New algorithms to compute the nearness symmetric solution of the matrix equation

Zhen-yun Peng*, Yang-zhi Fang, Xian-wei Xiao and Dan-dan Du

Abstract

In this paper we consider the nearness symmetric solution of the matrix equation

\[ AXB = C, \]

has been considered by many authors. Such as the generalized singular value decomposition method to compute the symmetric solutions, the reflexive and anti-reflexive solutions, the generalized reflexive solutions and the least-squares symmetric positive semidefinite solutions were studied by Chu (1989) (see also Dai 1990), Peng et al. (2007), Yuan et al. (2008) and Liao et al. (2003), respectively. The quotient singular value decomposition method to compute the least squares symmetric, skew-symmetric and positive semidefinite solutions were studied by Deng et al. (2003). The generalized inverse method to compute the reflexive solutions, the asymmetric positive solutions and the Hermitian part nonnegative definite solutions were considered by Cvetkovic-Ilic (2006), Arias et al. (2010) and Dragana et al. (2008), respectively. The matrix-form CGNE (Bjorck 2006) iteration method to compute the symmetric solutions, the skew-symmetric solutions and the least-squares symmetric solution were given by Peng et al. (2005), Huang et al. (2008) and Lei et al. (2007) (see also Peng 2005), respectively. The
matrix-form LSQR iteration method to compute the least-squares symmetric and anti-symmetric solutions were given by Qiu et al. (2007). The matrix-form BiCOR, CORS and GPBiCG iteration methods and the matrix-form CGNE iteration method to solve the extension from of the matrix Eq. (1) were studied by Hajarian (2015a, b) and Dehghan et al. (2010), respectively.

The problem of finding a nearest matrix in the symmetric solution set of the matrix Eq. (1) to a given matrix in the sense of the Frobenius norm, that is, finding $X$ such that

$$
\min_{X} \frac{1}{2}\|X - \tilde{X}\|^2 \quad \text{subject to } AXB = C, \quad X \in \text{SR}^{m \times n}
$$

is called the matrix nearness problem. The matrix nearness problem is initially proposed in the processes of test or the recovery of linear systems due to incomplete dates or revising dates. A preliminary estimate $\tilde{X}$ of the unknown matrix $X$ can be obtained by the experimental observation values and the information of static distribution. The matrix nearness problem (2) with unknown matrix $X$ being symmetric, skew-symmetric and generalized reflexive were considered by Liao et al. (2007) (see also Peng et al. 2005), Huang et al. (2008) and Yuan et al. (2008), respectively. The approaches taken in these papers include the generalized singular value decomposition method and the matrix-form CGNE iteration method. In addition, there are many important results on the discussions of the matrix nearness problem associated with the other matrix equations, we refer the readers to (Chu and Golub 2005; Deng and Hu 2005; Higham 1988; Jin and Wei 2004; Konstantinov et al. 2003; Penrose 1956) and references therein.

In this paper, we continue to consider the matrix nearness problem (2). By discussing the equivalent form of the matrix nearness problem (2), we derive some necessary and sufficient conditions for the matrix $X^*$ is a solution of the matrix nearness problem (2). Based on the idea of the alternating variable minimization with multiplier (AVMM) method (Bai and Tao 2015), we propose two iterative methods to compute the solution of the matrix nearness problem (2), and analyze the global convergence results of the proposed algorithms. Numerical comparisons with some existing methods are also given.

Throughout this paper the following notations are used. $R^{m \times n}$ and $\text{SR}^{m \times n}$ denote the set of $m \times n$ real matrices and the set of $n \times n$ real symmetric matrices. $I$ denote the identity matrix with size implied by context. $A^+$ denote the Moore–Penrose generalized inverse of the matrix $A$. Define the inner product in space $R^{m \times n}$ by $\langle A, B \rangle = tr(A^T B)$ for all $A, B \in R^{m \times n}$, then the associated norm is the Frobenius norm, and denoted by $\|A\|$.

**Iteration methods to solve the matrix nearness problem (2)**

In this section we first give the equivalent constrained optimization problems of the matrix nearness problem (2), and discuss the properties of the solutions of these constrained optimization problems. Then we propose iteration methods to compute the solution of the equivalent constrained optimization problems, and hence to compute the solution of the matrix nearness problem (2). Finally, we prove some convergence theorems of the proposed algorithms.

Obviously, the matrix nearness problem (2) is equivalent to the following constrained optimization problem
\[
\min_{X,Y} F(X, Y) = \frac{1}{2} \|X - \bar{X}\|^2 \quad \text{subject to} \quad AX - Y = 0, \quad YB - C = 0, \quad X \in SR^{n \times n} \tag{3}
\]

or
\[
\min_{X,Y} F(X, Y) = \frac{1}{2} \|X - \bar{X}\|^2 \quad \text{subject to} \quad XB - Y = 0, \quad AY - C = 0, \quad X \in SR^{n \times n} \tag{4}
\]

**Theorem 1** Matrix pair \([X^*;Y^*]\) is a solution of the constrained optimization problem (3) if and only if exists matrices \(M^* \in R^{m \times n}\) and \(N^* \in R^{m \times p}\) such that the following equalities (5–8) hold.

\[
(X^* - \bar{X} - A^T M^*) + (X^* - \bar{X} - A^T M^*)^T = 0, \tag{5}
\]

\[
M^* - N^* B^T = 0, \tag{6}
\]

\[
AX^* - Y^* = 0, \tag{7}
\]

\[
Y^* B - C = 0. \tag{8}
\]

**Proof** Assume that there exist matrices \(M^*\) and \(N^*\) such that the equalities (5–8) hold. Let

\[
\tilde{F}(X, Y) = F(X, Y) - \langle M^*, AX - Y \rangle - \langle N^*, YB - C \rangle.
\]

Then, for any matrices \(U \in SR^{n \times n}\) and \(V \in R^{m \times n}\), we have

\[
\tilde{F}(X^* + U, Y^* + V)
\]

\[
\begin{align*}
&= \frac{1}{2} \|X^* + U - \bar{X}\|^2 - \langle M^*, Au(X^* + U) - (Y^* + V) \rangle - \langle N^*, (Y^* + V)B - C \rangle \\
&= \tilde{F}(X^*, Y^*) + \frac{1}{2} \|U\|^2 + \langle U, X^* - \bar{X} \rangle - \langle M^*, AU - V \rangle - \langle N^*, VB \rangle \\
&= \tilde{F}(X^*, Y^*) + \frac{1}{2} \|U\|^2 + \frac{1}{2} \langle U, (X^* - \bar{X} - A^T M^*)^T \rangle \\
&\quad + \langle V, M^* - N^* B^T \rangle \\
&= \tilde{F}(X^*, Y^*) + \frac{1}{2} \|U\|^2 \geq F(X^*, Y^*)
\end{align*}
\]

This implies that the matrix pair \([X^*;Y^*]\) is a global minimizer of the matrix function \(\tilde{F}(X, Y)\). Since \(AX^* - Y^* = 0, Y^* B - C = 0\) and \(\tilde{F}(X, Y) \geq \tilde{F}(X^*, Y^*)\) hold for all \(X \in SR^{n \times n}\) and \(Y \in R^{m \times n}\), we have

\[
F(X, Y) \geq F(X^*, Y^*) + \langle M^*, AX - Y \rangle + \langle N^*, YB - C \rangle.
\]

Hence, \(F(X, Y) \geq F(X^*, Y^*)\) holds for all \(X \in SR^{n \times n}, Y \in R^{m \times n}\) with \(AX - Y = 0\) and \(YB - C = 0\). That is, the matrix pair \([X^*;Y^*]\) is a solution of the constrained optimization problem (3).

Conversely, if the matrix pair \([X^*;Y^*]\) is a global solution of the constrained optimization problem (3), then the matrix pair \([X^*;Y^*]\) certainly satisfies Karush–Kuhn–Tucker conditions of the constrained optimization problem (3). That is, there exist matrices \(M^* \in R^{m \times n}\) and \(N^* \in R^{m \times p}\) such that satisfy conditions (5–8). \(\square\)
Theorem 2 Matrix pair $[X^*; Y^*]$ is a solution of the constrained optimization problem (4) if and only if exists matrices $M^* \in \mathbb{R}^{n \times p}$ and $N^* \in \mathbb{R}^{m \times p}$ such that the following equalities (9–12) hold.

\[
(X^* - \tilde{X} - M^*B^T) + (X^* - \tilde{X} - M^*B^T)^T = 0,
\]

\[
M^* - A^TN^* = 0,
\]

\[
X^*B - Y^* = 0,
\]

\[
AY^* - C = 0.
\]

Proof The proof is similar to Theorem 1 and is omitted here. \ 

Let

\[
L_{\alpha,\beta}(X,Y,M,N) = \frac{1}{2}\|X - \tilde{X}\|^2 - \langle M, AX - Y \rangle - \langle N, YB - C \rangle + \frac{\alpha}{2}\|AX - Y\|^2
\]

\[
+ \frac{\beta}{2}\|YB - C\|^2.
\]

We propose an iteration method to solve the constrained optimization (3), and hence to solve the matrix nearness problem (2) as follows.

Algorithm 1

Step 1. Input the matrices $A, B, C, \tilde{X}$ and the tolerance $\varepsilon > 0$. Choose the initial matrices $Y_0, M_0, N_0$ and the parameters $\alpha, \beta > 0$. Set $k \leftarrow 0$.

Step 2. Exit if a stopping criterion has been met.

Step 3. Compute

\[
X_{k+1} = \arg \min_{X \in \mathbb{R}^{n \times n}} L_{\alpha,\beta}(X, Y_k, M_k, N_k),
\]

\[
Y_{k+1} = \arg \min_{Y \in \mathbb{R}^{m \times n}} L_{\alpha,\beta}(X_{k+1}, Y, M_k, N_k),
\]

\[
M_{k+1} = M_k - \alpha(AX_{k+1} - Y_{k+1}),
\]

\[
N_{k+1} = N_k - \beta(Y_{k+1}B - C),
\]

Set $k \leftarrow k + 1$ and go to Step 2.

Step 4. Let

\[
\tilde{L}_{\alpha,\beta}(X,Y,M,N) = \frac{1}{2}\|X - \tilde{X}\|^2 - \langle M, XB - Y \rangle - \langle N, AY - C \rangle + \frac{\alpha}{2}\|XB - Y\|^2
\]

\[
+ \frac{\beta}{2}\|AY - C\|^2.
\]

We similarly propose an iteration method to solve the constrained optimization (4), and hence to solve the matrix nearness problem (2) as follows.
Algorithm 2
Step 1. Input the matrices $A, B, C, \bar{X}$ and the tolerance $\varepsilon > 0$. Choose the initial matrices $Y_0, M_0, N_0$ and the parameters $\alpha, \beta > 0$. Set $k \leftarrow 0$.

Step 2. Exit if a stopping criterion has been met.

Step 3. Compute

$$X_{k+1} = \arg \min_{X \in \mathbb{SR}^{n \times n}} \frac{1}{2} \| X - \bar{X} \|^2 - (M_k, AX - Y_k) + \frac{\alpha}{2} \| AX - Y_k \|^2$$

where $S = \begin{bmatrix} \sqrt{\alpha} A \\ I \end{bmatrix} \in \mathbb{R}^{(m+n) \times n}$, $T = \begin{bmatrix} \sqrt{\alpha} Y_k + M_k / \sqrt{\alpha} \\ \bar{X} \end{bmatrix} \in \mathbb{R}^{(m+n) \times n}$. Analogously, $X_{k+1}$ in (19) can be expressed as

$$X_{k+1} = \arg \min_{X \in \mathbb{SR}^{n \times n}} \frac{1}{2} \| XS - T \|^2$$

where $S = [I, \sqrt{\alpha} B] \in \mathbb{R}^{n \times (n+p)}$, $T = [\bar{X}, \sqrt{\alpha} Y_k + M_k / \sqrt{\alpha}] \in \mathbb{R}^{n \times (n+p)}$.

To solve the problems (23) and (24), we give the following Lemma 1.

Lemma 1 (Sun 1988) Given matrices $B \in \mathbb{R}^{n \times n}$ and $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$, then the problem $\sigma_i > 0$ for $i = 1, \ldots, n \| X \Sigma - B \|^2 = \min$ have a unique least squares symmetric solution with the following expression

$$X = \Phi \circ (B \Sigma + \Sigma B^T)$$

where $\Phi_{ij} = \frac{1}{\sigma_i + \sigma_j}, \Phi = (\Phi_{ij}) \in \mathbb{R}^{n \times n}$, and $A \circ B = (a_i b_j)$ denotes the Hadamard product.

Noting that the matrix $S$ in (23) is full column rank, the singular value decomposition (SVD) of the matrix $S$ can be expressed as

$$S = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T$$
where \( \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n), \ \sigma_i > 0, \) and \( U = (U_1, U_2) \in R^{(m+n) \times (m+n)}, \ V \in R^{n \times n} \) are orthogonal matrices, \( U_1 \in R^{(m+n) \times n} \). Hence, \( X_{k+1} \) in (23) can be expressed as

\[
X_{k+1} = \arg \min_{X \in S R^{n \times n}} \frac{1}{2} \left\| U \left( \begin{array}{c} \Sigma \\ 0 \end{array} \right) V^T X - T \right\|^2
\]

\[
= \arg \min_{X \in S R^{n \times n}} \frac{1}{2} \left\| \left( \begin{array}{c} \Sigma \\ 0 \end{array} \right) V^T X V - U^T TV \right\|^2
\]

\[
= \arg \min_{X \in S R^{n \times n}} \frac{1}{2} \left\| \left( \begin{array}{c} \Sigma \\ 0 \end{array} \right) V^T X V - \left( \begin{array}{c} U_1^T \\ U_2^T \end{array} \right) TV \right\|^2
\]

\[
= \arg \min_{X \in S R^{n \times n}} \frac{1}{2} \left\| \Sigma V^T X V - U^T TV \right\|^2
\]

Let \( \tilde{T} = U_1^T TV \), we have by Lemma 1 that \( X_{k+1} \) in (23) can be expressed as

\[
X_{k+1} = V(\Phi \circ (\Sigma \tilde{T} + \tilde{T}^T \Sigma))V^T
\]

Analogously, the matrix \( S \) in (24) is full row rank, and the SVD of \( S \) can be expressed as

\[
S = P(\Sigma, 0)Q^T
\]

where \( \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n), \ \sigma_i > 0, \) and \( P \in R^{n \times n}, \ Q = (Q_1, Q_2) \in R^{(n+p) \times (n+p)} \) are the orthogonal matrices, \( Q_1 \in R^{(n+p) \times n} \), and \( X_{k+1} \) in (24) can be expressed as

\[
X_{k+1} = P(\Phi \circ (\Sigma \tilde{T} + \tilde{T}^T \Sigma))P^T
\]

where \( \tilde{T} = P^T TQ_1 \).

Then, we change our attention to compute \( Y_{k+1} \). By simply changing, \( Y_{k+1} \) in (15) can be expressed as

\[
Y_{k+1} = \arg \min_{Y \in R^{n \times n}} \frac{1}{2} \left\| Y \left( \sqrt{\alpha}I, \sqrt{\beta}B \right) - \left( \sqrt{\alpha}AX_{k+1} - M_k / \sqrt{\alpha}, \sqrt{\beta}C + N_k / \sqrt{\beta} \right) \right\|^2
\]

\[
= \left( \sqrt{\alpha}AX_{k+1} - M_k / \sqrt{\alpha}, \sqrt{\beta}C + N_k / \sqrt{\beta} \right) \left( \sqrt{\alpha}I, \sqrt{\beta}B \right)^+
\]

and \( Y_{k+1} \) in (20) can be expressed as

\[
Y_{k+1} = \arg \min_{Y \in R^{n \times p}} \frac{1}{2} \left\| \left( \sqrt{\alpha}I \right) \left( \sqrt{\beta}A \right) Y - \left( \sqrt{\alpha}X_{k+1}B - M_k / \sqrt{\alpha}, \sqrt{\beta}C + N_k / \sqrt{\beta} \right) \right\|^2
\]

\[
= \left( \sqrt{\alpha}I \right)^+ \left( \sqrt{\beta}A \right) \left( \sqrt{\alpha}X_{k+1}B - M_k / \sqrt{\alpha}, \sqrt{\beta}C + N_k + 1 / \sqrt{\beta} \right)
\]

Next, we discuss the global convergence of Algorithm 1 and 2. Note that Algorithm 2 is similar to Algorithm 1, we only discuss the convergence of Algorithm 1.

**Theorem 3** Let \((X^*, Y^*, M^*, N^*)\) be a saddle point for the Lagrange function

\[
L(X, Y, M, N) = \frac{1}{2} \left\| X - \tilde{X} \right\|^2 - \langle M, AX - Y \rangle - \langle N, YB - C \rangle
\]

of the constrained optimization problem (3), that is, the matrices \( X^*, Y^*, M^*, N^* \) satisfy conditions (5–8). Define
\[ S_{k+1} = AX_{k+1} - Y_{k+1}, \quad T_{k+1} = Y_{k+1}B - C, \quad \mu_k = \alpha \|Y_k - Y^*\|^2 \]
\[ + \frac{1}{\alpha} \|M_k - M^*\|^2 + \frac{1}{\beta} \|N_k - N^*\|^2, \]

then, the following inequality holds
\[ \mu_{k+1} \leq \mu_k - \beta \|T_{k+1}\|^2 - \alpha \|S_{k+1} + Y_{k+1} - Y_k\|^2. \] (25)

**Proof** Since \((X^*, Y^*, M^*, N^*)\) is a saddle point, we have by the saddle point theorem (Björck 2006) that
\[ L(X^*, Y^*, M, N) \leq L(X^*, Y^*, M^*, N^*) \leq L(X, Y, M^*, N^*), \]
for all \(X, Y, M, N\), where \(L(X, Y, M, N) = \frac{1}{2} \|X - \tilde{X}\|^2 - \langle N, YB - C \rangle)\), which called the Lagrange function of the constrained optimization problem (3). Hence we have
\[ L(X^*, Y^*, M^*, N^*) \leq L(X_{k+1}, Y_{k+1}, M^*, N^*) \]

Noting that \(AX^* - Y^* = 0, Y^*B - C = 0, S_{k+1} = AX_{k+1} - Y_{k+1}\) and \(T_{k+1} = Y_{k+1}B - C\), we know that the following inequality holds
\[ \frac{1}{2} \|X^* - \tilde{X}\|^2 - \frac{1}{2} \|X_{k+1} - \tilde{X}\|^2 \leq -\langle M^*, S_{k+1} \rangle - \langle N^*, T_{k+1} \rangle, \] (26)

Since \(X_{k+1}\) minimize the matrix function \(L_{\alpha, \beta}(X, Y_k, M_k, N_k)\), we have
\[ 0 = [X_{k+1} - \tilde{X} - A^TM_k + \alpha A^T(AX_{k+1} - Y_k)] + [X_{k+1} - \tilde{X} - A^TM_k + \alpha A^T(AX_{k+1} - Y_k)]^T \]
\[ = [X_{k+1} - \tilde{X} - A^TM_k + \alpha A^T(Y_k - Y_{k+1})] + [X_{k+1} - \tilde{X} - A^TM_k + \alpha A^T(Y_k - Y_{k+1})]^T, \] (27)

where the first equality is the first-order optimality condition of the problem (14), and the second equality is followed by Algorithm 1. This implies that
\[ X_{k+1} = \arg \min_{X \in \mathbb{R}^{n\times n}} \frac{1}{2} \|X - \tilde{X}\|^2 - \langle M_{k+1} - \alpha Y_{k+1} + \alpha Y_k, AX \rangle \]

Hence, we have
\[ \frac{1}{2} \|X_{k+1} - \tilde{X}\|^2 - \langle M_{k+1} - \alpha Y_{k+1} + \alpha Y_k, AX_{k+1} \rangle \]
\[ \leq \frac{1}{2} \|X^* - \tilde{X}\|^2 - \langle M_{k+1} - \alpha Y_{k+1} + \alpha Y_k, AX^* \rangle. \] (28)

Since \(Y_{k+1}\) minimize \(L_{\alpha, \beta}(X_{k+1}, Y, M_k, N_k)\), we have
\[ 0 = M_k - N_kB^T - \alpha (AX_{k+1} - Y_{k+1}) + \beta (Y_{k+1}B - C)B^T \]
\[ = M_{k+1} - N_{k+1}B^T, \] (29)

where the first equality is the first-order optimality condition of the problem (15), and the second equality is followed by Algorithm 1. This implies that
\[ Y_{k+1} = \arg \min_{Y \in \mathbb{R}^{m \times n}} \langle M_{k+1} - N_{k+1}B, Y \rangle \]

Hence, we have

\[ \left\langle M_{k+1} - N_{k+1}B^T, Y_{k+1} \right\rangle \leq \left\langle M_{k+1} - N_{k+1}B^T, Y^* \right\rangle \]  

(30)

Adding the inequalities (28) and (30), and using \( AX^* - Y^* = 0, Y^*B - C = 0 \), we know that the following inequality holds

\[ \frac{1}{2} \| X_{k+1} - \tilde{X} \|^2 - \frac{1}{2} \| X^* - \tilde{X} \|^2 \leq \left\langle M_{k+1}, S_{k+1} \right\rangle + \langle N_{k+1}, T_{k+1} \rangle - \alpha \langle Y_{k+1} - Y_k, S_{k+1} + Y_{k+1} - Y^* \rangle, \]  

(31)

Adding the inequalities (26) and (31), we have

\[ \left\langle M^* - M_{k+1}, S_{k+1} \right\rangle + \langle N^* - N_{k+1}, T_{k+1} \rangle + \alpha \langle Y_{k+1} - Y_k, S_{k+1} + Y_{k+1} - Y^* \rangle \leq 0. \]  

(32)

Noting that

\[
\begin{align*}
2\left\langle M^* - M_{k+1}, S_{k+1} \right\rangle &+ 2\langle N^* - N_{k+1}, T_{k+1} \rangle \\
&= 2\langle M^* - M_k, S_{k+1} \rangle + 2\langle N^* - N_k, T_{k+1} \rangle \\
&= \frac{2}{\alpha} \left\langle M^* - M_k, M_k - M_{k+1} \right\rangle + \frac{1}{\alpha} \| M_k - M_{k+1} \|^2 + 2\langle N^* - N_k, N_k - N_{k+1} \rangle \\
&+ \frac{2}{\beta} \left\langle N^* - N_k, N_k - N_{k+1} \right\rangle + \frac{1}{\beta} \| M_k - M_{k+1} \|^2 \\
&+ \alpha \| S_{k+1} \|^2 + \frac{1}{\beta} \left\langle N_k - N_{k+1}, N_k - N_{k+1} \right\rangle + \beta \| T_{k+1} \|^2 \\
&\text{and} \\
&\alpha \| S_{k+1} \|^2 + 2\alpha \langle Y_{k+1} - Y_k, S_{k+1} + Y_{k+1} - Y^* \rangle \\
&= \alpha \langle S_{k+1} + Y_{k+1} - Y_k \rangle^2 + \alpha \langle Y_{k+1} - Y_k \rangle^2 + 2\alpha \langle Y_{k+1} - Y_k, Y_k - Y^* \rangle \\
&= \alpha \langle S_{k+1} + Y_{k+1} - Y_k \rangle^2 + \alpha \langle Y_{k+1} - Y^* \rangle^2 - \| Y_k - Y^* \|^2, \\
\end{align*}
\]

We have by inequality (32) and the definition of \( \mu_k \) that

\[ \mu_{k+1} \leq \mu_k - \beta \| T_{k+1} \|^2 - \alpha \| S_{k+1} + Y_{k+1} - Y_k \|^2 \]

which means that the inequality (25) holds. The proof is completed. \( \square \)

Theorem 3 implies that the sequence \( \{\mu_k\} \) is a nonnegative monotone decreasing with low bounded. Hence, the limit of the sequence \( \{\mu_k\} \) exists which implies that the limit of the sequences \( \{Y_k\}, \{M_k\}, \{N_k\} \) exist, and \( S_{k+1} + Y_{k+1} - Y_k = AX_{k+1} - Y_k \rightarrow 0 \) and \( T_{k+1} = Y_{k+1}B - C \rightarrow 0 \) as \( k \rightarrow \infty \). Furthermore, \( S_{k+1} + Y_{k+1} - Y_k = AX_{k+1} - Y_k \rightarrow 0 \) as \( k \rightarrow \infty \) implies that the limit of the sequence \( \{X_k\} \) exists. Assume that \( X_k \rightarrow X^*, Y_k \rightarrow Y^*, M_k \rightarrow M^*, N_k \rightarrow N^* \) as \( k \rightarrow \infty \), then (9) and (10) are hold.
by taking limit, respectively, the Eqs. (27) and (29), and (11) and (12) are hold by
\[ S_{k+1} + Y_{k+1} - Y_k = AX_{k+1} - Y_k \to 0 \quad \text{and} \quad T_{k+1} = Y_{k+1}B - C \to 0 \quad \text{as} \quad k \to \infty. \]
Hence, we have by Theorem 1 that the matrix pair \([X^*; Y^*]\) is a solution of the con-
strained optimization problem (3), and hence is a solution of the matrix near-
ness problem (2). In addition, Note that the subjective function of the constrained
optimization problem (3) is a strictly convex functions and the constrained set
\[ \Omega = \left\{ [X; Y] \mid AX - Y = 0, YB - C = 0, X \in SR^{n \times n} \right\} \]
is closed and convex, we know
that matrix pair \([X^*; Y^*]\) is the unique solution of problem (3). Hence the sequence \([X_k]\)
generated by Algorithm 1 converges to the unique solution of the matrix nearness prob-
lem (2). These results can be described as the following Theorem 4.

**Theorem 4** Assume that \([X_k]\) is a sequence generated by Algorithm 1 with any initial
matrices \(Y_0, M_0, N_0\) and the parameters, then the sequence \(\alpha, \beta > 0\) \([X_k]\) converges to a
solution of the matrix nearness problem (2).

**Numerical experiments**

In this section, we compare Algorithm 1 and 2 with existing two methods proposed in
(Peng et al. 2005; Peng 2010), denoted, respectively, by CG and LSQR. Our computational
experiments were performed on an IBM ThinkPad T410 with 2.5 GHz and 3.0 RAM. All
tests were performed in MATLAB 7.1 with a 64-bit Windows 7 operating system.

In the implementation of Algorithm 1 and 2, we take parameters \(\alpha = \beta = 10\). The ini-
tial matrices \(Y_0, M_0, N_0\) in Algorithm 1 and 2, and \(X_0\) in Algorithm CG and LSQR are
chosen as zeros matrices. All of algorithms, the small tolerance \(\varepsilon = 10^{-8}\) and the ter-
mination criterion are chosen as \(\|AX_kB - C\| \leq \varepsilon\). In addition, the maximum iterations
numbers of the three methods is limited within 20000.

For the matrix nearness problem (2), the matrices \(A, B, \tilde{X}\) and \(C\) are given as follows (in
MATLAB style): \(A = \text{randn}(m, n), B = \text{randn}(n, p), \tilde{X} = \text{randn}(n, n), C = AX_0B\) with

| m, n, p | Algorithm 1 | Algorithm 2 | Algorithm CG | Algorithm LSQR |
|---------|-------------|-------------|---------------|---------------|
|         | IT         | CPU         | IT           | CPU          |
| 20, 20, 20 | 2807 1.0407 | 1992 0.9338 | 941 0.1603  | 816 0.1181 |
| 40, 40, 40 | 6709 60.3186 | 5085 52.1359 | 3445 25.3471 | 2925 21.6702 |
| 60, 60, 60 | 8109 74.2018 | 7882 67.0493 | 7203 45.1428 | 5987 43.2527 |
| 80, 80, 80 | 8246 88.6394 | 5503 73.6081 | 10974 86.6511 | 9773 83.7413 |
| 100, 80, 80 | 5795 67.6504 | 2081 23.6263 | 6681 59.9316 | 5636 54.6361 |
| 200, 80, 80 | 9445 130.5067 | 2503 12.9294 | 4721 60.9609 | 4913 55.7173 |
| 400, 80, 60 | 13,756 252.4951 | 2234 0.2925 | 4721 60.9609 | 4913 55.7173 |
| 80, 80, 400 | 25 0.2173  | 14,009 258.8992 | 2234 0.2925 | 4721 60.9609 |
| 100, 80, 100 | 332 3.9936  | 352 4.1774  | 2011 18.5992 | 1604 16.1305 |
| 200, 80, 200 | 39 0.7021  | 38 0.6241  | 531 7.8664  | 459 6.9108 |
| 200, 80, 400 | 21 0.5096  | 41 0.9568  | 299 11.5416 | 348 7.3581 |
| 400, 80, 200 | 41 1.0451  | 22 0.4941  | 415 8.7829  | 361 7.8676 |
\( X_0 = H + H^T \) and \( H = \text{randn}(n, n) \). Here the matrix \( C \) is chosen in this way to guarantee that the matrix nearness problem (2) is solvable.

In Table 1, we report the iteration CPU time (‘CPU’) and the iteration numbers (‘IT’) based on their average values of 10 repeated tests with randomly generated matrices \( A \), \( B \) and \( C \) for each problem size in each tests.

Based on the tests reported in Table 1 and many other performed unreported tests which show similar patterns, we have the following results: when \( m, p \gg n \), Algorithm 1 and 2 are more effective than Algorithms CG and LSQR. But when \( m \approx p \approx n \), Algorithm CG and LSQR are relatively more effective than Algorithms 1 and 2. When \( p < m \) and \( m \gg n \), Algorithm 2 is the most effective, and Algorithm 1 is the most effective when \( m < p \) and \( p \gg n \).

**Conclusions**

In this paper, we have considered the matrix nearness problem (2), i.e., finding the matrix nearness solution \( X^* \) of matrix equation \( AXB = C \) to a given matrix \( \tilde{X} \). By discussing equivalent form of the considered problem, we have derived some necessary and sufficient conditions for the matrix \( X^* \) is a solution of the considered problem. Based on the idea of the alternating variable minimization with multiplier method, we have proposed two iterative methods to compute the solution of the considered problem, and have analyzed global convergence results of the proposed algorithms. Numerical results illustrate proposed methods are more effective than existing two methods proposed in Peng et al. (2005) and Peng (2010).

**Authors’ contributions**

ZP conceived the study, carried out the analysis of the sequences and drafted the manuscript. YF, XX and DD participated in its design and coordination and helped to finalise the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

Research is supported by National Natural Science Foundation of China (11261014, 11271117, 11301107).

**Competing interests**

The authors declare that they have no competing interests.

Received: 25 January 2016   Accepted: 25 May 2016

**Published online:** 07 July 2016

**References**

Arias ML, Gonzalez MC (2010) Positive solutions to operator equations \( AXB = C \). Linear Algebra Appl 433:1194–1202

Bai ZZ, Tao M (2015) Rigorous convergence analysis of alternating variable minimization with multiplier methods for quadratic programming problems with equality constraints. BIT Numer Math. doi:10.1007/s10543-015-0563-2

Björck A (2006) Numerical methods in scientific computing, vol. I–II. SIAM, Philadelphia. http://www.mai.liu.se/akbjo/

Chu KE (1989) Symmetric solution of linear matrix equations by matrix decomposition. Linear Algebra Appl 119:35–50

Chu M, Golub G (2005) Inverse Eigenvalue Problems, Theory Algorithms and Applications. Oxford University Press, Oxford

Cvetković-Ilić DS (2006) The reflexive solutions of the matrix equation \( AXB = C \). Computers Math Appl 51:897–902

Dai H (1990) On the symmetric solutions of linear matrix equations. Linear Algebra Appl 131:1–7

Dehghan M, Hajarian M (2010) The general coupled matrix equations over generalized bisymmetric matrices. Linear Algebra Appl 432:1531–1552

Deng YB, Hu X (2005) On solutions of the matrix equation \( AX^2 + BY^2 = C \). J Comput Math 23:17–26

Deng YB, Hu XY, Zhang L (2003) Least squares solution of \( 8X^2 + T \) over symmetric, skewsymmetric, and positive semidefinite matrices \( X \). SIAM J Matrix Anal Appl 25:486–494

Dragana S, Cvetkovic-Ilić DS (2008) Re-nnd solutions of the matrix equation \( AXB = C \). J Aust Math Soc 84:63–72

Hajarian M (2015a) Developing BiCOR and CORS methods for coupled Sylvester-transpose and periodic Sylvester matrix equations. Appl Math Model 39:6073–6084
Hajarian M (2015b) Matrix GPBiCG algorithms for solving the general coupled matrix equations. IET Control Theory Appl 9:74–81
Higham NJ (1988) Computing a nearest symmetric positive semidefinite matrix. Linear Algebra Appl 103:103–118
Huang GX, Yin F, Guo K (2008) An iterative method for the skew-symmetric solution and the optimal approximate solution of the matrix equation $AXB = C$. J Comput Appl Math 212:231–244
Jin X, Wei Y (2004) Numerical Linear Algebra and its Applications. Science Press, Beijing/New York
Konstantinov M, Gu D, Mehrmann V, Petkov P (2003) Perturbation theory for matrix equations. Elsevier, Amsterdam
Lei Y, Liao AP (2007) A minimal residual algorithm for the inconsistent matrix equation $AXB = C$ over symmetric matrices. Appl Math Comput 188:499–513
Liao AP, Bai ZZ (2003) Least-squares solution of $AXB = D$ over symmetric positive semidefinite matrices $X$. J Comput Math 21:175–182
Liao AP, Lei Y (2007) Optimal approximate solution of the matrix equation $AXB = C$ over symmetric matrices. J Comput Math 25:543–552
Peng ZY (2005) An iterative method for the least squares symmetric solution of the linear matrix equation $AXB = C$. Appl Math Comput 170:711–723
Peng ZY (2010) A matrix LSQR iterative method to solve matrix equation $AXB = C$. Int J Comput Math 87:1820–1830
Peng XY, Hu XY, Zhang L (2005) An iteration method for the symmetric solutions and the optimal approximation solution of matrix equation $AXB = C$. Appl Math Comput 160:763–777
Peng XY, Hu XY, Zhang L (2007) The reflexive and anti-reflexive solutions of the matrix equation, $A^TXB = C$. J Comput Appl Math 200:749–760
Penrose R (1956) On best approximate solutions of linear matrix equations. Proc Cambridge Philos Soc 52:17–19
Qiu YY, Zhang ZY, Lu JF (2007) Matrix iterative solutions to the least squares problem of $BXA = F$ with some linear constraints. Appl Math Comput 185:284–300
Sun JG (1988) Two kinds of inverse eigenvalue problems for real symmetric matrices. J Comput Math 3:283–290
Yuan YX, Dai H (2008) Generalized reflexive solutions of the matrix equation $AXB = D$ and an associated optimal approximation problem. Computers Math Appl 56:1643–1649

Submit your manuscript to a SpringerOpen journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ➤ springeropen.com