SOME NEW TECHNIQUES IN THE DYNAMIC PROGRAMMING SOLUTION OF VARIATIONAL PROBLEMS

Richard Bellman

F-1148

August 6, 1957
SUMMARY

In previous papers, it has been shown that the functional equation technique of dynamic programming may be applied to yield the numerical solution of a wide class of variational problems of the type occurring in mathematical physics, engineering, and economics.

It was seen that the numerical solution of a problem involving \( N \) state variables depended upon the computation of sequences of functions of \( N \) variables. This fact made the method routine only for the case where \( N = 1 \) or \( 2 \), with grave difficulties arising in the general case.

In this paper, it is indicated how to overcome this difficulty for a large class of problems in which the underlying equations and the criterion function are linear, although the restraints on the forcing functions may be nonlinear, corresponding say to energy considerations.

The same methods are applicable to other classes of linear equations, and, in particular, to differential–difference equations, arising from time–lag problems, and to various classes of partial differential equations. These problems could not previously be treated by dynamic programming techniques in any usable fashion.

Finally, it is briefly indicated how the method of successive approximations may be combined with the foregoing techniques to reduce general variational problems, in which the
equations and criterion function are nonlinear, to sequences of problems which can be solved numerically by means of sequences of functions of one variable. There are a number of interesting and difficult convergence questions associated with this program which we do not discuss here.
SOME NEW TECHNIQUES IN THE
DYNAMIC PROGRAMMING SOLUTION OF VARIATIONAL PROBLEMS

Richard Bellman

1. INTRODUCTION

A variational problem that is encountered in many parts of pure and applied mathematics is that of determining the minimum or maximum of a functional of the form

\[ J(y) = \int_0^T F(x_1, x_2, \ldots, x_M, y_1, y_2, \ldots, y_M) \, dt \]

over all functions \( y_1, y_2, \ldots, y_M \) satisfying the relations

\[ \frac{dx_i}{dt} = G_i(x, y), \quad x_i(0) = c_i, \quad i = 1, 2, \ldots, N, \]

and constraints of the form

\[ \begin{align*}
(a) \quad & x_i(T) = b_i, \quad i = 1, 2, \ldots, R, \\
(b) \quad & R_j(x, y) \leq 0, \quad j = 1, 2, \ldots, k.
\end{align*} \]

For a variety of reasons which one discovers almost immediately, this problem, to the degree of generality stated above, presents formidable difficulties. These difficulties arise not only in connection with the analytic solution of the problem, cf. [9], [10], but also in connection with the apparently more modest demand for a numerical solution. We have discussed elsewhere various applications of the theory of dynamic programming to the numerical solution of classes of
variational problems of the foregoing type, [1], [2], [3], [4].

In this paper we wish to indicate some recent developments which greatly enlarge the scope of these methods presented in the cited references. In addition, these new developments, combined with the classical tool of successive approximations, enable us to attack systematically classes of problems formerly far beyond our powers.

We shall begin our discussion with a terminal control problem for a linear system with constant coefficients. The problem is that of maximizing a functional of the form

\[ J(y) = H(x_1(T), x_2(T), ..., x_k(T)) \]

over all functions \( y_1, y_2, ..., y_N \), where \( x \) and \( y \) are linked by linear relations of the form

\[ \frac{dx_i}{dt} = \sum_{j=1}^{N} a_{ij} x_j + \sum_{j=1}^{N} b_{ij} y_j, \quad x_i(0) = x_i, \quad i = 1, 2, ..., N \]

and the \( y_j \) satisfy restraints of the type

(a) \( m_j \leq y_j(t) \leq m_j', \quad 0 \leq t \leq T, \quad j = 1, 2, ..., N \),

(b) \( \int_0^T K_j(y_1, y_2, ..., y_N) dt \leq c_j, \quad j = 1, 2, ..., l. \)

Whereas the techniques of dynamic programming discussed previously convert this variational problem into one requiring the computation of a sequence of functions of \( N \) variables, the linearity of the defining equation in (5) enables us to
transform the problem into one involving a sequence of functions of \( k \) variables, where \( k \) is as in (4).

Since current digital computers do not take kindly to storage of functions of more than two variables, this is a very important reduction. This is particularly so since a large class of problems occurring in the fields of engineering and industrial economics may be formulated in the above terms with large \( N \) and \( k \) of one or two.

If \( H(x_1, x_2, \ldots, x_k) \) is linear,

\[
H = \sum_{j=1}^{k} a_j x_j,
\]

(7)

the problem may be still further reduced to the computation of sequences of functions of one variable, regardless of the value of \( k \). The same remark holds for the case where it is desired to maximize a linear functional of the form

\[
J = \int_0^T \left\{ \sum_{j=1}^{N} \beta_j x_j(t) \right\} dt.
\]

(8)

In exactly the same way we can handle discrete control problems where difference equations replace differential equations. As a matter of fact, these techniques were developed in connection with a discrete problem, the "caterer" problem, [5].

The same methods may be applied to the treatment of problems involving time lags and retarded control, cf. [7], in which case the equations corresponding to (5) become
\[ \frac{dx_i}{dt} = \sum_{j=1}^{N} a_{ij} x_j(t) + \sum_{j=1}^{N} a_{ij} x_j(t-\delta) + \sum_{j=1}^{N} b_{ij} y_j(t), \]

(9)

\[ y_1(t) = c_1(t), \quad 0 \leq t \leq \delta, \quad i = 1, 2, \ldots, N. \]

More complicated types of hereditary processes may also be treated by means of the techniques we shall present below. Problems of this kind could not formerly be treated by dynamic programming techniques because of their dependence upon functionals rather than functions.

The results stated so briefly in the foregoing paragraphs depend in an essential manner upon the linearity of the equations defining the process and upon the linearity of the criterion function. To extend these techniques to cover more general situations, we turn to that general factotum, the method of successive approximations. Using this basic technique of analysis, we show how a variety of apparently multidimensional processes can be reduced to computational processes involving sequences of functions of one variable. Once this has been done we have a feasible approach to these problems.

A large number of interesting and significant problems involving convergence of these methods, rapidity of convergence, stability, and so on, arise from these investigations. These will be discussed in detail at a subsequent date.

§2. DYNAMIC PROGRAMMING

In order to appreciate the improvement in technique
afforded by the methods we present here, let us sketch briefly the direct approach of dynamic programming which we have presented in the works cited above.

Considering the general problem posed in (1.1)-(1.3) define the function of the \(N\) variables \(c_1, c_2, \ldots, c_N\), and \(T\),

\[ f(c_1, c_2, \ldots, c_N; T) = \max_y J(y) \]

where the maximum is taken over all functions \(y_1(t)\) satisfying (1.3). Under appropriate assumptions concerning the continuous dependence of maximizing \(y_1\) upon the initial values \(c_1\) and \(T\), we obtain for \(f\) a nonlinear partial differential equation

\[ \frac{\partial f}{\partial t} = \max_{v_1} \left[ F(c_1, v) + \sum_{i=1}^{N} g_i(c_i, v) \frac{\partial f}{\partial c_i} \right], \quad f(c_1, c_2, \ldots, c_N, 0) = 0, \]

where the maximization is over quantities \(v_1\) satisfying the constraints

1. \(m_1 \leq v_1 \leq m_1'\)
2. \(R_j(c, v) \leq 0.\)

Provided that \(F\) and \(C\) are suitably differentiable, and assuming that the maximum is always assumed inside the region of variation, this nonlinear partial differential equation leads via characteristics to the usual Euler equations, cf. [1], [2].
\section{Computational Aspects}

Since the analytic solution of these problems, as mentioned above, is only rarely attained, we turn to computational techniques.

To determine $f(c_1, c_2, \ldots, c_N, T)$ we can either use (2.2) and any of a number of standard techniques for the numerical solution of partial differential equations of this type, or, as has turned out to be preferable, we can go over to a discrete version of the original continuous process. Analytically, this means that the original differential equations are replaced by difference equations. Thus, (1.2) becomes

\begin{align}
    x_1(t+\delta) &= x_1(t) + \delta a_1(x_1(t), x_2(t), \ldots, x_N(t)); \\
    y_1(t), \ldots, y_M(t), x_1(0) &= c_1,
\end{align}

with $t$ assuming only the values 0, 5, 25, ... .

The nonlinear partial differential equation is then replaced by the nonlinear recurrence relation

\begin{align}
    f(c, T) &= \max_v [\delta P(c; v) + f(c_1+\delta a_1, \ldots, c_N) + \delta g_N(c, v; T-C)].
\end{align}

In order to carry out the indicated process, we must be able to tabulate functions of $N$ variables. Consequently, at the present time this approach is only feasible if $N = 1$.
or 2. For \( N > 2 \), the memory requirements become prohibitive.

It is necessary then to develop some new techniques if we wish to utilize dynamic programming to solve large scale problems of the type arising in the engineering and economic spheres.

\section*{6. PRELIMINARIES ON LINEAR SYSTEMS}

In this section we shall mention some well-known results concerning the solution of vector-matrix systems of linear differential equations. These will be utilized in what follows. Proofs of the results cited here may be found in [6].

The linear system of (1.5) may be written, using an obvious vector-matrix notation, in the form

\begin{equation}
\frac{dx}{dt} = Ax + By, \quad x(0) = c.
\end{equation}

Consider first the case in which \( A \) is constant. The solution of (1) may then be written in the form

\begin{equation}
x = e^{At}c + \int_{0}^{t} e^{A(t-s)}By(s)ds.
\end{equation}

If \( A = A(t) \), a matrix dependent upon \( t \), then \( x \) may be written in the form

\begin{equation}
x = X(t)c + \int_{0}^{t} X(t)X^{-1}(s)By(s)ds,
\end{equation}

where \( X(t) \) is the matrix solution of

\begin{equation}
\frac{dX}{dt} = A(t)X, \quad X(0) = I.
\end{equation}
We shall utilize these representations in a crucial manner below.

45. TERMINAL CONTROL

Let us now turn to the problem of maximizing a given function \( H(x_1(T), x_2(T), \ldots, x_k(T)) \) of the terminal state of the system over all control functions \( y_i(t) \) which are related to the \( x_i \) by means of the linear equation in (4.1), and which are subject to the constraints

(a) \( m_1 \leq y_i(t) \leq m'_1, \quad 0 \leq t \leq T, \quad i = 1, 2, \ldots, N, \)

(b) \( \int_0^T G(y_1, y_2, \ldots, y_N) dt \leq k. \)

We wish to show that the numerical solution of a problem of this type can be made to depend upon a sequence of functions of \( k \) variables, rather than upon sequences of functions of \( N \) variables. We shall consider first the case where \( A = (a_{ij}) \) is a constant matrix.

We begin with the linear representation of (4.2), which yields a set of equations

\[ x_i(t) = z_i(t) + \int_t^T \left[ \sum_{j=1}^N x_{ij}(t-s)y_j(s) \right] ds, \]

\[ i = 1, 2, \ldots, N, \]

where \( z_i(t) \) is the \( i \)-th component of \( e^{At}c \), and \( x(t) = (x_{ij}(t)) \).

The problem we wish to consider may then be cast in the form of maximizing a functional of the type
\[ H(u_1 + \int_0^T \left[ \sum_{j=1}^N x_j(T-s) y_j(s) \right] ds, \ldots, u_k \]
\[ \quad + \int_0^T \left[ \sum_{j=1}^N x_{kj}(T-s) y_j(s) \right] ds, \]

where \( u_1, u_2, \ldots, u_k \) are given quantities, over all functions \( y_1, y_2, \ldots, y_N \) satisfying the constraints (1a) and (1b).

Let us then consider the sequence of functions
\[ f(u_1, u_2, \ldots, u_k; T), \]
implicitly dependent upon \( \lambda \), defined as follows
\[ f(u_1, u_2, \ldots, u_k; T) = \max_y \left[ H(u_1 + \int_0^T \ldots \right] ds, \ldots, u_k \]
\[ \quad + \int_0^T \ldots ds) - \lambda \int_0^T g(y_1, y_2, \ldots, y_N) ds, \]

where the functions \( y_1(t) \) are now constrained by (1a).

A motivation and discussion of the use of the Lagrange multiplier may be found in [8], and a numerical example in [11].

The value of the method lies in the fact that enables us to reduce multi-dimensional problems to sequences of lower dimensional problems.

To obtain a functional equation for \( f(u_1, u_2, \ldots, u_k; T) \), we proceed as follows. Suppose that the values of \( y_1(t), y_2(t), \ldots, y_N(t) \) have been determined over \([0, 5] \).

Then we may write
H(u_1 + \int_0^T \ldots ds, \ldots, u_k + \int_0^T \ldots ds)

- \lambda \int_0^T g(y_1, y_2, \ldots, y_N) ds

= H(u_1 + \int_0^\delta \ldots ds + \int_0^T \ldots ds, \ldots, u_k

+ \int_0^\delta \ldots ds + \int_0^T \ldots ds)

(5)

- \lambda \int_0^\delta g(y_1, y_2, \ldots, y_N) ds - \lambda \int_0^T g(y_1, y_2, \ldots, y_N) ds

= H(u_1 + \int_0^\delta \ldots ds

+ \int_0^{T-\delta} \left[ \sum_{j=1}^N x_{1j} (T-s) y_j(s+\delta) \right] ds, \ldots )

- \lambda \int_0^\delta g(y_1, y_2, \ldots, y_N) ds

- \lambda \int_0^{T-\delta} g(y_1(s+\delta), \ldots, y_N(s+\delta)) ds.

The principle of optimality cf. [1] then yields the

functional equation

f(u_1, u_2, \ldots, u_k; T) = \max_{y[0, \delta]} \left[ - \lambda \int_0^\delta g(y_1, y_2, \ldots, y_N) ds +

(5)

\int_0^{\delta} \left[ \sum_{j=1}^N x_{1j}(T-s) y_j(s) \right] ds, \ldots, u_k

+ \int_0^\delta \left[ \sum_{j=1}^N x_{kj}(T-s) y_j(s) \right] ds \right].
The maximum is now taken over all functions $y_1(s)$ defined over $0 \leq s \leq 5$, and satisfying the constraints $m_1' \leq y_1(s) \leq m_1$ in $[0,5]$.

For computational purposes, we may use the approximate relation

$$f(u_1, u_2, \ldots, u_k; T) = \max_v \left[ - \lambda \int G(v_1, v_2, \ldots, v_N) + \right.$$

$$\left. \sum_{j=1}^{N} x_{1j}(T)v_j, \ldots, u_k + \right.$$\n
$$\sum_{j=1}^{N} x_{kj}(T)v_j \right],$$

or we may start with a discrete version of the original process.

We have thus reduced the numerical solution of the variational problem to the determination of a sequence of functions of $k$ variables. If $k = 1$ or 2, we have a feasible method of solution.

**VI. TERMINAL CONTROL—VARIABLE COEFFICIENTS**

It is important in connection with our subsequent discussion of the use of successive approximations to consider the same problem for the case where $A$ is a variable matrix. Let us consider then the case where the equation governing the process has the form

$$\frac{dx}{dt} = A(t)x + B(t)y + \phi(t), \quad x(0) = c.$$  

As we know, the solution of this equation is given by the expression
\[ x = x(t)c + \int_0^t x(t)x^{-1}(s)B(s)y(s)ds + \]

\[ \lambda(t)x^{-1}(s)\varphi(s)ds. \]

Hence the components of \( x(T) \) has the form

\[ x_i(T) = u_i + \int_0^T \left[ \sum_{j=1}^N x_{ij}(T,s)y_j(s) \right] ds, \]

where the \( u_i \) are independent of \( y \).

In order to take account of the non-stationarity of the process, we count time backwards. In place of noting the time at which the process ends, we single out the time at which it begins. Fixing \( T \), we consider the function \( f(u_1, u_2, \ldots, u_k; r) \) defined by the relation

\[ f(u_1, u_2, \ldots, u_k; r) = \max_{y \in \mathbb{R}^N} \left[ H(u_1 + \right. \]

\[ \int_r^T \left[ \sum_{j=1}^N w_{ij}(T,s)y_j(s) \right] ds, \ldots, u_k + \]

\[ \int_r^T \left[ \ldots \right] ds) - \lambda \int_r^T G(y_1, y_2, \ldots, y_N)ds \right]. \]

Arguing as in the preceding section, we see that \( f \) satisfies the relation

\[ f(u_1, u_2, \ldots, u_k; r) = \max_{y \in \mathbb{R}^N} \left[ - \lambda \int_r^{r+6} G(y_1, y_2, \ldots, y_N)ds \right. \]

\[ + f(u_1 + \int_r^{r+5} \left[ \ldots \right] ds, \ldots, u_k \]
\[ + \int_{r}^{\infty} \ldots \text{ds}] \].

For computational purposes, this reduces to

\[ f(u_1, u_2, \ldots, u_k; r) = \max_v \left[ -\lambda \delta \mathcal{Q}(v_1, v_2, \ldots, v_N) + \right. \]

\[ f(u_1 + 5 \sum_{j=1}^{N} w_{ij}(T, r)v_j, \ldots, u_k + \]

\[ 5 \sum_{j=1}^{N} w_{kj}(T, r)v_j) \],

with \( f(u_1, u_2, \ldots, u_k; T) = 0. \)

7. TERMINAL CONTROL—LINEAR CRITERION

Let us now consider the case where \( H \) is a linear function. More generally, let us consider the problem of maximizing the inner product \((x(T), a)\) where \( a \) is a given vector. To simplify the notation, let us consider only the case where \( A \) is a constant matrix.

Using the representation for \( x(t) \) given in (4.2), we see that

\[ (x(T), a) = (e^{AT}c, a) + (0^T e^{A(T-s)}y(s) ds, a). \]

Neglecting the term \((e^{AT}c, a)\), which is independent of \( y \), we have the problem of maximizing

\[ J(y) = (0^T e^{A(T-s)}y(s) ds, a) \]
over all \( y_1 \) satisfying the constraints

(a) \( m_1 \leq y_1 \leq m'_1, \; 0 \leq t \leq T, \; i = 1, 2, \ldots, N, \)

(b) \( \int_0^T g(y_1, y_2, \ldots, y_N) \, ds \leq k. \)

\( (3) \)

Introduce the function

\( (4) \quad f(k, T) = \max_y J(y), \)

where the maximum is over all functions \( y_1 \) satisfying (3a) and (3b).

It is easy to see that \( f(k, T) \) satisfies the equation

\[
 f(k, T) = \max_{y[0,5]} \left[ \left( \int_0^5 e^{A(T-s)} y ds \right) a \right] \\
+ f(k - \int_0^5 g(y_1, y_2, \ldots, y_N) \, ds, T-5) \right].
\]

\( (5) \quad f(k, 0) = 0. \)

For computational purposes, we can use the relation

\[
 f(k, T) = \max_v \left[ 5(e^{AT} v, a) + f(k-5 g(v_1, v_2, \ldots, v_N)T-5) \right].
\]

We see then that in the case where the underlying equation is linear, the criterion function is linear, and there is only one constraint of the form in (3b), we can compute the solution using sequences of functions of one variable.

If there are two constraints of the type
we introduce a Lagrange multiplier and consider the problem of maximizing

\[ (x(T), a) - \lambda \int_0^T q(y_1, y_2, \ldots, y_N) \, ds. \]

For each value of \( \lambda \) we have a one-dimensional problem. As the parameter \( \lambda \) is varied, we range over a set of values of the constraint \( \int_0^T q(y_1, y_2, \ldots, y_N) \, ds. \)

**QM. TERMINAL CONTROL VS. GENERAL CONTROL**

The general control problem is that of maximizing a functional of the form

\[ J(y) = \int_0^T F(x, y) \, dt + H(x(T)), \]

subject to relations and constraints of the type described above.

If we introduce a new scalar variable \( x_{N+1}(t) \) by means of the relation

\[ \frac{dx_{N+1}}{dt} = F(x, y), \quad x_{N+1}(0) = 0, \]

we see that we once again have a terminal control problem.

In particular, the problem of maximizing a linear functional of the form
(3) \[ \int_0^T \left[ \sum_{j=1}^N [a_j x_j(t) + b_j y_j(t)] \right] dt, \]

is readily treated by the above methods.

**II. TIME LAG PROBLEMS**

As we have seen in the foregoing sections, the success of our methods rested upon the superposition principle. Given an equation of the form

\[ L(x) = y, \]

\[ x_{t=0} = c, \]

we were able to write the solution in the form

\[ x = x(t)c + T(y), \]

and thus avoid interaction between the initial conditions and the forcing function.

Put in these terms, it is clear that the same methods will handle the case where the linear operator \( L \) is represented by a differential–difference operator of the type given in (1.9). The requisite representation theorems may be found in [7].

**10. PARTIAL DIFFERENTIAL EQUATIONS**

Similarly, a variety of control problems associated with the heat equation
\( u_t = u_{xx} + v(x,t), \)

(1) \( u(x,0) = c(x), \quad 0 \leq x \leq 1 \)

\( u(0,t) = u(1,t) = 0, \quad t > 0, \)

may be treated.

We shall reserve a detailed exposition of problems of this type for a future date.

**11. SUCCESSIVE APPROXIMATIONS**

Let us now briefly, without entering into any rigorous discussion, which is non-trivial, indicate how the method of successive approximations may be combined with the foregoing techniques so as to reduce general control problems involving nonlinear differential equations and nonlinear criteria to sequences of computational problems involving functions of one variable.

Consider the problem of maximizing

(1) \( J(y) = H(x_1(T), x_2(T), \ldots, x_k(T)) \)

over all \( y \) subject to

(a) \( \frac{dx_1}{dt} = q_1(x,y), \quad x_1(0) = x_{11}, \quad 1 = 1, 2, \ldots, N, \)

(b) \( m_1 \leq y_1 \leq n_1', \quad i = 1, 2, \ldots, N, \)

(c) \( \int_0^T g(y_1, y_2, \ldots, y_N) dt \leq k. \)
Let \( y^0 \) = \( (y_1^0, y_2^0, \ldots, y_N^0) \) be an initial guess in policy space, and let \( x^0 = (x_1^0, x_2^0, \ldots, x_N^0) \) be the resultant \( x \)-values, determined by (2a).

Consider the new criterion function

\[
J_0(y) = H(x_1^0(T), x_2^0(T), \ldots, x_k^0(T)) + \sum_{i=1}^{k} [x_i(T) - x_i^0(T)] \frac{\partial H}{\partial x_i} (x_1^0(T), \ldots, x_k^0(T)),
\]

and the new system of differential equations

\[
\frac{dx_1}{dt} = a_1(x^0, y) + \sum_{j=1}^{N} (x_j - x_j^0) \frac{\partial f_1}{\partial x_j} (x^0, y),
\]

\( x_1^0 = c_1, \ 1 = 1, 2, \ldots, N. \)

The new variational problem is that of maximizing \( J_0(y) \) over all \( y \) satisfying (4) and the constraints (2b) and (2c).

Since the criterion function and the system of differential equations are linear, the methods outlined in the foregoing pages permit us to solve this problem by means of a sequence of one-dimensional functions.

Let the maximizing \( y(t) \) be called \( y^{(1)}(t) \). Proceeding as above, we determine an \( y^{(1)}(t) \), and formulate a new variational problem as in (3) and (4). Proceeding in this way we form a sequence of one-dimensional problems which we hope converges to the solution of the original problem.


