Positive semidefinite matrix approximation with a trace constraint

Kouhei Harada *

August 8, 2018

We propose an efficient algorithm to solve positive a semidefinite matrix approximation problem with a trace constraint. Without constraints, it is well known that positive semidefinite matrix approximation problem can be easily solved by one-time eigendecomposition of a symmetric matrix. In this paper, we confirmed that one-time eigendecomposition is also sufficient even if a trace constraint is included. Although an additional binary search is necessary, it is not computationally expensive.

Key words. positive semidefinite matrix approximation, nearest matrix, trace constraint, projection onto a simplex, simplex constraint

1 Introduction

In this short paper, we are interested in the problem (P):

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\left\|X - A\right\|_F^2, \\ \text{subject to} \quad & \mathrm{Tr}\left(X\right) = b, \\ & X \succeq 0, \end{aligned} \tag{1.1}$$

where A is a $n \times n$ symmetric real matrix, b is a nonnegative scalar, and $\|\cdot\|_F$ be a Frobenius norm of a matrix. Without the constraint (1.1), it is well known that the problem (P) becomes a positive semidefinite matrix approximation problem which can be solved efficiently by one-time eigendecomposition [7].

Positive semidefinite matrix approximation problems have been widely researched with additional constraints or more generalized objective. For example, if we add $X_{ii} = 1$ for all $i = 1, \cdots, n$, the problem becomes nearest correlation matrix problem which has

*NTTDATA Mathematical Systems Inc, 1F, Shinanomachi Renga-Kan, 35, Shinanomachi, Shinjuku-ku, Tokyo, JAPAN, email:harada@msi.co.jp

many applications in finance [8]. Positive semidefinite matrix approximation problems can also include rank constraint [4] or condition number constraint [9]. If the whole of the additional constraints is an affine subspace, general approaches are reviewed by [6]. In terms of generalized objective, weighted objective $\|H \circ (X - A)\|_F^2$, where H is a $n \times n$ symmetric matrix whose element is 0 or 1 and $\circ$ is a element-wise product of matrices, is possible [1]. However, positive semidefinite matrix approximation problems with a trace constraint have not been researched as a special case.

In this paper, we propose an algorithm to solve (P) with one-time eigendecomposition and an additional binary-search which are not computationally expensive.

This paper is organized as follows. In section 2, we explain some preliminary results for positive semidefinite matrix approximation problems. In section 3, we propose an algorithm which can efficiently solve (P). In section 4, we consider to what extent the proposed algorithm can be applied. In particular, we can see that (1.1) can be replaced with an inequality such as $\mathrm{Tr}\,(X) \leq b$, $b \leq \mathrm{Tr}\,(X)$, or $\underline{b} \leq \mathrm{Tr}\,(X) \leq \bar{b}$ without compromising its efficiency. In section 5, we consider an origin of the problem (P) and it is confirmed that (P) is essentially equivalent to a projection of $\lambda(A)$, eigenvalues of A, onto a simplex $\Delta_b = \{x | \sum_{i=1}^n x_i = b,\, x \geq 0\}$. In section 6, we consider an application of alternating projection method to the problem (P) and explain the difference from the proposed algorithm. In the final section, we give a conclusion of this paper.

We use the following notation throughout this paper. For a given $x \in \mathbb{R}$, $|x|_+$ and $|x|_-$ denote $\max(x, 0)$ and $\max(-x, 0)$ respectively. For a given $a \in \mathbb{R}^n$, $|a|_+$ and $|a|_-$ denote $(|a_1|_+, \cdots, |a_n|_+)$ and $(|a_1|_-, \cdots, |a_n|_-)$ respectively. Therefore, we have $\langle |a|_+, |a|_- \rangle = 0$. We define $\delta_{ij} = 1$ if $i = j$, otherwise 0. We denote $\mathbb{S}^n$ be a set of $n \times n$ real symmetric matrices and ${\mathbb{S}_+}^n$ be a set of $n \times n$ real positive semidefinite ones.

2 Preliminaries

It is well known that applying eigendecomposition for a given $A \in \mathbb{S}^n$, we obtain

$$A = \sum_{i=1}^n \lambda_i p_i p_i^T,$$

where $\lambda = (\lambda_1, \cdots, \lambda_n)$ are eigenvalues and $P = (p_1, \cdots, p_n)$ are corresponding eigenvectors, such that $\langle p_i, p_j \rangle = \delta_{ij}$.

Using this result, we can decompose every real symmetric matrix $A \in \mathbb{S}^n$ into positive and negative parts as follows:

$$A^+ = \sum_{i=1}^n |\lambda_i|_+ p_i p_i^T,\; A^- = \sum_{i=1}^n |\lambda_i|_- p_i p_i^T.$$

For this decomposition, it is well known that following properties hold.

Lemma 2.1. *Let A be a real symmetric matrix, the following hold.*

(i) $A^+ \succeq 0$, $A^- \succeq 0$, *and* $A = A^+ - A^-$.

(ii) $\langle A^+, A^- \rangle = 0$.

Proof. Item (i) is obvious. To prove item (ii), we can see that

$$\langle A^+, A^- \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} |\lambda_i|_+ |\lambda_j|_- \langle p_i p_i^T, p_j p_j^T \rangle = \sum_{i=1}^{n} |\lambda_i|_+ |\lambda_i|_- \langle p_i p_i^T, p_i p_i^T \rangle = 0.$$

We have completed the proof. □

Using the result, we can see that the solution of (P) without (1.1) is given by $X = A^+$.

Lemma 2.2. *For a given $A \in \mathbb{S}^n$, $X = A^+$ is an optimal solution of the following minimization problem.*

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} \|X - A\|_F^2, \\ \text{subject to} \quad & X \succeq 0. \end{aligned}$$

Proof. We have that

$$\begin{aligned} \frac{1}{2} \|X - A\|_F^2 &= \frac{1}{2} \left\|X - A^+ + A^-\right\|_F^2 \\ &= \frac{1}{2} \left\|X - A^+\right\|_F^2 + \langle X - A^+, A^- \rangle + \frac{1}{2} \left\|A^-\right\|_F^2 \\ &= \frac{1}{2} \left\|X - A^+\right\|_F^2 + \langle X, A^- \rangle + \frac{1}{2} \left\|A^-\right\|_F^2. \end{aligned}$$

We have $\langle X, A^- \rangle \geq 0$ because $X \succeq 0$ and $A^- \succeq 0$. Therefore, we obtain

$$\frac{1}{2} \|X - A\|_F^2 \geq \frac{1}{2} \left\|A^-\right\|_F^2, \tag{2.1}$$

where the equality holds at $X = A^+$. □

In fact, $X = A^+$ minimizes $\|X - A\|$ subject to $X \succeq 0$ for any matrix norm, see [4, Section 3].

3 Algorithm

In this section, we propose an algorithm to solve (P). Let us consider the Karush-Kuhn-Tucker(KKT) conditions of the problem (P) given as follows:

$$X - Z - A + yI = 0, \tag{3.1}$$

$$\langle I, X \rangle = b, \tag{3.2}$$

$$X \succeq 0, \quad Z \succeq 0, \quad \langle X, Z \rangle = 0, \tag{3.3}$$

where $y \in \mathbb{R}$ is the dual value of the constraint (1.1) and $Z \succeq 0$ is the dual matrix of the positive semidefinite constraint $X \succeq 0$.

Suppose that y is fixed. In this case, by means of Lemma 2.1, X and Z which satisfies (3.1) and (3.3) are obtained by

$$X = (A - yI)^+,\ Z = (A - yI)^-.$$

Therefore, what remains is to adjust a scalar y to satisfy (3.2), and hence we can apply binary search. This is the outline of the proposed algorithm.

The point is that eigenvectors of $A - yI$ are invariable for all $y \in \mathbb{R}$. Therefore, we do not have to decompose $A - yI$ again even if y is updated. Furthermore, the fact is that we do not even need to calculate $(A - yI)^+$ except for the case of the final step in the algorithm. Following lemma is easy to confirm but essential for the proposed algorithm.

Lemma 3.1. *For $A \in \mathbb{S}^n$, the following hold.*

$$\operatorname{Tr}\left((A - yI)^+\right) = \sum_{i=1}^{n} |\lambda_i - y|_+.$$

Proof. We have that

$$\operatorname{Tr}\left((A - yI)^+\right) = \langle I, (A - yI)^+\rangle = \langle I, \sum_{i=1}^{n} |\lambda_i - y|_+ p_i p_i^T\rangle = \sum_{i=1}^{n} |\lambda_i - y|_+.$$

We have completed the proof. □

It follows that only λ and y are necessary to judge whether (3.2) is satisfied or not. Now, we are ready to propose the algorithm, one-time eigendecomposition with binary search(OEBS), to solve (P).

Algorithm 1 One-time eigendecomposition with binary search

Inputs:
choose y_U and y_L sufficiently large and small.
choose ϵ sufficiently small. given $A \in \mathbb{S}^n$ and $b \in \mathbb{R}_+$.
Initialize:
decompose $A = \sum_{i=1}^{n} \lambda_i p_i p_i^T$ such that $p_i^T p_j = \delta_{ij}$. set $r \leftarrow +\infty$.
while $|r| > \epsilon$ **do**
set $y \leftarrow (y_L + y_U)/2$.
set $r \leftarrow \sum_{i=1}^{n} |\lambda_i - y|_+ - b$.
update $y_U \leftarrow y$ **if** $r \geq 0$ or $y_L \leftarrow y$ **otherwise**
end while
set $X \leftarrow \sum_{i=1}^{n} |\lambda_i - y|_+ p_i p_i^T$.
set $Z \leftarrow X - (A - yI)$.
return (X, y, Z)

We can see that the most computationally expensive part is the one-time eigendecomposition in its initial step, and what remains is just a simple binary search to obtain proper y, which is less expensive than eigendecomposition.

4 Extensions

In this section, we consider to what extent the proposed algorithm can be generalized. In the first subsection, we consider whether the trace condition (1.1) can be replaced with an inequality. In the second subsection, we consider the possibility that (1.1) can be replaced with more generalized equality constraint $\langle C, X\rangle = b$ for $C \in \mathbb{S}^n$.

4.1 Inequality

We first consider that whether $\mathrm{Tr}\,(X) = b$ can be replaced with $\mathrm{Tr}\,(X) \leq b$. We can easily confirm that the answer is yes. Let $y \geq 0$ be a dual value of $\mathrm{Tr}\,(X) \leq b$, the KKT conditions for the replaced problem (P) is given as follows:

$$X - Z - A + yI = 0, \tag{4.1}$$

$$\langle I, X\rangle \leq b, \quad y \geq 0, \quad y(\langle I, X\rangle - b) = 0, \tag{4.2}$$

$$X \succeq 0,\; Z \succeq 0, \quad \langle X, Z\rangle = 0. \tag{4.3}$$

Conditions (4.1) and (4.3) are equivalent to (3.1) and (3.3) respectively. Therefore, $X = (A - yI)^+$ and $Z = (A - yI)^-$ satisfies (4.1) and (4.3). The new point is that we have to search y which satisfies (4.2). If $y = 0$, the required condition in (4.2) is $\langle I, A^+\rangle \leq b$. Otherwise, $\langle I, X\rangle = b$ is required to satisfy (4.2). Putting these considerations altogether, we can extend the proposed algorithm below. By a similar argument,

---

Algorithm 2 OEBS for $\mathrm{Tr}\,(X) \leq b$

1: **Inputs:**
   choose y_U sufficiently large. choose ϵ sufficiently small.
   given $A \in \mathbb{S}^n$ and $b \in \mathbb{R}$.
2: **Initialize:**
   decompose $A = \sum_{i=1}^n \lambda_i p_i p_i^T$ such that $p_i^T p_j = \delta_{ij}$.
   set $r \leftarrow +\infty$, $y \leftarrow 0$, and $y_L \leftarrow 0$.
3: **Goto** 9 **if** $\sum_{i=1}^n |\lambda_i|_+ \leq b$.
4: **while** $|r| > \epsilon$ **do**
5:    set $y \leftarrow (y_L + y_U)/2$.
6:    set $r \leftarrow \sum_{i=1}^n |\lambda_i - y|_+ - b$.
7:    update $y_U \leftarrow y$ **if** $r \geq 0$ or $y_L \leftarrow y$ **otherwise**
8: **end while**
9: set $X \leftarrow \sum_{i=1}^n |\lambda_i - y|_+ p_i p_i^T$.
10: set $Z \leftarrow X - (A - yI)$.
11: **return** (X, y, Z)

---

replacing (1.1) with $\mathrm{Tr}\,(X) \geq b$ is also possible.

If we want to replace (1.1) with $\underline{b} \leq \mathrm{Tr}\,(X) \leq \bar{b}$ such that $\underline{b} < \bar{b}$, the situation becomes more complicated. Nevertheless, we can extend the proposed algorithm. Let $\underline{y} \geq 0$ and $\bar{y} \geq 0$ be dual values of $\underline{b} \leq \mathrm{Tr}\,(X)$ and $\mathrm{Tr}\,(X) \leq \bar{b}$ respectively. Then, the KKT conditions of the replaced (P) is given as follows:

$$X - Z - A + (\bar{y} - \underline{y})I = 0, \tag{4.4}$$

$$\langle I, X\rangle \leq \bar{b}, \quad \bar{y} \geq 0, \quad \bar{y}(\langle I, X\rangle - \bar{b}) = 0, \tag{4.5}$$

$$\langle I, X\rangle \geq \underline{b}, \quad \underline{y} \geq 0, \quad \underline{y}(\langle I, X\rangle - \underline{b}) = 0, \tag{4.6}$$

$$X \succeq 0, \quad Z \succeq 0, \quad \langle X, Z\rangle = 0. \tag{4.7}$$

Let $y = \bar{y} - \underline{y}$. Then, we have that $\bar{y} = |y|_+$ and $\underline{y} = |y|_-$ because $\underline{y}\bar{y} = 0$ by definition. Therefore, we can see that $X = (A - yI)^+$ and $Z = (A - yI)^-$ satisfies (4.4) and (4.7), and hence what remains is to search y which satisfies (4.5) and (4.6).

If $y = 0$, (4.5) and (4.6) are satisfied except the primal feasibility condition, $\underline{b} \leq \langle I, X\rangle \leq \bar{b}$. Therefore, we need to calculate $\langle I, X\rangle = \langle I, A^+\rangle = \sum_{i=1}^n |\lambda_i|_+$. Even if $\underline{b} \leq \sum_{i=1}^n |\lambda_i|_+ \leq \bar{b}$ does not hold, the value of $\sum_{i=1}^n |\lambda_i|_+$ suggests whether optimal y is positive or negative. If $\sum_{i=1}^n |\lambda_i|_+ > \bar{b}$ holds, first condition of (4.5) implies $\sum_{i=1}^n |\lambda_i|_+ \geq \bar{b} \geq \langle I, X\rangle$ and $\langle I, X\rangle = \langle I, (A - yI)^+\rangle = \sum_{i=1}^n |\lambda_i - y|_+$. Thus, we have $\sum_{i=1}^n |\lambda_i|_+ > \sum_{i=1}^n |\lambda_i - y|_+$ and it follows that $y > 0$. Similarly, we deduce $y < 0$ if $\sum_{i=1}^n |\lambda_i|_+ < \underline{b}$.

In the case of $y > 0$, it implies that $\bar{y} = y > 0$ and $\underline{y} = 0$. Therefore, $\langle I, X\rangle = \bar{b}$ is necessary and sufficient to hold (4.5) and (4.6), since $\underline{b} \leq \langle I, X\rangle \leq \bar{b}$ is automatically satisfied if $\langle I, X\rangle = \bar{b}$ holds. By a similar argument, in the case of $y < 0$, (4.5) and (4.6) come down to $\langle I, X\rangle = \underline{b}$. Putting these considerations altogether, we can extend the proposed algorithm as follows.

Algorithm 3 OEBS for $\underline{b} \leq \mathrm{Tr}\,(X) \leq \bar{b}$.

1: **Inputs:**
  choose y_U and y_L sufficiently large and small.
  choose ϵ sufficiently small. given $A \in \mathbb{S}^n$ and $b \in \mathbb{R}$.
2: **Initialize:**
  decompose $A = \sum_{i=1}^n \lambda_i p_i p_i^T$ such that $p_i^T p_j = \delta_{ij}$.
  set $r = +\infty$ and $r_0 \leftarrow \sum_{i=1}^n |\lambda_i|_+$.
3: **Goto** 11 **if** $\underline{b} \leq r_0 \leq \bar{b}$
4: set $b \leftarrow \bar{b}$ and update $y_L \leftarrow 0$ **if** $r_0 > \bar{b}$.
5: set $b \leftarrow \underline{b}$ and update $y_U \leftarrow 0$ **if** $r_0 < \underline{b}$.
6: **while** $|r| > \epsilon$ **do**
7:   set $y \leftarrow (y_L + y_U)/2$.
8:   set $r \leftarrow \sum_{i=1}^n |\lambda_i - y|_+ - b$.
9:   update $y_U \leftarrow y$ **if** $r \geq 0$ or $y_L \leftarrow y$ **otherwise**
10: **end while**
11: set $X \leftarrow \sum_{i=1}^n |\lambda_i - y|_+ p_i p_i^T$.
12: set $Z \leftarrow X - (A - yI)$.
13: **return** (X, y, Z)

4.2 Generalized Constraint

In this subsection, we consider the case that (1.1) is replaced with a more generalized affine constraint:

$$\langle C, X \rangle = b. \tag{4.8}$$

In this case, the framework of the proposed algorithm is still applicable, however it is not efficient anymore. Let us consider the KKT conditions in this case,

$$X - A - Z + yC = 0, \tag{4.9}$$

$$\langle C, X \rangle = b, \tag{4.10}$$

$$X \succeq 0, \quad Z \succeq 0, \quad \langle X, Z \rangle = 0. \tag{4.11}$$

Therefore we can confirm that $X = (A - yC)^+$ and $Z = (A - yC)^-$ satisfies (4.9) and (4.11) by a similar argument. Unfortunately, in calculating $\langle C, (A - yC)^+ \rangle$, we cannot use the eigenvalues of A directly and we have to decompose $A - yC$ every time whenever y is updated. Thus, the algorithm becomes expensive.

This observation suggests that the proposed algorithm is still efficient if $X = (A-yC)^+$ is easily calculated. For example, if A and C have the same eigenvectors, then C can be decomposed as

$$C = \sum_{i=1}^{n} \mu_i p_i p_i^T,$$

and hence $\mathrm{Tr}\left((A - yC)^+\right)$ can be easily calculated without decomposing $A - yC$. Precisely, by a similar argument of Lemma 3.1, we have that

$$\mathrm{Tr}\left((A - yC)^+\right) = \langle I, (A - yC)^+ \rangle = \langle I, \sum_{i=1}^{n} |\lambda_i - y\mu_i|_+ p_i p_i^T \rangle = \sum_{i=1}^{n} |\lambda_i - y\mu_i|_+.$$

Therefore, only one-time eigendecompositon is necessary.

5 Projection onto a simplex

In this section, we describe another aspect of the proposed algorithm. Let us consider the problem of obtaining a projection of a vector $a \in \mathbb{R}^n$ to a simplex $\Delta_b = \{x \in \mathbb{R}^n | \sum_{i=1}^n x_i = b, \, x \geq 0\}$. This problem can be written as a quadratic optimization problem (S) as follows:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} \|x - a\|^2, \\ \text{subject to} \quad & e^T x = b, && (5.1) \\ & x \geq 0, && (5.2) \end{aligned}$$

where $e = (1, \cdots, 1) \in \mathbb{R}^n$. We can see that the problem (P) is a generalization of (S) by replacing vectors to symmetric matrices. Furthermore, it is confirmed that an optimal solution of (S) provides an optimal solution of (P) if a are eigenvalues of the matrix A.

Proposition 5.1. *For a given $A \in \mathbb{S}^n$, Let $a \in \mathbb{R}^n$ be eigenvalues of the matrix A. If x^* be an optimal solution of the problem (S), then an optimal X^* for (P) is given by $X^* = \sum_{i=1}^n x_i^* p_i p_i^T$, where p_i is a eigenvector of A with respect to a_i.*

Proof. Let $y \in \mathbb{R}$ and $z \in \mathbb{R}_+^n$ be dual variables of (5.1) and (5.2) respectively. Then, the KKT conditions of (S) is given as follows:

$$x - z - a + ye = 0, \tag{5.3}$$

$$e^T x = b, \tag{5.4}$$

$$x \geq 0, \quad z \geq 0, \quad x^T z = 0. \tag{5.5}$$

Therefore, $(x^*, z^*) = (|a - y^* e|_+, |a - y^* e|_-)$ satisfies (5.3) and (5.5). In order to hold (5.4), we have $\sum_{i=1}^n x_i^* = \sum_{i=1}^n |a_i - y^*|_+ = b$. It follows that $\mathrm{Tr}\,(X) = \sum_{i=1}^n x_i^* = b$, and hence (3.2) hold. Let $Z^* = \sum_{i=1}^n z_i^* p_i p_i^T$. Then, conditions (3.1) and (3.3) hold by means of Lemma 2.1. Therefore, X^* is proved to be optimal for (P). □

To solve (S), an algebric procedure is explained at [3]. Another idea of applying a binary search is proposed by [2].

6 Alternating projection method

In this section, we consider the proposed algorithm from a viewpoint of alternating projection method which has been applied to the nearest correlation matrix problem by Higham [8]. This idea can be applicable to the problem (P). To explain the procedure, we denote Π_F as a projection operator onto a convex set F. We also explicitly define the subspace of symmetric matrices which satisfies the trace constraint:

$$\mathbb{T}_b = \{X \mid X \in \mathbb{S}^n,\ \mathrm{Tr}\,(X) = b\}.$$

In our setting, the alternating projection method is given below:

Algorithm 4 Alternating projection method for (P)

```
1: Inputs:
     choose ϵ sufficiently small. given A ∈ S^n and b ∈ R.
2: Initialize:
     set X_0 ← A.
3: for k = 0, 1, 2, ··· do
4:   set X̄_k ← Π_{S^n_+}(X_k).
5:   set ΔX_k ← Π_{T_b}(X̄_k) − X̄_k.
6:   break if ‖ΔX_k‖ ≤ ϵ
7:   update X_{k+1} ← X_k + ΔX_k.
8: end for
9: return X
```

Note that we have introduced the notation $\bar{X}_k$ to make the framework easy to understand. $\bar{X}_k$ is, of course, equivalent to X_k^+.

The computational difficulty of the method depends on the cost of calculating $\Pi_{\mathbb{S}^n_+}(X_k)$ and $\Pi_{\mathbb{T}_b}(\bar{X}_k)$. In our setting, they are not expensive. Moreover, even matrix calculations are not necessary. We can compress the alternating projection method into a more simplified form using λ, eigenvalues of A, as follows:

Algorithm 5 Compressed alternating projection method for (P)

Inputs:
  choose ϵ sufficiently small. given $A \in \mathbb{S}^n$ and $b \in \mathbb{R}$.
Initialize:
  decompose $A = \sum_{i=1}^n \lambda_i p_i p_i^T$. set $x^0 \leftarrow \lambda$.
for $k = 0, 1, 2, \cdots$ **do**
  set $\Delta x^k \leftarrow \frac{1}{n}(b - \sum_{i=1}^n |x_i^k|_+)e$.
  break if $\left\|\Delta x^k\right\| \leq \epsilon$.
  update $x^{k+1} \leftarrow x^k + \Delta x^k$.
end for
set $X \leftarrow \sum_{i=1}^n |x_i^k|_+ p_i p_i^T$.
return X.

In the compressed algorithm, it can be seen that the procedure starts from $y = 0$ and moving to an optimal y^* gradually, in other words, y never leaping over an optimal y^* during the procedure. Let

$$\Delta y^k = \frac{1}{n}\left(\sum_{i=1}^n |x_i^k|_+ - b\right),\ y^{k+1} = y^k + \Delta y^k,\ y^0 = 0,$$

then the sign of y^k is invariable, the sequence $\{y^k\}$ is monotonically increasing or decreasing, and we have $y^k \to y^*$. Furthermore, we have that $x^k = \lambda - y^k$. To show these properties, we use the following lemma which is easy to confirm.

Lemma 6.1. *For a arbitrary scalar $v \in \mathbb{R}$ and a nonnegative scalar $\Delta > 0$, the folloing hold.*

(i) $|v|_+ \leq |v + \Delta|_+ \leq |v|_+ + \Delta$.

(ii) $|v|_+ \geq |v - \Delta|_+ \geq |v|_+ - \Delta$.

Proposition 6.2. *Let $f(y) = \sum_{i=1}^n |\lambda_i - y|_+ - b$, for the generated sequence $\{y^k\}$, the following hold.*

(i) *If $f(0) \geq 0$, then $f(y^k) \geq f(y^{k+1}) \geq 0$ holds for all k. Otherwise, $f(y^k) \leq f(y^{k+1}) \leq 0$ holds for all k.*

(ii) *If $f(0) \geq 0$, then $y^{k+1} \geq y^k \geq 0$ holds for all k. Otherwise, $y^{k+1} \leq y^k \leq 0$ holds for all k.*

(iii) *The sequence* $\{y^k\}$ *converges to an optimal* y^*.

Proof. We prove item (i) by induction. If $f(y^k) \geq 0$ for some k, we have $\Delta y^k = f(y^k)/n \geq 0$. Therefore, by means of Lemma 6.1 item (ii), we have

$$\begin{aligned} f(y^{k+1}) = f(y^k + \Delta y^k) &= \sum_{i=1}^{n} |\lambda_i - y^k - \Delta y^k|_+ - b \\ &\geq \sum_{i=1}^{n} |\lambda_i - y^k|_+ - n\Delta y^k - b = 0. \end{aligned}$$

On the other hand, again by means of Lemma 6.1 item (ii), we have

$$\begin{aligned} f(y^{k+1}) = f(y^k + \Delta y^k) &= \sum_{i=1}^{n} |\lambda_i - y^k - \Delta y^k|_+ - b \\ &\leq \sum_{i=1}^{n} |\lambda_i - y^k|_+ - b = f(y^k). \end{aligned}$$

Therefore $f(y^k) \geq f(y^{k+1}) \geq 0$ holds. It implies that if $f(y^0) \geq 0$ holds, we have $f(y^k) \geq f(y^{k+1}) \geq 0$ for all k. By a similar argument, if $f(y^0) \leq 0$, $f(y^k) \leq f(y^{k+1}) \leq 0$ holds for all k. We have proved item (i).

Let us move on to item (ii). If $f(0) \geq 0$, item (i) implies that $\Delta y^k = f(y^k)/n \geq 0$ for all k. Therefore, we have $y^{k+1} = y^k + \Delta y^k \geq y^k = \sum_{j=0}^{k-1} \Delta y^k \geq 0$. Similarly, if $f(0) \leq 0$, we deduce $y^{k+1} \leq y^k \leq 0$ for all k. We have proved item (ii).

By means of item (ii), the sequence $\{y^k\}$ is nondecreasing or nonincreasing and bounded. Therefore, it converges to some $\bar{y} \in \mathbb{R}$. Assume that $\bar{y}$ is not an optimal y^*, this yields $f(\bar{y}) \neq 0$. On the other hand, $\Delta y^k = f(y^k)/n$ implies that $\lim_{k\to\infty} \Delta y^k \neq 0$. This is a contradiction. We have proved item (iii). □

The compressed alternating projection method can also be regarded as a variant of dual ascent method. Let $\theta(y)$ be a dual function of the problem (P) defined as follows:

$$\begin{aligned} L(X, y) &= \frac{1}{2} \|X - A\|_F^2 + y(\langle I, X\rangle - b), \\ \theta(y) &= \min_{X \succeq 0} L(X, y). \end{aligned}$$

Let $X(y) = \arg\min_{X\succeq 0} L(X, y)$, by definition, we have

$$\langle I, X(y)\rangle - b \in \partial\theta(y).$$

By means of Lemma 2.2, $X(y) = (A - yI)^+$. This yields

$$\sum_{i=1}^{n} |\lambda_i - y|_+ - b \in \partial\theta(y).$$

Therefore, an update y^k in the alternating projection method can be seen a dual method with stepsize $1/n$. The relationship between the two methods for a more generalized case is explained at [5].

7 Conclusion

We propose an algorithm which solves the positive semidefinite matrix approximation problems with a trace constraint efficiently. The proposed algorithm is a combination of one-time eigendecomposition and a binary search. Furthermore, it is essentially equivalent to a projection of eigenvalues onto a simplex.

Alternating projection methods and dual methods are also applicable to the problem. By exploiting the structure of the problem, only one-time eigendecomposition is necessary in every case and the difference between these methods is just how to use the information of $\sum_{i=1}^{n} |\lambda_i - y|_+ - b$ to reach an optimal y^*.

References

[1] Abdo Alfakih and Henry Wolkowicz. Matrix completion problems. In *Handbook of semidefinite programming*, pages 533–545. Springer, 2000.

[2] Mathieu Blondel, Akinori Fujino, and Naonori Ueda. Large-scale multiclass support vector machine training via euclidean projection onto the simplex. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 1289–1294. IEEE, 2014.

[3] Yunmei Chen and Xiaojing Ye. Projection onto a simplex. *arXiv preprint arXiv:1101.6081*, 2011.

[4] Achiya Dax. Low-rank positive approximants of symmetric matrices. *Advances in Linear Algebra & Matrix Theory*, 4(03):172, 2014.

[5] Didier Henrion and Jérôme Malick. Projection methods for conic feasibility problems: applications to polynomial sum-of-squares decompositions. *Optimization Methods & Software*, 26(1):23–46, 2011.

[6] Didier Henrion and Jérôme Malick. Projection methods in conic optimization. In *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 565–600. Springer, 2012.

[7] Nicholas J Higham. Computing a nearest symmetric positive semidefinite matrix. *Linear algebra and its applications*, 103:103–118, 1988.

[8] Nicholas J Higham. Computing the nearest correlation matrix, a problem from finance. *IMA journal of Numerical Analysis*, 22(3):329–343, 2002.

[9] Mirai Tanaka and Kazuhide Nakata. Positive definite matrix approximation with condition number constraint. *Optimization Letters*, 8(3):939–947, 2014.