ON A METHOD OF DERIVATION OF BELLMAN EQUATION

(OB ODNOM SPOSOBE VYVODA URAVNENIIA BELLMANA)

PMM Vol. 31, No. 1, 1967, pp. 145-147

K. A. LUR'E
(Leningrad)

(Received October 1, 1966)

We shall consider the following optimum problem. A system

$$\frac{dx^i}{dt} = g^i(t;\ x,\ u), \qquad x^i(t_0) = x_0^i, \qquad i = 1, \ldots, n \tag{1}$$

of ordinary differential equations and initial conditions is given, and $u(t) = (u^1(t), \ldots, u^m(t))$ is the control vector-function.

Moreover, a set of relations expressed in terms of finite equalities (*)

$$R^j(t;\ x,\ u) = 0, \quad j = 1, \ldots, r \leqslant m \tag{2}$$

is given and we assume that the matrix $\|\partial R^j / \partial u^i\|$ is of maximum rank.

Our aim is to define the control function $u(t)$ which makes the functional $J = S_T(x(T), T)$, a minimum. This functional is a function of a finite point $(x(T), T)$ on the phase trajectory in the (x, t)-space

$$S_T(x(T), T) = F(x^1(T), \ldots, x^n(T), T) \tag{3}$$

We know, that the stated problem can be solved by means of the Bellman [1] differential equation (usual summation convention is adopted here)

$$\frac{\partial S_T}{\partial t} = \max_u \left[-\frac{\partial S_T}{\partial x^i} g^i(t;\ x,\ u) \right], \qquad u \in A(t,\ x) \tag{4}$$

Here $A(t, x)$ denotes the aggregate of admissible controls, i.e. of controls satisfying the condition (2). Solving (4) with the boundary condition (3) we find the function $S_T(x, t)$, which on substitution into

$$u = \gamma(x,\ t,\ \partial S_T / \partial x^i) \tag{5}$$

obtained from the maximum condition

$$\frac{\partial}{\partial u} \left[\frac{\partial S_T}{\partial x^i} g^i(t;\ x,\ u) + \Gamma_j R^j(t;\ x,\ u) \right] = 0 \tag{6}$$

leads to Expression

$$u = \varphi(x,\ t) \tag{7}$$

describing the synthesis of the optimum control. Solutions of the system of equations

$$dx^i/dt = g^i(t;\ x,\ \varphi(x,\ t)) \tag{8}$$

*) The case of relations defined in terms of finite inequalities is easily reduced to the present case by introducing "auxiliary" controls.

minimize the functional (3) under the arbitrary initial conditions $x^i(t_0) = x^i_0$: at the same time (2) is satisfied identically, i. e. $R^j(t;\ x,\ \varphi(x,\ t)) \equiv 0$.

The argument leading to the Bellman equation (4) is well known. In fact, (4) represents a differential formulation of the principle of the optimum. Solution of (4) with given initial conditions, yields the synthesizing function (7), which on substitution into the right-hand side of (4), gives

$$\frac{\partial S_T}{\partial t} = -\frac{\partial S_T}{\partial x^i} g^i(t;\ x,\ \varphi(x,\ t)) \tag{9}$$

Let us now return to the problem stated at the beginning and let us consider some admissible control $U(x,\ t)$. We should note that the function U chosen is a function of two independent variables x and t.

Replacing u in (1) with $U(x,\ t)$, we arrive at the differential equations of admissible trajectories. With the specified initial state $(x_0,\ t_0)$, the admissible trajectory is defined uniquely.

Consider the functional $S_T(x_0,\ t_0)$ on the admissible trajectory. Its value is defined only by the final (at the instant $t = T$) position at the phase point, therefore this value is independent of the choice of the initial point on the admissible trajectory. This is expressed by

$$\frac{dS_T(x_0,\ t_0)}{dt_0} = 0 \tag{10}$$

which is valid along the admissible trajectory.

Performing the differentiation on (10) and taking (1) into account, we obtain

$$\frac{\partial S_T(x_0,\ t_0)}{\partial t_0} = -\frac{\partial S_T(x_0,\ t_0)}{\partial x_0^i} g^i(t_0;\ x_0,\ U(x_0,\ t_0))$$

or, discarding the zeros,

$$\frac{\partial S_T(x,\ t)}{\partial t} = -\frac{\partial S_T(x,\ t)}{\partial x^i} g^i(t;\ x,\ U(x,\ t)) \tag{11}$$

Up to now we have considered (11) only along the admissible trajectory: it should now be pointed out that any trajectory of (1) (with $U(x,\ t)$ replacing u) can be selected as admissible. Hence, by (10) $S_T(x_0,\ t_0)$, is maintained along any trajectory of (1), which in turn implies [2] that $S_T(x,\ t)$ satisfies, as a function of two variables, $(x,\ t)$, Equation (11) considered as a first order partial differential equation.

It should again be stressed that the coordinates of the initial point of the admissible trajectory appear as the argument of $S_T(x,\ t)$ and the function S itself is defined as some combination of coordinates of the final point of the trajectory. With the control chosen, the values of $S_T(x,\ t)$ are conserved during the motion along the trajectory, consequently the minimum of $S_T(x(T),\ T)$ coincides with the minimum of $S_T(x_0,\ t_0)$. We find, however, that, while the values of x_0 and t_0 are known, those of $x(T)$ are not. Therefore the value of $S_T(x_0,\ t_0)$ can be minimized in place of $S_T(x,(T),\ T)$ at the given point $(x_0,\ t_0)$ of the boundary at the admissible region of the xyt plane.

Thus we arrive at the following optimum problem for Equation (11) considered as a first order partial differential equation: to find a control function $u(x,\ t)$ of two independent variables acted upon by the constraints (2) such, that the solution $S_T(x,\ t)$, satisfying condition (3) when $t = T$, assumes a minimum possible value at the point $(x_0,\ t_0)$. The control is sought in the class of piecewise continuous functions and the solution $S_T(x,\ t)$ in the class of continuous functions of two independent variables.

Solution of this problem is obtained at once, using the method given previously [3]. Denoting $S_T(x, t)$ by z, we shall write the basic equation (11) (taking for simplicity the case of one coordinate x: the case of several coordinates can be dealt with in the analogous manner) as

$$z_t = -g\zeta, \quad z_x = \zeta$$

Let us introduce the Lagrangian multipliers ξ, η and Γ. Using the Hamiltonian function

$$H = -\xi g\zeta + \eta\zeta - \Gamma R$$

we shall construct Euler's equations

$$\partial\xi / \partial t + \partial\eta / \partial x = 0, \quad \eta - g\xi = 0 \quad \xi\zeta\partial g / \partial u + \Gamma\partial R / \partial u = 0$$

which, after eliminating η, become

$$\partial\xi / \partial t + g\partial\xi / \partial x + \xi\partial g / \partial x = 0 \tag{12}$$

On the characteristic $x = x(t)$, (12) assumes the form

$$\frac{d\xi}{dt} = -\xi\left[\frac{\partial g}{\partial x}\right]_{x=x(t)}$$

from which it follows that the sign of ξ on the characteristic is set by the sign of the initial value of ξ.

Weierstrass' function constructed uner the assumption that z is continuous on the boundary on the region of variation [4] is found to be equal

$$\Delta H = -\xi\,\frac{gt_\tau - x_\tau}{Gt_\tau - x_\tau}\,\zeta\,[g - G]$$

here g corresponds to the optimum control, while G to the admissible control; t_τ and x_τ are the direction cosines of the region of variation.

If z is minimized at $t = t_0$ and $x = x_0$, then the natural boundary condition [3] gives

$$\xi(\tau) = \delta(\tau - x_0) > 0$$

where δ is the delta function and τ denotes the length of arc of the boundary of the main region, traversed in the positive direction (in our case this region consists of a strip $t_0 \leq t \leq T$ on the plane (x, t): natural condition is postulated on the line $t = t_0$).

In fact, the given condition defines only the characteristic beginning at the point (x_0, t_0). Therefore the requirement that $\Delta H \geq 0$, is actually imposed and fulfilled only along this characteristic, where it is equivalent, as we shall show later, to the usual maximum condition $\zeta\,[g - G] \geqslant 0$.

Indeed, our previous argument implies that $\xi > 0$ along the characteristic and it only remains to show that the relation $(gt_\tau - x_\tau) / (Gt_\tau - x_\tau)$ is always positive. When $g = G$, the above relation is equal to unity: when $g \neq G$, then such directions (t_τ, x_τ) can always be found, which correspond to the negative value of the above relation. Such directions should be excluded from our investigation, since the corresponding characteristics arrive at the discontinuity (boundary of the region of variation) from both sides with the result, that the perturbations accumulate at the discontinuity instead of passing through it. The usual situation is, that the perturbations originating in the region of variation pass through its boundaries and enter the outer parts of the region, and this is only possible when their velocities of propagation (slopes of characteristics) are, simultaneously, either greater or smaller than the velocity of propagation of the discontinuity (slope of the boundary of the region of variation). The last requirement is equivalent to the demand that the expression quoted above, is positive.

We thus see, that in our problem a genuine restriction on the slope of the region of variation, appears. Weierstrass' condition thus assumes a form of an inequality $-\zeta[g - G] \geq 0$ valid along the characteristic originating at the initial point (x_0, t_0). Equation (11) is at the same time rewritten as

$$\frac{\partial S_T}{\partial t} = -\max_U \left[\frac{\partial S_T}{\partial x^i} g^i(t;\ x,\ U(x, t)) \right], \qquad U \in A(t,\ x)$$

valid along the optimum trajectory. This, on utilizing a well known argument [5] easily leads to Bellman equation.

BIBLIOGRAPHY

1. Berkovitz, L. D., Variational Methods in Problems of Control and Programing. J. Math. Anal. and Appl., No. 3, 1961.
2. Trikomi, F., Lektsii po uravneniiam v chastnykh proizvodnykh (Lessons on Partial Differential Equations). M. Izd. inostr. lit., (Russian translation). 1957.
3. Lur'e, K. A., Zadacha Maiera-Boltsa dlia kratnykh integralov i optimizatsiia povedeniia sistem s raspredelennymi parametrami (The Mayer-Bolza problem for multiple integrals and optimization of the performance of systems with distributed parameters). PMM Vol. 27, No. 5, 1963.
4. Lur'e, K. A., Optimal'noe upravlenie provodimost'iu zhidkosti, dvizhushcheisia po kanalu v magnitnom pole (Optimum control of conductivity of a fluid moving in a channel in a magnetic field). PMM Vol. 28, No. 2, 1964.
5. Rozonoer, L. I., Printsip maksimuma L. S. Pontriagina v teorii optimal'nykh sistem (Pontriagin maximum principle in the theory of optimum systems), III., Avtomatika i telemekhanika, Vol. 20, No. 12, 1959.

Translated by L. K.