AMO - Advanced Modeling and Optimization, Volume 16, Number 1, 2014

A descent alternating direction method for monotone variational inequalities with separable structure

Liu Jing[1]

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

Abstract. To solve a special class of variational inequalities with separable structure, this paper proposes a descent alternating direction method based on a new residual function. The most prominent characteristic of the method is that it is easily performed, in which, only some orthogonal projections and function evaluations are involved at each iteration, so its computational load is very tiny. Under mild conditions, the global convergence of the proposed descent method is proved.

Keywords. variational inequalities; descent direction; alternating direction method; global convergence.

AMS(2000) Subject Classification: 90C30

1 Introduction

This paper considers the following variational inequality(VI) problem with separate structure: Find $u^* \in \Omega$, such that

$$(u - u^*)^\top T(u^*) \geq 0, \quad u \in \Omega \tag{1}$$

where

$$u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, T(u) = \begin{pmatrix} f(x) \\ g(y) \\ h(z) \end{pmatrix},$$

$$\Omega = \{(x, y, z) | x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}, Ax + By + Cz = b\},$$

$\mathcal{X} \subseteq \mathcal{R}^{n_1}$, $Y \subseteq \mathcal{R}^{n_2}$ and $Z \subseteq \mathcal{R}^{n_3}$ are given nonempty closed convex sets; $f : \mathcal{X} \to R^{n_1}$, $g : \mathcal{Y} \to \mathcal{R}^{n_2}$ and $h : \mathcal{Z} \to \mathcal{R}^{n_3}$ are given continuous monotone operators; $A \in \mathcal{R}^{m\times n_1}$, $B \in \mathcal{R}^{m\times n_2}$ and $C \in \mathcal{R}^{m\times n_3}$ are given full-rank matrices; $b \in \mathcal{R}^m$ is a given vector. This problem arises frequently from many application fields, e.g., network economics, traffic assignment, game theoretic problems, etc.; see [1-3]

[1]This work is supported by the Foundation of Zhejiang Provincial Education Department Under Grant Y201225096.

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

and the references therein. Throughout, we assume that $m \geq \max\{n_1, n_2, n_3\}$; and that the solution of (1) (denoted by Ω^*) is nonempty.

By attaching a Lagrange multiplier vector $\lambda \in \mathcal{R}^m$ to the linear constraints $Ax + By + Cz = b$, (1) can be equivalently transformed into the following compact form, denoted by VI($\mathcal{W}, Q$): Find $w^* \in \mathcal{W}$, such that

$$(w - w^*)^\top Q(w^*) \geq 0, \quad \forall w \in \mathcal{W} \tag{2}$$

where

$$w = \begin{pmatrix} x \\ y \\ z \\ \lambda \end{pmatrix}, Q(w) = \begin{pmatrix} f(x) - A^\top \lambda \\ g(y) - B^\top \lambda \\ h(z) - C^\top \lambda \\ Ax + By + Cz - b \end{pmatrix}, \mathcal{W} = \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \times \mathcal{R}^m.$$

We denote by $\mathcal{W}^*$ the solution of VI($\mathcal{W}, Q$). Obviously, for any $(x^*, y^*, z^*) \in \Omega^*$, there exists $\lambda^* \in \mathcal{R}^m$ such that $w^* := (x^*, y^*, z^*, \lambda^*)$ is a solution of VI($\mathcal{W}, Q$). Therefore, $\mathcal{W}^*$ is nonempty under the assumption that Ω^* is nonempty. In addition, due to the monotonicity of f, g, h, the underlying function Q of VI($\mathcal{W}, Q$) is also monotone, thus $\mathcal{W}^*$ is convex([5]).

The alternating direction method (ADM) proposed by Gabay and Mercier[6] and Gabay[7] is an efficient method for solving structured VI($\mathcal{W}, Q$), which decomposes the original problems into a series subproblems with lower scale. In particular, for a given $w^k = (x^k, y^k, z^k, \lambda^k) \in \mathcal{W}$, the new iterate $w^{k+1} := \tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{z}^k, \tilde{\lambda}^k)$ is generated by the following procedure:

$$(x' - \tilde{x}^k)^\top \{f(\tilde{x}^k) - A^\top [\lambda^k - \beta(A\tilde{x}^k + By^k + Cz^k - b)]\} \geq 0, \quad \forall x' \in \mathcal{X}, \tag{3}$$

$$(y' - \tilde{y}^k)^\top \{g(\tilde{y}^k) - B^\top [\lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k + Cz^k - b)]\} \geq 0, \quad \forall y' \in \mathcal{Y}, \tag{4}$$

$$(z' - \tilde{z}^k)^\top \{h(\tilde{z}^k) - C^\top [\lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b)]\} \geq 0, \quad \forall z' \in \mathcal{Z}, \tag{5}$$

$$\lambda^{k+1} = \lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b), \tag{6}$$

where $\beta > 0$ is a given penalty parameter for the linear constraints $Ax + By + Cz = b$. Obviously, the procedure adopts the new information whenever possible and it only requires the function values $f(x), g(y)$ and $h(z)$. For the VI($\mathcal{W}, Q$) with two separable operators, by applying the classical proximal point algorithm (PPA) [8,9] to regularize the above auxiliary VIs, He et al.[11] proposed a proximal-based ADM(PADM), which only to solve strongly monotone VIs with lower dimensions instead of the monotone VIs (3)-(5). Then, Yuan[14] developed an improved PADM(IPADM) by performing an additional descent step at each iteration, which inherits all the advantages of ADM, PPA, and descent-type methods.

However, the subproblems in ADM, PADM or IPADM are still variational inequality problems, which is usually difficult to solve, except that they possess closed-form solution or can be solved efficiently. To overcome this difficulty, He and Zhou[12] firstly proposed a modified ADM(MADM) for a special linear VI, which only needs a projection onto the simple and calculate some matrix-vector products. Then, the method was extended for nonlinear monotone VI[10], the nonlinear co-coercive VI[15], and the monotone general structured VI[4]. Motivated by the ideas in [4,12], in this paper, we propose a descent ADM for monotone VI with three separable operators. It inherits the most advantages of the MADM. More specifically, at each iteration, the method only need to perform some orthogonal projection to simple sets and some function evaluations instead of solving sub-variational inequalities, such as (3)-(5).

The rest of the paper is organized as follows. We first give some basic concepts which are useful in the following analysis in Section 2. Then, in Section 3, we describe the descent ADM method(DADM) for structured variational inequalities, and the global convergence of the new method is proved. Some concluding remarks are address in Sections 4.

2 Preliminaries

In this section, we summarize some basic properties and related definitions that will be used in the following analysis. First, we denote $\|v\| = \sqrt{v^\top v}$ as the Euclidean norm of vector v. For a given nonempty closed convex set $\mathcal{V} \subseteq \mathcal{R}^n$, the projection under Euclidean norm, denoted by $P_{\mathcal{V}}(\cdot)$, is defined as

$$P_{\mathcal{V}}(v) := \text{argmin}\{\|v-u\| | u \in \mathcal{V}\}.$$

The following well-known results about the projection operator $P_{\mathcal{V}}(\cdot)$ will be used in this paper.

Lemma 2.1 Let $\mathcal{V}$ be a closed convex set in $\mathcal{R}^n$, then the following statements hold.

$$(v - P_{\mathcal{V}}(v))^\top (u - P_{\mathcal{V}}(v)) \leq 0, \forall v \in \mathcal{R}^n, u \in \mathcal{V}; \tag{7}$$

$$\|P_{\mathcal{V}}(v) - u\| \leq \|v - u\|, \forall v \in \mathcal{R}^n, u \in \mathcal{V}. \tag{8}$$

Proof. See [13].

It is well known that VI($\mathcal{W}, Q$) is equivalent to the projection equation

$$e(w,\beta) := w - P_{\mathcal{W}}[w - \beta Q(w)] = \begin{pmatrix} x - P_{\mathcal{X}}[x - \beta(f(x) - A^\top\lambda)] \\ y - P_{\mathcal{Y}}[y - \beta(g(y) - B^\top\lambda)] \\ z - P_{\mathcal{Z}}[z - \beta(h(z) - C^\top\lambda)] \\ \beta(Ax + By + Cz - b) \end{pmatrix} = 0,$$

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

Lemma 2.2 VI($\mathcal{W}, Q$) is equivalent to the equation

$$r(w,\beta) = 0,$$

where

$$r(w,\beta) := \begin{pmatrix} r_1(w,\beta) \\ r_2(w,\beta) \\ r_3(w,\beta) \\ r_4(w,\beta) \end{pmatrix} = \begin{pmatrix} x - P_{\mathcal{X}}[x - \beta(f(x) - A^\top(\lambda - (Ax+By+Cz-b)))] \\ y - P_{\mathcal{Y}}[y - \beta(g(y) - B^\top(\lambda - (Ax+By+Cz-b)))] \\ z - P_{\mathcal{Z}}[z - \beta(h(z) - C^\top(\lambda - (Ax+By+Cz-b)))] \\ \beta(Ax+By+Cz-b) \end{pmatrix}. \tag{9}$$

Proof. We only need to prove that the solutions of $e(w,\beta)=0$ and $r(w,\beta)=0$ coincide. First, if $w^* = (x^*,y^*,z^*,\lambda^*)$ is solution of $e(w,\beta)=0$, then we have $\beta(Ax^*+By^*+Cz^*-b)=0$. Since $\beta>0$, it follows that $Ax^*+By^*+Cz^*-b=0$. Thus $r(w^*,\beta)=e(w^*,\beta)=0$, which indicates that w^* is a solution of $r(w,\beta)=0$. On the other hand, if $w^*=(x^*,y^*,z^*,\lambda^*)$ is solution of $r(w,\beta)=0$, similarly we can deduce $e(w^*,\beta)=r(w^*,\beta)=0$, thus w^* is solution of $e(w,\beta)=0$. This completes the proof.

To make the following analysis more succinct, we denote $n = n_1+n_2+n_3$, and $\bar{\lambda} = \lambda - (Ax+By+Cz-b)$.

Remark 2.1 The definition of $r(w,\beta)$ is motivated by the $e(u,\beta)$ in [4], however, there is a little difference between them. If we define $r(w,\beta)$ as $e(u,\beta)$ completely, then $\bar{\lambda} = \lambda - \beta(Ax+By+Cz-b)$, however we cannot deduce the following monotonicity of $\|r(w,\beta)\|$ and $\|r(w,\beta)\|/\beta$ in this case.

Remark 2.2 The purpose of incorporating $Ax+By+Cz-b$ in $r_1(w,\beta), r_2(w,\beta)$ and $r_3(w,\beta)$ is to generate a descent direction of $\|w-w^*\|^2/2$ whenever $w \in \mathcal{W}$ is not a solution of VI$(\mathcal{W},Q)$(see Lemma 3.1).

Lemma 2.3 For any $w \in \mathcal{R}^{n+m}$ and $\tilde{\beta} \geq \beta > 0$, we have

$$\|r(w,\beta)\| \leq \|r(w,\tilde{\beta})\|, \tag{10}$$

and

$$\frac{\|r(w,\tilde{\beta})\|}{\tilde{\beta}} \leq \frac{\|r(w,\beta)\|}{\beta}. \tag{11}$$

Proof. Let $t := \frac{\|r(w,\tilde{\beta})\|}{\|r(w,\beta)\|}$, we only need to prove that $1 \leq t \leq \frac{\tilde{\beta}}{\beta}$. Note that it is equivalent to

$$(t-1)\left(t - \frac{\tilde{\beta}}{\beta}\right) \leq 0. \tag{12}$$

From (7), we have

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

it follows from (9) that

$$(r_1(w,\beta) - \beta(f(x) - A^\top\bar{\lambda})^\top(r_1(w,\tilde{\beta}) - r_1(w,\beta)) \geq 0.$$

Thus,

$$r_1(w,\beta)^\top(r_1(w,\tilde{\beta}) - r_1(w,\beta)) \geq \beta(f(x) - A^\top\bar{\lambda})^\top(r_1(w,\tilde{\beta}) - r_1(w,\beta)). \tag{13}$$

Similarly, we have

$$r_1(w,\tilde{\beta})^\top(r_1(w,\beta) - r_1(w,\tilde{\beta})) \geq \tilde{\beta}(f(x) - A^\top\bar{\lambda})^\top(r_1(w,\beta) - r_1(w,\tilde{\beta})). \tag{14}$$

Multiplying (13) and (14) by $\tilde{\beta}$ and β, respectively, and then adding them, we get

$$(\tilde{\beta}r_1(w,\beta) - \beta r_1(w,\tilde{\beta}))^\top(r_1(w,\tilde{\beta}) - r_1(w,\beta)) \geq 0. \tag{15}$$

For $r_2(w,\beta)$, $r_3(w,\beta)$, we also have

$$(\tilde{\beta}r_2(w,\beta) - \beta r_2(w,\tilde{\beta}))^\top(r_2(w,\tilde{\beta}) - r_2(w,\beta)) \geq 0. \tag{16}$$

$$(\tilde{\beta}r_3(w,\beta) - \beta r_3(w,\tilde{\beta}))^\top(r_3(w,\tilde{\beta}) - r_3(w,\beta)) \geq 0. \tag{17}$$

For $r_4(w,\beta)$, from (9), we get

$$\begin{aligned} & (\tilde{\beta}r_4(w,\beta) - \beta r_4(w,\tilde{\beta}))^\top(r_4(w,\tilde{\beta}) - r_4(w,\beta)) \\ = \; & (\tilde{\beta}\beta(Ax+By+Cz-b) - \beta\tilde{\beta}(Ax+By+Cz-b))^\top(r_4(w,\tilde{\beta}) - r_4(w,\beta)) \\ = \; & 0. \end{aligned} \tag{18}$$

Thus, from (15)-(18), we obtain

$$(\tilde{\beta}r(w,\beta) - \beta r(w,\tilde{\beta}))^\top(r(w,\tilde{\beta}) - r(w,\beta)) \geq 0.$$

Thus,

$$\begin{aligned} & \tilde{\beta}\|r(w,\beta)\|^2 + \beta\|r(w,\tilde{\beta})\|^2 \\ \leq \; & (\beta+\tilde{\beta})r(u,\beta)^\top r(w,\tilde{\beta}) \\ \leq \; & (\beta+\tilde{\beta})\|r(u,\beta)\| \cdot \|r(w,\tilde{\beta})\|. \end{aligned}$$

Dividing the above inequality by $\|r(w,\beta)\|^2$, we get

$$\tilde{\beta} + \beta t^2 \leq (\beta+\tilde{\beta})t.$$

Then the inequality (12) holds and the lemma is proved. This completes the proof.

3 Algorithm and convergence

For convenience, set $r_i = r_i(w,\beta), i = 1,2,3,4,\ F = f(x) - A^\top\bar{\lambda},\ G = g(y) - B^\top\bar{\lambda},\ H = h(z) - C^\top\bar{\lambda}$.

Lemma 3.1 Let $w^* = (x^*, y^*, z^*, \lambda^*) \in \mathcal{W}^*$ be an arbitrary solution of VI$(\mathcal{W}, Q)$, then for any $w \in \mathcal{W}$ and $\beta > 0$, we have

$$(w - w^*)^\top d(w, \beta) \geq \varphi(w, \beta), \tag{19}$$

where

$$d(w, \beta) := \begin{pmatrix} r_1 + \beta f(x - r_1) - \beta f(x) \\ r_2 + \beta g(y - r_2) - \beta g(y) \\ r_3 + \beta h(z - r_3) - \beta h(z) \\ r_4 - \beta A r_1 - \beta B r_2 - \beta C r_3 \end{pmatrix}, \tag{20}$$

and

$$\begin{aligned} \varphi(w, \beta) = {} & \|r_1\|^2 + \|r_2\|^2 + \|r_3\|^2 + \|r_4\|^2/\beta + \beta r_1^\top (f(x - r_1) - f(x)) \\ & + \beta r_2^\top (g(y - r_2) - g(y)) + \beta r_3^\top (h(z - r_3) - h(z)) - r_4^\top (A r_1 + B r_2 + C r_3). \end{aligned} \tag{21}$$

Proof. Since $x^* \in \mathcal{X}$, it follows from (3) that,

$$\{x - \beta F - P_{\mathcal{X}}[x - \beta F]\}^\top \{P_{\mathcal{X}}[x - \beta F] - x^*\} \geq 0,$$

i.e.,

$$r_1^\top (x - x^*) \geq \|r_1\|^2 + \beta (f(x) - A^\top \bar{\lambda})^\top (x - x^* - r_1). \tag{22}$$

As w^* is a solution of VI$(\mathcal{W}, Q)$, we get

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

i.e.,

$$\beta (x - x^* - r_1)^\top f(x^*) \geq \beta (x - x^* - r_1)^\top A^\top \lambda^*. \tag{23}$$

From the monotonicity of f, we have

$$\beta (f(x - r_1) - f(x^*))^\top (x - r_1 - x^*) \geq 0,$$

i.e.,

$$\beta f(x - r_1)^\top (x - x^*) \geq \beta r_1^\top f(x - r_1) + \beta (x - x^* - r_1)^\top f(x^*). \tag{24}$$

Adding (22)-(24), we get

$$\begin{aligned} & (x - x^*)^\top (r_1 + \beta f(x - r_1)) \\ \geq {} & \|r_1\|^2 + \beta (x - x^* - r_1)^\top (f(x) - A^\top \bar{\lambda}) + \beta r_1^\top f(x - r_1) + \beta (x - x^* - r_1)^\top A^\top \lambda^* \\ = {} & \|r_1\|^2 + \beta (x - x^*)^\top f(x) + \beta r_1^\top (f(x - r_1) - f(x)) + \beta (x - x^* - r_1)^\top A^\top (\lambda^* - \bar{\lambda}) \\ = {} & \|r_1\|^2 + \beta (x - x^*)^\top f(x) + \beta r_1^\top (f(x - r_1) - f(x)) + \beta (x - x^* - r_1)^\top A^\top (\lambda^* - \lambda + r_4/\beta). \end{aligned}$$

Consequently, we obtain

$$\begin{array}{rl} & (x-x^*)^\top (r_1+\beta f(x-r_1)-\beta f(x)) \\ \geq & \|r_1\|^2+\beta r_1^\top (f(x-r_1)-f(x))+\beta(Ax-Ax^*-Ar_1)^\top(\lambda^*-\lambda+r_4/\beta). \end{array} \tag{25}$$

In a similar way, we have

$$\begin{array}{rl} & (y-y^*)^\top (r_2+\beta g(y-r_2)-\beta g(y)) \\ \geq & \|r_2\|^2+\beta r_2^\top (g(y-r_2)-g(y))+\beta(By-By^*-Br_2)^\top(\lambda^*-\lambda+r_4/\beta). \end{array} \tag{26}$$

$$\begin{array}{rl} & (z-z^*)^\top (r_3+\beta h(z-r_3)-\beta h(z)) \\ \geq & \|r_3\|^2+\beta r_3^\top (h(z-r_3)-h(z))+\beta(Cz-Cz^*-Cr_3)^\top(\lambda^*-\lambda+r_4/\beta). \end{array} \tag{27}$$

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

$$\begin{array}{rl} & (x-x^*)^\top (r_1+\beta f(x-r_1)-\beta f(x))+(y-y^*)^\top (r_2+\beta g(y-r_2)-\beta g(y)) \\ & +(z-z^*)^\top (r_3+\beta h(z-r_3)-\beta h(z)) \\ \geq & \|r_1\|^2+\|r_2\|^2+\|r_3\|^2+\beta r_1^\top (f(x-r_1)-f(x))+\beta r_2^\top (g(y-r_2)-g(y)) \\ & +\beta r_3^\top (h(z-r_3)-h(z))-(r_4-\beta Ar_1-\beta Br_2-\beta Cr_3)^\top(\lambda-\lambda^*-r_4/\beta) \\ = & \|r_1\|^2+\|r_2\|^2+\|r_3\|^2+\|r_4\|^2/\beta+\beta r_1^\top (f(x-r_1)-f(x))+\beta r_2^\top (g(y-r_2)-g(y)) \\ & +\beta r_3^\top (h(z-r_3)-h(z))-(r_4-\beta Ar_1-\beta Br_2-\beta Cr_3)^\top(\lambda-\lambda^*) \\ & -r_4^\top (Ar_1+Br_2+Cr_3). \end{array}$$

The assertion of this lemma follows from the above inequality and the definitions of $d(w,\beta)$ and $\varphi(w,\beta)$. This completes the proof.

Lemma 3.2 For any $k\geq 0$, we have

$$\begin{array}{rl} & \|r_1\|^2+\|r_2\|^2+\|r_3\|^2+\|r_4\|^2/\beta-r_4^\top(Ar_1+Br_2+Cr_3) \\ \geq & (1-\beta\|A\|^2)\|r_1\|^2+(1-\beta\|B\|^2)\|r_2\|^2+(1-\beta\|C\|^2)\|r_3\|^2+\frac{1}{4\beta}\|r_4\|^2. \end{array} \tag{28}$$

Proof. It follows from Cauchy-Schwartz inequality that

$$\begin{array}{rl} & \|r_1\|^2+\dfrac{\|r_4\|^2}{3\beta}-r_4^\top Ar_1 \\ \geq & \|r_1\|^2+\dfrac{\|r_4\|^2}{3\beta}-\|Ar_1\|\|r_4\| \\ \geq & \|r_1\|^2+\dfrac{\|r_4\|^2}{3\beta}-\dfrac{1}{2}\left(2\beta\|Ar_1\|^2+\dfrac{\|r_4\|^2}{2\beta}\right) \\ \geq & (1-\beta\|A\|^2)\|r_1\|^2+\dfrac{1}{12\beta}\|r_4\|^2. \end{array}$$

Similarly, we get

$$\|r_2\|^2+\frac{\|r_4\|^2}{3\beta}-r_4^\top Br_2\geq(1-\beta\|B\|^2)\|r_2\|^2+\frac{1}{12\beta}\|r_4\|^2.$$

$$\|r_3\|^2 + \frac{\|r_4\|^2}{3\beta} - r_4^\top C r_3 \geq (1 - \beta\|C\|^2)\|r_3\|^2 + \frac{1}{12\beta}\|r_4\|^2.$$

Adding the above three inequalities, we obtain (28). The proof is completed.

Lemma 3.3 Suppose that f, g and h is continuous. If $w \in \mathcal{W}$ is not a solution of $\mathrm{VI}(\mathcal{W}, Q)$, then for any $\delta \in (0,1)$, there exist $\nu > 0$, such that for all $\beta \in (0, \nu]$,

$$\beta(\|f(x - r_1) - f(x)\| + \|g(y - r_2) - g(y)\| + \|h(z - r_3) - h(z)\|) \leq \delta\|r(w, \beta)\|. \tag{29}$$

Proof. Suppose that (29) doesn't hold, i.e., for any $\beta > 0$, we have

$$\beta(\|f(x - r_1) - f(x)\| + \|g(y - r_2) - g(y)\| + \|h(z - r_3) - h(z)\|) > \delta\|r(w, \beta)\|.$$

Let $\beta \to 0^+$ and taking the limit in the above inequality, we obtain

$$0 \geq \lim_{\beta \to 0^+} \delta \frac{\|r(w, \beta)\|}{\beta} \geq \delta\|r(w, 1)\|,$$

where the second inequality follows from (11). Thus, $\|r(w, 1)\| = 0$, which contradicts that w is not a solution of $\mathrm{VI}(\mathcal{W}, Q)$. The proof is completed.

Let $0 < \delta < 1$, $0 < \beta < \beta_U = \min\{(1-\delta)/\|A\|^2, (1-\delta)/\|B\|^2, (1-\delta)/\|C\|^2, 1/(4\delta)\}$, and

$$\tau = \min\{1 - \beta_U\|A\|^2 - \delta, 1 - \beta_U\|B\|^2 - \delta, 1 - \beta_U\|C\|^2 - \delta, 1/(4\beta_U) - \delta\},$$

obviously, from the range of the parameter β, we get $\tau > 0$. Then from Lemmas 3.1-3.3, we have

$$(w - w^*)^\top d(w, \beta) \geq \varphi(w, \beta) \geq \tau\|r(w, \beta)\|^2, \tag{30}$$

which indicates that $-d(w, \beta)$ is a descent direction of the function $\|w - w^*\|^2/2$. This motivates us to design the following algorithm.

Algorithm 3.1

Step 0: Given $\varepsilon > 0$. Choose $w^0 \in \mathcal{W}$ and $\gamma \in (0,2)$, $\beta > 0$, $\delta \in (0,1)$, $\delta_0 \in (0,1)$, $\mu \in (0,1)$, $0 < \beta_L < \beta_U$, and $\beta_0 \in (\beta_L, \beta_U)$, where β_U satisfies the above condition. Set $k := 0$;

Step 1: If $\|r(w^k, \beta_k)\| < \varepsilon$, then stop; otherwise, find the smallest nonnegative integer m_k, such that $\beta_k = \beta\mu^{m_k}$ satisfying

$$\begin{aligned}\beta_k(\|f(x^k - r_1(w^k, \beta_k)) - f(x^k)\| + \|g(y^k - r_2(w^k, \beta_k)) - g(y^k)\| \\ + \|h(z^k - r_3(w^k, \beta_k)) - h(z^k)\|) \leq \delta\|r(w^k, \beta_k)\|.\end{aligned} \tag{31}$$

Step 2: Calculate $d(w^k, \beta_k)$ and $\varphi(w^k, \beta_k)$ from (20) and (21), respectively, and the step size

$$\alpha_k = \frac{\varphi(w^k, \beta_k)}{\|d(w^k, \beta_k)\|^2}.$$

Step 3: Determine the new iterate:

$$w^{k+1} = P_{\mathcal{W}}[w^k - \gamma\alpha_k d(w^k, \beta_k)]. \tag{32}$$

Step 4: If

$$\beta(\|f(x^k - r_1(w^k,\beta_k)) - f(x^k)\| + \|g(y^k - r_2(w^k,\beta_k)) - g(y^k)\| + \|h(z^k - r_3(w^k,\beta_k)) - h(z^k)\|) \le \delta_0\|r(w^k,\beta_k)\|,$$

then set $\beta = \beta_k/\mu$, else set $\beta = \beta_k$. Set $k := k+1$, and go to Step 1.

It follows from (20) and (29) that

$$d(w^k,\beta_k) \le (1 + \delta + \beta_U(\|A\| + \|B\| + \|C\|))\|r(w^k,\beta_k)\|,$$

consequently, using (30) and the definition of α_k, we get

$$\alpha_k \ge \frac{\tau}{1 + \delta + \beta_U(\|A\| + \|B\| + \|C\|)} \doteq \upsilon. \tag{33}$$

This shows that α_k is lower bounded away from zero.

Theorem 3.1 Suppose that the operator f, g, h are continuous and monotone, the solution set $\mathcal{W}^*$ of VI$(\mathcal{W}, Q)$ is nonempty. Then the sequence of $\{w^k\} = \{(x^k, y^k, z^k, \lambda^k)\}$ generated by the algorithm is bounded and

$$\|w^{k+1} - w^*\|^2 \le \|w^k - w^*\|^2 - \gamma(2-\gamma)\upsilon\tau\|r(w^k,\beta_k)\|^2. \tag{34}$$

Proof. Let w^* be a solution of VI$(\mathcal{W}, Q)$. Then from (31), we have

$$\begin{aligned}
& \|w^{k+1} - w^*\| \\
\le\ & \|w^k - w^* - \gamma\alpha_k d(w^k,\beta_k)\| \\
=\ & \|w^k - w^*\|^2 - 2\gamma\alpha_k(w^k - w^*)^\top d(w^k,\beta_k) + \gamma^2\alpha_k^2\|d(w^k,\beta_k)\|^2 \\
\le\ & \|w^k - w^*\|^2 - 2\gamma\alpha_k\varphi(w^k,\beta_k) + \gamma^2\alpha_k\varphi(w^k,\beta_k) \\
\le\ & \|w^k - w^*\|^2 - \gamma(2-\gamma)\upsilon\tau\|r(w^k,\beta_k)\|^2,
\end{aligned}$$

where the first inequality follows from (8), the second inequality follows from Lemma 3 and the third inequality follows from (30) and (32). Since $\gamma \in (0,2)$, we have

$$\|w^{k+1} - w^*\| \le \|w^k - w^*\| \le \cdots \le \|w^0 - w^*\|..$$

Thus, the assertion of this theorem is right. This completes the proof.

From Theorem 3.1, the sequence $\{w^k\}$ generated by the proposed method is Fejér monotone with respect to $\mathcal{W}^*$. Now, we have already proved the convergence of the new method.

Theorem 3.2 Suppose that the assumptions in Theorem 3.1 hold. Then

(1) $\lim_{k\to\infty} \|r(w^k,\beta_k)\|/\beta_k = 0$.

(2) The whole sequence $\{w^k\}$ converges to a solution of VI$(\mathcal{W}, Q)$.

Proof. It follows from (33) that

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

and thus

$$\lim_{k\to\infty} \|r(w^k,\beta_k)\| = 0. \tag{35}$$

(1) Suppose that there is an infinite index set K_0, such that

$$\|r(w^k,\beta_k)\|/\beta_k \geq \varepsilon > 0. \quad \forall k \in K_0. \tag{36}$$

From (35),(36), we have,

$$\lim_{k\to\infty, k\in K_0} \beta_k = 0.$$

Let $\bar{\lambda}^k = \lambda^k - (Ax^k + By^k + Cz^k - b)$, $F^k = f(x^k) - A^\top\bar{\lambda}^k$, $G^k = g(y^k) - B^\top\bar{\lambda}^k$, and $H^k = h(z^k) - C^\top\bar{\lambda}^k$. Since $\{w^k\}$ is bounded, $\{F^k\}, \{G^k\}$ and $\{H^k\}$ are also bounded. Therefore, from the nonexpansivity of the projection operator, we have

$$\|x^k - P_{\mathcal{X}}[x^k - \beta_k F^k/\mu]\| \leq \beta_k \|F^k\|/\mu \to 0,$$

$$\|y^k - P_{\mathcal{Y}}[y^k - \beta_k G^k/\mu]\| \leq \beta_k \|G^k\|/\mu \to 0.$$

$$\|z^k - P_{\ddagger}[z^k - \beta_k H^k/\mu]\| \leq \beta_k \|H^k\|/\mu \to 0.$$

By the choice of β_k we know that (31) is not satisfied for $m_k - 1$. That is,

$$\begin{aligned}\|f(x^k) - f(P_{\mathcal{X}}[x^k - \beta_k F^k/\mu])\| + \|g(y^k) - g(P_{\mathcal{Y}}[y^k - \beta_k G^k/\mu])\| \\ + \|h(z^k) - H(P_{\mathcal{Z}}[z^k - \beta_k H^k/\mu])\| > \delta \frac{\|r(w^k, \beta_k/\mu)\|}{\beta_k/\mu}.\end{aligned}$$

Let $k \in K_0$ and set $k \to +\infty$, and we get

$$\begin{aligned} 0 \;\leftarrow\;& \|f(x^k) - f(P_X[x^k - \beta_k F^k/\mu])\| + \|g(y^k) - g(P_Y[y^k - \beta_k G^k/\mu])\| \\ & + \|h(z^k) - H(P_{\mathcal{Z}}[z^k - \beta_k H^k/\mu])\| \\ >\;& \delta \frac{\|r(w^k, \beta_k/\mu)\|}{\beta_k/\mu} \\ \geq\;& \mu\delta \frac{\|r(w^k,\beta_k)\|}{\beta_k}, \end{aligned}$$

where the last inequality follows from (10). The above inequality contradicts to (36). Thus, the assertion of (1) holds.

(2) We divide our proof into two cases. (I) Suppose that $\lim_{k\to\infty} \sup \beta_k > 0$, then there is $\epsilon_0 > 0$ and an infinite set K_1, such that $\beta_k \geq \epsilon_0$, if $k \in K_1$. From (10), we have $\|r(w^k, \beta_k)\| \geq \|r(w^k, \epsilon_0)\|$, if $k \in K_1$. Combining (35), we get

$$\|r(w^k, \epsilon_0)\| \to 0.$$

Since $\{w^k\}$ is bounded, it has a cluster point $\bar{w}$ such that$\|r(\bar{w}, \epsilon_0)\| = 0$, which implies that $\bar{w}$ is a solution of VI$(\mathcal{W}, Q)$.

(II) Suppose that $\lim_{k\to\infty} \beta_k = 0$, then for sufficiently large k, from (11), we have

$$\frac{\|r(w^k, \beta_k)\|}{\beta_k} \geq \|r(w^k, 1)\|.$$

From (1) of this theorem and the above inequality, we get

$$\|r(w^k, 1)\| \to 0.$$

Similarly, since $\{w^k\}$ is bounded, it has a cluster point $\bar{w}$ such that$\|r(\bar{w}, 1)\| = 0$, which implies that $\bar{w}$ is a solution of VI$(\mathcal{W}, Q)$.

From Theorem 1, we have

$$\|w^{k+1} - \bar{w}\| \leq \|w^k - \bar{w}\|.$$

Therefore the whole sequence $\{w^k\}$ converges to $\bar{w}$, a solution of VI$(\mathcal{W}, Q)$. This completes the proof.

4 Conclusions

In this paper, we have proposed a descent alternating direction method for solving structured variational inequalities with three separable operators, which only need to perform some orthogonal projections and calculate functional values in the solution process. Under mild conditions, we proved the global convergence of the proposed method.

References

[1] Bertsekas, D.P., Gafni, E.M.: Projection method for variational inequalities with applications to the traffic assignment problem. Math. Program. Stud. 17, 139C159 (1982).

[2] Bertsekas, D.P., Tsitsiklis, J.N.: Parallel and Distributed Computation, Numerical Methods. Prentice- Hall, Englewood Cliffs (1989).

[3] Bertsekas, D.P.: Constrained Optimization and Lagrange MultiplierMethods. Academic Press, Boston (1982).

[4] Bnouhachem, A., Xu, M. H., Khalfaoui, M., et al., A new alternating direction method for solving variational inequalities. Computers and Mathematics with Applications, 62(2), 626-634(2011).

[5] Facchinei, F., Pang J.S., Finite-Dimensional Variational Inequalities and Complementarity Problems, vol 1, Springer Series in Operations Research, Springer, New York (2003).

[6] Gabay, D., Mercier, B., A dual algorithm for the solution of nonlinear variational problems via finite element approximations. Comput. Math. Appl. 2, 17C40 (1976).

[7] Gabay, D., Applications of the method of multipliers to variational inequalities. In: Fortin, M., Glowinski, R. (eds.) Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems, pp. 299C331. North-Holland, Amsterdam (1983).

[8] Hager, W.W., Zhang H.C., Self-adaptive inexact proximal point methods. Comput. Optim. Appl., 39(2), 161-181(2008).

[9] Hager, W.W., Zhang H.C., Asymptotic convergence analysis of a new class of proximal point methods. SIAM J. Control Optim. 46(5), 1683-1704(2007)

[10] Han, D.R., A new hybrid generalized proximal point algorithm for variational inequality problems. Journal of Global Optimization, 26, 125-140 (2003).

[11] He, B.S., Parallel splitting augmented Lagangian methods for monotone structured variational inequalities. Comput. Optim. Appl., 42, 195-212(2009).

[12] He, B.S., Zhou, J., A modified alternating direction method for convex minimization problems. Applied Mathematics Letters, 13(2), 123-130(2000).

[13] Sun, M., Two new self-adaptive descent methods without line search for co-coercive structured variational inequality problems. Journal of Applied Mathematics and Computing, 35(1-2), 179-194(2011).

[14] Yuan, X.M., An improved proximal alternating direction method for monotone variational inequalities with separable structure. Comput. Optim. Appl., 49, 17-29(2011).

[15] Zhang, W., Han, D.R., A new alternating direction method for co-coercive variational inequality problems. Computers and Mathematics with Applications, 57(7), 1168-1178(2009).