Available online at www.sciencedirect.com
ScienceDirect
Acta Mathematica Scientia 2018,**38B**(2):479–496
Mathematica Acta Scientia 数学物理学报
http://actams.wipm.ac.cn

A NEW ADAPTIVE TRUST REGION ALGORITHM FOR OPTIMIZATION PROBLEMS*

Zhou SHENG (盛洲) *Gonglin YUAN* (袁功林)† *Zengru CUI* (崔曾如)
College of Mathematics and Information Science, Guangxi University, Nanning 530004, China
E-mail: *szhou03@live.com; glyuan@gxu.edu.cn; cuizengru@126.com*

Abstract It is well known that trust region methods are very effective for optimization problems. In this article, a new adaptive trust region method is presented for solving unconstrained optimization problems. The proposed method combines a modified secant equation with the BFGS updated formula and an adaptive trust region radius, where the new trust region radius makes use of not only the function information but also the gradient information. Under suitable conditions, global convergence is proved, and we demonstrate the local superlinear convergence of the proposed method. The numerical results indicate that the proposed method is very efficient.

Key words Optimization; trust region method; global convergence; local convergence

2010 MR Subject Classification 65K05; 90C26

1 Introduction

Consider the following minimization problem:

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

where $f:\mathbb{R}^n\to\mathbb{R}$ is a twice continuously differentiable function. This problem has appeared in many applications in the medical science, optimal control, functional approximation, curve fitting fields, and other areas of science and engineering. Many methods are presented to solve this problem (1.1), including the conjugate gradient method (see [1–8]), the quasi-Newton method (see [9–14]), and the trust region method (see [15–28]).

As is known, the trust region method plays an important role in the area of nonlinear optimization, and is among efficient methods for solving problem (1.1). The trust region methods that was firstly proposed by Powell in [29] used of an iterative structure. At each iterative point x_k, a trial step d_k was obtained by solving the following subproblem

$$\begin{aligned} \min_{d\in\mathbb{R}^n} \quad & m_k(d) = g_k^T d + \frac{1}{2} d^T B_k d, \\ \text{s.t.} \quad & \|d\| \le \Delta_k, \end{aligned} \tag{1.2}$$

*Received July 9, 2016; revised October 13, 2016. Supported by the National Natural Science Foundation of China (11661009), the Guangxi Science Fund for Distinguished Young Scholars (2015GXNSFGA139001), the Guangxi Natural Science Key Fund (2017GXNSFDA198046), and the Basic Ability Promotion Project of Guangxi Young and Middle-Aged Teachers (2017KY0019).

†Corresponding author

where $g_k = \nabla f(x_k)$, B_k is an $n \times n$ symmetric approximation matrix of $\nabla^2 f(x_k)$, and Δ_k is the trust region radius.

Solving the trust region subproblem (1.2) plays a key role for solving problem (1.1) in the trust region algorithm. In traditional trust region methods, at each iterative point x_k, the trust region radius Δ_k is independent of the variables g_k, B_k, and $\|d_{k-1}\|/\|y_{k-1}\|$, where

$$y_{k-1} = g_k - g_{k-1}.$$

However, these variables contain first- and second-order information, which we do not use. We determine the initial trust region radius and some artificial parameters to adjust it, and the choice of these parameters has an important influence on our numerical results. Therefore, Zhang [25] proposed a modified trust region algorithm by replacing Δ_k with $c^p \frac{\|g_k\|}{a_k}$, where

$$0 < c < 1, \quad a_k = \max\{\|H_k\|, 1\}, \quad H_k = \nabla^2 f(x_k),$$

and p is a positive integer. With this modified trust region subproblem, instead of updating Δ_k, p is adjusted. However, at every iteration, we need to calculate the Hessian of f. On the basis of the technique described in [25], Zhang, Zhang, and Liao [19] presented a new trust region subproblem, where the trust region radius uses $\Delta_k := c^p \|\widehat{B}_k^{-1}\| \|g_k\|$, the definitions of c, p, g_k remain the same as in [25], and $\widehat{B}_k$ is a positive definite matrix based on the Schnable and Eskow [30] modified Cholesky factorization. We see that there are two drawbacks in computing the inverse of the matrix $\widehat{B}_k$ and the Euclidean norm of $\widehat{B}_k^{-1}$ at each iterative point x_k. Cui and Wu [24] presented a trust region radius by replacing Δ_k with $\mu_k \|\widehat{B}_k^{-1}\| \|g_k\|$, where $\mu_k > 0$ and satisfies an update rule, and for the same reason as that found by Zhang, Zhang, and Liao [19], this algorithm requires the calculation of $\widehat{B}_k^{-1}$ at every iteration. Motivated by the first- and second-order information of g_k and $\|d_{k-1}\|/\|y_{k-1}\|$, Li [26] proposed a modified trust region radius that uses $\Delta_k := \frac{\|d_{k-1}\|}{\|y_{k-1}\|} \|g_k\|$, which confers the advantage of avoiding to computation of the matrix $\widehat{B}_k$ of the inverse and $\|\widehat{B}_k^{-1}\|$ at each iterative point x_k. At the same time, it can decrease the workload and computational time involved. However, the algorithm only uses the gradient information. Some authors have also applied adaptive trust region methods to solve nonlinear equations in [31–33].

The purpose of this article is to present an efficient adaptive trust region algorithm to solve (1.1). Motivated by the adaptive technique, the proposed method possesses the following nice properties: (i) the trust region radius uses not only the gradient value but also the function value, (ii) computing the matrix $\widehat{B}_k$ of the inverse and the value of $\|\widehat{B}_k^{-1}\|$ at each iterative point x_k is not required, and (iii) the associated workload and computational time involved are reduced, which is very important for medium-scale problems.

The remainder of this article is organized as follows. In the next section, we briefly review some basic results of a modified quasi-Newton secant equation and present an adaptive algorithm for solving problem (1.1). In Section 3, we prove the global convergence of the proposed method. In Section 4, we prove the superlinear convergence of the algorithm under suitable conditions. Numerical results are reported in Section 5. We conclude this article in Section 6.

Throughout this article, unless otherwise specified, $\|\cdot\|$ denotes the Euclidean norm of vectors or matrices.