A Power Penalty Method for Box Constrained Variational Inequality Problem

Hu Xizhen  
The Experiment Center  
The Second Artillery College  
Wuhan, China  
Huhu1218@126.com

Liao Sihong  
Training Department  
The Second Artillery College  
Wuhan, China

Yang Yang  
Training Department  
The Second Artillery College  
Wuhan, China

Hu Xiaobei  
Training Department  
The Second Artillery College  
Wuhan, China

Abstract—The variational inequality and the nonlinear complementarity problems has been well documented in the literature. Many methods such as nonsmooth Newton method and interior point method were studied in the variational inequality problem, but there was a limited study of penalty methods for the variational inequality problem. And a power penalty methods was presented for the box variational inequality problem in this paper. The box constrained variational inequality problem was proved to be equivalent to a nonlinear mixed complementarity problem and a new variational inequality problem, and present a novel power penalty approach to the new variational inequality problem in which the variational inequality problem is approximated by a nonlinear equation containing a power term. We show that the solution to the penalty equation converges to that of the VIP in the Euclidean norm when the function involved is Holder continuous and has a certain monotonicity property. Finally, we use a 3-dimensional vector-valued function defined to demonstrate the effectiveness of the algorithm.

Keywords: box constrained variational inequality; nonlinear mixed complementarity; power penalty method; monotone; Holder continuous

I. INTRODUCTION

The theory as well as the applications of both the variational inequality and the nonlinear complementarity problems has been well documented in the literature. Various extensions of these two problems have recently been introduced and studied by many authors including the traffic equilibrium problem[1-3], the spatial equilibrium problem and the Nash equilibrium problem. Variational inequality problems (VIP) are of fundamental importance in a wide range of mathematical and applied sciences, such as mathematical programming, traffic, engineering, economics and equilibrium problems, etc...to mention just a few. The theoretical foundation of VIP has been well studied and analyzed in the literature, and many algorithms have been proposed to solve the variational inequality problem. Many optimization algorithms and artificial neural networks developed to solve the variational inequality problems, because it has many important applications in wide variety of scientific and engineering fields including network economics, transportation science, game theory, military scheduling, automatic control, signal processing, regression analysis, structure design, mechanical design, electrical networks planning, and so on[4-9]. Christian Kanzow and Masao Fukushima present a method by using the D-gap function for the solution of the box constrained variational inequality problem in [10]. The method is a nonsmooth Newton method applied to formulation of as a system of nonsmooth equations involving the natural residual and show that the proposed algorithm is globally and fast locally convergent at last. There are interior point method for solving VIP such as in [11-14] and so on.

However, there was a limited study of penalty methods for VIP. Recently, Wang. and Yang [15] first presented a power penalty method for LCP in R^n based on approximating the LCP by a nonlinear equation, and Huang [16-18] developed it to NLCP and shown that the solution to the penalty equation converges to that of the NCP in the Euclidean norm at a rate of at least O(\lambda^{-k/2}). Inspired by their work, we develop a power penalty method for solving VIP, based on the idea in [16-18]. We first approximate the VIP by a nonlinear system of equations in which a power penalty term with a penalty constant \lambda > 1 and a power parameter k > 1 are contained. If F(x) obey the Assumption A1 and A2, we show that the solution to the penalty equation converges to that of the VIP in the Euclidean norm at a rate of at least O(\lambda^{-k/(k-1)}).

In this paper, we use || \cdot ||_p to denote the usual l_p-norm on R^n for any p > 1. When p = 2, it becomes the Euclidean norm. [u]_+ = \max\{u, 0\} and [u]_- = \min\{-u, 0\} and y'' = (y_1'', y_2'', \ldots, y_n'')^T for any y = (y_1, \ldots, y_n)^T and constant \sigma > 0. The outline of the paper is as follows. In section2, we briefly introduce three equivalent problems, prove their equality and uniqueness of the solution and analyze the penalty formulation and its convergence analysis in section 3.
II. BOX CONSTRAINED VARIATIONAL INEQUALITY

Problem 2.1 Find \( x \in K_i \) such that \( (y - x)^T F(x) \geq 0, \forall y \in R^n \), where \( F(x) \) is an \( n \)-dimensional vector-valued function defined on \( R^n \), \( a = (a_1, \cdots, a_n)^T \in R^n \) and \( b = (b_1, \cdots, b_n)^T \in R^n \) are given \( n \)-dimensional vectors. \( K_i = \{ x \in R^n : a \leq x \leq b \} \).

This problem is equivalent to the form of the nonlinear mixed complementarity problems discussed in [15]:

Theorem 2.1 Let \( K_i \) be \( (x \in R^n : a \leq x \leq b) \). A vector \( x \) solves the VI\((K_i,F)\) if and only if there exist vectors \( y, z \in R^n \) such that \((x^T,y^T,z^T)^T \in R^{3n}\) solves the following nonlinear mixed complementarity problems

Problem 2.2 Find \( x, y, z \in R^n \) such that
\[
\begin{align*}
F(x) - y + z &= 0 \\
a - x &\leq 0 \\
z &\leq 0 \\
z^T(a - x) &= 0
\end{align*}
\] and
\[
\begin{align*}
y &\leq 0 \\
x - b &\leq 0 \\
y^T(x - b) &= 0
\end{align*}
\]
Let
\[
t = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, H(t) = \begin{pmatrix} F(x) - y + z \\ x - b \\ a - x \end{pmatrix}
\]

Defining \( K = \{ (x^T,y^T,z^T)^T : x,y,z \in R^n, y \leq 0, z \leq 0 \} \), it is obvious that \( K \) is closed and convex cone in \( R^{3n} \). Using this \( K \), we define the following variational inequality problem corresponding to Problem 2.2:

Problem 2.3 Find \( t = (x^T,y^T,z^T)^T \) in \( K \), such that for all \( s \in K \)
\[
y \leq 0, z \leq 0, x \in R^n, H(t) \leq 0, (s - t)^T H(t) \geq 0
\]

Proposition 2.1 A vector \((x^T,y^T,z^T)^T \in R^{3n}\) is a solution to Problem 2.2 if and only if it is a solution to Problem 2.3.

Proof. If a vector \((x^T,y^T,z^T)^T \in R^{3n}\) is a solution to Problem 2.2, it is obvious that \((x^T,y^T,z^T)^T \in K \), \( F(x) - y + z = 0 \) and for any \((u^T,v^T,w^T)^T \in K \), we have
\[
\begin{pmatrix} u \\ v \\ w \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \end{pmatrix}^T \begin{pmatrix} F(x) - y + z \\ x - b \\ a - x \end{pmatrix} = (u - x)^T (F(x) - y + z) + (v - y)^T (x - b) + (w - z)^T (a - x)
\]

We notice that \( F(x) - y + z = 0 \), \( v \leq 0 \), \( w \leq 0 \), \( x - b \leq 0 \), \( a - x \leq 0 \), so
\[
\begin{pmatrix} u \\ v \\ w \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \end{pmatrix}^T \begin{pmatrix} F(x) - y + z \\ x - b \\ a - x \end{pmatrix} = v^T(x - b) + w^T(a - x) \geq 0
\]
Therefore, \((x^T,y^T,z^T)^T \) is a solution to Problem 2.3.

Conversely, if \((x^T,y^T,z^T)^T \) is a solution to Problem 2.3, we have \( y \leq 0, z \leq 0 \). We now need to prove \( x - b \leq 0 \). If it were not true, then there would exist at least an index \( i \), such that the \( i \)th component of \( x - b \) satisfies \( (x - b)_i > 0 \). Since \((u^T,v^T,w^T)^T \in K \) is arbitrary, we choose
\[
u = x, w = z, v = \begin{cases} y_j & j \neq i \\ y_j - \varepsilon, & j = i \end{cases}
\]
and for \( j = 1, 2, \ldots, n \) and arbitrary constant \( \varepsilon > 0 \).

Substituting this into \((s - t)^T H(t) \geq 0\) gives
\[
\begin{pmatrix} u \\ v \\ w \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \end{pmatrix}^T \begin{pmatrix} F(x) - y + z \\ x - b \\ a - x \end{pmatrix} = -\varepsilon (x - b)_i < 0
\]
This contradicts the fact that \((x^T,y^T,z^T)^T \) is a solution to Problem 2.3. Thus, we have \( x - b \leq 0 \). Similarly, we can prove that \( a - x \leq 0 \).

Next, we show that \( F(x) - y + z = 0 \). If it is not true, there must exist at least one index \( i \), such that \((F(x) - y + b)_i \neq 0 \). Choose
\[
v = y, w = z, u = \begin{cases} x_j & j \neq i \\ x_j - \varepsilon \text{sgn}(F(x) - y + b)_j, & j = i \end{cases}
\]
For \( j = 1, 2, \ldots, n \), where \text{sgn} denotes the sign function and constant \( \varepsilon > 0 \). Substituting this into (2.1) yields
\[
\begin{pmatrix} u \\ v \\ w \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \end{pmatrix}^T \begin{pmatrix} F(x) - y + z \\ x - b \\ a - x \end{pmatrix} = (u - x)^T (F(x) - y + z) + (v - y)^T (x - b) + (w - z)^T (a - x) = (u_i - x_i) (F(x) - y + b)_i - \varepsilon \text{sgn}[(F(x) - y + b)_i] (F(x) - y + b)_i < 0
\]
Which is impossible as \((x^T,y^T,z^T)^T \) is a solution to Problem 2.3. Therefore, \( F(x) - y + z = 0 \).

Finally, let us show that \( y^T(x - b) = 0 \). Since \((u^T,v^T,w^T)^T \in K \) is arbitrary, we choose \((u^T,v^T,w^T)^T \) as follows:
\[
(u^T,v^T,w^T)^T = (x^T,2y^T,z^T)^T
\]
So we have
\[
\begin{pmatrix} u \\ v \\ w \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \end{pmatrix}^T \begin{pmatrix} F(x) - y + z \\ x - b \\ a - x \end{pmatrix} = (u - x)^T (F(x) - y + z) + (v - y)^T (x - b) + (w - z)^T (a - x) = y^T(x - b) \geq 0
\]
And we choose \((u^T,v^T,w^T)^T \) as follows:
\[
(u^T,v^T,w^T)^T = \begin{pmatrix} x^T \\ \frac{1}{2} y^T \\ z^T \end{pmatrix}^T
\]
So we have
\[
\begin{pmatrix} u \\ v \\ w \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \end{pmatrix}^T \begin{pmatrix} F(x) - y + z \\ x - b \\ a - x \end{pmatrix} = (u - x)^T (F(x) - y + z) + (v - y)^T (x - b) + (w - z)^T (a - x) = -\frac{1}{2} y^T(x - b) \geq 0
\]
We can deduce \( y^T(x - b) = 0 \). We can prove that \( z^T(a - x) = 0 \) with the similar choose. This completes the proof of the proposition.

Before further discussion, it is necessary to impose the following assumptions on the nonlinear function \( F(x) \) in the Problem 2.1 which will be used in the rest of this paper.

A1. \( F(x) \) is Holder continuous on \( R^n \), i.e., there exist constants \( \beta > 0 \) and \( \gamma \in (0,1) \) such that
\[
||F(x) - F(y)||_2 \leq \beta ||x - y||^\gamma_2, \forall x, y \in R^n
\]
A2 \( F(x) \) is \( \xi \)-monotone, i.e., there exist constants 
\( a > 0 \) and \( \xi \in (1,2] \) such that
\[
(x - y)^T (F(x) - F(y)) \geq a \|x - y\|_2^\xi \tag{2.10}
\]

In the rest of this paper, we assume that Assumptions A1 and A2 are satisfied by \( F(x) \). Using these assumptions we are able to establish the continuity and the partial monotonicity of \( H(t) \) as given in the following theorem.

**Theorem 2.1** The function \( H(t) \) satisfies the following partial \( \xi \)-monotone property.
\[
(t_1 - t_2)^T (H(t_1) - H(t_2)) \geq a \|x_1 - x_2\|_2^\xi
\]
for any \( t_1 = (x_1^T, y_1^T, z_1^T)^T \in K \) and \( t_2 = (x_2^T, y_2^T, z_2^T)^T \in K \).

Proof. Let \( t_1 = (x_1^T, y_1^T, z_1^T)^T \in K \) and \( t_2 = (x_2^T, y_2^T, z_2^T)^T \in K \). be two arbitrary elements, then from Problem1.2, we have
\[
(t_1 - t_2)^T (H(t_1) - H(t_2)) = (x_1 - x_2)^T (F(x_1) - F(x_2)) + (y_1 - y_2)^T (x_1 - x_2) + (z_1 - z_2)^T (y_1 - y_2)
\]
\[
= (x_1 - x_2)^T (F(x_1) - F(x_2)) + (y_1 - y_2)^T (x_1 - x_2) + (z_1 - z_2)^T (y_1 - y_2)
\]
\[
= (x_1 - x_2)^T (F(x_1) - F(x_2)) + (y_1 - y_2)^T (x_1 - x_2) + (z_1 - z_2)^T (y_1 - y_2)
\]
\[
geq a \|x_1 - x_2\|_2^\xi
\]
by Assumption A2. Thus, we have proved theorem2.1.

Combining assumption A2 with \( K_1 \) is a closed and convex cone we see from Theorem 2.3.5 of [15] that Problem1.3, or equivalently Problem 2.2 and 2.3 has solutions. Moreover, for this particular problem , it is possible to show that solution is also unique, as given in the following theorem.

**Theorem 2.2** There exists a unique solution to Problem 2.3.

Proof. We concentrate on the uniqueness of the solution.
Suppose \( t_1 = (x_1^T, y_1^T, z_1^T)^T \in K \) and \( t_2 = (x_2^T, y_2^T, z_2^T)^T \in K \) are solutions to Problem 2.3. Then \( t_1 \) and \( t_2 \) satisfy
\[
(s - t_1)^T H(t_1) \geq 0
\]
\[
(l - t_2)^T H(t_2) \geq 0
\]
for any \( s, l \in K \). Replacing \( s \) and \( l \) with \( t_1 \) and \( t_2 \) respectively, adding the resulting inequalities up and rearranging the terms, we have
\[
(t_1 - t_2)^T (H(t_1) - H(t_2)) \leq 0
\]
Combining this inequality and (2.6) gives \( x_1 = x_2 = z_2 \).

Now we show that \( y \) and \( z \) are unique. For any \( i \in \{1,2,\ldots,n\} \), if \( (x_i - b_i) \neq 0 \), it is easy to see that \( y_i = 0 \), notice that \( F(x) - y = z \), we have \( y_i = y_i \). In the case that \( (x_i - b_i) = 0 \) for some \( i \in \{1,2,\ldots,n\} \), i.e. \( x_i = b_i \), we can deduce that \( z_i = 0 \) since \( a - x \geq 0 \) and \( (a - x_i) > 0 \), we have \( y_i = -F(x) \).

III. THE PENALTY FORMULATION AND ITS CONVERGENCE ANALYSIS

Let \( k > 0 \) be a fixed parameter. We propose the following penalty problem to approximate Problem 3:

**Problem3.1** Find \( (x_3^T, y_3^T, z_3^T)^T \in R^n \) with \( x_3, y_3, z_3 \in R^n \) such that
\[
(F(x_3) - y_3 + z_3) + \lambda \left( \begin{array}{c}
0 \\
0 \\
0
\end{array} \right) = 0
\]
where \( \lambda > 0 \) is the penalty parameter.

Clearly, (3.1) is a penalty equation which approximates Problem2.2. This equation contains a penalty term \( \lambda \left( \left[ y_{3+} \right]^k, [z_{3+}]^k \right)^T \) which penalizes the positive part of \( y_{3+}^k, z_{3+}^k \) when (3.10) is violated. It is easy to see from (3.1) that (2.10) is always satisfied by \( x^T, y^T, z^T \) because \( \lambda \left( \left[ y_{3+} \right]^k, [z_{3+}]^k \right)^T \geq 0 \).

We start our convergence analysis with the following lemma.

**Lemma 3.1** Let \( x_3^T, y_3^T, z_3^T \) be a solution (3.1). Then, there exists a positive constant \( C_1, C_2 \), independent of \( x_3, y_3, z_3, \lambda \) and \( k \), such that
\[
\|x_3\|_2 \leq C_1
\]
\[
\|y_3^T [z_3^T]^k \|_2 \leq C_2^k
\]
Using Lemmas 3.1, we ready to present and prove our main convergence results as given in the following theorem.

**Theorem3.1** Let \( x^T, y^T, z^T \) be the solutions to Problem 2.1 and 3.1, respectively. There exists positive constant \( C \), independent of \( x_3, y_3, z_3, \lambda \) and \( k \), such that
\[
\|x - x_3\|_2 \leq \frac{c}{\lambda (k+1)}
\]
Proof. We let \( C \) be a generic positive constant, independent of \( x_3, y_3, z_3, k \). We first show (3.4) in a similar way as that in [5], as given below.
Let \( t = (x_i^T, y_i^T, z_i^T)^T \),
we decompose \( t - t_3 \) into \( t - t_3 = t - (t_2 - t_3) = t + (t_2 - t_3) - t_3\),
\[
= r_2 - t_3
\]
where \( r_2 = t + [t_2]_+ \). Noticing \( t - r_2 = -[t_3]_+ \leq 0 \). We have \( t - r_2 \in K \). Note \( t \) is a solution to Problem2.2 and thus satisfies \( (s - t)H(t) \geq 0 \). Therefore, replacing \( s \) in \( (s - t)H(t) \geq 0 \) with \( t_2 \) gives
\[
r_2H(t_2) \geq 0
\]
Since \( t_2 \) satisfies (3.1), left-multiplying both sides of (3.1) by \( r_2^T \), we have
\[
r_2^T H(t_2) + \lambda r_2^T \left( \left[ y_{3+} \right]^k, [z_{3+}]^k \right)^T = 0
\]
Adding up both sides of (3.6) and (3.7) gives
\[
r_2^T H(t_2) - H(t_2) + \lambda r_2^T \left( \left[ y_{3+} \right]^k, [z_{3+}]^k \right)^T \geq 0
\]
Note that
\[
r_2^T \left( \left[ y_{3+} \right]^k, [z_{3+}]^k \right)^T = (t_2 + [t_2]_+) \left( \left[ y_{3+} \right]^k, [z_{3+}]^k \right)^T
\]
\[
= y^T [y_{3+}]^k + z^T [z_{3+}]^k \leq 0
\]
thus, (3.4) reduces to
\[
(r_2^T H(t_2) - H(t_2)) \leq 0
\]
Or equivalently
We consider the variational inequality Problem 2.3. Find $x \in K$ such that $(y-x)^T F(x) \geq 0$, $\forall y \in R^n$, where $F(x)$ is a vector-valued function defined on $R^n$, and $F(x) = \begin{pmatrix} x_1 + x_1^2 - 2 \\ x_2 + x_2^2 - 1 \end{pmatrix}$. $a = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $K = \{ x \in R^2 : a \leq x \leq b \}$.

We have already proved that $A$ vector $x$ solves $\text{VI}(K, F)$ if and only if there exist vectors $y, z \in R^n$ such that $(x^T, y^T, z^T) \in R^{3n}$ solves the following nonlinear mixed complementarity problems:

Find $x, y, z$, such that:

\[
\begin{align*}
F(x) + y - z &= 0 \\
x - a &\geq 0 \\
z &\geq 0 \\
y^T (x - a) &= 0 \\
(-b) - x &\geq 0 \\
y &\geq 0 \\
y^T (b - x) &= 0
\end{align*}
\]

It is easily tested that $F(x)$ has $\xi$-monotonicity property, and the above problem has a unique solution for $x \in \begin{pmatrix} 1 & 1 \end{pmatrix}$, $y \in \begin{pmatrix} 0 & 0 \end{pmatrix}$, $z \in \begin{pmatrix} 1 & 1 \end{pmatrix}$. The problem is to find a unique solution to the above problem: