Journal of Computational Mathematics, Vol.13, No.1, 1995, 20–31.

A NONMONOTONIC TRUST REGION TECHNIQUE FOR NONLINEAR CONSTRAINED OPTIMIZATION* 1)

Zhu De-tong
(*Shanghai Normal University, Shanghai, China*)

Abstract

In this paper, a nonmonotonic trust region method for optimization problems with equality constraints is proposed by introducing a nonsmooth merit function and adopting a correction step. It is proved that all accumulation points of the iterates generated by the proposed algorithm are Kuhn-Tucker points and that the algorithm is q-superlinearly convergent.

1. Introduction

Consider the nonlinear equality constrained optimization problem

$$\min f(x), \text{s.t. } c(x) = 01.1$$

where $f : R^n \to R^1$ and $c : R^n \to R^m$, $m \leq n$. Recently, reduced Hessian methods are proposed to solve this problem. Coleman and Conn[1], and Nocedal and Overton[6] proposed separately similar quasi-Newton methods using approximate reduced Hessian. However, such methods can not ensure global convergence and therefore are available only when the initial starts are good enough.

Two basic approaches, namely the line search and the trust region, have been developed in order to ensure global convergence towards local minima (see [4] and [5] for example). However, most of the methods based on these two approaches enforce a monotonic decrease of a certain merit function at each step, and this can considerably slow the convergence rate of the minimization process, especially in the presence of steep-sided valleys (see [4], [5]). More recently, the nonmonotonic line search technique for unconstrained optimization was proposed by Grippo, Lampariello and Lucidi[5]. Furthermore, the nonmonotonic technique has been developed into the trust region algorithm for unconstrained optimization[4]. The nonmonotonic idea motivates the study on the projected Hessian methods with trust region. In this paper, we describe and analyze improved projected methods with nonmonotonic trust region for problem (1.1),

* Received July 11, 1992.
1) The Project Supported by National Natural Science Foundation of China.

introducing a nondifferentiable penalty function as a merit function and employing a correction step which allows us to overcome phenomena similar to the Maratos effect.

Section 2 of this paper gives the improved projected Hessian method in association with the nonmonotonic trust region in more detail. In Section 3, we state the global convergence properties of the method, while in Section 4 we prove the local convergence rate of the algorithm.

2. Algorithm

We first introduce some standard notations. Let $\|\cdot\|$ be the Euclidean norm on R^n. Let $f : R^n \to R^1$ be twice continuously differentiable, with gradient $g : R^n \to R^1$ and Hessian matrix $\nabla^2 f$. Let $c : R^m \to R^n$ be the vector of twice continuously differentiable constrained function $c_i(x)$, for $i = 1, 2, \cdots, m$; with the gradient of $c_i(x)$ denoted by $a_i(x)$ and the Hessian matrix of $c_i(x)$ denoted by $\nabla^2 c_i(x)$.

In the sequel, we adopt the notations

$$f_k := f(x_k), \quad g_k := g(x_k), \quad H_k := H(x_k).$$

We first state the revised projected reduced Hession algorithms, in which, after a moving vector p_k is determined by using the two-sided projected Hessian technique of Nocedal and Overton [6], a correction step will also be taken to make the performance of the algorithm more satisfactory and to overcome the Maratos effect.

Let $A(x) = \nabla c(x)$ be the $n \times m$ matrix consisting of the column vectors $a_i(x)$, for $i = 1, \cdots, m$. Assume $A(x)$ has full column rank. Make a QR decomposition for $A(x)$:

$$A(x) = [Y(x), Z(x)]\,[\,R\,(x)0] = Y(x)R(x)2.1$$

where $[Y, Z]$ is an orthogonal matrix and $R(x) \in R^{m\times m}$ is a nonsingular upper triangular matrix. The columns of $Y(x) \in R^{n\times m}$ and $Z(x) \in R^{n\times t}$, $t = n - m$, form respectively a normalized basis of the range space $R(A(x))$ of $A(x)$ and the null space $N(A(x)^T)$ of $A(x)^T$, i.e.

$$A(x)^T Z(x) = 0.2.2$$

Clearly,

$$Y(x)^T Z(x) = 0,\ Y(x)^T Y(x) = I_m,\ Z(x)^T Z(x) = I_t,$$
$$Y(x)Y(x)^T + Z(x)Z(x)^T = I_n.2.3$$

Let

$$L(x, \lambda) = f(x) - \lambda^T c(x)2.4$$

be the Lagrangian of problem (1.1), where λ is the solution vector of the least-squares problem, called projective multiplier,

$$\min_{\lambda} \|A(x)\lambda - g(x)\|.$$