MEMORANDUM
RM-4474-PR
FEBRUARY 1965

DYNAMIC PROGRAMMING, GENERALIZED STATES, AND SWITCHING SYSTEMS
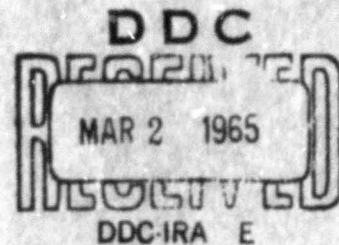
Richard Bellman

The RAND Corporation
SANTA MONICA • CALIFORNIA

DYNAMIC PROGRAMMING, GENERALIZED STATES, AND SWITCHING SYSTEMS

Richard Bellman

PREFACE

Part of the Project RAND research program consists
of basic supporting studies in mathematics. In this
Memorandum the author presents an alternative mathematical
approach to a type of general problem arising in the
study of control processes.

SUMMARY

Control problems associated with the linear vector system

$$\frac{dx}{dt} = Ax + f(t), \quad x(0) = c,$$

with $f(t)$, the control vector, subject to nonclassical constraints, have received a great deal of attention in recent years. In particular, let us cite the "bang—bang" process where each component of $f(t)$ is allowed to assume only two distinct values.

In this paper, we present an alternative formulation in dynamic programming terms which is independent of the dimension of x, the state vector. It is based upon an extension of the concept of state variable and has application to a number of systems with switching characteristics. In its simplest form, the approach was used in the study of adaptive control processes.

CONTENTS

DYNAMIC PROGRAMMING, GENERALIZED STATES, AND SWITCHING SYSTEMS

1. INTRODUCTION

Control problems associated with the linear vector system

$$(1.1) \qquad \frac{dx}{dt} = Ax + f(t), \quad x(0) = c,$$

with $f(t)$, the control vector, subject to nonclassical constraints, have received a great deal of attention in recent years. In particular, let us cite the "bang-bang" process where each component of $f(t)$ is allowed to assume only two distinct values; see [1,2,3]. Other references will be found in these sources.

Although there are many approaches with varying degrees of effectiveness now available, it cannot be said that the problem of numerical solution of problems of this genre has been completely resolved. The situation is, of course, even more unsatisfactory when the basic equation describing the system is nonlinear. In this paper we wish to make a contribution to the general problem by considering the case where $f(t)$ has only one nonzero component. We may consider that this type of problem arises in the case where the system is described by a scalar equation of the form

$$(1.2) \qquad u^{(N)} = g(u^{(N-1)}, \ldots, u, t, v(t)),$$

where $v(t) = \pm 1$. It should also be pointed out that
this particular control process can be used as the basis
of a method of successive approximations. We shall
return to this point below.

Control processes of general type, with or without
constraints, can readily be formulated in dynamic
programming terms; see [4], [5]. Numerical application
of this formulation is limited at the present time by the
limited rapid-access storage capacities of current
digital computers.

In what follows, we present an alternative formula-
tion in dynamic programming terms which is independent
of the dimension of x, the state vector. It is based
upon an extension of the concept of state variable and
has application to a number of systems with switching
characteristics. In its simplest form, the approach was
used in the study of adaptive control processes; see
[5], [6].

2. EXTENDED STATE

In the classical formulation of descriptive and
control processes, the state of the system is defined to
be the minimal set of data required to determine the
future behavior of the system; see [5], [7]. Let us now

expand this concept in the following manner. The

extended state of a system is an algorithm which permits

us to calculate the state.

The point of this is that specification of the

algorithm may require very little rapid—access storage.

On the other hand, time is required for the calculation.

Thus, as usual, we are trading time for rapid—access

storage. This idea has been used both in our previous

work in dynamic programming and in quasilinearization [8].

3. DISCRETE SWITCHING PROCESS

Consider the vector difference system

$$(3.1) \qquad x_{n+1} = g(x_n, y_n), \quad x_0 = c,$$

where y_n is a control vector subject at each time to

the condition that it belong to a constraint set R,

$y_n \in R$. Since we are thinking in terms of a digital

computer calculation, there is no loss of generality in

beginning with a discrete process. Let it be required

to choose the y_n in R so as to minimize

$$(3.2) \qquad \|x_N - z\|,$$

where $\|\cdots\|$ denotes some measure of the deviation of

x_N from z.

Writing

(3.3) $\qquad f_N(c) = \min_{y_n \in R} \|x_N - z\|$,

we readily obtain the recurrence relation

(3.4) $\qquad f_N(c) = \min_{y \in R} f_{N-1}(g(c,y)), \quad N \geq 1,$

$\qquad\qquad f_0(c) = \|c - z\|.$

If the dimension of x_n is large, this is not computationally feasible; see [9] for discussion.

4. ALTERNATIVE FORMULATION

Let us now consider the case where each y_n has only one nonzero component, say the first, which can assume only the values ± 1. Then a policy consists of a choice of $+1$ for T_1 stages, -1 for T_2 stages, and so on, or -1 for T_1 stages, $+1$ for T_2 stages, and so forth.

We therefore introduce the extended states

(4.1) $\qquad [+,T_1,T_2,\ldots,T_k], \quad [-,T_1,T_2,\ldots,T_k]$

at time $n = T_1 + T_2 + \cdots + T_k > 0$. The first state indicates that $+1$ has been used for T_1 stages, -1 for the next T_2 stages, and so on. We suppose that all

T_i are positive. With the aid of the equation in (3.1), we can now calculate the actual state in phase space.

Introduce the two functions

(4.2) $\quad f_N^+(T_1,T_2,\ldots,T_k)$ = distance from z at the end of N stages, starting in extended state $[+,T_1,T_2,\ldots,T_k]$, and using an optimal policy,

and $f_N(T_1,T_2,\ldots,T_k)$, defined similarly.

The principle of optimality now yields, in the usual fashion, the functional equations

(4.3) $\quad f_N^+(T_1,T_2,\ldots,T_k)$

$$= \min[f_{N-1}^+(T_1,T_2,\ldots,T_k+1),f_{N-1}^+(T_1,T_2,\ldots,T_k,1)],$$

for $N \geq 1$, with $f_0^+(T_1,T_2,\ldots,T_k) = g(T_1,T_2,\ldots,T_k)$, a quantity calculable using (3.1), and

(4.4) $\quad f_N^-(T_1,T_2,\ldots,T_k)$

$$= \min[f_{N-1}^-(T_1,T_2,\ldots,T_k+1),f_{N-1}^-(T_1,T_2,\ldots,T_k,1)],$$

for $N \geq 1$, with $f_0^-(T_1,T_2,\ldots,T_k)$ calculable.

The quantity $\min \|x_N - z\|$ is given by $\min[f_{N-1}^+(1), f_{N-1}^-(1)]$.

5. COMPUTATIONAL ASPECTS

If the dimension of x_n is large, and we restrict the number of switchings and the duration of the process suitably, it is easy to see that the formulation of Sec. 4 requires considerably less rapid—access storage than the usual formulation of Sec. 3.

Let us also point out that in the case where y_n has a number of nonzero components, we can use the foregoing procedure as a method of successive approximations; see the discussion of the general Hitchcock—Koopmans—Kantorovich scheduling problem in [9].

6. PRODUCTION PROCESSES

In a number of production and experimentation processes, the equation describing the system takes the form

$$(6.1) \qquad x_{n+1} = A_n x_n + y_n, \quad x_0 = c,$$

where the matrices A_n are to be chosen subject to constraints. These may be treated in the same way as above.

REFERENCES

1. Bellman, R., I. Glicksberg, and O. Gross, "On the 'Bang–bang' Control Problem," Q. Appl. Math., Vol. 14, 1956, pp. 11–18.

2. LaSalle, J. P., "Time Optimal Control Systems," Proc. Nat. Acad. Sci. USA, Vol. 45, 1959, pp. 573–577.

3. Bellman, R., I. Glicksberg, and O. Gross, Some Aspects of the Mathematical Theory of Control Processes, The RAND Corporation, R–313, 1958.

4. Bellman, R., Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

5. ———, Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, New Jersey, 1961.

6. ———, "A Problem in the Sequential Design of Experiments," Sankhya, Vol. 16, 1956, pp. 221–229.

7. Zadeh, L., and C. Desoer, Linear System Theory, The State Space Approach, McGraw–Hill Book Company, Inc., New York, 1963.

8. Bellman, R., and R. Kalaba, Quasilinearization and Nonlinear Boundary–value Problems, American Elsevier Publishing Company, New York, to appear.

9. Bellman, R., and S. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1962.