British Journal of Mathematics & Computer Science

22(3): 1-6, 2017; Article no.BJMCS.33880

ISSN: 2231-0851

Research on the Dual Problem of Trust Region Bundle Method

Jie Shen[1] and Ya-Li Gao[1*]

[1] *School of Mathematics, Liaoning Normal University, Dalian 116029, China.*

Authors' contributions

This work was carried out in collaboration by both authors. Author YLG designed the study, optimized the method, and wrote the first draft of the manuscript. Author JS proposed the concerned problem, and finished the final manuscript. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMCS/2017/33880
Editor(s):
(1) Dijana Mosic, Department of Mathematics, University of Nis, Serbia.
Reviewers:
(1) Xiaolan Liu, Sichuan University of Science & Engineering, China.
(2) Yasemin Sagiroglu, Karadeniz Technical University, Turkey.
Complete Peer review History: http://www.sciencedomain.org/review-history/19221

Received: ***2**nd* ***May 2017***
Accepted: ***19**th* ***May 2017***
Published: ***27**th* ***May 2017***

Original Research Article

Abstract

With the rapid development of science and technology as well as the cross-integration between the various disciplines, the nonsmooth optimization problem plays an increasingly important role in operational research. In this paper, we use the trust region method to study nonsmooth unconstrained optimization problems. Trust region subproblem is constructed to produce the next iteration point by using feasible set as constraint condition. As the number of iterations increases, the compression principle is used to control the elements in a bundle of information. And then the subproblem is studied by Lagrangian function and penalized bundle method [1]. The optimal solution and the relevant derivative conclusion are obtained by transforming the primal problem and dual problem into each other.

Keywords: Nonsmooth optimization; trust region bundle method; subgradient; dual problem.

[1] **Corresponding author: E-mail: 704756727@qq.com;*

1 Introduction

Nonsmooth optimization problem is also called non-differentiable optimization [2], because it does not have the property of continuous differentiability, the traditional differential theory and methods are no longer applicable. Bundle method is now recognized as one of the most efficient and promising methods to solve the problem of nonsmooth optimization[3][4][5][6][7][8], which guarantees a certain stability at the same time as the objective function descent. This method stores the acquired points in a bundle of information:$\mathbf{B} = \{f_i, y^i, s^i, i = 1, \cdots, k\} \bigcup \{x^k\}$, where x^k is the best objective function value[1]. In this paper, we mainly study the trust region bundle method, which uses the information in the black-box to construct the piecewise-linear affine model of the objective function in the primal problem, and the regard feasible set as a constraint condition to construct the next iteration point. With the increase of the number of iterations, we use the compression mechanism to control the size of subproblems. Through the idea of dual space [9][10][11][12], we study the Lagrangian function and dual problem of trust region subproblems, and describe the relationship between the original problem and the dual problem relationship, two important conclusions are obtained.

2 Preliminaries

In order to ensure the objective function decreases, the bundle method [1] remembers the point where the best decrease has been made so far. As the iteration goes along, the algorithm generate two sequences. One sequence is a sample point that defines the model, which is called the candidate point and we denote them by $\{y^k\}$. The other sequence is a sample point with an effective descent of the objective function, which is called the stability centers and we denote them by $\{x^k\}$(it is a subsequence of $\{y^k\}$). Assuming that the current stability center is $\{x^k\}$. The trust region bundle method is to find the point where the model function decrease most in the sphere centered at the stability center $\{x^k\}$, with κ_k as the radius of the sphere. We construct the trust region subproblem as follows:

$$\begin{cases} \min & \varphi_k(y) \\ s.t. & |y - x^k|_k^2 \leqslant \kappa_k, \end{cases} \tag{2.1}$$

where $\kappa > 0$ is the trust region radius, $\kappa_k \to 0$, as $k \to \infty$. The corresponding nominal decrease is defined by $\delta_{k+1} := f(x^k) - \varphi_k(y^{k+1})$. The linearization errors of f at x^k is defined by $(0 \leqslant) e_i := f(x^k) - f_i - \langle s^i, x^k - y^i \rangle, i = 1, 2, \cdots k$. As the number of iterations increases, the number of elements in the bundle of information becomes more and more, we use compression mechanism to keep the elements in the information bundle to np_k. Now, we define the piecewise-linear model function as

$$\varphi_k(y) = f(x^k) + \max_{i=1,\cdots,np_k} \{-e_i + \langle s^i, y - x^k \rangle\}.$$

For all $x \in R^n$, the norms of primal space and the corresponding dual space are denoted by $|x|_k^2 = \langle M_k x, x \rangle = x^T M_k^T x$ and $\|x\|_k^2 = \langle x, M_k^{-1} x \rangle = x^T M_k^{-1} x$, where M_k is a positive definite matrix.

3 Subproblem of Trust Region Bundle Method

Theorem 1 Given the parameter $\kappa_k > 0$, let y^{k+1} be the optimal solution to (2.1), Then

$$y^{k+1} = x^k - \frac{1}{2\beta} M_k^{-1} \hat{s}^k, \ \textit{where}\ \hat{s}^k = \sum_{i=1}^{np_k} \bar{\alpha}_i s^i, \tag{3.1}$$

and $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \cdots, \bar{\alpha}_{np_k})$ is an optimal solution to the following problem

$$\begin{cases} \min_{\alpha_i,\beta} & \frac{1}{4\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2 + \beta\kappa_k + \sum_{i=1}^{np_k} \alpha_i e_i \\ \textit{subject to} & \alpha \in \Delta_k := \{z \in [0,1]^{np_k} : \sum_{i=1}^{np_k} z_i = 1\} \\ & \beta \in R^1 \ . \end{cases} \tag{3.2}$$

Proof. Use an extra variable r to write the equivalent form of problem (2.1) as follows

$$\begin{cases} \min & r \\ s.t. & r \geqslant f(x^k) - e_i + \langle s^i, y - x^k \rangle, i = 1, 2, \cdots, np_k \\ & |y - x^k|_k^2 \leqslant \kappa_k \ . \end{cases} \tag{3.3}$$

The corresponding Lagrangian function is, for $\alpha \in R_+^{np_k}$, $\beta \in R^1$,

$$L(y, r, \alpha, \beta) = r + \sum_{i=1}^{np_k} \alpha_i (f(x^k) - e_i + \langle s^k, y - x^k \rangle - r) + \beta(|y - x^k|_k^2 - \kappa_k) \ ,$$

that is,

$$L(y, r, \alpha, \beta) = (1 - \sum_{i=1}^{np_k} \alpha_i) r + \sum_{i=1}^{np_k} \alpha_i (f(x^k) - e_i + \langle s^k, y - x^k \rangle) + \beta(|y - x^k|_k^2 - \kappa_k) \ .$$

In view of the strong convexity of objective function, (2.1) has the unique solution y^{k+1}. Therefore, there exists an optimal multiplier $(\bar{\alpha}, \beta)$ related to y^{k+1}. (y^{k+1}, α, β) can be obtained by solving the primal problem (3.3) or its dual problem.

$$\min_{(y,r) \in R^n \times R} \ \max_{\alpha \in R_+^{np_k}, \beta \in R^1} L(y, r, \alpha, \beta) \equiv \max_{\alpha \in R_+^{np_k}, \beta \in R^1} \ \min_{(y,r) \in R^n \times R} L(y, r, \alpha, \beta).$$

The above problem is equivalent to the problem (3.3), they have the same finite optimal value. However, interior programming is an unconstrained minimization problem, we have $1 - \sum_{i=1}^{np_k} \alpha_i = 0$, So y^{k+1} and $(\bar{\alpha}, \beta)$ are solved by primal and dual problems, respectively,

$$\min_{y \in R^n} \ \max_{\alpha \in R_+^{np_k}, \beta \in R^1} L(y, \alpha, \beta) \equiv \max_{\alpha \in R_+^{np_k}, \beta \in R^1} \ \min_{y \in R^n} L(y, \alpha, \beta),$$

where

$$\begin{aligned} L(y, \alpha, \beta) &= \sum_{i=1}^{np_k} \alpha_i (f(x^k) - e_i + \langle s^k, y - x^k \rangle) + \beta(|y - x^k|_k^2 - \kappa_k) \\ &= f(x^k) + \sum_{i=1}^{np_k} (-\alpha_i e_i + \langle \alpha_i s^i, y - x^k \rangle) + \beta((y - x^k)^T M_k (y - x^k) - \kappa_k). \end{aligned}$$

Consider the dual problem, for each $\alpha \in \Delta_k$, $\beta \in R^1$, the optimality conditions of $y(\alpha.\beta) = \arg\min_y L(y, \alpha, \beta)$ is $\nabla_y L(\alpha, y(\alpha, \beta)) = 0$, i.e.,

$$\nabla_y L(y, \alpha, \beta) = \sum_{i=1}^{np_k} \alpha_i s^i + 2\beta M_k (y - x^k) = 0. \tag{3.4}$$

Particularly, when $\alpha = \bar{\alpha}$, $y(\alpha, \beta) = y^{(k+1)}$, $y^{(k+1)} = x^k - \frac{1}{2\beta} M_k^{-1} \hat{s}^k$ holds. Next we prove that $\bar{\alpha}$ is the solution to the problem (3.2). Multiply (3.4) on both sides by M_k^{-1} and by $\frac{1}{2\beta} \sum_{i=1}^{np_k} \alpha_i s^i$,

$$0 = 2\beta(y(\alpha, \beta) - x^k) + M_k^{-1} \sum_{i=1}^{np_k} \alpha_1 s^i = \sum_{i=1}^{np_k} \alpha_i \langle s^i, y(\alpha, \beta) - x^k \rangle + \frac{1}{2\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2.$$

Further more, multiply (3.4) by $y(\alpha, \beta) - x^k$,

$$0 = 2\beta |y(\alpha, \beta) - x^k|_k^2 + \sum_{i=1}^{np_k} \alpha_i \langle s^i, y(\alpha, \beta) - x^k \rangle.$$

Thus, we have

$$0 = 2\beta |y(\alpha, \beta) - x^k|_k^2 + \sum_{i=1}^{np_k} \alpha_i \langle s^i, y(\alpha, \beta) - x^k \rangle = \sum_{i=1}^{np_k} \alpha_i \langle s^i, y(\alpha, \beta) - x^k \rangle + \frac{1}{2\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2. \quad (3.5)$$

From (3.5) we obtain

$$2\beta |y(\alpha, \beta) - x^k|_k^2 = \frac{1}{2\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2. \quad (3.6)$$

Using formula (3.6), further operations can be obtained

$$\begin{aligned} L(y(\alpha, \beta), \alpha, \beta) &= f(x^k) + \sum_{i=1}^{np_k} (-\alpha_i e_i + \langle \alpha_i s^i, y - x^k \rangle) + \beta(|y - x^k|_k^2 - \kappa_k) \\ &= f(x^k) + \frac{1}{4\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2 - \beta\kappa_k - \sum_{i=1}^{np_k} \alpha_i e_i + \langle \sum_{i=1}^{np_k} \alpha_i s^i, y - x^k \rangle \\ &= f(x^k) + \frac{1}{4\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2 - \beta\kappa_k - \sum_{i=1}^{np_k} \alpha_i e_i + \langle s^k, -\frac{1}{2\beta} M_k^{-1} s^k \rangle \\ &= f(x^k) - (\frac{1}{4\beta} \|\hat{s}^k\|_k^2 + \beta\kappa_k + \sum_{i=1}^{np_k} \alpha_i e_i). \end{aligned}$$

Altogether, $\bar{\alpha}$ is the solution to

$$\max_{\alpha \in \Delta_k, \beta \in R^1} L(y(\alpha, \beta), \alpha, \beta), i.e., \min_{\alpha \in \Delta_k, \beta \in R^1} \{ \frac{1}{4\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2 + \beta\kappa_k + \sum_{i=1}^{np_k} \alpha_i e_i \}.$$

Theorem 2 For trust region subproblem (2.1), the following conclusions hold:

(*a*) $\delta_{k+1} = \varepsilon_k + \frac{1}{4\beta} \|\hat{s}^k\|_k^2 + \beta\kappa_k$, where $\varepsilon_k := \sum_{i=1}^{np_k} \bar{\alpha}_i e_i$;

(*b*) Set $\gamma_k = \varepsilon_k - \frac{1}{4\beta} \|\hat{s}^k\|_k^2 + \beta\kappa_k$, if $\gamma_k \geqslant 0$, we have $\hat{s}^k \in \partial_{\gamma_k} f(x^k)$.

Proof. (*a*) Because there is no duality gap, the optimal solution of the primal problem (2.1) is equal to the optimal solution of the dual problem. According to the nominal decrease, we have $\varphi_k(y^{k+1}) = f(x^k) - \{ \frac{1}{4\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2 + \beta\kappa_k + \sum_{i=1}^{np_k} \alpha_i e_i \}$, and

$$\delta_{k+1} = \frac{1}{4\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2 + \beta\kappa_k + \sum_{i=1}^{np_k} \alpha_i e_i = \frac{1}{4\beta} \| \sum_{i=1}^{np_k} \alpha_i s^i \|_k^2 + \beta\kappa_k + \varepsilon_k.$$

(b) Note that $f \geqslant \varphi_k$, for any $y \in R^n$, $f(y) \geqslant \varphi_k(y) \geqslant \varphi_k(y^{k+1}) + \langle \hat{s}^k, y - y^{k+1} \rangle$. Use (3.1), the above inequality can be rewritten as

$$\begin{aligned} f(y) &\geqslant \varphi_k(y^{k+1}) + \langle \hat{s}^k, y - x^k \rangle - \langle \hat{s}^k, y^{k+1} - x^k \rangle \\ &= f(x^k) + \langle \hat{s}^k, y - x^k \rangle - (f(x^k) - \varphi_k(y^{k+1}) - \frac{1}{2\beta} \|\hat{s}^k\|_k^2) \\ &= f(x^k) + \langle \hat{s}^k, y - x^k \rangle - (\varepsilon_k + \beta \kappa_k - \frac{1}{4\beta} \|\hat{s}^k\|_k^2) \\ &= f(x^k) + \langle \hat{s}^k, y - x^k \rangle - \gamma_k. \end{aligned}$$

The conclusion $\hat{s}^k \in \partial_{\gamma_k} f(x^k)$ is obtained.

4 Conclusion

In this paper, we propose a new trust region bundle method to solve the problem of unconstrained nonsmooth optimization problem.By using the idea of penalized bundle method, we study the primal and dual problems of trust region subproblems respectively.

Acknowledgement

Supported by the National Natural Science Foundation of China (Grant 11301246) and Guidance Plan Project of Natural Science Foundation of Liaoning Province (Grant 20170540573)

Competing Interest

Authors have declared that no competing interests exist.

References

[1] Frédéric Bonnans J, Charles gilbert J, Claude Lemaréhal, Claudia A.Sagastizábal. Bundle Methods. The Quest for Descent[J]. Numerical Optimization. 2003;Π: 137-146.

[2] Yan Gao. Nonsmooth Optimization[M]. Operations Research and Management Science Series; 2008.

[3] Makela MM. Survey of bundle methods for nonsmooth. Optimization Methods and Software. 2001;7(1):1-29.

[4] Welington Oliveira, Mikhail Solodov. A doubly stabilized Bundle Method For nonsmooth convex Optimization; 2013.

[5] Claudia Sagastizábal. Composite proximal bundle method[J]. Mathematical Programming; 2012.

[6] Sagastizábal C, Hare W. A Redistributed Proximal Bundle Method for Nonconvex Optimization. SIAM Journal on Optimazation; 2010.

[7] Xiao-Qian Liu. An Approximate Bundle Method for Non-convex Optimization and Its convergence[D]. Liaoning Normal University; 2016.

[8] Jie Shen, Dan Li, Li-Ping Pang, Gaohang Yu. A Cutting Plane and Level Stabilization Bundle Method with Inexact Data for Minimizing Nonsmooth Nonconvex Functions[J]. Abstract and Applied Analysis; 2014.

[9] Jie Shen, Shui-Tian Cao˙ Research on the Dual Problem of Composite Proximal[J]. Journal of Jilin Normal University(Natural Science Edition). 2013;:1-4.

[10] Xuan Li. Research on the dual problem of several bundle methods for convex optimization problems[D]. Liaoning Normal University; 2015.

[11] Jie Shen, Xuan Li,Na Li. Research on the dual problem of doubly stabilized bundle method [J]. Journal of Jilin Normal University (Natural Science Edition). 2014;35(3):64-67.

[12] Jie Shen, Xiao-Qian Liu. An Approximate Bundle Method for Non-convex Optimization and Its Dual Problem[J]. Journal of Chongqing Normal University(Natural Science Edition). 2016;(1):1-5.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/19221