



---

Soviet-era science, translated into English

# FINDING SADDLE POINTS ON POLYHEDRA

1970

SovietRxiv

---

View the original and related papers at <https://sovietrxiv.org/items/ru-197001.84211>

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

**Abstract**

**Full Text**

UDC 519.8

**V. F. Dem'yanov**

## FINDING SADDLE POINTS ON POLYHEDRA

*(Presented by Academician L. V. Kantorovich, 13 X 1969)*

For finding saddle points, various methods were proposed in (1-4). Below we consider the case in which the sets of variation of the variables are polyhedra. Many other sets can often be described sufficiently well by polyhedra.

Let, in the Euclidean space  $E_n$ , the set

$$\Omega_1 = \{x \mid (A_i, x) + a_i \leq 0, i \in \overline{1, N_1}\},$$

and in the space  $E_m$ , the set

$$\Omega_2 = \{y \mid (B_j, y) + b_j \leq 0, j \in \overline{1, N_2}\}.$$

Without loss of generality we assume that

$$\|A_i\| = 1, i \in \overline{1, N_1}; \quad \|B_j\| = 1, j \in \overline{1, N_2},$$

and that the sets  $\Omega_1$  and  $\Omega_2$  are bounded.

On  $\Omega_1 \times \Omega_2$  there is given a twice continuously differentiable function  $f(x, y)$ , strictly convex in  $x$  and strictly concave in  $y$  on  $\Omega_1 \times \Omega_2$ , i.e., there exist  $M_1 > 0$  and  $M_2 > 0$  such that for all  $[x, y] \in \Omega_1 \times \Omega_2$  and all  $V \in E_n, W \in E_m$ , the inequalities

$$\left(V, \frac{\partial^2 f(x, y)}{\partial x^2} V\right) \geq M_1 \|V\|^2, \quad - \left(W, \frac{\partial^2 f(x, y)}{\partial y^2} W\right) \geq M_2 \|W\|^2.$$

It is required to find a saddle point of the function  $f$  on  $\Omega_1 \times \Omega_2$ , i.e., a point  $[x^*, y^*] \in \Omega_1 \times \Omega_2$  satisfying the inequalities

$$f(x^*, y) \leq f(x^*, y^*) \leq f(x, y^*) \tag{1}$$

for all  $x \in \Omega_1, y \in \Omega_2$ .

Let  $x \in \Omega_1$ ,  $y \in \Omega_2$ . Introduce the index sets

$$Q_1(x) = \{i \mid i \in \overline{1, N_1}, (A_i, x) + a_i = 0\},$$

$$Q_2(y) = \{j \mid j \in \overline{1, N_2}, (B_j, y) + b_j = 0\},$$

as well as the cones

$$\Gamma_1^+(x) = \left\{ g \mid g \in E_n, g = - \sum_{i \in Q_1(x)} \alpha_i A_i, \alpha_i \geq 0 \right\},$$

$$\Gamma_2^+(y) = \left\{ q \mid q \in E_m, q = - \sum_{j \in Q_2(y)} \beta_j B_j, \beta_j \geq 0 \right\}.$$

If  $Q_1(x) = \emptyset$ , then  $\Gamma_1^+(x) = \{0\}$ , and if  $Q_2(y) = \emptyset$ , then  $\Gamma_2^+(y) = \{0\}$ .

We also introduce the functions

$$d_1(x, y) = \min_{z \in \Gamma_1^+(x)} \|z - \partial f(x, y)/\partial x\| = \|z_1(x, y) - \partial f(x, y)/\partial x\|,$$

$$d_2(x, y) = \min_{z \in \Gamma_2^+(y)} \|z + \partial f(x, y)/\partial y\| = \|z_2(x, y) + \partial f(x, y)/\partial y\|.$$

We note that for any  $[x, y] \in \Omega_1 \times \Omega_2$  the points  $z_1(x, y)$  and  $z_2(x, y)$  are unique.

If  $d_1(x, y) > 0$ , then the direction

$$g(x, y) = \|z_1(x, y) - \partial f(x, y)/\partial x\|^{-1} (z_1(x, y) - \partial f(x, y)/\partial x)$$

is the direction of steepest descent of the function  $f_1(z) \equiv f(z, y)$  at the point  $z = x$ ; and if  $d_2(x, y) > 0$ , then the direction

$$q(x, y) = \|z_2(x, y) + \partial f(x, y)/\partial y\|^{-1} (z_2(x, y) + \partial f(x, y)/\partial y)$$

is the direction of steepest descent of the function  $f_2(z) \equiv -f(x, z)$  at the point  $z = y$ .

From Theorem 1 of paper (2) the following can be proved.

**Theorem 1.** *In order that the point  $[x^*, y^*] \in \Omega_1 \times \Omega_2$  be a saddle point of the function  $f$  on the set  $\Omega_1 \times \Omega_2$ , it is necessary and sufficient that*

$$d_1(x^*, y^*) = d_2(x^*, y^*) = 0. \quad (2)$$

Geometrically, the necessary condition (2) means that at the saddle point there must be

$$\partial f(x^*, y^*)/\partial x \in \Gamma_1^+(x^*), \quad \partial f(x^*, y^*)\partial y \in \Gamma_2^+(y^*). \quad (3)$$

For any  $\varepsilon \geq 0$ , consider the sets

$$Q_{1\varepsilon}(x) = \{i \mid i \in \overline{1, N_1}, -\varepsilon \leq (A_i, x) + a_i \leq 0\},$$

$$Q_{2\varepsilon}(y) = \{j \mid j \in \overline{1, N_2}, -\varepsilon \leq (B_j, y) + b_j \leq 0\}$$

and the cones

$$\Gamma_{1\varepsilon}^+(x) = \left\{ g \mid g \in E_n, \quad g = - \sum_{i \in Q_{1\varepsilon}(x)} \alpha_i A_i, \quad \alpha_i \geq 0 \right\},$$

$$\Gamma_{2\varepsilon}^+(y) = \left\{ q \mid q \in E_m, \quad q = - \sum_{j \in Q_{2\varepsilon}(y)} \beta_j B_j, \quad \beta_j \geq 0 \right\}.$$

Let

$$d_{1\varepsilon}(x, y) = \min_{z \in \Gamma_{1\varepsilon}^+(x)} \|z - \partial f(x, y)/\partial x\| = \|z_{1\varepsilon}(x, y) - \partial f(x, y)/\partial x\| \equiv \|g_\varepsilon(x, y)\|,$$

$$d_{2\varepsilon}(x, y) = \min_{z \in \Gamma_{2\varepsilon}^+(y)} \|z + \partial f(x, y)/\partial y\| = \|z_{2\varepsilon}(x, y) + \partial f(x, y)/\partial y\| \equiv \|q_\varepsilon(x, y)\|.$$

It can be shown that there is always a representation of the point  $z_{1\varepsilon}(x, y)$  in the form

$$z_{1\varepsilon}(x, y) = - \sum_{i \in Q_{1\varepsilon}(x)} \alpha_i(x, y) A_i,$$

such that if  $\alpha_i(x, y) > 0$ , then necessarily

$$i \in \overline{Q}_{1\varepsilon}(x, y) \equiv \{i \mid i \in Q_{1\varepsilon}(x), (A_i, g_\varepsilon(x, y)) = 0\}.$$

Similarly, the existence of such a representation of the point  $z_{2\varepsilon}(x, y)$  in the form

$$z_{2\varepsilon}(x, y) = - \sum_{j \in Q_{2\varepsilon}(y)} \beta_j(x, y) B_j,$$

is shown, so that for those  $j$  for which  $\beta_j(x, y) > 0$ , one has

$$j \in \overline{Q}_{2\varepsilon}(x, y) \equiv \{j \mid j \in Q_{2\varepsilon}(y), (B_j, q_\varepsilon(x, y)) = 0\}.$$

For the remaining  $i \in Q_{1\varepsilon}(x)$  and  $j \in Q_{2\varepsilon}(y)$  we have

$$(A_i, g_\varepsilon(x, y)) \equiv (A_i, z_{1\varepsilon}(x, y) - \partial f(x, y)/\partial x) \leq 0,$$

$$(B_j, q_\varepsilon(x, y)) \equiv (B_j, z_{2\varepsilon}(x, y) + \partial f(x, y)/\partial y) \leq 0.$$

Of course, there may also exist other representations of the points  $z_{1\varepsilon}(x, y)$  and  $z_{2\varepsilon}(x, y)$  (for example, if among the vectors  $\{A_i\}$  (or  $\{B_j\}$  there are linearly dependent ones). A point  $[x, y] \in \Omega_1 \times \Omega_2$  will be called an  $\varepsilon$ -saddle point of the function  $f$  on the set  $\Omega_1 \times \Omega_2$  if

$$d_{1\varepsilon}(x, y) = d_{2\varepsilon}(x, y) = 0. \quad (4)$$

Let us now consider the function

$$d_\varepsilon(x, y) = \frac{1}{2} [d_{1\varepsilon}^2(x, y) + d_{2\varepsilon}^2(x, y)].$$

It is clear that condition (4) is equivalent to the condition  $d_\varepsilon(x, y) = 0$ .

Fix  $\varepsilon > 0$ . We describe the following method of successive approximations for finding an  $\varepsilon$ -saddle point.

As a first approximation choose an arbitrary point  $[x_1, y_1] \in \Omega_1 \times \Omega_2$ . Suppose that  $[x_k, y_k] \in \Omega_1 \times \Omega_2$  has already been found. If  $d_\varepsilon(x_k, y_k) = 0$ , then the point  $[x_k, y_k]$  is an  $\varepsilon$ -saddle point, and the process terminates. If  $d_\varepsilon(x_k, y_k) > 0$ , then consider the rays

$$x_{k\alpha} = x_k + \alpha \left( z_{1\varepsilon}(x_k, y_k) - \frac{\partial f(x_k, y_k)}{\partial x} \right) \equiv x_k + \alpha g_k,$$

$$y_{k\alpha} = y_k + \alpha \left( z_{2\varepsilon}(x_k, y_k) + \frac{\partial f(x_k, y_k)}{\partial y} \right) \equiv y_k + \alpha q_k,$$

find  $\alpha_k > 0$  such that

$$d_\varepsilon(x_{k\alpha_k}, y_{k\alpha_k}) = \min_{\alpha \geq 0} d_\varepsilon(x_{k\alpha}, y_{k\alpha}), \quad x_{k\alpha} \in \Omega_1, y_{k\alpha} \in \Omega_2,$$

and set

$$x_{k+1} = x_{k\alpha_k}, \quad y_{k+1} = y_{k\alpha_k}.$$

It is clear that  $(x_{k+1}, y_{k+1}) \in \Omega_1 \times \Omega_2$ ,  $d_\varepsilon(x_{k+1}, y_{k+1}) \leq d_\varepsilon(x_k, y_k)$ . We then proceed analogously.

Thus we construct a sequence  $\{[x_k, y_k]\} \subset \Omega_1 \times \Omega_2$ . If this sequence contains a finite number of points, then the last point obtained, by construction, is an  $\varepsilon$ -saddle point of the function  $f$  on  $\Omega_1 \times \Omega_2$ . Otherwise, the following holds.

**Theorem 2.** *Every limit point of the sequence  $\{[x_k, y_k]\}$  is an  $\varepsilon$ -saddle point of the function  $f$  on the set  $\Omega_1 \times \Omega_2$ .*

If we are interested in finding a saddle point (and not an  $\varepsilon$ -saddle point), then we can use the following algorithm.

Fix any  $\varepsilon_1 > 0$  and, using the method described above with  $\varepsilon = \varepsilon_1$ , in a finite number of steps find a point  $[x_1, y_1] \in \Omega_1 \times \Omega_2$  such that

$$d_{\varepsilon_1}(x_1, y_1) \leq a\varepsilon_1,$$

where  $a > 0$  is any fixed number independent of  $k$ . Now set  $\varepsilon_2 = \frac{1}{2}\varepsilon_1$  and take as the first approximation the point  $[x_1, y_1]$  obtained; applying the basic algorithm again with  $\varepsilon = \varepsilon_2$ , in a finite number of steps we obtain a point  $[x_2, y_2]$  such that  $d_{\varepsilon_2}(x_2, y_2) \leq a\varepsilon_2$ . We continue analogously. It is not difficult to show that the sequence  $\{[x_k, y_k]\}$  tends to a saddle point of the function  $f$  on  $\Omega_1 \times \Omega_2$ .

**Remark.** If the function  $f(x, y)$  is convex-concave, then instead of it one may consider the strictly convex-concave function

$$F(x, y) = f(x, y) + cx^2 - dy^2,$$

where  $c > 0$  and  $d > 0$  are arbitrary numbers. For small  $c$  and  $d$ , the function  $F(x, y)$  does not differ much from  $f(x, y)$ .

Leningrad State University  
named after A. A. Zhdanov

Received  
7 X 1969

## REFERENCES

1. V. F. Demyanov, DAN, 177, No. 1, 21 (1967).
2. V. F. Demyanov, Vestn. Leningrad Univ., 19, 25 (1967).
3. D. M. Danskin, *Theory of Max-Min*, Moscow, 1963, p. 123.
4. J. Robinson, in: *Matrix Games*, Moscow, 1961, p. 110.

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

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