# ON THE CONVERGENCE OF MINIMIZING SEQUENCES IN CONSTRAINED EXTREMUM PROBLEMS

MATHEMATICS

1966

SovietRxiv

## Abstract

## Full Text

UDC 513.88 : 519.3

*MATHEMATICS*

E. S. LEVITIN, B. T. POLYAK

# ON THE CONVERGENCE OF MINIMIZING SEQUENCES IN CONSTRAINED EXTREMUM PROBLEMS

*(Presented by Academician L. V. Kantorovich, 1 X 1965)*

**1°.** Let $Q$ be a subset of a Banach space $E$, and let $f(x)$ be a functional defined on $E$. A sequence $x^n \in E$ will be called a **generalized minimizing sequence** (g.m.s.) for $f(x)$ on $Q$ if:

$$\text{a) } \rho(x^n, Q) = \inf_{x\in Q} \|x^n - x\| \xrightarrow[n\to\infty]{} 0; \qquad \text{b) } f(x^n) \xrightarrow[n\to\infty]{} f^* = \inf_{x\in Q} f(x).$$

Such an extension of the notion of a minimizing sequence in comparison with the usual definition (when all $x^n \in Q$) is natural in constrained extremum problems, since in many minimization methods the successive approximations, generally speaking, do not satisfy the constraints. The following assertions concerning g.m.s. $x^n$ for $f(x)$ on $Q$ are obvious. If $f(x)$ is continuous, $Q$ is closed, and $x^n$ converges to $x^*$, then $x^* \in Q$ and $f(x^*) = f^*$. If, moreover, $f(x)$ is weakly lower semicontinuous, $Q$ is convex, and $x^n$ converges weakly to $x^*$, then $x^* \in Q$ and $f(x^*) = f^*$.

Below we give theorems guaranteeing strong convergence of a g.m.s. These results are adjacent to the work ($^1$), whose terminology we use.

**Theorem 1.** *Let $Q$ be a closed convex subset of a reflexive space $E$; let $f(x)$ be a uniformly quasiconvex lower semicontinuous functional, bounded below on $Q$, and suppose that either $Q$ is bounded or the sets $\{x : f(x) \le \lambda\}$ are bounded for all $\lambda$. Then every g.m.s. converges to the (unique) minimum point of $f(x)$ on $Q$.*

Let us note an important special case of Theorem 1. We shall call a functional **uniformly convex** if there exists a function $\delta(\tau)$, $\delta(\tau) > 0$ for $\tau > 0$ (which may be assumed monotone), such that

$$f\left(\frac{x+y}{2}\right) \le \frac{1}{2}f(x) + \frac{1}{2}f(y) - \delta(\|x-y\|)$$

for all $x, y$. For a differentiable $f(x)$ this condition is equivalent to the following: for every $x \in E$ there exists $c \in E^*$ such that $f(x+y) \geq f(x)+(c,y)+\delta(\|y\|)$ for any $y$. It turns out that such a functional grows no more slowly than a quadratic one: $f(x) \geq a\|x\|^2$, $a > 0$, for $\|x\| \geq R$. Consequently, $f(x)$ is bounded below on every set, and $\{x : f(x) \leq \lambda\}$ is bounded for any $\lambda$. Therefore, in this case Theorem 1 is applicable.

**Theorem 2.** *If $Q$ is a closed uniformly convex set and $f(x)$ is a weakly lower semicontinuous functional attaining a unique minimum on $Q$ at the boundary point $x^*$, then every g.m.s. converges to $x^*$.*

**Remark 1.** The conditions of Theorem 2 are certainly satisfied if $f(x)$ is a convex continuous functional all of whose supporting functionals on $Q$ are nonzero (in particular, if $f(x)$ is a nonzero linear functional).

**Remark 2.** It can be shown that every uniformly convex set distinct from $E$ is bounded.

From Theorem 2 the following result can be obtained, not directly connected with extremal problems.

**Theorem 3.** *Let $Q$ be a closed uniformly convex set, let $x^n$ converge weakly to $x^*$, and let*

$$\lim_{n\to\infty} \rho(x^n, Q) = 0.$$

*Then $x^* \in Q$, and if $x^*$ is a boundary point of $Q$, then*

$$\lim_{n\to\infty} \|x^n - x^*\| = 0.$$

**Corollary.** *If $E$ is a uniformly convex space, $x^n$ converges weakly to $x^*$, and*

$$\lim_{n\to\infty} \|x^n\| \leq \|x^*\|,$$

*then*

$$\lim_{n\to\infty} \|x^n - x^*\| = 0.$$

**Remark.** The condition of uniform convexity of $Q$ in Theorem 3 can be replaced by one or another local condition. For example: for every $x \in Q$ the ball

$$(x^* + x)/2 + z \in Q$$

whenever

$$\|z\| \leq \delta(\|x - x^*\|),$$

or: there exists a $c \in E^*$ such that

$$(c, x - x^*) \geq \delta(\|x - x^*\|), \quad x \in Q.$$

The set $U$ considered below in Theorem 4 often occurs in problems of optimal control. It is not uniformly convex (it does not even have interior points);

nevertheless, it possesses a number of analogous properties. Let us call a point $z^* \in M$ **regular** for $M \subset E^r$ if there exists $c \in E^r$ such that

$$(c, z^*) < (c, z)$$

for all $z \in M$. For example, if $M$ is a polyhedron, then its vertices are regular points; if $M$ is strictly convex, then all boundary points are regular.

**Theorem 4.** *Let*

$$U = \{u(t) \in L_2^r(0,T) : u(t) \in M(t) \text{ for almost all } t \in (0,T)\},$$

*where $M(t)$ is given by a function $\varphi(u,t)$,*

$$M(t) = \{u \in E^r : \varphi(u,t) \le 0\},$$

*continuous and convex in $u$ and measurable in $t$, and*

$$u\,\varphi(u,t) \to \infty \quad \text{as } \|u\|_{E^r} \to \infty$$

*uniformly in $t$. Let $u^*(t) \in U$, and let $u^*(t)$ be a regular point of $M(t)$ for almost all $t$. Then every sequence $u^n(t)$, weakly converging to $u^*(t)$ and such that*

$$\lim_{n\to\infty} \rho(u^n, U) = 0,$$

*converges to $u^*(t)$ in $L_2^r$.*

**2°.** Let us now consider applications of the theorems proved to linear problems of optimal control. It is required to minimize the functional

$$f(u) = \int_0^T \varphi(x(t), u(t), t)\,dt + \Phi(x(T)), \tag{1}$$

where

$$x(t) = (x_1(t), \dots, x_n(t)), \quad u(t) = (u_1(t), \dots, u_r(t)), \quad r \le n,$$

$$dx(t)/dt = A(t)x(t) + B(t)u(t), \quad x(0) = x^0. \tag{2}$$

Here $A(t), B(t)$ are matrices depending continuously on $t$, respectively $n \times n$ and $n \times r$. The solution is sought in the class of controls $u(t) \in L_2^r(0,T)$ under the presence of some constraints on $u(t)$. Below three special cases of this problem are presented.

**2.1.** Let the constraints be given by any set of conditions of the following form:

a) $u(t) \in U$, where $U$ is defined in Theorem 4;

b) $x(t) \in D(t)$ for all $t \in [0,T]$, where $D(t)$ for every $t$ is a closed convex set in $E^n$;

c)

$$\int_0^T F(u(t),t)\,dt \le 0,$$

where $F(u,t)$ is continuous and convex in $u$ and measurable in $t$.

Then they define in $L_2^r$ a closed convex set $Q$. We shall assume that it is nonempty. Further, let for any $x^1, x^2 \in E^n$, $u^1, u^2 \in E^r$, $t \in [0,T]$,

$$\varphi\left(\frac{x^1+x^2}{2}, \frac{u^1+u^2}{2}, t\right) \le \frac{1}{2}\varphi(x^1,u^1,t) + \frac{1}{2}\varphi(x^2,u^2,t) - \delta(\|u^1-u^2\|_{E^r}), \quad (3)$$

where $\delta(\tau) > 0$ for $\tau > 0$, and $\Phi(x)$ is a continuous convex function in $E^n$.

Then $f(u)$ is a uniformly convex functional, all the conditions of Theorem 1 are satisfied, and in this case any m.s. converges to the unique solution of the problem.

**Remark 1.** We have given only the simplest condition (3) of uniform convexity of $f(u)$. More precise conditions can be formulated analogously to how this was done in ($^1$), Theorem 10, for the simplest variational problem.

**Remark 2.** If $Q = Q_1 \cap Q_2$, $Q_1, Q_2$ are closed convex sets and

$$\lim_{n\to\infty} \rho(x^n, Q_i) = 0, \qquad i = 1,2,$$

then, generally speaking, it is not necessary that

$$\lim_{n\to\infty} \rho(x^n, Q) = 0.$$

However, this is certainly true if $Q_1$ (or $Q_2$) is uniformly convex, or if $Q_1^0 \cap Q_2 \ne \emptyset$ and $Q$ is bounded. This should be kept in mind when checking condition a) in the definition of an m.s., if the admissible set is given by a collection of constraints.

**2.2.** Let all constraints have the form

$$\int_0^T F_i(u(t),t)\,dt \le 0, \qquad i = 1, \dots, k.$$

where $F_i(u,t)$ are measurable in $t$, continuous and uniformly convex in $u$ for each $t$ (with functions $\delta_i(\tau)$ independent of $t$). The set $Q$ determined by such constraints is uniformly convex and closed in $L_2^r$. Suppose it is nonempty (for example, if $F_i(0,t) \le 0$, then $u(t) \equiv 0 \in Q$). Suppose $\varphi(x,u,t)$ is continuous in $\{x,u\}$, convex in $u$, and measurable in $t$, and $\Phi(x)$ is continuous. It can be shown that in this case $f(u)$ is weakly lower semicontinuous. We now give a condition guaranteeing that the optimal control $u^*(t)$ lies on the boundary of $Q$.

Let us note that, under additional assumptions on the smoothness of $\varphi$ and $\Phi$, the functional $f(u)$ is differentiable, and its gradient is $f'(u) = \varphi_u - B^*\psi$, where

$\psi(t)$ is the solution of the system $d\psi/dt = -A^*\psi + \varphi_x,\ \psi(T) = -\Phi'(x(T))$. We shall say that system (2) is **nondegenerate** if, for any nonzero solution $p(t)$ of the system $dp/dt = -A^*p$, the function $B^*p$ does not vanish on any interval from $(0, T)$. It turns out that, if the system is nondegenerate and along the optimal trajectory any one of the conditions is fulfilled: a) $\varphi_u \equiv 0,\ \Phi'(x(T)) \neq 0$, b) $\varphi_u \equiv 0,\ \varphi_x \neq 0$ on no interval, then $f'(u^*) \neq 0$. Therefore $u^*(t)$ cannot be an interior point of $Q$. Thus, if all these conditions are satisfied and $u^*(t)$ is unique, then Theorem 2 is applicable and any m.s. $u^n(t)$ converges to $u^*(t)$ in $L_2^r$. If, in particular, $\varphi(x, u, t)$ is convex in $\{x, u\}$ and $\Phi(x)$ is convex, then $f(u)$ is convex, and the assumption of uniqueness of $u^*(t)$ is certainly fulfilled.

**2.3.** Let the only constraint have the form $u(t) \in U$, where $U$ is defined in Theorem 4. Suppose that, for any nonzero solution $\psi(t)$ of the system $d\psi/dt = -A^*\psi$, the function $(B^*\psi, u)$ attains a unique maximum with respect to $u$ on $M(t)$ for almost all $t \in (0, T)$. This condition is satisfied, for example, if $M(t)$ is strictly convex for all $t$, and system (2) is nondegenerate, or if $M(t) \equiv M$ is a polyhedron and the general-position condition (2) is fulfilled. Finally, suppose that $\varphi \equiv 0$, and that $\Phi(x)$ is convex and continuous, with $\Phi'(x(T)) \neq 0$ for all attainable $x(T)$. Then, using the maximum principle (2) and applying Theorem 4, we obtain that every m.s. $u^n(t)$ converges to the unique solution $u^*(t)$ in $L_2^r$.

**3°.** Of course, in the general case any minimizing sequence need not converge strongly (for example, in the problem of minimizing

$$\int_0^1 x^2(t)\,dt$$

or $x^2(1)$, where $x(0) = 0,\ dx/dt = u$, $u(t)$ satisfies the constraints

$$|u(t)| \leq 1$$

or

$$\int_0^1 u^2(t)\,dt \leq 1,$$

the sequence $u^n(t) = \sin nt$ is minimizing, but converges only weakly to the solution $u^*(t) \equiv 0$. However, there exists a very general device that makes it possible to obtain strong

a convergent minimizing sequence. The idea of this device (regularization) is due to A. N. Tikhonov (3). Our method of constructing the regularizing functional differs somewhat from that proposed by A. N. Tikhonov ($g(x)$ is continuous in the metric of the original space; the set $\{g(x) \leq \lambda\}$ is not assumed to be compact).

Let $E$ be a reflexive Banach space, $f(x)$ a lower semicontinuous quasiconvex functional, and $Q$ a closed bounded convex set.

**Theorem 5.** *If $g(x)$ is a nonnegative uniformly convex functional, then for every $a_n > 0$ there exists $x^n$, a point of minimum of $f(x) + a_n g(x)$ on $Q$;*

*moreover, as* $a_n \to +0$*, the sequence* $x^n$ *is minimizing and converges to* $x^*$*, a point of minimum of* $f(x)$ *on* $Q$ *(that one of them for which* $g(x)$ *is minimal).*

**Proof.** Since $f(x) + a_n g(x)$ is a strictly quasiconvex functional, it attains a unique minimum on $Q$ at the point $x^n$ (([1]), Theorem 2). Let $\tilde{x}$ be an arbitrary point of minimum of $f(x)$ on $Q$; then

$$f(x^n) + a_n g(x^n) \leq f(\tilde{x}) + a_n g(\tilde{x}) \leq f(x^n) + a_n g(\tilde{x}),$$

i.e. $g(x^n) \leq g(\tilde{x})$. Further,

$$f(\tilde{x}) \leq f(x^n) \leq f(\tilde{x}) + a_n \left[g(\tilde{x}) - g(x^n)\right] \leq f(\tilde{x}) + a_n g(\tilde{x}).$$

Hence $f(x^n) \to \inf f(x)$. From $\{x^n\}$ one can choose a subsequence weakly converging to some $x^*$, and it must be that

$$f(x^*) = \inf_{x \in Q} f(x).$$

But since $g(x^n) \leq g(\tilde{x})$, it follows that $g(x^*) \leq g(\tilde{x})$ for all $\tilde{x}$. Such a point $x^*$ (by the strict convexity of $g(x)$ and the convexity of $\{\tilde{x}\}$) is unique; therefore the whole sequence $x^n$ converges weakly to $x^*$. By the uniform convexity of $g(x)$,

$$\delta(\|x^n - x^*\|) \leq \frac{1}{2} g(x^n) + \frac{1}{2} g(x^*) - g\left(\frac{x^n + x^*}{2}\right) \leq g(x^*) - g\left(\frac{x^n + x^*}{2}\right),$$

but from the weak lower semicontinuity of $g(x)$,

$$\lim_{n \to \infty} g\left(\frac{x^n + x^*}{2}\right) \geq g(x^*),$$

i.e.

$$\delta(\|x^n - x^*\|) \to 0, \qquad \text{whence also} \qquad \|x^n - x^*\| \to 0.$$

In optimal-control problems, as a regularizing functional one may take, for example,

$$\int_0^T \|u(t)\|_{E'}^2 \, dt.$$

Moscow State University
named after M. V. Lomonosov

Received
25 IX 1965

## CITED LITERATURE

[1] B. T. Polyak, *Dokl. Akad. Nauk SSSR*, **166**, No. 2 (1966).
[2] L. V. Pontryagin, V. G. Boltyansky, R. V. Gamkrelidze, E. F. Mishchenko,

*The Mathematical Theory of Optimal Processes*, Moscow, 1961.
[3] A. N. Tikhonov, *Dokl. Akad. Nauk SSSR*, **162**, No. 4 (1965).

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

*Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.*