Journal of Computational Mathematics, Vol.21, No.2, 2003, 229–236.

A SELF-ADAPTIVE TRUST REGION ALGORITHM*1)

Long Hei†
(*Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing 100080, China*)
(*Department of Industrial Engineering and Management Sciences Northwestern University C230, 2145 Sheridan Road Evanston, Illinois 60208, USA*)

Abstract

In this paper we propose a self-adaptive trust region algorithm. The trust region radius is updated at a variable rate according to the ratio between the actual reduction and the predicted reduction of the objective function, rather than by simply enlarging or reducing the original trust region radius at a constant rate. We show that this new algorithm preserves the strong convergence property of traditional trust region methods. Numerical results are also presented.

Key words: Trust region, Unconstrained optimization, Nonlinear optimization.

1. Introduction

In this paper we study a new type of trust region method for solving the following unconstrained optimization problem:

$$\min_{x\in\Re^n} \quad f(x). \tag{1.1}$$

A trust region method calculates a trial step d_k by solving the trust region subproblem

$$\min_{d\in\Re^n} \quad \Phi_k(d) := g_k^T d + \tfrac{1}{2} d^T B_k d \tag{1.2}$$

$$s.t. \qquad ||d||_2 \le \Delta_k, \tag{1.3}$$

where $g_k = \nabla f(x_k)$ is the gradient of the objective function at the current approximate solution x_k, B_k is an $n \times n$ symmetric matrix approximating the Hessian of $f(x)$, and $\Delta_k > 0$ is the current trust region radius. Compared with the line search methods, one of the most important advantages of trust region methods is that B_k is allowed to be indefinite.

After obtaining a trial step d_k, which is an exact or approximate solution of subproblem (1.2)-(1.3), a trust region method computes the ratio ρ_k between the actual reduction in the objective function and the predicted reduction in the quadratic model of the objective function, that is,

$$\rho_k := \frac{Ared_k}{Pred_k} \tag{1.4}$$

$$= \frac{f(x_k) - f(x_k + d_k)}{\Phi_k(0) - \Phi_k(d_k)}. \tag{1.5}$$

Then the trust region radius Δ_k is updated according to the value of ρ_k. The common method for updating Δ_k is to enlarge it by a constant time (say, double):

$$\Delta_{k+1} = \beta_1 \Delta_k \qquad (\beta_1 > 1), \tag{1.6}$$

* Received July 1, 2000.
1) Supported partially by Chinese NSF grant.
† Email address: lhei@neruda.ece.nwu.edu.

if ρ_k is *satisfactory* enough, and to reduce it by a constant fraction (say, a half):

$$\Delta_{k+1} = \beta_2 \Delta_k \qquad (0 < \beta_2 < 1), \tag{1.7}$$

in the case ρ_k is not *satisfactory* enough. Now, if the trial step d_k is successful, one then accepts this step, and sets $x_{k+1} = x_k + d_k$; otherwise, the step d_k is rejected.

The trust region algorithm stated above is often used to solve problem (1.1). It converges globally and superlinearly. However, in the course of updating the trust region radius Δ_k, we do not make full use of the ratio ρ_k. In fact, the value of ρ_k, in some degree, reflects the extent to which the quadratic model $\Phi_k(d)$ approximates the objective function $f(x)$. Our goal in this paper is to design an algorithm in which Δ_k is updated at a variable rate according to the value of ρ_k directly.

2. Ideal Trust Region and R-Function

Now let us reconsider the idea of trust region algorithms. At the current solution x_k, if the trial step d_k is successful and the ratio ρ_k is satisfactory, one accepts the trial step and enlarges the trust region radius. On the contrary, if the trial step d_k is not successful and the ratio ρ_k is not satisfactory, d_k is rejected and Δ_k is reduced. As the ratio between the actual reduction $Ared_k$ and the predicted reduction $Pred_k$, ρ_k reflects the extent to which we are satisfied with the solution d_k of the subproblem (1.2)-(1.3), or to say, the extent to which the quadratic model $\Phi_k(d)$ approximates the original objective function $f(x)$.

We now think about the extreme case when ρ_k is $+\infty$, which means the computed step d_k is very successful. At this time we may, from the idealized point of view, enlarge the trust region radius Δ_k greatly, even to $+\infty$. In the other extreme case, say, when ρ_k is $-\infty$, which implies that the trial step d_k is so *bad* that it causes the objective function value to rise rapidly, it is then reasonable for us to imagine the trust region radius Δ_k should be reduced to a very small value, even near 0. These ideal cases, which we call *ideal trust region*, inspire us to study the following type of functions of ρ_k, named *R-function*:

Definition *Any one-dimensional function $R_\eta(t)$ that is defined in $\Re = (-\infty, +\infty)$ with the parameter $\eta \in (0,1)$ is an* R-function *if and only if it satisfies:*
*(**i**) $R_\eta(t)$ is non-decreasing in $(-\infty, +\infty)$;*
*(**ii**)*

$$\lim_{t \to -\infty} R_\eta(t) = \beta \qquad (\textit{where } \beta \in [0,1) \textit{ is a small constant}); \tag{2.1}$$

*(**iii**)*

$$R_\eta(t) \le 1 - \gamma_1 \quad (\textit{for all } t < \eta, \textit{ where } \gamma_1 \in (0, 1-\beta) \textit{ is a constant}); \tag{2.2}$$

*(**iv**)*

$$R_\eta(\eta) = 1 + \gamma_2 \qquad (\textit{where } \gamma_2 \in (0, +\infty) \textit{ is a constant}); \tag{2.3}$$

*(**v**)*

$$\lim_{t \to +\infty} R_\eta(t) = M \qquad (\textit{where } M \in (1+\gamma_2, +\infty) \textit{ is a constant}). \tag{2.4}$$

From this definition we can easily see some properties of R-functions:

Theorem 2.2. *An R-function $R_\eta(t)$ (where $\eta \in (0,1)$) satisfies:*

$$0 < \beta \le R_\eta(t) \le 1 - \gamma_1 < 1, \qquad \forall t \in (-\infty, \eta); \tag{2.5}$$

$$1 < 1 + \gamma_2 \le R_\eta(t) \le M < +\infty, \qquad \forall t \in [\eta, +\infty). \tag{2.6}$$