JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 93 (1998) 107–122

Inexact generalized Newton methods for second order C-differentiable optimization

Dingguo Pu [1], Jianzhong Zhang *

Department of Mathematics, City University of Hong Kong, Hong Kong

Received 6 December 1996; received in revised form 9 March 1998

Abstract

In this paper we define second order C-differentiable functions and second order C-differential operators, describe their some properties and propose an inexact generalized Newton method to solve unconstrained optimization problems in which the objective function is not twice differentiable, but second order C-differentiable. We prove that the algorithm is linearly convergent or superlinearly convergent including the case of quadratic convergence depending on various conditions on the objective function and different values for the control parameter in the algorithm. © 1998 Elsevier Science B.V. All rights reserved.

AMS classification: 90C30

Keywords: Inexact generalized Newton method; Second order C-differentiable; Convergence and convergence rate

1. Introduction

Consider the nonlinear programming problem

$$\min f(x), \quad x \in \mathbb{R}^n. \tag{1}$$

If $f(x)$ is twice differentiable, a classical algorithm for finding a solution to problem (1) is the Newton method (see [4]). Given an initial guess x_0, we compute a sequence of steps $\{s_k\}$ and iterates $\{x_k\}$ as follows:

Step 0. Give x_0 and let $k=0$.

* Corresponding author. E-mail: mazhang@cityu.edu.hk.

[1] This research is supported by Hong Kong Research Grant Council (Competitive Earmarked Research Grant CITYU-9040159).

PII S0377-0427(98)00064-8

Step 1. Solve

$$\nabla^2 f(x_k)s_k = -g_k, \tag{2}$$

where $\nabla^2 f(x_k)$ is the Hessian of $f(x)$ at x_k and $g_k = g(x_k) = \nabla f(x_k)$, the gradient of $f(x)$ at x_k.

Step 2. If $g(x_k) = 0$ then stop, otherwise let $x_{k+1} = x_k + s_k$ and $k = k + 1$, goto Step 1.

The Newton method is attractive because it converges rapidly from any sufficiently close initial guess. Indeed, it is often taken as a standard convergent method because one way of characterizing superlinear convergence is that the step should approach the Newton step asymptotically in both magnitude and direction [5].

However, if the number of variables is large, solving the Newton equation may be prohibitively expensive. For this reason, special methods have been developed to solve large-scale problems. One of them is inexact Newton method (see [2, 3]) in which we solve Eq. (2) only approximately. A natural stopping rule would be based on the size of the relative residual $\|r_k\|/\|g_k\|$ (throughout this paper the vector norms are Euclidean), where the residual r_k is given by

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

and s_k is the step actually computed (i.e., the approximate solution to Eq. (2)). Such inexact Newton methods may offer a trade-off between the accuracy with which the Newton equations are solved and the amount of work per iteration. An important question is what level of accuracy is required to preserve the rapid local convergence of Newton's method.

Specifically, we consider the class of inexact Newton methods which compute an approximate solution to the Newton equations in some unspecified manner such that

$$\frac{\|r_k\|}{\|g_k\|} \leqslant \eta_k, \tag{4}$$

where the nonnegative forcing sequence $\{\eta_k\}$ is used to control the level of accuracy. To be precise, an inexact Newton method is any variation of Newton method in which (3) replaces (2) and for a given initial guess x_0, a sequence $\{x_k\}$ of approximations to x_* is generated.

If f is not twice differentiable, Newton method cannot be used. Some generalized Newton methods for solving nonsmooth equations

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

have been developed in recent years (see [7, 9, 11]), which are based upon Clarke's generalized Jacobian $\partial F(x)$ or B-differential $\partial_B F(x)$ as well as semismoothness. [6] proposed a class of inexact generalized Newton methods for nonsmooth equations and proved their convergence.

However, exact calculus rules do not hold for Clarke's generalized Jacobians and B-differentials. For example, $\partial f + \partial g \neq \partial(f+g)$ in general. This causes some trouble in implementing these methods. Qi [10] introduced new tools, C-differential operator and C-differentiability, to ease this difficulty and to extend further the applicable area of generalized Newton methods. Qi gives in [10] the following definition:

Definition 1.1. Suppose that $T : \mathbb{R}^n \to \mathbb{R}^{m\times n}$ is a set-valued operator, i.e., for any $x \in \mathbb{R}^n$, any $V \in T(x)$ is an $m \times n$ matrix. We say that the function $F : \mathbb{R}^n \to R^m$ is C-differentiable at $x \in \mathbb{R}^n$ if

(1) $T(y)$ is nonempty and compact for any y in a neighborhood of x;

(2) T is upper semicontinuous at x;

(3) for any $V \in T(x+d)$,

$$F(x+d) = F(x) + Vd + \mathrm{o}(\|d\|).$$

And call T a C-differential operator of F. If furthermore,

(4) for any $V \in T(x+d)$,

$$F(x+d) = F(x) + Vd + \mathrm{O}(\|d\|^2),$$

we say that F is strongly C-differentiable at x.

In this paper we will extend the above definition to the second order case and propose an inexact generalized Newton method to solve some unconstrained optimization problems (1) in which the objective functions are not twice differentiable, but second order C-differentiable. We will prove that the algorithm is linearly convergent or quadratically convergent under some mild conditions. Comparing with the methods in [6], we discuss the method in this paper for different type of objective functions and its convergence under some different conditions. For example, we do not assume that all η_k are small enough in the method of this paper.

Finally, we propose a globally convergent inexact generalized Newton method for second order C-differentiable optimization problems in which we do not need additional conditions in proving its convergence.

The paper is organized as follows: in Section 2, we first define second order C-differentiable functions and second order C-differential operators, then describe some of their properties. In Section 3, we discuss uniformly C-convex functions. Then, we will propose a generalized Newton method and prove its convergence in Section 4. Finally, in Section 5 we introduce a globally converegent inexact generalized Newton method.

2. Second order C-differentiable functions

In this section we first extend the definition of C-differentiability in [10] for nonsmooth equations to the second order C-differentiable functions and second order C-differential operators for non-twice continuously differentiable functions, and then discuss their properties.

Definition 2.1. A first order differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is said second order C-differentiable at x with a second order C-differential operator (or called C2 operator for short) T if the gradient g of f is C-differentiable at x with T. Furthermore, we say that f is second order C-differentiable at x with a C2 operator T and degree ρ, $\rho > 1$, if f is second order C-differentiable at x with T and for any $V \in T(x+d)$,

$$g(x+d) = g(x) + Vd + \mathrm{O}(\|d\|^{\rho}).$$

We say that f is second order C-differentiable in D with a C2 operator T (and degree ρ) if f is second order C-differentiable at each point $x \in D$ with T (and degree ρ).

Remark 1. The second order C-differential operator is defined as a set-valued operator in a neighborhood of x, not only at a single point, even when we discuss the second order C-differentiability at x.

Remark 2. We define the C-differentiability of f without using local Lipschitz and directional differentiability assumptions.

Remark 3. If f is second order C-differentiable then it may have various C2-differential operators, e.g., if T is a second order C-differential operator of f then we can define other second order C-differential operators $\bar{T}$ by one of the following ways:

1. let D be any finite set in $\mathbb{R}^n$ and define $\bar{T}(x) \supset T(x)$, $\bar{T}(x)$ can be any bounded set for $x \in D$ and $\bar{T}(x) = T(x)$ for $x \notin D$;
2. $\bar{T}(x) = co\{T(x)\}$ for all x;
3. $\bar{T}(x) = \{\lim_{x_k \to x} T(x_k)\}$.

Remark 4. If T_1 and T_2 are two second order C-differential operators of f and $\bar{T} = T_1 \cup T_2$, then $\bar{T}$ is also a C2-operator of f.

The concept of second order C-differentiability has some relations with the requirement for the gradient g to be semismooth or Lipschitz continuous. Semismoothness was originally introduced by Mifflin in 1977. Convex functions, smooth functions and subsmooth functions are examples of semismooth functions. For a locally Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}^n$ we say that F is semismooth at x if

$$\lim_{\substack{V \in \partial F(x+th') \\ h' \to h, t \downarrow 0}} \{Vh'\}$$

exists for any $h \in \mathbb{R}^n$, where $\partial F(x+th')$ is the generalized Jacobian of F at $x + th'$ in the sense of [1]. When F is differentiable at x_k and $x_k \to x$, we know that

$$\partial F(x) = co\left\{\lim_{x_k \to x} \nabla F(x_k)\right\}.$$

If F is semismooth at x then for any $V \in \partial F(x+h)$, when $h \to 0$, $F(x+h) - F(x) = Vh + o(\|h\|)$. Define F as strongly semismooth at x if for any $V \in \partial F(x+h)$, when $h \to 0$, $F(x+h) - F(x) = Vh + O(\|h\|^2)$ (see [9]).

If the gradient function g is Lipschitzian then we know that $\nabla g_k = \nabla^2 f(x_k)$ exists almost everywhere, and we can define

$$\partial_B^2 f(x) = \left\{\lim_{x_k \to x} \nabla^2 f(x_k)\right\} \tag{6}$$

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

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

Using the above results we obtain the following proposition immediately.

Proposition 2.2. *Suppose the gradient function g of f is locally Lipschitzian.*

If g is a semismooth function, then f is second order C-differentiable with a second order C-differential operator $\partial^2 f$ (or $\partial_B^2 f$).

If g is strongly semismooth, then f is second order C-differentiable with degree 2 and the associated second order C-differential operator is $\partial^2 f$ (or $\partial_B^2 f$).

However, it remains unknown that if f is second order C-differentiable and g is directionally differentiable, whether g must be semismooth.

We now show other properties about the second order C-differentiable functions.

Proposition 2.3. *If g is C-differentiable at x with a C2 operator T, then g is Hölder-continuous with exponent 1 at x, i.e., there exists a $\delta > 0$, a neighborhood $B_\delta(x) = \{z \mid \|x - z\| \leqslant \delta\}$ of x and a constant $L > 0$ such that for any $y \in B_\delta(x)$,*

$$\|g(y) - g(x)\| \leqslant L\|y - x\|. \tag{8}$$

Proof. If this proposition is not true, then there exists a sequence of points $\{x_k\}$, $x_k \to x$ and

$$\|x_k - x\| = o(\|g_k - g(x)\|). \tag{9}$$

On the other hand Definition 2.1 implies

$$g_k - g(x) = V_k(x_k - x) + o(\|x_k - x\|), \tag{10}$$

where $V_k \in T(x_k)$. Condition (1) of Definition 2.1 implies that $\|V_k\|$ is bounded and therefore, there exists a C_0 such that for sufficiently large k,

$$\|g_k - g(x)\| \leqslant C_0\|x_k - x\|. \tag{11}$$

The contradiction between (9) and (11) proves this proposition. □

Proposition 2.4. *If g is C-differentiable in domain D with a C2 operator T, then g is locally Lipschitzian at any $x \in int\, D$.*

Proof. For any $x \in int\, D$, the local boundedness of T implies that there exists a closed neighborhood $B_\delta(x)$ of x, $x \in int\, B_\delta(x) \subset D$, and a constant $L > 0$ such that for any $y \in int\, B_\delta(x)$ and any $V_y \in T(y)$, $L > 2\|V_y\|$. As f is second order C-differentiable in D, for any $y \in B_\delta(x)$ and any $V_{y+d} \in T(y + d)$,

$$g(y + d) - g(y) = V_{y+d}d + o(\|d\|). \tag{12}$$

For each $y \in int\, B_\delta(x)$ there exists an open neighborhood $B_{\delta(y)}(y)$ of y such that when $y+d \in B_{\delta(y)}(y) \subset B_\delta(x)$ the following inequality holds.

$$\|g(y+d)-g(y)\| \leqslant L\|d\|. \tag{13}$$

Now for any z and $z+d$ in $int\, B_\delta(x)$, $[z, z+d]$ is a closed set and

$$[z, z+d] \subset \bigcup_{y \in [z, z+d]} B_{\delta(y)}(y). \tag{14}$$

By real analysis, there exist finitely many $B_{\delta(y_i)}(y_i)$, $y_i \in [z, z+d]$ such that

$$[z, z+d] \subset \bigcup_i B_{\delta(y_i)}(y_i). \tag{15}$$

In fact we can find out p points z_i, $z_i = z + t_i d$, $0 = t_0 < t_1 < t_2 < \cdots < t_p = 1$, each pair of successive points z_i and z_{i+1} are in the same $B_{\delta(y_i)}(y_i)$ and $y_i \in [z_i, z_{i+1}]$. Hence by (13),

$$\begin{aligned}\|g(z+d)-g(z)\| &\leqslant \sum_{i=0}^{p-1} [\|g(z_{i+1})-g(y_i)\| + \|g(y_i)-g(z_i)\|] \\ &\leqslant \sum_{i=0}^{p-1} L(t_{i+1}-t_i)\|d\| = L\|d\|,\end{aligned}$$

i.e., the proposition is true. □

The semicontinuity of $T(x)$ implies the following proposition (see [10])

Proposition 2.5. *Assume that g is C-differentiable at x_* with a C2 operator T. If $\|V_*^{-1}\|$, where $V_* \in T(x_*)$, is bounded, then there exists a $\delta > 0$ such that for all $x \in B_\delta(x_*)$ and any $V_x \in T$, $\|V_x^{-1}\|$ is bounded.*

We can further define the concept of uniformly second order C-differentiable.

Definition 2.6. Assume that f is second order C-differentiable at x with a C2 operator T. f is called uniformly second order C-differentiable at x with a C2 operator T if for any given $\varepsilon > 0$, there exists a neighborhood $B_\delta(x)$ of x such that for any y satisfying $y+d \in B_\delta(x)$, $(y+d \neq x)$ and any $V_{y+d} \in T(y+d)$ the following result holds.

$$\|g(y+d)-g(y)-V_{y+d}d\| \leqslant \varepsilon\|d\|. \tag{16}$$

3. C2-differentiable convex functions

In this section, we discuss second order C-differentiable convex function.

Definition 3.1. We say that f is a second order C-differentiable convex function, or called a C2-convex function for short, at x with a C2 operator T (and degree ρ) if f is second order C-differentiable at x with a C2 operator T (and degree ρ) and any $V \in T$ is positive semidefinite; f is said to be a second order uniformly C-differentiable convex function, or called uniformly C2-convex for short, at x (or in D) with a C2 operator T (and degree ρ), if f is second order C-differentiable at x (or in D) with a C2 operator T (and degree ρ) and all $V_x \in T(x)$ (and any $x \in D$) are uniformly positive definite.

Proposition 3.2. *If f is uniformly C2-convex at x_* with a C2 operator T, then there exist $\varepsilon > 0$ and a neighborhood $B_\delta(x_*)$ of x_* such that for any $x \in B_\delta(x_*)$ and any $V \in T$, V is uniformly positive definite and*

$$y^{\mathrm{T}} V y \geqslant \varepsilon \|y\|^2, \quad \forall y \in \mathbb{R}^n. \tag{17}$$

Proof. Suppose the proposition is not true, then there exists a sequence of points $\{x_k\}$ satisfying

1. all x_k are in a neighborhood $B_\delta(x_*)$ of x_* which meets the assumption (1) of Definition 2.1 and $x_k \to x_*$;
2. each $T(x_k)$ includes a V_k whose smallest eigenvalue $\lambda_1(V_k)$ is less than ε_k, where $\varepsilon_k > 0$ and $\varepsilon_k \to 0$.

Because $\{V_k\}$ is bounded it has a convergent subsequence. For simplicity assume that $\{V_k\}$ converges to a matrix V. V is not positive definite because for each k, $\lambda_1(V_k) < \varepsilon_k$. But the upper semicontinuity of T implies that $V \in T(x_*)$ and hence V is positive definite. This contradiction implies that the proposition is true. □

Proposition 3.3. *If f is uniformly C2-convex at x_*, then f is strictly convex at x_*.*

Proof. Proposition 3.2 implies that if f is uniformly C2-convex at x_* with a C2 operator T, then there are an $\varepsilon > 0$ and a $\delta > 0$ such that for any $x \in B_\delta(x_*)$, any $V_x \in T(x)$ and any $z \in \mathbb{R}^n$

$$z^{\mathrm{T}} V_x z \geqslant \varepsilon \|z\|^2. \tag{18}$$

Because f is second order C-differentiable at x_*, we know that there exists a δ_1, $0 < \delta_1 \leqslant \delta$ such that for any $x \in B_{\delta_1}(x_*)$ and any $V_x \in T(x)$,

$$(x - x_*)^{\mathrm{T}}(g(x) - g(x_*) - V_x(x - x_*)) \geqslant -\tfrac{1}{2}\varepsilon \|x - x_*\|^2. \tag{19}$$

So, we obtain that for any $x \in B_{\delta_1}(x_*)$ and any $V_x \in T(x)$,

$$\begin{aligned} &f(x) - f(x_*) - (x - x_*)^{\mathrm{T}} g(x_*) \\ &= \int_0^1 ((x - x_*)^{\mathrm{T}} g(x_* + t(x - x_*) - g(x_*)))\,\mathrm{d}t \\ &\geqslant \int_0^1 (t(x - x_*)^{\mathrm{T}} V_{x_* + t(x - x_*)}(x - x_*))\,\mathrm{d}t - \varepsilon \|x - x_*\|^2/2 \int_0^1 t\,\mathrm{d}t \\ &\geqslant \varepsilon \|x - x_*\|^2/4 > 0. \end{aligned} \tag{20}$$

This completes the proof of the proposition. □

Propositions 3.2 and 3.3 immediately imply the following results.

Proposition 3.4. *If f is uniformly C2-convex at x_* with a C2 operator T and $g(x_*)=0$, then x_* is a locally strict minimum point of f.*

Proposition 3.5. *Assume that $f(x)$ is uniformly C2-convex in D with a C2 operator T, $x_* \in D$ and $g(x_*)=0$. If x_0 is sufficiently close to x_*, then there exists a neighborhood $B_\delta(x_*)$ of x_* such that the level set*

$$S(x)=\{x \mid f(x) \leqslant f(x_0), \text{ and } x \in D\} \tag{21}$$

belongs to $B_\delta(x_*)$.

4. An algorithm and its convergence properties

In this section we assume that function f is second order C-differentiable at x with a C2 operator T.

Algorithm 4.1 (*Inexact generalized Newton algorithm*). Given a nonnegative sequence $\{\eta_k\}$ and an initial guess x_0, we compute a sequence of steps $\{s_k\}$ and iterates $\{x_k\}$ as follows:

Step 0. Give x_0 and let $k=0$.

Step 1. Solve

$$V_k s_k = -g_k + r_k, \tag{22}$$

where $V_k \in T(x_k)$ and

$$\frac{\|r_k\|}{\|g_k\|} \leqslant \eta_k. \tag{23}$$

Step 2. If $g(x_k)=0$ then stop, otherwise let $x_{k+1}=x_k+s_k$ and $k=k+1$, goto Step 1.

Lemma 4.2. *Assume that f is second order C-differentiable at x_* with a C2 operator T and $g(x_*)=0$. Let Algorithm 4.1 be implemented with $\eta_k \leqslant \eta_{\max} < t < 1$. If x_k is sufficiently close to x_*, then the following inequality holds:*

$$\|V_k(x_{k+1}-x_*)\| \leqslant t \|V_k(x_k - x_*)\|. \tag{24}$$

Proof. The definition of C-differentiable and $g(x_*)=0$ imply that if x_k is sufficiently close to x_*, then

$$\|V_k(x_k - x_*) - g_k\| \leqslant (t-\eta_{\max})\|x_k - x_*\|/(4M),$$

and hence we have

$$\begin{aligned}\|V_k(x_{k+1}-x_*)\| &= \|V_k(x_k - x_*) + V_k s_k\| \\ &\leqslant \|V_k(x_k - x_*) - g_k\| + \|r_k\|\end{aligned}$$

$$\leqslant \|V_k(x_k - x_*) - g_k\| + \eta_k\|g_k\|$$

$$\leqslant (1+\eta_{\max})\|V_k(x_k - x_*) - g_k\| + \eta_{\max}\|V_k(x_k - x_*)\|$$

$$\leqslant (t-\eta_{\max})\|x_k - x_*\|/(2M) + \eta_{\max}\|V_k(x_k - x_*)\|$$

$$\leqslant (t+\eta_{\max})\|V_k(x_k - x_*)\|/2 \tag{25}$$

$$\leqslant t\|V_k(x_k - x_*)\|, \tag{26}$$

i.e., Lemma 4.1 holds. □

Lemma 4.3. *Assume that f is second order C-differentiable at x_* with a C2 operator T and $x_k \to x_*$. Let $\{V_k\}$ be a sequence of matrices satisfying $V_k \in T(x_k)$ and let S be the set of accumulation points of the sequence $\{V_k\}$, Then, for any given $\varepsilon > 0$, there exists a K such that for any $k > K$, we can find a $V_{k*} \in S$ satisfying*

$$\|V_k - V_{k*}\| < \varepsilon. \tag{27}$$

Proof. If this lemma is not true, then there exist an $\varepsilon > 0$ and a sequence $\{V_{k(i)}\}$, where $V_{k(i)} \in T(x_{k(i)})$ such that

$$\|V_{k(i)} - V_*\| \geqslant \varepsilon \quad \text{for all } V_* \in S. \tag{28}$$

Let V_* be an accumulation point of the sequence $\{V_{k(i)}\}$, then for some large i, $\|V_{k(i)} - V_*\| < \varepsilon$ which contradicts (28). □

Theorem 4.4. *Assume that f is second order C-differentiable at x_* with a C2 operator T, $g(x_*) = 0$ and Algorithm 4.1 is implemented with $\eta_k \leqslant \eta_{\max} < t < 1$ to produce $\{x_k\}$ and the associated matrices used are $\{V_k\}$. Let $M_1 = max\{\|V_{*1} - V_{*2}\| \mid V_{*1}, V_{*2} \in S\}$ (S is given in last lemma) and $\|V_k^{-1}\| \leqslant M$ ($M > 1$ is a constant) for all k. If $\eta_k^{-1} \geqslant (1 + M_1 M)/\eta_{\max}$ for sufficiently large k and x_0 is in a sufficiently small neighborhood $B_\delta(x_*)$, then the sequence $\{x_k\}$ converges to x_*. Moreover, the convergence is linear in the sense that for sufficiently large k,*

$$\|V_k(x_{k+1} - x_*)\| \leqslant t\|V_{k-1}(x_k - x_*)\|. \tag{29}$$

Remark. If S consists of only one V_*, i.e. $V_k \to V_*$, then $M_1 = 0$ and hence no extra condition on η_k is required.

Proof. Because f is second order C-differentiable at x_* with a C2 operator T, there exist a $\delta_1 > 0$ and a constant $M_0 > 1$ such that for any $x \in B_{\delta_1} = B_{\delta_1}(x_*)$ and any $V_x \in T(x)$, $\|V_x\| < M_0$ and

$$\|g(x) - V_x(x - x_*)\| \leqslant (t - \eta_{\max})\|x - x_*\|/(4MM_0). \tag{30}$$

We choose δ_1 small enough such that if $x_k \in B_{\delta_1}$ then Lemma 4.1 holds, and hence

$$\|x_{k+1} - x_*\| \leqslant t\|V_k\|\|V_k^{-1}\|\|x_k - x_*\| \leqslant tM_0M\|x_k - x_*\|. \tag{31}$$

The conditions of this theorem and Lemma 4.2 ensure that there exists a K such that for all $k>K$, $\eta_k^{-1} \geqslant (1+M_1M)/\eta_{\max}$, and also there would be a $V_{k*} \in S$ satisfying

$$\|V_k - V_{k*}\| \leqslant (t-\eta_{\max})/(4MM_0). \tag{32}$$

Let $\delta = \delta_1/(M^{K+1}M_0^{K+1})$ and choose $x_0 \in B_\delta(x_*)$, then starting with $k=0$, we can use the induction method to prove from $x_k \in B_{\delta_1}$ that for any $k \leqslant K$,

$$\|x_{k+1} - x_*\| \leqslant tM_0M\|x_k - x_*\| \leqslant t^kM_0^kM^k\|x_0 - x_*\| < t^k\delta_1/(MM_0), \tag{33}$$

from which we know that

$$x_{k+1} \in B_{\delta_1} \quad \text{and} \quad \|V_k(x_{k+1} - x_*)\| \leqslant \delta_1/M, \quad \text{for } k \leqslant K. \tag{34}$$

We now consider the case $k>K$. If $x_k \in B_{\delta_1}$, then by (30), (32) and the definition of M_1, we can obtain

$$\begin{aligned}
\|V_k(x_{k+1} - x_*)\| &= \|V_k(x_k - x_* + s_k)\| \\
&\leqslant \|V_k(x_k - x_*) - g_k\| + \|V_ks_k + g_k\| \\
&\leqslant \|V_k(x_k - x_*) - g_k\| + \eta_k\|g_k\| \\
&\leqslant (1+\eta_k)\|V_k(x_k - x_*) - g_k\| + \eta_k\|V_k(x_k - x_*)\| \\
&\leqslant (t-\eta_{\max})\|x_k - x_*\|/(2M) + \eta_k[(\|V_k - V_{k*}\| \\
&\quad + \|V_{k*} - V_{(k-1)*}\| + \|V_{(k-1)*} - V_{k-1}\|)\|x_k - x_*\| + \|V_{k-1}(x_k - x_*)\|] \\
&\leqslant (t-\eta_{\max})\|V_{k-1}(x_k - x_*)\| + \eta_kM_1\|x_k - x_*\| + \eta_k\|V_{k-1}(x_k - x_*)\| \\
&\leqslant (t-\eta_{\max})\|V_{k-1}(x_k - x_*)\| + \eta_kM_1M\|V_{k-1}(x_k - x_*)\| + \eta_k\|V_{k-1}(x_k - x_*)\| \\
&\leqslant t\|V_{k-1}(x_k - x_*)\|.
\end{aligned} \tag{35}$$

Since

$$\|V_k(x_{k+1} - x_*)\| \leqslant \delta_1/M \Rightarrow x_{k+1} \in B_{\delta_1},$$

starting with $k=K$ (see (34)), we can use induction method to obtain from (35) that, for all $k>K$,

$$\|V_{k-1}(x_k - x_*)\| \leqslant t^{(k-K+1)}\|V_K(x_{K+1} - x_*)\|. \tag{36}$$

So, it is assured that if $x_0 \in B_\delta(x_*)$, then all $x_k \in B_{\delta_1}$ and $x_k \to x_*$ by (33) and (36), and also (29) holds for $k>K$. $\square$

Recall that Proposition 3.4 shows that if $f(x)$ is uniformly C2-convex at x_* and $g(x_*)=0$, then x_* is a locally strict minimum point of $f(x)$. Furthermore, for uniformly C2-convex function, the algorithm has the following convergence property. Note that the assumptions $\eta_k^{-1} \geqslant (1+M_1M)/\eta_{\max}$ and $\|V_k^{-1}\| \leqslant M$ in Theorem 4.4 are removed when f is uniformly C2-convex.

Theorem 4.5. *Suppose f is uniformly second order C-differentiable and uniformly C-convex at x_* with a C2 operator T, $g(x_*)=0$, and the parameter η_k in Algorithm* 4.1 *meets the condition*

$\eta_k \leqslant \eta_{\max} < t < 1$. *If* x_0 *is sufficiently close to* x_*, *then the sequence of inexact generalized Newton iterates* $\{x_k\}$ *converges to* x_*. *Moreover, for sufficiently large* k *the convergence is linear in the sense that*

$$\|V_{k+1}(x_{k+1} - x_*)\| \leqslant [(1+2t)/(2+t)]\ \|V_k(x_k - x_*)\|. \tag{37}$$

Proof. Since f is uniformly C2-convex and uniformly second order C-differential, by Proposition 3.2 and Definition 2.3, there exist $\delta > 0$ and constant $M_1 > 0$ such that for any y, $y+d \in B_\delta(x_*)$ and any $V_{y+d} \in T(y+d)$, we have that

$$\|V_y\| \leqslant M_1 \quad \text{and} \quad \|V_y^{-1}\| \leqslant M_1. \tag{38}$$

and

$$\|g(y+d) - g(y) - V_{y+d}d\| \leqslant \min\{(1-t),\ t(t-\eta_{\max})\}\|d\|/(3M_1). \tag{39}$$

Without loss of generality, we may assume that δ is chosen sufficiently small such that if $x_k \in B_\delta(x_*)$, then (24) holds. Now if $x_k \in B_\delta(x_*)$ then (38) and (39) imply

$$\begin{aligned}\|x_k - x_*\| &\leqslant M_1\|V_k(x_k - x_*)\|\\ &\leqslant M_1[\|V_k(x_k - x_*) - g_k\| + \|g_k\|]\\ &\leqslant (1-t)\|x_k - x_*\| + M_1\|g_k\|,\end{aligned}$$

or $\|x_k - x_*\| \leqslant M_1\|g_k\|/t$. So, if x_k, $x_{k+1} \in B_\delta(x_*)$, then

$$\|s_k\| = \|x_{k+1} - x_* - (x_k - x_*)\| \leqslant M_1(\|g_{k+1}\| + \|g_k\|)/t. \tag{40}$$

If we replace $y, y+d$ and V_{y+d} in (39) by x_{k+1}, x_k and V_k respectively, then

$$\begin{aligned}\|g_{k+1}\| &= \|g_k + V_k s_k - (g_k - g_{k+1} - V_k(-s_k))\|\\ &\leqslant \|V_k s_k + g_k\| + t(t-\eta_{\max})\|s_k\|/(3M_1)\\ &\leqslant \|r_k\| + (t-\eta_{\max})(\|g_k\| + \|g_{k+1}\|)/3\\ &\leqslant \eta_{\max}\|g_k\| + (t-\eta_{\max})(\|g_k\| + \|g_{k+1}\|)/3\\ &= (t+2\eta_{\max})\|g_k\|/3 + (t-\eta_{\max})\|g_{k+1}\|/3.\end{aligned}$$

So,

$$\|g_{k+1}\| \leqslant \|g_k\|\ [(t+2\eta_{\max})/3]/[1-(t-\eta_{\max})/3] \leqslant t\|g_k\|, \tag{41}$$

where the last inequality is obtained due to the fact

$$3t - t^2 + t\eta_{\max} = t + 2\eta_{\max} + (2-t)(t-\eta_{\max}) > t + 2\eta_{\max}.$$

Let $y = x_*$ and $y + d = x_k$ or x_{k+1} in (39). Then

$$\begin{aligned}
&(2+t)\|V_{k+1}(x_{k+1}-x_*)\|/3 \\
&\quad = \|V_{k+1}(x_{k+1}-x_*)\| - (1-t)\|V_{k+1}(x_{k+1}-x_*)\|/3 \\
&\quad \leqslant \|V_{k+1}(x_{k+1}-x_*)\| - (1-t)\|x_{k+1}-x_*\|/(3M_1) \\
&\quad \leqslant \|g_{k+1}\| \leqslant t\|g_k\| \\
&\quad \leqslant t\|V_k(x_k-x_*)\| + (1-t)\|x_k-x_*\|/(3M_1) \\
&\quad \leqslant (1+2t)\|V_k(x_k-x_*)\|/3.
\end{aligned}$$

To summarize, we have proved that if $x_k, x_{k+1} \in B_\delta$, then

$$\|V_{k+1}(x_{k+1}-x_*)\| \leqslant \tilde{t}\|V_k(x_k-x_*)\|, \tag{42}$$

where $\tilde{t} = (1+2t)/(2+t)$. Clearly, $0 < \tilde{t} < 1$. Now if $\|x_0 - x_*\| \leqslant \delta/M_1^4$, we know from (24) that $\|x_1 - x_*\| \leqslant \delta/M_1^2$ and $\|x_2 - x_*\| \leqslant \delta$, which ensures that $\|V_2\| \leqslant M_1$ and $\|V_2^{-1}\| \leqslant M_1$. Then we can obtain $\|x_2 - x_*\| \leqslant \delta/M_1^2$ by using (42). In this way we can use induction method to obtain that all $\|x_k - x_*\| \leqslant \delta/M_1^2$ by (24) and (42). Moreover, the following inequality holds

$$\|V_k(x_k-x_*)\| \leqslant \tilde{t}^k\|V_0(x_0-x_*)\|. \tag{43}$$

It is clear that $x_k \to x_*$ and (37) holds. □

Theorem 4.6. *Suppose that f is second order C-differentiable and uniformly C2-convex at x_* with degree ρ and a C2 operator T, and the parameter η_k in Algorithm 4.1 meets the condition $\eta_k \leqslant \eta_{\max} < t < 1$. If $g(x_*) = 0$, x_0 is sufficiently close to x_*, and choose $\eta_k \leqslant M_2\|g_k\|^{\rho-1}$, $\rho > 1$, M_2 a constant, then the sequence of inexact generalized Newton iterates $\{x_k\}$ converges to x_*. Moreover, the order of convergence is ρ in the sense that*

$$\|x_{k+1} - x_*\| = O(\|x_k - x_*\|^\rho). \tag{44}$$

Proof. Because $f(x)$ is uniformly C2-convex at x_* with degree ρ and $g(x_*) = 0$, Proposition 3.2 implies that there exist $\delta > 0$ and constant $M_1 > 0$ such that for any $x \in B_\delta(x_*)$ and any $V_x \in T(x)$,

$$\|V_x\| \leqslant M_1 \quad \text{and} \quad \|V_x^{-1}\| \leqslant M_1. \tag{45}$$

By the definition of C2-differentiable with degree ρ and Proposition 2.2, we may reduce δ if necessary so that if $x_k \in B_\delta(x_*)$ then (46) and (47) below hold.

$$\|g_k - V_k(x_k - x_*)\| \leqslant M_3\|x_k - x_*\|^\rho \leqslant t\|x_k - x_*\|/(3M_1), \tag{46}$$

where $M_3>0$ is a constant, and

$$\eta_k \leqslant M_2\|g_k\|^{\rho-1} \leqslant L^{\rho-1}M_2\|x_k-x_*\|^{\rho-1} \leqslant t/(3M_1^2). \tag{47}$$

Now for such $x_k \in B_\delta(x_*)$ and any $V_k \in T(x_k)$, by using (23), (47) and (46),

$$\begin{aligned}\|r_k\| &\leqslant \eta_k\|g_k\| \\ &\leqslant \eta_k(\|V_k(x_k-x_*)\| + \|g_k - V_k(x_k-x_*)\|) \\ &\leqslant t\|V_k\|\|x_k-x_*\|/(3M_1^2) + t\|x_k-x_*\|/(3M_1) \\ &\leqslant 2t\|x_k-x_*\|/(3M_1).\end{aligned} \tag{48}$$

By (22), (46) and (48) we obtain

$$\begin{aligned}\|V_k(x_{k+1}-x_*)\| &= \|V_k(x_k-x_*)+V_ks_k\| \\ &\leqslant \|V_k(x_k-x_*)-g_k\| + \|r_k\| \leqslant t\|x_k-x_*\|/M_1,\end{aligned} \tag{49}$$

and hence

$$\begin{aligned}\|x_{k+1}-x_*\| &\leqslant \|V_k^{-1}\|\ \|V_k(x_{k+1}-x_*)\| \\ &\leqslant t\|V_k^{-1}\|\ \|x_k-x_*\|/M_1 \leqslant t\|x_k-x_*\|.\end{aligned} \tag{50}$$

(50) implies that if $\|x_k-x_*\|\leqslant\delta$ then $\|x_{k+1}-x_*\|\leqslant t\delta$. If x_0 is chosen in $B_\delta(x_*)$ then for all k, $x_k \in B_\delta(x_*)$ and $\|x_k-x_*\|\leqslant t^k\|x_0-x_*\|$. So $x_k-x_*\to 0$. Moreover, we have

$$\begin{aligned}\|V_k(x_{k+1}-x_*)\| &\leqslant \|V_k(x_k-x_*)-g_k\| + \|r_k\| \\ &\leqslant \|V_k(x_k-x_*)-g_k\| + \eta_k\|g_k\| \\ &\leqslant (1+\eta_k)\|V_k(x_k-x_*)-g_k\| + \eta_k\|V_k(x_k-x_*)\| \\ &= \mathrm{O}(\|x_k-x_*\|^\rho),\end{aligned}$$

where the last step is obtained from (46) and (47). This completes the proof of the theorem. □

Notice that the requirement for η_k in Theorem 4.6 is implementable when we use the algorithm. In fact we can take $\eta_k=\min\{\eta_{\max}, M_2\|g_k\|^{\rho-1}\}$.

Corollary 4.7. *If the assumptions in Theorem* 4.6 *hold and* $\rho=2$, *then Algorithm* 4.1 *is quadratically convergent.*

Notice that if in Theorems 4.5 and 4.6 we assume $\{V_k^{-1}\}$ is bounded, then the assumption of uniform C2-convexity can be removed.

5. A globally convergent method

In Section 4, we have given an inexact generalized Newton method and discussed its convergence. How can we find a point x_1 which is sufficiently close to x_* with $g_* = 0$? We may use some stabilization techniques, such as the methods with decreasing gradient norms or decreasing function values, to obtain globally convergent methods. In this section we propose such a method briefly. For a more detailed discussion, see [8].

Algorithm 5.1. Give x_0, $0<\theta_1<1$, $0<\theta_2<1$, $0<\theta_3\leqslant\theta_4<1$ and let $k=0$.

Step 1. Solve

$$V_k s_k = -g_k + r_k, \tag{51}$$

where $V_k \in T(x_k)$ and

$$\frac{\|r_k\|}{\|g_k\|} \leqslant \eta_k. \tag{52}$$

Step 2. If $f(x_k) - f(x_k + s_k) \geqslant \theta_1 \min\{1, \|g_k\|\}\|g_k\|\|s_k\|$, then let $x_{k+1} = x_k + s_k$, goto Step 5, otherwise goto Step 3.

Step 3. If $-g_k^{\mathrm{T}} s_k \geqslant \theta_2 \min\{1, \|g_k\|\}\|s_k\|\|g_k\|$, then let $\beta_k = 0$, otherwise let

$$\beta_k = -g_k^{\mathrm{T}} s_k/\|g_k\|^2 - \theta_2 \min\{1, \|g_k\|\}\|s_k\|/\|g_k\|. \tag{53}$$

Let $d_k = s_k + \beta_k g_k$.

Step 4. Choose α_k so that

$$f(x_k) - f(x_k + \alpha_k d_k) > -\theta_3 \alpha_k g_k^{\mathrm{T}} d_k, \tag{54}$$

$$|d_k^{\mathrm{T}} g(x_k + \alpha_k d_k)| < -\theta_4 g_k^{\mathrm{T}} d_k. \tag{55}$$

Let $x_{k+1} = x_k + \alpha_k d_k$.

Step 5. If $g_{k+1} = 0$ then stop, otherwise let $k = k+1$, goto Step 1.

Theorem 5.2. *Assume that f is second order C-differentiable at x_* with a C2 operator T and is bounded below. If Algorithm 5.1 is implemented with $\eta_k \leqslant \eta_{\max} < t < 1$, x_* is an accumulation point of the sequence $\{x_k\}$ obtained by the algorithm, then $g(x_*) = 0$.*

Proof. Because f is second order C-differentiable at x_* with a C2 operator T, there exist a $\delta_1 > 0$ and a constant $M > 1$ such that $\|V_x\| < M$ for all $x \in B_{\delta_1}(x_*)$ and $V_x \in T(x)$.

As $f(x_k)$ is bounded below and monotonically decreasing as k increases, we have $f(x_k) - f(x_{k+1}) \geqslant 0$ for all k and $f(x_k) - f(x_{k+1}) \to 0$.

If Algorithm 5.1 is implemented with $\eta_k \leqslant \eta_{\max} < t < 1$, then we have

$$\|V_k s_k\| \geqslant \|g_k\| - \|r_k\| \geq (1 - \eta_{\max})\|g_k\|. \tag{56}$$

(56) implies that if $x_k \in B_{\delta_1}(x_*)$, then

$$\|s_k\| \geqslant M^{-1}(1 - \eta_{\max})\|g_k\|. \tag{57}$$

If $g(x_*) \neq 0$ then there exist a $\delta \leqslant \delta_1$ and an $\varepsilon > 0$ such that $x_k \in B_\delta(x_*)$ and $\|g_k\| \geqslant \varepsilon > 0$ for infinitely many k because x_* is an accumulation point of the sequence $\{x_k\}$. Now for these k, consider the following two possible cases:

Case 1. If Step 3 and Step 4 are not implemented, then

$$\begin{aligned} f(x_k) - f(x_k + s_k) &\geqslant \theta_1 \min\{1, \|g_k\|\}\|g_k\|\|s_k\| \\ &\geqslant \theta_1 M^{-1}(1 - \eta_{\max}) \min\{1, \varepsilon\}\varepsilon^2. \end{aligned} \tag{58}$$

Case 2. If Steps 3 and 4 are implemented, then it is clear that, whether $\beta_k = 0$ or not, $-g_k^{\mathrm{T}} d_k \geqslant \theta_2 \min\{1, \|g_k\|\}\|g_k\|\|s_k\|$. Eq. (55) implies

$$-(1 - \theta_4) d_k^{\mathrm{T}} g_k \leqslant -d_k^{\mathrm{T}}(g_k - g_{k+1}) \leqslant \alpha_k L \|d_k\|^2. \tag{59}$$

On the other hand, (53) implies that $\beta_k \|g_k\| \leqslant 2\|s_k\|$ and $\|d_k\| \leqslant 3\|s_k\|$. So, we obtain

$$\begin{aligned} \alpha_k \|d_k\| &\geqslant -(1 - \theta_4) d_k^{\mathrm{T}} g_k / (L\|d_k\|) \\ &\geqslant (1 - \theta_4)\theta_2 \min\{1, \|g_k\|\}\|g_k\|\|s_k\| / (L\|d_k\|) \\ &\geqslant (1 - \theta_4)\theta_2 \min\{1, \|g_k\|\}\|g_k\| / (3L). \end{aligned} \tag{60}$$

Eqs. (54) and (60) imply

$$\begin{aligned} f(x_k) - f(x_k + \alpha_k d_k) &> -\theta_3 \alpha_k g_k^{\mathrm{T}} d_k \\ &\geqslant \alpha_k \theta_3 \theta_2 \min\{1, \|g_k\|\}\|g_k\|\|s_k\| \\ &\geqslant \alpha_k \theta_3 \theta_2 \min\{1, \|g_k\|\}\|g_k\|\|d_k\|/3 \\ &\geqslant (1 - \theta_4)\theta_3 (\theta_2 \min\{1, \|g_k\|\}\|g_k\|)^2 / (9L) \\ &\geqslant (1 - \theta_4)\theta_3 (\theta_2 \min\{1, \varepsilon\})^2 \varepsilon^2 / (9L). \end{aligned} \tag{61}$$

In both cases $f(x_k) - f(x_{k+1})$ is bigger than a constant, which contradicts $f(x_k) - f(x_{k+1}) \to 0$, as $k \to \infty$. This theorem is true. □

Combining Theorem 4.6 and Theorem 5.2, the following theorem holds clearly.

Theorem 5.3. *Suppose that f is second order C-differentiable and uniformly C2-convex at x_* with degree ρ and a C2 operator T, and the parameter η_k in Algorithm 5.1 meets the condition $\eta_k \leqslant \eta_{\max} < t < 1$. If we choose $\eta_k \leqslant M_2 \|g_k\|^{\rho - 1}$, where $\rho > 1$ and M_2 is a constant, then the sequence $\{x_k\}$, generated by Algorithm 5.1, converges to x_*. Moreover, the order of convergence is ρ in the sense that*

$$\|x_{k+1} - x_*\| = \mathrm{O}(\|x_k - x_*\|^\rho). \tag{62}$$

Similar to the short remark at the end of Section 4, if we assume that $\{V_k^{-1}\}$ is bounded in Theorem 5.3, then the assumption of uniformly C2-convexity can be removed.

References

[1] F.H. Clarke, Optimization and nonsmooth analysis, SIAM, Philadelphia, 1990.
[2] R.S. Dembo, S.C. Eisenstat, T. Steihaug, Inexact Newton methods, SIAM J. Numer. Anal., 14 (1982) 400–408.
[3] R.S. Dembo, T. Steihaug, Truncated-Newton algorithms for large-scale unconstrained optimization, Math. Programming 26 (1983) 190–212.
[4] J.E. Dennis Jr., R.B. Schnabel, Numerical methods for unconstrained optimization and nonlinear equations, Prentice-Hall, Englewood Cliffs, NJ, 1983.
[5] J.E. Dennis Jr., Jorge J. Moré, Quasi-Newton method, motivation and theory, SIAM Rev. 19 (1977) 46–89.
[6] J.M. Martinez, L. Qi, Inexact Newton methods for solving nonsmooth equations, J. Comput. Appl. Math. 60 (1995) 127–145.
[7] J.S. Pang, L. Qi, Nonsmooth equations: motivation and algorithms, SIAM J. Optim. 3 (1993) 443–465.
[8] D. Pu, J. Zhang, Globally convergent inexact generalized Newton methods for non-twice differentiable optimization problems, Research Report MA-97-08, City University of Hong Kong, 1997. Presented at ISMP97, Lausanne, 1997.
[9] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Math. Oper. Res. 18 (1993) 227–244.
[10] L. Qi, C-differential operators, C-differentiability and generalized Newton methods, AMR 96/5, Applied Mathematics Report, University of New South Wales, 1996.
[11] L. Qi, J. Sun, A nonsmooth version of Newton's method, Math. Programming 58 (1993) 353–368.