ON FINDING A MINIMAX ON A CONSTRAINED SET

MATHEMATICS

1970

SovietRxiv

Abstract

Full Text

UDC 519.8

MATHEMATICS

V. F. DEM' YANOV

ON FINDING A MINIMAX ON A CONSTRAINED SET

(Presented by Academician L. V. Kantorovich, 29 IV 1969)

1°. Let

$$\varphi(X) = \max_{i \in 1,\dots,N} f_i(X), \tag{1}$$

where $X \in E_n$, and the functions f_i $(i = 1, \dots, N)$ are continuously differentiable on a bounded open set $S \subset E_n$ containing a convex compact set Ω. It is required to find

$$\min_{X \in \Omega} \varphi(X).$$

As established in ($^{1-5}$), the function φ is continuous and directionally differentiable.

Theorem 1 (see ($^{1-3, 6}$)). *In order that at a point $Y \in \Omega$ the function φ attain its minimal value on Ω, it is necessary (and, in the case of convexity of φ on Ω, sufficient) that*

$$\min_{Z \in \Omega} \max_{i \in R(Y)} \left(\frac{\partial f_i(Y)}{\partial Y}, Z - Y \right) = 0, \tag{2}$$

where $R(X) = \{i \mid i = 1, \dots, N,\ f_i(X) = \varphi(X)\}$.

A point Y satisfying (2) is called a stationary point of the function φ on the set Ω.

Let $\gamma(X)$ be the cone of feasible directions at the point $X \in \Omega$; let $\Gamma(X)$ be the closure of $\gamma(X)$, and let $\Gamma^+(X)$ be the cone conjugate to $\Gamma(X)$.

Then condition (2) means that at the minimum point one must have

$$L(Y) \cap \Gamma^+(Y) \neq \Lambda, \tag{3}$$

where $L(Y)$ is the convex hull of the set

$$H(Y) = \{\partial f_i(Y)/\partial Y,\ i \in R(Y)\}.$$

Let

$$d(X) = \min_{\substack{V \in L(X) \\ Z \in \Gamma^+(X)}} \|V - Z\|.$$

If $d(X) = 0$, then the point X is stationary; if, however, $d(X) = \|V(X) - Z(X)\| > 0$, then the direction $g(X) = d^{-1}(X)(Z(X) - V(X))$ (clearly, $\|g(X)\| = 1$) is a direction of steepest descent of the function φ at the point $X \in \Omega$ on the set Ω. Moreover, such a $g(X)$ is unique and $g(X) \in \Gamma(X)$.

2°. We now consider the case where

$$\Omega = \{X \mid X \in E_n,\ h_j(X) \leq 0,\ j = 1, \dots, N_1\}, \tag{4}$$

where the functions $h_j(X)$ $(j = 1, \dots, N_1)$ are continuously differentiable and convex on S, and Slater' s condition is satisfied,

$$\min_{X \in E_n} \max_{j \in 1, \dots, N_1} h_j(X) < 0, \tag{5}$$

i.e., there exists a strictly interior point of the set Ω. Without loss of generality we assume that the set Ω is bounded.

In (7–10) various methods of successive approximations were obtained for finding stationary points of φ on Ω.

To obtain other algorithms, we use the necessary condition (3). In the present case

$$\Gamma^+(X) = \left\{ g \mid g = -\sum_{j \in Q(X)} \alpha_j \frac{\partial h_j(X)}{\partial X}, \quad \alpha_j \geqslant 0 \right\}.$$

If $Q(X) = \Lambda$, then $\Gamma^+(X) = \{0\}$.

By virtue of condition (5), there exists $\mu_0 > 0$ such that, for any X such that

$$-\mu_0 \leqslant \max_{j \in 1, \dots, N} k_j(X) h_j(X) \leqslant 0, \tag{6}$$

we have

$$\min_{\|g\|\leqslant 1}\max_{j\in Q_{\mu_0}(X)}\left(k_j(X)\frac{\partial h_j(X)}{\partial X},\,g\right)=\max_{j\in Q_{\mu_0}(X)}\left(k_j(X)\frac{\partial h_j(X)}{\partial X},\,q(X)\right)\leqslant -a_0<0, \tag{7}$$

where

$$Q_{\mu_0}(X)=\{j\mid j\in 1,\dots,N_1,\ -\mu_0\leqslant k_j(X)h_j(X)\leqslant 0\},\quad k_j(X)=\left\|\frac{\partial h_j(X)}{\partial X}\right\|^{-1}.$$

It is clear that $k_j(X)\leqslant K,\ j\in Q_{\mu_0}(X)$, for all X satisfying (6). Then there exists $a_0>0$ such that, for any $X\in\Omega$, one has $X+\alpha q(X)\in\Omega$ for $\alpha\in[0,a_0]$.

We now describe a method of successive approximations which is a generalization of the steepest descent method (11). Let $\varepsilon'>0,\ \mu'>0$ (and $\mu'\leqslant\mu_0$), $\rho'>0$ be fixed. Introduce the function

$$d_{\varepsilon\mu}(X)=\min_{\substack{V\in L_\varepsilon(X)\\ Z\in\Gamma_\mu^+(X)}}\|V-Z\|,$$

where

$$\varepsilon\geqslant 0,\quad \mu\geqslant 0,\quad L_\varepsilon(X)=\operatorname{co}H_\varepsilon(X),\quad H_\varepsilon(X)=\{\partial f_i(X)/\partial X,\ i\in R_\varepsilon(X)\},$$

$$[R_\varepsilon(X)]=\{i\mid i\in 1,\dots,N,\ \varphi(X)-f_i(x)\leqslant\varepsilon\},$$

$$\Gamma_\mu^+(X)=\left\{g\mid g=-\sum_{j\in Q_\mu(X)}\alpha_j\frac{\partial h_j(X)}{\partial X},\quad \alpha_j\geqslant 0\right\},$$

$$Q_\mu(X)=\{j\mid j\in 1,\dots,M_1,\ -\mu\leqslant k_j(X)h_j(X)\leqslant 0\},$$

$$k_j(X)=\|\partial h_j(X)/\partial X\|^{-1}.$$

Set $\Gamma_\mu^+(X)=\{0\}$ if $Q_\mu(X)=\Lambda$.

If $d_{\varepsilon\mu}(X)>0$, then $d_{\varepsilon\mu}(X)=\|V_{\varepsilon\mu}(X)-Z_{\varepsilon\mu}(X)\|$, and the vector $g_{\varepsilon\mu}(X)=d_{\varepsilon\mu}^{-1}(X)(Z_{\varepsilon\mu}(X)-V_{\varepsilon\mu}(X))$ is unique, moreover $g_{\varepsilon\mu}(X)\in\Gamma(\overline{X})$.

As the first approximation, take an arbitrary point $X_1 \in \Omega$. Suppose that $X_k \in \Omega$ has already been found. If $d(X_k) = 0$, then the point X_k is a stationary point of the function φ on Ω, and the process terminates. If $d(X_k) > 0$, set $\varepsilon_{k1} = \varepsilon'$, $\mu_{k1} = \mu'$, $\rho_{k1} = \rho'$, and find $d_{\varepsilon_{k1}\mu_{k1}}(X_k)$.

If $d_{\varepsilon_{k1}\mu_{k1}}(X_k) \geqslant \rho_{k1}$, then set $\varepsilon_k = \varepsilon_{k1}$, $\mu_k = \mu_{k1}$, $\rho_k = \rho_{k1}$, $g_k = g_{\varepsilon_{k1}\mu_{k1}}(X)$, $\bar{d}_k = d_{\varepsilon_{k1}\mu_{k1}}(X_k)$. Otherwise, if $d_{\varepsilon_{k1}\mu_{k1}}(X_k) < \rho_{k1}$, take $\varepsilon_{k2} = \frac{1}{2}\varepsilon_{k1}$, $\mu_{k2} = \frac{1}{2}\mu_{k1}$, $\rho_{k2} = \frac{1}{2}\rho_{k1}$, again find $d_{\varepsilon_{k2}\mu_{k2}}(X_k)$, and so continue until we find the smallest r_k such that

$$d_{\varepsilon_{kr_k}\mu_{kr_k}}(X_k) \geqslant \rho_{kr_k}. \tag{8}$$

(such a finite r_k will necessarily be found, since $d(X_k) > 0$), and set

$$\varepsilon_k = \varepsilon_{kr_k}, \quad \mu_k = \mu_{kr_k}, \quad \rho_k = \rho_{kr_k},$$
$$g_k = g_{\varepsilon_{kr_k}\mu_{kr_k}}(X_k), \quad \bar{d}_k = d_{\varepsilon_{kr_k}\mu_{kr_k}}(X_k).$$

It is clear that

$$g_k \in \Gamma(X_k), \quad (\partial f_i(X_k)/\partial X, g_k) \leq -\bar{d}_k \quad (i \in R_{\varepsilon_k}(X_k)),$$

$$(\partial h_j(X_k)/\partial X, g_k) \leq 0 \quad (j \in Q_{\mu_k}(X_k)).$$

Further, to correct the direction g_k, one may proceed in one of the following ways.

Method 1. Let

$$K = \max_{X \in \Omega} \|\partial f_j(x)/\partial X\|.$$

Take the ray

$$X_{k\alpha} = X_k + \alpha\left(g_k + \frac{\bar{d}_k}{2K} q_k\right).$$

Then for $j \in Q_{\mu_k}(X_k)$ we have

$$K_j(X_k)h_j(X_{k\alpha}) \leq K_j(X_k)h_j(X_k) - \alpha\frac{\bar{d}_k}{2K}a_0 + o(\alpha),$$

$$f_i(X_{k\alpha}) \leq f_i(X_k) - \frac{1}{2}\alpha\bar{d}_k + o(\alpha) \quad (i \in R_{\varepsilon_k}(X_k)),$$

i.e., for sufficiently small α it will turn out that $X_{k\alpha} \in \Omega$, $\varphi(X_{k\alpha}) < \varphi(X_k)$.

Method 2. Let

$$\Gamma^+_{\mu\xi}(X) = \mathrm{co}\{g \mid g = \alpha Z,\ \|Z - \|V\|^{-1}V\| \le \xi,\ \alpha \ge 0,\ V \ne 0,\ V \in \Gamma^+_\mu(X)\}.$$

Find the maximal ξ_k such that

$$d_{\varepsilon_k \mu_k \xi_k}(X_k) \equiv \min_{\substack{V \in L_{\varepsilon_k}(X_k) \\ Z \in \Gamma^+_{\mu_k \xi_k}(X_k)}} \|V - Z\| = \|V'_k - Z'_k\| \ge \frac{1}{2} d_k.$$

It is clear that, by condition (5), $\xi_k > 0$ if $\bar{d}_k > 0$. Let

$$g'_k = (Z'_k - V'_k)/\|V'_k - Z'_k\|.$$

It is not difficult to obtain that

$$(\partial f_i(X_k)/\partial X, g'_k) \le -\frac{1}{2} d_k \quad (i \in R_{\varepsilon_k}(X_k)),$$

$$K_j(X_k)(\partial h_j(X_k)/\partial X, g'_k) \le -\xi_k \quad (j \in Q_{\mu_k}(X_k)),$$

i.e., the direction g'_k is feasible.

Consider now the ray $X_{k\alpha} = X_k + \alpha g'_k$. Thus, applying Method 1 or 2, we have $X_{k\alpha}$. Find $a_k \ge 0$ such that

$$\varphi(X_{k\alpha_k}) = \min_{\substack{\alpha \ge 0 \\ X_{k\alpha} \in \Omega}} \varphi(X_{k\alpha})$$

and set $X_{k+1} = X_{k\alpha_k}$ (a_k is finite, since the set Ω is bounded). It is clear that $X_{k+1} \subset \Omega$, $\varphi(X_{k+1}) < \varphi(X_k)$, if $d(X_k) > 0$.

We then continue analogously. Thus we construct a sequence of points $\{X_k\} \subset \Omega$. If this sequence consists of a finite number of points, then its last obtained point is a stationary point of the function φ on Ω. Otherwise the following is true.

Theorem 2. *Every limit point of the sequence $\{X_k\}$ is a stationary point of the function φ on the set Ω.*

Remark 1. If all functions h_j are linear, then one need not find and use the vector $q(X_k)$ (or q'_k), setting

$$X_{k\alpha} = X_k + \alpha g_k.$$

Remark 2. Let

$$\Omega = \{X \mid h_j(X) \le 0,\ j \in 1, \dots, N_1;\ (A_j, X) + b_j = 0,\ j \in N_1 + 1, \dots, N_2\}.$$

In this case condition (5) is replaced by the condition

$$\min_{X\in\Omega_1} \max_{j\in 1,\dots,N_1} h_j(X) < 0,$$

where $\Omega_1 = \{X \mid (A_j, X) + b_j = 0,\ j \in N_1 + 1, \dots, N_2\}$, and the cone $\Gamma^+(X)$ is defined as follows:

$$\Gamma^+(X) = \left\{ g \,\middle|\, g = -\sum_{j\in Q(X)} \alpha_j \frac{\partial h_j(X)}{\partial X} + \sum_{j=N_1+1}^{N_2} \beta_j A_j;\ \alpha_j \geqslant 0,\ -\infty < \beta_j < \infty \right\}.$$

The cone $\Gamma^+_\mu(X)$ is defined analogously.

Remark 3. As in ($^{1-3}$), the results obtained can be extended to the case when $\varphi(X) = \max_{Y\in\Omega} f(X, Y)$, where $\Omega_1 \in E_m$ is some compact set.

Remark 4. All auxiliary problems whose solution is necessary at each step may be solved approximately.

Remark 5. It is not difficult to prove that $r_k \to \infty$ as $k \to \infty$. To reduce the amount of computation required, note that $\varepsilon_k, \mu_k, \rho_k$ may be found from the formulas:

$$\varepsilon_k = \varepsilon_{kr_k} = \varepsilon_{k-1}/2^{r_k-1}, \qquad \mu_k = \mu_{kr_k} = \mu_{k-1}/2^{r_k-1}, \qquad \rho_k = \rho_{kr_k}/2^{r_k-1},$$

where r_k is the smallest integer (not necessarily nonnegative) for which (8) holds. In addition, it must be that $\varepsilon_{kr_k} \leqslant \varepsilon'$.

Leningrad State University
named after A. A. Zhdanov

Received
21 IV 1969

CITED LITERATURE

[1] V. F. Demyanov, Vestn. LGU, No. 7 (1966).
[2] V. F. Demyanov, Kibernetika, No. 6 (1966); No. 3 (1967).
[3] V. F. Demyanov, A. M. Rubinov, SIAM J. Control, 6, No. 1 (1968).
[4] B. N. Pshenichnyi, Kibernetika, No. 6 (1967).
[5] J. M. Danskin, *The Theory of Max-Min and its Application to Weapons Allocation Problems*, N. Y., 1967.
[6] Yu. B. Germeyer, Kibernetika, No. 3 (1967).
[7] G. Zoutendijk, *Methods of Feasible Directions*, IL, 1963.
[8] S. I. Zukhovitskii, R. A. Polyak, M. E. Primak, DAN, 163, No. 2 (1965).
[9] L. I. Abdeeva, S. I. Zukhovitskii, *Linear and Convex Programming*, "Nauka," 1964.
[10] V. F. Demyanov, J. Computer and Systems Sciences, 3, 342 (1968).
[11] L. V. Kantorovich, DAN, 28, No. 1 (1940).

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

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