Journal of Computational Mathematics, Vol.22, No.4, 2004, 509–522.

A CONSTRAINED OPTIMIZATION APPROACH FOR LCP [*1)]

Ju-liang Zhang Jian Chen
(*Department of Management Science and Engineering, School of Economics and Management, Tsinghua University, Beijing 100084, China*)

Xin-jian Zhuo
(*School of Information Engineering, Beijing University of posts and Telecommunication, Beijing 100876, China*)

Abstract

In this paper, LCP is converted to an equivalent nonsmooth nonlinear equation system $H(x, y) = 0$ by using the famous NCP function–Fischer-Burmeister function. Note that some equations in $H(x, y) = 0$ are nonsmooth and nonlinear hence difficult to solve while the others are linear hence easy to solve. Then we further convert the nonlinear equation system $H(x, y) = 0$ to an optimization problem with linear equality constraints. After that we study the conditions under which the K–T points of the optimization problem are the solutions of the original LCP and propose a method to solve the optimization problem. In this algorithm, the search direction is obtained by solving a strict convex programming at each iterative point. However, our algorithm is essentially different from traditional SQP method. The global convergence of the method is proved under mild conditions. In addition, we can prove that the algorithm is convergent superlinearly under the conditions: M is P_0 matrix and the limit point is a strict complementarity solution of LCP. Preliminary numerical experiments are reported with this method.

Mathematics subject classification: 90C30, 65K05.
Key words: LCP, Strict complementarity, Nonsmooth equation system, P_0 matrix, Superlinear convergence.

1. Introduction

Consider the following linear complementarity problem (LCP)

$$\begin{aligned} y = Mx + q, \\ x \geq 0,\ y \geq 0,\ x^T y = 0, \end{aligned} \tag{1}$$

where $M \in R^{n\times n}$, $x, y \in R^n$ and $x \geq 0$ $(y \geq 0)$ means that $x_i \geq 0$ $(y_i \geq 0)$. In this paper, we assume that the solution set of (1) is nonempty. Let X denote the solution set of (1). For convenience, we sometimes use $w = (x, y)$ for $(x^T, y^T)^T$.

LCP has many applications in economic and engineering, see [11, 16, 23] for survey. A lot of experts studied the problem. At present, numerous algorithms were proposed for the problem, for example, interior method (see [33] and references therein), nonsmooth Newton method (see [13, 15, 19, 21, 27]) and smoothing method (see [3, 4, 6, 28] and [8] for survey).

Since the work by Mangasarian [25] it has been well known that by means of a suitable function $\phi : R^2 \to R$, the system

$$a \geq 0,\ b \geq 0,\ ab = 0 \tag{2}$$

[*] Received January 27, 2002; final revised March 27, 2003.
[1)] This work is supported in part by National Science Foundation of China (Grant No.70302003, 10171055 and 70071015) and China Postdoctoral Science Foundation.

can be transformed into an equivalent nonlinear equation

$$\phi(a,b) = 0. \tag{3}$$

In this case, function ϕ is named as NCP-function. Then (1) can be reformulated as the following equivalent nonlinear equation system

$$\Phi(x) = \begin{pmatrix} \phi(x_1,(Mx)_1) \\ \vdots \\ \phi(x_n,(Mx)_n) \end{pmatrix}, \tag{4}$$

or

$$H(x,y) = \begin{pmatrix} \phi(x_1,y_1) \\ \vdots \\ \phi(x_n,y_n) \\ y = Mx - q \end{pmatrix}. \tag{5}$$

Many methods have been proposed to solve (4) or (5) or to minimize their natural residual

$$\Psi_1(x) = \frac{1}{2}\|\Phi(x)\|^2 \quad \text{or} \quad \Psi_2(x,y) = \frac{1}{2}\|H(x,y)\|^2,$$

see [13, 18, 17, 20, 15, 14]. In this paper, we are concerned about formulation (5). Generally speaking, (5) is nonsmooth and nonlinear, hence it is not easy to solve. However, in (5), the first n components are nonsmooth and nonlinear and difficult to solve while the last n components are linear and easy to handle. Therefore, it is reasonable to handle the first part which consists of the n nonsmooth components and the second part which consists of the n linear equations separately. Based on this idea, we transform further (5) into the following equivalent minimization problem

$$\begin{array}{rl} \min_{(x,y)\in R^{2n}} & \Psi(w) = \Psi(x,y) = \frac{1}{2}\sum_{i=1}^{n}\phi(x_i,y_i)^2 \\ \text{s.t.} & y - Mx - q = 0. \end{array} \tag{6}$$

Then we propose an SQP(Sequential Quadratic Programming) type method to solve (6). However, the method is different from the traditional SQP methods. The search direction is obtained by solving the following convex programming at each iterative point

$$\begin{array}{rl} \min_{dw\in R^{2n}} & \theta(dw) = \frac{1}{2}\|Vdw + \phi(x,y)\|_2^2 + \frac{1}{2}\mu\|dw\|_2^2, \\ \text{s.t.} & (-M, I_n)dw = -y + Mx + q, \end{array} \tag{7}$$

where $dw = (dx, dy)$ and $V^T \in \partial\phi(x,y)$, which is a generalized Jacobian of $\phi(w) = \phi(x,y) = \begin{pmatrix} \phi(x_1,y_1) \\ \vdots \\ \phi(x_n,y_n) \end{pmatrix}$ at w and $\mu = \|H(w)\|^\delta$ $(\delta = (0,2])$ and $I_n \in R^{n\times n}$ is the identical matrix. The motivation of using (7) to generate search direction is from the recent results in [12, 30]. Note that (7) is a strict convex quadratical programming, it has the unique solution. Throughout the paper, we shall only use the famous Fischer-Burmeister function defined by

$$\phi(a,b) = \sqrt{a^2+b^2} - a - b, \ (a,b \in R). \tag{8}$$

It has many promised properties and attracted the attention of many researchers [17, 13, 15, 2], see [18] for a survey of its applications.

The paper is organized as follows. In Section 2, we state the algorithm model and its global convergence. in Section 3 we analyze the local convergence properties of the algorithm. In Section 4, Numerical results on some problems are reported. In Section 5, some discussions and conclusions are given.

2. Algorithm Model and Global Convergence

As mentioned in Section 1, we exploit the famous Fischer-Burmeister function defined as

$$\phi(a,b)=\sqrt{a^2+b^2}-a-b. \tag{9}$$

Then (1) can be converted to the following equivalent nonlinear equation system

$$H(w)=H(x,y)=\begin{pmatrix}\phi(x_1,y_1)\\ \vdots\\ \phi(x_n,y_n)\\ y-Mx-q\end{pmatrix}=0. \tag{10}$$

For $\phi(a,b)$ and $H(w)$, we have the following lemmas.

Lemma 2.1 [18,19]**.** *Function ϕ has the following properties:*

(1) $\phi(a,b)=0 \iff a\ge 0,\ b\ge 0,\ ab=0$;

(2) ϕ is Lipschitz continuous with modulus $L=1+\sqrt{2}$, i.e., $|\phi(\omega)-\phi(\omega')|\le L|\omega-\omega'|$ for all $\omega,\ \omega'\in R^2$;

(3) ϕ is directionally differentiable;

(4) ϕ is strongly semismooth on R^2;

(5) ϕ is continuously differentiable on $R^2\backslash(0,0)$;

(6) The generalized gradient $\partial\phi(a,b)$ of ϕ at $(a,b)\in R^2$ equals to

$$\partial\phi(a,b)=\begin{cases}\{(a/\sqrt{a^2+b^2}-1,\, b/\sqrt{a^2+b^2}-1)\} & \text{if } (a,b)\neq(0,0),\\ \{(\xi-1,\zeta-1)\} & \text{if } (a,b)=(0,0),\end{cases}$$

where (ξ,ζ) is any vector satisfying $\sqrt{\xi^2+\zeta^2}\le 1$.

Lemma 2.2. *$H(w)$ has the following properties:*

(1) $H(x^*,y^*)=0 \iff (x^*,y^*)$ *solves (1)*;

(2) $H(w)$ is Lipschitz continuous on R^{2n}, i.e., there exists $L_1>0$ such that

$$\|H(w)-H(w')\|\le L_1\|w-w'\|,\ \forall w,w'\in R^{2n};$$

(3) $H(w)$ is strongly semismooth on R^{2n};

(4) If $X\neq\emptyset$, then there exists $c_1>0$ such that

$$dist(w,X)\le c_1\|H(w)\|,\ \forall w\in B(X,1),$$

where $dist(w,X)=\min\{\|w-w'\|,w'\in X\}$, and $B(X,1)=\{w|dist(w,X)\le 1\}$.

Proof. (1)–(3) follow from Lemma 2.1 and (4) follows from Theorem 2.4 in [18].

If we defined $\psi : R^2 \to R$ and $\Psi(x,y) : R^{2n} \to R$ as follows

$$\psi(a,b) = \frac{1}{2}\phi(a,b)^2,\ a,b \in R;$$

$$\Psi(x,y) = \frac{1}{2}\sum_{i=1}^{n}\phi(x_i,y_i)^2,\ x,y \in R^n,$$

then we have the following lemma

Lemma 2.3 *[19, 20]Functions ψ and Ψ are continuously differentiable on R^2, R^{2n} respectively. Moreover, the following properties are valid for all $a,b \in R$:*

(i) $\nabla_1\psi(a,b) = \nabla_2\psi(a,b) = 0 \iff \psi(a,b) = 0;$

(ii) $\nabla_1\psi(a,b) = \nabla_2\psi(a,b) = 0 \iff \nabla_1\psi(a,b)\nabla_2\psi(a,b) = 0;$

(iii) $\nabla_1\psi(a,b)\nabla_2\psi(a,b) \geq 0.$

As pointed out in Section 1, we are interested in solving problem

$$\begin{array}{rl} \min_{(x,y)\in R^{2n}} & \Psi(x,y) \\ \text{s.t.} & y - Mx - q = 0. \end{array} \tag{11}$$

Obviously (x^*,y^*) solves (1) if and only if (x^*,y^*) solves (11). However, the algorithm we proposed in this paper converges to a K–T point of (11). The first question needed to be answered is what conditions guarantee that a K–T point of (11) is a global solution of (11). First, we have the following lemma.

Lemma 2.4[19]**.** *Let M be P_0 matrix. Furthermore, let vectors $v,\ u \in R^n$ such that $u_iv_i \geq 0$ for all $i = 1,\ldots,n$ and $u_iv_i = 0$ implies $u_i = v_i = 0$ for all $i = 1,\ldots,n$. Then*

$$u + Mv = 0 \textit{ if and only if } u = v = 0.$$

Lemma 2.5. *If M is P_0 matrix, then*

$$w^* = (x^*,y^*) \textit{ solves (1)} \iff w^* \textit{ is a K–T point of (11).}$$

Proof. If w^* is a solution of (1), then w^* is a global minimal of (11). Hence w^* is a K–T point of (11).

Conversely, if $w^* = (x^*,y^*)$ is a K–T point of (11), then there exists $\lambda \in R^n$ such that

$$\nabla\Psi(x^*,y^*) + \begin{pmatrix} -M^T \\ I_n \end{pmatrix}\lambda = 0, \tag{12}$$

$$y^* - Mx^* - q = 0, \tag{13}$$

here $\nabla\Psi(x,y)$ denotes the gradient of Ψ at (x,y).

Note that

$$\nabla\Psi(x^*,y^*) = \begin{pmatrix} \nabla_1\phi(x_1^*,y_1^*) \\ \vdots \\ \nabla_1\phi(x_n^*,y_n^*) \\ \nabla_2\phi(x_1^*,y_n^*) \\ \vdots \\ \nabla_2\phi(x_n^*,y_n^*) \end{pmatrix} = \begin{pmatrix} \nabla_1\Psi(x^*,y^*) \\ \nabla_2\Psi(x^*,y^*) \end{pmatrix},$$

therefore (12) implies that

$$\lambda = -\nabla_2 \Psi(x^*, y^*). \tag{14}$$

It follows from (12) and (14) that

$$\nabla_1 \Psi(x^*, y^*) + M^T \nabla_2 \Psi(x^*, y^*) = 0.$$

By Lemma 2.3 (ii)(iii) and Lemma 2.4, we have

$$\nabla_1 \Psi(x^*, y^*) = \nabla_2 \Psi(x^*, y^*) = 0. \tag{15}$$

Then follows from Lemma 2.3 (i), (15) and the structure of $\nabla_1 \Psi(x, y)$ and $\nabla_2 \Psi(x, y)$ that $\Psi(x^*, y^*) = 0$. By (13) and Lemma 2.1 (i), we know that (x^*, y^*) solves (1).

Now we propose an SQP(Sequential Quadratic Programming) method to solve (11). Note that the constraints in this problem are linear and it is easy to obtain a feasible solution. Hence the initial point is a feasible point. Moreover, each iterative point $w^k = (x^k, y^k)$ is kept feasible. In addition, the search direction at iterative point w^k is obtained by solving that following convex quadratic programming:

$$\begin{array}{rl} \min_{dw \in R^{2n}} & \theta_k(dw) = \frac{1}{2}\|V^k dw + \phi(x^k, y^k)\|^2 + \frac{1}{2}\mu_k \|dw\|^2 \\ \text{s.t.} & dy - M dx = 0, \end{array} \tag{16}$$

where $dw = (dx, dy)$, $\phi(w) = \phi(x, y) = \begin{pmatrix} \phi(x_1, y_1) \\ \vdots \\ \phi(x_n, y_n) \end{pmatrix}$, ${V^k}^T \in \partial \phi(x^k, y^k)$ is a generalized Jacobian of $\phi(w)$ at $w^k = (x^k, y^k)$ and $\mu_k = \|H(w^k)\|^\delta = \|\phi(w^k)\|^\delta$ ($\delta \in (0, 2]$). Clearly, problem (16) is a strict convex quadratic programming. Therefore, it has the unique solution. Furthermore, $dw^k = (dx^k, dy^k)$ is the solution of (16) if and only if there exists $\lambda_k \in R^n$ such that

$$({V^k}^T V^k + \mu_k I_{2n}) dw^k + \begin{pmatrix} -M^T \\ I_n \end{pmatrix} \lambda_k = -{V^k}^T \phi(w^k) = -\nabla \Psi(w^k), \tag{17}$$

$$dy^k - M dx^k = 0, \tag{18}$$

where $I_{2n} \in R^{2n \times 2n}$, $I_n \in R^{n \times n}$ are identical matrices.

Now we state our algorithm formally.

Algorithm 2.1.

step 0. Choose parameters $\gamma \in (0, 1)$, $\alpha \in (0, 1)$, $\beta \in (0, 1)$, $\delta \in (0, 2]$ and initial point (x^0, y^0) satisfying $y^0 = Mx^0 + q$. Set $\mu_0 = \|H(x^0, y^0)\|^\delta$ and k:=0;

step 1. Solve (16) to obtain $dw^k = (dx^k, dy^k)$. If $dw^k = 0$, stop;

step 2. If

$$\|\phi(w^k + dw^k)\| \le \gamma \|\phi(w^k)\| \tag{19}$$

holds, then $w^{k+1} = w^k + dw^k$, go to step 4. Otherwise, go to step 3;

step 3. Let m_k be the smallest nonnegative integer satisfying the following formula

$$\Psi(w^k + \beta^{m_k} dw^k) - \Psi(w^k) \le \alpha \beta^{m_k} \nabla \Psi(w^k)^T dw^k.$$

Set $w^{k+1} = w^k + \beta^{m_k} dw^k$;

step 4. Set $\mu_k = \|H(w^k)\|^\delta$, $k := k + 1$, go to step 1.

Remark. (i) It follows from the definition of the algorithm that $y^k = Mx^k + q$ for all $k = 1, 2, \ldots$.

(ii) If Algorithm 2.1 stops at iterative point w^k, then w^k is a K–T point of (11) by (17) and Remark (i). Hence $w^k = (x^k, y^k)$ is a solution of (1) if M is P_0 matrix.

(iii) Since

$$\begin{aligned} \nabla\Psi(w^k)^T dw^k &= -\left((V^{k^T} V^k + \mu_k I_{2n})dw^k + \begin{pmatrix} -M^T \\ I_n \end{pmatrix} \lambda_k\right)^T dw^k \\ &= -dw^{k^T}((V^{k^T} V^k + \mu_k I_{2n})dw^k \\ &\leq -\mu_k \|dw^k\|^2 < 0, \end{aligned}$$

Algorithm 2.1 is well defined in step 3.

(iv) By [18], the sequence $w^k = (x^k, y^k)$ generated by Algorithm 2.1 is bounded if $X \neq \emptyset$ and $M \in R_0$.

In the remainder of this section, we prove that the algorithm is convergent globally. To this end, we assume that the sequence $\{w^k\}$ generated by Algorithm 2.1 is infinite and bounded.

Theorem 2.1. *Suppose that the sequence $\{w^k\}$ is generated by Algorithm 2.1, then any cluster point of $\{w^k\}$ is a K–T point of (11).*

Proof. By Remark (iii) and step 2, step 3, we know that $\{\Psi(w^k)\}$ is a monotonically decreasing sequence. Note that $\mu_k = \|H(x^k, y^k)\|^\delta = \|\phi(x^k, y^k)\|^\delta = (\Psi(x^k, y^k))^{\frac{\delta}{2}}$, then $\{\mu_k\}$ is monotonically decreasing and bounded from below. Hence it is convergent. If $\mu_k \to 0$, then $H(w^k) \to 0$. Therefore any limit point w^* of $\{w^k\}$ is a solution of (1). So it is a K–T point of (11). If $\lim_{k\to\infty} \mu_k = \bar{\mu} > 0$, then we have

$$\begin{aligned} \nabla\Psi(w^k)^T dw^k &= -\left((V^{k^T} V^k + \mu_k I_{2n})dw^k + \begin{pmatrix} -M^T \\ I_n \end{pmatrix} \lambda_k\right)^T dw^k \\ &= -dw^{k^T}((V^{k^T} V^k + \mu_k I_{2n})dw^k \\ &\leq -\bar{\mu} \|dw^k\|^2 < 0. \end{aligned}$$

It is similar to the standard arguments that we can prove that $dw^k \to 0$. So let w^* be a cluster point of $\{w^k\}$ and $\{w^k\}_{k\in\mathcal{K}}$ converge to w^*. It follows from Lemma 2.1 (6) that $\{V^k\}_{k\in\mathcal{K}}$ is bounded. Without loss of generality, let $\lim_{k\to\infty, k\in\mathcal{K}} V^k = V^*$. The problem

$$\begin{aligned} \min_{dw\in R^{2n}} \quad & \theta(dw) = \tfrac{1}{2}\|V^* dw + \phi(x^*, y^*)\|^2 + \tfrac{1}{2}\bar{\mu}\|dw\|^2 \\ \text{s.t.} \quad & dy - M dx = 0, \end{aligned}$$

has the unique solution $dw^* = 0$. Hence there exists $\lambda^* \in R^n$ such that

$$\begin{pmatrix} -M^T \\ I_n \end{pmatrix} \lambda^* + V^{*^T} \phi(w^*) = 0. \tag{20}$$

From [10], we know that $V^{*^T} \in \partial\phi(w^*)$. So

$$V^{*^T} \phi(w^*) = \nabla\Psi(w^*). \tag{21}$$

It follows from $y^k - Mx^k - q = 0$, $\forall k$ that

$$y^* - Mx^* - q = 0. \tag{22}$$

(20)-(22) imply that $w^* = (x^*, y^*)$ is a K–T point of problem (11).

3. Local Convergence

In order to analyze the local convergence properties, we need the following assumption:

Assumption 3.1. $\{w^k\} \to w^*$, where w^* is a solution of (1) and satisfies strict complement condition, i.e., $x_i^* + y_i^* > 0$, for all $i = 1, \ldots, n$.

By Assumption 3.1, we know that there exists positive integer $K_0 >$ such that

$$x_i^k + y_i^k > 0, \ \forall i = 1, \ldots, n,$$

and

$$(x^k, y^k) \in B(X, 1)$$

for all $k \geq K_0$. Hence follows from Lemma 2.1 (5), Lemma 2.2 (4), Assumption 3.1 and the definition of $\phi(x, y)$ that we have for all $k \geq K_0$

$$\partial\phi(x^k, y^k) = \{\nabla\phi(x^k, y^k)\}, \tag{23}$$

and

$$dist(w^k, X) \leq c_1\|H(w^k)\| = c_1\|\phi(w^k)\|. \tag{24}$$

Therefore

$$V^{k^T} = \nabla\phi(x^k, y^k). \tag{25}$$

In what follows, we assume that $k \geq K_0$.

Let $\bar{w}^k$ denote a vector such that

$$\|w^k - \bar{w}^k\| = dist(w^k, X), \ \bar{w}^k \in X. \tag{26}$$

Note that such $\bar{w}^k$ always exists even though the set X is nonconvex. It follows from Lemma 2.1 (4) and the structure of $\phi(w)$ that $\phi(w)$ is strongly semismooth, i.e., there exists $L_2 > 0$ such that

$$\|\phi(w') - \phi(w) - V(w' - w)\| \leq L_2\|w' - w\|^2, \ \forall V^T \in \partial\phi(w'). \tag{27}$$

First we give several lemmas.

Lemma 3.1. *Suppose that Assumption 3.1 holds and* $\{w^k\}$ *is generated by Algorithm 2.1. If* $w^k \in N\left(w^*, \frac{1}{2}\right) = \{w \| \|w - w^*\| \leq \frac{1}{2}\}$, *then*

$$\|dw^k\| \leq c_2 dist(w^k, X),$$

$$\|V^k dw^k + \phi(w^k)\| \leq c_3(dist(w^k, X))^{1+\frac{\delta}{2}}$$

where $c_2 = \sqrt{c_1^\delta L_2^2 + 1}$, *and* $c_3 = \sqrt{L_1^\delta + L_2^2}$,

Proof. Note that $\bar{w}^k - w^k$ is a feasible solution of (11), then

$$\theta_k(dw^k) \leq \theta_k(\bar{w}^k - w^k). \tag{28}$$

Since $w^k \in N(w^*, \frac{1}{2})$, then

$$\|\bar{w}^k - w^*\| \leq \|\bar{w}^k - w^k\| + \|w^k - w^*\| \leq 2\|w^k - w^*\| \leq 1. \tag{29}$$

So $\bar{w}^k \in B(X, 1)$. By Lemma 2.2 (2) (4) and (26), we know

$$\mu_k = \|H(w^k)\|^\delta \geq \frac{1}{c_1^\delta}\|\bar{w}^k - w^k\|, \tag{30}$$

$$\mu_k = \|H(w^k)\|^\delta = \|H(w^k) - H(\bar{w}^k)\|^\delta \leq L_1^\delta\|w^k - \bar{w}^k\|^\delta. \tag{31}$$

From the definition of θ_k and (27), (28)–(31), we have

$$\begin{array}{ll} & \|dw^k\|^2 \\ \leq & \frac{2}{\mu_k}\theta_k(dw^k) \\ \leq & \frac{2}{\mu_k}\theta_k(\bar{w}^k - w^k) \\ = & \frac{1}{\mu_k}(\|V^k(\bar{w}^k - w^k) + \phi(w^k)\|^2 + \mu_k\|\bar{w}^k - w^k\|^2) \\ = & \frac{1}{\mu_k}(\|\phi(w^k) - V^k(w^k - \bar{w}^k) - \phi(\bar{w}^k)\|^2 + \mu_k\|\bar{w}^k - w^k\|^2) \\ \leq & \frac{1}{\mu_k}(L_2^2\|\bar{w}^k - w^k\|^4 + \mu_k\|\bar{w}^k - w^k\|^2) \\ \leq & (c_1^\delta L_2^2 + 1)\|\bar{w}^k - w^k\|^2. \end{array}$$

Let $c_2 = \sqrt{c_1^\delta L_2^2 + 1}$, then the first equation is obtained.

Now we prove the second equation. It is similar to the first equation we can prove that

$$\begin{array}{rcl} \|V^k dw^k + \phi(w^k)\|^2 & \leq & \theta_k(dw^k) \\ & \leq & \theta_k(\bar{w}^k - w^k) \\ & \leq & L_2^2\|\bar{w}^k - w^k\|^4 + \mu_k\|\bar{w}^k - w^k\|^2 \\ & \leq & (L_2^2 + L_1^\delta)\|\bar{w}^k - w^k\|^{2+\delta}. \end{array}$$

Let $c_3 = \sqrt{L_2^2 + L_1^\delta}$, we obtain the second equation.

Lemma 3.2. *Suppose that Assumption 3.1 holds. If* $w^k,\ w^k + dw^k \in N(w^*, 1) = \{w \mid \|w - w^*\| \leq 1\}$*, then*

$$dist((w^k + dw^k), X) \leq c_4(dist(w^k, X))^{1+\frac{\delta}{2}}.$$

Especially, there exists a constant $b_3 > 0$ *such that*

$$dist(w^k, X) \leq b_3 \Rightarrow dist((w^k + dw^k), X) \leq \frac{1}{2} dist(w^k, X).$$

Proof. Since $\phi(w)$ is twice continuously differentiable at w^k for $k \geq K_0$ by Assumption 3.1 and Lemma 2.1, there exist $K_1 \geq K_0$ and $L_3 > 0$ such that

$$\|\phi(w^k + dw^k) - \phi(w^k) - \nabla\phi(w^k)^T dw^k\| \leq L_3\|dw^k\|^2,\ \forall k \geq K_1. \tag{32}$$

From Lemma 3.1 and (25), we know that for all $k \geq K_1$

$$\|dw^k\| \leq c_2 dist(w^k, X), \tag{33}$$

$$\|\phi(w^k) + \nabla\phi(w^k)^T dw^k\| = \|V^k dw^k + \phi(w^k)\| \leq c_3 dist(w^k, X)^{1+\frac{\delta}{2}}. \tag{34}$$

Then by Lemma 2.2 (4), (32)–(34), we have

$$\begin{array}{rcl} \frac{1}{c_1} dist((w^k + dw^k), X) & \leq & \|H(w^k + dw^k)\| \\ & = & \|\phi(w^k + dw^k)\| \\ & \leq & \|\phi(w^k) + \nabla\phi(w^k) dw^k\| + L_3\|dw^k\|^2 \\ & \leq & c_3 dist(w^k, X)^{1+\frac{\delta}{2}} + L_3 c_2^{1+\frac{\delta}{2}} dist(w^k, X)^{1+\frac{\delta}{2}} \\ & \leq & (c_3 + L_3 c_2^{1+\frac{\delta}{2}}) dist(w^k, X)^{1+\frac{\delta}{2}}. \end{array}$$

Let $c_4 = c_1(c_3 + L_3 c_2^{1+\frac{\delta}{2}})$, we know that the conclusion holds.

Lemma 3.3. *Suppose that Assumption 3.1 holds. Then there exists a positive integer* $\bar{K} \geq K_1$ *such that (19) holds for all* $k \geq \bar{K}$*, i.e., the iteration formula is as follows*

$$w^{k+1} = w^k + dw^k.$$

Proof. Let $r = \min\left\{\frac{1}{2(1+c_2)}, \frac{1}{2c_4}\right\}$. Since w^* satisfies $\phi(w^*) = 0$ and $\phi(w)$ is continuous, there exists a positive integer $\bar{K} \geq K_1$ by Assumption 3.1 such that

$$\|\phi(w^{\bar{K}})\|^{\frac{\delta}{2}} \leq \frac{\gamma}{c_4 L_1 c_1^{1+\frac{\delta}{2}}}, \tag{35}$$

and

$$\|w^{\bar{K}} - w^*\| \leq r. \tag{36}$$

Now we prove that the following statements hold for all $k \geq \bar{K}$:
(i) (19) holds;
(ii) w^k, $w^k + dw^k \in N(w^*, 1)$;
(iii) $w^{k+1} = w^k + dw^k$.
We prove these conclusions by induction.
When $k = \bar{K}$, since

$$\begin{aligned} \|w^k + dw^k - w^*\| &\leq \|w^k - w^*\| + \|dw^k\| \\ &\leq r + c_2 dist(w^k, X) \\ &\leq r + c_2\|w^k - w^*\| \\ &\leq (1 + c_2)r \\ &\leq 0.5, \end{aligned}$$

(ii) holds.
Let $\hat{w}^k \in X$ such that

$$\|(w^k + dw^k) - \hat{w}^k\| = dist(w^k + dw^k, X).$$

It is similar to (29) that we can prove that $\hat{w}^k \in N(w^*, 1)$. Then by Lemma 3.2, Lemma 2.2 and (35), we have

$$\begin{aligned} \|\phi(w^k + dw^k)\| &= \|H(w^k + dw^k)\| \\ &= \|H(w^k + dw^k) - H(\hat{w}^k)\| \\ &\leq L_1 dist(w^k + dw^k, X) \\ &\leq L_1 c_4 dist(w^k, X)^{1+\frac{\delta}{2}} \\ &\leq L_1 c_4 c_1^{1+\frac{\delta}{2}} \|H(w^k)\|^{1+\frac{\delta}{2}} \\ &= L_1 c_4 c_1^{1+\frac{\delta}{2}} \|\phi(w^k)\|^{1+\frac{\delta}{2}} \\ &\leq \gamma\|\phi(w^k)\|. \end{aligned} \tag{37}$$

So (i) holds. Therefore (iii) holds by the definition of the algorithm.

Now we assume that (i) (ii) (iii) hold for $k = \bar{K},\ \bar{K}+1,\ \ldots, l$. We need prove that (i) (ii) (iii) hold for $k = l + 1$.

Obviously, $w^{k+1} \in N(w^*, 1)$, $\forall k = \bar{K},\ \bar{K}+1,\ \ldots, l$. From assumption, (ii) (iii) and Lemma 3.2, we know that for all $k = \bar{K},\ \bar{K}+1,\ \ldots, l$

$$\begin{aligned} & dist(w^k, X) \leq c_4 dist(w^{k-1}, X)^{1+\frac{\delta}{2}} \leq \ldots \leq c_4^{(1+\frac{\delta}{2})^{k-\bar{K}}-1} \|w^{\bar{K}} - \bar{w}^{\bar{K}}\|^{(1+\frac{\delta}{2})^{k-\bar{K}}} \\ \leq\ & c_4^{(1+\frac{\delta}{2})^{k-\bar{K}}-1} \|w^{\bar{K}} - w^*\|^{(1+\frac{\delta}{2})^{k-\bar{K}}} \leq r(\tfrac{1}{2})^{(1+\frac{\delta}{2})^{k-\bar{K}}-1}. \end{aligned}$$

Hence, by Lemma 3.1, we have that for all $k = \bar{K},\ \bar{K}+1,\ \ldots, l$

$$\|dw^k\| \leq c_2 dist(w^k, X) \leq c_2 r(\frac{1}{2})^{(1+\frac{\delta}{2})^{k-\bar{K}}-1}.$$

Then we have

$$\begin{array}{rcl} \|w^{l+1}+dw^{l+1}-w^*\| & \leq & \|w^{\bar{K}}-w^*\|+\sum_{k=\bar{K}}^{l+1}\|dw^k\| \\ & \leq & r+c_2 r\sum_{k=\bar{K}}^{l+1}(\frac{1}{2})^{(1+\frac{\delta}{2})^{k-\bar{K}}-1} \\ & \leq & r+c_2 r\sum_{k=1}^{\infty}(\frac{1}{2})^{(1+\frac{\delta}{2})^{k-\bar{K}}-1} \\ & \leq & (1+c_2)r \\ & \leq & \frac{1}{2}. \end{array}$$

Then (ii) holds.

Since $\{\|\phi(w^k)\|\}$ is monotonically decreasing,

$$\|\phi(w^{l+1})\| \leq \ldots \leq \|\phi(w^{\bar{K}})\| \leq \frac{\gamma}{c_4 L_1 c_1^{1+\frac{\delta}{2}}}.$$

It is similar to (37) that we can prove that

$$\|\phi(w^{l+1}+dw^{l+1})\| \leq \gamma\|\phi(w^{l+1})\|.$$

Hence (i) holds. So (iii) holds by the definition of the algorithm.

Combining Lemma 3.2 and Lemma 3.3, we have the following theorem.

Theorem 3.1. *Suppose that Assumption 3.1 holds and $\{w^k\}$ is generated by Algorithm 2.1. Then $\{dist(w^k,X)\}$ converges to 0 superlinearly. If $\delta=2$, then $\{dist(w^k,X)\}$ converges to 0 quadratically.*

Proof. By Lemma 3.3, for all $k \geq \bar{K}$, iteration formula is as follows

$$w^{k+1}=w^k+dw^k,$$

and

$$w^k,\ w^k+dw^k \in N(w^*,1).$$

The conclusion follows from Lemma 3.2.

4. Implementation and Numerical Experiments

In this section, we test our algorithm's efficiency on some typical test problems. The program was written in MATLAB and runs in MATLAB 6.0 environment. However, we do not solve directly problem (16) to obtain the search direction. We consider the following equivalent unconstrained convex optimization

$$\min_{dx\in R^n} 0.5dx^T\left((I_n,M)({V^k}^T V^k+\mu_k I_{2n})\left(\begin{array}{c} I_n \\ M \end{array}\right)\right)dx+\phi(x^k)V^k\left(\begin{array}{c} I_n \\ M \end{array}\right)dx. \quad (38)$$

Note that (38) equals to the following equation

$$\left((I_n,M)({V^k}^T V^k+\mu_k I_{2n})\left(\begin{array}{c} I_n \\ M \end{array}\right)\right)dx+(I_n,M)V^k\phi(x^k)=0, \quad (39)$$

we solve equation (39) to obtain dx, then let $dy=Mdx$. In this way, we obtain the search direction. At each iterative point, we obtain the search direction by solving a system of linear equations. Since the system of linear equations is symmetric positive definite, the computation is less. The parameters are chosen as follows $\gamma=0.9,\ \alpha=0.1,\ \beta=0.5,\ \delta=1$. The stop criterion is $\|dw\| \leq 10^{-10}$. The numerical results are summarized in Table 1 and the test problems are introduced as follows.

LCP1: $M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $q = (-1, -1)$. This problem is given in Cottle et [11], the initial point is $(0, \cdots, 0)$.

LCP2: $M = \begin{pmatrix} 0 & -1 & 2 \\ 2 & 0 & -2 \\ -1 & 1 & 0 \end{pmatrix}$, $q = (-3, 6, -1)$. This problem is given in Cottle et [11], the initial point is $(0, \cdots, 0)$.

LCP3: $M = \begin{pmatrix} 0 & 0 & 10 & 20 \\ 0 & 0 & 30 & 15 \\ 10 & 20 & 0 & 0 \\ 30 & 15 & 0 & 0 \end{pmatrix}$, $q = (-1, -1, -1, -1)$.This problem is given in Cottle et [11], the initial point is $(0, \cdots, 0)$.

LCP4: $M = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 2 & \cdots & 2 \\ 0 & 0 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$, $q = -e$, $n = 16$.This linear complementarity problem is one for which Murty has shown that Lemke's complementary pivot algorithm is known to run in a number of pivots exponential in the number of variables in the problem (see [26]). The initial point is $(0, \cdots, 0)$.

LCP5: $M = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 & 2 \\ 0 & 1 & 2 & \cdots & 2 & 2 \\ 0 & 0 & 1 & \cdots & 2 & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 2 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$, $q = -(1, \cdots, 1, 0)$. This problem is given in Chen and Ye [7], the initial point is $(0, \cdots, 0)$.

LCP6: $M = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}$, $q = (1, 0, -1)$. This problem is from Yamashita, Dan and Fukushima [31]. The initial point is $(0, \cdots, 0)$.

LCP7: $M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}$, $q = (0, -1, 0)$. This problem is from Yamashita, Dan and Fukushima [31]. The initial point is $(0, \cdots, 0)$.

LCP8: $M = \begin{pmatrix} 4 & 2 & 2 & 1 \\ 2 & 4 & 0 & 1 \\ 2 & 0 & 2 & 2 \\ -1 & -1 & -2 & 0 \end{pmatrix}$, $q = (-8, -6, -4, 3)$. This problem is from Yamashita and Fukushima [32]. The initial point is $(0, \cdots, 0)$.

LCP9: $M = \begin{pmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{pmatrix}$, $q = (0, 0, 0, 0)$.This problem is from Yamashita and Fukushima [32]. The initial point is $(1, \cdots, 1)$.

LCP10: $M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$, $q = (0, 0, 1)$. This problem is from Chen and Ye [7]. The initial point is $(1, \cdots, 1)$.

Table 1: *Dim.* is the dimension n of the problem, *No. of the iter.* is the number of the iterations and *Residual* is $\|\phi(x,y)\|$.

Problem	Dim.	No. of Iter.	Residual
LCP1	2	8	1.2e-13
LCP2	3	7	5.8e-15
LCP3	4	9	7.9e-15
LCP4	16	35	1.1e-12
LCP5	100	26	2.7e-13
LCP5	300	42	1.3e-14
LCP6	3	8	1.6e-14
LCP7	3	8	2.7e-19
LCP8	4	20	1.3e-14
LCP9	4	30	5.2e-12
LCP10	3	10	4.0e-12
LCP11	3	10	4.3e-17
LCP12	300	19	3.8e-13
LCP12	500	22	1.1e-11
LCP13	300	21	2.1e-17
LCP13	500	24	1.3e-11

LCP11: $M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 1 \end{pmatrix}$, $q = (0,0,1)$. This problem is from Zhao and Li [34]. The initial point is $(1,\cdots,1)$.

LCP12: $M = \begin{pmatrix} 4 & -2 & 0 & \cdots & 0 & 0 \\ 1 & 4 & -2 & \cdots & 0 & 0 \\ 0 & 1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -2 \\ 0 & 0 & 0 & \cdots & 1 & 4 \end{pmatrix}$, $q = -e$. This problem is from Ahn [1]. The initial point is $(0,\cdots,0)$.

LCP13: $M = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 4 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 4 \end{pmatrix}$, $q = -e$. This problem is from Geiger and Kanzow [20]. The initial point is $(0,\cdots,0)$.

From Table 1, we note that the algorithm can solve these problems. For some problems, for example LCP 5, LCP 12 and LCP 13, the method solves them fast. However, for some problems, the algorithm is worse. For example, for LCP 4, the method in [22] is very efficient while our method solves the problem with the number of the iterations as twice as the dimension. During the experiment, we observe that the algorithm converges fast even though the solution is degenerate. However, there is no common knowledge on the choice for δ. For some problems, the lager δ is, the better the algorithm performs, whereas for other problems, the smaller δ is , the better the algorithm performs.

5. Conclusion

In this paper, we propose a new method for LCP. The conditions guaranteeing the global convergence of the algorithm are mild. Furthermore, we prove that the algorithm is superlinearly convergent under the condition that M is P_0 and one of the cluster points of sequence generated by the algorithm is strict complementarity. We know that Yamashita and Fukushima in [29] obtained the same results. Our algorithm is different from the algorithm in [29]. Here we use Fischer–Burmeister function, which performs efficiently in practice. The essential difference between our algorithm and algorithm in [29] is that they applied LMM to (4) directly while we convert (5) into equivalent constrained optimization (11). Furthermore, we report numerical results. Numerical experiments show that the performance of the algorithm is notable.

References

[1] Ahn, B.H., Iterative methods for linear complementarity problem with upperbounds and lowerbounds, *Math. Prog.*, **26** (1983), 265-315.

[2] Chen, B., Chen, X., Kanzow, C., A penalized Fischer-Burmeister NCP-function: Theoretical investigation and numerical results, *Math. Prog.*, **88** (2000), 211-216.

[3] Chen, B., Chen, X., A globally and locally superlinear comtinuation-smoothing method for P_0+R_0 and monotone NCP, *SIAM J. Optim.*, **9** (1999), 624-645.

[4] Chen, B., Harker, P.T., A non-interior-point comtinuation method for linear complementarity problems, *SIAM J. Matrix Anal. Appl.*, **14** (1993), 1168-1190.

[5] Chen, B., Xiu, N., A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing function, *SIAM J. Optim.*, **9** (1999), 605-623.

[6] Chen, B., Xiu, N., A superlinear noninterior one-step continuation method for monotone LCP in the absence of strict complementarity, *Journal of Optimization Theory and Application*, **108** (2001), 317-332.

[7] Chen, X., Ye, Y., On smoothing methods for the P_0 matrix linear complementarity problem, *SIAM J. Optim.*, **11** (2000), 341-363.

[8] Chen, X., Smoothing methods for complementarity problems and their applications: A survey, *Journal of the Operation Reseach Society of Japan*, **43** (2000), 32-47.

[9] Chen, C., Mangasarian, O.L., Smoothing methods for convex inequalities and linear complementarity problems, *Math. Prog.*, **71** (1995), 51-69.

[10] Clark, F.H., Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[11] Cottle, R.W., Pang, J.-S., Stone, R.W., The Linear Complementarity Problem, Academic, New York, NY, 1992.

[12] Dan, H., Yamashita, N., Fukushima, M., Convergence Properties of the inexact Levenberg-Marquardt method under local error bound conditions, Technical Report 2001-003, Department of Applied Mathematics and Physics, Kyoto University, (Jannuary 2001).

[13] De Luca, T., Fancchinei, F., Kanzow, C., A semismooth equation approach to the solution of nonlinear complementarity problems, *Math. Prog.*, **75** (1996), 407-439.

[14] Fachinei, F., Soares, A., A new merit function for nonlinear complementarity problems and a related algorithm, *SIAM J. Optim.*, **7** (1997), 225-247.

[15] Facchinei, F., Kanzow, C., A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems, *Math. Prog.*, **76** (1997), 493-512.

[16] Ferris, M.C., Pang, J.-S, Engineering and economic application of complementarity problems, *SIAM J. Review*, **39** (1997), 669-713.

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

[18] Fischer, A., An NCP-function and its use for the solution of complementarity problems, In D.Z., Du, L., Qi and R.S., Womersley eds, Recent Advences in Nonsmooth Optimization, Kluwer Academic Publishers, (1995), 88-105.

[19] Fischer, A., Solution of monotone complementarity problems with locally lipschitzian functions, *Math. Prog.*, **76** (1997), 513-532.

[20] Geiger, C., Kanzow, K., On the resolution of monotone complementarity problems, *Computational Optimization and Applications*, **5** (1996), 155-173.

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

[22] He, B.S., A modified projection and contraction method for a class of linear complementarity problem, *Journal of Computational Mathematics*, **14**:1 (1996), 54-63.

[23] Isac, G., Complementarity Problems, Lecture note in Mathematics, Vol. 1528, Springer-Verlay, Berlin, Heidelberg, 1995.

[24] Kanzow, C., Some noninterior continuation methods for linear complementarity problems, *SIAM J. Matrix Anal. Appl.*, **17** (1996), 851-868.

[25] Mangasarian, O.L., Equivalence of complemetarity to a system of nonlinear equations, *SIAM J. on Applied Mathematics*, **61** (1976), 89-92.

[26] Murty, K.G., Linear Complementarity, Linear and Nonlinear Programming, Helderman-Verlag, Berlin, 1988.

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

[28] Qi, L., Sun, D., Zhou, G., A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained ariational inequalities, *Math. Prog.*, **87** (2000), 1-35.

[29] Yamashita, N., Fukushima, M., On the rate of convergence of the Levenberg-Marquardt method, *Computing,* **15** (2001), 239-249.

[30] Yamashita, N., Dan, H., Fukushima, M., On the identification of degenerate indices in the nonlinear complementarity problem with the proximal point algorithm, Technical Report, Department of Applied Mathematics and Physics, Kyoto University (2001).

[31] Yamashita, N., Fukushima, M., Modified Newton methods for solving a semismooth reformulation of monotone complementarity, *Math. Prog.*, **76** (1997), 469-491.

[32] Yamashita, N., Fukushima, M., The proximal point algorithm with genuine superlinear convergence for the monotone complementarity problem, *SIAM J. Optim.* (to appear)

[33] Yoshise, A., Complementarity problems, In: Terlaky, T., eds. Interior point methods of mathematical programming, Kluwer Academic Publishers, 1996.

[34] Zhao, Y., Li, D., A global and locally superlinearly convergent non-interior-point algorithm for P_0 LCPs, Tech. Report, Department of Systems Engineering and Engineering Management, The Chinese University of HongKong, 2001.