

# On the $O(1/k)$ Convergence Rate of He's Alternating Directions Method for a Kind of Structured Variational Inequality Problem

Haiwen Xu<sup>1</sup>

<sup>1</sup> School of Computer Science, Civil Aviation Flight University of China, Guanghan, China

Correspondence: Haiwen Xu, School of Computer Science, Civil Aviation Flight University of China, Guanghan, 618307, China. E-mail: xuhaiwen.dream@163.com

Received: March 3, 2015 Accepted: March 17, 2015 Online Published: March 27, 2015

doi:10.5539/jmr.v7n2p69 URL: <http://dx.doi.org/10.5539/jmr.v7n2p69>

*The author was supported by Joint Fund of National Natural Science Foundation of China and Civil Aviation Administration of China (Grant No. U1233105).*

## Abstract

The alternating directions method for a kind of structured variational inequality problem (He, 2001) is an attractive method for structured monotone variational inequality problems. In each iteration, the subproblems are convex quadratic minimization problem with simple constraints and a well-conditioned system of nonlinear equations that can be efficiently solved using classical methods. Researchers have recently described the convergence rate of projection and contraction methods for variational inequality problems and the original ADM and its linearized variant. Motivated and inspired by research into the convergence rate of these methods, we provide a simple proof to show the  $O(1/k)$  convergence rate of alternating directions methods for structured monotone variational inequality problems (He, 2001).

**Keywords:** Structured variational inequality problems, monotone operators, alternating directions methods, convex quadratic minimization problems, convergence rate

## 1. Introduction

Let  $f$  be a continuous and monotone mapping from  $\mathbf{R}^n$  onto itself and let  $K \subset \mathbf{R}^m$  be a nonempty closed convex cone with vertex at the origin. Consider the variational inequality problem  $\text{VI}(S, f)$  (He, 2001): find  $x^* \in S$ , such that

$$\text{VI}(S, f) \quad (x - x^*)^T f(x^*) \geq 0, \quad \forall x \in S, \quad (1)$$

where  $A \in \mathbf{R}^{m \times n}$ ,  $b \in \mathbf{R}^m$ ,  $S = \{x \in \mathbf{R}^n | Ax - b \in K\}$ .

In fact, by attaching a Lagrange multiplier vector  $y \in Y$  to the constraint  $Ax - b \in K$ , we can obtain the structured variational inequality  $\text{VI}(\Omega, F)$  (He, 2001): find  $u^* \in \Omega$  such that

$$\text{VI}(\Omega, F) \quad (u - u^*)^T F(u^*) \geq 0, \quad \forall u \in \Omega, \quad (2)$$

where

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad F(u) = \begin{pmatrix} f(x) - A^T y \\ Ax - b \end{pmatrix}, \quad \Omega = \mathbf{R}^n \times Y,$$

and  $Y$  is a proper subset of  $\mathbf{R}^m$ .

Throughout we assume that the solution set of  $\text{VI}(\Omega, F)$ , denoted by  $\Omega^*$ , is nonempty. Let  $\|\cdot\|$  denote the Euclidean norm and  $P_K(y)$  the projection of  $y$  to the closed convex set  $K$ , that is,

$$P_K(y) = \arg \min\{\|y - x\| | x \in K\}, \quad \forall y \in \mathbf{R}^n.$$

The alternating directions method (ADM) is a classical method for solving variational inequality problems first introduced in the mid-1970s by Gabay and Mercier, and Glowinski and Marrocco, although similar ideas emerged in the mid-1950s (Douglas & Rachford, 1956, Peaceman & Rachford 1955). Much development of the ADM has been carried out in the areas of variational inequality problems and convex programming to yield simple single-processor algorithms (Boyd, Parikh, Chu, Peleato & Eckstein, 2010, Eckstein & Bertsekas, 1992, Gabay & Mercier, 1976, Gabay, 1983, Glowinski, 1984, He, 2001, He, Liao, Han & Yang, 2002, Peaceman & Rachford, 1955, Tao & Yuan, 2011).

To solve the structured variational inequality  $VI(\Omega, F)$ , the ADM of Gabay and Mercier produces the new iterate  $u^{k+1} = (x^{k+1}, y^{k+1})^T$  from a given  $u^k = (x^k, y^k)^T \in \Omega$  via the following steps (Gabay & Mercier, 1976, Gabay, 1983):

Step 1. Find  $x^{k+1}$  such that

$$(x' - x^{k+1})^T \{f(x^{k+1}) - A^T P_Y[y^k - (Ax^{k+1} - b)]\} \geq 0, \forall x' \in \mathbb{R}^n. \quad (3)$$

Step 2. The multipliers  $y^{k+1}$  are updated according to

$$y^{k+1} = P_Y[y^k - (Ax^{k+1} - b)]. \quad (4)$$

Subproblem (3) can solve a variational inequality problem or a nonlinear equation that includes  $f(x^{k+1})$  and  $P_Y[y^k - (Ax^{k+1} - b)]$ . Moreover, subproblem (4) is also a variational inequality problem. Naturally, it is difficult to solve the structured variational inequality  $VI(\Omega, F)$  using the original ADM.

He recently provided a modified ADM for  $VI(\Omega, F)$  that only solves a convex quadratic minimization problem with simple constraints and a system of strongly monotonic nonlinear equations in each iteration (He, 2001). This approach is obviously easier than solving (3) and (4) in many situations. Subsequently, we review the He's ADM and related results (He, 2001).

**The He's ADM** (He, 2001):

Given  $\gamma \in (0, 2)$ , and  $x^0 \in \mathbb{R}^n$ .

Step 1. Compute  $y^k$  such that

$$(y' - y)^T \{(Ax^k - b) - A[f(x^k) - A^T y]\} \geq 0, \forall y' \in Y. \quad (5)$$

Step 2. If  $\|f(x^k) - A^T y^k\| > \epsilon$ , update  $x^{k+1}$  according to

$$x^{k+1} + f(x^{k+1}) = x^k + f(x^k) - \gamma[f(x^k) - A^T y^k]. \quad (6)$$

Subproblem (5) is a symmetric linear variational inequality and it can be solved according to the convex quadratic minimization problem

$$\min\{\frac{1}{2}y^T M y + q^T y | y \in Y\},$$

where  $M = AA^T$  and  $q = (Ax^k - b) - A f(x^k)$ . In particular, if  $Y = \mathbb{R}_+^m$ , (5) is equivalent to the symmetric linear complementarity problem

$$y \geq 0, AA^T y + q \geq 0, y^T (AA^T y + q) = 0,$$

which can be solved by the pivotal Lemke method (Lemke, 1965). Numerical experiments showed that (5) can be solved within a finite number of steps (He, 2001). Moreover, (6) can be solved by some triangular decomposition to  $I + \nabla f(x)$  while solving  $x + f(x) - c_k = 0$  using Newton methods (Eckstein & Bertsekas, 1992), where

$$c_k = x^k + f(x^k) - \gamma[f(x^k) - A^T y^k].$$

In fact,  $f(x)$  is a monotone operator and then  $x + f(x)$  is a strictly monotone operator, so (6) is a well-conditioned system of nonlinear equations. Naturally, the subproblems in the He's ADM are easier than those in the classical ADM.

Nesterov introduced a method for solving a convex programming problem and provided a simple proof of the  $O(1/k^2)$  convergence rate of the method (Nesterov, 1983), which is derived from the Nemirovski non-relaxational

minimizing sequence. Nemirovski subsequently showed that the extragradient method has the  $O(1/t)$  convergence rate for variational inequalities with Lipschitz continuous monotone operators (Nemirovski, 2005). He recently described the convergence rate of projection and contraction methods under a unified conceptual framework (He, 2011). It is worth noting that the  $O(1/t)$  convergence rate of the ADM is given in the literature (He, 2012, Tao & Yuan, 2011). Much research has been devoted to this topic (Han & Yuan, 2014, Li, 2014, Shen, Chen & Cai, 2013, Jia, Guo & Cai, 2013, Yuan, 2012). Motivated and inspired by the encouraging progress for the convergence rate, we consider the  $O(1/k)$  convergence rate of He’s ADM.

The remainder of the paper is organized as follows. Section 2 presents some basic definitions and fundamental lemmas that are useful for the convergence rate analysis. In Section 3, the convergence rate of He’s ADM (He, 2001) is proved. Finally, conclusions are drawn in Section 4.

**2. Preliminaries**

In this section, we provide some important lemmas for subsequent analysis. The proofs are widely available in the literature (Facchinei & Pang, 2003, He, 2001).

**Definition 1** *Let  $f$  be a continuous mapping from  $R^n$  onto itself. Then  $f(x)$  is a monotone operator if and only if  $(x-y)^T[f(x) - f(y)] \geq 0, \forall x, y \in R^n$ .*

In fact, if  $f(x)$  is a monotone operator, then

$$F(u) = \begin{pmatrix} f(x) - A^T y \\ Ax - b \end{pmatrix}$$

is a monotone operator.

According to the relationship between a variational inequality and a fixed point problem, we can easily obtain the following lemma.

**Lemma 1** *When  $\Omega = R^n \times Y$  is a closed and convex set,  $u^*$  is the solution of  $VI(F, \Omega)$  if and only if*

$$u^* = P_{\Omega}[u^* - F(u^*)].$$

In fact, the solution of structured variational inequality  $VI(F, \Omega)(2)$  is the zero point of

$$e(u) := u - P_{\Omega}[u - F(u)] = \begin{pmatrix} f(x) - A^T y \\ y - P_Y[y - (Ax - b)] \end{pmatrix}.$$

Subsequently, we recall some previous results (He, 2001) for proof of the convergence rate.

**Lemma 2** (He, 2001, Remark 1) *For a given  $u^k = (x^k, y^k)^T \in R^n \times Y$ , let  $\{u^k\}$  be produced by the He’s ADM. Then we have*

$$\frac{1}{1 + \|A\|} \|e(u^k)\| \leq \|f(x^k) - A^T y^k\| \leq \|e(u^k)\|, \forall \text{ integer } k \geq 0.$$

By Lemma , taking  $\|f(x^k) - A^T y^k\| \leq \epsilon$  as a stopping criterion is reasonable.

**Lemma 3** (He, 2001, Proposition 2) *For a given  $u^* = (x^*, y^*)^T \in \Omega^*, 0 < \gamma < 2$ , let  $\{u^k\}$  be produced by the He’s ADM. Then we have*

$$\|(x^{k+1} - x^*) + [f(x^{k+1}) - f(x^*)]\|^2 \leq \|(x^k - x^*) + [f(x^k) - f(x^*)]\|^2 - \gamma(2 - \gamma)\|f(x^k) - A^T y^k\|^2.$$

The most important properties of  $\|f(x^k) - A^T y^k\|^2$  are summarized in the following. For integrity of the proof, we provide a proof of the following lemma. Moreover, for convenience, we denote  $x^k, y^k, x^{k+1}$ , and  $y^{k+1}$  by  $x, y, \tilde{x}$ , and  $\tilde{y}$ , respectively.

**Lemma 4** (He, 2001, Proposition 4) *Let  $u = (x, y)^T$  be the sequence produced by the He's ADM,  $0 < \gamma < 2$ . Then we have*

$$\|f(\tilde{x}) - A^T \tilde{y}\|^2 \leq \|f(x) - A^T y\|^2 - \frac{2-\gamma}{\gamma} \| [f(\tilde{x}) - A^T \tilde{y}] - [f(x) - A^T y] \|^2.$$

*Proof.* By a series of computations, we can obtain

$$\begin{aligned} \|f(\tilde{x}) - A^T \tilde{y}\|^2 &= \|(f(x) - A^T y) + [(f(\tilde{x}) - A^T \tilde{y}) - (f(x) - A^T y)]\|^2 \\ &= \|f(x) - A^T y\|^2 + \|(f(\tilde{x}) - A^T \tilde{y}) - (f(x) - A^T y)\|^2 \\ &\quad + 2(f(x) - A^T y)^T [(f(\tilde{x}) - A^T \tilde{y}) - (f(x) - A^T y)]. \end{aligned}$$

Hence, we need only to show that

$$\gamma(f(x) - A^T y)^T [(f(x) - A^T y) - (f(\tilde{x}) - A^T \tilde{y})] \geq \|(f(x) - A^T y) - (f(\tilde{x}) - A^T \tilde{y})\|^2.$$

According to (5), we have

$$(\tilde{y} - y)^T [(Ax - b) - A(f(x) - A^T y)] \geq 0 \quad (7)$$

and

$$(y - \tilde{y})^T [(A\tilde{x} - b) - A(f(\tilde{x}) - A^T \tilde{y})] \geq 0. \quad (8)$$

Adding (7) and (8), we get

$$(\tilde{y} - y)^T \{A(x - \tilde{x}) - A[(f(x) - A^T y) - (f(\tilde{x}) - A^T \tilde{y})]\} \geq 0,$$

i.e.,

$$(A^T \tilde{y} - A^T y)^T (x - \tilde{x}) + (A^T y - A^T \tilde{y})^T [f(x) - A^T y - (f(\tilde{x}) - A^T \tilde{y})] \geq 0. \quad (9)$$

Note that  $(A^T \tilde{y} - A^T y)^T (x - \tilde{x})$  can be changed to

$$\begin{aligned} &(A^T \tilde{y} - A^T y)^T (x - \tilde{x}) \\ &= ((f(x) - A^T y) - (f(\tilde{x}) - A^T \tilde{y}) + (f(\tilde{x}) - f(x)))^T (x - \tilde{x}) \\ &= (x - \tilde{x})^T [(f(x) - A^T y) - (f(\tilde{x}) - A^T \tilde{y})] + (f(\tilde{x}) - f(x))^T (x - \tilde{x}). \end{aligned} \quad (10)$$

Furthermore,  $(A^T y - A^T \tilde{y})^T [f(x) - A^T y - (f(\tilde{x}) - A^T \tilde{y})]$  can be changed to

$$\begin{aligned} &(A^T y - A^T \tilde{y})^T [f(x) - A^T y - (f(\tilde{x}) - A^T \tilde{y})] \\ &= [f(x) - f(\tilde{x})]^T [(f(x) - A^T y) - (f(\tilde{x}) - A^T \tilde{y})] \\ &\quad - \|(f(x) - A^T y) - (f(\tilde{x}) - A^T \tilde{y})\|^2. \end{aligned} \quad (11)$$

Since  $f$  is monotone, we have

$$[f(x) - f(\tilde{x})]^T (x - \tilde{x}) \leq 0. \quad (12)$$

According to (9), (10), (11) and (12), we obtain

$$\begin{aligned} &([f(x) - f(\tilde{x})]^T (x - \tilde{x}) + (f(x) - f(\tilde{x}))^T [f(x) - A^T y - (f(\tilde{x}) - A^T \tilde{y})]) \\ &- \|(f(x) - A^T y) - (f(\tilde{x}) - A^T \tilde{y})\|^2 \\ &\geq (A^T \tilde{y} - A^T y)^T (x - \tilde{x}) + (A^T y - A^T \tilde{y})^T [f(x) - A^T y - (f(\tilde{x}) - A^T \tilde{y})] \\ &\geq 0. \end{aligned}$$

With  $(x - \tilde{x}) + (f(x) - f(\tilde{x})) = \gamma(f(x) - A^T y)$ , we get

$$\gamma(f(x) - A^T y)^T [f(x) - A^T y - (f(\tilde{x}) - A^T \tilde{y})] \geq \|f(x) - A^T y - (f(\tilde{x}) - A^T \tilde{y})\|^2.$$

Hence the lemma is proved.

### 3. Convergence Rate

In this section, we analyze the  $O(1/k)$  convergence rate of He's ADM.

**Theorem 1** Let  $\{u^k\}$  be produced by the He's ADM,  $0 < \gamma < 2$ . Then for all integer  $k \geq 0$ , we have

$$\|e(u^k)\|^2 \leq \frac{1}{(k+1)\gamma(2-\gamma)}(1 + \|A\|^2)\|(x^0 - x^*) + [f(x^0) - f(x^*)]\|^2.$$

**Proof.** According to Lemma 2, we have

$$(k+1)\gamma(2-\gamma)\|e(u^k)\|^2 \leq (k+1)\gamma(2-\gamma)(1 + \|A\|^2)\|f(x^k) - A^T y^k\|^2.$$

Applying Lemma 4, we obtain

$$\|f(x^{k+1}) - A^T y^{k+1}\|^2 \leq \|f(x^k) - A^T y^k\|^2, \forall \text{ integer } k \geq 0,$$

so it immediately follows that

$$(k+1)\gamma(2-\gamma)(1 + \|A\|^2)\|f(x^k) - A^T y^k\|^2 \leq \gamma(2-\gamma)(1 + \|A\|^2) \sum_{t=0}^k \|f(x^t) - A^T y^t\|^2.$$

Furthermore, by Lemma 3 we can obtain

$$\begin{aligned} \gamma(2-\gamma) \sum_{t=0}^k \|f(x^t) - A^T y^t\|^2 &\leq \gamma(2-\gamma) \sum_{t=0}^{\infty} \|f(x^t) - A^T y^t\|^2 \\ &\leq \|(x^0 - x^*) + [f(x^0) - f(x^*)]\|^2. \end{aligned}$$

Thus, we have

$$\|e(u^k)\|^2 \leq \frac{1}{(k+1)\gamma(2-\gamma)}(1 + \|A\|^2)\|(x^0 - x^*) + [f(x^0) - f(x^*)]\|^2, 0 < \gamma < 2.$$

In fact, when we take  $\|e(u^k)\|^2 \leq \epsilon$  or  $\|f(x^k) - A^T y^k\|^2 \leq \epsilon$  as the stopping criterion, by Lemma 1 and Lemma 2 and Theorem 1, we obtain the  $O(1/k)$  convergence rate of He's ADM.

### 4. Conclusions

We proved the  $O(1/k)$  convergence rate of He's ADM. This result can be considered an improvement and refinement of previous results (He, 2001).

### Acknowledgements

The author was supported by Joint Fund of National Natural Science Foundation of China and Civil Aviation Administration of China (Grant No. U1233105) and Science Foundation of the Civil Aviation Flight University of China (Grant No. J2010-45).

### References

- Boyd, S., Parikh, N., Chu, E., Peleato, B., & Eckstein, J. (2010). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learning*, 3, 1-122. <http://dx.doi.org/10.1561/22000000016>
- Chen, C., Shen, Y., & You, Y. (2013). On the Convergence Analysis of the Alternating Direction Method of Multipliers with Three Blocks. *Abstract and Applied Analysis*. <http://dx.doi.org/10.1155/2013/183961>
- Douglas, J., & Rachford, H. H. (1956). On the numerical solution of the heat conduction problem in 2 and 3 space variables. *Transactions of the American Mathematical Society*, 82, 421-439. <http://dx.doi.org/10.1090/S0002-9947-1956-0084194-4>

- Eckstein, J., & Bertsekas, D. P. (1992). On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program*, 55, 293-318. <http://dx.doi.org/10.1007/BF01581204>
- Facchinei, F., & Pang, J. S. (2003). Finite-Dimensional Variational Inequalities and Complementarity Problems, Vol. I and II. *Springer Series in Operations Research*. Springer Verlag, New York.
- Gabay, D., & Mercier, B. (1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximations. *Computers & Mathematics with Applications*, 2, 17-40. [http://dx.doi.org/10.1016/0898-1221\(76\)90003-1](http://dx.doi.org/10.1016/0898-1221(76)90003-1)
- Gabay, D. (1983). Applications of the method of multipliers to variational inequalities, Augmented Lagrange Methods: Applications to the Solution of Boundary-valued Problems, M. Fortin, R. Glowinski, eds., *North Holland, Amsterdam*, 299-331.
- Glowinski, R. (1984). Numerical Methods for Nonlinear Variational Problems. Springer, New York. <http://dx.doi.org/10.1007/978-3-662-12613-4>
- He, B. S. (2001). On the different structured variational inequalities and the related alternating directions methods. manuscript, Nanjing University.
- He, B. S., Liao, L. Z., Han, D. R., & Yang, H. (2002). A new inexact alternating directions method for monotone variational inequalities. *Math. Program*, 92, 103-118. <http://dx.doi.org/10.1007/s101070100280>
- He, B. S., & Yuan, X. M. (2012). On the  $O(1/t)$  convergence rate of the alternating direction method, *SIAM Journal on Numerical Analysis*, 50(2), 700-709. <http://dx.doi.org/10.1137/110836936>
- He, B. S. (2011). On the  $O(1/t)$  convergence rate of the projection and contraction methods for variational inequalities with Lipschitz continuous monotone operators. [http://www.optimization-online.org/DB\\_HTML/2011/07/3097.html](http://www.optimization-online.org/DB_HTML/2011/07/3097.html)
- Han, D., & Yuan, X. (2013). Local Linear Convergence of the alternating direction method of multipliers for quadratic programs, *SIAM Journal on numerical analysis*, 51(6), 3446-3457. <http://dx.doi.org/10.1137/120886753>
- Jia, Z., Guo, K., & Cai, X. (2013). Convergence Analysis of Alternating Direction Method of Multipliers for a Class of Separable Convex Programming, *Abstract and Applied Analysis*. <http://dx.doi.org/10.1155/2013/680768>
- Lemke, C. E. (1965). Bimatrix equilibrium points and mathematical programming. *Management Science*, 11, 681-689. <http://dx.doi.org/10.1287/mnsc.11.7.681>
- Li, M., & You, Y. (2013). Convergence Analysis of the Relaxed Proximal Point Algorithm. *Abstract and Applied Analysis*. <http://dx.doi.org/10.1155/2013/912846>
- Nemirovski, A. (2005). Prox-method with rate of convergence  $O(1/t)$  for variational inequality with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 15, 229-251. <http://dx.doi.org/10.1137/S1052623403425629>
- Nesterov, Y. E. (1983). A method for unconstrained convex minimization problem with the rate of convergence  $O(1/k^2)$ . *Doklady an SSSR*, 269, 543-547.
- Peaceman, D. W., & Rachford, H. H. (1955). The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, 3, 28-41. <http://dx.doi.org/10.1137/0103003>
- Tao, M., & Yuan, X. M. (2011). Recovering low-rank and sparse components of matrices from incomplete and noisy observations. *SIAM Journal on optimization*, 21(1), 57-81. <http://dx.doi.org/10.1137/100781894>
- Tao, M., & Yuan, X. M. (2012). On the  $O(1/t)$  convergence rate of alternating method with logarithmic-quadratic proximal regularization. *SIAM Journal on optimization*, 22(4), 1431-1448. <http://dx.doi.org/10.1137/110847639>
- Yuan, S., & Xu, M. (2013). On the  $O(1/t)$  convergence rate of YeCYuans modified alternating direction method of multipliers. *Abstract and Applied Analysis*.
- Yuan, X. M. (2012). Alternating direction methods for covariance selection models. *Journal of scientific Computing*, 5(2), 261-273. <http://dx.doi.org/10.1007/s10915-011-9507-1>

### **Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).