Journal of Computational Mathematics, Vol.18, No.3, 2000, 251–264.

A GLOBALLY DERIVATIVE-FREE DESCENT METHOD FOR NONLINEAR COMPLEMENTARITY PROBLEMS*[1)]

Hou-duo Qi
(*Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing, 100080, China*)
Yu-zhong Zhang
(*Institute of Operation Research, QuFu Normal University, ShanDong, 273165, China*)

Abstract

Based on a class of functions, which generalize the squared Fischer-Burmeister NCP function and have many desirable properties as the latter function has, we reformulate nonlinear complementarity problem (NCP for short) as an equivalent unconstrained optimization problem, for which we propose a derivative-free descent method in monotone case. We show its global convergence under some mild conditions. If F, the function involved in NCP, is R_0-function, the optimization problem has bounded level sets. A local property of the merit function is discussed. Finally, we report some numerical results.

Keywords: Complementarity problem, NCP-function, unconstrained minimization method, derivative-free descent method.

1. Introduction

Consider the nonlinear complementarity problem (NCP for short), which is to find an $x \in \Re^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0, \tag{1}$$

where $F : \Re^n \to \Re^n$ and the inequalities are taken componentwise. This problem have many important applications in various fields. [13, 7, 5].

Due to the less storage in computation, derivative-free descent method, which means the search direction used does not involve the Jacobian matrix of F, is popular in finding solutions of nonlinear complementarity problems. We briefly view some (not all) progress in such setting. In 1992, Fukushima [3] reformulates variational inequality problem, which includes NCP as its special element, into a constrained minimization problem through regularized gap function

$$f(x) = \max_{y \geq 0}\{(x-y)^T F(x) - \frac{1}{2\alpha}||x-y||^2\}, \qquad \alpha > 0.$$

* Received October 29, 1996.
1) This work is supported by the National Natural Science Foundation of China.

and propose a descent method for monotone case with global convergence. In 1993, Mangasarian and Solodov [12] reformulate NCP as an equivalent unconstrained minimization problem through the implicit Lagrangian function

$$\Psi(x) = \sum_{i=1}^{n} \psi_i(x), \tag{2}$$

where $\psi_i(x) = \phi_{MS}(x_i, F_i(x))$ and $\phi_{MS} : \Re^2 \to \Re^1$ is defined by

$$\phi_{MS}(a,b) = ab + \frac{1}{2\alpha}\left(\max{}^2(0, a-\alpha b) - a^2 + \max{}^2(0, b - \alpha a) - b^2\right), \qquad \alpha > 1.$$

Yamashita and Fukushima [19] propose a descent method for such reformulation with strong monotonicity and show the global convergence. Geiger and Kanzow [4] consider another kind of function (2) with $\psi_i(x) = \phi_{FB}^2(x_i, F_i(x))$, where $\phi_{FB} : \Re^2 \to \Re^1$ is Fischer-Burmeister NCP function defined by

$$\phi_{FB}(a,b) = \sqrt{a^2+b^2} - (a+b).$$

Also a descent method is described in [4] for monotone case. Recently, Luo and Tseng [10] study a new class of merit functions Ψ defined by

$$\Psi(x) = \psi_0(x^T F(x)) + \sum_{i=1}^{n} \psi_i(-F_i(x), x_i),$$

where $\psi_0 : \Re 1 \to [0,\infty)$ and $\psi_1, \cdots, \psi_n : \Re^2 \to [0,\infty)$ are continuous functions that are zero on the nonpositive orthant only. For such merit functions, descent methods are proposed for monotone case.

We reconsider the function ϕ_{MS}, the two parts in brackets, i.e, $(\max^2\{a - \alpha b, 0\} - a^2)$ and $(\max^2\{b-\alpha a, 0\} - b^2)$, can be viewed as relative errors with respect to a and b and the penalty parameter α. ϕ_{MS} is sum of the both parts with the third one. We note that the quantity $(\sqrt{a^2+\alpha b^2} - a)$ can be viewed as a relative error connected to $(\sqrt{a^2+b^2} - a)$ and α. $(\sqrt{\beta a^2 + b^2} - b)$ can be viewed in similar way. We consider the product of the both parts.

$$\phi(a,b) = (\sqrt{a^2+\alpha b^2} - a)(\sqrt{\beta a^2 + b^2} - b), \qquad \alpha > 0, \quad \beta > 0.$$

The function ϕ is originally proposed by Peng in [15] and its elementary properties are discussed therein. It is interesting to note that ϕ is a generalization of ϕ_{FB}^2, since for $\alpha = \beta = 1$, we have

$$\phi(a,b) = (\sqrt{a^2+\alpha b^2} - a)(\sqrt{\beta a^2+b^2} - b) = \frac{1}{2}(\sqrt{a^2+b^2} - a - b)^2 = \frac{1}{2}\phi_{FB}^2.$$

Moreover, if $\alpha\beta = 1$, and let $c = \sqrt{\alpha} b$, we have

$$\begin{aligned}\phi(a,b) &= \frac{1}{\sqrt{\alpha}}(\sqrt{a^2+c^2} - a)(\sqrt{a^2+c^2} - c) \\ &= \frac{1}{\sqrt{\alpha}}\phi_{FB}^2(a,b).\end{aligned}$$

Hence in this case, $\phi(a,b)$ reduces to the squared Fischer-Burmeister NCP function except a constant factor. The function ϕ has the following properties:

(P1) $\phi(a,b)=0 \Longleftrightarrow a\geq 0, \quad b\geq 0, \quad ab=0,$

(P2) $\phi(a,b)\geq 0 \quad$ for all $(a,b)\in\Re^2$.

With the (P1) and (P2) in hand, we can reformulate NCP as an equivalent unconstrained minimization problem

$$\min \Psi(x)=\sum_{i=1}^{n}\phi(x_i,F_i(x)) \quad \text{subject to} \quad x\in\Re^n. \tag{3}$$

The paper is organized as follows: In addition to some definitions, section 2 includes the detailed study on the properties of ϕ defined in (3). In section 3, we reformulate NCP as an equivalent minimization problem, and show that any stationary point of the objective function is a solution of NCP if $F'(x)$ is P_0-matrix. It is also showed that the level set is bounded if F is R_0-function. In section 4, In monotone case we propose a globally descent method without using any derivative information of F. In section 5, we discuss a local property of the merit function Ψ near strict complementarity solutions. Numerical results are included in section 6.

2. Properties of ϕ and Definitions

In this section, we show ϕ has some desirable properties in addition to (P1) and (P2). We note that the function ϕ is continuously differentiable. In fact, the partial-derivative of ϕ with respect to a and b have the following expressions: If $(a,b)=(0,0)$, then $\nabla\phi(0,0)=0$, and if $(a,b)\neq(0,0)$, then

$$\frac{\partial\phi}{\partial a}(a,b)=(\sqrt{a^2+\alpha b^2}-a)\left(\frac{\beta a}{\sqrt{\beta a^2+b^2}}+\frac{b-\sqrt{\beta a^2+b^2}}{\sqrt{a^2+\alpha b^2}}\right) \tag{4}$$

$$\frac{\partial\phi}{\partial b}(a,b)=(\sqrt{\beta a^2+b^2}-b)\left(\frac{\alpha b}{\sqrt{a^2+\alpha b^2}}+\frac{a-\sqrt{a^2+\alpha b^2}}{\sqrt{\beta a^2+b^2}}\right). \tag{5}$$

Proposition 2.1 [15]

$$\phi(a,b)=0 \quad \textit{if and only if} \quad \nabla\phi(a,b)=0.$$

Lemma 2.2 *For $(a,b)\in\Re^2_+$, we have*

$$\frac{\partial\phi}{\partial a}(a,b)\frac{\partial\phi}{\partial b}(a,b)\geq 0.$$

Proof. Due to the symmetric role of a and b in (4) and (5), it is suffice to prove the inequality below

$$\frac{\partial\phi}{\partial a}(a,b)\geq 0 \quad \text{for all } (a,b)\in\Re^2_+.$$

We only consider $0 \neq (a,b) \in \Re^2_+$, it follows that

$$\begin{aligned}
\frac{\partial \phi}{\partial a}(a,b) &= (\sqrt{a^2+\alpha b^2}-a)\left(\frac{\beta a}{\sqrt{\beta a^2+b^2}}+\frac{b-\sqrt{\beta a^2+b^2}}{\sqrt{a^2+\alpha b^2}}\right)\\
&= (\sqrt{a^2+\alpha b^2}-a)\frac{\beta a\sqrt{a^2+\alpha b^2}+b\sqrt{\beta a^2+b^2}-(b^2+\beta a^2)}{\sqrt{a^2+\alpha b^2}\sqrt{\beta a^2+b^2}}\\
&= (\sqrt{a^2+\alpha b^2}-a)\frac{\beta(a\sqrt{a^2+\alpha b^2}-a^2)+b(\sqrt{\beta a^2+b^2}-b)}{\sqrt{a^2+\alpha b^2}\sqrt{\beta a^2+b^2}}\\
&\geq 0.
\end{aligned}$$

□

Proposition 2.3 *If the two positive parameters α and β satisfy the condition*

$$\alpha\beta \leq 1.$$

Then for any $(a,b) \in \Re^2$, one has

$$\frac{\partial \phi}{\partial a}(a,b)\frac{\partial \phi}{\partial b}(a,b) \geq 0.$$

Proof. We assume the contrary that there exists a pair $(a,b) \in \Re^2$, which satisfies

$$\frac{\partial \phi}{\partial a}(a,b)\frac{\partial \phi}{\partial b}(a,b) < 0. \tag{6}$$

Without loss of generality, we assume that

$$\frac{\partial \phi}{\partial a}(a,b) > 0, \quad \frac{\partial \phi}{\partial b}(a,b) < 0. \tag{7}$$

By (7) and (4), we get

$$\frac{\beta a}{\sqrt{\beta a^2+b^2}} > \frac{\sqrt{\beta a^2+b^2}-b}{\sqrt{a^2+\alpha b^2}} \geq 0, \tag{8}$$

and hence $a > 0$. Combining (5), the second inequality of (7) and (8), we have

$$\begin{aligned}
0 &> \frac{\alpha b}{\sqrt{a^2+\alpha b^2}}+\frac{a-\sqrt{a^2+\alpha b^2}}{\sqrt{\beta a^2+b^2}}\\
&= \frac{\alpha b}{\sqrt{a^2+\alpha b^2}}+\frac{\sqrt{\beta a^2+b^2}-b}{\beta\sqrt{a^2+\alpha b^2}}-\frac{\sqrt{a^2+\alpha b^2}}{\sqrt{\beta a^2+b^2}} \qquad \text{(by 8)}\\
&= \frac{\alpha\beta b\sqrt{\beta a^2+b^2}+(\beta a^2+b^2)-b\sqrt{\beta a^2+b^2}-(a^2+\alpha b^2)\beta}{\beta\sqrt{a^2+\alpha b^2}\sqrt{\beta a^2+b^2}}\\
&= \frac{(\alpha\beta-1)b\sqrt{\beta a^2+b^2}+(1-\alpha\beta)b^2}{\beta\sqrt{a^2+\alpha b^2}\sqrt{\beta a^2+b^2}}\\
&= \frac{(\alpha\beta-1)(b\sqrt{\beta a^2+b^2}-b^2)}{\beta\sqrt{a^2+\alpha b^2}\sqrt{\beta a^2+b^2}}.
\end{aligned} \tag{9}$$

Since $\alpha\beta \le 1$, then by (9), we have $\alpha\beta < 1$ and

$$b\sqrt{\beta a^2 + b^2} - b^2 > 0.$$

It follows from the above inequality that $b > 0$, hence we obtain that $(a, b) \in \Re^2_+$. By Lemma 2.2, we get

$$\frac{\partial \phi}{\partial a}(a,b)\frac{\partial \phi}{\partial b}(a,b) \ge 0,$$

which contradicts the assumed inequality (6). This completes the proof. □

Proposition 2.4 [17] *If $\alpha\beta \le 1$, then the following equivalent relationships hold*

$$\phi(a,b) = 0 \Longleftrightarrow \frac{\partial \phi}{\partial a}(a,b) = 0 \Longleftrightarrow \frac{\partial \phi}{\partial b}(a,b) = 0 \Longleftrightarrow \frac{\partial \phi}{\partial a}(a,b)\frac{\partial \phi}{\partial b}(a,b) = 0.$$

At the end of this section, we recall some definitions for F. Let $I = \{1, 2, \cdots, n\}$.

Definition 2.1 *A function $F : \Re^n \to \Re^n$ is said to be*
(a) *monotone if*

$$(F(x) - F(y))^T(x - y) \ge 0, \qquad \forall x, y \in \Re^n,$$

(b) *strongly monotone with modulus $\mu > 0$ if*

$$(F(x) - F(y))^T(x - y) \ge \mu||x - y||^2, \qquad \forall x, y \in \Re^n.$$

(c) *P_0-function, if for any $x, y \in \Re^n$, with $x \ne y$, there exists $i \in I$ such that $x_i \ne y_i$ and*

$$[x_i - y_i][F_i(x) - F_i(y)] \ge 0,$$

(d) *R_0-function, if it has the following property: For any sequence $x^1, x^2, \cdots$ in $\Re^n$ with $||x^k|| \to \infty$ and*

$$\lim_{k\to\infty}\inf \frac{\min\{F_i(x^k) : i \in I\}}{||x^k||} \ge 0, \qquad \lim_{k\to\infty}\inf \frac{\min\{x_i^k : i \in I\}}{||x^k||} \ge 0.$$

There is an index $i_0 \in I$ such that $\limsup_{k\to\infty} F_{i_0}(x^k) = \infty$ *and* $\limsup_{k\to\infty} x_{i_0}^k = \infty$.

A matrix $M \in \Re^{n\times n}$is said to be a P_0-matrix, if for any $0 \ne x \in \Re^n$, there exists an index $i \in I, x_i \ne 0$ such that $x_i(Mx)_i \ge 0$.

If F is continuously differentable in $\Re^n$, then the following facts are known.

(a') F is monotone if and only if the Jacobian $F'(x)$ is positive semidefinite for all $x \in \Re^n$.

(b') F is strongly monotone if and only if $F'(x)$ is uniformly positive definite, i.e.,

$$d^TF'(x)d \ge \mu||d||^2 \quad \text{for some } \mu > 0 \text{ and all } x, d \in \Re^n.$$

(c') F is P_0 function if and only if $F'(x)$ is P_0 matrix for all $x \in \Re^n$.

Also if F is strongly monotone, then it is R_0-function.

3. Unconstrained Reformulation

Throughout this section and the forthcoming one, we assume $\alpha\beta \le 1$. We mainly consider the unconstrained problem (3). Let

$$p(x) = \left(\cdots, \frac{\partial\phi}{\partial a}(x_i, F_i(x)), \cdots\right) \in \Re^n, \tag{10}$$

$$q(x) = \left(\cdots, \frac{\partial\phi}{\partial b}(x_i, F_i(x)), \cdots\right) \in \Re^n. \tag{11}$$

Then the gradient of Ψ is given by

$$\nabla\Psi(x) = p(x) + \nabla F(x)q(x). \tag{12}$$

Here $\nabla F(x)$ is the transpose of Jacobian matrix of F at x. Since the proof of the following result is similar to that in [4] and [17], we omit the proof.

Theorem 3.5 *Assume that NCP is solvable, if F is continuously differentiable and a P_0-function, then a vector $x^* \in \Re^n$ is a solution of* (1) *if and only if x^* is a stationary point of Ψ.*

In order to prove that Ψ has bounded level sets under some assumptions, we need the following Lemma, which is on the limit behavior of ϕ.

Lemma 3.6 *For any sequence $\{(a^k, b^k)\} \subset \Re^2$, the following equivalent relationship holds*

$$\phi(a^k, b^k) \to \infty \Longleftrightarrow a^k \to -\infty \text{ or } b^k \to -\infty \text{ or } [a^k \to +\infty \text{ and } b^k \to +\infty]. \tag{13}$$

Proof. The if part is obvious. To show the only if part, it suffices to show if both $a^k \to -\infty$ and $b^k \to -\infty$ do not occur, then the third case must occur. If not, without loss of generality, we assume that $a^k \to +\infty$ and $|b^k| < L$ for some $L > 0$. Then it follows that

$$\begin{aligned}
\infty = \lim_{k\to\infty} \phi(a^k, b^k) &= \lim_{k\to\infty} \frac{\alpha\beta(a^k)^2(b^k)^2}{\sqrt{(a^k)^2 + \alpha(b^k)^2} + a^k)(\sqrt{\beta(a^k)^2 + (b^k)^2} + b^k)} \\
&= \lim_{k\to\infty} \frac{\alpha\beta(b^k)^2}{\sqrt{1 + \alpha(b^k)^2/(a^k)^2} + 1)(\sqrt{\beta + (b^k)^2/(a^k)^2} + b^k/a^k)} \\
&\le \frac{1}{2}\alpha\sqrt{\beta}L^2 < +\infty.
\end{aligned}$$

This contradiction completes the proof. □

(13) has an equivalent description below. There exists a constant $M > 0$ such that the following relationship hold uniformly in k.

$$\{\phi(a^k, b^k)\} \text{ is bounded } \Longleftrightarrow -M < \min\{a^k, b^k\} < M. \tag{14}$$

The relation (14) has been well used in [18].

Theorem 3.7 *Suppose that F is R_0-function. Let $x^0 \in \Re^n$ be any given vector and $\mathcal{L}(x^0) := \{x \in \Re^n | \ \Psi(x) \le \Psi(x^0)\}$ be the corresponding level set. Then $\mathcal{L}(x^0)$ is compact.*

Proof. Assume the contrary that there is a sequence $\{x^k\} \subset \mathcal{L}(x^0)$ such that $\lim_{k\to\infty} ||x^k|| = \infty$. Since $\{\Psi(x^k)\}$ is bounded, by Lemma 3.2 we have, for some $L > 0$

$$\liminf_{k\to\infty} \min\{F_1(x^k), \cdots, F_n(x^k)\} \ge -L \quad \text{and} \quad \liminf_{k\to\infty} \min\{x_1^k, \cdots, x_n^k\} \ge -L.$$

Dividing all sides of the above inequalities by $||x^k||$, we obtain that the sequence $\{x^k\}$ satisfies the conditions of R_0 function. Hence there is an index $i_0 \in I$ such that $x_{i_0}^k \to +\infty$ and $F_{i_0}(x^k) \to +\infty$ for some subsequence of $\{x^k\}$. Again from Lemma 3.2, we get for this subsequence $\Psi(x^k) \to \infty$, which contradicts the boundedness of $\Psi(x^k)$. Hence $\mathcal{L}(x^0)$ is compact. □

If we apply descent methods to search minimizers of Ψ, the generated sequence $\{x^k\} \subset \Re^n$ will remain in the bounded level set $\mathcal{L}(x^0)$ if F is an R_0-function. We will give such a descent method in the next section, and show its global convergence for monotone complementarity problems.

4. A Descent Method

Throughout this section, we assume F is continuously differentiable and monotone. Let x^k be a iterative point which is not a solution of NCP and

$$d^k = -q(x^k). \tag{15}$$

Then $d^k \ne 0$ by Proposition 2.4. If F is monotone, then $F'(x)$ is positive semidefinite and from (12) and Proposition 2.4, we have

$$\nabla\Psi(x^k)^T d^k = -(p(x^k))^T q(x^k) - (q(x^k))^T \nabla F(x^k) q(x^k) < 0. \tag{16}$$

(16) means that if $\nabla\Psi(x^k)^T d^k = 0$, then $d^k = 0$. If F is strongly monotone, with modulus $\mu > 0$, then we obtain similarly that

$$\nabla\Psi(x^k)^T d^k \le -\mu||d^k||^2.$$

Now we present a descent algorithm for monotone NCP.

Algorithm

S.1 Choose $x^0 \in \Re^n$, $\epsilon > 0$, $0 < \gamma < 1$, $0 < \delta < 1$, let $k_0 = 0$.

S.2 If $\Psi(x^k) \le \epsilon$, stop: x^k is an approximate solution of NCP.

S.3 Define the search direction d^k by (15).

S.5 Compute the steplength $\lambda_k = \gamma^{m_k}$, where m_k is the smallest nonnegative integer satisfying Armijo-type line search rule

$$\Psi(x^k + \gamma^{m_k} d^k) \le \Psi(x^k) - \gamma^{2m_k}\delta||d^k||^2. \tag{17}$$

S.6 Set $x^{k+1} = x^k + \lambda_k d^k$, $k := k+1$ and go to **S.1**.

Algorithm above is also used by Geiger and Kanzow [4]. Note that it does not involve any Jacobian information of F, hence it is a derivative-free algorithm. Such algorithm has its advantage over Jacobian-based ones when the evaluation of Jacobian is time-consuming task. Now we show Algorithm is well defined for monotone NCP.

Lemma 4.8 *For any iterative point x^k, which is not a solution of NCP, then there exists a finite integer m satisfying the line search condition* (17).

Proof. Since x^k is not a solution, then d^k is a descent direction by (16). If there exist no nonnegative integer m satisfying (17), then for any $m > 0$, we have

$$\Psi(x^k + \gamma^m d^k) > \Psi(x^k) - \gamma^{2m}\delta||d^k||^2.$$

Dividing the both sides of the above inequality, and let $m \to \infty$, we get

$$\nabla\Psi(x^k)^T d^k \geq 0,$$

contradicting the fact that d^k is a descent direction. Hence Algorithm is well defined. □

Theorem 4.9 *Let $\{x^k\} \subset \Re^n$ be a sequence generated by Algorithm, then any cluster point of the sequence is a solution of NCP. Moreover if F is strongly monotone, then x^k must converge to the unique solution of NCP.*

Proof. Assume x^* is a cluster point of $\{x^k\}$. Without loss of generality, we assume that $x^k \to x^*$. Since the continuity of $\nabla\phi(x)$, $\{d^k\}$ is bounded and we let $d^k \to d^*$ for some subsequence. Here we denote, for simplicity, the subsequence as $\{x^k\}$. Let $N = \{1, 2, \cdots\}$ and

$$\lambda := \inf\{\lambda_k |\ k \in N\}.$$

If $\lambda > 0$, then

$$\Psi(x^k) - \Psi(x^{k+1}) \geq \lambda^2\delta||d^k||^2.$$

Since the right hand of the above inequality converges to zero, we have

$$\lim_{k\to\infty} ||d^k|| = ||d^*|| = 0,$$

which implies that $q(x^*) = 0$. By Proposition 2.4, x^* must solve NCP.
If $\lambda = 0$, then the sequence $\{m_k\}$ is unbounded. We assume $m_k \to \infty$ for $k \in N_1 \subset N$. Obviously, we have from (16) that

$$\nabla\Psi(x^*)^T d^* \leq 0. \tag{18}$$

By the line search rule, we have

$$\Psi(x^k + \frac{\lambda_k}{\gamma} d^k) - \Psi(x^k) \geq \gamma^{2m_k - 2}\delta||d^k||^2. \tag{19}$$

Dividing the both sides of (19) by (λ_k/γ) and taking limit, we obtain

$$\nabla\Psi(x^*)^T d^* \geq 0. \tag{20}$$

It follows from (18) and (20) that $\nabla\Psi(x^*)^T d^* = 0$, this implies $d^* = 0$. Hence x^* solves NCP. Moreover if F is strongly monotone, NCP has a unique solution, say x^*. Then the whole sequence $\{x^k\}$ must converge to x^*. □

Remarks: If monotone NCP is unsolvable, then the generated sequence must be unbounded and has no accumulation points. We note that the sequence $\{\Psi(x^k)\}$ is monotonically decreasing and nonnegative, Qi in [18] has shown that the direction sequence $\{d^k\}$ is bounded even $\{x^k\}$ is unbounded. This property is useful in studying asymptotic behavior of the unbounded sequence $\{x^k\}$.

5. A Local Property Near Strict Complementarity Solutions

We observe that $-q(x)$ in section 4 functions as a descent direction for monotone NCP under the condition $\alpha\beta \leq 1$. The fact essentially relys on the property in Proposition 2.3. This property may be violated if the condition $\alpha\beta \leq 1$ does not hold. For example, let $\alpha = 100$, $\beta = 1$, $a = 1, b = -1$, we have

$$\frac{\partial\phi}{\partial a}\frac{\partial\phi}{\partial b}|(1,-1) < 0.$$

On the other hand, from computational point of view, when iterative point is near a solution, we hope increase α and β large enough in order for stability of computation. Under such consideration, we hope that $-q(x)$, without any restriction on α and β, can also function as a descent direction near a solution. We show it is true near a strict complementarity solution. We also estimate the radius of the neighborhood, in which $-q(x)$ is a descent direction. Let

$$\mathrm{sign}(t) = \begin{cases} 1 & t > 0, \\ -1 & t < 0. \end{cases}$$

We say $(a^*, b^*) \in \Re^2$ satisfies the complementarity condition provided that $a^* \geq 0, b^* \geq 0, a^*b^* = 0$. (a^*, b^*) is said to satisfy the strict complementarity condition if it satisfies complementarity condition with in addition $a^* + b^* > 0$. Since in this section, we mainly verify a descent property, we only consider points, which do not satisfy complementarity condition. Throughout this section, we only assume $\alpha > 0, \beta > 0$ without any other restriction.

Proposition 5.10 *Let $(a^*, b^*) \in \Re^2$ satisfy the strict complementarity condition. then the following property holds near (a^*, b^*) provided that (a, b) does not satisfy the complementarity condition.*

$$\frac{\partial\phi}{\partial a}(a,b)\frac{\partial\phi}{\partial b}(a,b) > 0. \tag{21}$$

Proof. Note that (a^*, b^*) satisfies the strict complementarity condition. Let $a \to a^*, b \to b^*$. Since the symmetric role of a and b in ϕ, we only consider the case $a^* > 0, b^* = 0$. Obviously $a > 0$. If $b \to 0^+$, then (21) is true by Lemma 2.2. Now we assume $b \to 0^-$, we have

$$\begin{aligned}\frac{\partial \phi}{\partial a}(a,b) &= (\sqrt{a^2+\alpha b^2}-a)\left(\frac{\beta a}{\sqrt{\beta a^2+b^2}}+\frac{b-\sqrt{\beta a^2+b^2}}{\sqrt{a^2+\alpha b^2}}\right) \\ &= (\sqrt{a^2+\alpha b^2}-a)\frac{\beta a(\sqrt{a^2+\alpha b^2}-a)+b\sqrt{\beta a^2+b^2}-b^2}{\sqrt{\beta a^2+b^2}\sqrt{a^2+\alpha b^2}}.\end{aligned}$$

Let

$$\Delta := \beta a(\sqrt{a^2+\alpha b^2}-a)+b\sqrt{\beta a^2+b^2}-b^2 = \Delta_1+\Delta_2+\Delta_3.$$

$$\begin{aligned}\Delta_1 &= \frac{\alpha\beta a b^2}{\sqrt{a^2+\alpha b^2}+a} = O(b^2), \\ \Delta_2 &= b\sqrt{\beta a^2+b^2} = O(b).\end{aligned}$$

Hence

$$\operatorname{sign}\left(\frac{\partial\phi}{\partial a}(a,b)\right) = \operatorname{sign}(\Delta) = \operatorname{sign}(\Delta_1+\Delta_2+\Delta_3) = \operatorname{sign}(b) = -1. \tag{22}$$

On the other hand, for $a > 0, b < 0$ we obviously have

$$\operatorname{sign}\left(\frac{\phi}{\partial b}(a,b)\right) = -1. \tag{23}$$

Then (22) and (23) lead to the property (21). □

In the following, we estimate the radius of the neighborhood, in which property (21) holds. We again assume $a^* > 0, b^* = 0$. By the proof procedure above, we need only guarantee (22) in this neighborhood. Let

$$\delta = \min\left\{\frac{1}{3\alpha\sqrt{\beta}},\ \frac{1}{2}\right\} a^*,$$

and let $B(\delta)$ denote the ball of radius δ centered at (a^*, b^*), i.e.,

$$B(\delta) = \left\{(a,b) \in \Re^2 \,\middle|\, (a-a^*)^2+b^2 \le \delta^2\right\}.$$

Let $(a,b) \in B(\delta)$, then

$$\sqrt{a^2+\alpha b^2}+a \ge a^*, \quad a \le \frac{3}{2}a^*, \quad \sqrt{\beta a^2+b^2} \ge \frac{\sqrt{\beta}}{2}a^*.$$

Thus

$$\begin{aligned}\Delta_1 &\le \frac{3\alpha\beta}{2}b^2, \\ \Delta_2 &\ge \frac{\sqrt{\beta}}{2}a^*|b|.\end{aligned}$$

When $b < 0$, (22) holds since

$$\begin{aligned} \Delta < \Delta_1 + \Delta_2 &\le \frac{3\alpha\beta}{2} b^2 - \frac{\sqrt{\beta}}{2} a^* |b| \\ &= \frac{\alpha\beta}{2} |b| \left(|b| - \frac{1}{3\sqrt{\beta}\alpha} a^* \right) \\ &\le \frac{\alpha\beta}{2} |b| \left(|b| - \delta a^* \right) \\ &\le 0. \end{aligned}$$

Similarly, if $a^* = 0, b^* > 0$, the radius δ should be set by

$$\delta = \min \left\{ \frac{1}{3\sqrt{\alpha\beta}},\ \frac{1}{2} \right\} b^*,$$

in order that (22) holds in this neighborhood. Let $x^* \in \Re^n$ is a strict complementarity solution of NCP, i.e.,

$$x^* \ge 0, \quad F(x^*) \ge 0, \quad (x^*)^T F(x^*) = 0, \quad \text{and} \quad x^* + F(x^*) > 0.$$

Let

$$\delta^* = \min_{1 \le i \le n} \left\{ x_i^* + F_i(x^*) \right\}. \tag{24}$$

Then we have the following descent property.

Theorem 5.11 *Let F be continuously differentiable and monotone. Let x^* be a strict complementarity solution of NCP. Then, if x is not a solution of NCP, $-q(x)$ defined by (11) is a descent direction of Ψ in the neighborhood $N((x^*, F(x^*)), \epsilon^*)$, where*

$$N((x^*, F(x^*)), \epsilon^*) = \left\{ (x, F(x)) \in \Re^{n \times n} \mid \quad ||(x, F(x)) - (x^*, F(x^*))|| \le \epsilon^* \right\},$$

and

$$\epsilon^* = \min \left\{ \frac{1}{3\sqrt{\alpha\beta}},\ \frac{1}{3\alpha\sqrt{\beta}},\ \frac{1}{2} \right\} \delta^*.$$

Proof. Let $(x, F(x)) \in N((x^*, F(x^*)), \epsilon^*)$. Then for any index $i \in K$, $(x_i, F_i(x))$ must satisfy one of the following inequalities,

$$||x_i - x_i^*||^2 + ||F_i(x) - F_i(x^*)||^2 \le \left(\min \left\{ \frac{1}{3\alpha\sqrt{\beta}},\ \frac{1}{2} \right\} \delta^* \right)^2, \tag{25}$$

$$||x_i - x_i^*||^2 + ||F_i(x) - F_i(x^*)||^2 \le \left(\min \left\{ \frac{1}{3\sqrt{\alpha\beta}},\ \frac{1}{2} \right\} \delta^* \right)^2. \tag{26}$$

Then by analysis before the Theorem, we obtain

$$p_i(x) q_i(x) \ge 0.$$

Since x is not a solution, then there exists an index i such that $(x_i, F_i(x))$ does not satisfy the complementarity condition, but does satisfy either (25) or (26). Hence for such i, we have by Proposition 5.10,

$$p_i(x) q_i(x) > 0.$$

This shows that $q(x) \neq 0$. By similar proof as for (16), we claim that $-q(x)$ is a descent direction for Ψ. $\square$

Remark. In fact, Ψ has another local property near a strict complementarity solution x^* without any restriction on α and β. If F is twice continuously differentiable and the gradients of active sets are linearly independent, i.e., $\nabla F_i(x^*)(i \in I^* = \{i | x_i^* > 0\})$ and $e_i(i \notin I^*)$ are linearly independent, where e_i is the i^{th} colum of the $n-$dimensional identity matrix. Then the Hessian matrix $H(x)$ of Ψ near x^* is positive definite. By this local property, we can use quasi-Newton methods such as BFGS method to minimize the function Ψ [15, 16, 17].

6. Numerical Results

Since Algorithm presented in section 3 involves no gradient information of the function F, we expect no rapid convergence. Also we note that even the direction $d(x^k)$ is not a descent direction when F is not monotone. To overcome this drawback, we modify the direction selection procedure (S.3) as following

(S.3') Let $d^k = -q(x^k)$. If $(d^k)^T \nabla \Psi(x^k) \geq 0$, then let $d^k = -\nabla \Psi(x^k)$.

We denote Algorithm with (S.3') by Alg. A. In fact, if d^k is not a descent direction we use negative gradient direction of Ψ. As pointed out in section 5, d^k can be a descent direction when iteration points are near strict complementarity solutions. So Alg. A may fail in solving some hard problems. Alg. B is an implementation of the BFGS method for minimizing Ψ, with the identity as the initial scaling matrix and the step size determined by Armijo rule (S.5). The parameters used are: $\alpha = \beta = 10$, $\epsilon = 10^{-14}$, $\delta = 0.1$ and $\gamma = 0.5$. The initial point is $x^0 = (1, \cdots, 1)^T$. We point out that the BFGS methods for minimizing the function Ψ are implemented in [15, 16, 17] with the stepsizes determined by Wolfe-conditions and with various choices α and β. We report numerical results on test problems selected from literatures.

Problem 1. (Kojima-Shindo [14]) (n=4) This is a nonmonotone NCP.

Problem 2. (Lemke [5]) (n=30) This is monotone LCP.

Problem 3. (Hock-Schittkowski problem 76 [6]) (n=7) This is a LCP and comes from a quadratic programming problem.

Problem 4. (Luo-Tseng PSDLCP [10]) (n=30) $F(x) = Mx + q$ where $M = AA^T$ and every entry of the 30×10 matrix A is uniformly generated from $[-1, 1]$, and $q = -M\bar{x} + \bar{y}$, where each entry of $\bar{x}$ is uniformly generated from $[0, 10]$ and each of $\bar{y}$ is zero if the corresponding entry of $\bar{x}$ is zero and otherwise is uniformly generated from $[0, 10]$. So $\bar{x}$ is typically a degenerate solution.

The numerical results are summarized in the following table, in which iter. res. and val. stand for the number of iteration, natural residual $|| \min\{x^k, F(x^k)\}||$ and the final value of Ψ upon termination respectively.

Problem	Alg. A			Alg. B		
	iter.	res.	val.	iter.	res.	val.
Prob. 1	274	2.646e-8	9.748e-15	41	5.911e-8	3.564e-15
Prob. 2	>10000	—	—	27	2.818e-8	9.83e-15
Prob. 3	325	1.308e-8	7.617e-15	38	1.928e-8	8.617e-15
Prob. 4	> 10000	—	—	189	3.451e-8	7.548e-15

We point out that we can not find solutions of Problems 2 and 4 by Alg. A within 10000 iterations. This may due to that problem 2 is one for which Lemke's method is known to run in exponential time and problem 4 is randomly generated such that it has degenerate solutions.

Acknowledements The first author thanks professor Yaxiang Yuan for his encouragements and helps. Especially he told the author that ϕ, in fact, equals to the following form, which involves one parameter t.

$$\phi_t(a,b) = (\sqrt{a^2+tb^2}-a)(\sqrt{a^2+b^2}-b) \qquad t>0.$$

It is easy to see the equivalence. Let $\alpha\beta = t$ and $c = \sqrt{\beta}a$ in ϕ, then we have

$$\begin{aligned} \phi(a,b) &= (\sqrt{a^2+\frac{t}{\beta}b^2}-a)(\sqrt{\beta a^2+b^2}-b) \\ &= \frac{1}{\sqrt{\beta}}(\sqrt{\beta a^2+tb^2}-\sqrt{\beta}a)(\sqrt{\beta a^2+b^2}-b) \\ &= \frac{1}{\sqrt{\beta}}(\sqrt{c^2+tb^2}-c)(\sqrt{c^2+b^2}-b) \\ &= \phi_t(c,b). \end{aligned}$$

We note that ϕ_t destroys the symmetric role of α and β. The author would like to thank Jiming Peng for his constant help and discussion on this subject.

References

[1] F. Facchinei and C. Kanzow, On unconstrained and constrained stationary points of the implicit Lagrangian, *J. Optim. Theory Appl.*, 92(1997), 493-512.

[2] A. Fischer, An NCP-function and its use for the solution of complementarity problems, in D.Z. Du, L. Qi and Womersly eds., *Recent Advances in Nonsmooth Optimization*, World Scientific, 1995, 261-289.

[3] M. Fukushima, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, *Math. Prog.*, 53(1992), 99-110.

[4] C. Geiger and C. Kanzow, On the resolution of monotone complementarity problems, *Comput. Optim. Appl.*, 5(1996), 155-173.

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

[6] W. Hock and K. Schittkowski, Test examples for nonlinear programming codes, Lecture notes in Economics and Mathematical System 187, Springer-Verlag, Berlin, Germany 1981.

[7] G. Isac, *Complementarity Problems*, Lecture Notes in Mathematics, Springer-Verlag, New York, 1992.

[8] C. Kanzow, Nonlinear complementarity as unconstrained optimization, *J. Optim. Theory Appl.,* 88(1996), 139-155.

[9] C. Kanzow, Global convergence properties of some iterative methods for linear complementarity problems, *SIAM J. Optim.,* 2 (1996), 324-341.

[10] Z.Q. Luo and P. Tseng, A new class of merit functions for nonlinear complementarity problems, *In M. C. Ferris and J. S. Pang (eds.), Complementarity and Variational Problems: State of the Art, SIAM Philadelphia, PA, pp. 204-225, 1997.*

[11] O.L. Mangasarian, Equivalence of the complementarity problem to a system of nonlinear equations, *SIAM J. Appl. Math.,* 31(1976), 89-92.

[12] O.L. Mangasarian and M.V. Solodov, Nonlinear complementarity as unconstrained and constrained minimization, *Math. Prog.*, 62B, 277-297.

[13] J.S. Pang, Complementarity problems, in R. Horst and P. Pardulos eds., *Handbook on Global Optimization*, Klumer Academic Publishers, Norwell, Massachusetts, 1994.

[14] J.S. Pang and S.A. Gabriel, *NE/SQP*: a robust algorithm for the nonlinear complementarity problem, *Math. Oper. Res.,* 60(1993), 295-337.

[15] J.M. Peng, Unconstrained optimization methods for nonlinear complementarity problem, *J. Comput. Math.,* 13(1995), 259-266.

[16] J.M. Peng, Unconstrained methods for generalized nonlinear complementarity and variational inequality problems, *J. Comput. Math.,* 14(1996), 99-107.

[17] H.D. Qi and J.M. Peng, A new unconstrained optimization approach to nonlinear complementarity problems, Technical Report, Institute of Computational Mathematics and Scientific/Engineering Computing, Academia Sinica, Beijing, China, July, 1996.

[18] H.D. Qi, On the minimizing and stationary sequences of a new merit function for complementarity problems, *J. Optim. Theory Appl.*, 102 (1999), 411-431.

[19] N. Yamashita and M. Fukushima, On Stationary points of the Implicit Lagrangian for Nonlinear Complementarity problems, *J. Optim. Theory Appl.,* 84(1995), 653-663.