MONOTONE CONVERGENCE IN DYNAMIC PROGRAMMING AND THE CALCULUS OF VARIATIONS

Richard Bellman

P-513

Revised 3 September 1954

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

Summary

In the present paper we show that the method of "approximation in policy space," developed in the theory of dynamic programming, yields monotone convergence in the calculus of variations.

MONOTONE CONVERGENCE IN DYNAMIC PROGRAMMING AND THE CALCULUS OF VARIATIONS

Richard Bellman

§1. Introduction

In [1] we outlined some applications of the functional equation approach of the theory of dynamic programming to the characterization of extremal curves and eigenvalues in the calculus of variations. A more detailed account of this new formalism will be found in [2] and [3].

The purpose of the present note is to show that another important concept in the theory of dynamic programming, that of "approximation in policy space," may also be utilized to yield some interesting results in the calculus of variations. As we shall show below, this idea leads to the solution of variational problems by iterative techniques which yield monotone approximation, and indeed monotone convergence, as we shall show elsewhere.

To illustrate this new concept, we shall consider first a discrete dynamic programming problem, and then present the analogous treatment of a continuous version, namely the maximization of $J(y) = \int_0^T F(x,y,t)dt$, subject to $\dot{x} = G(x,y,t)$, $x(0) = c$. Following this we shall discuss the application of this technique to the eigenvalue problem associated with the equation $\ddot{u} + \lambda^2\phi(t)u = 0$, $u(0) = u(1) = 0$.

Finally, we shall sketch briefly a method that may be used to demonstrate convergence of the iterative procedure.

§2. Monotone Convergence in Dynamic Programming

Let us consider the functional equation

$$f(x) = \max_{0 \le y \le x} \left[g(y) + h(x-y) + f(ay + b(x-y)) \right] \equiv T(f) \qquad (2.1)$$

where $0 < a, b < 1$, which arises in connection with various types of multi-stage allocation processes. It is readily shown that if $g(x)$ and $h(x)$ are continuous in x over an interval $[0,c]$, with $g(0) = h(0) = 0$, then, starting with any initial function $f_0(x)$ which is continuous over $[0,c]$ and zero at $x = 0$, the iterative procedure $f_{n+1}(x) = T(f_n)$ yields the unique solution of (2.1) which is continuous at $x = 0$.

To obtain monotone convergence, we approximate first in policy space. A policy is, with reference to (2.1), a choice of $y = y(x)$ with $0 \le y(x) \le x$. Let $y_0 = y_0(x)$ be an initial policy and let $f_0(x)$ be computed by recurrence from the functional equation

$$f_0(x) = g(y_0) + h(x-y_0) + f_0(ay_0 + b(x-y_0)), \quad f_0(0)=0 \qquad (2.2)$$

It is immediately clear that $f_1(x)$ as determined by

$$f_1(x) = T(f_o) \tag{2.3}$$

is greater than or equal to $f_o(x)$ for $x \geq 0$, $f_1(x) \geq f_o(x)$. From this it follows inductively that $f_{n+1}(x) = T(f_n)$ is greater than or equal to $f_n(x)$ for all $x \geq 0$. Hence, we have monotone convergence.

§3. Monotone Approximation in the Maximization of $\int_0^T F(x,y,t)dt$

As in [1] and [3], we write

$$f(a,c,T) = \operatorname*{Max}_y \int_a^{a+T} F(x,y,t)dt \tag{3.1}$$

where $dx/dt = G(x,y,t)$, $x(a) = c$. The function f satisfies the functional equation

$$f_T = \operatorname*{Max}_v \left[F(c,v,a) + G(c,v,a)f_c + f_a \right] \tag{3.2}$$

As in our previous notes, we shall avoid all discussion of necessary or sufficient conditions for these equations and present only the basic formalism.

An approximation in policy space is now a choice of y as a function of x, a, and T, which is to say $v = y(a)$ as a function of c, a, and T. Let v_o represent an initial choice, and let $f_o(a,c,T)$ denote the function obtained in this way. Then $f_o(a,c,T) = \int_a^{a+T} F(x_o,y_o,t)dt$, $\dot{x}_o = G(x_o,y_o,t)$, $x_o(a) = c$, and f_o satisfies the partial differential equation

$$f_{oT} = F(c,v_o,a) + G(c,v_o,a)f_{oc} + f_{oa} \tag{3.3}$$

$$f_o(a,c,0) \equiv 0$$

A further approximation, v_1, to an optimal policy is now determined by the condition that v_1 maximize the function of v given by

$$H(v,f_o) = F(c,v,a) + G(c,v,a)f_{oc} + f_{oa} \tag{3.4}$$

Let f_1 be the function determined by v_1, satisfying the equation $f_{1T} = H(v_1,f_1)$. Similarly we determine v_2 by the condition that it maximize $H(v,f_1)$, and so on, obtaining in this way two sequences of functions $\{v_n\}$ and $\{f_n\}$.

Let us now demonstrate the essential result that the sequence $\{f_n\}$ is monotone increasing in n for all a, c, and $T \geq 0$. We have

$$\begin{aligned} f_{1T} - f_{oT} &= H(v_1,f_1) - H(v_o,f_o) \\ &= H(v_1,f_o) - H(v_o,f_o) + H(v_1,f_1) - H(v_1,f_o) \end{aligned} \tag{3.5}$$

or

$$(f_1-f_o)_T = A(c,T,a) + (f_1-f_o)_c\, B(c,T,a) + (f_1-f_o)_a \tag{3.6}$$

where, from the manner in which v_1 was determined, $A \geq 0$.

From this, and the boundary condition at $T = 0$ for f_o and f_1, it follows readily that $f_1 - f_o \geq 0$ for all a, c, and $T \geq 0$. The same argument shows that $f_{n+1} \geq f_n$.

§4. Sketch of a Convergence Proof

The above argument yields monotone approximation with discussing convergence. One approach to a proof of convergence is to consider the corresponding discrete problem of maximizing $J(y) = \sum_{k=0}^{N} F(x_k, y_k, k)$, subject to $x_{k+1} - x_1 = G(x_k, y_k, k)$, $x_1 = c$, which yields the functional equation for $f(a,c,N)$

$$f(a,c,N+1) = \operatorname*{Max}_{y} \left[F(c,y,a) + f(a+1, c+G(c,y,a), N) \right] \qquad (4.1)$$

and use a limiting process.

§5. Monotone Convergence in Eigenvalue Problems

The problem of determining the values of λ^2 which permit nontrivial solution of

$$u'' + \lambda^2 \phi(t) u = 0, \quad u(0) = u(1) = 0 \qquad (5.1)$$

is, under slight restrictions on $\phi(t)$, equivalent to the problem of determining the relative maxkma of $\int_0^1 \phi(t) u^2 dt$ subject to the constraints $\int_0^1 u'^2 dt = 1$, $u(0) = u(1) = 0$. To attack this problem by the functional equation method outlined above, we consider the more general problem of determining the maximum of

$$J(u) = \int_a^1 \phi(t)u^2 dt + k \int_a^1 (1-t)u\phi(t)dt \qquad (5.2)$$

Setting $\underset{u}{\text{Max}}\, J(u) = f(a,k)$, we obtain (see [ ], [ ]) the equation

$$f_a = \underset{w}{\text{Min}} \left[(f-kf_k/2)w^2 - w\left[(2+k\psi(a)/1-a)\right]\right] \qquad (5.3)$$

where

$$\psi(a) = \int_a^1 (1-t)^2\phi(t)dt \qquad (5.4)$$

and $w = u'(a)$. A choice of a policy is a choice of $w = w(a,k)$. The method of successive approximations used above may again be employed and the proof of the monotonicity is essentially as before.

BIBLIOGRAPHY

1. Bellman, R., "Dynamic Programming and a New Formalism in the Calculus of Variations," Proc. Nat. Acad. Sci., April, 1954, Vol. 40, pp. 231-235.

2. ————————, "Dynamic Programming and the Calculus of Variations—I," The RAND Corporation Paper No. P-495, March, 1954.

3. ________, "Dynamic Programming and the Calculus of Variations—II," The RAND Corporation Paper No. P-512, April, 1954.

4. ________, An Introduction to the Theory of Dynamic Programming, The RAND Corporation Report No. R-245, 1953.

5. ________, "Some Problems in the Theory of Dynamic Programming," Econometrica, Vol. 22, No. 1 (January 1954), pp. 37-48.

6. ________, "Computational Problems in the Theory of Dynamic Programming," The RAND Corporation Paper No. P-423, presented to the Symposium on Applied Mathematics, Santa Monica, 1954.

BLANK PAGE