Research Article

Proximal Alternating Direction Method with Relaxed Proximal Parameters for the Least Squares Covariance Adjustment Problem

Minghua Xu,1 Yong Zhang,1 Qinglong Huang,1 and Zhenhua Yang2

1 School of Mathematics and Physics, Changzhou University, Jiangsu 213164, China
2 College of Science, Nanjing University of Posts and Telecommunications, Jiangsu 210003, China

Correspondence should be addressed to Minghua Xu; xuminghua@cczu.edu.cn
Received 13 June 2013; Accepted 27 July 2013; Published 21 January 2014

Copyright © 2014 Minghua Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider the problem of seeking a symmetric positive semidefinite matrix in a closed convex set to approximate a given matrix. This problem may arise in several areas of numerical linear algebra or come from finance industry or statistics and thus has many applications. For solving this class of matrix optimization problems, many methods have been proposed in the literature. The proximal alternating direction method is one of those methods which can be easily applied to solve these matrix optimization problems. Generally, the proximal parameters of the proximal alternating direction method are greater than zero. In this paper, we conclude that the restriction on the proximal parameters can be relaxed for solving this kind of matrix optimization problems. Numerical experiments also show that the proximal alternating direction method with the relaxed proximal parameters is convergent and generally has a better performance than the classical proximal alternating direction method.

1. Introduction

This paper concerns the following problem:

\[ \min_X \left\{ \frac{1}{2} \| X - C \|_F^2 \mid X \in S^+_n \cap S_B \right\}, \quad (1) \]

where \( C \in R^{n \times n} \) is a given symmetric matrix,

\[ S^+_n = \left\{ X \in R^{n \times n} \mid X^T = X, X \succeq 0 \right\}, \]

\[ S_B = \left\{ X \in R^{n \times n} \mid \text{Tr} (A_i X) = b_i, i = 1, 2, \ldots, p, \right\} \]

\[ \text{Tr} (G_j X) \leq d_j, j = 1, 2, \ldots, m \right\}, \]

matrices \( A_i \in R^{n \times n} \) and \( G_j \in R^{n \times n} \) are symmetric and scalars, \( b_j \) and \( d_j \) are the problem data, \( X \succeq 0 \) denotes that \( X \) is a positive semidefinite matrix, \( \text{Tr} \) denotes the trace of a matrix, and \( \| \cdot \|_F \) denotes the Frobenius norm; that is,

\[ \| X \|_F = \left( \text{Tr} (X^T X) \right)^{1/2} = \left( \sum_{i,j=1}^n X_{ij}^2 \right)^{1/2}, \quad (3) \]

and \( S^+_n \cap S_B \) is nonempty. Throughout this paper, we assume that the Slater’s constraint qualification condition holds so that there is no duality gap if we use Lagrangian techniques to find the optimal solution to problem (1).

Problem (1) is a type of matrix nearness problem, that is, the problem of finding a matrix that satisfies some properties and is nearest to a given one. Problem (1) can be called the least squares covariance adjustment problem or the least squares semidefinite programming problem and solved by many methods [1–4]. In a least squares covariance adjustment problem, we make adjustments to a symmetric matrix so that it is consistent with prior knowledge or assumptions and a valid covariance matrix [2,5,6]. The matrix nearness problem has many applications especially in several areas of numerical linear algebra, finance industry, and statistics in [6]. A recent survey of matrix nearness problems can be found in [7]. It is clear that the matrix nearness problem considered here is a convex optimization problem. It thus follows from the strict feasibility and coercivity of the objective function that the minimum of (1) is attainable and unique.
In the literature of interior point algorithms, $S^+_n$ is called the semidefinite cone and the related problem (1) belongs to the class of semidefinite programming (SDP) and second-order cone programming (SOCP) [8]. In fact, it is possible to reformulate problem (1) into a mixed SDP and SOCP as in [3, 9]:

\[
\begin{align*}
\min & \quad t \\
\text{s.t.} & & \langle A_i, X \rangle = b_i, \quad i = 1, 2, \ldots, p, \\
& & \langle G_j, X \rangle \leq d_j, \quad j = 1, 2, \ldots, m, \\
& & t \geq \|X - C\|_F, \\
& & X \in S^n_+,
\end{align*}
\]

where $\langle X, Y \rangle = \text{Tr}(X^T Y)$.

Thus, problem (1) can be efficiently solved by standard interior-point methods such as SeDuMi [10] and SDPT3 [11] when the number of variables (i.e., entries in the matrix $X$) is modest, say under 1000 (corresponds to $n$ around 32) and the number of equality and inequality constraints is not too large (say, 5,000) [2, 3, 12].

Specially, let

\[
S_B = \{ X \in R^{n \times n} \mid \text{Diag}(X) = e \},
\]

where $\text{Diag}(X)$ is the vector of diagonal elements of $X$ and $e$ is the vector of 1s. Then problem (1) can be viewed as the nearest correlation matrix problem. For the nearest correlation matrix problem, a quadratically convergent Newton algorithm was presented recently by Qi and Sun [13], and Zhang, Han, Li, Yuan, and Bauschke and Borwein showed that the new iterate $\tilde{x}$ in (14) and (15) of the proximal ADM with relaxed proximal parameters are greater than zero. In this paper, we will show that the restriction on the proximal parameters can be relaxed while the proximal ADM is used to solve problem (1). Numerical experiments also show that the proximal ADM with the relaxed proximal parameters generally has a better performance than the classical proximal ADM.

The paper is organized as follows. In Section 2, we give some preliminaries about the proximal alternating direction method. In Section 3, we convert the problem (1) to a structured variational inequality and apply the proximal ADM to solve it. The basic analysis and convergent results of the proximal ADM with relaxed proximal parameters are built in Section 4. Preliminary numerical results are reported in Section 5. Finally, we give some conclusions in Section 6.

### 2. Proximal Alternating Direction Method

In order to introduce the proximal ADM, we first consider the following structured variational inequality problem which includes two separable subvariational inequality problems: find $(x, y) \in \Omega$ such that

\[
\begin{align*}
(x' - x)^T f(x) & \geq 0, \\
(y' - y)^T g(y) & \geq 0,
\end{align*}
\]

where

\[
\Omega = \{(x, y) \mid Ax + By = b, x \in \mathcal{X}, y \in \mathcal{Y}\},
\]

$f : R^n \rightarrow R^n$ and $g : R^m \rightarrow R^m$ are monotone; that is,

\[
\begin{align*}
(x - x)^T (f(x) - f(\tilde{x})) & \geq 0, \quad \forall \tilde{x}, x \in R^n, \\
(y - y)^T (g(y) - g(\tilde{y})) & \geq 0, \quad \forall \tilde{y}, y \in R^m,
\end{align*}
\]

where $A \in R^{k \times n}, B \in R^{k \times m},$ and $b \in R^l; \mathcal{X} \subset R^n$ and $\mathcal{Y} \subset R^m$ are closed convex sets. Studies of such variational inequality can be found in Glowinski [20], Glowinski and Le Tallec [21], Eckstein and Fukushima [22–24], He and Yang [25], He et al. [26], and Xu [27].

By attaching a Lagrange multiplier vector $\lambda \in R^l$ to the linear constraint $Ax + By = b$, problem (6)-(7) can be explained as the following form (see [20, 21, 24]): find $w = (x, y, \lambda) \in \mathcal{W}$ such that

\[
\begin{align*}
(x' - x)^T f(x) - A^T \lambda & \geq 0, \\
(y' - y)^T g(y) - B^T \lambda & \geq 0, \quad \forall w' = (x', y', \lambda') \in \mathcal{W},
\end{align*}
\]

where

\[
\mathcal{W} = \mathcal{X} \times \mathcal{Y} \times R^l.
\]

For solving (9)-(10), Gabay [28] and Gabay and Mercier [29] proposed the ADM method. In the classical ADM method, the new iterate $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1}) \in \mathcal{W}$ is generated from a given triple $w^k = (x^k, y^k, \lambda^k) \in \mathcal{W}$ via the following procedure.

First, $\lambda^{k+1}$ is found by solving the following problem:

\[
\begin{align*}
(x' - x)^T f(x) - A^T \left[ \lambda^k - \beta (Ax + By^k - b) \right] & \geq 0, \\
& \forall x' \in \mathcal{X},
\end{align*}
\]
where $x \in \mathcal{X}$. Then, $y^{k+1}$ is obtained by solving
\[
(y' - y)^T \left\{ g(y) - B^T \left[ \lambda^k - \beta \left( Ax^{k+1} + By^{k+1} - b \right) \right] \right\} \geq 0,
\forall y' \in \mathcal{Y},
\tag{12}
\]
where $y \in \mathcal{Y}$. Finally, the multiplier is updated by
\[
\lambda^{k+1} = \lambda - \beta \left( Ax^{k+1} + By^{k+1} - b \right),
\tag{13}
\]
where $\beta > 0$ is a given penalty parameter for the linearly constraint $Ax + By - b = 0$. Most of the existing ADM methods require that the subvarational inequality problems (II)-(12) should be solved exactly at each iteration. Note that the involved subvariational inequality problem (II)-(12) may not be well-conditioned without strongly monotone assumptions on $f$ and $g$. Hence, it is difficult to solve these subvariational inequality problems exactly in many cases. In order to improve the condition of solving the subproblem by the ADM, some proximal ADMs were proposed (see, e.g., [26, 27, 30–34]). The classical proximal ADM is one of the attractive ADMs. From a given triple $w^k = (x^k, y^k, \lambda^k) \in \mathcal{W}$, the classical proximal ADM produces the new iterate $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1}) \in \mathcal{W}$ by the following procedure.

First, $x^{k+1}$ is obtained by solving the following variational inequality problem:
\[
(x' - x)^T \left\{ f(x) - A^T \left[ \lambda^k - \beta \left( Ax + By^k - b \right) \right] \right\} + r \left( x - x^k \right) \geq 0, \quad \forall x' \in \mathcal{X},
\tag{14}
\]
where $r > 0$ is the given proximal parameter and $x \in \mathcal{X}$. Then, $y^{k+1}$ is found by solving
\[
(y' - y)^T \left\{ g(y) - B^T \left[ \lambda^k - \beta \left( Ax^{k+1} + By^{k+1} - b \right) \right] \right\} + s \left( y - y^k \right) \geq 0, \quad \forall y' \in \mathcal{Y},
\tag{15}
\]
where $s > 0$ is the given proximal parameter and $y \in \mathcal{Y}$. Finally, the multiplier is updated by
\[
\lambda^{k+1} = \lambda^k - \beta \left( Ax^{k+1} + By^{k+1} - b \right),
\tag{16}
\]

In this paper, we will conclude that problem (1) can be solved by the proximal ADM and the restriction on the proximal parameters $r > 0$, $s > 0$ can be relaxed as $r > -1/2$, $s > -1/2$ when the proximal ADM is applied to solve problem (1). Our numerical experiments later also show that the numerical performance of the proximal ADM with smaller value of proximal parameters is generally better than the proximal ADM with comparatively larger value of proximal parameters.

### 3. Converting Problem (1) to a Structured Variational Inequality

In order to solve the problem (1) with proximal ADM, we convert problem (1) to the following equivalent one:
\[
\begin{align*}
\min_{x,y} & \quad \frac{1}{2} \| X - C \|_F^2 + \frac{1}{2} \| Y - C \|_F^2 \\
\text{s.t.} & \quad X - Y = 0,
\end{align*}
\tag{17}
\]
Following the KKT condition of (17), the solution to (17) can be found by finding $w = (X, Y, A) \in \mathcal{W}$ such that
\[
\begin{align*}
\left\langle X' - X, (X - C) - A \right\rangle & \geq 0, \\
\left\langle Y' - Y, (Y - C) + A \right\rangle & \geq 0, \quad \forall w' = (X', Y', A') \in \mathcal{W}, \\
X - Y & = 0,
\end{align*}
\tag{18}
\]
where
\[
\mathcal{W} = S_a \times S_b \times R_{n \times n}.
\tag{19}
\]

It is easy to see that problem (18)-(19) is a special case of the structured variational inequality (9)-(10) and thus can be solved by proximal ADM. For given $w^k = (X^k, Y^k, \Lambda^k) \in \mathcal{W}$, it is fortunate that the $w^{k+1} = (X^{k+1}, Y^{k+1}, \Lambda^{k+1})$ can be exactly obtained by the proximal ADM in the following way:
\[
X^{k+1} = P_{S_a} \left\{ \frac{1}{1 + \beta + r} \left( C + r X^k + \beta Y^k + \Lambda^k \right) \right\},
\tag{20}
\]
\[
Y^{k+1} = P_{S_b} \left\{ \frac{1}{1 + \beta + s} \left( C + \beta X^{k+1} + s Y^k - \Lambda^k \right) \right\},
\tag{21}
\]
\[
\Lambda^{k+1} = \Lambda^k - \beta \left( X^{k+1} - Y^{k+1} \right),
\tag{22}
\]
where the projection of $v$ on a nonempty closed convex set $S$ of $R_{n \times n}$ under Frobenius norm, denoted by $P_S(v)$, is the unique solution to the following problem; that is,
\[
P_S(v) = \arg \min_u \left\{ \| u - v \|_F^2 \mid u \in S \right\}.
\tag{23}
\]

It follows that the solution to
\[
\min \left\{ \frac{1}{2} \| Z - X \|_F^2 \mid Z \in S_a \right\}
\tag{24}
\]
is called the projection of $X$ on $S_a$ and denoted by $P_{S_a}(X)$. Using the fact that matrix Frobenius norm is invariant under unitary transform, it is known (see [35]) that
\[
P_{S_a}(X) = Q \tilde{A} Q^T,
\tag{25}
\]
where
\[
Q^T X Q = \text{diag}(\lambda_1, \ldots, \lambda_n)
\tag{26}
\] is the symmetric Schur decomposition of $X$ ($Q = (q_1, \ldots, q_n)$ is an orthogonal matrix whose column vector $q_i$, $i = 1, \ldots, n$, is...
is the eigenvector of \( X \), and \( \lambda_i, \ i = 1, \ldots, n \), is the related
eigenvalue),
\[
\tilde{\lambda} = \text{diag} (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n), \ \ \tilde{\lambda}_i = \max (\lambda_i, 0).
\] (27)
In order to obtain the projection \( P_{S_p}(X) \), we need to solve the
following quadratic program:
\[
\min_Z \frac{1}{2} \| Z - X \|_F^2 \quad \text{s.t.} \quad \begin{align*}
\text{Tr} (A_i Z) &= b_i, \ i = 1, 2, \ldots, p, \\
\text{Tr} (G_j Z) &\leq d_j, \ j = 1, 2, \ldots, m.
\end{align*} (28)
\]

The dual problem of (28) can be written as
\[
\min_v \frac{1}{2} v^T H v + q^T v \quad \text{s.t.} \quad v \in \mathbb{R}^p \times \mathbb{R}^m, (29)
\]
where \( H \) is positive semidefinite and \( H \) and \( q \) have the
following form, respectively:

Further, letting
\[
F (w^{k+1}) = \begin{pmatrix}
X_{k+1} - C - \Lambda_{k+1} \\
y_{k+1} - X_{k+1} - \Lambda_{k+1}
\end{pmatrix},
\]
\[
d_1 (w^k, w^{k+1}) = \begin{pmatrix}
\frac{1}{\beta} I_n & 0 & 0 \\
0 & (s + \beta) I_n & 0 \\
0 & 0 & \frac{1}{\beta} I_n
\end{pmatrix} \begin{pmatrix}
X_{k+1} - X_k \\
y_{k+1} - y_k \\
\Lambda_{k+1} - \Lambda_k
\end{pmatrix}, (34)
\]
where \( I_n \in \mathbb{R}^{n \times n} \) is the unit matrix, and
\[
d_2 (w^k, w^{k+1}) = F (w^{k+1}) - \beta \begin{pmatrix}
I_n \\
0
\end{pmatrix} (y_k - y_{k+1}), (35)
\]
then we can get the following lemmas.

**Lemma 1.** Let \( \{w^k\} \) be the sequence generated by applying the
proximal ADM to problem (18)-(19) and let \( w^* \in \mathscr{W}^* \) be any
solution to problem (18)-(19), then one has
\[
\left\langle w^{k+1} - w^*, d_2 (w^k, w^{k+1}) \right\rangle \]
\[
\geq - \left\langle \Lambda^k - \Lambda^{k+1}, Y_k - Y_{k+1} \right\rangle + \| X_{k+1} - X_k \|_F^2 \\ + \| y_{k+1} - y^* \|_F^2, (36)
\]

4. Main Results

Let \( \{w^k\} \) be the sequence generated by applying the
procedure (14)-(16) to problem (18)-(19); then for any \( w^j = (X^j, Y^j, \Lambda^j) \in \mathscr{W} \), we have that
\[
\begin{align*}
\left\langle X^j - X^{k+1}, X^{k+1} - C - \Lambda^{k+1} - \beta (Y^k - Y^{k+1}) \right\rangle \\
+ r \left( X^{k+1} - X^j \right) &\geq 0, \\
\left\langle Y^j - Y^{k+1}, Y^{k+1} - C + \Lambda^{k+1} + s (Y^{k+1} - Y^j) \right\rangle &\geq 0, \\
\Lambda^{k+1} &= \Lambda^k - \beta (X^k - Y^k).
\end{align*} (33)
\]
Proof. From (22) and (35), we have
\[
\langle w_{k+1} - w^*, d_2(w_k, w_{k+1}) \rangle = -\langle \Lambda^k - \Lambda^{k+1}, Y^k - Y^{k+1} \rangle
+ \langle w_{k+1} - w^*, F(w_{k+1}) \rangle.
\] (37)

Since (9) and \( w^* \) are a solution to problem (18)-(19) and \( X^{k+1} \in S^*_1, Y^{k+1} \in S_B \), we have
\[
\langle w_{k+1} - w^*, F(w^*) \rangle \geq 0.
\] (38)

From (38), it follows that
\[
\begin{align*}
\langle w_{k+1} - w^*, F(w_{k+1}) \rangle + F(w^*) & \geq 0. \\
\text{Thus, we have} & \\
\langle w_{k+1} - w^*, F(w_{k+1}) \rangle & \geq \langle w_{k+1} - w^*, F(w_{k+1}) \rangle - F(w^*) \\
& = \langle X^{k+1} - X^*, X^{k+1} - X^* - (\Lambda^{k+1} - \Lambda^*) \rangle \\
& + \langle Y^{k+1} - Y^*, Y^{k+1} - Y^* + (\Lambda^{k+1} - \Lambda^*) \rangle \\
& + \langle \Lambda^{k+1} - \Lambda^*, X^{k+1} - X^* - (Y^{k+1} - Y^*) \rangle \\
& = \langle X^{k+1} - X^*, X^{k+1} - X^* \rangle + \langle Y^{k+1} - Y^*, Y^{k+1} - Y^* \rangle \\
& = \|X^{k+1} - X^*\|_F^2 + \|Y^{k+1} - Y^*\|_F^2. \\
\text{Substituting (40) into (37), we get the assertion of this lemma.} \\
\end{align*}
\]

Lemma 2. Let \( \{w^k\} \) be the sequence generated by applying the proximal ADM to problem (18)-(19) and let \( w^* \in \mathcal{W}^* \) be any solution to problem (18)-(19); then one has
\[
\begin{align*}
\langle w^k - w^*, G_0(w^k - w^{k+1}) \rangle & \geq \langle w^k - w^{k+1}, G_0(w^k - w^{k+1}) \rangle - \langle \Lambda^k - \Lambda^{k+1}, Y^k - Y^{k+1} \rangle \\
& + \|X^{k+1} - X^*\|_F^2 + \|Y^{k+1} - Y^*\|_F^2. \\
\text{Thus, we have} & \\
\langle w^k - w^*, d_1(w^k, w^{k+1}) \rangle & \geq \langle w^k - w^*, d_1(w^k, w^{k+1}) \rangle \\
& \geq \langle w^k - w^*, d_2(w^k, w^{k+1}) \rangle \\
& \geq -\langle \Lambda^k - \Lambda^{k+1}, Y^k - Y^{k+1} \rangle + \|X^{k+1} - X^*\|_F^2 \\
& + \|Y^{k+1} - Y^*\|_F^2. \\
\text{From the above inequality, we get} & \\
\langle w^k - w^*, G_0(w^k - w^{k+1}) \rangle & \geq \langle w^k - w^{k+1}, G_0(w^k - w^{k+1}) \rangle \\
& - \langle \Lambda^k - \Lambda^{k+1}, Y^k - Y^{k+1} \rangle + \|X^{k+1} - X^*\|_F^2 \\
& + \|Y^{k+1} - Y^*\|_F^2. \\
Hence, (41) holds and the proof is completed. \qed
\]

Theorem 3. Let \( \{w^k\} \) be the sequence generated by applying the proximal ADM to problem (18)-(19) and let \( w^* \in \mathcal{W}^* \) be any solution to problem (18)-(19); then one has
\[
\|w^{k+1} - w^*\|_F^2 \leq \|w^k - w^*\|_G^2 - \langle w^k - w^{k+1}, M(w^k - w^{k+1}) \rangle,
\] (46)

where
\[
G = \begin{pmatrix}
(r + 1) I_n & 0 & 0 \\
0 & (1 + s + \beta) I_n & 0 \\
0 & 0 & 1 - I_n
\end{pmatrix},
\]
\[
M = \begin{pmatrix}
\frac{1}{n} + r I_n & 0 & 0 \\
0 & \frac{1}{n} + s + \beta I_n & 0 \\
0 & 0 & - I_n
\end{pmatrix},
\]
and \( \|w\|_G^2 = \langle w, Gw \rangle \).

Proof. From (41), we have
\[
\|w^{k+1} - w^*\|_G^2 \\
\leq \|w^k - w^* - (w^k - w^{k+1})\|_G^2 \\
\leq \|w^k - w^*\|_G^2 - 2\|w^k - w^{k+1}\|_G^2.
\]
\[ + 2 \langle \Lambda^k - \Lambda^{k+1}, Y^k - Y^{k+1} \rangle - 2 \| X^{k+1} - X^* \|_F^2 \\
- 2 \| Y^{k+1} - Y^* \|_F^2 + \| w^k - w^{k+1} \|_{G_0}^2 \]
\[ = \| w^k - w^* \|_{G_0}^2 - \| w^k - w^{k+1} \|_{G_0}^2 \]
Rearranging the inequality above, we find that
\[ + 2 \langle \Lambda^k - \Lambda^{k+1}, Y^k - Y^{k+1} \rangle - 2 \| X^{k+1} - X^* \|_F^2 \\
- 2 \| Y^{k+1} - Y^* \|_F^2 \]

Using the Cauchy-Schwarz Inequality on the last term of the right-hand side of (49), we obtain
\[ \| X^{k+1} - X^* \|_F^2 + \| X^k - X^* \|_F^2 \geq \frac{1}{2} \| X^{k+1} - X^k \|_F^2, \]
\[ \| Y^{k+1} - Y^* \|_F^2 + \| Y^k - Y^* \|_F^2 \geq \frac{1}{2} \| Y^{k+1} - Y^k \|_F^2. \]
Substituting (50) into (49), we get
\[ \| w^{k+1} - w^* \|_G^2 \leq \| w^k - w^* \|_G^2 \]
\[ + \langle w^k - w^{k+1}, M (w^k - w^{k+1}) \rangle. \]
Thus, the proof is completed.

Based on the Theorem 3, we get the following lemma.

**Lemma 4.** Let \( \{w^k\} \) be the sequence generated by applying proximal ADM to problem (18)-(19), \( w^* \in \mathbb{R}^n \) any solution to problem (18)-(19), \( r > -1/2 \), and \( s > -1/2 \); then one has the following.

1. The sequence \( \| w^k - w^* \|_G^2 \) is nonincreasing;
2. The sequence \( \{w^k\} \) is bounded;
3. \( \lim_{k \to \infty} \| w^{k+1} - w^k \|_G^2 = 0 \);
4. \( G \) and \( M \) are both symmetric positive-definite matrices.

**Proof.** Since
\[ \left( \frac{1}{2} + s + \beta \right) I_n - I_n - \frac{1}{\beta} I_n \]
\[ \left( \frac{1}{2} + s + \beta \right) I_n - I_n - \frac{1}{\beta} I_n \]
it is easy to check that if \( r > -1/2 \), \( s > -1/2 \), then \( G \) and \( M \) are symmetric positive-definite matrices.

Let \( \tau > 0 \) be the smallest eigenvalue of matrix \( M \). Then, from (46), we have
\[ \| w^{k+1} - w^* \|_G^2 \leq \| w^k - w^* \|_G^2 - \tau \| w^k - w^{k+1} \|_F^2. \]
Following (53), we immediately have that \( \| w^k - w^* \|_G^2 \) is nonincreasing and thus the sequence \( \{w^k\} \) is bounded. Moreover, we have
\[ \| w^{k+1} - w^* \|_G^2 \leq \| w^0 - w^* \|_G^2 - \tau \sum_{j=0}^{k} \| w^j - w^{j+1} \|_F^2. \]
So, we get
\[ \sum_{j=0}^{k} \| w^j - w^{j+1} \|_F^2 < \infty, \quad \forall k > 0, \]
then
\[ \lim_{k \to \infty} \| w^k - w^{k+1} \|_F^2 = 0. \]
Thus, the proof is completed.

Following Lemma 4, now we are in the stage of giving the main convergence results of proximal ADM with \( r > -1/2 \) and \( s > -1/2 \) for problem (18)-(19).

**Theorem 5.** Let \( \{w^k\} \) be the sequence generated by applying proximal ADM to problem (18)-(19), \( r > -1/2 \), and \( s > -1/2 \); then \( \{w^k\} \) converges to a solution point of (18)-(19).

**Proof.** Since the sequence \( \{w^k\} \) is bounded (see point (2) of Lemma 4), it has at least one cluster point. Let \( w^\infty \) be a cluster point of \( \{w^k\} \) and the subsequence \( \{w^{k^*}\} \) converges to \( w^\infty \). It follows from (33) that
\[ \lim_{j \to \infty} \langle X' - X^{k^*+1}, X^{k^*+1} - C - \Lambda^{k^*+1} - \beta \left( Y^{k^*} - Y^{k^*+1} \right) + r \left( X^{k^*+1} - X^* \right) \rangle \geq 0, \]
Following point (3) of Lemma 4, we have
\[
\begin{align*}
\langle X' - X_\infty, X_\infty - C - \Lambda \rangle + \langle Y' - Y_\infty, Y_\infty - C - \Lambda \rangle & \geq 0, \\
\langle X' - X_\infty, X_\infty - Y_\infty \rangle + \langle Y' - Y_\infty, Y_\infty - X_\infty \rangle & \geq 0, \forall w' \in \mathcal{W},
\end{align*}
\]
(57)
This means that \( w_\infty \) is a solution point of (18)-(19). Since \( \{w^k\} \) converges to \( w_\infty \), we have that, for any given \( \varepsilon > 0 \), there exists an integer \( N > 0 \) such that
\[
\|w^k - w_\infty\|_G < \varepsilon, \quad \forall k \geq N.
\]
(59)
Furthermore, using the inequality (53), we have
\[
\|w^k - w_\infty\|_G^2 < \|w^k - w^j\|_G^2, \quad \forall k \geq k_j.
\]
(60)
Combining (59) and (60), we get that
\[
\|w^k - w_\infty\|_G^2 < \varepsilon, \quad \forall k > N.
\]
(61)
This implies that the sequence \( \{w^k\} \) converges to \( w_\infty \). So the proof is completed.

5. Numerical Experiments

In this section, we implement the proximal ADM to solve the problem (1) and show the numerical performances of proximal ADM with different proximal parameters. Additionally, we compare the classical ADM (i.e., the proximal ADM with proximal parameters \( r = 0 \) and \( s = 0 \)) with the alternating projections method proposed by Higham [6] numerically and show that the alternating projections method is not equivalent to proximal ADM with zero proximal parameters. All the codes were written in Matlab 7.1 and run on IBM notebook PC R400.

Example 6. In the first numerical experiment, we set the \( C_1 \) as an \( n \times n \) matrix whose entries are generated randomly in \([-1, 1]\). Let \( C = (C_1 + C_1^T)/2 \) and further let the diagonal elements of \( C \) be \( 1 \) that is, \( C_{ii} = 1, i = 1, 2, \ldots, n \). In this test example, we simply let \( S_p \) be in the form of (31) and
\[
H_L = (l_{ij}) \in R^{n,n},
\]
\[
l_{ij} = \begin{cases} 
-0.5, & i \neq j, \\
1, & i = j, 
\end{cases} \quad i, j = 1, 2, \ldots, n,
\]
\[
H_U = (u_{ij}) \in R^{n,n},
\]
(62)
Moreover, let \( X^0 = \text{eye}(n), Y^0 = \text{eye}(n), \Lambda^0 = \text{zeros}(n), \beta = 4, \) and \( \varepsilon = 10^{-6} \), where \( \text{eye}(n) \) and \( \text{zeros}(n) \) are both the Matlab functions. For different problem size \( n \) and different proximal parameters \( r \) and \( s \), Table 1 shows the computational results. There, we report the number of iterations (It.) and the computing time in seconds (CPU.) it takes to reach convergence. The stopping criterion of the proximal ADM is
\[
\|w^{k+1} - w^k\|_G < \varepsilon,
\]
where \( \|X\|_\infty = \max(\max(\text{abs}(X))) \) is the maximum absolute value of the elements of the matrix \( X \).

Remark 7. Note that if the proximal parameters are equal to zero, that is, \( r = 0 \) and \( s = 0 \), then the proximal ADM is the classical ADM.

Example 8. All the data are the same as in Example 6 except that \( C_1 \) is an \( n \times n \) matrix whose entries are generated randomly in \([-1000, 1000]\),
\[
H_L = (l_{ij}) \in R^{n,n},
\]
\[
l_{ij} = \begin{cases} 
-500, & i \neq j, \\
1000, & i = j, 
\end{cases} \quad i, j = 1, 2, \ldots, n,
\]
\[
H_U = (u_{ij}) \in R^{n,n},
\]
(64)
The computational results are reported in Table 2.

Example 9. Let \( S_p \) be in the form of (31) and \( l_{ij} = 0, u_{ij} = +\infty, \) \( i, j = 1, 2, \ldots, n \). Assume that \( C, X_0, Y_0, \Lambda_0, \beta, \varepsilon \), and the stopping criterion are the same as those in Example 6, but the diagonal elements of matrix \( C \) are replaced by
\[
C_{ii} = \alpha + (1 - \alpha) \times \text{rand}, \quad i = 1, 2, \ldots, n,
\]
(65)
where \( \alpha \in (0, 1) \) is a given number, rand is the Matlab function generating a number randomly in \([0, 1]\). In the following numerical experiments, we let \( \alpha = 0.2 \). For different problem size \( n \) and different proximal parameters \( r \) and \( s \), Table 3 shows the number of iterations and the computing time in seconds it takes to reach convergence.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( r = -0.3, s = -0.3 )</th>
<th>( r = 0, s = 0 )</th>
<th>( r = 3, s = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>It.</td>
<td>CPU.</td>
<td>It.</td>
</tr>
<tr>
<td>100</td>
<td>31</td>
<td>0.292</td>
<td>34</td>
</tr>
<tr>
<td>200</td>
<td>33</td>
<td>1.346</td>
<td>39</td>
</tr>
<tr>
<td>300</td>
<td>38</td>
<td>4.265</td>
<td>41</td>
</tr>
<tr>
<td>400</td>
<td>40</td>
<td>9.872</td>
<td>43</td>
</tr>
<tr>
<td>500</td>
<td>39</td>
<td>15.83</td>
<td>45</td>
</tr>
</tbody>
</table>
Abstract and Applied Analysis

Table 2: Numerical results of Example 8.

<table>
<thead>
<tr>
<th>n</th>
<th>r = −0.3, s = −0.3</th>
<th>r = 0, s = 0</th>
<th>r = 3, s = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lt. CPU.</td>
<td>lt. CPU.</td>
<td>lt. CPU.</td>
</tr>
<tr>
<td>100</td>
<td>49 0.476</td>
<td>54 0.551</td>
<td>116 1.837</td>
</tr>
<tr>
<td>200</td>
<td>51 2.197</td>
<td>57 2.334</td>
<td>128 5.430</td>
</tr>
<tr>
<td>300</td>
<td>59 6.614</td>
<td>61 8.108</td>
<td>136 15.25</td>
</tr>
<tr>
<td>400</td>
<td>56 12.74</td>
<td>63 14.51</td>
<td>140 31.65</td>
</tr>
<tr>
<td>500</td>
<td>58 23.90</td>
<td>66 26.90</td>
<td>147 59.98</td>
</tr>
</tbody>
</table>

Table 3: Numerical results of Example 9.

<table>
<thead>
<tr>
<th>n</th>
<th>r = −0.3, s = −0.3</th>
<th>r = 0, s = 0</th>
<th>r = 3, s = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lt. CPU.</td>
<td>lt. CPU.</td>
<td>lt. CPU.</td>
</tr>
<tr>
<td>100</td>
<td>32 0.282</td>
<td>35 0.288</td>
<td>70 0.566</td>
</tr>
<tr>
<td>200</td>
<td>33 1.295</td>
<td>36 1.397</td>
<td>72 4.006</td>
</tr>
<tr>
<td>300</td>
<td>34 3.745</td>
<td>37 4.156</td>
<td>73 8.285</td>
</tr>
<tr>
<td>400</td>
<td>34 7.885</td>
<td>37 8.571</td>
<td>73 16.73</td>
</tr>
<tr>
<td>500</td>
<td>34 14.07</td>
<td>37 15.42</td>
<td>74 29.87</td>
</tr>
</tbody>
</table>

Example 11. Let $C_1$ be an $n \times n$ matrix whose entries are generated randomly in $[−0.5, 0.5], C = (C_1 + C_1^T)/2,$ and let the diagonal elements of $C$ be 1. And let

$$S_R = \{ X \in R^{n \times n} \mid X = X^T, X_{ij} = e_{ij}, (i, j) \in B_c, X_{ij} \geq l_{ij}, (i, j) \in B_l, X_{ij} \leq u_{ij}, (i, j) \in B_u \},$$

where $B_c, B_l, B_u$ are subsets of $\{(i, j) \mid 1 \leq i, j \leq n\}$ denoting the indexes of such entries of $X$ that are constrained by equality, lower bounds, and upper bounds, respectively. In this test example, we let the index sets $B_c, B_l, B_u$ be the same as in Example 5.4 of [3]; that is, $B_c = \{(i, i) \mid i = 1, 2, \ldots, n\}$ and $B_l, B_u \subset \{(i, j) \mid 1 \leq i < j \leq n\}$ consist of the indices of min$(\tilde{n}_r, n - i)$ randomly generated elements at the $r$th row of $X, i = 1, 2, \ldots, n$ with $\tilde{n}_r = 5$ and $\tilde{n}_r = 10,$ respectively. We take $e_{ij} = 1$ for $(i, i) \in B_c, l_{ij} = -0.1$ for $(i, j) \in B_l,$ and $u_{ij} = 0.1$ for $(i, j) \in B_u.$

Moreover, let $X_0, Y_0, \Lambda_0, \beta, \epsilon,$ and the stopping criterion be the same as those in Example 6. For different problem size $n$, different proximal parameters $r$ and $s$, and different values of $\tilde{n}_r$, Tables 5(a) and 5(b) show the number of iterations and the computing time in seconds it takes to reach convergence, respectively.

Numerical experiments show that the proximal ADM with relaxed parameters is convergent. Moreover, we draw the conclusion that the proximal ADM with smaller value of proximal parameters generally converges more quickly than the proximal ADM with comparatively larger value of proximal parameters to solve the problem (I).

Example 12. In this test example, we apply the proximal ADM with $r = 0, s = 0,$ (i.e., the classical ADM) to solve the nearest correlation matrix problem, that is, problem (I) with $S_R$ in the form of (5), and compare the classical ADM numerically with the alternating projections method (APM) [6]. The APM computes the nearest correlation matrix to a symmetric $C \in R^{n \times n}$ by the following process:

$$\Delta S_0 = 0, Y_0 = C;$$

for $k = 1, 2, \ldots$

$$R_k = X_{k-1} - \Delta S_{k-1};$$

$$X_k = P_{S_R}(R_k);$$

$$\Delta S_k = X_k - R_k;$$

$$Y_k = P_{S_R}(X_k);$$

end.

In this numerical experiment, the stopping criterion of the APM is

$$\max \{ \| X_k - X_{k-1} \|_{\text{max}}, \| Y_k - Y_{k-1} \|_{\text{max}} \| X_k - Y_k \|_{\text{max}} \} < \epsilon. \quad (67)$$

Let the matrix $C$ and the initial parameters of classical ADM be the same as those in Example 6. Table 6(a) reports the numerical performance of proximal ADM and the APM for computing the nearest correlation matrix to $C.$
Further, let $C_1$ be an $n \times n$ matrix whose entries are generated randomly in $[0, 1]$ and $C = (C_1 + C_1^T)/2$. The other data are the same as above. Table 6(b) reports the numerical performance of the classical ADM and the APM for computing the nearest correlation matrix to the matrix $C$. Numerical experiments show that the classical ADM generally exhibits a better numerical performance than the APM for the test problems above.

6. Conclusions

In this paper, we apply the proximal ADM to a class of matrix optimization problems and find that the restriction of proximal parameters can be relaxed. Moreover, numerical experiments show that the proximal ADM with relaxed parameters generally has a better numerical performance in solving the matrix optimization problem than the classical proximal alternating direction method.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank the referees very sincerely for their valuable suggestions and careful reading of their paper. This research is financially supported by a research Grant from the Research Grant Council of China (Project no. 10971095).

References


Submit your manuscripts at http://www.hindawi.com