ON A CERTAIN FORM OF OPTIMALITY CONDITIONS

MATHEMATICS

1969

SovietRxiv

Abstract

Full Text

UDC 519.35

MATHEMATICS

Yu. P. KRIVENKOV

ON A CERTAIN FORM OF OPTIMALITY CONDITIONS

(Presented by Academician S. L. Sobolev, 27 I 1969)

The use of the mathematical theory of optimal processes in economic studies is hindered by its insufficient adaptability to the solution of large multidimensional problems and of problems involving constraints on phase coordinates (see ([1], [2])). Pursuing the aim of bringing it to such a degree of formalism as would be sufficiently convenient for machine analysis and for the solution of multi-dimensional dynamic problems of mathematical economics, the present paper describes a certain form of optimality conditions.

To prove the necessity of these conditions, in the course of developing a generalized variational calculus (see ([3], [4])), a special apparatus was developed, called the generalized Lagrange method, which makes it possible to write down the totality of optimality conditions with greater completeness. This apparatus, applied to the proof of the existence of a system of adjoint functions for a certain fairly general problem in the theory of optimal processes, made it possible, first, to note new properties of these functions and, second, to formulate a form of the optimality principle which may be useful by virtue of its convenience for machine implementation, since it is written only in terms of relations of inequality type.

The sufficiency of the optimality principle for the problem posed in the theory of optimal processes is proved by a method which, despite certain limitations, is suitable for nonlocalized problems, can be applied without special allowance for unconditional boundaries arising from phase constraints, and also without the restrictive consideration of time as a phase coordinate.

I. **Problem A.** In a space $\mathfrak{Z}$ of type B, determine an element z ensuring the condition

$$\max\{f(z) : r(z) = 0,\ \psi(z) \geqslant 0\},$$

in which $f(z)$ is a functional, and $r(z)$ and $\psi(z)$ are abstract functions belonging to the classes $(z \to \mathfrak{R})$ and $(z \to \mathfrak{P})$, where $\mathfrak{R}$ and $\mathfrak{P}$ are spaces of type B. The space $\mathfrak{P}$ is partially ordered by means of a certain convex cone $\mathfrak{C} \subset \mathfrak{P}$. The notation $\psi' \geqslant \psi''$ is used for any $\psi', \psi'' \in \mathfrak{P}$ satisfying the condition $\psi' - \psi'' \in \mathfrak{C}$.

Variations. The difference $z - z_0 + \bar{z}$, where $z, z_0 \in \mathfrak{Z}$ and z_0 is an admissible element of the problem, will be called a **variation of the element** z_0. For variations $\bar{z}$ of the element z_0, we distinguish the following sets: a) the set $\mathfrak{L}$ of variations admissible with respect to the equality-type constraint, if $r(z_0 + \varepsilon\bar{z}) = o(\varepsilon)$, where ε is small; b) the set $\mathfrak{C}$ of variations admissible with respect to the inequality-type constraint, if for all sufficiently small $z \in \mathfrak{Z}$ and small $\varepsilon > 0$ the inequality $\psi(z_0 + \varepsilon(\hat{z} + \bar{z})) > 0$ holds.

Assumptions. A_1. A solution of Problem A exists.

A_2. The abstract functions $f(z)$, $r(z)$, and $\psi(z)$ on $\mathfrak{Z}$ have Gateaux derivatives, which are linear operators on variations of the argument (see ([5]), p. 54).

A$_3$. The set $\mathfrak{L}$ is a subspace of the space $\mathfrak{Z}$, and $\mathfrak{C}$ is a convex cone. Moreover, the intersection $\mathfrak{L} \cap \mathfrak{C}$ is nonempty.

We note that the subspace $\mathfrak{L}$ is defined by the equality

$$\frac{\partial}{\partial z} r(z_0)\bar{z} = 0,$$

and sufficiently small variations $\bar{z}$ satisfying the inequality

$$\frac{\partial}{\partial z} p(z_0)\bar{z} + p(z_0) > 0$$

belong to the cone $\mathfrak{C}$.

An arbitrary $z \in \mathfrak{Z}$, representable in the form $z = z_0 + \bar{z}$, where $\bar{z} \in \mathfrak{L} \cap \mathfrak{C}$, will be called an **admissible tested value of the element** z_0.

Lemma 1. *A necessary condition for a maximum of $\bar{f}(z)$ at the element z_0 among admissible elements belonging to a neighborhood of z_0 is that $f(z)$ have a maximum at the element z_0 among any sufficiently small neighborhood of its admissible tested values.*

Incompatibility. We shall use the notation $\psi' > \psi''$ if $\psi' \geq \psi''$ and $\psi' - \psi''$ does not coincide with the vertex of $\mathfrak{C}$. In the product

$$\sigma = E^1 \times \mathfrak{R} \times \mathfrak{P},$$

where E^1 is the one-dimensional real space, for the admissible element z_0 and for some neighborhood $\mathfrak{Z}^*$ of zero in the space $\mathfrak{Z}$, consider the convex sets:

$$\bar{\sigma} = \{\mathfrak{z}_0, \mathfrak{z}_1, \mathfrak{z}_2 : \mathfrak{z}_0 > 0,\ \mathfrak{z}_1 = 0,\ \mathfrak{z}_2 > 0\},$$

$$\sigma^*(z_0, \mathfrak{Z}^*) = \Big\{\mathfrak{z}_0, \mathfrak{z}_1, \mathfrak{z}_2 : \mathfrak{z}_0 \leq \frac{\partial}{\partial z} f(z_0)\bar{z},\ \mathfrak{z}_1 = \frac{\partial}{\partial z} r(z_0)\bar{z},$$
$$\mathfrak{z}_2 \leq \frac{\partial}{\partial z} p(z_0)\bar{z} + p(z_0),\ \bar{z} \in \mathfrak{Z}^*\Big\}.$$

Lemma 2. *A necessary condition for a maximum of $f(z)$ at the element z_0 among the admissible elements of the problem is the incompatibility condition for the sets $\bar{\sigma}$ and $\sigma^*(z_0, \mathfrak{Z}^*)$, i.e. $\bar{\sigma} \cap \sigma^*(z_0, \mathfrak{Z}^*) = \emptyset$ for some $\mathfrak{Z}^* \subset \mathfrak{Z}$.*

Lemma 3. *There exists a functional $\lambda(\mathfrak{z})$ defined on σ which is strictly positive on $\bar{\sigma}$ and positive on $\sigma^*(z_0, \mathfrak{Z}^*)$.* Sets of type B of linear continuous functionals $\rho = \langle \rho, r \rangle$, and $\pi = \langle \pi, \psi \rangle$, defined respectively on $\mathfrak{R}$ and $\mathfrak{P}$, will be denoted by P and Π.

We shall call the functional π nonnegative if it is nonnegative on $\mathfrak{C}$, and strictly positive if it is positive on the cone $\mathfrak{C}$ except at its vertex.

Theorem 1. *For problem A, a necessary condition for optimality of the element z_0 among some neighborhood is the existence of such linear continuous functionals $\rho \in P$ and $\pi \in \Pi$, of which π is strictly positive, that $\langle \pi, \psi(z_0) \rangle = 0$ and, for all z,*

$$\frac{\partial}{\partial z} \mathcal{F}(z_0, \rho, \pi)\bar{z} = 0, \quad \text{where} \quad \mathcal{F}(z, \rho, \pi) = f(z) + \langle \rho, r(z) \rangle + \langle \pi, p(z) \rangle.$$

Proof. The general expressions for linear continuous functionals defined on σ and strictly positive on $\bar{\sigma}$ and nonnegative on $\sigma^*(z_0, \mathfrak{Z}^*)$ have, respectively, the form:

$$\alpha \mathfrak{z}_0 + \langle \gamma, \mathfrak{z}_1 \rangle + \langle \vartheta, \mathfrak{z}_2 \rangle,$$

$$\beta \mathfrak{z}_0 + \beta \frac{\partial}{\partial z} f(z_0)\bar{z} + \left\langle \mu, \mathfrak{z}_1 - \frac{\partial r}{\partial z}(z_0)\bar{z} \right\rangle + \left\langle \delta, \mathfrak{z}_2 - \frac{\partial}{\partial z} p(z_0)\bar{z} - p(z_0) \right\rangle,$$

where $\alpha > 0$, $\beta \geq 0$, and $\gamma, \mu, \vartheta, \delta$ are linear continuous functionals, of which ϑ is strictly positive and δ is positive.

Thus, the functional $\lambda(\mathfrak{z})$, which exists according to Lemma 3, has two representations. Consequently, successively putting $\bar{z} = 0$, $\mathfrak{z}_0 = 0$, $\mathfrak{z}_1 = 0$, $\mathfrak{z}_2 = 0$, we obtain:

$$\alpha = \beta > 0, \qquad \gamma = \mu, \qquad \vartheta = \delta.$$

Introducing the notation $\mu/\alpha = \rho$, $\delta/\alpha = \pi$, we obtain what was required by the theorem.

II. Consider the problem of the theory of optimal processes to which the following reduce: by expanding the vector of phase coordinates–a problem with parameters–and by the substitution $t = t_1 + \tau(t_2 - t_1)$–a problem with movable endpoints.

Problem B. On the interval $[t_1, t_2]$, determine vector-functions $x(t)$ and $u(t)$ that provide

$$\max\{G : \ dx/dt = f, \ g = 0, \ h \geq 0\},$$

where G is a functional of $x(t_1), x(t_2)$; f, g, h are vector-functions of $x(t), u(t), t, x(t_1)$,

$x(t_2)$. We shall consider the regular case. It is characterized by the fact that, for any finite $x(t)$, t, ξ, η, the set $\{u:\ g=0,\ h\geqslant 0\}$ is bounded.

We shall seek solutions of problem B in the class of absolutely continuous $x(t)$ and bounded measurable $u(t)$.

Let us make the assumptions:

B_1. An optimal solution $x^0(t)$, $u^0(t)$ of problem B exists.

B_2. The functional $G(\xi,\eta)$ is continuously differentiable in a neighborhood of $\xi^0=x^0(t_1)$, $\eta^0=x^0(t_2)$.

B_3. The functions f, g, h, regarded as functions of x, u, t, ξ, η, for any absolutely continuous $x(t)$, bounded measurable $u(t)$, and any ξ and η taken, respectively, from neighborhoods of $x^0(t)$, $u^0(t)$, $x^0(t_1)$, $x^0(t_2)$, together with their first-order derivatives with respect to x, u, ξ, η, are summable with respect to t on $[t_1,t_2]$.

B_4. The conditions $g=0$ and $h\geqslant 0$ ensure, in a neighborhood of the optimal solution, a nonempty set of admissible solutions.

Theorem 2. *For problem B, under assumptions B_1–B_4, there exist vector functions $\psi(t)$, $\omega(t)$, and $\varepsilon(t)$ defined on $[t_1,t_2]$ (of which $\psi(t)$ is absolutely continuous, while $\omega(t)$ and $\varepsilon(t)$ are measurable and bounded almost everywhere), satisfying, at the optimal values $x(t)$, $u(t)$, the conditions*

$$\frac{d\psi}{dt}+\psi f'_x+\omega g'_x+\varepsilon h'_x=0; \tag{1}$$

$$\psi f'_u+\omega g'_u+\varepsilon h'_u=0; \tag{2}$$

$$G'_\xi+\psi(t_1)+\int_{t_1}^{t_2}\left(\psi f'_\xi+\omega g'_\xi+\varepsilon h'_\xi\right)dt=0; \tag{3}$$

$$G'_\eta-\psi(t_2)+\int_{t_1}^{t_2}\left(\psi f'_\eta+\omega g'_\eta\varepsilon h'_\eta\right)dt=0; \tag{4}$$

$$\varepsilon(t)h=0; \tag{5}$$

$$\varepsilon(t)\geqslant 0, \tag{6}$$

in which the expressions (1), (2), (5), and (6) are fulfilled almost everywhere on $[t_1,t_2]$; $\xi=x(t_1)$, $\eta=x(t_2)$. The vector quantities x, u, t, g, h are written as columns, and $\psi(t)$, $\omega(t)$, $\varepsilon(t)$ as rows. The rule for differentiating a vector with respect to a vector is the usual one.

Proof. On the completed space $\mathfrak{Z}$ of sets $z = (x(t), u(t))$, consider $\mathfrak{f}(z) = G$; $\mathfrak{t}(z) = (\mathfrak{t}_1(z), \mathfrak{t}_2(z))$, where $\mathfrak{t}_1(z) = f - dx/dt$, $\mathfrak{t}_2(z) = g$, and $\mathfrak{p}(z) = h$. Taking into account B_1–B_4 and the fact that $\mathfrak{t}_1(z)$, $\mathfrak{t}_2(z)$, $\mathfrak{p}(z)$ are expressed through summable functions, we note that the spaces conjugate to them are the spaces of integral functionals $\rho = (\rho_1, \rho_2)$ and π with measurable almost everywhere bounded kernels $\psi(t)$, $\omega(t)$, and $\varepsilon(t)$ (see (6), p. 190).

Carrying out the variation of the generalized Lagrange function with respect to $x(t)$, $u(t)$, $x(t_1)$, $x(t_2)$, on the basis of theorem 1 we obtain equations (1)–(4), fulfilled in a generalized sense. At the same time, equation (1) implies the absolute continuity of $\psi(t)$.

The expressions (5), (6), describing the properties of the function, also follow from the assertions of theorem 1, as was required to prove.

From the theorem proved there follows the existence of adjoint functions and, in particular, of the continuous function $\psi(t)$, first, outside the conception of jump conditions and, second, both with special consideration and without special consideration of the unconditional boundary, consisting in replacing explicit or implicit constraints of the form $l(x, t) \geqslant 0$, which in the case $l(x(t), t) = 0$ form an unconditional boundary, by the constraint of the form

$$\frac{\partial l}{\partial x} f + \frac{\partial l}{\partial t} \geqslant 0$$

(see ([1,2]);

III. Let us formulate the content of Theorem 2 in the form of a certain geometric principle.

In the space Z of column vectors z, composed of the coordinates of the vectors x and u corresponding to an admissible solution $x(t)$ and $u(t)$ for some $t \in [t_1, t_2]$, consider the sets $L = \{z : g = 0\}$ and $H = \{z : h \geqslant 0\}$. The aggregate of normals to the hypersurfaces determined by the rows of the vector conditions $g = 0$ and $h = 0$ will be written in the form of matrix derivatives $-\partial g/\partial z$, $-\partial h/\partial z$. Here all normals $-\partial h/\partial z$ are directed outward from H. In Z consider the convex cone

$$k(t) = \{-\varepsilon\, \partial h/\partial z - \omega\, \partial g/\partial z : \varepsilon h = 0,\ \varepsilon \geqslant 0\}$$

and the vector

$$\varphi(t) = \{\partial \psi/dt + \psi\, \partial f/\partial x : \varphi\, \partial f/\partial u\},$$

which we shall call, respectively, the local adjoint cone and the local optimizing vector for the admissible solution at the point z. Analogously, in the space Y of column vector-functions y, composed of the coordinates of the vector-functions $x(t)$ and $u(t)$, consider the general adjoint cone

$$k^* = \left\{ -\int_{t_1}^{t_2} \left(\varepsilon \frac{\partial h}{\partial y} + \omega \frac{\partial g}{\partial y} \right) dt : \varepsilon(t) h = 0,\ \varepsilon(t) \geqslant 0 \right\}$$

and the general optimizing vector

$$\chi = \left\{ \frac{\partial G}{\partial \xi} + \psi(t_1) + \int_{t_1}^{t_2} \psi \frac{\partial f}{\partial \xi}\, dt;\ \frac{\partial G}{\partial \eta} - \psi(t_2) + \int_{t_1}^{t_2} \psi \frac{\partial f}{\partial \eta}\, dt \right\}_{\substack{\xi = x(t_1),\\ \eta = x(t_2)}}.$$

Then the principle of optimality is formulated as follows:

Theorem 3. *Under the adopted assumptions, the optimal solution has the property that, for almost all* $t \in [t_1, t_2]$, $\varphi(t) \in K(t)$ *and* $\chi \in K^*$.

IV. **Sufficiency of the principle of optimality.** Let us additionally assume:

$\mathbf{B}_5$**.** The functions G, f, g, h are polynomial expressions of second order with respect to $x(t), u(t), x(t_1), x(t_2)$, respectively. Moreover, for any absolutely continuous $x(t)$ and measurable bounded $u(t)$, the second derivatives of f, g, h are measurable and almost everywhere bounded in t on $[t_1, t_2]$.

Theorem 4. *Let conditions* B_1–B_5 *be fulfilled. Then any admissible solution* $x(t), u(t)$ *of problem* B *will be optimal if it, together with some aggregate of vector-functions* $\psi(t), \omega(t)$, *and* $\varepsilon(t)$, *satisfies the principle of optimality and the condition*

$$\zeta = D^2_{\xi,\eta} G + \int_{t_1}^{t_2} D^2_{x,y,\xi,\eta} [\psi f + \omega g + \varepsilon h]\, dt \leqslant 0,$$

in which $\xi = x(t_1)$, $\eta = x(t_2)$, *and* $D^2_{x,y,\ldots,z}$ *denotes the superposition of the operators*

$$D_{x,y,\ldots,z} = \frac{\partial}{\partial x} \Delta x + \frac{\partial}{\partial y} \Delta y + \cdots + \frac{\partial}{\partial z} \Delta z.$$

Corollary 1. *In the case when* G, f, g, h *are linear in* $x(t), u(t), x(t_1), x(t_2)$, *the condition* $\zeta \leqslant 0$ *is always fulfilled.*

Corollary 2. *If* f *and* g *are linear functions of* $x(t), u(t), x(t_1), x(t_2)$, *then, for the condition* $\zeta \leqslant 0$ *to be fulfilled, it is sufficient that the functions* G *and* h *be concave downward on any subset.*

Moscow Institute of Physics and Technology

Received
14 I 1969

REFERENCES

1. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, 1961.

2. V. P. Anorov, *Automation and Telemechanics*, No. 3 and 4 (1967).

3. A. Ya. Dubovitskii, A. A. Milyutin, DAN, **149**, No. 4 (1963).

4. A. Ya. Dubovitskii, A. A. Milyutin, *Zhurn. Vychisl. Mat. i Mat. Fiz.*, **5**, No. 3 (1965).

5. P. Levy, *Concrete Problems of Functional Analysis*, "Nauka," 1967.

6. L. V. Kantorovich, G. P. Akilov, *Functional Analysis in Normed Spaces*, 1959.

Note: Figure translations are in progress. See original paper for figures.