Journal of Computational Mathematics
Vol.34, No.4, 2016, 421–436.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1601-m2015-0399

A RETROSPECTIVE TRUST REGION ALGORITHM WITH TRUST REGION CONVERGING TO ZERO*

Jinyan Fan
School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University, Shanghai 200240, China
Email: jyfan@sjtu.edu.cn
Jianyu Pan
Department of Mathematics, Shanghai Key Laboratory of PMMP, East China Normal University, Shanghai 200241, China
Email: jypan@math.ecnu.edu.cn
Hongyan Song
School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract

We propose a retrospective trust region algorithm with the trust region converging to zero for the unconstrained optimization problem. Unlike traditional trust region algorithms, the algorithm updates the trust region radius according to the retrospective ratio, which uses the most recent model information. We show that the algorithm preserves the global convergence of traditional trust region algorithms. The superlinear convergence is also proved under some suitable conditions.

Mathematics subject classification: 65K05, 65K10, 90C26, 90C30.
Key words: Retrospective trust region algorithm, Unconstrained optimization, Superlinear convergence.

1. Introduction

We consider the unconstrained optimization problem

$$\min_{x\in R^n} f(x), \tag{1.1}$$

where $f : R^n \to R$ is a continuous and differentiable function. Trust region methods are well-known methods for (1.1) (cf. [8, 10–12, 15–18]). At the k-th iteration, we construct the trust region subproblem

$$\begin{aligned} \min_{x\in R^n} \quad & m_k(x) := f(x_k) + g_k^T(x-x_k) + \frac{1}{2}(x-x_k)^T B_k (x-x_k) \\ s.t. \quad & \|x-x_k\|_2 \le \Delta_k, \end{aligned} \tag{1.2}$$

where $g_k = \nabla f(x_k)$ is the gradient of $f(x)$ at x_k, B_k is the Hessian matrix $\nabla^2 f(x_k)$ or its approximation, and Δ_k is the trust region radius. Suppose the optimal solution of (1.2) is x_k^*. Let $d_k = x_k^* - x_k$. Define the ratio of the actual reduction of the objective function to the predicted reduction:

$$r_k = \frac{Ared_k}{Pred_k} = \frac{f(x_k) - f(x_k+d_k)}{m_k(x_k) - m_k(x_k+d_k)}.$$

* Received September 21, 2015 / Revised version received December 14, 2015 / Accepted January 26, 2016 / Published online June 28, 2016 /

It plays an important role in deciding whether to accept the trial step d_k and how to adjust the trust region radius Δ_k. The next iterate x_{k+1} is chosen as

$$x_{k+1} = \begin{cases} x_k, & \text{if } r_k < p_0, \\ x_k + d_k, & \text{otherwise,} \end{cases} \tag{1.3}$$

where $0 \le p_0 < 1$ is a constant. In traditional trust region algorithms, the trust region radius for the next iteration is updated as

$$\Delta_{k+1} \in \begin{cases} [p_2\|d_k\|, p_3\Delta_k], & \text{if } r_k \le p_1, \\ [\Delta_k, p_4\Delta_k], & \text{otherwise,} \end{cases} \tag{1.4}$$

where p_i are positive constants satisfying $p_0 < p_1 < 1$ and $0 < p_2 < p_3 < 1 < p_4$.

As we know, the trust region radius will be larger than a positive constant in the traditional trust region algorithm when the algorithm converges. Note that the trial step will converge to zero, it suffices for the convergence that the trust region radius is larger than the distance from the iterate to the solution. In [5], Fan and Yuan proposed a novel trust region algorithm for (1.1) with the trust region converging to zero, in which the trust region radius is taken as

$$\Delta_{k+1} = \mu_{k+1}\|g_{k+1}\|, \tag{1.5}$$

where μ_{k+1} is computed according to the ratio r_k as follows:

$$\mu_{k+1} = \begin{cases} c_0\mu_k, & \text{if } r_k < p_0, \\ c_1\mu_k, & \text{if } r_k \ge p_1 \text{ and } \|d_k\| > \frac{1}{2}\Delta_k, \\ \mu_k, & \text{otherwise,} \end{cases} \tag{1.6}$$

where $0 \le p_0 < p_1 < 1$ and $0 < c_0 < 1 < c_1$ are positive constants.

Recently, Bastin et al. [1] proposed a retrospective trust region method for (1.1). The iterate is updated according to the ratio r_k as traditional trust region methods, but the trust region radius is no longer updated according to r_k. If the iteration is unsuccessful (i.e. the trial step is not accepted), the trust region radius is reduced. Otherwise if the iteration is successful (i.e. the trial step is accepted), the trust region radius is updated according to the retrospective ratio defined as

$$\tilde{r}_k = \frac{Ared_k}{\tilde{P}red_k} = \frac{f(x_k) - f(x_k + d_k)}{m_{k+1}(x_k) - m_{k+1}(x_k + d_k)}, \tag{1.7}$$

where the predicted reduction uses the model $m_{k+1}(x)$ at the new iterate rather than the model $m_k(x)$ at the preceding iterate. If $\tilde{r}_k$ is larger than a positive constant, the trust region radius is increased, otherwise it is unchanged or reduced. In retrospect, it might be more natural to update the trust region radius Δ_{k+1} according to how well $m_{k+1}(x)$ predicts the value of the objective function at the $(k+1)$-th iteration. This exploits the most relevant information on the model's quality at the current iterate rather than at the previous iterate.

Motivated by the retrospective idea, in this paper, we present a retrospective trust region algorithm with the trust region converging to zero for the unconstrained optimization problem. The iterate is updated according to the ratio r_k as usual, and the trust region radius is still computed by (1.5). But the parameter μ_k is no longer updated according to r_k. If the iteration