Inexact alternating direction method for variational inequality problems with linear equality constraints

Min Sun, Jing Liu

Abstract—In this article, a new inexact alternating direction method (ADM) is proposed for solving a class of variational inequality problems. At each iteration, the new method firstly solves the resulting subproblems of ADM approximately to generate an temporal point \( \tilde{x}^k \), and then the multiplier \( \tilde{y}^k \) is updated to get the new iterate \( y^{k+1} \). In order to get \( x^{k+1} \), we adopt a new descent direction which is simple compared with the existing prediction-correction type ADMs. For the inexact ADM, the resulting proximal subproblem has closed-form solution when the proximal parameter and inexact term are chosen appropriately. We show the efficiency of the inexact ADM numerically by some preliminary numerical experiments.

Keywords—variational inequality problems, alternating direction method, global convergence

I. INTRODUCTION

Let \( S \subseteq \mathbb{R}^n \) be a nonempty closed convex set and \( f(\cdot) \) be a continuous mapping from \( S \subseteq \mathbb{R}^n \) into itself. The variational inequality (VI) problem, denoted by \( \text{VI}(f, S) \), is to find \( x^* \in S \), such that

\[
    (x - x^*)^\top f(x^*) \geq 0, \quad \forall x \in S,
\]

where ‘\(\top\)’ denotes the standard inner product. In this paper, we consider the VI\((f, S)\) which feasible set \( S \) has the following structure

\[
    S = \{ x \in \mathbb{R}^n | Ax = b, x \in \mathcal{X} \},
\]

where \( A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m \) and \( \mathcal{X} \subseteq \mathbb{R}^n \) is a nonempty closed convex set. For wide applications of VI\((f, S)\) with linear equality constraints, see e.g.[1-3,6].

By attaching a Lagrangian multiplier \( y \in \mathbb{R}^m \) to the linear constraint \( Ax = b \), VI\((1)-(2)\) can be expressed in the structured variational inequalities of the following form, denoted by VI\((F, U)\): find \( u^* \in U \), such that

\[
    (u - u^*)^\top F(u^*) \geq 0 \quad \forall u \in U,
\]

where

\[
    u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad F(u) = \begin{pmatrix} f(x) - A^\top y \\ Ax - b \end{pmatrix}, \quad U = \mathcal{X} \times \mathbb{R}^m.
\]

To solve VI\((F, U)\), the classical alternating direction method (ADM)[4,5,8,9] finds a new iterate \((x^{k+1}, y^{k+1}) \subseteq \mathcal{X} \times \mathbb{R}^m\) from a given tuple \((x^k, y^k) \subseteq \mathcal{X} \times \mathbb{R}^m\) via the following procedure:

Find \( \tilde{x}^k \in \mathcal{X} \) such that

\[
    (x' - \tilde{x}^k)^\top \{ f(\tilde{x}^k) - A^\top [y^k - \beta(A\tilde{x}^k - b)] \} \geq 0, \quad \forall x' \in \mathcal{X},
\]

and take \( \tilde{x}^k \) as \( x^{k+1} \).

Find \( \tilde{y}^k \) such that

\[
    y^k = y^k - \beta(A\tilde{x}^k - b),
\]

and take it as \( y^{k+1} \), where \( \beta > 0 \) is a given penalty parameter for the linear constraint \( Ax = b \).

In fact, the subproblem (4) is not tractable, since it involves the following implicit projection equation

\[
    \tilde{x}^k = P_X \{ \tilde{x}^k - [f(\tilde{x}^k) - A^\top [y^k - \beta(A\tilde{x}^k - b)] + r(\tilde{x}^k - x^k) + \xi^k] \},
\]

in the sense that the unknown variable \( \tilde{x}^k \) appears on both sides of the above equation. Thus, in the new ADM to be proposed in this paper, we introduce a proximal term \( r(\tilde{x}^k - x^k) \) and an admissible error \( \xi^k \) into (4), and then solve the resulting subproblem. That is:

\[
    (x' - \tilde{x}^k)^\top \{ f(\tilde{x}^k) - A^\top [y^k - \beta(A\tilde{x}^k - b)] + r(\tilde{x}^k - x^k) + \xi^k \} \geq 0, \quad \forall x' \in \mathcal{X}.
\]

Of course, the above subproblem becomes complex. However, if we choose a good ‘inexact term’, then (6) will become very easy to compute. For simplification, we may assume that \( r = 1/\beta \). Then, with this choice of \( r \), the proximally regularized subproblem (6) can be rewritten into:

\[
    (x' - \tilde{x}^k)^\top \{ \beta(f(\tilde{x}^k) - A^\top [y^k - \beta(A\tilde{x}^k - b)] + (\tilde{x}^k - x^k) + \beta\xi^k) \} \geq 0, \quad \forall x' \in \mathcal{X}.
\]

A suitable choice of \( \xi^k \) will lead to an explicit projection of (7), such as if we set \( \xi^k = f(\tilde{x}^k) - f(x) + \beta A^\top A(x - \tilde{x}^k) \).

Then with this particular \( \xi^k \), the subproblem (7) reduces to:

\[
    (x' - \tilde{x}^k)^\top \{ \beta(f(x) - A^\top [y^k - \beta(A\tilde{x}^k - b)] + (\tilde{x}^k - x^k) \} \geq 0, \quad \forall x' \in \mathcal{X}.
\]

Obviously, the VI\((8)\) has the closed-form solution given by the following projection:

\[
    \tilde{x}^k = P_X \{ x^k - \beta(f(x^k) - A^\top [y^k - \beta(Ax^k - b)]) \}.
\]

Despite the obvious simplicity, this idea raises immediately the question: How can we ensure the convergence of the sequence generated by (9) and (5) to a solution of VI\((F, U)\). In this paper, motivated by [7,10], we give an answer to the above question. In order to ensure the convergence of the inexact ADM, firstly, we give the condition on \( \xi^k \), and then a correction step is adopted to generate the new iterate.

The rest of the paper is organized as follows: In Section 2, we give some basic concepts we will use in the following analysis. In Section 3, we describe the inexact ADM in details, and the global convergence of the new method is proved. We report some preliminary numerical results in Section 4 and some conclusions are drawn in the last section.

Min Sun: Department of Mathematics and Information Science, Zaozhuang University, Shandong 271610, China. e-mail: ziyouxiaozhou@163.com

Jing Liu: School of Mathematics and Statistics, Zhejiang University of Finance and Economics, Hangzhou 310018, China. e-mail: bjl8899@163.com
II. PRELIMINARIES

In this section, we give some basic properties and related definitions which will be used in the following discussion. The Euclidean norm of \( v \in \mathbb{R}^n \) is defined by \( \| v \| = \sqrt{v^\top v} \), and let \( G \in \mathbb{R}^{n \times n} \) be a symmetric positive definite, then the G-norm of a vector \( z \in \mathbb{R}^n \) is denoted by \( \| z \|_G := z^\top G z \). The definition of projection operator which is defined as a mapping from \( \mathbb{R}^n \) to a nonempty closed convex subset \( K \):

\[
P_K[x] := \text{argmin}\{\| x - y \| : y \in K \}, \quad \forall x \in \mathbb{R}^n.
\]

The following well known properties of the projection operator will be used bellow.

**Lemma 2.1.** Let \( K \) be a nonempty closed convex subset of \( \mathbb{R}^n \). For any \( x, y \in \mathbb{R}^n \) and any \( z \in K \), the following properties hold.

\[
(x - P_K[x])^\top (z - P_K[x]) \leq 0, \quad \forall x \in \mathbb{R}^n, z \in K.
\]

**Assumption**

(A1) The solution set of problem VI\((F, U)\), denoted by \( U^* \), is nonempty.

(A2) The underlying mapping \( f \) is monotone, i.e.,

\[
(x - y)^\top (f(x) - f(y)) \geq 0, \quad \forall x, y \in \mathbb{R}^n.
\]

(A3) \( \mathcal{X} \) is a simple closed convex set. That is, the projection onto the set is simple to carry out.

It is easy to prove that \( F \) is also monotone when \( f \) is monotone, thus the solution set of VI\((F, U)\) is convex under Assumption (A2).

III. ALGORITHM AND GLOBAL CONVERGENCE

The error term \( \xi_k \) in (6) allows the sub-VI to be solved approximately. In this paper, we use the following inexact criterion:

\[
\| \xi_k \| \leq c_r \| x^k - \tilde{x}^k \|, \quad \text{with } v \in (0, 1).
\]

Now, we describe our algorithm detailed as follows:

**The Inexact ADM**

**Step 0.** Let \( r, \beta > 0, v \in (0, 1) \). Given \( \varepsilon > 0 \), choose \( u^0 = (x^0, y^0)^\top \in \mathcal{U} \) and set \( k = 0 \).

**Step 1.** Find \( \tilde{x}^k \in \mathcal{X} \) (with fixed \( x^k, y^k \)) by (6), where \( \xi_k \) satisfies (13).

**Step 2.** The new iterate is produced by: Find \( y^{k+1} = \tilde{y}^k \) (with fixed \( \tilde{x}^k, y^k \)) by (5), and \( x^{k+1} \) is updated by:

\[
x^{k+1} = \tilde{x}^k + \frac{1}{r} \xi^k.
\]

**Step 3.** Convergence verification: If \( \| x^k - \tilde{x}^k \| + \| y^k - y^{k+1} \| < \varepsilon \), then stop; otherwise, set \( k := k + 1 \) and goto Step 1.

Note that \( \| x^k - \tilde{x}^k \| + \| y^k - y^{k+1} \| = 0 \) if and only if \( x^k = \tilde{x}^k, y^k = y^{k+1} \). Then, from (5) (6) and (14), we have that \( u^k \) is actually a solution of VI\((F, U)\), which means the iteration will be terminated. Thus, the stopping condition in Step 3 is reasonable.

**Remark 3.1** The updating formula (14) can be rewritten as

\[
x^{k+1} = x^k - d(x^k, \tilde{x}^k, \xi^k),
\]

where

\[
d(x^k, \tilde{x}^k, \xi^k) = (x^k - \tilde{x}^k) - \frac{1}{r} \xi^k,
\]

which is referred to as the search direction of \( x \) in the inexact ADM.

**Remark 3.2** If we set \( r = 1/\beta \) and \( \xi_k = f(x^k) - f(\tilde{x}^k) + \beta A^\top A(x^k - \tilde{x}^k) \) in (6), then (6) has the closed-form solution (9). Now, we illustrate that the condition (13) on the inexact term is well-defined. In fact, if we assume that \( f(x) \) is Lipschitz continuous and \( L_f \) is the Lipschitz constant of \( f(x) \). Obviously, when \( r \) satisfies

\[
r \geq L_f + \beta \| A^\top A \|,
\]

it follows that

\[
\| \xi^k \| \leq (L_f + \beta \| A^\top A \|) \| x^k - \tilde{x}^k \| \leq rv \| x^k - \tilde{x}^k \|,
\]

which guarantees the condition (13).

**Lemma 3.1.** If \( \tilde{u}^k = (\tilde{x}^k, y^k) \) is generated by (5)-(6) from a given \( u^k = (x^k, y^k) \), then for any \( u^* = (x^*, y^*) \in U^* \), we have

\[
r(x^* - \tilde{x}^k)^\top d(x^*, \tilde{x}^k, \xi^k) + (y^* - y^k)^\top (A \tilde{x}^k - b) \geq 0.
\]

**Proof.** Since \( u^* \in U^* \) and \( \tilde{x}^k \in \mathcal{X} \), we have

\[
(x^* - \tilde{x}^k)^\top (f(x^*) - A^\top y^*) \geq 0.
\]

Adding (18) and (19), and by the monotonicity of operator \( f \), we get

\[
(x^* - \tilde{x}^k)^\top (A^\top (y^* - y^k) - A^\top (y^k - y^{k+1})) + rd(x^k, \tilde{x}^k, \xi^k) \geq 0.
\]

Combining (20) and (21), we have

\[
r(x^* - \tilde{x}^k)^\top d(x^*, \tilde{x}^k, \xi^k) + (y^* - y^k)^\top (A \tilde{x}^k - b) \geq (A \tilde{x}^k - b)^\top (y^k - y^{k+1})
\]

\[
= \beta \| A \tilde{x}^k - b \|^2 (\text{using (5)}).
\]

The assertion of this lemma follows from the above inequality directly. The proof is complete.

Now, we are ready to prove that \(-d_1(u^k, \tilde{u}^k) \) defined in (9) and \(-d_2(u^k, \tilde{u}^k) \) defined in (10) are two descent directions of the merit function \( \frac{1}{2} \| u - u^* \|_v^2 \) at \( u = u^k \), though \( u^* \in U^* \) is unknown.
Lemma 3.2. Let \( \{u^k\} = \{(x^k, y^k)\} \) be the sequence generated by the inexact ADM. Then, for any \( u^* = (x^*, y^*) \in \mathcal{U}^* \), we have

\[
\|u^k - u^*\|_G^2 \geq \|u^{k+1} - u^*\|_G^2 + 2\beta \|Ax^k - b\|^2 + r(1 - \|v\|^2)\|x^k - x^*\|^2,
\]

where

\[
G_{(n+m)\times(n+m)} = \begin{pmatrix} rI_n & 0 \\ 0 & I_m/\beta \end{pmatrix}
\]
is a positive definite matrix.

Proof. It follows from (15) that

\[
2r(x^k - x^*)^T d(x^k, \bar{x}^k, \xi^k) = \left( \|x^k - x^*\|_{I_n}^2 - \|x^{k+1} - x^*\|_{I_n}^2 \right) + \|d(x^k, \bar{x}^k, \xi^k)\|^2,
\]

By (5),

\[
2(y^k - y^*)^T(Ax^k - b) = \left( \|y^k - y^*\|_{\frac{3}{2}I_n}^2 - \|y^{k+1} - y^*\|_{\frac{3}{2}I_n}^2 \right) + \beta \|Ax^k - b\|^2.
\]

Substituting (24) and (25) in (18) and using the matrix \( G \), we have

\[
\left( \|u^k - u^*\|^2_G - \|u^{k+1} - u^*\|^2_G + r\|d(x^k, \bar{x}^k, \xi^k)\|^2 \right) + 2\beta \|Ax^k - b\|^2 \geq 2\beta \|Ax^k - b\|^2 + 2r(x^k - \bar{x}^k)^T d(x^k, \bar{x}^k, \xi^k).
\]

By a manipulation, we have

\[
\|u^k - u^*\|^2_G - \|u^{k+1} - u^*\|^2_G \geq \beta \|Ax^k - b\|^2 + 2r(x^k - \bar{x}^k)^T d(x^k, \bar{x}^k, \xi^k)
\]

Then, from (13) and (16), we get

\[
2r(x^k - \bar{x}^k)^T d(x^k, \bar{x}^k, \xi^k) = r\|d(x^k, \bar{x}^k, \xi^k)\|^2
\]

and consequently

\[
\left\{ \begin{array}{c} (x - x^\infty)^T [f(x^\infty) - A^Ty^\infty] \geq 0, \forall x \in \mathcal{X}; \\
Ax^\infty - b = 0, \end{array} \right.
\]

which implies that \( x^\infty \in \mathcal{U}^* \), i.e., \( y^\infty \) is a solution of VI(\( F, \mathcal{U} \)).

Now, we have to show that the sequence \( \{u^k\} \) actually converges to \( u^\infty \). Suppose that \( \hat{u} \) is another cluster point of \( \{u^k\} \). Then, we have

\[
\delta := \|u^\infty - \hat{u}\|_G > 0.
\]

Because \( u^\infty \) is a cluster point of the sequence \( \{u^k\} \), there is a \( k_0 > 0 \) such that

\[
\|u^{k_0} - u^\infty\|_G \leq \frac{\delta}{2}.
\]

On the other hand, since \( \{\|u^k - u^\infty\|_G\} \) is non-increased, we have \( \|u^k - u^\infty\|_G \leq k \leq k_0 \), and it follows that

\[
\|u^k - \hat{u}\|_G \geq \|u^\infty - \hat{u}\|_G - \|u^k - u^\infty\|_G \geq \frac{\delta}{2}, \forall k \geq k_0,
\]

which is a contradiction of the fact that \( \hat{u} \) is a cluster point of \( \{u^k\} \). This contradiction indicates that the sequence \( \{u^k\} \) converges its unique cluster point \( u^\infty \), which is a solution of VI(\( F, \mathcal{U} \)). This completes the proof.

IV. Numerical results

In this section, we present an example to show the applicability and robustness of the proposed method. The example used here is the first test problem in paper[9], which mapping \( f \) are taken as

\[
f(x) = Mx + \rho C(x) + q,
\]

where \( M \) is an \( R_{\mathbb{F}}^{5x5} \) asymmetric positive matrix and \( C_i(x) = \arctan((x - 2)/i), i = 1, 2, \cdots, 5 \). The parameter \( \rho \) is used to vary the degree of asymmetry and nonlinearity, and the data of this example are illustrate as follows:

\[
M = \begin{pmatrix}
0.726 & -0.949 & 0.266 & -1.193 & -0.504 \\
1.645 & 0.678 & 0.333 & -0.217 & -1.443 \\
-1.016 & -0.225 & 0.709 & 0.934 & 1.007 \\
1.063 & 0.567 & -1.144 & 0.550 & -0.548 \\
-0.259 & 1.453 & -1.073 & 0.509 & 1.026
\end{pmatrix}
\]

and

\[
q = (5.308, 0.008, -0.938, 1.024, -1.312)^T.
\]

The problem has a unique solution \( x^* = (2, 2, 2, 2, 2)^T \). The parameters used in the inexact ADM were set as \( \beta = 0.05, r = 1/\beta \) for \( \rho = 10 \), and \( \beta = 0.01, r = 1/\beta \) for \( \rho = 20 \). The stop parameter \( \varepsilon = 10^{-6} \). The results for \( \rho = 10 \) and \( \rho = 20 \) are listed in Table 1 and Table 2, respectively. In this table, 'Num. of Iter.' denotes the number of iterations and 'CPU Time' denotes the CPU time in seconds. The codes for this problem is given in the following box:
x=[25 0 0 0 0]; time=cputime; M=[0.726 -0.949 0.266 -1.193 -0.504; 1.645 0.678 0.333 -0.217 -1.443; -1.016 -0.225 0.769 0.934 1.007; 1.063 0.567 -1.144 0.550 -0.548; -0.259 1.453 -1.073 0.509 1.026]; q=[5.308 0.008 -0.938 1.024 -1.312]; A=[1 1 1 1 1]; b=10; rho=20; ka=5; A=A*ka; b=b*ka; beta=0.01; r=1/beta; eps=10e-6; y=0; k=0; f=M*x+rho*atan(x-[2 2 2 2 2])+q; xwan=max(x-beta*(f-A*(y-beta*(A*x-b))),zeros(5,1)); ywan=y-beta*(A*xwan-b); error=norm(x-xwan)+norm(y-ywan); while error>=eps fwan=M*xwan+rho*atan(xwan-[2 2 2 2 2])+q; xi=f-fwan+beta*A'*A*(x-xwan); x=xwan+xi/r; y=ywan; k=k+1; f=M*x+rho*atan(x-[2 2 2 2 2])+q; xwan=max(x-beta*(f-A*(y-beta*(A*x-b))),zeros(5,1)); ywan=y-beta*(A*xwan-b); error=norm(x-xwan)+norm(y-ywan); end time=cputime-time xing=[2,2,2,2,2]; error1=norm(x-xing)