SUCCESSIVE APPROXIMATIONS FOR FINDING SADDLE POINTS

MATHEMATICS

1967

SovietRxiv

Abstract

Full Text

UDC 519.8

MATHEMATICS

V. F. DEM' YANOV

SUCCESSIVE APPROXIMATIONS FOR FINDING SADDLE POINTS

(Presented by Academician L. V. Kantorovich on 9 I 1967)

1°. Statement of the problem. Of great interest is the problem of finding saddle points of functions ([1–3]). The problem of finding the minimax of a continuously differentiable function was considered in ([4]).

Let $f(X, Y)$ be a function continuously differentiable on $\Omega_X \times \Omega_Y$, and let the sets $\Omega_X \subset E_n$ and $\Omega_Y \subset E_m$ be convex, closed, and bounded.

Let (X^*, Y^*) be a saddle point of the function $f(X, Y)$ on $\Omega_X \times \Omega_Y$, i.e., for all $X \in \Omega_X,\ Y \in \Omega_Y$,

$$f(X, Y^*) \leq f(X^*, Y^*) \leq f(X^*, Y); \tag{1}$$

then

$$f(X^*, Y^*) = \max_{X \in \Omega_X} f(X, Y^*) = \min_{Y \in \Omega_Y} f(X^*, Y). \tag{2}$$

We shall call the function $f(X, Y)$ **concave-convex** if, for every fixed $Y \in \Omega_Y$, the function $f_Y(X) \equiv f(X, Y)$ is concave in X on Ω_X, and for every fixed $X \in \Omega_X$, the function $f_X(Y) \equiv f(X, Y)$ is convex in Y on Ω_Y.

It is required to find a saddle point of the function $f(X, Y)$ on $\Omega_X \times \Omega_Y$.

As is not difficult to see, the following is true (for example, see ([5])).

Theorem 1. *In order that the point (X^*, Y^*) be a saddle point of the function $f(X, Y)$ on the set $\Omega_X \times \Omega_Y$, it is necessary (and if $f(X, Y)$ is concave-convex on $\Omega_X \times \Omega_Y$, also sufficient) that*

$$\max_{X \in \Omega_X} \left(\frac{\partial f(X^*, Y^*)}{\partial X} \right)^* (X - X^*) = \min_{Y \in \Omega_Y} \left(\frac{\partial f(X^*, Y^*)}{\partial Y} \right)^* (Y - Y^*) = 0. \tag{3}$$

Corollary. *If $\Omega_X = E_n,\ \Omega_Y = E_m$, then condition (3) is replaced by the condition*

$$\partial f(X^*, Y^*)/\partial X = \partial f(X^*, Y^*)/\partial Y = 0. \tag{4}$$

A point $(X^*, Y^*) \in \Omega_X \times \Omega_Y$ satisfying (3), or respectively (4) (if $\Omega_X = E_n,\ \Omega_Y = E_m$), is called a **stationary point** of the function $f(X, Y)$ on the set $\Omega_X \times \Omega_Y$.

2°. Let $\Omega_X = E_n,\ \Omega_Y = E_m$. Consider the systems of differential equations

$$dX(t)/dt \equiv \dot{X}(t) = \partial f(X, Y)/\partial X; \tag{5}$$

$$\dot{X}(0) = X_0 \in E_n; \tag{6}$$

$$dY(t)/dt \equiv \dot{Y}(t) = \partial f(X, Y)/\partial Y; \tag{7}$$

$$Y(0) = Y_0 \in E_m. \tag{8}$$

By $X(t, X_0, Y_0),\ Y(t, X_0, Y_0)$ we denote the solutions of systems (5), (7) with the initial conditions (6), (8).

Suppose that the function $f(X, Y)$ is twice continuously differentiable and strictly concave-convex on $E_n \times E_m$. Then the matrices

$-\partial^2 f/\partial X^2,\ \partial^2 f/\partial Y^2$ are strictly positive definite, i.e., for any finite $(z_1, z_2) \in E_n \times E_m$ and for any $(X, Y) \in E_n \times E_m$

$$-z_1^* \left(\frac{\partial^2 f(X, Y)}{\partial X^2} z_1 \right) \geq m_1(X, Y) \|z_1\|^2, \qquad m_1(X, Y) > 0; \tag{9}$$

$$z_2^* \left(\frac{\partial^2 f(X, Y)}{\partial Y^2} z_2 \right) \geq m_2(X, Y) \|z_2\|^2, \qquad m_2(X, Y) > 0. \tag{10}$$

For any bounded set $S \subset E_n \times E_m$ there exist $m_1 > 0$ and $m_2 > 0$, depending on S, such that $m_1(X, Y) \geq m_1 > 0,\ m_2(X, Y) \geq m_2 > 0$ for all $(X, Y) \in S$.

By $M(X_0, Y_0) \subset E_n \times E_m$ we denote the set

$$\{(X, Y) \mid F(X, Y) \leq F(X_0, Y_0)\},$$

where

$$F(X,Y) = \frac{1}{2}\left[(\partial f(X,Y)/\partial X)^2 + (\partial f(X,Y)/\partial Y)^2\right].$$

Under the assumptions made, the following is valid.

Theorem 2. *If the set $M(X_0,Y_0)$ is bounded, then the solutions $X(t,X_0,Y_0)$, $Y(t,X_0,Y_0)$ of systems (5) and (7) converge to the unique saddle point.*

Systems (5), (7) give a "continuous" method for finding saddle points in the whole space. On the basis of this "continuous" method one can construct a number of discrete methods for finding saddle points. Let us consider one of the possible such methods.

Take arbitrary $(X_1,Y_1) \in E_n \times E_m$. Suppose that $M(X_1,Y_1)$ is bounded. Let (X_k,Y_k) have been found. Consider the rays

$$X_{k\alpha} = X_k + \alpha G_{Xk}, \qquad \alpha \in [0,\infty),$$

$$Y_{k\beta} = Y_k - \beta G_{Yk}, \qquad \beta \in [0,\infty),$$

where $G_{Xk} = \partial f(X_k,Y_k)/\partial X, \quad G_{Yk} = \partial f(X_k,Y_k)/\partial Y$.

We have

$$F(X_{k\alpha},Y_{k\beta}) = F(X_k,Y_k) + \alpha A_k + (\beta-\alpha)B_k - \beta C_k + O_k(\alpha,\beta), \tag{11}$$

where

$$A_k = \left(\frac{\partial f(X_k,Y_k)}{\partial X}\right)^* \left(\frac{\partial^2 f(X_k,Y_k)}{\partial X^2}\frac{\partial f(X_k,Y_k)}{\partial X}\right),$$

$$B_k = \left(\frac{\partial f(X_k,Y_k)}{\partial X}\right)^* \left(\frac{\partial^2 f(X_k,Y_k)}{\partial X\,\partial Y}\frac{\partial f(X_k,Y_k)}{\partial Y}\right),$$

$$C_k = \left(\frac{\partial f(X_k,Y_k)}{\partial Y}\right)^* \left(\frac{\partial^2 f(X_k,Y_k)}{\partial Y^2}\frac{\partial f(X_k,Y_k)}{\partial Y}\right),$$

$$\frac{O_k(\alpha,\beta)}{\sqrt{\alpha^2+\beta^2}} \longrightarrow 0 \quad \underset{\beta\to+0}{\overset{\alpha\to+0}{\longrightarrow}} \quad \text{uniformly in } k,$$

and moreover

$$A_k < 0, \quad \text{if } G_{Xk} \neq 0,$$

$$C_k > 0, \quad \text{if } G_{Yk} \neq 0.$$

If $B_k < 0$, set $\beta = 2\alpha$; if $B_k \geq 0$, set $\beta = \frac{1}{2}\alpha$. Then, if $2F(X_k, Y_k) = G_{Xk}^2 + G_{Yk}^2 > 0$, for sufficiently small α and β we will have $F(X_{k\alpha}, Y_{k\beta}) < F(X_k, Y_k)$.

Find $\alpha_k \in [0, \infty)$ from the condition

$$F(X_{k\alpha_k}, Y_{k\beta(\alpha_k)}) = \min_{\alpha \in [0,\infty)} F(X_{k\alpha}, Y_{k\beta(\alpha)})$$

and set

$$X_{k+1} = X_{k\alpha_k}, \qquad Y_{k+1} = Y_{k\beta(\alpha_k)}.$$

We proceed analogously further.

It can be shown that $X_k \underset{k\to\infty}{\longrightarrow} X^*$, $Y_k \underset{k\to\infty}{\longrightarrow} Y^*$, and (X^*, Y^*) is a saddle point of the function $f(X, Y)$ on $E_n \times E_m$.

Remark. As in the ordinary gradient method, one need not seek, at each step, the minimum of $F(X_{k\alpha}, Y_{k\beta})$ on $[0, \infty)$, but may set $X_{k+1} = X_{k\alpha_k}$, $Y_{k+1} = Y_{k\beta(\alpha_k)}$, where $\alpha_k \in [\varepsilon_0, \varepsilon_1]$, $\varepsilon_1 > \varepsilon_0 > 0$, are certain fixed quantities independent of k.

3°. Let $\Omega_X \subset E_n$, $\Omega_Y \subset E_m$ be strictly convex, bounded and closed sets. Then consider the functions

$$\psi(X, Y) = \max_{z \in \Omega_X} (\partial f(X, Y)/\partial X)^*(z - X), \tag{12}$$

$$\varphi(X, Y) = \min_{z \in \Omega_Y} (\partial f(X, Y)/\partial Y)^*(z - Y). \tag{13}$$

For all $(X, Y) \in \Omega_X \times \Omega_Y$, $\psi(X, Y) \geq 0$; $\varphi(X, Y) \leq 0$. Since Ω_X, Ω_Y are strictly convex sets, for fixed (X, Y) there exists a unique point $\theta_1(X, Y) \in \Omega_X$ and a unique point $\theta_2(X, Y) \in \Omega_Y$ such that

$$\psi(X, Y) = (\partial f(X, Y)/\partial X)^*(\theta_1(X, Y) - X),$$

$$\varphi(X, Y) = (\partial f(X, Y)/\partial Y)^*(\theta_2(X, Y) - Y).$$

The vector functions $\theta_1(X, Y)$ and $\theta_2(X, Y)$ are continuous on $\Omega_X \times \Omega_Y$. Consider the systems of differential equations

$$dX(t)/dt \equiv \dot{X}(t) = \theta_1(X(t), Y(t)) - X(t); \tag{14}$$

$$X(0) = X_0; \tag{15}$$

$$dY(t)/dt \equiv \dot{Y}(t) = \theta_2(X(t), Y(t)) - Y(t); \tag{16}$$

$$Y(0) = Y_0. \tag{17}$$

If $X_0 \in \Omega_X$, $Y_0 \in \Omega_Y$, then the solutions $X(t) \equiv X(t, X_0, Y_0)$, $Y(t) \equiv Y(t, X_0, Y_0)$ of the systems (14), (16) (the solutions exist and are continuous by Peano' s theorem) belong, for $t \in [0, \infty)$, respectively to the sets Ω_X and Ω_Y.

Theorem 3. *If $f(X, Y)$ is a strictly concave-convex function, then the solutions of the systems (14), (16) for $(X_0, Y_0) \in \Omega_X \times \Omega_Y$ converge to the unique saddle point.*

On the basis of the "continuous" method (14), (16) for finding a saddle point, one can develop discrete methods for searching for saddle points. We give one of them.

Take arbitrary $X_1 \in \Omega_X$, $Y_1 \in \Omega_Y$. Suppose X_k, Y_k have been found ($X_k \in \Omega_X$, $Y_k \in \Omega_Y$). Let $\theta_{1k} = \theta_1(X_k, Y_k)$, $\theta_{2k} = \theta_2(X_k, Y_k)$. If $H(X_k, Y_k) = 0$, then the point (X_k, Y_k) is a saddle point, and the process is finished. If, however, $H(X_k, Y_k) > 0$, then consider the segment in Ω_X $X_{k\alpha} = X_k + \alpha(\theta_{1k} - X_k)$, $\alpha \in [0, 1]$, $X_{k\alpha} \in \Omega_X$, and the segment in Ω_Y $Y_{k\beta} = Y_k + \beta(\theta_{2k} - Y_k)$, $\beta \in [0, 1]$, $Y_{k\beta} \in \Omega_Y$. We have

$$h_1(\alpha, \beta) \equiv H(X_{k\alpha}, Y_{k\beta}) = H(X_k, Y_k) + \alpha A_k +$$

$$+(\beta - \alpha)B_k - \beta C_k + O_k(\alpha, \beta),$$

where

$$A_k = (\theta_{1k} - X_k)^* \frac{\partial^2 f(X_k, Y_k)}{\partial X^2} (\theta_{1k} - X_k),$$

$$B_k = (\theta_{2k} - Y_k)^* \frac{\partial^2 f(X_k, Y_k)}{\partial Y \, \partial X} (\theta_{1k} - X_k),$$

$$C_k = (\theta_{2k} - Y_k)^* \frac{\partial^2 f(X_k, Y_k)}{\partial Y^2} (\theta_{2k} - Y_k),$$

$$O_k(\alpha,\beta)/\sqrt{\alpha^2+\beta^2} \xrightarrow[\alpha\to+0,\ \beta\to+0]{} 0 \quad \text{uniformly in } k.$$

If $B_k < 0$, then we set $\beta = 2\alpha$. In this case we consider the function $h_2(\alpha) \equiv h_1(\alpha, 2\alpha)$, find $\alpha_k \in [0, {}^1/_2]$ such that $h_2(\alpha_k) = \min_{\alpha\in[0,{}^1/_2]} h_2(\alpha)$, and set

$$X_{k+1} = X_k + \alpha_k(\theta_{1k} - X_k), \qquad Y_{k+1} = Y_k + 2\alpha_k(\theta_{2k} - Y_k).$$

If, on the other hand, $B_k \geq 0$, then we set $\beta = {}^1/_2\alpha$ and consider the function $h_3(\alpha) \equiv h_1(\alpha, {}^1/_2\alpha)$. We find $\alpha_k \in [0,1]$ such that $h_3(\alpha_k) = \min_{\alpha\in[0,1]} h_3(\alpha)$ and set

$$X_{k+1} = X_k + \alpha_k(\theta_{1k} - X_k), \qquad Y_{k+1} = Y_k + {}^1/_2\alpha_k(\theta_{2k} - Y_k).$$

It is clear that in both cases $X_{k+1} \in \Omega_X$, $Y_{k+1} \in \Omega_Y$, and if $H(X_k, Y_k) > 0$, then $H(X_{k+1}, Y_{k+1}) < H(X_k, Y_k)$.

Thus we construct the sequences $\{X_k\}$, $\{Y_k\}$. The sequence $\{H_k\}$, $H_k = H(X_k, Y_k)$, is monotonically decreasing and therefore converges. Let $H^* = \lim_{k\to\infty} H_k$. Then $H_k \geq H^*$.

Theorem 4. *The sequences $\{X_k\}$, $\{Y_k\}$ constructed above converge to a saddle point of the function $f(X,Y)$ on the set $\Omega_X \times \Omega_Y$.*

Leningrad State University
named after A. A. Zhdanov

Received
28 XII 1966

REFERENCES

1. K. J. Arrow, L. Hurwicz, H. Uzawa, *Studies in Linear and Nonlinear Programming*, IL, 1962.
2. D. Robinson, in: *Matrix Games*, Moscow, 1961.
3. V. A. Volkonskii, *Economics and Mathematical Methods*, 1, No. 2 (1965).
4. V. F. Dem' yanov, *Cybernetics*, 2, No. 6, Kiev (1965).
5. V. F. Dem' yanov, *Cybernetics*, 1, No. 6, Kiev (1965).

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