Journal of Southeast University (English Edition) Vol. 21, No. 1, pp. 119 - 122 Mar. 2005 ISSN 1003—7985

Some studies on finding the nearest volume-preserving matrix

Mu Jianfei Huang Jianguo

(Department of Mathematics, Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: Finding the nearest volume-preserving matrix for a given matrix is studied. A matrix equation is first obtained, which is a necessary condition for the solution to the problem. Then the equation is solved by the singular value decomposition method. Some additional results are also provided to further characterize the solution. Using these results, a numerical algorithm is introduced and a numerical test is given to illustrate the effectiveness of the algorithm.

Key words: volume-preserving matrix; matrix nearness problem; singular value decomposition

We consider the following nearest volume-preserving matrix problem: For any given matrix $\boldsymbol{A} \in \mathbf{R}^{n\times n}$, find a matrix $\hat{\boldsymbol{X}} \in S(n)$ such that

$$\hat{\boldsymbol{X}} = \arg \min_{\boldsymbol{X}\in S(n)} \|\boldsymbol{A} - \boldsymbol{X}\|_{\mathrm{F}}^2 \tag{1}$$

where $S(n)$ denotes the set of all real volume-preserving matrices of order n,

$$S(n) \equiv \{\boldsymbol{A} \in \mathbf{R}^{n\times n}: \det(\boldsymbol{A}) = 1\}$$

which is a Lie group with the Lie algebra[1]:

$$s(n) \equiv \{\boldsymbol{A} \in \mathbf{R}^{n\times n}: \mathrm{tr}(\boldsymbol{A}) = 0\}$$

Such kinds of problems have wide applications in signal processing and computer vision (see Refs. [2 - 7]). The problem considered here can be viewed as a "dual" problem of the one considered in Ref. [7], in the sense that the constraint sets of the two problems are $S(n)$ and $s(n)$ respectively, and the latter one is just the Lie algebra of the former one. On the other hand, observing that $S(n)$ consists of all volume-preserving affine transformations which have practical applications[8], when we obtain an approximate transformation by measurements, we have to solve problem (1) to get the modified one with a physical sense. We refer to Refs. [3, 6, 9] for more details along this line.

In this paper, we first derive a matrix equation which should be satisfied by the solution to problem (1). We then solve the equation by the singular value decomposition method. We have also obtained a result characterizing the solution of (1) in detail. Finally, a numerical algorithm is provided by means of the previous results and a numerical test is given to illustrate the effectiveness of the algorithm.

1 Necessary Conditions for Solutions of (1)

In this section, we first give a necessary condition for the solution of (1).

Lemma 1 Let $\hat{\boldsymbol{X}} \in S(n)$ solve the least square problem (1). Then it must satisfy the following matrix equation:

$$\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X} - \boldsymbol{A}) = \lambda \boldsymbol{I}_n \qquad \boldsymbol{X} \in S(n) \tag{2}$$

where $\lambda \in \mathbf{R}$ is some real constant and $\boldsymbol{I}_n$ denotes the $n \times n$ identity matrix.

Proof It is evident for any matrix $\boldsymbol{W} \in s(n)$ and any $t \in \mathbf{R}$ that $\det(\mathrm{e}^{t\boldsymbol{W}}) = 1$., Hence, the smooth curve $\{\hat{\boldsymbol{X}}\mathrm{e}^{t\boldsymbol{W}}\}_{t\in\mathbf{R}}$ is on the Lie group $S(n)$. Noting that $\hat{\boldsymbol{X}}$ is the solution of (1), we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{A} - \hat{\boldsymbol{X}}\mathrm{e}^{t\boldsymbol{W}}\|_{\mathrm{F}}^2 \Big|_{t=0} = 0$$

which together with some simple computation implies

$$\langle \boldsymbol{A} - \hat{\boldsymbol{X}}, \hat{\boldsymbol{X}}\boldsymbol{W}\rangle = 0$$

i. e.,

$$\langle \hat{\boldsymbol{X}}^{\mathrm{T}}(\boldsymbol{A} - \hat{\boldsymbol{X}}), \boldsymbol{W}\rangle = 0 \tag{3}$$

Moreover, since $\boldsymbol{Z} - \frac{1}{n}\mathrm{tr}(\boldsymbol{Z})\boldsymbol{I}_n \in s(n)$ for any $\boldsymbol{Z} \in \mathbf{R}^{n\times n}$, it follows from (3) that

$$\langle \hat{\boldsymbol{X}}^{\mathrm{T}}(\boldsymbol{A} - \hat{\boldsymbol{X}}), \boldsymbol{Z} - \frac{1}{n}\mathrm{tr}(\boldsymbol{Z})\boldsymbol{I}_n\rangle = 0$$

i. e.,

$$\langle \hat{\boldsymbol{X}}^{\mathrm{T}}(\boldsymbol{A} - \hat{\boldsymbol{X}}) - \frac{1}{n}\mathrm{tr}(\hat{\boldsymbol{X}}^{\mathrm{T}}(\boldsymbol{A} - \hat{\boldsymbol{X}}))\boldsymbol{I}_n, \boldsymbol{Z}\rangle = 0$$

which yields asserted result due to the arbitrariness of $\boldsymbol{Z}$.

We next consider the solution to the matrix equation (2). To avoid unnecessary complexity, we assume in what follows that $\boldsymbol{A}$ is nonsingular, and so admits the singular value decomposition[10]:

$$\boldsymbol{A} = \boldsymbol{P}_1\boldsymbol{\Lambda}\boldsymbol{P}_2 \tag{4}$$

where $\boldsymbol{P}_1, \boldsymbol{P}_2 \in \boldsymbol{O}(n)$; $\boldsymbol{\Lambda} = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_n)$, $\mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_n > 0$; $\boldsymbol{O}(n)$ denotes the set of all $n \times n$ orthogonal matrices, and $\mathrm{diag}(\tau_1, \tau_2, \ldots, \tau_n)$ is a diagonal matrix with the i-th diagonal entry τ_i.

Lemma 2 Let the $n \times n$ matrix $\boldsymbol{A}$ be nonsingular with the singular value decomposition (4). Then

Received 2004-06-02.

Biography: Mu Jianfei (1961—), male, associate professor, jfmu@sjtu.edu.cn.

there exist some constants $\tau_1, \tau_2, \ldots, \tau_n$ such that

$$\boldsymbol{X} = \boldsymbol{P}_1 \operatorname{diag}(\tau_1, \tau_2, \ldots, \tau_n)\ \boldsymbol{P}_2 \tag{5}$$

solves the matrix equation (2). Moreover, each solution of Eq. (2) must be described in this form by means of a proper singular value decomposition of $\boldsymbol{A}$.

Proof Let $\boldsymbol{X}$ be a solution of Eq. (2). It follows from Eq. (2) that $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}$ is symmetric,

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{X} \tag{6}$$

Since the singular values $\mu_1, \mu_2, \ldots, \mu_n$ of $\boldsymbol{A}$ may be equal, we rewrite the singular value decomposition (4) in block form for later use.

$$\boldsymbol{A} = \boldsymbol{P}_1 \boldsymbol{\Lambda} \boldsymbol{P}_2 \tag{7}$$

where $\boldsymbol{P}_1, \boldsymbol{P}_2 \in \boldsymbol{O}(n)$, $\boldsymbol{\Lambda} = \operatorname{diag}(\sigma_1 \boldsymbol{I}_{n_1}, \sigma_2 \boldsymbol{I}_{n_2}, \ldots, \sigma_k \boldsymbol{I}_{n_k})$, $\sigma_1 > \sigma_2 > \ldots > \sigma_k > 0$.

Let $\boldsymbol{X}_1 = \boldsymbol{P}_1{}^{\mathrm{T}} \boldsymbol{X} \boldsymbol{P}_2^{\mathrm{T}}$. Then Eq. (6) can be recast as

$$(\boldsymbol{\Lambda}\boldsymbol{X}_1)^{\mathrm{T}} = \boldsymbol{\Lambda}\boldsymbol{X}_1$$

which indicates $\boldsymbol{\Lambda}\boldsymbol{X}_1 = \boldsymbol{V}$ is symmetric, hence Eq. (2) becomes

$$\boldsymbol{V}\boldsymbol{\Lambda}^{-2}\boldsymbol{V} - \boldsymbol{V} = \lambda \boldsymbol{I}_n$$

or equivalently,

$$\boldsymbol{\Lambda}^{-2}\boldsymbol{V} = \boldsymbol{I}_n + \lambda \boldsymbol{V}^{-1}$$

So we know $\boldsymbol{\Lambda}^{-2}\boldsymbol{V}$ is symmetric, which gives

$$\boldsymbol{\Lambda}^{-2}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda}^{-2} \tag{8}$$

Now let us partition the matrix $\boldsymbol{V}$ in the following block form compatible with the block structure of $\boldsymbol{\Lambda}$:

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{V}_{11} & \cdots & \boldsymbol{V}_{1k} \\ \vdots & & \vdots \\ \boldsymbol{V}_{k1} & \cdots & \boldsymbol{V}_{kk} \end{bmatrix} \qquad \boldsymbol{V}_{ij} \in \mathbf{R}^{n_i \times n_j}$$

Then Eq. (8) is equivalent to

$$(\sigma_i^{-2} - \sigma_j^{-2})\ \boldsymbol{V}_{ij} = \boldsymbol{0} \qquad i, j = 1, 2, \ldots, k$$

which leads to

$$\boldsymbol{V}_{ij} = \boldsymbol{0} \qquad i \neq j$$

$$\boldsymbol{V}_{ii} = \boldsymbol{V}_{ii}^{\mathrm{T}} \in \mathbf{R}^{n_i \times n_i} \qquad 1 \leqslant i \leqslant k$$

On the other hand, since $\boldsymbol{V}_{ii}$ is symmetric, we have the spectral decomposition:

$$\boldsymbol{V}_{ii} = \boldsymbol{Q}_{ii}\boldsymbol{\Delta}_{ii}\boldsymbol{Q}_{ii}{}^{\mathrm{T}} \qquad \boldsymbol{Q}_{ii} \in \boldsymbol{O}(n_i)$$

where $\boldsymbol{\Delta}_{ii}$ is diagonal. We define

$$\boldsymbol{Q}_1 \equiv \boldsymbol{P}_1 \begin{bmatrix} \boldsymbol{Q}_{11} & & \\ & \ddots & \\ & & \boldsymbol{Q}_{kk} \end{bmatrix}$$

$$\boldsymbol{Q}_2 \equiv \begin{bmatrix} \boldsymbol{Q}_{11} & & \\ & \ddots & \\ & & \boldsymbol{Q}_{kk} \end{bmatrix}^{\mathrm{T}} \boldsymbol{P}_2$$

and

$$\boldsymbol{\Delta} \equiv \begin{bmatrix} \sigma_1^{-1}\boldsymbol{\Delta}_{11} & & \\ & \ddots & \\ & & \sigma_k^{-1}\boldsymbol{\Delta}_{kk} \end{bmatrix}$$

Then, it is clear that $\boldsymbol{Q}_1, \boldsymbol{Q}_2 \in \boldsymbol{O}(n)$,

$$\boldsymbol{A} = \boldsymbol{P}_1\boldsymbol{\Lambda}\boldsymbol{P}_2 = \boldsymbol{Q}_1\boldsymbol{\Lambda}\boldsymbol{Q}_2$$

and

$$\boldsymbol{X} = \boldsymbol{P}_1\boldsymbol{X}_1\boldsymbol{P}_2 = \boldsymbol{P}_1\boldsymbol{\Lambda}^{-1}\boldsymbol{V}\boldsymbol{P}_2 =$$

$$\boldsymbol{P}_1\boldsymbol{\Lambda}^{-1} \begin{bmatrix} \boldsymbol{Q}_{11}\boldsymbol{\Delta}_{11}\boldsymbol{Q}_{11}^{\mathrm{T}} & & \\ & \ddots & \\ & & \boldsymbol{Q}_{kk}\boldsymbol{\Delta}_{kk}\boldsymbol{Q}_{kk}^{\mathrm{T}} \end{bmatrix} \boldsymbol{P}_2 = \boldsymbol{Q}_1\boldsymbol{\Delta}\boldsymbol{Q}_2$$

Now, if $\boldsymbol{A}$ and $\boldsymbol{X}$ are written in the above forms, we can easily see that Eq. (2) is equivalent to $\boldsymbol{\Delta}^2 - \boldsymbol{\Delta}\boldsymbol{\Lambda} = \lambda \boldsymbol{I}_n$. Hence, $\boldsymbol{X} = \boldsymbol{P}_1\boldsymbol{\Delta}\ \boldsymbol{P}_2$ also solves Eq. (2) if $\boldsymbol{\Delta}$ meets the previous equation. Note also that $\boldsymbol{A} = \boldsymbol{Q}_1\boldsymbol{\Lambda}\boldsymbol{Q}_2$ is a singular value decomposition of $\boldsymbol{A}$. Therefore, each solution of Eq. (2) can be described in the form Eq. (5) by means of a proper singular value decomposition of $\boldsymbol{A}$.

2 Further Discussions

It follows from lemma 1 and lemma 2 that each solution $\hat{\boldsymbol{X}}$ of the constrained optimization problem (1) can always be expressed in the form (5). Therefore, in order to solve problem (1), it suffices to evaluate all the values of $\|\boldsymbol{A} - \boldsymbol{X}\|_{\mathrm{F}}^2$ with $\boldsymbol{X}$ given by Eq. (5). In this case, we have by the basic property of Frobenius norm that

$$g(\tau) \equiv \|\boldsymbol{A} - \boldsymbol{X}\|_{\mathrm{F}}^2 = \|\boldsymbol{\Lambda} - \operatorname{diag}(\tau_1, \tau_2, \ldots, \tau_n)\|_{\mathrm{F}}^2 = \sum_{i=1}^{n} (\mu_i - \tau_i)^2 \tag{9}$$

where, as shown in lemma 2, $\{\mu_i\}_{i=1}^n$ are the singular values of $\boldsymbol{A}$ enumerated in descending order. Moreover, the constraint condition $\det(\boldsymbol{X}) = 1$ amounts to

$$\prod_{i=1}^{n} \tau_i = \det(\boldsymbol{Q}_1\boldsymbol{Q}_2) = \operatorname{sgn}(\det(\boldsymbol{A})) \tag{10}$$

where $\operatorname{sgn}(\cdot)$ is a sign function defined by

$$\operatorname{sgn}(x) = \begin{cases} 1 & x \geqslant 0 \\ -1 & x < 0 \end{cases}$$

Thus problem (1) can be reformulated as

$$\begin{aligned} &\min_{\tau}\ g(\tau) \\ &\text{subject to } \prod_{i=1}^{n} \tau_i = \operatorname{sgn}(\det(\boldsymbol{A})) \end{aligned} \tag{11}$$

Lemma 3 Let $\hat{\tau} = \{\hat{\tau}_1, \hat{\tau}_2, \ldots, \hat{\tau}_n\}$ solve the constrained optimization problem (11). Then there exists some real constant λ such that

$$\hat{\tau}_i = \frac{\mu_i \pm \sqrt{\mu_i^2 + \lambda}}{2} \qquad i = 1, 2, \ldots, n \tag{12}$$

Proof By the method of Lagrange multipliers, there exists some real constant $\frac{\operatorname{sgn}(\det(\boldsymbol{A}))}{2}\lambda$ such that $\hat{\tau}$ satisfies

$$\partial_{\tau_i}\left[\sum_{i=1}^{n}(\mu_i - \tau_i)^2 - \frac{\operatorname{sgn}(\det(\boldsymbol{A}))}{2}\lambda \cdot \left(\prod_{i=1}^{n}\tau_i - \operatorname{sgn}(\det(\boldsymbol{A}))\right)\right] = 0 \qquad i = 1, 2, \ldots, n$$

or equivalently,

$$\tau_i(\mu_i - \tau_i) = -\frac{\mathrm{sgn}(\det(\boldsymbol{A}))}{4}\lambda\prod_{i=1}^{n}\tau_i = -\frac{\lambda}{4}$$

which yields the desired expression (12) immediately.

Lemma 3 provides an important property characterizing the solution of problem (11), but there are still $2n$ possible combinations, each of which may result in the desired solution. The next result is given to simplify this difficulty.

Lemma 4 Let $\hat{\tau} = \{\hat{\tau}_1, \hat{\tau}_2, \ldots, \hat{\tau}_n\}$ solve the constrained optimization problem (11). Then the following statements hold:

① $\hat{\tau}_1 \geqslant \hat{\tau}_2 \geqslant \ldots \geqslant \hat{\tau}_n$;

② If there exists some $\hat{\tau}_i$ such that $\hat{\tau}_i = \frac{\mu_i - \sqrt{\mu_i^2 + \lambda}}{2}$, then for every $j \geqslant i$, $\mu_j^2 + \lambda > 0$, we have

$$\hat{\tau}_j = \frac{\mu_j - \sqrt{\mu_j^2 + \lambda}}{2}$$

③ There do not exist two different singular values $\mu_i > \mu_j (i < j)$ such that

$$\hat{\tau}_i = \frac{\mu_i - \sqrt{\mu_i^2 + \lambda}}{2}, \quad \hat{\tau}_j = \frac{\mu_j - \sqrt{\mu_j^2 + \lambda}}{2}$$

Proof

1) If statement ① is not true, then there exist at least two components $\hat{\tau}_i, \hat{\tau}_j (i < j)$ of $\hat{\tau}$, such that $\hat{\tau}_i < \hat{\tau}_j$. We then exchange the values of these two components to obtain a new admissible choice $\tilde{\tau}$. It is easy to see that $g(\tilde{\tau}) < g(\hat{\tau})$, which leads to a contradiction. The desired result then follows.

2) If statement ② is not true, then there exists some $j_0 > i$ such that $\mu_{j_0}^2 + \lambda > 0$ and

$$\hat{\tau}_{j_0} = \frac{\mu_{j_0} + \sqrt{\mu_{j_0}^2 + \lambda}}{2}$$

Thus the above statement gives

$$\hat{\tau}_i \geqslant \hat{\tau}_{j_0}$$

i. e. ,

$$\mu_i - \mu_{j_0} \geqslant \sqrt{\mu_i^2 + \lambda} + \sqrt{\mu_{j_0}^2 + \lambda}$$

or equivalently,

$$-(\lambda + \mu_i\mu_{j_0}) \geqslant \sqrt{\mu_i^2 + \lambda}\sqrt{\mu_{j_0}^2 + \lambda}$$

That means $\lambda + \mu_i\mu_{j_0} \leqslant 0$, which together with $\mu_{j_0} \leqslant \mu_i$ yields

$$\lambda + \mu_{j_0}^2 \leqslant \lambda + \mu_i\mu_{j_0} \leqslant 0$$

We then obtain a contradiction. This completes the proof of statement ② of lemma 4.

3) If statement ③ is not true, then there exist two different singular values $\mu_{i_0} > \mu_{j_0} (i_0 < j_0)$ such that

$$\hat{\tau}_{i_0} = \frac{\mu_{i_0} - \sqrt{\mu_{i_0}^2 + \lambda}}{2}, \hat{\tau}_{j_0} = \frac{\mu_{j_0} - \sqrt{\mu_{j_0}^2 + \lambda}}{2}$$

Thus it follows from statement ① just proved that $\hat{\tau}_{i_0} \geqslant \hat{\tau}_{j_0}$, which results in $\lambda \geqslant 0$ by some simple calculation. When $\lambda = 0$, $\hat{\tau}_{i_0} = 0$ which contradicts the admissible condition (10). When $\lambda > 0$, it is clear that $\hat{\tau}_{i_0} < 0$, $\hat{\tau}_{j_0} < 0$. We then change the components of $\hat{\tau}$ at the positions i_0 and j_0 into $-\hat{\tau}_{i_0}$, $-\hat{\tau}_{j_0}$, respectively, and that gives us a new admissible choice $\tilde{\tau}$. It is evident that $g(\tilde{\tau}) < g(\hat{\tau})$. This is a contradiction. The desired result is therefore verified.

By virtue of lemmas 1 to 4, we can obtain the following result directly.

Theorem 1 Let the nonsingular matrix $\boldsymbol{A}$ admit the singular value decomposition (4) and assume that its smallest singular value has multiplicity n_0, i. e. , $\mu_{n-n_0+1} = \ldots = \mu_n$. Then there exists some nonnegative integer k ($0 \leqslant k \leqslant n_0$) and some real constant λ satisfying the equation

$$\prod_{i=1}^{n-k}\frac{\mu_i + \sqrt{\mu_i^2 + \lambda}}{2}\left(\frac{\mu_n - \sqrt{\mu_n^2 + \lambda}}{2}\right)^k = \mathrm{sgn}(\det(\boldsymbol{A})) \tag{13}$$

such that the following matrix solves problem (1):

$$\hat{\boldsymbol{X}} = \boldsymbol{P}_1\boldsymbol{D}\boldsymbol{P}_2$$

with

$$\boldsymbol{D} = \mathrm{diag}\left(\frac{\mu_1 + \sqrt{\mu_1^2 + \lambda}}{2}, \ldots, \frac{\mu_{n-k} + \sqrt{\mu_{n-k}^2 + \lambda}}{2}, \frac{\mu_n - \sqrt{\mu_n^2 + \lambda}}{2}, \ldots, \frac{\mu_n - \sqrt{\mu_n^2 + \lambda}}{2}\right)$$

Proof According to lemma 1 and lemma 2, problem (1) can be converted into problem (11), the solution of which should take the form (12) by lemma 3. Furthermore, by lemma 4 only the following combinations may solve problem (11):

$$\tau_1 = \frac{\mu_1 + \sqrt{\mu_1^2 + \lambda}}{2}, \ldots, \tau_{n-l} = \frac{\mu_{n-l} + \sqrt{\mu_{n-l}^2 + \lambda}}{2},$$

$$\tau_{n-l+1} = \frac{\mu_n - \sqrt{\mu_n^2 + \lambda}}{2}, \ldots, \tau_n = \frac{\mu_n - \sqrt{\mu_n^2 + \lambda}}{2}$$

where $0 \leqslant l \leqslant n_0$. We then obtain the result of theorem 1 easily.

3 Numerical Results

We first transfer Eq. (13) into

$$f(\lambda) \equiv \prod_{i=1}^{n-k}\frac{\mu_i + \sqrt{\mu_i^2 + \lambda}}{2}\left(\frac{\sqrt{\mu_n^2 + \lambda} - \mu_n}{2}\right)^k - (-1)^k\mathrm{sgn}(\det(\boldsymbol{A})) = 0 \tag{14}$$

It is clear that $f(\lambda)$ is strictly increasing as $\lambda \in (0, +\infty)$. Hence, only if $f(0) < 0$, Eq. (14) has one (only one) solution in $(0, +\infty)$.

According to theorem 1, we then have the following algorithm for solving problem (1):

① Compute the singular value decomposition of $\boldsymbol{A}$ to obtain $\boldsymbol{P}_1, \boldsymbol{P}_2$ and $\boldsymbol{\Lambda}$ used in Eq. (4).

② Let n_0 be the multiplicity of the smallest singular value of $\boldsymbol{A}$. For $k = 0, 1, \ldots, n_0$, compute the whole set of solutions of Eq. (14) in the interval $[-\mu_n^2, 0]$; if $f(0) < 0$, then also compute the unique solution of Eq. (14) in $(0, +\infty)$.

③ For each λ obtained above, compute the related τ by the formula given in theorem 1, and then compute $g(\tau)$ by Eq. (9). Determine those λ which correspond to the minimal value of $g(\tau)$ just obtained.

④ For each of these τ, compute $\hat{\boldsymbol{X}} = \boldsymbol{P}_1 \boldsymbol{D} \boldsymbol{P}_2$, where $\boldsymbol{D}$ is a diagonal matrix with the i-th entry taking the value τ_i, the i-th component of τ. Then $\hat{\boldsymbol{X}}$ solves problem (1).

We next use the algorithm to solve a concrete problem. We randomly generate a matrix as follows:

$$\boldsymbol{A} = \begin{bmatrix} 0.846\,2 & 0.672\,1 & 0.681\,3 \\ 0.525\,2 & 0.838\,1 & 0.379\,5 \\ 0.202\,6 & 0.019\,6 & 0.831\,8 \end{bmatrix}$$

The related singular value decomposition is

$$\boldsymbol{P}_1 = \begin{bmatrix} 0.733\,6 & -0.055\,7 & -0.677\,3 \\ 0.579\,0 & -0.470\,6 & 0.665\,8 \\ 0.355\,8 & 0.880\,6 & 0.313\,0 \end{bmatrix}$$

$$\boldsymbol{P}_2 = \begin{bmatrix} 0.576\,0 & 0.569\,3 & 0.586\,7 \\ -0.172\,4 & -0.617\,0 & 0.767\,9 \\ -0.799\,1 & 0.543\,4 & 0.257\,2 \end{bmatrix}$$

$$\boldsymbol{\Lambda} = \mathrm{diag}(1.730\,9, 0.671\,9, 0.200\,4)$$

In this case, $n_0 = 1$. By computation we also find that only as $k = 0$ Eq. (14) is solvable and $\lambda = 0.908\,2$, $\tau = \{1.853\,4, 0.919\,0, 0.587\,1\}$, $g(\tau) = 0.225\,6$. Thus

$$\hat{\boldsymbol{X}} = \boldsymbol{P}_1 \boldsymbol{D} \boldsymbol{P}_2 = \begin{bmatrix} 1.109\,7 & 0.589\,4 & 0.656\,1 \\ 0.380\,3 & 1.090\,2 & 0.398\,4 \\ 0.093\,5 & -0.024\,0 & 1.055\,6 \end{bmatrix}$$

solves problem (1) with the above matrix $\boldsymbol{A}$.

References

[1] Warner F W. *Foundations of differentiable manifolds and Lie groups* [M]. Berlin: Springer, 1983.

[2] Andersson L E, Elfving T. A constrained procrustes problem [J]., *SIAM J Matrix Anal Appl*, 1997, **18**: 124 – 139.

[3] Arun K S. A unitarily constrained total least squares problem in signal processing [J]. *SIAM J Matrix Anal Appl*, 1992, **13**(3): 729 – 745.

[4] Higham N J. Computing a nearest symmetric positive semidefinite matrix [J]. *Linear Algebra and Its Applications*, 1988, **103**: 103 – 118.

[5] Higham N J. The symmetric procrustes problem [J]. *BIT*, 1988, **28**(1): 133 – 143.

[6] Umeyama S. Least-squares estimation of transformation parameters between two point patterns [J]. *IEEE Trans Pattern Anal Mach Intelligence*, 1991, **13**(4): 376 – 380.

[7] Zietak K. On approximation problems with zero-trace matrices [J]. *Linear Algebra and Its Applications*, 1996, **247**: 169 – 183.

[8] Arnold V I. *Mathematical methods of classical mechanics* [M]. Berlin: Springer, 1978.

[9] Cyganski D, Orr J A. Applications of tensor theory to object recognition and orientation determination [J]. *IEEE Trans Pattern Anal Mach Intelligence*, 1985, **7**(6): 663 – 673.

[10] Golub G H, van Loan C F. *Matrix computations* [M]. Baltimore: The Johns Hopkins University Press, 1989.

求解最近保体矩阵问题的若干研究

沐建飞　　黄建国

(上海交通大学数学系, 上海 200240)

摘要: 研究了求解给定矩阵的最近保体矩阵问题. 首先导出该问题解所必须满足的一个矩阵方程, 然后用奇异值分解方法求解该矩阵方程; 并获得了该问题解的其他更进一步的刻画条件. 利用这些结果建立了一个求解算法, 并通过数值算例说明了该算法的有效性.

关键词: 保体矩阵; 矩阵拟合问题; 奇异值分解

中图分类号: O241.1