Operations Research Letters 23 (1998) 151–161

Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities

Bingsheng He[a], Hai Yang[b,*]

[a] *Department of Mathematics, Nanjing University, Nanjing, 210093, People's Republic of China*

[b] *Department of Civil Engineering, The Hong Kong University of Science & Technology, Clear Water Bay, Kowloon, Hong Kong, People's Republic of China*

Received 1 June 1997; received in revised form 1 July 1998

Abstract

Variational inequalities have important applications in mathematical programming. The alternative direction methods are suitable and often used in the literature in solving large-scale, linearly constrained variational inequalities arising in transportation research. In this paper, we present a few inequalities associated with the alternative direction method of multipliers given by Gabay and Mercier. The inequalities are helpful in understanding the algorithm. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Monotone variational inequality; Decomposition; Method of multipliers; Convergence properties

1. Introduction

Many equilibrium problems arising in network economics [1], transportation research [2] and regional science [3] can be described uniformly by variational inequalities. Much effort has been focused on developing decomposition methods for solving large and real VI problems. In the decomposition methods, the original large problem will be solved advantageously by solving a series of low-dimensional subproblems. The aim of this paper is to extend the framework of one of the decomposition methods (the method of multipliers [4]) and study its convergence properties.

The mathematical form of VI consists in finding a vector $u^* \in \Omega$ such that

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

where Ω 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 separable structure, namely (e.g. [1–3,5]),

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

* Corresponding author.

PII: S0167-6377(98)00044-3

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

where $\mathscr{X} \subset \mathscr{R}^n$ and $\mathscr{Y} \subset \mathscr{R}^m$ are given closed convex sets, $f : \mathscr{X} \to \mathscr{R}^n$ and $g : \mathscr{Y} \to \mathscr{R}^m$ are given monotone operators, $A \in \mathscr{R}^{r\times n}$, $B \in \mathscr{R}^{r\times m}$ 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 explained as a *mixed variational inequality* (VI with equality restriction $Ax + By = b$ and unrestricted variable λ):

$$\text{Find } w^* \in \mathscr{W}, \text{ such that } (w - w^*)^{\mathrm{T}} Q(w^*) \geqslant 0, \quad \forall w \in \mathscr{W}, \tag{4}$$

where

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

For convenience, we denote problem (4)–(5) as MVI($\mathscr{W}, Q$).

Typically, problems in network economics are quite large and are often solved by decomposition methods [1–3]. The decomposition method that was originally proposed by Gabay [6] and Gabay and Mercier [4] is used frequently in the literature [7,8]. At each iteration of this method, the new iterate $\tilde{w} = (\tilde{x}, \tilde{y}, \tilde{\lambda}) \in \mathscr{X} \times \mathscr{Y} \times \mathscr{R}^r$ is generated from a given triple $w = (x, y, \lambda) \in \mathscr{X} \times \mathscr{Y} \times \mathscr{R}^r$ by the following procedure: First, $\tilde{x}$ is obtained (with y and λ held fixed) by solving

$$(x' - \tilde{x})^{\mathrm{T}}(f(\tilde{x}) - A^{\mathrm{T}}[\lambda - \beta(A\tilde{x} + By - b)]) \geqslant 0, \quad \forall x' \in \mathscr{X}$$

and then $\tilde{y}$ is produced (with $\tilde{x}$ and λ held fixed) by solving

$$(y' - \tilde{y})^{\mathrm{T}}(g(\tilde{y}) - B^{\mathrm{T}}[\lambda - \beta(A\tilde{x} + B\tilde{y} - b)]) \geqslant 0, \quad \forall y' \in \mathscr{Y}.$$

Finally, the multipliers are updated by

$$\tilde{\lambda} = \lambda - \gamma\beta(A\tilde{x} + B\tilde{y} - b),$$

where $\gamma \in (0, (1+\sqrt{5})/2)$ and $\beta > 0$ are given constants (β is a penalty parameter for the linearly constrained equation $Ax + By - b = 0$). This decomposition method is referred to as a *method of multiplier* in literature [6], and the convergence proof can be found in [4,8] (for $B = I$) and [9,10] (for general B).

In this paper, besides extending the framework, we establish some inequalities with the method of multipliers that are important for understanding the algorithm. For any solution point of MVI($\mathscr{W}, Q$), denoted by $w^* = (x^*, y^*, \lambda^*)$, we show that the generated sequence $\{w^k = (x^k, y^k, \lambda^k)\}$ satisfies

$$\left\| \begin{matrix} A(x^{k+1} - x^*) \\ B(y^{k+1} - y^*) \\ \lambda^{k+1} - \lambda^* \end{matrix} \right\|_G^2 \leqslant \left\| \begin{matrix} A(x^k - x^*) \\ B(y^k - y^*) \\ \lambda^k - \lambda^* \end{matrix} \right\|_G^2 - c\|e(w^{k+1})\|^2, \tag{6}$$

where G is a positive-definite matrix, $c > 0$ is a constant and $\|e(w)\|$ is an error bound for MVI($\mathscr{W}, Q$), which measures how much w fails to be a solution point. Consequently, we prove that

$$\lim_{k\to\infty} e(w^k) = 0 \quad \text{and} \quad \lim_{k\to\infty} \left\| \begin{matrix} A(x^k - x^*) \\ B(y^k - y^*) \\ \lambda^k - \lambda^* \end{matrix} \right\|_G = 0. \tag{7}$$

The following notation is used in this paper. For any real matrix M and vector v, we denote the 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 vector z will be denoted by $||z||$, i.e., $||z|| = \sqrt{z^{\mathrm{T}}z}$. G denotes a symmetric positive definite matrix and $||z||_G$ denotes $(z^{\mathrm{T}}Gz)^{1/2}$.

2. Assumptions and the method

For the problem under consideration, we make the following standard assumptions:

Assumption A. 1. The solution set of MVI($\mathscr{W}, Q$), denoted by $\mathscr{W}^*$, is nonempty.

2. f is monotone with respect to $\mathscr{X}$, and g is monotone with respect to $\mathscr{Y}$; both f and g are continuous operators.

3. The sets $\mathscr{X}$ and $\mathscr{Y}$ are convex, closed, and bounded.

We say f is monotone with respect to $\mathscr{X}$ if

$$(x' - x)^{\mathrm{T}}(f(x') - f(x)) \geqslant 0, \quad \forall x', x \in \mathscr{X}. \tag{8}$$

A set is bounded if there is a number $K > 0$ such that the norm of any vectors of this set is not greater than K.

2.1. Description of the decomposition method

Step 0: *Given* $\varepsilon > 0$, $\gamma \in (0, (1+\sqrt{5})/2)$, $0 < \beta_1 \leqslant \beta_2, \ldots$ (*or* $\beta_1 \geqslant \beta_2 \geqslant, \ldots, \geqslant \tau > 0$), $y^0 \in \mathscr{Y}$ *and* $\lambda^0 \in \mathscr{R}^r$, *set* $k = 0$.

Step 1: *find* $x^{k+1} \in \mathscr{X}$, *such that*

$$(x' - x^{k+1})^{\mathrm{T}}(f(x^{k+1}) - A^{\mathrm{T}}[\lambda^k - \beta_{k+1}(Ax^{k+1} + By^k - b)]) \geqslant 0, \quad \forall x' \in \mathscr{X}. \tag{9}$$

Step 2: *find* $y^{k+1} \in \mathscr{Y}$, *such that*

$$(y' - y^{k+1})^{\mathrm{T}}(g(y^{k+1}) - B^{\mathrm{T}}[\lambda^k - \beta_{k+1}(Ax^{k+1} + By^{k+1} - b)]) \geqslant 0, \quad \forall y' \in \mathscr{Y}. \tag{10}$$

Step 3: *Update*

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

Step 4: *Convergence verification*: *If* $||Ax^{k+1} + By^{k+1} - b||^2 + ||By^{k+1} - By^k||^2 < \varepsilon$, *stop*; *otherwise set* $k := k+1$, *and go to Step* 1.

Remark 1. In contrast to the original method of multipliers by Gabay [6] and Gabay and Mercier [4], here B is not necessarily an identity matrix; in addition, the penalty parameter β is not necessarily fixed.

Remark 2. In view of the variant version of the method of multipliers used by Nagurney et al. [2,3], we do not consider any matrix pseudo-inverse. In addition, we allow that $\{\beta_k\}$ is a positive sequence, either monotone increasing or monotone decreasing.

We make the assumption that $\mathscr{X}$ (resp., $\mathscr{Y}$) is bounded to guarantee that the solution set of the sub-variational inequality in Step 1 (resp., in Step 2) is nonempty. In many cases, without loss of generality, it is reasonable to assume that the sets $\mathscr{X}$ and $\mathscr{Y}$ are bounded (for example, in transportation network policy modeling, shipments, derivations on a particular link, tax rates and the subsidies should in practice be limited). With the boundness assumption of $\mathscr{X}$ (resp., $\mathscr{Y}$), the sub-variational inequality in Step 1 (resp., in Step 2) has a unique solution if f (resp., g) is strongly monotone with respect to $\mathscr{X}$ (resp., $\mathscr{Y}$) or Matrix A (resp., Matrix

B) has a full column rank. Here, f is said to be strongly monotone with respect to $\mathscr{X}$ if there is a $\delta > 0$ such that

$$(x' - x)^{\mathrm{T}}(f(x') - f(x)) \geqslant \delta ||x' - x||^2, \quad \forall x', x \in \mathscr{X}.$$

We say that Matrix A has a full column rank means that Matrix $A^{\mathrm{T}}A$ is positive definite. In this paper, we focus our attention only on the framework of the decomposition method rather than the sub-VIs for which a number of solution methods can be found in, for instance, [11–23]. Although the sub-VIs are solved numerically, we assume that the exact solutions of the sub-VIs in each iteration can be obtained without difficulty.

3. Convergence analysis

Since the early works of Eaves [24] and others (e.g., see [1]), it has been well known that VI(Ω, F) is equivalent to the following projection equation:

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

where $P_{\Omega}(\cdot)$ denotes the projection on Ω. Based on the equivalence of the variational inequality and its related projection equation, solving MVI($\mathscr{W}, Q$) is equivalent to finding a zero point of

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

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

Let $\{(x^k, y^k, \lambda^k)\}$ be the sequence generated by the decomposition method described in Section 2. In the sequel convergence analysis, for convenience, we shall sometimes write briefly $x, y, \lambda, \beta, \tilde{x}, \tilde{y}, \tilde{\lambda}$ and $\tilde{\beta}$ instead of $x^k, y^k, \lambda^k, \beta_k, x^{k+1}, y^{k+1}, \lambda^{k+1}$ and β_{k+1}, respectively.

Lemma 1. *If $A\tilde{x} + B\tilde{y} - b = 0$ and $B(y - \tilde{y}) = 0$, then $\tilde{w} = (\tilde{x}, \tilde{y}, \tilde{\lambda})$ is a solution of MVI($\mathscr{W}, Q$).*

Proof. From Eqs. (9)–(11) and the assumption of this lemma, it follows that

$$(x' - \tilde{x})^{\mathrm{T}}(f(\tilde{x}) - A^{\mathrm{T}}\lambda) \geqslant 0, \quad \forall x' \in \mathscr{X},$$

$$(y' - \tilde{y})^{\mathrm{T}}(g(\tilde{y}) - B^{\mathrm{T}}\lambda) \geqslant 0, \quad \forall y' \in \mathscr{Y}$$

and

$$\tilde{\lambda} = \lambda.$$

Therefore $\tilde{w}$ satisfies Eqs. (6) and (7), and is a solution of MVI($\mathscr{W}, Q$). □

Note that, based on the equivalence of the solutions of the variational inequality and the projection equation, the generated $\tilde{x}$ (from Eq. (9)) and $\tilde{y}$ (from Eq. (10)) satisfy

$$\tilde{x} = P_{\mathscr{X}}\{\tilde{x} - [f(\tilde{x}) - A^{\mathrm{T}}(\lambda - \tilde{\beta}(A\tilde{x} + By - b))]\} \tag{12}$$

and

$$\tilde{y} = P_{\mathscr{Y}}\{\tilde{y} - [g(\tilde{y}) - B^{\mathrm{T}}(\lambda - \tilde{\beta}(A\tilde{x} + B\tilde{y} - b))]\}, \tag{13}$$

respectively. Remember that

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

Using Eqs. (12), (13) and $||P_{\mathscr{W}}(w) - P_{\mathscr{W}}(w')|| \leqslant ||w - w'||$, where $\mathscr{W} = \mathscr{X} \times \mathscr{Y} \times \mathscr{R}^r$, we have

$$||e(\tilde{w})|| = \left\| \begin{matrix} P_{\mathscr{X}}\{\tilde{x} - [f(\tilde{x}) - A^{\mathrm{T}}(\lambda - \tilde{\beta}(A\tilde{x} + By - b))]\} - P_{\mathscr{X}}\{\tilde{x} - [f(\tilde{x}) - A^{\mathrm{T}}\tilde{\lambda}]\} \\ P_{\mathscr{Y}}\{\tilde{y} - [g(\tilde{y}) - B^{\mathrm{T}}(\lambda - \tilde{\beta}(A\tilde{x} + B\tilde{y} - b))]\} - P_{\mathscr{Y}}\{\tilde{y} - [g(\tilde{y}) - B^{\mathrm{T}}\tilde{\lambda}]\} \\ A\tilde{x} + B\tilde{y} - b \end{matrix} \right\|$$

$$\leqslant \left\| \begin{matrix} A^{\mathrm{T}}[\lambda - \tilde{\lambda} - \tilde{\beta}(A\tilde{x} + By - b)] \\ B^{\mathrm{T}}[\lambda - \tilde{\lambda} - \tilde{\beta}(A\tilde{x} + B\tilde{y} - b)] \\ A\tilde{x} + B\tilde{y} - b \end{matrix} \right\|.$$

Furthermore, substituting $\lambda = \tilde{\lambda} + \gamma\tilde{\beta}(A\tilde{x} + B\tilde{y} - b)$ in the above inequality, we get

$$||e(\tilde{w})|| \leqslant \left\| \begin{matrix} (\gamma - 1)\tilde{\beta}A^{\mathrm{T}} \\ (\gamma - 1)\tilde{\beta}B^{\mathrm{T}} \\ I_r \end{matrix} \right\| \cdot ||A\tilde{x} + B\tilde{y} - b|| + ||\tilde{\beta}A^{\mathrm{T}}|| \cdot ||B(y - \tilde{y})||.$$

Therefore, for fixed γ and bounded $\{\beta_k\}$, there is a constant $v > 0$, such that

$$v||e(\tilde{w})||^2 \leqslant ||A\tilde{x} + B\tilde{y} - b||^2 + ||B(y - \tilde{y})||^2. \tag{14}$$

Since solving problem MVI($\mathscr{W}, Q$) is equivalent to finding a zero point of $e(w)$, based on the above analysis the stopping criterion in Step 4 is reasonable and we can focus our attention on showing

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

Next, we will first prove two lemmas, which make the proof of Theorem 1 precise.

Lemma 2. *For any $w^* = (x^*, y^*, \lambda^*) \in \mathscr{W}^*$, we have*

$$(\lambda - \lambda^*)^{\mathrm{T}}(A\tilde{x} + B\tilde{y} - b) \geqslant \tilde{\beta}||A\tilde{x} + B\tilde{y} - b||^2 + \tilde{\beta}(A\tilde{x} - Ax^*)^{\mathrm{T}}(By - B\tilde{y}). \tag{15}$$

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

$$(\tilde{x} - x^*)^{\mathrm{T}}(f(x^*) - A^{\mathrm{T}}\lambda^*) \geqslant 0, \tag{16}$$

$$(\tilde{y} - y^*)^{\mathrm{T}}(g(y^*) - B^{\mathrm{T}}\lambda^*) \geqslant 0 \tag{17}$$

and

$$Ax^* + By^* - b = 0. \tag{18}$$

On the other hand, from Eqs. (9) and (10), it follows that

$$(x^* - \tilde{x})^{\mathrm{T}}(f(\tilde{x}) - A^{\mathrm{T}}[\lambda - \tilde{\beta}(A\tilde{x} + By - b)]) \geqslant 0 \tag{19}$$

and

$$(y^* - \tilde{y})^{\mathrm{T}}(g(\tilde{y}) - B^{\mathrm{T}}[\lambda - \tilde{\beta}(A\tilde{x} + B\tilde{y} - b)]) \geqslant 0. \tag{20}$$

Adding Eqs. (16) and (19), and using the fact that operator f is monotonic, we get

$$(\tilde{x} - x^*)^{\mathrm{T}}(A^{\mathrm{T}}[(\lambda - \lambda^*) - \tilde{\beta}(A\tilde{x} + By - b)]) \geqslant 0. \tag{21}$$

Similarly, adding Eqs. (17) and (20), and using the fact that operator g is monotonic, it follows that

$$(\tilde{y} - y^*)^{\mathrm{T}}(B^{\mathrm{T}}[(\lambda - \lambda^*) - \tilde{\beta}(A\tilde{x} + B\tilde{y} - b)]) \geqslant 0. \tag{22}$$

Combining Eqs. (21) and (22) and using Eq. (18), we get the assertion of this lemma. □

Lemma 3. *For $k \geqslant 2$ we have*

$$\tilde{\beta}(A\tilde{x}+B\tilde{y}-b)^{\mathrm{T}}(By-B\tilde{y}) \geqslant (1-\gamma)\beta(Ax+By-b)^{\mathrm{T}}(By-B\tilde{y}). \tag{23}$$

Proof. By setting $y' = y$ in Eq. (10) we get

$$(y-\tilde{y})^{\mathrm{T}}(g(\tilde{y}) - B^{\mathrm{T}}[\lambda - \tilde{\beta}(A\tilde{x}+B\tilde{y}-b)]) \geqslant 0. \tag{24}$$

Similarly, taking $k := k-1$ and $y' = \tilde{y}$ in Eq. (10) we have

$$(\tilde{y}-y)^{\mathrm{T}}(g(y) - B^{\mathrm{T}}[\lambda^{k-1} - \beta(Ax+By-b)]) \geqslant 0. \tag{25}$$

By adding Eqs. (24) and (25) and using the fact that operator g is monotonic, we obtain

$$(\tilde{y}-y)^{\mathrm{T}}B^{\mathrm{T}}([\lambda - \tilde{\beta}(A\tilde{x}+B\tilde{y}-b)] - [\lambda^{k-1} - \beta(Ax+By-b)]) \geqslant 0. \tag{26}$$

Substituting $\lambda = \lambda^{k-1} - \gamma\beta(Ax+By-b)$ in Eq. (26), the assertion of this lemma follows immediately. □

Using the above lemmas, we prove the main theorem in this paper:

Theorem 1. *Let $w^* = (x^*, y^*, \lambda^*) \in \mathscr{W}^*$ be a solution point of $MVI(\mathscr{W}, Q)$. We have for $\gamma \leqslant 1$*

$$\begin{aligned}
&\|\tilde{\lambda}-\lambda^*\|^2 + \gamma\tilde{\beta}^2\|B(\tilde{y}-y^*)\|^2 + \gamma(1-\gamma)\tilde{\beta}^2\|A\tilde{x}+B\tilde{y}-b\|^2 \\
&\quad \leqslant \|\lambda-\lambda^*\|^2 + \gamma\tilde{\beta}^2\|B(y-y^*)\|^2 + \gamma(1-\gamma)\beta^2\|Ax+By-b\|^2 \\
&\qquad - \gamma\tilde{\beta}^2(\|A\tilde{x}+B\tilde{y}-b\|^2 + \gamma\|B(y-\tilde{y})\|^2)
\end{aligned} \tag{27}$$

and for $\gamma \geqslant 1$

$$\begin{aligned}
&\|\tilde{\lambda}-\lambda^*\|^2 + \gamma\tilde{\beta}^2\|B(\tilde{y}-y^*)\|^2 + (\gamma-1)\tilde{\beta}^2\|A\tilde{x}+B\tilde{y}-b\|^2 \\
&\quad \leqslant \|\lambda-\lambda^*\|^2 + \gamma\tilde{\beta}^2\|B(y-y^*)\|^2 + (\gamma-1)\beta^2\|Ax+By-b\|^2 \\
&\qquad - (1+\gamma-\gamma^2)\tilde{\beta}^2(\|A\tilde{x}+B\tilde{y}-b\|^2 + \gamma\|B(y-\tilde{y})\|^2).
\end{aligned} \tag{28}$$

Proof. Using Eq. (11) and

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

we get

$$\|\lambda-\lambda^*\|^2 = \|\tilde{\lambda}-\lambda^*\|^2 - \gamma^2\tilde{\beta}^2\|A\tilde{x}+B\tilde{y}-b\|^2 + 2\gamma\tilde{\beta}(\lambda-\lambda^*)^{\mathrm{T}}(A\tilde{x}+B\tilde{y}-b). \tag{29}$$

Similarly, we have

$$\gamma\tilde{\beta}^2\|B(y-y^*)\|^2 = \gamma\tilde{\beta}^2\|B(\tilde{y}-y^*)\|^2 - \gamma\tilde{\beta}^2\|By-B\tilde{y}\|^2 + 2\gamma\tilde{\beta}^2(By-By^*)^{\mathrm{T}}(By-B\tilde{y}). \tag{30}$$

Thus, combining Eqs. (29) and (30), we get

$$\begin{aligned}
\|\lambda-\lambda^*\|^2 + \gamma\tilde{\beta}^2\|B(y-y^*)\|^2 &= \|\tilde{\lambda}-\lambda^*\|^2 + \gamma\tilde{\beta}^2\|B(\tilde{y}-y^*)\|^2 \\
&\quad + 2\gamma\tilde{\beta}(\lambda-\lambda^*)^{\mathrm{T}}(A\tilde{x}+B\tilde{y}-b) + 2\gamma\tilde{\beta}^2(By-By^*)^{\mathrm{T}}(By-B\tilde{y}) \\
&\quad - \gamma^2\tilde{\beta}^2\|A\tilde{x}+B\tilde{y}-b\|^2 - \gamma\tilde{\beta}^2\|B(y-\tilde{y})\|^2.
\end{aligned} \tag{31}$$

Note that

$$2\gamma\tilde{\beta}(\lambda-\lambda^*)^{\mathrm{T}}(A\tilde{x}+B\tilde{y}-b)+2\gamma\tilde{\beta}^2(By-By^*)^{\mathrm{T}}(By-B\tilde{y})$$
$$\geqslant 2\gamma\tilde{\beta}^2||A\tilde{x}+B\tilde{y}-b||^2+2\gamma\tilde{\beta}^2(A\tilde{x}+By-b)^{\mathrm{T}}(By-B\tilde{y})$$
$$=2\gamma\tilde{\beta}^2||A\tilde{x}+B\tilde{y}-b||^2+2\gamma\tilde{\beta}^2||B(y-\tilde{y})||^2+2\gamma\tilde{\beta}^2(A\tilde{x}+B\tilde{y}-b)^{\mathrm{T}}(By-B\tilde{y})$$
$$\geqslant 2\gamma\tilde{\beta}^2||A\tilde{x}+B\tilde{y}-b||^2+2\gamma\tilde{\beta}^2||B(y-\tilde{y})||^2+2\gamma(1-\gamma)\beta\tilde{\beta}(Ax+By-b)^{\mathrm{T}}(By-B\tilde{y}). \quad (32)$$

The first inequality in Eq. (32) is based on Lemma 2 and $Ax^*+By^*=b$, and the last one is obtained from Lemma 3. Substituting Eq. (32) in Eq. (31), we get

$$||\lambda-\lambda^*||^2+\gamma\tilde{\beta}^2||B(y-y^*)||^2 \geqslant ||\tilde{\lambda}-\lambda^*||^2+\gamma\tilde{\beta}^2||B(\tilde{y}-y^*)||^2+\gamma(2-\gamma)\tilde{\beta}^2||A\tilde{x}+B\tilde{y}-b||^2$$
$$+\gamma\tilde{\beta}^2||B(y-\tilde{y})||^2-2\beta\tilde{\beta}|\gamma(1-\gamma)(Ax+By-b)^{\mathrm{T}}(By-B\tilde{y})|. \quad (33)$$

If $\gamma\leqslant 1$, using Cauchy–Schwarz inequality, we have

$$-2\beta\tilde{\beta}|\gamma(1-\gamma)(Ax+By-b)^{\mathrm{T}}(By-B\tilde{y})|\geqslant -\gamma(1-\gamma)\beta^2||Ax+By-b||^2-\gamma(1-\gamma)\tilde{\beta}^2||B(y-\tilde{y})||^2. \quad (34)$$

Substituting Eq. (34) in Eq. (33), we derive the first conclusion of this theorem. When $\gamma\geqslant 1$, again using Cauchy–Schwarz inequality, we have

$$-2\beta\tilde{\beta}|\gamma(1-\gamma)(Ax+By-b)^{\mathrm{T}}(By-B\tilde{y})|\geqslant -(\gamma-1)\beta^2||Ax+By-b||^2-(\gamma-1)\gamma^2\tilde{\beta}^2||By-B\tilde{y}||^2. \quad (35)$$

The second conclusion follows directly after we substitute Eq. (35) in Eq. (33). □

Theorem 2. *Let $\{(x^k,y^k,\lambda^k)\}$ be the sequence generated by the decomposition method described in Section 2. We have*

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

Proof. Note that $1+\gamma-\gamma^2>0$ for all $\gamma\in(0,(1+\sqrt{5})/2)$. We prove the theorem only for the case where $1\leqslant\gamma<(1+\sqrt{5})/2$ (the part of proof for $0<\gamma\leqslant 1$ is the same). First, if $\{\beta_k\}$ is monotonically increasing, then by using $\beta\leqslant\tilde{\beta}$ in Eq. (28), we have

$$(1+\gamma-\gamma^2)(||A\tilde{x}+B\tilde{y}-b||^2+\gamma||B(y-\tilde{y})||^2)\leqslant(\gamma||B(y-y^*)||^2+(\gamma-1)||Ax+By-b||^2)$$
$$-(\gamma||B(\tilde{y}-y^*)||^2+(\gamma-1)||A\tilde{x}+B\tilde{y}-b||^2)$$
$$+\frac{1}{\tilde{\beta}^2}(||\lambda-\lambda^*||^2-||\tilde{\lambda}-\lambda^*||^2). \quad (36)$$

It follows that

$$\sum_{s=1}^{k}(1+\gamma-\gamma^2)(||Ax^{s+1}+By^{s+1}-b||^2+\gamma||B(y^s-y^{s+1})||^2)$$
$$\leqslant(\gamma||B(y^1-y^*)||^2+(\gamma-1)||Ax^1+By^1-b||^2)+\frac{1}{\beta_2^2}||\lambda^1-\lambda^*||^2$$
$$+\sum_{s=2}^{k}\left(\frac{1}{\beta_{s+1}^2}-\frac{1}{\beta_s^2}\right)||\lambda^s-\lambda^*||^2. \quad (37)$$

Consequently, again using $\beta_{s+1} \geqslant \beta_s$, from the above inequality, we have

$$\begin{aligned}
&\sum_{s=1}^{k} (||Ax^{s+1} + By^{s+1} - b||^2 + \gamma||B(y^s - y^{s+1})||^2) \\
&\quad \leqslant \frac{1}{1+\gamma-\gamma^2} \left(\gamma||B(y^1 - y^*)||^2 + (\gamma - 1)||Ax^1 + By^1 - b||^2 + \frac{1}{\beta_2^2}||\lambda^1 - \lambda^*||^2 \right) \\
&\quad \leqslant +\infty.
\end{aligned} \tag{38}$$

Secondly, if $\beta_1 \geqslant \beta_2 \geqslant \cdots \geqslant \tau > 0$, then from Eq. (28) we have

$$\begin{aligned}
&(1+\gamma-\gamma^2)\tilde{\beta}^2(||A\tilde{x} + B\tilde{y} - b||^2 + \gamma||B(y - \tilde{y})||^2) \\
&\quad \leqslant ||\lambda - \lambda^*||^2 + \gamma\tilde{\beta}^2||B(y - y^*)||^2 + (\gamma - 1)\beta^2||Ax + By - b||^2 \\
&\qquad - (||\tilde{\lambda} - \lambda^*||^2 + \gamma\tilde{\beta}^2||B(\tilde{y} - y^*)||^2 + (\gamma - 1)\tilde{\beta}^2||A\tilde{x} + B\tilde{y} - b||^2).
\end{aligned} \tag{39}$$

It follows that

$$\begin{aligned}
&\sum_{s=1}^{k} (1+\gamma-\gamma^2)\beta_{s+1}^2(||Ax^{s+1} + By^{s+1} - b||^2 + \gamma||B(y^s - y^{s+1})||^2) \\
&\quad \leqslant ||\lambda^1 - \lambda^*||^2 + \gamma\beta_2^2||B(y^1 - y^*)||^2 + (\gamma - 1)\beta_1^2||Ax^1 + By^1 - b||^2 \\
&\qquad + \sum_{s=2}^{k}(\beta_{s+1}^2 - \beta_s^2)||B(y^s - y^*)||^2.
\end{aligned} \tag{40}$$

Again, using $\beta_1 \geqslant \beta_2 \geqslant \cdots \geqslant \tau$,

$$\begin{aligned}
&\sum_{s=1}^{k}(||Ax^{s+1} + By^{s+1} - b||^2 + \gamma||B(y^s - y^{s+1})||^2) \\
&\quad \leqslant \frac{1}{(1+\gamma-\gamma^2)\tau}(||\lambda^1 - \lambda^*||^2 + \gamma\beta_2^2||B(y^1 - y^*)||^2 + (\gamma - 1)\beta_1^2||Ax^1 + By^1 - b||^2) \\
&\quad < +\infty.
\end{aligned} \tag{41}$$

Thus, from Eqs. (38) and (41) we conclude that the theorem holds. □

As a direct consequence of Eq. (14) and Theorem 2, we have

$$\lim_{k\to\infty} e(w^k) = 0,$$

and thus the method is convergent.

Now we discuss some convergence properties for the case $\beta_k \equiv \beta > 0$. Note that in this case the result in Theorem 1 can be simplified as

$$\left\| \begin{matrix} A(\tilde{x} - x^*) + B(\tilde{y} - y^*) \\ B(\tilde{y} - y^*) \\ \tilde{\lambda} - \lambda^* \end{matrix} \right\|_{G_0}^2 \leqslant \left\| \begin{matrix} A(x - x^*) + B(y - y^*) \\ B(y - y^*) \\ \lambda - \lambda^* \end{matrix} \right\|_{G_0}^2 - c_0 \left\| \begin{matrix} A\tilde{x} + B\tilde{y} - b \\ B(y - \tilde{y}) \end{matrix} \right\|^2, \tag{42}$$

where G_0 is a positive definite diagonal matrix and $c_0 > 0$ is a constant. For example, in the case $\gamma = 1$, from Eq. (27) (or Eq. (28)) we have

$$\begin{aligned}
&||\tilde{\lambda} - \lambda^*||^2 + \beta^2||B(\tilde{y} - y^*)||^2 + \tfrac{1}{2}\beta^2||A\tilde{x} + B\tilde{y} - b||^2 \\
&\quad \leqslant ||\lambda - \lambda^*||^2 + \beta^2||B(y - y^*)||^2 - \beta^2(\tfrac{1}{2}||A\tilde{x} + B\tilde{y} - b||^2 + ||B(y - \tilde{y})||^2) \\
&\quad \leqslant ||\lambda - \lambda^*||^2 + \beta^2||B(y - y^*)||^2 + \tfrac{1}{2}\beta^2||Ax + By - b||^2 \\
&\qquad - \tfrac{1}{2}\beta^2(||A\tilde{x} + B\tilde{y} - b||^2 + ||B(y - \tilde{y})||^2).
\end{aligned}$$

Using $Ax^* + By^* = b$, this can be written in the form of Eq. (42) with

$$G_0 = \begin{pmatrix} \tfrac{1}{2}\beta^2 I & & \\ & \beta^2 I & \\ & & I \end{pmatrix} \quad \text{and} \quad c_0 = \tfrac{1}{2}\beta^2.$$

Furthermore, let

$$G = Z^{\mathrm{T}} G_0 Z \quad \text{with} \quad Z = \begin{pmatrix} I & I & \\ & I & \\ & & I \end{pmatrix}.$$

Then G is positive definite and Eq. (42) can be denoted as

$$\left\| \begin{matrix} A(\tilde{x} - x^*) \\ B(\tilde{y} - y^*) \\ \tilde{\lambda} - \lambda^* \end{matrix} \right\|_G^2 \leqslant \left\| \begin{matrix} A(x - x^*) \\ B(y - y^*) \\ \lambda - \lambda^* \end{matrix} \right\|_G^2 - c_0 \left\| \begin{matrix} A\tilde{x} + B\tilde{y} - b \\ B(y - \tilde{y}) \end{matrix} \right\|^2. \tag{43}$$

In addition, by using Eq. (14), we have a constant $c := c_0 v > 0$ and

$$\left\| \begin{matrix} A(x^{k+1} - x^*) \\ B(y^{k+1} - y^*) \\ \lambda^{k+1} - \lambda^* \end{matrix} \right\|_G^2 \leqslant \left\| \begin{matrix} A(x^k - x^*) \\ B(y^k - y^*) \\ \lambda^k - \lambda^* \end{matrix} \right\|_G^2 - c||e(w^{k+1})||^2. \tag{44}$$

Remember that $||e(w)||$ measures how much w fails to be in $\mathscr{W}^*$. Inequality (44) states that, if $||e(w^{k+1})||$ is not too small, then we profitted nicely from the previous iteration; conversely, if we have a very small profit from the previous iteration, it implies that $||e(w^{k+1})||$ is already very small and w^{k+1} is a 'sufficiently good' approximation of a $w^* \in \mathscr{W}^*$. In fact, from Eq. (44) we have the following theorem.

Theorem 3. *Let $\{(x^k, y^k, \lambda^k)\}$ be the sequence generated by the decomposition method with the constant penalty parameter $\beta_k \equiv \beta > 0$. We have*

$$\lim_{k\to\infty} \left\| \begin{matrix} A(x^k - x^*) \\ B(y^k - y^*) \\ \lambda^k - \lambda^* \end{matrix} \right\|_G^2 = 0.$$

Proof. First, it is easy to check that sequence $\{w^k\}$ is in a closed and bounded set, say $\overline{\mathscr{W}}$. Assume that

$$\lim_{k\to\infty} \left\| \begin{matrix} A(x^k - x^*) \\ B(y^k - y^*) \\ \lambda^k - \lambda^* \end{matrix} \right\|_G^2 = \delta_0 > 0. \tag{45}$$

It follows from Eq. (45) that there is a $\delta > 0$ such that

$$\lim_{k\to\infty} \operatorname{dist}(w^k, \mathscr{W}^*) \geqslant \delta, \quad \text{where} \quad \operatorname{dist}(w, \mathscr{W}^*) := \inf\{\|w - w^*\| \mid w^* \in \mathscr{W}^*\}.$$

Then

$$\{w^k\} \subset S := \{w \in \mathscr{W} \mid w \in \bar{\mathscr{W}} \ \ \& \ \operatorname{dist}(w, \mathscr{W}^*) \geqslant \delta\}$$

and S is a closed and bounded set. From assumption (45), $S \cap \mathscr{W}^* = \emptyset$. Since $e(w)$ is continuous, we have

$$\min\{c\|e(w)\|^2 \mid w \in S\} := \varepsilon_0 > 0. \tag{46}$$

From Eqs. (44) and (45), there is a $k_0 > 0$ such that for all $k > k_0$

$$\left\| \begin{matrix} A(x^k - x^*) \\ B(y^k - y^*) \\ \lambda^k - \lambda^* \end{matrix} \right\|_G^2 < \delta_0 + \frac{\varepsilon_0}{2}. \tag{47}$$

On the other hand, from Eqs. (44), (46) and (47) it follows that

$$\left\| \begin{matrix} A(x^{k+1} - x^*) \\ B(y^{k+1} - y^*) \\ \lambda^{k+1} - \lambda^* \end{matrix} \right\|_G^2 \leqslant \left\| \begin{matrix} A(x^k - x^*) \\ B(y^k - y^*) \\ \lambda - \lambda^* \end{matrix} \right\|_G^2 - c\|e(\tilde{w}^{k+1})\|^2 \leqslant \delta_0 - \frac{\varepsilon_0}{2}, \tag{48}$$

which contradicts Eq. (45), and the theorem is proved. □

As straightforward consequences of Theorem 3, the following corollaries emerge.

Corollary 1. *Let $\{(x^k, y^k, \lambda^k)\}$ be the sequence generated by the decomposition method with the constant penalty parameter $\beta_k \equiv \beta > 0$. Then we have $\lim_{k\to\infty}\|\lambda^k - \lambda^*\| = 0$.*

Corollary 2. *Let $\{w^k = (x^k, y^k, \lambda^k)\}$ be the sequence generated by the decomposition method with the constant penalty parameter $\beta_k \equiv \beta > 0$. If matrices A and B have full column rank, then $\lim_{k\to\infty}\|w^k - w^*\| = 0$.*

Acknowledgements

The authors would like to thank the referees for their careful reading of an earlier version of this paper and for their helpful suggestions that improved the presentation of this work. The first author gratefully acknowledges the support of NSFC grant 19671041 and the second author acknowledges the support of UGC grant RI94/95. EG01.

References

[1] A. Nagurney, Network Economics, A Variational Inequality Approach, Kluwer Academic Publishers, Dordrecht, 1993.
[2] A. Nagurney, P. Ramanujam, Transportation network policy modeling with goal targets and generalized penalty functions, Transportation Sci. 30 (1996) 3–13.
[3] A. Nagurney, S. Thore, J. Pan, Spatial market policy modeling with goal targets, Oper. Res. 44 (1996) 393–406.
[4] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite-element approximations, Computer and Mathematics with Applications 2 (1976) 17–40.
[5] S. Lawphongpanich, D.W. Hearn, Benders decomposition for variational inequalities, Math. Programming 48 (1990) 231–247.

[6] D. Gabay, Applications of the method of multipliers to variational inequalities, in: M. Fortin, R. Glowinski (Eds.), Augmented Lagrange Methods: Applications to the Solution of Boundary-valued Problems, North-Holland, Amsterdam, Netherlands, 1983, pp. 299–331.

[7] M. Fortin, R. Glowinski (Eds.), Augmented Lagrangian Methods: Applications to the solution of Boundary-Valued Problems, North-Holland, Amsterdam, 1983.

[8] R. Glowinski, Numerical Methods for Nonlinear Variational Problems, Springer, New York, 1984.

[9] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979) 964–979.

[10] P. Tseng, Applications of splitting algorithm to decomposition in convex programming and variational inequalities, Technical Report, Laboratory for Information and Decision Systems, M.I.T., November 1988.

[11] M. Fukushima, A relaxed projection method for variational inequalities, Math. Programming 35 (1986) 58–70.

[12] P.T. Harker, J.S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications, Math. Programming 48 (1990) 161–220.

[13] P.T. Harker, B. Xiao, Newton's method for the nonlinear complementarity problem: a B-differentiable equation approach, Math. Programming 48 (1990) 339–357.

[14] B.S. He, A projection and contraction method for a class of linear complementarity problem and its application in convex quadratic programming, Appl. Math. Optim. 25 (1992) 247–262.

[15] B.S. He, A new method for a class of linear variational inequalities, Math. Programming 66 (1994) 137–144.

[16] B.S. He, Solving a class of linear projection equations, Numer. Math. 68 (1994) 71–80.

[17] B.S. He, A class of new methods for monotone variational inequalities, Appl. Math. Optim. 35 (1997) 69–76.

[18] D. Kinderlehhrer, G. Stampacchia, An Introduction to Variational Inequalities and Their Applications, Academic Press, New York, 1980.

[19] J.S. Pang, D. Chan, Iterative methods for variational and complementarity problems, Math. Programming 24 (1982) 284–313.

[20] D. Sun, A projection and contraction method for nonlinear complementarity problem and its extensions, Math. Numer. Sinica 16 (1994) 183–194.

[21] D. Sun, A class of iterative methods for solving nonlinear projection equations, J. Optim. Theory Appl. 91 (1996) 123–140.

[22] B. Xiao, P.T. Harker, A nonsmooth Newton method for variational inequalities, I: theory, Math. Programming 65 (1994) 151–194.

[23] B. Xiao, P.T. Harker, A nonsmooth Newton method for variational inequalities, II: numerical results, Math. Programming 65 (1994) 195–216.

[24] B.C. Eaves, On the basic theorem of complementarity, Math. Programming 1 (1971) 68–75.