PARAMETER-RELATED PROJECTION-BASED ITERATIVE ALGORITHM FOR A KIND OF GENERALIZED POSITIVE SEMIDEFINITE LEAST SQUARES PROBLEM

CHENGJIN LI

College of Mathematics and Informatics
Fujian Normal University
Fuzhou, 350007, P. R. China

Abstract. A projection-based iterative algorithm, which is related to a single parameter (or the multiple parameters), is proposed to solve the generalized positive semidefinite least squares problem introduced in this paper. The single parameter (or the multiple parameters) projection-based iterative algorithms converges to the optimal solution under certain condition, and the corresponding numerical results are shown too.

1. Introduction. Let $\|\cdot\|_2$, $\|\cdot\|_F$ be the $2-$norm of vector and the $F-$norm of matrix, and we use $S^r_+, S^r$ and $S^r_s$ to denote the $r-$dimension square matrix set consisting of the symmetric positive semidefinite matrix, the symmetric matrix and the anti-symmetric matrix, respectively. The generalized positive semidefinite least squares problem we are concerned in this paper is

$$
\min_{X \in \mathbb{R}^{n \times n}} (1/2)\|G X H - F\|_F^2 \quad \text{s.t.} \quad X \in \mathcal{K} := \mathcal{K}_1 + \mathcal{K}_2,
$$

where $\mathcal{K}_1 = Q_r S^r_+ Q_r^\top$ with $r \in \{1, 2, \cdots, n\}$, $\mathcal{K}_2$ is any subspace of $\mathbb{R}^{n \times n}$ satisfying $\mathcal{K}_1 \perp \mathcal{K}_2$, and $G \in \mathbb{R}^{m \times n}$, $H \in \mathbb{R}^{n \times p}$, $F \in \mathbb{R}^{m \times p}$, $Q_r \in \mathbb{R}^{n \times r}$ are all given parameter matrices with $G$ full column rank, $H$ full row rank, $Q = [Q_r, \bar{Q}_r] \in \mathbb{R}^{n \times n}$ being an orthonormal matrix.

As described in [7], the symmetric positive semidefinite cone, the nonsymmetric positive semidefinite cone and the conditional symmetric positive semidefinite cone are all special cases of $\mathcal{K}$. Hence, Problem (1) can be found in a wide range of application fields, and you can refer the papers [1, 14, 17, 4, 6, 15, 12] for the more details. Moreover, there are several methods can be used to solve some special cases of Problem (1), such as the gradient projection method [5], the proximal point-type method [2, 7], the predictor-corrector algorithm [1], the interior-point method [16, 6], the GSVD method [9] and the semi-proximal ADMM [8], the Newton-type method [13], the semismooth newton-CG method [11].

2020 Mathematics Subject Classification. Primary: 65K05; Secondary: 90C30.

Key words and phrases. Projection, iterative algorithm, positive semidefinite, least squares problem.

This work was supported by the National Natural Science Foundations of China (11301080, 11526053), Science Foundation of Fujian Province of China (2016J05003), the Foundation of the Education Department of Fujian Province of China (JA15106), and the Project of Nonlinear analysis and its applications (IRT1206).
then the parameter matrices $Q$ to Problem (1) is constructed. An equivalent problem of Problem (1) is:

\[ \text{minimize} \quad \frac{1}{2} \| \tilde{X} \|_F^2, \quad \text{s.t.} \quad X \in K := K_1 + K_2, \]

where $\alpha > 0$. It is clear that Problem (3) has the same solution with Problem (1) and the first-order derivative of $\Phi_\alpha(X)$ is $\alpha G^\top (GXH - F)H^\top$. Therefore, the first-order optimal condition (or KKT condition) of Problem (3) is reformulated as [3]:

\[
\begin{align*}
\{ & \alpha G^\top (GXH - F)H^\top - Y = 0_{n \times n}, \\
& K \ni X \perp Y \in K^*,
\end{align*}
\]

where $\langle \cdot, \cdot \rangle$ and $K^*$ denote respectively inner product and the adjoint cone of $K = \{ Y \in R^{n \times n} | \langle Y, Z \rangle \geq 0 \text{ for all } Z \in K \}$. Let $K_3$ is the smallest subspace in $R^{n \times n}$ which containing $K_1$, which yields

\[
K_3 = \left( \begin{array}{cc} S^r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{array} \right),
\]

and

\[
K_3^\perp = \left( \begin{array}{c} 0_{r \times r} \\ R^{(n-r) \times r} \end{array} \right) \oplus S_n^{n-r},
\]

where $K_3^\perp$ is the orthogonal complement subspace of $K_3$ in $R^{n \times n}$. Clearly, the relation

\[
K_2 \subseteq K_3^\perp
\]

holds. With the help of (6), the formula of $K^*$ can be shown in the following lemma.

**Lemma 2.1.**

\[
K^* = \left( \begin{array}{cc} S_+^r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{array} \right) + K_{2,3}^\perp.
\]
where $\tilde{K}_{2,3}$ is the orthogonal complement subspace of the subspace $K_2$ in $K_{3}$. 

Proof. The only thing we need to prove is that the both sides of (7) are included in each other. According to the definition of $K$, the element in $K$ is formulated as

$$X = \begin{pmatrix} X_1 & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{pmatrix} + X_2,$$

where $X_1 \in S_r^+$ and $X_2 \in K_2$. The orthogonality between $K_1$ and $K_2$ together with the formula (2) deduces that

$$X_2^{11} \in S_r^+,$$

where $X_2^{11}$ is the $r \times r$ principal submatrix of $X_2$.

For each $Y \in K^r \subseteq \mathbb{R}^{n \times n}$, we suppose

$$Y = \begin{pmatrix} Y_{11} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{pmatrix} + \begin{pmatrix} 0_{r \times r} & Y_{21}^T \\ Y_{21} & Y_{22} \end{pmatrix} + Y_2,$$

with $Y_{11} \in S_r, Y_{21} \in \mathcal{R}^{(n-r) \times r}, Y_{22} \in S^{n-r}$ and $Y_2 \in S_r^n$. The definition of $K^r$ together with (8) and (10) yields

$$0 \leq \langle Y, X \rangle$$

$$= \langle \left( \begin{pmatrix} X_1 & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{pmatrix} + X_2, \right) \begin{pmatrix} Y_{11} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{pmatrix} + \begin{pmatrix} 0_{r \times r} & Y_{21}^T \\ Y_{21} & Y_{22} \end{pmatrix} + Y_2 \rangle$$

$$= \langle Y_{11}, X_1 \rangle + \langle \begin{pmatrix} 0_{r \times r} & Y_{21}^T \\ Y_{21} & Y_{22} \end{pmatrix} + Y_2, X_2 \rangle$$

(11)

hold for all $X_1 \in S_r^+$ and $X_2 \in K_2$, where the second equation come from (9). And therefore, we have

$$Y_{11} \in S_r^+$$

(12)

by considering the $X_2$ in (11) substituted by $0_{n \times n}$. Similarly, if $X_1$ in (11) is substituted by $0_{r \times r}$, the following relation

$$\begin{pmatrix} 0_{r \times r} & Y_{21}^T \\ Y_{21} & Y_{22} \end{pmatrix} + Y_2 \in K_{2}$$

(13)

holds from the fact that $K_2$ is a subspace of $K_{3}$. It follows from (13) and (5) that

$$\begin{pmatrix} 0_{r \times r} & Y_{21}^T \\ Y_{21} & Y_{22} \end{pmatrix} + Y_2 \in \tilde{K}_{2,3}$$

which together with (12) deduces

$$Y \in \left( \begin{pmatrix} Y_{11} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{pmatrix} + \tilde{K}_{2,3}, S_r^+ \begin{pmatrix} 0_{r \times (n-r)} \\ 0_{(n-r) \times r} \end{pmatrix} \begin{pmatrix} 0_{r \times (n-r)} \\ 0_{(n-r) \times (n-r)} \end{pmatrix} \right) + \tilde{K}_{2,3}.$$

Conversely, for each

$$Y = \begin{pmatrix} Y_{11} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{pmatrix} + \tilde{Y}_2 \in \left( S_r^+ \begin{pmatrix} 0_{r \times (n-r)} \\ 0_{(n-r) \times r} \end{pmatrix} \begin{pmatrix} 0_{r \times (n-r)} \\ 0_{(n-r) \times (n-r)} \end{pmatrix} \right) + \tilde{K}_{2,3},$$

we have

$$\tilde{Y}_2 = \begin{pmatrix} 0_{r \times r} & Y_{21}^T \\ Y_{21} & Y_{22} \end{pmatrix} + Y_2$$
with \( Y_{11} \in S^r, Y_{21} \in \mathcal{R}^{(n-r) \times r}, Y_{22} \in S^{n-r} \), \( Y_2 \in S^n_+ \) by considering \( \tilde{Y}_2 \in \mathcal{K}^{2,3} \subseteq \mathcal{K}^+ \) and (5). It is deduced from (11) that the inequality \(<X, Y> \geq 0\) holds for all \( X \in \mathcal{K} \), which implies \( Y \in \mathcal{K}^* \).

From Lemma 2.1, we know that \( \mathcal{K} \neq \mathcal{K}^* \) in generally, which implies the properties related to the self-adjoint cone cannot be directly used for \( \mathcal{K} \). But a useful property of the self-adjoint cone still holds for \( \mathcal{K} \), and the corresponding proof is shown in the next lemma.

**Lemma 2.2.** The second condition in (4) is equivalent to
\[
X = \Pi_{\mathcal{K}}(X - Y),
\]
where \( \Pi \) denotes the projection operator.

**Proof.** For simplicity, we suppose \( X = X_1 + X_2, Y = Y_1 + Y_2 \) with \( X_1, Y_1 \in \mathcal{K}_3 \) and \( X_2, Y_2 \in \mathcal{K}^+_2 \).

On one hand, it is easily deduced from the definition of \( \mathcal{K}^+_3 \) and (6) that
\[
\Pi_{\mathcal{K}}(X - Y) = \arg\min_{Q \in \mathcal{K}} \| X - Y - Q \|_F^2 \\
= \arg\min_{Q \in \mathcal{K}_1} \| X_1 - Y_1 - Q_1 \|_F^2 + \arg\min_{Q_2 \in \mathcal{K}_2} \| X_2 - Y_2 - Q_2 \|_F^2 \\
= \Pi_{\mathcal{K}_1}(X_1 - Y_1) + \Pi_{\mathcal{K}_2}(X_2 - Y_2),
\]
which implies the condition (14) is equivalent to
\[
X_1 = \Pi_{\mathcal{K}_1}(X_1 - Y_1) \quad \text{and} \quad X_2 = \Pi_{\mathcal{K}_2}(X_2 - Y_2).
\]

On the other hand, the formula of \( \mathcal{K} \) in (1) together with (7) deduces that the second condition in (4) is equivalent to
\[
X_1, Y_1 \in \mathcal{K}_1, \quad X_2 \in \mathcal{K}_2, \quad Y_2 \in \mathcal{K}^+_2, \quad <X_1, Y_1> = 0.
\]
By considering the self-adjointness of \( S^+_n \) and the formula of \( \mathcal{K}_1 \), we know[3]
\[
X_1, Y_1 \in \mathcal{K}_1, \quad <X_1, Y_1> = 0 \iff X_1 = \Pi_{\mathcal{K}_1}(X_1 - Y_1)
\]
holds for all \( X_1, Y_1 \in \mathcal{K}_3 \). Hence, the equivalence of (15) and (16) is dependent on whether the relation
\[
X_2 = \Pi_{\mathcal{K}_2}(X_2 - Y_2) \iff X_2 \in \mathcal{K}_2, \quad Y_2 \in \mathcal{K}^+_2
\]
holds for all \( X_2, Y_2 \in \mathcal{K}^+_2 \), and it is obviously.

From Lemma 2.2, we know the KKT condition (4) can be reformulated as
\[
X = \Pi_{\mathcal{K}}[X - \alpha G^T(GXH - F)H^T],
\]
which together with the equivalence between Problem (1) and Problem (3) implies that the unique optimal solution of the strictly convex optimization problem—Problem (1) is exactly the solution of (17).

Since the projection operator \( \Pi_{\mathcal{K}}(\cdot) \) is nonexpansive, the following iteration process
\[
X_{k+1} = \Pi_{\mathcal{K}}[X_k - \alpha G^T(GX_kH - F)H^T] \\
= \Pi_{\mathcal{K}}[(X_k - \alpha G^T GX_kHH^T) + \alpha G^TFH^T]
\]
is convergent if the positive parameter \( \alpha \) can be chosen properly. Furthermore, the fixed point of the iteration process (18) is the solution of (17) if the iterative sequence \( \{X_k\} \) generated by (18) converges.
In order to determine a proper value of $\alpha$, we transform the iterative point $X_k$ into $x_k$, where the $x_k$ is the column vector obtained by stacking the columns of $X_k$ on top of each other. Obviously, the following relations hold for all $k$:

\[
\|x_{k+1} - x_k\|_2 = \|X_{k+1} - X_k\|_F
\]
\[
= \left\| \Pi_k \left[ (X_k - \alpha G^T G X_k H H^T) + \alpha G^T F H^T \right] - \Pi_k \left[ (X_{k-1} - \alpha G^T G X_{k-1} H H^T) + \alpha G^T F H^T \right] \right\|_F
\]
\[
\leq \| (X_k - X_{k-1}) - \alpha G^T G (X_k - X_{k-1}) H H^T \|_F
\]
\[
= \left\| I_{n^2} - \alpha (H H^T) \otimes (G^T G) \right\| x_k - x_{k-1} \|_2
\]
\[
\leq \sigma \| x_k - x_{k-1} \|_2,
\]
where $'\otimes'$ denotes Kronecker product and $\sigma \geq 0$ is the largest singular value of $I_{n^2} - \alpha (H H^T) \otimes (G^T G)$. It is clear that the iterative sequence $\{X_k\}$ converges to the fixed point of (18)/(or the solution of (17)) if $\sigma < 1$.

For the simplicity of the following discussion, the notations $\lambda_{\min}^G(\lambda_{\max}^G)$ and $\lambda_{\min}^H(\lambda_{\max}^H)$ are used to denote the least eigenvalue (the largest eigenvalue) of $G^T G$ and $H H^T$, respectively. It follows from the property of Kronecker product that $\lambda_{\min}^G \lambda_{\min}^H$ and $\lambda_{\max}^G \lambda_{\max}^H$. By considering the full-rank-assumption mentioned in Section 1, we know $G^T G$, $H H^T$ are both symmetric positive definite, so is $(H H^T) \otimes (G^T G)$, which implies $\sigma_{\min} > 0$ and $\sigma_{\max} > 0$.

It is easy to deduce that $\sigma = \max\{|1 - \alpha \sigma_{\min}|, |1 - \alpha \sigma_{\max}|\} \geq 0$, which together with the fact $\sigma_{\max} > \sigma_{\min} > 0$ implies $\alpha \in (0, 2/\sigma_{\max}) \Rightarrow \sigma < 1$ hold. With the help of the above discussion, we claim the main conclusion in this section.

**Theorem 2.3.** The iteration process (18) converges if the corresponding parameter $\alpha \in (0, 2/\sigma_{\max})$.

In order to improve the convergence rate of (18), the parameter $\alpha$ should be chosen such that the non-negative number $\sigma$ as small as possible. Specifically, $\alpha \in (1/\sigma_{\max}, 2/\sigma_{\max})$ is a good choice, and we define

\[
\alpha = \begin{cases} \frac{2}{\sigma_{\max}^2 + \sigma_{\max}}, & \text{if } \frac{\sigma_{\min}}{\sigma_{\max}} \leq \frac{1}{2}, \\ \frac{1}{\sigma_{\max}^2}, & \text{otherwise}. \end{cases}
\]

**Algorithm 2.4.** Single parameter projection-based iterative algorithm

(S.0) The parameter $\alpha$ is defined by (20), and set the initial iteration point $\tilde{X} = 0_{n \times n}$ and the tolerance $\varepsilon$.

(S.1) The new iteration point $X$ is updated by

\[
X = \Pi_k [\tilde{X} - \alpha G^T (G \tilde{X} H - F) H^T],
\]

and go to (S.2).

(S.2) If $\|X - \tilde{X}\|_F/\|X\|_F < \varepsilon$, stop; Otherwise update $\tilde{X} = X$, go to (S.1).

3. **Multiple parameters projection-based iterative algorithm.** From what was discussed in Section 2, we know the convergent rate of Algorithm 2.4 is dependent on the parameter $\alpha$. In order to improve the efficiency of the projection-based
iterative algorithm, two sequences of functions, which have the same minimum point as \( \Phi_\alpha(X) \), are constructed in this section:

\[
\Phi_\alpha(X) := (\alpha^2_k/2)\|GXH - F\|^2_F \quad \text{and} \quad \Phi_\beta(X) := (\beta^2_k/2)\|GXH - F\|^2_F,
\]

(21)

where \( \alpha_k \neq 0 \neq \beta_k \) for all \( k \). And, for each \( k \geq 1 \), we suppose the feasible iteration points \( X_{k-1} \) and \( X_k \) have already obtained, then it is easy to see the following inequalities

\[
\begin{align*}
\Phi_\alpha(X) & \leq \Phi_\alpha(X_{k-1}) + \alpha^2_k \langle G^\top (GX_{k-1}H - F)H^\top, X - X_{k-1} \rangle + \frac{\alpha^2_k \sigma_{\max}}{2} \|X - X_{k-1}\|^2_F, \\
\Phi_\beta(X) & \geq \Phi_\beta(X_k) + \beta^2_k \langle G^\top (GX_kH - F)H^\top, X - X_k \rangle + \frac{\beta^2_k \sigma_{\min}}{2} \|X - X_k\|^2_F.
\end{align*}
\]

(22)

hold for all \( X \in \mathbb{R}^{n \times n} \). The inequalities in (22) are deduced from Tayler expression and the fact that \( \sigma_{\min}\|X\|^2_F \leq \|G^\top GXH^\top\|^2_F \leq \sigma_{\max}\|X\|^2_F \).

With the help of (22), a series of new objective functions are defined as follows:

\[
\begin{align*}
\Phi_k(X) & = \lambda_k \left[ \Phi_\alpha(X_{k-1}) + \alpha^2_k \langle G^\top (GX_{k-1}H - F)H^\top, X - X_{k-1} \rangle + \frac{\alpha^2_k \sigma_{\max}}{2} \|X - X_{k-1}\|^2_F \right] + (1 - \lambda_k) \left[ \Phi_\beta(X_k) + \beta^2_k \langle G^\top (GX_kH - F)H^\top, X - X_k \rangle + \frac{\beta^2_k \sigma_{\min}}{2} \|X - X_k\|^2_F \right],
\end{align*}
\]

(23)

where \( \lambda_k \in \Theta_k := \{ t \in \mathbb{R} \mid t \alpha^2_k + (1 - t)\beta^2_k > 0 \} \). The definition of \( \Theta_k \) together with \( \alpha_k \neq 0, \beta_k \neq 0 \) implies \( (0, 1) \subseteq \Theta_k \) and the value \( \lambda_k \alpha^2_k + (1 - \lambda_k)\beta^2_k \) is positive, where \( \lambda_k \alpha^2_k + (1 - \lambda_k)\beta^2_k \) is the parameter of the function \( \lambda_k \Phi_\alpha(X) + (1 - \lambda_k)\Phi_\beta(X) \). It is easy to see from (22) and (23) that \( \Phi_k(X) \) can be regarded as an approximating function of the convex function \( \lambda_k \Phi_\alpha(X) + (1 - \lambda_k)\Phi_\beta(X) \).

The new iteration point of the multiple parameters projection-based iterative algorithm is determined by the solution of the KKT condition of the following approximating problem

\[
\min_{X \in \mathbb{K}} \Phi_k(X).
\]

(24)

Obviously, the \( F \)-derivative of \( \Phi_k(X) \) is

\[
G^\top G[\lambda_k \alpha^2_k X_{k-1} + (1 - \lambda_k)\beta^2_k X_k]HH^\top - [\lambda_k \alpha^2_k + (1 - \lambda_k)\beta^2_k]G^\top FH^\top + [\lambda_k \alpha^2_k \sigma_{\max} + (1 - \lambda_k)\beta^2_k \sigma_{\min}]X - [\lambda_k \alpha^2_k \sigma_{\max} X_{k-1} + (1 - \lambda_k)\beta^2_k \sigma_{\min} X_k],
\]

which together with (17) deduce that the KKT condition of (24) is

\[
X = \Pi_{\mathcal{K}} \{ X \in \mathbb{H} \mid G[\lambda_k \alpha^2_k X_{k-1} + (1 - \lambda_k)\beta^2_k X_k]HH^\top + [\lambda_k \alpha^2_k + (1 - \lambda_k)\beta^2_k]G^\top FH^\top - [\lambda_k \alpha^2_k \sigma_{\max} + (1 - \lambda_k)\beta^2_k \sigma_{\min}]X + [\lambda_k \alpha^2_k \sigma_{\max} X_{k-1} + (1 - \lambda_k)\beta^2_k \sigma_{\min} X_k] \}.
\]

(25)
After substituting $X$ in the left and right of equation (25) by $X_{k+1}$ and $X_k$ respectively, the new iterative process can be formulated as

$$X_{k+1} = \Pi_K \left\{ \lambda_k \alpha_k^2 \sigma_{max} X_{k-1} + (1 - \lambda_k \alpha_k^2 \sigma_{max}) X_k \right\} -$$

$$\left[ \lambda_k \alpha_k^2 + (1 - \lambda_k) \beta_k^2 \right] G^\top \left[ G \left( \frac{\lambda_k \alpha_k^2}{\lambda_k \alpha_k^2 + (1 - \lambda_k) \beta_k^2} X_{k-1} + \frac{(1 - \lambda_k) \beta_k^2}{\lambda_k \alpha_k^2 + (1 - \lambda_k) \beta_k^2} X_k \right) H + F \right] H^\top \right\}. \quad (26)$$

Similar to (18), two linear combinations of $X_{k-1}$ and $X_k$ in (26) should be equivalent, and we use $Y_k$ to denote these linear combinations, i.e.,

$$Y_k := \lambda_k \alpha_k^2 \sigma_{max} X_{k-1} + (1 - \lambda_k \alpha_k^2 \sigma_{max}) X_k \quad (27)$$

$$= \frac{\lambda_k \alpha_k^2}{\lambda_k \alpha_k^2 + (1 - \lambda_k) \beta_k^2} X_{k-1} + \frac{(1 - \lambda_k) \beta_k^2}{\lambda_k \alpha_k^2 + (1 - \lambda_k) \beta_k^2} X_k. \quad \text{(29)}$$

Clearly, the second equality in (27) holds if

$$\lambda_k \alpha_k^2 \sigma_{max} = \frac{\lambda_k \alpha_k^2}{\lambda_k \alpha_k^2 + (1 - \lambda_k) \beta_k^2},$$

or

$$\lambda_k \alpha_k^2 + (1 - \lambda_k) \beta_k^2 = \frac{1}{\sigma_{max}}, \quad (28)$$

which together with $\sigma_{max} > 0$ implies $\lambda_k \in \Theta_k$. Hence, the iterative process (26) can be reformulated as

$$X_{k+1} = \Pi_K \left[ Y_k - \frac{1}{\sigma_{max}} G^\top (GY_k H + F) H^\top \right] \quad (29)$$

with $\lambda_k \in \Theta_k$ if (28) holds. Based on what was discussed above, the multiple parameters projection-based iterative algorithm is constructed as follows.

**Algorithm 3.1. Multiple parameters projection-based iterative algorithm**

(S.0) Set $X_0 \in K$, $\alpha_0 > 0$, $k = 0$, $\beta_0$, $\lambda_0$ are determined by $\alpha_0$ under the condition (28) with two additional conditions.

(S.1) Update $Y_k$ and $X_{k+1}$ by using (27) and (29), respectively.

(S.2) $\alpha_{k+1}$ is updated by $\beta_k$ in some way, and $\beta_{k+1}$ are determined by $\alpha_{k+1}$ under the condition (28) with two additional conditions.

(S.3) Stop if $\|X_k - X_{k+1}\|_F / \|X_{k+1}\|_F < \varepsilon$; Otherwise, $k = k + 1$ and go to (S.1).

The different choices of the parameters sequence $\{\alpha_k\}$, $\{\beta_k\}$, $\{\lambda_k\}$ imply the different multiple parameters projection-based iterative algorithms, so the convergence of Algorithm 3.1 is dependent on the additional conditions defined in (S.0).

As an example of the suitable addition conditions, we assume that

$$\sqrt{\sigma_{min}} \alpha_k - \lambda_k \alpha_k \sigma_{max} (\sqrt{\sigma_{min}} \alpha_k^2 + \beta_k) = 1 \quad (30)$$

and

$$\lambda_k = 1 - \frac{1}{\sqrt{\sigma_{min}} \beta_k} \quad (31)$$

hold for all $k$. 
On one hand, substituting (31) in (28) yields
\[
\left(1 - \frac{1}{\sigma_{\min}}\right)\left(\sigma_{\min}\alpha_k\right)^2 + \sqrt{\sigma_{\min}\beta_k} = \frac{\sigma_{\min}}{\sigma_{\max}},
\]
which implies
\[
\beta_k = \frac{1}{2\sigma_{\min}}\left(\frac{\sigma_{\min}}{\sigma_{\max}} - (\sigma_{\min}\alpha_k)^2 + \sqrt{\left(\frac{\sigma_{\min}}{\sigma_{\max}} - (\sigma_{\min}\alpha_k)^2 + 4(\sigma_{\min}\alpha_k)^2\right)}\right).
\]
On the other hand, it is transformed from (30) that
\[
-\lambda_k\alpha_k^2\sigma_{\max} = \sqrt{\sigma_{\min}\alpha_k}(1 - \sqrt{\sigma_{\min}\alpha_k})
\]
which together with (27) deduce
\[
Y = X_k + \frac{\sigma_{\min}\alpha_k(1 - \sqrt{\sigma_{\min}\alpha_k})}{\sigma_{\min}\alpha_k^2 + \sqrt{\sigma_{\min}\beta_k}}(X_k - X_{k-1}).
\]

Based on (33) and (35), Algorithm 3.1 can be rewritten as follows.

**Algorithm 3.2.** A special case of Algorithm 3.1

(S.0) Set the initial parameter $\alpha$, and the parameter $\beta$ is defined by (33). Set again the initial iteration point $X$ with the tolerance $\varepsilon$, and let $Y = X$.

(S.1) Update $\tilde{X} = X$, and compute
\[
X = \Pi_K[Y - (1/\sigma_{\max})G^T(GY - F)]H^T], \quad \alpha = \beta,
\]
\[
\beta = \frac{1}{2\sigma_{\min}}\left(\frac{\sigma_{\min}}{\sigma_{\max}} - (\sigma_{\min}\alpha_k)^2 + \sqrt{\left(\frac{\sigma_{\min}}{\sigma_{\max}} - (\sigma_{\min}\alpha_k)^2 + 4(\sigma_{\min}\alpha_k)^2\right)}\right),
\]
\[
Y = X + \frac{\sigma_{\min}\alpha_k(1 - \sqrt{\sigma_{\min}\alpha_k})}{\sigma_{\min}\alpha_k^2 + \sqrt{\sigma_{\min}\beta_k}}(X - \tilde{X});
\]

Go to (S.2).

(S.2) If $\|X - \tilde{X}\|_F/\|X\|_F < \varepsilon$, stop; Otherwise, go to (S.1).

For the convergence theorem of Algorithm 3.2, you can refer Theorem 2.2.3 in [10], because Algorithm 3.2 become the gradient projection method in [5] (or [10]) after defining the new iteration point sequence $\{u_k\}$ by
\[
u_k = \sqrt{\sigma_{\min}\alpha_k} \quad \text{and} \quad u_{k+1} = \sqrt{\sigma_{\min}\beta_k}
\]
for each $k$.

4. **Numerical results.** In this section, Algorithm 2.4 and Algorithm 3.2 are both performed on matlab R2013b, and the corresponding numerical results reported later are obtained from a PC with 3.82G memory and Intel(R) Core(TM) i5-3337U 1.8GHz CPU running on win64-bit Windows 8.

The notations ‘tim1’ and ‘tim2’ in this section are used to denote the time (second) spent by Algorithm 2.4 and Algorithm 3.2, respectively. We choose the initial iteration point $X = 0_{n \times n}$, the tolerance $\varepsilon = 1e-8$ and the initial parameter $\alpha = 1/(2\sqrt{\sigma_{\min}})$ for Algorithm 3.2. Both algorithms are stopped when the stopping condition $\|X - \tilde{X}\|_F/\|X\|_F < \varepsilon$ hold or the number of iterations is more than $1e8$. The ‘type’-column in the following table is used to distinguish the different kind of feasible region $K$. More specifically, by T1, T2 and T3, we denote the case of $K$.
being $S_n^+$, the nonsymmetric positive semidefinite cone, the conditional symmetric positive semidefinite cone, respectively.

The parameter matrices $G$, $H$ and $F$ are all generated randomly, i.e., $num = \text{randi}(100, [1, 3])$, $m = 3 + \max(num)$, $n = 3 + \min(num)$, $p = 3 + \text{median}(num)$, $r = \text{randi}(n - 1, 1)$, $H = \text{randn}(n, p)$, $G = \text{randn}(m, n)$ and $F = \text{randn}(m, p)$. For each fix parameters $m, n, p, r$, only the average spending time of the 10 running times of these two algorithms for solving Problem (1) with different $G, H, F$ are shown in the following table.

Comparing of Algorithm 2.4 and Algorithm 3.2

| Parameters $(m, n, p, (r))$ | Spending time type | Spending time | Parameters $(m, n, p, (r))$ | Spending time type |
|-----------------------------|------------------|--------------|-----------------------------|------------------|
| T1                          | 62.1             | 15.6         | T1                          | 0.017            | 0.016         |
| (53,39,48)                  | 2.13             | 1.95         | (103,61,85)                 | 1.82             | 1.09          |
| (101,33,61)                 | 0.25             | 0.19         | (102,78,90)                 | 8.95             | 5.61          |
| (102,19,95)                 | 0.040            | 0.039        | (93,20,92)                  | 0.041            | 0.038         |
| (66,48,53)                  | 10.1             | 3.28         | (58,18,45)                  | 0.05             | 0.05          |
| (62,43,68)                  | 0.009            | 0.008        | (49,41,42)                  | T2              | 115.6         | 8.67          |
| (71,11,24)                  | 0.04             | 0.02         | (94,44,92)                  | 0.94             | 0.28          |
| (92,43,64)                  | 1.99             | 0.45         | (98,19,93)                  | 0.05             | 0.03          |
| (93,7,48)                   | 0.014            | 0.012        | (73,4,30)                   | T2              | 0.008         | 0.007         |
| (68,64,72)                  | 32.9             | 3.99         | (78,39,62)                  | 1.54             | 0.34          |
| (94,16,85,(6))              | 0.03             | 0.02         | (91,43,83,(21))             | T3              | 0.90          | 0.20          |
| (60,36,53,(25))             | 1.18             | 0.28         | (54,11,24,(9))              | 0.04             | 0.02          |
| (31,8,10,(7))               | 0.12             | 0.03         | (86,9,83,(1))               | 0.019            | 0.015         |
| (78,55,69,(7))              | 20.7             | 0.77         | (97,17,33,(15))             | 0.08             | 0.04          |
| (98,45,51,(3))              | 10.3             | 0.44         | (42,10,36,(9))              | 0.023            | 0.019         |

From the data shown in above table, we know the multiple parameters projection-based iterative algorithm with the proper parameters is more efficient than the single parameter projection-based iterative algorithm.

5. Conclusions. A kind of parameter-related projection-based iterative method for the generalized positive semidefinite least squares problem is introduced in this work, and the efficiency of these algorithm are verified by the corresponding numerical results in Section 4. For the multiple parameters projection-based iterative algorithm, there are many different choices of the parameters $\alpha_k, \beta_k$ and $\lambda_k$. Hence, for the high convergent rate, how to construct the suitable additional conditions in Algorithm 3.1 is worth further researching.

REFERENCES

[1] B. Akesson, J. Jorgensen, N. Poulsen and S. Jorgensen, A generalized autocovariance least-squares method for Kalman filter tuning, Journal of Process Control, 18 (2008), 769–779.
[2] J. C. Allwright, Positive semidefinite matrices: Characterization via conical hulls and least-squares solution of a matrix equation, SIAM Journal on Control and Optimization, 26 (1988), 537–556.
[3] Z. X. Chan and D. F. Sun, Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming, SIAM Journal on Optimization, 19 (2008), 370–396.
[4] H. Dai and P. Lancaster, Linear matrix equations from an inverse problem of vibration theory, Linear Algebra and Its Applications, 246 (1996), 31–47.
520  CHENGJIN LI

[5] N. Gillis and P. Sharma, A semi-analytical approach for the positive semidefinite procrustes problem, Linear Algebra and Its Applications, 540 (2018), 112–137.

[6] N. Krislock, J. Lang, J. Varah, D. K. Pai and H. P. Seidel, Local compliance estimation via positive semidefinite constrained least squares, IEEE transactions on Robotics, 20 (2004), 1007–1011.

[7] C. J. Li, S. G. Zhang and H. H. Wu, The proximal point iterative algorithm for the generalized semidefinite least squares problem, ACTA Mathematicae Applicatae Sinica,(in Chinese), 42 (2019), 371–384.

[8] X. Li, D. F. Sun and K. C. Toh, A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions, Mathematical Programming, 155 (2016), 333–373.

[9] A. Liao and Z. Bai, Least-squares solution of AXB=D over symmetric positive semidefinite matrices X, Journal of Computational Mathematics, 21 (2003), 175–182.

[10] Y. Nesterov, Introductory Lectures On Convex Optimization: A Basic Course, Springer Science and Business Media, 2004.

[11] H. D. Qi, Conditional quadratic semidefinite programming: examples and methods, Journal of Operations Research Society of China, 2 (2014), 143–170.

[12] H. D. Qi, A convex matrix optimization for the additive constant problem in multidimensional scaling with application to locally linear embedding, SIAM Journal on Optimization, 26 (2016), 2564–2590.

[13] H. D. Qi and D. Sun, A quadratically convergent Newton method for computing the nearest correlation matrix, SIAM Journal on Matrix Analysis and Applications, 28 (2006), 360–385.

[14] A. Rinnan, M. Andersson, C. Ridder and B. Engelsen, Recursive weighted partial least squares(rPLS): An efficient variable selection method using PLS, Journal of Chemometrics, 28 (2014), 439–447.

[15] M. L. Stein, Spatial variation of total column ozone on a global scale, The Annals of Applied Statistics, 1 (2007), 191–210.

[16] K. C. Toh, M. J. Todd and R. H. Tutuncu, SDPT3 — a Matlab software package for semidefinite programming, Optimization Methods and Software, 11 (1999), 545–581.

[17] K. G. Woodgate, Efficient stiffness matrix estimation for elastic structures, Computers and Structures, 69 (1998), 79–84.

Received March 2020; 1st revision July 2020; Final revision September 2020.

E-mail address: chengjin982968163.com