Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 183961, 7 pages
http://dx.doi.org/10.1155/2013/183961

Research Article

On the Convergence Analysis of the Alternating Direction Method of Multipliers with Three Blocks

Caihua Chen,[1] Yuan Shen,[2] and Yanfei You[3]

[1] *International Center of Management Science and Engineering, School of Management and Engineering, Nanjing University, Nanjing 210093, China*
[2] *School of Applied Mathematics, Nanjing University of Finance & Economics, Nanjing 210023, China*
[3] *Department of Mathematics, Nanjing University, Nanjing 210093, China*

Correspondence should be addressed to Caihua Chen; cchenhuayx@gmail.com

Received 4 July 2013; Accepted 5 September 2013

Academic Editor: Xu Minghua

We consider a class of linearly constrained separable convex programming problems whose objective functions are the sum of three convex functions without coupled variables. For those problems, Han and Yuan (2012) have shown that the sequence generated by the alternating direction method of multipliers (ADMM) with three blocks converges globally to their KKT points under some technical conditions. In this paper, a new proof of this result is found under new conditions which are much weaker than Han and Yuan's assumptions. Moreover, in order to accelerate the ADMM with three blocks, we also propose a relaxed ADMM involving an additional computation of optimal step size and establish its global convergence under mild conditions.

1. Introduction

In various fields of applied mathematics and engineering, many problems can be equivalently formulated as a separable convex optimization problem with two blocks; that is, given two closed convex functions $f_i : \mathfrak{R}^{n_i} \to \mathfrak{R} \cup \{+\infty\}$, $i = 1, 2$, to find a solution pair (x_1^*, x_2^*) of the following problem:

$$\begin{aligned} \min \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & A_1 x_1 + A_2 x_2 = b, \end{aligned} \tag{1}$$

where A_i is a matrix in $\mathfrak{R}^{p\times n_i}$, $i = 1, 2$, and b is a vector in $\mathfrak{R}^p$. The classical alternating direction method of multipliers (ADMM) [1, 2] applied to problem (1) yields the following scheme:

$$\begin{aligned} x_1^{k+1} &= \arg\min_{x_1\in\mathfrak{R}^{n_1}} f_1(x_1) - \left\langle A_1^T\lambda^k, x_1\right\rangle \\ &\quad + \frac{\beta}{2}\left\| A_1x_1 + A_2x_2^k - b\right\|^2, \\ x_2^{k+1} &= \arg\min_{x_2\in\mathfrak{R}^{n_2}} f_2(x_2) - \left\langle A_2^T\lambda^k, x_2\right\rangle \\ &\quad + \frac{\beta}{2}\left\| A_1x_1^{k+1} + A_2x_2 - b\right\|^2, \\ \lambda^{k+1} &= \lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} - b\right), \end{aligned} \tag{2}$$

where λ^k is a Lagrangian multiplier and $\beta > 0$ is a penalty parameter. Possibly due to its simplicity and effectiveness, the ADMM with two blocks has received continuous attention both in theoretical and application domains. We refer to [3–8] for theoretical results on ADMM with two blocks and [9–13] for its efficient applications in high-dimensional statistics, compressive sensing, finance, image processing, and engineering, to name just a few.

In this paper, we concentrate on the linearly constrained convex programming problem with three blocks:

$$\begin{aligned} \min \quad & f_1(x_1) + f_2(x_2) + f_3(x_3) \\ \text{s.t.} \quad & A_1 x_1 + A_2 x_2 + A_3 x_3 = b, \end{aligned} \tag{3}$$

where $f_3 : \mathfrak{R}^{n_3} \rightarrow \mathfrak{R} \cup \{+\infty\}$ is a closed convex function and A_3 is a matrix in $\mathfrak{R}^{p\times n_3}$. For solving (3), a nature idea is to extend the ADMM with two blocks to the ADMM with three blocks in which the next iteration $(x_2^{k+1}, x_3^{k+1}, \lambda^{k+1})$ is updated by

$$\left(x_2^{k+1}, x_3^{k+1}, \lambda^{k+1}\right) := \left(\tilde{x}_2^k, \tilde{x}_3^k, \tilde{\lambda}^k\right), \tag{4}$$

where

$$\begin{aligned}
\tilde{x}_1^k &= \arg\min_{x_1\in\mathfrak{R}^{n_1}} f_1(x_1) - \langle A_1^T\lambda^k, x_1\rangle \\
&\quad + \frac{\beta}{2}\left\|A_1x_1 + A_2x_2^k + A_3x_3^k - b\right\|^2, \\
\tilde{x}_2^k &= \arg\min_{x_2\in\mathfrak{R}^{n_2}} f_2(x_2) - \left\langle A_2^T\lambda^k, x_2\right\rangle \\
&\quad + \frac{\beta}{2}\left\|A_1\tilde{x}_1^k + A_2x_2 + A_3x^k - b\right\|^2, \\
\tilde{x}_3^k &= \arg\min_{x_3\in\mathfrak{R}^{n_3}} f_3(x_3) - \left\langle A_3^T\lambda^k, x_3\right\rangle \\
&\quad + \frac{\beta}{2}\left\|A_1\tilde{x}_1^k + A_2\tilde{x}_2^k + A_3x_3 - b\right\|^2, \\
\tilde{\lambda}^k &= \lambda^k - \beta\left(A_1\tilde{x}_1^k + A_2\tilde{x}_2^k + A_3\tilde{x}_3^k - b\right).
\end{aligned} \tag{5}$$

Similar to the ADMM with two blocks, the ADMM with three blocks has found numerous applications in a broad spectrum of areas, such as doubly nonnegative cone programming [14], high-dimensional statistics [15, 16], imaging science [17], and engineering [18]. Even though its numerical efficiency is clearly seen from those applications, the theoretical treatment of ADMM with three blocks is challenging and the convergence of the ADMM is still open given only the convex assumptions of the objective function. To alleviate this difficulty, the authors of [19, 20] proposed prediction-correction type methods to solve the general separable convex programming; however, numerical results show that the direct ADMM outperforms its variants substantially. Therefore, it is of great significance to investigate the theoretical performance of the ADMM with three blocks even only to provide sufficient conditions to guarantee the convergence. To the best of our knowledge, there exist only two works aiming to attack the convergence problem of the direct ADMM with three blocks. By using an error bound analysis method, Hong and Luo [21] proved the linear convergence of the ADMM with m blocks for sufficiently small β subject to some technical conditions. However, the sufficiently small requirement on β makes the algorithm difficult to implement. In [22], Han and Yuan employed a contractive analysis method to establish the convergence of ADMM under the strongly convex assumptions of f_i and the parameter β less than a threshold depending on all the strongly convex moduli. In this paper, we firstly prove the convergence of ADMM with three blocks under two conditions weaker than those of [22]. In our conditions, the threshold on the parameter β only relies on the strongly convex moduli of f_2 and f_3, and furthermore f_1 is not necessarily strongly convex in one of our conditions. Also, the restricted range of β in this paper is shown to be at least three times as big as that of [22].

In order to accelerate the ADMM with three blocks, we also propose a relaxed ADMM with three blocks which involves an additional computation of optimal step size. Specifically, with the triple $(x_2^k, x_3^k, \lambda^k)$, we first generate a predictor $(\tilde{x}_2^k, \tilde{x}_3^k, \tilde{\lambda}^k)$ according to (5) and then obtain $(x_2^{k+1}, x_3^{k+1}, \lambda^{k+1})$ in the next iteration by

$$\begin{aligned}
x_2^{k+1} &= x_2^k - \gamma\alpha_k^*\left(x_2^k - \tilde{x}_2^k\right), \\
x_3^{k+1} &= x_3^k - \gamma\alpha_k^*\left(x_3^k - \tilde{x}_3^k\right), \\
\lambda^{k+1} &= \lambda^k - \gamma\alpha_k^*\left(\lambda^k - \tilde{\lambda}^k\right),
\end{aligned} \tag{6}$$

where $\gamma \in (0, 2)$ and α_k^* is special step size defined in (43). The convergence of the relaxed ADMM is also established under mild conditions. We should mention that it is possible to modify the analyses given in this paper to be problems with more than three blocks of separability. But this is not the focus of this paper.

The remaining parts of this paper are organized as follows. In Section 2, we list some preliminaries on the strongly convex function, subdifferential, and the ADMM with three blocks. In Section 3, we first show the contractive property of the distance between the sequence generated by ADMM with three blocks and the solution set and then prove the convergence of ADMM under certain conditions. In Section 4, we extend the direct ADMM with three blocks to the relaxed ADMM with an optimal step size and establish its convergence under suitable conditions. We conclude our paper in Section 5.

Notation. For any positive integer m, let I_m be the $m \times m$ identity matrix. We use $\|\cdot\|$ and $\|\cdot\|_2$ to denote the vector Euclidean norm and the spectral norm of matrices (defined as the maximum singular value of matrices). For any symmetric matrix $S \in \mathfrak{R}^{n\times n}$, we write $\|x\|_S^2 = x^TSx$ for any $x \in \mathfrak{R}^n$. G and M are two $(n_2 + n_3 + p) \times (n_2 + n_3 + p)$ matrices defined by

$$\begin{aligned}
G &:= \begin{pmatrix} \beta A_2^TA_2 & 0 & 0 \\ 0 & \beta A_3^TA_3 & 0 \\ 0 & 0 & \dfrac{I}{\beta} \end{pmatrix}, \\
M &:= \begin{pmatrix} 2\beta A_2^TA_2 & 0 & 0 \\ 0 & \beta A_3^TA_3 & 0 \\ 0 & 0 & \dfrac{I}{\beta} \end{pmatrix},
\end{aligned} \tag{7}$$

respectively. For given $x_1 \in \mathfrak{R}^{n_1}$, $x_2 \in \mathfrak{R}^{n_2}$, $x_3 \in \mathfrak{R}^{n_3}$, and $\lambda \in \mathfrak{R}^p$, we frequently use u and v to denote the joint vectors of x_2, x_3, λ and x_1, x_2, x_3, λ, respectively; that is,

$$u = \left[x_2^T, x_3^T, \lambda^T\right]^T, \qquad v = \left[x_1^T, x_2^T, x_3^T, \lambda^T\right]^T, \tag{8}$$

while $\tilde{u}$ and $\tilde{v}$ are the joint vectors corresponding to $\tilde{x}_2$, $\tilde{x}_3$, $\tilde{\lambda}$ and $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$, $\tilde{\lambda}$.

2. Preliminaries

Throughout this paper, we assume f_i, $i = 1, 2, 3$, are strongly convex functions with modulus $\mu_i \geq 0$; that is

$$f_i\left((1-\alpha)z + \alpha z'\right) \leq (1-\alpha) f_i(z) + \alpha f_i\left(z'\right) - \frac{1}{2}\mu_i \alpha (1-\alpha) \left\|z - z'\right\|^2, \quad \forall z,\ z' \in \mathfrak{R}^{n_i}, \tag{9}$$

for each i. Note that f_i is a strongly convex function with modulus 0 being equivalent to the convexity of f_i. Let x be a point of $\mathrm{dom}(f_i)$; the subdifferential of f_i at x is defined by

$$\partial f_i(x) := \left\{x^* \mid f(z) \geq f(x) + \langle x^*, z - x\rangle, \forall z\right\}. \tag{10}$$

From Proposition 6 in [23], we know that, for each i, ∂f_i is strongly monotone with modulus μ_i which means

$$\langle z_1 - z_2, x_1 - x_2\rangle \geq \mu_i \|z_1 - z_2\|^2 \geq 0, \quad \forall x_1, x_2, z_1 \in \partial f_i(x_1),\ z_2 \in \partial f_i(x_2). \tag{11}$$

The next lemma introduced in [22] plays a key role in the convergence analysis of the ADMM and the relaxed ADMM with three blocks.

Lemma 1. *Let $(x_1^*, x_2^*, x_3^*, \lambda^*)$ be any KKT point of problem (3). Let $\tilde{v}^k$ be generated by (5) from given u^k. Then, one has*

$$\begin{aligned}\left\langle \tilde{u}^k - u^*, G\left(u^k - \tilde{u}^k\right)\right\rangle &\geq \sum_{i=1}^{3} \mu_i \|\tilde{x}_i^k - x_i^*\|^2 + \left\langle \lambda^k - \tilde{\lambda}^k, \sum_{i=2}^{3} A_i\left(x_i^k - \tilde{x}_i^k\right)\right\rangle \\ &\quad + \beta\left\langle A_3\left(\tilde{x}_3^k - x_3^*\right), A_2\left(\tilde{x}_2^k - x_2^k\right)\right\rangle.\end{aligned} \tag{12}$$

3. The ADMM with Three Blocks

In this section, we first investigate the contractive property of the distance between the sequence generated by ADMM with three blocks and the solution set under the condition that $0 < \beta \leq \min\{\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2\}$.

Lemma 2. *Let $v^* = (x_1^*, x_2^*, x_3^*, \lambda^*)$ be a KKT point of problem (3) and let the sequence $\{v^k = (x_1^k, x_2^k, x_3^k, \lambda^k)\}$ be generated by the ADMM with three blocks. Then, it holds that*

$$\begin{aligned}\left\|u^{k+1} - u^*\right\|_M^2 &\leq \left\|u^k - u^*\right\|_M^2 - \beta\left\|A_3(x_3^{k+1} - x_3^k)\right\|^2 \\ &\quad - \beta\left\|A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\right\|^2 \\ &\quad - 2\mu_1\left\|x_1^{k+1} - x_1^*\right\|^2 \\ &\quad - 2\left\|x_2^{k+1} - x_2^*\right\|_{\mu_2 I_{n_2} - \beta A_2^T A_2}^2 \\ &\quad - 2\left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2.\end{aligned} \tag{13}$$

Proof. Since x_3^j minimizes $f_3(\cdot) - \langle A_3^T \lambda^j, \cdot\rangle$, we deduce from the first order optimality condition that

$$A_3^T \lambda^j \in \partial f_3\left(x_3^j\right), \quad j = 0, 1, \ldots, k. \tag{14}$$

By (14) and the monotonicity of $\partial f_3(\cdot)$ (11), it is easily seen that

$$\left\langle x_3^k - x_3^{k+1}, A_3^T \lambda^k - A_3^T \lambda^{k+1}\right\rangle \geq 0. \tag{15}$$

Then for each k,

$$\begin{aligned}\left\langle u^{k+1} - u^*, G\left(u^k - u^{k+1}\right)\right\rangle &\geq \sum_{i=1}^{3} \mu_i \left\|x_i^{k+1} - x_i^*\right\|^2 + \left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_i^k - x_2^{k+1}\right)\right\rangle \\ &\quad + \beta\left\langle A_3\left(x_3^{k+1} - x_3^*\right), A_2\left(x_2^{k+1} - x_2^k\right)\right\rangle \\ &\geq \sum_{i=1}^{2} \mu_i \left\|x_i^{k+1} - x_i^*\right\|^2 + \left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2 \\ &\quad + \left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_2^k - x_2^{k+1}\right)\right\rangle \\ &\quad - \frac{\beta}{4}\left\|A_2\left(x_2^{k+1} - x_2^k\right)\right\|^2,\end{aligned} \tag{16}$$

where the last "$\geq$" follows from the elementary inequality

$$\langle x, y\rangle \geq -\|x\|^2 - \frac{1}{4}\|y\|^2. \tag{17}$$

Since

$$\left\|A_3\left(x_3^{k+1} - x_3^k\right)\right\|^2 \leq 2\left\|A_3(x_3^{k+1} - x_3^*)\right\|^2 + 2\left\|A_3\left(x_3^k - x_3^*\right)\right\|^2, \tag{18}$$

by direct computations, we further obtain that

$$\begin{aligned}\left\|u^k - u^*\right\|_G^2 &\geq \left\|u^{k+1} - u^*\right\|_G^2 \\ &\quad + \left\|u^{k+1} - u^k\right\|_G^2 + 2\mu_1\left\|x_1^{k+1} - x_1^*\right\|^2 \\ &\quad + 2\left\|x_2^{k+1} - x_2^*\right\|_{\mu_2 I_{n_2} - (\beta/2) A_2^T A_2}^2 \\ &\quad + 2\left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2 \\ &\quad + 2\left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_2^k - x_2^{k+1}\right)\right\rangle \\ &\quad - \beta\left\|A_2\left(x_2^k - x_2^*\right)\right\|^2,\end{aligned} \tag{19}$$

which, together with $G = M - \begin{pmatrix} \beta A_2^T A_2 & & \\ & 0 & \\ & & 0 \end{pmatrix}$, implies

$$\begin{aligned} \left\|u^k - u^*\right\|_M^2 \geq \left\|u^{k+1} - u^k\right\|_G^2 & \\ + \left\|u^{k+1} - u^*\right\|_M^2 + 2\mu_1 \left\|x_1^{k+1} - x_1^*\right\|^2 & \\ + 2\left\|x_2^{k+1} - x_2^*\right\|_{\mu_2 I_{n_2} - \beta A_2^T A_2}^2 & \\ + 2\left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2 & \\ + 2\left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_2^k - x_2^{k+1}\right)\right\rangle . & \end{aligned} \tag{20}$$

Note that

$$\begin{aligned} \left\|x_2^k - x_2^{k+1}\right\|_{\beta A_2^T A_2}^2 + 2\left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_2^k - x_2^{k+1}\right)\right\rangle & \\ + \frac{1}{\beta}\left\|\lambda^k - \lambda^{k+1}\right\|^2 & \\ = \beta\left\|A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\right\|^2. & \end{aligned} \tag{21}$$

We complete the proof of this lemma. □

With the above preparation, we are ready to prove the convergence of the ADMM with three blocks for solving (3) given the following conditions.

Theorem 3. *Let* $\{v^k = (x_1^k, x_2^k, x_3^k, \lambda^k)\}$ *be the sequence generated by the ADMM with three blocks. Then* $\{v^k\}$ *converges to a KKT point of problem* (3) *if either of the following conditions holds:*

(i) $\mu_1 > 0$ *and* $0 < \beta \leq \min\{\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2\}$;

(ii) A_1 *is of full column rank,* $0 < \beta < \mu_2\|A_2\|_2^2$, *and* $\beta \leq \mu_3\|A_3\|_2^2$.

Proof. By the inequality (13), it follows that the sequence $\{A_2 x_2^k, A_3 x_3^k, \lambda^k\}$ is bounded. Recall that

$$A_1 x_1^{k+1} = \frac{\lambda^k - \lambda^{k+1}}{\beta} - A_2 x_2^{k+1} - A_3 x_3^{k+1} + b. \tag{22}$$

Hence $\{A_1 x_1^k\}$ is also bounded. Moreover, from (13) we see immediately that

$$\begin{aligned} +\infty > \sum_{k=1}^{\infty} \beta\left\|A_3 (x_3^{k+1} - x_3^k)\right\|^2 & \\ + \beta\left\|A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\right\|^2 & \\ + \sum_{k=1}^{\infty} 2\mu_1\left\|x_1^{k+1} - x_1^*\right\|^2 & \\ + 2\left\|x_2^{k+1} - x_2^*\right\|_{\mu_2 I_{n_2} - \beta A_2^T A_2}^2 & \\ + 2\left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2 . & \end{aligned} \tag{23}$$

According to the condition that $0 < \beta \leq \min\{\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2\}$, we know

$$\begin{aligned} & \sum_{k=1}^{\infty} \| A_3 \left(x_3^{k+1} - x_3^k\right) \|^2 < \infty, \\ & \sum_{k=1}^{\infty} \| A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\|^2 < +\infty, \\ & \sum_{k=1}^{\infty} \mu_1 \| x_1^{k+1} - x_1^*\|^2 < +\infty, \\ & \sum_{k=1}^{\infty} \left\|x_2^{k+1} - x_2^*\right\|_{\mu_2 I_{n_2} - \beta A_2^T A_2}^2 < +\infty, \\ & \sum_{k=1}^{\infty} \left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2 < +\infty. \end{aligned} \tag{24}$$

It therefore holds that

$$\begin{aligned} & \lim_{k\to\infty} \left\|A_3\left(x_3^{k+1} - x_3^k\right)\right\|^2 = 0, \\ & \lim_{k\to\infty} \left\|A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\right\|^2 = 0, \end{aligned} \tag{25}$$

$$\begin{aligned} & \lim_{k\to\infty} \mu_1 \left\|x_1^{k+1} - x_1^*\right\|^2 = 0, \\ & \lim_{k\to\infty} \left\|x_2^{k+1} - x_2^*\right\|_{\mu_2 I_{n_2} - \beta A_2^T A_2}^2 = 0, \\ & \lim_{k\to\infty} \left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2 = 0. \end{aligned} \tag{26}$$

Therefore, the sequence $\{\mu_1\|x_1^k\|^2, \|x_2^k\|_{\mu_2 I_{n_2} - \beta A_2^T A_2}^2, \|x_3^k\|_{\mu_3 I_{n_3} - \beta A_2^T A_2}^2\}$ is bounded, which, together with the boundedness of $\{A_1 x_1^k, A_2 x_2^k, A_3 x_3^k, \lambda^k\}$, implies that $\{x_2^k, x_3^k, \lambda^k\}$ is bounded, and $\{x_1^k\}$ is bounded given the condition $\mu_1 > 0$ or A_1 is of full column rank. Moreover, since

$$\begin{aligned} \left\|x_3^{k+1} - x_3^k\right\|^2 = \left\|A_3 x_3^{k+1} - A_3 x_3^k\right\|^2 & \\ + \left\|x_3^{k+1} - x_3^k\right\|_{\mu_3 I_{n_3} - A_3^T A_3}^2, & \end{aligned} \tag{27}$$

by the first equality in (25) and the third equality in (26), it holds that

$$\lim_{k\to\infty} \left\|x_3^{k+1} - x_3^k\right\| = 0. \tag{28}$$

We proceed to prove the convergence of ADMM by considering the following two cases.

Case 1 ($\mu_1 > 0$ and $\beta \leq \min(\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2)$). In this case, the sequence $\{x_1^k\}$ converges to x_1^* and then

$$\lim_{k\to\infty} \left\|A_2 x_2^{k+1} - A_2 x_2^k\right\| = 0, \qquad \lim_{k\to\infty} \left\|\lambda^{k+1} - \lambda^k\right\| = 0. \tag{29}$$

By the second equality in (26), we deduce from (29) that

$$\lim_{k\to\infty} \left\|x_2^{k+1} - x_2^k\right\| = 0. \tag{30}$$

Since $\{x_2^k, x_3^k, \lambda^k\}$ is bounded, there exist a triple $(x_2^\infty, x_3^\infty, \lambda^\infty)$ and a subsequence $\{n_k\}$ such that

$$\lim_{k\to\infty} x_2^{n_k} = x_2^\infty, \qquad \lim_{k\to\infty} x_2^{n_k} = x_2^\infty, \qquad \lim_{k\to\infty} \lambda^{n_k} = \lambda^\infty, \tag{31}$$

which by combining (25), (29) with given conditions, implies

$$\begin{gathered} \lim_{k\to\infty} x_2^{n_k+1} = x_2^\infty, \qquad \lim_{k\to\infty} x_2^{n_k+1} = x_2^\infty, \\ \lim_{k\to\infty} \lambda^{n_k+1} = \lambda^\infty. \end{gathered} \tag{32}$$

Note that

$$\begin{gathered} 0 \in \partial f_1\left(x_1^{k+1}\right) - A_1^T\lambda^{k+1} + A_1^T A_2\left(x_2^k - x_2^{k+1}\right) \\ + A_1^T A_3\left(x_3^k - x_3^{k+1}\right), \\ 0 \in \partial f_2\left(x_2^{k+1}\right) - A_2^T\lambda^{k+1} + A_2^T A_3\left(x_3^k - x_3^{k+1}\right), \\ 0 \in \partial f_3\left(x_3^{k+1}\right) - A_3^T\lambda^{k+1}, \\ \lambda^{k+1} = \lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k+1}\right). \end{gathered} \tag{33}$$

Then, by taking the limits on both sides of (33), using (25) and (29), and invoking the upper semicontinuous of $\partial f_1(\cdot)$, $\partial f_2(\cdot)$, and $\partial f_3(\cdot)$ [24], one can immediately write

$$\begin{gathered} 0 \in \partial f_1\left(x^*\right) - A_1^T\lambda^\infty, \\ 0 \in \partial f_2\left(x_2^\infty\right) - A_2^T\lambda^\infty, \\ 0 \in \partial f_3\left(x_3^\infty\right) - A_3^T\lambda^\infty, \\ A_1 x^* + A_2 x_2^\infty + A_3 x_3^\infty = b, \end{gathered} \tag{34}$$

which indicates $(x_1^*, x_2^\infty, x_3^\infty, \lambda^\infty)$ is a KKT point of problem (3). Hence, the inequality (13) is also valid if $(x_1^*, x_2^*, x_3^*, \lambda^*)$ is replaced by $(x_1^*, x_2^\infty, x_3^\infty, \lambda^\infty)$. Then it holds that

$$\begin{aligned} 2\beta&\left\|A_2 x_2^{k+1} - A_2 x_2^\infty\right\|^2 + \beta\left\|A_3 x_3^{k+1} - A_3 x_3^\infty\right\|^2 \\ &+ \frac{1}{\beta}\left\|\lambda^{k+1} - \lambda^\infty\right\|^2 \leq 2\beta\left\|A_2 x_2^k - A_2 x_2^\infty\right\|^2 \\ &+ \beta\left\|A_3 x_3^k - A_3 x_3^\infty\right\|^2 + \frac{1}{\beta}\left\|\lambda^k - \lambda^\infty\right\|^2, \end{aligned} \tag{35}$$

which yields

$$\begin{gathered} \lim_{k\to\infty} \left\|x_2^k - x_2^\infty\right\|_{A_2^T A_2}^2 = 0, \\ \lim_{k\to\infty} \left\|x_3^k - x_3^\infty\right\|_{A_3^T A_3}^2 = 0, \end{gathered} \tag{36}$$

$$\lim_{k\to\infty} \lambda^k = \lambda^\infty. \tag{37}$$

By adding the last two equalities in (26) to (36), we know

$$\lim_{k\to\infty} x_2^k = x_2^\infty, \qquad \lim_{k\to\infty} x_3^k = x_3^\infty. \tag{38}$$

Therefore, we have shown that the whole sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ converges to $(x_1^*, x_2^\infty, x_3^\infty, \lambda^\infty)$ under condition (i) in Theorem 3.

Case 2 (A_1 is of full column rank, $0 < \beta < \mu_2/\|A_2\|_2^2$, and $\beta \leq \mu_3/\|A_3\|_2^2$). In this case, the sequence $\{x_2^k\}$ converges to x_2^* and $\{x_1^k\}$ is bounded. From the second equality in (25) and (28), we have

$$\begin{gathered} \lim_{k\to\infty} \| A_1 x_1^{k+1} - A_1 x_1^k \| = 0, \\ \lim_{k\to\infty} \| \lambda^k - \lambda^{k+1} \| = 0. \end{gathered} \tag{39}$$

Since A_1 is of full column rank, it therefore holds that

$$\lim_{k\to\infty} \| x_1^{k+1} - x_1^k \| = 0. \tag{40}$$

Let $(x_1^\infty, x_3^\infty, \lambda^\infty)$ be a cluster point of the sequence $\{x_1^k, x_3^k, \lambda^k\}$. Following a similar proof in Case 1, we are able to show $(x_1^\infty, x_2^*, x_3^\infty, \lambda^\infty)$ is a KKT point of problem (3) and the whole sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ converges to this point. □

Remark 4 (see [22]). the authors proved the convergence of the ADMM under the conditions that f_1, f_2, and f_3 are strongly convex and $0 < \beta < \min_{1\leq i\leq 3}\{\mu_i/3\|A_i\|_2^2\}$. Our result improves the upper bound $\min_{1\leq i\leq 3}\{\mu_i/3\|A_i\|_2^2\}$ by $\min\{\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2\}$. Moreover, in our condition (ii), the strongly convexity assumption is only imposed on f_2 and f_3 while f_1 is not necessarily strongly convex with positive modulus.

4. The Relaxed ADMM with Three Blocks

For the ADMM with two blocks, Ye and Yuan [25] developed a variant of alternating direction method with an optimal step size. Numerical results demonstrated that an additional computation on the optimal step size would improve the efficiency of the new variant of ADMM. In this section, by adopting the essential idea of Ye and Yuan [25], we propose a relaxed ADMM with three blocks to accelerate the ADMM via an optimal step size. For notational simplicity, we write

$$\Phi\left(u^k,\tilde{u}^k\right) := \frac{3\beta}{4}\left\|A_2(x_2^k-\tilde{x}_2^k)\right\|^2 + \beta\left\|A_3(x_3-\tilde{x}_3^k)\right\|^2 + \frac{1}{\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2 + \left\langle \lambda^k-\tilde{\lambda}^k, A_2\left(x_2^k-\tilde{x}_2^k\right)+A_3\left(x_3^k-\tilde{x}_3^k\right)\right\rangle. \tag{41}$$

With $u^k=(x_2^k,x_3^k,\lambda^k)$, the new iterate of extended ADMM is produced by

$$u^{k+1}=u^k-\gamma\alpha^*\left(u^k-\tilde{u}^k\right), \quad \gamma\in(0,2), \tag{42}$$

where $\tilde{u}^k$ is the solution of (5) and α^* is defined by

$$\alpha^* := \frac{\Phi\left(u^k,\tilde{u}^k\right)}{\|u^k-\tilde{u}^k\|_G^2}. \tag{43}$$

Lemma 5. *Let the sequence $\{u^k\}$ be generated by the relaxed ADMM with three blocks. Then, if $0<\beta\le\mu_3/\|A_3\|_2^2$, the following statements are valid:*

(i) $\Phi(u^k,\tilde{u}^k)\ge(1/6)\|u^k-u^{k+1}\|_G^2$ *and thus* $\alpha^*\ge 1/6$;

(ii) $\|u^{k+1}-u^*\|_G^2 \le \|u^k-u^*\|_G^2-(1/36)\gamma(2-\gamma)\|u^k-\tilde{u}^k\|_G^2-(1/3)\gamma\mu_1\|\tilde{x}_1^k-x_1^*\|^2-(1/3)\gamma\mu_2\|\tilde{x}_2^k-x_2^*\|^2-(1/3)\gamma\|\tilde{x}_3^k-x_3^*\|^2_{\mu_3I-\beta A_3^TA_3}$.

Proof. By direct computations to $\Phi(u^k,\tilde{u}^k)$, we know that

$$\begin{aligned}\Phi\left(u^k,\tilde{u}^k\right) &= \frac{3\beta}{4}\left\|A_2\left(x_2^k-\tilde{x}_2^k\right)\right\|^2+\beta\left\|A_3(x_3-\tilde{x}_3^k)\right\|^2 + \frac{1}{\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2 + \left\langle \lambda^k-\tilde{\lambda}^k, A_2\left(x_2^k-\tilde{x}_2^k\right)+A_3\left(x_3^k-\tilde{x}_3^k\right)\right\rangle \\ &\ge \frac{3\beta}{4}\left\|A_2(x_2^k-\tilde{x}_2^k)\right\|^2+\beta\left\|A_3(x_3-\tilde{x}_3^k)\right\|^2 + \frac{1}{\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2-\frac{\beta}{2}\left\|A_2\left(x_2^k-\tilde{x}_2^k\right)\right\|^2 \\ &\quad -\frac{1}{2\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2-\frac{3\beta}{4}\left\|A_3(x_3^k-\tilde{x}_3^k)\right\|^2 - \frac{1}{3\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2 = \frac{\beta}{4}\left\|A_2(x_2^k-\tilde{x}_2^k)\right\|^2 \\ &\quad + \frac{\beta}{4}\left\|A_3(x_3^k-\tilde{x}_3^k)\right\|^2+\frac{1}{6\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2,\end{aligned} \tag{44}$$

where the second inequality follows Cauchy inequality. It therefore holds that

$$\Phi\left(u^k,\tilde{u}^k\right)\ge\frac{1}{6}\left\|u^k-\tilde{u}^k\right\|_G^2, \tag{45}$$

which completes the proof of the first part. By Lemma 1 and the elementary inequality (17), it can be easily verified that

$$\begin{aligned}&\left\langle u^k-u^*, G\left(u^k-\tilde{u}^k\right)\right\rangle \\ &\quad\ge \Phi\left(u^k,\tilde{u}^k\right)+\mu_1\left\|\tilde{x}_1^k-x_1^*\right\|^2+\mu_2\left\|\tilde{x}_2^k-x_2^*\right\|^2 + \left\|\tilde{x}_3^k-x_3^*\right\|^2_{\mu_3I_{n_3}-\beta A_3^TA_3}\end{aligned} \tag{46}$$

and then

$$\begin{aligned}\left\|u^{k+1}-u^*\right\|_G^2 &= \left\|u^k-u^*-\gamma\alpha^*\left(u^k-\tilde{u}^k\right)\right\|_G^2 \\ &\le \left\|u^k-u^*\right\|_G^2-\gamma(2-\gamma)\left(\alpha^*\right)^2 \times\left\|u^k-\tilde{u}^k\right\|_G^2-2\gamma\alpha^*\mu_1\left\|\tilde{x}_1^k-x_1^*\right\|^2 \\ &\quad -2\mu_2\gamma\alpha^*\left\|\tilde{x}_2^k-x_2^*\right\|^2 -2\gamma\alpha^*\left\|\tilde{x}_3^k-x_3^*\right\|^2_{\mu_3I_{n_3}-\beta A_3^TA_3}.\end{aligned} \tag{47}$$

This, together with the fact that $\alpha^*\ge1/6$, completes the proof. □

Based on the above inequality, we are able to prove the following convergence result of the relaxed ADMM with three blocks. Since the proof is in line with that of Theorem 3, we omit it.

Theorem 6. *Let $\{v^k=(x_1^k,x_2^k,x_3^k,\lambda^k)\}$ be the sequence generated by the relaxed ADMM. Then $\{v^k\}$ converges to a KKT point of problem (3) under the conditions that $0<\beta\le\mu_3/\|A_3\|_2^2$ and A_1, A_2, and A_3 are of full column rank.*

5. Conclusion Remarks

In this paper, we take a step to investigate the ADMM for separable convex programming problems with three blocks. Based on the contractive analysis of the distance between the sequence and the solution set, we establish theoretical results to guarantee the global convergence of ADMM with three blocks under weaker conditions than those employed in [22]. By adopting the essential idea of [25], we also present a relaxed ADMM with an optimal step size to accelerate the ADMM and prove its convergence under mild assumptions.

Acknowledgment

The first author is supported by the Natural Science Foundation of Jiangsu Province and the National Natural Science Foundation of China under Project no. 71271112. The second author is supported by university natural science research fund of jiangsu province under grant no. 13KJD110002.

References

[1] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximations," *Computational Mathematics with Applications*, vol. 2, pp. 17–40, 1976.

[2] R. Glowinski and A. Marrocco, "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires," vol. 9, no. R-2, pp. 41–76, 1975.

[3] J. Eckstein and D. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 3, pp. 293–318, 1992.

[4] D. Gabay, "Chapter ix applications of the method of multipliers to variational inequalities," *Studies in Mathematics and Its Applications*, vol. 15, pp. 299–331, 1983.

[5] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer, New York, NY, USA, 1984.

[6] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, vol. 9, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 1989.

[7] B. He, L. Liao, D. Han, and H. Yang, "A new inexact alternating directions method for monotone variational inequalities," *Mathematical Programming*, vol. 92, no. 1, pp. 103–118, 2002.

[8] B. He and X. Yuan, "On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method," *SIAM Journal on Numerical Analysis*, vol. 50, no. 2, pp. 700–709, 2012.

[9] C. Chen, B. He, and X. Yuan, "Matrix completion via an alternating direction method," *IMA Journal of Numerical Analysis*, vol. 32, no. 1, pp. 227–245, 2012.

[10] M. Fazel, T. K. Pong, D. F. Sun, and P. Tseng, "Hankel matrix rank minimization with applications to system identification and realization," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 946–977, 2012.

[11] B. He, M. Xu, and X. Yuan, "Solving large-scale least squares semidefinite programming by alternating direction methods," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 1, pp. 136–152, 2011.

[12] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for tvl1-l2 signal reconstruction from partial fourier data," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 288–297, 2010.

[13] J. Yang and Y. Zhang, "Alternating direction method algorithms for l1-problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.

[14] 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.

[15] 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.

[16] J. Yang, D. Sun, and K. Toh, "A proximal point algorithm for logdeterminant optimization with group Lasso regularization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 857–293, 2013.

[17] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 763–770, 2010.

[18] K. Mohan, P. London, M. Fazel, D. Witten, and S. I. Lee, "Node-based learning of multiple gaussian graphical models," http://arxiv.org/abs/1303.5145.

[19] B. 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, no. 4, pp. 573–596, 2013.

[20] B. He, M. Tao, and X. Yuan, "Alternating direction method with Gaussian back substitution for separable convex programming," *SIAM Journal on Optimization*, vol. 22, no. 2, pp. 313–340, 2012.

[21] M. Hong and Z. Luo, "On the linear convergence of the alternating direction Method of multipliers," http://arxiv.org/abs/1208.3922.

[22] D. Han and X. Yuan, "A note on the alternating direction method of multipliers," *Journal of Optimization Theory and Applications*, vol. 155, no. 1, pp. 227–238, 2012.

[23] R. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[24] R. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, no. 28, Princeton University Press, Princeton, NJ, USA, 1970.

[25] C. Ye and X.-M. Yuan, "A descent method for structured monotone variational inequalities," *Optimization Methods & Software*, vol. 22, no. 2, pp. 329–338, 2007.