

An iterative method for the skew-symmetric solution and the optimal approximate solution of the matrix equation $AXB = C$ [☆]

Guang-Xin Huang^{a,*}, Feng Yin^{b,*}, Ke Guo^a

^aCollege of Information and Management, Chengdu University of Technology, Chengdu 610059, PR China

^bDepartment of Mathematics, Sichuan University of Science and Engineering, Zigong 643000, PR China

Received 27 May 2006; received in revised form 24 August 2006

Abstract

In this paper, an iterative method is constructed to solve the linear matrix equation $AXB = C$ over skew-symmetric matrix X . By the iterative method, the solvability of the equation $AXB = C$ over skew-symmetric matrix can be determined automatically. When the equation $AXB = C$ is consistent over skew-symmetric matrix X , for any skew-symmetric initial iterative matrix X_1 , the skew-symmetric solution can be obtained within finite iterative steps in the absence of roundoff errors. The unique least-norm skew-symmetric iterative solution of $AXB = C$ can be derived when an appropriate initial iterative matrix is chosen. A sufficient and necessary condition for whether the equation $AXB = C$ is inconsistent is given. Furthermore, the optimal approximate solution of $AXB = C$ for a given matrix X_0 can be derived by finding the least-norm skew-symmetric solution of a new corresponding matrix equation $A\tilde{X}B = \tilde{C}$. Finally, several numerical examples are given to illustrate that our iterative method is effective.

© 2006 Elsevier B.V. All rights reserved.

MSC: 65F15; 65F20

Keywords: Iterative method; Skew-symmetric solution; Least-norm skew-symmetric solution; Optimal approximate solution

1. Introduction

In this paper the following notations are considered and used. Let $R^{m \times n}$ denote the set of all $m \times n$ real matrices, $SSR^{n \times n}$ denote the set of all $n \times n$ real skew-symmetric matrices in $R^{n \times n}$. We denote by the superscripts T and + the transpose and Moore–Penrose generalized inverse of matrices, respectively. In matrix space $R^{m \times n}$, define inner product as: $\langle A, B \rangle = \text{tr}(B^T A)$ for all $A, B \in R^{m \times n}$, $\|A\|$ represents the Frobenius norm of A . $R(A)$ represents the column space of A . $\text{vec}(\cdot)$ represents the vector operator. i.e., $\text{vec}(A) = (a_1^T, a_2^T, \dots, a_n^T)^T \in R^{mn}$ for the matrix $A = (a_1, a_2, \dots, a_n) \in R^{m \times n}$, $a_i \in R^m$, $i = 1, 2, \dots, n$. $A \otimes B$ stands for the Kronecker product of matrices A and B . For more notations we refer to [3].

In this paper, we will consider the following two problems.

[☆] This work is supported by the Key Science Fund and the Youngth Foundation of Chengdu University of Technology.

* Corresponding authors.

E-mail addresses: huangx@cdut.edu.cn (G.-X. Huang), fyin@suse.edu.cn (F. Yin).

Problem I. For given matrices $A \in R^{m \times n}$, $B \in R^{n \times p}$, $C \in R^{m \times p}$, find matrix $X \in SSR^{n \times n}$ such that

$$AXB = C. \tag{1}$$

Problem II. When Problem I is consistent, let S_E denote the set of skew-symmetric solutions of Problem I, for a given matrix $X_0 \in R^{n \times n}$, find $\hat{X} \in S_E$ such that

$$\|\hat{X} - X_0\| = \min_{X \in S_E} \|X - X_0\|. \tag{2}$$

In fact, Problem II is to find the optimal approximately skew-symmetric solution to a given matrix $X_0 \in R^{n \times n}$.

Problem I has been considered in the case of special solution structures, e.g. symmetric, triangular or diagonal solution X . We refer to Dai [5], Eric Chu [2], Don [4], Magnus [6], Morris and Odell [8], Bjerhammer [1] for more details. Mitra [7] considered common solutions to a pair of linear matrix equations $A_1XB_1 = C_1$, $A_2XB_2 = C_2$. In these literatures, the problem was discussed by using matrices decomposition such as the singular value decomposition (SVD), the generalized SVD (GSVD), the quotient SVD (QSVD) and the canonical correlation decomposition (CCD). However, these methods are difficult to be applied to solve Problem II, and the representation of solutions of Problems I and II are complicated. Huang and Yin [11,12] recently solve the constrained inverse eigenproblem and associated approximation problem for anti-Hermitian R -symmetric matrices and the matrix inverse problem and its optimal approximation problem for R -symmetric matrices. Recently, Peng et al. [9] have constructed an iteration method to solve the linear matrix equation $AXB = C$ over symmetric matrix X . Peng [10] has shown an iterative method to solve the minimum Frobenius norm residual problem: $\min \|AXB - C\|$ with unknown symmetric matrix X . In this paper, we will consider the skew-symmetric solution of the linear matrix equation $AXB = C$, we construct an iterative method by which the solvability of Problem I can be determined automatically, the solution can be obtained within finite iterative steps when Problem I is consistent, and the solution of Problem II can be obtained by finding the least-norm skew-symmetric solution of a new matrix equation $A\tilde{X}B = \tilde{C}$.

This paper is organized as follows. In Section 2, we will solve Problem I by construct an iterative method, i.e., for an arbitrary initial matrix $X_1 \in SSR^{n \times n}$, if there exists a positive integer k , such that $R_k \neq \mathbf{0}$ and $Q_k = \mathbf{0}$, then Problem I is inconsistent, where R_k and Q_k ($k = 1, 2, \dots$) are defined in Algorithm 2.1. If Problem I is consistent, then for an arbitrary initial matrix $X_1 \in SSR^{n \times n}$, we can obtain a solution $\bar{X} \in SSR^{n \times n}$ of Problem I within finite iterative steps in the absence of roundoff errors. Let $X_1 = A^T H^T B^T - BHA$, where H is an arbitrary matrix in $R^{p \times m}$, or more especially, let $X_1 = \mathbf{0} \in SSR^{n \times n}$, we can obtain the unique least-norm solution of Problem I. Then in Section 3, we give the optimal approximate solution of Problem II by finding the least-norm skew-symmetric solution of a corresponding new matrix equation. In Section 4, several numerical examples are given to illustrate the application of our iterative methods.

2. The solution of Problem I

In this section, we will first introduce an iterative method to solve Problem I, then prove that it is convergent.

Algorithm 2.1. Step 1: Input matrices $A \in R^{m \times n}$, $B \in R^{n \times p}$, $C \in R^{m \times p}$;

Step 2: Choose an arbitrary matrix $X_1 \in SSR^{n \times n}$, compute

$$R_1 = C - AX_1B,$$

$$P_1 = A^T R_1 B^T,$$

$$Q_1 = \frac{1}{2}(P_1 - P_1^T),$$

$$k := 1;$$

Step 3: If $R_1 = \mathbf{0}$, or $R_1 \neq \mathbf{0}$ and $Q_1 = \mathbf{0}$, then stop; Else go to step 4;

Step 4: Compute

$$X_{k+1} = X_k + \frac{\|R_k\|^2}{\|Q_k\|^2} Q_k,$$

$$R_{k+1} = C - AX_{k+1}B,$$

$$P_{k+1} = A^T R_{k+1} B^T,$$

$$Q_{k+1} = \frac{1}{2}(P_{k+1} - P_{k+1}^T) + \frac{\text{tr}(P_{k+1}Q_k)}{\|Q_k\|^2} Q_k;$$

Step 5: If $R_{k+1} = \mathbf{0}$, or $R_{k+1} \neq \mathbf{0}$ and $Q_{k+1} = \mathbf{0}$, then stop; Else, let $k := k + 1$, go to step 4.

Obviously, it can be seen that $Q_i \in SSR^{n \times n}$, $X_i \in SSR^{n \times n}$, where $i = 1, 2, \dots$.

Lemma 2.1. For the sequences $\{R_i\}$, $\{P_i\}$ and $\{Q_i\}$ generated in Algorithm 2.1, we have

$$\text{tr}(R_{i+1}^T R_j) = \text{tr}(R_i^T R_j) + \frac{\|R_i\|^2}{\|Q_i\|^2} \text{tr}(Q_i P_j). \tag{3}$$

Proof. Noting that $Q_i^T = -Q_i$, by Algorithm 2.1, we have

$$\begin{aligned} \text{tr}(R_{i+1}^T R_j) &= \text{tr}[(C - AX_{i+1}B)^T R_j] \\ &= \text{tr} \left[\left(C - A \left(X_i + \frac{\|R_i\|^2}{\|Q_i\|^2} Q_i \right) B \right)^T R_j \right] \\ &= \text{tr} \left[(C - AX_i B)^T R_j - \frac{\|R_i\|^2}{\|Q_i\|^2} (A Q_i B)^T R_j \right] \\ &= \text{tr}(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} \text{tr}(B^T Q_i^T A^T R_j) \\ &= \text{tr}(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} \text{tr}(Q_i^T A^T R_j B^T) \\ &= \text{tr}(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} \text{tr}(Q_i^T P_j) \\ &= \text{tr}(R_i^T R_j) + \frac{\|R_i\|^2}{\|Q_i\|^2} \text{tr}(Q_i P_j). \quad \square \end{aligned}$$

Lemma 2.2. For the sequences $\{R_i\}$ and $\{Q_i\}$ generated by Algorithm 2.1, and $k \geq 2$, we have

$$\text{tr}(R_i^T R_j) = 0, \quad \text{tr}(Q_i^T Q_j) = 0, \quad i, j = 1, 2, \dots, k, i \neq j. \tag{4}$$

Proof. Since $\text{tr}(R_i^T R_j) = \text{tr}(R_j^T R_i)$ and $\text{tr}(Q_i^T Q_j) = \text{tr}(Q_j^T Q_i)$ for all $i, j = 1, 2, \dots, k$, we only need prove that $\text{tr}(R_i^T R_j) = 0, \text{tr}(Q_i^T Q_j) = 0$ for all $1 \leq j < i \leq k$. We prove the conclusion by induction and two steps are required.

Step 1: we will show that

$$\text{tr}(R_{i+1}^T R_i) = 0, \quad \text{tr}(Q_{i+1}^T Q_i) = 0, \quad i = 1, 2, \dots, k - 1. \tag{5}$$

To prove this conclusion, we also use induction.

For $i = 1$, by Lemma 2.1 and Algorithm 2.1, noting that $Q_1^T = -Q_1$, it follows that

$$\begin{aligned}
 \operatorname{tr}(R_2^T R_1) &= \operatorname{tr}(R_1^T R_1) + \frac{\|R_1\|^2}{\|Q_1\|^2} \operatorname{tr}(Q_1 P_1) \\
 &= \|R_1\|^2 + \frac{\|R_1\|^2}{\|Q_1\|^2} \operatorname{tr} \left(\frac{Q_1 P_1}{2} + \frac{(Q_1 P_1)^T}{2} \right) \\
 &= \|R_1\|^2 + \frac{\|R_1\|^2}{\|Q_1\|^2} \operatorname{tr} \left(\frac{Q_1 P_1}{2} + \frac{P_1^T Q_1^T}{2} \right) \\
 &= \|R_1\|^2 + \frac{\|R_1\|^2}{\|Q_1\|^2} \operatorname{tr} \left(Q_1 \frac{P_1 - P_1^T}{2} \right) \\
 &= \|R_1\|^2 + \frac{\|R_1\|^2}{\|Q_1\|^2} \operatorname{tr}(Q_1 Q_1) \\
 &= \|R_1\|^2 - \frac{\|R_1\|^2}{\|Q_1\|^2} \operatorname{tr}(Q_1^T Q_1) = 0
 \end{aligned}$$

and

$$\begin{aligned}
 \operatorname{tr}(Q_2^T Q_1) &= \operatorname{tr} \left[\left(\frac{P_2 - P_2^T}{2} + \frac{\operatorname{tr}(P_2 Q_1)}{\|Q_1\|^2} Q_1 \right)^T Q_1 \right] \\
 &= -\operatorname{tr} \left(\frac{P_2 - P_2^T}{2} Q_1 \right) + \frac{\operatorname{tr}(P_2 Q_1)}{\|Q_1\|^2} \operatorname{tr}(Q_1^T Q_1) \\
 &= -\operatorname{tr} \left(\frac{P_2 Q_1 + (Q_1 P_2)^T}{2} \right) + \operatorname{tr}(P_2 Q_1) \\
 &= -\operatorname{tr} \left(\frac{P_2 Q_1 + P_2 Q_1}{2} \right) + \operatorname{tr}(P_2 Q_1) = 0.
 \end{aligned}$$

Assume (5) holds for $i = s - 1$, i.e., $\operatorname{tr}(R_s^T R_{s-1}) = 0$, $\operatorname{tr}(Q_s^T Q_{s-1}) = 0$. Noting that $Q_s^T = -Q_s$, by Lemma 2.1 we have

$$\begin{aligned}
 \operatorname{tr}(R_{s+1}^T R_s) &= \operatorname{tr}(R_s^T R_s) + \frac{\|R_s\|^2}{\|Q_s\|^2} \operatorname{tr}(Q_s P_s) \\
 &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \operatorname{tr}(Q_s^T P_s) \\
 &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \operatorname{tr} \left[\frac{Q_s^T P_s + (P_s Q_s^T)^T}{2} \right] \\
 &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \operatorname{tr} \left(Q_s^T \frac{P_s - P_s^T}{2} \right) \\
 &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \operatorname{tr} \left[Q_s^T \left(Q_s - \frac{\operatorname{tr}(P_s Q_{s-1})}{\|Q_{s-1}\|^2} Q_{s-1} \right) \right] \\
 &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \operatorname{tr} \left[\left(Q_s^T Q_s \right) - \frac{\operatorname{tr}(P_s Q_{s-1})}{\|Q_{s-1}\|^2} \operatorname{tr}(Q_s^T Q_{s-1}) \right] \\
 &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \operatorname{tr}(Q_s^T Q_s) = 0
 \end{aligned}$$

and

$$\begin{aligned}
 \text{tr}(Q_{s+1}^T Q_s) &= \text{tr} \left[\left(\frac{P_{s+1} - P_{s+1}^T}{2} + \frac{\text{tr}(P_{s+1} Q_s)}{\|Q_s\|^2} Q_s \right)^T Q_s \right] \\
 &= \text{tr} \left[\left(-\frac{P_{s+1} - P_{s+1}^T}{2} + \frac{\text{tr}(P_{s+1} Q_s)}{\|Q_s\|^2} Q_s^T \right) Q_s \right] \\
 &= \text{tr} \left(-\frac{P_{s+1} Q_s - P_{s+1}^T Q_s}{2} \right) + \frac{\text{tr}(P_{s+1} Q_s)}{\|Q_s\|^2} \text{tr}(Q_s^T Q_s) \\
 &= -\text{tr} \left[\frac{P_{s+1} Q_s + (Q_s P_{s+1})^T}{2} \right] + \text{tr}(P_{s+1} Q_s) \\
 &= -\text{tr} \left[\frac{P_{s+1} Q_s + (P_{s+1} Q_s)^T}{2} \right] + \text{tr}(P_{s+1} Q_s) \\
 &= -\text{tr}(P_{s+1} Q_s) + \text{tr}(P_{s+1} Q_s) = 0.
 \end{aligned}$$

Hence, (5) holds for $i = s$. Therefore, (5) holds by the principle of induction.

Step 2: Assume that $\text{tr}(R_s^T R_j) = 0, \text{tr}(Q_s^T Q_j) = 0, j = 1, 2, \dots, s - 1$, then we show that

$$\text{tr}(R_{s+1}^T R_j) = 0, \quad \text{tr}(Q_{s+1}^T Q_j) = 0, \quad j = 1, 2, \dots, s. \tag{6}$$

In fact, by Lemma 2.1 we have

$$\begin{aligned}
 \text{tr}(R_{s+1}^T R_j) &= \text{tr}(R_s^T R_j) + \frac{\|R_s\|^2}{\|Q_s\|^2} \text{tr}(Q_s P_j) \\
 &= \frac{\|R_s\|^2}{\|Q_s\|^2} \text{tr} \left[\frac{Q_s P_j}{2} + \frac{(Q_s P_j)^T}{2} \right] \\
 &= \frac{\|R_s\|^2}{\|Q_s\|^2} \text{tr} \left(Q_s \frac{P_j - P_j^T}{2} \right) \\
 &= \frac{\|R_s\|^2}{\|Q_s\|^2} \text{tr} \left[Q_s \left(Q_j - \frac{\text{tr}(P_j Q_{j-1})}{\|Q_{j-1}\|^2} Q_{j-1} \right) \right] \\
 &= \frac{\|R_s\|^2}{\|Q_s\|^2} \left[\text{tr}(Q_s Q_j) - \frac{\text{tr}(P_j Q_{j-1})}{\|Q_{j-1}\|^2} \text{tr}(Q_s Q_{j-1}) \right] \\
 &= \frac{\|R_s\|^2}{\|Q_s\|^2} \left[-\text{tr}(Q_s^T Q_j) + \frac{\text{tr}(P_j Q_{j-1})}{\|Q_{j-1}\|^2} \text{tr}(Q_s^T Q_{j-1}) \right] = 0
 \end{aligned}$$

and

$$\begin{aligned}
 \text{tr}(Q_{s+1}^T Q_j) &= \text{tr} \left[\left(\frac{P_{s+1} - P_{s+1}^T}{2} + \frac{\text{tr}(P_{s+1} Q_s)}{\|Q_s\|^2} Q_s \right)^T Q_j \right] \\
 &= \text{tr} \left[\left(-\frac{P_{s+1} - P_{s+1}^T}{2} + \frac{\text{tr}(P_{s+1} Q_s)}{\|Q_s\|^2} Q_s^T \right) Q_j \right] \\
 &= \text{tr} \left(-\frac{P_{s+1} - P_{s+1}^T}{2} Q_j \right) + \frac{\text{tr}(P_{s+1} Q_s)}{\|Q_s\|^2} \text{tr}(Q_s^T Q_j)
 \end{aligned}$$

$$\begin{aligned}
 &= -\frac{1}{2}\text{tr}(P_{s+1}Q_j) + \frac{1}{2}\text{tr}(P_{s+1}^T Q_j) \\
 &= -\frac{1}{2}\text{tr}(P_{s+1}Q_j) + \frac{1}{2}\text{tr}[(P_{s+1}^T Q_j)^T] \\
 &= -\frac{1}{2}\text{tr}(P_{s+1}Q_j) + \frac{1}{2}\text{tr}(Q_j^T P_{s+1}) \\
 &= -\frac{1}{2}\text{tr}(P_{s+1}Q_j) - \frac{1}{2}\text{tr}(Q_j P_{s+1}) \\
 &= -\text{tr}(P_{s+1}Q_j) \\
 &= -\text{tr}(Q_j P_{s+1}).
 \end{aligned}$$

Noting that $\text{tr}(R_{s+1}^T R_j) = 0$, $\text{tr}(R_{s+1}^T R_{j+1}) = 0$, by Lemma 2.1 we have

$$\begin{aligned}
 \text{tr}(Q_{s+1}^T Q_j) &= -\text{tr}(Q_j P_{s+1}) \\
 &= \frac{\|Q_j\|^2}{\|R_j\|^2} [\text{tr}(R_j^T R_{s+1}) - \text{tr}(R_{j+1}^T R_{s+1})] \\
 &= \frac{\|Q_j\|^2}{\|R_j\|^2} [\text{tr}(R_{s+1}^T R_j) - \text{tr}(R_{s+1}^T R_{j+1})] = 0.
 \end{aligned}$$

By the principle of induction, (6) holds. Noting that (4) is implied in steps 1 and 2 by the principle of induction, we complete the proof. \square

Lemma 2.3. Suppose \bar{X} be an arbitrary solution of Problem I, i.e., $A\bar{X}B = C$ and $\bar{X}^T = -\bar{X}$, then

$$\text{tr}[(\bar{X} - X_k)Q_k] = -\|R_k\|^2, \quad k = 1, 2, \dots, \tag{7}$$

where the sequences $\{X_k\}$, $\{R_k\}$, $\{Q_k\}$ are generated by Algorithm 2.1.

Proof. We proof the conclusion by induction.

For $k = 1$,

$$\begin{aligned}
 \text{tr}[(\bar{X} - X_1)Q_1] &= \text{tr}\left[(\bar{X} - X_1)\frac{P_1 - P_1^T}{2}\right] \\
 &= \text{tr}\left[\frac{(\bar{X} - X_1)P_1}{2} + \frac{(\bar{X} - X_1)^T P_1^T}{2}\right] \\
 &= \text{tr}\left[\frac{(\bar{X} - X_1)P_1}{2} + \frac{P_1(\bar{X} - X_1)}{2}\right] \\
 &= \text{tr}[(\bar{X} - X_1)P_1] \\
 &= \text{tr}[(\bar{X} - X_1)A^T(C - AX_1B)B^T] \\
 &= \text{tr}[B(C - AX_1B)^T A(\bar{X} - X_1)^T] \\
 &= -\text{tr}[B(C - AX_1B)^T A(\bar{X} - X_1)] \\
 &= -\text{tr}[A(\bar{X} - X_1)B(C - AX_1B)^T] \\
 &= -\text{tr}[(A\bar{X}B - AX_1B)R_1^T] \\
 &= -\text{tr}[(C - AX_1B)R_1^T] \\
 &= -\text{tr}(R_1 R_1^T) = -\text{tr}(R_1^T R_1) = -\|R_1\|^2.
 \end{aligned}$$

Assume (7) holds for $k = s$. Since

$$\begin{aligned}\operatorname{tr}[(\bar{X} - X_{s+1})Q_s] &= \operatorname{tr}\left[(\bar{X} - X_s - \frac{\|R_s\|^2}{\|Q_s\|^2}Q_s)Q_s\right] \\ &= \operatorname{tr}[(\bar{X} - X_s)Q_s] - \frac{\|R_s\|^2}{\|Q_s\|^2}\operatorname{tr}(Q_s Q_s) \\ &= -\|R_s\|^2 + \frac{\|R_s\|^2}{\|Q_s\|^2}\operatorname{tr}(Q_s^T Q_s) \\ &= 0,\end{aligned}$$

then by Algorithm 2.1, we have

$$\begin{aligned}\operatorname{tr}[(\bar{X} - X_{s+1})Q_{s+1}] &= \operatorname{tr}\left[(\bar{X} - X_{s+1})\left(\frac{P_{s+1} - P_{s+1}^T}{2} + \frac{\operatorname{tr}(P_{s+1}Q_s)}{\|Q_s\|^2}Q_s\right)\right] \\ &= \operatorname{tr}\left[\frac{(\bar{X} - X_{s+1})P_{s+1} + (\bar{X} - X_{s+1})^T P_{s+1}^T}{2}\right] + \frac{\operatorname{tr}(P_{s+1}Q_s)}{\|Q_s\|^2}\operatorname{tr}[(\bar{X} - X_{s+1})Q_s] \\ &= \operatorname{tr}\left[\frac{(\bar{X} - X_{s+1})P_{s+1} + P_{s+1}(\bar{X} - X_{s+1})}{2}\right] \\ &= \operatorname{tr}[(\bar{X} - X_{s+1})P_{s+1}] \\ &= \operatorname{tr}[(\bar{X} - X_{s+1})A^T R_{s+1} B^T] \\ &= \operatorname{tr}[(\bar{X} - X_{s+1})A^T(C - AX_{s+1}B)B^T] \\ &= \operatorname{tr}[B(C - AX_{s+1}B)^T A(\bar{X} - X_{s+1})^T] \\ &= -\operatorname{tr}[B(C - AX_{s+1}B)^T A(\bar{X} - X_{s+1})] \\ &= -\operatorname{tr}[A(\bar{X} - X_{s+1})B(C - AX_{s+1}B)^T] \\ &= -\operatorname{tr}[(A\bar{X}B - AX_{s+1}B)(C - AX_{s+1}B)^T] \\ &= -\operatorname{tr}[(C - AX_{s+1}B)(C - AX_{s+1}B)^T] \\ &= -\operatorname{tr}(R_{s+1}R_{s+1}^T) \\ &= -\operatorname{tr}(R_{s+1}^T R_{s+1}) = -\|R_{s+1}\|^2.\end{aligned}$$

Therefore, (7) holds for $k = s + 1$. By the principle of induction the proof is completed. \square

Theorem 2.1. Suppose that Problem I is consistent, then for an arbitrary initial matrix $X_1 \in SSR^{n \times n}$, a solution of Problem I can be obtained with finite iteration steps in the absence of roundoff errors and the minimum of the steps marked t_0 is within $\min(mp, n^2)$.

Proof. If $R_i \neq \mathbf{0}$, $i = 1, 2, \dots, mp$, by Lemma 2.3 we have $Q_i \neq \mathbf{0}$, $i = 1, 2, \dots, mp$, then we can compute X_{mp+1} , R_{mp+1} by Algorithm 2.1.

By Lemma 2.2, we have

$$\operatorname{tr}(R_{mp+1}^T R_i) = 0, \quad i = 1, 2, \dots, mp$$

and

$$\text{tr}(R_i^T R_j) = 0, \quad i, j = 1, 2, \dots, mp, \quad i \neq j.$$

It can be seen that the set of R_1, R_2, \dots, R_{mp} is an orthogonal basis of the matrix space $R^{m \times p}$, which implies that $R_{mp+1} = \mathbf{0}$, i.e., X_{mp+1} is a solution of Problem I.

When Problem I is consistent, we can verify that the solution of Problem I can be obtained within t_0 iterative steps, where $t_0 = \min(mp, n^2)$. In fact, if $n^2 \leq mp$ and if $R_i \neq \mathbf{0}, i = 1, 2, \dots, n^2$, then $Q_i \neq \mathbf{0}, i = 1, 2, \dots, n^2$, and we can compute $X_{n^2+1}, R_{n^2+1}, Q_{n^2+1}$ by Algorithm 2.1. Similar to the previous proof, we have $Q_{n^2+1} = \mathbf{0}$, and then by Lemma 2.3, we have $R_{n^2+1} = \mathbf{0}$, i.e., X_{n^2+1} is a solution of Problem I. \square

From Theorem 2.1, we can easily obtain the following result.

Theorem 2.2. *The necessary and sufficient conditions of the inconsistency of Problem I is that there exists a positive integer k , such that $R_k \neq \mathbf{0}$ and $Q_k = \mathbf{0}$ in the process of Algorithm 2.1.*

Proof. *Sufficiency:* If there exists a positive integer k , such that $R_k \neq \mathbf{0}$ and $Q_k = \mathbf{0}$, by Lemma 2.3, it is easy to see that Problem I is inconsistent.

Necessity: The inconsistency of Problem I implies that $R_i \neq \mathbf{0}$ for all positive integer i . If $Q_i \neq \mathbf{0}$ for all positive integer i , then Problem I has solutions by Theorem 2.1, which contradict to the inconsistency of Problem I. Therefore, there exists a positive integer number k , such that $R_k \neq \mathbf{0}$ and $Q_k = \mathbf{0}$. \square

From Theorem 2.1 and 2.2, we get the conditions when Algorithm 2.1 can be terminated. To show the least-norm skew-symmetric solution of Problem I, we first introduce the following result.

Lemma 2.4 (See Peng et al. [9, Lemma 2.4]). *Suppose that the consistent system of linear equation $My = b$ has a solution $y_0 \in R(M^T)$, then y_0 is the least-norm solution of the system of linear equations.*

By Lemma 2.4, the following result can be obtained.

Theorem 2.3. *Suppose that Problem I is consistent. If we choose the initial iterative matrix $X_1 = A^T H^T B^T - BHA$, where H is an arbitrary matrix in $R^{p \times m}$, especially, let $X_1 = \mathbf{0} \in SSR^{n \times n}$, we can obtain the unique least-norm skew-symmetric solution of Problem I within finite iterative steps in the absence of roundoff errors by using Algorithm 2.1.*

Proof. By Algorithm 2.1 and Theorem 2.1, if we let $X_1 = A^T H^T B^T - BHA$, where H is an arbitrary matrix in $R^{p \times m}$, we can obtain the solution X^* of Problem I within finite iterative steps in the absence of roundoff errors, and the solution X^* can be represented that $X^* = A^T Y^T B^T - BYA$.

In the sequel, we will prove that X^* is just the least-norm solution of Problem I.

Consider the following system of matrix equations:

$$\begin{cases} AXB = C, \\ B^T XA^T = -C^T. \end{cases} \tag{8}$$

If Problem I has a solution $X_0 \in SSR^{n \times n}$, then $X_0^T = -X_0, AX_0B = C$, and

$$B^T X_0 A^T = (AX_0^T B)^T = (-AX_0 B)^T = -(AX_0 B)^T = -C^T.$$

Hence, the systems of matrix equations (8) also has a solution X_0 .

Conversely, if the systems of matrix equations (8) has a solution $\bar{X} \in R^{n \times n}$, such that $A\bar{X}B = C$, $B^T\bar{X}A^T = -C^T$, let $X_0 = (\bar{X} - \bar{X}^T)/2$, then $X_0 \in SSR^{n \times n}$, and

$$\begin{aligned} AX_0B &= \frac{1}{2}A(\bar{X} - \bar{X}^T)B \\ &= \frac{1}{2}(A\bar{X}B - A\bar{X}^TB) \\ &= \frac{1}{2}[A\bar{X}B - (B^T\bar{X}A^T)^T] \\ &= \frac{1}{2}[C - (-C^T)^T] = C. \end{aligned}$$

Therefore, X_0 is a solution of Problem I.

So the solvability of Problem I is equivalent to that of the systems of matrix equations (8), and the solution of Problem I must be the solution of the systems of matrix equations (8).

Let S'_E denote the set of all solutions of the systems of matrix equations (8), then we know that $S_E \subset S'_E$, where S_E is the set of all solutions of Problem I. In order to prove that X^* is the least-norm solution of Problem I, it is enough to prove that X^* is the least-norm solution of the systems of matrix equations (8). Denote $\text{vec}(X) = x$, $\text{vec}(X^*) = x^*$, $\text{vec}(Y^T) = y_1$, $\text{vec}(Y) = y_2$, $\text{vec}(C) = c_1$, $\text{vec}(C^T) = c_2$, then the systems of matrix equations (8) is equivalent to the systems of linear equations

$$\begin{bmatrix} B^T \otimes A \\ A \otimes B^T \end{bmatrix} x = \begin{bmatrix} c_1 \\ -c_2 \end{bmatrix}. \tag{9}$$

Noting that

$$\begin{aligned} x^* &= \text{vec}(A^TY^TB^T - BYA) \\ &= (B \otimes A^T)y_1 - (A^T \otimes B)y_2 \\ &= [B \otimes A^T \quad A^T \otimes B] \begin{bmatrix} y_1 \\ -y_2 \end{bmatrix} \in R \left(\begin{bmatrix} B^T \otimes A \\ A \otimes B^T \end{bmatrix}^T \right), \end{aligned}$$

by Lemma 2.4 we know that X^* is the least-norm solution of the systems of linear equations (9). Since vector operator is isomorphic, X^* is the unique least-norm solution of the systems of matrix equations (8), then X^* is the unique least-norm solution of Problem I. \square

3. The solution of Problem II

In this section, we will show that the optimal approximate solution of Problem II for a given matrix can be derived by finding the least-norm skew-symmetric solution of a new corresponding matrix equation $A\tilde{X}B = \tilde{C}$.

For a given matrix $X_0 \in R^{n \times n}$, since symmetric matrix and a skew-symmetric matrix are orthogonal each other, we have

$$\begin{aligned} \|X - X_0\|^2 &= \left\| X - \left(\frac{X_0 + X_0^T}{2} + \frac{X_0 - X_0^T}{2} \right) \right\|^2 \\ &= \left\| \left(X - \frac{X_0 + X_0^T}{2} \right) - \frac{X_0 - X_0^T}{2} \right\|^2 \\ &= \left\| X - \frac{X_0 + X_0^T}{2} \right\|^2 + \left\| \frac{X_0 - X_0^T}{2} \right\|^2 \end{aligned}$$

for any $X \in SSR^{n \times n}$.

When Problem I is consistent, the set of solutions of Problem I denoted by S_E is not empty, then linear equation

$$AXB = C$$

is equivalent to the following equation:

$$A \left(X - \frac{X_0 - X_0^T}{2} \right) B = C - A \frac{X_0 - X_0^T}{2} B.$$

Let $\tilde{X} = X - (X_0 - X_0^T)/2$, $\tilde{C} = C - A[(X_0 - X_0^T)/2]B$, then Problem II is equivalent to finding the least-norm skew-symmetric solution \tilde{X}^* of the matrix equation

$$A\tilde{X}B = \tilde{C}. \tag{10}$$

By using Algorithm 2.1, let initially iterative matrix $\tilde{X}_1 = A^T H^T B^T - BHA$, or more especially, let $\tilde{X}_1 = \mathbf{0} \in R^{n \times n}$, we can obtain the unique least-norm solution \tilde{X}^* of the matrix equation (10), then we can obtain the solution \hat{X} of Problem II, and \hat{X} can be represented that $\hat{X} = \tilde{X}^* + (X_0 - X_0^T)/2$.

4. Examples for the iterative methods

In this section, we will show several numerical examples to illustrate our results. All the tests are performed by MATLAB 6.5.1.

Example 1. Consider the skew-symmetric solution of the equation $AXB = C$, where

$$A = \begin{pmatrix} 1 & 3 & -5 & 7 & -9 \\ 2 & 0 & 4 & 6 & -1 \\ 0 & -2 & 9 & 6 & -8 \\ 3 & 6 & 2 & 27 & -13 \\ -5 & 5 & -22 & -1 & -11 \\ 8 & 4 & -6 & -9 & -19 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 0 & 8 & -5 & 4 \\ -1 & 5 & 0 & -2 & 3 \\ 4 & -1 & 0 & 2 & 5 \\ 0 & 3 & 9 & 2 & -6 \\ -2 & 7 & -8 & 1 & 11 \end{pmatrix},$$

$$C = \begin{pmatrix} 171 & -537 & 74 & -29 & -281 \\ 142 & -278 & 212 & -92 & -150 \\ 196 & -523 & -59 & -111 & 24 \\ 661 & -1507 & 922 & -234 & -1003 \\ -39 & -192 & -207 & 186 & -227 \\ -165 & -292 & -1154 & 76 & 422 \end{pmatrix}.$$

We will find the skew-symmetric solution of the matrix equation $AXB = C$ by using Algorithm 2.1. Because of the influence of the error of calculation, the residual R_i is usually unequal to zero in the process of the iteration, where $i = 1, 2, \dots$. For any chosen positive number ε , however small enough, e.g., $\varepsilon = 1.0000e - 010$, whenever $\|R_k\| < \varepsilon$, stop the iteration, and X_k is regarded to be a solution of the matrix equation $AXB = C$. Choose an initially iterative

matrix $X_1 \in SSR^{5 \times 5}$, such as

$$X_1 = \begin{pmatrix} 0 & 0 & 1 & -3 & 0 \\ 0 & 0 & -2 & -1 & 3 \\ -1 & 2 & 0 & -1 & 0 \\ 3 & 1 & 1 & 0 & -4 \\ 0 & -3 & 0 & 4 & 0 \end{pmatrix},$$

by Algorithm 2.1, we have

$$X_{14} = \begin{pmatrix} 0 & 2.0000 & -1.0000 & -2.0000 & -0.0000 \\ -2.0000 & 0 & 2.0000 & 1.0000 & -4.0000 \\ 1.0000 & -2.0000 & 0 & -1.0000 & -0.0000 \\ 2.0000 & -1.0000 & 1.0000 & 0 & -4.0000 \\ 0.0000 & 4.0000 & 0.0000 & 4.0000 & 0 \end{pmatrix},$$

$$\|R_{14}\| = 3.2646e - 011 < \varepsilon.$$

So we obtain a skew-symmetric solution of the matrix equation $AXB = C$ as follows:

$$X = \begin{pmatrix} 0 & 2.0000 & -1.0000 & -2.0000 & -0.0000 \\ -2.0000 & 0 & 2.0000 & 1.0000 & -4.0000 \\ 1.0000 & -2.0000 & 0 & -1.0000 & -0.0000 \\ 2.0000 & -1.0000 & 1.0000 & 0 & -4.0000 \\ 0.0000 & 4.0000 & 0.0000 & 4.0000 & 0 \end{pmatrix}.$$

Let

$$X_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

by Algorithm 2.1, we have

$$X_{14} = \begin{pmatrix} 0 & 2.0000 & -1.0000 & -2.0000 & -0.0000 \\ -2.0000 & 0 & 2.0000 & 1.0000 & -4.0000 \\ 1.0000 & -2.0000 & 0 & -1.0000 & -0.0000 \\ 2.0000 & -1.0000 & 1.0000 & 0 & -4.0000 \\ 0.0000 & 4.0000 & 0.0000 & 4.0000 & 0 \end{pmatrix},$$

$$\|R_{14}\| = 9.8875e - 011 < \varepsilon.$$

So we obtain a skew-symmetric solution of the matrix equation $AXB = C$ as follows:

$$X = \begin{pmatrix} 0 & 2.0000 & -1.0000 & -2.0000 & -0.0000 \\ -2.0000 & 0 & 2.0000 & 1.0000 & -4.0000 \\ 1.0000 & -2.0000 & 0 & -1.0000 & -0.0000 \\ 2.0000 & -1.0000 & 1.0000 & 0 & -4.0000 \\ 0.0000 & 4.0000 & 0.0000 & 4.0000 & 0 \end{pmatrix}.$$

Example 2. Consider the least-norm solution of the equation $AXB = C$ in Example 1. Let

$$H = \begin{pmatrix} 9 & 0 & -2 & 5 & 4 & 3 \\ 8 & 4 & 3 & 0 & 1 & 1 \\ 3 & 0 & 1 & 6 & 2 & 5 \\ 2 & 5 & 2 & 8 & -5 & -3 \\ -6 & 0 & -7 & 1 & 0 & 2 \end{pmatrix}$$

and

$$X_1 = A^T H^T B^T - BHA.$$

By using Algorithm 2.1, we have

$$X_{17} = \begin{pmatrix} 0 & 2.0000 & -1.0000 & -2.0000 & -0.0000 \\ -2.0000 & 0 & 2.0000 & 1.0000 & -4.0000 \\ 1.0000 & -2.0000 & 0 & -1.0000 & 0.0000 \\ 2.0000 & -1.0000 & 1.0000 & 0 & -4.0000 \\ 0.0000 & 4.0000 & -0.0000 & 4.0000 & 0 \end{pmatrix},$$

$$\|R_{17}\| = 8.1162e - 011 < \varepsilon.$$

So we obtain the least-norm solution of the matrix equation $AXB = C$ as follows:

$$X = \begin{pmatrix} 0 & 2.0000 & -1.0000 & -2.0000 & -0.0000 \\ -2.0000 & 0 & 2.0000 & 1.0000 & -4.0000 \\ 1.0000 & -2.0000 & 0 & -1.0000 & 0.0000 \\ 2.0000 & -1.0000 & 1.0000 & 0 & -4.0000 \\ 0.0000 & 4.0000 & -0.0000 & 4.0000 & 0 \end{pmatrix}.$$

Example 3. Consider the skew-symmetric solution of the equation $AXB = C$, where

$$A = \begin{pmatrix} 1 & -1 & 0 & 3 \\ -1 & -3 & -4 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -1 & 0 & -1 \\ -3 & 0 & 1 & -1 \\ 0 & -2 & 4 & 1 \\ 1 & -2 & 1 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 20 & 3 & -22 & 2 \\ 24 & 24 & -72 & 6 \\ 16 & -18 & 28 & -2 \end{pmatrix}.$$

Let $\varepsilon = 1.0000e - 005$ and initial matrix $X_1 = \mathbf{0}$. By using Algorithm 2.1, we have

$$\|R_6\| = 1.0408e + 003, \quad \|Q_6\| = 1.9143e - 008 < \varepsilon.$$

Therefore, there is no skew-symmetric solution for the matrix equation $AXB = C$ by Theorem 2.2.

Example 4. Let S_E denote the set of all skew-symmetric solutions of the matrix equation $AXB = C$, where the matrices A, B and C are mentioned in Example 1. Suppose

$$X_0 = \begin{pmatrix} 1 & 0 & 4 & -1 & 0 \\ 5 & 3 & 2 & 7 & 4 \\ -1 & -2 & 0 & -1 & 0 \\ 2 & 6 & 1 & 8 & -4 \\ 0 & 3 & 1 & 4 & 2 \end{pmatrix},$$

we will find $\widehat{X} \in S_E$, such that

$$\|\widehat{X} - X_0\| = \min_{X \in S_E} \|X - X_0\|,$$

i.e., find the optimal approximate solution to the matrix X_0 in S_E . Let $\widetilde{X}_0 = (X_0 - X_0^T)/2$, $\widetilde{X} = X - \widetilde{X}_0$, $\widetilde{C} = C - A\widetilde{X}_0B$, by the method mentioned in Section 3, we can obtain the least-norm skew-symmetric solution \widetilde{X}^* of the matrix equation $A\widetilde{X}B = \widetilde{C}$ by choosing the initial iteration matrix $\widetilde{X}_1 = \mathbf{0}$, and \widetilde{X}^* is that

$$\widetilde{X}_{14}^* = \begin{pmatrix} 0 & 4.5000 & -3.5000 & -0.5000 & -0.0000 \\ -4.5000 & 0 & 0.0000 & 0.5000 & -4.5000 \\ 3.5000 & -0.0000 & 0 & -0.0000 & 0.5000 \\ 0.5000 & -0.5000 & 0.0000 & 0 & 0.0000 \\ 0.0000 & 4.5000 & -0.5000 & -0.0000 & 0 \end{pmatrix},$$

$$\|R_{14}\| = 7.0170e - 011 < \varepsilon = 1.0000e - 010$$

and

$$\widehat{X} = \widetilde{X}_{14}^* + X_0 = \begin{pmatrix} 0 & 2.0000 & -1.0000 & -2.0000 & -0.0000 \\ -2.0000 & 0 & 2.0000 & 1.0000 & -4.0000 \\ 1.0000 & -2.0000 & 0 & -1.0000 & -0.0000 \\ 2.0000 & -1.0000 & 1.0000 & 0 & -4.0000 \\ 0.0000 & 4.0000 & 0.0000 & 4.0000 & 0 \end{pmatrix}.$$

Acknowledgements

The authors are grateful to the anonymous referee for the constructive and helpful comments and Prof. Lothar Reichel for the communication.

References

[1] A. Bjerhammer, Rectangular reciprocal matrices with special reference to geodetic calculations, Kung. Tekn. Hogsk. Handl. Stockholm 45 (1951) 1–86.
 [2] K.W. Eric Chu, Symmetric solutions of linear matrix equations by matrix decompositions, Linear Algebra Appl. 119 (1989) 35–50.
 [3] G.H. Golub, C.F. Van Loan, Matrix Computations, John Hopkins University Press, Baltimore, MD, 1996.
 [4] F.J. Henk Don, On the symmetric solution of a linear matrix equation, Linear Algebra Appl. 93 (1988) 1–7.

- [5] D. Hua, On the symmetric solutions of linear matrix equations, *Linear Algebra Appl.* 131 (1990) 1–7.
- [6] J.R. Magnus, L-structured matrices and linear matrix equation, *Linear Multilinear Algebra Appl.* 14 (1983) 67–88.
- [7] S.K. Mitra, Common solutions to a pair of linear matrix equations $A_1 X B_1 = C_1$, $A_2 X B_2 = C_2$, *Proc. Cambridge Philos. Soc.* 74 (1973) 213–216.
- [8] G.R. Morris, P.L. Odell, Common solutions for n matrix equations with applications, *J. Assoc. Comput. Mach.* 15 (1968) 272–274.
- [9] Y.X. Peng, X.Y. Hu, L. Zhang, An iteration method for the symmetric solutions and the optimal approximation solution of the matrix equation $A X B = C$, *Appl. Math. Comput.* 160 (2005) 763–777.
- [10] Z.Y. Peng, An iterative method for the least squares symmetric solution of the matrix equation $A X B = C$, *Appl. Math. Comput.* 170 (2005) 711–723.
- [11] G.X. Huang, F. Yin, Constrained inverse eigenproblem and associated approximation problem for anti-Hermitian R-symmetric matrices, *Appl. Math. Comput.* (2006), doi: 10.1016/j.amc.2006.07.11.
- [12] G.-X. Huang, F. Yin, Matrix inverse problem and its optimal approximation problem for R-symmetric matrices, *Appl. Math. Comput.* (2006), doi: 10.1016/j.amc.2006.11.157.