ON ONE APPROACH TO MINIMIZATION PROBLEMS

MATHEMATICS

1969

SovietRxiv

Abstract

Full Text

UDC 519.3

MATHEMATICS

Yu. M. DANILIN

ON ONE APPROACH TO MINIMIZATION PROBLEMS

(Presented by Academician A. N. Tikhonov on 18 III 1969)

Suppose it is required to find the minimum value of a nonlinear functional $f(x)$ on a closed convex bounded set of a reflexive Banach space. In what follows, when speaking of the n-th derivative of a functional, we shall mean the strong derivative ([1]).

Theorem 1. Let Q be a closed convex bounded set of a reflexive Banach space E; let $f(x)$ be a functional n times differentiable on Q, and suppose that

$$\|f^{(n)}(x) - f^{(n)}(y)\| \leq R\|x - y\|, \qquad x, y \in Q. \tag{1}$$

Suppose, furthermore, that the functional

$$f_k(x) = f'(x^k)(x - x^k) + \frac{1}{2}f''(x^k)(x - x^k)^2 + \dots + \frac{1}{n!}f^{(n)}(x^k)(x - x^k)^n$$

is convex for every $k \geq 0$; $\bar{x}^k$ is a point of minimum of $f_k(x)$ on Q; $\alpha_k = \beta_k\gamma_k$, where

$$\beta_k = \min\left\{1, \left[\frac{-f_k(\bar{x}^k)}{\|\bar{x}^k - x^k\|^{n+1}}\right]^{1/n}\right\},$$

and $0 < \gamma \leq \gamma_k \leq \frac{1}{\beta_k}$ is chosen from the condition

$$f(x^{k+1}) - f(x^k) \leq \varepsilon\beta_k\gamma_k f_k(\bar{x}^k), \qquad 0 < \varepsilon < 1.$$

Then, if

$$x^{k+1} = x^k + \alpha_k(\bar{x}^k - x^k), \tag{2}$$

then: 1) $f(x^{k+1}) \leq f(x^k)$; 2) $\lim_{k\to\infty} f_k(\bar{x}^k) = 0$; 3) if $f(x)$ is convex, then

$$\lim_{k\to\infty} f(x^k) = \inf_{x\in Q} f(x) = f(x^*).$$

Remark. For $n > 1$, instead of condition (1) it is sufficient to restrict oneself to the requirement $\|f^{(n)}(x)\| \leq M,\ x \in Q$; in this case

$$\beta_k = \min\left\{1, \left[\frac{-f_k(\bar{x}^k)}{\|\bar{x}^k - x^k\|^n}\right]^{1/(n-1)}\right\}.$$

It is not difficult to see that many known minimization methods (for example, gradient methods in unconstrained problems, the conditional-gradient method, Newton' s method, method ([2])) follow directly from Theorem 1. The approach proposed in Theorem 1 can also be used to justify methods for minimizing a functional $f(x, y)$ under the condition $P(x, y) = 0$ (see Theorem 3 below).

We note that, in contrast to previously used algorithms for choosing the parameter α_k (see ([2–4])), the algorithm proposed in Theorem 1 is not connected with computing the minimum of the function $f(\alpha)$ in the direction of motion and does not require knowledge of constants characterizing the functional being minimized. As a consequence, it makes it possible to avoid complications arising from an inaccurate determination of the point of minimum of $f(\alpha)$ and to reduce the time spent on solving the problem. In second-order methods (Theorem 2) the proposed algorithm for choosing the step length makes it possible, moreover, to obtain a higher estimate of the rate of convergence in comparison with the known ones ([2]).

Theorem 2. Let Q be a closed convex bounded set of a Hilbert space E; let $f(x)$ be a twice differentiable functional on Q, and suppose that

$$m\|y\|^2 \leq (f''(x), y, y) \leq M\|y\|^2, \qquad m > 0. \tag{3}$$

Then, if $\bar{x}^k$ is a point of minimum of the quadratic functional

$$f_k(x) = (f'(x^k), x - x^k) + \frac{1}{2}(f''(x^k)(x - x^k), x - x^k)$$

on the set Q, and $0 < \alpha \leq \alpha_k \leq 1$ is determined from the condition

$$f(x^{k+1}) - f(x^k) \leq \varepsilon\alpha_k f_k(\bar{x}^k), \qquad 0 < \varepsilon < 1,$$

then for the sequence (2) assertions 1)–3) of Theorem 1 hold, and, in addition: 4) there exists a number $N(\varepsilon)$ such that, for $k \geq N(\varepsilon)$, $\alpha_k \equiv 1$ and

$$\|x^{N+p} - x^*\| \leq C\lambda_N\lambda_{N+1}\cdots\lambda_{N+p},$$

$$\lambda_{N+i} < 1 \quad \text{for any } i \geq 0, \qquad \lambda_n \to 0 \quad \text{as } n \to \infty;$$

5) if, along with (3), the condition

$$\|f''(x) - f''(z)\| \leq R\|x - z\|, \qquad x, z \in Q,$$

is satisfied, then there exists a number $L(\varepsilon)$ such that, for $k \geq L(\varepsilon)$,

$$\alpha_k \equiv 1, \qquad \delta = \frac{2R}{m}\|x^{L+1} - x^L\| < 1, \qquad \|x^{L+p} - x^*\| \leq \frac{m}{2R}\sum_{i=p}^{\infty}\delta^{2^i}.$$

Let us also consider the following problem: to find the minimal value of the functional $f(x, y)$ on the set

$$A = \{(x, y) : \ x \in Q, \ P(x, y) = 0\}.$$

Here $Q \subset E_1$ is a closed convex bounded set; $y \in E_2$; $P \in (E_1 \times E_2) \to E_3$ is a differentiable nonlinear operator satisfying the requirements

$$\|P_y^{-1}(x, y)\| \leq N_1, \qquad \|P_x(x, y)\| \leq N_2, \qquad (x, y) \in A, \tag{4}$$

E_1, E_2, E_3 are Hilbert spaces. By virtue of (4), the equation $P(x, y) = 0$ determines a differentiable function $y = y(x)$, and

$$y'(x^k) = -P_y^{-1}(x^k, y^k)P_x(x^k, y^k).$$

Theorem 3. Let $f(x, y)$ be a twice differentiable function in $E_1 \times E_2$, and let

$$\|f_{xx}(x, y)\| \leq M, \qquad \|f_{xy}(x, y)\| \leq M, \qquad \|f_{yy}(x, y)\| \leq M$$

for all $(x, y) \in A$; let the operator $P(x, y)$ satisfy conditions (4), and let $y'(x)$ satisfy a Lipschitz condition with constant L for all $x \in Q$. Suppose, further,

$$\begin{aligned} f_k(x, y) = {} & (f_x(x^k, y^k), x - x^k) + (f_y(x^k, y^k), y - y^k) \\ & + \frac{1}{2}\Big[(f_{xx}(x^k, y^k)(x - x^k), x - x^k) + (f_{yy}(x^k, y^k)(y - y^k), y - y^k)\Big] \\ & + (f_{xy}(x^k, y^k)(x - x^k), y - y^k) \end{aligned}$$

is a function convex in (x, y); $(\bar{x}^k, \bar{y}^k_A)$ is a point of minimum of $f_k(x, y)$ on the set

$$A_\Delta = \{(x, y) : \ x \in Q, \ P_x(x^k, y^k)(x - x^k) + P_y(x^k, y^k)(y - y^k) = 0\}.$$

Then, if $\alpha_k = \beta_k\gamma_k$, where

$$\beta_k = \min\left\{1, \frac{-f_k(\bar{x}^k, \bar{y}^k_A)}{\|\bar{x}^k - x^k\|^2}\right\},$$

and $0 < \gamma \le \gamma_k \le 1/\beta_k$ is chosen from the condition

$$f(x^{k+1}, y^{k+1}) - f(x^k, y^k) \le \varepsilon \beta_k \gamma_k f_k(\bar{x}^k, \bar{y}_A^k), \qquad 0 < \varepsilon < 1,$$

then for the sequence

$$x^{k+1} = x^k + \alpha_k(\bar{x}^k - x^k), \qquad y^{k+1} = y(x^{k+1})$$

the following assertions hold: 1) $f(x^{k+1}, y^{k+1}) \le f(x^k, y^k)$; 2)

$$\lim_{k\to\infty} (\bar{x}^k, \bar{y}_A^k) = 0;$$

3) if $f(x, y(x))$ is a convex (in x) function, then

$$\lim_{k\to\infty} f(x^k, y^k) = f(x^*, y(x^*)) = \inf_{(x,y)\in A} f(x, y).$$

The author expresses gratitude to B. M. Budak and F. P. Vasil' ev for discussing the results and for useful remarks.

Institute of Cybernetics
Academy of Sciences of the Ukrainian SSR
Kiev

Received
10 III 1969

References

1. L. A. Lyusternik, V. I. Sobolev, *Elements of Functional Analysis*, Nauka, 1965.

2. M. N. Yakovlev, DAN, 156, No. 3 (1964).

3. V. F. Demyanov, A. M. Rubinov, Vestn. LGU, issue 19 (1964).

4. E. S. Levitin, B. T. Polyak, Zhurnal vychislitel' noi matematiki i matematicheskoi fiziki, 6, No. 5 (1966).

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