# MATHEMATICS

1965

SovietRxiv

## Abstract

## Full Text

# MATHEMATICS

**V. F. Dem' yanov, A. M. Rubinov**

# ON THE PROBLEM OF MINIMIZING A SMOOTH FUNCTIONAL UNDER CONVEX CONSTRAINTS

*(Presented by Academician V. I. Smirnov on 25 VI 1964)*

Consider a Banach space $X$ in which a convex closed bounded set $\Omega$ is specified. Suppose that on $\Omega$ there is given a smooth functional $f$. We are interested in the question of finding points at which the functional $f$ attains a minimum on $\Omega$ (at least a local one). Let us first note that if $f$ has gradient $F$, then for $x, \tilde{x} \in X$ the relation

$$f(\tilde{x}+\alpha(x-\tilde{x})) = f(\tilde{x}) + \alpha(x-\tilde{x}, F\tilde{x}) + o(\alpha). \tag{1}$$

holds. If, moreover, the operator $F$ has a derivative $F'$, then the formulas

$$f(\tilde{x}+\alpha(x-\tilde{x})) = f(\tilde{x}) + \alpha(x-\tilde{x}, F\tilde{x}) + \frac{\alpha^2}{2}(x-\tilde{x}, F'\tilde{x}(x-\tilde{x})) + o(\alpha^2); \tag{2}$$

$$f(\tilde{x}+\alpha(x-\tilde{x})) = f(\tilde{x})+\alpha(x-\tilde{x}, F\tilde{x})+\frac{\alpha^2}{2}(x-\tilde{x}, F'(\tilde{x}+\theta(x-\tilde{x}))(x-\tilde{x})), \quad 0 \leqslant \theta \leqslant 1. \tag{3}$$

are valid.

We now formulate a necessary condition for a minimum.

**Theorem 1.** *Let the functional $f$ have gradient $F$. In order that a local minimum be attained at a point $y \in \Omega$, it is necessary that*

$$\min_{x\in\Omega}(x-y, Fy) = 0. \tag{4}$$

**Proof.** From formula (1) it follows that for $x \in \Omega$

$$f(y+\alpha(x-y)) - f(y) = \alpha(x-y, Fy) + o(\alpha). \tag{5}$$

For small $\alpha \in [0,1]$ the left-hand side of (5) is nonnegative. Since the sign of the right-hand side for sufficiently small $\alpha$ is determined by the first term, $(x-y, Fy) \geqslant 0$. But for $x = y$ we have $(x-y, Fy) = 0$, whence the validity of the theorem follows.

We shall call a point $y \in \Omega$ **stationary** if (4) is satisfied for it. If the stationary point $y$ is an interior point of $\Omega$, then it is easy to show that $Fy = 0$; in other words, in this case the point $y$ is a critical point of the functional $f$.

Suppose now that $y$ is a stationary point, but $(x-y, Fy) \neq 0$ for $x \in \Omega$. In this case it follows from (5) that at the point $y$ there cannot be a local maximum, and the functional either attains a minimum at it or has no extremum. Let us note a sufficient condition for a local minimum.

**Theorem 2.** *Let $f$ have differentiable gradient $F$, and let $y$ be a stationary point. Then, if the operator $F'y$ is positive definite (in the sense that for $x \in \Omega$, $(x-y, F'y(x-y)) \geqslant m\|x-y\|^2$), then a local minimum is attained at the point $y$.*

**Proof.** From (2) it follows that

$$f(y+\alpha(x-y)) - f(y) =$$

$$= \alpha(x-y, Fy) + {}^1\!/_2\alpha^2(x-y, F'y(x-y)) + o(\alpha^2).$$

Since $(x-y, Fy) \geqslant 0$, we have

$$f(y+\alpha(x-y)) - f(y) \geqslant \frac{1}{2}\alpha^2 m\|x-y\|^2 + o(\alpha^2)$$

and, consequently, for sufficiently small $\alpha$,

$$f(y+\alpha(x-y)) > f(y)$$

for all $x \in \Omega$.

**Theorem 3.** *Let $f$ have a differentiable gradient $F$, and let $y$ be a stationary point. If, in the intersection of some neighborhood of the point $y$ with $\Omega$, the functional $f$ is convex, then a local minimum is attained at the point $y$.*

**Proof.** Setting in (3) $\alpha = 1$, $\tilde{x} = y$, we have

$$f(x)-f(y) = (x-y, Fy)+\frac{1}{2}(x-y, F'(y+\Theta(x-y))(x-y)), \qquad 0 \leqslant \Theta \leqslant 1. \ (6)$$

It is easy to show that, owing to the stationarity of $y$ and the convexity of the functional $f$, the right-hand side of (6) is positive, which proves the theorem.

Now let $x \in \Omega$. Denote by $\bar{x}$ one of the elements at which

$$\min_{z\in\Omega}(z, Fx)$$

is attained. In the new notation, stationarity of the point $y$ means that $y$ is a solution of the equation

$$(\bar{x} - x, Fx) = 0. \tag{7}$$

We shall give a method of successive approximations for solving equation (7).

Take an arbitrary element $x_0 \in \Omega$, find $\bar{x}_0$, and, for $\alpha \in [0, 1]$, consider the function $g_0(\alpha) = f(x_0 + \alpha(\bar{x}_0 - x_0))$. Let this function attain its minimum on $[0, 1]$ at the point $\alpha_0$. Put $x_1 = x_0 + \alpha_0(\bar{x}_0 - x_0)$. It is clear that $f(x_1) = g_0(\alpha) \leqslant g_0(0) = f(x_0)$. Starting from the point $x_1$, in the same way we construct $x_2$, and so on. As a result we have constructed the sequences

$$\begin{gathered} x_0,\ x_1,\ x_2, \dots, \\ \bar{x}_0,\ \bar{x}_1,\ \bar{x}_2, \dots \end{gathered} \tag{8}$$

Moreover, $f(x_0) \geqslant f(x) \geqslant \dots$.

**Theorem 4.** *Let the differentiable functional $f$ be bounded below on the compact set $\Omega$. Then*

$$\lim(\bar{x}_n - x_n, Fx_n) = 0.$$

**Corollary 1.** If $\Omega$ is compact, then all limit points of the sequence (8) are stationary.

**Corollary 2.** If $\Omega$ is weakly compact, and $F$ is a completely continuous operator, then all limit points (in the weak topology) of the sequence (8) are stationary.

Suppose now that $f$ is a convex functional. Applying Lagrange' s formula, we have

$$f(x) - f(x_n) = (x - x_n, F(x_n + \Theta(x - x_n))), \qquad 0 \leqslant \Theta \leqslant 1.$$

From the convexity of the functional $f$ it follows, as is easy to show, that the function

$$s(\alpha) = (x - x_n, F(x_n + \alpha(x - x_n)))$$

is nondecreasing; hence

$$(x-x_n, F(x_n+\Theta(x-x_n))) \geqslant (x-x_n, Fx_n)$$

and, therefore,

$$f(x)-f(x_n) \geqslant (x-x_n, Fx_n).$$

Let $y \in \Omega$ be an element at which $f$ attains a global minimum on $\Omega$. From the preceding inequality it follows that

$$f(y)-f(x_n) = \min_{x\in\Omega}(f(x)-f(x_n)) \geqslant \min_{x\in\Omega}(x-x_n, Fx_n) = (\bar{x}_n - x_n, Fx_n),$$

whence

$$0 \leqslant f(x_n)-f(y) \leqslant (x_n-\bar{x}_n, Fx_n). \tag{9}$$

From Theorem 4 we now obtain that

$$f(x_n) \to f(y) = \min_{x\in\Omega} f(x).$$

Inequalities (9) give an a posteriori estimate of convergence.

As we learned from conversations with V. V. Khomenyuk, these inequalities had been obtained by him earlier from other considerations.

Let us note that if $f$ is a quadratic functional in a Hilbert space $H$:

$$f(x) = (Ax, x) + (b, x),$$

where $A$ is a positive definite operator, $b \in H$, then it can be shown that

$$y = \lim x_n.$$

The method set forth can be used to solve a number of practically important problems, in particular:

1. Problems of convex programming. We note that if the constraints in a convex programming problem are linear, then the algorithm obtained for this case is close in idea to one of G. Zoutendijk' s algorithms (1).
2. Problems of optimal control of automatic control systems. For the case of linear systems and quadratic functionals, the method is set forth in (2,3).

Leningrad State
University named after A. A. Zhdanov

Received
19 VI 1964

## CITED LITERATURE

[1] G. Zoutendijk, *Methods of Feasible Directions*, Moscow, 1963. [2] V. F. Demyanov, *Prikl. matem. i mekh.*, 27, issue 3 (1963). [3] V. F. Demyanov, *Avtomatika i telemekh.*, 25, No. 1 (1964).

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