A Projective Separating Plane Method with Additional Clipping for Non-Smooth Optimization

EVGENIYA A. VORONTSOVA
Far Eastern Federal University
Department of Mathematical Methods in Economics
Suhanova st. 8, 690091 Vladivostok
RUSSIAN FEDERATION
vorontsovaea@gmail.com

Abstract: This paper is devoted to constructing a definite efficient scheme for non-smooth optimization. A separating plane algorithm with additional clipping is proposed. The algorithm is used for solving the unconstrained non-smooth convex optimization problem. The latter problem can be reformulated as the computation of the value of a conjugate function at the origin. The algorithm convergence is proved. Encouraging results of numerical experiments are presented.

Key–Words: Non-smooth optimization, convex optimization, separating plane algorithms, black-box minimization

1 Introduction

Let $x = \{x^1, \ldots, x^n\}$ be the vector in n-dimensional real linear space $\mathbb{R}^n$. We consider the solvable unconstrained minimization problem

$$\min_{x \in \mathbb{R}^n} f(x) = f(x^*) \tag{1}$$

where $f(x)$ is a non-smooth convex objective function from $\mathbb{R}^n$ into $\mathbb{R}^1$. The objective function $f(x)$ is given by a black-box oracle.

Problems of kind (1) occur frequently in engineering, mechanics, control theory [1], economics [2] and many other fields. So there are many various areas of applications for methods solving these problems. Moreover, the area of large-scale optimization will have more advantages of any improvements in the methods for minimization of non-differentiable functions.

Research in this direction has resulted in several efficient methods [2-9], and others. Depending on what information about the objective function is available, optimization methods can be divided into two categories: white-box and black-box minimization. In white-box model we have some definite facts about the objective function's structure and/or behaviour. For example, a very interesting such approach for constructing efficient schemes for non-smooth optimization is proposed in [10]. The objective functions should have explicit max-structure. Vice versa, we know nothing about the function's structure in black-box model. In this paper we propose the last approach, namely, the black-box oracle model of the objective function. At each point $x \in \mathbb{R}^n$ we can receive the objective function value $f(x)$ and one subgradient g from the subdifferential set $\partial f(x)$.

Well-known subgradient methods [2] were the first numerical schemes for non-smooth convex optimization (see [11] for historical comments). These methods were under intense study since 1960s. At present these methods remain competitive over the low complexity of each iteration. However, subgradient methods have the slow rate of convergence. It was proved in the case black-box oracle model of the objective function that subgradient methods' complexity can not be better than $\mathbb{O}(\frac{1}{\varepsilon^2})$ iterations, where $\varepsilon > 0$ is the desired accuracy for the objective value [3]. So subgradient-type methods with space dilation [2] have been proposed. Then further development of non-smooth optimization methods went along the lines of cutting-plane methods, descent methods with bundle-type approach [4], descent methods with proximal-type approach (for example, [12], [13]), and the ellipsoid method ([2], [6]).

In this paper we present another class of non-smooth convex optimization methods with solving the corresponding Fenchel dual problem. The suggested algorithm for solving the problem (1) belongs to a type of separating plane algorithms [7]–[8] and improves them. As the next sections will show, the idea to replace the problem (1) with solving the corresponding Fenchel dual problem leads to some important merits such as increasing the rate of convergence.

The rest of this paper is organized as follows. The algorithm is proposed in Sect. 2. In Sect. 3, we present a convergence analysis for the proposed algorithm. In Sect. 4, we test the performance of the proposed algorithm and compare it with the standard SPA [8]. Sect. 5 concludes the paper.

2 Algorithm

The basic idea of separating plane algorithms (SPA) is to use the next to trivial identity of convex analysis [8, 14]:

$$\min_x f(x) = f(x^*) = -\max_x \{x \cdot 0 - f(x)\} = -f^*(0)$$

where $f^*(g) = \sup_x \{xg - f(x)\}$ is the Fenchel-Moreau conjugate of the function $f(x)$.

In this way the problem (1) can be reformulate as a problem of computing $f^*(0)$ (see Figure 1). The optimal point x^* can be obtained as a subgradient of f^*: $x^* \in \partial f^*(0)$.

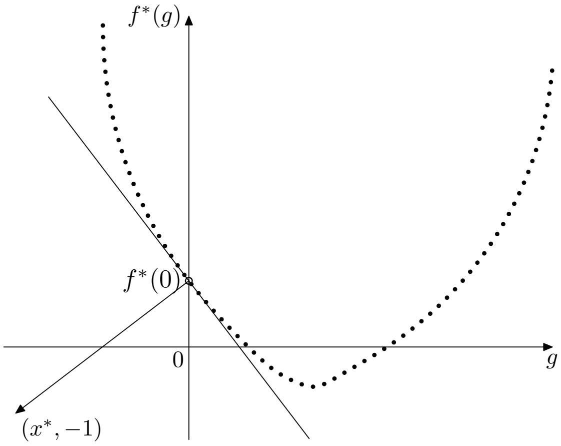

Figure 1: Graphical interpretation of the problem (1)

The SPA algorithms construct sequences of outer and inner approximations of the epigraph of f^* (epi $f^* = \{(\nu, g) : \nu \geq f^*(g)\}$). At each iteration of the algorithm the approximations are gradually refined. Eventually we obtain converging lower and upper bounds for $f^*(0)$.

As the SPA algorithm [8] has no guarantee of monotony, the following changes are suggested. At each iteration we execute an additional step which removes the upper part of epi f^*:

$$\sup_{\substack{(g,\,\varepsilon)\,\in\,\mathrm{epi}\,f^* \\ \varepsilon\,\leq\,\bar{v}}} \{gx - \varepsilon\}, \tag{2}$$

where estimate $\bar{v}$ is a solution of a linear-programming problem

$$\bar{v} = \min_{(0,\,\tau)\,\in\,\mathrm{co}\,((g_k,\,f^*(g_k)),\,k=1,\,2,\,\ldots)+0\times\mathbf{R}_+} \tau. \tag{3}$$

The SPA with additional clipping is illustrated in Figure 2.

Let us consider the set

$$\{\mathrm{co}\,((g_k,\, f^*(g_k)),\, k = 1, 2, \ldots) + 0 \times \mathbf{R}_+\} \tag{4}$$

where co S denotes the convex hull of a set S (the intersection of all convex sets which contain the set S). It is clear that the set (4) is contained in epi $f^*(g)$.

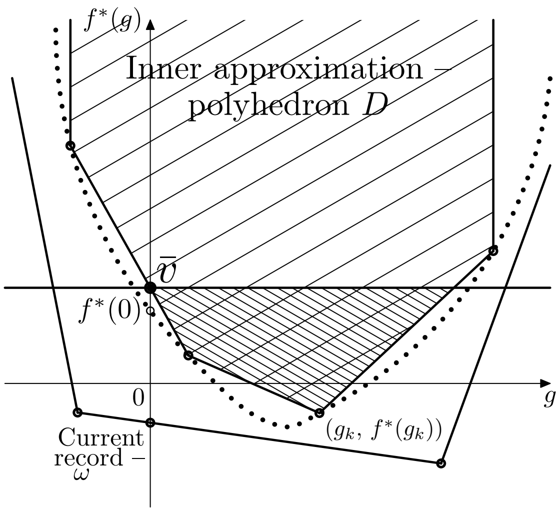

Figure 2: kth iteration of separating plane algorithm with additional clipping

Lagrangian function for the problem (2) is

$$L(g,\, \lambda) = gx - \varepsilon + \lambda(\bar{v} - \varepsilon).$$

Further, we reduce (2) to a line-search problem:

$$\sup_{(g,\,\varepsilon)\,\in\,\mathrm{epi}\,f^*;\ \varepsilon\,\leq\,\bar{v}} \{gx - \varepsilon\}$$

$$= \sup_{\varepsilon\,\geq\,f^*(g)}\ \inf_{\lambda\,\geq\,0} L(g,\, \lambda)$$

$$= \sup_{g} \inf_{\lambda \geq 0} \{gx - f^*(g) + \lambda(\bar{v} - f^*(g))\}$$

$$= \inf_{\lambda \geq 0} \{\lambda \bar{v} + \sup_{g} \{gx - (\lambda + 1) f^*(g)\}\}$$

$$= \inf_{\lambda \geq 0} \left\{ \lambda \bar{v} + (1 + \lambda) f\left(\frac{x}{1 + \lambda} \right) \right\}.$$

Finally, after the change of variables and excluding an independent of θ summand $-\bar{v}$:

$$\inf_{0<\theta\leq 1} \frac{1}{\theta}\Big(f(\theta x) + \bar{v} \Big) = \inf_{0<\theta\leq 1} \varphi(\theta). \qquad (5)$$

It is easy to prove that if $f(x)$ is a convex function, then the function of the single variable $\varphi(\theta)$ is a convex function too.

It is suggested to solve the one-dimensional non-smooth optimization problem (5) by means of a new modification of line-search algorithm for non-smooth convex optimization (see [15] for the original scheme). Convergence of this algorithm is quadratic.

Finally we come to the following **algorithmic scheme**.

Step 0. Initialization:

Set iteration counter $k := 0$, Given a starting point $x_0 \in$ **dom** f of minimizing sequence.

Step 1. Compute

$$\inf_{0 \in U_k(\omega)} \omega = \omega_k$$

where U_k is the k-th outer approximation of the epigraph of f^*. The latter problem can be solved recurrently:

$$\begin{aligned} \omega_k &= \inf_{0 \in U_k(\omega)} \omega \\ &= \inf_{0 \in U_{k-1}(\omega) \cap \{(g,\omega) \,|\, gx_{k-1} - \omega \leq f(x_{k-1})\}} \omega \\ &= \max\{ \inf_{0 \in U_{k-1}(\omega)} \omega, \inf_{\omega \geq -f(x_{k-1})} \omega \} \\ &= \max\{\omega_{k-1}, -f(x_{k-1})\}. \end{aligned} \qquad (6)$$

And $\omega_0 := -\infty$ for $k = 0$.

Step 2. Determine z. The point z is a projection of a point $(0, \omega_k)$ onto the polyhedron D. As mentioned above, the polyhedron D is inner approximations of the epigraph of f^* .

Step 3. Update.

Compute

$$x_k := -z(1:n) \,/\, z(n+1).$$

Then the last component of a vector

$$\bar{z} = -z \,/\, z(n+1) = (x_k, -1)$$

is equal to -1. That is needed (see Figure 1).

Step 4. Determine a cutting level of the upper part of epi f^*, *i.e.*, $\bar{v}$. The value $\bar{v}$ is found by solving the LP problem (3). If (3) has no solution, then go to the *Step 7.*

Step 5. Solve the one-dimensional non-smooth minimization problem (5). By θ_k denote the computed at the kth iteration solution of (5).

Step 6. Update 2. Compute $x_k := \theta_k x_k$.

Step 7. Add a pair $(g_k \in \partial f(x_k), f^*(g_k))$ to the inner approximation, *i.e.*, the polyhedron D.

Step 8. If stopping criterion is satisfied, then **quit**. Else increase iteration counter k and go to the *Step 1.*

3 Convergence Analysis

In this section we present a simple convergence analysis for the SPA with additional clipping.

Throughout the section, $\|.\|$ denotes the Euclidean norm of vectors.

Theorem 1 *Let $f(x)$ be a convex finite function, $f(0) = 0$, $\omega_* = -\min f(x) << \Omega < \infty$. Then $\lim\limits_{k\to\infty} \omega_k = \omega_*$.*

Proof: Using induction on k, it can be shown that for any k, we have $\omega_k \leq f^*(0)$.

Indeed, basis step: $\omega_0 = -\infty < f^*(0)$. According to (6),

$$\omega_k = \max\{\omega_{k-1}, -f(x_{k-1})\}$$

Inductive step. Besides,

$$\begin{aligned} &-f(x_{k-1}) = 0 \cdot x_{k-1} - f(x_{k-1}) \\ &\leq \sup_{x} \{0 \cdot x - f(x)\} = f^*(0) \end{aligned}$$

Let us consider two cases.

Case 1: $\omega_k = -f(x_{k-1})$. Then $\omega_k \leq f^*(0)$.

Case 2: $\omega_k = \omega_{k-1}$. By the inductive assumption, $\omega_k \leq f^*(0)$, which was to be proved.

Since $\omega_k \le f^*(0)$ and $\bar{z}_k + (0,\, \omega_k) \in \mathrm{co}\{D_k\}$, we only need to show that

$$\|\bar{z}_k\| \to 0, \text{ when } \quad k \to \infty$$

In order to prove decrease of the norm $\|\bar{z}_k\|$, let us consider several cases.

Case 1: The problem (3) at *Step 4* has no solution. Then the SPA with additional clipping turns into the standard SPA. The convergence of the latter algorithm was proved in [7].

Case 2: The problem (3) at *Step 4* has a solution. By $\bar{x}_k$ denote a solution of (2), $\bar{x}_k = \theta_k x_k$, $0 < \theta_k \le 1$. In this case, z_k can be represented as follows:

$$\bar{z}_k = -r_k(\theta_k x_k,\, -1)$$

Two cases can occur, depending on whether or not a current record of the objective function $-\omega_k$ changes at the k-th iteration.

Case 2.1: $\omega_k = \omega_{k-1}$. Then the projection is found from the same point:

$$\|\bar{z}_k\|^2 = \min_{z+(0,\,\omega_k)\,\in\,\mathrm{co}\{D'_k\}} \|z\|^2$$
$$= \min_{z+(0,\,\omega_{k-1})\,\in\,\mathrm{co}\{D'_{k-1}\}} \|z\|^2$$

Here D'_k is a polyhedron that is computed at the kth iteration, after the clipping (2).

And the following inequality holds:

$$\|\bar{z}_k\|^2 \le \min_{\lambda\in[0,\,1]} \|\bar{z}_{k-1} + \lambda\left((g_k,\, f^*(g_k)) - \bar{z}_{k-1}\right)\|^2.$$

The solution of the problem

$$\min_{\lambda\in[0,\,1]} \|\bar{z}_{k-1} + \lambda\left((g_k,\, f^*(g_k)) - \bar{z}_{k-1}\right)\|^2$$

is a projection of the minimum of a one-dimensional quadratic function on the interval $[0,\,1]$:

$$\lambda^* = \min\, \left\{\left(\bar{z}_{k-1} - (g_k,\, f^*(g_k))\right)\bar{z}_{k-1}\,/\right.$$
$$\left./\,\|(g_k,\, f^*(g_k)) - \bar{z}_{k-1}\|^2,\, 1\right\}.$$

It now follows that

$$\|\bar{z}_k\|^2 \le \|\bar{z}_{k-1} + \lambda\left((g_k,\, f^*(g_k)) - \bar{z}_{k-1}\right)\|^2$$

for any

$$\lambda \le \frac{\left(\bar{z}_{k-1} - (g_k,\, f^*(g_k))\right)\bar{z}_{k-1}}{\|(g_k,\, f^*(g_k)) - \bar{z}_{k-1}\|^2}. \tag{7}$$

Expanding an expression in (7) yields

$$\begin{aligned}&\|\bar{z}_k\|^2 \le \|\bar{z}_{k-1}\|^2\\ &-2\lambda\Big(\left(\bar{z}_{k-1} - (g_k,\, f^*(g_k))\right) \times \bar{z}_{k-1}\\ &\qquad -\frac{\lambda}{2}\|(g_k,\, f^*(g_k)) - \bar{z}_{k-1}\|^2\Big).\end{aligned} \tag{8}$$

Using (7), we get

$$\begin{aligned}&\left(\bar{z}_{k-1} - (g_k,\, f^*(g_k))\right)\bar{z}_{k-1}\\ &-\frac{\lambda}{2}\,\|(g_k,\, f^*(g_k)) - -\bar{z}_{k-1}\|^2\\ &\ge \frac{\lambda}{2}\,\|(g_k,\, f^*(g_k)) - \bar{z}_{k-1}\|^2 > 0\end{aligned}$$

for $\lambda \ne 0$.

To conclude the proof of decrease of the norm $\|\bar{z}_k\|$ in the case 2.1, it remains to combine (8) and the latter inequality: $\|\bar{z}_k\|^2 < \|\bar{z}_{k-1}\|^2$.

Case 2.2: $\omega_k = -f(x_{k-1}) > \omega_{k-1}$.
We have

$$\|\bar{z}_k\|^2 = \min_{z+(0,\,\omega_k)\,\in\,\mathrm{co}\{D'_k\}} \|z\|^2 \le \|\bar{z}_\lambda\|^2, \tag{9}$$

where

$$\bar{z}_\lambda = \frac{\Omega\, -\, \omega_k}{\Omega\, -\, \omega_{k-1}}\bar{z}_{k-1} < \bar{z}_{k-1}.$$

It follows easily that the latter inequality in (9) holds for such $\bar{z}_\lambda$. Therefore $\|\bar{z}_k\|^2 < \|\bar{z}_{k-1}\|^2$.

Now it follows from the monotonicity that there is a limit $\lim\limits_{k\to\infty} \|\bar{z}_k\| = \rho$.

Let us prove that $\rho = 0$. Assume the converse. Then $\|\bar{z}_k\| \ge \tau r_k$ for any $\tau > 0$.

Besides, it can be proved that

$$\lim_{k\to\infty}\left((g_{k+1}, f^*(g_{k+1})) - (0,\, \omega_k)\right)\bar{z}_k - \|\bar{z}_k\|^2 = 0.$$

Indeed, assume the converse. Suppose that

$$\left((g_{k'+1}, f^*(g_{k'+1})) - (0,\, \omega_{k'})\right)\bar{z}_{k'} \le \|\bar{z}_{k'}\|^2 - \gamma$$

with $\gamma > 0$ for some subsequence.

Then

$$\begin{aligned}&\|\bar{z}_{k'+1}\|^2 =\\ &\min_{z+(0,\omega_{k'})\in\mathrm{co}\left\{(g_i, f^*(g_i)), i=0,1,\cdots,k'+1\},(0,\Omega)\right\}} \|\bar{z}\|^2\\ &\le \min_{z\,+\,(0,\,\omega_{k'})\,=\,\lambda\tilde{g}_{k'}\,+\,(1-\lambda)(g_{k'+1},\, f^*(g_{k'+1}))} \|\bar{z}\|^2,\end{aligned}$$

with $\lambda \in [0, 1]$.

Since $\tilde{g}_{k'} = (0, \omega_{k'}) + \bar{z}_{k'}$, it follows that

$$\begin{aligned}
& \|\bar{z}_{k'+1}\|^2 \\
\le\ & \min_{\lambda \in [0,1]} \big\{ \|\lambda \bar{z}_{k'} + (1-\lambda)((g_{k'+1}, f^*(g_{k'+1})) \\
& \qquad - (0, \omega_{k'}))\|^2 \big\} \\
=\ & \min_{\lambda \in [0,1]} \big\{ \lambda^2 \|\bar{z}_{k'}\|^2 \\
& + 2\lambda(1-\lambda)\bar{z}_{k'}((g_{k'+1}, f^*(g_{k'+1})) - (0, \omega_{k'})) \\
& + (1-\lambda)^2 \|((g_{k'+1}, f^*(g_{k'+1})) - (0, \omega_{k'}))\|^2 \big\} \\
\le\ & \min_{\lambda \in [0,1]} \big\{ \lambda \|\bar{z}_{k'}\|^2 + 2\lambda(1-\lambda)\|\bar{z}_{k'}\|^2 \\
& - 2\gamma\lambda(1-\lambda) \\
& + (1-\lambda)^2 \|((g_{k'+1}, f^*(g_{k'+1})) - (0, \omega_{k'}))\|^2 \big\} \\
=\ & \min_{\lambda \in [0,1]} \big\{ (2\lambda - \lambda^2)\|\bar{z}_{k'}\|^2 - 2\gamma\lambda(1-\lambda) \\
& + (1-\lambda)^2 \|((g_{k'+1}, f^*(g_{k'+1})) - (0, \omega_{k'}))\|^2 \big\} \\
=\ & \min_{\lambda \in [0,1]} \big\{ \|\bar{z}_{k'}\|^2 - 2\gamma\lambda(1-\lambda) \\
& \quad + (1-\lambda)^2 \big(\|(g_{k'+1}, f^*(g_{k'+1})) - (0, \omega_{k'})\|^2 \\
& \quad - \|\bar{z}_{k'}\|^2 \big) \big\} \\
\le\ & \|\bar{z}_{k'}\|^2 - 2\gamma\lambda(1-\lambda) + (1-\lambda)^2 \delta^2
\end{aligned}$$

for any $\lambda \in [0, 1]$.

If we replace λ by $(\delta^2 + \gamma)/(\delta^2 + 2\gamma) > 0$, we obtain

$$\|\bar{z}_{k'+1}\|^2 \le \|\bar{z}_{k'}\|^2 - \gamma^2(\delta^2 + \gamma)/(\delta^2 + 2\gamma). \quad (10)$$

Taking the limit as $k' \to \infty$ in (10), we have a contradiction. Therefore

$$\lim_{k\to\infty} \big((g_{k+1}, f^*(g_{k+1})) - (0, \omega_k)\big)\bar{z}_k - \|\bar{z}_k\|^2 = 0.$$

Using

$$0 \le (\bar{z}_k + (0, \omega_k)\bar{z}_k - (g_{k+1}, f^*(g_{k+1}))\bar{z}_k \to 0$$

as $k \to \infty$, we get

$$(\bar{z}_k + (0, \omega_k))\bar{z}_k - (g_{k+1}, f^*(g_{k+1}))\bar{z}_k \le r_k^2 \varepsilon^2$$

for any $\varepsilon > 0$ and sufficiently large k. Hence,

$$\begin{aligned}
& (0, \omega_{k+1})\bar{z}_k \ge (g_{k+1}, f^*(g_{k+1}))\bar{z}_k \\
\ge\ & (0, \omega_k)\bar{z}_k + \|\bar{z}_k\|^2 - r_k\varepsilon \\
\ge\ & (0, \omega_k)\bar{z}_k + r_k^2\tau^2 - r_k^2\varepsilon^2 \\
\ge\ & (0, \omega_k)\bar{z}_k + r_k^2\tau^2/2
\end{aligned}$$

for any $\varepsilon \le \tau/\sqrt{2}$. That is,

$$r_k(0, \omega_{k+1}) \ge r_k(0, \omega_k) + r_k^2\tau^2/2.$$

Further,

$$f^*(0) \ge (0, \omega_{k+1}) \ge (0, \omega_k) r_k \tau^2 \ge (0, \omega_k) + \delta$$

where $\delta \ge r_k\tau^2 \ge 0$. But this is impossible as $k \to \infty$. This contradiction proves the equality $\lim_{k\to\infty} \|\bar{z}_k\| = 0$. □

4 Numerical Experiments

We conclude this paper with the results of numerical experiments. Numerical experiments demonstrated quite satisfactory computational performance of the separating plane algorithm with additional clipping. Moreover, the algorithm described above is compared with the standard SPA [8]. The both codes were written by the author in Octave programming language [16] under a Linux operating system.

4.1 Half-and-half function

The half-and-half function was created by *Lewis and Overton* [17] to analyze some optimization method behavior when minimizing a non-smooth function. The objective function to be minimized is given by

$$f(x) = \sqrt{x^T A x} + x^T B x, \ n = 8. \quad (11)$$

The matrix A is with all elements zero, except for ones on the diagonal at odd numbered locations; the matrix B is diagonal with elements $B(i, i) = 1/i^2$ for $i = 1, ..., 8$. Initial point x_0 is randomly chosen for every test.

To show the efficiency of the SPA with additional clipping, we adopt the performance profiles introduced in [18] to evaluate the number of iterations. The performance profile for a method is a nondecreasing, piecewise constant function, continuous from the right at each breakpoint. These profiles are very convenient for the performance evaluation of optimization methods. The relative efficiency of each method can be directly seen from graphs: the higher the particular curve, the better the corresponding method.

The results (see Figures 3–5) show that the SPA with additional clipping performs best, and this method is a more and more efficient with respect to computational times when the tolerance is increased.

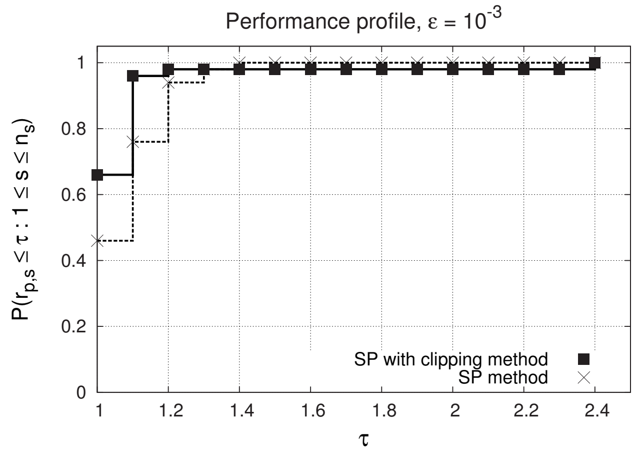

Figure 3: Performance profiles, the tolerance $\varepsilon = 10^{-3}$, the objective function (11)

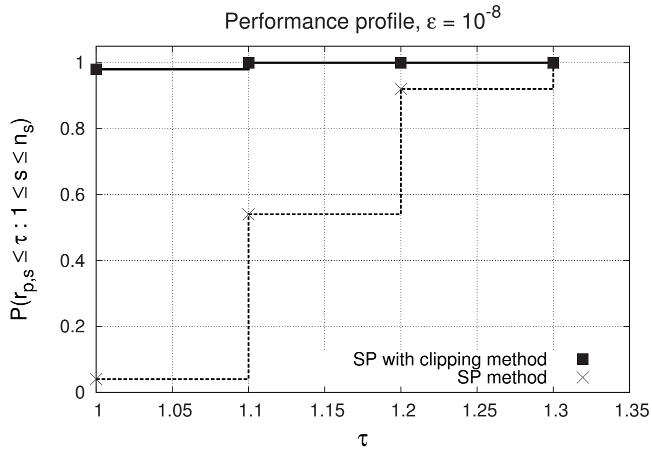

Figure 4: Performance profiles, the tolerance $\varepsilon = 10^{-8}$, the objective function (11)

4.2 MAXQUAD problem

Further, tests were made with the MAXQUAD function [19]:

$$f(x) = \max\{x^{\mathrm{T}} B_k\, x + b_k x \mid k = 1, ..., 5\}, \qquad x \in \mathbb{R}^n, \quad n = 10; \quad (12)$$

where

$$B_k(i,\, j) = e^{i/j}\cos(i * j)\sin(k),\; i < j;$$
$$B_k(i,\, i) = \frac{i}{n}|\sin(k)| + \sum_{j \neq i} |B_k(i, j)|;$$
$$b_k(i) = e^{i/k}\sin(i * k).$$

Matrices B_k are symmetrical. As is known, a symmetric diagonally dominant matrix with real non-negative diagonal entries is positive semi-definite.

Figure 6 shows speeds of descent to the minimum for the both methods. It is seen that SPA with clipping is better after ~ 40th iteration.

5 Conclusions

The projective separating plane method with additional clipping for non-smooth optimization is presented in this paper. Both the theoretical substantiation and investigation of the method and the results of numerical experiments suggest that this method is effective and widely applicable to non-smooth optimization problems with convex objective functions.

Acknowledgements: The author would like to thank Professor Evgeni A. Nurminski (Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences) for his help and attention. The author is also grateful to the anonymous referees for useful suggestions.

Funding: This work has been partially supported by RFBR [grant number 13-07-12010].

References:

[1] F. H. Clarke, Y. S. Ledyaev, R. J. Stern and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.

[2] N. Z. Shor, *Nondifferentiable Optimization and Polynomial Problems*, Kluwer Academic Publishers, Dordrecht, 1998.

[3] J.-L. Goffin, On the convergence rate of subgradient optimization methods, *Math. Program.* 13, 1977, pp. 329–347.

[4] J.-B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms II: Advanced theory and bundle methods. Fundamental Principles of Mathematical Sciences 306*, Springer-Verlag, Berlin, 1993.

[5] E. Polak and D. Q. Mayne, Algorithm models for nondiffentiable optimization, *SIAM J. Contr. Optim.* 23, No. 3, 1985, pp. 477–491.

[6] D. B. Yudin and A. S. Nemirovskii, Information complexity and effective methods for solving

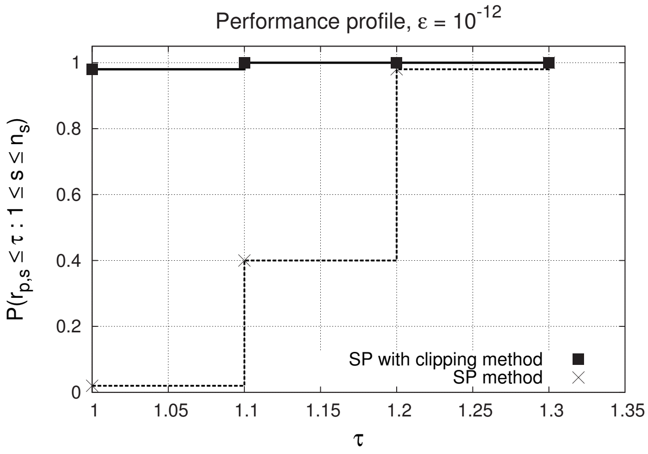

Figure 5: Performance profiles, the tolerance $\varepsilon = 10^{-12}$, the objective function (11)

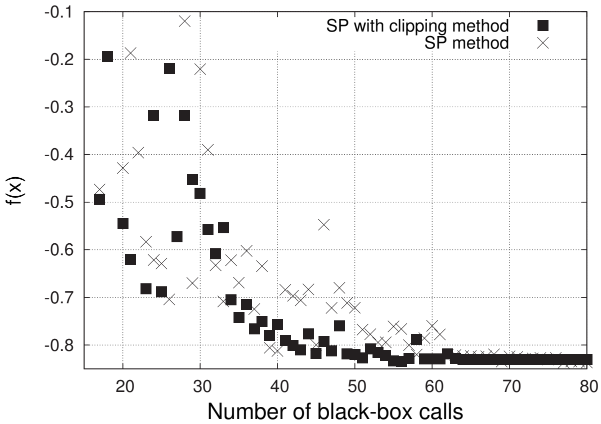

Figure 6: Computational results, the tolerance $\varepsilon = 10^{-15}$, the objective function (12)

convex extremum problems. *Ekonom. i Matem. Metody* 12(2), 1976, pp. 357–369.

[7] E. A. Nurminski, *Numerical Methods of Convex Optimization*, Nauka, Moscow, 1991 (in Russian).

[8] E. A. Nurminski, Separating plane algorithms for convex optimization, *Math. Program.* 76, 1997, pp. 373–391.

[9] A. Ben-Tal and A. Nemirovski, Non-Euclidean restricted memory level method for large-scale convex optimization, *Math. Program.*, 102, 2005, pp. 407–456.

[10] Yu. Nesterov, Smooth minimization of non-smooth functions, *Math. Program.* Ser. A 103, 2005, pp. 127–152.

[11] B. T. Polyak, *Introduction to Optimization*, Optimizaton Software, New York, 1987.

[12] K. C. Kiwiel, Proximity control in bundle methods for convex nondifferentiable minimization, *Math. Program.* 46, 1990, pp. 105–122.

[13] S. Kim, K.-N. Chang and J.-Y. Lee, A descent method with linear programming subproblems for nondifferentiable convex optimization, *Math. Program.* 71, 1995, pp. 17–28.

[14] E. A. Vorontsova, Separating plane algorithm with additional clipping for convex optimization, III International conference *Optimization and applications (OPTIMA-2012)* Proceedings, Costa da Caparica, Portugal, September 23-30, 2012, pp. 254-255.

[15] E. A. Nurminski, A quadratically convergent line-search algorithm for piecewise smooth convex optimization, *Optimization Methods and Software* 6, 1995, pp. 59–80.

[16] GNU Octave Page. URL: http://www.gnu.org/software/octave/index.html.

[17] R. Mifflin and C. Sagastizbal, A Science Fiction Story in Nonsmooth Optimization Originating at IIASA, *Documenta Math.* Extra Volume: Optimization Stories, 2012, pp. 291–300.

[18] E. Dolan and J. More, Benchmarking optimization software with performance profiles, *Math. Program.* 91, 2002, pp. 201–213.

[19] C. Lemarechal and R. Mifflin (eds.), Nonsmooth Optimization, *Proceedings of the IIASA Workshop* March 28 – April 8, 1977, Vol. 3, Oxford, Pergamon Press, 1978.