 8** *Utilizing the function above that models the amount of Prozac in the bloodstream, complete the following:*

a. *Find the amount of Prozac in the bloodstream after one week.*
b. *Estimate the half-life (the amount of time required to decrease the original amount by one half) of Prozac in the bloodstream.*
c. *Estimate how long it takes until the level drops to 10mg.*

***Question 9** *Equation (5) provided a function that yields drug level as a function of time. You were able to iterate to find the answer to Question 7. What other technique could we use to determine the day at which Prozac level reached 10mg? What answer did you get?*

Question 10 *Under ideal conditions, bacteria flourish in a given cadet's sink. However, it is Tuesday afternoon and you are preparing for WAMI. There are currently 1000 bacteria in your sink. Your friend from another regiment measured how bacteria remained in his sink after he sprayed it at 2400 hours the night before his last WAMI. Below is the table that he made.*

Hour	Bacteria
0	1000
1	500
2	250
3	125
4	62.5
5	31.25

a. How many bacteria would be left after 8 hours?

b. Given an inspection that starts at 0630, when would you have to spray to ensure there were no living bacteria in your sink at the beginning of inspection?

c. Your squad leader is a Chemistry major. He begins to inspect your room and sees that there are 5 bacteria remaining in your sink. When did you spray?

Question 11 *Go to the web or any other reference and find the population data of your favorite country (other than the US). Plot the data, and predict the population for the years 2010, 2020 and 3000. What type of model did you use? What assumptions did you make? Discuss how good you think your predictions are and why?*

Question 12 *The cost of a US first-class postage stamp was 29 cents in 1990 and was 39 cents in 2006; predict when the cost of this stamp will be $1 using an exponential model. **HINT** You may want to scale the years so you are not raising a number to the 1990 power.*

Question 13 *The population in Orange County, New York in 1990 was 307,647 and increased to 341,367 in 2000. Using this information and assuming an exponential model, what do you predict the population will be in 2010?*

THIS PAGE INTENTIONALLY LEFT BLANK

1.7 Linear and Exponential Data Fitting

1.7.1 The Nature of the Data

When information or data is collected, it frequently contains some amount of error. Therefore, data with an inherently linear relationship may not appear to fit this relationship perfectly. This can also be said of data which truly have attributes lending them to the use of an exponential model. In other words, although we have found the correct underlying relationship, sometimes our models may not appear to fit the data perfectly. Other times, we may select a model which appears to fit the data well but it isn't the best choice for prediction. This lesson will challenge you to think about the data so we can begin choosing the best type of model, based upon the nature of the data.

We've discussed the *nature of the data* previously in class. Our intuition regarding the *nature of the data* is that it would be what the data would look like if there was no randomness associated with the collection, as if we lived in a "perfect world." However, we know that there is some degree of randomness in almost everything that happens in the world. For example, a person may think that firing a bullet from a rifle exactly the same way will result in hitting the same spot on the target. But, how many times do two bullets go through the same hole? Not often. Even if we take the human completely out of the loop and fire the rifle from mounts on the ground, the bullets won't go through the same hole. Why? The answer is in the randomness that exists in the world, possibly: wind resistance, tiny abnormalities in the bullet affecting trajectory, wear on the rifle barrel, percentage of powder igniting, and many more.

The result is that the data we see most likely cannot be *perfectly* modeled with any reasonable model. But, we know that we can model the trajectory of the bullet, determining where it will strike, with a parabolic function (thanks to Sir Isaac Newton), because we understand that the nature of the data is a parabolic trend.

> The **nature of the data** consists of the underlying attributes of the data which describe the pattern it will take both within and outside of the collected region.

In addition to examining our knowledge of the nature of the data, we can apply quantitative tests to determine what we may expect the nature of the data to be.

- Is the rate of change of the data constant (or near constant)?
    - If yes, try a linear model.

- If not, then try an exponential model, unless the data follow a cycle, which we will address in later lessons.

1.7.2 Data Fitting

In section 1.7.1, we discussed using the nature of the data to determine what type of model we will use to model data we are given. It is now time to fully develop the model that we selected. Finally, we will plot our models against the data and apply the "eyeball," SSE, and r^2 tests to determine if a particular model is good or not.

Example 1 *In doing some research on Mexico, we come across the data in Table 1.11 regarding Mexico's population. We are interested in modeling the population as a function of time but need to determine which function will best fit the data.*[1]

Year	Population (millions)
2000	100
2001	102
2002	104.04
2003	106.12
2004	108.24
2005	110.41

Table 1.11: Population of Mexico

Step 1: *Transform the problem.* We are **given** a table of data, we must examine the table.

- Define the variables:
  - Independent variable (input): year
  - Dependent variable (output): population

- Nature of the data: Apply the quantitative measures of the nature of the data to gain an idea of which type of function may be most appropriate. See Table 1.12.

Mexico Population		
Year	**Population (Millions)**	**Rate of Change**
2000	100	
2001	102	2
2002	104.04	2.04
2003	106.12	2.08
2004	108.24	2.12
2005	110.41	2.17

Table 1.12: Quantitative Measure of the Nature of the Data

[1] Problem adapted from *Functions Modeling Change: a Preparation for Calculus*, 3rd Edition. Connally, Hughes-Hallett, Gleason, et al., 2007. p. 23.

Notice that the data are increasing, as is the rate; therefore, we need a function that is increasing at an increasing rate, modeled by an exponential function. Let's graph the data to verify our conjecture. See Figure 1.34.

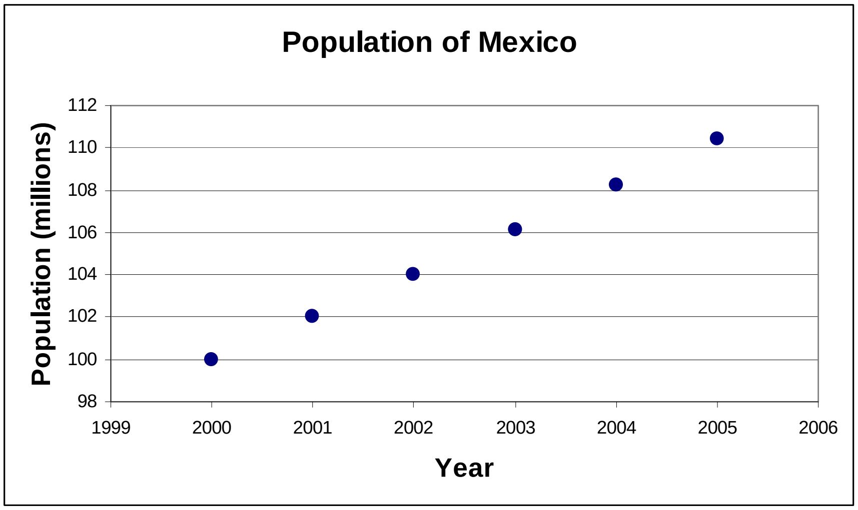


Figure 1.34: Scatterplot of Table 1.11 Data.

The data in Figure 1.34 seems to exhibit a *linear* pattern. But our quantitative measure of the nature of the data and our visual measure seem to conflict. We know we will have to **find** the model that fits through the data best.

- Assumptions
  - Based on the scatterplot in Figure 1.34, we will assume the plot of the data appears to increase at a constant rate. We'll assume a linear model might be a good choice.
  - Since the data do not increase in a perfectly linear manner (or, the rate of change would be constant), we'll also assume a generalized exponential model might model the data. We select this because we can develop a model that incorporates a bend in the data and a vertical shift.
  - The data accurately represent what will happen to the population in the coming years.

Our **plan** is to use the model development and evaluation techniques that we have used to this point in the course to find the best model to determine Mexican population in the future.

Step 2. *Solve the problem using appropriate solution techniques.* To **solve**, we will develop a linear and an exponential model, then evaluate the two to decide which is best.

- Linear model development. First, we'll try the linear model. In general, our linear model will be:

$$Population(year) = a(year) + d$$

- Linear Model Parameter Estimation: Using modeling skills developed to date, we need to estimate the parameters (a and d) for a linear model. To estimate two parameters, we need two data points that represent the data well.

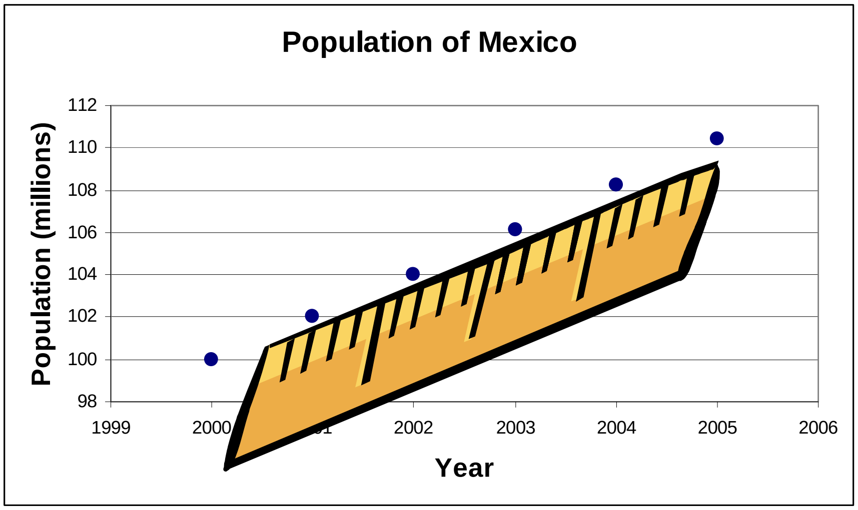


Figure 1.35: Determining Representative Data Points for Parameter Estimation

- The apparent linear trend of the data, seen in Figure 1.35, indicates that any of the data will represent it well; therefore, we'll select the first and last data points to estimate our parameters.

- Now, as we have many times in the past, we'll create two linear equations using the data points to estimate our parameter values. Our two equations are:

$$100 = a(2000) + d$$
$$110.41 = a(2005) + d$$

```
Solve[{100 == a * 2000 + d, 110.41 == a * 2005 + d}, {a, d}]

{{a → 2.082, d → -4064.}}
```

Figure 1.36: Using *Mathematica* to Solve Two Simultaneous Linear Equations

Using the solution from Figure 1.36 yields the final model:

$$Population(year) = 2.082(year) - 4064$$
$$Domain : \{year \mid 2000 \le year \le 2005\}$$
$$Range : \{population \mid 100 \le population \le 110.41$$

The model is plotted in Figure 1.37; by visual inspection, it fits the data well.

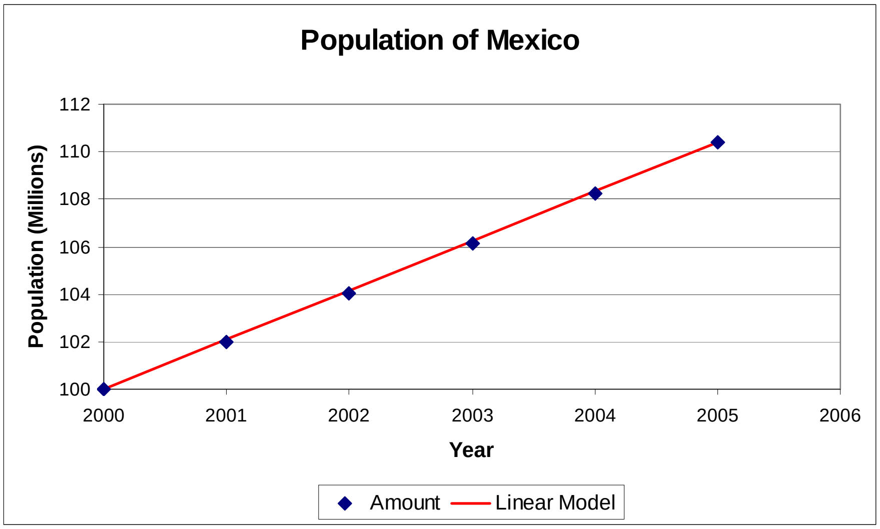


Figure 1.37: Mexican Population Data Overlaid with Linear Model

- Exponential model development. Now that we've developed a linear model, it is time to develop our generalized exponential model, of the form:

$$Population(year) = a(b)^{year} + d$$

- As we saw earlier in the example, each of the points seems representative. It make sense to solve for all three parameters if we can, so we'll use the second, fourth, and sixth points to form three equations in three unknowns and let *Mathematica* do the work for us. See Figure 1.38 for the results.

- **NOTE** The independent variable in the equation (year) is scaled by subtracting 2000 from each actual year value. It is common practice to scale the independent variable when working with exponential functions because raising a number to the 2000th power is very often problematic.

```
In[1]:= Solve[{102 == a * b^1 + d, 106.12 == a * b^3 + d, 110.41 == a * b^5 + d}, {a, b, d}]

Out[1]= {{d → 2.15059, a → -97.851, b → -1.02042}, {d → 2.15059, a → 97.851, b → 1.02042}}
```

Figure 1.38: Using *Mathematica* to Solve for Three Parameters

- Since the b parameter needs to be greater than 1 (to ensure an increasing exponential function), we choose the second solution returned by *Mathematica*: $a = 97.851$, $b = 1.02042$, and $d = 2.15059$.

Using the second solution from Figure 1.38, yields the final model:

$$Population(year) = 97.851(1.02042)^{year} + 2.15059$$
$$Domain: \{year \mid 0 \le year \le 5\}$$
$$Range: \{population \mid 100 \le population \le 110.41$$

The exponential model is plotted in Figure 1.39, to see how well it appears to fit the data.

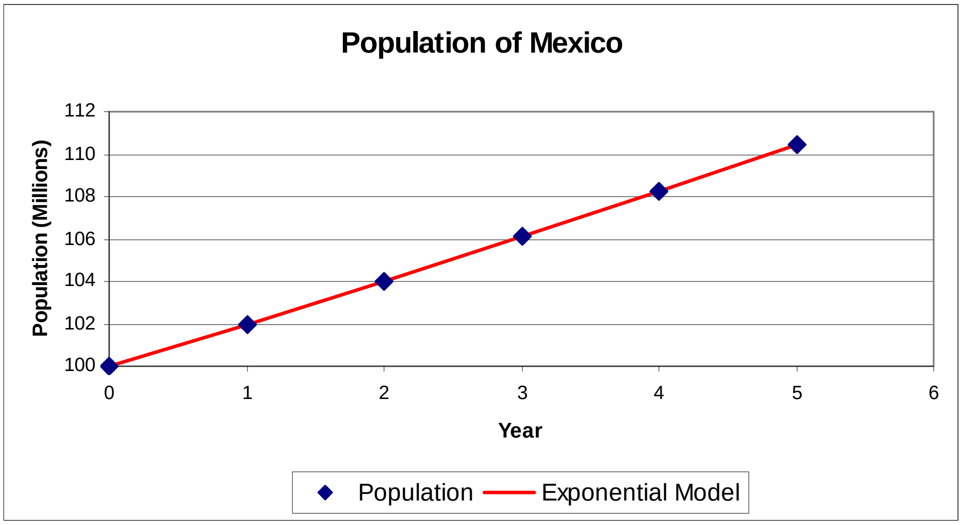


Figure 1.39: Mexican Population Data Overlaid with Exponential Model

Step 3. *Interpret the results of the solution.* The estimated linear model certainly seems to fit the data as well as the estimated exponential model. Overall, the linear model seems to be our best model – it fits almost as well as the exponential and has one fewer parameter (it's simpler). Though it may not seem necessary to compute sum of squared error and r^2 in this case, we do it anyway for two reasons.

- Quantitative, objective justifications are useful to back-up qualitative, subjective observations.
- Valuable practice in the computation of SSE is useful for cases that are not so clear cut.

Recall from Section 1.5 that we compute SSE by using the following formula:

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^{\,2} .$$

Derive the error at each point by subtracting the predicted value from your model from the actual value of the data. Square each of these values, summing these squared provides SSE. See the results in Figure 1.40.

Year	Population	Linear Model	Squared Mean Error	SST	Model Error	Squared Model Error	SSE			
2000	100	100	26.368225	75.83	0.00000000	0.00000000	0.04572000		a	2.082
2001	102	102.082	9.828225		0.08200000	0.00672400			b	
2002	104.04	104.164	1.199025		0.12400000	0.01537600	**R-Sq**		c	
2003	106.12	106.246	0.970225		0.12600000	0.01587600	0.99939709		d	-4064
2004	108.24	108.328	9.641025		0.08800000	0.00774400				
2005	110.41	110.41	27.825625		0.00000000	0.00000000				
Mean:	105.14									

Year	Population	Exponential Model	Squared Mean Error	SST	Model Error	Squared Model Error	SSE			
0	100	100.00159	26.368225	75.83	0.00159000	0.00000253	0.00001203		a	97.851
1	102	101.9997074	9.828225		-0.00029258	0.00000009			b	1.02042
2	104.04	104.0386264	1.199025		-0.00137360	0.00000189	**R-Sq**		c	
3	106.12	106.1191801	0.970225		-0.00081990	0.00000067	0.99999984		d	2.15059
4	108.24	108.2422187	9.641025		0.00221871	0.00000492				
5	110.41	110.4086098	27.825625		-0.00139023	0.00000193				
Mean:	105.14									

Figure 1.40: Computation of Linear and Power Sum of Squared Error

As anticipated through visual inspection, the linear and exponential models have similar SSE and r^2. In fact, the linear model accounts for nearly 99.94% of the variation in the data whereas the exponential model accounts for greater than 99.99%.

The nature of the data would appear to support an exponential model, but when we use the other tests at our disposal; we realize that we can use a simpler, two-parameter linear model instead of the more complicated three-parameter exponential model. So, it seems both of our models are very good. So, which do we use? Good question! Let's see if we can determine what to do as we **reflect**.

We have determined the following: The linear model would probably prove useful in extrapolating for a short time period either prior to 2000 or after 2005 but wouldn't be useful for the long-term, because it does not fit the nature of the data. The problem is identifying an appropriate domain for our model. How far out can we consider the linear model useful?

This question is an important question that highlights the fact that modeling has components that are "science," as in the development and evaluation of the models, and "art," as in the interpretation of the models and their relevance over certain domains. In fact, our final model may be a piecewise function that enables us to choose different functions over different ranges of the domain.

Question 1 *In Example 1, over what domain would a linear function be most appropriate? An exponential function?*

Question 2 *The following table contains data on the population of two countries in millions. One of the countries experiences relatively constant growth between 1950 and 2000 while the other does not.*[2]

a. *Develop two models (one linear and one generalized exponential model) to predict each country's population over time. Outline your problem solving process ensuring you include your assumptions and parameter estimation process. Discuss the fit of your models.*

b. *Which of the two countries exhibits non-constant growth?*

c. *Using the model you believe to be best for each country, estimate the populations in 1993.*

Year	1950	1960	1970	1980	1990	2000
Country A	8.2	9.8	12.4	15.1	14.7	23.9
Country B	7.5	9.9	12.5	14.9	17.2	19.2

Population Data for Countries A and B

Question 3 *Go back to one of the models that you have worked with in class, the one you found most interesting. Estimate the parameters of the model to achieve a better fit. Communicate your results in terms of the steps you took to develop the model, evaluate the model, and conduct a sensitivity analysis.*

[2] Problem adapted from *Functions Modeling Change: a Preparation for Calculus*, 3rd Edition. Connally, Hughes-Hallett, Gleason, et al., 2007, p. 25.

Challenge Question.

In Section 1.4, we considered data pertaining to the launch of the Space Shuttle Challenger. This data, on the percent expansion of an o-ring, appeared to indicate a linear relationship between the surrounding temperature and the percent expansion. Unfortunately, our attempts to model this data failed to identify that the space shuttle should not have launched.

Now, let's consider the *nature of this data*. O-rings are made of rubber (picture an o-ring as a rubber band). When heated, rubber expands allowing it to stretch beyond the expansion it would be capable of at room temperature but with a limitation - eventually, the rubber will break. When cooled, rubber becomes brittle and is capable of very little expansion before breaking. Therefore, if we were to describe an o-ring's expansion based on the temperature of its surroundings, we would realize that at low temperatures, the o-ring would not be capable of much expansion. As the temperature increased, the o-ring would expand quickly but, over the long term, the o-ring would reach its maximum expansion and would eventually break. Graphically, o-ring expansion would resemble something like the graph shown in Figure 1.41.

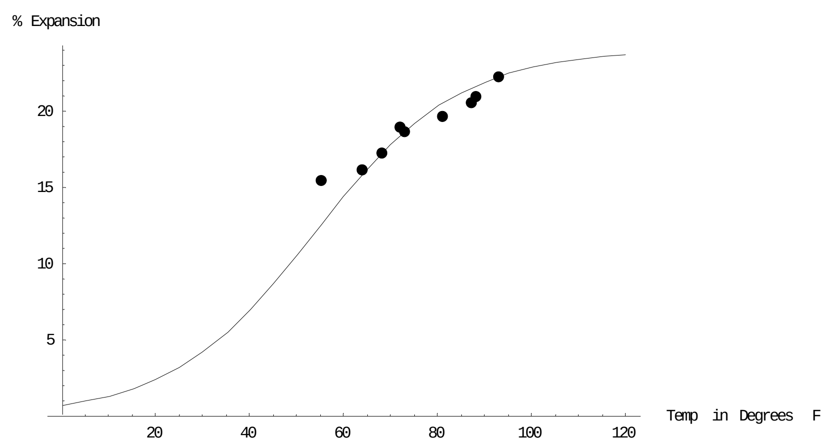


Figure 1.41: Nature of the Data: O-ring Expansion

Develop, fully evaluate, and communicate a model that would effectively represent the nature of the data illustrated in Figure 1.41 and predict the explosion of the Challenger.

THIS PAGE INTENTIONALLY LEFT BLANK

1.8 Generalized Power Function

Power functions can be extremely useful models that describe many interesting phenomena in our world. Shortly after the apple fell on his head, Sir Isaac Newton discovered that gravity could be represented and modeled using a power function, namely a quadratic function of the form $y=x^2$.

Recall our discussions of rates of change in sections 1.3 and 1.6. Up until now, if the nature of the data we were presented with demonstrated a constant rate of change, we could use a two-parameter linear model. If the data have curves in them (non-constant rate of change), we could use a three-parameter exponential function to model the data. In this section, you will learn about another model that is useful in modeling data with curves – the four-parameter power function.

In the linear case, we saw that the rate of change was constant and its sign determined if the line was increasing or decreasing. We will see many examples of power functions that increase or decrease and their rates of change that also increase or decrease. Figure 1.42 catalogs the four possibilities of power functions and their corresponding changing rates. Each of the pieces of the black circle represents a piece of a function. The adjacent box explains the behavior that is graphically depicted by the corresponding segment, i.e., how the function and its rate of change are changing. Use this diagram as a tool to help analyze new functions throughout this text. **NOTE** This is the same figure as Figure 1.26 in Section 1.6 on exponential functions.

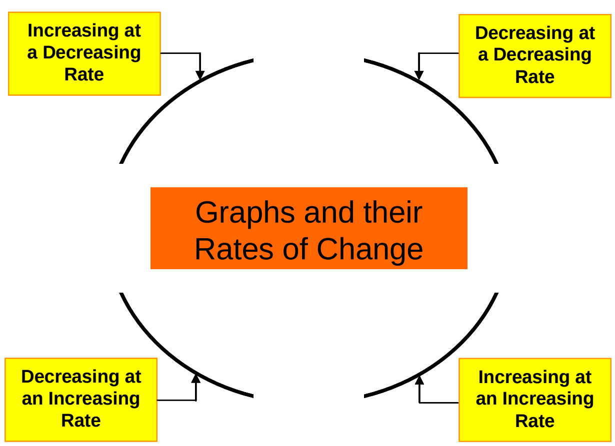


Figure 1.42: Graphs and rates of change.

1.8.1 Three Cases of Power Functions

We will consider three different cases of the power function

$$f(x)=x^b \qquad (1)$$

where x is the independent variable and b is a parameter that determines the shape of the power function. Notice there is a distinct difference between the shapes of the graphs when b is even or odd. We will show each of these two standard functions in each of the three cases. There are many forms of the power function other than the three that we will discuss, but all power functions are characterized by the *independent variable* being raised to a constant power.

Case 1 $b = n$, **where** n **is a positive integer**

The parabola

$$f(x) = x^2$$

is a power function. As characterized above, our independent variable is raised to a constant power (in this case, the power is 2).

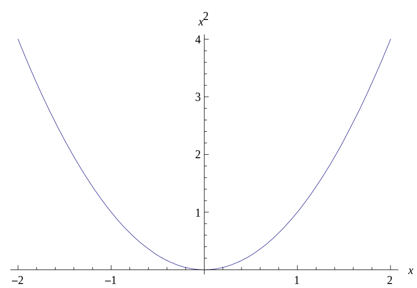


Figure 1.43: Plot of Power Function x^2

Another example of Case 1 is the cubic function

$$f(x) = x^3$$

which has parameter $b = 3$.

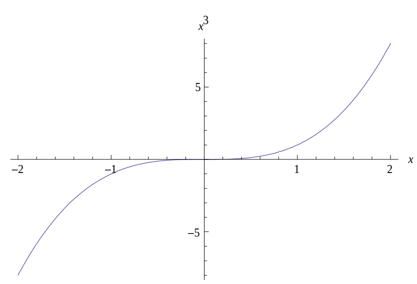


Figure 1.44: Plot of Power Function x^3

Note that if $b = 1$ we have the function $f(x) = x$, which is linear. Therefore, the linear function is a particular type of power function.

Question 1 *Plot x^4 by hand. How do the plots of x^2, x^3, and x^4 compare to each other? What can you conclude about the effect of the b parameter on Equation 1?*

Question 2 *What are the domain and range of functions that fall into Case 1?*

Question 3 *What can you say about the rates of change of the functions in Case 1 ?*

Question 4 *The kinetic energy (KE) from a moving billiard or pool ball can be measured by the relationship:*

$$KE = \frac{1}{2}mv^2$$

a. *Given a constant mass, what does doubling the velocity (v) do to the amount of kinetic energy?*

b. *How could you describe the rate of change of this function?*

An example of a function in Case 1 is the relationship for the area of a circle with radius *r*, $A = \pi r^2$. This relationship is an example of a power function where b=2, the radius, *r,* is the independent variable, and A or $A(r)$ is the dependent variable.

Thus far, we've considered the case where our exponent is a positive integer. We will discuss the two other cases that may prove useful in modeling.

Case 2 $b = 1/n$, **where** n **is a positive integer** (e.g., $f(x) = x^{1/2} = \sqrt{x}$)

This type of power function is frequently referred to as a **root function.** *Root functions* involve taking a root of the *independent variable* and exhibit one of the two general shapes shown in Figures 1.45 and 1.46.

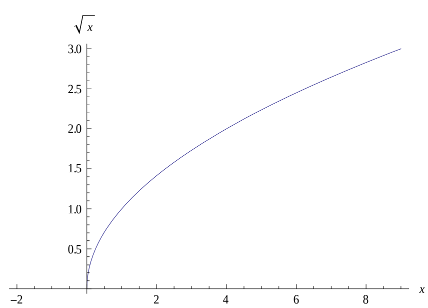


Figure 1.45: Plot of Power Function $\sqrt{x} = x^{\frac{1}{2}}$

Power functions whose n values are even numbers like 2, 4, and 6 are called **even root functions**. The domains of even root functions are limited to $[0, \infty)$.

Question 5 *How do the plots of* $\sqrt{x}$, $\sqrt[4]{x}$, *and* $\sqrt[6]{x}$ *compare to each other? What are your conclusions with respect to their rates of change?*

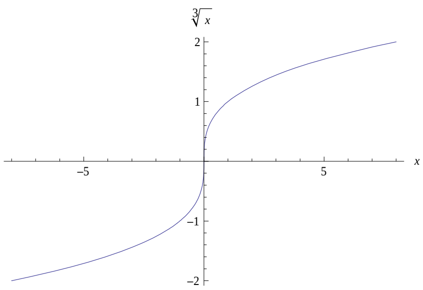


Figure 1.46: Plot of Power Function $\sqrt[3]{x} = x^{\frac{1}{3}}$

Power functions whose n values are odd numbers like 3, 5, and 7 are called **odd root functions**. The domains of odd root functions consist of all real numbers.

Question 6 *How do the plots of $\sqrt[3]{x}$, $\sqrt[5]{x}$, and $\sqrt[7]{x}$ compare to each other? What are your conclusions with respect to their rate of change?*

Case 3 $b = n$, **where *n* is a negative integer**

This is known as a **reciprocal function**. Its characteristic shape is a hyperbola as shown in Figures 1.47 and 1.48.

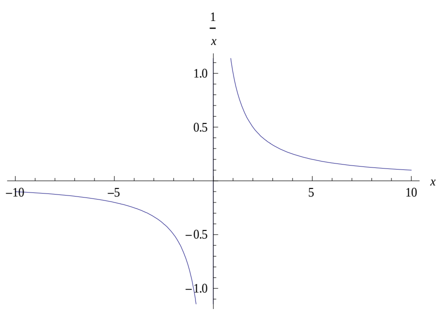


Figure 1.47: Plot of Power Function $\dfrac{1}{x} = x^{-1}$

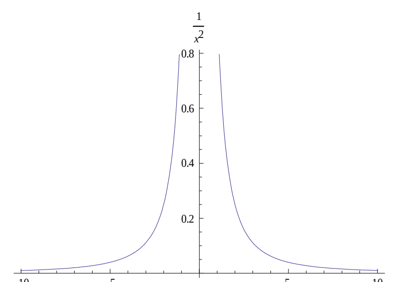


Figure 1.48: Plot of Power Function $\dfrac{1}{x^2} = x^{-2}$

Each of these two graphs have horizontal and vertical asymptotes that separate it from the previous cases. The vertical asymptote is the *y*-axis or the line $x = 0$.

The horizontal asymptote for each of these functions is the *x*-axis or $y = 0$.

Question 7 *What are the domains and ranges of Case 3 power functions?*

Question 8 *How do the plots of $\dfrac{1}{x}$, $\dfrac{1}{x^3}$, and $\dfrac{1}{x^5}$ compare to each other? What can you say about the rates of change of these functions on the different intervals in their domains?*

Question 9 *How do the plots of* $\frac{1}{x^2}$, $\frac{1}{x^4}$, and $\frac{1}{x^6}$ *compare to each other? What can you say about the rates of change of these functions on the different intervals in their domains?*

This function below is often used with a limited domain to describe real world phenomena. It is Boyle's Law from chemistry and physics which states that for a constant temperature, *C*, the volume of gas, *V*, is inversely proportional to the pressure, *P*.

$$V(P) = C^{\circ} P^{-1} = \frac{C^{\circ}}{P}$$

Question 10 *Using Figure 1.42, the circle diagram, give an example of a power or polynomial function and its domain that satisfies each of the four cases.*

1.6.3 The Generalized Power Function

Regardless of the type of function we're working with (our readings have covered the *linear, exponential, and power families),* functions contain *parameters* or constant values that influence the shape and location of their graphs. Experimentation with graphing in *Mathematica* should have resulted in some generalities regarding the *b* parameter for power functions. However, most data that follows a power function trend does not go through the origin, so we find modeling with the form in equation (1) to be too restrictive. We use the more general form

$$f(x) = a(x+c)^b + d$$

where *a*, *b, c*, and *d* are parameters that can assist us in modeling data that follows a power trend. We call this function the *generalized power function*.

Question 11 *Begin with the function* $f(x)=ax^2$. *Select a positive and negative whole number and decimal value for the a parameter. What effect can you conclude that this parameter has upon the generalized power function?*

Question 12 *Begin with the function* $f(x)=(x+c)^2$. *Select a positive and negative whole number for the c parameter. What effect can you conclude that this parameter has upon the generalized power function?*

Question 13 *Begin with the function* $f(x)=x^2+d$. *Select a positive and negative whole number for the d parameter. What effect can you conclude that this parameter has upon the generalized power function?*

When b is a non-negative integer (Case 1), then the power function $f(x)$ is also called a polynomial. A **polynomial** function is a linear combination of power functions with non-negative integer powers. They can all be written in the form:

$$f(x) = a_b x^b + a_{b-1}x^{b-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

where b is a non-negative *integer*. The **coefficients** or numbers appearing before each term, $a_0, a_1, a_2, \ldots, a_b$, are constants within the polynomial.

For example,

$$f(x) = 3(x-6)^3 + 4 \;=\; 3x^3 - 54x^2 + 324x - 644$$

after expanding the polynomial. So, the coefficients of the polynomial are $a_3 = 3, a_2 = -54, a_1 = 324,$ and $a_0 = -644.$ However, not all polynomials can be written in the generalized power function form. For example,

$$f(x) = x^3 + x^2 + x + 1$$

cannot be written in the form $f(x) = a(x+c)^b + d$. There are many applications that can be modeled using polynomials, but we will restrict our work to this generalized power function for polynomials. We do this because the general polynomial

$$f(x) = a_b x^b + a_{b-1}x^{b-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

has $b+1$ parameters and would require $b+1$ points to find each parameter. Solving for so many parameters is time consuming and often leads to low predictive power (recall our discussion in section 1.5).

1.8.4 The Role of Parameters

Understanding the role each parameter plays in our families of functions contributes to our ability to create useable *mathematical models*. Initially, understanding function *parameters* assists in the selection of an appropriate function family. Next, knowledge of *parameters* enables us to improve our *mathematical model* which ultimately results in a model that can be used for prediction.

When modeling using a power function, the first step is to choose an appropriate b value based on the general shape of the curve. After estimating or assuming b, it is important to estimate the other parameters. First, consider the role of the d parameter. For linear functions, it describes the y-intercept. If we assume that $x = 0$ in the generalized power function, we get $y = ac^b + d$, so d is not the y-intercept for a generalized power function (unless a or c is also 0). Look at the graphs in Figures 1.49 through 1.51 to see how the d parameter changes them.

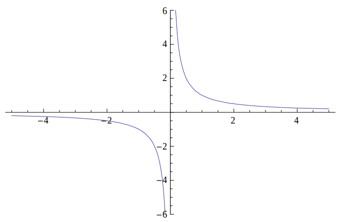


Figure 1.49: Plot of Power Function $\frac{1}{x}$

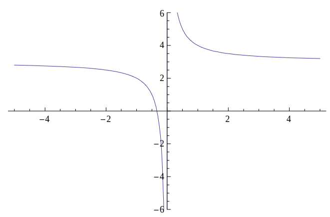


Figure 1.50: Plot of $\frac{1}{x}+3$

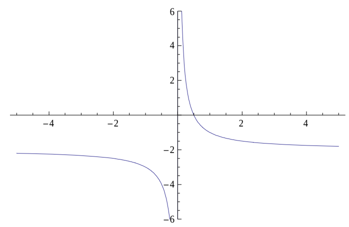


Figure 1.51: Plot of $\frac{1}{x}-2$

Generalization 1. We can shift any graph of a function upward or downward by adding a parameter to the function:

$$y = f(x) + d$$

If d is positive, we will shift or translate the function's graph upward. If d is negative, we will shift or translate the function's graph downward. ****NOTE**** This is also the same first generalization we made for the linear and exponential functions.

Next, consider the c parameter. In Figures 1.52-1.54, we see that this parameter creates a horizontal shift as opposed to the vertical shift of the d parameter. In these figures, $f(x)=\sqrt{x},\ f(x+2)=\sqrt{x+2},\ f(x-3)=\sqrt{x-3}$. You may expect that a positive value of c will move the function to the right and a negative to the left. Notice that the actual shift is opposite: positive c values shift the function left and negative c values shift the function right.

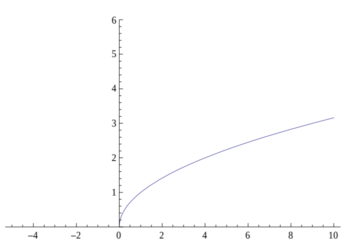


Figure 1.52: Plot of $\sqrt{x}$

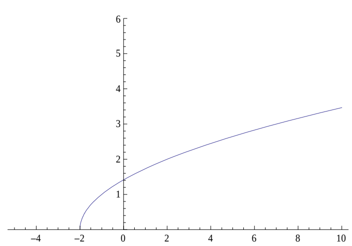


Figure 1.53: Plot of $\sqrt{x+2}$

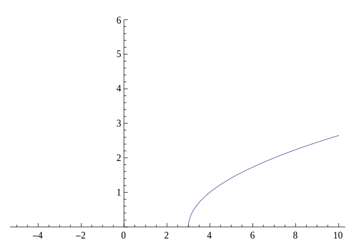


Figure 1.54: Plot of $\sqrt{x-3}$

Generalization 2. We can shift any graph of a function to the right or left by adding a parameter to the function as:
$y = f(x+c)$

If c is positive, we will shift or translate the function's graph to the left. If c is negative, we will shift or translate the function's graph to the right.

Estimating the a parameter in $f(x)=a(x+c)^2+d$ is not as easy as the other three parameters. We often estimate the b parameter based on shape (s-curve versus u-shaped, …), then estimate the c and d parameters using horizontal and vertical shifts, respectively (often using the coordinates of the first data point). We can then solve for a using a representative point from the data that is different from the first data point.

Question 14 *What happens if you change the 'a' parameter? Like in the above examples, experiment with different values using Mathematica.*

Example 1 **Case 1 of the Power Function.** *Suppose a paratrooper jumps from a helicopter at a height of 1000 meters. During his descent, he radios you and reports altitude. Five seconds into his descent, he is at an altitude of 878 meters.*

Sir Isaac Newton described freefall motion by a quadratic (raised to the second power) power function. Using the data provided by the paratrooper, determine the specific power function that describes the paratrooper's descent.

REAL WORLD PROBLEM: Model the paratrooper's descent using a generalized power function.

Step 1. Transform problem into a mathematical model.

a. **Given:** Generalized power function: $Alt(t) = a(t+c)^b + d$, $b = 2$
The points $(0,1000)$ and $(5,878)$.
Variable declaration: $Alt(t)$ is the altitude of the paratrooper after t seconds.

b. **Find:** Parameters *a, b* (given), *c, d:* the final model of *Alt (t).*

c. Solution **Plan**: Use parameter estimation techniques to find *b, c*, and *d* and then use a representative point to find *a*.

Step 2. **Solve** using appropriate solution techniques (algebraic manipulation).

A good estimation for the *c* and *d* parameters are the *x* and *y* coordinates of the first data point, respectively. Therefore, assume $c = 0$ and assume $d = 1000$. Our new function looks like

$$Alt(t) = a(t+0)^2 + 1000 = at^2 + 1000.$$

To solve for *a*, we use another representative data point. Let's use the second point $(5,878)$.

$$Alt(5) = a(5s)^2 + 1000m = 878m$$

$$a = -4.88\frac{m}{s^2}$$

$$Alt(t) = -4.88t^2 + 1000$$

The domain of this model has a time restriction. Because it doesn't make sense to model earlier than our first recording, we will limit the lower bound of the independent value to zero seconds. Furthermore, we need to limit the upper bound on our domain value because eventually the jumper will hit the ground. Once this happens, time will continue but our model will no longer make reasonable predictions. To see this, let's try it. What if we wanted to see what our model predicts 60 seconds into the paratroopers flight?

$$Alt(60) = -4.88(60)^2 + 1000 = -17568 + 1000 = -16568m$$

So, for what range of t should our model be valid? It should work until the jumper hits the ground or has an altitude of zero. Using this information, we can algebraically solve for the time when this happens.

$$0 = -4.88(t)^2 + 1000$$

$$4.88t^2 = 1000$$

$$t^2 = 204.82$$

$$t = \pm 14.315 \sec$$

Where does this leave us? We previously established that the lower bound of our domain is zero, eliminating the -14.315 seconds. Therefore, our domain consists of all times between 0 and 14.315 seconds and our range consists of values between 0 meters and 1000 meters.

Step 3. Communicate and **reflect** upon results.

The paratrooper fell 1000 meters in 14.315 seconds according to our model. This means that the paratrooper fell at nearly 70 meters per second on average or about 157 miles per hour. This may seem fast, but this is close to the falling speed due to gravity.

Question 15 *What luck! It turns out the paratrooper's altimeter is digital and it recorded his jump. The collected data are in Table 1.13 and are linked on the course website. Using two new data points develop another function modeling the paratrooper's decent. How does your new model compare to your old? Which is better? Why?*

Time (sec)	0	1	2	3	4	5	6	7	8	9	10
Altitude (m)	1000	995.01	978.57	953.6	917.3	873.96	810.77	741.92	660.8	576.5	456

Table 1.13: Altimeter Data on Paratrooper Jump

Example 2 Case 2 of the Power Function. *Data have been collected on weightlifters' body weights and their corresponding squatting, benching and dead-lifting strengths. We'd like to model weightlifter benching strength as a function of body weight. The data is provided in Table 1.14 and are linked on the course website.*

Weight	Squat	Bench	Deadlift
114	330	180	415
123	415	245	425
132	495	285	540
148	535	335	584
165	610	390	660
181	670	473	725
198	700	500	730
220	722	510	705
242	720	529	755
275	730	525	700

Table 1.14: Weightlifting Data

REAL WORLD PROBLEM: Model Weightlifting data

Step 1. Transform problem into a mathematical model.

a. **Given:** Data on weightlifters ranging in weight from 114 to 275 pounds. A plot of the data in Figure 1.55 suggests the shape of an inverted parabola or perhaps a root function.

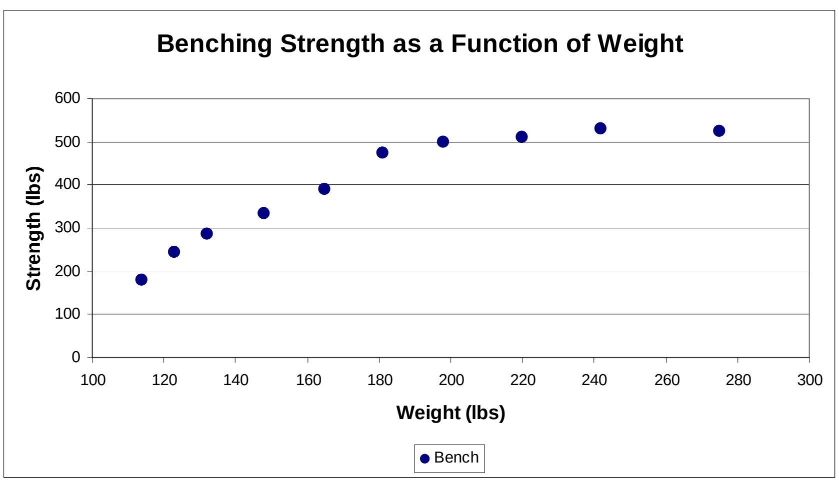


Figure 1.55: Graph of Weightlifting Data

b. **Find:** We are trying to model how much a weightlifter can bench given his or her body weight. In essence, we would like to use our independent variable, *Weight,* to predict our dependent variable, *Benching Strength.*

c. **Assume:**

Question 16 *What must we assume to solve this problem?*

d. Solution **Plan**: Since we are trying to model *Benching Strength* as a function of *Body Weight, Body Weight* is the independent variable and *Benching Weight* is the dependent variable. In general, because our data exhibit a curve that increases at a decreasing rate, our model will be:

$$Strength(Weight) = a(Weight + c)^{\frac{1}{n}} + d$$

Next, we'll need to estimate values for the parameters (a, $b = 1/n$ c, and d) of our *root-function power model*. We can use three representative points to find the a, c, and d parameters after making an assumption about the b parameter. Because we know the b parameter must be between zero and one, let's assume $b = \frac{1}{2}$.

Step 2. **Solve** using appropriate solution techniques (algebraic manipulation).

We assumed that $b = \frac{1}{2}$, now we need three representative data points to estimate the a, c and d parameters. Using the second, fifth and eighth data points yield the following three equations:

$$245 = a(123 + c)^{1/2} + d$$

$$390 = a(165 + c)^{1/2} + d$$

$$510 = a(220 + c)^{1/2} + d$$

Thus, after using *Mathematica* to solve for a, c and d, our model is

$$Strength(weight) = 39.6363(weight - 107.702)^{1/2} + 89.9702 .$$

Its domain and range are:

$$Domain: \{weight \mid 114 \le weight \le 400\} = [114, 400]$$

$$Range: \{strength \mid 180 \le strength \le 785.359\} = [180, 785.359]$$

Certainly, these numbers are not exact but we are trying to identify that our model has limitations. Here, we are saying that we believe our model is valid for individuals weighing between 114 and 400 lbs since those weighing less or more are unlikely to be lifters. Also, we are saying there are limits on the amount a person can bench which depend on our domain constraints. Perhaps you feel this upper value is high and would like to see it be lower. That's ok.

Sometimes models don't have exact bounds on their domains and ranges or in the course of study you realize revision to your initial estimates is necessary.

Step 3. Communicate and **reflect** upon results.

Our calculations show that our model predicts our second, fifth and eighth data points. This should be the case since these are the three data points we used to solve for the parameters. Notice that the model goes through these points in Figure 1.56.

Now that we have developed a model, we can use it to verify weightlifting strength within our data set; recall that we referred to this as *interpolation*. Often, a model which predicts well within the data is also useful outside of the collected range of data; recall that we referred to this as *extrapolation*. Therefore, we should use our model to predict a data point within the data and also one outside of the data to test how useable our model appears to be.

$$Strength(165) = 39.6363(165 - 107.702)^{\frac{1}{2}} + 89.9702 = 389.999$$

$$Strength(350) = 39.6363(350 - 107.702)^{\frac{1}{2}} + 89.9702 = 706.946$$

Our model appears to predict fairly well for the given data but appears too high for prediction outside of the range of our data. Figure 1.56 is a plot of our data with our model overlaid on the same axes.

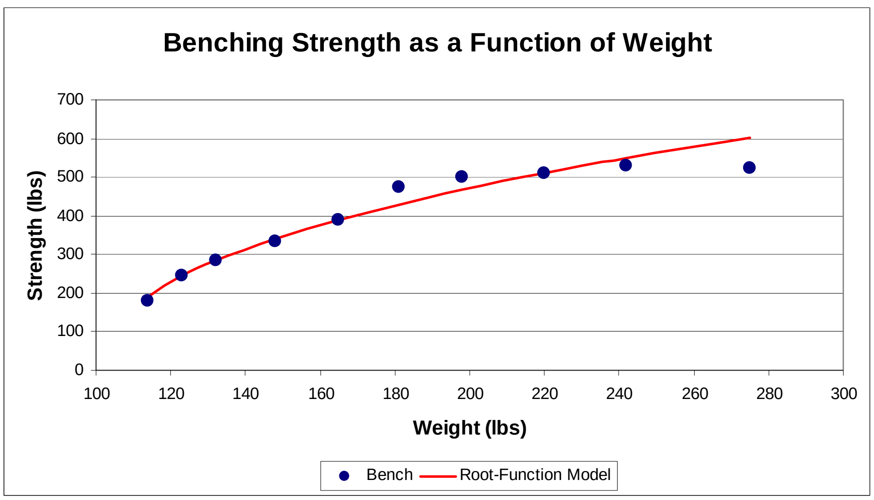


Figure 1.56: Weightlifting Problem – Plot of Model and Data

Question 17 *Based on your weight, what does this model expect you to bench?*

Question 18 *Based on the amount you can bench, what does this model predict you should weigh?*

Question 19 *Do a sensitivity analysis on our assumption that* $b = \frac{1}{2}$ *to see how this affected our model. Is the new b value a better representation of the data?*

1.9 Exponential and Power Function Data Fitting

This section will further explore the methodology behind modeling a data set with exponential and power functions. We will explore cases where we have no insight into the data as well as cases where the problem statements allow us to gain some predictive knowledge as to which model may fit the problem best.

Example 1 *Assume that you are a data analyst for a research company. Your company receives contracts from numerous outside agencies that require an unbiased evaluation of data sets that they have collected. In order to remain completely unbiased, you specify that you want to see their raw data – and only their raw data. You are totally unaware of what agency you are working with, unaware of the experiment that the data was collected from, and unaware of any preconceived ideas of the model with which the contracted agency would like the data modeled. The data you receive is represented in Table 1.15.*

x	$f(x)$
-2	0.02
-1	0.08
0	1.7
1	2.9
2	4.2
3	7.4
4	12.5
5	21.4
6	39.0
7	57.2
8	88.1

Table 1.15: Example 1 (Raw Data)

Step 1: *Transform the problem.* We are **given** a table of data, we must examine the table.

- Define the variables:
  - Independent variable (input): x
  - Dependent variable (output): $f(x)$

- Nature of the data: Apply the quantitative measures of the nature of the data to gain an idea of which type of function may be most appropriate. See Table 1.16.

x	*f(x)*	AVG RoC
-2	0.02	
-1	0.08	0.06
0	1.7	1.62
1	2.9	1.2
2	4.2	1.3
3	7.4	3.2
4	12.5	5.1
5	21.4	8.9
6	39.0	17.6
7	57.2	18.2
8	88.1	30.9

Table 1.16: Quantitative Measure of the Nature of the Data

Notice that the data are increasing, as is the rate. Therefore, we need a function that is increasing at an increasing rate. We may model these data by either an exponential or a power function. Let's graph the data to verify our conjecture. See Figure 1.57.

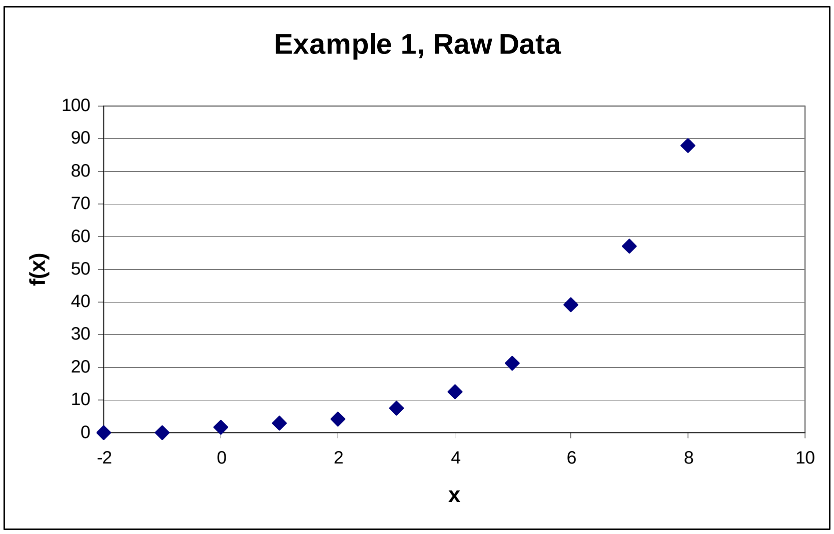


Figure 1.57: Scatterplot of Table 1.12 Data.

The data in Figure 1.57 seems to exhibit a pattern consistent with either an exponential or a power function. We know that we will have to **find** the model that fits through the data best. Because we have no knowledge as to what the data may do outside of the domain of the data in the given table, it would be wise to find a "good" exponential model **AND** a "good" power model. By completing both models, we can recommend to the agency that contracted our services which model fits the experiment from which the data came.

- Assumptions
  - Using an exponential function model of the form $f(x) = ab^x + d$, we can assume that the parameter $d = 0$ because there appears to be a horizontal asymptote at $f(x) = 0$.
  - Based on a power function model of the form $f(x) = a(x+c)^b + d$, a reasonable assumption seems to be that the vertex is at the first data point. Because the first data point is left of the origin by two units we will assume $c = 2$. Since the function value at the first data point is 0.02, we will assume that $d = 0.02$.

Our **plan** is to use the model development and evaluation techniques that we have used up to this point in the course to find the best models (exponential and power) for the data in Table 1.16.

Step 2. *Solve the problem using appropriate solution techniques.* To **solve**, we will develop a "good" exponential and a "good" power model. We will provide both options to our client.

- Exponential model development.

$$Model(x) = ab^x$$

- Exponential Model Parameter Estimation: Using modeling skills developed to date, we need to estimate the parameters (a and b) for an exponential model. To estimate two parameters, we need two data points that represent the data well. Let's try the fourth and tenth points: $(1, 2.9)$ and $(7, 57.2)$.

- Using Mathematica to solve for a and b yields,

```
Solve[{2.9 == a (b)^1, 57.2 == a (b)^7}, {a, b}]

{{a → -1.76427, b → -1.64374}, {a → -0.882135 - 1.5279 i, b → -0.82187 + 1.42352 i},
 {a → -0.882135 + 1.5279 i, b → -0.82187 - 1.42352 i},
 {a → 0.882135 - 1.5279 i, b → 0.82187 + 1.42352 i},
 {a → 0.882135 + 1.5279 i, b → 0.82187 - 1.42352 i}, {a → 1.76427, b → 1.64374}}
```

- To ensure exponential growth, we need to choose a and b to be positive. Therefore, our initial exponential model becomes:

$$Model(x) = 1.764(1.643)^x$$
$$Domain: \{x \mid x \in \Re\}$$
$$Range: \{Model(x) \mid 0 \le Model(x)\}$$

Figure 1.58 is a plot of the model, to see how well it appears to fit the data.

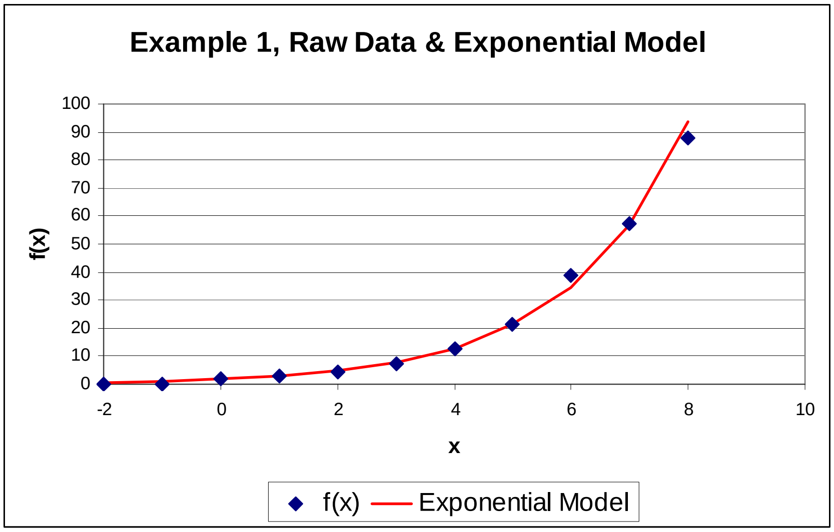


Figure 1.58: Table 1.16 Data Overlaid with Exponential Model

- Power model development.

$$Model2(x) = a\,(x+2)^b + 0.2$$

- Power Model Parameter Estimation: Using modeling skills developed to date, we need to estimate the parameters (a and b) for a power model. To estimate two parameters, we need two data points that represent the data well. Let's try the same points that we used for the exponential model: $(1, 2.9)$ and $(7, 57.2)$.

- Using Mathematica to solve for a and b yields,

```
Solve[{2.9 == a (1 + 2.9)^b + 0.02, 57.2 == a (7 + 2.9)^b + 0.02},
 {a, b}]

Solve::ifun :
  Inverse functions are being used by Solve, so some solutions may not be found;
    use Reduce for complete solution information. >>

Solve::svars : Equations may not give solutions for all "solve" variables. >>

{{a → 57.18 9.9^(-1. b)}}
```

- Note that this solution still shows parameter a in terms of parameter b, meaning we still must make an assumption for the b parameter.

- In this form, it is difficult to intuitively make a good estimation of what value b should take on. However, if we consider the basic power function cases, we **can** make a solid *estimation* for b. We see that the data increases more steeply than we would expect in a quadratic function. The bottom of the "cup" is also more flat than we would expect in a quadratic, both characteristics of a function with a higher power. Let's assume that the parameter $b = 3$. Using the tenth point and $b = 3$ yields:

$$57.2 = a(7+2)^3 + 0.02$$

$$a = \frac{(57.2 - 0.02)}{9^3} \approx .0784$$

- Therefore, our initial exponential model becomes:

$$Model2(x) = .0784(x+2)^3 + 0.02$$
$$Domain: \{x \mid x \in \Re\}$$
$$Range: \{Model(x) \mid 0 \le Model(x)\}$$

Let's plot our models, Figure 1.59, to see how well they appear to fit the data.

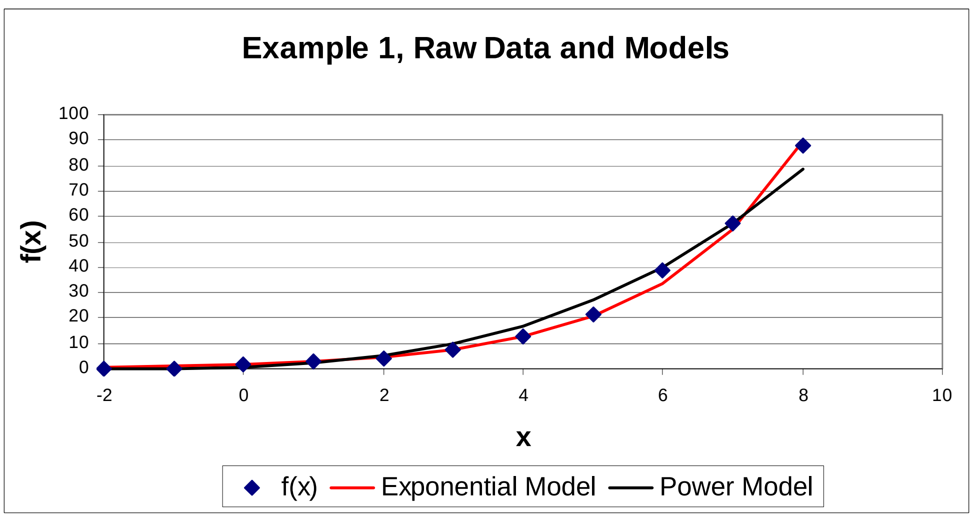


Figure 1.59: Table 1.12 Data Overlaid with Exponential & Power Models

Step 3. *Interpret the results of the solution.* Both models fit the trend of the data. The SSE for the exponential model is approximately 51.6 whereas the SSE for the power model is approximately 153.5. Currently, the exponential model seems to be the better model. Let's refine each model as best we can to provide the client with the best possible model for each type of function.

By manually adjusting the parameters for the exponential and power models, both SSEs can be reduced. With the exponential model, it seems that little can be done to lower SSE. In fact, we can only make minor adjustments to the a and b parameters to make the model less steep. The refined model is:

$$Model(x) = 1.76(1.635)^x .$$

The power model also can be transformed only by a little. If we make the curve steeper, it appears that it will fit the data a little better. To do this, we can change the exponent from a 3 to a 4. After doing that, we must decrease the a parameter because simply changing the b makes the curve too steep. Remember, refining models is an iterative process. The refined power model is:

$$Model2(x) = .00888(x+2)^4 + 0.02 .$$

You can see that, especially with power and exponential models, if good estimations are made at the beginning of the modeling process, then small parameter adjustments may be all that are necessary to improve models.

The refined exponential and power models have SSEs of approximately 50.04 and 23.46, respectively. The plot of both models follows.

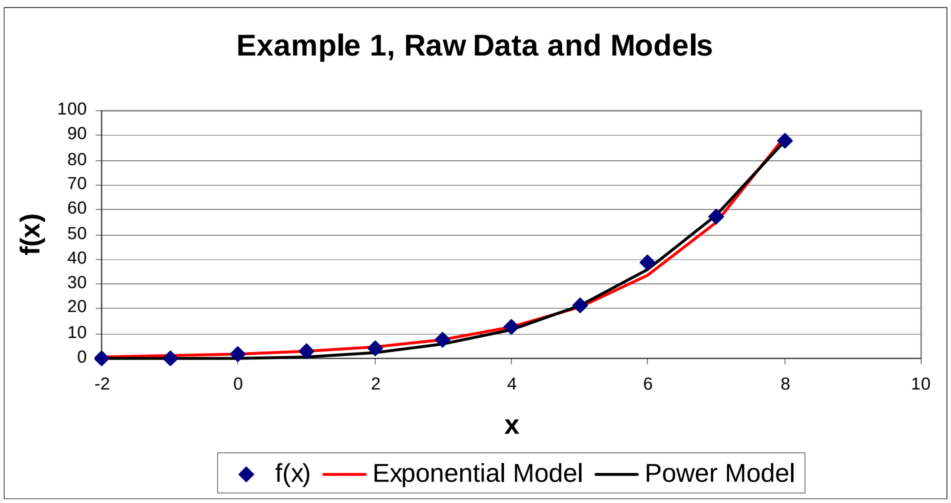


Figure 1.60: Comparison of Power and Exponential Models

We have determined the following: both models fit the trend of the data for the domain within the data set and both models increase at an increasing rate as the domain values increase past the values in the data set. The biggest difference in the models is for domain values that are less than -2. The exponential model will continue to decrease toward 0 as the domain values become more negative whereas the power model will start to “turn up” as the domain values become more negative. For domain values less than -2, the power model will become the reflection of the values greater than -2 about the line $x = -2$. Does this consideration affect your decision on which model to recommend? If you know where the data came from it might; if there were possibly negative values less than two then you would have to consider if the corresponding y values increased or decreased.

The two models that we developed appear to be good models, but are they the best possible models? Do they have the minimum possible SSE? The next subsection will help us to answer that very question!

The data for the following question is also linked on the course website.

Question 1 *Estimate your best possible exponential and power models for the following average salary data for professional basketball players in the years 1984-2004. Which model better represents the nature of the data? Is either model better for extrapolating outside of the domain of the data? Explain your conclusions.*

Year	AVG Salary
1984-85	$330,000
1985-86	$382,000
1986-87	$431,000
1987-88	$502,000
1988-89	$575,000
1989-90	$717,000
1990-91	$927,000
1991-92	$1,100,000
1992-93	$1,300,000
1993-94	$1,500,000
1994-95	$1,800,000
1995-96	$2,000,000
1996-97	$2,300,000
1997-98	$2,600,000
1998-99	$3,000,000
1999-2000	$3,600,000
2000-01	$4,200,000
2001-02	$4,500,000
2002-03	$4,546,000
2003-04	$4,917,000

Table 1.17: NBA Salary Data

1.9.1 Minimizing Sum of Squared Error

We have discovered a measure to *quantify* how far your model deviates from the actual data, the Sum of Squared Error. Since can quantify a model's "goodness," it is natural to want to develop the *best possible* model. Let's use technology to help us find better models, more quickly than we could find through trial and error.

Previously, we were intent on finding a model that predicted the value of the dependent variable of a situation, given the value of the independent variable. The models we have been exploring in this section resulted in the fitting of a data set with the following two models:

$$\begin{aligned} &ExponentialModel: \quad Model(x) = 1.76(1.635)^x \\ &PowerModel: \quad Model2(x) = .00888(x+2)^4 + 0.02 \end{aligned}$$

Using the techniques mentioned earlier in this section, we determined the SSE of the exponential model to be about 50.04 and the power model to be about 23.46 (feel free to calculate this yourself, for additional practice). Since SSE is a relative measure, we also know that the power model we estimated appears to be a better model than the exponential; it certainly fits the data better over the domain of the data set. Are we done? Is this good enough? Can we do better?

You could probably continue to adjust the parameter values parameters to get smaller and smaller SSE values, but this would be tedious and we would not know if we truly had the best combination of parameter values to call it the "BEST" model. As it turns out, we have a powerful ally in our quest to find a low SSE in *Excel*. This ally is a tools add-in called Solver. What does Solver do you ask? Well... it solves! *(If you dig deep into the help files of Excel, you find that it uses something called the Simplex method to solve/optimize linear equations and a reduced gradient algorithm to solve/optimize non-linear equations – but that is a* touch *beyond the scope of this course).*

Before we begin to use Solver, we must first load it onto your laptop. To load Solver, you must first open up *Excel*. In your Tools pull down menu, select Add-Ins. Another screen will appear with several options as to which Add-In you wish to install. You must select the Solver Add-In as shown in Figure 1.61. Ensure you also select the first Add-in, the Analysis ToolPak.

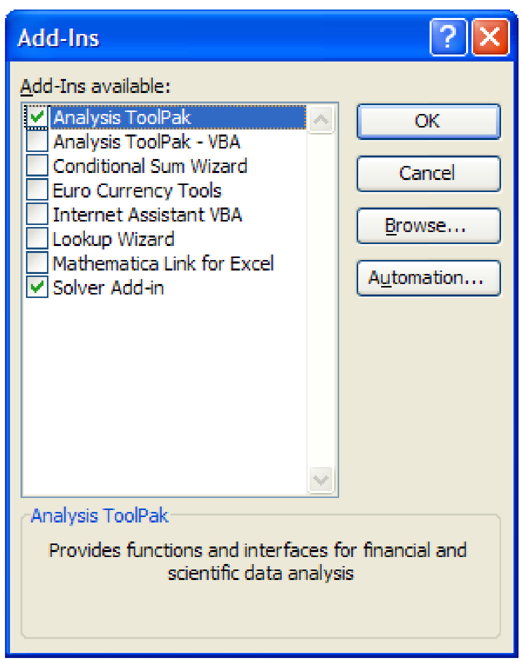


Figure 1.61: Add-In Dialog Box

Once we have Solver loaded, we can work on the problem at hand, that is, finding the "BEST" or minimum SSE.

In the Tools menu, you now have an option to run Solver. When you select this application, you get the Solver dialog box, shown in Figure 1.62.

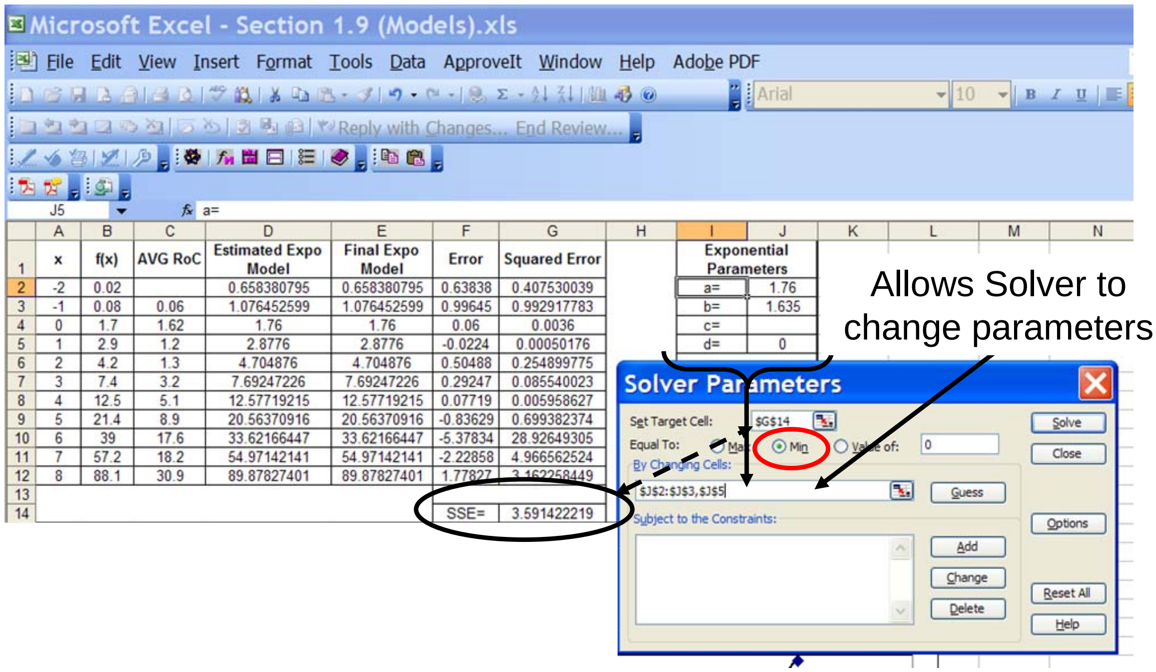


Figure 1.62: Solver Dialog Box

When you were adjusting the parameter values, what was your goal? To minimize the SSE, right? Solver's first input is to set the target cell. The target cell is the cell that you want to minimize, maximize, or go to a specific value. In this case, we wish to minimize the cell that contains the SSE. As a matter of practice, do not have Solver attempt to achieve a value of 0, this is most likely an unachievable goal.

When you were adjusting the parameter values to get a low SSE, you were changing specific cells to help you achieve that goal. In order for Solver to find a "best" solution, we must tell it which cells on the spreadsheet it can change in order to find the minimum SSE. In this case, we want to allow Solver to change the parameter values. Now we just need to tell it to Solve. You may notice that in Figure 1.64, there is a $ before each reference in the Solver parameter box. Take some time to look up **absolute referencing** in the *Excel* help files to Figure out what this means. *Your use of absolute referencing is vitally important to your modeling efforts.*

Solver will usually find its best solution quite quickly. Before Solver writes the solution it obtained onto your spreadsheet, you will be asked if you want to change to the new solution that was just obtained, or restore the original values with which you started. For this example, we will accept the Solver values. See Figure 1.63 for the *better* exponential solution. Our eyeball test shows that that the final model appears to be a closer fit to the data. Note that the already small SSE was cut by more than two-thirds. Both of our model evaluation tests agree: the final model fits the data better than the estimated model.

M	N	O	P	Q	R	S	T	U	V
x	f(x)	Final Exponential Model	Squared Mean Error	Model Error	Squared Model Error				
-2	0.02	-0.69681766	453.6125	0.716818	0.513828				
-1	0.08	-0.05649564	451.0604	0.136496	0.018631		a	2.805963	
0	1.7	0.93219453	384.8731	0.767805	0.589525		b	1.544051	
1	2.9	2.45878307	339.2294	0.441217	0.194672		c	0	
2	4.2	4.81591438	293.0321	-0.61591	0.379351		d	-1.87377	
3	7.4	8.45544652	193.7158	-1.05545	1.113967				
4	12.5	14.0750716	77.76033	-1.57507	2.48085				
5	21.4	22.752062	0.006694	-1.35206	1.828072				
6	39	36.1497821	312.6467	2.850218	8.123742				
7	57.2	56.8365519	1287.505	0.363448	0.132095				
8	88.1	88.7779897	4459.811	-0.67799	0.45967				
	21.31818	SST:	8253.253	SSE:	15.8344				
				r^2:	0.998081				

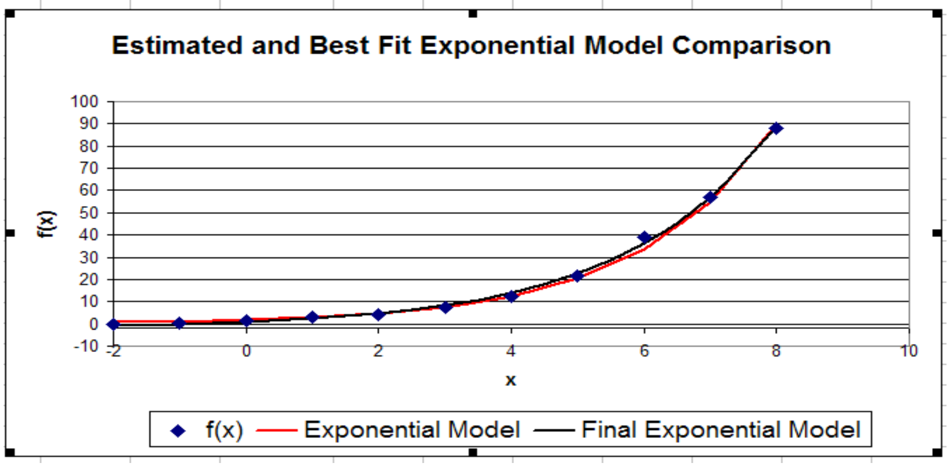


Figure 1.63: Model Post-Solver

After using Solver, the exponential model of *best* fit is:

$$\begin{aligned} &Model(x) = 2.806(1.544)^x - 1.874 \\ &Domain: \quad \{x \mid -2 \le x \le 8\} \\ &Range: \quad \{y \mid 0 \le y \le 88.1\} \end{aligned}$$

Are we done? Can we do better with our model? Run Solver again to see if it results in an even more refined answer. Occasionally, with models that are a little more complex, Solver will not always find an optimal solution. It is best to give Solver a couple of chances to see if it "converges" to an answer within its specified tolerance. Another important point is that often, Solver needs initial parameter values that are reasonable before it can optimize.

Since starting with different parameter values may result in different SSEs for complex models, care should be taken in selecting the initial start or "guess" values. You should select start values that allow a reasonably good fit between your model and the data which is why the parameter estimation techniques you've been taught thus far are of vital importance. In other words, if you put "garbage" in, you get "garbage" out.

The data for all of the following problems are linked on the course website.

Question 2 *Develop a best fit model for the power function modeled in Example 1. **HINT** Do not allow Solver to optimize the b parameter, the exponent. The result will be a decimal (root) function which will provide undefined answers for negative values of (x+c). You may manually change the exponent, say, from a 3 to a 4. Is the power model or exponential better? Why?*

Question 3 *The following table represents the average life expectancy for females in the United States[1] since 1900. Model the data with an exponential and power model. Which model best reflects the trend of the data for the 20th century? Which model is better for extrapolating female life expectancies into the 21st century? Explain.*

Year	1900	1909	1919	1929	1939	1949	1959	1969	1979	1989	2002
Life Expectancy (US Females)	50.7	53.2	57.4	60.9	65.9	71	73.2	74.6	77.6	78.8	79.9

***Question 4** *LoggerPro, a computer-based data collection tool used by the Mathematics Department at USMA, collected the following potential measurements within a capacitor. If you are interested in what a capacitor is and how it works, check out http://electronics.howstuffworks.com/capacitor.htm.*

Model the data with an exponential and power model. Which model best fits the data within the specified domain? Which model do you think would be best to extrapolate data for domain values that are greater than those provided? Explain.

Time (s)	Potential (V)	Time (s)	Potential (V)	Time (s)	Potential (V)
0	1.5995	30	0.2271	60	0.0464
2	1.4286	32	0.2027	62	0.0415
4	1.2527	34	0.1783	64	0.0317
6	1.0867	36	0.1587	66	0.0317
8	0.9499	38	0.1441	68	0.0269
10	0.8327	40	0.1197	70	0.0269
12	0.7350	42	0.1148	72	0.0366
14	0.6374	44	0.1001	74	0.0220
16	0.5592	46	0.0904	76	0.0220
18	0.4957	48	0.0855	78	0.0317
20	0.4322	50	0.0708	80	0.0269
22	0.3785	52	0.0659	82	0.0269
24	0.3346	54	0.0562	84	0.0171
26	0.2906	56	0.0611	86	0.0269
28	0.2662	58	0.0513	88	0.0220

[1] Source: <http://www.ncseonline.org/nle/crsreports/05mar/RL32792.pdf>

THIS PAGE INTENTIONALLY LEFT BLANK

1.10 Trigonometric Functions

1.10.1 Properties of Trigonometric Functions

We have discussed several different types of models so far in this course. We've discussed those that increase or decrease at a constant rate, linear models. We've also discussed generalized exponential and power that may increase at an increasing or decreasing rate, or decrease and an increasing or decreasing rate. However, in addition to phenomena that may be modeled with these types of functions, there are many phenomena that can be considered ***periodic***. In other words, the phenomena that you observe repeat themselves again and again. For example, over the course of a year, you could record the sunrise and sunset times in your hometown to compare them to historical data. Ultimately, you would discover that, for every date, the sunrise and sunset time would be very similar, year after year. (The sunrise time on January 1, 2006 is roughly equal to the sunrise time on January 1, 2007).

To model events in the world that are periodic, we need functions that are different than the others we've used: trigonometric functions. Trigonometric (or, trig) functions are often used in modeling real-world phenomena such as vibrations, waves, elastic motion, and other quantities that vary in a periodic manner. Here is a simple example of trigonometric behavior with which you may be familiar. Since our armed forces have a technological advantage over the majority of our adversaries, it is common for the U.S. Military to conduct its combat operations at night. Despite our advanced war fighting capabilities, even night vision technology relies on some ambient light to allow night operations. Consider the graph in Figure 1.64. This graph records the percent of visible moon (ambient light) between June 1, 2005 and July 11, 2005.[1]

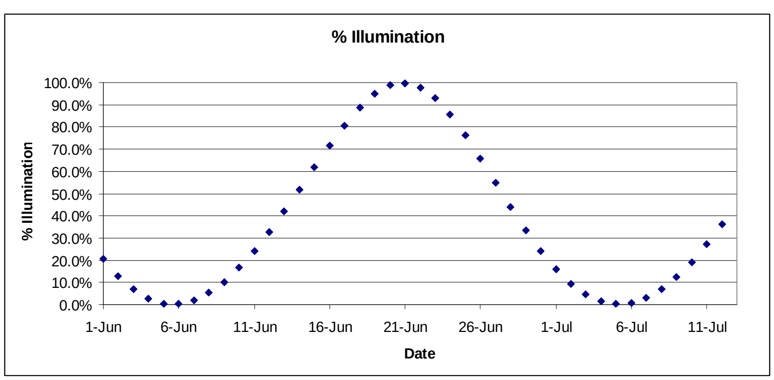


Figure 1.64: Plot of Moon Phase Data from 1 June - 12 July 2005

[1] The Moon's Phase - http://imagiware.com/astro/moon.cgi Downloaded from the Web 2 June 2005.

What do you notice about how much of the moon is visible? When would you recommend conducting night combat operations?

We will limit our discussion of trigonometric functions and their properties to the sine and cosine functions. Let's consider the basic sine function $f(x) = \sin(x)$, which means the sine of the angle whose radian measure is *x*.

Plot this function in *Mathematica* for values of *x* between 0 and 7 as shown in Figure 1.65. *Mathematica* evaluates trigonometric functions in radians.

```
In[1]:= Plot[Sin[x], {x, 0, 7}]
```

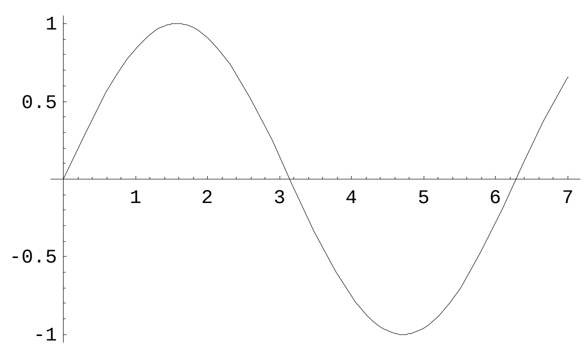


Figure 1.65: *Mathematica* Code to Plot *Sin(x)* Between 0 and 7

Question 1 *How long does it take the sine wave to complete one cycle-time it takes to move from peak to peak or from trough to trough? How tall is the function (peak to valley)? Where does $f(x) = \sin(x)$ cross the y-axis, i.e., what is $f(0)$?*

The general form of the sine function that allows us to adjust the amplitude, *a*; frequency, *b*; horizontal shift, *c*; and the vertical shift, *d*, is:

$$f(x) = a * sin(b(x+c)) + d \quad (1)$$

Question 2 *Plot the following functions by hand: $f(x) = 5\sin(x)$ and $f(x) = \sin(x) + 5$. What is the difference? How do the parameters a and d affect the basic sine function?*

Question 3 *Now that you have discovered the effects of the a and d parameters what affect do the b and c parameters in Equation (1) have on the function? Determine the answer to this question by plotting*

$$f(x) = \sin(2x) \textit{ and } g(x) = \sin(x + \frac{\pi}{2}).$$

After you have answered Question 3, go to the website below and adjust the parameters to test your hypotheses regarding the effects of changing each of the parameters. See if you can find the parameter values that make the function fit the illumination data best.

http://www.dean.usma.edu/departments/math/MRCW/MA103/Trig/live˙graph.html

You should have concluded that the parameters affect the function in four ways, four ways that are similar to generalizations 1 and 2 in section 1.8.

- *a* parameter: controls the extreme range values of the function, the values that the function *oscillates* between (a vertical stretch)
- *b* parameter: affects the distance along the independent variable axis that it takes to complete one cycle(the shape of the graph, a horizontal stretch)
- *c* parameter: horizontal shift
- *d* parameter: vertical shift

We will discuss a mathematical way to calculate each of these so that you can accurately estimate the values from a dataset or a graph before minimizing your model's error using Solver.

Amplitude of a Trigonometric Function: Consider again the graph of the moon phase data. Between what values does the percentage of illumination oscillate? If we define the **amplitude** of a function as *half the distance between the peak and trough of the function*, what is the amplitude of the moon phase?

> The **amplitude** of a function is half the distance between the peak and trough of the function.

We notice that the greatest value (according to the data) is 99.7% and the smallest value is 0.2%. Half the distance between these two values is:

$$Amplitude = \frac{1}{2}\left(99.7\% - 0.2\%\right) = 49.75\%.$$

This makes the amplitude, or the a parameter in our model, 49.75%.

Frequency and Period of a Trigonometric Function: Mathematicians call the parameter b in Equation 1 the **frequency** of the sine function. *Frequency* tells us *the number of complete cycles that occur between* $x = 0$ *and* $x = 2\pi$ (One complete repetition of the pattern in the function is called a cycle). You are probably most familiar with $f(x) = sin(x)$ where the *frequency* or b parameter is one. As shown in Figure 1.66, for the basic sine function there is one cycle between $x = 0$ and $x = 2\pi$ (about 6.28).

> The **frequency** of a function is the number of complete cycles that occur between $x = 0$ and $x = 2\pi$.

Often when attempting to model data with trigonometric functions, it is easier to model the data by considering the period instead of the frequency.

> The **period** of a function is defined as the time it takes to complete one cycle.

The relationship between *period* and *frequency* is as follows:

$$frequency = \frac{2\pi}{period}.$$

Thus, often you may see Equation (1) written as

$$f(x) = a\sin\left(\frac{2\pi}{period}(x+c)\right)+d .$$

Since the *frequency* of the basic sine function is one, this indicates its period is 2π. What is the period of the moon phase?

Date	% Illumination	Date	% Illumination	Date	% Illumination
1-Jun	20.8%	15-Jun	61.8%	29-Jun	33.5%
2-Jun	13.0%	16-Jun	71.5%	30-Jun	24.1%
3-Jun	6.9%	17-Jun	80.6%	1-Jul	16.0%
4-Jun	2.7%	18-Jun	88.6%	2-Jul	9.5%
5-Jun	0.5%	19-Jun	94.8%	3-Jul	4.6%
6-Jun	0.2%	20-Jun	98.7%	4-Jul	1.5%
7-Jun	1.9%	21-Jun	99.7%	5-Jul	0.2%
8-Jun	5.3%	22-Jun	97.8%	6-Jul	0.7%
9-Jun	10.3%	23-Jun	93.0%	7-Jul	3.0%
10-Jun	16.7%	24-Jun	85.6%	8-Jul	6.9%
11-Jun	24.3%	25-Jun	76.3%	9-Jul	12.3%
12-Jun	32.8%	26-Jun	65.8%	10-Jul	19.2%
13-Jun	42.1%	27-Jun	54.8%	11-Jul	27.2%
14-Jun	51.9%	28-Jun	43.8%	12-Jul	36.2%

Table 1.18: Moon phase data from 1 June – 12 July 2005

Notice in Table 1.18 that illumination is at its lowest value on 6 June and again on 5 July. The moon phase completes its cycle in just about one month (30 days to be exact). This makes the period 30 days. Now, we can determine the *frequency* of our data.

$$frequency = \frac{2\pi}{30}$$

or around 0.2094 Cycles per 2π days. Thus the b parameter in our model would be 0.2094.

<u>Horizontal Shift of a Trigonometric Function:</u> How do we model **horizontal shift**? Hopefully you discovered from your experimentation earlier that the c parameter *shifts our graph left and right depending on positive or negative values*. More specifically, if c is negative, the function shifts to the right by c units. Similarly, if c is positive, the function shifts to the left by c units.

> The **horizontal shift** of a function is the distance left or right from the vertical axis. This definition assumes the vertical axis is used as the "start point" for a basic trigonometric function such as $f(x) = \sin(x)$ or $g(h) = \cos(h)$.

Question 4 *Suppose that we define Day 0 as June 1, 2005. Which direction do we need to move our model to match the data? How much do we need to shift this model in that direction?*

Vertical Shift of a Trigonometric Function: Mathematicians, by convention, use the *x*-axis as a baseline to refer to functions that are shifted vertically. We notice the midline of the moon phase data is shifted upwards from the *x*-axis. If we define the **vertical shift** of a function as: *the distance that the midline of the function is shifted upwards or downwards from the x-axis*, what is the vertical shift of the illumination data?

> The **vertical shift** is the distance that the midline of the function is shifted upwards or downwards from the *x*-axis.

We noted before that the amplitude of the function was 49.75%. We use the amplitude to help us figure out the vertical shift of the function. Since the amplitude is half the distance between the peaks and valleys of the data, we could consider this the "midline" of the data. How much is this "midline" shifted upward from the *x*-axis?

$$49.75\% + 0.2\% = 49.95\%$$

This makes the vertical shift, or the *d* parameter, 49.95%. See Figure 1.66 for a graphical interpretation of the vertical shift.

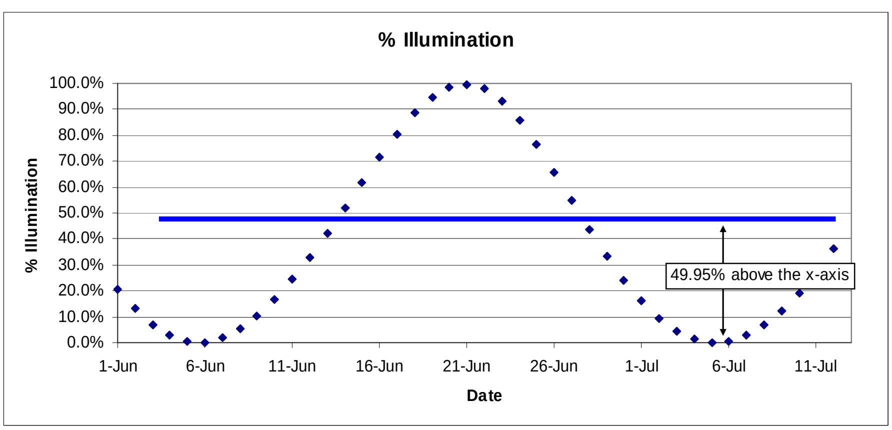


Figure 1.66: Illumination Data Demonstrating Vertical Shift

Question 5 *Recreate Figure 1.64 in Excel. The data for this is shown in Table 1.15 and is posted on the course website.*

Question 6 *Now that you've seen how to estimate the parameters for the moon phase data using a sine function, determine and calculate new parameters for a cosine model of the same data set.*

Question 7 *In a laboratory environment, scientists use an instrument called an oscilloscope to convert physical phenomena such as sound and light waves to electric pulses. The data in Table 1.19 represents extracted readings from the oscilloscope for a particular wave.*

Length (in Å)	Amplitude (in Volts)
0.0	-1.65
0.1	-0.93
0.2	0.00
0.3	0.93
0.4	1.65
0.5	1.99
0.6	1.87
0.7	1.33
0.8	0.48
0.9	-0.48
1.0	-1.33
1.1	-1.87
1.2	-1.99
1.3	-1.65
1.4	-0.93
1.5	0.00
1.6	0.93
1.7	1.65
1.8	1.99
1.9	1.87
2.0	1.33

Table 1.19: Oscilloscope Readings for a Given Wave

a. Estimate the wavelength (or period) of the wave. Given the information provided in Figure 1.67, where on the electromagnetic spectrum does this wave lie?

b. Estimate the amplitude, vertical and horizontal shift of a sinusoidal model that fits the given data. Calculate the frequency of the sinusoidal model using your estimate of wavelength above.

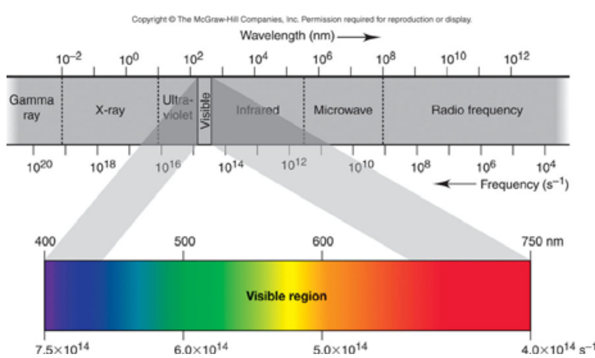


Figure 1.67: Region of the Electromagnetic Spectrum

***Question 8** *Given the data in Table 1.20 on the sunrise in Highland Falls, NY, develop a mathematical model to predict sunrise at any given day during the year (linked on the MA103 website).*

Date	Day of Year	Sunrise Time	Date	Day of Year	Sunrise Time	Date	Day of Year	Sunrise Time
1-Jan	1	722	7-May	127	546	10-Sep	253	631
15-Jan	15	720	21-May	141	532	24-Sep	267	645
29-Jan	29	710	4-Jun	155	524	8-Oct	281	700
12-Feb	43	655	18-Jun	169	522	22-Oct	295	715
26-Feb	57	636	2-Jul	183	527	5-Nov	309	632
12-Mar	71	613	16-Jul	197	536	19-Nov	323	649
26-Mar	85	550	30-Jul	211	549	3-Dec	337	704
9-Apr	99	626	13-Aug	225	603	17-Dec	351	716
23-Apr	113	605	27-Aug	239	617	31-Dec	365	722

Table 1.20: Sunrise data for Highland Falls, New York

a. *Plot the day of year versus sunrise time. How does your data look? How would you expect your data to look? What happens (each year) between 26 March and 9 April? Sometimes data doesn't come to us in an immediately usable format – see if you can "clean-up" the data to make it more usable.*

b. *Estimate the period of the data. Calculate the frequency.*

c. *Estimate the amplitude of the data.*

d. *Estimate the vertical shift of the data.*

e. *Estimate the horizontal shift of the data.*

f. *Put this all together to predict the sunrise time on July 5. For what domain and range is your model valid?*

1.10.2 Modeling with Trigonometric Functions

So far, we have introduced the properties and attributes of trigonometric functions. Hopefully, by now you not only have some ideas about cyclic events, but also have developed the ability to estimate the parameters of trigonometric occurrences. The ability to estimate the parameters of these functions is a key to modeling periodic behavior well.

As we begin our discussion of trigonometric modeling, it is important to note that we will only consider the sine and cosine functions when modeling or data fitting.

Example 1 *Let's use the moon phase data, Table 1.21, to develop a mathematical model that predicts the percent illumination for the day of the month.*

Date	% Illumination	Date	% Illumination	Date	% Illumination
1-Jun	20.8%	15-Jun	61.8%	29-Jun	33.5%
2-Jun	13.0%	16-Jun	71.5%	30-Jun	24.1%
3-Jun	6.9%	17-Jun	80.6%	1-Jul	16.0%
4-Jun	2.7%	18-Jun	88.6%	2-Jul	9.5%
5-Jun	0.5%	19-Jun	94.8%	3-Jul	4.6%
6-Jun	0.2%	20-Jun	98.7%	4-Jul	1.5%
7-Jun	1.9%	21-Jun	99.7%	5-Jul	0.2%
8-Jun	5.3%	22-Jun	97.8%	6-Jul	0.7%
9-Jun	10.3%	23-Jun	93.0%	7-Jul	3.0%
10-Jun	16.7%	24-Jun	85.6%	8-Jul	6.9%
11-Jun	24.3%	25-Jun	76.3%	9-Jul	12.3%
12-Jun	32.8%	26-Jun	65.8%	10-Jul	19.2%
13-Jun	42.1%	27-Jun	54.8%	11-Jul	27.2%
14-Jun	51.9%	28-Jun	43.8%	12-Jul	36.2%

Table 1.21: Moon phase data from 1 June – 12 July 2005

Step 1. Transform the Problem. We are **given** the data in Table 1.21 and must **find** the best model to predict the percent of illumination for any given day in a month. Our **plan** will be to review a plot of the data (Figure 1.65, page 94) estimate the parameters of the model using the procedures developed earlier in the section, then find the best fit model by minimizing the SSE. We can assume that we will use a trigonometric model for two reasons: first, we know that the nature of moon phases is cyclic, requiring a trig model of the form $f(x) = a * sin(b(x+c)) + d$; second, the graph of the data, Figure 1.65, demonstrates a cyclic trend that reinforces what we thought about the nature of the phases of the moon.

Recall that a represents the *amplitude* or *half of the distance between the peak and valley of the function,* b represents the *frequency* or *the number of cycles that occur between x=0 and x=2π,* c represents the *horizontal shift,* and d represents the *vertical shift*.

Step 2. Solve Using Appropriate Techniques

We will begin the solution process by estimating our parameters. To calculate the *amplitude* of the moon phase data, we find the highest and lowest illumination values from Table 1.21.

$$Amplitude = a = \frac{1}{2}\left(99.7\% - 0.2\%\right) = 49.75\% \, .$$

To calculate the *frequency* of the data we see how long the data takes to make 1 cycle; here it takes approximately 30 days. When looking at the table, notice that the illumination on 1 June is 20.8%. After that date, the illumination falls until it reaches a low on the 6^{th} of June and then climbs until it reaches its peak on the 21^{st}. In this case, to estimate the period of the data, we want to see how long it takes the data to drop to its lowest point, rise to its highest point and then fall until it reaches its approximate starting value. This occurs between 30 June and 1 July. Therefore, we say the period of the illumination data is approximately 30 days.

$$Frequency = b = \frac{2\pi}{period} = \frac{2\pi}{30\,days} = 0.2094$$

Determining the *horizontal shift* for the sine model was part of your homework from last lesson. You should have used a plot of the function and seen that we needed to shift the function to the right about 12.5 days to model the scenario using a sine function. Where does the 12.5 days come from?

Remember, $\sin(0) = 0$. This point is halfway between the function's first peak and the preceding trough, both along the *x*-axis and along the *y*-axis. See Figure 1.68. In Figure 1.69 we have shifted the function to the right two units. The center along the *y*-axis between the *first* peak and the preceding trough occurs at $x = 2$, which is how far we have shifted the function to the right.

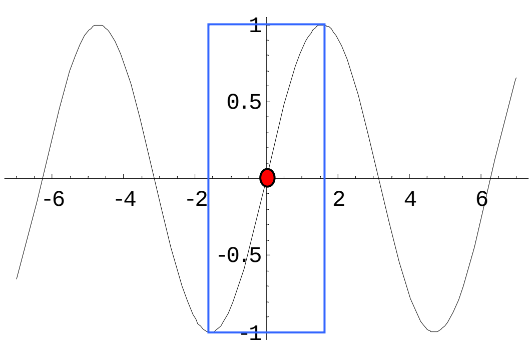


Figure 1.68: Graph of $f(x) = \sin(x)$ from -2π *to* 2π

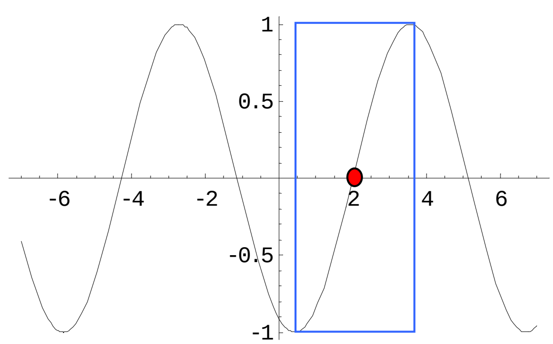


Figure 1.69: Graph of $f(x) = \sin(x - 2)$ from -2π *to* 2π

Seeing the *horizontal shift* quickly becomes more complex when a function exhibits both a vertical as well as a horizontal shift. Figure 1.70 illustrates that the same procedure is used to determine the *horizontal shift* for the case including both shifts. We need to find the midpoint between the *first* peak and the preceding trough to determine how far our function has been shifted. Here, we can see our midpoint is located at $x = 4$, indicating our function is shifted 4 units to the right.

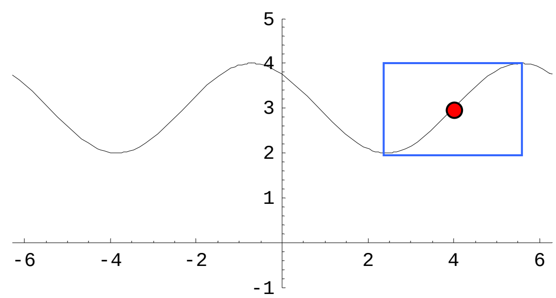


Figure 1.70: Graph of $f(x) = \sin(x-4) + 3$ from -2π *to* 2π

In our moon phase data, the *horizontal shift* is the *number of days our data is shifted to the right of the dependent variable's axis.* We see in Figure 1.71 that the first peak occurs on 21 June and the preceding trough occurs on 6 June yielding a horizontal difference of 15 days. Half of this distance is 7.5 days and occurs between the 13th and 14th of the month. See Figure 1.71 for a graphical representation of the *sine model horizontal shift*. As you can see, the center of the peak and trough is shifted 12.5 units to the right of the vertical axis.

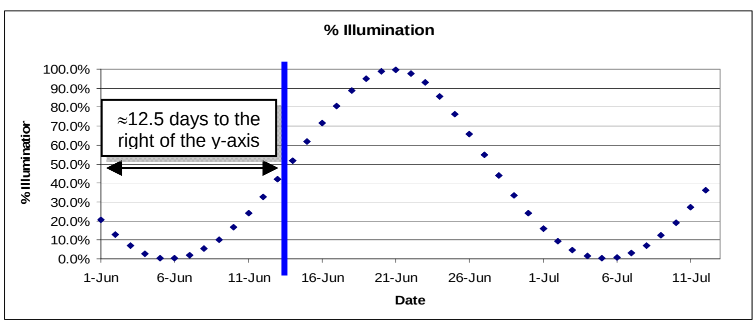


Figure 1.71: Illumination Data Demonstrating Horizontal Shift – Sine Model

$$\textit{Horizontal Shift}_{\text{sine\_model}} = c = 1 - 13.5 = -12.5$$

Finally, we need to determine the *vertical shift* of our data. Recall that the *vertical shift* is *the distance that the midline of the function is shifted upwards or downwards from the x-axis.*

$$\textit{Vertical Shift} = \textit{Amplitude} + \textit{trough} = 49.75\% + 0.2\% = 49.95\%$$

<u>Put it all Together:</u> We would like to use the sine function to model our illumination data. Now that we have determined estimates for each of our parameter values, let's create our mathematical model.

$$illum(day) = 49.75 * \sin(0.2094 * (day - 12.5)) + 49.95 \quad (2)$$

$$\text{Domain}: \{day \mid -\infty < day < \infty\} = (-\infty, \infty)$$
$$\text{Range}: \{illum \mid 0 \le illum \le 100\} = [0, 100]$$

Equation 2 is our mathematical model for predicting the percent of illumination based upon the day of the month.

We now have an estimation for our model. Figure 1.72 shows how it compares to a graph of our data points.

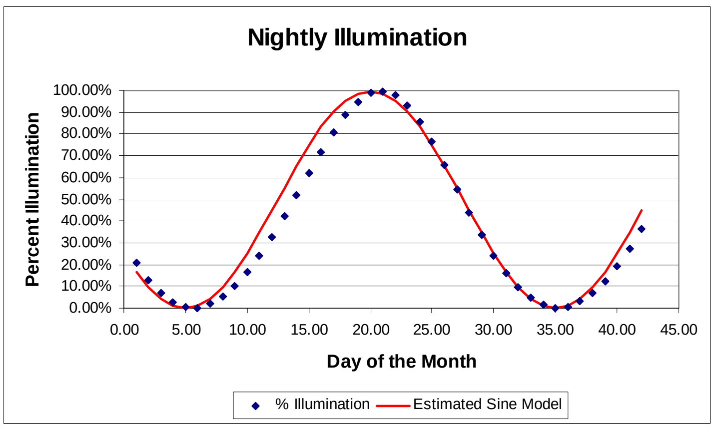


Figure 1.72: Graph of Moon Phase Data Overlaid With Sine Model

The model looks good; it passes the eyeball test, so we seem to have a good model. Figure 1.73 shows the results after using Solver to minimize the SSE, and develop the "best fit" model. This yields the model:

$$illum(day) = 48.91 * \sin(0.2132 * (day - 13.26)) + 47.06$$

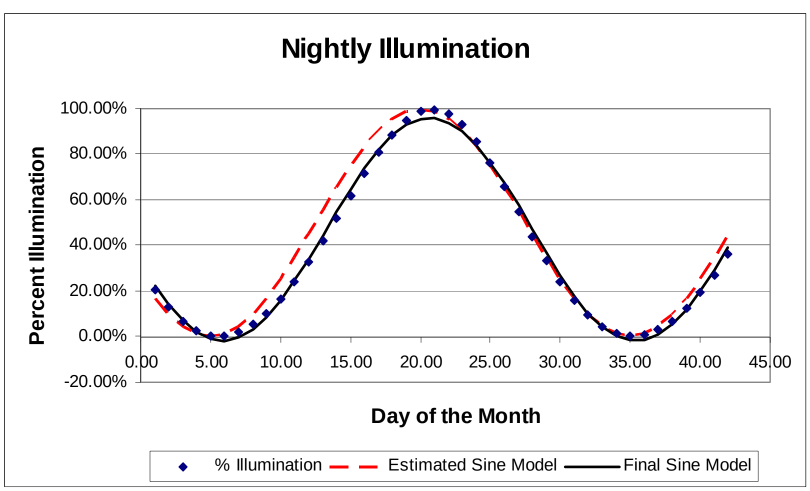


Figure 1.73: Illustration of the Best Fit Sine Model

Step 3. Interpret the Result: Reflect and Communicate.

In reflecting about the choices we have made for our model, we must challenge our assumptions to ensure they are the best for the situation. We can do this by conducting a sensitivity analysis. One major assumption we used in the development of our model is that a sine function would most accurately represent the data. What if we chose cosine?

What is the difference between sine and cosine? Let's consider the most basic sine and cosine functions, $f(x) = \sin(x)$ and $g(x) = \cos(x)$. Both functions share the same generalized form,

$$f(x) = a * trig\_function\ (b*(x+c)) + d\ .$$

We see that in the most basic cases, $a = 1$, $b = 1$, $c = 0$, $d = 0$. What's the difference between the functions? Consider where the functions cross the y-axis. The sine function crosses at $y = 0$ (the midpoint of the function) and the cosine at $y = 1$ (the peak of the function). It appears that if we shift the sine graph left $\frac{\pi}{2}$ we would have similar graph to the cosine graph. So, the difference in modeling the data with the two functions must be in the horizontal shift.

Note that the cosine function's first peak from the y-axis is at $f(0)$. Looking at the moon phase data, let's estimate how much that first peak of our data is shifted to the right of the dependent axis using Figure 1.74.

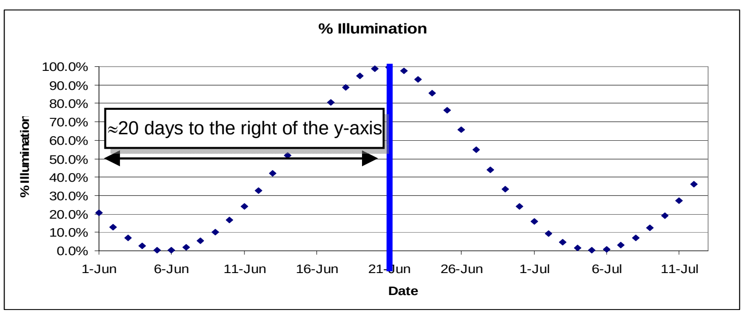


Figure 1.74: Illumination Data Demonstrating Horizontal Shift – Cosine Model

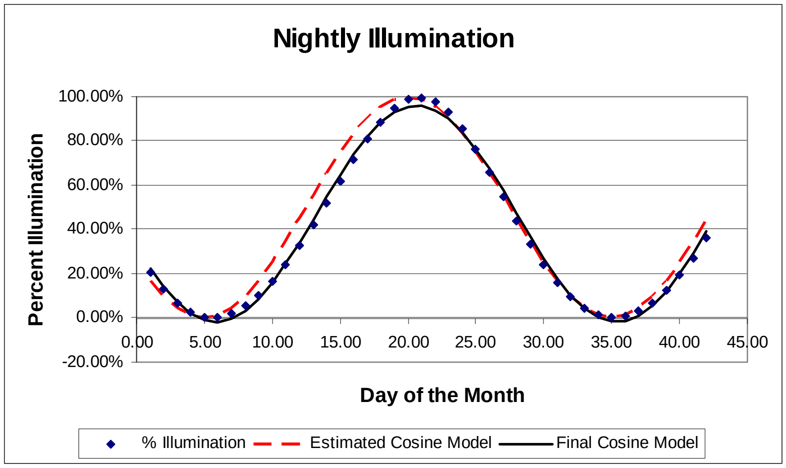
$$Horizontal\ Shift_{\text{cosine\_model}} = c = 1 - 21 = -20$$

We conclude that for our cosine model, we should set our *horizontal shift* parameter, *c,* equal to -20 days. Our estimated cosine model is:

$$illum(day) = 49.75 * \cos(0.2094 * (day - 20)) + 49.95 \qquad (3)$$

$$Domain: \qquad \{day \,| \quad \begin{array}{l} \text{Domain}: \{day \,|\, -\infty < day < \infty\} = (-\infty, \infty) \\ \text{Range}: \{illum \,|\, 0 \le illum \le 100\} = [0, 100] \end{array}$$

After using Solver to minimize the sum of squared error and plotting the model in Figure 1.75, we see that the final model appears to fit the data just as well as the sine model.



1.75: Illustration of the Best Fit Cosine Model

In fact, the sum of squared errors for each model is identical. What we have just seen is that the sine and cosine models are simply shift versions of each other. If you data "begins" at the function's midline, it may be easier to use the sine function. If the data "begins" at the functions peak or trough, it may be easier to use a cosine function.

Question 9 *What does the model in Example 1 predict for the percent illumination on day 20 (June 20th)? Compare this prediction to the actual value found in Table 1.15. How well do you feel our model fits the data? What improvements would you suggest?*

Question 10 *The thermostat at Sara's home in Washington D.C. is set at 70°F. Whenever the temperature drops to 68°, roughly every 30 minutes, the furnace comes on and stays on until the temperature reaches 72°.*[2]

a. *Use your knowledge of trigonometric functions to estimate parameters and create a model to predict the temperature of Sara's home as a function of time.*

b. *Assume your barracks thermostat is set the same way. What modifications would you make to your model to accurately reflect the temperature of your barracks?*

c. *Now consider the modifications you would make to your model to accurately predict the temperature in a home located in southern Florida. What would they be?*

d. *Think about the nature of the data. Is using a trigonometric function to model this situation appropriate? Explain. (Hint: Think about the rates at which the temperature increases and decreases.)*

[2] This problem is adapted from Functioning in the Real World, A Precalculus Experience by Gordon, Gordon, Tucker, and Siegel, p. 510.

***Question 11** *Ever since your summer vacation of salmon fishing in Alaska you have wanted to go back and get the "one that got away." You are planning a fishing trip for 8 June – 12 June 2008. You know that the best fishing occurs at high tide, but the problem is that you only have tidal tables that run through 31 May 2006. What will be the best time(s) to fish on 9 June? An extract of the tide table is shown in Table 1.16. What assumptions did you make in arriving at your conclusion? (Data for Question 3 located on course website)*

Time of Day	Tide Height Observed	Time of Day	Tide Height Observed
0:00	6.94	13:00	4.23
1:00	5.06	14:00	2.64
2:00	3.36	15:00	1.57
3:00	2.25	16:00	1.32
4:00	1.93	17:00	1.96
5:00	2.41	18:00	3.4
6:00	3.55	19:00	5.38
7:00	5.05	20:00	7.44
8:00	6.46	21:00	9.08
9:00	7.4	22:00	9.93
10:00	7.62	23:00	9.81
11:00	7.08	0:00	8.72
12:00	5.86		

Table 1.16: Tidal data for Seward, Alaska on May 31, 2006

Chapter 2

Discrete Dynamical Systems

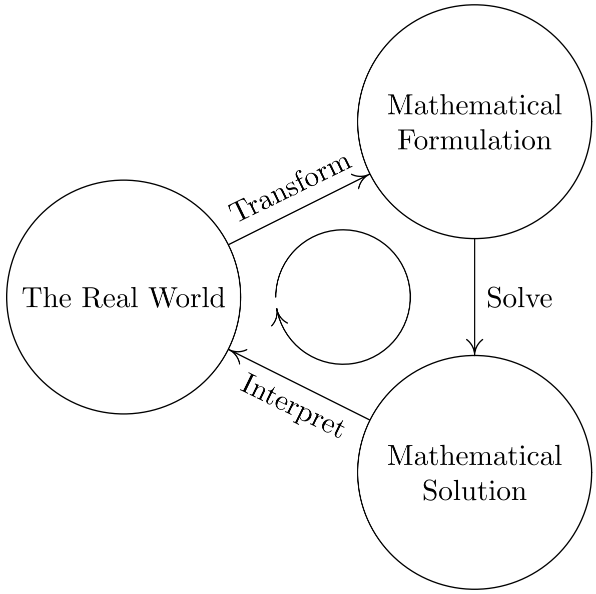


Figure 2.1: The modeling triangle

It is worthwhile to look at the "modeling triangle" shown in Figure 2.1. The heart of this book is using mathematics as a language. We transform problems from the real world into mathematical problems, solve those problems mathematically, and then interpret the mathematical results back in the real world. We start in the circle at the left, in the real world, and then we follow the arrow up and to the right transforming, or expressing, what we see in the real world into the language of mathematics. Then we follow the arrow downward using all the tools of mathematics – including numerical and algebraic calculations and graphics – to obtain mathematical results. Next we follow the arrow back to the real world interpreting our mathematical results back in the real world. Because we are interested in using mathematics to solve real and important problems, the arrows from and to the real world are as important as the mathematical manipulations represented by the arrow on the right from the mathematical formulation to the mathematical solution. In fact, these parts, transforming real problems into mathematical expressions and interpreting mathematical results back in the real world, of modeling and problem-solving often require a better understanding of mathematics than the purely mathematical solutions.

Building models is an iterative process. We begin with simple models to get some traction on a problem and then, as we compare our models with the real world, we cycle several times through the modeling triangle, progressively building better and better models.

In Chapter 1 we worked with **descriptive models** – models that use functions to describe a phenomenon. Such models are valuable because they help us organize information and in many cases even make predictions. We are, however, even more interested in understanding and explaining why things happen. In this chapter we begin to look at **explanatory models** – models that tell us something about the underlying mechanisms involved in real world phenomenon.

Successful modeling often involves both descriptive and explanatory modeling. One of the great achievements of humankind is understanding our solar system. We are all familiar, *now*, with diagrams showing our solar system with the planets revolving about the sun and the moons revolving about the planets. But, these diagrams are based on our models of the solar system. Before the advent of space travel, there were no photographs or movies giving us this bird's eye view of the solar system. Names like Ptolemy, Copernicus, Galileo, Brahe, and Kepler are associated with the first huge leap in understanding – building a succession of descriptive models that organized this data into a model with the sun at the center of the solar system. Names like Newton and Einstein are associated with more explanatory models – the laws of gravity, the relationship between force and acceleration, and the theory of relativity. In this chapter we look at some explanatory models.

2.1 Introduction to Discrete Dynamical Systems

This chapter is about explanatory modeling with "discrete dynamical systems." The word "dynamic" refers to situations where the quantities of interest are changing over time. In this subsection we contrast "discrete" dynamical systems and "continuous" dynamical systems. The first two questions you might ask are:

- What are "discrete" dynamical systems and how is this kind of modeling different from other kinds of modeling?
- When should I use "discrete" dynamical systems and when should I use other kinds of modeling?

We begin this discussion with two examples. The first example is best modeled using "discrete" dynamical systems and the second is best modeled using "continuous" dynamical systems. Looking at these two models side-by-side helps answer the two questions above.

Example 1 *A country, named Tankerland, is building up supplies of oil for a possible military engagement. Every month a tanker arrives with 500,000 tonnes*[1] *of oil that are*

[1] Oil is often measured in metric tons, or tonnes – 1000 kilograms.

added to the reserves. The current reserve has 1,500,000 tonnes and the first tanker is scheduled to arrive at the end of the first month. The graph on the left side of Figure 2.2 shows the oil reserves in Tankerland over the next year.

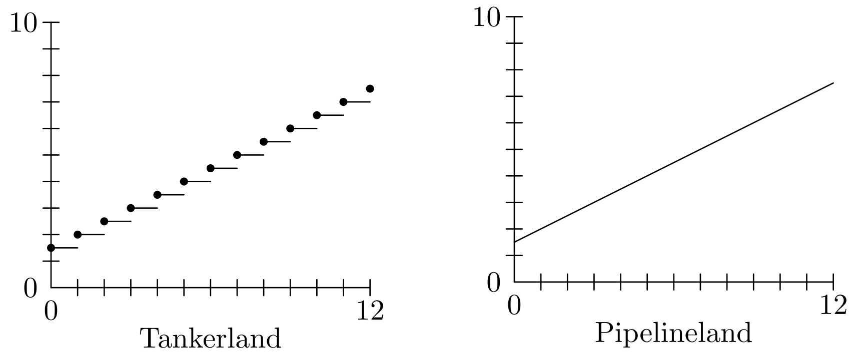


Figure 2.2: Tankers and a pipeline (reserves in millons of tonnes)

Example 2 *A neighboring country, named Pipelineland, is building up supplies of oil for a possible military engagement. This country has its own oil wells and is able to deliver oil via a pipeline directly to its reserves at the rate of 500,000 tonnes per month. The current reserve has 1,500,000 tonnes. The graph on the right side of Figure 2.2 shows the size of the reserves in Pipelineland for the next year.*

Even though 500,000 tonnes of oil are being added to the reserves each month in both Tankerland and Pipelineland, the two graphs look quite different. The tankers deliver oil at discrete times in big "chunks" and the pipeline delivers oil in a continuous stream. As a result, the oil in the reserves jumps at the end of each month when oil is delivered by tankers and it rises continuously when oil is delivered by pipeline.

For the first example we use a **discrete** dynamical system. We will use the letter n to denote time. The values of n will be $0, 1, 2, \ldots 12$. We use the notation p_n to denote the amount of oil in the reserves during month n up to but not including the very last day. Because oil is delivered on the very last day of each month, the amount of oil in the reserves jumps on the very last day of each month. This gives us the model

$$\begin{aligned} p_0 &= 1.5 \text{ million tonnes} \\ p_1 &= p_0 + 0.5 = 2.0 \text{ million tonnes} \end{aligned}$$

$$\begin{aligned} p_2 &= p_1 + 0.5 = 2.5 \text{ million tonnes} \\ &\vdots \\ p_{12} &= p_{11} + 0.5 \text{ million tonnes} \end{aligned}$$

The ordered collection of numbers, $p_0, p_1, \ldots p_{12}$ is called a **sequence**. The individual numbers that make up this sequence are called the **terms** of this sequence. Notice there is a pattern in the calculations above. Starting with the first term (or **initial value**), $p_0 = 1.5$, we compute each subsequent term by

$$p_n = p_{n-1} + 0.5$$

This equation, or pattern, is called a **recursion equation**. One way to read this particular recursion equation is "each month we add 0.5 million tonnes of oil to the previous amount in the reserves." Notice the parallel between the mathematical statement "$p_n = p_{n-1} + 0.5$" and the English statement "each month a tanker delivers 0.5 metric tonnes of oil." This parallel is the reason why the adjective "explanatory" is applied to this model.

For the second example we use the **continuous** model

$$p(t) = 1.5 + 0.5t \text{ million tonnes}$$

The function $p(t)$ represents the amount of oil in the reserves t months from now.

We use the words **discrete** and **continuous** to distinguish between models in which the quantity of interest changes at discrete times and models in which it changes continuously. Because the idea of change is so important, we often write a model like the one for Tankerland in the form of a **difference equation**

$$p_n - p_{n-1} = 0.5 \text{ million tonnes}$$

to emphasize the change from one month to the next. For a continuous model the same idea is written as a **differential equation**

$$\frac{dp}{dt} = 0.5 \text{ million tonnes.}$$

The symbol $\frac{dp}{dt}$ on the left side of this equation denotes the **derivative**. The derivative is the instantaneous rate of change of a continuously changing quantity. For example, if $p(t)$ is the location of a car at time t, then its derivative $\frac{dp}{dt}$, represents the car's velocity. We study derivatives and continuous dynamical systems later in this book.

Notice the parallel between the mathematical statement "$\frac{dp}{dt} = 0.5$" and the English statement "oil is flowing into the reserves at the rate of 0.5 metric tonnes per month." This parallel is the reason why the adjective "explanatory" is applied to this model.

In summary, we will eventually use two different kinds of dynamical systems:

- **Discrete dynamical systems**
  - The quantity of interest changes at distinct times and is described by a sequence – $p_0, p_1, \ldots$.
  - A graph of the quantity of interest looks like the graph for Tankerland in Figure 2.2 on page 115 and jumps at distinct times.
  - The way in which the quantity of interest *changes* is described by a **difference equation** – for example,

    $$p_n - p_{n-1} = 0.5$$

    or by a **recursion equation** – for example,

    $$p_n = p_{n-1} + 0.5.$$

    Notice that the recursion equation and the difference equation express exactly the same information. The difference equation is often read "present - past = change" and focuses our attention on the difference between the present and the past. This particular recursion equation might be read "present = past + change" and focuses our attention on how we determine the present based on the past.
  - An **initial value** that tells us the starting value of the quantity of interest.
- **Continuous dynamical systems**
  - The quantity of interest changes continuously and is described by a function – $p(t)$.
  - A graph of the quantity of interest looks like the graph for Pipelineland in Figure 2.2 on page 115 and shows continuous change with no jumps.

- ◦ The way in which the quantity of interest changes is described by a **differential equation** – for example,

$$\frac{dp}{dt} = 0.5.$$

- ◦ An **initial value** that tells us the starting value of the quantity of interest.

The Tankerland example, Example 1 on page 115, is an example of the simplest kind of discrete dynamical system or the simplest kind of sequence – an **arithmetic sequence**. An arithmetic sequence can be described in many different, equivalent ways –

- An initial value and a recursion equation

$$p_0 = a, \qquad p_n = p_{n-1} + d$$

- An initial value and a difference equation

$$p_0 = a, \qquad p_n - p_{n-1} = d$$

- A formula

$$p_n = a + nd.$$

A formula like this last formula that enables us to compute each term p_n directly is often called a **closed form**, or **analytic**, solution.

We often use a formula to describe a model – for example, in Tankerland the amount of oil in the reserves after the n-th delivery is given by the formula $p_n = 1.5 + 0.5n$ and in Pipelineland the amount of oil in the reserves is given by the formula $p(t) = 1.5 + 0.5t$. Except for using the notation n in one formula and the notation t in the other, these two formulas look identical. But, we must be careful and remember the underlying models. For example, if war broke out between these two countries each country might make disrupting the other country's oil supply lines an immediate priority. Tankerland might bomb Pipelineland's pipelines and Pipelineland might destroy the tankers bringing oil to Tankerland. If the war began on the next to last day of any month, Tankerland would be at a decided disadvantage with almost 0.5 million fewer tonnes of oil in reserve than Pipelineland.

Figure 2.3: M1091 Fuel/Water Tanker

Question 1 *The M1091 fuel/water tanker (Figure 2.3.[2]) is designed to support soldiers on the battlefield. It can transport 1500 gallons of either water or fuel. You are in charge of a supply station that has been established in a secure area to supply water to troops on the front lines. You have 95,000 gallons of water in a tank and every day a convoy of ten M1091 tankers leaves your station with a full load of water. Develop a model for the water in your tank. Draw a graph for this model. When will you need to be resupplied?*

Question 2 *Water can be pumped from the M1091 at a rate of 100 gallons per minute. Develop a model showing how much water is in an M1091 starting at the moment it starts pumping water out. Draw a graph for this model. How long will it take to completely empty each tanker?*

Question 3 *You are posted at a mobile command center. Every morning you receive an update from a unit in the field reporting how much fuel they have. This morning they reported that they have 200,000 gallons of fuel. You know that every day they have reported having 15,000 gallons less fuel than the day before. You post the most recent report every morning immediately after it comes in at 0700. Develop a model for the number posted. Draw a graph for this model.*

Draw a graph for a continuous model that you think is plausible for the actual amount of fuel the unit has, taking into account the fact that they are using fuel throughout the day. Your model should take into account the idea that the rate at which the unit is using fuel varies over the course of each day. For example, fuel might be used at a higher rate when many personnel are out on patrol.

[2]Source: www.globalsecurity.org.

2.2 Discrete Dynamical Systems, Gasoline, and Money

2.2.1 The Price of Gasoline

Figure 2.4: Buying gasoline, January and May 2008

These words are being written just one hour after the author paid over \$60.00 to fill up his car's gasoline tank. In July 2002 the average price of regular gasoline was \$1.36 per gallon. In July 2007 the average of regular gasoline was \$2.93. The left side of Figure 2.4 shows a photograph made in January 2008 and the right side shows a photograph made in May 2008. By the time you read these words the prices in that photograph may seem like the "good old days." Table[3] 2.1 provides some historical data on the price of regular gasoline.

Date	Price	Date	Price	Date	Price
Jul 01, 1991	\$1.104	Jul 07, 1997	\$1.169	Jul 07, 2003	\$1.448
Jul 06, 1992	\$1.147	Jul 06, 1998	\$1.041	Jul 05, 2004	\$1.835
Jul 05, 1993	\$1.086	Jul 05, 1999	\$1.110	Jul 04, 2005	\$2.189
Jul 04, 1994	\$1.097	Jul 03, 2000	\$1.606	Jul 03, 2006	\$2.873
Jul 03, 1995	\$1.169	Jul 02, 2001	\$1.384	Jul 02, 2007	\$2.933
Jul 01, 1996	\$1.219	Jul 01, 2002	\$1.357	Jun 16, 2008	\$4.082

Table 2.1: Midyear retail price of regular gasoline per gallon

[3]Source: United States Energy Information Administration, Department of Energy.

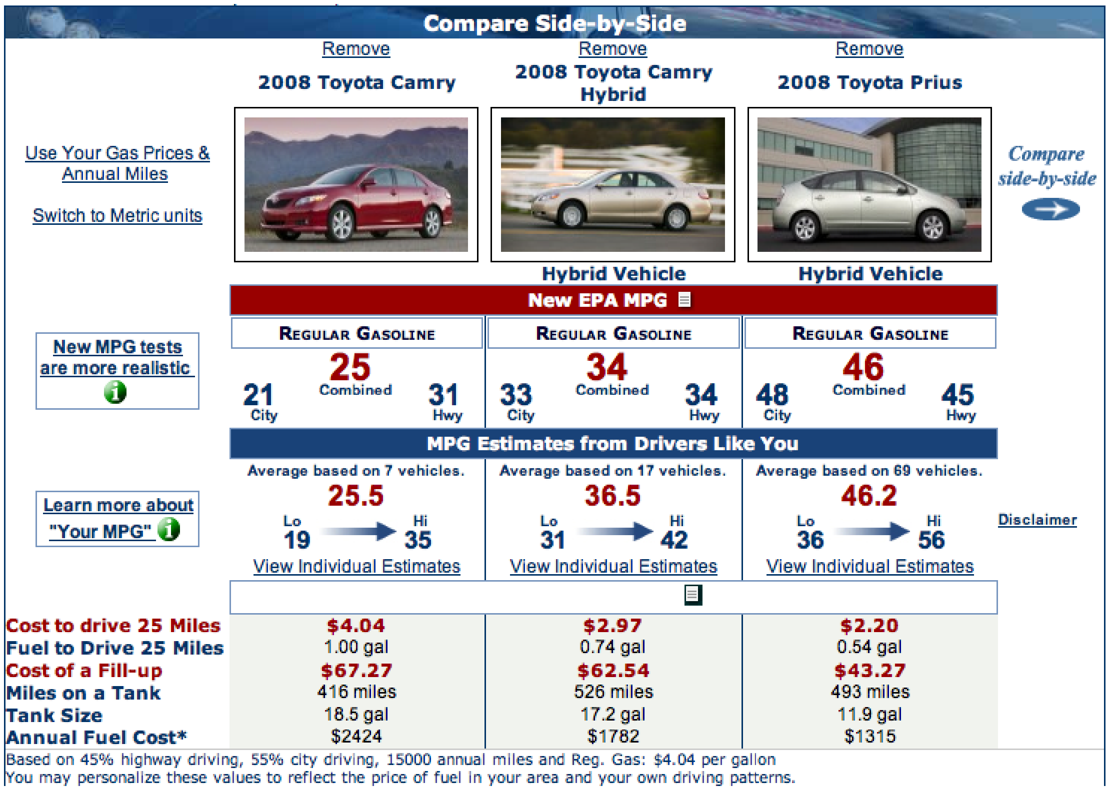


Figure 2.5: Comparing the conventional and hybrid Toyota Camries and the Toyota Prius

So you want to buy a car and you'd probably like to drive it too. You're going to have many choices: Will you buy a car that is really fuel efficient, like the Toyota Prius; one that is less fuel efficient, like the Toyota Camry; a compromise like the hybrid version of Toyota Camry; or a gas-guzzling SUV? Will you wait a few years and save money so that you can pay cash or will you buy a car now and take out a loan? Will you finance the car through the dealer? How about buying a very inexpensive used car now and waiting a few years for even more fuel efficient cars before buying a new car?

The manufacturer's suggested retail price (MSRP) for the Toyota hybrid Camry is \$25,200 and for a conventional Camry the MSRP ranges from \$19,620 - \$28,120. There are two versions of the Toyota Prius. One has an MSRP of \$21,500 and the other has an MSRP of \$23,770. So, the purchase prices of the three cars are all in the same range. Figure 2.5[4] compares the fuel economy of these three vehicles side-by-side. Fuel cost estimates are

[4] United States Department of Energy and Environmental Protection Agency – http://www.fueleconomy.gov/feg/gasprices/index.shtml Downloaded from the Web June 2008

based on regular gasoline at \$4.04 per gallon. You've always wanted a Camry but every time you pass a gas station, that 46 miles per gallon for combined city and highway driving for the Prius sounds better and better. It is 84% better than the 25 miles per gallon for combined city and highway driving for the conventional Camry.

The purpose of this chapter is to build tools that can help you make better choices. We begin with an example that poses a specific question.

Example 1 *Estimate the amount you would spend for gasoline over the next five years if you buy the conventional Toyota Camry. Note that Table*[5] *2.1 on page 120 provides some historical data on the price of regular gasoline.*

Like most real world problems, this problem requires some information that is not readily available. For example, to solve this problem you need to know how much you will drive for each of the next five years and the price of gasoline for each of the next five years. As you work on real problems like this you will often need to make some plausible assumptions about information that is not readily available. This author drives about 18,000 miles per year and we will use that figure in this example. Whenever you make assumptions like this, you must also examine how your results are affected by your assumptions. This is often called **sensitivity analysis**. For example, if you drive much less than this author does then you would spend much less for gasoline then he would. More precisely, if you drove 25% fewer miles per year then you would spend 25% less for gasoline.

Predicting the price of gasoline over the next five years is more difficult. From Table 2.1 on page 120 we see that from July 1997 to July 2007 the price of gasoline rose from \$1.169 per gallon to \$2.933 per gallon, or \$1.764. This rise occurred over a period of ten years, so the average price increase per year is roughly \$0.18. We will assume that this rate of increase will continue. Once again note that if our assumption underestimates the rate at which the price of gasoline rises then our estimates for the amount you will spend on gasoline will be too low and if our assumption overestimates the rate at which the price of gasoline rises then our estimates for the amount you will spend on gasoline will be too high. The price of gasoline rose much more dramatically between July 2, 2007 and June 16, 2008. In fact, it rose more than one dollar in slightly less than one year. Thus, our assumption probably underestimates the rate at which the price is rising.

With these assumptions we are ready to attack this problem. The first step is to estimate the amount of gasoline you expect to buy each year over the next five years. Since the conventional Toyota Camry averages 25 miles per gallon (Figure 2.5 on page

[5] Source: United States Energy Information Administration, Department of Energy.

121) and we estimate you will drive 18,000 miles per year, our estimate for the number of gallons required each year for this vehicle is

$$\frac{18,000}{25} = 720 \text{ gallons.}$$

Now we must estimate the price of gasoline over the next five years. Our estimate will be based on a sequence $p_{2007}, p_{2008}, p_{2009}, p_{2010}, p_{2011}, p_{2012}$ that gives the price each year starting in the year 2007 and continuing for the next five years. This will give us a rough estimate because the price of gasoline varies considerably over the course of each year. It tends to be higher in weeks when people are doing lots of driving and lower at other times. In midyear 2007 the price of regular gasoline was \$2.93 per gallon. Thus, $p_{2007} = \$2.93$. All three of the vehicles we are considering use regular gasoline. We estimated above that the price of regular gasoline will rise by \$0.18 each year. Thus, we estimate

$$\begin{aligned}
p_{2007} &= \$2.93 \\
p_{2008} &= \$2.93 + \$0.18 = \$3.11 \\
p_{2009} &= \$3.11 + \$0.18 = \$3.29 \\
p_{2010} &= \$3.29 + \$0.18 = \$3.47 \\
p_{2011} &= \$3.47 + \$0.18 = \$3.65 \\
p_{2012} &= \$3.65 + \$0.18 = \$3.83
\end{aligned}$$

Notice that this is a discrete dynamical system with the initial value $p_{2007} = \$2.93$ and the recursion equation

$$p_n = p_{n-1} + \$0.18.$$

Now we can estimate the amount you will spend on gasoline over each of the next five years by multiplying the price per gallon for each year by the number of gallons you expect to buy each year. Then we total the figures for the five years to get the total estimated cost of gasoline, \$12,362.40. See Table 2.2 on page 124.

It is worthwhile to compare this figure, \$12,362.40, with the purchase prices of the cars under consideration to get some idea of its impact on the total cost of owning and operating a car. These cars all cost in the range \$19,000 - \$26,000, so the total cost of gasoline for the conventional Toyota Camry for five years is roughly half the purchase

Year	Gallons	Price per Gallon	Total
2008	720	\$3.11	\$2,109.60
2009	720	\$3.29	\$2,368.80
2010	720	\$3.47	\$2,498.40
2011	720	\$3.65	\$2,628.00
2012	720	\$3.83	\$2,757.60
Total			\$12,362.40

Table 2.2: Gasoline expenditures

price of the vehicles under consideration. Comparisons like this and even more careful comparisons are important. For example, if a hybrid costs more than a comparable conventional vehicle then it is important to know whether the savings in the cost of gasoline for a hybrid will offset its higher cost. Because of the importance of this kind of comparison, Consumer Reports ran a story including not only the purchase price and cost of gasoline but also the resale value and other factors for each vehicle in its April 2006 issue. Cadets and faculty in MA 103 found an error in Consumer Reports' original analysis. See http://www.msnbc.msn.com/id/11637968/ and http://www.consumerreports.org/cro/cars/new-cars/resource-center/fuel-economy/high-cost-of-hybrid-vehicles-406/overview/index.htm. The online version is the corrected version but the print version contains the original error.

Question 1 *Using the same assumptions we made above, estimate the total cost of gasoline over the next five years for a Toyota Prius.*

Question 2 *Using the same assumptions we made above, estimate the total cost of gasoline over the next five years for a hybrid Toyota Camry.*

Question 3 *One of the reasons that the price of gasoline is rising is that worldwide consumption of gasoline is rising. You can personally decrease your own gasoline consumption by 45% if you buy a Toyota Prius rather than a conventional Toyota Camry. Suppose that everyone in the United States made the decision to lower their consumption of gasoline and other petroleum products. How would this affect the model above and our conclusions?*

Question 4 *Choose three vehicles that you might consider buying. Choose one vehicle that gets relatively poor gas mileage, one vehicle that gets very good gas mileage, and one vehicle whose gas mileage is between the other two. You can find mileage estimates from http:fueleconomy.gov. Compare the cost of gasoline for these three vehicles over the next five years using your own estimates for how many miles you expect to drive.*

Question 5 *Between 1997 and 2007 the price of regular gasoline rose by an average of 9.6% each year. Compare the total cost for gasoline over the next five years for the conventional Toyota Camry, the hybrid Toyota Camry, and the Toyota Prius assuming that you drive 18,000 miles per year and that the price of gasoline continues to rise by 9.6% each year.*

Question 6 *The preceding questions were all written before the dramatic gasoline price increases in the first half of 2008. Look back at your answers and conclusions. How would they be changed in light of more recent data?*

2.2.2 Borrowing Money

Example 2 *Suppose that you have decided on a car that costs \$23,000. The car dealer offers you a special financing package. She will give you a \$1,000 discount on the price of the car if you finance it through the dealership and borrow the full purchase price (\$22,000 after the discount). You must pay \$559.99 per month for 48 months. Your credit union is willing to loan you \$23,000 at an annual percentage rate of 5% (or 0.416667% per month) if you make monthly payments of at least \$400.00. If you finance through the credit union, you will not receive the \$1,000 discount that the dealer offered you.*

Which way should you finance your purchase?

Discrete dynamical systems are ideal for studying this kind of problem. We will use a sequence $p_0, p_1, p_2, \ldots p_{48}$ to keep track of what would happen if we financed the car through the credit union. The first term, p_0, of the sequence is the amount of the loan.

$$p_0 = \$23,000$$

As usual the value, \$23,000, of the first term, p_0, is called the initial value. The remaining terms $p_1, p_2, \ldots p_{48}$ represent the loan balance after each of 48 monthly payments. After one month the bank charges interest at the rate of 0.416667% and then deducts the amount of the payment. Suppose you pay \$559.99 each month, the same monthly payment you would make if you financed the car through the dealer. Then we would compute the balance, p_1, after the first monthly payment by

$$p_1 = p_0 + \underbrace{0.00416667 p_0}_{\text{interest}} - \underbrace{\$559.99}_{\text{payment}} = \$22,535.84.$$

and we could compute the balance, p_2, after the second monthly payment by

$$p_2 = p_1 + \underbrace{0.00416667p_1}_{\text{interest}} - \underbrace{\$559.99}_{\text{payment}} = \$22,069.75.$$

We can continue computing the balance after each monthly payment by

$$p_n = p_{n-1} + \underbrace{0.00416667p_{n-1}}_{\text{interest}} - \underbrace{\$559.99}_{\text{payment}},$$

for $n = 1, 2, 3, \ldots 48$. This equation is the recursion equation and describes how each term, p_n, of the sequence is computed from the previous term, p_{n-1}. Table 2.3 shows the results of these calculations.

Term	**Balance**	**Term**	**Balance**	**Term**	**Balance**	**Term**	**Balance**
0	\$23,000.00						
1	\$22,535.84	13	\$16,812.78	25	\$10,796.92	37	\$4,473.28
2	\$22,069.75	14	\$16,322.85	26	\$10,281.92	38	\$3,931.92
3	\$21,601.72	15	\$15,830.87	27	\$9,764.77	39	\$3,388.32
4	\$21,131.74	16	\$15,336.84	28	\$9,245.47	40	\$2,842.45
5	\$20,659.80	17	\$14,840.75	29	\$8,724.00	41	\$2,294.30
6	\$20,185.89	18	\$14,342.60	30	\$8,200.36	42	\$1,743.87
7	\$19,710.01	19	\$13,842.37	31	\$7,674.54	43	\$1,191.14
8	\$19,232.14	20	\$13,340.06	32	\$7,146.52	44	\$636.12
9	\$18,752.29	21	\$12,835.65	33	\$6,616.31	45	\$78.78
10	\$18,270.43	22	\$12,329.14	34	\$6,083.89	46	-\$480.88
11	\$17,786.57	23	\$11,820.52	35	\$5,549.25	47	-\$1,042.88
12	\$17,300.69	24	\$11,309.79	36	\$5,012.38	48	-\$1,607.21

Table 2.3: Loan balances

Notice that after the 46^{th} payment the loan balance is negative. When we add interest to p_{45} we get

$$p_{45} + 0.00416667p_{45} = \$79.11.$$

So the 46^{th} payment should have been only \$79.11 and would have paid off the loan. There was no need for payments 47 and 48. This is a much better way to finance the vehicle.

```
In[4]:= Clear[p]

        p[0] = 23000

        p[n_] := p[n] = p[n - 1] + (0.05 / 12) * p[n - 1] - 559.99

        TableForm[Table[{n, p[n]}, {n, 0, 48}]]

Out[5]= 23000

Out[7]//TableForm=
        0   23000
        1   22535.8
        2   22069.8
        3   21601.7
        4   21131.7
        5   20659.8
        6   20185.9
        7   19710.
        8   19232.1
        9   18752.3
        10  18270.4
```

Figure 2.6: A *Mathematica* screenshot with some underlined code

Because computing all the loan balances after all 48 payments is time-consuming and tedious, we usually use either *Mathematica* or a spreadsheet to do the work. See Figure 2.6. Notice that we have underlined the code corresponding to the three key elements of the model in the *Mathematica* notebook.

- The second line in the *Mathematica* notebook corresponds to the initial value $p_0 = \$23,000$. In this line we underlined the code corresponding to the initial value 23 000. This is the code you will change for different initial values. Notice that in *Mathematica* we do not use commas or dollar signs in numbers. *Mathematica* automatically puts a small space where a comma would normally be. You do not type the space.
- The third line in the *Mathematica* notebook corresponds to the recursion equation

$$p_n = p_{n-1} + \left(\frac{0.05}{12}\right) p_{n-1} - 559.99.$$

  In this line we underlined the code that corresponds to the right hand side of the recursion equation. This is the code you will change for different recursion equations.
- The fourth line of the *Mathematica* notebook prints a table showing the terms p_n for $n = 0, 1, 2, \ldots 48$. We have underlined the code that specifies which terms are to be included in the table. This is the code you will change if you are interested

in different terms. To save space we only show part of the output in Figure 2.6 on page 127.

You can use this *Mathematica* notebook to study other models like this by changing the underlined code in the notebook.

Question 7 *You had decided that you could afford payments of $559.99 per month for 48 months but it was a stretch. Could you lower your payments to $539.99 per month and still pay off the loan in 48 months?*

Question 8 *Recall that the credit union will give you a loan at this rate if you make monthly payments of at least $400.00. How long would it take you to pay off this loan if you made the minimum monthly payments?*

Question 9 *Suppose that you have been considering some options that would bring the price of this vehicle up to $24,000. Can you buy the vehicle with these options and still pay $559.99 per month for 48 months?*

Question 10 *The Federal Reserve Bank has just lowered interest rates in response to a sagging economy. As a result, you can get a new car loan at the annual rate of 4.5% per year. How would this affect your car buying options?*

2.2.3 Geometric Sequences

We begin this subsection with an example.

Example 3 *Suppose that your rich uncle gave you a gift of $10,000 the day that you were born and that the money was invested in a fund that pays 6% interest each year with the interest credited on your birthday. How much money would you have after interest was credited on your* 21^{st} *birthday?*

We can express this situation with the recursion equation and initial value

$$p_n = p_{n-1} + 0.06 p_{n-1} = 1.06\ p_{n-1}, \quad p_0 = \$10,000$$

A sequence, like the one produced by this recursion equation, in which the terms change by the same *percentage* amount each time, is called a **geometric** sequence. Notice that

$$\begin{aligned}
p_1 &= 1.06p_0 \\
p_2 &= 1.06p_1 = 1.06(1.06p_0) = (1.06)^2p_0 \\
p_3 &= 1.06p_2 = 1.06((1.06)^2p_0 = (1.06)^3p_0 \\
&\vdots \\
p_n &= (1.06)^n p_0.
\end{aligned}$$

More generally, we can write a geometric sequence using a recursion equation of the form

$$p_n = Rp_{n-1}$$

or by the formula

$$p_n = R^n p_0.$$

As mentioned earlier, a formula like this that expresses the terms p_n by a function

$$p_n = f(p_n)$$

is called a **closed form** or **analytic** solution. For our example, the closed form solution is

$$p_n = (1.06)^n \$10,000.$$

On your 21^{st} birthday you will have

$$p_{21} = (1.06)^{21}\$10,000 = \$33,995.60.$$

The next example is somewhat more complicated and cannot be described be a geometric sequence.

Example 4 *Suppose that your rich uncle gives you a gift of $50,000. Suppose you deposit the gift in a bank account that earns interest at the rate of 5% per year and that each year on the anniversary of the gift after the interest is posted you withdraw $5,000 and use it for a vacation. How long will you be able to continue taking vacations using your uncle's gift?*

We use a_n to denote the money in your bank account at the beginning of the n^{th} year after your uncle's gift. Thus,

$$\begin{aligned} a_1 &= \$50,000 \\ a_n &= a_{n-1} + 0.05a_{n-1} - \$5,000 \\ &= 1.05a_{n-1} - \$5,000. \end{aligned}$$

This leads to Table 2.4. Note that this is not a geometric sequence and so does not have a closed form solution of the form $p_n = R^n p_0$.

Year	Amount	Year	Amount
1	$50,000.00	9	$26,127.23
2	$47,500.00	10	$22,433.59
3	$44,875.00	11	$18,555.27
4	$42,118.75	12	$14,483.03
5	$39,224.69	13	$10,207.18
6	$36,185.92	14	$5,717.54
7	$32,995.22	15	$1,003.42
8	$29,644.98	16	-$3,946.41

Table 2.4: The amount of money in your bank account

Notice that $a_{16} = -\$3,946.41$, which means that after withdrawing $5,000 at the end of the 15^{th} year for your 15^{th} vacation your balance would be negative – your account was $3,946.41 short of what you needed for your 15^{th} vacation. Thus, you will be able to take 14 full vacations and one much smaller vacation.

Question 11 *The model that we used in the preceding example,*

$$\begin{aligned} a_1 &= \$50,000 \\ a_n &= a_{n-1} + 0.05a_{n-1} - \$5,000 \\ &= 1.05a_{n-1} - \$5,000, \end{aligned}$$

reflected the fact that you withdrew $5,000 each year after the interest was paid. How would you change this model if you withdrew $5,000 each year before the interest was paid?

Question 12 *Suppose that you deposit \$3,000 each year on your birthday in a bank account that earns interest at the rate of 4% per year. How much money would you have immediately after making your* 10^{th} *deposit? What assumptions did you make regarding when your bank paid interest? Use a discrete dynamical system to model this situation and answer the questions.*

Question 13 *Suppose that you deposit \$250 each month on the first of the month in a bank account that earns interest at the rate of 4% per year. Suppose that the bank pays interest monthly (at the rate of 1/3% per month). How much money would you have immediately after making your* 120^{th} *deposit? What assumptions did you make regarding when your bank paid interest? Use a discrete dynamical system to model this situation and answer the questions.*

Question 14 *In this question we consider the possibility of setting up a permanent colony on the moon. Suppose that we sent one rocket to the moon each year with 75 colonists for twenty years and that the birthrate on the moon was 5% – that is, each year the number of babies born was 5% of the current population. Assume that the death rate was 1% per year. What would the population of the colony be after twenty years? What assumptions did you make? Use a discrete dynamical system to model this situation and answer the questions.*

Question 15 *Someone dumped 100 tons of garbage on some vacant land. You just purchased the land and plan to build a development on the land. You must remove the garbage. Your trucks can remove 5 tons of garbage each day. If you start removing the garbage now you want to know how much garbage will remain each day. Model this situation with a discrete dynamical system. Find a closed form solution for this discrete dynamical system. Use your closed form solution to determine how long it will take to remove all the garbage.*

Question 16 *Someone dumped 100 tons of a pollutant into a spring-fed lake. You have just purchased the lake and some surrounding land. You plan to use the lake for fish-farming but before you do so you must remove 95% of the pollution. Each week you can remove 5% of the water in the lake and it will be replaced by unpolluted water from the spring. Thus, each week you can lower the amount of pollution in the lake by 5%. Model this situation with a discrete dynamical system. Find a closed form solution for this discrete dynamical system. Use your closed form solution to determine how long it will take to remove 95% of the pollution.*

MONEY MARKET CERTIFICATE RATES

Term	Dividend Rate	APY
6-Month	3.48%	3.50%
1-Year	3.92%	4.00%
2-Year	4.07%	4.15%
3-Year	4.41%	4.50%
4-Year	4.64%	4.75%
5-Year	4.74%	4.85%
7-Year	4.88%	5.00%

$1,000 minimum deposit
Rates are accurate as of 1/31/2008
APY = Annual Percentage Yield

Figure 2.7: Pentagon Federal Credit Union money market rates (31 January 2008)

2.2.4 A Multibillion Dollar Question – Discrete vs. Continuous

Figure 2.7 shows a typical advertisement for money market interest rates. Notice that there are two different interest rates – the "dividend rate" and the "APY", or "annual percentage yield." Actually, if you dig a bit deeper, you will find there are more than two – and it is not just money market certificates but loans and virtually any investment that have a multitude of different "rates" and "yields." This is not a finance course so we will only scratch the surface of this topic. We will, however, discuss the most significant mathematical issue involved – the same issue that is involved in discrete versus continuous dynamical systems.

The fundamental mathematical[6] issue underlying the two rates in Figure 2.7 is when or how often interest (or dividends) is added to the account, or "compounded." We begin with a question. Answer it before going on to the next page.

Question 17 *Suppose that you invest $12,000 in a bank account that earns 4% interest per year with the interest paid annually. How much money will be in the account after ten years?*

[6] You may wonder why we attach the qualifying adjective "mathematical" to the noun "issue." The reason is that there are a number of other equally important issues that are not purely mathematical. For example, when you take out a loan or buy stock there may be various fees involved.

Many investments pay interest, or compound, more often than once a year. The next example looks at one common possibility – quarterly compounding.

Example 5 *Suppose you invest \$12,000 in a bank account that earns 4% interest per year with interest paid quarterly (every three months). How much money will be in the account after ten years?*

Banks often pay interest quarterly or monthly. Some even pay interest daily. Since the interest rate is 4% each year, the bank pays 1% each quarter. We model this situation as follows.

First, we let $n = 0, 1, 2, \ldots 40$ denote the number of quarters after the initial investment. Then,

$$\begin{aligned} p_0 &= \$12,000 \\ p_n &= 1.01 p_{n-1} \end{aligned}$$

Table 2.5 shows the calculations for this example – \$12,000 invested at an annual interest rate of 4% with the interest compounded (paid) quarterly. Table 2.6 on page 134 shows the calculations for Question 17 – \$12,000 invested at an annual interest rate of 4% with the interest compounded yearly.

n	p_n	n	p_n	n	p_n	n	p_n
0	\$12,000.00						
1	\$12,120.00	11	\$13,388.02	21	\$14,788.70	31	\$16,335.93
2	\$12,241.20	12	\$13,521.90	22	\$14,936.59	32	\$16,499.29
3	\$12,363.61	13	\$13,657.12	23	\$15,085.96	33	\$16,664.28
4	\$12,487.25	14	\$13,793.69	24	\$15,236.82	34	\$16,830.92
5	\$12,612.12	15	\$13,931.63	25	\$15,389.18	35	\$16,999.23
6	\$12,738.24	16	\$14,070.94	26	\$15,543.08	36	\$17,169.23
7	\$12,865.62	17	\$14,211.65	27	\$15,698.51	37	\$17,340.92
8	\$12,994.28	18	\$14,353.77	28	\$15,855.49	38	\$17,514.33
9	\$13,124.22	19	\$14,497.31	29	\$16,014.05	39	\$17,689.47
10	\$13,255.47	20	\$14,642.28	30	\$16,174.19	40	\$17,866.36

Table 2.5: Bank account balance – compounded quarterly

Notice that if the interest is compounded quarterly, then, after ten years, there is \$17,866.36 in the bank account; whereas, if the interest is only compounded yearly, the balance is \$17,762.93 (as calculated in Question 17).

0	\$12,000.00
1	\$12,480.00
2	\$12,979.20
3	\$13,498.37
4	\$14,038.30
5	\$14,599.83
6	\$15,183.83
7	\$15,791.18
8	\$16,422.83
9	\$17,079.74
10	\$17,762.93

Table 2.6: Bank account balance – compounded annually

Even after the end of the first year the investment that was compounded quarterly is worth more than the investment that was compounded annually – \$12,487.25 as compared to \$12,480.00. In effect, the investment that was compounded quarterly has earned 4.06% interest in one year as computed by:

$$\frac{\$12,487.25 - \$12,000}{\$12,000} \times 100\% = 4.06\%$$

This is called the **effective annual interest rate**. Bankers use several different terms for variations on these ideas. Sometimes they use the term **annual percentage yield**. See Figure 2.7 on page 132 at the beginning of this subsection. We can calculate the effective annual interest rate another way. Recall that our model for this investment is based on the recursion equation

$$p_n = 1.01 p_{n-1}.$$

Thus, we see that for any quarter k,

$$\begin{aligned}
p_{k+1} &= 1.01 p_k \\
p_{k+2} &= 1.01 p_{k+1} = (1.01^2) p_k \\
p_{k+3} &= 1.01 p_{k+2} = (1.01^3) p_k \\
p_{k+4} &= 1.01 p_{k+3} = (1.01^4) p_k
\end{aligned}$$

and each year (each four quarters) the investment is multiplied by $(1.01)^4 = 1.0406$. This is the same thing as adding 4.06% interest.

Question 18 *Analyze an investment of \$12,000 for ten years in a bank account earning 4% interest compounded monthly. How much is this investment worth at the end of ten years? What is the effective annual interest rate?*

Question 19 *Analyze an investment of \$12,000 for ten years in a bank account earning 4% interest compounded daily. How much is this investment worth at the end of ten years? What is the effective annual interest rate?*

The following questions all involve loans that are repaid in one lump sum rather than in monthly or yearly payments.

Question 20 *Suppose you borrow \$12,000 for a period of ten years at an interest rate of 12% per year and the interest is compounded monthly. How much do you have to repay at the end of ten years? What is the effective annual interest rate?*

Question 21 *Suppose you borrow \$12,000 for a period of ten years at an interest rate of 12% per year and the interest is compounded daily. How much do you have to repay at the end of ten years? What is the effective annual interest rate?*

Question 22 *Suppose you borrow \$12,000 for a period of ten years at an interest rate of 18% per year and the interest is compounded monthly. How much do you have to repay at the end of ten years? What is the effective annual interest rate?*

Question 23 *Suppose you borrow \$12,000 for a period of ten years at an interest rate of 18% per year and the interest is compounded daily. How much do you have to repay at the end of ten years? What is the effective annual interest rate?*

Not surprisingly, banks want to earn as much money as possible from their loans. Suppose that a loan is advertised as having an interest rate of 18% per year. If the interest is compounded every quarter, or four times per year, then each quarter the interest rate is $18\%/4 = 4.5\%$. Thus, if we let p_0 be the initial loan and p_n be the amount owed (with no payments) after n quarters we see that

$$\begin{aligned} p_1 &= \left(1+\frac{0.18}{4}\right)p_0 \\ p_2 &= \left(1+\frac{0.18}{4}\right)^2 p_0 \\ p_3 &= \left(1+\frac{0.18}{4}\right)^3 p_0 \\ p_4 &= \left(1+\frac{0.18}{4}\right)^4 p_0 \approx 1.1925 p_0 \end{aligned}$$

and after four quarters (one year) the effective interest rate is 19.25%. Thus, by compounding the interest quarterly instead of annually the bank would earn an additional 1.25% interest each year.

If the bank compounded the interest monthly instead of quarterly we would have a monthly interest rate of $18\%/12 = 1.5\%$ and

$$\begin{aligned} p_1 &= \left(1+\frac{0.18}{12}\right)p_0 \\ p_2 &= \left(1+\frac{0.18}{12}\right)^2 p_0 \\ &\vdots \\ p_{12} &= \left(1+\frac{0.18}{12}\right)^{12} p_0 \approx 1.1956 p_0 \end{aligned}$$

and after 12 months (one year) the effective interest rate is 19.56%. Thus, by compounding the interest monthly instead of annually the bank would earn an additional 1.56% interest each year.

If the bank compounded the interest daily then we would have

$$\begin{aligned} p_1 &= \left(1+\frac{0.18}{365}\right)p_0 \\ p_2 &= \left(1+\frac{0.18}{365}\right)^2 p_0 \end{aligned}$$

$$\vdots$$

$$p_{365} = \left(1 + \frac{0.18}{365}\right)^{365} p_0 \approx 1.197164 p_0$$

and after 365 days (one year) the effective interest rate is 19.72%. Thus, by compounding the interest daily instead of annually the bank would earn an additional 1.72% interest each year.

You might wonder what would happen if the bank compounded the interest "infinitely" often or "continuously." Of course, compounding the interest infinitely often seems impossible but you could approximate the result by compounding the interest every hour, or every minute, or every second. You would discover that the effective interest rate eventually approaches a limit, about 19.721736%. Later in this book we will talk about continuous dynamical systems. Continuous compounding, or compounding "infinitely often," is modeled by the differential equation

$$\frac{dp}{dt} = 0.18p$$

and has the closed form solution

$$p(t) \approx (1.1972136)^t p_0.$$

Thus, a credit card with an advertised rate of 18% per year can actually earn up to 19.72% depending on how often they compound interest. The difference between 18% and 19.72% is huge if you are making billions of dollars worth of loans.

We close this section by listing explicitly five steps that are very useful for developing a model. These steps are all part of the *transform* arrow in the modeling triangle (Figure 2.1 on page 113) – expressing a situation in the real world mathematically. Following these steps carefully can help you organize your model development and avoid some common mistakes. Notice that we list these steps using a bulleted list rather than a numbered list because you may not always follow these five steps in exactly the order below.

- **Define the variables of interest:** These variables identify the quantities of interest – that is, the quantities that you are modeling or tracking. For example, you might be interested in the amount of money in a bank account, the average personal income of the residents of a particular country, or the amount of oil in worldwide known oil reserves. As part of identifying and defining the variables of interest, you need to describe the units in which they are measured. For example, if you are tracking the average personal income in a particular country you need to specify whether income is measured in local currency or in some common currency[7]. In addition, you need to specify whether or not income is measured in an inflation adjusted currency.

- **Identify the domain:** In most of the models in this book, we are tracking how the variables of interest change over time. This is not, however, always the case. For example, we might be interested in tracking how the amount of dissolved oxygen in the ocean changes as the depth changes. If we are tracking a particular variable, say p over time or, for example, as depth changes, we use a subscript to denote the value of that variable at a particular time or depth. That is, we use notation like p_n for the value of p at time n or at depth n. We need to specify the units in which n is measured and the domain of n. For example, if we are tracking the amount of money owed on a car loan over a period of 60 months, we would use the notation p_n to denote the amount of money owed after n monthly payments. The subscript n would be measured in payments or months and n would have the values $n = 0, 1, 2, 3, \ldots 60$. The number p_0 would represent the initial amount borrowed; the number p_1 would represent the amount owed one month after the loan was taken out and after the first payment; the number p_2 would represent the amount owed two months after the loan was taken out and after the second payment; and the number $p_{60} = \$0.00$ would represent the final loan balance 60 months after the loan was taken out and after the final payment. Notice that the first payment on the loan is made one month after the loan is taken out and the last payment is made 60 months after the loan is taken out. This is the most common practice but it is possible to set up different loan payment schedules. Whatever schedule is used is an important part of the model.

- **List assumptions:** This is often a key part of building a model and in many cases assumptions are implicit rather than explicit – that is, people may not state explicitly all the assumptions that underlie the model. It is easy to go overboard listing assumptions but it is also easy to err in the other direction, ignoring assumptions that are important. For example, if we are modeling the price of gasoline we might assume that political conditions in oil-producing countries are stable. This is a very significant assumption. On the other hand, we might also assume that we are not visited by an interplanetary group of oil traders based on the planet Neptune that offers us regular and reliable oil deliveries in any quantity at a price of \$8.00 per barrel. This is a frivolous assumption.

[7] Personal income is a tricky business. Two commonly used measures are U.S. dollars in a fixed year and "purchasing power parity." The former is good if we are primarily interested in how residents of a particular country compete on an international market – for example, can they afford to buy gasoline or travel to international destinations or can they afford to send their kids to college in the United States or Europe. The latter is useful if we are interested in local standards of living – how much food can they buy or what kind of housing can they afford.

- **State the initial value:** If, for example, we were tracking the amount owed on a loan of \$15,000, the initial value would be $p_0 = \$15,000$.
- **Determine the difference equation or the recursion equation that describes how each term of the sequence is computed from the preceding term:** If, for example, the loan was paid off in *monthly* installments of \$350 at an *annual* interest rate of 6%, this would be

$$p_n = p_{n-1} + \left(\frac{0.06}{12}\right) p_{n-1} - \$350.$$

or

$$p_n = p_{n-1} + 0.005 p_{n-1} - \$350.$$

We use three slightly different but equivalent forms to describe how each term is computed from the previous term.

- The first form focuses on the difference between the current term and the preceding term – for example,

$$\underbrace{p_n}_{\text{current}} - \underbrace{p_{n-1}}_{\text{preceding}} = \underbrace{0.005p_{n-1} - \$350}_{\text{difference}}.$$

- The second form expresses the current term as the present term plus some difference – for example,

$$\underbrace{p_n}_{\text{current}} = \underbrace{p_{n-1}}_{\text{preceding}} + \underbrace{0.005p_{n-1} - \$350}_{\text{difference}}.$$

- The third form expresses the current term as a function of the preceding term – for example,

$$\underbrace{p_n}_{\text{current}} = 1.0005 \underbrace{p_{n-1}}_{\text{preceding}} - \$350.$$

2.3 Linear Discrete Dynamical Systems

This section looks at linear discrete dynamical systems. Linear discrete dynamical systems are important for three reasons.

- They are easy to use.
- They can be used to model many important situations.
- They illustrate in simple form some of the most interesting ideas that come up when we study and apply discrete dynamical systems.

Throughout this chapter we use two different kinds of graphs to help us visualize discrete dynamical systems given by an initial condition

$$p_0 = \underline{\hspace{2em}}$$

and a recursion equation

$$p_n = f(p_{n-1}).$$

- **Time series graphs** focus our attention on the sequence $p_0, p_2, \ldots p_n$.
- **Fundamental graphs** focus our attention on the recursion equation – that is, on the way the value of the quantity p changes from one term to the next.

In the two subsections below we illustrate these two kinds of graphs with the same discrete dynamical system.

$$p_n = 0.75p_{n-1} + 100, \quad p_0 = 50.$$

2.3.1 Time Series Graphs

Table 2.7 on page 141 shows the first thirty terms of the sequence generated by this discrete dynamical system and Figure 2.8 on the same page shows exactly the same information as a graph. This is an example of a **time series** graph. It focuses our attention on the values

n	p_n	n	p_n	n	p_n
0	50.00				
1	137.50	11	385.22	21	399.17
2	203.13	12	388.91	22	399.38
3	252.34	13	391.68	23	399.53
4	289.26	14	393.76	24	399.65
5	316.94	15	395.32	25	399.74
6	337.71	16	396.49	26	399.80
7	353.28	17	397.37	27	399.85
8	364.96	18	398.03	28	399.89
9	373.72	19	398.52	29	399.92
10	380.29	20	398.89	30	399.94

Table 2.7: $p_n = 0.75p_{n-1} + 100, \quad p_0 = 50$

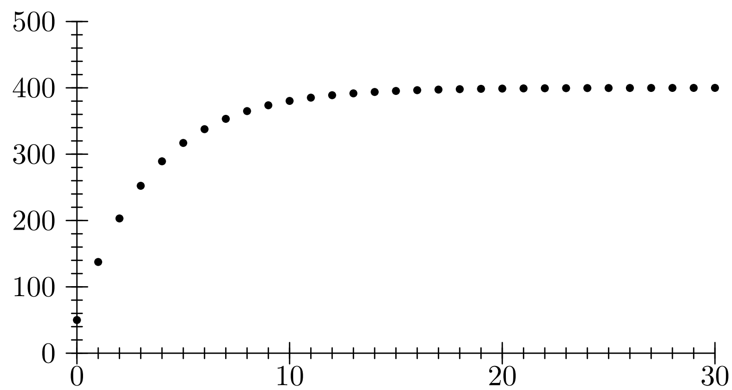


Figure 2.8: Time series graph for $p_n = 0.75p_{n-1} + 100$

of the quantity p over a period of time – in this case over the period $n = 0, 1, \ldots 30$. Figure 2.9 on page 142 shows *Mathematica* code that you can use as a template for making time series graphs by modifying the underlined code.

Question 1 *Draw a time series graph for the discrete dynamical system*

$$p_n = 0.90p_{n-1} + 100, \quad p_0 = 50.$$

Question 2 *Draw a time series graph for the discrete dynamical system*

$$p_n = 1.10p_{n-1} + 100, \quad p_0 = 50.$$

```
Clear[p]

p[0] = 50;

p[n_] := p[n] = 0.75 * p[n - 1] + 100

ListPlot[Table[{n, p[n]}, {n, 0, 30}], PlotStyle → {PointSize[0.015]}, PlotRange → {0, 500}]
```

Figure 2.9: *Mathematica* code for making time series graphs

2.3.2 Fundamental Graphs

Figure 2.10 is the fundamental graph for the same discrete dynamical system,

$$p_n = 0.75p_{n-1} + 100, \quad p_0 = 50,$$

that we used as an example in the previous subsection.

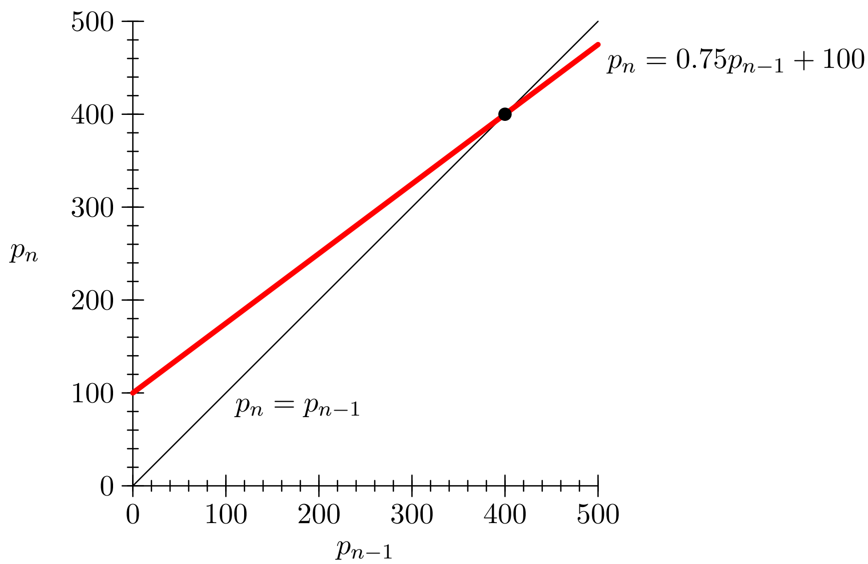


Figure 2.10: Fundamental graph for $p_n = 0.75p_{n-1} + 100$

The fundamental graph focuses on the recursion equation – that is, on the way the value of the quantity p changes from one term to the next. Think of the recursion equation,

$$p_n = 0.75p_{n-1} + 100,$$

as a function whose input (or independent variable) is denoted p_{n-1} and whose output (or dependent variable) is denoted p_n. In other words this is a machine for computing p_n from p_{n-1}. We have labeled the axes in Figure 2.10 on page 142 using this notation. The graph of the function

$$p_n = 0.75p_{n-1} + 100$$

is a thick red (if you have a color copy) line. Notice it is just a linear function whose y-intercept is 100 and whose slope is 0.75.

We also draw a second graph, the graph of the function

$$p_n = p_{n-1},$$

on the fundamental graph. This graph is a thin black line. This graph represents the world's most boring recursion equation – the one that represents an absolutely unchanging situation – where $p_0 = p_1 = p_2 = p_3 = \cdots$ and so on forever. By comparing these two graphs on the fundamental graph we can see some interesting things.

- If the thick red graph, the graph of $p_n = 0.75p_{n-1} + 100$, is above the thin black graph, the graph of $p_n = p_{n-1}$, then p_n will be greater than p_{n-1}. In other words the quantity p will increase. Notice that in Figure 2.10 on page 142 whenever $p_{n-1} < 400$ the thick red graph is above the thin black graph and p_n will be larger than p_{n-1}. That is why the terms $p_0, p_1, p_2, \ldots$ in Figure 2.8 on page 141 keep increasing – each term is below 400 and therefore the next term is larger.

- If the thick red graph, the graph of $p_n = 0.75p_{n-1} + 100$, is below the thin black graph, the graph of $p_n = p_{n-1}$, then p_n will be less than p_{n-1}. In other words the quantity p will decrease. Notice that in Figure 2.10 on page 142 whenever $400 < p_{n-1}$ the thick red graph is below the thin black graph and p_n will be less than p_{n-1}.

- At points where the two graphs cross, for example $p_{n-1} = 400$ in Figure 2.10 on page 142, the value of p will not change. As an example, notice that if we plug this value $p_{n-1} = 400$ into the recursion equation

  $$p_n = 0.75p_{n-1} + 100$$

  we get

  $$p_n = 0.75(400) + 100 = 300 + 100 = 400$$

  Points like this are called *equilibrium points.* They are extremely important and the subject of Section 2.7.

Figure 2.11 shows *Mathematica* code that you can use as a template for making fundamental graphs by modifying the underlined code.

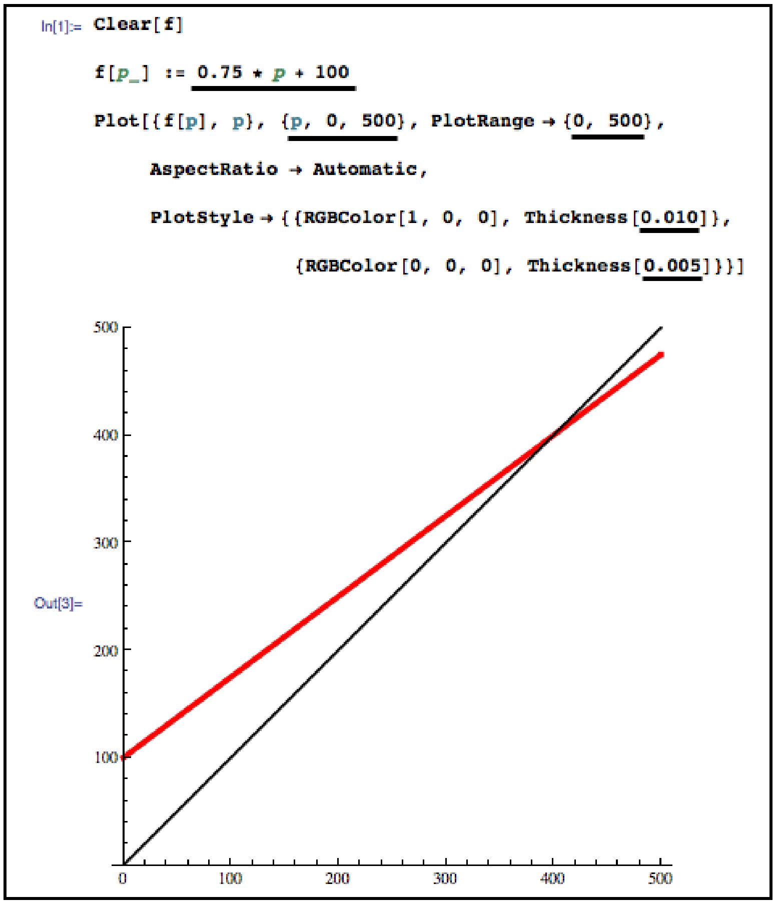


Figure 2.11: *Mathematica* code for making fundamental graphs

Question 3 *Draw a fundamental graph for the discrete dynamical system*

$$p_n = 1.10p_{n-1} - 25, \quad p_0 = 300.$$

Discuss what this fundamental graph tells us about how the quantity p changes.

Question 4 *Draw a fundamental graph for the discrete dynamical system*

$$p_n = 1.10p_{n-1} + 25, \quad p_0 = 300.$$

Discuss what this fundamental graph tells us about how the quantity p changes.

We now turn to the primary topic of this section – linear discrete dynamical systems.

2.3.3 Linear Discrete Dynamical Systems

We begin with some applications that come from very different places but that use exactly the same mathematics.

Example 1 *Barren Island is a miserable and inhospitable island off the coast of a lush and much more hospitable mainland. There is a colony of 1200 birds on the island. Because the island is so inhospitable, the population would drop by 20% each year if the island were isolated. Each year, however, 500 birds arrive from the mainland and make their homes on the island. Model this situation and predict the future bird population on the island.*

We follow the steps discussed at the end of the section 2.2 as we model this situation.

- Defining the variables of interest: The variable of interest is the bird population on Barren Island. We use the notation p_n to denote the bird population measured in numbers of birds on the island in the year n.
- Identifying the domain: Because the bird population varies over the course of each year, it needs to be measured at a particular time each year, say May 15. Time is measured in years. Typically we start with the first year for which we have reliable data or we start with the year in which the model is built. The domain for population models is usually a few tens of years but this depends on the use to which the model will be put and how confident we are in the long-term usefulness of the model. As described above, we use the notation n to denote the year and let $n = 1, 2, \ldots$ with $n = 1$ being the current year.
- Stating the initial value: The initial value is

$$p_1 = 1,200.$$

- Determining the recursion equation: The recursion equation is

$$p_n = p_{n-1} - \underbrace{0.20p_{n-1}}_{\text{natural drop}} + \underbrace{500}_{\text{immigration}}$$

Notice that one term represents the natural population drop of 20% that would occur if there were no immigration and another term represents the immigration. Note also that we are assuming that the immigration occurs after the natural population drop of 20%. With a little algebra this becomes

$$\begin{aligned} p_n &= p_{n-1} - 0.20p_{n-1} + 500 \\ &= 0.80p_{n-1} + 500 \end{aligned}$$

This model and any model that can be written in the form

$$p_n = mp_{n-1} + b,$$

where m and b are constants, is called a **linear discrete dynamical system**. The adjective "linear" comes from the fact that the function

$$p_n = mp_{n-1} + b$$

that appears as the recursion equation is linear. Its graph, the fundamental graph, is a straight line. If the constant b is zero then this is called a **homogeneous linear discrete dynamical system** and is just a geometric model. If the constant b is not zero then this is called a **nonhomogeneous linear discrete dynamical system.** Figure 2.12 on page 147 and Table 2.8 on page 147 show some predictions made by this model.

Notice that because this is a linear system, the fundamental graph is a straight line. The time series graph, Figure 2.12, however, is not a straight line.

- Listing assumptions:

This model has many assumptions. For example, we are assuming that the climate on the island and immigration patterns remain stable, so that the same recursion equation applies during the lifetime of the model.

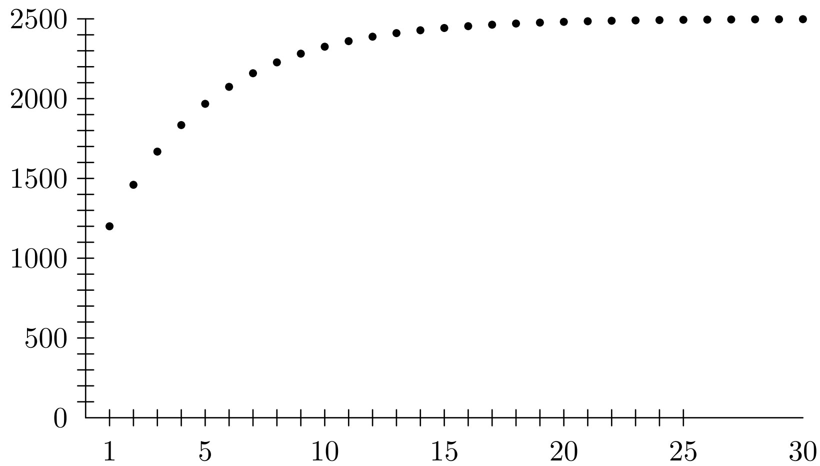


Figure 2.12: Birds on Barren Island

Year	Population	Year	Population	Year	Population
1	1200.00	11	2360.41	21	2485.01
2	1460.00	12	2388.33	22	2488.01
3	1668.00	13	2410.66	23	2490.41
4	1834.40	14	2428.53	24	2492.33
5	1967.52	15	2442.83	25	2493.86
6	2074.02	16	2454.26	26	2495.09
7	2159.21	17	2463.41	27	2496.07
8	2227.37	18	2470.73	28	2496.86
9	2281.90	19	2476.58	29	2497.49
10	2325.52	20	2481.27	30	2497.99

Table 2.8: Birds on Barren Island

Example 2 *You are saving[8] to buy a car. You currently have \$4,000[9] and plan to save an additional \$200 each month for 24 months. You are keeping this money in a bank account that earns interest at the annual rate of 3% compounded monthly. How much will you have after 24 months?*[10]

The annual interest rate of 3% results in a monthly rate of 0.25%. We assume that each month you add \$200 to the account after interest is paid. We model this using the

[8]**Defining the variables of interest:** The variable of interest p_n is the amount of money in a savings account.

[9]**Stating the initial value:** The initial value is $p_0 = \$4,000$.

[10]**Identifying the domain:** The domain is $n = 0, 1, 2, \dots 24$, measured in months after now.

linear model[11]

$$\begin{aligned} p_0 &= \$4,000 \\ p_n &= p_{n-1} + 0.0025p_{n-1} + \$200 \\ &= 1.0025p_{n-1} + \$200 \end{aligned}$$

where p_n denotes the amount of money in your new car fund after n additional payments and $n = 0, 1, 2, \ldots 24$. The results[12] are shown in Table 2.9. Notice that after 24 payments you will have \$9,187.59 in your new car fund.

Month	Balance	Month	Balance
0	\$4,000.00		
1	\$4,210.00	13	\$6,771.33
2	\$4,420.53	14	\$6,988.26
3	\$4,631.58	15	\$7,205.73
4	\$4,843.16	16	\$7,423.74
5	\$5,055.26	17	\$7,642.30
6	\$5,267.90	18	\$7,861.41
7	\$5,481.07	19	\$8,081.06
8	\$5,694.77	20	\$8,301.26
9	\$5,909.01	21	\$8,522.02
10	\$6,123.78	22	\$8,743.32
11	\$6,339.09	23	\$8,965.18
12	\$6,554.94	24	\$9,187.59

Table 2.9: New Car Fund

As you answer the questions below, go through the five modeling steps explicitly.

Question 5 *A particular area has a thriving rabbit population. If the rabbits were left to themselves, their population would increase by 30% each year. However, people from a nearby town hunt the rabbits for food. If the current rabbit population is 5,000 rabbits and each year the townspeople kill 2,000 rabbits, describe what will happen. What would happen if the townspeople killed only 1,000 rabbits each year? What advice would you give the townspeople?*

Question 6 *Three insurance companies are trying to sell you life insurance policies. All three policies offer the same benefits and coverage. The premium for one policy is currently*

[11]**Determining the recursion equation:** The recursion equation is $p_n = 1.0025p_{n-1} + \$200$.
[12]**Listing assumptions:** We are assuming that all payments are made on schedule.

\$1,500 per year and will rise by 1% per year. The premium for the second policy is currently \$1,200 per year and will rise by \$100 per year. The premium for the third policy is currently \$1350 and will rise by 2% per year. Compare the premiums for the three policies.

Question 7 *A storage facility currently has 400 kilograms of a particular radioactive isotope. Left by itself this isotope will decay at the rate of 15% per year. Each year an additional 150 kilograms of the isotope is placed in storage. How many kilograms of the isotope will be in the storage facility in ten years?*

Question 8 *You are about to buy some coffee. One stand sells coffee that is 175 degrees Fahrenheit and another sells coffee that is 195 degrees Fahrenheit. Suppose that coffee is safe to drink when its temperature is 143 degrees Fahrenheit and that you prefer coffee whose temperature is above 125 degrees Fahrenheit. The dining room is 68 degrees Fahrenheit. You have a choice between a small cup of coffee and a large mug. If you buy a small cup, then each minute it will lose 10% of the difference between its temperature and room temperature. If you buy a large mug, it will only lose 4% of this difference each minute. Discuss the pros and cons of your four choices.*

Question 9 *The parents of a new baby decide that each year they will save some money to help with her college expenses. Because they don't have much money when the baby is born but expect their incomes to rise, they plan to deposit \$1,000 times the child's age in years into the bank each year. Thus, on the baby's first birthday they will start the child's education fund with \$1,000. On the child's second birthday, they will add \$2,000 and so forth. Assume that the account pays interest at the rate of 6% per year. How much money will the child have after the deposit is made on her* 18^{th} *birthday. Is your model a linear model?*

2.4 Supply and Demand

Figure 2.13: Buying fresh tomatoes in the summer

In this section we will see one way in which economists use discrete dynamical systems to understand our complex economy. We begin by looking at models in which prices are determined by supply and demand and we conclude this section by using what we've learned to help us understand what might happen when the price of raw materials rises. This kind of modeling can, for example, provide some insight into the ripple effects of rising oil prices.

Toward the end of the summer, tomatoes are not only wonderful; they are cheap. However, when tomatoes are out-of-season, the quality is low and the price is high. Although it doesn't seem fair to have to pay more for inferior tomatoes, this is exactly what is predicted by the Law of Supply and Demand.

The Law of Supply and Demand is the dream of free market economists. It is used to describe how prices are determined in a situation with many independent producers and many independent buyers.

We will illustrate the Law of Supply and Demand with a hypothetical product called Byties. Byties are baked each night by many independent bakers and are sold each morning in the town square to many independent buyers. Each baker decides how many Byties to bake or even whether to bake any Byties at all on the basis of the price of Byties on the

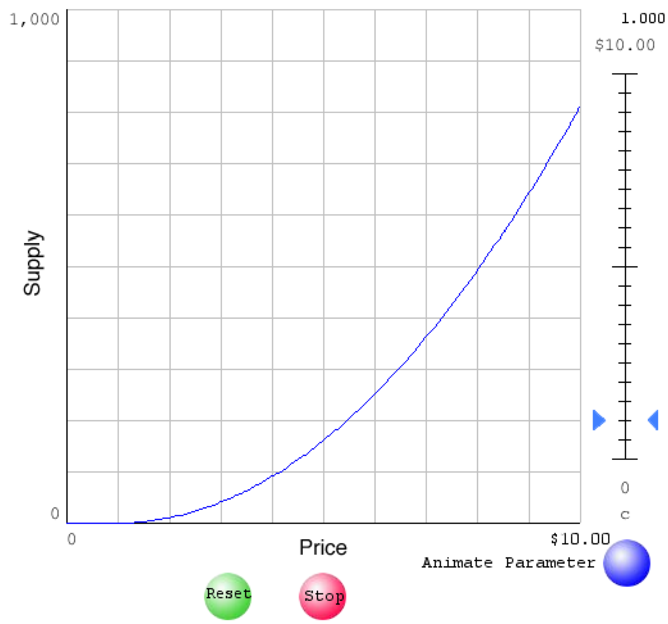


Figure 2.14: An example of a supply function

preceding day. When the price is high, they bake many Byties and when the price is low they only bake a few. If the price is so low that bakers can't recover the costs of ingredients, they won't bake any Byties at all. The number of Byties bakers bake is called the supply of Byties. This relationship between the price of Byties and the supply of Byties is called the supply function because the supply (the dependent variable) of Byties is determined by their price (the independent variable). Figure 2.14 shows an example of a supply function, the function

$$S(p) = \begin{cases} 0, & \text{if } p \leq 1; \\ 10\,(p-1)^2\,, & \text{if } 1 < p. \end{cases}$$

Notice that when the price is higher, so is the supply. This matches our intuition since we would expect people to be willing to bake more Byties when their profit would be higher. Although this is generally what we would expect, we can also imagine other scenarios. For example, when the price dropped, producers might actually bake more Byties in an attempt to make up in volume what they are losing on individual sales.

The supply function $S(p) = 10(p-1)^2$ shown in Figure 2.14 is the kind of supply function one would expect if the cost of ingredients was $1.00 because bakers won't make

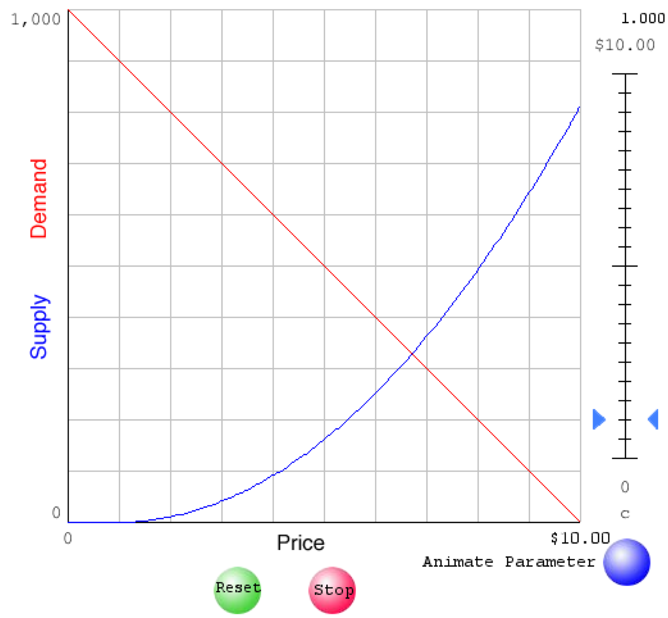


Figure 2.15: An example of a supply and a demand function

any profit at all unless the price is above \$1.00 and, thus, the supply function is zero unless the price is greater than \$1.00. Since we want to be able to study what happens when the cost of ingredients changes, we are interested in the supply function below, in which the fixed cost \$1.00 has been replaced by a parameter c representing a cost that might possibly change. When $c = \$1.00$ this supply function is exactly the same as the one above.

$$S(p) = \begin{cases} 0, & \text{if } p \leq c; \\ 10\,(p-c)^2\,, & \text{if } c < p. \end{cases}$$

There is a live version of Figure 2.14 (page 151) on the web. Click here[13] and then follow the instructions to experiment with different values of the parameter c that represents the cost of ingredients.

Each day the buyers decide whether to go to the town square and how many Byties to buy based on the price the preceding day. This is called the demand for the product and the relationship between the price for a product and the demand is called the demand function. Figure 2.15 shows an example of a demand function, the function

[13] http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/liveSupply.html

$$D(p) = \begin{cases} 100(10-p), & \text{if } p \le 10; \\ 0, & \text{if } 10 < p, \end{cases}$$

added to our earlier graph.

Notice that as the price rises, the demand falls. Again, this is what we would generally expect but we can imagine other scenarios. For example, sometimes people actually buy more when a product is more expensive because they equate price with quality. There are many examples of this phenomenon – for example, people sometimes avoid cheap hotels and motels because they assume the quality is lower.

Notice that when the price is relatively low the demand is larger than the supply and when the price is relatively high the demand is lower than the supply. There is a price, slightly below \$7.00, at which the supply and demand are equal. On the graph this is the point at which the graphs of the supply function and the demand function intersect. This price is called the **equilibrium price.** This is the ideal price because the bakers are selling all the Byties they bake and the buyers are able to buy all the Byties they want to buy. We can determine the equilibrium price algebraically as shown below.

$$\begin{aligned} D(p) &= S(p) \\ 100(10-p) &= 10(p-1)^2 \\ 1000 - 100p &= 10p^2 - 20p + 10 \\ 10p^2 + 80p - 990 &= 0 \\ p^2 + 8p - 99 &= 0 \end{aligned}$$

Applying the quadratic formula we see that the equilibrium price is roughly \$6.72. Note that the quadratic formula yields two roots but the second one is negative and irrelevant in this situation.

Click here[14] for a live version of Figure 2.15 on the Web. Follow the instructions to see how different values for the parameter[15] c affect the equilibrium price.

Question 1 *Suppose that the cost of ingredients rises to \$2.00. What happens to the equilibrium price?*

[14]http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/liveDemand.html

[15]Recall that this parameter represents the cost of ingredients.

The change in the equilibrium price is only the beginning of the story. When you answered the question above, you probably discovered that the equilibrium price went up. At the new and higher equilibrium price the demand will be lower than it was at the old equilibrium price. This will affect bakers because they will sell fewer Byties. They might also be affected in other ways. When the cost of ingredients went up their profit per Bytie might go down. Whether and how much it goes down depends on how much the selling price (the new equilibrium price) has risen. The next question asks you to look at the effects of the rise in the cost of ingredients on the various players – bakers and buyers. This question is part of the interpret step in the modeling triangle. We are interested in what our mathematical model can tell us about the real world. For example, if you are a baker you want to know whether you should look for another job and this model can shed some light on that question.

Question 2 *What other effects are caused by the rise in the cost of ingredients? Think about what happens from the point-of-view of the bakers. Think about what happens from the point-of-view of the buyers.*

Question 3 *Consider the supply and demand functions:*

$$\begin{aligned} S(p) &= 10,000(p-2) \\ D(p) &= 5,000(20-p) \end{aligned}$$

Find the equilibrium price.

Question 4 *Consider the supply and demand functions:*

$$\begin{aligned} S(p) &= 10,000(p-c) \\ D(p) &= 5,000(20-p) \end{aligned}$$

where c is a parameter representing the cost of ingredients. How does the equilibrium price depend on the cost of ingredients?

Now, we are ready to talk about how the marketplace works and how prices change – that is, we will look at a dynamic model in which prices change rather than a static market in which the price somehow has found its equilibrium. We will illustrate this discussion with an example – the familiar baked good, Byties. Suppose that Byties are wonderful when they are fresh but stale Byties aren't very good. As a result, bakers must sell their Byties each day because day-old Byties aren't worth very much. There are many products like this. For example, an empty airplane seat becomes stale as soon as the plane takes off.

Early each morning bakers bring their Byties to the town square. A few minutes later customers begin to arrive and a few hours later that same day there are three possibilities that will affect the price of Byties the next day:

- If the price is the equilibrium price, then everybody is happy. The bakers sell all their Byties and the customers are able to buy all the Byties they want.
- If the price is above the equilibrium price, then there are more Byties than the customers want to buy and bakers are faced with unsold Byties that are about to become stale. In this situation, customers begin to bargain and bakers begin to put their remaining Byties "on sale." As a result, the price of Byties falls.
- If the price is below the equilibrium price, then there are more customers than Byties and the bakers are able to raise their prices for the remaining Byties.

One simple model that captures this general behavior is given by a recursion equation of the form

$$p_n = p_{n-1} + k\left(D\left(p_{n-1}\right) - S\left(p_{n-1}\right)\right),$$

where k is a positive constant whose value is determined by the nature of the marketplace and the bakers and customers. Notice that today's price, p_n, is determined by the supply and the demand based on yesterday's price, p_{n-1}. The quantity $\left(D\left(p_{n-1}\right) - S\left(p_{n-1}\right)\right)$ is the difference between the demand and the supply based on yesterday's price. Notice that, since k is a positive constant, if the demand is greater than the supply then the price will go up and, if the demand is lower than the supply, then the price will go down.

Example 1 *Suppose the supply and demand functions for Byties are*

$$\begin{aligned} S(p) &= 10,000(p-2) \\ D(p) &= 5,000(20-p) \end{aligned}$$

This leads to the recursion equation

$$
\begin{aligned}
p_n &= p_{n-1} + k\left(D\left(p_{n-1}\right) - S\left(p_{n-1}\right)\right) \\
p_n &= p_{n-1} + k\left(5,000\left(20 - p_{n-1}\right) - 10,000\left(p_{n-1} - 2\right)\right) \\
p_n &= p_{n-1} + k\left(100,000 - 5,000p_{n-1} - 10,000p_{n-1} + 20,000\right) \\
p_n &= p_{n-1} + k\left(120,000 - 15,000p_{n-1}\right) \\
p_n &= p_{n-1} + 15,000k\left(8 - p_{n-1}\right)
\end{aligned}
$$

We can rewrite this last equation as

$$p_n = (1 - 15,000k)p_{n-1} + 120,000k$$

and we see that this is a linear discrete dynamical system.

Now, suppose that the value of the constant k is 0.00004 and that the initial price is \$4.00. Thus, our complete model is

$$
\begin{aligned}
p_0 &= \$4.00 \\
p_n &= p_{n-1} + 15,000(0.00004)\left(8 - p_{n-1}\right) \\
&= 0.40p_{n-1} + 4.8
\end{aligned}
$$

Table 2.10 on page 157 shows the results computed using this model for the next ten days.

The value of the constant k represents the volatility of the marketplace and depends on customs and the nature of both the bakers and the customers. Some markets are very quick to adjust prices – both customers and sellers like to bargain. Others are slower to adjust prices. Notice that in this example, the initial price was below the equilibrium price but within nine days it was within a fraction of a cent of the equilibrium price. The following questions are also part of the interpret step in the modeling triangle.

Question 5 *Experiment with Example 1. What happens if the initial price is above the equilibrium price?*

Day	Price
1	\$4.00
2	\$6.40
3	\$7.36
4	\$7.74
5	\$7.90
6	\$7.96
7	\$7.98
8	\$7.99
9	\$8.00
10	\$8.00

Table 2.10: The price of Byties

Question 6 *Continue your experimentation with Example 1.*

- *What happens if the marketplace is less volatile – for example, what happens if* $k = 0.00002$?
- *What happens if the marketplace is more volatile – for example, what happens if* $k = 0.00008$?
- *See if you can find a general description of what happens with various values of the volatility constant* k.

The last question looks at some of the downstream or "ripple" effects of a rise in the cost of ingredients. This kind of modeling is important when we try to analyze the effects of rises in oil prices.

Question 7 *Example 1 uses a supply function that would be appropriate if the cost of the ingredients for Byties was \$2.00. Suppose that the price is at equilibrium and the cost of the ingredients for Byties suddenly rises to \$4.00. Describe what happens. How does the volatility of the marketplace affect what happens? Describe the effects of the rise in the cost of ingredients on both bakers and buyers.*

2.5 Long Term Behavior and Limits

2.5.1 Long Term Behavior and Limits

In this section we develop notation and terminology for discussing the long term behavior of sequences generated by discrete dynamical systems. We begin with an example.

Example 1 *Consider the discrete dynamical system*

$$p_n = 0.75p_{n-1} + 100, \quad p_0 = 50$$

n	p_n	n	p_n	n	p_n
0	50.00				
1	137.50	11	385.22	21	399.17
2	203.13	12	388.91	22	399.38
3	252.34	13	391.68	23	399.53
4	289.26	14	393.76	24	399.65
5	316.94	15	395.32	25	399.74
6	337.71	16	396.49	26	399.80
7	353.28	17	397.37	27	399.85
8	364.96	18	398.03	28	399.89
9	373.72	19	398.52	29	399.92
10	380.29	20	398.89	30	399.94

Table 2.11: $p_n = 0.75p_{n-1} + 100, \quad p_0 = 50$

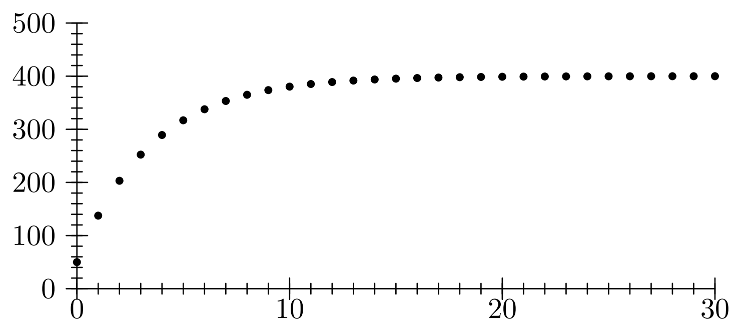


Figure 2.16: $p_n = 0.75p_{n-1} + 100$

The first thirty terms of the resulting sequence are shown in Table 2.11 and Figure 2.16. Notice that the sequence is getting closer and closer to the number 400. In this situation we say that the sequence **converges** to the number 400. We make this a formal definition.

Definition 1 *If after a long time (that is, for large n) the terms of a sequence $p_1, p_2, p_3, \ldots$ are getting arbitrarily close to a number L then we say that the sequence is* **converging** *to the number L and write*

$$\lim_{n\to\infty} p_n = L.$$

This is sometimes read "the limit of the sequence $p_1, p_2, p_3, \ldots$ is L."

Notice that usually the sequence never actually reaches the limit. It just gets very, very close.

Example 2 *Consider the discrete dynamical system given by the initial value and recursion equation*

$$p_1 = 200, \quad p_n = 0.80 p_{n-1}.$$

The first 20 terms of this sequence are shown graphically in Figure 2.17.

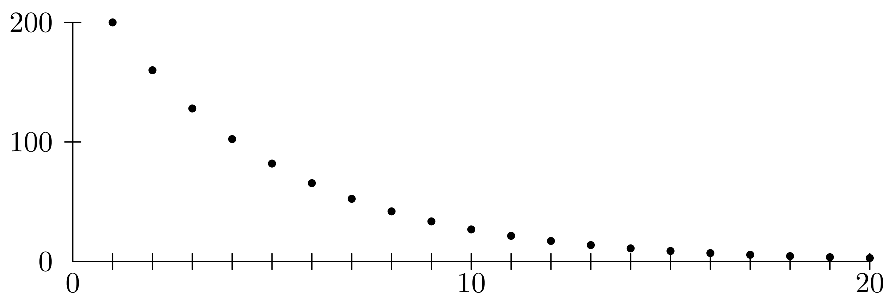


Figure 2.17: Example 2

Because this sequence is a geometric sequence it has a closed form solution

$$p_n = 200(0.80)^n$$

We can use this closed form solution to calculate some additional sample terms.

$$\begin{aligned} p_{30} &= 200(0.80)^{29} = 0.3095 \\ p_{50} &= 200(0.80)^{49} = 3.56 \times 10^{-3} \\ p_{100} &= 200(0.80)^{99} = 5.09 \times 10^{-8} \end{aligned}$$

Notice that for very large values of n, p_n is getting close to zero. So we say that this sequence converges to zero and write

$$\lim_{n\to\infty} p_n = 0.$$

We can see why this sequence converges to zero by looking at its closed form solution

$$p_n = 200(0.80)^{n-1}.$$

Because (0.80) is less than one, as n gets bigger and bigger $(0.80)^n$ gets smaller and smaller and thus $p_n = 200(0.80)^n$ also gets smaller and smaller.

Question 1 *Consider the discrete dynamical system given by the initial value and recursion equation*

$$p_1 = 50, \quad p_n = 0.80p_{n-1} + 200.$$

Does the sequence produced by this discrete dynamical system converge? If so, determine its limit.

Question 2 *Consider the discrete dynamical system given by the initial value and recursion equation*

$$p_1 = 500, \quad p_n = 0.80p_{n-1} + 200.$$

Does the sequence produced by this discrete dynamical system converge? If so, determine its limit.

Question 3 *Consider the discrete dynamical system given by the initial value and recursion equation*

$$p_1 = 5,000, \quad p_n = 0.80p_{n-1} + 200.$$

Does the sequence produced by this discrete dynamical system converge? If so, determine its limit.

Table 2.12 on page 161 looks at the first twenty terms of the sequence produced by the discrete dynamical system

$$p_n = 0.5p_{n-1} + 200, \quad p_0 = 50.$$

n	p	p	p	p
0	50	50.0	50.00	50.000
1	225	225.0	225.00	225.000
2	313	312.5	312.50	312.500
3	356	356.3	356.25	356.250
4	378	378.1	378.13	378.125
5	389	389.1	389.06	389.063
6	395	394.5	394.53	394.531
7	397	397.3	397.27	397.266
8	399	398.6	398.63	398.633
9	399	399.3	399.32	399.316
10	400	399.7	399.66	399.658
11	400	399.8	399.83	399.829
12	400	399.9	399.91	399.915
13	400	400.0	399.96	399.957
14	400	400.0	399.98	399.979
15	400	400.0	399.99	399.989
16	400	400.0	399.99	399.995
17	400	400.0	400.00	399.997
18	400	400.0	400.00	399.999
19	400	400.0	400.00	399.999
20	400	400.0	400.00	400.000

Table 2.12: The discrete dynamical system $p_n = 0.5p_{n-1} + 200, \quad p_0 = 50$

All four columns labeled p in this table have exactly the same entries except that the entries are rounded differently. In the first of these columns the entries are rounded to the nearest integer; in the second they are rounded to one digit past the decimal point; in the third to two digits; and in the fourth to three digits. Notice that this sequence appears to converge to the limit 400. It appears that the sequence actually reaches its limit but this is an illusion created by roundoff. Notice in the first p column, where we round to the nearest integer, the sequence appears to reach its limit by $n = 10$. But, in the second, where we round to one digit past the decimal point, we have to wait until $n = 13$ before the sequence appears to reach its limit. In the third column, rounding to two digits past the decimal point, we must wait until $n = 17$ and in the fourth column, rounding to three digits, we must wait until $n = 20$. If we round the entries to more and more digits to the right of the decimal point we see that the sequence never actually reaches its limit.

Example 3 *Consider the discrete dynamical system given by the initial value and recursion equation*

$$p_0 = 1000, \quad p_n = 1.3p_{n-1} - 100.$$

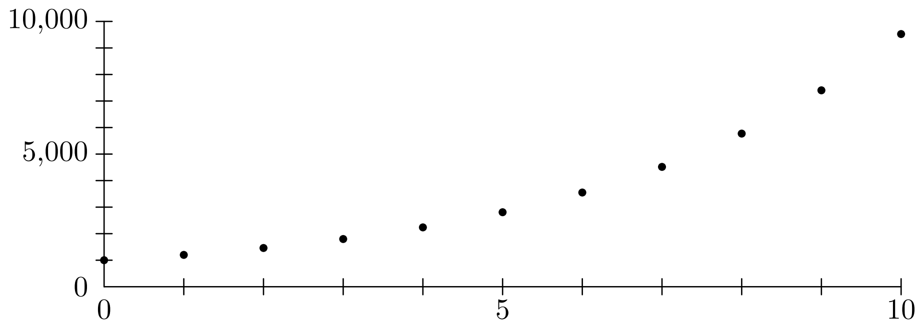


Figure 2.18: Example 2

Figure 2.18 shows the first few terms of the sequence produced by this discrete dynamical system graphically. The terms appear to be getting larger and larger without any bound. In this situation we say the sequence **diverges to** $+\infty$. We make this a formal definition.

Definition 2 *If after a very long time (that is, for large* n*) the terms of a sequence* $p_1, p_2, p_3, \ldots$ *are getting very large and positive without any bound then we say that the sequence is* **diverging to** $+\infty$ *and write*

$$\lim_{n\to\infty} p_n = +\infty.$$

Sometimes this is read "The limit of the sequence $p_1, p_2, p_3, \ldots$ *is* $+\infty$*."*

If after a very long time (that is, for large n*) the terms of a sequence* $p_1, p_2, p_3, \ldots$ *are getting very large and negative without any bound then we say that the sequence is* **diverging to** $-\infty$ *and write*

$$\lim_{n\to\infty} p_n = -\infty.$$

Sometimes this is read "The limit of the sequence $p_1, p_2, p_3, \ldots$ *is* $-\infty$*."*

These ideas are very important for modeling. If, for example, a sequence $p_1, p_2, p_3, \ldots,$ of prices converges to a limit L – that is, if

$$\lim_{n\to\infty} p_n = L$$

then over the long term prices are settling down to the number L. But, if the sequence diverges to $+\infty$ – that is, if

$$\lim_{n\to\infty} p_n = +\infty$$

then we have run-away inflation for this product.

Question 4 *Consider the discrete dynamical system given by the initial value and recursion equation*

$$p_0 = 10,000, \quad p_n = 1.2p_{n-1} - 3000.$$

Does the sequence produced by this discrete dynamical system converge? Does it diverge? If so, express your answer using limit notation.

Question 5 *Consider the discrete dynamical system given by the initial value and recursion equation*

$$p_0 = 10,000, \quad p_n = 1.2p_{n-1} - 1000.$$

Does the sequence produced by this discrete dynamical system converge? Does it diverge? If so, express your answer using limit notation.

Question 6 *Consider the discrete dynamical system given by the initial value and recursion equation*

$$p_0 = 10,000, \quad p_n = 1.2p_{n-1} - 2000.$$

Does the sequence produced by this discrete dynamical system converge? Does it diverge? If so, express your answer using limit notation.

2.5.2 Supply and Demand, Revisited

In the last section we studied supply and demand models and, as one example, looked at the recursion equation

$$p_n = (1 - 15,000k)p_{n-1} + 120,000k$$

that described how prices changed. The parameter k is particularly important. It represents the volatility of the market – how rapidly prices respond to market conditions. In the last section we looked at a number of examples and questions with different values of k and with different initial values. Look back at this work as you answer the following question.

Question 7 *How does the value of the volatility constant, k, affect the long-term behavior of prices in this example?*

Question 8 *How does the initial value affect the long-term behavior of prices in this example?*

Question 9 *A large lake with many fish is fed by water draining from the surrounding area and is drained by a river. The water draining from the surrounding area contains a certain pollutant. As a result, the water in the lake contains the same pollutant. After a heavy rain, runoff from the surrounding area causes the level of this pollutant to rise dramatically. You are part of a team that has been studying this lake for the local board of health. You have determined that during periods of normal rainfall the level of pollution in the lake changes according to the recursion equation*

$$p_n = 0.8p_{n-1} + 15 \text{ ppm}$$

where n is time in weeks. There was a very heavy rain three days ago but now the weather has returned to normal. As a result of the heavy rain and the following runoff, the level of pollution in the lake is now 200 ppm.[16] *Table 2.13 on page 165 shows your predictions based on this recursion equation for pollution levels for the next 60 weeks.*

Notice that apparently

$$\lim_{n\to\infty} p_n = 75 \text{ ppm}.$$

This observation agrees with historical records of the pollution level in the lake. Unfortunately for the fishermen and their customers this pollution level is uncomfortably high. Fish caught in this water are fairly safe to eat but when the pollution level rises much above this level the fish can pose a danger. The preponderance of medical opinion is that people should not eat fish caught in water where the pollution level is above 90 ppm. Thus,

[16] Parts per million.

n	p	n	p	n	p
0	200.0000				
1	175.0000	21	76.4412	41	75.0208
2	155.0000	22	76.1529	42	75.0166
3	139.0000	23	75.9223	43	75.0133
4	126.2000	24	75.7379	44	75.0106
5	115.9600	25	75.5903	45	75.0085
6	107.7680	26	75.4722	46	75.0068
7	101.2144	27	75.3778	47	75.0054
8	95.9715	28	75.3022	48	75.0044
9	91.7772	29	75.2418	49	75.0035
10	88.4218	30	75.1934	50	75.0028
11	85.7374	31	75.1547	51	75.0022
12	83.5899	32	75.1238	52	75.0018
13	81.8719	33	75.0990	53	75.0014
14	80.4976	34	75.0792	54	75.0011
15	79.3980	35	75.0634	55	75.0009
16	78.5184	36	75.0507	56	75.0007
17	77.8147	37	75.0406	57	75.0006
18	77.2518	38	75.0325	58	75.0005
19	76.8014	39	75.0260	59	75.0004
20	76.4412	40	75.0208	60	75.0003

Table 2.13: Predicted Pollution Levels

immediately after a major rainfall like this one, the fisherman must suspend their fishing. When will it be safe for the fisherman to resume fishing?

The fishermen believe that medical opinion is too strict and that the fish would be safe if the level of pollution were below 100 ppm. If we accept their argument when will it be safe for them to resume fishing?

New studies give some evidence that this pollutant is more dangerous than previously thought. As a result, a new regulation has been proposed that prohibits fishing in lakes when the level of pollution is above 80 ppm. If this new regulation is accepted when will it be safe for the fisherman to resume fishing?

Question 10 *Your parents are building up their savings by depositing \$12,000 each year in an investment account that earns interest at the rate of 5.5% each year. Their current balance in this account is \$350,000. During your trip home on Thanksgiving you built the following model that they can use to predict the balance in their retirement account.*

$$p_n = 1.055p_{n-1} + \$12,000, \quad p_0 = \$350,000$$

and told them that

$$\lim_{n\to\infty} p_n = +\infty.$$

Needless to say the idea of having an infinite amount of money to spend in their retirement was rather exciting but you pointed out to them that ∞ *was a long time away.*

Suppose they would like to have a nest egg of \$1,000,000 when they retire. When can they retire?

Suppose they would like to have a nest egg of \$2,000,000 when they retire. When can they retire?

Suppose they would like to have a nest egg of \$3,000,000 when they retire. When can they retire?

In practice, retirement plans must consider the impact of inflation. Suppose your parents believe they could retire comfortably now if they had a nest egg of \$2,000,000 but they believe that inflation will average 2% per year for the foreseeable future. When can they retire?

2.6 Life in a Finite Habitat – Logistic Models

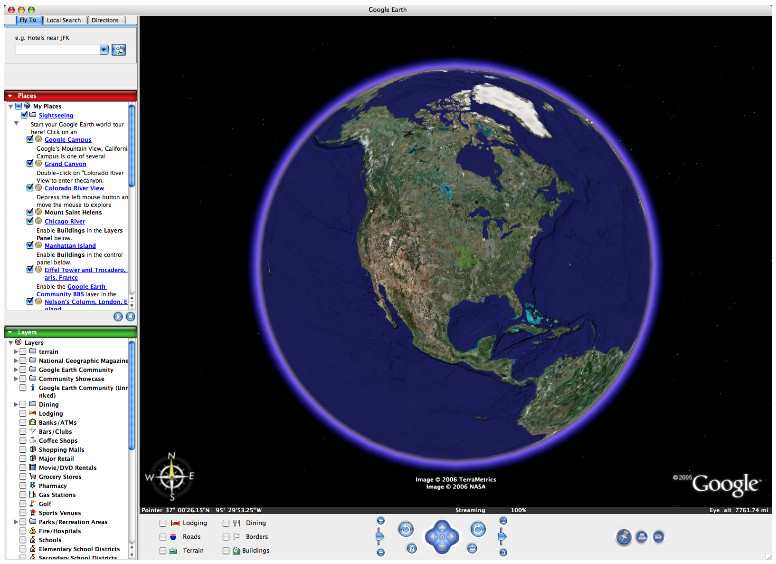


Figure 2.19: A finite habitat

2.6.1 Logistic Models

In this section we are interested in populations that live in a finite habitat, like a lake or the Earth. The simplest models are often expressed informally using words like "the population is increasing at the rate of 3% per year" or "the population is decreasing at the rate of 2% per year." These two models can be expressed by the recursion equations

$$p_n = 1.03p_{n-1} \quad \text{or} \quad p_n = 0.98p_{n-1}.$$

These kinds of population models are of the form

$$p_n = mp_{n-1}$$

where $p_0 > 0$ and $m > 0$. They produce geometric sequences and have closed-form solutions of the form

$$p_n = m^n p_0.$$

These kinds of models are often called **exponential models**. There are three possibilities for the long term behavior of exponential models.

- If $m = 1$ then $p_n = p_{n-1}$, so $p_0 = p_1 = p_2 = \cdots$ and the initial population never changes.
- If $m < 1$ then the population converges to zero – that is,

$$\lim_{n\to\infty} p_n = 0.$$

  In other words the population dies out.

- If $m > 1$ then the population diverges to $+\infty$ – that is,

$$\lim_{n\to\infty} p_n = +\infty.$$

  In other words, the population increases without any bound. Because exponential models with $m > 1$ grow rapidly without any bound, the word "exponential" is sometimes used in informal language as a synonym for rapid and unbounded growth.

The three possibilities for the behavior of these models are shown in Figure 2.20.

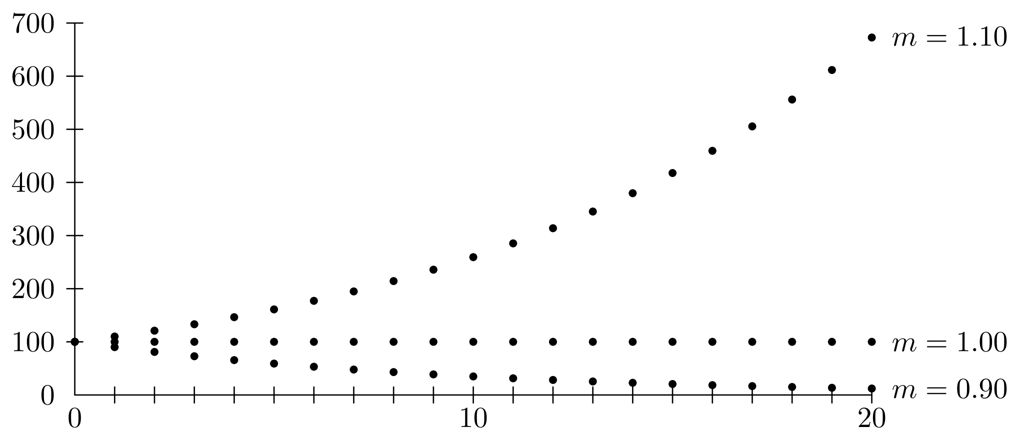


Figure 2.20: Three possibilities four exponential models with $p_0 = 100$

None of these possibilities describe the populations that we see around us every day. Usually, when a new species is introduced into a habitat, it either dies out, in which case we don't see it, or it grows rapidly until it reaches a point where the available resources and the competition prevent it from growing further. Thus, real populations often have two equilibrium points – zero and an equilibrium that is determined by the size of the habitat and its available resources – food, water, shelter, and so forth. Most of the models we have looked at so far are linear models – that is, they can be described by difference equations of the form

$$p_n = mp_{n-1} + b.$$

Figure 2.21 shows the fundamental graph for a linear model. Notice there is only one equilibrium point. Thus, linear models can't model many of the population situations we see in the natural world.

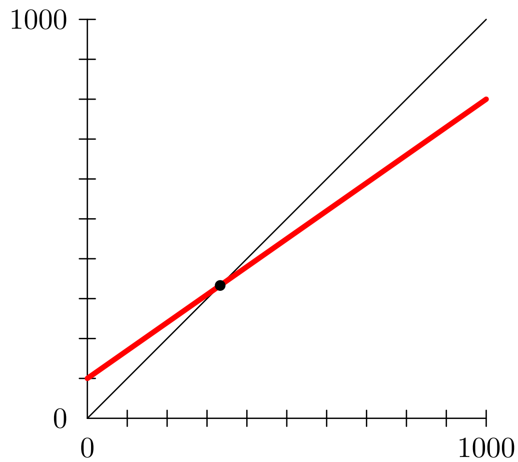


Figure 2.21: The fundamental graph for a linear model

In this section we look at some models that are not linear – that is, they are **nonlinear**. We start with linear models that can be written either using a difference equation of the form

$$p_n - p_{n-1} = Rp_{n-1}$$

or, equivalently, using a recursion equation

$$p_n = mp_{n-1}.$$

These two forms are just different ways of writing the same model. Notice that we can rewrite the difference equation form

$$p_n - p_{n-1} = Rp_{n-1}$$

as

$$p_n = \underbrace{(1+R)}_{m} p_{n-1}.$$

We know that these models are exponential models and have closed-form solutions of the form

$$p_n = (1+R)^n p_0.$$

The constant R in the finite difference equation

$$p_n - p_{n-1} = Rp_{n-1}$$

says that the population rises or falls by the same percentage from one term to the next no matter what is going on. The number R is unaffected by weather, the food supply, the water supply, or any other event. To build a more realistic model, we must replace the constant factor R by a factor that depends on what is occurring in the habitat. This factor might depend on many different things – food, water, the current population, predators, and so forth. In this section we will work with models for habitats that are relatively stable. That is, the food and water supply is relatively constant; the weather is relatively constant; the predator situation is relatively constant; and so forth. The factor that replaces R will depend on only one thing – the population itself. One example is a nonlinear difference equation of the form

$$p_n - p_{n-1} = \underbrace{R\left(1 - \frac{p_{n-1}}{C}\right)}_{\text{replaces the constant } R} p_{n-1}$$

When we use the simple constant factor R, the population always changes by the same percentage but when we use the factor $R\left(1 - \frac{p_{n-1}}{C}\right)$, the percentage by which the population changes depends on the population. When the population is low, it grows by a large percentage but, as the population rises, it grows by smaller percentages as a larger population shares the same food, water, shelter, and other resources.

The models in this family are called **logistic models**. Notice that when p_{n-1} is very small, the quantity P_{n-1}/C is very small and the righthand side

$$R\left(1-\frac{p_{n-1}}{C}\right)p_{n-1}$$

of the difference equation is very close to Rp_{n-1}. So, when the population is very low, a logistic model behaves like an exponential model.

When $p_{n-1}=C$,

$$R\left(1-\frac{p_{n-1}}{C}\right)=R\left(1-\frac{C}{C}\right)=0.$$

So, the population does not change – that is, the rate of increase is zero. The number C is called the **carrying capacity** of the habitat.

When $p_{n-1}<C$,

$$R\left(1-\frac{p_{n-1}}{C}\right)>0.$$

So, the rate at which the population increases is positive and the population rises.

On the other hand, when $p_{n-1}>C$,

$$R\left(1-\frac{p_{n-1}}{C}\right)<0.$$

So, the rate at which the population increases is negative and the population drops.

Example 1 *Consider the logistic model*

$$p_n-p_{n-1}=0.25\left(1-\frac{p_{n-1}}{500}\right)p_{n-1}$$

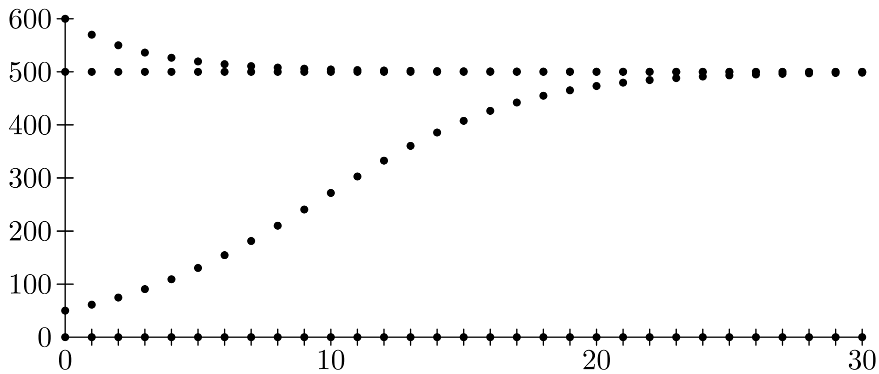


Figure 2.22: Four initial values for a logistic model

The carrying capacity for this model is 500. Figure 2.22 shows what happens with four different initial values.

- $p_0 = 50$.
- $p_0 = 500$.
- $p_0 = 600$.
- $p_0 = 0$.

When the initial value is greater than zero and below the carrying capacity, the population rises and, in this example, approaches the carrying capacity. When the initial value is equal to the carrying capacity, then it stays there. When the initial value is above the carrying capacity, the population drops and, in this example, approaches the carrying capacity. When the initial value is zero, the population stays at zero.

Figure 2.23 on page 174 shows the fundamental graph for this same model. Notice the two equilibrium points. Compare this figure with Figure 2.21 on page 170. Note that there are two equilibrium points in Figure 2.23 but only one in Figure 2.21.

Question 1 *Investigate the logistic model*

$$p_n - p_{n-1} = \left(1 - \frac{p_{n-1}}{2000}\right) p_{n-1}$$

with different initial values. Find the equilibrium values of this model. Be sure to choose initial values that illustrate the full range of possible behaviors. For example, you should

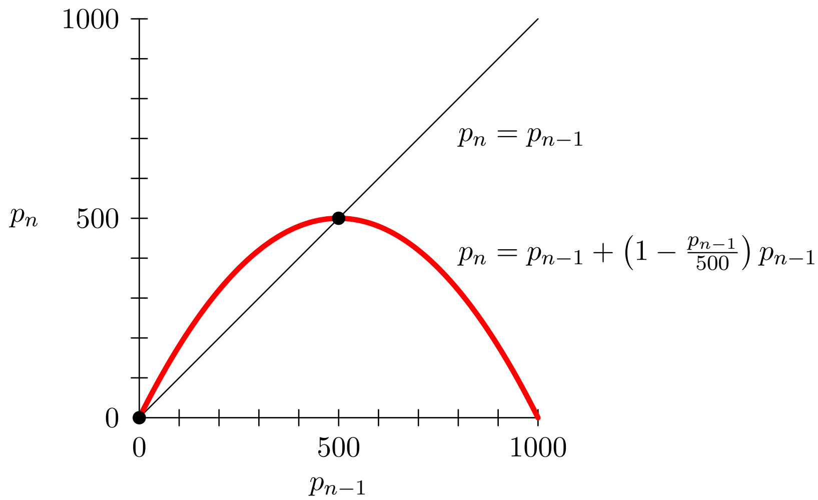


Figure 2.23: The fundamental graph for the recursion relation $p_n = p_{n-1} + \left(1 - \frac{p_{n-1}}{500}\right) p_{n-1}$

try some initial values that are below the carrying capacity and some that are above the carrying capacity. What happens if $p_0 = 0$*?*

Question 2 *Investigate the logistic model*

$$p_n - p_{n-1} = \left(1 - \frac{p_{n-1}}{5000}\right) p_{n-1}$$

with different initial values. Determine the equilibrium values of this model. Be sure to choose initial values that illustrate the full range of possible behaviors. You should try some initial values that are below the carrying capacity and some that are above the carrying capacity.

Question 3 *Investigate the logistic model*

$$p_n - p_{n-1} = \left(2 - \frac{p_{n-1}}{1000}\right) p_{n-1}$$

with different initial values. Determine the equilibrium values of this model. Be sure to choose initial values that illustrate the full range of possible behaviors. You should try some initial values that are below the carrying capacity and some that are above the carrying capacity.

Question 4 *Investigate the logistic model*

$$p_n - p_{n-1} = \left(2.5 - \frac{p_{n-1}}{1000}\right) p_{n-1}$$

with different initial values. Determine the equilibrium values of this model. Be sure to choose initial values that illustrate the full range of possible behaviors. You should try some initial values that are below the carrying capacity and some that are above the carrying capacity.

Question 5 *Start with the logistic model*

$$p_n - p_{n-1} = \left(1 - \frac{p_{n-1}}{2000}\right) p_{n-1}$$

and modify it to include immigration of 200 individuals per year. Investigate the resulting model with different initial values. Describe your results. Be sure to choose initial values that exhibit the full range of possible behavior.

Question 6 *The forest near a town has a colony of small animals whose population can be modeled by the logistic equation*

$$p_n - p_{n-1} = \left(1 - \frac{p_{n-1}}{50,000}\right) p_{n-1}.$$

The population is currently 50,000. The townspeople start killing these small animals for food at the rate of 400 per year. Describe what happens as a result. What is the maximum sustainable rate at which townspeople could kill these animals for food.

2.6.2 A More Complicated Example

In this subsection we look at a more complicated example that illustrates a fairly common behavior. Many species rely on cooperation – for example, some species hunt in packs and may be unable to catch sufficient prey individually. We consider the following recursion equation.

$$p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right).$$

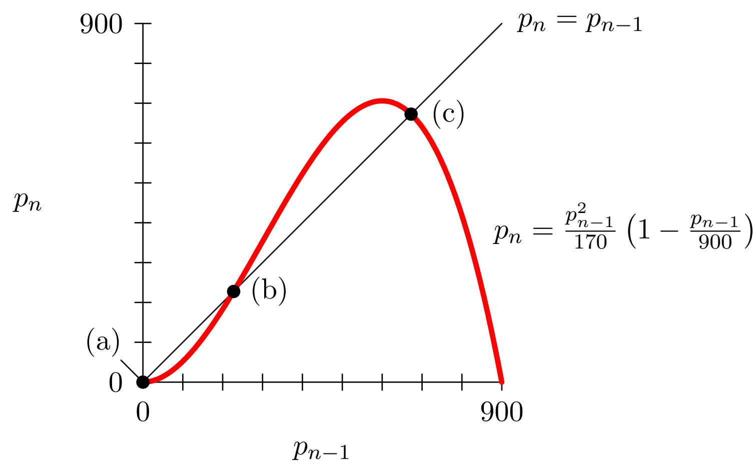


Figure 2.24: The fundamental graph for the recursion relation $p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right)$

The fundamental graph for this recursion relation is shown in Figure 2.24. We can learn a lot from this graph. First, notice that there are three equilibrium points. Roughly speaking, they are at $p = 0$, $p \approx 228$, and $p \approx 672$. Notice the following.

- Between the first two equilibrium points, labeled (a) and (b), the graph of

$$p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right)$$

  is below the graph of $p_n = p_{n-1}$. If the population is in this range it will fall and eventually die out. This is exactly what we would expect for a species that relies on cooperation. If the population is too low then it will fall.

- If the population is between the second and third equilibrium points, labeled (b) and (c), then the graph of

$$p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right)$$

  is above the graph of $p_n = p_{n-1}$. This means that if the population is in this range it will rise. This is similar to the behavior we saw for the logistic model. The third equilibrium point, labeled (c), acts like the carrying capacity.

- If the population is above the third equilibrium point, labeled (c), then the graph of

$$p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right)$$

is below the graph of $p_n = p_{n-1}$. This means that if the population is in this range it will fall. This is what we would expect when the population exceeds the carrying capacity.

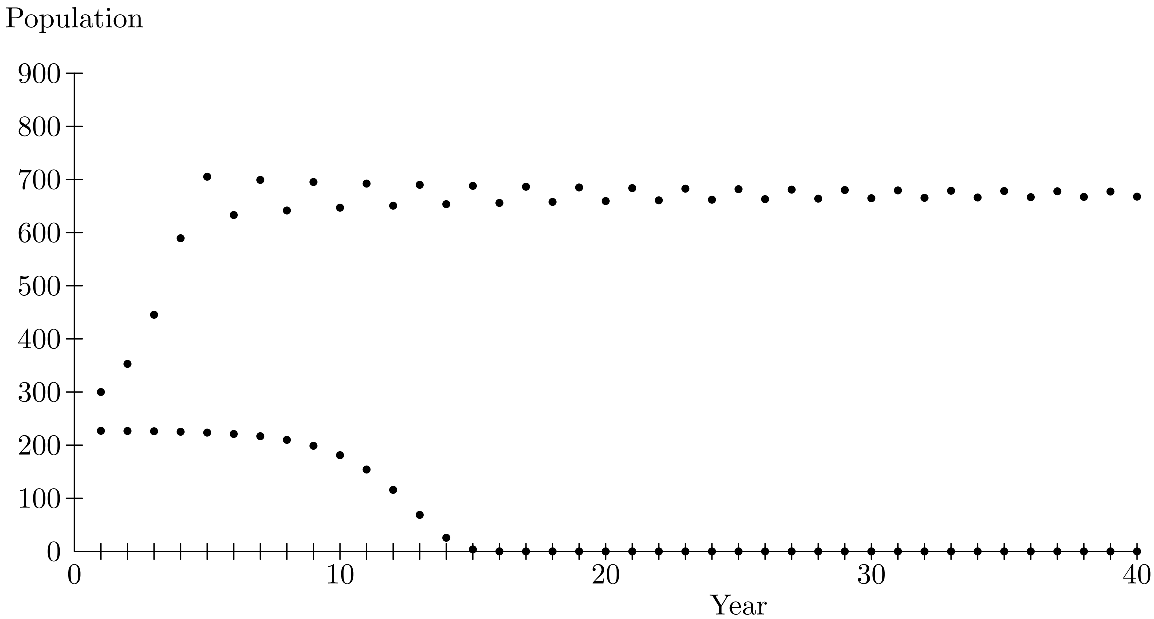


Figure 2.25: Two initial value problems for the recursion relation $p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right)$

Figure 2.25 shows what happens with two different initial values. One initial value, $p_1 = 227$, is just slightly below the second equilibrium point. Notice that with this initial value the population dies out. The second initial value, $p_1 = 300$, is slightly above the second equilibrium point. Notice that with this initial value the population approaches the third equilibrium point, the one that acts like the carrying capacity.

Question 7 *Try other initial value problems using the recursion equation:*

$$p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right)$$

Describe what you see.

2.7 Equilibrium Points

Equilibrium points and the long term behavior of models are intimately related. We have seen many examples of different kinds of long term behavior. In this section we introduce some terminology and we study the relationship between equilibrium points and long term behavior.

2.7.1 Equilibrium Points

Definition 1 *An* **equilibrium point** *for a recursion equation*

$$p_n = f(p_{n-1})$$

is a solution of the equation

$$p = f(p).$$

We often use the notation p_* to denote an equilibrium point and write this equation as

$$p_* = f(p_*).$$

The reason equilibrium points are important is that if p_* is an equilibrium point and if any $p_{n-1} = p_*$ then

$$p_n = f(p_{n-1}) = f(p_*) = p_*.$$

So that if a sequence starts at or ever hits an equilibrium point it stays there.

We first met equilibrium points in section 2.3 and our first example involved the recursion equation

$$p_n = 0.75p_{n-1} + 100.$$

The fundamental graph for this recursion equation is shown in Figure 2.26 on page 179. From this fundamental graph we can see that the equilibrium point is roughly $p_* = 400$. We can also find this equilibrium point algebraically by solving the equation $f(p_*) = p_*$. For this example, since

$$p_n = f(p_{n-1}) = 0.75p_{n-1} + 100,$$

we solve the equation

$$\begin{aligned} p_* &= 0.75p_* + 100 \\ 0.25p_* &= 100 \\ p_* &= \frac{100}{0.25} \\ p_* &= 400 \end{aligned}$$

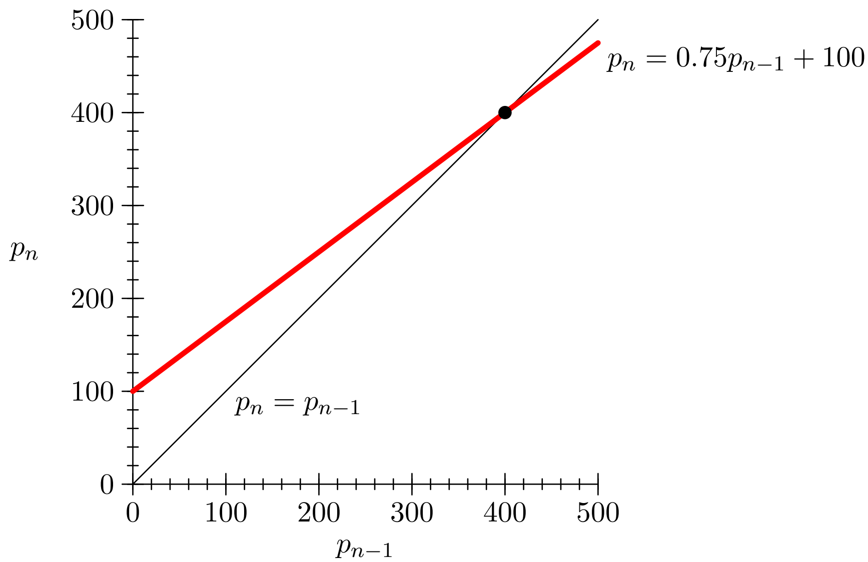


Figure 2.26: Fundamental graph for $p_n = 0.75p_{n-1} + 100$

Figure 2.27 and Table 2.14 on page 180 look at one initial value problem for this recursion equation. An initial value problem is a recursion equation together with an initial value. Notice that for this initial value problem the sequence p_n converges to the equilibrium point. Often, as in this case, the sequence never actually reaches the equilibrium point. It just gets closer and closer. You may want to look back to Table 2.12 on page 161 and to the discussion on the same page.

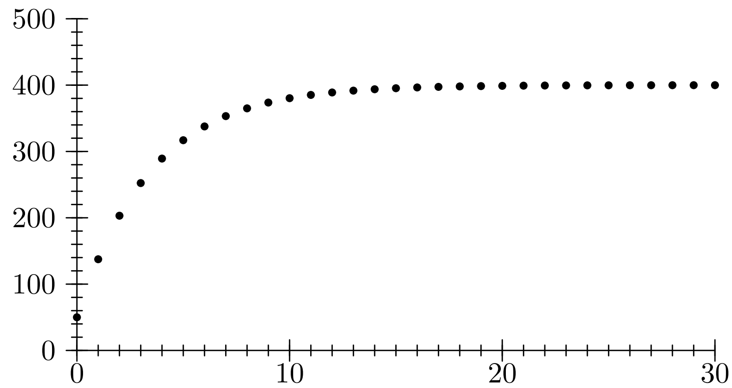


Figure 2.27: Time series graph for $p_n = 0.75p_{n-1} + 100$

n	p_n	n	p_n	n	p_n
0	50.00				
1	137.50	11	385.22	21	399.17
2	203.13	12	388.91	22	399.38
3	252.34	13	391.68	23	399.53
4	289.26	14	393.76	24	399.65
5	316.94	15	395.32	25	399.74
6	337.71	16	396.49	26	399.80
7	353.28	17	397.37	27	399.85
8	364.96	18	398.03	28	399.89
9	373.72	19	398.52	29	399.92
10	380.29	20	398.89	30	399.94

Table 2.14: $p_n = 0.75p_{n-1} + 100, \quad p_0 = 50$

Question 1

- *Find the equilibrium point of the recursion equation* $p_n = 0.65p_{n-1}+200$ *algebraically. Then verify the answer numerically and graphically.*
- *Describe the long term behavior of the initial value problem*

$$p_n = 0.65p_{n-1} + 200, \quad p_0 = 50.$$

- *Describe the long term behavior of the initial value problem*

$$p_n = 0.65p_{n-1} + 200, \quad p_0 = 1000.$$

Question 2

- *Find the equilibrium point of the recursion equation*

$$p_n = 1.2p_{n-1} - 200.$$

- *Describe the long term behavior of the initial value problem*

$$p_n = 1.2p_{n-1} - 200, \quad p_0 = 900.$$

- *Describe the long term behavior of the initial value problem*

$$p_n = 1.2p_{n-1} - 200, \quad p_0 = 1000.$$

- *Describe the long term behavior of the initial value problem*

$$p_n = 1.2p_{n-1} - 200, \quad p_0 = 1100.$$

You probably noticed a big difference between your answers to Question 1 and Question 2. For both initial conditions in Question 1 it turns out that

$$\lim_{n\to\infty} p_n = p_*$$

That is, the sequence converges to the equilibrium point. In fact, for any initial value you try you will discover exactly the same thing. This kind of equilibrium point is called a **globally attracting equilibrium point** because wherever you start you are pulled in to the equilibrium point. You may wonder about the adverb "globally." This equilibrium point pulls you in no matter where you start. Many equilibrium points are **attracting** but might not pull you in if you do not start sufficiently close. You have seen examples of this behavior earlier and will see them again. Sometimes we add the adverb "locally" and speak of a **locally attracting equilibrium point** to emphasize that you might not be pulled in to the equilibrium point if you do not start sufficiently close to it. The adverb "locally" is not necessary. An **attracting equilibrium point** with no qualifying adverb is locally attracting. Sometimes we just add the adverb "locally" for emphasis.

We sometimes use another term, "stable equilibrium," instead of "attracting equilibrium" for these equilibrium points. The two terms "stable equilibrium" and "attracting equilibrium" are synonymous. The adjective "stable" emphasizes a very important quality of these equilibrium points – if you are on a stable equilibrium point and something happens to move you a small distance away then you will be pulled back. When you drive a vehicle, you illustrate this kind of behavior. If you hit a rut or a pothole and are deflected from your path straight ahead, you subconsciously correct and steer back toward your original course.

For Question 2, however, the results are different. For this recursion equation we see that

- If $p_0 > p_*$ then

$$\lim_{n\to\infty} p_n = +\infty$$

- If $p_0 < p_*$ then

$$\lim_{n\to\infty} p_n = -\infty$$

- If $p_0 = p_*$ then

$$\lim_{n\to\infty} p_n = p_*$$

This equilibrium point is an example of a **repelling equilibrium**. These equilibrium points get their name from the fact that unless you start out exactly on the equilibrium point it pushes you away. In this example, the pushing away is extreme – all the way to $+\infty$ or $-\infty$. Many repelling equilibrium points do not push you that far away.

We sometimes use another adjective, "unstable," instead of "repelling" for these equilibrium points. The two terms "unstable" and "repelling" are synonyms in mathematics. The adjective "unstable" emphasizes a very important quality of these equilibrium points – if you are on an unstable equilibrium point and something happens to move you a small distance away then you get pushed further away. Staying on unstable equilibrium points is a delicate balancing act. In fact, you rarely see things in an unstable equilibrium. For example, you can, in theory, balance a baseball bat in the palm of your hand so that it points straight upward but this is an unstable situation and the slightest breath of air will knock the bat away from this equilibrium and it will fall over.

Question 3 *Find the equilibrium points for the model given by the recursion equation*

$$p_n = 0.80p_{n-1} + 100.$$

Investigate what happens with this recursion equation and several different initial values.

Question 4 *Find the equilibrium points for the model given by the difference equation*

$$p_n - p_{n-1} = 0.30p_{n-1} - 200.$$

Describe what happens with the model given by this difference equation and the initial value $p_0 = 100$*. Describe what happens with the model given by this difference equation and the initial value* $p_0 = 1000$*.*

Question 5 *Find the equilibrium points for the model given by the difference equation*

$$p_n - p_{n-1} = -0.30p_{n-1} + 200.$$

Describe what happens with the model given by this difference equation and the initial value $p_0 = 100$*. Describe what happens with the model given by this difference equation and the initial value* $p_0 = 1000$*.*

2.7.2 Equilibrium Points and Linear Discrete Dynamical Systems

Linear recursion equations,

$$p_n = mp_{n-1} + b,$$

have one equilibrium point unless $m = 1$. We can see this from the fundamental graph. Graphically, equilibrium points of a recursion equation,

$$p_n = f(p_{n-1}),$$

are points at which the graph of the function

$$p_n = f(p_{n-1})$$

intersects the graph of the function

$$p_n = p_{n-1}$$

and both the graph of

$$p_n = mp_{n-1} + b \quad \text{and} \quad p_n = p_{n-1}$$

are straight lines. Straight lines always intersect in exactly one point *unless they are parallel.* These two straight lines are parallel if $m = 1$.

We can also see this algebraically. To find the equilibrium point of a recursion equation

$$p_n = mp_{n-1} + b$$

we solve the equation

$$\begin{aligned} p_* &= mp_* + b \\ p_* - mp_* &= b \\ p_*(1-m) &= b \\ p_* &= \frac{b}{1-m} \end{aligned}$$

Thus, once again we see that there is one equilibrium point

$$p_* = \frac{b}{1-m}$$

unless $m = 1$. Notice that if $m = 1$ then the equation $p_* = mp_* + b$ becomes $p_* = p_* + b$, or $0 = b$, so this equation has no solutions unless $b = 0$, in which case every point is an equilibrium point.

2.7.3 Equilibrium points and nonlinear discrete dynamical systems

Our next example is a nonlinear discrete recursion equation – a logistic equation.

Example 1 *Consider the difference equation*

$$p_n - p_{n-1} = \left(1 - \frac{p_{n-1}}{500}\right) p_{n-1}$$

This can also be written as a recursion equation,

$$p_n = p_{n-1} + \left(1 - \frac{p_{n-1}}{500}\right) p_{n-1}$$

or

$$p_n = \left(2 - \frac{p_{n-1}}{500}\right) p_{n-1}.$$

We find the equilibrium points of this recursion equation with a bit of algebra

$$\begin{aligned} p_* &= p_* + \left(1 - \frac{p_*}{500}\right) p_* \\ 0 &= \left(1 - \frac{p_*}{500}\right) p_*, \end{aligned}$$

which leads to the two solutions $p_* = 0$ or $p_* = 500$. This recursion equation has two equilibrium points, zero and 500. We can see the same thing from the fundamental graph, Figure 2.28 on page 186.

Table 2.15 on page 186 and Figure 2.29 on page 187 show the results of using this recursion equation with four different initial values: $p_0 = 0$; $p_0 = 200$; $p_0 = 500$; and $p_0 = 700$. Notice that 500 is a locally attracting equilibrium point.

Question 6 *Find the equilibrium points for the model given by the difference equation*

$$p_n - p_{n-1} = \left(1 - \frac{p_{n-1}}{1000}\right) p_{n-1} - 20.$$

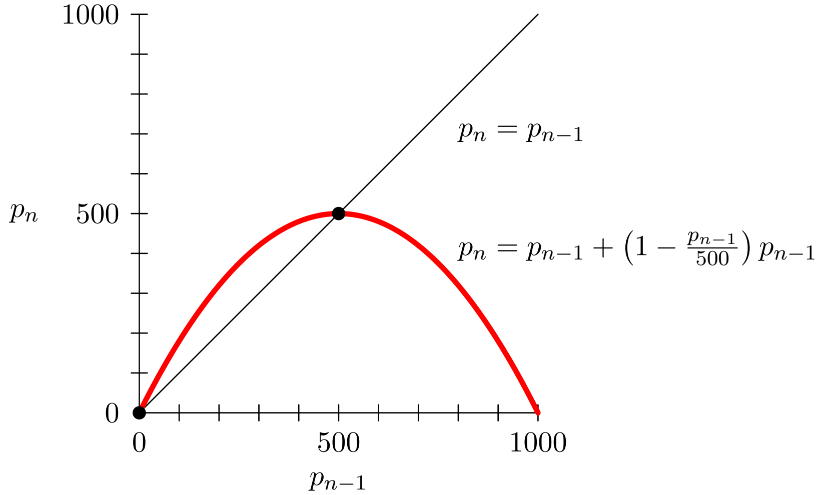


Figure 2.28: The fundamental graph for $p_n = p_{n-1} + \left(1 - \frac{p_{n-1}}{500}\right) p_{n-1}$

n	$p_0 = 0$	$p_0 = 200$	$p_0 = 500$	$p_0 = 700$
0	0.00	200.00	500.00	700.00
1	0.00	320.00	500.00	420.00
2	0.00	435.20	500.00	487.20
3	0.00	491.60	500.00	499.67
4	0.00	499.86	500.00	500.00
5	0.00	500.00	500.00	500.00

Table 2.15: Example 1

Investigate what happens with the model given by this difference equation and several different initial values. Try the initial values: $p_0 = 20$, $p_0 = 50$, $p_0 = 700$, *and* $p_0 = 990$.

Question 7 *Find the equilibrium points for the model given by the difference equation*

$$p_n - p_{n-1} = \left(1 - \frac{p_{n-1}}{1000}\right) p_{n-1} - 200.$$

Investigate what happens with this difference equation and several different initial values. Try the initial values: $p_0 = 200$, $p_0 = 300$, $p_* = 700$, *and* $p_* = 800$.

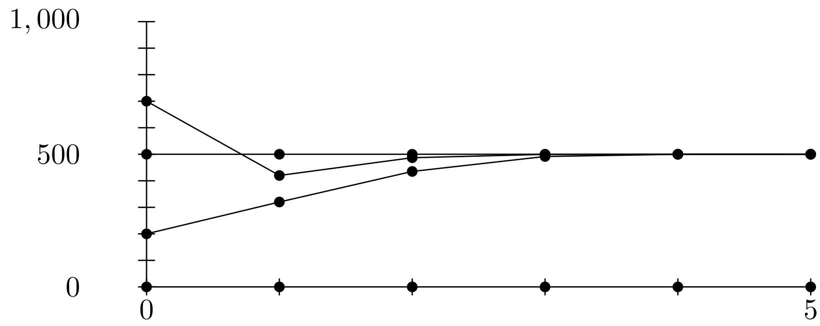


Figure 2.29: Four time series for the recursion relation $p_n = \left(2 - \frac{p_{n-1}}{500}\right) p_{n-1}$

Question 8 *Find the equilibrium points for the model given by the recursion equation*

$$p_n = \frac{6p_{n-1}}{3^{\left(\frac{p_{n-1}}{500}\right)}}.$$

Investigate what happens with this recursion equation and several different initial values.

Example 2 *Find the equilibrium points for the recursion equation*

$$p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right).$$

Classify these equilibrium points – that is, determine whether each one is attracting, repelling, or neither.

Figure 2.30 on page 188 shows the fundamental graph for this recursion equation. From this graph we can see that there are three equilibrium points – at zero, roughly 210, and roughly 620. We can determine these equilibrium points more precisely algebraically by solving the equation

$$p_* = \frac{{p_*}^2}{170}\left(1 - \frac{p_*}{900}\right).$$

Since p_* is a factor of both sides of this equation, $p_* = 0$ is one solution. Now, if p_* is not zero, we can divide both sides of this equation by p_*, leaving us with

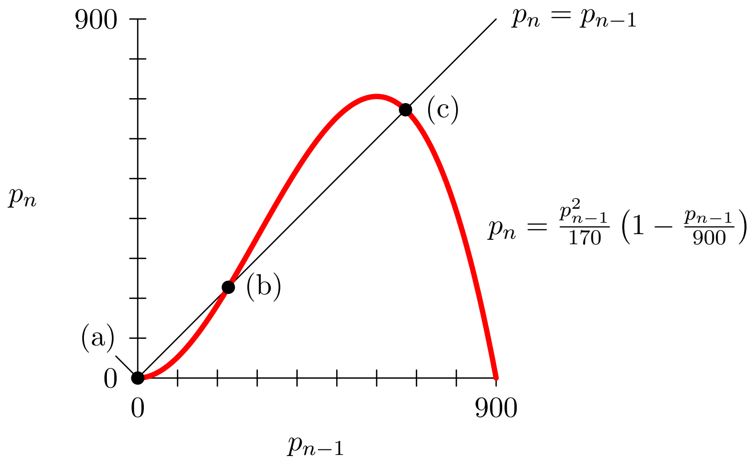


Figure 2.30: The fundamental graph for the recursion relation $p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right)$

$$\frac{p_*}{170}\left(1 - \frac{p_*}{900}\right) = 1,$$

which we solve as follows

$$\begin{aligned}
\frac{p_*}{170}\left(1 - \frac{p_*}{900}\right) &= 1 \\
p_*(900 - p_*) &= (170)(900) \\
900p_* - p_*^2 &= 153,000 \\
p_* - 900p + 153,000 &= 0 \\
p_* &= \frac{900 \pm \sqrt{900^2 - 4(153,000)}}{2} \\
p_* &= \frac{900 \pm \sqrt{810,000 - 612,000}}{2} \\
p_* &= \frac{900 \pm \sqrt{198,000}}{2}
\end{aligned}$$

This yields two more equilibrium points $p_* \approx 672.49$ and $p_* \approx 227.51$. Together with our first equilibrium point $p_* = 0$ we have a total of three equilibrium points $p_* = 0$, $p_* \approx 672.49$ and $p_* \approx 227.51$. A little experimentation (See Figure 2.31 on page 189) provides some evidence that the equilibrium points $p_* = 0$ and $p_* \approx 672.49$ are attracting and the equilibrium point $p_* \approx 227.5$ is repelling.

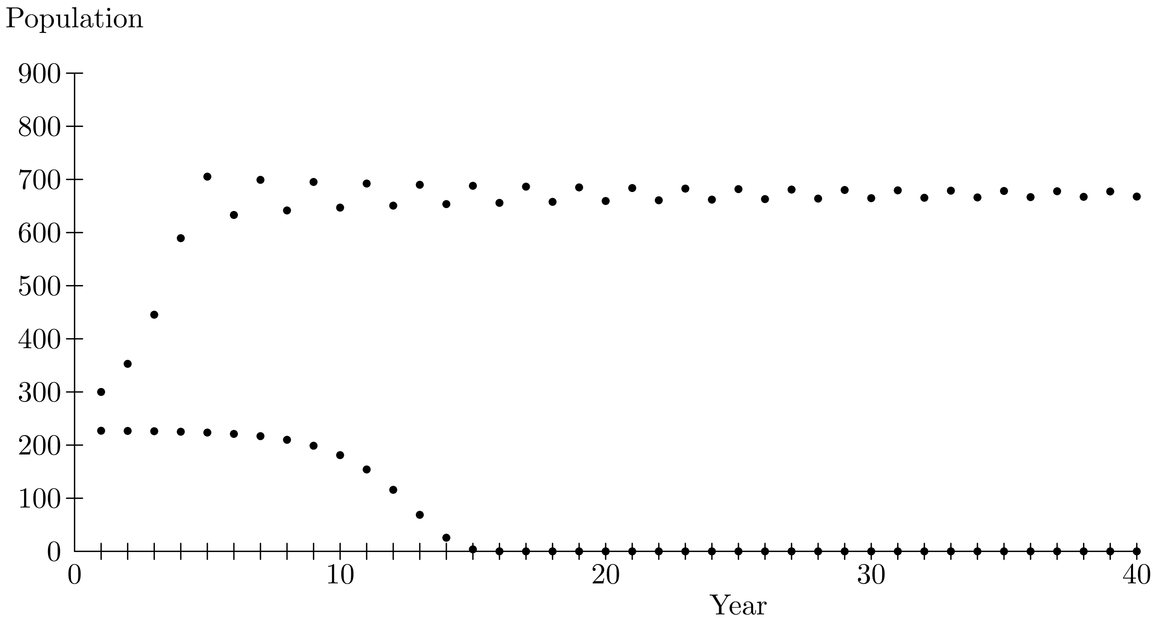


Figure 2.31: Two initial value problems for the recursion relation $p_n = \frac{p_{n-1}^2}{170}\left(1 - \frac{p_{n-1}}{900}\right)$

Question 9 *Find and classify the equilibrium points of the recursion equation*

$$p_n = 2.0\left(1 - \frac{p_{n-1}}{1000}\right)p_{n-1}.$$

Describe the long term behavior of this recursion equation with the initial value $p_0 = 50$.

Question 10 *Find and classify the equilibrium points of the recursion equation*

$$p_n = 2.5\left(1 - \frac{p_{n-1}}{1000}\right)p_{n-1}.$$

Describe the long term behavior of this recursion equation with the initial value $p_0 = 50$.

Question 11 *Find and classify the equilibrium points of the recursion equation*

$$p_n = 3.0\left(1 - \frac{p_{n-1}}{1000}\right)p_{n-1}.$$

Describe the long term behavior of this recursion equation with the initial value $p_0 = 50$.

Question 12 *Find and classify the equilibrium points of the recursion equation*

$$p_n = 3.2\left(1 - \frac{p_{n-1}}{1000}\right)p_{n-1}.$$

Describe the long term behavior of this recursion equation with the initial value $p_0 = 50$.

Question 13 *Find and classify the equilibrium points of the recursion equation*

$$p_n = 3.4\left(1 - \frac{p_{n-1}}{1000}\right)p_{n-1}.$$

Describe the long term behavior of this recursion equation with the initial value $p_0 = 50$.

Question 14 *Find and classify the equilibrium points of the recursion equation*

$$p_n = 3.6\left(1 - \frac{p_{n-1}}{1000}\right)p_{n-1}.$$

Describe the long term behavior of this recursion equation with the initial value $p_0 = 50$.

Question 15 *Find and classify the equilibrium points of the recursion equation*

$$p_n = 3.8\left(1 - \frac{p_{n-1}}{1000}\right)p_{n-1}.$$

Describe the long term behavior of this recursion equation with the initial value $p_0 = 50$.

Question 16 *Find and classify the equilibrium points of the recursion equation*

$$p_n = 4.0\left(1 - \frac{p_{n-1}}{1000}\right)p_{n-1}.$$

Describe the long term behavior of this recursion equation with the initial value $p_0 = 50$.

2.8 The Linear Stability Theorem

One of the most important themes in mathematics is the interplay between experimentation and theory. Understanding always begins with experimentation. Theory is largely organizing, confirming, and giving names to the things we observe during experimentation. In the last few sections we have analyzed many examples of linear recursion equations

$$p_n = mp_{n-1} + b$$

and we have seen that, unless $m = 1$, there is one equilibrium point

$$p_* = \frac{b}{1-m}.$$

Furthermore, this equilibrium always seems to be either globally attracting or globally repelling. Our goal in this section is to prove a theorem, called the Linear Stability Theorem, that confirms and organizes our observations. To see how to prove this theorem we begin with two examples.

Example 1 *A small colony of birds lives on a barren island a few miles off the coast of a lush mainland. The island is so inhospitable that, left on its own, the population of the colony would drop by 20% each year and could be described by the recursion equation*

$$p_n = 0.80p_{n-1}.$$

Each year, however, 500 birds from the mainland get lost and wind up on the barren island. This leads to the recursion equation

$$p_n = 0.80p_{n-1} + 500.$$

Figure 2.32 on page 193 shows this model's predictions for the next 20 years if the initial population of the colony is $p_0 = 5,000$ birds.

Notice that this model has one equilibrium point, $p_* = 2500$, and experimentation suggests that it is globally attracting.

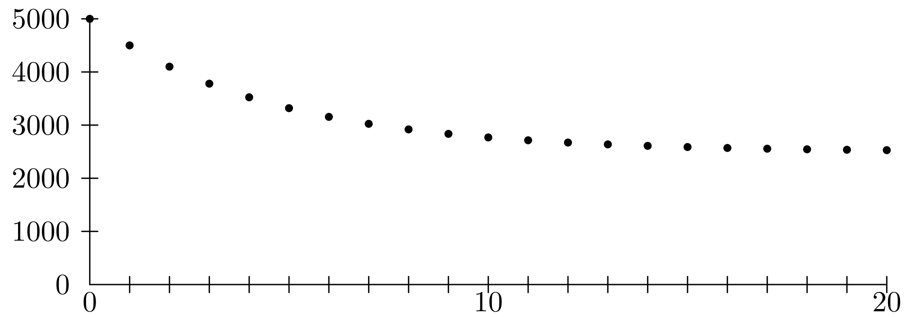


Figure 2.32: A barren island

Example 2 *IASA, the International Aeronautics and Space Administration, plans to establish a colony on Mars. IASA will only be able to send a single spaceship to Mars and will choose colonists who are expected to have lots of children. In fact, if there were no interference, the population of the colony would be expected to grow by 10% every year. This would lead to the recursion equation:*

$$p_n = 1.10p_{n-1}.$$

Unfortunately, Martian leaders have been monitoring the weak signals from earth television programs that have reached Mars. Because the signals are so weak, the general Martian populace has not yet been exposed to these programs and Martian leaders are determined to protect their culture from the influence of these programs. They were able to annihilate an earlier expedition sent from earth. IASA has determined that the Martians can kill 100 colonists every year. Thus, they believe that the colony's population change can be described by the recursion equation:

$$p_n = 1.10p_{n-1} - 100.$$

Notice that this model has one equilibrium point, $p_* = 1000$, and it appears to be repelling. See Figure 2.33 on page 194.

Notice that both of these examples have linear recursion equations. We want to determine when the equilibrium point of a linear recursion equation is attracting and when it is repelling. To do this it is helpful to look not just at the terms of a sequence produced by the recursion equation but also at the difference between each term and the equilibrium value. We are interested in this difference because we are interested in whether the terms

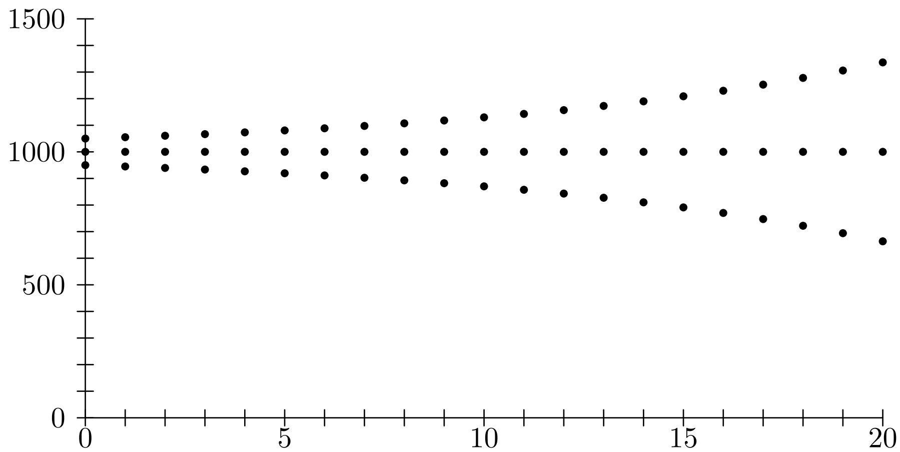


Figure 2.33: Population of Mars over time

are getting close to the equilibrium value and, if so, how fast they are approaching the equilibrium value.

Table 2.16 shows at the first few terms for Example 1. This table has an extra column that gives the difference between each term and the equilibrium point, $p_* = 2500$. Notice that the differences between the terms and the equilibrium point appears to be falling by 20% from term to term.

n	p_n	$p_n - p_*$
0	5000	2500
1	4500	2000
2	4100	1600
3	3780	1280

Table 2.16: The first few terms of $p_n = 0.80p_{n-1} + 500$ and the differences $p_n - p_*$

Question 1 *For each of the following discrete dynamical systems make a table like Table 2.16.*

a. $p_n = 0.50p_{n-1} + 200, \quad p_0 = 600.$

b. $p_n = 0.50p_{n-1} + 200, \quad p_0 = 200.$

c. $p_n = 0.50p_{n-1} + 200, \quad p_0 = 0.$

d. $p_n = 0.75p_{n-1} + 200, \quad p_0 = 1000.$

e. $p_n = 0.75p_{n-1} + 200, \quad p_0 = 500.$

f. $p_n = 0.75p_{n-1} + 200, \quad p_0 = 0.$

This experimentation motivates us to do some algebra that enables us to generalize our observations. We compute the difference between each term p_n and the equilibrium point

$$p_* = \frac{b}{1-m}$$

and compare it to the difference between the preceding term and the equilibrium point as shown below.

$$\begin{aligned}
p_n - p_* &= mp_{n-1} + b - p_* \\
&= mp_{n-1} + b - \left(\frac{b}{1-m}\right) \\
&= mp_{n-1} + m\left(\frac{b}{m} - \frac{b}{m(1-m)}\right) \\
&= m\left(p_{n-1} + \frac{b}{m} - \frac{b}{m(1-m)}\right) \\
&= m\left(p_{n-1} + \frac{b(1-m) - b}{m(1-m)}\right) \\
&= m\left(p_{n-1} + \frac{b - bm - b}{m(1-m)}\right) \\
&= m\left(p_{n-1} - \frac{b}{1-m}\right) \\
&= m\left(p_{n-1} - p_*\right)
\end{aligned}$$

Thus,

$$\begin{aligned}
p_1 - p_* &= m\left(p_0 - p_*\right) \\
p_2 - p_* &= m\left(p_1 - p_*\right) = m\left(m\left(p_0 - p_*\right)\right) = m^2\left(p_0 - p_*\right) \\
p_3 - p_* &= m\left(p_2 - p_*\right) = m\left(m^2\left(p_0 - p_*\right)\right) = m^3\left(p_0 - p_*\right) \\
&\vdots \\
p_n - p_* &= m^n\left(p_0 - p_*\right)
\end{aligned}$$

Now, one last algebraic step gives us an analytic solution, or closed-form, solution

$$\begin{aligned} p_n - p_* &= m^n (p_0 - p_*) \\ p_n &= p_* + m^n (p_0 - p_*) \,. \end{aligned}$$

We can now state and prove two theorems that find and classify the equilibrium point of a linear recursion equation and determine the long term behavior of a linear discrete dynamical system. The first helps us find equilibrium points and the second helps us classify them

Theorem 1 *If $m \neq 1$ then the linear model*

$$p_n = mp_{n-1} + b$$

has exactly one equilibrium point

$$p_* = \frac{b}{1-m}.$$

If $m = 1$ then we are looking at the recursion equation

$$p_n = p_{n-1} + b$$

and

- if $b = 0$ then every point is an equilibrium point since the equation is $p_n = p_{n-1}$
- if $b \neq 0$ then there are no equilibrium points since the equation

  $$p_* = p_* + b$$

  has no solutions.

Theorem 2 (Linear Stability Theorem) *if $m \neq 0$ then the long-term behavior of the linear model*

$$p_n = mp_{n-1} + b$$

is determined as follows.

- *If $|m| < 1$, then for every initial value p_0,*

  $$\lim_{n\to+\infty} p_n = p_*.$$

  Thus, p_*, is **globally attracting.**

- *If $m > 1$ and*

  - *if $p_0 = p_*$, then*
  $$\lim_{n\to+\infty} p_n = p_*.$$
  - *if $p_0 > p_*$, then*
  $$\lim_{n\to+\infty} p_n = +\infty.$$
  - *if $p_0 < p_*$, then*
  $$\lim_{n\to+\infty} p_n = -\infty.$$

  Thus, p_*, is **globally repelling.**

- *If $m < -1$ and*

  - *if $p_0 = p_*$, then*
  $$\lim_{n\to+\infty} p_n = p_*.$$
  - *if $p_0 \neq p_*$, then the sequence $p_0, p_1, p_2, \ldots$ oscillates wildly, bouncing back-and-forth from below the equilibrium point to above the equilibrium point and getting further and further away from the equilibrium point.*

  Thus, p_* is **globally repelling.**

Because this is a long theorem with many parts, in the next few pages we will look at an example and a series of questions relating to this theorem.

Example 3 *The first clause of this theorem states that*

- *If* $|m| < 1$ *then for every initial value* p_0*,*

$$\lim_{n \to +\infty} p_n = p_*.$$

In this situation the equilibrium point, p_**, is* **globally attracting.**

This part of our theorem (and, indeed, every part) follows from the closed form solution that we obtained earlier in this section,

$$p_n = p_* + m^n(p_0 - p_*).$$

For this part, we observe that, if $|m| < 1$, then for large values of n, m^n is very small. Thus, $m^n(p_0 - p_*)$ is also very small and $p_* + m^n(p_0 - p_*)$ is very close to p_*.

Example 1 on page 192 illustrates this clause of the Linear Stability Theorem.

Question 2 *Discuss, with examples, each bullet and subbullet of the Linear Stability Theorem. Your discussion should be similar to the discussion in Example 3.*

Question 3 *Consider the recursion equation from Example 1 on page 155.*

$$p_n = p_{n-1} + 15,000k(8 - p_{n-1})$$

where $p_0, p_1, p_2, \ldots$ *is a sequence of prices that change according to the Law of Supply and Demand.*

- *Find the equilibrium point for this recursion equation.*
- *For what values of the constant* k *is this equilibrium point globally attracting?*
- *For what values of the constant* k *do prices oscillate wildly?*

Does this help us understand the behavior of models based on supply and demand? Your answer should build on the features of the marketplace that are represented by the value of the constant k*.*

Question 4 *Experiment with the nonlinear models*

$$\text{(a)} \quad p_n = p_{n-1} + 1.2p_{n-1}\left(1 - \frac{p_{n-1}}{1,000}\right)$$

$$\text{(b)} \quad p_n = p_{n-1} + 1.8p_{n-1}\left(1 - \frac{p_{n-1}}{1,000}\right)$$

$$\text{(c)} \quad p_n = p_{n-1} + 2.2p_{n-1}\left(1 - \frac{p_{n-1}}{1,000}\right)$$

$$\text{(d)} \quad p_n = p_{n-1} + 2.4p_{n-1}\left(1 - \frac{p_{n-1}}{1,000}\right)$$

$$\text{(e)} \quad p_n = p_{n-1} + 2.6p_{n-1}\left(1 - \frac{p_{n-1}}{1,000}\right)$$

How do you think Theorems 1 and 2 in this section would need to be modified to apply to nonlinear models?

Question 5 *Summarize our work with linear models by filling in the blanks in Table 2.17.*

value of m	p_0	**Long term behavior of p_n**
$m = 0$		
$m < -1$	If $p_0 \neq p_*$	
$-1 < m < 0$	If $p_0 \neq p_*$	
$0 < m < 1$		
$1 < m$	If $p_0 \neq p_*$ If $p_0 < p_*$ If $p_* < p_0$	
$\|m\| < 1$		
$\|m\| > 1$	If $p_0 \neq p_*$	

Table 2.17: Long-term behavior for $p_n = mp_{n-1} + b$

Question 6 *For the models in Question 1 on page 194, use the Linear Stability Theorem to find*

$$\lim_{n\to\infty} p_n.$$

2.9 Closed-Form Solutions

Given a discrete dynamical system we can always compute the terms of the corresponding sequence by starting with the initial value and then applying the recursion equation repeatedly, once for each subsequent term. As we have seen, sometimes there is another way — we can compute the terms directly using a closed-form solution. The terms "closed-form solution," "analytic solution," and "algebraic solution" are all synonymous. Both ways of computing the terms are useful. The first way is more direct because it mirrors the underlying mechanism – how things change. More importantly, it may be quite difficult or even impossible to find a closed-form solution. If we do have a closed-form solution, however, then

- We can find any term with one calculation instead of first calculating all the preceding terms.
- It is often easy to see the long-term behavior from the closed-form solution.

Closed-form solutions depend on both the initial value and the recursion relation. We often use the term **initial value problem**, or **IVP**, to refer to this combination of an initial value and a recursion equation. A closed-form solution must satisfy both the recursion equation and the initial value. In fact, we often find a closed-form solution in two steps.

- First, we find a **general solution** that satisfies the recursion equation. The general solution usually has one parameter, or constant, that can have different values.
- Second, we find a **particular solution** by choosing the value of the parameter or constant that satisfies the initial value.

2.9.1 Verifying general and particular solutions

It is often useful to check or "verify" a solution that we have found. Verifying a general solution to a recursion equation or a particular solution to an initial value problem relies on basically the same idea as verifying a solution to an equation – we substitute the proposed solution into the original equation and check to see if it satisfies the original equation.

Example 1 *Verify that $x = 2$ is a solution to the equation*

$$x^2 - 4x + 4 = 0.$$

Solution:

$$\begin{aligned}
x^2 - 4x + 4 &\quad = \quad 0 \\
(2)^2 - 4(2) + 4 &\quad ?=? \quad 0 \\
4 - 8 + 4 &\quad ?=? \quad 0 \\
0 &\quad = \quad 0 \quad \text{YES!!}
\end{aligned}$$

Substituting a number for a variable in an equation is easy. Substituting a closed-form solution is a bit harder. The following example illustrates the idea.

Example 2

- *Verify that the formula $p_n = C(2^n)$ is a general solution to the recursion equation*

$$p_n = 2p_{n-1}.$$

- *Verify that the formula $p_n = 100(2^n)$ is a particular solution to the initial value problem*

$$p_n = 2p_{n-1}, \quad p_1 = 200.$$

Solution:

To verify the formula $p_n = C(2^n)$ is a general solution to the recursion equation we want to substitute it into the recursion equation

$$p_n = 2p_{n-1}.$$

First, we need to look at the proposed closed form solution and substitute $(n-1)$ for n.

$$\text{Original}: \quad p_n = C(2^n)$$

$$\text{After the substitution}: \quad p_{n-1} = C(2^{(n-1)}).$$

Now we have a formula for p_{n-1} in addition to our original formula for p_n. We substitute these into the recursion equation

$$\begin{gathered} p_n = 2p_{n-1} \\ \underbrace{C(2^n)}_{\text{replaces } p_n} \ ?{=}?2 \ \underbrace{C(2^{(n-1)})}_{\text{replaces } p_{n-1}} \end{gathered}$$

and now do a bit of algebra

$$\begin{aligned} C(2^n) &\quad ?{=}? \quad 2C(2^{(n-1)}) \\ C(2^n) &\quad ?{=}? \quad C2(2^{(n-1)}) \\ C(2^n) &\quad = \quad C(2^n) \quad \text{YES!!!} \end{aligned}$$

To verify that $p_n = 100(2^n)$ is a solution to the IVP

$$p_n = 2p_{n-1}, \quad p_1 = 200$$

we need to do two things.

- Verify that it satisfies the recursion equation.
- Verify that it satisfies the initial value.

Because the proposed particular solution, $p_n = 100(2^n)$, is obtained from the general solution, $p_n = C(2^n)$, by setting the constant $C = 100$, we already know that the first condition is satisfied. Thus, we only need to check the second condition

$$\begin{aligned} p_1 &\quad = \quad 200 \\ 100(2^1) &\quad ?{=}? \quad 200 \\ 200 &\quad = \quad 200 \quad \text{YES!!!} \end{aligned}$$

Sometimes we are given a proposed particular solution directly without being given a general solution. The following example illustrates this.

Example 3 *Verify that*

$$p_n = 2 - \frac{1}{2^n}$$

is a closed form solution for the initial value problem (IVP)

$$p_n = p_{n-1} + \frac{1}{2^n}, \quad p_0 = 1.$$

Since this is a possible particular solution, we must check both the recursion equation and the initial value. We usually check the initial value first because it is easiest.

$$\begin{array}{rcl} p_0 & = & 1 \\ 2 - \dfrac{1}{2^0} & ?=? & 1 \\ 2 - \dfrac{1}{1} & ?=? & 1 \\ 2 - 1 & ?=? & 1 \\ 1 & = & 1 \quad \text{YES!!!} \end{array}$$

Now we turn our attention to the recursion equation. First we substitute $(n-1)$ for n in the original closed-form solution

$$p_n = 2 - \frac{1}{2^n}$$

to obtain

$$p_{n-1} = 2 - \frac{1}{2^{(n-1)}}$$

Next we substitute our formulas for p_n and p_{n-1} into the recursion equation

$$p_n = p_{n-1} + \frac{1}{2^n}$$

to obtain

$$\underbrace{2-\frac{1}{2^n}}_{\text{substituted for } p_n} \quad ?=? \quad \underbrace{2-\frac{1}{2^{(n-1)}}}_{\text{substituted for } p_{n-1}} + \frac{1}{2^n}$$

and now with a bit of algebra

$$\begin{aligned}
2-\frac{1}{2^n} \quad &?=? \quad 2-\frac{1}{2^{(n-1)}}+\frac{1}{2^n} \\
2-\frac{1}{2^n} \quad &?=? \quad 2-\frac{2}{2^n}+\frac{1}{2^n} \\
2-\frac{1}{2^n} \quad &?=? \quad 2-\frac{2-1}{2^n} \\
2-\frac{1}{2^n} \quad &?=? \quad 2-\frac{1}{2^n} \quad \text{YES!!!}
\end{aligned}$$

In many cases we have a general closed-form solution for a recursion equation and want to find a particular closed-form solution given an initial condition. The following example illustrates this.

Example 4 *A particular lake is fed by runoff from the surrounding agricultural area and drained by a river. You are an analyst for the department of public works and have been tracking the level of a particular pollutant in the lake. The level always goes up after a particularly heavy rainfall and then returns to normal as the weather returns to normal. The recursion equation that governs this is*

$$p_n = 0.80p_{n-1} + 0.05$$

with n measured in days and p_n in parts per million. You have computed the equilibrium point for this recursion equation. It is $p_* = 0.25$. You have also found a general solution for this recursion equation. It is

$$p_n = 0.25 + C(0.8^n).$$

Two days ago there was a tremendous rain. You just measured the level of pollution in the lake this morning and found it to be 1.85 *parts per million. Find a particular solution for this IVP. Use your particular solution to determine when the level of pollution will drop below 0.50 parts per million.*

Solution:

We let $n = 0$ be today. We substitute the initial value $p_0 = 1.85$ into the closed form solution

$$p_n = 0.25 + C(0.8^n)$$

to obtain

$$1.85 = 0.25 + C(0.8^0)$$

and with a bit of algebra

$$\begin{aligned} 1.85 &= 0.25 + C(0.8^0) \\ 1.60 &= C(0.8^0)) \\ 1.60 &= C(1) \\ 1.60 &= C \end{aligned}$$

Thus, the particular solution we seek is

$$p_n = 0.25 + 1.60(0.8^n).$$

To find out when the pollution level will reach 0.50 parts per million we solve the equation

$$\begin{aligned} 0.25 + 1.60(0.8^n) &= 0.50 \\ 1.60(0.8^n) &= 0.25 \\ 0.8^n &= \frac{0.25}{1.60} \end{aligned}$$

$$
\begin{aligned}
n\log(0.8) &= \log\left(\frac{0.25}{1.60}\right) \\
n &= \frac{\log\left(\frac{0.25}{1.60}\right)}{\log(0.8)} \\
n &\approx 8.32
\end{aligned}
$$

So after 9 days the level of pollution will be below 50 parts per million. We can check this by computing

$$p_9 = 0.25 + 1.60(0.8^9) \approx 0.4647$$

and

$$p_8 = 0.25 + 1.60(0.8^8) \approx 0.5184.$$

Question 1 *Check if*

$$p_n = \frac{n(n+1)}{2}.$$

is a closed-form solution of the initial value problem:

$$p_1 = 1, \quad p_n = p_{n-1} + n.$$

Question 2 *Check if*

$$p_n = C + 0.90^n$$

is a general closed-form solution of the recursion equation

$$p_n = 0.90 * p_{n-1} + 100.$$

If so, find the particular solution of the IVP

$$p_n = 0.90 * p_{n-1} + 100, \quad p_1 = 600.$$

Question 3 *Check if*

$$p_n = 200 + C(0.90^n)$$

is a general closed-form solution of the recursion equation

$$p_n = 0.90 * p_{n-1} + 100.$$

If so, find the particular solution of the IVP

$$p_n = 0.90 * p_{n-1} + 100, \quad p_1 = 600.$$

Question 4 *Check if*

$$p_n = 1000 + C(0.90^n)$$

is a general closed-form solution of the recursion equation

$$p_n = 0.90 * p_{n-1} + 100.$$

If so, find the particular solution of the IVP

$$p_n = 0.90 * p_{n-1} + 100, \quad p_1 = 600.$$

Example 5 *When the economy slows people often talk about an economic stimulus package to help the economy recover. In early 2008 politicians agreed rapidly on a politically easy solution – temporary tax cuts. But, there are other options. In this example we examine one of those options. The nation's infrastructure is in bad shape and one possibility would have been to spend the money on repairing the infrastructure instead of temporary tax cuts. This would have created jobs directly. In addition, it would have created some jobs* *indirectly because people newly employed repairing our infrastructure would buy more and would, thus, create new jobs. These new jobs would, in turn, create more new jobs, and so forth. Suppose that each job that is created creates in turn R new jobs. Normally, $R < 1$. Suppose that we create A jobs directly by an economic stimulus package focused on repairing our infrastructure. If we only consider the jobs created directly we have a first estimate of the total number of jobs created*

$$p_1 = A \text{ jobs.}$$

If we add the jobs created indirectly by the jobs created directly we get a second estimate

$$p_2 = \underbrace{p_1}_{\text{original jobs}} + \underbrace{AR}_{\text{first indirect jobs}}.$$

But now the first indirect jobs create more indirect jobs giving us a third estimate

$$p_3 = p_2 + \underbrace{AR^2}_{\text{second indirect jobs}}.$$

This leads us to the recursion equation

$$p_n = p_{n-1} + AR^{(n-1)}.$$

Analyze the impact of creating A new jobs repairing the nation's infrastructure.

We will begin by verifying a proposed closed-form solution of the IVP

$$p_n = p_{n-1} + AR^{(n-1)}, \quad p_1 = A.$$

In this section we will not discuss how this solution was found. The proposed solution is

$$p_n = A\left(\frac{1 - R^n}{1 - R}\right).$$

First, we check the initial condition

$$\begin{aligned}
A &\ ?\!=\!?\ p_1 \\
A &\ ?\!=\!?\ A\left(\frac{1-R^1}{1-R}\right) \\
A &\ ?\!=\!?\ A(1) \\
A &= A \quad \text{YES!!!}
\end{aligned}$$

Next we check the recursion equation. First we substitute $(n-1)$ for n in the closed-form solution

$$p_n = A\left(\frac{1-R^n}{1-R}\right).$$

to get

$$p_{n-1} = A\left(\frac{1-R^{(n-1)}}{1-R}\right).$$

and now substituting in the recursion equation and doing some algebra, we see

$$\begin{aligned}
p_n &= p_{n-1} + AR^{(n-1)} \\
A\left(\frac{1-R^n}{1-R}\right) &\ ?\!=\!?\ A\left(\frac{1-R^{(n-1)}}{1-R}\right) + AR^{(n-1)} \\
\left(\frac{1-R^n}{1-R}\right) &\ ?\!=\!?\ \left(\frac{1-R^{(n-1)}}{1-R}\right) + R^{(n-1)} \\
\left(\frac{1-R^n}{1-R}\right) &\ ?\!=\!?\ \left(\frac{1-R^{(n-1)}}{1-R}\right) + \left(\frac{R^{(n-1)}(1-R)}{1-R}\right) \\
\left(\frac{1-R^n}{1-R}\right) &\ ?\!=\!?\ \left(\frac{1-R^{(n-1)}}{1-R}\right) + \left(\frac{R^{(n-1)}-R^n}{1-R}\right) \\
\left(\frac{1-R^n}{1-R}\right) &\ ?\!=\!?\ \left(\frac{1-R^{(n-1)}+R^{(n-1)}-R^n}{1-R}\right)
\end{aligned}$$

$$\left(\frac{1-R^n}{1-R}\right) \quad = \quad \left(\frac{1-R^n}{1-R}\right) \quad \text{YES!!!}$$

So our proposed closed-form solution is correct. Job creation is not instantaneous. Suppose that it takes one month for each newly created job to create R new jobs. Then if we create A new jobs now after one year the total number of direct and indirect new jobs will be

$$p_{13} = A\left(\frac{1-R^{13}}{1-R}\right)$$

Suppose that we assume that $R = 0.75$ – that is, that each new job creates indirectly 0.75 additional new jobs. Then

$$p_{13} = A\left(\frac{1-(0.75)^{13}}{1-0.75}\right) \approx 3.90A$$

The factor 3.90 is called the job multiplier.

If we considered all the new jobs created over the long term by this investment in our infrastructure we would get

$$\lim_{n\to\infty} p_n = 4.0A.$$

Notice that the vast majority of the new jobs that are created, are created in the first year.

Question 5 *Explain why in the example above that*

$$\lim_{n\to\infty} p_n = 4.0A.$$

Question 6 *We made a lot of assumptions in the example above. These include*

- $R = 0.75$.
- *It takes one month for each job to generate R new jobs.*

We did not make any assumptions about how many jobs would be created by an economic stimulus package.

The economic stimulus package passed in early 2008 was $168,000,000,000. Make your own assumption about how many jobs this money would have created if it were invested in repairing our infrastructure. Make your own assumptions about the value of R and about how long it takes each new job to create R additional new jobs. Based on your assumptions, how many jobs would have been created in the first year? How many jobs would have eventually been created? Explain how you arrived at your assumptions.

2.9.2 Guess-and-check

We can often find a closed-form solution for an initial value problem by the method of **guess-and-check.** To see this, consider the following example.

Example 6 *Find a closed-form solution for the initial value problem:*

$$p_1 = 1, \quad p_n = p_{n-1} + (2n-1).$$

Table 2.18 shows the first few terms of this sequence.

n	p_n
1	1
2	4
3	9
4	16
5	25
6	36
7	49
8	64
9	81
10	100

Table 2.18: The first few terms of the sequence $p_1 = 1,\ p_n = p_{n-1} + (2n-1)$

Looking at Table 2.18 we can make a guess, or conjecture that: $p_n = n^2$. We can check whether this guess is correct as follows.

- First, we need to check that it gives the correct initial value. The conjectured closed-form solution $p_n = n^2$ predicts that $p_1 = 1^2 = 1$ and thus, agrees with the actual initial value. Our conjectured formula passes this first test.
- Next, we need to check that our guess meets the requirements of the recursion equation. According to our guess,
  - $p_n = n^2$
  - $p_{n-1} = (n-1)^2$

  and substituting these into the recursion equation,

  $$p_n = p_{n-1} + (2n-1),$$

  we check

$$\begin{aligned} \underbrace{n^2}_{\text{replaces } p_n} \quad &?=? \quad \underbrace{(n-1)^2}_{\text{replaces } p_{n-1}} + (2n-1)? \\ n^2 \quad &?=? \quad n^2 - 2n + 1 + 2n - 1? \\ n^2 \quad &?=? \quad n^2 \quad \text{YES!!} \end{aligned}$$

Thus, our guess passes both tests and is a correct closed-form solution of the initial value problem. These two steps – guess and check – are a powerful method for determining closed-form solutions for initial value problems. This method is called the **guess-and-check method**.

Question 7 *Use the guess-and-check method to find a closed-form solution for the initial value problem*

$$p_0 = 1, \quad p_n = p_{n-1} + (2n+1).$$

Question 8 *Use the guess-and-check method to find a closed-form solution for the initial value problem*

$$p_0 = 0, \quad p_n = p_{n-1} + (2n+1).$$

Question 9 *Use the guess-and-check method to find a closed-form solution for the initial value problem*

$$p_0 = 0, \quad p_n = p_{n-1} + \frac{1}{2^n}.$$

What is the long-term behavior of this sequence?

Question 10 *Check if*

$$p_n = \frac{1 - r^{n+1}}{1 - r}$$

is a closed-form solution of the initial value problem

$$p_0 = 1, \quad p_n = p_{n-1} + r^n.$$

Question 11 *Consider the recursion equation*

$$p_n = p_{n-1} + 37.$$

1. *Guess a general solution of this recursion equation by computing a few terms and looking for a pattern.*
2. *Verify that your guess is a general solution.*
3. *Find a particular solution to the IVP*

$$p_n = p_{n-1} + 37, \quad p_1 = 150.$$

4. *Find a particular solution to the IVP*

$$p_n = p_{n-1} + 37, \quad p_0 = 150.$$

Question 12 *Show that*

$$p_n = CR^n + \frac{b}{1 - R}$$

is a general solution of the recursion equation

$$p_n = Rp_{n-1} + b.$$

Then, find the particular solution of the IVP

$$p_n = Rp_{n-1} + b \quad p_1 = 2000.$$

Chapter 3

Vectors, Matrices, and Systems of Linear Equations

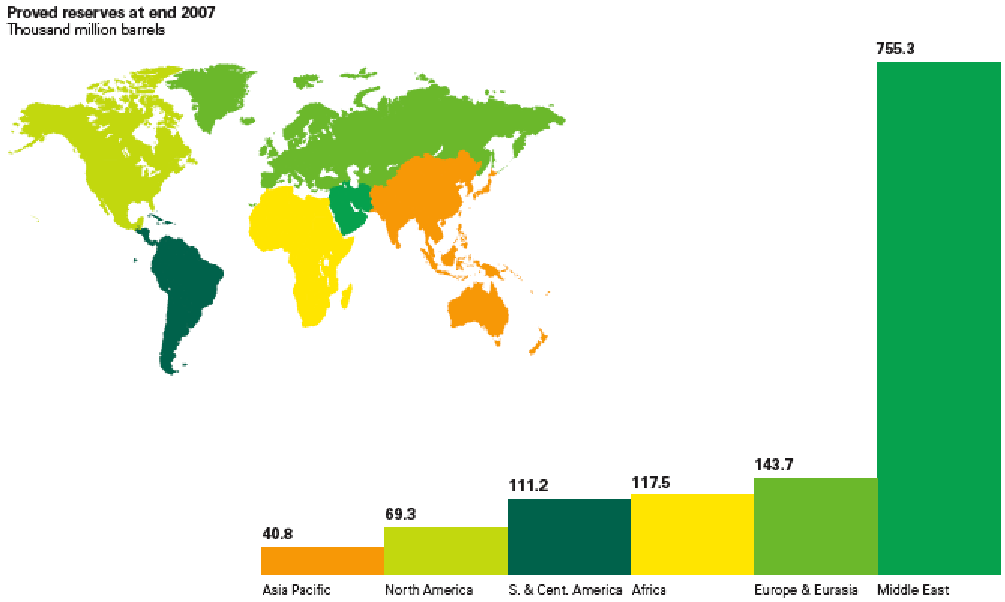


Figure 3.1: Known oil reserves (end of 2007). BP Statistical Review of World Energy, 2008

In the last chapter we focused on modeling situations in which we kept track of one quantity of interest but there are many situations in which we need to keep track of many quantities. For example, Figure 3.1 shows proven oil reserves at the end of the year 2007 broken down by region. It is not enough to keep track of the total oil reserves for the whole world. We need to know where the oil reserves are located.

This chapter is about vectors and matrices – two tools that are enormously powerful when we need to keep track of many quantities. In this chapter we also study systems of linear equations. These systems can often be used to describe or at least approximate the relationships among quantities. The mathematics we develop in this chapter for working with many changing quantities and linear relationships is so important and so powerful for modeling and simulation that modern computers have powerful GPUs, or Graphics Processing Units, that can perform trillions of operations per second based on this mathematics.

Region	**Oil Reserves** (Thousand million barrels)
Asia Pacific	40.8
North America	69.3
South and Central America	111.2
Africa	117.5
Europe and Eurasia	143.7
Middle East	755.3

Table 3.1: Known oil reserves (end of 2007).

3.1 Introduction to Vectors

Figure 3.1 on page 217 shows the world's known oil reserves at the end of the year 2007. Because the location of the oil is important, we are interested in showing oil reserves by region, as in Figure 3.1 and in Table 3.1, rather than just a single number giving the world's total known oil reserves.

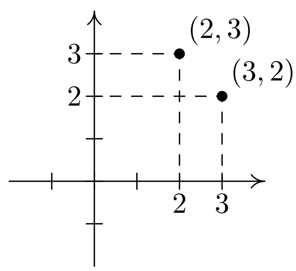


Figure 3.2: The points $(3, 2)$ and $(3, 2)$

There are many other situations in which we need lists of numbers rather than a single number. For example, a point on a plane is described by a list of two numbers giving its x- and y-coordinates as shown in Figure 3.2. Notice that the order of the elements of a list is important. The points $(2, 3)$ and $(3, 2)$ are different points. Similarly, a point in three-dimensional space can be described by a list of three numbers giving its x- y- and z-coordinates.

Lists of numbers – like the lists of two numbers used to describe a point in the plane, the lists of three numbers used to describe points in three dimensional space, or the list of six numbers used above to describe the world's known oil reserves at the end of 2007 – are called **vectors** and are written as described in the following definition.

Definition 1 *An n-***dimensional vector** *is an ordered list of n numbers. We use a symbol like $\vec{v}$ with an arrow above a letter to denote a vector and we use the same letter with subscripts to denote the entries in the vector. We use the symbols "$\langle$" and "$\rangle$" to enclose the entries in the vector. For example, if $\vec{v} = \langle 3, -2, 4\rangle$ is a three-dimensional vector then $v_1 = 3, v_2 = -2$, and $v_3 = 4$. The order of the entries in a vector is important.*

You may have seen two- and three-dimensional vectors before used to represent points in a plane or in three-dimensional space but there are many situations in which much longer lists – that is, much higher dimensional vectors – are used. For example, the data in Table 3.1 on page 218 can be written as the six-dimensional vector,

$$\langle 40.8, 69.3, 111.2, 117.5, 143.7, 755.3\rangle.$$

Modern sound recordings are made using vectors with 44,100 entries[1] for each second of sound, and digital photographs often use vectors that have 3,000,000 or more entries. Black-and-white photographs use one entry for each pixel and color photographs use three entries for each pixel, one each for the red, blue, and green components of the color. The tools we develop in this chapter are used not only to record and process real sound and images, they are also used in movies like Star Wars and in electronic games to create completely artificial worlds. One of the most important uses of these tools in the military is to create simulations that allow us to rehearse dangerous operations. Both electronic games and virtual simulations rely on trillions of vector and matrix operations per second. Because these operations are so important and because we need so many of them, modern graphics processing units are designed for extremely high-speed vector and matrix operations.

Example 1 *In order to combat an insurgency in a particular country, it is important to keep track of the number of insurgents in different parts of the country. Analysts have divided the country into six sectors. Each week, analysts record the situation as a list of six numbers – each number giving the estimated number of insurgents in one of the six sectors. Table 3.2 on page 220 is an example of such a list for one particular week.*

We can represent this information by the vector

$$\vec{v} = \langle 160, 182, 231, 119, 158, 318\rangle$$

[1] Sound is caused by rapid variations in air pressure and each entry in the vectors that record sound represents the difference between the air pressure at a particular time and the average air pressure.

sector	**number**
1	160
2	182
3	231
4	119
5	158
6	318

Table 3.2: Insurgents in each sector

Example 2 *Three insurgent groups (let's call them a, b, and c) have members in the six sectors given by the following three vectors.*

$$\begin{aligned}\vec{a} &= \langle 250, 140, 175, 86, 90, 315\rangle \\ \vec{b} &= \langle 115, 230, 214, 100, 65, 75\rangle \\ \vec{c} &= \langle 77, 60, 329, 113, 413, 115\rangle\end{aligned}$$

We can compute the total number of insurgents in each sector as follows.

- *Sector 1:* $250 + 115 + 77 = 442$.
- *Sector 2:* $140 + 230 + 60 = 430$.
- *Sector 3:* $175 + 214 + 329 = 718$.
- *Sector 4:* $86 + 100 + 113 = 299$.
- *Sector 5:* $90 + 65 + 413 = 568$.
- *Sector 6:* $315 + 75 + 115 = 505$.

We can express the total number of insurgents in each sector as a vector.

$$\vec{t} = \langle 442, 430, 718, 299, 568, 505\rangle$$

Notice that each entry in the vector $\vec{t}$ is obtained by adding the corresponding entries of each of the three vectors, $\vec{a}$, $\vec{b}$, and $\vec{c}$.

Because we often need to add the corresponding entries of two or more vectors in this way, we introduce an operation called **vector addition**, or the **vector sum**, as described in the definition below.

Definition 2 *Suppose that* $\vec{u} = \langle u_1, u_2, \dots u_n \rangle$ *and* $\vec{v} = \langle v_1, v_2, \dots v_n \rangle$ *are two vectors of the same dimension. Then, we define their* **vector sum**, $\vec{u} + \vec{v}$, *by*

$$\vec{u} + \vec{v} = \langle u_1 + v_1, u_2 + v_2, \dots u_n + v_n \rangle$$

For example,

$$\langle 1, 2, 3 \rangle + \langle 2, -4, 6 \rangle = \langle 3, -2, 9 \rangle$$
$$\langle 1, 2, 3, 4 \rangle + \langle 0, 0, 0, 0 \rangle = \langle 1, 2, 3, 4 \rangle$$
$$\langle 1, 2 \rangle + \langle 3, 4, 5 \rangle = \textbf{not defined!}$$

The sum $< 1, 2 > + < 3, 4, 5 >$ *is not defined because the two vectors do not have the same dimension.*

Similarly, we define **vector subtraction** *by*

$$\vec{u} - \vec{v} = \langle u_1 - v_1, u_2 - v_2, \dots u_n - v_n \rangle.$$

Question 1 *New intel reports that the numbers of insurgents given by the three vectors,* $\vec{a}, \vec{b}$, *and* $\vec{c}$, *in Example 2 seriously underestimates the true strength of each group. This intel reports that the figures for the group represented by* $\vec{a}$ *should be multiplied by 1.5 and the figures for each of the other two groups should be doubled. Find three vectors that describe the revised estimated number of insurgents in each sector for each of the three groups and then use these vectors to find a vector describing the total number of insurgents from all three groups in each sector.*

Your first step in answering Question 1 above was probably multiplying the entries of the vector $\vec{a}$ by 1.5 as shown below.

$$1.5\langle 250, 140, 175, 86, 90, 315 \rangle = \langle 375, 210, 262.5, 129, 135, 472.5 \rangle$$

Because we often need to multiply all the entries in a vector by the same number, we introduce an operation called **scalar multiplication** as described in the definition below.

Definition 3 *Suppose that $\vec{u} = \langle u_1, u_2, \ldots u_n \rangle$ is an n-dimensional vector and that c is a real number, or scalar. Then, we define the* **scalar multiple***, $c\vec{u}$, of the vector $\vec{u}$ multiplied by the* **scalar** *c to be*

$$c\vec{u} = \langle cu_1, cu_2, \ldots cu_n \rangle$$

For example,

$$3\langle 1, 2, 3 \rangle = \langle 3, 6, 9 \rangle.$$

Question 2 *According to the BP Statistical Review of World Energy, 2008, the total known reserves of oil at the end of 2007 were 1,237.8 billion[2] barrels, or GbO. Table 3.1 on page 218 breaks this total down by area of the World. Represent this information as a vector.*

According to the same source, oil production by area in 2007 was given by Table 3.3. These figures are in thousands of barrels per day.

Region	**Oil Production (Thousands of Barrels per Day)**
Asia Pacific	7,907
North America	13,665
South and Central America	6,633
Africa	10,318
Europe and Eurasia	17,835
Middle East	25,176

Table 3.3: 2007 Oil production.

If no new oil reserves are discovered and oil production continues at the present rate for the next five years, what will the known reserves for each sector be at the end of 2008? at the end of 2009? at the end of 2010? Use the vector operations described in Definitions 2 and 3 to answer these questions.

We have already mentioned one of the most common uses of two-dimensional vectors – to represent points in the plane. For example, Figure 3.3 on page 223 shows the point represented by the vector $\vec{v} = \langle 3, 2 \rangle$. This is the point whose x-coordinate is 3 and whose y-coordinate is 2. Figure 3.3 shows both this point and its vector representation. Notice

[2]In the United States and this book "billion" means a thousand million or 10^9. In some parts of the world "billion" means a million million or 10^{12}.

that vectors and points are different things. A vector may represent many different things, one of which is a point. For this reason we use the notation (u_1, u_2) to denote a point and the notation $\langle u_1, u_2 \rangle$ to denote a vector.

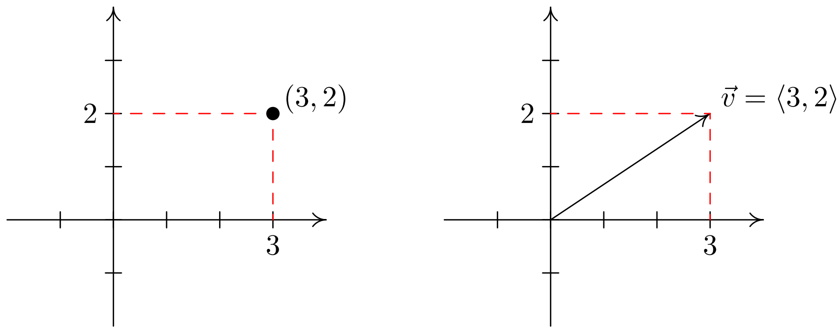


Figure 3.3: A point $(3, 2)$ and its vector representation $\vec{v} = \langle 3, 2 \rangle$

Another very common use of vectors is to represent **motion**, or **displacement**, from one point to another. For example, the motion or displacement required to reach the point $(3, 1)$ from the point $(0, -1)$ is represented by the vector $\vec{v} = \langle 3, 2 \rangle$. Figure 3.4 shows several displacements represented by the same vector $\vec{v} = \langle 3, 2 \rangle$. Notice that the same vector is used to represent the displacement between any two points with the same difference in x-coordinates and the same difference in y-coordinates. When we talk about the displacement required to go from one point to another, we sometimes use the words "initial point" for the starting point and "final point" for the ending point.

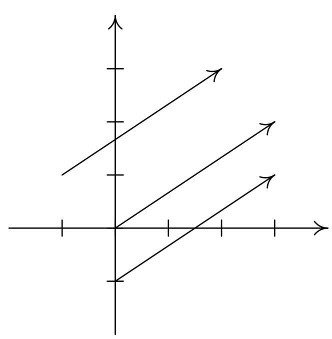


Figure 3.4: Several displacements represented by the vector $\vec{v} = \langle 3, 2 \rangle$

If the point A is represented by the vector $\vec{a}$ and the point B is represented by the vector $\vec{b}$, then the displacement required to reach the point B from the point A is just $\vec{b} - \vec{a}$. We sometimes denote this vector $\overrightarrow{AB}$.

Question 3 *If $A = (3, 2)$ and $B = (5, 7)$ find $\overrightarrow{AB}$ both numerically and graphically.*

Question 4 *If $A = (5, 7)$ and $B = (3, 2)$ find $\overrightarrow{AB}$ both numerically and graphically.*

Scalar multiplication has a simple geometric meaning as shown in Figure 3.5. Multiplying a vector by a positive number bigger than one stretches it. Multiplying a vector by a positive number less than one shrinks it. Multiplying a vector by a negative number reverses its direction and can also shrink it or stretch it.

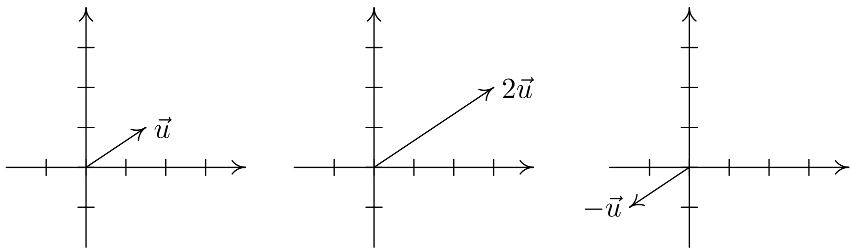


Figure 3.5: The geometric meaning of scalar multiplication

Vector addition also has a simple geometric meaning as shown in Figure 3.6 on page 225. If the vectors $\vec{u}$ and $\vec{v}$ represent the sides of a parallelogram, then the vector $\vec{u} + \vec{v}$ represents the diagonal shown in Figure 3.6.

Similarly, the vector operation $\vec{v} - \vec{u}$ has a simple geometric meaning as shown in Figure 3.7 on page 225. If $\vec{u}$ and $\vec{v}$ represent two sides of a triangle, then the vector $\vec{v} - \vec{u}$ represents the third side going from the point $\vec{u}$ to the point $\vec{v}$. You can also think of $\vec{v} - \vec{u}$ as the motion or displacement required to reach the point represented by $\vec{v}$ from the point represented by $\vec{u}$. This is shown clearly in Figure 3.7.

3.1.1 Applications

You have probably seen movies like Star Wars in which computer generated spaceships move across the screen. These movies are made using the two ideas we have just discussed. Suppose you want to move a spaceship from one point, represented by a vector $\vec{u}$, to another point, represented by a vector $\vec{v}$. This can be done by creating a series of frames in each of

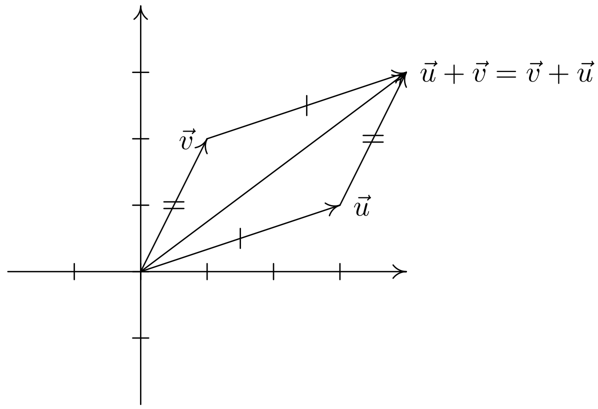


Figure 3.6: The geometric meaning of vector addition

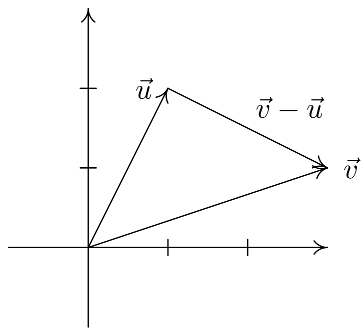


Figure 3.7: The geometric meaning of vector subtraction

which the spaceship has moved a little further along the desired path. When these frames are projected in a rapid sequence we see the illusion of motion. In Figure 3.8 on page 226, a series of dots represent different positions of the space ship as it moves. Each of the dots represents one of the points in the following list.

- $\vec{u} = \vec{u} + 0.0(\vec{v} - \vec{u})$. This dot is labeled $\vec{u}$ and is slightly larger than the unlabeled dots.
- $\vec{u} + 0.1(\vec{v} - \vec{u})$.
- $\vec{u} + 0.2(\vec{v} - \vec{u})$.
- $\vec{u} + 0.3(\vec{v} - \vec{u})$.

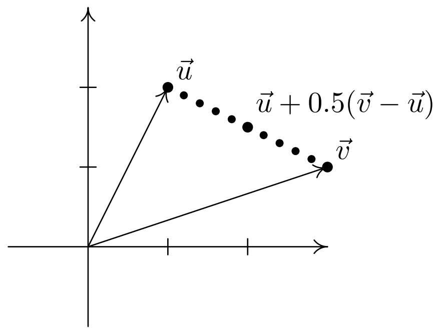


Figure 3.8: Positions of a spaceship

- $\vec{u} + 0.4(\vec{v} - \vec{u})$.
- $\vec{u} + 0.5(\vec{v} - \vec{u})$. This dot is labeled $\vec{u} + 0.5(\vec{v} - \vec{u})$ and is slightly larger than the unlabeled dots.
- $\vec{u} + 0.6(\vec{v} - \vec{u})$.
- $\vec{u} + 0.7(\vec{v} - \vec{u})$.
- $\vec{u} + 0.8(\vec{v} - \vec{u})$.
- $\vec{u} + 0.9(\vec{v} - \vec{u})$.
- $\vec{v} = \vec{u} + 1.0(\vec{v} - \vec{u})$. This dot is labeled $\vec{v}$ and is slightly larger than the unlabeled dots.

Question 5 *How would you use vectors and vector operations to find the point halfway between the points represented by $\vec{u}$ and $\vec{v}$?*

Question 6 *How would you use vectors and vector operations to find the point two-thirds of the way from the point represented by $\vec{u}$ to the point represented by $\vec{v}$?*

Click here[3] to open a *Mathematica* notebook showing how these operations create the illusion of motion – moving a black dot from the position marked by a red dot to the position marked by a blue dot. See Figure 3.9 on page 227. To see the animation, evaluate the cell.

[3]http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/moving-dot.nb

```
Clear[a, b]
a = {1, 4};
b = {5, 2};
ptA = Graphics[{PointSize[0.05], RGBColor[1, 0, 0], Point[a]}];
ptB  = Graphics[{PointSize[0.05], RGBColor[0, 0, 1], Point[b]}];
Animate[
  c = a + t * (b - a);
  ptC = Graphics[{PointSize[0.03], RGBColor[0, 0, 0], Point[c]}];
  Show[{ptA, ptB, ptC}, PlotRange → {{-1, 6}, {-1, 6}}, Axes → True,
      AspectRatio → Automatic], {t, 0, 1, 0.01}]
```

Figure 3.9: *Mathematica* screenshot

The key line in this *Mathematica* notebook is the line in large type. The number t is a parameter used in computing the vector

$$\vec{c} = \vec{a} + t(\vec{b} - \vec{a})$$

As t moves from zero to one, the point represented by the vector $\vec{c}$ moves from the point represented by the vector $\vec{a}$ (when $t = 0$) to the point represented by the vector $\vec{b}$ (when $t = 1$). This *Mathematica* notebook creates an animation by showing a sequence of still pictures, or frames, with the point represented by the vector $\vec{c}$ moving by a small amount from each frame to the next frame.

Question 7 *Note the vectors $\vec{u}$ and $\vec{v}$ in Figure 3.10 on page 228. Sketch each of the following vectors on Figure 3.10. Label your answers on the figure.*

- $2\vec{u}$.
- $\vec{v} + 2\vec{u}$.
- $2\vec{u} - \vec{v}$.
- $-\vec{u}$.
- $-\vec{v}$.
- $-(\vec{u} + \vec{v})$.

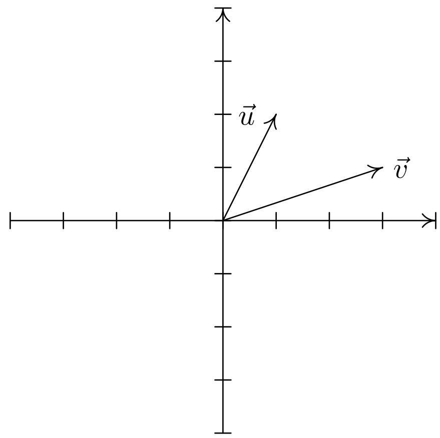


Figure 3.10: Question 7

Question 8 *The food that astronauts bring with them is carefully planned. They have a choice of meals but there are many conditions that must be met. Suppose that astronauts have a choice of ten possible meals and that each of three astronauts chooses meals represented by one of the following vectors*

$$\begin{aligned}\vec{a} &= < a_1, a_2, a_3, \ldots, a_{10} > \\ \vec{b} &= < b_1, b_2, b_3, \ldots, b_{10} > \\ \vec{c} &= < c_1, c_2, c_3, \ldots, c_{10} >\end{aligned}$$

for each week on the International Space Station (ISS). For example, a_1 represents the number of packages of meal 1 brought by astronaut A for each week.

Use vector operations to compute the following.

a. *A vector representing the total weekly food supply for all three astronauts.*

b. *Vectors representing the total food supply for each of the three astronauts for ten weeks.*

c. *A vector representing the total food supply for all three astronauts for ten weeks. Describe two different ways of finding this vector. Does it make any difference which way you use?*

d. *Suppose the calorie content of each meal is given by the vector* $\vec{v} =< v_1, v_2, v_3, \ldots, v_{10} >$ *– that is,* v_1 *is the calorie content of meal 1,* v_2 *is the calorie content of meal 2, and so forth. How would you compute the total weekly calorie content for each astronaut?*

e. *Suppose the mass of each meal is given by the vector* $\vec{m} =< m_1, m_2, m_3, \ldots, m_{10} >$*. How would you compute the total mass of the weekly food supply for each astronaut?*

In practice, astronauts, especially those in crews made up of astronauts from different countries, often wind up trading meals on board the ISS just like kids trade the school lunches they bring from home.

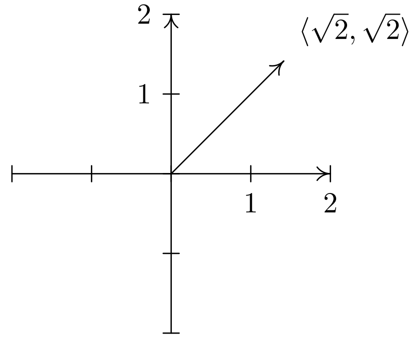


Figure 3.11: The vector $\langle\sqrt{2},\sqrt{2}\rangle$

3.2 Magnitude, Direction, and the Dot Product

You may have heard vectors described as having magnitude and direction – for example, the vector shown in Figure 3.11 can be described as a vector whose magnitude is 2 and whose direction is to the northeast, or whose direction makes an angle of 45° or $\pi/4$ radians with the x-axis. Up to now we have been using notation like $\langle\sqrt{2},\sqrt{2}\rangle$ to describe this vector. These two descriptions are just two ways of describing exactly the same thing.

In this section we are interested in going back-and-forth between a purely numeric description of a vector as a list of numbers and a more geometric description of a vector in terms of magnitude and direction. We will also develop a new vector operation called the **dot product**, or **scalar product** that is especially useful for using mathematics (and computers) to describe and create geometric images.

The **length**, **magnitude**, or **norm** (these are all synonyms) of a two-dimensional vector can be computed using the Pythagorean Theorem.

Definition 1 *The* **length**, **magnitude**, *or* **norm** *of a vector $\vec{u}=\langle u_1,u_2\rangle$ is written $|\vec{u}|$ and can be computed (See Figure 3.12 on page 231) by*

$$|\vec{u}|=\sqrt{u_1^2+u_2^2}.$$

In fact, a very similar computation works in three dimensions.

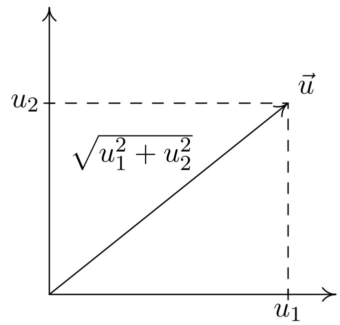


Figure 3.12: The magnitude of a vector

Definition 2 *If $\vec{u} = \langle u_1, u_2, u_3 \rangle$ then*

$$|\vec{u}| = \sqrt{u_1^2 + u_2^2 + u_3^2}.$$

This leads to the even more general, Definition 3, for an n-dimensional vector.

Definition 3 *If $\vec{u} = \langle u_1, u_2, \ldots u_n \rangle$ then*

$$|\vec{u}| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}.$$

Example 1 *If the displacement required to fly from one place to another is $\vec{u} = \langle 150, 200 \rangle$, measured in kilometers, then the length of the trip is given by*

$$|\vec{u}| = \sqrt{150^2 + 200^2} = 250 \text{ kilometers}$$

Question 1 *Let $\vec{u} = \langle 3, -2, 1 \rangle$. Find $|\vec{u}|$.*

Question 2 *Find $|\langle 3, -4 \rangle|$.*

Next we want to describe the **direction** of a vector. In two dimensions we use angles. Mathematicians usually describe angles in terms of radians (or sometimes degrees) measured counter-clockwise from the positive x-axis. In everyday life, however, angles are usually described in terms of degrees measured from the north. In three dimensions we use various ways to describe direction. Astronomers talk about azimuth and declination and in the military we often talk about azimuth and elevation.

We are going to develop a method that works in all dimensions, not just two and three dimensions. Think about sitting at the origin in the middle of a circle of radius 1. Think of this circle as a dial on which you are going to mark a direction. The different points on this circle or dial represent the different directions you might take when you walk away from the origin.

Now, look at Figure 3.13. This figure shows a vector $\langle 3, 4\rangle$. It also shows the unit circle centered at the origin. The dot on this circle marks the point on the dial that indicates the direction of the vector $\langle 3, 4\rangle$.

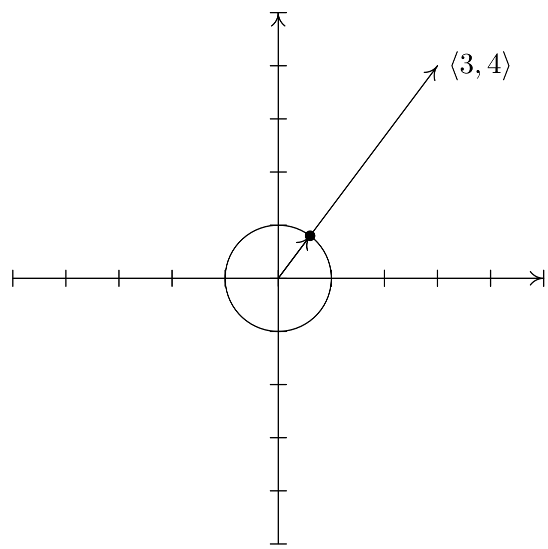


Figure 3.13: Describing direction by a unit vector

We can think of this dot as the tip of a vector. Since this vector has the same direction as the vector $\vec{v} = \langle 3, 4\rangle$, it must be a scalar multiple of $\vec{v}$ and, since it has magnitude 1, it must be the vector

$$\vec{u} = \frac{1}{|\vec{v}|}\ \vec{v}.$$

Vectors like the vector $\vec{u}$, that have length 1, are important because they capture the idea of direction. In view of this importance, we make the following definition.

Definition 4 *A vector $\vec{u}$ that has length 1 is called a* **unit vector.** *We often use the notation $\hat{u}$ instead of $\vec{u}$ to emphasize that a vector is a unit vector.*

Notice that if $\vec{v}$ is any vector then the vector

$$\hat{u} = \frac{1}{|\vec{v}|}\ \vec{v}$$

is a unit vector and we can write

$$\vec{v} = |\vec{v}|\ \hat{u}.$$

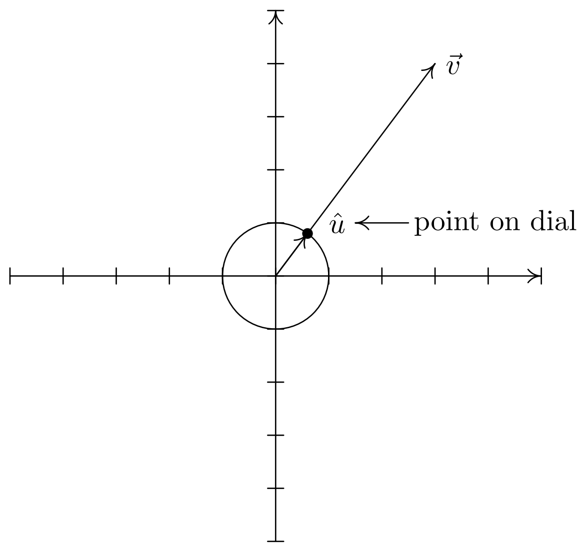


Figure 3.14: A vector using a unit vector and magnitude

We think of the unit vector, $\hat{u}$, as indicating the direction of the vector $\vec{v}$. Thus, we have written $\vec{v}$ in terms of its magnitude, $|\vec{v}|$, and its direction, $\hat{u}$. See Figure 3.14. The key idea is that the unit vector, $\hat{u}$, represents a point on a dial centered at the origin and that points on this dial indicate direction from the origin. In three dimensions, we get exactly the same picture except that now the dial is spherical rather than circular. In higher dimensions the dial is harder to visualize but the idea is the same – we can write any vector in terms of its magnitude and a unit vector that indicates its direction. More precisely, we have written it as the product of its length and its direction, as represented by a unit vector.

Question 3 *Express each of the following vectors in terms of its magnitude and a unit vector indicating its direction.*

a. $\vec{v} = \langle 4, 3 \rangle$

b. $\vec{v} = \langle -4, 3 \rangle$

c. $\vec{v} = \langle 4, -3 \rangle$

d. $\vec{v} = \langle -4, -3 \rangle$

e. $\vec{v} = \langle 1, 2 \rangle$.

f. $\vec{v} = \langle 1, 2, 3 \rangle$.

g. $\vec{v} = \langle 1, 2, 3, 4 \rangle$.

We often write two-dimensional vectors in the form

$$\vec{v} = \langle v_1, v_2 \rangle = \langle |\vec{v}| \cos\theta, |\vec{v}| \sin\theta \rangle = |\vec{v}| \langle \cos\theta, \sin\theta \rangle$$

based on Figure 3.15.

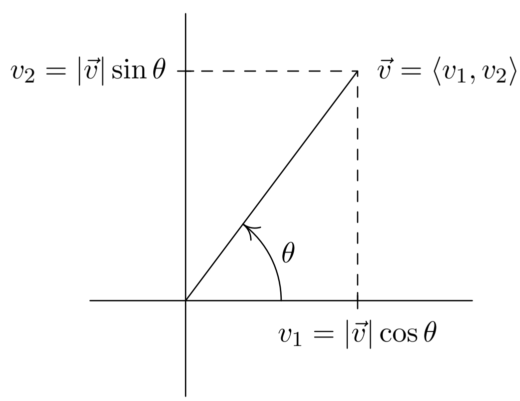


Figure 3.15: Expressing direction by angle

Question 4 *Show that the vector* $\langle \cos\theta, \sin\theta \rangle$ *is a unit vector.*

Our new description is used in many ways. One example is aiming devices that allow a pilot to aim at a target by looking at it. An on-board computer determines the direction of the vector from the pilot's eye to the target by sensing where the pilot is looking. This direction is expressed as a unit, three-dimensional vector because targets live in our three-dimensional world. Then, a ranging device is used to determine the magnitude of the vector from the pilot's eye to the target.

Our next goal is another useful vector operation, called the **dot product** or **scalar product**. We begin with an example.

Example 2 *In this example we examine oil consumption. For this discussion we divide the world into six groups of countries. The first three groups each contain just a single country – the United States, China, and India. We divide the remaining countries into three groups by income. This division is used by the World Bank. The fourth group is high income countries excluding the United States. The fifth group is middle income countries excluding China. The sixth group is low income countries excluding India. The 2004 population of each group is given in Table 3.4.*

United States	China	India	Other High	Other Middle	Other Low
293,027,571	1,298,847,624	1,065,070,607	707,306,600	1,709,730,000	1,258,362,000

Table 3.4: 2004 population

The 2004 oil consumption per capita (that is, per person) for these six groups is given in barrels in Table 3.5.

United States	China	India	Other High	Other Middle	Other Low
25.51	1.90	0.88	15.78	4.08	1.85

Table 3.5: 2004 per capita oil consumption in barrels

We can represent this information by two six-dimensional vectors.

$$\vec{p} = \langle 293027571,\ 1298847624,\ 1065070607,\ 707306600,\ 1709730000,\ 1258362000\rangle$$

and

$$\vec{q} = \langle 25.51,\ 1.90,\ 0.88,\ 15.78,\ 4.08,\ 1.85\rangle$$

Suppose that we want to compute the total oil consumption for the world in 2004. We need to multiply the population in each group by the per capita oil consumption in the same group to get the total oil consumption in each group. Then we need to add these six figures. In other words we need to compute

$$(p_1 \times q_1) + (p_2 \times q_2) + (p_3 \times q_3) + (p_4 \times q_4) + (p_5 \times q_5) + (p_6 \times q_6)$$

The answer is roughly, 3.13×10^{10} *or slightly more than 31 billion barrels.*

The computation in the example above is extremely common. For example, suppose that you buy four boxes of cereal at \$2.50 per box and two gallons of milk at \$3.00 per gallon. You can represent your purchases by the vector $\vec{q} = \langle 4, 2 \rangle$ and the prices by the vector $\langle \$2.50, \$3.00 \rangle$. Then the total price you must pay is

$$(4 \times \$2.50) + (2 \times \$3.00) = \$16.00.$$

This leads us to define a new vector operation, called the **scalar product** or the **dot product**.

Definition 5 *Suppose that*

$$\vec{u} = \langle u_1, u_2, \ldots, u_n \rangle$$

and

$$\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle$$

are two n*-dimensional vectors. We define their* **scalar product**, *or* **dot product**, *denoted* $\vec{u} \cdot \vec{v}$ *by*

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i.$$

The two terms "scalar product" and "dot product" are synonyms.

Note that the scalar *product* we developed in this section is different from the scalar *multiplication* we discussed in the previous section and that we use the notation $\vec{u} \cdot \vec{v}$ for the *scalar product* of two n-dimensional *vectors* and the notation $c\vec{u}$ for the *scalar multiple* of the *scalar* c and the *vector* $\vec{u}$. This notation can be confusing because in the past you have used the notation $x \cdot y$ and xy as synonyms for the product of two real numbers x and y. When we deal with vectors the two notations are not synonymous.

Question 5 *For each of the following, identify whether the product is a scalar multiplication or a scalar product and if the operation is legal. If it is legal then compute the answer. If not, then state why it is not.*

1. $4\langle 1, 2, 3, 4\rangle$
2. $\langle 1, 2, 3, 4\rangle \cdot \langle 3, 2, 1\rangle$.
3. $6 \cdot \langle 3, 2, 1\rangle$.
4. $\langle 1, 2, 3, 4\rangle \cdot \langle -1, 2, 1, -2\rangle$.
5. $(\langle 1, 2, 3\rangle \cdot \langle 3, 2, 1\rangle) \cdot \langle 4, 2, 4\rangle$.
6. $(\langle -1, 2, -2\rangle \cdot \langle 2, -1, 2\rangle)\langle 2, 1, 3, -1\rangle$.

We have been talking about vectors using both geometric ideas and numeric computations. For example, the magnitude of a vector is a geometric idea and we can compute it numerically. The connection between geometry and numerical computation is at the heart of computer-based simulations and movies like Star Wars. Our next theorem establishes another important connection between geometric ideas and numerical computations. Its proof requires a trigonometric identity:

$$\cos(\alpha - \beta) = \cos\alpha \cos\beta + \sin\alpha \sin\beta. \qquad (1)$$

Theorem 1 *Suppose that $\vec{u}$ and $\vec{v}$ are two, two-dimensional vectors. Then*

$$\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}| \cos\theta$$

where θ is the angle between the two vectors. *See Figure 3.16 on page 238.*

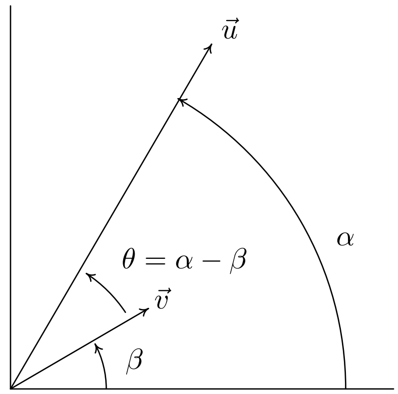


Figure 3.16: The angle between two vectors

Proof

Recall from Figure 3.15 on page 234 that we can write

$$\vec{u} = |\vec{u}|\ \langle \cos\alpha, \sin\alpha \rangle$$

and

$$\vec{v} = |\vec{v}|\ \langle \cos\beta, \sin\beta \rangle$$

$$\vec{u} \cdot \vec{v} = |\vec{u}|\ \langle \cos\alpha, \sin\alpha \rangle \cdot |\vec{v}|\ \langle \cos\beta, \sin\beta \rangle = |\vec{u}||\vec{v}|(\cos\alpha\cos\beta + \sin\alpha\sin\beta) = |\vec{u}||\vec{v}|\cos(\alpha - \beta) = |\vec{u}||\vec{v}|\cos\theta$$

by the trigonometric identity Equation (1). This proves the theorem since θ, the angle between the vectors $\vec{u}$ and $\vec{v}$, is $\alpha - \beta$. See Figure 3.16. ∎

Although the proof above assumed that the vectors were two-dimensional vectors, the theorem is true for three-dimensional and higher dimensional vectors. We state this more general theorem but do not give a proof.

Theorem 2 *Suppose that $\vec{u}$ and $\vec{v}$ are two, n-dimensional vectors. Then*

$$\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}|\cos\theta$$

where θ is the angle between the two vectors. See Figure 3.16.

Two vectors, $\vec{u}$ and $\vec{v}$, are perpendicular if the angle, θ, between them is a right angle. This implies that $\cos\theta = 0$ and $\vec{u} \cdot \vec{v} = 0$ leading to the following theorem.

Theorem 3 *Two vectors $\vec{u}$ and $\vec{v}$ are* **perpendicular** *if and only if*

$$\vec{u} \cdot \vec{v} = 0.$$

We use the word **orthogonal** as a synonym for perpendicular. When two vectors are perpendicular we write $\vec{u} \perp \vec{v}$.

Definition 6 *If $\vec{u}$ and $\vec{v}$ are two nonzero vectors, then we say they are* **parallel** *if there is a real number x such that $\vec{v} = x\vec{u}$.*

Question 6 *Determine which of the following pairs of vectors are perpendicular.*

- $\langle 3, 2\rangle$ *and* $\langle -4, 6\rangle$.
- $\langle 1, 2, -6\rangle$ *and* $\langle 1, 3, 1\rangle$.
- $\langle 1, 2, -6\rangle$ *and* $\langle 2, 2, 1\rangle$.

Question 7 *Determine which of the following pairs of vectors are parallel.*

- $\langle 1, 2, 4\rangle$ *and* $\langle 2, 4, 6\rangle$.
- $\langle 2, 0, 4\rangle$ *and* $\langle -4, 0, -8\rangle$.
- $\langle 1, 2, 3\rangle$ *and* $\langle 2, 3, 1\rangle$.

Question 8 *For each of the following pairs of vectors, determine whether they are perpendicular, parallel, or neither.*

- $\langle 2, 4, 8\rangle$ *and* $\langle 3, 6, 12\rangle$.
- $\langle 1, 2, 3\rangle$ *and* $\langle 2, -7, 4\rangle$.
- $\langle 1, 2, 3\rangle$ *and* $\langle -2, 7, -4\rangle$.

- $\langle 1, 2, 3\rangle$ *and* $\langle 2, 7, -4\rangle$.

We have established some powerful connections in this section between purely algebraic operations like the algebraic definition of the dot product

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

and the purely algebraic computation of length

$$|\vec{v}| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

and very geometric ideas like the length of a vector and the angle between two vectors. This connection is captured in the equation

$$\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}| \cos\theta$$

where θ is the angle between $\vec{u}$ and $\vec{v}$.

Because computers cannot measure angles using protractors as people do, this equation is one of the key equations involved in computer-based simulations. Computers can compute the angle between two vectors algebraically by

$$\begin{aligned}
|\vec{u}||\vec{v}| \cos\theta &= \vec{u} \cdot \vec{v} \\
\cos\theta &= \frac{\vec{u} \cdot \vec{v}}{|\vec{u}||\vec{v}|} \\
\theta &= \arccos\left(\frac{\vec{u} \cdot \vec{v}}{|\vec{u}||\vec{v}|}\right) \\
\theta &= \arccos\left(\frac{u_1 v_1 + u_2 v_2 + \cdots + u_n v_n}{\sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}}\right)
\end{aligned}$$

Of course, people can compute the angle between two vectors in the same way.

Question 9 *Find the angle between the vectors* $\langle 2,3\rangle$ *and* $\langle 5,-3\rangle$*.*

Question 10 *Find the angle between the vectors* $\langle 2,1,4\rangle$ *and* $\langle 2,-3,7\rangle$*.*

The vector operations – scalar multiplication, vector addition, and the dot product, or scalar product – are based on the familiar operations of addition and multiplication of real numbers. For this reason, many of the familiar properties of real number operations carry over to these vector operations.

Addition is commutative: Given any two real numbers, a and b, the two sums, $a+b$ and $b+a$, are the same. As a result if $\vec{u}$ and $\vec{v}$ are any two vectors then

$$\vec{u}+\vec{v}=\vec{v}+\vec{u}.$$

This property is called the **commutative property of addition** and we say that "vector addition is commutative."

Addition is associative: Given any three real numbers, a, b, and c, $a+(b+c)=(a+b)+c$. This is called the **associate property of addition.** It carries over to vector addition – if $\vec{u}$, $\vec{v}$, and $\vec{w}$ are any three vectors, then

$$\vec{u}+(\vec{v}+\vec{w})=(\vec{u}+\vec{v})+\vec{w}.$$

Multiplication is also commutative and associative – that is, given any three real numbers a, b, and c

$$ab=ba \text{ and } a(bc)=(ab)c.$$

As a result, given any two real numbers, a and b, and any vector $\vec{u}$,

$$a(b\vec{u})=(ab)\vec{u} \text{ and } a(b\vec{u})=b(a\vec{u}).$$

Multiplication is distributive across addition: Given any three real numbers, a, b and c, $a(b+c)=ab+ac$. As a result, given any real numbers, a and b, and any two vectors, $\vec{u}$ and $\vec{v}$,

$$a(\vec{u}+\vec{v})=a\vec{u}+a\vec{v}.$$

and

$$(a+b)\vec{u} = a\vec{u} + b\vec{u}.$$

Scalar product properties: Finally, given any real number, a, and any three vectors, $\vec{u}$, $\vec{v}$, and $\vec{w}$,

$$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u},$$

$$a(\vec{u} \cdot \vec{v}) = (a\vec{u}) \cdot \vec{v} = \vec{u} \cdot (a\vec{v}),$$

and

$$\vec{u} \cdot (\vec{v} + \vec{w}) = (\vec{u} \cdot \vec{v}) + (\vec{u} \cdot \vec{w}).$$

These properties are summarized below.

- $\vec{u} + \vec{v} = \vec{v} + \vec{u}$. This property is called the **commutative property of vector addition.**
- $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$. This property is called the **associative property of vector addition.**
- $a(b\vec{u}) = (ab)\vec{u}$ and $a(b\vec{u}) = b(a\vec{u})$. This property does not seem to have a name.
- $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$. This property is another **distributive property.**
- $(a+b)\vec{u} = a\vec{u} + b\vec{u}$. This property is another **distributive property.**
- $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$. This property is called the **commutative property of the dot product.**
- $a(\vec{u} \cdot \vec{v}) = (a\vec{u}) \cdot \vec{v} = \vec{u} \cdot (a\vec{v})$. This is another property without a name.
- $\vec{u} \cdot (\vec{v} + \vec{w}) = (\vec{u} \cdot \vec{v}) + (\vec{u} \cdot \vec{w})$. This property is another **distributive property.**

Question 11 *Compute the following wherever possible. If a particular computation is not possible state why.*

a. $\langle 1, 3, -6\rangle + 3\langle 3, 1, -4\rangle$.

b. $\langle 3, -2, 3\rangle \cdot \langle -2, 4, 6\rangle$.

c. $\langle 6, -1, 2\rangle + \langle 3, 4, 2, 1\rangle$.

d. $\langle 1, 2, 3\rangle \cdot \langle 5, 4, 3, 2\rangle$.

e. $0.25\langle 12, 4, 8\rangle + 0.33\langle 6, 9, 18\rangle$.

f. $\vec{u} + 2(\vec{v} - \vec{u})$.

Question 12 *Show that if $\vec{u} \perp \vec{v}$ and $\vec{u} \perp \vec{w}$ then $\vec{u} \perp (\vec{v} + \vec{w})$.*

Question 13 *Show that if $\vec{u}$, $\vec{v}$, and $\vec{w}$ are three nonzero vectors, then, if $\vec{u}$ and $\vec{v}$ are parallel and $\vec{v}$ and $\vec{w}$ are parallel, $\vec{u}$ and $\vec{w}$ are parallel.*

3.3 Matrix Algebra, I

We have been studying vectors, or lists[4] of numbers. Vectors are important because there are many situations in which we must keep track of lists of numbers rather than just a single number. The following application is a good example illustrating the power of vectors.

Application 1 *Sound recordings are vectors. In the old days, sound recordings were represented physically by magnetic signals on tapes. Nowadays, most sound recordings are digital with up to 44,100 numbers for each recorded second. Digital sound recordings are just very high dimensional vectors. For example, a ten second recording would be a vector*

$$\vec{v} = \langle v_1, v_2, \dots v_{441000} \rangle$$

You can physically add two sounds of the same length together by playing them at the same time and you can physically multiply a sound recording by a real number by running its signal through an amplifier. If the real number is negative, then you also have to run it through a circuit called an inverter. If the sound is recorded digitally, you can accomplish the same operations by using vector addition and scalar multiplication. If you have a pair of noise-cancelling headphones, you have a nice example of how this is used. Noise-cancelling headphones have a microphone that captures the noise around you. The signal captured by this microphone is inverted and added to the signal that reaches your ears. This effectively cancels the noise and allows you to listen to your music even in noisy environments.

There are other situations in which we must keep track of tables of numbers instead of lists of numbers. In these situations we use **matrices** instead of vectors. Tables typically have several rows and several columns. If a table, or matrix, has n rows and k columns we say it is an $(n \times k)$-matrix, read an "n by k matrix." We sometimes say the matrix has "dimension n by k." We use the following notation:

Definition 1 *We use a capital letter – for example, A – to denote a matrix and the same lower case letter with subscripts to denote its entries. For example, the entry in the i^{th} row and j^{th} column of the matrix A is denoted a_{ij}. We use square brackets – "[" and "]" – to enclose the entries in a matrix.*

[4] Recall that lists are ordered lists.

Example 1 *Let*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

This is a (2×3)*-matrix and its entries are*

$$\begin{aligned} a_{11} &= 1 \\ a_{12} &= 2 \\ a_{13} &= 3 \\ a_{21} &= 4 \\ a_{22} &= 5 \\ a_{23} &= 6 \end{aligned}$$

Application 2 *Images from digital cameras can be represented by matrices. Such images are made up of very small points or pixels arranged in a rectangle. The simplest images are monochrome, or black-and-white. A typical medium resolution image might have 800 rows with 1200 pixels in each row. If a monochrome image is made up of* n *rows of pixels and each row has* k *pixels, then it can be represented by a matrix that has* n *rows and* k *columns. Usually, each entry in such a matrix is a real number between zero and one, with zero being black and one being the brightest white that can be displayed by a particular device. Mathematically, we often ignore this restriction. It really only comes into play when the image is displayed. Color images can be represented by three matrices, each of whose entries represent the intensity of one of the three colors – red, green, and blue – for one of the pixels.*

In this section, we look at two operations – multiplying a matrix by a scalar and matrix addition – that are very similar to the two vector operations – multiplying a vector by a scalar and vector addition – from the last section. These operations are very easy and have the same properties that we saw in the last section. They are also powerful. Later in this section we look at an example showing how these operations are used in image processing.

Definition 2 *Suppose that*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}$$

is a matrix with n rows and k columns, or an $(n \times k)$-matrix, and that c is a real number. We define the matrix cA by

$$cA = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1k} \\ ca_{21} & ca_{22} & \cdots & ca_{2k} \\ \vdots & \vdots & & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nk} \end{bmatrix}$$

Definition 3 *Suppose that*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}$$

and

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{bmatrix}$$

are both matrices with n rows and k columns. Note: They must both have the same number of rows and the same number of columns. We define the sum $A + B$ by

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1k} + b_{1k} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2k} + b_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{nk} + b_{nk} \end{bmatrix}$$

Example 2

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 2 & -1 & 3 \\ -1 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 6 \\ 3 & 8 & 4 \end{bmatrix}$$

Note that the two matrices added in this example both have the same dimension, (2×3)*, and the result has the same dimension.*

Example 3

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} = \text{undefined.}$$

Note that in this example the attempt to add the two matrices fails because they have different dimensions – one has dimension (2×3) *and the other has dimension* (2×2)*.*

The two operations, scalar multiplication of matrices and matrix addition, defined above, have the familiar properties –

- If A and B are both $(n \times k)$-matrices then $A + B = B + A$.
- If A, B, and C are all $(n \times k)$-matrices then $A + (B + C) = (A + B) + C$.
- If A is an $(n \times k)$-matrix and a and b are real numbers then $a(bA) = (ab)A = b(aA)$.
- If A and B are both $(n\times k)$-matrices and a is a real number then $a(A+B) = aA+aB$.
- If A is an $(n \times k)$-matrix and a and b are real numbers then $(a + b)A = aA + bA$.

Question 1 *Using the matrices below*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 0 & 6 \\ 2 & 1 & -3 \end{bmatrix} \qquad C = \begin{bmatrix} -1 & 3 \\ 3 & -2 \end{bmatrix}$$

compute each of the following when it is possible. If it is not possible state why.

a. $2A + 2B$

b. $2(A + B)$

c. $3A - B$

d. $A + C$

e. $C + A$.

f. $A + (B - A)$

The two pictures in Figure 3.17 were taken at Lake Minnewaska State Park in New York State. They show Awosting Falls in the spring and in the winter. Pictures like this are often used in movies to suggest the passage of time. At the end of one scene, the camera might focus on the picture on the left – Lake Awosting in spring. Then, that picture might fade out at the same time as the right hand picture – Lake Awosting in winter – fades in and music suggests the passage of time.

Figure 3.17: Awosting Falls in spring and winter

Figure 3.9 on page 227 showed how we can use vector operations to create the illusion of motion. To move a dot, for example, from the point represented by the vector $\vec{a}$ to the point represented by the vector $\vec{b}$ we created a series of frames with the dot at the point

$$\vec{c} = \vec{a} + t(\vec{b} - \vec{a})$$

and as t goes from 0 to 1 the dot moves from the point represented by the vector $\vec{a}$ to the point represented by the vector $\vec{b}$. We use the same idea together with matrix operations to create a smooth transition from Awosting Falls in the spring to Awosting Falls in the winter. We start with a matrix A that represents our picture of Awosting Falls in the spring and a matrix B that represents our picture of Awosting Falls in the winter. Then we look at the matrices

$$C = A + t(B - A)$$

as t varies from 0 to 1. Typically, we would generate 20 or 30 frames and then show them in rapid succession to create a smooth transition. Figure 3.18 shows three of these frames.

Figure 3.18: Three intermediate frames in a transition from spring to winter

We close this section with an application that continues an application from Section 3.1 and some questions that will motivate a new idea in the next section.

Application 3 *In order to combat an insurgency in a particular country, it is important to keep track of the number of insurgents in different parts of the country. This task is complicated by the fact that the insurgents are highly mobile and move around a lot. Analysts have divided the country into six sectors and they estimate that the number of insurgents in each sector during the current week is given by Table 3.6 on page 250.*

Intel indicates that each week insurgents move from one sector to another according to Table 3.7 on page 250. The entry in the row labeled "to sector i" and the column labeled "from sector j" indicates the fraction of the insurgents in sector j that move from sector j to sector i each week. For example, each week 17% of the insurgents in sector 4 move from

sector	number
1	160
2	182
3	231
4	119
5	158
6	318

Table 3.6: Insurgents in each sector (current week)

sector 4 to sector 1 and each week 8% of the insurgents in sector 5 move from sector 5 to sector 2.

to sector	from sector					
	1	2	3	4	5	6
1	0.40	0.10	0.05	0.17	0.05	0.05
2	0.04	0.30	0.15	0.07	0.08	0.15
3	0.10	0.10	0.45	0.12	0.06	0.08
4	0.08	0.07	0.09	0.42	0.04	0.04
5	0.10	0.15	0.05	0.14	0.40	0.06
6	0.12	0.09	0.07	0.06	0.05	0.60

Table 3.7: Movements among sectors from one week to the next week

Question 2 *Based on Tables 3.6 and 3.7, how many insurgents would you expect to be in each sector the week after the current week?*

Question 3 *Based on Tables 3.6 and 3.7, how many insurgents would you expect to be in each sector two weeks after the current week?*

Question 4 *Based on Tables 3.6 and 3.7, how many insurgents would you think would have been in each sector the week prior to the current week?*

Used together, matrices and vectors are powerful modeling tools in situations where there are many quantities of interest. We have already seen several examples. The questions below ask you to use these tools to build some models based on the data shown, graphically, in Figure 3.19 on page 251 and, numerically, in Table 3.8 on page 252. This set of data is refered to as a **population pyramid** because it presents a snapshot of population broken down by age and gender and the graphs often look like pyramids. This data and similar data is available at http://www.census.gov/ipc/www/idbnew.html. You can see a lot from data like this – for example, there are a lot of people whose age is between 40 and 59. This age group will put a strain on social security and medicare over the next 20 - 30 years.

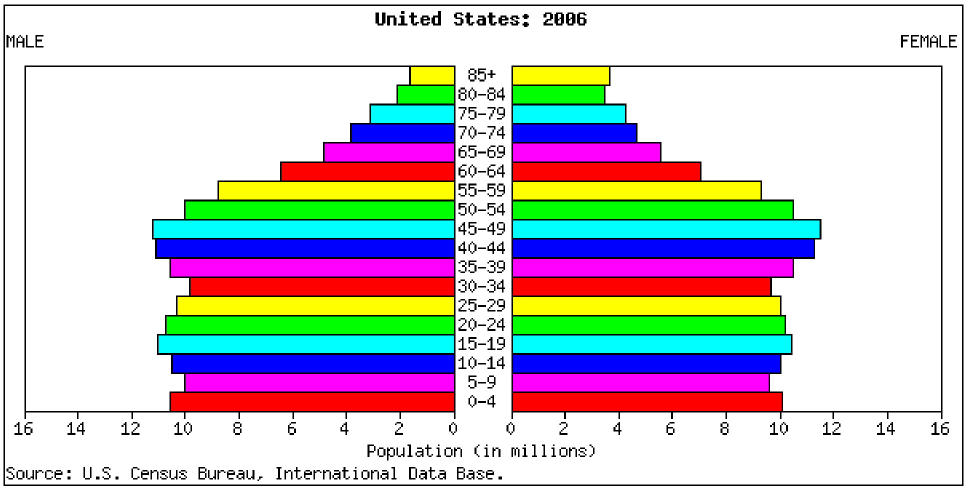


Figure 3.19: United States population by gender and age group midyear 2006

Question 5

a. *Determine a vector whose entries represent the male population in each age group.*

b. *Determine a vector whose entries represent the female population in each age group.*

c. *Determine a vector whose entries represent the total population in each age group.*

Question 6 *Pick another country in which you are interested and get similar data from the Web site http://www.census.gov/ipc/www/idbnew.html. For the country you choose –*

a. *Determine a vector whose entries represent the male population in each age group.*

b. *Determine a vector whose entries represent the female population in each age group.*

c. *Determine a vector whose entries represent the total population in each age group.*

From Age	To Age	Male	Female
0	4	10,544,578	10,094,346
5	9	10,034,348	9,597,204
10	14	10,516,921	10,024,322
15	19	11,012,552	10,451,316
20	24	10,754,502	10,194,491
25	29	10,336,863	10,022,829
30	34	9,833,159	9,659,910
35	39	10,542,520	10,485,644
40	44	11,106,208	11,243,245
45	49	11,215,604	11,512,111
50	54	10,025,751	10,470,636
55	59	8,754,898	9,310,852
60	64	6,440,788	7,062,450
65	69	4,833,624	5,543,152
70	74	3,819,597	4,698,878
75	79	3,096,612	4,252,239
80	–	3,792,455	7,159,610

Table 3.8: United States population by gender and age group midyear 2006

Question 7 *During every presidential year, there is a great deal of discussion about social security and medicare. The year 2008 is unlikely to be an exception. Make a rough estimate for the fraction of the population that is retired in each age group and gender. Based on your estimates, how many people were retired in 2006?*

Question 8 *For the country you chose in Question 6, make a rough estimate for the fraction of the population that is retired in each age group and gender. Based on your estimates how many people were retired in 2006 in that country?*

Question 9 *Suppose that you would like to predict the population of the United States by age group and gender in the year 2011 – that is, five years after the data presented in Figure 3.19 on page 251 and Table 3.8. Begin discussing this problem with words. What factors are important as you try to make a prediction? What data would you like? As you attack this question, consider the following steps which are useful in any modeling problem. These steps can help with the arrow in the modeling triangle (Figure 3.20 on page 253 going from the real world to the mathematical formulation – expressing a real world problem mathematically or transforming a real world problem into a mathematical one.*

- *Identify the quantities that are* **known.**
- *Identify the quantities that are* **unknown.**

- *Identify the* **relationships** *between the unknown quantities and the known quantities.*
- *Identify additional unknown quantities or data whose determination would be useful.*

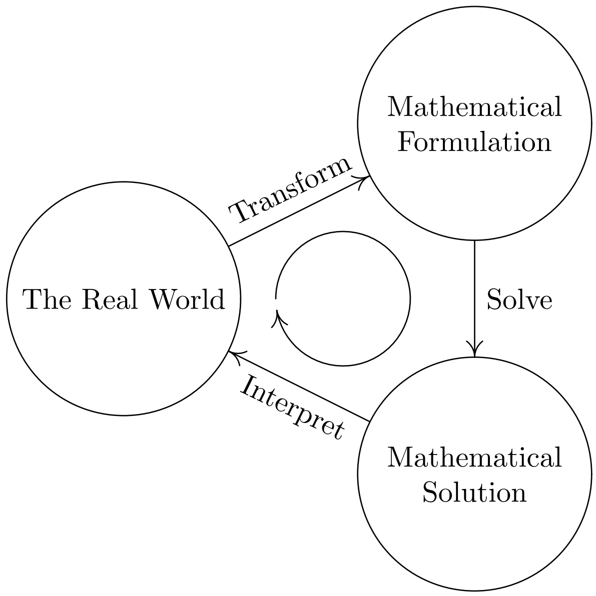


Figure 3.20: The modeling triangle

As one example, the people in each age group and gender who survive for five years will move to the next age group and the same gender. One crucial piece of additional data is the fraction in each age group and gender that will survive for five years.

3.4 Matrix Algebra, II

We begin this section with a matrix operation that will prove to be extremely powerful. It is called **Matrix Multiplication.**

Definition 1 *Suppose that A is an $(n \times k)$-matrix and that B is a $(k \times p)$-matrix. Notice that the number of columns in the matrix A must match the number of rows in the matrix B.*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kp} \end{bmatrix}$$

We define a matrix C called the product of A and B in that order and written $AB = C$ as follows.

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{np} \end{bmatrix}$$

where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} = \sum_{r=1}^{k} a_{ir}b_{rj}.$$

For example,

$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2k}b_{k1}.$$

Notice that the entry, c_{ij}, in the i^{th} row and j^{th} column of the matrix $C = AB$ is computed using the entries from the i^{th} row of the matrix A and the j^{th} column of the matrix B. It is the dot product of the i^{th} row of the matrix A and the j^{th} column of the matrix B. That is why the matrix A must have the same number of columns as the matrix B has rows.

The matrix C is an $(n \times p)$-matrix – it has the same number of rows as the first matrix, A, and the same number of columns as the second matrix, B.

Example 1

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 \times (-1) + 2 \times 4 + 3 \times 2 & 1 \times 3 + 2 \times (-2) + 3 \times (-1) \\ 4 \times (-1) + 5 \times 4 + 6 \times 2 & 4 \times 3 + 5 \times (-2) + 6 \times (-1) \end{bmatrix} = \begin{bmatrix} 13 & -4 \\ 28 & -4 \end{bmatrix}$$

Example 2

$$\begin{bmatrix} 2 & 1 & -3 \\ -2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 & 4 \\ 6 & -2 & 3 & -2 \\ 3 & 2 & -1 & 5 \end{bmatrix} = \begin{bmatrix} -3 & -4 & 8 & -9 \\ 30 & -2 & 3 & 6 \end{bmatrix}$$

Example 3

$$\begin{bmatrix} 0 & 2 & 1 & 4 \\ 6 & -2 & 3 & -2 \\ 3 & 2 & -1 & 5 \end{bmatrix} \begin{bmatrix} 2 & 1 & -3 \\ -2 & 3 & 4 \end{bmatrix} = \text{undefined.}$$

This product is undefined because the number of columns in the first matrix does not match the number of rows in the second matrix.

Question 1 *Four of the six entries in the matrix multiplication below have been worked out. Fill in the two missing entries.*

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 & 4 \\ 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 7 & \\ 6 & & 12 \end{bmatrix}$$

Question 2 *Let*

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

Find the two matrix products AB and BA. If either or both of these products don't exist, say so.

sector	number
1	160
2	182
3	231
4	119
5	158
6	318

Table 3.9: Insurgents

Example 4

Recall that, in the last section, we looked at insurgents in six sectors of a particular country. Table 3.9 describes the current number of insurgents in each sector. We rewrite the information in this table as a (6×1)*-matrix*

$$P = \begin{bmatrix} 160 \\ 182 \\ 231 \\ 119 \\ 158 \\ 318 \end{bmatrix}.$$

Table 3.10 describes the fraction of insurgents in each sector that moved from that sector to each of the six sectors from one week to the next.

to sector	from sector					
	1	2	3	4	5	6
1	0.40	0.10	0.05	0.17	0.05	0.05
2	0.04	0.30	0.15	0.07	0.08	0.15
3	0.10	0.10	0.45	0.12	0.06	0.08
4	0.08	0.07	0.09	0.42	0.04	0.04
5	0.10	0.15	0.05	0.14	0.40	0.06
6	0.12	0.09	0.07	0.06	0.05	0.60

Table 3.10: Movements of insurgents among sectors

We rewrite the information in this table as a (6×6)*-matrix*

$$T = \begin{bmatrix} 0.40 & 0.10 & 0.05 & 0.17 & 0.05 & 0.05 \\ 0.04 & 0.30 & 0.15 & 0.07 & 0.08 & 0.15 \\ 0.10 & 0.10 & 0.45 & 0.12 & 0.06 & 0.08 \\ 0.08 & 0.07 & 0.09 & 0.42 & 0.04 & 0.04 \\ 0.10 & 0.15 & 0.05 & 0.14 & 0.40 & 0.06 \\ 0.12 & 0.09 & 0.07 & 0.06 & 0.05 & 0.60 \end{bmatrix}.$$

Based on the information in these two tables, we want to determine how many insurgents will be in each sector next week. Consider the product

$$\begin{aligned} TP &= \begin{bmatrix} 0.40 & 0.10 & 0.05 & 0.17 & 0.05 & 0.05 \\ 0.04 & 0.30 & 0.15 & 0.07 & 0.08 & 0.15 \\ 0.10 & 0.10 & 0.45 & 0.12 & 0.06 & 0.08 \\ 0.08 & 0.07 & 0.09 & 0.42 & 0.04 & 0.04 \\ 0.10 & 0.15 & 0.05 & 0.14 & 0.40 & 0.06 \\ 0.12 & 0.09 & 0.07 & 0.06 & 0.05 & 0.60 \end{bmatrix} \begin{bmatrix} 160 \\ 182 \\ 231 \\ 119 \\ 158 \\ 318 \end{bmatrix} \\ &= \begin{bmatrix} (0.40 \times 160) + (0.10 \times 182) + (0.05 \times 231) + (0.17 \times 119) + (0.05 \times 158) + (0.05 \times 318) \\ (0.04 \times 160) + (0.30 \times 182) + (0.15 \times 231) + (0.07 \times 119) + (0.08 \times 158) + (0.15 \times 318) \\ (0.10 \times 160) + (0.10 \times 182) + (0.45 \times 231) + (0.12 \times 119) + (0.06 \times 158) + (0.08 \times 318) \\ (0.08 \times 160) + (0.07 \times 182) + (0.09 \times 231) + (0.42 \times 119) + (0.04 \times 158) + (0.04 \times 318) \\ (0.10 \times 160) + (0.15 \times 182) + (0.05 \times 231) + (0.14 \times 119) + (0.40 \times 158) + (0.06 \times 318) \\ (0.12 \times 160) + (0.09 \times 182) + (0.07 \times 231) + (0.06 \times 119) + (0.05 \times 158) + (0.60 \times 318) \end{bmatrix} \\ &= \begin{bmatrix} 137.78 \\ 164.32 \\ 187.35 \\ 115.30 \\ 153.79 \\ 257.59 \end{bmatrix}. \end{aligned}$$

Lets look at the calculation we would need to make to determine how many insurgents will be in sector 1 next week. The insurgents that will be in sector 1 next week come from the six sectors. There are currently 160 insurgents in sector 1, and according to the table, 40% of them will remain in sector 1. Thus, sector 1 contributes $0.40 \times 160 = 64$ *insurgents to sector 1 next week. Notice that this is just the first term in the computation below for the first element of the product* TP*.*

$$(0.40 \times 160) + (0.10 \times 182) + (0.05 \times 231) + (0.17 \times 119) + (0.05 \times 158) + (0.05 \times 318)$$

Sector 2 currently has 182 insurgents and, according to the table, 10% of them will move to sector 1 next week. Thus, sector 2 will contribute $0.10 \times 182 = 18.2$ *insurgents to sector*

1 next week. Notice this is just the second term in the computation above. This analysis continues in the same way. The contribution that each of the six sectors makes to the number of insurgents in sector 1, next week, is computed by one of the six terms in the computation above of the first element of the product TP. The computation of the number of insurgents in sector 2, next week, corresponds to the computation of the second element of the product TP; the computation of the number of insurgents in sector 3, next week, corresponds to the computation of the third element of the product TP; and so forth. This is not an accident. The definition of the matrix product was invented to apply to situations like this.

Thus, next week the number of insurgents in each sector is given by Table 3.11[5] *and the following week by Table 3.12.*

sector 1	sector 2	sector 3	sector 4	sector 5	sector 6
137.78	164.32	187.35	115.35	153.79	257.59

Table 3.11: Number of insurgents in each sector the next week

sector 1	sector 2	sector 3	sector 4	sector 5	sector 6
121.090	141.926	158.194	104.288	140.914	213.601

Table 3.12: Number of insurgents in each sector the following week

Definition 2 *For each n, we define an $(n \times n)$-matrix, I_n, called the* **identity matrix** *of dimension n by*

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Sometimes, when the dimension of an identity matrix is obvious from the context, we simply write I rather than I_n.

[5]Notice that mathematically we have fractional insurgents. In practice, of course, fractional people are not realistic.

Question 3 *Write the matrix* I_3.

Question 4 *Compute*

$$I_3 \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Question 5 *Compute*

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} I_3.$$

Question 6 *Suppose that* A *is an* $(n \times n)$*-matrix. What is* $I_n A$*? What is* AI_n*?*

Question 7 *Suppose that* A *is an* $(n \times k)$*-matrix. What is* $I_n A$*? What is* $I_k A$*? What is* AI_n*? What is* AI_k*?*

In the problems above you should have noticed that if A is an $(n \times k)$-matrix then $I_n A = A$ and $AI_k = A$ but that $I_k A$ and AI_n are undefined unless $n = k$.

Question 8 *Suppose that* X *is the matrix*

$$X = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and that A *is any* (3×3)*-matrix. Describe* XA*. Describe* AX*.*

Question 9 *Suppose that* X *is the matrix*

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

and that A *is any* (3×3)*-matrix. Describe* XA*. Describe* AX*.*

Question 10 *Suppose that X is the matrix*

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

and that A is any (3×3)-matrix. Describe XA. Describe AX.

Question 11 *Suppose that X is the $(n \times n)$-matrix*

$$X = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & x_n \end{bmatrix}.$$

A matrix like X that has zeros every place but on the diagonal is called a **diagonal matrix**. *Suppose that A is an $(n \times n)$-matrix. Describe XA. Describe AX.*

Definition 3 *It is often useful to write a vector $\vec{x} = \langle x_1, x_2, \ldots, x_n \rangle$ as a matrix. We can do this in two ways – as a* **row-vector**:

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$$

or as a **column-vector**:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Question 12 *Consider the system of equations*

$$\begin{aligned} 3x + 5y + 2z &= 2 \\ 2x - 5y + 6z &= 5 \\ 5x + 3y - 3z &= 3. \end{aligned}$$

Show that this system of equations can be written as the single equation below by carrying out the multiplication

$$\begin{bmatrix} 3 & 5 & 2 \\ 2 & -5 & 6 \\ 5 & 3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix}.$$

Question 13 *Consider the system of n equations in k unknowns*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nk}x_k &= b_n. \end{aligned}$$

Show that this system of equations can be written as the following single equation by carrying out the multiplication

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Writing a system of equations as a single equation using matrices is extremely powerful. This form for a system of equations is called the **matrix form.** The matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}$$

is called the **coefficient matrix** for this system of equations and the matrix or column-vector

$$B = \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

is called the **constant matrix** or sometimes the **constant vector** for this system of equations.

Although matrix multiplication has many of the properties we might expect, it lacks one property that we might expect. In most cases, it is *not commutative.* In fact, if A and B are two matrices, then there are many different possibilities:

- The product AB is defined but the product BA is not defined.
- Both products AB and BA are defined but they have different dimensions.
- Both products AB and BA are defined and they have the same dimensions but they are different.
- Both products AB and BA are defined and they are the same.

Question 14 *Give examples showing that each of the possibilities above can occur.*

Fortunately many of the properties we have come to expect do carry over to matrix multiplication.

- If A is an $(n \times k)$-matrix, B is a $(k \times p)$-matrix, and C is a $(p \times q)$-matrix, then,

  $$A(BC) = (AB)C.$$

  We say that matrix multiplication is **associative**.

- If A is an $(n \times k)$-matrix and B and C are $(k \times p)$-matrices, then,

  $$A(B+C) = (AB) + (AC).$$

  If A and B are $(n \times k)$-matrices and C is a $(k \times p)$-matrix, then,

  $$(A+B)C = (AC) + (BC).$$

  We say that matrix multiplication is **distributive**.

- If A is an $(n \times k)$-matrix, then,

  $$AI_k = A \text{ and } I_n A = A.$$

- If A is an $(n \times k)$-matrix, B is a $(k \times p)$-matrix, and c is a real number, then,

  $$c(AB) = (cA)B = A(cB).$$

Question 15 *Illustrate each of the properties above with an example.*

3.5 The Inverse of a Matrix

Figure 3.21: A photograph of Washington DC from Google Earth

The world is awash in digital imagery. Millions of people post photographs on the Web. Anyone with Internet access can obtain those photographs as well as aerial photographs like the one shown in Figure 3.21 from Google Earth. We are interested not just in the photographs themselves but in the reality that was photographed – for example, we are less interested in the fact that two features shown on a photograph are six centimeters apart on the photograph than in the fact that they are 2 kilometers apart on the ground.

We can measure the location of features in a photograph directly using a ruler but we need to translate its coordinates on the photograph to its coordinates on the ground. Matrices and vectors are exactly the tools we need to transform photograph coordinates to ground coordinates or ground coordinates to photograph coordinates. In this section, we develop some of the ideas involved for features that are on the ground. We continue to develop these ideas later in this chapter.

We keep track of the location of a particular feature in two ways – first, its location on the ground and, second, its location on the photograph. On the ground we pick a point as the origin and measure every other point by its relationship with the origin. For example, in Washington, D.C. we might choose the Washington Monument as the origin.

When we look at an aerial photograph, we usually choose the center of the photograph as the origin and describe points in terms of how many centimeters they are above or below and left or right of this origin. Figure 3.22 on page 265 shows coordinate axes added to Figure 3.21. Notice the origin is in the center of the photograph.

Figure 3.22: Axes added to a photograph of Washington D.C. from Google Earth

In summary, if we are looking at a particular object, we can describe its location using a vector, $\vec{x} = \langle x_1, x_2 \rangle$, that corresponds to its location on the ground or using a vector, $\vec{y} = \langle y_1, y_2 \rangle$, that corresponds to the location of its image on the photograph – for example,

- My favorite mountain is located at the point $(10, 20)$ – 10 kilometers east and 20 kilometers north of City Hall in Smithtown.
- My favorite mountain is located 6 centimeters below and 4 centimeters to the right of the center of a particular aerial photograph of Smith County.

Because troops on the ground want to use the information from aerial photographs, we need to be able to translate back-and-forth between these two descriptions. The way in which this translation is done depends on the location and position of the camera when it took the picture and the lens that was used. We will be working with photographs that were taken by a camera pointing directly downward toward a point on the ground. Using geometry, one can show that in this situation the translation can always be described by a pair of equations,

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + b_1 \\ y_2 &= a_{21}x_1 + a_{22}x_2 + b_2 \end{aligned}$$

or, using matrices and vectors,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

or, more compactly,

$$\vec{y} = A\vec{x} + \vec{b}$$

where we write the vectors $\vec{x}$, $\vec{y}$, and $\vec{b}$ as column-vectors –

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Because the coordinates of the origin on the ground are $\vec{x} =< 0, 0 >$, the vector $\vec{b}$ is the location on the photograph of the origin on the ground. That is,

$$A \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

In Section 3.6 we will see how to determine the entries in the matrix A. We say that the **transformation** $\vec{y} = A\vec{x} + \vec{b}$ enables us to translate ground coordinates into photograph coordinates.

Example 1 *An analyst is working from an aerial photograph of a mountainous area near the Afghan-Pakistan border. An operative on the ground measures locations on the ground in kilometers relative to a distinctive rock that is easily recognizable from the air. The analyst measures locations on the photograph in centimeters relative to the center of the photograph. The transformation from coordinates, $\vec{x}$, on the ground as measured by the operative to coordinates, $\vec{y}$, on the photograph as measured by the analyst is given by*

$$\vec{y} = A\vec{x} + \vec{b}$$

where

$$\vec{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad A = \begin{bmatrix} 2.5 & -2.5 \\ 2.5 & 2.5 \end{bmatrix}.$$

For example, the distinctive rock that the operative is using as his origin is located on the photograph at the point

$$\vec{y} = \begin{bmatrix} 2.5 & -2.5 \\ 2.5 & 2.5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Question 1 *The operative on the ground reports suspicious activity at a cave located at a point 2.5 kilometers north and 1.6 kilometers east of the distinctive rock. Where is this cave on the analyst's photograph?*

Question 2 *The operative has provided a list of locations that he considers suspicious. Translate these locations into photograph coordinates.*

a. $\langle -2, 3 \rangle$ *or, written as a column-vector,* $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

b. $\langle 3, -5 \rangle$.

c. $\langle 7, -2 \rangle$.

Notice that we use a variety of different notations to represent variations of the same basic idea. For example, a point, $(3, 4)$, might be represented by a vector, $\langle 3, 4 \rangle$, or a row-vector, $[3, 4]$, or a column-vector,

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Now, suppose our analyst has spotted a point on the photograph that he believes is suspicious. He wants to describe it to the operative on the ground. This problem is the "inverse" or reverse of the problems we discussed above. Now we want to translate from photograph coordinates to ground coordinates. Given our formula

$$\vec{y} = A\vec{x} + \vec{b},$$

we want to find a formula that enables us to compute $\vec{x}$ if we know $\vec{y}$.

This problem looks very much like problems we routinely solve involving numbers. If we have an equation,

$$y = 2x + 3$$

we first subtract 3 from both sides to get

$$y - 3 = 2x$$

and then multiply both sides by $\frac{1}{2}$ to get

$$\left(\frac{1}{2}\right)(y - 3) = x.$$

This works for any equation of the form

$$y = mx + b.$$

A little algebra shows that

$$x = \left(\frac{1}{m}\right)(y - b).$$

This enables us to determine x given a value for y.

We often say that the function

$$x = \left(\frac{1}{m}\right)(y - b)$$

is the inverse of the function

$$y = mx + b.$$

Our goal is to find an inverse for the function, or transformation,

$$\vec{y} = A\vec{x} + \vec{b}.$$

We can start by trying to mimic the algebra we used above to find the inverse of the function $y = mx + b$. We begin with

$$\vec{y} = A\vec{x} + \vec{b}$$

and we subtract $\vec{b}$ from both sides to get

$$\vec{y} - \vec{b} = A\vec{x}$$

but now for the next step we need something like $\frac{1}{A}$. Unfortunately, however, $\frac{1}{A}$ doesn't make sense because we can't divide the real number 1 by a matrix. If "something" like $\frac{1}{A}$ exists then we call it the **inverse** of the matrix A. We do not use the notation $\frac{1}{A}$ for the inverse of a matrix. We use A^{-1} instead. This terminology mirrors the terminology we use for numbers – the number $\frac{1}{2}$ is called the multiplicative inverse of the number 2 because

$$\frac{1}{2} \times 2 = 1.$$

and we sometimes write 2^{-1} instead of $\frac{1}{2}$.

The only matrices that can possibly have inverses are square matrices – that is matrices with the same number of rows and columns. Even if a matrix is square, it might not have an inverse. We use the following terminology.

Definition 1 *Suppose that A is an $(n \times n)$-matrix, or a square matrix. If there is another $(n \times n)$-matrix B such that*

$$AB = BA = I,$$

then we say that the matrix A is **invertible**, *or* **nonsingular**, *and call the matrix B the* **inverse** *of the matrix A. Recall that I represents the identity matrix. We write*

$$B = A^{-1}.$$

The words "invertible" and "nonsingular" are synonyms. If a matrix does not have an inverse then we say it is "noninvertible" or "singular."

Question 3 *Show that the matrix*

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

is invertible and that its inverse is the matrix

$$B = \begin{bmatrix} \frac{1}{2} & 0 \\ \\ 0 & \frac{1}{3} \end{bmatrix}.$$

Question 4 *Show that the matrix*

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

is invertible and its inverse is the matrix

$$\begin{bmatrix} -2 & 1 \\ \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$

Question 5 *Recall that we said a matrix B was the inverse of a matrix A if $AB = BA = I$. We said that a matrix could have an inverse only if it was square. Explain why a matrix that is not square cannot have an inverse.*

Question 6 *Consider a matrix of the form*

$$A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}.$$

Show that this matrix has an inverse and find the inverse. Do you need to make any assumptions to show that this matrix has an inverse?

Question 7 *Consider a matrix of the form*

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}.$$

Show that this matrix has an inverse and find the inverse. Do you need to make any assumptions to show that this matrix has an inverse?

Question 8 *Suppose that A is a diagonal matrix – that is, that the only nonzero entries in A are on the diagonal.*

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & a_{nn} \end{bmatrix}$$

Show that if all of the diagonal entries in A are nonzero then A is invertible and find its inverse. Why must we assume that all the diagonal entries are nonzero?

Question 9 *Suppose that*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is a (2×2)-matrix and that

$$a_{11}a_{22} - a_{12}a_{21} \neq 0.$$

Show that the matrix

$$B = \begin{bmatrix} \frac{a_{22}}{a_{11}a_{22}-a_{12}a_{21}} & -\frac{a_{12}}{a_{11}a_{22}-a_{12}a_{21}} \\ -\frac{a_{21}}{a_{11}a_{22}-a_{12}a_{21}} & \frac{a_{11}}{a_{11}a_{22}-a_{12}a_{21}} \end{bmatrix}.$$

is the inverse of A. This formula for the inverse of a (2×2) is called **Cramer's Rule.**

Question 10 *Why do we have to assume that* $a_{11}a_{22} - a_{12}a_{21} \neq 0$*?*

Definition 2 *The quantity*

$$a_{11}a_{22} - a_{12}a_{21}$$

is called the **determinant** *of* A*.*

As we have seen, some square matrices have inverses and others do not. For example, a (2×2)-matrix has an inverse if and only if its determinant is nonzero. Recall that if a square matrix does not have an inverse we say that it is **noninvertible**, or **singular.**

Question 11 *Does the matrix*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

have an inverse? If so, find its inverse.

Question 12 *Does the matrix*

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

have an inverse? If so, find its inverse.

Question 13 *Does the matrix*

$$A = \begin{bmatrix} 2 & 6 \\ 3 & 7 \end{bmatrix}$$

have an inverse? If so, find its inverse.

Now, we can solve the problem that started this discussion of inverses. We begin with the formula

$$\vec{y} = A\vec{x} + \vec{b}$$

and we subtract $\vec{b}$ from both sides to get

$$A\vec{x} = \vec{y} - \vec{b}.$$

Now, *if the matrix A is invertible*, we multiply both sides of this equation by A^{-1} to get

$$\begin{aligned} A^{-1}A\vec{x} &= A^{-1}(\vec{y} - \vec{b}) \\ I\vec{x} &= A^{-1}(\vec{y} - \vec{b}) \\ \vec{x} &= A^{-1}(\vec{y} - \vec{b}) \end{aligned}$$

and this formula is exactly what we need to transform coordinates, $\vec{y}$, on the aerial photograph to coordinates, $\vec{x}$, on the ground.

The following example continues Example 1.

Example 2 *An analyst is working from an aerial photograph of a mountainous area near the Afghan-Pakistan border. An operative on the ground measures locations on the ground in kilometers relative to a distinctive rock that is easily recognizable from the air. The analyst measures locations on the photograph in centimeters relative to the center of the photograph. The transformation from coordinates, $\vec{x}$, on the ground as measured by the operative to coordinates, $\vec{y}$, on the photograph as measured by the analyst is given by*

$$\vec{y} = A\vec{x} + \vec{b}$$

where

$$\vec{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad A = \begin{bmatrix} 2.5 & -2.5 \\ 2.5 & 2.5 \end{bmatrix}.$$

The analyst spots a particular feature on the photograph that he believes might be an IED. It is located on the photograph at the point $\vec{y} = \langle 5, 6 \rangle$. He wants to tell the operative on the ground where it is located on the ground.

Using the work above we know that the transformation from photograph coordinates $\vec{y}$ to ground coordinates $\vec{x}$ is given by

$$\vec{x} = A^{-1}(\vec{y} - \vec{b}).$$

Using Cramer's Rule we see that

$$A^{-1} = \begin{bmatrix} 2.5 & -2.5 \\ 2.5 & 2.5 \end{bmatrix}^{-1} = \begin{bmatrix} 0.2 & 0.2 \\ -0.2 & 0.2 \end{bmatrix}.$$

So the possible IED is located on the ground at the point

$$\begin{bmatrix} 0.2 & 0.2 \\ -0.2 & 0.2 \end{bmatrix} \left(\begin{bmatrix} 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right) = \begin{bmatrix} 0.2 & 0.2 \\ -0.2 & 0.2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0.0 \end{bmatrix}.$$

Question 14 *The analyst spots suspicious activity at a cave whose photograph coordinates are* $\langle 5.25, 14.25 \rangle$. *Where is this cave on the ground? Compare this question with Question 1.*

Question 15 *The analyst has a list of the photograph coordinates of locations that he considers suspicious. Transform the photograph coordinates into ground coordinates.*

a. $\langle -9.5, 6.5 \rangle$.

b. $\langle 23, -1 \rangle$.

c. $\langle 25.5, 16.5 \rangle$.

Compare this question with Question 2.

Question 16 *The analyst has identified three points on the aerial photograph that are suspicious. He wants to send the operative the ground coordinates of those three points. Their coordinates on the aerial photograph are given below. Find the corresponding ground coordinates.*

a. $\langle 3, 2 \rangle$.

b. $\langle -4, 7 \rangle$.

c. $\langle -7, -8 \rangle$.

Question 17 *Look back at Question 4 in Section 3.3. Can you answer that question now using the work we've done in this section?*

Question 18 *You might wonder what would happen if the matrix A was not invertible. Explore this question by looking at the equation*

$$\vec{y} = A\vec{x} + \vec{b}$$

with

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

a. *First show that A is not invertible.*

b. *Try to find $\vec{x}$ given*

$$\vec{y} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

c. *Try to find $\vec{x}$ given*

$$\vec{y} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$$

d. *From your experience with photographs do you believe that in a real problem the matrix A can be non-invertible?*

3.6 Solving Systems of Linear Equations

In Section 3.5 we discussed how matrices and vectors could be used for translating back-and-forth between ground coordinates, $\vec{x}$, and coordinates, $\vec{y}$, on a photograph using a transformation of the form

$$\vec{y} = A\vec{x} + \vec{b}.$$

The vector $\vec{b}$ was easy to determine. It was just the location on the photograph (in photograph coordinates) of the origin on the ground. We promised to show how to determine the matrix A in this section. We will fulfill that promise but, first, we begin with a simpler example and develop skills that will enable us to solve a large number of important problems including the problem of determining the matrix A from Section 3.5.

Example 1 *You are preparing 10 gallons of a rust retarding spray made up of 4 gallons of agent P and 6 gallons of agent Q. You can order the spray to be custom made but it would be very expensive and there are two much cheaper pre-mixed sprays that can be bought in bulk. One is called mixture X and has 25% agent P and 75% agent Q. The other is called mixture Y and has 80% agent P and 20% agent Q. Can you save money by buying these pre-mixed sprays? If so, how many gallons of each should you buy?*

As we discuss this problem, it is worthwhile to use the steps that that were discussed at the end of Section 3.3.

- Identify the quantities that are known.
- Identify the quantities that are unknown.
- Identify the relationships between the unknown quantities and the known quantities.
- Identify additional quantities or data whose determination would be useful.

We follow these four steps explicitly below.

- This problem provides us with a great deal of known information. We know that we want "10 gallons of a rust retarding spray made up of 4 gallons of agent P and 6 gallons of agent Q." We also know that "mixture X has 25% agent P and 75% agent Q" and that "mixture Y has 80% agent P and 20% agent Q."

- We need to determine how many gallons of mixture X and how many gallons of mixture Y to buy. We use the notation x for the unknown number of gallons of mixture X we should buy and y for the unknown number of gallons of mixture Y we should buy.

- If we purchase x gallons of mixture X and y gallons of mixture Y, then the amounts of agents P and Q that we get are given by the equations

$$\begin{aligned} p &= 0.25x + 0.80y \\ q &= 0.75x + 0.20y \end{aligned}$$

  where p represents the number of gallons of agent P and q represents the number of gallons of agent Q. Because we want 4 gallons of agent P and 6 gallons of agent Q we need to solve the equations

$$\begin{aligned} 0.25x + 0.80y &= 4 \\ 0.75x + 0.20y &= 6 \end{aligned}$$

  These equations are the key relationships between our unknown and known quantities. We can write these equations very compactly using matrices and vectors

$$A\vec{x} = \vec{b}$$

  where A is the matrix

$$A = \begin{bmatrix} 0.25 & 0.80 \\ 0.75 & 0.20 \end{bmatrix}$$

  and $\vec{x}$ and $\vec{b}$ are the vectors

$$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

- Although there are additional quantities of interest, for example, the prices of the two mixtures, this particular problem asks only for the values of x and y.

Now, if the matrix A is invertible, we can solve this system of equations using its inverse as follows.

$$\begin{aligned} A\vec{x} &= \vec{b} \\ A^{-1}A\vec{x} &= A^{-1}\vec{b} \\ I\vec{x} &= A^{-1}\vec{b} \\ \vec{x} &= A^{-1}\vec{b}. \end{aligned}$$

Question 1 *Use Cramer's Rule from page 271, Section 3.5, to find the inverse of the matrix A in Example 1 above. Then, complete the solution of the problem in Example 1.*

If you have solved these kinds of equations before, this may seem like a new and unnecessarily complicated way to do something that is easy. This new way, however, enables us to work with systems of hundreds of equations with hundreds of unknowns as easily as systems of two equations with two unknowns.

Definition 1 *A system of n linear equations with k unknowns is a system of equations of the form*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nk}x_k &= b_n. \end{aligned}$$

Using vectors and matrices we can write a system of n linear equations in k unknowns more compactly as follows:

$$A\vec{x} = \vec{b}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

The technique we developed to solve the example can be applied in this more general situation.

Theorem 4 *Suppose that we have a system of n linear equations in n unknowns.*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{12}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

Notice that the coefficient matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

is a square matrix. If this matrix has an inverse, then the solution of this system of equations is given by

$$\vec{x} = A^{-1}\vec{b}$$

where

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Proof

$$\begin{aligned} A\vec{x} &= \vec{b} \\ A^{-1}A\vec{x} &= A^{-1}\vec{b} \\ I\vec{x} &= A^{-1}\vec{b} \\ \vec{x} &= A^{-1}\vec{b} \quad \blacksquare \end{aligned}$$

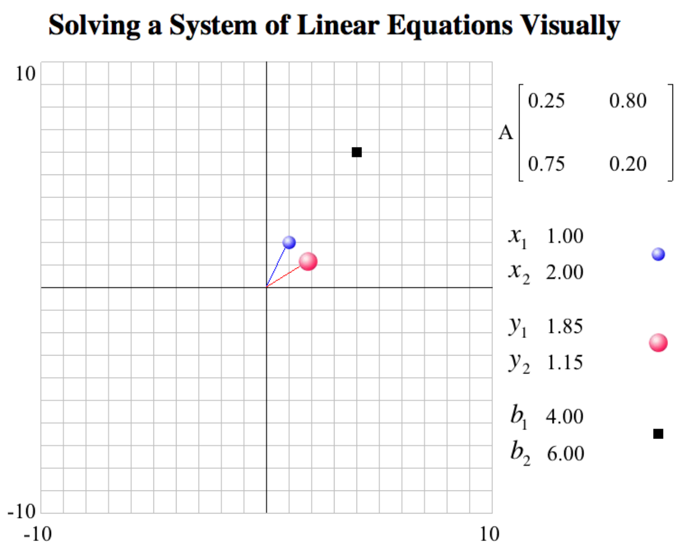


Figure 3.23: Solving a system of linear equations visually

We can gain some insight into these kinds of problems by looking at them visually – see Figure 3.23 and click here[6] to open a new window with a live copy of Figure 3.23. The small blue dot represents a possible purchase that you might make of the two mixtures from Example 1. In Figure 3.23 the blue dot is at the point $(1, 2)$ representing a purchase of one gallon of mixture X and 2 gallons of mixture Y. The larger red dot is at the point that represents the amounts of agents P and Q in this purchase – 1.85 gallons of agent P and 1.15 gallons of agent Q.

Notice that Figure 3.23 uses different notation than was used in Example 1. You can translate between the two notations as follows.

[6]http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/solve-visually.html

- The two unknowns are the numbers of gallons of mixture X and mixture Y to be purchased. In Example 1 we used the notation x and y for these two unknowns. In Figure 3.23 we use the notation x_1 and x_2 for these same unknowns.

- In Example 1 we used the notation p and q for the number of gallons of agents P and Q that we receive if we purchase x gallons of mixture x and y gallons of mixture Y. in Figure 3.23 we use the notation y_1 and y_2 for these same quantities.

- Our goal is 4 gallons of agent P and 6 gallons of Agent Q. We used the notation $b_1 = 4$ and $b_2 = 6$ for this goal in both Example 1 and Figure 3.23.

In the live diagram that you just opened, you can drag the blue dot around, trying out different possible purchases graphically. As you move the blue dot, the red dot will also move, indicating the resulting amounts of agents P and Q of each possible purchase of mixtures X and Y. Try to hit the target result – four gallons of agent P and 6 gallons of agent Q, represented by the black square. You should get roughly the same answer we obtained earlier. You probably won't be able to get exactly the same answer because of the limited resolution of the screen as seen in Figure 3.24.

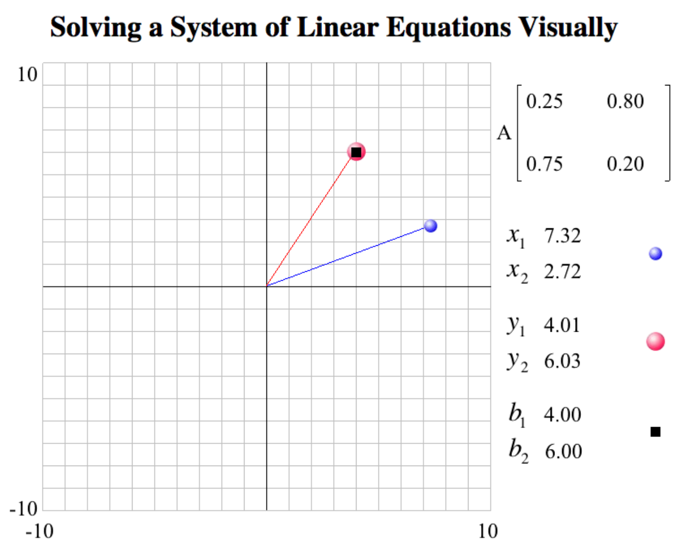


Figure 3.24: Solving a system of linear equations visually with limited resolution

The words "input" and "output" are often used in this situation. Your "input" is your order or purchase – x gallons of mixture X and y gallons of mixture Y. The "output" of your input or purchase is the resulting amounts of agents P and Q.

Question 2 *Use Theorem 4 on page 279 and Cramer's Rule to solve the following system of linear equations.*

$$\begin{aligned} 2x_1 + 3x_2 &= 5 \\ 3x_1 - x_2 &= 2 \end{aligned}$$

Question 3 *Write the following system of equations in the form $A\vec{x} = \vec{b}$.*

$$\begin{aligned} x_1 - 2x_2 + x_3 - 4x_4 + x_5 &= 10 \\ 2x_1 + 3x_2 - x_3 - x_4 + x_5 &= 8 \\ -x_1 + x_2 - 2x_3 + x_4 - x_5 &= 6 \\ 3x_1 - 3x_2 + x_3 + x_4 + x_5 &= 12 \\ -2x_1 - 3x_2 - x_3 + x_4 - 4x_5 &= -3 \end{aligned}$$

Question 4 *Write the following system of equations in the form $A\vec{x} = \vec{b}$.*

$$\begin{aligned} x_1 - 2x_2 - 4x_4 + x_5 &= 10 \\ 2x_1 + 3x_2 - x_3 - x_4 + x_5 &= 8 \\ -x_1 + x_2 - 2x_3 + x_4 &= 6 \\ -3x_2 + x_3 + x_4 + x_5 &= 12 \\ -2x_1 - 3x_2 - x_3 + x_4 - 4x_5 &= -3 \end{aligned}$$

Click here[7] to open the live diagram we used earlier. You can change the entries for the matrix A and for the vector $\vec{b}$ by clicking on the entry that you want to change and then editing it in the usual way. Use this live diagram for the questions below.

[7] http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/solve-visually.html

Question 5 *Try to solve graphically the system of equations*

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 7. \end{aligned}$$

Describe what happens. Then try to solve this same system of equations using Theorem 4 on page 279. What happens?

Question 6 *Try to solve graphically the system of equations*

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 8. \end{aligned}$$

Describe what happens. Then try to solve this same system of equations using Theorem 4 on page 279. What happens?

Question 7 *Try to solve graphically the system of equations*

$$\begin{aligned} 4x_1 - 3x_2 &= 7 \\ 2x_1 + x_2 &= 11. \end{aligned}$$

Question 8 *Try to solve graphically the system of equations*

$$\begin{aligned} 4x_1 - 3x_2 &= 3 \\ 2x_1 + x_2 &= 2. \end{aligned}$$

We are now prepared for the promised solution of our problem from Section 3.5. We have some information from an operative on the ground and we have some aerial imagery and we want to be able to convert back-and-forth between ground coordinates and coordinates on a photograph. Because our photograph was made from a point directly overhead the area of interest with the camera pointing straight downward toward the ground, we know that this conversion can be based on a transformation of the form

$$\vec{y} = A\vec{x} + \vec{b}$$

where $\vec{y}$ represents coordinates on the photograph and $\vec{x}$ represents coordinates on the ground. In Section 3.5, we saw that the vector $\vec{b}$ was the location on the photograph of the landmark used as the origin for the ground coordinates. In order to find the entries in the matrix A, we need two additional landmarks that appear in the photograph and whose ground coordinates are known. We begin with an example.

Example 2 *You are part of a team of analysts supporting a local group based in an area with high insurgent activity and in which it is difficult to distinguish friend from foe. The local group describes its territory using a large monument as their origin. You have a detailed, large high resolution photograph of the area in which this group is operating. The monument is easily distinguishable on the photograph. You have placed a transparent grid on top of the photograph and identify points on the photograph using grid coordinates*

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

On the photograph the grid coordinates of the monument are

$$\vec{y} = \begin{bmatrix} 1.34 \\ 2.42 \end{bmatrix}.$$

Because the local group is using this monument as its origin, the local group describes the monument as having ground coordinates

$$\vec{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

There are two bridges crossing a large river that flows through the area. One bridge is a railroad bridge and is located 2.3 kilometers east and 1.2 kilometers north of the monument. Thus, its ground coordinates are

$$\vec{x} = \begin{bmatrix} 2.3 \\ 1.2 \end{bmatrix}.$$

This rail bridge is easily identifiable on your photograph and its grid coordinates on the photograph are

$$\vec{y} = \begin{bmatrix} 5.079 \\ 3.499 \end{bmatrix}.$$

The second bridge carries a road and is located 1.6 kilometers west and 1.9 kilometers north of the monument. Thus, its ground coordinates are

$$\vec{x} = \begin{bmatrix} -1.6 \\ 1.9 \end{bmatrix}.$$

This road bridge is easily identifiable on your photograph and its grid coordinates on the photograph are

$$\vec{y} = \begin{bmatrix} -0.446 \\ 5.690 \end{bmatrix}.$$

Based on this information, we need to determine how to translate from ground coordinates to grid coordinates on the photograph using a formula of the form

$$\vec{y} = A\vec{x} + \vec{b}.$$

That is, we must determine the vector, $\vec{b}$*, and the matrix,* A*. Note that once we have solved this problem we can also translate from grid coordinates on the photograph to local ground coordinates using the formula*

$$\vec{x} = A^{-1}(\vec{y} - \vec{b}).$$

We summarize the available information in the table below.

Landmark	**Ground**		**Photograph**	
	x_1	x_2	y_1	y_2
Monument	0	0	1.34	2.42
Rail bridge	2.3	1.2	5.079	3.499
Road bridge	-1.6	1.9	-0.446	5.690

Table 3.13: Known information

In Section 3.5, we saw that the vector $\vec{b}$ was the location on the photograph of the landmark used as the origin for the ground coordinates. Thus,

$$\vec{b} = \begin{bmatrix} 1.34 \\ 2.42 \end{bmatrix}.$$

Our job is to find the entries in the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

We refer back to this example in the following paragraphs. In general, here is our situation.

- We know the vector $\vec{b}$. In our example

$$\vec{b} = \begin{bmatrix} 1.34 \\ 2.42 \end{bmatrix}.$$

- We know the ground coordinates $\vec{p}$ and $\vec{q}$ of two additional landmarks. We can locate these landmarks on the photograph and we measure their (grid) coordinates on the photograph. Call these measured (grid) coordinates $\vec{u}$ for the landmark at $\vec{p}$ and $\vec{v}$ for the landmark at $\vec{q}$.

  In our example,

$$\vec{p} = \begin{bmatrix} 2.3 \\ 1.2 \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} 5.079 \\ 3.499 \end{bmatrix}; \quad \vec{q} = \begin{bmatrix} -1.6 \\ 1.9 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} -0.446 \\ 5.690 \end{bmatrix}.$$

- We have four unknowns – the entries: a_{11}, a_{12}, a_{21}, and a_{22} in the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

We need to adjust our thinking a bit here. Notice there are four unknowns – the entries: a_{11}, a_{12}, a_{21}, and a_{22} in the matrix A. We usually use letters like x and y for unknowns but in this case we are using the letter a because our unknowns also happen to be elements of the matrix A.

Now, we are ready to look at our system of equations. Using matrix and vector operations we get two equations

$$\begin{aligned}\vec{u} &= A\vec{p} + \vec{b} \\ \vec{v} &= A\vec{q} + \vec{b}\end{aligned}$$

If we carry out the matrix operations, we obtain

$$\begin{aligned}u_1 &= a_{11}p_1 + a_{12}p_2 + b_1 \\ u_2 &= a_{21}p_1 + a_{22}p_2 + b_2\end{aligned}$$

from the first equation and

$$\begin{aligned}v_1 &= a_{11}q_1 + a_{12}q_2 + b_1 \\ v_2 &= a_{21}q_1 + a_{22}q_2 + b_2\end{aligned}$$

from the second equation. Note again that the unknowns are the entries: a_{11}, a_{12}, a_{21}, and a_{22} in the matrix A. That is, our four equations with the unknowns marked are

$$\begin{aligned}u_1 &= \underbrace{a_{11}}_{\text{unknown}} p_1 + \underbrace{a_{12}}_{\text{unknown}} p_2 + b_1 \\ u_2 &= \underbrace{a_{21}}_{\text{unknown}} p_1 + \underbrace{a_{22}}_{\text{unknown}} p_2 + b_2 \\ v_1 &= \underbrace{a_{11}}_{\text{unknown}} q_1 + \underbrace{a_{12}}_{\text{unknown}} q_2 + b_1 \\ v_2 &= \underbrace{a_{21}}_{\text{unknown}} q_1 + \underbrace{a_{12}}_{\text{unknown}} q_2 + b_2.\end{aligned}$$

All the other symbols in the system of equations above represent known quantities.

With a little rearranging, we obtain

$$\begin{aligned}
\underbrace{a_{11}}_{\text{unknown}} p_1 + \underbrace{a_{12}}_{\text{unknown}} p_2 &= u_1 - b_1 \\
\underbrace{a_{21}}_{\text{unknown}} p_1 + \underbrace{a_{22}}_{\text{unknown}} p_2 &= u_2 - b_2 \\
\underbrace{a_{11}}_{\text{unknown}} q_1 + \underbrace{a_{12}}_{\text{unknown}} q_2 &= v_1 - b_1 \\
\underbrace{a_{21}}_{\text{unknown}} q_1 + \underbrace{a_{12}}_{\text{unknown}} q_2 &= v_2 - b_2
\end{aligned}$$

or, using matrix and vector notation,

$$\underbrace{\begin{bmatrix} p_1 & p_2 & 0 & 0 \\ 0 & 0 & p_1 & p_2 \\ q_1 & q_2 & 0 & 0 \\ 0 & 0 & q_1 & q_2 \end{bmatrix}}_{\text{known}} \underbrace{\begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix}}_{\text{unknown}} = \underbrace{\begin{bmatrix} u_1 - b_1 \\ u_2 - b_2 \\ v_1 - b_1 \\ v_2 - b_2 \end{bmatrix}}_{\text{known}}.$$

Hence,

$$\begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & 0 & 0 \\ 0 & 0 & p_1 & p_2 \\ q_1 & q_2 & 0 & 0 \\ 0 & 0 & q_1 & q_2 \end{bmatrix}^{-1} \begin{bmatrix} u_1 - b_1 \\ u_2 - b_2 \\ v_1 - b_1 \\ v_2 - b_2 \end{bmatrix},$$

and if the matrix

$$W = \begin{bmatrix} p_1 & p_2 & 0 & 0 \\ 0 & 0 & p_1 & p_2 \\ q_1 & q_2 & 0 & 0 \\ 0 & 0 & q_1 & q_2 \end{bmatrix}$$

is invertible, we have fulfilled our promise from Section 3.5.

If we apply this procedure to our example we obtain[8]

$$\begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 2.3 & 1.2 & 0 & 0 \\ 0 & 0 & 2.3 & 1.2 \\ -1.6 & 1.9 & 0 & 0 \\ 0 & 0 & -1.6 & 1.9 \end{bmatrix}^{-1} \begin{bmatrix} 3.739 \\ 1.079 \\ -1.806 \\ 3.270 \end{bmatrix} = \begin{bmatrix} 1.473970 \\ 0.290715 \\ -0.297917 \\ 1.470170 \end{bmatrix}$$

and, thus, the matrix A is

$$A = \begin{bmatrix} 1.473970 & 0.290715 \\ -0.297917 & 1.470170 \end{bmatrix}.$$

[8]We used *Mathematica* to find the inverse of the matrix W. The code is **Inverse[W] // MatrixForm**

3.7 Row Reduction - Solving Systems of Linear Equations

In Questions 5 and 6 in Section 3.6, we looked at the following two systems of linear equations

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 7 \end{aligned}$$

and

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 8 \end{aligned}$$

and tried to solve them as discussed in the last section using the inverse of the coefficient matrix

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}.$$

You should have discovered that this coefficient matrix has no inverse. Thus, the method we developed in the previous section fails. What does this mean for the two systems of equations? In the following pages we develop a way to analyze systems of equations like these and answer this question.

We can gain some insight into this question by looking at the basic problem of solving a system of two linear equations in two unknowns graphically. We use the two systems of equations above as examples.

Example 1 *We begin by writing the system of equations from Question 5 in Section 3.6,*

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 7, \end{aligned}$$

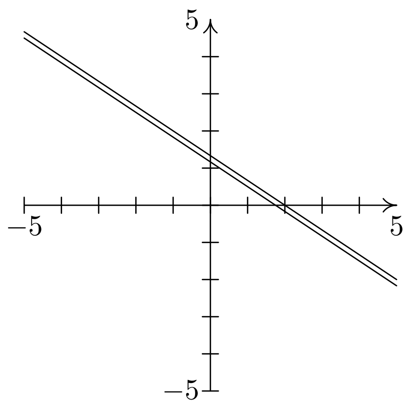
using the letters x and y instead of x_1 and x_2 for our two unknowns.

$$\begin{aligned} 2x + 3y &= 4 \\ 4x + 6y &= 7 \end{aligned}$$

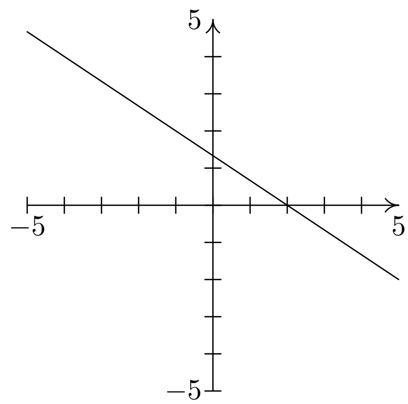
With a bit of algebra we can rewrite these as

$$
\begin{aligned}
y &= -\frac{2}{3}x + \frac{4}{3} \\
y &= -\frac{2}{3}x + \frac{7}{6}
\end{aligned}
$$

These are the equations of two straight lines. The solution(s) of these equations are points where the two straight lines intersect. The two lines are parallel since their slopes are the same. That means that the two lines either lie on top of each other or they never meet. Because their y-intercepts are different the two lines are parallel and never intersect. See the left side of Figure 3.25. Thus, this system of equations has no solutions.



$2x + 3y = 4$
$4x + 6y = 7$



$2x + 3y = 4$
$4x + 6y = 8$

Figure 3.25: Graphs for two systems of equations

Example 2 *Next we write the system of equations from Question 6 in Section 3.6,*

$$
\begin{aligned}
2x_1 + 3x_2 &= 4 \\
4x_1 + 6x_2 &= 8,
\end{aligned}
$$

using the letters x and y instead of x_1 and x_2 for our two unknowns.

$$\begin{aligned} 2x + 3y &= 4 \\ 4x + 6y &= 8 \end{aligned}$$

With a bit of algebra we can rewrite these as

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

These are the equations of two straight lines. The solution(s) of these equations are points where the two straight lines intersect. The two lines are parallel since their slopes are the same. That means that the two lines either lie on top of each other or they never meet. Because their y-intercepts are the same the two lines lie on top of each other. See the right side of Figure 3.25 on page 291. This system of equations has an infinite number of solutions.

The method we developed in the last two examples works for any system of two linear equations in two unknowns. Suppose that we have two linear equations in two unknowns

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2. \end{aligned}$$

With a bit of algebra we can rewrite these equations as

$$\begin{aligned} y &= -\left(\frac{a_{11}}{a_{12}}\right)x + \frac{b_1}{a_{12}} \\ y &= -\left(\frac{a_{21}}{a_{22}}\right)x + \frac{b_2}{a_{22}}. \end{aligned}$$

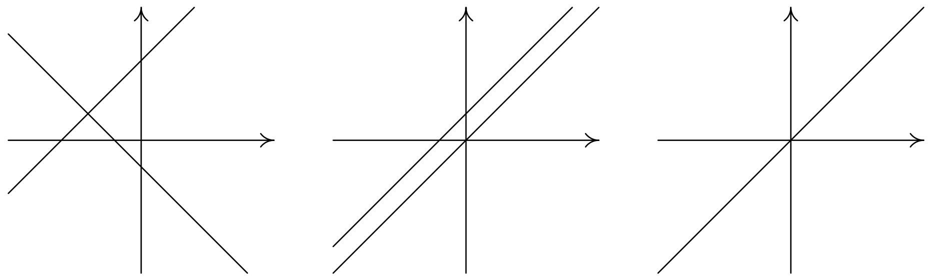


Figure 3.26: Three possibilities for two straight lines

Each of these equations is a linear equation whose graph is a straight line. Figure 3.26 shows three possibilities for a pair of straight lines:

- The two lines intersect in a single point. This point is the solution of the two equations.
- The two lines are parallel and do not intersect. In this case the two equations do not have a solution.
- The two lines are actually the same line. That is, the system of equations is redundant. In this case there are an infinite number of solutions. Any point on the single line is a solution.

Notice that the two lines

$$\begin{aligned} y &= -\left(\frac{a_{11}}{a_{12}}\right)x + \frac{b_1}{a_{12}} \\ y &= -\left(\frac{a_{21}}{a_{22}}\right)x + \frac{b_2}{a_{22}} \end{aligned}$$

are parallel if they have the same slope – that is, if and only if

$$\begin{aligned} -\left(\frac{a_{11}}{a_{12}}\right) &= -\left(\frac{a_{21}}{a_{22}}\right) \\ a_{11}a_{22} &= a_{21}a_{21} \\ a_{11}a_{22} - a_{12}a_{21} &= 0 \end{aligned}$$

Thus, we see again the importance of the determinant, $a_{11}a_{22} - a_{21}a_{21}$. The two lines are not paralllel and, thus, intersect in a unique solution if and only if the determinant is nonzero. Recall also that the coefficient matrix is singular, or noninvertible, if and only if its determinant is zero. In summary, when the coefficient matrix of a system of 2 equations in 2 unknowns is singular, or noninvertible, then there are two possibilities.

- The system has no solutions.
- The system has an infinite number of solutions.

Graphing equations helps to understand systems of two linear equations in two unknowns but it does not help for systems with more linear equations and more unknowns. In this section, we develop a more powerful method that works for systems with many linear equations and many unknowns. We illustrate this technique using the same systems of equations we have been discussing.

One way to solve the first system of equations,

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 7, \end{aligned}$$

is to multiply the first equation by 2. This gives us the pair of equations

$$\begin{aligned} 4x_1 + 6x_2 &= 8 \\ 4x_1 + 6x_2 &= 7. \end{aligned}$$

Now if we subtract the second equation from the first equation we get

$$0 = 1.$$

Of course, $0 \neq 1$, so the original two equations are contradictory. This means that the original two equations cannot have any solutions.

If we apply the same procedure to the second system of equations,

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 8, \end{aligned}$$

first multiplying the first equation by 2 to obtain

$$\begin{aligned} 4x_1 + 6x_2 &= 8 \\ 4x_1 + 6x_2 &= 8 \end{aligned}$$

and then subtracting the second equation from the first equation we get

$$0 = 0.$$

In this case, our two equations are not contradictory – they are redundant. The second equation, $4x_1 + 6x_2 = 8$, is just twice the first equation, $2x_1 + 3x_2 = 4$, and doesn't give us any new information.

In this situation, not only is there a solution, there are infinitely many solutions – for example,

$$\begin{aligned} x_1 &= 2 \\ x_2 &= 0 \end{aligned}$$

is one solution and

$$\begin{aligned} x_1 &= -1 \\ x_2 &= 2 \end{aligned}$$

is another.

We can determine any number of these infinite solutions by rewriting our single equation,

$$2x_1 + 3x_2 = 4,$$

and solving for x_2 as

$$x_2 = \frac{4}{3} - \frac{2}{3}x_1.$$

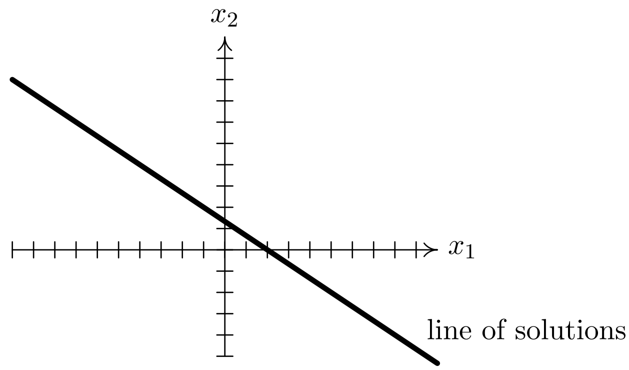


Figure 3.27: Infinitely many solutions

If x_1 is any number, we can use this equation to compute the corresponding value of x_2 that gives us another solution. For example, if $x_1 = 6$ then

$$\begin{aligned} x_2 &= \frac{4}{3} - \frac{2}{3}x_1 \\ x_2 &= \frac{4}{3} - \frac{2}{3}(6) = -\frac{8}{3} \end{aligned}$$

and $x_1 = 6, x_2 = -8/3$ is another solution.

The thick line in Figure 3.27 shows all the possible solutions for the original system of equations,

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 8. \end{aligned}$$

The Method of Row Reduction

The method of row reduction systematically applies a series of steps to change (or reduce) a system of equations like

$$\begin{aligned} 2x_1 + 3x_2 &= 8 \\ 3x_1 - x_2 &= 1 \end{aligned}$$

to a form like

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 2 \end{aligned}$$

Each step is the result of applying one of three basic operations to a system of equations. These basic operations are

- We can multiply any equation by a non-zero constant.
- We can interchange any two equations (change their order).
- We can add or subtract a multiple of any equation to any other equation.

These operations are called **elementary operations**.

Each of these operations changes a system of equations into a new system of equations that has exactly the same solutions.

Example 3 *Suppose we start with the system of equations*

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \quad (E1) \\ 2x_1 + 3x_2 + 4x_3 &= 5 \quad (E2) \\ x_1 - 2x_2 + 3x_3 &= 1 \quad (E3) \end{aligned}$$

The first operation allows us to multiply any equation by any nonzero constant. Like all elementary operations, this operation does not change the solutions of the system of equations. For example, if we multiply the second equation by 3 we would obtain

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 6x_1 + 9x_2 + 12x_3 &= 15 \\ x_1 - 2x_2 + 3x_3 &= 1. \end{aligned}$$

We use the notation $3E_2 \to E_2$ *as shorthand for this elementary operation. In words, this says "Replace Equation E2 with three times Equation E2."*

Example 4 *Suppose we start with the system of equations*

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \quad (E1) \\ 2x_1 + 3x_2 + 4x_3 &= 5 \quad (E2) \\ x_1 - 2x_2 + 3x_3 &= 1 \quad (E3) \end{aligned}$$

The second operation allows us to interchange any two equations. Like all elementary operations, this operation does not change the solutions of the system of equations. For example, if interchange the first and third equations we would obtain

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 1 \\ 2x_1 + 3x_2 + 4x_3 &= 5 \\ x_1 - x_2 + x_3 &= 2. \end{aligned}$$

We use the notation $E_1 \leftrightarrow E_3$ as shorthand for this elementary operation. In words, this says "Replace Equation E3 with Equation E1 and replace Equation E1 with Equation E3."

Example 5 *Suppose we start with the system of equations*

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \quad (E1) \\ 2x_1 + 3x_2 + 4x_3 &= 5 \quad (E2) \\ x_1 - 2x_2 + 3x_3 &= 1 \quad (E3) \end{aligned}$$

The third operation allows us to add any multiple of one equation to another. Like all elementary operations, this operation does not change the solutions of the system of equations. For example, if we add -2 times the first equation to the third equation we obtain

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 2x_1 + 3x_2 + 4x_3 &= 5 \\ -x_1 \qquad + x_3 &= -3 \end{aligned}$$

Note that this operation changes the third equation and leaves the first equation unchanged.

We use the notation $-2E_1 + E_3 \to E_3$ as shorthand for this elementary operation. In words, this says, "Replace Equation E3 with the sum of -2 times Equation E1 and Equation E3."

Recall that our goal is to use these elementary operations to change (or reduce) a system of equations like

$$\begin{aligned} 2x_1 + 3x_2 &= 8 \\ 3x_1 - x_2 &= 1 \end{aligned}$$

to a form like

$$\begin{aligned} x_1 &= \text{something} \\ x_2 &= \text{something.} \end{aligned}$$

For the system of equations

$$\begin{aligned} 2x_1 + 3x_2 &= 8 \\ 3x_1 - x_2 &= 1 \end{aligned}$$

we would use the following steps to reach our goal.

1. Start
$$\begin{aligned} 2x_1 + 3x_2 &= 8 \\ 3x_1 - x_2 &= 1. \end{aligned}$$

2. Because we want the first equation to begin with x_1, we multiply the first equation by $\frac{1}{2}$. ($\frac{1}{2}E_1 \to E_1$)
$$\begin{aligned} x_1 + \frac{3}{2}x_2 &= 4 \\ \\ 3x_1 - x_2 &= 1. \end{aligned}$$

3. Because we do not want the variable x_1 to appear in the second equation, we subtract three times the first equation from the second equation. ($E_2 - 3E_1 \to E_2$)

$$\begin{aligned} x_1 + \frac{3}{2}x_2 &= 4 \\ -\frac{11}{2}x_2 &= -11. \end{aligned}$$

4. Because we want the second equation to begin with x_2, we multiply the second equation by $-\frac{2}{11}$. ($-\frac{2}{11}E_2 \to E_2$)

$$\begin{aligned} x_1 + \frac{3}{2}x_2 &= 4 \\ x_2 &= 2. \end{aligned}$$

5. Because we do not want the variable x_2 to appear in the first equation, we subtract $\frac{3}{2}$ times the second equation from the first equation. ($E_1 - \frac{3}{2}E_2 \to E_1$)

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 2. \end{aligned}$$

These same steps can be written in a more compact form using matrices. We use "augmented matrices" to represent systems of equations. The left side of each augmented matrix represents the coefficients of the equations and the right side represents the constant term on the right side of each equation. For example, the pair of equations

$$\begin{aligned} 2x_1 + 3x_2 = 4 \\ 5x_1 + 6x_2 = 7 \end{aligned} \quad \text{is written as the augmented matrix} \quad \left[\begin{array}{cc|c} 2 & 3 & 4 \\ 5 & 6 & 7 \end{array}\right].$$

Many people, including us, draw a vertical bar between the two sides. In this compact form we call the elementary operations **elementary row operations.** Notice we use almost the same notation to abbreviate elementary row operations as we used earlier to abbreviate elementary (equation) operations. We just use R to abbreviate "row" instead of E to abbreviate "equation." In this compact form we write the same series of elementary operations we wrote earlier as –

1. $\left[\begin{array}{cc|c} 2 & 3 & 8 \\ 3 & -1 & 1 \end{array}\right].$

2. $\left[\begin{array}{cc|c} 1 & \frac{3}{2} & 4 \\ 3 & -1 & 1 \end{array}\right] \quad \frac{1}{2}R_1 \to R_1.$

3. $\left[\begin{array}{cc|c} 1 & \frac{3}{2} & 4 \\ 0 & -\frac{11}{2} & -11 \end{array}\right] \quad R_2 - 3R_1 \to R_2.$

4. $\left[\begin{array}{cc|c} 1 & \frac{3}{2} & 4 \\ 0 & 1 & 2 \end{array}\right] \quad -\frac{2}{11}R_2 \to R_2.$

5. $\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right] \quad R_1 - \frac{3}{2}R_2 \to R_1.$

Notice the last matrix above represents the system of equations

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 2. \end{aligned}$$

The method we have developed in this section is called **row reduction**. Although we can write this method using equations, we usually write it using augmented matrices as shown above.

Question 1 *Use the augmented matrix form of row reduction to solve the following systems of equations. Once you have your solution in augmented matrix form, convert the output back into a system of equations. If the system has one solution find that solution. If it has an infinite number of solutions find two example solutions.*

a. $\begin{aligned} x_1 - x_2 &= 0 \\ x_1 + x_2 &= 4 \end{aligned}$ **b.** $\begin{aligned} x_1 + x_2 &= 2 \\ 2x_1 + 2x_2 &= 5 \end{aligned}$

c. $\begin{aligned} 2x_1 + 3x_2 &= 3 \\ 4x_1 + 6x_2 &= 6 \end{aligned}$ **d.** $\begin{aligned} 2x_1 + 3x_2 &= 3 \\ 6x_1 + 9x_2 &= 6 \end{aligned}$

Question 2 *Use the augmented matrix form of row reduction to solve the following systems of equations. Once you have your solution in augmented matrix form, convert the output back into a system of equations. If the system has one solution find that solution. If it has an infinite number of solutions find two example solutions.*

a.
$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ x_1 + x_2 + x_3 &= 4 \\ 3x_1 - x_2 + 3x_3 &= 6 \end{aligned}$$

b.
$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ x_1 + x_2 + x_3 &= 4 \\ 3x_1 - x_2 + 3x_3 &= 8 \end{aligned}$$

c.
$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ 2x_1 + 2x_2 + 2x_3 &= 6 \\ 3x_1 + 3x_2 + 3x_3 &= 8 \end{aligned}$$

d.
$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ 2x_1 + 2x_2 + 2x_3 &= 6 \\ 3x_1 + 3x_2 + 3x_3 &= 9 \end{aligned}$$

Question 3 *One of the most important questions in homeland security is distinguishing rapidly, accurately, and non-intrusively between friends and foes. For example, in crowd situations there may be uniformed friends, undercover friends, innocent bystanders, and foes. We have some control over the attire of uniformed and undercover friends but no control over the attire of innocent bystanders and foes. Foes, in particular, might be wearing copies of uniforms or of the clothes of innocent bystanders. Copies are normally designed to look like the originals but they are sometimes made using different dyes and can be distinguished from the originals using multi-spectral analysis. We look for color characteristics at different wavelengths. Two pieces of material that look alike to the human eye may look very different using multi-spectral analysis.*

Suppose that we are looking at two particular wavelengths of light – A and B, and that neither wavelength is visible to the naked eye. Light flow is commonly measured in lumens (lm), or lux (lumens per square meter). A piece of material will reflect percentages p_1 *and* p_2 *of wavelength A and B light, respectively. As spectators at a sporting event pass through a security checkpoint, they are illuminated by two brief, invisible flashes of light. The first flash has 30 lumens of wavelength A and 60 lumens of wavelength B. The second flash has 75 lumens of wavelength A and 25 lumens of wavelength B. The total amount of light from the first flash reflected by the shirt worn by a particular spectator is 15 lumens. The total amount of light reflected by the same shirt from the second light is 20 lumens. Determine what percentage of wavelength A light is reflected by the shirt and what percentage of wavelength light B is reflected by the shirt. These percentages would then be compared with known characteristics of different fabrics and dyes to determine if the spectator poses a risk to the rest of the crowd.*

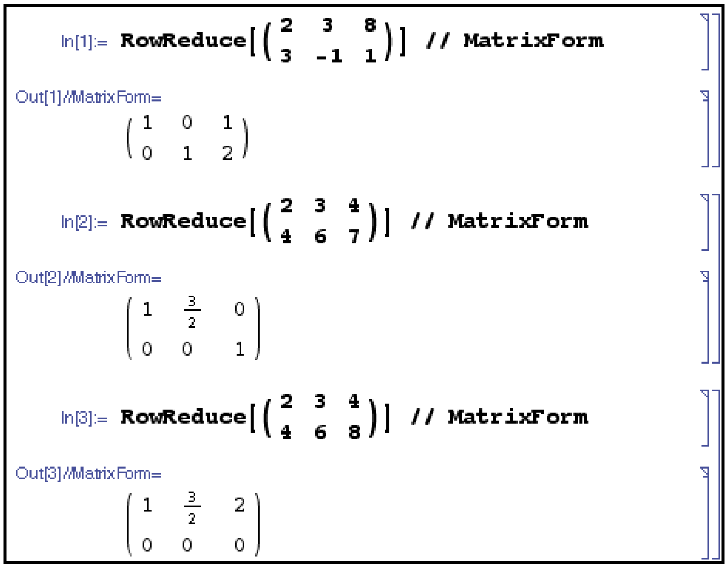


Figure 3.28: The **RowReduce** command in *Mathematica*

3.8 Mathematica's RowReduce and How Many Solutions?

Mathematica has a built in procedure **RowReduce** that performs row reduction with just a few keystrokes. You may have seen a similar procedure using the terminology **Row Reduced Echelon Form** on a graphing calculator. See Figure 3.28, which shows the *Mathematica* procedure applied to three of the systems of equations from the previous section.

$$\begin{aligned} 2x_1 + 3x_2 &= 8 \\ 3x_1 - x_2 &= 1 \end{aligned}$$

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 7 \end{aligned}$$

and

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 8. \end{aligned}$$

First, notice that when we apply row reduction to the system of equations

$$\begin{aligned} 2x_1 + 3x_1 &= 8 \\ 3x_1 - x_2 &= 1 \end{aligned}$$

we get the result

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

which corresponds to the solution

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 2 \end{aligned}$$

Next, notice that when we apply row reduction to the system of equations

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 7 \end{aligned}$$

we get the result

$$\left[\begin{array}{cc|c} 1 & \frac{3}{2} & 0 \\ 0 & 0 & 1 \end{array}\right],$$

and the second row of this augmented matrix corresponds to the equation

$$0 = 1.$$

This is obviously false – a contradiction!! If you look back at the original equations

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 7 \end{aligned}$$

and multiply the first equation by 2 you obtain

$$4x_1 + 6x_2 = 8,$$

which contradicts the second equation

$$4x_1 + 8x_2 = 7.$$

This means that the original equations contained contradictory information and have no solutions.

Whenever the augmented matrix produced by RowReduce has a row with all zeros on the left side and a nonzero entry in the rightmost column, then the system of equations is contradictory and does not have a solution.

Finally, notice that when we apply row reduction to the system of equations

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 8 \end{aligned}$$

we get the result

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

The second row in this augmented matrix corresponds to the equation

$$0 = 0.$$

This is obvious and provides no substantive information. If you look back at the original equations

$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 4x_1 + 6x_2 &= 8 \end{aligned}$$

and multiply the first equation by two you get

$$4x_1 + 6x_2 = 8,$$

which is the second equation. Thus, these two equations are really slight variations of the same equation. In this situation we say the equations are **redundant**, or repetitious.

Whenever the matrix produced by RowReduce has a row with all zeros, then some of the equations are redundant.

It is possible for a system of equations to be both redundant and contradictory. In this case, it will not have any solutions. For example, the system of equations

$$\begin{aligned} x + y + z &= 1 \\ 2x + 2y + 2z &= 2 \\ x + y + z &= 2 \end{aligned}$$

is both redundant and contradictory. The first two equations are redundant but the first and third equations are contradictory. This system of equations has no solutions. See Figure 3.29 on page 308. Notice the second row of the result produced by **RowReduce** corresponds to the equation

$$0 = 1,$$

indicating that this system of equations is contradictory. The third line corresponds to the equation

$$0 = 0,$$

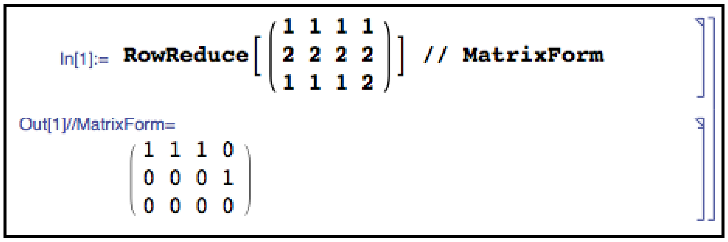


Figure 3.29: The **RowReduce** command in *Mathematica*

indicating that this system of equations is redundant. Because this system of equations is contradictory it has no solutions even though it is also redundant.

If a system of n equations in n unknowns is redundant and is not contradictory, then it has an infinite number of solutions.

Example 1 *Use* **RowReduce** *to determine whether the system of equations*

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

has one solution, no solutions, or an infinite number of solutions. If it has one solution find that solution. If it has an infinite number of solutions find two example solutions.

From Figure 3.30 on page 309 we see that this system has one solution

$$\begin{aligned} x &= \frac{29}{8} \\ y &= -\frac{15}{8} \\ z &= -\frac{19}{8}. \end{aligned}$$

Note that we used the letters x, y, and z for the unknowns instead of x_1, x_2, and x_3. We can use whichever letters are most convenient.

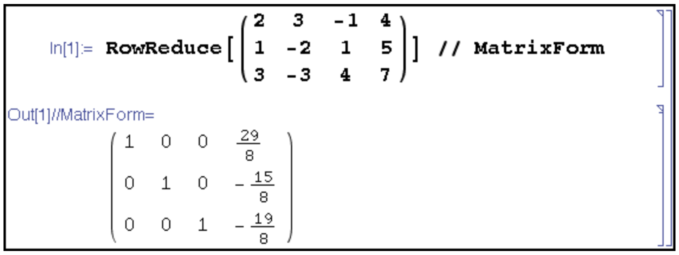


Figure 3.30: Example 1

Question 1 *Use* **RowReduce** *to attempt to solve the system of equations*

$$\begin{aligned} 2x_1 + x_2 - 3x_3 &= 4 \\ x_1 - 2x_2 + 2x_3 &= 2 \\ 3x_1 - x_2 - x_3 &= 5. \end{aligned}$$

Interpret the result of applying **RowReduce**.

Example 2 *Use* **RowReduce** *to determine whether the system of equations*

$$\begin{aligned} 2x_1 + x_2 - 3x_3 &= 4 \\ x_1 - 2x_2 + 2x_3 &= 2 \\ 3x_1 - x_2 - x_3 &= 6 \end{aligned}$$

has one solution, no solutions or an infinite number of solutions. If it has one solution find that solution. If it has an infinite number of solutions find two example solutions.

From Figure 3.31 on page 310 we see that this system has an infinite number of solutions. The three rows of the result of **RowReduce** shown in Figure 3.31 correspond to the three equations

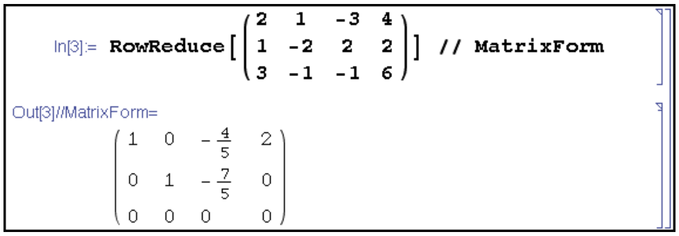


Figure 3.31: Example 2

$$
\begin{aligned}
x_1 - \left(\frac{4}{5}\right) x_3 &= 2 \\
x_2 - \left(\frac{7}{5}\right) x_3 &= 0 \\
0 &= 0.
\end{aligned}
$$

The third equation is redundant and of no use. We can rewrite the first two equations as

$$
\begin{aligned}
x_1 &= 2 + \left(\frac{4}{5}\right) x_3 \\
x_2 &= \left(\frac{7}{5}\right) x_3
\end{aligned}
$$

Now, we can let x_3 have any value and compute x_1 and x_2 from these two equations to get a solution of the original set of equations – for example, if $x_3 = 1$, the computations

$$
\begin{aligned}
x_1 &= 2 + \left(\frac{4}{5}\right) x_3 = 2 + \left(\frac{4}{5}\right)(1) = \frac{14}{5} \\
x_2 &= \left(\frac{7}{5}\right) x_3 = \left(\frac{7}{5}\right)(1) = \frac{7}{5}
\end{aligned}
$$

give us the solution

$$
\begin{aligned}
x_1 &= \frac{14}{5} \\
x_2 &= \frac{7}{5} \\
x_3 &= 1
\end{aligned}
$$

We can get another solution by letting $x_3 = 2$ as follows

$$
\begin{aligned}
x_1 &= 2 + \left(\frac{4}{5}\right) x_3 = 2 + \left(\frac{4}{5}\right)(2) = \frac{18}{5} \\
x_2 &= \left(\frac{7}{5}\right) x_3 = \left(\frac{7}{5}\right)(2) = \frac{14}{5}
\end{aligned}
$$

giving us the solution

$$
\begin{aligned}
x_1 &= \frac{14}{5} \\
x_2 &= \frac{7}{5} \\
x_3 &= 1.
\end{aligned}
$$

We can repeat these calculations as often as we want to get as many solutions as we want.

Question 2 *Use* **RowReduce** *to attempt to solve the system of equations*

$$
\begin{aligned}
x_1 + x_2 - x_3 - x_4 &= 2 \\
x_1 + x_2 + x_3 + x_4 &= 4 \\
x_1 + x_2 &= 3 \\
x_3 + x_4 &= 1.
\end{aligned}
$$

Interpret the result of **RowReduce***.*

We often describe the solutions for the system of equations in Example 2 as a **one parameter** family of solutions because in Example 2 any value for the one parameter x_3 gives us a solution. The solutions of the system of equations in Question 2 are often described as a **two parameter** family of solutions because we get a solution for any choice of values for two parameters.

Question 3 *Use* **RowReduce** *to solve the following systems of equations. Once you have your* **RowReduce** *output, convert the output back into a system of equations. If the system has one solution find that solution. If it has an infinite number of solutions find two example solutions.*

a.
$$\begin{aligned} x_1+2x_2+3x_3-x_4 &= 1\\ 2x_1-x_2+2x_3-x_4 &= 2\\ x_1-2x_2-3x_3+x_4 &= -1\\ 2x_1-2x_2-x_3+5x_4 &= 1 \end{aligned}$$

b.
$$\begin{aligned} x_1-x_2+x_3-x_4 &= 0\\ x_1+x_2+x_3+x_4 &= 4\\ x_1-x_2-x_3+3x_4 &= 2\\ x_1-2x_2-x_3+5x_4 &= 3 \end{aligned}$$

c.
$$\begin{aligned} x_1-x_2+x_3 &= 0\\ x_1+x_2+x_3 &= 4\\ 3x_1-x_2+3x_3 &= 6 \end{aligned}$$

d.
$$\begin{aligned} x_1-x_2+x_3 &= 0\\ x_1+x_2+x_3 &= 4\\ 3x_1-x_2+3x_3 &= 8 \end{aligned}$$

e.
$$\begin{aligned} x_1+x_2+x_3 &= 3\\ 2x_1+2x_2+2x_3 &= 6\\ 3x_1+3x_2+3x_3 &= 8 \end{aligned}$$

f.
$$\begin{aligned} x_1+x_2+x_3 &= 3\\ 2x_1+2x_2+2x_3 &= 6\\ 3x_1+3x_2+3x_3 &= 9 \end{aligned}$$

g.
$$\begin{aligned} x_1+6x_2+2x_4 &= 23\\ 2x_1+4x_2+2x_3 &= 20\\ 2x_1+x_2+5x_3+2x_4 &= 23\\ x_1+x_2+x_3+x_4 &= 11 \end{aligned}$$

h.
$$\begin{aligned} x_1+6x_2+2x_4 &= 23\\ 2x_1+4x_2+2x_3 &= 20\\ 2x_1+x_2+5x_3+2x_4 &= 23\\ x_1+2x_2+x_3 &= 11 \end{aligned}$$

i.
$$\begin{aligned} x_1+5x_2-9x_3 &= -64\\ 2x_1+4x_2-2x_3+x_4 &= -22\\ 3x_1+4x_2-x_3+x_4 &= -16\\ x_2+5x_3+4x_3-8x_4 &= -16 \end{aligned}$$

j.
$$\begin{aligned} x_1+6x_2+2x_4 &= 23\\ 2x_1+4x_2+2x_3 &= 20\\ 2x_1+x_2+5x_3+2x_4 &= 23\\ x_1+2x_2+x_3 &= 10 \end{aligned}$$

Question 4 *An important issue in homeland security is distinguishing rapidly, accurately, and non-intrusively between friends and foes. For example, in crowd situations there may be uniformed friends, undercover friends, innocent bystanders, and foes. We have some control over the attire of uniformed and undercover friends but no control over the attire of innocent bystanders and foes. Foes, in particular, might be wearing copies of uniforms or of the clothes of innocent bystanders. Copies are normally designed to look like the originals but they are sometimes made using different dyes and can be distinguished from the originals using multi-spectral analysis. We look for color characteristics at different* *wavelengths. Two pieces of material that look alike to the human eye may look very different using multi-spectral analysis.*

Suppose that we are looking at two particular wavelengths of light – A and B, and that neither wavelength is visible to the naked eye. Light flow is commonly measured in lumens (lm), or lux (lumens per square meter). A piece of material will reflect percentages p_1 *and* p_2 *of wavelength A and B light, respectively. As spectators at a sporting event pass through a security checkpoint, they are illuminated by two brief, invisible flashes of light. The first flash has 30 lumens of wavelength A and 60 lumens of wavelength B. The second flash has 75 lumens of wavelength A and 25 lumens of wavelength B. The total amount of light from the first flash reflected by the shirt worn by a particular spectator is 15 lumens. The total amount of light reflected by the same shirt from the second light is 20 lumens. Determine what percentage of wavelength A light is reflected by the shirt and what percentage of wavelength light B is reflected by the shirt. These percentages would then be compared with known characteristics of different fabrics and dyes to determine if the spectator poses a risk to the rest of the crowd.*

This problem appeared at the end of the previous section but now we are solving it using **RowReduce***.*

3.9 Networks

Today's world is heavily impacted by the use of networks. The food we eat was brought from the field or farm that the ingredients were grown to a manufacturing plant to the Mess Hall via a transportation network. We communicate each day through a communication network, known as the Internet. Many people transmit text and data messages through a communication network with cell phones. You interact each day inside a social network.

Through understanding what networks are and how they may be used, we gain a greater understanding of the world around us. In fact, knowledge of networks is being used today for military purposes in the Global War on Terrorism. Research done by 1LT Julie Paynter (Jorgenson), a 2006 USMA graduate and math department alumnus, is currently contributing to understanding terrorist networks and determining whether or not an unknown author is violent or non-violent. [1] To lay the groundwork for our exploration of networks, let us define our terms.

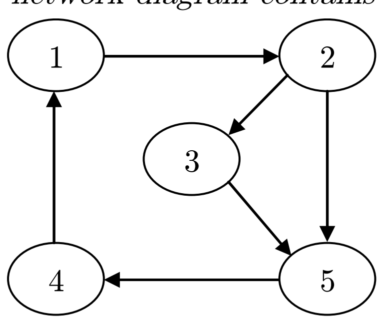
3.9.1 Basic Definitions

A *graph* or *network* is made up of *nodes* and *arcs.*

Definition 1 A ***node*** (also called a ***vertex***) is a transfer point within a network, such as a facility or intersection.

Definition 2 An ***arc*** consists of an ordered pair of nodes and is a representation of possible flow between nodes or vertices. The *initial node* is the first node in the pair, representing the start point of the flow. The *terminal node* is the final node in the pair, representing the end of the arc. Arc $x_{j,k}$ denotes flow on arc x from node j to node k.

Example 1 *The following network diagram contains nodes $N = \{1, 2, 3, 4, 5\}$.*



The corresponding arcs are $A = \{(1, 2), (2, 3), (2, 5), (3, 5), (5, 4), (4, 1)\}$. Notice that the arrowhead on the arc shows the direction of flow and corresponds to the order that the vertices occur within the label of each arc. The label on arc (2, 3) denotes flow from node 2 to node 3. The arrowhead points from node 2 to node 3.

[1] "West Point Cadet Interns at the Hugh Downs School," *Communication Matters*, 2006, Available from http://www.asu.edu/clas/communication/about/news/newsletter/documents/NewsletterMar06.pdf; Internet, Accessed April 1, 2008.

3.9.2 Flow Within a Network

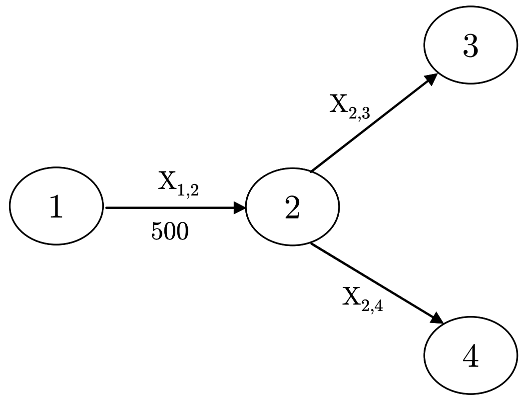


Figure 3.32: A Y-intersection

Figure 3.32 shows a diagram of a Y-intersection of three roads –one road enters the intersection from the west and two roads leave it, one toward the northeast and one toward the southeast. All three roads are one way roads, as indicated by the arrows. Suppose that, on the average during each morning's commute, 500 vehicles enter the intersection from the west; $X_{2,3}$ vehicles leave it toward the northeast; and $X_{2,4}$ vehicles leave it toward the southeast.

Question 1 Write an equation that describes the relationship between $X_{2,3}$ and $X_{2,4}$.

Question 2 Is it possible to have a correct mathematical solution to the equation you developed in the preceding question that does not make intuitive sense? Explain your answer.

The answer to Question 1 is the basic relationship upon which network flow analysis is based: what comes in must come out. If your conjecture to question 1 was that

$$X_{1,2} = X_{2,3} + X_{2,4} \text{ (or, } 500 = X_{2,3} + X_{2,4})$$

you were correct!

Figure 3.33 on page 326 shows the junctions and one way roads in a town center. The average number of vehicles entering and leaving the town center each hour has been determined by data collection devices and is shown by the numbers in Figure 3.33.

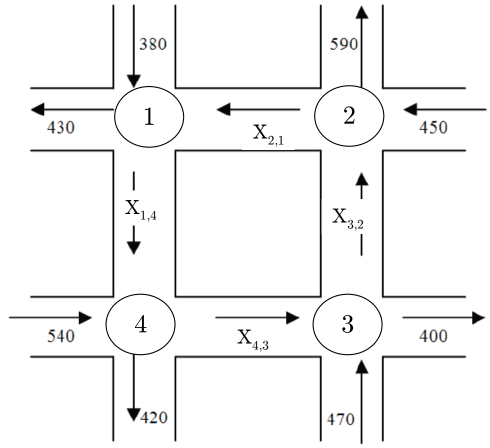


Figure 3.33: One way roads in the town center

Example 2 *Determine the average number of vehicles traveling each hour along each of the four one way roads in the town center – that is, the road from Junction 1 to Junction 4; the road from Junction 4 to Junction 3; the road from Junction 3 to Junction 2; and the road from Junction 2 to Junction 1. Explain your response. If you could recommend the city collect some more data (so that you could complete your answer to this question), where would you recommend they collect the data?*

The first step in solving this problem is to transform the word problem into a mathematical model that we can solve. We must set up a system of equations that will enable the solution for each of the unknown variables. We begin by establishing one equation for each node: ***what goes in must come out****.*

$$\begin{aligned}
&Node1: && X_{2,1} + 380 = 430 + X_{1,4} \\
&Node2: && X_{3,2} + 450 = 590 + X_{2,1} \\
&Node3: && X_{4,3} + 470 = X_{3,2} + 400 \\
&Node4: && X_{1,4} + 540 = X_{4,3} + 420
\end{aligned}$$

A goal in solving a system of equations with many variables should be to put the system of equations in matrix-vector form to solve using the inverse method or row reduction. To do that, put the above system of equations in standard form, with all the variables on one side, with each variable in its own column.

$$\begin{array}{llllll}
Node1: & -X_{1,4} & +X_{2,1} & & & = 430 - 380 = 50 \\
Node2: & & -X_{2,1} & +X_{3,2} & & = 590 - 450 = 140 \\
Node3: & & & -X_{3,2} & +X_{4,3} & = 400 - 470 = -70 \\
Node4: & X_{1,4} & & & -X_{4,3} & = 420 - 540 = -120
\end{array}$$

The system of equations in standard form enables us to put the equations in matrix-vector form:

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} X_{1,4} \\ X_{2,1} \\ X_{3,2} \\ X_{4,3} \end{bmatrix} = \begin{bmatrix} 50 \\ 140 \\ -70 \\ -120 \end{bmatrix}$$

With our system of equations in matrix-vector form, a quick check of the determinant in Mathematica (Figure 3.34) shows that the coefficient matrix is singular, therefore we cannot use the inverse method to solve the system of equations. There are either infinite or no solutions; row reduction is the next solution technique to try.

```
A = ( -1  1  0  0
       0 -1  1  0
       0  0 -1  1
       1  0  0 -1 )

Det[A]

0
```

Figure 3.34: Use of Mathematica to Check Singularity of a Coefficient Matrix

After forming the augmented matrix from the above system of equations, the use of Mathematica's RowReduce command provides the output in Figure 3.35, showing that there are an infinite number of solutions to the problem.

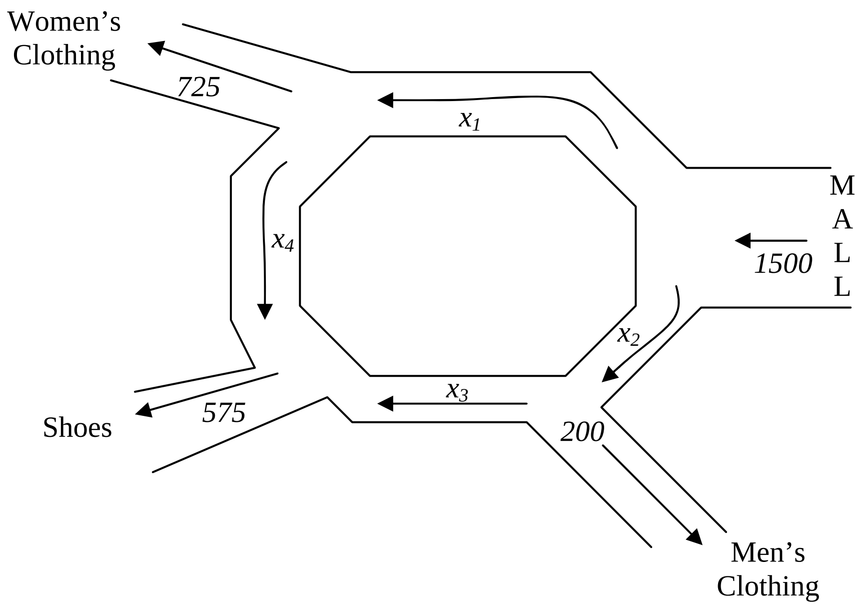
```
In[5]:= AugmentedA = ( -1  1  0  0   50
                        0 -1  1  0  140
                        0  0 -1  1  -70
                        1  0  0 -1 -120 )

In[6]:= RowReduce[AugmentedA] // MatrixForm

Out[6]//MatrixForm=
        ( 1 0 0 -1 -120
          0 1 0 -1 -70
          0 0 1 -1 70
          0 0 0  0   0 )
```

Figure 3.35: Mathematica Row Reduction output

We see that each of the variables can be found by using simple algebra, if we know (or assume) a value for $X_{4,3}$. Therefore, I would suggest that the city collect more data, specifically the number of vehicles that transit road $X_{4,3}$. When this number is established, the flow across all other roads can be found using substitution.

Question 3 Suppose that the city must complete some repairs on a bridge that is immediately to the south of intersection 4. During the repairs, only half of the normal traffic will be able to pass. What impact, if any, will that have on the traffic exiting the city north of intersection 2? What advice would you give to the town to help them cope with traffic while this road is closed?

Question 4 Observe the following diagram of the flow of customers through a Department store. Assume 1500 customers enter the Mall, 725 go to the Women's Clothing Department, 575 to the Shoe Department, and 200 to the Men's Clothing Department. Let the variable adjacent to each arrow represent the number of customers walking along that arrow.

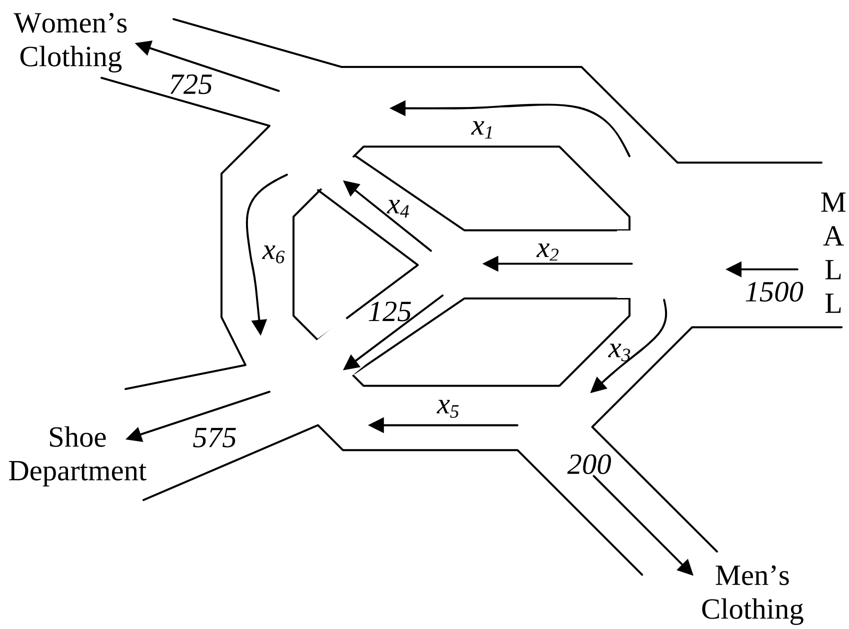


a. Using the variables x_1, x_2, x_3, *and* x_4, set up a system of equations that models the flow of customers through the department store above.

b. How many customers pass along each of the internal arcs above? Clearly indicate the process used in determining your answers. If there are no solutions, explain why; if there is one solution, provide it; and if there are infinitely many solutions, give two.

c. Why might management be interested in knowing the number of customers that are travelling along these paths?

Question 5 Assume now that the diagram in **Question 5** is a simplified model of what actually exists. The following diagram is more realistic.



How many customers pass through each of the points indicated with a variable (x_1, x_2, x_3, x_4, x_5 *and* x_6)? Clearly indicate the process used and interpret your results. If there are no solutions explain why, if there is one solution, provide it and if there are infinite solutions, give two.

3.10 Images – Transformations and Animations

3.10.1 Transformations of Still Images

Vectors and matrices are often used to manipulate images and create special effects in movies like Star Wars. At the end of this section, you will be able to move and rotate images and reflect (or flip) them, as in mirrors. You will also be able to make them larger or smaller or stretch them horizontally and vertically by different factors. You will even be able to create animations or movies involving these same effects. Figure 3.32 shows a simple example – flipping an image upside down – of what we will be able to do. We start with an image like the one on the left side of Figure 3.32 and manipulate it in some way, producing a new image like the one on the right side of Figure 3.32.

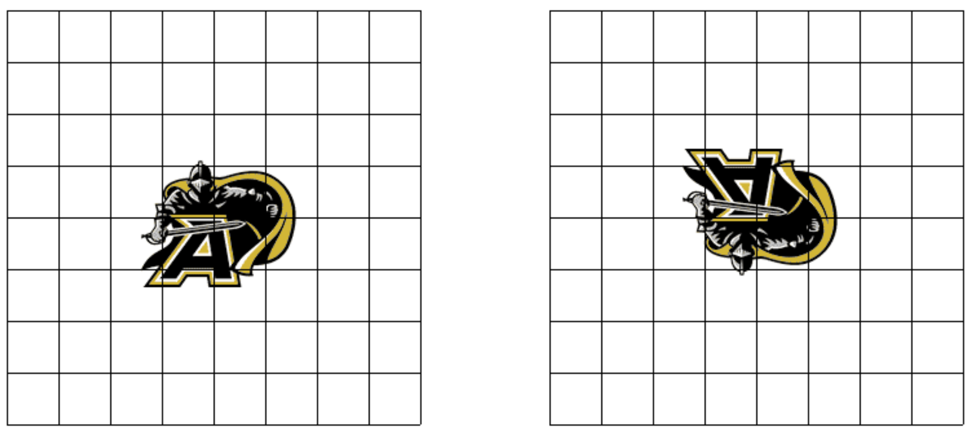


Figure 3.32: Flipping an image upside down

The key to all of this is representing points in an original image by column-vectors,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

and points on a television, movie or computer screen, or on a page where the new image will appear, by column-vectors,

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Then each point, represented by a vector $\vec{x}$, from the original image is placed at the point represented by the vector $\vec{y} = A\vec{x}$ on the computer screen where A is a matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Notice that we can also write this as shown below.

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 \\ y_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned}$$

In the example shown in Figure 3.32, the matrix A is

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

so that

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= -x_2. \end{aligned}$$

Thus the x-coordinate (the first coordinate), denoted x_1, of a point in the original image is unchanged in the new image and the y-coordinate (the second coordinate), denoted x_2, of a point in the original image is multiplied by -1 before it is placed in the new image. In effect, this flips the original image upside-down.

Click here[9] to open a new window with a live version of Figure 3.33 on page 322. The two axes in this live figure run from -4 to 4. You can change the entries (highlighted in salmon) in the matrix A in the usual way by selecting them with your mouse and editing them. After you've changed the entries, press the **Play** button, and notice the effect on the image on the right side of the figure.

Question 1 *Produce the image shown in Figure 3.34 on page 323 by changing the entries in matrix A. Note that you can see Figure 3.34 on the left side of the screen by clicking the* **Question 1** *button in the upper right corner.*

[9] http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/image-transformations-1/imageAnimation.html

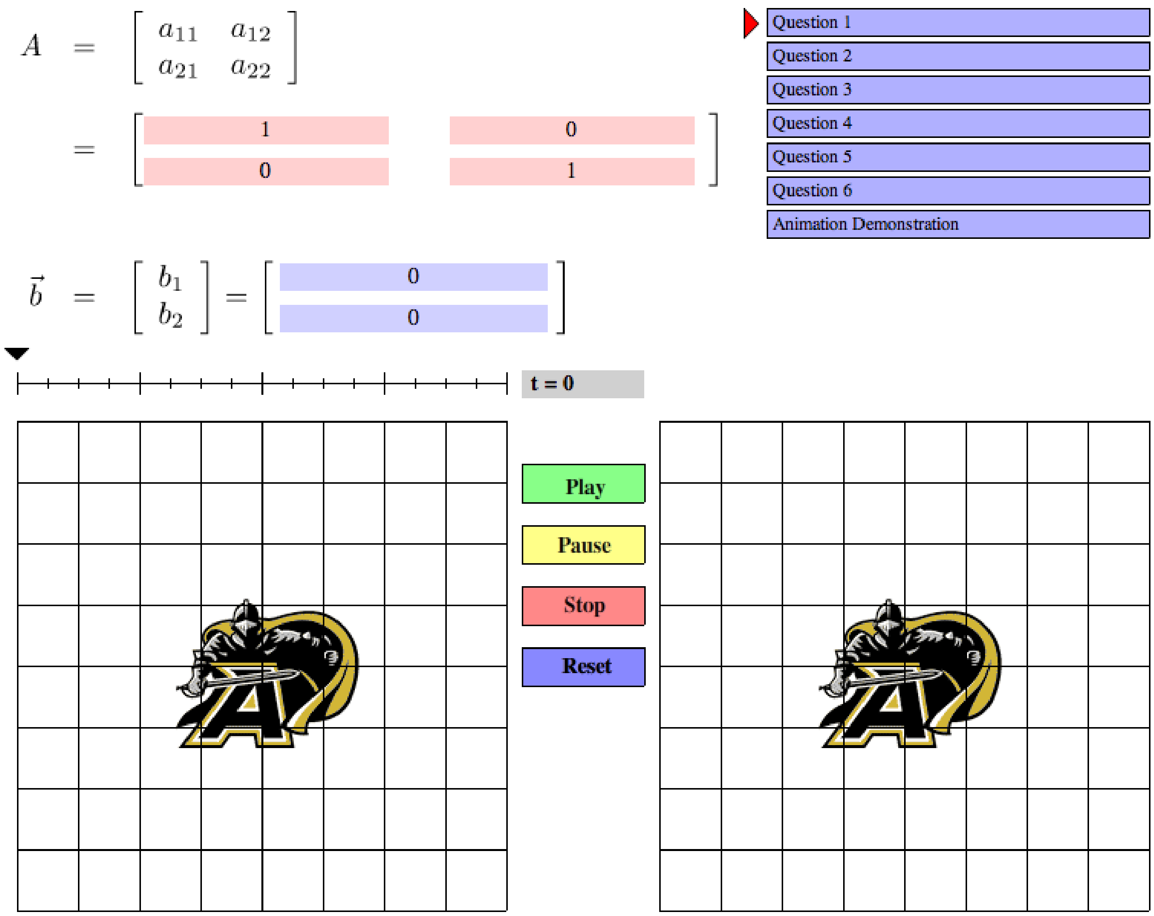


Figure 3.33: Screenshot of a live figure

Question 2 *Produce the image shown in Figure 3.35 on page 323 by changing the entries in matrix A. Note that you can see Figure 3.35 on the left side of the screen by clicking the* **Question 2** *button in the upper right corner.*

Question 3 *Produce the image shown in Figure 3.36 on page 324 by changing the entries in matrix A. Note that you can see Figure 3.36 on the left side of the screen by clicking the* **Question 3** *button in the upper right corner.*

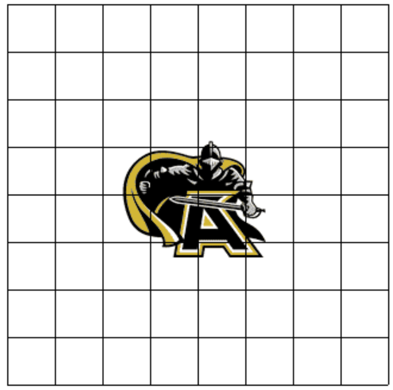


Figure 3.34: Question 1

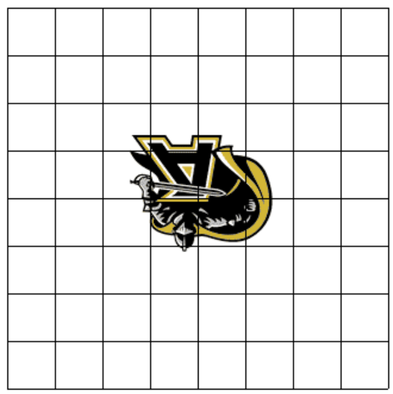


Figure 3.35: Question 2

To produce different effects, we need to know a bit more about how the image manipulation

$$\vec{y} = A\vec{x} \qquad \text{or} \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

affects the image at $\vec{y}$ of each point $\vec{x}$ in the original image. First, notice that these image manipulations always leave the point $(0,0)$ unchanged. Two additional good points to look at from the original image are the points $(1,0)$ and $(0,1)$, which we write as column-vectors

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

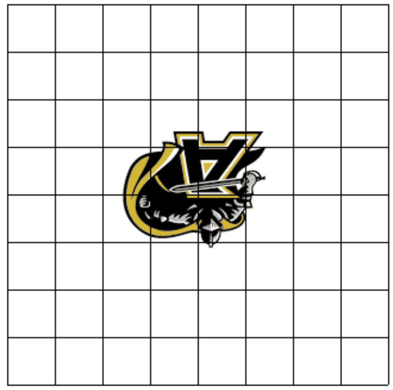


Figure 3.36: Question 3

Notice that when

$$\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

then

$$\vec{y} = A\vec{x} \quad \text{becomes} \quad \vec{y} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}.$$

So, whatever part of the original image was located at the point $(1, 0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is now placed at the point given by the first column of the matrix A in the new image.

Similarly, when

$$\vec{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

then

$$\vec{y} = A\vec{x} \quad \text{becomes} \quad \vec{y} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}.$$

So whatever part of the original image was located at the point $(0,1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is now placed at the point given by the second column of the matrix A in the new image.

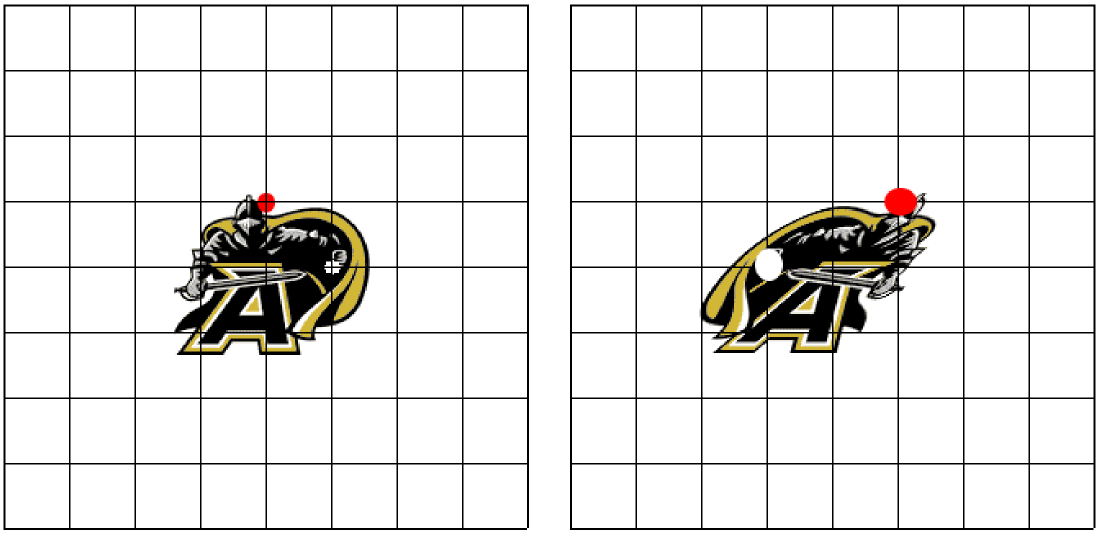


Figure 3.37: Two key points

You can produce any image transformation that you want by determining the new coordinates of the point $(1,0)$ and using these new coordinates as the first column of the matrix A and then determining the new coordinates of the point $(0,1)$ and using these new coordinates as the second column of the matrix A. The two key points $(1,0)$ and $(0,1)$ are marked on the left side of Figure 3.37. The point $(1,0)$ is marked by a white dot and the point $(0,1)$ is marked by a red dot on both pictures. Recall that the axes run from -4 to 4. In the example shown on the right side of Figure 3.37 the point $(1,0)$ goes to the point $(-1,0)$ and the point $(0,1)$ goes to the point $(1,1)$. Thus, the required matrix is

$$A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Question 4 *Using the usual live figure,*[10] *produce the image shown in Figure 3.38 on page 326 by changing the entries in matrix A. Note that you can see Figure 3.38 on the left side of the screen by clicking the* **Question 4** *button in the upper right corner.*

[10]http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/image-transformations-1/imageAnimation.html

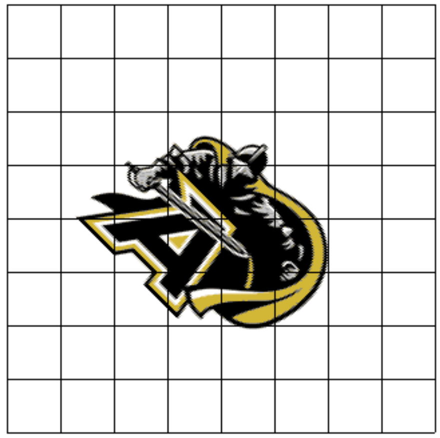


Figure 3.38: Question 4

Question 5 *Produce the image shown in Figure 3.39 by changing the entries in matrix A. Note that you can see Figure 3.39 on the left side of the screen by clicking the* **Question 5** *button in the upper right corner.*

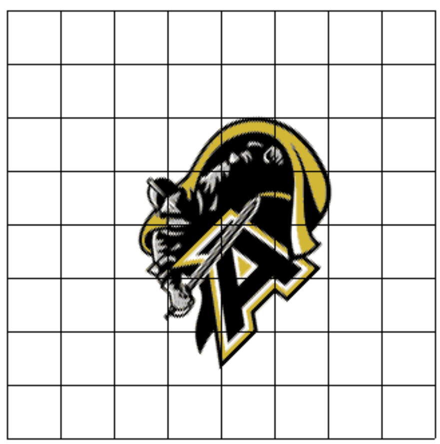


Figure 3.39: Question 5

3.10.2 Rotation of Images

One very common image manipulation is rotation – for example, Figure 3.40 shows a rotation of 30 degrees ($\pi/6$ radians) in the counterclockwise direction.

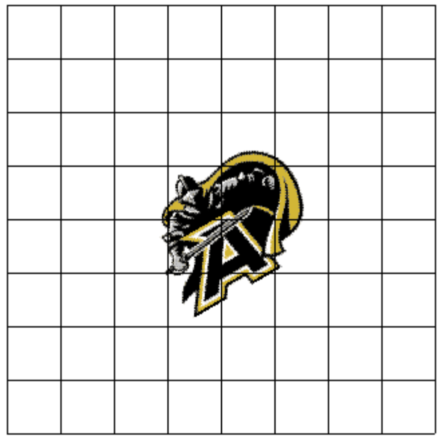


Figure 3.40: Rotation by 30 degrees counterclockwise

We can use Figure 3.41 to see how to accomplish a rotation counterclockwise by an angle θ. First look at what happens to the point $(1, 0)$. After it is rotated by the angle θ it is still 1 unit away from the origin and by looking at the shaded triangle in Figure 3.41 we see that its coordinates are $(\cos\theta, \sin\theta)$ and, thus, the first column of the matrix we seek is

$$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

Similarly, by looking at the cross-hatched triangle in Figure 3.41 on page 328 we see that the new coordinates of the point $(0, 1)$ are $(-\sin\theta, \cos\theta)$ and, thus the second column of the matrix we seek is

$$\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}.$$

Thus, we can rotate an image counterclockwise by an angle of θ by using the matrix

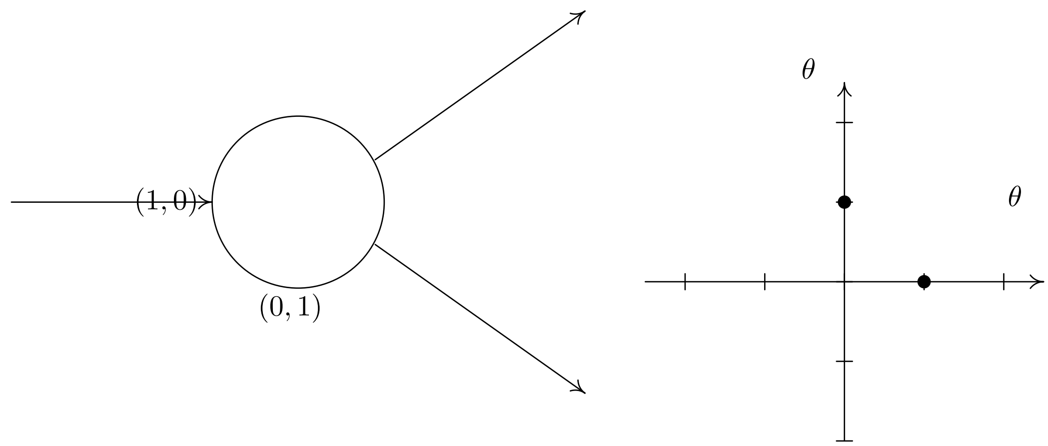


Figure 3.41: Counterclockwise rotation by θ radians

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Use the following rules when you enter algebraic expressions in our live figures.

- Use ***** to denote multiplication.
- Use **pi** to denote π.
- Use **cos** for cosine and **sin** for sine.
- To take the square root of a number use something like **sqrt(2)**.

So far, we have changed images in various ways while keeping them in the center of the screen. If we want to move an image as well, we need to use transformations given by

$$\vec{y} = A\vec{x} + \vec{b} \qquad \text{or} \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

where the vector

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

determines where the center of the original image is placed on the screen, since

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Figure 3.42 shows how we can move an image and reflect it about the independent axis. Click here[11] to open the usual live figure. Recall that the axes run from -4 to 4. The matrix A and the vector $\vec{b}$ can be changed by editing their entries in the usual way. After you change them click the **Play** button to see the effect on the new image. Figure 3.43 shows another example.

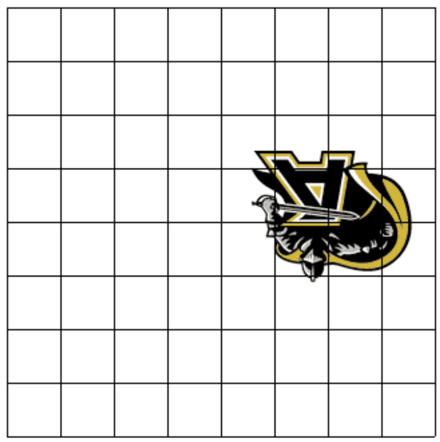


Figure 3.42: Moving an image and reflecting it about the independent axis.

[11] http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/image-transformations-1/imageAnimation.html

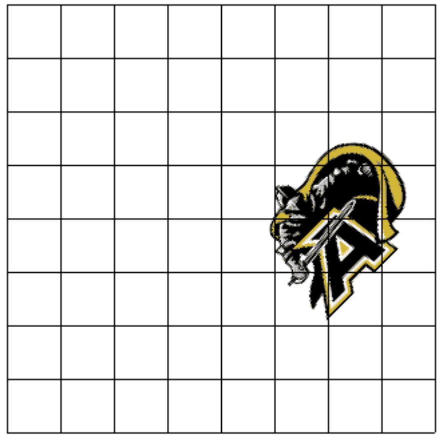


Figure 3.43: Moving an image and rotating it

Question 6 *Using the usual live figure,*[12] *produce the image shown in Figure 3.44 by changing the entries in the matrix A and the vector $\vec{b}$.*

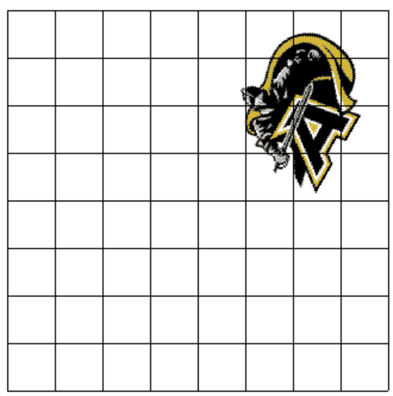


Figure 3.44: Question 6

Question 7 *As an in-class exercise, break into teams. Each team should create an interesting image manipulation and then challenge the other teams to reproduce it.*

[12] http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/image-transformations-1/imageAnimation.html

3.10.3 Animations

So far we have manipulated static, or still, images. Now we want to create animations. In the usual live figure[13] enter the matrix

$$A = \begin{bmatrix} 1-t & 0 \\ 0 & 1-t \end{bmatrix}$$

and click the **Play** button.

This matrix is a bit different from the matrices we have used so far. Two of the entries in this matrix are algebraic expressions involving the variable t rather than numbers. When you click the green **Play** button, the variable t increases from 0 to 1 in small steps. As it does, the values of the entries in the matrix A change and this causes the new image to change as well. The result is an animation showing the original image full-sized (when $t = 0$) and then fading away to nothing (when $t = 1$).

Click here[14] to open a new window with another live figure. This window has a series of questions that challenge you to produce various animations. Click on each of the buttons on the upper right corner and then click the **Play** button to see the animations you are to reproduce. The challenges are shown in the left side of the window. Your challenge is to reproduce them on the right side by changing the entries in the matrix A and the vector $\vec{b}$.

Question 8 *As an in-class exercise, break into teams. Each team should create an interesting animation and then challenge the other teams to reproduce it.*

[13] http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/image-transformations-1/imageAnimation.html

[14] http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_III/image-transformations-2/imageAnimation.html

Chapter 4

Discrete Dynamical Systems with Many Variables

4.1 Discrete Dynamical Sytems with Many Variables

We begin this section with an example.

Example 1 *You are in charge of a fleet of humvees during an extended engagement. The first week you have 150 humvees, all of which are fully operational. At the end of each week, 15% of the humvees that were operational at the beginning of the week are in the shop for repairs and 5% are so damaged that they can only be used for spare parts. At the end of each week, 75% of the humvees that were in the shop for repairs at the beginning of the week are now fully operational and 25% are only good for spare parts. Each week you receive five fully operational humvees from Supply. Make a table showing how many humvees are in each of the following three categories each week for the first 40 weeks – fully operational, in the shop for repairs, only good for spare parts.*

We keep track of three different classes of humvees – fully operational humvees, humvees that are in for repairs, and humvees that are good only for parts. We use the notation w_n for fully operational humvees (working); the notation r_n for humvees that are in for repairs; and s_n for humvees that are good only for spare parts. Since we have 150 humvees the first week and they are all fully operational, the initial values of these sequences are

$$w_0 = 150, \quad r_0 = 0, \quad \text{and} \quad s_0 = 0.$$

We use the letter n for the domain and note that $n = 0, 1, 2, 3, \ldots$ with $n = 1$ the end of the 1^{st} week of the engagement.

The following recursion equations describe how the situation changes from one week to the next.

$$\begin{aligned} w_n &= 0.80w_{n-1} + 0.75r_{n-1} + 5 \\ r_n &= 0.15w_{n-1} \\ s_n &= 0.05w_{n-1} + 0.25r_{n-1} + s_{n-1} \end{aligned}$$

Using matrices and vectors, we can write this as

$$\begin{bmatrix} w_n \\ r_n \\ s_n \end{bmatrix} = \begin{bmatrix} 0.80 & 0.75 & 0 \\ 0.15 & 0 & 0 \\ 0.05 & 0.25 & 1 \end{bmatrix} \begin{bmatrix} w_{n-1} \\ r_{n-1} \\ s_{n-1} \end{bmatrix} + \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}.$$

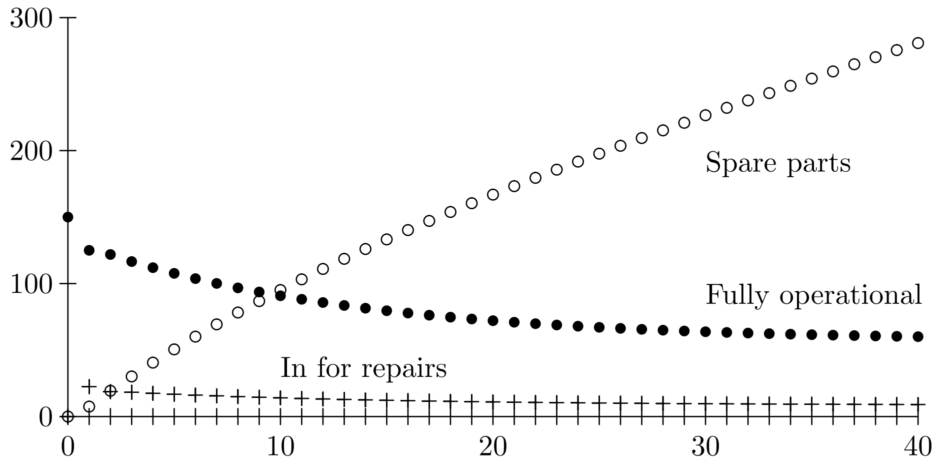


Figure 4.1: Humvees

Figure 4.1 and, on page 334, Table 4.1 show the predictions made by this model for the first 40 weeks. Click here[1] for a spreadsheet with this same model.

Question 1 *Describe the long term behavior of this model.*

[1]http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_IV/humvees.xls

Week	Fully Operational	In for Repairs	Parts
0	150.00	0.00	0.00
1	125.00	22.50	7.50
2	121.88	18.75	19.38
3	116.56	18.28	30.16
4	111.96	17.48	40.55
5	107.68	16.79	50.52
6	103.74	16.15	60.11
7	100.11	15.56	69.33
8	96.76	15.02	78.23
9	93.67	14.51	86.82
10	90.82	14.05	95.13
11	88.19	13.62	103.18
12	85.77	13.23	111.00
13	83.54	12.87	118.60
14	81.48	12.53	125.99
15	79.58	12.22	133.20
16	77.83	11.94	140.23
17	76.22	11.67	147.11
18	74.73	11.43	153.84
19	73.36	11.21	160.43
20	72.10	11.00	166.90
21	70.93	10.81	173.26
22	69.85	10.64	179.51
23	68.86	10.48	185.66
24	67.95	10.33	191.72
25	67.11	10.19	197.70
26	66.33	10.07	203.61
27	65.61	9.95	209.44
28	64.95	9.84	215.21
29	64.34	9.74	220.91
30	63.78	9.65	226.57
31	63.26	9.57	232.17
32	62.79	9.49	237.72
33	62.35	9.42	243.24
34	61.94	9.35	248.71
35	61.57	9.29	254.14
36	61.22	9.23	259.54
37	60.90	9.18	264.91
38	60.61	9.14	270.25
39	60.34	9.09	275.57
40	60.09	9.05	280.86

Table 4.1: Humvees

Question 2 *What would be the long term behavior of starting with 300 humvees instead of 150?*

Question 3 *What would be the long term behavior of starting with 150 humvees and receiving ten new humvees each week instead of five?*

We will leverage the matrix and vector description of this model,

$$\begin{bmatrix} w_n \\ r_n \\ s_n \end{bmatrix} = \begin{bmatrix} 0.80 & 0.75 & 0 \\ 0.15 & 0 & 0 \\ 0.05 & 0.25 & 1 \end{bmatrix} \begin{bmatrix} w_{n-1} \\ r_{n-1} \\ s_{n-1} \end{bmatrix} + \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}.$$

First, we introduce some notation.

$$\vec{H}_n = \begin{bmatrix} w_n \\ r_n \\ s_n \end{bmatrix},$$

$$M = \begin{bmatrix} 0.80 & 0.75 & 0 \\ 0.15 & 0 & 0 \\ 0.05 & 0.25 & 1 \end{bmatrix}, \text{ and}$$

$$\vec{b} = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}.$$

With this notation we can express this model as

$$\vec{H}_n = M\vec{H}_{n-1} + \vec{b}.$$

Question 4 *In Example 1 we kept track of fully operational humvees, humvees in for repairs, and humvees being used for parts. How would you modify this model to keep track of only fully operational humvees and humvees in for repairs? Express your model using two sequences and also using matrices and vectors.*

Question 5 *Find an equilibrium for the model in Question 4. What do we mean by an equilibrium in this situation?*

Question 6 *The population of a particular country is divided into two groups. One group lives in cities and the other group lives in the countryside. Currently, there are 10,000,000 people living in the countryside and 5,000,000 living in the cities. Suppose each year 75% of the people living in the countryside remain there, 20% move to the cities and 5% die.*[2] *Suppose each year 90% of the people living in the cities remain in the cities, 5% move to the countryside and 5% die. Suppose each year 20,000 people immigrate into this country and all move into the countryside. Build a model for this situation and describe the long term population for this country. Express your model with two sequences and also with matrices and vectors.*

Question 7 *Two pizzerias – Tony's and Mario's – supply all the pizzas to the students living in Collegetown. Suppose each week each of the students orders one pizza or some other fast food. Suppose the first week of the semester 2,000 pizzas are ordered from Tony's and 2,000 pizzas are ordered from Mario's. Suppose each week 75% of Mario's customers from the previous week remain faithful to Mario and order a pizza from Mario, 15% switch to Tony's and 10% switch to some other fast food. Suppose each week 85% of Tony's customers remain loyal to Tony, 10% switch to Mario's, and 5% switch to some other fast food. Suppose each week 50 new students move into Collegetown and they all buy their first pizza from Tony's. If the semester is 15 weeks long, how many customers does each pizzeria have in the* 15^{th} *week? Use sequence notation to express this situation and answer the questions. Express the same model using matrices and vectors. What is the total number of pizzas sold by each pizzeria during the 15 week semester?*

Now we are ready to continue the example with which we started this section.

Example 2 *You are in charge of a fleet of humvees during an extended engagement. The first week of the engagement you have 150 humvees, all of which are fully operational. At the end of each week, 15% of the humvees that were operational at the beginning of the week are in the shop for repairs and 5% are so damaged that they can only be used for spare parts. At the end of each week, 75% of the humvees that were in the shop for repairs at the beginning of the week are now fully operational and 25% are only good for spare parts. Each week you receive five fully operational humvees from Supply. Predict how many humvees are in each of the three categories – fully operational, in the shop for repairs, only good for spare parts – each week over the long term .*

If we keep track of only fully operational humvees and humvees that are in for repairs, we obtain the following model.

[2] 5% is actually the difference between the birth rate and the death rate.

$$\begin{aligned} w_0 &= 150 \\ r_0 &= 0 \\ w_n &= 0.80w_{n-1} + 0.75r_{n-1} + 5 \\ r_n &= 0.15w_{n-1} \end{aligned}$$

The recursion equations for this model can be written

$$\begin{bmatrix} w_n \\ r_n \end{bmatrix} = \begin{bmatrix} 0.80 & 0.75 \\ 0.15 & 0 \end{bmatrix} \begin{bmatrix} w_{n-1} \\ r_{n-1} \end{bmatrix} + \begin{bmatrix} 5 \\ 0 \end{bmatrix}.$$

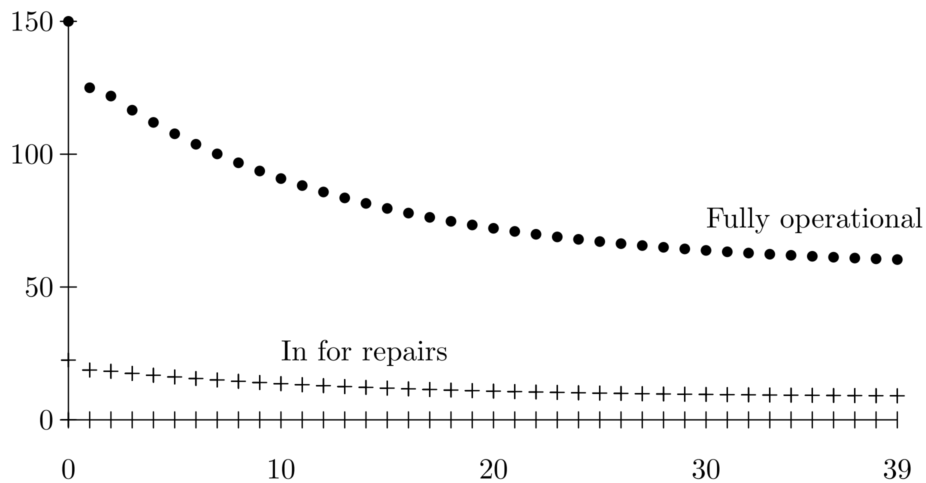


Figure 4.2: Humvees

Table 4.2 on page 338 and Figure 4.2 show the predictions made by this model for the first 40 weeks of the engagement. Notice that the number of fully operational humvees appears to be settling down at about 60 and the number of humvees that are in for repairs seems to be settling down at about 9. This might lead us to guess that there is an equilibrium point for this model.

Week	**Fully Operational**	**In for Repairs**
0	150.00	0.00
1	125.00	22.50
2	121.88	18.75
3	116.56	18.28
4	111.96	17.48
5	107.68	16.79
6	103.74	16.15
7	100.11	15.56
8	96.76	15.02
9	93.67	14.51
10	90.82	14.05
11	88.19	13.62
12	85.77	13.23
13	83.54	12.87
14	81.48	12.53
15	79.58	12.22
16	77.83	11.94
17	76.22	11.67
18	74.73	11.43
19	73.36	11.21
20	72.10	11.00
21	70.93	10.81
22	69.85	10.64
23	68.86	10.48
24	67.95	10.33
25	67.11	10.19
26	66.33	10.07
27	65.61	9.95
28	64.95	9.84
29	64.34	9.74
30	63.78	9.65
31	63.26	9.57
32	62.79	9.49
33	62.35	9.42
34	61.94	9.35
35	61.57	9.29
36	61.22	9.23
37	60.90	9.18
38	60.61	9.14
39	60.34	9.09

Table 4.2: Humvees

Definition 1 *An equilibrium point for a model with two variables described by recursion equations,*

$$\begin{aligned} x_n &= f(x_{n-1}, y_{n-1}) \\ y_n &= g(x_{n-1}, y_{n-1}), \end{aligned}$$

is a pair of values x_* *and* y_* *such that*

$$\begin{aligned} x_* &= f(x_*, y_*) \\ y_* &= g(x_*, y_*). \end{aligned}$$

In other words, if the sequences x_n *and* y_n *start at these values, then they stay there.*

For our example, we look for equilibrium values by solving the pair of equations

$$\begin{aligned} w_* &= 0.80w_* + 0.75r_* + 5 \\ r_* &= 0.15w_* \end{aligned}$$

Substituting the second equation into the first equation, we obtain

$$\begin{aligned} w_* &= 0.80w_* + 0.75(0.15w_*) + 5 \\ w_* &= 0.80w_* + 0.1125w_* + 5 \\ 0.0875w_* &= 5 \\ w_* &= 57.1429 \end{aligned}$$

and, using the second of our two original equations,

$$r_* = 0.15 w_* = 8.5714.$$

Recall that for linear recursion equations,

$$p_n = m p_{n-1} + b,$$

we showed that, if $m \neq 1$, there is a unique equilibrium point given by the equation

$$p_* = \frac{b}{1-m}.$$

Our example gives us reason to hope that there might be a similar theorem for a system of linear equations. There is.

Theorem 1 *Suppose that*

$$\vec{p}_n = M\vec{p}_{n-1} + \vec{b}$$

is a linear system and that the matrix $I - M$ is non-singular.[3] *Then there is unique equilibrium point*

$$\vec{p}_* = (I - M)^{-1}\vec{b}.$$

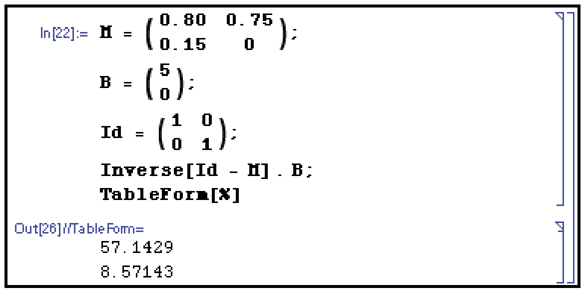
Proof

We must solve the equation

$$\begin{aligned}
\vec{p}_* &= M\vec{p}_* + \vec{b} \\
I\vec{p}_* &= M\vec{p}_* + \vec{b} \\
I\vec{p}_* - M\vec{p}_* &= \vec{b} \\
(I - M)\vec{p}_* &= \vec{b} \\
\vec{p}_* &= (I - M)^{-1}\vec{b} \quad \blacksquare
\end{aligned}$$

[3] Recall a square matrix is non-singular if it has an inverse.

Notice how similar this theorem is to the theorem for single linear recursion equations. Figure 4.3 shows how this theorem may be used with *Mathematica* to find equilibrium points for linear systems. Also note that these are the same equilibrium points that we found earlier.



```
In[22]:= M = (0.80  0.75
              0.15  0   );
         B = (5
              0);
         Id = (1  0
               0  1);
         Inverse[Id - M] . B;
         TableForm[%]
Out[26]//TableForm=
         57.1429
         8.57143
```

Figure 4.3: A *Mathematica* notebook for finding the equilibrium point

Question 8 *Look for an equilibrium point for our original humvee model in which we kept track of vehicles that were only good for parts as well as fully operational vehicles and vehicles that were in for repairs. Think about this model, its long term behavior, and its equilibrium point.*

Question 9 *Look for an equilibrium point for the model you developed for Question 7 on page 336.*

Question 10 *Look for an equilibrium point for the model you developed for Question 6 on page 336.*

It is worthwhile introducing some terminology to describe the kinds of models we have been developing in the section.

Definition 2 *A* **discrete dynamical system**, *or* **initial value problem**, *with many variables consists of*

- *A set of initial conditions describing the initial or starting value for each variable of interest. In our first example these were*

$$w_0 = 150, \quad r_0 = 0, \quad \text{and} \quad s_0 = 0.$$

- *A set of recursion equations describing how the variables of interest change. In our first example these were*

$$\begin{aligned} w_n &= 0.80w_{n-1} + 0.75r_{n-1} + 5 \\ r_n &= 0.15w_{n-1} \\ s_n &= 0.05w_{n-1} + 0.25r_{n-1} + s_{n-1} \end{aligned}$$

Definition 3 *A* **linear discrete dynamical system with many variables** *is a discrete dynamical system with many variables that can be described using matrix and vector notation in the form*

$$\begin{aligned} \vec{p}_0 &= \vec{r} \\ \vec{p}_n &= A\vec{p}_{n-1} + \vec{b} \end{aligned}$$

Our first example was a linear discrete dynamical system with many variables and could be described in the form

$$\begin{aligned} \vec{p}_0 &= \begin{bmatrix} w_0 \\ r_0 \\ s_0 \end{bmatrix} = \begin{bmatrix} 150 \\ 0 \\ 0 \end{bmatrix} \\ \vec{p}_n &= \begin{bmatrix} w_n \\ r_n \\ p_n \end{bmatrix} = \begin{bmatrix} 0.80 & 0.75 & 0 \\ 0.15 & 0 & 0 \\ 0.05 & 0.25 & 1 \end{bmatrix} \begin{bmatrix} w_{n-1} \\ r_{n-1} \\ p_{n-1} \end{bmatrix} + \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

Definition 4 *A linear discrete dynamical system with the recursion equation*

$$\vec{p}_n = A\vec{p}_{n-1} + \vec{b}$$

is said to be

- **homogeneous** *if*

$$\vec{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- **nonhomogenous** *if*

$$\vec{b} \neq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Figure 4.4: A roulette wheel

4.2 Roulette and Markov Chains

4.2.1 Roulette – A Question of Strategy

Figure 4.4 shows a roulette wheel. It has a total of 38 slots – 18 are red, 18 are black, and two are green. A player often bets on either black or red. These bets are "even money" bets. The player places the amount he wishes to bet on the table. If the ball lands in a slot whose color matches the color on which the player bet, then the croupier puts an amount equal to the player's bet on top of his bet and pushes the pile to the player. If the ball lands in a slot of a different color, then the croupier rakes in the player's bet. For example, if the player starts with \$30.00 and bets \$10.00 on one spin of the wheel on black or red then after that spin he will have either \$40.00 if he wins or \$20.00 if he loses. The player's chances of winning on each spin of the wheel are 18/38 because 18 of the slots match the color on which the bet was placed.

Suppose that a player has \$30.00 and wants to win an additional \$30.00 to give himself a total of \$60.00. We want to examine and contrast two of his many possible strategies.

- **The aggressive strategy:** The player strides confidently up to the table and places a single bet of \$30.00 on the first spin of the wheel. He either wins or loses. If he loses he smiles bravely and leaves. If he wins he smiles triumphantly, pockets his \$60.00, and leaves. With this strategy his chances of winning are 18/38 or 47.37%.
- **The conservative strategy:** The player walks hesitantly up to the table and places

a bet of \$10.00 on the first spin of the wheel. Whatever happens, he places another bet of \$10.00 on the next spin of the wheel. He continues in this way, betting \$10.00 on each spin of the wheel, until he either reaches his goal of \$60.00 or he goes broke.

This is an example of a common kind of choice that people often face. For example, investors must often decide whether to place all their money in a single investment or to diversify their holdings, placing smaller amounts in each of several investments.

Answer the questions below based on your intuition.

Question 1 *Do you think the player is more likely to win using the aggressive strategy or more likely to win using the conservative strategy?*

Question 2 *Do you think it makes much difference which strategy the player uses?*

We already know that the probability that the player wins with the aggressive strategy is 18/38, or roughly 47%. Thus, our real problem is determining the probability that the player wins with the conservative strategy.

One way you might attack this question is by experimentation. You could try the conservative strategy many times and keep track of how often you win. We will do this in class for this lesson before looking at other, more efficient, and more effective ways of tackling these questions.

While the player is playing the conservative strategy, there are seven possible situations or **states** in which he might find himself – according to how much money he has.

- **State 1:** \$0.00. He is broke and has lost the game. He is no longer playing.
- **State 2:** \$10.00. He is still playing.
- **State 3:** \$20.00. He is still playing.
- **State 4:** \$30.00. He is still playing.
- **State 5:** \$40.00. He is still playing.
- **State 6:** \$50.00. He is still playing.
- **State 7:** \$60.00. He has reached his goal and has won. He is no longer playing.

This situation involves uncertainty and probability – we cannot determine whether the player will win or lose with the conservative strategy but we can determine how likely he is to win or lose. Similarly, we cannot determine how much money the player will have after two spins of the roulette wheel (that is, what state he is in) but we can determine how likely he is to be in each state. We use a number, p, between zero and one to describe the likelihood or **probability** of an event. If $p = 1$ then the event is certain – that is, it has already occurred or it is certain that it will occur. If $p = 0$ then the event either did not occur or it is certain that it will not occur. If, for example, $p = 0.25$, then the event will occur $1/4$ of the time. For example, if you flip a fair (or balanced) coin then the probability of heads is $1/2$, or if you roll a six-sided die then the probability of coming up with a four is $1/6$.

In our situation, we need to keep track of seven probabilities – the probability of each of the seven states. For this purpose, we use seven-dimensional vectors, called **probability vectors**,

$$\vec{p} = \langle p_1, p_2, p_3, p_4, p_5, p_6, p_7 \rangle.$$

Each of the seven entries in this vector is a number between zero and one indicating the probability of the corresponding state. Because there are only seven possible states,

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7 = 1.$$

When the player starts playing he has \$30.00 and is in state 4. Thus, the initial or starting probability vector is

$$\vec{p}_0 = \langle 0, 0, 0, 1, 0, 0, 0 \rangle.$$

Notice the subscript 0 in the notation $\vec{p}_0$. This subscript indicates that we are talking about the probability after zero spins of the wheel. We will use the notation $\vec{p}_1$ for the probability vector after one spin of the wheel; the notation $\vec{p}_2$ for the probability vector after two spins of the wheel; and $\vec{p}_n$ for the probability vector after n spins of the wheel. Don't confuse the notation p_1 and $\vec{p}_1$. In this discussion the notation $\vec{p}_1$ refers to the specific vector of probabilities after one spin of the wheel. We used the notation p_1 above to denote the first element of a general probability vector, $\vec{p}$.

Question 3 *Find the probability vector $\vec{p}_1$ that describes what we can expect after one spin of the wheel.*

Question 4 *Find the probability vector $\vec{p}_2$ that describes what we can expect after two spins of the wheel.*

Question 5 *Find the probability vector $\vec{p}_3$ that describes what we can expect after three spins of the wheel.*

Figure 4.5 on page 348 is the key to analyzing the conservative strategy. It describes how a player moves from one state to another on each spin of the wheel. It is called a **transition diagram.** The circles on the left side of this transition diagram show the states that the player might be in before the spin and the circles on the right side show the states that the player might be in after the spin. The arrows indicate the possible changes and their probabilities. As one example, if a player is in state 1 then he is broke and no longer playing, so he will remain in that state. Notice there is only one arrow leading from state 1 and that arrow goes to the same state, state 1, and has probability 1. As another example, suppose that a player is in state 3 and has \$20.00. Then he will either win or lose \$10.00 when the wheel is spun. Thus, after the spin he will be in either state 2 (\$10.00) or state 4 (\$30.00). There are two arrows leading from state 3 on the left and they go to states 2 and 4. Because the probability of winning on one spin of the wheel is 18/38, the probability on the arrow going from state 3 to state 4 is 18/38. Because the probability of losing on each spin is 20/38, the probability on the arrow going from state 3 to state 2 is 20/38. This same information is shown in Table 4.3. This table is called the **transition table**.

	from state 1	from state 2	from state 3	from state 4	from state 5	from state 6	from state 7
to state 1	1	20/38	0	0	0	0	0
to state 2	0	0	20/38	0	0	0	0
to state 3	0	18/38	0	20/38	0	0	0
to state 4	0	0	18/38	0	20/38	0	0
to state 5	0	0	0	18/38	0	20/38	0
to state 6	0	0	0	0	18/38	0	0
to state 7	0	0	0	0	0	18/38	1

Table 4.3: The transition table for the conservative strategy

We can capture the same information in the **transition matrix**

$$T = \begin{bmatrix} 1 & 20/38 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 20/38 & 0 & 0 & 0 & 0 \\ 0 & 18/38 & 0 & 20/38 & 0 & 0 & 0 \\ 0 & 0 & 18/38 & 0 & 20/38 & 0 & 0 \\ 0 & 0 & 0 & 18/38 & 0 & 20/38 & 0 \\ 0 & 0 & 0 & 0 & 18/38 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 18/38 & 1 \end{bmatrix}$$

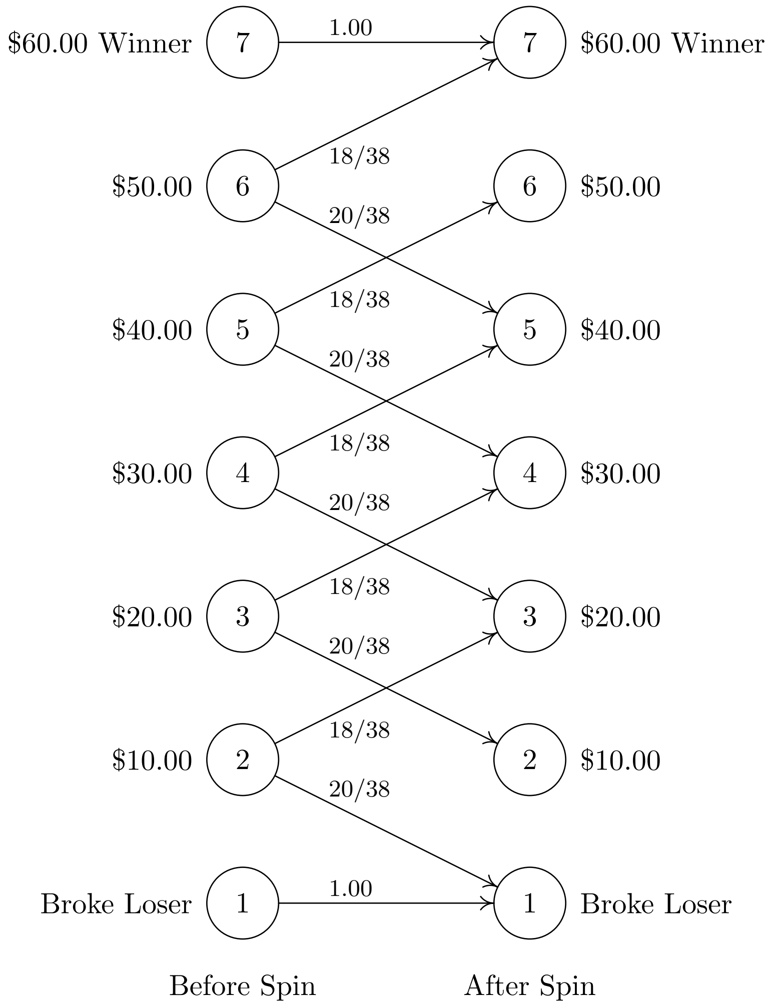


Figure 4.5: Transition diagram for one spin of the wheel

Vectors, matrices, and matrix multiplication were designed to handle this kind of problem. Suppose that

$$\vec{p} = \langle p_1, p_2, p_3, p_4, p_5, p_6, p_7 \rangle$$

denotes the probability vector before a particular spin of the wheel and that

$$\vec{q} = \langle q_1, q_2, q_3, q_4, q_5, q_6, q_7 \rangle$$

denotes the probability vector after the same spin of the wheel.

Notice that

$$\begin{aligned}
q_1 &= p_1 + (20/38)p_2 \\
q_2 &= (20/38)p_3 \\
q_3 &= (18/38)p_2 + (20/38)p_4 \\
q_4 &= (18/38)p_3 + (20/38)p_5 \\
q_5 &= (18/38)p_4 + (20/38)p_6 \\
q_6 &= (18/38)p_5 \\
q_7 &= (18/38)p_6 + p_7
\end{aligned}$$

This is exactly what we get if we perform the matrix multiplication

$$\vec{q} = T\vec{p}$$

where we think of the vectors $\vec{p}$ and $\vec{q}$ as column-vectors.

Recall that we use the notation $\vec{p}_n$ for the probability vector after n spins of the wheel playing the conservative strategy. Using this notation, we can write this information as

$$\vec{p}_n = T\vec{p}_{n-1}$$

with our starting condition

$$\vec{p}_0 = \langle 0, 0, 0, 1, 0, 0, 0 \rangle.$$

Question 6 *Compute $\vec{p}_1$ using matrix multiplication (use Mathematica) as described above. Compare your answer with your answer to Question 3.*

Question 7 *Compute $\vec{p}_2$ using matrix multiplication (use Mathematica) as described above. Compare your answer with your answer to Question 4.*

Question 8 *Find the probability vector $\vec{p}_{20}$ that describes what we can expect after 20 spins of the wheel playing the conservative strategy. What is the probability that a player playing the conservative strategy will have won after twenty spins of the wheel? What is the probability that a player playing the conservative strategy will have lost after twenty spins of the wheel? What is the probability that a player playing the conservative strategy will still be playing after twenty spins of the wheel?*

Question 9 *Compare the conservative and aggressive strategies.*

Question 10 *Another player has \$50.00. She plans to bet \$50.00 on either red or black on each spin of the wheel until she either goes broke or reaches \$150.00. What is her probability of winning? Before answering this question using the techniques above, make a rough guess. When you are done, compare your rough guess with your answer.*

Question 11 *Another player has \$50.00. She plans to bet \$50.00 on either red or black on each spin of the wheel until she either goes broke or reaches \$200.00. What is her probability of winning? Before answering this question using the techniques above, make a rough guess. When you are done, compare your rough guess with your answer.*

4.2.2 Markov Chains

Our analysis of the roulette question was based on a powerful mathematical idea called **Markov Chains**. The following definition summarizes the elements of a Markov Chain.

Definition 1 *A* **Markov Chain** *consists of the following.*

- *A set S of n* **states**. *In the roulette problem the set S had seven states.*

- *An* $n \times n$ **transition matrix**

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{bmatrix}$$

  The element t_{ij} *gives the probability of moving from state* j *to state* i *on each turn or play.*

- *An* **initial probability vector**

$$\vec{p}_0 = \langle p_1, p_2, \ldots p_n \rangle$$

  that we think of as a column vector.

With this information we define a sequence of probability vectors $\vec{p}_0, \vec{p}_1, \vec{p}_2, \ldots \vec{p}_n, \ldots$ by

$$\vec{p}_n = T\vec{p}_{n-1}.$$

Notice that $\vec{p}_n$ is the probability vector that describes the probability of being in each state after n turns or plays.

The following observations are important.

- Because the vectors $\vec{p}_n$ are all probability vectors, their entries are all between zero and one and the sum of their entries is always one.
- Because each column of the transition matrix T gives the probability of moving on one turn from one state to each of the other states, the entries in T are all between zero and one and the sum of the entries in each column is one.

Notice that a Markov Chain is also a homogeneous linear discrete system with many variables. See page 342.

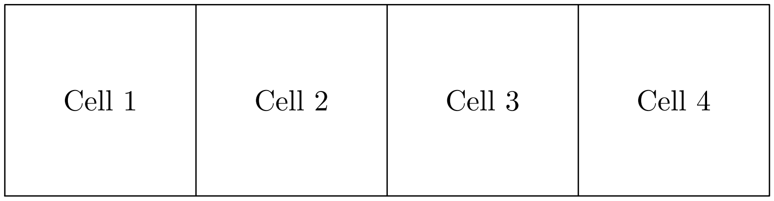


Figure 4.6: A four-celled organism

Example 1 *Figure 4.6 shows a diagram of a four-celled organism. A chemical is injected into the leftmost cell – the cell labeled "Cell 1." The molecules of this chemical move randomly between adjacent cells. Every second the probability that a given molecule moves from one cell to each of the adjacent cells is* $1/5$*. Determine the distribution of the chemical in this organism after ten seconds. Determine the distribution of the chemical in this organism after 100 seconds.*

We can set this up as a Markov Chain problem. There are four states, corresponding to Cells 1, 2, 3, and 4. The initial probability vector is $\vec{p}_0 = \langle 1, 0, 0, 0 \rangle$ because the chemical is injected initially into Cell 1. The transition matrix is

$$T = \begin{bmatrix} 0.80 & 0.20 & 0 & 0 \\ 0.20 & 0.60 & 0.20 & 0 \\ 0 & 0.20 & 0.60 & 0.20 \\ 0 & 0 & 0.20 & 0.80 \end{bmatrix}$$

Figure 4.7 on page 353 shows a *Mathematica* notebook used to study this problem. Notice that after ten seconds roughly 37.4% of the chemical is in Cell 1; roughly 29.9% is in Cell 2; roughly 19.8% is in Cell 3 and roughly 12.9% is in Cell 4. After 100 seconds, roughly 25% of the chemical is in each cell.

Question 12 *Figure 4.8 on page 354 shows another four-celled organism. A chemical is injected into Cell 1. The molecules of this chemical move randomly between cells that share a side. Every second the probability that a given molecule moves from one cell to each of the cells with which it shares a side is* $1/5$*. Determine the distribution of the chemical in this organism after ten seconds. Determine the distribution of the chemical in this organism after 100 seconds.*

```
In[11]:= Clear[p]
```

$$\texttt{p[0]} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix};$$

$$\texttt{T} = \begin{pmatrix} 0.80 & 0.20 & 0 & 0 \\ 0.20 & .60 & 0.20 & 0 \\ 0 & 0.20 & 0.60 & 0.20 \\ 0 & 0 & 0.20 & 0.80 \end{pmatrix};$$

```
p[n_] := p[n] = T.p[n - 1]

MatrixForm[p[10]]
```

Out[15]//MatrixForm=

$$\begin{pmatrix} 0.374266 \\ 0.299333 \\ 0.197644 \\ 0.128757 \end{pmatrix}$$

```
In[16]:= MatrixForm[p[100]]
```

Out[16]//MatrixForm=

$$\begin{pmatrix} 0.250002 \\ 0.250001 \\ 0.249999 \\ 0.249998 \end{pmatrix}$$

Figure 4.7: Using *Mathematica* to study a four-celled organism

Question 13 *Collegetown has three pizza restaurants – Tony's, Maria's, and Papa Dave's. At the beginning of the year they each have 1/3 of the students as customers. Each week during the year 1/5 of Tony's customers switch to Maria's and 1/5 switch to Papa Dave's. The remainder stay with Tony's. Each week during the year 1/6 of Maria's customers switch to Tony's; 1/10 switch to Papa Dave's and the remainder stay with Maria's. Each week during the year 1/2 of Papa Dave's customers switch to Tony's; 1/3 switch to Maria's and the remainder stay with Papa Dave's. What percentage of of the students are customers at each pizza restaurant after ten weeks? What percentage of of the students are customers at each pizza restaurant after 25 weeks?*

4.2.3 A Closed Form Solution for a Homogeneous Linear Discrete Dynamical System with Many Variables

We have seen several examples of homogeneous linear discrete dynamical systems with many variables. If we use the notation

$$\vec{p}_n = A\vec{p}_{n-1}, \quad \vec{p}_0 = \vec{c}$$

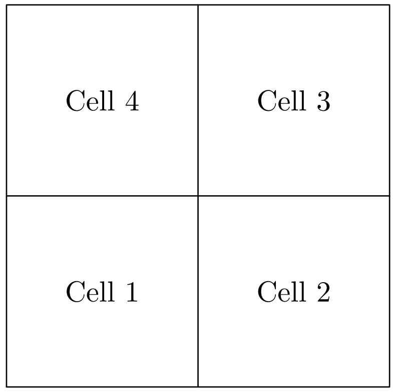


Figure 4.8: Another four-celled organism

then notice that

$$
\begin{aligned}
\vec{p}_1 &= A\vec{p}_0 = A\vec{c} \\
\vec{p}_2 &= A\vec{p}_1 = A(A\vec{c}) = A^2\vec{c} \\
\vec{p}_3 &= A\vec{p}_2 = A(A^2\vec{c})) = A^3\vec{c} \\
&\vdots \\
\vec{p}_n &= A^n\vec{c}
\end{aligned}
$$

giving us a closed from solution

$$\vec{p}_n = A^n\vec{c}.$$

Note that

$$A^n = \underbrace{A \cdot A \cdots A}_{n-\text{times}}.$$

and $\vec{c}$ is the initial value – that is, $\vec{p}_0 = \vec{c}$.

Later in this chapter we will develop closed form solutions that make it easier to understand long term behavior. For now, however, note that one way to gain some understanding of long term behavior is by looking at the matrices A^n for very large values of n. Although this involves lots of matrix multiplication, the combination of computer power and a bit of cleverness can make the calculations easy. Consider the transition matrix

$$T = \begin{bmatrix} 1 & 20/38 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 20/38 & 0 & 0 & 0 & 0 \\ 0 & 18/38 & 0 & 20/38 & 0 & 0 & 0 \\ 0 & 0 & 18/38 & 0 & 20/38 & 0 & 0 \\ 0 & 0 & 0 & 18/38 & 0 & 20/38 & 0 \\ 0 & 0 & 0 & 0 & 18/38 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 18/38 & 1 \end{bmatrix}$$

from our analysis of the conservative roulette strategy. We can compute

$$\begin{aligned} T^2 &= T \cdot T \\ T^4 &= (T^2) \cdot (T^2) \\ T^8 &= (T^4) \cdot (T^4) \\ &\vdots \\ T^{1024} &= (T^{512}) \cdot (T^{512}) \end{aligned}$$

so with just ten matrix multiplications we can compute T^{1024} instead of the 1023 matrix multiplications it would take it we computed $\underbrace{T \cdot T \cdot T \cdots T \cdot T}_{1024 \text{ factors}}$. The result is

$$T^{1024} = \begin{bmatrix} 1.000000 & 0.873977 & 0.733952 & 0.578369 & 0.405499 & 0.213420 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.126023 & 0.266048 & 0.421631 & 0.594501 & 0.786580 & 1.000000 \end{bmatrix}$$

This matrix can be used to determine what happens after 1024 spins of the wheel, since

$$\vec{p}_{1024} = T^{1024} \vec{p}_0$$

Recall that our gambler started with \$30.00 and was in state 4 with probability 1. Thus

$$\vec{p}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

and after 1024 spins

$$\vec{p}_{1024} = \begin{bmatrix} 1.000000 & 0.873977 & 0.733952 & 0.578369 & 0.405499 & 0.213420 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.126023 & 0.266048 & 0.421631 & 0.594501 & 0.786580 & 1.000000 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.578369 \\ 0.000000 \\ 0.000000 \\ 0.000000 \\ 0.000000 \\ 0.000000 \\ 0.421631 \end{bmatrix}.$$

Thus, with probability 57.8% our gambler is broke and with probability 42.2% our gambler has won and now has \$60.00.

Question 14 *Suppose the gambler starts with \$20.00. What is the probability that after many spins of the wheel the gambler has won and what is the probability that after many spins of the wheel the gambler has lost?*

Question 15 *Suppose the gambler starts with \$40.00. What is the probability that after many spins of the wheel the gambler has won and what is the probability that after many spins of the wheel the gambler has lost?*

4.3 Age Dependent Population Models and Leslie Matrices

4.3.1 A Simple Example

We begin this section with a simple example.

Example 1 *You are advising the mayor of a small town that was formerly under enemy control. Retreating enemy forces poisoned a lake on which the town depended for food. You have neutralized the poisons in the lake and now need to restock the lake to replace the fish that were killed.*

You have consulted with biologists and have learned that one good way to model the population for this particular species of fish is to divide the population into two groups – fish under one year old and fish that are one year or more in age. Each year starting with the year 2008 we will use a two-dimensional column-vector to represent the fish population. We will use the notation $\vec{p}_{2008}$ for the vector representing the population in the current year, $\vec{p}_{2009}$ for the vector representing the population in the year 2009, and so forth. You have stocked the lake with six tons of fish, all very young and under the age of one year. Thus,

$$\vec{p}_{2008} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}.$$

Notice the first entry in this vector represents the fact that there are 6 tons of fish under the age of one year and the second entry represents the fact that there are no fish that are aged one year or more.

Your advisor reports that each year the reproduction rate for fish under the age of one year is 0.50. This means that next year you can expect to have 3 tons of fish under the age of one year. She also reports that each year the survival rate for fish that are under one year is 0.60. This means that next year you can expect to have 3.6 tons of fish that are one year old or more. Thus,

$$\vec{p}_{2009} = \begin{bmatrix} 3.0 \\ 3.6 \end{bmatrix}$$

For subsequent years you also need to know the reproduction and survival rates for fish that are one year or more old. Your biologist reports that the reproduction rate for fish in this age group is 0.60 and the survival rate is 0.40. Thus, we can model this situation by letting

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix}$$

and

$$\vec{p}_n = A\vec{p}_{n-1}.$$

Question 1 *Express the recursion equation above using a pair of equations and the notation*

$$\vec{p}_n = \langle a_n, b_n \rangle \quad \text{and} \quad \vec{p}_{n-1} = \langle a_{n-1}, b_{n-1} \rangle.$$

Question 2 *Explain where each entry in the matrix A comes from. Be careful. Notice that two of the entries, a_{12} and a_{21}, are the same, namely 0.60. This is because the survival rate for age group 1 is the same as the fertility rate for age group 2. Which of these two entries represents the survival rate for age group 1 and which represents the fertility rate for age group 2?*

Notice that this is a homogeneous linear discrete dynamical system with many variables and

$$\begin{aligned}
\vec{p}_{2009} &= A\vec{p}_{2008} \\
\vec{p}_{2010} &= A\vec{p}_{2009} = A(A\vec{p}_{2008}) = A^2\vec{p}_{2008} \\
\vec{p}_{2011} &= A\vec{p}_{2010} = A(A^2\vec{p}_{2008}) = A^3\vec{p}_{2008} \\
&\vdots \\
\vec{p}_n &= A^{n-2008}\vec{p}_{2008},
\end{aligned}$$

giving us a closed-form solution for this problem.

Figure 4.9 on page 360 can provide some visual insight into the long term behavior of this model. Each dot in this figure represents the fish population for one year. Each dot has two coordinates – an x-coordinate and a y-coordinate. The x-coordinate is the fish population under age one year and the y-coordinate is the fish population aged one year and above. Thus, $\vec{p}_{2008}$ is represented by the point $(6, 0)$ and p_{2009} by the point $(3.0, 3.6)$.

Looking at Figure 4.9, we see that after the first few years the populations appear to be marching out along a straight line. We can also look at this model using Table 4.4 on page 360. You can modify this spreadsheet[4] to produce graphs like Figure 4.9 and tables

[4]http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_IV/age-dependent-population.xls

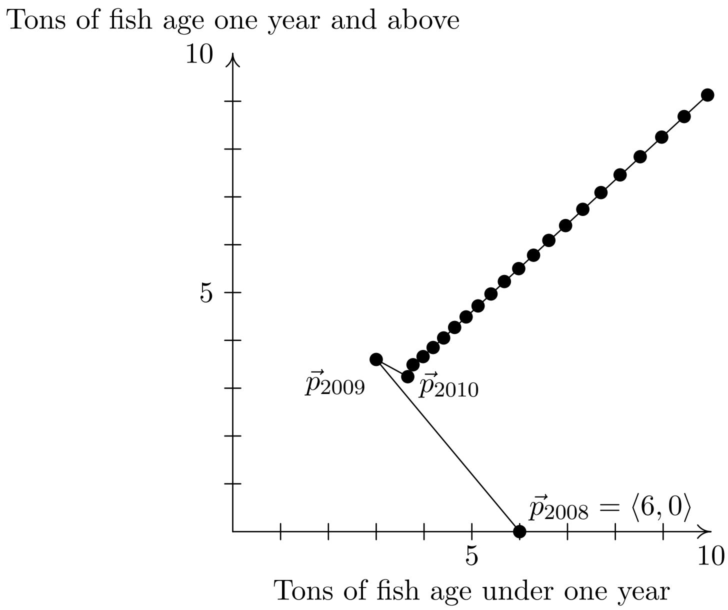


Figure 4.9: Representing the Fish Population Visually

Year	**Age** < 1	$1 \leq$ **Age**
2008	6.00000	0.00000
2009	3.00000	3.60000
2010	3.66000	3.24000
2011	3.77400	3.49200
2012	3.98220	3.66120
2013	4.18782	3.85380
2014	4.40619	4.05421
2015	4.63562	4.26540
2016	4.87705	4.48753
2017	5.13104	4.72124
2018	5.39827	4.96712
2019	5.67941	5.22581
2020	5.97519	5.49797
2021	6.28638	5.78430
2022	6.61377	6.08555
2023	6.95821	6.40248
2024	7.32060	6.73592
2025	7.70185	7.08673
2026	8.10296	7.45580
2027	8.52496	7.84410

Table 4.4: Twenty Years

like Table 4.4. Now, look at the population for three consecutive years after a few years have passed,

$$\vec{p}_{2016} = \begin{bmatrix} 4.87705 \\ 4.48753 \end{bmatrix} \qquad \vec{p}_{2017} = \begin{bmatrix} 5.13105 \\ 4.72124 \end{bmatrix} \qquad \vec{p}_{2018} \begin{bmatrix} 5.39827 \\ 4.96712 \end{bmatrix}.$$

A little arithmetic shows that

$$\vec{p}_{2017} = 1.05208\ \vec{p}_{2016} \quad \text{and} \quad \vec{p}_{2018} = 1.05208\ \vec{p}_{2017}.$$

If you look at the next few years, you will see that this pattern appears to continue.

$$\vec{p}_n = 1.05208\vec{p}_{n-1}$$

This is why the vectors seem to be marching out along a straight line in Figure 4.9. This gives us a lot of information about the long term behavior of the fish population in this lake. According to this model, the total fish population will continue to rise at the rate of roughly 5.21% per year and the percentage of the population that is one year or younger will stabilize at about 52%. In practice, of course, this model cannot continue to predict the population forever because the lake is finite and can only support a finite fish population.

4.3.2 Real People and Real Countries

Figure 4.10 on page 362 shows the estimated United States population in the year 2000 and predictions for the year 2025 broken down by age and gender. This data was obtained from the International Data Base maintained by the United States Census Bureau. Figure 4.11 on page 363 shows the same information for India obtained from the same source. These figures are called **population pyramids**. Information like this can be tremendously important for public policy decisions. For example, if you compare the percentage of the population aged 70 and over in the year 2000 with the same percentage in the year 2025 you will see immediately why people are worried about skyrocketing costs for Social Security and Medicare.

The population pyramid for India in the year 2000 is a typical example of a population pyramid for an underdeveloped country. Because survival rates are so low, the population drops rapidly as age increases. The population pyramid for India for the year 2025 is a typical population pyramid for a developing country. As the public health improves, survival rates improve and we see smaller population decreases as age increases. You can

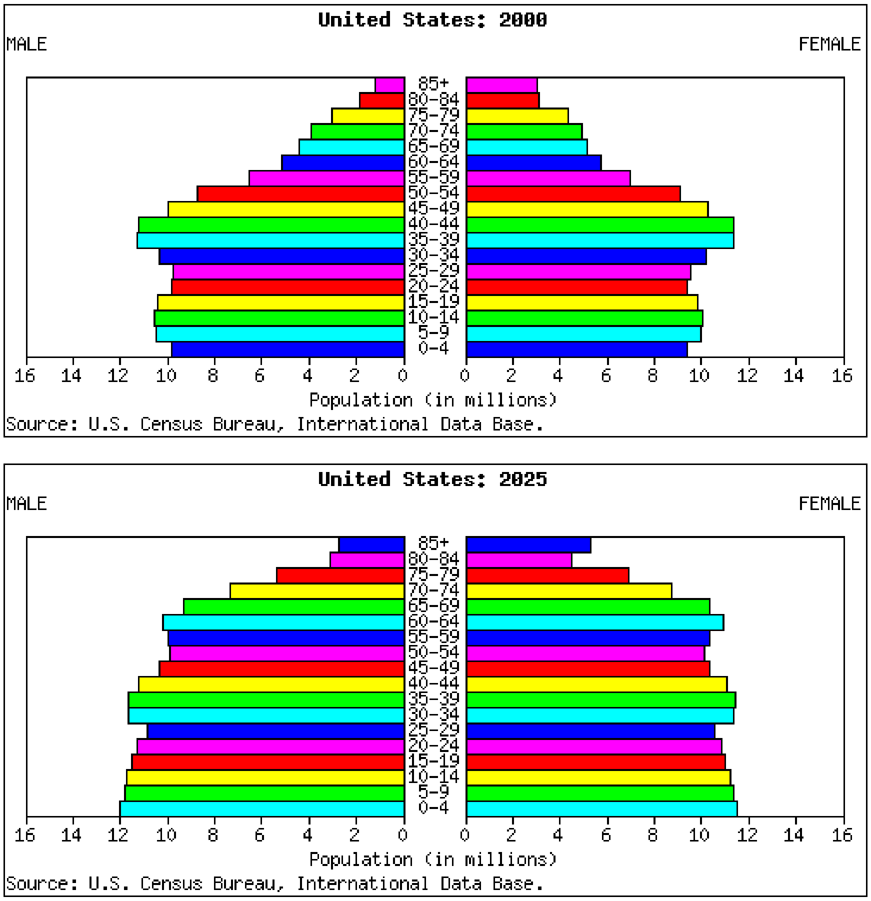


Figure 4.10: Population Pyramids for the United States 2000 and 2025

see this effect clearly by comparing the population pyramids for India for the years 2000 and 2025. Note that the population pyramids for the United States and India for the year 2025 are based on predictions not on actual facts. So demographers are predicting that over the next twenty years India will develop rapidly.

As we build population models we consider many factors. The two most important factors are fertility rates and survival rates. Immigration and emigration are also important. In this section we build models that include survival rates and fertility rates but not immigration and emigration. Thus, our models will be better for countries that have low rates of immigration and emigration than for countries that have higher rates of immigration and emigration. We will also simplify our models by breaking population down by age but not by gender.

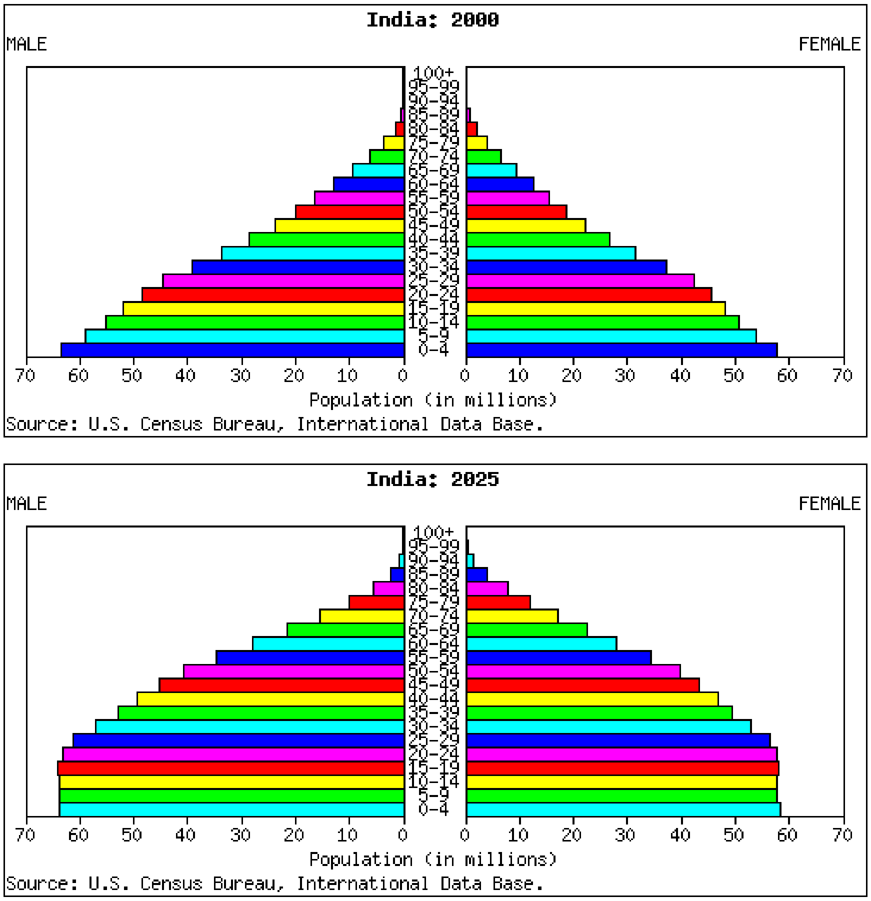


Figure 4.11: Population Pyramids for India 2000 and 2025

Our models will have 17 age groups. The first age group has everyone from birth through just before their 5^{th} birthday. We write this as age group 0–4. The second age group has everyone from their 5^{th} birthday through just before their 10^{th} birthday. This age group is written as age group 5–9. The last age group has everyone whose has passed their 80^{th} birthday and is written as age group 80^{+}. For each year we write the population as a 17-dimensional vector. For example, Table 4.5 on page 364 has the population figures for the United States in midyear 2008 and the same information is represented by the vector

$$\vec{p}_{2008} = \langle 21,009,914, \quad 20,155,574, \quad 19,981,256, \quad \ldots 11,409,264 \rangle.$$

0-4	21,009,914
5-9	20,155,574
10-14	19,981,265
15-19	21,728,978
20-24	21,186,421
25-29	21,161,376
30-34	19,531,264
35-39	20,909,399
40-44	21,428,750
45-49	22,858,209
50-54	21,463,268
55-59	18,580,896
60-64	15,139,163
65-69	11,321,863
70-74	8,732,349
75-79	7,226,693
80+	11,409,264

Table 4.5: 2008 United States Total Midyear Population by Age Group

We will build a model that looks like

$$\vec{p}_{n+5} = A\vec{p}_n \qquad \text{or} \qquad \vec{p}_{n+5} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,17} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,17} \\ \vdots & \vdots & & \vdots \\ a_{17,1} & a_{17,2} & \cdots & a_{17,17} \end{bmatrix} \vec{p}_n$$

where the vectors $\vec{p}_n$ and $\vec{p}_{n+5}$ are written as column vectors. Instead of estimating the population for every year, we will work with five year intervals. That is why we write $\vec{p}_{n+5} = A\vec{p}_n$ instead of $\vec{p}_{n+1} = A\vec{p}_n$.

The first entry (age group 0-4) in the vector $\vec{p}_{n+5}$ is based on the first row of the matrix A and the vector $\vec{p}_n$. Because we are building a model with no immigration or emigration, the people in age group 0-4 will come from births. The entries in the first row of the matrix A are birth rates. Because we are working with five year intervals, these are birth rates per person for five years. Recall also that we are lumping males and females together. In the first row the entry $a_{1,k}$ is the average number of children born in each five year period to each person. For example, if every female in age group 5 (that's 20-24 years old) had three children in five years and if the percentage of women in that age group was 50% then $a_{1,5}$ would be 1.5. Usually $a_{1,1}$ is zero because people aged 0-4 years are too young to have children within five years. Usually $a_{1,2}$ is small but not zero because some of the women who start in that age group will be 14 years old after five years and will have had some children. Toward the end of the first row the entries $a_{1,k}$ are smaller because older people tend to have fewer children. The largest entries in the first row are in the middle of the row.

Most of the remaining entries in the matrix are zero. The only nonzero entries are the entry $a_{17,17}$ and the entries $a_{k+1,k}$, for $k = 1, 2, \ldots 16$. The entry $a_{17,17}$ is the five year survival rate for people aged 80^+. People in that age group who survive remain in the same age group. The entries $a_{k+1,k}$ for $k = 1, 2, \ldots 16$ are the five year survival rates for age group k because all the people in those age groups who survive for five years move up one age group. Thus our matrix looks like

$$A = \begin{bmatrix} 0 & a_{1,2} & a_{1,3} & a_{1,4} & \cdots & a_{1,15} & a_{1,16} & a_{1,17} \\ a_{2,1} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & a_{3,2} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & a_{4,3} & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{5,4} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & a_{17,16} & a_{17,17} \end{bmatrix}.$$

The matrix A is called a **Leslie** matrix.

Because demographers work with computers they can handle 81-dimensional vectors and $(81 \times 81)-$dimensional matrices easily. Thus, they usually use 81 age groups rather than 17 age groups. To save paper they often report their results using 17 age groups. Some countries break populations into different age groups. For example, Figure 4.11 on page 363 has 21 age groups.

To save paper we will continue this section with a simpler model, using five age groups: 0-19, 20-39, 40-59, 60-79, and 80^+. We will use a plausible but made-up initial population vector $\vec{p}_{2008} = \langle 10,000,000, \quad 9,800,000, \quad 9,600,000, \quad 9,000,000, \quad 8,400,000 \rangle$ and made up but plausible fertilty rates and survival rates to give us the Leslie matrix

$$A = \begin{bmatrix} 0.00 & 0.1 & 0.4 & 0.5 & 0.1 \\ 0.95 & 0 & 0 & 0 & 0 \\ 0 & 0.96 & 0 & 0 & 0 \\ 0 & 0 & 0.90 & 0 & 0 \\ 0 & 0 & 0 & 0.85 & 0.50 \end{bmatrix}$$

and the model

$$\vec{p}_{n+20} = A\vec{p}_n, \qquad \vec{p}_{2008} = \begin{bmatrix} 10,000,000 \\ 9,800,000 \\ 9,600,000 \\ 9,000,000 \\ 8,400,000 \end{bmatrix}.$$

Question 3 *Describe the recursion equation above using five equations and the notation*

$$\vec{p}_{n+20} = \langle a_{n+20}, b_{n+20}, c_{n+20}, d_{n+20}, e_{n+20} \rangle \quad \text{and} \quad \vec{p}_n = \langle a_, b_n, c_n, d_n, e_n \rangle.$$

Question 4 *Using the same model predict the population in the year 2028. Find the percentage of the population in each age group in that year.*

Question 5 *Using the same model predict the population in the year 2048. Find the percentage of the population in each age group in that year.*

Question 6 *Using the same model predict the population in the year 2068. Find the percentage of the population in each age group in that year.*

Question 7 *Using the same model predict the population in the year 2208. Find the percentage of the population in each age group in that year.*

Question 8 *Using the same model predict the population in the year 2228. Find the percentage of the population in each age group in that year.*

Question 9 *Using the same model predict the total population in the year 2028.*

Question 10 *Using the same model predict the percentage rise in total population in each 20 year period starting in the year 2008.*

Question 11 *Use the techniques we developed in subsection 4.2.3 to investigate the long term behavior of this model.*

The age dependent population models we have developed in this section are homogeneous linear discrete dynamical systems with many variables. See page 342 for this definition.

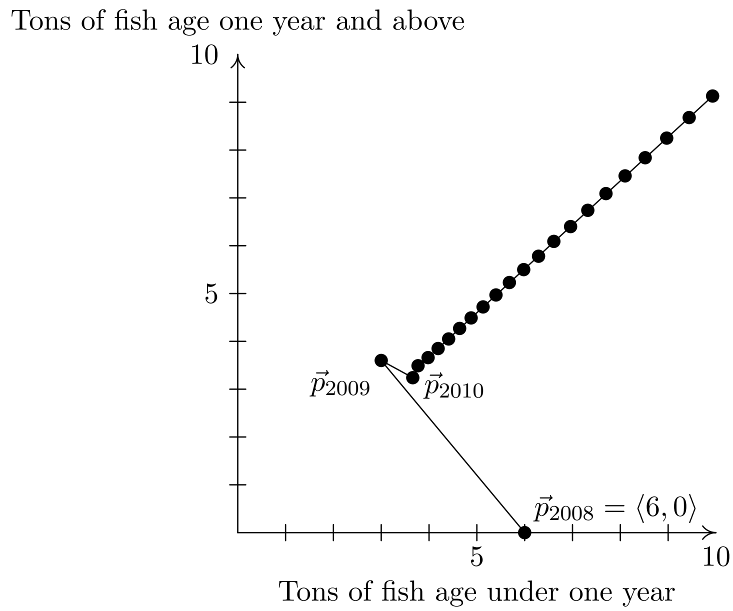


Figure 4.12: Representing the Fish Population Visually

4.4 Eigenvalues and Eigenvectors, I

In this section we begin the development of one of the most powerful tools for understanding systems like the ones we've been studying. This tool goes by the name of "eigenvalues and eigenvectors." We begin by recalling Example 1 from Section 4.3 on page 358. You may want to look back at that example to recall the context. In this section we focus on the mathematics. Look at Figure 4.12 above and Table 4.6 on page 369. Both of these refer to the model

$$\vec{p}_n = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix} \vec{p}_{n-1}, \qquad \vec{p}_{2008} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}.$$

Notice that in Figure 4.12 the sequence of vectors $\vec{p}_{2008}, \vec{p}_{2009}, \vec{p}_{2010}, \ldots$ appears, after a somewhat shaky start, to be marching out along a straight line. That is, except for the first few vectors in this sequence, they all seem to have about the same direction and the only thing that is changing is their magnitude. We can also see this from Table 4.6 on page 369. Notice that, for example,

$$\vec{p}_{2027} = \begin{bmatrix} 8.52496 \\ 7.84410 \end{bmatrix} = 1.05208 \begin{bmatrix} 8.10296 \\ 7.45580 \end{bmatrix} = 1.05208\vec{p}_{2026}.$$

Year	**Age** < 1	$1 \leq$ **Age**
2008	6.00000	0.00000
2009	3.00000	3.60000
2010	3.66000	3.24000
2011	3.77400	3.49200
2012	3.98220	3.66120
2013	4.18782	3.85380
2014	4.40619	4.05421
2015	4.63562	4.26540
2016	4.87705	4.48753
2017	5.13104	4.72124
2018	5.39827	4.96712
2019	5.67941	5.22581
2020	5.97519	5.49797
2021	6.28638	5.78430
2022	6.61377	6.08555
2023	6.95821	6.40248
2024	7.32060	6.73592
2025	7.70185	7.08673
2026	8.10296	7.45580
2027	8.52496	7.84410

Table 4.6: Twenty Years

Since

$$\vec{p}_n = A\vec{p}_{n-1}$$

where

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix},$$

this means that

$$\vec{p}_{2027} = A\vec{p}_{2026} = 1.05208\vec{p}_{2026}.$$

This vector, $\vec{p}_{2026}$ and the number 1.05208, give us an example of an associated eigenvector and eigenvalue – multiplying $\vec{p}_{2026}$ by the matrix A (on the left) is the same as multiplying it by the scalar 1.05208. Here is a more formal definition.

Definition 1 *If we have an $(n \times n)$-matrix A and there is a nonzero vector $\vec{v}$ and a constant*[5] *λ such that*

$$A\vec{v} = \lambda\vec{v},$$

then we say that $\vec{v}$ is an **eigenvector** *for the matrix A and λ is its associated* **eigenvalue***.*

If $\vec{v}$ is an eigenvector for a matrix A then multiplying $\vec{v}$ by A (on the left) has the same effect as multiplying it by the associated eigenvalue λ – that is stretching or shrinking it by λ.[6] Click here[7] to open a new window with a live diagram that you can use to see what eigenvalues and eigenvectors mean visually. This diagram will enable you to explore the possible eigenvalues and eigenvectors of any (2×2)-matrix. It is initially set up to explore the matrix

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix}.$$

You should start by exploring this matrix but later on you may want to explore other matrices by editing the entries in the usual way. Notice the blue dot at the point

$$\vec{x} = \begin{bmatrix} 6.00 \\ 0.00 \end{bmatrix}$$

and the red dot at the point

$$\vec{y} = A\vec{x} = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix} \begin{bmatrix} 6.00 \\ 0.00 \end{bmatrix} = \begin{bmatrix} 3.00 \\ 3.60 \end{bmatrix}.$$

Question 1 *Click and drag the blue dot to try different possible values for the vector $\vec{x}$. The red dot will automatically move to show the new vector $\vec{y} = A\vec{x}$. Play with this for a bit, moving the vector $\vec{x}$ around to see what happens to the vector $\vec{y}$. See if you can find a vector $\vec{x}$ such that the vector $\vec{y}$ points in the same direction as the vector $\vec{x}$. See if you can find a vector $\vec{x}$ such that the vector $\vec{y}$ points in the opposite direction of the vector $\vec{x}$. Congratulations!! You've just found your first two eigenvectors.*

---

[5]it is traditional to use the Greek letter "lambda" in this situation.

[6]If λ is negative the direction of $\vec{v}$ is also reversed.

[7]http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_IV/population-growth.html

When the vector $\vec{y}$ points in the same direction as the vector $\vec{x}$, then

$$\vec{y} = A\vec{x} = \lambda\vec{x},$$

where λ is a positive number. That is, $\vec{y}$ is a positive multiple of $\vec{x}$.

When the vector $\vec{y}$ points in the opposite direction of the vector $\vec{x}$, then

$$\vec{y} = A\vec{x} = \lambda\vec{x},$$

where λ is a negative number. That is, $\vec{y}$ is a negative multiple of $\vec{x}$.

This is the geometric idea underlying eigenvalues and eigenvectors. If $\vec{x}$ is an eigenvector of the matrix A and its associated eigenvalue is positive, then the vector $\vec{y} = A\vec{x}$ points in the same direction as the vector $\vec{x}$. If $\vec{x}$ is an eigenvector and its associated eigenvalue is negative, then the vector $\vec{y} = A\vec{x}$ points in the opposite direction. The magnitude of the eigenvalue tells us how much the eigenvector is stretched or shrunk. Use the same live diagram to estimate the eigenvalues of the matrix

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix}.$$

Example 1 *We closed section 4.3 with a slightly more complicated age dependent population model. We used five age groups: 0-19, 20-39, 40-59, 60-79, and* 80^{+}*. This example is similar to that example but with different numbers. We use the same initial population vector*

$$\begin{bmatrix} 10,000,000 \\ 9,800,000 \\ 9,600,000 \\ 9,000,000 \\ 8,400,000 \end{bmatrix}$$

and the Leslie matrix

$$A = \begin{bmatrix} 0.00 & 0.20 & 0.50 & 0.60 & 0.20 \\ 0.95 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.96 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.90 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.85 & 0.70 \end{bmatrix}.$$

This gives us the model

$$\vec{p}_{n+20} = A\vec{p}_n, \qquad \vec{p}_{2008} = \begin{bmatrix} 10,000,000 \\ 9,800,000 \\ 9,600,000 \\ 9,000,000 \\ 8,400,000 \end{bmatrix}.$$

Beware that this model has a twist that might be confusing – we are working with 20 year intervals. Calculating predicted population through the year 2428 we see that

$$\vec{p}_{2408} = \begin{bmatrix} 10,997,900 \\ 9,328,600 \\ 7,994,750 \\ 6,423,990 \\ 13,001,100 \end{bmatrix} \quad \text{and} \quad \vec{p}_{2428} = \begin{bmatrix} 12,317,700 \\ 10,448,000 \\ 8,955,460 \\ 7,195,270 \\ 14,561,200 \end{bmatrix}.$$

If we divide each entry in the vector $\vec{p}_{2428}$ by the corresponding entry in the vector $\vec{p}_{2408}$ we see that

$$\begin{aligned} \frac{12,317,700}{10,997,900} &= 1.120 \\ \frac{10,448,000}{9,328,600} &= 1.120 \\ \frac{8,955,460}{7,994,750} &= 1.120 \\ \frac{7,195,270}{6,423,990} &= 1.120 \\ \frac{14,561,200}{13,001,100} &= 1.120 \end{aligned}$$

and

$$\vec{p}_{2428} = \begin{bmatrix} 12,317,700 \\ 10,448,000 \\ 8,955,460 \\ 7,195,270 \\ 14,561,200 \end{bmatrix} = A\vec{p}_{2408} = A \begin{bmatrix} 10,997,900 \\ 9,328,600 \\ 7,994,750 \\ 6,423,990 \\ 13,001,100 \end{bmatrix} = 1.120 \begin{bmatrix} 10,997,900 \\ 9,328,600 \\ 7,994,750 \\ 6,423,990 \\ 13,001,100 \end{bmatrix}.$$

This gives us another example of an associated eigenvalue and eigenvector

$$\lambda = 1.120, \qquad \vec{v} = \begin{bmatrix} 10,997,900 \\ 9,328,600 \\ 7,994,750 \\ 6,423,990 \\ 13,001,100 \end{bmatrix}.$$

The following theorem has important practical implications for these kinds of models.

Theorem 2 *If A is an $(n \times n)$-matrix and $\vec{v}$ is an eigenvector for A with associated eigenvalue, λ, then if c is any constant, the vector $c\vec{v}$ is also an eigenvector and has the same associated eigenvalue.*

Proof

$$A(c\vec{v}) = cA\vec{v} = c(\lambda\vec{v}) = \lambda(c\vec{v}) \ \blacksquare$$

Example 2 *The vector $\vec{v} = \langle 1, 1 \rangle$ is an eigenvector of the matrix*

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

with an associated eigenvalue 3, since

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The scalar multiple $\langle 20, 20\rangle$ of $\langle 1, 1\rangle$ is also an eigenvector with the same associated eigenvalue since

$$A\begin{bmatrix} 20 \\ 20 \end{bmatrix} = \begin{bmatrix} 60 \\ 60 \end{bmatrix} = 3\begin{bmatrix} 20 \\ 20 \end{bmatrix}.$$

Recall the significance of equilibrium points – when you start at an equilibrium point you stay there. As a consequence of the preceding theorem, eigenvectors have a similar property. When you start at an eigenvector, each subsequent term is also an eigenvector and is obtained by multiplying the preceding term by the associated eigenvalue. For example, we began this section with the model

$$\vec{p}_n = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix}\vec{p}_{n-1}, \qquad \vec{p}_{2008} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

that described the fish population in a particular lake.

We saw that

$$\vec{p}_{2026} = \begin{bmatrix} 8.10296 \\ 7.45580 \end{bmatrix}$$

was an eigenvector[8] with associated eigenvalue $\lambda = 1.05208$. This implies that

$$\begin{aligned} \vec{p}_{2027} &= \lambda\vec{p}_{2026} \\ \vec{p}_{2028} &= \lambda\vec{p}_{2027} = \lambda^2\vec{p}_{2026} \\ \vec{p}_{2029} &= \lambda\vec{p}_{2028} = \lambda^3\vec{p}_{2026} \\ &\vdots \\ \vec{p}_{2026+n} &= \lambda^n\vec{p}_{2026}, \end{aligned}$$

since by Theorem 2 the vectors $\lambda\vec{p}_{2025}, \lambda^2\vec{p}_{2026}\ldots$ are all eigenvectors with the same eigenvalue.

[8] This might not be exactly an eigenvector but it is very, very close. Close enough so the difference is less than rounding error.

Remember this discussion is focused on long term behavior. The formula

$$\vec{p}_{2026+n} = \lambda^n \vec{p}_{2026}$$

only applies when n is positive and we are looking at the long term results for our original model. We cannot, for example, use this formula to compute $\vec{p}_{2008}$ by using $n = -18$.

This associated eigenvector and eigenvalue tell us two things about the long term behavior of this particular fish population.

- Because the eigenvalue is $\lambda = 1.05280$, the population increases by 5.280% each year. This increase is the same in both age groups.
- Because the eigenvector is $(8.10296, 7.45580)$ the age distribution in the long term is

$$\begin{aligned} \text{age group 1} &= \frac{8.10296}{8.10296 + 7.45580} = 52.08\% \\ \text{age group 2} &= \frac{7.45580}{8.10296 + 7.45580} = 47.92\% \end{aligned}$$

  This kind of information is very important. The rate at which the population is growing determines how quickly food and energy supplies, for example, must grow. The age distribution impacts how much the expenses associated with certain age groups – school for young children, for example, and health care for the elderly – affect the working population.

If an eigenvector has the associated eigenvalue $\lambda = 1$ then it is an equilibrium point because multiplying a vector by 1 leaves it unchanged.

Question 2 *Analyze the long term behavior of Example 1 using eigenvectors and eigenvalues as we did in our discussion of the two age fish population model.*

Question 3 *Consider the model*

$$\vec{p}_n = \begin{bmatrix} 0.60 & 0.70 \\ 0.75 & 0.35 \end{bmatrix} \vec{p}_{n-1}, \qquad \vec{p}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Compute the first 20 terms. Based on your computations find an eigenvector and eigenvalue. Describe the long term behavior using the eigenvector and eigenvalue that you found.

Question 4 *Consider the model*

$$\vec{p}_n = \begin{bmatrix} 0.60 & 0.60 \\ 0.75 & 0.35 \end{bmatrix} \vec{p}_{n-1}, \qquad \vec{p}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Compute the first 20 terms. Based on your computations find an eigenvector and eigenvalue. Describe the long term behavior using the eigenvector and eigenvalue that you found.

Question 5 *Consider the model*

$$\vec{p}_n = \begin{bmatrix} 0.60 & 0.50 \\ 0.75 & 0.35 \end{bmatrix} \vec{p}_{n-1}, \qquad \vec{p}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Compute the first 20 terms. Based on your computations find an eigenvector and eigenvalue. Describe the long term behavior using the eigenvector and eigenvalue that you found.

Question 6 *Consider the model*

$$\vec{p}_n = \begin{bmatrix} 0.60 & 0.40 \\ 0.75 & 0.35 \end{bmatrix} \vec{p}_{n-1}, \qquad \vec{p}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Compute the first 20 terms. Based on your computations find an eigenvector and eigenvalue. Describe the long term behavior using the eigenvector and eigenvalue that you found.

Question 7 *Consider the model*

$$\vec{p}_n = \begin{bmatrix} 0.60 & 0.40 \\ 0.75 & 0.20 \end{bmatrix} \vec{p}_{n-1}, \qquad \vec{p}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Compute the first 20 terms. Based on your computations find an eigenvector and eigenvalue. Describe the long term behavior using the eigenvector and eigenvalue that you found.

4.5 Eigenvalues and Eigenvectors, II

In this section we see how to find eigenvectors and eigenvalues "by hand" and look at examples with (2×2)-matrices. In the next section we see how to find eigenvalues and eigenvectors using *Mathematica*.

Notice that in our discussion of eigenvalues and eigenvectors we have been writing vectors as column-vectors. Also, recall the idea of an identity matrix. An identity matrix is an $(n \times n)$-matrix of the form

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Recall that if $\vec{x}$ is an n-dimensional column-vector and I is an $(n \times n)$ identity matrix, then

$$I\vec{x} = \vec{x}.$$

In addition, recall that if

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is a (2×2)-matrix then we can compute its determinant, $a_{11}a_{22}-a_{12}a_{21}$, and that the matrix A is invertible (non-singular) if its determinant is nonzero and non-invertible (singular) if its determinant is zero. For $(n \times n)$-matrices the determinant is harder to compute but serves the same function. Thus, we can tell if an $(n \times n)$-matrix,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

is invertible by computing its determinant and seeing if it is nonzero. In practice, determinants are usually computed using a computer or calculator rather than by hand.

Now, we develop two methods for finding eigenvalues and eigenvectors. The first method is "by hand" and is developed in this section. We will develop this method for (2×2)-matrices, although it can be used for larger matrices with some difficulty. The second method uses *Mathematica* and is developed in the next section. This method can find eigenvalues and eigenvectors of very large $(n \times n)$-matrices with ease.

Suppose that we are interested in the long term behavior of a particular model and want to find the eigenvalues of a (2×2)-matrix,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

We are looking for numbers, λ, and nonzero vectors, $\vec{v}$, such that

$$A\vec{v} = \lambda\vec{v}.$$

Suppose that we want to determine if a particular number, λ, is an eigenvalue. If it is an eigenvalue, then there is a nonzero vector $\vec{v}$ such that

$$A\vec{v} = \lambda\vec{v}.$$

Now, we can use a little matrix algebra to rewrite the above equation as

$$\begin{aligned} A\vec{v} &= \lambda\vec{v} \\ A\vec{v} &= \lambda(I\vec{v}) \\ A\vec{v} &= (\lambda I)\vec{v} \\ A\vec{v} - (\lambda I)\vec{v} &= \vec{0} \\ (A - \lambda I)\vec{v} &= \vec{0}, \end{aligned}$$

where

$$\vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The equation $(A - \lambda I)\vec{v} = \vec{0}$ always has at least one solution, namely $\vec{0}$. But, because by definition eigenvectors are nonzero, we need a nonzero solution. If the matrix $(A - \lambda I)$ had an inverse then we could write

$$\begin{aligned}
(A - \lambda I)\vec{v} &= \vec{0} \\
(A - \lambda I)^{-1}(A - \lambda I)\vec{v} &= (A - \lambda I)^{-1}\vec{0} \\
I\vec{v} &= \vec{0} \\
\vec{v} &= \vec{0}
\end{aligned}$$

and, since the only solution is the zero vector, there would be no eigenvectors corresponding to the number λ. That is, the number, λ, would not be an eigenvalue. Thus, in order to have an eigenvector with eigenvalue λ, the matrix $(A - \lambda I)$ must be singular (or noninvertible). Recall that a matrix is singular if and only if its determinant is zero. Thus, we are looking for numbers λ such that

$$\det(A - \lambda I) = 0.$$

This gives us the equation

$$\det(A - \lambda I) = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0.$$

This equation,

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{12} = 0,$$

is called the **characteristic equation** of the matrix A.

Solving the characteristic equation is the first step in finding eigenvalues and eigenvectors. This step gives us the eigenvalues. We still must find the eigenvectors. Let's look at an example.

Example 1 *Find the eigenvectors and eigenvalues of the matrix*

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

First, we look at and solve the characteristic equation for this matrix.

$$\begin{aligned}
\det(A - \lambda I) &= 0 \\
\det\left(\begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix}\right) & \\
(1-\lambda)(1-\lambda) - 4 &= 0 \\
(1-\lambda)^2 &= 4 \\
1-\lambda &= \pm 2 \\
\lambda &= 1 \pm 2
\end{aligned}$$

and we see that there are two solutions, $\lambda = -1$ and $\lambda = 3$. These are the eigenvalues of the matrix A.

To find the eigenvectors associated with the first eigenvalue, $\lambda = -1$, we must solve the equation

$$\begin{aligned}
A\vec{v} &= (-1)\vec{v} = -\vec{v} \\
\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} &= -\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\
\begin{bmatrix} v_1 + 2v_2 \\ 2v_1 + v_2 \end{bmatrix} &= \begin{bmatrix} -v_1 \\ -v_2 \end{bmatrix}
\end{aligned}$$

giving us the pair of equations

$$\begin{aligned}
v_1 + 2v_2 &= -v_1 \\
2v_1 + v_2 &= -v_2.
\end{aligned}$$

These equations are both equivalent to $v_1 = -v_2$.[9] Thus, our two original equations were redundant and we see that any vector of the form $\vec{v} =< t, -t >$ is an eigenvector associated with the eigenvalue -1.

[9] For example, to show that the first equation, $v_1 + 2v_2 = -v_1$ is equivalent to the equation $v_1 = -v_2$ subtract $2v_2$ from both sides of the first equation get the equation $v_1 = -v_1 - 2v_2$, then add v_1 to both sides to get the equation $2v_1 = -2v_2$, and finally divide both sides by 2 to get $v_1 = -v_2$.

To find the eigenvectors associated with the second eigenvalue, $\lambda = 3$, we must solve the equation

$$\begin{aligned} A\vec{v} &= 3\vec{v} \\ \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} &= 3 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ \begin{bmatrix} v_1 + 2v_2 \\ 2v_1 + v_2 \end{bmatrix} &= \begin{bmatrix} 3v_1 \\ 3v_2 \end{bmatrix} \end{aligned}$$

giving us the pair of equations

$$\begin{aligned} v_1 + 2v_2 &= 3v_1 \\ 2v_1 + v_2 &= 3v_2. \end{aligned}$$

But these equations are again redundant and are both equivalent to $v_1 = v_2$. Thus, we see that any vector of the form $\vec{v} =< t, t >$ is an eigenvector associated with the eigenvalue 3.

We have now determined the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

We often report these results by saying – "There are two eigenvalues. One is $\lambda = -1$ and is associated with the eigenvector $< 1, -1 >$. The other is $\lambda = 3$ and is associated with the eigenvector $< 1, 1 >$. Because any multiple of an eigenvector is another eigenvector (see Theorem 2) with the same eigenvalue, we understand from this sentence that every vector of the form $< t, -t >$ is an eigenvector associated with the eigenvalue -1 and that every vector of the form $< t, t >$ is an eigenvector associated with the eigenvalue 3. For example, $< 1, 1 >$, $< 2, 2 >$, $< 25, 25 >$ are all eigenvectors associated with the eigenvalue 3 and $< 1, -1 >$, $< 2, -2 >$, and $< 25, -25 >$ are all eigenvectors associated with the eigenvalue -1.

Question 1 *Verify that any vector of the form $\vec{v} =< t, t >$ really is an eigenvector of the matrix A and its associated eigenvalue is 3 in the example above by computing $A\vec{v}$.*

Also, verify that any vector of the form $\vec{v} =< t, -t >$ really is an eigenvector of the matrix A and its associated eigenvalue is -1 in the example above by computing $A\vec{v}$.

The example we have just worked is a prototype for finding the eigenvalues and eigenvectors of a (2×2)-matrix. You should be able to solve the following problems using the process from that example. You may, however, be in for some surprises. If you've closed the window with the live diagram that we used earlier, you can click here[10] to reopen it. You may want to use this live diagram to explore eigenvalues and eigenvectors of the matrices in Questions 3-6. Match your solutions to what you see when you explore each matrix using this live graph.

For each of the following questions, verify your eigenvectors $\vec{v}$ and eigenvalues λ by computing $A\vec{v}$.

Question 2 *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix}.$$

Question 3 *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.50 & 0.40 \end{bmatrix}.$$

Question 4 *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 0.50 & 0.60 \\ -0.60 & 0.50 \end{bmatrix}.$$

Question 5 *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 0.60 & 0.00 \\ 0.00 & 0.60 \end{bmatrix}.$$

[10] http://www.dean.usma.edu/departments/math/courses/MA103/MRCW_text/Block_IV/population-growth.html

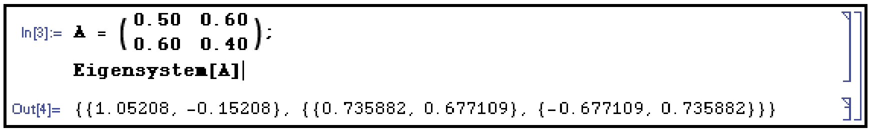


Figure 4.13: Finding Eigenvectors and Eigenvalues with *Mathematica*

4.6 Eigenvalues and Eigenvectors, III

4.6.1 Using *Mathematica* to Find Eigenvalues and Eigenvectors

Because eigenvalues and their associated eigenvectors are so important, *Mathematica* has a procedure, **Eigensystem**, that computes the eigenvectors of an $(n \times n)$-matrix and their associated eigenvalues. Figure 4.13 shows how this procedure is used to find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix}.$$

The output of **Eigensystem** is a list with two items. The first item is a list of the eigenvalues. Notice in Figure 4.13 the list of eigenvalues is $\{1.05208, -0.15208\}$, the same two eigenvalues you should have found in Question 2 of the preceding section. The second item in the output of **Eigensystem** is another list – containing the eigenvectors associated with the eigenvalues in the first list. Notice in Figure 4.13 that the two vectors may be different from the ones that you found because any multiple of an eigenvector is an eigenvector with the same associated eigenvalue. Recall Theorem 2 on page 373. For example, suppose that for a particular problem you found an eigenvector $< 1, 2, 3 >$ and *Mathematica* found an eigenvector (with the same eigenvalue) $< 3, 6, 9 >$. These two answers are really the same since

$$< 3, 6, 9 > = 3 < 1, 2, 3 > .$$

There is a subtle difference between "numbers" like 0.60 and 3/5. Numbers that are written with decimals are subject to round-off error but numbers that are written as fractions are not. For example, the number 0.33 is different from the number 1/3. Even when round-off error appears to be absent, using computers can introduce round-off error. For example, there is no error when we write 3/5 as 0.60 in the decimal system but computers use the binary system and 1/5 is not exact in the binary system. Because fractions are exact and decimals may not be exact it is sometimes better to use fractions instead of decimals in *Mathematica* when that is possible.

Question 1 *Check that the eigenvectors you found for the matrix*

$$\begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix}$$

are multiples of the eigenvectors found by Mathematica.

The first three questions below ask you to find the eigenvalues and eigenvectors for the same matrices as the questions at the end of the preceding section. This time you should use the *Mathematica* procedure **Eigensystem**.

Question 2 *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.50 & 0.40 \end{bmatrix}.$$

Question 3 *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 0.50 & 0.60 \\ -0.60 & 0.50 \end{bmatrix}.$$

Question 4 *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 0.60 & 0 \\ 0 & 0.60 \end{bmatrix}.$$

Question 5 *Find the eigenvalues and eigenvectors of the matrix*

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

We conclude this section with an example and a series of questions that will lead to a powerful tool for understanding the long term behavior of systems.

Example 1 *This example involves a fish population model with three age groups – age under one year, age between one year and two years, and age over two years. Suppose these fish are introduced into a lake in the hope that they will thrive and eventually provide a food source for a nearby village. We use the notation a_n, b_n, and c_n for the number of thousands of fish in each age group in the lake during year n, starting with $n = 0$ being the year in which the fish are introduced into the lake. Suppose that the initial population in each age group is given by*

$$a_0 = 5, \quad b_0 = 0, \text{ and } \quad c_0 = 0.$$

Based on experience with the same species in other lakes, scientists believe that the survival rates for each age group are:

- *Each year 70% of the population in the first age group will survive and reach the second age group the next year.*
- *Each year 90% of the population in the second age group will survive and reach the third age group the next year.*
- *Each year 40% of the population in the third age group will survive and remain in the third age group the next year. This estimate is based on allowing villagers to catch and keep 20% of the fish in the third age group. They will not be allowed to keep any fish they catch in either of the first two age groups.*

Based on the same experience, scientists also believe that the reproduction rates will be:

- *On average, each year 10% of the fish in the first age group will have a child.*
- *On average, each year 80% of the fish in the second age group will have a child.*
- *On average, each year 30% of the fish in the third age group will have a child.*

This information can be expressed as the following model

$$
\begin{aligned}
a_n &= 0.10a_{n-1} + 0.80b_{n-1} + 0.30c_{n-1} \\
b_n &= 0.70a_{n-1} \\
c_n &= \qquad\qquad\quad 0.90b_{n-1} + 0.40c_{n-1}
\end{aligned}
$$

We can also express this model using vectors and matrices as

$$
\vec{p}_n = \begin{bmatrix} a_n \\ b_n \\ c_n \end{bmatrix}
$$

$$
\vec{p}_n = \begin{bmatrix} 0.10 & 0.80 & 0.30 \\ 0.70 & 0.00 & 0.00 \\ 0.00 & 0.90 & 0.40 \end{bmatrix} \vec{p}_{n-1}
$$

with the initial condition

$$
\vec{p}_0 = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}.
$$

Using the *Mathematica* procedure **Eigensystem** we find the following eigenvalues and associated eigenvectors.

$$
\lambda_1 = 0.989686 \quad \vec{v}_1 = \begin{bmatrix} 0.612497 \\ 0.433216 \\ 0.661190 \end{bmatrix},
$$

$$
\lambda_2 = 0.553571 \quad \vec{v}_2 = \begin{bmatrix} -0.498544 \\ 0.630418 \\ -0.595002 \end{bmatrix},
$$

$$
\lambda_3 = 0.0638848 \quad \vec{v}_3 = \begin{bmatrix} -0.031913 \\ -0.349681 \\ 0.936325 \end{bmatrix}.
$$

Question 6 *Predict the fish population in this lake by computing* $\vec{p}_1, \vec{p}_2, \ldots \vec{p}_{20}$. *Do you notice any relationship between your results and the eigenvalues and eigenvectors of the matrix.*

Question 7 *Suppose that the initial fish population in the lake was* $\vec{v}_1$. *Describe what would happen over the long term.*

Question 8 *Suppose that the initial fish population in the lake was* $\vec{v}_2$. *Describe what would happen over the long term.*

Question 9 *Suppose that the initial fish population in the lake was* $\vec{v}_3$. *Describe what would happen over the long term.*

Question 10 *Suppose that the initial fish population in the lake was* $\vec{v}_1 + \vec{v}_2 + \vec{v}_3$. *Describe what would happen over the long term.*

Year	$0 \leq$ age < 1	$1 \leq$ age < 2	$2 \leq$ age	**Total**	**Total Population Ratio**
0	5.000000	0.000000	0.000000	5.000000	
1	0.500000	3.500000	0.000000	4.000000	0.800000
2	2.850000	0.350000	3.150000	6.350000	1.587500
3	1.510000	1.995000	1.575000	5.080000	0.800000
4	2.219500	1.057000	2.425500	5.702000	1.122441
5	1.795200	1.553650	1.921500	5.270350	0.924298
6	1.998890	1.256640	2.166885	5.422415	1.028853
7	1.855267	1.399223	1.997730	5.252220	0.968613
8	1.904224	1.298687	2.058393	5.261303	1.001730
9	1.846889	1.332957	1.992175	5.172021	0.983030
10	1.848707	1.292823	1.996531	5.138061	0.993434
11	1.818088	1.294095	1.962153	5.074336	0.987597
12	1.805731	1.272662	1.949546	5.027939	0.990857
13	1.783566	1.264011	1.925214	4.972792	0.989032
14	1.767130	1.248496	1.907696	4.923322	0.990052
15	1.747819	1.236991	1.886725	4.871535	0.989481
16	1.730392	1.223473	1.867982	4.821847	0.989800
17	1.712212	1.211275	1.848319	4.771806	0.989622
18	1.694736	1.198549	1.829475	4.722760	0.989722
19	1.677155	1.186316	1.810484	4.673954	0.989666
20	1.659913	1.174008	1.791877	4.625799	0.989697

Table 4.7: Three Age Group Fish Model

4.7 Analytic Solutions of Systems of Recursion Equations

As you worked on Questions 6 - 10 in the preceding section, you may have constructed a table like Table 4.7. This table computes the population in each age group for twenty years and also looks at the total population each year. In addition, it computes the ratio of each year's total population to the previous year's total population. Notice, after twenty years, this ratio is very close to the eigenvalue λ_1. Furthermore, if we look at the population in the 20^{th} year,

$$p_{20} = \begin{bmatrix} 1.659913 \\ 1.174008 \\ 1.791877 \end{bmatrix},$$

and compare this vector term-by-term with the eigenvector associated with λ_1,

$$\vec{v}_1 = \begin{bmatrix} 0.612497 \\ 0.433216 \\ 0.661190 \end{bmatrix},$$

we see that

$$\frac{1.659913}{0.612497} = 2.710075 \quad \frac{1.174008}{0.433216} = 2.709983 \quad \frac{1.791877}{0.661190} = 2.710079.$$

These ratios are all roughly the same. If we round off to three digits past the decimal point, we get 2.710. Thus, $\vec{p}_{20}$ is roughly 2.710 $\vec{v}_1$. Since $\vec{v}_1$ is an eigenvector associated with the eigenvalue $\lambda_1 = 0.989686$ and any multiple of an eigenvector is also an eigenvector with the same associated eigenvalue, we see that, after a long time, each year the population vector is very close to an eigenvector.

Based in part on this example, it looks as if eigenvalues and eigenvectors might help us understand these kinds of models. We can explore this connection and answer some of the questions we have been considering by using matrix-vector notation for this linear discrete dynamical system with many variables,

$$\vec{p}_n = A\vec{p}_{n-1}$$

where

$$A = \begin{bmatrix} 0.10 & 0.80 & 0.30 \\ 0.70 & 0.00 & 0.00 \\ 0.00 & 0.90 & 0.40 \end{bmatrix}.$$

We continue our exploration by asking what would happen if the initial condition was one of the eigenvectors.

- Notice that, if $p_0 = \vec{v}_1$ then

$$\begin{aligned} p_1 &= A\vec{p}_0 = A\vec{v}_1 = \lambda_1\vec{v}_1 \\ p_2 &= A\vec{p}_1 = A(\lambda_1\vec{v}_1) = \lambda_1^2\vec{v}_1 \\ p_3 &= A\vec{p}_2 = A(\lambda_1^2\vec{v}_1) = \lambda_1^3\vec{v}_1 \\ &\vdots \\ p_n &= \lambda_1^n\,\vec{v}_1 \end{aligned}$$

- Notice that, if $p_0 = \vec{v}_2$ then

$$\begin{aligned} p_1 &= A\vec{p}_0 = A\vec{v}_2 = \lambda_2\vec{v}_2 \\ p_2 &= A\vec{p}_1 = A(\lambda_2\vec{v}_2) = \lambda_2^2\vec{v}_2 \\ p_3 &= A\vec{p}_2 = A(\lambda_2^2\vec{v}_2) = \lambda_2^3\vec{v}_2 \\ &\vdots \\ p_n &= \lambda_2^n\ \vec{v}_2 \end{aligned}$$

- Notice that, if $p_0 = \vec{v}_3$ then

$$\begin{aligned} p_1 &= A\vec{p}_0 = A\vec{v}_3 = \lambda_3\vec{v}_3 \\ p_2 &= A\vec{p}_1 = A(\lambda_3\vec{v}_3) = \lambda_3^2\vec{v}_3 \\ p_3 &= A\vec{p}_2 = A(\lambda_3^2\vec{v}_3) = \lambda_3^3\vec{v}_3 \\ &\vdots \\ p_n &= \lambda_3^n\ \vec{v}_3 \end{aligned}$$

Finally, for the last question at the end of the previous section we see that if $p_0 = \vec{v}_1 + \vec{v}_2 + \vec{v}_3$, then

$$\begin{aligned} p_1 &= A\vec{p}_0 = A(\vec{v}_1 + \vec{v}_2 + \vec{v}_3) = A\vec{v}_1 + A\vec{v}_2 + A\vec{v}_3 = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \lambda_3\vec{v}_3 \\ p_2 &= A\vec{p}_1 = A(\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \lambda_3\vec{v}_3) = A\lambda_1\vec{v}_1 + A\lambda_2\vec{v}_2 + A\lambda_3\vec{v}_3 = \lambda_1^2\vec{v}_1 + \lambda_2^2\vec{v}_2 + \lambda_3^2\vec{v}_3 \\ &\vdots \\ p_n &= \lambda_1^n\ \vec{v}_1 + \lambda_2^n\ \vec{v}_2 + \lambda_3^n\ \vec{v}_3 \end{aligned}$$

Thus, we have a closed-form solution to the IVP

$$\vec{p}_0 = \vec{v}_1 + \vec{v}_2 + \vec{v}_3, \qquad \vec{p}_n = A\vec{p}_{n-1}$$

namely,

$$\vec{p}_n = \lambda_1^n\ \vec{v}_1 + \lambda_2^n\ \vec{v}_2 + \lambda_3^n\ \vec{v}_3 = (0.989686)^n\vec{v}_1 + (0.553571)^n\vec{v}_2 + (0.0638848)^n\vec{v}_3.$$

We can see the long term behavior of this IVP from this closed-form solution. Because

$$\lim_{n\to\infty}(0.989686)^n = 0, \quad \lim_{n\to\infty}(0.553571)^n = 0, \quad \text{and} \quad \lim_{n\to\infty}(0.0638848)^n = 0,$$

we see that

$$\lim_{n\to\infty}\vec{p}_n = \lim_{n\to\infty}\left((0.989686)^n\vec{v}_1 + (0.553571)^n\vec{v}_2 + (0.0638848)^n\vec{v}_3\right) = \vec{0}.$$

We can even be a bit more precise. Because the eigenvalues $\lambda_2 = 0.553571$ and $\lambda_3 = 0.0638848$ are so much smaller than the eigenvalue $\lambda_1 = 0.989686$, the second and third terms in $\vec{p}_n$ die out much more quickly than the first term.

$$\vec{p}_n = \underbrace{(0.989686)^n}_{\text{dies out slowly}}\vec{v}_1 + \underbrace{(0.553571)^n}_{\text{dies out quickly}}\vec{v}_2 + \underbrace{(0.0638848)^n}_{\text{dies out quickly}}\vec{v}_3.$$

This means that after many years the population is close to

$$\vec{p}_n \approx (0.989686)^n\vec{v}_1.$$

Notice that we use the symbol $\approx$ instead of the symbol $=$ because this is only an approximate formula for $\vec{p}_n$. Moreover, this formula only applies after many years.

Table 4.8 on page 394 compares the direct calculations for the IVP

$$\vec{p}_0 = \vec{v}_1 + \vec{v}_2 + \vec{v}_3 = \begin{bmatrix} 0.612497 \\ 0.433216 \\ 0.661190 \end{bmatrix} + \begin{bmatrix} -0.498544 \\ 0.630418 \\ -0.595002 \end{bmatrix} + \begin{bmatrix} -0.031913 \\ -0.349681 \\ 0.936325 \end{bmatrix} = \begin{bmatrix} 0.082040 \\ 0.713953 \\ 1.002513 \end{bmatrix}, \qquad \vec{p}_n = A\vec{p}_{n-1}$$

with the closed-form solution

$$\vec{p}_n = \lambda_1^n\ \vec{v}_1 + \lambda_2^n\ \vec{v}_2 + \lambda_3^n\ \vec{v}_3 = (0.989686)^n\vec{v}_1 + (0.553571)^n\vec{v}_2 + (0.0638848)^n\vec{v}_3$$

and the approximation

$$\vec{p}_n \approx (0.989686)^n\vec{v}_1.$$

n	**Population by Age Group a = age**								
	Original Model			**Closed Form Solution**			**Approximation**		
	$0 \le a < 1$	$1 \le a < 2$	$2 \le a$	$0 \le a < 1$	$1 \le a < 2$	$2 \le a$	$0 \le a < 1$	$1 \le a < 2$	$2 \le a$
0	0.082040	0.713953	1.002513	0.082040	0.713953	1.002513	0.612497	0.433216	0.661190
1	0.880120	0.057428	1.043563	0.328161	0.755390	0.384812	0.606180	0.428748	0.654370
2	0.447023	0.616084	0.469110	0.447023	0.616084	0.469110	0.599928	0.424326	0.647621
3	0.678303	0.312916	0.742120	0.509160	0.526800	0.540252	0.593740	0.419949	0.640942
4	0.540799	0.474812	0.578473	0.540799	0.474812	0.578472	0.587616	0.415618	0.634331
5	0.607471	0.378560	0.658720	0.555639	0.444102	0.596859	0.581555	0.411331	0.627789
6	0.561211	0.425230	0.604191	0.561211	0.425230	0.604191	0.575557	0.407089	0.621314
7	0.577562	0.392847	0.624384	0.561679	0.412932	0.605427	0.569621	0.402890	0.614905
8	0.559349	0.404294	0.603316	0.559350	0.404294	0.603316	0.563746	0.398735	0.608563
9	0.560365	0.391545	0.605191	0.555498	0.397699	0.599382	0.557931	0.394622	0.602286
10	0.550829	0.392255	0.594466	0.550830	0.392255	0.594467	0.552177	0.390552	0.596075
11	0.547227	0.385581	0.590816	0.545736	0.387467	0.589037	0.546482	0.386524	0.589927
12	0.540432	0.383059	0.583349	0.540432	0.383059	0.583349	0.540845	0.382537	0.583842
13	0.535495	0.378302	0.578093	0.535039	0.378881	0.577548	0.535267	0.378592	0.577820
14	0.529619	0.374847	0.571709	0.529620	0.374847	0.571710	0.529746	0.374687	0.571861
15	0.524352	0.370733	0.566046	0.524212	0.370911	0.565879	0.524283	0.370822	0.565963
16	0.518836	0.367046	0.560078	0.518836	0.367047	0.560079	0.518875	0.366998	0.560125
17	0.513544	0.363185	0.554373	0.513502	0.363240	0.554322	0.513523	0.363212	0.554348
18	0.508214	0.359481	0.548616	0.508215	0.359481	0.548616	0.508227	0.359466	0.548631
19	0.502991	0.355750	0.542979	0.502978	0.355767	0.542964	0.502985	0.355759	0.542972
20	0.497793	0.352094	0.537367	0.497794	0.352094	0.537367	0.497797	0.352089	0.537372

Table 4.8: Comparing the Model, the Closed-Form Solution, and an Approximation

Notice that, as expected, the calculations using the closed-form solution yield the same results as the the direct calculations for the IVP but that the approximation is only a reasonably good approximation after a number of years have passed.

Question 1 *The model we have been using as an example was based on the assumption that villagers would be allowed to catch and keep 20% of the fish in the third age group. As we have seen so far, it looks like this is a recipe for long term disaster because the fish population will slowly die out. Suppose that villagers are only allowed to catch and keep 15% of the fish in the third age group. Analyze this situation in the same way that we analyzed the first situation. Notice that for this new situation, we will use the matrix*

$$B = \begin{bmatrix} 0.10 & 0.80 & 0.30 \\ 0.70 & 0.00 & 0.00 \\ 0.00 & 0.90 & 0.45 \end{bmatrix}$$

instead of the matrix A. You may want to refer back to Example 1 which starts on Page 386 and to Questions 6 - 10 on Page 388. Note that this is the same example we discussed at the beginning of this section.

In the last few pages we found a solution of the IVP

$$\vec{p}_0 = \vec{v}_1 + \vec{v}_2 + \vec{v}_3, \qquad \vec{p}_n = A\vec{p}_{n-1}$$

and you may have found a solution of the IVP

$$\vec{p}_0 = \vec{v}_1 + \vec{v}_2 + \vec{v}_3, \qquad \vec{p}_n = B\vec{p}_{n-1}$$

as you answered Question 1 – BUT we really want solutions for actual initial values, not just for ones that are conveniently chosen – like the eigenvectors, $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$, and their sum, $\vec{v}_1 + \vec{v}_2 + \vec{v}_3$. Now, we show how we can use the work we've done so far to find a closed-form solution for real IVPs. We illustrate this by looking at our original IVP,

$$\vec{p}_0 = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \qquad \vec{p}_n = A\vec{p}_{n-1}.$$

The key idea is to find numbers c_1, c_2, and c_3 such that

$$\vec{p}_0 = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3.$$

This procedure is called "vector decomposition."

Example 1 *Suppose that*

$$\vec{p}_0 = \begin{bmatrix} 8 \\ 14 \end{bmatrix}, \quad \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Then we have the vector decomposition

$$\vec{p}_0 = 2\vec{v}_1 + 3\vec{v}_2.$$

If we can decompose our initial condition into a linear combination of eigenvectors, then we can obtain a powerful closed-form solution by noticing that

$$\begin{aligned}
\vec{p}_0 &= c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 \\
\vec{p}_1 &= A\vec{p}_0 \\
&= A(c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3) \\
&= c_1A\vec{v}_1 + c_2A\vec{v}_2 + c_3A\vec{v}_3 \\
&= c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2 + c_3\lambda_3\vec{v}_3 \\
\vec{p}_2 &= A\vec{p}_1 \\
&= A(c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2 + c_3\lambda_3\vec{v}_3) \\
&= c_1\lambda_1A\vec{v}_1 + c_2\lambda_2A\vec{v}_2 + c_3\lambda_3A\vec{v}_3 \\
&= c_1\lambda_1^2\vec{v}_1 + c_2\lambda_2^2\vec{v}_2 + c_3\lambda_3^2\vec{v}_3 \\
&\vdots \\
\vec{p}_n &= (\lambda_1)^n c_1\vec{v}_1 + (\lambda_2)^n c_2\vec{v}_2 + (\lambda_3)^n c_3\vec{v}_3.
\end{aligned}$$

Now we want to find the vector decomposition of our initial condition. To do this we must solve the equation

$$\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = c_1\begin{bmatrix} 0.612497 \\ 0.433216 \\ 0.661190 \end{bmatrix} + c_2\begin{bmatrix} -0.498544 \\ 0.630418 \\ -0.595002 \end{bmatrix} + c_3\begin{bmatrix} -0.031913 \\ -0.349681 \\ 0.936325 \end{bmatrix},$$

which leads to the system of equations

$$\begin{aligned}
0.612497c_1 - 0.498544c_2 - 0.031913c_3 &= 5 \\
0.433216c_1 + 0.630418c_2 - 0.349681c_3 &= 0 \\
0.661190c_1 - 0.595002c_2 + 0.936325c_3 &= 0
\end{aligned}$$

This system of equations can easily be solved using *Mathematica*. We obtain

$$\begin{aligned} c_1 &= 3.33448 \\ c_2 &= -5.55583 \\ c_3 &= -5.8852. \end{aligned}$$

Thus,

$$\vec{p}_0 = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} = 3.33448 \begin{bmatrix} 0.612497 \\ 0.433216 \\ 0.661190 \end{bmatrix} - 5.55583 \begin{bmatrix} -0.498544 \\ 0.630418 \\ -0.595002 \end{bmatrix} - 5.8852 \begin{bmatrix} -0.031913 \\ -0.349681 \\ 0.936325 \end{bmatrix},$$

Putting this all together, we see that

$$\begin{aligned} \vec{p}_n &= \lambda_1^n\ (3.33448)\vec{v}_1 - \lambda_2^n\ (5.55583)\vec{v}_2 - \lambda_3^n\ (5.8852)\vec{v}_3 \\ &= (0.989686)^n(3.33448) \begin{bmatrix} 0.612497 \\ 0.433216 \\ 0.661190 \end{bmatrix} - (0.553571)^n(5.55583) \begin{bmatrix} -0.498544 \\ 0.630418 \\ -0.595002 \end{bmatrix} - \\ &\quad (0.0638848)^n(5.8852) \begin{bmatrix} -0.031913 \\ -0.349681 \\ 0.936325 \end{bmatrix} \end{aligned}$$

Question 2 *Use either Excel or Mathematica to check the closed-form solution above by comparing the results of this closed-form solution to the direct calculations from the original IVP. See Table 4.7 on page 390.*

Recall Example 1 from Section 4.3 on page 358. You may want to look back at that example to recall the context. Look at Figure 4.14 and Table 4.9 on page 398. Both of these refer to the model

$$\vec{p}_n = A\vec{p}_{n-1}, \qquad \vec{p}_{2008} = \begin{bmatrix} 6 \\ 0 \end{bmatrix},$$

where

$$A = \begin{bmatrix} 0.50 & 0.60 \\ 0.60 & 0.40 \end{bmatrix}.$$

Now, we want to use the method we developed in the last example to solve this IVP.

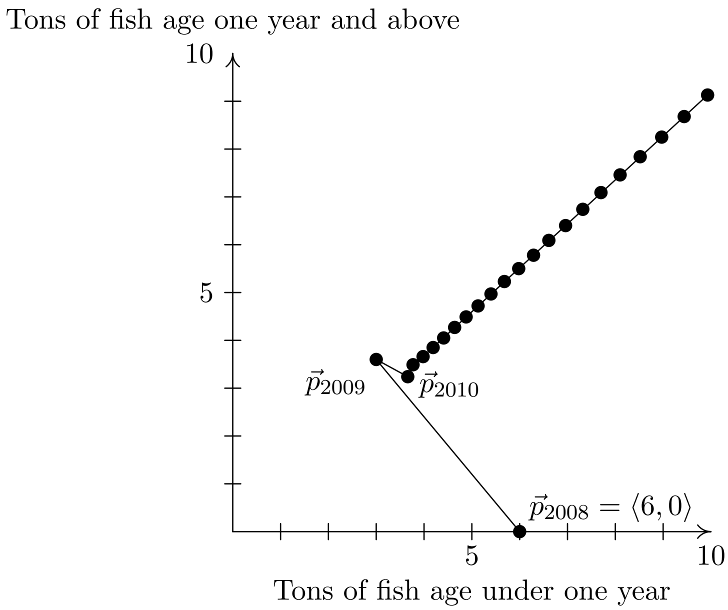


Figure 4.14: Representing the Fish Population Visually

Year	**Age < 1**	$1 \leq$ **Age**
2008	6.00000	0.00000
2009	3.00000	3.60000
2010	3.66000	3.24000
2011	3.77400	3.49200
2012	3.98220	3.66120
2013	4.18782	3.85380
2014	4.40619	4.05421
2015	4.63562	4.26540
2016	4.87705	4.48753
2017	5.13104	4.72124
2018	5.39827	4.96712
2019	5.67941	5.22581
2020	5.97519	5.49797
2021	6.28638	5.78430
2022	6.61377	6.08555
2023	6.95821	6.40248
2024	7.32060	6.73592
2025	7.70185	7.08673
2026	8.10296	7.45580
2027	8.52496	7.84410

Table 4.9: Twenty Years

```
In[1]:= A = (0.50  0.60
             0.60  0.40);
        Eigensystem[A]
Out[2]= {{1.05208, -0.15208}, {{0.735882, 0.677109}, {-0.677109, 0.735882}}}
```

Figure 4.15: Using *Mathematica* to find the Eigenvalues and Eigenvectors of A

Figure 4.15 shows the eigenvalues and eigenvectors of our matrix A using *Mathematica*,

$$\lambda_1 = 1.05208 \quad \vec{v}_1 = \begin{bmatrix} 0.735882 \\ 0.677109 \end{bmatrix}, \qquad \lambda_2 = -0.15208 \qquad \vec{v}_2 = \begin{bmatrix} -0.677109 \\ 0.735882 \end{bmatrix}.$$

Now, we want to write our initial condition in the form

$$\vec{p}_{2008} = c_1\vec{v}_1 + c_2\vec{v}_2.$$

That is, we want to find numbers c_1 and c_2 such that

$$\begin{bmatrix} 6.0 \\ 0.0 \end{bmatrix} = c_1 \begin{bmatrix} 0.735882 \\ 0.677109 \end{bmatrix} + c_2 \begin{bmatrix} -0.677109 \\ 0.735882 \end{bmatrix}.$$

This gives us the equations

$$\begin{aligned} 0.735882c_1 - 0.677109c_2 &= 6.0 \\ 0.677109c_1 + 0.735882c_2 &= 0. \end{aligned}$$

These equations are easily solved using *Mathematica*. We obtain

$$\begin{aligned} c_1 &= 4.4153 \\ c_2 &= -4.06266. \end{aligned}$$

Putting this all together we see that

$$\begin{aligned}\vec{p}_n &= (\lambda_1)^{(n-2008)} c_1 \vec{v}_1 + (\lambda_2)^{(n-2007)} c_2 \vec{v}_2 \\ &= (1.05208)^{(n-2008)}(4.4153) \begin{bmatrix} 0.735882 \\ 0.677109 \end{bmatrix} + \\ &\quad (-0.15208)^{(n-2008)}(-4.06266) \begin{bmatrix} -0.677109 \\ 0.735882 \end{bmatrix}.\end{aligned}$$

Just as in our first example, this form tells us a lot about the long term behavior of this population. Since $|\lambda_2| < 1$, the term

$$(\lambda_2)^{(n-2008)} \vec{v}_2 = (-0.15208)^{(n-2008)}(-4.06266) \begin{bmatrix} -0.677109 \\ 0.735882 \end{bmatrix}$$

becomes very small as n increases. Since $|\lambda_1| > 1$, the term

$$(\lambda_1)^{(n-2008)} \vec{v}_1 = (1.05208)^{(n-2008)}(4.4153) \begin{bmatrix} 0.735882 \\ 0.677109 \end{bmatrix}$$

keeps growing as n increases. This explains the behavior we saw in Figure 4.14 and Table 4.9.

Question 3 *The village elders propose allowing the villagers to catch and keep 15% of the fish that are aged one year and above each year. What advice would you give them?*

Question 4 *Recall that, for this example, we were given the information*

- *Each year the reproduction rate for fish under the age of one year is 0.50 and the reproduction rate for fish aged one year or more is 0.60.*
- *Each year the survival rate for fish aged one year or less is 0.60 and the survival rate for fish aged one year or more is 0.40.*

Suppose that you are given a choice of stocking the lake with this fish or with another species with the following characteristics.

- *Each year the reproduction rate for fish under the age of one year is 0.60 and the reproduction rate for fish aged one year or more is 0.80.*
- *Each year the survival rate for fish aged one year or less is 0.50 and the survival rate for fish aged one year or more is 0.20.*

Develop a model for this species and use it to compare the two species. Which would you choose?

4.8 Eigenvalues, Eigenvectors, Closed-Form Solutions, and Long Term Behavior

In this section we summarize the work we've done in the last few sections in the form of some tools and theorems that collectively describe a procedure that can be used to study the behavior, especially the long term behavior, of many models like the ones we've been studying. We begin with a basic IVP of the form

$$\vec{p}_0 = \vec{b}, \qquad \vec{p}_k = A\vec{p}_{k-1}$$

where the vectors are n-dimensional vectors and the matrix A is an $(n \times n)$-matrix. The vector $\vec{b}$ is called the initial value.

Our most important tool is the *Mathematica* procedure **Eigensystem** that we use to find the eigenvalues and eigenvectors of the matrix A. In general, the subject of eigenvalues and eigenvectors is somewhat complex. An $(n \times n)$-matrix can have up to n eigenvalues and some of them may be complex numbers.

Example 1 *Let*

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

We find the eigenvalues of this matrix by solving the equation

$$\begin{aligned}
\det(A - \lambda I) &= 0 \\
\begin{vmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{vmatrix} &= 0 \\
(1-\lambda)^2 + 1 &= 0 \\
(1-\lambda)^2 &= -1 \\
1-\lambda &= \pm i \\
\lambda &= 1 \pm i.
\end{aligned}$$

Thus, this matrix has two complex eigenvalues and no real eigenvalues. If you are not familiar with complex numbers, then the real take-away from this example is that this matrix has no real eigenvalues.

Because time is short in this course, we study only situations in which the matrix A has n distinct, real eigenvalues. Its eigenvalues and associated eigenvectors may be found using the *Mathematica* procedure **Eigensystem**. Note that for each eigenvalue, any scalar multiple of an associated eigenvector is also an associated eigenvector.

The following theorem summarizes what we have seen in the preceding two sections.

Theorem 3 *Suppose that*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

has n distinct real eigenvalues and associated eigenvectors

$$\lambda_1, \vec{v}_1; \quad \lambda_2, \vec{v}_2; \quad \ldots \lambda_n, \vec{v}_n.$$

Now, suppose that $\vec{b}$ is any vector. Then, there are numbers c_1, c_2, $\ldots c_n$ such that

$$\vec{b} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$$

and these numbers can be found using the Mathematica **Solve** *procedure as shown in Figure 4.16 on page 404.*

Figure 4.16 shows how we can use the *Mathematica* **Eigensystem** and **Solve** procedures efficiently for the kinds of problems we have been discussing. In particular, this figure illustrates some techniques that can be used to save some typing and to avoid the ugly details of the way that *Mathematica* presents the results of the procedure **Eigensystem**.

- The line **work = Eigensystem[A]** computes and saves the eigenvalues and eigenvectors of the matrix A in the format that *Mathematica* uses to return these values. The line **Out[2]** shows this format.
- The line **lambda[i_] := work[[1]][[i]]** defines a function that picks out the i^{th} eigenvalue. The lines **lambda[1]**, **lambda[2]**, and **lambda[3]** then print out the three eigenvalues.

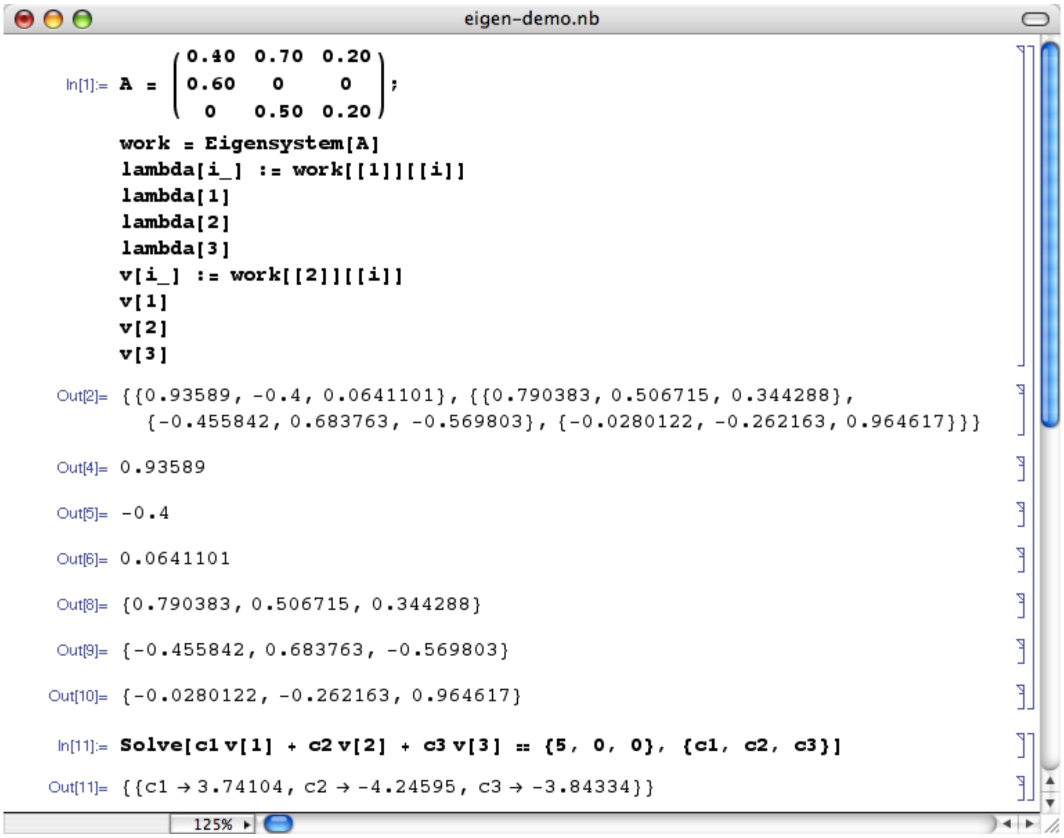


Figure 4.16: Using *Mathematica*

- The line **v**[**i**_] **:=** **work**[[**2**]][[**i**]] defines a function that picks out the i^{th} eigenvector, the one associated with λ_i. The lines **v**[**1**], **v**[**2**], and **v**[**3**] print out the three eigenvectors.

- The line **Solve**[**c1 v**[1] + **c2 v**[**2**] + **c3 v**[**3**] **== {5, 0, 0}, c1, c2, c3**] solves the equation

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 =< 5, 0, 0 > .$$

  Notice that we use row-vector notation in the equation above and in Figure 4.16.

Putting this all together, if the $(n \times n)$-matrix A has n distinct real eigenvalues, then we have a step-by-step procedure for finding a closed-form solution of the IVP

$$\vec{p}_n = A\vec{p}_{n-1}, \quad \vec{p}_0 = \vec{b}$$

given by

1. Find the eigenvalues and eigenvectors of the matrix A. Check to be sure that there are n distinct, real eigenvalues. Let $\lambda_1, \lambda_2, \ldots \lambda_n$ denote the eigenvalues and $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n$ denote the associated eigenvectors. Thus,

$$A\vec{v}_1 = \lambda_1\vec{v}_1,\ A\vec{v}_2 = \lambda_2\vec{v}_2,\ \ldots,\ A\vec{v}_n = \lambda_n\vec{v}_n.$$

2. Solve the equation

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = \vec{b}.$$

3. Now, we have

$$\begin{aligned}
\vec{p}_0 &= \vec{b} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \\
\vec{p}_1 &= A\vec{p}_0 = A(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n) = \lambda_1 c_1\vec{v}_1 + \lambda_2 c_2\vec{v}_2 + \cdots + \lambda_n c_c\vec{v}_n \\
\vec{p}_2 &= A\vec{p}_1 = A(\lambda_1 c_1\vec{v}_1 + \lambda_2 c_2\vec{v}_2 + \cdots + \lambda_n c_n\vec{v}_n) = (\lambda_1)^2 c_1\vec{v}_1 + (\lambda_2)^2 c_2\vec{v}_2 + \cdots + (\lambda_n)^2 c_n\vec{v}_n \\
&\vdots \\
\vec{p}_k &= (\lambda_1)^k c_1\vec{v}_1 + (\lambda_2)^k c_2\vec{v}_2 + \cdots + (\lambda_n)^k c_n\vec{v}_n.
\end{aligned}$$

The next theorem is the first step in understanding the long term behavior of an IVP.

Theorem 4 *Consider the IVP*

$$\vec{p}_k = A\vec{p}_{k-1}, \quad \vec{p}_0 = \vec{b}$$

If the matrix A has n distinct real eigenvalues all of which have absolute value less than one, then

$$\lim_{k\to\infty} \vec{p}_k = \vec{0}.$$

This means that the populations in each of the three age groups go to zero.

Proof

$$\lim_{k\to\infty} \vec{p}_k = \lim_{k\to 0}((\lambda_1)^k c_1\vec{v}_1 + (\lambda_2)^k c_2\vec{v}_2 + \cdots + (\lambda_n)^k c_n\vec{v}_n).$$

and, since all the eigenvalues have absolute values less than one,

$$\begin{aligned} \lim_{k\to\infty} (\lambda_1)^k &= 0 \\ \lim_{k\to\infty} (\lambda_2)^k &= 0 \\ &\vdots \\ \lim_{k\to\infty} (\lambda_n)^k &= 0 \end{aligned}$$

and we see that

$$\lim_{k\to\infty} \vec{p}_k = \vec{0}. \blacksquare$$

Example 2 *This example is another population model with three age groups. We are interested in a particular species in a particular habitat and we divide the population into three groups – those aged less than one year; those whose age is greater than or equal to one year but less than two years; and those whose age is greater than or equal to two years. Suppose the initial population is 5 million individuals in age group 1 and none in either of the other two age groups. Thus,*

$$\vec{p}_0 = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}.$$

Notice that the entries in the vector $\vec{p}_0$ give the number of millions of individuals in each age group.

Suppose that the fertility rates are 0.40 for age group 1; 0.70 for age group 2; and 0.20 for age group 3 and that the survival rates are 0.60 for age group 1; 0.50 for age group 2; and 0.20 for age group 3. Thus,

$$\vec{p}_k = A\vec{p}_{k-1},$$

where

$$A = \begin{bmatrix} 0.40 & 0.70 & 0.20 \\ 0.60 & 0 & 0 \\ 0 & 0.50 & 0.20 \end{bmatrix}.$$

Using *Mathematica* (See Figure 4.16) we see that the eigenvalues and eigenvectors are

$$\lambda_1 = 0.93589, \vec{v}_1 = \begin{bmatrix} 0.790383 \\ 0.506715 \\ 0.344288 \end{bmatrix}; \quad \lambda_2 = -0.4, \vec{v}_2 = \begin{bmatrix} -0.455842 \\ 0.683763 \\ -0.569803 \end{bmatrix}; \quad \lambda_3 = 0.0641101, \vec{v}_3 = \begin{bmatrix} -0.0280122 \\ -0.2621630 \\ 0.9646170 \end{bmatrix}.$$

Also, using vector decomposition, we see that

$$\vec{p}_0 = 3.74104\ \vec{v}_1 - 4.24595\ \vec{v}_2 - 3.84334\ \vec{v}_3$$

and obtain the closed-form solution

$$\vec{p}_k = (0.93589)^k\ 3.74104\ \vec{v}_1 - (-0.4)^k\ 4.24595\ \vec{v}_2 - (0.0641101)^k\ 3.84334\ \vec{v}_3.$$

Since all of the eigenvalues have absolute value less than 1, we see that

$$\lim_{k\to\infty} \vec{p}_k = \vec{0}.$$

We can say a bit more. Looking more closely at the closed-form solution, we see that the second and third terms die out more quickly than the first term

$$\vec{p}_k = \underbrace{(0.93589)^k}_{\text{dies out slowly}}\ 3.74104\ \vec{v}_1 - \underbrace{(-0.4)^k}_{\text{dies out quickly}}\ 4.24595\ \vec{v}_2 - \underbrace{(0.0641101)^k}_{\text{dies out very quickly}}\ 3.84334\ \vec{v}_3.$$

This means that after some years we can approximate the population by

$$\vec{p}_k \approx (0.93589)^k\ 3.74104\ \vec{v}_1.$$

n	**Population by Age Group** (a = **age**)								
	Original Model			**Closed Form Solution**			**Approximation**		
	$0 \le a < 1$	$1 \le a < 2$	$2 \le a$	$0 \le a < 1$	$1 \le a < 2$	$2 \le a$	$0 \le a < 1$	$1 \le a < 2$	$2 \le a$
0	5.00000	0.00000	0.00000	5.00000	0.00000	0.00000	2.95685	1.89564	1.28800
1	2.00000	3.00000	0.00000	2.00000	3.00000	0.00000	2.76729	1.77411	1.20542
2	2.90000	1.20000	1.50000	2.90000	1.20000	1.50000	2.58988	1.66037	1.12814
3	2.30000	1.74000	0.90000	2.30000	1.74000	0.90000	2.42384	1.55393	1.05582
4	2.31800	1.38000	1.05000	2.31800	1.38000	1.05000	2.26845	1.45430	0.98813
5	2.10320	1.39080	0.90000	2.10320	1.39080	0.90000	2.12302	1.36107	0.92478
6	1.99484	1.26192	0.87540	1.99484	1.26192	0.87540	1.98691	1.27381	0.86549
7	1.85636	1.19690	0.80604	1.85636	1.19690	0.80604	1.85953	1.19215	0.81001
8	1.74158	1.11382	0.75966	1.74159	1.11382	0.75966	1.74032	1.11572	0.75808
9	1.62824	1.04495	0.70884	1.62824	1.04495	0.70884	1.62875	1.04419	0.70948
10	1.52453	0.97694	0.66424	1.52453	0.97694	0.66424	1.52433	0.97725	0.66399
11	1.42652	0.91472	0.62132	1.42652	0.91472	0.62132	1.42660	0.91460	0.62142
12	1.33517	0.85591	0.58162	1.33517	0.85591	0.58162	1.33514	0.85596	0.58158
13	1.24953	0.80110	0.54428	1.24953	0.80110	0.54428	1.24955	0.80109	0.54430
14	1.16944	0.74972	0.50941	1.16944	0.74972	0.50941	1.16944	0.74973	0.50940
15	1.09446	0.70167	0.47674	1.09446	0.70167	0.47674	1.09447	0.70166	0.47675
16	1.02430	0.65668	0.44618	1.02430	0.65668	0.44618	1.02430	0.65668	0.44618
17	0.95863	0.61458	0.41757	0.95863	0.61458	0.41758	0.95863	0.61458	0.41758
18	0.89717	0.57518	0.39080	0.89717	0.57518	0.39081	0.89717	0.57518	0.39081
19	0.83965	0.53830	0.36575	0.83966	0.53830	0.36575	0.83966	0.53830	0.36575
20	0.78582	0.50379	0.34230	0.78583	0.50379	0.34230	0.78583	0.50379	0.34230

Table 4.10: Comparing the Model, the Closed Form Solution, and an Approximation

Table 4.10 compares the results of direct calculations using the original model, calculations using the closed-form solution, and calculations using the approximate model $\vec{v}_k \approx (0.93589)^k\ 3.74104\ \vec{v}_1$. As expected, the first two sets of calculations yield the same results (except for an occasional small round-off error) and the third set of calculations is close to the others only after some years have passed.

The next theorems continue our list of deductions[11] that can be made from our work so far.

[11] By "deductions" we mean conclusions that follow from this work. This is an example of "deductive reasoning," which can be contrasted with "inductive reasoning" in which conclusions are based on examples and experiments.

Theorem 5 *Consider the IVP*

$$\vec{p}_k = A\vec{p}_{k-1}, \quad \vec{p}_0 = \vec{b}.$$

If the matrix A has n distinct real eigenvalues, one of which is 1 and the rest of which have absolute value less than one, then

$$\lim_{k\to\infty} \vec{p}_k = \vec{v}$$

where $\vec{v}$ is an eigenvector associated with the eigenvalue 1.

Proof

We will assume that $\lambda_1 = 1$. If this wasn't true, we could just renumber the eigenvalues and eigenvectors. Now using our closed form solution,

$$\vec{p}_k = \lambda_1^k c_1 \vec{v}_1 + \lambda_2^k c_2 \vec{v}_2 + \cdots + \lambda_n^k c_n \vec{v}_n,$$

we see that

$$\lim_{k\to\infty} \vec{p}_k = \lim_{k\to 0}(\lambda_1^k c_1 \vec{v}_1 + \lambda_2^k c_2 \vec{v}_2 + \cdots + \lambda_n^k c_n \vec{v}_n),$$

and, since $\lambda_1 = 1$ and all the remaining eigenvalues have absolute value less than one,

$$\begin{aligned}
\lim_{k\to\infty} \lambda_1^k &= 1 \\
\lim_{k\to\infty} \lambda_2^k &= 0 \\
&\vdots \\
\lim_{k\to\infty} \lambda_n^k &= 0
\end{aligned}$$

and we see that

$$\lim_{k\to\infty} \vec{p}_k = c_1 \vec{v}_1.$$

Since any multiple of an eigenvector is an eigenvector, this completes the proof. ■

Theorem 6 *Consider the IVP*

$$\vec{p}_k = A\vec{p}_{k-1}, \quad \vec{p}_0 = \vec{b}.$$

If the matrix A has n distinct, real eigenvalues, one of which is bigger than 1 and the rest of which have absolute value less than 1, then, as usual, we can find a closed-form solution

$$\vec{p}_k = c_1\lambda_1^k\vec{v}_1 + c_2\lambda_2^k\vec{v}_2 + \cdots + c_n\lambda_n^k\vec{v}_k.$$

Suppose that λ_1 is the eigenvalue that is bigger than 1. Notice that, after many years, we see

$$\vec{p}_k = c_1 \underbrace{\lambda_1^k}_{\text{grows}} \vec{v}_1 + c_2 \underbrace{\lambda_2^k}_{\text{dies out}} \vec{v}_2 + \cdots + c_n \underbrace{\lambda_n^k}_{\text{dies out}} \vec{v}_k.$$

So, the vectors p_k will continue to grow and look more and more like $c_1\lambda_1^k\vec{v}_1$ after many years[12]*.*

Proof

The proof is left to the reader. ■

Although we haven't covered all of the possibilities (For example, there might be several eigenvalues that are bigger than one.), we have covered the possibilities that occur most frequently. The other possibilities can be analyzed using the same ideas we used to analyze the possibilities that we did cover.

Question 1 *Using the techniques we've developed in the last few sections, determine the long term behavior of the IVP*

$$\vec{p}_k = \begin{bmatrix} 0.4 & 0.2 \\ 0.6 & 0.8 \end{bmatrix} \vec{p}_{k-1}, \quad \vec{p}_0 = \begin{bmatrix} 0.80 \\ 0.20 \end{bmatrix}.$$

Question 2 *Using the techniques we've developed in the last few sections, determine the long term behavior of the IVP*

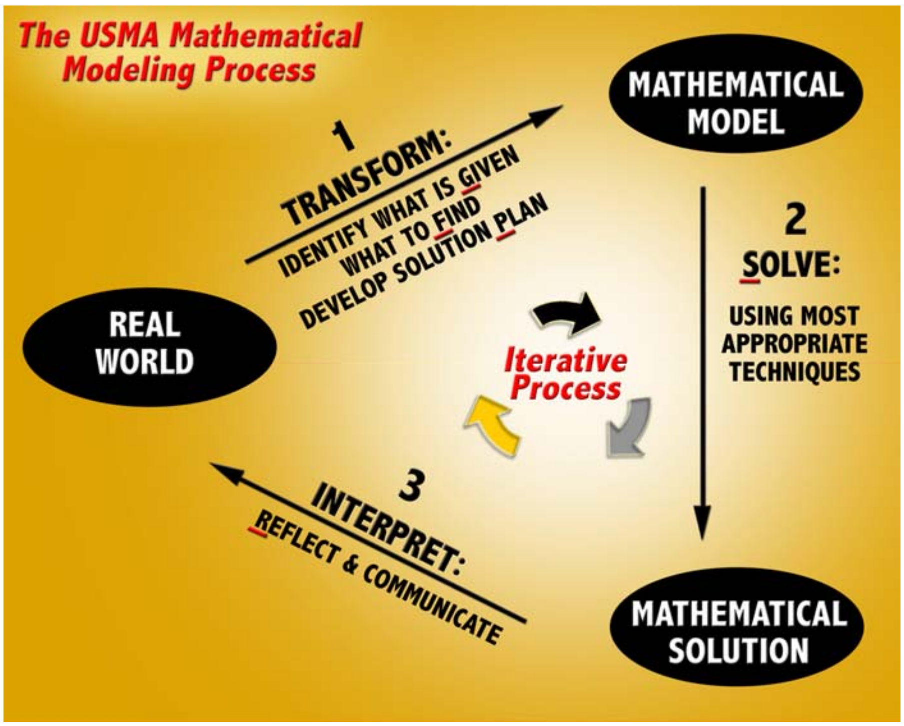
$$\vec{p}_k = \begin{bmatrix} 0.4 & 0.3 \\ 0.6 & 0.8 \end{bmatrix} \vec{p}_{k-1}, \quad \vec{p}_0 = \begin{bmatrix} 0.80 \\ 0.20 \end{bmatrix}.$$

[12] Note that if $c_1 = 0$ then $\lim_{k\to\infty} \vec{p}_k = 0$.

"Experience with mathematical models of thought builds mathematical power – a capacity of mind of increasing value in the technological age that enables one to read critically, to identify fallacies, to detect bias, to assess risk, and to suggest alternatives. Mathematics empowers us to understand better the information-laden world in which we live." National Research Council, 1989

"Mathematics is the door and the key to the sciences." Roger Bacon

"Difficulties strengthen the mind, as labor does the body." Seneca



"All wish to possess knowledge, but few, comparatively speaking, are willing to pay the price." – Juvenal

"The mind, when stretched by a new idea, never returns to its original shape." Oliver Wendell Homes

"In the end, though, the success of the National Security Agency -- and by extension, the national security of the United States -- hinges on the health of American math education."-James R. Schatz, *National Security Agency*

"There are no shortcuts to any place worth going." – Unknown