Modified Simplex Imbeddings Method in Convex Non-differentiable Optimization *

Anton Kolosnitcyn

Melentiev Energy Systems Institute,
Lermontov st. 130, Irkutsk, Russia
ankolos25@mail.ru

Abstract. We consider a new interpretation of the modified simplex imbeddings method. The main construction of this method is a simplex which contains a solution of convex non-differentiable problem. A cutting plane drawn through the simplex center is used to delete a part of the simplex without the solution. The most interesting feature of this method is the convergence estimation which depends only on the quantity of simplex vertices that are cut off by the cutting hyperplane. The more vertices are cut off by the cutting hyperplane, the higher rate of method convergence. We consider the special technique of constructing the simplex containing the set of points defining the truncated simplex. Such approach let us attribute the problem of constructing the minimal volume simplex to structural optimization problems that have quite efficient interior-point schemes for finding the optimal solution. The results of numerical experiment are also given in this paper.

Keywords: modified simplex imbeddings method, structural optimization, self-concordant barrier.

1 Introduction

We consider the problem of convex optimization in the next form

$$\begin{array}{l} f_0(x) \to \min, \\ f_i(x) \le 0,\ x \in \mathbb{R}^n, \end{array} \tag{1}$$

where $f_i(x)$, $i = 1, ..., m$ – convex not necessarily differentiable functions.

To solve this problem we can use quite extensive set of methods. One can notice that subgradient method was applied to solving convex non-differentiable problems historically first. The information about these methods can be found in [1], [2]. The main idea of subgradient methods is based on the using of arbitrary subgradient instead of gradient in the scheme of gradient method. However, in this case we cannot guarantee the relaxation sequence of approximations, but what we can obtain is monotonic decrease of the distance to the minimum point. One more feature of subgradient method

* The reported study was supported by RFBR, research project No. 15-07-08986.

In: A. Kononov et al. (eds.): DOOR 2016, Vladivostok, Russia, published at http://ceur-ws.org

is related to the rule of step length choosing. It has to be chosen as the tending to zero sequence. The main advantage of subgradient methods is the extremely simplicity of its using, if we know the simple way of subgradient determination. However, such methods are useless in practical applying due to the low rate of convergence and absence of reliable stop criteria. Nevertheless, it is suitable for rude approximation to the problem solution.

We can add reasonable stop criteria using group of piecewise linear approximation methods [3]. These methods collect the information about function value, subgradient value and approximation points from each iteration. This data allow us to construct piecewise linear approximation from below the objective function. In this case we come across quite low rate of convergence except some particular problems, moreover we need to solve linear programming problem at each iteration with growing input data. This method doesn't have quite robust characteristic due to the big gap between the behavior of function values and the sequence of approximation points .

Group of the bundle methods [4] is based on the piecewise linear approximation methods and includes a stabilization quadratic term to make method more robust. Additional term provides closeness of the current approximation value to the previous approximation. We can notice that it is possible to get superlinear rate of convergence for some classes of non-differential functions.

Important class of subgradient methods is represented by the space dilation methods in two variants: dilation in subgradient direction and dilation in two consistent subdifferentials residual direction [5], [6]. The main idea of space dilation method in subgradient direction is based on the attempt to decrease the sinus of the angle between the antisubgradient and minimum point direction. Sinus of the described angle for ravine functions tends to one. It prevents from getting geometric rate of convergence for subgradient method. Space dilation of the arguments in subgradient direction allows to get the geometric rate of convergence for certain cases.

Notice that particular case of space dilation method in subgradient direction is well known as the ellipsoid method and got wide popularity due to the Khachiyan work. Having used this method he constructed first polynomial algorithm for solving linear inequality systems [7]. Similar algorithm was developed in [8].

Space dilation methods in two consistent subdifferentials residual direction, named also r-algorithms, allow to convert obtuse angle between subgradients into sharp angle in extension dimension. It provides a decrease direction of function. Notice that we can get the quadratic rate of convergence using the r-algorithms for convex differentiable optimization problems.

It is applied analogues of conjugate gradient methods among different schemes of subgradient methods of solving convex non-differentiable problems. There are two main approaches of these methods [9]. First of them includes algorithms that collect subgradient package obtained from the previous iterations of method. This package is used for constructing the descent direction of objective function as the solution of the convex optimization problem. Different conditions define our further actions. We can restart algorithm, in other words, we change the starting point, or we can continue collecting subgradients, or finally take a step to the next point with certain step length defined through the one dimensional optimization problem. The main difficulty that appears in this approach is the size of subgradient package. It increases indefinitely if we increase the accuracy of the solution. The second approach of using the conjugate subgradient methods is related to the applying the analogue of the Polak-Ribiere formula. It is used for finding the descent direction of objective function. Such variant of conjugate subgradient methods demonstrates quite good results for smooth optimization problems. The association of two described approaches is considered in[10], and it is made the restriction of the volume of computational costs.

It is important to notice one more group of methods that are based on the center of feasible set [11]. In cutting plane methods we define the center of feasible set and construct the cutting hyperplane to except the part of feasible set that doesn't contain the points improving the objective function. We continue working with another part of the feasible set, find its center and construct one more cutting hyperplane. We continue this procedure up to getting the optimal point with required accuracy.

In [2] it is shown that the most effective method to construct cutting hyperplane is the gravity center method. However, finding the gravity center in a convex set is NP-hard problem that makes this method inapplicable in practical using. Notice some important substitutions for gravity center that are quite easy to find: center of inscribed ellipsoid, volumetric center and analytical center. In [11] it is shown that methods which are based on these types of centers have the polynomial difficulty.

In this paper we suggest the modification of simplex imbeddings method which is related to the cutting plane methods [12]. This method has a nonstandard estimation of rate convergence that depends on only the amount of cut off simplex vertices. One of the main principle of the method is constructing the minimal volume simplex containing a given truncated simplex. It is important to notice that in [12] the authors don't discuss the uniqueness of constructed minimal volume simplex. That is why we suggest method modification that uses the special technique of constructing the minimal volume simplex containing the set of points that define the truncated simplex. As we can obtain minimal volume simplices with different edges using different approaches of simplex constructing, we will compare three methods based on different simplex constructing techniques.

2 The Main Idea of Simplex Imbeddings Method

Recall that we need to solve the problem (1) using simplex imbeddings method. Suppose that we have the start simplex S_0 on the step $k = 0$, and this simplex contains the feasible set of the problem (1). We find the center $x^{c,0}$ of the simplex S_0 and construct the cutting hyperplane $L = \{x : g^T \left(x - x^{c,0}\right) = 0\}$ through the center, where $g \in \mathbb{R}^n$ is the subgradient of objective function. Then we move to the next step $k = k + 1$ and immerse the part of the simplex that contains the solution to the problem into the new simplex S_1 which has the minimal volume. Repeating this procedure we construct simplices that have less volume than previous ones and localize the problem solution consistently. We stop the method when the simplex volume becomes quite small.

We will need some important definitions that are given below.

Definition 1. *The simplex* $S \subset \mathrm{I\!R}^n$ *is the set in the next form*

$$S = \left\{ x \in R^n : \, x = x^0 + \sum_{i=1}^{n} \sigma_i \left(x^i - x^0 \right), \; \sigma_i \geq 0, \; \sum_{i=1}^{n} \sigma_i \leq 1 \right\},$$

where x^0 *is the simplex vertex,* $(x^1 - x^0, ..., x^n - x^0)$ *are the simplex edges that form basis in the* $\mathrm{I\!R}^n$*.*

Definition 2. *The point* x^c *is called the center of the simplex* S *and it is defined by the next expression*

$$x^c = \frac{1}{n+1} \left(x^1 + ... + x^n \right) \quad .$$

Definition 3. *The volume of the simplex* S *is defined by the next formula*

$$V(S) = \frac{1}{n!} \left| det \left(\bar{X} \right) \right|,$$

where $\bar{X}$ *is the* $n \times n$ *dimension matrix. The columns of this matrix are represented by the vectors* $x^1, x^2, ..., x^n$*.*

Definition 4. *We will call each hyperplane in the form*

$$L = \left\{ x : g^T \left(x - x^c \right) = 0 \right\}$$

as the cutting hyperplane passing through the center x^c *of the simplex* S*.*

We also need to define the base simplex imbeddings method as the method which is described in [12] and doesn't use additional minimax problems to construct resulting cutting hyperplanes described in [13].

3 Constructing the Minimal Volume Simplex. Rate of Convergence Estimation

We need to describe some key characteristics of the simplex imbeddings method. The immersion procedure of truncated simplex into the new minimal volume simplex is the important principle of the method providing the convergence to optimum problem solution. Information about constructing the minimum volume simplex containing truncated simplex is given in [12]. This approach is used as the element of the base simplex imbeddings method.

Now we concentrate our attention on the method convergence estimation. This estimation is formulated in [12] as the theorem.

Theorem 1. *Let the set* $S \subset \mathrm{I\!R}^n$ *be the n-dimensional simplex,* x^c *is the center of the simplex,* $S_G = \left\{ x \in S, \;\; g^T \left(x - x^c \right) \leq 0 \right\}$ *is truncated simplex. Simplex* S_G *can be immersed into the simplex* S^* *and the following relation between the volumes* $V(S)$ *and* $V(S^*)$ *of the simplices* S *and* S^* *is fulfilled:*

$$q_k^* = \frac{V\left(S^*\right)}{V(S)} \leq \begin{cases} \frac{1}{2} & k_l = 1; \\ \left(\frac{k_l}{k_l+1}\right)^{k_l} \left(\frac{k_l}{k_l-1}\right)^{k_l-1}, & 2 \leq k_l \leq n, \end{cases} \qquad (2)$$

where k_l *is the amount of saved vertices.*

The rate of convergence estimation (2) depends only on the amount of cut off simplex vertices. If it is cut off n vertices of the simplex we obtain the analogue of dichotomy method [3].

In [13] the author describes the modification of simplex imbeddings method that uses the feature of rate convergence estimation. The main idea of this modification is the introduction of several cutting hyperplanes and consideration its linear combination to construct only one resulting hyperplane [14]. Such technique was applied to the special class of convex optimization problems that was related to the polyhedral programming problems [15]. It was used the parametric description of functions subdifferentials to find the resulting hyperplane by means of solving the special minimax problems. Such hyperplanes cut off as many verticies of simplex as possible that let increase the rate of method convergence.

In the next section we will describe the idea of one more modification that uses the special principle of minimal volume simplex constructing.

4 The Main Idea of Simplex Imbeddings Method Modification

In [11] it is described the approach of constructing the minimal volume ellipsoid containing certain points. To obtain such ellipsoid containing the points $a^i \in \mathbb{R}^n$, $i = 1, ..., m$ we need to solve the following optimization problem:

$$\begin{aligned} &\tau \to min, \\ s.t. \ &-\ln det\, H \le \tau, \\ &\|Ha^i - v\| \le 1,\ i = 1, ..., m, \end{aligned} \tag{3}$$

where H is $(n \times n)$-symmetric positive semi-definite matrix, $v \in \mathbb{R}^n$, $\tau \in \mathbb{R}^1$.

Indeed, we can obtain the minimal volume ellipsoid by means of solving the problem (3) if we have some set of points like in the figure (1).

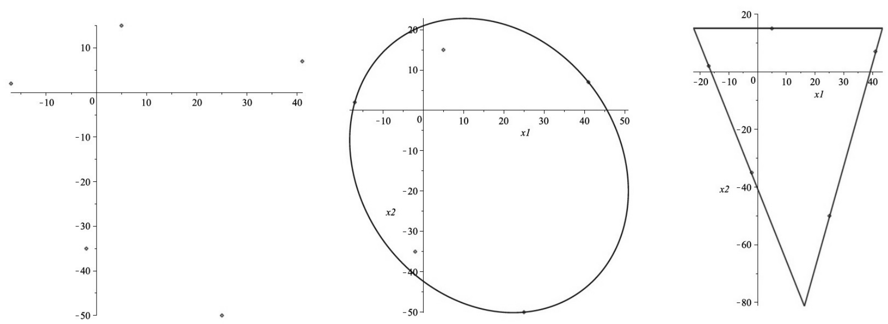

Fig. 1. Minimal volume objects

The problem (3) is related to the structural optimization problems [11]. It means that described problem has the self-concordant barrier for the feasible set. It let us apply the interior-point schemes for solving the problem (3). Such conclusion is very important for us as we can use very efficient and fast schemes for solving such problems.

We modified the approach of constructing the minimal volume ellipsoid and applied it to the constructing minimal volume simplex containing the set of defined points. If we have the set of defined points a^i, $i = 1, ..., m$ we can obtain minimal volume simplex containing all of these points by means of solving further optimization problem:

$$\begin{aligned} & \tau \to min, \\ s.t. \quad & -\ln det\, H \le \tau, \\ & Ha^i - v = u^i, \\ & u^i \ge 0, \ \textstyle\sum_{i=1}^m u^i \le 1, \\ & i = 1, ..., m, \end{aligned} \tag{4}$$

where $u^i \in \mathbb{R}^n$, $i = 1, ..., m$.

The problem (4) differs from (3) by the fact of matrix H doesn't suppose to be symmetric that leads to the complication of the problem (4). Nevertheless such approach of constructing the minimal volume simplex is applicable in practical using as we can estimate the number of verticies in truncated simplex. If k is the amount of saved verticies in the simplex, then we obtain $k(n+1-k)$ verticies in truncated simplex. Eventually the number of verticies in truncated simplex does not exceed $\lceil (n+1)/2 \rceil^2 \sim O(n^2)$. Such amount of points are easily calculated on modern computers.

In the figure (1) it is shown the example of the minimal volume simplex containing certain set of points. Thus we can transform the base simplex imbeddings method and use new principle of constructing the minimal volume simplex containing the certain points. Such technique is quite useful in terms of the applying efficient schemes of interior points methods that can enhance the rate of method convergence. The algorithm of simplex imbeddings method is described in [12] in details. In the next section we will give the results of numerical experiment using this algorithm with substitution of constructing minimal volume simplex technique to suggested modified principle of finding new simplex containing certain set of points.

5 Numerical experiment

Preliminary testing was carried out on the unconstrained problem of convex optimization

$$F(x) = \sum_{i=1}^{m} \left(\alpha_i \left| a_i^T x - b_i \right| + r_i \right) \to \min, \ \ x \in R^n \tag{5}$$

with the optimal point x^* that was known beforehand. The main idea of the test problem (5) is given in [13] in details. We solved the problems (5) only for dimension $n = 2$ due to the complexity of some constraints realization in the algorithm.

The calculations were carried out by means of program complex GAMS [16] on the computer with further configuration: AMD FX-8350/4.0 GHz processor, 8 GB of operative memory. The testing results are represented in the table 1 for three methods.

Table 1. Results of numerical experiment

m	k_B	T_B	k_{M_1}	T_{M_1}	k_{M_2}	T_{M_2}
100	45	0.28	39	1.22	48	0.33
500	48	0.36	48	1.83	52	0.39
1000	53	0.52	46	1.74	53	0.52
5000	47	3.06	54	4.56	44	2.74
10000	50	11.1	40	9.58	52	13.62
50000	45	376.3	36	292	51	417.8
100000	50	1993	38	1470	57	3040.9

First of them is the base simplex imbeddings method, described in [12]. The second method is the modified simplex imbeddings method that uses special technique of constructing the minimal volume simplex arround defined set of points. We will call it as the first method modification. And finally the third method uses the resulting cutting hyperplane for constructing minimal volume simplex. This method is described in [13]. We will call it as the second method modification. We took the following designations: m is the number of summands in the objective function, k_B is the number of iterations in the base simplex imbeddings method, T_B is the execution time of base method (in sec.), k_{M_1} is the number of iterations in the first method modification, T_{M_1} is the execution time of the first method modification (in sec.), k_{M_2} is the number of iterations in the second method modification, T_{M_2} is the execution time of the second method modification (in sec.). We give the average results of problems series, which contain 5 problems for each number of summands. The solution accuracy was equal to $\varepsilon = 10^{-3}$.

We can conclude that different approaches of constructing the minimal volume simplex give different results in the realization of methods algorithms. All three methods give close results considering the number of iterations, but for a significant number of summands the first method modification works faster. Moreover this modification is interesting in terms of the possibility of using the self-concordant barrier for optimization problem that give us quite good chance to improve the rate of method convergence.

References

1. Shor N.Z.: Minimization Methods for Non-differentiable Functions. Springer-Verlag, Berlin (1985)
2. Polyak B.T.: Introduction to Optimization. Optimization Software Inc. (1987)
3. Izmailov A.F., Solodov M.V.: Numerical Methods of Optimization (in Russian). Fizmatlit, Moscow (2003)
4. Lemarechal C., Nemirovskii A., Nesterov Ju. New Variants of Bundle Methods. Mathematical Programming, Volume 69, Numbers 1-3, pp. 111–147 (1995)
5. Shor N.Z.: Cutting Plane Method with Dimension Extension for Solving Convex Optimization Problems (in Russian). Kibernetika, Number 1, pp. 94–95 (1977)
6. Shor N.Z. Zhurbenko N.G.: Minimization Method with Dimension Extension Operation in Residual of Two Consistent Gradients Direction (in Russian). Kibernetika, Number 3, pp. 51–59 (1971)

7. Khachiyan L.G.: Polynomial Algorithms in Linear Programming (in Russian). Zh. Vychisl. Mat. Mat. Fiz., Volume 20, Number 1, pp. 51–68 (1980)
8. Nemirovski A.S., Yudin D.B.: Information Complexity and Efficient Methods for Solving Convex Extremal Problems (in Russian). Nauka, Moscow (1978)
9. Wolfe Ph.: A Method of Conjugate Subgradients for Minimizing Nondifferentiable functions. Nondifferentiable Optimization. Mathematical Programming Studies. Volume 3, Berlin, Heidelberg: Springer, pp. 145–173 (1975)
10. Nurminskii E.A., Tien D.: Method of Conjugate Subgradients with Constrained Memory. Automation and Remote Control, Volume 75, Issue 4, pp. 646–656 (2014)
11. Nesterov Yu.E.: Introductory Lectures on Convex Programming: A Basic Course. Springer Science + Business Media, New York (2004)
12. Antsiferov E.G., Bulatov V.P.: An Algorithm of Simplex Imbeddings in Convex Programming (in Russian). Zh. Vychisl. Mat. Mat. Fiz., Volume 27, Number 3, pp. 377–384 (1987)
13. Kolosnitcyn A.V.: Using of Modified Simplex Imbeddings Method for Solving Special Class of Convex Non-differentiable Optimization Problems. IIGU Ser. Matematika, Number 11, pp. 54–68 (2015)
14. Apekina Y.V., Khamisov O.V.: A Modified Simplex Immersions Method with Simultaneous Introduction of Several Intersecting Planes. Izv. Vyssh. Uchebn. Zaved. Mat., Number 12, pp. 16–24 (1997)
15. Filimonov A.B., Filimonov N.B.: Polyhedral Programming: Elements of Theoty and Applications. Information Technologies, Number 11, pp. 2–12 (1999)
16. Rosenthal R. GAMS A users guide. GAMS development corporation. Available at: www.gams.com (accessed 25.04.2016).