# A Trust Region Method for Solving Generalized Complementarity Problems

Houyuan Jiang[a,1], Masao Fukushima[b,2], Liqun Qi[a,1] and Defeng Sun[a,1]

[a] School of Mathematics
The University of New South Wales
Sydney, Australia

[b] Graduate School of Information Science
Nara Institute of Science and Technology
Ikoma, Nara, Japan

November 30, 1995

**Abstract.** Based on a semismooth equation reformulation using Fischer's function, a trust region algorithm is proposed for solving the generalized complementarity problem (GCP). The algorithm uses a generalized Jacobian of the function involved in the semismooth equation and adopts the squared natural residual of the semismooth equation as a merit function. The proposed algorithm is applicable to the nonlinear complementarity problem, because the latter problem is a special case of the GCP. Global convergence and, under a nonsingularity assumption, local Q-superlinear (or quadratic) convergence of the algorithm are established. Moreover, calculation of a generalized Jacobian is discussed and numerical results are presented.

**Key Words.** Generalized complementarity problem, nonlinear complementarity problem, semismooth equation, trust region method, global and superlinear convergence.

**AMS (MOS) subject classifications.** 90C33, 65K10.

## 1. Introduction

The generalized complementarity problem, denoted by GCP, is to find a vector $x \in \Re^n$ such that

$$F(x) \geq 0, \quad G(x) \geq 0, \quad F(x)^T G(x) = 0, \tag{1.1}$$

where $F, G : \Re^n \to \Re^n$ are assumed to be continuously differentiable.

When $G(x) \equiv x - E(x)$, where $E : \Re^n \to \Re^n$, the GCP reduces to the so-called implicit complementarity problem [5, 23, 22]. In particular, if $G(x) \equiv x$, the GCP reduces to the nonlinear complementarity problem, denoted by NCP, which is a general framework for optimality

[1] The research of this author is supported by the Australian Research Council.

[2] The research of this author is supported in part by the Science Research Grant-in-Aid from the Ministry of Education, Science and Culture, Japan.

conditions of mathematical programs as well as variational inequalities. Moreover, if $F$ is an affine function, i.e., $F(x) = Mx + q$ with a matrix $M \in \Re^{n\times n}$ and a vector $q \in \Re^n$, then the NCP reduces to the linear complementarity problem (LCP), which in turn contains linear and quadratic programming problems as special cases.

There are many Newton-based methods for solving the NCP. We do not intend to give a short survey of numerical methods for solving the NCP. The interested reader is referred to two survey papers [14, 24], and [1, 11, 17] for recent progress. In order to enlarge the domain of convergence of Newton-based methods, line search is usually used on certain merit functions derived from different reformulations of the NCP. A trust region strategy is used in [21] for solving a bound-constrained nonlinear least square reformulation of the NCP. Global convergence of this trust region method was established, and under the strict complementarity condition and other conditions, superlinear convergence was also given.

In this paper, we shall propose a trust region method for solving an unconstrained least squares reformulation of the GCP. It is surely applicable to the NCP since the GCP is a generalization of the NCP. The main difference between the proposed trust region method and the classical trust region methods lies in the updating rule for the trust region radius at the beginning of each iteration. More precisely, at the beginning of each iteration, the trust region radius is always set greater than a fixed (small) positive constant, rather than solely updated from the final trust region radius of the last iteration as in the classical trust region methods. This type of updating rule for the trust region radius has been used in recent literature [10, 13, 18] for optimization problems and systems of nonsmooth equations. We will show that the proposed special strategy of updating trust region radii enables us to recover local superlinear convergence under some conditions, in spite of the fact that the functions involved in the system of equations are only semismooth. (For globally and superlinearly convergent trust region methods for systems of smooth equations, see [20].)

The remainder of the paper is organized as follows. In the next section, the GCP will be converted into a system of semismooth equations and an unconstrained differentiable minimization problem. These two reformulations are equivalent to the GCP in a certain sense. In Section 3, a trust region method is proposed for solving the GCP based on these two reformulations. In Section 4, we discuss regularity conditions which ensure that a stationary point of the reformulated unconstrained minimization problem is a solution of the GCP. Section 5 is devoted to proving global convergence of the algorithm. In Section 6, local superlinear convergence of the algorithm will be established under some nonsingularity condition. In Section 7, we present a method for calculating a generalized Jacobian which is needed in the algorithm. Numerical results are presented in Section 8. We conclude our paper by giving some remarks in the final section.

A few words about our notation. The inner product of vectors $x, y \in \Re^n$ is denoted by $x^T y$. Let $\|\cdot\|$, $\|\cdot\|_\infty$ and $\|\cdot\|_2$ denote any fixed norm, the $l_\infty$ norm and the Euclidean norm, respectively, in a space of appropriate dimension. The Jacobian of a vector function $F$ at a point $x$ is denoted by $\nabla F(x)^T$. If $M = (M_{ij})$ is an $n \times n$ matrix and $\alpha, \beta$ are two subsets of $\{1, \ldots, n\}$, then $M_{\alpha\beta}$ denotes the submatrix of $M$ consisting of elements $M_{ij}$ with $i \in \alpha$ and $j \in \beta$.

## 2. Reformulation and Preliminaries

Throughout this paper, we assume that $F$ and $G$ are continuously differentiable on $\Re^n$.

Let $\phi : \Re^2 \to \Re$ be defined by

$$\phi(a, b) := \sqrt{a^2 + b^2} - a - b. \tag{2.1}$$

A basic property of this function is that

$$\phi(a, b) = 0 \Longleftrightarrow a \geq 0, b \geq 0, ab = 0.$$

From this property, the GCP (1.1) can readily be recast as the system of nonsmooth equations

$$H(x) = \begin{pmatrix} H_1(x) \\ \vdots \\ H_n(x) \end{pmatrix} := \begin{pmatrix} \phi(F_1(x), G_1(x)) \\ \vdots \\ \phi(F_n(x), G_n(x)) \end{pmatrix} = 0 \tag{2.2}$$

in the sense that $x$ solves (1.1) if and only if $x$ solves (2.2). The function $\phi$ was first introduced by Fischer [7] (but attributed to Burmeister) and later used to study various methods for solving the NCP and related problems, see a survey paper [8] for details. Recently the function $\phi$ has been used to reformulate the GCP [16]. (See also [30] for other reformulations of the GCP.)

Note that $H$ is locally Lipschitz on $\Re^n$ and Fréchet-differentiable on the set $\Omega$, where

$$\Omega := \{x \in \Re^n \mid F_i(x)^2 + G_i(x)^2 > 0, i = 1, \ldots, n\}.$$

However, $H$ is not necessarily differentiable at $x \notin \Omega$. Let $\partial H(x)$ denote the Clarke's generalized Jacobian of $H$ at $x \in \Re^n$ [2], which can be defined as the convex hull of the set $\partial_B H(x)$ [27], where

$$\partial_B H(x) = \{V \in \Re^{n \times n} \mid V = \lim_{x^k \to x} \nabla H(x^k)^T, H \text{ is differentiable at } x^k \text{ for all } k\}.$$

Similar to the discussions of [6, 15], we have for any $x \in \Re^n$

$$\partial H(x) \subseteq \mathcal{D}^F(x) \nabla F(x)^T + \mathcal{D}^G(x) \nabla G(x)^T, \tag{2.3}$$

where $\mathcal{D}^F(x)$ and $\mathcal{D}^G(x)$ are sets of $n \times n$ diagonal matrices such that, for each pair $(D^F, D^G) = (\text{diag}\,\{D_1^F, \ldots, D_n^F\}, \text{diag}\,\{D_1^G, \ldots, D_n^G\}) \in \mathcal{D}^F(x) \times \mathcal{D}^G(x)$, the following conditions are satisfied

$$(D_i^F + 1)^2 + (D_i^G + 1)^2 \leq 1, \quad i = 1, \ldots, n. \tag{2.4}$$

In particular, if $F_i(x)^2 + G_i(x)^2 > 0$ for all $i = 1, \ldots, n$, then we have $\mathcal{D}^F(x) = \{D^F\}$ and $\mathcal{D}^G(x) = \{D^G\}$ with

$$D_i^F = \frac{F_i(x)}{\sqrt{F_i(x)^2 + G_i(x)^2}} - 1, \quad D_i^G = \frac{G_i(x)}{\sqrt{F_i(x)^2 + G_i(x)^2}} - 1, \quad i = 1, \ldots, n, \tag{2.5}$$

and $\nabla H(x)$ exists and is given by

$$\nabla H(x)^T = D^F \nabla F(x)^T + D^G \nabla G(x)^T.$$

Define a merit function $\Phi : \Re^n \to \Re$ as

$$\Phi(x) := \frac{1}{2}\sum_{i=1}^{n} \phi(F_i(x), G_i(x))^2 = \frac{1}{2}\|H(x)\|_2^2.$$

When the GCP has a solution, solving (1.1) is equivalent to finding a global minimum point of the unconstrained minimization problem

$$\min_{x \in \Re^n} \Phi(x). \tag{2.6}$$

A favourable property of $\Phi$ is that it is continuously differentiable on the whole space $\Re^n$, although $H$ itself is not continuously differentiable in general [16]. It is easy to verify that for any $x \in \Re^n$ and any $V \in \partial H(x)$

$$\nabla\Phi(x) = V^T H(x) = \nabla F(x) D^F H(x) + \nabla G(x) D^G H(x), \tag{2.7}$$

where $D^F$ and $D^G$ are diagonal matrices satisfying (2.4) such that $V = D^F \nabla F(x)^T + D^G \nabla G(x)^T$. The unconstrained minimization reformulation (2.6) as well as the semismooth equation reformulation (2.2) will be the basis of the proposed trust region method for solving the GCP.

We now introduce some definitions which will be useful in the subsequent analysis. Let $E : \Re^n \to \Re^n$ be locally Lipschitz around $x \in \Re^n$. $E$ is said to be semismooth at $x$ if

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

exists for any $h \in \Re^n$ [28]. It has been proved [28] that $E'(x;h)$, the directional derivative of $E$ at $x$ along the direction $h$, exists for any $h \in \Re^n$ if $E$ is semismooth at $x$. Moreover, $E$ is said to be strongly semismooth at $x$ if $E$ is semismooth at $x$ and for any $V \in \partial E(x+h)$, $h \to 0$,

$$Vh - E'(x;h) = O(\|h\|^2).$$

A matrix $M \in \Re^{n \times n}$ is called

a $P$-matrix if its principal minors are all positive;

a $P_0$-matrix if its principal minors are all nonnegative;

an $S_0$-matrix if

$$\{x \in \Re^n \mid x \geq 0, x \neq 0, Mx \geq 0\} \neq \emptyset.$$

It is known [3] that every $P$-matrix is a $P_0$-matrix and every $P_0$-matrix is an $S_0$-matrix.

## 3. Algorithm

The proposed trust region algorithm for solving the GCP (1.1) is now formally stated.

**Algorithm**:

Step 0. Let $\alpha_1$, $\alpha_2$, $\alpha_3$, $\alpha_4$, $\rho_1$, $\rho_2$, $\Delta_{\min}$ and $\Delta_1$ be such that $0 < \alpha_1 < \alpha_2 < 1 < \alpha_3 < \alpha_4$, $0 < \rho_1 < \rho_2 < 1$, $\Delta_{\min} > 0$ and $\Delta_1 > 0$. Let $x^1 \in \Re^n$ be a starting point. Set $k := 1$.

Step 1. If $\nabla\Phi(x^k) = 0$, stop. Otherwise, let $\hat{\Delta} := \max\{\Delta_{\min}, \Delta_k\}$ and choose $V_k \in \partial_B H(x^k)$.

Step 2. Let $\hat{s}$ be a solution of the minimization problem

$$\begin{aligned} \min \quad & \nabla\Phi(x^k)^T s + \frac{1}{2} s^T V_k^T V_k s \\ \text{s.t.} \quad & \|s\| \le \hat{\Delta}. \end{aligned} \tag{3.1}$$

Step 3. Let

$$\hat{r} := \frac{\Phi(x^k + \hat{s}) - \Phi(x^k)}{\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k)}.$$

If $\hat{r} \ge \rho_1$, then let $s^k := \hat{s}$, $x^{k+1} := x^k + s^k$, $\delta_k := \hat{\Delta}$. For a fixed $\alpha \in [\alpha_3, \alpha_4]$, let

$$\Delta_k := \begin{cases} \hat{\Delta} & \text{if} \quad \rho_1 \le \hat{r} < \rho_2, \\ \alpha\hat{\Delta} & \text{if} \quad \hat{r} \ge \rho_2. \end{cases}$$

Let $k$ be replaced by $k+1$ and return to Step 1. Otherwise, choose $\Delta \in [\alpha_1\hat{\Delta}, \alpha_2\hat{\Delta})$. Let $\hat{\Delta} := \Delta$ and repeat Step 2.

**Remarks**. (i) In Step 2 of the algorithm, the minimization problem (3.1) may have several solutions. From the first equality in (2.7), it is easy to see that, for any $s \in \Re^n$

$$\frac{1}{2}\|H(x^k) + V_k s\|_2^2 - \Phi(x^k) = \nabla\Phi(x^k)^T s + \frac{1}{2} s^T V_k^T V_k s. \tag{3.2}$$

(ii) The main difference between the proposed trust region method and the classical trust region methods lies in the updating rule for the trust region radius at the beginning of each iteration. More precisely, at the beginning of each iteration $k$, the trust region radius is always set greater than the fixed positive constant $\Delta_{\min}$, rather than solely updated from the final trust region radius of iteration $k-1$ as in the classical trust region methods. This type of updating rule for the trust region radius is also used in recent literature [10, 13, 18]. As we shall see, this special strategy enables us not only to establish global convergence but also to recover local superlinear convergence of the algorithm under some conditions, despite that (2.2) is a system of nonsmooth equations.

(iii) Note that $\delta_k$ does not play any role in the algorithm. But it will be useful in proving convergence theorems in the subsequent sections.

**Proposition 3.1** *The algorithm cannot cycle between Step 2 and Step 3 infinitely, provided that $x^k$ is not a stationary point of $\Phi$.*

**Proof**. Suppose the algorithm does cycle for some $k$ and $x^k$ is not a stationary point of $\Phi$, i.e., $\nabla\Phi(x^k) \ne 0$. Then the continuous differentiability of $\Phi$ implies

$$\begin{aligned} \Phi(x^k + \hat{s}) - \Phi(x^k) &= \nabla\Phi(x^k)^T\hat{s} + o(\|\hat{s}\|) \\ &\le \nabla\Phi(x^k)^T\hat{s} + \frac{1}{2}\hat{s}^T V_k^T V_k\hat{s} + o(\|\hat{s}\|) \\ &= \frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k) + o(\|\hat{s}\|), \end{aligned} \tag{3.3}$$

where the last equality follows from (3.2). Let $s^* = -\frac{\nabla\Phi(x^k)}{\|\nabla\Phi(x^k)\|}$. Clearly, $\tau\hat{\Delta}s^*$ is a feasible solution of (3.1) for any $\tau \in [0,1]$. Since $\hat{s}$ solves (3.1), it follows from (3.2) that for $0 \le \tau \le 1$

$$\begin{aligned}\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k) &\le \tau\hat{\Delta}\nabla\Phi(x^k)^T s^* + \frac{1}{2}\tau^2\hat{\Delta}^2(s^*)^T V_k^T V_k s^* \\ &= -\tau\hat{\Delta}\|\nabla\Phi(x^k)\| + \frac{1}{2}\tau^2\hat{\Delta}^2(s^*)^T V_k^T V_k s^*.\end{aligned}$$

By the assumption that the algorithm cycles between Steps 2 and 3, both $\|\hat{s}\|$ and $\hat{\Delta}$ tend to zero, while $\hat{r} < \rho_1$ is maintained. Consequently, there exist $\tau^* > 0$ and $\Delta^* > 0$ such that when $\hat{\Delta} \le \Delta^*$

$$\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k) \le -\tau^*\hat{\Delta}\|\nabla\Phi(x^k)\|.$$

This, together with the definition of $\hat{r}$ and the relation (3.3), implies that

$$\hat{r} \ge 1 - \frac{o(\|\hat{s}\|)}{\tau^*\hat{\Delta}\|\nabla\Phi(x^k)\|} \ge 1 - \frac{o(\|\hat{\Delta}\|)}{\tau^*\hat{\Delta}\|\nabla\Phi(x^k)\|} \to 1 \qquad \text{as} \quad \hat{\Delta} \to 0,$$

which indicates that we eventually have $\hat{r} \ge \rho_1$. This is a contradiction and the desired result follows. □

# 4. Regularity Condition

Most unconstrained minimization methods normally generate a sequence converging to a local minimizer or a stationary point rather than a global minimizer. As discussed in Section 2, if the GCP has a solution, then solving the GCP is equivalent to finding a global minimizer of the unconstrained minimization problem (2.6). It is therefore crucial to study under what conditions a stationary point of (2.6) becomes a a solution of the GCP. To this end, let us introduce some index sets: $\mathcal{C}(x), \mathcal{R}(x), \mathcal{P}(x)$ and $\mathcal{N}(x)$, which stand for Complementarity indices, Residual indices, Positive indices and Negative indices, respectively, i.e.,

$$\begin{aligned}\mathcal{C}(x) &:= \{i \in \{1,\ldots,n\} \mid F_i(x) \ge 0,\ G_i(x) \ge 0,\ F_i(x)G_i(x) = 0\},\\ \mathcal{R}(x) &:= \{i \in \{1,\ldots,n\} \mid i \notin \mathcal{C}(x)\},\\ \mathcal{P}(x) &:= \{i \in \mathcal{R}(x) \mid F_i(x) > 0,\ G_i(x) > 0\},\\ \mathcal{N}(x) &:= \mathcal{R}(x) \setminus \mathcal{P}(x).\end{aligned}$$

In the sequel, we shall simply denote these sets as $\mathcal{C}, \mathcal{R}, \mathcal{P}$ and $\mathcal{N}$; the point $x$ under consideration will always be clear from the context.

It is easy to verify the following result using (2.3), (2.4) and (2.5). The proof is omitted.

**Proposition 4.1** *Let diagonal matrices $D^F$ and $D^G$ satisfy (2.4) such that $D^F\nabla F(x)^T + D^G\nabla G(x)^T \in \partial H(x)$. Then for each $x \in \Re^n$, we have the following relations:*

$$\begin{aligned}(D^F H(x))_i > 0 &\iff (D^G H(x))_i > 0 \iff i \in \mathcal{P}\\ (D^F H(x))_i = 0 &\iff (D^G H(x))_i = 0 \iff i \in \mathcal{C}\\ (D^F H(x))_i < 0 &\iff (D^G H(x))_i < 0 \iff i \in \mathcal{N}.\end{aligned}$$

Note that, since $D^F$ and $D^G$ are diagonal matrices, we have $(D^F H(x))_i = D_i^F H_i(x)$ and $(D^G H(x))_i = D_i^G H_i(x)$ for each $i = 1, \ldots, n$.

**Definition 4.1** *The GCP is said to be regular at a point $x$ if for any two vectors $z^1 \neq 0$, $z^2 \neq 0$ in $\Re^n$ satisfying*

$$\begin{array}{lll} z^1_{\mathcal{C}} = 0, & z^1_{\mathcal{P}} > 0, & z^1_{\mathcal{N}} < 0, \\ z^2_{\mathcal{C}} = 0, & z^2_{\mathcal{P}} > 0, & z^2_{\mathcal{N}} < 0, \end{array}$$

*there exists a vector $y \in \Re^n$ such that*

$$y^T(\nabla F(x) z^1 + \nabla G(x) z^2) \neq 0.$$

*Moreover, a stationary point $x$ of $\Phi$ is called a regular stationary point if the GCP is regular at $x$.*

It is not hard to see that, when the GCP reduces to the NCP, the regularity condition introduced here is slightly weaker than that defined in [4]. Under the present regularity condition, the following theorem establishes the equivalence between solutions of the GCP and stationary points of the merit function $\Phi$.

**Theorem 4.1** *$x$ is a solution of the GCP if and only if $x$ is a regular stationary point of $\Phi$.*

**Proof**. If $x$ is a solution of the GCP, then $x$ is a stationary point of $\Phi$, and hence, $\mathcal{P} = \mathcal{N} = \emptyset$. By definition, $x$ is a regular point of $\Phi$.

Conversely, suppose $x$ is a regular stationary point of $\Phi$. Let $D^F$ and $D^G$ satisfy the conditions of Proposition 4.1. Let $z^1 = D^F H(x)$ and $z^2 = D^G H(x)$. Then

$$0 = \nabla\Phi(x) = \nabla F(x) D^F H(x) + \nabla G(x) D^G H(x) = \nabla F(x) z^1 + \nabla G(x) z^2. \tag{4.1}$$

If $x$ is not a solution of the GCP, it follows from Proposition 4.1 that $z^1 \neq 0$, $z^2 \neq 0$. By the definition of regularity and Proposition 4.1, for $z^1$ and $z^2$, there exists a vector $y \in \Re^n$ such that

$$y^T(\nabla F(x) z^1 + \nabla G(x) z^2) \neq 0,$$

which contradicts (4.1). Therefore $x$ is a solution of the GCP. □

Next we present a sufficient condition for ensuring regularity of the GCP. The proof is based on a result of [4] for the NCP. Moreover, when the GCP happens to be the NCP, we reobtain the result in [4].

Let $D \in \Re^{|\mathcal{R}| \times |\mathcal{R}|}$ denote a diagonal matrix with diagonal elements $D_1, \ldots, D_{|\mathcal{R}|}$ defined by

$$D_i = \begin{cases} 1 & \text{if } i \in \mathcal{P} \\ -1 & \text{if } i \in \mathcal{N}. \end{cases}$$

Evidently, $DD$ is the $|\mathcal{R}| \times |\mathcal{R}|$ identity matrix. Using this notation, we establish the following proposition.

**Proposition 4.2** *Assume that $\nabla G(x)$ is invertible and $D(\nabla F(x)^T(\nabla G(x)^T)^{-1})_{\mathcal{RR}}D$ is an $S_0$-matrix. Then the GCP is regular at $x$.*

**Proof**. Since $D(\nabla F(x)^T(\nabla G(x)^T)^{-1})_{\mathcal{RR}}D$ is an $S_0$-matrix, there exists a vector $y^1_{\mathcal{R}} \neq 0$ such that

$$y^1_{\mathcal{R}} \geq 0, \quad D(\nabla F(x)^T(\nabla G(x)^T)^{-1})_{\mathcal{RR}}Dy^1_{\mathcal{R}} \geq 0. \tag{4.2}$$

Let vectors $y^2 \in \Re^n$ and $y \in \Re^n$ be such that

$$y^2_{\mathcal{C}} = 0, \quad y^2_{\mathcal{R}} = Dy^1_{\mathcal{R}}, \tag{4.3}$$

$$y = (\nabla G(x)^{-1})^T y^2. \tag{4.4}$$

Clearly,

$$y^2_{\mathcal{R}} \neq 0, \quad y^2_{\mathcal{P}} \geq 0, \quad y^2_{\mathcal{N}} \leq 0. \tag{4.5}$$

For any two vectors $z^1 \neq 0$, $z^2 \neq 0$ in $\Re^n$ satisfying

$$\begin{array}{lll} z^1_{\mathcal{C}} = 0, & z^1_{\mathcal{P}} > 0, & z^1_{\mathcal{N}} < 0, \\ z^2_{\mathcal{C}} = 0, & z^2_{\mathcal{P}} > 0, & z^2_{\mathcal{N}} < 0, \end{array}$$

it follows from (4.3), (4.4) and (4.5) that

$$y^T\nabla G(x)z^2 = (y^2)^T z^2 = (y^2_{\mathcal{R}})^T z^2_{\mathcal{R}} > 0. \tag{4.6}$$

By the definition of $y$, $y^2$ and $z^1$, we have

$$\begin{array}{rcl} y^T\nabla F(x)z^1 & = & (y^2)^T\nabla G(x)^{-1}\nabla F(x)z^1 \\ & = & (y^2_{\mathcal{R}})^T(\nabla G(x)^{-1}\nabla F(x))_{\mathcal{RR}}z^1_{\mathcal{R}} \\ & = & (y^1_{\mathcal{R}})^T D(\nabla G(x)^{-1}\nabla F(x))_{\mathcal{RR}}DDz^1_{\mathcal{R}}. \end{array}$$

Then (4.2) and the fact that $Dz^1_{\mathcal{R}} > 0$ imply that

$$y^T\nabla F(x)z^1 \geq 0,$$

which together with (4.6) yields

$$y^T(\nabla F(x)z^1 + \nabla G(x)z^2) > 0.$$

Therefore the GCP is regular at $x$. □

It has been proved in [16] that any stationary point of $\Phi$ is a solution of the GCP if $\nabla G(x)^T$ is invertible and $\nabla F(x)^T(\nabla G(x)^T)^{-1}$ is a $P_0$-matrix. It is not difficult to see that this result is a consequence of Proposition 4.2 by using the fact that a matrix $M$ is a $P_0$-matrix if and only if the matrix $DMD$ is a $P_0$-matrix for any nonsingular diagonal matrix $D$, and the fact that any $P_0$-matrix is an $S_0$-matrix.

## 5. Global Convergence

We now suppose that the algorithm generates an infinite sequence $\{x^k\}$, i.e., the stopping test in Step 1 of the algorithm is never fulfilled. Let $c > 0$ be a constant such that

$$\|y\| \leq c\|y\|_2$$

for all $y \in \Re^n$. We first present a standard result.

**Lemma 5.1** *Let $\hat{s}$ be a solution of (3.1). Then*

$$\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k) \leq -\frac{1}{2}\|\nabla\Phi(x^k)\|_2 \min\left\{\frac{\hat{\Delta}}{c}, \frac{\|\nabla\Phi(x^k)\|_2}{\|V_k^T V_k\|_2}\right\}.$$

**Proof.** Suppose $\tilde{s}$ is a solution of the following minimization problem

$$\begin{aligned} \min \quad & \nabla\Phi(x^k)^T s + \tfrac{1}{2} s^T V_k^T V_k s \\ \text{s.t.} \quad & \|s\|_2 \leq \frac{\hat{\Delta}}{c}. \end{aligned}$$

Then from Theorem 4 in [26],

$$\frac{1}{2}\|H(x^k) + V_k\tilde{s}\|_2^2 - \Phi(x^k) \leq -\frac{1}{2}\|\nabla\Phi(x^k)\|_2 \min\left\{\frac{\hat{\Delta}}{c}, \frac{\|\nabla\Phi(x^k)\|_2}{\|V_k^T V_k\|_2}\right\}. \tag{5.1}$$

Since $\|\tilde{s}\| \leq c\|\tilde{s}\|_2 \leq \hat{\Delta}$, $\tilde{s}$ is a feasible solution of (3.1). Since $\hat{s}$ solves (3.1), the desired result follows from (3.2) and (5.1). □

**Lemma 5.2** *Suppose $x^*$ is the limit of a subsequence $\{x^k\}_{k\in K}$. If $x^*$ is not a stationary point of $\Phi$, then there exist $\hat{k}$ and $\Delta > 0$ such that for all $k \geq \hat{k}$ $(k \in K)$*

$$\hat{r} := \frac{\Phi(x^k + \hat{s}) - \Phi(x^k)}{\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k)} \geq \rho_1,$$

*whenever $\hat{\Delta} \in (0, \Delta)$, where $\hat{s}$ is a solution of (3.1).*

**Proof**. First note that

$$\hat{r} = 1 + \frac{-\frac{1}{2}\hat{s}^T V_k^T V_k \hat{s} + o(\|\hat{s}\|_2)}{\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k)}. \tag{5.2}$$

Since $x^*$ is not a stationary point and $\partial H$ is upper semicontinuous, there exist positive constants $\beta_1$ and $\beta_2$ such that

$$\|\nabla\Phi(x^k)\|_2 \geq \beta_1, \quad \frac{\|\nabla\Phi(x^k)\|_2}{\|V_k^T V_k\|_2} \geq \beta_2 \tag{5.3}$$

for all sufficiently large $k \in K$. By Lemma 5.1,

$$\frac{1}{2}\|H(x^k) + V_k\hat{s}\|_2^2 - \Phi(x^k) \leq -\frac{1}{2}\beta_1 \min\left\{\frac{\hat{\Delta}}{c}, \beta_2\right\} \tag{5.4}$$

for all sufficiently large $k \in K$. Then the desired result is a direct consequence of (5.2), (5.3) and (5.4). □

**Lemma 5.3** *Suppose $x^*$ is the limit of a subsequence $\{x^k\}_{k\in K}$. If $x^*$ is not a stationary point of $\Phi$, then*

$$\liminf_{k\in K,\ k\to\infty} \delta_k > 0,$$

*where $\delta_k$ is defined as in the algorithm.*

**Proof**. By Lemma 5.2, there exists a $\Delta > 0$ such that $\hat{r} \geq \rho_1$ whenever $\hat{\Delta} < \Delta$, at each iteration $k \in K$ sufficiently large. Thus, by the updating rule of the trust region radius in the algorithm, we have $\delta_k \geq \alpha_1\Delta$ for all sufficiently large $k \in K$. The desired result follows. □

Now we are ready to establish a global convergence theorem for the proposed trust region algorithm.

**Theorem 5.1** *Let $\{x^k\}$ be generated by the algorithm. Then any accumulation point of $\{x^k\}$ is a stationary point of $\Phi$. Moreover, it is a solution of the GCP if the regularity condition holds at this point.*

**Proof**. Suppose $x^*$ is an accumulation point of $\{x^k\}$, say $\lim_{k\in K,\ k\to\infty} x^k = x^*$. If $x^*$ is not a stationary point of $\Phi$, then Lemmas 5.1 and 5.3 imply that there exist $\beta_1 > 0$, $\beta_2 > 0$, $\delta > 0$ and $\hat{k}$ such that (5.4) holds and $\delta_k \geq \delta$ for all $k \geq \hat{k}$ $(k \in K)$. By the algorithm, Lemma 5.1, (5.4) and the fact that $\Phi$ is nonnegative, we obtain

$$\begin{aligned}
\Phi(x^1) &\geq \sum_{k=1}^{\infty} \left[\Phi(x^k) - \Phi(x^{k+1})\right] \\
&\geq \sum_{k=1}^{\infty} \rho_1 \left[\Phi(x^k) - \frac{1}{2}\|H(x^k) + V_k s^k\|_2^2\right] \\
&\geq \sum_{\substack{k\geq\hat{k}\\ k\in K}} \rho_1 \left[\Phi(x^k) - \frac{1}{2}\|H(x^k) + V_k s^k\|_2^2\right] \\
&\geq \rho_1 \sum_{\substack{k\geq\hat{k}\\ k\in K}} \frac{1}{2}\beta_1 \min\left\{\frac{\Delta_k}{c}, \beta_2\right\} \\
&\geq \rho_1 \sum_{\substack{k\geq\hat{k}\\ k\in K}} \frac{1}{2}\beta_1 \min\left\{\frac{\delta_k}{c}, \beta_2\right\} \\
&\geq \frac{1}{2}\rho_1\beta_1 \sum_{\substack{k\geq\hat{k}\\ k\in K}} \min\left\{\frac{\delta}{c}, \beta_2\right\} \;=\; \infty.
\end{aligned}$$

This is impossible. Therefore, $x^*$ is a stationary point of $\Phi$. □

We now turn to the case where $\{x^k\}$ does not necessarily have an accumulation point.

**Theorem 5.2** *Let $\{x^k\}$ be generated by the algorithm. If $\{V_k\}$ is bounded, then $\{\nabla\Phi(x^k)\}$ is not bounded away from zero, that is,*

$$\liminf_{k\to\infty} \|\nabla\Phi(x^k)\| = 0. \tag{5.5}$$

**Proof**. Suppose (5.5) does not hold. Then the boundedness of $\{V_k\}$ and Lemma 5.1 imply the existence of $\beta_1$ and $\beta_2$ such that (5.4) holds for all $k$. By the algorithm and (5.4), we obtain

$$\begin{aligned}
\Phi(x^1) &\geq \sum_{k=1}^{\infty} \left[\Phi(x^k) - \Phi(x^{k+1})\right] \\
&\geq \sum_{k=1}^{\infty} \rho_1 \left[\Phi(x^k) - \frac{1}{2}\|H(x^k) + V_k s^k\|_2^2\right] \\
&\geq \rho_1 \sum_{k=1}^{\infty} \frac{1}{2}\beta_1 \min\left\{\frac{\Delta_k}{c}, \beta_2\right\} \\
&\geq \frac{1}{2}\rho_1\beta_1 \sum_{k=1}^{\infty} \min\left\{\frac{\delta_k}{c}, \beta_2\right\}.
\end{aligned}$$

This implies

$$\sum_{k=1}^{\infty} \delta_k < \infty,$$

and hence,

$$\sum_{k=1}^{\infty} \|x^{k+1} - x^k\| < \infty.$$

This implies that $\{x^k\}$ converges to a point $x^*$. By Theorem 5.1, $x^*$ is a stationary point of $\Phi$, i.e, $\nabla\Phi(x^*) = 0$. This contradicts the assumption that $\liminf_{k\to\infty} \|\nabla\Phi(x^k)\| > 0$. □

**Remark**. If $\nabla F$ and $\nabla G$ are bounded on $\Re^n$, then $\partial H$ is bounded on $\Re^n$, and the boundedness assumption on $\{V_k\}$ is satisfied.

This theorem says that the sequence $\{x^k\}$ generated by the algorithm contains a stationary subsequence $\{x^k\}_{k\in K}$ in the sense that $\lim_{k\to\infty, k\in K} \|\nabla\Phi(x^k)\| = 0$, even if $\{x^k\}$ is unbounded. However, a stationary sequence is not necessarily a minimizing sequence of $\Phi$ in general. Conditions under which any stationary sequence of $\Phi$ is a minimizing sequence have been studied in [12] for the NCP.

## 6. Superlinear Convergence

In this section, we shall be concerned with the rate of convergence of the algorithm. It is known [28, 27] that semismoothness and certain nonsingularity conditions at a solution of the system of nonsmooth equations play a crucial role in establishing superlinear convergence of some generalized Newton methods. Recall that $H$ is said to be $BD$-regular at a point $x$ if all the elements in $\partial_B H(x)$ are nonsingular [27]. We now derive some sufficient conditions for ensuring $BD$-regularity at a solution of the GCP.

Suppose $\nabla G(\bar{x})^T$ is invertible at a solution $\bar{x}$ of the GCP. Let

$$M = \begin{pmatrix} M_{\alpha\alpha} & M_{\alpha\beta} & M_{\alpha\gamma} \\ M_{\beta\alpha} & M_{\beta\beta} & M_{\beta\gamma} \\ M_{\gamma\alpha} & M_{\gamma\beta} & M_{\gamma\gamma} \end{pmatrix} := \nabla F(\bar{x})^T (\nabla G(\bar{x})^T)^{-1},$$

where

$$\begin{aligned}\alpha &:= \{i \in \{1,\ldots,n\} \mid F_i(\bar{x}) = 0, G_i(\bar{x}) > 0\},\\ \beta &:= \{i \in \{1,\ldots,n\} \mid F_i(\bar{x}) = 0, G_i(\bar{x}) = 0\},\\ \gamma &:= \{i \in \{1,\ldots,n\} \mid F_i(\bar{x}) > 0, G_i(\bar{x}) = 0\}.\end{aligned}$$

The GCP is said to be $R$-regular at $\bar{x}$ if $M_{\alpha\alpha}$ is nonsingular and the Schur-complement of $M_{\alpha\alpha}$ in the matrix

$$\begin{pmatrix} M_{\alpha\alpha} & M_{\alpha\beta} \\ M_{\beta\alpha} & M_{\beta\beta} \end{pmatrix},$$

i.e., $M_{\beta\beta} - M_{\beta\alpha}M_{\alpha\alpha}^{-1}M_{\alpha\beta}$, is a $P$-matrix, see [29].

**Proposition 6.1** *Let $\bar{x}$ be a solution of the GCP. Suppose $\nabla G(\bar{x})^T$ is invertible. If the GCP is $R$-regular at $\bar{x}$, then any $V \in \partial H(\bar{x})$ is nonsingular.*

**Proof**. Since $\nabla G(\bar{x})^T$ is invertible, by (2.3), any $V \in \partial H(\bar{x})$ can be represented as

$$\begin{aligned}V &= D^F \nabla F(\bar{x})^T + D^G \nabla G(\bar{x})^T\\ &= \left(D^F \nabla F(\bar{x})^T (\nabla G(\bar{x})^T)^{-1} + D^G\right) \nabla G(\bar{x})^T\\ &= (D^F M + D^G)\nabla G(\bar{x})^T,\end{aligned}$$

where $D^F$ and $D^G$ satisfy the condition (2.4) and $M = \nabla F(\bar{x})^T(\nabla G(\bar{x})^T)^{-1}$. Then $V$ is nonsingular if and only if $D^F M + D^G$ is nonsingular. By some standard analysis (see, e.g., [6]), the nonsingularity of $D^F M + D^G$ can be deduced from the fact that the GCP is $R$-regular at $x$. □

**Corollary 6.1** *Suppose $\nabla G(\bar{x})^T$ is invertible. If $\nabla F(\bar{x})^T(\nabla G(\bar{x})^T)^{-1}$ is a $P$-matrix, then any $V \in \partial H(\bar{x})$ is nonsingular.*

**Proof**. Note that every principal submatrix of a $P$-matrix is a $P$-matrix, hence nonsingular, and the Schur-complement of every principal submatrix of a $P$-matrix is a $P$-matrix. Therefore, the GCP is $R$-regular at $\bar{x}$. The desired result follows from Proposition 6.1. □

**Lemma 6.1 (a)** *$H$ is semismooth on $\Re^n$ if $F$ and $G$ are continuously differentiable on $\Re^n$;*
**(b)** *$H$ is strongly semismooth on $\Re^n$ if $\nabla F$ and $\nabla G$ are Lipschitz continuous on $\Re^n$.*

**Proof**. Apparently, $H$ is (strongly) semismooth at $x$ if and only if each component of $H$ is (strongly) semismooth at $x$. Therefore it suffices to prove that $H_1$ is (strongly) semismooth at any $x \in \Re^n$ under different assumptions.

Note that $H_1$ can be regarded as a composition of the function $\phi : \Re^2 \to \Re$ defined by (2.1) and the function $h : \Re^n \to \Re^2$ with $h(x) = (F_1(x), G_1(x))$, i.e., $H_1(x) = \phi(h(x))$. It is known that the composition of semismooth functions are semismooth [19, Theorem 5], and the composition of strongly semismooth functions are strongly semismooth [9, Theorem 5.7]. By Lemma 5.6 in [9], $\phi$ is strongly semismooth, hence semismooth on $\Re^2$. Therefore (a) holds if $F$ and $G$ are semismooth on $\Re^n$, and (b) holds if $F$ and $G$ are strongly semismooth on $\Re^n$. On the

other hand, the semismoothness of $F$ and $G$ follows from the continuous differentiability of $F$ and $G$, and the strong semismoothness of $F$ and $G$ follows from the Lipschitz continuity of the Jacobians of $F$ and $G$. □

**Lemma 6.2** *Let $\{x^k\}$ be generated by the algorithm and $x^*$ be an accumulation point of $\{x^k\}$. Suppose $x^*$ is a solution of the GCP. If the BD-regularity condition holds at $x^*$, then there exists an open neighbourhood $\mathcal{N}(x^*)$ of $x^*$ such that when $x^k \in \mathcal{N}(x^*)$, we have*

$$x^{k+1} = x^k - V_k^{-1} H(x^k) \in \mathcal{N}(x^*).$$

**Proof**. By the $BD$-regularity condition and Theorem 5.1, $V_k^{-1} H(x^k) \to 0$ as $x^k \to x^*$. Therefore, when $x^k$ is sufficiently close to $x^*$, we have $\|V_k^{-1} H(x^k)\| \leq \Delta_{\min}$ and $V_k^{-1} H(x^k)$ is a solution of (3.1) if $\hat{\Delta} \geq \Delta_{\min}$. On the other hand, we have $\|H(x^k - V_k^{-1} H(x^k))\| = o(\|H(x^k)\|)$ as $x^k \to x^*$ by Theorem 3.1 in [27]. This implies that $\hat{r} \geq \rho_1$ holds for some $\hat{\Delta} \geq \Delta_{\min}$, when $x^k$ is sufficiently close to $x^*$. This in turn implies that, by the updating rule of the trust region radius, $\Delta_k \geq \Delta_{\min}$ for all $k$ large enough. Therefore we have $x^{k+1} = x^k - V_k^{-1} H(x^k)$ for all $k$ sufficiently large, and $\|x^{k+1} - x^*\| = o(\|x^k - x^*\|)$ as $x^k \to x^*$ from Theorem 3.1 in [27]. The conclusion follows easily. □

We are now in a position to present rate of convergence results for the proposed algorithm.

**Theorem 6.1** *Suppose that the BD-regularity condition holds at an accumulation point $x^*$ of the sequence $\{x^k\}$ generated by the algorithm. Then the entire sequence $\{x^k\}$ converges to $x^*$ Q-superlinearly if $F$ and $G$ are continuously differentiable on $\Re^n$. Moreover, the rate of convergence is Q-quadratic if $\nabla F$ and $\nabla G$ are Lipschitz continuous on $\Re^n$.*

**Proof**. Since $x^*$ is an accumulation point of $\{x^k\}$, there exists a subsequence $\{x^k\}_{k \in K}$ such that $\lim_{k \in K, k \to \infty} x^k = x^*$. By Lemma 6.2, there exists $\hat{k} \in K$ such that when $k \geq \hat{k}$,

$$x^{k+1} = x^k - V_k^{-1} H(x^k).$$

It follows that the algorithm reduces to the generalized Newton method considered in [27]. Therefore, Q-superlinear convergence is guaranteed by Theorem 3.1 in [27]. Note that $H$ is strongly semismooth if $\nabla F$ and $\nabla G$ are Lipschitz continuous. This implies Q-quadratic convergence again by Theorem 3.1 in [27]. □

## 7. Computation of a Generalized Jacobian

In the algorithm, we assumed that a generalized Jacobian of $H$ is available at any point $x$. We now present a method for calculating a generalized Jacobian of $H$ at $x$.

Define

$$I(x) := \{i \in \{1, \ldots, n\} \mid F_i(x)^2 + G_i(x)^2 = 0,\ \|\nabla F_i(x)\| + \|\nabla G_i(x)\| > 0\}.$$

If $I(x) \neq \emptyset$, then we may assume, without loss of generality, that $I(x) = \{1, \ldots, k\}$ for some $k \leq n$ and that $\nabla F_i(x) \neq 0$ for each $i \in I(x)$. Now we shall construct a vector $d \in \Re^n$ such that

$$\nabla F_i(x)^T d \neq 0 \quad \text{for } i \in I(x). \tag{7.1}$$

First let $d := \nabla F_1(x)$ and

$$J(x) := \{i \in I(x) \mid \nabla F_i(x)^T d = 0\}.$$

If $J(x) = \emptyset$, then $d$ satisfies (7.1). Otherwise, choose $j \in J(x)$, let

$$\bar{d} := \frac{\min_{i \in I(x) \setminus J(x)} |\nabla F_i(x)^T d|}{2 \max_{i \in J(x)} \|\nabla F_i(x)\| \|\nabla F_j(x)\|} \nabla F_j(x),$$

and put $\hat{d} := d + \bar{d}$. Then it is clear that $\nabla F_j(x)^T \hat{d} \neq 0$. Moreover, for $i \in I(x) \setminus J(x)$, we have

$$\nabla F_i(x)^T \hat{d} = \nabla F_i(x)^T d + \frac{\min_{i \in I(x) \setminus J(x)} |\nabla F_i(x)^T d|}{2 \max_{i \in J(x)} \|\nabla F_i(x)\| \|\nabla F_j(x)\|} \nabla F_i(x)^T \nabla F_j(x).$$

Since the absolute value of the second term on the right-hand side is never greater than that of the first term, we have $\nabla F_i(x)^T \hat{d} \neq 0$ for $i \in I(x) \setminus J(x)$. Define $\hat{J}(x)$ by

$$\hat{J}(x) := \{i \in I(x) \mid \nabla F_i(x)^T \hat{d} = 0\}.$$

The above arguments indicate that $\hat{J}(x) \subseteq J(x) \setminus \{j\}$. Put $d := \hat{d}$ and $J(x) := \hat{J}(x)$. If $J(x) = \emptyset$, then $d$ satisfies (7.1). Otherwise, choose an index from $J(x)$ and repeat the above procedure. After at most $k$ steps, we will have $J(x) = \emptyset$ and a vector $d$ satisfying (7.1).

For small $t > 0$, let

$$y(t) := x + td + \hat{d}(t),$$

where $\|\hat{d}(t)\| = o(t)$ and $d$ is a vector satisfying (7.1). Then $H$ is differentiable at $y(t)$ with appropriately chosen $\hat{d}(t)$. Letting $t$ tend to zero, we obtain a matrix $V := \lim_{t \to 0^+} \nabla H(y(t))^T$, which is an element of $\partial_B H(x)$ with the form $D^F \nabla F(x)^T + D^G \nabla G(x)^T$, where $D^F = \mathrm{diag}\{D_1^F, \ldots, D_n^F\}$ and $D^G = \mathrm{diag}\{D_1^G, \ldots, D_n^G\}$ are determined by

$$D_i^F = \begin{cases} \dfrac{F_i(x)}{\sqrt{F_i(x)^2 + G_i(x)^2}} - 1 & \text{if } F_i(x)^2 + G_i(x)^2 > 0, \\ \xi_i & \text{if } F_i(x) = G_i(x) = 0,\ \|\nabla F_i(x)\| + \|\nabla G_i(x)\| = 0, \\ \dfrac{\nabla F_i(x)^T d}{\sqrt{(\nabla F_i(x)^T d)^2 + (\nabla G_i(x)^T d)^2}} - 1 & \text{if } F_i(x) = G_i(x) = 0,\ \|\nabla F_i(x)\| + \|\nabla G_i(x)\| > 0, \end{cases}$$

$$D_i^G = \begin{cases} \dfrac{G_i(x)}{\sqrt{F_i(x)^2 + G_i(x)^2}} - 1 & \text{if } F_i(x)^2 + G_i(x)^2 > 0, \\ \eta_i & \text{if } F_i(x) = G_i(x) = 0,\ \|\nabla F_i(x)\| + \|\nabla G_i(x)\| = 0, \\ \dfrac{\nabla G_i(x)^T d}{\sqrt{(\nabla F_i(x)^T d)^2 + (\nabla G_i(x)^T d)^2}} - 1 & \text{if } F_i(x) = G_i(x) = 0,\ \|\nabla F_i(x)\| + \|\nabla G_i(x)\| > 0, \end{cases}$$

for $i = 1, \ldots, n$, and $\xi_i$ and $\eta_i$ are some constants. Note that we have not specified any fixed numbers for $D_i^F$ and $D_i^G$ if $F_i(x) = G_i(x) = 0$ and $\|\nabla F_i(x)\| + \|\nabla G_i(x)\| = 0$. In practice, however, this does not cause any problem in calculating a generalized Jacobian $V \in \partial_B H(x)$. In fact, as shown easily, $\|\nabla F_i(x)\| + \|\nabla G_i(x)\| = 0$ implies that the $i$-th row of any $V \in \partial_B H(x)$ becomes a zero vector.

## 8. Numerical Experiments

In this section, we present some numerical experiments for the proposed algorithm. We have chosen the $l_\infty$-norm in the constraint set of the subproblem (3.1), so that (3.1) becomes a linear least square problem with box constraints. The algorithm was implemented in Fortran and run on a DEC 5000 workstation. The subproblem (3.1) was solved by the NAG subroutine E04NCF. Throughout the computational experiments, the parameters used in the algorithm were set as $\Delta_{\min} = 1.0$, $\Delta_1 = 100$, $\rho_1 = 10^{-4}$, $\rho_2 = 0.75$. The trust region radius is updated as follows: If $\hat{r} \geq \rho_2$, then $\Delta_k := 2\hat{\Delta}$; if $\hat{r} < \rho_1$, then $\hat{\Delta} := \hat{\Delta}/2$. We used

$$\min\{\| \min\{F(x), G(x)\}\|_\infty, \|\nabla \Phi(x)\|_2\} \leq 10^{-6}$$

as a stopping rule, where $\min\{F(x), G(x)\}$ denotes the vector with components $\min\{F_i(x), G_i(x)\}$, $i = 1, \ldots, n$. Note that the second term of the left hand side of the above stopping rule is used as a safeguard against the case that an accumulation point of the sequence generated by the algorithm is a mere stationary point of $\Phi$, which is not a solution of the NCP or the GCP. We remark that the algorithm actually terminated by satisfying the condition $\| \min\{F(x), G(x)\}\|_\infty \leq 10^{-6}$ for all of the test problems reported below. The numerical results are summarized in Table 1 for different test problems. Table 1 shows, for each problem, the number of variables in the problem denoted by **Dim**, the starting point denoted by **Start**, the number of iterations denoted by **Iter** (which is also equal to the number of Jacobian evaluations for the functions $F$ and $G$), the number of function evaluations for the functions $F$ and $G$ denoted by **NF**, and the final value of $\Phi$ at the obtained solution denoted by **Res**. In the following, we just give some brief descriptions of tested problems since most of those test problems are widely used. We remark that the first nine problems are NCP problems and the last two problems are implicit complementarity problems. The interested reader is referred to the cited references for the source and more details of each test problem.

**Lemke problem** [17]: Here $n = 30$. This problem is a linear complementarity problem. The problem is known to run in exponential time by the well known Lemke's method.

*Starting points*: (a) $(0, \ldots, 0)$, (b) $(1, \ldots, 1)$.

**Kojima-Shindo problem** [4, 15, 25]: Here $n = 4$ and $F$ is not a $P_0$-function. This problem has two solutions: $(\sqrt{6}/2, 0, 0, 0.5)$ and $(1, 0, 3, 0)$.

*Starting points*: (a) $(0, 0, 0, 0)$, (b) $(1, 1, 1, 1)$.

**Kojima-Josephy problem** [4]: Here $n = 4$ and $F$ is not a $P_0$-function. This problem is a modification of the Kojima-Shindo problem and has a unique solution: $(\sqrt{6}/2, 0, 0, 0.5)$.

*Starting points*: (a) $(0, 0, 0, 0)$, (b) $(1, 1, 1, 1)$.

**Watson problem** [15, 17]: Here $n = 5$ and $F$ is the gradient of a strictly convex function. This problem has a unique degenerate solution: $(0, 0, 1, 2, 3)$.

*Starting points*: (a) $(0, 0, 0, 0, 0)$, (b) $(1, 1, 1, 1, 1)$.

**Nash equilibrium problem** [4, 25]: Here $n = 10$. The function $F$ is a $P$-function on the strictly positive orthant, but not twice differentiable.

*Starting points*: (a) $(0.5, \ldots, 0.5)$, (b) $(1, \ldots, 1)$.

**Modified Mathiesen problem** [4, 15]: Here $n = 4$. This is a slight modification of the Mathiesen example of a Walrasian equilibrium model. This problem has infinitely many solutions: $(\lambda, 0, 0, 0)$ for $\lambda \in [0, 3]$

*Starting points*: (a) $(0.5, 0.5, 0.5, 0.5)$, (b) $(1, 1, 1, 1)$.

**Spatial price equilibrium problem** [4, 25]: Here $n = 42$ and $F$ is a $P$-function. This problem has a unique solution.

*Starting points*: (a) $(0, \ldots, 0)$, (b) $(1, \ldots, 1)$.

**Traffic equilibrium problem** [4, 25]: Here $n = 50$. This is a traffic equilibrium problem with elastic demands.

*Starting points*: (a) $x_1, x_2, x_3, x_{10}, x_{11}, x_{20}$, $x_{21}$, $x_{22}$, $x_{29}$, $x_{30}$, $x_{40}$ and $x_{45}$ are ones, $x_{39}$, $x_{41}$, $x_{42}$, $x_{43}$, $x_{44}$, $x_{46}$ and $x_{47}$ are $7, 6, 7, 7, 10, 7, 6$, respectively, and all other elements of $x$ are zeros, (b) $(0, \ldots, 0)$, (c) $(1, \ldots, 1)$.

**Walrasian equilibrium problem** [4, 25]: Here $n = 14$. It has been observed from our computational experience that $\|\nabla\Phi(x)\|$ is still relatively large even when $\Phi(x)$ is very close to zero.

*Starting points*: (a) $(1, 1, 1, 1, 1, 1, 0.4, 1, 0, 3, 0, 0, 1, 0)$,
(b) $(0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.4, 1, 0, 3, 0, 0, 1, 0)$,
(c) $(1, \ldots, 1)$.

**Outrata-Zowe first problem** [22]: Here $n = 4$. It is an implicit complementarity problem with both $F$ and $G$ being linear functions. We used the same three starting points as in [22].

*Starting points*: (a) $(0, 0, 0, 0)$, (b) $(-0.5, -0.5, -0.5, -0.5)$, (c) $(-1, -1, -1, -1)$.

**Outrata-Zowe second problem** [22]: Here $n = 4$. This problem is a modification of the last problem in which $F$ is unchanged, but $G$ is a nonlinear function.

*Starting points*: (a) $(0, 0, 0, 0)$, (b) $(-0.5, -0.5, -0.5, -0.5)$, (c) $(-1, -1, -1, -1)$.

## 9. Conclusions

In this paper, we have proposed a trust region method for solving the generalized complementarity problem by using both semismooth equation and differentiable optimization reformulations. The special trust region updating rule enables us to establish not only global convergence but also local superlinear convergence of the algorithm under some suitable conditions. The preliminary numerical results presented shows the viability of this method. A possible future topic is to extend it to the variational inequality problem.

## References


[1] C. Chen and O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, *Computational Optimization and Applications*, to appear.

[2] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[3] R.W. Cottle, J.-S. Pang and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, New York, 1992.

[4] T. De Luca, F. Facchinei and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, Technical Report 01.95, Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza", Rome, Italy, January 1995.

[5] I.C. Dolcetta and U. Mosco, Implicit complementarity problems and quasi-variational inequalities, in: R.W. Cottle, F. Giannessi, J.-L. Lions, eds., *Variational Inequalities and Complementarity Problems: Theory and Applications*, Wiley, New York, 1980, pp. 75-87.

[6] F. Facchinei and J. Soares, A new merit function for nonlinear complementarity problems and a related algorithm, Technical Report 15.94, Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza", Roma, Italy, 1994, *SIAM Journal on Optimization*, to appear.

[7] A. Fischer, A special Newton-type optimization method, *Optimization* 24 (1992) 269-284.

[8] A. Fischer, An NCP-function and its use for the solution of complementarity problems, in: D. Du, L. Qi and R. Womersley, eds., *Recent Advances in Nonsmooth Optimization*, World Scientific Publishers, 1995, pp. 88-105.

[9] A. Fischer, Solution of monotone complementarity problems with locally Lipschitzian functions, Preprint MATH-NM-09-1995, Institute for Numerical Mathematics, Technical University of Dresden, Dresden, Germany, May 1995.

[10] A. Friedlander, J.M. Martínez and A. Santos, A new algorithm for bound constrained minimization, *Journal of Applied Mathematics and Optimization* 30 (1994) 235-266.

[11] M. Fukushima, Merit functions for variational inequality and complementarity problems, in: G. Di Pillo and F. Giannessi, eds., *Nonlinear Optimization and Applications*, Plenum Press, New York, to appear.

[12] M. Fukushima and J.-S. Pang, Minimizing and stationary sequences of merit functions for complementarity problems and variational inequalities, Report TR-IS-95026, Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan, 1995.

[13] S.A. Gabriel and J.-S. Pang, A trust region method for constrained nonsmooth equations, in: W.W. Hager, D.W. Hearn and P.M. Pardalos, eds., *Large Scale Optimization: State of the Art*, Kluwer Academic Publishers, Boston, 1994, pp. 159-186.

[14] P.T. Harker and J.-S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications, *Mathematical Programming* 48 (1990) 161-220.

[15] H. Jiang and L. Qi, A new nonsmooth equations approach to nonlinear complementarity problems, Technical Report AMR 94/30, School of Mathematics, The University of New South Wales, Australia, 1994 (Revised July 1995).

[16] C. Kanzow and M. Fukushima, Equivalence of the generalized complementarity problem to differentiable unconstrained minimization, *Journal of Optimization Theory and Applications*, to appear.

[17] Z.-Q. Luo and P. Tseng, A new class of merit functions for the nonlinear complementarity problem, Technical Report, Department of Mathematics, University of Washington, Seattle, 1995.

[18] J.M. Martínez and A. Santos, A trust region strategy for minimization on arbitrary domains, *Mathematical Programming* 68 (1995) 267-302.

[19] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, *SIAM Journal on Control and Optimization* 15 (1977) 959-972.

[20] J.J. Moré, Recent developments in algorithms and software for trust region methods, in: A. Bachem, M. Grötschel and B. Korte, eds., *Mathematical Programming: The State of the Art*, Springer-Verlag, Berlin, 1983, pp. 258-287.

[21] J.J. Moré, Global methods for nonlinear complementarity problems, Preprint MCS-P429-0494, Argonne National Laboratory, 1994.

[22] J.V. Outrata and J. Zowe, A Newton method for a class of quasi-variational inequalities, *Computational Optimization and Applications*, to appear.

[23] J.-S. Pang, The implicit complementarity problem, in: O.L. Mangasarian, S.M. Robinson, and R.R. Meyer, eds., *Nonlinear Programming 4*, Academic Press, New York, 1981, pp. 487-518.

[24] J.-S. Pang, Complementarity problems, in: R. Horst and P. Pardalos, eds., *Handbook of Global Optimization*, Kluwer Academic Publishers, Boston, 1995, pp. 271-338.

[25] J.-S. Pang and S.A. Gabriel, NE/SQP: A robust algorithm for the nonlinear complementarity problem, *Mathematical Programming* 60 (1993) 295-337.

[26] M.J.D. Powell, Convergence properties of a class of minimization algorithms, in: O.L. Mangasarian, R.R., Meyer and S.M. Robinson, eds., *Nonlinear Programming 2*, Academic Press, New York, 1975, pp. 1-27.

[27] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Mathematics of Operations Research* 18 (1993) 227-244.

[28] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Mathematical Programming* 58 (1993) 353-368.

[29] S.M. Robinson, Generalized equations, in: A. Bachem, M. Grötschel and B. Korte, eds., *Mathematical Programming: The State of the Art*, Springer-Verlag, Berlin, 1983, pp. 346-367.

[30] P. Tseng, N. Yamashita and M. Fukushima, Equivalence of complementarity problems to differentiable minimization: A unified approach, *SIAM Journal on Optimization*, to appear.

| Problem | Dim | Start | Iter | NF | Res |
|---|---|---|---|---|---|
| Lemke | 30 | (a) | 17 | 85 | $1.091891\times10^{-16}$ |
| | 30 | (b) | 4 | 5 | $1.834318\times10^{-17}$ |
| Kojima-Shindo | 4 | (a) | 10 | 28 | $3.090314\times10^{-17}$ |
| | 4 | (b) | 7 | 19 | $3.482776\times10^{-20}$ |
| Kojima-Josephy | 4 | (a) | 7 | 17 | $1.621855\times10^{-16}$ |
| | 4 | (b) | 7 | 19 | $9.042811\times10^{-25}$ |
| Watson | 5 | (a) | 21 | 22 | $1.029203\times10^{-15}$ |
| | 5 | (b) | 17 | 26 | $5.265319\times10^{-21}$ |
| Nash equilibrium | 10 | (a) | 9 | 10 | $5.029806\times10^{-21}$ |
| | 10 | (b) | 8 | 9 | $9.613329\times10^{-20}$ |
| Modified Mathiesen | 4 | (a) | 5 | 33 | $1.086752\times10^{-15}$ |
| | 4 | (b) | 8 | 60 | $5.742932\times10^{-22}$ |
| Spatial price | 42 | (a) | 19 | 25 | $2.479347\times10^{-13}$ |
| equilibrium | 42 | (b) | 20 | 23 | $2.538253\times10^{-13}$ |
| Traffic equilibrium | 50 | (a) | 12 | 37 | $1.091435\times10^{-08}$ |
| | 50 | (b) | 18 | 43 | $1.099669\times10^{-08}$ |
| | 50 | (c) | 16 | 26 | $2.415396\times10^{-13}$ |
| Walrasian equilibrium | 14 | (a) | 58 | 823 | $1.490584\times10^{-12}$ |
| | 14 | (b) | 52 | 710 | $1.718279\times10^{-12}$ |
| | 14 | (c) | 55 | 845 | $1.747844\times10^{-12}$ |
| Outrata-Zowe 1 | 4 | (a) | 5 | 12 | $7.649951\times10^{-18}$ |
| | 4 | (b) | 4 | 12 | $9.714795\times10^{-15}$ |
| | 4 | (c) | 5 | 6 | $3.434292\times10^{-24}$ |
| Outrata-Zowe 2 | 4 | (a) | 5 | 12 | $1.047405\times10^{-18}$ |
| | 4 | (b) | 4 | 12 | $4.886036\times10^{-15}$ |
| | 4 | (c) | 5 | 6 | $7.054095\times10^{-22}$ |

Table 1: Numerical results for the algorithm