THE CONVERGENCE RATE ANALYSIS OF THE SYMMETRIC ADMM FOR THE NONCONVEX SEPARABLE OPTIMIZATION PROBLEMS

ZHENG JIA
Department of Information and Computing Science, School of Mathematics and Statistics
Nanjing University of Information Science and Technology, Nanjing 210044, China

XUE GAO AND XINGJU CAI∗
School of Mathematical Sciences, Key Laboratory for NSLSCS of Jiangsu Province
Nanjing Normal University, Nanjing 210023, China

DEREN HAN
LMIB of the Ministry of Education, School of Mathematical Sciences
Beihang University, Beijing 100191, China

(Communicated by Jinyan Fan)

Abstract. The symmetric alternating direction method of multipliers is an efficient algorithm, which updates the Lagrange multiplier twice at each iteration and the variables are treated in a symmetric manner. Considering that the convergence range of the parameters plays an important role in the implementation of the algorithm. In this paper, we analyze the convergence rate of the symmetric ADMM with a more relaxed parameter range for solving the two block nonconvex separable optimization problem under the assumption that the generated sequence is bounded. Two cases are considered. In the first case, both components of the objective function are nonconvex, we prove the convergence of the augmented Lagrangian function sequence, and establish the $O(1/\sqrt{k})$ worst-case complexity measured by the difference of two consecutive iterations. In the second case, one component of the objective function is convex and the error bound condition is assumed, then we can prove that the iterative sequence converges locally to a KKT point in a R-linear rate; and an auxiliary sequence converges in a Q-linear rate. Furthermore, a practical inexact symmetric ADMM with relative error criteria is proposed, and the associated convergence analysis is established under the same conditions.

1. Introduction. In this paper, we consider the following optimization problem

$$\min \ f(x) + g(y)$$
$$\text{s.t.} \ \ Ax - By - c = 0, \tag{1}$$

where $f : \mathcal{R}^n \to \mathcal{R} \cup \{+\infty\}$ is a proper lower semi-continuous function; $g : \mathcal{R}^m \to \mathcal{R}$ is a smooth function and $\nabla g$ is Lipschitz continuous with modulus $L > 0$; $A \in \mathcal{R}^{m \times n}$

2010 Mathematics Subject Classification. 90C26, 49M27, 90C33.

Key words and phrases. Symmetric alternating direction method of multipliers, nonconvex minimization, inexact, relative error criteria, sublinear convergence, linear convergence.

∗ Corresponding author: Xingju Cai.
is a matrix, $B \in \mathbb{R}^{m \times m}$ is an invertible matrix and $c \in \mathbb{R}^m$ is a vector. This nonconvex program captures numerous applications [2, 11, 12, 18, 21].

If we apply the Peaceman-Rachford splitting method (PRSM) to the dual of (1), we can obtain the following scheme [3, 13, 16]

\[
\begin{aligned}
  x^{k+1} &= \arg \min_x L_{\beta}(x, y^k, \lambda^k), \\
  \lambda^{k+\frac{1}{2}} &= \lambda^k - \beta(Ax^{k+1} - By^k - c), \\
  y^{k+1} &= \arg \min_y L_{\beta}(x^{k+1}, y, \lambda^{k+\frac{1}{2}}), \\
  \lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \beta(Ax^{k+1} - By^{k+1} - c),
\end{aligned}
\]  

(2)

where $L_{\beta}(\cdot)$ denotes the augmented Lagrangian function of (1) defined by

\[
L_{\beta}(x, y, \lambda) = f(x) + g(y) - \lambda^T(Ax - By - c) + \frac{\beta}{2} \|Ax - By - c\|^2, 
\]  

(3)

$\lambda \in \mathbb{R}^m$ is the Lagrange multiplier associated with the linear constraints and $\beta > 0$ is a penalty parameter. This scheme can be considered as the symmetric version of ADMM, in the sense that the Lagrange multiplier is updated twice at each iteration and the variables are treated in a symmetric manner. However, it was commented in [6] that the sequence generated by the symmetric ADMM (2) is not necessarily convergent even if both $f$ and $g$ are convex functions. Some efforts are devoted to slightly modify the symmetric ADMM to guarantee its convergence.

For the case that both $f$ and $g$ are convex in (1), [6] suggested updating the Lagrange multiplier more conservatively and obtained the symmetric ADMM scheme with smaller step size:

\[
\begin{aligned}
  x^{k+1} &= \arg \min_x L_{\beta}(x, y^k, \lambda^k), \\
  \lambda^{k+\frac{1}{2}} &= \lambda^k - \alpha \beta(Ax^{k+1} - By^k - c), \\
  y^{k+1} &= \arg \min_y L_{\beta}(x^{k+1}, y, \lambda^{k+\frac{1}{2}}), \\
  \lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \alpha \beta(Ax^{k+1} - By^{k+1} - c),
\end{aligned}
\]  

(4)

where the parameter $\alpha \in (0, 1)$. To further relax the requirement on the parameters such that the algorithm can achieve better performance, [5] took different step sizes to update the Lagrange multiplier, i.e., they assigned two different step sizes in (2),

\[
\begin{aligned}
  x^{k+1} &= \arg \min_x L_{\beta}(x, y^k, \lambda^k), \\
  \lambda^{k+\frac{1}{2}} &= \lambda^k - \alpha \beta(Ax^{k+1} - By^k - c), \\
  y^{k+1} &= \arg \min_y L_{\beta}(x^{k+1}, y, \lambda^{k+\frac{1}{2}}), \\
  \lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \theta \beta(Ax^{k+1} - By^{k+1} - c),
\end{aligned}
\]  

(5)

where $\alpha$ and $\theta$ are two constants belonging to the following domain

\[
D = \left\{ (\alpha, \theta) \mid \alpha \in (0, 1), \theta \in \left(0, \frac{1 - \alpha + \sqrt{(1 + \alpha)^2 + 4(1 - \alpha^2)}}{2}\right) \right\}. 
\]  

(6)

The $O(1/k)$ sublinear convergence rate in the ergodic and nonergodic sense were given in [5] when $f$ and $g$ are convex functions. [7] also considered the scheme (5), and the constants $\alpha$ and $\theta$ are restricted in the following domain

\[
\mathcal{D} = \left\{ (\alpha, \theta) \mid \theta \in \left(0, \frac{1 + \sqrt{5}}{2}\right), \alpha \in (-1, 1), \alpha + \theta > 0, |\alpha| < 1 + \theta + \theta^2 \right\}. 
\]  

(7)
For the case where $f$ and/or $g$ are nonconvex in (1), [19] recently proved that each bounded sequence generated by (5) converges to a critical point of the augmented Lagrangian function when $\theta = 1$ and $\alpha \in (-1,1)$. As far as we know, no other paper discussed the range of the parameters $\alpha$ and $\theta$ under the nonconvex case.

It is worth noting that the range of parameters $\alpha$ and $\theta$ in (5) play important roles in the implementation of the algorithm. The original choice $\alpha = \theta = 1$ is succinct and efficient in most cases, but the convergence is not guaranteed even in the convex case. It is now a common sense that the condition imposed should be as relaxable as possible. In this paper, we propose a new domain for the choice of these two parameters

$$
\tilde{D} = \left\{ (\alpha, \theta) \mid \theta \in (0, 2), \alpha + \theta > 0, 1 > \frac{\theta - 1}{1 + \alpha}, \frac{1}{2} > \frac{\alpha \theta}{\alpha + \theta} + \frac{6\alpha^2(\theta - 1)^2}{(1 + \alpha - |\theta - 1|)^2(\alpha + \theta)} \right\}
$$

We illustrate the domain $\tilde{D}$ in the grey shaded area of Figure 1. It should be pointed out that even for the case that both $f$ and $g$ are convex, the maximum range of the parameter $\theta$ in [5] and [6] is $(0, \frac{1+\sqrt{5}}{2})$ (see (6) and (7)), and here ours is $(0, 2)$. Furthermore, the convergence range of the parameters in [19] is also contained in the domain $\tilde{D}$ as a special case.

Our main purpose is to establish the convergence and the linear convergence rate of the symmetric ADMM (5) with a more relaxed parameter range for the nonconvex optimization problem (1). We consider two cases. In the first case, both components of the objective function are nonconvex. We prove the convergence of the augmented Lagrangian function sequence \( \{L_\beta(x^k, y^k, \lambda^k)\} \), and the $O(1/\sqrt{k})$ iteration complexity measured by the difference of two consecutive iterations. We emphasize that though we can not prove the convergence of the iterative sequence, our results do not rely on any additional conditions such as the KL property of the augmented Lagrangian function [19] (see Theorem 3.7). In the second case, we assume that the nonsmooth function $f$ in (1) is convex and the error bound condition is satisfied, and then we prove that each bounded sequence generated by the symmetric ADMM (5) converges locally to a KKT point of (1) in a R-linear rate (see Theorem 4.4); and an auxiliary sequence converges in a Q-linear rate (see Theorem 4.3).

In applications, it is desirable that we can use approximate solution of $x$ and $y$ minimization problems in (5), and the accuracy criterion should be flexible. We thus propose an inexact version of the symmetric ADMM (5). Specially, the two subproblems in the symmetric ADMM are allowed to be solved inexactly under the relative error criteria, where we just need choose two parameters to control the accuracy. In many cases, one iterative step is enough to obtain a satisfactory solution [8, 9]. Under the same assumptions as those in the second case, we also establish (see Theorem 5.7-5.8) the R-linear and the Q-linear convergence rate of the algorithm.

The rest of this paper is organized as follows. Section 2 presents some basic notations and preliminary materials. In Section 3, under the assumption that the sequence generated by symmetric ADMM (5) is bounded, we establish the convergence of the augmented Lagrangian function sequence \( \{L_\beta(x^k, y^k, \lambda^k)\} \) and the sublinear convergence rate of the generated sequence. And we also give some sufficient conditions to guarantee the boundness of the sequence generated by (5). In Section 4, we prove that the iterative sequence generated by symmetric ADMM (5) converges locally to a KKT point of (1) in a R-linear rate, and an auxiliary sequence converges in a Q-linear rate, when assuming that one component function of the objective function in (1) is convex and the error bound condition is satisfied. In
Section 5, we propose an inexact version of the symmetric ADMM, and establish
the linear convergence rate based on the error bound condition. Finally, we make
some conclusions in Section 6.

2. Preliminaries. In this section, we summarize some notations and preliminaries
to be used for further analysis.

Throughout this paper, we use $\mathbb{R}^n$ to denote the $n$-dimensional Euclidean space,
with its standard inner product denoted by $\langle \cdot, \cdot \rangle$. The Euclidean norm is denoted
by $\| \cdot \|$. For any subset $\Omega \subset \mathbb{R}^n$ and any point $x \in \mathbb{R}^n$, the distance from $x$ to $\Omega$,
denoted by $\text{dist}(x, \Omega)$, is defined as
$$\text{dist}(x, \Omega) = \inf_{y \in \Omega} \| y - x \|.$$ 
When $\Omega$ is closed and convex, we use $P_\Omega(x)$ to denote the unique closest point on
$\Omega$ to $x$. For a matrix $A \in \mathbb{R}^{m \times n}$, we use $A^T$ to denote its transpose. For a given
sequence $\{z^k : k \geq 0\}$, let $\{\Delta z^k\}$ be the sequence defined by
$$\Delta z^k = z^k - z^{k-1}, \quad k \geq 1.$$ 
So $\Delta x^k = x^k - x^{k-1}$, $\Delta y^k = y^k - y^{k-1}$ and $\Delta \lambda^k = \lambda^k - \lambda^{k-1}$, $k \geq 1$. The domain
of an extended-real-valued function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined as
$$\text{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}.$$ 
We say that $f$ is proper, if $\text{dom} f \neq \emptyset$. Such a function is closed if it is lower
semicontinuous.

Definition 2.1. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function.

1. The Fréchet subdifferential, or regular subdifferential, of $f$ at $x \in \text{dom} f$, written as $\hat{\partial} f(x)$, is the set of vectors $x^* \in \mathbb{R}^n$ that satisfy
$$\lim_{y \to x, y \neq x} \inf \frac{f(y) - f(x) - \langle x^*, y - x \rangle}{\| y - x \|} \geq 0.$$ 
When $x \notin \text{dom} f$, we set $\hat{\partial} f(x) = \emptyset$.

2. The limiting-subdifferential, or simply the subdifferential, of $f$ at $x \in \text{dom} f$, written as $\partial f(x)$, is defined as follows:
$$\partial f(x) = \{ x^* \in \mathbb{R}^n \mid \exists x_n \to x, f(x_n) \to f(x), x_n^* \in \hat{\partial} f(x_n), \text{ with } x_n^* \to x^* \}. $$ 
A necessary condition for $x \in \mathbb{R}^n$ to be a minimizer of $f$ is
$$0 \in \partial f(x). $$ (9)
And a point which satisfies (9) is called the critical point or stationary point. The
critical points set of $f$ is denoted by $\text{crit} f$.

Definition 2.2. we say that $(x^*, y^*, \lambda^*)$ is a critical point of the augmented La-
grangian function $L_\beta(\cdot)$ defined by (3), if it satisfies:
$$\begin{cases}
A^T \lambda^* \in \partial f(x^*), \\
\nabla g(y^*) = -B^T \lambda^*, \\
Ax^* - By^* - c = 0.
\end{cases} $$ (10)
The critical points set of $L_\beta(\cdot)$ is denoted by $\text{crit} L_\beta$. 
Obviously, a critical point of the augmented Lagrangian function \( L_\beta(\cdot) \) is corresponding to a KKT point of problem (1). In this paper, we assume that there is at least a KKT point of problem (1).

For a proper closed convex function \( h : \mathcal{R}^n \to \mathcal{R} \cup \{+\infty\} \), the subdifferential of \( h \) at \( x \in \text{dom} h \) is given by
\[
\partial h(x) = \{ \xi \in \mathcal{R}^n \mid h(u) - h(x) - \langle \xi, u - x \rangle \geq 0, \ \forall u \in \mathcal{R}^n \}.
\]
From the definition we can find that, the set \( \partial h(x) \) is closed and convex. We use Prox\(_x h(v) \) to denote the proximal operator of a proper closed convex function \( h \) at any \( v \in \mathcal{R}^n \), i.e.,
\[
\text{Prox}_x h(v) = \arg \min_{x \in \mathcal{R}^n} \left\{ h(x) + \frac{1}{2\gamma} \|x - v\|^2 \right\}.
\]
Obviously, this operator is well defined for any \( v \in \mathcal{R}^n \), and we refer the readers to [17] for properties of the proximal operator. According to the definition of the proximal operator, we have the following lemma.

**Lemma 2.3.** Suppose that \( f : \mathcal{R}^n \to \mathcal{R} \cup \{+\infty\} \) is a proper convex lower semi-continuous function, \( g : \mathcal{R}^m \to \mathcal{R} \) is a possibly nonconvex function and \( \nabla g \) is Lipschitz continuous with modulus \( L > 0 \), then solving (10) is equivalent to finding a zero point of
\[
e(w, \gamma) = \begin{pmatrix}
e_x(w, \gamma) := x - \text{Prox}_x f(x + \gamma A^T \lambda) \\
e_y(w, \gamma) := \gamma(\nabla g(y) + B^T \lambda) \\
e_\lambda(w, \gamma) := \gamma(Ax - By - e)
\end{pmatrix},
\]
where \( w := (x, y, \lambda) \) and \( \gamma > 0 \) is an arbitrary but fixed parameter.

Considering that \( e(w^*, \gamma) = 0 \) when \( w^* = (x^*, y^*, \lambda^*) \in \text{crit} L_\beta \), so \( e(w, \gamma) \) can be viewed as a measurement of the distance from \( w \) to the critical points set \( \text{crit} L_\beta \). Some useful properties respecting to \( e(w, \gamma) \) will be summarized in the following lemma. Since the proof has been given in [10], we omit it here.

**Lemma 2.4.** Suppose that \( f : \mathcal{R}^n \to \mathcal{R} \cup \{+\infty\} \) is a proper convex lower semi-continuous function, \( g : \mathcal{R}^m \to \mathcal{R} \) is a possibly nonconvex function and \( \nabla g \) is Lipschitz continuous with modulus \( L > 0 \). If \( w \in \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^m \) is not a critical point of \( L_\beta(\cdot) \), and \( \bar{\gamma} \geq \gamma > 0 \), we have
\[
\|e(w, \bar{\gamma})\| \geq \|e(w, \gamma)\|
\]
and
\[
\frac{\|e(w, \bar{\gamma})\|}{\bar{\gamma}} \leq \frac{\|e(w, \gamma)\|}{\gamma}.
\]
In Section 4, we further assume that \( f : \mathcal{R}^n \to \mathcal{R} \cup \{+\infty\} \) is a convex function in (1), then we can establish local linear convergence of the iterative sequence \( \{w^k\} \) and the auxiliary sequence \( \{\mathcal{F}_k\} \) (defined in (25)) under the following two assumptions, which have been used in the convergence analysis of many algorithms [14, 15, 20].

Now we give the two assumptions used in the following analysis. The first one is the error bound condition.

**Assumption 2.5.** [error bound condition] For any \( \zeta \geq \inf_{w \in \mathcal{R}^{2m+n}} L_\beta(w) \), there exists scalars \( \varepsilon > 0 \) and \( \tau > 0 \) such that
\[
\text{dist}(w, \text{crit} L_\beta) \leq \tau \|e(w, 1)\|,
\]
whenever \( \|e(w, 1)\| \leq \varepsilon \) and \( L_\beta(w) \leq \zeta \).
Remark 2.6. Note that Lemma 2.4 implies that for any fixed $\gamma > 0$,
\[\|e(w, 1)\| \leq \max \left\{ \gamma, \frac{1}{\gamma} \right\} \|e(w, \gamma)\|.
\]
Therefore, the Assumption 2.5 holds for any fixed $\gamma > 0$, i.e.,
\[\text{dist}(w, \text{crit } L_\beta) \leq \tau \max \left\{ \gamma, \frac{1}{\gamma} \right\} \|e(w, \gamma)\|,
\]
and we take $\gamma \equiv 1$ in the following analysis without loss of generality.

The second assumption concerning the separation of stationary values.

Assumption 2.7. For any $\bar{w} = (\bar{x}, \bar{y}, \bar{\lambda}) \in \text{crit } L_\beta$, there exists $\delta > 0$ so that $L_\beta(\bar{w}) = L_\beta(\tilde{w})$, whenever $\tilde{w} = (\tilde{x}, \tilde{y}, \tilde{\lambda}) \in \text{crit } L_\beta$ and $\|\bar{w} - \tilde{w}\| \leq \delta$.

This assumption means that the isocost surfaces of $L_\beta$ restricted to the critical points set $\text{crit } L_\beta$ are properly separated.

Remark 2.8. Assumption 2.7 is automatically satisfied from the KL property (see [1] Remark 2.5 (d)), which is a mild condition.

The following descent lemma for smooth function is useful for the convergence analysis.

Lemma 2.9. Let $g: \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function with $\nabla g$ is Lipschitz continuous with modulus $L > 0$, then for any $x, y \in \mathbb{R}^n$, we have
\[|g(y) - g(x) - \langle \nabla g(x), y - x \rangle| \leq \frac{L}{2} \|y - x\|^2.
\]

Following, we will consider the convergence results of the symmetric ADMM (5). For the sake of brevity, we set the matrix $B$ in (1) to be an identity matrix $I$.

3. The sublinear convergence of the symmetric ADMM. In this section, we discuss the convergence rate of the sequence generated by the symmetric ADMM (5) for solving (1). Throughout this paper, we make the following assumptions.

Assumption 3.1. Assume the following holds:
(1) $\alpha + \theta > 0$, $1 - \left| \frac{\theta - 1}{1 + \alpha} \right| > 0$;
(2) $Q = \frac{1}{2} - \frac{\alpha \theta}{\alpha + \theta} - \frac{6 \alpha^2 (\theta - 1)^2}{(1 + \alpha - |\theta - 1|)^2 (\alpha + \theta)} > 0$;
(3) $\beta$ satisfies
\[Q \beta^2 - \frac{L}{2} \beta - \frac{3L^2 (\alpha + \theta)}{(1 + \alpha - |\theta - 1|)^2} > 0.
\]

Remark 3.2. In order to obtain an intuitive understanding of the relationships between $\alpha$ and $\theta$ in Assumption 3.1, we give the corresponding Figure 1. Obviously, the assumptions can be satisfied in the grey shaded area intersected by the blue curve and red curve.
Lemma 3.3. For the sequence \( \{x^k, y^k, \lambda^k\} \) generated by the symmetric ADMM (5), for any \( k \geq 0 \), we have

(i) \( L_\beta(x^{k+1}, y^k, \lambda^k) \leq L_\beta(x^k, y^k, \lambda^k) \);

(ii) \( L_\beta(x^{k+1}, y^{k+1}, \lambda^k) \leq L_\beta(x^{k+1}, y^k, \lambda^k) + \alpha \beta (Ax^{k+1} - y^k - c, \Delta y^{k+1}) - \frac{\beta - L}{2} \| \Delta y^{k+1} \|^2 \);

(iii) \( L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) = L_\beta(x^{k+1}, y^{k+1}, \lambda^k) - \langle \Delta \lambda^{k+1}, Ax^{k+1} - y^{k+1} - c \rangle \).

Proof. (i) According to the update of \( x \) in iteration (5), we have

\[ L_\beta(x^{k+1}, y^k, \lambda^k) \leq L_\beta(x^k, y^k, \lambda^k). \]

(ii) Since \( \nabla g \) is Lipschitz continuous with modulus \( L > 0 \), it follows from Lemma 2.9 that

\[ L_\beta(x^{k+1}, y^k, \lambda^k) \geq L_\beta(x^{k+1}, y^{k+1}, \lambda^k) + \langle \nabla_y L_\beta(x^{k+1}, y^{k+1}, \lambda^k), y^k - y^{k+1} \rangle + \frac{\beta - L}{2} \| y^k - y^{k+1} \|^2. \] (11)

According to the update of \( y \) in iteration (5), we have

\[ 0 = \nabla g(y^{k+1}) + \lambda^{k+\frac{1}{2}} - \beta (Ax^{k+1} - y^{k+1} - c). \]

Thus

\[ \nabla_y L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) = \nabla g(y^{k+1}) + \lambda^{k+\frac{1}{2}} - \beta (Ax^{k+1} - y^{k+1} - c) \]

\[ = \nabla g(y^{k+1}) + \lambda^{k+\frac{1}{2}} - \beta (Ax^{k+1} - y^{k+1} - c) + (\lambda^k - \lambda^{k+\frac{1}{2}}) \]

\[ = \alpha \beta (Ax^{k+1} - y^k - c), \]

(12)

where the last equality follows from \( \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha \beta (Ax^{k+1} - y^k - c) \). Substitute (12) into (11), then we have

\[ L_\beta(x^{k+1}, y^{k+1}, \lambda^k) \leq L_\beta(x^{k+1}, y^k, \lambda^k) + \alpha \beta (Ax^{k+1} - y^k - c, y^{k+1} - y^k) - \frac{\beta - L}{2} \| y^k - y^{k+1} \|^2. \]

Combining with the identity \( \Delta y^{k+1} = y^{k+1} - y^k \), (ii) is proved.
Lemma 3.4. For any given \((y^0, \lambda^0) \in \mathbb{R}^m \times \mathbb{R}^m\), let \(\Delta y^0 = y^0 - y^{-1}\) and \(\Delta \lambda^0 = \lambda^0 - \lambda^{-1}\) be such that

\[
\frac{\theta - 1}{\alpha + 1} \Delta \lambda^0 + \frac{\alpha \beta (\theta - 1)}{\alpha + 1} \Delta y^0 = \frac{\theta + \alpha}{\alpha + 1} \lambda^0 + \frac{\alpha + \theta}{\alpha + 1} \nabla g(y^0),
\]

i.e., \(\Delta \lambda^0\) and \(\Delta y^0\) are the solution of the equality (13). Then, for the sequence \(\{x^k, y^k, \lambda^k\}\) generated by the symmetric ADMM (5), we have

\[
\Delta \lambda^{k+1} = -\frac{\theta - 1}{\alpha + 1} \Delta \lambda^k + u^k, \quad \forall k \geq 0,
\]

where

\[
u^k = -\frac{\theta + \alpha}{\alpha + 1} (\nabla g(y^{k+1}) - \nabla g(y^k)) + \frac{\alpha \beta (\theta - 1)}{\alpha + 1} (\Delta y^{k+1} - \Delta y^k).
\]

Proof. According to the update of \(\lambda\) in iterative (5), we have

\[
\lambda^{k+1} = \lambda^k - (\theta \beta (Ax^{k+1} - y^{k+1} - c) + \alpha \beta (Ax^{k+1} - y^k - c)) = \lambda^k - (\alpha + \theta) \beta Ax^{k+1} + \alpha \beta (y^k + c) + \theta \beta (y^{k+1} + c).
\]

So,

\[
(\alpha + \theta) \beta Ax^{k+1} = \lambda^{k+1} = \lambda^k + \alpha \beta (y^k + c) + \theta \beta (y^{k+1} + c).
\]

The update of \(y\) in iterative (5) shows that

\[
0 = \nabla g(y^{k+1}) + \lambda^{k+1} - \beta (Ax^{k+1} - y^k - c) = \nabla g(y^{k+1}) + \lambda^{k+1} + (\theta - 1) \beta (Ax^{k+1} - y^{k+1} - c).
\]

Substituting (16) into (17), we can deduce

\[
0 = \nabla g(y^{k+1}) + \frac{\alpha + 1}{\alpha + \theta} \lambda^{k+1} + \frac{\theta - 1}{\alpha + \theta} \lambda^k + \frac{\alpha \beta (\theta - 1)}{\alpha + \theta} (y^k - y^{k+1}).
\]

Thus we get

\[
-\frac{\alpha + 1}{\alpha + \theta} \lambda^{k+1} = \nabla g(y^{k+1}) + \frac{\theta - 1}{\alpha + \theta} \lambda^k - \frac{\alpha \beta (\theta - 1)}{\alpha + \theta} (y^k - y^{k+1}).
\]

By taking \(k := k\) and \(k := k - 1\) in the above equality respectively, and then making subtraction, we can get

\[
-\frac{\alpha + 1}{\alpha + \theta} \Delta \lambda^{k+1} = (\nabla g(y^{k+1}) - \nabla g(y^k)) + \frac{\theta - 1}{\alpha + \theta} \Delta \lambda^k - \frac{\alpha \beta (\theta - 1)}{\alpha + \theta} (\Delta y^{k+1} - \Delta y^k).
\]

And hence, in view of (15), we deduce the equality (14) for every \(k \geq 1\).

Also, rearranging (19), we have

\[
\lambda^{k+1} = \frac{\alpha + \theta}{\alpha + 1} \nabla g(y^{k+1}) - \frac{\theta - 1}{\alpha + 1} \lambda^k + \frac{\alpha \beta (\theta - 1)}{\alpha + 1} (y^{k+1} - y^k).
\]
Together with (15), and take \( k = 0 \), we get
\[
\Delta\lambda^1 = \lambda^1 - \lambda^0 = \frac{\alpha + \theta}{\alpha + 1} \nabla g(y^1) - \frac{\theta + \alpha}{\alpha + 1} \lambda^0 + \frac{\alpha \beta (\theta - 1)}{\alpha + 1} \Delta y^1,
\]
\[
= - \frac{\theta + \alpha}{\alpha + 1} \lambda^0 - \frac{\alpha + \theta}{\alpha + 1} \nabla g(y^0) + a^0 + \frac{\alpha \beta (\theta - 1)}{\alpha + 1} \Delta y^0.
\]
Considering the definition of \( \Delta y^0 \) in (13), we deduce that (14) also holds for \( k = 0 \).

\[ \square \]

**Remark 3.5.** Let \((y^0, \lambda^0) \in \mathbb{R}^m \times \mathbb{R}^m\) and the constants \( \theta, \alpha \) be given. Then, the equality (13) in terms of \( \Delta\lambda^0 = \lambda^0 - \lambda^{-1} \) and \( \Delta y^0 = y^0 - y^{-1} \) must have solution. In fact that

(i) when \( \frac{\theta - 1}{\alpha + 1} = 0 \) and the choice of the initial point \((y^0, \lambda^0)\) satisfies the equality
\[ 0 = \frac{\theta + \alpha}{\alpha + 1} \lambda^0 + \frac{\alpha \beta}{\alpha + 1} \nabla g(y^0) \], then \( \Delta y^0 \) and \( \Delta \lambda^0 \) can be any vectors;
(ii) when \( \frac{\theta - 1}{\alpha + 1} \neq 0 \), then the nonhomogeneous systems of equations (13) must have solutions.

Then we have the following lemma which is very important for our convergence analysis.

**Lemma 3.6.** For the sequence \( \{x^k, y^k, \lambda^k\} \) generated by symmetric ADMM (5), under Assumption 3.1, for any \( k \geq 0 \), we have
\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^{k+1}\|^2 + c_2 \|\Delta y^{k+1}\|^2
\]
\[
\leq L_\beta(x^k, y^k, \lambda^k) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^k\|^2 + c_2 \|\Delta y^k\|^2 - a_1 \|\Delta y^{k+1}\|^2 - \left( \frac{c_1}{2} - \frac{c_1}{2 + \sigma} \right) \|\Delta \lambda^{k+1}\|^2,
\]
where \( \gamma = \left( \frac{1}{\theta - 1} \right) \left( \frac{\theta - 1}{\alpha + 1} \right) \), \( a_1 = \left( \frac{1}{2} - \frac{\alpha \beta}{\alpha + 1} \right) \beta - \frac{\alpha^2 + 2 \alpha \beta + \beta^2}{\alpha + 1} \), \( \beta > 0 \), \( c_1 = \frac{\beta^2 + 2 \beta}{\alpha + 1} \). (20)

**Proof.** Together with (i), (ii) and (iii) of Lemma 3.3, we have
\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) \leq L_\beta(x^k, y^k, \lambda^k) + \alpha \beta \left( A x^{k+1} - y^{k+1} - c, \Delta y^{k+1} \right)
\]
\[
- \left( \Delta \lambda^{k+1}, A x^{k+1} - y^{k+1} - c \right) - \frac{\beta - L}{2} \|\Delta y^{k+1}\|^2.
\]
Substituting (16) into (21), we have
\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1})
\]
\[
\leq L_\beta(x^k, y^k, \lambda^k) - \frac{\beta - L}{2} \|\Delta y^{k+1}\|^2
\]
\[
+ \alpha \beta \left( \frac{\lambda^k - \lambda^{k+1}}{\alpha + \theta} + \frac{\theta (y^{k+1} - y^k)}{\alpha + \theta}, \Delta y^{k+1} \right) - \left( \Delta \lambda^{k+1}, \frac{\lambda^k - \lambda^{k+1}}{\alpha + \theta} + \frac{\alpha (y^k - y^{k+1})}{\alpha + \theta} \right)
\]
\[
= L_\beta(x^k, y^k, \lambda^k) - \left( \frac{\beta - L}{2} - \frac{\alpha \beta \theta}{\alpha + \theta} \right) \|\Delta y^{k+1}\|^2 + \frac{1}{\alpha + \theta} \|\Delta \lambda^{k+1}\|^2.
\]
(22)
Combining (22) and (14), we have

\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2} \|\Delta \lambda^{k+1}\|^2 \\
\leq L_\beta(x^k, y^k, \lambda^k) - \left(\frac{\beta - L}{2} - \frac{\alpha \beta \theta}{\alpha + \theta}\right) \|\Delta y^{k+1}\|^2 \\
+ \left(\frac{c_1}{2} + \frac{1}{\alpha + \theta}\right) \left(\left(\frac{\theta - 1}{\alpha + 1}\right)^2 + \left(1 + \frac{1}{t}\right) \|u^k\|^2\right).
\]

where \(c_1\) is a non-negative constant. Together with the relation \(\|s_1 + s_2\|^2 \leq (1 + t)\|s_1\|^2 + (1 + \frac{1}{t}) \|s_2\|^2, \forall s_1, s_2 \in \mathbb{R}^n, t > 0,\) the above inequality yields

\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2} \|\Delta \lambda^{k+1}\|^2 \\
\leq L_\beta(x^k, y^k, \lambda^k) - \left(\frac{\beta - L}{2} - \frac{\alpha \beta \theta}{\alpha + \theta}\right) \|\Delta y^{k+1}\|^2 \\
+ \left(\frac{c_1}{2} + \frac{1}{\alpha + \theta}\right) \left(\left(\frac{\theta - 1}{\alpha + 1}\right)^2 + \left(1 + \frac{1}{t}\right) \|u^k\|^2\right).
\]

Taking \((t + 1)\left|\frac{\theta - 1}{1 + \alpha}\right| = 1, c_1 = \frac{\left|\frac{\theta - 1}{1 + \alpha}\right|}{(\alpha + \theta)(\frac{\theta - 1}{\alpha + 1} - \frac{\theta - 1}{1 + \alpha})} \geq 0, \sigma > 0,\) we can get

\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2} \|\Delta \lambda^{k+1}\|^2 \\
\leq L_\beta(x^k, y^k, \lambda^k) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^{k}\|^2 \\
- \left(\frac{\beta - L}{2} - \frac{\alpha \beta \theta}{\alpha + \theta}\right) \|\Delta y^{k+1}\|^2 + \frac{2}{2 + \sigma} \left|\frac{1 + \alpha}{\theta - 1}\right| \beta\gamma \|u^k\|^2,
\]

where \(\gamma = \left(\frac{1 + \alpha}{\theta - 1}\right) - 1\left(\frac{2}{2 + \sigma} - \frac{\theta - 1}{1 + \alpha}\right).\) Rearranging the inequality, we have

\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^{k+1}\|^2 \\
\leq L_\beta(x^k, y^k, \lambda^k) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^{k}\|^2 - \left(\frac{c_1}{2} - \frac{c_1}{2 + \sigma}\right) \|\Delta y^{k+1}\|^2 \\
- \left(\frac{\beta - L}{2} - \frac{\alpha \beta \theta}{\alpha + \theta}\right) \|\Delta y^{k+1}\|^2 + \frac{2}{2 + \sigma} \left|\frac{1 + \alpha}{\theta - 1}\right| \beta\gamma \|u^k\|^2. \tag{23}
\]

According to (15) and the fact that \(\nabla g\) is Lipschitz continuous with modulus \(L > 0,\) we have

\[
\|u^k\|^2 \leq 3 \left(\frac{\alpha + \theta}{\alpha + 1}\right)^2 L^2 \|\Delta y^{k+1}\|^2 + 3\alpha^2 \beta^2 \left(\frac{\theta - 1}{\alpha + 1}\right)^2 \left(\|\Delta y^{k+1}\|^2 + \|\Delta y^k\|^2\right), \tag{24}
\]
where the inequality follows from the relation \((s_1 + s_2 + s_3)^2 \leq 3s_1^2 + 3s_2^2 + 3s_3^2, \forall s_1, s_2, s_3 \in \mathcal{R}\). Combining (23) and (24), we can deduce
\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2+\sigma}\|\Delta \lambda^{k+1}\|^2 + c_2\|\Delta y^{k+1}\|^2
\leq L_\beta(x^k, y^k, \lambda^k) + \frac{c_1}{2+\sigma}\|\Delta \lambda^k\|^2 + c_2\|\Delta y^k\|^2 + \left(\frac{3\alpha^2 \beta \frac{2}{2+\sigma} \left|\frac{\theta - 1}{1+\alpha}\right|}{(\alpha + \theta)\gamma} - c_2\right)\|\Delta y^k\|^2
+ \left(\frac{3L^2(\alpha + \theta)\frac{2}{2+\sigma} \left|\frac{1+\alpha}{1+\alpha}\right|}{\beta (1+\alpha)^2\gamma} + \frac{3\alpha^2 \beta \frac{2}{2+\sigma} \left|\frac{\theta - 1}{1+\alpha}\right|}{(\alpha + \theta)\gamma} + c_2\right)\|\Delta y^{k+1}\|^2
- \left(\frac{\beta - L}{2} - \frac{\alpha \beta \theta}{\alpha + \theta}\right)\|\Delta y^{k+1}\|^2 - \left(\frac{c_1}{2} - \frac{c_1}{2+\sigma}\right)\|\Delta \lambda^k\|^2
- a_1\|\Delta y^{k+1}\|^2 - a_2\|\Delta y^k\|^2 - \left(\frac{c_1}{2} - \frac{c_1}{2+\sigma}\right)\|\Delta \lambda^{k+1}\|^2,
\]
where \(\gamma = \left(\left|\frac{1+\alpha}{1+\alpha}\right| - 1\right)\left(\frac{2}{2+\sigma} - \left|\frac{\theta - 1}{1+\alpha}\right|\right)\).
\(a_1 = \frac{\beta - L}{2} - \frac{\alpha \beta \theta}{\alpha + \theta} - \frac{3L^2(\alpha + \theta)\frac{2}{2+\sigma} \left|\frac{1+\alpha}{1+\alpha}\right|}{\beta (1+\alpha)^2\gamma} - 3\alpha^2 \beta \frac{2}{2+\sigma} \left|\frac{\theta - 1}{1+\alpha}\right|}{(\alpha + \theta)\gamma} - c_2\), \(c_2\) is a non-negative constant, and \(a_2 = c_2 - \frac{3\alpha^2 \beta \frac{2}{2+\sigma} \left|\frac{\theta - 1}{1+\alpha}\right|}{(\alpha + \theta)\gamma}\). Under the Assumption 3.1, there must exist a non-negative \(c_2\) such that \(a_1 > 0\) and \(a_2 \geq 0\). For brevity, we take \(c_2 = \frac{3\alpha^2 \beta \frac{2}{2+\sigma} \left|\frac{\theta - 1}{1+\alpha}\right|}{(\alpha + \theta)\gamma}\) \(\geq 0\) and the lemma is proved.

The following regularization of the augmented Lagrangian function \(F_\beta : \mathcal{R}^n \times \mathcal{R}^m \times \mathcal{R}^m \times \mathcal{R}^m \to \mathcal{R}\) will play an important role in the convergence analysis of the nonconvex symmetric ADMM algorithm:
\[
F_\beta(x, y, \lambda, \hat{\lambda}) = L_\beta(x, y, \lambda) + \frac{c_1}{2+\sigma}\|\lambda - \hat{\lambda}\|^2 + c_2\|y - \hat{y}\|^2,
\]
where \(c_1\), \(c_2\), and \(\sigma\) are defined in Lemma 3.6. For every \(k \geq 0\), we denote
\[
F_k := F_\beta(x^k, y^k, \lambda^k, y^{k-1}, \lambda^{k-1}) = L_\beta(x^k, y^k, \lambda^k) + \frac{c_1}{2+\sigma}\|\Delta \lambda^k\|^2 + c_2\|\Delta y^k\|^2. (25)
\]

We are now ready to prove the main result of this section.

**Theorem 3.7.** Suppose that \(A \in \mathcal{R}^{m \times n}\) is of full column rank, and Assumption 3.1 is satisfied. Let \(\{w^k := (x^k, y^k, \lambda^k)\}\) be the sequence generated by the symmetric ADMM (5) which is assumed to be bounded. Then,
(i) the sequences \(\{L_\beta(w^k)\}\) and \(\{F_k\}\) are convergent, and we have \(\lim_{k \to +\infty} L_\beta(w^k) = \lim_{k \to +\infty} F_k\).
(ii) for every \(k \geq 1\), there exists \(j \leq k\) such that
\[
\|\Delta x^j\| \leq \sqrt{\frac{C_1 (F_0 - F^*)}{k}},
\]
\[
\|\Delta y^j\| \leq \sqrt{\frac{F_0 - F^*}{a_1 k}},
\]
\[
\|\Delta \lambda^j\| \leq \sqrt{\frac{\mathcal{F}_0 - \mathcal{F}^*}{C_1 \left( \frac{c_1}{2} - \frac{c_1}{2 + \sigma} \right) k}},
\]

(27)

where the parameters \( \mathcal{F}^* = \lim_{k \to \infty} \mathcal{F}_k \), \( C_1 = \frac{8}{(\alpha + \theta)^2 \beta^2 \lambda_{\text{min}}(A^T A)} \left( \frac{1}{2} - \frac{1}{2 + \sigma} \right) + \frac{4\alpha^2 + 4\theta^2}{(\alpha + \theta)^2 \lambda_{\text{min}}(A^T A) a_1} \).

and \( a_1, c_1 \) are defined in Lemma 3.6.

**Proof.** (i) We will first show that \( \{L_{\beta}(w^k)\} \) is bounded from below, which will imply that \( \{\mathcal{F}_k\} \) is bounded from below as well. Since \( \{w^k\} \) is bounded, it has at least one accumulation point. Let \( w^* \) be an accumulation point of \( \{w^k\} \) and let \( \{w^k_j\} \) be the subsequence converging to it, i.e., \( w^{k_j} \to w^* \). Since \( f \) is the lower semicontinuous and \( g \) is continuous, \( L_{\beta}(\cdot) \) is lower semicontinuous, and hence

\[
L_{\beta}(w^*) \leq \liminf_{j \to +\infty} L_{\beta}(w^{k_j}).
\]

Consequently, \( \{L_{\beta}(w^{k_j})\} \) is bounded from below. And the \( \{\mathcal{F}_k\} \) is bounded from below as well, which, together with the fact that \( \{\mathcal{F}_k\} \) is nonincreasing, means that \( \{\mathcal{F}_k\} \) is convergent. Then \( \{\mathcal{F}_k\} \) is convergent. Considering that the parameters \( a_1, \sigma > 0 \), inequality (20) implies

\[
\|y^{k+1} - y^k\| \to 0, \|\lambda^{k+1} - \lambda^k\| \to 0.
\]

(28)

According to the definition of \( \mathcal{F}_k \) in (25), we know that \( \{L_{\beta}(w^k)\} \) is also convergent, and

\[
\lim_{k \to +\infty} L_{\beta}(w^k) = \lim_{k \to +\infty} \mathcal{F}_k.
\]

(ii) Set \( \mathcal{F}^* = \lim_{k \to +\infty} \mathcal{F}_k \). Further, the inequality (20) implies

\[
\sum_{j=1}^{k} a_1 \|\Delta y^j\|^2 + \sum_{j=1}^{k} \left( \frac{c_1}{2} - \frac{c_1}{2 + \sigma} \right) \|\Delta \lambda^{k+1}\|^2 \leq \mathcal{F}_0 - \mathcal{F}_k < \mathcal{F}_0 - \mathcal{F}^* < +\infty,
\]

which, in particular, implies that

\[
\sum_{j=1}^{k} \|\Delta y^j\|^2 \leq \frac{\mathcal{F}_0 - \mathcal{F}^*}{a_1},
\]

(29)

and

\[
\sum_{j=1}^{k} \|\Delta \lambda^{k+1}\|^2 \leq \frac{\mathcal{F}_0 - \mathcal{F}^*}{\frac{1}{2} - \frac{1}{2 + \sigma}}.
\]

(30)

Then we will show that inequality (26) holds. By rewriting (16), we have

\[
Ax^{k+1} = -\frac{1}{(\alpha + \theta)\beta} \Delta \lambda^{k+1} + \frac{\alpha}{\alpha + \theta} y^k + \frac{\theta}{\alpha + \theta} y^{k+1} + c.
\]

It is obvious that

\[
A(x^{k+1} - x^k) = \frac{1}{(\alpha + \theta)\beta} (\Delta \lambda^k - \Delta \lambda^{k+1}) + \frac{\alpha}{\alpha + \theta} \Delta y^k + \frac{\theta}{\alpha + \theta} \Delta y^{k+1}.
\]
Thus
\[
\|\Delta x^{k+1}\|^2 \leq \frac{1}{\lambda_{\text{min}}(ATA)} \|A\Delta x^{k+1}\|^2 \\
\leq \frac{4}{(\alpha + \theta)^2 \beta^2 \lambda_{\text{min}}(ATAT)} (\|\Delta \lambda^k\|^2 + \|\Delta \lambda^{k+1}\|^2) \\
+ \frac{4\alpha^2}{(\alpha + \theta)^2 \lambda_{\text{min}}(ATAT)} \|\Delta y^k\|^2 + \frac{4\theta^2}{(\alpha + \theta)^2 \lambda_{\text{min}}(ATAT)} \|\Delta y^{k+1}\|^2.
\]
(31)

Using (29), (30) and (31), we obtain
\[
\sum_{j=1}^k \|\Delta x^j\|^2 \leq C_1 (\mathcal{F}_0 - \mathcal{F}^*),
\]
where \(C_1 = \frac{8}{(\alpha + \theta)^2 \beta^2 \lambda_{\text{min}}(ATAT)(\frac{3}{2} - \frac{\alpha - 1}{1+\alpha})} + \frac{4\alpha^2 + 4\theta^2}{(\alpha + \theta)^2 \lambda_{\text{min}}(ATAT)\sigma_1}. \) This completes the proof.

\[\square\]

The above theorem established the \(O(1/\sqrt{k})\) sublinear convergence rate of the sequence generated by the symmetric ADMM (5) for solving nonconvex separable optimization problems, when both component functions of the objective function are nonconvex. Next, we give some sufficient conditions for guarantee that the sequence \(\{(x^k, y^k, \lambda^k)\}\) generated by the symmetric ADMM (5) is bounded.

**Lemma 3.8.** Let \(\{(x^k, y^k, \lambda^k)\}\) be the sequence generated by the symmetric ADMM (5), suppose that
\[
\inf_y \left\{ g(y) - \frac{3}{2\beta} \|\nabla g(y)\|^2 \right\} = g^* > -\infty,
\]
and \(2(\alpha + \theta) \frac{\alpha - 1}{1+\alpha} > 3(\theta - 1)^2 \left(1 - \frac{\alpha - 1}{1+\alpha}\right)\). If one of the following statements is true:
(i) function \(f\) is coercive;
(ii) \(A \in \mathbb{R}^{m \times n}\) is of full column rank, function \(f\) is bounded from below and \(g\) is coercive,
then the sequence \(\{(x^k, y^k, \lambda^k)\}\) is bounded.

**Proof.** From Lemma 3.6, we have that for all \(k \geq 0\)
\[
\mathcal{F}_{k+1} + a_1 \|\Delta y^{k+1}\|^2 \leq \mathcal{F}_k, \quad a_1 > 0,
\]
which shows that \(\{\mathcal{F}_k\}\) is monotonically decreasing. Consequently, we have
\[
\mathcal{F}_0 \geq \mathcal{F}_{k+1} + a_1 \|\Delta y^{k+1}\|^2 \\
= L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2+\sigma} \|\Delta \lambda^{k+1}\|^2 + (a_1 + c_2) \|\Delta y^{k+1}\|^2 \\
= f(x^{k+1}) + g(y^{k+1}) - \langle \lambda^{k+1}, Ax^{k+1} - y^{k+1} - c \rangle \\
+ \frac{\beta}{2} \|Ax^{k+1} - y^{k+1} - c\|^2 + \frac{c_1}{2+\sigma} \|\Delta \lambda^{k+1}\|^2 + (a_1 + c_2) \|\Delta y^{k+1}\|^2 \\
= f(x^{k+1}) + g(y^{k+1}) + \frac{\beta}{2} \|Ax^{k+1} - y^{k+1} - c\|^2 - \frac{1}{\beta} \lambda^{k+1}\|^2 \\
- \frac{1}{2\beta} \|\lambda^{k+1}\|^2 + \frac{c_1}{2+\sigma} \|\Delta \lambda^{k+1}\|^2 + (a_1 + c_2) \|\Delta y^{k+1}\|^2, \quad \forall k \geq 0.
\]
According to (18), we have
\[
\lambda^{k+1} = -\nabla g(y^{k+1}) + \frac{\theta - 1}{\alpha + \theta} \Delta \lambda^{k+1} + \frac{\alpha \beta (\theta - 1)}{\alpha + \theta} \Delta y^{k+1}.
\] (33)

Substituting (33) into (32), we have
\[
x_0 \geq f(x^{k+1}) + \left( \frac{\mu}{2} \|\nabla g(y^{k+1})\|^2 + \frac{\beta}{2} \|Ax^{k+1} - y^{k+1} - c - \frac{1}{\beta} \lambda^{k+1}\|^2 \right.
\]
\[
+ \left( \frac{c_1}{2 + \sigma} - \frac{3(\theta - 1)^2}{2(\alpha + \theta)^2} \right) \|\Delta \lambda^{k+1}\|^2 + \left( a_1 + c_2 - \frac{3\alpha^2 \beta (\theta - 1)^2}{2(\alpha + \theta)^2} \right) \|\Delta y^{k+1}\|^2, \quad \forall k \geq 0.
\] (34)

which follows from the relation \((s_1 + s_2 + s_3)^2 \leq 3s_1^2 + 3s_2^2 + 3s_3^2, \forall s_1, s_2, s_3 \in R\).

Now we will prove the boundedness of \(\{(x^k, y^k, \lambda^k)\}\) under each of the two scenarios.

(i) Suppose \(f\) is coercive. It follows from (34) that
\[
x_0 - g^* \geq f(x^{k+1}) + \frac{3}{2\beta} \|\nabla g(y^{k+1})\|^2 + \frac{\beta}{2} \|Ax^{k+1} - y^{k+1} - c - \frac{1}{\beta} \lambda^{k+1}\|^2
\]
\[
+ \left( \frac{c_1}{2 + \sigma} - \frac{3(\theta - 1)^2}{2(\alpha + \theta)^2} \right) \|\Delta \lambda^{k+1}\|^2 + \left( a_1 + c_2 - \frac{3\alpha^2 \beta (\theta - 1)^2}{2(\alpha + \theta)^2} \right) \|\Delta y^{k+1}\|^2, \quad \forall k \geq 0.
\]

By choosing \(\beta\) sufficiently large, we can guarantee \(\frac{c_1}{2 + \sigma} - \frac{3(\theta - 1)^2}{2(\alpha + \theta)^2} > 0\) and \(a_1 + c_2 - \frac{3\alpha^2 \beta (\theta - 1)^2}{2(\alpha + \theta)^2} > 0\). Considering that \(f\) is coercive, we obtain that the sequences \(\{x^k\}\), \(\{Ax^k - y^k - c - \frac{1}{\beta} \lambda^k\}\), \(\{\Delta y^k\}\) and \(\{\Delta \lambda^k\}\) are bounded. Rearranging (16), we get
\[
(\alpha + \theta) \beta y^k = (\alpha + \theta) \beta Ax^k + \Delta \lambda^k + \alpha \beta \Delta y^k - (\alpha + \theta) \beta c,
\]
which shows that \(\{y^k\}\) is also bounded. Therefore, \(\{\lambda^k\}\) is bounded and hence \(\{w^k\}\) is bounded.

(ii) Suppose \(A\) is of full column rank, \(f\) is bounded from below and \(g\) is coercive. Again thanks to (34), we see that for all \(k \geq 0\)
\[
x_0 - f(x^{k+1}) - \frac{1}{2} \geq \frac{1}{2} \|g(y^{k+1})\|^2 + \frac{\beta}{2} \|Ax^{k+1} - y^{k+1} - c - \frac{1}{\beta} \lambda^{k+1}\|^2
\]
\[
+ \left( \frac{c_1}{2 + \sigma} - \frac{3(\theta - 1)^2}{2(\alpha + \theta)^2} \right) \|\Delta \lambda^{k+1}\|^2 + \left( a_1 + c_2 - \frac{3\alpha^2 \beta (\theta - 1)^2}{2(\alpha + \theta)^2} \right) \|\Delta y^{k+1}\|^2.
\]

By choosing \(\beta\) sufficiently large, we can guarantee \(\frac{c_1}{2 + \sigma} - \frac{3(\theta - 1)^2}{2(\alpha + \theta)^2} > 0\) and \(a_1 + c_2 - \frac{3\alpha^2 \beta (\theta - 1)^2}{2(\alpha + \theta)^2} > 0\). Since \(g\) is coercive, it follows that the sequences \(\{y^k\}\), \(\{Ax^k - y^k - c - \frac{1}{\beta} \lambda^k\}\), \(\{\Delta y^k\}\) and \(\{\Delta \lambda^k\}\) are bounded. From (16), we obtain the boundedness of \(\{Ax^k\}\). Considering that \(A\) is of full column rank, we deduce that \(\{x^k\}\) is bounded. Consequently, \(\{\lambda^k\}\) is bounded and hence \(\{w^k\}\) is bounded.

4. The linear convergence of the symmetric ADMM. In this section, we further assume that the component function \(f\) in (1) is convex, and prove that the iterative sequence generated by symmetric ADMM converges locally to a KKT point of the nonconvex optimization problem in a R-linear rate; and the auxiliary sequence converges in a Q-linear rate.

Considering the optimality condition of iterative process (5), we have
\[
\begin{cases}
0 \in \partial f(x^{k+1}) - A^T \lambda^{k+1} - \beta A^T (Ax^{k+1} - y^k - c), \\
0 = \nabla g(y^{k+1}) + \lambda^{k+1} + (\theta - 1) \beta (Ax^{k+1} - y^{k+1} - c), \\
\lambda^{k+1} = \lambda^k - \theta \beta (Ax^{k+1} - y^{k+1} - c) - \alpha \beta (Ax^{k+1} - y^k - c).
\end{cases}
\] (35)
Lemma 4.1. Let \( \{w^k\} \) be the sequence generated by (5), then we have
\[
\|e(w^{k+1}, 1)\| \leq \sqrt{b_1\|\lambda^k - \lambda^{k+1}\|^2 + b_2\|y^{k+1} - y^k\|^2}, \quad \forall k \geq 0, \tag{36}
\]
where \( b_1, b_2 > 0. \)

Proof. In order to get the inequality (36), observing the definition of \( e(w, 1) \) and the nonexpansiveness property of the proximal operator, we have
\[
\|e_x(w^{k+1}, 1)\| = \|\lambda^k - \lambda^{k+1}\| = \|\lambda^k + A^T\lambda^{k+1} - y^{k+1} - c\| = \|\lambda^k + A^T\lambda^{k+1} - \lambda^{k+1} + A^T\lambda^{k+1} - y^{k+1} - c\|
\leq \|A\|\|\lambda^k - \lambda^{k+1}\| + \|y^{k+1} - y^{k+1}\| = \|A\|\|\lambda^k - \lambda^{k+1}\| + \|y^{k+1} - y^{k+1}\|,
\]
where the second equality follows from (35) and the property of the proximal operator, the first inequality follows from the nonexpansive property of the proximal operator, and the third equality follows from (16). Since \( 0 = \nabla g(y^{k+1}) + \lambda^{k+1} + (\theta - 1)\beta(Ax^{k+1} - y^{k+1} - c), \) we can deduce
\[
\|e_y(w^{k+1}, 1)\| = \|\nabla g(y^{k+1}) + \lambda^{k+1}\|
\leq \|\lambda^k - \lambda^{k+1}\| + \|y^{k+1} - y^{k+1}\|,
\]
where the third equality follows from (16). Since \( \lambda^{k+1} = \lambda^k - \theta\beta(Ax^{k+1} - y^{k+1} - c) = \alpha\beta(Ax^{k+1} - y^{k+1} - c), \) we get
\[
\|e_x(w^{k+1}, 1)\| = \|Ax^{k+1} - y^{k+1} - c\|
\leq \|\lambda^k - \lambda^{k+1}\| + \|y^{k+1} - y^{k+1}\|,
\]
where the second equality follows from (16). Combining (37), (38) and (39), we have
\[
\|e(w^{k+1}, 1)\| = \sqrt{\|e_x(w^{k+1}, 1)\|^2 + \|e_y(w^{k+1}, 1)\|^2 + \|e_x(w^{k+1}, 1)\|^2}
\leq \sqrt{b_1\|\lambda^k - \lambda^{k+1}\|^2 + b_2\|y^{k+1} - y^k\|^2},
\]
where \( b_1 := 2\left(1 - \frac{1}{\alpha + \theta}\right)^2\|A\|^2 + \frac{2(1 - \theta)^2}{(\alpha + \theta)^2} + \frac{2}{(\alpha + \theta)^2}\|A\|^2 \) and \( b_2 := \frac{2\alpha^2\beta^2(1 - \theta)^2}{(\alpha + \theta)^2} + \frac{2\alpha^2}{(\alpha + \theta)^2} \) in the last inequality.

Under the Assumption 2.5 and Assumption 2.7, we can prove the following lemma.

Lemma 4.2. Suppose that \( A \in \mathbb{R}^{m \times n} \) is of full column rank, and Assumption 3.1 is satisfied. Let \( \{w^k\} \) be the sequence generated by the symmetric ADMM (5) which is assumed to be bounded. Then, under the Assumption 2.5 and Assumption 2.7,
we have that

(i) \( \lim_{k \to +\infty} \text{dist}(w^k, \text{crit}L_\beta) = 0. \)

(ii) there exists a positive integer \( K_3 \), and for any \( k \geq K_3 \)

\[ L_\beta(\bar{w}^k) = \lim_{k \to +\infty} L_\beta(w^k) = \lim_{k \to +\infty} \inf_{k \in \mathbb{N}} F_k, \]

where \( \bar{w}^k \in \text{crit}L_\beta \) such that \( \|\bar{w}^k - w^k\| = \text{dist}(w^k, \text{crit}L_\beta) \).

**Proof.** (i) From Theorem 3.7, we know that \( \{F_k\} \) and \( \{L_\beta(w^k)\} \) are convergent, and \( \lim_{k \to +\infty} L_\beta(w^k) = \lim_{k \to +\infty} F_k \). Due to the parameter \( \sigma > 0 \), inequality (20) implies

\[ \|y^{k+1} - y^k\| \to 0, \|\lambda^{k+1} - \lambda^k\| \to 0. \quad (40) \]

Let \( \zeta = F_0 = F_\beta(x^0, y^0, \lambda^0, y^{-1}, \lambda^{-1}) \). Since \( \{F_k\} \) is nonincreasing by Lemma 3.6, we must have \( L_\beta(w^k) \leq F_k \leq F_0 = \zeta \) for all \( k \). Combining the inequalities (31), (36) and (40), we can deduce

\[ \|w^{k+1} - w^k\| \to 0, \|e(w^{k+1}, 1)\| \to 0. \quad (41) \]

In view of this and Assumption 2.5, there exist \( \tau > 0 \) and a positive integer \( K_1 \), so that for all \( k \geq K_1 \), we have

\[ \text{dist}(w^k, \text{crit}L_\beta) \leq \tau \|e(w^k, 1)\|, \forall k \geq K_1, \quad (42) \]

and then

\[ \lim_{k \to +\infty} \text{dist}(w^k, \text{crit}L_\beta) = 0. \]

(ii) For each \( k \), let \( \bar{w}^{k+1} = (\bar{x}^{k+1}, \bar{y}^{k+1}, \bar{\lambda}^{k+1}) \in \text{crit}L_\beta \), such that \( \|\bar{w}^{k+1} - w^{k+1}\| = \text{dist}(w^{k+1}, \text{crit}L_\beta) \). Then we have \( \|w^{k+1} - \bar{w}^{k+1}\| \to 0 \). According to (41), we can further deduce

\[ \|w^k - \bar{w}^{k+1}\| \to 0. \]

Under the Assumption 2.7, i.e., the separation of stationary values, there exists \( \delta > 0 \) such that \( L_\beta(\bar{w}) = L_\beta(\bar{w}) \), whenever \( \bar{w} = (\bar{x}, \bar{y}, \bar{\lambda}) \in \text{crit}L_\beta \) and \( \|\bar{w} - \bar{w}\| \leq \delta \). There must exist a positive integer \( K_2 \geq K_1 \) and a constant \( L_\beta^\infty \) such that

\[ L_\beta(w^{k+1}) = L_\beta(\bar{w}^k) = L_\beta^\infty, \forall k \geq K_2. \quad (43) \]

Following, we will analyze the value of the constant \( L_\beta^\infty \).

First, we will show that any accumulation point of \( \{w^k\} \) is a stationary point of \( L_\beta \) (i.e., the accumulation point belongs to \( \text{crit}L_\beta \)). Let \( w^* = (x^*, y^*, \lambda^*) \) be a cluster point of \( \{w^k = (x^k, y^k, \lambda^k)\} \), then there exist a subsequence \( \{w^{k^j} = (x^{k^j}, y^{k^j}, \lambda^{k^j})\} \) such that \( w^{k^j} \to w^* \). Since \( \|w^{k+1} - w^k\| \to 0 \), it is easy to check \( (x^{k^j+1}, y^{k^j+1}, \lambda^{k^j+1}) \) also converges to \( (x^*, y^*, \lambda^*) \). Considering the continuity of \( \nabla g \) and the closeness of \( \partial f \), by taking the limit in (35) along the sequence \( (x^{k^j+1}, y^{k^j+1}, \lambda^{k^j+1}) \), we have

\[
\begin{cases}
A^T \lambda^* \in \partial f(x^*), \\
\nabla g(y^*) = -\lambda^*, \\
Ax^* - y^* - c = 0.
\end{cases}
\]

Then, \( (x^*, y^*, \lambda^*) \) is a critical point of \( L_\beta \), hence \( w^* \in \text{crit}L_\beta \).

Following we will consider the value of \( L_\beta \) on the set of accumulation points. Observing the first subproblem of the iteration process (5), \( x^{k+1} \) is the minimizer for the the variable \( x \), so we have

\[ L_\beta(x^{k+1}, y^k, \lambda^k) \leq L_\beta(x^*, y^k, \lambda^k). \quad (44) \]
Theorem 4.3. Under the same assumption in Lemma 4.2, the sequence convergence of the sequence \( \{F_k\} \) with respect to \( y \) and \( \lambda \), we have

\[
\limsup_{j \to +\infty} L_\beta(x^{k_j+1}, y^{k_j}, \lambda^{k_j}) \leq \limsup_{j \to +\infty} L_\beta(x^*, y^*, \lambda^*) = L_\beta(x^*, y^*, \lambda^*). 
\]

Since \( \|w^{k+1} - w^k\| \to 0 \) and the boundedness of sequence, we obtain

\[
\limsup_{j \to +\infty} L_\beta(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}) \leq L_\beta(x^*, y^*, \lambda^*). 
\] (45)

On the other hand, from the lower semicontinuity of \( L_\beta(\cdot) \), we have

\[
L_\beta(x^*, y^*, \lambda^*) \leq \liminf_{j \to +\infty} L_\beta(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}). 
\] (46)

Combining (45) and (46), we have

\[
\lim_{j \to +\infty} L_\beta(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}) = L_\beta(x^*, y^*, \lambda^*). 
\]

Considering that the convergence of whole sequence \( \{L_\beta(w^k)\} \) have been proved in (1), and the subsequence \( L_\beta(w^{k_j+1}) \to L_\beta(w^*) \), so we can deduce \( L_\beta(w^k) \to L_\beta(w^*) \) and \( F_k \to F_\beta(w^*, y^*, \lambda^*) = L_\beta(w^*) \). Therefore, \( L_\beta(\cdot) \) is constant on the set of accumulation points. Obviously, due to the sequence \( \{F_k\} \) is nonincreasing, we have \( L_\beta(w^k) = F_\beta(w^*, y^*, \lambda^*) = \lim_{k \to +\infty} F_k = \inf_{k \in N} F_k \).

Due to \( \|w^* - \bar{w}^{k_j+1}\| \leq \|w^* - \bar{w}^{k_j}\| + \|\bar{w}^{k_j} - \bar{w}^{k_j+1}\| \), we can deduce \( \|w^* - \bar{w}^{k_j+1}\| \to 0 \). According to the Assumption 2.7, there exists \( \delta > 0 \) such that \( L_\beta(w^*) = L_\beta(\bar{w}^{k_j+1}) \) when \( \|w^* - \bar{w}^{k_j+1}\| \leq \delta \), i.e., there exists a positive integer \( K_3 \geq K_2 \) such that \( L_\beta(w^*) = L_\beta(\bar{w}^{k_j+1}), \forall k_{j+1} \geq K_3 \). Together with (43), we can deduce that the value of \( L_\beta^\infty \), i.e.,

\[
L_\beta(\bar{w}^k) = L_\beta^\infty = L_\beta(w^*) = F_\beta(w^*, y^*, \lambda^*) = \inf_{k \in N} F_k, \forall k \geq K_3.
\]

Now, we are ready to prove the Q-linear convergence of \( \{F_k\} \) and R-linear convergence of the sequence \( \{w^k\} \).

\textbf{Theorem 4.3.} Under the same assumption in Lemma 4.2, the sequence \( \{F_k\} \) is Q-linearly convergent.

\textit{Proof.} Recall that \( \|\bar{w}^{k+1} - w^{k+1}\| = \text{dist}(w^{k+1}, \text{crit}L_\beta) \), then we have \( \|\bar{y}^{k+1} - y^{k+1}\| \leq \text{dist}(w^{k+1}, \text{crit}L_\beta), \|\bar{x}^{k+1} - x^{k+1}\| \leq \text{dist}(w^{k+1}, \text{crit}L_\beta), \) and \( A\bar{x}^{k+1} - B\bar{y}^{k+1} - c = 0 \). Together with (35), for all \( k \geq 0 \), we have
\[ L_\beta(w^{k+1}) - L_\beta(w^{k+1}) \]
\[ = \Big( f(x^{k+1}) + g(y^{k+1}) - \langle \lambda^{k+1}, Ax^{k+1} - y^{k+1} - c \rangle + \frac{\alpha}{2} \| Ax^{k+1} - y^{k+1} - c \|^2 \Big) \]
\[ + f(x^{k+1}) + g(y^{k+1}) - \langle \lambda^{k+1}, Ax^{k+1} - y^{k+1} - c \rangle + \frac{\alpha}{2} \| Ax^{k+1} - y^{k+1} - c \|^2 \]
\[ \leq \langle \lambda^{k} - \beta(Ax^{k+1} - y^{k+1}), Ax^{k+1} + Az^{k+1} \rangle - \langle \lambda^{k+1} + (\theta - 1)\beta(Ax^{k+1} - y^{k+1} - c), y^{k+1} - y^{k+1} \rangle \]
\[ + \frac{L}{2} \| y^{k+1} - y^{k+1} \| - \langle \lambda^{k+1}, Ax^{k+1} - y^{k+1} - c \rangle + \frac{\alpha}{2} \| Ax^{k+1} - y^{k+1} - c \|^2, \]
\[ = \langle \lambda^{k+1} + y^{k+1} + Ax^{k+1} - y^{k+1} - c \rangle + \langle \lambda^{k}, Ax^{k+1} - y^{k+1} \rangle - (\theta - 1)\beta(Ax^{k+1} - y^{k+1} - c), y^{k+1} - y^{k+1} \rangle \]
\[ + \frac{\alpha}{2} \| Ax^{k+1} - y^{k+1} - c \|^2 + \frac{L}{2} \| y^{k+1} - y^{k+1} \|^2, \]
where the first inequality follows from the convexity of \( f \) and the Lipschitz continuity of \( \nabla g \). Considering that \( (\alpha + \theta)\beta(Ax^{k+1} = \lambda^{k+1} + c, Ax^{k+1} - y^{k+1} + c) \), we can obtain the following two equalities
\[ Ax^{k+1} - y^{k+1} = \frac{1}{(\alpha + \theta)\beta} (\lambda^{k} - \lambda^{k+1}) + \frac{\alpha}{\alpha + \theta} (y^{k} - y^{k+1}) + \frac{\theta}{\alpha + \theta} (y^{k+1} - y^{k+1}), \]
\[ Ax^{k+1} - y^{k+1} = \frac{1}{(\alpha + \theta)\beta} (\lambda^{k} - \lambda^{k+1}) + \frac{\alpha}{\alpha + \theta} (y^{k} - y^{k+1}). \]
Due to \( L_\beta(w^{k+1}) - L_\beta(w^{k+1}) \)
\[ \leq t_1 \langle \lambda^{k} - \lambda^{k+1}, y^{k} - y^{k+1} \rangle + t_2 \langle \lambda^{k} - \lambda^{k+1}, y^{k+1} - y^{k+1} \rangle + t_3 \langle y^{k} - y^{k+1}, y^{k+1} - y^{k+1} \rangle \]
\[ + t_4 \| \lambda^{k} - \lambda^{k+1} \|^2 + t_5 y^{k+1} - y^{k+1} \|^2 + \frac{L}{2} \| y^{k+1} - y^{k+1} \|^2, \]
where the parameters are \( t_1 = \frac{\alpha}{\alpha + \theta} \left[ \frac{\alpha}{\alpha + \theta} + \frac{\alpha}{\alpha + \theta} \right], t_2 = \left( \frac{1}{\alpha + \theta} \right) \left[ \frac{\alpha}{\alpha + \theta} + \frac{\alpha}{\alpha + \theta} \right], t_3 = \frac{\alpha\theta}{\alpha + \theta}, t_4 = \frac{\alpha}{\alpha + \theta}, t_5 = \frac{\alpha}{\alpha + \theta}, \) and \( t_7 = \frac{\alpha\theta}{2(\alpha + \theta)}. \) Observe the equality \( \text{dist}^2(w^{k+1}, \text{crit} L_\beta) = \| y^{k+1} - y^{k+1} \|^2, \) and setting \( t_1 = \frac{1}{2} (t_1 + \)
of \(L\) generated by the symmetric ADMM (5).

Theorem 4.4. Under the same assumption in Lemma 4.2, the sequence \(\{w^k\}\) generated by the symmetric ADMM (5) is R-linearly convergent to a critical point of \(L_\beta(\cdot)\).
Proof. In view of Theorem 4.3, the sequence \( \{F_k\} \) is Q-linearly convergent. We obtain from (20) that

\[
\|y^{k+1} - y^k\|^2 \leq \frac{1}{a_1} (F_k - F_{k+1}) \\
\leq \frac{1}{a_1} \left( F_k - \inf_{k \in N} F_k \right),
\]

(53)

\[
\|
\lambda^{k+1} - \lambda^k\|^2 \leq \frac{1}{c_1} \left( \frac{1}{2} - \frac{c_2}{2+\sigma} \right) (F_k - F_{k+1}) \\
\leq \frac{1}{c_1} \left( \frac{1}{2} - \frac{c_2}{2+\sigma} \right) \left( F_k - \inf_{k \in N} F_k \right),
\]

(54)

where the last inequalities in (53) and (54) follow from the fact that the sequence \( \{F_k\} \) is nonincreasing and converges to \( \inf_{k \in N} F_k \). Using the above inequality and the fact that the sequence \( \{F_k\} \) is Q-linearly convergent, we see that there exist \( 0 < q < 1 \) and \( M_1, M_2 > 0 \) such that

\[
\|y^{k+1} - y^k\| \leq M_1 q^k, \quad \forall k \geq 0,
\]

(55)

\[
\|
\lambda^{k+1} - \lambda^k\| \leq M_2 q^k, \quad \forall k \geq 0.
\]

(56)

This, together with (31), yields

\[
\|x^{k+1} - x^k\| \leq \frac{4 \left( \frac{1}{q} + 1 \right)}{(\alpha + \theta)^2 \beta \lambda_{\min}(A^T A)} M_2 q^k + \frac{4 \alpha^2}{q} + 4 \theta^2 M_1 q^k \\
= \left( \frac{4 M_2 \left( \frac{1}{q} + 1 \right)}{(\alpha + \theta)^2 \beta \lambda_{\min}(A^T A)} + \frac{4 M_1 \left( \frac{\alpha^2}{q} + \theta^2 \right)}{(\alpha + \theta)^2 \lambda_{\min}(A^T A)} \right) q^k, \quad \forall k \geq 0.
\]

(57)

Combining (55), (56) and (57), we obtain

\[
\|w^{k+1} - w^k\| \leq \bar{M} q^k, \quad \forall k \geq 0,
\]

where \( \bar{M} = M_1 + M_2 + \frac{4 M_2 \left( \frac{1}{q} + 1 \right)}{(\alpha + \theta)^2 \beta \lambda_{\min}(A^T A)} + \frac{4 M_1 \left( \frac{\alpha^2}{q} + \theta^2 \right)}{(\alpha + \theta)^2 \lambda_{\min}(A^T A)} > 0 \). Consequently, for any \( m_2 > m_1 \geq 1 \), we have

\[
\|w^{m_2} - w^{m_1}\| \leq \sum_{k=m_1}^{m_2-1} \|w^{k+1} - w^k\| \leq \frac{\bar{M}}{1 - q} q^{m_1},
\]

this implies that \( \{w^k\} \) is a Cauchy sequence and hence convergent. Denoting its limit by \( \bar{w} \) and passing to the limit as \( m_2 \to \infty \) in the above relation, we see further that

\[
\|w^{m_1} - \bar{w}\| \leq \frac{\bar{M}}{1 - q} q^{m_1}.
\]

This means that the sequence \( \{w^k\} \) is R-linearly convergent to its limit, which is a stationary point of \( L_{\beta} \) according to Theorem 4.3. This completes the proof. \( \square \)
 Remark 5.1. In the Algorithm 1, the choice of $d_1^{k+1}$ in Step 1 is as follows:

$$d_1^{k+1} = \xi^{k+1} - \eta^{k+1},$$

where $\eta^{k+1} = A^T \lambda^k - \beta A^T (Ax^{k+1} - By^k - c)$ and $\xi^{k+1} = P_{\partial f(x^{k+1})}(\eta^{k+1})$. 

5. The linear convergence of the inexact symmetric ADMM. In this section, we give an inexact version of the symmetric ADMM (5) for solving the two block nonconvex separable problems (1). When the component function $f$ in (1) is convex, we establish the convergence properties and convergence rate of the inexact symmetric ADMM under the error bound condition (described in Assumption 2.5) and the separability of stationary values (described in Assumption 2.7). Specially, both subproblems in (5) are allowed to be solved inexactly by certain relative error criteria, the scheme is as follows

$$\begin{align*}
\lambda^{k+\frac{1}{2}} &= \lambda^k - \alpha \beta (Ax^{k+1} - y^k - c), \\
y^{k+1} &\approx y^{k+1}_{\text{exact}} \in \arg\min_y L_\beta(x^{k+1}, y, \lambda^{k+\frac{1}{2}}), \\
\lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \theta \beta (Ax^{k+1} - y^{k+1} - c).
\end{align*}$$

(58)

Obviously, if we take $x^{k+1} = x^{k+1}_{\text{exact}}$ and $y^{k+1} = y^{k+1}_{\text{exact}}$ for each $k \geq 0$, the corresponding algorithm (58) is the symmetric ADMM (5). In the following, we present an inexact symmetric ADMM with relative error criteria, i.e., Algorithm 1, where the accuracy of the approximate solution to each of the subproblems only needs the subgradient information at the candidate solutions, and only two non-negative constants, i.e., $\sigma_1, \sigma_2$, are needed to control the error tolerance.

**Algorithm 1** An inexact symmetric ADMM with relative error criteria

Let $\beta > 0$ satisfied the Assumption 3.1 (3), and $\sigma_1, \sigma_2 \geq 0$ be given parameters. Choose $x^0 \in \text{dom} f$, $y^0 \in \text{dom} g$, $\lambda^0 \in \mathbb{R}^m$. For $k = 0, 1, ...$

**Step 1.** Compute $x^{k+1} \approx \arg\min_x L_\beta(x, y^k, \lambda^k)$, such that

$$\left| \langle d_1^{k+1}, x^k - x^{k+1} \rangle \right| + \|d_1^{k+1}\|^2 \leq \sigma_1 \|x^k - x^{k+1}\|^2,$$

with $d_1^{k+1} \in \partial_x L_\beta(x^{k+1}, y^k, \lambda^k)$.

**Step 2.** Compute

$$\lambda^{k+\frac{1}{2}} = \lambda^k - \alpha \beta (Ax^{k+1} - y^k - c).$$

**Step 3.** Compute $y^{k+1} \approx \arg\min_y L_\beta(x^{k+1}, y, \lambda^{k+\frac{1}{2}})$, such that

$$\left| \langle d_2^{k+1}, y^k - y^{k+1} \rangle \right| + \frac{1}{\beta} \|d_2^{k+1}\|^2 + \|d_2^{k+1}\|^2 \leq \sigma_2 \|y^k - y^{k+1}\|^2,$$

with $d_2^{k+1} = \nabla_y L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+\frac{1}{2}})$.

**Step 4.** Compute

$$\lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \theta \beta (Ax^{k+1} - y^{k+1} - c).$$

**Remark 5.1.** In the Algorithm 1, the choice of $d_1^{k+1}$ in Step 1 is as follows:

$$d_1^{k+1} = \xi^{k+1} - \eta^{k+1},$$

where $\eta^{k+1} = A^T \lambda^k - \beta A^T (Ax^{k+1} - By^k - c)$ and $\xi^{k+1} = P_{\partial f(x^{k+1})}(\eta^{k+1})$. 

Considering the optimality condition of iteration in Algorithm 1, we have
\[
\begin{aligned}
\begin{cases}
d^{k+1}_1 & \in \partial f(x^{k+1}) - AT\lambda^k + \beta AT(Ax^{k+1} - y^k - c), \\
\lambda^{k+\frac{1}{2}} & = \lambda^k - \alpha\beta(Ax^{k+1} - y^k - c), \\
d^{k+1}_2 & = \nabla g(y^{k+1}) + \lambda^{k+\frac{1}{2}} - \beta(Ax^{k+1} - y^{k+1} - c), \\
\lambda^{k+1} & = \lambda^{k+\frac{1}{2}} - \theta\beta(Ax^{k+1} - y^{k+1} - c).
\end{cases}
\end{aligned}
\]

Using the second and fourth equalities and rearranging terms, we obtain
\[
\begin{aligned}
\begin{cases}
d^{k+1}_1 & = \partial f(x^{k+1}) - AT\lambda^k + \beta AT(Ax^{k+1} - y^k - c), \\
d^{k+1}_2 & = \nabla g(y^{k+1}) + \lambda^{k+1} + \beta(Ax^{k+1} - y^{k+1} - c), \\
\lambda^{k+1} & = \lambda^k - \alpha\beta(Ax^{k+1} - y^k - c) - \theta\beta(Ax^{k+1} - y^{k+1} - c).
\end{cases}
\end{aligned}
\]

Lemma 5.2. For the sequence \(\{x^k, y^k, \lambda^k\}\) generated by Algorithm 1, the value of the augmented Lagrangian function \(L_\beta\) defined in (3).

(i) \(L_\beta(x^{k+1}, y^k, \lambda^k) \geq L_\beta(x^{k+1}, y^{k+1}, \lambda^k) - \langle d^{k+1}_1, \Delta x^{k+1} \rangle + \frac{\beta}{2}\lambda_{\min}(ATA)\|\Delta x^{k+1}\|^2;\)

(ii) \(L_\beta(x^{k+1}, y^{k+1}, \lambda^k) \geq L_\beta(x^{k+1}, y^{k+1}, \lambda^k) - \alpha\beta(Ax^{k+1} - y^{k+1} - c, \Delta y^{k+1}) - \langle d^{k+1}_2, \Delta y^{k+1} \rangle + \frac{\beta}{2}\lambda_{\min}(ATA)\|\Delta y^{k+1}\|^2;\)

(iii) \(L_\beta(x^{k+1}, y^{k+1}, \lambda^k) = L_\beta(x^{k+1}, y^{k+1}, \lambda^k) - \langle \Delta \lambda^{k+1}, Ax^{k+1} - y^{k+1} - c \rangle, \forall k \geq 0.\)

Proof. (i) Considering the convexity of \(f\) and the update of \(x\) in Algorithm 1, we have
\[
\begin{aligned}
L_\beta(x^{k+1}, y^k, \lambda^k) & \geq L_\beta(x^{k+1}, y^{k+1}, \lambda^k) + \langle \nabla L_\beta(x^{k+1}, y^{k}, \lambda^k), x^{k+1} - x^{k+1} \rangle + \frac{\beta}{2}\lambda_{\min}(ATA)\|x^{k+1} - x^{k+1}\|^2 \\
& = L_\beta(x^{k+1}, y^{k}, \lambda^k) + \langle d^{k+1}_1, x^{k+1} - x^{k+1} \rangle + \frac{\beta}{2}\lambda_{\min}(ATA)\|x^{k+1} - x^{k+1}\|^2.
\end{aligned}
\]

Combining with the identity \(\Delta x^{k+1} = x^{k+1} - x^k\), (i) is proved.

(ii) The proof is similar to the analysis in Lemma 3.3 (ii). The equality (12) becomes
\[
\nabla_y L_\beta(x^{k+1}, y^{k+1}, \lambda^k) = \alpha\beta(Ax^{k+1} - y^{k+1} - c) + d^{k+1}_2.
\]

Substitute (62) into (11), we can deduce
\[
\begin{aligned}
L_\beta(x^{k+1}, y^k, \lambda^k) & \geq L_\beta(x^{k+1}, y^{k+1}, \lambda^k) + \frac{\beta - L}{2}\|y^k - y^{k+1}\|^2 \\
& - \alpha\beta(Ax^{k+1} - y^{k+1} - c, y^{k+1} - y^{k+1}) - \langle d^{k+1}_2, y^{k+1} - y^{k+1} \rangle - \langle d^{k+1}_2, y^{k+1} - y^{k+1} \rangle.
\end{aligned}
\]

Thus (ii) is proved with the identity \(\Delta y^{k+1} = y^{k+1} - y^{k}\).

(iii) The proof of the result is the same as that in Lemma 3.3 (iii).

The following lemma provides the new recursive relation of the sequence \(\{\Delta \lambda^k\}\) in Algorithm 1, which is different from the exact case (Lemma 3.4).

Lemma 5.3. For any given \((y^0, \lambda^0) \in \mathcal{R}^m \times \mathcal{R}^m\), let \(\Delta y^0\) and \(\Delta \lambda^0\) satisfy the equality (13). Then, for the sequence \(\{(x^k, y^k, \lambda^k)\}\) generated by Algorithm 1, we have
\[
\Delta \lambda^{k+1} = -\frac{\theta - 1}{\alpha + 1}\Delta \lambda^k + u^k + \frac{\alpha + \theta}{\alpha + 1}\Delta d^{k+1}_2, \quad \forall k \geq 0,
\]
where \(u^k\) is defined as (15).
Proof. Using the same technique as that in Lemma 3.4, the equality (17) is changed to
\[ d_2^{k+1} = \nabla g(y^{k+1}) + \lambda^{k+1} + (\theta - 1)\beta(Ax^{k+1} - y^{k+1} - c). \]  
(64)

Substituting (16) into (64), we have
\[ d_2^{k+1} - \frac{\alpha + 1}{\alpha + \theta} \lambda^{k+1} = \nabla g(y^{k+1}) + \frac{\theta - 1}{\alpha + \theta} \lambda^k - \frac{\alpha \beta (\theta - 1)}{\alpha + \theta} (y^{k+1} - y^k). \]  
(65)

Thus we can get
\[ \Delta d_2^{k+1} - \frac{\alpha + 1}{\alpha + \theta} \Delta \lambda^{k+1} = \left( \nabla g(y^{k+1}) - \nabla g(y^k) \right) + \frac{\theta - 1}{\alpha + \theta} \Delta \lambda^k - \frac{\alpha \beta (\theta - 1)}{\alpha + \theta} \left( \Delta y^{k+1} - \Delta y^k \right). \]

And hence, in view of (15), we deduce the equality (63) for every k ≥ 1.

Similar to Lemma 3.4, we can deduce (63) also holds for k = 0.

Now we turn to the convergence analysis of Algorithm 1. The following lemma plays the same role as Lemma 3.6, and the proof technique is similar.

Lemma 5.4. For the sequence \{x^k, y^k, \lambda^k\} generated by Algorithm 1, under the Assumption 3.1, we have
\[ F_{k+1} \leq F_k - \tilde{a}_1 \|\Delta y^{k+1}\|^2 - \left( \frac{C_1}{2} - \frac{C_1}{2 + \sigma} \right) \|\Delta \lambda^{k+1}\|^2 - \tilde{a}_2 \|\Delta x^{k+1}\|^2, \forall k \geq 0, \]  
(66)

where \( C_1 = \frac{\|\theta + 1\|}{(\alpha + \theta)(\frac{\theta + 1}{\alpha + \theta} - 1)} \geq 0, \tilde{a}_2 = \frac{\sigma}{2} \gamma_{\min}(A^T A) - \sigma_1 > 0, \gamma = \left( \frac{1 + \gamma}{\alpha + \gamma} - 1 \right) \left( \frac{\theta + 1}{\alpha + \theta} - \frac{1}{\alpha + \gamma} \right), \)
\[ c_2 = \frac{6\alpha^2 \beta - \frac{\theta + 1}{\alpha + \gamma}}{(\alpha + \gamma)^2} \geq 0, \tilde{a}_1 = \left( \frac{1}{2} - \frac{\alpha \theta}{\alpha + \gamma} - \frac{12\alpha^2 \frac{\theta + 1}{\alpha + \gamma}}{(\alpha + \gamma)^2} \right) \beta - \frac{6\alpha^2 \beta - \frac{\theta + 1}{\alpha + \gamma}}{(\alpha + \gamma)^2} \frac{1}{\beta} - \frac{\sigma}{2} (1 + \tau_0) \sigma_2 > 0, \tau_0 = \frac{2(\alpha + \gamma) \frac{\theta + 1}{\alpha + \gamma}}{(1 + \gamma)^2} > 0, \]  
and \( \sigma > 0. \)

Proof. Similar to the proof of Lemma 3.6, combining (i), (ii) and (iii) in Lemma 5.2, we have
\[ L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) \leq L_\beta(x^k, y^k, \lambda^k) + \alpha \beta \left( \langle Ax^{k+1} - y^k - c, \Delta y^{k+1} \rangle + \langle \Delta \lambda^{k+1}, Ax^{k+1} - y^{k+1} - c \rangle + \frac{\beta - L}{2} \|\Delta y^{k+1}\|^2 + \|\Delta x^{k+1}\|^2 + \|d_1^{k+1}, \Delta x^{k+1}\| \right). \]  
(67)

Substituting (16) into (67), we can deduce
\[ L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) \leq L_\beta(x^k, y^k, \lambda^k) - \left( \frac{\beta - L}{2} - \frac{\alpha \beta \theta}{\alpha + \gamma} \sigma_2 \right) \|\Delta y^{k+1}\|^2 - \left( \frac{\beta}{2} \gamma_{\min}(A^T A) - \sigma_1 \right) \|\Delta x^{k+1}\|^2 \]
\[ + \frac{1}{(\alpha + \gamma)\beta} \|\Delta \lambda^{k+1}\|^2, \]  
(68)

which also using the inequalities (59) and (60). Considering the equality (63) and the relation \( \|s_1 + s_2\|^2 \leq (1 + t)\|s_1\|^2 + (1 + \frac{1}{2})\|s_2\|^2, \forall s_1, s_2 \in \mathbb{R}^n, t > 0, \) the
inequality (68) becomes
\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2} \|\Delta \lambda^{k+1}\|^2 \\
\leq L_\beta(x^k, y^k, \lambda^k) - \left( \frac{\beta - L}{2} - \frac{\alpha \theta}{\alpha + \theta - \sigma_2} \right) \|\Delta y^{k+1}\|^2 - \left( \frac{c_1}{2} - \frac{c_1}{2 + \sigma} \right) \|\Delta \lambda^{k+1}\|^2 \\
+ \left( \frac{c_1}{2} + \frac{1}{(\alpha + \theta)\beta} \right) \left( \frac{\theta}{\alpha + 1} \right)^2 (1 + t) \|\Delta \lambda^k\|^2 + \left( 1 + \frac{1}{t} \right) \left\|u_k + \frac{\alpha + \theta}{\alpha + 1} \Delta d^{k+1}_2 \right\|^2.
\]
(69)

where \(c_1\) is a non-negative constant. Taking \((t + 1)\left|\frac{\theta - 1}{\alpha + \theta}\right| = 1\), \(c_1 = \frac{|\frac{\theta - 1}{\alpha + \theta}|}{(\alpha + \theta)\beta}\) ≥ 0, \(\sigma > 0\), we rearrange the inequality (69)
\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^{k+1}\|^2 \\
\leq L_\beta(x^k, y^k, \lambda^k) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^k\|^2 - \left( \frac{c_1}{2} - \frac{c_1}{2 + \sigma} \right) \|\Delta \lambda^{k+1}\|^2 \\
- \left( \frac{\beta - L}{2} - \frac{\alpha \theta}{\alpha + \theta - \sigma_2} \right) \|\Delta y^{k+1}\|^2 - \left( \frac{\beta}{2} \lambda_{\min}(A^T A) - \sigma_1 \right) \|\Delta x^{k+1}\|^2 \\
+ \frac{2}{2 + \sigma} \frac{1 + \alpha}{\theta - 1} \left( \frac{\alpha}{\alpha + \theta} \right)^2 \|\Delta \theta_2^{k+1}\|^2.
\]
(70)

Combining (70) with (24), we get
\[
L_\beta(x^{k+1}, y^{k+1}, \lambda^{k+1}) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^{k+1}\|^2 + c_2 \|\Delta y^{k+1}\|^2 \\
\leq L_\beta(x^k, y^k, \lambda^k) + \frac{c_1}{2 + \sigma} \|\Delta \lambda^k\|^2 + c_2 \|\Delta y^k\|^2 \\
- \tilde{a}_1 \|\Delta y^{k+1}\|^2 - a_2 \|\Delta y^k\|^2 - \left( \frac{c_1}{2} - \frac{c_1}{2 + \sigma} \right) \|\Delta \lambda^{k+1}\|^2 - \tilde{a}_2 \|\Delta x^{k+1}\|^2,
\]

where \(c_2\) is a non-negative constant, \(\gamma = (\|\frac{\alpha}{\alpha + \theta}\| - 1) \left( \frac{\theta}{\alpha + \theta} - \|\frac{\theta - 1}{\alpha + \theta}\| \right)\), \(\tau_0 = \frac{2(\alpha + \theta)}{(\alpha + \theta)^3 \gamma} \left( \frac{\theta - 1}{\alpha + \theta}\right) > 0\).

\[\tilde{a}_1 = \left( \frac{1}{2} - \frac{\alpha \theta}{\alpha + \theta - \sigma_2} \right) \left( \frac{6\alpha^2 \beta}{(\alpha + \theta)^2} \right) \|\Delta \theta^{k+1}\|^2 - \frac{6\alpha^2 \beta \gamma}{(\alpha + \theta)^2} \|\Delta \theta^{k+1}\|^2 - \frac{1}{\beta} - c_2 - \frac{\beta}{2} - (1 + \tau_0) \sigma_2,
\]
a_2 = c_2 - \frac{6\alpha^2 \beta \gamma}{(\alpha + \theta)^2} \|\Delta \theta^{k+1}\|^2,\] and \(\tilde{a}_2 = \frac{\beta}{2} \lambda_{\min}(A^T A) - \sigma_1\). Under the Assumption 3.1, there must exist a non-negative \(c_2\) such that \(\tilde{a}_1 > 0\), \(\tilde{a}_2 > 0\), and \(a_2 \geq 0\) can be satisfied. For brevity, we take \(c_2 = \frac{6\alpha^2 \beta \gamma}{(\alpha + \theta)^2} \|\Delta \theta^{k+1}\|^2 \geq 0\). Together with the definition of \(F_k\), the lemma is proved.

Following we will give a similar result as Lemma 4.1 and a brief proof will be given.

**Lemma 5.5.** Let \(\{w^k\}\) be the sequence generated by Algorithm 1, then we have
\[
\|e(w^{k+1}, 1)\| \leq \sqrt{b_1 \|\lambda^k - \lambda^{k+1}\|^2 + \tilde{b}_2 \|y^{k+1} - y^k\|^2 + 3\sigma_1 \|x^{k+1} - x^k\|^2}, \quad \forall k \geq 0,
\]
(71)
where \(b_1, \tilde{b}_2 > 0\).
Proof. Different from the inequality (37) in Lemma 4.1, we have 
\[
\|e_x(w^{k+1}, 1)\| = \|x^{k+1} - \text{Prox}_f(x^{k+1} + A^T x^{k+1})\|
\]
\[
= \|\text{Prox}_f(x^{k+1} + A^T x^{k+1} - \beta A^T (A x^{k+1} - y_k - c)) - \text{Prox}_f(x^{k+1} + A^T x^{k+1})\|
\]
\[
\leq \|x^{k+1} + A^T (\lambda - \lambda^{k+1}) - \beta A^T (A x^{k+1} - y_k - c)\|
\]
\[
= \|x^{k+1} + A^T (1 - \frac{1}{\alpha + \theta})(\lambda - \lambda^{k+1}) + \frac{\theta}{\alpha + \theta} A^T (y_k - y^{k+1})\|
\]
\[
\leq \sqrt{\sigma_1}\|x^{k+1} - x_k\| + \|A\|\|\lambda - \lambda^{k+1}\| + \|\lambda\|\|y_k - y^{k+1}\|.
\] (72)
where the second equality follows from (61), the third equality follows from (16), and the last inequality follows from (59). Since \(d^{k+1}_2 = \nabla g(y^{k+1}) + \lambda^{k+1} + (\theta - 1)\beta(A x^{k+1} - y^{k+1} - c)\) in Algorithm 1, it follows that 
\[
\|e_y(w^{k+1}, 1)\| = \|(1 - \theta)\beta(A x^{k+1} - y^{k+1} - c) + d^{k+1}_2\|
\]
\[
\leq |1 - \theta| \beta \left| 1 - \frac{\lambda - \lambda^{k+1}}{\alpha + \theta} \right| + |\frac{\alpha - \lambda^{k+1}}{\alpha + \theta}| + |d^{k+1}_2|
\]
\[
\leq \left| 1 - \frac{\lambda - \lambda^{k+1}}{\alpha + \theta} \right| \|y^{k+1} - y^{k+1}\| + \|y^{k+1} - y^{k+1}\| + \sqrt{\sigma_2}\|y^{k+1} - y^{k+1}\|,
\] (73)
where the last inequality follows from (60). Combining (72), (73) and (39), we have 
\[
\|e(w^{k+1}, 1)\| \leq \sqrt{\sigma_1}\|\lambda^{k+1} - \lambda^{k+1}\|^2 + \tilde{b}_2\|y^{k+1} - y^{k+1}\|^2 + 3\sigma_1\|x^{k+1} - x_k\|^2,
\]
where parameters \(b_1 := 3\left(1 - \frac{1}{\alpha + \theta}\right)^2\) \(\|A\|^2 + \frac{2(1 - \theta)^2}{\alpha + \theta} + \frac{2}{\alpha + \theta^2}\), \(\tilde{b}_2 := \frac{2\sigma_1^2}{\alpha + \theta} + \frac{2\sigma_2^2}{\alpha + \theta^2}\) in the last inequality.

Under the Assumption 2.5 and Assumption 2.7, we can prove the following lemma similar to Lemma 4.2.

**Lemma 5.6.** Suppose that \(A \in \mathbb{R}^{m \times n}\) is of full column rank, and Assumption 3.1 is satisfied. Let \(\{w^k\}\) be the sequence generated by Algorithm 1 which is assumed to be bounded. Then, under the Assumption 2.5 and Assumption 2.7, we have that 
(i) \(\lim\limits_{k \to +\infty} \text{dist}(w^k, \text{crit} L_\beta) = 0\).
(ii) there exists a positive integer \(K_3\), for any \(k \geq K_3\), it holds that 
\[
L_\beta(w^k) = \lim\limits_{k \to +\infty} L_\beta(w^k) = \lim\limits_{k \to +\infty} F_k = \inf\limits_{k \in N} F_k,
\]
where \(w^k \in \text{crit} L_\beta\) so that \(\|w^k - w^k\| = \text{dist}(w^k, \text{crit} L_\beta)\).

Proof. (i) \(\{F_k\}\) is convergent due to the boundness of \(\{w^k\}\). Considering that the parameters \(\tilde{a}_1, \tilde{a}_2, \sigma > 0\), inequality (66) implies 
\[
\|x^{k+1} - x_k\| \to 0, \|y^{k+1} - y_k\| \to 0, \|\lambda^{k+1} - \lambda^k\| \to 0.
\]
Thus, we have \(\|w^{k+1} - w^k\| \to 0\), and \(\|e(w^{k+1}, 1)\| \to 0\). Since \(\{F_k\}\) is nonincreasing by Lemma 5.4, together with the Assumption 2.5, similar to the proof of Lemma 4.2 (i), we can also deduce (42) for Algorithm 1. Then \(\lim\limits_{k \to +\infty} \text{dist}(w^k, \text{crit} L_\beta) = 0\).

(ii) For each \(k\), let \(\tilde{w}^{k+1} = (\tilde{x}^{k+1}, \tilde{y}^{k+1}, \tilde{\lambda}^{k+1}) \in \text{crit} L_\beta\), such that \(\|\tilde{w}^{k+1} - w^{k+1}\| = \text{dist}(w^{k+1}, \text{crit} L_\beta)\). The same proof as Lemma 4.2 (ii), we can also deduce the equality (43).
Due to the relative error criteria (59) and (60), we get
\[ \|d_2^{k+1} - d_2^k\|^2 \to 0, \quad \|d_2^k\|^2 \to 0, \quad \|d_2^k\|^2 \to 0. \]
Similar to the analysis in Lemma 4.2 (ii), considering the continuity of \( \nabla g \) and the closeness of \( \partial f \), together with \( d_1^k \to 0 \), \( d_2^k \to 0 \), we can deduce that any accumulation point \( w^* = (x^*, y^*, \lambda^*) \) of \( \{w^k\} \) is a critical point of \( L_\beta \) by taking the limit in (61), hence \( w^* \in \text{crit} L_\beta \).

Following we will consider the value of \( L_\beta \) on the set of accumulation points. Considering the convexity of the augmented Lagrangian function \( L_\beta(x, y^k, \lambda^k) \) with respect to \( x \), we have
\[ L_\beta(x^{k+1}, y^k, \lambda^k) + \langle d_1^{k+1}, x^* - x^{k+1} \rangle \leq L_\beta(x^*, y^k, \lambda^k). \] (74)

Setting \( k = k_j \) in the inequality (74), together with the continuity of \( L_\beta(\cdot) \) with respect to \( y \) and \( \lambda \), we have
\[ \limsup_{j \to +\infty} L_\beta(x^{k_j+1}, y^{k_j}, \lambda^{k_j}) + \limsup_{j \to +\infty} (d_1^{k_j+1}, x^* - x^{k_j+1}) \leq \limsup_{j \to +\infty} L_\beta(x^*, y^{k_j}, \lambda^{k_j}) = L_\beta(x^*, y^*, \lambda^*). \]

Since \( \|w^{k+1} - w^k\| \to 0 \), \( d_1^k \to 0 \), and the boundedness of sequence, we obtain
\[ \limsup_{j \to +\infty} L_\beta(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}) \leq L_\beta(x^*, y^*, \lambda^*). \] (75)
Together with the lower semicontinuous of \( L_\beta \), i.e.,
\[ L_\beta(x^*, y^*, \lambda^*) \leq \liminf_{j \to +\infty} L_\beta(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}), \]
we have
\[ \lim_{j \to +\infty} L_\beta(x^{k_j+1}, y^{k_j+1}, \lambda^{k_j+1}) = L_\beta(x^*, y^*, \lambda^*). \]
The following proof is the same as Lemma 4.2 (ii), here we omit it and just give the result
\[ L_\beta(w^k) = L_\beta^\infty = L_\beta(w^*) = F_\beta(w^*, y^*, \lambda^*) = \inf_{k \in \mathbb{N}} F_k, \quad \forall k \geq K_3. \]

Now, we are ready to prove the Q-linear convergence of \( \{F_k\} \) and R-linear convergence of the sequence \( \{w^k\} \). In order to simplify the proof, we denote the last terms in formulas (47)-(49) as \( D_1, D_2 \) and \( D_3 \) respectively.

**Theorem 5.7.** Under the same assumption in Lemma 5.6, the sequence \( \{F_k\} \) is Q-linearly convergent.

**Proof.** Together with (61), we have
\[ L_\beta(w^{k+1}) - L_\beta(w^k) \leq D_1 + \langle d_1^{k+1}, x^{k+1} - x^k \rangle + \langle d_2^{k+1}, y^{k+1} - y^k \rangle, \quad \forall k \geq 0. \]
Considering the equalities \( (\alpha + \theta) Bx^{k+1} = \lambda^k - \lambda^{k+1} + \alpha \beta (y^k + c) + \theta (y^{k+1} + c) \) and \( A x^{k+1} - B y^{k+1} = 0 \), and using the Cauchy-Schwarz inequality, we get
\[ L_\beta(w^{k+1}) - L_\beta(w^k) \leq D_2 + \frac{1}{2} \|d_1^{k+1}\|^2 + \frac{1}{2} \|d_2^{k+1} - x^{k+1} - x^k\|^2 + \frac{1}{2} \|d_2^{k+1}\|^2 + \frac{1}{2} \|y^{k+1} - y^k\|^2, \quad \forall k \geq 0. \]
We rearrange the above inequality as follows, which corresponding to the inequality (50) in Theorem 4.3.
\[ L_\beta(w^{k+1}) - L_\beta(w^k) \]
\[ \leq D_3 + \frac{\sigma_1}{2} \|x^{k+1} - x^k\|^2 + \frac{\sigma_2}{2} \|y^{k+1} - y^k\|^2 + \frac{1}{2} \|x^{k+1} - x^k\|^2 + \frac{1}{2} \|y^{k+1} - y^k\|^2 \]
\[ \leq \frac{\sigma_1}{2} \|x^{k+1} - x^k\|^2 + \tau_1 \|\lambda^{k+1} - \lambda^k\|^2 + \tau_2 \|y^{k+1} - y^k\|^2 + \tau_3 \text{dist}^2(w^{k+1}, \text{crit} L_\beta), \quad \forall k \geq 0. \] (76)
where the second inequality is due to $\text{dist}^2(w^{k+1}, \text{crit}L_\beta) = \|w^{k+1} - w^{k+1}\|^2$ and the positive parameters are $\tau_1 = \frac{1}{2}((t_1 + t_2 + t_4) + t_6) \geq 0$, $\tau_2 = \frac{1}{2}((t_3 + t_4 + t_5) + t_1 + t_5 + t_7 + \frac{c_2}{4}) \geq 0$, $\tau_3 = t_1 + t_5 + \frac{1}{2}(t_2 + t_3 + L) + \frac{1}{4} \geq 0$. According to the result in Lemma 5.6 (ii), for $k \geq K_3$, we can deduce the following inequality from (76)

$$L_\beta(w^{k+1}) - \inf_{k \in N} F_k \leq (\tau_1 + \tau_3 \tau^2 \beta_1)\|\lambda^{k+1} - \lambda^k\|^2 + (\tau_2 + \tau_3 \tau^2 \beta_2)\|y^{k+1} - y^k\|^2 + 3\tau_3 \tau^2 \sigma_1 \|x^{k+1} - x^k\|^2, \forall k \geq K_3,$$

which using (71) and (42). According to the formula of $F_{k+1}$, we have

$$F_{k+1} - \inf_{k \in N} F_k \leq \tau_0 (\|\lambda^{k+1} - \lambda^k\|^2 + \|y^{k+1} - y^k\|^2 + \|x^{k+1} - x^k\|^2), \forall k \geq K_3,$$

where the positive parameter $\tau_0 = \max \left\{ \tau_1 + \tau_3 \tau^2 \beta_1 + \frac{c_1}{2 + \sigma}, \tau_2 + \tau_3 \tau^2 \beta_2 + c_2, 3\tau_3 \tau^2 \sigma_1 \right\}$ in the last inequality. Combing (77) with (66) and $F_{k+1} - \inf_{k \in N} F_k \geq 0$, we obtain further that

$$\begin{aligned}
&\left( F_{k+1} - \inf_{k \in N} F_k \right) - \left( F_k - \inf_{k \in N} F_k \right) \\
\leq &- \tilde{a}_1 \|\Delta y^{k+1}\|^2 - \left( \frac{c_1}{2} - \frac{c_1}{2 + \sigma} \right) \|\Delta \lambda^{k+1}\|^2 - \tilde{a}_2 \|\Delta x^{k+1}\|^2 \\
\leq &- \sigma_0 (\|\Delta y^{k+1}\|^2 + \|\Delta \lambda^{k+1}\|^2 + \|\Delta x^{k+1}\|^2) \\
\leq &- \frac{\sigma_0}{\tau_0} \left( F_{k+1} - \inf_{k \in N} F_k \right),
\end{aligned}$$

where the positive parameter $\sigma_0 = \min \left\{ a_1, \frac{\sigma}{2}, -\frac{c_1}{2 + \sigma}, \tilde{a}_2 \right\}$ in the second inequality, and the last inequality is due to (77). Reorganizing (78), we see that for all sufficiently large $k$,

$$0 \leq F_{k+1} - \inf_{k \in N} F_k \leq \frac{1}{1 + \frac{\sigma_0}{\tau_0}} \left( F_k - \inf_{k \in N} F_k \right),$$

which implies that the sequence $\{F_k\}$ is $Q$-linearly convergent. This completes the proof.

$\square$

**Theorem 5.8.** Under the same assumption in Lemma 5.6, the sequence $\{w^k\}$ generated by Algorithm 1 is $R$-linearly convergent to a critical point of $L_\beta(\cdot)$

**Proof.** The proof is similar to that in Theorem 4.4, we omit it here.

### 6. Conclusions

In this paper, we studied the convergence rate of the symmetric ADMM (5) for two block nonconvex separable optimization problems. When both component functions of the objective function are nonconvex, we established the $O(1/\sqrt{k})$ iteration complexity for the symmetric ADMM, under the assumption of the boundedness of the generated sequence. We also gave some sufficient conditions guaranteeing the boundedness of the iterative sequence generated by the symmetric ADMM. When the nonsmooth component function of the objective function is convex, we proved that the iterative sequence generated by the symmetric ADMM converges locally to a KKT point of (1) in a $R$-linear rate and the auxiliary sequence $\{F_k\}$ converges in a $Q$-linear rate, under the error bound condition. Finally, an inexact version with relative error criteria was proposed, and related convergence analysis was conducted.
Acknowledgments This work is supported by National Natural Science Foundation of China (Grant No. 11625105, 11431002, 11801279, 11871279, and 11571178), Natural Science Foundation of Jiangsu Province (Grant No. BK20180782), and the Startup Foundation for Introducing Talent of NUIST (Grant No. 2017r059).

REFERENCES

[1] H. Attouch, J. Bolte, B. F. Svaiter and A. Soubeyran, Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods, Mathematical Programming, 137 (2013), 91–129.
[2] L. E. Gibbons, D. W. Hearn, P. M. Pardalos and M. V. Ramana, Continuous characterizations of the maximum clique problem, Mathematics of Operations Research, 22 (1997), 754–768.
[3] R. Glowinski and A. Marroco, Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité d’une classe de problèmes de Dirichlet non linéaires, Revue Française D’Automatique, Informatique, Recherche Opérationnelle. Analyse Numérique, 9 (1975), 41–76.
[4] R. Glowinski, T. W. Pan and X. C. Tai, Some facts about operator-splitting and alternating direction methods, Splitting Methods in Communication, Imaging, Science, and Engineering, (2016), 19–94.
[5] Y. Gu, B. Jiang and D. R. Han, A semi-proximal-based strictly contractive Peaceman-Rachford splitting method, arXiv:1506.02221, 2015.
[6] B. S. He, H. Liu, Z. R. Wang and X. M. Yuan, A strictly contractive Peaceman-Rachford splitting method for convex programming, SIAM Journal on Optimization, 24 (2014), 1011–1040.
[7] B. S. He, F. Ma and X. M. Yuan, Convergence study on the symmetric version of admm with larger step sizes, SIAM Journal on Imaging Sciences, 9 (2016), 1467–1501.
[8] D. R. Han, A new hybrid generalized proximal point algorithm for variational inequality problems, Journal of Global Optimization, 26 (2003), 125–140.
[9] D. R. Han, Inexact operator splitting methods with selfadaptive strategy for variational inequality problems, Journal of Optimization Theory and Applications, 132 (2007), 227–243.
[10] Z. H. Jia, X. Gao, X. J. Cai and D. R. Han, Local Linear Convergence of the Alternating Direction Method of Multipliers for Nonconvex Separable Optimization Problems, Manuscript, 2018.
[11] W. Kaplan, A test for copositive matrices, Linear Algebra and its Applications, 313 (2000), 203–206.
[12] J. Liu, J. H. Chen and J. P. Ye, Large-scale sparse logistic regression, Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Paris, (2009), 547–556.
[13] P. L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM Journal on Numerical Analysis, 16 (1979), 964–979.
[14] Z. Q. Luo and P. Tseng, Error bound and convergence analysis of matrix splitting algorithms for the affine variational inequality problem, SIAM Journal on Optimization, 2 (1992), 43–54.
[15] Z. Q. Luo and P. Tseng, Error bounds and convergence analysis of feasible descent methods: A general approach, Annals of Operations Research, 46 (1993), 157–178.
[16] D. H. Peaceman and H. H. Rachford, The numerical solution of parabolic and elliptic differential equations, Journal of the Society for Industrial and Applied Mathematics, 3 (1955), 28–41.
[17] N. Parikh and S. Boyd, Proximal algorithms, Foundations and Trends in Optimization, 1 (2014), 127–239.
[18] H. Villálio, Quadratic programming criteria for copositive matrices, Linear Algebra and its Applications, 119 (1989), 163–182.
[19] Z. M. Wu, M. Li, David Z. W. Wang and D. R. Han, A symmetric alternating direction method of multipliers for separable nonconvex minimization problems, Asia-Pacific Journal of Operational Research, 34 (2017), 1750030, 27pp.
[20] B. Wen, X. J. Chen and T. K. Pong, Linear convergence of proximal gradient algorithm with extrapolation for a class of nonconvex nonsmooth minimization problems, SIAM Journal on Optimization, 27 (2017), 124–145.
[21] Z. B. Xu, X. Y. Chang, F. M. Xu and H. Zhang, $L_{1/2}$ regularization: A thresholding representation theory and a fast solver, *IEEE Transactions on Neural Networks and Learning Systems*, 23 (2012), 1013–1027.

Received July 2019; revised October 2019.

E-mail address: jiazehui90@126.com
E-mail address: 15380823957@163.com
E-mail address: caixingju@njnu.edu.cn
E-mail address: handr@buaa.edu.cn