ELSEVIER

Journal of Computational and Applied Mathematics 146 (2002) 1–10

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

An iterative method for solving semismooth equations ☆

Xiaojiao Tong[a, *], Dong-Hui Li[b,1], Liqun Qi[c]

[a]Department of Mathematics, Nanjing University, 210093 Nanjing, China
[b]Institute of Applied Mathematics, Hunan University, Changsha, China
[c]Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

Received 21 November 2000; received in revised form 17 November 2001

Abstract

In this paper, we combine trust region technique with line search technique to develop an iterative method for solving semismooth equations. At each iteration, a trust region subproblem is solved. The solution of the trust region subproblem provides a descent direction for the norm of a smoothing function. By using a backtracking line search, a steplength is determined. The proposed method shares advantages of trust region methods and line search methods. Under appropriate conditions, the proposed method is proved to be globally and superlinearly convergent. In particular, we show that after finitely many iterations, the unit step is always accepted and the method reduces to a smoothing Newton method. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Semismooth equation; Smoothing function; Trust region method; Line search

1. Introduction

The concept of semismoothness was introduced in [8] for real-valued functions and extended in [11,13] to vector-valued functions. A locally Lipschitzian function $H : \mathbb{R}^n \to \mathbb{R}^m$ is said to be semi-

☆ The work was partly supported by the Research Grant Council of Hong Kong, Polyll 5077/00P, and the National Natural Science Foundation of China, No. 10171030.

* Corresponding author. Tel.: +86-25-3594068; fax: +86-25-3604447.

E-mail addresses: luoket@public.cs.hn.cn (X. Tong), dhli@mail.hunu.edu.cn (Dong-Hui Li), maqilq@polyu.edu.hk (L. Qi).

[1] The work was done while he was visiting Hong Kong Polytechnic University.

PII: S0377-0427(02)00413-2

smooth at a point z if the following limit:

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

exists for any $h\in\mathbb{R}^n$, where $\partial H(z)$ is the generalized Jacobian of H at z in the sense of Clarke [2]. If H is semismooth at z, then for any $V\in\partial H(z+h)$, as $h\to 0$, it holds that [13]

$$H(z+h)-H(z)-Vh=\mathrm{o}(\|h\|).$$

When $H:\mathbb{R}^n\to\mathbb{R}^n$ is semismooth, the nonlinear equation

$$H(z)=0 \tag{1.1}$$

is called a semismooth equation.

Many practical problems such as nonlinear complementarity problems, the KKT systems of variational inequality problems and nonlinear programming problems can be reformulated as semismooth equations. In recent years, nonsmooth Newton methods and smoothing Newton methods for solving semismooth equations have received much attention. Under certain conditions, these methods possess global and superlinear convergence properties. We refer to [1,3,4,6] for recent progress on nonsmooth Newton methods and smoothing Newton methods.

In this paper, we aim to develop a trust region-type method [9] for solving semismooth equation (1.1). Trust region methods for solving nonsmooth equations have been studied in [5,12]. The method proposed in [5] is devoted to solving a semismooth equation reformulation for generalized complementarity problems. In their method, the subproblem is the following minimization problem:

$$\begin{aligned}&\min\ \nabla\Phi(z_k)^{\mathrm{T}}d+\tfrac{1}{2}d^{\mathrm{T}}V_k^{\mathrm{T}}V_kd\\ &\text{s.t.}\quad \|d\|\leqslant\Delta_k,\end{aligned}$$

where $\Phi(z)=\frac{1}{2}\|H(z)\|^2$ is continuously differentiable, $V_k\in\partial_BH(z_k)$ and

$$\partial_BH(z)=\left\{V\in R^{n\times n}\,|\,V=\lim_{z^k\to x}\nabla H(z^k)^{\mathrm{T}},\ H \text{ is differentiable at } z^k \text{ for all } k\right\}.$$

The method proposed in [12] is based on a smooth plus nonsmooth decomposition of H, i.e., $H=p+q$, where p is smooth and q is locally Lipschitz and relatively small. The trust region subproblem is the following minimization problem:

$$\min\{\|H(z_k)+p'(z_k)d\|_2\colon \|d\|\leqslant\Delta_k\}.$$

Unlike the above two methods, in this paper, we develop a new trust region-type method for solving general semismooth equation (1.1). The method is based on the recently developed smoothing technique. A function $H^{\varepsilon}:\mathbb{R}^n\to\mathbb{R}^n$ is said to be a smoothing function of the nonsmooth function H if it is continuously differentiable and satisfies for any $\varepsilon\geqslant 0$

$$\|H^{\varepsilon}(z)-H(z)\|\leqslant\mu\varepsilon,\quad \forall z, \tag{1.2}$$

where $\mu>0$ is a constant. The parameter ε is called smoothing parameter. Based on the smoothing function H^{ε}, we propose a trust region-type method where the trust region subproblem is the following minimization problem:

$$\min\{\|H^{\varepsilon_k}(z_k)+\nabla H^{\varepsilon_k}(z_k)d\|_2\colon \|d\|\leqslant\Delta_k\}. \tag{1.3}$$

Another difference between the proposed method and the existing trust region methods is the rule for adjusting the trust region radius. In an ordinary trust region method, at each step, if the solution $\hat{d}_k$ of the subproblem does not make $\|H(z_k + \hat{d}_k)\|$ smaller than $\|H(z_k)\|$ sufficiently, then the trust region radius Δ_k is decreased and the subproblem is solved again. This process is repeated until $\|H(z_k + \hat{d}_k)\|$ is less than $\|H(z_k)\|$ sufficiently. Thus, at each step of an ordinary trust region method, the subproblem needs to be solved many times. In our method, we adjust the trust region radius in a different way. Specifically, at each step, we only solve the subproblem once. If $\hat{d}_k$ is not accepted, then we use a line search strategy. The solution of subproblem (1.3) provides a descent direction of the function $\|H^{\varepsilon_k}\|$ at z_k. We then use a backtracking line search to determine a steplength λ_k. The next iterate is obtained by letting $z_{k+1} = z_k + \lambda_k d_k$. In this sense, the proposed method is actually a combination of a trust region method with a line search method. This combination of the trust region technique and the line search technique was introduced in [10] for solving unconstrained optimization problems. We extend this technique to develop an iterative method for solving semismooth equations. The advantage of the proposed method is two fold. First, it shares the advantages of trust region methods. The trust region subproblem always has a solution whether $\nabla H^{\varepsilon_k}(z_k)$ is nonsingular or singular. Second, at each step, the subproblem is solved once only. Under appropriate conditions, we prove the global convergence of the proposed method. Moreover, we show that the proposed method actually reduces to a smoothing Newton method with unit steplength after finitely many iterations. Consequently, it is superlinearly convergent.

In Section 2, we state the steps of the algorithm and show that the algorithm is well defined. In Sections 3 and 4, we prove the global and superlinear convergence of the proposed method.

Throughout this paper, without specification, the norm of a vector or a matrix is l_2 norm. For a real-valued function f, we use ∇f to denote its gradient, and for a vector-valued function F, $\nabla F(z)$ stands for the Jacobian of F at z.

2. Algorithm

In this section, we give the steps of the algorithm and prove its well-definiteness. Let $H^{\varepsilon} : \mathbb{R}^n \to \mathbb{R}^n$ be a smoothing function of H that satisfies (1.2). For the sake of simplicity, we use H^k to stand for the abbreviation of H^{ε_k}. Let

$$W(z) = \tfrac{1}{2}\|H(z)\|^2,$$

$$W^k(z) = \tfrac{1}{2}\|H^k(z)\|^2.$$

The basic idea of our method has been described in the last section. We give the steps of the algorithm as follows.

Algorithm 2.1

Step 0: Given some constants $\alpha_1 > 1$, $0 < \eta_1 < \eta_2 < 1$, $\alpha, \eta \in (0, \frac{1}{2})$, $\Delta_0 > 0$, $\beta \in (0,1)$. Chosen an initial point $z_0 \in \mathbb{R}^n$ and an initial $\varepsilon_0 \in (0, (\alpha/2\mu)\|H(z_0)\|)$. Let $k{:=}0$.

Step 1: Solve the following trust region subproblem to obtain a trial step d_k.

$$\begin{cases} \min\ Q^k(d) = \frac{1}{2}\|H^k(z_k) + \nabla H^k(z_k)d\|^2, \\ \text{s.t.}\quad \|d\| \leqslant \Delta_k. \end{cases} \tag{2.1}$$

Step 2: Compute

$$r_k = \frac{W^k(z_k) - W^k(z_k + d_k)}{W^k(z_k) - Q^k(d_k)}.$$

If $r_k \geqslant \eta_1$, set $z_{k+1} = z_k + d_k$ and go to Step 4.

Step 3: Let i_k be the smallest nonnegative integer i such that

$$W^k(z_k) - W^k(z_k + \beta^i d_k) \geqslant -\eta\beta^i d_k^{\mathrm{T}}(\nabla H^k(z_k)^{\mathrm{T}} H^k(z_k)). \tag{2.2}$$

Let $z_{k+1} := z_k + \beta^{i_k} d_k$ and $\Delta_{k+1} := \Delta_k$ and go to Step 5.

Step 4: Choose

$$\Delta_{k+1} := \begin{cases} \Delta_k & \text{if } \eta_1 \leqslant r_k < \eta_2, \\ a_1 \Delta_k & \text{if } r_k \geqslant \eta_2. \end{cases} \tag{2.3}$$

Step 5: If $\|H(z_{k+1})\| = 0$, stop and output z_{k+1}.

Step 6: If $\|H(z_{k+1})\| \geqslant \alpha\|H(z_k)\| + \mu\alpha^{-1}\varepsilon_k$, set $\varepsilon_{k+1} = \varepsilon_k$. Otherwise choose a ε_{k+1} satisfying

$$\begin{cases} \varepsilon_{k+1} \leqslant \min\{\frac{1}{2}\varepsilon_k, \frac{\alpha}{2\mu}\|H(z_{k+1})\|, W(z_{k+1})\}, \\ \mathrm{dist}(\nabla H^{k+1}(z_{k+1}), \partial_C H(z_{k+1})) \leqslant \bar{\mu}\varepsilon_k, \end{cases} \tag{2.4}$$

where $\bar{\mu} > 0$ is a constant and

$$\partial_C H(z) = \partial H_1(z) \times \partial H_2(z) \times \cdots \times \partial H_n(z).$$

Let $k := k + 1$ and go to Step 1.

It is clear from Algorithm 2.1 that at each step, trust region subproblem (2.1) is solved once only. The sequence $\{\varepsilon_k\}$ is nonincreasing. Define

$$K = \{0\} \cup \{k \,|\, \|H(z_k)\| < \alpha\|H(z_{k-1})\| + \mu\alpha^{-1}\varepsilon_{k-1}\}. \tag{2.5}$$

Then it is easy to deduce the following inequalities:

$$\mu\varepsilon_k \leqslant \alpha\|H(z_k)\|, \quad \forall k, \tag{2.6}$$

$$\varepsilon_k \leqslant \tfrac{1}{2}\varepsilon_{k-1}, \quad \forall k \in K, \tag{2.7}$$

$$\mathrm{dist}(\nabla H^k(z_k), \partial_C H(z_k)) \leqslant \bar{\mu}\varepsilon_{k-1}. \tag{2.8}$$

The purpose of (2.8) is to make $\nabla H^k(z_k)$ satisfy the so-called Jacobian consistence property [1].

The next lemma shows that any solution of (2.1) is a descent direction of W^k at z_k unless that z_k is a solution of (1.1).

Lemma 2.1. *Let* $\{z_k\}$ *and* $\{d_k\}$ *be generated by Algorithm* 2.1. *Then we have*

$$W^k(z_k) - Q^k(d_k) \geqslant \frac{1}{2}\|\nabla H^k(z_k)^{\mathrm{T}} H^k(z_k)\| \min\left\{\Delta_k, \frac{\|\nabla H^k(z_k)^{\mathrm{T}} H^k(z_k)\|}{\|\nabla H^k(z_k)^{\mathrm{T}} \nabla H^k(z_k)\|}\right\} \tag{2.9}$$

and

$$d_k^{\mathrm{T}}(\nabla H^k(z_k)^{\mathrm{T}} H^k(z_k)) \leqslant -\frac{1}{2}\|\nabla H^k(z_k)^{\mathrm{T}} H^k(z_k)\| \min\left\{\Delta_k, \frac{\|\nabla H^k(z_k)^{\mathrm{T}} H^k(z_k)\|}{2\|\nabla H^k(z_k)^{\mathrm{T}} \nabla H^k(z_k)\|}\right\}. \tag{2.10}$$

Proof. Inequality (2.9) can be obtained in a way similar to the proof of Lemma 1 in [12], and inequality (2.10) can be proved in a way similar to the proof of Lemma 2.4 in [10]. □

Inequality (2.10) particularly implies that any solution of the trust region subproblem provides a descent direction of function W^k at z_k. Therefore for each k, we can find a finite integer i_k such that (2.2) holds. On the other hand, if H satisfies the Jacobian consistency property [1], then inequalities in (2.4) are satisfied for all sufficiently small ε_{k+1}. Consequently, Algorithm 2.1 is well-defined.

3. Global convergence

In this section, we prove the global convergence of Algorithm 2.1. We first introduce the following blanket assumptions:

Assumption 1. The level set

$$L_0 = \left\{ z \in \mathbb{R}^n \mid \|H(z)\| \leqslant \frac{2-\alpha}{1-2\alpha} \|H(z_0)\| \right\}$$

is bounded.

Assumption 2. For each $\varepsilon > 0$ and any $z \in L_0$, the matrix $\nabla H^\varepsilon(z)$ is nonsingular.

Lemma 3.1. *Let* $\{z_k\}$ *be generated by Algorithm* 2.1. *Then we have* $\{z_k\} \subset L_0$.

Proof. Let $K = \{k_0 = 0 < k_1 < k_2 < \cdots\}$. From Algorithm 2.1, we have $\varepsilon_k = \varepsilon_{k_j}$ for each k satisfying $k_j \leqslant k < k_{j+1}$. By Steps 2 and 3, we have for each $k_j \leqslant k < k_{j+1}$

$$\|H^k(z_k)\| = \|H^{k_j}(z_k)\| \leqslant \|H^{k_j}(z_{k-1})\| \leqslant \cdots \leqslant \|H^{k_j}(z_{k_j})\|. \tag{3.1}$$

Since $k_{j+1} \in K$, we deduce

$$\begin{aligned} \|H(z_{k_{j+1}})\| &< \alpha\|H(z_{k_{j+1}-1})\| + \mu\alpha^{-1}\varepsilon_{k_{j+1}-1} \\ &\leqslant \alpha\|H^{k_j}(z_{k_{j+1}-1})\| + \alpha\|H^{k_j}(z_{k_{j+1}-1}) - H(z_{k_{j+1}-1})\| + \mu\alpha^{-1}\varepsilon_{k_j} \\ &\leqslant \alpha\|H^{k_j}(z_{k_j})\| + \mu(\alpha + \alpha^{-1})\varepsilon_{k_j} \\ &\leqslant \alpha\|H(z_{k_j})\| + \alpha\|H(z_{k_j}) - H^{k_j}(z_{k_j})\| + \mu(\alpha + \alpha^{-1})\varepsilon_{k_j} \\ &\leqslant \alpha\|H(z_{k_j})\| + \mu(2\alpha + \alpha^{-1})\varepsilon k_j, \end{aligned} \tag{3.2}$$

where the third inequality follows form (1.2) and (3.1). Note that $\alpha < \frac{1}{2}$, $\varepsilon_{k_j} \leqslant \frac{1}{2}\varepsilon_{k_j-1} \leqslant \cdots \leqslant (\frac{1}{2})^j\varepsilon_0$. By the use of Lemma 7 in [7], we deduce from (3.2)

$$\begin{aligned} \|H(z_{k_j})\| &\leqslant \left(\|H(z_0)\| + \frac{1}{(1/2)-\alpha}\mu(2\alpha + \alpha^{-1})\varepsilon_0 \right) \left(\frac{1}{2}\right)^j \\ &= \left(1 + \frac{1+2\alpha^2}{1-2\alpha}\right)\left(\frac{1}{2}\right)^j \|H(z_0)\|. \end{aligned}$$

This together with (1.2) and (3.1) implies

$$\begin{aligned}
\|H(z_k)\| &\leqslant \|H^{k_j}(z_k)\| + \mu\varepsilon_{k_j} \\
&\leqslant \|H^{k_j}(z_{k_j})\| + \mu\varepsilon_{k_j} \\
&\leqslant \|H(z_{k_j})\| + 2\mu\varepsilon_{k_j} \\
&\leqslant \left(1 + \frac{1+2\alpha^2}{1-2\alpha}\right)\left(\frac{1}{2}\right)^j \|H(z_0)\| + 2\mu\left(\frac{1}{2}\right)^j \varepsilon_0 \\
&\leqslant \left(1 + \frac{1+2\alpha^2}{1-2\alpha} + \alpha\right)\left(\frac{1}{2}\right)^j \|H(z_0)\| \\
&= \frac{2-\alpha}{1-2\alpha}\left(\frac{1}{2}\right)^j \|H(z_0)\| \\
&\leqslant \frac{2-\alpha}{1-2\alpha} \| H(z_0)\|.
\end{aligned}$$

This implies $\{z_k\} \subset L_0$. □

It is not difficult to get the following lemma from the proof of Lemma 3.1.

Lemma 3.2. *If the index set K is infinite, then we have*

$$\lim_{k\to\infty} \varepsilon_k = 0, \tag{3.3}$$

$$\lim_{k\to\infty} \operatorname{dist}(\nabla H^k(z_k), \partial_C H(z_k)) = 0, \tag{3.4}$$

and

$$\lim_{k\to\infty} \| H(z_k)\| = 0. \tag{3.5}$$

The following theorem shows the global convergence of Algorithm 2.1.

Theorem 3.1. *Let $\{z_k\}$ be generated by Algorithm* 2.1. *Then we have*

$$\lim_{k\to\infty} \|H(z_k)\| = 0. \tag{3.6}$$

In particular, every accumulation point of $\{z_k\}$ is a solution of (1.1).

Proof. From Lemma 3.2, we only need to prove that the index set K is infinite. For the sake of contradiction, we assume that K is finite. Then by Step 6 of Algorithm 2.1, there exists an integer $m_1 > 0$ such that for all $k > m_1$

$$\|H(z_{k+1})\| \geqslant \alpha\|H(z_k)\| + \mu\alpha^{-1}\varepsilon_k$$

and $\varepsilon_k = \varepsilon_{m_1} \triangleq \bar{\varepsilon} > 0$ for all $k \geqslant m_1$. This implies $H^k = H^{m_1} \triangleq \bar{H}$ for all $k \geqslant m_1$. Consequently, we have for all $k > m_1$

$$\|H(z_k)\| \geqslant \alpha\|H(z_{k-1})\| + \mu\alpha^{-1}\varepsilon_{k-1} = \alpha\|H(z_{k-1})\| + \mu\alpha^{-1}\bar{\varepsilon} \geqslant \mu\alpha^{-1}\bar{\varepsilon}.$$

This implies

$$\|\bar{H}(z_k)\| \geqslant \|H(z_k)\| - \|H(z_k) - \bar{H}(z_k)\| \geqslant \mu\alpha^{-1}\bar{\varepsilon} - \mu\bar{\varepsilon} = \mu(\alpha^{-1} - 1)\bar{\varepsilon} > 0. \tag{3.7}$$

If there is an index set K' such that $\{\|\nabla\bar{H}(z_k)^{\mathrm{T}}\bar{H}(z_k)\|\}_{K'}$ tends to zero, then by Assumption 2, any accumulation point $\bar{z}$ of $\{z_k\}_{K'}$ satisfies $\bar{H}(\bar{z}) = 0$. This contradicts (3.7). The contradiction shows that there is a constant $\delta > 0$ such that

$$\|\nabla\bar{H}(z_k)^{\mathrm{T}}\bar{H}(z_k)\| \geqslant \delta \tag{3.8}$$

holds for all k. Let $M_2 > 0$ be an upper bound of $\{\|\nabla\bar{H}(z_k)\|\}$. In the following, we deduce a contradiction by discussing two cases.

Case 1. Step 3 is used for finitely many k only. By Step 2 of Algorithm 2.1, there exists an integer $m_2 > 0$ such that the inequality $r_k \geqslant \eta_1$ holds for all $k > m_2$. It then follows from Lemma 2.1 and the definition of r_k that

$$\begin{aligned}\bar{W}(z_k) - \bar{W}(z_{k+1}) &\geqslant \frac{1}{2}\eta_1\|\nabla\bar{H}(z_k)^{\mathrm{T}}\bar{H}(z_k)\| \min\left\{\Delta_k, \frac{\|\nabla\bar{H}(z_k)^{\mathrm{T}}\bar{H}(z_k)\|}{\|\nabla\bar{H}(z_k)^{\mathrm{T}}\nabla\bar{H}(z_k)\|}\right\}\\ &\geqslant \frac{1}{2}\eta_1\delta\min\{\Delta_k, M_2^{-2}\delta\}.\end{aligned}$$

This implies $\Delta_k \to 0$ as $k \to \infty$. However, by the steps of Algorithm 2.1, the sequence $\{\Delta_k\}$ is nondecreasing. So, we get a contradiction.

Case 2. The line search step is used for infinitely many k. Let K_1 be the set of indices k at which Step 3 is used. From Algorithm 2.1 we have $\Delta_k \geqslant \Delta_0 > 0$. Therefore, it follows from Lemma 2.1, (3.8) and (2.2) that the following inequality holds for all $k \in K_1$.

$$\bar{W}(z_k) - \bar{W}(z_{k+1}) \geqslant \tfrac{1}{2}\eta\beta^{i_k}\delta\min\{\Delta_0, \tfrac{1}{2}M_2^{-2}\delta\}.$$

Since $\{\bar{W}(z_k)\}$ is nonincreasing, taking limits in both sides of the above inequality, we get $\beta^{i_k} \to 0$ as $k \to \infty$ with $k \in K_1$.

On the other hand, however, by the mean-value theorem, for any nonnegative integer i, there is a constant $\bar{\theta}_k \in (0,1)$ such that

$$\begin{aligned}\bar{W}(z_k) - \bar{W}(z_k + \beta^i d_k) &= -\beta^i\nabla\bar{W}(z_k + \bar{\theta}_k\beta^i d_k)^{\mathrm{T}}d_k\\ &= -\beta^i\nabla\bar{W}(z_k)^{\mathrm{T}}d_k - \beta^i(\nabla\bar{W}(z_k + \bar{\theta}_k\beta^i d_k) - \nabla\bar{W}(z_k))^{\mathrm{T}}d_k\\ &= -\beta^i d_k^{\mathrm{T}}(\nabla\bar{H}(z_k)^{\mathrm{T}}\bar{H}(z_k)) + \mathrm{o}(\beta^i)\\ &= -\eta\beta^i d_k^{\mathrm{T}}(\nabla\bar{H}(z_k)^{\mathrm{T}}\bar{H}(z_k)) - (1-\eta)\beta^i d_k^{\mathrm{T}}(\nabla\bar{H}(z_k)^{\mathrm{T}}\bar{H}(z_k)) + \mathrm{o}(\beta^i)\\ &\geqslant -\eta\beta^i d_k^{\mathrm{T}}(\nabla\bar{H}(z_k)^{\mathrm{T}}\bar{H}(z_k)) + (1-\eta)\tfrac{1}{2}\beta^{i_k}\delta\min\{\Delta_0, \tfrac{1}{2}M_2^{-2}\delta\} + \mathrm{o}(\beta^i),\end{aligned}$$

where the last inequality follows from (2.10) and (3.8). By the line search rule, the last inequality shows that $\{\beta^{i_k}\}_{K_1}$ is bounded away from zero. This contradicts the fact that $\beta^{i_k} \to 0$. The proof is then complete. □

4. Superlinear convergence

In this section, we show the superlinear convergence of Algorithm 2.1. First, we note that Lemma 3.1 and Theorem 3.1 have shown that under the conditions of Assumptions 1 and 2 the sequence of $\{z_k\}$ generated by Algorithm 2.1 has at least one accumulation point, and every accumulation point of $\{z_k\}$ is a solution of (1.1). Moreover, we have

$$\lim_{k\to\infty} \varepsilon_k = \lim_{k\to\infty} \|H(z_k)\| = \lim_{k\to\infty} \|H^k(z_k)\| = 0. \tag{4.1}$$

This together with (2.8) implies

$$\lim_{k\to\infty} \operatorname{dist}(\nabla H^k(z_k), \partial_C H(z_k)) = 0. \tag{4.2}$$

It is also easy to see that the index set K defined by (2.5) is infinite. Moreover, by Step 6 of Algorithm 2.1, we have for each $k\in K$, $\varepsilon_k = \mathrm{O}(\|W(z_k)\|) = \mathrm{o}(\|H(z_k)\|)$. Let z^* be an accumulation point of $\{z_k\}$ and K_0 be an infinite subset of K such that

$$\lim_{k\in K_0} z_k = z^*.$$

To show the superlinear convergence of Algorithm 2.1, we need the following additional assumption.

Assumption 3. Every matrix of $\partial_C H(z^*)$ is nonsingular.

Under assumptions Assumptions 1–3, we have the following lemma.

Lemma 4.1. *Assume that Assumptions* 1–3 *hold, then* $d_k = -\nabla H^k(z_k)^{-1} H^k(z_k)$ *is the unique solution of trust region subproblem* (2.1) *for sufficiently large* $k\in K_0$. *Moreover, the trial step* d_k *can be accepted by Step* 2 *in Algorithm* 2.1, *i.e.*, $z_{k+1} = z_k + d_k$ *with* $k\in K_0$ *large enough.*

Proof. Denote $s_k = -\nabla H^k(z_k)^{-1} H^k(z_k)$. It is clear from (4.2) that there is a constant $M > 0$ such that the inequality $\|\nabla H^k(z_k)^{-1}\| \leqslant M$ holds for all $k\in K_0$ sufficiently large. On the other hand, from Algorithm 2.1 we have $\Delta_k \geqslant \Delta_0 > 0$. Then for large enough $k\in K_0$ it follows

$$\|s_k\| = \| - \nabla H^k(z_k)^{-1} H^k(z_k)\| \leqslant \Delta_k,$$

which shows that the Newton direction is a feasible point of (2.1). Therefore, the solution of (2.1) is $d_k = s_k$.

Next, we prove that d_k can be accepted by Step 2 in Algorithm 2.1. Let $V_k \in \partial_C H(z_k)$ satisfy

$$\operatorname{dist}(\nabla H^k H^k(z_k), \partial_C H(z_k)) = \|\nabla H^k(z_k) - V_k\|.$$

It follows from semismoothness of $H(z)$, the updating rule of ε_k in Algorithm 2.1 and (4.2) that

$$\begin{aligned}\|z_k + d_k - z^*\| &= \|z_k - \nabla H^k(z_k)^{-1} H^k(z_k) - z^*\| \\ &= \|(-\nabla H^k(z_k)[H^k(z_k) - \nabla H^k(z_k)(z_k - z^*)]\| \\ &\leqslant M[\|H(z_k) - H(z^*) - V_k(z_k - z^*)\| + \|H^k(z_k) - H(z_k)\|\end{aligned}$$

$$
\begin{aligned}
&\quad + \|(\nabla H^k(z_k) - V_k)(z_k - z^*)\|] \\
&\leqslant M[\mathrm{o}(\|z_k - z^*\|) + \mu\varepsilon_k + \|\nabla H^k(z_k) - V_k\|\|z_k - z^*\|] \\
&\leqslant \mathrm{o}(\|z_k - z^*\|) + \frac{\mu M}{2}\|H(z_k)\|^2 \\
&= \mathrm{o}(\|z_k - z^*\|),
\end{aligned} \tag{4.3}
$$

where the last equality follows from Lipschitz property of $H(z)$. This together with Assumption 3 implies that when $k \in K_0$ is sufficiently large, it holds that

$$
\|H(z_k + d_k)\| = \mathrm{o}(\|H(z_k)\|). \tag{4.4}
$$

Consequently, we get

$$
\|H^k(z_k + d_k)\| \leqslant \|H(z_k + d_k)\| + \mu\varepsilon_k = \mathrm{o}(\|H(z_k)\|). \tag{4.5}
$$

We also have

$$
\begin{aligned}
\|H^k(z_k)\| &\geqslant \|H(z_k)\| - \|H^k(z_k) - H(z_k)\| \\
&\geqslant \|H(z_k)\| - \mu\varepsilon_k \geqslant (1 - \alpha)\|H(z_k)\|.
\end{aligned} \tag{4.6}
$$

Notice that $Q^k(d_k) = 0$, we deduce from (4.5) and (4.6) that

$$
\begin{aligned}
|1 - r_k| &= \left| \frac{W^k(z_k + d_k) - Q^k(d_k)}{W^k(z_k) - Q_k(d_k)} \right| = \frac{\|H^k(z_k + d_k)\|^2}{\|H^k(z_k)\|^2} \\
&\leqslant \frac{\mathrm{o}(\|H(z_k)\|^2)}{(1-\alpha)^2\|H(z_k)\|^2} \to 0, \quad (k \to \infty,\ k \in K_0).
\end{aligned}
$$

This implies that, for $k \in K_0$ sufficiently large, we have $r_k \geqslant \eta_1$. Hence, d_k can be accepted by Step 2 in Algorithm 2.1. This completes the proof of lemma. □

It is easy to see from the proof of Lemma 4.1 that when $k \in K_0$ is sufficiently large, (4.3) means

$$
\|z_{k+1} - z^*\| = \mathrm{o}(\|z_k - z^*\|) \tag{4.7}
$$

We now establish the following superlinear convergence theorem for Algorithm 2.1.

Theorem 4.1. *Let the conditions of Assumptions* 1–3 *hold. Then the whole sequence* $\{z_k\}$ *generated by Algorithm* 2.1 *converges to* z^* *superlinearly.*

Proof. It follows from the proof of Lemma 4.1 that

$$
\|z_{k+1} - z_k\|_{k\in K_0} = \|d_k\|_{k \in K_0} \to 0, \quad (k \to \infty).
$$

In a way similar to the proof of Theorem 3.2 in [1], we can show that the whole sequence $\{z_k\}$ converges to z^*. Moreover, (4.7) shows that the convergent rate is suplinear. □

We conclude the paper with some remarks on Assumptions 1–3. For some semismooth equations and related smoothing functions arising from the nonlinear complementarity problem and the box constrained variational inequality problem, Chen, Qi and Sun [1] presented sufficient conditions to guarantee Assumptions 1–3. We also refer to a recent book edited by Fukushima and Qi [4] for a comprehensive study on semismooth equations and smoothing methods.

References

[1] X. Chen, L. Qi, D. Sun, Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities, Math. Comp. 67 (1998) 519–540.
[2] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
[3] A. Fischer, A special Newton-type optimization method, Optimization 24 (1992) 269–284.
[4] M. Fukushima, L. Qi (Eds.), Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers, Boston, MA, 1999.
[5] H. Jiang, M. Fukushima, L. Qi, D. Sun, A trust region method for solving generalized complementarity problem, SIAM J. Optim. 8 (1998) 140–157.
[6] C. Kanzow, H. Pieper, Jacobian smoothing method for nonlinear complementarity problems, SIAM J. Optim. 9 (1999) 343–373.
[7] D.H. Li, L. Qi, J. Tam, S.Y. Wu, A smoothing Newton method for semi-infinite programming, Journal of Global Optimization, to appear.
[8] R. Mifflin, Semismooth and semiconvex function in constrained optimization, SIAM J. Control Optim. 15 (1977) 957–972.
[9] J.J. More, D.C. Sorensen, Computing trust region step, SIAM J. Sci. Statist. Comput. 4 (1983) 553–572.
[10] J. Nocedal, Y.X. Yuan, Combining trust region and line search techniques, in: Y. Yuan (Ed.), Advances in Nonlinear Programming, Kluwer Academic Publishers, Boston, MA, 1998, pp. 153–175.
[11] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equation, Math. Oper. Res. 18 (1993) 227–244.
[12] L. Qi, Trust region algorithms for solving nonsmooth equation, SIAM J. Optim. 5 (1995) 219–230.
[13] L. Qi, J. Sun, A nonsmooth version of Newton's method, Math. Programming 58 (1993) 353–367.