NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.
AN INVERSE PROBLEM IN DYNAMIC PROGRAMMING AND AUTOMATIC CONTROL

Richard Bellman and Robert Kalaba

PREPARED FOR:
UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA
AN INVERSE PROBLEM IN
DYNAMIC PROGRAMMING AND
AUTOMATIC CONTROL

Richard Bellman and Robert Kalaba

This research is sponsored by the United States Air Force under Project RAND—contract No. AF 49(638)-700 monitored by the Directorate of Development Planning, Deputy Chief of Staff, Research and Development, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.
PREFACE

Part of the Project RAND research program consists of basic supporting studies in mathematics. The mathematical research presented here deals with an inverse problem in automatic control theory.
SUMMARY

The authors investigate the inverse of a basic problem of automatic control theory, namely, being given an optimal policy and the equations and inequalities representing the constraints of the problem, they seek the criterion function or functions that lead to this policy; methods of solution are found in some simple cases.
CONTENTS

PREFACE ................................. 11
SUMMARY ................................. iii

Section
  1. INTRODUCTION ...................... 1
  2. SCALAR CASE ....................... 2
  3. THE FIRST-ORDER EQUATION FOR $f_T$ .... 3
  4. MULTIDIMENSIONAL CASE .......... 4
  5. DISCUSSION ......................... 5

REFERENCES ............................ 6
1. INTRODUCTION

One of the basic problems of modern control theory is that of minimizing a functional

\[ J(x,y) = \int_0^T g(x,y) dt, \]

where \( x \) and \( y \) are \( N \)-dimensional vector functions subject to the relations

\[ \begin{align*}
(1.2) (a) & \quad x(0) = c, \\
(b) & \quad \frac{dx}{dt} = h(x,y), \\
(c) & \quad r_i(x,y) \leq 0, \quad i = 1,2,...,k.
\end{align*} \]

Problems of this nature may be attacked by means of classical variational techniques and extensions \[1\], \[2\], or by means of dynamic programming. The latter emphasizes the physical concept of feedback control, since it attempts to determine directly the optimal policy \( v = v(c,T) \) as a solution of the equation

\[ f_T = \max_v \left[ g(c,v) + (\text{grad}_f, h) \right], \]

where \( r_i(c,v) \leq 0, \quad i = 1,2,...,k \), and

\[ f(c,T) = \min J(x,y). \]
In this Memorandum we initiate investigation of an inverse problem of some importance: "Given the optimal policy \( v = v(c,T) \) and the descriptive equations of (1.2), what criterion function possesses this optimal policy?" The study of inverse problems forms a chapter in the calculus of variations, with emphasis laid on determining Lagrangians having a given family of curves as Euler arcs. (See [4], [5].)

In the simple case where no constraints are present and \( h(x,y) = y \), we obtain a simple linear partial differential equation which, in principle, determines \( f \) and \( g \). In some cases, this leads to an effective analytic solution.

2. SCALAR CASE

Consider the problem of minimizing

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

with \( u(0) = c \), where \( u(t) \) is a scalar function. Writing \( f(c,T) \) for the minimum value, we obtain in the usual fashion [3] the equation

\[
f_T = \min_v [g(c,v) + vf_c].
\]

This in turn leads to the two equations.
(2.3) \[ f_T = g(c,v) + vf_c , \]
\[ 0 = g_v + f_c , \]
or
(2.4) \[ f_T = g - vg_v , \]
\[ f_c = - g_v . \]

Eliminating \( f \) by partial differentiation, we have the equation

(2.5) \[ (g - vg_v)_c = - (g_v)_T , \]

which, in principle, determines \( g \), given the function \( v(c,T) \).

This equation in \( g \) involves second partial derivatives with respect to \( v \) and is not immediately tractable.

3. THE FIRST-ORDER EQUATION FOR \( f_T \)

Returning to (2.3), let us proceed in the following fashion. Write

(3.1) \[ (f_T - vf_c)_T = g(c,v)_T = g_v v_T = - v_T f_c , \]
upon taking account of the second equation of (2.3).

Hence,

(3.2) \[ f_{TT} - v_T f_c - v_{f_c T} = - v_T f_c , \]
or

(3.3) \((f_T)_T - v(f_T)_c = 0\),

a first-order linear equation for \(f_T\), given the function \(v = v(c,T)\). Once \(f\) is determined, we obtain \(g\) from (3.3).

As is immediately seen, the determination of \(f_T\) is quite simple if we ask for a criterion function which yields an optimal policy of the form \(v(c,T) = \varphi(c)k(T)\).

4. MULTIDIMENSIONAL CASE

Consider the multidimensional problem. The basic equation is

(4.1) \(f_T = \min_v [g(c,v) + (\text{grad } f,v)]\).

In terms of the components, we obtain the equations

(4.2) \(f_T = g(c,v) + \sum_{i=1}^{N} f_{ci} v_i\),

\[0 = g_{v_i} + f_{ci},\]

which yield, in exactly the same fashion as before, the equation

(4.3) \((f_T)_T = \sum_{i=1}^{N} (f_T)_{ci} v_i\).
Again, the determination of $f_T$, given that $v(c,T) = \varphi(c)k(T)$ (where $k(T)$ is a scalar and $\varphi(c)$ is a vector), is reduced to a conventional problem in first-order partial differential equations which can be treated by means of characteristics or otherwise.

5. DISCUSSION

Finally, let us note that if one of the significant functions of a variational problem is given, it is frequently possible to form equations for the others in terms of it. For example, if the integrand $g(c,v)$ is given, then it is possible to find a first-order partial differential equation for the optimal control function $v(c,T)$. In fact, upon expanding, we see that equation (2.5) yields the desired relationship. These observations have interesting consequences in analytical dynamics and control theory which will be discussed elsewhere.
REFERENCES


