

NEWTON AND QUASI-NEWTON METHODS FOR A CLASS OF NONSMOOTH EQUATIONS AND RELATED PROBLEMS*

DEFENG SUN[†] AND JIYE HAN[†]

Abstract. The paper presents concrete realizations of quasi-Newton methods for solving several standard problems including complementarity problems, special variational inequality problems, and the Karush–Kuhn–Tucker (KKT) system of nonlinear programming. A new approximation idea is introduced in this paper. The Q -superlinear convergence of the Newton method and the quasi-Newton method are established under suitable assumptions, in which the existence of $F'(x^*)$ is not assumed. The new algorithms only need to solve a linear equation in each step. For complementarity problems, the QR factorization on the quasi-Newton method is discussed.

Key words. nonsmooth equations, Newton method, quasi-Newton method, Q -superlinear convergence

AMS subject classifications. 90C30, 90C33

PII. S1052623494274970

1. Introduction. In recent years, many authors have considered various forms of Newton methods for solving nonsmooth equations (NE) (see, e.g., [17, 18, 19, 20, 11, 12, 13, 21, 22, 23, 26]). Some authors have also considered the application of the quasi-Newton methods to nonsmooth equations. In Kojima and Shindo [11], the quasi-Newton method was applied to piecewise smooth equations. When the iteration sequence moves to a new C^1 -piece, a new approximate starting matrix is needed. Ip and Kyparisis [9] considered the local convergence of quasi-Newton methods directly applied to B-differentiable equations (in the sense of Robinson [25]). The superlinearly convergent theorems are established under the assumption that F is strongly F-differentiable [15] at the solution.

The main object of this paper is to construct a practical quasi-Newton method for nonsmooth equations, especially for those which are of concrete background. In order to complete this, we first give a slight modification of the generalized Newton method [21, 22, 13]. Based on the modified generalized Newton method, we give a quasi-Newton method for solving a class of nonsmooth equations, which arises from the complementarity problem, variational inequality problem, the Karush–Kuhn–Tucker (KKT) system of nonlinear programming, and related problems. In each step, we only need to solve a linear equation. The Q -superlinear convergence is established under mild conditions.

The characteristics of the quasi-Newton method for solving (4.12) established in section 4 include the following: (i) without assuming the existence of $F'(x^*)$, we prove the Q -superlinearly convergent property; (ii) only one approximate starting matrix is needed; and (iii) from the QR factorization of the k th iterate matrix we need at most $O((I(k) + 1)n^2)$ arithmetic operations to get the QR factorization of the $(k + 1)$ th iterate matrix (for the definition of $I(k)$, see (5.8)).

The remainder of this paper is organized as follows. In section 2, we give some preliminaries on nonsmooth functions. In section 3, we propose a modified generalized

* Received by the editors September 30, 1994; accepted for publication (in revised form) November 14, 1995. This work was supported by the National Natural Science Foundation of China under project 19371084.

<http://www.siam.org/journals/siopt/7-2/27497.html>

[†] Institute of Applied Mathematics, Academia Sinica, Beijing 100080, P. R. China (sun@maths.unsw.edu.au, jyhan%amath3@amath3.amt.ac.cn).

Newton method. In section 4, we give a quasi-Newton method for solving a class of nonsmooth equations. In section 5, we discuss the implementation of the quasi-Newton method for the nonlinear complementarity problem. The KKT system of variational inequality problems with upper and lower bounds are discussed in section 6. The computational results are given in section 7.

2. Preliminaries. In general, assume that $F : R^n \rightarrow R^m$ is locally Lipschitzian. In order to reduce the nonsingularity assumption of the generalized Newton method [22], the concept $\partial_B F(x)$ was introduced by Qi [21]:

$$(2.1) \quad \partial_B F(x) = \left\{ \lim_{\substack{x^k \rightarrow x \\ x^k \in D_F}} F'(x^k) \right\},$$

where D_F is the set where F is differentiable. Let ∂F be the generalized Jacobian of F in the sense of Clarke [4]. Then $\partial F(x)$ is the convex hull of $\partial_B F(x)$,

$$(2.2) \quad \partial F(x) = \text{conv } \partial_B F(x).$$

For $m = 1$, $\partial_B F(x)$ was introduced by Shor [28]. Here, we denote

$$(2.3) \quad \partial_b F(x) = \partial_B F_1(x) \times \partial_B F_2(x) \times \cdots \times \partial_B F_m(x).$$

When $m = 1$, $\partial_b F(x) = \partial_B F(x)$.

We say that F is *semismooth* at x if

$$(2.4) \quad \lim_{\substack{V \in \partial F(x+th') \\ h' \rightarrow h, t \downarrow 0}} \{Vh'\}$$

exists for any $h \in R^n$. Semismoothness was originally introduced by Mifflin [14] for functionals. Convex functions, smooth functions, and piecewise linear functions are examples of semismooth functions. Scalar productions and sums of semismooth functions are still semismooth functions (see [14]). In [23], Qi and Sun extended the definition of semismooth functions to $F : R^n \rightarrow R^m$. It was proved in [23] that F is semismooth at x if and only if all its component functions are so.

Condition (2.4) is stronger than the assumption that for any $h \in R^n$,

$$(2.5) \quad \lim_{\substack{V \in \partial F(x+th) \\ t \downarrow 0}} \{Vh\}$$

exists. Under the latter assumption, Qi and Sun [Proposition 2.1, 22] proved that the classical derivative

$$F'(x; h) = \lim_{t \downarrow 0} \frac{F(x + th) - F(x)}{t}$$

exists and is equal to the limit in (2.5); i.e.,

$$(2.6) \quad F'(x; h) = \lim_{\substack{V \in \partial F(x+th) \\ t \downarrow 0}} \{Vh\}.$$

If the right-hand side limit in (2.6) is uniformly convergent for all h with unit norm, then from Theorem 2.3 of [22] we have that F is semismooth at x . In [13], Kummer discussed sufficient and necessary conditions for the convergence of the Newton

method based on generalized derivatives. One of the conditions for guaranteeing convergence (see Theorem 2 of [13]) is (specialized to the fourth case discussed in [13]) that for any $V \in \partial F(x + h)$, $h \rightarrow 0$,

$$(2.7) \quad F(x + h) - F(x) - Vh = o(\|h\|).$$

Since F is locally Lipschitz continuous, from [27] we know that if $F'(x; h)$ exists, then $F'(x; h)$ coincides with the B -derivative of F at x ; i.e.,

$$(2.8) \quad \lim_{h \rightarrow 0} \frac{F(x + h) - F(x) - F'(x; h)}{\|h\|} = 0.$$

So, if $F'(x; h)$ exists, then (2.7) implies that for any $V \in \partial F(x + h)$, $h \rightarrow 0$,

$$(2.9) \quad Vh - F'(x; h) = o(\|h\|).$$

Again, (2.9) implies the semismoothness of F at x from Theorem 2.3 of [22]. But in [13], Kummer also discussed the case that $F'(x; h)$ may not exist. In this paper we will only consider the case that $F'(x; h)$ exists. Under the existence assumption of $F'(x; h)$, similar to the above discussion from Theorem 2.3 of [22], we can prove that in finite dimensional space the condition (CA*) in Theorem 2 of [13] implies (2.9) (by assuming $F(x) = 0$), which is essentially equivalent to the semismoothness of F at x . Semismoothness is a useful tool in proving the Q -superlinear convergence of the generalized Newton method for nonsmooth equations [21, 22, 23]. We also need it in this paper. In addition, Kummer [13] discussed the approximation of Newton matrices and errors when solving the auxiliary problems. In this paper we will put our main attention on constructing concrete quasi-Newton methods for solving special nonsmooth equations and will not discuss the inexact solution of the subproblems.

LEMMA 2.1 (see [22]). *Suppose that $F : R^n \rightarrow R^m$ is a locally Lipschitzian function and semismooth at x . Then*

(1) *for any $V \in \partial F(x + h)$, $h \rightarrow 0$,*

$$Vh - F'(x; h) = o(\|h\|);$$

(2) *for any $h \rightarrow 0$,*

$$F(x + h) - F(x) - F'(x; h) = o(\|h\|).$$

In the rest of this paper, let $\|\cdot\|$ denote the l_2 vector norm or its induced matrix norm.

LEMMA 2.2. *Suppose that $F : R^n \rightarrow R^m$ is a locally Lipschitzian function. If all $V \in \partial_b F(x)$ are nonsingular, then there exists a positive constant β such that*

$$\|V^{-1}\| \leq \beta$$

for any $V \in \partial_b F(x)$. Furthermore, there exists a neighborhood $N(x)$ of x such that for any $y \in N(x)$, all $W \in \partial_b F(y)$ are nonsingular and satisfy

$$(2.10) \quad \|W^{-1}\| \leq \frac{10}{9}\beta.$$

Proof. From the definition of $\partial_b F$ we can easily know that $\partial_b F(\cdot)$ is bounded and closed in a neighborhood of x . Then the proof of the theorem is similar to that of [21, 22]. We omit the detail here. \square

3. Newton method for nonsmooth equations. Suppose that $F : R^n \rightarrow R^n$ is locally Lipschitzian. We are interested in finding a solution of the equation

$$(3.1) \quad F(x) = 0.$$

Qi and Sun [22], Qi [21], and Kummer [13] considered various forms of the Newton method for solving (3.1) when F is not F -differentiable. Here we will consider the following slightly modified Newton method

$$(3.2) \quad x^{k+1} = x^k - V_k^{-1}F(x^k), \quad k = 0, 1, \dots,$$

where $V_k \in \partial_b F(x^k)$. This method is useful to establish the superlinear convergence of quasi-Newton methods given in section 4. Similar to that of [21, 22], we can give the following convergence theorem.

THEOREM 3.1. *Suppose that x^* is a solution of (3.1), F is locally Lipschitzian and semismooth at x^* , and all $V_* \in \partial_b F(x^*)$ are nonsingular. Then the iteration method (3.2) is well defined and converges to x^* Q -superlinearly in a neighborhood of x^* .*

Proof. By Lemma 2.2, (3.2) is well defined in a neighborhood of x^* for the first step $k = 0$. Since $V_k \in \partial_b F(x^k)$, the i th row V_k^i of V_k satisfies

$$V_k^i \in \partial_B F_i(x^k).$$

From the semismoothness of F we know that F_i is semismooth at x^* . By Lemma 2.1,

$$V_k^i(x^k - x^*) - F'_i(x^*; x^k - x^*) = o(\|x^k - x^*\|), \quad i = 1, \dots, n.$$

Therefore,

$$(3.3) \quad V_k(x^k - x^*) - F'(x^*; x^k - x^*) = o(\|x^k - x^*\|).$$

From Lemma 2.1 and (3.3) we have

$$\begin{aligned} \|x^{k+1} - x^*\| &= \|x^k - x^* - V_k^{-1}F(x^k)\| \\ &\leq \|V_k^{-1}[F(x^k) - F(x^*) - F'(x^*; x^k - x^*)]\| \\ &\quad + \|V_k^{-1}[V_k(x^k - x^*) - F'(x^*; x^k - x^*)]\| \\ &= o(\|x^k - x^*\|). \end{aligned} \quad \square$$

From the theoretical point of view, there is no need to allow Newton matrices in $\partial_b F(\cdot)$ only since, due to the semismoothness assumptions, even each matrix of $\text{conv } \partial_b F(\cdot)$ could be used. The latter would lead to more general statements than those in Theorem 3.1. On the other hand, from the computational point of view, the assumption that all matrices $V \in \text{conv } \partial_b F(x)$ are nonsingular is too strong and not necessary. So here we only restrict $V \in \partial_b F(x)$ and will not discuss the more general case that $V \in \text{conv } \partial_b F(x)$. See [20] and section 6 for further discussions on the nonsingularity assumption of $V \in \partial_b F(x)$. For general statements on Newton methods for nonsmooth equations, see Qi and Sun [22] and Kummer [13].

4. Quasi-Newton method for nonsmooth equations and its specializations. In this section, we will first consider a quasi-Newton method for general nonsmooth equations and then discuss its specializations to a class of nonsmooth equations and related problems.

Consider the following quasi-Newton method:

$$(4.1) \quad x^{k+1} = x^k - V_k^{-1}F(x^k), \quad V_k \in R^{n \times n}, \quad k = 0, 1, \dots$$

THEOREM 4.1. *Suppose that $F : R^n \rightarrow R^n$ is a locally Lipschitzian function in the open convex set $D \subset R^n$ and $x^* \in D$ is a solution of $F(x) = 0$. Suppose that F is semismooth at x^* and all $W_* \in \partial_b F(x^*)$ are nonsingular. There exist positive constants ε, Δ such that if $x^0 \in D, \|x^0 - x^*\| \leq \varepsilon$, and there exists $W_k \in \partial_b F(x^k)$ such that*

$$(4.2) \quad \|V_k - W_k\| \leq \Delta,$$

then the sequence of points generated by (4.1) is well defined and converges to x^* Q -linearly in a neighborhood of x^* .

Proof. From Lemma 2.2, there exists a positive constant β such that $\|W_*^{-1}\| \leq \beta$ for all $W_* \in \partial_b F(x^*)$ and there exists a neighborhood $N_0(x^*) (\subseteq D)$ of x^* such that

$$\|W^{-1}\| \leq \frac{10}{9}\beta$$

for any $y \in N_0(x^*), W \in \partial_b F(y)$. Choose $\Delta > 0$ such that

$$(4.3) \quad 6\beta\Delta \leq 1.$$

Recall that a map is semismooth at x^* if and only if each of its components is semismooth at x^* . So from (1) and (2) of Lemma 2.1, for any $W^i \in \partial_b F_i(x), x \rightarrow x^*$,

$$(4.4) \quad \|F_i(x) - F_i(x^*) - W^i(x - x^*)\| = o(\|x - x^*\|).$$

Therefore, for any $W \in \partial_b F(x), x \rightarrow x^*$, we have

$$(4.5) \quad \|F(x) - F(x^*) - W(x - x^*)\| = o(\|x - x^*\|).$$

Then we can choose a positive constant ε small enough such that for any $x \in N(x^*) = \{y \mid \|y - x^*\| \leq \varepsilon\} \subseteq N_0(x^*), W \in \partial_b F(x)$, we have

$$(4.6) \quad \|F(x) - F(x^*) - W(x - x^*)\| \leq \Delta\|x - x^*\|.$$

If $\|x^k - x^*\| \leq \varepsilon$, then $W_k \in \partial_b F(x^k)$ is nonsingular and $\|W_k^{-1}\| \leq \frac{10}{9}\beta$. By Theorem 2.3.2 of Ortega and Rheinboldt [15], V_k is invertible and

$$(4.7) \quad \|V_k^{-1}\| \leq \frac{\|W_k^{-1}\|}{1 - \|W_k^{-1}(W_k - V_k)\|} \leq \frac{\frac{10}{9}\beta}{1 - \frac{5}{27}} < \frac{3}{2}\beta.$$

Then when $\|x^k - x^*\| \leq \varepsilon$, we have

$$(4.8) \quad \begin{aligned} \|x^{k+1} - x^k\| &= \|x^k - V_k^{-1}F(x^k) - x^*\| \\ &\leq \|V_k^{-1}\| \|F(x^k) - F(x^*) - V_k(x^k - x^*)\| \\ &\leq \|V_k^{-1}\| [\|F(x^k) - F(x^*) - W_k(x^k - x^*)\| \\ &\quad + \|V_k - W_k\| \|x^k - x^*\|]. \end{aligned}$$

Substituting (4.2), (4.6), and (4.7) into (4.8) gives

$$\begin{aligned}
 \|x^{k+1} - x^*\| &\leq \frac{3}{2}\beta[\Delta\|x^k - x^*\| + \Delta\|x^k - x^*\|] \\
 &\leq 3\beta\Delta\|x^k - x^*\| \\
 (4.9) \qquad &\leq \frac{1}{2}\|x^k - x^*\|.
 \end{aligned}$$

This shows that the sequence of points generated by (4.1) is well defined and converges to x^* Q -linearly in a neighborhood of x^* . \square

In [20], Pang and Qi extended Theorem 2.2 in Dennis and Moré [5] to nonsmooth equations. Here, we can do a similar extension and point out that some quasi-Newton methods belong to our frame form.

THEOREM 4.2. *Suppose that $F : R^n \rightarrow R^n$ is a locally Lipschitzian function in the open convex set $D \subset R^n$. Assume that F is semismooth at some $x^* \in D$ and all $W_* \in \partial_b F(x^*)$ are nonsingular. Let $\{V_k\}$ be a sequence of nonsingular matrices in $R^{n \times n}$, and suppose for some x^0 in D that the sequence of points generated by (4.1) remains in D and satisfies $x^k \neq x^*$ for all k , and $\lim_{k \rightarrow \infty} x^k = x^*$. Then $\{x^k\}$ converges Q -superlinearly to x^* , and $F(x^*) = 0$ if and only if there exists $W_k \in \partial_b F(x^k)$ such that*

$$(4.10) \qquad \lim_{k \rightarrow \infty} \frac{\|(V_k - W_k)s^k\|}{\|s^k\|} = 0,$$

where $s^k = x^{k+1} - x^k$.

Proof. Write $e^k = x^k - x^*$. Then both sequence $\{e^k\}$ and $\{s^k\}$ converge to zero. From (4.1) we have

$$\begin{aligned}
 F(x^*) &= [F(x^k) + W_k s^k] - [F(x^k) - F(x^*) - W_k e^k] - W_k e^{k+1} \\
 &= [F(x^k) + V_k s^k] + [(V_k - W_k)s^k] \\
 &\quad - [F(x^k) - F(x^*) - W_k e^k] - W_k e^{k+1} \\
 (4.11) \qquad &= [(V_k - W_k)s^k] - [F(x^k) - F(x^*) - W_k e^k] - W_k e^{k+1}.
 \end{aligned}$$

From the semismoothness of F at x^* and (4.5) we know that the term in the second square bracket approaches zero as $k \rightarrow \infty$. So if (4.10) holds, then $H(x^*) = 0$. From Lemma 2.2, $\{\|W_k^{-1}\|\}$ is bounded. Thus, from (4.5), (4.10), (4.11), and the boundedness of $\{\|W_k^{-1}\|\}$, we have

$$\|e^{k+1}\| \leq o(\|s^k\|) + o(\|e^k\|) \leq o(\|e^k\|) + o(\|e^{k+1}\|),$$

which means that

$$\lim_{k \rightarrow \infty} \frac{\|e^{k+1}\|}{\|e^k\|} = 0.$$

Conversely, suppose that $H(x^*) = 0$ and $\{x^k\}$ converges Q -superlinearly to x^* . Then reversing the above discussion easily establishes condition (4.10). \square

As applications to Theorems 4.1 and 4.2, we will first consider the following nonsmooth equations, which arise from complementarity problems, special variational inequality problems, and the KKT system of nonlinear programming:

$$(4.12) \quad F(x) = x - P_X[x - f(x)] = 0,$$

where $f : R^n \rightarrow R^n$ is a continuously differentiable function, $P_Y(\cdot)$ is the orthogonal projection operator onto a nonempty closed convex set Y , and $X = \{x \in R^n \mid l \leq x \leq u\}$, where $l, u \in \{R \cup \{\infty\}\}^n$. To solve equation (4.12) is the original motivation in investigating nonsmooth equations. When $f \in C^1$, F is a semismooth function. The results of the Newton method for solving (4.12) are fruitful, but not for the quasi-Newton method. In this section, we will give a new quasi-Newton method for solving equation (4.12).

QUASI-NEWTON METHOD (BROYDEN'S CASE [1]).

Given $f : R^n \rightarrow R^n$, $x^0 \in R^n$, $A_0 \in R^{n \times n}$

Do for $k = 0, 1, \dots$:

Define

$$f^k(x) = f(x^k) + A_k(x - x^k)$$

$$(4.13) \quad F^k(x) = x - P_X[x - f^k(x)]$$

Choose $V_k \in \partial_b F^k(x^k)$

Solve $V_k s^k + F(x^k) = 0$ for s^k

$$x^{k+1} = x^k + s^k$$

$$y^k = f(x^{k+1}) - f(x^k)$$

$$(4.14) \quad A_{k+1} = A_k + \frac{(y^k - A_k s^k) s^{kT}}{s^{kT} s^k}.$$

For any matrix $B \in R^{n \times n}$, let B^i be the i th row of B . For an arbitrary function $f \in C^1$, if $V \in \partial_b F(x)$, then V satisfies

$$(4.15) \quad V^i = \begin{cases} I^i & \text{if } x_i - f_i(x) < l_i \text{ (or } > u_i), \\ \lambda_i I^i + (1 - \lambda_i) f'_i(x) & \text{if } x_i - f_i(x) = l_i \text{ (or } = u_i), \\ f'_i(x) & \text{if } l_i < x_i - f_i(x) < u_i, \end{cases}$$

where $\lambda_i \in \{0, 1\}$ and I is the unit matrix of $R^{n \times n}$. On the other hand, any V of the above form is an element of $\partial_b F(x)$.

COROLLARY 4.1. *Suppose that $f : R^n \rightarrow R^n$ is continuously differentiable, x^* is a solution of (4.12), $f'(x)$ is Lipschitz continuous in a neighborhood of x^* , and the Lipschitz constant is γ . Suppose that all $W_* \in \partial_b F(x^*)$ are nonsingular. There exist positive constants ε, δ such that if $\|x^0 - x^*\| \leq \varepsilon$ and $\|A_0 - f'(x^*)\| \leq \delta$, then the*

sequence $\{x^k\}$ generated by the quasi-Newton method (Broyden's case) is well defined and converges Q -superlinearly to x^* .

Proof. First we prove the Q -linear convergence of $\{x^k\}$. Choose ε and Δ as in the proof of Theorem 4.1 and restrict ε to be small enough such that for any $y \in N(x^*) = \{x \mid \|x - x^*\| \leq \varepsilon\}$, we have

$$(4.16) \quad \|f'(y) - f'(x^*)\| \leq \gamma \|y - x^*\|,$$

$$(4.17) \quad 3\gamma\varepsilon \leq \Delta.$$

Denote $\delta := \Delta/2$. From the definition of $F^k(x)$ and (4.15), the j th row V_k^j of V_k satisfies

$$(4.18) \quad V_k^j = \begin{cases} I^j & \text{if } x_j^k - f_j^k(x^k) < l_j \text{ (or } > u_j), \\ \lambda_j^k I^j + (1 - \lambda_j^k) A_k^j & \text{if } x_j^k - f_j^k(x^k) = l_j \text{ (or } = u_j), \\ A_k^j & \text{if } l_j < x_j^k - f_j^k(x^k) < u_j, \end{cases}$$

where $\lambda_j^k \in \{0, 1\}$. For such constants λ_j^k we define a companion matrix W_k such that the j th row W_k^j of W_k satisfies

$$(4.19) \quad W_k^j = \begin{cases} I^j & \text{if } x_j^k - f_j^k(x^k) < l_j \text{ (or } > u_j), \\ \lambda_j^k I^j + (1 - \lambda_j^k) f_j'(x^k) & \text{if } x_j^k - f_j^k(x^k) = l_j \text{ (or } = u_j), \\ f_j'(x^k) & \text{if } l_j < x_j^k - f_j^k(x^k) < u_j. \end{cases}$$

From $f(x^k) = f^k(x^k)$ and (4.19) we get

$$W_k \in \partial_b F(x^k).$$

From (4.18) and (4.19) for any $x \in R^n$ we get

$$|(W_k^j - V_k^j)x| \leq |(A_k^j - f_j'(x^k))x|,$$

which means that

$$(4.20) \quad \|(W_k - V_k)x\| \leq \|(A_k - f'(x^k))x\|.$$

Thus,

$$(4.21) \quad \begin{aligned} \|W_k - V_k\| &\leq \|A_k - f'(x^k)\| \\ &\leq \|A_k - f'(x^*)\| + \|f'(x^k) - f'(x^*)\|. \end{aligned}$$

The local Q -linear convergence proof consists of showing by induction that

$$(4.22) \quad \|A_k - f'(x^*)\| \leq (2 - 2^{-k})\delta,$$

$$(4.23) \quad \|V_k - W_k\| \leq \Delta.$$

For $k = 0$, (4.22) is trivially true. The proof of (4.23) is identical to the proof at the induction step, so we omit it here.

Now assume that (4.22) and (4.23) hold for $k = 0, 1, \dots, i - 1$. From the proof of Theorem 4.1, for $k = 0, 1, \dots, i - 1$, we have

$$(4.24) \quad \|e^{k+1}\| \leq \frac{1}{2}\|e^k\|.$$

For $k = i$, we have from Lemma 8.2.1 of [6] (also see [5]), (4.24), and the induction hypothesis that

$$(4.25) \quad \begin{aligned} \|A_i - f'(x^*)\| &\leq \|A_{i-1} - f'(x^*)\| + \frac{\gamma}{2}(\|e^i\| + \|e^{i-1}\|) \\ &\leq (2 - 2^{-(i-1)})\delta + \frac{3\gamma}{4}\|e^{i-1}\|. \end{aligned}$$

From (4.24) and $\|e^0\| \leq \varepsilon$ we get

$$\|e^{i-1}\| \leq 2^{-(i-1)}\|e^0\| \leq 2^{-(i-1)}\varepsilon.$$

Substituting this into (4.25) and using (4.17) gives

$$\begin{aligned} \|A_i - f'(x^*)\| &\leq (2 - 2^{-(i-1)})\delta + \frac{3\gamma}{4}\varepsilon \cdot 2^{-(i-1)} \\ &\leq (2 - 2^{-(i-1)} + 2^{-i})\delta = (2 - 2^{-i})\delta, \end{aligned}$$

which verifies (4.22).

To complete the induction, we verify (4.23). Substituting (4.22) into (4.21) for $k = i$ and using $\|e^0\| \leq \varepsilon$, (4.16), (4.17), and (4.24) gives

$$\begin{aligned} \|W_i - V_i\| &\leq (2 - 2^{-i})\delta + 2^{-i}\varepsilon\gamma \\ &= (2 - 2^{-i})\frac{\Delta}{2} + \frac{1}{3} \cdot 2^{-i}\Delta \\ &< \Delta. \end{aligned}$$

This proves (4.23). So the Q -linear convergence follows from Theorem 4.1.

Next we will prove the Q -superlinear convergence of $\{x^k\}$ under the assumptions. Let $E_k = A_k - f'(x^*)$. From the last part of the proof of Theorem 8.2.2 of [6] (also see [5]) we get

$$(4.26) \quad \lim_{k \rightarrow \infty} \frac{\|E_k s^k\|}{\|s^k\|} = 0.$$

From (4.20) and (4.16), we have

$$(4.27) \quad \begin{aligned} \|(V_k - W_k)s^k\| &\leq \|(A_k - f'(x^k))s^k\| \\ &\leq \|(A_k - f'(x^*))s^k\| + \|(f'(x^k) - f'(x^*))s^k\| \\ &\leq \|E_k s^k\| + \gamma\|e^k\|\|s^k\|. \end{aligned}$$

Substituting (4.26) into (4.27) and using the linear convergence of $\{x^k\}$ gives

$$\lim_{k \rightarrow \infty} \frac{\|(V_k - W_k)s^k\|}{\|s^k\|} = 0,$$

which, from Theorem 4.2, means that $\{x^k\}$ converges to x^* Q -superlinearly. \square

Recall that when X is the nonnegative orthant, i.e., $X = R_+^n$, $F(x)$ defined by (4.12) is essentially equivalent to the function $H(x)$ in [9] and [17]. In [9], Ip and Kyparisis discussed the convergence properties of quasi-Newton methods directly applied to nonsmooth equations. For nonlinear complementarity problems, they described the sufficient conditions to guarantee the convergence of the quasi-Newton method (see Theorem 5.2 of [9]). A restrictive assumption in [9] is that F is strongly F -differentiable at x^* . This condition, which restricts the class f to which Theorem 5.2 of [9] applies, is satisfied if $f'_i(x^*) = I^i$ for all $i \in \{j | f_j(x^*) = x_j^*, j = 1, \dots, n\}$. Here, to guarantee the convergence of our new quasi-Newton method, we need the nonsingularity of $\partial_b F(x^*)$ instead of needing the existence and invertibility of $F'(x^*)$. For nonlinear complementarity problems, the nonsingularity assumption of $\partial_b F(x^*)$ is equivalent to the b -regularity assumption in [19]. For a detailed discussion on b -regularity, see [19].

Next we consider the following nonsmooth equation:

$$(4.28) \quad F(x) = \min(f(x), g(x)) = 0,$$

where $f, g : R^n \rightarrow R^n$ are continuously differentiable and the “min” operator denotes the componentwise minimum of two vectors. Such a system arises from nonsmooth partial differentiable equations [3, 2, 15] and implicit complementarity problems (see, e.g., [16]). When $g(x) = x$, (4.28) is the function $H(x)$ discussed in [9] and [17] and is equivalent to (4.12) for $X = R_+^n$. Here we will give a new quasi-Newton method (Broyden’s case) for solving (4.28). In particular, the new resulting method with $g(x) = x$ coincides with the quasi-Newton method for solving (4.12) with $X = R_+^n$. In both methods, the concept $\partial_b F(\cdot)$ has an important role.

QUASI-NEWTON METHOD (BROYDEN’S CASE [1]).

Given $x^0 \in R^n, A_0, B_0 \in R^{n \times n}$

Do for $k = 0, 1, \dots$:

Define

$$f^k(x) = f(x^k) + A_k(x - x^k)$$

$$g^k(x) = g(x^k) + B_k(x - x^k)$$

$$F^k(x) = \min(f^k(x), g^k(x))$$

Choose $V_k \in \partial_b F^k(x^k)$

Solve $V_k s^k + F(x^k) = 0$ for s^k

$$x^{k+1} = x^k + s^k$$

$$y^k = f(x^{k+1}) - f(x^k)$$

$$z^k = g(x^{k+1}) - g(x^k)$$

$$A_{k+1} = A_k + \frac{(y^k - A_k s^k) s^{kT}}{s^{kT} s^k}$$

$$B_{k+1} = B_k + \frac{(z^k - B_k s^k) s^{kT}}{s^{kT} s^k}.$$

COROLLARY 4.2. *Suppose that $f, g : R^n \rightarrow R^n$ are continuously differentiable, x^* is a solution of (4.28), $f'(x), g'(x)$ are Lipschitz continuous in a neighborhood of x^* , and the common Lipschitz constant is γ . Suppose that all $W_* \in \partial_b F(x^*)$ are nonsingular. There exist positive constants ε, δ such that if $\|x^0 - x^*\| \leq \varepsilon, \|A_0 - f'(x^*)\| \leq \delta$, and $\|B_0 - g'(x^*)\| \leq \delta$, then the sequence $\{x^k\}$ generated by the quasi-Newton method (Broyden's case) is well defined and converges Q-superlinearly to x^* .*

Proof. The proof is similar to that of Corollary 4.1. Here we only give an outline of the proof. It is not difficult to give the detail.

Choose ε and Δ as in the proof of Theorem 4.1 and restrict ε to be small enough such that for any $y \in N(x^*) = \{x \mid \|x - x^*\| \leq \varepsilon\}$, we have

$$(4.29) \quad \|f'(y) - f'(x^*)\| \leq \gamma \|y - x^*\|, \|g'(y) - g'(x^*)\| \leq \gamma \|y - x^*\|,$$

$$(4.30) \quad 6\gamma\varepsilon \leq \Delta.$$

Denote $\delta := \Delta/4$. From the definition of $F^k(x)$ there exists $\lambda_j^k \in \{0, 1\}$ such that the j th row V_k^j of V_k satisfies

$$(4.31) \quad V_k^j = \begin{cases} A_k^j & \text{if } f_j^k(x^k) < g_j^k(x^k), \\ \lambda_j^k A_k^j + (1 - \lambda_j^k) B_k^j & \text{if } f_j^k(x^k) = g_j^k(x^k), \\ B_k^j & \text{if } f_j^k(x^k) > g_j^k(x^k). \end{cases}$$

For such constants λ_j^k we define a companion matrix W_k such that the j th row W_k^j of W_k satisfies

$$(4.32) \quad W_k^j = \begin{cases} f_j'(x^k) & \text{if } f_j^k(x^k) < g_j^k(x^k), \\ \lambda_j^k f_j'(x^k) + (1 - \lambda_j^k) g_j'(x^k) & \text{if } f_j^k(x^k) = g_j^k(x^k), \\ g_j'(x^k) & \text{if } f_j^k(x^k) > g_j^k(x^k). \end{cases}$$

From $f(x^k) = f^k(x^k), g(x^k) = g^k(x^k)$, and the definition of $\partial_b F(x^k)$, we get

$$W_k \in \partial_b F(x^k).$$

From (4.31) and (4.32), for any $x \in R^n$ we get

$$(4.33) \quad \|(V_k - W_k)x\| \leq \|(A_k - f'(x^k))x\| + \|(B_k - g'(x^k))x\|.$$

Thus,

$$\begin{aligned}
 \|W_k - V_k\| &\leq \|A_k - f'(x^k)\| + \|B_k - g'(x^k)\| \\
 &\leq \|A_k - f'(x^*)\| + \|f'(x^k) - f'(x^*)\| \\
 (4.34) \quad &\quad + \|B_k - g'(x^*)\| + \|g'(x^k) - g'(x^*)\|.
 \end{aligned}$$

The local Q -linear convergence proof consists of showing by induction that

$$\|A_k - f'(x^*)\| \leq (2 - 2^{-k})\delta, \quad \|B_k - g'(x^*)\| \leq (2 - 2^{-k})\delta,$$

$$\|V_k - W_k\| \leq \Delta.$$

The induction proof is similar to that of Corollary 4.1. We omit it here.

To prove the Q -superlinear convergence of $\{x^k\}$, let $E_k = A_k - f'(x^*)$ and $H_k = B_k - g'(x^*)$. From the last part of the proof of Theorem 8.2.2 of [6] (also see [5]) we get

$$(4.35) \quad \lim_{k \rightarrow \infty} \frac{\|E_k s^k\|}{\|s^k\|} = 0, \quad \lim_{k \rightarrow \infty} \frac{\|H_k s^k\|}{\|s^k\|} = 0.$$

From (4.33) and (4.29), we have

$$\begin{aligned}
 \|(V_k - W_k)s^k\| &\leq \|(A_k - f'(x^k))s^k\| + \|(B_k - g'(x^k))s^k\| \\
 (4.36) \quad &\leq \|E_k s^k\| + \gamma \|e^k\| \|s^k\| + \|H_k s^k\| + \gamma \|e^k\| \|s^k\|.
 \end{aligned}$$

Thus, from (4.35), (4.36), and the linear convergence of $\{x^k\}$, we get

$$\lim_{k \rightarrow \infty} \frac{\|(V_k - W_k)s^k\|}{\|s^k\|} = 0,$$

which, from Theorem 4.2, means that $\{x^k\}$ converges to x^* Q -superlinearly. \square

In [21], Qi discussed a Newton method for solving (4.28) and provided a method to compute $\partial_B F$. Here, by using the concept $\partial_b F$, we give a quasi-Newton method. The main condition to guarantee the local Q -superlinear convergence is the nonsingularity assumption of $\partial_b F(x^*)$. When $g(x) = x$, this nonsingularity assumption is exactly the b -regularity in [19].

5. Implementation of the quasi-Newton method. The implementation of the quasi-Newton method discussed in section 4 for solving equation (4.12) has no difference to the smooth case except for the implementation of the QR factorization of the iterate matrix V_k . The entire QR factorization of V_k costs $O(n^3)$ arithmetic operations. If we do this in every step, then the advantage of quasi-Newton method loses a lot. In this section, we will show how to update the QR factorization of V_k into the QR factorization of V_{k+1} at most in $O((I(k) + 1)n^2)$ operations (see (5.8) for the definition of $I(k)$). For simplicity, we will assume that $X = R_+^n$.

For a given vector $x \in R^n$, denote the index sets

$$\alpha(x) = \{i : x_i > f_i(x)\},$$

$$\beta(x) = \{i : x_i = f_i(x)\},$$

$$\gamma(x) = \{i : x_i < f_i(x)\}.$$

Suppose for each k that we choose $V_k \in \partial_b F^k(x^k)$ such that the i th row V_k^i of V_k satisfies

$$(5.1) \quad V_k^i = \begin{cases} A_k^i & \text{if } i \in \alpha(x^k), \\ I^i & \text{if } i \in \beta(x^k) \cup \gamma(x^k). \end{cases}$$

Denote a matrix \bar{V}_k such that its i th row \bar{V}_k^i satisfies

$$(5.2) \quad \bar{V}_k^i = \begin{cases} A_{k+1}^i & \text{if } i \in \alpha(x^k), \\ I^i & \text{if } i \in \beta(x^k) \cup \gamma(x^k). \end{cases}$$

From (5.1), (5.2), and (4.14), we get

$$(5.3) \quad \bar{V}_k = V_k + \frac{(\bar{y}^k - V_k s^k) s^{kT}}{s^{kT} s^k},$$

where \bar{y}^k satisfies

$$(5.4) \quad \bar{y}_i^k = \begin{cases} y_i^k & \text{if } i \in \alpha(x^k), \\ s_i^k & \text{if } i \in \beta(x^k) \cup \gamma(x^k). \end{cases}$$

It is well known that we can update the QR factorization of V_k into the QR factorization of \bar{V}_k in $O(n^2)$ operations (see, e.g., [7, 8]).

The i th row V_{k+1}^i of V_{k+1} satisfies

$$(5.5) \quad V_{k+1}^i = \begin{cases} A_{k+1}^i & \text{if } i \in \alpha(x^{k+1}), \\ I^i & \text{if } i \in \beta(x^{k+1}) \cup \gamma(x^{k+1}). \end{cases}$$

Therefore,

$$(5.6) \quad V_{k+1} = \bar{V}_k + \Delta \bar{V}_k,$$

where $\Delta \bar{V}_k$ satisfies

$$(5.7) \quad \Delta \bar{V}_k^i = \begin{cases} 0 & \text{if } i \in \alpha(x^k) \cap \alpha(x^{k+1}), \\ 0 & \text{if } i \in \{\beta(x^k) \cup \gamma(x^k)\} \cap \{\beta(x^{k+1}) \cup \gamma(x^{k+1})\}, \\ V_{k+1}^i - \bar{V}_k^i & \text{otherwise.} \end{cases}$$

Denote

$$(5.8) \quad I(k) = n - (|\alpha(x^k) \cap \alpha(x^{k+1})| + |\{\beta(x^k) \cup \gamma(x^k)\} \cap \{\beta(x^{k+1}) \cup \gamma(x^{k+1})\}|).$$

Since the number of the nonzero rows of $\Delta\bar{V}_k$ is at most $I(k)$, we can update the QR factorization of \bar{V}_k into the QR factorization of V_{k+1} at most in $O(I(k)n^2)$ operations (see, e.g., [7, 8]).

Therefore, we get the following theorem.

THEOREM 5.1. *The cost of updating the QR factorization of V_k into the QR factorization of V_{k+1} is at most $O((I(k) + 1)n^2)$ arithmetic operations.*

Joseph [10] considered the quasi-Newton method for solving generalized equations (see Robinson [24]). For nonlinear complementarity problems, in every step his method needs to solve a linear complementarity problems, which requires more cost than solving a linear equation. Kojima and Shindo [11] extended the quasi-Newton method to piecewise smooth equations. They applied the classical Broyden's method as the points x^k stayed within a given C^1 -piece. When the points x^k arrived at a new piece, a new starting matrix was used and it was needed to perform the entire QR factorization (or other factorizations) in $O(n^3)$ operations in general. Thus a potentially large number of matrices need to be stored and need to be performed to get an entire QR factorization (or other factorizations). Here, our method needs only one approximate matrix, and except for the first step we only need less effort to solve a linear equation, which may be solved in much less than $O(n^3)$ operations. The smaller the measure of $I(k)$ is, the less computing effort is needed in the $(k + 1)$ th step (note that $I(k)$ is related to the nonsmoothness of F). Ip and Kyparisis [9] discussed the local convergence of the classical Broyden's quasi-Newton method for solving nonsmooth equations. Although the form used in [9] is very simple, the convergence remains open without assuming the existence of $F'(x^*)$.

6. The KKT system of variational inequality problems. For a given closed set $X \subseteq R^n$ and a mapping $f : X \rightarrow R^n$, the variational inequality problem which is denoted by $VI(X, f)$ is to find a vector $x^* \in X$ such that

$$(x - x^*)^T f(x^*) \geq 0 \quad \text{for all } x \in X.$$

If $X = R_+^n$, then $VI(X, f)$ is equivalent to the complementarity problem which is to find $x^* \in R_+^n$ such that

$$f(x^*) \in R_+^n \text{ and } x^{*T} f(x^*) = 0.$$

When f is a gradient mapping, say $f(x) = \nabla\theta(x)$ for some real-valued function θ , $VI(X, f)$ is equivalent to the problem of finding a stationary point for the following minimization problem:

$$\begin{aligned} &\text{minimize } \theta(x) \\ &\text{subject to } x \in X. \end{aligned}$$

Here we shall assume that X has the form

$$(6.1) \quad X = \{x \in R^n \mid g(x) \leq 0, h(x) = 0, l \leq x \leq u\},$$

where $g : R^n \rightarrow R^m$ and $h : R^n \rightarrow R^p$ are assumed to be twice continuously differentiable, and $l, u \in \{R \cup \{\infty\}\}^n$. By introducing multipliers $(\lambda, \mu, v, w) \in R^{m+p+2n}$ corresponding to the constraints in X , the (VI) Lagrangian (vector-valued) function (see, e.g., Tobin [29]) can be defined by

$$L(x, \lambda, \mu, v, w) = f(x) + \sum_{i=1}^m \nabla g_i(x) \lambda_i + \sum_{j=1}^p \nabla h_j(x) \mu_j - v + w.$$

If $l_i = -\infty$ (or $u_i = +\infty$) for some i , the corresponding v_i (w_i , respectively) is absent in the above formula. Then the KKT system of VI(X, f) can be written as

$$(6.2) \quad \begin{cases} L(x, \lambda, \mu, v, w) = 0, \\ \lambda \geq 0, -g(x) \geq 0, \text{ and } \lambda^T g(x) = 0, \\ -h(x) = 0, \\ v \geq 0, x - l \geq 0, \text{ and } v^T(x - l) = 0, \\ w \geq 0, u - x \geq 0, \text{ and } w^T(x - u) = 0. \end{cases}$$

Define

$$\tilde{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \nabla g_i(x) \lambda_i + \sum_{j=1}^p \nabla h_j(x) \mu_j$$

and

$$(6.3) \quad H(x, \lambda, \mu) = \begin{pmatrix} x - P_{[l,u]}[x - \tilde{L}(x, \lambda, \mu)] \\ \lambda - P_{R_+^n}[\lambda - (-g(x))] \\ -h(x) \end{pmatrix}.$$

Suppose that $(x^*, \lambda^*, \mu^*, v^*, w^*) \in R^{n+m+p+2n}$ is a solution of the KKT system (6.2), then (x^*, λ^*, μ^*) satisfies $H(x^*, \lambda^*, \mu^*) = 0$; conversely, if $(x^*, \lambda^*, \mu^*) \in R^{n+m+p}$ is a solution of $H(x, \lambda, \mu) = 0$, then $(x^*, \lambda^*, \mu^*, v^*, w^*)$ is a solution of the KKT system (6.2), where v^*, w^* are defined as

$$(6.4) \quad v^* = P_{R_+^n}[\tilde{L}(x^*, \lambda^*, \mu^*)] \text{ and } w^* = P_{R_+^n}[-\tilde{L}(x^*, \lambda^*, \mu^*)].$$

So finding a solution of the KKT system of VI is equivalent to solving $H(x, \lambda, \mu) = 0$. Let $z = (x, \lambda, \mu)$, $K = [l, u] \times R_+^n \times R^p$, and

$$\tilde{f}(z) = \begin{pmatrix} \tilde{L}(z) \\ -g(x) \\ -h(x) \end{pmatrix}.$$

Then $H(x, \lambda, \mu) = 0$ can be written as

$$(6.5) \quad H(z) = z - P_K[z - \tilde{f}(z)] = 0,$$

which is a special form of (4.12).

Now suppose that z^* is a solution of $H(z) = 0$ and f is continuously differentiable at x^* ; we will discuss a sufficient condition on the nonsingularity assumption of $\partial_b H(z^*)$. Let

$$I(z^*) = \{i \mid 1 \leq i \leq m, g_i(x^*) = 0\},$$

$$I^+(z^*) = \{i \in I(z^*) \mid \lambda_i^* > 0\},$$

$$G^+(z^*) = \{d \in R^n \mid \nabla g_i(x^*)^T d = 0 \text{ for } i \in I^+(z^*) \\ \text{and } \nabla h_i(x^*)^T d = 0 \text{ for } i = 1, \dots, p\},$$

and

$$R(z^*) = \{d \in R^n \mid d_i = 0 \text{ if } x_i^* = l_i \text{ (or } u_i) \text{ and } (\tilde{L}(z^*))_i \neq 0 \text{ for } i = 1, \dots, n\}.$$

THEOREM 6.1. *Suppose that z^* is a solution of $H(z) = 0$ and that it satisfies $d^T \nabla_{xx}^2 \tilde{L}(z^*) d > 0$ for all $d \in G^+(z^*) \cap R(z^*) \setminus \{0\}$. If $\{\nabla g_i(x^*), i \in I(z^*)\}$ and $\{\nabla h_i(x^*), i = 1, \dots, p\}$ are linearly independent, then all $V \in \partial_b H(z^*)$ are nonsingular.*

Proof. Combining (4.15) and the proof of Theorem 4.1 in Robinson [24], we can get the result. \square

7. Numerical examples. In this section, we report computational results obtained for two small nonlinear complementarity problems using the above Newton method and quasi-Newton method. For the quasi-Newton method, the initial matrices are generated by the difference approximation method. In Table 1, “N” and “QN” represent the Newton method and quasi-Newton method, respectively, and “P 1” and “P 2” represent Problem 1 and Problem 2, respectively.

Problem 1 (a nondegenerate nonlinear complementarity problem [10, 9]). Consider the following problem: find $x \in R^4$ such that $x \geq 0$, $f(x) \geq 0$, and $x^T f(x) = 0$, where $f : R^4 \rightarrow R^4$ is given by

$$f_1(x) = 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6, \\ f_2(x) = 2x_1^2 + x_1 + x_2^2 + 3x_3 + 2x_4 - 2, \\ f_3(x) = 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1, \\ f_4(x) = x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3.$$

This problem has the solution

$$x^* = \left(\frac{1}{2}\sqrt{6} \approx 1.2247, 0, 0, 0.5\right), \quad f(x^*) = \left(0, 2 + \frac{1}{2}\sqrt{6} \approx 3.2247, 5, 0\right).$$

Since $\beta(x^*) = \emptyset$, x^* is nondegenerate (see [9]) and it is easy to check that $F'(x^*)$ (here $\partial_b F'(x^*) = \{F'(x^*)\}$) is nonsingular.

Problem 2 (a degenerate nonlinear complementarity problem [11, 9]). Consider the following problem: find $x \in R^4$ such that $x \geq 0$, $f(x) \geq 0$, and $x^T f(x) = 0$, where $f : R^4 \rightarrow R^4$ is given by

$$f_1(x) = 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6, \\ f_2(x) = 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2, \\ f_3(x) = 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9, \\ f_4(x) = x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3.$$

TABLE 1

Results for Problems 1 and 2, where D = degenerate solution and ND = nondegenerate solution.

Algorithm	Starting point	Number of Iterations		sum of $I(k)$	
		P 1	P 2	P 1	P 2
N	(1,0,0,0)	3	3(D)		
QN	(1,0,0,0)	4	4(D)	0	2
N	(1,0,1,0)	4	1(ND)		
QN	(1,0,1,0)	5	1(ND)	1	0
N	(1,0,0,1)	4	4(D)		
QN	(1,0,0,1)	5	5(D)	1	2
N	(1,0.2,0.5,1)	4	4(D)		
QN	(1,0.2,0.5,1)	6	6(D)	0	2
N	(1,0,1,-1)	3	3(D)		
QN	(1,0,1,-1)	5	5(D)	1	2
N	(1.5,-0.5,4.5,-1.0)	4	4(D)		
QN	(1.5,-0.5,4.5,-1.0)	6	6(D)	1	0
N	(1.1,-0.1,3.1,-0.1)	4	3(ND)		
QN	(1.1,-0.1,3.1,-0.1)	5	4(ND)	1	0
N	(0.85,0.2,0.5,1)	4	5(D)		
QN	(0.85,0.2,0.5,1)	7	7(D)	1	2

This problem has the following two solutions:

$$x_D^* = \left(\frac{1}{2}\sqrt{6} \approx 1.2247, 0, 0, 0.5 \right), \quad f(x_D^*) = \left(0, 2 + \frac{1}{2}\sqrt{6} \approx 3.2247, 0, 0 \right),$$

and

$$x_{ND}^* = (1, 0, 3, 0), \quad f(x_{ND}^*) = (0, 31, 0, 4).$$

Since $\beta(x_{ND}^*) = \emptyset$ for the solution x_{ND}^* , it is a nondegenerate solution (see [9]). On the other hand, $\beta(x_D^*) = \{3\}$ for the solution x_D^* , so it is a degenerate solution (see [9]). It is easy to check that $\partial_b F(x_{ND}^*)$ and $\partial_b F(x_D^*)$ are nonsingular.

From Table 1 we see that even for Problem 2 when the starting point is close to a solution, the sequence will converge to the corresponding solution no matter whether it is degenerate or not.

In this paper two small examples are used to show the effectiveness of the Newton method and the quasi-Newton method for solving some nonsmooth equations. More examples are needed to show the efficiency of the above algorithms. For problem (4.12) with a general convex set X , especially when X is a polyhedral set, how to construct appropriate Newton methods and quasi-Newton methods is our further research topic.

Acknowledgments. We are grateful to two referees for their helpful suggestions and comments on this paper, to Prof. Jong-Shi Pang for his useful comments on an earlier version of this paper, and to Prof. Liqun Qi and Dr. Xiaojun Chen for their helpful suggestions on this subject. Finally, we thank Prof. Masakazu Kojima for his help in revising this paper.

REFERENCES

[1] C. G. BROYDEN, *A class of methods for solving nonlinear simultaneous equations*, Math. Comp., 19 (1965), pp. 577–593.
 [2] X. CHEN AND T. YAMAMOTO, *On the convergence of some quasi-Newton methods for solving nonlinear equations with nondifferentiable operators*, Computing, 49 (1992), pp. 87–94.

- [3] X. CHEN AND L. QI, *A parameterized Newton method and a quasi-Newton method for solving nonsmooth equations*, *Comput. Optim. Appl.*, 3 (1994), pp. 157–179.
- [4] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983.
- [5] J. E. DENNIS AND J. J. MORÉ, *A characterization of superlinear convergence and its application to quasi-Newton methods*, *Math. Comp.*, 28 (1974), pp. 549–560.
- [6] J. E. DENNIS AND R. B. SCHNABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice–Hall, Englewood Cliffs, NJ, 1983.
- [7] P. E. GILL AND M. MURRAY, *Quasi-Newton methods for unconstrained optimization*, *J. Inst. Math. Appl.*, 9 (1972), pp. 91–108.
- [8] G. H. GOLUB AND C. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1983.
- [9] C.-M. IP AND T. KYPARISIS, *Local convergence of quasi-Newton methods for B-differentiable equations*, *Math. Programming*, 56 (1992), pp. 71–89.
- [10] N. H. JOSEPHY, *Quasi-Newton Methods for Generalized Equations*, Technical summary report 1966, Mathematical Research Center, University of Wisconsin, Madison, WI, 1979.
- [11] M. KOJIMA AND S. SHINDO, *Extensions of Newton and quasi-Newton methods to systems of PC^1 equations*, *J. Oper. Res. Soc. Japan*, 29 (1986), pp. 352–374.
- [12] B. KUMMER, *Newton's method for non-differentiable functions*, in *Advances in Mathematical Programming*, J. Guddat, B. Bank, H. Hollatz, P. Kall, D. Klatt, B. Kummer, K. Lommatzsch, L. Tammer, M. Vlach, and K. Zimmerman, eds., *Academi Verlag*, Berlin, 1988, pp. 114–125.
- [13] B. KUMMER, *Newton's method based on generalized derivatives for nonsmooth functions: Convergence analysis*, in *Lecture Notes in Economics and Mathematical Systems 382: Advances in Optimization*, W. Oettli and D. Pallaschke, eds., Springer, Berlin, 1992, pp. 171–194.
- [14] M. MIFFLIN, *Semismooth and semiconvex functions in constrained optimization*, *SIAM J. Control Optim.*, 15 (1977), pp. 957–972.
- [15] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
- [16] J.-S. PANG, *The implicit complementarity problem*, in *Nonlinear Programming 4*, O. L. Mangasarian, S. M. Robinson, and P. R. Meyer, eds., Academic Press, New York, 1981, pp. 487–518.
- [17] J.-S. PANG, *Newton's method for B-differentiable equations*, *Math. Oper. Res.*, 15 (1990), pp. 311–341.
- [18] J.-S. PANG, *A B-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems*, *Math. Programming*, 51 (1991), pp. 101–131.
- [19] J.-S. PANG AND S. A. GABRIEL, *NE/SQP: A robust algorithm for the nonlinear complementarity problem*, *Math. Programming*, 60 (1993), pp. 295–337.
- [20] J.-S. PANG AND L. QI, *Nonsmooth equations: Motivation and algorithms*, *SIAM J. Optim.*, 3 (1993), pp. 443–465.
- [21] L. QI, *Convergence analysis of some algorithms for solving nonsmooth equations*, *Math. Oper. Res.*, 18 (1993), pp. 227–244.
- [22] L. QI AND J. SUN, *A nonsmooth version of Newton's method*, *Math. Programming*, 58 (1993), pp. 353–368.
- [23] L. QI AND J. SUN, *A Nonsmooth Version of Newton's Method and an Interior Point Algorithm for Convex Programming*, Applied Mathematics Preprint 89/33, School of Mathematics, The University of New South Wales, Sydney, Australia, 1991.
- [24] S. M. ROBINSON, *Strongly regular generalized equations*, *Math. Oper. Res.*, 5 (1980), pp. 43–62.
- [25] S. M. ROBINSON, *Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity*, *Math. Programming Study*, 30 (1987), pp. 45–66.
- [26] S. M. ROBINSON, *Newton's methods for a class of nonsmooth functions*, *Set-Valued Analysis*, 2 (1994), pp. 292–305.
- [27] A. SHAPIRO, *On concepts of directional differentiability*, *J. Optim. Theory Appl.*, 66 (1990), pp. 477–487.
- [28] N. Z. SHOR, *A class of almost-differentiable functions and a minimization method for functions of this class*, *Kibernetik*, 4 (1972), pp. 65–70.
- [29] R. L. TOBIN, *Sensitivity analysis for variational inequalities*, *J. Optim. Theory Appl.*, 48 (1986), pp. 191–204.