Journal of Computational Mathematics, Vol.13, No.3, 1995, 250–258.

GENERALIZED NEWTON'S METHOD FOR LC[1] UNCONSTRAINED OPTIMIZATION *1)

W. Sun
(*Department of Mathematics, Nanjing University, Nanjing, China*)

Abstract

In this paper the generalized Newton's method for LC^1 unconstrained optimization is investigated. This method is an extension of Newton's method for the smooth optimization. Some basic concepts are introduced according to Clarke(1983). We give optimality conditions for this kind of optimization problems. The local and the global convergence with exact line search are established under the condition of semismoothness.

1. Introduction

In this paper we consider the generalized Newton's method for LC^1 unconstrained optimization problems. The LC^1 optimization problem is one type of nonsmooth optimization problems which exists extensively in optimization problems. Many examples of this kind of problems are given in Qi (1994), Hiriart-Urruty etc.(1984), Rockafellar(1987) and Fletcher(1980). Polak etc.(1983) researched some extensions of smooth optimization methods to nonsmooth optimization. In this paper we use the second order information of the objective function and extend Newton's method for smooth optimization to LC^1 optimization problems.

The classical Newton's method for smooth optimization problem is

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k), \tag{1}$$

where $f : R^n \to R$ is twice continuously differentiable. However, if the objective function f is not twice continuously differentiable, the iterative formula (1) cannot be used. Many authors discussed the Newton's method for nonsmooth equations (for example, see Pang(1990) , Qi and Sun (1993), Qi(1993), and Robinson (1988)). Here, in the generalized Newton's method for LC^1 optimization problem we use the generalized Hessian instead of Hessian in (1).

Consider

$$\min\ f(x),\ x \in R^n, \tag{2}$$

* Received December 24, 1993.
1) This Work was Supported by the Australian Research Council.

where $f : D \subset R^n \to R$ is a LC^1 function, D is a open subset of R^n, i.e., f is differentiable and its derivative function is locally Lipshitzian on D.

The generalized Newton's method for LC^1 optimization (2) is defined by

$$x_{k+1} = x_k - V_k^{-1}\nabla f(x_k),\ V_k \in \partial^2 f(x_k). \tag{3}$$

where $\partial^2 f(x_k)$ is a generalized Hessian at x_k.

In this paper we give some basic concepts about LC^1 optimization in Section 2 and optimality conditions for LC^1 optimization problem (2) in Section 3. In Section 4 the local convergence and the global convergence under the exact line search are established.

2. Basic Concepts

Consider

$$\min\ f(x),\ x \in R^n,\ f \in LC^1, \tag{4}$$

i.e., f is differentiable and ∇f is locally Lipschitzian. According to Rademacher's Theorem, ∇f is differentiable almost everywhere. Let $D_{\nabla f}$ be the set of points where ∇f is differentiable. Let $\partial^2 f(x)$ be the generalized Hessian in Clarke's sense. Then

$$\partial^2 f(x) = \text{co}\{\lim_{x_i \in D_{\nabla f}, x_i \to x} \nabla^2 f(x_i)\}, \tag{5}$$

where $\partial^2 f(x)$ is the convex hull of all $n \times n$ matrices obtained as the limit of a sequence of the form $\nabla^2 f(x_i)$, where $x_i \to x$ and $x_i \in D_{\nabla f}$. $\partial^2 f(x)$ is a nonempty convex compact subset of $R^{n\times n}$ according to Clarke (1983). Let

$$\partial_B^2 f(x) = \{\lim_{x_i \in D_{\nabla f}, x_i \to x} \nabla^2 f(x_i)\}, \tag{6}$$

then

$$\partial^2 f(x) = \text{co}\{\partial_B^2 f(x)\}. \tag{7}$$

From the mean value theorem 2.6.5 of Clarke (1983), we have

$$\nabla f(y) - \nabla f(x) \in \text{co}\{\partial^2 f([x,y])(y-x)\},\ x, y \in D \subset R^n. \tag{8}$$

The second order generalized directional derivative of f at x in the direction h for unconstrained optimization is defined as follows.

Definition 1. *The second order generalized directional derivative of f at x in the direction h is defined by*

$$f^{\circ\circ}(x;h) = \limsup_{x' \to x, t\downarrow 0} \frac{f^\circ(x'+th;h) - f^\circ(x';h)}{t}, \tag{9}$$

where $f^\circ(x;h)$ denotes the generalized directional derivative of f at x in the direction h. If $f \in LC^1$, then

$$f^{\circ\circ}(x;h) = \limsup_{x' \to x, t\downarrow 0} \frac{[\nabla f(x'+th) - \nabla f(x')]^T h}{t}. \tag{10}$$

Obviously, we have the following basic properties on $f^{\circ\circ}(x;h)$ as Clarke (1983) proved.