ISSN 0005-1179, Automation and Remote Control, 2015, Vol. 76, No. 11, pp. 1966–1975. © Pleiades Publishing, Ltd., 2015.
Original Russian Text © I.Ya. Zabotin, R.S. Yarullin, 2015, published in Avtomatika i Telemekhanika, 2015, No. 11, pp. 76–88.

SYSTEM ANALYSIS AND OPERATIONS RESEARCH

Cutting-Plane Method Based on Epigraph Approximation with Discarding the Cutting Planes

I. Ya. Zabotin and R. S. Yarullin

Kazan (Volga Region) Federal University, Kazan, Russia
e-mail: IYaZabotin@mail.ru, YarullinRS@gmail.com

Received January 18, 2015

Abstract—Propose a method for solving a mathematical programming problem from the class of cutting methods. In our method, on each step the epigraph of the objective function is embedded into a specifically constructed polyhedral set, and on this set an auxiliary linear function is minimized in order to construct the iteration point. Proposed method does not require that each approximation set is embedded in the previous ones. This feature lets us periodically discard additional constraints that form the approximation sets obtained during the solution process. Prove the method's convergence and discuss possible implementations.

DOI: 10.1134/S0005117915110065

1. INTRODUCTION

The class of cutting methods for function minimization is rather wide (see, e.g., [1–11]). However, investigations dealing with methods from the mentioned class still attract the attention of researchers in both continuous and discrete optimization.

Cutting methods are characterized by the following. To construct iteration points, such methods sequentially embed either the set of constraints for the original problem or the objective function's epigraph into some sets with a simpler structure. Each of the embedding, or, as they are usually called, approximating sets is constructed from the previous one by cutting off some subset of it, usually with hyperplanes.

A major problem that arises in numerical implementation of such methods is as follows. As the number of iterations grows, the number of cutting planes also increases unboundedly. This means that there are more and more inequalities that define approximating sets, and so the computational complexity of solving the problems of finding iteration points also raises from iteration to iteration.

Previously, we have proposed one approach to constructing cutting algorithms with embedding the admissible set that could periodically discard already used cutting planes [4]. The idea considered in [4] is extended in this work to cutting algorithms that use the embedding operation not for the constraint region but for the epigraph of the objective function of the problem. The developed in this work cutting method is based on the ideas of a known level-set method for conditional minimization of convex functions [11]. One characteristic feature of the proposed method is that it can update the sets where the epigraph is embedded by periodically discarding any number of any previously constructed cutting planes.

2. PROBLEM SETTING

Let D be a convex bounded closed set in the n-dimensional Euclidean space R_n, $f(x)$ be a function convex in R_n. Solve the problem

$$\min\{f(x) : x \in D\}. \tag{1}$$

Let $f^* = \min\{f(x) : x \in D\}$, $X^* = \{x \in D : f(x) = f^*\}$, $x^* \in X^*$, $\partial f(x)$ be the subdifferential of function $f(x)$ at point $x \in R_n$. Set $K = \{0, 1, \ldots\}$, $\mathrm{epi}\,(f, D) = \{(x, \gamma) \in R_{n+1} : x \in D, \gamma \geqslant f(x)\}$.

3. THE CUTTING METHOD AND DISCUSSION

The proposed method for solving problem (1) construct a sequence of approximations $\{x_i\}$, $i \in K$, $\{z_k\}$, $k \in K$, and works as follows.

Choose a convex closed set $M_0 \subset R_{n+1}$ such that $\mathrm{epi}\,(f, D) \subset M_0$. Define numbers ε_0, α_0, β_{-1} satisfying conditions

$$\varepsilon_0 \geqslant 0, \quad \alpha_0 \leqslant f^* \leqslant \beta_{-1}.$$

Let $\delta_0 = +\infty$, $i = 0$, $k = 0$.

1. Find a solution (y_i, γ_i), where $y_i \in R_n$, $\gamma_i \in R_1$, of the following problem:

$$\min\{\gamma : (x, \gamma) \in M_i,\ x \in D,\ \gamma \geqslant \alpha_i\}. \tag{2}$$

If

$$f(y_i) = \gamma_i, \tag{3}$$

then y_i is a solution of problem (1), and the process is over.

2. Choose a number λ_i defined by condition

$$0 < \lambda_i \leqslant \overline{\lambda} < 1, \tag{4}$$

let

$$\beta_i = \min\{\beta_{i-1}, \delta_i\}, \tag{5}$$
$$l_i = (1 - \lambda_i)\gamma_i + \lambda_i \beta_i. \tag{6}$$

Construct the set U_i as follows. The set U_i includes each of the points $x \in D$ for which for some $\gamma_x \leqslant l_i$ it holds that $(x, \gamma_x) \in M_i$.

3. Choose a point $x_i \in U_i$.

4. If the inequality

$$f(x_i) - \gamma_i > \varepsilon_k, \tag{7}$$

then let

$$Q_i = M_i \tag{8}$$

and go to step 5. Otherwise let $i_k = i$,

$$z_k = x_{i_k}, \quad \sigma_k = \gamma_{i_k}, \tag{9}$$

choose a convex closed set $Q_i \subset R_{n+1}$ such that

$$(x^*, f^*) \in Q_i, \tag{10}$$

specify $\varepsilon_{k+1} \geqslant 0$, and increment k by one.

5. Let

$$M_{i+1} = Q_i \bigcap \{(x, \gamma) \in R_{n+1} : f(x_i) + \langle a_i, x - x_i \rangle \leqslant \gamma\}, \tag{11}$$

where $a_i \in \partial f(x_i)$.

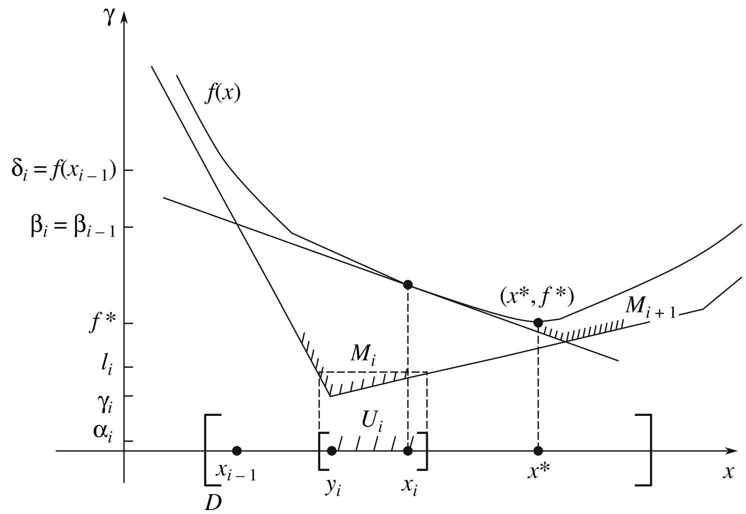

Figure.

6. Let

$$\alpha_{i+1} = \gamma_i, \tag{12}$$
$$\delta_{i+1} = f(x_i). \tag{13}$$

Increment i by one and go to 1.

Before we discuss the properties of sequences constructed with this algorithm, we begin with some remarks.

First of all we note that we will justify that the admissible set (2) and set U_i are nonempty, and justify the stopping criterion used in step 1 of the method, at a later point.

The figure illustrates the ways to construct auxiliary set $U_i \subset D$ and the set M_{i+1} based on M_i on the ith step ($i > 0$).

If D is a polyhedral set and M_0, Q_i, $i \in K$, are chosen to be such that M_{i+1}, $i \in K$, of the form (11), are also polyhedral, then (2) for all $i \in K$ is a linear programming problem. Let us discuss the ways to define the sets M_0, Q_i.

Remark 1. There are a lot of possibilities for the choice of M_0. We first note that we can let

$$M_0 = R_{n+1}.$$

In this case the pair (y_0, γ_0), where y_0 is any point of D and $\gamma_0 = \alpha_0$, can be taken as a solution of problem (2) for $i = 0$. We can define the set M_0 with one linear inequality

$$\langle c, x\rangle - \gamma \leqslant \langle c, u\rangle - f(u), \tag{14}$$

where $u \in R_n$, $c \in \partial f(u)$, or with a group of such inequalities. If

$$f(x) = \max_{j\in J} f_j(x), \tag{15}$$

where J is a finite set of indices, $f_j(x)$, $j \in J$, is a function convex in R_n, then we can let

$$M_0 = \mathrm{epi}\,(f_{j_0}, R_n), \tag{16}$$

where $j_0 \in J$, or similar to (14) we can define M_0 with an inequality $\langle a_{j_0}, x\rangle - \gamma \leqslant \langle a_{j_0}, u\rangle - f_{j_0}(u)$, where $a_{j_0} \in \partial f_{j_0}(u)$. If the function $f_{j_0}(x)$ is linear than choice (16) is convenient from the practical point of view.

Remark 2. If we have no information on the value of f^*, we can choose the number α_0 as, for instance, a solution of the minimization problem for variable γ under constraint (14), where $u \in D$, and according to condition $x \in D$. We can also take α_0 as the minimal value of function $f_{j_0}(x)$, $j_0 \in J$ on the set D if $f(x)$ has form (15). It is easy to find such a value if D is a polyhedron and the function $f_{j_0}(x)$ is linear.

It is natural to call the sets M_i approximating for the objective function's epigraph, and planes $f(x_i) + \langle a_i, x - x_i \rangle = \gamma$ or $\langle a_i, x \rangle - \gamma = \langle a_i, x_i \rangle - f(x_i)$ that construct on the ith iteration the next approximating set M_{i+1} by (11), cutting planes.

The approximation quality for the epigraph of the function with the sets M_i is defined in the neighborhood of points x_i by the value $f(x_i) - \gamma_i$. On iterations when x_i satisfies inequality (7), approximation quality is considered to be insufficient to fix the point z_k. On such iterations, the sets Q_i are chosen as (8), and then according to (11) in the construction of the set M_{i+1} cutting planes are accumulated. At some step $i = i_k$, as we show below, the difference $f(x_i) - \gamma_i$ satisfies an inequality opposite to (7). In this case, approximation quality becomes satisfactory, the point z_k becomes fixed in form (9), and due to the virtually arbitrary choice of the set $Q_i = Q_{i_k}$ we can update the set M_{i_k+1}.

Let us now proceed to discussing how to specify the sets Q_i in such a way that would let us update the approximating sets. The sets Q_i, due to condition (10) and inclusion

$$(x^*, f^*) \in M_i, \tag{17}$$

which is proven in Lemma 1 (see Appendix), can be specified for all $i \in K$ as (8), regardless of whether inequalities (7) hold. But in this case the number of cutting planes that define approximating sets unboundedly from step to step, and auxiliary problems (2) become computationally very hard. We now give the main conceptual remark regarding the definition of sets Q_i based on which there appears a possibility to periodically update approximating sets by discarding any number of cutting planes.

Remark 3. For those pairs of indices i, k for which inequalities

$$f(x_i) - \gamma_i \leqslant \varepsilon_k \tag{18}$$

hold, condition (10) allows for a lot of possibilities in the choice of the sets $Q_i = Q_{i_k}$, and therefore in the choice of the sets M_{i+1}. In case of (18) we can let, for instance,

$$Q_i = Q_{i_k} = R_{n+1}. \tag{19}$$

Then M_{i+1} is only defined by inequality $f(x_i) + \langle a_i, x - x_i \rangle \leqslant \gamma$, i.e., in case of (19) all cutting planes obtained by step $i = i_k$ are not used in the construction of M_{i+1}. Further, if condition (18) holds then the sets $Q_i = Q_{i_k}$ can also be specified as

$$Q_i = M_{r_i}, \tag{20}$$

where $0 \leqslant r_i \leqslant i = i_k$, since for all $r_i = 0, \ldots, i$ due to (17) inclusion (10) holds. According to (19), (20), for every $i = i_k$ we can discard in the construction of M_{i_k+1} any number of any cutting planes accumulated by step i_k.

As we show in Theorem 3 (see below), for every $k \in K$, if numbers ε_k are chosen to be positive there exists a point (x_i, γ_i) satisfying (18), and, consequently, there is a possibility to define Q_i as (19) or (20).

Let us discuss the relation of our method with previously known approaches of Kelley [7, 9] and the level-set method [7, 11]. Suppose that on the preliminary step of the method the set M_0 is chosen to be different from R_{n+1} and is defined by only one inequality of form (14). Let, moreover, that in the method we set $\varepsilon_0 = 0$, i.e., in step 4 for all $i \in K$ we define $Q_i = M_i$, points z_k are not constructed, and numbers ε_k, $k \geqslant 1$, are not chosen. If for this choice of M_0 and ε_0 we choose approximation x_i for all $i \geqslant 1$ as the projection of point x_{i-1} to the set U_i, then the method will coincide with the level-set method. Note that if $M_0 \neq R_{n+1}$ and $Q_{i_k} \neq R_{n+1}$, $k \in K$, then there is no need in the numbers α_i and constraint $\gamma \geqslant \alpha_i$ in (2).

If under our assumptions for all $i \in K$ we let in the method $x_i = y_i$, the algorithm will coincide with Kelley's method. Note that such a choice of points x_i on step 3 of the method is possible since Lemma 5 shows inclusions $y_i \in U_i$ for all $i \in K$. (All lemmas are given in the Appendix.)

Since the proposed method has a possibility for periodic updates of approximating sets, we have thus justified the technique of including discard procedures for accumulated cutting planes both in Kelley's method and in the method of levels.

Lemma 2 proves a conceptual property of the sequence $\{\alpha_i\}$, $i \in K$, constructed by the method, and based on this property and Lemma 1 we prove in Lemma 3 that auxiliary problems (2) are feasible for all $i \in K$.

Lemma 3 easily implies Lemma 4, which shows a property of sequence $\{\gamma_i\}$, $i \in K$, constructed with our method. With Lemma 4 we can also justify the following optimality criterion for point y_i.

Theorem 1. *Suppose that for some $i \in K$ equality* (3) *holds. Then $y_i \in X^*$.*

(See the proof in the Appendix.)

4. CONVERGENCE OF THE PROPOSED METHOD

We begin by stating a result that deals with the case when in the construction of $\{x_i\}$, $i \in K$, starting from some index, the sets Q_i are chosen as (8).

Theorem 2. *Suppose that sequences $\{x_i\}$, $\{\gamma_i\}$, $i \in K$, have been constructed by the method with the condition that starting from some index $\tilde{i} > 0$ the sets Q_i are chosen by* (8). *Then it holds that*

$$\lim_{i \in K} f(x_i) = f^*, \quad \lim_{i \in K} \gamma_i = f^*.$$

Note that by Theorem 2 and a known theorem [12, p. 74] the sequence $\{x_i\}$, $i \in K$, converges to the set X^*.

Statements of Theorem 2 have been proven under the assumption that in the method for all $i \in K$

$$M_{i+1} = M_i \bigcap \{(x, \gamma) \in R_{n+1} : f(x_i) + \langle a_i, x - x_i \rangle \leqslant \gamma\}. \tag{21}$$

In this case, in the construction of sequences $\{x_i\}$, $\{\gamma_i\}$, $i \in K$, the number of cutting planes increases unboundedly, and there are no updates of approximating sets. Note also that if we let $\varepsilon_0 = 0$ in the method then sets Q_i and M_{i+1} for all $i \in K$ will have the form (8) and (21) respectively, and none of the points $z_k = x_{i_k}$ will be fixed.

Considering these remarks we will pass to research the sequences $\{x_i\}$, $\{\gamma_i\}$, $i \in K$, in the construction of which we applied (9) and updated the sets M_{i+1} with a corresponding choice of the sets Q_i infinitely often. In other words, we investigate the properties of sequences $\{z_k\}$, $\{\sigma_k\}$, $k \in K$.

The existence of sequences $\{z_k\}$, $\{\sigma_k\}$, $k \in K$, under the additional condition on the choice of the numbers ε_k is justified by the following statement.

Theorem 3. *Suppose that sequences* $\{x_i\}$, $\{\gamma_i\}$, $i \in K$, *are constructed by the method under the assumption that*

$$\varepsilon_k > 0 \quad \forall k \in K. \tag{22}$$

Then for each $k \in K$ *there exists an index* $i = i_k$ *for which equalities* (9) *hold.*

The fact that $\{f(z_k)\}$, $\{\sigma_k\}$, $k \in K$, converge to the solution is justified under one more additional assumption on the sequence $\{\varepsilon_k\}$, $k \in K$, by the following.

Theorem 4. *Suppose that in the method, the numbers* ε_k, $k \in K$, *are chosen by* (22), *and, moreover,*

$$\varepsilon_k \to 0, \quad k \to \infty. \tag{23}$$

Then sequences $\{z_k\}$, $\{\sigma_k\}$, $k \in K$, *satisfy*

$$\lim_{k \in K} f(z_k) = f^*, \quad \lim_{k \in K} \sigma_k = f^*. \tag{24}$$

Let us briefly describe how one can choose the sequence $\{\varepsilon_k\}$, $k \in K$. Note that numbers ε_k, $k \geqslant 1$, just like the number ε_0, can be specified at the preliminary step of the method. However, in this case $\{\varepsilon_k\}$, $k \in K$, will not be adapted to the minimization process. Therefore, the method allows to choose ε_k, $k \geqslant 1$, in the process of constructing the approximations z_k. For instance, letting ε_0 be arbitrarily large, we can choose on step 4

$$\varepsilon_{k+1} = \tau_k(f(z_k) - \sigma_k), \quad k \geqslant 0, \tag{25}$$

where $\tau_k \in (0, 1)$. For such a sequence $\{\varepsilon_k\}$, $k \in K$, both (22) and condition (23) hold if $\tau_k \to 0$, $k \in K$.

In conclusion we represent the estimates dealing with the accuracy of our solutions for the problem. According to (A.3), for all $i \in K$, including $i = i_k$, it holds that $\gamma_i \leqslant f^* \leqslant f(x_i)$. Consequently, due to (9)

$$0 \leqslant f(z_k) - f^* \leqslant f(z_k) - \sigma_k \leqslant \varepsilon_k, \quad k \in K. \tag{26}$$

Under the additional assumption that function $f(x)$ is strongly convex with strong convexity constant μ for all $k \in K$ it holds that

$$||z_k - x^*|| \leqslant \sqrt{\frac{\varepsilon_k}{\mu}},$$

which follows from the well-known inequality $\mu||z_k - x^*||^2 \leqslant f(z_k) - f(x^*)$ (see, e.g., [12, p. 182]) and estimate (26).

To verify that the method works, we have conducted a number of numerical experiments on test examples with the number of variables ranging from 2 to 50. Objective functions in these examples were convex quadratic, and constraints is defined by parallelepipeds. Each example was solved without procedures for discarding cutting planes if we let $\varepsilon_0 = 0$, and with discarding if we chose ε_k in the form (25). A part of the test problems was solved by choosing approximations x_i as $x_i = y_i$ and as projections of points x_{i-1} on the sets U_i.

Iterations with indices $i = i_k$ have employed different methods for updating approximating sets. In particular, the choice of $Q_{i_k} = M_0$ (i.e., full update of the approximating sets) has proven to be not very efficient. The best method appears to be discarding all additional constraints that are "inactive" at point x_{i_k}. Under this method of discarding, we achieved a given accuracy of solutions 2–3 times faster as compared to choosing the sets Q_i in the form (8) for all i. The higher was the problem dimension, the greater was the difference in time.

5. CONCLUSION

To solve convex programming problems, we have presented a cutting method with epigraph approximation that admits periodic updates of approximating sets by discarding cutting planes. We have described several properties of the method, proposed implementations for it, compared our approach with known cutting methods. On test examples, we have compared several algorithms for the proposed approach. Experimental results have proven that the procedures for updating approximating sets that we have developed and included in the method are of practical importance. Namely, a given accuracy of solutions in all test problems was achieved significantly faster if the update procedures were used.

APPENDIX

Lemma 1. *Inclusion* (17) *holds for all* $i \in K$.

Proof of Lemma 1. For $i = 0$ inclusion (17) holds by the choice of M_0. Suppose now that (17) holds for every fixed $i = l \geqslant 0$. Let us show that (17) holds for $i = l + 1$, then the statement of the Lemma will be proven. Indeed, due to (8), (10) and the induction hypothesis the following inclusion holds:

$$(x^*, f^*) \in Q_l. \tag{A.1}$$

Besides, $f(x^*) - f(x_l) \geqslant \langle a_l, x^* - x_l \rangle$. Therefore, $f(x_l) + \langle a_l, x^* - x_l \rangle \leqslant f^*$, and by (11), (A.1) $(x^*, f^*) \in M_{l+1}$. This completes the proof of the lemma.

Lemma 2. *Sequence* $\{\alpha_i\}$, $i \in K$, *constructed by the proposed method satisfies*

$$\alpha_i \leqslant f^* \tag{A.2}$$

for all $i \in K$.

Proof of Lemma 2. Inequality (A.2) follows from the form of constraints in problem (2) and equality (12) together with inclusions (17).

Lemma 3. *The point* (x^*, f^*) *satisfies constraints of problem* (2) *for all* $i \in K$.

Proof of Lemma 3 follows from Lemmas 1, 2.

Lemma 4. *The sequence* $\{\gamma_i\}$, $i \in K$, *constructed by the method satisfies the following inequalities:*

$$\gamma_i \leqslant f^* \quad \forall i \in K. \tag{A.3}$$

Proof of Lemma 4 follows from Lemma 3.

Proof of Theorem 1. Due to (3), (A.3) $f(y_i) \leqslant f^*$. On the other hand, $f(y_i) \geqslant f^*$ since $y_i \in D$. Thus, $f(y_i) = f^*$, and the statement is proven.

Lemma 5. *For every* $i \in K$, *the set* U_i *is nonempty.*

Proof of Lemma 5. Fix an arbitrary index $i = r \in K$. To prove the statement it suffices to respresent that there exists a point $\overline{x}$ and a number $\overline{\gamma}$ such that

$$\overline{x} \in D, \quad \overline{\gamma} \leqslant l_r, \quad (\overline{x}, \overline{\gamma}) \in M_r. \tag{A.4}$$

According to (5), (13), $\beta_r \geqslant f^*$. Therefore, due to (A.3), (6), (4) it holds that

$$l_r \geqslant \gamma_r. \tag{A.5}$$

Further, due to (2) inclusions $y_r \in D$ and $(y_r, \gamma_r) \in M_r$ hold. Thus, letting $\overline{x} = y_r$, $\overline{\gamma} = \gamma_r$, we obtain from this together with (A.5) relations (A.4). This completes the proof of the lemma.

Proof of Theorem 2. We begin by showing that

$$\lim_{i \in K} (\beta_i - \gamma_i) = 0. \tag{A.6}$$

Note that due to conditions (5), (13) our choice of numbers β_i, and inequalities (A.3), $\beta_i - \gamma_i \geqslant 0$ for all $i \in K$. Suppose that equality (A.6) does not hold. Then there exist an infinite subset of indices $K_1 \subset K$ and number $\varepsilon > 0$ such that for all $i \in K_1$, $i \geqslant \tilde{i}$, satisfies inequality

$$\beta_i - \gamma_i \geqslant \varepsilon. \tag{A.7}$$

We choose from sequence $\{x_i\}$, $i \in K_1 \subset K$, a converging subsequence $\{x_i\}$, $i \in K_2 \subset K_1$, and assume that $\overline{x}$ is its limit point.

Fix the indices i', $i'' \in K_2$ such that

$$i'' > i' \geqslant \tilde{i}. \tag{A.8}$$

Since due to (5), (13) it holds that $\beta_{i'+1} \leqslant \delta_{i'+1} = f(x_{i'})$, and the sequence $\{\beta_i\}$, $i \in K$, is monotone decreasing, we get that

$$f(x_{i'}) \geqslant \beta_{i''}. \tag{A.9}$$

Following [11] and taking into account (A.9), (6), we get the following relations:

$$f(x_{i'}) - (1 - \lambda_{i''})(\beta_{i''} - \gamma_{i''}) \geqslant \beta_{i''} - (1 - \lambda_{i''})(\beta_{i''} - \gamma_{i''}) = l_{i''}. \tag{A.10}$$

Since by step 3 of the method $x_{i''} \in U_{i''}$, there exists a number γ'' such that the following inequality holds

$$\gamma'' \leqslant l_{i''} \tag{A.11}$$

together with inclusion $(x_{i''}, \gamma'') \in M_{i''}$. But due to (8), (11), (A.8) $M_{i''} \subset T_{i'} = \{(x, \gamma) \in R_{n+1} : f(x_{i'}) + \langle a_{i'}, x - x_{i'} \rangle \leqslant \gamma\}$, so $(x_{i''}, \gamma'') \in T_{i'}$. This together with (A.11) implies that

$$f(x_{i'}) + \langle a_{i'}, x_{i''} - x_{i'} \rangle \leqslant l_{i''}. \tag{A.12}$$

Then, taking into account inequalities (A.10), we have that

$$f(x_{i'}) - (1 - \lambda_{i''})(\beta_{i''} - \gamma_{i''}) \geqslant f(x_{i'}) + \langle a_{i'}, x_{i''} - x_{i'} \rangle. \tag{A.13}$$

Since $x_i \in D$, $i \in K$, and the set D is bounded, there exists (see, e.g., [9, p. 121]) $\theta < +\infty$ such that

$$||a|| \leqslant \theta \quad \forall a \in \partial f(x_i), \quad i \in K. \tag{A.14}$$

Then (A.13), (A.14) imply that $(1 - \lambda_{i''})(\beta_{i''} - \gamma_{i''}) \leqslant \theta ||x_{i''} - x_{i'}||$, and since $\beta_{i''} - \gamma_{i''} \geqslant \varepsilon$ by (A.7), in view of (4) we have that

$$\varepsilon \leqslant \theta \frac{||x_{i''} - x_{i'}||}{1 - \overline{\lambda}}. \tag{A.15}$$

Let us now choose for every $i \in K_2$ an index $p_i \in K_2$ such that $p_i \geqslant i + 1$. Then due to (A.15) $\varepsilon \leqslant \theta \frac{||x_{p_i} - x_i||}{1 - \overline{\lambda}}$, $i \in K_2$. The last inequality contradicts our choice of ε since $x_i \to \overline{x}$, $x_{p_i} \to \overline{x}$ over $i \in K_2$. Thus, Eq. (A.6) is proven.

Further, we prove that the following equality holds:

$$\lim_{i\in K}(f(x_i)-\gamma_i)=0. \tag{A.16}$$

Note that by (A.3) and inclusion $x_i \in D$ it holds for all $i \in K$ that $f(x_i)-\gamma_i \geqslant 0$. Suppose that equality (A.16) does not hold. Then there exist subsequences $\{x_i\}$, $\{\gamma_i\}$, $i \in K' \subset K$, sequences $\{x_i\}$, $\{\gamma_i\}$, $i \in K$, respectively, and a number $\varepsilon' > 0$ such that

$$f(x_i)-\gamma_i \geqslant \varepsilon'$$

for all $i \in K'$, $i \geqslant \tilde{i}$.

Let $\{x_i\}$, $i \in K'' \subset K'$, be a converging subsequence of sequence $\{x_i\}$, $i \in K'$, and let $\overline{x}$ be its limit point. Fix the indices i', $i'' \in K''$ from condition (A.8). Since $f(x_{i'}) \geqslant \gamma_{i'} + \varepsilon'$, (A.12) implies that $l_{i''} \geqslant \gamma_{i'} + \varepsilon' + \langle a_{i'}, x_{i''} - x_{i'}\rangle$. Then, taking into account (A.14), we get the following inequality:

$$l_{i''} \geqslant \gamma_{i'} + \varepsilon' - \theta||x_{i''} - x_{i'}||.$$

On the other hand,

$$l_{i''} = \gamma_{i''} + \lambda_{i''}(\beta_{i''} - \gamma_{i''}) \leqslant \beta_{i''} + \lambda_{i''}(\beta_{i''} - \gamma_{i''}) \leqslant \beta_{i'} + \lambda_{i''}(\beta_{i'} - \gamma_{i''}).$$

Therefore,

$$\varepsilon' - \theta||x_{i''} - x_{i'}|| \leqslant \lambda_{i''}(\beta_{i'} - \gamma_{i''}) + \beta_{i'} - \gamma_{i'}. \tag{A.17}$$

Due to (2) $\alpha_{i+1} \leqslant \gamma_{i+1}$, and by (12) $\alpha_{i+1} = \gamma_i$ for all $i \in K$, i.e., sequence $\{\gamma_i\}$, $i \in K$, is monotone increasing. Consequently, $\gamma_{i'} \leqslant \gamma_{i''}$, and (A.17) together with (4) imply that

$$\varepsilon' \leqslant \theta||x_{i''} - x_{i'}|| + (1+\overline{\lambda})(\beta_{i'} - \gamma_{i'}). \tag{A.18}$$

Now fix for each $i \in K''$ an index $p_i \in K''$ such that $p_i \geqslant i+1$. Then by (A.18)

$$\varepsilon' \leqslant \theta||x_{p_i} - x_i|| + (1+\overline{\lambda})(\beta_i - \gamma_i) \quad \forall i \in K''. \tag{A.19}$$

Since $x_i \to \overline{x}$ and $x_{p_i} \to \overline{x}$ for $i \to \infty$, $i \in K''$, and due to (A.6) it holds that $\lim_{i\in K''}(\beta_i - \gamma_i) = 0$, we get from (A.19) inequality $\varepsilon' \leqslant 0$, which contradicts the choice of ε'. This proves equality (A.16).

Now (A.16) and (A.3) imply the statement of the Theorem.

Proof of Theorem 3. 1. Let $k = 0$. If $f(x_0) - \gamma_0 \leqslant \varepsilon_0$, then by step 4 of the method $i_0 = 0$, $z_0 = x_{i_0} = x_0$, $\sigma_0 = \gamma_0$, and equalities (9) hold for $k = 0$. Therefore, we will assume that $f(x_0) - \gamma_0 > \varepsilon_0$. Then we show that there exists an index $i = i_0 > 0$ for which the following inequality holds:

$$f(x_{i_0}) - \gamma_{i_0} \leqslant \varepsilon_0. \tag{A.20}$$

Assume the opposite, i.e., $f(x_i) - \gamma_i > \varepsilon_0$ for all $i \in K$, $i > 0$. Then the sets Q_i, $i \in K$, $i > 0$, have form (8), i.e., conditions of Theorem 2 hold, and consequently Eq. (A.16) also holds. It implies inequality $\varepsilon_0 \leqslant 0$ which contradicts condition (22). Thus, the existence of an index $i_0 > 0$ that satisfies (A.20) has been proven, and equalities (9) hold for $k = 0$.

2. Suppose now that (9) hold for some fixed $k \geqslant 0$, i.e., there exists an index $i = i_k$ satisfying condition (18) for a given k. Let us show that there exists an index $i_{k+1} > i_k$ such that

$$f(x_{i_{k+1}}) - \gamma_{i_{k+1}} \leqslant \varepsilon_{k+1}, \tag{A.21}$$

then $z_{k+1} = x_{i_{k+1}}$, $\sigma_{k+1} = \gamma_{i_{k+1}}$, and the Lemma will be proven. If we assume the opposite, i.e., assume that for all $i \in K$, $i > i_k$ it holds that $f(x_i) - \gamma_i > \varepsilon_{k+1}$, then, similar to the first part of the proof, together with equality (A.16) we will get a contradiction with the choice of ε_{k+1}. Thus, there does indeed exist an index $i_{k+1} > i_k$ satisfying (A.21), and this completes the proof of the theorem.

Proof of Theorem 4. Due to (9) $f(x_{i_k}) - \gamma_{i_k} \leqslant \varepsilon_k$ or

$$f(z_k) \leqslant \sigma_k + \varepsilon_k \quad \forall k \in K. \tag{A.22}$$

Besides, by Lemma 4 and inclusion $z_k \in D$, $k \in K$, it holds that

$$\sigma_k \leqslant f^* \leqslant f(z_k) \quad \forall k \in K. \tag{A.23}$$

Therefore $f^* \leqslant f(z_k) \leqslant f^* + \varepsilon_k$, $k \in K$. Then due to (23)

$$f^* \leqslant \varliminf_{k \in K} f(z_k) \leqslant \varlimsup_{k \in K} f(z_k) \leqslant f^*,$$

and the first of equalities (24) is proven.

Further, due to the same inequalities (A.22), (A.23) $f^* - \varepsilon_k \leqslant \sigma_k \leqslant f^*$, $k \in K$. This together with (23) implies the second equality of (24), and this completes the proof of the theorem.

REFERENCES

1. Bulatov, V.P., *Metody pogruzheniya v zadachakh optimizatsii* (Immersion Methods in Optimization Problems), Novosibirsk: Nauka, 1977.
2. Bulatov, V.P. and Khamisov, O.V., Cutting Methods in E^{n+1} for Global Optimization Problems on a Class of Functions, *Zh. Vychisl. Mat. Mat. Fiz.*, 2007, vol. 47, no. 11, pp. 1830–1842.
3. Zabotin, I.Ya., On Some Immersion-Severance Algorithms for Mathematical Programming Problems, *Izv. Irkutsk. Gos. Univ., Ser. Mat.*, 2011, vol. 4, no. 2, pp. 91–101.
4. Zabotin, I.Ya. and Yarullin, R.S., On One Approach to Constructing Branch and Bound Algorithms with Discarding Separating Planes, *Izv. Vyssh. Uchebn. Zaved., Mat.*, 2013, no. 3, pp. 74–79.
5. Zangwill, W.I., *Nonlinear Programming: A Unified Approach*, Englewood Cliffs: Prentice Hall, 1969. Translated under the title *Nelineinoe programmirovanie*, Moscow: Sovetskoe Radio, 1973.
6. Kolokolov, A.A., Regular Partitions and Cuts in Integer Programming, *Sib. Zh. Issled. Oper.*, 1994, vol. 1, no. 2, pp. 18–39.
7. Nesterov, Yu.E., *Vvedenie v vypukluyu optimizatsiyu* (Introduction to Convex Optimization), Moscow: MTsNMO, 2010.
8. Nurminskii, E.A., The Separating Planes Method with Bounded Memory for Convex Nonsmooth Optimization Problems, *Vychisl. Metod. Programm.*, 2006, vol. 7, pp. 133–137.
9. Polyak, B.T., *Vvedenie v optimizatsiyu* (Introduction to Optimization), Moscow: Nauka, 1983.
10. Kelley, J.E., The Cutting-Plane Method for Solving Convex Programs, *SIAM J.*, 1960, vol. 8, no. 4, pp. 703–712.
11. Lemarechal, C., Nemirovskii, A., and Nesterov, Yu., New Variants of Bundle Methods, *Math. Program.*, 1995, vol. 69, pp. 111–148.
12. Vasil'ev, F.P., *Chislennye metody resheniya ekstremal'nykh zadach* (Numerical Methods for Extremal Problems), Moscow: Nauka, 1988.

This paper was recommended for publication by B.T. Polyak, a member of the Editorial Board