J. Appl. Math. & Informatics Vol. **28**(2010). No. 3 - 4. pp. 569 - 582
Website: http://www.kcam.biz

THE PERIODIC JACOBI MATRIX PROCRUSTES PROBLEM

JIAO-FEN LI* AND XI-YAN HU

ABSTRACT. The following "Periodic Jacobi Procrustes" problem is studied: find the Periodic Jacobi matrix X which minimizes the Frobenius (or Euclidean) norm of $AX - B$, with A and B as given rectangular matrices. The class of Procrustes problems has many application in the biological, physical and social sciences just as in the investigation of elastic structures. The different problems are obtained varying the structure of the matrices belonging to the feasible set. Higham has solved the orthogonal, the symmetric and the positive definite cases. Andersson and Elfving have studied the symmetric positive semidefinite case and the (symmetric) elementwise nonnegative case. In this contribution, we extend and develop these research, however, in a relatively simple way. Numerical difficulties are discussed and illustrated by examples.

AMS Mathematics Subject Classification: 65F30; 65H15
Key words and phrases : Periodic Jacobi matrix; procrustes problem; matrix nearness problem; singular value decomposition.

1. Introduction

Let $A \in R^{m\times n}$ be a data matrix obtained by performing a certain set of experiments, and let $B \in R^{m\times n}$ be another matrix obtained by performing the same set of experiments all over again. We are interested in the solving constrained least-squares *rectangular* matrix problems. More precisely, we consider the following constrained approximation problems:

$$\begin{cases} \min & \|AX - B\|^2 \\ s.t & \\ & X \in \mathcal{P}, \end{cases} \tag{1}$$

with $m > n$ and $\mathcal{P} \in R^{m\times n}$ has a particular pattern. In $R^{m\times n}$ we define inner

Received October 31, 2009. Revised November 9, 2009. Accepted December 10, 2009.
*Corresponding author. Research supported by the National Natural Science Foundation of China(Grant No. 10571047) and Doctorate Foundation of the Ministry of Education of China(Grant No. 20060532014).

product as:

$$\langle A,\ B\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}$$

for all $A = (a_{ij}),\ B = (b_{ij}) \in R^{m\times n}$. The norm of a matrix A generated by this inner product is the Frobenius norm, denoted by $\|A\|$.

A variety of matrix approximation problems is contained in the class (1). With no constraints on X ($\mathcal{P} = R^{n\times n}$) a standard least squares problem is obtained, having a solution $X = A^{+}B$, where A^{+} is the Moore-Penrose inverse of A (see, for example, [2, Ch.6]).

Replacing $\mathcal{P}$ by other subspaces of $R^{n\times n}$ yields other types of constrained Procrustes problems. Taking for $\mathcal{P}$ the set of orthogonal matrices yields the orthogonal Procrustes problem, which arises in a variety of applications, for example in psychometrics, in multidimensional scaling and factor analysis (see, for example [5, 6, 7, 8]). The problem is readily solved through the singular value decomposition (SVD) of B^TA: the solution is in fact the orthogonal polar factor of B^TA. Taking for $\mathcal{P}$ the set of symmetric matrices yields the symmetric Procrustes problem. This problem arises in the investigation of elastic structures wherein vectors f_i of observed forces are postulated to be related to vectors d_i of observed displacements according to $Xf_i = d_i$, where X is the symmetric strain (or flexibility) matrix [5, 9]. Higham [3] analyzes this problem by using SVD and gives a stable method to compute a solution. Higham also verifies that any solution of the normal equations (a special case of the Sylvester equations) yields a solution of the problem, and also notes that for $G(= A^TB + B^TA)$ positive semidefinite the solution of the normal equation is also definite. If $\mathcal{P}$ is the closed convex cone of symmetric positive semidefinite matrices or of symmetric elementwise nonnegative matrices, the corresponding constrained Procrustes problems reduces to problems studied by Andersson and Elfving [4].

When $m = n$ and $A = I$, a special case of (1) is the *matrix nearness problem*

$$\min_{X\in\mathcal{P}} \|X - B\|^2. \tag{2}$$

Problem (2) arises in statistics and mathematical economics. A classical example in statistics is the problem of finding the nearest symmetric positive definite patterned matrix to a sample covariance matrix. Patterned covariance matrices arise frequently from the models in physical and social sciences, see Hu and Olkin [13]. The symmetric positive semidefinite case has analyzed by Higham [11]. And an important nearness problem is to find the nearest normal matrix; see [10]. For more, we refer the reader to Refs [12, 14, 15, 16] and their references.

In this paper we explore the problem obtained from (1) when one imposes what is perhaps the simplest constraint on X, that of Periodic Jacobi. We will refer to this problem as the "Periodic Jacobi Procrustes" (PJP) problem, by analogy with the version mentioned above, that is

$$\begin{cases} \min & \|AX - B\|^2 \\ s.t & \\ & X \in \mathcal{T}, \end{cases} \qquad (3)$$

where $\mathcal{T} \in R^{n\times n}$ is the subspace of Periodic Jacobi matrices.

Let us recall that a matrix $T = (t_{ij})$ is said to be a Periodic Jacobi matrix, if it has a special symmetric structure

$$T = \begin{pmatrix} a_1 & b_1 & & & & b_n \\ b_1 & a_2 & b_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & b_{n-2} & a_{n-1} & b_{n-1} \\ b_n & & & & b_{n-1} & a_n \end{pmatrix}$$

Periodic Jacobi matrices play an important role in numerous applications in the theory of continued fractions, Padé and Hermite-Padé approximation (see [17, 18]). The complement of the essential spectrum of the operator associated to any such matrix determines, except for isolated points, the region of convergence of the Chebyshev continued fractions whose parameters are asymptotically periodic and the limits coincide with the elements of the periodic Jacobi matrix.

The plan of the paper is as follows. In section 2 we analyze the PJP problem by combining with the SVD. The general solution is derived by using the first order necessary condition and the second order necessary condition of the multidimensional function minimization problem. A lemma is presented to verify that the problem (3) has unique solution if the coefficient matrix A is full column rank, which is in accordance with the Lemma 1.1 presented in [4]. The cases of general Periodic Jacobi is analyzed in section 3, the unique solution is obtained by assuming the coefficient matrix A is full column rank. In section 4, we give a numerical example to show the efficiency of conclusion established in this paper.

This work extends the treatments in [3] and [4], however, in a totally different way even if periodic jacobi matrix is a special symmetric matrix. The main contributions being using of the SVD and using analytic methods to solve and analyze the PJP problem. We should remark that the theory and algorithms presented here are easily adapted to the "Jacobi Procrustes", "tridiagonal Procrustes", "five-diagonal Procrustes" problems, in which the constraint in (1) is taken to be that of Jacobi, tridiagonal, or five-diagonal.

2. The periodic Jacobi problem

In this section we characterize the feasibility set T of (3).

Given the matrices $A, B \in R^{m\times n}$, the subspace $\mathcal{T}$ can be represented by

$$\mathcal{T} \;= \left\{ X \in R^{n\times n} : X = \sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right\},$$

where $a_k, b_l \in R$ and $G_k, P_l \in R^{n\times n}$ are matrices defined as follows

$$(G_k)_{ij} = \begin{cases} 1, & \text{if } i = j = k; \\ 0, & \text{otherwise.} \end{cases} \quad \underset{1\le l\le n-1}{(P_l)_{ij}} = \begin{cases} 1, & \text{if } i = l, j = l+1; \\ 1, & \text{if } i = l+1, j = l; \\ 0, & \text{otherwise.} \end{cases}$$

$$(P_n)_{ij} = \begin{cases} 1, & if\ i = 1, j = n; \\ 1, & if\ i = n, j = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Clearly they form a basis for the subspace $\mathcal{T}$.

2.1 The transformed problem

We will transform the problem (3) in a simpler one. To do this transformation let A be the matrix have the following singular value decomposition

$$A = P \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} Q^T, \tag{4}$$

where $P \in R^{m\times m}$ and $Q \in R^{n\times n}$, are orthogonal matrices and $\Sigma = diag(\sigma_1, \sigma_2, \cdots, \sigma_n)$, where $\sigma_i > 0$ $(1 \le i \le n)$ is the singular value of A.

Using the invariance of the Frobenius norm under orthogonal transformation we have

$$\begin{aligned} \|AX - B\|^2 &= \|P \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} Q^T X - B\|^2 = \| \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} Q^T X - P^T B\|^2 \\ &= \| \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} Q^T X - C\|^2 = \|(\Sigma Q^T)X - C_1\|^2 + \|C_2\|^2 \\ &= \|T - C_1\|^2 + \|C_2\|^2 \end{aligned}$$

with

$$T = \Sigma Q^T X \in R^{n\times n}, \quad C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = P^T B, \quad C_1 \in R^{n\times n}.$$

then the problem (3) is equivalent to

$$\begin{cases} \min & \|T - C_1\|^2 \\ s.t & \\ & T \in \mathcal{T}' \end{cases}, \tag{5}$$

where $\mathcal{T}'$ is the following subspace:

$$\mathcal{T}' = \left\{ T \in R^{n\times n} : T = \Sigma Q^T \left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right) \right\}.$$

2.2 Characterization of the solution on $\mathcal{T}'$

We introduce the notation $P_U(A)$ meaning the projection of the matrix A on the set U. The following theorem characterizes the projection on the subspace of matrices $\mathcal{T}'$.

Theorem 1. *If $C_1 \in R^{n\times n}$, then the general solution of the problem* (5) *is*

$$P_{\mathcal{T}'}(C_1) = \Sigma Q^T\left(\sum_{k=1}^{n} a_k^* G_k + \sum_{l=1}^{n} b_l^* P_l\right),$$

$$\underset{1\le k\le n}{a_k^*} = \begin{cases} \dfrac{\sum_{i=1}^{n} \sigma_i (C_1)_{i,k} Q_{k,i}}{\sum_{i=1}^{n} \sigma_i^2 Q_{k,i}^2}, & \sum_{i=1}^{n} \sigma_i^2 Q_{k,i}^2 \neq 0; \\ \text{arbitrary}, & \text{otherwise.} \end{cases}$$

$$\underset{1\le l\le n-1}{b_l^*} = \begin{cases} \dfrac{\sum_{i=1}^{n} \sigma_i \left[(C_1)_{i,l} Q_{l+1,i} + (C_1)_{i,l+1} Q_{l,i}\right]}{\sum_{i=1}^{n} \sigma_i^2 (Q_{l+1,i}^2 + Q_{l,i}^2)}, & \sum_{i=1}^{n} \sigma_i^2 (Q_{l+1,i}^2 + Q_{l,i}^2) \neq 0, \\ \text{arbitrary}, & \text{otherwise.} \end{cases}$$

$$b_n^* = \begin{cases} \dfrac{\sum_{i=1}^{n} \sigma_i \left[(C_1)_{i,1} Q_{n,i} + (C_1)_{i,n} Q_{1,i}\right]}{\sum_{i=1}^{n} \sigma_i^2 (Q_{1,i}^2 + Q_{n,i}^2)}, & \sum_{i=1}^{n} \sigma_i^2 (Q_{1,i}^2 + Q_{n,i}^2) \neq 0, \\ \text{arbitrary}, & \text{otherwise.} \end{cases}$$

Proof. The objective function $f : R^{2n} \to R$ is given by

$$\begin{aligned} f(a,b) &= f\{a_1, \cdots, a_n, b_1, \cdots, b_{n-1}, b_n\} \\ &= \frac{1}{2}\|T - C_1\|^2 = \frac{1}{2}\|\Sigma Q^T X - C_1\|^2 \\ &= \frac{1}{2}\left\|\Sigma Q^T\left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l\right) - C_1\right\|^2 \\ &= \frac{1}{2}\sum_{i,j=1}^{n}\left[\left(\Sigma Q^T\left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l\right)\right)_{i,j} - (C_1)_{i,j}\right]^2 \end{aligned}$$

Since f is twice continuously differentiable, we compute $\dfrac{\partial f}{\partial a_p}(a,b)$ for all p such that $1 \le p \le n$

$$\begin{aligned} \frac{\partial f}{\partial a_p} &= \frac{\partial}{\partial a_p}\left\{\frac{1}{2}\sum_{i,j=1}^{n}\left[\left(\Sigma Q^T\left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l\right)\right)_{i,j} - (C_1)_{i,j}\right]^2\right\} \\ &= \sum_{i,j=1}^{n}\left[\left(\left(\Sigma Q^T\left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l\right)\right)_{i,j} - (C_1)_{i,j}\right)\right. \\ &\qquad \left.\frac{\partial}{\partial a_p}\left(\Sigma Q^T\left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l\right)\right)_{i,j}\right]. \end{aligned} \tag{6}$$

The factor in the right hand side of (6) can be written as

$$\begin{aligned}
&\frac{\partial}{\partial a_p}\left(\Sigma Q^T\left(\sum_{k=1}^n a_k G_k + \sum_{l=1}^n b_l P_l\right)\right)_{i,j} \\
&= \frac{\partial}{\partial a_p}\left[\sum_{s=1}^n\left((\Sigma Q^T)_{i,s}\cdot\left(\sum_{k=1}^n a_k G_k + \sum_{l=1}^n b_l P_l\right)_{s,j}\right)\right] \\
&= \sum_{s=1}^n\left[(\Sigma Q^T)_{i,s}\cdot\frac{\partial}{\partial a_p}\left(\sum_{k=1}^n a_k G_k + \sum_{l=1}^n b_l P_l\right)_{s,j}\right] \\
&= \sum_{s=1}^n(\Sigma Q^T)_{i,s}(G_p)_{s,j} = (\Sigma Q^T G_p)_{i,j},
\end{aligned}$$

and (6) becomes in

$$\frac{\partial f}{\partial a_p} = \sum_{i,j=1}^n\left[\left(\left(\Sigma Q^T\left(\sum_{k=1}^n a_k G_k + \sum_{l=1}^n b_l P_l\right)\right)_{i,j} - (C_1)_{i,j}\right)\cdot(\Sigma Q^T G_p)_{i,j}\right]. \tag{7}$$

Having in mind that

$$\left(\Sigma Q^T\left(\sum_{k=1}^n a_k G_k + \sum_{l=1}^n b_l P_l\right)\right)_{i,j}$$

is the inner product between the i^{th} row of (ΣQ^T) and the j^{th} column of $\left(\sum_{k=1}^n a_k G_k + \sum_{l=1}^n b_l P_l\right)$, we obtain

$$\begin{aligned}
&\left(\Sigma Q^T\left(\sum_{k=1}^n a_k G_k + \sum_{l=1}^n b_l P_l\right)\right)_{i,j} \\
&= \begin{cases} \sigma_i(a_1 Q_{1,i} + b_1 Q_{2,i} + b_n Q_{n,i}) & j = 1; \\ \sigma_i(b_{j-1}Q_{j-1,i} + a_j Q_{j,i} + b_j Q_{j+1,i}) & j \neq 1 \text{ and } j \neq n; \\ \sigma_i(b_n Q_{1,i} + b_{n-1}Q_{n-1,i} + a_n Q_{n,i}) & j = n \end{cases}
\end{aligned} \tag{8}$$

Similarly, the element $(\Sigma Q^T G_p)_{i,j}$, is the inner product between the i^{th} row of ΣQ^T and the j^{th} column of G_p. Then

$$(\Sigma Q^T G_p)_{i,j} = \begin{cases} \sigma_i Q_{p,i} & j = p; \\ 0 & j \neq p \end{cases}.$$

So (7) becomes in

$$\frac{\partial f}{\partial a_p} = \sum_{i=1}^n\left[\sigma_i\left(\Sigma Q^T\left(\sum_{k=1}^n a_k G_k + \sum_{l=1}^n b_l P_l\right)\right)_{i,p}\cdot Q_{p,i}\right] - \sum_{i=1}^n \sigma_i (C_1)_{i,p} Q_{p,i}. \tag{9}$$

And together with (8), the first sum item in the right hand side of (9) can be written as

$$\sum_{i=1}^{n}\left[\sigma_i\left(\Sigma Q^T\left(\sum_{k=1}^{n}a_kG_k+\sum_{l=1}^{n}b_lP_l\right)\right)_{i.p}\cdot Q_{p.i}\right]$$

$$=\left\{\begin{array}{l}\sum_{i=1}^{n}\sigma_i^2(a_1Q_{1.i}+b_1Q_{2.i}+b_nQ_{n.i})Q_{1.i}, p=1;\\ \sum_{i=1}^{n}\sigma_i^2(b_{p-1}Q_{p-1,i}+a_pQ_{p,i}+b_pQ_{p+1,i})Q_{p,i}, p\neq 1,\ \text{and } p\neq n;\\ \sum_{i=1}^{n}\sigma_i^2(b_nQ_{1.i}+b_{n-1}Q_{n-1,i}+a_nQ_{n.i})Q_{n,i}, p=n;\end{array}\right. \tag{10}$$

Defining $\sum_{i=1}^{n}(Q_{h,i}Q_{g,i})\sigma_i^2=\nu^T\theta$, with

$$\nu^T=(Q_{h,1}Q_{g,1},Q_{h,2}Q_{g,2},\cdots,Q_{h,n}Q_{g,n});\quad \theta=(\sigma_1^2,\sigma_2^2,\cdots,\sigma_n^2)^T,$$

we have that the sum of all the components of ν is the inner product between two columns of the orthogonal matrix Q and then it is zero when $h\neq g$. Therefore for $h\neq g$ we have

$$0=\left(\sum_{i=1}^{n}\nu_i\right)\sigma_n^2\leq\nu^T\theta=\sum_{i=1}^{n}\nu_i\sigma_i^2\leq\left(\sum_{i=1}^{n}\nu_i\right)\sigma_1^2=0,$$

then $\nu^T\theta=0$ for all $m\neq n$ and $\nu^T\theta=\sum_{i=1}^{n}\sigma_i^2Q_{m,i}^2$ for $m=n$. So (10) becomes in

$$\sum_{i=1}^{n}\left[\sigma_i\left(\Sigma Q^T\left(\sum_{k=1}^{n}a_kG_k+\sum_{l=1}^{n}b_lP_l\right)\right)_{i,p}\cdot Q_{p,i}\right]=\sum_{i=1}^{n}\sigma_i^2a_pQ_{p,i}^2,$$

for all $p=1,\cdots,n$. Then (9) becomes $\dfrac{\partial f}{\partial a_p}=\sum_{i=1}^{n}\sigma_i^2a_pQ_{p,i}^2-\sum_{i=1}^{n}\sigma_i(C_1)_{i,p}Q_{p,i}$.

Therefore, from the first order necessary condition $\dfrac{\partial f}{\partial a_p}=0,\quad p=1,\cdots,n$, we obtain

$$a_p^*=\frac{\sum_{i=1}^{n}\sigma_i(C_1)_{i,p}Q_{p,i}}{\sum_{i=1}^{n}\sigma_i^2Q_{p,i}^2}. \tag{11}$$

Now we observe that the denominator of (11), it is non zero when $rank(A)=n$, for all singular value of A are positive, then $\sum_{i=1}^{n}\sigma_i^2Q_{p,i}^2=0$ implies $\sigma_iQ_{p,i}=0\Rightarrow$ $[Q_{p.1},Q_{p.2},\cdots,Q_{p.n}]=0$, which is contradicted with Q is orthogonal. But it is possible to be zero if $rank(A)=k<n$, it just let the first k elements of the p^{th}

of Q are all zeros and the rest elements at least have one non-zero element. So we must consider the case of the denominator of (11) is zero, we let a_p^* to be arbitrary if $\sum_{i=1}^{n} \sigma_i^2 Q_{p,i}^2 = 0$.

Similar as $\frac{\partial f}{\partial a_p}$, we compute $\frac{\partial f}{\partial b_q}$ for all $1 \le q \le n$, we have

$$\begin{aligned}\frac{\partial f}{\partial b_q} = \sum_{i,j=1}^{n} \Bigg[& \left(\left(\Sigma Q^T \left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right) \right)_{i,j} - (C_1)_{i,j} \right) \\ & \cdot \frac{\partial}{\partial b_q} \left(\Sigma Q^T \left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right) \right)_{i,j} \Bigg]. \end{aligned} \tag{12}$$

We also have $\frac{\partial}{\partial b_q} \left(\Sigma Q^T \left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right) \right)_{i,j} = (\Sigma Q^T P_q)_{i,j}$, and expression $\left(\Sigma Q^T \left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right) \right)_{i,j}$ as in (8).

So, (12) becomes in

$$\begin{aligned}\frac{\partial f}{\partial b_q} = & \sum_{i,j=1}^{n} \left[\left(\Sigma Q^T \left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right) \right)_{i,j} \cdot (\Sigma Q^T P_q)_{i,j} \right] \\ & - \sum_{i,j=1}^{n} (C_1)_{i,j} (\Sigma Q^T P_q)_{i,j}. \end{aligned} \tag{13}$$

Noting that, the element $(\Sigma Q^T P_q)_{i,j}$, is the inner product between the i^{th} row of ΣQ^T and the j^{th} column of j of P_q. Then

$$(\Sigma Q^T P_q)_{i,j} = \begin{cases} \sigma_i Q_{q+1,i}, & j = q; \\ \sigma_i Q_{q,i}, & j = q+1; \quad \text{for } \ 1 \le q \le n-1, \\ 0, & \text{else.} \end{cases} \tag{14}$$

In particular when $q = n$

$$(\Sigma Q^T P_n)_{i,j} = \begin{cases} \sigma_i Q_{n,i}, & j = 1; \\ \sigma_i Q_{1,i}, & j = n; \\ 0, & \text{else.} \end{cases} \tag{15}$$

When $j = q$ and $1 \le q \le n-1$, combining with (8) and (14), the first sum item in the right hand side of (13) can be written as

$$\sum_{i=1}^{n} \left[\left(\Sigma Q^T \left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right) \right)_{i,q} \cdot (\Sigma Q^T P_q)_{i,q} \right]$$

$$
\begin{aligned}
&= \begin{cases} \sum\limits_{i=1}^{n} \sigma_i^2 (a_1 Q_{1.i} + b_1 Q_{2.i} + b_n Q_{n.i}) Q_{2.i}, & q = 1; \\ \sum\limits_{i=1}^{n} \sigma_i^2 (b_{q-1} Q_{q-1.i} + a_q Q_{q.i} + b_q Q_{q+1.i}) Q_{q+1.i}, & 1 < q \leq n-1. \end{cases} \\
&= \sum_{i=1}^{n} \sigma_i^2 b_q Q_{q+1.i}^2, \text{ for all } 1 \leq q \leq n-1.
\end{aligned} \tag{16}
$$

When $j = q+1$ and $1 \leq q \leq n-1$

$$
\sum_{i=1}^{n} \left[\left(\Sigma Q^T \left(\sum_{k=1}^{n} a_k G_k + \sum_{l=1}^{n} b_l P_l \right) \right)_{i,q+1} \cdot (\Sigma Q^T P_q)_{i,q+1} \right]
$$

$$
\begin{aligned}
&= \begin{cases} \sum\limits_{i=1}^{n} \sigma_i^2 (b_q Q_{q,i} + a_{q+1} Q_{q+1,i} + b_{q+1} Q_{q+2,i}) Q_{q,i}, & 1 \leq q < n-1. \\ \sum\limits_{i=1}^{n} \sigma_i^2 (b_n Q_{1,i} + b_{n-1} Q_{n-1,i} + a_n Q_{n,i}) Q_{n-1,i}, & q = n-1. \end{cases} \\
&= \sum_{i=1}^{n} \sigma_i^2 b_q Q_{q,i}^2, \text{ for all } 1 \leq q \leq n-1.
\end{aligned} \tag{17}
$$

And the second sum item

$$
\sum_{i,j=1}^{n} (C_1)_{i,j} (\Sigma Q^T P_q)_{i,j} = \sum_{i=1}^{n} \sigma_i (C_1)_{i,q} Q_{q+1,i} + \sum_{i=1}^{n} \sigma_i (C_1)_{i,q+1} Q_{q,i}.
$$

Therefore, (13) becomes

$$
\begin{aligned}
&\frac{\partial f}{\partial b_q} \\
&= \sum_{i=1}^{n} \sigma_i^2 b_q Q_{q+1,i}^2 + \sum_{i=1}^{n} \sigma_i^2 b_q Q_{q,i}^2 - \left(\sum_{i=1}^{n} \sigma_i (C_1)_{i,q} Q_{q+1,i} + \sum_{i=1}^{n} \sigma_i (C_1)_{i,q+1} Q_{q,i} \right),
\end{aligned}
$$

for all $q = 1, \cdots, n-1$. Therefore, from the first order necessary condition $\dfrac{\partial f}{\partial b_q} = 0, \quad q = 1, \cdots, n-1$, we obtain

$$
b_q^* = \frac{\sum\limits_{i=1}^{n} \sigma_i \left[(C_1)_{i,q} Q_{q+1,i} + (C_1)_{i,q+1} Q_{q,i} \right]}{\sum\limits_{i=1}^{n} \sigma_i^2 (Q_{q+1,i}^2 + Q_{q,i}^2)}. \tag{18}
$$

Similarly, the denominator of (18) is non zero if $rank(A) = n$, but it is possible to be zero if $rank(A) = k < n$, it just let the first k elements of the q^{th} row and $(q+1)^{th}$ row of Q are all zeros, and the rest elements of the two row each have at least one non-zero element. So we define b_q^* to be arbitrary if the the denominator is zero.

Especially, when $p = n$, combining with (8) and (15), the first sum item in the right hand side of (13) can be written as

$$\sum_{i=1}^{n}\left[\sigma_i\left(\Sigma Q^T\left(\sum_{k=1}^{n} a_k G_k+\sum_{l=1}^{n} b_l P_l\right)\right)_{i,1}\cdot Q_{n,i}\right]$$
$$+\sum_{i=1}^{n}\left[\sigma_i\left(\Sigma Q^T\left(\sum_{k=1}^{n} a_k G_k+\sum_{l=1}^{n} b_l P_l\right)\right)_{i,n}\cdot Q_{1,i}\right]$$
$$=\sum_{i=1}^{n}\sigma_i^2(a_1Q_{1,i}+b_1Q_{2,i}+b_nQ_{n,i})Q_{n,i}$$
$$+\sum_{i=1}^{n}\sigma_i^2(b_nQ_{1,i}+b_{n-1}Q_{n-1,i}+a_nQ_{n,i})Q_{1,i}$$
$$=\sum_{i=1}^{n}\sigma_i^2 b_n(Q_{1,i}^2+Q_{n,i}^2).$$

And the second sum item

$$\sum_{i,j=1}^{n}(C_1)_{i,j}(\Sigma Q^T P_q)_{i,j}=\sum_{i=1}^{n}\sigma_i(C_1)_{i,1}Q_{n,i}+\sum_{i=1}^{n}\sigma_i(C_1)_{i,n}Q_{1,i}.$$

Therefore, when $q=n$ (13) becomes

$$\frac{\partial f}{\partial b_n}=\sum_{i=1}^{n}\sigma_i^2 b_n(Q_{1,i}^2+Q_{n,i}^2)-\left(\sum_{i=1}^{n}\sigma_i(C_1)_{i,1}Q_{n,i}+\sum_{i=1}^{n}\sigma_i(C_1)_{i,n}Q_{1,i}\right).$$

Therefore, from the first order necessary condition $\frac{\partial f}{\partial b_n}=0$, we obtain

$$b_n^*=\frac{\sum_{i=1}^{n}\sigma_i[(C_1)_{i,1}Q_{n,i}+(C_1)_{i,n}Q_{1,i}]}{\sum_{i=1}^{n}\sigma_i^2(Q_{1,i}^2+Q_{n,i}^2)}. \tag{19}$$

Similarly, the denominator of (19) is non zero if $rank(A)=n$, and we let b_n^* be arbitrary if the denominator is zero.

Now, we need to verify that $a_p^*(1\le p\le n)$, $b_q^*(1\le q\le n-1)$, b_n^* is a minimizer of f . We compute

$$\frac{\partial^2 f}{\partial a_p^2}=\sum_{i=1}^{n}\sigma_i^2Q_{p,i}^2\ge 0, \frac{\partial^2 f}{\partial b_q^2}=\sum_{i=1}^{n}\sigma_i^2Q_{q+1,i}^2\ge 0, \frac{\partial^2 f}{\partial b_n^2}=\sum_{i=1}^{n}\sigma_i^2(Q_{1,i}^2+Q_{n,i}^2)\ge 0,$$

it is easy to verify that the above three inequalities can not be 0 simultaneously for Q is orthogonal, especially, the strict inequality are hold when $rank(A)=n$, and we also have

$$\frac{\partial^2 f}{\partial a_p a_k}=0,\quad if\quad k\ne p,\quad \frac{\partial^2 f}{\partial b_a b_l}=0,\quad if\quad l\ne q;$$

$$\frac{\partial^2 f}{\partial a_p b_q}=0,\quad \frac{\partial^2 f}{\partial a_p b_n}=0,\quad \frac{\partial^2 f}{\partial b_q a_p}=0,\quad \frac{\partial^2 f}{\partial b_q b_n}=0,\quad \frac{\partial^2 f}{\partial b_n a_p}=0,\quad \frac{\partial^2 f}{\partial b_n b_q}=0.$$

It implies that the Hessian matrix $\nabla^2 f(a^*, b^*)$ is positive semidefinite, especially, it is positive definite when $rank(A) = n$, therefore $a_p^*(1 \leq p \leq n)$, $b_q^*(1 \leq q \leq n-1)$, b_n^* is the minimizer of f . $\square$

Noting that from Theorem 1, the coefficients a_p^*, b_q^*, b_n^* are only dependent on the three factors P Σ and Q in the singular value decomposition of A. But unlike the diagonal factor Σ, the left and the right orthogonal factors P, and Q are never uniquely determined, the degree of nonuniqueness depends on the multiplicities of the singular values. Similar as Theorem 3.11' in [1], we can characterize the set of all possible left and right orthogonal factors in a singular value decomposition as follows:

Lemma 1. *Let $A \in R^{m\times n}$ be given, suppose that the distinct nonzero singular values of A are $\sigma_1 > \cdots > \sigma_k > 0$, with respective multiplicities $\mu_1, \cdots, \mu_k \geq 1$. Let $\mu_1 + \cdots + \mu_k = r$ and let $A = Pdiag(\Sigma, 0_{m-r,n-r})Q$ be a given singular value decomposition with $\Sigma = diag(\sigma_1 I_{\mu_1}, \cdots, \sigma_k I_{\mu_k}) \in R^{r\times r}$. Let $\hat{P} \in R^{m\times m}$ and $\hat{Q} \in R^{n\times n}$ be given orthogonal matrices. Then $A = \hat{P}\Sigma\hat{W}^T$ if and only if there are orthogonal matrices $U_i \in R^{\mu_i\times\mu_i}$, $i = 1, \cdots, k$, $\tilde{V} \in R^{(m-r)\times(m-r)}$, and $\tilde{W} \in R^{(n-r)\times(n-r)}$ such that*

$$\hat{P} = P[U_1 \oplus \cdots \oplus U_k \oplus \tilde{V}] \quad \text{and} \quad \hat{Q} = Q[U_1 \oplus \cdots \oplus U_k \oplus \tilde{W}]. \qquad (20)$$

In the next, together with lemma 1, we will prove that the uniqueness of the solution of Problem (3) when the coefficient matrix A is full column rank, and the solution is not unique if $rank(A) < n$, further, it is associated with the matrices $\tilde{V}$ and $\tilde{W}$ which were characterized in lemma 1. We first have the following well-known result

Lemma 2. *Let H be the linear space of real matrices with a fixed dimension and C be a closed convex and nonempty cone in H. Assume that f is convex and coercive[1]. Then Problem $\min_{X\in C} f(X)$ has a solution. If in addition f is strictly convex, the solution is unique.*

Suppose that the coefficient matrix A is full column rank and has another form of singular value decomposition $A = \hat{P}\begin{bmatrix}\Sigma\\0\end{bmatrix}\hat{Q}$, where

$$\hat{P} = P\begin{pmatrix} U_1 & & & \\ & \ddots & & \\ & & U_k & \\ & & & \tilde{V}\end{pmatrix} \quad \text{and} \quad \hat{Q} = Q\begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_k\end{pmatrix}$$

[1]On H we introduce the inner product $\langle X, Y\rangle = trace(XY^T)$, X, $Y \in H$. The function f is called coercive on C if

$$\min_{\|X\|\to\infty, X\in C} f(X) = \infty.$$

Partitioning $P = [P_1, P_2]$, $\hat{P} = [\hat{P}_1, \hat{P}_2]$, where $P_1 \in R^{m\times n}, \hat{P}_1 \in R^{m\times n}$, then

$$\hat{P}_1 = P_1 \begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_k \end{pmatrix}$$

Then by Theorem 1, we can obtain another solution $\hat{a}_p^*(1 \le p \le n)$, $\hat{b}_q^*(1 \le q \le n-1)$ and $\hat{b}_n^*$, which are associated with matrices $\hat{Q}$ and $\hat{C}_1(= \hat{P}_1^T B)$, the concrete expressions are similar as (11), (18), (19). But by straightforward computations, we have

$$\begin{aligned}&\sum_{i=1}^{n} \sigma_i \hat{Q}_{f,i} (\hat{C}_1)_{i,h} \\ &= [\sigma_1 Q_{f,1}, \cdots, \sigma_n Q_{f,n}] \begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_k \end{pmatrix} \begin{pmatrix} U_1^T & & \\ & \ddots & \\ & & U_k^T \end{pmatrix} P_1^T \begin{pmatrix} B_{1,h} \\ \vdots \\ B_{n,h} \end{pmatrix} \\ &= \sum_{i=1}^{n} \sigma_i Q_{f,i} (C_1)_{i,h}.\end{aligned}$$

$$\begin{aligned}&\sum_{i=1}^{n} \sigma_i^2 \hat{Q}_{f,i}^2 \\ &= [\sigma_1 Q_{f,1}, \cdots, \sigma_n Q_{f,n}] \begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_k \end{pmatrix} \begin{pmatrix} U_1^T & & \\ & \ddots & \\ & & U_k^T \end{pmatrix} \begin{pmatrix} \sigma_1 Q_{f,1} \\ \vdots \\ \sigma_n Q_{f,n} \end{pmatrix} \\ &= \sum_{i=1}^{n} \sigma_i^2 Q_{f,i}^2.\end{aligned}$$

Therefore, it is easy to see that $\hat{a}_p^* = a_p(1 \le p \le n)$, $\hat{b}_q^* = b_q(1 \le q \le n-1)$ and $\hat{b}_n^* = b_n$, which implies that when matrix A is full column rank, the solution of (3) is unique, which is in accordance with the lemma 2. We also know from the above discussion that if $rank(A) < n$, then the solutions are associated with the matrices $\tilde{V}$ and $\tilde{W}$ defined by (20), which implies the solution is not unique.

3. A numerical algorithm for solving problem (3)

Based on Theorem 1, we can establish an algorithm for finding the solution X of problem (3). To this end, we let $A,\ B \in R^{m\times n}$ $(m \ge n)$. The following algorithm computes the solution X of problem (3).

Algorithm 1.
Input: $A,\ B \in R^{m\times n}$ $(m \ge n)$.
Output: X.
Begin:
Step 1: Find the SVD of the matrix A by (4) and then partition matrix

$P = [P_1, P_2],\ P_1 \in R^{m\times n}$, determine the matrix $C_1 = P_1^T B$.

Step 2: Compute $a_p^*(1 \le p \le n)$, $b_q^*(1 \le q \le n)$ by Theorem 1.

Step 3: Compute the solution X of problem (3). If $rank(A) = n$, the solution is the unique solution of the problem.

End

In the following, we will give a numerical example with coefficient matrices to verify our algorithm. All the tests are performed using MATLAB 7.0 which has a machine precision 10^{-12}. Above all, we consider a general example.

Example Let matrices

$$A = \begin{pmatrix} -4.1 & 4.9 & 4.3 & 6.2 & 6.8 & 0.5 & 1.5 & 8.9 \\ 3.0 & 2.1 & 9.3 & 3.7 & 0.9 & -3.6 & -6.7 & 2.7 \\ 8.7 & 6.4 & 6.8 & 5.7 & 0.3 & -6.3 & -6.9 & 2.5 \\ 0.1 & 3.2 & 2.1 & 4.5 & 6.1 & -7.1 & 7.2 & -8.6 \\ -7.6 & -9.6 & 8.3 & 0.4 & -6.0 & 6.9 & 4.7 & -2.3 \\ -9.7 & 7.2 & 6.2 & 0.2 & 0.1 & 0.8 & -5.5 & -8.0 \\ 9.9 & -4.1 & -1.3 & -3.1 & 0.1 & 4.5 & 1.2 & 9.0 \\ 7.8 & -7.4 & -2.1 & 0 & 1.9 & 4.4 & -4.5 & 2.3 \\ -4.3 & -2.6 & 6.2 & 3.8 & 5.8 & 3.5 & 7.1 & 2.3 \end{pmatrix},$$

$$B = \begin{pmatrix} 1.04 & 0.67 & 2.49 & 3.56 & 2.21 & 1.71 & 1.59 & 1.57 \\ 0.64 & 0.80 & 2.80 & 3.00 & 0.02 & -1.36 & -1.38 & 0.11 \\ 1.07 & 1.04 & 2.99 & 2.85 & -0.28 & -1.97 & -1.83 & 0.89 \\ -1.25 & 0.40 & 1.58 & 2.56 & 0.23 & 1.12 & -0.87 & -0.88 \\ -0.97 & -0.32 & 1.14 & 0.81 & 0.58 & 0.56 & 1.54 & -0.98 \\ -1.88 & 0.95 & 1.76 & 1.37 & 0.21 & -0.63 & -2.02 & -3.98 \\ 2.06 & -0.37 & -1.18 & -1.00 & 0.42 & 0.93 & 2.12 & 3.64 \\ 0.88 & -0.71 & -0.90 & -0.11 & 1.21 & 0.47 & 0.11 & 1.04 \end{pmatrix}.$$

where $rank(A) = 8$ (full column rank). Compute the SVD of A, we obtain the singular values of A as follow

(25.6126, 24.2777, 19.2189, 16.4819, 12.6274, 7.5266, 4.5672, 2.3510).

Then by using Algorithm 1, we obtain the unique solution of problem (3) is

$$X = \begin{pmatrix} 0.1386 & -0.0243 & 0 & 0 & 0 & 0 & 0 & 0.1982 \\ -0.0243 & 0.1213 & 0.0426 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0426 & 0.1777 & 0.1256 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1256 & 0.2238 & 0.1458 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1458 & 0.2186 & 0.1452 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1452 & 0.2089 & 0.1588 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.1588 & 0.2073 & 0.1495 \\ 0.1982 & 0 & 0 & 0 & 0 & 0 & 0.1495 & 0.2201 \end{pmatrix}.$$

By concrete computations, we can further get $\|AX - B\| = 2.9358$.

REFERENCES

1. R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.
2. G.H. Golub and C.F. Van Loan, *Matrix computations*, 2nd ed. Johns Hopkins University Press, Baltimore, Mayland, 1989.
3. N.J. higham, *The symmetric prorustes problem*, BIT **28** (1988), 133-143.
4. L.E. Andersson and T. Elfving, *A constrained procrustes problem*, SIAM J. Matrix Anal. Appl. **18** (1997), 124-139.
5. J.E. Brock, *Optimal matrices describing linear systems*, AIAA Journal **6** (1968), 1292-1296.
6. J.C. Gower, *Multivariable analysis: Ordination, multidimensional scaling and allied topics*, in: Handbook of Applicable Mathematics, Vol. VI: Statistics, ed. E.H. Lloyd (Wiley, Chichster, 1984) pp. 727-781.
7. B. Green, *The orthogonal approximation of an oblique structure in factor analysis*, Psychometrika **17** (1952), 429-440.
8. P. Schonemann, *A generalized solution of the orthogonal Procrustes problem*, Psychometrika **31** (1966), 1-10.
9. H.J. Larson, *Least-squares estimation of the components of a symmetric matrix*, Technometrics **8** (1966), 360-362.
10. A. Ruhe, *Cloest normal matrix finally found!*, BIT **27** (1987), 585-598.
11. N.J. higham, *Computing a nearest symmetric positive semi-definite matrix*, Linear Algebra Appl. **103** (1988), 103-118.
12. M.G. Eberle and M.C. Maciel, *Finding the closest Toeplitz matrix*, Comput. Appl. Math. **1** (2003), 1-18.
13. H. Hu and I. Olkin, *A numerical procedure for finding the positive definite matrix cloest to a patterned matrix*, Stat. Prob. Letters, **12** (1991), 511-515.
14. F.Z. Zhou, X.Y. Hu and L. Zhang, *The solvability conditions for the inverse problems of symmetric-ortho-symmetric matrices*, Appl. Math. Comput. **154** (2004), 153-166.
15. D.X. Dong, L.Zhang and X.Y. Hu, *The solvability conditions for the inverse problem of bisymmetric nonnegative definite matrices*, J. Comput. Math. **6** (2000), 597-608.
16. Y. Lei, A.P. Liao and L. Zhang, *Minimization problem for symmetric orthogonal anti-symmetric matrices*, J. Comput. Math. **25** (2007), 211-220.
17. S.A. Andrea and T.G. Berry, *Continued fractions and periodic Jacobi matrices*, Linear Algebr. Appl. **161** (1992), 117-134.
18. B. Beckermann and V. A. Kaliaguine, *The diagonal of the Padé table and the approximation of the Weyl function of second order difference operators*, Constr. Approx. **13** (1997), 481-510.
19. A. Almendral Vázquez, *The Spectrum of a Periodic Complex Jacobi Matrix Revisited*, J. Approx. Theo. **105** (2000), 344-351.

Jiao-fen Li received the M.S., and Ph.D. degrees from the college of Mathematics and Econometrics, Hunan University, Changsha, P.R. China, in 2007, 2009, respectively. His research interests focus on Numerical algebra, Inverse problem, Iteration algorithms for linear matrix equation and Model updating problem.

College of Mathematics and Econometrics, Hunan University, Changsha 410082, P.R. China
e-mail: `lixiaogui1290@163.com`