A Modified Alternating Direction Method for Convex Quadratically Constrained Quadratic Semidefinite Programs

Jie Sun*and Su Zhang†

March 4, 2009

Abstract

We propose a modified alternate direction method for solving convex quadratically constrained quadratic semidefinite optimization problems. The method is a first order method, therefore requires much less computational effort per iteration than the second-order approaches such as the interior point methods or the smoothing Newton methods. At each iteration only a single inexact metric projection onto the positive semidefinite cone is required. We prove the global convergence of this method.

Key Words: Alternating direction method, Convex cone programming, Quadratic semidefinite optimization

---

*Department of Decision Sciences and Singapore-MIT Alliance, National University of Singapore, Email:jsun@nus.edu.sg, Fax: +65-6779-2621

†Department of Decision Sciences, National University of Singapore, Email:zhangsu@nus.edu.sg.

1. Introduction

Let $\mathbb{S}^n$ be the space of all $n \times n$ real symmetric matrices. We consider a special class of nonlinear optimization problems defined in $\mathbb{S}^n$, called *Convex Quadratically Constrained Quadratic Semidefinite Programs* (CQCQSDPs), as follows.

$$\begin{aligned} \min \quad & q_0(X) \equiv \frac{1}{2} \langle X, Q_0(X) \rangle + \langle B_0, X \rangle \\ \text{s.t.} \quad & q_i(X) \equiv \frac{1}{2} \langle X, Q_i(X) \rangle + \langle B_i, X \rangle + c_i \le 0, \quad i = 1, \cdots, m \\ & X \succeq 0, \end{aligned} \tag{1.1}$$

where $Q_i : \mathbb{S}^n \to \mathbb{S}^n, i = 0, 1, ..., m$, is a self-adjoint positive semidefinite linear operator; $X, B_i \in \mathbb{S}^n$, and $c_i \in \mathbb{R}$ is a scalar. In addition, $\langle \cdot, \cdot \rangle$ denotes the Frobenius inner product between square matrices, i.e., $\langle U, V \rangle = \text{Trace}(U^T V)$. We denote by $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ the cones of positive semidefinite matrices and positive definite matrices, respectively. By writing $X \succeq 0$ ($X \succ 0$) we mean that $X \in \mathbb{S}^n_+$ ($X \in \mathbb{S}^n_{++}$). Preliminary examples of $Q(X)$ include the symmetrized Kronecker product $U \otimes V(X) = (VXU^T + UX^TV^T)/2$, where U and V are both $n \times n$ symmetric matrices, and the Hadamard product $H \circ X$ defined as $(H \circ X)_{ij} = H_{ij}X_{ij}$, etc.

Let **vec** be an isometry identifying $\mathbb{S}^n$ with $\mathbb{R}^{n^2}$ so that $\langle B, X \rangle = \mathbf{vec}\,(B)^T \mathbf{vec}\,(X)$. Let the matrix representation of operator Q under this isometry be $\bar{Q}$. Then for any X, we have $\mathbf{vec}\,(Q(X)) = \bar{Q}\,\mathbf{vec}\,(X)$. Since Q is self-adjoint and positive semidefinite, $\bar{Q}$ is a symmetric positive semidefinite matrix.

Problem (1.1) is a convex programming problem in the symmetric matrix space. In a sense, it is also the most basic nonlinear semidefinite optimization problem. It has a number of important applications in engineering and management. For example, in order to find a positive semidefinite matrix that best approximates the solution to the matrix equation system

$$\langle A_i, X \rangle = b_i, \qquad \forall i = 1, ..., m,$$

we need to solve the matrix least-square problem

$$\min \sum_{i=1}^{m} \| \langle A_i, X \rangle - b_i \|^2 \quad \text{s.t. } X \succeq 0, \tag{1.2}$$

which is in the form of Problem (1.1). Another example is the following nearest covariance matrix problem in finance,

$$\begin{aligned} \min \quad & \|X - C\|^2 \\ \text{s.t.} \quad & \langle A_i, X \rangle = b_i, \quad i = 1, ..., p, \\ & \langle A_i, X \rangle \le b_i, \quad i = p+1, ..., m, \\ & X \text{ is a covariance matrix.} \end{aligned}$$

Since all covariance matrices are positive semidefinite with rank one, relaxing the rank requirement, we can substitute the last constraint by $X \succeq 0$. The resulting relaxation problem then becomes a special case of (1.1).

In [1], Beck studied quadratic matrix programming of order r which may not be convex. He constructed a special semidefinite relaxation and its dual and showed that under some mild conditions strong duality holds for the relaxed problem with at most r constraints. However, Beck's models does not include the semidefinite cone constraint. Therefore, it is essentially a vector optimization model, rather than a semidefinite optimization problem like (1.1).

A special case of (1.1) is the convex quadratic SDP (CQSDP) problem where the quadratic term only appears in the objective and the constraints are linear, together with the semidefinite cone constraint. In [14], a theoretical primal-dual potential reduction algorithm was proposed for CQSDP problems by Nie and Yuan. The authors suggested to use the conjugate gradient method to compute an approximate search direction. Subsequent works include Qi and Sun [16] and Toh [19]. Qi and Sun used a Lagrangian dual approach. Toh introduced an inexact primal-dual path-following method with three classes of pre-conditioners for the augmented equation for fast convergence under suitable nondegeneracy assumptions. Unfortunately, all these new methods can not be readily extended to solve CQCQSDP because the KKT conditions of (1.1) can no longer be written as a linear system as in CQSDP.

In two recent papers, Malick [13] and Boyd and Xiao [2], respectively applied classical quasi-Newton methods (in particular, the BFGS method) and the projected gradient method to the dual problem of linearly constrained quadratic semidefinite programs with continuously differentiable objective function. More recently, Gao and Sun [9] designed an inexact smoothing Newton method to solve a reformulated semismooth system with two level metric projection operators and showed high efficiency of the proposed method in numerical experiments. Again, all these algorithms are not applicable to CQCQSDP problem because of the existence of quadratic constraints.

In the field of general nonlinear semidefinite optimization, we notice the smoothing Newton algorithm of Sun, Sun, and Qi [18] and the augmented Lagrangian algorithm of Sun, Sun, and Zhang [17]. Although both algorithms could be used to solve CQCQSDP, they are not specifically designed for CQCQSDP. Therefore, they may not be able to take advantage of the specific structure of the problem. It is also noted that the smoothing Newton method is a second-order algorithm while the proposed algorithm in this paper is a first-order method. Therefore, in each iteration, the smoothing Newton method requires much more computations than our algorithm. Similar comparison can be also made between the augmented Lagrangian method and our algorithm. It should be noted that the successive linearization method of Li and Sun [12] is also a first order method, but it requires to solve a quadratic sub-problem at each iteration. Besides, it is targeted for general nonlinear semidefinite programs.

The general advantage of the first-order algorithms are twofold. First, this type of methods are relatively simple to implement, thus they are useful in finding an approximate solution of the problems, which may become the "first phase" of a hybrid second-order algorithm (e.g., a Newton method). Secondly, first-order methods usually require much less computation per iteration, therefore might be suitable for relatively large problems.

This paper is focused on convergence analysis of an alternating direction method (ADM) for CQCQSDP. The ADM has been an effective first-order approach for solving large optimization problems with vector variables. It was probably first considered by Gabay [7] and Gabay and Mercier [8]. Further studies of such methods can be found, for instance, in [3, 6, 10, 11]. The inexact versions of the ADM was proposed by Eckstein and Bertsekas [6] and Chen and Teboulle [3], respectively. He et al. [10] generalized the framework and proposed a new inexact ADM method with flexible conditions for structured monotone variational inequalities (VI). All of the work above, however, was devoted to vector optimization problems. It appears to be new to apply the idea of ADM to develop a method for solving quadratic semidefinite programming problems.

In [20], Yu applied an exact ADM for a special case of (1.1), in which both the objective function and the general constraints are linear. His main idea is to reformulate the primal-dual optimality conditions of SDP as a projection equation. However, there is a significant difficulty in applying Yu's idea to CQCQSDP because one has to solve a nonlinear variational problem on semidefinite cone at each iteration, which is very expensive. In this paper we propose a specially designed ADM for CQCQSDP. The new method has the advantage of being simple and cheap in computation. At each iteration only one projection onto the semidefinite cone is necessary. Furthermore, this projection is allowed to be inexact. Under some mild conditions, we prove the convergence of this method.

The paper is organized as follows. In Section 2 we reformulate Problem (1.1) as a variational inequality problem and present a specialization of the ADM, called the "original ADM". We then simplify the original ADM into a "modified ADM", which takes into account the special structure of the reformulated variational inequality problem. In Section 3 we prove the convergence of the modified ADM. Section 4 is devoted to the inexact case and its convergence. Finally we make some concluding remarks in Section 5.

2. The Algorithms

Recall that $q_i(X) \equiv \frac{1}{2}\langle X, Q_i(X)\rangle + \langle B_i, X\rangle + c_i \le 0$. By introducing artificial constraints

$$Y_i = X \text{ and } \Omega_i = \{Y_i : q_i(Y_i) \le 0\}, \quad \forall i = 1, \cdots, m, \tag{2.1}$$

we may rewrite (1.1) equivalently as

$$\begin{array}{ll} \min & q_0(X) \\ \text{s.t.} & X = Y_i,\ Y_i \in \Omega_i,\ i = 1, \cdots, m \\ & X \succeq 0. \end{array} \tag{2.2}$$

The Lagrange dual of problem (2.2) is

$$\max_{\lambda_i} \min_{X\succeq 0, Y_i \in \Omega_i} L(X, Y_1, \cdots, Y_m, \lambda_1, \cdots, \lambda_m) := q_0(X) - \sum_{i=1}^{m} \langle \lambda_i, X - Y_i\rangle$$

Notice that the Lagrange multipliers $\lambda_i,\ i = 1, ..., m$, are symmetric matrices. It is well known that under mild constraint qualifications (e.g., Slater' condition), strong duality holds and hence, X^* is a solution of (2.2) if and only if there exists $\lambda_i^* \in S^n,\ i = 1, ..., m$ such that $(X^*, Y_i^*, \lambda_i^*)$ satisfies

$$\left\{ \begin{array}{l} \left\langle X - X^*, Q_0(X^*) + B_0 - \sum\limits_{i=1}^{m} \lambda_i^* \right\rangle \ge 0,\ \forall\ X \in \mathbb{S}_+^n \\ \langle Y_i - Y_i^*, \lambda_i^*\rangle \ge 0,\ \forall\ Y_i \in \Omega_i,\ i = 1, \cdots, m \\ X^* = Y_i^*,\ i = 1, \cdots, m \end{array} \right. \tag{2.3}$$

Problem (2.3) is a variational inequality problem with a special structure. The variables (X, Y_i, λ_i) are symmetric matrices, the underlying sets $\mathbb{S}_+^n$ and Ω_i are convex and non-polyhedral.

In order to solve (2.3) by an alternate direction method, we need to compute the metric projection of a matrix onto Ω_i and $\mathbb{S}^n_+$. The projection onto Ω_i can be computed in a similar way as computing the Euclidean projection of a vector onto an ellipsiod in the real vector space. Therefore, the computation of this projection can be very fast, see, for example, [4] for the corresponding algorithms. The projection onto $\mathbb{S}^n_+$ repuires a full eigenvalue decomposition. To handle potetially large problems, we allow this computation to be performed inexactly. The alternating direction methods, when applied to problem (2.3), can be separated into three steps. The interested reader can refer to [11] for the motivation of these steps.

Algorithm 2.1. The Original Alternating Direction Method for CQCQSDP

Step 1. $(X^k, Y_i^k, \lambda_i^k) \to (X^{k+1}, Y_i^k, \lambda_i^k)$, *where*

$$\left\langle X - X^{k+1}, Q_0(X^{k+1}) + B_0 - \sum_{i=1}^{m} \left(\lambda_i^k - \beta_i \left(X^{k+1} - Y_i^k \right) \right) \right\rangle \geq 0, \ \forall X \succeq 0 \tag{2.4}$$

where β_i, $i = 1, \cdots, m$ *are certain positive scalars.*

Step 2. $(X^{k+1}, Y_i^k, \lambda_i^k) \to (X^{k+1}, Y_i^{k+1}, \lambda_i^k)$, $i = 1, \cdots, m$, *where*

$$\left\langle Y_i - Y_i^{k+1}, \lambda_i^k - \beta_i \left(X^{k+1} - Y_i^{k+1} \right) \right\rangle \geq 0, \ \forall \ Y_i \in \Omega_i \text{ and certain } \beta_i > 0. \tag{2.5}$$

Step 3. $(X^{k+1}, Y_i^{k+1}, \lambda_i^k) \to (X^{k+1}, Y_i^{k+1}, \lambda_i^{k+1})$, $i = 1, \cdots, m$, *where*

$$\lambda_i^{k+1} = \lambda_i^k - \beta_i \left(X^{k+1} - Y_i^{k+1} \right) \tag{2.6}$$

At Step 1 and Step 2 we should solve two variational inequalities. In the following, we will convert them to simple projection operations. Firstly, it is easy to see that (2.5) is equivalent to the following nonlinear equation

$$Y_i^{k+1} = P_{\Omega_i} \left[Y_i^{k+1} - \alpha_i \left(\lambda_i^k - \beta_i \left(X^{k+1} - Y_i^{k+1} \right) \right) \right],$$

where α_i can be any positive number. Thus by choosing $\alpha_i = \frac{1}{\beta_i}$, the right hand side item Y_i^{k+1} is cancelled. That is, in order to solve (2.5), we only have to compute

$$Y_i^{k+1} = P_{\Omega_i} \left[X^{k+1} - \frac{1}{\beta_i} \lambda_i^k \right]. \tag{2.7}$$

However, it does not work for (2.4) since it is generally impossible to select an α so that the right hand side X^{k+1} is cancelled in

$$X^{k+1} = P_{S^n_+} \left[X^{k+1} - \alpha \left(Q_0(X^{k+1}) + B_0 - \sum_{i=1}^{m} \left(\lambda_i^k - \beta_i \left(X^{k+1} - Y_i^k \right) \right) \right) \right].$$

We therefore suggest the following approximate approach. Let

$$R(X^k, X^{k+1}) \equiv Q_0(X^{k+1}) - Q_0(X^k) - \gamma \left(X^{k+1} - X^k \right)$$

for certain constant γ.

Note that

$$\begin{aligned} X^{k+1} &= P_{S_+^n}\left[X^{k+1}-\alpha\left(Q_0(X^{k+1})+B_0-\sum_{i=1}^m\left(\lambda_i^k-\beta_i\left(X^{k+1}-Y_i^k\right)\right)-R(X^k,X^{k+1})\right)\right] \\ &= P_{S_+^n}\left[X^{k+1}-\alpha\left(\left(\sum_{i=1}^m\beta_i+\gamma\right)X^{k+1}+B_0-\sum_{i=1}^m\left(\lambda_i^k+\beta_iY_i^k\right)-\gamma X^k+Q_0(X^k)\right)\right]. \end{aligned} \tag{2.8}$$

We choose γ so that $\gamma \geq \lambda_{\max}(Q_0)$, where $\lambda_{\max}(Q_0)$ is the largest eigenvalue of Q_0. Setting

$$\alpha = \left(\sum_{i=1}^m \beta_i + \gamma\right)^{-1} \quad \text{and} \quad D = B_0 - \sum_{i=1}^m \left(\lambda_i^k + \beta_i Y_i^k\right) - \gamma X^k + Q_0(X^k),$$

we obtain

$$X^{k+1} = P_{S_+^n}\left[-\alpha D\right]. \tag{2.9}$$

In summary, the modified alternating direction method is given as follows.

Algorithm 2.2. The Modified Alternating Direction Method for CQCQSDP

Step 1. $(X^k, Y_i^k, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^k, \lambda_i^k)$, *where*

$$X^{k+1} = P_{S_+^n}\left[-\left(\sum_{i=1}^m \beta_i + \gamma\right)^{-1} D\right]. \tag{2.10}$$

Step 2. $(X^{k+1}, Y_i^k, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^{k+1}, \lambda_i^k)$, $i = 1, \cdots, m$, *where* $\beta_i > 0$ *and*

$$Y_i^{k+1} = P_{\Omega_i}\left[X^{k+1} - \frac{1}{\beta_i}\lambda_i^k\right]. \tag{2.11}$$

step 3. $(X^{k+1}, Y_i^{k+1}, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^{k+1}, \lambda_i^{k+1})$, $i = 1, \cdots, m$, *where*

$$\lambda_i^{k+1} = \lambda_i^k - \beta_i\left(X^{k+1} - Y_i^{k+1}\right). \tag{2.12}$$

3. Convergence Results

Proposition 3.1. *The sequence* $\{X^k, Y_i^k, \lambda_i^k\}$ *generated by the modified alternating direction method for CQCQSDP satisfies*

$$\sum_{i=1}^m \frac{1}{\beta_i}\left\langle \lambda_i^{k+1} - \lambda_i^*, \lambda_i^k - \lambda_i^{k+1}\right\rangle + \sum_{i=1}^m \beta_i \left\langle Y_i^{k+1} - Y_i^*, Y_i^k - Y_i^{k+1}\right\rangle + \left\langle X^{k+1} - X^*, R(X^k, X^{k+1})\right\rangle \geq 0, \tag{3.1}$$

where (X^*, Y^*, λ^*) *are defined as in (2.3).*

Proof. Using (2.3) and (2.5), we have

$$\left\langle Y_i^{k+1} - Y_i^*, \lambda_i^* - \lambda_i^{k+1}\right\rangle \geq 0. \tag{3.2}$$

Similarly, we get

$$\left\langle Y_i^{k+1} - Y_i^k, \lambda_i^k - \lambda_i^{k+1} \right\rangle \geq 0. \tag{3.3}$$

Note that (2.8) can be written equivalently as

$$\left\langle X - X^{k+1}, Q_0(X^{k+1}) + B_0 - \sum_{i=1}^{m} \lambda_i^{k+1} - \sum_{i=1}^{m} \beta_i \left(Y_i^k - Y_i^{k+1} \right) - R(X^k, X^{k+1}) \right\rangle \geq 0, \ \forall X \in S_+^n.$$

Setting $X = X^*$ in it, we obtain

$$\left\langle X^{k+1} - X^*, -Q_0(X^{k+1}) - B_0 + \sum_{i=1}^{m} \lambda_i^{k+1} + \sum_{i=1}^{m} \beta_i \left(Y_i^k - Y_i^{k+1} \right) + R(X^k, X^{k+1}) \right\rangle \geq 0. \tag{3.4}$$

Let $X = X^{k+1}$ in inequality (2.3). Then

$$\left\langle X^{k+1} - X^*, Q_0(X^*) + B_0 - \sum_{i=1}^{m} \lambda_i^* \right\rangle \geq 0 \tag{3.5}$$

Adding (3.4) and (3.5) together, it follows that

$$\begin{aligned} & \left\langle \sum_{i=1}^{m} \left(\lambda_i^{k+1} - \lambda_i^* \right), X^{k+1} - X^* \right\rangle + \left\langle \sum_{i=1}^{m} \beta_i \left(Y_i^k - Y_i^{k+1} \right), X^{k+1} - X^* \right\rangle \\ + & \left\langle X^{k+1} - X^*, R(X^k, X^{k+1}) \right\rangle \geq \left\langle X^{k+1} - X^*, Q_0(X^{k+1}) - Q_0(X^*) \right\rangle \geq 0. \end{aligned} \tag{3.6}$$

It follows from (3.2), (3.3) and (3.6) that

$$\begin{aligned} & \left\langle \sum_{i=1}^{m} \left(\lambda_i^{k+1} - \lambda_i^* \right), X^{k+1} - X^* \right\rangle + \left\langle \sum_{i=1}^{m} \beta_i \left(Y_i^k - Y_i^{k+1} \right), X^{k+1} - X^* \right\rangle \\ & + \left\langle X^{k+1} - X^*, R(X^k, X^{k+1}) \right\rangle + \sum_{i=1}^{m} \left\langle Y_i^{k+1} - Y_i^*, \lambda_i^* - \lambda_i^{k+1} \right\rangle \\ & + \sum_{i=1}^{m} \left\langle Y_i^{k+1} - Y_i^k, \lambda_i^k - \lambda_i^{k+1} \right\rangle \\ = & \sum_{i=1}^{m} \frac{1}{\beta_i} \left\langle \lambda_i^{k+1} - \lambda_i^*, \lambda_i^k - \lambda_i^{k+1} \right\rangle + \sum_{i=1}^{m} \beta_i \left\langle Y_i^{k+1} - Y_i^*, Y_i^k - Y_i^{k+1} \right\rangle \\ & + \left\langle X^{k+1} - X^*, R(X^k, X^{k+1}) \right\rangle \geq 0. \end{aligned}$$

The proof is completed. □

The following is the main theorem.

Theorem 3.2. *The sequence $\{X^k\}$ generated by the Modified Alternating Direction method converges to a solution point X^* of system (2.3).*

Proof. We denote

$$W \equiv \begin{pmatrix} X \\ Y_i \\ \lambda_i \end{pmatrix}, \quad \text{and} \quad G \equiv \begin{pmatrix} \gamma I - Q_0 & 0 & 0 \\ 0 & H & 0 \\ 0 & 0 & H^{-1} \end{pmatrix},$$

where H is a diagonal matrix with β_i on its diagonal. clearly, G is positive definite. We define the G-inner product of W and W' as

$$\left\langle W, W' \right\rangle_G \equiv \sum_{i=1}^{m} \frac{1}{\beta_i} \left\langle \lambda_i, \lambda_i^{'} \right\rangle + \sum_{i=1}^{m} \beta_i \left\langle Y_i, Y_i^{'} \right\rangle + \left\langle X, \gamma X^{'} - Q_0(X^{'}) \right\rangle$$

and the associated G-norm as

$$\|W\|_G \equiv \left(\sum_{i=1}^{m} \frac{1}{\beta_i} \|\lambda_i\|^2 + \sum_{i=1}^{m} \beta_i \|Y_i\|^2 + \|X\|^2_{\gamma I - Q_0} \right)^{\frac{1}{2}},$$

where $\|X\|^2_{\gamma I - Q_0} = \mathbf{vec}\,(X)^T\,(\gamma I - Q_0)\;\mathbf{vec}\,(X) = \gamma \|X\|^2 - \langle X, Q_0(X) \rangle$. Observe that solving the optimal condition (2.3) for problem (2.2) is equivalent to finding a zero point of the residual function

$$\|e(W)\|_G := \left\| \begin{array}{c} X - P_{S^n_+}\left[X - \alpha \left(Q_0(X) + B_0 - \sum\limits_{i=1}^{m} \lambda_i \right) \right] \\ Y_i - P_{\Omega_i}\left[Y_i - \alpha_i \lambda_i \right] \\ \beta_i \left(X - Y_i \right) \end{array} \right\|_G .$$

Then we have from (2.8) that

$$\begin{aligned} X^{k+1} &= P_{S^n_+}\left[X^{k+1} - \alpha \left(Q_0(X^{k+1}) + B_0 - \sum_{i=1}^{m} \left(\lambda_i^k - \beta_i \left(X^{k+1} - Y_i^k \right) \right) - R(X^k, X^{k+1}) \right) \right] \\ &= P_{S^n_+}\left[X^{k+1} - \alpha \left(Q_0(X^{k+1}) + B_0 - \sum_{i=1}^{m} \lambda_i^{k+1} - \sum_{i=1}^{m} \beta_i \left(Y_i^k - Y_i^{k+1} \right) + R(X^k, X^{k+1}) \right) \right] \end{aligned}$$

which implies that

$$\left\| e(W^{k+1}) \right\|_G \;\le\; \underset{\text{(non-expansion of projection)}}{\left\| \begin{array}{c} \alpha \left(\sum\limits_{i=1}^{m} \beta_i \left(Y_i^{k+1} - Y_i^k \right) + R(X^k, X^{k+1}) \right) \\ 0 \\ \lambda_i^k - \lambda_i^{k+1} \end{array} \right\|_G} \le \delta \left\| W^k - W^{k+1} \right\|_G \tag{3.7}$$

where δ is a positive constant depending on α, β_is, and γ.

Note that (3.1) can be written as

$$\begin{aligned} & \left\langle W^{k+1} - W^*, W^k - W^{k+1} \right\rangle_G \ge 0 \\ \Rightarrow\; & \left\langle W^k - W^*, W^k - W^{k+1} \right\rangle_G \ge \left\| W^k - W^{k+1} \right\|_G^2 \end{aligned}$$

Thus

$$\begin{aligned} \left\| W^{k+1} - W^* \right\|_G^2 &= \left\| \left(W^k - W^* \right) - \left(W^k - W^{k+1} \right) \right\|_G^2 \\ &= \left\| W^k - W^* \right\|_G^2 - 2 \left\langle W^k - W^*, W^k - W^{k+1} \right\rangle_G + \left\| W^k - W^{k+1} \right\|_G^2 \\ &\le \left\| W^k - W^* \right\|_G^2 - \left\| W^k - W^{k+1} \right\|_G^2 \\ &\le \left\| W^k - W^* \right\|_G^2 - \frac{1}{\delta^2} \left\| e(W^{k+1}) \right\|_G^2 \end{aligned} \tag{3.8}$$

From the above inequality, we have

$$\left\|W^{k+1}-W^*\right\|_G^2 \leq \left\|W^k-W^*\right\|_G^2 \leq \cdots \leq \left\|W^0-W^*\right\|_G^2. \tag{3.9}$$

That is, the sequence $\{W^k\}$ is bounded. Thus there exists at least one cluster point of $\{W^k\}$.

It also follows from (3.8) that

$$\sum_{k=0}^{\infty} \frac{1}{\delta^2} \left\|e(W^{k+1})\right\|_G^2 < +\infty.$$

This implies that

$$\lim_{k\to\infty} \left\|e(W^k)\right\|_G = 0.$$

Let $\bar{W}$ be a cluster point of $\{W^k\}$, and let $\{W^{k_j}\}$ be a corresponding subsequence converging to $\bar{W}$. Therefore,

$$\left\|e(\bar{W})\right\|_G = \lim_{j\to\infty} \left\|e(W^{k_j})\right\|_G = 0,$$

which means that $\bar{W}$ is a zero point of the residual function. Therefore $\bar{W}$ satisfies (2.3). Setting $W^* = \bar{W}$ in (3.9), we have

$$\left\|W^{k+1}-\bar{W}\right\|_G^2 \leq \left\|W^k-\bar{W}\right\|_G^2, \ \forall k \geq 0.$$

Thus, the sequence $\{W^k\}$ has a unique cluster point and

$$\lim_{k\to\infty} W^k = \bar{W}.$$

This completes the proof. □

4. The Inexact Case

The modified alternating direction method requires to compute the projection onto semidefinite cone. However, there seems to be little justification of the effort required to do it exactly. In fact, inspired by Dembo et al. [5] and Pang [15], we could develop an inexact version of the modified alternating direction method.

The Inexact Version of the Modified Alternating Direction Method.

Given a nonnegative sequence $\{\eta_k\}$ satisfies $\sum\limits_{k=0}^{\infty} \eta_k^2 < +\infty$.

In each iteration, we solve (2.10) inexactly as follows

$$X^{k+1} = P_{S_+^n}\left[X^{k+1} - \alpha\left(Q_0(X^{k+1}) + B_0 - \sum_{i=1}^{m}\left(\lambda_i^k - \beta_i\left(X^{k+1}-Y_i^k\right)\right) - R(X^k, X^{k+1}) + \Theta_k(X^{k+1})\right)\right] \tag{4.1}$$

such that

$$\left\|\Theta_k(X^{k+1})\right\|_{(\gamma I - Q_0))^{-1}} \leq \eta_k \left\|e(W^{k+1})\right\|_G. \tag{4.2}$$

We next show that the sequence $\{X^k, Y_i^k, \lambda_i^k\}$ generated by the inexact version will satisfy a weak contractive property.

Proposition 4.1. *Let* $\{X^k, Y_i^k, \lambda_i^k\}$ *be the sequence generated by the inexact version of the modified alternating direction method. Then there is a* $k_0 \geq 0$*, such that*

$$\left\|W^{k+1} - W^*\right\|_G^2 \leq \left(1 + 4\delta^2\eta_k^2\right) \left\|W^k - W^*\right\|_G^2 - \frac{1}{4\delta^2} \left\|e(W^{k+1})\right\|_G^2, \ \forall k \geq k_0, \tag{4.3}$$

where δ *is the positive constant from inequality (3.7).*

Proof. Since in the inexact method

$$\Big\langle X - X^{k+1}, Q_0(X^{k+1}) + B_0 - \sum_{i=1}^m \lambda_i^{k+1} - \sum_{i=1}^m \beta_i \left(Y_i^k - Y_i^{k+1}\right)$$
$$-R(X^k, X^{k+1}) + \Theta_k(X^{k+1})\Big\rangle \geq 0, \quad \forall X \in S_+^n.$$

Similar to the proof of Theorem 1, by adding inequalities (3.2), (3.3) and (3.5) to the inequality above, we obtain

$$\sum_{i=1}^m \frac{1}{\beta_i} \left\langle \lambda_i^{k+1} - \lambda_i^*, \lambda_i^k - \lambda_i^{k+1}\right\rangle + \sum_{i=1}^m \beta_i \left\langle Y_i^{k+1} - Y_i^*, Y_i^k - Y_i^{k+1}\right\rangle$$
$$+ \ \left\langle X^{k+1} - X^*, R(X^k, X^{k+1})\right\rangle + \left\langle X^{k+1} - X^*, -\Theta_k(X^{k+1})\right\rangle \geq 0$$

According to the definitions of $\langle \cdot, \cdot \rangle_G$ and $\| \cdot \|_G$, the inequality above can be written as

$$\left\langle W^{k+1} - W^*, W^k - W^{k+1}\right\rangle_G + \left\langle X^{k+1} - X^*, -\Theta_k(X^{k+1})\right\rangle \geq 0,$$

which implies

$$\left\langle W^k - W^*, W^k - W^{k+1}\right\rangle_G \geq \left\|W^k - W^{k+1}\right\|_G^2 + \left\langle X^{k+1} - X^*, \Theta_k(X^{k+1})\right\rangle.$$

Thus,

$$\begin{aligned}
\left\|W^{k+1} - W^*\right\|_G^2 &= \left\|\left(W^k - W^*\right) - \left(W^k - W^{k+1}\right)\right\|_G^2 \\
&= \left\|W^k - W^*\right\|_G^2 - 2\left\langle W^k - W^*, W^k - W^{k+1}\right\rangle_G + \left\|W^k - W^{k+1}\right\|_G^2 \\
&\leq \left\|W^k - W^*\right\|_G^2 - \left\|W^k - W^{k+1}\right\|_G^2 - 2\left\langle X^{k+1} - X^*, \Theta_k(X^{k+1})\right\rangle \\
&= \left\|W^k - W^*\right\|_G^2 - \left\|W^k - W^{k+1}\right\|_G^2 - 2\left\langle X^{k+1} - X^k, \Theta_k(X^{k+1})\right\rangle \\
&\quad -2\left\langle X^k - X^*, \Theta_k(X^{k+1})\right\rangle.
\end{aligned} \tag{4.4}$$

Note that

$$\begin{aligned}
-2\left\langle X^k - X^*, \Theta_k(X^{k+1})\right\rangle &\leq 4\delta^2\eta_k^2 \left\|X^k - X^*\right\|_{\gamma I - Q_0}^2 + \frac{1}{4\delta^2\eta_k^2} \left\|\Theta_k(X^{k+1})\right\|_{(\gamma I - Q_0)^{-1}}^2 \\
&\leq 4\delta^2\eta_k^2 \left\|W^k - W^*\right\|_G^2 + \frac{1}{4\delta^2} \left\|e(W^{k+1})\right\|_G^2
\end{aligned} \tag{4.5}$$

Using (4.2) and $\sum_{k=0}^{\infty} \eta_k^2 < +\infty$, it is easy to show that there is a $k_0 \geq 0$ such that for all $k \geq k_0$

$$\begin{aligned}
-2\left\langle X^{k+1} - X^k, \Theta_k(X^{k+1})\right\rangle &\leq 4\delta^2\eta_k^2 \left\|X^{k+1} - X^k\right\|_{\gamma I - Q_0}^2 + \frac{1}{4\delta^2\eta_k^2} \left\|\Theta_k(X^{k+1})\right\|_{(\gamma I - Q_0)^{-1}}^2 \\
&\leq \frac{1}{4} \left\|W^{k+1} - W^k\right\|_G^2 + \frac{1}{4\delta^2} \left\|e(W^{k+1})\right\|_G^2.
\end{aligned} \tag{4.6}$$

Substituting (4.5) and (4.6) into inequality (4.4), we obtain

$$\begin{aligned}\left\|W^{k+1}-W^*\right\|_G^2 &\leq \left(1+4\delta^2\eta_k^2\right)\left\|W^k-W^*\right\|_G^2-\frac{3}{4}\left\|W^k-W^{k+1}\right\|_G^2+\frac{1}{2\delta^2}\left\|e(W^{k+1})\right\|_G^2\\ &\leq \left(1+4\delta^2\eta_k^2\right)\left\|W^k-W^*\right\|_G^2-\frac{3}{4\delta^2}\left\|e(W^{k+1})\right\|_G^2+\frac{1}{2\delta^2}\left\|e(W^{k+1})\right\|_G^2\\ &= \left(1+4\delta^2\eta_k^2\right)\left\|W^k-W^*\right\|_G^2-\frac{1}{4\delta^2}\left\|e(W^{k+1})\right\|_G^2.\end{aligned}$$

The proof is completed. □

The following theorem will prove the convergence of the inexact method.

Theorem 4.2. *The sequence $\{X^k\}$ generated by the inexact version of the modified alternating direction method for CQCQSDP converges to a solution point X^*.*

Proof. Note that under the assumption $\sum_{k=0}^{\infty}\eta_k^2<+\infty$ the product $\prod_{k=0}^{\infty}(1+4\delta^2\eta_k^2)$ is bounded. We denote

$$C_s\equiv 4\delta^2\sum_{k=0}^{\infty}\eta_k^2,\ \text{ and }\ C_p\equiv\prod_{k=0}^{\infty}\left(1+4\delta^2\eta_k^2\right).$$

Let $\hat{W}$ be a solution. From Proposition 4.1, we have for all $k\geq k_0$

$$\begin{aligned}\left\|W^k-\hat{W}\right\|_G^2 &\leq \left(\prod_{l=k_0}^{k}\left(1+4\delta^2\eta_l^2\right)\right)\left\|W^{k_0}-\hat{W}\right\|_G^2\\ &\leq C_p\left\|W^{k_0}-\hat{W}\right\|_G^2\end{aligned}$$

Therefore, there exists a constant $\tau>0$, such that

$$\left\|W^k-\hat{W}\right\|_G^2\leq\tau,\ \forall k\geq 0. \tag{4.7}$$

It follows that the sequence $\{W^k\}$ is bounded and therefore it has at least a cluster point. From (4.3) and (4.7), we get

$$\begin{aligned}\frac{1}{4\delta^2}\sum_{k=k_0}^{\infty}\left\|e(W^{k+1})\right\|_G^2 &\leq \left\|W^{k_0}-\hat{W}\right\|_G^2+4\delta^2\sum_{k=k_0}^{\infty}\left(\eta_k^2\left\|W^k-\hat{W}\right\|_G^2\right)\\ &\leq (1+C_s)\,\tau.\end{aligned}$$

It follows that

$$\lim_{k\to\infty}\left\|e(W^k)\right\|_G=0.$$

Let W^* be a cluster point of $\{W^k\}$. Suppose the subsequence $\{W^{k_j}\}$ converges to W^*. Then

$$\|e(W^*)\|_G=\lim_{j\to\infty}\left\|e(W^{k_j})\right\|_G=0$$

and W^* is a solution. Since W^* is a solution, for all $k\geq k_0$, $l\geq 0$, we have

$$\left\|W^{k+l}-W^*\right\|_G^2\leq C_p\left\|W^k-W^*\right\|_G^2. \tag{4.8}$$

It follows from (4.8) that the sequence $\{W^k\}$ has a unique cluster point and

$$\lim_{k\to\infty}W^k=W^*.$$

□

5. Concluding Remarks

We propose a modified alternating direction method for solving convex quadratically constrained quadratic semidefinite problems. The advantage of the proposed method is that it does not require to solve sub-variational inequality problems on semidefinite cone; instead, in each iteration, it requires only one projection onto semidefinite cone plus m easy vector projections. The convergence of the method is analyzed and it is shown that if the problem has an optimal solution at all, then the method will produce a sequence that converge to a solution. The method can be relaxed to allow inexact projection onto the semidefinite cone, while preserves the same convergence properties.

The proposed modified method does not require second order information and they are very easy to implement. For problems with a large number of quadratic constraints, the vector projections in Step 2 are readily parallelizable. These additional features appear to be attractive in solving large-scale convex quadratically constrained quadratic semidefinite programs.

References

[1] A. Beck, Quadratic matrix programming. *SIAM J. Optim.* 17, 1224-1238 (2007).

[2] S. Boyd and L. Xiao, Least-squares covariance matrix adjustment. *SIAM J. Matrix Anal. Appl.* 27, 532-546 (2005).

[3] G. Chen and M. Teboulle, A proximal-based decomposition method for convex minimization problems. *Math. Program.* 64, 81-101 (1994).

[4] Y.H. Dai, Fast algorithms for projection on an ellipsoid. *SIAM J. Optim.* 16, 986-1006 (2006).

[5] R.S. Dembo, S.C. Eisenstat, and T. Steihaug, The linear complementary problem. Academic Press, New York, 1982.

[6] J. Eckstein and D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.* 55, 293-318 (1992).

[7] D. Gabay, Applications of the method of multipliers to variational inequalities. In: Fortin, M., Glowinski, R., eds., *Augmented Lagrangian methods: Applications to the numerical solution of Boundary-Value Problem,* pp. 299-331. North-Holland, Amsterdam, The Netherlands, 1983.

[8] D. Gabay and B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximations. *Compu. Math. Appl.* 2, 17-40 (1976).

[9] Y. Gao and D. Sun, Calibrating least squares covariance matrix problems with equality and inequality constraints. *Working paper,* National University of Singapore, 2008.

[10] B.S. He, L.Z. Liao, D.R. Han, and H. Yang, A new inexact alternating directions method for monotone variational inequalities. *Math. Program.* 92, 103-118 (2002).

[11] S. Kontogiorgis and R.R. Meyer, A variable-penalty alternating directions method for convex optimization. *Math. Program.* 83, 29-53 (1998).

[12] C. Li and W. Sun, On filter-successive linearization methods for nonlinear semidefinite programming, to appear in *China Science, Ser. A* (2007).

[13] J. Malick, A dual approach to semidefinite least-squares problems. *SIAM J. Matrix Anal. Appl.* 26, 272-284 (2004).

[14] J.W. Nie and Y.X. Yuan, A predictor- corrector algorithm for QSDP combining Dikin-type and Newton centering steps. *Ann. Oper. Res.* 103, 115-133 (2001).

[15] J.-S. Pang, Inexact methods for the nonlinear complementarity problem. *Math. Program.* 36, 54-71 (1986).

[16] H. Qi and D. Sun, A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM J. Matrix Anal. Appl.* 28, 360-385 (2006).

[17] D. Sun, J. Sun, and L. Zhang, Rates of convergence of the augmented Lagrangian method for nonlinear semidefinite programming. *Math. Program.* 114, 349-391 (2008).

[18] D. Sun, J. Sun, and L. Qi, A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems. *SIAM J. Optim.* 14, 783-806 (2004).

[19] K. Toh, An inexact primal-dual path following algorithm for convex quadratic SDP. *Math. Program.* 112, 221-254 (2007).

[20] Z.S. Yu, Solving semidefinite programming problems via alternating direction methods. *J. Compu. Appl. Math.* 193, 437-445 (2006).