

RESEARCH

Open Access



A smoothing inexact Newton method for variational inequalities with nonlinear constraints

Zhili Ge^{1,2*}, Qin Ni¹ and Xin Zhang³

*Correspondence:

gezhiyly66@126.com

¹College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China

²Basic Sciences Department, Nanjing Polytechnic Institute, Nanjing, 210048, China

Full list of author information is available at the end of the article

Abstract

In this paper, we propose a smoothing inexact Newton method for solving variational inequalities with nonlinear constraints. Based on the smoothed Fischer-Burmeister function, the variational inequality problem is reformulated as a system of parameterized smooth equations. The corresponding linear system of each iteration is solved approximately. Under some mild conditions, we establish the global and local quadratic convergence. Some numerical results show that the method is effective.

Keywords: variational inequalities; nonlinear constraints; inexact Newton method; global convergence; local quadratic convergence

1 Introduction

We consider the variational inequality problem (VI for abbreviation), which is to find a vector $x^* \in \Omega$ such that

$$\text{VI}(\Omega, F) \quad (x - x^*)^\top F(x^*) \geq 0, \quad \forall x \in \Omega, \quad (1)$$

where Ω is a nonempty, closed and convex subset of \mathcal{R}^n and F is a continuous differentiable mapping from \mathcal{R}^n into \mathcal{R}^n . In this paper, without loss of generality, we assume that

$$\Omega := \{x \in \mathcal{R}^n \mid g(x) \geq 0\}, \quad (2)$$

where $g : \mathcal{R}^n \rightarrow \mathcal{R}^m$ and $g_i : \mathcal{R}^n \rightarrow \mathcal{R}$, ($i \in I = \{1, 2, \dots, m\}$) are twice continuously differentiable concave functions. When $\Omega = \mathcal{R}_+^n$, VI reduces to the nonlinear complementarity problem (NCP for abbreviation)

$$x^* \in \mathcal{R}_+^n, \quad F(x^*) \in \mathcal{R}_+^n, \quad x^{*\top} F(x^*) = 0. \quad (3)$$

Variational inequalities have important applications in mathematical programming, economics, signal processing, transportation and structural analysis [1–3]. So, there are various numerical methods which have been studied by many researchers; e.g., see [4].

A popular way to solve the $\text{VI}(\Omega, F)$ is to reformulate (1) to a nonsmooth equation via a KKT system of variational inequalities and an NCP-function. It is well known that the

KKT system of $VI(\Omega, F)$ can be given as follows:

$$\begin{cases} F(x) - \nabla g(x)^\top \lambda = 0, \\ g(x) - z = 0, \\ \lambda \geq 0, z \geq 0, \quad \lambda^\top z = 0, \end{cases} \tag{4}$$

and the NCP-function $\phi(a, b)$ is defined by the following condition:

$$\phi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0. \tag{5}$$

Then problem (1) and (2) is equivalent to the following nonsmoothing equation:

$$\begin{pmatrix} F(x) - \nabla g(x)^\top \lambda \\ g(x) - z \\ \phi(\lambda, z) \end{pmatrix} = 0. \tag{6}$$

Hence, problem (1) and (2) can be translated into (6).

We all know that the smoothing method is a fundamental approach to solve the nonsmooth equation (6). Recently, there has been strong interest in smoothing Newton methods for solving NCP [5–12]. The idea of this method is to construct a smooth function to approximate $\phi(\lambda, z)$. In the past few years, there have been many different smoothing functions which were employed to smooth equation (6). Here, we define

$$H(\mu, x, \lambda, z) := \begin{pmatrix} \mu \\ F(x) - \nabla g(x)^\top \lambda \\ g(x) - z \\ \Phi(\mu, \lambda, z) \end{pmatrix}, \tag{7}$$

where

$$\Phi(\mu, \lambda, z) := \begin{pmatrix} \varphi(\mu, \lambda_1, z_1) \\ \vdots \\ \varphi(\mu, \lambda_m, z_m) \end{pmatrix} \tag{8}$$

and

$$\varphi(\mu, a, b) = a + b - \sqrt{a^2 + b^2 + \mu^2}, \quad \forall (\mu, a, b) \in \mathbb{R}^3. \tag{9}$$

It follows from equations (4)-(9) that $H(\mu, x, \lambda, z) = 0$ is equivalent to $\mu = 0$ and (x, λ, z) is the solution of (6). Thus, we may solve the system of smoothing equation $H(\mu, x, \lambda, z) = 0$ and reduce μ to zero gradually while iteratively solving the equation.

Based on the above symmetric perturbed Fischer-Burmeister function (9), Chen et al. [13] proposed the first globally and superlinearly convergent smoothing Newton method. They dealt with general box constrained variational inequalities. And Rui et al. [14] proposed an inexact Newton-GMRES method for a large-scale variational inequality problem under the assumption of linear inequality constraints.

In reality, variational inequalities with nonlinear constraints are more attractive. These problems have wide applications in economic networks [15], image restoration [3, 16] and so on. So, in this paper, under the framework of smoothing Newton method, we propose a new inexact Newton method for solving $VI(\Omega, F)$ with nonlinear constraints, which extends the scope of constraints. We also prove the global and local quadratic convergence and present some numerical results which show the efficiency of the proposed method.

Throughout this paper, we always assume that the solution set of problem (1) and (2), denoted by Ω^* , is nonempty. \mathcal{R}_+ and \mathcal{R}_{++} mean the nonnegative and positive real sets. Symbol $\|\cdot\|$ stands for the 2-norm.

The rest of this paper is organized as follows. In Section 2, we summarize some useful properties and definitions. In Section 3, we describe the inexact Newton method formally and then prove its local quadratic convergence. We also give global convergence in Section 4. In Section 5, we report our numerical results. Finally, we give some conclusions in Section 6.

2 Preliminaries

In this section, we denote some basic definitions and properties that will be used in the subsequent sections.

Definition 2.1 The operator F is monotone if, for any $u, v \in \mathcal{R}^n$,

$$(u - v)^\top (F(u) - F(v)) \geq 0;$$

F is strongly monotone with modulus $\mu > 0$ if, for any $u, v \in \mathcal{R}^n$,

$$(u - v)^\top (F(u) - F(v)) \geq \mu \|u - v\|^2;$$

F is Lipschitz continuous with a positive constant $L > 0$ if, for any $u, v \in \mathcal{R}^n$,

$$\|F(u) - F(v)\| \leq L \|u - v\|.$$

The following lemma gives some properties of H and its corresponding Jacobian.

Lemma 2.1 Let $H(\mu, x, \lambda, z)$ be defined by (7). Assume that F is continuously differentiable and strongly monotone, g is twice continuously differentiable concave, $(\mu^*, x^*, \lambda^*, z^*)$ in $\mathcal{R}_+ \times \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^m$ is the solution of $H(\mu, x, \lambda, z) = 0$, the rows of $\nabla g(x^*)$ are linearly independent and (λ^*, z^*) satisfies the strict complementarity condition. Then

- (i) $H(\mu, x, \lambda, z)$ is continuously differentiable on $\mathcal{R}_+ \times \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^m$.
- (ii) $\nabla H(\mu^*, x^*, \lambda^*, z^*)$ is nonsingular, where

$$\nabla H(\mu, x, \lambda, z) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \nabla F(x) - \nabla^2 g(x)^\top \lambda & -\nabla g(x)^\top & 0 \\ 0 & \nabla g(x) & 0 & -I \\ D_\mu & 0 & D_\lambda & D_z \end{pmatrix}, \tag{10}$$

$$D_\mu = \text{vec} \left(-\frac{\mu}{\sqrt{\lambda_i^2 + z_i^2 + \mu^2}} \right),$$

$$D_\lambda = \text{diag}\left(1 - \frac{\lambda_i}{\sqrt{\lambda_i^2 + z_i^2 + \mu^2}}\right) \quad (i = 1, \dots, m),$$

$$D_z = \text{diag}\left(1 - \frac{z_i}{\sqrt{\lambda_i^2 + z_i^2 + \mu^2}}\right) \quad (i = 1, \dots, m) \quad \text{and}$$

$$\nabla^2 g(x)^\top \lambda = \sum_{j=1}^m \nabla^2 g_j(x)^\top \lambda_j.$$

Proof It is not hard to show that H is continuously differentiable on $\mathcal{R}_{++} \times \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^m$. From $H(\mu^*, x^*, \lambda^*, z^*) = 0$, by (7) we get that $\mu^* = 0$ easily. Since (λ^*, z^*) satisfies the strict complementarity condition, i.e., λ_i^*, z_i^* are not equal to 0 at the same time, we have that H is also continuously differentiable on $(\mu^*, x^*, \lambda^*, z^*)$. That is, (i) holds.

Now, we prove (ii). Let $q = (q_1, q_2, q_3, q_4)^\top \in \mathcal{R}_+ \times \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^m$, $q_1 = (q_{11})$, $q_2 = (q_{21}, q_{22}, \dots, q_{2n})$, $q_3 = (q_{31}, q_{32}, \dots, q_{3m})$, $q_4 = (q_{41}, q_{42}, \dots, q_{4m})$, and

$$\nabla H(\mu^*, x^*, \lambda^*, z^*) \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix} = 0. \tag{11}$$

Hence, we have

$$\begin{cases} q_1 = 0, & (12) \\ (\nabla F(x^*) - \nabla^2 g(x^*)^\top \lambda^*)q_2 - \nabla g(x^*)^\top q_3 = 0, & (13) \\ \nabla g(x^*)q_2 - q_4 = 0, & (14) \\ D_{\lambda^*}q_3 + D_{z^*}q_4 = 0, & (15) \end{cases}$$

where

$$D_{\lambda^*} = \text{diag}\left(1 - \frac{\lambda_i^*}{\sqrt{\lambda_i^{*2} + z_i^{*2}}}\right), \quad D_{z^*} = \text{diag}\left(1 - \frac{z_i^*}{\sqrt{\lambda_i^{*2} + z_i^{*2}}}\right).$$

We can observe $q_1 = 0$ easily by (12).

Next, we discuss formula (15). The full form of (15) can be described as follows:

$$\begin{pmatrix} (1 - \frac{\lambda_1^*}{\sqrt{\lambda_1^{*2} + z_1^{*2}}})q_{31} + (1 - \frac{z_1^*}{\sqrt{\lambda_1^{*2} + z_1^{*2}}})q_{41} \\ (1 - \frac{\lambda_2^*}{\sqrt{\lambda_2^{*2} + z_2^{*2}}})q_{32} + (1 - \frac{z_2^*}{\sqrt{\lambda_2^{*2} + z_2^{*2}}})q_{42} \\ \vdots \\ (1 - \frac{\lambda_m^*}{\sqrt{\lambda_m^{*2} + z_m^{*2}}})q_{3m} + (1 - \frac{z_m^*}{\sqrt{\lambda_m^{*2} + z_m^{*2}}})q_{4m} \end{pmatrix} = 0. \tag{16}$$

According to the strict complementarity condition of (λ^*, z^*) , we have $\lambda_i^* \geq 0, z_i^* \geq 0, \lambda_i^* \top z_i^* = 0$ and λ_i^*, z_i^* are not equal to 0 at the same time. If $z_i^* = 0$, then $\lambda_i^* > 0$, and

$$\left(1 - \frac{\lambda_i^*}{\sqrt{\lambda_i^{*2} + z_i^{*2}}}\right) = 0, \quad \left(1 - \frac{z_i^*}{\sqrt{\lambda_i^{*2} + z_i^{*2}}}\right) = 1.$$

From (16) we get that $q_{4i} = 0$ and $q_{3i}^\top q_{4i} = 0$. Similarly, if $\lambda_i^* = 0$, then $z_i^* > 0$. We get that $q_{3i} = 0$ and $q_{3i}^\top q_{4i} = 0$. Hence, $q_3^\top q_4 = 0$.

Multiplying the equation of (14) by q_3^\top on the left-hand side and using $q_3^\top q_4 = 0$, we have

$$q_3^\top \nabla g(x^*) q_2 = 0. \tag{17}$$

Multiplying the equation of (13) by q_2^\top on the left-hand side and using (17), we have

$$q_2^\top (\nabla F(x^*) - \nabla^2 g(x^*)^\top \lambda^*) q_2 = 0. \tag{18}$$

Meanwhile, because F is strongly monotone, we have that $\nabla F(x^*)$ is a positive definite matrix. Besides, since g is concave and λ^* is nonnegative, we have that $\nabla^2 g(x^*)^\top \lambda^* = \sum_{j=1}^m \nabla^2 g_j(x^*)^\top \lambda_j^*$ are nonpositive definite matrices, which implies that $(\nabla F(x^*) - \nabla^2 g(x^*)^\top \lambda^*)$ is a positive definite matrix. So, we get that $q_2 = 0$ by (18).

Substituting $q_2 = 0$ into (13) and using the rows of $\nabla g(x^*)$ are linearly independent, we get $q_3 = 0$. Substituting $q_2 = 0$ into (14), we get $q_4 = 0$. Hence we have $q = 0$, which implies that $\nabla H(\mu^*, x^*, \lambda^*, z^*)$ is nonsingular. This completes the proof. \square

3 The inexact algorithm and its convergence

We are now in the position to describe our smoothing inexact Newton method formally by using the smoothed Fischer-Burmeister function (9) to solve the variational inequalities with nonlinear constraints. We also show that this method has local quadratic convergence.

Algorithm 3.1 (Inexact Newton method)

Step 0. Let $w = (x, \lambda, z)$ and $(\mu^0, w^0) \in \mathcal{R}_{++} \times \mathcal{R}^{n+2m}$ be an arbitrary point. Choose $\mu^0 > 0$ and $\gamma \in (0, 1)$ such that $\gamma \mu^0 < \frac{1}{2}$.

Step 1. If $\|H(\mu^k, w^k)\|^2 = 0$, then stop.

Step 2. Compute $(\Delta \mu^k, \Delta w^k)$ by

$$H(\mu^k, w^k) + \nabla H(\mu^k, w^k) \begin{pmatrix} \Delta \mu^k \\ \Delta w^k \end{pmatrix} = \begin{pmatrix} \rho_k \mu^0 \\ r^k \end{pmatrix}, \tag{19}$$

where $\rho_k = \rho(\mu^k, w^k) := \gamma \min\{1, \|H(\mu^k, w^k)\|^2\}$, and $r^k \in \mathcal{R}^{n+2m}$ such that $\|r^k\| \leq \rho_k \mu^0$.

Step 3. Set $\mu^{k+1} = \mu^k + \Delta \mu^k$ and $w^{k+1} = w^k + \Delta w^k$. Set $k := k + 1$ and go to Step 1.

Remark 1

- (1) In theory, we use $\|H(\mu^k, w^k)\|^2 = 0$ as a termination of Algorithm 3.1. In practice, we use $\|H(\mu^k, w^k)\|^2 \leq \varepsilon$ as a termination rule, where ε is a pre-set tolerance error.
- (2) It is obvious that we have $\rho_k \leq \gamma \|H(\mu^k, w^k)\|^2$.
- (3) From (7) and (19), we have $\mu^{k+1} = \rho_k \mu^0 > 0$ for any $k \geq 0$.

Now, we are ready to analyze the convergence. The quadratic convergence of Algorithm 3.1 is given below.

Theorem 3.1 Assume that (μ^*, w^*) satisfies $H(\mu^*, w^*) = 0$. Suppose that $H(\mu, w)$ satisfies the condition of Lemma 2.1 and $\nabla H(\mu, w)$ is Lipschitz continuous with the constant L . Then we have the following conclusions:

- (1) *There exists a set $D \subset \mathcal{R}_+ \times \mathcal{R}^{n+2m}$ which contains (μ^*, w^*) such that for any $(\mu^0, w^0) \in D$, the iterate points (μ^k, w^k) generated by Algorithm 3.1 are well defined, remain in D and converge to (μ^*, w^*) ;*
- (2)

$$\|(\mu^{k+1} - \mu^*, w^{k+1} - w^*)\| \leq \beta \|(\mu^k - \mu^*, w^k - w^*)\|^2, \tag{20}$$

where $\beta := (L + 16\gamma\mu^0 \|\nabla H(\mu^*, w^*)\|^2) \|\nabla H(\mu^*, w^*)^{-1}\|$.

Proof According to Theorem 5.2.1 in [17] and Lemma 2.1, we give the proof in detail.

Denote

$$u = \begin{pmatrix} \mu \\ w \end{pmatrix}, \quad u^* = \begin{pmatrix} \mu^* \\ w^* \end{pmatrix}, \quad \Delta u = \begin{pmatrix} \Delta \mu \\ \Delta w \end{pmatrix}, \quad v^k = \begin{pmatrix} \rho_k \mu^0 \\ r^k \end{pmatrix}.$$

By Step 2 of Algorithm 3.1, we have

$$\|v^k\| \leq \rho_k \mu^0 + \|r^k\| \leq 2\rho_k \mu^0 \leq 2\gamma\mu^0 \|H(u^k)\|^2. \tag{21}$$

According to Lemma 2.1, we get that $\nabla H(u^*)$ is nonsingular. Then there exist a positive constant $\bar{t} < \frac{1}{\beta}$ and a neighborhood $N(u^*, \bar{t})$ of u^* such that $L\bar{t} \leq \|\nabla H(u^*)\|$, and for any $u \in N(u^*, \bar{t})$, we have that $\nabla H(u)$ is nonsingular and

$$\|\nabla H(u)\| - \|\nabla H(u^*)\| \leq \|\nabla H(u) - \nabla H(u^*)\| \leq L\|u - u^*\|, \tag{22}$$

where the first inequality follows from the triangle inequality and the second inequality follows from the Lipschitz continuity. Hence we have

$$\|\nabla H(u)\| \leq 2\|\nabla H(u^*)\| \tag{23}$$

for any $u \in N(u^*, \bar{t})$. Similarly, by the perturbation relation (3.1.20) in [17], we know that $\nabla H(u)$ is nonsingular and

$$\|\nabla H(u)^{-1}\| \leq \frac{\|\nabla H(u^*)^{-1}\|}{1 - \|\nabla H(u^*)^{-1}(\nabla H(u) - \nabla H(u^*))\|} \leq 2\|\nabla H(u^*)^{-1}\|. \tag{24}$$

Besides, for any $t \in [0, 1]$, we have

$$u^* + t(u - u^*) \in N(u^*, \bar{t}) \text{ and } H(u) - H(u^*) = \int_0^1 \nabla H[u^* + t(u - u^*)](u - u^*) dt.$$

From $\|H(u^*)\| = 0$, we have

$$\|H(u)\| \leq \int_0^1 \|\nabla H(u^* + t(u - u^*))\| \|u - u^*\| dt \leq 2\|\nabla H(u^*)\| \|u - u^*\|. \tag{25}$$

According to Algorithm 3.1, for any $u^k \in N(u^*, \bar{t}), k \geq 0$, we have

$$\begin{aligned} &u^{k+1} - u^* \\ &= u^k - u^* - \nabla H(u^k)^{-1}(H(u^k) - v^k) \end{aligned}$$

$$\begin{aligned}
 &= \nabla H(u^k)^{-1}(\nabla H(u^k)(u^k - u^*) - (H(u^k) - H(u^*)) + v^k) \\
 &= \nabla H(u^k)^{-1} \\
 &\quad \times \left(\int_0^1 (\nabla H(u^k) - \nabla H(u^* + t(u^k - u^*))) (u^k - u^*) dt + v^k \right). \tag{26}
 \end{aligned}$$

Taking norm of both sides, we get

$$\begin{aligned}
 &\|u^{k+1} - u^*\| \\
 &\leq \|\nabla H(u^k)^{-1}\| \left(\int_0^1 L(1-t)\|u^k - u^*\|^2 dt + \|v^k\| \right) \\
 &\leq 2\|\nabla H(u^*)^{-1}\| \left(\frac{1}{2}L\|u^k - u^*\|^2 + 2\gamma\mu^0\|H(u^k)\|^2 \right) \\
 &\leq 2\|\nabla H(u^*)^{-1}\| \left(\frac{1}{2}L\|u^k - u^*\|^2 + 8\gamma\mu^0\|\nabla H(u^*)\|^2\|u^k - u^*\|^2 \right) \\
 &= (L + 16\gamma\mu^0\|\nabla H(u^*)\|^2)\|\nabla H(u^*)^{-1}\|\|u^k - u^*\|^2, \tag{27}
 \end{aligned}$$

where the first inequality follows from the Lipschitz continuity, the second inequality follows from (21), and the third inequality follows from (25).

According to the definition of β and the condition of $\bar{t} < \frac{1}{\beta}$, we get that u^k converges to u^* . Besides, (20) also holds. This completes the proof. \square

4 The global inexact algorithm and its convergence

Now, we start our globally convergent method by using the global technique in Algorithm 3.1. We choose a merit function $h(\mu, w) = \frac{1}{2}\|H(\mu, w)\|^2$ and modify $(\Delta\mu^k, \Delta w^k)$ such that

$$-(\Delta\mu^k, \Delta w^k)^\top \nabla h(\mu^k, w^k) \geq \delta \|(\Delta\mu^k, \Delta w^k)\| \|\nabla h(\mu^k, w^k)\|. \tag{28}$$

We use line search to find a step-length $t^k \in (0, 1]$ such that

$$h(\mu^k + t^k \Delta\mu^k, w^k + t^k \Delta w^k) \leq h(\mu^k, w^k) + \bar{\rho} t^k \nabla h(\mu^k, w^k)^\top (\Delta\mu^k, \Delta w^k), \tag{29}$$

$$\nabla h(\mu^k + t^k \Delta\mu^k, w^k + t^k \Delta w^k)^\top (\Delta\mu^k, \Delta w^k) \geq \bar{\sigma} \nabla h(\mu^k, w^k)^\top (\Delta\mu^k, \Delta w^k) \tag{30}$$

and

$$\mu^k + t^k \Delta\mu^k \in \mathcal{R}_{++}, \quad w^k + t^k \Delta w^k \in \mathcal{R}^{n+2m}, \tag{31}$$

where $\bar{\rho} \in (0, 0.5), \bar{\sigma} \in (\bar{\rho}, 1), \delta \in (0, 1)$.

Algorithm 4.1 (Global inexact Newton method)

Step 0. Choose $(\mu^0, w^0) \in \mathcal{R}_{++} \times \mathcal{R}^{n+2m}$ to be an arbitrary point. Choose $\gamma \in (0, 1)$ such that $\gamma\mu^0 < \frac{1}{2}$. Choose $\bar{\rho} \in (0, 0.5), \bar{\sigma} \in (\bar{\rho}, 1), \delta \in (0, 1)$.

Step 1. If $\|H(\mu^k, w^k)\|^2 = 0$, then stop.

Step 2. Find $(\Delta\mu^k, \Delta w^k)$ by solving (19). If (28) is not satisfied, then choose τ_k and compute

$$(\Delta\mu^k, \Delta w^k) = -(\nabla H(\mu^k, w^k))^\top \nabla H(\mu^k, w^k) + \tau_k I)^{-1} \nabla h(\mu^k, w^k), \tag{32}$$

such that (28) is satisfied.

Step 3. Find a step-length $t^k \in (0, 1]$ satisfying (29)-(31).

$\mu^{k+1} = \mu^k + t^k \Delta\mu^k, w^{k+1} = w^k + t^k \Delta w^k$. Set $k := k + 1$ and go to Step 1.

Remark 2 In Step 2, if (28) is not satisfied, then the technique in [18], pp.264-265, is used to choose τ_k . From Lemma 3.1 in [18], it is not difficult to find t^k that can satisfy (29)-(31).

In order to obtain the global convergence of Algorithm 4.1, throughout the rest of this paper, we define the level set $\mathcal{L}(\mu^0, w^0) = \{(\mu, w) | h(\mu, w) \leq h(\mu^0, w^0)\}$ for $(\mu^0, w^0) \in \mathcal{R}_{++} \times \mathcal{R}^{n+2m}$.

Theorem 4.1 Suppose that $\nabla H(\mu, w)$ is Lipschitz continuous in $\mathcal{L}(\mu^0, w^0)$. Then we have

$$\lim_{k \rightarrow \infty} \nabla h(\mu^k, w^k) = 0.$$

Proof The proof follows Theorem 3.2 in [18] and condition (28). □

Theorem 4.2 Let $(\mu^0, w^0) \in \mathcal{R}_{++} \times \mathcal{R}^{n+2m}$, $H(\mu, x, \lambda, z)$ be defined by (7). Assume that $\nabla H(\mu, w)$ is Lipschitz continuous in $\mathcal{L}(\mu^0, w^0)$, $t^k = 1$ is admissible and (28) is satisfied for all $k \geq k_0$, $\nabla H(\mu^*, w^*)$ is nonsingular where k_0 is sufficiently great, and (μ^*, w^*) is a limited point of $\{(\mu^k, w^k)\}$ generated by Algorithm 4.1. Then the sequence $\{(\mu^k, w^k)\}$ converges to (μ^*, w^*) quadratically.

Proof From Theorem 4.1, we have

$$\lim_{k \rightarrow \infty} \nabla h(\mu^k, w^k) = 0,$$

where $\nabla h(\mu^k, w^k) = \nabla H(\mu^k, w^k)H(\mu^k, w^k)$. That is, the sequence $\{(\mu^k, w^k)\}$ is convergent. Since $\nabla H(\mu^*, w^*)$ is nonsingular and (μ^*, w^*) is a limited point of $\{(\mu^k, w^k)\}$ generated by Algorithm 4.1, we have

$$\lim_{k \rightarrow \infty} H(\mu^k, w^k) = 0.$$

According to the assumption that there exists k_0 such that $t^k = 1$ is admissible and (28) is satisfied for all $k \geq k_0$, $\{(\mu^k, w^k)\}$ can be generated by Algorithm 3.1 for $k > k_0$. We can get the conclusion from Theorem 3.1 directly. This completes the proof. □

5 Numerical results

In this section, we present some numerical results for Algorithm 4.1. All codes are written in Matlab and run on a RTM i5-3210M personal computer. In the algorithm, we choose $\gamma = 0.001$. We also use $\|H(\mu^k, w^k)\| \leq 10^{-5}$ as the stopping rule for all examples.

It is not easy to find proper test examples for the variational inequalities with nonlinear constraints. Hence, we modify some test examples in references and solve them by Algorithm 4.1.

Table 1 Numerical results for Example 5.1 with $x^0 = (0, 1, 1)$

k	Example 5.1			$\ H(\mu^k, w^k)\ $
	x_1	x_2	x_3	
1	1.6663	0.1527	0.7730	4.0941
2	1.6460	0.3416	0.0000	2.9087
3	1.6431	0.3212	0.0000	1.7601
4	1.2659	0.3957	0.0000	0.8018
5	1.0386	0.1050	0.0007	0.2139
6	1.0019	0.0174	0.0008	0.0301
7	1.0021	0.0005	0.0000	0.0043
8	1.0003	0.0000	0.0000	7.2097e-04
9	1.0000	0	0	1.8416e-05
10	1.0000	0.0000	0.0000	9.6147e-06

Example 5.1 (see [19]) Let

$$F(x) = \begin{pmatrix} 2x_1 + 0.2x_1^3 - 0.5x_2 + 0.1x_3 - 4 \\ -0.5x_1 + x_2 + 0.1x_2^3 + 0.5 \\ 0.5x_1 - 0.2x_2 + 2x_3 - 0.5 \end{pmatrix}$$

and

$$g(x) = -x_1^2 - 0.4x_2^2 - 0.6x_3^2 + 1.$$

It is verified that the problem has the solution $x^* = (1, 0, 0)$ easily. The initial point is $x^0 = (0, 1, 1)$ and $\mu^0 = 0.2$.

Example 5.2 This example is derived from [20]. Because the original problem is an optimization problem, we give its form of variational inequalities by the optimality condition, i.e.,

$$F(x) = \begin{pmatrix} 2x_1 - 5 \\ 2x_2 - 5 \\ 4x_3 - 21 \\ 2x_4 + 7 \end{pmatrix}, \quad g(x) = \begin{pmatrix} -x_1^2 - x_2^2 - x_3^2 - x_4^2 - x_1 + x_2 - x_3 + x_4 + 8 \\ -x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 + 10 \\ -2x_1^2 - x_2^2 - x_3^2 - 2x_1 + x_2 + x_4 + 5 \end{pmatrix}.$$

The solution of Example 5.2 is $x^* = (0, 1, 2, -1)$. The initial point is $x^0 = (0, 0, 0, 0)$ and $\mu^0 = 0.2$.

In Tables 1-2, ‘ k ’ means the number of iterations, ‘ $\|H(\mu^k, w^k)\|$ ’ means the 2-norm of $H(\mu^k, w^k)$. From Tables 1-2, we can observe that Algorithm 4.1 can find the solution in a smaller number of iterations for the above two examples. In order to further show the efficiency of Algorithm 4.1, we give other two examples where the dimension of the problems is from 100 to 1,000.

In the following tests, we solve Δw of the linear systems by using GMRES(m) package with $m = 10$, allowing a maximum of 100 cycles (2,000 iterations). And we choose μ^0 as a random number in $(0, 5)$.

Example 5.3 We consider the problem with nonlinear constraints. The problem is derived from [21] with different sizes. Based on the linear constraints of the original problem,

Table 2 Numerical results for Example 5.2 with $x^0 = (0, 0, 0, 0)$

k	Example 5.2				$\ H(\mu^k, w^k)\ $
	x_1	x_2	x_3	x_4	
1	0.3330	1.0661	4.0792	-4.3984	0.7137
2	3.7727	3.2371	5.7749	-2.3704	0.4308
3	1.8029	2.4677	2.9939	-1.3529	0.1151
4	0.6067	1.7571	2.2497	-0.9003	0.0399
5	0.4095	1.1936	2.1080	-0.7000	0.0076
6	0.0093	1.0500	2.1182	-0.8442	0.0045
7	0.0009	0.9994	2.0014	-0.9991	7.8889e-04
8	0.0002	1.0000	2.0005	-1.0000	3.2619e-05
9	0.0000	1.0000	2.0000	-1.0000	1.3066e-06

Table 3 Numerical results for Example 5.3

n	Example 5.3		$\ H(\mu^k, w^k)\ $
	No.it	CPU	
100	8	0.3120	7.7321e-06
200	8	0.4992	2.9356e-06
300	9	0.7956	2.5000e-06
400	9	1.1700	1.3077e-06
600	9	2.1840	1.8843e-06
800	9	3.8220	2.1879e-06
1,000	9	5.6004	2.0789e-06

we also add some nonlinear constraints to the problem. In this example,

$$F(x) = \begin{pmatrix} 4 & -2 & \dots & 0 & 0 \\ 1 & 4 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 4 & -2 \\ 0 & 0 & \dots & 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}, \quad g(x) = \begin{pmatrix} x_1 + 10 \\ x_2 + 10 \\ \dots \\ x_n + 10 \\ 100 - x_1^2 \\ 100 - x_2^2 \\ \dots \\ 100 - x_n^2 \end{pmatrix}.$$

Example 5.4 The example is the NCP. $F(x) = D(x) + Mx + q$. The components of $D(x)$ are $D_j(x) = d_j \cdot \arctan(x_j)$, where d_j is a random variable in $(0, 5)$. The matrix $M = A^T A + B$, where A is an $n \times n$ matrix whose entries are randomly generated in the interval $(-5, 5)$, and the skew-symmetric matrix B is generated in the same way. The vector q is generated from a uniform distribution in the interval $(-50, 50)$.

In Tables 3-4, ‘ n ’ means the dimension of problems, ‘No.it’ means the number of iterations, ‘CPU’ means the cpu time in seconds. ‘ $\|H(\mu^k, w^k)\|$ ’ means the 2-norm of $H(\mu^k, w^k)$. From Tables 3-4, we find that Algorithm 4.1 is robust to the different sizes for these two problems. Moreover, the iterative number is insensitive to the size of problems in our algorithm. In other words, our algorithm is more effective for two problems.

6 Conclusions

Based on the framework of smoothing Newton method, we propose a new smoothing inexact Newton algorithm for variational inequalities with nonlinear constraints. Under

Table 4 Numerical results for Example 5.4

<i>n</i>	Example 5.4		
	No.it	CPU	$\ H(\mu^k, w^k)\ $
100	8	0.4992	3.5558e-07
200	8	1.1388	3.0105e-06
300	8	2.3244	4.7536e-06
400	8	4.5552	7.7267e-06
600	9	13.1665	8.8521e-07
800	9	25.3346	1.4537e-06
1,000	9	41.4027	2.0736e-06

some mild conditions, we establish the global and local quadratic convergence. Furthermore, we also present some preliminary numerical results which show efficiency of the algorithm.

Acknowledgements

The first author is supported by Funding of Jiangsu Innovation Program for Graduate Education (KYZZ15_0087) and the Fundamental Research Funds for the Central Universities. The second author is supported by NSFC (11471159; 11571169; 61661136001) and the Natural Science Foundation of Jiangsu Province (BK20141409).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The authors contributed equally and significantly in writing this paper. All authors read and approved the final manuscript.

Author details

¹College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China. ²Basic Sciences Department, Nanjing Polytechnic Institute, Nanjing, 210048, China. ³School of Arts and Science, Suqian College, Suqian, 223800, China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 15 March 2017 Accepted: 22 June 2017 Published online: 10 July 2017

References

- Ferris, MC, Pang, JS: Engineering and economic applications of complementarity problems. *SIAM Rev.* **39**, 669-713 (1997)
- Fischer, A: Solution of monotone complementarity problems with locally Lipschitzian functions. *Math. Program.* **76**, 513-532 (1997)
- Li, YY, Santosa, F: A computational algorithm for minimizing total variation in image restoration. *IEEE Trans. Image Process.* **5**, 987-995 (1996)
- Facchinei, F, Pang, JS: *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volumes I and II*. Springer, Berlin (2003)
- Chen, BT, Harker, PT: Smooth approximations to nonlinear complementarity problems. *SIAM J. Optim.* **7**, 403-420 (1997)
- Chen, BT, Xiu, NH: A global linear and local quadratic noninterior continuation method for nonlinear complementarity problems based on Mangasarian smoothing functions. *SIAM J. Optim.* **9**, 605-623 (1999)
- Qi, LQ, Sun, DF, Zhou, GL: A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities. *Math. Program.* **87**, 1-35 (2000)
- Qi, LQ, Sun, DF: Improving the convergence of non-interior point algorithms for nonlinear complementarity problems. *Math. Comput.* **229**, 283-304 (2000)
- Tseng, P: *Error Bounds and Superlinear Convergence Analysis of Some Newton-Type Methods in Optimization. Nonlinear Optimization and Related Topics*. Springer, Berlin (2000)
- Ma, CF, Chen, XH: The convergence of a one-step smoothing Newton method for P0-NCP base on a new smoothing NCP-function. *Comput. Appl. Math.* **216**, 1-13 (2008)
- Zhang, J, Zhang, KC: A variant smoothing Newton method for P0-NCP based on a new smoothing function. *J. Comput. Appl. Math.* **225**, 1-8 (2009)
- Rui, SP, Xu, CX: A smoothing inexact Newton method for nonlinear complementarity problems. *J. Comput. Appl. Math.* **233**, 2332-2338 (2010)
- Chen, XJ, Qi, LQ, Sun, DF: Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities. *Math. Comput.* **67**, 519-540 (1998)

14. Rui, SP, Xu, CX: A globally and locally superlinearly convergent inexact Newton-GMRES method for large-scale variational inequality problem. *Int. J. Comput. Math.* **3**, 578-587 (2014)
15. Toyasaki, F, Daniele, P, Wakolbinger, T: A variational inequality formulation of equilibrium models for end-of-life products with nonlinear constraints. *Eur. J. Oper. Res.* **236**, 340-350 (2014)
16. Ng, MK, Weiss, P, Yuan, XM: Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods. *SIAM J. Sci. Comput.* **32**, 2710-2736 (2010)
17. Dennis, JE Jr, Schnabel, RB: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice Hall, Philadelphia (1983)
18. Nocedal, J, Wright, SJ: *Numerical Optimization*, 2nd edn. Springer, New York (1999)
19. Fukushima, M: A relaxed projection method for variational inequalities. *Math. Program.* **35**, 58-70 (1986)
20. Charalambous, C: Nonlinear least pth optimization and nonlinear programming. *Math. Program.* **12**, 195-225 (1977)
21. Ahn, BH: Iterative methods for linear complementarity problems with upperbounds on primary variables. *Math. Program.* **26**, 295-315 (1983)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
