J Optim Theory Appl (2007) 134: 107–117
DOI 10.1007/s10957-007-9192-2

Proximal Alternating Directions Method for Structured Variational Inequalities

M.H. Xu

Published online: 1 May 2007

Abstract In the alternating directions method, the relaxation factor $\gamma \in (0, \frac{\sqrt{5}+1}{2})$ by Glowinski is useful in practical computations for structured variational inequalities. This paper points out that the same restriction region of the relaxation factor is also valid in the proximal alternating directions method.

Keywords Structured variational inequalities · Proximal point algorithm · Alternating directions method

1 Introduction

The problem treated in this paper is the following variational inequality:

$$(x' - x)^T f(x) \geq 0, \quad \forall u' \in \Omega, \tag{1a}$$

$$(y' - y)^T g(y) \geq 0, \quad \forall u' \in \Omega, \tag{1b}$$

where

$$u = (x, y) \in \Omega, \quad \Omega = \{(x, y) | x \in \mathcal{X}, y \in \mathcal{Y}, Ax + By = b\}, \tag{2}$$

Communicated by X.Q. Yang.

The research was supported by the NSFC of China Grant 10571083 and MOEC Grant 20060284001. The author thanks the anonymous referees for valuable suggestions.

M.H. Xu (✉)
Department of Information Science, Jiangsu Polytechnic University, Changzhou, Jiangsu, People's Republic of China
e-mail: xuminghua@jpu.edu.cn

M.H. Xu
Department of Mathematics, Nanjing University, Nanjing, China

$\mathcal{X}$ and $\mathcal{Y}$ are given nonempty closed convex subsets of $\mathcal{R}^n$ and $\mathcal{R}^m$, $A \in \mathcal{R}^{l\times n}$, $B \in \mathcal{R}^{l\times m}$ are given matrices, $b \in \mathcal{R}^l$ is a given vector, and $f : \mathcal{X} \to \mathcal{R}^n$, $g : \mathcal{Y} \to \mathcal{R}^m$ are given monotone operators. Studies and applications of such problems can be found in Glowinski [1], Glowinski and Le Tallec [2], Eckstein and Fukushima [3–5] and He and Yang [6]. By attaching a Lagrange multiplier vector $\lambda \in \mathcal{R}^l$ to the linear constraints $Ax + By = b$, the problem (1–2) can be rewritten in the following form:

$$(x' - x)^T \{f(x) - A^T\lambda\} \geq 0, \quad \forall w' \in \mathcal{W}, \tag{3a}$$

$$(y' - y)^T \{g(y) - B^T\lambda\} \geq 0, \quad \forall w' \in \mathcal{W}, \tag{3b}$$

$$Ax + By - b = 0, \tag{3c}$$

where

$$w = (x, y, \lambda) \in \mathcal{W}, \quad \mathcal{W} = \mathcal{X} \times \mathcal{Y} \times \mathcal{R}^l. \tag{4}$$

Problem (3–4) is referred as SVI (structured variational inequality).

For solving SVI problems, from a given $v^k = (y^k, \lambda^k) \in \mathcal{Y} \times \mathcal{R}^l$, the alternating directions method (ADM, [1, 2] and [7–11]) produces the new iterate $v^{k+1} = (y^{k+1}, \lambda^{k+1}) \in \mathcal{Y} \times \mathcal{R}^l$ via the following procedure:

First, x^{k+1} is obtained by solving the following problem:

$$(x' - x)^T \{f(x) - A^T[\lambda^k - \beta(Ax + By^k - b)]\} \geq 0, \quad \forall x' \in \mathcal{X}, \tag{5}$$

where $x \in \mathcal{X}$. Then, y^{k+1} is produced by solving

$$(y' - y)^T \{g(y) - B^T[\lambda^k - \beta(Ax^{k+1} + By - b)]\} \geq 0, \quad \forall y' \in \mathcal{Y}, \tag{6}$$

where $y \in \mathcal{Y}$. Finally, the multipliers are updated via

$$\lambda^{k+1} = \lambda^k - \gamma\beta(Ax^{k+1} + By^{k+1} - b), \tag{7}$$

where $\beta > 0$ is a given penalty parameter for the linearly constrained equation $Ax + By - b = 0$ and $\gamma \in (0, \frac{\sqrt{5}+1}{2})$ is a relaxation factor. This method is referred to as the alternating directions method (ADM). Most ADM papers use $\gamma = 1$. To the best of our knowledge, the restriction region $(0, \frac{\sqrt{5}+1}{2})$ was suggested first by Glowinski [1] and the numerical experiments in [12] show that γ should be larger than 1 for fast convergence.

In order to improve the condition of the subproblem of the alternating directions method, some proximal alternating directions methods [13, 14] were proposed. The classical proximal alternating directions method is one of the attractive ADMs. For a given triplet $w^k = (x^k, y^k, \lambda^k) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{R}^l$, the classical proximal alternating directions method produces the new iterate $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1}) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{R}^l$ by the following procedure:

First, x^{k+1} is obtained by solving the following problem:

$$(x' - x)^T \{f(x) - A^T[\lambda^k - \beta(Ax + By^k - b)] + r(x - x^k)\} \geq 0, \quad \forall x' \in \mathcal{X}, \tag{8}$$

where $x \in \mathcal{X}$. Then, y^{k+1} is produced by solving

$$(y' - y)^T\{g(y) - B^T[\lambda^k - \beta(Ax^{k+1} + By - b)] \\ + s(y - y^k)\} \geq 0, \quad \forall y' \in \mathcal{Y}, \tag{9}$$

where $y \in \mathcal{Y}$. Finally, the multiplier is updated via

$$\lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b), \tag{10}$$

where $r > 0, s > 0$ are given proximal parameters and $\beta > 0$ is a given penalty parameter for the linearly constrained equation $Ax + By - b = 0$. This method is referred to as the QP alternating directions method because the proximal term $r(x - x^k)$ in (8) (resp. $s(y - y^k)$ in (9)) is the derivative of a quadratic function.

Since numerical experiments show that the adaptive relaxation factor γ can speed up the convergence of the alternating directions method, it is natural to introduce a relaxation factor γ in the QP alternating directions method and replace (10) by

$$\lambda^{k+1} = \lambda^k - \gamma\beta(Ax^{k+1} + By^{k+1} - b). \tag{11}$$

In this way, we get a more general QP alternating directions method. For convenience, we call the resulting method the QP alternating directions method (QPADM). In this paper, we prove that the restriction region of the relaxation factor γ in the alternating directions method is also valid in QPADM. Similar to the alternating directions method, it is possible to speed up the convergence of QPADM by choosing an adaptive relaxation factor γ.

The paper is organized as follows. In the next section, we present the QP alternating directions method. Section 3 shows some contractive properties. In Sect. 4, we prove the convergence of the proposed method. Finally, we give some conclusive remarks. In the rest of this paper, we let $H \in \mathcal{R}^{l\times l}$, $R \in \mathcal{R}^{n\times n}$ and $S \in \mathcal{R}^{m\times m}$ be positive definite and let $R = \mathrm{diag}(r_1, r_2, \ldots, r_n)$ and $S = \mathrm{diag}(s_1, s_2, \ldots, s_m)$.

2 Proposed Method

For convenience, we make the standard assumptions to guarantee that the problem under consideration is solvable and that the QPADM is well defined.

Assumption A. $f(x)$ is monotone with respect to $\mathcal{X}$ and $g(y)$ is monotone with respect to $\mathcal{Y}$.

Assumption B. The solution set of SVI, denoted by $\mathcal{W}^*$, is nonempty.

QP Alternating Directions Method (QPADM)

Step 0. Given $\varepsilon > 0$, $\gamma \in (0, \frac{1+\sqrt{5}}{2})$, $w^0 = (x^0, y^0, \lambda^0) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{R}^l$. Set $k = 0$.

Step 1. Find $x^{k+1} \in \mathcal{X}$ (with fixed x^k, y^k and λ^k) which is a solution of the following variational inequality:

$$(x' - x)^T\{f_k(x) + R(x - x^k)\} \geq 0, \quad x' \in \mathcal{X}, \tag{12}$$

where

$$f_k(x) = f(x) - A^T[\lambda^k - H(Ax + By^k - b)], \quad x \in \mathcal{X}. \tag{13}$$

Step 2. Find $y^{k+1} \in \mathcal{Y}$ (with fixed x^{k+1}, y^k and λ^k) which is a solution of the following variational inequality:

$$(y' - y)^T\{g_k(y) + S(y - y^k)\} \geq 0, \quad y' \in \mathcal{Y}, \tag{14}$$

where

$$g_k(y) = g(y) - B^T[\lambda^k - H(Ax^{k+1} + By - b)], \quad y \in \mathcal{Y}. \tag{15}$$

Step 3. Set

$$\lambda^{k+1} = \lambda^k - \gamma H(Ax^{k+1} + By^{k+1} - b). \tag{16}$$

Step 4. Verify the convergence. If $\|w^k - w^{k+1}\|_\infty < \varepsilon$, then stop; else, set $k := k + 1$ and go to Step 1.

Here, we use the relaxation factor $\gamma \in (0, \frac{1+\sqrt{5}}{2})$ in (16). To the best of our knowledge, this factor was first used by Glowinski [1] for solving variational inequality arising from partial differential equations.

Lemma 2.1 *Let $q(z) \in \mathcal{R}^n$ be a monotone mapping of z with respect to a nonempty closed convex set $\mathcal{Z}$ and let $Q \in \mathcal{R}^{n\times n}$ be a symmetric positive-definite matrix. For given $z^k \in \mathcal{Z}$, the variational inequality*

$$(z' - z)^T\{q(z) + Q(z - z^k)\} \geq 0, \quad \forall z' \in \mathcal{Z}, \tag{17}$$

where $z \in \mathcal{Z}$, has a unique solution z. Moreover, for $z \in \mathcal{Z}$ and any $z' \in \mathcal{Z}$, we have

$$(z' - z)^T q(z) \geq \frac{1}{2}\left(\|z - z'\|_Q^2 - \|z^k - z'\|_Q^2\right) + \frac{1}{2}\|z^k - z\|_Q^2. \tag{18}$$

Proof The first assertion is due to the fact that $q(z) + Q(z - z^k)$ is a strictly monotone mapping of z with respect to $\mathcal{Z}$. Note that the following is an identity:

$$(z - z')^T Q(z - z^k) = \frac{1}{2}\left(\|z - z'\|_Q^2 - \|z^k - z'\|_Q^2\right) + \frac{1}{2}\|z^k - z\|_Q^2. \tag{19}$$

It follows from (19) and (17) that

$$(z' - z)^T q(z) \geq \frac{1}{2}\left(\|z - z'\|_Q^2 - \|z^k - z'\|_Q^2\right) + \frac{1}{2}\|z^k - z\|_Q^2. \tag{20}$$

Hence, (18) holds and the proof is completed. □

3 Some Contractive Properties

The purpose of this section is to show that, for any $w^* = (x^*, y^*, \lambda^*) \in \mathcal{W}^*$, the sequence $\{w^k\}$ generated by the proposed method satisfies

$$\begin{aligned}&\big(\|w^{k+1} - w^*\|_M^2 + \|y^{k+1} - y^k\|_S^2\big)\\&\quad\le \big(\|w^k - w^*\|_M^2 + \|y^k - y^{k-1}\|_S^2\big) - \|w^k - w^{k+1}\|_N^2,\end{aligned}$$

where M and N are positive definite matrices. The proof is based on the following three lemmas. The first lemma is due to applying Lemma 2.1 to the QP systems in Step 1 and Step 2 of the proposed method.

Lemma 3.1 *Let x^{k+1} and y^{k+1} be generated by (12–15) from a given $w^k = (x^k, y^k, \lambda^k)$. Then, for any $w^* = (x^*, y^*, \lambda^*) \in \mathcal{W}^*$, we have*

$$\begin{aligned}&(\lambda^k - \lambda^*)^T(Ax^{k+1} + By^{k+1} - b)\\&\quad\ge \frac{1}{2}\big(\|x^{k+1} - x^*\|_R^2 - \|x^k - x^*\|_R^2\big) + \frac{1}{2}\|x^k - x^{k+1}\|_R^2\\&\qquad+ \frac{1}{2}\big(\|y^{k+1} - y^*\|_S^2 - \|y^k - y^*\|_S^2\big) + \tfrac{1}{2}\|y^k - y^{k+1}\|_S^2\\&\qquad+ \frac{1}{2}\big(\|B(y^{k+1} - y^*)\|_H^2 - \|B(y^k - y^*)\|_H^2\big)\\&\qquad+ \frac{1}{2}\|B(y^k - y^{k+1})\|_H^2 + \|Ax^{k+1} + By^{k+1} - b\|_H^2\\&\qquad+ (Ax^{k+1} + By^{k+1} - b)^T H(By^k - By^{k+1}).\end{aligned}\tag{21}$$

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

$$(x^{k+1} - x^*)^T(f(x^*) - A^T\lambda^*) \ge 0,\tag{22}$$
$$(y^{k+1} - y^*)^T(g(y^*) - B^T\lambda^*) \ge 0,\tag{23}$$

and

$$Ax^* + By^* - b = 0.$$

On the other hand, applying Lemma 2.1 to (12–13) (namely, by setting $z^k = x^k$, $z = x^{k+1}$, $q = f_k$, $Q = R$ and $z' = x^*$ in (18)) and using

$$f_k(x^{k+1}) = f(x^{k+1}) - A^T[\lambda^k - H(Ax^{k+1} + By^{k+1} - b) - H(By^k - By^{k+1})],$$

it follows that

$$\begin{aligned}&(x^* - x^{k+1})^T\{f(x^{k+1}) - A^T[\lambda^k - H(Ax^{k+1} + By^{k+1} - b) - H(By^k - By^{k+1})]\}\\&\quad\ge \frac{1}{2}\big(\|x^{k+1} - x^*\|_R^2 - \|x^k - x^*\|_R^2\big) + \frac{1}{2}\|x^k - x^{k+1}\|_R^2.\end{aligned}\tag{24}$$

Adding (22) and (24), and using the monotonicity of operator f, we get

$$(x^{k+1}-x^*)^T\{A^T(\lambda^k-\lambda^*)-A^TH(Ax^{k+1}+By^{k+1}-b)-A^TH(By^k-By^{k+1})\}$$
$$\geq \frac{1}{2}\big(\|x^{k+1}-x^*\|_R^2-\|x^k-x^*\|_R^2\big)+\frac{1}{2}\|x^k-x^{k+1}\|_R^2. \tag{25}$$

Similarly, applying Lemma 2.1 to (14–15) and using

$$g_k(y^{k+1})=g(y^{k+1})-B^T[\lambda^k-H(Ax^{k+1}+By^{k+1}-b)],$$

we obtain

$$(y^*-y^{k+1})^T\{g(y^{k+1})-B^T[\lambda^k-H(Ax^{k+1}+By^{k+1}-b)]\}$$
$$\geq \frac{1}{2}\big(\|y^{k+1}-y^*\|_S^2-\|y^k-y^*\|_S^2\big)+\frac{1}{2}\|y^k-y^{k+1}\|_S^2. \tag{26}$$

Adding (23) and (26), and using the monotonicity of operator g, it follows that

$$(y^{k+1}-y^*)^T\{B^T(\lambda^k-\lambda^*)-B^TH(Ax^{k+1}+By^{k+1}-b)\}$$
$$\geq \frac{1}{2}\big(\|y^{k+1}-y^*\|_S^2-\|y^k-y^*\|_S^2\big)+\frac{1}{2}\|y^k-y^{k+1}\|_S^2. \tag{27}$$

Combining (25) and (27) and using $Ax^*+By^*=b$, we get

$$\begin{aligned}&(\lambda^k-\lambda^*)^T(Ax^{k+1}+By^{k+1}-b)+(Ax^{k+1}-Ax^*)^TH(By^{k+1}-By^k)\\ &\quad\geq \frac{1}{2}\big(\|x^{k+1}-x^*\|_R^2-\|x^k-x^*\|_R^2\big)+\frac{1}{2}\|x^k-x^{k+1}\|_R^2\\ &\qquad+\frac{1}{2}\big(\|y^{k+1}-y^*\|_S^2-\|y^k-y^*\|_S^2\big)+\frac{1}{2}\|y^k-y^{k+1}\|_S^2\\ &\qquad+\|Ax^{k+1}+By^{k+1}-b\|_H^2.\end{aligned} \tag{28}$$

Note the following identity:

$$(By^{k+1}-By^*)^TH(By^{k+1}-By^k)$$
$$=\frac{1}{2}\big(\|B(y^{k+1}-y^*)\|_H^2-\|B(y^k-y^*)\|_H^2\big)+\frac{1}{2}\|B(y^k-y^{k+1})\|_H^2. \tag{29}$$

Adding (28) and (29), and using $Ax^*+By^*=b$, we get the result of Lemma 3.1. □

Now, we observe the updated form for λ^{k+1} in Step 3. Using (16), we have the following identity:

$$2(\lambda^k-\lambda^*)^T(Ax^{k+1}+By^{k+1}-b)$$
$$=\big(\|\lambda^k-\lambda^*\|_{(\gamma H)^{-1}}^2-\|\lambda^{k+1}-\lambda^*\|_{(\gamma H)^{-1}}^2\big)+\gamma\|Ax^{k+1}+By^{k+1}-b\|_H^2. \tag{30}$$

Then, we get the following lemma:

Lemma 3.2 *Let $w^{k+1}=(x^{k+1},y^{k+1},\lambda^{k+1})$ be generated by (12–16) from given $w^k=(x^k,y^k,\lambda^k)$. Then, for any $w^*=(x^*,y^*,\lambda^*)\in\mathcal{W}^*$, we have*

$$\begin{aligned}\|w^k-w^*\|_G^2 \ge \|w^{k+1}-w^*\|_G^2 &+ \big(\|x^k-x^{k+1}\|_R^2+\|y^k-y^{k+1}\|_S^2\big)\\ &+\|B(y^k-y^{k+1})\|_H^2+(2-\gamma)\|Ax^{k+1}+By^{k+1}-b\|_H^2\\ &+2(Ax^{k+1}+By^{k+1}-b)^T H(By^k-By^{k+1}),\end{aligned}\tag{31}$$

where

$$G=\begin{pmatrix} R & & \\ & S+B^T HB & \\ & & (\gamma H)^{-1}\end{pmatrix}\tag{32}$$

is a positive-definite block-diagonal matrix.

Proof Substituting (30) in (21), we get

$$\begin{aligned}&\big(\|\lambda^k-\lambda^*\|_{(\gamma H)^{-1}}^2-\|\lambda^{k+1}-\lambda^*\|_{(\gamma H)^{-1}}^2\big)+\gamma\|Ax^{k+1}+By^{k+1}-b\|_H^2\\ &\ge \big(\|x^{k+1}-x^*\|_R^2-\|x^k-x^*\|_R^2\big)+\|x^k-x^{k+1}\|_R^2\\ &\quad+\big(\|y^{k+1}-y^*\|_S^2-\|y^k-y^*\|_S^2\big)+\|y^k-y^{k+1}\|_S^2\\ &\quad+\big(\|B(y^{k+1}-y^*)\|_H^2-\|B(y^k-y^*)\|_H^2\big)\\ &\quad+\|B(y^k-y^{k+1})\|_H^2+2\|Ax^{k+1}+By^{k+1}-b\|_H^2\\ &\quad+2(Ax^{k+1}+By^{k+1}-b)^T H(By^k-By^{k+1}).\end{aligned}$$

Thus,

$$\begin{aligned}&\big(\|x^k-x^*\|_R^2-\|x^{k+1}-x^*\|_R^2\big)+\big(\|y^k-y^*\|_S^2-\|y^{k+1}-y^*\|_S^2\big)\\ &\quad+\big(\|\lambda^k-\lambda^*\|_{(\gamma H)^{-1}}^2-\|\lambda^{k+1}-\lambda^*\|_{(\gamma H)^{-1}}^2\big)\\ &\quad+\big(\|B(y^k-y^*)\|_H^2-\|B(y^{k+1}-y^*)\|_H^2\big)\\ &\ge \|x^k-x^{k+1}\|_R^2+\|y^k-y^{k+1}\|_S^2+\|B(y^k-y^{k+1})\|_H^2\\ &\quad+(2-\gamma)\|Ax^{k+1}+By^{k+1}-b\|_H^2\\ &\quad+2(Ax^{k+1}+By^{k+1}-b)^T H(By^k-By^{k+1}).\end{aligned}\tag{33}$$

The assertion (31) is only a compact form of (33). □

Further, let us treat the last term on the right-hand side of (31). Note that, from (14) and (15), we have

$$(y^k-y^{k+1})^T\{g(y^{k+1})-B^T[\lambda^k-H(Ax^{k+1}+By^{k+1}-b)]+S(y^{k+1}-y^k)\}\ge 0.$$

Due to the same reason in the $(k-1)$th iteration, it follows that

$$(y^{k+1}-y^k)^T\{g(y^k)-B^T[\lambda^{k-1}-H(Ax^k+By^k-b)]+S(y^k-y^{k-1})\}\ge 0.$$

From the above two inequalities, we get

$$\begin{aligned}(y^k - y^{k+1})^T\{S(y^{k-1} - y^k) + B^T[(\lambda^{k-1} - \lambda^k) \\ - H(Ax^k + By^k - b) + H(Ax^{k+1} + By^{k+1} - b)]\} \\ \geq \|y^k - y^{k+1}\|_S^2 + (y^k - y^{k+1})^T(g(y^k) - g(y^{k+1})).\end{aligned} \tag{34}$$

Since

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

and

$$(g(y^k) - g(y^{k+1}))^T(y^k - y^{k+1}) \geq 0,$$

we obtain

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

Therefore,

$$\begin{aligned}&(Ax^{k+1} + By^{k+1} - b)^T HB(y^k - y^{k+1}) \\ &\quad\geq (1-\gamma)(Ax^k + By^k - b)^T HB(y^k - y^{k+1}) \\ &\qquad + \frac{1}{2}\|y^k - y^{k+1}\|_S^2 - \frac{1}{2}\|y^{k-1} - y^k\|_S^2.\end{aligned} \tag{36}$$

Substituting (36) in (31), we obtain the following lemma:

Lemma 3.3 *Let $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ be generated by (12–16) for given $w^k = (x^k, y^k, \lambda^k)$. Then, for any $w^* = (x^*, y^*, \lambda^*) \in \mathcal{W}^*$, we have*

$$\begin{aligned}&\left(\|w^k - w^*\|_G^2 + \|y^k - y^{k-1}\|_S^2\right) \\ &\quad\geq \left(\|w^{k+1} - w^*\|_G^2 + \|y^{k+1} - y^k\|_S^2\right) \\ &\qquad + \left(\|x^k - x^{k+1}\|_R^2 + \|y^k - y^{k+1}\|_S^2\right) \\ &\qquad + \|B(y^k - y^{k+1})\|_H^2 + (2-\gamma)\|Ax^{k+1} + By^{k+1} - b\|_H^2 \\ &\qquad + 2(1-\gamma)(Ax^k + By^k - b)^T H(By^k - By^{k+1}).\end{aligned} \tag{37}$$

Finally, we treat the last term of the right-hand side of (37). By choosing $\tau > 0$ and using the Cauchy-Schwarz inequality on the last term of the right-hand side of (37), we obtain

$$\begin{aligned}&2(1-\gamma)(Ax^k + By^k - b)^T H(By^k - By^{k+1}) \\ &\quad\geq -\tau^2\|Ax^k + By^k - b\|_H^2 - \frac{(1-\gamma)^2}{\tau^2}\|B(y^k - y^{k+1})\|_H^2.\end{aligned} \tag{38}$$

Substituting (38) in (37), we derive

$$\begin{aligned}
&\|w^k - w^*\|_G^2 + \tau^2\|Ax^k + By^k - b\|_H^2 + \|y^k - y^{k-1}\|_S^2 \\
&\quad \ge \|w^{k+1} - w^*\|_G^2 + \tau^2\|Ax^{k+1} + By^{k+1} - b\|_H^2 + \|y^{k+1} - y^k\|_S^2 \\
&\qquad + \big(\|x^k - x^{k+1}\|_R^2 + \|y^k - y^{k+1}\|_S^2\big) + \left(1 - \frac{(1-\gamma)^2}{\tau^2}\right)\|B(y^k - y^{k+1})\|_H^2 \\
&\qquad + (2 - \gamma - \tau^2)\|Ax^{k+1} + By^{k+1} - b\|_H^2. \qquad (39)
\end{aligned}$$

Based on (39), we have the main theorem of this paper.

Theorem 3.1 *For any $w^* = (x^*, y^*, \lambda^*) \in \mathcal{W}^*$, the sequence $\{w^k\}$ generated by the proposed method satisfies*

$$\begin{aligned}
&\big(\|w^{k+1} - w^*\|_M^2 + \|y^{k+1} - y^k\|_S^2\big) \\
&\quad \le \big(\|w^k - w^*\|_M^2 + \|y^k - y^{k-1}\|_S^2\big) - \|w^k - w^{k+1}\|_N^2, \qquad (40)
\end{aligned}$$

where

$$M = G + \tau^2 \begin{pmatrix} A^T \\ B^T \\ 0 \end{pmatrix} H(A, B, 0) \qquad (41)$$

and

$$N = \begin{pmatrix} R & 0 & 0 \\ 0 & S + \big(1 - \frac{(1-\gamma)^2}{\tau^2}\big)B^T H B & 0 \\ 0 & 0 & (2 - \gamma - \tau^2)\gamma^{-2}H^{-1} \end{pmatrix}. \qquad (42)$$

With some adaptive values of the parameter τ, the matrices M and N are positive definite when $\gamma \in (0, \frac{\sqrt{5}+1}{2})$.

Proof Rewriting $Ax^k + By^k - b$ in (39) as $A(x^k - x^*) + B(y^k - y^*)$, and expressing the first $Ax^{k+1} + By^{k+1} - b$ of (39) as $A(x^{k+1} - x^*) + B(y^{k+1} - y^*)$ and the second as $(\gamma H)^{-1}(\lambda^k - \lambda^{k+1})$, (40) can be obtained by a simple manipulation.

Through simply inspection, we have that if $\tau = \frac{\sqrt{5}-1}{2}$, $\gamma \in [\frac{3-\sqrt{5}}{2}, \frac{\sqrt{5}+1}{2})$ or $\tau = 1$, $\gamma \in (0, 1)$, then $(1 - \frac{(1-\gamma)^2}{\tau^2}) \ge 0$ and $(2 - \gamma - \tau^2) > 0$. Since $[\frac{3-\sqrt{5}}{2}, \frac{\sqrt{5}+1}{2}) \bigcup (0, 1) = (0, \frac{\sqrt{5}+1}{2})$ and G is positive definite, the results of the theorem are proved. □

4 Convergence of the Proposed Method

The following lemma plays an important role in the convergence analysis of the proposed method.

Lemma 4.1 *Let $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ be generated by (12–16) for a given $w^k = (x^k, y^k, \lambda^k)$. Then, for any $w = (x, y, \lambda) \in \mathcal{W}$, we have*

$$(x - x^{k+1})^T f_k(x^{k+1}) \ge (x^k - x^{k+1})^T R(x - x^{k+1}), \qquad (43)$$

$$(y - y^{k+1})^T g_k(y^{k+1}) \geq (y^k - y^{k+1})^T S(y - y^{k+1}). \tag{44}$$

Now, we are ready to prove the convergence of the proposed method.

Theorem 4.1 *The sequence $\{w^k\}$ generated by the proposed method converges to some w^∞ which is a solution of the SVI.*

Proof It follows from (40) that $\{w^k\}$ is a bounded sequence and

$$\lim_{k\to\infty} \|w^k - w^{k+1}\|_N = 0.$$

Consequently,

$$\lim_{k\to\infty} \|x^k - x^{k+1}\| = 0, \qquad \lim_{k\to\infty} \|y^k - y^{k+1}\| = 0 \tag{45}$$

and

$$\lim_{k\to\infty} \|Ax^{k+1} + By^{k+1} - b\| = \lim_{k\to\infty} \|(\gamma H)^{-1}(\lambda^k - \lambda^{k+1})\| = 0. \tag{46}$$

Since (see (13) and (15))

$$\begin{aligned} f_k(x^{k+1}) = f(x^{k+1}) - A^T\lambda^{k+1} + A^T HB(y^k - y^{k+1}) \\ + (1-\gamma)A^T H(Ax^{k+1} + By^{k+1} - b) \end{aligned}$$

and

$$g_k(y^{k+1}) = g(y^{k+1}) - B^T\lambda^{k+1} + (1-\gamma)B^T H(Ax^{k+1} + By^{k+1} - b),$$

it follows from (43–46) that

$$\lim_{k\to\infty} (x - x^{k+1})^T\{f(x^{k+1}) - A^T\lambda^{k+1}\} \geq 0, \quad \forall x \in \mathcal{X}, \tag{47a}$$

$$\lim_{k\to\infty} (y - y^{k+1})^T\{g(y^{k+1}) - B^T\lambda^{k+1}\} \geq 0, \quad \forall y \in \mathcal{Y}. \tag{47b}$$

Because $\{w^k\}$ is bounded, it has at least one cluster point. Let w^∞ be a cluster point of $\{w^k\}$ and let the subsequence $\{w^{k_j}\}$ converges to w^∞. It follows from (46) and (47) that

$$\lim_{j\to\infty} (x - x^{k_j})^T\{f(x^{k_j}) - A^T\lambda^{k_j}\} \geq 0, \quad \forall x \in \mathcal{X},$$

$$\lim_{j\to\infty} (y - y^{k_j})^T\{g(y^{k_j}) - B^T\lambda^{k_j}\} \geq 0, \quad \forall y \in \mathcal{Y},$$

$$\lim_{j\to\infty} (Ax^{k_j} + By^{k_j} - b) = 0.$$

Consequently,

$$\begin{aligned} &(x - x^\infty)^T\{f(x^\infty) - A^T\lambda^\infty\} \geq 0, \quad \forall x \in \mathcal{X}, \\ &(y - y^\infty)^T\{g(y^\infty) - B^T\lambda^\infty\} \geq 0, \quad \forall y \in \mathcal{Y}, \\ &Ax^\infty + By^\infty - b = 0. \end{aligned}$$

This means that w^∞ is a solution of the SVI. Note that the inequality (40) is true for all solution points of the SVI. Hence, we have

$$\begin{aligned}\|w^{k+1}-w^\infty\|_M^2 + \|y^{k+1}-y^k\|_S^2 \\ \le \|w^k - w^\infty\|_M^2 + \|y^k - y^{k-1}\|_S^2, \quad \forall k \ge 0.\end{aligned} \tag{48}$$

Because

$$\lim_{k\to\infty} \|y^k - y^{k+1}\| = 0,$$

for any given $\varepsilon > 0$, there exists $l_0 > 0$ such that

$$\|y^k - y^{k+1}\|_S^2 < \frac{\varepsilon}{2}, \quad \forall k \ge l_0. \tag{49}$$

Since $w^{k_j} \to w^\infty$ for $j \to \infty$, for the ε given above, there exists $k_l > l_0$ such that

$$\|w^{k_l} - w^\infty\|_M^2 < \frac{\varepsilon}{2}. \tag{50}$$

Therefore, for any $k > k_l$, it follows from (48–50) that

$$\|w^k - w^\infty\|_M^2 \le \|w^{k_l} - w^\infty\|_M^2 + \|y^{k_l} - y^{k_l-1}\|_S^2 \le \varepsilon.$$

This implies that the sequence $\{w^k\}$ converges to w^∞ which is a solution of the SVI. □

References

1. Glowinski, R.: Numerical Methods for Nonlinear Variational Problems. Springer, New York (1984)
2. Glowinski, R., Le Tallec, P.: Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics. SIAM Studies in Applied Mathematics, Philadelphia (1989)
3. Eckstein, J.: Some saddle-function splitting methods for convex programming. Optim. Methods Softw. **4**, 75–83 (1994)
4. Eckstein, J., Fukushima, M.: Some reformulation and applications of the alternating direction method of multipliers. In: Hager, W.W. (ed.), Large Scale Optimization: State of the Art. Kluwer Academic, Dordrecht (1994)
5. Fukushima, M.: Application of the alternating direction method of multipliers to separable convex programming problems. Comput. Optim. Appl. **2**, 93–111 (1992)
6. He, B.S., Yang, H.: Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities. Oper. Res. Lett. **23**, 151–161 (1998)
7. Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite-element approximations. Comput. Math. Appl. **2**, 17–40 (1976)
8. Gabay, D.: Applications of the method of multipliers to variational inequalities. In: Fortin, M., Glowinski, R. (eds.) Augmented Lagrange Methods: Applications to the Solution of Boundary-valued Problems, pp. 299–331. North Holland, Amsterdam (1983)
9. Kontogiorgis, S., Meyer, R.R.: A variable-penalty alternating directions method for convex optimization. Math. Program. **83**, 29–53 (1998)
10. Teboulle, M.: Convergence of proximal-like algorithms. SIAM J. Optim. **7**, 1069–1083 (1997)
11. Tseng, P.: Applications of splitting algorithm to decomposition in convex programming and variational inequalities. SIAM J. Control Optim. **29**, 119–138 (1991)
12. He, B.S., Wang, S.L., Yang, H.: A modified variable-penalty alternating directions method for monotone variational inequalities. J. Comput. Math. **21**, 495–504 (2003)
13. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
14. Martinet, B.: Regularization d'inequations variationelles par approximations sucessives. Rev. Francaise Inform. Rech. Oper. **4**, 154–159 (1970)