A New Alternating Direction Method for Structured Variational Inequalities in Engineering Modeling

Jing Liu[1]

School of Mathematics and Statistics, Zhejiang University of Finance and Economics, Hangzhou, 310018, China.

Abstract. Recently, the alternating direction method(ADM) has gained tremendous interest in the area of applied sciences, such as the image processing, and matrix decomposition. In this paper, we propose a new alternating direction method for structured variational inequalities, which only needs functional values in the solution process. We give a new residual function $r(u, \beta)$, and based on it a new descent direction $d(u, \beta)$ is obtained. Under Lipschitz continuity of the underlying function $f(\cdot)$, its global convergence is proved. Some computational results are given to illustrate its efficiency.

Keywords. Variational inequalities; Alternating direction method; Descent direction.

AMS(2000) Subject Classification: 90C25, 90C30

1 Introduction

Let $A \in \mathcal{R}^{m\times n}$ and $b \in \mathcal{R}^m$, $\mathcal{X} \subset \mathcal{R}^n$ be a nonempty closed convex set and f be continuous mapping from $\mathcal{R}^n$ into itself. In this paper, we focus on the following constrained variational inequalities: find $x^* \in S$, such that

$$(x - x^*)^\top f(x^*) \geq 0, \quad \forall x \in S, \tag{1}$$

where

$$S = \{v \in \mathcal{R}^n | Av = b, v \in \mathcal{X}\} \tag{2}$$

or

$$S = \{v \in \mathcal{R}^n | Av \leq b, v \in \mathcal{X}\} \tag{3}$$

By attaching a Lagrangian multiplier vector $y \in \mathcal{Y} = \mathcal{R}^m$ and $y \in \mathcal{Y} = \mathcal{R}^m_+$ to the linear equality constraints $Av = b$ and $Av \leq b$, respectively, problem (1) can be converted into the following equivalent form, denoted by VI(F,$\mathcal{U}$):

$$\text{Find} \;\; u^* \in \mathcal{U}, \;\text{such that}\;\; (u - u^*)^\top F(u^*) \geq 0 \quad \forall u \in \mathcal{U}, \tag{4}$$

[1]This work was supported by the Shandong Provincial the Natural Science Foundation (Grant No. ZR2010AL005)

***AMO-Advanced Modeling and optimization. ISSN: 1841-4311**

where

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, F(u) = \begin{pmatrix} f(x) - A^\top y \\ Ax - b \end{pmatrix}, \text{ and } \mathcal{U} = \mathcal{X} \times \mathcal{Y}.$$

A typical method for solving VI(F,$\mathcal{U}$) is the decomposition method proposed by Gabay[1] and Gabay and Mercier[2], which is attractive for large-scale problems since it decompositions the original problem to a series of small-scale VI problems. However, solving the sub-VI exactly could be computationally intensive by itself. To overcome this difficulty, He and Zhou[3], Han[4-6], Sun[7] proposed some modified alternating direction methods which only needs some projections onto simple sets and calculate some matrix-vector products. Most recently, Abdellah et al.[8] extended this type method to general variational inequalities and they claimed that an efficient alternating direction method is constructed and showed its global convergence. Recently, the alternating direction methods (ADM) have gained impressive applications in a wide range of applied science problems, such as the image processing and matrix decomposition. In [10], the authors have applied ADM to solve the constrained total-variation image restoration and reconstruction problem.

Motivated by the above research, in this paper, we propose a new alternating direction method for VI(F,$\mathcal{U}$). Note that in our method, we only needs functional values. Since we don't use the concrete structure of $\mathcal{Y}$, our method can be extended to solve VI(F,$\mathcal{U}$) with a proper set $\mathcal{Y}$, however, the methods in [3-6] can't be extended to solve such problems.

The rest of this paper is organized as follows. In the next section, some basic concepts about variational inequalities are presented. In Section 3, we describe the new alternating direction method in details, and its global convergence is also analyzed. We report some preliminary computational results in Section 4 and some conclusions are given in Section 5.

2 Preliminaries

We first give the definition of projection operator which is defined as a mapping from $\mathcal{R}^n$ to a nonempty closed convex subset $\mathcal{K}$:

$$P_{\mathcal{K}}[x] := \operatorname{argmin}\{\|x-y\| | y \in \mathcal{K}\}, \quad \forall x \in \mathcal{R}^n.$$

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

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

$$(x - P_{\mathcal{K}}[x])^\top (z - P_{\mathcal{K}}[x]) \le 0, \quad \forall x \in \mathcal{R}^n, z \in \mathcal{K}. \tag{5}$$

$$\|P_{\mathcal{K}}[x] - P_{\mathcal{K}}[y]\|^2 \le \|x-y\|^2 - \|P_{\mathcal{K}}[x] - x + y - P_{\mathcal{K}}[y]\|^2, \quad \forall x, y \in \mathcal{R}^n. \tag{6}$$

It follows from (6) that

$$\|P_{\mathcal{K}}[x] - P_{\mathcal{K}}[y]\| \leq \|x - y\|, \quad \forall x, y \in \mathcal{R}^n. \tag{7}$$

Definition 2.1 (a) The underlying mapping f is said to be monotone on $\mathcal{R}^n$ if

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

(b) The underlying mapping f is said to be Lipschitz continuous on $\mathcal{R}^n$ if there exists a constant $L > 0$ such that

$$\|f(x) - f(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathcal{R}^n.$$

It is well known from Eaves [9] that VI(F,$\mathcal{U}$) is equivalent to the following projection equation

$$u = P_{\mathcal{U}}[u - \beta F(u)], \tag{8}$$

where $\beta > 0$ is an arbitrary but fixed parameter. Let

$$e(u, \beta) := \begin{pmatrix} e_1(u, \beta) \\ e_2(u, \beta) \end{pmatrix} = \begin{pmatrix} x - P_{\mathcal{X}}[x - \beta(f(x) - A^\top y)] \\ y - P_{\mathcal{Y}}[y - \beta(Ax - b)] \end{pmatrix} \tag{9}$$

denote the residual error of the projection equation, then solving VI(F,$\mathcal{U}$) is equivalent to finding the zero points of the residual function $e(u, \beta)$. We define

$$r(u, \beta) := \begin{pmatrix} r_1(u, \beta) \\ r_2(u, \beta) \end{pmatrix} = \begin{pmatrix} x - P_{\mathcal{X}}[x - \beta(f(x) - A^\top y)] \\ y - P_{\mathcal{Y}}[y - \beta(A(x - r_1(u, \beta)) - b)] \end{pmatrix}. \tag{10}$$

From $e_1(u, \beta) = r_1(u, \beta)$, it is obvious that finding zeros of $e(u, \beta)$ is equivalent to finding zeros of $r(u, \beta)$.

We make the following standard assumptions throughout this paper:

Assumptions. • f is a monotone and Lipschitz continuous mapping on $\mathcal{X}$.

• The solution set of problem VI(F,$\mathcal{U}$) , denoted by $\mathcal{U}^*$, is nonempty.

• $\mathcal{X}$ is a simple closed convex set. That is, the projection onto the set is simple to carry out(e.g., $\mathcal{X}$ is the nonnegative orthant $\mathcal{R}^n_+$, or more generally, a box).

3 Main results

In this section, we describe our new alternating direction method formally, and prove its global convergence. Firstly, setting $r_i = r_i(u, \beta), i = 1, 2$.

Lemma 3.1. Let $u^* = (x^*, y^*) \in \mathcal{U}^*$ be an arbitrary solution of VI(F,$\mathcal{U}$), and the function f be a monotone and Lipschitz continuous function, then for any $u = (x, y) \in R^{n+m}$, we have

$$(u - u^*)^\top d(u, \beta) \geq \phi(u, \beta),$$

where

$$d(u,\beta) := \begin{pmatrix} r_1 + \beta A^\top r_2 - \beta f(x) + \beta f(x - r_1) \\ r_2 \end{pmatrix},$$

and

$$\phi(u,\beta) = \|r_1\|^2 + \|r_2\|^2 - \beta r_1^\top (f(x) - f(x - r_1)) + \beta r_2^\top A r_1.$$

Proof. Setting $x := x - \beta(f(x) - A^\top y)$ and $z := x^*$ in (5), we have

$$\{x - \beta(f(x) - A^\top y) - P_{\mathcal{X}}[x - \beta(f(x) - A^\top y)]\}^\top \{P_{\mathcal{X}}[x - \beta(f(x) - A^\top y)] - x^*\} \geq 0,$$

combining with (10), we have,

$$[r_1 - \beta(f(x) - A^\top y)]^\top (x - x^* - r_1) \geq 0,$$

that is

$$r_1^\top (x - x^*) \geq \|r_1\|^2 + \beta[f(x) - A^\top y]^\top (x - x^* - r_1). \tag{11}$$

From the assumption that u^* is a solution of VI(F,$\mathcal{U}$), we get

$$\beta(P_{\mathcal{X}}[x - \beta(f(x) - A^\top y)] - x^*)^\top (f(x^*) - A^\top y^*) \geq 0.$$

i.e.,

$$\beta(x - x^* - r_1)^\top (f(x^*) - A^\top y^*) \geq 0. \tag{12}$$

On the other hand, using the monotonicity of f, we obtain

$$\beta(f(P_{\mathcal{X}}[x - \beta(f(x) - A^\top y)]) - f(x^*))^\top (P_{\mathcal{X}}[x - \beta(f(x) - A^\top y)] - x^*) \geq 0,$$

i.e.,

$$\beta f(x - r_1)^\top (x - x^*) \geq \beta r_1^\top (f(x - r_1) - f(x^*)) + \beta(x - x^*)^\top f(x^*). \tag{13}$$

Then adding (11)-(13), it follows that

$$\begin{aligned} & (x - x^*)^\top (r_1 + \beta f(x - r_1)) \\ \geq\ & \|r_1\|^2 + \beta(x - x^* - r_1)^\top [f(x) - f(x^*) - A^\top (y - y^*)] \\ & + \beta r_1^\top (f(x - r_1) - f(x^*)) + \beta(x - x^*)^\top f(x^*) \\ =\ & \|r_1\|^2 + \beta(x - x^* - r_1)^\top [f(x) - A^\top (y - y^*)] + \beta r_1^\top f(x - r_1) \\ =\ & \|r_1\|^2 + \beta(x - x^*)^\top f(x) - \beta r_1^\top (f(x) - f(x - r_1)) \\ & - \beta(x - x^* - r_1)^\top A^\top (y - y^*). \end{aligned} \tag{14}$$

Similarly, setting $x := y - \beta(A(x - r_1(u,\beta)) - b)$ and $z := y^*$ in (5), we have

$$\{y - \beta(A(x - r_1) - b) - P_{\mathcal{Y}}[y - \beta(A(x - r_1) - b)]\}^\top \{P_{\mathcal{Y}}[y - \beta(A(x - r_1) - b)] - y^*\} \geq 0,$$

i.e.,

$$\{r_2 - \beta[A(x - r_1) - b]\}^\top (y - y^* - r_2) \geq 0. \tag{15}$$

From the assumption that u^* is a solution of VI(F,$\mathcal{U}$) again, we get

$$\beta\{P_\mathcal{Y}[y - \beta(A(x - r_1) - b)] - y^*\}^\top (Ax^* - b) \geq 0,$$

i.e.,

$$\beta(y - y^* - r_2)^\top (Ax^* - b) \geq 0. \tag{16}$$

From (15)+(16), we have

$$(y - y^*)^\top r_2 \geq \|r_2\|^2 + \beta(y - y^* - r_2)^\top A(x - x^* - r_1). \tag{17}$$

Then, we can get the assertion of this lemma follows directly from (14) and (17). The proof is completed.

Lemma 3.2. Let f be Lipschitz continuous with constant L and $\phi(u, \beta)$ be defined as in Lemma 3.1, $\beta < 2/(2L + \|A\|)$, then if u is not a solution of VI(F,$\mathcal{U}$), we have

$$(u - u^*)^\top d(u, \beta) \geq \phi(u, \beta) \geq \tau\|r(u, \beta)\|^2 > 0,$$

where $\tau = 1 - \beta(L + \frac{\|A\|}{2}) > 0$.

Proof. The assertion of this lemma is easy to get from Cauchy-Schwartz Inequality and the Lipschitz continuity of f. This completes the proof.

Remark 3.1 Lemma 3.2 shows that $-d(u, \beta)$ is a descent direction of the distance function $\|u - u^*\|^2/2$ whenever u is not a solution of VI(F,$\mathcal{U}$). This fact has motivated us to construct the following algorithm.

Algorithm 3.1 Improved Alternating Direction Method

Step 0: Given $\varepsilon > 0$. Choose $u^0 \in \mathcal{U}$, $\gamma \in (1, 2)$, $\beta \in (0, 2/(2L + \|A\|))$ and set $k := 0$;

Step 1: If $\|r(u^k, \beta)\| < \varepsilon$, then stop; otherwise, goto Step 2.

Step 2: Calculate $d(u^k, \beta)$ from Lemma 3.2 and the optimal step size

$$\rho_k = \tau\|r(u^k, \beta)\|^2/\|d(u^k, \beta)\|^2.$$

Step 3: Calculate the new iterate $u^{k+1} = P_\mathcal{U}[u^k - \gamma\rho_k d(u^k, \beta)]$. Set $k := k + 1$, go to Step 1.

Remark 3.2 If $\mathcal{Y}$ is a proper subset of $\mathcal{R}^m$, our method is also effective.

Remark 3.3 From Lemma 3.2, we can adopt an Armijo line search procedure to remove the global Lipschitz of $f(\cdot)$, because the Lipschitz constant L is difficult to evaluate even if $f(\cdot)$ is an affine mapping.

Now, we begin to investigate convergence of the proposed method.

Theorem 3.1. Let u^* be a solution of VI(F,$\mathcal{U}$) and let $\{u^k\} = \{(x^k, y^k)\}$ be the sequence obtained from the Algorithm 3.1. Then $\{u^k\}$ is bounded and

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \gamma\tau\rho_k(2 - \gamma)\|r(u^k, \beta)\|^2.$$

Proof. From the nonexpansivity of the projection operator and $u^* \in \mathcal{U}$, it follows that

$$\begin{aligned} & \|u^{k+1} - u^*\|^2 \\ \leq\ & \|u^k - \gamma\rho_k d(u^k, \beta) - u^*\|^2 \\ =\ & \|u^k - u^*\|^2 - 2\gamma\rho_k (u^k - u^*)^\top d(u^k, \beta) + \gamma^2\rho_k^2 \|d(u^k, \beta)\|^2 \\ \leq\ & \|u^k - u^*\|^2 - 2\gamma\rho_k\tau \|r(u^k, \beta)\|^2 + \gamma^2\rho_k^2 \|d(u^k, \beta)\|^2 \\ =\ & \|u^k - u^*\|^2 - \gamma\tau\rho_k(2-\gamma)\|r(u^k, \beta)\|^2, \end{aligned}$$

where the second inequality follows from Lemma 3.2 and the last equality follows from the definition of ρ_k. Thus, $\{u^k\}$ is bounded from the above inequality.

Lemma 3.3 Suppose that $\beta < 2/(2L + \|A\|)$. Then for any $k \geq 0$, there is $\varpi > 0$, such that

$$\rho_k \geq \varpi.$$

Proof. The proof is simple, thus is omitted.

Now, we give the convergence of Algorithm 3.1.

Theorem 3.2 Suppose that the assumptions in Theorem 3.1 hold. Then, the whole sequence $\{u^k\}$ converges to a solution of VI(F,$\mathcal{U}$).

Proof. From Theorem 3.1 and Lemma 3.3, we have

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \gamma\tau\varpi(2-\gamma)\tau\|r(u^k, \beta)\|^2.$$

Thus

$$\sum_{k=0}^{\infty} \|r(u^k, \beta)\|^2 < \infty,$$

which means that

$$\lim_{k\to\infty} \|r(u^k, \beta)\| = 0. \tag{18}$$

Since $\{u^k\}$ is also bounded, it has at least one cluster point. Let $u^\infty = (x^\infty, y^\infty)$ be a cluster point of the sequence $\{u^k\}$ and the subsequence $\{u^{k_j}\}$ converges to u^∞. Thus, from (18), we have

$$\|r(u^\infty, \beta)\| = 0,$$

which implies that $u^\infty \in \mathcal{U}^*$. Setting $u^* := u^\infty$, we again have

$$\|u^{k+1} - u^\infty\| \leq \|u^k - u^\infty\|,$$

and the whole sequence $\{u^k\}$ converges to u^∞, a solution of VI(F,$\mathcal{U}$). This completes the proof.

4 Preliminary Computational Results

In this section, we illustrate the efficiency of our methods. The example used here is the test problem in paper[5], which constraint set S and the mapping f are taken, respectively, as

$$S = \{x \in R^5_+ | \sum_{i=1}^{5} x_i = 10\},$$

and

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

where M is an $R^{5\times5}$ asymmetric positive matrix and $C_i(x) = \arctan(x_i-2), i = 1, 2, \cdots, 5$. The parameter ρ is used to vary the degree of asymmetry and nonlinearity. The data of example are illustrate as follows:

$$M = \begin{pmatrix} 0.726 & -0.949 & 0.266 & -1.193 & -0.504 \\ 1.645 & 0.678 & 0.333 & -0.217 & -1.443 \\ -1.016 & -0.225 & 0.769 & 0.943 & 1.007 \\ 1.063 & 0.587 & -1.144 & 0.550 & -0.548 \\ -0.256 & 1.453 & -1.073 & 0.509 & 1.026 \end{pmatrix}$$

and

$$q = (5.308,\ 0.008\ , -0.938,\ 1.024,\ \ -1.312)^{\top}.$$

In the experiment, we take the stopping criterion $\varepsilon = 10^{-6}$ as the initial point. All programs are coded in Matlab 7.1. 'IN' denotes the number of iterations and 'CPU' denotes the CPU time in seconds. The problem has a unique solution $x^* = (2,\ 2,\ 2,\ 2,\ 2)'$. We take $\beta = 0.06, \gamma = 1.96$ when $\rho = 10$ and $\beta = 0.006, \gamma = 1.98$ when $\rho = 20$. The parameters used in Han's method are same as those in [5]. The iteration numbers and the computational time for $\rho = 10$ and $\rho = 20$ are given in Table 1 and 2, respectively.

The results in the Table 1 and Table 2 indicate that the performance of Algorithm 3.1 is better than Han's method.

5 Conclusions

In this paper, we present a new alternating direction method for monotone structured VI(f, S). Total computational cost of the method is very tiny provided that the projection is easy to implement. Thus, the new method is applicable in practice.

Table 1: Numerical results for $\rho = 10$.

Starting point	Method	IN	CPU
(0 2.5 2.5 2.5 2.5)	Han's method	29	0.01
	Algorithm 3.1	18	0.01
(0 0 0 0 0)	Han's method	36	0.02
	Algorithm 3.1	17	0.01
(25 0 0 0 0)	Han's method	49	0.02
	Algorithm 3.1	36	0.01
(10 0 10 0 10)	Han's method	66	0.02
	Algorithm 3.1	26	0.01

Table 2: Numerical results for $\rho = 20$.

Starting point	Method	IN	CPU
(0 2.5 2.5 2.5 2.5)	Han's method	203	0.05
	Algorithm 3.1	53	0.01
(0 0 0 0 0)	Han's method	209	0.05
	Algorithm 3.1	42	0.01
(25 0 0 0 0)	Han's method	205	0.05
	Algorithm 3.1	56	0.01
(10 0 10 0 10)	Han's method	120	0.02
	Algorithm 3.1	44	0.01

References

[1] Gabay D., Applications of the method of multipliers to variational inequalities. In Fortin M, Glowinski R. eds, Augmented Lagrangian Methods: Applications to the Solution of Boundary-Valued Problems. North-Holland, Amsterdam, 299-331, 1983.

[2] Gabay D., Mercier B., A dual algorithm for the solution of nonlinear variational problems via finite-element approximations, *Computers and Mathematics with Applications*, **2(1)** (1976), 17-40.

[3] He B.S., Zhou J., A modified alternating direction method for convex minimization problems. *Appl. Math. Lett.*, **13**, 2000, 122-130.

[4] Han D.R., Hong K.Lo., New alternating direction method for a class of nonlinear variational inequality problems. *J. Optim. Theory Appl.*, **112(3)**, 2002, 549-560.

[5] Han D.R., A Modified Alternating Direction Method for Variational Inequality Problems. *Applied Mathematics and Optimization*, **45**, 2002, 63-74.

[6] Han D.R., A Proximal decomposition algorithm for Variational Inequality Problems. *Journal of Computational and Applied Mathematics*, **161**, 2003, 231-244.

[7] Sun M., A Descent Decomposition Method for Solving a Class of Variational Inequalities in Equilibrium Modeling. *ICIC Express Letters*, **5(7)**, 2011, 2327-2332.

[8] Abdellah B, Muhammad M.A., Mohamed K., Sheng Z.H., An alternating direction method for general variational inequalities. *J. Appl. Math. Comput.* , doi:10.1007/s12190-011-0495-y, 2011.

[9] Eaves B.C., On the basic theorem of complementarity. *Mathematical Programming*, **1**, 1971, 68-75.

[10] Ng M., Weiss P.A., and Yuan X.M., Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods. *SIAM J. Sci. Comput.*, **32**, 2010, 2710-2736.