Journal of Global Optimization (2006) 35: 343–366
DOI 10.1007/s10898-005-3842-4

Boundedness and Regularity Properties of Semismooth Reformulations of Variational Inequalities*

This paper is dedicated to Alex Rubinov on the occasion of his 65th birthday

LIQUN QI
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: maqilq@polyu.edu.hk)

(Accepted 6 October 2005)

Abstract. The Karush-Kuhn-Tucker (KKT) system of the variational inequality problem over a set defined by inequality and equality constraints can be reformulated as a system of semismooth equations via an nonlinear complementarity problem (NCP) function. We give a sufficient condition for boundedness of the level sets of the norm function of this system of semismooth equations when the NCP function is metrically equivalent to the minimum function; and a sufficient and necessary condition when the NCP function is the minimum function. Nonsingularity properties identified by Facchinei, Fischer and Kanzow, 1998, *SIAM J. Optim.* 8, 850–869, for the semismooth reformulation of the variational inequality problem via the Fischer-Burmeister function, which is an irrational regular pseudo-smooth NCP function, hold for the reformulation based on other regular pseudo-smooth NCP functions. We propose a new regular pseudo-smooth NCP function, which is piecewise linear-rational and metrically equivalent to the minimum NCP function. When it is used to the generalized Newton method for solving the variational inequality problem, an auxiliary step can be added to each iteration to reduce the value of the merit function by adjusting the Lagrangian multipliers only.

Mathematics Subject Classifications. Primary: 65H10; Secondary: 90C33.

Key words: generalized Newton method, global convergence, Lagrangian multipliers, quadratic convergence, variational inequalities

1. Introduction

Consider the variational inequality problem VI(X, F), which is to find a vector $x^* \in X$ such that for all $x \in X$,

$$F(x^*)^T (x - x^*) \geqslant 0, \tag{1.1}$$

*This work is supported by the Research Grant Council of Hong Kong.

where $X \subseteq \Re^n$ is a closed set and F is a continuously differentiable function from $\Re^n$ to $\Re^n$. Suppose that X is defined by

$$X = \{x \in \Re^n : g(x) \leqslant 0, \;\; h(x) = 0\}, \tag{1.2}$$

where g and h are twice continuously differentiable functions from $\Re^n$ to $\Re^p$ and $\Re^q$, respectively. Let $N = n + p + q$. The Karush-Kuhn-Tucker (KKT) system of VI(X, F) is:

$$\begin{aligned} F(x) + \sum_{j=1}^{p} u_j \nabla g_j(x) + \sum_{j=1}^{q} v_j \nabla h_j(x) = 0, \\ u \geqslant 0, g(x) \leqslant 0, u^T g(x) = 0, \\ h(x) = 0. \end{aligned} \tag{1.3}$$

For ease of presentation, we will let $z^T = (x^T, u^T, v^T)$. The KKT system plays a central role in the theory and algorithms for the variational inequality problem and the constrained nonlinear programming problem in the case that $F = \nabla f$ for some f mapping from $\Re^n$ to $\Re$ [22].

We may reformulate the KKT system (1.3) as a nonsmooth equation problem $H(z) = 0$, via an nonlinear complementary problem (NCP) function. Let

$$N_0 := \{(a, 0) : a \geqslant 0\} \cup \{(0, b) : b \geqslant 0\}.$$

DEFINITION 1.1 NCP function. A function $\psi : \Re^2 \to \Re$ is called an NCP function if $\phi(a, b) = 0$ if and only if $(a, b) \in N_0$.

Using an NCP function ϕ, we may reformulate the KKT system (1.3) to a system of nonsmooth equations:

$$\begin{aligned} & F(x) + \sum_{j=1}^{p} u_j \nabla g_j(x) + \sum_{j=1}^{q} v_j \nabla h_j(x) = 0, \\ & \phi(u_j, -g_j(x)) = 0, \quad j = 1, \ldots, p, \\ & h(x) = 0. \end{aligned} \tag{1.4}$$

We may denote it as

$$H(z) = 0, \tag{1.5}$$

where H maps from $\Re^N$ to $\Re^N$. We use θ, defined by

$$\theta(z) = \frac{1}{2}\|H(z)\|^2,$$

to denote its norm function. For some NCP functions, θ can be continuously differentiable. Assume that θ is continuously differentiable. Then we say that z^* is a stationary point of θ if $\nabla\theta(z^*) = 0$.

The simplest NCP function is the minimum function $\phi_{\min}$, defined by

$$\phi_{\min}(a, b) = \min\{a, b\}.$$

The minimum function is a piecewise linear function.

Another well-known NCP function is *the Fischer-Burmeister function* [12, 13], defined by

$$\psi_{FB}(a, b) := a + b - \sqrt{a^2 + b^2}. \tag{1.6}$$

This form of the Fischer-Burmeister function is the original form multiplied by -1. The Fischer-Burmeister NCP function is irrational and not piecewise smooth.

Both the minimum function and the Fischer-Burmeister function are strongly semismooth functions. Based upon this property, we may construct generalized Newton methods for solving (1.4), which is locally quadratically convergent [21, 23, 27].

If we use the Fischer-Burmeister function, then the norm function θ is continuously differentiable and the generalized Newton direction is a descent direction of θ at z^k [3, 17]. Based upon these nice properties of the Fischer-Burmeister function, globally and superlinearly convergent algorithms have been constructed for solving the NCP, with strong numerical evidence of the efficiency of these algorithms [3, 4, 8–10, 14–30]. In particular, Ferris and Munson [11] presented results for a 60 million variable problem.

The nice properties of the Fischer-Burmeister function are shared by some other NCP functions. Qi [24] defined a class of NCP functions, called regular pseudo-smooth NCP functions, and showed that they share the nice properties of the Fischer-Burmeister function. Qi [24] also proposed to use single-valued *generalized derivative functions* to replace set-valued B-subdifferentials and Clarke's generalized Jacobians, in solving pseudo-smooth equations. For single-valued generalized derivatives, also see [1, 26].

The generalized Newton method was also proposed to solve the KKT Equation (1.4). Facchinei et al. [5, 6] proposed to solve the KKT Equation (1.4) for VI(X, F), by generalized Newton methods or inexact Newton methods, via the Fischer-Burmeister function reformulation. Qi and Jiang [25] discussed properties of the KKT Equation (1.4) based on both the minimum function and the Fischer-Burmeister function. Their discussion was for constrained nonlinear programs, but actually is also true for VI(X, F). Facchinei et al. [7] identified regularity properties of the KKT Equation (1.4) based on the Fischer-Burmeister function.

However, the existing results of generalized Newton methods for solving the KKT Equation (1.4) are not very satisfactory. For example, a typical global convergence theorem claims that each accumulation point of the iterates generated by the generalized Newton method is a stationary point of the merit function θ. If the level sets of θ are bounded, then such an accumulation point exists. What are conditions for boundedness of the level sets of θ? No answer is available to this question in the literature.

The generalized Newton method for solving (1.4) is an infeasible method. When θ tends to zero, in which way x and u will approach feasibility? Can $\theta(z)$ provide a bound for such near feasibility of x and u?

The regularity properties of the KKT Equation (1.4) based on the Fischer-Burmeister function, identified in [7], are very significant to convergence analysis of generalized Newton methods. However, the Fischer-Burmeister function is irrational. The Lagrangian multipliers u_i are linear in the KKT system (1.3), but become irrational in the KKT Equation (1.4) if the Fischer-Burmeister function is used. Can we construct some NCP functions, which are linear in a substantial part of the plane, yet have all the nice properties of the Fischer-Burmeister function, when they are used in the KKT Equation (1.4)?

We try to answer these questions in this paper.

In Section 2, we show that if the NCP function ϕ is metrically equivalent to the minimum function $\phi_{\min}$, then the square root of $\theta(z)$ multiplying with a constant will give a bound of near feasibility of x and u. In this case, we give a sufficient condition for boundedness of the level sets of θ. We also give a sufficient and necessary condition for boundedness of the level sets of θ when $\phi = \phi_{\min}$ or $-\phi_{\min}$ or $|\phi_{\min}|$ or $-|\phi_{\min}|$. In Section 3, we briefly review and slightly extend the definition of regular pseudo-smooth NCP functions, introduced in [24], such that a nonsymmetric regular pseudo-smooth NCP function can be proposed later. We show in Section 4 that all the nice properties of the Fischer-Burmeister function, when it is used in the KKT Equation (1.4), hold for other regular pseudo-smooth NCP functions. Then, in Section 5, we propose a new regular pseudo-smooth NCP function, which is piecewise linear-rational. It is linear when $|b| \geqslant |a|$, and is equal to the minimum function when $b \geqslant |a|$. When $|b| < |a|$, it is rational. We show that this new NCP function is metrically equivalent to the minimum function. Generalized Newton and Gauss-Newton methods for solving the KKT Equation (1.4) are described in Section 6. In Section 7, we show that when the new NCP function is used, an auxiliary step can be added to each iteration to reduce the value of the merit function by adjusting the Lagrangian multipliers only. This may speed the method and eventually fix the Lagrangian multipliers at zero for inactive constraints at the optimal solution which the iterates converge to. As Conn et al. [2] noted: "In nonlinear optimization problems with expensive function and gradient evaluations, it is desirable to extract as much improvement as possible at each iteration of an algorithm. When the objective function contains a subset of variables that occurs in a predictable function form, a second, computationally relatively inexpensive, update can be applied to these variables following a classical optimization step. The additional step provides a further reduction in the objective function and can lead to superior optimization efficiency." This justifies the introduction of the new NCP function in Section 5 and the additional step in Section 7. Some concluding remarks are drawn in Section 8.

2. Boundedness of Level Sets

DEFINITION 2.1 Metrical Equivalence. Suppose that ϕ is an NCP function. If there exist $c_1 > 0$ and $c_2 > 0$ such that for all $(a, b) \in \Re^2$, we have

$$\frac{1}{c_1}|\phi_{\min}(a,b)| \leqslant |\phi(a,b)| \leqslant c_2|\phi_{\min}(a,b)|, \tag{2.1}$$

then we say that ϕ is metrically equivalent to the minimum function $\phi_{\min}$.

Tseng [29] established boundedness of the level sets of the norm function of the semismooth reformulation of the strong monotone NCP problem via the minimum function. He also proved the Fischer-Burmeister function is metrically equivalent to the minimum function. For the NCP problem, (2.1) plays a critical role to ensure boundedness of the level sets of the norm function, which in turn will guarantee that there exists an accumulation point of the sequence generated by a descent method for solving the problem.

Let $P := \{1, 2, \ldots, p\}$ and $Q := \{1, 2, \ldots, q\}$.

For any $\epsilon > 0$, let

$$X_\epsilon := \{x \in \Re^n : g_i(x) \leqslant c_1\epsilon, i \in P; |h_i(x)| \leqslant \epsilon, i \in Q\}$$

and

$$Z_\epsilon := \{z \in \Re^N : \theta(z) \leqslant \frac{1}{2}\epsilon^2\}.$$

LEMMA 2.1. *Let $\epsilon > 0$. Suppose that ϕ is an NCP function which is metrically equivalent to the minimum function with $c_1 > 0$ and $c_2 > 0$ in* (2.1). *Suppose that $z \in Z_\epsilon$. Then we have*

$$x \in X_\epsilon, \tag{2.2}$$

$$\left\| F(x) + \sum_{j=1}^{p} u_j \nabla g_j(x) + \sum_{j=1}^{q} v_j \nabla h_j(x) \right\| \leqslant \epsilon, \tag{2.3}$$

and for $i \in P$, *either*

$$u_i \geqslant -c_1\epsilon, \quad \text{and} \quad |g_i(x)| \leqslant c_1\epsilon, \tag{2.4}$$

or

$$|u_i| \leqslant c_1\epsilon, \quad \text{and} \quad g_i(x) \leqslant c_1\epsilon. \tag{2.5}$$

Proof. Suppose that the assumptions of the theorem hold. Clearly, we have

$$|h_i(x)| \leqslant \epsilon$$

for $i \in Q$. By (2.1), for $i \in P$, we have

$$\left|\min\{u_i, -g_i(x)\}\right| \leqslant c_1\epsilon.$$

It is easy to see that if $u_i \geqslant -g_i(x)$, then (2.4) holds and if $u_i \leqslant -g_i(x)$, then (2.5) holds.

This also proves (2.2). It is obvious that (2.3) holds. □

This shows that ϵ and $c_1\epsilon$ give bounds of near feasibility of x and u. We now discuss boundedness properties of X_ϵ.

In general, even if X_0 is bounded, X_ϵ can be unbounded for arbitrarily small $\epsilon > 0$. For example, let $n = 1$, $p = 1$, $q = 0$ and $g(x) = x(x+1)e^{-x}$. Then $X_0 = [-1, 0]$ is bounded, while X_ϵ is unbounded for arbitrarily small $\epsilon > 0$. However, when g_i is convex for each $i \in P$ and h_i is affine for each $i \in Q$, we have the following theorem.

THEOREM 2.1. *If* g_i *is convex for each* $i \in P$ *and* h_i *is affine for each* $i \in Q$, *then* X_ϵ *is bounded for any* $\epsilon > 0$ *as long as* X_0 *is bounded.*

Proof. This can be seen by the fact that in this case X_0 and X_ϵ for any $\epsilon > 0$ have the same recession cone

$$\{v \in \Re^n : \nabla g_i(x)^T v \leqslant 0, \forall x \in \Re^n, i \in P; \nabla h_i(x)^T v = 0, \forall i \in Q\}.$$

See [28]. □

We now discuss conditions for boundedness of Z_ϵ. For $\epsilon > 0$, let $\bar{X}_\epsilon$ be a subset of X_ϵ such that for each $x \in \bar{X}_\epsilon$, there exist $u \in \Re^p$ and $v \in \Re^q$ satisfying (2.3), and for each $i \in P$, either (2.4) and (2.5) holds.

For each $x \in \mathrm{cl}\ \bar{X}_\epsilon$, let

$$A_\epsilon(x) = \{i \in P : |g_i(x)| \leqslant c_1 \epsilon\}$$

and

$$B_\epsilon(x) = P \setminus A_\epsilon(x).$$

DEFINITION 2.2 ϵ-Mangasarian-Fromovitz Condition. Let $x \in \mathrm{cl}\ \bar{X}_\epsilon$. We say that the ϵ-Mangasarian-Fromovitz condition holds at x if there are no nonnegative numbers $\alpha_i, i \in A_\epsilon(x)$ and real numbers $\beta_i, i \in Q$, where at least one of α_i and β_i is nonzero, such that

$$\sum_{i \in A_\epsilon(x)} \alpha_i \nabla g_i(x) + \sum_{i \in Q} \beta_i \nabla h_i(x) = 0.$$

THEOREM 2.2. *Let $\epsilon > 0$. Suppose that ϕ is an NCP function which is metrically equivalent to the minimum function with $c_1 > 0$ and $c_2 > 0$ in* (2.1). *If $\bar{X}_\epsilon$ is bounded and the ϵ-Mangasarian-Fromovitz condition holds at all $x \in \mathrm{cl}\bar{X}_\epsilon$, then Z_ϵ is bounded.*

Proof. Assume that Z_ϵ is not bounded. Then there exist a sequence $\{z^k \in Z_\epsilon : k = 1, 2, \dots\}$ such that $\|z^k\| \to \infty$. Let

$$z^k = \begin{pmatrix} x^k \\ u^k \\ v^k \end{pmatrix}.$$

By Lemma 2.1, $x^k \in \bar{X}_\epsilon$. Since $\bar{X}_\epsilon$ is bounded, without loss of generality, we may assume that $x^k \to x^*$ and $x^* \in \mathrm{cl}\ \bar{X}_\epsilon$. Since $\|z^k\| \to \infty$,

$$\left\| \begin{pmatrix} u^k \\ v^k \end{pmatrix} \right\| \to \infty.$$

Without loss of generality, we may also assume that

$$\frac{\begin{pmatrix} u^k \\ v^k \end{pmatrix}}{\left\| \begin{pmatrix} u^k \\ v^k \end{pmatrix} \right\|} \to \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \neq 0,$$

where $\alpha \in \Re^p$ and $\beta \in \Re^q$. We now have

$$\left\| F(x^k) + \sum_{j=1}^{p} u_j^k \nabla g_j(x^k) + \sum_{j=1}^{q} v_j^k \nabla h_j(x^k) \right\| \leqslant \epsilon.$$

Dividing

$$F(x^k) + \sum_{j=1}^{p} u_j^k \nabla g_j(x^k) + \sum_{j=1}^{q} v_j^k \nabla h_j(x^k)$$

by

$$\left\| \begin{pmatrix} u^k \\ v^k \end{pmatrix} \right\|$$

and letting $k \to \infty$, we have

$$\sum_{j=1}^{p} \alpha_j \nabla g_j(x^*) + \sum_{j=1}^{q} \beta_j \nabla h_j(x^*) = 0.$$

Since $z^k \in Z_\epsilon$, by Lemma 2.1,

$$u_i^k \geqslant -c_1 \epsilon.$$

This implies that $\alpha_i \geqslant 0$ for all $i \in P$. For $i \in B_\epsilon(x^*)$, we have $i \in B_\epsilon(x^k)$ for k big enough and

$$|u_i^k| \leqslant c_1 \epsilon.$$

Thus, $\alpha_i = 0$ for $i \in B_\epsilon(x^*)$. We now have

$$\sum_{i \in A_\epsilon(x^*)} \alpha_i \nabla g_i(x^*) + \sum_{i \in Q} \beta_i \nabla h_i(x^*) = 0,$$

all $\alpha_i \geqslant 0$, and not all α_i and β_i are zero. This contradicts the assumption that the ϵ-Mangasarian-Fromovitz condition holds at x^*. Hence, Z_ϵ must be bounded. □

LEMMA 2.2. *Let $\epsilon > 0$. Suppose that ϕ is an NCP function which is metrically equivalent to the minimum function with $c_1 = c_2 = 1$ in* (2.1) *i.e., $\phi = \phi_{\min}$ or $-\phi_{\min}$ or $|\phi_{\min}|$ or $-|\phi_{\min}|$. If Z_ϵ is bounded, then $\bar{X}_\epsilon$ is compact and the ϵ-Mangasarian-Fromovitz condition holds at all $x \in \bar{X}_\epsilon$.*

Proof. In this case, it is easy to show that if Z_ϵ is bounded, then $\bar{X}_\epsilon$ is bounded. Let $\{x^k\} \subset \bar{X}_\epsilon$ and $x^k \to x^*$. Then there are $\{u^k\}$ and $\{v^k\}$ such that $z^k = ((x^k)^T, (u^k)^T, (v^k)^T)^T \in Z_\epsilon$. Since Z_ϵ is bounded, without loss of generality, we may assume that $z^k \to z^*$. It is easy to see that Z_ϵ is closed. Hence, $z^* \in Z_\epsilon$. Therefore, $x^* \in \bar{X}_\epsilon$. This shows that $\bar{X}_\epsilon$ is compact.

Assume that the ϵ-Mangasarian-Fromovitz condition does not hold at one particular $x^* \in \bar{X}_\epsilon$. Then there are $\alpha \in \Re^p$ and $\beta \in \Re^q$ such that not both α and β are zero,

$$\sum_{i \in A_\epsilon(x^*)} \alpha_i \nabla g_i(x^*) + \sum_{i \in Q} \beta_i \nabla h_i(x^*) = 0,$$

with $\alpha_i = 0$ for $i \in B_\epsilon(x^*)$ and $\alpha_i \geqslant 0$ for $i \in A_\epsilon(x^*)$. Since $x^* \in \bar{X}_\epsilon$, there will be $u^* \in \Re^p$ and $v^* \in \Re^q$ such that

$$z^* = \begin{pmatrix} x^* \\ u^* \\ v^* \end{pmatrix} \in Z_\epsilon.$$

It is now not difficult to see that

$$\begin{pmatrix} x^* \\ u^* + t\alpha \\ v^* + t\beta \end{pmatrix} \in Z_\epsilon$$

for all $t \geqslant 0$. This contradicts the boundedness of Z_ϵ. Hence, the theorem holds. □

Combining Theorem 2.1 and Lemma 2.2, we have the following theorem.

THEOREM 2.3. *Let $\epsilon > 0$. Suppose that ϕ is an NCP function which is metrically equivalent to the minimum function with $c_1 = c_2 = 1$ in* (2.1) *i.e., $\phi = \phi_{\min}$ or $-\phi_{\min}$ or $|\phi_{\min}|$ or $-|\phi_{\min}|$. Then Z_ϵ is bounded, if and only if $\bar{X}_\epsilon$ is compact and the ϵ-Mangasarian-Fromovitz condition holds at all $x \in \bar{X}_\epsilon$.*

If X_ϵ is bounded, then $\bar{X}_\epsilon$ is bounded. Theorem 2.1 gives a sufficient condition for boundedness of X_ϵ. A further question is: If $\bar{X}_0$ is bounded, in what condition will $\bar{X}_\epsilon$ also be bounded when ϵ is small?

3. Regular Pseudo-Smooth NCP Functions

We use $\|\cdot\|$ to denote the 2-norm in this paper. Let $T: \Re^n \to \Re^m$ be a locally Lipschitzian vector function. By Rademacher's theorem, T is differentiable almost everywhere. Let Ω_T denote the set of points where T is differentiable. Then the B-subdifferential of T at $x \in \Re^n$ is defined to be

$$\partial_B T(x) = \left\{ \lim_{\substack{x^k \to x \\ x^k \in \Omega_T}} \nabla T(x^k)^T \right\}, \tag{3.1}$$

while Clarke's generalized Jacobian of T at x is defined to be

$$\partial T(x) = \operatorname{conv} \partial_B T(x), \tag{3.2}$$

(see Qi [23]). T is called *semismooth* at x if T is directionally differentiable at x and for all $V \in \partial T(x+d)$ and $d \to 0$,

$$T'(x; d) = Vd + o(\|d\|); \tag{3.3}$$

T is called *strongly semismooth* at x if T is semismooth at x and for all $V \in \partial T(x+d)$ and $d \to 0$,

$$T'(x; d) = Vd + O(\|d\|^2); \tag{3.4}$$

T is called a (strongly) semismooth function if it is (strongly) semismooth everywhere. Here, $O(\|d\|)$ stands for a vector function $e : \Re^n \to \Re^n$, satisfying

$$\lim_{d \to 0} \frac{e(d)}{\|d\|} = 0,$$

while $O(\|d\|^2)$ stands for a vector function $e : \Re^n \to \Re^n$, satisfying

$$\|e(d)\| \leqslant M \|d\|^2$$

for all d satisfying $\|d\| \leqslant \delta$, and some $M > 0$ and $\delta > 0$.

We first summarize the definitions of pseudo-smooth functions, generalized derivative functions and their properties, discussed in [24].

DEFINITION 3.1 Pseudo-smooth function. Let $\psi : \Re^2 \to \Re$ be a strongly semismooth function. Denote P_ψ as the set of points where ψ takes zero. Let E_ψ be the extreme point set of P_ψ. We say ψ is *a pseudo-smooth function* if it is smooth everywhere in $\Re^2 \setminus E_\psi$.

For an NCP function ϕ, E_ϕ is a singleton, consisted by the origin only. Let $\bar{\Re}^2 := \Re^2 \setminus \{(0,0)^T\}$.

If for any (a, b) and $k > 0$,

$$\phi(ka, kb) = k\phi(a, b),$$

then we say that ϕ is positively homogeneous.

PROPOSITION 3.1. *Suppose that* $\phi : \Re^2 \to \Re$ *is a continuous positively homogeneous NCP function. Suppose that* ϕ *is smooth and piecewise twice continuously differentiable* (PC^2) *in* $\bar{\Re}^2$, *and* $A_\phi = \{\nabla\phi(a,b)^T : (a,b) \in \bar{\Re}^2\}$ *is bounded. Then* ϕ *is a pseudo-smooth NCP function, and* $\partial_B\phi(0,0) = A_\phi$.

We may use the generalized Newton method to solve (1.4). But it may need some work to determine a matrix in $\partial H(z)$ or $\partial_B H(z)$. To exploit the structure of (1.4) and properties of a pseudo-smooth NCP function, we may use the generalized derivative function for a pseudo-smooth NCP function, which was introduced in [24]. Let $\phi : \Re^2 \to \Re$ be a pseudo-smooth NCP function. We say that a function $\xi : \Re^2 \to \Re^2$ is *a generalized derivative function* of ϕ, if it satisfies

$$\xi(a,b) = \nabla\phi(a,b) \tag{3.5}$$

for $(a,b) \neq (0,0)$, and

$$\xi(0,0)^T \in \partial_B\psi(0,0). \tag{3.6}$$

Let

$$L(z) = F(x) + \sum_{j=1}^{p} u_j \nabla g_j(x) + \sum_{j=1}^{q} v_j \nabla h_j(x) = 0.$$

Then, the generalized derivative function $G : \Re^N \to \Re^{N\times N}$ of H has the following form:

$$G(z) = \begin{pmatrix} \nabla_x L(z) & \nabla g(x)^T & \nabla h(x)^T \\ -E(z)\nabla g(x) & D(z) & 0 \\ \nabla h(x) & 0 & 0 \end{pmatrix},$$

where $D(z)$ is a $p \times p$ diagonal matrix whose ith diagonal element is the first element of $\xi(u_i, -g_i(x))$, and $E(z)$ is a $p \times p$ diagonal matrix whose ith diagonal element is the second element of $\xi(u_i, -g_i(x))$.

Consider the following generalized Newton method for solving $H(z) = 0$:

$$z^{k+1} := z^k + d^k, \tag{3.7}$$

where d^k is a solution of

$$H(z^k) + G(z^k)d^k = 0. \tag{3.8}$$

EXAMPLE 3.1. Let $\phi:\Re^2\to\Re$ be the Fischer-Burmeister function, defined by (1.6). Let

$$\xi(a,b):=\begin{cases}\begin{pmatrix}0\\1\end{pmatrix} & \text{if } a=b=0,\\ \begin{pmatrix}1-\dfrac{a}{\sqrt{a^2+b^2}}\\ 1-\dfrac{b}{\sqrt{a^2+b^2}}\end{pmatrix} & \text{if } (a,b)\in\bar{\Re}^2.\end{cases} \tag{3.9}$$

Then ξ is a generalized derivative function of the Fischer-Burmeister function ϕ_{FB}.

Assume that H is defined by (1.4), and that G is a generalized derivative function of H. The following two theorems and one proposition are similar to Theorem 3.1, Proposition 3.1 and Theorem 3.2 of [24] for the NCP problem. Their proofs are also similar. Thus, we omit their proofs here.

THEOREM 3.1. *1*

(i) *H_j is continuously differentiable at z with $\nabla H_j(z)=G_j(z)^T$, except when $j=n+i$, $i\in P$ and $(u_i,-g_i(x))=0$.*

(ii) *For any fixed $z\in\Re^N$, we have*

$$H(z+d)=H(z)+G(z+d)d+o(\|d\|) \tag{3.10}$$

as $d\to 0$. If furthermore ∇F, $\nabla^2 g$ and $\nabla^2 h$ are locally Lipschitz around z, then we have

$$H(z+d)=H(z)+G(z+d)d+O(\|d\|^2) \tag{3.11}$$

as $d\to 0$.

(iii) *The norm function θ is continuously differentiable in $\Re^N$ with*

$$\nabla\theta(z)^T=H(z)^TG(z) \tag{3.12}$$

for all $z\in\Re^N$.

(iv) *If $H(z)\neq 0$ and*

$$H(z)+G(z)d(z)=0$$

has a solution $d(z)$, then the generalized Newton direction $d(z)$ is a descent direction of θ.

(v) *If z^* is a stationary point of θ and $G(z^*)$ is nonsingular, then z^* is a solution of $H(z)=0$.*

PROPOSITION 3.2. Let $z^* \in \Re^N$ and $A(z^*)$ be the set of all $N \times N$ matrices W such that

$$W = \begin{pmatrix} \nabla_x L(z^*) & \nabla g(x^*)^T & \nabla h(x^*)^T \\ -E\nabla g(x^*) & D & 0 \\ \nabla h(x^*) & 0 & 0 \end{pmatrix},$$

where D is a $p \times p$ diagonal matrix whose ith diagonal element is the first element of $\xi(u_i^*, -g_i(x^*))$, and E is a $p \times p$ diagonal matrix whose ith diagonal element is the second element of $\xi(u_i^*, -g_i(x^*))$. If all the matrices in $A(z^*)$ are nonsingular, then there are a positive number c and a neighborhood $N(z^*)$ of z^* such that for all $z \in N(z^*)$, $G(z)$ is nonsingular and

$$\|G(z)^{-1}\| \leqslant c.$$

We say that H is *A-regular* at z^* if all the matrices in $A(z^*)$ are nonsingular. Note that $G(z^*)$ is only one element of $A(z^*)$. Hence, this condition is stronger than nonsingularity of $G(z^*)$. Then, we have the following theorem for quadratic convergence of the generalized Newton method (3.7)–(3.8).

THEOREM 3.2. *If z^* is a stationary point of θ and H is A-regular at z^*, then z^* is a solution of $H(z) = 0$, the generalized Newton method* (3.7) – (3.8) *is well-defined in a neighborhood of z^* and the sequence $\{z^k\}$ converges to z^* Q-superlinearly if z^0 is in this neighborhood. If furthermore ∇F, $\nabla^2 g$ and $\nabla^2 h$ are locally Lipschitz around z^*, then this convergence is Q-quadratic.*

We then summarize the definition of a regular pseudo-smooth NCP function and its properties, also discussed in [24].

Let

$$N_+ := \{(a, b) : (a, b) > 0\}$$

and

$$N_- := \Re^2 \setminus (N_+ \cup N_0).$$

DEFINITION 3.2 Regular pseudo-smooth NCP function. Suppose that ϕ: $\Re^2 \to \Re$ is an NCP function, satisfying the conditions of Proposition 3.1, i.e., ϕ is a continuous positively homogeneous NCP function, smooth and PC^2 in $\bar{\Re}^2$, and $A_\phi = \{\nabla\phi(a, b)^T : (a, b) \in \bar{\Re}^2\}$ is bounded. Then ϕ is a pseudo-smooth NCP function, and $\partial_B \phi(0, 0) = A_\phi$. We say that ϕ is a *regular pseudo-smooth NCP function* if ϕ also satisfies the following conditions:

(i) $\phi(a,b)>0$ for all $(a,b)\in N_+$ and $\phi(a,b)<0$ for all $(a,b)\in N_-$;
(ii) $\nabla\phi(a,b)\geqslant 0$ for all $(a,b)\in\bar{\Re}^2$;
(iii) $\nabla\phi(a,0)=(0,\alpha)^T$ for all $a>0$, where $\alpha>0$;
(iv) $\nabla\phi(0,b)=(1,0)^T$ for all $b>0$;
(v) the first element of $\nabla\phi(a,b)$ is positive for all (a,b) not in the ray of $\{(a,0):a\geqslant 0\}$.

In the original definition of the regular pseudo-smooth NCP function, $\alpha=1$. In order to study some nonsymmetric NCP function, we do not fix $\alpha=1$. The Fischer-Burmeister function ϕ_{FB} is a regular pseudo-smooth NCP function. Other regular pseudo-smooth NCP functions include the (normalized) Tseng-Luo NCP function, the (normalized) Kanzow-Kleinmichel NCP function, the ratio generated NCP function, etc [24]. All these regular pseudo-smooth NCP functions are symmetric, i.e., $\phi(a,b)\equiv\phi(b,a)$.

4. Nonsingularity Conditions

DEFINITION 4.1 Partially positive definiteness condition and linear independence condition. Suppose that $z^T=(x^T,u^T,v^T)\in\Re^N$. Let $P=\{1,\cdots,p\}$,

$$I(z)=\{j:j\in P,\, g_j(x)=0,\, u_j\geqslant 0\},$$

$$I_0(z)=\{j\in I(z):u_j=0\},$$

$$I_1(z)=\{j\in I(z):u_j>0\}$$

and

$$G(z)=\{d\in\Re^n:\nabla g_j(x)^T d=0\ \text{ for } j\in I_1(z),\, \nabla h(x)^T d=0\}.$$

The KKT system (1.3) is said to satisfy the partially positive definiteness condition at a point $z\in\Re^N$ if $d^T\nabla_x L(z)d>0$ for all $d\in G(z)\setminus\{0\}$.The KKT system (1.3) is said to satisfy the linear independence condition at a point $z\in\Re^N$ if $\{\nabla g_j(x),\, j\in I(z),\, \nabla h_j(x),\, j=1,\ldots,q\}$ are linearly independent.

THEOREM 4.1. *Let ϕ in* (1.4) *be relaxed to be any regular pseudo-smooth NCP function satisfying Definition 3.2. If the KKT system* (1.3) *satisfies the partially positive definiteness condition and the linear independence condition at a point $z\in\Re^N$, then H, defined by* (1.4) *is A-regular at z.*

Proof. By the definition of A-regularity of H, it suffices to show that any $N \times N$ matrix W, with the form

$$W = \begin{pmatrix} \nabla_x L(z) & \nabla g(x)^T & \nabla h(x)^T \\ -E\nabla g(x) & D & 0 \\ \nabla h(x) & 0 & 0 \end{pmatrix},$$

is nonsingular, where $D = \text{diag}\{\alpha_1, \ldots, \alpha_p\}$, $E = \text{diag}\{\beta_1, \ldots, \beta_p\}$,

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} = \nabla\phi(u_j, -g_j(x))$$

if $(u_j, -g_j(x)) \neq 0$, and

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \in A_\phi$$

if $(u_j, -g_j(x)) = 0$. By Definition 3.2,

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \geqslant 0 \tag{4.1}$$

and

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \neq 0. \tag{4.2}$$

Suppose that

$$W \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = 0, \tag{4.3}$$

where $d_1 \in \Re^n, d_2 \in \Re^p, d_3 \in \Re^q$. Use d_{2j} to denote components of d_2. Then (4.3) implies

$$\nabla_x L(z) d_1 + \nabla g(x)^T d_2 + \nabla h(x)^T d_3 = 0, \tag{4.4}$$

$$-\beta_j \nabla g_j(x) d_1 + \alpha_j d_{2j} = 0 \tag{4.5}$$

for $j \in P$ and

$$\nabla h(x) d_1 = 0. \tag{4.6}$$

Let $P_1 = \{j \in P, \alpha_j > 0, \beta_j > 0\}$, $P_2 = \{j \in P, \beta_j = 0\}$ and $P_3 = \{j \in P, \alpha_j = 0\}$. Then $I_1(z) \subseteq P_3$. By (4.5), (4.1) and (4.2), $d_{2j} = 0$ if $j \in P_2$,

$$\nabla g_j(x) d_1 = 0 \tag{4.7}$$

if $j \in P_3$ and

$$\nabla g_j(x) d_1 = \nu_j d_{2j} \tag{4.8}$$

where $\nu_j = \frac{\alpha_j}{\beta_j} > 0$ if $j \in P_1$. Multiplying (4.4) by d_1^T, by (4.6), (4.7) and (4.8),

$$d_1^T \nabla_x L(z) d_1 + \sum_{j \in P_1} \nu_j d_{2j}^2 = 0.$$

Since $I_1(z) \subseteq P_3$, by (4.6) and (4.7), $d_1 \in G(z)$. Since $\nu_j > 0$ for $j \in P_1$, by the partially positive definiteness condition, $d_1 = 0$ and $d_{2j} = 0$ for $j \in P_1$. Now (4.4) yields

$$\sum_{j \in P_3} \nabla g_j(x)^T d_{2j} + \nabla h(x)^T d_3 = 0.$$

Note that $P_3 \subseteq I(z)$. By the linear independence condition, $d_3 = 0$ and $d_{2j} = 0$ for $j \in P_3$. Hence, $d = 0$. This shows that W is nonsingular. Therefore, H is A-regular at z. This completes the proof. □

Remark 4.1. Note that z can be any point in $\Re^N$. Thus, if the KKT system (1.3) satisfies the partially positiveness condition and the linear independence condition at z^k, the Equation (3.8) is solvable. The result in the special case that ϕ is the Fischer-Burmeister function was established by Facchine et al. [7], in a slightly different form. They established some more general results in that special case. We do not go to that detail. In the nonlinear programming case, if z is a solution of the KKT system (1.3), then the partially positive definiteness condition becomes the strong second-order sufficiency condition.

5. A New Regular Pseudo-Smooth NCP Function

We now present a new NCP function $\phi : \Re^2 \to \Re$, defined by

$$\phi(a, b) = \begin{cases} a & \text{if } b \geqslant |a|, \\ 2b - \dfrac{b^2}{a} & \text{if } a \geqslant |b| \quad \text{and } a > 0, \\ 2a + 2b + \dfrac{b^2}{a} & \text{if } a \leqslant -|b| \quad \text{and } a < 0, \\ a + 4b & \text{if } b \leqslant -|a|. \end{cases}$$

Then

$$\nabla\phi(a,b)=\begin{cases}\begin{pmatrix}1\\0\end{pmatrix} & \text{if } b\geqslant|a| \quad \text{and } b>0,\\ \begin{pmatrix}\frac{b^2}{a^2}\\2-\frac{2b}{a}\end{pmatrix} & \text{if } a\geqslant|b| \quad \text{and } a>0,\\ \begin{pmatrix}2-\frac{b^2}{a^2}\\2+\frac{2b}{a}\end{pmatrix} & \text{if } a\leqslant-|b| \quad \text{and } a<0,\\ \begin{pmatrix}1\\4\end{pmatrix} & \text{if } b\leqslant-|a| \quad \text{and } b<0,\end{cases}.$$

and

$$A_\phi=\partial_B\phi(0,0)=\left\{\begin{pmatrix}t^2\\2-2t\end{pmatrix}:|t|\leqslant1\right\}\bigcup\left\{\begin{pmatrix}2-t^2\\2-2t\end{pmatrix}:|t|\leqslant1\right\}.$$

By Definitions 1.1 and 3.2, as well as Proposition 3.1, we see that the following theorem holds.

THEOREM 5.1. *ϕ is a regular pseudo-smooth NCP function in the extended sense with $\alpha=2$.*

For our new NCP function, we also have the following theorem.

THEOREM 5.2. *ϕ is metrically equivalent to the minimum function with $c_1=1$ and $c_2=5$, i.e., for all $(a,b)\in\Re^2$, we have*

$$|\phi_{\min}(a,b)|\leqslant|\phi(a,b)|\leqslant5|\phi_{\min}(a,b)|. \tag{5.1}$$

Proof. For $b\geqslant|a|$, we have

$$\phi(a,b)=\phi_{\min}(a,b)=a.$$

For $a\leqslant-|b|$ and $a<0$, we have $\phi_{\min}(a,b)=a$ and $|b|\leqslant|a|$. Thus,

$$|\phi(a,b)|\leqslant5|a|=5|\phi_{\min}(a,b)|$$

and

$$|\phi(a,b)|\geqslant2|a|-2|b|+\left|\frac{b^2}{a}\right|=|a|\cdot\left|2-2\left|\frac{b}{a}\right|+\left|\frac{b}{a}\right|^2\right|\geqslant|a|=|\phi_{\min}(a,b)|.$$

For $a \geqslant |b|$ and $a > 0$, we have $\phi_{\min}(a, b) = b$ and $|b| \leqslant |a|$. Thus,

$$|\phi(a, b)| \leqslant 3|b| = 3|\phi_{\min}(a, b)|$$

and

$$|\phi(a, b)| \geqslant 2|b| - \left|\frac{b^2}{a}\right| \geqslant |b| = |\phi_{\min}(a, b)|.$$

For $b \leqslant -|a|$, we have $\phi_{\min}(a, b) = b$ and $|b| \geqslant |a|$. Thus,

$$|\phi(a, b)| \leqslant 5|b| = 5|\phi_{\min}(a, b)|$$

and

$$|\phi(a, b)| \geqslant 4|b| - |a| \geqslant 3|b| \geqslant |\phi_{\min}(a, b)|.$$

Putting the four cases together, we have (5.1). □

6. Generalized Damped Newton and Gauss-Newton Methods for Solving KKT Equations

We may solve (1.5) by the following generalized damped Newton method.

ALGORITHM 1 (Generalized Damped Newton Method).

Step 1. Let $z^0 \in \Re^N$, $\sigma, \rho \in (0, 1)$, $\eta > 0$, $p > 2$ and $k = 0$.
Step 2. If $H(z^k) = 0$, stop. Otherwise, let d^k be a solution of

$$H(z^k) + G(z^k)d = 0. \tag{6.1}$$

If (6.1) is not solvable, or if $\nabla\theta(z^k)^T d^k > -\eta\|d^k\|^p$, set $d^k = -\nabla\theta(z^k)$.
Step 3. If

$$\theta(z^k + d^k) \leqslant \sigma\theta(z^k),$$

set $\bar{z}^k = z^k + d^k$ and go to Step 6.
Step 4. Let $\alpha_k = \rho^{j_k}$, where j_k is the smallest nonnegative integer j such that

$$\theta(z^k + \rho^j d^k) - \theta(z^k) \leqslant \sigma\rho^j \nabla\theta(z^k)^T d^k,$$

where ρ^j means the jth power of ρ.
Step 5. Let $\bar{z}^k = z^k + \alpha_k d^k$. Go to Step 6.

Step 6. Let $z^{k+1} = \bar{z}^k$ and $k = k+1$. Go to Step 2.

This algorithm is a generalization of Algorithm 1 of [24] for the NCP problem. By a similar argument, we may have the following convergence theorem.

THEOREM 6.1. *Assume that z^* is an accumulation point of $\{z^k\}$ generated by Algorithm 1. Then z^* is a stationary point of θ. If $G(z^*)$ is nonsingular, then z^* is a solution of $H(z)=0$. If H is A-regular at z^*, then $\{z^k\}$ converges to z^* Q-superlinearly. If furthermore ∇F, $\nabla^2 g$ and $\nabla^2 h$ are locally Lipschitz around z^*, then this convergence is Q-quadratic.*

We may also use the Gauss-Newton technique. The following is a generalized damped Gauss-Newton method, which is a generalization of Algorithm 2 in [24] for the NCP problem.

ALGORITHM 2 (Generalized Damped Gauss-Newton Method).

Step 1. Let $z^0 \in \Re^N, \sigma \in (0, \frac{1}{2}), \rho \in (0, 1), \beta_0 > 0, k = 0$.

Step 2. If $G(z^k)^T H(z^k) = 0$, stop. Otherwise, let d^k be a solution of

$$G(z^k)^T H(z^k) + [G(z^k)^T G(z^k) + \beta_k I]d = 0.$$

Step 3. Let $\alpha_k = \rho^{j_k}$, where j_k is the smallest nonnegative integer j such that

$$\theta(z^k + \rho^j d^k) - \theta(z^k) \leqslant \sigma \rho^j \nabla\theta(z^k)^T d^k,$$

where ρ^j means the jth power of ρ.

Step 4. Choose $\beta_{k+1} > 0$. Let $\bar{z}^k = z^k + \alpha_k d^k$.

Step 5. Let $z^{k+1} = \bar{z}^k$ and $k := k+1$. Go to Step 2.

Similar to Theorem 7.2 of [24], we have the following global and quadratic convergence theorem for this algorithm.

THEOREM 6.2. *Let $\beta_k = \min\{\theta(z^k), \|\nabla\theta(z^k)\|\}$ in Algorithm 2. Then Algorithm 2 is well-defined. Assume that z^* is an accumulation point of $\{z^k\}$ generated by Algorithm 2. If H is A-regular at z^*, then z^* is a solution of $H(z) = 0$, and $\{z^k\}$ converges to z^* Q-superlinearly. If furthermore ∇F, $\nabla^2 g$ and $\nabla^2 h$ are locally Lipschitz around z^*, then this convergence is Q-quadratic.*

7. An Additional Step

We may change Step 6 in Algorithm 1 and Step 5 of Algorithm 2 as follows.

Suppose that $\bar{z}^k$ is given. First we may reduce the value of θ with respect to v_i by fixing other variables in z at $\bar{z}^k$. This can be done when ϕ is a general regular pseudo-smooth NCP function. We may do it for $i=1,\ldots,p$, sequentially. For each i, this turns out to reduce the value of a convex quadratic polynomial

$$\frac{1}{2}\|\nabla h_i(\bar{x}^k)\|^2 v_i^2 + \nabla h_i(\bar{x}^k)^T\left[F(\bar{x}^k) + \sum_{j=1}^{p}\bar{u}_j^k \nabla g_j(\bar{x}^k) + \sum_{\substack{j=1\\ j\neq i}}^{q}\bar{v}_j^k \nabla h_j(\bar{x}^k)\right]v_i.$$

Thus, if $\nabla h_i(\bar{x}^k)\neq 0$, we may replace $\bar{v}_i^k$ in $\bar{z}^k$ with

$$\bar{v}_i^k = \frac{-\nabla h_i(\bar{x}^k)^T\left[F(\bar{x}^k) + \sum_{j=1}^{p}\bar{u}_j^k \nabla g_j(\bar{x}^k) + \sum_{\substack{j=1\\ j\neq i}}^{q}\bar{v}_j^k \nabla h_j(\bar{x}^k)\right]}{\|\nabla h_i(\bar{x}^k)\|^2}, \qquad (7.1)$$

for $i=1,\ldots,p$, sequentially. In this way we obtain a new $\bar{z}^k$ with a reduced value of θ. We may minimize θ with respect to the vector v. But this implies that we need to solve a system of linear equations. It is not worth doing that as we only wish to use minor efforts as (7.1) to reduce θ.

We now assume that ϕ is the new NCP function introduced in Section 5. Since ϕ is linear in half of the plane, we may also reduce the value of θ with respect to u_i if $i\in I_1$ or $i\in I_2$. Here,

$$I_1=\{i: -g_i(\bar{x}^k)\geqslant|\bar{u}_i^k|>0\}$$

and

$$I_2=\{i: g_i(\bar{x}^k)\geqslant|\bar{u}_i^k|,\ g_i(\bar{x}^k)>0\}$$

I_1 does not include the case $g_i(\bar{x}^k)<0$ and $u_i^k=0$ as this case is a standard ith part of the solution of the KKT system, thus we do not want to update u_i in this case.

For $i\in I_1$, we may reduce the value of θ with respect to u_i by fixing other variables in z at $\bar{z}^k$, subject to the constraint

$$-g_i(\bar{x}^k)\geqslant|u_i|. \qquad (7.2)$$

This turns out to reduce the value of a uniformly convex quadratic polynomial

$$\frac{1}{2}\left[1+\|\nabla g_i(\bar{x}^k)\|^2\right]u_i^2+\nabla g_i(\bar{x}^k)^T\left[F(\bar{x}^k)+\sum_{\substack{j=1\\ j\neq i}}^{p}\bar{u}_j^k\nabla g_j(\bar{x}^k)+\sum_{j=1}^{q}\bar{v}_j^k\nabla h_j(\bar{x}^k)\right]u_i,$$

over the constraint (7.2). If

$$\frac{1}{2}\left[1+\|\nabla g_i(\bar{x}^k)\|^2\right](\bar{u}_i^k)^2+\nabla g_i(\bar{x}^k)^T \times\left[F(\bar{x}^k)+\sum_{\substack{j=1\\ j\neq i}}^{p}\bar{u}_j^k\nabla g_j(\bar{x}^k)+\sum_{j=1}^{q}\bar{v}_j^k\nabla h_j(\bar{x}^k)\right]\bar{u}_i^k\geqslant 0, \tag{7.3}$$

then we simply replace $\bar{u}_i^k$ in $\bar{z}^k$ with

$$\bar{u}_i^k=0. \tag{7.4}$$

This means that we fix $u_i=0$ in this case. This will eventually fix the Lagrangian multipliers at zero for inactive constraints at the optimal solution which the iterates converge to.

Otherwise, let

$$\hat{u}_i^k=\frac{-\nabla g_i(\bar{x}^k)^T\left[F(\bar{x}^k)+\sum_{\substack{j=1\\ j\neq i}}^{p}\bar{u}_j^k\nabla g_j(\bar{x}^k)+\sum_{j=1}^{q}\bar{v}_j^k\nabla h_j(\bar{x}^k)\right]}{1+\|\nabla g_i(\bar{x}^k)\|^2}. \tag{7.5}$$

Thus, we may replace $\bar{u}_i^k$ in $\bar{z}^k$ with

$$\bar{u}_i^k=\begin{cases} g_i(\bar{x}^k) & \text{if } \hat{u}_i^k<g_i(\bar{x}^k),\\ \hat{u}_i^k & \text{if } g_i(\bar{x}^k)\leqslant\hat{u}_i^k\leqslant -g_i(\bar{x}^k),\\ -g_i(\bar{x}^k) & \text{if } \hat{u}_i^k>-g_i(\bar{x}^k).\end{cases} \tag{7.6}$$

We may use (7.4) or (7.6) to update $\bar{u}_i^k$, depending upon if (7.3) holds or not, for $i\in I_1$ sequentially.

We may deal with u_i for $i\in I_2$ similarly, though we do not try to fix $u_i=0$, as $g_i(\bar{x}^k)>0$ means infeasibility.

For $i\in I_2$, we may reduce the value of θ with respect to u_i by fixing other variables in z at $\bar{z}^k$, subject to the constraint

$$g_i(\bar{x}^k)\geqslant |u_i|. \tag{7.7}$$

This turns out to reduce the value of a uniformly convex quadratic polynomial

$$\frac{1}{2}\left[1+\|\nabla g_i(\bar{x}^k)\|^2\right]u_i^2+\left\{-4g_i(\bar{x}^k)+\nabla g_i(\bar{x}^k)^T \left[F(\bar{x}^k)+\sum_{\substack{j=1\\ j\neq i}}^{p}\bar{u}_j^k\nabla g_j(\bar{x}^k)+\sum_{j=1}^{q}\bar{v}_j^k\nabla h_j(\bar{x}^k)\right]\right\}u_i,$$

over the constraint (7.7). Let

$$\hat{u}_i^k = \frac{4g_i(\bar{x}^k) - \nabla g_i(\bar{x}^k)^T \left[F(\bar{x}^k) + \sum_{\substack{j=1 \\ j\neq i}}^{p} \bar{u}_j^k \nabla g_j(\bar{x}^k) + \sum_{j=1}^{q} \bar{v}_j^k \nabla h_j(\bar{x}^k) \right]}{1 + \|\nabla g_i(\bar{x}^k)\|^2}.$$

Thus, we may replace $\bar{u}_i^k$ in $\bar{z}^k$ with

$$\bar{u}_i^k = \begin{cases} -g_i(\bar{x}^k) & \text{if } \hat{u}_i^k < -g_i(\bar{x}^k), \\ \hat{u}_i^k & \text{if } -g_i(\bar{x}^k) \leqslant \hat{u}_i^k \leqslant g_i(\bar{x}^k), \\ g_i(\bar{x}^k) & \text{if } \hat{u}_i^k > g_i(\bar{x}^k), \end{cases} \tag{7.8}$$

for $i \in I_2$ sequentially.

Thus, we may update v_i sequentially by (7.1) for $i = 1, \ldots, q$, u_i sequentially by (7.4) or (7.6) for $i \in I_1$, and by (7.8) for $i \in I_2$. This will reduce the value of θ with minor calculations. The updating on u_i are based upon the partially linearity of ϕ. Then we can let $z^{k+1} = \bar{z}^k, k := k+1$ and go to Step 2 in Algorithms 1 and 2.

Remark 7.1. For $i \in I_1$, using (7.6) to update $\bar{u}_i^k$ will not change the sign of $\bar{u}_i^k$. In this case, (7.3) does not hold. Thus,

$$\nabla g_i(\bar{x}^k)^T \left[F(\bar{x}^k) + \sum_{\substack{j=1 \\ j\neq i}}^{p} \bar{u}_j^k \nabla g_j(\bar{x}^k) + \sum_{j=1}^{q} \bar{v}_j^k \nabla h_j(\bar{x}^k) \right] \bar{u}_i^k < 0,$$

i.e., the signs of

$$\nabla g_i(\bar{x}^k)^T \left[F(\bar{x}^k) + \sum_{\substack{j=1 \\ j\neq i}}^{p} \bar{u}_j^k \nabla g_j(\bar{x}^k) + \sum_{j=1}^{q} \bar{v}_j^k \nabla h_j(\bar{x}^k) \right]$$

and the original $\bar{u}_i^k$ are different. By (7.5), this implies that $\hat{u}_i^k$ and the original $\bar{u}_i^k$ have the same sign. By (7.6), this updating will not change the sign of $\bar{u}_i^k$.

8. Concluding Remarks

In this paper, we give a sufficient condition for boundedness of the level sets of the norm function of the semismooth reformulation of the KKT system of the variational inequality problem, when the NCP function used in the reformulation is metrically equivalent to the minimum function; and a sufficient and necessary condition when the NCP function is the minimum function. We show that nonsingularity properties identified by Facchinei, et al. [7] for the semismooth reformulation of the variational inequality problem via the Fischer-Burmeister function, which is an irrational regular pseudo-smooth NCP function, hold for the reformulation based on other regular pseudo-smooth NCP functions. We then propose a new regular pseudo-smooth NCP function, which is piecewise linear-rational and metrically equivalent to the minimum NCP function. When it is used to the generalized Newton method for solving the variational inequality problem, an auxiliary step can be added to each iteration to reduce the value of the merit function by adjusting the Lagrangian multipliers only.

Acknowledgements

The author is thankful to Paul Tseng and Terry Rockafellar for the discussion on boundedness of X_ϵ, and to Yufei Yang for his comments on Section 4.

References

1. Chen X., Nashed, Z. and Qi, L. (2000), Smoothing methods and semismooth methods for nondifferentiable operator equations, *SIAM J. Numer. Anal.* 38, 1200–1216.
2. Conn, A.R., Vicente, L.N. and Visweswariah, C. (1999), Two-step algorithms for non-linear optimization with structure applications, *SIAM J. Optim.* 9, 924–947.
3. De Luca, T., Facchinei, F. and Kanzow, C. (1996), A semismooth equation approach to the solution of nonlinear complementarity problems, *Math. Prog.* 75, 407–439.
4. De Luca, T., Facchinei, F. and Kanzow, C. (2000), A theoretical and numerical comparison of some semismooth algorithms for complementarity problems, *Comput. Optim. Appl.* 16, 173–205.
5. Facchinei, F. Fischer, A. and Kanzow, C. (1995), A Semismooth Newton Method for Variational Inequalities: Theoretical Results and Preliminary Numerical Results, Technical Report 102, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany.
6. Facchinei, F., Fischer, A. and Kanzow, C. (1996), Inexact Newton methods for semismooth equations with applications to variational inequality problems, In: Di Pillo, G. and Giannessi, F. (eds.), *Nonlinear Optimization and Applications*, Plenum Press, New York, pp. 125–139.
7. Facchinei, F., Fischer, A. and Kanzow, C. (1998), Regularity properties of a semismooth reformulation of variational inequalities, *SIAM J. Optim.* 8, 850–869.
8. Facchinei, F. and Kanzow, C. (1997), A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems, *Math. Prog.* 76, 493–512.
9. Facchinei, F. and Pang, J.S. (2003), *Finite-Dimensional Variational Inequalities and Complementarity Problems, I-II*, Springer-Verlag, New York.
10. Facchinei, F. and Soares, J. (1997), A new merit function for nonlinear complementarity problems and a related algorithm, *SIAM J. Optim.* 7, 225–247.
11. Ferris, M.C. and Munson, T.S. (2000), Semismooth support vector machines, Data Mining Institute Technical Report 00-09, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin.

12. Fischer, A. (1992), A special Newton-type optimization method, *Optimization* 24, 269–284.
13. Fischer, A. (1995), An NCP-function and its use for the solution of complementarity problems, In: Du, D, Qi, L. and Womersley, R. (eds.), *Recent Advances in Nonsmooth Optimization*, World Scientific Publishers, New Jersey, pp. 88–105.
14. Fischer, A. (1997), Solution of monotone complementarity problems with locally Lipschitzian functions, *Math. Prog.* 76, 513–532.
15. Jiang, H. (1999), Global convergence analysis of the generalized Newton and Gauss-Newton methods of the Fischer-Burmeister equation for the complementarity problem, *Math. Oper. Res.* 24, 529–543.
16. Jiang, H., Fukushima, M., Qi, L. and Sun, D. (1998), A trust region method for solving generalized complementarity problems, *SIAM J. Optim.* 8, 140–157.
17. Jiang, H. and Qi, L. (1997), A new nonsmooth equations approach to nonlinear complementarity problems, *SIAM J. Control Optim.* 35, 178–193.
18. Jiang, H., Qi, L., Chen, X. and Sun, D. (1996), Semismoothness and superlinear convergence in nonsmooth optimization and nonsmooth equations, In: Di Pillo, G. and Giannessi, F. (eds.), *Nonlinear Optimization and Applications*, Plenum Press, New York, 197–212.
19. Jiang, H. and Ralph, D. (1999), Global and local superlinear convergence analysis of Newton-type methods for semismooth equations with smooth least squares, In: Fukushima, M. and Qi, L. (eds.), *Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Kluwer Academic Publishers, Nowell, MA, pp. 181–210.
20. Munson, T.S., Facchinei, F., Ferris, M.C., Fischer, A. and Kanzow, C. (2001), The semismooth algorithm for large scale complementarity problems, *INFORMS J. Comput.* 13, 294–311.
21. Pang, J.S. and Qi, L. (1993), Nonsmooth equations: Motivation and algorithms, *SIAM J. Optim.* 3, 443–465.
22. Qi, H., Qi, L. and Sun, D. (2004), Solving KKT systems via the trust region and the conjugate gradient methods, *SIAM J. on Optim.* 14, 439–463.
23. Qi, L. (1993), Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.* 18, 227–244.
24. Qi, L. (1999), Regular pseudo-smooth NCP and BVIP functions and globally and quadratically convergent generalized Newton methods for complementarity and variational inequality problems, *Math. Oper. Res.* 24, 440–471.
25. Qi, L. and Jiang, H. (1997), Semismooth Karush-Kuhn-Tucker equations and convergence analysis of Newton methods and quasi-Newton methods for solving these equations, *Math. Oper. Res.* 22, 301–325.
26. Qi, L., Ralph, D. and Zhou, G. (2000), Semiderivative functions and reformulation methods for solving complementarity and variational inequality problems, In: Di Pillo, G. and Giannessi, F. (eds.), *Nonlinear Optimization and Related Topics*, Kluwer Academic Publisher, Nowell, MA, 317–350.
27. Qi, L. and Sun, J. (1993), A nonsmooth version of Newton's method, *Math. Prog.* 58, 353–368.
28. Rockafellar, R.T. (1970), *Convex Analysis*. Princeton, NJ.
29. Tseng, P. (1996), Global behaviour of a class of merit functions for the nonlinear complementarity problem, *J. Optim. Theory Appl.* 89, 17–37.
30. Yamashita, N. and Fukushima, M. (1997), Modified Newton methods for solving semismooth reformulations of monotone complementarity problems, *Math. Prog.* 76, 469–491.