MATHEMATICS

S. I. Zukhovitskii, R. A. Polyak, M. E. Primak

1965

SovietRxiv

Abstract

Full Text

MATHEMATICS

S. I. Zukhovitskii, R. A. Polyak, M. E. Primak

A NUMERICAL METHOD FOR SOLVING A CONVEX PROGRAMMING PROBLEM IN A HILBERT SPACE

(Presented by Academician N. N. Bogolyubov on 14 I 1965)

1. Let a convex functional be given in a Hilbert space H

$$f_0(x) \tag{1}$$

and let there be a bounded domain Ω, defined by the inequalities

$$f_j(x) \leqslant 0, \qquad j \in J = \{1, \dots, p\}, \tag{2}$$

where $f_j(x)$ are convex functionals in H. We shall assume that the domain Ω contains interior points and that the functionals $f_0(x)$ and $f_j(x)$ possess Fréchet gradients $h_0(x)$ and $h_j(x)$, which are continuous operators acting from H into H.

We pose the problem of minimizing $f_0(x)$ in the domain Ω, i.e., of finding a vector (point) x^* such that

$$f_0(x^*) = \min_{x \in \Omega} f_0(x). \tag{3}$$

An analogous problem is considered in ([1,2]).

In the present work, for solving problem (1)–(3), an algorithm of steepest descent is constructed, in which each time, in order to find the direction of descent, a quadratic programming problem in a finite-dimensional space is solved. As in ([3,4]), at each step a certain parameter is introduced, which eliminates the possibility of "jamming" and determines the accuracy of the computations.

2. As the initial approximation we take an arbitrary vector $x^{(0)} \in \Omega$ and choose a sufficiently small number $\delta_1 > 0$. Define the set of indices

$$J(x^{(0)};\delta_1) = \{j \in J \mid -\delta_1 < f_j(x^{(0)}) \leqslant 0\}.$$

The direction of steepest descent $z^{(1)}$ is defined from the requirement that, when moving in this direction, both the functional $f_0(x)$ and the functionals $f_j(x)$, $j \in J(x^{(0)};\delta_1)$, decrease, and that the quantity

$$\min_{j\in J'(x^{(0)};\delta_1)} \left|(h_j(x^{(0)}), z^{(1)})\right|,$$

i.e., the smallest rate of decrease of the functionals $f_j(x)$, $j \in J'(x^{(0)};\delta_1) = J(x^{(0)};\delta_1) \cup \{0\}$, be as large as possible. Taking into account the negativity of the quantities $(h_j(x^{(0)}), z)$, $j \in J'(x^{(0)};\delta_1)$, the definition of the vector of steepest descent is reduced to finding a direction $z^{(1)}$ such that

$$\max_{j\in J'(x^{(0)};\delta_1)} (h_j(x^{(0)}), z^{(1)}) = \min_{\|z\|\leqslant 1} \max_{j\in J'(x^{(0)};\delta_1)} (h_j(x^{(0)}), z). \tag{4}$$

Since the orthogonal complement to the subspace spanned by the vectors $h_j(x^{(0)})$ plays no role in (4), we conclude that the direction of steepest descent has the form

$$z = \sum_{j\in J'(x^{(0)};\delta_1)} \xi_j h_j(x^{(0)}),$$

and the problem of finding the direction of steepest descent is reduced, therefore, to finding

$$\min_{\|z\|\leqslant 1} \max_{i\in J'(x^{(0)};\delta_1)} \sum_{j\in J'(x^{(0)};\delta_1)} (h_i(x^{(0)}),\, h_j(x^{(0)}))\xi_j =$$

$$= \min_{\|z\|\leqslant 1} \max_{i\in J'(x^{(0)};\delta_1)} \sum_{j\in J'(x^{(0)};\delta_1)} a_{ij}\xi_j, \quad \text{where } a_{ij} = (h_i(x^{(0)}),\, h_j(x^{(0)})),$$

which is equivalent to the problem of minimizing the linear form

$$u = \xi \tag{5}$$

subject to the constraints

$$\sum_{j\in J'(x^{(0)};\delta_1)} a_{ij}\xi_j \leqslant \xi, \quad i \in J'(x^{(0)};\delta_1),$$

$$\|z\| = \left\| \sum_{j \in J'(x^{(0)};\delta_1)} \xi_j h_j(x^{(0)}) \right\| \leqslant 1. \tag{6}$$

Let $\min u = u_1 < 0$. To determine the approximation step t_1, we find the smallest positive root of the equations

$$f_0(x^{(0)} + tz^{(1)}) = f_0(x^{(0)}) + \frac{1}{2} u_1 t,$$

$$f_j(x^{(0)} + tz^{(1)}) = 0, \quad j \in J.$$

If $u_1 < -\delta_1$, then we set $\delta_2 = \delta_1$, and if $-\delta_1 \leqslant u_1 < 0$, then we set $\delta_2 = \delta_1/2$. In the case $u_1 = 0$, we solve a problem of the type (5)–(6), replacing $J(x^{(0)};\delta_1)$ by the set $J(x^{(0)};0)$, i.e., considering only those surfaces $f_j(x) = 0$ for which $f_j(x^{(0)}) = 0$. If in this case too $\min u = 0$, then the problem is solved. Otherwise we continue the computations from the new approximation

$$x^{(1)} = x^{(0)} + t_1 z^{(1)}.$$

3. Let us outline the proof of convergence of the algorithm. First we shall verify that $\lim_{k\to\infty} \delta_k = 0$. Indeed, assuming the contrary, that $\lim_{k\to\infty} \delta_k = \delta > 0$, we obtain

$$\infty > \sum_{k=1}^{\infty} [f_0(x^{(k-1)}) - f_0(x^{(k)})] \geqslant \sum_{k=1}^{\infty} \frac{|u_k|}{2} t_k \geqslant \frac{\delta}{2} \sum_{k=1}^{\infty} t_k,$$

i.e. the series $\sum_{k=1}^{\infty} t_k$ converges, and since

$$\|x^{(n)} - x^{(m)}\| = \left\| \sum_{i=m+1}^{n} t_i z^{(i)} \right\| \leqslant \sum_{i=m+1}^{n} t_i,$$

the sequence $\{x^{(k)}\}$ converges strongly. Let $\lim x^{(k)} = \tilde{x}$. It is not difficult to show that the sequence $\{z^{(k)}\}$ of descent directions is also strongly convergent. Let $\lim_{k\to\infty} z^{(k)} = \tilde{z}$. We may assume that, starting from some index k_0, one and the same collection $J_0 \supset J(\tilde{x};\delta)$ of boundary surfaces $f_j(x) = 0$ participates in the definition of the descent direction. Therefore, from the assumption $\delta > 0$ it follows that

$$\tilde{u} = \max_{j \in J'(\tilde{x};\delta)} (h_j(\tilde{x}),\, \tilde{z}) \leqslant -\delta,$$

and there exists $\tilde{t} > 0$ such that

$$f_0(\tilde{x} + \tilde{t}\tilde{z}) < f_0(\tilde{x}) + \frac{1}{2}\tilde{u}\tilde{t}; \tag{7}$$

$$f_j(\tilde{x} + \tilde{t}\tilde{z}) < 0, \quad j \in J. \tag{8}$$

Consequently, for $x^{(k)}, z^{(k)}$ from sufficiently small neighborhoods of the points $\tilde{x}$ and $\tilde{z}$, an inequality of the type (7)–(8) will hold, i.e. $t_k > \tilde{t}$, which contradicts the convergence of the series $\sum_{k=1}^{\infty} t_k$. Consequently, $\lim_{k\to\infty} \delta_k = 0$.

Let $x^{(k)} \xrightarrow{\text{sl}} x^*$. Then $x^* \in \Omega$, and the assumption that x^* is not a solution of problem (1)–(3) contradicts the tendency of δ_k to zero. It can also be shown that

$$f_0(x^*) = \lim_{k\to\infty} f_0(x^{(k)}).$$

4. Let us give some examples of realizations of problem (1)–(3). Suppose a system of differential equations is given

$$\frac{dx}{dt} = A(t)\,x(t) + B(t)\,u(t) \tag{9}$$

with the initial condition $x(0) = x^{(0)}$.

a) Consider problem (5) of finding, among controls $u(t) \in U$, a control $u^*(t)$ such that the corresponding solution $x^*(t)$ of system (9) minimizes the convex functional $f_0(x) = \|x(t)\|$.

The domain U may be given, for example, by the inequality

$$f_1(u) = \|u(t)\| - 1 \le 0.$$

b) For a fixed value $t = \tilde{t}$ and a fixed point $x^{(1)}$ of the phase space X, one must find a control

$$u^*(t) \in U = \left\{ u(t) \;\middle|\; \sum_{i=1}^{n} \int_0^{\tilde{t}} u_i^2(s)\,ds \le 1 \right\}$$

such that the minimum of the functional is attained

$$f_0(u) = \|x(\tilde{t}) - x^{(1)}\|.$$

Kyiv State
Pedagogical Institute
named after A. M. Gorky

Ukrainian Road-Transport
Scientific-Research Institute

Received
28 XII 1964

CITED LITERATURE

1. V. F. Dem' yanov, A. M. Rubinov, *Vestn. Leningrad Univ.*, issue 4, No. 19 (1964).

2. B. N. Pshenichnyi, *Zhurn. Vychisl. Matem. i Matem. Fiz.*, 5, No. 1, 98 (1965).

3. S. I. Zukhovitskii, R. A. Polyak, M. E. Primak, *DAN*, 153, No. 5 (1963).

4. T. Zoytendeyk, *Methods of Feasible Directions*, Moscow, 1963.

5. R. Bellman, I. Glicksberg, O. Gross, *Some Questions of the Mathematical Theory of Control Processes*, IL, 1962.

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