Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 680768, 8 pages
http://dx.doi.org/10.1155/2013/680768

Research Article

Convergence Analysis of Alternating Direction Method of Multipliers for a Class of Separable Convex Programming

Zehui Jia,[1] Ke Guo,[2] and Xingju Cai[1]

[1] *School of Mathematical Science and Key Laboratory for NSLSCS of Jiangsu Province, Nanjing Normal University, Nanjing, Jiangsu 210023, China*
[2] *College of Mathematics and Information, China West Normal University, Nanchong, Sichuan 637009, China*

Correspondence should be addressed to Xingju Cai; caixingju@njnu.edu.cn

Received 19 July 2013; Accepted 30 July 2013

Academic Editor: Xu Minghua

The purpose of this paper is extending the convergence analysis of Han and Yuan (2012) for alternating direction method of multipliers (ADMM) from the strongly convex to a more general case. Under the assumption that the individual functions are composites of strongly convex functions and linear functions, we prove that the classical ADMM for separable convex programming with two blocks can be extended to the case with more than three blocks. The problems, although still very special, arise naturally from some important applications, for example, route-based traffic assignment problems.

1. Introduction

In this paper, we consider the convex programming with separable functions:

$$\min \left\{ \sum_{i=1}^{m} f_i(x_i) \mid \sum_{i=1}^{m} A_i x_i = b,\ x_i \in \mathcal{X}_i,\ i = 1, 2, \ldots, m \right\}, \tag{1}$$

where $f_i : \mathcal{R}^{n_i} \to \mathcal{R} \cup \{+\infty\}$ $(i = 1, 2, \ldots, m)$ are closed proper convex functions (not necessarily smooth); $A_i \in \mathcal{R}^{l \times n_i}$ $(i = 1, 2, \ldots, m)$; $\mathcal{X}_i \subseteq \mathcal{R}^{n_i}$ $(i = 1, 2, \ldots, m)$ are closed convex sets; $b \in \mathcal{R}^l$ and $\sum_{i=1}^{m} n_i = n$. Throughout the paper, we assume that the solution set of (1) is nonempty.

For the special case of (1) with $m = 2$,

$$\begin{aligned} \min \{ & f_1(x_1) + f_2(x_2) \mid \\ & A_1 x_1 + A_2 x_2 = b,\ x_i \in \mathcal{X}_i,\ i = 1, 2 \}, \end{aligned} \tag{2}$$

the problem has been studied extensively. Among lots of numerical methods, one of the most popular methods is the alternating direction method of multipliers (ADMM) which was presented originally in [1, 2]. The iterative scheme of ADMM for (2) is as follows:

$$\begin{aligned} x_1^{k+1} &= \arg\min \Big\{ f_1(x_1) - \left(\lambda^k\right)^T A_1 x_1 \\ &\qquad + \frac{\beta}{2} \left\| A_1 x_1 + A_2 x_2^k - b \right\|^2 \mid x_1 \in \mathcal{X}_i \Big\}; \\ x_2^{k+1} &= \arg\min \Big\{ f_2(x_2) - \left(\lambda^k\right)^T A_2 x_2 \\ &\qquad + \frac{\beta}{2} \left\| A_1 x_1^{k+1} + A_2 x_2 - b \right\|^2 \mid x_2 \in \mathcal{X}_2 \Big\}; \\ \lambda^{k+1} &= \lambda^k - \beta \left(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b \right), \end{aligned} \tag{3}$$

where λ^k is Lagrange multiplier associated with the linear constraints and $\beta > 0$ is the penalty parameter. The convergence of ADMM for (2) was also established under the condition that the involved functions are convex and the constrained sets are convex too.

While there are diversified applications whose objective function is separable into $m \geq 3$ individual convex functions without coupled variables, such as traffic problems, the problem of recovering the low-rank, sparse components of matrices from incomplete and noisy observation in [3], the constrained total-variation image restoration and reconstruction problem in [4, 5], and the minimal surface PDE problem in [6], it is thus natural to extend ADMM from 2 blocks to m blocks, resulting in the iterative scheme:

$$\begin{aligned}
x_1^{k+1} &= \arg\min \Big\{ f_1(x_1) - (\lambda^k)^T A_1 x_1 \\
&\qquad + \frac{\beta}{2} \| A_1 x_1 + A_2 x_2^k + \cdots \\
&\qquad + A_m x_m^k - b \|^2 \mid x_1 \in \mathcal{X}_1 \Big\}; \\
x_2^{k+1} &= \arg\min \Big\{ f_2(x_2) - (\lambda^k)^T A_2 x_2 \\
&\qquad + \frac{\beta}{2} \| A_1 x_1^{k+1} + A_2 x_2 + \cdots \\
&\qquad + A_m x_m^k - b \|^2 \mid x_2 \in \mathcal{X}_2 \Big\}; \\
&\vdots \\
x_m^{k+1} &= \arg\min \Big\{ f_m(x_m) - (\lambda^k)^T A_m x_m \\
&\qquad + \frac{\beta}{2} \| A_1 x_1^{k+1} + A_2 x_2^{k+1} \cdots \\
&\qquad + A_m x_m - b \|^2 \mid x_m \in \mathcal{X}_m \Big\}; \\
\lambda^{k+1} &= \lambda^k - \beta \left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + \cdots + A_m x_m^{k+1} - b \right).
\end{aligned} \tag{4}$$

Unfortunately, the convergence of the natural extension is still open under convex assumption, and the recent convergence results [7] are under the assumption that all the functions involved in the objective functions are strongly convex. This lack of convergence has inspired some ADM-based methods, for example, prediction-correction type method [3, 8–11], that is, the iterate $x_1^{k+1}, x_2^{k+1}, \ldots, x_m^{k+1}$ is regarded as a prediction, and the next iterate is a correction for it. However, the numerical results show that the algorithm (4) always performs better than these variants. Recently, Han and Yuan [7] show that the global convergence of the extension of ADMM for $m \geq 3$ is valid if the involved functions are further assumed to be strongly convex. This result does not answer the open problem regarding the convergence of the extension of ADMM under the convex assumption, but it makes a key progress towards this objective.

In this paper, we consider the separable convex optimization problem (1) where each individual function f_i is the combination of a strongly convex function g_i and a linear transform B_i. That is, (1) takes the following form:

$$\min \left\{ \sum_{i=1}^m g_i(B_i x_i) \mid \sum_{i=1}^m A_i x_i = b,\ x_i \in \mathcal{X}_i,\ i = 1, 2, \ldots, m \right\}, \tag{5}$$

where $g_i : \mathcal{R}^{s_i} \to \mathcal{R} \cup \{+\infty\}$ $(i = 1, 2, \ldots, m)$ are closed proper strongly convex function with the modulus μ_i (not necessarily smooth); $A_i \in \mathcal{R}^{l \times n_i}$ $(i = 1, 2, \ldots, m)$; $\mathcal{X}_i \subseteq \mathcal{R}^{n_i}$ $(i = 1, 2, \ldots, m)$ are closed convex sets; $b \in \mathcal{R}^l$ and $\sum_{i=1}^m n_i = n$; $B_i \in \mathcal{R}^{s_i \times n_i}$ $(i = 1, 2, \ldots, m)$, where B_i may not have full column rank (if B_i has full column rank, the composite function is strongly convex and reduces to the case considered in [7]). Note that although (5) is very special, it arises frequently from many applications. One example is under the route-based traffic assignment problem [12], where g_i is the link traffic cost, B_i is the link-path incidence matrix, and x is the path follow vector.

In the following, we abuse a little the notation and still write g_i with f_i; that is, the problem under consideration is

$$\min \left\{ \sum_{i=1}^m f_i(B_i x_i) \mid \sum_{i=1}^m A_i x_i = b,\ x_i \in \mathcal{X}_i,\ i = 1, 2, \ldots, m \right\}, \tag{6}$$

where $f_i : \mathcal{R}^{s_i} \to \mathcal{R} \cup \{+\infty\}$ $(i = 1, 2, \ldots, m)$ are closed proper strongly convex function with the modulus μ_i (not necessarily smooth).

The rest of the paper is organized as follows. In the next section, we list some necessary preliminary results that will be used in the rest of the paper. We then describe the algorithm formally and analyze its global convergence under reasonable conditions in Section 3. We complete the paper with some conclusions in Section 4.

2. Preliminaries

In this section, we summarize some basic concepts and their properties that will be useful for further discussion.

Let $\|\cdot\|_p$ denote the standard definition of the l^p-norm, and particularly, let $\|\cdot\| = \|\cdot\|_2$ denote the Euclidean norm. For a symmetric and positive definite matrix G, we denote $\|\cdot\|_G$ the G-norm, that is, $\|x\|_G = \sqrt{x^T G x}$. If G is the product of a positive parameter β and the identity matrix I, that is, $G = \beta I$, we use the simpler notation: $\|\cdot\|_G = \|\cdot\|_\beta$.

Let $f : \mathcal{R}^n \to \mathcal{R} \cup \{+\infty\}$. If the domain of f denoted by $\operatorname{dom} f = \{x \in \mathcal{R}^n \mid f(x) < +\infty\}$ is not empty, then f is said to be proper. If for any $x \in \mathcal{R}^n$ and $y \in \mathcal{R}^n$, we have

$$f(tx + (1-t)y) \leq t f(x) + (1-t) f(y), \quad \forall t \in [0,1], \tag{7}$$

then f is said to be convex. Furthermore, f is said to be strongly convex with the modulus $\mu > 0$ if and only if

$$f(tx + (1-t)y) \leq t f(x) + (1-t) f(y) - \frac{1}{2}\mu t (1-t) \|x - y\|^2, \quad \forall t \in [0,1]. \tag{8}$$

A set-valued operator T defined on $\mathscr{R}^n$ is said to be monotone if and only if

$$(u-\tilde{u})^T(w-\tilde{w}) \geq 0, \quad \forall w \in Tu,\ \forall \tilde{w} \in T\tilde{u}, \tag{9}$$

and T is said to be strongly monotone with modulus $\mu > 0$ if and only if

$$(u-\tilde{u})^T(w-\tilde{w}) \geq \mu\|u-\tilde{u}\|^2, \quad \forall w \in Tu,\ \forall \tilde{w} \in T\tilde{u}. \tag{10}$$

Let $\Gamma_0(\mathscr{R}^n)$ denote the set of closed proper convex functions from $\mathscr{R}^n$ to $\mathscr{R} \cup \{+\infty\}$. For any $f \in \Gamma_0(\mathscr{R}^n)$, the subdifferential of f which is the set-valued operator, defined by

$$\partial f : x \longmapsto \Big\{\xi \in \mathscr{R}^n \mid (y-x)^T\xi + f(x) \leq f(y), \forall y \in \operatorname{dom} f\Big\}, \tag{11}$$

is monotone. Moreover, if f is strongly convex function with the modulus μ, ∂f is strongly monotone with the modulus μ.

Let F be a mapping from a set $\Omega \subset \mathscr{R}^n \to \mathscr{R}^n$. Then F is said to be co-coercive on Ω with modulus $\gamma > 0$, if

$$(u-v)^T(F(u)-F(v)) \geq \gamma\|F(u)-F(v)\|^2, \quad \forall u, v \in \Omega. \tag{12}$$

Throughout the paper, we make the following assumptions.

Assumption 1. (i) $n_i\|B_i x_i\| \geq \|A_i\|\|x_i\|$, $\forall x_i \in \mathscr{R}^{n_i}$, $i \in \{1,2,\ldots,m\}$; (ii) the solution set of (1) is nonempty.

Remark 2. Assumption 1 is a little restrictive. However, some problems can satisfy it. A remarkable one is the following route-based traffic assignment problem.

Consider a transportation network $G(\mathscr{N}, E)$, where $\mathscr{N}$ is the set of nodes. We denote the set of links by $\mathscr{A}$, and the number of the element of $\mathscr{A}$ by $N_{\mathscr{A}}$, respectively. Let RS denote the set of origin-destination (O-D) pairs. For an O-D pair $\mathrm{rs} \in \mathrm{RS}$, let q^{rs} be its traffic demand; let P^{rs} be the set of routes connecting rs, and $p \in P^{\mathrm{rs}}$; $\mathscr{N}^{\mathrm{rs}}$ denotes the number of the routes connecting rs; let h_p^{rs} be the route flow on p. The feasible route flow vector $h = (p \in P^{\mathrm{rs}} \mid \mathrm{rs} \in \mathrm{RS})$ is thus given by

$$\begin{aligned} H &= \left\{h \mid \sum_{p\in P^{\mathrm{rs}}} h_p^{\mathrm{rs}} = q^{\mathrm{rs}},\ h_p^{\mathrm{rs}} \geq 0,\ \forall p \in P^{\mathrm{rs}},\ \mathrm{rs} \in \mathrm{RS}\right\} \\ &= \Big\{h \mid e^T\left(h_1^{\mathrm{rs}}, h_2^{\mathrm{rs}}, \ldots, h_{N^{\mathrm{rs}}}^{\mathrm{rs}}\right) = q^{\mathrm{rs}}, \\ &\qquad h_p^{\mathrm{rs}} \geq 0, \forall p \in P^{\mathrm{rs}},\ \mathrm{rs} \in \mathrm{RS}\Big\}. \end{aligned} \tag{13}$$

Define E as the link-route incidence matrix such that

$$\delta_p^a = \begin{cases} 1, & \text{if } p \text{ contains link } a \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

Then, link flow f_a can be written as

$$\begin{aligned} f_a &= \sum_{\mathrm{rs}\in\mathrm{RS}} \sum_{p\in P^{\mathrm{rs}}} \delta_p^a h_p^{\mathrm{rs}}, \quad \forall a \in \mathscr{A}, \\ F &= EH = \{f \mid f = Eh,\ h \in H\}. \end{aligned} \tag{15}$$

By denoting the link cost function as $C_a(f)$ and for the additive case, the route cost function as $C_p(h)$, they can be related by

$$C_p h = \sum_{a\in\mathscr{A}} \delta_p^a C_a(f). \tag{16}$$

The user equilibrium traffic assignment problem can be formulated as a VI: find $f^* \in F$ such that

$$(f-f^*)^T C(f^*) \geq 0, \quad \forall f \in F, \tag{17}$$

or equivalently, find $h^* \in H$ such that

$$(h-h^*)^T E^T C(Eh^*) \geq 0, \quad \forall h \in H, \tag{18}$$

where $C = \{C_a\}$ is the vector of the link cost function.

In general, it is easy to show that e is a row of E and E is not a full column rank (if E is, then the above variational inequality is strongly monotone).

For simplicity, in the following, we only consider the case for $m = 3$. Notice that for $m \geq 3$, it can be proved similarly following the processing of $m = 3$.

3. The Method

In this section, we consider the following convex minimization problem with linear constraint, where the objective function is in the form of the sum of three individual functions without coupled variable:

$$\begin{aligned} &\min \quad f_1(B_1x_1) + f_2(B_2x_2) + f_3(B_3x_3) \\ &\text{s.t.} \quad A_1x_1 + A_2x_2 + A_3x_3 = b, \quad x_i \in \mathscr{X}_i,\ i = 1,2,3, \end{aligned} \tag{19}$$

where $f_i : \mathscr{R}^{s_i} \to \mathscr{R} \cup \{+\infty\}$ $(i = 1,2,3)$ are closed proper strongly convex function with the modulus μ_i (not necessarily smooth); $B_i \in \mathscr{R}^{s_i\times n_i}$ $(i = 1,2,3)$, $A_i \in \mathscr{R}^{l\times n_i}$ $(i = 1,2,3)$; $\mathscr{X}_i \subseteq \mathscr{R}^{n_i}$ $(i = 1,2,3)$ are closed convex sets; $b \in \mathscr{R}^l$ and $\sum_{i=1}^3 n_i = n$.

The iterative scheme of ADMM for problem (19) is as follows:

$$\begin{aligned} x_1^{k+1} &= \arg\min \Big\{ f_1(B_1x_1) - (\lambda^k)^T A_1x_1 \\ &\quad + \frac{\beta}{2}\Big\| A_1x_1 \\ &\quad + A_2x_2^k + A_3x_3^k - b\Big\|^2 \mid x_1 \in \mathcal{X}_1 \Big\}, \\ x_2^{k+1} &= \arg\min \Big\{ f_2(B_2x_2) - (\lambda^k)^T A_2x_2 \\ &\quad + \frac{\beta}{2}\Big\| A_1x_1^{k+1} \\ &\quad + A_2x_2 + A_3x_3^k - b\Big\|^2 \mid x_2 \in \mathcal{X}_2 \Big\}, \\ x_3^{k+1} &= \arg\min \Big\{ f_3(B_3x_3) - (\lambda^k)^T A_3x_3 \\ &\quad + \frac{\beta}{2}\Big\| A_1x_1^{k+1} \\ &\quad + A_2x_2^{k+1} + A_3x_3 - b\Big\|^2 \mid x_3 \in \mathcal{X}_3 \Big\}, \\ \lambda^{k+1} &= \lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^{k+1} - b\right), \end{aligned} \tag{20}$$

where λ^k is the Lagrangian multiplier associated with the linear constraints and $\beta > 0$ is the penalty parameter.

4. Convergence

In this section, we prove the convergence of the extended ADMM for problem (19). As the assumptions aforementioned, by invoking the first-order necessary and sufficient condition for convex programming, we easily see that the problem (19) under the condition is characterized by the following variational inequality (VI): find $u^* \in \mathcal{U}$ and $\xi_i^* \in \partial f_i(B_ix_i^*)$ such that

$$(u - u^*)^T Q(u^*) \geq 0, \quad \forall u \in \mathcal{U}, \tag{21}$$

where

$$\begin{gathered} u := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix}, \qquad Q(u) := \begin{pmatrix} B_1^T\xi_1 - A_1^T\lambda \\ B_2^T\xi_2 - A_2^T\lambda \\ B_3^T\xi_3 - A_3^T\lambda \\ A_1x_1 + A_2x_2 + A_3x_3 - b \end{pmatrix}, \\ \mathcal{U} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times \mathcal{R}^l. \end{gathered} \tag{22}$$

We denote the VI (21)-(22) by MVI$(\mathcal{U}, Q)$.

Similarly, in [7], we propose an easily implementable stopping criterion for executing (20):

$$\max\left\{\max_{1\leq i\leq 3}\left\|A_ix_i^k - A_ix_i^{k+1}\right\|, \left\|\sum_{i=1}^3 A_ix_i^k - b\right\|\right\} \leq \epsilon, \tag{23}$$

and its rationale can be seen in the following lemma.

Lemma 3 (see [7]). *If $\sum_{i=1}^3 A_ix_i^k - b = 0$ and $A_ix_i^k = A_ix_i^{k+1}$ $(i = 1, 2, 3)$, then $(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}, \lambda^{k+1})$ is a solution of MVI$(\mathcal{U}, Q)$.*

Lemma 3 implies that the iterate $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ is a solution of MVI$(\mathcal{U}, Q)$ when the inequality (23) holds with $\epsilon = 0$. Some techniques of establishing the error bounds in [13] can help us analyze how precisely the iterate satisfies the optimality conditions when the proposed stopping criterion is satisfied with a tolerance $\epsilon > 0$.

Lemma 4. *Let $(x_1^*, x_2^*, x_3^*, \lambda^*)$ be the solution of the problem* (19), *and let λ^* be a corresponding Lagrange multiplier associated with the linear constraint. Then, the sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ generated by* (20) *satisfies*

$$\begin{aligned} &\left(\lambda^k - \lambda^*\right)^T \left(\sum_{i=1}^3 A_ix_i^{k+1} - b\right) \\ &\quad \geq \sum_{i=1}^3 \left(x_i^{k+1} - x_i^*\right)^T \left(B_i^T\xi_i^{k+1} - B_i^T\xi_i^*\right) \\ &\qquad + \beta\left\|\sum_{i=1}^3 A_ix_i^{k+1} - b\right\|^2 + \beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T \\ &\qquad \times \left[A_2x_2^k - A_2x_2^{k+1} + \left(A_3x_3^k - A_3x_3^{k+1}\right)\right] \\ &\qquad + \beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T \left(A_3x_3^k - A_3x_3^{k+1}\right). \end{aligned} \tag{24}$$

Proof. By invoking the first-order optimality condition for the x_i^{k+1}-related subproblem in (20), for any $x_i \in \mathcal{X}_i$, $i = 1, 2, 3$, we get

$$\begin{aligned} &\left(x_1 - x_1^{k+1}\right)^T \Big\{B_1^T\xi_1^{k+1} \\ &\qquad - A_1^T\left[\lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^k + A_3x_3^k - b\right)\right]\Big\} \\ &\quad \geq 0, \\ &\left(x_2 - x_2^{k+1}\right)^T \Big\{B_2^T\xi_2^{k+1} - A_2^T \\ &\qquad \times \left[\lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^k - b\right)\right]\Big\} \\ &\quad \geq 0, \end{aligned}$$

$$
\begin{aligned}
&\left(x_3 - x_3^{k+1}\right)^T \left\{B_3^T \xi_3^{k+1} - A_3^T \right.\\
&\qquad \left.\times \left[\lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k+1} - b\right)\right]\right\}\\
&\quad \geq 0.
\end{aligned}
\tag{25}
$$

Setting $x_i = x_i^*$ $(i = 1, 2, 3)$ in (25), we have

$$
\begin{aligned}
&\left(x_1^* - x_1^{k+1}\right)^T \left\{B_1^T \xi_1^{k+1} \right.\\
&\qquad \left. - A_1^T \left[\lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^{k} + A_3 x_3^{k} - b\right)\right]\right\}\\
&\quad \geq 0,\\
&\left(x_2^* - x_2^{k+1}\right)^T \left\{B_2^T \xi_2^{k+1} - A_2^T \right.\\
&\qquad \left.\times \left[\lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k} - b\right)\right]\right\}\\
&\quad \geq 0,\\
&\left(x_3^* - x_3^{k+1}\right)^T \left\{B_3^T \xi_3^{k+1} - A_3^T \right.\\
&\qquad \left.\times \left[\lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k+1} - b\right)\right]\right\}\\
&\quad \geq 0.
\end{aligned}
\tag{26}
$$

On the other hand, setting $(x_1, x_2, x_3) = (x_1^{k+1}, x_2^{k+1}, x_3^{k+1})$ in (21), it follows that

$$
\begin{pmatrix} x_1^{k+1} - x_1^* \\ x_2^{k+1} - x_2^* \\ x_3^{k+1} - x_3^* \end{pmatrix}^T
\begin{pmatrix} B_1^T \xi_1^* - A_1^T \lambda^* \\ B_2^T \xi_2^* - A_2^T \lambda^* \\ B_3^T \xi_3^* - A_3^T \lambda^* \end{pmatrix} \geq 0.
\tag{27}
$$

Adding (26) and (27), we obtain

$$
\begin{aligned}
&\left(x_1^{k+1} - x_1^*\right)^T \left\{\left(B_1^T \xi_1^* - B_1^T \xi_1^{k+1}\right) - A_1^T\left(\lambda^* - \lambda^k\right)\right.\\
&\qquad \left. - \beta A_1^T \left(A_1 x_1^{k+1} + A_2 x_2^{k} + A_3 x_3^{k} - b\right)\right\} \geq 0,\\
&\left(x_2^{k+1} - x_2^*\right)^T \left\{\left(B_2^T \xi_2^* - B_2^T \xi_2^{k+1}\right) - A_2^T\left(\lambda^* - \lambda^k\right)\right.\\
&\qquad \left. - \beta A_2^T \left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k} - b\right)\right\} \geq 0,\\
&\left(x_3^{k+1} - x_3^*\right)^T \left\{\left(B_3^T \xi_3^* - B_3^T \xi_3^{k+1}\right) - A_3^T\left(\lambda^* - \lambda^k\right)\right.\\
&\qquad \left. - \beta A_3^T \left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k+1} - b\right)\right\}\\
&\quad \geq 0.
\end{aligned}
\tag{28}
$$

With the rearrangement of the above inequalities, we derive that

$$
\begin{aligned}
&\left(x_1^{k+1} - x_1^*\right)^T A_1^T \left(\lambda^k - \lambda^*\right)\\
&\quad \geq \left(x_1^{k+1} - x_1^*\right)^T \left(B_1^T \xi_1^{k+1} - B_1^T \xi_1^*\right)\\
&\qquad + \beta\left(A_1 x_1^{k+1} - A_1 x_1^*\right)^T \left(\sum_{i=1}^{3} A_i x_i^{k+1} - b\right)\\
&\qquad + \beta\left(A_1 x_1^{k+1} - A_1 x_1^*\right)^T\\
&\qquad \times \left[\left(A_2 x_2^k - A_2 x_2^{k+1}\right) + \left(A_3 x_3^k - A_3 x_3^{k+1}\right)\right],\\
&\left(x_2^{k+1} - x_2^*\right)^T A_2^T \left(\lambda^k - \lambda^*\right)\\
&\quad \geq \left(x_2^{k+1} - x_2^*\right)^T \left(B_2^T \xi_2^{k+1} - B_2^T \xi_2^*\right)\\
&\qquad + \beta\left(A_2 x_2^{k+1} - A_2 x_2^*\right)^T \left(\sum_{i=1}^{3} A_i x_i^{k+1} - b\right)\\
&\qquad + \beta\left(A_2 x_2^{k+1} - A_2 x_2^*\right)^T \left(A_3 x_3^k - A_3 x_3^{k+1}\right),\\
&\left(x_3^{k+1} - x_3^*\right)^T A_3^T \left(\lambda^k - \lambda^*\right)\\
&\quad \geq \left(x_3^{k+1} - x_3^*\right)^T \left(B_3^T \xi_3^{k+1} - B_3^T \xi_3^*\right)\\
&\qquad + \beta\left(A_3 x_3^{k+1} - A_3 x_3^*\right)^T \left(\sum_{i=1}^{3} A_i x_i^{k+1} - b\right).
\end{aligned}
\tag{29}
$$

Adding the above inequalities (29), we have

$$
\begin{aligned}
&\left(\lambda^k - \lambda^*\right)^T \left(\sum_{i=1}^{3} A_i x_i^{k+1} - b\right)\\
&\quad \geq \sum_{i=1}^{3} \left(x_i^{k+1} - x_i^*\right)^T \left(B_i^T \xi_i^{k+1} - B_i^T \xi_i^*\right) + \beta \left\| \sum_{i=1}^{3} A_i x_i^{k+1} - b \right\|^2\\
&\qquad + \beta\left(A_1 x_1^{k+1} - A_1 x_1^*\right)^T \left[\left(A_2 x_2^k - A_2 x_2^{k+1}\right)\right.\\
&\qquad\qquad \left. + \left(A_3 x_3^k - A_3 x_3^{k+1}\right)\right]\\
&\qquad + \beta\left(A_2 x_2^{k+1} - A_2 x_2^*\right)^T \left(A_3 x_3^k - A_3 x_3^{k+1}\right).
\end{aligned}
\tag{30}
$$

The proof is complete. □

Hereafter, we define a matrix which will make the notation of proof more succinct. More specifically, let

$$
M = \begin{pmatrix}
2\beta A_1^T A_1 & 0 & 0 & 0 \\
0 & 2\beta A_2^T A_2 & 0 & \\
0 & 0 & 2\beta A_3^T A_3 & 0 \\
0 & 0 & 0 & \dfrac{1}{\beta} I
\end{pmatrix}.
\tag{31}
$$

Obviously, M is a positive semidefinite matrix, only for analysis convenience; we denote

$$\|u\|_M^2 = 2\beta\left(\left\|A_1x_1\right\|^2 + \left\|A_2x_2\right\|^2 + \left\|A_3x_3\right\|^2\right) + \|\lambda\|_{1/\beta}^2. \tag{32}$$

Lemma 5. *Let $u^* = (x_1^*, x_2^*, x_3^*, \lambda^*)$ be a solution of MVI($\mathcal{U}$, Q), and let the sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ be generated by* (20). *Then, one has*

$$\begin{aligned}\left\|u^{k+1} - u^*\right\|_M^2 \le{}& \left\|u^k - u^*\right\|_M^2 \\ &+ \sum_{i=1}^{3} 3\beta\left\|A_ix_i^{k+1} - A_ix_i^*\right\|^2 \\ &- 2\sum_{i=1}^{3}\mu_i\left\|B_ix_i^{k+1} - B_ix_i^*\right\|^2.\end{aligned} \tag{33}$$

Proof. From (20) and Lemma 4, we have

$$\begin{aligned}\left\|\lambda^{k+1} - \lambda^*\right\|_{1/\beta}^2 ={}& \left\|\lambda^k - \lambda^* - \beta\left(\sum_{i=1}^{3}A_ix_i^{k+1} - b\right)\right\|_{1/\beta}^2 \\ ={}& \left\|\lambda^k - \lambda^*\right\|_{1/\beta}^2 \\ &- 2\left(\lambda^k - \lambda^*\right)^T\left(\sum_{i=1}^{3}A_ix_i^{k+1} - b\right) \\ &+ \beta\left\|\sum_{i=1}^{3}A_ix_i^{k+1} - b\right\|^2 \\ \le{}& \left\|\lambda^k - \lambda^*\right\|_{1/\beta}^2 \\ &- 2\sum_{i=1}^{3}\left(x_i^{k+1} - x_i^*\right)^T\left(B_i^T\xi_i^{k+1} - B_i^T\xi_i^*\right) \\ &- \beta\left\|\sum_{i=1}^{3}A_ix_i^{k+1} - b\right\|^2 \\ &- 2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T \\ &\times\left(\sum_{i=2}^{3}\left(A_ix_i^k - A_ix_i^{k+1}\right)\right) \\ &- 2\beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T\left(A_3x_3^k - A_3x_3^{k+1}\right).\end{aligned} \tag{34}$$

Since

$$\begin{aligned}\left\|\sum_{i=1}^{3}\left(A_ix_i^{k+1} - A_ix_i^*\right)\right\|^2 ={}& \sum_{i=1}^{3}\left\|A_i\left(x_i^{k+1} - x_i^*\right)\right\|^2 \\ &+ \sum_{i\ne j}\left(A_i\left(x_i^{k+1} - x_i^*\right)\right)^T A_j\left(x_j^{k+1} - x_j^*\right),\end{aligned} \tag{35}$$

and $A_1x_1^* + A_2x_2^* + A_3x_3^* = b$, we can get

$$\begin{aligned}-\beta\left\|\sum_{i=1}^{3}A_ix_i^{k+1} - b\right\|^2 ={}& -\beta\sum_{i=1}^{3}\left\|A_i\left(x_i^{k+1} - x_i^*\right)\right\|^2 \\ &- \beta\sum_{i\ne j}\left(A_i\left(x_i^{k+1} - x_i^*\right)\right)^T A_j\left(x_j^{k+1} - x_j^*\right).\end{aligned} \tag{36}$$

Using Cauchy-Schwarz inequality, we have

$$\begin{aligned}&-2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(\sum_{i=2}^{3}\left(A_ix_i^k - A_ix_i^{k+1}\right)\right) \\ &\quad- 2\beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T\left(A_3x_3^k - A_3x_3^{k+1}\right) \\ &= -2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(A_2x_2^k - A_2x_2^*\right) \\ &\quad+ 2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(A_2x_2^{k+1} - A_2x_2^*\right) \\ &\quad- 2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(A_3x_3^k - A_3x_3^*\right) \\ &\quad+ 2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(A_3x_3^{k+1} - A_3x_3^*\right) \\ &\quad- 2\beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T\left(A_3x_3^k - A_3x_3^*\right) \\ &\quad+ 2\beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T\left(A_3x_3^{k+1} - A_3x_3^*\right) \\ &\le 2\beta\left\|A_1x_1^{k+1} - A_1x_1^*\right\|^2 + \beta\left\|A_2x_2^{k+1} - A_2x_2^*\right\|^2 \\ &\quad+ \beta\left\|A_2\left(x_2^k - x_2^*\right)\right\|^2 + 2\beta\left\|A_3\left(x_3^k - x_3^*\right)\right\|^2 \\ &\quad+ \beta\sum_{i\ne j}\left(A_i\left(x_i^{k+1} - x_i^*\right)\right)^T A_j\left(x_j^{k+1} - x_j^*\right).\end{aligned} \tag{37}$$

Substituting (36) and (37) into (34), we get

$$\begin{aligned}\left\|\lambda^{k+1}-\lambda^*\right\|_{1/\beta}^2 \le \left\|\lambda^k-\lambda^*\right\|_{1/\beta}^2 + 2\beta\sum_{i=1}^3\left\|A_i\left(x_i^k-x_i^*\right)\right\|^2 \\ + \beta\sum_{i=1}^3\left\|A_i\left(x_i^{k+1}-x_i^*\right)\right\|^2 \\ - 2\sum_{i=1}^3\left(x_i^{k+1}-x_i^*\right)^T\left(B_i^T\xi_i^{k+1}-B_i^T\xi_i^*\right).\end{aligned} \tag{38}$$

Since f_i is strongly convex, from the strong monotonicity of the subdifferential mapping ∂f_i (with the modulus μ_i), then we have

$$\begin{aligned}&\left(x_i^{k+1}-x_i^*\right)^T\left(B_i^T\xi_i^{k+1}-B_i^T\xi_i^*\right) \\ &\quad = \left(B_ix_i^{k+1}-B_ix_i^*\right)^T\left(\xi_i^{k+1}-\xi_i^*\right) \ge \mu_i\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2,\end{aligned} \tag{39}$$

where $\xi_i^* \in \partial f_i(B_ix_i^*)$, $\xi_i^{k+1} \in \partial f_i(B_ix_i^{k+1})$, for any $i \in \{1,2,3\}$.

By using the notion of $\|u^{k+1}-u^*\|_M^2$, from (38) we have

$$\begin{aligned}&\left\|u^{k+1}-u^*\right\|_M^2 \\ &\quad = \left\|\lambda^{k+1}-\lambda^*\right\|_{1/\beta}^2 \\ &\qquad + 2\beta\sum_{i=1}^3\left\|A_i(x_i^{k+1}-x_i^*)\right\|^2 \\ &\quad \le \left\|\lambda^k-\lambda^*\right\|_{1/\beta}^2 + 2\beta\sum_{i=1}^3\left\|A_i(x_i^k-x_i^*)\right\|^2 \\ &\qquad + 3\beta\sum_{i=1}^3\left\|A_i(x_i^{k+1}-x_i^*)\right\|^2 - 2\sum_{i=1}^3\mu_i\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2 \\ &\quad \le \left\|u^k-u^*\right\|_M^2 + \sum_{i=1}^3 3\beta\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2 \\ &\qquad - 2\sum_{i=1}^3\mu_i\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2.\end{aligned} \tag{40}$$

The proof is complete. □

Theorem 6. *Under Assumption 1, for any*

$$0 < \beta < \min_{1\le i\le 3}\left\{\frac{2\mu_i}{3n_i^2}\right\}, \tag{41}$$

the sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ generated by (20) converges to a solution of MVI($\mathcal{U}$, Q).

Proof. From Lemma 5, we have

$$\begin{aligned}\left\|u^{k+1}-u^*\right\|_M^2 \le \left\|u^k-u^*\right\|_M^2 + \sum_{i=1}^3 3\beta\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2 \\ - 2\sum_{i=1}^3\mu_i\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2,\end{aligned} \tag{42}$$

where

$$0 < \beta < \min_{1\le i\le 3}\left\{\frac{2\mu_i}{3n_i^2}\right\}. \tag{43}$$

From Assumption 1, it follows that

$$\begin{aligned}\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2 &\le \left\|A_i\right\|^2\left\|x_i^{k+1}-x_i^*\right\|^2 \\ &\le n_i^2\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2, \quad i=1,2,3.\end{aligned} \tag{44}$$

Consequently,

$$\begin{aligned}&\left\|u^{k+1}-u^*\right\|_M^2 \le \left\|u^k-u^*\right\|_M^2 \\ &\qquad + \sum_{i=1}^3\left(3\beta-\frac{2\mu_i}{n_i^2}\right)\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2.\end{aligned} \tag{45}$$

From (45), we have

$$\left\|u^{k+1}-u^*\right\|_M^2 \le \left\|u^k-u^*\right\|_M^2 \le \cdots \le \left\|u^0-u^*\right\|_M^2 < +\infty, \tag{46}$$

which means that the generated sequence $\{u^k\}$ is bounded.

Furthermore, it follows that

$$\begin{aligned}&\sum_{k=0}^{+\infty}\left\{\sum_{i=1}^3\left(2\frac{\mu_i}{n_i^2}-3\beta\right)\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2\right\} \\ &\quad \le \sum_{k=0}^{+\infty}\left\{\left\|u^k-u^*\right\|_M - \left\|u^{k+1}-u^*\right\|_M\right\} < +\infty,\end{aligned} \tag{47}$$

which means that

$$\lim_{k\to+\infty}\sum_{i=1}^3\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2 = 0. \tag{48}$$

Therefore, we have

$$\lim_{k\to+\infty}\left\|\sum_{i=1}^3 A_ix_i^{k+1}-b\right\|^2 = 0. \tag{49}$$

Since $\|A_i\|$ is nonzero and bounded, from (48) we have

$$\lim_{k\to+\infty}\left\|x_i^{k+1}-x_i^*\right\| = 0, \quad \forall i = 1,2,3. \tag{50}$$

Since $\{u^k\}$ is bounded, $\{\lambda^k\}$ has at least one cluster point, say $\overline{\lambda}$. Let $\{\lambda^{k_j}\}$ be the corresponding subsequence that converges to $\bar{\lambda}$. Taking a limit along this subsequence in (25) and (49), we obtain $\xi_i^* \in \partial f_i(B_i x_i^*)$,

$$\left(x_i - x_i^*\right)^T \left(B_i^T \xi_i^* - A_i^T \bar{\lambda}\right) \geq 0, \quad \forall x_i \in \mathcal{X}_i,\ i = 1, 2, 3,$$
$$\sum_{i=1}^{3} A_i x_i^* - b = 0, \tag{51}$$

which follows that $\bar{\lambda}$ is an optimal Lagrange multiplier. Since λ^* is arbitrary, we can set $\lambda^* = \bar{\lambda}$ in (46) and conclude that the whole generated sequence converges to a solution of MVI$(\mathcal{U}, Q)$. □

5. Conclusions

In this paper, we extend the convergence analysis of the ADMM for the separable convex optimization problem with strongly convex functions to the case in which the individual functions are composites of strongly convex functions with a linear transform. Under further assumptions, we established the global convergence of the algorithm.

It should be admitted that although some problems arising from applications such as traffic assignment fall into our analysis, the problems considered here are too special. Thus, it is far away to solve the open problem of convergence of the ADMM with more than three blocks.

Acknowledgments

Xingju Cai was supported by the National Natural Science Foundation of China (NSFC) Grants nos. 11071122 and 11171159 and by the Doctoral Fund of Ministry of Education of China no. 20103207110002.

References

[1] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers and Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[2] D. Gabay, "Applications of the method of multipliers to variational inequalities," in *Augmented Lagrangian Methods: Applications to Numerical Solution of Boundary-Value Problems*, M. Fortin and R. Glowinski, Eds., pp. 299–331, North-Holland Publisher, Amsterdam, The Netherland, 1983.

[3] M. Tao and X. Yuan, "Recovering low-rank and sparse components of matrices from incomplete and noisy observations," *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 57–81, 2011.

[4] M. K. Ng, P. Weiss, and X. Yuan, "Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2710–2736, 2010.

[5] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.

[6] Z. Wen, D. Goldfarb, and W. Yin, "Alternating direction augmented Lagrangian methods for semidefinite programming," *Mathematical Programming Computation*, vol. 2, no. 3-4, pp. 203–230, 2010.

[7] D. Han and X. Yuan, "A note on the alternating direction method of multipliers," *Journal of Optimization Theory and Applications*, vol. 155, pp. 227–238, 2012.

[8] D. R. Han, X. M. Yuan, W. X. Zhang, and X. J. Cai, "An ADM-based splitting method for separable convex programming," *Computational Optimization and Applications*, vol. 54, pp. 343–369, 2013.

[9] B. S. He, M. Tao, and X. M. Yuan, "Alternating direction method with Gaussian back substitution for separable convex programming," *SIAM Journal on Optimization*, vol. 22, pp. 313–340, 2012.

[10] B. S. He, M. Tao, M. H. Xu, and X. M. Yuan, "Alternating directions based contraction method for generally separable linearly constrained convex programming problems," *Optimization*, vol. 62, pp. 573–596, 2013.

[11] B. S. He, M. Tao, and X. M. Yuan, "A splitting method for separable convex programming," *IMA Journal of Numerical Analysis*. In press.

[12] D. Han and H. K. Lo, "Solving non-additive traffic assignment problems: a descent method for co-coercive variational inequalities," *European Journal of Operational Research*, vol. 159, no. 3, pp. 529–544, 2012.

[13] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementary Problems. Volume I and II*, Springer, New York, NY, USA, 2003.