Introduction to Dynamic Systems

Theory, Models, and Applications

David G. Luenberger

Introduction to Dynamic Systems

Theory, Models, and Applications

David G. Luenberger

Stanford University

John Wiley & Sons

New York Chichester Brisbane Toronto

Library of Congress Cataloging in Publication Data

Luenberger, David G., 1937–
Introduction to dynamic systems.

Includes bibliographical references and index.
1. System analysis. 2. Differential equations.
3. Control theory. I. Title. II. Title: Dynamic Systems.

QA402.L84 003 78-12366
ISBN 0-471-02594-1

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Preface

This book is an outgrowth of a course developed at Stanford University over the past five years. It is suitable as a self-contained textbook for second-level undergraduates or for first-level graduate students in almost every field that employs quantitative methods. As prerequisites, it is assumed that the student *may* have had a first course in differential equations and a first course in linear algebra or matrix analysis. These two subjects, however, are reviewed in Chapters 2 and 3, insofar as they are required for later developments.

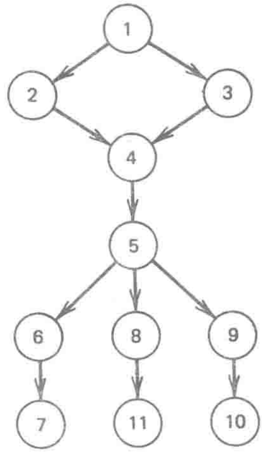
The objective of the book, simply stated, is to help one develop the ability to analyze real dynamic phenomena and dynamic systems. This objective is pursued through the presentation of three important aspects of dynamic systems: (1) the *theory*, which explores properties of mathematical representations of dynamic systems, (2) example *models*, which demonstrate how concrete situations can be translated into appropriate mathematical representations, and (3) *applications*, which illustrate the kinds of questions that might be posed in a given situation, and how theory can help resolve these questions. Although the highest priority is, appropriately, given to the orderly presentation of the theory, significant samples of all three of these essential ingredients are contained in the book.

The organization of the book follows theoretical lines—as the chapter titles indicate. The particular theoretical approach, or style, however, is a blend of the traditional approach, as represented by many standard textbooks on differential equations, and the modern state-space approach, now commonly used as a setting for control theory. In part, this blend was selected so as to broaden the scope—to get the advantages of both approaches; and in part it was dictated by the requirements of the applications presented. It is recognized, however, that (as in every branch of mathematics) the root ideas of dynamic systems transcend any particular mathematical framework used to describe those ideas. Thus, although the theory in this book is presented within a certain framework, it is the intent that what is taught about dynamic systems is richer and less restrictive than the framework itself.

The content of the book is, of course, partly a reflection of personal taste, but in large portion it was selected to directly relate to the primary objective of developing the ability to analyze real systems, as stated earlier. The theoretical material in Chapters 2 through 5 is quite standard, although in addition to theory these chapters emphasize the relation between theory and analysis. Dominant eigenvector analysis is used as an extended illustration of this relationship. Chapter 6 extends the classical material of linear systems to the special and rich topic of positive systems. This chapter, perhaps more than any other, demonstrates the intimate relation between theory and intuition. The topic of Markov chains, in Chapter 7, has traditionally been treated most often as a distinct subject. Nevertheless, although it does have some unique features, a great deal of unity is achieved by regarding this topic as a branch of dynamic system theory. Chapter 8 outlines the concepts of system control—from both the traditional transform approach and the state-space approach. Chapters 9 and 10 treat nonlinear systems, with the Liapunov function concept serving to unify both the theory and a wide assortment of applications. Finally, Chapter 11 surveys the exciting topic of optimal control—which represents an important framework for problem formulation in many areas. Throughout all chapters there is an assortment of examples that not only illustrate the theory but have intrinsic value of their own. Although these models are abstractions of reality, many of these are "classic" models that have stood the test of time and have had great influence on scientific development. For developing effectiveness in analysis, the study of these examples is as valuable as the study of theory.

The book contains enough material for a full academic year course. There is room, however, for substantial flexibility in developing a plan of study. By omitting various sections, the book has been used at Stanford as the basis for a six-month course. The chapter dependency chart shown below can be used to plan suitable individual programs. As a further aid to this planning, difficult sections of the book that are somewhat tangential to the main development are designated by an asterisk*.

An important component of the book is the set of problems at the end of the chapters. Some of these problems are exercises, which are more or less straightforward applications of the techniques discussed in the chapter; some are extensions of the theory; and some introduce new application areas. A few



Chapter Dependency Chart (A chapter depends on all chapters leading to it in the chart.)

of each type should be attempted from each chapter. Especially difficult problems are marked with an asterisk*.

The preparation of this book has been a long task that could not have been completed without the help of many individuals. Many of the problems and examples in the book were developed jointly with teaching assistants and students. I wish to acknowledge the Department of Engineering-Economic Systems at Stanford which provided the atmosphere and resources to make this project possible. I wish to thank my family for their help, encouragement, and endurance. I wish to thank Lois Goularte who efficiently typed the several drafts and helped organize many aspects of the project. Finally, I wish to thank the scores of students, visitors, and colleagues who read primitive versions of the manuscript and made many valuable individual suggestions.

DAVID G. LUENBERGER
Stanford, California
January 1979

Contents

chapter 1.
Introduction

1.1 DYNAMIC PHENOMENA

The term *dynamic* refers to phenomena that produce time-changing patterns, the characteristics of the pattern at one time being interrelated with those at other times. The term is nearly synonymous with *time-evolution* or *pattern of change*. It refers to the unfolding of events in a continuing evolutionary process.

Nearly all observed phenomena in our daily lives or in scientific investigation, have important dynamic aspects. Specific examples may arise in (a) a physical system, such as a traveling space vehicle, a home heating system, or in the mining of a mineral deposit; (b) a social system, such as the movement within an organizational hierarchy, the evolution of a tribal class system, or the behavior of an economic structure; or (c) a life system, such as that of genetic transference, ecological decay, or population growth. But while these examples illustrate the pervasiveness of dynamic situations and indicate the potential value of developing the facility for representing and analyzing dynamic behavior, it must be emphasized that the general concept of dynamics transcends the particular origin or setting of the process.

Many dynamic systems can be understood and analyzed intuitively, without resort to mathematics and without development of a general theory of dynamics. Indeed, we often deal quite effectively with many simple dynamic situations in our daily lives. However, in order to approach unfamiliar complex situations efficiently, it is necessary to proceed systematically. Mathematics can provide the required economy of language and conceptual framework.

With this view, the term *dynamics* soon takes on somewhat of a dual meaning. It is, first, as stated earlier, a term for the time-evolutionary phenomena in the world about us, and, second, it is a term for that part of mathematical science that is used for the representation and analysis of such phenomena. In the most profound sense the term refers simultaneously to both aspects: the real, the abstract, and the interplay between them.

Although there are endless examples of interesting dynamic situations arising in a spectrum of areas, the number of corresponding general forms for mathematical representation is relatively small. Most commonly, dynamic systems are represented mathematically in terms of either differential or difference equations. Indeed, this is so much the case that, in terms of pure mathematical content, at least the elementary study of dynamics is almost synonymous with the theory of differential and difference equations. It is these equations that provide the structure for representing time linkages among variables.

The use of either differential or difference equations to represent dynamic behavior corresponds, respectively, to whether the behavior is viewed as occurring in continuous or discrete time. Continuous time corresponds to our usual conception, where time is regarded as a continuous variable and is often viewed as flowing smoothly past us. In mathematical terms, continuous time of this sort is quantified in terms of the continuum of real numbers. An arbitrary value of this continuous time is usually denoted by the letter t. Dynamic behavior viewed in continuous time is usually described by differential equations, which relate the derivatives of a dynamic variable to its current value.

Discrete time consists of an ordered sequence of points rather than a continuum. In terms of applications, it is convenient to introduce this kind of time when events and consequences either occur or are accounted for only at discrete time periods, such as daily, monthly, or yearly. When developing a population model, for example, it may be convenient to work with yearly population changes rather than continuous time changes. Discrete time is usually labeled by simply indexing, in order, the discrete time points, starting at a convenient reference point. Thus time corresponds to integers 0, 1, 2, and so forth, and an arbitrary time point is usually denoted by the letter k. Accordingly, dynamic behavior viewed in discrete time is usually described by equations relating the value of a variable at one time to the values at adjacent times. Such equations are called *difference equations.*

1.2 MULTIVARIABLE SYSTEMS

The term *system*, as applied to general analysis, was originated as a recognition that meaningful investigation of a particular phenomenon can often only be achieved by explicitly accounting for its environment. The particular variables of interest are likely to represent simply one component of a complex, consisting of perhaps several other components. Meaningful analysis must consider the entire system and the relations among its components. Accordingly, mathematical models of systems are likely to involve a large number of interrelated variables—and this is emphasized by describing such situations as *multivariable systems.* Some examples illustrating the pervasiveness and importance of multivariable phenomena arise in consideration of (a) the migration patterns of population between various geographical areas, (b) the simultaneous interaction of various individuals in an economic system, or (c) the various age groups in a growing population.

The ability to deal effectively with large numbers of interrelated variables is one of the most important characteristics of mathematical system analysis. It is necessary therefore to develop facility with techniques that help one clearly think about and systematically manipulate large numbers of simultaneous relations. For one's own thinking purposes, in order to understand the essential elements of the situation, one must learn, first, to view the whole set of relations as a unit, suppressing the details; and, second, to see the important detailed interrelations when required. For purposes of manipulation, with the primary objective of computation rather than furthering insight, one requires a systematic and efficient representation.

There are two main methods for representing sets of interrelations. The first is vector notation, which provides an efficient representation both for computation and for theoretical development. By its very nature, vector notation suppresses detail but allows for its retrieval when required. It is therefore a convenient, effective, and practical language. Moreover, once a situation is cast in this form, the entire array of theoretical results from linear algebra is available for application. Thus, this language is also well matched to mathematical theory.

The second technique for representing interrelations between variables is by use of diagrams. In this approach the various components of a system are represented by points or blocks, with connecting lines representing relations between the corresponding components. This representation is exceedingly helpful for visualization of essential structure in many complex situations; however, it lacks the full analytical power of the vector method. It is for this reason that, although both methods are developed in this book, primary emphasis is placed on the vector approach.

Most situations that we investigate are both dynamic and multivariable. They are, accordingly, characterized by several variables, each changing with time and each linked through time to other variables. Indeed, this combination of multivariable and time-evolutionary structure characterizes the setting of the modern theory of dynamic systems.

That most dynamic systems are both time-evolutionary and multivariable implies something about the nature of the mathematics that forms the basis for their analysis. The mathematical tools are essentially a combination of differential (or difference) equations and vector algebra. The differential (or difference) equations provide the element of dynamics, and the vector algebra provides the notation for multivariable representation. The combination and interplay between these two branches of mathematics provides the basic foundation for all analysis in this book. It is for this reason that this introductory chapter is followed first by a chapter on differential and difference equations and then by a chapter on matrix algebra.

1.3 A CATALOG OF EXAMPLES

As in all areas of problem formulation and analysis, the process of passing from a "real world" dynamic situation to a suitable abstraction in terms of a mathematical model requires an expertise that is refined only through experience. In any given application there is generally no single "correct" model; rather, the degree of detail, the emphasis, and the choice of model form are subject to the discretionary choice of the analyst. There are, however, a number of models that are considered "classic" in that they are well-known and generally accepted. These classic models serve an important role, not only as models of the situation that they were originally intended to represent, but also as examples of the degree of clarity and reality one should strive to achieve in new situations. A proficient analyst usually possesses a large mental catalog of these classic models that serve as valuable reference points—as well-founded points of departure.

The examples in this section are in this sense all classic, and as such can form the beginnings of a catalog for the reader. The catalog expands as one works his way through succeeding chapters, and this growth of well-founded examples with known properties should be one of the most important objectives of one's study. A diverse catalog enriches the process of model development.

The first four examples are formulated in discrete time and are, accordingly, defined by difference equations. The last two are defined in continuous time and thus result in differential equations. It will be apparent from a study of the examples that the choice to develop a continuous-time or a discrete-time model of a specific phenomenon is somewhat arbitrary. The choice is usually resolved on the basis of data availability, analytical tractability, established convention in the application area, or simply personal preference.

Example 1 (Geometric Growth). A simple growth law, useful in a wide assortment of situations (such as describing the increase in human or other

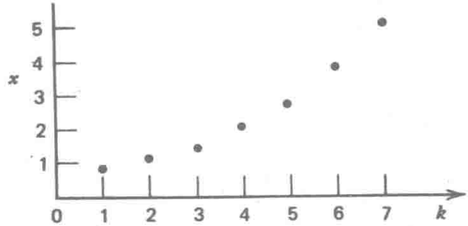


Figure 1.1. Geometric growth.

populations, the growth of vegetation, accumulated publications in a scientific field, consumption of raw materials, the accumulation of interest on a loan, etc.), is the linear law described by the difference equation

$$x(k+1)=ax(k)$$

The value $x(k)$ represents the magnitude of the variable (e.g., population) at time instant k. The parameter a is a constant that determines the rate of growth. For positive growth, the value of a must be greater than unity—then each successive magnitude is a fixed factor larger than its predecessor.

If an initial magnitude is given, say $x(0)=1$, the successive values can be found recursively. In particular, it is easy to see that $x(1)=a$, $x(2)=a^2$, and, in general, $x(k)=a^k$ for $k=0,1,2,\ldots$. A typical pattern of growth resulting from this model is shown in Fig. 1.1.

The growth pattern resulting from this simple linear model is referred to as *geometric growth* since the values grow as the terms of a geometric series. This form of growth pattern has been found to agree closely with empirical data in many situations, and there is often strong accompanying theoretical justification for the model, at least over a range of values.

Example 2 (Cohort Population Model). For many purposes (particularly in populations where the level of reproductive activity is nonuniform over a normal lifetime) the simple growth model given above is inadequate for comprehensive analysis of population change. More satisfactory models take account of the age distribution within the population. The classical model of this type is referred to as a *cohort population model.*

The population is divided into age groups (or cohorts) of equal age span, say five years. That is, the first group consists of all those members of the population between the ages of zero and five years, the second consists of those between five and ten years, and so forth. The cohort model itself is a discrete-time dynamic system with the duration of a single time period corresponding to the basic cohort span (five years in our example). By assuming that the male and female populations are identical in distribution, it is possible to simplify the model by considering only the female population. Let $x_i(k)$ be the (female) population of the ith age group at time period k. The groups are indexed sequentially from 0 through n, with 0 representing the lowest age group and n the largest. To describe system behavior, it is only necessary to describe how these numbers change during one time period.

First, aside from the possibility of death, which will be considered in a moment, it is clear that during one time period the cohorts in the ith age group simply move up to the $(i+1)$th age group. To account for the death rate of individuals within a given age group, this upward progression is attenuated by a survival factor. The net progression can be described by the simple equations

$$x_{i+1}(k+1) = \beta_i x_i(k), \qquad i = 0, 1, \ldots, n-1 \tag{1-1}$$

where β_i is the survival rate of the ith age group during one period. The factors β_i can be determined statistically from actuarial tables.

The only age group not determined by the equation above is $x_0(k+1)$, the group of individuals born during the last time period. They are offspring of the population that existed in the previous time period. The number in this group depends on the birth rate of each of the other cohort groups, and on how large each of these groups was during the previous period. Specifically,

$$x_0(k+1) = \alpha_0 x_0(k) + \alpha_1 x_1(k) + \alpha_2 x_2(k) + \cdots + \alpha_n x_n(k) \tag{1-2}$$

where α_i is the birth rate of the ith age group (expressed in number of female offspring per time period per member of age group i). The factor α_i also can be usually determined from statistical records.

Together Eqs. (1-1) and (1-2) define the system equations, determining how $x_i(k+1)$'s are found from $x_i(k)$'s. This is an excellent example of the combination of dynamics and multivariable system structure. The population system is most naturally visualized in terms of the variables representing the population levels of the various cohort groups, and thus it is a multivariable system. These variables are linked dynamically by simple difference equations, and thus the whole can be regarded as a composite of difference equations and multivariable structure.

Example 3 (National Economics). There are several simple models of national economic dynamics.* We present one formulated in discrete time, where the time between periods is usually taken as quarters of full years. At each time period there are four variables that define the model. They are

$$Y(k) = \text{National Income or National Product}$$
$$C(k) = \text{Consumption}$$
$$I(k) = \text{Investment}$$
$$G(k) = \text{Government Expenditure}$$

* See the notes and references for Sect. 4.8, at the end of Chapter 4.

The variable Y is defined to be the National Income: the total amount earned during a period by all individuals in the economy. Alternatively, but equivalently, Y can be defined as the National Product: the total value of goods and services produced in the economy during the period. Consumption C is the total amount spent by individuals for goods and services. It is the total of every individual's expenditure. The Investment I is the total amount invested in the period. Finally, G is the total amount spent by government during the period, which is equal to the government's current revenue. The basic national accounting equation is

$$Y(k) = C(k) + I(k) + G(k) \tag{1-3}$$

From an income viewpoint, the equation states that total individual income must be divided among consumption of goods and services, investment, or payments to the government. Alternatively, from a national product viewpoint, the total aggregate of goods and services produced must be divided among individual consumption, investment, or government consumption.

In addition to this basic definitional equation, two relationships are introduced that represent assumptions on the behavior of the economy. First, it is assumed that consumption is a fixed fraction of national income. Thus,

$$C(k) = mY(k) \tag{1-4}$$

for some m. The number m, which is restricted to the values $0 < m < 1$, is referred to as the *marginal propensity to consume.* This equation assumes that on the average individuals tend to consume a fixed portion of their income.

The second assumption concerning how the economy behaves relates to the influence of investment. The general effect of investment is to increase the productive capacity of the nation. Thus, present investment will increase national income (or national product) in future years. Specifically, it is assumed that the increase in national income is proportional to the level of investment. Or,

$$Y(k+1) - Y(k) = rI(k) \tag{1-5}$$

The constant r is the *growth factor*, and it is assumed that $r > 0$.

The set of equations (1-3), (1-4), and (1-5) defines the operation of the economy. Of the three equations, only the last is dynamic. The first two, (1-3) and (1-4), are *static*, expressing relationships among the variables that hold at every k. These two static equations can be used to eliminate two variables from the model. Starting with

$$Y(k) = C(k) + I(k) + G(k)$$

substitution of (1-4) produces

$$Y(k) = mY(k) + I(k) + G(k)$$