Constructing Mixed Algorithms on the Basis of Some Bundle Method

Rashid Yarullin(✉)

Kazan (Volga region) Federal University, Kazan, Russia
yarullinrs@gmail.com

Abstract. In the paper, a method is proposed for minimizing a nondifferentiable convex function. This method belongs to a class of bundle methods. In the developed method it is possible to periodically produce discarding all previously constructed cutting planes that form the model of the objective function. These discards are applied when approximation of the epigraph of the objective function is sufficiently good in the a neighborhood of the current iteration point, and the quality of this approximation is estimated by using the model of the objective function. It is proposed an approach for constructing mixed minimization algorithms on the basis of the developed bundle method with involving any relaxation methods. The opportunity to mix the developed bundle method with other methods is provided as follows. In the proposed method during discarding the cutting planes the main iteration points are fixed with the relaxation condition. Any relaxation minimization method can be used to build these points. Moreover, the convergence of all such mixed algorithms will be guaranteed by the convergence of the developed bundle method. It is important to note that the procedures for updating cutting planes introduced in the bundle method will be transferred to mixed algorithms. The convergence of the proposed method is investigated, its properties are discussed, an estimate of the accuracy of the solution and estimation of the complexity of finding an approximate solution are obtained.

Keywords: Nondifferentiable optimization · Mixed algorithms · Bundle methods · Cutting planes · Sequence of approximations · Convex functions

1 Introduction

Nowadays a lot of different methods have been developed for solving nonlinear programming problems. Each of these optimization methods has its own disadvantages and advantages. In this regard, for solving practical problems these methods are used in a complex manner in order to accelerate the convergence of the optimization process. Namely, at each step to find the next approximation there are opportunities to choose any minimization method among other methods which allows to construct descent direction from the current point faster. The

Y. Kochetov et al. (Eds.): MOTOR 2020, CCIS 1275, pp. 150–163, 2020.
https://doi.org/10.1007/978-3-030-58657-7_14

algorithm that is formed as a result of applying various optimization methods is called mixed (e.g., [1,2]).

In this paper, based on the ideas of [2], an approach is proposed for constructing mixed algorithms on the basis of some proximal bundle method which is characterized by the possibility to periodically discard the cutting planes.

2 Problem Setting

Let $f(x)$ be a convex function defined in an n-dimensional Euclidian space, $\partial f(x)$, $\partial_\epsilon f(x)$ be a subdifferential and an ϵ-subdifferential of the function $f(x)$ at x respectively.

Suppose $f^* = \min\{f(x) : x \in \mathrm{R^n}\}$, $X^* = \{x \in \mathrm{R^n} : f(x) = f^*\} \neq \emptyset$, $X^*(\varepsilon) = \{x \in \mathrm{R^n} : f(x) \leq f^* + \varepsilon\}$, $\varepsilon > 0$, $K = \{0, 1, \dots\}$, $L(y) = \{x \in \mathrm{R^n} : f(x) \leq f(y)\}$, where $y \in \mathrm{R^n}$. Denote by $\lceil \chi \rceil$ the least integer no less than $\chi \in \mathrm{R^1}$. It is assumed that the set $L(y)$ is bounded for any $y \in \mathrm{R^n}$. Fix an arbitrary point $x^* \in X^*$.

It is required to find a point from the set $X^*(\varepsilon)$ with given $\varepsilon > 0$ for a finite number of iterations.

3 Minimization Method

First, consider an auxiliary procedure $\pi = \pi(\bar{x}, \bar{\xi}, \bar{\theta}, \bar{\mu})$ with the following input parameters:

$$\bar{x} \in \mathrm{R^n}, \quad \bar{\xi} > 0, \quad \bar{\theta} \in (0, 1), \quad \bar{\mu} > 0.$$

Step 0. Define initial parameters $k = 0$, $x_k = \bar{x}$.

Step 1. Choose a subgradient $s_k \in \partial f(x_k)$. Assign $i = 0$, $s_{k,i} = s_k$, $x_{k,i} = x_k$,

$$\hat{f}_{k,i}(y) = f(x_{k,i}) + \langle s_{k,i}, y - x_{k,i} \rangle. \tag{1}$$

Step 2. Find a point

$$x_{k,i+1} = \arg\min\{\hat{f}_{k,i}(y) + \frac{\bar{\mu}}{2}\|y - x_k\|^2 : y \in \mathrm{R^n}\}. \tag{2}$$

Step 3. Compute a parameter

$$\delta_{k,i} = f(x_k) - \hat{f}_{k,i}(x_{k,i+1}) - \frac{\bar{\mu}}{2}\|x_{k,i+1} - x_k\|^2. \tag{3}$$

Step 4. If the inequality

$$\delta_{k,i} \leq \bar{\xi}, \tag{4}$$

is fulfilled, then the process of finding sequence is stopped, and the point

$$\hat{x} = \arg\min\{f(x_{k,j}) : 0 \leq j \leq i + 1\} \tag{5}$$

is a result of the procedure.

Step 5. If the condition

$$f(x_{k,i+1}) \le f(x_k) - \bar{\theta}\delta_{k,i} \tag{6}$$

is fulfilled, then choose a point $x_{k+1} \in \mathrm{R}^{\mathrm{n}}$ according to the inequality

$$f(x_{k+1}) \le f(x_{k,i+1}), \tag{7}$$

fix a number $i_k = i$, increase the value of k by one, and go to Step 1. Otherwise, go to the next step.

Step 6. Choose a subgradient $s_{k,i+1} \in \partial f(x_{k,i+1})$, assign

$$\hat{f}_{k,i+1}(y) = \max\{\hat{f}_{k,i}(y), f(x_{k,i+1}) + \langle s_{k,i+1}, y - x_{k,i+1}\rangle\}, \tag{8}$$

and go to Step 2 with incremented i.

Consider some remarks concerning the procedure π.

Remark 1. For some $k \ge 0$, $i \ge 0$ on the basis of (1), (8) it is not difficult to obtain the equality

$$\hat{f}_{k,i}(y) = \max_{0 \le j \le i}\{f(x_{k,j}) + \langle s_{k,j}, y - x_{k,j}\rangle\}. \tag{9}$$

The function $\hat{f}_{k,i}(y)$ is a model of the convex function $f(x)$. Since the model $\hat{f}_{k,i}(y)$ is the maximum of linear (hence convex) functions, then the function $\hat{f}_{k,i}(y)$ is convex.

One of the main problems arising in the numerical implementation of bundle and cutting methods is the unlimited growth of the count of cutting planes which are used to find iteration points. Currently, several approaches are proposed to discard cutting planes for bundle methods (e. g., [3,5,6]). These approaches are realized according to the aggregation technique of cutting planes proposed in [3] as follows. At the initial step of any bundle method, a storage of cutting planes (called a bundle) is formed and its size is set. Then the overflow of this storage is checked at each step. If the storage of the cutting planes is full, then the procedure is started for discarding the cutting planes in two stages. All inactive cutting planes are discarded at the first stage, and if the first stage does not allow to allocate free spaces in the plane storage, then the second stage is performed. At the second stage any active cutting plane is removed from the storage to free space and one aggregated cutting plane is added which is constructed as a convex combination of active and inactive cutting planes. Note that the application of such an aggregation technique allows approximating the subdifferential of the objective function at the current point and construct some e-subgradient. However, the quality of the approximation of the epigraph of the objective function at the current iteration point is deteriorated after performing the second stage of the procedure for discarding the cutting planes.

A different approach was developed for cutting plane methods for periodically discarding cutting planes in [7–9]. This approach is based on some criteria for estimating the quality of approximating sets formed by cutting planes in a neighborhood of current iteration points. In particular, in [8] the quality of the approximation is estimated by the proximity of the current iteration point to a feasible set of the initial problem, and in [9] the quality is estimated by the assessment of the proximity of the current iteration value to the optimal value. After obtaining sufficiently good approximation sets the proposed approach allows to use update procedures such that it is possible to periodically discard an arbitrary number of any previously constructed cutting planes. Namely, both full and partial updating of approximating sets is permissible. In the case of using partial updating it is possible to leave, for example, only active cutting planes or $n+1$-last cutting planes.

In this paper, the procedure π is proposed, where cutting planes are discarded based on the approach developed for the cutting plane methods. Namely, at Step 5 of the procedure π there is the possibility of periodically discarding all cutting planes as follows. In the neighborhood of the point $x_{k,i+1}$ the approximation quality of the epigraph of the function $f(x)$ is evaluated by the model $\hat{f}_{k,i}(x)$. If inequality (6) is fulfilled for some $k \geq 0$, $i \geq 0$, then the approximation quality is enough good, and there is a full update of the model of the function $f(x)$ by discarding cutting planes. Otherwise, the model of the convex function $\hat{f}_{k,i}(x)$ is refined and cutting planes are not discarded.

Based on the procedure π the bundle method will be constructed below. Note that at Step 5 of the procedure π during discarding cutting planes basic points x_k, $k \in K$ are determined. In the process of constructing these points can be used any relaxation minimization methods. It is important to note that convergence of such mixed algorithms is guaranteed by the convergence of the proposed bundle method even if the mentioned relaxation methods included in mixed algorithms are heuristic.

Lemma 1. *Let $S \subset \mathrm{R}^{\mathrm{n}}$ be a bounded closed set, $\tau \geq 0$. Then the set*

$$B(\tau, S) = \bigcup_{v \in S} \{y \in \mathrm{R}^{\mathrm{n}} : \|y - v\| \leq \tau\} \tag{10}$$

is bounded.

Proof. Since the set S is bounded, then there exists a number $\tau' > 0$ such that for any $v \in S$ the inequality

$$\|v\| \leq \tau' \tag{11}$$

is defined. Now suppose that the set $B(\tau, S)$ is not bounded. Then for any $\omega > 0$ there exists a point $y \in B(\tau, S)$ such that $\|y\| > \omega$. Fix any sequence of positive numbers $\{\omega_k\}$, $k \in K$, such that $\omega_k \to +\infty$, $k \in K$. Due to unboundedness of the set $B(\tau, S)$ there is a sequence of points $\{y_k\}$, $k \in K$, such that

$$y_k \in B(\tau, S), \quad \|y_k\| > \omega_k, \quad k \in K. \tag{12}$$

Moreover, in accordance with construction of points $\{y_k\}$, $k \in K$, for each $k \in K$ there exists a point $v_k \in S$ satisfying the condition

$$\|y_k - v_k\| \leq \tau.$$

Hence, from (11), (12) we have

$$\omega_k < \|y_k \pm v_k\| \leq \|y_k - v_k\| + \|v_k\| \leq \tau + \tau'.$$

The obtained inequality $\omega_k \leq \tau + \tau'$ contradicts the assumption $\omega_k \to +\infty$. The lemma is proved.

Lemma 2. *Suppose for some $k \geq 0$, $i \geq 0$ the points x_k, $x_{k,0}$, $x_{k,1}$, ..., $x_{k,i}$ and the model $\hat{f}_{k,i}(y)$ are constructed by the procedure π. Then we obtain*

$$\epsilon_{k,j} = f(x_k) - f(x_{k,j}) - \langle s_{k,j}, x_k - x_{k,j}\rangle \geq 0, \quad 0 \leq j \leq i, \tag{13}$$

$$\hat{f}_{k,i}(y) = f(x_k) + \max_{0\leq j\leq i}\{\langle s_{k,j}, y - x_k\rangle - \epsilon_{k,j}\}.$$

Proof. Since the function $f(x)$ is convex and $s_{k,j} \in \partial f(x_{k,j})$, $0 \leq j \leq i$, then using definition of a subgradient it is not difficult to obtain (13). Further, taking account (9) and (13) we have

$$\begin{aligned}\hat{f}_{k,i}(y) &= \max_{0\leq j\leq i}\{f(x_{k,j}) + \langle s_{k,j}, x - x_{k,j}\rangle \pm \epsilon_{k,j}\}\\ &= \max_{0\leq j\leq i}\{\langle s_{k,j}, y - x_{k,j}\rangle - \epsilon_{k,j} + f(x_k) - \langle s_{k,j}, x_k - x_{k,j}\rangle\}\\ &= f(x_k) + \max_{0\leq j\leq i}\{\langle s_{k,j}, y - x_k\rangle - \epsilon_{k,j}\}.\end{aligned}$$

The lemma is proved.

The following theorem is proved in [3, p. 144].

Theorem 1. *Suppose for some $k \geq 0$, $i \geq 0$ the point $x_{k,i+1}$ is constructed according to (2) by the procedure π. Then*

$$x_{k,i+1} = x_k - \frac{\hat{s}_{k,i}}{\bar{\mu}}, \tag{14}$$

where

$$\hat{s}_{k,i} = \sum_{j=0}^{i} \hat{\alpha}_{k,i}^j s_{k,j}, \tag{15}$$

and the vector $\hat{\alpha}_{k,i} = (\hat{\alpha}_{k,i}^0, \hat{\alpha}_{k,i}^1, \ldots, \hat{\alpha}_{k,i}^i) \in \mathrm{R}^{\mathrm{i}+1}$ is a solution of the following problem:

$$\min_{\alpha=(\alpha^0,\alpha^1,\ldots,\alpha^i)\in\mathrm{R}^{\mathrm{i}+1}} \frac{1}{2\bar{\mu}}\|\sum_{j=0}^{i}\alpha^j s_{k,j}\|^2 + \sum_{j=0}^{i}\alpha^j\epsilon_{k,j}, \tag{16}$$

$$s.t. \quad \alpha = (\alpha^0, \alpha^1, \ldots, \alpha^i) \geq 0, \quad \sum_{j=0}^{i}\alpha^j = 1. \tag{17}$$

Moreover, the following expressions

$$\delta_{k,i} = \hat{\epsilon}_{k,i} + \frac{1}{2\bar{\mu}}\|\hat{s}_{k,i}\|^2, \tag{18}$$

$$\hat{s}_{k,i} \in \partial_{\hat{\epsilon}_{k,i}} f(x_k), \tag{19}$$

$$\hat{s}_{k,i} \in \partial \hat{f}_{k,i}(x_{k,i+1}) \tag{20}$$

are valid, where

$$\hat{\epsilon}_{k,i} = \sum_{j=0}^{i} \hat{\alpha}_{k,i}^{j} \epsilon_{k,j}. \tag{21}$$

From inclusion (19) it follows

Lemma 3. *Suppose the points x_k, $x_{k,0}, \ldots, x_{k,i+1}$ and the corresponding subgradients s_k, $s_{k,0}, \ldots, s_{k,i+1}$ are constructed for some $k \geq 0$, $i \geq 0$ by the proposed procedure π. Then for any point $y \in \mathrm{R}^{\mathrm{n}}$ the inequality*

$$f(x_k) - f(y) \leq \langle \hat{s}_{k,i}, x_k - y \rangle + \hat{\epsilon}_{k,i} \tag{22}$$

is fulfilled, where $\hat{s}_{k,i}$, $\hat{\epsilon}_{k,i}$ are defined according to (15), (21) respectively.

Lemma 4. *Suppose that the stopping criterion (4) is fulfilled for some $k \geq 0$, $i \geq 0$. Then the following estimate holds:*

$$f(\hat{x}) - f^* \leq \bar{\rho}\sqrt{2\bar{\mu}\bar{\xi}} + \bar{\xi}, \tag{23}$$

where $\bar{\rho} > 0$ is the diameter of the set $L(\bar{x})$.

Proof. Note that the equality $f(x_0) = f(\bar{x})$ is fulfilled in accordance with Step 0 of the procedure π, and from (6), (7) we have $f(x_k) \leq f(\bar{x})$. Consequently, $x_k \in L(\bar{x})$. Moreover, in view of condition (5) the inequality $f(\hat{x}) \leq f(x_k)$ is defined. Hence and from inequality (22) under $y = x^*$ the estimate holds

$$f(\hat{x}) - f^* \leq \|\hat{s}_{k,i}\| \|x_k - x^*\| + \hat{\epsilon}_{k,i}. \tag{24}$$

Further, according to the stopping criterion (4) and equality (18) we obtain

$$\|\hat{s}_{k,i}\| \leq \sqrt{2\bar{\mu}\delta_{k,i}} \leq \sqrt{2\bar{\mu}\bar{\xi}},$$

$$\hat{\epsilon}_{k,i} \leq \delta_{k,i} \leq \bar{\xi}.$$

Hence and from (24), $x^* \in L(\bar{x})$, $x_k \in L(\bar{x})$ it follows the estimate (23). The lemma is proved.

To prove finiteness of the procedure π let's show that values $\delta_{k,i}$, $\|x_{k,i+1} - x_k\|$ are bounded.

Lemma 5. *Suppose that for some $k \geq 0$, $i \geq 0$ the points x_k, $x_{k,i+1}$ are constructed, the subgradient s_k is fixed, the number $\delta_{k,i}$ is computed by the procedure π. Then the following expressions*

$$\|x_{k,i+1} - x_k\| \leq \frac{2\|s_k\|}{\bar{\mu}}, \tag{25}$$

$$0 \le \delta_{k,i} \le \frac{2\|s_k\|^2}{\bar{\mu}}, \tag{26}$$

$$f(x_k) - \delta_{k,i} = \hat{f}_{k,i}(x_{k,i+1}) + \langle \hat{s}_{k,i}, y - x_{k,i+1} \rangle + \frac{\bar{\mu}}{2}\|y - x_k\|^2 - \frac{\bar{\mu}}{2}\|y - x_{k,i+1}\|^2 \tag{27}$$

are fulfilled, where $y \in \mathrm{R^n}$.

Proof. Note that according to (9) for all $j = 0, \dots, i$, we have

$$\hat{f}_{k,i}(y) \ge f(x_{k,j}) + \langle s_{k,j}, y - x_{k,j} \rangle, \tag{28}$$

where $y \in \mathrm{R^n}$, and from Step 1 of the procedure π it follows that $x_{k,0} = x_k$, $s_{k,0} = s_k$. Hence from formula (9) with $y = x_{k,i+1}$, $j = 0$, we obtain

$$f(x_k) - \hat{f}_{k,i}(x_{k,i+1}) \le \langle s_k, x_k - x_{k,i+1} \rangle \le \|s_k\| \|x_k - x_{k,i+1}\|. \tag{29}$$

Moreover, from (3) it follows

$$\frac{\bar{\mu}}{2}\|x_{k,i+1} - x_k\|^2 \le f(x_k) - \hat{f}_{k,i}(x_{k,i+1}). \tag{30}$$

Hence combining inequalities (29), (30) we prove (25).

Further, according to Lemma 2 for all $j = 0, \dots, i$ we get $\epsilon_{k,j} \ge 0$, therefore, in view of (21) the inequality $\hat{\epsilon}_{k,i} \ge 0$ is determined. Hence and from (18) taking into account $\bar{\mu} > 0$ it follows that $\delta_{k,i} \ge 0$. Moreover, in accordance with (3), (29) we have

$$\delta_{k,i} \le f(x_k) - \hat{f}_{k,i}(x_{k,i+1}) \le \|s_k\| \|x_k - x_{k,i+1}\|.$$

Using the last inequality, (25) and $\delta_{k,i} \ge 0$ expression (26) is obtained.

Let's turn to obtain inequality (27). For any $y \in \mathrm{R^n}$ it is determined

$$\|y \pm x_{k,i+1} - x_k\|^2 = \|y - x_{k,i+1}\|^2 + \|x_{k,i+1} - x_k\|^2 + 2\langle y - x_{k,i+1}, x_{k,i+1} - x_k \rangle.$$

Then multiplying the last equality by $\bar{\mu}/2$ and taking into account (14) we get

$$\frac{\bar{\mu}}{2}\|x_{k,i+1} - x_k\|^2 = \frac{\bar{\mu}}{2}\|y - x_k\|^2 - \frac{\bar{\mu}}{2}\|y - x_{k,i+1}\|^2 + \langle \hat{s}_{k,i}, y - x_{k,i+1} \rangle. \tag{31}$$

Moreover, from (3) it follows

$$f(x_k) - \delta_{k,i} = \hat{f}_{k,i}(x_{k,i+1}) + \frac{\bar{\mu}}{2}\|x_{k,i+1} - x_k\|^2.$$

Now substituting $\bar{\mu}/2\|x_{k,i+1} - x_k\|^2$ by (31) in the last equality we obtain (27). The lemma is proved.

Corollary 1. *Suppose that conditions of Lemma 5 are defined,* $S \subset \mathrm{R^n}$ *is bounded closed set satisfying the inclusion*

$$L(\bar{x}) \subset S. \tag{32}$$

Then there exists numbers $\eta = \eta(S) > 0$, $\zeta = \zeta(\eta) > 0$ *such that the inequalities*

$$\|s_k\| \leq \eta, \tag{33}$$

$$\|x_{k,i+1} - x_k\| \leq \frac{2\eta}{\bar{\mu}}, \tag{34}$$

$$\delta_{k,i} \leq \frac{2\eta^2}{\bar{\mu}}, \tag{35}$$

$$\|s_{k,i+1}\| \leq \zeta \tag{36}$$

are fulfilled.

Proof. Since inclusion (32) is fulfilled according to conditions of the corollary and we have $x_k \in L(\bar{x})$, $s_k \in \partial f(x_k)$ by construction, then in view of boundness of the set S there exists a number $\eta = \eta(S) > 0$ (e. g., [4, p. 121]) such that inequality (33) is determined. Moreover, taking into account inequality (33) from (25), (26) it follows (34), (35).

Further, since the set S is bounded and closed, then according to Lemma 1 the set $B(2\eta/\bar{\mu}, S)$ is bound too. Moreover, from the inclusion $x_k \in L(\bar{x}) \subset S$ and inequality (34) we have $x_{k,i+1} \in B(2\eta/\bar{\mu}, S)$. Therefore, taking into account $s_{k,i+1} \in \partial f(x_{k,i+1})$ there exists a number $\zeta = \zeta(\eta) > 0$ (e. g., [4, p. 121]) such that inequality (36) is determined. The assertion is proved.

Lemma 6. *Suppose that by the proposed procedure* π *for some* $\bar{k} \geq 0$, $\bar{i} \geq 2$ *the points* $x_{\bar{k}} = x_{\bar{k},0}$,

$$x_{\bar{k},1}, x_{\bar{k},2}, \ldots, x_{\bar{k},\bar{i}+1} \tag{37}$$

are constructed, the subgradients $s_{\bar{k}} = s_{\bar{k},0}$,

$$s_{\bar{k},1}, s_{\bar{k},2}, \ldots, s_{\bar{k},\bar{i}+1} \tag{38}$$

are chosen, and according to (3) the numbers

$$\delta_{\bar{k},0}, \delta_{\bar{k},1}, \ldots, \delta_{\bar{k},\bar{i}} \tag{39}$$

are computed. Then for each $i = 0, \ldots, \bar{i} - 2$ *it is determined that*

$$\delta_{\bar{k},i} - \delta_{\bar{k},i+1} \geq \frac{\bar{\mu}(1-\bar{\theta})^2}{2(\|s_{\bar{k},i+2}\| + \|s_{\bar{k},i+1}\|)^2}\delta^2_{\bar{k},i+1}. \tag{40}$$

Proof. According to Step 5 of the procedure π for each $l = 0, \ldots, \bar{i} - 1$ it is determined

$$f(x_{\bar{k},l+1}) > f(x_{\bar{k}}) - \bar{\theta}\delta_{\bar{k},l}, \tag{41}$$

and in view of equality (14) the vectors

$$\hat{s}_{\bar{k},0}, \hat{s}_{\bar{k},1}, \ldots, \hat{s}_{\bar{k},\bar{i}}$$

correspond to points (37). Choose an arbitrary index i such that $0 \le i \le \bar{i} - 2$. Then using definition of a subgradient of a convex function and taking into account (20) we have

$$\hat{f}_{\bar{k},i}(x_{\bar{k},i+1}) \le \hat{f}_{\bar{k},i}(x_{\bar{k},i+2}) + \langle \hat{s}_{\bar{k},i}, x_{\bar{k},i+1} - x_{\bar{k},i+2} \rangle. \tag{42}$$

Moreover, according to (8) for any $y \in \mathrm{R}^{\mathrm{n}}$ it is defined

$$\hat{f}_{\bar{k},i}(y) \le \hat{f}_{\bar{k},i+1}(y).$$

Hence under $y = x_{\bar{k},i+2}$ and from (42) it follows that

$$\hat{f}_{\bar{k},i}(x_{\bar{k},i+1}) \le \hat{f}_{\bar{k},i+1}(x_{\bar{k},i+2}) + \langle \hat{s}_{\bar{k},i}, x_{\bar{k},i+1} - x_{\bar{k},i+2} \rangle,$$

and taking into account for the $(i+1)$-th element the last inequality has the form

$$\hat{f}_{\bar{k},i}(x_{\bar{k},i+1}) \le f(x_{\bar{k}}) - \delta_{\bar{k},i+1} - \frac{\bar{\mu}}{2}\|x_{\bar{k},i+2} - x_{\bar{k}}\|^2 + \langle \hat{s}_{\bar{k},i}, x_{\bar{k},i+1} - x_{\bar{k},i+2} \rangle. \tag{43}$$

Now using equality (27) from Lemma 5 under $k = \bar{k}$, $y = x_{\bar{k},i+2}$ it is obtained

$$f(x_{\bar{k}}) - \delta_{\bar{k},i} + \frac{\bar{\mu}}{2}\|x_{\bar{k},i+2} - x_{\bar{k},i+1}\|^2 =$$
$$= \hat{f}_{\bar{k},i}(x_{\bar{k},i+1}) + \langle \hat{s}_{\bar{k},i}, x_{\bar{k},i+2} - x_{\bar{k},i+1} \rangle + \frac{\bar{\mu}}{2}\|x_{\bar{k},i+2} - x_{\bar{k}}\|^2.$$

Hence and from (43) it follows that

$$\frac{\bar{\mu}}{2}\|x_{\bar{k},i+2} - x_{\bar{k},i+1}\|^2 \le \delta_{\bar{k},i} - \delta_{\bar{k},i+1}. \tag{44}$$

On the other hand, from (3), (9) (for the $(i+1)$-th element) we get

$$\delta_{\bar{k},i+1} \le f(x_{\bar{k}}) - f(x_{\bar{k},i+1}) - \langle s_{\bar{k},i+1}, x_{\bar{k},i+2} - x_{\bar{k},i+1} \rangle,$$

and from inequality (41) under $l = i + 1$ it follows that

$$-\bar{\theta}\delta_{\bar{k},i+1} < f(x_{\bar{k},i+2}) - f(x_{\bar{k}}).$$

Now summing the last two inequalities it is determined that

$$\begin{aligned}(1 - \bar{\theta})\delta_{\bar{k},i+1} &\le f(x_{\bar{k},i+2}) - f(x_{\bar{k},i+1}) - \langle s_{\bar{k},i+1}, x_{\bar{k},i+2} - x_{\bar{k},i+1} \rangle \\ &\le (\|s_{\bar{k},i+2}\| + \|s_{\bar{k},i+1}\|)\|x_{\bar{k},i+2} - x_{\bar{k},i+1}\|.\end{aligned}$$

Hence and from (44) we obtain (40). The lemma is proved.

Theorem 2. *Let $S \subset \mathrm{R}^{\mathrm{n}}$ be a bounded closed set satisfied condition (32). Then complexity of the procedure π is equal to*

$$\lceil \frac{f(\bar{x}) - f^*}{\bar{\theta}\bar{\xi}} \rceil \lceil 1 + \frac{16\eta^2\zeta^2}{\bar{\mu}^2(1 - \bar{\theta})^2\bar{\xi}^2} \rceil, \tag{45}$$

where $\eta = \eta(S) > 0$, $\zeta = \zeta(\eta) > 0$.

Proof. First, let's estimate the number of iterations of the procedure π by k. Assume that in the procedure π there is a loop in relation to k. In this case, it is constructed a sequence $\{x_k\}$, $k \in K$, such that according to Steps 4, 5 of the procedure π for each $k \in K$ the following conditions hold:

$$\Delta_k > \bar{\xi}, \tag{46}$$

$$f(x_{k+1}) \leq f(x_k) - \bar{\theta}\Delta_k, \tag{47}$$

where $\Delta_k = \delta_{k,i_k}$. Now summing the last inequality by k from 0 to $n \geq 0$ we have

$$\sum_{k=0}^{n} \bar{\theta}\Delta_k \leq \sum_{k=0}^{n}(f(x_k) - f(x_{k+1})) \leq f(x_0) - f^*.$$

Hence under $n \to +\infty$ we obtain $\Delta_k \to 0$ which contradicts condition (46). Consequently, there exists a number $k' \geq 0$ such that the criterion

$$\Delta_{k'} \leq \bar{\xi}$$

is fulfilled.

Further, let's consider two cases to estimate the value k'.

1) Suppose that condition (4) is determined under $k = k' = 0$ and $i \geq 0$. Then it is clear that the number of iterations k' does not exceed the value of the first multiplier of valuation (45).
2) Suppose that criterion is fulfilled under $k = k' > 0$ and $i \geq 0$. Then according to Steps 4, 5 of the procedure π and in view of (46), (47) we have

$$\sum_{p=0}^{k'-1} \bar{\theta}\bar{\xi} \leq \sum_{p=0}^{k'-1}(f(x_p) - f(x_{p+1})) \leq f(x_0) - f^*.$$

Hence taking into account $x_0 = \bar{x}$ (in accordance with Step 0 of the procedure π) it is obtained that

$$k' \leq \lceil \frac{f(\bar{x}) - f^*}{\bar{\theta}\bar{\xi}} \rceil. \tag{48}$$

Now let's obtain a complexity of the procedure π in relation to i while k is fixed. Suppose that the point x_k is constructed under some $k \geq 0$ by the procedure π, and there is a loop in relation to i, i. e. for each $i \in K$ conditions (4), (6) are not fulfilled simultaneously. Then there is a sequence $\{\delta_{k,i}\}$, $i \in K$, constructed by the procedure π such that according to Lemma 6 for each $i \in K$ it is determined

$$\frac{\bar{\mu}(1-\bar{\theta})^2}{2(\|s_{k,i+2}\| + \|s_{k,i+1}\|)^2}\delta_{k,i+1}^2 \leq \delta_{k,i} - \delta_{k,i+1}.$$

Hence taking into account (36) from Corollary 1 we get

$$\frac{\bar{\mu}(1-\bar{\theta})^2}{8\zeta^2}\delta_{k,i+1}^2 \leq \delta_{k,i} - \delta_{k,i+1}.$$

After summing the last inequality by i from 0 to $n \geq 0$ we get

$$\sum_{i=0}^{n} \frac{\bar{\mu}(1-\bar{\theta})^2}{8\zeta^2} \delta_{k,i+1}^2 \leq \sum_{i=0}^{n} (\delta_{k,i} - \delta_{k,i+1}) \leq \delta_{k,0}.$$

Hence from $n \to +\infty$ it follows that $\delta_{k,i} \to 0$, $i \in K$. Therefore, there exists a number $i' \in K$ such that the inequality

$$\delta_{k,i'} \leq \bar{\xi}$$

is fulfilled.

To estimate i' consider the following cases.

1) Suppose that it is defined either criterion (4) or condition (6) for some $k \geq 0$, $i' = i \leq 1$. Then i' does not exceed the value of the second multiplier of variable (45).
2) Assume that any condition of (4), (6) is fulfilled for some $k \geq 0$, $i' = i \geq 2$. Then according to Lemma 6, stopping criterion (4) and inequalities (36), (35) from Corollary 1 we get

$$\sum_{j=0}^{i'-2} \frac{\bar{\mu}(1-\bar{\theta})^2}{8\zeta^2} \bar{\xi}^2 \leq \sum_{j=0}^{i'-2} (\delta_{k,j} - \delta_{k,j+1}) \leq \delta_{k,0} \leq \frac{2\eta^2}{\bar{\mu}}.$$

Therefore, the estimate

$$i' \leq \lceil 1 + \frac{16\eta^2\zeta^2}{\bar{\mu}^2(1-\bar{\theta})^2\bar{\xi}^2} \rceil$$

is obtained. Further, taking into account the last estimate and (48) the theorem is proved. Now let's propose a method which permits to find a point allowed to find a point from the set $X^*(\varepsilon)$ under the determined $\varepsilon > 0$ for a finite number of iterations.

Step 0. Assign $t = 0$. Choose a point $z_t \in \mathrm{R}^{\mathrm{n}}$. Determine parameters $\kappa > 0$, $\sigma \in (0,1)$, $\mu > 0$, $\theta \in (0,1)$.
Step 1. Compute $\xi_t = \kappa\sigma^t$.
Step 2. Find a point $z_{t+1} = \pi(z_t, \xi_t, \theta, \mu)$.
Step 3. Increase the value of t by one, and go to Step 1.

Remark 2. According to Steps 0, 4, 5 of the procedure π and Step 2 of the proposed method for each $t \in K$ we obtain

$$f(z_{t+1}) \leq f(z_t). \tag{49}$$

Therefore, the constructed sequence $\{f(z_t)\}$, $k \in K$, is non-increasing.

Theorem 3. *Suppose the sequence* $\{z_t\}$, $t \in K$, *is constructed by the proposed method. Then for each* $t \in K$ *it holds*

$$z_t \in L(z_0), \tag{50}$$

$$f(z_{t+1}) - f^* \leq \rho\sqrt{2\mu\xi_t} + \xi_t, \tag{51}$$

where $\rho > 0$ *is a diameter of the set* $L(z_0)$.

Proof. In accordance with Theorem 2 the procedure π is finite for each $t \in K$, and as already noted in Remark 2 for each $t \in K$ inequality (49) is fulfilled. Consequently, for each $t \in K$ we obtain inclusion (50).

In view of Lemma 4, Step 4 of the procedure π and Step 2 of the proposed method for each $t \in K$ we have

$$f(z_{t+1}) - f^* \leq \varrho_t\sqrt{2\mu\xi_t} + \xi_t,$$

where $\varrho_t > 0$ is a diameter of the set $L(z_t)$. Since for each $t \in K$ inequality (49) is fulfilled, then $L(z_t) \subset L(z_0)$, $t \in K$. Therefore, there is a constant $\rho > 0$ such that estimate (51) is determined for each $t \in K$.

Theorem 4. *Let* $\varepsilon > 0$ *and* $\rho > 0$ *be a diameter of the set* $L(z_0)$. *Then the complexity of the procedure of finding* ε*-solution by the proposed method is equal to*

$$\lceil 2\log_\sigma \varepsilon - \log_\sigma \kappa - 2\log_\sigma \hat{\rho}\rceil \lceil \frac{(f(z_0) - f^*)\hat{\rho}^2}{\theta\varepsilon^2}\rceil \lceil 1 + \frac{16\eta^2\zeta^2\hat{\rho}^4}{\mu^2(1-\theta)^2\varepsilon^4}\rceil, \tag{52}$$

where $\hat{\rho} = \rho\sqrt{2\mu} + \sqrt{\xi_0}$, $\eta = \eta(L(z_0)) > 0$, $\zeta = \zeta(\eta) > 0$.

Proof. From inequality (51) of Theorem 3 for each $t \in K$ it follows that

$$f(z_{t+1}) - f^* \leq \xi_t^{1/2}(\rho\sqrt{2\mu} + \sqrt{\xi_t}).$$

Since according to Step 1 of the proposed method we have $\xi_t \leq \xi_0$, $\xi_t \to 0$, $t \in K$, then there exists a number $t' \in K$ such that for each $t \geq t'$ the expression

$$f(z_{t+1}) - f^* \leq \xi_t^{1/2}(\rho\sqrt{2\mu} + \sqrt{\xi_t}) \leq \sqrt{\kappa\sigma^t}(\rho\sqrt{2\mu} + \sqrt{\xi_0}) \leq \varepsilon \tag{53}$$

is defined.

If $t' = 0$, then the number of iterations in relations to t does not exceed the first multiplier of value (52). In this connection assume that $t' > 0$. Then from (53) under $t = t'$ it follows

$$t' \leq \lceil 2\log_\sigma \varepsilon - \log_\sigma \kappa - 2\log_\sigma(\rho\sqrt{2\mu} + \sqrt{\xi_0})\rceil, \tag{54}$$

and for each $p < t'$ the inequality

$$\frac{1}{\xi_p} \leq \frac{\hat{\rho}^2}{\varepsilon^2}. \tag{55}$$

is fulfilled.

Further, since for each $t \in K$ inclusion (50) is determined and $L(z_0)$ is a bounded closed set, then according to Theorem 2 under $S = L(z_0)$ there exists numbers $\eta = \eta(L(z_0))$, $\zeta = \zeta(\eta) > 0$ such that for each $t < t'$ complexity of finding the point z_{t+1} on basis of the point z_t by the procedure π equals

$$\lceil \frac{f(z_t) - f^*}{\theta \xi_t} \rceil \lceil 1 + \frac{16\eta^2\zeta^2}{\mu^2(1-\theta)^2\xi_t^2} \rceil.$$

Hence and from (54), (55), $f(z_{t+1}) \leq f(z_t)$, $t < t'$ it follows that general complexity of the proposed method equals

$$\sum_{j=0}^{t'-1} \lceil \frac{f(z_j) - f^*}{\theta \xi_j} \rceil \lceil 1 + \frac{16\eta^2\zeta^2}{\mu^2(1-\theta)^2\xi_j^2} \rceil \leq \sum_{j=0}^{t'-1} \lceil \frac{(f(z_0) - f^*)\hat{\rho}^2}{\theta \varepsilon^2} \rceil \lceil 1 + \frac{16\eta^2\zeta^2\hat{\rho}^4}{\mu^2(1-\theta)^2\varepsilon^4} \rceil \leq$$

$$\lceil 2\log_\sigma \varepsilon - \log_\sigma \kappa - 2\log_\sigma(\rho\sqrt{2\mu} + \sqrt{\xi_0}) \rceil \lceil \frac{(f(z_0) - f^*)\hat{\rho}^2}{\theta \varepsilon^2} \rceil \lceil 1 + \frac{16\eta^2\zeta^2\hat{\rho}^4}{\mu^2(1-\theta)^2\varepsilon^4} \rceil.$$

The theorem is proved.

4 Conclusion

The bundle method is proposed for minimizing a convex function. To control the count of cutting planes the developed method updates the model of the objective function in case of obtaining good approximation quality of the epigraph in the neighborhood of the current iteration point. Moreover, at the moment of discarding cutting planes there are opportunities to involve any minimization method. The convergence of the proposed method is proved. Estimation of the complexity of finding an ε-solution is equal to $O(\varepsilon^{-6})$.

References

1. Zangwill, W.I.: Nonlinear Programming: A Unified Approach. Prentice-Hall, Englewood Cliffs (1969)
2. Zabotin, I.Y., Yarullin, R.S.: A cutting method and construction of mixed minimization algorithms on its basis. Uchen. Zap. Kazansk. Univ. Ser. Fiz.-Matem. Nauki. **156**(4), 14–24 (2014). (in Russian)
3. Bonnans, J., Gilbert, J., Lemarechal, C., Sagastizabal, C.: Numerical Optimization: Theoretical and Practical Aspects, 2nd edn. Springer, Heidelberg (2003). https://doi.org/10.1007/978-3-662-05078-1
4. Polyak, B.T.: Introduction to Optimization. Nauka, Moscow (1983). [in Russian]
5. de Oliveira, W., Eckstein, J.: A bundle method for exploiting additive structure in difficult optimization problems. Technical report (2015)
6. Kiwiel, K.C.: Efficiency of proximal bundle methods. J. Optim. Theory Appl. **104**, 589–603 (2000)

7. Shulgina, O.N., Yarullin, R.S., Zabotin, I.Ya.: A cutting method with approximation of a constraint region and an epigraph for solving conditional minimization problems. Lobachevskii J. Math. **39**(6), 847–854 (2018). https://doi.org/10.1134/S1995080218060197
8. Zabotin, I.Y., Yarullin, R.S.: One approach to constructing cutting algorithms with dropping of cutting planes. Russ. Math. (Iz. VUZ). **57**(3), 60–64 (2013)
9. Zabotin, I.Ya., Yarullin, R.S.: Cutting-plane method based on epigraph approximation with discarding the cutting planes. Autom. Rem. Control **76**(11), 1966–1975 (2015). https://doi.org/10.1134/S0005117915110065