

Iterative solutions to the linear matrix equation

$$AXB + CX^T D = E$$

Yongxin Yuan, Jiashang Jiang

Abstract—In this paper the gradient based iterative algorithm is presented to solve the linear matrix equation $AXB + CX^T D = E$, where X is unknown matrix, A, B, C, D, E are the given constant matrices. It is proved that if the equation has a solution, then the unique minimum norm solution can be obtained by choosing a special kind of initial matrices. Two numerical examples show that the introduced iterative algorithm is quite efficient.

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

I. INTRODUCTION

LINEAR matrix equations play an important role in linear system theory. For example, the Sylvester matrix equation $AX + XB = C$ can be used to solve many control problems such as pole assignment[1], robust pole assignment[2], eigenstructure assignment[3] and fault detection. Its special forms include the well-known Lyapunov matrix equation $AX + XA^T = -C$ which has very important applications in stability analysis of linear systems[4]. Due to their wide applications, over the past several decades, the problem of searching for analytical and numerical solutions to Sylvester-class equations has been well investigated in the literature, for example, [5-9] and the references therein. Though analytical solutions are theoretically appealing, they may suffer computing problems in practice[10]. On the other hand, for some problems such as stability analysis and robustness analysis, approximate solutions are sufficient[7]. Furthermore, if the coefficient matrices of the matrix equations have uncertainties, it is generally impossible to obtain analytical solutions[7,8,11,12]. Therefore, it is important to investigate numerical solutions to Sylvester-class matrix equations.

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

$$AXB + CX^T D = E, \quad (1)$$

$$\sum_{i=1}^s A_i X B_i + \sum_{j=1}^t C_j X^T D_j = E, \quad (2)$$

Yongxin Yuan: School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, P R China. e-mail: yuanyx_703@163.com

Jiashang Jiang: School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, P R China. e-mail: jjiasang@163.com

where $X \in \mathbf{R}^{m \times n}$ is unknown matrix, $A \in \mathbf{R}^{p \times m}, B \in \mathbf{R}^{n \times q}, C \in \mathbf{R}^{p \times n}, D \in \mathbf{R}^{m \times q}, E \in \mathbf{R}^{p \times q}$ and $A_i \in \mathbf{R}^{p \times m}, B_i \in \mathbf{R}^{n \times q}, C_j \in \mathbf{R}^{p \times n}, D_j \in \mathbf{R}^{m \times q}$ for $i = 1, \dots, s; j = 1, \dots, t$ are the given constant matrices. We observe that Xie et al.[13, 14] have considered the iterative solutions of Eqs.(1) and (2) by using the hierarchical identification principle, 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 the gradient based iterative algorithms to solve Eqs.(1) and (2) 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. 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 $(A, B) = \text{trace}(B^T A)$, 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 $A = [a_1, a_2, \dots, a_n]$, where $a_i, i = 1, \dots, n$, is the i -th column vector of A , the stretching function $\text{vec}(A)$ is defined as $\text{vec}(A) = [a_1^T, a_2^T, \dots, a_n^T]^T$.

II. THE ITERATIVE SOLUTIONS TO THE MATRIX EQUATIONS (1) AND (2)

To begin with, we first give some lemmas.

Lemma 1: [7, 15, 16, 17]. 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

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

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

Lemma 2: [18]. 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: [19]. 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: [19]. 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: [18]. Let $Z \in \mathbf{R}^{n \times m}$ be any matrix. Then

$$P(m, n)\text{vec}(Z^T) = \text{vec}(Z),$$

where $P(m, n)$ is uniquely determined by the integers m and n . Moreover, the matrix $P(m, n)$ is of the following properties.
a). For two arbitrary integers m and n , $P(m, n)$ has the following explicit form

$$P(m, n) = \begin{bmatrix} E_{11}^T & E_{12}^T & \cdots & E_{1n}^T \\ E_{21}^T & E_{22}^T & \cdots & E_{2n}^T \\ \vdots & \vdots & \ddots & \vdots \\ E_{m1}^T & E_{m2}^T & \cdots & E_{mn}^T \end{bmatrix}_{mn \times mn},$$

where E_{ij} , $i = 1, \dots, m$; $j = 1, \dots, n$ is an $m \times n$ matrix with the element at position (i, j) being 1 and the others being 0.
b). $P(m, n)$ is an orthogonal matrix, i.e.,

$$P(m, n)P^T(m, n) = P^T(m, n)P(m, n) = I_{mn}.$$

c). Let m, n, p and q be four integers and $A \in \mathbf{R}^{m \times n}$, $B \in \mathbf{R}^{p \times q}$. Then

$$P(m, p)(B \otimes A) = (A \otimes B)P(n, q).$$

Using Lemma 2 and Lemma 5, we know that the equation of (1) is equivalent to

$$M\text{vec}(X) = \text{vec}(E), \quad (3)$$

where

$$M = B^T \otimes A + (D^T \otimes C)P(m, n).$$

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

$$\begin{cases} X_k = X_{k-1} + \mu [A^T(E - AX_{k-1}B - CX_{k-1}^T D)B^T \\ \quad + D(E^T - B^T X_{k-1}^T A^T - D^T X_{k-1} C^T)C], \\ 0 < \mu < \frac{1}{\|A\|^2 \cdot \|B\|^2 + \|C\|^2 \cdot \|D\|^2}, \end{cases}$$

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

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

$$\text{vec}(X_k) = \text{vec}(X_{k-1}) + \mu M^T (\text{vec}(E) - M\text{vec}(X_{k-1})). \quad (5)$$

From (5), Lemma 2 and Lemma 5, we obtain

$$\begin{aligned} \text{vec}(X_k) &= \text{vec}(X_{k-1}) + \mu [(B \otimes A^T)\text{vec}(E) \\ &\quad + P(m, n)(D \otimes C^T)\text{vec}(E) - (BB^T \otimes A^T A)\text{vec}(X_{k-1}) \\ &\quad - (BD^T \otimes A^T C)P(m, n)\text{vec}(X_{k-1}) \\ &\quad - P(m, n)(DB^T \otimes C^T A)\text{vec}(X_{k-1}) \\ &\quad - P(m, n)(DD^T \otimes C^T C)P(m, n)\text{vec}(X_{k-1})] \\ &= \text{vec}(X_{k-1}) + \mu [\text{vec}(A^T E B^T) + \text{vec}(D E^T C) \\ &\quad - \text{vec}(A^T A X_{k-1} B B^T) - \text{vec}(A^T C X_{k-1}^T D B^T) \\ &\quad - \text{vec}(D B^T X_{k-1}^T A^T C) - \text{vec}(D D^T X_{k-1} C^T C)]. \end{aligned}$$

Thus we have

$$X_k = X_{k-1} + \mu [A^T(E - AX_{k-1}B - CX_{k-1}^T D)B^T + D(E^T - B^T X_{k-1}^T A^T - D^T X_{k-1} C^T)C], \quad (6)$$

Observe that

$$\begin{aligned} \|M\|^2 &\leq (\|B^T \otimes A\| + \|(D^T \otimes C)P(m, n)\|)^2 \\ &= (\|B^T \otimes A\| + \|D^T \otimes C\|)^2 \\ &= (\|B\|\|A\| + \|D\|\|C\|)^2 \\ &\leq 2(\|A\|^2\|B\|^2 + \|C\|^2\|D\|^2). \end{aligned}$$

According to Lemma 1, Theorem 1 is proven.

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

$$\begin{aligned} &\text{vec}(A^T J B^T + D J^T C) \\ &= (B \otimes A^T + P(m, n)(D \otimes C^T))\text{vec}(J) \\ &\subset R(B \otimes A^T + P(m, n)(D \otimes C^T)) \\ &= R(M^T). \end{aligned}$$

It is obvious that if we choose

$$X_0 = A^T J B^T + D J^T C, \quad (7)$$

where J is an arbitrary matrix, then all X_k generated by the equation of (6) satisfy

$$\text{vec}(X_k) \subset R(M^T), \quad k = 1, 2, \dots$$

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

$$MM^+ \text{vec}(E) = \text{vec}(E). \quad (8)$$

By Lemma 4, we have proved the following result.

Theorem 2: Suppose that the condition (8) is satisfied. If we choose the initial matrix by (7), where J is an arbitrary matrix, or especially, $X_0 = 0$, then the iterative solution X_k obtained by the gradient iterative algorithm (4) converges to the unique minimum Frobenius norm solution X^* of Eq.(1).

The proposed algorithm can be applied to the generalized matrix equation (2).

Define \tilde{M} as

$$\tilde{M} = \sum_{i=1}^s (B_i^T \otimes A_i) + \sum_{j=1}^t (D_j^T \otimes C_j) P(m, n).$$

Theorem 3: Let $E \in \mathbf{R}^{p \times q}$ and $A_i \in \mathbf{R}^{p \times m}, B_i \in \mathbf{R}^{n \times q}, C_j \in \mathbf{R}^{p \times n}, D_j \in \mathbf{R}^{m \times q}$ for $i = 1, \dots, s; j = 1, \dots, t$, and suppose that the condition $\tilde{M} \tilde{M}^+ \text{vec}(E) = \text{vec}(E)$ is satisfied. If we choose the initial matrix $X_0 = \sum_{i=1}^s A_i^T J B_i^T + \sum_{j=1}^t D_j J^T C_j$, where $J \in \mathbf{R}^{p \times q}$ is an arbitrary matrix, or especially, $X_0 = 0$, then the gradient based iterative algorithm

$$\begin{cases} X_k = X_{k-1} + \mu \left[\sum_{i=1}^s A_i^T \Phi B_i^T + \sum_{j=1}^t D_j \Phi^T C_j \right], \\ \Phi = E - \sum_{l=1}^s A_l X_{k-1} B_l - \sum_{r=1}^t C_r X_{k-1}^T D_r, \\ 0 < \mu < \frac{1}{\sum_{i=1}^s \|A_i\|^2 \cdot \|B_i\|^2 + \sum_{j=1}^t \|C_j\|^2 \cdot \|D_j\|^2}, \end{cases}$$

converges to the unique minimum Frobenius norm solution X^* of Eq.(2).

III. 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 matrix equation $AXB + CX^T D = E$ with

$$A = \begin{bmatrix} 2 & 5 \\ 4 & -7 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & -3 \\ 1 & 2 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

and

$$E = \begin{bmatrix} 317 & 9 \\ 41 & 27 \end{bmatrix}.$$

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

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 7 & 5 \\ 4 & 3 \end{bmatrix}.$$

Taking $X_0 = 0$, we apply the gradient based algorithm in (4) to compute X_k . Choose $\mu = 2.4678e - 004$, the iterative solutions X_k are shown in Table 1, where $\delta := \|X_k - X\|/\|X\|$ is the relative error. The relative error δ versus k with $\mu = 2.4678e - 004$ is shown in Fig.1.

From Table 1 and Fig.1, 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.

TABLE I
THE ITERATIVE SOLUTION ($\mu = 2.4678e - 004$)

k	x_{11}	x_{12}	x_{21}	x_{22}	δ
1	1.3474	1.0797	2.1603	0.6473	0.7537
10	6.4796	4.6498	3.8452	2.7616	0.0692
20	6.9669	4.9763	3.9896	2.9725	0.0050
40	7.0015	4.9995	4.0002	2.9964	3.9496e-004
60	7.0005	4.9998	4.0001	2.9989	1.2366e-004
63	7.0004	4.9999	4.0001	2.9991	1.0393e-004
64	7.0004	4.9999	4.0001	2.9991	9.8082e-005
Solution	7	5	4	3	

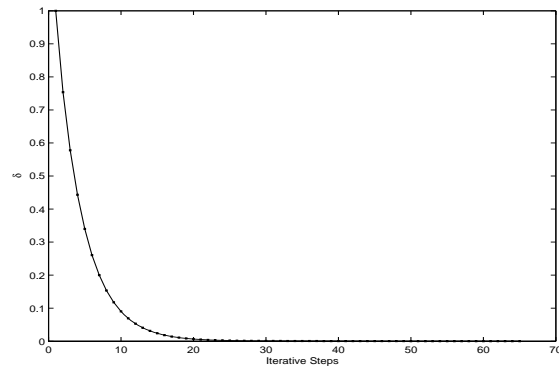


Fig. 1. The relative errors δ versus k of the gradient based algorithm

Example 2. Consider the matrix equation $AXB + CX^T D = E$ with

$$A = \begin{bmatrix} 2 & 1 \\ -3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -9 \\ 1 & -3 \end{bmatrix},$$

$$C = \begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix}$$

and

$$E = \begin{bmatrix} 14 & 0 \\ -28 & 0 \end{bmatrix}.$$

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 (4) to compute X_k . The iterative solutions X_k are shown in Table 2, where $r := \|E - AX_k B - CX_k^T D\|$. The error r versus k with $\mu = 6.7114e - 004$ is shown in Fig.2. This implies that the algorithm in (4) can be used to solve the minimum norm solution of the equation $AXB + CX^T D = E$.

REFERENCES

- [1] G. R. Duan. Solutions to matrix equation $AV + BW = VF$ and their application to eigenstructure assignment in linear systems. IEEE Transactions on Automatic Control, 38 (1993) 276-280.
- [2] S. P. Bhattacharyya, E. De Souza. Pole assignment via Sylvester's equation. Systems and Control Letters, 1 (1972) 261-263.

TABLE II
THE ITERATIVE SOLUTION ($\mu = 6.7114e - 004$)

k	x_{11}	x_{12}	x_{21}	x_{22}	r
1	0.2067	-0.0564	-0.0940	-0.0940	29.0779
10	0.8592	-0.9805	-0.3077	-0.7360	16.1807
50	1.8897	-1.9859	0.0174	-1.3021	2.0102
100	2.0280	-2.1093	0.1115	-1.3555	0.1614
150	2.0390	-2.1191	0.1193	-1.3596	0.0130
161	2.0395	-2.1195	0.1196	-1.3598	0.0075
162	2.0395	-2.1195	0.1196	-1.3598	0.0071

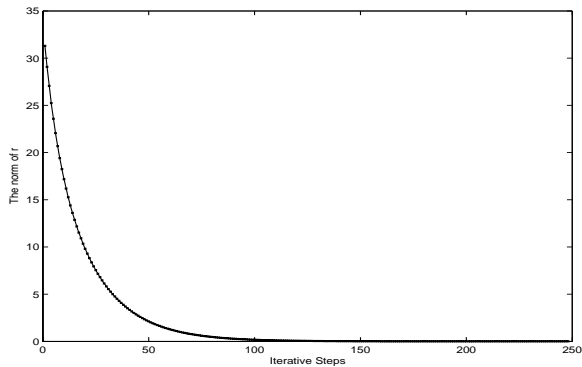


Fig. 2. The errors r versus k of the gradient based algorithm

equations. SIAM Journal on Control and Optimization, 44 (2006) 2269-2284.

- [16] 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.
- [17] J. Ding, Y. Liu, F. Ting. Iterative solutions to matrix equations of the form $A_i X B_i = F_i$. Computers and Mathematics with Applications, 59 (2010) 3500-3507.
- [18] P. Lancaster, M. Tismenetsky. The Theory of Matrices. 2nd Edition. London: Academic Press, 1985.
- [19] A. Ben-Israel, T. N. E. Greville. Generalized Inverses. Theory and Applications (second ed). New York: Springer, 2003.

- [3] G. R. Duan. On the solution to Sylvester matrix equation $AV + BW = EVF$. IEEE Transactions on Automatic Control, 41 (1996) 612-614.
- [4] K. Zhou, J. Doyle, K. Glover. Robust and optimal control. Prentice-Hall, 1996.
- [5] B. Zhou, G. R. Duan. An explicit solution to the matrix equation $AX - XF = BY$. Linear Algebra and its Applications, 402 (2005) 345-366.
- [6] B. Zhou, G. R. Duan. A new solution to the generalized Sylvester matrix equation $AV - EVF = BW$. Systems & Control Letters, 55 (2006) 193-198.
- [7] F. Ding, T. Chen. Iterative least squares solutions of coupled Sylvester matrix equations. Systems & Control Letters, 54 (2005) 95-107.
- [8] H. Mukaidani, H. Xu, K. Mizukami. New iterative algorithm for algebraic Riccati equation related to H_∞ control problem of singularly perturbed systems. IEEE Transactions on Automatic Control, 46 (2001) 1659-1666.
- [9] I. Borno, Z. Gajic. Parallel algorithm for solving coupled algebraic Lyapunov equations of discrete-time jump linear systems. Computers & Mathematics with Applications, 30 (1995) 1-4.
- [10] C. T. Kelley. Iterative Methods for Linear and Nonlinear Equations. SIAM, Philadelphia, 1995.
- [11] T. Mori, A. Derese. A brief summary of the bounds on the solution of the algebraic matrix equations in control theory. International Journal of Control, 39 (1984) 247-256.
- [12] M. Mrabti, M. Benseddik. Unified type non-stationary Lyapunov matrix equations—simultaneous eigenvalue bounds. Systems & Control Letters, 18 (1995) 73-81.
- [13] L. Xie, Y. Liu, H. Yang. Gradient based and least squares based iterative algorithms for matrix equations $AXB + CX^T D = F$. Applied Mathematics and Computation, 217 (2010) 2191-2199.
- [14] L. Xie, J. Ding, F. Ding. Gradient based iterative solutions for general linear matrix equations. Computers and Mathematics with Applications, 58 (2009) 1441-1448.
- [15] F. Ding, T. Chen. On iterative solutions of general coupled matrix