Mathematical Modeling, Proof of Concept

Mathematical Modeling

Gerhard Dangelmayr and Michael Kirby

Department of Mathematics, Colorado State University, Fort Collins, Colorado 80524

E-mail address: gerhard@math.colostate.edu, kirby@math.university.edu

This project is supported in part by NSF Grant DUE-#0126650. ©2002.

Contents

Preface vii

Chapter 1. Mathematical Modeling 1
1.1. Examples of Modeling 1
1.2. The Modeling Process 3
1.3. Purpose of this Course 4

Chapter 2. Qualitative Modeling with Functions 7
2.1. Modeling Species Propagation 7
2.2. Supply and Demand 8
2.3. Modeling with Proportion and Scale 11
2.4. Dimensional Analysis 18

Chapter 3. Data Fitting 27
3.1. Linear Least Squares 27
3.2. Splines 33

Chapter 4. Difference Equations 39
4.1. Analytical Solutions 41
4.2. Population Modeling 51
4.3. Empirical Modeling 54

Chapter 5. Linear Programming 65
5.1. Examples of Linear Programs 66
5.2. Geometric Solution of a 2D Linear Program 67
5.3. Sensitivity Analysis 68
5.4. Data Fitting and the Uniform Approximation 70

Chapter 6. Nonlinear Programming 75
6.1. Unconstrained Optimization in One Dimension 76
6.2. Unconstrained Optimization in Higher Dimensions 77
6.3. Constrained Optimization and Lagrange Multipliers 79

Chapter 7. Simulation Modeling 83
7.1. The Tire Distributor Problem 83
7.2. Blackjack Strategy 85

Chapter 8. Final Course Projects 91
8.1. Quarterback Passer Ratings 92
8.2. Baseball Simulation 95
8.3. Predicting the Stock Market 97
8.4. Population Biology 99

Appendix A. Matlab Code for Data Fitting 101
A.1. Mammalian Heart Rate Problem 101
A.2. Least Squares with Normal Equations 103
A.3. Least Squares with Overdetermined System 105
A.4. Non-Newtonian Fish 106
A.5. Preditor or Prey? 106
A.6. Tire Distributor 107
A.7. Blackjack 109

Appendix B. Solutions 113

Appendix. Bibliography 115

Appendix. Index 117

Preface

This is the first draft of a set of course notes in mathematical modeling. These materials are the result of the National Science Foundation DUE-0126650 award entitled "Mathematical Modeling Program for Undergraduates in Science, Mathematics, Engineering & Technology".

These notes are incomplete and not intended for general dissemination at this time. If you would like to learn more about this project please contact one of the authors.

Gerhard Dangelmayr and Michael Kirby

CHAPTER 1

Mathematical Modeling

Mathematical modeling is becoming an increasingly important subject as computers expand our ability to translate mathematical equations and formulations into concrete conclusions concerning the world, both natural and artificial, that we live in.

1.1. Examples of Modeling

Here we do a quick tour of several examples of the mathematical process. We present the models as finished results as opposed to attempting to develop the models.

1.1.1. Modeling with Difference Equations. Consider the situation in which a variable changes in discrete time steps. If the current value of the variable is a_n then the predicted value of the variable will be a_{n+1}. A mathematical model for the evolution of the (still unspecified) quantity a_n could take the form

$$a_{n+1} = \alpha a_n + \beta$$

In words, the new value is a scalar multiple of the old value offset by some constant β. This model is common, e.g., it is used for modeling bank loans. One might amend the model to make the dependence depend on more terms and to include the possibility that every iteration the offset can change, thus,

$$a_{n+1} = \alpha_1 a_n + \alpha_2 a_n^2 + \beta_n$$

This could correpsond to, for example, a population model where the the migration levels change every time step. In some instances, it is clear that information required to predict a new value goes back further than the current value, e.g.,

$$a_{n+1} = a_n + a_{n-1}$$

Note now that two initial values are required to evolve this model. Finally, it may be that the form of the difference equations are unknown and the model must be written

$$a_{n+1} = f(a_n, a_{n-1}, a_{n-M-1})$$

Determining the nature of f and the step M is at the heart of model formulation with difference equations. Often observed data can be employed to assist in this effort.

1.1.2. Modeling with Ordinary Differential Equations. Although modeling with ordinary differential equations shares many of the ideas of modeling with the difference equations discussed above, there are many fundamental differences. At the center of these differences is the assumption that time is a continuous variable.

One of the simplest differential equations is also an extremely important model, i.e.,

$$\frac{dx}{dt} = \alpha x$$

In words, the rate of change of the quantity x depends on the amount of the quantity. If $\alpha > 0$ then we have exponential growth. If $\alpha < 0$ the situation is exponetial decay. Of course additional terms can be added that fundamentally alter the evolution of $x(t)$. For example

$$\frac{dx}{dt} = \alpha_1 x + \alpha_2 x^2$$

The model formulation again requires the development of the appropriate right-hand side.

In the above model the value x on the right hand side is implicitly assumed to be evaluated at the time t. It may be that there is evidence that the instantaneous rate of change at time t is actually a function of a previous time, i.e.,

$$\frac{dx}{dt} = f(x(t)) + g(x(t-\tau))$$

This is referred to as a delay differential equation.

1.1.3. Modeling with Partial Differential Equation. In the previous sections on modeling the behaviour of a variable as a function of time we assumed that there was only one independent variable. Many situations arise in practice where the number of independent variables is larger than two. For spatio-temporal models we might have time and space (hence the name!), e.g.,

$$\frac{\partial f}{\partial t} = \alpha \frac{\partial^2 f}{\partial x^2}$$

or

$$\frac{\partial^2 f}{\partial t^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

1.1.4. Optimization. In many modeling problems the goal is to compute the "best" solution. This may correspond to maximizing profit in a company, or minimizing loss in a conflict. It is no surprise that optimization techniques take a central seat in the mathematical modeling literature.

Now one may allow $x \in \mathbb{R}^n$ and require that

$$x^* = \arg\min f(x)$$

The quantity $f(x)$ is referred to as the *objective function* while the vector x consists of decision variables. Because x sits in $\mathbb{R}^n$ the problem is referred to as unconstrained.

Alternatively, one might require that the solution x have all positive components. If we refer to this set as S then the optimization problem is constrained

$$x^* = \arg\min_{x \in S} f(x)$$

If the objective function as well as the equations that define the constraint set are linear, than the optimization problem is called a *linear programming problem*. Otherwise, the problem is referred to as a nonlinear programming problem. As we shall see, solution methods for linear and nonlinear programming problems are very different.

1.1.5. Modeling with Simulations. Many problems may afford a mathematical formulation yet be analytically intractable. In these situations a computer can implement the mathematics literally and repetitively often times to extreme advantage.

1.1.5.1. *Simulating Games.*

- What is the probability that you can win a game of solitaire?
- What is the best strategy for playing blackjack?
- Given a baseball team consisting of certain players, in what order should they hit?

On the other hand, computer simulations can be employed to model evolution equations. Applications in the realm of fluid dynamics and weather prediction are well established. A striking new example of such simulation modeling is attempting to model electrical activity in the brain.

1.1.6. Function Fitting: Data Modeling. Often data is available from a process to assist in the modeling. How can functions be computed that reflect the relationships between variables in the data. Produce a model

$$y = f(x; w)$$

and using the set of input output pairs compute the parameters w. In some cases the form of f may be guessed. In other cases a model free approach can be used.

1.2. The Modeling Process

The goal in all modeling problems is *added value*. Something novel must be learned from the modeling process or one has completed an exercise in futility, or mathematical wheel spinning, depending on your perspective. There are many obvious questions the answers to which have inherent added value. For example:

- Should a stock be bought or sold?
- Is the earth becoming warmer?
- Does creating a law have a positive or negative societal effect?
- What is the most valuable property in monopoly?

Clearly this is a very small start to an extremely long list.

1.2.1. An Algorithm for Modeling? The modeling process has a sequence of common steps that serve as an abstraction for the modeler:

- Identify the problem and questions.
- Identify the relevant variables in a problem.
- Simplify until tractable.
- Relate these variables mathematically.
- Solve.
- Does the solution provide added value?
- Tweak model and compare solutions.

1.3. Purpose of this Course

The primary goal of this course is to assist the student in developing an understanding of why mathematics is useful as a language for characterizing the interaction among variables.

The arena for mathematical modeling is so large that there can be a trade-off between learning the specifics of a particular problem and the mathematical tools for solving it. Here we emphasize not the special knowledge often required to make expert conclusions concerning models, but rather the efficacy of certain mathematics for constructing models.

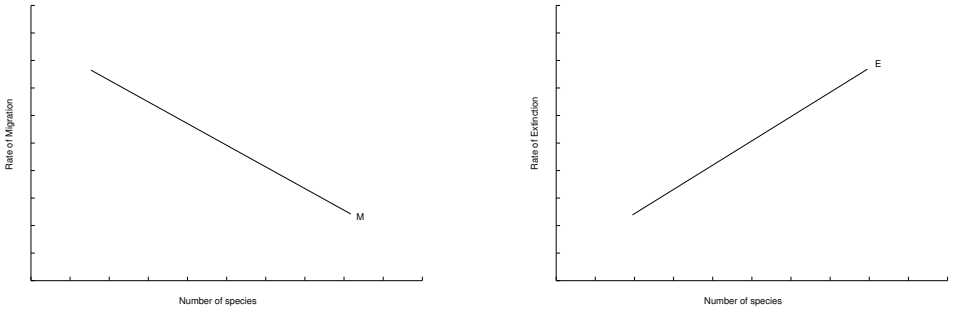
Problems

PROBLEM 1.1. Name three problems that might be modeled mathematically. Why do you think mathematics may provide a key to each solution. What is the added value in each case?

PROBLEM 1.2. Consider the differential equation

$$\frac{dx}{dt} = x$$

Translate this model to a difference equation. Compare the solutions and discuss.



(a) Migration curve.

(b) Extinction curve

FIGURE 1. Qualitative form of the migration and extinction curves.

CHAPTER 2

Qualitative Modeling with Functions

It is often surprising that very simple mathematical modeling ideas can produce results with added value. Indeed, the solutions may be elegant and provide quality of understanding that obviates further exploration by more technical or complex means. In this chapter we explore a few simple approaches to qualitatively modeling phenomena with well-behaved functions.

2.1. Modeling Species Propagation

This problem concerns the factors that influence the number of species existing on an island. The discussion is adapted from [**?**].

One might speculate that factors affecting the number of species could include

- Distance of the island from the mainland
- Size of the island

Of course limiting ourselves to these influences has the dual effect of making a tractable model that needs to be recognized as omitting many possible factors.

The number of species may increase due to new species discovering the island as a suitable habitat. We will refer to this as the *migration rate*. Alternatively, species may become extinct due to competition. We will refer to this as the *extinction rate*. This discussion will be simplified by employing an aggregate total for the number of species and not attempting to distinguish the nature of each species, i.e., birds versus plants.

Now we propose some basic modeling assumptions that appear reasonable.

The migration rate of new species decreases as the number of species on the island increases.

The argument for this is straight forward. The more species on an island the smaller the number of new species there is to migrate. See Figure 1 (a) for a qualitative picture.

The extinction rate of species increases as the number of species on the island increases.

Clearly the more species there are the more possibilities there are for species to die out. See Figure 1 (b) for a qualitative picture.

If we plot the extinction rate and the migration rate on a single plot we identify the point of intersection as an equilibrium, i.e., the migration is exactly offset by the extinction and the number of species on the island is a constant. We will assume in this discussion that we are considering islands for which the number of species is roughly constant over time, i.e., they are in a state of equilibrium.

Now we consider whether this simple model provides any added value. In particular, can it be used to address our questions posed at the outset.

First, what is the effect of the distance of the island from the mainland on the number of bird species? One can characterize this effect by a shift in the migration curve. The further the island is away from the mainland, the less likely a species is to successfully migrate. Thus the migration curve is shifted down for *far* islands and shifted up for *near* islands. Presumably, this distance of the island from the mainland has no impact on the extinction curve. Thus, by examining the shift in the equilibrium, we may conclude that the number of species on an island decreases as the island's distance from the mainland increases. See Figure 2.

Note in this model we assume that the time-scales are small enough that new species are not developed via evolution. While this may seem reasonable there is evidence that in some extreme climates, such as those found in the Galapagos Islands, variation may occur over shorter periods. There have been 140 different species of birds

2.2. Supply and Demand

In this section we sketch a well-known concept in economics, i.e., supply and demand. We shall see that relatively simple laws, when taken together, afford interesting insight into the relationship between producers and consumers. Furthermore, we may use this framework to make predictions such as

- What is the impact of a tax on the sale price?
- What is the impact an increase in employees wages on sales price? Can the owner of the business pass this increase on to the consumer?

Law of Supply: An increase in the price of a commodity will result in an increase of the amount supplied.

Law of Demand: If the price of a commodity increases, then the quantity demanded will decrease.

Thus, we may model the supply curve qualitatively by a monotonically increasing function. For simplcity we may assume a straight line with positive slope. Analogously, we may model the demand curve qualitatively by a monotonically decreasing function, which again we will take as a straight line.

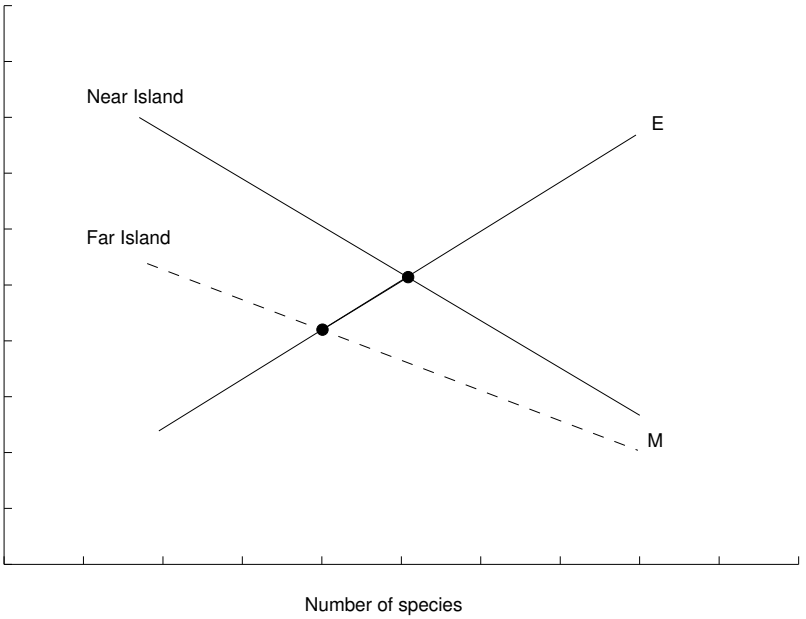


FIGURE 2. The effect of distance of the island from the mainland is to shift the migration curve. Consequently the equilibrium solution dictates a smaller number of species will be supported for islands that are farther away from the mainland.

A flat demand curve may be interpreted as consumers being very sensitive to the price of a commodity. If the price goes up just a little, then the quantity in demand goes down significantly. Steep and flat supply and demand curves all have similar qualitative interpretations (see the problems).

2.2.1. Market Equilibrium. Given a supply curve and a demand curve we may plot them on the same axis and note their point of intersection (q_*, p_*). This point is special for the following reason:

- The seller is willing to supply q_* at the price p_*
- The demand is at the price p_* is q_*

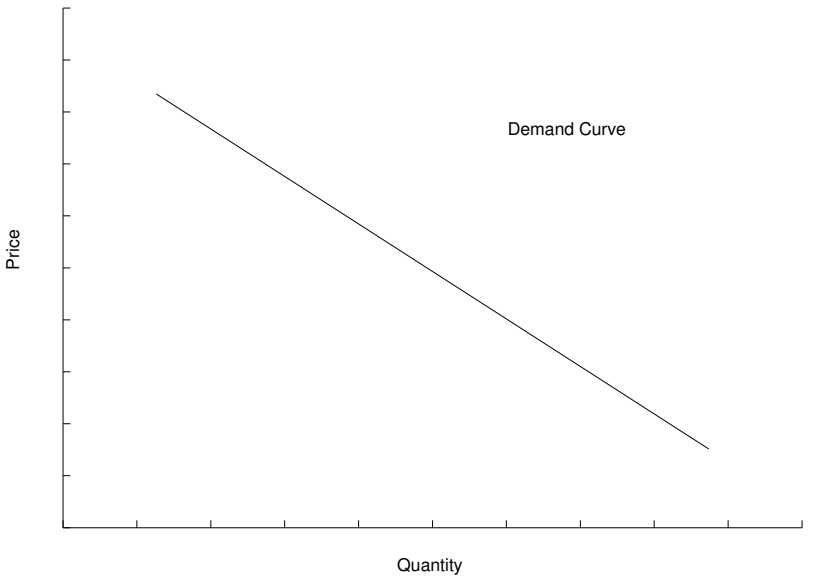
So both the supplier(s) and the purchaser(s) are happy economically speaking.

2.2.2. Market Adjustment. Of course, in general markets do not exist in the perfect economic utopia described above. We may model the market adjustment as a sequence of points on the demand and supply curves.

Based on market research it is estimated that consumers will demand a quantity q_1 at a price p_1. The supply and demand curves will permit a prediction of how the market will evolve. For simplicity, we will assume that the initial point (q_1, p_1) is on the demand curve to the right of the equilibrium point.

At the price p_1 the supplier looks to his supply curve and proposes to sell a reduced quantity q_2. Thus we move from right to left horizontally. Note that moving vertically to the supply curve would not make sense as this would correspond to offering the quantity q_1 at an increased price. These goods will not sell at this price.

(a) Supply curve



(b) Demand curve

FIGURE 3. (a) Qualitative form of supply and demand curves.

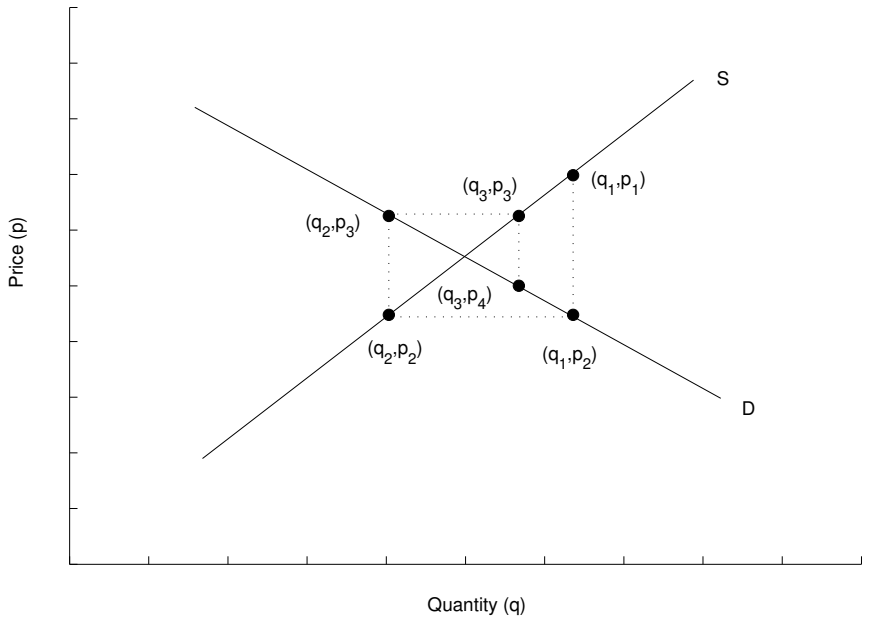


FIGURE 4. The cobweb model illustrating a sequence of market adjustments.

From the point (q_1, p_2) the consumer will respond to the new reduced quantity q_2 by being willing to pay more. This corresponds to moving vertically upward to the new point (q_2, p_2) on the supply curve.

Now the supplier adjusts to the higher price being paid in the market place by increasing the quantity produced to q_3. This process then continues, in theory, until an equilbrium is reached. It is possible that this will never happen, at least not without a basic adjustment to the shape of either the supply of demand curves, for example through cost cutting methods such as improved efficiency, or layoffs.

2.2.3. Taxation. The effect of a new tax on a product is to shift the demand curve down because consumers will not be willing to pay as much for the product (before the tax). Note that this leads to a new equilibrium point which reduces the price paid to the seller per item and reduces the quantity supplied by the producer. Thus one may conclude from this picture that the effect of a tax on alcohol is to reduce consumption as well as profit for the supplier. See Figure 5.

2.3. Modeling with Proportion and Scale

In the previous sections we have considered how simple functions may be employed to qualitatively model various situations and produce added value. Now we turn to considerations that assist in determining the nature of these functional dependencies in more complex terms.

2.3.1. Proportion. If a quantity y is *proportional* to a quantity x then we write

$$y \propto x$$

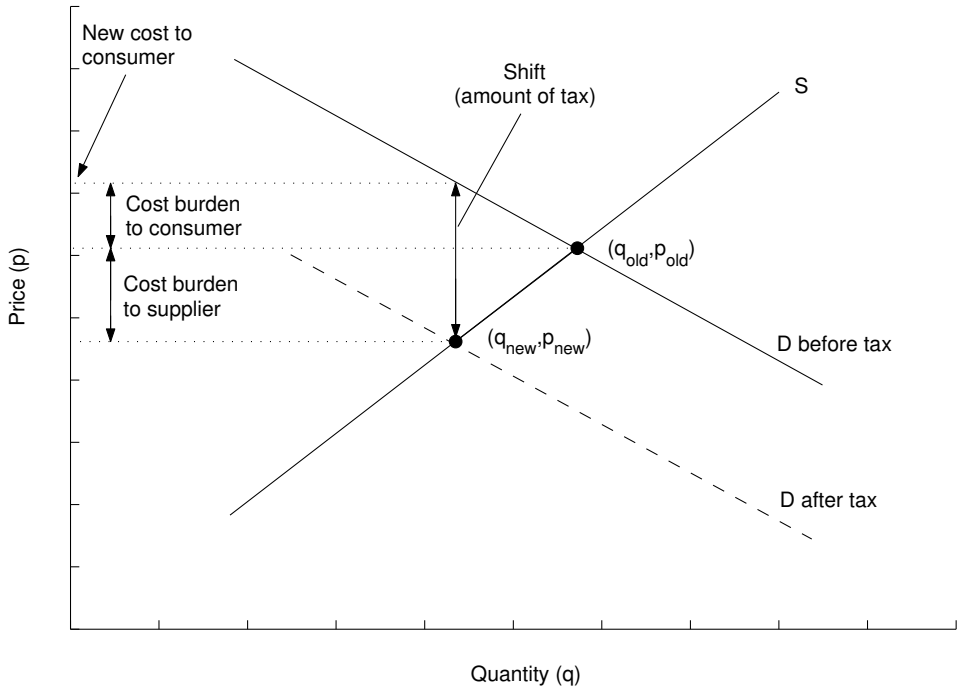


FIGURE 5. A tax corresponds to a downwards shift in the demand curve.

by which is meant

$$y = kx$$

for some constant of proportionality k.

EXAMPLE 2.1. In 1678 Robert Hooke proposed that the restoring force F of a spring is proportional to its elongation e, i.e.,

$$F \propto e$$

or,

$$F = ke$$

where k is the *stiffness* of the spring.

Note that the property of proportionality is symmetric, i.e.,

$$y \propto x \to x \propto y \tag{2.1}$$

and transitive, i.e.,

$$y \propto x \quad \text{and} \quad z \propto y \to z \propto x \tag{2.2}$$

EXAMPLE 2.2. If $y = kx + b$ where k, b are constants, then

$$y \not\propto x$$

but

$$y - b \propto x$$

Inverse proportion. If $y \propto 1/x$ then y is said to be *inversely* proportional to x.

EXAMPLE 2.3. If y varies inversely as the square-root of x then

$$y = \frac{k}{\sqrt{x}}$$

Joint Variation. The volume of a cylinder is given by

$$V = \pi r^2 h$$

where r is the radius and h is the height. The volume is said to vary *jointly* with r^2 and h, i.e.,

$$V \propto r^2 \quad \text{and} \quad V \propto h$$

EXAMPLE 2.4. The volume of a given mass of gas is proportional to the temperature and inversely proportional to the pressure, i.e., $V \propto T$ and $V \propto 1/P$, or,

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

EXAMPLE 2.5. Frictional drag due to the atmosphere is jointly proportional to the surface area S and the velocity v of the object.

Superposition of Proportions. Often a quantity will vary as the sum of proportions.

EXAMPLE 2.6. The stopping distance of a car when an emergency situation is encountered is the sum of the reaction time of the driver and the amount of time it takes for the breaks to dissipate the energy of the vehicle. The reaction distance is proportional to the velocity. The distance traveled once the breaks have been hit is proportional to the velocity squared. Thus,

$$\text{stopping distance} = k_1 v + k_2 v^2$$

EXAMPLE 2.7. Numerical error in the computer estimation of the center difference formula for the derivative is given by

$$e(h) = \frac{c_1}{h} + c_2 h^2$$

where the first term is due to roundoff error (finite precision) and the second term is due to truncation error. The value h is the distance δx in the definition of the derivative.

2.3.1.1. *Direct Proportion.* If

$$y \propto x$$

we say y varies in *direct* proportion to x. This is not true, for example, if $y \propto r^2$. On the other hand, we may construct a direct proportion via the obvious change of variable $x = r^2$. This simple trick always permits the investigation of the relationship between two variables such as this to be recast as a direct proportion.

2.3.2. Scale. Now we explore how the size of an object can be represented by an appropriate length scale if we restrict our attention to replicas that are *geometrically similar*. For example, a rectangle with sides l_1 and w_1 is geometrically similar to a rectangle with sides l_2 and w_2 if

$$\frac{l_1}{l_2} = \frac{w_1}{w_2} = k \tag{2.3}$$

As the ratio $\kappa = l_1/w_1$ characterizes the geometry of the rectangle it is referred to as the *shape factor*. If two objects are geometrically similar, then it can be

(a) Plot of y against r for $y = r^2$

(b) Plot of y against r^2

(c) Plot of y against r for $y = kr(r+1)$

(d) Plot of y against $r(r+1)$

FIGURE 6. Simple examples of how a proportion may be converted to a direct proportion.

shown that they have the same shape factor. This follows directly from multiplying Equation (2.3) by the factor l_2/w_1, i.e.,

$$\frac{l_1}{w_1} = \frac{l_2}{w_2} = k\frac{l_2}{w_1}$$

2.3.2.1. *Characteristic Length.* Characteristic length is useful concept for characterizing a family of geometrically similar objects. We demonstrate this with an example.

Consider the area of a rectangle of side l and width w where l and w may vary under the restriction that the resulting rectangle be geometrically similar to the rectangle with length l_1 and width w_1. An expression for the area of the varying

(a) (b) (c)

FIGURE 7. Examples of characteristic lengths for the rectangle.

triangle can be simplified as a consequence of the constraint imposed by geometric similarity. To see this

$$\begin{aligned} A &= lw \\ &= l(\frac{w_1 l}{l_1}) \\ &= \kappa l^2 \end{aligned}$$

where $\kappa = w_1/l_1$, i.e., the shape factor.

EXAMPLE 2.8. Watering a farmer's rectangular field requires an amount of area proportional to the area of the field. If the characteristic length of the field is doubled, how much additional water q will be needed, assuming the new field is geometrically similar to the old field? Solution: $q \propto l^2$, i.e., $q = kl^2$. Hence

$$q_1 = kl_1^2$$

$$q_2 = kl_2^2$$

Taking the ratio produces

$$\frac{q_1}{q_2} = \frac{l_1^2}{l_2^2}$$

Now if $q_2 = 100$ acre feet of water are sufficient for a field of length $l_2 = 100$, how much water will be required for a field of length $l_1 = 200$? Sol.

$$q_1 = q_2 \frac{l_1^2}{l_2^2} = 100 \frac{200^2}{100^2} = 400 \,\text{acre feet} \quad \square$$

EXAMPLE 2.9. Why are gymnasts typically short? It seems plausible that the ability A, or natural talent, of gymnast would be proportional to strength and inversely proportional to weight, i.e.,

$$A \propto \text{strength}$$

and

$$A \propto \frac{1}{\text{weight}}$$

and taken jointly

$$A \propto \frac{\text{strength}}{\text{weight}}$$

One model for strength is that the strength of a limb is proportional to the cross-sectional area of the muscle. The weight is proportional to the volume (assuming constant density of the gynmast). Now, assuming all gymnasts are geometrically similar with characteristic length l

$$\text{strength} \propto \text{muscle area} \propto l^2$$

and

$$\text{weight} \propto \text{volume} \propto l^3$$

so the ability A follows

$$A \propto \frac{l^2}{l^3} \propto \frac{1}{l}$$

So shortness equates to a talent for gymnastics.

□

EXAMPLE 2.10. Proportions and terminal velocity. Consider a uniform density spherical object falling under the influence of gravity. The object will travel will constant (terminal) velocity if the accelerating force due to gravity $F_g = mg$ is balance exactly by the decelerating force due to atmospheric friction $F_d = kSv^2$; S is the cross-sectional surface area and v is the velocity of the falling object. Our equilibrium condition is then

$$F_g = F_d$$

Since surface area satisfies $S \propto l^2$ it follows $l \propto S^{1/2}$. Given uniform density $m \propto w \propto l^3$ so it follows $l \propto m^{1/3}$. Combining proportionalities

$$m^{1/3} \propto S^{1/2}$$

from which it follows by substitution into the force equation that

$$m \propto m^{2/3} v^2$$

or, after simplifying,

$$v \propto m^{1/6}$$

□

EXAMPLE 2.11. In this example we will attempt to model observed data displayed in Table 1 that relates the heart rate of mammals to there body weight. From the table we see that we would like to relate the heart rate as a function of body weight. Smaller animals have a faster heart rate than larger ones. But how do we estimate this proportionality?

We begin by assuming that all the energy E produced by the body is used to maintain heat loss to the environment. This heat loss is in turn proportional to the surface area s of the body. Thus,

$$E \propto s$$

The energy available to the body is produced by the process of respiration and is assumed to be proportional to the oxygen available which is in turn proportional to the blood flow B through the lungs. Hence, $B \propto s$ If we denote the pulse rate as r we may assume

$$B \propto rV$$

where V is the volume of the heart.

mammal	body weight (g)	pulse rate
shrew[2]	3.5	782
pipistrelle bat[1]	4	660
bat[2]	6	588
mouse[1]	25	670
hamster[2]	103	347
kitten[2]	117	300
rat[1]	200	420
rat[2]	252	352
guinea pig[1]	300	300
guinea pig[1]	437	269
rabbit[2]	1,340	251
rabbit[1]	2,000	205
oppossum[2]	2,700	187
little dog[1]	5,000	120
seal[2]	22,500	100
big dog[1]	30,000	85
goat[2]	33,000	81
sheep[1]	50,000	70
human[1]	70,000	72
swine[2]	100,000	70
horse[2]	415,000	45
horse[1]	450,000	38
ox[1]	500,000	40
elephant[1]	3,000,000	48

TABLE 1. Superscript 1 data source A.J. Clark; superscript 2 data source Altman and Dittmer.

We still need to incorporate the body weight w into this model. If we take W to be the weight of the heart assuming constant density of the heart it follows

$$W \propto V$$

Also, if the bodies are assumed to be geometrically similar then $w \propto W$ so by transitivity $w \propto V$ and hence

$$B \propto rw$$

Using the geometric similarity again we can relate the body surface area s to its weight w. From characteristic length scale arguments

$$v^{1/3} \propto s^{1/2}$$

so

$$s \propto w^{2/3}$$

from which we have $rw \propto w^{2/3}$ or

$$r = kw^{-1/3}$$

To validate this model we plot $w^{-1/3}$ versus r for the data Table 8. We see that for the larger animals with slower heart rates that this data appears linear and suggests this rather crude model actually is supported by the data. For much

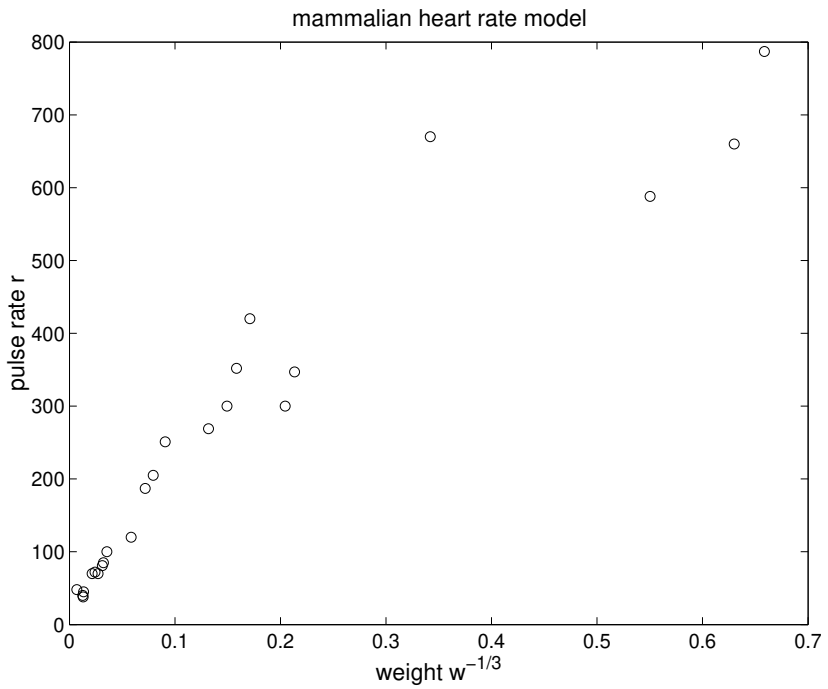


FIGURE 8. Testing the model produced by proportionality. For the model to fit, the data should sit on a straight line eminating from the origin.

smaller animals there appear to be factors that this model is not capturing and the data falls off the line.

2.4. Dimensional Analysis

In this chapter we have explored modeling with functions and proportion. In some instances, such as the mammalian heart rate, it is possible to cobble enough information together to actually extract a model; in particular, to identify the functional form for the relationship between the dependent and independent variables. Now we turn to a surprisingly powerful and simple tool known as *dimensional analysis*[1].

Dimensional analysis operates on the premise that equations contain terms that have units of measurement and that the validity of these equations, or laws, are not dependent on the system of measurement. Rather these equations relate variables that have inherent physical dimensions that are derived from the fundamental dimensions of *mass, length* and *time*. We label these dimensions generically as M, L and T, respectively.

As we shall see, dimensional analysis provides an effective tool for mathematical modeling in many situations. In particular, some benefits include

- determination of the form of a joint proportion
- reduce number of variables in a model

[1]This dimension should not be confused with the usual notion of geometric dimension.

- enforcement of dimensional consistency
- ability to study scaled versions of models

2.4.1. Dimensional homogeneity. An equation is said to be *dimensionally homogeneous* if all the terms in the equation have the same physical dimension.

EXAMPLE 2.12. All the laws of physics are dimensionally homogeneous. Consider Newton's law

$$F = ma$$

The units on the right side are

$$M \cdot \frac{L}{T^2}$$

so we conclude that the physical dimension of a force must be MLT^{-2}. □

EXAMPLE 2.13. The equation of motion of a linear spring with no damping is

$$m\frac{d^2x}{dt^2} + kx = 0$$

What are the units of the spring constant? Dimensionally we can recast this equation as

$$MLT^{-2} + M^aL^bT^cL = 0$$

Matching exponents for each dimension permits the calculation of a, b and c.

$$\begin{aligned} M: &\quad a = 1 \\ L: &\quad 1 = b + 1 \\ T: &\quad -2 = c \end{aligned}$$

Thus we conclude that the spring constant has the dimensions MT^{-2}. □

EXAMPLE 2.14. Let v be velocity, t be time and x be distance. The model equation

$$v^2 = t^2 + \frac{x}{t}$$

is dimensionally inconsistent.

EXAMPLE 2.15. An angle may be defined by the formula

$$\theta = \frac{s}{r}$$

where the arclength s subtends the angle θ and r is the radius of the circle. Clearly this angle is dimensionless.

2.4.2. Discovering Joint Proportions. If in the formulation of a problem we are able to identify a dependent and one or more independent variables, it is often possible to identify the form of a joint proportion. The form of the proportion is actually constrained by the fact that the equations must be dimensionally consistent.

EXAMPLE 2.16. In this problem we consider the drag force F_D on an airplane. As our model we propose that this drag force (dependent variable) is proportional to the independent variables

- cross-sectional area A of airplane
- velocity v of airplane
- density ρ of the air

As a joint proportion we have

$$F_D = kA^a v^b \rho^c$$

where a, b and c are unknown exponents. As a consequence of dimensional consistency we have

$$\begin{aligned} MLT^{-2} &= (M^0L^0T^0)(L^2)^a(\frac{L}{T})^b(\frac{M}{L^3})^c \\ &= M^cL^{2a-3c+b}T^{-b} \end{aligned}$$

From the M exponent we conclude $c = 1$. From the T exponent $b = 2$ and from the L exponent it follows that $1 = 2a - 3c + b$, whence $a = 1$. Thus the only possibility for the form of this joint proportion is

$$F_D = kAv^2\rho$$

Note that if the density of were a constant it would be appropriate to simplify this dependency as

$$F = \tilde{k}Av^2$$

but now the constant $\tilde{k}$ actually has dimensions. □

2.4.3. Procedure for Nondimensionalization. Consider the nonlinear model for a pendulum

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\sin\theta$$

Based on the terms in this model we may express the solution very generally as a relationship between these included terms, i.e.,

$$\phi(\theta, g, l, t) = 0$$

Note that the angle in this model is dimensionless but the other variables all have dimensions. We can convert this equation into a new equation where none of the terms have dimensions. This will be referred to, for obvious reasons, as a dimensional form of the model.

To accomplish this, let

$$\tau = \frac{t}{\sqrt{l/g}}.$$

The substitution of variables may be accomplished by noting that

$$\frac{d^2\theta}{dt^2} = \frac{d^2\theta}{\frac{l}{g}d\tau^2}$$

Thus, after cancelation, the dimensionless form for the nonlinear pendulum model is

$$\frac{d^2\theta}{d\tau^2} = -\sin\theta$$

Now the solution has the general form

$$f(\theta, \tau) = 0,$$

or equivalently,

$$f(\theta, \sqrt{\frac{l}{g}}t) = 0$$

This is a special case of a more general theory.

The Buckingham π-theorem. Any dimensionally homogeneous equation with physical variables $x_1, \ldots, x_m$ expressed

$$\phi(x_1, \ldots, x_m)$$

may be rewritten in terms of its associated dimensionless variables $\pi_1, \ldots, \pi_n$ as

$$f(\pi_1, \ldots, \pi_n) = 0$$

where

$$\pi_k = x_1^{a_{k1}} \ldots x_m^{a_{km}}$$

2.4.4. Modeling with Dimensional Analysis. Now we consider two examples of the application of the ideas described above concerning dimensional analysis. In each of these examples there is more than one dimensionless parameter and it is appropriate to apply the Buckingham π-theorem.

2.4.4.1. *The Pendulum.* In this example the goal is to understand how the period of a pendulum depends on the other parameters that describe the nature of the pendulum. The first task is to identify this set of parameters that act as the independent variables on which the period P depends.

Obvious candidates include From this list we are motivated to write

variable	symbol	dimensions
mass	m	M
length	l	L
gravity	g	LT^{-2}
angle	θ_0	$M^0L^0T^0$
period	P	T

TABLE 2. Parameters influencing the motion of a simple pendulum.

$$P = \phi(m, l, g, \theta_0)$$

As we shall see, attempting to establish the form of ϕ directly is unnecessarily complicated. Instead, we pursue the idea of dimensional analysis.

To begin this modeling procedure, we compute the values of a, b, c, d and e that make the quantity

$$\pi = m^a l^b g^c \theta_0^d P^e$$

a dimensionless parameter. Again, this is done by equating exponents on the fundamental dimensions

$$M^0L^0T^0 = M^aL^b(LT^{-2})^c(M^0L^0T^0)^dT^e$$

From M^0: $0 = a$.
From L^0: $0 = b + c$.
From T^0: $0 = -2c + e$.
From this we may conclude that

$$\pi = m^0 l^{-c} g^c \theta_0^d P^{2c}$$

or, after collecting terms,

$$\pi = \theta_0^d \left(\frac{gP^2}{l}\right)^c$$

where π is dimensionless for any values of d and c. Thus we have found a complete set of dimensionless parameters

$$\pi_1 = \theta_0$$

and

$$\pi_2 = \sqrt{\frac{g}{l}}P$$

Since the period P of the pendulum is based on dimensionally consistent physical laws we may apply the Buckingham π-theorem. In general,

$$f(\pi_1, \pi_2) = 0$$

which we rewrite as

$$\pi_2 = h(\pi_1)$$

which now becomes

$$P = \sqrt{\frac{l}{g}}h(\theta_0)$$

We may draw two immediate conclusions from this model.

- The period depends on the square root of the length of the pendulum.
- The period is independent of the mass

Of course we have not really shown these conclusions to be "true". But now we have something to look for that can be tested. We could test these assertions and if they contradict our model then we would conclude that we are missing an important factor that governs the period of the pendulum. Indeed, as we have neglected drag forces due to friction it seems our model will have limited validity.

The functional form of h may now be reasonable calculated as there is only one independent variable θ_0. If we select several different initial displacements $\theta_0(i)$ and measure the period for each one we have a set of domain–range values

$$h(\theta_0(i)) = P_i\sqrt{\frac{g}{l}}$$

to which a data fitting procedure may now be applied.

2.4.4.2. *The damped pendulum.* We assumed that there was no damping of this pendulum above due to air resistance. We can include a drag force F_D by augmenting the list of relevant parameters to

$$m, l, g, \theta_0, P, F_D$$

Now our dimensionless parameter takes the form

$$\pi = m^a l^b g^c \theta_0^d P^e F_D^f$$

Converting to dimensions

$$M^0L^0T^0 = M^aL^b(LT^{-2})^c(M^0L^0T^0)^dT^e(MLT^{-2})^f$$

As

$$0 = a + f$$

it is no longer possible to immediately conclude that $a = 0$. In fact, it is not. (See problems).

variable	symbol	dimensions
velocity	v	LT^{-1}
density	ρ	ML^{-3}
gravity	g	LT^{-2}
radius	l	L
viscosity	μ	$ML^{-1}T^{-1}$

TABLE 3. Parameters influencing the motion of a fluid around a submerged body.

2.4.4.3. *Fluid Flow.* Consider the parameters governing the motion of an oil past a spherical ball bearing. Let's assume they include:

The dimensionless combination has the form

$$\pi = v^a \rho^b l^c g^d \mu^e$$

Using the explicit form of the physical dimensions for each term we have

$$M^0 L^0 T^0 = (LT^{-1})^a (ML^{-3})^b (L)^c (LT^{-2})^d (ML^{-1}T^{-1})^e$$

Again, matching exponents

$$\begin{aligned} M: &\quad 0 = b + e \\ L: &\quad 0 = a - 3b + c + d - e \\ T: &\quad 0 = -a - 2d - e \end{aligned}$$

Sinc there are three equations and five unknowns the system is said to be undetermined. Given these numbers, we anticipate that there we can solve for three variable in terms of the other two. Of course, we can solve in terms of *any* of the two variables. For example,

$$\begin{aligned} a &= -2d - e \\ b &= -e \\ c &= d - e \end{aligned}$$

Plugging these constraints into our expression for π gives

$$\pi = \left(\frac{v^2}{lg}\right)^{-d} \left(\frac{\rho l v}{\mu}\right)^{-e}$$

Thus, our two dimensionless parameters are the *Froude number*

$$\pi_1 = \frac{v^2}{lg}$$

and the *Reynolds number*

$$\pi_2 = \frac{v\rho l}{\mu}$$

For further discussion see Giordano, Wells and Wilde, UMAP module 526.

Problems

PROBLEM 2.1. By drawing a new graph, show the effect of the size of the island on the

- extinction curve
- migration curve

Now predict how island size impacts the number of species on the island. Does this seem reasonable?

PROBLEM 2.2. Give an example of a commodity that does not obey the

- law of supply
- law of demand

and justify your claim.

PROBLEM 2.3. Translate into words the qualitative interpretation of the slope of the supply and demand curves. In particular, what is the meaning of a

- flat supply curve?
- steep supply curve?
- steep demand curve?

PROBLEM 2.4. Consider the table of market adjustments below. Assuming the first point is on the demand curve, compute the equations of both the demand and supply curve. Using these equations, find the missing values A, B, C, D. What is the equilibrium point? Do you think the market will adjust to it?

quantity	price
3	0.7
0.14	0.7
0.14	0.986
0.1972	0.986
$A =?$	$B =?$
$C =?$	$D =?$

PROBLEM 2.5. Using the cobweb plot show an example of a market adjustment that oscillates wildly out of control. Can you describe a qualitative feature of the supply and demand curves that will ensure convergence to an equilibrium?

PROBLEM 2.6. Consider the effect of a price increase on airplane fuel (kerosene) on the airline industry. What effect does this have on the supply curve? Will the airline industry be able to pass this cost onto the flying public? How does your answer differ if the demand curve is flat versus steep?

PROBLEM 2.7. Prove properties 2.1 and 2.2.

PROBLEM 2.8. Is the temperature measured in degrees Fahrenheit proportional to the temperature measured in degrees centigrade?

PROBLEM 2.9. Consider the Example 2.6 again. Demonstrate the proportionalities stated. For the case of the breaking distance equate the work done by the breaks to the dissipated kinetic energy of the car.

PROBLEM 2.10. Items at the grocery store typically come in various sizes and the cost per unit is generally smaller for larger items. Model the cost per unit weight by considering the superposition of proportions due to the costs of

- production
- packaging
- shipping

the product. What predictions can you make from this model. This problem was adapted from Bender [**?**].

PROBLEM 2.11. Go to your nearest supermarket and collect data on the cost of items as a function of size. Do these data behave in a fashion predicted by your model in the previous problem?

PROBLEM 2.12. In this problem take the diagonal of a rectangle as it's length scale l. Show by direct calculation that this can be used to measure the area, i.e.,

$$A = \alpha l^2$$

Determine the constant of proportionality α in terms of the shape factor of the rectangle.

PROBLEM 2.13. Consider a radiator designed as a spherical shell. If the characteristic length of the shell doubles (assume the larger radiator is geometrically similar to the smaller radiator) what is the effect on the amount of heat loss? What if the design of the radiator is a parallelpiped instead?

PROBLEM 2.14. How does the argument in Example 2.10 change if the falling object is not spherical but some other irregular shape?

PROBLEM 2.15. Extend the definition of geometric similarity for

- parallelpipeds
- irregularly shaped objects

Can you propose a computer algorithm for testing whether two objects are geomtrically similar?

PROBLEM 2.16. Consider the force on a pendulum due to air friction modeled by

$$F_D = \kappa v^2$$

Determine the units of κ.

PROBLEM 2.17. Newton's law of gravitation states that

$$F = \frac{G m_1 m_2}{r^2}$$

where F is the force between two objects of masses m_1, m_2 and r is the distance between them.

a) What is the physical dimension of G?
b) Compute two dimensionless products π_1 and π_2 and show explicitly that they satisfy the Buckingham π-theorem.

PROBLEM 2.18. This problem concerns the pendulum example described in subsection 2.4.4.1. Repeat the analysis to determine the dimensionless parameter(s) but now omit the gravity term g. Discuss.

PROBLEM 2.19. This problem concerns the pendulum example described in subsection 2.4.4.1. Repeat the analysis for determining all the dimensionless parameters but now include a parameter κ associated with the drag force of the form $F_D = \kappa v$. Hint: first compute the dimensions of κ.

PROBLEM 2.20. Convert the equation governing the distance traveled by a projectile

$$\frac{d^2x}{dt^2} = \frac{-gR^2}{(x+R)^2}$$

with the initial conditions

$$x(0) = 0 \qquad \frac{dx}{dt}(0) = V$$

to non-dimensional form using the transformation

$$y = \frac{x}{R} \qquad \tau = \frac{Vt}{R}$$

Verify the resulting equation is dimensionless. See [**?**] for further details.

PROBLEM 2.21. Reconsider the example in subsection 2.4.4.3. Instead of solving for a, b, c in terms of d, e solve for c, d, e in terms of a and b. Show that now

$$\pi^{'}{}_1 = \frac{v}{\sqrt{lg}}$$

and

$$\pi^{'}{}_2 = \frac{\rho l^{3/2} g^{1/2}}{\mu}$$

Show also that both $\pi_1^{'}$ and $\pi_2^{'}$ can be written in terms of π_1 and π_2.

PROBLEM 2.22. Consider an object with surface area A traveling with a velocity v through a medium with kinematic viscosity μ and density ρ.

a) Assuming the effect of μ is small compute the drag force due to the density F_ρ.
b) Assuming the effect of ρ is small compute the drag force due to the kinematic viscosity F_μ.
c) Compute the dimensionless ratio of these drag forces and discuss what predictions you can make.

PROBLEM 2.23. How does the required power P of a helicopter engine depend on the length of the rotors l? The rotors are pushing air so presumable the density ρ as well as the weight of the helicopter $w = mg$ are variables that affect the power requirement. Draw a sketch of your result plotting P versus l. See [**?**] for more discussion of this problem.

CHAPTER 3

Data Fitting

In this chapter the model building is *empirical* in nature, i.e., the relationship between the dependent and independent variables is found by direct examination of data related to the process.

Data fitting problems have several common elements. The *model* has the general form

$$y = f(x_1, \ldots, x_N; w_1, \ldots, w_M)$$

and the *parameters* w_i are determined empirically from the *observations*

$$\{(x^{(i)}, y^{(i)})\}_{i=1}^{P}$$

by requiring f to be such that

$$y^{(i)} = f(x_1^{(i)}, \ldots, x_N^{(i)}; w)$$

or at least that the *error function* $E(w)$ defined by

$$S(w) = \sum_i (y^{(i)} - f(x_1^{(i)}, \ldots, x_N^{(i)}; w_1, \ldots, w_M))^2$$

be *small*. This error function seeks to minimize the sum of squares of the *residuals*. As we shall see, other error functions are possible but least squares is certainly the most widely used and we will focus on this approach at the outset.

EXAMPLE 3.1. A *Radial basis function* model has the form

$$f(x; w, c) = w_0 + \sum_{k=1}^{N_c} w_k \phi(\|x - c_k\|)$$

where the w_k are the weights and the c_k are the centers of the basis functions. An example of a radial basis function is

$$\phi(r) = \exp(-r^2)$$

The norm $\|\cdot\|$ is generally taken to be the Euclidean distance.

3.1. Linear Least Squares

In this section we begin by revisiting an example from the previous chapter followed by a general formulation of linear least squares and some simple extentions to exponential fits.

3.1.1. The Mammalian Heart Revisited. Recall from Example 2.11 in Section 2.3.2.1 that a sequence of proportionalities produced the model

$$r = kw^{-1/3}$$

where w is the body weight of a mammal and r is its heart rate. The data on Figure 8 corresponds to observations

$$\{(w_i^{-1/3}, r_i)\}$$

collected for various measured rates and weights. The residual error for the ith measurement is

$$\epsilon_i = r_i - kw_i^{-1/3}$$

and the total squared error is

$$E = \sum_{i=1}^{P} \epsilon_i^2$$

We rewrite this error as a function of the unknown slope parameter k as

$$E(k) = \sum_{i=1}^{P} (r_i - kw_i^{-1/3})^2$$

To minimize E as a function of k we compute the derivative of E w.r.t. k, i.e.,

$$\frac{dE}{dk} = \sum_{i=1}^{P} 2(r_i - kw_i^{-1/3}) \cdot (-w_i^{-1/3}) = 0$$

From which it follows that

$$k = \frac{\sum_{i=1}^{P} r_i w_i^{-1/3}}{\sum_{i=1}^{P} w_i^{-2/3}}$$

Thus we can obtain an estimate for the slope of the line empirically from the data.

3.1.2. General Formulation. In this section we focus our attention to one of the most widely used models

$$f(x; m, b) = mx + b$$

To clean-up the notation we now use subsripts to label points for domain data that is one dimensional; we used superscripts in the previous section when the dimension of the domain could exceed one. For a set of observations $\{(x_i, y_i)\}$, $i = 1, \ldots, P$, the total squared error is given by

$$E(m, b) = \sum_{i=1}^{P} (y_i - mx_i - b)^2 \tag{3.1}$$

Now because there are two parameters that determine the error function the necessary condition for a minimum is now

$$\frac{\partial E}{\partial m} = 0$$
$$\frac{\partial E}{\partial b} = 0$$

Solving the above equations gives the slope of the line as

$$m = \frac{(\sum y_i)(\sum x_i) - P \sum y_i x_i}{(\sum x_i)^2 - P \sum x_i^2}$$

FIGURE 1. Linear least squares. The line $y = mx+b$ is determined such that the residuals ϵ_i^2 are minimized.

and its intercept to be

$$b = \frac{-(\sum y_i)(\sum x_i^2) + (\sum x_i)(\sum y_i x_i)}{(\sum x_i)^2 - P\sum x_i^2}$$

3.1.2.1. *Interpolation Condition.* In this section we present another route to the equations for m and b produced in the previous section. Again, the input data is taken as $\{x_i\}$, the output data is $\{y_i\}$ and the model equation is $y = mx + b$. Applying the *interpolation condition* for each observation we have

$$\begin{aligned} y_1 &= mx_1 + b \\ y_2 &= mx_2 + b \\ y_3 &= mx_3 + b \\ &\vdots \\ y_P &= mx_P + b \end{aligned}$$

In terms of matrices

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_P \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_P \end{pmatrix}$$

In terms of matrices we can summarize the above as

$$Xb = y$$

We can reveal the relationship between the previous approach using calculus and this approach with the interpolation condition by hitting both sides of the above matrix equation with the transpose X^T

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_P \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_P \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_P \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_P \end{pmatrix}$$

In terms of the matrices,

$$X^T X b = X^T y$$

Multiplying out produces the equations that are seen to be the same as those in the above section, i.e.,

$$\begin{pmatrix} P & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

In linear algebra these equations are referred to as the *normal* equations.

There are many algorithms in the field of numerical linear algebra developed precisely for solving the problem

$$Xb = y$$

We will consider these more in the problems.

3.1.3. Exponential Fits. We have already seen models of the form

$$y = kx^n$$

where n was given. How about if n is unknown? Can it be determined empirically from the data? Note now that the computation of the derivative of the error function w.r.t. n is now quite complicated. This problem is resolved by converting it to a linear least squares problem now in terms of logarithms. Specifically,

$$\begin{aligned} \ln y &= \ln(kx^n) \\ &= \ln k + \ln x^n \\ &= \ln k + n \ln x \end{aligned}$$

This is now seen to be a linear least squares problem

$$y' = nx' + k'$$

where we have made the substitutions $y' = \ln y, k' = \ln k$ and $x' = \ln x$. Now one can apply the standard least squares solution to determine n and k'. The value of k can be found as well by

$$k = \exp(k')$$

3.1.4. Fitting Data with Polynomials. In the previous section we consider the basic linear model $f(x; c_0, c_1) = c_0 + c_1 x$. The simplest extension to this is the second order polynomial

$$f(x; c_0, c_1, c_2) = c_0 + c_1 x + c_2 x^2$$

Note first that adding the term $c_2 x^2$ will change the least square fit values of the coefficients c_0, c_1 obtain from the linear model and hence all the coefficients c_0, c_1 and c_2 must be computed. The least squares procedure follows along lines similar

to the previous section. We assume a set of observations $\{(x_i, y_i)\}$ and define the sum of the squares of the model residuals to be the error function, i.e.,

$$E(c_0, c_1, c_2) = \sum_{i=1}^{P} (y_i - c_0 - c_1 x_i - c_2 x_i^2)^2 \tag{3.2}$$

Now requiring

$$\frac{\partial E}{\partial c_0} = 0$$
$$\frac{\partial E}{\partial c_1} = 0$$
$$\frac{\partial E}{\partial c_2} = 0$$

The resulting necessary conditions, writtin in terms of matrices, are then

$$\begin{pmatrix} P & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum y_i x_i^2 \end{pmatrix} \tag{3.3}$$

As for the linear model, it is possible to solve for the parameters c_0, c_1 and c_2 analytically, i.e., in closed form. It is less cumbersome to write Equation (3.3) as

$$Xc = z$$

where c is the column vector made up of the elements (c_0, c_1, c_2) and z is the column vector comprised of the elements $(\sum y_i, \sum x_i y_i, \sum y_i x_i^2)$ and X is the 3×3 matrix on the left of Equation (3.3). Now a computer package can be used to easily solve the resulting matrix equation.

3.1.4.1. *Lagrange Polynomials.* Consider the first degree polynomial defined by

$$P_1(x) = \frac{x - x_2}{x_1 - x_2} y_1 + \frac{x - x_1}{x_2 - x_1} y_2$$

By construction, it is clear that

$$P_1(x_1) = y_1$$

and

$$P_1(x_2) = y_2$$

So we have found the unique line passing through the points $\{(x_1, y_1), (x_2, y_2)\}$.

This procedure may be continued analogously for more points. A second degree polynomial passing through three prescribed points is given by

$$P_2(x) = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} y_1 + \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} y_3$$

By construction, it again may be verified that

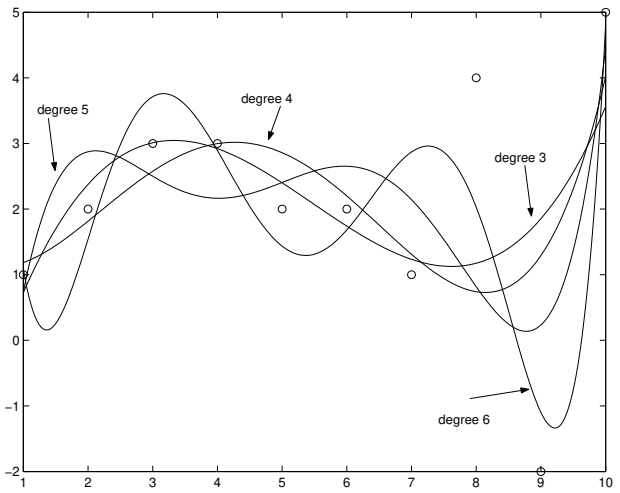
$$P_2(x_1) = y_1,$$
$$P_2(x_2) = y_2,$$

and

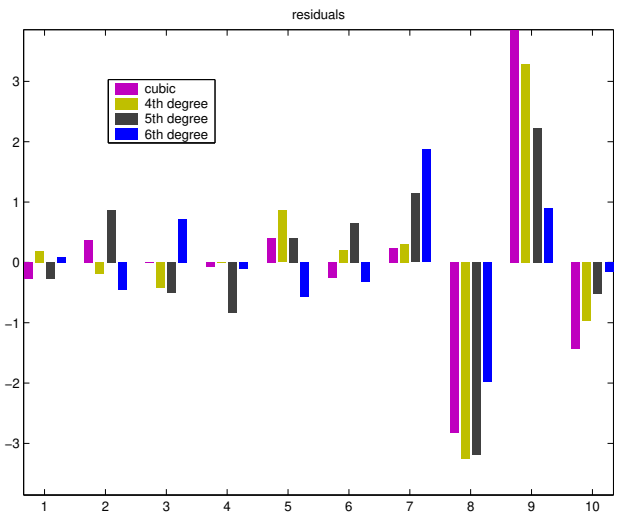
$$P_2(x_3) = y_3.$$

A pattern has emerged that makes it apparent that this simple procedure may be employed to fit a polynomial of degree n through a set of $n + 1$ points.

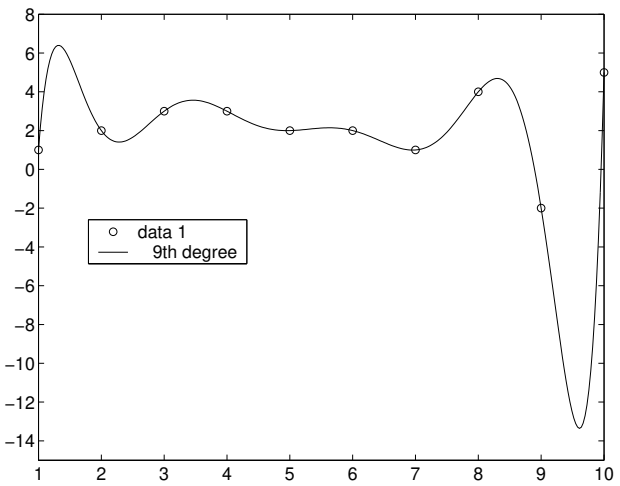
3.1.5. Interpolation versus Least Squares.



(a) Degree 3, 4, 5, 6 polynomials fit to a data set.



(b) Residual plot of degree 3, 4, 5, 6 polynomial fits.



(c) Degree 9 polynomial fit.

FIGURE 2. Comparitive polynomial approximation to a 10 point data set.

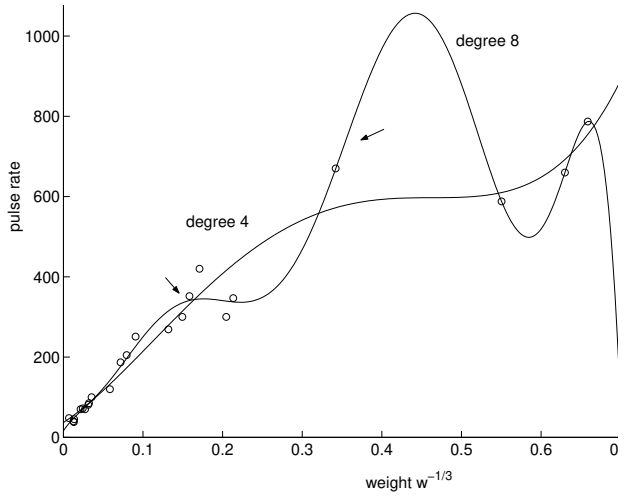


FIGURE 3. Degree 4 and degree 8 polynomials applied to the data $(w_i^{-1/3}, r_i)$.

3.1.5.1. *Dangers of Higher Order Polynomials.* It is clear from the preceding section that we can always find a polynomial of degree n to exactly fit (i.e., satisfy the interpolation condition) $n+1$ points. So why not simply use high order polynomials? To answer this question consider the model

$$f(x) = c_{20}x^{20} + c_{19}x_{19} + \cdots + c_1x_1 + c_0$$

If one of the coefficients is perturbed by a small value ϵ the resulting model may predict wildly different results. For example, let

$$g(x) = (c_{20} + \epsilon)x^{20} + c_{19}x_{19} + \cdots + c_1x_1 + c_0$$

Then

$$g(x) = \epsilon x^{20} + f(x)$$

So, even if ϵ is small the difference between $f(x)$ and $g(x)$ is potentially very large. This is a manifestation of ill-conditioning of high degree polynomials, i.e., small changes in parameters may result in large changes in the function.

See Figure 3 for an example of the oscillations that appear with higher order polynomials.

3.2. Splines

One obvious procedure for reducing the need for higher order polynomials is to restrict each polynomial for the description of limited contiguous data subsets. This is the central idea behind *splines*. A spline is a piecewise defined function

$$S(x) = \begin{cases} S_1(x) & \text{if} \quad x_1 \le x < x_2, \\ S_2(x) & \text{if} \quad x_2 \le x < x_3, \\ \vdots & \vdots \\ S_j(x) & \text{if} \quad x_j \le x < x_{j+1}, \\ \vdots & \vdots \\ S_n(x) & \text{if} \quad x_n \le x < x_{n+1} \end{cases} \tag{3.4}$$

that satisfies the interpolation conditions for all the data points

$$S(x_j) = y_j = S_j(x_j) \tag{3.5}$$

Furthermore, the piecewise defined models may be joined by enforcing *auxilary conditions* to be described. We begin with the simplest case.

3.2.1. Linear Splines. For linear splines the piecewise function takes the form

$$S_j(x) = c_0^j + c_1^j x$$

and this is valid over the interval $x_j \leq x < x_{j+1}$. (See Figure 4.)

For $x_1 \leq x < x_2$ the function $S_1(x)$ is the line passing through the points (x_1, y_1) and (x_2, y_2). The interpolation condition $S_1(x_1) = y_1$ requires

$$c_0^1 + c_1^1 x_1 = y_1.$$

The matching condition

$$S_1(x_2) = S_2(x_2) = y_2$$

requires

$$c_0^1 + c_1^1 x_2 = y_2$$

This system of two equations in two unknowns has the solutions

$$c_0^1 = \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1}$$

and

$$c_1^1 = \frac{-y_1 + y_2}{x_2 - x_1}$$

Similarly for $S_2(x)$

$$S_2(x) = c_0^2 + c_1^2 x$$

The parameters c_0^2 and c_1^2 are then determined by conditions

$$\begin{cases} S_2(x_2) = y_2 \\ S_2(x_3) = S_3(y_3) = y_3 \end{cases} \tag{3.6}$$

which can be found to be

$$c_0^2 = \frac{x_3 y_2 - x_2 y_3}{x_3 - x_2}$$

and

$$c_1^2 = \frac{-y_2 + y_3}{x_3 - x_2}$$

In general, it follows

$$\begin{cases} S_j(x_j) = y_j \\ S_j(x_{j+1}) = S_{j+1}(x_{j+1}) = y_{j+1} \end{cases} \tag{3.7}$$

which can be found to be

$$c_0^j = \frac{x_{j+1} y_j - x_j y_{j+1}}{x_{j+1} - x_j}$$

and

$$c_1^j = \frac{-y_j + y_{j+1}}{x_{j+1} - x_j}$$

FIGURE 4. Linear splines.

3.2.2. Cubic Splines. Notice that with linear splines that the function matches at interpolation points, i.e.,

$$S_j(x_{j+1}) = S_{j+1}(x_{j+1})$$

but that in general the derivative does not match, i.e.,

$$S_j^{'}(x_{j+1}) = S_{j+1}^{'}(x_{j+1})$$

This potential problem may be overcome by employing cubic splines (quadratic splines do not provide enough parameters).

For cubic splines the piecewise function takes the form

$$S_j(x) = c_0^j + c_1^j x + c_2^j x^2 + c_3^j x^3$$

and this is valid over the interval $x_j \leq x < x_{j+1}$. To simplify notation we will only consider two segments of $S(x)$

$$S(x) = \begin{cases} S_1(x) & \text{if} \quad x_1 \leq x < x_2, \\ S_2(x) & \text{if} \quad x_2 \leq x < x_3 \end{cases} \tag{3.8}$$

Now

$$S_1(x) = c_0^1 + c_1^1 x + c_2^1 x^2 + c_3^1 x^3$$

and

$$S_2(x) = c_0^2 + c_1^2 x + c_2^2 x^2 + c_3^2 x^3$$

And we observe that there are 8 parameters.

Three equations are obtained by employing the interpolation conditions

$$\begin{cases} S_1(x_1) = y_1 \\ S_2(x_2) = y_2 \\ S_2(x_3) = y_3 \end{cases} \tag{3.9}$$

The matching condition for the function value is

$$S_1(x_2) = S_2(x_2) = y_2$$

Given there are 8 parameters and only 4 equations specified we require 4 more relations.

We will require that first and second derivatives match at the interior points,

$$\begin{cases} S_1^{'}(x_2) = S_2^{'}(x_2) \\ S''_1(x_2) = S''_2(x_2) \end{cases} \tag{3.10}$$

The additional two parameters may be obtained by applying conditions on the derivatives at the endpoints x_1 and x_3. One possibility is to require

$$\begin{cases} S_1^{''}(x_1) = 0 \\ S_2^{''}(x_3) = 0 \end{cases} \tag{3.11}$$

Since no data matching is involved these are called *natural splines*.

Problems

PROBLEM 3.1. Consider the model

$$f(x; c_0, c_1) = c_0 x^{-3/2} + c_1 x^{5/2}$$

Use the least squares approach to determine equations for c_0 and c_1 in terms of the available data.

PROBLEM 3.2. Consider the model

$$f(w; a, n) = a w^n$$

for the raw mammalian heart data provided in Table 1. Find a and n using the least squares approach to fit the data. You may also use the least squares code provided in the Appendix (either from Section A.2 or Section A.3) but you will first need to take appropriate transforms of the data.

PROBLEM 3.3. Write MATLAB code to fit a second order polynomial

$$f(x; c_0, c_1, c_2) = c_0 + c_1 x + c_2 x^2$$

to the *linearized* mammalian heart data consisting of components $\{(w_i^{-1/3}, r_i)\}$. Compute the total squared error $E(c_0, c_1, c_2)$ given by Equation (3.2) and compare with the error term $E(m, b)$ given by Equation (3.1). You may modify the code provided in the appendices for this problem.

PROBLEM 3.4. Derive the matrix Equation (3.3).

PROBLEM 3.5. Rederive the matrix Equation (3.3) using the interpolation approach of Subsection 3.1.2.1.

PROBLEM 3.6. Derive the 4×4 system of equations required to fit the model

$$f(x; c_0, c_1, c_2, c_3) = c_0 + c_1 x + c_2 x^2 + c_3 x^3$$

Put this system into matrix form and compare with the result in Equation (3.3). Can you extend this pattern to write down the matrix form of the least squares equations for the model

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_9 x^9$$

No exact derivation is required for this last part.

PROBLEM 3.7. Reread Section 3.1.4.1 and propose a formula for a polynomial of degree 3 $P_3(x)$ such that the interpolation conditions $P_3(x_1) = y_1$, $P_3(x_2) = y_2$, $P_3(x_3) = y_3$, and $P_3(x_4) = y_4$ are satisfied.

PROBLEM 3.8. Find the linear spline through points $\{(0, 0), (2, 1), (3, -2), (5, 2)\}$.

PROBLEM 3.9. Apply MATLAB's cubic spline routine to the Fort Collin's daily temperature data provided on odd days in September 2002; see Table 1. Use this model to predict the temperature on the even days. Compare your predictions with the actual temperature values provided in Table 2. Plot your results.

Day	5pm temperature
1	87.0
3	83.3
5	89.7
7	82.3
9	65.6
11	59.8
13	65.5
15	77.1
17	69.1
19	63.8
21	51.1

TABLE 1. Fort Collins' temperatures on odd days in September, 2002. All temperatures recorded at 5pm. Use this data to build spline model.

Day	5pm temperature
2	80.3
4	87.1
6	90.1
8	79.2
10	64.4
12	60.2
14	71.2
16	73.9
18	58.4
20	73.5

TABLE 2. Fort Collins' temperatures on even days in September, 2002. All temperatures recorded at 5pm. Use this data to test spline model.

CHAPTER 4

Difference Equations

One of the most exciting areas of modeling concerns predicting temporal evolution. The main question that is posed in this setting is how do quantities of interest change over time? This type of problem is everywhere to be found, for example in areas as diverse as science, engineering and finance.

In this chapter we make the mathematical assumption that the time variable n is incremented discretely and corresponds to the integers $\{0, 1, 2, 3, 4, \dots\}$. The value of the variable of interest is then a sequence $\{a_1, a_2, a_3, a_4, \dots\}$. Now the problem of modeling is to determine an equation of the form

$$a_{n+1} = a_n + \Delta a_n$$

and this is done by estimating how the variable a_n changes as n is incremented from time n to time $n + 1$.

We proceed by developing this topic along the following three complementary lines:

- numerical solutions
- analytical solutions
- qualitative behavior
- modeling techniques

As the terminolgy suggests, numerical approaches to difference equations will involve direct computation of these sequences via computer. In contrast, analytical solution methods seek closed form solutions; these are available only in limited circumstances. The topic of modeling will treat empirical and qualitiative approaches for constructing difference equations. We will consider the development of the models based on the qualitative approaches presented in Chapter 2 as well as the more quantitative data fitting approaches of Chapter 3.

EXAMPLE 4.1. Let's motivate the material in this chapter by first considering a simple difference equation

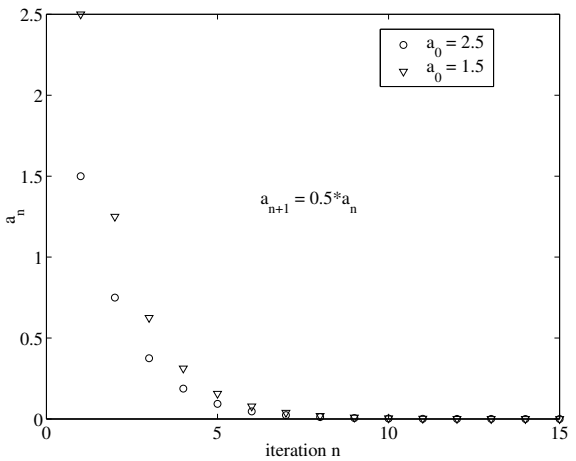
$$a_{n+1} = \alpha a_n + \beta$$

We see the numerically simulated solutions for various values of α and β in Figure 1. In Figure 1 a) we see that the solution decays to zero while in Figure 1 b) the solution tends to the value 2. In Figure Figure 1 c) the solution goes off to $+\infty$. Thus, a noticable feature for all of these solutions is the long term bahavior. Qualitatively we say the solution either blows up or has an *equilibrium solution*. An equilibrium solution $\overline{a}$ is defined by the property that

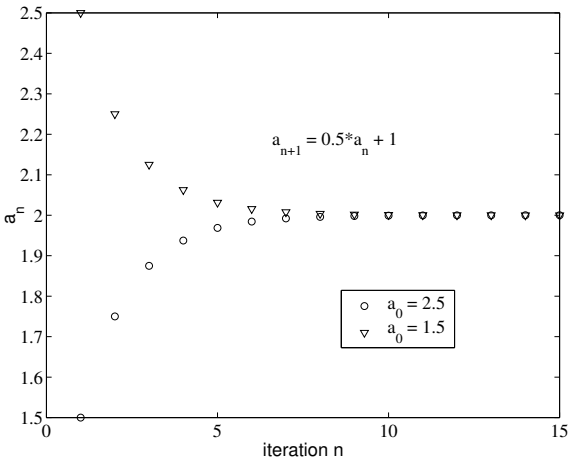
$$\overline{a} = a_{n+1} = a_n \quad \forall n$$

Both the equations

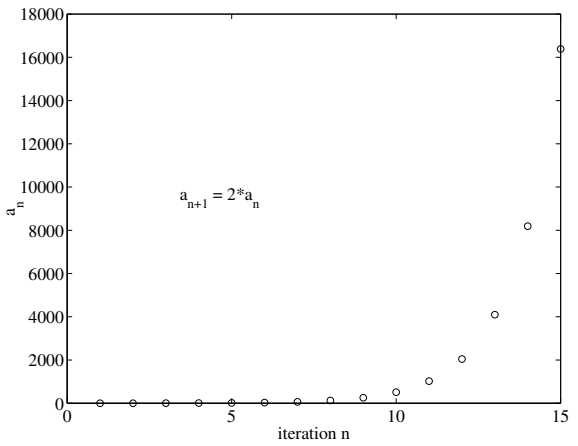
$$a_{n+1} = \frac{a_n}{2}$$



(a)



(b)



(c)

FIGURE 1. A comparison of the numerical solutions for some simple linear difference equations.

and

$$a_{n+1} = 2a_n$$

are readily seen to have $\overline{a} = 0$ as an equilibrium solution but the numerical solution for these problems is qualitatively very different. In the first equation the equilibrium is stable and in the second it is unstable. The equation

$$a_{n+1} = \frac{a_n}{2} + 1$$

has the shifted equilibrium $\overline{a} = 2$ and it appears to be stable. We shall see that the analytical, or closed form solution, for this problem is given by

$$a_n = c\alpha^n + \frac{\beta}{1-\alpha}$$

where c is an arbitrary constant. Although the stability of a solution may be easily determined from analytical solution, typically this is not a viable approach to determining stability as analytical solutions are generally only available for linear problems.

4.1. Analytical Solutions

In this section we will develop methods for constructing closed form *analytical* solutions for linear difference equations of the form

$$x_{n+1} = \alpha_1 x_n + \alpha_2 x_{n-1} + \cdots + \alpha_k x_{n-k} + \beta_n$$

The equation is said to be *nonhomogeneous* due to the presence of the β_n term that is independent of x_n. If $\beta_n = 0$ for all n then the equation is said to be *homogeneous*.

The order of a difference equation is the difference between the largest time index $n+1$ and the smallest time index $n-k$. So the equation above is a $k+1$th order difference equation.

A difference equation is said to be *linear* if all the terms x_n occur linearly, i.e., there are no nonlinear functions of these quantities. Note that, by definition, an equation that does not satisfy the property of being linear is referred to as *nonlinear*.

Example 4.2.

$$x_{n+2} = x_{n+1} + x_n$$

is a second order, linear, homogeneous difference equation.

Example 4.3.

$$x_{n+1} = sin(x_n) + 1$$

is a first order, nonlinear, nonhomogeneous difference equation.

4.1.1. First Order Difference Equations. Possibly the simplest nontrivial linear difference equation has the form

$$x_{n+1} = \alpha x_n \tag{4.1}$$

Implementing the iteration

$$\begin{aligned} x_1 &= \alpha x_0 \\ x_2 &= \alpha x_1 = \alpha^2 x_0 \\ x_3 &= \alpha x_2 = \alpha^3 x_0 \\ &\vdots \\ x_n &= \alpha^n x_0 \end{aligned}$$

Thus we see a general pattern that affords the writing of the *solution* of the difference equation in general form.

Note that any expression of the form $c\alpha^n$ is a solution to the difference equation $x_{n+1} = \alpha x_n$. This can be verified independently of the above solution procedure. All that we need to do to verify that we have a solution to a difference equation is to plug it in and see if the resulting equation is true. So, for example, let's test $h_n = c\alpha^n$. The left-hand side of Equation (4.1) is then $h_{n+1} = c\alpha^{n+1}$. The right hand side is $\alpha c\alpha^n$ which is clearly an identical expression to the left-hand side.

Note that c is an arbitrary constant in the sense that the above solution $h_n = c\alpha^n$ did not require the value of c to be specified. Since there is an arbitrary parameter that may be used to characterize the set of solutions as a family of solutions, we refer to $h_n = c\alpha^n$ as the *general solution*. It is crucial that a general solution have enough parameters to specify which member of the family corresponds to a given initial condition. In this case it is simple to verify that one parameter is enough. To see this let $n = 0$, then it follows $h_0 = c$. Thus, to select one member of the family of solutions associated with a given initial condition we take $c = h_0$. The solution

$$h_n = h_0\alpha^n$$

is referred to as a *particular solution* as it contains no free parameters.

EXAMPLE 4.4. Let

$$x_{n+1} = 3x_n$$

This equation has the general solution

$$x_n = c3^n$$

EXAMPLE 4.5. Let

$$x_{n+1} = 3x_n + 2n + 1$$

This equation has a particular solution of

$$p_n = -n - 1$$

and a general solution

$$x_n = c3^n - n - 1$$

It is easy to verify each of these claims.

To test that an expression is a particular solution plug it in the equation. First, the left-hand side

$$p_{n+1} = -(n+1) - 1 = -n - 2$$

The right hand side of the original difference equation is then

$$3p_n + 2n + 1 = 3(-n - 1) + 2n + 1 = -n - 2$$

Since the left and right hand sides are the same we have verified that $p_n = -n - 1$ is a particular solution. Note that it is not a general solution.

Now we verify that the expression $x_n = c3^n - n - 1$ is a general solution to this problem, i.e., the expression is a solution with the added feature that the free parameter x may be found for any initial value x_0. First compute the left-hand side

$$x_{n+1} = c3^{n+1} - (n+1) - 1 = c3^{n+1} - n - 2$$

The right-hand side is

$$3x_n + 2n + 1 = 3(c3^n - n - 1) + 2n + 1 = c3^{n+1} - n - 2$$

so they are equal and we have verified the solution. Now, to show that this is a general solution we apply an arbitrary initial condition $x_0 = a$. Evaluating the general solution to the difference equation at $x_0 = a$ gives

$$a = c3^0 - 0 - 1$$

or $c = a + 1$. Thus, for any initial condition $x_0 = a$ we have the solution

$$x_n = (a+1)3^n - n - 1$$

and hence we have verified that we started with the general solution.

Consider the modest embellishment on the simple difference equation considered to this point in the form

$$x_{n+1} = \alpha x_n + \beta_n \tag{4.2}$$

where again α is a constant and β_n is independent of x_n. Due to this additional term Equation 4.2 is referred to as a *nonhomogeneous* equation and the Equation (4.1) is its associated *homogeneous* equation. Note one always obtains an associated homogeneous equation by setting the nonhomogeneous term to zero.

The distinction between homogeneous and nonhomogeneous equations is central to the solution procedure for equations of this type. This solution procedure will also rely on the fact the Equations (4.1) and (4.2) are linear.

Now let's consider the general situation of determining an analytical solution to a linear nonhomogeneous equation

$$x_{n+1} = \alpha x_n + \beta_n$$

Let h_n be the general solution to the associated homogeous equation

$$x_{n+1} = \alpha x_n$$

Again, this means that h_n is effectively a family of solutions in that it has a free parameter such that any solution of the homogeneous equation can be expressed by appropriately selecting this parameter.

Also, let p_n be a particular solution to the nonhomogeneous problem. Now we are thinking of a solution

$$p_{n+1} = \alpha p_n + \beta_n$$

where p_n containes no undetermined coefficients.

THEOREM 4.6. *The general solution of the nonhomogeneous equation*

$$x_{n+1} = \alpha x_n + \beta_n$$

is given by

$$x_n = h_n + p_n$$

where h_n is the general solution to the associated homogeneous problem $x_{n+1} = \alpha x_n$ and p_n is a particular solution to the non-homogeneous problem.

First we show that the expression

$$x_n = h_n + p_n$$

is a solution to the nonhomogeneous problem. To prove this is suffices to plug this expression into the equation

$$x_{n+1} = \alpha x_n + \beta$$

This gives

$$\begin{aligned} h_{n+1} + p_{n+1} &= \alpha h_n + (\alpha p_n + \beta_n) \\ &= \alpha(h_n + p_n) + \beta_n \\ &= \alpha x_n + \beta_n \end{aligned}$$

Now we must show that every solution to the non-homogeneous problem may be writtin this way. If this is true, then the general solution to the homogeneous problem could be written

$$h_n = x_n - p_n$$

The proof proceeds by checking that the right hand side is indeed a solution to the nonhomogeneous problem. Thus we test the truth of the statement

$$x_{n+1} - p_{n+1} = \alpha(x_n - p_n)$$

Rearranging produces

$$(x_{n+1} - \alpha x_n) - (p_{n+1} - \alpha p_n) = 0$$

but this is just

$$\beta_n - \beta_n = 0$$

EXAMPLE 4.7. Compute the analytical form for the general solution to the problem

$$x_{n+1} = \alpha x_n + \beta$$

where α and β are constants (independent of n) and $\alpha \neq 1$. Compute the particular solution associated with the initial condition $x_0 = 1$.

- Step 1. Determine the general solution to the homogeneous problem

  $$x_{n+1} = \alpha x_n$$

  From our previous work we know that this is

  $$h_n = c\alpha^n$$

- Step 2. Find a particular solution to the nonhomogeneous equation. Let $p_n = A$ and see if we can determine the coefficient A. Plugging into the nonhomogeneous problem

  $$A = \alpha A + \beta$$

  Hence,

  $$A = \frac{\beta}{1 - \alpha}$$

- Step 3. The general solution may be written as $x_n = h_n + p_n$, or

  $$x_n = c\alpha^n + \frac{\beta}{1 - \alpha}$$

- Step 4. The application of the initial condition may now proceed.

$$x_0 = 1 = c + \frac{\beta}{1-\alpha}$$

So the solution to the initial value problem is then

$$x_n = (1 - \frac{\beta}{1-\alpha})\alpha^n + \frac{\beta}{1-\alpha}$$

EXAMPLE 4.8. Compute the analytical form for the general solution to the problem

$$x_{n+1} = 3x_n + 2n + 1$$

where α and β are constants (independent of n) and $\alpha \neq 1$. Compute the particular solution associated with the initial condition $x_0 = 1$.

- Step 1. Determine the general solution to the homogeneous problem

$$x_{n+1} = \alpha x_n$$

This is

$$h_n = c3^n$$

- Step 2. Find a particular solution to the nonhomogeneous equation. Let $p_n = An + B$ and see if we can determine the coefficients A, B. Plugging into the nonhomogeneous problem

$$A(n+1) + B = 3(An + B) + 2n + 1$$

Hence,

$$(-2n + 1)A - 2B = 2n + 1$$

Letting $n = 0$ it follows $A - 2B = 1$ and letting $n = 1$ it follows $-A - 2B = 3$ from which we conclude that $A = B = -1$. So $p_n = -n - 1$. (Check this.)

- Step 3. The general solution may be written as $x_n = h_n + p_n$, or

$$x_n = c3^n - n - 1$$

- Step 4. Note no initial condition was specified in this problem so we are done.

EXAMPLE 4.9. You purchase a home for \$200,000 at 0.5% monthly interest. Assuming the term of the load is 30 years, compute your monthly payment. How much will you owe after 10 years?

We begin by defining the amount owed at the end of month end as x_n. The change in the amount owed each month increases due to interest and decreases as a consequence of the monthly payment p. As a difference equation we have

$$x_{n+1} - x_n = 0.005x_n - p$$

This has the general solution

$$x_n = 1.005^n c + 200p$$

where c is a constant. Applying the initial condition

$$x_0 = 200000$$

and final condition

$$x_{360} = 0$$

(after 30 years, or 360 months, you should owe exactly nothing) it follows that

$$c = -39820.21$$

and the monthly payment is

$$p = \$1199.10$$

giving the general solution

$$x_n = 1.005^n(-39820.21) + 239820.21$$

so at the end of ten years (or 120 months)

$$x_{120} = \$167,371.45$$

A possibly surprising amount is still owed after 1/3 the term of the loan.

EXAMPLE 4.10. A muder victim is discovered in an office building that is maintained at 68 degrees F. Given the medical examiner found the body temperature to be 88 degrees F at 8am and that one hour later the body temperature was 86 degrees F, at what time was the crime committed?

Let T_n be the temperature of the body and the change in temperature $\Delta T_n = T_{n+1} - T_n$. We propose the model (actually Newton's Law of Cooling) that the rate of change of the temperature of the body is proportional to the difference between the room temperature and the temperature of the body, i.e.,

$$\Delta T_n \propto 68 - T_n$$

In other words,

$$\Delta T_n = k(68 - T_n)$$

and since we know the body temperature is decreasing it must follow that $k > 0$. The difference equation that arises from this model is then

$$T_{n+1} = T_n + k(68 - T_n)$$

which has the general solution

$$T_n = c(1 - k)^n + 68$$

Setting $T_0 = 98.6$ (where 0 is the time the crime was committed) we see

$$T_0 = 98.6 = 68 + c$$

or $c = 30.6$ so the particular solution of interest is

$$T_n = 68 + 30.6(1 - k)^n$$

If we define n_1 as the time the body was observed initially by the medical examiner and the time one hour later as $n_1 + 1$ we have the equations

$$T_{n_1} = 88 = 68 - 30.6(1 - k)^{n_1}$$

and

$$T_{n_1+1} = 86 = 68 - 30.6(1 - k)^{n_1+1}$$

These two equations may be solved to give $k = 1/10$ and $n_1 = 4.036$. So the crime was committed just before 4am.

4.1.2. Second Order Equations. We begin by considering the second order linear homogeneous difference equation

$$x_{n+2} + \alpha x_{n+1} + \beta x_n = 0 \tag{4.3}$$

It is readily verified that this equation has solutions of the form

$$x_n = \lambda^n$$

Upon substitution into Equation (4.3) we obtain the *auxiliary* equation

$$\lambda^2 + \alpha\lambda + \beta = 0$$

This quadratic equation has solutions that break down into three cases: i) both solutions real and distinct, ii) one real double solution, and iii) a pair of complex solutions as

$$\lambda_\pm = \frac{-\alpha \pm \sqrt{\alpha^2 - 4\beta}}{2}$$

Case i). $\alpha^2 - 4\beta > 0$. Two real roots.
In this case

$$\lambda_+ = \frac{-\alpha + \sqrt{\alpha^2 - 4\beta}}{2}$$

and

$$\lambda_- = \frac{-\alpha - \sqrt{\alpha^2 - 4\beta}}{2}$$

are both real and the solution is

$$h_n = c_1(\lambda_+)^n + c_2(\lambda_-)^n$$

Since the equation is linear we know that the superposition of solutions is again a solution. Notice that there are now two free parameters c_1 and c_2 to accommodate the two initial conditions x_0 and x_1 required for a second order difference equation.

EXAMPLE 4.11.

$$x_{n+2} = x_{n+1} + x_n$$

The auxiliary equation is now

$$\lambda^2 - \lambda - 1 = 0$$

The solutions to this quadratic are

$$\lambda_\pm = \frac{1 \pm \sqrt{5}}{2}$$

Thus, the general solution to the homogeneous problem is

$$h_n = c_1\left(\frac{1+\sqrt{5}}{2}\right)^n + c_2\left(\frac{1-\sqrt{5}}{2}\right)^n$$

If we select $h_0 = h_1 = 1$ we have the *Fibonocci sequence* $\{1, 1, 2, 3, 5, 8, 13, \dots\}$. Employing this pair of initial conditions it is easily shown that the particular solution is

$$h_n = \left(\frac{\sqrt{5}+1}{2\sqrt{5}}\right)\left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{\sqrt{5}-1}{2\sqrt{5}}\right)\left(\frac{1-\sqrt{5}}{2}\right)^n$$

You might impress your friends by telling them the 50th number in this sequence $h_{50} = 20365011074$. It is also apparent that these numbers increase exponentially fast.

Case ii). $\alpha^2 - 4\beta = 0$. One real (double) root.

In this case

$$\lambda_+ = \lambda_- = -\frac{\alpha}{2}$$

so we only have one solution while we require two for the general solution of a second order difference equation.

It is not hard to verify that in this instance the second solution is actually

$$x_n = n(\frac{-\alpha}{2})^n$$

. (See Exercise 4.9). Now the general solution to this homogeneous equation is

$$h_n = c_1\big(-\frac{\alpha}{2}\big)^n + c_2 n\big(-\frac{\alpha}{2}\big)^n$$

EXAMPLE 4.12.

$$x_{n+2} + 2x_{n+1} + x_n = 0$$

The auxiliary equation is

$$\lambda^2 + 2\lambda + 1$$

which has the solution $\lambda = -1$. Thus, the general solution to this homogeneous problem is

$$h_n = c_1(-1)^n + c_2 n(-1)^n$$

Case iii). $\alpha^2 - 4\beta < 0$. Two complex roots.

The solution to the auxiliary equation is again

$$\lambda_\pm = \frac{-\alpha \pm \sqrt{\alpha^2 - 4\beta}}{2}$$

Based on the fact $\alpha^2 - 4\beta < 0$ we rewrite this as

$$\lambda_\pm = \frac{-\alpha}{2} \pm i\frac{\sqrt{4\beta - \alpha^2}}{2}$$

where $i = \sqrt{-1}$.[1]

We could now write the solution

$$h_n = c_1\left(\frac{-\alpha}{2} + i\frac{\sqrt{4\beta - \alpha^2}}{2}\right)^n + c_1\left(\frac{-\alpha}{2} - i\frac{\sqrt{4\beta - \alpha^2}}{2}\right)^n$$

but this form would not provide much insight. Instead we employ Demoivre's theorem that states

$$\exp(inx) = \cos(nx) + i\sin(nx)$$

To exploit this formula we need to recall that each solution to the auxiliary equation can be written in its complex polar form

$$z = x + iy = r\exp(i\theta)$$

where $x = r\cos\theta$ and $y = r\sin\theta$. Thus, we take

$$x = \frac{-\alpha}{2}, \quad \text{and} \quad y = \frac{\sqrt{4\beta - \alpha^2}}{2}$$

To compute the polar form we need r and θ. Recall

$$r^2 = x^2 + y^2$$

[1]Unlike the previous cases we now assume familiarity with basic complex numbers.

so

$$\begin{aligned} r^2 &= (\frac{-\alpha}{2})^2 + (\frac{\sqrt{4\beta - \alpha^2}}{2})^2 \\ &= \beta \end{aligned}$$

So

$$r = \sqrt{\beta}$$

The angle satisfies

$$\tan\theta = \frac{y}{x} = \frac{\sqrt{4\beta - \alpha^2}}{-\alpha}$$

In polar form, the solution is

$$h_n = c_1 r^n \exp(in\theta) + c_2 r^n \exp(-in\theta)$$

EXAMPLE 4.13. Find the general solution to the homogeneous difference equation

$$x_{n+2} + 2x_{n+1} + 5x_n = 0$$

The auxiliary equation gives the solutions

$$\lambda_{\pm} = -2 \pm i$$

If we write these in polar form we have

$$h_n = 5^{n/2}(c_1 \exp(in\theta) + c_2 \exp(-in\theta))$$

where $\tan\theta = 1/2$. Simplifying this to the associated real valued form we have

$$h_n = 5^{n/2}(c_1 \cos(n\theta) + c_2 \sin(n\theta))$$

where we have used the facts that $\exp(in\theta) = \cos(n\theta) + i\sin(n\theta)$ and that the real and imaginary parts of a complex solution are real solutions (see problems).

4.1.2.1. *The Cobweb Model Revisited.* Consider a supply curve

$$p = m_s q + b_s$$

and a demand curve

$$p = m_d q + b_d$$

Here we derive a formula for the values (q_n, p_n) that are the iterations along the supply and demand curves that either converge to an economic equilibrium or spiral out of control. Let the starting point on the demand curve be (q_0, p_0). The next iteration is then given by

$$(q_1, p_1) = (\frac{p_0 - b_s}{m_s}, p_0)$$

Similarly,

$$(q_2, p_2) = (q_1, m_d q_1 + b_d),$$

$$(q_3, p_3) = (\frac{p_2 - b_s}{m_s}, p_2)$$

and

$$(q_4, p_4) = (q_3, m_d q_3 + b_d)$$

Thus, we have established the following pattern:

$$(q_{2n}, p_{2n}) = (q_{2n-1}, m_d q_{2n-1} + b_d)$$

and

$$(q_{2n+1}, p_{2n+1}) = (\frac{p_{2n} - b_s}{m_s}, p_{2n})$$

It is now possible to create a second order difference equation for both q_n and p_n. Since

$$q_{2n+1} = \frac{p_{2n} - b_s}{m_s}$$

it follows, upon substituting for p_{2n} that

$$q_{2n+1} = \frac{(m_d q_{2n-1} + b_d) - b_s}{m_s}$$

or,

$$q_{2n+1} = \frac{m_d}{m_s} q_{2n-1} + \frac{b_d - b_s}{m_s}$$

4.1.2.2. *Nonhomogeneous Second Order Equations.* Consider the equation

$$q_{2n+1} = \alpha q_{2n-1} + \beta$$

This is a nonhomogeneous second order difference equation whose general solution is given using the theory of the first order case, i.e.,

$$x_n = h_n + p_n$$

The associated homogeneous equation is given by

$$q_{2n+1} = \alpha q_{2n-1}$$

which has the auxiliary equation

$$\lambda^2 = \alpha$$

so the general solution to the homogeneous problem is

$$h_n = c_1 \alpha^{n/2} + c_2 (-\alpha^{1/2})^n$$

As the nonhomogeneous term is a constant we first search for a particular solution of the form $p_n = A$. This must be an equilibrium solution, if it exists. Solving for A then gives

$$A = \alpha A + \beta$$

or

$$A = \frac{\beta}{1 - \alpha}$$

In terms of the original variables the general solution to the nonhomogeneous equation is

$$q_n = c_1 \left(\frac{m_d}{m_s} \right)^{n/2} + c_2 \left(-(\frac{m_d}{m_s})^{1/2} \right)^n + \frac{b_d - b_s}{m_s - m_d}$$

It is clear from our previous work that this equation will only converge if

$$|\frac{m_d}{m_s}| < 1$$

Note also that if this condition holds then the quantity supplied

$$q_n \to \frac{b_d - b_s}{m_s - m_d}$$

the market equilibrium.

4.2. Population Modeling

In this section we will consider a sequence of modifications of a population model that characterize the modeling process and illustrate how including or deleting terms in equations can have dramatic effects on the predictive powers of a model.

The simplest model for population growth makes the assumption that there is no competition for resources such as nutrients or habitat. This exponential growth is readily captured by the simple difference equation

$$p_{n+1} - p_n = \Delta p_n = kp_n \tag{4.4}$$

where the growth constant $k > 0$ reflects the rate of reproduction. One would assume that for rabbits this constant would be larger than for elephants. Actual values for k must be determined empirically from the data using a data fitting technique such as least squares.

If instead of simply taking $k > 0$ in Equation (4.4) we could have modeled both the birth rate k_b and the death rate k_d such that

$$p_{n+1} - p_n = k_b p_n - k_d p_n \tag{4.5}$$

Clearly now we may write

$$k = k_b - k_d$$

and as we would expect, if $k > 0$ the model predicts that the population grows exponentially fast and if $k < 0$ then the population decays exponentially fast.

The effect of adding to a population via immigration or subtracting via emmigration is capture by

$$p_{n+1} - p_n = kp_n + \beta_n \tag{4.6}$$

where β_n is the net flux of population. Now we might expect that growth rates could be offset by immigration or emmigration. For example $k < 0$ but $\beta_n = \beta$ can produce a positive equilibrium population.

Obviously ignoring competition for finite resources places significant limitations on this model. It will work well where the assumptions hold true but when the effects of competition for resources become important it will not capture them. To model competition we may argue as follows: competition occurs when there is interaction between two members of a species and the total amount of competition is the number of ways we can select subsets of 2 from a population p which is

$$\text{number of pairwise interactions} \propto \frac{p(p-1)}{2}$$

where we have divided by two to compute the number of combinations rather than permutations. Now we may modify the model to incorporate competition as

$$p_{n+1} - p_n = k_1 p_n - k_2 p_n(p_n - 1) \tag{4.7}$$

again ignoring effects due to migration. Here we are assuming $k_2 > 0$ and use the negative sign to reflect the fact that competition reduces the population. This equation can be simplified to

$$p_{n+1} - p_n = c_1 p_n - c_2 p_n^2 \tag{4.8}$$

This is the well-known *logistic* difference equation for population growth and it appears to correspond well to the growth of bacteria in agar jelly, for example.

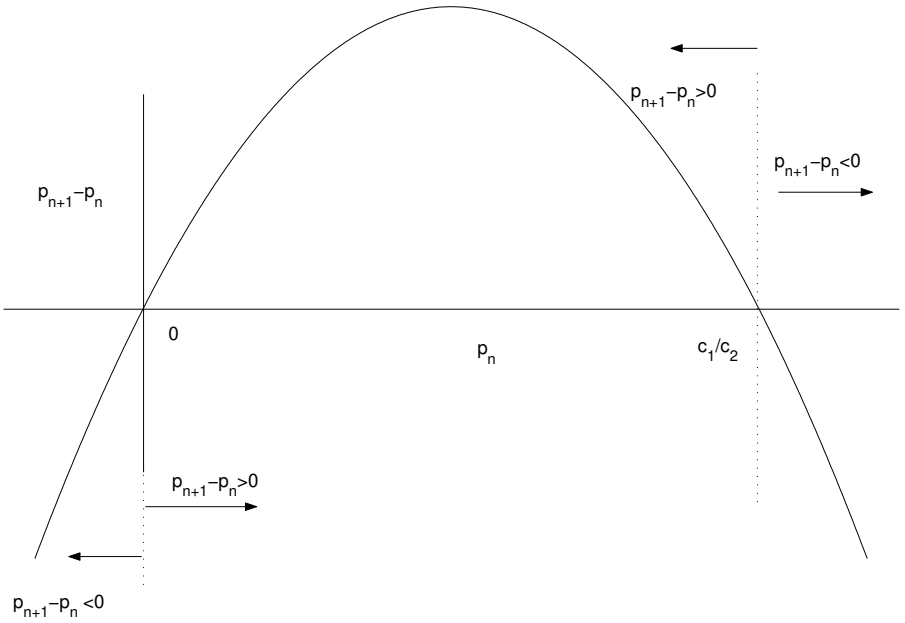


FIGURE 2. A plot of the change in population $p_{n+1} - p_n$ as a function of the population p_n.

Superficially we see that the difference between the model that does not model competition and the one that does is a quadratic term. A more fundamental difference is that Equation (4.4) is linear while Equation (4.8) is nonlinear. The only fixed point for Equation (4.4) is $\overline{p} = 0$. For Equation (4.8) there are now two fixed points $\overline{p_1} = 0$ and $\overline{p_2} = \frac{c_1}{c_2}$; see Figure 2. From the plot of $p_{n+1} - p_n$ it is clear that this new model predicts that the population will be limited, i.e., it can't grow unbounded to ∞ because as soon as $p_n > c_1/c_2$ then $\Delta p_n < 0$ so the population must decrease.

One is initially tempted to conclude that the equilibrium point $\overline{p_1} = 0$ is unstable while the equilibrium point $\overline{p_2} = c_1/c_2$ is stable. As we shall see in the numerical simulations this can be true, but for certain values of c_1 and c_2 the situation can be much more complicated including periodic and even chaotic solutions!

4.2.1. Systems of Equations and Competing Species. Now consider two species A and B whose populations are denoted a_n and b_n, respectively. If we assume that these species have infinite resources and compete neither with themselves or each other then we would propose the simple *system* of difference equations

$$a_{n+1} - a_n = k_1 a_n$$

$$b_{n+1} - b_n = k_2 b_n$$

This system is said to be *uncoupled* as the values of a_n do not influence b_n and, similarly, the values of b_n do not influence a_n.

If species B eats the same kind of food species A does, but species A does not eat the same kind of food species B does we have the model

$$a_{n+1} - a_n = g_1 a_n - c_1 a_n b_n$$
$$b_{n+1} - b_n = g_2 b_n$$

If species A and B both like each others food we would employ the model

$$a_{n+1} - a_n = g_1 a_n - c_1 a_n b_n$$
$$b_{n+1} - b_n = g_2 b_n - c_2 a_n b_n$$

See Figure 3 for a numerical simulation of this system. Note that this nonlinear system does not have a closed form solution. For the parameters selected we see that even though species A initially has a lower population it appears to grow without bound while population B becomes extinct. Here we may conclude that species A is more fit than species B and consequently survives.

If species A and B compete both with each other and with themselves the population model would then become

$$a_{n+1} - a_n = g_1 a_n - c_1 a_n b_n - k_1 a_n^2$$
$$b_{n+1} - b_n = g_2 b_n - c_2 a_n b_n - k_2 b_n^2$$

Notice that if population B becomes extinct while A survives then the model reduces to the logistic difference equation for a single species.

4.2.1.1. *Predator Prey Model.* Now consider modeling the interaction between natural predators and their prey. A classic example of this relationship is given by foxes and rabbits. The population of foxes and rabbits are intimately linked given that the rabbits are the food supply for the foxes. When the population of rabbits increases one can predict an associated, though possibly time lagged, increase in the number of foxes. Conversely, when the number of rabbits decreases the less food there is for the foxes. Of course an increase in the number of foxes will result in more rabbits being eaten and thus a reduction in the rabbit population.

Let's develop a model for this situation. First, denote the fox population by f_n and the rabbit population by r_n. If we assume that in the absence of rabbits the fox population becomes extinct we have the model

$$\Delta f_n = -g_1 f_n$$

where the constant $g_1 > 0$. If rabbits are available, then they should contribute positively to a change in the fox population. It seems reasonable to assume that the increase in the fox population will be proportional to the number of fox and rabbit interactions which is given by the product $f_n r_n$. Thus, in the presence of rabbits we may model the change in the fox population to be

$$\Delta f_n = -g_1 f_n + c_1 f_n r_n$$

where the constant $c_1 > 0$

Now the rabbits should multiply in the absence of foxes

$$\Delta r_n = g_2 r_n$$

where the constant $g_2 > 0$. The impact of the foxes on the rabbits is presumably also proportional to the number of interactions but now this reduces the rabbit population.

$$\Delta r_n = g_2 r_n - c_2 f_n r_n$$

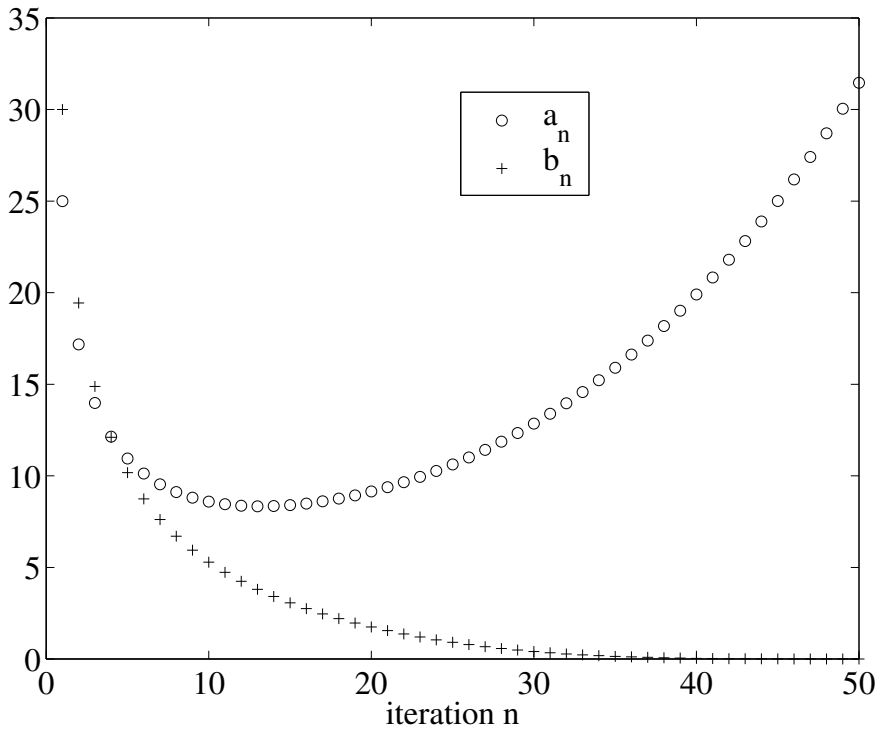


FIGURE 3. Competition for finite resources. We have selected the initial conditions $a_1 = 25$ and $b_1 = 30$. In addition the parameters were chosen to be $g_1 = .047, c_1 = .012, g_2 = .023, c_2 = .015$.

In summary we have the model

$$f_{n+1} = (1 - g_1)f_n + c_1 f_n r_n \tag{4.9}$$

$$r_{n+1} = (1 + g_2)r_n - c_2 f_n r_n \tag{4.10}$$

Note that we have omitted the competition amongst the foxes for the rabbits as well as the competition amongst the rabbits for their food. This is easily captured by extending the above system to

$$f_{n+1} = (1 - g_1)f_n + c_1 f_n r_n - d_1 f_n^2 \tag{4.11}$$

$$r_{n+1} = (1 + g_2)r_n - c_2 f_n r_n - d_2 r_n^2 \tag{4.12}$$

See Figure 4 for a simulation of the above equations. Note that the predicted oscillation is in fact there, however it is damped and the solution proceeds to a stable equilibrium.

4.3. Empirical Modeling

One may imagine that true observations, e.g., from populations in nature, will not be precise due to limitations in counting species in the wild. Thus, the data will contain what we refer to as a unknown *noise* component. In general, model selection can be arrived at by

(1) Collect observations to build models

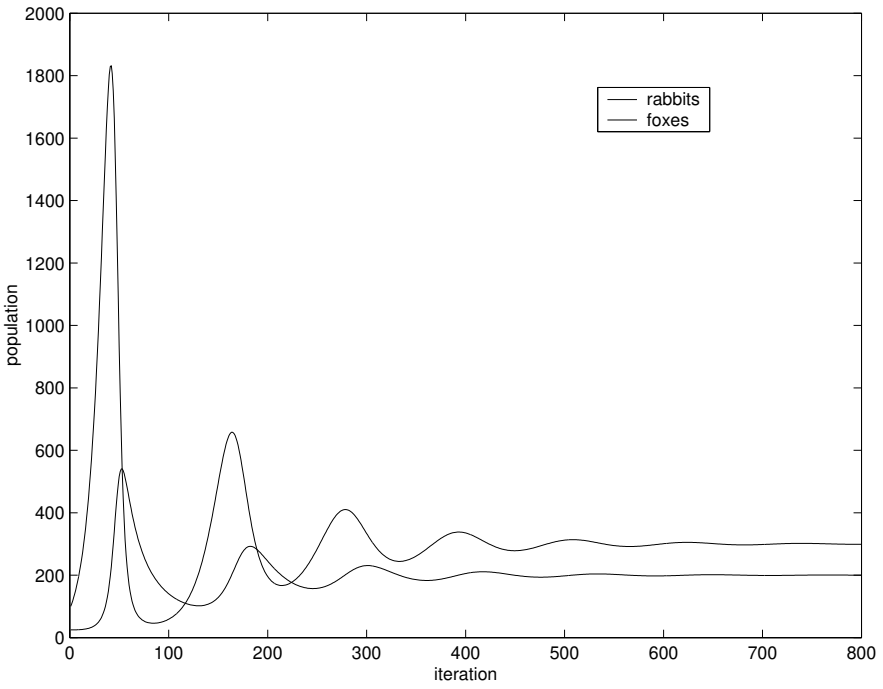


FIGURE 4. Simulation of preditor prey equations. We have selected the initial conditions $f_1 = 25$ and $r_1 = 100$. In addition the parameters were chosen to be $g_1 = 0.01, c_1 = 0.0001, g_2 = 0.1, c_2 = 0.0005$, $d_1 = 0.0001$ and $d_2 = 0$.

(2) Propose models, e.g., predator prey or competing species
(3) Compute model coefficients in each case
(4) Compare models through validation and testing

Now we present the method of least squares as a means to determine our unknown model coefficients.

4.3.1. Non-Newtonian Fish? Recall that Newton's Law of Cooling states that the temperature change in a body is proportional to the difference between the temperature of the body T_n and the surrounding temperature M, i.e., as a difference equation

$$\Delta_n = k(M - T_n)$$

After repeatedly overcooking a certain kind of fish based on this law a frustrated cook has decided to take science into her own hands. She speculates that the actual law of cooking for this fish has the more general form

$$\Delta_n = k(M - T_n)^\alpha$$

and that for certain types of foods, call them *Non-Newtonian foods*, that $\alpha \neq 1$.

To test her hypothesis, our cook measures the temperature of a fish every minute until it approaches the temperature of the oven which is set to 425 degrees F. The results of her data collection are shown in Figure 5.

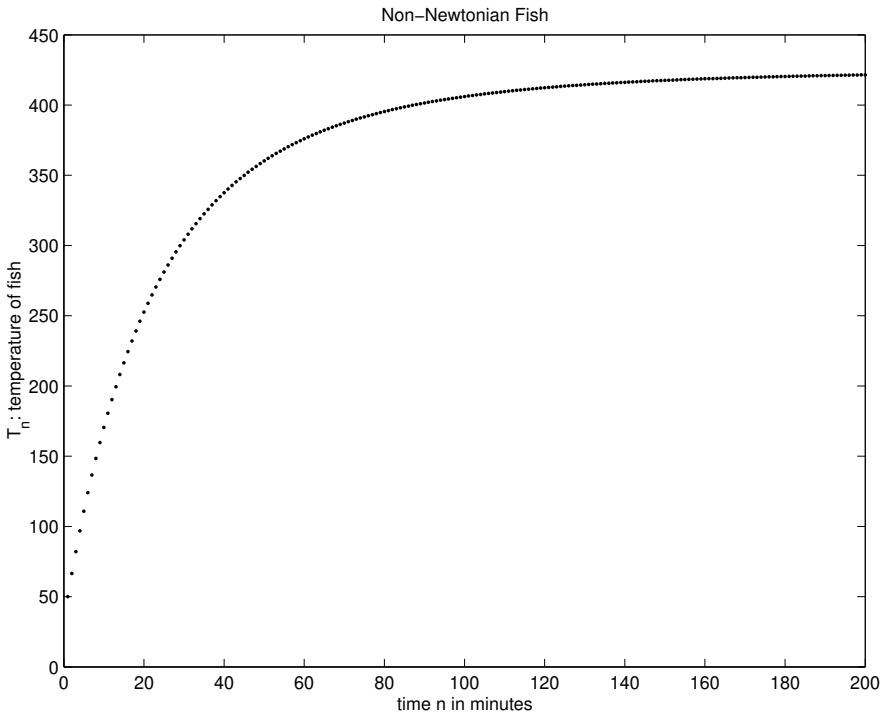


FIGURE 5. Observations of a Non-Newtonian fish. These are (synthetic) measurements of the temperature of the fish as a function of time.

Now Δ_n is known since T_n is known for $n = 1, \ldots, 200$. Thus, for any α and k we can compute a model error of

$$E(\alpha, k) = \sum_n \big(\Delta_n - k(425 - T_n)^\alpha\big)^2$$

We recall from our previous work with least squares that computing α and k requires differentiating the error term $E(\alpha, k)$ with respect to α and k. For this particular model it is simpler to employ a logarithmic transformation

$$y_n = \ln \Delta T_n$$

$$b = \ln k$$

$$x_n = \ln(425 - T_n)$$

giving

$$E(\alpha, b) = \sum_n (y_n - b - \alpha x_n)^2$$

Differentiating these with respect to α and b and setting the results equal to zero produces the equations

$$\begin{pmatrix} \sum_n x_n^2 & \sum_n x_n \\ \sum_n x_n & P \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} \sum_n y_n x_n \\ \sum_n y_n \end{pmatrix}$$

Solving these equations using *only the first 101 observations* $T_0, T_1, \ldots, T_{100}$ and the MATLAB code in Appenix A.4 produces the results

$$\alpha = 1.25$$

and

$$k = 0.01$$

4.3.2. Predator or Prey? Assume that the data in Figure 6 is provided. The goal is to see if we can calculate the coefficients of the preditor prey equations that will reproduce this data. Thus, given the tentative model

$$\Delta f_n = -g_1 f_n + c_1 f_n r_n$$
$$\Delta r_n = g_2 r_n - c_2 f_n r_n$$

the points $\{f_n, r_n\}$ are now observations while the equation coefficients $\{g_1, g_2, c_1, c_2\}$ are to be determined.

The least squares error is now

$$E(g_1, c_1, g_2, c_2) = \sum_n (\Delta f_n + g_1 f_n - c_1 f_n r_n)^2 + \sum_n (\Delta r_n - g_2 f_n + c_2 f_n r_n)^2 \tag{4.13}$$

Setting

$$\frac{\partial E}{\partial g_1} = \frac{\partial E}{\partial c_1} = \frac{\partial E}{\partial g_2} = \frac{\partial E}{\partial c_2} = 0$$

produces the necessary conditions for a minimum error. Taking the uncoupled equations for c_1 and g_1 we have

$$\begin{pmatrix} -\sum_n f_n^2 & \sum_n f_n^2 r_n \\ -\sum_n f_n^2 r_n & \sum_n f_n^2 r_n^2 \end{pmatrix} \begin{pmatrix} g_1 \\ c_1 \end{pmatrix} = \begin{pmatrix} \sum_n (\Delta f_n) f_n \\ \sum_n (\Delta f_n) f_n r_n \end{pmatrix} \tag{4.14}$$

These must be solved simultaneously with the uncoupled conditions for c_2 and g_2, i.e.,

$$\begin{pmatrix} -\sum_n r_n^2 & -\sum_n r_n^2 f_n \\ \sum_n r_n^2 f_n & -\sum_n f_n^2 r_n^2 \end{pmatrix} \begin{pmatrix} g_2 \\ c_2 \end{pmatrix} = \begin{pmatrix} \sum_n (\Delta r_n) r_n \\ \sum_n (\Delta r_n) f_n r_n \end{pmatrix} \tag{4.15}$$

Solving these equations produces the exact coefficients that were used to generate the data. In principal, this procedure may be applied to direct observations from nature. One may conclude if a model *fits* the data and, if so, which species plays which role, i.e., by examining the computed signs of c_1, c_2, g_1 and g_2 one may infer which species is the predator and which is the prey. See the MATLAB code for these equations in Appendix A.5.

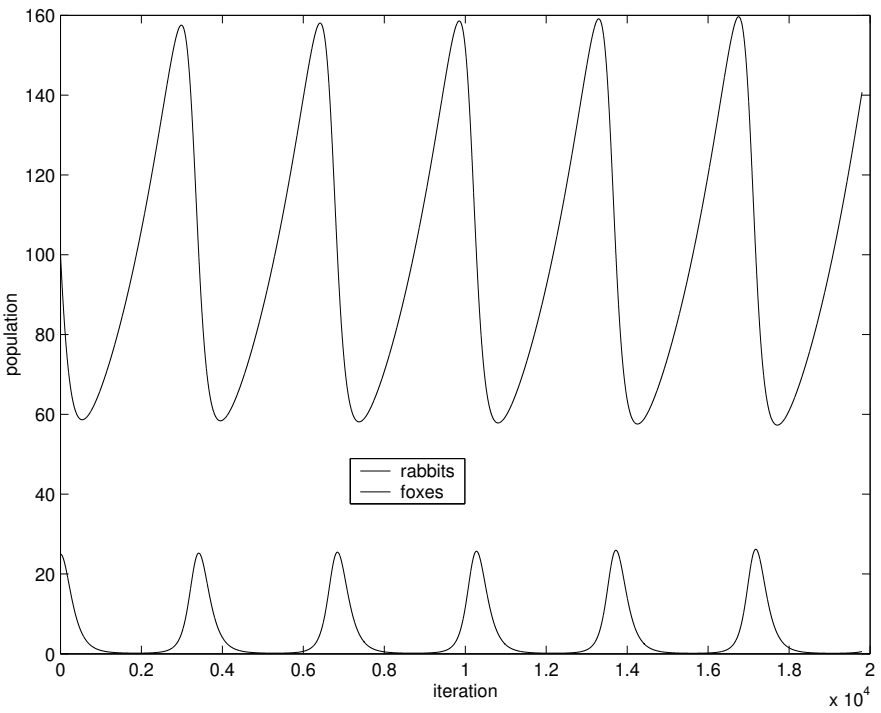


FIGURE 6. Simulation of preditor prey equations. We have selected the initial conditions $f_0 = 25$ and $r_0 = 100$. In addition the parameters were chosen to be $g_1 = .01, g_2 = .0005, c_1 = .0001, c_2 = .0001, d_1 = 0.0, d_2 = 0.0$

Problems

PROBLEM 4.1. Consider the following equations and identify as

- linear or nonlinear
- homogeneous or nonhomogenous
- which order

a) $x_{n+1}^2 + x_n = 1.$
b) $x_{n+1} = x_{n-1} + 2$
c) $x_{n+1} = \sin(x_{n-1})$
d) $x_{n+3} = x_{n+1} + x_{n-3} + n^2$

PROBLEM 4.2. Determine particular solutions to the following equations

a) $x_{n+1} = x_n + 1$
b) $x_{n+1} = 5x_n + n^2$
c) $x_{n+1} = \frac{x_n}{2} + 6^n$

PROBLEM 4.3. Show that the real and imaginary parts of a complex solution to a linear difference equation are also solutions to the same difference equation.

PROBLEM 4.4. Determine general solutions to the following equations

a) $x_{n+1} = x_n + 1$
b) $x_{n+1} = 5x_n + n^2$
c) $x_{n+1} = \frac{x_n}{2} + 6^n$

d) $x_{n+1} = \frac{x_n}{2} + 4n^2 + 2n + 1$

PROBLEM 4.5. Assume the temperature of a roast in the oven increase at a rate proportional to the difference between the oven (set to 400 degrees F) and the roast. If the roast enters the oven at 50 degrees F and is measured one hour later to be at 90 when should the table be set if the eating temperature is 166 degrees F? Hint: write down the difference equation and solve analytically.

PROBLEM 4.6. *Computer.* This question concerns numerically exploring the solutions of the equation

$$p_{n+1} = p_n + \alpha p_n(1 - p_n)$$

First determine all the *equilibrium* solutions of this difference equation by setting $\overline{p} = p_{n+1} = p_n$. Now investigate the stability of these equilibrium numerically. Consider the initial conditions

- $p_0 = 0$
- $p_0 = 0.0001$
- $p_0 = 2$

Numerically simulate the difference equation using the following values of α

- $\alpha = .1$
- $\alpha = .7$
- $\alpha = 1.2$

Describe your results and comment on the stability of the equilibrium you found. Provide plots of all your results. It will make your comparisons easier if you plot all the results for one value of α on a single graph.

PROBLEM 4.7. *Computer.* This question concerns numerically exploring the solutions of the equation

$$p_{n+1} = p_n + 0.1p_n(1 - p_n)(2 - p_n)$$

First determine all the *equilibrium* solutions of this difference equation. Numerically simulate the difference equation using the following initial conditions

- $p_0 = 0$
- $p_0 = 0.0001$
- $p_0 = .9999$
- $p_0 = 1$
- $p_0 = 1.0001$
- $p_0 = 1.9999$
- $p_0 = 2$
- $p_0 = 2.0001$

Describe your results and comment on the stability of the equilibrium you found. Provide plots of all your results. It will make your comparisons easier if you plot all the results on a single graph.

PROBLEM 4.8. *Computer.* Simulate the fourth order difference equation

$$p_{n+4} = \sin(p_{n+3} + p_{n+2} + p_{n+1} - p_n) + 2$$

and compare to the related equation

$$p_{n+4} = \sin(p_{n+2} + p_{n+1} - p_n) + 2$$

using the initial coditions $p_1 = 6, p_2 = 1, p_3 = 2.5, p_4 = -3$. Explore modifications to these difference equations and see if you can find any interesting behavior. For

example, what is the effect of varying the nonhomogeneous term? Plot your results in each case for 100 iterations.

PROBLEM 4.9. *Computer.* Consider the system of difference equations

$$x_{n+1} = 0.3x_n + 0.8y_n$$
$$y_{n+1} = 0.7x_n + 0.2y_n$$

Simulate these equations numerically for a variety of initial conditions and attempt to determine any stable equilibrium. Verify that you have actually determined an equilibrium solution by substituting into the original system. (Note that the equilibrium solution in this problem actually depends on the initial condition.)

a) How do the solutions change if you modify the first coefficient from 0.30 to 0.31?
b) How do the solutions change if you modify the first coefficient from 0.30 to 0.31 *and* modify the 0.7 coefficient to 0.69.
c) Compare the results in part a) and b). Can you explain?

PROBLEM 4.10. Consider the difference equation

$$x_{n+2} + \alpha x_{n+1} + \beta x_n = 0$$

where it is assumed that $\alpha^2 - 4\beta = 0$. Show that $x_n = (-\frac{\alpha}{2})^n n$ is a solution.

PROBLEM 4.11. Find the linear second order nonhomogeneous difference equation relating the price p_{2n+2} and p_{2n} in the cobweb model. Solve this equation and produce a convergence criterion. What does the equilibrium price converge to? Check your result by computing the point of intersection of the supply and demand curves.

PROBLEM 4.12. Determine analytical solutions to the following difference equations assuming in each case that $x_1 = 1$ and $x_0 = -1$. Plot your results.

a) $x_{n+2} + 3x_{n+1} + x_n = 0$
b) $10x_{n+2} + x_{n+1} + x_n = 0$
b) $x_{n+2} + \sqrt{3}x_{n+1} + \frac{3}{4}x_n = 0$

PROBLEM 4.13. Extend the population model with pairwise competition to include competition among groups of three. Furthermore, assume that the competition among groups of three is more intense than competition between pairs. Identity the new equilibrium solution(s). Use a plot of Δp_n versus p_n to argue whether the model predicts a bounded population.

PROBLEM 4.14. Consider a clam population that obeys the logistic difference Equation (4.8). Modify this equation to account for constant harvesting of the clams. By computing the new equilibrium points of the population model describe the impact of harvesting on the clam population.

PROBLEM 4.15. Consider three species A, B, C and the evolution of their populations a_n, b_n and c_n.

- Species A eats B and C
- Species B eats neither A nor B
- Species C eats only A.
- Species B eats waste products produced by species A and B.
- The population of both species A and B increase in the absence of other species.
- The population of species C decreases in the absence of A and B.

- Species C is competes with itself for food while this is not true for species A and B.

Write down a system of three coupled difference equations modeling the populations of the three species.

PROBLEM 4.16. *Computer.* Provide a model for the bee colony population data in Table 1. What does your model predict the long-term population to be?

day	1	2	3	4	5	6	7	8	9	10
number	20	25	60	85	111	146	177	182	184	171
day	11	12	13	14	15	16	17	18	19	20
number	179	167	161	146	159	154	162	166	166	168

TABLE 1. Bee colony population data.

PROBLEM 4.17. *Computer.* Find all the equilibrium solutions of the logistic map

$$x_{n+1} = \lambda x_n(1 - x_n)$$

as a function of λ. Letting $x_0 = 0.2$ numerically iterate this difference equation for 200 iterations for the following values of λ:

- $\lambda = 2$
- $\lambda = 3.2$
- $\lambda = 3.8282$
- $\lambda = 3.83$

Plot your results x_n as a function of n for each case and comment. Does this seem like a reasonable model for a population?

PROBLEM 4.18. *Computer.* Use a least squares approach to determine k in Newton's Law of cooling

$$T_{n+1} = T_n + k(M - T_n)$$

using the data generated by our empirical fish model

$$T_{n+1} = T_n + 0.01(M - T_n)^{1.25}$$

First generate 200 points using this equation and compute k based on these points. Now predict the next 200 points and calculate the error. If a fish is well cooked at 170 degrees F how long does each model predict it will take to cook the fish? Use the values $M = 425$ and $T_0 = 50$.

PROBLEM 4.19. *Computer.* Using the data provided in Table 2 estimate via least squares the coeffients c_1, c_2, d_1, d_2 in the model

$$a_{n+1} = a_n + c_1 a_n + d_1 b_n$$

$$b_{n+1} = b_n + c_2 a_n + d_2 b_n$$

Include your equations for the unknown coefficients in your write-up.

PROBLEM 4.20. Extend the Equations (4.14) and (4.15) provided for computing the coefficients c_1, c_2, g_1, g_2 for the predator-prey model with no intra-species competition given by Equation (4.9) to the case of Equations 4.11 where intraspecies competition is accounted for. Your equations should now provide estimates for $c_1, c_2, g_1, g_2, d_1, d_2$

n	a_n	b_n
1	15.00	45.00
2	30.00	30.00
3	24.00	36.00
4	26.40	33.60
5	25.44	34.56
6	25.82	34.18
7	25.67	34.33
8	25.73	34.27
9	25.71	34.29
10	25.72	34.28
11	25.71	34.29

TABLE 2. Did this data come from a linear system?

PROBLEM 4.21. Consider the differential equation for the unforced damped nonlinear pendulum

$$\frac{d^2x}{dt^2} + \alpha\frac{dx}{dt} + \sin x = 0$$

where $x(t)$ represents the angular displacement from the equilibrium in radians. Using the expressions for the numerical estimates of the derivatives

$$\frac{d^2x}{dt^2} = \frac{x_{n+1} + x_{n-1} - 2x_n}{(\Delta t)^2}$$

and

$$\frac{dx}{dt} = \frac{x_n - x_{n-1}}{\Delta t}$$

where $x_n \equiv x(n\Delta t)$.

a) Show that the differential equation can be approximated by the second order difference equation

$$x_{n+1} = (2 - \alpha\Delta t)x_n + (\alpha\Delta t - 1)x_{n-1} - (\Delta t)^2 \sin(x_n) \tag{4.16}$$

b) Simulate this difference equation for 1000 iterations using the values $\Delta t = 0.05, \alpha = 0.1, x_1 = 0, x_2 = 0.0001$ and plot your result. Repeat this calculation for $x_1 = 2, x_2 = 2.0001$ and compare your results.

c) Redo this simulation using the *small angle approximation* $\sin x = x$, i.e., simulate

$$x_{n+1} = (2 - \alpha\Delta t)x_n + (\alpha\Delta t - 1)x_{n-1} - (\Delta t)^2 x_n \tag{4.17}$$

using the values $\Delta t = 0.05, \alpha = 0.1, x_1 = 0, x_2 = 0.0001$ and plot your result. Again, repeat this calculation for $x_1 = 2, x_2 = 2.0001$ and compare your results with those found in part (c).

d) Rewrite the second order Equation (4.16) as a system of two first order equations via the substitution $y_{n+1} = x_n$ and determine all equilibria. Note that the equilibria can also be determined directly from Equation (4.16).

e) By computing the eigenvalues of the Jacobian matrix of this system, ascertain which equilibria are stable and unstable. Discuss.

PROBLEM 4.22. Repeat parts (d) and (e) of Problem 4.20 for the small angle approximation Equation (4.17) and compare.

PROBLEM 4.23. Analytically solve the linear difference equation from the previous problem

$$x_{n+1} = (2 - \alpha\Delta t)x_n + (\alpha\Delta t - 1)x_{n-1} - (\Delta t)^2 x_n$$

and compare with your numerical simulation above. For simplicity you may take $\Delta t = 0.05, \alpha = 0.1, x_1 = 2, x_2 = 2.0001$.

PROBLEM 4.24. Analytically solve the linear nonhomogeneous difference equation

$$x_{n+1} = (2 - \alpha\Delta t)x_n + (\alpha\Delta t - 1)x_{n-1} - (\Delta t)^2 x_n + 0.01\sin(n/50)$$

Simulate this problem numerically and compare with your analytical solution for 2000 iterations. Can you identify a transient (i.e., a term that goes to zero) and steady state (persisent) components of your solution? Again, for simplicity you may take $\Delta t = 0.05, \alpha = 0.1, x_1 = 2, x_2 = 2.0001$. Hint: combine your solution to the homogeneous problem found above with a particular solution of the form

$$p_n = A\cos(n/50) + B\sin(n/50)$$

Solve for the undetermined coefficients A and B.

CHAPTER 5

Linear Programming

Linear programming, like its nonlinear counterpart, is a method for making decisions based on solving a mathematical optimization problem. The general field of linear programming has been a major area of applied mathematical research in the last 50 years. A combination of new algorithms, e.g., the simplex method, and widely available computing power now make this an indispensible tool for the mathematical modeler.

We begin our discussion of linear programming by presenting the basic mathematical formulation and terminology in general terms. We will follow this with a number of examples of problems that may be formulated in terms of linear programs. Our goal here is to obtain an abstract understanding of what a linear program is and to develop an intuition that will assist the modeler in assessing whether linear programming is the right tool for a given problem.

Consider a linear function of the variables $x = (x_1, \ldots, x_n)$

$$f(x) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

where the parameters a_i are known. We seek to pick the values of all the x_i, referred to as *decision variables*, so as to maximize $f(x)$ which is referred to as the *objective function*. Clearly picking each $x_i = \infty$ (or even just one) would provide a maximum, albeit meaningless. The interest arises when the values of the x_i are constrained, e.g.,

$$c_{11}x_1 + \cdots + c_{1n}x_n \leq b_1$$

Based on the application many constraints are possible so we write

$$c_{i1}x_1 + \cdots + c_{in}x_n \leq b_i$$

for $i = 1, \ldots, m$. Note that these constraints are also linear in the decision variables. We may interpret this system of constraints geometrically as defining a region, i.e., a continuum of points such that all the constraints are simultaneaously satisfied. This region is referred to as the *feasible set* S. So now we may view the optimization problem as one to find the maximum value of the objective function over the feasible set S.

In summary, a linear programming problem is defined as

$$\max a^T x$$

subject to the constraints

$$(Cx)_i \leq b_i$$

where $(C)_{ij} = c_{ij}$.

5.1. Examples of Linear Programs

In this section we survey a variety of applications that fit exactly into the formulation of the abstract linear program.

5.1.1. Red or White? A winemaker would like to decide how many bottles of red wine and how many bottles of white wine to produce. Given his expertise is in red wine making he can sell a bottle of red wine for $12 while he can only sell a bottle of white wine for $7. Clearly the winemaker would seek to maximize his profits, and, having recently completed a course in mathematical modeling, proceeds to construct the objective function

$$f(x_1, x_2) = 12x_1 + 7x_2$$

where the decision variables are the number of bottles of red wine to produce x_1 and the number of bottles of white wine to produce, i.e., x_2.

Aging wine in wooden or glass-lined vats is an integral component of the production process but do to limited space the wine must be aged for a limited time. The wine maker has determined that red wine should be aged two years per bottle and white wine one year per bottle and his facilities allow that each batch is limited to 10,000 bottle-years (5 bottles of red and 3 bottles of white require a total of 13 bottle years ripening time). Thus the winemaker formulates a constraint

$$2x_1 + x_2 \leq 10000$$

Also the volume of grapes that may be processed is limited and it takes 3 gallons of grapes to make a bottle of red wine and two gallons of grapes to make a bottle of white wine. Furthermore, the winery can only process a total of 16,000 gallons of grapes for each batch. Thus, the winemaker produces the additional constraint

$$3x_1 + 2x_2 \leq 16000$$

Now the winemaker would like to determine how many bottles of each wine to produce as well as how much money he will expect to make. Note that we must also require that negative bottles of wine are not allowed so

$$x_1 \geq 0$$

and

$$x_2 \geq 0$$

5.1.2. How Many Fish? A child with a new 29 gallon fish tank asks her daddy to put as many fish in the tank as possible. Sensing that too many fish is not a good thing, the dad asks the pet shop owner how many fish can go into a tank. The answer was more complex than anticipated. "You can put one inch of fish in per gallon of water." The little girl then added that she wanted only the big orange fish (Gouramis) and the small stripey fish (Zebra Dannios).

As the child seeks to maximize the total number of fish her objective function is

$$f(x_1, x_2) = x_1 + x_2$$

where x_1 is the number of Gouramis and x_2 is the number of Zebra Dannios.

Additionally, a full grown Gourami is two inches long while a Danio is just one inch long. The constraint of not exceeding 29 inches of total fish length can now be written

$$2x_1 + x_2 \leq 29$$

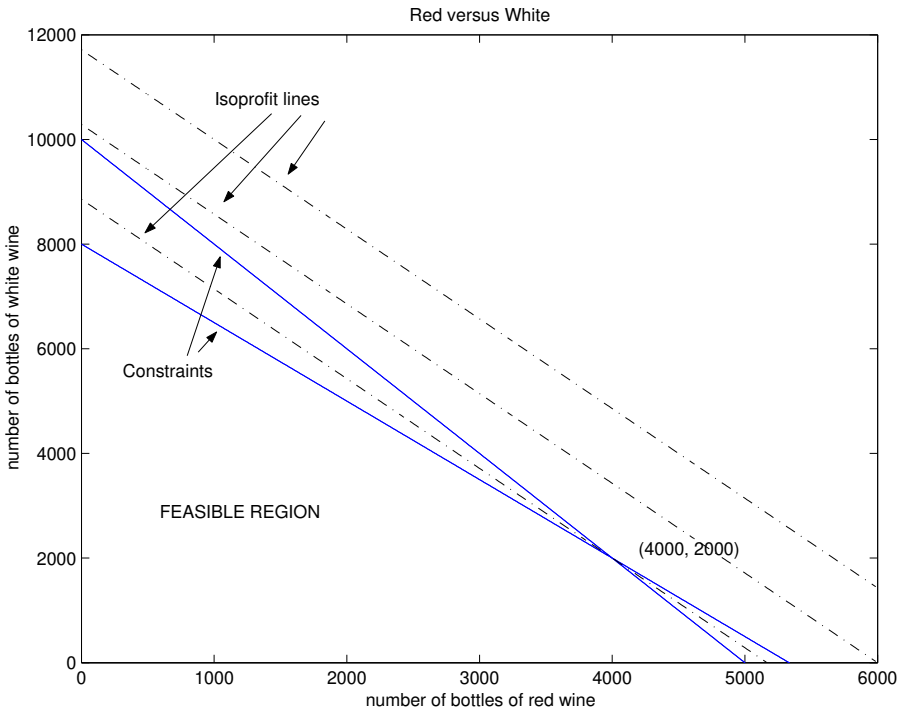


FIGURE 1. Geometric picture of the linear programming problem.

Danios are very active fish and actually require twice as much food as Gouramis. Each Dannio eats 4 grams/day of fish flakes while the slower Gourami eats 2 grams/day. The dad decides that he would perfer not to go broke buying fish food and thus wants to limit the tank to 50 grams/day. Thus, we have the constraint

$$2x_1 + 4x_2 \leq 50$$

The pet shop owner adds, by the way, that Danios need to live in schools of at least 5 fish or they don't do well. Thus

$$x_2 \geq 5$$

Additionaly, the little girl stipulates that she must have at least two Gouramis as they are known to kiss (hence the term Kissing Gouramis) so we add

$$x_2 \geq 2$$

How many Gouramis and Dannios can the little girl have in her tank?

5.2. Geometric Solution of a 2D Linear Program

Let us now solve the winemaker's linear programming problem using graphical techniques. Recalling the problem:

- Objective function: $f(x_1, x_2) = 12x_1 + 7x_2$
- Constraint 1: $3x_1 + 2x_2 \leq 16000$
- Constraint 2: $2x_1 + x_2 \leq 10000$
- Constraint 3: $x_1 \geq 0$

- Constraint 4: $x_2 \geq 0$

First, let us identify the feasible set. Again, this is the intersection of all the regions defined by the constraints. (Note that this set is independent of the objective function.) The boundary of the first constraint is defined by the equality

$$x_2 = 8000 - \frac{3}{2}x_1$$

The constraint may be viewed as a half-plane with this line dividing the region of allowed points from the unallowed points. It is easy to identify which region is the allowed region by considering a single point. For example, is the origin a point that satisfies the first constraint? Since the answer is clearly yes we know that the set of points that satisfies constraint 1 consists of the halfplane defined by $x_2 = 8000 - \frac{3}{2}x_1$ that contains origin.

Similarly, the second constraint defines a halfplane of points containing the origin and bounded by the line

$$x_2 = 10000 - 2x_1$$

The intersection of contraints 3 and 4 is the first quadrant of the x_1x_2 plane.

The intersection of all of these constraints as shown in Figure 1 constitutes the feasible set. Now we must pick the point in the feasible set that maximizes the objective function.

We can define an *isoprofit line* to be

$$12x_1 + 7x_2 = c$$

For all points on this line the profit is the same. We can see that as c decreases the line shifts towards the origin. So the goal is to pick the isoprofit line with the largest value of c such that x_1, x_2 is a point in the feasible set. Graphically we see that the first point the descending isoprofit line will touch is the vertex of the intersection of constraints 1 and 2. This is easily calculated to be $(4000, 2000)$.

Thus, the solution to the winemaker's linear programming problem is that he should produce 4000 bottles of red and 2000 bottles of white and that this will lead to a maximum profit of \$62,000.

5.3. Sensitivity Analysis

Often the coefficients in a linear programming model are known only approximately. Thus, it is interesting to know what the impact of modifying the terms present in the model. How is the objective function impacted? How does the optimal solution change? These questions are the subject of *sensitivity analysis*.

5.3.1. Price Sensitivity. First we examine how changing the price of a bottle of white wine impacts the optimal solution. Letting the price of the white wine be a variable w we now have the linear program

- Objective function: $f(x_1, x_2) = 12x_1 + wx_2$
- Constraint 1: $3x_1 + 2x_2 \leq 16000$
- Constraint 2: $2x_1 + x_2 \leq 10000$
- Constraint 3: $x_1 \geq 0$
- Constraint 4: $x_2 \geq 0$

From our graphical solution we know that any isoprofit line with slope between -2 and -3/2 will produce the same optimal solution of $(4000, 2000)$. Since the slope of the isoprofit line is $-12/w$ this condition is

$$-2 < \frac{12}{w} < -\frac{3}{2}$$

from which we conclude that the price of the white wine may vary as

$$6 < w < 8$$

with the solution unchanged as $(4000, 2000)$. Further examination produces Table **??**. The double arrows here mean that any point on the isoprofit curve containing these points produces the same profit.

cost of white wine	optimal solution
$6 < w < 8$	(4000, 2000)
$w = 6$	$(4000, 2000) \leftrightarrow (5000, 0)$
$w = 8$	$(4000, 2000) \leftrightarrow (0, 8000)$
$w < 6$	(5000, 0)
$w > 8$	(0, 8000)

TABLE 1. The effect of pricing the white wine on the optimal solution.

5.3.2. Resource Sensitivity. Now we left the number of gallons of grapes α and the number of bottle-years storage capacity β be variable. Now the linear program becomes

- Objective function: $f(x_1, x_2) = 12x_1 + 7x_2$
- Constraint 1: $3x_1 + 2x_2 \leq \alpha$
- Constraint 2: $2x_1 + x_2 \leq \beta$
- Constraint 3: $x_1 \geq 0$
- Constraint 4: $x_2 \geq 0$

Now the relative values of α and β determine the geometry of the solution. For example, if $\alpha/2 > \beta$ then constraint 1 becomes irrelevant. When the intersection of constraints 1 and 2 determines the optimal solution it is readily shown that

$$x_1 = -\alpha + 2\beta$$

and

$$x_2 = 2\alpha - 3\beta$$

Now the optimal solution to the objective function can be expressed as

$$f(x_1, x_2) = 2\alpha + 3\beta$$

Hence if α is increased by one unit then $f(x_1, x_2)$ is increased by 2, while if β is increased by one unit then $f(x_1, x_2)$ is increased by 3. So if a winemaker considers expanding his winery he realizes that the cost of increasing grape processing must be less than \$2 and the expense of increasing wine storage must be less than \$3. Otherwise expansion will lose money.

5.3.3. Constraint Coefficient Sensitivity. Now we consider the problem of adjusting one of the coefficients in one of the constraint equations. In particular consider the amount of time γ we age a bottle of red wine to be allowed to vary.

- Objective function: $f(x_1, x_2) = 12x_1 + 7x_2$
- Constraint 1: $3x_1 + 2x_2 \leq 16000$
- Constraint 2: $\gamma x_1 + x_2 \leq 10000$
- Constraint 3: $x_1 \geq 0$
- Constraint 4: $x_2 \geq 0$

To simplify the discussion, let's examine the effect of reducing the amount of time we age the red wine from 2 years to 1.95 years. The solution to the resulting linear program suggests that now 4444 bottles of red can be sold while 1333 bottles of white can be sold for a total profit of \$62,659, increasing the income by almost \$700. Of course, for this to be advisable it must be true that all the bottles of this "younger" red wine can still be sold at the same price, i.e., the taste has not suffered enough to reduce its popularity.

5.4. Data Fitting and the Uniform Approximation

Now let us reconsider the topic of fitting a model to a data set. This can also be put into the framework of a linear program. Again, if we have a data set of domain (input) values $\{x_i\}$ and associated range (output) values $\{y_i\}$ we would like to determine the parameters $\{w_1, \ldots, w_k\}$ such that

$$y_i = f(x_i; w_1, \ldots, w_k)$$

An alternative to requiring that the sum of the squares of the residuals be zero is to simply minimize the maximum residual. This approach is known as the *uniform approximation*, or *Chebyshev* criterion. Since large negative residuals would make this meaningless we minimize the maximum absolute value of the residual.

To implement this idea, we first compute each residual ϵ_i as

$$\epsilon_i = y_i - f(x_i; w_1, \ldots, w_k)$$

and from all these determine the largest

$$\epsilon_{\max} = \max_i |\epsilon_i|$$

which will serve as our objective function. So the programming problem is

$$\min_i \epsilon_{\max}$$

where based on the definition of $\epsilon_{\max}$ we have the side constraints

$$|\epsilon_i| \leq \epsilon_{\max}$$

So, for a linear model, $f(x) = ax + b$ we have

$$\epsilon_i = y_i - ax_i - b$$

so the constraints become

$$|y_i - ax_i - b| \leq \epsilon_{\max}$$

or

$$-\epsilon_{\max} \leq y_i - ax_i - b \leq \epsilon_{\max}$$

Thus the linear program is to

$$\min \epsilon_{\max}$$

subject to the constraints

$$-\epsilon_{max} - y_i + ax_i + b \leq 0$$

$$-\epsilon_{\max} + y_i - ax_i - b \leq 0$$

where $i = 1, \ldots, P$.

In matrix notation we may rewrite the constraints as

$$\begin{pmatrix} -x_1 & -1 & -1 \\ x_1 & +1 & -1 \\ \vdots & \vdots & \vdots \\ -x_1 & -1 & -1 \\ x_1 & +1 & -1 \\ \vdots & \vdots & \vdots \\ -x_P & -1 & -1 \\ x_P & +1 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ \epsilon_{\max} \end{pmatrix} \leq \begin{pmatrix} -y_1 \\ +y_1 \\ \vdots \\ -y_i \\ +y_i \\ \vdots \\ -y_P \\ +y_P \end{pmatrix}$$

Now solving this linear program to implement the uniform approximation approach on the uniform noise data produces the linear model equation

$$y = 0.9937x - 0.275$$

while the least squares error criterion produces the model

$$y = 0.8923x - 0.5466$$

For the uniform noise data the squared error was found to be 11.543 for the uniform approximation model while for the least squares model it is 9.998. Please see the errors in Table 2 and the comparitive plot of the two models in Figure 2.

x_i	y_i	ϵ_i uniform	ϵ_i least squares
1	2.3525	1.6338	0.9136
2	0.0786	-1.6338	-2.2527
3	3.7251	1.0191	0.5016
4	3.5179	-0.1818	-0.5979
5	6.3272	1.6338	1.3190
6	6.0113	0.3242	0.1108
7	7.8379	1.1571	1.0451
8	7.7156	0.0411	0.0305
9	8.2185	-0.4496	-0.3589
10	8.7586	-0.9032	-0.7111

TABLE 2. The data to be fit comprise the first two columns. The pointwise error for the uniform approximation and least squares approximation are in columns three and four, respectively. The underlined errors are those with maximum magnitude for each model. As expected, the uniform approximation has a smaller maximum error.

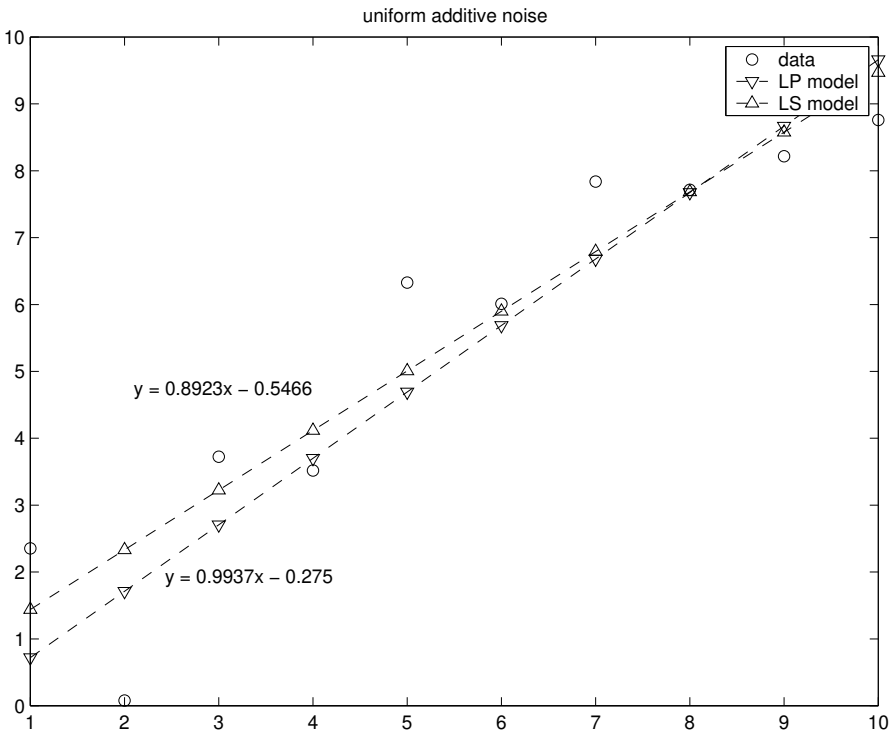


FIGURE 2. Least squares and uniform approximation to a linear trend with uniformly distributed additive noise.

5.4.1. Error Model Selection? Now we have seen that there is no unique way to compute the coefficients that fit a given model to data. The model is dependent on the way we measure the error (note that if our model is exact–e.g., the interpolation condition is satisfied–then the coefficients are unique and the error is zero). So the natural question arises: given a collection of data what is the appropriate error measure. The answer to this question lies partly in the nature of the data. If your data is very accurate, then a uniform approximation is indeed appropriate. If your data contains statistical outliers or lots of noise, a least squares approximation may be more robust.

The ultimate decision factor in what error term to use is in the added value of the model. If the predictive value of the model is superior in one error measure than another the choise is clear. Establishing superiority can often be challenging in practice.

PROBLEM 5.1. Think of an optimization problem that can be written as a linear program with two decision variables. Specify the objective function as well as the constraints. Avoid constructing a problem the solution to which is that one of the decision variables is zero. Can you extend your problem to more decision variables and constraints?

PROBLEM 5.2. Solve the question of how many Zebra Dannios and Gouramis should be purchased for the fish tank modeled in section 5.1.2.

PROBLEM 5.3. A new burger chain, the EcoliExpress, has has two new products: the large 1/3 pound "Big Whoopie" burger and the smaller 1/4 pound "Wimpy Whoopie" burger. It has been determined in test market trials that the Big Whoopie can be sold at a profit of 45 cents per burger and the Wimpy Whoopie at a profit of 25 cents. Furthermore, a chain knows that it can sell all its burgers if it uses 100 pounds of meat per week. In addition, the preparation time for a Big Whoopie is two minutes and for a Wimpy Whoopie is one minute and the chain has one employee working 40 hours per week preparing both types of Whoopies. Assuming the owner of the EcoliExpress wishes to maximize profits formulate a solution using linear programming. Using this model, answer the following:

a) How many Big Whoopies and Wimpy Whoopies should be sold?
b) Assuming the unit profit of the Big Whoopie is fixed at 45 cents, for what range of prices of the Wimpy Whoopie is the solution in a) optimal?
c) Assuming the unit profit of the Wimpy Whoopie is fixed at 25 cents, how large does the unit profit for the Big Whoopie have to be to justify making only this type of burger?
d) What should the cost of meat be (per pound) to justify purchasing additional quantities? Hint: the profit must increase.

PROBLEM 5.4. Fit the model $y = ax$ to the data in Table 3 using the uniform approximation criterion. Identify the decision variables and the objective function and graphically solve this linear programming problem. Finally, draw a graph of the model with the errors labeled for each point and comment.

x_i	8	16	32
y_i	10	14	31

TABLE 3. Data for model in Problem 5.4.

PROBLEM 5.5. If in Problem 5.4 the values y_i where each increased by one unit how would the coefficient a change? Consider the three cases separately.

PROBLEM 5.6. Consider the data in Table 4 and the model

$$y = ax^2 + bx + c$$

Write down the linear program to compute a, b, c according to uniform approximation criterion developed in Section 5.4. Identify the decision variables and reformulate the linear program in matrix notation.

PROBLEM 5.7. *Data Fitting.* Construct a linear program to fit the model $y = ax + b$ to the data set $\{(x_i, y_i)\}$ using the l_1 error criterion

$$\min_{a,b} \sum_i |y_i - ax_i - b|$$

x_i	1	2	3	4
y_i	2.3	5.3	8.7	16.1

TABLE 4. Data for model in Problem 5.4.

CHAPTER 6

Nonlinear Programming

A nonlinear programming problem has the form

$$\min f(x)$$

subject to the inequality constraints

$$g_i(x) \leq 0$$

for $i = 1, \ldots, p$ and the equality constraints

$$h_i(x) = 0$$

for $i = 1, \ldots, q$.

EXAMPLE 6.1. Given a fixed area of cardboard A construct a box of maximum volume. The nonlinear program for this is

$$\min xyz$$

subject to

$$2xy + 2xz + 2yz = A$$

EXAMPLE 6.2. Consider the problem of determining locations for two new high schools in a set of P subdivisions N_j. Let w_{1j} be the number of students going to school A and w_{2j} be the number of students going to school B from subdivision N_j. Assume that the student capacity of school A is c_1 and the capacity of school B is c_2 and that the total number of students in each subdivision is r_j. We would like to minimize the total distance traveled by all the students given that they may attend either school A or B. It is possible to construct a nonlinear program to determine the locations (a, b) and (c, d) of high schools A and B, respectively assuming the location of each subdivision N_i is model as a single point denoted (x_i, y_i).

$$\min \sum_{j=1}^{P} w_{1j}\left((a - x_j)^2 + (b - y_j)^2\right)^{\frac{1}{2}} + w_{2j}\left((c - x_j)^2 + (d - y_j)^2\right)^{\frac{1}{2}}$$

subject to the constraints

$$\sum_j w_{ij} \leq c_i$$

$$w_{1j} + w_{2j} = r_j$$

for $j = 1, \ldots, P$.

EXAMPLE 6.3. Neural networks have provided a new tool for approximating functions where the functional form is unknown. The approximation takes on the form

$$f(x) = \sum_j b_j \sigma(a_j x - \alpha_j) - \beta$$

and the corresponding sum of squares error term is

$$E(a_j, b_j, \alpha_j, \beta) = \sum_i \big(y_i - f(x_i)\big)^2$$

The problem of minimizing the error function is, in this instance, an unconstrained optimization problem. An efficient means for computing the gradient of E is known as the *backpropogation* algorithm.

6.1. Unconstrained Optimization in One Dimension

Here we begin by considering a significantly simplified (but nonetheless important) nonlinear programming problem, i.e., the domain and range of the function to be minimized are one-dimensional and there are no constraints. A necessary condition for a minimum of a function was developed in calculus and is simply

$$f'(x) = 0$$

Note that higher derivative tests can determine whether the function is a max or a min, or the value $f(x + \delta)$ may be compared to $f(x)$.

Note that if we let

$$g(x) = f'(x)$$

then we may convert the problem of finding a minimum or maximum of a function to the problem of finding a zero.

6.1.1. Bisection Algorithm. Let x^* be a root, or zero, of $g(x)$, i.e., $g(x^*) = 0$. If an initial bracket $[a, b]$ is known such that $x^* \in [a, b]$, then a simple and robust approach to determining the root is to bisect this interval into two intervals $[a, c]$ and $[c, b]$ where c is the midpoint, i.e.,

$$c = \frac{a + b}{2}$$

If

$$g(a)g(c) < 0$$

then we conclude

$$x^* \in [a, c]$$

while if

$$g(b)g(c) < 0$$

then we conclude

$$x^* \in [b, c]$$

This process may now be iterated such that the size of the bracket (as well as the actual error of the estimate) is being divided by 2 every iteration.

6.1.2. Newton's Method. Note that in the bisection method the actual value of the function $g(x)$ was only being used to determine the correct bracket for the root. Root finding is accelerated considerably by using this function information more effectively.

For example, image we were seeking the root of a function that was a straight line, i.e., $g(x) = ax + b$ and our initial guess for the root was x_0. If we extend this straight line from the point x_0 it is easy to determine where it crosses the axis, i.e.,

$$ax_1 + b = 0$$

so $x_1 = -b/a$. Of course, if the function were truly linear then no first guess would be required. So now consider the case that $g(x)$ is nonlinear but may be approximated locally about the point x_0 by a line. Then the point of intersection of this line with the x-axis is an estimate, or second guess, for the root x^*. The linear approximation comes from Taylor's theorem, i.e.,

$$g(x) = g(x_0) + g'(x_0)(x - x_0) + \frac{1}{2}g''(x_0)(x - x_0)^2 + \dots$$

So the linear approximation to $g(x)$ about the point x_0 can be written

$$l(x) = g(x_0) + g'(x_0)(x - x_0)$$

If we take x_1 to be the root of the linear approximation we have

$$l(x_1) = 0 = g(x_0) + g'(x_0)(x_1 - x_0)$$

Solving for x_1 gives

$$x_1 = x_0 - \frac{g(x_0)}{g'(x_0)}$$

or at the nth iteration

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}$$

The iteration above is for determining a zero of a function $g(x)$. To determine a maximum or minimum value of a function f we employ condition that $f'(x) = 0$. Now the iteration is modified as as

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

6.2. Unconstrained Optimization in Higher Dimensions

Now we consider the problem of minimizing (or maximizing) a scalar function of many variables, i.e., defined on a vector field. We consider the extension of Newton's method presented in the previous section as well as a classical approach known as steepest descent.

6.2.1. Taylor Series in Higher Dimensions. Before we extend the search for extrema to higher dimensions we present Taylor series for functions of more than one domain variable. To begin, the Taylor series for a function of two variables is given by

$$\begin{aligned} g(x,y) =& g(x^{(0)}, y^{(0)}) + \frac{\partial g}{\partial x}(x - x^{(0)}) + \frac{\partial g}{\partial y}(y - y^{(0)}) \\ &+ \frac{\partial^2 g}{\partial x^2}\frac{(x - x^{(0)})^2}{2} + \frac{\partial^2 g}{\partial y^2}\frac{(y - y^{(0)})^2}{2} + \frac{\partial^2 g}{\partial x \partial y}(x - x^{(0)})(y - y^{(0)}) \\ &+ \text{higher order terms} \end{aligned}$$

In n variables $x = (x_1, \dots, x_n)$ the Taylor series expansion becomes

$$g(x) = g(x^{(0)}) + \nabla g(x^{(0)})(x - x^{(0)}) + \frac{1}{2}(x - x^{(0)})^T Hg(x^{(0)})(x - x^{(0)}) + \cdots$$

where the Hessian matrix is defined as

$$\left(Hg(x^{(0)})\right)_{ij} = \left.\frac{\partial^2 g}{\partial x_i \partial x_j}\right|_{x = x^{(0)}}$$

and the gradient is written as a row vector, i.e.,

$$\big(\nabla g(x^{(0)})\big)_i = \left.\frac{\partial g}{\partial x_i}\right|_{x=x^{(0)}}$$

6.2.2. Roots of a Nonlinear System. We saw that Newton's method could be used to develop an iteration for determining the zeros of a scalar function. We can extend those ideas for determining roots of the nonlinear system

$$\begin{aligned} g_1(x_1,\ldots,x_n) &= 0\\ g_2(x_1,\ldots,x_n) &= 0\\ &\vdots\\ g_n(x_1,\ldots,x_n) &= 0 \end{aligned}$$

or, more compactly,

$$g(x) = 0.$$

Now we apply Taylor's theorem to each component $g_i(x_1,\ldots,x_n)$ individually, i.e., retaining only the first order terms we have the linear approximation to g_i about the point $x^{(0)}$ as

$$l_i(x) = g_i(x^{(0)}) + \nabla g_i(x^{(0)})(x - x^{(0)})$$

for $i = 1,\ldots,n$. We can write these components together as a vector equation

$$l(x) = g(x^{(0)}) + \nabla g(x^{(0)})(x - x^{(0)})$$

where now

$$\big(\nabla g(x^{(0)})\big) = \left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^{(0)}}$$

As in the scalar case we base our iteration on the assumption that

$$l(x^{(k+1)}) = 0$$

Hence,

$$g(x^{(k)}) + \nabla g(x^{(k)})(x^{(k+1)} - x^{(k)}) = 0$$

and given $x^{(k)}$ we may determine the next iterate $x^{(k+1)}$ by solving an $n \times n$ system of equations.

6.2.3. Newton's Method. In this chapter we are interested in minimizing functions of several variables. Analogously with the scalar variable case we may modify the above root finding method to determine maximum (or minima) of a function $f(x_1,\ldots,x_n)$. To compute an extreme point we require that $\nabla f = 0$ hence we set

$$g(x) = \big(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n}\big)^T$$

Substituting

$$g_i(x) = \frac{\partial f}{\partial x_i}$$

into

$$g(x^{(k)}) + \nabla g(x^{(k)})(x^{(k+1)} - x^{(k)}) = 0$$

produces

$$\nabla f(x^{(k)}) + Hf(x^{(k)})(x^{(k+1)} - x^{(k)}) = 0$$

where

$$\big(Hf(x^{(k)})\big)_{ij} = \frac{\partial g_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Again we have a linear system for $x^{(k+1)}$.

6.2.4. Steepest Descent. Another form for Taylor's formula in n-variables is given by

$$f(x+th) = f(x) + t\nabla f(x)h + \frac{t^2}{2}h^T Hf(x)h$$

where again $(\nabla f(x))_i = \partial f/\partial x_i$ and $(Hf(x))_{ij} = \partial^2 f/\partial x_i \partial x_j$. Now t is a scalar and $x+th$ is a ray emanating from the point x in the direction h. We can compute the derivative of the function f w.r.t. t as

$$\frac{df}{dt}(x+th) = \nabla f(x)h + th^T Hf(x)f$$

Evaluating the derivative at the point $t = 0$ gives

$$\frac{df}{dt}(x) = \nabla f(x)h$$

This quantity, known as the directional derivative, indicates how the function is changing at the point x in the direction h. Recall from calculus that the direction of maximum increase (decrease) of a function is in the direction of the gradient (negative gradient). This is readily seen from the formula for the directional derivative using the identity

$$\nabla f(x)h = \|\nabla f(x)\|\|h\|\cos(\theta)$$

We can assume without loss of generality that h is of unit length, i.e., $\|h\| = 1$. So the quantity on the right is a maximum when the vectors h and $\nabla f(x)$ point in the same direction so $\theta = 0$.

This observation may be used to develop an algorithm for unconstrained function minimization. With an appropriate choice of the scalar step-size α, the iterations

$$x^{(k+1)} = x^{(k)} - \alpha\nabla f(x^{(k)}) \tag{6.1}$$

will converge (possibly slowly) to a minimum of the function $f(x)$.

6.3. Constrained Optimization and Lagrange Multipliers

Consider the constrained minimization problem

$$\min f(x)$$

subject to

$$c_i(x) = 0$$

$i = i, \ldots, p$. It can be shown that a necessary condition for a solution to this problem is provided by solving

$$\nabla f = \lambda_1 \nabla c_1 + \cdots + \lambda_p \nabla c_p$$

where the λ_i are referred to as Lagrange multipliers. Consider the case of f, c being functions of two variables and consider their level curves. A little thought reveals that the minimum value of f on the constraint c is given when the gradient of each function is parallel.

We demonstrate a solution via this procedure by recalling our earlier example.

EXAMPLE 6.4. Given a fixed area of cardboard A construct a box of maximum volume. The nonlinear program for this is

$$\min xyz$$

subject to

$$2xy + 2xz + 2yz = A$$

Now $f(x, y, z) = xyz$ and $c(x, y, z) = 2xy + 2yz + 2xz - A$. Substituting these functions into our condition gives

$$\nabla f = \lambda \nabla c$$

produces the system of equations

$$\begin{aligned} yz - \lambda(2y + 2z) &= 0 \\ xz - \lambda(2x + 2z) &= 0 \\ xy - \lambda(2y + 2x) &= 0 \end{aligned}$$

The solution to this system may be computed analytically to be

$$x = y = z = \sqrt{\frac{A}{6}}$$

In this special case the nonlinear system could be solved by hand. Typically this is not the case and one must resort to numerical techniques such as Newton's method to solve the resulting $(n + m) \times (n + m)$ system

$$g(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = 0.$$

PROBLEM 6.1. Extend Example 6.2 for a collection of S schools.

PROBLEM 6.2. Show how Newton's method for root finding can be used to calculate $\sqrt{3}$. Use Matlab to compute the sequence that converges to this value. What is the effect of changing the initial condition?

PROBLEM 6.3. Minimize the objective function

$$f(x_1, x_2) = 7x_1^2 + 2x_1x_2 + x_2^2$$

using 50 iterations of

a) Newton's method
b) Steepest Descent

Plot the value of the iterates for each method on the same graph. You may experiment with the value of α in Equation (6.1). Hint start small.

PROBLEM 6.4. Assume a farmer has L feet of fencing for a rectangular area with lengths x and y. Determine these lengths such that the enclosed area is a maximum.

CHAPTER 7

Simulation Modeling

It is not unusual that the complexity of a phenomenon or system makes a direct mathematical attack time-consuming, or worse, intractable. An alternative modeling approach consists of the literal execution of rules by the computer. Such *simulation* approaches occur in great variety but share the common feature that a computer is the central vehicle for knowledge, or process discovery. Here are some problems where simulation modeling appears to be successful, if not indispensible:

- Develop strategies in games with simple rules but stochastic components such as blackjack, checkers, and solitaire.
- Numerically approximate solutions to complex models such as the Navier-Stokes equations of fluid dynamics. (These simulations are typically deterministic).
- Modeling phenomenon with inherent probabilistic components such as processing queues, traffic problems, and inventory problems.

This above list, while certainly not complete, gives an indication of the range of problems that lend themselves naturally to simulation modeling.

7.1. The Tire Distributor Problem

Consider the inventory problem confronting the distributor of a commodity with random demand. To be concrete we will consider the situation of a car tire distributor. Our assumption is that this distributor supplies tires to a large number of clients and in turn purchases its supply of tires from a major tire producer. The distributor itself does not fabricate tires but relies on deliveries from the factory. Based on observed daily demand it is up to the distributor to determine a quantity X of tires to be delivered at an integer interval N in days.

The problem is driven by costs. It is the desire of the distributor to select an X and N to minimize total costs. These costs are assumed to consist of two components:

- tire delivery costs
- interest costs on money borrowed to pay for tires in stock

The tire delivery costs may be modeled in several different ways. We will assume that the factory delivers and charges by the truckload and each truckload can contain up to 1,000 tires. Furthermore, the delivery charges for a truckload with no discount if the truck is not full. Hence, if a truckload costs $\$\alpha$, then the delivery costs are $d(X)$ are

$$d(X) = \alpha\left(\left[\frac{X-1}{1000}\right] + 1\right)$$

where the notation $[a]$ means the the number a rounded down to the nearest integer. So, for example,

$$\left[\frac{1500}{1000}\right] = 1$$

The second cost is due to the interest charged by the bank for the load the company has to purchase the tires. Note that even if the distributor had enough cash flow to not need to borrow, we would still view the investment in the tires sitting in the warehouse as a cost as this money could presumably be invested to earn a dividend.

Let's assume the interest rate per day per tire is β. If x_n tires are in stock at the end of day n then we are assessed the interest fee βx_n so that the total cost of interest over an interval N is

$$I = \sum_{n=1}^{N} \beta x_n$$

This problem is complicated by the fact that the number of tires purchased by customers varies randomly from day to day as shown in Table 1. Furthermore, the average demand was calculated to be 997 tires/day. Thus, the distributor must attempt to simulate this statistical demand in the model. To accomplish this we need to develop a demand subroutine that will mimic the observed daily demand. This may be achieved by associating the demand intervals with segments of the unit interval the length of which is determined by the actual frequence of demand. So, for example, the demand interval $0 \leq x_n < 100$ occurs 12 days out of 365 (presumably due to holidays). So, we associate the interval

$$I_1 = [0, 12/365]$$

to the probability that the daily demand will be in the interval $0 \leq x_n < 100$ tires. Since we are mapping the interval [0,1] to daily demand ranges we must require that these subintervals be nonoverlapping. Thus, the fact that the interval $100 \leq x_n < 300$ occurred on exactly 4 days means that we should reserve 4/365 of the unit interval for this demand range, i.e., $I_2 = [12/365, 16/365]$. We may develop the rest of the intervals in a similar fashion, i.e.,

$$I_3 = [16/365, 43/365]$$

$$I_4 = [43/365, 86/365]$$

and so on.

Now we have partitioned the unit interval, i.e.,

$$[0, 1] = I_1 \cup I_2 \cup \cdots \cup I_{11}$$

Thus, if we pick a random number $z \in [0, 1]$ (always picking this number uniformly from the interval) we can map that number to an appropriate daily demand interval.

For example, if our uniform random number generator returns $z = .0768$ then we select the interval containing this point, i.e., I_3. Now the question remains how to pick an actual daily demand? If we are in interval I_3 we only know that the demand should be between 300 and 500 tires. So we may select this demand randomly from the integers in this interval. This model we have constructed for simulating the demand can be simply tested by running the model for 100 years and seeing if we reproduce the demands (now averaged over the 100 years). The

results of doing this, shown in the last column of table 1 suggest this model is rather good.

daily demand	frequency in days	cumulative distribution	simulated freq.
$0 \leq x_n < 100$	12	12	12.22
$100 \leq x_n < 300$	4	16	4.04
$300 \leq x_n < 500$	27	43	26.67
$500 \leq x_n < 700$	43	86	43.05
$700 \leq x_n < 900$	48	134	47.20
$900 \leq x_n < 1100$	67	201	66.31
$1100 \leq x_n < 1300$	78	279	78.92
$1300 \leq x_n < 1500$	55	334	55.77
$1500 \leq x_n < 1700$	22	356	21.56
$1700 \leq x_n < 190$	7	363	7.24
$1900 \leq x_n < 2100$	2	365	2.02

TABLE 1. Number of days certain quantities of tires were demanded. Total number of days of collected data is 365.

Summary of Tire Distributor Simulation

We run the simulation for 365 days and compute the average daily cost. This calculation is repeated 100 times and the average cost is now averaged again. This produces a more reliable stochastic estimate of the cost.

- Make an initial delivery of tires of size deliver_quantity.
- Compute the stochastic demand and subtract sales from stock NUM_TIRES
- If the stock is zero then add a penalty
- Accrue interest costs every iteration.
- If delivery_interval counter indicates delivery then increment NUM_TIRES by delivery_quantity.

7.2. Blackjack Strategy

Blackjack is a poker game pitting the Dealer, or Bank, against one or more players. For simplicity we will assume there is only one player, as in video blackjack. The object of the game is to score higher than the dealer without going over 21 points. If this occurs the player is paid the value of his bet. If the Dealer has a higher score than the player the Dealer collects the bet. Ties result in no loss of bet, or a *push*.

Card Values. The 10, Jack, Queen, and King are all valued at 10 while the cards from 2 through 9 are valued as indicated. The score of a hand is obtained by adding the values of the individual cards in the hand. The Ace is worth 1 or 11 points. The player can choose the value of the Ace whereas the dealer must always take the high value unless the total is over 21.

Blackjack occurs when a total of 21 is obtained with the first two cards in a hand (a black Jack is not necessary). If the Dealer has Blackjack and the Player has 21, the Dealer wins. If the Dealer has 21 and the Player has Blackjack, the Player wins. If a Player wins with a Blackjack then he is paid 1.5 times the bet placed. A total of 21 with more than two cards is not black jack.

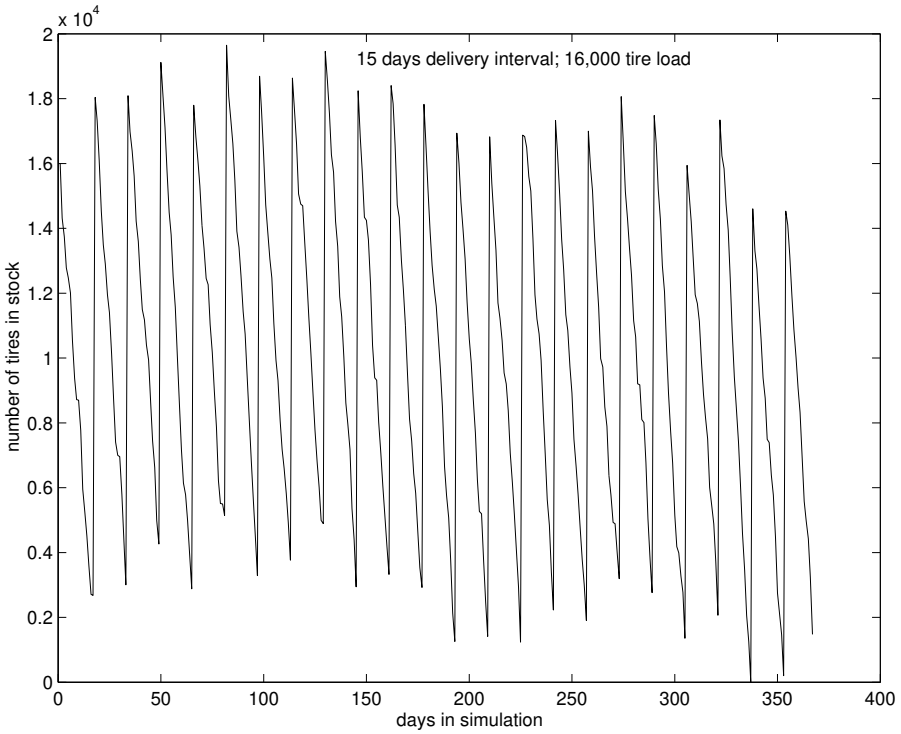


FIGURE 1. The average daily cost for this simulation averaged over 365 days was $228. This simulation has no days without tires and does not tend to accumulate tires that would be subject to interest.

Rules of Play. The dealer distributes two cards to the Player and to himself. The Dealer shows the value of one of his cards to the player. The Player then requests cards one at a time until he decides to stay pat (receive no more cards) or goes bust (exceeds 21). If the Player stands pat then the Dealer plays. The Dealer has no choices in how he plays his cards but must follow a specific set of rules.

Dealer Play. Dealer must stand pat on 17. Dealer must continue to take cards until his total is 17 or higher An Ace in the dealer's hand is always counted as 11 unless counting it as 1 prevents the dealer from going over. Thus if the dealer holds (ace,7) then he must stand. If the dealer holds (10,6,Ace) where the Ace is the third card picked then the dealer holds 17 and must stand.

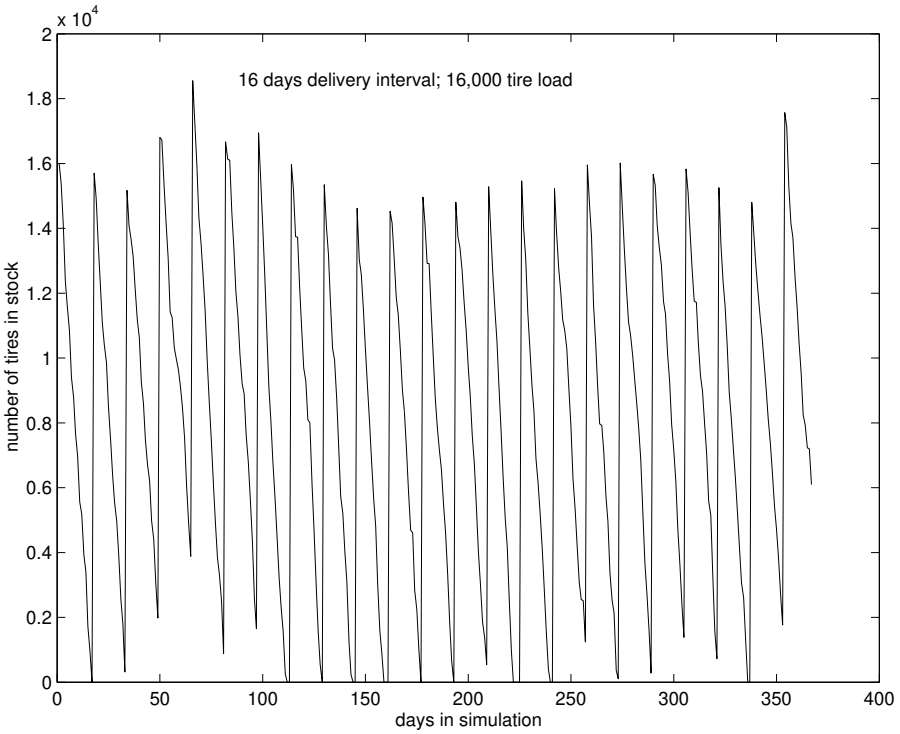


FIGURE 2. The average daily cost for this simulation averaged over 365 days was \$703. The increased cost is due to number of days (19) where the stock went to zero.

PROBLEM 7.1. The tire factory that supplies our distributor is overstocked and has decided to discount shipping such that the first truck costs the usual amount $\alpha = \$400$, the second truck costs $\$\alpha/2$, the third truck costs $\$\alpha/3$ and so on. Modify the tire distributor problem to account for this and determine new optimal values of delivery_amount and delivery_interval. Are your answers what you would expect?

PROBLEM 7.2. Modify the ordering of a new shipment of tires to occur when the stock is at 20% of the delivery load. How does this effect the optimal delivery amount and interval? Does the smallest average daily cost go down?

PROBLEM 7.3. Modify your code so that the number of tires in stock can't excede $T_{\max} = 17,000$ tires. Include graphs of your new simulations to demonstrate this limited capacity. What is the new optimal delivery_interval and delivery_quantity?

PROBLEM 7.4. Modify the ordering of a new shipment of tires to occur when the stock is at 20% of the delivery load. How does this effect the optimal delivery amount and interval? Does the smallest average daily cost go down?

PROBLEM 7.5. Now assume that the placement of orders for new shipments of tires have a random component based on the routines of the three rotating managers. Assume manager I works 30% of the shifts and that he will place the

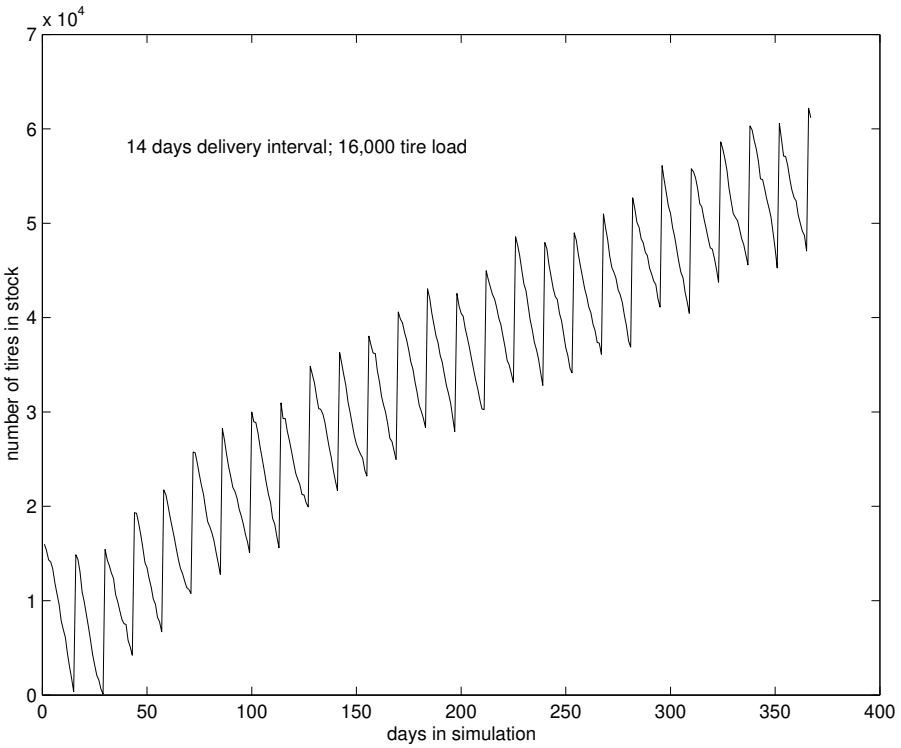


FIGURE 3. The daily cost for this simulation averaged over 365 days was $447. This increased cost is apparently due to the interest costs on the increasing stock.

order when stock dips under 35% *of the delivery_quantity*; manager II works 50 % of the shifts and he places orders will when the stock dips below 20% of the delivery_quantity; manager III works 20% of the shifts and he places orders when the stock is just under 5% of the delivery_quantity. Modify your code to account for this. What is the new optimal deliver_quantity? (Note: there is now no delivery_interval.) Hand in your modified code for grading.

PROBLEM 7.6. In the simulation code provided the days without tires are all charged at the same flat rate. Modify the code to account for the fact that some tires may have been sold before the stock ran out. For example, if there were 700 tires at the beginning of the day and 800 were sold the penalty for not filling 100 orders should be smaller than if the whole day were spent without tires.

PROBLEM 7.7. Adapt the Blackjack computer program in the Appendices to exploit knowledge of the Dealer's showing card when deciding when to stop getting new cards. Run the simulation 1000 times (i.e., 2000 decks of cards) picking my_stay_value based on the value of the Dealer's face card and compare the results. Compare your winnings with the provided code that does not consider the Dealer's face card. You should win a higher percentage after the modification.

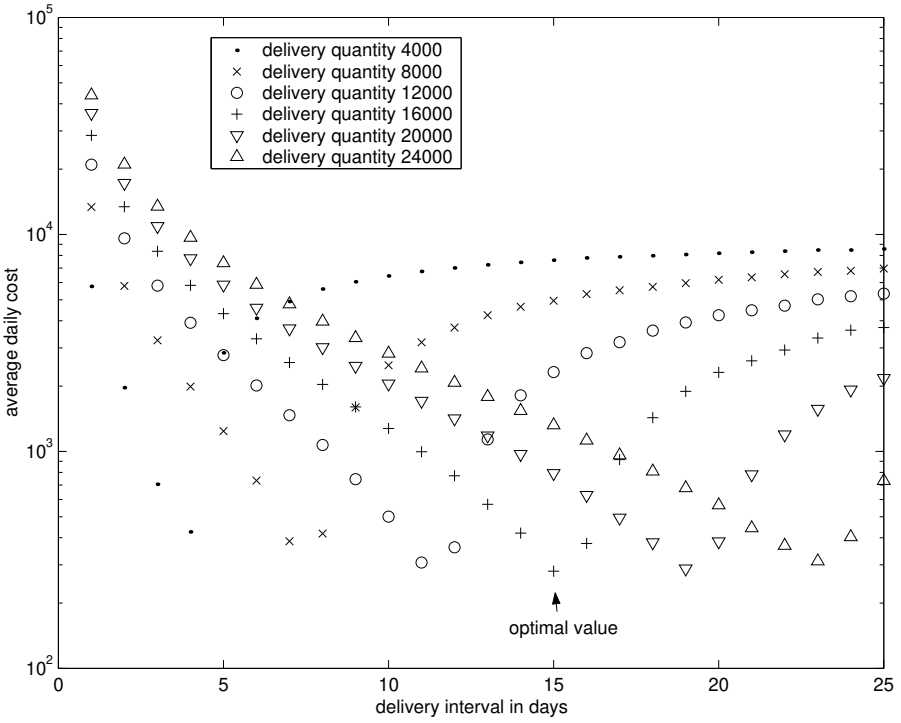


FIGURE 4. Average daily cost for a range of delivery intervals and delivery quantities. The minumum costs are themselves averaged over 365 day cycles and in turn these cycles are rerun 100 times and the result of the minimum daily cost averaged. The minimum average cost in these curves is (rounded to the nearest dollar) \$415, \$416, \$354, \$331, \$340 and \$341. A deliver interval of 15 days and quanity 16,000 tires appears optimal from our simulations.

PROBLEM 7.8. Experiment with the player's strategy for deciding whether to select an ace as a one or eleven based on the Dealer's showing card. Can you improve over your results in the previous problem?

CHAPTER 8

Final Course Projects

The goal of the final project is to allow you time to develop and study a mathematical model in greater depth. These projects will be done in groups of 3-4 and a single grade will be awarded to each group. It is the responsibility of each group to make sure that all group members participate and contribute to the project. It is permitted to partition the work among the group members but each group member is responsible to read and evaluate the work produced by their group. Again, there is only one grade for the project.

The grade assessment of the final project will emphasize presentation. Your group should spend a signifant fraction of its efforts on organizing your results and presenting them in a way that clearly demonstrated the utility of your work. All reports must be typed and graphs imported into the text. Careful labeling of graphs is important. Do not include more graphs than necessary to make your point. Include all MATLAB code as an appendix.

Given the different nature of each project there are many appropriate ways to write up your results. Consider the summary below as a general guideline.

General structure of the project write-up

(1) Introduction and General Discussion: why is the problem interesting? What do you hope to learn by employing a mathematical model. What are the goals of the investigation?
(2) Presentation of Mathematical Model including assumptions. Discussion of algorithm in code.
(3) Detailed description of results including numerics, graphs and interpretations.
(4) Process Discovery:
    (a) What did you learn? what was the added value of the mathematical model? Here the emphasis is on demonstrating that you discovered something that you did not know before implementing your model.
    (b) Sources of errors in the model. How believable are the results?
(5) Conclusions and general discussion:
    (a) Concise summary of what was discovered.
    (b) Summary of weaknesses and strengths of approach.
    (c) How might the model be inproved?

8.1. Quarterback Passer Ratings

NFL passer ratings are based on the formula

$$x_1 = \text{completion percentage}$$
$$x_2 = \text{yards/attempts}$$
$$x_3 = \text{touchdown percentage}$$
$$x_4 = \text{interception percentage}$$
$$r(x_1, x_2, x_3, x_4) = c_0 + c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$

Apparently, the actual values of the constants $c_0, \ldots, c_4$ were kept secret by the NFL until some clever mathematical modelers showed how they could be calculated.[1]

a) Using the odd numbered quartbacks in the list compute the constants c_i for $i = 0, \ldots, 4$ and hence determine the NFL quarterback rating system using the method of least squares. Predict the even numbered quarterbacks ratings predicted by your formula. Now calculate errors between the formula and the raw data given in the table. Discuss these errors carefully.

b) Argue that the raw data
   - total yards
   - total completions
   - total touchdowns
   - total interceptions

   could be used to obtain an equivalent formula.

c) It has been observed that most passers would improve their rating by completing an extra pass for no gain. Do you believe this? How much do the ratings of Montana, Elway and Vick's change if they throw one more pass for no gain? How many passes can Elway throw for no gain and still improve his rating? Answer the same question for Montana. Can you answer this last question analytically? What recommendations might you make to a quarterback as good ways to improve a rating and why?

d) Assuming the NCAA uses the same formula but different weighting constants, apply the same least squares approach to compute the NCAA passer rating formula. Estimate Griese and Manning's NCAA passer rating. Who are the top five rated quarterbacks in the NFL using the NCAA formula on the NFL data?

e) Johny Unitas, Bart Starr, Fran Tarkenton and Joe Namath had careers before the new NFL system was introduced in 1973. Estimate their passer ratings using the NFL formula you have found. Where do they rank amongst modern QBs?

f) Apply the NFL rating system to the NCAA numbers of Joey Harrington, Major Applewhite and David Carr. Comment.

g) How would you modify the formula to extend the passer rating to a more general quarterback rating? Do the top five quarterbacks change with the your new ratings system?

[1]This project is an extension of the UMAP module 765 "How Does the NFL Rate Passers", by Roger Johnson, 1997. Additional data has been added and updated.

NFL PASSING STATISTICS	ATT	COMP	YDS	TDS	INTS	RATING
Warner, Kurt	1581	1063	13864	101	62	99.6*
Young, Steve	4065	2622	32678	229	103	97.6
Montana, Joe	5391	3409	40551	273	139	92.3
Marino, Dan	7989	4763	58913	408	235	87.3
Favre, Brett	5820	3550	41363	307	183	87.0*
Graham, Otto	2626	1464	23584	174	135	86.6
Manning, Peyton	2639	1638	19343	130	94	85.9
Brunell, Mark	3473	2089	24753	136	85	84.8*
Griese, Brian	1600	991	11152	69	49	84.4*
Kelly, Jim	4779	2874	35467	237	175	84.4
Staubach, Roger	2958	1685	22700	153	109	83.4
Johnson, Brad	2380	1466	16379	92	68	83.07
Gannon, Rich	3295	1949	22256	145	88	83.06
Aikman, Troy	4011	2479	28346	141	115	82.8
Lomax, Neil	3153	1817	22771	136	90	82.7
Jurgensen, Sonny	4262	2433	32224	255	189	82.6
Dawson, Len	3741	2136	28711	239	183	82.6
Anderson, Ken	4475	2654	32838	197	160	81.9
Vick, Michael	365	200	2606	11	5	81.8*
Kosar, Bernie	3365	1994	23301	124	87	81.8
White, Danny	2950	1761	21959	155	132	81.7
O'Donnell, Neil	2862	1650	19026	104	57	81.6
Cunningham, Randall	3875	2177	27082	190	119	81.6
Krieg, Dave	5311	3105	38147	261	199	81.5
Esiason, Boomer	5205	2969	37920	247	184	81.1
Moon, Warren	6786	3972	49097	290	232	81.0
Beuerlein, Steve	3174	1810	22932	142	104	81.0*
Chandler, Chris	2587	1494	18526	119	90	80.9
Elway, John	7250	4123	51475	300	226	79.9
Bledsoe, Drew	4950	2820	32865	184	147	77.2*
Carr, David	294	158	1900	9	11	68.4*
Harrington, Joey	347	173	1932	11	13	61.8*
Unitas, Johny	5186	2830	40239	290	253	
Namath, Joe	3762	1886	27663	173	220	
Tarkenton, Fran	6467	3686	47003	342	266	
Starr, Bart	3149	1808	24718	152	138	

NCAA	ATT	COM	INT	YRDS	TDS	NCAA RATING
McMahon, Jim	1060	653	34	9536	84	156.9
Testaverde, Vinny	674	413	25	6058	48	152.9
Young, Steve	908	592	33	7733	56	149.8
Aikman, Troy	637	401	18	5436	40	149.7
Bosco, Robbie	997	638	36	8400	66	149.4
Hartlieb, Chuck	716	461	17	6269	34	148.9
White, Danny	649	345	36	5932	59	148.9
Long, Chuck	1072	692	46	9210	64	147.8

Ware, Andre	1074	660	28	8202	75	143.4
Kosar, Bernie	743	463	29	5971	40	139.8
Elway, John	1246	774	39	9349	77	139.3
Theismann, Joe	509	290	35	4411	31	136.1
Davis, Steve	209	83	17	1973	21	135.9
Fusina, Chuck	664	371	32	5382	37	132.7
Jones, Bert	418	221	16	3255	28	132.7
Everett, Jim	923	550	30	7158	40	132.5
Flutie, Doug	1270	677	54	10579	67	132.2
Marino, Dan	1084	626	64	7905	74	129.7
Plunkett, Jim	962	530	47	7544	52	129.0
Montana, Joe	515	268	25	4121	25	127.3
Esiason, Boomer	850	461	27	6259	42	126.1
Blackledge, Todd	658	341	41	4812	41	121.4
Kelly, Jim	676	376	28	5228	32	127.9
Manning, Archie	761	402	40	4753	31	108.2
Kramer, Tommy	1036	507	52	6197	37	100.9
Dickey, Lynn	994	501	65	6208	29	99.41
Griese, Brian	606	355	18	4383	33	??
Manning, Peyton	1354	851	33	11201	90	??

2002 Draft	ATT	COM	YRDS	TDS	INT
Applewhite, Major	1065	611	8353	60	28
Carr, David	934	587	7849	70	23
Harrington, Joey	928	512	6911	59	23

*still active in NFL

8.2. Baseball Simulation

In this project pick a closely contested world series and resimulate the best of seven series from 1,000 to 1,000,000 times to determine if the "better team" won. To make your simulation as realistic as possible you will want to collect the actual year statistics for each starting player (it is okay to omit pitching statistics) including percentage

- singles hit (over season)
- doubles hit
- triples hit
- homeruns
- walks
- strikeouts

It will be more realistic if you also include hit by pitches.

You make modify the code below for your simulation. It currently is designed for a single player hiotting 2% triples, 4% doubles etc... Extend this code to accomodate 9 players on each team. Follow the rules of baseball and give each team 27 outs. Compare the scores of each team after 27 outs to determine the winner. Play extra innings if necessary to determine a winner of a game. Play a best of seven series and determine the winner. Repeat to assess who would win the higher percentage of seven game series.

Does changing the lineup order for either team impact these results significantly?

Discuss the utility of your results for predicting winners of sporting events via simulation.

8.2.1. Basic Matlab code for baseball simulation.

```
%simulate the top of an inning

%assume identical batters
triple = 0
double = 0
single = 0
homerun = 0
outs = 0
runs = 0
base1 = 0 %set to one of player on first base, else 0
base2 = 0
base3 = 0

while outs < 3

%what is the result of the at bat?
x = rand(1);
if x <= .02
   triple = triple + 1;
   if base1 == 1
      runs = runs + 1;
      base1=0;%reset first base to be empty
```

```
   end
   if base2 == 1
      runs = runs + 1;
      base2=0%/reset second base to be empty
   end
   if base3 == 1
      runs = runs +1;
      %third base has new runner so don't reset
   end

elseif x > .02 & x <=.06
   double = double + 1;
   if base3 == 1
      runs = runs + 1;
      base3=0;%reset base to be empty
   end
   if base2 == 1
      runs = runs + 1;
      %/second base has new runner
   end
   if base1 == 1
      base3 = 1; %reset third base to be empty
      base1 = 0;%send player on first to third
   end


elseif x > .06 & x <=.16
   homerun = homerun + 1;
   runs = runs + 1
   if base1 == 1
      runs = runs + 1;
      base1=0;%reset first base to be empty
   end
   if base2 == 1
      runs = runs + 1;
      base2=0%/reset second base to be empty
   end
   if base3 == 1
      runs = runs +1;
      base3 = 0; %reset third base to be empty
   end
elseif x > .16 & x <=.28
    single = single +1;%note this includes walks and hit-by-pitch
    if base3 == 1
      runs = runs + 1;
      base3=0;%reset first base to be empty
   end
   if base2 == 1
```

```
     base3=1;
      base2=0;%/reset second base to be empty
   end
   if base1 == 1
      base2 = 1; %reset third base to be empty
   end
   base1 = 1; %send player to first
else
   outs = outs + 1;
end

end%with while
single
double
triple
homerun
outs
runs
```

8.3. Predicting the Stock Market

This problem involves the application least squares and a radial basis function network to the prediction of exchange rate time-series. Employ a time lag of the data such that your RBF model is a mapping $\tilde{f} : U \in \mathbb{R}^3 \to \mathbb{R}$, i.e., approximate the function

$$x_{n+1} = f(x_n, x_{n-1}, x_{n-2})$$

where x_n is the value of the provided time series at time n. Specifically,

$$\tilde{f}(x) = w_0 + \sum_{m=1}^{N} w_m \phi(\|x - c_m\|)$$

where the function ϕ is the the radial basis function. You may select

$$\phi(r) = r^3$$

or

$$\phi(r) = \exp(-r^2)$$

By applying the interpolation condition

$$y^{(i)} = \tilde{f}(x^{(i)})$$

show that the weights w_m may be found by solving a least squares problem of the form

$$y = \Phi w$$

Select the number N of cluster centers c_m to vary from 10 to 100 using random subset selection of the data. Build your RBF model on the first one hundred points using any seast squares approach to solve for the weights, e.g., the MATLAB backslash routine. Test your results for each set of centers by attempting to predict the next 100 values using both

(1) One-step prediction of x_{n+1} where all actual values $x_n, x_{n-1}, \ldots$ known exactly, i.e.,

$$\tilde{x}_{n+1} = f(x_n, x_{n-1}, x_{n-2})$$

(2) Iterated prediction of x_{n+1} where the output is used to predict future values via

$$\tilde{x}_{n+1} = f(\tilde{x}_n, \tilde{x}_{n-1}, \tilde{x}_{n-2})$$

Calculate the means-square error in each case. You may find it interesting to see how your results change using more data and more centers.

8.4. Population Biology

Develop a system of difference equations for the interacting species A, B and C where

(1) Species A is a predator for species B and C (A eats B and C)
(2) Species B is a predator for species C
(3) Species C is not a predator for A or B
(4) Species A has no food supply besides species B and C
(5) Species B has no food supply besides species C
(6) Species C has a plentiful food supply (not species A or B)

Include all the constants of proportionality and describe their importance. Pick various values of these constants and simulate the populations. Can you find any stable or unstable equilibrium? Can you find any period solutions? Balance your numerical arguments with analytical ones, e.g., compute the eigenvalues of the Jacobian to test for asymptotic stability or instability. You may use your Newton's method (or steepest descent) code to find equilibria.

How will the model change if all species suddenly become predators for all other species? How does this impact your solutions? Can you think of any systems that might be described by this model?

APPENDIX A

Matlab Code for Data Fitting

A.1. Mammalian Heart Rate Problem

```
File: ls_mammals.m
--------------------Start of actual file contents-----------------------
%LEAST SQUARES ANALYSIS OF MAMMALIAN HEART RATE

%w body weights
%r corresponding heart rates
%data is a row vector
w = [3.5 4 6 25 103 117 200 252 300 437 1340 2000 2700 5000
     22500 30000 33000 50000  70000 100000 415000 450000 500000 3000000];
r = [787 660 588 670 347 300 420 352 300 269 251 205 187 120 100
     85 81 70 72 70 45 38 40 48];

x1 = w.^(-1/3) %raise each component of vector w to -1/3 power.
x2 = w.^(-2/3) %the result of this operation is a vector the same size as w
```

Now calculate the slope given by the formula:

$$k = \frac{\sum_{i=1}^{P} r_i w_i^{-1/3}}{\sum_{i=1}^{P} w_i^{-2/3}}$$

We will use the variables numerator and denominator to split up the calcution in the obvious fashion. The numerator is expressed as the vector dot product of r, the row vector of heart rates, and x1 as found above.

```
numerator = r*x1';%apply transpose operator ' to x1 to compute dot product.
denominator = sum(x2);%compute the sum of each component
k1 = numerator/denominator;
```

Now we reproduce the above calculation reproducing all the steps but by using a different data set to compute the slope. It would be more effiecient to pass the data to a subroutine rather than repeat all the code. We examine this in the next section.

```
%%Now build the model on the first 2/3 of the data (16 points)
ws = w(8:24)%the notation 8:24 is equivalent to [8 9 10 11 12 .... 24]
rs = r(8:24)

x1 = ws.^(-1/3)%raise each component of vector w to -1/3 power.
```

```
x2 = ws.^(-2/3)

numerator = rs*x1';%apply transpose operator ' to x1 to compute dot product.
denominator = sum(x2);%compute the sum of each component
k2 = numerator/denominator;%see formula in section 3.1.1

hold on
plot(w.^(-1/3),r,'o')%plot raw data
plot(w.^(-1/3),k1*w.^(-1/3),'--x')%plot first model
plot(w.^(-1/3),k2*w.^(-1/3),'--v')%plot second model
title('mammalian heart rate model')
xlabel('weight w^{(-1/3)}')
ylabel('pulse rate')
legend('raw data','least squares fit (all data)', 'least squares fit 2/3 data')
-----------------------End of actual file contents--------------------------------
```

A.2. Least Squares with Normal Equations

This program consists of two parts: a subroutine called ls_normal.m and a driver called run_ls.m.

```
File:  ls_normal.m
----------------------Start of actual file contents------------------------
%Input:
%   x is a column vector of domain (input) variables
%   y is a column vector of range (output) variables
%
%Output:
%   m is the slope of the line
%   b is the intercept of the line

function [m,b] = ls_normal (x,y)

P = size(x,1)%how many points are there in this column vector?
```

Now we compute the terms required in the evaluation of m and b in the normal equations. Recall

$$m = \frac{(\sum y_i)(\sum x_i) - P\sum y_i x_i}{(\sum x_i)^2 - P\sum x_i^2}$$

$$b = \frac{-(\sum y_i)(\sum x_i^2) + (\sum x_i)(\sum y_i x_i)}{(\sum x_i)^2 - P\sum x_i^2}$$

We set sy $= \sum y_i$, dp_xy $= \sum y_i x_i$ and x_sq $= \sum x_i^2$.

```
sy = sum(y);%sum y_i (scalar)
sx = sum(x);%sum x_i (scalar)
dp_xy = x'*y; %y dot product x (scalar)
x_sq = sum(x.*x);%sum x_i^2 term (scalar)
denom = sx^2 - P*x_sq; %(scalar)

m = (sy*sx-P*dp_xy)/denom
b = (-sy*x_sq+sx*dp_xy)/denom

%plot results
plot(x,y,'o')
hold
plot(x,m*x+b,'--v')
legend('data','model')
-----------------------End of actual file contents--------------------------------
```

The above subroutine is called using the following driver:

```
File: run_ls.m
----------------------Start of actual file contents------------------------
w = [3.5 4 6 25 103 117 200 252 300 437 1340 2000 2700 5000
     22500 30000 33000 50000 70000 100000 415000 450000 500000 3000000];
r = [787 660 588 670 347 300 420 352 300 269 251 205 187 120 100
     85 81 70 72 70 45 38 40 48];

x = w.^(-1/3)'%note that the transpose operator ' turns the row vec into col vec.
y = r'% the subroutine expects column vectors by design.

[m,b] = ls_normal (x,y)
-----------------------End of actual file contents--------------------------------
```

A.3. Least Squares with Overdetermined System

```
File: ls_interp.m
----------------------Start of actual file contents-----------------------
%Input:
%   x is a column vector of domain (input) variables
%   y is a column vector of range (output) variables
%
%Output:
%   m is the slope of the line
%   b is the intercept of the line

function [m,b] = ls_interp(x,y)

%Compute the matrix X
```

Recall the equation we are solving in this problem:

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_P \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_P \end{pmatrix}$$

In terms of matrices we can summarize the above as

$$\mathrm{Xvec} = \mathrm{y}$$

This equation will be solved via matlab's backslash routine.

```
P = size(x,1)%how big is the data set?

c1 = ones(P,1)%create a column vector of ones of length P

X = [c1 x]% construct the "interpolation matrix"

vec = X\y% solve the least squares problem

b = vec(1)%obtain the first component (intercept)
m = vec(2)%obtain the slope

%plot results
plot(x,y,'o')
hold
plot(x,m*x+b,'--v')
legend('data','model')
----------------------End of actual file contents--------------------------
```

A.4. Non-Newtonian Fish

See the discussion in Section 4.3.1.

```
File: fishfit.m
----------------------Start of actual file contents------------------------
y = log(T(2:101)-T(1:100))%use only the first 100 points
x = log(425-T)%


m11 = x(1:100)*x(1:100)'

m12 = sum(x(1:100))
m21 = m12
m22 = 100
r1 = (x(1:100))*y'
r2 = sum(y(1:100))

R = [r1;r2]

M = [m11 m12; m21 m22]

XX = inv(M)*R%this vector contains the results [b alpha]
k = exp(XX(2))%transforming back to original model equation
-----------------------End of actual file contents--------------------------------
```

A.5. Preditor or Prey?

See the discussion in Section 4.3.2.

```
File: predpreyfit.m
----------------------Start of actual file contents------------------------
a = f(1:19800);%init data representing the population of species A
b = r(1:19800);%init data representing the population of species B

%%FOXES matrix equation coefs
m11 = -sum(a*a')
m12 = sum(a.*a.*b)
m21 = -m12
m22 = sum(a.*a.*b.*b)

z1 = (f(2:19801)-a)*a'
z2 = sum((f(2:19801)-a).*a.*b)

M = [m11 m12; m21 m22];%now compute the matrix by assembling the components
R1 = [z1;z2]%this is the RHS of equation for (g_1, c_1)

cg1 = inv(M)*R1%solve for (g_1, c_1)


%%rabbits equation coefs
```

```
m11 = sum(b*b')
m12 = -sum(b.*b.*a)
m21 = -m12
m22 = -sum(a.*a.*b.*b)

z1 = (r(2:19801)-b)*b'
z2 = sum((r(2:19801)-b).*a.*b)

M = [m11 m12; m21 m22];
R1 = [z1;z2]

cg2 = inv(M)*R1%solve for (g_2, c_2)
----------------------End of actual file contents-------------------------------
```

A.6. Tire Distributor

See the discussion in Chapter 7 Section 7.1.

File: flatdemand.m

```
---------------------Start of actual file contents-----------------------

function [num_tires] = flatdemand

u = rand(1);
u1 = rand(1);
if u < 12/365
    num_tires = floor(u1*100);
elseif u >= 12/365 & u < 16/365
    num_tires = 100 + floor(u1*200);
elseif u >= 16/365 & u < 43/365
    num_tires = 300 + floor(u1*200);
elseif u >= 43/365 & u < 86/365
    num_tires = 500 + floor(u1*200);
elseif u >= 86/365 & u < 134/365
    num_tires = 700 + floor(u1*200);
elseif u >= 134/365 & u < 201/365
    num_tires = 900 + floor(u1*200);
elseif u >= 201/365 & u < 279/365
    num_tires = 1100 + floor(u1*200);
elseif u >= 279/365 & u < 334/365
    num_tires = 1300 + floor(u1*200);
elseif u >= 334/365 & u < 356/365
    num_tires = 1500 + floor(u1*200);
elseif u >= 356/365 & u < 363/365
    num_tires = 1700 + floor(u1*200);
else u >= 363/365 & u <1
    num_tires = 1900 + floor(u1*200);
end
----------------------End of actual file contents-------------------------------
```

File: simulate.m

```
----------------------Start of actual file contents-----------------------


function [AVE_DAILY_COST] = simulate(num_runs, delivery_interval, delivery_quantity)

NUM_DAYS = 365
TRUCK_CAPACITY=4000;%tires
truck_charge = 400;%delivery cost per truck
delivery_charge = truck_charge*(floor((delivery_quantity-1)/TRUCK_CAPACITY)+1);
penalty = 1000;%10 dollar penalty per day * 1000 tires

for i = 1:num_runs
    day_counter = 0;
    %Assume there is a delivery at the outset
    NUM_TIRES = delivery_quantity;%init
    COST =delivery_charge;
    interest_rate = 0.01;
    days_without_tires =0;
    delivery_counter = 0;

    daily_inventory(1) = NUM_TIRES;

    while day_counter <= NUM_DAYS;
        if delivery_counter == delivery_interval;%add delivery charge
            NUM_TIRES = NUM_TIRES + delivery_quantity;
            COST = COST + delivery_charge;
            delivery_counter = 0;
        end
        NUM_TIRES = NUM_TIRES - flatdemand; %sell tires for day

        if NUM_TIRES <=0;%out of stock?
            days_without_tires = days_without_tires + 1;
            NUM_TIRES = 0;
        end
        COST = COST + interest_rate*NUM_TIRES;%add charges for unsold tires
        day_counter = day_counter + 1;
        daily_inventory(day_counter+1) = NUM_TIRES;
        delivery_counter = delivery_counter + 1;
    end
    %add daily charge due interest at end of day for remaining tires
    cost_penalty = days_without_tires*10000;
    AVE_DAILY_COST(i) = (COST + cost_penalty)/NUM_DAYS;
end
----------------------End of actual file contents-------------------------------
```

```
File: run_simulate.m
---------------------Start of actual file contents-----------------------

num_runs = 100;
delivery_quantity = 4000; %number to have delivered in each shipment (X in notes)


for j=1:6
    for i = 1:25
    delivery_interval = i;
    DAILY_COST(:,i,j) = simulate(num_runs, delivery_interval, delivery_quantity);
    end
    delivery_quantity = delivery_quantity + 4000
end

for j=1:6
    for i = 1:25
    AVES(i,j) = sum(DAILY_COST(:,i,j))/num_runs
    end
end



semilogy(1:25,AVES(:,1),'.',1:25,AVES(:,2),'x',1:25,AVES(:,3),'o',1:25,AVES(:,4),'+',1:25,AVE
legend('delivery quantity 4000','delivery quantity 8000','delivery quantity 12000','delivery
ylabel('average daily cost')
xlabel('delivery interval in days')
----------------------End of actual file contents-------------------------------
```

A.7. Blackjack

```
File: blackjack.m
---------------------Start of actual file contents-----------------------
%blackjack simulation
%bet one dollar on each hand; start with $100
my_money = 100;
iwin=0;%counter for number of wins
ilose = 0;%counter for number of losses
ties = 0;%counter for number of ties
hand = 0;%counter for number of hands played.
my_stay_value = 14; %don't take another card if hand is worth 14 or more.

%deal cards from 2 decks
for i = 1:100
%note 11= jack, 12 = queen, 13 = king, and 14 = ace.
d1 = [2 2 2 2 3 3 3 3 4 4 4 4 5 5 5 5 6 6 6 6 7 7 7 7 8 8 8 8 9 9 9 9 10 10 10 10
      11 11 11 11 12 12 12 12 13 13 13 13 14 14 14 14];
deck = [d1 d1];
numcards = size(deck,2);
```

```
bigindexset = 1:numcards;
permuted_indices = bigindexset(randperm(numcards));
shuffled_cards = deck(permuted_indices);


%One game (of two decks)
card_counter = 1;
while card_counter < numcards - 10%dont let deck run out!
   %new hand
   hand = hand +1;
   my_count = 0;
   dealer_count = 0;

   %make initial deal of two cards to Player and Dealer.
   my_count = my_card_value(card_counter, shuffled_cards, my_count);
   card_counter = card_counter+1;%advance deck index
   dealer_count = dealer_card_value(card_counter, shuffled_cards, dealer_count);
   card_counter = card_counter+1;
   my_count = my_card_value(card_counter, shuffled_cards, my_count);
   card_counter = card_counter+1;
   dealer_count = dealer_card_value(card_counter, shuffled_cards, dealer_count);
   card_counter=card_counter+1;

   while my_count < my_stay_value  & card_counter < numcards
      my_count = my_card_value(card_counter, shuffled_cards, my_count);
      card_counter = card_counter +1;
   end

   while dealer_count < 17  & card_counter < numcards & my_count < 22
      dealer_count = dealer_card_value(card_counter, shuffled_cards, dealer_count);
      card_counter = card_counter +1;
   end

%who wins?
   if my_count > 21% I am bust
      my_money = my_money - 1;
      ilose = ilose +1;
   elseif dealer_count > 21%dealer is bust
      my_money = my_money + 1;
      iwin = iwin +1;
   elseif my_count == dealer_count
      %push--my winnings don't change
      ties = ties +1;
   elseif my_count > dealer_count
      my_money = my_money +1;
      iwin = iwin +1;
   else
      my_money = my_money -1;
```

```
      ilose = ilose + 1;
   end%if
   %construct an array that computes a running fraction of losses
   perclose(hand) = ilose/(iwin+ilose+ties);
   end%while
end%for

   iwin
   ilose
   ties
   plot(perclose)
   my_money
----------------------End of actual file contents-------------------------------
File: dealer_card_value.m
---------------------Start of actual file contents------------------------

   function newcount = dealer_card_value(card_counter, shuffled_cards, dealer_count)

   card_value = shuffled_cards(card_counter);
   if card_value >= 10 & card_value < 14
newcount = dealer_count + 10;
   elseif card_value < 10
      newcount = dealer_count + card_value;
   else%card is an ace and has value of 11 for dealer unless he goes bust.
      bigcount = 11 + dealer_count;
      smallcount = 1 + dealer_count;
      if bigcount <22
         newcount = bigcount;
      else
         newcount = smallcount;
      end
   end
-----------------------End of actual file contents-------------------------------
File: dealer_card_value.m
---------------------Start of actual file contents------------------------

   function newcount = my_card_value(card_counter, shuffled_cards, my_count)

   card_value = shuffled_cards(card_counter);
   if card_value >= 10 & card_value < 14
newcount = my_count + 10;
   elseif card_value < 10
      newcount = my_count + card_value;
   else%card is an ace and has value one or 11.
      bigcount = 11 + my_count;
      smallcount = 1 + my_count;
      if bigcount >18 & bigcount <22
         newcount = bigcount;
```

```
      else
         newcount = smallcount;
      end
   end
-----------------------End of actual file contents--------------------------------
```

APPENDIX B

Solutions

Solution to Problem 4.22: Analytically solve the linear difference equation from the previous problem

$$x_{n+1} = (2 - \alpha\Delta t)x_n + (\alpha\Delta t - 1)x_{n-1} - (\Delta t)^2 x_n$$

and compare with your numerical simulation above. For simplicity you may take $\Delta t = 0.05, \alpha = 0.1, x_1 = 2, x_2 = 2.0001$.

Let's write the problem as

$$x_{n+1} + ax_n + bx_{n-1} = 0$$

where

$$a = \alpha\Delta t + (\Delta)^2 - 2 = -1.9925$$

and

$$b = 1 - \alpha\Delta t = 0.995$$

The associated characteristic equation is

$$\lambda^2 + a\lambda + b = 0$$

which has solutions (after factoring out i)

$$\lambda_{\pm}^n = \big(-\frac{a}{2} \pm i\frac{\sqrt{4b - a^2}}{2}\big)^n$$

Observe that since $a^2 - 4b = -0.00994375 < 0$ the roots are imaginary. Thus, our solutions can be written in the imaginary form

$$x_n = c_1\lambda_+^n + c_2\lambda_-^n$$

To reformulate the solution as a real solution we compute λ_+ in polar form. Let

$$\lambda_+ = x + iy = -a/2 + i\frac{\sqrt{4b - a^2}}{2}$$

so $x = -a/2$ and $y = \sqrt{4b - a^2}/2$. We showed in Section 4.1.2 that given

$$x = r\cos\theta, y = r\sin\theta$$

the

$$r = \sqrt{b} = 0.99749686716300, \quad \theta = \tan^{-1}(\frac{\sqrt{4b - a^2}}{-\alpha}) = 0.05000513115822$$

Using

$$\lambda_+^n = (r\exp(i\theta))^n = r^n\exp(in\theta) = r^n\cos(n\theta) + ir^n\sin(n\theta)$$

and recalling that both the imaginary and real parts are solutions we obtain

$$x_n = c_1 r^n\cos(n\theta) + c_2 r^n\sin(n\theta)$$

To simplify the calculation (try it both ways to see why) we will shift the indices of the initial conditions to $0, 1$ so

$$x_0 = c_1$$
$$x_1 = c_1 r\cos(\theta) + c_2 r\sin\theta$$

which gives

$$c_2 = (x_1 \tfrac{a}{2} x_0)\frac{2}{\sqrt{4b-a^2}} = 0.15242931201143$$

So the analytical solution to the homogeneous problem is

$$x_n = 2(0.995)^{n/2}\cos(0.05n) + 0.1524(0.995)^{n/2}\sin(0.05n)$$

□

Solution to Problem 4.23: Analytically solve the linear nonhomogeneous difference equation

$$x_{n+1} = (2-\alpha\Delta t)x_n + (\alpha\Delta t - 1)x_{n-1} - (\Delta t)^2 x_n + 0.01\sin(n/50)$$

Simulate this problem numerically and compare with your analytical solution for 2000 iterations. Can you identify a transient (i.e., a term that goes to zero) and steady state (persisent) components of your solution? Again, for simplicity you may take $\Delta t = 0.05, \alpha = 0.1, x_1 = 2, x_2 = 2.0001$. Hint: combine your solution to the homogeneous problem found above with a particular solution of the form

$$p_n = A\cos(n/50) + B\sin(n/50)$$

Solve for the undetermined coefficients A and B.

Plugging $p_n = A\cos(n/50) + B\sin(n/50)$ into the nonhomogeneous problem we have

$$A\cos\frac{n+1}{50} + B\sin\frac{n+1}{50} + a(A\cos\frac{n}{50} + B\sin\frac{n}{50}) + b(A\cos\frac{n-1}{50} + B\sin\frac{n-1}{50}) = .01\sin\frac{1}{50}$$

Letting $n = 1$ gives

$$A(\cos\frac{2}{50} + a\cos\frac{1}{50} + b) + B\sin(\frac{2}{50} + a\sin\frac{1}{50}) = .01\sin\frac{1}{50}$$

and letting $n = 2$ gives

$$A(\cos\frac{3}{50} + a\cos\frac{2}{50} + b\cos\frac{1}{50}) + B(\sin\frac{3}{50} + a\sin\frac{2}{50} + b\sin\frac{1}{50}) = .01\sin\frac{1}{50}$$

Recall the values of $a = -1.9925$ and $b = 0.995$ now we can solve to find

$$A = -0.22601165839997, \quad B = 4.74885158590701$$

So the solution is the sum of the homogenous part from above (before the constants are evaluated!!) and this particular solution

$$x_n = c_1(0.995)^{n/2}\cos(.05n) + c_2(0.995)^{n/2}\sin(.05n) - 0.2260\cos(n/50) + 4.7489\sin(n/50)$$

Setting $x_0 = 2$ and $x_1 = 2.0002$ one gets

$$c_1 = 2.22601165839997, \quad c_2 = -1.73625725291089$$

Note that if you use $x_1 = 2$ and $x_2 = 2.0002$ you should obtain

$$c_1 = 2.22108702188432, c_2 = -1.63986237742405$$

Bibliography

[A] T. Aoki, *Calcul exponentiel des opérateurs microdifferentiels d'ordre infini.* I, Ann. Inst. Fourier (Grenoble) **33** (1983), 227–250.

[B] R. Brown, *On a conjecture of Dirichlet*, Amer. Math. Soc., Providence, RI, 1993.

[D] R. A. DeVore, *Approximation of functions*, Proc. Sympos. Appl. Math. vol. 36, Amer. Math. Soc., Providence, RI, 1986, pp. 34–56.

Index

Absorbing barrier, 4
Adjoint partial differential operator, 20
A-harmonic function, 16, 182
A^*-harmonic function, 182

Boundary condition, 20, 22
  Dirichlet, 15
  Neumann, 16
Boundary value problem
  the first, 16
  the second, 16
  the third, 16
Bounded set, 19

Diffusion
  coefficient, 1
  equation, 3, 23
Dirichlet
  boundary condition, 15
  boundary value problem, 16

Elliptic
  boundary value problem, 14, 158
  partial differential equation, 14
  partial differential operator, 19

Fick's law, 1
Flux, 1
Formally adjoint partial differential operator, 20
Fundamental solution
  conceptional explanation, 12
  general definition, 23
  temporally homogeneous case, 64, 112

Genuine solution, 196
Green function, 156
Green's formula, 21

Harnack theorems
  first theorem, 185
  inequality, 186
  lemma, 186
  second theorem, 187
  third theorem, 187
Helmhotz decomposition, 214
Hilbert-Schmidt expansion theorem, 120

Initial-boundary value problem, 22
Initial condition, 22
Invariant measure (for the fundamental solution), 167

Maximum principle
  for A-harmonic functions, 183
  for parabolic differential equations, 65
  strong, 83

Neumann
  boundary condition, 16
  boundary value problem, 16
  function, 179

One-parameter semigroup, 113

Parabolic initial-boundary value problem, 22
Partial differential equation
  of elliptic type, 14
  of parabolic type, 22
Positive definite kernel, 121

Reflecting barrier, 4
Regular (set), 19
Removable isolated singularity, 191
Robin problem, 16

Semigroup property (of fundamental solution), 64, 113
Separation of variables, 131
Solenoidal (vector field), 209
Strong maximum principle, 83
Symmetry (of fundamental solution), 64, 112

Temporally homogeneous, 111

Vector field with potential, 209

Weak solution
  of elliptic equations, 195
  of parabolic equation, 196
  associated with a boundary condition, 204