JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 106, No. 2, pp. 337–356, AUGUST 2000

# Alternating Direction Method with Self-Adaptive Penalty Parameters for Monotone Variational Inequalities[1]

B. S. HE,[2] H. YANG,[3] AND S. L. WANG[4]

Communicated by R. Glowinski

**Abstract.** The alternating direction method is one of the attractive approaches for solving linearly constrained separate monotone variational inequalities. Experience on applications has shown that the number of iterations depends significantly on the penalty parameter for the system of linear constraint equations. While the penalty parameter is a constant in the original method, in this paper we present a modified alternating direction method that adjusts the penalty parameter per iteration based on the iterate message. Preliminary numerical tests show that the self-adaptive adjustment technique is effective in practice.

**Key Words.** Monotone variational inequalities, alternating direction method, variable penalty parameters.

## 1. Introduction

Variational inequality problem consists of finding a vector $u^* \in \Omega$ such that

$$(\mathrm{VI}(\Omega, F)) \quad (u-u^*)^T F(u^*) \geq 0, \qquad \forall u \in \Omega, \tag{1}$$

where $\Omega$ is a nonempty closed convex subset of $\mathscr{R}^l$ and $F$ is a continuous mapping from $\mathscr{R}^l$ to itself. In practice, many VI problems have the following

[1]This first author was supported by NSFC Grant 19971040. The second author was supported in part by UGC Grant RI94/95 EG01.

[2]Professor, State Key Lab. for Novel Software Technology, Department of Mathematics, Nanjing University, Nanjing, China.
[3]Associate Professor, Department of Civil Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China.
[4]PhD Student, Department of Mathematics, Nanjing University, Nanjing, China.

separable structure (e.g., Ref. 1):

$$u=\begin{bmatrix} x \\ y \end{bmatrix}, \qquad F(u)=\begin{bmatrix} f(x) \\ g(y) \end{bmatrix}, \tag{2}$$

$$\Omega=\{(x,y)\,|\,x\in\mathscr{X},\ y\in\mathscr{Y},\ Ax+By=b\}, \tag{3}$$

where $\mathscr{X}\subset\mathscr{R}^n$ and $\mathscr{Y}\subset\mathscr{R}^m$ are given closed convex sets, $f\colon\mathscr{X}\to\mathscr{R}^n$ and $g\colon\mathscr{Y}\to\mathscr{R}^m$ are given monotone operators, $A\in\mathscr{R}^{r\times n}$ and $B\in\mathscr{R}^{r\times m}$ are given matrices, and $b$ is a given vector in $\mathscr{R}^r$.

By attaching a Lagrange multiplier vector $\lambda\in\mathscr{R}^r$ to the linear constraints $Ax+By=b$, the problem under consideration can be treated as a mixed variational inequality (VI with equality restriction $Ax+By=b$ and unrestricted variable $\lambda$): Find $w^*\in\mathscr{W}$, such that

$$(\mathrm{MVI}(\mathscr{W},Q))\quad (w-w^*)^TQ(w^*)\geq 0, \qquad \forall w\in\mathscr{W}, \tag{4}$$

where

$$w=\begin{bmatrix} x \\ y \\ \lambda \end{bmatrix}, \qquad Q(w)=\begin{bmatrix} f(x)-A^T\lambda \\ g(y)-B^T\lambda \\ Ax+By-b \end{bmatrix}, \qquad \mathscr{W}=\mathscr{X}\times\mathscr{Y}\times\mathscr{R}^r. \tag{5}$$

Problem (4)–(5) will be considered in this paper.

Typically, problems in applications [for example, network economics (Ref. 1) and nonlinear mechanics (Refs. 2–4)] are quite large and are often solved by alternating direction methods. The alternating direction method, originally proposed by Gabay (Ref. 5) and by Gabay and Mercier (Ref. 6), is used frequently in the literature (Refs. 2–3). At each iteration of this method, the new iterate

$$w^{k+1}=(x^{k+1},y^{k+1},\lambda^{k+1})\in\mathscr{X}\times\mathscr{Y}\times\mathscr{R}^r$$

is generated from a given triple

$$w^k=(x^k,y^k,\lambda^k)\in\mathscr{X}\times\mathscr{Y}\times\mathscr{R}^r$$

by the following procedure: First, with $y^k$ and $\lambda^k$ held fixed, $x^{k+1}$ is obtained by solving

$$(x'-x^{k+1})^T\{f(x^{k+1})-A^T[\lambda^k-\beta(Ax^{k+1}+By^k-b)]\}\geq 0, \qquad \forall x'\in\mathscr{X}; \tag{6}$$

then, with $x^{k+1}$ and $\lambda^k$ held fixed, $y^{k+1}$ is produced by solving

$$(y'-y^{k+1})^T\{g(y^{k+1})-B^T[\lambda^k-\beta(Ax^{k+1}+By^{k+1}-b)]\}\geq 0, \qquad \forall y'\in\mathscr{Y}. \tag{7}$$

Finally, the multipliers are updated by

$$\lambda^{k+1} = \lambda^k - \gamma\beta(Ax^{k+1} + By^{k+1} - b), \tag{8}$$

where $\gamma \in (0, (1+\sqrt{5})/2)$ and $\beta > 0$ are given constants: $\gamma$ is the relaxation factor and $\beta$ is a penalty parameter for the linearly constrained equation $Ax + By - b = 0$. This method is referred to as a method of multipliers in the literature (Ref. 5), and the convergence proof can be found in Refs. 3, 6 for $B = I$ and in Refs. 7–8 for general $B$. Further studies and applications of such methods can be found in Glowinski (Ref. 3), Glowinski and Le Tallec (Ref. 4), and Eckstein and Fukushima (Ref. 9).

Experience on applications (Refs. 2, 10, 11) has shown that, if the fixed penalty $\beta$ is chosen too small or too large, the solution time can increase significantly. In order to improve such methods, some researchers suggested to use methods that replace the constant $\beta$ in (6)–(8) by a variable penalty sequence $\{\beta_k\}$. For example, Nagurney et al. (Ref. 1) utilized the method in which $\{\beta_k\}$ is monotonically increasing. In a more recent paper (Ref. 11), instead of the sequence $\{\beta_k\}$, Kontogiorgis and Meyer took a sequence of penalty symmetric positive-definite matrices $\{H_k\}$. The convergence was proved under the assumption that $\gamma \equiv 1$, the eigenvalues of $\{H_k\}$ are uniformly bounded from below away from zero, and with finitely many exceptions, the eigenvalues of $H_k - H_{k+1}$ are nonnegative. Most recently, He and Yang (Ref. 12) studied some convergence properties of the alternating direction method. The convergence was proved for the cases that $\gamma \in (0, (\sqrt{5}+1)/2)$ is fixed and the positive sequence $\{\beta_k\}$ is either monotone decreasing bounded from below away from zero or monotone increasing (Ref. 12).

In all the existing variants of the alternating direction method, the convergence theorem was proved under a monotonicity assumption on the sequence of the penalty parameters $\{\beta_k\}$ [resp. penalty matrices $\{H_k\}$]. For such methods, it is necessary to choose a proper starting penalty parameter $\beta_0$. In most cases, this is difficult to realize.

In this paper, we continue the research by relaxing some restrictions. Namely, we give a modified method that allows the parameter $\beta_k$ to either increase or decrease. In particular, we present some self-adaptive rules for adjusting the penalty sequence $\{\beta_k\}$ and prove convergence under such modifications.

The paper is organized as follows. In Section 2, we give the description of the proposed method and discuss how to adjust the parameter $\beta_k$ which satisfies the conditions and leads to fast convergence. The main theorem, which allows us to vary the parameter $\beta$, is proved in Section 3. In Section 4, based on the main theorem, we give some restrictions on the sequence $\{\beta_k\}$ and prove convergence under such conditions. Finally, in Section 5, we give results of preliminary numerical tests to demonstrate the advantage of the proposed modified method.

The following notation is used in this paper. For any real matrix $M$ and vector $v$, we denote transposition by $M^T$ and $v^T$, respectively. Superscripts such as in $v^k$ refer to specific vectors and are usually iteration indices. The Euclidean norm of the vector $z$ will be denoted by $\|z\|$; i.e.,

$$\|z\| = \sqrt{z^T z}.$$

## 2. Alternating Direction Method with Variable Penalty Parameter

It has been well known (e.g., see Refs. 1 and 13) that VI$(\Omega, F)$ is equivalent to the following projection equation:

$$(\text{PE}(\Omega, F)) \quad u = P_\Omega[u - F(u)],$$

where $P_\Omega(\cdot)$ denotes the projection on $\Omega$. The equivalence between VI$(\Omega, F)$ and PE$(\Omega, F)$ is a basic tool in our analysis. Hence, solving MVI$(\mathscr{W}, Q)$ is equivalent to finding a zero point of

$$e(w) := w - P_{\mathscr{W}}[w - Q(w)]. \tag{9}$$

Since the projection mapping is nonexpansive, and thus $e(w)$ is continuous, $\|e(w)\|$ can be viewed as a function measuring how much $w$ fails to be a solution of MVI$(\mathscr{W}, Q)$. We give the description of the method.

**Alternating Direction Method with Variable Penalty Parameter.**

Step 0. Given $\epsilon > 0$, $\gamma \in (0, (1+\sqrt{5})/2)$, $\beta_1 > 0$, $y^0 \in \mathscr{Y}$, and $\lambda^0 \in \mathscr{R}^r$, set $k = 0$.

Step 1. Find $x^{k+1} \in \mathscr{X}$ such that

$$(x' - x^{k+1})^T \{f(x^{k+1}) - A^T[\lambda^k - \beta_{k+1}(Ax^{k+1} + By^k - b)]\} \geq 0, \quad \forall x' \in \mathscr{X}. \tag{10}$$

Step 2. Find $y^{k+1} \in \mathscr{Y}$ such that

$$(y' - y^{k+1})^T \{g(y^{k+1}) - B^T[\lambda^k - \beta_{k+1}(Ax^{k+1} + By^{k+1} - b)]\} \geq 0, \quad \forall y' \in \mathscr{Y}. \tag{11}$$

Step 3. Update the multiplier,

$$\lambda^{k+1} = \lambda^k - \gamma\beta_{k+1}(Ax^{k+1} + By^{k+1} - b). \tag{12}$$

Step 4. Convergence verification: if $\|e(w^{k+1})\|_\infty < \epsilon$, stop.

Step 5. Adjust the penalty parameter $\beta > 0$ (this will be specified later); set $k := k+1$, and go to Step 1.

For convenience, we state the basic assumptions to guarantee that the problem under consideration is solvable and the method is well defined.

**Assumption A1.** The solution set of MVI($\mathscr{W}, Q$), denoted by $\mathscr{W}^*$, is nonempty.

**Assumption A2.** Problems (10) and (11) are solvable.

We focus our attention on only the framework of the alternating direction method with variable penalty parameter, rather than the subVIs (10) and (11) for which a number of solution methods can be found for instance in Ref. 13. Although the subVIs are solved numerically, we assume that the exact solutions of the subVIs in each iteration can be obtained without difficulty.

Now, we consider how to adjust the penalty parameter $\beta_k$ for fast convergence. Recall that solving MVI($\mathscr{W}, Q$) is equivalent to finding a zero point of $e(w)$ and

$$e(w) = \begin{bmatrix} e_x(w) \\ e_y(w) \\ e_\lambda(w) \end{bmatrix} = \begin{bmatrix} x - P_{\mathscr{X}}\{x - [f(x) - A^T\lambda]\} \\ y - P_{\mathscr{Y}}\{y - [g(y) - B^T\lambda]\} \\ Ax + By - b \end{bmatrix}. \tag{13}$$

For simplicity, let

$$w^k = (x^k, y^k, \lambda^k) \in \mathscr{X} \times \mathscr{Y} \times \mathscr{R}^r$$

be generated from a given triple

$$w^{k-1} = (x^{k-1}, y^{k-1}, \lambda^{k-1}) \in \mathscr{X} \times \mathscr{Y} \times \mathscr{R}^r$$

by (10)–(12) with $\gamma = 1$. Using the equivalence of the solutions of the variational inequality and the projection equation, we find that

$$y^k = P_{\mathscr{Y}}\{y^k - [g(y^k) - B^T\lambda^k]\},$$

and thus $e_y(w^k) = 0$. It follows that

$$\|e(w^k)\|^2 = \|e_x(w^k)\|^2 + \|e_\lambda(w^k)\|^2.$$

This consideration offers us a message on how to choose the penalty parameter $\beta$. For the sake of balance, we should adjust the penalty parameter $\beta$ such that $\|e_x(w)\| \approx \|e_\lambda(w)\|$. In other words, for an iterate $w = (x, y, \lambda)$, if

$\|e_x(w)\| \ll \|e_\lambda(w)\|$, we should increase $\beta$ in the next iteration; conversely, we decrease $\beta$ when $\|e_x(w)\| \gg \|e_\lambda(w)\|$. This is the basic idea in our modified alternating direction method. In detail, we consider the following different techniques for adjusting $\beta_k$.

Given $\mu \in (0, 1)$ and a nonnegative sequence $\{\tau_k\}$ that satisfies

$$\sum_{k=0}^{\infty} \tau_k < \infty, \tag{14}$$

we consider three strategies below.

Strategy S1. $\{\beta_k\}$ is monotonically increasing,

$$\beta_{k+1} = \begin{cases} \beta_k(1+\tau_k), & \text{if } \|x^k - P_{\mathscr{X}}[x^k - (f(x^k) - A^T\lambda^k)]\| < \mu\|Ax^k + By^k - b\|, \\ \beta_k, & \text{otherwise.} \end{cases}$$

Strategy S2. $\{\beta_k\}$ is monotonically decreasing,

$$\beta_{k+1} = \begin{cases} \beta_k/(1+\tau_k), & \text{if } \mu\|x^k - P_{\mathscr{X}}[x^k - (f(x^k) - A^T\lambda^k)]\| > \|Ax^k + By^k - b\|, \\ \beta_k, & \text{otherwise.} \end{cases}$$

Strategy S3. $\{\beta_k\}$ is a self-adaptive variable,

$$\beta_{k+1} = \begin{cases} \beta_k(1+\tau_k), & \text{if } \|x^k - P_{\mathscr{X}}[x^k - (f(x^k) - A^T\lambda^k)]\| < \mu\|Ax^k + By^k - b\|, \\ \beta_k/(1+\tau_k), & \text{if } \mu\|x^k - P_{\mathscr{X}}[x^k - (f(x^k) - A^T\lambda^k)]\| > \|Ax^k + By^k - b\|, \\ \beta_k, & \text{otherwise.} \end{cases}$$

In any cases, if $\beta_{k+1} \neq \beta_k$, then

$$\beta_{k+1} = (1+\tau_k)\beta_k \quad \text{or} \quad \beta_{k+1} = \beta_k/(1+\tau_k).$$

Under condition (14), we have

$$\prod_{i=1}^{\infty} (1+\tau_i) < +\infty.$$

Hence, the sequence $\{\beta_k\}$ is both upper bounded and bounded below away from zero; that is, we have

$$B_L := \inf_k\{\beta_k\} > 0, \qquad B_U := \sup_k\{\beta_k\} < +\infty. \tag{15}$$

## 3. Main Theorem

The task of this section is to establish a theorem that ensures the convergence of the proposed method with the different strategies for adjusting the penalty parameters mentioned in the last section. To start our discussion, let us observe the difference between $\|\lambda^k - \lambda^*\|^2$ and $\|\lambda^{k+1} - \lambda^*\|^2$. From [see (12)]

$$\lambda^k = \lambda^{k+1} + \gamma\beta_{k+1}(Ax^{k+1} + By^{k+1} - b),$$

$$\|\lambda^k - \lambda^*\|^2 = \|\lambda^{k+1} - \lambda^*\|^2 - \|\lambda^k - \lambda^{k+1}\|^2 + 2(\lambda^k - \lambda^*)^T(\lambda^k - \lambda^{k+1}),$$

we get

$$\|\lambda^k - \lambda^*\|^2 = \|\lambda^{k+1} - \lambda^*\|^2 - \gamma^2\beta_{k+1}^2\|Ax^{k+1} + By^{k+1} - b\|^2 + 2\gamma\beta_{k+1}(\lambda^k - \lambda^*)^T(Ax^{k+1} + By^{k+1} - b). \tag{16}$$

The following lemma provides a desirable property of the last term of (16).

**Lemma 3.1.** For any $w^* = (x^*, y^*, \lambda^*) \in \mathscr{W}^*$, we have

$$\begin{aligned}&(\lambda^k - \lambda^*)^T(Ax^{k+1} + By^{k+1} - b)\\ &\geq \beta_{k+1}\|Ax^{k+1} + By^{k+1} - b\|^2\\ &\quad + \beta_{k+1}(Ax^{k+1} - Ax^*)^T(By^k - By^{k+1}).\end{aligned} \tag{17}$$

**Proof.** Since $w^* \in \mathscr{W}^*$, $x^{k+1} \in \mathscr{X}$, and $y^{k+1} \in \mathscr{Y}$, we have

$$(x^{k+1} - x^*)^T[f(x^*) - A^T\lambda^*] \geq 0, \tag{18}$$

$$(y^{k+1} - y^*)^T[g(y^*) - B^T\lambda^*] \geq 0, \tag{19}$$

$$Ax^* + By^* - b = 0. \tag{20}$$

On the other hand, from (10) and (11), it follows that

$$(x^* - x^{k+1})^T\{f(x^{k+1}) - A^T[\lambda^k - \beta_{k+1}(Ax^{k+1} + By^k - b)]\} \geq 0, \tag{21}$$

$$(y^* - y^{k+1})^T\{g(y^{k+1}) - B^T[\lambda^k - \beta_{k+1}(Ax^{k+1} + By^{k+1} - b)]\} \geq 0. \tag{22}$$

Adding (18) and (21), and using the monotonicity of the operator $f$, we get

$$(x^{k+1} - x^*)^T\{A^T[(\lambda^k - \lambda^*) - \beta_{k+1}(Ax^{k+1} + By^k - b)]\} \geq 0. \tag{23}$$

Similarly, adding (19) and (22), and using the monotonicity of the operator $g$, it follows that

$$(y^{k+1} - y^*)^T\{B^T[(\lambda^k - \lambda^*) - \beta_{k+1}(Ax^{k+1} + By^{k+1} - b)]\} \geq 0. \tag{24}$$

Now combining (23) and (24) and using (20), we get the assertion of this lemma. □

From the result of Lemma 3.1, equation (16) becomes

$$\|\lambda^k-\lambda^*\|^2=\|\lambda^{k+1}-\lambda^*\|^2+\gamma(2-\gamma)\beta_{k+1}^2\|Ax^{k+1}+By^{k+1}-b\|^2 \\ +2\gamma\beta_{k+1}^2(Ax^{k+1}-Ax^*)^T[B(y^k-y^{k+1})]. \tag{25}$$

In addition, we notice that

$$\|B(y^k-y^*)\|^2=\|B(y^{k+1}-y^*)\|^2+\|B(y^k-y^{k+1})\|^2 \\ +2[B(y^{k+1}-y^*)]^T[B(y^k-y^{k+1})]. \tag{26}$$

It would be useful in the following discussion to observe the difference between

$$\|\lambda^k-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^k-y^*)\|^2,$$
$$\|\lambda^{k+1}-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^{k+1}-y^*)\|^2.$$

In fact, using $Ax^*+By^*=b$, it follows from (25) and (26) that

$$\begin{aligned}&\|\lambda^k-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^k-y^*)\|^2\\ &\geq\|\lambda^{k+1}-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^{k+1}-y^*)\|^2\\ &\quad+\gamma(2-\gamma)\beta_{k+1}^2\|Ax^{k+1}+By^{k+1}-b\|^2+\gamma\beta_{k+1}^2\|B(y^k-y^{k+1})\|^2\\ &\quad+2\gamma\beta_{k+1}^2(Ax^{k+1}+By^{k+1}-b)^TB(y^k-y^{k+1}).\end{aligned} \tag{27}$$

Now, we observe the last term in (27).

**Lemma 3.2.** For $k\geq 2$, we have

$$\begin{aligned}&\beta_{k+1}(Ax^{k+1}+By^{k+1}-b)^TB(y^k-y^{k+1})\\ &\geq(1-\gamma)\beta_k(Ax^k+By^k-b)^TB(y^k-y^{k+1}).\end{aligned} \tag{28}$$

**Proof.** By setting $y'=y^k$ in (11), we get

$$(y^k-y^{k+1})^T\{g(y^{k+1})-B^T[\lambda^k-\beta_{k+1}(Ax^{k+1}+By^{k+1}-b)]\}\geq 0. \tag{29}$$

Similarly, taking $k:=k-1$ and $y'=y^{k+1}$ in (11), we have

$$(y^{k+1}-y^k)^T\{g(y^k)-B^T[\lambda^{k-1}-\beta_k(Ax^k+By^k-b)]\}\geq 0. \tag{30}$$

By adding (29) and (30) and using the monotonicity of the operator $g$, we obtain

$$(y^{k+1}-y^k)^TB^T\{[\lambda^k-\beta_{k+1}(Ax^{k+1}+By^{k+1}-b)] \\ -[\lambda^{k-1}-\beta_k(Ax^k+By^k-b)]\}\geq 0. \tag{31}$$

Substituting

$$\lambda^k = \lambda^{k-1} - \gamma\beta_k(Ax^k + By^k - b)$$

in (31), the assertion of this lemma follows immediately. □

Recall that, as in Glowinski Ref. 3, we restrict $\gamma \in (0, (1+\sqrt{5})/2)$, in the method described. It is easy to verify that

$$1 + \gamma - \gamma^2 > 0, \qquad \forall \gamma \in (0, (1+\sqrt{5})/2).$$

Let

$$T = 2 - (1/2)(1 + \gamma - \gamma^2); \tag{32}$$

it follows that

$$T - \gamma = (1/2)(\gamma^2 - 3\gamma + 3) = (1/2)[(\gamma - 3/2)^2 + 3/4] \geq 3/8.$$

Now, we are in the stage to prove the main theorem of this paper.

**Theorem 3.1.** Let $w^* = (x^*, y^*, \lambda^*) \in \mathscr{W}^*$ be a solution point of MVI($\mathscr{W}$, $Q$), let $\gamma \in (0, (1+\sqrt{5})/2)$, and let $T$ be defined as in (32). Then, we have

$$\begin{aligned}
&\|\lambda^{k+1} - \lambda^*\|^2 + \gamma\beta_{k+1}^2\|B(y^{k+1} - y^*)\|^2 \\
&\quad + \gamma(T-\gamma)\beta_{k+1}^2\|Ax^{k+1} + By^{k+1} - b\|^2 \\
&\leq \|\lambda^k - \lambda^*\|^2 + \gamma\beta_{k+1}^2\|B(y^k - y^*)\|^2 \\
&\quad + \gamma(T-\gamma)\beta_k^2\|Ax^k + By^k - b\|^2 - (1/3)(1 + \gamma - \gamma^2)\gamma\beta_{k+1}^2 \\
&\quad \times \{\|Ax^{k+1} + By^{k+1} - b\|^2 + \|B(y^k - y^{k+1})\|^2\}.
\end{aligned} \tag{33}$$

**Proof.** It follows from (27) and Lemma 3.2 that

$$\begin{aligned}
&\|\lambda^k - \lambda^*\|^2 + \gamma\beta_{k+1}^2\|B(y^k - y^*)\|^2 \\
&\geq \|\lambda^{k+1} - \lambda^*\|^2 + \gamma\beta_{k+1}^2\|B(y^{k+1} - y^*)\|^2 \\
&\quad + \gamma(2-\gamma)\beta_{k+1}^2\|Ax^{k+1} + By^{k+1} - b\|^2 + \gamma\beta_{k+1}^2\|B(y^k - y^{k+1})\|^2 \\
&\quad + 2\beta_k\beta_{k+1}\gamma(1-\gamma)(Ax^k + By^k - b)^T(By^k - By^{k+1}).
\end{aligned} \tag{34}$$

Using the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
&2\beta_k\beta_{k+1}|\gamma(1-\gamma)(Ax^k + By^k - b)^T(By^k - By^{k+1})| \\
&\geq -\gamma(T-\gamma)\beta_k^2\|Ax^k + By^k - b\|^2 \\
&\quad - [(1-\gamma)^2/(T-\gamma)]\gamma\beta_{k+1}^2\|B(y^k - y^{k+1})\|^2.
\end{aligned} \tag{35}$$

Substituting (35) in (34), we derive

$$\begin{aligned}
&\|\lambda^k-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^k-y^*)\|^2+\gamma(T-\gamma)\beta_k^2\|Ax^k+By^k-b\|^2\\
&\geq\|\lambda^{k+1}-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^{k+1}-y^*)\|^2\\
&\quad+\gamma(T-\gamma)\beta_{k+1}^2\|Ax^{k+1}+By^{k+1}-b\|^2\\
&\quad+\gamma\beta_{k+1}^2\{(2-T)\|Ax^{k+1}+By^{k+1}-b\|^2\\
&\qquad\qquad+[1-(1-\gamma)^2/(T-\gamma)]\|B(y^k-y^{k+1})\|^2\}.
\end{aligned}\tag{36}$$

Using (32), it follows that

$$1-(1-\gamma)^2/(T-\gamma)=(1+\gamma-\gamma^2)/(3-3\gamma+\gamma^2).$$

For $\gamma\in(0,(1+\sqrt{5})/2)$, we have

$$3/4\leq\gamma^2-3\gamma+3=(\gamma-3/2)^2+3/4\leq 3,$$

and hence,

$$1-(1-\gamma)^2/(T-\gamma)\geq(1/3)(1+\gamma-\gamma^2).\tag{37}$$

Substituting

$$2-T=(1/2)(1+\gamma-\gamma^2)$$

and (37) in (36), we get

$$\begin{aligned}
&\|\lambda^k-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^k-y^*)\|^2+\gamma(T-\gamma)\beta_k^2\|Ax^k+By^k-b\|^2\\
&\geq\|\lambda^{k+1}-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^{k+1}-y^*)\|^2\\
&\quad+\gamma(T-\gamma)\beta_{k+1}^2\|Ax^{k+1}+By^{k+1}-b\|^2+(1/3)(1+\gamma-\gamma^2)\gamma\beta_{k+1}^2\\
&\quad\times\{\|Ax^{k+1}+By^{k+1}-b\|^2+\|B(y^k-y^{k+1})\|^2\}.
\end{aligned}\tag{38}$$

The conclusion of this theorem is proved. □

## 4. Convergence

Let $\{w^k\}=\{(x^k,y^k,\lambda^k)\}$ be the sequence generated by the alternating direction method in which the sequence $\{\beta_k\}$ satisfies condition (15). For convergence analysis, we demonstrate that we need only focus on showing that

$$\lim_{k\to\infty}\{\|Ax^{k+1}+By^{k+1}-b\|^2+\|B(y^k-y^{k+1})\|^2\}=0.$$

Note that, based on the equivalence of the solutions of the variational inequality and the projection equation, $x^{k+1}$ generated from (10) and $y^{k+1}$ generated from (11) satisfy

$$x^{k+1} = P_{\mathscr{X}}\{x^{k+1} - [f(x^{k+1}) - A^T(\lambda^k - \beta_{k+1}(Ax^{k+1} + By^k - b))]\}, \tag{39}$$

$$y^{k+1} = P_{\mathscr{Y}}\{y^{k+1} - [g(y^{k+1}) - B^T(\lambda^k - \beta_{k+1}(Ax^{k+1} + By^{k+1} - b))]\}, \tag{40}$$

respectively. Recall that

$$e(w^{k+1}) = \begin{bmatrix} e_x(w^{k+1}) \\ e_y(w^{k+1}) \\ e_\lambda(w^{k+1}) \end{bmatrix} = \begin{bmatrix} x^{k+1} - P_{\mathscr{X}}\{x^{k+1} - [f(x^{k+1}) - A^T\lambda^{k+1}]\} \\ y^{k+1} - P_{\mathscr{Y}}\{y^{k+1} - [g(y^{k+1}) - B^T\lambda^{k+1}]\} \\ Ax^{k+1} + By^{k+1} - b \end{bmatrix}, \tag{41}$$

and hence,

$$\|e(w^{k+1})\| \le \|e_x(w^{k+1})\| + \|e_y(w^{k+1})\| + \|e_\lambda(w^{k+1})\|. \tag{42}$$

Replacing the first $x^{k+1}$ in $e_x(w^{k+1})$ by the right-hand side of (39) and using

$$\|P_{\mathscr{X}}(v) - P_{\mathscr{X}}(v')\| \le \|v - v'\|$$

and (12), we have

$$\begin{aligned} &\|e_x(w^{k+1})\| \\ &\le \|A^T[\lambda^k - \lambda^{k+1} - \beta_{k+1}(Ax^{k+1} + By^k - b)]\| \\ &= \|A^T[\lambda^k - \lambda^{k+1} - \beta_{k+1}(Ax^{k+1} + By^{k+1} - b)] + \beta_{k+1}A^TB(y^k - y^{k+1})\| \\ &\le \|\beta_{k+1}A^T\|\{|\gamma - 1| \cdot \|Ax^{k+1} + By^{k+1} - b\| + \|B(y^k - y^{k+1})\|\}. \end{aligned} \tag{43}$$

Similarly, replacing the first $y^{k+1}$ in $e_y(w^{k+1})$ by the right-hand side of (40), and using the non-expansion of $P_{\mathscr{Y}}(\cdot)$ and (12), we obtain

$$\begin{aligned} \|e_y(w^{k+1})\| &\le \|B^T[\lambda^k - \lambda^{k+1} - \beta_{k+1}(Ax^{k+1} + By^k - b)]\| \\ &\le \|(\gamma - 1)\beta_{k+1}B^T\| \cdot \|Ax^{k+1} + By^{k+1} - b\|. \end{aligned} \tag{44}$$

Combining (43) and (44), we get

$$\begin{aligned} \|e(w^{k+1})\| \le &\{1 + |\gamma - 1|\beta_{k+1}(\|A^T\| + \|B^T\|)\}\|Ax^{k+1} + By^{k+1} - b\| \\ &+ \beta_{k+1}\|A^T\| \cdot \|B(y^k - y^{k+1})\|. \end{aligned} \tag{45}$$

Therefore, for bounded $\{\beta_k\}$, there is a constant $c_0 > 0$ such that

$$\|e(w^{k+1})\|^2 \le c_0\{\|Ax^{k+1} + By^{k+1} - b\|^2 + \|B(y^k - y^{k+1})\|^2\}. \tag{46}$$

Since solving MVI($\mathscr{W}, Q$) is equivalent to finding a zero point of $e(w)$, we need only prove that

$$\lim_{k\to\infty} \{\|Ax^{k+1} + By^{k+1} - b\|^2 + \|B(y^k - y^{k+1})\|^2\} = 0.$$

In fact, it is easy to check from (10)–(12) that, if

$$Ax^{k+1} + By^{k+1} - b = 0 \quad \text{and} \quad B(y^k - y^{k+1}) = 0,$$

then $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ is a solution of MVI($\mathscr{W}, Q$).

Based on Theorem 3.1, we study convergence under the following conditions.

**Condition C1.** $\inf\{\beta_k\}_1^\infty = B_L > 0$ and $\sum_{k=1}^\infty \eta_k^2 < +\infty$, where

$$\eta_k = \begin{cases} \sqrt{(\beta_{k+1}/\beta_k)^2 - 1}, & \text{if } k \in K_I, \\ 0, & \text{otherwise,} \end{cases} \tag{47a}$$

$$K_I := \{k \mid \beta_{k+1} > \beta_k\}. \tag{47b}$$

**Condition C2.** $\sum_{k=1}^\infty \theta_k^2 < +\infty$, where

$$\theta_k = \begin{cases} \sqrt{1 - (\beta_k/\beta_{k+1})^2}, & \text{if } k \in K_D, \\ 0, & \text{otherwise,} \end{cases} \tag{48a}$$

$$K_D := \{k \mid \beta_{k+1} < \beta_k\}. \tag{48b}$$

**Theorem 4.1.** Let $\{w^k\} = \{(x^k, y^k, \lambda^k)\}$ be the sequence generated by the proposed method, and let $w^* = (x^*, y^*, \lambda^*) \in \mathscr{W}^*$ be a solution. If the sequences $\{\beta_k\}$ satisfies Condition C1 or Condition C2, then the method is convergent.

**Proof.** First, we consider the case that the sequence $\{\beta_k\}$ satisfies Condition C1. It follows from (47) that $\sum_{k=0}^\infty \eta_k^2 < \infty$ and the product $\prod_{k=0}^\infty (1 + \eta_k^2)$ is bounded. Denote

$$C_s := \sum_{i=1}^\infty \eta_i^2, \qquad C_p := \prod_{i=1}^\infty (1 + \eta_i^2). \tag{49}$$

Note that the result of Theorem 3.1 [see (33)] can be rewritten as

$$\begin{aligned} &\|\lambda^{k+1} - \lambda^*\|^2 + \gamma\beta_{k+1}^2\|B(y^{k+1} - y^*)\|^2 + \gamma(T-\gamma)\beta_{k+1}^2\|Ax^{k+1} + By^{k+1} - b\|^2 \\ &\le \|\lambda^k - \lambda^*\|^2 + (\beta_{k+1}^2/\beta_k^2)\gamma\beta_k^2\|B(y^k - y^*)\|^2 + \gamma(T-\gamma)\beta_k^2\|Ax^k + By^k - b\|^2 \\ &\quad - (1/3)(1 + \gamma - \gamma^2)\gamma\beta_{k+1}^2\{\|Ax^{k+1} + By^{k+1} - b\|^2 + \|B(y^k - y^{k+1})\|^2\}. \end{aligned} \tag{50}$$

Under the condition (47), we have

$$\beta_{k+1}^2/\beta_k^2 \leq 1+\eta_k^2.$$

It follows that

$$\begin{aligned}
&\|\lambda^{k+1}-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^{k+1}-y^*)\|^2+\gamma(T-\gamma)\beta_{k+1}^2\|Ax^{k+1}+By^{k+1}-b\|^2\\
&\leq(1+\eta_k^2)\{\|\lambda^k-\lambda^*\|^2+\gamma\beta_k^2\|B(y^k-y^*)\|^2+\gamma(T-\gamma)\beta_k^2\|Ax^k+By^k-b\|^2\}\\
&\quad-(1/3)(1+\gamma-\gamma^2)\gamma\beta_{k+1}^2\{\|Ax^{k+1}+By^{k+1}-b\|^2+\|B(y^k-y^{k+1})\|^2\},
\end{aligned}\tag{51}$$

and hence,

$$\begin{aligned}
&\|\lambda^{k+1}-\lambda^*\|^2+\gamma\beta_{k+1}^2\|B(y^{k+1}-y^*)\|^2+\gamma(T-\gamma)\beta_{k+1}^2\|Ax^{k+1}+By^{k+1}-b\|^2\\
&\leq\prod_{i=1}^{k}(1+\eta_i^2)\{\|\lambda^1-\lambda^*\|^2+\gamma\beta_1^2\|B(y^1-y^*)\|^2+\gamma(T-\gamma)\beta_1^2\|Ax^1+By^1-b\|^2\}\\
&\leq C_p\{\|\lambda^1-\lambda^*\|^2+\gamma\beta_1^2\|B(y^1-y^*)\|^2+\gamma(T-\gamma)\beta_1^2\|Ax^1+By^1-b\|^2\}.
\end{aligned}\tag{52}$$

Therefore, there exists a constant $C>0$, such that

$$\|\lambda^k-\lambda^*\|^2+\gamma\beta_k^2\|B(y^k-y^*)\|^2+\gamma(T-\gamma)\beta_k^2\|Ax^k+By^k-b\|^2\leq C,\qquad\forall k.\tag{53}$$

From (51) amd (53), we get

$$\begin{aligned}
&\sum_{i=1}^{\infty}(1/3)(1+\gamma-\gamma^2)\gamma\beta_{i+1}^2\{\|Ax^{i+1}+By^{i+1}-b\|^2+\|B(y^i-y^{i+1})\|^2\}\\
&\leq\|\lambda^1-\lambda^*\|^2+\gamma\beta_1^2\|B(y^1-y^*)\|^2+\gamma(T-\gamma)\beta_1^2\|Ax^1+By^1-b\|^2\\
&\quad+\sum_{i=1}^{\infty}\eta_i^2\{\|\lambda^i-\lambda^*\|^2+\gamma\beta_i^2\|B(y^i-y^*)\|^2+\gamma(T-\gamma)\beta_i^2\|Ax^i+By^i-b\|^2\}\\
&\leq(1+C_s)C.
\end{aligned}\tag{54}$$

Since $\inf\{\beta_k\}_1^\infty\geq B_L>0$, it follows from (54) that

$$\lim_{k\to\infty}\{\|Ax^{k+1}+By^{k+1}-b\|^2+\|B(y^k-y^{k+1})\|^2\}=0.$$

Now, we turn to study the case that $\{\beta_k\}$ satisfies Condition C2. For this purpose, we rewrite the result of Theorem 3.1 [see (33)] as

$$\begin{aligned}
&(1/\beta_{k+1}^2)\|\lambda^{k+1}-\lambda^*\|^2+\gamma\|B(y^{k+1}-y^*)\|^2\\
&\quad+\gamma(T-\gamma)\|Ax^{k+1}+By^{k+1}-b\|^2\\
&\leq(\beta_k^2/\beta_{k+1}^2)(1/\beta_k^2)\|\lambda^k-\lambda^*\|^2+\gamma\|B(y^k-y^*)\|^2\\
&\quad+(\beta_k^2/\beta_{k+1}^2)\gamma(T-\gamma)\|Ax^k+By^k-b\|^2-(1/3)(1+\gamma-\gamma^2)\gamma\\
&\quad\times\{\|Ax^{k+1}+By^{k+1}-b\|^2+\|B(y^k-y^{k+1})\|^2\}.
\end{aligned}\tag{55}$$

Note that, under condition (48), we have

$$\beta_k^2/\beta_{k+1}^2 \leq (1+\theta_k^2).$$

We define

$$D_s := \sum_{i=0}^{\infty} \theta_i^2, \qquad D_p := \prod_{i=0}^{\infty} (1+\theta_i^2).$$

It follows from (55) and (48) that

$$\begin{aligned}
&(1/\beta_{k+1}^2)\|\lambda^{k+1}-\lambda^*\|^2 + \gamma\|B(y^{k+1}-y^*)\|^2 \\
&\quad + \gamma(T-\gamma)\|Ax^{k+1}+By^{k+1}-b\|^2 \\
&\leq \left[\prod_{i=1}^{k}(1+\theta_i^2)\right]\{(1/\beta_1^2)\|\lambda^1-\lambda^*\|^2 + \gamma\|B(y^1-y^*)\|^2 \\
&\qquad\qquad + \gamma(T-\gamma)\|Ax^1+By^1-b\|^2\} \\
&\leq D_p\{(1/\beta_1^2)\|\lambda^1-\lambda^*\|^2 + \gamma\|B(y^1-y^*)\|^2 \\
&\qquad + \gamma(T-\gamma)\|Ax^1+By^1-b\|^2\};
\end{aligned} \tag{56}$$

hence, there exists a constant $D>0$, such that

$$\begin{aligned}
&(1/\beta_k^2)\|\lambda^k-\lambda^*\|^2 + \gamma\|B(y^k-y^*)\|^2 \\
&+ \gamma(T-\gamma)\|Ax^k+By^k-b\|^2 \leq D, \qquad \forall k.
\end{aligned} \tag{57}$$

From (55) and (57), we get

$$\begin{aligned}
&\sum_{i=1}^{\infty}(1/3)(1+\gamma-\gamma^2)\gamma\{\|Ax^{i+1}+By^{i+1}-b\|^2 + \|B(y^i-y^{i+1})\|^2\} \\
&\leq (1/\beta_1^2)\|\lambda^1-\lambda^*\|^2 + \gamma\|B(y^1-y^*)\|^2 + \gamma(T-\gamma)\|Ax^1+By^1-b\|^2 \\
&\quad + \sum_{i=1}^{\infty}\theta_i^2\{(1/\beta_i^2)\|\lambda^i-\lambda^*\|^2 + \gamma_i^2\|B(y^i-y^*)\|^2 + \gamma(T-\gamma)\|Ax^i+By^i-b\|^2\} \\
&\leq (1+D_s)D,
\end{aligned} \tag{58}$$

and therefore,

$$\lim_{k\to\infty}\{\|Ax^{k+1}+By^{k+1}-b\|^2 + \|B(y^k-y^{k+1})\|^2\} = 0.$$

The theorem is proved. □

**Remark 4.1.** It is clear that, if the sequence $\{\beta_k\}$ is monotone decreasing and bounded below away from zero, then $\eta_k \equiv 0$ and $\{\beta_k\}$ satisfies Condition C1. Conversely, if the positive sequence $\{\beta_k\}$ is monotone increasing, then $\theta_k \equiv 0$ and $\{\beta_k\}$ satisfies condition C2. Hence, the proposed method using either parameter adjusting Strategy S1 or S2 in Section 2 is convergent.

**Remark 4.2.** The sequence $\{\beta_k\}$ generated by adjusting Strategy S3 in Section 2 satisfies both Conditions C1 and C2. Note that, for

$$k \in K_I = \{k \mid \beta_{k+1} > \beta_k\},$$
$$k \in K_D = \{k \mid \beta_{k+1} < \beta_k\},$$

we have

$$\eta_k^2 = 2\tau_k + \tau_k^2,$$
$$\theta_k^2 = 2\tau_k + \tau_k^2,$$

respectively. It follows from (14) that

$$\sum_{k=1}^{\infty} \eta_k^2 < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \theta_k^2 < \infty.$$

Thus, the sequence $\{\beta_k\}$ satisfies both Conditions C1 and C2.

## 5. Numerical Experiments

In this section, we present the results of some numerical experiments. Our main interest is in showing the efficiency of the alternating direction method with self adaptive penalty parameters.

For this purpose, we consider the following problem:

$$\min\{c^T x \mid x \in \Omega_1 \cap \Omega_2\}, \tag{59}$$

where

$$\Omega_1 = \{x \mid \|x\| \le r_1, x \in R^n\},$$
$$\Omega_2 = \{x \mid \|x - b\| \le r_2, x \in R^n\}.$$

We test the problems with dimension $n = 1000$. In order to guarantee the feasibility of the problem, we should have

$$\|b\| \le r_1 + r_2.$$

The test problems of the form (59) are generated randomly. First, the components of $b$ are uniformally distributed in (0, 10). We take

$$r_1 = 0.5\|b\| \quad \text{and} \quad r_2 = 0.6\|b\|,$$

Table 1. Description of various methods.

| | | |
|---|---|---|
| Method F | $\beta_k \equiv \beta_1$. | |
| Method 1, Strategy S1 | $\beta_{k+1} = \begin{cases} \beta_k * 2, & \text{if } \|e_x(w^k)\| < 0.1 * \|e_\lambda(w^k)\| \text{ and } k \le k_{\max}, \\ \beta_k, & \text{otherwise.} \end{cases}$ | |
| Method 2, Strategy S2 | $\beta_{k+1} = \begin{cases} \beta_k * 0.5, & \text{if } 0.1 * \|e_x(w^k)\| > \|e_\lambda(w^k)\| \text{ and } k \le k_{\max}, \\ \beta_k, & \text{otherwise.} \end{cases}$ | |
| Method 3, Strategy S3 | $\beta_{k+1} = \begin{cases} \beta_k * 2, & \text{if } \|e_x(w^k)\| < 0.1 * \|e_\lambda(w^k)\| \text{ and } k \le k_{\max}, \\ \beta_k * 0.5, & \text{if } 0.1 * \|e_x(w^k)\| > \|e_\lambda(w^k)\| \text{ and } k \le k_{\max}, \\ \beta_k, & \text{otherwise.} \end{cases}$ | |

and thus the problem is well defined. The components of $c$ are uniformly distributed in $(-K, K)$. To cast (59) in a format suitable for the alternating direction method, we introduce an auxiliary vector $y$ and rewrite the problem as follows:

$$\min \quad c^T x, \tag{60a}$$

$$\text{s.t.} \quad x + y - b = 0, \tag{60b}$$

$$x \in \odot_{r1}, \qquad y \in \odot_{r2}, \tag{60c}$$

where $\odot_r$ denotes a ball centered on zero point with radius $r$. For this problem, by using the alternating direction method, the subVIs (10) and (11) are equivalent to finding $x^{k+1}$ and $y^{k+1}$ such that

$$x^{k+1} = P_{\odot_{r_1}}\{(1-\beta_{k+1})x^{k+1} + \lambda^k + \beta_{k+1}(b - y^k) - c\},$$

$$y^{k+1} = P_{\odot_{r_2}}\{(1-\beta_{k+1})y^{k+1} + \lambda^k + \beta_{k+1}(b - x^{k+1})\},$$

respectively. Note that the solution of problem as

$$v = P_{\odot_r}[(1-\beta)v + q]$$

can be given explicitly by

$$v = \begin{cases} q/\beta, & \text{if } \|q\| \le \beta r, \\ rq/\|q\|, & \text{otherwise.} \end{cases}$$

Hence, for this test problem, we can solve the subVIs (10) and (11) in each iteration exactly.

We test the problems by using the original alternating direction method and the methods using different penalty parameter adjusting strategies. For simplicity, we take the relaxation factor $\gamma \equiv 1$ and use the methods listed in Table 1.

Method F with fixed parameter is the original alternating direction method as described in Refs. 2–5. Methods 1, 2, 3 belong to the modified strategies of this paper, in which we use the penalty parameter adjusting Strategies S1, S2, S3 as described in Section 2. Here, we take $\mu = 0.1$, $k_{\max} = 50$, and

$$\tau_k = \begin{cases} 1, & \text{if } k \le k_{\max}, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, the sequence $\{\tau_k\}$ satisfies (14). As starting point, we take

$$y^0 = (0, \dots, 0)^T \quad \text{and} \quad \lambda^0 = (0, \dots, 0)^T.$$

The iteration was stopped as soon as

$$\|e(w^k)\|_\infty \le 10^{-8}.$$

All codes are written in Matlab and run on a P-II 400 Personal Computer. Clearly, the CPU time used for solving a problem cannot be too precise, but is still a useful quantity for comparison purposes.

**5.1. First Set of Tests.** In our first set of tests, problem (60) has fixed $K = 10$; i.e., the components of $c \in R^{1000}$ are uniformly distributed in $(-10, 10)$. We tested the different methods with different starting penalty parameters. The results are given in Table 2.

Table 2. Test results for the same problem with different starting penalties.

| | Method F | | Method 1 | | Method 2 | | Method 3 | |
|---|---|---|---|---|---|---|---|---|
| $\beta_1$ | $N_{it}$ | $T$[sec] | $N_{it}$ | $T$[sec] | $N_{it}$ | $T$[sec] | $N_{it}$ | $T$[sec] |
| $10^{-5}$ | — | — | 51 | 0.33 | — | — | 49 | 0.27 |
| $10^{-4}$ | — | — | 47 | 0.28 | — | — | 46 | 0.27 |
| $10^{-3}$ | — | — | 44 | 0.22 | — | — | 43 | 0.22 |
| $10^{-2}$ | — | — | 39 | 0.22 | — | — | 42 | 0.27 |
| $10^{-1}$ | 1136 | 5.82 | 36 | 0.20 | 1106 | 6.48 | 38 | 0.22 |
| 1 | 122 | 0.61 | 34 | 0.22 | 118 | 0.72 | 35 | 0.22 |
| 10 | 36 | 0.22 | 40 | 0.28 | 50 | 0.28 | 35 | 0.16 |
| $10^2$ | 69 | 0.33 | 65 | 0.33 | 39 | 0.27 | 39 | 0.22 |
| $10^3$ | 287 | 1.38 | 285 | 1.59 | 40 | 0.22 | 42 | 0.28 |
| $10^4$ | 2525 | 11.81 | 2376 | 12.58 | 43 | 0.28 | 45 | 0.28 |
| $10^5$ | — | — | — | — | 46 | 0.28 | 46 | 0.27 |
| $10^6$ | — | — | — | — | 50 | 0.23 | 52 | 0.32 |
| $10^7$ | — | — | — | — | 57 | 0.27 | 57 | 0.33 |
| $10^8$ | — | — | — | — | 56 | 0.33 | 56 | 0.33 |

—means that the number of iteration is greater than 10,000 and the CPU time is in excess of 100 sec.

In Method F, the penalty parameter $\beta_k$ is fixed. Indeed, the number of iteration is dependent significantly on the parameter $\beta$. For this problem, it seems that $\beta \approx 10$ is the best choice for the method. In general, we do not have a priori knowledge on how to choose a fixed $\beta$ for a particular problem.

If we use Method 1, the penalty sequence $\{\beta_k\}$ is monotonically increasing. For this method, the starting penalty parameter $\beta_1$ should be relatively small. Even if the starting penalty parameter is too small, using Strategy S1, the method will find a proper parameter in a few iterations. As demonstrated in Table 2 for this test problem, a large $\beta_1$ (such as $\beta_1 > 1000$) may lead to a slow convergence or cause convergence to fail.

In Method 2, the penalty sequence $\{\beta_k\}$ is monotonically decreasing. By using this method, the starting penalty parameter $\beta_1$ should be relatively large. The method will reduce the penalty parameter automatically when the starting parameter is too large. A small $\beta_1$ (such as $\beta_1 < 0.01$) may cause convergence to fail.

Method 3 belongs to the proposed method with Strategy S3. Compared with the Methods 1 and 2, we find out that Method 3 is most effective and flexible. The method converges quickly and the number of iterations is almost independent of the starting penalty parameter $\beta_1$.

**5.2. Second Set of Tests.** Note that Problem (60) is invariant under multiplication of $c$ by a positive factor $K$. In our second set of tests, first we let the components of $c \in R^{1000}$ be uniformly distributed in $(-1, 1)$; then, we multiplied $c$ by a positive factor $K$. We take $\beta_1 = 1$ as the starting penalty parameter in all test methods, and let $K$ range from 0.00001 to 10,000. The test results are depicted in Table 3.

Table 3. Test results for different problems with the same starting penalty.

| | Method F | | Method 1 | | Method 2 | | Method 3 | |
|---|---|---|---|---|---|---|---|---|
| $K$ | $N_{it}$ | $T$[sec] | $N_{it}$ | $T$[sec] | $N_{it}$ | $T$[sec] | $N_{it}$ | $T$[sec] |
| $10^{-5}$ | — | — | — | — | 47 | 0.25 | 51 | 0.27 |
| $10^{-4}$ | 25701 | 117.49 | 24800 | 130.78 | 46 | 0.28 | 49 | 0.27 |
| $10^{-3}$ | 2506 | 11.42 | 2515 | 13.23 | 42 | 0.22 | 51 | 0.28 |
| $10^{-2}$ | 90 | 1.26 | 278 | 1.49 | 40 | 0.22 | 53 | 0.33 |
| $10^{-1}$ | 63 | 0.33 | 64 | 0.33 | 36 | 0.21 | 48 | 0.28 |
| 1 | 35 | 0.16 | 41 | 0.27 | 49 | 0.27 | 45 | 0.27 |
| 10 | 121 | 0.61 | 37 | 0.22 | 118 | 0.60 | 38 | 0.22 |
| $10^{2}$ | 1136 | 5.65 | 38 | 0.22 | 1110 | 0.31 | 37 | 0.17 |
| $10^{3}$ | 11164 | 55.87 | 52 | 0.33 | 11475 | 65.64 | 51 | 0.33 |
| $10^{4}$ | — | — | 43 | 0.27 | — | — | 44 | 0.22 |

—means that the number of iteration is greater than 100,000 and the CPU time is in excess of 500 sec.

For the same problem, the number of iterations of Method F is dependent significantly on the individual magnitude of $K$. For $K < 0.01$, the starting penalty parameter $\beta_1 = 1$ in Method 1 is too large. Conversely, for $K > 10$, $\beta_1 = 1$ in Method 2 is too small. Again, using Method 3, the number of iterations is almost independent of the range of $K$.

## 6. Conclusions

In this paper, we have proposed a modified alternating direction method for solving separate monotone variational inequalities. The method presented extends the original one by allowing the penalty parameter for the system of constrained equations to be varied in each iteration. We have suggested strategies for adjusting the penalty parameter and convergence results have been established. The preliminary numerical tests show that the proposed method is effective in practice.

## References


1. NAGURNEY, A., and RAMANUJAM, P., *Transportation Network Policy Modeling with Goal Targets and Generalized Penalty Functions*, Transportation Science, Vol. 30, pp. 3–13, 1996.
2. FORTIN, M., and GLOWINSKI, R., Editors, *Augmented Lagrangian Methods: Applications to the Solution of Boundary-Valued Problems*, North-Holland, Amsterdam, Holland, 1983.
3. GLOWINSKI, R., *Numerical Methods for Nonlinear Variational Problems*, Springer Verlag, New York, NY, 1984.
4. GLOWINSKI, R., and LE TALLEC, P., *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM Studies in Applied Mathematics, Philadelphia, Pennsylvania, 1989.
5. GABAY, D., *Applications of the Method of Multipliers to Variational Inequalities*, Augmented Lagrange Methods: Applications to the Solution of Boundary-Valued Problems, Edited by M. Fortin and R. Glowinski, North Holland, Amsterdam, Holland, pp. 299–331, 1983.
6. GABAY, D., and MERCIER, B., *A Dual Algorithm for the Solution of Nonlinear Variational Problems via Finite-Element Approximations*, Computer and Mathematics with Applications, Vol. 2, pp. 17–40, 1976.
7. LIONS, P. L., and MERCIER, B., *Splitting Algorithms for the Sum of Two Nonlinear Operators*, SIAM Journal on Numerical Analysis, Vol. 16, pp. 964–979, 1979.
8. TSENG, P., *Applications of Splitting Algorithm to Decomposition in Convex Programming and Variational Inequalities*, SIAM Journal on Control and Optimization, Vol. 29, pp. 119–138, 1991.

9. ECKSTEIN, J., and FUKUSHIMA, M., *Some Reformulation and Applications of the Alternating Direction Method of Multipliers*, Large Scale Optimization: State of the Art, Edited by W. W. Hager et al., Kluwer Academic Publishers, Amsterdam, Holland, pp. 115–134, 1994.
10. FUKUSHIMA, M., *Application of the Alternating Direction Method of Multipliers to Separable Convex Programming Problems*, Computational Optimization and Applications, Vol. 2, pp. 93–111, 1992.
11. KONTOGIORGIS, S., and MEYER, R. R., *A Variable-Penalty Alternating Directions Method for Convex Optimization*, Mathematical Programming, Vol. 83, pp. 29–53, 1998.
12. HE, B. S., and YANG, H., *Some Convergence Properties of a Method of Multipliers for Linearly Constrained Monotone Variational Inequalities*, Operations Research Letters, Vol. 23, pp. 151–161, 1998.
13. HARKER, P. T., and PANG, J. S., *Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms, and Applications*, Mathematical Programming, Vol. 48, pp. 161–220, 1990.
14. KINDERLEHRER, D., and STAMPACCHIA, G., *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York, NY, 1980.