

# A New Technique for Estimating the Lower Bound of the Trust-Region Subproblem\*

Xinlong Luo

School of Information and Communication Engineering,  
 Beijing University of Posts and Telecommunications, Beijing, China  
 E-mail: luoxinlong@gmail.com, luoxinlong@bupt.edu.cn

Received January 23, 2011; revised February 10, 2011; accepted February 12, 2011

## Abstract

Trust-region methods are popular for nonlinear optimization problems. How to determine the predicted reduction of the trust-region subproblem is a key issue for trust-region methods. Powell gave an estimation of the lower bound of the trust-region subproblem by considering the negative gradient direction. In this article, we give an alternate way to estimate the same lower bound of the trust-region subproblem.

**Keywords:** Trust-Region Method, Unconstrained Optimization, Trust-Region Subproblem

## 1. Introduction

There are many methods for solving an unconstrained optimization problem

$$\min_{x \in \mathcal{R}^n} f(x), \quad (1)$$

where  $f$  is a smooth function. Line search methods [1,2] and trust-region methods [3] are the two popular classes of methods for the problem (1). In this article, we consider the trust-region method for this problem. The trust-region method for the problem (1) need to solve the following trust-region subproblem at every step:

$$\min_{s \in \mathcal{R}^n} \mathbf{q}_k(s) \equiv \mathbf{g}_k^T s + \frac{1}{2} s^T \mathbf{B}_k s, \quad (2)$$

$$\text{subject to } \|s\| \leq \Delta_k, \quad (3)$$

where  $\mathbf{g}_k = \nabla f(x_k)$  is the gradient of the objective function  $f(\cdot)$  at the current approximation solution  $x_k$ ,  $\mathbf{B}_k$  is an  $n$  by  $n$  symmetric matrix which approximates the Hessian matrix of  $f$ , and  $\Delta_k > 0$  is a trust-region radius.

It is a key issue to estimate the lower bound of the trust-region subproblem (2) and (3) for analyzing the global convergence of the trust-region methods for the problem (1) [1-5]. Powell [5] obtained a lower bound of

the subproblem (2) and (3) by considering the quadratic model  $\mathbf{q}_k(s)$  along the negative gradient direction  $-\mathbf{g}_k$ . In the next section we give an alternate way to obtain the same estimation of its lower bound. Throughout the paper  $\|\cdot\|$  denotes the Euclidean vector norm or its induced matrix norm.

## 2. A New Estimation Technique

In this section, we give an alternate way to estimate the lower bound and upper bound of  $\mathbf{q}_k(0) - \mathbf{q}_k(s_k)$ , respectively, where  $s_k$  is the solution of the subproblem (2) and (3). Firstly, we give the well-known properties of the trust-region problem [6,7]:

**Lemma 2.1** Vector  $s_k$  is a solution of the problem (2)-(3) iff  $\|s_k\| \leq \Delta_k$  and there exists  $\lambda_k \geq 0$  to satisfy

$$(\mathbf{B}_k + \lambda_k \mathbf{I})s_k = -\mathbf{g}_k, \quad (4)$$

$$(\mathbf{B}_k + \lambda_k \mathbf{I}) \succeq 0, \quad (5)$$

$$\lambda_k (\Delta_k - \|s_k\|) = 0, \quad (6)$$

where  $\mathbf{B}_k \in \mathcal{R}^{n \times n}$  is a symmetric matrix.

This lemma can be proved by the KKT (Karush-Kuhn-Tucker) conditions for the constraint optimization problem [4].

Powell [5] obtained an estimation of the lower bound of  $\mathbf{q}_k(0) - \mathbf{q}_k(s)$  in the trust region  $\|s\| \leq \Delta_k$  as follows:

**Lemma 2.2** (Powell, 1975 [5]) If  $s_k$  is a solution of

\*This research was supported in part by Grant 2009CB320401 from National Basic Research Program of China, and Grant YBWL2010152 from Huawei Technologies Co., Ltd., and Grant BUPT2009RC0118 from the Fundamental Research Funds for the Central Universities.

(2) and (3), then

$$\mathbf{q}_k(0) - \mathbf{q}_k(s_k) \geq \frac{1}{2} \min\{\Delta_k, \|\mathbf{g}_k\|/\|\mathbf{B}_k\|\} \|\mathbf{g}_k\|. \quad (7)$$

The estimation of the predicted reduction (7) is a key property to analyze the convergence of the trust-region method. Here, we are interested in the question whether we can obtain the analogous result of (7) if we directly consider the solution of (4) and (5). Our motivation is that we obtain the solution of the trust-region subproblem (2) and (3) by solving (4)-(6), and we need to directly consider the search direction  $s_k$  of (4) and its restricted step [2] or the parameter  $\lambda_k$  [8-11] for some trust region methods. Therefore, we give an alternate way to estimate the lower bound of  $\mathbf{q}_k(0) - \mathbf{q}_k(s_k)$  by directly considering the solution (4)-(6) of the trust-region subproblem (2) and (3).

**Lemma 2.3** *If  $\lambda_k \geq 0$  such that (5) is satisfied and  $s_k$  is a solution of (4), then*

$$\mathbf{q}_k(0) - \mathbf{q}_k(s_k) \geq \frac{1}{2} \|\mathbf{g}_k\| \min\{\|s_k\|, \|\mathbf{g}_k\|/\|\mathbf{B}_k\|\}. \quad (8)$$

**Proof.** From (4) we get

$$s_k^T \mathbf{B}_k s_k = -\lambda_k s_k^T s_k - \mathbf{g}_k^T s_k. \quad (9)$$

Combining the above equality and (4)-(5), we have

$$\begin{aligned} \mathbf{q}_k(0) - \mathbf{q}_k(s_k) &= -\frac{1}{2} s_k^T \mathbf{B}_k s_k - \mathbf{g}_k^T s_k = \frac{1}{2} \lambda_k s_k^T s_k - \frac{1}{2} \mathbf{g}_k^T s_k \\ &= \frac{1}{2} \lambda_k \|s_k\|^2 + \frac{1}{2} \mathbf{g}_k^T (\mathbf{B}_k + \lambda_k \mathbf{I})^+ \mathbf{g}_k \\ &\geq \frac{1}{2} \left( \lambda_k \|s_k\|^2 + \frac{1}{\lambda_k + \|\mathbf{B}_k\|} \|\mathbf{g}_k\|^2 \right), \end{aligned} \quad (10)$$

where  $(\mathbf{B}_k + \lambda_k \mathbf{I})^+$  is the generalized inverse of  $\mathbf{B}_k + \lambda_k \mathbf{I}$ . Now we consider the properties of the function

$$\varphi(\lambda) \equiv \lambda \|s_k\|^2 + \frac{1}{\lambda + \|\mathbf{B}_k\|} \|\mathbf{g}_k\|^2. \quad (11)$$

It is not difficult to know that the function  $\varphi(\lambda)$  is convex when  $\lambda + \|\mathbf{B}_k\| > 0$ , since

$$\varphi''(\lambda) = 2 \|\mathbf{g}_k\|^2 / (\lambda + \|\mathbf{B}_k\|)^3 \geq 0. \text{ Thus, the function } \varphi(\lambda)$$

attains its minimizer  $\varphi(\lambda_{min})$  when  $\lambda_{min}$  satisfies

$$\varphi'(\lambda_{min}) = 0 \text{ and } \lambda \geq -\|\mathbf{B}_k\|, \text{ i.e.}$$

$$\varphi(\lambda_{min}) = 2 \|\mathbf{g}_k\| \|s_k\| - \|\mathbf{B}_k\| \|s_k\|^2, \quad (12)$$

when

$$\lambda_{min} = \|\mathbf{g}_k\|/\|s_k\| - \|\mathbf{B}_k\|, \text{ and } \lambda_{min} > -\|\mathbf{B}_k\|. \quad (13)$$

We will prove this property by distinguishing two cases separately, namely  $\lambda_{min}$  is nonnegative or negative.

When  $\|\mathbf{g}_k\|/\|s_k\| \geq \|\mathbf{B}_k\|$ , from (13), we have  $\lambda_{min} \geq 0$ . In this case, combining  $\lambda_k \geq 0$  with (10)-(13), we obtain

$$\begin{aligned} \mathbf{q}_k(0) - \mathbf{q}_k(s_k) &\geq \lambda \|s_k\|^2 + \frac{1}{\lambda + \|\mathbf{B}_k\|} \|\mathbf{g}_k\|^2 \\ &= \varphi(\lambda) \geq \varphi(\lambda_{min}) \\ &= \frac{1}{2} (2 \|\mathbf{g}_k\| \|s_k\| - \|\mathbf{B}_k\| \|s_k\|^2) \\ &\geq \frac{1}{2} \|\mathbf{g}_k\| \|s_k\|. \end{aligned} \quad (14)$$

The other case is  $\|\mathbf{g}_k\|/\|s_k\| < \|\mathbf{B}_k\|$ , which gives  $\lambda_{min} < 0$  from (13). Since the function  $\varphi(\lambda)$  is monotonically increasing for all  $\lambda \geq \lambda_k \geq 0$  when  $\|\mathbf{g}_k\|/\|s_k\| < \|\mathbf{B}_k\|$ , from (10) and (11), we obtain

$$\begin{aligned} \mathbf{q}_k(0) - \mathbf{q}_k(s_k) &\geq \frac{1}{2} \left( \lambda_k \|s_k\|^2 + \frac{1}{\lambda_k + \|\mathbf{B}_k\|} \|\mathbf{g}_k\|^2 \right) \\ &= \frac{1}{2} \varphi(\lambda_k) \geq \frac{1}{2} \varphi(0) = \frac{1}{2} \|\mathbf{g}_k\|^2 / \|\mathbf{B}_k\|. \end{aligned} \quad (15)$$

Combining (14) with (15), we obtain

$$\mathbf{q}_k(0) - \mathbf{q}_k(s_k) \geq \frac{1}{2} \|\mathbf{g}_k\| \min\{\|\mathbf{g}_k\|/\|\mathbf{B}_k\|, \|s_k\|\}. \quad (16)$$

When  $\lambda_k = 0$  which shows the constraint is inactive, from (4) and (5), and the definition of  $q_k(s)$ , we have

$$\mathbf{q}_k(0) - \mathbf{q}_k(s_k) = \frac{1}{2} \mathbf{g}_k^T \mathbf{B}_k^+ \mathbf{g}_k \geq \frac{1}{2} \|\mathbf{g}_k\|^2 / \|\mathbf{B}_k\| \quad (17)$$

which shows that (8) is also true.  $\square$

### 3. Conclusions

In this article, we give an alternate way to estimate the lower bound of the trust-region method. This new technique can be applied to analyze the convergence of trajectory-following methods for unconstrained optimization problems [9-11]. The interested future works is that this new technique is applied to estimate the lower bound of the trust-region subproblem of the constraint optimization.

### 4. Acknowledgments

We would like to acknowledge Jiawang Nie and Hongchao Zhang for their constructive suggestions.

### 5. References

- [1] J. E. Dennis and R. B. Schnabel, "Numerical Methods for Unconstrained Optimization and Nonlinear Equations," Prentice-Hall, Inc., Englewood Cliffs, 1983.

- [2] R. Fletcher, "Practical Methods of Optimization," 2nd Edition, John Wiley & Sons, Chichester, 1987
- [3] A. R. Conn, N. Gould and P. L. Toint, "Trust-Region Methods," SIAM, Philadelphia, 2000.  
doi:10.1137/1.9780898719857
- [4] J. Nocedal and S. J. Wright, "Numerical Optimization," Springer-Verlag, New York, 1999. doi:10.1007/b98874
- [5] M. J. D. Powell, "Convergence Properties of a Class of Minimization Algorithms," In: O. L. Mangasarian, R. R. Meyer and S. M. Robinson Eds., *Nonlinear Programming 2*, Academic Press, New York, 1975, pp. 1-27.
- [6] D. M. Gay, "Computing Optimal Locally Constrained Steps," *SIAM Journal on Scientific and Statistical Computing*, Vol. 42, No. 2, 1981, pp. 186-197.  
doi:10.1137/0902016
- [7] J. J. Moré and D. C. Sorensen, "Computing a Trust Region Step," *SIAM Journal on Scientific and Statistical Computing*, Vol. 4, No. 3, 1983, pp. 553-572.  
doi:10.1137/0904038
- [8] D. J. Higham, "Trust Region Algorithms and Time Step Selection," *SIAM Journal on Numerical Analysis*, Vol. 37, No. 1, 1999, pp. 194-210.  
doi:10.1137/S0036142998335972
- [9] X.-L. Luo, C. T. Kelley, L.-Z. Liao and H. W. Tam, "Combining Trust Region Techniques and Rosenbrock Methods to Compute Stationary Points," *Journal of Optimization Theory and Applications*, Vol. 140, No. 2, 2009, pp. 265-286. doi:10.1007/s10957-008-9469-0
- [10] X.-L. Luo, L.-Z. Liao and H. W. Tam, "Convergence Analysis of the Levenberg-Marquardt Method," *Optimization Methods and Software*, Vol. 22, No. 4, 2007, pp. 659-678. doi:10.1080/10556780601079233
- [11] X.-L. Luo, "A Second-Order Pseudo-Transient Method for Steady-State Problems," *Applied Mathematics Computation*, Vol. 216, No. 6, 2010, pp. 1752-1762.  
doi:10.1016/j.amc.2009.12.029