International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:5, No:4, 2011

Iterative solutions to some linear matrix equations

Jiashang Jiang, Hao Liu, Yongxin Yuan

Abstract—In this paper the gradient based iterative algorithms are presented to solve the following four types linear matrix equations: (a) $AXB = F$; (b) $AXB = F,\ CXD = G$; (c) $AXB = F$ s. t. $X = X^T$; (d) $AXB+CYD = F$, where X and Y are unknown matrices, A, B, C, D, F, G are the given constant matrices. It is proved that if the equation considered has a solution, then the unique minimum norm solution can be obtained by choosing a special kind of initial matrices. The numerical results show that the proposed method is reliable and attractive.

Keywords—matrix equation, iterative algorithm, parameter estimation, minimum norm solution.

I. Introduction

MATRIX equations are often encountered in many systems and control applications, such as Lyapunov matrix equations, Sylvester matrix equations and so on. Traditional methods convert such matrix equations into their equivalent forms by using the Kronecker product and stretching function, however, which involve the inversion of the associated large matrix and result in increasing computation and excessive computer memory. In recent years iterative approaches for solving matrix equations and recursive identification for parameter estimation have received much attention, e.g.,[1-6]. For example, Dehghan and Hajarian studied the finite iterative algorithm for the reflexive solutions of the generalized coupled Sylvester matrix equations [7]; Mukaidani et al. gave a numerical algorithm for finding solution of cross-coupled algebraic Riccati equations [8]; Zhou and Duan studied the explicit solutions to generalized Sylvester matrix equations [9, 10]; Ding and Chen presented a gradient based and a least-squares based iterative algorithms for generalized Sylvester matrix equations and general coupled matrix equations [11, 12].

Our main contribution in this paper is to provide a gradient based iterative algorithm to solve the following matrix equations:

$$AXB = F, \tag{1}$$

$$AXB = F,\ CXD = G, \tag{2}$$

$$AXB = F \ \text{ s. t. } \ X = X^T, \tag{3}$$

$$AXB + CYD = F, \tag{4}$$

where X and Y are unknown matrices, A, B, C, D, F, G are the given constant matrices. We observe that Ding et al.[13, 14] have considered the iterative solutions of Eqs.(1) and (2), but their algorithms can work well on the condition that the matrix equation considered should have the unique solution, which seems a rigorous requirement. In this paper, we present gradient based iterative algorithms to solve Eqs.(1)-(4) and prove that if the equation considered has a solution, then the unique minimum norm solution can be obtained by choosing a special kind of initial matrices. The numerical results show that the proposed method is reliable and attractive.

Throughout this paper, we shall adopt the following notation. $\mathbf{R}^{m\times n}$ denotes the set of all $m \times n$ real matrices. A^T, A^+ and $R(A)$ stand for the transpose, Moore-Penrose generalized inverse and the column space of the matrix A, respectively. $\lambda_{\max}(M^TM)$ denotes the maximum eigenvalue of M^TM. I_n represents the identity matrix of order n. For $A, B \in \mathbf{R}^{m\times n}$, an inner product in $\mathbf{R}^{m\times n}$ is defined by $(A, B) = \text{trace}(B^TA)$, then $\mathbf{R}^{m\times n}$ is a Hilbert space. The matrix norm $\|\cdot\|$ induced by the inner product is the Frobenius norm. Given two matrices $A = [a_{ij}] \in \mathbf{R}^{m\times n}$ and $B \in \mathbf{R}^{p\times q}$, the Kronecker product of A and B is defined by $A \otimes B = [a_{ij}B] \in \mathbf{R}^{mp\times nq}$. Also, for an $m \times n$ matrix $A = [a_1, a_2, \cdots, a_n]$, where $a_i, i = 1, \cdots, n$, is the i-th column vector of A, the stretching function $\text{vec}(A)$ is defined as $\text{vec}(A) = [a_1^T, a_2^T, \cdots, a_n^T]^T$.

II. Preliminary Considerations

To begin with, we first give some lemmas.

Lemma 1: [11, 12, 13]. If the linear equation system $Mx = b$, where $M \in \mathbf{R}^{m\times n}, b \in \mathbf{R}^m$, has a unique solution x^*, then for any initial vector $x_0 \in \mathbf{R}^n$, the gradient based iterative algorithm

$$\begin{cases} x_k = x_{k-1} + \mu M^T(b - Mx_{k-1}), \\ 0 < \mu < \frac{2}{\lambda_{\max}(M^TM)} \ \text{ or } \ 0 < \mu < \frac{2}{\|M\|^2}, \end{cases}$$

yields $\lim_{k\to\infty} x_k = x^*$.

Lemma 2: [15]. Let $D \in \mathbf{R}^{m\times n}, H \in \mathbf{R}^{n\times l}, J \in \mathbf{R}^{l\times s}$. Then

$$\text{vec}(DHJ) = (J^T \otimes D)\text{vec}(H).$$

Lemma 3: [16]. If $L \in \mathbf{R}^{m\times q},\ b \in \mathbf{R}^m$, then $Ly = b$ has a solution $y \in \mathbf{R}^q$ if and only if $LL^+b = b$. In this case, the general solution of the equation can be described as $y = L^+b + (I_q - L^+L)z$, where $z \in \mathbf{R}^q$ is an arbitrary vector.

Lemma 4: [16]. Suppose that the consistent linear equation $Ax = b$ has a solution $x \in R(A^T)$, then x is the unique minimum Frobenius norm solution of the linear equation.

Lemma 5: [17]. Let $f(x, y) = \sum_{i,j=0}^{K} c_{ij}x^iy^j$ be a real coefficient binary polynomial. For $A \in \mathbf{R}^{m\times m}, B \in \mathbf{R}^{n\times n}$, define a matrix polynomial as $f(A, B) = \sum_{i,j=0}^{K} c_{ij}A^i \otimes B^j$, where $A^0 = I_m, B^0 = I_n$. If the eigenvalues of A and B are, respectively, ξ_i and μ_j, $i = 1, \cdots, m;\ j = 1, \cdots, n$, then

Jiashang Jiang: School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, P R China. e-mail: jjiashang@163.com
Hao Liu: Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P R China. e-mail: hliu@nuaa.edu.cn
Yongxin Yuan: School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, P R China. e-mail: yuanyx_703@163.com

the eigenvalues of $f(A,B)$ are $f(\xi_i,\mu_j),\ i=1,\cdots,m;\ j=1,\cdots,n.$

Lemma 6: The equation of $AXB=F$ has a symmetric solution X if and only if the matrix equations

$$\left\{\begin{array}{l} AXB=F, \\ B^TXA^T=F^T, \end{array}\right. \tag{5}$$

are consistent.

Proof. If the equation of $AXB=F$ has a symmetric solution X^*, then $AX^*B=F$, and $(AX^*B)^T=B^TX^*A^T=F^T$. That is to say, X^* is a solution of (5).

Conversely, if the matrix equations of (5) has a solution, say, $X=U$. Let $X^*=\frac{1}{2}(U+U^T)$, then X^* is a symmetric matrix, and

$$AX^*B=\frac{1}{2}(AUB)+\frac{1}{2}(AU^TB)=\frac{1}{2}F+\frac{1}{2}(F^T)^T=F.$$

Hence, X^* is a symmetric solution of $AXB=F$.

III. The solution of the matrix equation $AXB=F$

Using Lemma 2, we know that the equation of (1) is equivalent to

$$(B^T\otimes A)\mathrm{vec}(X)=\mathrm{vec}(F). \tag{6}$$

Theorem 1: Suppose that $A\in\mathbf{R}^{m\times n}, B\in\mathbf{R}^{p\times q}$ and $F\in\mathbf{R}^{m\times q}$. If the equation of (1) has a unique solution X^*, then for any initial matrix X_0, the gradient based iterative algorithm

$$\left\{\begin{array}{l} X_k=X_{k-1}+\mu A^T(F-AX_{k-1}B)B^T, \\ 0<\mu<\frac{2}{\lambda_{\max}(A^TA)\cdot\lambda_{\max}(BB^T)} \quad \text{or} \quad 0<\mu<\frac{2}{\|A\|^2\cdot\|B\|^2}, \end{array}\right. \tag{7}$$

yields $\lim_{k\to\infty}X_k=X^*$.

Proof. Applying Lemma 1 to Eq.(6), we have the gradient based iterative algorithm for the equation of (1) described as follows.

$$\begin{array}{r} \mathrm{vec}(X_k)=\mathrm{vec}(X_{k-1})+\mu(B^T\otimes A)^T\left(\mathrm{vec}(F)\right. \\ \left.-(B^T\otimes A)\mathrm{vec}(X_{k-1})\right). \end{array} \tag{8}$$

From (8) and Lemma 2, we can easily obtain

$$X_k=X_{k-1}+\mu A^T(F-AX_{k-1}B)B^T. \tag{9}$$

By Lemma 5, we know that

$$\begin{array}{rcl} \lambda_{\max}\left((B^T\otimes A)^T(B^T\otimes A)\right) & = & \lambda_{\max}\left(BB^T\otimes A^TA\right) \\ =\lambda_{\max}(A^TA)\cdot\lambda_{\max}(BB^T) & \leq & \|A\|^2\cdot\|B\|^2. \end{array}$$

According to Lemma 1, Theorem 1 is proven.

Now, assume that $J\in\mathbf{R}^{m\times q}$ is an arbitrary matrix, then we have

$$\mathrm{vec}(A^TJB^T)=(B\otimes A^T)\mathrm{vec}(J)\subset R(B\otimes A^T).$$

It is obvious that if we choose

$$X_0=A^TJB^T, \tag{10}$$

where J is an arbitrary matrix, then all X_k generated by the equation of (9) satisfy

$$\mathrm{vec}(X_k)\subset R(B\otimes A^T),\ k=1,2,\cdots.$$

It follows from Lemma 3 that the equation of (1) has a solution if and only if

$$(B^T\otimes A)(B^T\otimes A)^+\mathrm{vec}(F)=\mathrm{vec}(F),$$

which implies that

$$AA^+FBB^+=F. \tag{11}$$

By Lemma 4, we have proved the following result.

Theorem 2: Suppose that the condition (11) is satisfied. If we choose the initial matrix by (10), where J is an arbitrary matrix, or especially, $X_0=0$, then the iterative solution $\{X_k\}$ obtained by the gradient iterative algorithm (7) converges to the unique minimum Frobenius norm solution X^* of Eq.(1).

IV. The solution of the matrix equations $AXB=F, CXD=G$

Using Lemma 2, we know that the equations of (2) are equivalent to

$$M\mathrm{vec}(X)=\left[\begin{array}{c}\mathrm{vec}(F)\\ \mathrm{vec}(G)\end{array}\right], \tag{12}$$

where

$$M=\left[\begin{array}{c}B^T\otimes A\\ D^T\otimes C\end{array}\right].$$

Theorem 3: Suppose that $A\in\mathbf{R}^{m\times n}, B\in\mathbf{R}^{p\times q},\ C\in\mathbf{R}^{f\times n}, D\in\mathbf{R}^{p\times t},\ F\in\mathbf{R}^{m\times q}$ and $G\in\mathbf{R}^{f\times t}$. If the equation of (2) has a unique solution X^*, then for any initial matrix X_0, the gradient based iterative algorithm

$$\left\{\begin{array}{l} X_k=X_{k-1}+\mu\left[A^T(F-AX_{k-1}B)B^T\right. \\ \qquad\left.+C^T(G-CX_{k-1}D)D^T\right], \\ 0<\mu<\frac{2}{\lambda_{\max}(A^TA)\cdot\lambda_{\max}(BB^T)+\lambda_{\max}(C^TC)\cdot\lambda_{\max}(DD^T)} \\ \qquad \text{or} \quad 0<\mu<\frac{2}{\|A\|^2\cdot\|B\|^2+\|C\|^2\cdot\|D\|^2}, \end{array}\right. \tag{13}$$

yields $\lim_{k\to\infty}X_k=X^*$.

Proof. Applying Lemma 1 to Eq.(12), we have the gradient based iterative algorithm for the equation of (2) described as follows.

$$\mathrm{vec}(X_k)=\mathrm{vec}(X_{k-1})+\mu M^T\left(\left[\begin{array}{c}\mathrm{vec}(F)\\ \mathrm{vec}(G)\end{array}\right]-M\mathrm{vec}(X_{k-1})\right). \tag{14}$$

From (14) and Lemma 2, we can easily obtain

$$\begin{array}{r} X_k=X_{k-1}+\mu\left[A^T(F-AX_{k-1}B)B^T\right. \\ \left.+C^T(G-CX_{k-1}D)D^T\right]. \end{array} \tag{15}$$

By Lemma 5, we know that

$$\begin{array}{rcl} \lambda_{\max}\left(M^TM\right) & = & \lambda_{\max}\left(BB^T\otimes A^TA+DD^T\otimes C^TC\right) \\ & = & \lambda_{\max}(A^TA)\cdot\lambda_{\max}(BB^T) \\ & + & \lambda_{\max}(C^TC)\cdot\lambda_{\max}(DD^T) \\ & \leq & \|A\|^2\cdot\|B\|^2+\|C\|^2\cdot\|D\|^2. \end{array}$$

According to Lemma 1, the proof is complete.

Now, assume that $J\in\mathbf{R}^{m\times q}$ and $L\in\mathbf{R}^{f\times t}$ are arbitrary matrices, then we have

$$\mathrm{vec}(A^TJB^T+C^TLD^T)=M^T\left[\begin{array}{c}\mathrm{vec}(J)\\ \mathrm{vec}(L)\end{array}\right]\subset R(M^T).$$

It is obvious that if we choose

$$X_0 = A^T J B^T + C^T L D^T, \tag{16}$$

where J, L are arbitrary matrices, then all X_k generated by the equation of (15) satisfy

$$\mathrm{vec}(X_k) \subset R(M^T),\ k = 1, 2, \cdots.$$

It follows from Lemma 3 that the equation of (2) has a solution if and only if

$$MM^+ \begin{bmatrix} \mathrm{vec}(F) \\ \mathrm{vec}(G) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(F) \\ \mathrm{vec}(G) \end{bmatrix}. \tag{17}$$

By Lemma 4, we have proved the following result.

Theorem 4: Suppose that the condition (17) is satisfied. If we choose the initial matrix by (16), where J, L are arbitrary matrices, or especially, $X_0 = 0$, then the iterative solution $\{X_k\}$ obtained by the gradient iterative algorithm (13) converges to the unique minimum Frobenius norm solution X^* of Eq.(2).

The proposed algorithm can be applied to the generalized matrix equations:

$$\begin{cases} A_1 X B_1 = F_1, \\ A_2 X B_2 = F_2, \\ \cdots\cdots\cdots\cdots \\ A_s X B_s = F_s. \end{cases} \tag{18}$$

Define $\tilde{M}, \tilde{b}$ as

$$\tilde{M} = \begin{bmatrix} B_1^T \otimes A_1 \\ B_2^T \otimes A_2 \\ \cdots\cdots\cdots \\ B_s^T \otimes A_s \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} \mathrm{vec}(F_1) \\ \mathrm{vec}(F_2) \\ \cdots\cdots \\ \mathrm{vec}(F_s) \end{bmatrix}.$$

Theorem 5: Let $A_i \in \mathbf{R}^{m_i \times n}, B_i \in \mathbf{R}^{p \times q_i}$ and $F_i \in \mathbf{R}^{m_i \times q_i},\ i = 1, 2, \cdots, s$, and suppose that the condition $\tilde{M}\tilde{M}^+\tilde{b} = \tilde{b}$ is satisfied. If we choose the initial matrix $X_0 = \sum_{i=1}^{s} A_i^T J_i B_i^T$, where $J_i,\ i = 1, 2, \cdots, s$, are arbitrary matrices, or especially, $X_0 = 0$, then the gradient based iterative algorithm

$$\begin{cases} X_k = X_{k-1} + \mu \left(\sum_{i=1}^{s} A_i^T (F_i - A_i X_{k-1} B_i) B_i^T \right), \\ 0 < \mu < \frac{2}{\sum_{i=1}^{s} \lambda_{\max}(A_i^T A_i) \cdot \lambda_{\max}(B_i B_i^T)} \\ \quad \text{or } \ 0 < \mu < \frac{2}{\sum_{i=1}^{s} \|A_i\|^2 \cdot \|B_i\|^2}, \end{cases}$$

converges to the unique minimum Frobenius norm solution X^* of Eq.(18).

V. The symmetric solution of the matrix equation $AXB = F$

Using Lemma 2, we know that the equations of (5) are equivalent to

$$N\mathrm{vec}(X) = \begin{bmatrix} \mathrm{vec}(F) \\ \mathrm{vec}(F^T) \end{bmatrix}, \tag{19}$$

where

$$N = \begin{bmatrix} B^T \otimes A \\ A \otimes B^T \end{bmatrix}.$$

Theorem 6: Suppose that $A \in \mathbf{R}^{m \times n}, B \in \mathbf{R}^{n \times q}$ and $F \in \mathbf{R}^{m \times q}$. If the equation of (3) has a unique symmetric solution X^*, then for any initial symmetric matrix X_0, the gradient based iterative algorithm

$$\begin{cases} X_k = X_{k-1} + \mu \left[A^T (F - A X_{k-1} B) B^T \right. \\ \qquad \left. + B(F^T - B^T X_{k-1} A^T) A \right], \\ 0 < \mu < \frac{1}{\lambda_{\max}(A^T A) \cdot \lambda_{\max}(B B^T)} =: \mu_0 \\ \quad \text{or } \ 0 < \mu < \frac{1}{\|A\|^2 \cdot \|B\|^2}, \end{cases} \tag{20}$$

yields $\lim_{k \to \infty} X_k = X^*$.

Proof. Applying Lemma 1 to Eq.(19), we have the gradient based iterative algorithm for the equation of (3) described as follows.

$$\mathrm{vec}(X_k) = \mathrm{vec}(X_{k-1}) + \mu N^T \left(\begin{bmatrix} \mathrm{vec}(F) \\ \mathrm{vec}(F^T) \end{bmatrix} - N\mathrm{vec}(X_{k-1}) \right). \tag{21}$$

From (21) and Lemma 2, we can easily obtain

$$\begin{aligned} X_k = X_{k-1} + \mu \big(& A^T (F - A X_{k-1} B) B^T \\ & + B(F^T - B^T X_{k-1} A^T) A \big). \end{aligned} \tag{22}$$

By Lemma 5, we know that

$$\begin{aligned} \lambda_{\max}\left(N^T N\right) &= \lambda_{\max}\left(B B^T \otimes A^T A + A^T A \otimes B B^T\right) \\ &= 2\lambda_{\max}(A^T A) \cdot \lambda_{\max}(B B^T) \\ &\le 2\|A\|^2 \cdot \|B\|^2. \end{aligned}$$

According to Lemma 1, the proof is complete.

Now, assume that $J \in \mathbf{R}^{m \times q}$ is an arbitrary matrix, then we have

$$\mathrm{vec}(A^T J B^T + B J^T A) = N^T \begin{bmatrix} \mathrm{vec}(J) \\ \mathrm{vec}(J^T) \end{bmatrix} \subset R(N^T).$$

It is obvious that if we choose

$$X_0 = A^T J B^T + B J^T A, \tag{23}$$

where J is an arbitrary matrix, then all X_k generated by the equation of (20) satisfy

$$X_k^T = X_k, \quad \mathrm{vec}(X_k) \subset R(N^T),\ k = 1, 2, \cdots.$$

It follows from Lemma 3 and Lemma 6 that the equation of (3) has a solution if and only if

$$NN^+ \begin{bmatrix} \mathrm{vec}(F) \\ \mathrm{vec}(F^T) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(F) \\ \mathrm{vec}(F^T) \end{bmatrix}. \tag{24}$$

By Lemma 4, we have proved the following result.

Theorem 7: Suppose that the condition (24) is satisfied. If we choose the initial matrix by (23), where J is an arbitrary matrix, or especially, $X_0 = 0$, then the iterative solution $\{X_k\}$ obtained by the gradient iterative algorithm (20) converges to the unique minimum Frobenius norm symmetric solution X^* of Eq.(3).

The proposed algorithm can be used to solve the symmetric solution of the generalized matrix equations:

$$\begin{cases} A_1 X B_1 = F_1, \\ A_2 X B_2 = F_2, \\ \cdots\cdots\cdots\cdots \\ A_s X B_s = F_s, \end{cases} \quad \text{s. t. } \ X^T = X. \tag{25}$$

Define $\tilde{N}, \tilde{g}$ as

$$\tilde{N} = \begin{bmatrix} B_1^T \otimes A_1 \\ A_1 \otimes B_1^T \\ B_2^T \otimes A_2 \\ A_2 \otimes B_2^T \\ \cdots\cdots\cdots \\ B_s^T \otimes A_s \\ A_s \otimes B_s^T \end{bmatrix}, \quad \tilde{g} = \begin{bmatrix} \mathrm{vec}(F_1) \\ \mathrm{vec}(F_1^T) \\ \mathrm{vec}(F_2) \\ \mathrm{vec}(F_2^T) \\ \cdots\cdots \\ \mathrm{vec}(F_s) \\ \mathrm{vec}(F_s^T) \end{bmatrix}.$$

Theorem 8: Let $A_i \in \mathbf{R}^{m_i \times n}, B_i \in \mathbf{R}^{n \times q_i}$ and $F_i \in \mathbf{R}^{m_i \times q_i}$, $i = 1, 2, \cdots, s$, and suppose that the condition $\tilde{N}\tilde{N}^+\tilde{g} = \tilde{g}$ is satisfied. If we choose the initial matrix $X_0 = \sum_{i=1}^s (A_i^T J_i B_i^T + B_i J_i^T A_i)$, where $J_i,\ i = 1, 2, \cdots, s$, are arbitrary matrices, or especially, $X_0 = 0$, then the gradient based iterative algorithm

$$\begin{cases} X_k = X_{k-1} + \mu \left(\sum_{i=1}^s A_i^T (F_i - A_i X_{k-1} B_i) B_i^T \right. \\ \qquad \left. + \sum_{i=1}^s B_i (F_i^T - B_i^T X_{k-1} A_i^T) A_i \right), \\ 0 < \mu < \frac{1}{\sum_{i=1}^s \lambda_{\max}(A_i^T A_i) \cdot \lambda_{\max}(B_i B_i^T)} \\ \quad \text{or} \ \ 0 < \mu < \frac{1}{\sum_{i=1}^s \|A_i\|^2 \cdot \|B_i\|^2}, \end{cases}$$

converges to the unique minimum Frobenius norm symmetric solution X^* of Eq.(25).

VI. The solution of the matrix equation $AXB + CYD = F$

Using Lemma 2, we know that the equation of (4) is equivalent to

$$P \begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(Y) \end{bmatrix} = \mathrm{vec}(F), \qquad (26)$$

where

$$P = \begin{bmatrix} B^T \otimes A & D^T \otimes C \end{bmatrix}.$$

Theorem 9: Suppose that $A \in \mathbf{R}^{m \times n}, B \in \mathbf{R}^{p \times q},\ C \in \mathbf{R}^{m \times e}, D \in \mathbf{R}^{h \times q}$ and $F \in \mathbf{R}^{m \times q}$. If the equation of (4) has a unique solution pair (X^*, Y^*), then for any initial matrices X_0 and Y_0, the gradient based iterative algorithm

$$\begin{cases} X_k = X_{k-1} + \mu \left[A^T (F - AX_{k-1}B - CY_{k-1}D) B^T \right], \\ Y_k = Y_{k-1} + \mu \left[C^T (F - AX_{k-1}B - CY_{k-1}D) D^T \right], \\ 0 < \mu < \frac{2}{\lambda_{\max}(AA^T) \cdot \lambda_{\max}(B^T B) + \lambda_{\max}(CC^T) \cdot \lambda_{\max}(D^T D)} \\ \text{or} \ \ 0 < \mu < \frac{2}{\|A\|^2 \cdot \|B\|^2 + \|C\|^2 \cdot \|D\|^2}, \end{cases} \qquad (27)$$

yields $\lim_{k \to \infty} X_k = X^*$ and $\lim_{k \to \infty} Y_k = Y^*$.

Proof. Applying Lemma 1 to Eq.(26), we have the gradient based iterative algorithm for the equation of (4) described as follows.

$$\begin{bmatrix} \mathrm{vec}(X_k) \\ \mathrm{vec}(Y_k) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(X_{k-1}) \\ \mathrm{vec}(Y_{k-1}) \end{bmatrix} + \mu P^T \left(\mathrm{vec}(F) - P \begin{bmatrix} \mathrm{vec}(X_{k-1}) \\ \mathrm{vec}(Y_{k-1}) \end{bmatrix} \right). \qquad (28)$$

From (28) and Lemma 2, we can easily obtain

$$X_k = X_{k-1} + \mu \left[A^T (F - AX_{k-1}B - CY_{k-1}D) B^T \right], \qquad (29)$$

$$Y_k = Y_{k-1} + \mu \left[C^T (F - AX_{k-1}B - CY_{k-1}D) D^T \right]. \qquad (30)$$

By Lemma 5, we know that

$$\begin{aligned} \lambda_{\max}\left(P^T P\right) &= \lambda_{\max}\left(PP^T\right) \\ &= \lambda_{\max}\left(B^T B \otimes AA^T + D^T D \otimes CC^T\right) \\ &= \lambda_{\max}(AA^T) \cdot \lambda_{\max}(B^T B) \\ &+ \lambda_{\max}(CC^T) \cdot \lambda_{\max}(D^T D) \\ &\leq \|A\|^2 \cdot \|B\|^2 + \|C\|^2 \cdot \|D\|^2. \end{aligned}$$

According to Lemma 1, the proof is complete.

Now, assume that $J \in \mathbf{R}^{m \times q}$ is an arbitrary matrix, then we have

$$\begin{bmatrix} \mathrm{vec}(A^T J B^T) \\ \mathrm{vec}(C^T J D^T) \end{bmatrix} = P^T \mathrm{vec}(J) \subset R(P^T).$$

It is obvious that if we choose

$$X_0 = A^T J B^T, \quad Y_0 = C^T J D^T, \qquad (31)$$

where J is an arbitrary matrix, then all X_k and Y_k generated by the equations of (29) and (30) satisfy

$$\begin{bmatrix} \mathrm{vec}(X_k) \\ \mathrm{vec}(Y_k) \end{bmatrix} \subset R(P^T),\ k = 1, 2, \cdots.$$

It follows from Lemma 3 that the equation of (4) has a solution if and only if

$$PP^+ \mathrm{vec}(F) = \mathrm{vec}(F). \qquad (32)$$

By Lemma 4, we have proved the following result.

Theorem 10: Suppose that the condition (32) is satisfied. If we choose the initial matrices by (31), where J is an arbitrary matrix, or especially, $X_0 = 0, Y_0 = 0$, then the iterative solution $\{X_k\}$ and $\{Y_k\}$ obtained by the gradient iterative algorithm (27) converges to the unique minimum Frobenius norm solution (X^*, Y^*) of Eq.(4).

The proposed algorithm can be applied to the generalized matrix equation:

$$\sum_{i=1}^s A_i X_i B_i = F, \qquad (33)$$

where $A_i \in \mathbf{R}^{m \times n_i}, B_i \in \mathbf{R}^{p_i \times q},\ i = 1, 2, \cdots$, and $F \in \mathbf{R}^{m \times q}$ are known matrices. Define $\tilde{P}$ as

$$\tilde{P} = \begin{bmatrix} B_1^T \otimes A_1, & B_2^T \otimes A_2, & \cdots, & B_s^T \otimes A_s \end{bmatrix}.$$

Theorem 11: Let $A_i \in \mathbf{R}^{m \times n_i}, B_i \in \mathbf{R}^{p_i \times q},\ i = 1, 2, \cdots$, and $F \in \mathbf{R}^{m \times q}$. Suppose that the condition $\tilde{P}\tilde{P}^+ \mathrm{vec}(F) = \mathrm{vec}(F)$ is satisfied. If we choose the initial matrix $X_i^{(0)} = A_i^T J B_i^T,\ i = 1, 2, \cdots, s$, where J is an arbitrary matrix, or especially, $X_i^{(0)} = 0,\ i = 1, 2, \cdots, s$, then the gradient based iterative algorithm

$$\begin{cases} X_i^{(k)} = X_i^{(k-1)} + \mu \left[A_i^T (F - \sum_{i=1}^s A_i X_i^{(k-1)} B_i) B_i^T \right], \\ \ i = 1, 2, \cdots, s, \\ 0 < \mu < \frac{2}{\sum_{i=1}^s \lambda_{\max}(A_i A_i^T) \cdot \lambda_{\max}(B_i^T B_i)} \\ \text{or} \ \ 0 < \mu < \frac{2}{\sum_{i=1}^s \|A_i\|^2 \cdot \|B_i\|^2}, \end{cases}$$

converges to the unique minimum Frobenius norm solution $(X_1^*, X_2^*, \cdots, X_s^*)$ of Eq.(33).

TABLE I
THE ITERATIVE SOLUTION ($\mu = 0.047$)

k	x_{11}	x_{12}	x_{22}	δ
1	1.4229	1.0989	2.1689	0.3487
2	1.6742	0.6781	1.3437	0.1282
10	1.8999	0.7000	1.5999	5.6296e-005
11	1.9000	0.7000	1.6000	2.1895e-005
19	1.9000	0.7000	1.6000	1.2411e-008
20	1.9000	0.7000	1.6000	4.9009e-009
Solution	1.9	0.7	1.6	

TABLE II
THE ITERATIVE SOLUTION ($\mu = 1/240$)

k	x_{11}	x_{12}	x_{22}	r
1	0.7782	1.0239	0.9829	4.2224
2	0.9198	1.2102	1.1618	0.7686
5	0.9511	1.2515	1.2014	0.0046
6	0.9513	1.2517	1.2016	8.4356e-004
11	0.9513	1.2517	1.2016	1.6853e-007
12	0.9513	1.2517	1.2016	3.0675e-008
13	0.9513	1.2517	1.2016	5.5833e-009

VII. TWO NUMERICAL EXAMPLES

In this section, we will give two numerical examples to illustrate the proposed algorithms and the test is performed using MATLAB 6.5.

Example 1. Consider the following matrix equation:

$$AXB = F \text{ s. t. } X^T = X$$

with

$$A = \begin{bmatrix} 0.95 & -0.6 \\ 0.9 & 1.2 \end{bmatrix},$$

$$B = \begin{bmatrix} 1.13 & -1.72 \\ -2.26 & -1.44 \end{bmatrix},$$

$$F = \begin{bmatrix} 2.2317 & -1.9574 \\ -2.8815 & -8.058 \end{bmatrix}.$$

We can easily see that the equation has unique solution and the exact solution is

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} = \begin{bmatrix} 1.9 & 0.7 \\ 0.7 & 1.6 \end{bmatrix}.$$

Taking $X_0 = 0$, we apply the gradient based algorithm in (20) to compute $\{X_k\}$. Fig.1 shows the effect of changing μ on the iterative steps. From the plot, an optimum value of μ may be obtained and a good compromise value of μ would be $\mu = 0.047$. With which μ, the iterative solutions X_k are shown in Table 1, where $\delta := \|X_k - X\|/\|X\|$ is the relative error. The errors δ versus k with different convergence factors are shown in Fig.2. From Table 1 and Fig.2, it is clear that the error δ becomes smaller and smaller and goes to zero within several iterations. This indicates that the gradient based iterative algorithm is effective.

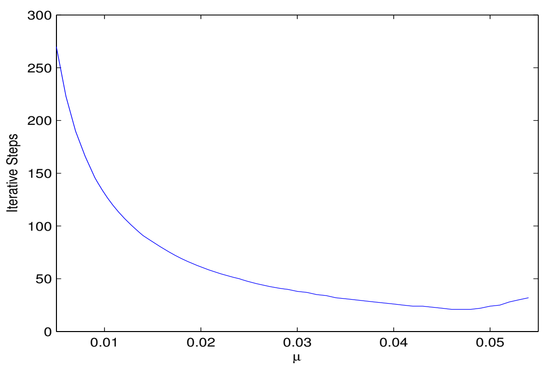

Fig.1 Variation of the iterative steps versus μ

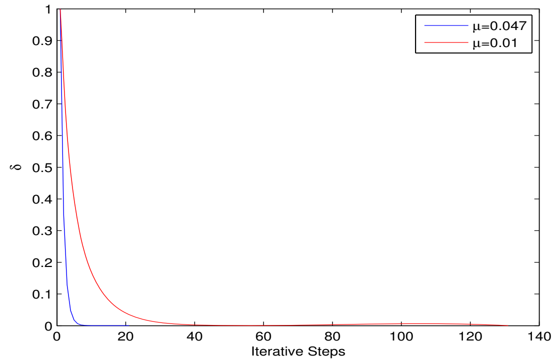

Fig.2 The relative errors δ versus k of the gradient based algorithm

Example 2. Consider the matrix equation $AXB = F$ s. t. $X^T = X$ with

$$A = \begin{bmatrix} 0.95 & 0.60 \\ 2.85 & 1.80 \end{bmatrix},$$

$$B = \begin{bmatrix} 1.13 & 0.72 \\ 2.26 & 1.44 \end{bmatrix},$$

$$F = \begin{bmatrix} 6.1867 & 3.942 \\ 18.56 & 11.826 \end{bmatrix}.$$

Observe that the equation has many solutions, that is, the solution is not unique. Choosing initial iterative matrix $X_0 = 0$, we apply the gradient based algorithm in (20) to compute $\{X_k\}$. The iterative solutions X_k are shown in Table 2, where $r := \|F - AX_kB\|$. This implies that the algorithm in (20) can be used to solve the minimum norm symmetric solution of the equation $AXB = F$.

VIII. CONCLUDING REMARKS

This paper presents gradient based iterative algorithms for solving some linear matrix equations. The analysis indicates that if the equation considered has a solution, then the iterative solutions given by the gradient based iterative algorithm converges fast to its exact solution or the unique minimum norm solution by choosing a special kind of initial matrices. The approach is demonstrated by two numerical examples and reasonable results are produced.

ACKNOWLEDGMENT

This research was supported by the National Natural Science Foundation of China (No.10926130).

REFERENCES

[1] A. Kilicman, Z. Al Zhour, Vector least-squares solutions for coupled singular matrix equations, Journal of Computational and Applied Mathematics, 206 (2007) 1051-1069.

[2] F. Ding, L. Qiu, T. Chen, Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems, Automatica, 45 (2009) 324-332.

[3] F. Ding, T. Chen, Performance analysis of multi-innovation gradient type identification methods, Automatica, 43 (2007) 1-14.

[4] F. Ding, P. X. Liu, G. Liu, Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises, Signal Processing, 89 (2009) 1883-1890.

[5] F. Ding, P. X. Liu, H. Z. Yang, Parameter identification and intersample output estimation for dual-rate systems, IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 38 (2008) 966-975.

[6] F. Ding, H. Z. Yang, F. Liu, Performance analysis of stochastic gradient algorithms under weak conditions, Science in China Series F-Information Sciences, 51 (2008) 1269-1280.

[7] M. Dehghan, M. Hajarian, An iterative algorithm for the reflexive solutions of the generalized coupled Sylvester matrix equations and its optimal approximation, Applied Mathematics and Computation, 202 (2008) 571-588.

[8] H. Mukaidani, S. Yamamoto, T. Yamamoto, A numerical algorithm for finding solution of cross-coupled algebraic Riccati equations, IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences, E91A (2008) 682-685.

[9] B. Zhou, G. R. Duan, Solutions to generalized Sylvester matrix equation by Schur decomposition, International Journal of Systems Science, 38 (2007) 369-375.

[10] B. Zhou, G. R. Duan, On the generalized Sylvester mapping and matrix equations, Systems & Control Letters 57 (3) (2008) 200-208.

[11] F. Ding, T. Chen, Iterative least squares solutions of coupled Sylvester matrix equations, Systems & Control Letters, 54 (2005) 95-107.

[12] F. Ding, T. Chen, On iterative solutions of general coupled matrix equations, SIAM Journal on Control and Optimization, 44 (2006) 2269-2284.

[13] F. Ding, P. X. Liu, J. Ding. Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle. Applied Mathematics and Computation, 197 (2008) 41-50.

[14] J. Ding, Y. Liu, F. Ting, Iterative solutions to matrix equations of the form $A_iXB_i = F_i$, Computers and Mathematics with Applications, 59 (2010) 3500-3507.

[15] P. Lancaster, M. Tismenetsky, The Theory of Matrices. 2rd Edition. London: Academic Press, 1985.

[16] A. Ben-Israel, T. N. E. Greville, Generalized Inverses. Theory and Applications (second ed). New York: Springer, 2003.

[17] H. Dai, The Theory of Matrices. Beijing: Science Press, 2002.