

# ON THE MINIMAX PROBLEM

V. F. Dem'yanov

1969

SovietRxiv

---

View the original and related papers at <https://sovietsrxiv.org/items/ru-196901.83093>

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

**Abstract**

**Full Text**

UDC 519.8

**MATHEMATICS**

V. F. DEMYANOV

## ON THE MINIMAX PROBLEM

(Presented by Academician L. V. Kantorovich, 16 XII 1968)

§ 1. Let the function  $f(X, Y)$  be continuously differentiable with respect to  $X$  on  $E_n \times \Omega$ , where  $\Omega \subset E_m$ , and let the intersection of any compact set with  $\Omega$  be closed. Put

$$\varphi(X) = \sup_{Y \in \Omega} f(X, Y), \quad R_\varepsilon(X) = \{Y \mid \varphi(X) - f(X, Y) \leq \varepsilon\}.$$

Suppose that, for some set  $\Omega_1 \subset E_n$ , there exists  $\bar{\varepsilon} > 0$  such that the set  $\bigcup_{X \in \Omega_1} R_{\bar{\varepsilon}}(X)$  is bounded (and then it is also closed). In this case, for every  $X \in \Omega_1$  there exists a bounded closed set  $R(X) = R_0(X) = \{Y \mid \varphi(X) = f(X, Y)\}$ . Then

$$\varphi(X) = \max_{Y \in \Omega} f(X, Y).$$

As shown in <sup>(1-3)</sup>, the function  $\varphi$  is differentiable in any direction  $g$  ( $\|g\| < \infty$ ) at every interior point of  $\Omega_1$ , and

$$\frac{\partial \varphi(X)}{\partial g} = \lim_{\alpha \rightarrow 0+} \frac{\varphi(X + \alpha g) - \varphi(X)}{\alpha} = \max_{Y \in R(X)} \left( \frac{\partial f(X, Y)}{\partial X}, g \right). \quad (1)$$

From (1) it immediately follows that

**Theorem 1.** *In order that at a point  $X$  (let  $X \in \Omega_1$ ) the function  $\varphi$  attain its minimum value on  $E_n$ , it is necessary (and, in the case of convexity of  $\varphi$ , also sufficient) that*

$$\Psi(X) = \min_{g \in G} \partial \varphi(X) / \partial g = 0, \quad (2)$$

where  $G \subset E_n$  is a compact set containing the origin of coordinates as an interior point.

In (2), the unit sphere is taken as  $G$ .

A point satisfying (2) is called a stationary point of the function  $\varphi$  on  $E_n$ . If for some  $G$  at the point  $X$  it turns out that  $\Psi(X) = 0$ , then this will also be so for any other  $G$ .

Since the function  $\Psi(X)$ , generally speaking, is discontinuous, condition (2) cannot be used directly for constructing algorithms for minimizing the function  $\varphi$ . For example, the method of steepest descent, constructed using the necessary condition (2), does not, in general, lead to a stationary point. Therefore let us consider the function

$$\Psi_\varepsilon(X) = \min_{\|g\| \leq 1} \max_{Y \in R_\varepsilon(X)} (\partial f(X, Y) / \partial X, g) \quad (3)$$

(we shall consider only  $\varepsilon \in [0, \bar{\varepsilon}]$ ).

**Theorem 2.** *In order that at a point  $X \in \Omega_1$  the function  $\varphi$  attain its minimum value on  $E_n$ , it is necessary (and, in the case of convexity of  $\varphi$ , also sufficient) that*

$$D(X) = \inf_{\varepsilon \in [0, \bar{\varepsilon}]} h(\varepsilon) \Psi_\varepsilon(X), \quad (4)$$

where  $h(\varepsilon)$  is any continuous strictly increasing function on  $[0, \bar{\varepsilon}]$ , with  $h(0) = 0$ .

Conditions (2) and (4) are equivalent: (4) follows from (2), and (2) follows from (4).

The function  $D(X)$  is continuous; therefore it can be used to construct methods of successive approximations for minimizing  $\varphi$ .

Consider the following method of successive approximations. As the first approximation choose any point  $X_1 \in E_n$  such that the set  $M(X_1) = \{X / \varphi(X) \leq \varphi(X_1)\} \subset \Omega_1$  is bounded. Suppose  $X_k$  has already been found and  $X_k \in M(X_1)$ . If  $D(X_k) = 0$ , then the point  $X_k$  is stationary, and the process of searching for a minimum stops. If, however,  $D(X_k) < 0$  (positive  $D$  cannot occur), then find any  $\varepsilon_k \in [0, \bar{\varepsilon}]$  such that

$$h(\varepsilon_k) \Psi_{\varepsilon_k}(X_k) \leq \frac{1}{s_1} D(X_k),$$

where  $s_1 > 1$  is independent of  $k$ .

Let the vector  $g_k \in G$  be such that

$$\max_{Y \in R_{\varepsilon_k}(X_k)} \left( \frac{\partial f(X_k, Y)}{\partial X}, g_k \right) \leq \frac{1}{s_2} \Psi_{\varepsilon_k}(X_k),$$

where  $s_2 > 1$  is also independent of  $k$ . Clearly,  $\|g_k\| > 0$ . Now consider the ray

$$X_{k\alpha} = X_k - \alpha g_k, \quad \alpha \geq 0,$$

and find  $\alpha_k \in [0, \infty)$  such that

$$\varphi(X_k) - \varphi(X_{k\alpha_k}) \geq [\varphi(X_k) - a_k]/s_3,$$

where

$$a_k = \min_{\alpha \in [0, \infty)} \varphi(X_{k\alpha}), \quad s_3 > 1$$

and  $s_3$  is independent of  $k$ . Put  $X_{k+1} = X_{k\alpha_k}$ . It is clear that if  $D(X_k) < 0$ , then

$$\varphi(X_{k+1}) < \varphi(X_k), \quad X_{k+1} \in M(X_1).$$

Thus we construct a sequence  $\{X_k\}$ . If this sequence consists of a finite number of points, then its rightmost endpoint is a stationary point. If, on the other hand, this sequence consists of an infinite number of terms, then the following holds.

**Theorem 3.** *Every limit point of the sequence  $\{X_k\}$  is a stationary point of the function  $\varphi$  on  $E_n$ .*

**Remark 1.** Applying the method set out above to the problem of mathematical programming and choosing the corresponding function  $h(\varepsilon)$  and set  $G$ , one can obtain algorithms close to the algorithms presented in <sup>(4, 5)</sup>.

§ 2. Let the function  $f(X, Y)$  be continuously differentiable with respect to  $X$  on  $\Omega_1 \times \Omega_2$ , where  $\Omega_1$  is a convex compact subset of  $E_n$ , and  $\Omega_2$  is a compact subset of the space  $E_m$ . The function  $\varphi$  and the set  $R_\varepsilon(X)$  are defined as above. It is required to find

$$\min_{X \in \Omega_1} \varphi(X).$$

**Theorem 4.** *In order that the function  $\varphi$  attain at a point  $X \in \Omega_1$  its minimum value on  $\Omega_1$ , it is necessary (and, in the case of convexity of  $\varphi$ , also sufficient) that*

$$\Psi_1(X) = \min_{Z \in \Omega_1} \max_{Y \in R(X)} (\partial f(X, Y)/\partial X, Z - X) = 0. \quad (5)$$

A point  $X \in \Omega_1$  satisfying (5) will be called a **stationary point** of the function  $\varphi$  on the set  $\Omega_1$ .

Since the function  $\Psi_1(X)$  is not continuous, let us introduce into consideration the function

$$\Psi_{1\varepsilon}(X) = \min_{Z \in \Omega_1} \max_{Y \in R_\varepsilon(X)} (\partial f(X, Y)/\partial X, Z - X), \quad \varepsilon > 0$$

and formulate the necessary condition in the following form.

**Theorem 5.** In order that at the point  $X \in \Omega_1$  the function  $\varphi$  attain its minimum on  $\varphi_1$ , it is necessary (and, in the case of convexity of  $\varphi$ , also sufficient) that

$$D_1(X) = \inf_{\varepsilon \in [0, \bar{\varepsilon}]} h(\varepsilon) \Psi_{1\varepsilon}(X) = 0. \quad (6)$$

The function  $h(\varepsilon)$  in (6) satisfies the conditions formulated in Theorem 2, and  $\bar{\varepsilon} > 0$  is any fixed number. Conditions (5) and (6) are equivalent. The function  $D_1(X)$  is continuous on  $\Omega_1$  and therefore can be used for the development of successive-approximation methods. Let us describe one such possible algorithm.

As a first approximation an arbitrary  $X_1 \in \Omega_1$  is chosen. Suppose that  $X_k \in \Omega_1$  has already been found. If  $D_1(X_k) = 0$ , then the point  $X_k$  is stationary, and the process stops. If, however,  $D_1(X_k) < 0$ , then we proceed as follows.

Find  $\varepsilon_k \in [0, \bar{\varepsilon}]$  such that

$$h(\varepsilon_k) \Psi_{1\varepsilon_k}(X_k) \leq \frac{1}{s_1} D_1(X_k)$$

and a point  $z_k \in \Omega_1$

$$\max_{Y \in R_{\varepsilon_k}(X_k)} (\partial f(X_k, Y)/\partial X, Z_k - Y_k) \leq \frac{1}{s_2} \Psi_{1\varepsilon_k}(X_k).$$

Now consider the segment

$$X_{k\alpha} = \alpha X_k + (1 - \alpha) Z_k; \quad \alpha \in [0, 1]; \quad X_{k\alpha} \in \Omega_1$$

and find  $\alpha_k \in [0, 1]$  such that

$$\varphi(X_k) - \varphi(X_{k\alpha_k}) \geq [\varphi(X_k) - a_k]/s_3, \quad a_k = \min_{\alpha \in [0, 1]} \varphi(X_{k\alpha}).$$

Above,  $s_1, s_2, s_3$  are fixed, independent of  $k$ , and  $s_i > 1$  ( $i = 1, 2, 3$ ). Put  $X_{k+1} = X_{k\alpha_k}$ . Clearly,  $X_{k+1} \in \Omega_1$  and  $\varphi(X_{k+1}) < \varphi(X_k)$  if  $D(X_k) < 0$ . We then continue analogously. If the sequence  $\{X_k\}$  constructed in this way consists

of a finite number of points, then its rightmost point is stationary; otherwise the following is true.

**Theorem 6.** Any limit point of the sequence  $\{X_k\}$  constructed by the method described above is a stationary point of the function  $\varphi$  on the set  $\Omega_1$ .

**Remark 2.** The algorithms described above may be applied to solve a number of practically important problems, for example, optimal-control problems in the presence of constraints on phase coordinates.

§ 3. The algorithms constructed above are first-order methods, since they use only the concept of the first directional derivative of the function  $\varphi$ . It may be useful to find the second directional derivative of  $\varphi$  in the direction  $g$ .

Let the function  $f(X, Y)$  be given and twice continuously differentiable with respect to  $X$  and  $Y$  on the open bounded set  $S_1 \times S_2$ , where  $S_1 \subset E_n$ ,  $S_2 \subset E_m$ . Form the function

$$\varphi(X) = \max_{Y \in \Omega} f(X, Y),$$

where  $\Omega$  is a closed bounded set,  $\Omega \subset S_2$ . The function  $\varphi$  is defined on  $S_1$ . Fix  $g \in E_n$ . Introduce into consideration the sets  $R(X)$  and

$$R_2(X, g) = \{Y \mid Y \in R(X), (\partial f(X, Y)/\partial X, g) = \max_{Z \in R(X)} (\partial f(X, Z)/\partial X, g)\}. \quad (7)$$

We shall call a vector  $V \in E_m$  an admissible direction in the broad sense at the point  $Y \in \Omega$  if, for every  $\varepsilon > 0$ , there exist a point  $V_\varepsilon \in S_\varepsilon(V)$  and a number  $\alpha_\varepsilon \in (0, \varepsilon]$  such that  $Y + \alpha_\varepsilon V_\varepsilon \in \Omega$ . Here  $S_\varepsilon(V) = \{Z \mid Z \in E_m, \|Z - V\| \leq \varepsilon\}$ . We denote by  $M(Y)$  the cone of admissible directions in the broad sense. It is closed.

Suppose that, for all  $Y \in R(X)$ ,

$$\max_{V \in Q(X, Y)} \left( \frac{\partial^2 f(X, Y)}{\partial Y^2} V, V \right) \leq -m(X) \|V\|^2, \quad (8)$$

where  $m(X) > 0$ ,  $Q(X, Y) = \{V \mid V \in M(Y), (V, \partial f(X, Y)/\partial Y) = 0\}$ .

Then the following is true.

**Theorem 7.** The function  $\varphi$  is twice differentiable with respect to the direction  $g$  at the point  $X$ , and

$$\partial^2 \varphi(X)/\partial g^2 = \max_{Y \in R_2(X, g)} \max_{V \in Q(X, Y)} \left\{ \left( \frac{\partial^2 f(X, Y)}{\partial X^2} g, g \right) + \right.$$

$$+ 2 \left( \frac{\partial^2 f(X, Y)}{\partial Y \partial X} g, V \right) + \left( \frac{\partial^2 f(X, Y)}{\partial Y^2} V, V \right) \Bigg\}. \quad (9)$$

Thus,

$$\varphi(X + \alpha g) = \varphi(X) + \alpha \partial \varphi(X) / \partial g + \frac{1}{2} \alpha^2 \partial^2 \varphi(X) / \partial g^2 + O_g(\alpha^2). \quad (10)$$

In particular, if the point  $Y \in R(X)$  is an interior point of  $\Omega$ , then  $Q(X, Y) = E_m$ , and condition (8) means strict negative definiteness of the matrix  $A(X, Y) = \partial^2 f(X, Y) / \partial Y^2$ . In this case there exists an inverse matrix  $A^{-1}(X, Y)$ , and for this  $Y$  the inner maximum in (9) is attained at

$$V = - \left( \frac{\partial^2 f(X, Y)}{\partial Y^2} \right)^{-1} \left( \frac{\partial^2 f(X, Y)}{\partial Y, \partial X} g \right).$$

If, on the other hand,  $Y \in R(X)$  is an isolated point of the set  $\Omega$ , then  $Q(X, Y) = \{0\}$ .

**Remark 3.** Condition (8) is not very restrictive, since, by virtue of the fact that  $Y \in R(X)$  (i.e.,  $f(X, Y) = \max_{Z \in \Omega} f(X, Z)$ ), one must have

$$\max_{V \in Q(X, Y)} (A(X, Y)V, V) \leq 0.$$

Leningrad State University  
named after A. A. Zhdanov

Received  
13 XI 1968

## REFERENCES

1. V. F. Dem'yanov, *Vestn. Leningrad. Univ.*, No. 7, 21 (1966).
2. V. F. Dem'yanov, *Kibernetika*, No. 6, 58 (1966); No. 3, 62 (1967).
3. B. N. Pshenichnyi, *Kibernetika*, No. 6, 54 (1967).
4. G. Zoutendijk, *Methods of Feasible Directions*, IL, 1963.
5. S. I. Zukhovitskii, R. A. Polyak, M. E. Primak, *DAN*, 163, No. 2, 282 (1965).

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

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