Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 598563, 10 pages
http://dx.doi.org/10.1155/2014/598563

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 Yang[2]

[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

Academic Editor: Abdellah Bnouhachem

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:

$$\min_X \left\{ \frac{1}{2} \|X - C\|_F^2 \mid X \in S_+^n \cap S_B \right\}, \tag{1}$$

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

$$S_+^n = \left\{ X \in R^{n\times n} \mid X^T = X, X \succeq 0 \right\},$$

$$S_B = \{ X \in R^{n\times n} \mid \mathrm{Tr}\left(A_i X\right) = b_i, i = 1, 2, \ldots, p, \quad \mathrm{Tr}\left(G_j X\right) \le d_j, j = 1, 2, \ldots, m \}, \tag{2}$$

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

$$\|X\|_F = \left(\mathrm{Tr}\left(X^T X\right)\right)^{1/2} = \left(\sum_{i,j=1}^{n} X_{ij}^2 \right)^{1/2}, \tag{3}$$

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]:

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & \langle A_i, X\rangle = b_i, \quad i = 1,2,\ldots,p, \\ & \langle G_j, X\rangle \le d_j, \quad j = 1,2,\ldots,m, \\ & t \ge \|X - C\|_F, \\ & X \in S_+^n, \end{aligned} \tag{4}$$

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

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

$$S_B = \{X \in R^{n\times n} \mid \operatorname{Diag}(X) = e\}, \tag{5}$$

where $\operatorname{Diag}(X)$ 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 improved by Borsdorf and Higham [1]. For problem (1) with equality and inequality constraints, one difficulty in finding an efficient method for solving this problem is the presence of the inequality constraints. In [3], Gao and Sun overcome this difficulty by reformulating the problem as a system of semismooth equations with two level metric projection operators and then design an inexact smoothing Newton method to solve the resulting semismooth system. For the problem (1) with large number of equality and inequality constraints, the numerical experiments in [14] show that the alternating direction method (hereafter alternating direction method is abbreviated as ADM) is more efficient in computing time than the inexact smoothing Newton method which additionally requires solving a large system of linear equations at each iteration. The ADM has many applications in solving optimization problems [15, 16]. Papers written by Zhang, Han, Li, Yuan, and Bauschke and Borwein show that the ADM can be applied to solve convex feasibility problems [17–19].

The proximal ADM is a class of ADM type methods which can also be easily applied to solve the matrix optimization problems. Generally, the proximal parameters (i.e., the parameters r and s in (14) and (15)) of the proximal ADM 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

$$\begin{aligned} (x' - x)^T f(x) \ge 0, \\ (y' - y)^T g(y) \ge 0, \end{aligned} \quad \forall (x', y') \in \Omega, \tag{6}$$

where

$$\Omega = \{(x, y) \mid Ax + By = b, x \in \mathcal{X}, y \in \mathcal{Y}\}, \tag{7}$$

$f : R^{n_1} \to R^{n_1}$ and $g : R^{n_2} \to R^{n_2}$ are monotone; that is,

$$\begin{aligned} (\tilde{x} - x)^T (f(\tilde{x}) - f(x)) \ge 0, \quad \forall \tilde{x}, x \in R^{n_1}, \\ (\tilde{y} - y)^T (g(\tilde{y}) - g(y)) \ge 0, \quad \forall \tilde{y}, y \in R^{n_2}, \end{aligned} \tag{8}$$

$A \in R^{l\times n_1}$, $B \in R^{l\times n_2}$, and $b \in R^l$; $\mathcal{X} \subset R^{n_1}$ and $\mathcal{Y} \subset R^{n_2}$ 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

$$\begin{aligned} & (x' - x)^T [f(x) - A^T\lambda] \ge 0 \\ & (y' - y)^T [g(y) - B^T\lambda] \ge 0, \quad \forall w' = (x', y', \lambda') \in \mathcal{W}, \\ & \quad Ax + By - b = 0, \end{aligned} \tag{9}$$

where

$$\mathcal{W} = \mathcal{X} \times \mathcal{Y} \times R^l. \tag{10}$$

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, x^{k+1} is found by solving the following problem:

$$(x' - x)^T \{f(x) - A^T [\lambda^k - \beta(Ax + By^k - b)]\} \ge 0, \quad \forall x' \in \mathcal{X}, \tag{11}$$

where $x \in \mathcal{X}$. Then, y^{k+1} is obtained by solving

$$\left(y' - y\right)^T \left\{g(y) - B^T\left[\lambda^k - \beta\left(Ax^{k+1} + By - b\right)\right]\right\} \geq 0, \quad \forall y' \in \mathcal{Y}, \tag{12}$$

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 subvariational inequality problems (11)-(12) should be solved exactly at each iteration. Note that the involved subvariational inequality problem (11)-(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:

$$\begin{aligned}&\left(x' - x\right)^T \left\{f(x) - A^T\left[\lambda^k - \beta\left(Ax + By^k - b\right)\right]\right. \\ &\qquad \left. + r\left(x - x^k\right)\right\} \geq 0, \quad \forall x' \in \mathcal{X},\end{aligned} \tag{14}$$

where $r > 0$ is the given proximal parameter and $x \in \mathcal{X}$. Then, y^{k+1} is found by solving

$$\begin{aligned}&\left(y' - y\right)^T \left\{g(y) - B^T\left[\lambda^k - \beta\left(Ax^{k+1} + By - b\right)\right]\right. \\ &\qquad \left. + s\left(y - y^k\right)\right\} \geq 0, \quad \forall y' \in \mathcal{Y},\end{aligned} \tag{15}$$

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:

$$\begin{aligned}\min_{X,Y} \quad & \frac{1}{2}\|X - C\|_F^2 + \frac{1}{2}\|Y - C\|_F^2 \\ \text{s.t.} \quad & X - Y = 0, \\ & X \in S_+^n, \quad Y \in S_B.\end{aligned} \tag{17}$$

Following the KKT condition of (17), the solution to (17) can be found by finding $w = (X, Y, \Lambda) \in \mathcal{W}$ such that

$$\begin{aligned}&\left\langle X' - X, (X - C) - \Lambda\right\rangle \geq 0, \\ &\left\langle Y' - Y, (Y - C) + \Lambda\right\rangle \geq 0, \qquad \forall w' = \left(X', Y', \Lambda'\right) \in \mathcal{W}, \\ &\qquad X - Y = 0,\end{aligned} \tag{18}$$

where

$$\mathcal{W} = S_+^n \times S_B \times R^{n\times n}. \tag{19}$$

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:

$$X^{k+1} = P_{S_+^n}\left\{\frac{1}{1+\beta+r}\left(C + rX^k + \beta Y^k + \Lambda^k\right)\right\}, \tag{20}$$

$$Y^{k+1} = P_{S_B}\left\{\frac{1}{1+\beta+s}\left(C + \beta X^{k+1} + sY^k - \Lambda^k\right)\right\}, \tag{21}$$

$$\Lambda^{k+1} = \Lambda^k - \beta\left(X^{k+1} - Y^{k+1}\right), \tag{22}$$

where the projection of v on a nonempty closed convex set S of $R^{m\times n}$ 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

$$\min\left\{\frac{1}{2}\|Z - X\|_F^2 \mid Z \in S_+^n\right\} \tag{24}$$

is called the projection of X on S_+^n and denoted by $P_{S_+^n}(X)$. Using the fact that matrix Frobenius norm is invariant under unitary transform, it is known (see [35]) that

$$P_{S_+^n}(X) = Q\widetilde{\Lambda}Q^T, \tag{25}$$

where

$$Q^T XQ = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \tag{26}$$

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 the eigenvector of X, and λ_i, $i = 1, \ldots, n$, is the related eigenvalue),

$$\widetilde{\Lambda} = \operatorname{diag}\left(\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n\right), \quad \widetilde{\lambda}_i = \max\left(\lambda_i, 0\right). \tag{27}$$

In order to obtain the projection $P_{S_B}(X)$, we need to solve the following quadratic program:

$$\begin{aligned} \min_{Z} \quad & \frac{1}{2}\|Z - X\|_F^2 \\ \text{s.t.} \quad & \operatorname{Tr}\left(A_i Z\right) = b_i, \quad i = 1, 2, \ldots, p, \\ & \operatorname{Tr}\left(G_j Z\right) \le d_j, \quad j = 1, 2, \ldots, m. \end{aligned} \tag{28}$$

The dual problem of (28) can be written as

$$\begin{aligned} \min_{v} \quad & \frac{1}{2} v^T H v + q^T v \\ \text{s.t.} \quad & v \in R^p \times R_+^m, \end{aligned} \tag{29}$$

where H is positive semidefinite and H and q have the following form, respectively:

$$H = \begin{pmatrix} \operatorname{Tr}\left(A_1 A_1^T\right) & \cdots & \operatorname{Tr}\left(A_1 A_p^T\right) & \operatorname{Tr}\left(A_1 G_1^T\right) & \cdots & \operatorname{Tr}\left(A_1 G_m^T\right) \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \operatorname{Tr}\left(A_p A_1^T\right) & \cdots & \operatorname{Tr}\left(A_p A_p^T\right) & \operatorname{Tr}\left(A_p G_1^T\right) & \cdots & \operatorname{Tr}\left(A_p G_m^T\right) \\ \operatorname{Tr}\left(G_1 A_1^T\right) & \cdots & \operatorname{Tr}\left(G_1 A_p^T\right) & \operatorname{Tr}\left(G_1 G_1^T\right) & \cdots & \operatorname{Tr}\left(G_1 G_m^T\right) \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \operatorname{Tr}\left(G_m A_1^T\right) & \cdots & \operatorname{Tr}\left(G_m A_p^T\right) & \operatorname{Tr}\left(G_m G_1^T\right) & \cdots & \operatorname{Tr}\left(G_m G_m^T\right) \end{pmatrix}, \quad q = \begin{pmatrix} b_1 - \operatorname{Tr}\left(A_1 X\right) \\ \vdots \\ b_p - \operatorname{Tr}\left(A_p X\right) \\ d_1 - \operatorname{Tr}\left(G_1 X\right) \\ \vdots \\ d_m - \operatorname{Tr}\left(G_m X\right) \end{pmatrix}. \tag{30}$$

Problem (29) is often a medium-scale quadratic programming (QP) problem. A variety of methods for solving the QP are commonly used, including interior-point methods and active set algorithm (see [36, 37]).

Particularly, if S_B is the following special case:

$$S_B = \left\{X \in R^{n\times n} \mid X^T = X, H_L \le X \le H_U\right\}, \tag{31}$$

where $H \ge 0$ expresses that each element of H is nonnegative, H_L and H_U are given $n \times n$ symmetric matrices, and $X \le H_U$ means that $H_U - X \ge 0$; then $P_{S_B}(X)$ is easy to be carried out and is given by

$$P_{S_B}(X) = \min\left(\max\left(X, H_L\right), H_U\right), \tag{32}$$

where $\max(X, Y)$ and $\min(X, Y)$ compute the element-wise maximum and minimum of matrix X and Y, respectively.

4. Main Results

Let $\{w^k\}$ be the sequence generated by applying the procedure (14)–(16) to problem (18)-(19); then for any $w' = (X', Y', \Lambda') \in \mathscr{W}$, we have that

$$\begin{aligned} &\left\langle X' - X^{k+1}, X^{k+1} - C - \Lambda^{k+1} - \beta\left(Y^k - Y^{k+1}\right) + r\left(X^{k+1} - X^k\right)\right\rangle \ge 0, \\ &\left\langle Y' - Y^{k+1}, Y^{k+1} - C + \Lambda^{k+1} + s\left(Y^{k+1} - Y^k\right)\right\rangle \ge 0, \\ &\Lambda^{k+1} = \Lambda^k - \beta\left(X^{k+1} - Y^{k+1}\right). \end{aligned} \tag{33}$$

Further, letting

$$F\left(w^{k+1}\right) = \begin{pmatrix} X^{k+1} - C - \Lambda^{k+1} \\ Y^{k+1} - C + \Lambda^{k+1} \\ X^{k+1} - Y^{k+1} \end{pmatrix},$$

$$d_1\left(w^k, w^{k+1}\right) = \begin{pmatrix} rI_n & 0 & 0 \\ 0 & (s+\beta) I_n & 0 \\ 0 & 0 & \frac{1}{\beta} I_n \end{pmatrix} \begin{pmatrix} X^k - X^{k+1} \\ Y^k - Y^{k+1} \\ \Lambda^k - \Lambda^{k+1} \end{pmatrix}, \tag{34}$$

where $I_n \in R^{n\times n}$ is the unit matrix, and

$$d_2\left(w^k, w^{k+1}\right) = F\left(w^{k+1}\right) - \beta \begin{pmatrix} I_n \\ -I_n \\ 0 \end{pmatrix} \left(Y^k - Y^{k+1}\right), \tag{35}$$

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*

$$\begin{aligned} &\left\langle w^{k+1} - w^*, d_2\left(w^k, w^{k+1}\right)\right\rangle \\ &\quad \ge -\left\langle \Lambda^k - \Lambda^{k+1}, Y^k - Y^{k+1}\right\rangle + \left\|X^{k+1} - X^*\right\|_F^2 \\ &\quad + \left\|Y^{k+1} - Y^*\right\|_F^2. \end{aligned} \tag{36}$$

Proof. From (22) and (35), we have

$$\left\langle w^{k+1}-w^{*},d_{2}\left(w^{k},w^{k+1}\right)\right\rangle = -\left\langle \Lambda^{k}-\Lambda^{k+1},Y^{k}-Y^{k+1}\right\rangle + \left\langle w^{k+1}-w^{*},F\left(w^{k+1}\right)\right\rangle. \quad (37)$$

Since (9) and w^* are a solution to problem (18)-(19) and $X^{k+1} \in S_+^n$, $Y^{k+1} \in S_B$, we have

$$\left\langle w^{k+1}-w^{*},F\left(w^{*}\right)\right\rangle \geq 0. \quad (38)$$

From (38), it follows that

$$\left\langle w^{k+1}-w^{*},F\left(w^{k+1}\right)-F\left(w^{k+1}\right)+F\left(w^{*}\right)\right\rangle \geq 0. \quad (39)$$

Thus, we have

$$\begin{aligned}
&\left\langle w^{k+1}-w^{*},F\left(w^{k+1}\right)\right\rangle \\
&\geq \left\langle w^{k+1}-w^{*},F\left(w^{k+1}\right)-F\left(w^{*}\right)\right\rangle \\
&= \left\langle X^{k+1}-X^{*},X^{k+1}-X^{*}-\left(\Lambda^{k+1}-\Lambda^{*}\right)\right\rangle \\
&\quad + \left\langle Y^{k+1}-Y^{*},Y^{k+1}-Y^{*}+\left(\Lambda^{k+1}-\Lambda^{*}\right)\right\rangle \\
&\quad + \left\langle \Lambda^{k+1}-\Lambda^{*},X^{k+1}-X^{*}-\left(Y^{k+1}-Y^{*}\right)\right\rangle \\
&= \left\langle X^{k+1}-X^{*},X^{k+1}-X^{*}\right\rangle + \left\langle Y^{k+1}-Y^{*},Y^{k+1}-Y^{*}\right\rangle \\
&= \left\|X^{k+1}-X^{*}\right\|_F^2 + \left\|Y^{k+1}-Y^{*}\right\|_F^2.
\end{aligned} \quad (40)$$

Substituting (40) into (37), we get the assertion of this lemma. □

Lemma 2. *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*

$$\begin{aligned}
&\left\langle w^{k}-w^{*},G_0\left(w^{k}-w^{k+1}\right)\right\rangle \\
&\geq \left\langle w^{k}-w^{k+1},G_0\left(w^{k}-w^{k+1}\right)\right\rangle - \left\langle \Lambda^{k}-\Lambda^{k+1},Y^{k}-Y^{k+1}\right\rangle \\
&\quad + \left\|X^{k+1}-X^{*}\right\|_F^2 + \left\|Y^{k+1}-Y^{*}\right\|_F^2,
\end{aligned} \quad (41)$$

where

$$G_0 = \begin{pmatrix} rI_n & 0 & 0 \\ 0 & (s+\beta)I_n & 0 \\ 0 & 0 & \frac{1}{\beta}I_n \end{pmatrix}. \quad (42)$$

Proof. It follows from (33) that

$$\left\langle w'-w^{k+1},d_{2}\left(w^{k},w^{k+1}\right)-d_{1}\left(w^{k},w^{k+1}\right)\right\rangle \geq 0, \quad \forall w' \in \mathscr{W}. \quad (43)$$

Thus, we have

$$\begin{aligned}
&\left\langle w^{k+1}-w^{*},d_{1}\left(w^{k},w^{k+1}\right)\right\rangle \\
&\geq \left\langle w^{k+1}-w^{*},d_{2}\left(w^{k},w^{k+1}\right)\right\rangle \\
&\geq -\left\langle \Lambda^{k}-\Lambda^{k+1},Y^{k}-Y^{k+1}\right\rangle + \left\|X^{k+1}-X^{*}\right\|_F^2 \\
&\quad + \left\|Y^{k+1}-Y^{*}\right\|_F^2.
\end{aligned} \quad (44)$$

From the above inequality, we get

$$\begin{aligned}
&\left\langle w^{k}-w^{*},G_0\left(w^{k}-w^{k+1}\right)\right\rangle \\
&\geq \left\langle w^{k}-w^{k+1},G_0\left(w^{k}-w^{k+1}\right)\right\rangle \\
&\quad - \left\langle \Lambda^{k}-\Lambda^{k+1},Y^{k}-Y^{k+1}\right\rangle + \left\|X^{k+1}-X^{*}\right\|_F^2 \\
&\quad + \left\|Y^{k+1}-Y^{*}\right\|_F^2.
\end{aligned} \quad (45)$$

Hence, (41) holds and the proof is completed. □

Theorem 3. *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*

$$\left\|w^{k+1}-w^{*}\right\|_G^2 \leq \left\|w^{k}-w^{*}\right\|_G^2 - \left\langle w^{k}-w^{k+1},M\left(w^{k}-w^{k+1}\right)\right\rangle, \quad (46)$$

where

$$G = \begin{pmatrix} (r+1)I_n & 0 & 0 \\ 0 & (1+s+\beta)I_n & 0 \\ 0 & 0 & \frac{1}{\beta}I_n \end{pmatrix},$$
$$M = \begin{pmatrix} \left(\frac{1}{2}+r\right)I_n & 0 & 0 \\ 0 & \left(\frac{1}{2}+s+\beta\right)I_n & -I_n \\ 0 & -I_n & \frac{1}{\beta}I_n \end{pmatrix}, \quad (47)$$

and $\|w\|_G^2 = \langle w, Gw\rangle$.

Proof. From (41), we have

$$\begin{aligned}
&\left\|w^{k+1}-w^{*}\right\|_{G_0}^2 \\
&= \left\|w^{k}-w^{*}-\left(w^{k}-w^{k+1}\right)\right\|_{G_0}^2 \\
&\leq \left\|w^{k}-w^{*}\right\|_{G_0}^2 - 2\left\|w^{k}-w^{k+1}\right\|_{G_0}^2
\end{aligned}$$

$$
\begin{aligned}
&+2\left\langle \Lambda^k-\Lambda^{k+1},Y^k-Y^{k+1}\right\rangle-2\left\|X^{k+1}-X^*\right\|_F^2\\
&-2\left\|Y^{k+1}-Y^*\right\|_F^2+\left\|w^k-w^{k+1}\right\|_{G_0}^2\\
&=\left\|w^k-w^*\right\|_{G_0}^2-\left\|w^k-w^{k+1}\right\|_{G_0}^2\\
&+2\left\langle \Lambda^k-\Lambda^{k+1},Y^k-Y^{k+1}\right\rangle-2\left\|X^{k+1}-X^*\right\|_F^2\\
&-2\left\|Y^{k+1}-Y^*\right\|_F^2.
\end{aligned}
\tag{48}
$$

Rearranging the inequality above, we find that

$$
\begin{aligned}
\left\|w^{k+1}-w^*\right\|_G^2\le&\left\|w^k-w^*\right\|_G^2-\left\langle w^k-w^{k+1},\begin{pmatrix} rI_n & 0 & 0\\ 0 & (s+\beta)I_n & -I_n\\ 0 & -I_n & \frac{1}{\beta}I_n\end{pmatrix}\left(w^k-w^{k+1}\right)\right\rangle-\left(\left\|X^{k+1}-X^*\right\|_F^2+\left\|X^k-X^*\right\|_F^2\right)\\
&-\left(\left\|Y^{k+1}-Y^*\right\|_F^2+\left\|Y^k-Y^*\right\|_F^2\right).
\end{aligned}
\tag{49}
$$

Using the Cauchy-Schwarz Inequality on the last term of the right-hand side of (49), we obtain

$$
\begin{aligned}
\left\|X^{k+1}-X^*\right\|_F^2+\left\|X^k-X^*\right\|_F^2&\ge\frac{1}{2}\left\|X^{k+1}-X^k\right\|_F^2,\\
\left\|Y^{k+1}-Y^*\right\|_F^2+\left\|Y^k-Y^*\right\|_F^2&\ge\frac{1}{2}\left\|Y^{k+1}-Y^k\right\|_F^2.
\end{aligned}
\tag{50}
$$

Substituting (50) into (49), we get

$$
\begin{aligned}
\left\|w^{k+1}-w^*\right\|_G^2\le&\left\|w^k-w^*\right\|_G^2\\
&-\left\langle w^k-w^{k+1},M\left(w^k-w^{k+1}\right)\right\rangle.
\end{aligned}
\tag{51}
$$

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)*, $w^*\in\mathscr{W}^*$ 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) $\lim_{k\to\infty}\|w^{k+1}-w^k\|_F^2=0$;

(4) *G and M are both symmetric positive-definite matrices.*

Proof. Since

$$
\begin{vmatrix}\left(\dfrac{1}{2}+s+\beta\right)I_n & -I_n\\ -I_n & \dfrac{1}{\beta}I_n\end{vmatrix}=\frac{((1/2)+s)}{\beta},
\tag{52}
$$

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

$$
\left\|w^{k+1}-w^*\right\|_G^2\le\left\|w^k-w^*\right\|_G^2-\tau\left\|w^k-w^{k+1}\right\|_F^2.
\tag{53}
$$

Following (53), we immediately have that $\|w^k-w^*\|_G^2$ is nonincreasing and thus the sequence $\{w^k\}$ is bounded. Moreover, we have

$$
\left\|w^{k+1}-w^*\right\|_G^2\le\left\|w^0-w^*\right\|_G^2-\tau\sum_{j=0}^{k}\left\|w^j-w^{j+1}\right\|_F^2.
\tag{54}
$$

So, we get

$$
\sum_{j=0}^{k}\left\|w^j-w^{j+1}\right\|_F^2<\infty,\quad \forall k>0,
\tag{55}
$$

then

$$
\lim_{k\to\infty}\left\|w^k-w^{k+1}\right\|_F^2=0.
\tag{56}
$$

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^∞ be a cluster point of $\{w^k\}$ and the subsequence $\{w^{k_j}\}$ converges to w^∞. It follows from (33) that

$$
\begin{aligned}
\lim_{j\to\infty}&\left\langle X'-X^{k_j+1},X^{k_j+1}-C-\Lambda^{k_j+1}-\beta\left(Y^{k_j}-Y^{k_j+1}\right)\right.\\
&\left.+r\left(X^{k_j+1}-X^{k_j}\right)\right\rangle\ge0,
\end{aligned}
$$

Table 1: Numerical results of Example 6.

n	$r=-0.3, s=-0.3$		$r=0, s=0$		$r=3, s=3$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	31	0.292	34	0.331	72	0.764
200	33	1.346	39	1.570	84	3.364
300	38	4.265	41	5.746	90	9.991
400	40	9.872	43	9.919	94	22.03
500	39	15.83	45	18.39	98	39.91

$$\lim_{j\to\infty}\left\langle Y'-Y^{k_j+1},Y^{k_j+1}-C+\Lambda^{k_j+1}+s\left(Y^{k_j+1}-Y^{k_j}\right)\right\rangle \geq 0,\quad \forall w'\in\mathscr{W},$$
$$\lim_{j\to\infty}\Lambda^{k_j+1}=\Lambda^{k_j}-\beta\left(X^{k_j+1}-Y^{k_j+1}\right). \tag{57}$$

Following point (3) of Lemma 4, we have

$$\begin{gathered}\left\langle X'-X^{\infty},X^{\infty}-C-\Lambda^{\infty}\right\rangle\geq 0,\\ \left\langle Y'-Y^{\infty},Y^{\infty}-C+\Lambda^{\infty}\right\rangle\geq 0,\quad\forall w'\in\mathscr{W},\\ X^{\infty}-Y^{\infty}=0.\end{gathered} \tag{58}$$

This means that w^{∞} is a solution point of (18)-(19). Since $\{w^{k_j}\}$ converges to w^{∞}, we have that, for any given $\varepsilon>0$, there exists an integer $N>0$ such that

$$\left\|w^{k_j}-w^{\infty}\right\|_G^2<\varepsilon,\quad\forall k_j\geq N. \tag{59}$$

Furthermore, using the inequality (53), we have

$$\left\|w^{k}-w^{\infty}\right\|_G^2<\left\|w^{k_j}-w^{\infty}\right\|_G^2,\quad\forall k\geq k_j. \tag{60}$$

Combining (59) and (60), we get that

$$\left\|w^{k}-w^{\infty}\right\|_G^2<\varepsilon,\quad\forall k>N. \tag{61}$$

This implies that the sequence $\{w^k\}$ converges to w^{∞}. 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_B be in the form of (31) and

$$\begin{gathered}H_L=\left(l_{ij}\right)\in R^{n\times n},\\ l_{ij}=\begin{cases}-0.5, & i\neq j\\ 1, & i=j,\end{cases}\quad i,j=1,2,\ldots,n,\\ H_U=\left(u_{ij}\right)\in R^{n\times n},\\ u_{ij}=\begin{cases}0.5, & i\neq j\\ 1, & i=j,\end{cases}\quad i,j=1,2,\ldots,n.\end{gathered} \tag{62}$$

Moreover, let $X^0=\text{eye}(n)$, $Y^0=\text{eye}(n)$, $\Lambda^0=\text{zeroes}(n)$, $\beta=4$, and $\varepsilon=10^{-6}$, where eye(n) and zeroes(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

$$\left\|w^{k+1}-w^{k}\right\|_{\max}<\varepsilon, \tag{63}$$

where $\|X\|_{\max}=\max(\max(\text{abs}(X)))$ 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 $[-1000,1000]$,

$$\begin{gathered}H_L=\left(l_{ij}\right)\in R^{n\times n},\\ l_{ij}=\begin{cases}-500, & i\neq j\\ 1000, & i=j,\end{cases}\quad i,j=1,2,\ldots,n,\\ H_U=\left(u_{ij}\right)\in R^{n\times n},\\ u_{ij}=\begin{cases}500, & i\neq j\\ 1000, & i=j,\end{cases}\quad i,j=1,2\ldots,n.\end{gathered} \tag{64}$$

The computational results are reported in Table 2.

Example 9. Let S_B 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, Λ_0, β, ε, and the stopping criterion are the same as those in Example 6, but the diagonal elements of matrix C are replaced by

$$C_{ii}=\alpha+(1-\alpha)\times\text{rand},\quad i=1,2,\ldots,n, \tag{65}$$

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 2: Numerical results of Example 8.

n	$r=-0.3, s=-0.3$		$r=0, s=0$		$r=3, s=3$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	49	0.476	54	0.551	116	1.837
200	51	2.197	57	2.334	128	5.430
300	59	6.614	61	8.108	136	15.25
400	56	12.74	63	14.51	140	31.65
500	58	23.90	66	26.90	147	59.98

Table 3: Numerical results of Example 9.

n	$r=-0.3, s=-0.3$		$r=0, s=0$		$r=3, s=3$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	32	0.282	35	0.288	70	0.566
200	33	1.295	36	1.397	72	4.006
300	34	3.745	37	4.156	73	8.285
400	34	7.885	37	8.571	73	16.73
500	34	14.07	37	15.42	74	29.87

Table 4: Numerical results of Example 10.

5	$r=-0.3, s=-0.3$		$r=0, s=0$		$r=3, s=3$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	32	0.259	35	0.300	70	0.557
200	33	1.306	36	1.424	72	2.880
300	33	3.750	37	4.087	72	7.958
400	34	7.799	37	8.546	74	16.98
500	34	13.96	37	16.10	74	30.77

Table 5: (a) Numerical results of Example 11 with $\hat{n}_r = 5$. (b) Numerical results of Example 11 with $\hat{n}_r = 10$.

(a)

n	$r=-0.3,\ s=-0.3$		$r=0,\ s=0$		$r=1,\ s=1$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	22	0.293	25	0.354	34	0.448
200	25	2.119	28	2.425	40	3.436
300	27	7.141	30	8.024	44	11.64
400	29	17.40	31	18.59	46	27.32
500	30	34.17	33	37.45	48	53.84

(b)

n	$r=-0.3,\ s=-0.3$		$r=0,\ s=0$		$r=1,\ s=1$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	23	0.309	25	0.342	33	0.439
200	24	2.029	27	2.305	38	3.162
300	27	7.150	29	7.801	42	11.29
400	28	16.68	31	18.47	45	26.60
500	29	32.73	32	36.37	47	53.06

Example 10. All the data are the same as in Example 9 except that $\alpha = 0$. The computational results are reported in Table 4.

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

$$S_B = \{X \in R^{n\times n} \mid X = X^T, X_{ij} = e_{ij}, (i,j) \in \mathscr{B}_e, \quad X_{ij} \geq l_{ij}, (i,j) \in \mathscr{B}_l, \quad X_{ij} \leq u_{ij}, (i,j) \in \mathscr{B}_u\}, \tag{66}$$

where $\mathscr{B}_e$, $\mathscr{B}_l$, $\mathscr{B}_u$ 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 $\mathscr{B}_e$, $\mathscr{B}_l$, and $\mathscr{B}_u$ be the same as in Example 5.4 of [3]; that is, $\mathscr{B}_e = \{(i,i) \mid i = 1,2,\ldots,n\}$ and $\mathscr{B}_l, \mathscr{B}_u \subset \{(i,j) \mid 1 \leq i < j \leq n\}$ consist of the indices of $\min(\hat{n}_r, n-i)$ randomly generated elements at the ith row of X, $i = 1,2,\ldots,n$ with $\hat{n}_r = 5$ and $\hat{n}_r = 10$, respectively. We take $e_{ii} = 1$ for $(i,i) \in \mathscr{B}_e$, $l_{ij} = -0.1$ for $(i,j) \in \mathscr{B}_l$, and $u_{ij} = 0.1$ for $(i,j) \in \mathscr{B}_u$.

Moreover, let X_0, Y_0, Λ_0, β, ε, 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 $\hat{n}_r$, 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 (1).

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 (1) with S_B 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 $k = 1, 2, \ldots$
$R_k = Y_{k-1} - \Delta S_{k-1};$
$X_k = P_{S_+^n}(R_k);$
$\Delta S_k = X_k - R_k;$
$Y_k = P_{S_B}(X_k);$
end.

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

$$\max\{\|X_k - X_{k-1}\|_{\max}, \|Y_k - Y_{k-1}\|_{\max}, \|X_k - Y_k\|_{\max}\} < \varepsilon. \tag{67}$$

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.

Table 6: (a) Numerical results of Example 12. (b) Numerical results of Example 12.

(a)

n	ADM		APM	
	It.	CPU.	It.	CPU.
100	28	0.381	47	0.743
200	33	2.878	59	5.443
300	36	9.462	70	20.68
400	38	22.50	81	54.38
500	39	43.32	89	114.7

(b)

n	ADM		APM	
	It.	CPU.	It.	CPU.
100	27	0.634	42	0.582
200	30	2.590	59	5.428
300	32	8.524	65	19.36
400	34	20.34	75	50.79
500	35	39.43	86	111.6

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

[1] R. Borsdorf and N. J. Higham, "A preconditioned Newton algorithm for the nearest correlation matrix," *IMA Journal of Numerical Analysis*, vol. 30, no. 1, pp. 94–107, 2010.

[2] S. Boyd and L. Xiao, "Least-squares covariance matrix adjustment," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 2, pp. 532–546, 2005.

[3] Y. Gao and D. Sun, "Calibrating least squares semidefinite programming with equality and inequality constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1432–1457, 2009.

[4] S. Gravel and V. Elser, "Divide and concur: a general approach to constraint satisfaction," *Physical Review E*, vol. 78, Article ID 036706, 2008.

[5] N. J. Higham, "Computing a nearest symmetric positive semidefinite matrix," *Linear Algebra and its Applications*, vol. 103, pp. 103–118, 1988.

[6] N. J. Higham, "Computing the nearest correlation matrix—a problem from finance," *IMA Journal of Numerical Analysis*, vol. 22, no. 3, pp. 329–343, 2002.

[7] N. J. Higham, "Matrix nearness problems and applications," in *Applications of Matrix Theory*, M. Gover and S. Barnett, Eds., vol. 22, pp. 1–27, Oxford University Press, Oxford, UK, 1989.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[9] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.

[10] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11/12, no. 1–4, pp. 625–653, 1999.

[11] R. H. Tütüncü, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming*, vol. 95, no. 2, pp. 189–217, 2003.

[12] J. Malick, "A dual approach to semidefinite least-squares problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 1, pp. 272–284, 2004.

[13] H. Qi and D. Sun, "A quadratically convergent Newton method for computing the nearest correlation matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 2, pp. 360–385, 2006.

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

[15] P. M. Pardalos and M. G. C. Resende, *Handbook of Applied Optimization*, Oxford University Press, Oxford, UK, 2002.

[16] P. M. Pardalos, T. M. Rassias, and A. A. Khan, *Nonlinear Analysis and Variational Problems*, vol. 35 of *Springer Optimization and Its Applications*, Springer, New York, NY, USA, 2010, In honor of George Isac, Edited by Panos M. Pardalos, Themistocles M. Rassias and Akhtar A. Khan.

[17] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Review*, vol. 38, no. 3, pp. 367–426, 1996.

[18] W. Zhang, D. Han, and Z. Li, "A self-adaptive projection method for solving the multiple-sets split feasibility problem," *Inverse Problems*, vol. 25, no. 11, 2009.

[19] W. Zhang, D. Han, and X. Yuan, "An efficient simultaneous method for the constrained multiple-sets split feasibility problem," *Computational Optimization and Applications*, vol. 52, no. 3, pp. 825–843, 2012.

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

[21] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, vol. 9 of

SIAM Studies in Applied Mathematics, SIAM, Philadelphia, Pa, USA, 1989.

[22] J. Eckstein, "Some saddle-function splitting methods for convex programming," *Optimization Methods and Software*, vol. 4, pp. 75–83, 1994.

[23] J. Eckstein and M. Fukushima, "Some reformulations and applications of the alternating direction method of multipliers," in *Large Scale Optimization: State of the Art*, W. W. Hager, D. W. Hearn, and P. M. Pardalos, Eds., pp. 115–134, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994.

[24] M. Fukushima, "Application of the alternating direction method of multipliers to separable convex programming problems," *Computational Optimization and Applications*, vol. 1, no. 1, pp. 93–111, 1992.

[25] B. He and H. Yang, "Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities," *Operations Research Letters*, vol. 23, no. 3-5, pp. 151–161, 1998.

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

[27] M. H. Xu, "Proximal alternating directions method for structured variational inequalities," *Journal of Optimization Theory and Applications*, vol. 134, no. 1, pp. 107–117, 2007.

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

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

[30] O. Güler, "New proximal point algorithms for convex minimization," *SIAM Journal on Optimization*, vol. 2, no. 4, pp. 649–664, 1992.

[31] W. W. Hager and H. Zhang, "Asymptotic convergence analysis of a new class of proximal point methods," *SIAM Journal on Control and Optimization*, vol. 46, no. 5, pp. 1683–1704, 2007.

[32] W. W. Hager and H. Zhang, "Self-adaptive inexact proximal point methods," *Computational Optimization and Applications*, vol. 39, no. 2, pp. 161–181, 2008.

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

[34] M. Teboulle, "Convergence of proximal-like algorithms," *SIAM Journal on Optimization*, vol. 7, no. 4, pp. 1069–1083, 1997.

[35] W. K. Glunt, "An alternating projections method for certain linear problems in a Hilbert space," *IMA Journal of Numerical Analysis*, vol. 15, no. 2, pp. 291–305, 1995.

[36] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, NY, USA, 1999.

[37] N. Narendra, "A new polynomial time algorithm for linear programming," *Combinatorica*, vol. 4, pp. 373–395, 1987.