1

Introduction

1.1 System Models

This book analyzes mathematical models for systems and explores techniques for optimizing systems described by these models. We use the term system in its broad sense; by a **system** we mean a collection of things which are related in such a way that it makes sense to think of them as a whole. Examples of systems are an electric motor, an automobile, a transportation system, and a city. Each of these systems is part of a larger system. Small systems are usually well understood; large, complex systems are not.

Rational decision making concerning the design and operation of a system is always based upon a model of that system. A **model** of a system is a simpler system that behaves sufficiently like the system of interest to be of use in predicting the behavior of the system. The choice of appropriate model depends upon the complexity of the system, the available resources, and the questions that need to be answered by the model. Many decisions are based upon nothing more than the conceptual model which the decision maker develops by observing the operation of other systems. In this book we concern ourselves with a more quantitative class of models, mathematical models.

Most systems can be thought of (or modeled) as an operation on the system inputs (or independent variables) which produces the system outputs (or dependent variables); we state this input-output relationship symbolically by means of the following mathematical equation:

$$\mathbf{T}\mathbf{x} = \mathbf{y} \tag{1.1}$$

In this equation $\mathbf{x}$ represents the set of inputs to the system and $\mathbf{y}$ the set of outputs of the system.* The symbol $\mathbf{T}$ represents the operation which the system performs on the inputs; thus $\mathbf{T}$ is a mathematical model of the system.

*See Section 2.3 for a more complete discussion of inputs and outputs.

In order for a model of a system to be conceptually simple, it must be abstract. The more details we include explicitly in the model, the more complicated it becomes. The more details we make implicit, the more abstract it becomes. Thus if we seek conceptual simplicity, we cannot avoid abstraction. The model $\mathbf{T}$ of (1.1) epitomizes this simplicity and abstraction.

The generality of the model given in (1.1) allows it to be applied to many different systems. In the simplest of situations $\mathbf{T}$ might represent a simple economic transaction: let p be the unit price of a particular commodity; then (1.1) means $\mathbf{y} = p\mathbf{x}$, where $\mathbf{x}$ is the quantity purchased and $\mathbf{y}$ is the total cost of the purchase. At the other extreme, $\mathbf{T}$ might represent a large city. Figure 1.1 shows the system output $\mathbf{y}$ that might result from a given input $\mathbf{x}$; obviously, many pertinent variables are not explicit in Figure 1.1.

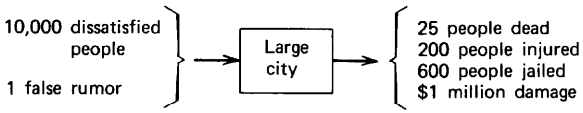


Figure 1.1. A conceptual model of a large city.

Equation (1.1) is the focus of this book. The first five chapters are devoted to a detailed analysis of (1.1) for models $\mathbf{T}$ which are linear.* By decomposing linear models into smaller, simpler pieces we develop an intuitive feel for their properties and determine the practical computational difficulties which can arise in using linear models. Chapter 6 treats the least-square optimization of systems that can be represented by linear models. The analysis and optimization of systems that are described by nonlinear models are considered in Chapters 7-8.

We emphasize linear models because most known analytical results pertain only to linear models. Furthermore, most of the successful techniques for analyzing and optimizing nonlinear systems consist in the repetitive application of linear techniques (Chapter 7-8). We dwell extensively on the two most frequently used linear models-linear algebraic equations and linear differential equations. These models are the most frequently used because they are well understood and relatively easy to deal with. In addition, they are satisfactory models for a large number of practical systems.

Throughout the text we explore the computational implications of the analytical techniques which we develop, but we do not develop computer

*See Section 2.4 for the definition of a linear system.

algorithms. We do not discuss stochastic systems; we treat systems with stochastic inputs only by means of examples.

System Questions

Questions concerning a system usually fall into one of the following categories:

1. *System operation:* in terms of (1.l), given the model $\mathbf{T}$ and the input $\mathbf{x}$, find the output $\mathbf{y}$.
2. *System inversion:* given the model $\mathbf{T}$ and output $\mathbf{y}$, find the input $\mathbf{x}$.
3. *System synthesis* or *identification:* given several different choices of input $\mathbf{x}$ and the corresponding output $\mathbf{y}$ for each input, determine a suitable system model $\mathbf{T}$. (If the system is to be identified, the inputs and outputs are measurements from a real system. If the system is to be synthesized, $\mathbf{T}$ would be chosen to provide some desired input-output relationship.)
4. *System optimization:* pick the input $\mathbf{x}$, the output $\mathbf{y}$, or the system $\mathbf{T}$ so that some criterion is optimized.

Note that we have expressed these questions in terms of the system model rather than in terms of the system itself. Although experimentation with actual systems may be appropriate in certain circumstances, these questions are usually explored by means of a model. We discuss the modeling process briefly in Section 1.4. We also examine in Chapter 6 some techniques for making an optimum choice of model parameters once a model structure has been established. However, we do not dwell extensively on techniques for obtaining good models. Rather, we work with the models themselves, assuming that they are good models for the systems they represent. Questions 1 and 2 are treated in Chapters 1, 2, 4, and 5 for linear algebraic equation models and in Chapters 2-5 for linear differential equation models. Question 4 is treated in Chapters 6-8. We do not consider question 3.*

The concepts explored in this book apply directly to any field which uses equations to represent systems or portions of systems. Although we focus on linear algebraic equations and linear differential equations, we also demonstrate the applicability of the concepts to partial differential equations and difference equations; we include equations which are probabilistic, "time-varying," and nonlinear. Our examples pertain to models and optimization in such fields as automatic control, electric power, circuits, statistical communications, coding, heat flow, economics, operations research, etc.

*See Sage [1.10] for a discussion of identification.

1.2 Approach

All students of science and engineering have noticed occasional similarities between the physical laws of different fields. For instance, gravitational attraction, electrostatic attraction, and magnetic attraction all obey an inverse-square law. Electrical resistance to the flow of current has its analogue in the resistance of materials to the conduction of heat. Not only does the physical world tend to repeat itself; it also tends toward simplicity and economy. Most natural phenomena can be explained by simple differential relationships: the net force on a rigid object is proportional to its acceleration; the rate of flow of heat is proportional to the gradient of the temperature distribution.

If we put a number of simple relationships together to describe the motion of a nonrigid object (fuel in a rocket) or the heat flow in an irregular nonhomogeneous object (a nuclear reactor), then nature appears complicated. The human mind is not good at thinking of several things at once. The development of large-scale digital computers has provided the capability for solving complex sets of equations; it has made system study a reality. However, the engineer, the designer of a system, still must conceive of the variables and interactions in the system to such an extent that he can describe for a computer what it is he wants to know. He needs simple conceptual models for systems.

We can simplify models for complex systems by stretching our imagination in a search for analogies. For instance, the multiplication of an electrical current by a resistance to determine a voltage has an analogue in the differentiation of a current and then multiplication by an inductance; both actions are operations on a current to yield a voltage. This analogy suggests that we think of differentiation as analogous to multiplication by a number. By reducing the number of "different concepts" necessary to understand the parts of a system, such analogies help the system designer to achieve greater economy of thought; he can conceive of the system in simpler terms, hopefully gaining insight in the process. William K. Linvill [1.7] has coined the term "portable concept" to describe a concept that is transferable from one setting to another. This book is concerned with *portable mathematical concepts.* The purpose of exploring such concepts is to enhance the ability of the reader to model systems, understand them, synthesize them, and optimize them. Our basic premise is that this ability is enhanced by an intuitive understanding of the models and optimization techniques that have proved useful in many settings in the past. By an intuitive understanding, we mean the type of "intuitive feel" that an engineer obtains by applying and reapplying a concept to many different situations.

It would seem, then, that we must fully absorb most of mathematics. However, much of the mathematical literature is directed toward the

modeling and optimization of pathological cases, those cases for which "standard" models or techniques are insufficient. Because techniques for handling these cases are new, it is appropriate that they be the focus of the current literature. Yet this emphasis on exceptional cases can distort our perspective. In maximizing a function, we should not become so concerned about nondifferentiability of functions that we forget to try setting the derivative equal to zero. Rather than try to explore *all* cases, we focus on well-behaved systems. By making analogies, we organize the most common models and optimization techniques into a framework which contains only a relatively few fundamental concepts. The exceptional cases can be more clearly understood in comparison to this basic framework.

The importance of learning the *structure* of a subject is stressed by Bruner [1.1]: "Grasping the structure of a subject is understanding it in a way that permits many other things to be related to it meaningfully… the transfer of principles is dependent upon mastery of the structure of the subject matter… . Perhaps the most basic thing that can be said about human memory, after a century of intensive research, is that unless detail is placed into a structured pattern, it is rapidly forgotten." In order to simplify and unify the concepts used in model analysis and optimization, we organize fundamental mathematical principles into a mnemonic structure—a structure which draws extensively on geometrical analogies as an aid to the memory. We also develop a mathematical language suitable for communicating these structural concepts.

The first half of this book is concerned with models and their analysis. Mathematically speaking, this is the subject of algebra-the use of symbols to express quantitative concepts and their relations. In the latter half of the book we turn to geometry-the measurement and comparison of quantitative concepts—in order to further analyze models and to optimize their parameters and inputs. Because the bulk of known analytical results are concerned with linear models, these models necessarily dominate our discussions. Our emphasis is on geometrical insight rather than mathematical theorems. We reach deep into the mathematical literature for concepts. We try to be rigorously correct. Yet we develop concepts by means of analogies and simple examples rather than proofs, in order to nurture the intuition of the reader. We concern ourselves with the practical aspects of computation. To engineers the material seems like mathematics; to mathematicians it seems like engineering.

1.3 Portable Concepts

To illustrate the portability of the mathematical model (1.1) we compare the two most common mathematical models: *(a)* a set of linear algebraic equations; and *(b)* a linear differential equation. The following algebraic

equations might represent the relationship between the voltages and the currents in a resistive circuit:

$$\begin{aligned} 2\xi_1 + 3\xi_2 &= \eta_1 \\ \xi_1 + \xi_2 &= \eta_2 \end{aligned} \tag{1.2}$$

Such a set of equations is often expressed in the matrix form:

$$\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \tag{1.3}$$

In the form (1.3), we can interpret the set of equations as an operation (matrix multiplication) on the pair of variables ξ_1 and ξ_2 to obtain the pair of quantities η_1 and η_2. The relationship (1.2) between the pairs of variables can also be expressed in terms of the "inverse equations":

$$\begin{aligned} \xi_1 &= -\eta_1 + 3\eta_2 \\ \xi_2 &= \eta_1 - 2\eta_2 \end{aligned} \tag{1.4}$$

Equations (1.4) can be verified by substitution into (1.2). The coefficients in (1.4) indicate what must be done to the "right-hand side" variables in order to determine the solution to (1.2). Equations (1.4) can be expressed in the "inverse matrix" form:

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} -1 & 3 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \tag{1.5}$$

In Section 1.5 we explore in detail the process of solving or inverting equations such as (1.2). In Chapter 2 we begin the discussion of algebraic equation models in a manner which is consistent with the notation of (1.1). Chapters 4 and 5 are, to a great extent, devoted to analyzing these models.

The angular velocity $\omega(t)$ of a particular loaded dc motor, initially at rest, can be expressed in terms of its armature voltage $u(t)$ as

$$\frac{d\omega(t)}{dt} + \omega(t) = u(t), \qquad \omega(0) = 0 \tag{1-6}$$

We can think of the differential equation and boundary condition as an abstract operation on ω to obtain u. Equation (1.6) also can be expressed in the inverse form:

$$\omega(t) = \int_0^t e^{-(t-s)} u(s)\, ds \tag{1.7}$$

That the integral equation (1.7) is, in fact, the solution to (1.6) is easily

verified for a *particular* armature voltage, say, $u(t) = e^{2t}$, by evaluating $\omega(t)$, then substituting it into (1.6). We can think of (1.7) as an abstract "integral" operation on u to determine ω; this is the "inverse" of the "differential" operation in (1.6). These two abstract operations and techniques for determining the inverse operation are the subject of Chapter 3. The analysis of these abstract operations carries into Chapters 4 and 5.

The algebraic equations (1.2) and the differential equation with its boundary condition (1.6) have much in common. We must not let details cloud the issue; in each case, an "input" is affecting an "output" according to certain (linear) principles. We can think of the pair of variables ξ_1 and ξ_2 and the function ω as each constituting a single "vector" variable. The analogy between these entities is carried further in the comparison of Figure 1.2, wherein the pair of variables ξ_1, ξ_2 is treated as a "discrete" function. This analogy is discussed further in Section 2.1. It seems evident that concepts are more clearly portable if they are abstracted-stripped of their details.

A Portable Optimization Concept

We again employ the analogy between a "discrete vector" variable and a "continuous vector" variable to discuss the portability of an optimization

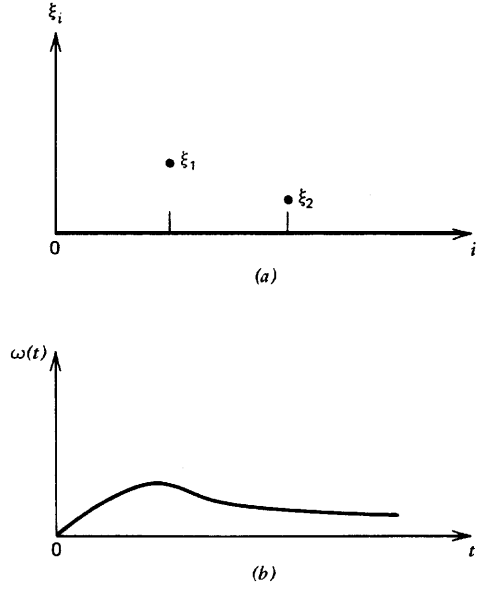


Figure 1.2. Vector variables plotted as functions: (a) discrete variables of (1.2); (b) continuous variable of (1.6).

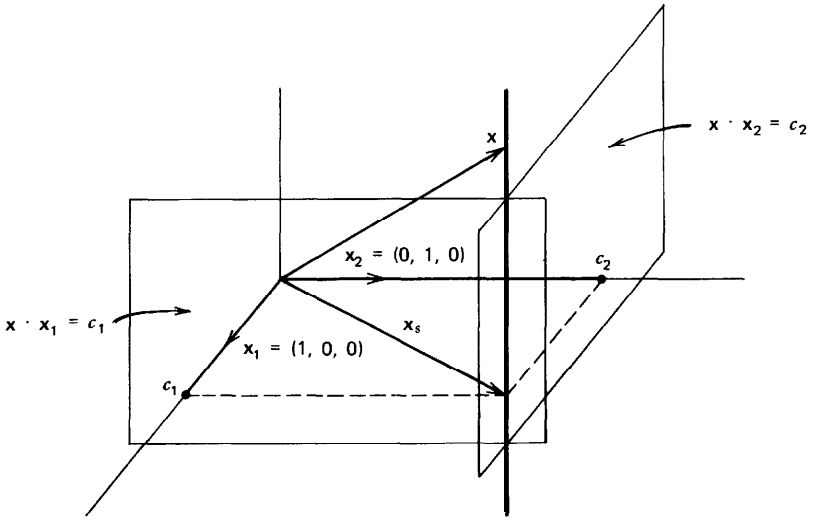


Figure 1.3. **A** vector of minimum length.

concept. Figure 1.3 shows the locus of all vectors $\mathbf{x}$ in a three-dimensional space which lie in the intersection of two planes. We seek that vector $\mathbf{x}$ which is of minimum length. The solution vector $\mathbf{x}_s$ is perpendicular to the line which constitutes the locus of the candidate vectors $\mathbf{x}$.

Using the standard notation of analytic geometry, we think of the vector $\mathbf{x}$ as $\mathbf{x} = (\xi_1, \xi_2, \xi_3)$. The plane that is perpendicular to the vector $\mathbf{x}_1$ can be expressed mathematically in terms of the dot product of vectors as $\mathbf{x} \cdot \mathbf{x}_1 = \xi_1 = c_1$. Similarly, the second plane consists in vectors $\mathbf{x}$ which satisfy $\mathbf{x} \cdot \mathbf{x}_2 = c_2$. Since $\mathbf{x}_s$ must be perpendicular to the intersection of the planes, it must be some combination of the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ that determine the planes; that is, $\mathbf{x}_s = d_1\mathbf{x}_1 + d_2\mathbf{x}_2$ for some constants d_1 and d_2. Substituting $\mathbf{x}_s$ into the equations that determine the planes, we obtain a pair of algebraic equations in d_1 and d_2:

$$\begin{aligned} \mathbf{x}_s \cdot \mathbf{x}_1 &= d_1\mathbf{x}_1 \cdot \mathbf{x}_1 + d_2\mathbf{x}_2 \cdot \mathbf{x}_1 = c_1 \\ \mathbf{x}_s \cdot \mathbf{x}_2 &= d_1\mathbf{x}_1 \cdot \mathbf{x}_2 + d_2\mathbf{x}_2 \cdot \mathbf{x}_2 = c_2 \end{aligned} \tag{1.8}$$

Since the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are perpendicular and of unit length, then

$$\mathbf{x}_1 \cdot \mathbf{x}_1 = \mathbf{x}_2 \cdot \mathbf{x}_2 = 1, \qquad \mathbf{x}_1 \cdot \mathbf{x}_2 = \mathbf{x}_2 \cdot \mathbf{x}_1 = 0, \qquad d_i = c_i,$$

and

$$\mathbf{x}_s = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 = (c_1, c_2, 0)$$

The geometric minimization problem described above is simple. By using geometric notions, we have found the vector $\mathbf{x}$ which satisfies two linear equations and for which the quantity $\xi_1^2+\xi_2^2+\xi_3^2$ (the length of $\mathbf{x}$ squared) is minimum. The same geometric principles can be used to solve other, more complicated, problems wherein linear equations must be satisfied and a quadratic quantity minimized. For instance, the angular position $\phi(t)$ of the shaft of the dc motor of (1.6) satisfies

$$\frac{d^2\phi(t)}{dt^2}+\frac{d\phi(t)}{dt}=u(t) \tag{1.9}$$

Suppose we seek that armature voltage function u(t) that will drive the motor shaft from one position to another in a fixed time, while consuming a minimum amount of energy; that is, let $\phi(0)=\dot{\phi}(0)=0$, $\phi(1)=1$, $\dot{\phi}(1)=0$, and pick u to minimize $\int_0^1 u^2(t)\,dt$. In our search for a technique for solving this problem we should not cloud the issue by thinking about techniques for solving differential equations. Equation (1.9) is linear; the quantity to be minimized is quadratic. Chapter 6 is devoted to solving such problems by using analogues of the planes and perpendicular vectors of Figure 1.3.

1.4 System Modeling

The rationale for modeling a system is a desire to determine how to design and/or operate a system without experimenting with actual systems. If a system is large, experimenting is usually very time consuming, extremely expensive, and often socially unacceptable. A designer uses models to predict the performance characteristics of a system or to aid in modifying the design of the system so that it meets a desired set of specifications. He will probably be interested in the degree of stability of the system, its accuracy, and its speed of response to commands. The designer also uses models to predict the nature of the interaction of the system with other systems. For example, he may wish to predict the effect of the system or of a particular system operating policy on the environment or on a related energy distribution system. Or he may wish to predict the performance of the system in the presence of extraneous inputs (noise) or sudden changes in load. The reliability of the system and the sensitivity of the system performance to changes in the environment are also important.

Types of Models

A single system has many models. One or more models of the system pertain to its electrical behavior, others to its thermal behavior, still others

to its mechanical behavior. An investigation of the social or economic characteristics of the system requires additional models.

Physical models are appropriate in many situations. One example of such a model is a scale model of a building or bridge. The conceptual representation of a rocket by a solid cylinder is another example. In most system studies, a *mathematical model* for the system (or part of the system) facilitates analysis. An appropriate mathematical model usually can be derived more easily from a simplified physical model than from the original system. The resulting mathematical model usually consists of a set of algebraic and/or differential equations. Often these equations can be solved (for given system inputs) on a digital, analogue, or hybrid computer.[†] In some instances, the distributed nature of the system requires a mathematical model consisting of partial differential equations, and computer solutions are difficult to obtain even if the equations are linear.

The behavior of some systems fluctuates randomly with time. For such systems (or portions of systems) it is common to build a discrete-event *simulation model*,t Rather than predicting the precise behavior of the system, such a model simulates the behavior numerically in a manner that is statistically correct. For instance, we might be interested in the flow of customers through a set of checkout counters. A simple physical model of such a customer service system consists of a single checkout counter, where customers arrive, wait for service, are served, then leave; arrival times and service times are random with known statistics. By means of a digital computer, we would generate a random sequence of arrivals (with correct statistical properties). We would also determine a service time for each customer by an appropriate random number generation process. Then we would observe the simulated flow of customers over time. The simulation would predict not only the average flow through the system, but also the frequency of occurrence of various queue lengths and waiting times. Thus the dynamic performance of certain types of systems can be predicted by digital simulation.

As a practical matter, a model should contain no more detail than is necessary to accomplish the purposes of the model. One is seldom sure of the accuracy of a model. Yet if a model is accurate enough to improve one's decision-making capability, it serves a useful purpose. Generally speaking, the more complex the model is, the more expensive will be the process of developing and using the model. In the extreme, the most accurate model is a copy of the system itself.

*Special computer programs have been developed to facilitate the solving of certain classes of equations. One example is MATLAB®; it is effective in solving linear algebraic and linear differential equations.

†Specialized computer languages have been developed to facilitate discrete-event simulation. Examples are ARENA®, SIMSCRIPT®, and GPSS®.

Unfortunately, it is probable that some complex systems will never be represented in sufficient detail by manageable mathematical models. Yet a *conceptual model* can be applied in situations where it is difficult to obtain meaningful quantitative models; for example, the principle of negative feedback (with its beneficial effects on stability and sensitivity) often is applied successfully without the use of a mathematical model. The system concepts that are associated with mathematical models serve as a guide to the exploration of complex systems. By the use of specific models for small subsystems, by computer analysis of the combined subsystem models, and by the application of model concepts (such as feedback) to the whole system, we can better understand large systems.

The Modeling Process

The process of modeling can be divided into two closely related steps: (1) establishing the model structure and (2) supplying the data. We focus primarily on the first step. However, we cannot ignore the second; it is seldom useful to establish a model structure for which we cannot obtain data.

We begin the modeling process by examining the system of interest. In many complex systems, even the boundaries of the system are not clear. The motivation for modeling such a system is usually a desire to solve a problem, to improve an unsatisfactory situation, or to satisfy a felt need. We must describe the system and the manner in which it performs in a simple fashion, omitting unnecessary detail. As we begin to understand better the relationship between the system and the problem which motivates study of the system, we will be able to establish suitable boundaries for the system.

Suppose a housing official of a large city is concerned because the number of vacant apartments in his city cycles badly, some times being so high as to seriously depress rental rates, other times being so low as to make it difficult for people to find or afford housing.* What is the reason for the cycling? To answer this question, we need to explore the "housing system." Should we include in "the system" the financial institutions which provide capital? The construction industry and labor unions which affect new construction? The welfare system which supports a significant fraction of low-income housing? Initially, we would be likely to concern ourselves only with the direct mechanisms by which vacant apartments are generated (new construction, people moving out, etc.) and eliminated (new renters).

Should the model account for different sizes of apartments? Different styles? Different locations? Seasonal variations in the number of vacan-

*The idea for this example was obtained from Truxal [1.1], Chapter 21.

ties? A model that accounts for all these factors would require detailed data (as a function of time) for each factor. These data are not likely to be readily available. Rather, obtaining the data would require the cooperation of many apartment managers and an extensive data-taking operation over at least a l-year period. A more likely approach, at least initially, would be to develop a simple model which predicts the average number of vacancies (of any type) in the city in a l-year period. Data concerning this quantity are probably available for at least a large fraction of the large apartment complexes in the city.

Once the approximate extent of the system and the approximate degree of detail of the model have been determined, the course of model development usually progresses through the following steps:

1. Development of a simple physical model.
2. Derivation of a mathematical model of the physical model.
3. Obtaining of data from which model parameters are determined.
4. Validation of the model.

In deriving a model for a system it usually helps to visualize the behavior of the unfamiliar system in terms of the behavior of familiar systems which are similar. It is for this reason that we start with a simple physical model. The physical model of the system is likely to be conceptual rather than actual. It is a simple abstraction which retains only the essential characteristics of the original system. In the case of the apartment vacancy model introduced above, a simple physical model might consist of a set of identical empty boxes (vacant apartments). At l-year intervals some number of boxes is added by construction or renters moving out; another number of boxes is removed by new renters. See Figure 1.4.

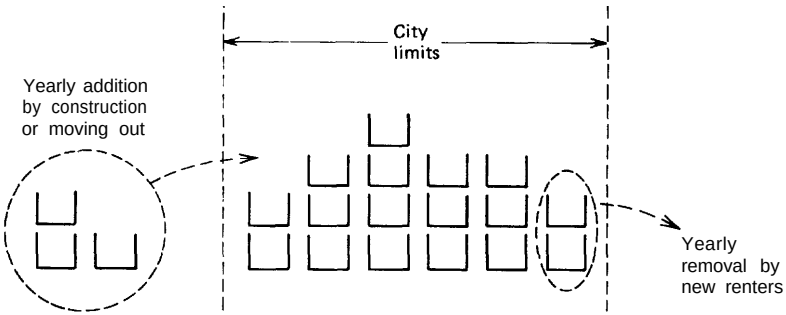


Figure 1.4. Simple physical model of apartment vacancies.

A mathematical model of a system is usually easier to derive from a simple physical model than from the system itself. In most instances the mathematical model consists of algebraic and/or differential equations. The mathematical model must be kept simple in order that it be solvable analytically or by means of practical computer techniques. Generally, the model simplifications that reduce data requirements also reduce the complexity of the mathematical model. For example, in the housing system described above, the aggregation of the various types of apartments into a single type greatly reduces the number of variables in the mathematical model. Other simplifying approximations which may be appropriate in some situations are (1) ignoring interaction between the system and its environment; (2) neglecting uncertainty and noise; (3) lumping distributed characteristics; and (4) assuming linearity and time invariance. Sage [1.10] describes some techniques that are useful in identifying the structure and parameter values of those systems that act in a linear fashion.

Mathematical Model of Apartment Vacancies

In order to demonstrate the logical thought process entailed in the derivation of a mathematical model, we derive a mathematical model of the physical model of apartment vacancies illustrated in Figure 1.4.

We expect that the number of "apartment construction starts" in a given year is approximately equal to the apparent need for new apartments. We formalize this statement by postulating the following relationship:

$$\begin{aligned} S(n) &= \alpha(V_d - V(n)), & V(n) \leqslant V_d \\ &= 0, & V(n) \geqslant V_d \end{aligned} \tag{1.10}$$

where $S(n)$ = number of apartment construction starts in year n;

$V(n)$ = average number of vacant apartments during the l-year period centered at the beginning of year n.

Underlying (1.10) is the assumption that the people who build apartments feel that the city should have approximately V_d vacancies. The proportionality factor α and the number of vacancies V_d should be selected in such a manner that (1.10) most nearly describes recent historical data for the city.

Of course, actual apartment completions lag behind the starts by an appreciable time. We formalize this statement by the equation

$$C(n) = S(n - l) \tag{1.11}$$

where $C(n)$ is the number of completions in year n, and l is the average

construction time. A suitable value for the lag l should be determined from historical data.

Let $R(n)$ denote the number of new apartments rented during year n. We can include in $R(n)$ the families who move out of apartments during the year [$R(n)$ can be negative]. From Figure 1.4, it is apparent that

$$\Delta V(n) = C(n) - R(n) \tag{1.12}$$

where $\Delta V(n) = V(n+1) - V(n)$, the increase in vacant apartments during the l-year period.

The empirical relations (1.10)-(1.11) and the logical statement (1.12) can be related pictorially by means of a **block diagram.** A block diagram is a conceptual tool which is useful for clarifying the structure of a model or for portraying sequences of events. It dramatizes cause and effect relationships. A block diagram of the mathematical model (1.10)-(1.12) is shown in Figure 1.5. Each block in the diagram displays one of the relationships in the mathematical model. *

Figure 1.5 establishes the model structure. In order to determine the values of the model parameters and to validate the model, we need historical data for each variable in the model. The data that we need in order to pick appropriate values for the parameters α, V_d, and l are historical values of yearly starts $S(n)$, yearly completions $C(n)$, and yearly average vacancies $V(n)$. We would probably pick the values of α, V_d, and l by the least-square data-fitting process known as *linear regression* (see Section 6.1).

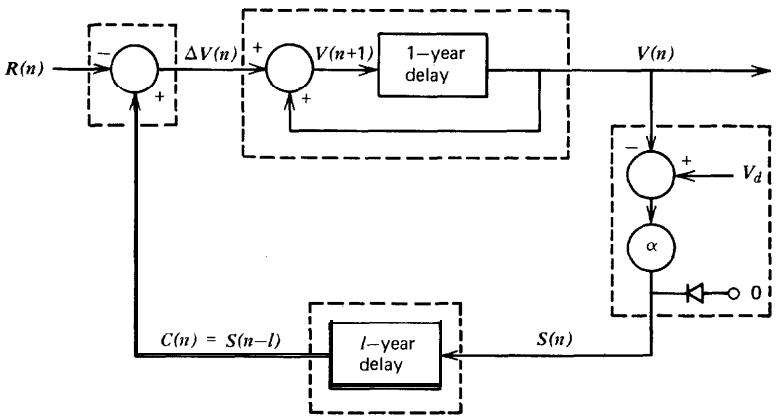


Figure 1.5. Block diagram model of apartment vacancies.

*See Cannon [1.2] for a detailed discussion of block diagrams and their use.

After parameter values have been determined, we need to verify that the mathematical model is a sufficiently good representation of the actual apartment vacancy system. In order to validate the model, we need historical values of the model input $R(n)$ and output $V(n)$. Since we required data for $V(n)$ previously, the only additional data needed are a corresponding set of yearly rentals $R(n)$ (new rentals minus renters moving out). We use the input data $R(n)$ for a sequence of years together with the mathematical model to obtain a predicted sequence of values of $V(n)$. The model is validated if the predicted values of $V(n)$ agree sufficiently with the corresponding historical values of $V(n)$. If the model were verified to be accurate to a certain precision for historical data, we would feel confident that it would exhibit approximately the same accuracy in predicting future apartment vacancies. A housing official would probably be satisfied if the predicted vacancies were within 10% of the actual average vacancies. Of course, predictions of future values of $V(n)$ have to be based on assumed future values of $R(n)$. If future values of $R(n)$ cannot be predicted with reasonable confidence, then another model must be developed to relate the demand for apartments $R(n)$ to those variables which affect demand.

If the data do not validate the model to a sufficient degree, then the model structure must be modified; additional factors must be accounted for. Specifically, the number of apartment construction starts $S(n)$ is likely to depend not only on the demand for housing $R(n)$, but also on the number of uncompleted housing starts (starts from the previous $l-1$ years). The number of starts $S(n)$ is also likely to depend on the availability of capital at a favorable interest rate. Thus an improved apartment vacancy model would probably have more than one input variable.

Once a validated model has been obtained, it can be used to aid city officials in determining an appropriate housing policy. City officials can affect the number of apartment vacancies by modifying the variables which are inputs to the model. Demand for apartments $R(n)$ can be affected by adjusting tax rates, rent subsidies, urban renewal plans, etc. If the final model includes interest rate as an input, this interest rate can be affected by means of interest rate subsidies.

Suppose that low interest capital has been plentiful, and there has been an overabundance of housing. Specifically, suppose recent historical data indicate that the best values for the parameters of the model in Figure 1.5 are $V_d = 1000$ apartments, $\alpha = 0.5$, and $l = 2$ years, and that reasonable initial conditions are $V(0) = 1500$ vacancies, and $S(-2) = S(-1) = 0$ apartments. Suppose that as a result of a new rent subsidy program we expect the future demand to be $R(n) = 500$ apartments, $n = 0,1,2,\ldots$. According to the mathematical model of (1.10)-(1.12) and Figure 1.5, the new rent subsidy program will cause the apartment vacancies in the city to exhibit the behavior shown in Table 1.1 and Figure 1.6.

Table 1.1 *Apartment Vacancies Predicted by Figure 1.5*

n	$V(n)$	$S(n)$	$C(n)$	$R(n)$	$\Delta V(n)$	$V(n+1)$
0	1500	0	0	500	−500	1000
1	1000	0	0	500	−500	500
2	500	250	0	500	−500	0
3	0	500	0	500	−500	−500
4	−500	750	250	500	−250	−750
5	−750	875	500	500	0	−750
6	−750	875	750	500	250	−500
7	−500	750	875	500	375	−125
8	−125	563	875	500	375	250
9	250	375	750	500	250	500
10	500	250	563	500	63	563
11	563	219	375	500	−125	438
12	438	281	250	500	−250	188

According to Figure 1.6, the model predicts that severe housing shortages will result from the new housing policy. If the model is correct, and if social pressures make the rent subsidy program mandatory, then the city officials must compensate for the policy by encouraging builders to expand the available housing. (Perhaps this expansion could be encouraged by publicizing the predicted housing shortage, or by having the city assume some of the risk of investment in new construction.)

If the model has not been carefully validated, however, the predictions that result from the model should be used with caution. The fact that builders themselves might predict future housing shortages is ignored in

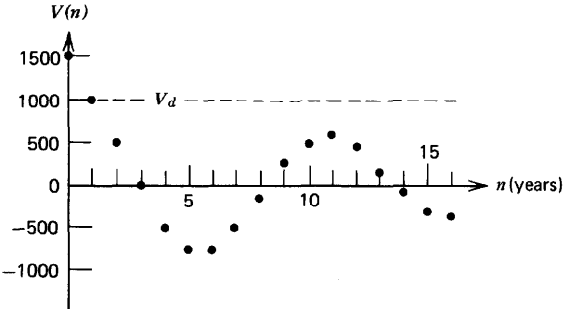


Figure 1.6. Apartment vacancies predicted by Figure 1.5.

(1.10). Thus this simple model of the relationship between vacancies and construction starts should probably be modified to more accurately describe the process by which builders decide to start new construction. Then the determination of model parameter values and the validation of the model should be repeated for the new model before it is used to predict the effect of housing policies.

The modeling process we have described has been used extensively to describe such situations as the flow of electric power in large transmission line networks and the growth of competing species in ecosystems. It is apparent that the same modeling process can be used to describe the relationships among the variables in many other types of systems. For example, it is suitable for describing the response of an eye pupil to variations in light intensity, the response of a banking system to market fluctuations, or the response of the people of a given country to variations in the world price of oil. It is in the social, economic, and biological fields that system modeling is likely to have its greatest impact in the future.

1.5 Solution of Linear Algebraic Equations

To this point our discussion has been of an introductory nature. The development of vector space concepts and the vector space language begins in Chapter 2. We now explore briefly, in a matrix format, the process of solving sets of linear algebraic equations, in order that we be able to use such sets of equations in the examples of Chapter 2 and later chapters. In this discussion we emphasize practical techniques for computing solutions to sets of linear algebraic equations and for computing the inverses of square matrices.

Models of most systems eventually lead to the formation and solution of sets of linear algebraic equations. For example, it is common practice to replace the derivatives in a differential equation by finite differences, thereby producing a set of linear algebraic equations which can be solved by a digital computer. The solution of nonlinear equations almost always requires linearization and, again, involves solution of linear algebraic equations (Chapter 8). Thus simultaneous algebraic equations are fundamental to practical analysis.

There is a wide variety of methods for solving a set of linear algebraic equations.* The design of *practical* computer algorithms which will obtain accurate solutions in an efficient manner calls upon most of the concepts of this book: spectral analysis, least-square optimization, orthogonalization, iteration, etc. Frequently, the sets of equations that arise in practice

*See Forsythe [1.6].

are nearly degenerate; that is, they border on being unsolvable by computers which have finite accuracy. Furthermore, the number of equations can be large; finite-difference approximations for partial differential equations sometimes involve more than 100,000 equations (P&C 2.17). Thus the solution of linear algebraic equations constitutes one of the easiest, and yet one of the most difficult problems.

Any set of linear algebraic equations can be written in the form

$$\begin{aligned} a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n &= \eta_1 \\ \vdots \qquad\qquad\qquad\qquad &\vdots \\ a_{m1}\xi_1 + a_{m2}\xi_2 + \cdots + a_{mn}\xi_n &= \eta_m \end{aligned} \tag{1.13}$$

Equation (1.13) easily fits the symbolic structure of the basic system model (1.1). Suppose we define $\mathbf{x} \overset{\Delta}{=} \{\xi_1, \xi_2, \ldots, \xi_n\}$ and $\mathbf{y} \overset{\Delta}{=} \{\eta_1, \eta_2, \ldots, \eta_m\}$ as the unknown inputs and known outputs, respectively, of the model, $\mathbf{T}$. Our immediate goal is to clarify the manner in which $\mathbf{T}$, by way of the coefficients a_{ij}, relates $\mathbf{x}$ to $\mathbf{y}$. Associated with (1.13) are three basic questions:

1. Do the equations possess a solution $\mathbf{x}$ for each given $\mathbf{y}$; that is, are the equations consistent?
2. Is the solution unique; that is, are there enough independent equations to determine $\mathbf{x}$?
3. What is the solution (or solutions)?

It is appropriate to ask the same questions concerning (1.1). Although the third question may appear to be the most pertinent for a specific problem, the answers to the other two give valuable insight into the structure of the model and its applicability to the situation it is supposed to represent. Such insight is generally the real reason for solving the equations, and certainly the prime purpose of our present analysis.

We rephrase the problem in matrix notation in order to separate the information about the system $\{a_{ij}\}$ from the information about the "state" or "condition" of the system (the variables $\{\xi_i\}$, $\{\eta_j\}$).

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} \tag{1.14}$$

Matrix multiplication is defined in such a way that (1.13) and (1.14) are equivalent.* The notation of (1.14) is close to the abstract symbolism of (1.1). In order to be more direct concerning the meaning of $\mathbf{T}$, we redefine $\mathbf{x}$ and $\mathbf{y}$ as the column matrices:

$$\mathbf{x} \triangleq \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \qquad \mathbf{y} \triangleq \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix}$$

Then (1.14) states

$$\mathbf{Ax} = \mathbf{y} \tag{1.15}$$

where $\mathbf{A}$ is the $m \times n$ matrix of equation coefficients. The system $\mathbf{T}$ can be defined explicitly by $\mathbf{Tx} \triangleq \mathbf{Ax}$; that is, the abstract operation of the system model $\mathbf{T}$ on the "vector" $\mathbf{x}$ is multiplication of $\mathbf{x}$ by the matrix $\mathbf{A}$.

Typical of the classical methods of solution of (1.15) is Cramer's formula (Appendix 1):

$$\xi_i = \frac{\det(\mathbf{A}(i))}{\det(\mathbf{A})}$$

where $\mathbf{A}(i)$ is the matrix $\mathbf{A}$ with its ith column replaced by $\mathbf{y}$. The formula applies only when $\mathbf{A}$ is square $(m = n)$ and $\det(\mathbf{A}) \neq 0$. The method indicates that for square $\mathbf{A}$, $\det(\mathbf{A}) \neq 0$ is a necessary and sufficient condition to guarantee a unique solution $\mathbf{x}$ to (1.15).

The most efficient scheme for evaluating a determinant requires approximately $n^3/3$ multiplications (Appendix 1 and P&C 1.3). Thus solution for $\mathbf{x}$ using Cramer's formula requires $(n + 1)n^3/3$ multiplications. Compared with other techniques, Cramer's formula is not a practical tool for analyzing linear equations.

Row Reduction

Ordinary elimination of variables forms the basis for an efficient method of solution to (1.15). In point of fact, it is the basis for most computer algorithms for solving sets of linear algebraic equations. In essence, the method consists in successively adding some multiple of one equation to another until only one variable remains in each equation; then we obtain

*See Appendix 1 for a brief introduction to matrices and determinants.

the unknowns by inspection. For example:

$$\begin{aligned} \xi_1 + 2\xi_2 &= 2 \\ 3\xi_1 + 4\xi_2 &= 6 \end{aligned} \quad \rightarrow \quad \begin{aligned} \xi_1 + 2\xi_2 &= 2 \\ -2\xi_2 &= 0 \end{aligned} \quad \rightarrow$$

$$\begin{aligned} \xi_1 + 2\xi_2 &= 2 \\ \xi_2 &= 0 \end{aligned} \quad \rightarrow \quad \begin{aligned} \xi_1 &= 2 \\ \xi_2 &= 0 \end{aligned}$$

The elimination method reduces to an automatable procedure (or algorithm) which requires no creative decision making by the user. Since the unknowns are unaffected by the procedure, they need not be written down; the above elimination process is expressed in matrix notation by

$$\begin{pmatrix} 1 & 2 & \vdots & 2 \\ 3 & 4 & \vdots & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & \vdots & 2 \\ 0 & -2 & \vdots & 0 \end{pmatrix} \rightarrow$$

$$\begin{pmatrix} 1 & 2 & \vdots & 2 \\ 0 & 1 & \vdots & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & \vdots & 2 \\ 0 & 1 & \vdots & 0 \end{pmatrix}$$

The first matrix in this elimination process is $(\mathbf{A} \vdots \mathbf{y})$; we call it the **augmented matrix** (we augmented $\mathbf{A}$ with $\mathbf{y}$). We refer to the matrix version of this elimination process as row reduction of the matrix $(\mathbf{A} \vdots \mathbf{y})$. Specifically, **row reduction of a matrix $\mathbf{B}$** consists in systematically operating on the rows of $\mathbf{B}$ as if they were equations until (a) the first nonzero element in each row is 1; (b) each column which contains the leading 1 for some row has all its other entries 0; and (c) the leading l's are in an order which descends from the left, with all zero rows at the bottom. We need the last requirement only to make the row-reduced matrix unique. We call the row-reduced matrix the **echelon form** (or Hermite normal form) of $\mathbf{B}$.

There are two basic techniques for row reducing a matrix. In **Gauss-Jordan elimination** we complete the operations on each column, obtaining a single 1 with all other elements 0, before concerning ourselves with succeeding columns (Example 1). In **Gaussian elimination** we first eliminate all elements below the main diagonal, one column at a time, thereby making the matrix "upper triangular." We then eliminate elements above the diagonal by a process commonly called "back substitution." In Example 2 the first three steps demonstrate the triangularization, the last two the back substitution. Although the two methods are similar, Gaussian elimination is 33% more efficient than Gauss-Jordan elimination for large sets of equations (say, $n > 5$); Gaussian elimination requires about $n^3/3$ multiplications to row reduce $(\mathbf{A} \vdots \mathbf{y})$ for an $n \times n$ matrix $\mathbf{A}$. Gauss-Jordan

elimination requires about $n^3/2$ multiplications. Both methods are far superior to Cramer's formula for solving linear algebraic equations (P&C 1.3).

Example 1. Gauss Jordan Elimination

$$\begin{pmatrix}1&2&2&1\\2&3&5&1\\3&2&5&1\end{pmatrix}\longrightarrow\begin{pmatrix}\textcircled{1}&2&2&1\\0&-1&1&-1\\0&-4&-1&-2\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&\textcircled{0}&4&-1\\0&\textcircled{1}&-1&1\\0&\textcircled{0}&-5&2\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&\textcircled{0}&\frac{3}{5}\\0&1&\textcircled{0}&\frac{3}{5}\\0&0&\textcircled{1}&-\frac{2}{5}\end{pmatrix}$$

Example 2. Gaussian Elimination

$$\begin{pmatrix}1&2&2&1\\2&3&5&1\\3&2&5&1\end{pmatrix}\longrightarrow\begin{pmatrix}\textcircled{1}&2&2&1\\0&-1&1&-1\\0&-4&-1&-2\end{pmatrix}\longrightarrow\begin{pmatrix}1&2&2&1\\0&\textcircled{1}&-1&1\\0&\textcircled{0}&-5&2\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&2&2&1\\0&1&\textcircled{-1}&1\\0&0&\textcircled{1}&-\frac{2}{5}\end{pmatrix}\longrightarrow\begin{pmatrix}1&2&\textcircled{0}&\frac{9}{5}\\0&1&\textcircled{0}&\frac{3}{5}\\0&0&1&-\frac{2}{5}\end{pmatrix}\longrightarrow\begin{pmatrix}1&\textcircled{0}&0&\frac{3}{5}\\0&1&0&\frac{3}{5}\\0&0&1&-\frac{2}{5}\end{pmatrix}$$

In the row reduction of small matrices by hand, the number of multiplications is of less concern than is accuracy. To guard against errors during row reduction of a matrix B, we can add a "check" column whose ith element is the sum of the elements in the ith row of **B**. Throughout the row-reduction process the ith element in the check column should remain equal to the sum of all other elements in the ith row; wherever it is not equal to that sum, one of the elements in that. row is in error. Because adding fractions by hand is complicated, we can avoid fractions by not forcing nonzero elements to be 1 until the last step in the row-reduction process.

Example 3. Row Reduction by Hand

$$(\mathbf{B}\ \vdots\ \text{check column})\triangleq\left(\begin{array}{ccc:c}3&1&2&6\\4&2&1&7\end{array}\right)\longrightarrow\left(\begin{array}{ccc:c}12&4&8&24\\12&6&3&21\end{array}\right)$$

$$\longrightarrow\left(\begin{array}{ccc:c}12&4&8&24\\0&2&-5&-3\end{array}\right)\longrightarrow\left(\begin{array}{ccc:c}6&2&4&12\\0&2&-5&-3\end{array}\right)\longrightarrow\left(\begin{array}{ccc:c}6&0&9&15\\0&2&-5&-3\end{array}\right)$$

$$\longrightarrow\left(\begin{array}{ccc:c}1&0&\frac{3}{2}&\frac{5}{2}\\0&1&-\frac{5}{2}&-\frac{3}{2}\end{array}\right)$$

If we are interested in the solution to a set of equations $\mathbf{Ax} = \mathbf{y}$ as a function of $\mathbf{y}$, we can carry an unspecified $\mathbf{y}$ through the row-reduction process.

Example 4. Row Reduction with an Unspecified Column

$$(\mathbf{A} \vdots \mathbf{y}) \triangleq \left(\begin{array}{ccc:c} 1 & 2 & 2 & \eta_1 \\ 2 & 3 & 5 & \eta_2 \\ 3 & 2 & 5 & \eta_3 \end{array}\right) \longrightarrow \left(\begin{array}{ccc:c} 1 & 2 & 2 & \eta_1 \\ 0 & -1 & 1 & \eta_2 - 2\eta_1 \\ 0 & -4 & -1 & \eta_3 - 3\eta_1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc:c} 1 & 0 & 4 & -3\eta_1 + 2\eta_2 \\ 0 & 1 & -1 & 2\eta_1 - \eta_2 \\ 0 & 0 & -5 & 5\eta_1 - 4\eta_2 + \eta_3 \end{array}\right) \longrightarrow \left(\begin{array}{ccc:c} 1 & 0 & 0 & \eta_1 - \frac{6}{5}\eta_2 + \frac{4}{5}\eta_3 \\ 0 & 1 & 0 & \eta_1 - \frac{1}{5}\eta_2 - \frac{1}{5}\eta_3 \\ 0 & 0 & 1 & -\eta_1 + \frac{4}{5}\eta_2 - \frac{1}{5}\eta_3 \end{array}\right)$$

The solution to the equations represented by the matrix $(\mathbf{A} \vdots \mathbf{y})$ of Example 4 can be expressed

$$\mathbf{x} = \begin{pmatrix} \eta_1 - \frac{6}{5}\eta_2 + \frac{4}{5}\eta_3 \\ \eta_1 - \frac{1}{5}\eta_2 - \frac{1}{5}\eta_3 \\ -\eta_1 + \frac{4}{5}\eta_2 - \frac{1}{5}\eta_3 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{6}{5} & \frac{4}{5} \\ 1 & -\frac{1}{5} & -\frac{1}{5} \\ -1 & \frac{4}{5} & -\frac{1}{5} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}$$

Clearly, the final coefficients on the variables $\{\eta_i\}$ constitute the inverse matrix $\mathbf{A}^{-1}$. The coefficients which multiply these variables during the row reduction keep a record of the elimination operations on the rows of $\mathbf{A}$. The variables $\{\eta_i\}$ merely serve to keep the coefficients separated. The row reduction of Example 4 was, in effect, performed on $(\mathbf{A} \vdots \mathbf{I})$ to obtain $(\mathbf{I} \vdots \mathbf{A}^{-1})$, where $\mathbf{I}$ is the identity matrix; that is,*

$$\left(\begin{array}{ccc:ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 2 & 3 & 5 & 0 & 1 & 0 \\ 3 & 2 & 5 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 1 & -\frac{6}{5} & \frac{4}{5} \\ 0 & 1 & 0 & 1 & -\frac{1}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & -1 & \frac{4}{5} & -\frac{1}{5} \end{array}\right)$$

Row reduction is an efficient method for computing $\mathbf{A}^{-1}$. Yet in most instances, computation of $\mathbf{A}^{-1}$ is, in itself, inefficient, Computing $\mathbf{A}^{-1}$ by using Gaussian elimination on $(\mathbf{A} \vdots \mathbf{I})$ requires $\frac{4}{3}n^3$ multiplications for an $n \times n$ matrix $\mathbf{A}$ (P&C 1.3). Since this is four times the number of multiplications needed to find the solution $\mathbf{x}$ for a given $\mathbf{y}$, we find the inverse only when we actually need it—when we are interested in the properties of the system model (the set of equations) and of the matrix $\mathbf{A}$ which represents it.

*In Appendix 1, $\mathbf{A}^{-1}$ is defined as a matrix which satisfies $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. In P&C 1.4 we find that if such a matrix exists, the row reduction of $(\mathbf{A} \vdots \mathbf{I})$ will produce it.

Many system models lead to matrices which are not square; there can be more equations than unknowns; there can be fewer. Even if the matrix is square, its inverse need not exist. Yet for any $m \times n$ matrix $\mathbf{A}$, row reduction of $(\mathbf{A} \vdots \mathbf{I})$ yields complete information about the equation $\mathbf{Ax} = \mathbf{y}$, including answers to the questions of existence and uniqueness of the solutions (P&C 1.1, 1.2).

Example 5. Solution by Row Reduction-a Nonsquare Matrix. Suppose we obtain the following equations from three independent measurements of some quantity

$$\xi_1 + \xi_2 = 1.2$$
$$\xi_1 + \xi_2 = 1.3$$
$$\xi_1 + \xi_2 = 1.2$$

Then

$$(\mathbf{A} \vdots \mathbf{I}) = \left(\begin{array}{ccc:ccc} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

which we row reduce to

$$\left(\begin{array}{ccc:ccc} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{array}\right)$$

We interpret the row reduced matrix to mean

$$\xi_1 + \xi_2 = \eta_1$$
$$0 = \eta_2 - \eta_1$$
$$0 = \eta_3 - \eta_1$$

Unless $\eta_1 = \eta_2 = \eta_3$, the equations allow no solution. In our example the equations are not consistent; $\eta_1 = \eta_3 = 1.2$, but $\eta_2 = 1.3$. If the equations were consistent, the row-reduced equations indicate that the solution would not be unique; for example, if η_2 were 1.2, the solution would be

$$\xi_1 + \xi_2 = \eta_1$$

Row and Column Interpretations

We have, to this point, viewed the matrix multiplication in (1.14) as the operation of the system on $\mathbf{x}$ to produce $\mathbf{y}$. This interpretation is expressed in (1.15). We now suggest two more interpretations that will be useful

throughout our discussions of modeling. It is apparent from (1.14) and (1.15) that the columns of the matrix $\mathbf{A}$ are in some sense similar to $\mathbf{y}$; they both contain the same number (m) of elements. We call them **column vectors of A**, and denote the jth column vector by $\mathbf{A}_{(j)}$. Again, the rows of $\mathbf{A}$ are similar to $\mathbf{x}$, both containing n elements; we denote the ith **row vector of A** by $\mathbf{A}^{(i)}$. If we focus on the column vectors of $\mathbf{A}$, (1.14) becomes

$$\xi_1\mathbf{A}_{(1)} + \xi_2\mathbf{A}_{(2)} + \cdots + \xi_n\mathbf{A}_{(n)} = \mathbf{y} \tag{1.16}$$

That is, $\mathbf{y}$ is a simple combination of the column vectors of $\mathbf{A}$; the elements of $\mathbf{x}$ specify the combination. We will make use of this column vector interpretation in Section 2.2 and thereafter.

Changing our focus to the row vectors of $\mathbf{A}$, (1.14) becomes

$$\begin{aligned} \mathbf{A}^{(1)}\mathbf{x} &= \eta_1 \\ \mathbf{A}^{(2)}\mathbf{x} &= \eta_2 \\ &\vdots \qquad \vdots \\ \mathbf{A}^{(m)}\mathbf{x} &= \eta_m \end{aligned} \tag{1.17}$$

Each element of $\mathbf{y}$ is determined by the corresponding row vector of $\mathbf{A}$. By this interpretation, we are merely focusing separately on each of the equations of (1.13). We can use the geometrical pictures of analytic geometry to help develop a physical feel for the individual algebraic equations of (1.17). Suppose

$$\mathbf{A}\mathbf{x} = \begin{pmatrix} 2 & 1 \\ 2 & 1+\epsilon \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \tag{1.18}$$

where ϵ is some constant. The 2×1 matrix $\mathbf{x}$ and the 1×2 matrices $\mathbf{A}^{(i)}$ are each equivalent to a vector (or arrow) in a plane. We simply pick coordinate axes and associate with each element of $\mathbf{x}$ or $\mathbf{A}^{(i)}$ a component along one of the axes. Thus we can represent (1.18) geometrically as in Figure 1.7. The vectors $\mathbf{x}$ such that

$$\mathbf{A}^{(1)}\mathbf{x} = \text{a constant}$$

terminate on a line perpendicular to the vector $\mathbf{A}^{(1)}$. The solution $\mathbf{x}$ to the pair of equations lies at the intersection of the lines $\mathbf{A}^{(1)}\mathbf{x} = 2$ and $\mathbf{A}^{(2)}\mathbf{x} = 3$. Since the lines in Figure 1.7 have a well-defined intersection, the equations of (1.18) possess a well-defined (unique) solution. However, if $\epsilon \to 0$, $\mathbf{A}^{(2)} \to \mathbf{A}^{(1)}$ and the system becomes degenerate; the lines become parallel, the equations become inconsistent, and there is no solution (intersection). If

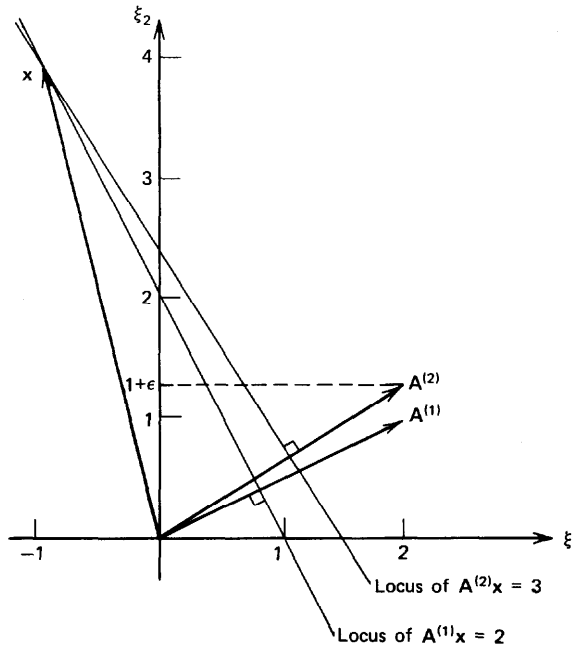


Figure 1.7. Row vector interpretation of (1.18) for $\epsilon = 0.25$.

the numbers on the right side of (1.18) were equal, the lines would overlap, the equations would be consistent, but the solution would not be unique—any $\mathbf{x}$ terminating on the common line would satisfy both equations.

The geometrical example of (1.18) and Figure 1.7 introduces a significant computational difficulty which exists in nearly degenerate systems of equations. Slight changes in the numbers on the right side of (1.18) result in slight shifts in the positions of the lines in Figure 1.7. Slight changes in the equation coefficients cause slight tilts in these lines. If ϵ is nearly zero, the lines are nearly parallel, and slight perturbations in the line positions or angles cause large swings in the intersection (or solution) $\mathbf{x}$. A solution to a matrix equation which is very sensitive to small changes (or errors) in the data is called an **unstable solution.** A matrix (or the corresponding set of equations) which leads to an unstable solution is said to be **ill-conditioned.** Assume the matrix is normalized so that the magnitude of its largest element is approximately one. Then the magnitudes of the elements of the inverse matrix indicate the degree of sensitivity of the solution $\mathbf{x}$ of (1.14) to errors in the data, $\{ a_{ij} \}$ or $\{ \eta_i \}$. In Section 6.6 we define a condition number which indicates the size of the largest elements of the inverse. A

very large condition number implies that the matrix is ill-conditioned. The size of $\det(\mathbf{A})$ is another indication of the ill-conditioning of the equations; as the equations become more degenerate, $\det(\mathbf{A})$ must approach zero (P&C 1.6). However, $\det(\mathbf{A})$ is not an absolute measure of ill-conditioning as is the condition number.

Numerical Error

There are two fundamental sources of error in the solution to a set of linear algebraic equations, measurement error and computer roundoff. When the data that are used to make up a set of equations come from physical measurements, these data usually contain empirical error. Even if the data are exact, however, the numbers are rounded by the computer; the data can be represented only to a finite number of significant digits. Thus inaccuracies in the equation data are the rule, not the exception. As computations are carried out, further rounding occurs. Although individual inaccuracies are slight, their cumulative effect can be disastrous if handled carelessly.

The following example demonstrates that slight errors in the data can be vastly magnified by straightforward use of row-reduction techniques. Let

$$(\mathbf{A} \vdots \mathbf{y}) = \left(\begin{array}{ccc:c} 2 & 1 & 3 & 1 \\ 2 & 1.01 & 1 & 2 \\ 2 & 3 & 2 & 3 \end{array}\right) \tag{1.19}$$

Suppose the element a_{22} is in error by 0.5%; that is, $a_{22} = 1.01 \pm 0.005$. Elimination operations on the first column reduces (1.19) to

$$\left(\begin{array}{ccc:c} 2 & 1 & 3 & 1 \\ 0 & 0.01 & -2 & 1 \\ 0 & 2 & -1 & 2 \end{array}\right) \tag{1.20}$$

where the subtraction of two nearly equal numbers has magnified the error at the element in question to about 50%, that is, the new element in row 2, column 2, is 0.01 ± 0.005. Were we to use this element to eliminate the other elements in column 2, we would propagate this 50% error throughout the matrix; that is, we would obtain

$$\left(\begin{array}{ccc:c} 2 & 0 \mp 0.5 & 203 \pm 100 & -99 \mp 50 \\ 0 & 1 \pm 0.5 & -200 \mp 100 & 100 \pm 50 \\ 0 & 0 \mp 1 & 399 \pm 200 & -198 \mp 100 \end{array}\right) \tag{1.21}$$

Further computations would be meaningless. Fortunately, we do not need to divide by the inaccurate element. We merely interchange rows 2 and 3

in (1.20) to obtain

$$\left(\begin{array}{ccc:c} 2 & 1 & 3 & 1 \\ 0 & 2 & -1 & 2 \\ 0 & 0.01 & -2 & 1 \end{array}\right) \tag{1.22}$$

This interchange is equivalent to writing the equations in a different order. We now use the larger and more accurate element "2" of row 2, column 2 to eliminate the other elements in column 2:

$$\left(\begin{array}{ccc:c} 4 & 0 & 7 & 0 \\ 0 & 2 & -1 & 2 \\ 0 & 0 \pm 0.005 & -1.995 & 0.99 \end{array}\right) \tag{1.23}$$

The element moved into position for elimination of other elements in its column is called a pivot. The process of interchanging rows to avoid division by relatively small (and therefore inaccurate) numbers is called **pivoting** or **positioning for size.** We also can move the inaccurate element from row 2, column 2 of (1.20) by interchanging *columns* 2 and 3 if we change the order of the variables ξ_2 and ξ_3 which multiply these columns. This column interchange is also used in pivoting. All good computer algorithms for solving sets of linear algebraic equations or for inverting square matrices use some form of pivoting to minimize the magnification and propagation of errors in the data. Scaling of the equations is also an important part of these algorithms.

The matrix of (1.19) is not ill-conditioned. It is apparent, therefore, that we must compute solutions carefully, regardless of the conditioning of the equations, if we are to avoid magnification of errors. If the equations are ill-conditioned, however, careful computing (scaling and pivoting) and the use of double precision arithmetic (additional significant digits) are crucial. Furthermore, division by small numbers is inevitable at some point in the process of solving ill-conditioned equations, and errors *will* be magnified. An iterative technique for improving the computed solution to a set of ill-conditioned equations is described in P&C 1.5.

If a set of equations is very ill-conditioned, it may be that the underlying system is degenerate. Perhaps the matrix would be singular, were it not for empirical error in the data. (That is, perhaps ϵ should be zero in (1.18) and Figure 1.7.) Then in order to completely solve the set of equations, we not only need to compute a particular solution $\mathbf{x}$ as described above, but we also need to estimate the full set of "near solutions" (the locus of the "nearly-overlapping" lines of Figure 1.7 for $\epsilon = 0$). We describe a technique for computing this set of "near solutions" in Section 2.4. Further informa-

tion on the solution of linear algebraic equations is contained in Forsythe and Moler [1.4] and Forsythe [1.5].

1.6 Problems and Comments

*1.1 *Exploring matrix equations by row reduction:* let $\mathbf{A}$ be an $m \times n$ matrix. Row reduction of $(\mathbf{A} \vdots \mathbf{y})$ for an unspecified column vector $\mathbf{y} = (\eta_1 \cdots \eta_m)^T$, or the equivalent row reduction of $(\mathbf{A} \vdots \mathbf{I})$ for an $m \times m$ matrix $\mathbf{I}$, determines the conditions which must be satisfied by $\mathbf{y}$ in order for the equation $\mathbf{Ax} = \mathbf{y}$ to have a solution; the set of vectors $\mathbf{y}$ for which a solution $\mathbf{x}$ exists is called the **range of A**. The same row reduction determines the set of solutions $\mathbf{x}$ for $\mathbf{y} = (0 \cdots 0)^T$; this set of solutions is referred to as the **nullspace of A**. If the nullspace of $\mathbf{A}$ contains nonzero vectors, the solutions to $\mathbf{Ax} = \mathbf{y}$ cannot be unique. Let the matrix equation be

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 2 & 1 & 1 & 3 \\ 4 & 5 & 3 & 9 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$$

(*a*) Row reduce $(\mathbf{A} \vdots \mathbf{I})$.

(*b*) Determine the range of $\mathbf{A}$; that is, determine the relationships that must exist among the elements $\{\eta_i\}$ of $\mathbf{y}$ in order for the matrix equation $\mathbf{Ax} = \mathbf{y}$ to have a solution.

(*c*) Determine the nullspace of $\mathbf{A}$.

(*d*) Determine the solutions $\mathbf{x}$ for the specified right-hand side $\mathbf{y}$.

(*e*) Give an example of a matrix equation that is both inconsistent and underdetermined; that is, an equation for which $\mathbf{y}$ is not in the range of $\mathbf{A}$ and for which the nullspace of $\mathbf{A}$ is nonzero.

1.2 Use the row-reduction technique to determine the solutions to the following sets of equations:

(*a*)
$$\begin{aligned} \xi_1 + 6\xi_2 - 18\xi_3 &= 0 \\ -4\xi_1 \qquad + 5\xi_3 &= 0 \\ -3\xi_1 + 6\xi_2 - 13\xi_3 &= 0 \\ -7\xi_1 + 6\xi_2 \quad - 8\xi_3 &= 0 \end{aligned}$$

$$(b)\qquad \begin{aligned} 2\xi_1+3\xi_2+4\xi_3&=9\\ \xi_1+\xi_2+\xi_3&=3\\ 3\xi_1+2\xi_2+2\xi_3&=7 \end{aligned}$$

$$(c)\qquad \begin{aligned} 2\xi_1+3\xi_2+4\xi_3&=9\\ 3\xi_1+4\xi_2+5\xi_3&=12\\ 4\xi_1+3\xi_2+3\xi_3&=10\\ 5\xi_1+5\xi_2+6\xi_3&=10 \end{aligned}$$

$$(d)\qquad \begin{aligned} \xi_1-2\xi_3&=\eta_1\\ 2\xi_1+2\xi_2&=\eta_2\\ 2\xi_1-4\xi_3&=\eta_3\\ \xi_1+\xi_2+3\xi_3&=\eta_4 \end{aligned}$$

1.3 *Efficiency of computations:* the number of multiplications performed during a computation is a measure of the efficiency of a computational technique. Let $\mathbf{A}$ be an invertible $n \times n$ matrix. Determine the number of multiplications required:

(*a*) To compute $\mathbf{A}^{-1}$ by Gaussian elimination;

(*b*) To compute $\mathbf{A}^{-1}$ by Gauss-Jordan elimination;

(*c*) To compute $\det(\mathbf{A})$, using Gaussian elimination to triangularize $\mathbf{A}$ (Example 2, Appendix 1).

Determine the number of multiplications required to solve $\mathbf{Ax} = \mathbf{y}$ for a specific vector $\mathbf{y}$ by:

(*d*) Cramer's rule [Hint: use the answer to (*c*)].

(*e*) The computation in (*a*) and the multiplication $\mathbf{A}^{-1}\mathbf{y}$;

(*f*) Direct row reduction of $(\mathbf{A} \vdots \mathbf{y})$.

1.4 *Elementary matrices:* the row reduction of an $m \times n$ matrix $\mathbf{A}$ consists in performing elementary operations on the rows of $\mathbf{A}$. Each such operation is equivalent to the multiplication of $\mathbf{A}$ by a simple $m \times m$ matrix which we refer to as an **elementary matrix.**

(*a*) For $m = 5$, find the elementary matrices corresponding to the following:

(1) the multiplication of row 3 by a constant c;

(2) the addition of row 4 to row 1;

(3) the interchange of row 3 with row 5.

(*b*) Every elementary matrix is invertible. Find the inverses of the elementary matrices determined in (*a*).

(*c*) The row reduction of $(\mathbf{A} \vdots \mathbf{I})$ is equivalent to multiplication of $(\mathbf{A} \vdots \mathbf{I})$ by an invertible matrix $\mathbf{B}$ (a product of elementary matrices). Show that if $\mathbf{A}$ is square and $(\mathbf{A} \vdots \mathbf{I})$ can be row reduced to the form $(\mathbf{I} \vdots \mathbf{B})$, then $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, and therefore $\mathbf{B} = \mathbf{A}^{-1}$.

1.5 *Iterative improvement of solutions:* the solution to the matrix equation $\mathbf{Ax} = \mathbf{y}$ can be obtained by Gaussian elimination. As a result of roundoff, the computed solution $\mathbf{x}_1$ is usually in error. Denote the error by $\mathbf{x} - \mathbf{x}_1$, where $\mathbf{x}$ is the exact solution. A computable measure of the error is the residual $r_1 \overset{\Delta}{=} \mathbf{y} - \mathbf{Ax}_1$. If we could solve exactly for $(\mathbf{x} - \mathbf{x}_1)$ in the equation $\mathbf{A}(\mathbf{x} - \mathbf{x}_1) = \mathbf{y} - \mathbf{Ax}_1 = r_1$, we could obtain the exact solution. We solve $\mathbf{Az}_1 = r_1$ by Gaussian elimination to obtain a correction $\mathbf{z}_1$; $\mathbf{x}_2 \overset{\Delta}{=} \mathbf{x}_1 + \mathbf{z}_1$ is an improved solution. By repeating the improvement process iteratively, we obtain an approximate solution which is accurate to the number of significant digits used in the computation. However, the residuals $\mathbf{r}_k = \mathbf{y} - \mathbf{Ax}_k$ must be computed to double precision; otherwise the corrections, $\mathbf{z}_k$, will not be improvements. See Forsythe and Moler [1.4, p. 49]. Let

$$\mathbf{A} = \begin{pmatrix} 2.1 & 1.9 \\ 1.9 & 2.0 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} 1.2 \\ 1.3 \end{pmatrix}$$

To five figures, the solution to $\mathbf{Ax} = \mathbf{y}$ is $\mathbf{x} = (-0.11864\ 0.76271)^{\mathrm{T}}$.

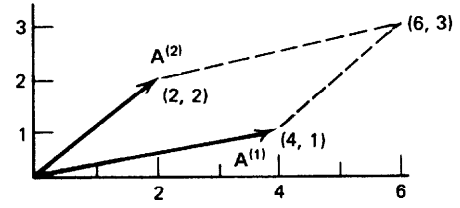
(*a*) Compute an approximate solution $\mathbf{x}_1$ by Gaussian elimination, rounding all computations to three significant digits (slide rule accuracy).

(*b*) Find the residual r_1 by hand computation to *full* accuracy.

(*c*) Round r_1 to three significant digits, if necessary, and compute the correction $\mathbf{z}_1$. Find $\mathbf{x}_2 = \mathbf{x}_1 + \mathbf{z}_1$.

1.6 *Determinants and volumes:* using a natural correspondence between row vectors and arrows in a plane, we associate a parallelogram with the rows of every real 2 *x* 2 matrix $\mathbf{A}$. For example,

$$\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 2 & 2 \end{pmatrix} \Leftrightarrow$$



(*a*) Show that the area of the above parallelogram is equal to the determinant of the matrix A which is associated with it.

(*b*) For the right-hand coordinate system shown above, we define the *sign of the area* to be positive if $\mathbf{A}^{(1)}$ turns counterclockwise inside the parallogram in order to reach $\mathbf{A}^{(2)}$; if $\mathbf{A}^{(1)}$ turns clockwise, the area is negative. Show graphically that the area of the above parallelogram obeys the following properties of determinants:

(1) The value of $\det(\mathbf{A})$ is not changed if we add to one row of $\mathbf{A}$ a multiple of another row of $\mathbf{A}$;

(2) The sign of $\det(\mathbf{A})$ is reversed if we interchange two rows of $\mathbf{A}$;

(3) If we multiply one row of $\mathbf{A}$ by c, then $\det(\mathbf{A})$ is multiplied by c;

(4) If the rows of $\mathbf{A}$ are dependent (i.e., one is a multiple of the other), then $\det(\mathbf{A}) = 0$.

(*c*) The geometrical interpretation of $\det(\mathbf{A})$ can be extended to $n \times n$ matrices by defining n-dimensional spaces, n-dimensional parallelepipeds, and signed volumes. See Martin and Mizel [1.9]. Since $\det(\mathbf{A}^T) = \det(\mathbf{A})$, the volume of the parallelopiped described by the columns of $\mathbf{A}$ equals the volume described by the rows of $\mathbf{A}$. Verify graphically that the geometrical interpretation of determinants extends to 3×3 matrices.

1.7 *Partitioned matrices:* it is sometimes useful to partition a matrix into an array of submatrices. If $\mathbf{P}$ and $\mathbf{Q}$ are conformable, we can form the partitions

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix} \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix}$$

in a manner which allows us to express $\mathbf{PQ}$ as

$$\mathbf{PQ} = \begin{pmatrix} \mathbf{P}_{11}\mathbf{Q}_{11} + \mathbf{P}_{12}\mathbf{Q}_{21} & \mathbf{P}_{11}\mathbf{Q}_{12} + \mathbf{P}_{12}\mathbf{Q}_{22} \\ \mathbf{P}_{21}\mathbf{Q}_{11} + \mathbf{P}_{22}\mathbf{Q}_{21} & \mathbf{P}_{21}\mathbf{Q}_{12} + \mathbf{P}_{22}\mathbf{Q}_{22} \end{pmatrix}$$

(*a*) Assume that $\mathbf{A}$ is an invertible matrix. The following factorization can be verified by the block multiplication described above:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$$

(b) Show that for any submatrix $\mathbf{P}$ of appropriate dimensions,

$$\begin{vmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{P} & \mathbf{I} \end{vmatrix} = 1$$

Use this result with (a) to show that

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}|$$

(c) Use (a) to show that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{O} \\ \mathbf{O} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{pmatrix}$$

The number of multiplications required to compute the determinant or the inverse of an $n \times n$ matrix can be reduced by a factor of eight (if n is large) by use of the partitioning schemes in (b) or (c), respectively.

1.7 References

[1.1] Bruner, Jerome S., *The Process of Education,* Harvard University Press, Cambridge, Mass., 1960.

[1.2] Cannon, Robert H., Jr., *Dynamics of Physical Systems,* McGraw-Hill, New York, 1969.

[1.3] Forrester, Jay W., *Urban Dynamics,* M.I.T. Press, Cambridge, Mass., 1969.

*[1.4] Forsythe, George E. and Cleve B. Moler, *Computer Solution of Linear Algebraic Systems,* Prentice-Hall, Englewood Cliffs, N.J., 1967.

[1.5] Forsythe, George E., "Today's Computational Methods of Linear Algebra," *SIAM Rev.,* **9,** 3 (July 1967), 489-515.

[1.6] Forsythe, George E., "Solving Linear Algebraic Equations Can be Interesting," *Bull. Am. Math. Soc.,* ***59*** (1953), 299-329.

[1.7] Linvill, William K., "Models and Model Construction," *IRE Trans. Educ.,* **E-5,** 2 (June 1962), 64-67.

[1.8] Meier, Robert C., William T. Newell, and Harold L. Pazer, *Simulation in Business and Economics,* Prentice-Hall, Englewood Cliffs, N.J., 1969.

[1.9] Martin, Allan D. and Victor J. Mizel, *Introduction to Linear Algebra,* McGraw-Hill, New York, 1966.

[1.10] Sage, Andrew P. and James L. Melsa, *System Identification,* Academic Press, New York, 1971.

[1.11] Truxal, John G., *Introduction to Systems Engineering,* McGraw-Hill, New York, 1972.