Journal of Computational Mathematics, Vol.20, No.3, 2002, 289–300.

GLOBALLY CONVERGENT INEXACT GENERALIZED NEWTON METHODS WITH DECREASING NORM OF THE GRADIENT*[1)]

Ding-guo Pu
(*Department of Mathematics, Tongji University, Shanghai 200331, China*)

Abstract

In this paper, motivated by the Martinez and Qi methods[1], we propose one type of globally convergent inexact generalized Newton methods to solve unconstrained optimization problems in which the objective functions are not twice differentiable, but have LC gradient. They make the norm of the gradient decreasing. These methods are implementable and globally convergent. We prove that the algorithms have superlinear convergence rates under some mile conditions.

The methods may also be used to solve nonsmooth equations.

Key words: Nonsmooth optimization, Inexact Newton method, Generalized Newton method, Global convergence, Superlinear rate.

1. Introduction

Let $f:\ R^n \to R$ be a nonlinear function and $g = \bigtriangledown f$. The usual Newton method for solving nonlinear equations $g(x) = 0$ involves making a first order approximation at the current trial point x_k:

$$\bigtriangledown^2 f(x_k)s_k + g_k = 0 \tag{1}$$

where $g_k = g(x_k)$. We then solve this equation calling the solution x_{k+1}, namely $x_{k+1} = x_k - \bigtriangledown^2 f(x_k)^{-1} g_k$ and repeat the process until obtaining a solution of the original optimization problem.

Newton method is attractive because it converges rapidly from any sufficiently good initial guess. Indeed, it is often taken as a standard convergent method, since one way of characterizing superlinear convergence is that the step should approach Newton step asymptotically in both magnitude and direction (Dennis and Moré [2]).

If the number of variables is large, or if second derivative information is difficult to compute, Newton method may be prohibitively expensive to use. If f is not twice differentiable, Newton method can not be used. For this reason, special methods have been developed to solve large-scale problems. One method is the inexact Newton method, see [3] and [4]. That is, we do not need to solve the Newton equation (1) accurately and as long as the magnitude of the residual vector for an approximate solution s_k:

$$r_k = \bigtriangledown^2 f(x_k)s_k + g_k \tag{2}$$

is asymptotically smaller comparing with $\|g_k\|$, inexact Newton method can work well. In order to stabilize the behavior of inexact Newton method, Eisenstat and Walker in [5] introduced and analyzed the following class of methods: For $k = 0$ step 1 until convergence do

* Received April 2, 1999; Final revised February 23, 2000.
[1)] This research is supported by Ministry of Education P.R.C. and Asia-Pacific Operations Research Center (APORC).

1. Obtain a s_k such that

$$\frac{\|g_k + \nabla^2 f(x_k) s_k\|}{\|g_k\|} = \frac{\|r_k\|}{\|g_k\|} \leq \eta_k, \tag{3}$$

and

$$\|g(x_k + s_k)\| \leq [1 - \theta(1 - \eta_k)]\|g_k\|. \tag{4}$$

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

On the other hand, Newton method has been extended to non-twice differentiable case. Particularly, in recent years some superlinearly convergent generalized Newton methods for solving nonsmooth equations

$$F(x) = 0 \tag{5}$$

have been developed which are based upon Clarke's generalized Jacobian $\partial F(x)$ (see [6]) or B-differentials $\partial_B F(x)$ which used to replace $\nabla^2 f_k$ in (1) or (2), see e.g. [7], [8], [9] and [1].

Stimulated by the progress in these two aspects, in this paper we propose one type of globally convergent inexact generalized Newton methods to solve unconstrained optimization problems in which the objective functions are not twice differentiable, but have LC gradient. Comparing with the globally convergent method in [1], we construct the methods without using the iteration function and prove their convergence and superlinearly convergence under some weaker conditions.

We assume that function f has LC gradient, *i.e.*, there exists an $L > 0$ such that, for any $x, y \in R^n$,

$$\|g(x) - g(y)\| \leq L\|x - y\|. \tag{6}$$

If $g(x)$ is Lipschitzian then Rademacher's theorem implies that $\nabla g(x) = \nabla^2 f(x)$ exists almost everywhere and we can define

$$\partial_B^2 f(x) = \{\lim_{x_k \to x} \nabla^2 f(x_k)\}, \tag{7}$$

where the limit is taken for the x_k at which $f(x_k)$ is twice differentiable. Similar to Clarke's generalized Jacobian we define

$$\partial^2 f(x) = co\{\partial_B^2 f(x)\}. \tag{8}$$

We have (see [6])

(**R1**) $\partial^2 f(x)$ is nonempty, compact and bounded;

(**R2**) $\partial^2 f(x)$ is upper semicontinuous at x;

(**R3**) If $f(x)$ is uniformly convex at x_* and $g_* = g(x_*) = 0$, then x_* is a locally strict minimum point of $f(x)$.

We shall propose one type of globally convergent inexact generalized Newton methods for functions which have LC gradient. These methods are implementable. We shall prove that under some mild assumptions the algorithms are linearly convergent, and they are superlinearly convergent or even quadratically convergent for uniformly convex functions.

The paper is organized as follows. In Section 2, we propose the inexact generalized Newton algorithm which makes the norms of the gradients decreasing, and prove global convergence and superlinear convergence of the algorithm. We also point out that the method can be also used to solve the nonsmooth equations. In Section 3, we discuss the main assumption of the paper and show that the functions with semismooth gradients or C-differentiable gradients must meet the assumption. Section 4 lists some applications for the algorithm and gives some numerical results.

Throughout this paper the vector norms are Euclidean.