Journal of Computational Mathematics, Vol.14, No.3, 1996, 195–202.

A RESTRICTED TRUST REGION METHOD WITH SUPERMEMORY FOR UNCONSTRAINED OPTIMIZATION*1)

L.P. Sun
(*Department of Mathematics, Nanjing University, Jiangsu, China*)

Abstract

A new method for unconstrained optimization problems is presented. It belongs to the class of trust region method, in which the descent direction is sought by using the trust region steps within the restricted subspace. Because this subspace can be specified to include information about previous steps, the method is also related to a supermemory descent method without performing multiple dimensional searches. Trust region methods have attractive global convergence property. Supermemory information has good scale independence property. Since the method possesses the characteristics of both the trust region methods and the supermemory descent methods, it is endowed with rapidly convergence. Numerical tests illustrate this point.

1. Introduction

In unconstrained optimization the basic problem considered is

$$\text{Min } f(x) \tag{1.1}$$

where $f(x) : R^n \to R$ is a real differentiable function. Many algorithms have been proposed for solving (1.1). The supermemory descent method is one of them. Its main idea is to combine a descent direction with the displacements generated by previous iterations for obtaining a new search direction. the typical form of the method is shown by Wolfe and Viazminsky[14]. That is, for the kth iteration, calculate α_k, $\beta_k^{(i)}$, s_k and x_{k+1} from

$$f\left(x_k + \alpha_k p_k + \sum_{i=1}^{m} \beta_k^{(i)} s_{k-1}\right) = \min_{\alpha,\beta^{(i)}} f\left(x_k + \alpha_k p_k + \sum_{i=1}^{m} \beta_k^{(i)} s_{k-1}\right),$$

$$s_k = \alpha_k p_k + \sum_{i=1}^{m} \beta_k^{(i)} s_{k-1}, \tag{1.2}$$

and

$$x_{k+1} = x_k + s_k.$$

* Received January 14, 1991.
1) The Project Supported by National Natural Sciences Foundation of China.

where p_k is a basic search direction and m is the number of memory terms. For a quadratic function with positive definite Hessian matrix, the iteration (1.2) with exact line search has the finite step termination property. Choosing different p_k, we obtain different supermemory descent algorithm: supermemory gradient methods, supermemory quasi-Newton methods, etc. Numerical experience show that it is more rapidly convergent than quasi-Newton methods, in general. the major weakness in this class of methods is the computational labour required to perform the (m+1)-dimentional search at each iteration. In order to overcome this defect, Sun[13] constructed a kind of supermemory descent algorithm that does not require the multiple dimentional search. But the method requires that the objective function possesses fairly strong quadratic properties in the neighbourhood of the iterative points to ensure convergence.

On the other hand, trust region methods is an effective way to overcome the difficulty caused by non-positive definite Hessian matrices in Newton's method. The basic idea is that the step is restricted by region of validity of the Taylor series. Given $x_k \in R^n$, consider the subproblem

$$\begin{cases} \text{Minimize} \quad \varphi_k(s) = f_k + g_k^T s + \frac{1}{2} s^T B_k s \\ \text{Subject to} \quad \| s \|_2 \leq \Delta_k \end{cases} \tag{1.3}$$

where B_k is an approximation to the Hessian matrix $\bigtriangledown^2 f(x)$ at x_k and Δ_k is the trust radius. The iteration consists of solving (1.3), and then comparing the actual reduction of the objective function

$$\text{ared}_k = f_k - f(x_k + s_k) \tag{1.4}$$

to the reduction predicted by the quadratic model

$$\text{pred}_k = f_k - \varphi(s_k). \tag{1.5}$$

If the reduction is satisfactory, then the step can be taken and a large trust region tried. Otherwise the trust region is reduced and the minor iteration is repeated.

The motivation for the idea of this paper is to find a means whereby the potential of a good quasi-Newton algorithm is exploited. The scheme suggested is one in which the descent step is sought by using trust region steps within restricted subspace. Because each subspace can be specified to include information about previous steps, the method is also related to a supermemory descent method but avoids the need for performing multiple dimensional searches. Information of this kind may be useful in providing local geometry information. Trust region methods have attractive global convergence property. Supermemory information has good scale independence property. Since the method possesses the characteristics of both the trust region methods and the supermemory descent methods, it is endowed with rapidly convergence. In Section 2 we specify the restricted trust region method. In Section 3 we discuss a rule for constructing the subspace. In Section 4 the convergence of the method is proved. In Section 5 numerical results are presented.