MATHEMATICS

1965

SovietRxiv

Abstract

Full Text

MATHEMATICS

I. I. Eremin

A Relaxation Method for Solving Systems of Inequalities with Convex Functions on the Left-Hand Sides

(Presented by Academician I. M. Vinogradov on 14 VIII 1964)

In the present note a relaxation method is considered that makes it possible to find at least one solution (or at least one nonnegative solution) of the system of inequalities

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

in which $f_j(x)$ are convex smooth functions defined on the space $R^{(n)}$.

I. We preface the formulation and proof of the main assertions with a number of lemmas.

Lemma 1. *Let P be a half-space of the space $R^{(n)}$, corresponding to the inequality $(e, x - q) \leqslant 0$, and let p be an element of $R^{(n)}$ not lying in P, whose projection onto P coincides with q.*

Then the inequality $|y - p_\lambda| < |y - p|$ holds, where $p_\lambda = p + \lambda(q - p)$, $\lambda \in (0, 2)$, and y is an arbitrary element of P (see, for example, ([1])).

Here and below (x, y) denotes the scalar product of the vectors x and y.

A sequence $\{p_k\}$ of elements of the space $R^{(n)}$ is said to be Fejér with respect to a set $N \subset R^{(n)}$ if $|y - p_{k+1}| < |y - p_k|$ for $k = 1, 2, \dots$ and arbitrary $y \in N$. It is easily established (see ([1])) that if the sequence $\{p_k\}$ is Fejér with respect to N and p', p'' are any two limit points of this sequence, then N lies in the plane that is the geometric locus of points equidistant from p' and p''. Let us note here that it follows from this: 1) if the set N is convex and n-dimensional, then the sequence $\{p_k\}$ has a unique limit point; 2) if at least one of the limit points of the sequence $\{p_k\}$ belongs to N, then it is the unique limit point of the sequence under consideration.

Let M be some (nonempty) convex closed set in $R^{(n)}$; let $d(x)$ be a continuous convex function such that $\{y \mid d(y) \leqslant 0\} = M$; and let $e(x)$ be a vector function possessing the property: $|e(y)| \neq 0$ for $y \notin M$.

For an arbitrarily chosen $p_0 \in R^{(n)}$ define the sequence

$$p_0, p_1, \dots, p_k, \dots \tag{2}$$

by the relation

$$p_{k+1} = \begin{cases} p_k - \lambda_k \dfrac{d(p_k)}{|e_k|^2} e_k, & d(p_k) > 0, \\ p_k, & d(p_k) \leqslant 0; \end{cases} \tag{3}$$

here $\lambda_k \in (0, \beta] \subset (0, 2)$, and e_k is an abbreviated notation for the vector $e(p_k)$.

We make the following assumptions:

$(*)$ If $d(p_k) > 0$, then the half-space P_k corresponding to the inequality

$$(e_k, x - q_k) \leqslant 0, \quad \text{where} \quad q_k = p_k - [d(p_k)/|e_k|^2] e_k,$$

contains the set M.

$$(**) \quad \sup_k |e_k| = c < +\infty.$$

Lemma 2 (basic). *The sequence $\{p_k\}$, defined by relation (3) under conditions $(*)$ and $(**)$, converges to some vector $p' \in R^{(n)}$; if, moreover, $\operatorname{Inf}_k \lambda_k > 0$, then $p' \in M$.*

Proof. If for some $k = N$ we have $d(p_N) \leqslant 0$, then $p' = p_N \in M$, i.e. in this case the lemma is true. We may therefore assume, successively, that $d(p_k) > 0$ for $k = 0, 1, 2, \dots$. Under this assumption relation (3) takes the form

$$p_{k+1} = p_k + \lambda_k (q_k - p_k). \tag{4}$$

But then, by Lemma 1, the sequence (2) is Fejér with respect to M and, consequently, is bounded.

If $\operatorname{Inf}_k d(p_k) = 0$, then at least one of the limit points of the sequence (2) belongs to M. But this, in view of the second part of the remark made above, entails the convergence of the sequence (2). We shall prove its convergence also in the case when $\operatorname{Inf}_k d(p_k) = \delta > 0$. Let $0 < \varepsilon < \delta$. By direct verification we see that equality (4) is transformed into the form

$$p_{k+1} = p_k + \lambda_k^{(\varepsilon)} (q_k^{(\varepsilon)} - p_k), \tag{5}$$

where

$$\lambda_k^{(\varepsilon)} = \frac{\lambda_k d(p_k)}{d(p_k) - \varepsilon}, \qquad q_k^{(\varepsilon)} = q_k + \frac{\varepsilon}{|e_k|^2} e_k.$$

By choosing a suitable $\varepsilon = \varepsilon^0 > 0$, one can satisfy the condition

$$\lambda_k^{(\varepsilon_0)} \in (0, \beta'] \subset (0, 2). \tag{6}$$

Introducing the notation

$$M^\varepsilon = \{x \mid \rho(x, M) \leqslant \varepsilon\},$$

where $\rho(x, M)$ is the distance function from the point x to M, and $\varepsilon > 0$, $P_k^{(\varepsilon_0)}$ for the half-space corresponding to the inequality

$$(e_k, x - q_k^{(\varepsilon_0)}) \equiv (e_k, x - q_k) - \varepsilon_0 \leqslant 0,$$

and taking into account that $P_k \supset M$, we have

$$P_k^{(\varepsilon_0)} \supset M^{\varepsilon_0/|e_k|} \supset \bigcap_k M^{\varepsilon_0/|e_k|} = M^{\mathrm{Inf}_k\, \varepsilon_0/|e_k|} = M^{\varepsilon_0/c}.$$

By virtue of (5), (6), and Lemma 1, it follows that the sequence (2) is Fejér with respect to $M^{\varepsilon_0/c}$. Since the set $M^{\varepsilon_0/c}$ is, obviously, n-dimensional, it follows from item 1) of the remark made above that the sequence (2) converges.

Now the proof of the lemma is completed without difficulty. Indeed, if $\mathrm{Inf}_k\, \lambda_k > 0$, then, passing to the limit in the expression

$$d(p_k) = \frac{|e_k|}{\lambda_k} |p_{k+1} - p_k|,$$

which follows from (3), we obtain $d(p') = 0$ (p' is a limit element of the sequence (2)), i.e. $p' \in M$. The lemma is proved.

II. By the symbol $gf_j(x)$ we shall denote the gradient of the function $f_j(x)$ at the point x, i.e.

$$gf_j(x) = (\partial f_j(x)/\partial x_1, \dots, \partial f_j(x)/\partial x_n).$$

Lemma 3. *Let $f(x)$ be a convex smooth function defined on $R^{(n)}$, and let p be an element for which $f(p) > 0$ and $|gf(p)| \neq 0$. The half-space of the space $R^{(n)}$ corresponding to the inequality*

$$(gf(p), x - q) \leqslant 0, \qquad \text{where} \qquad q = p - \frac{f(p)}{|gf(p)|^2} gf(p),$$

contains the set

$$M(f) = \{x \mid f(x) \leqslant 0\}.$$

Proof. As is known, the requirement of convexity of the function $f(x)$ is equivalent to the requirement that the inequality

$$(gf(y), x - y) \leqslant f(x) - f(y)$$

hold for any $x, y \in R^{(n)}$.

Putting $y = p$, we transform it to the form $(gf(p), x - q) \leqslant f(x)$. This inequality evidently gives the assertion of the lemma.

Denote by $\Delta(p)$, $p \in R^{(n)}$, the set

$$\{i \mid \max_j f_j(p) = f_i(p)\}.$$

Theorem 1. *The sequence $\{p_k\}$, defined by relation (3) under the conditions*

a) $\mathrm{Inf}_k \lambda_k > 0$;

b) $d(x) = \max_j f_j(x)$;

c) $e_k = gf_{j_k}(p_k), \quad j_k \in \Delta(p_k)$

converges to one of the solutions of system (1), if the latter is consistent.

Remark. If system (1) is consistent, then the number $|gf_{j_k}(p_k)| = |e_k|$, which stands in the denominator in relation (3), cannot be equal to zero. Indeed, otherwise the vector p_k would deliver for the function $f_{j_k}(x)$ a minimum value on $R^{(n)}$ equal to $f_{j_k}(p_k) > 0$. But this contradicts the fact that any solution of system (1) delivers for the function $f_{j_k}(x)$ a nonpositive value.

Let $s(x) = \{i \mid f_j(x) > 0\}$ and $c_j > 0$ $(j = 1, 2, \dots, m)$ be a system of constants.

Theorem 2. *The sequence $\{p_k\}$, defined by relation (3) under the conditions*

a) $\mathrm{Inf}_k \lambda_k > 0$;

b)

$$d(x) = \begin{cases} \sum_{j \in s(x)} c_j f_j(x), & s(x) \text{ is nonempty}, \\ 0, & s(x) \text{ is empty}; \end{cases}$$

c)

$$e_k = \sum_{j \in s(p_k)} c_j g f_j(p_k)$$

converges to one of the solutions of system (1), if the latter is consistent.

The proofs of Theorems 1 and 2 are obtained by a direct application of Lemmas 2 and 3 (where the role of M is played here by the set of solutions of system (1)).

Remark. When applied to a system of linear inequalities, Theorems 1 and 2 lead to the methods described in papers ([1],[2]).

III. Let us consider the question of finding nonnegative solutions of system (1). If $x = (x_1, \dots, x_n) \in R^{(n)}$, then by x^+ we shall mean the vector whose negative coordinates have been replaced by zeros.

Theorem 3. *If system (1) has at least one nonnegative solution, then the sequence* $\{p_k\}$*, defined (for an arbitrary nonnegative vector* $p_0 \in R^{(n)}$*) by the relation*

$$p_{k+1} = \begin{cases} \left[p_k - \lambda_k \dfrac{d(p_k)}{|e_k|^2} e_k\right]^+, & d(p_k) > 0, \\ p_k, & d(p_k) \leqslant 0, \end{cases}$$

where the functions $d(x)$*,* $e(x)$ *and the numbers* λ_k *are subject to the conditions of Theorem 1 (or to the conditions of Theorem 2), converges to one of such solutions.*

The proof of Theorem 3 differs only in some details from the proofs of Theorems 1 and 2.

Sverdlovsk Branch
of the Mathematical Institute named after V. A. Steklov
of the Academy of Sciences of the USSR

Received
29 VI 1964

REFERENCES

[1] S. Agmon, Canad. J. Math., 6, No. 3, 382 (1954).
[2] Yu. I. Merzlyakov, Zhurn. vychisl. matem. i matem. fiz., 2, No. 3, 482 (1962).

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