DYNAMIC PROGRAMMING AND MODERN CONTROL THEORY

Richard Bellman
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CALIFORNIA

Robert Kalaba
THE RAND CORPORATION
SANTA MONICA, CALIFORNIA

ACADEMIC PRESS New York and London
Preface

The scope of this book is a bit more ambitious than the title would indicate. Our aim has been to present an introduction to the mathematical theory of processes. The processes studied in physics, engineering, economics, biology, and operations research possess a bewildering array of special features: the state vector may be finite or infinite dimensional; the transition from one state to another may be deterministic or stochastic; it may be possible to influence the course of the process or not; competitive influences may be present or not; the controller may be able to learn about unknown aspects of the process or not. It is essential for the successful analyst to know how to incorporate any or all of these features, and others, into mathematical models of the situations which he wishes to study. Furthermore, he must have an awareness of the capabilities and limitations of modern computing machines and of the interfaces between theoretical formulation and numerical solution. In one sense, the building of realistic mathematical models is simple. Realistic models, however, have an unpleasant habit of quickly exceeding the memory and speed limitations of current computers. Progress can be made when adroit simplifications of the mathematical model permit a meaningful computational study or when analytical advances permit a class of processes, hitherto thought too complex, to be examined.

Chapter I is devoted to the general concept of multistage processes, both deterministic and stochastic, with time taken to be both continuous and discrete. This enables us to introduce the idea of recurrence relations and observe them in action in various settings. In Chapter II we discuss multistage decision processes, in which the objective is to make a sequence of decisions during the course of a process so as to maximize the gain from the process. The theoretical structure here is that of dynamic programming. There are short excursions into the calculus of variations, feedback control, and mathematical physics.

Chapter III examines some computational aspects of dy-
namic programming. Various theoretical questions concerning stability of numerical methods and the reduction of memory requirements are discussed and a FORTRAN program for an equipment replacement process is provided. In Chapter IV we turn to some communication and control processes for which we can obtain reasonably explicit analytical results. These are valuable as springboards from which to jump into the consideration of more realistic models using successive approximations.

The last chapter introduces the new and exciting topic of adaptive control. In such processes the controller simultaneously influences the course of the process and learns about unknown aspects of the process. This study is still in its infancy but holds much promise for the future.

Throughout we have provided many problems, potentially making the book useful for both classroom and individual study. Our hope is that this book will stimulate others to use the methods given here in still other fields and to devise new methods for gaining insight into complex processes via a combination of modern computational devices and modern mathematical theories.

November, 1965
Santa Monica, California

Richard Bellman
Robert Kalaba
Contents

Preface ............................................................................................................. v

Chapter I
Multistage Processes

1. Introduction ................................................................................................ 1
2. Systems in Flux .......................................................................................... 1
3. Multistage Process .................................................................................... 3
4. Reduction of Data ....................................................................................... 5
5. Independence of Past ................................................................................ 6
6. Recurrence Relations ................................................................................ 7
7. Infinite Processes ....................................................................................... 9
8. Processes with Stop Rules ......................................................................... 11
9. Time-Dependent or Nonstationary Processes ........................................... 12
10. Discussion .................................................................................................. 14
11. Continuous Multistage Processes ............................................................. 14
12. A Trajectory Process ............................................................................... 16
13. Inhomogeneous Atmosphere ................................................................... 19
14. Causality ................................................................................................... 20
15. Stochastic Processes ............................................................................... 21
16. Correlation ............................................................................................... 22
17. Noisy Observations ............................................................................... 23
18. Hidden Variables ..................................................................................... 24
19. Induced Processes .................................................................................... 24
20. Indirect and Induced Processes ................................................................. 26
21. More General Stochastic Processes ......................................................... 26
22. Conclusion ............................................................................................... 27
Bibliography and Comments ......................................................................... 28

Chapter II
Multistage Decision Processes

1. Introduction ................................................................................................ 31
2. Multistage Decision Process ..................................................................... 32
3. Policy ......................................................................................................... 33
4. Separable Criteria ..................................................................................... 34
5. Principle of Optimality ............................................................................. 35
6. Illustrations ............................................................................................... 35
7. Alternate Derivation of Results ................................................................. 37
Contents

8. What Constitutes a Solution? ........................................... 37
9. Continuous Multistage Decision Process ......................... 38
10. Calculus of Variations ............................................... 39
11. Geometric Aspects .................................................. 41
12. Stochastic Multistage Decision Processes ...................... 42
13. Feedback Control .................................................... 43
14. Analysis of the Equations .......................................... 44
15. Successive Approximations ......................................... 45
16. Approximation in Policy Space .................................... 46
17. Quasilinearization .................................................... 47
18. Analytic and Computational Aspects .............................. 48
19. Fermat's Principle and the Eikonal Equation .................... 48
20. Shortest Paths through Networks .................................. 50
    Bibliography and Comments ....................................... 55

Chapter III
Computational Aspects

1. Introduction .......................................................... 57
2. Digital Computers .................................................. 58
3. Computational Solution of Dynamic Programming
    Processes ............................................................. 58
4. Approximation in Function Space .................................. 60
5. Approximation in Policy Space .................................... 61
6. Computational Feasibility .......................................... 61
7. Polynomial Approximation ......................................... 62
8. Numerical Stability ................................................ 64
9. Discussion ........................................................... 65
10. Equipment Replacement Policy .................................... 66
    Bibliography and Comments ....................................... 71

Chapter IV
Analytic Results in Control and Communication Theory

1. Introduction .......................................................... 73
2. Feedback Control .................................................. 74
3. Scalar Case .......................................................... 74
4. Discussion ........................................................... 76
5. Stochastic Version .................................................. 77
6. Discussion ........................................................... 78
Contents

7. Multidimensional Case—Deterministic .................................. 78
8. The Continuous Case......................................................... 80
9. Prediction Theory.......................................................... 84
10. Communication of Information.......................................... 85
11. The Efficient Gambler...................................................... 86
12. Dynamic Programming Formulation.................................... 87
13. Discussion........................................................................ 89
    Bibliography and Comments............................................... 90

Chapter V

Adaptive Control Processes

1. Introduction....................................................................... 93
2. Linear Control Process of Adaptive Type............................. 94
3. Analytic Formulation......................................................... 95
4. Analytic Aspects.................................................................. 97
5. Discussion.......................................................................... 98
6. A Priori Distribution Function.......................................... 98
7. The Adaptive Gambler....................................................... 99
8. The Two-Armed Bandit Problem........................................ 100
9. A General Formulation..................................................... 102
10. The Estimation Problem................................................... 103
11. From Functionals to Functions.......................................... 103
12. More General Adaptive Processes..................................... 104
    Bibliography and Comments............................................... 104

Author Index........................................................................ 107
Subject Index........................................................................ 109
DYNAMIC PROGRAMMING
AND MODERN
CONTROL THEORY
Chapter I

Multistage Processes

1. Introduction

In this chapter, we wish to introduce a number of fundamental mathematical concepts which we shall use throughout the book in the analysis of various types of control and decision processes. The term “process” will be precisely defined and categorized in various ways: finite, infinite, unbounded; deterministic, stochastic; discrete, continuous; stationary and nonstationary. Using the basic notions of state variables and transformations, the power, versatility, and flexibility of recurrence relations and other types of functional equations will be illustrated.

2. Systems in Flux

We are primarily interested in the study of systems, a term as yet undefined, but with a meaningful intuitive flavor, and
how they behave in the course of time. This point of view permits us to introduce the idea of a "process" in a simple and natural way. For the moment, let us fix our attention on systems. Purely analytically, we conceive of a system as a state vector \( x(t) \), and a rule for determining its value at any time \( t \). For the sake of simplicity, we shall in this volume deal only with finite-dimensional vectors,

\[
x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ \vdots \\ x_N(t) \end{pmatrix}.
\] (2.1)

Each component of \( x(t) \) measures a different property of the system. The number \( N \) is called the dimension of the system.

Classically, the rules for determining the values of \( x(t) \) have been of the following type:

\[
\frac{dx}{dt} = g(x(t)) \quad \text{(differential equations)}
\]

\[
x(t) = g(x(t - 1)) \quad \text{(difference equations)} \tag{2.2}
\]

\[
\frac{dx}{dt} = h(x(t), x(t - 1)) \quad \text{(differential-difference equations)}
\]

\[
x(t) = u(t) + \int g(x(s), t) \, ds \quad \text{(integral equations)}
\]

together with various combinations of types.

Partial differential equations introduce infinite-dimensional vectors and thus determine more general systems. Although the methods we use can be applied to study these systems of more complex structure, in the text, we will be resolutely finite dimensional. Only by following this course can we preserve the introductory character of this monograph.

It is important at this point to emphasize that what we call the state of a system is not uniquely determined by the physical properties of the underlying real system being studied. In no sense is it to be considered an intrinsic characteristic. On the contrary, it hinges strongly on the particular, and often tortuous, road followed in the mathematical formulation.
3. Multistage Process

It depends, or should depend, upon what we wish to know about the physical process, what we can observe or measure, the accuracy of these observations, and, generally, upon the scientific and mathematical developments to date. It is impossible to overemphasize the significance of flexibility in the application of mathematical techniques to the scientific scene. There are, or should be, no chosen analytic methods, nor any chosen formulations.

In the past, most analytic work in applied mathematics has been in connection with functional equations of the type shown above. This work has produced results of lasting importance and there is little doubt that formulations in classical terms will continue to play a fundamental role in modern control theory and elsewhere. However, as new problems, and indeed new types of problems, have emerged from the scientific revolution we are experiencing, it has become quite clear that classical theories have severe limitations. New methods, together with extensions and generalizations of old methods, are required to answer the new and more exacting questions of modern science and technology and, in particular, to make fuller use of that scientific factotum, the Sorcerer's Apprentice, the digital computer. This remarkable device, even in its infancy, with scant understanding on our part of its use and potential, has already significantly and indeed irrevocably altered the ground rules of mathematics and science.

Let us then proceed to introduce some simple, yet powerful, mathematical methods which will permit us to formulate and analyze a number of problems which cannot be readily discussed along the lines of conventional theory. In this chapter, we preserve close contact with classical ideas; in the following chapter, we pioneer.

3. Multistage Process

Having introduced a purely mathematical definition of a "system," let us now make precise another intuitive concept—that of process. For the sake of generality, and in order not to be tied subliminally to older notions, we shall replace the
symbol \( x(t) \) by the symbol \( p \), and think of \( p \) as a point in a set, or space, \( R \). In all of the discussions of this volume, \( R \) will be a set of points in \( M \)-dimensional space, sometimes an \( M \)-dimensional cube, sometimes the lattice points of this cube. For example, \( R \) may consist of the positive integers, \( 1, 2, \ldots, M, \) and 0. It need not, however, be of any of the foregoing types in more general situations.

Consider next a function \( T(p) \), a transformation with the property that the transformed point \( p_1 = T(p) \) belongs to \( R \) for all \( p \) in \( R \). Intuitively, \( p \) represents the initial state of a system, \( p_1 = T(p) \) the state one time unit later, and, generally, the set of vectors

\[
[p, p_1, p_2, \ldots, p_n, \ldots],
\]

where \( p_0 = p, p_{n+1} = T(p_n), n = 0, 1, 2, \ldots \), represents the time history of a system observed at the discrete times \( n = 0, 1, 2, \ldots \), the successive states of the system. We could also write \( p_n = T^n(p) \), indicating the \( n \)-fold application of the transformation \( T \), and occasionally we will use this notation.

We call this infinite set of vectors appearing in (3.1) a multistage process, and, in turn, they define the term. A convenient notation is \([p, T(p)]\), or merely \([p, T]\). More precisely, it is a multistage process of discrete, deterministic type. Since, however, it is the only multistage process we presumably know about at the moment, we shall use the simpler term. In discussing other types of multistage processes, we shall add qualifying adjectives.

**EXERCISES**

1. Let \( p \) be a scalar, a one-dimensional vector, and let the transformation \( T \) have the form \( T(p) = rp \), where \( r \) is a real number. Depending upon the value of \( r \), determine the limiting behavior of \( p_n \) as \( n \to \infty \).

2. Let \( p \) be a two-dimensional vector, \( p = (x_1, x_2) \), and let the transformation \( T \) be a \( 2 \times 2 \) matrix of the form

\[
T = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.
\]
4. Reduction of Data

Show that geometrically $T$ represents a rotation through an angle $\theta$ in the $(x_1, x_2)$-plane.

3. What is the limiting behavior of $\Sigma_{n=1}^{N} T^n(p)/N$ as $N \to \infty$ as a function of $\theta$?

4. Let $u$ be a solution of the scalar differential equation $du/dt = au$, and let $T$ be the transformation induced and defined by expressing $u(1)$ in terms of $u(0)$. What is the analytic form of $T$? What is the limiting behavior of $T^n$ as $n \to \infty$, as a function of $a$?

5. Consider the corresponding problem for the case where $x$ is a vector satisfying the vector differential equation $dx/dt = Ax, x(0) = c$, where $A$ is a constant matrix. (See R. Bellman, *Introduction to Matrix Analysis*. New York: McGraw-Hill Book Company, Inc., 1960. Chapter 10.)

6. Consider the linear differential equation $u' + au = f(t), u(0) = c$, and write the solution in the form $u = L(t, c, f)$, indicating explicitly the dependence upon $t, c$, and the forcing function $f(t)$. Show that

   (a) $L(t, c, f) = L_0(t, c) + L_0(t, f)$, where
   (b) $L_0(t, c_1 + c_2) = L_0(t, c_1) + L_0(t, c_2),$
   (c) $L_0(t, f_1 + f_2) = L_0(t, f_1) + L_0(t, f_2).

7. A transformation from a function to a number is called a functional; a transformation from a function to a function is called an operation. The quantity $L_0(T, f)$ is a linear functional of $f$. Show that $u^2(T)$ is a quadratic functional having the form

$$u^2(T) = a(T)c^2 + 2c \int_0^T b(t, T)f(t) dt$$
$$+ \int_0^T \int_0^T k(t_1, t_2, T)f(t_1)f(t_2) dt_1 dt_2,$$

and determine the functions $a(T), b(t, T)$, and $k(t_1, t_2, T)$.

4. Reduction of Data

In most cases, a multistage process as personified by (3.1) tells far more than we want to know, and usually far more than we can absorb or profitably analyze. It is ironical in science that in order to understand, we must throw away information. We cannot, at least at this level of our intellectual development, grapple with a high order of complexity. Consequently, we must simplify.

To begin with, let us deal only with a section of the multistage process, an $N$-stage process, the set of vectors
I. Multistage Processes

\[ [p, p_1, p_2, \ldots, p_N], \quad (4.1) \]

where, as before, \( p_{k+1} = T(p_k), k = 0, 1, 2, \ldots, N - 1 \). We wish to examine the properties of various scalar functions of this process. If we allow these functions, \( \{g(p, p_1, \ldots, p_N)\} \), to have a completely general form, there is, as one might imagine, little of importance that can be said. In order to obtain interesting and useful results, we must introduce some structure into the function \( g \), or, equivalently, some structure into the multistage process. Observe, once again, that this is a superimposed structure, not necessarily intrinsic to the basic physical structure. For example, we may consider functions of the following forms:

\[
\sum_{i=0}^{N} h(p_i), \quad (4.2a)
\]

\[
\prod_{i=0}^{N} h(p_i), \quad (4.2b)
\]

\[
\max_{0 \leq i \leq N} h(p_i), \quad (4.2c)
\]

\[
\sum_{i=0}^{N-1} h(p_i, p_{i+1}). \quad (4.2d)
\]

In many cases, and indeed in most cases, the physical processes themselves will suggest useful functions to study. Each of the foregoing functions has many physical interpretations, and each arises from the consideration of interesting and significant physical processes.

5. Independence of Past

Before deriving any analytic consequences from the nature of a multistage process, let us discuss some conceptual aspects. We have assumed that at the initial time, the system is specified by the state vector \( p \), that at the subsequent observation, one time unit later, the state vector is \( p_1 = T(p) \), at the next observation, the state vector is \( p_2 = T(p_1) \), and so on. An immediate consequence of this assumption is that the states subsequent to the \( k \)th point in time depend only on \( p_k \), the state at time \( k \).
We do not require any knowledge of the past history of the system in order to determine the future. We have thus a clear-cut notion of the terms "past," "present," and "future," as far as multistage processes are concerned.

Another way of putting this is the statement that as far as $p_N$ is concerned, it is either the $N$th state in an $N$-stage process starting in initial state $p$ or the $(N - k)$th state in an $(N - k)$-stage process starting in state $p_k$ at time $k$. Symbolically, this may be expressed by the equation

$$T^N = T^{N-k}(T^k). \quad (5.1)$$

This relation may be considered an analytical statement of causality, or equivalently, of the fact that the future is uniquely determined by the present. In classical terms, this is a uniqueness theorem.

6. Recurrence Relations

Let us turn now to the analysis and an examination of the possibility of the computational determination of the associated functions of (4.2). Consider first the function

$$\sum_{i=0}^{N} h(p_i), \quad (6.1)$$

where $h(p)$ is a prescribed function, and observe that the function is completely determined by the initial state $p$ and the number of stages $N$, once we have decided upon the rule whereby $p_k$ is obtained from $p_{k-1}$, which is to say, once we have prescribed the transformation $T$. Since the analytic transliteration of the phrase "depends upon" is "is a function of," we introduce the sequence of functions $\{f_N(p)\}$ defined by the relation

$$f_N(p) = \sum_{i=0}^{N} h(p_i)$$

$$= h(p) + h(T(p)) + h(T^2(p)) + \cdots + h(T^N(p)), \quad (6.2)$$

where we allow $N$ to assume the values $N = 0, 1, 2, \ldots$, and let $p$ range over $R$. 
Consider for $N \geq 1$ the partial sum
\[ h(T(p)) + h(T^2(p)) + \cdots + h(T^N(p)). \] (6.3)
We see that this is equal to
\[ h(p_1) + h(T(p_1)) + \cdots + h(T^{N-1}(p_1)), \] (6.4)
and thus, $f_{N-1}(p_1)$ by definition of the sequence of functions \{f_N(p)\}.

It follows that $f_N(p)$ satisfies the functional equation (in this case, a recurrence relation)
\[ f_N(p) = h(p) + f_{N-1}(T(p)), \quad N \geq 1, \] (6.5)
\[ f_0(p) = h(p). \]

Similarly, if we turn to the function of (4.2b) and write
\[ f_N(p) = \prod_{i=0}^{N} h(p_i), \] (6.6)
we obtain the functional equation
\[ f_N(p) = h(p)f_{N-1}(T(p)). \] (6.7)

We leave to the reader the derivation of the relation
\[ f_N(p) = \max \{h(p), f_{N-1}(T(p))\}, \quad N \geq 1, \] (6.8)
for the function of $N$ and $p$ defined by
\[ f_N(p) = \max_{0 \leq i \leq N} h(p_i), \] (6.9)
arising from (4.2c); similarly, the relation
\[ f_N(p) = h(p, T(p)) + f_{N-1}(T(p)), \] (6.10)
may be established for the function of (4.2d).

**EXERCISES**

1. Consider the infinite geometric series $S = 1 + r + r^2 + \cdots$. Write $S = 1 + r(1 + r + r^2 + \cdots)$ and hence show that $S = 1/(1 - r)$ if $|r| < 1$.

2. Consider the infinite continued fraction
7. Infinite Processes

\[ S = 1 + \frac{1}{1 + \frac{1}{1 + \cdots}}. \]

Show formally that \( S = 1 + 1/S \) and hence that \( S = (1 + \sqrt{5})/2 \), provided that the continued fraction converges. Show that the expression above may be defined meaningfully as the limit of the finite continued fraction.

3. Consider the infinite continued fraction

\[ S(x) = 1 + \frac{x}{1 + \frac{x^2}{1 + \frac{x^4}{1 + \cdots}}}. \]

Show that \( S(x) = 1 + \frac{x}{1 + S(x^2)} \).

4. If

\[ S = \sqrt{1 + \sqrt{2 + \sqrt{4 + \cdots + \sqrt{2^{2n} + \cdots}}}}, \]

show that \( S = 2 \), assuming convergence. Once again, show that the formal infinite expression may be defined as the limit of the finite expression.

5. Consider the scalar relation \( u_{n+1} = au_n + b \), \( u_0 = c \). Write \( f_N(c) = \sum_{k=0}^{N} u_k^2 \). Show that \( f_N(c) \) is a quadratic polynomial in \( c \), \( f_N(c) = u_N c^2 + 2v_N c + w_N \). Using the recurrence relation \( f_N(c) = c^2 + f_{N-1}(ac + b) \), determine recurrence relations for the coefficients \( u_N \), \( v_N \), and \( w_N \), and examine the asymptotic behavior as \( N \to \infty \).

6. Generalize these results to the case where \( x_{n+1} = Ax_n + b \), \( x_0 = c \), with \( x_n \) a vector and \( A \) a matrix, and \( f_N(c) = \sum_{n=0}^{N} (x_n, x_n) \).

7. Consider the scalar relation \( u_{n+1} = (au_n + b)/(cu_n + d) \), \( u_0 = v \). Write \( u_N = (a_N v + b_N)/(c_N v + d_N) \), obtain recurrence relations for \( a_N \), \( b_N \), \( c_N \), and \( d_N \), and examine the asymptotic behavior as \( N \to \infty \).

8. Show that iteration of the transformation \( T(u) = (au + b)/(cu + d) \) is equivalent to raising the matrix \( A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) to successive powers.

7. Infinite Processes

The relations obtained in Section 6 simplify considerably if we consider infinite processes, which is to say processes in
I. Multistage Processes

which \( N = \infty \), the original multistage process. Proceeding formally and ignoring questions of convergence which are not relevant at the moment, write

\[
f(p) = \sum_{i=0}^{\infty} h(p_i).
\]  

(7.1)

Observe that the infinite number of stages removes any dependence upon \( N \), and thus removes one variable, a most important consideration. Then, arguing as before, we have

\[
f(p) = h(p) + \sum_{i=1}^{\infty} h(p_i),
\]  

(7.2)

\[
f(p) = h(p) + f(T(p)).
\]

Similarly, if

\[
f(p) = \max_{i} h(p_i),
\]  

(7.3)

we have, using the obvious relation \( \max [x_1, x_2, \ldots] = \max [x_1, \max [x_2, \ldots]] \),

\[
f(p) = \max [h(p), f(T(p))].
\]  

(7.4)

At this point, we can agree that infinite processes have no physical reality. Nonetheless, they represent very convenient approximations to the more complex finite-stage processes of long duration. We can equally take the converse of the argument and assert that all processes are unbounded, with finite processes representing a convenient approach to the more difficult infinite processes. The essential point is to recognize that we have at the time of formulation a choice of both conceptual and analytic format, and then to exploit this freedom. There are advantages and disadvantages to each.

In many cases, the function \( f_N(p) \) does not approach a limit as \( N \to \infty \). Rather, we have an "average" behavior, or steady-state behavior, expressed by an asymptotic relation of the form

\[
\frac{f_N(p)}{N} \sim g_1(p)
\]  

(7.5)

as \( N \to \infty \), or, perhaps, a relation of the form
8. Processes with Stop Rules

\[ f_N(p) \sim g_3(p)^N h(p). \] (7.6)

Generally, the physical background tells us the type of limiting behavior to expect, and the functions \( g_1 \) and \( g_2 \) have physical, engineering, and economic significance. We shall, nonetheless, avoid any discussion of these very interesting, important, and difficult matters which require, in general, detailed and sophisticated analysis.

8. Processes with Stop Rules

In a number of important situations, we are confronted by finite processes in which the number of stages is not fixed in advance, but itself depends upon the initial state; i.e., \( N = N(p) \). We then have associated functions of the form

\[ f(p) = \sum_{i=0}^{N(p)} h(p_i). \] (8.1)

A first simple example of this might arise from a process in which we are using resources at each stage with the agreement that the process stops as soon as our resources are exhausted. Analytically, this condition might express itself by means of a constraint of the form

\[ \sum_{i}^{N} k(p_i) \leq a. \] (8.2)

Another example of some significance arises from the study of trajectory processes. An object continues in motion until it comes within an agreed-upon distance of some other object. Analytically, if we introduce a metric into \( R \), a distance \( d(p, q) \) between any two elements \( p \) and \( q \) in \( R \), we stipulate that the process continued until \( d(T^N p, q) \leq \epsilon \), where \( q \) and \( \epsilon \) are given in advance.

Processes of this general structure play an important role in modern engineering and economics.

Let us consider the trajectory process where the process ends as soon as \( d(p_N, q) \leq \epsilon \). If we introduce the function

\[ f(p) = \sum_{i=0}^{N} h(p_i), \] (8.3)
where \( N = N(p) \) is determined by the foregoing "stop rule," we have

\[
\begin{align*}
f(p) = h(p) & \quad \text{if} \quad d(p, q) \leq \epsilon, \\
f(p) = h(p) + f(T(p)) & \quad \text{if} \quad d(p, q) > \epsilon. 
\end{align*}
\] (8.4)

In the consideration of processes of this nature, the recurrence relation (functional equation) technique is often of great utility. These processes are examples of unbounded processes, which are not necessarily infinite processes.

9. Time-Dependent or Nonstationary Processes

So far we have assumed that the form of \( T(p) \) was independent of the stage. In other words, the transformation \( T \) was kept unchanged over time. Processes of this type are called stationary. Suppose that we now consider the more general situation where

\[
p_{n+1} = T_n(p_n),
\] (9.1)

which is to say, where the transformation is time-dependent. In this case, the relation of (5.1) is not valid. In order to obtain an analogous result, we must take account of the time at which the process starts. A general nonstationary process has the form

\[
[p_m, p_{m+1}, \ldots, p_n, \ldots],
\] (9.2)

where

\[
p_{m+1} = T_m(p_m), \quad p_{m+2} = T_{m+1}(p_{m+1}), \ldots
\] (9.3)

We may then just as well suppose that the associated functions are also dependent on time and consider, therefore, associated functions of the form

\[
\sum_{k=m}^{N} h_k(p_k), \quad \text{(9.4a)}
\]

\[
\prod_{k=m}^{N} h_k(p_k), \quad \text{(9.4b)}
\]

\[
\max_{m \leq k \leq N} h_k(p_k). \quad \text{(9.4c)}
\]
9. Time-Dependent or Nonstationary Processes

To apply the functional equation approach of the preceding sections, we observe that the state of the system may be taken to be the vector $p$ plus the time at which the process begins. Hence, to treat the function

$$\sum_{k=m}^{N} h_k(p_k)$$  \hspace{1cm} (9.5)

by means of the functional technique used previously, we write $p_m = p$ and indicate the dependence on both $p$ and $m$ by introducing the function

$$f_{N,m}(p) = \sum_{k=m}^{N} h_k(p_k),$$  \hspace{1cm} (9.6)

satisfying the relation

$$f_{N,m}(p) = h_m(p) + f_{N,m+1}(T_m(p)).$$

If $N$ is fixed throughout our discussion, we can simplify the notation by introducing the new functions

$$\phi_m(p) = f_{N,m}(p),$$  \hspace{1cm} (9.7)

and write

$$\phi_m(p) = h_m(p) + \phi_{m+1}(T_m(p)), \quad 0 \leq m < N,$$

$$\phi_N(p) = h_N(p).$$  \hspace{1cm} (9.8)

EXERCISES

1. Consider a one-dimensional system $S$ specified by the state variable $x_n$, $n = 0, 1, 2, \ldots$, and the relation $x_{n+1} = ax_n$, $x_0 = c$, $a \geq 1$, $c > 0$. Consider a new system, a controlled system, specified by the equation $x_{n+1} = ax_n + y(x_n)$, $x_0 = c$, where $y(x) = 0$, $0 \leq x \leq 1$, $y(x) = -1$, $x > 1$. Determine the behavior of $x_n$ as $n \to \infty$ both graphically and analytically. For what values of $a$ does $x_n \to \infty$ as $n \to \infty$?

2. Let $f_n(c)$ denote the value of $x_n$ at time $n$. Show that

$$f_{n+1}(c) = f_n(ac + y(c)), \quad n \geq 0,$$

$$f_0(c) = c.$$
10. Discussion

Up to this point, we have been following a classical route, laid out by Poincaré, Birkhoff, Hadamard, and others, connecting transformations at discrete intervals with the theory of iteration. Our subsequent results will proceed in several different directions, continuous processes, decision processes of deterministic, stochastic, and adaptive type, and processes of convenient analytic aspect.

11. Continuous Multistage Processes

So far we have been supposing, albeit tacitly, that the system is observable only at a discrete set of times, and, therefore, as far as we are concerned, transforms only at these times. Let us now consider the case where the state of the system changes continually over time. To do this, we use the formalism set up for a discrete process and then let the time intervals between observations go to zero.

Take the times of observation to be \( t = 0, \Delta, 2\Delta, \ldots \), and let the transformation \( T(p) \) have the form

\[
T(p) = p + S(p)\Delta + o(\Delta),
\]

(11.1)

where \( \Delta \) is a small quantity which we will ultimately let approach zero. Furthermore, let the associated function have the form \( g(p)\Delta + g(T(p))\Delta + \cdots + g(T^n(p))\Delta \). In other words, we suppose that only small changes in state and associated function are allowed over small time intervals. Then, if we let \( t = n\Delta \), we have the recurrence relation

\[
f_{t+\Delta}(p) = g(p)\Delta + f_t(p + S(p)\Delta + o(\Delta)).
\]

(11.2)

Consider the \( M \)-dimensional case where

\[
p = \begin{pmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_M
\end{pmatrix}, \quad S(p) = \begin{pmatrix}
  S_1(x_1, x_2, \ldots, x_M) \\
  \vdots \\
  S_M(x_1, x_2, \ldots, x_M)
\end{pmatrix}.
\]

(11.3)

Then, using the familiar notations of vector analysis,
11. Continuous Multistage Processes

\[
\text{grad } f = \begin{pmatrix}
\frac{\partial f}{\partial x_1} \\
\vdots \\
\frac{\partial f}{\partial x_M}
\end{pmatrix}, \\
(f, h) = \sum_{i=1}^{M} f_i h_i,
\]

we have from (11.2),

\[
f_i(p) + \frac{\partial f}{\partial t} \Delta + o(\Delta) = g(p)\Delta + f_i(p)
\]

\[
+ (\text{grad } f, S(p))\Delta + o(\Delta),
\]

assuming that the function \( f \) possesses the necessary partial derivatives. Passing to the limit as \( \Delta \to 0 \), we obtain the partial differential equation

\[
\frac{\partial f}{\partial t} = g(p) + (\text{grad } f, S(p)).
\]

So far we have proceeded formally. If we now wish to define a continuous process starting from first principles, we can proceed in various ways. We can first of all make the foregoing analysis rigorous by showing that the discrete sequences \( \{f_n(\Delta)\}, \{p_n\} \), converge in various fashions as \( \Delta \to 0 \). Secondly, we can define a continuous process by means of a function \( p(t) \), the state variable, satisfying the differential equation

\[
p'(t) = S(p), \quad p(0) = p_0,
\]

and the associated function by means of the partial differential equation

\[
\frac{\partial f}{\partial t} = g(p) + (\text{grad } f, S(p)), \quad f(0) = f_0.
\]

There is now the task of showing that these are meaningful definitions in the sense of existence and uniqueness of solutions of these equations, with various desired properties.

These are nontrivial and often quite difficult matters. It is for this reason that we prefer in what follows to concentrate upon discrete processes. Multistage decision processes of discrete type involve us in a sufficient number of conceptual, analytic, and computational problems without the burden of
answering often extraneous delicate existence and uniqueness questions connected with continuous processes.

EXERCISES

1. Consider the following process. A particle moves in the interval \([0, 1]\) until it hits a boundary point, at which time it reverses direction and moves in the other direction until it hits the other boundary point, and so on. The particle starts at a point \(c\) moving to the right according to the law \(du/dt = a > 0\). When it moves to the left, it moves according to the law \(du/dt = -b < 0\). Find an expression for the position of the particle at time \(t\).

2. Suppose that the equations of motion are \(du/dt = au\), \(du/dt = -bu\), respectively, and solve the same problem.

3. For each set of equations of motion, consider the expression \(v(t) = u(t)/t\). Does \(\lim_{t \to \infty} v(t)\) exist as \(t \to \infty\)? If it exists, as a function of \(c\), determine its properties and, in particular, obtain a functional equation.

12. A Trajectory Process

In order to illustrate some of the preceding ideas, let us consider a simple trajectory process. Suppose that we throw a stone straight up with a velocity \(v\) and assume that it then moves solely under the influence of gravity. The conventional treatment of this process is straightforward enough.

\[
\begin{align*}
\text{FIG. 1-1} \\
\end{align*}
\]

Let \(x(t)\) denote the distance above ground at time \(t\). Neglecting such realistic effects as air resistance, etc., the equation of motion of the point particle is
12. A Trajectory Process

\[ \frac{d^2x}{dt^2} = -g \]  

(acceleration is due to gravity and directed downward), with the initial conditions

\[ x(0) = 0, \quad x'(0) = v, \]  

where \( v \geq 0 \).

The explicit solution is

\[ x = vt - \frac{gt^2}{2}. \]  

From this, we see that the maximum altitude is attained at \( t = v/g \) and that the maximum altitude is given by

\[ x_{\text{max}} = \frac{v^2}{2g}. \]  

The particle concludes its flight at the time \( t = 2v/g \), when it hits the ground.

Simple enough, but there are some loose ends. Suppose that we had merely asked for the maximum altitude. Could we determine it without the superfluous information concerning the rest of the flight? In this case, there seems little point, but in more complex processes this economy of effort can be significant, particularly, if the only point of the original investigation is the determination of the dependence of maximum altitude upon initial velocity.

Let us employ the preceding ideas to answer the foregoing question. We begin with the observation that the maximum altitude depends upon the initial velocity \( v \). Let us then introduce the function \( f(v) \), the maximum altitude attained starting with initial velocity \( v \). We then consider the moving particle as generating a continuous multistage process in which the state of the system is \( v \). Over an infinitesimal time interval \( \Delta \), the transformation affecting the state is

\[ T(v) = v - g\Delta. \]  

Hence, we have the relation

\[ f(v) = v\Delta + f(v - g\Delta) + o(\Delta). \]
I. Multistage Processes

\[
\begin{align*}
\{ f(v) \} & \quad \rightarrow \quad \{ f(v-g\Delta) \} \\
0 & \quad \rightarrow \quad v\Delta
\end{align*}
\]

FIG. 1-2. Additive Nature of Maximum Altitude.

Expanding the right-hand and letting \( \Delta \to 0 \), we obtain the equation

\[ 0 = v - gf''(v). \quad (12.7) \]

This is a differential equation for \( f(v) \) with the initial condition \( f(0) = 0 \). Integrating, we obtain the relation of (12.4).

EXERCISES

1. A dog is pursuing a rabbit, as indicated in the figure below.

Let the speed of the dog be \( v_d \) and that of the rabbit be \( v_r \) and assume that \( v_d > v_r \). Suppose that at each instant of time the rabbit points along the \( x \)-axis and the dog points towards the rabbit.

Obtain the differential equation for the path of dog and compare the "solution" in classical terms, i.e., \( x(t) \) as a function of \( t \), with the "solution" in terms of the original line-of-sight pursuit process as far as a computational solution is concerned.

2. Let \( f(r, d) \) denote the time at which the dog catches the rabbit. Show that

\[ f(r, d) = \Delta + f\left( r + \Delta, d - \frac{\Delta d^2}{\sqrt{r^2 + d^2}} \right) + o(\Delta), \]
whence, letting $\Delta \to 0$,

$$0 = 1 + \frac{\partial f}{\partial r} - \frac{d^2}{\sqrt{r^2 + d^2}} \frac{\partial f}{\partial d}$$

with $f(r, 0) = r/(v_d - v_r)$. Similarly, derive a partial differential equation for the quantity $g(r, d)$, the position along the x-axis where the dog overtakes the rabbit.

3. Use the technique of characteristics to find $f(r, d)$ and $g(r, d)$ in explicit analytic terms.

13. Inhomogeneous Atmosphere

Let us now assume that in place of the equation of (12.1), we have an equation of the form

$$\frac{d^2 x}{dt^2} = h(x, x'), \quad x(0) = 0, \quad x'(0) = v, \quad (13.1)$$

an equation derived in accordance with the supposition that we take account of an air resistance dependent upon position and velocity. As a consequence of this inhomogeneity of the atmosphere, we must introduce the additional state variable $c$, equal to the altitude at which the process starts.

We consider then the two state variables $c$ and $v$ with the transformation law

$$c_1 = T_1(c, v) = c + v\Delta,$$

$$v_1 = T_2(c, v) = v + h(c, v)\Delta, \quad (13.2)$$

to terms in $o(\Delta)$. If we introduce the function

$$f(c, v) = \text{the maximum altitude attained starting at altitude } c \text{ with velocity } v,$$

then, as before,

$$f(c, v) = v\Delta + f(c + v\Delta, v + h(c, v)\Delta) + o(\Delta). \quad (13.3)$$

In the limit as $\Delta \to 0$, we are led to the partial differential equation

$$0 = v + v \frac{\partial f}{\partial c} + h(c, v) \frac{\partial f}{\partial v}. \quad (13.5)$$

In general, this cannot be solved analytically. A numerical solution may be obtained in a number of ways, using the
initial condition \( f(c, 0) = 0 \), but it should be pointed out there are a number of points that arise that require experience and skill. See the discussion of computational aspects in Chapter 3. Alternatively, we can employ the theory of characteristics to obtain a numerical solution.

14. Causality

The equations obtained above are all particular examples of relations obtained from the principle of causality, or, equivalently, of determinism. Let the state of a system at time \( t \) be represented by \( f(c, t) \), where \( c \) is the state at time \( t = 0 \).

We can think of the system as starting in state \( c \) at \( t = 0 \) and evolving for a time \( s + t \), in which case its terminal state will be \( f(c, s + t) \), or we can think of it as starting in state \( c \) at time \( t = 0 \), evolving for a time \( t \), in which case its new state will be \( f(c, t) \), and then continuing for an additional time \( s \), in which case its terminal state will be \( f(f(c, t), s) \). We have already presented the discrete version of this argument in Section 5.

\[ f(c, s + t) = f(f(c, s), t). \quad (14.1) \]

From this relation, many further results can be obtained, as, for example, those of Section 11.

EXERCISES

1. Use the uniqueness of the solution of \( u' = au, u(0) = c \), to obtain the functional equation for the exponential function \( e^{a(s+t)} = e^{as} e^{at} \).
2. Use the corresponding result for the vector-matrix equation $x' = Ax$, $x(0) = c$, to obtain the matrix analog, $e^{A(s+t)} = e^{As}e^{At}$.

3. Similarly, starting with the fact that $\sin t$ and $\cos t$ are solutions of $u'' + u = 0$, establish the addition formulas $\sin(s + t) = \sin s \cos t + \cos s \sin t$, $\cos(s + t) = \cos s \cos t - \sin s \sin t$.

15. Stochastic Processes

So far we have assumed that the result of the transformation $T$ is to take the state vector $p$ into the state vector $p_1$, where $p_1$ is uniquely determined by $p$. In many cases, we must face the fact that $T$ is not completely known. This means that the conceptually simple deterministic models we have been using must be replaced by more sophisticated ones. To circumvent partially the roadblock of ignorance, we use the classical theory of probability.

In place of the frank statement that $p_1$ is unknown, we suppose that $T$ is a stochastic transformation which produces a random vector $p_1$ whose probability distribution is determined by $p$. The set of vectors $[p, p_1, p_2, \ldots]$ now define a discrete multistage process of stochastic type.

It follows that the associated functions will also be random variables. To meet the problem of introducing numerical values, we make the fundamental assumption that we shall be content with expected values in dealing with functions of these random variables. This is the standard procedure in the study of stochastic processes.

To illustrate these ideas, let us begin with perhaps the simplest case where $p_k$ is determined by the relation

$$p_k = T(p_{k-1}, r_k), \quad k = 1, 2, \ldots, \quad (15.1)$$

where $p_0 = p$ and the $r_k$ are independent random variables with the identical probability distribution $dG(r)$. Suppose that we wish to determine the expected value of

$$g(p) + g(p_1) + \cdots + g(p_N). \quad (15.2)$$

As before, we start with the observation that this expected value depends upon the initial state $p$ and the number of stages $N$. Let us then write for $N \geq 1$, 

I. Multistage Processes

\[ f_N(p) = \exp \left[ g(p) + g(p_1) + \cdots + g(p_N) \right]. \quad (15.3) \]

Here, the notation \( \exp_r \) indicates that the expected value is to be taken with respect to the random variables \( r_1, r_2, \ldots, r_N \). Arguing as before, the assumed independence of the \( r_k \) permits us to conclude that

\[ f_N(p) = \exp_{r_1} \left[ g(p) + f_{N-1}(T(p, r_1)) \right] \]
\[ = g(p) + \int f_{N-1}(T(p, r_1)) \, dG(r_1), \quad (15.4) \]

with \( f_0(p) = g(p) \).

EXERCISES

1. Let \( u_{n+2} = au_n + r_n \) where the \( r_n \) are random independent variables with the common distribution function \( dG(r) \). Write \( f_N(c) = \exp \sum_{n=0}^{N} u_n^2 \).
   Show that \( f_N(c) \) is a quadratic polynomial in \( c \), and use the foregoing equation to obtain recurrence relations for its coefficients. Show that \( f_N(c) \) depends only on the first and second moments of \( dG(r) \), the quantities \( \int r \, dG(r) \), \( \int r^2 \, dG(r) \), and study its asymptotic behavior as \( N \to \infty \).

2. Generalize to the vector-matrix case where \( x_{n+1} = Ax_n + r_n \), \( x_0 = c \).

16. Correlation

In many situations, the assumption of independence of the \( r_k \) is unwarranted. A first step in the direction of taking account of dependence is to suppose that the distribution function for \( r_k \) depends only upon the value of \( r_{k-1} \), i.e., \( dG(r_k) = dG(r_k, r_{k-1}) \).

In this case we can choose as state variables \( p \), the present state of the system, and \( r_0 \), the previous value of the random variable \( r \). The analog of (15.4) is

\[ f_N(p, r_0) = g(p) + \int f_{N-1}(T(p, r_1)) \, dG(r_1, r_0). \quad (16.1) \]

There are many other ways of introducing interdependence of the \( r_k \). The assumptions made above are most convenient for the functional equation approach we are employing.
EXERCISE

Consider the equation $u_{n+1} = au_n + r_n$, where the distribution of $r_n$ is $dG(r_n, r_{n-1})$. Let $u_0 = c$, and $r_{-1} = b$. Set $f_N(c, b) = \exp\left(\sum_{n=0}^{N} u_n^2\right)$. Show that $f_N(c, b) = u_N(b)c^2 + 2v_N(b)c + w_N(b)$, and obtain recurrence relations, for $u_N$, $v_N$, $w_N$ from a recurrence relation for $f_N(c, b)$. Study the asymptotic behavior as $N \to \infty$.

17. Noisy Observations

As soon as we admit uncertainty anywhere into our philosophy, we must allow for it everywhere. Suppose that we take account of the fact that there may be significant uncertainties in the observation of the state of the system. A simple way of doing this is to assume that when we observe $p$, we know that the state of the system is actually $p + r$ where $r$ is a random variable with a known distribution.

We can then conceive of two processes, one $[p, p_1, p_2, \ldots]$, the actual state of the system, and the second $[q, q_1, q_2, \ldots]$, the observations made on the system, which is to say the information we actually possess. The following relations hold:

$$p = q,$$
$$p_1 = T(p), \quad q_1 = p_1 + r_1,$$
$$p_2 = T(p_1), \quad q_2 = p_2 + r_2,$$
$$\vdots$$
$$p_n = T(p_{n-1}), \quad q_n = p_n + r_n.$$  \hspace{1cm} (17.1)

If we use the observations as the states of the system, we see from the foregoing that we are considering a stochastic multi-stage process $[q, q_1, q_2, \ldots]$ with

$$q_1 = T(q) - r_1,$$
$$q_2 = T(q_1 - r_1) - r_2,$$
$$\vdots$$
$$q_n = T(q_{n-1} - r_{n-1}) - r_n.$$  \hspace{1cm} (17.2)
a process with interdependence among the \( r_n \) of the type described above. It is thus easy to derive functional equations for the associated functions.

18. Hidden Variables

If \( p_1 = T(p, q) \), where \( q \) is a vector whose very existence is unknown to us, we can be led to think of \( T \) as a random transformation in the following way. Fixing \( p \) and repeating the transformation process—i.e., observing a set of similar systems—we find a set of values

\[
p_{11} = T(p, q_1), \quad p_{12} = T(p, q_2), \ldots \quad (18.1)
\]

in general, different values of \( p_1 \) corresponding to the same value of \( p \), but to different values of \( q, q_1 \) for the first system, \( q_2 \) for the second system, and so on. The possible lack of reproducibility of \( p_1 \) for the same value of \( p \) may lead us to believe that determinism does not hold. Whether or not one wishes to believe in deterministic processes as fundamental with stochastic processes solely as a mathematical device introduced to handle unknown or "hidden" variables, or to believe that nature is basically stochastic with determinism the result of averaging, again a mathematical device, is a matter of personal philosophy which fortunately has little effect upon the analytic models that arise.

It is important, however, to emphasize once again that one has a great deal of leeway in constructing mathematical models of physical phenomena, and it would be unfortunate to let scientific dogmatism prevent the formulation and consideration of convenient mathematical models.

19. Induced Processes

Starting with a process of any of the foregoing types, we can derive other processes in a variety of ways. To begin with, we can take advantage of the structure of the process to simplify the analytic form of the associated function or to reduce the dimension of the process. Let us illustrate the first of these
19. Induced Processes

points with a process already discussed in the exercises. Suppose that

\[ u_{n+1} = au_n + b, \quad u_0 = c, \]  
\[ f_N = \sum_{n=0}^{N} u_n^2. \]  

(19.1)

It is easy to see that \( f_N = p_N + 2q_N c + r_N c^2 \) where \( p_N, q_N, r_N \) are independent of \( c \). Furthermore, the transformation of the original state variable induces a transformation on the variables \( (p_N, q_N, r_N) \) by means of the recurrence relation

\[ f_{N+1}(c) = c^2 + f_N(ac + b), \]  

(19.2)

namely

\[ p_{N+1} = p_N + 2bq_N + b^2 r_N, \]
\[ q_{N+1} = aq_N + abr_N, \]
\[ r_{N+1} = a^2 r_N + 1. \]  

(19.3)

In place of (19.2), we can use these transformations to study the original process. This is a significant decrease in dimensionality and increase in analytic flexibility, since we replace a function \( f_N(c) \), defined for \(-\infty < c < \infty\), by a triple \([p_N, q_N, r_N]\). We shall return to this point in Chapters 3 and 4.

In a similar vein, suppose we begin with the process \( x(t) \) determined by the vector equation

\[ \frac{dx}{dt} = Ax(t), \quad x(0) = c, \]  

(19.4)

where \( x \) is an \( N \)-dimensional vector, and take as the associated function \( g(x_1(T), \ldots, x_k(T)) \), a function of the first \( k \) components of \( x \). Since the solution of (19.4) is

\[ x(t) = e^{At}c, \]  

(19.5)

we see that, as far as this associated function is concerned, new state vectors are the first \( k \) components of \( e^{AT}c \), a reduction in dimensionality from \( N \) to \( k \). In other words, we can condense the desired information concerning the system into a set of \( k \) numbers. This can be quite important in applications, from both the analytic and computational point of view.
20. Indirect and Induced Processes

We have assumed, in what has gone before, that \( p_1 \), the new state, is obtained by means of the transformation \( T \) applied to the former state \( p \). In a number of processes, the situation is not as direct.

Suppose, for example, we have a process \( \{x_n\} \) with the transformation rule

\[
x_{n+1} = Ax_n, \quad x_0 = c.
\]  

(20.1)

Here \( x_n \) is an \( N \)-dimensional vector with complex components. Let \( x_{ni}, i = 1, 2, \ldots, N \), denote the \( N \) components of \( x_n \) and form the new set of \( N \) quantities \(|x_{ni}|^2, i = 1, 2, \ldots, N\). If \( A \) is a unitary matrix, so that

\[
\sum_{i=1}^{N} |x_{ni}|^2 = \sum_{i=1}^{N} |c_i|^2,
\]

we can normalize \( c \) so that

\[
\sum_{i=1}^{N} |c_i|^2 = 1,
\]

and thus suppose that the quantities \(|x_{ni}|^2\) are probabilities.

The set \([|x_{n1}|^2, \ldots, |x_{nN}|^2]\) constitutes a new process, with a transformation law outside the class we have so far been considering. Transformations of this type, strange as they may seem, are the primary transformations of quantum mechanics. We shall, however, not deal with them here, nor with the very intriguing new types of control processes they create.

21. More General Stochastic Processes

As soon as we allow uncertainty to raise its inquiring head, we introduce many extremely interesting new types of processes, and, correspondingly, new types of functional equations. Let us briefly indicate various ways in which we can modify deterministic processes in order to obtain significant stochastic processes.

In the first place, instead of allowing the transformations to act upon the system at prescribed times, we may suppose that
the time between transformations is itself a random variable. This brings us into the domain of branching processes which includes, as special cases, renewal processes. Processes of this type are of importance in nuclear physics, in modern inventory theory, in control theory, and in the study of learning processes.

Secondly, we can consider a stochastic process, \([p, T(p, r)]\), in which there is a probability at each stage that the actual state of the system cannot be observed. In practice, this can occur as a consequence of an interruption of communications or a failure of a sensing device.

Processes of this more general nature force us, frequently, to replace a finite-dimensional state variable by an infinite-dimensional one, based upon a probability distribution. Various techniques are then used to replace the infinite-dimensional vector by a finite-dimensional one. Thus, for example, we may consider only the first and second moments, or, equivalently, the means and variances, as the new state variables.

As mentioned before, we have avoided any systematic discussion of processes involving infinite-dimensional state vectors.

22. Conclusion

The purpose of this chapter has been to introduce the reader to the fundamental ideas of the mathematical theory of processes and systems and to have him gain some feeling of familiarity with functional equations and recurrence relations. Our objective in doing so was to ease the "culture shock" of the introduction of the concepts and methods of dynamic programming. Since we had only this limited objective in mind, we did not allow ourselves to pursue any of the many interesting questions concerned with the classification of processes, their analysis, and/or their computational solution. In the study of stochastic processes, these inquiries lead to the concepts of sufficient statistics, and asymptotic sufficient statistics. Oddly enough, the corresponding concepts, despite their obvious importance, have never been formally studied for deterministic processes.
I. Multistage Processes

Much remains to be done in these areas and they afford fertile fields for research for the young scientist. A number of significant papers and books for further reading are given in the references.

BIBLIOGRAPHY AND COMMENTS

§1. For a precise and detailed discussion of the concept of "system," with many illustrative and illuminating examples, see

§2. There are many books currently available which discuss differential, difference and integral equations. For the theory of differential-difference equations, see

§3. For a detailed account of the abstract theory of classical mechanics, see
References to the work of Poincaré, Hadamard, Levi-Civita, and others will be found there.

§5. See the comments below pertaining to Section 17.

§7. A systematic investigation of when limiting behavior exists would lead into a number of the classical fields of analysis, stability theory, ergodic behavior, asymptotic theory, summability, and so on. Let us merely mention

§10. The reader interested in pursuing the theme of iteration may refer to

§11. Much of contemporary effort in partial differential equations is

A more sophisticated introduction to the study of continuous multistage processes would be by way of the theory of continuous groups of transformations, the theory of Sophus Lie. For an introductory account, see E. L. Ince, *Ordinary Differential Equations*, Dover Publications, New York, 1944.


§18. The theme of “hidden variables” is one that has excited much controversy since the development of quantum mechanics. See
I. Multistage Processes


where many additional references may be found.

§19. See

§20. See the article

for a detailed discussion.

§21. See the book by Harris referred to above, and

For "interrupted" control processes, see

§22. The reader interested in these matters may wish to consult


Chapter II

Multistage Decision Processes

1. Introduction

We are now ready to begin the study of multistage decision processes, which is to say, the theory of dynamic programming. These processes are mathematically important in their own right as natural and far reaching generalizations of the transformations treated in classical analysis. In addition, as we shall see, they occupy a fundamental position in modern control theory.

To introduce the theory of dynamic programming, we add the concept of a "decision" to the concept of a multistage process. As before, the text will in the main be restricted to a presentation of the abstract ideas, while the exercises will contain a number of examples illustrating the methods, together with references for further reading.
An important new idea is that of "policy." This, in analytic terms, leads to a type of successive approximations not found in classical analysis, namely approximation in policy space. This, in turn, as we show, leads to the technique of quasilinearization, a very powerful and versatile approach to many types of nonlinear functional equations.

In Chapter 3 we shall consider the use of dynamic programming as a computational algorithm capable of yielding numerical answers to numerical questions. In Chapter 4 we discuss a particular class of multistage decision processes which can be treated in some analytic detail.

2. Multistage Decision Process

To define a multistage decision process of the simplest type, discrete and deterministic, we start with the previously established notion of a multistage process \([p, T(p)]\). To recall, this is the set of vectors \([p, p_1, p_2, \ldots, p_n, \ldots]\) with \(p_{n+1} = T(p_n), p_0 = p\). We now enlarge this concept by taking the transformation \(T\) to depend on another vector as well, \(T = T(p, q)\).

We suppose that we have sufficient influence over the process so that at each stage we can choose the value of \(q\) from a set of allowable vectors \(S(q)\). Let \(q_i\) be the choice at the \(i\)th stage so that

\[
\begin{align*}
p_1 &= T(p, q_0), \\
p_2 &= T(p_1, q_1), \\
&\quad \quad \quad \quad \vdots \\
p_{n+1} &= T(p_n, q_n).
\end{align*}
\] (2.1)

The vector \(q_i\) is called the decision vector or decision variable and a choice of \(q_i\) is called a decision. As in the outside world, a decision is equivalent to a transformation. We shall be concerned with processes in which the \(q_i\) are chosen so as to maximize a prescribed scalar function of the state and decision variable

\[
R(p, p_1, p_2, \ldots; q_0, q_1, \ldots). 
\] (2.2)
What was formerly the "associated function" is now the criterion function, or return function.

An $N$-stage decision process (of discrete deterministic type, the only kind we know about at the moment) is the set of vectors

$$[p, p_1, p_2, \ldots, p_N; q_0, q_1, \ldots, q_N], \tag{2.3}$$

where $p_{n+1} = T(p_n, q_n)$ for each $n$.

3. Policy

Let us now introduce the important concept of a policy. As indicated above, the values $q_k$ are to be chosen at each stage, in a fashion which, in view of the general nature of the function $R$, is dependent upon the current state of the system, past and future states, and past and future decisions as well. We shall be concerned with criteria $R$ possessing a structure which permits us to focus our attention solely upon the past and present history of the process in a search for values of $q$. In several sections that follow we will exhibit criterion functions possessing this most important property. With these assumptions, we can restrict ourselves to functions of the form

$$q_k = q_k(p, p_1, p_2, \ldots, p_k; q_0, q_1, \ldots, q_{k-1}). \tag{3.1}$$

This function is called the policy function or simply the policy. As one might imagine, in the study of control processes consideration of policies possessing the above structure is not a severe restriction since, in general, decision-making involves only a knowledge of the past. It is well, however, to make all assumptions explicit.

A policy which maximizes the function $R$ is called an optimal policy. Our treatment of multistage decision processes will hinge upon the determination of optimal policies.

There still remains too much generality in the expression in (3.1). We wish to concentrate upon policies which have the simpler form

$$q_k = q_k(p_k), \tag{3.2}$$
a function of the current state and stage of the process. This additional simplification will be a consequence of a still further particularization of the structure of the criterion function $R$, a structure which permits us to make optimal decisions with a knowledge only of the current state of the system.

We have previously emphasized the fact that the state of a system is not an intrinsic property but merely a superimposed mathematical representation. By the simple device of introducing a new state

$$\pi_k = [p_k, p_{k-1}, \ldots],$$  \hspace{1cm} (3.3)

in a higher-dimensional space, we can always write

$$q_k = q_k(\pi_k).$$  \hspace{1cm} (3.4)

Hence, a functional relation such as (3.2) is meaningful only in operational terms, in terms of analytic and computational ease and feasibility. For it to be significant the dimension of $p_k$ must be small and the functional relation of some simplicity.

We shall return to this point in Chapter 3 in connection with the computational aspects of dynamic programming.

4. Separable Criteria

Fortunately, a number of important criteria possess this vital property of divorcing the past from the present. In general, it is easy to predict this property from the very nature of the original multistage decision process. In some situations, however, a little thought or a reformulation is required. Examples of this property which we shall verify below, are provided by the following criteria:

$$\sum_{k=0}^{N} g(p_k, q_k),$$  \hspace{1cm} (4.1a)

$$g(p_N) \text{ (Terminal control)},$$  \hspace{1cm} (4.1b)

$$\max_{k \geq 0} g(p_k, q_k).$$  \hspace{1cm} (4.1c)

Observe that this last criterion is associated with an infinite-stage process.
5. Principle of Optimality

Let us henceforth suppose that we are dealing with decision processes whose formulation permits us to restrict our examination of policies to those which depend only on current states. In this special, but extremely important, case the optimal policy is characterized very simply:

**Principle of Optimality.** *An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*

To illustrate this intuitive result, suppose that we have an \(N\)-stage process, possessing the required separation of past and future, starting in state \(p\), with an optimal sequence of decisions \(q_0, q_1, \ldots, q_{N-1}\), so that \(p_1 = T(p, q_0), p_2 = T(p_1, q_1), \) and so on. Then \(q_1, q_2, \ldots, q_{N-1}\) must represent an optimal sequence of decisions for the \((N - 1)\)-stage process starting in state \(p_1\). Using this observation, we shall derive equations which permit us to study, both analytically and computationally, the optimal policy and the maximum return.

6. Illustrations

Consider the problem of maximizing the return function

\[
R(p, p_1, \ldots, q_0, q_1, \ldots) = \sum_{k=0}^{N} g(p_k, q_k). \tag{6.1}
\]

To dispose quickly of the question of the existence of an actual maximum (rather than say a supremum), we can either restrict \(p\) and \(q\) to range over finite sets of values, or impose reasonable constraints of continuity on the functions \(g(p, q)\) and \(T(p, q)\). Since a computational solution requires finiteness of some form, there is little loss in making the first assumption. In any case, there is seldom any difficulty on this point in connection with \(N\)-stage processes. Consequently, we shall refer the reader to other sources for discussions of existence of optimal policies
II. Multistage Decision Processes

for general multistage processes, and restrict our attention to the formal aspects.

Let then the maximum value of $R$, dependent only upon the initial state $p$ and the number of stages $N$, be denoted by $f_N(p)$. In other words,

$$f_N(p) = \text{the total } N\text{-stage return obtained starting in state } p \text{ using an optimal policy.}$$ \hfill (6.2)

The principle of optimality permits us to conclude that regardless of the initial decision $q_0$, we have, for $N \geq 1$,

$$g(p, q_0) + [g(p_1, q_1) + \cdots + g(p_N, q_N)] = g(p, q_0) + f_{N-1}(T(p, q_0)).$$ \hfill (6.3)

Since this holds for all initial decisions $q_0$, to obtain $f_N(p)$, the maximum return, we must maximize over $q_0$. Thus we obtain the basic functional equation

$$f_N(p) = \max_{q_0} [g(p, q_0) + f_{N-1}(T(p, q_0))],$$ \hfill (6.4)

$N \geq 1$, with

$$f_0(p) = \max_{q_0} g(p, q_0).$$ \hfill (6.5)

Similarly, for the terminal control problem of (4.1b) we have

$$f_0(p) = g(p),$$ \hfill (6.6)

$$f_N(p) = \max_{q_0} f_{N-1}(T(p, q_0)).$$

For the third return function, of (4.1c), we write

$$f(p) = \max_{q_0} \max_{k \geq 0} g(p_k, q_k),$$ \hfill (6.7)

and thus, since $\max_{k \geq 1} u_k = \max [u_1, \max_{k \geq 2} u_k]$,

$$f(p) = \max_{q_0} \max_{k \geq 0} [g(p), f(T(p, q_0))]
= \max_{q_0} [g(p), \max_{q_0} f(T(p, q_0))].$$ \hfill (6.8)

In connection with the equation in (6.8), it is important to note that there can be some difficulty as far as existence and uniqueness of the solution is concerned. Generally, we can
either circumvent or answer these questions, or do both, by using in place of (4.1c) the return function
\[
\max_{0 \leq b \leq N} g(p_b, q_b),
\]
associated with an \(N\)-stage process. We shall elaborate this point in Section 15.

7. Alternate Derivation of Results

For those who may, with some justice, distrust the principle of optimality at this early stage, let us present a proof of (6.4) along conventional lines. We have
\[
\max_{[q_0, q_1, \ldots, q_N]} R = \max_{q_0} \max_{[q_1, q_2, \ldots, q_N]} R.
\]
Hence
\[
f_N(p) = \max_{[q_0, q_1, \ldots, q_N]} [g(p, q_0) + g(p_1, q_1) + \cdots + g(p_N, q_N)]
\]
\[
= \max_{q_0} \max_{[q_1, q_2, \ldots, q_N]} [g(p, q_0) + \cdots]
\]
\[
= \max_{q_0} [g(p, q_0) + \max_{[q_1, q_2, \ldots, q_N]} [g(p_1, q_1)
\]
\[
+ g(p_2, q_2) + \cdots + g(p_N, q_N)]]
\]
\[
= \max_{q_0} [g(p, q_0) + f_{N-1}(T(p, q_0))].
\]
Throughout we have used only the obvious properties of the maximum operation and the separable structure of the return function \(R\).

In dealing with more complex processes, it is usually very much simpler to use the concepts of a multistage decision process than to rely upon analytic manipulation.

8. What Constitutes a Solution?

By means of the functional equation technique of dynamic programming, we have converted the problem of determining a sequence of decisions that maximize a criterion function into that of solving a functional equation. Using the equation of (7.2) to make our discussion concrete, let us examine what we mean by "solution."
II. Multistage Decision Processes

In classical context, a solution means an algorithm for the determination of the sequence \( \{f_N(p)\} \) which is of simpler nature than the recurrence relation of (7.2). In the context of dynamic programming, we have more flexibility. A solution can be given in terms of the sequence \( \{f_N(p)\} \), the sequence of return functions, or the sequence \( \{g_N(p)\} \), the sequence of policy functions. It is easy to see that each sequence determines the other. Note, however, that there is only one sequence of optimal return functions, although there may be many optimal policies which yield the same maximum return.

In Section 11, this duality will be examined from the geometric point of view.

9. Continuous Multistage Decision Process

To introduce the important and useful concept of a continuous multistage decision process, we proceed as before from the discrete to the continuous by means of a limiting process. Let \( \Delta \) be an infinitesimal, let the return function be

\[
\sum_{k=0}^{N} g(p_k, q_k) \Delta,
\]

and take the transformation to have the form

\[
T(p, q) = p + S(p, q) \Delta.
\]

Let decisions be made at times 0, \( \Delta, 2\Delta, \ldots \), and set \( N\Delta = T \). Then if we let \( f_N(p) = f(p, T) \) denote the maximum of the return function, we obtain in a familiar fashion

\[
f(p, T) = \max_q [g(p, q) + f(p + S(p, q)\Delta, T - \Delta)].
\]

Expanding in powers of \( \Delta \) and passing to the limit as \( \Delta \to 0 \), we obtain formally the partial differential equation

\[
\frac{\partial f}{\partial T} = \max_q [g(p, q) + (\text{grad } f, S(p, q))],
\]

with the initial condition \( f(p, 0) = 0 \).

We shall illustrate this approach in the next section in connection with some problems from the calculus of variations. There are many points of rigor that require investigation in
connection with the foregoing formalism. These can be disposed of in several different ways, albeit with some ingenuity and analytic skill required.

10. Calculus of Variations

A most interesting and significant example of a continuous multistage decision process is furnished by the calculus of variations. Let us consider one of the simplest problems that can arise, that of maximizing an integral

\[ J(u) = \int_{0}^{T} g(u, u') \, dt \]  \hspace{1cm} (10.1)

over all functions \( u(t) \) defined over \([0, T]\), for which the integral exists, and satisfying the initial condition \( u(0) = c \). Proceeding formally, and assuming that all the conditions for the existence of a maximum are met, we introduce the function

\[ f(c, T) = \max_{u} J(u). \]  \hspace{1cm} (10.2)

At any particular time \( t \), the state variables are \( c = u(t) \), and \( T - t \), the time remaining in the process. A decision involves a choice of a direction, the value of \( u'(t) \). Assuming, as we shall, that we are dealing with "smooth" processes, defined in terms of the existence say of second derivatives of \( u(t) \), a choice of \( u'(t) \) determines \( u(t) \) over an infinitesimal interval \([t, t + \Delta]\).

Consider then the problem of determining the initial slope, \( u'(0) \), which we denote by \( v \).

Writing

\[ \int_{0}^{T} = \int_{0}^{\Delta} + \int_{\Delta}^{T}, \]  \hspace{1cm} (10.3)
and making the approximation
\[ \int_0^\Delta g(u, u') \, dt = g(c, v)\Delta + o(\Delta), \quad (10.4) \]
we see that the principle of optimality yields the relation
\[ f(c, T) = \max_v [g(c, v)\Delta + f(c + v\Delta, T - \Delta)] + o(\Delta). \quad (10.5) \]

Expanding in powers of \( \Delta \) and passing to the limit as \( \Delta \to 0 \), we obtain the partial differential equation
\[ \frac{\partial f}{\partial T} = \max_v \left[ g(c, v) + v \frac{\partial f}{\partial c} \right], \quad (10.6) \]
with \( f(c, 0) = 0 \). This determines both the maximum value of the integral and the policy function \( v = v(c, T) \).

Observe that this is an initial value problem, whereas the usual formulation by way of the Euler equation
\[ \frac{\partial g}{\partial u} - \frac{d}{dt} \left( \frac{\partial g}{\partial u'} \right) = 0, \quad (10.7) \]
is a two-point boundary-value problem. The solution (or solutions) is determined by the initial condition \( u(0) = c \) and the variational condition
\[ \frac{\partial g}{\partial u'} \bigg|_{t=T} = 0. \quad (10.8) \]

Consider now the situation where there is a constraint on the rate of change of \( u \), say
\[ |u'(t)| \leq k, \quad 0 \leq t \leq T. \quad (10.9) \]

This introduces a good deal of complication into the classical formulation. As far as the dynamic programming approach is concerned, it suffices to replace (10.6) by
\[ \frac{\partial f}{\partial T} = \max_{|v| \leq k} \left[ g(c, v) + v \frac{\partial f}{\partial c} \right], \quad f(c, 0) = 0. \quad (10.10) \]
As we shall indicate in Chapter 3, the imposition of constraints of this type actually simplifies the computational solution.
11. Geometric Aspects

EXERCISES

1. If \( f(c, T) = \min_w \int_0^T [u^2 + v^2] \, dt \), where \( u' = au + w \) \((w \equiv w(t))\), \( u(0) = c \), show that
   \[
   f_T = \min_v [c^2 + v^2 + (ac + v)f_c], \quad f(c, 0) = 0,
   \]
   and thus that
   \[
   f_T = c^2 + acf_c - \frac{f_c^2}{4}.
   \]

2. Show that \( f(c, T) = c^2 g(T) \), where
   \[
   g'(T) = 1 + 2ag(T) - g^2(T), \quad g(0) = 0.
   \]
   Hence, determine \( f(c, T) \) and \( v(c, T) \) explicitly, and thus their asymptotic behavior as \( T \to \infty \). Verify the solution using the Euler equation.

3. Show that the problem of maximizing \( J(u) = \int_0^T g(u, u', t) \, dt, u(a) = c \), leads to the nonlinear partial differential equation
   \[
   -\frac{\partial f}{\partial a} = \max_v \left[ g(c, v, a) + v \frac{\partial f}{\partial c} \right], \quad f(T, c) = 0,
   \]
   for the function \( f(a, c) = \max_u J(u) \).

4. Show that the problem of maximizing \( J(u) = \int_0^T g(u, v) \, dt \), over \( v \), where \( du/dt = h(u, v) \), \( u(0) = c \), leads to the equation
   \[
   \frac{\partial f}{\partial T} = \max_v \left[ g(c, v) + h(c, v) \frac{\partial f}{\partial c} \right].
   \]

5. Consider the problem of maximizing \( J(u) = \int_0^T g(u, v) \, dt \), over \( v \) where \( du/dt = h(u, v) \), \( u(0) = c \), and \( \int_0^T k(u, v) \, dt \leq a \), in two ways: first by taking \( a \) to be a state variable, and secondly by using a Lagrange multiplier. Obtain a relation between the two formulations.

6. Obtain a partial differential equation for the smallest characteristic value of \( u'' + \lambda \varphi(t)u = 0 \), \( u(0) = 0 \), \( u(a) = 0 \), by considering the associated variational problem of minimizing \( \int_0^a u'^2 \, dt \), subject to \( \int_0^a \varphi(t)u^2 \, dt = 1 \), \( u(a) = c \), \( u(0) = 0 \).

11. Geometric Aspects

It is instructive to compare the dynamic programming approach with the classical approach in the treatment of the
II. Multistage Decision Processes

foregoing variational problem. Classically one seeks a curve $u = u(t)$ defined over the given interval $[0, T]$ which maximizes. The unknown function $u$ is regarded as a point in function space. In our approach, at each point, we seek a direction which is optimal; the solution is obtained in the form of a policy, a set of instructions for carrying out the process.

![Diagram of classical and dynamic programming approaches](image)

In geometric parlance, we can say that the classical view is that of a curve as a locus of points, while dynamic programming considers a curve to be an envelope of tangents. Hence the two theories are dual to each other, a fact which manifests itself constantly. This duality and equivalence remains valid, however, only for deterministic processes.

It follows that use of the two approaches jointly will be more powerful and more enlightening than use of one method or the other alone. This holds true for both analytic and computational aspects, where, in both areas, dual methods play a fundamental role.

12. Stochastic Multistage Decision Processes

Let us now allow stochastic effects, but restrain our discussion only to processes of discrete type in order to keep the mathematical level moderate. It suffices for the purposes of illustration to take the transformation to have the form $T = T(p, q, r)$, and to suppose that we agree to maximize an expected value. This expected value need not necessarily be of the return function itself, but can be the average of a function of the return function. The exercises will contain some examples of this.
Suppose we wish to maximize the expected value of

$$\sum_{k=0}^{N} g(p_k, q_k)$$  \hspace{1cm} (12.1)$$

where \( p_0 = p \) and \( p_{k+1} = T(p_k, q_k, r_k) \), under the assumption that the \( r_k \) are independent random variables with a common distribution function \( dG(r) \). Proceeding as before, let \( f_N(p) \) denote the maximum value. Then

$$f_0(p) = \max_{q_0} g(p, q_0)$$  \hspace{1cm} (12.2)$$

and, for \( N \geq 1 \), we obtain, using the principle of optimality,

$$f_N(p) = \max_{q_0} \left[ g(p, q_0) + \int f_{N-1}(T(p, q_0, r_0)) \, dG(r_0) \right].$$  \hspace{1cm} (12.3)$$

It is important to observe that both deterministic and stochastic decision processes are treated by means of the same formalism, and, furthermore, that the functional equation technique handles multistage processes of decision type or not in the same fashion.

13. Feedback Control

It is essential to make precise the structure of the process we have been considering. We suppose that we observe that the system is in state \( p \), that we then choose \( q_0 \), that \( r_0 \) is then chosen from the distribution function \( dG(r) \), and finally that \( p_1 = T(p, q_0, r_0) \) is then generated, and observed. Using the knowledge of the new state, \( q_1 \) is chosen; \( r_1 \) is then chosen from the distribution function \( dG(r) \), and so on. This is the basic concept of feedback control.

In the deterministic case, whether the \( q_k \) are chosen sequentially, as above, or all at the same time, as in the usual formulation, is a matter of analytic or computational convenience. The maximum returns and optimal policies are the same. This fortuitous equivalence (the duality of Euclidean geometry in particular cases; see Section 11) unfortunately obscures the difference between these quite distinct types of optimizations. The fact that in the deterministic case, the feedback aspects can
be ignored prevented a development of an adequate mathematical theory for general multistage decision processes for some time.

In the stochastic case, the two approaches correspond to quite different processes, and indeed to extremes of a family of processes. In the case treated above, we suppose that we can observe the state of the system at each stage. On the other hand, if we can never observe the states of the system once the process begins, then we are forced to choose the vectors $q_0$, $q_1$, $\ldots$, $q_N$ at the outset and to maximize the function $\exp R$ over these vectors.

We see that feedback control involves a choice of functions, $\{q_k(p)\}$, the policy functions. Hence, it is on a more sophisticated level than the usual maximization over vectors $q$. Nonetheless, it is easier to deal with analytically and computationally.

Let us point out, finally, we can generate an infinite class of stochastic control processes by varying the types of information concerning the state of the process that we allow the decision-maker. As pointed out before, most of these induce infinite-dimensional decision processes, where the state is a probability distribution.

14. Analysis of the Equations

Once various classes of decision processes have been formulated in precise analytic terms, we face the problem of obtaining useful information from the equations. Specifically, we can ask the following questions:

1. For what classes of processes can we obtain simple analytic expressions for the return functions and policies?

2. Under what conditions can we deduce structural characteristics of the optimal policy?

3a. When can we obtain numerical solutions with the aid of digital computers?

3b. Under what conditions can we find simple asymptotic or steady-state policies and return functions?
Since a great deal has been done in these directions, detailed answers to any of these questions would take us too far afield and defeat the purposes of an introductory account. We shall, however, supply a number of references in the exercises and in the bibliography at the end of the chapter. Furthermore, we shall make some general remarks below, and consider a number of important special processes in Chapters 4 and 5.

15. Successive Approximations

Since we have principally been concerned with finite-stage processes, most of the functional equations have taken the form of a nonlinear recurrence relation.

\[ f_N(p) = \max_q [g(p, q) + f_{N-1}(T(p, q))], \quad N \geq 1, \quad (15.1) \]

with \( f_0(p) = \max_q g(p, q) \) or, in any case, a known function. It follows that no questions of existence or uniqueness of the sequence of return functions \( \{f_n(p)\} \) arise in the consideration of equations of this type. Given \( f_0 \), we determine \( f_1 \), then \( f_2 \), and so on.

If we consider, as an approximation to the solution of (15.1) for the case where \( N \gg 1 \), the case of an infinite number of stages, we obtain the equation

\[ f(p) = \max_q [g(p, q) + f(T(p, q))], \quad (15.2) \]

an equation which arises naturally in the consideration of various generalized trajectory and control processes.

The classical approach to the demonstration of a solution of an equation of this type, under appropriate hypotheses concerning the functions \( g(p, q) \) and \( T(p, q) \), is based on the method of successive approximations. This is usually carried out in the following fashion. We guess an initial approximation, \( f_0(p) \), and then proceed recurrently to determine the subsequent approximation to \( f(p) \),

\[ f_1(p) = \max_q [g(p, q) + f_0(T(p, q))], \]

\[ . \]

\[ . \]

\[ f_{n+1}(p) = \max_q [g(p, q) + f_n(T(p, q))]. \quad (15.3) \]
In a number of processes, it is not difficult to establish the convergence of the sequence \( \{f_n(p)\} \) to a solution of (15.2). Details of these proofs under various assumptions will be found in the references given at the end of the chapter.

16. Approximation in Policy Space

In the theory of dynamic programming we possess another type of approximation, approximation in policy space. As we have pointed out in Section 8, a functional equation of the type appearing in (15.2) determines two functions, the return function \( f(p) \), and the policy function \( q(p) \). The fact that each determines the other is the key to the method we now discuss. Let us begin by approximating to \( q(p) \) with the function \( q_0(p) \).

Next, determine \( f_0(p) \) by means of the equation

\[
f_0(p) = g(p, q_0) + f(T(p, q_0))
= g(p, q_0) + g(T(p, q_0), q_0') + \cdots,
\]

(16.1)

where \( q_0' = q_0(T(p, q_0)) \), and so on. In other words, \( f_0(p) \) is the total return obtained using the policy \( q_0(p) \).

Let us improve this policy by determining the function \( q_1(p) \) which maximizes \( g(p, q) + f_0(T(p, q)) \), and then determining \( f_1(p) \) by means of the equation

\[
f_1(p) = g(p, q_1) + f_1(T(p, q_1)).
\]

(16.2)

Since \( q_1 = q_1(p) \), we can write this in the form

\[
f_1(p) = g(p) + f_1(h(p)),
\]

(16.3)

where \( g \) and \( h(p) \) are known functions, \( h(p) = T(p, q_1(p)) \). Under plausible assumptions concerning these functions, we can solve for \( f_1(p) \) by means of an iterative procedure. Thus

\[
f_1(p) = g(p) + g(h(p)) + g(h^{(2)}(p)) + \cdots.
\]

(16.4)

In other words, \( f_1(p) \) is obtained using the policy \( q_1(p) \). Since

\[
f_0(p) = g(p, q_0) + f_0(T(p, q_0))
\leq \max_q [g(p, q) + f_0(T(p, q))]
= g(p, q_1) + f_0(T(p, q_1)),
\]

(16.5)
we see that
\[
\begin{align*}
f_0(p) & \leq g(p) + f_0(h(p)) \\
& \leq g(p) + g(h(p)) + g(h^{(2)}(p)) + \cdots. \quad (16.6)
\end{align*}
\]
Hence, again under reasonable conditions we can demonstrate that \( f_0(p) \leq f_1(p) \).

Continuing in this way, we obtain a new policy \( q_\alpha(p) \) from the function \( f_1(p) \), and, inductively, a sequence of approximations to \( f(p) \), \( \{f_n(p)\} \), which is monotone increasing in \( n \), \( f_1 \leq f_2 \leq \cdots \leq f_n < f \).

One advantage of this approach lies in the fact that in many processes there exist intuitive approximations in policy space which can easily be used in this fashion. Furthermore, and this is a most important point as far as applications are concerned, the foregoing yields a systematic technique for improving existing policies.

17. Quasilinearization

One important application of this idea of approximation in policy space is to the theory of quasilinearization. We begin with an equation of the form
\[
L(f(p)) = N(f(p)), \quad (17.1)
\]
where \( L \) is a linear operator and \( N \) is nonlinear, and attempt to transform it to an equation of the type (15.2). To do this we wish to express the right-hand side as a maximum. Suppose, for the sake of simplicity, that \( N(u) \) is a scalar function which is convex in \( u \). Then, as is easy to see, either geometrically or analytically,
\[
N(u) = \max_v [N(v) + (u - v)N'(v)]. \quad (17.2)
\]
Hence (17.1) can be written
\[
L(f) = \max_q [N(q) + (f - q)N'(q)]. \quad (17.3)
\]
This equation looks as if it is associated with a multistage decision process and we can proceed as if it actually were. Employing approximation in policy space, very rapidly con-
verging approximations to the solution of (17.1) can often be obtained. There are many important applications of this technique described in a number of references given at the end of the chapter.

18. Analytic and Computational Aspects

Can the equations derived in the preceding pages be used to obtain analytic and computational results? The answer is a qualified "yes." As usual in mathematical theories, it depends upon the problem. In the fourth chapter we shall consider some important types of processes where quite satisfactory analytic progress can be made; in the following chapter, Chapter 3, we shall discuss computational techniques which permit us to handle general classes of dynamic programming processes, granted the use of a modern computer.

As will be seen, there is nothing routine about the application of digital computers to numerical solution. A great deal of analytic skill and ingenuity is generally required, and, in the main, far more understanding of both the scientific background and mathematical techniques is required. Only if we can produce an algorithm for obtaining numerical results can we feel that a problem has been completely resolved.

19. Fermat’s Principle and the Eikonal Equation

As an interesting application of the preceding ideas, let us consider the propagation of light in an inhomogeneous two-dimensional medium. Our aim is to show that by applying the principle of optimality in conjunction with Fermat’s principle of least time we are led in a simple and direct fashion to the basic equation of geometrical optics, the eikonal equation.

We suppose that the medium is characterized by an index of refraction which is a function of position, \( n = n(x, y) \). Then if we set

\[
v(x, y) = \frac{c}{n(x, y)},
\]

(19.1)

\( v \) is the speed of light at the point \( (x, y) \), where \( c \) is the speed of light in a vacuum. According to Fermat’s principle, in
passing from one point \((x, y)\) to another point \((x_0, y_0)\) in the medium, a light ray will select a path of least time. Let us denote the time that is consumed by a particle of light in passing from a generic point \((x, y)\) to a fixed point \((x_0, y_0)\), according to the foregoing principle, by \(t(x, y)\). If the particle leaves the point \((x, y)\) and travels along a line making an angle \(\theta\) with the \(x\)-axis for a distance \(\Delta\), it will arrive at the point \((x + \Delta \cos \theta, y + \Delta \sin \theta)\) after a time \(\Delta/v(x, y)\), where we neglect terms in \(\Delta\) of order greater than one. It follows that

\[
t(x, y) = \min_{\theta} \left\{ \frac{\Delta}{v(x, y)} + t(x + \Delta \cos \theta, y + \Delta \sin \theta) + o(\Delta) \right\}. \tag{19.2}
\]

From this we see that

\[
0 = \min_{\theta} \left\{ \frac{\Delta}{v(x, y)} + \Delta \cos \theta + \Delta \sin \theta + o(\Delta) \right\}, \tag{19.3}
\]
or upon dividing by \(\Delta\) and letting \(\Delta\) tend to zero,

\[
-v^{-1}(x, y) = \min_{\theta} \left\{ t_x \cos \theta + t_y \sin \theta \right\}. \tag{19.4}
\]

This is a novel form of the eikonal equation.

The choice of \(\theta\) which minimizes the expression in brackets in the above equation is a solution of the equation

\[-t_x \sin \theta + t_y \cos \theta = 0, \tag{19.5}\]
or

\[\tan \theta = \frac{t_y}{t_x}. \tag{19.6}\]

Upon substituting this choice of \(\theta\) into Eq. (19.4) we obtain our desired result,

\[t_x^2 + t_y^2 = v^{-2}, \tag{19.7}\]

which can be rewritten in the form

\[t_x^2 + t_y^2 = \frac{n^2(x, y)}{c^2}. \tag{19.8}\]

Equation (19.8) is the eikonal equation.

Next we observe that if a disturbance is initiated at the point
(x₀, y₀) at time zero, then at time T the wave front of this disturbance will have as its equation

\[ T = t(x, y). \] (19.9)

The slope of the ray passing through the point (x, y) is \( t_y/t_z \), according to Eq. (19.6), and the slope of the wave front is \(-t_x/t_y\). This shows that the rays and wave fronts are orthogonal in the case considered.

**EXERCISES**

1. Discuss the case of an inhomogeneous and anisotropic medium for which \( n = n(x, y, y') \). In particular find the relationship between the rays and wave fronts.
2. Find the equations of the rays in the isotropic case.
3. Examine the case in which \( n(x, y) = y \). Show that the rays are arcs of circles with their centers on the x-axis.
4. Discuss the concept of approximation in policy space as it applies to Eq. (19.4), an alternative form of the eikonal equation.

**20. Shortest Paths through Networks**

Consider a network consisting of \( N \) nodes numbered 1, 2, \ldots, \( N \), and interconnecting links. Let the time that it takes to traverse the link \((i, j)\) be \( t_{ij} > 0 \). In particular, note that \( t_{ij} \) need not equal \( t_{ji} \). We wish to determine a path through the network which connects two given nodes and which involves a total time of traverse at least as short as that yielded by any other path connecting the given nodes. Clearly this problem is of importance in the routing of airplanes and automobiles through transportation networks and of messages through communication networks. In addition it is of significance in the study of time-optimal feedback control systems. To see this, we interpret the \( N \) nodes as the possible states of a system
and the links as transformations from one state to another. Frequently, we wish to transform a system from a given initial state into a desired terminal state in minimal time. This problem may also be viewed as a discrete version of the optical problem involving Fermat's principle considered in Section 19. Lastly we note that by interpreting the numbers $t_{ij}$ as fuel, or any other resources consumed, we can regard our problem as that of transforming a system from a given initial state into a desired terminal state in the most efficient manner, with the costs of the individual transformations additive.

Let us consider the node $N$ to be the desired terminal node and let us introduce the quantities

$$u_i = \text{the time to transform the system from the initial state } i \text{ to the desired terminal state } N \text{ along a shortest path, } i = 1, 2, \ldots, N. \quad (20.1)$$

We set

$$u_N = 0. \quad (20.2)$$

Use of the principle of optimality yields the basic system of nonlinear algebraic equations

$$u_i = \min_{j \neq i} (t_{ij} + u_j), \quad i = 1, 2, \ldots, N - 1, \quad (20.3)$$

$$u_N = 0.$$

These may be considered to be a discrete form of the eikonal equation of Section 19.

In contrast to many of the other systems of equations derived from the principle of optimality no sequential method of determining the unknowns presents itself naturally. Hence we shall use a method of successive approximations. First, though, we shall establish the uniqueness of the solution of Eqs. (20.3), in order to insure that sequences which converge to a solution converge to the same solution, namely, the desired minimal times. Notice, though, that while the quantities $u_1, u_2, \ldots, u_N$ are unequivocally defined, the paths for achieving these values may not be.

Let $u_1, u_2, \ldots, u_N$ and $U_1, U_2, \ldots, U_N$ be two different solutions of the system (20.3). Let $m$ be an index for which the
difference \( U_i - u_j, j = 1, 2, \ldots, N \), is a maximum. We wish to show that this maximum difference is zero. Let

\[
U_m = t_{mn} + U_n \leq t_{mr} + U_r, \quad (20.4)
\]
\[
u_m = t_{mr} + u_r. \quad (20.5)
\]

From this it follows that

\[
U_m - \nu_m \leq U_r - \nu_r, \quad (20.6)
\]
and since \( m \) is an index for which the difference \( U_m - \nu_m \) is a maximum, equality must hold in both Eqs. (20.6) and (20.4). Furthermore it is clear that

\[
m \neq r. \quad (20.7)
\]

In a similar fashion we can find a node \( s \), with

\[
s \neq m, \quad s \neq r, \quad (20.8)
\]
for which

\[
U_m - \nu_m = U_r - \nu_r = U_s - \nu_s. \quad (20.9)
\]

Ultimately, however, since there are only a finite number of nodes, we must come upon the node \( N \) for which

\[
U_N - \nu_N = 0. \quad (20.10)
\]

This completes the proof of uniqueness. Next we turn to the numerical solution via an application of approximation in policy space. Simultaneously, we shall establish both the existence of a solution of Eqs. (20.3) and exhibit an effective means for calculating this solution.

As our initial approximation, \( u_i^{(0)}, i = 1, 2, \ldots, N \), let us set

\[
u_i^{(0)} = t_{iN}, \quad i = 1, 2, \ldots, N, \quad t_{NN} = 0, \quad (20.11)
\]
which corresponds to the policy of transforming the system directly from its initial state into the desired terminal state. If there is no direct link from \( i \) to \( N \), so that \( t_{iN} = \infty \), we set \( t_{iN} = 10^{30} \) or some other suitably large number. This simple device avoids difficult topological questions of connectivity. The next approximation, \( u_i^{(1)} \), is obtained via the formulas

\[
u_i^{(1)} = \min_{j \neq i} \{ t_{ij} + u_j^{(0)} \}, \quad i = 1, 2, \ldots, N - 1, \quad (20.12)
\]
\[
u_N^{(1)} = 0. \quad (20.13)
\]
20. Shortest Paths through Networks

The operations indicated in formula (20.12) are readily carried out computationally. Values of one row of the matrix \((t_{ij})\) are required as are values of the initial approximation \(u_i^{(0)}, j = 1, 2, \ldots, N\). The minimization is done by a direct comparison of the sums involved, one after the other. Thus, both computing times and high-speed storage requirements are small. We proceed from the \(k\)th to the \((k + 1)\)st approximation by means of the relationships

\[
u_i^{(k+1)} = \min_{j \neq i} \{t_{ij} + u_j^{(k)}\}, \quad i = 1, 2, \ldots, N - 1, \quad (20.14)\]

\[
u_N^{(k+1)} = 0. \quad (20.15)\]

Formulas (20.12), (20.13), (20.14), and (20.15) have a simple physical interpretation. The quantity \(u_i^{(0)}\) represents the minimal time to pass from node \(i\) to node \(N\) with at most zero intermediate nodes. The quantity \(u_i^{(1)}\) is the minimal time to pass from node \(i\) to node \(N\) with at most one intermediate node. Similarly, \(u_i^{(k)}\) is the minimal time to pass from node \(i\) to node \(N\) with at most \(k\) intermediate nodes. With this physical interpretation in mind we see immediately that the sequence of approximations is monotone decreasing; i.e.,

\[0 \leq u_i^{(k+1)} \leq u_i^{(k)}, \quad i = 1, 2, \ldots, N. \quad (20.16)\]

This is an illustration of the efficacy of approximation in policy space. Let us furthermore note that this approximation procedure is finite in nature. Since an optimal path from \(i\) to \(N\) can have no loops, it can contain no more than \(N - 2\) intermediate nodes. It follows that the quantities \(u_i^{(N-2)}, i = 1, 2, \ldots, N - 1\), must satisfy the equations (20.3), which means that there are at most \(N - 2\) iterations.

If we wish to determine the optimal trajectories from any initial state to any terminal state, we could apply the technique just discussed \(N\) times. Alternatively, we can introduce the quantities

\[
u_{ij}^{(k)} = \text{the minimal time to transform a system from state } i \text{ into state } j \text{ using a path with at most } k \text{ intermediate states.} \quad (20.17)\]
II. Multistage Decision Processes

From the principle of optimality there follow the equations

\[ u^{(k+1)}_{im} = \min_{j \neq i} \{ t_{ij} + u^{(k)}_{jm} \}, \quad k = 0, 1, 2, \ldots, N - 3. \quad (20.18) \]

For computational purposes, though, it may be advantageous to leapfrog ahead by means of the formula

\[ u^{(2k+1)}_{ij} = \min_{j \neq i} \{ u^{(k)}_{im} + u^{(k)}_{mj} \} \quad (20.19) \]

which permits us to determine the sequence of matrices

\( (u^{(0)}_{ij}), (u^{(1)}_{ij}), (u^{(3)}_{ij}), (u^{(7)}_{ij}), (u^{(15)}_{ij}), \ldots. \)

In various contexts it is desirable to determine preferred suboptimal paths. Let us show how to determine second shortest paths. Let

\[ v_i = \text{the length of a second shortest path from} \quad i \quad \text{to} \quad N, \quad i = 1, 2, \ldots, N. \quad (20.20) \]

We observe that if we first pass from state \( i \) to state \( j \), then the continuation from state \( j \) to the terminal state \( N \) must be via a shortest, or second shortest, path. Using the notation

\[ \min_2 (a_1, a_2, \ldots, a_k) = \text{the second smallest of} \quad a_1, a_2, \ldots, a_k, \text{assuming} \]

that they are not all equal, \( (20.21) \)

we may write

\[ v_i = \min_2 (t_{ij} + u_{ij}, t_{ij} + v_j), \quad i = 1, 2, \ldots, N - 1, \quad (20.22) \]

\[ v_N = 0. \]

Assuming that the \( u_i \) have already been determined, we can use the foregoing procedures to determine the \( v_i \).

**EXERCISES**

1. Let the number of nodes in a network be \( N = 3, 4, \) and \( 5, \) and let \( t_{ij} = i + j \). Find the shortest paths from node 1 to node \( N \) in each case.

2. In many situations only a few of all the possible links will be present. How does this fact modify the equations given above?

3. Suppose that we wish to travel from \( i \) to \( N \) in such a way that the largest of the numbers \( t_{ij} \) on the links which we traverse is as small as possible. Show that this leads to the functional equations

\[ s_i = \min_{j \neq i} \max (t_{ij}, s_j), \quad i = 1, 2, \ldots, N - 1, \]

\[ s_N = 0. \]
BIBLIOGRAPHY AND COMMENTS

§1. Detailed accounts of the theory and application of dynamic programming will be found in

§8. For a further discussion of these ideas, see

§10. For detailed discussions of interconnections between dynamic programming and the calculus of variations, see
In this last paper, Berkovitz discusses variational problems subject to constraints, and derivation of the Pontryagin maximum principle from results of Valentine and others in the calculus of variations. For a further discussion of the connection between dynamic programming and the Pontryagin maximum principle, see
For a discussion of interrelations between dynamic programming, Fermat's principle, and some ideas of Pascal, see
II. Multistage Decision Processes


§12. See


For a treatment of stochastic variational processes of continuous type, see


§13. For a history of the feedback control concept, see


§16. For some applications of approximation in policy space, see the books by Bellman and Dreyfus, and Howard, referred to above.

§17. See


§19. See


§20. See


Chapter III

Computational Aspects

1. Introduction

In the previous two chapters, we have indicated the application of the analytic formalism of the theory of dynamic programming to the study of general multistage decision processes. Let us now turn to the question of the feasibility of numerical solution by means of this approach, and, in particular, to computational solution by means of a digital computer. We shall emphasize only a few of the high points, referring the interested reader to other works where detailed and comprehensive treatments may be found.

A point which cannot be overemphasized is that every numerical procedure involves a control process in which the various errors attendant upon arithmetical operations must be kept within feasible bounds. Thus, dynamic programming techniques can be used in the numerical solution of the classical
functional equations, ordinary and partial differential equations, and in the computational solution of the functional equations of dynamic programming themselves. We shall give some references to applications of this kind, but we will not pursue these ideas in depth.

2. Digital Computers

A digital computer is a device for doing arithmetic, which is to say, the carrying out of the four fundamental operations, addition, subtraction, multiplication, and division. Its prominent position in modern science is a consequence of the fact that, unlike a human being, it can perform these operations accurately and extremely rapidly. At the present time, two ten digit numbers can be multiplied together in about one hundred thousandth of a second. A major disadvantage of the digital computer is that, unlike a human being, every problem to which it can be applied must be reduced, in often a quite tedious fashion, to a sequence of arithmetical problems.

Observe that we do not say that a digital computer is "merely" an arithmetic device. The capacity to perform billions of arithmetic operations quickly and accurately profoundly affects every aspect of modern life and philosophy itself. A significant change in quantity is a significant change in quality.

There are other types of computers such as analog computers which do not possess this limitation to arithmetic operations. In return, however, they are, at present, restricted in their use to the consideration of quite specialized classes of problems. This situation can easily change in the near future.

In what follows, we wish to sketch the way in which digital computers may be applied to determine optimal policies and return functions.

3. Computational Solution of Dynamic Programming Processes

Consider a typical recurrence relation

\[ f_N(p) = \max_q [h(p, q) + f_{N-1}(T(p, q))], \quad (3.1) \]

for \( N \geq 1 \), with \( f_0(p) \) given.
To calculate both the return functions and the policy functions recurrently, we proceed in the following fashion. Consider the equation for $f_i(p)$,

$$f_i(p) = \max_q [h(p, q) + f_0(T(p, q))].$$  \hspace{1cm} (3.2)

In order to calculate the values of $f_i(p)$, it is necessary to have the values of $h(p, q)$, $T(p, q)$, and $f_0(p)$ stored in the computer, or procedures for generating these values. For the moment, let us suppose we are merely storing values. Subsequently, we shall discuss a more sophisticated approach.

Since it is impossible to store all values of $f_i(p)$, if $p$ runs over an infinite set, and inconvenient even if this set of $p$-values is finite, but large, we store a discrete set of values, $\{f_i(p_j)\}$, $i = 1, 2, \ldots, M$. Other values of $f_i(p)$ will be obtained by means of extrapolation or interpolation, where required.

To determine the maximum over $q$ in (3.2), we, in an analogous fashion, restrict $q$ to a finite set of values $\{q_j\}$, $j = 1, 2, \ldots, R$, and proceed by straightforward comparison of values. We evaluate the expression $h(p_1, q) + f_0(T(p_1, q))$ for $q = q_1, q_2, \ldots, q_R$, and choose the value of $q$, say $q_1$, which maximizes. Maximization over a discrete set is an arithmetic operation which can readily be carried out by digital computer. If several values yield the maximum, we store them all.

The foregoing procedure determines $f_i(p_1)$ and $q_1(p_1)$. Repeating this procedure for each of the states $p_i$, $i = 1, 2, \ldots, M$, yields both the return function $f_i(p)$ and the policy function $q_i(p)$ for this set of states.

Replacing (3.2) by the relation

$$f_2(p) = \max_q [h(p, q) + f_1(T(p, q))],$$  \hspace{1cm} (3.3)

we obtain similarly $f_2(p)$ and $q_2(p)$, again for the finite set $p = p_1, p_2, \ldots, p_M$. Thus, recurrently, we compute the sequences $\{f_n(p)\}$ and $\{q_n(p)\}$.

To begin with, let us observe that the foregoing procedure is iterative. We employ the same procedure at each stage to compute $f_{n+1}(p)$, given $f_n(p)$. Consequently, the set of computer
instructions is relatively small and easily written down in one of the new simplified computer languages such as FORTRAN or ALGOL. This iterative, or repetitive, property lends itself ideally to digital computer usage.

Note also that once \( f_1(p) \) has been used to determine \( f_2(p) \) and \( q_2(p) \), via (3.3), it can be washed out of the fast memory. To determine \( f_3(p) \) and \( q_3(p) \), we need only \( f_2(p) \). This is an extremely important fact, since, as we shall see, in these days of relatively limited computers, rapid-access storage considerations are crucial to the success of dynamic programming.

Finally, observe that if we had a constraint of the form \( q \in S(p) \), then the number of possible \( q \)-values over which to search at each stage is diminished. This reduces the computing time. Hence, the more constraints on the policy function, the easier it is to apply dynamic programming techniques. This is an important advantage of these techniques as compared to classical techniques which handle constraints in an awkward fashion.

4. Approximation in Function Space

A number of important processes lead to functional equations of the form

\[
f(p) = \max_q [h(p, q) + f(T(p, q))], \tag{4.1}
\]
as we have indicated in Chapter 2. To reduce the computational solution of equations of this form to the level of the equations of recurrence type treated in Section 3, we invoke the classical method of successive approximations.

Let \( f_0(p) \) now represent an initial approximation, deduced from mathematical or physical considerations, and consider the sequence \( \{f_n(p)\} \) generated by the relation

\[
f_{n+1}(p) = \max_q [h(p, q) + f_n(T(p, q))], \tag{4.2}
\]

\( n = 0, 1, 2, \ldots \). Under various hypotheses concerning the functions \( h(p, q) \), \( T(p, q) \), and \( f_0(p) \), it can be shown that the sequence converges to a solution of (4.1). Generally, the same
assumptions can be used to establish the fact that (4.1) has a unique solution.

5. Approximation in Policy Space

A resource possessed by dynamic programming, but non-existent in classical analysis, is approximation in policy space. As we have emphasized before, the return function is determined by the optimal policy and, conversely, the optimal policy is determined by the return function.

In the previous section, we followed the classical path in approximating to the return function. Let us once again note that we can also employ approximation in policy space. This has been discussed in Section 16 of Chapter 2 and illustrated in Section 20 of that chapter.

The analytic and computational efficacy of this technique depend upon a combination of understanding of the underlying decision process and ingenuity cum serendipity.

6. Computational Feasibility

Let us now examine the algorithm presented in Section 3 in some detail. Is it a feasible computational technique as far as present-day, and even future, digital computers are concerned? We are thinking in terms of rapid access storage of the order of $10^6$, or, at most $10^8$.

Consider the case where $p$ is a vector of dimension 2, $p = (x, y)$, so that $f(p) = f(x, y)$. Let $x$ and $y$ range over the interval $0 \leq x, y \leq 1$, and suppose that we store the values of $f$ at 100 $x$-points and 100 $y$-points, a total of $10^4$ points. This is within our current capabilities. Thus, if $p$ is one-dimensional or two-dimensional, we can consider ourselves in possession of a routine computational solution.

If $p$ is a vector of higher dimension, say 4, the routine approach described above involves the storage of $10^8$ values. This exceeds our expectations, looking say 25 years into the future.* More adroit techniques must be applied if we wish to treat these processes now.

* We are being conservative; ten years is a better estimate. In 25 years, we can reasonably expect fadv memories of $10^{10}$. 
7. Polynomial Approximation

In the most important applications of dynamic programming, the function \( f(p) \) is much more than a set of numbers associated with a set of allowable state vectors. It is a function possessing structure which implies, above all, a "smooth" function of \( p \). In other words, there is a high correlation between values of \( f(p) \) for different values of \( p \).

It follows that in place of merely storing the set of values \( \{ f(p_i) \}, i = 1, 2, \ldots, M \), it is more sensible to think in terms of storing various rules or prescriptions for reproducing the values of \( f(p) \). Let us focus our attention on one of the simplest of these techniques, polynomial approximation. Reference to more sophisticated techniques such as "differential approximation" will be found at the end of the chapter.

To keep the notation under control, consider the case where \( p \) is one-dimensional and constrained to an interval such as \([0, 1]\). Write \( f(p) = f(x) \) where \( 0 \leq x \leq 1 \).

The approach presented above in Section 3 recreates the function \( f(x) \) when needed by means of interpolation, starting with a set of values such as \( f(k/M), k = 0, 1, 2, \ldots, M \). Let us instead ask whether or not we can obtain an approximate representation for \( f(x) \) in the form

\[
 f(x) \cong \sum_{k=0}^{P} a_k x^k. \tag{7.1}
\]

If so, we can store the values of \( f(x) \) in terms of the \( P + 1 \) coefficients \( a_0, a_1, \ldots, a_p \), together with instructions for the evaluation of a polynomial in \( x \) with these coefficients.

If \( f(x) \) is indeed a smooth function, relatively few coefficients will suffice. The simplest way to determine the coefficients is to use mean-square approximation. We demand that the \( a_k \) minimize the quadratic form

\[
 \int_0^1 \left( f(x) - \sum_{k=0}^{P} a_k x^k \right)^2 \, dx. \tag{7.2}
\]
In order to avoid the algebraic and numerical problems associated with this formulation, let us in place of (7.1) use the approximation

$$f(x) \approx \sum_{k=0}^{P} b_k \phi_k(x), \quad (7.3)$$

where the $\phi_k(x)$ are the orthonormal polynomials over $[0, 1].$ * In this case, the requirement of mean-square approximation yields the relations

$$b_k = \int_{0}^{1} f(x) \phi_k(x) \, dx. \quad (7.4)$$

We now face the difficulty of obtaining the $b_k$ numerically. To do this we use numerical quadrature,

$$b_k = \sum_{i=1}^{M} w_i f(x_i) \phi_k(x_i). \quad (7.5)$$

It follows that only the values $\{f(x_i)\}$ need be stored.

Detailed discussions of the determination and evaluation of the $\phi_k(x)$, and of numerical quadrature, will be found at the end of the chapter.

At first sight, we would have appeared to have gained a great deal by reducing the storage of $f(x)$ to the storage of the values $\{f(x_i)\}$. We certainly do save as far as rapid access storage is concerned. But—we have traded time for storage. A great deal of time is consumed in evaluating the coefficients $b_k$ and then in evaluating $f(x)$, using (7.3), at other $x$-values.

This shows up in a particularly serious way when we try to circumvent the curse of dimensionality in this fashion.

Consider, for example, the applicability of these approximation techniques to functions of higher dimension. Suppose that $f(p)$ is a function of four variables, say $f(x_1, x_2, x_3, x_4)$. If then we write

$$f(x_1, x_2, x_3, x_4) \approx \sum_{k,l,m,n=0}^{P} a_{klmn} \phi_k(x_1) \phi_l(x_2) \phi_m(x_3) \phi_n(x_4), \quad (7.6)$$

* To be precise, they are the shifted Legendre polynomials.
where the $a_{kmn}$ are once again determined by the requirement of mean-square fit, we see that $(P + 1)^4$ coefficients are required to store $f(p)$. If $P = 4$, this is $5^4 = 625$; if $P = 9$, this is $10^4 = 10,000$. These figures represent a considerable improvement over the requirements for storage of $f(p)$ at lattice points of a four-dimensional grid.

We can considerably reduce the time required for evaluation of $f$ by using lower degree approximations over subregions. Thus, for example, we can divide the region $0 \leq x_1, x_2, x_3, x_4 \leq 1$ into 256 regions and using an approximation by a quadratic form in each.

Work in this crucial area of approximation has just begun, and there are many unsolved problems that plague us. Our aim in this section has been to indicate that a number of promising techniques exist, and that the reward for ingenuity is high.

8. Numerical Stability

In using the method sketched in Section 2, numerical error arises in the following ways:

\[ g(p, q) \text{ and } T(p, q) \text{ can introduce round-off errors in their calculation.} \]  
\[ (8.1a) \]

\[ f_{n-1}(T(p, q)) \text{ may require an interpolation in its calculation.} \]  
\[ (8.1b) \]

The search for the maximum over $q$ may introduce an error since we allow only a finite set of allowable $q$-values.

\[ (8.1c) \]

It follows that the computational determination of the sequence $\{f_n(p)\}$ generates a new multistage process $\{\phi_n(p)\}$ governed by the relation

\[ \phi_n(p) = \max_q [g(p, q) + \phi_{n-1}(T(p, q))] + u_n(p), \]  
\[ (8.2) \]

where $u_n(p)$ represents the accumulation of the different errors mentioned above.

We hope that by taking sufficiently many $p_i$ and $q_i$ we can keep $|u_n(p)|$ sufficiently small, and that this, in turn, implies
9. Discussion

that \(|f_n(p) - \phi_n(p)|\) is small. If so, our computational process is numerically stable; if not, any results obtained numerically must be regarded with a jaundiced eye. Let us now demonstrate quite simply that we do have stability in the calculation of \(\{f_n(p)\}\).

Let \(q\) be a maximizing value in the original functional equation

\[
f_n(p) = \max_q [g(p, q) + f_{n-1}(T(p, q))],
\]

and \(Q\) a maximizing value in (8.2). Then

\[
\begin{align*}
  f_n(p) &= g(p, q) + f_{n-1}(T(p, q)) \geq g(p, Q) + f_{n-1}(T(p, Q)), \\
  \phi_n(p) &= g(p, Q) + \phi_{n-1}(T(p, Q)) + u_n(p) \\
  &\geq g(p, q) + \phi_{n-1}(T(p, q)) + u_n(p).
\end{align*}
\]

Hence,

\[
\begin{align*}
  f_n(p) - \phi_n(p) &\geq f_{n-1}(T(p, Q)) - \phi_{n-1}(T(p, Q)) - u_n(p) \\
  &\leq f_{n-1}(T(p, q)) - \phi_{n-1}(T(p, q)) - u_n(p),
\end{align*}
\]

and thus,

\[
\begin{align*}
  |f_n(p) - \phi_n(p)| &\leq \max \{|f_{n-1}(p, Q) - \phi_{n-1}(T(p, Q))|, \\
  &|f_{n-1}(T(p, q) - \phi_{n-1}(T(p, q))| + |u_n(p)|.
\end{align*}
\]

Let us now suppose that we have taken care that \(|u_n(p)| \leq \varepsilon\) for \(p \in R\). Then

\[
\max_R |f_n(p) - \phi_n(p)| \leq \max_R |f_{n-1}(p) - \phi_{n-1}(p)| + \varepsilon,
\]

and thus

\[
\max_R |f_n(p) - \phi_n(p)| \leq n\varepsilon.
\]

Linear growth of maximum error is stability in the computer trade.

9. Discussion

In this brief summary of the computational aspects of dynamic programming, we have omitted a number of very interesting and important topics such as reformulation in terms of functions of smaller dimension, Lagrange multipliers, search techniques, stochastic approximation and, generally, any more
than superficial account of approximation methods. References to these and related matters will be found at the end of
the chapter.

Our aim so far in this chapter has been to sketch the basic
techniques involved in numerical solution of the functional
equations of dynamic programming. In the section that fol-
lows, we will enter into some details.

10. Equipment Replacement Policy

Many interesting multistage decision processes of surprising
mathematical depth arise in the control of manufacturing
enterprises. Let us focus our attention on the area of equipment
replacement. Suppose that a certain piece of equipment is
characterized by a purchase price \( p \), and a net annual return
function \( n(t) \), where

\[
n(t) = \text{net return from operating the equipment from}
\]

\[
\text{age } t \text{ to age } t + 1, \, t = 0, 1, 2, \ldots \quad \text{(10.1)}
\]

This function will be a nonincreasing function of \( t \). We suppose
that the equipment is of a specialized nature and thus has no
resale value. At the beginning of each year, a decision is to be
made either to keep, or replace the equipment. Our aim is to
determine the replacement policy which results in the maximum
total return over an \( N \)-year period of operation. More precisely,
we wish to solve the feedback control problem; i.e., we wish
to determine whether to keep or replace equipment that is \( t \)
years old when \( K \) years remain in the process, for all integers
\( K \) and \( t \).

We introduce, as usual, the optimal return function \( f_{K}(t) \),

\[
f_{K}(t) = \text{the net return over a } K\text{-year process}
\]

\[
\text{beginning with equipment on hand that}
\]

\[
is \, t \text{ years old and employing an optimal}
\]

\[
equipment replacement policy, \, K = 1, \, 2, \ldots ; \, t = 0, 1, 2, \ldots \quad \text{(10.2)}
\]

Use of the principle of optimality leads us directly to the
functional equation
10. Equipment Replacement Policy

\[ f_K(t) = \max [n(t) + f_{K-1}(t + 1), -p + n(0) + f_{K-1}(1)], \]
\[ K = 2, 3, \ldots, \quad t = 0, 1, 2, \ldots \quad (10.3) \]

In addition, for the one-stage process we have
\[ f_1(t) = \max [n(t), -p + n(0)]. \quad (10.4) \]

The first term in the bracket on the right-hand side of Eq. (10.3) represents the immediate return from a decision to keep the equipment plus the greatest return from the remainder of the process beginning with a piece of equipment that is one year older. The second term arises from buying a new piece of equipment, operating it for one year, and entering the remainder of the process with a one-year-old piece of equipment.

Now let us specialize the situation further and produce a FORTRAN program which will yield both the optimal policy and returns. We assume that
\[ n(t) = 10 - t, \quad t = 0, 1, 2, \ldots, 10, \quad (10.5) \]
\[ n(t) = 0, \quad t = 11, 12, \ldots, \quad (10.6) \]

Equations (10.3) and (10.4) become
\[ f_1(t) = \max [n(t), 0] = n(t), \quad (10.7) \]
\[ f_K(t) = \max [n(t) + f_{K-1}(t + 1), f_{K-1}(1)]. \quad (10.8) \]

Below we reproduce a listing of a FORTRAN IV program and the print-out which it yielded.

The program is similar to those used in solving much more complex problems. Statements through number 41 are used for certain initial purposes, such as producing the function \( n(t) \) which is stored as the vector XN(J), \( J = 1, 2, \ldots, 12 \). The variable \( K \) is an index which keeps track of the number of stages remaining. The Do-loop beginning “DO 8 NT = 1,11” embraces most of the actual computing. The gains for making the decision “keep” or “replace,” XKEEP and XRPL, with a machine that is NT years old and \( K \) stages remaining, are computed. The larger is selected, its value is noted and called.
III. Computational Aspects

$JOB 
2967, EQUIP, K0160, 1, 20, 20, C
$IBJOB EQUIP MAP
$IBFTC MAIN LIST
C EQUIPMENT REPLACEMENT
DIMENSION F(12, 2), NQ(12), XN(12)
DO 1 J=1, 11
E=J-1
1 XN(J)=10.0-E
XN(12)=0.0
DO 3 J=1, 12
2 F(J, 1)=XN(J)
3 NQ(J)=1
PRINT4, (F(J, 1), NQ(J), J=1, 6)
4 FORMAT(1HO 6(E13.3, 13))
K=1
41 K=K+1
DO 8 NT=1, 11
XKEEP= XN(NT)+F(NT+1, 1)
XRPL=F(2, 1)
IF(XKEEP-XRPL) 7, 6, 5
5 F(NT, 2)=XKEEP
NQ(NT)=1
GOTO 8
6 F(NT, 2)=XKEEP
NQ(NT)=2
GOTO 8
7 F(NT, 2)=XRPL
NQ(NT)=0
8 CONTINUE
F(12, 2)=F(2, 1)
9 NQ(12)=0
PRINT4, (F(I, 2), NQ(I), I=1, 6)
DO 10 I=1, 12
10 F(I, 1)=F(I, 2)
IF(K-25) 11, 12, 12
11 GOTO 41
12 CALLEXIT
13 END
$ENTRY MAIN
$IBSYS ENDJOB
| 0.100E 02 | 0.900E 01 | 0.800E 01 | 0.700E 01 | 0.600E 01 |
| 0.190E 02 | 0.170E 02 | 0.150E 02 | 0.130E 02 | 0.110E 02 |
| 0.270E 02 | 0.240E 02 | 0.210E 02 | 0.180E 02 | 0.170E 02 |
| 0.360E 02 | 0.330E 02 | 0.300E 02 | 0.260E 02 | 0.240E 02 |
| 0.460E 02 | 0.350E 02 | 0.320E 02 | 0.310E 02 | 0.300E 02 |
| 0.550E 02 | 0.410E 02 | 0.390E 02 | 0.370E 02 | 0.360E 02 |
| 0.510E 02 | 0.480E 02 | 0.450E 02 | 0.430E 02 | 0.410E 02 |
| 0.580E 02 | 0.540E 02 | 0.510E 02 | 0.480E 02 | 0.480E 02 |
| 0.660E 02 | 0.600E 02 | 0.560E 02 | 0.550E 02 | 0.540E 02 |
| 0.700E 02 | 0.650E 02 | 0.630E 02 | 0.610E 02 | 0.600E 02 |
| 0.750E 02 | 0.720E 02 | 0.690E 02 | 0.670E 02 | 0.660E 02 |
| 0.820E 02 | 0.780E 02 | 0.750E 02 | 0.730E 02 | 0.720E 02 |
| 0.880E 02 | 0.840E 02 | 0.810E 02 | 0.790E 02 | 0.780E 02 |
| 0.940E 02 | 0.900E 02 | 0.870E 02 | 0.850E 02 | 0.840E 02 |
| 0.100E 03 | 0.960E 02 | 0.930E 02 | 0.910E 02 | 0.900E 02 |
| 0.106E 03 | 0.102E 03 | 0.990E 02 | 0.970E 02 | 0.960E 02 |
| 0.112E 03 | 0.108E 03 | 0.105E 03 | 0.103E 03 | 0.102E 03 |
| 0.118E 03 | 0.114E 03 | 0.111E 03 | 0.109E 03 | 0.108E 03 |
| 0.124E 03 | 0.120E 03 | 0.117E 03 | 0.115E 03 | 0.114E 03 |
| 0.130E 03 | 0.126E 03 | 0.123E 03 | 0.121E 03 | 0.120E 03 |
| 0.136E 03 | 0.132E 03 | 0.129E 03 | 0.127E 03 | 0.126E 03 |
| 0.142E 03 | 0.138E 03 | 0.135E 03 | 0.133E 03 | 0.132E 03 |
| 0.148E 03 | 0.144E 03 | 0.141E 03 | 0.139E 03 | 0.138E 03 |
| 0.154E 03 | 0.150E 03 | 0.147E 03 | 0.145E 03 | 0.144E 03 |
| 0.160E 03 | 0.156E 03 | 0.153E 03 | 0.151E 03 | 0.150E 03 |
F(NT,2), and the decision which produced it is recorded. We set NQ(NT) = 1, if the correct decision is to keep; 0, if the correct decision is to replace; and 2 if the gains from both decisions are equal. The statements

\[ \text{DO 10, I = 1,12} \]
\[ 10 \text{ F(I,1) = F(I,2)} \]

shift the currently computed values of the optimal return function F(I,2) into temporary storage F[I,2] so that the next time we go through the loop, we will be set to calculate the optimal returns from a process that is one year longer in duration. The calculation is set to produce results for processes that are 25 years or less in duration.

The print-out, called for by the "Print" statements gives both the optimal returns and the optimal decisions. The returns are specified in the E-notation. For example 0.100 E 02 is $0.1 \times 10^2 = 10$. The entries in the tenth row show that $f_{10}(0) = 70$, and with a zero year old machine and 10 years remaining the correct decision is "Keep."

**EXERCISES**

1. Refer to the print-out from the problem discussed in the text. Consider a ten-stage process beginning with a five-year-old machine. Determine an optimal replacement procedure.

2. Observe that for the tabulated processes longer than eleven in duration, the increase in the optimal return for each additional year of operation is 6. Comment on this. State a theorem using this observation and establish it via induction.

3. In general computer programs for solving dynamic programming problems involve three nested loops. One refers to the decisions, one to the states, and one to the stages. Rewrite the program given in the text to show this clearly.

4. Write a program for solving another problem discussed in the text.
BIBLIOGRAPHY AND COMMENTS

§1–§6. For detailed discussions of this material, together with extensive applications, see


§7. For further discussions and applications of these techniques, see the book mentioned above, and


For applications to the numerical solution of partial differential equations, see


For an application of dynamic programming to the fitting of functions by polygonal functions, see


———, “A Line Segment Curve-fitting Algorithm Related to
III. Computational Aspects


Polynomial approximation is a particular example of approximation to a function $f(x)$ by means of the solution of a linear differential equation of the form $u^{(N)} + b_1u^{(N-1)} + \cdots + b_Nu = 0$. For a discussion and application of these techniques, see Bellman, R., and R. Kalaba, *Quasilinearization and Nonlinear Boundary-value Problems*, American Elsevier Publishing Company, Inc., New York, 1965.

§8. For a general discussion of stability theory, see


§10. For further discussion, see
Chapter IV

Analytic Results
in Control and Communication Theory

1. Introduction

In the first two chapters, we presented an abstract discussion of both multistage processes and multistage decision processes, and in the third chapter we considered, albeit briefly, the computational feasibility of dynamic programming algorithms. In this chapter we wish to examine in some detail processes arising in control and communication theory of such a structure that both policy and return functions can be obtained in simple and explicit form by means of the functional equation technique of dynamic programming.

We will treat both deterministic and stochastic processes in order to prepare the way for a study of adaptive processes. Having laid the groundwork by means of an analysis of par-
ticular processes, we can then proceed in the succeeding chapter to a formulation of general multistage decision processes of adaptive type.

2. Feedback Control

Let us now particularize the general feedback control process of deterministic type. We assume a linear relation

$$x_{n+1} = Ax_n + y_n, \quad x_0 = c,$$  \hspace{1cm} (2.1)

where $x_n$ and $y_n$ are $N$-dimensional vectors, state, and policy vectors respectively, and $A$ is an $N \times N$ matrix. We shall occasionally call $y_n$ the control vector.

We suppose that the $y_n$ are to be chosen so as to minimize a quadratic criterion function, say

$$R_N(x, y) = \sum_{k=0}^{N} [(x_k, Bx_k) + (y_k, y_k)] + (x_N, x_N).$$  \hspace{1cm} (2.2)

This is not the most general criterion function of this type, but it will serve to illustrate the general techniques we shall employ. If $B = 0$, we have an example of what is commonly called a "terminal control" process.

The results obtained are certainly of intrinsic interest, and possess many immediate applications. Their most important function, however, is to serve as useful starting points for the investigation of more complex processes by way of approximations in both policy and function space.

3. Scalar Case

In order to illustrate the kinds of results obtainable, without the obstructions of notation and calculation, let us begin with the scalar case. Having observed the analytic structure of the solution, we will find it quite easy to derive analogous results for the vector-matrix case.

Let $\{u_n\}$ be a scalar sequence determined by the relation

$$u_{n+1} = au_n + v_n, \quad u_0 = c,$$  \hspace{1cm} (3.1)

and let the $v_n$ be chosen so as to minimize the scalar criterion function
3. Scalar Case

\[ \sum_{n=0}^{N-1} (u_n^2 + v_n^2) + u_N^2. \] (3.2)

Denote the value of this minimum by \( f_N(c) \). Then, in the usual fashion, we obtain the recurrence relation

\[ f_N(c) = \min_v \left[ c^2 + v^2 + f_{N-1}(ac + v) \right], \] (3.3)

for \( N \geq 1 \), with \( f_0(c) = c^2 \).

It is easy to see, inductively or otherwise, that \( f_N(c) \) is a quadratic in \( c \). Since \( f_N(c) \) is clearly even in \( c \) and \( f_N(0) = 0 \), it follows that

\[ f_N(c) = u_N c^2, \] (3.4)

where \( u_N \) is a constant depending upon \( N \), but not \( c \). Substituting in (3.3), we have the relation

\[ u_N c^2 = \min_v \left[ c^2 + v^2 + u_{N-1}(ac + v)^2 \right], \] (3.5)

or

\[ u_N = \min_v \left[ 1 + v^2 + u_{N-1}(a + v)^2 \right]. \] (3.6)

The minimum is determined by the equation

\[ v + u_{N-1}(a + v) = 0, \] (3.7)

whence

\[ v = -\frac{u_{N-1}a}{1 + u_{N-1}}. \] (3.8)

Substituting this relation in (3.6) and simplifying, we obtain the recurrence relation

\[ u_N = 1 + \frac{a^2 u_{N-1}}{(1 + u_{N-1})}, \] (3.9)

with \( u_0 = 1 \). The optimal policy is given by the equation

\[ v_{N-1} = -\frac{au_{N-1}c}{1 + u_{N-1}}. \] (3.10)

We have explicitly obtained the policy and return functions, and thus the complete solution of the original optimization problem. In the exercises we will indicate how one can go further and obtain \( u_N \) and \( v_N \) explicitly as functions of \( N \).
EXERCISES

1. Write, using (3.9),
\[ u_N = (1 + a^2) - \frac{a^2}{(1 + u_{N-1})}, \]
and thus show that the sequence \( \{u_n\} \) is monotone increasing to a limit \( \bar{u} \) determined as the positive solution of
\[ u^2 = 1 + a^2 u. \]

2. Show that \( u_N \) can be determined explicitly in terms of the solution of a second order linear difference equation and thus obtain an analytic expression for \( u_N \).

3. Show that \( u_N - u = O(\lambda^N) \), for some \( \lambda \) with \( 0 < \lambda < 1 \), as \( N \to \infty \), and determine \( \lambda \).

4. Consider the case where the criterion function has the form
\[ \Sigma_{n=0}^{N-1} (u_n^2 + v_n^2) + \lambda(u_N - b)^2 \]
and study the dependence of the solution on \( \lambda \) as \( \lambda \to \infty \).

5. Consider the case where the criterion function has the form
\[ \Sigma_{n=0}^{N-1} v_n^2 + \lambda(u_N - b)^2. \] As \( \lambda \to \infty \), does the optimal control policy approach the optimal control policy for the problem of minimizing \( \Sigma_{n=0}^{N-1} v_n^2 \) under the terminal constraint \( u_N = b \)?

6. Show that \( f_N(c) \) is monotone in \( N \) for fixed \( c \) directly from the definition in (3.3).

7. Taking \( v_0 = -c \), show that \( f_N(c) \leq 2c^2 \), and, hence, using Exercise 6, that \( f_N(c) \) converges as \( N \to \infty \).

4. Discussion

There are several points worth noting about the preceding solution. In the first place, the total cost of the process remains finite as \( N \to \infty \). Secondly, optimal control is linear, which is to say at each stage, \( v \) is directly proportional to \( u \). Thirdly, the constant of proportionality as given by (3.10) approaches a constant, independent of the number of stages remaining, as \( N \to \infty \), and, indeed, quite rapidly. All of these properties carry over to the general multidimensional process formulated in Section 2.
5. Stochastic Version

Let us now consider a stochastic version of the preceding process. In place of (3.1), we write

\[ u_{n+1} = au_n + v_n + r_n, \quad r_0 = c, \quad (5.1) \]

where the \( r_n \) are independent random variables with given distribution functions. The \( v_n \) are to be chosen, using feedback control as discussed in Section 13 of Chapter 2, so as to minimize the expected value of

\[ \sum_{n=0}^{N-1} (u_n^2 + v_n^2) + u_N^2. \quad (5.2) \]

Let \( f_N(c) \) denote the minimum expected value. As before, we obtain the recurrence relation

\[ f_N(c) = \min_v \left[ c^2 + v^2 \right] + \int f_{N-1}(ac + v + r) \, dG(r), \quad (5.3) \]

assuming for the sake of simplicity that \( dG(r) \) is the common distribution function for the random variables. Once again, it is easily seen inductively or otherwise, that each \( f_N(c) \) is quadratic in \( c \). Hence, we may write

\[ f_N(c) = u_N c^2 + v_N c + w_N \quad (5.4) \]

where \( u_N \), \( v_N \), and \( w_N \) are independent of \( c \).

Substituting in (5.3), we have

\[ u_N c^2 + v_N c + w_N = \min_v \left[ (c^2 + v^2) + (u_{N-1}(ac + v + r)^2 \right. \\
\left. + v_{N-1}(ac + v) + w_{N-1} \right] dG(r). \quad (5.5) \]

We can now calculate the minimizing value of \( v \) directly in terms of \( u_{N-1} \) and \( v_{N-1} \) and thus obtain a recurrence relation for the triple \([u_N, v_N, w_N]\) in terms of \([u_{N-1}, v_{N-1}, w_{N-1}]\). We shall not pursue the details here since the calculation is straightforward and best left as an exercise for the reader.
EXERCISES

1. When is $v_N$ identically zero?

2. Obtain the analogs of the results stated in Exercises 2 and 3 at the end of Section 3.

3. Consider the case where the $r_n$ are not independent random variables, beginning with the case where the distribution function for $r_n$ has the form $dG(r_n r_{n-1})$, which is to say, it depends on the value of $r_{n-1}$. What is the asymptotic behavior of the control function in this case?

6. Discussion

Observe that the distribution function for $r_n$ enters only in the form of its first and second moments. Hence, any two distribution functions with the same two moments would lead to the same return function and optimal policy, although the two realizations of the process might look quite different.

It is clear then that the introduction of random variables has not materially affected either the policy function, or the return function. Consequently, stochastic control theory based upon linear equations and quadratic criteria must be viewed with a certain amount of reservation. Superficially, it appears to be more sophisticated than deterministic control theory, but in reality it is not. Only when more effective measures of deviation are employed, together with nonlinear equations, do stochastic features represent a significant and radical departure from the deterministic treatment.

7. Multidimensional Case—Deterministic

Using a modicum of matrix notation, we can now handle the multidimensional version of the deterministic feedback control process described in Section 2. Let $x_n$ and $y_n$ be $M$-dimensional vectors and $A$ an $M \times M$ matrix. Given the transformation relation

$$x_{n+1} = Ax_n + y_n, \quad x_0 = c,$$

(7.1)

where $x_n$ represents the state vector at time $n$ and $y_n$ the control
vector, we wish to choose the $y_n$ so as to minimize the criterion function
\[ R_N = \sum_{k=0}^{N} [(x_k, Bx_k) + (y_k, y_k)] + (x_N, x_N), \quad (7.2) \]
where, as usual, $(x, y)$ denotes the inner product, that is,
\[ (x, y) = \sum_{i=1}^{M} x_i y_i \quad (7.3) \]
if $x_i$ and $y_i$ are respectively the $i$th components of $x$ and $y$.

Setting $f_N(c) = \min_{\{y\}} R_N$, we have the recurrence relation
\[ f_N(c) = \min_{y} [(c, Bc) + (y, y) + f_{N-1}(Ac + y)], \quad (7.4) \]
$N \geq 1$, with $f_0(c) = (c, c)$.

Observing that $f_N(c)$ is itself a quadratic form in $c$, a result obtained either intuitively, inductively, or directly from the original variational problem, we write $f_N(c) = (c, Q_N c)$. Substituting in (7.4), we obtain the relation
\[ (c, Q_N c) = \min_{y} [(c, Bc) + (y, y) \]
\[ + (Ac + y, Q_{N-1}(Ac + y))], \quad (7.5) \]
with $Q_0 = I$.

The usual variational approach yields the linear control law
\[ y(I + Q_{N-1}) + Q_{N-1}Ac = 0, \quad (7.6) \]
or
\[ y = -(I + Q_{N-1})^{-1}Q_{N-1}Ac. \quad (7.7) \]

The value at the minimum is easily calculated, noting that
\[ (y, y) + (Ac + y, Q_{N-1}(Ac + y)) = (Ac, Q_{N-1}Ac) \]
\[ + (y, y + Q_{N-1}y + Q_{N-1}Ac) + (y, Q_{N-1}Ac). \quad (7.8) \]
By virtue of (7.6) and (7.7) this reduces to
\[ = (Ac, Q_{N-1}Ac) - ((I + Q_{N-1})^{-1}Q_{N-1}Ac, Q_{N-1}Ac) \]
\[ = (Ac, Q_{N-1}Ac) - ((I + Q_{N-1})^{-1}(I + Q_{N-1}Ac, Q_{N-1}Ac)) \]
\[ + ((I + Q_{N-1})^{-1}Ac, Q_{N-1}Ac) \]
\[ = ((I + Q_{N-1})^{-1}Ac, Q_{N-1}Ac). \]
Hence, "equating" coefficients, we obtain from (7.5)
\[ Q_N = B + A'Q_{N-1}(I + Q_{N-1})^{-1}A, \]  
with \( Q_0 = I \). This permits the recursive calculation of the \( Q_N \).
Writing (7.9) in the form
\begin{align*}
Q_N &= B + A'(I + Q_{N-1})(I + Q_{N-1})^{-1}A - A'(I + Q_{N-1})^{-1}A \\
&= B + A'A - A'(I + Q_{N-1})^{-1}A,  
\end{align*}
(7.10)
we see that \( Q_N \) is uniformly bounded and monotonically increasing in the sense that \( Q_N - Q_{N-1} \) is positive definite.

**EXERCISES**

1. Show that if \( \{ Q_N \} \) is a bounded sequence of positive definite matrices with the property that \( Q_N - Q_{N-1} \) is a nonnegative definite, then \( Q_N \) converges as \( N \to \infty \).

2. Show that \( Q_N \) satisfying (7.10) can be obtained by solving a linear difference system of order \( 2M \).

3. Consider the stochastic process generated by
\[ x_{n+1} = Ax_n + y_n + r_n, \quad x_0 = c, \]
where the \( r_n \) are independent random vectors. Obtain the analogs of the foregoing results.

**8. The Continuous Case**

The corresponding results for continuous control processes are also of some elegance. Consider the problem of minimizing the functional
\[ J(y) = \int_0^T [(x, Bx) + (y, y)] \, dt \]  
(8.1)
over all \( y \) where the vectors \( x \) and \( y \) are connected by the relation
\[ \frac{dx}{dt} = Ax + y, \quad x(0) = c. \]  
(8.2)
Writing \( f(c, T) = \min_y J(y) \), we have, arguing as before,
\[ f(c, T) = \min_v \left[ [(c, Bc) + (v, v)]\Delta \right. \\
\left. + f(c + (Ac + v)\Delta, T - \Delta) + o(\Delta) \right] \]  
(8.3)
and thus in the limit as $\Delta \to 0$, the partial differential equation
\[
\frac{\partial f}{\partial T} = (c, Bc) + \min_v [(v, v) + (Ac + v, \text{grad} f)],
\] (8.4)
where, as usual,
\[
\text{grad } f = \begin{pmatrix}
\frac{\partial f}{\partial c_1} \\
\vdots \\
\frac{\partial f}{\partial c_M}
\end{pmatrix},
\] (8.5)
with $c_1, c_2, \ldots, c_M$ the components of $c$. The optimal control is given by
\[
v = -\frac{\text{grad } f}{2}.
\] (8.6)
Writing $f(c, T) = (c, Q(T)c)$, we have $\text{grad } f = 2Q(T)c$. Hence, the optimal control is linear,
\[
v = -Q(T)c,
\] (8.7)
and (8.4) yields the Riccati equation
\[
Q'(T) = \frac{B - Q^2(T)}{4} + \frac{A'Q + QA}{2},
\] (8.8)
with $Q(O) = 0$.

**EXERCISES**

1. Show that (8.8) can be reduced to a linear equation by means of the substitution $Q = YZ^{-1}$, where $Y$ and $Z$ satisfy linear differential equations.
2. Show in two ways that $Q(T)$ approaches a limit as $T \to \infty$.
3. The preceding techniques can be used to study the variation of Green's functions of linear functional equations. Let us begin with the second order linear differential equation
\[
u'' + q(x)u = v(x), \quad a < x < 1, \quad u(a) = u(1) = 0,
\]
and leave the steps as a series of exercises for the reader. Show that the solution of the foregoing equation may be written in the form
IV. Analytic Results in Control and Communication Theory

\[ u(x) = \int_{a}^{1} k(x, y, a)v(y) \, dy, \]

where \( k \) is constructed by means of solutions of the homogeneous equation. The kernel \( k(x, y, a) \) is called the Green’s function associated with the particular boundary conditions.

4. Determine \( k(x, y, a) \) for \( q(x) \) constant.

5. Show that the Euler equation associated with the problem of minimizing

\[ J(u) = \int_{a}^{1} [q(x)u'^2 - u'^2 - 2vu] \, dx \]

is the second order linear equation \( u'' - q(x)u = v \), with end conditions determined by the conditions on \( u \).

6. If \( \min_{a} J(u) = f(a, c) \), where \( u(a) = c \), show that

\[ \frac{\partial f}{\partial a} = q(a)c^2 - 2v(a)c + \frac{\partial f/\partial c}{4}. \]

7. Show, using the Euler equation, that

\[ f(a, c) = \int_{a}^{1} \int_{a}^{1} K(x, y, a)v(x)v(y) \, dx \, dy + 2cL_\lambda(v) + c^2g(a), \]

where \( L_\lambda(v) \) is a linear functional of \( v \).

8. Substituting in the equation of Exercise 6, obtain the equation

\[ \frac{\partial K}{\partial a}(x, y, a) = \frac{\partial K}{\partial y}(x, a, a) \frac{\partial K}{\partial x}(a, y, a). \]

9. Use the equation \( u'' + (q(x) + \lambda r(x))u = v(x) \) to obtain an analogous formula for the resolvent function and thus formulas for the dependence of the characteristic values and functions on the endpoint \( a \).


10. Consider the Fredholm integral equation

\[ u(x) + v(x) + \int_{a}^{T} k(x, y)u(y) \, dy = 0, \quad 0 \leq a \leq T, \]

with the unique solution

\[ u(x) = -v(x) + \int_{a}^{T} Q(x, y, a)v(y) \, dy. \]

Using the dynamic programming techniques to minimize the quadratic functional
8. The Continuous Case

\[ J(u) = \int_a^T u^2(x) \, dx + 2 \int_a^T u(x)v(x) \, dx + \int_a^T \int_a^T K(x, y)u(x)u(y) \, dx \, dy \]

and derive the equation

\[ \frac{\partial Q}{\partial a}(x, y, a) = Q(a, x, a)Q(a, y, a). \]

(See

Further applications of this technique may be found in

and in Chapter 9 of

11. Consider the problem of minimizing the quadratic functional

\[ J(u) = \int_0^T (u^2 + u'^2) \, dt, \]

under the conditions \( u(0) = c, \ |u'| \leq m \) for \( 0 \leq t \leq T \). Write \( f(c, T) = \min_u J(u) \) and show that

\[ f_T = \min_{|v| \leq m} [c^2 + v^2 + v_T]. \]

12. Obtain the analytic form of \( f(c, T) \) and the form of the optimal policy by determining \( f(c, T) \) first for small \( T \), and then increasing the value of \( T \) until the constraint is met.

(For a systematic approach to variational problems with constraints by means of the Neyman-Pearson lemma, see the monograph

For an approach using classical variational techniques and extensions, see

IV. Analytic Results in Control and Communication Theory

13. Consider the problem of minimizing the functional

\[ J(u) = \int_0^T \left( u'^2 + g(u) \right) dt, \quad u(0) = c. \]

Let \( u_n \) be an \( n \)th approximation, and let \( u_{n+1} \) be determined as the function minimizing

\[ J_n(u) = \int_0^T \left( u'^2 + g(u_n) + (u - u_n)g'(u_n) + \frac{(u - u_n)^2}{2} g''(u_n) \right) dt, \]

with \( u(0) = c. \) Under what conditions does the sequence \( \{u_n\} \) converge to the function \( u(t) \) minimizing \( J(u) \)?


9. Prediction Theory

A fundamental objective of science is the prediction of the future on the basis of past and present knowledge. Since the classical methods based upon differential equations usually require more knowledge than is available, different approaches have been developed over the past years based upon probabilistic considerations.

Let us consider, in particular, the problem of obtaining an estimate for \( a_n \) for \( n > 0 \) given the sequence \( \{a_n\} \) for \( n = 0, -1, -2, \ldots \). More generally, given two sequences \( \{a_k\} \) and \( \{b_k\} \), how do we obtain the best estimation for the \( b_k \) given only the \( a_k \) and various correlations?

One way of approaching this problem, which is both analytically tractable and intuitive from the electrical engineering point of view, is to ask for a third sequence \( \{u_k\} \) with the property that

\[ b_k \approx \sum_{l=0}^{M} u_k a_{k-l}. \tag{9.1} \]

Following Kolmogorov and Wiener, it is reasonable to ask that the \( u_l \) be determined so as to minimize the quadratic expression

\[ \sum_{k=0}^{N} \left( b_k - \sum_{l=0}^{M} u_l a_{k-l} \right)^2. \tag{9.2} \]
Limiting forms of this problem are more easily handled; either

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N} \left( b_k - \sum_{l=0}^{N} u_i a_{k-l} \right)^2$$

(9.3)

or

$$\int_0^T \left( b(t) - \int_0^T a(t - s) u(s) \, ds \right)^2 \, dt.$$  

(9.4)

As might be suspected from the preceding discussions, variational problems of this genre can be readily and systematically discussed by means of the functional equation technique of dynamic programming applied in the preceding sections. The interested reader is referred to a number of references to be found at the end of the chapter.

10. Communication of Information

Let us turn to another interesting area of modern research, that centering about the communication of information. We can decompose the process into five parts:

(a) The origination of data.
(b) The encoding of the data for transmission.
(c) The communication of the encoded data.
(d) The decoding of the signal received, after transmission.
(e) The use of the decoded signal.

The process can be considerably complicated by introducing feedback loops for purposes of verification.

In the past, and well into the present, there has been an unfortunate, but understandable, tendency to study these parts individually with little or no acknowledgment of the existence of the other parts. At best, this "splendid isolation" yields incomplete results; at worst, it produces thoroughly misleading results. This is particularly the case when results, such as those pertaining to idealized channel capacity, have been taken out of context and applied willy-nilly with no understanding of the basic processes or objectives.

It is essential in the treatment of these complex problems to take account of the interaction of the different parts of the communication system described above. To illustrate an
approach which can be used, we shall discuss a problem originally posed and partially resolved by Kelly. Applying dynamic programming techniques, we can resolve the original problem and extend it in a number of ways. To simplify the presentation, we shall ignore the encoding and decoding aspects. The essential point in that communication of information is for the purposes of decision-making.

11. The Efficient Gambler

Suppose that an unscrupulous gambler receives prior information concerning the outcomes of sporting events over a telephone with a great deal of background noise. Since time is short, he receives merely the words "win" or "lose" and proceeds to act on this information. Because of the difficulties of communication, there is a probability that he hears the wrong word. With this in mind, and knowing the probability of hearing an incorrect word, how much should he wager on each event?

It is clear that his betting policy must depend strongly upon the criterion function he uses measuring the utility of money. If, for example, he wishes to maximize the expected value of his winnings, then he wagers his total fortune whenever he thinks that the probability of having correct information is greater than \( \frac{1}{2} \). This is clearly a very risky policy, since there is an appreciable probability of total ruin over a sequence of wagers.

Consequently, it seems wiser to use a hedging criterion, one that prevents total ruin, no matter what happens. We can ensure a more cautious attitude by adopting a different criterion. For example, the gambler may decide to bet so as to maximize the expected value of the logarithm of the total amount of money he possesses at the end of his wagering.

To see how that affects the betting policy, suppose that the probability that the information is correct is \( p \); that he possesses \( x \) dollars initially and that he wants to make a wager \( y \) which will maximize the expected value of the logarithm of the total
amount of money he has after the wager has been paid off one way or the other.

This expected value has the form

$$E(y) = p \log (x + y) + (1 - p) \log (x - y), \quad (11.1)$$

as we see upon enumerating outcomes. A simple application of calculus shows that the maximum is attained at $y = (2p - 1)x$, if $p > \frac{1}{2}$, and at $y = 0$ otherwise. Let us consider, then, the interesting case where $p > \frac{1}{2}$.

The maximum value of $E(y)$ is given by

$$E_{\max} = \log x + \log 2 + p \log p + q \log q. \quad (11.2)$$

The first two terms represent what the expected value would have been with perfect information, $p = 1$. The second two terms, which are negative, represent the cost of misinformation and thus can be used to denote the effect of noisy transmission.

Let us next consider a multistage betting process in which the gambler receives information, makes a wager, wins or loses, receives a second signal, and so on. What does his optimal policy look like now?

12. Dynamic Programming Formulation

Proceeding in the usual fashion, let

$$f_N(x) = \text{the expected value of the logarithm of the final capital obtained from an N-stage wagering process, starting with the final amount } x \text{ and using an optimal policy.} \quad (12.1)$$

The principle of optimality yields the recurrence relation

$$f_N(x) = \max_{0 \leq y \leq x} [pf_{N-1}(x + y) + (1 - p)f_{N-1}(x - y)], \quad (12.2)$$

$N \geq 2$, with, as calculated above,

$$f_1(x) = \log x + K, \quad (12.3)$$

where

$$K = \begin{cases} 
\log 2 + p \log p + (1 - p) \log (1 - p), & p > \frac{1}{2}, \\
0, & p \leq \frac{1}{2}. \end{cases} \quad (12.4)$$
The recurrence relation of (12.2) remains the same regardless of the hedging function we use. The initial function \( f_1(x) \) determines the structure of the optimal policy. The rather surprising property of the logarithm is that it allows us to obtain a simple explicit solution of (12.2), starting with the initial value of (12.3).

Let us show inductively that

\[
f_N(x) = \log x + NK,
\]

for \( N \geq 1 \), and that the optimal policy is independent of \( N \), and has precisely the form given above, namely,

\[
y = (2p - 1)x, \quad p > \frac{1}{2}, \\
y = 0, \quad p \leq \frac{1}{2}.
\]

Assume that the result holds for \( N \). Then

\[
f_{N+1}(x) = \max_{0 \leq y \leq x} \left[ p(\log(x + y) + NK) \\
+ (1 - p)[\log(x - y) + NK] \right]
\]

\[
= NK + \max_{0 \leq y \leq x} \left[ p \log(x + y) + (1 - p)\log(x - y) \right]
\]

\[
= NK + \log x + K = \log x + (N + 1)K.
\]

This establishes both assertions.

**EXERCISES**

1. Consider the case in which the transmission properties change with time, so that at the \( k \)th stage the probability of correct transmission is \( p_k \), with \( p_k > \frac{1}{2} \). Show that the maximum expected value of the logarithm of the capital at the end of \( N \) stages is given by the expression

\[
\log x + N \log 2 + \sum_{k=1}^{N} (p_k \log p_k + (1 - p_k) \log(1 - p_k)).
\]

2. Consider the case where the probability of correct transmission of the \( k \)th signal depends upon whether or not the \((k - 1)\)st signal was transmitted correctly.

3. Consider the situation where \( M \) different signals can be transmitted at any stage and the gambler possesses the following data:
13. Discussion

\( p_{ij} \) = the conditional probability that the \( j \)-signal was emitted at
the source if the \( i \)-signal is received by the observer.

\( q_i \) = the probability that the observer at any stage will receive
an \( i \)-signal.

\( r_i \) = the return from a winning bet of one unit on signal \( j \).

The gambler receives a \( j \)-signal and bets a quantity \( z_i, \Sigma_{i=1}^{M} z_i \leq x \)
that the original signal was actually \( i \).

4. Consider the case of a continuum of signals, \(-\infty < v < \infty\), where we
are given the conditional probability \( dG(u, v) \) that a signal with label
between \( v \) and \( v + dv \) was sent if \( u \) was received, and the other relevant
quantities.

(For these and other results, see
R. Bellman and R. Kalaba, "Dynamic Programming and Statistical
pp. 749–751.

———, "On Communication Processes Involving Learning and
Random Duration," 1958 IRE National Convention Record, Part 4,

———, "Dynamic Programming and Adaptive Processes: Mathe-
matical Foundations," IRE Trans. on Automatic Control, Vol. AC-5,
1960, pp. 5–10.


5. What does the gambler do if \( p < \frac{1}{2} \)?

13. Discussion

An immediate important observation is that the use of the
criterion function \( \log x_N \) permits us to separate the effect of
initial resources from the effect of noisy transmission. Under
these conditions we can forget about the use of the signals and
examine the communication channel properties by itself. It also
turns out that this very function \( p \log p + (1 - p) \log (1 - p) \)
arises in the study of some specialized coding problems as a
result of an important structural property and in statistical
mechanics in the form of entropy.

A few interesting applications of the entropy function have
been made to the study of the transmission of information,
and a far greater number of ill-advised and misleading appli-
cations. The foregoing analysis shows how closely the entropy
function is connected with the criterion function \( \log x_N \). Were
we to use a different criterion and to add other realistic features, we would obtain an entirely different measure of the noisy properties of the communication channel, and as a matter of fact, we could not separate the influence of the initial resources from the influence of the communication channel.

BIBLIOGRAPHY AND COMMENTS

§1—§3. See Chapter X in


§4. See


§5. For stochastic control processes, see


Bibliography and Comments


§7. See


For a discussion of time-lag processes, see


§8. Results of this type were first given in


§9. See the Appendix by N. Levinson in

N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series and Engineering Applications*, John Wiley and Sons, New York, 1942; see also

Chapter Nine of “Matrix Analysis,” page 155, and the paper by Kalman referred to above. Kalman has systematically applied the theory of dynamic programming to engineering design and control processes and obtained very detailed results. For some of his later results, see


The approach given here in Section 12 was given in the papers cited at the end of Section 12. For a further discussion of these matters, see J. Marshak, *Remarks on the Economics of Information*, Cowles Foundation Discussion Paper No. 70, April 1959.

A great deal of attention has been focused on control processes governed by linear differential equations of the form \[ x' = Ax + y, \]
where the components of \( y \) satisfy the constraints \( |y_i| \leq m_i, \]
\[ 0 \leq t \leq T, \]
and are to be chosen so as to minimize the time required to drive \( x \) to the origin. The mathematical depth of these questions is quite surprising and a number of powerful analytic tools have been applied to their study. See R. Bellman, I. Glicksberg, and O. Gross, "On the 'Bang-bang' Control Problem," *Q. Appl. Math.*, Vol. 14, 1956, pp. 11–18.


Chapter V

Adaptive Control Processes

1. Introduction

Let us now proceed cautiously to grapple with some of the more realistic, and certainly more complex and intriguing, versions of uncertainty and their associated control processes. The processes we shall discuss are quite new mathematically. They are, however, new scientifically only in the sense that the majority of mathematicians, engineers, and scientists have steadfastly until very recently refused even to admit their existence. Nonetheless, they have always been uninvited guests at the scientific banquet table.

In our study of deterministic control processes, we made the traditional assumptions that state variables are identifiable and observable, that the set of possible decisions is known, that cause and effect relations are known, and that the objectives
of the decision process under study are well defined. The mathematical results obtained in this fashion are approximations of greater or less validity. As far as applications are concerned, it is a matter of "caveat emptor," with no scientific Food and Drug Act requiring the strict labeling of the contents of all panaceas.

The first step away from completely deterministic concepts, and a step of considerable import, is the classical theory of probability with its introduction of random variables. In this chapter, we want to indicate the existence of higher levels of uncertainty and to show that the functional equation technique of dynamic programming can be used to obtain analytic formulations, even in these more complicated situations.

In order to orient the reader in these new directions with as little motion sickness as possible, we shall treat in some detail some relatively simple adaptive processes, analogs of those already considered in a deterministic and stochastic setting. Following this, we shall re-examine our techniques and indicate approaches to general adaptive control processes. The corresponding multistage decision processes lead to difficult analyses and a major effort is required to employ a computer to obtain numerical solutions. This area contains many problems worthy of research.

2. Linear Control Process of Adaptive Type

In Chapter 4, we considered the problem of minimizing the criterion function

$$R_N = \sum_{n=0}^{N} (u_n^2 + v_n^2)$$

(2.1)

over all \(v_n\), where

$$u_{n+1} = au_n + v_n, \quad u_0 = c.$$  \hspace{1cm}  (2.2)

Functional equation techniques were employed to obtain the policy function and return function.

We then continued to the problem of minimizing the expected value of \(R_N\), employing feedback control, where \(u_n\)
and \( v_n \) are now stochastic variables connected by the relation

\[
    u_{n+1} = au_n + v_n + r_n, \quad u_0 = c,
\]

(2.3)

with the \( r_n \) random variables. The easiest case is that where the \( r_n \) are independent random variables with a common known distribution function, but we can also treat, without undue difficulty, the case where there are "nearest-neighbor" correlations; i.e., the distribution of \( r_n \) depends upon the values of the previous \( r_{n-1}, r_{n-2}, \ldots, r_{n-k} \).

Let us now venture further into the unknown by assuming that the distribution function is known to exist, but its precise form is unknown. How do we proceed from here?

Since these are processes which we have not discussed previously, we must once again make precise what we mean by an optimal policy, and, indeed, what we mean by an expected value. Our aim is to provide a uniform analytical formulation of the control processes arising in situations of this nature. Observe that we said "... a uniform formulation ...," not "... the uniform formulation ...." As we shall indicate subsequently, there are many paths we can take, each with certain attractive and unattractive features. It is perhaps a little distracting for the student to have all of this freedom, but this is a part of the price of venturing into the unknown.

3. Analytic Formulation

We begin, as usual, by simplifying. In place of a distribution function of unknown form, we shall suppose that the analytic structure is known, but that certain parameters are unknown. Thus, for example, we may use a Gaussian distribution function with unknown mean and variance, or a Poisson distribution with unknown mean, or a general Pearson distribution with unknown parameters, and so on.

Let us consider first perhaps the simplest case where \( r \) has a binomial distribution. The variable \( r \) assumes the value +1 with probability \( p \) and -1 with probability \( (1 - p) \). As pointed out in the preceding paragraph, however, we do not possess any knowledge concerning the numerical value of \( p \), apart from
the obvious fact that $0 < p < 1$. We will, however, estimate $p$ at each stage on the basis of the set of values of $r$ which have occurred over the previous stages. This is a reasonable procedure.

In processes of this type, we face the problem of learning and performing simultaneously. In place of the word "learning" which has overdetermined psychological implications, we shall use the term "adaptation," which is sufficiently vague at the present to be useful. Hence, processes of this nature will be called *adaptive control processes*.

Simplifying once again, we shall suppose that it is sufficient merely to record the number of $+1$'s and the number of $-1$'s which have occurred in the past, or, equivalently, reverting to coin-tossing, the number of heads and the number of tails. What is here a mathematical assumption may often be a realistic constraint in the sense that frequently in practice only part of the over-all information is available to the decision maker, or, it may be too expensive or time-consuming to collect more extensive data. On the other hand, it may turn out that no further information concerning the process is necessary. We shall come back to this point of "sufficient statistics" below.

Let us agree to estimate $p$ by means of the formula

$$p_{mn} = \frac{m + 1}{m + n + 2}$$  \hspace{1cm} (3.1)

if $m + 1$'s and $n - 1$'s have been observed. Again, the question of *why* this particular estimation procedure has been used, must and will be discussed.

We turn next to the calculation of expected values. Our fundamental assumption is that we compute averages, expected values, as if the estimated probabilities were the actual probabilities in a stochastic control process of the type we have already considered. This stochastic control process is the one described in (2.3) with a *known* distribution function for $r$. An optimal policy is now defined as a policy which minimizes the expected value of the criterion function.
Finally, let us observe that the concept of state variable has been materially broadened. In addition to the current physical state, we must also include some of the past history of the process. One of the advantages of the dynamic programming approach is that it automatically extends to these more general situations. We shall dwell on this idea at some length in subsequent sections.

With these preliminaries disposed of, let us turn to an analytic formulation. As before, we take our criterion function to have the simple analytic form

\[ R_N = \sum_{n=0}^{N-1} (u_n^2 + v_n^2) + u_N^2. \]  

(3.2)

Introduce the function

\[ f_N(c, m, n) = \text{the expected value of } R_N \text{ obtained using} \]

an optimal policy, starting in state \( c \) with \( m + 1 \)'s and \( n - 1 \)'s observed.  

(3.3)

The principle of optimality yields the functional equation

\[ f_N(c, m, n) = \min_v \left[ (c^2 + v^2) + p_{mn}f_{N-1}(ac + v + 1, m + 1, n) \right. \]

\[ + \left. (1 - p_{mn})f_{N-1}(ac + v - 1, m, n + 1) \right], \]  

(3.4)

for \( N \geq 1 \), with \( f_0(c, m, n) = c^2 \), and \( p_{mn} \) calculated as in (3.1).

We have thus reduced the study of an adaptive control process to the analytic and computational solution of a functional equation.

4. Analytic Aspects

It is easily seen, as before, that

\[ f_N(c, m, n) = u_N(m, n)c^2 + v_N(m, n)c + w_N(m, n) \]  

(4.1)

where \( u_N, v_N, \) and \( w_N \) are independent of \( c \).

Substituting, we have the equation

\[ u_N(m, n)c^2 + v_N(m, n)c + w_N(m, n) \]

\[ = \min_v \left[ (c^2 + v^2) + p_{mn}(u_{N-1}(m + 1, n)c^2 \right. \]

\[ + v_{N-1}(m + 1, n)c + w_{N-1}(m + 1, n)) \]

\[ + (1 - p_{mn})(u_{N-1}(m, n + 1)c^2 + v_{N-1}(m, n + 1)c \]

\[ + w_{N-1}(m, n + 1)) \right]. \]  

(4.2)
From this, we see once again that the optimal control is linear, and we can derive recurrence relations for the triple $[u_N(m, n)/v_N(m, n), w_N(m, n)]$. We leave this to the reader to carry out as an exercise. Let us in fairness warn him, however, that it is more than an exercise to determine the asymptotic behavior of these quantities as $N \to \infty$.

5. Discussion

Having made a number of assumptions in the formulation, we wish to institute some checks, some controls, in our mathematical experiments. For example, we would like to demonstrate that with probability one the foregoing process converges to the stochastic control process with known $p$. By convergence, we would mean that the optimal policy of one process converges to the optimal policy of the other, as well as convergence of the return functions. Although results of this nature are intuitively clear, no rigorous justification has been given as yet, and there is ample room for research in this area. As the reader will see, it is not a simple matter to use the recurrence relations obtained from (4.2) for either analytic or computational purposes.

6. A Priori Distribution Function

Let us now consider the apparently more difficult case where we possess an a priori distribution function for $p$, say $dG(p)$, a probability for a probability.

How shall we modify this distribution function on the basis of observation of the process? We agree, once again, subject to debate and discussion, to use a Bayes technique. If a $+1$ is observed for $r$, then we replace $dG$ by the new distribution function

$$dG_+ = \frac{p \, dG(p)}{\int_0^1 p \, dG}, \quad (6.1)$$

while if a $-1$ is observed, we use the new distribution function

$$dG_- = \frac{(1 - p) \, dG(p)}{\int_0^1 (1 - p) \, dG}. \quad (6.2)$$
What is interesting is that once again it is sufficient merely to keep track of the number of plus and minus values of \( r \). It is not necessary specifically to indicate the dependence upon \( dG(r) \). If we take \( dG(r) = dr \), the functional equation is precisely that of (3.3), and the rule for calculating \( p_{mn} \) reduces to that given in Section 3, (3.1).

There are various ways of justifying the transformations of (6.1) and (6.2) based upon the theory of games and statistical decision. However, these justifications require still further justification, and so on. We may just as well introduce our assumptions in a convenient form.

7. The Adaptive Gambler

Let us now return to the problem of a gambler using an imperfect communication channel, the problem to which we devoted some space in Chapter 4. Suppose now that he does not even know the probability that the signal will be transmitted incorrectly, but he does know that the probability exists and is fixed over time. Let us formulate the problem of optimal wagering in the foregoing terms.

As before, we can postulate an a priori distribution function for \( p \), \( dG(p) \), and agree, consistent with the formulas in Section 6, to transform \( dG(p) \) into

\[
dG_{m,n}(p) = \frac{p^m(1-p)^n}{\int_0^1 p^m(1-p)^n \, dG(p)}
\]  

(7.1)

after \( m \) successful wagers and \( n \) unsuccessful wagers have been observed. If we define

\[
f_N(x, m, n) = \text{the expected value of } \log x_N, \text{ starting}
\]

with a quantity \( x \) and the information that there have been \( m \) successes and \( n \) failures to date, using an optimal policy,  

(7.2)

then we obtain the equation

\[
f_N(x, m, n) = \max_{0 \leq y \leq x} \left[ p_{mn} f_{N-1}(x + y, m + 1, n) + (1 - p_{mn}) f_{N-1}(x - y, m, n + 1) \right],
\]

(7.3)

with
V. Adaptive Control Processes

\[ f_l(x, m, n) = \max_{0 \leq y \leq x} \left[ p_{mn} \log (x + y) + (1 - p_{mn}) \log (x y) \right] \]  

(7.4)

It is easy, as before, to show inductively that

\[ f^*_N(x, m, n) = \log x + c_N(m, n), \]  

(7.5)

where \( c_N(m, n) \) satisfies a linear recurrence relation of rather complicated form.

Nevertheless, the optimal policy still has the simple intuitive form

\[ y = (2p_{mn} - 1)x, \quad \text{if} \quad p_{mn} > \frac{1}{2}, \]  
\[ = 0, \quad \text{otherwise}, \]  

(7.6)

where \( p_{mn} = \int_0^1 p \, dG_{mn} \). Thus, the structure of the optimal policy has been left unchanged even with the addition of greater uncertainty.

EXERCISES

1. If \( dG(p) = p^{a-1}(1 - p)^{b-1} \, dp/B(a, b) \), where \( B(a, b) \) is the Beta function, show that

\[ p_{mn} = \frac{m + a}{m + a + n + b}, \]

and that after \( m \) wins and \( n \) losses one wagers the fraction \((m + a) - (n + b))/(m + a) + (n + b) \) of one's capital.

2. If the criterion function is \( X_N a^2, \ a > 0 \), what are the return functions and the optimal policy?

8. The Two-Armed Bandit Problem

Continuing in this vein, let us now discuss a problem which arose in the work of W. R. Thompson on the testing of new drugs. It has since acquired the picturesque name given above, by virtue of the fact that in place of drugs, we consider slot machines.*

Suppose that we have two slot machines in front of us, one with known properties and one with unknown properties.

* A consequence of the proximity of Los Angeles to Las Vegas rather than Rochester, Minnesota.
When the handle on the first machine is pulled, there is a known probability $s$ of receiving a dollar; when the second machine is played, there is a fixed, but unknown, probability of success, $p$.

The process assumes the following form. We try the second machine a number of times to be determined by the outcomes, and then decide to use the first machine from then on. It is intuitively clear, and can be easily demonstrated, that having once decided to use the machine with known characteristics, we never return to the machine with unknown properties. However, we may never use the first machine.

To simplify the analysis, and to illustrate the use of a criterion function we have not so far employed, let us suppose that we wish to maximize the expected value of

$$R = \sum_{n=0}^{\infty} a^n z_n,$$

where $0 < a < 1$ and $z_n$ represents the return obtained on the $n$th trial. If, as before,

$$f_{mn} = \text{the expected return obtained using an optimal policy after } m \text{ successes and } n \text{ failures have been observed using the second machine},$$

and $dG(p)$ represents an a priori distribution function for $p$, we readily obtain the functional equation

$$f_{mn} = \max \left[ \frac{(1 + af_{m+1,n}) p_{mn} + a(1 - p_{mn}) f_{m,n+1}}{s/(1 - a)} \right],$$

where, once again,

$$p_{mn} = \frac{\int_0^1 p^{m+1}(1 - p)^n dG(p)}{\int_0^1 p^m(1 - p)^n dG(p)}.$$

The relation of (8.3) can be used for a further study of the structural properties of the optimal policy, or used to obtain the computational solution. Let us warn the reader that the determination of the structure of the optimal policy is not easy.
9. A General Formulation

Now that we have discussed three particular cases, in Section 2–8, we are ready to present a general formulation of the type of adaptive control process we have been considering. We suppose that

\[ p_1 = T(p, q, r), \]  

(9.1)

where \( p \) as usual is the state variable, \( q \) is the decision variable, and \( r \) is a random variable with a fixed, but unknown, distribution function. Let \( dG(r) \) be an a priori estimate for this unknown distribution function. The physical state is \( p \), but the state of the control process is the couple \((p, G)\).

Our aim is to use a control policy which will maximize the expected value of the criterion function

\[ R_N = g(p, q_1, r_1) + g(p_1, q_2, r_2) + \cdots + g(p_{N-1}, q_N, r_N). \]  

(9.2)

In order to make this problem precise, let us make explicit our assumptions:

We can observe the state of the system at each stage.  \hspace{1cm} (9.3a)  
At each stage, we regard the a priori estimate as the actual distribution function. Expected values are obtained on this basis. \hspace{1cm} (9.3b)  
We have a systematic procedure for modifying the a priori distribution function as the process unfolds; this procedure itself may be an adaptive one. \hspace{1cm} (9.3c)

As a result of the choice of \( q_1 \), we then have the transformations

\[ p_1 = T(p, q_1, r_1), \]

\[ dG_1(r) = S(r; p, q_1, r_1, G). \]  

(9.4)

In words, the new distribution function depends upon the old distribution function, the value of \( r_1 \), if observed, the original state \( p \), the new state \( p_1 \) and the decision \( q_1 \).

Let

\[ f_N(p, G) = \text{the expected value of } R_N \text{ obtained using an optimal policy, starting in the state } (p, G). \]  

(9.5)
Then the principle of optimality yields the functional equation

\[ f_N(p, G) = \max_{q_1} \left[ \int [g(p, q_1, r_1) + f_{N-1}(T(p, q_1, r_1), G_1)] dG(r_1) \right], \quad (9.6) \]

where \( G_1 \) is as in (9.4), for \( N \geq 2 \), and for \( N = 1 \), we have

\[ f_1(p, G) = \max_{q_1} \left[ \int g(p, q_1, r_1) dG(r_1) \right]. \quad (9.7) \]

It is not too difficult to write down these relations, but it is quite difficult to obtain either analytic or numerical results from these equations.

10. The Estimation Problem

Let us emphasize that by virtue of assumption (9.3c), we have bypassed one of the fundamental difficulties associated with adaptive processes—namely, the estimation problem. An essential part of the adaptive control process is the learning process, entailing the optimal use of information as it is acquired. How should one modify an a priori guess on the basis of experience?

We have employed a Bayes estimation procedure because of its simplicity and intuitive character. There is no reason to suppose that this is an optimal estimation procedure in all cases, and, as a matter of fact, one can be sure that it is not. References to some work in this area will be found at the end of the chapter.

11. From Functionals to Functions

The relation in (9.6) can be used to establish existence and uniqueness theorems, to determine the structure of optimal policies, and to study asymptotic behavior. It is not as it stands useful for numerical purposes, because of the appearance of functionals, functions of functions.

To this end, we introduce the familiar concept of "sufficient statistics." We suppose that \( G(r) \) is a member of a family of distributions, \( G(r, a) \), where \( a \) is now a finite-dimensional
vector, and that the result of the estimation procedure implicit in (9.3c) and (9.4) is to replace $a$ by $a_1$, where $a_1$ depends upon $a, r_1, p$, and $q_1$.

In this case, $f_N(p, G)$ is replaced by $f_N(p, a)$, and (9.6) becomes

$$f_N(p, a) = \max_{q_1} \left[ \int \left[ g(p, q_1, r_1) + f_{N-1}(T(p, q_1, r_1), a_1) \right] dG(r_1) \right],$$

(11.1)

where $a_1 = H(a, p, q_1, r_1)$, with $H$ a fixed function.

This is the technique we employed in the three examples we discussed.

12. More General Adaptive Processes

We have by no means exhausted the reaches of uncertainty. To begin with, we can consider the case where the transformation of (9.1) has fixed, but unknown, parameters in it, say

$$p_1 = T(p, q, r, b),$$

(12.1)

again with an a priori distribution function. Secondly, we can suppose that $p$ itself cannot be determined accurately at each stage, with the result that only a distribution function for $p$ is known. Thirdly, we can assume that the objective is not clearly defined initially. Finally, we can introduce the problems of determining whether or not stochastic influences are present, or whether there are hidden variables, or whether hostile influences are present.

It is clear that there is little point in attempting to formulate at this stage of control theory what we mean by the general adaptive control process. It is, however, clear that there is unbounded range for imagination and talent in these areas and that practically everything remains to be done.

* * *

BIBLIOGRAPHY AND COMMENTS

§1. For further discussions and references, see the following books:


§2. Processes of this type were discussed by


§7. These results were given in


§8. See


Author Index

Aczel, J., 29
Adorno, D. S., 90
Aoki, M., 84
Aris, R., 55
Arrow, K. J., 30, 55
Azen, S., 71

Bateman, H., 56
Beckenbach, E. F., 56
Beckwith, R., 91
Bellman, R., 28, 29, 30, 55, 56, 71, 72, 82, 83, 89, 90, 91, 92, 104, 105
Berkovitz, L. D., 55, 83
Bertram, J. E., 91
Bharucha-Reid, A. T., 29
Birkhoff, G. D., 14, 28
Bode, H. W., 56
Boltyanskii, V. G., 83
Brock, P., 30, 55
Bucy, R., 90

Clark, C., 30
Cooke, K. L., 28, 72
Cox, H., 90
Craft, C., 30
deBroglie, L., 30
deBruijn, N. G., 28
Desoer, C. A., 28, 55
Dreyfus, S., 41, 55, 56, 71, 90
Eisenhart, L. P., 29
Fel'dbaum, A. A., 105
Florentin, J. J., 90
Forsythe, G., 29
Fortet, R., 56
Freimer, M., 105
Gamkrelidze, R. V., 83
Gibson, J. E., 91
Glicksberg, I., 83, 92
Author Index

Gluss, B., 71
Gross, O., 83, 92

Hadamard, J., 14, 28, 29
Halmos, P. R., 28
Harris, T. E., 28, 30
Hille, E., 29
Howard, R., 55, 56

Ince, E. L., 29
Juncosa, M., 72

Kagiwada, H., 71
Kalaba, R., 29, 30, 56, 71, 72, 89, 105
Kalman, R. E., 91, 92
Karlin, S., 30, 55
Kelly, J., 86, 92
Koopcke, R. W., 92
Kolmogorov, A., 84
Kotkin, B., 71

Kramer, J. D., 91
Kushner, H. J., 56

Lanczos, C., 50
LaSalle, J. P., 92
Lehman, R. S., 82
Levi-Civita, 28
Levinson, N., 91
Lie, S., 29

Malcolm, D., 30
Marshak, J., 92
McCracken, D. D., 72

Merriam, C. W. III, 91
Mischenko, E. F., 83
Montroll, E., 30
Ono, T., 90
Osborn, H., 83

Pascal, 55
Pearson, J. D., 90
Poincaré, H., 14, 28
Pontryagin, L. S., 55, 83
Prestrud, M., 71

Ricciardi, F., 30
Roberts, S. M., 55
Rozonoer, L. I., 55
Sawaragi, Y., 90
Scarf, H., 30, 55
Schwepp, F. C., 56
Sunahara, Y., 90

Thompson, W. R., 100, 105
Tou, J. T., 90, 105

Vadhanaphuti, B., 90
Valentine, F. A., 55
Wald, A., 105
Wasow, W., 29
Westcott, J. H., 90
Wiener, N., 84, 91
Wing, G. M., 29

Zadeh, L. A., 28
Subject Index

adaptive control, 29, 89, 90, 91, 93, 96, 104, 105
adaptive gambler, 99
adaptive processes, 89, 104
addition formulas, 21
approximation of curves, 71
approximation in function space, 60
approximation methods, 90
approximation in policy space, 32, 46, 47, 50, 52, 56, 61
a priori distribution function, 98
asymptotic control theory, 90
asymptotic methods, 28
asymptotic sufficient statistics, 27
asymptotic theory, 28, 90

bang-bang control, 92
Bayes estimation, 103
betting policy, 86
branching processes, 27, 28
business game, 30

calculus of variations, 39, 41, 55
causality, 7, 20
characteristic functions, 82
characteristic value, 41, 82
chemical engineering, 55
chemical reactions, 92
communication channel, 90
communication networks, 50
communication theory, 73
computational aspects, 48, 57
computational feasibility, 61
computational solution, 71
constraints, 40, 55, 83
continuous groups, 29
continuous multistage decision process, 38
continuous multistage processes, 14
control, 55, 73
control processes, 83, 92
control theory, 56
controlled system, 13
correlation, 22
criterion function, 33
curse of dimensionality, 63
decision variable, 32
decision vector, 32
determinism, 20
differential approximation, 62
differential difference, 28
differential-difference equations, 72
differential equations, 72
digital computer, 3, 29, 48, 57, 58, 92
digital control systems, 91
discrete processes, 15
duality, 42
dynamic programming, 31, 55, 71, 90, 105
dynamical systems, 28
efficient gambler, 86
eikonal equation, 48, 49, 51, 56
entropy function, 89
equipment replacement policy, 66, 72
ergodic behavior, 28
estimation, 90, 103
Euler equation, 40, 41, 82
expected value, 42
extrapolation, 91
feedback control, 43, 74, 91
Fermat's principle, 48, 51, 55, 56
filtering theory, 91
finite-difference methods, 29
finite-stage processes, 10
FORTRAN program, 67, 72
Fredholm integral equation, 82
Fredholm resolvent, 83
functional, 5, 103
functional approximations, 71
functional equations, 29
general stochastic processes, 26
geometric aspects, 41

graph theory, 56
Green's function, 81, 82, 83
hidden variables, 24, 29
Huygen's principle, 50
independence of past, 6
index of refraction, 48
indirect processes, 26
induced processes, 24
infinite processes, 9, 10
information, 44, 72, 85, 86, 89, 92
information rate, 92
inhomogeneous atmosphere, 19
initial value problem, 40
intelligent machines, 30
interpolation, 91
interrupted control processes, 30
invariant imbedding, 29, 71
inventory, 30, 55
inventory theory, 27
iteration, 28

kth best policies, 56

Lagrange multiplier, 41, 65
learning, 89, 96, 105
learning processes, 27
least squares fitting, 71
linear prediction, 91
linear systems, 91

Markov chains, 30
Markov processes, 29
Markovian decision processes, 55
mathematical experimentation, 30
mathematical model-making, 30
matrix analysis, 83, 90
maximum altitude, 17
maximum operation, 56
maximum range, 29
monotone convergence, 56
multistage betting process, 87
Subject Index

multistage decision processes, 15, 31, 32
multistage processes, 2, 4

network, 50
Neyman-Pearson lemma, 83
noisy observations, 23
nonlinear boundary-value problems, 56
nonlinear circuit theory, 30
nonlinear recurrence relation, 45
numerical stability, 64

operation, 5
optimal policy, 33
optimum digital control systems, 90

partial differential equations, 71
policy, 32, 33
policy function, 33
polynomial approximation, 62, 71
Pontryagin maximum principle, 55
positivity, 56
prediction theory, 84
principle of optimality, 35, 37, 51, 55
probability distribution, 44
problem-solving, 30, 55
process, 1, 3
production, 30, 55

quasilinearity, 56
quasilinearization, 32, 47, 56, 72

radiative transfer, 71
recurrence relations, 7
reduction of data, 5
reduction of dimensionality, 30, 72
renewal processes, 27
resolvents, 81
return function, 33
Riccati equation, 81

routing, 56
sampling systems, 91
search techniques, 65
second shortest path, 54
self-organizing systems, 30
semigroups, 29
separable criteria, 34
sequential analysis, 105
sequential design, 105
shortest path through networks, 50
smoothing, 91
solution, 37
stable, 65
stability, 65
stability theory, 28, 71
state, 2
state vector, 2
stationary, 12
statistical communication theory, 89, 105
steady-state behavior, 10
stochastic approximation, 65
stochastic control, 56, 78
stochastic multistage decision processes, 42
stochastic processes, 21, 29
stochastic systems, 90
stochastic transformation, 21
stochastic variational processes, 56
stop rules, 11
successive approximations, 45, 51, 60
sufficient statistics, 27
systems, 1, 28

terminal control, 34, 74
time-dependent processes, 12, 71
time-lag processes, 91
time-optimal control, 92
time-optimal feedback control systems, 50
trajectories, 84
trajectory process, 11, 16
transport theory, 29
transportation networks, 50
two-armed bandit, 100
two-point boundary-value problem, 40, 91
uncertainty, 93
uniqueness, 20
wave front, 50
Wiener integrals, 30