Journal of Mathematical Research with Applications
May, 2017, Vol. 37, No. 3, pp. 351–366
DOI:10.3770/j.issn:2095-2651.2017.03.013
Http://jmre.dlut.edu.cn

Gradient Based Iterative Solutions for Sylvester-Conjugate Matrix Equations

Hailong SHEN, Cheng PENG, Xinhui SHAO*, Tie ZHANG
Department of Mathematics, Northeastern University, Liaoning 110819*, P. R. China*

Abstract This paper presents a gradient based iterative algorithm for Sylvester-conjugate matrix equations with a unique solution. By introducing a relaxation parameter and applying the hierarchical identification principle, an iterative algorithm is constructed to solve Sylvester matrix equations. By applying a real representation of a complex matrix as a tool and using some properties of the real representation, convergence analysis indicates that the iterative solutions converge to the exact solutions for any initial values under certain assumptions. Numerical examples are given to illustrate the efficiency of the proposed approach.

Keywords Sylvester-conjugate matrix equations; iterative solutions; convergence; relaxation parameter

MR(2010) Subject Classification 65F08; 65F10

1. Introduction

In the matrix algebra fields, iterative approaches for solving matrix equations and recursive identifications have attracted a lot of attention from many researchers since Huang proposed an iterative method for solving the linear matrix equation over skew-symmetric matrix. Ding derived iterative solutions of various Sylvester matrix equations by extending the well-known Jacobi and Gauss Seidel iterations for in [1–4]. In [5], Niu proposed a relaxed gradient based on algorithm for solving Sylvester equations by introducing a relaxation parameter. In [6,7] and [8–11], Song and Wu respectively gave iterative algorithms to solve Sylvester-conjugate and Sylvester-transpose equations. In [12,13], Xie gave different algorithms for Sylvester matrix equations. In [14], Wang constructed an iterative algorithm of generalized Sylvester matrix equation. Efficiently numerical algorithms were presented based on the hierarchical identification principle which regards the unknown matrix as the system parameter matrix to be identified.

In this paper, we firstly discuss the Sylvester matrix equation $AXB + C\overline{X}D = F$ with the unknown matrix X. Obviously, the matrix equations $AX - \overline{X}B = C$ and $X - A\overline{X}B = F$ are included as special cases. Due to these facts, $AXB + C\overline{X}D = F$ can be regarded as a general form of Sylvester matrix equation. By introducing a relaxation parameter and applying the hierarchical identification principle, we propose a gradient based algorithm with a relaxation

Received May 4, 2016; Accepted November 23, 2016
Supported by the National Natural Science Foundation of China (Grant No. 11071033).
* Corresponding author
E-mail address: xinhui1002@126.com (Xinhui SHAO)

parameter for finding the iterative solution to the Sylvester-conjugate matrix equation and give the convergence properties of the algorithm. By applying a real representation of a complex matrix as a tool and using some properties of the real representation, the range of the convergence factor is provided to guarantee that the iterative solution converges to the exact solution. No matter what value of parameter is determined in its range, it is different from the algorithm given by Wu in [10]. In addition, a more general case is also considered. Convergence analysis and numerical examples are given to illustrate the effectiveness of the algorithm.

Through out this paper, $I[i,j]$ denotes the set $[i,i+1,\dots,j]$ for two integers $i<j$. The symbols $\overline{A}$, A^T, A^H, $\mathrm{tr}(A)$, $\|A\|$, $\|A\|_2$ denote the transpose, the trace, the Frobenius norm. For a matrix $X=[x_1x_2\cdots x_n]\in\mathbb{R}^{m\times n}$, vec(X) is the column stretching operation of X,

$$\mathrm{vec}(X)=[x_1^Tx_2^T\cdots x_n^T]^T.$$

For two matrices M and N, $M\in\mathbb{R}^{r\times m}$, $M\in\mathbb{R}^{n\times s}$, MXN is their Kronecker product. For matrices M, N, X with appropriate dimension,

$$\mathrm{vec}(MXN)=(N^T\otimes M)\mathrm{vec}(X)$$

and $\|A\|=\|\mathrm{vec}(A)\|$ for an arbitrary matrix A.

2. Preliminaries

In this section, some important conclusions which play important roles in the following section are provided.

Lemma 2.1 ([1]) *Consider the equation*

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

where $A\in\mathbb{C}^{p\times m}$, $B\in\mathbb{C}^{n\times q}$, $F\in\mathbb{C}^{p\times q}$ are the given known matrices and $X\in\mathbb{C}^{m\times n}$ is unknown to be determined.

The iterative algorithm of the equation is written as

$$X(k+1)=X(k)+\mu A^H(F-AX(k)B)B^H \quad \text{with} \quad 0<\mu<\frac{2}{\|A\|_2^2\|B\|_2^2}.$$

If (1) has the unique solution X_*, then iterative solution $X(k)$ converges to the unique solution, that is $\lim_{k\to\infty}X(k)=X_*$.

Now, we introduce a real representation of a complex matrix. The concept was firstly used in [14]. Let $A\in\mathbb{C}^{m\times n}$. Then A can be uniquely written as $A=A_1+A_2i$ with $A_1,A_2\in\mathbb{R}^{m\times n}$, $i=\sqrt{-1}$. Define real representation as

$$A_\sigma=\begin{bmatrix}A_1 & A_2\\ A_2 & -A_1\end{bmatrix}\in\mathbb{R}^{2m\times 2n},$$

A_σ is called the real representation of the matrix A. For an $n\times n$ complex matrix A, define $A_\sigma^i=(A_\sigma)^i$, and

$$P_j=\begin{bmatrix}I_j & 0\\ 0 & I_j\end{bmatrix},\quad Q_j=\begin{bmatrix}0 & I_j\\ -I_j & 0\end{bmatrix},$$

where I_j is the $j \times j$ identity matrix. The real representation possesses the following properties.

Lemma 2.2 ([14]) *The properties of the real representation*

(i) If $A, B \in \mathbb{C}^{m\times n}$, $a \in \mathbb{R}$, *then*

$$\begin{cases} (A+B)_\sigma = A_\sigma + B_\sigma, \\ (aA)_\sigma = aA_\sigma, \\ P_m A_\sigma P_n = \overline{A}_\sigma. \end{cases}$$

(ii) If $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times r}$, $C \in \mathbb{C}^{r\times p}$, *then*

$$\begin{cases} (AB)_\sigma = A_\sigma P_n B_\sigma = A_\sigma \overline{B}_\sigma P_r, \\ (ABC)_\sigma = A_\sigma \overline{B}_\sigma C_\sigma. \end{cases}$$

(iii) If $A \in \mathbb{C}^{m\times n}$, *then* $A_m A_\sigma A_n = A_\sigma$.

(iv) If $A \in \mathbb{C}^{m\times n}$, *then* $\left((A^T)_\sigma\right)^T$.

Lemma 2.3 ([10]) *Given a complex, we have*

(i) $\|A_\sigma\|^2 = 2\|A\|^2$,

(ii) $\|A_\sigma\|_2 = \|A\|_2$.

Lemma 2.4 ([10]) *For two square matrices* $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, *if* $\mathrm{tr}(A) + \mathrm{tr}(B)$ *is real, then*

$$\mathrm{tr}(A) + \mathrm{tr}(B) = \overline{\mathrm{tr}(A) + \mathrm{tr}(B)} = \mathrm{tr}(\overline{A}) + \mathrm{tr}(\overline{B}).$$

3. The matrix equation

In this section, we consider the following equation,

$$AXB + C\overline{X}D = F, \tag{2}$$

where $A, C \in \mathbb{C}^{m\times r}$, $B, D \in \mathbb{C}^{s\times n}$ and $F \in \mathbb{C}^{m\times n}$ are the given constant matrices, and $X \in \mathbb{C}^{r\times s}$ is the matrix to be solved. Define the matrices by applying the hierarchical identification principle. Let

$$F_1 = F - C\overline{X}D, \tag{3}$$

and

$$F_2 = \overline{F} - \overline{AXB}. \tag{4}$$

So

$$F_2 = \overline{C}X\overline{D}.$$

By introducing a relaxation parameter, we can define the following respective relaxed recursive forms,

$$X_1(k) = X_1(k-1) + \omega\mu A^H (F_1 - AX_1(k-1)B)B^H, \tag{5}$$

$$X_2(k) = X_2(k-1) + (1-\omega)\mu \overline{C}^H (F_2 - \overline{C}X_2(k-1)\overline{D})\overline{D}^H, \tag{6}$$

where ω is a relaxation parameter satisfying $0 < \omega < 1$, it controls the relative importance of the two residual matrices. Substituting (3) and (4) into (5) and (6), respectively, we have

$$X_1(k) = X_1(k-1) + \omega\mu A^H\big(F - C\overline{X}D - AX_1(k-1)B\big)B^H, \tag{7}$$

$$X_2(k) = X_2(k-1) + (1-\omega)\mu\overline{C}^H\big(\overline{F} - \overline{AXB} - \overline{C}X_2(k-1)\overline{D}\big)\overline{D}^H. \tag{8}$$

The exact solution on the right-hand sides is unknown. So the variable X is replaced with its estimate $X(k-1)$ in (7) and (8), we can follow that

$$X_1(k) = X_1(k-1) + \omega\mu A^H\big(F - C\overline{X_1(k-1)}D - AX_1(k-1)B\big)B^H, \tag{9}$$

$$X_2(k) = X_2(k-1) + (1-\omega)\mu\overline{C}^H\big(\overline{F} - \overline{AX_2(k-1)B} - \overline{C}X_2(k-1)\overline{D}\big)\overline{D}^H. \tag{10}$$

As the approximate solution $X(k)$ is wanted rather than $X_1(k)$ and $X_2(k)$, so we propose the following balanced strategy to form the k-th approximate solution.

$$X_1(k) = X(k-1) + \omega\mu A^H\big(F - C\overline{X(k-1)}D - AX(k-1)B\big)B^H, \tag{11}$$

$$X_2(k) = X(k-1) + (1-\omega)\mu\overline{C}^H\big(\overline{F} - \overline{AX(k-1)B} - \overline{C}X(k-1)\overline{D}\big)\overline{D}^H, \tag{12}$$

$$X(k) = (1-\omega)X_1(k) + \omega X_2(k). \tag{13}$$

The iterative algorithm can be written as

$$\begin{aligned} X(k) =& X(k-1) + (1-\omega)\omega\mu\overline{C}^H\big(\overline{F} - \overline{AX(k-1)B} - \overline{C}X(k-1)\overline{D}\big)\overline{D}^H + \\ & (1-\omega)\omega\mu A^H\big(F - C\overline{X(k-1)}D - AX(k-1)B\big)B^H. \end{aligned} \tag{14}$$

Theorem 3.1 *If the equation in (2) has a unique solution X_*, then for any initial value $X(0)$, the iterative solution $X(k)$ given in (14) converges to X_*, i.e., $\lim_{k\to\infty} X(k) = X_*$, if*

$$0 < \mu < \frac{2}{\omega(1-\omega)\big\|(\overline{B})_\sigma \otimes (A^H)_\sigma + (D_\sigma P_n) \otimes (C^T)_\sigma P_m\big\|_2^2}.$$

Proof Define error matrices

$$\widetilde{X}(k) = X(k) - X_*, \tag{15}$$

$$\widetilde{X}_i(k) = X_i(k) - X_*. \tag{16}$$

Let

$$Z(k-1) = A\widetilde{X}(k-1)B + C\overline{\widetilde{X}(k-1)}D. \tag{17}$$

Using (11), (12), (15), (16), we have

$$\begin{aligned} \widetilde{X}(k) =& \widetilde{X}(k-1) + (1-\omega)\omega\mu\overline{C}^H\big(\overline{A\widetilde{X}(k-1)B} + \overline{C}\widetilde{X}(k-1)\overline{D}\big)\overline{D}^H - \\ & (1-\omega)\omega\mu A^H\big(A\widetilde{X}(k-1)B - C\overline{\widetilde{X}(k-1)}D\big)B^H. \end{aligned} \tag{18}$$

Using (17), (18), it is easy to get

$$\widetilde{X}(k) = \widetilde{X}(k-1) - (1-\omega)\omega\mu\overline{C}^H\overline{Z(k-1)}\overline{D}^H - (1-\omega)\omega\mu A^H Z(k-1)B^H.$$

Considering the fact

$$\|X\|^2 = \mathrm{tr}(X^H X), \quad \mathrm{tr}(AB) = \mathrm{tr}(BA),$$

we have

$$\|\widetilde{X}(k)\|^2 = \mathrm{tr}\big[\widetilde{X}^H(k)\widetilde{X}(k)\big],$$

$$\begin{aligned}\|\widetilde{X}(k)\|^2 =& \|\widetilde{X}(k-1)\|^2 + (1-\omega)^2\omega^2\mu^2\big\|\big(\overline{C}^H\overline{Z(k-1)}\overline{D}^H + A^H Z(k-1)B^H\big)\big\|^2 - \\ &(1-\omega)\omega\mu\,\mathrm{tr}\big[\widetilde{X}(k-1)\big(\overline{C}^H\overline{Z(k-1)}\overline{D}^H + A^H Z(k-1)B^H\big)^H\big] - \\ &(1-\omega)\omega\mu\,\mathrm{tr}\big[\widetilde{X}^H(k-1)\big(\overline{C}^H\overline{Z(k-1)}\overline{D}^H + A^H Z(k-1)B^H\big)\big].\end{aligned}$$

Considering Lemmas 2.2 and 2.3, one has

$$\begin{aligned}&\big\|\big(\overline{C}^H\overline{Z(k-1)}\overline{D}^H + A^H Z(k-1)B^H\big)\big\|^2 \\ &= \frac{1}{2}\big\|\big(\overline{C}^H\overline{Z(k-1)}\overline{D}^H + A^H Z(k-1)B^H\big)_\sigma\big\|^2 \\ &= \frac{1}{2}\big\|\big((\overline{C}^H)_\sigma P_m(\overline{Z(k-1)})_\sigma P_n(\overline{D}^H)_\sigma + (A^H)_\sigma(\overline{Z(k-1)})_\sigma(B^H)_\sigma\big)_\sigma\big\|^2 \\ &= \frac{1}{2}\big\|\big[\big(P_n(\overline{D}^H)_\sigma\big)^T \otimes (\overline{C}^H)_\sigma P_m + \big((B^H)_\sigma\big)^T \otimes (A^H)_\sigma\big]\mathrm{vec}\big(\big(\overline{Z(k-1)}\big)_\sigma\big)\big\|^2 \\ &= \frac{1}{2}\big\|\big[(D_\sigma P_n)\otimes(C^T)_\sigma P_m + (\overline{B})_\sigma\otimes(A^H)_\sigma\big]\mathrm{vec}\big(\big(\overline{Z(k-1)}\big)_\sigma\big)\big\|^2 \\ &\le \frac{1}{2}\big\|\big[(D_\sigma P_n)\otimes(C^T)_\sigma P_m + (\overline{B})_\sigma\otimes(A^H)_\sigma\big]\big\|_2^2\big\|\mathrm{vec}\big(\big(\overline{Z(k-1)}\big)_\sigma\big)\big\|^2 \\ &= \big\|Z(k-1)\big\|^2\big\|(\overline{B})_\sigma\otimes(A^H)_\sigma + (D_\sigma P_n)\otimes(C^T)_\sigma P_m\big\|_2^2.\end{aligned}$$

Denote

$$\Big\|(\overline{B})_\sigma\otimes(A^H)_\sigma + (D_\sigma P_n)\otimes(C^T)_\sigma P_m\Big\|_2^2 = \pi.$$

By using Lemma 2.4, we have

$$\begin{aligned}&\mathrm{tr}\Big[\overline{D}\widetilde{X}(k-1)\overline{C}\,\overline{Z(k-1)}^H + \overline{C}^H\widetilde{X}^H(k-1)\overline{D}^H\overline{Z(k-1)}\Big] \\ &\quad= \mathrm{tr}\Big[D\overline{\widetilde{X}(k-1)}CZ^H(k-1) + C^H\overline{\widetilde{X}(k-1)}^H D^H Z(k-1)\Big].\end{aligned}$$

Then

$$\begin{aligned}&\mathrm{tr}\Big[\widetilde{X}^H(k-1)\Big(\overline{C}^H\overline{Z(k-1)}\overline{D}^H + A^H Z(k-1)B^H\Big)\Big] + \\ &\quad\mathrm{tr}\Big[\widetilde{X}(k-1)\Big(\overline{C}^H\overline{Z(k-1)}\overline{D}^H + A^H Z(k-1)B^H\Big)^H\Big] \\ &= \mathrm{tr}\Big[\overline{D}\widetilde{X}(k-1)\overline{C}\,\overline{Z(k-1)}^H + A\widetilde{X}(k-1)Z^H(k-1)B\Big] + \\ &\quad\mathrm{tr}\Big[\overline{C}^H\widetilde{X}^H(k-1)\overline{D}^H\overline{Z(k-1)} + A^H\widetilde{X}^H(k-1)B^H Z(k-1)\Big] \\ &= \mathrm{tr}\Big[D\overline{\widetilde{X}(k-1)}CZ^H(k-1) + C^H\overline{\widetilde{X}(k-1)}^H D^H Z(k-1)\Big] + \\ &\quad\mathrm{tr}\Big[A\widetilde{X}(k-1)BZ^H(k-1) + A^H\widetilde{X}^H(k-1)B^H Z(k-1)\Big] \\ &= 2\mathrm{tr}\Big[Z^H(k-1)Z(k-1)\Big] = 2\|Z(k-1)\|^2.\end{aligned}$$

Obviously,

$$\|\widetilde{X}(k)\|^2 \le \|\widetilde{X}(k-1)\|^2 - \|Z(k-1)\|^2\mu\omega(1-\omega)[2 - \pi\omega(1-\omega)\mu]$$

$$\leq \|\widetilde{X}(0)\|^2 - \sum_{i=0}^{k-1} \|Z(i)\|^2 \mu\omega(1-\omega)[2 - \pi\omega(1-\omega)\mu].$$

Thus,

$$\sum_{i=0}^{k-1} \|Z(i)\|^2 \mu\omega(1-\omega)[2 - \pi\omega(1-\omega)\mu] \leq \|\widetilde{X}(0)\|^2 < \infty.$$

If the convergence factor is chosen to satisfy the following inequality

$$0 < \mu < \frac{2}{\omega(1-\omega)\pi},$$

we have

$$\sum_{i=0}^{k-1} \|Z(ki)\|^2 < \infty.$$

Hence, $\|Z(k)\|^2 \to 0$. And,

$$Z(k) \to 0,\ A\widetilde{X}(k-1)B + C\overline{\widetilde{X}(k-1)}D \to 0,$$

$$\widetilde{X}(k) \to 0,\ \text{as } k \to \infty.$$

This completes the proof of Theorem 3.1. □

Corollary 3.2 ([2]) *If the equation in (2) has a unique solution X_*, then for any initial value $X(0)$, the iterative solution $X(k)$ given in (14) converges to X_*, i.e., $\lim_{k\to\infty} X(k) = X_*$, if*

$$0 < \mu < \frac{1}{\omega(1-\omega)\big(\|B\|_2^2\|A\|_2^2 + \|D\|_2^2\|C\|_2^2\big)}$$

Proof For two arbitrary matrices, $\|A \otimes B\| = \|A\|\|B\|$. Then

$$\begin{aligned}
&\Big\|\big(\overline{B}\big)_\sigma \otimes \big(A^H\big)_\sigma + \big(D_\sigma P_n\big) \otimes \big(C^T\big)_\sigma P_m\Big\|_2^2 \\
&\leq \Big(\Big\|\big(\overline{B}\big)_\sigma \otimes \big(A^H\big)_\sigma\Big\|_2 + \Big\|\big(D_\sigma P_n\big) \otimes \big(C^T\big)_\sigma P_m\Big\|_2\Big)^2 \\
&= \Big(\Big\|\big(\overline{B}\big)_\sigma\Big\|_2\Big\|\big(A^H\big)_\sigma\Big\|_2 + \Big\|\big(D_\sigma P_n\big)\Big\|_2\Big\|C_\sigma^T P_m\Big\|_2\Big)^2 \\
&= \Big(\big\|B\big\|_2\big\|A\big\|_2 + \big\|D\big\|_2\big\|C\big\|_2\Big)^2 \\
&\leq 2\Big(\big\|B\big\|_2^2\big\|A\big\|_2^2 + \big\|D\big\|_2^2\big\|C\big\|_2^2\Big)
\end{aligned}$$

$$\frac{1}{\omega(1-\omega)\Big(\big\|B\big\|_2^2\big\|A\big\|_2^2 + \big\|D\big\|_2^2\big\|C\big\|_2^2\Big)} < \frac{2}{\omega(1-\omega)\Big\|\big(\overline{B}\big)_\sigma \otimes \big(A^H\big)_\sigma + \big(D_\sigma P_n\big) \otimes \big(C^T\big)_\sigma P_m\Big\|_2^2}.$$

This completes the proof of Corollary 3.2. □

4. A general case

In this section, we consider a more general Sylvester matrix equation. Such a class of equations are in the form of

$$\sum_{j=1}^{m} A_j X B_j + \sum_{i=1}^{n} C_i \overline{X} D_i = F,$$

which can be written as

$$\sum_{j=1}^{p} A_j X B_j + \sum_{i=1}^{p} C_i \overline{X} D_i = F, \quad p = \max(m, n), \tag{19}$$

where $A_j, C_i \in \mathbb{C}^{m\times r}, B_j, D_i \in \mathbb{C}^{s\times n}, j \in I[1,p], i \in I[1,p],\ F \in \mathbb{C}^{m\times n}$ are the given constant matrices, and $X \in \mathbb{C}^{r\times s}$ is the matrix to be solved. The equation (19) can be written as

$$F_{i+p} = \overline{C}_i X \overline{D}_i, \quad i \in I[1,p],$$

$$F_j = A_j X B_j, \quad j \in I[1,p].$$

Let

$$F_j = F - \sum_{l=1}^{p} A_l X B_l - \sum_{l=1}^{p} C_l \overline{X} D_l + A_j X B_j, \quad j \in I[1,p], \tag{20}$$

and

$$\overline{F}_{i+p} = F - \sum_{l=1}^{p} A_l X B_l - \sum_{l=1}^{p} C_l \overline{X} D_l + C_i \overline{X} D_i, \quad i \in I[1,p]. \tag{21}$$

By introducing a relaxation parameter, we can define the following respective relaxed recursive forms,

$$X_j(k) = X_j(k-1) + (1-\omega)\mu A_j^H \big(F_j - A_j X_j(k-1) B_j\big) B_j^H, \quad j \in I[1,p], \tag{22}$$

$$X_{i+p}(k) = X_{i+p}(k-1) + \omega\mu \overline{C}_i^H \big(F_{i+p} - \overline{C}_i X_{i+p}(k-1) \overline{D}_i\big) \overline{D}_i^H, \quad i \in I[1,p], \tag{23}$$

where ω is a relaxation parameter satisfying $0 < \omega < 1$, it controls the relative importance of the two residual matrices. Substitute (20) and (21) into (22) and (23). For $i \in I[1,p]$, $j \in I[1,p]$ we have

$$X_j(k) = X_j(k-1) + (1-\omega)\mu A_j^H \Big(F - \sum_{l=1}^{p} A_l X B_l - \sum_{l=1}^{p} C_l \overline{X} D_l + A_j X B_j - A_j X_j(k-1) B_j\Big) B_j^H,$$

$$X_{i+p}(k) = X_{i+p}(k-1) + \omega\mu \overline{C}_i^H \Big(\overline{F} - \sum_{l=1}^{p} \overline{A_l X B_l} - \sum_{l=1}^{p} \overline{C}_l X \overline{D}_l + \overline{C}_i X \overline{D}_i - \overline{C}_i X_{i+p}(k-1) \overline{D}_i\Big) \overline{D}_i^H.$$

The exact solution X on the right-hand sides is unknown, so the variable X is replaced with $X(k-1)$ in above two matrices. Hence, for $i \in I[1,p]$, $j \in I[1,p]$, we can get

$$X_j(k) = X_j(k-1) + (1-\omega)\mu A_j^H \Big(F - \sum_{l=1}^{p} A_l X_j(k-1) B_l - \sum_{l=1}^{p} C_l \overline{X_j(k-1)} D_l\Big) B_j^H,$$

$$X_{i+p}(k) = X_{i+p}(k-1) + \omega\mu \overline{C}_i^H \Big(\overline{F} - \sum_{l=1}^{p} \overline{A_l X_{i+p}(k-1) B_l} - \sum_{l=1}^{p} \overline{C}_l X_{i+p}(k-1) \overline{D}_l\Big) \overline{D}_i^H.$$

As the approximate solution $X(k)$ is wanted rather than $X_i(k)$ and $X_j(k)$, so we propose the following balanced strategy to form the k-th approximate solution. For $i \in I[1,p]$, $j \in I[1,p]$,

$$X_{i+p}(k) = X(k-1) + \omega\mu \overline{C}_i^H \Big(\overline{F} - \sum_{l=1}^{p} \overline{A_l X(k-1) B_l} - \sum_{l=1}^{p} \overline{C}_l X(k-1) \overline{D}_l\Big) \overline{D}_i^H. \tag{24}$$

$$X_j(k) = X(k-1) + (1-\omega)\mu A_j^H\Big(F - \sum_{l=1}^{p} A_l X(k-1) B_l - \sum_{l=1}^{p} C_l \overline{X(k-1)} D_l\Big) B_j^H, \tag{25}$$

$$X(k) = \frac{1}{p}\Big[(1-\omega)\sum_{i=1}^{p} X_{i+p}(k) + \omega \sum_{j=1}^{p} X_j(k)\Big]. \tag{26}$$

For $i \in I[1,p]$, $j \in I[1,p]$, the iterative algorithm can be written as

$$\begin{aligned} X(k) =& X(k-1) + \frac{\omega\mu(1-\omega)}{p}\sum_{j=1}^{p} A_j^H\Big(F - \sum_{l=1}^{p} A_l X(k-1) B_l - \sum_{l=1}^{p} C_l \overline{X(k-1)} D_l\Big) B_j^H + \\ & \frac{\omega\mu(1-\omega)}{p}\sum_{i=1}^{p} \overline{C}_i^H\Big(\overline{F} - \sum_{l=1}^{p} \overline{A_l X(k-1) B_l} - \sum_{l=1}^{p} \overline{C}_l X(k-1) \overline{D}_l\Big) \overline{D}_i^H. \end{aligned}$$

Theorem 4.1 *If the equation in (19) has a unique solution X_*, then for any initial value $X(0)$, the iterative solution $X(k)$ given in (26) converges to X_*, i.e., $\lim_{k\to\infty} X(k) = X_*$, if*

$$0 < \mu < \frac{2p}{\| \sum\limits_{i=1}^{p} \big((D_i)_\sigma P_n\big) \otimes \big((C_i^T)_\sigma P_m\big) + \sum\limits_{j=1}^{p} \big(\overline{B}_j\big)_\sigma \otimes \big(A_j^H\big)_\sigma \|_2^2 (1-\omega)\omega}.$$

Proof For $i \in I[1,p]$, $j \in I[1,p]$, define error matrices as follows

$$\widetilde{X}(k) = X(k) - X_*, \tag{27}$$

$$\widetilde{X}_i(k) = X_i(k) - X_*. \tag{28}$$

Using (24), (25), (27), (28), we get

$$\begin{aligned} \widetilde{X}_j(k) =& \widetilde{X}(k-1) + (1-\omega)\mu A_j^H\Big(F - \sum_{l=1}^{p} A_l X(k-1) B_l - \sum_{l=1}^{p} C_l \overline{X(k-1)} D_l\Big) B_j^H \\ =& \widetilde{X}(k-1) - (1-\omega)\mu A_j^H\Big(\sum_{l=1}^{p} A_l \widetilde{X}(k-1) B_l + \sum_{l=1}^{p} C_l \overline{\widetilde{X}(k-1)} D_l\Big) B_j^H, \\ \widetilde{X}_{i+p}(k) =& \widetilde{X}(k-1) + \omega\mu \overline{C}_i^H\Big(\overline{F} - \sum_{l=1}^{p} \overline{A_l X(k-1) B_l} - \sum_{l=1}^{p} \overline{C}_l X(k-1) \overline{D}_l\Big) \overline{D}_i^H \\ =& \widetilde{X}(k-1) - \omega\mu \overline{C}_i^H\Big(\sum_{l=1}^{p} \overline{A_l \widetilde{X}(k-1) B_l} + \sum_{l=1}^{p} \overline{C}_l \widetilde{X}(k-1) \overline{D}_l\Big) \overline{D}_i^H, \end{aligned}$$

$$\widetilde{X}(k) = \frac{1}{p}\Big[(1-\omega)\sum_{i=1}^{p} \widetilde{X}_{i+p}(k) + \omega \sum_{j=1}^{p} \widetilde{X}_j(k)\Big].$$

Denote

$$\begin{aligned} Z(k-1) =& F - \sum_{l=1}^{p} A_l X(k-1) B_l - \sum_{l=1}^{p} C_l \overline{X(k-1)} D_l \\ =& \Big(\sum_{l=1}^{p} A_l \widetilde{X}(k-1) B_l + \sum_{l=1}^{p} C_l \overline{\widetilde{X}(k-1)} D_l\Big). \end{aligned}$$

Then

$$\begin{aligned}X(k) =&\widetilde{X}(k-1)-\frac{\omega\mu(1-\omega)}{p}\sum_{j=1}^{p}A_j^H\Big(\sum_{l=1}^{p}A_l\widetilde{X}(k-1)B_l+\sum_{l=1}^{p}C_l\overline{\widetilde{X}(k-1)}D_l\Big)B_j^H-\\ &\frac{\omega\mu(1-\omega)}{p}\sum_{i=1}^{p}\overline{C}_i^H\Big(\sum_{l=1}^{p}\overline{A_l\widetilde{X}(k-1)B_l}+\sum_{l=1}^{p}\overline{C}_l\widetilde{X}(k-1)\overline{D}_l\Big)\overline{D}_i^H\\ =&\widetilde{X}(k-1)-\frac{\omega\mu(1-\omega)}{p}\Big(\sum_{i=1}^{p}\overline{C}_i^H\overline{Z(k-1)}\overline{D}_i^H+\sum_{j=1}^{p}A_j^HZ(k-1)B_j^H\Big).\end{aligned}$$

Consider the fact $\|X\|^2=\mathrm{tr}\big(X^HX\big)$, $\mathrm{tr}(AB)=\mathrm{tr}(BA)$. We have

$$\begin{aligned}\big\|\widetilde{X}(k)\big\|^2 =&\mathrm{tr}\big[\widetilde{X}^H(k)\widetilde{X}(k)\big]\\ =&\big\|\widetilde{X}(k-1)\big\|^2+\big[\frac{\omega\mu(1-\omega)}{p}\big]^2\Big\|\sum_{i=1}^{p}\overline{C}_i^H\overline{Z(k-1)}\overline{D}_i^H+\sum_{j=1}^{p}A_j^HZ(k-1)B_j^H\Big\|^2-\\ &\frac{\omega\mu(1-\omega)}{p}\mathrm{tr}\Big[\widetilde{X}^H(k-1)\Big(\sum_{i=1}^{p}\overline{C}_i^H\overline{Z(k-1)}\overline{D}_i^H+\sum_{j=1}^{p}A_j^HZ(k-1)B_j^H\Big)\Big]-\\ &\frac{\omega\mu(1-\omega)}{p}\mathrm{tr}\Big[\Big(\sum_{i=1}^{p}\overline{C}_i^H\overline{Z(k-1)}\overline{D}_i^H+\sum_{j=1}^{p}A_j^HZ(k-1)B_j^H\Big)^H\widetilde{X}(k-1)\Big].\end{aligned}$$

Then, consider Lemma 2.4,

$$\begin{aligned}&\mathrm{tr}\Big[\widetilde{X}^H(k-1)\Big(\sum_{i=1}^{p}\overline{C}_i^H\overline{Z(k-1)}\overline{D}_i^H+\sum_{j=1}^{p}A_j^HZ(k-1)B_j^H\Big)\Big]+\\ &\quad\mathrm{tr}\Big[\Big(\sum_{i=1}^{p}\overline{C}_i^H\overline{Z(k-1)}\overline{D}_i^H+\sum_{j=1}^{p}A_j^HZ(k-1)B_j^H\Big)^H\widetilde{X}(k-1)\Big]\\ &=\mathrm{tr}\Big[\overline{Z(k-1)}\sum_{i=1}^{p}\overline{C}_i^H\widetilde{X}^H(k-1)\overline{D}_i^H+Z(k-1)\sum_{j=1}^{p}A_j^H\widetilde{X}^H(k-1)B_j^H\Big]+\\ &\quad\mathrm{tr}\Big[\sum_{i=1}^{p}\overline{D}_i\widetilde{X}(k-1)\overline{C}_i\overline{Z(k-1)}^H+\sum_{j=1}^{p}A_j\widetilde{X}(k-1)B_jZ^H(k-1)\Big]\\ &=\mathrm{tr}\Big[Z(k-1)\sum_{j=1}^{p}A_j^H\widetilde{X}^H(k-1)B_j^H+\sum_{j=1}^{p}A_j\widetilde{X}(k-1)B_jZ^H(k-1)\Big]+\\ &\quad\mathrm{tr}\Big[\sum_{i=1}^{p}D_i\overline{\widetilde{X}(k-1)}C_iZ^H(k-1)+Z(k-1)\sum_{i=1}^{p}C_i^H\overline{\widetilde{X}(k-1)}^HD_i^H\Big]\\ &=\mathrm{tr}\big[Z^H(k-1)Z(k-1)+Z(k-1)Z^H(k-1)\big]\\ &=2\big\|Z(k-1)\big\|^2.\end{aligned}$$

And considering the formula $\mathrm{vec}\big(MXN\big)=\big(N^T\otimes M\big)\mathrm{vec}\big(X\big)$, we get

$$\Big\|\Big(\sum_{i=1}^{p}\overline{C}_i^H\overline{Z(k-1)}\overline{D}_i^H+\sum_{j=1}^{p}A_j^HZ(k-1)B_j^H\Big)\Big\|^2$$

$$
\begin{aligned}
&= \frac{1}{2}\Big\|\Big(\sum_{i=1}^{p}\overline{C}_i^H\overline{Z(k-1)}\overline{D}_i^H+\sum_{j=1}^{p}A_j^H Z(k-1)B_j^H\Big)_\sigma\Big\|^2 \\
&= \frac{1}{2}\Big\|\sum_{i=1}^{p}\big(\overline{C}_i^H\big)_\sigma P_m\big(\overline{Z(k-1)}\big)_\sigma P_n\big(\overline{D}_i^H\big)_\sigma+\sum_{j=1}^{p}\big(A_j^H\big)_\sigma\big(\overline{Z(k-1)}\big)_\sigma\big(B_j^H\big)_\sigma\Big\|^2 \\
&= \frac{1}{2}\Big\|\Big[\sum_{i=1}^{p}\big(P_n\big(\overline{D}_i^H\big)_\sigma\big)^T\otimes\big(\big(\overline{C}_i^H\big)_\sigma P_m\big)+\sum_{j=1}^{p}\big(\big(\overline{B}_j^H\big)_\sigma\big)^T\otimes\big(\big(\overline{A}_j^H\big)_\sigma\big)\Big]\mathrm{vec}\big(\big(\overline{Z(k-1)}\big)_\sigma\big)\Big\|^2 \\
&= \frac{1}{2}\Big\|\Big[\sum_{i=1}^{p}\big((D_i)_\sigma P_n\big)\otimes\big((C_i^T)_\sigma P_m\big)+\sum_{j=1}^{p}\big(\overline{B}_j\big)_\sigma\otimes\big(A_j^H\big)_\sigma\Big]\mathrm{vec}\big(\big(\overline{Z(k-1)}\big)_\sigma\big)\Big\|^2 \\
&\le \frac{1}{2}\Big\|\sum_{i=1}^{p}\big((D_i)_\sigma P_n\big)\otimes\big((C_i^T)_\sigma P_m\big)+\sum_{j=1}^{p}\big(\overline{B}_j\big)_\sigma\otimes\big(A_j^H\big)_\sigma\Big\|_2^2\Big\|\mathrm{vec}\big(\big(\overline{Z(k-1)}\big)_\sigma\big)\Big\|^2 \\
&= \Big\|\sum_{i=1}^{p}\big((\overline{D}_i)_\sigma P_n\big)\otimes\big((C_i^T)_\sigma P_m\big)+\sum_{j=1}^{p}\big(\overline{B}_j\big)_\sigma\otimes\big(A_j^H\big)_\sigma\Big\|_2^2\big\|Z(k-1)\big\|^2.
\end{aligned}
$$

Denote

$$\pi=\Big\|\sum_{i=1}^{p}\big((D_i)_\sigma P_n\big)\otimes\big((C_i^T)_\sigma P_m\big)+\sum_{j=1}^{p}\big(\overline{B}_j\big)_\sigma\otimes\big(A_j^H\big)_\sigma\Big\|_2^2.$$

One has

$$
\begin{aligned}
\big\|\widetilde{X}(k)\big\|^2 &\le\big\|\widetilde{X}(k-1)\big\|^2-\frac{2(1-\omega)\mu\omega}{p}\big\|Z(k-1)\big\|^2+\big[\frac{(1-\omega)\mu\omega}{p}\big]^2\pi\big\|Z(k-1)\big\|^2 \\
&\le\big\|\widetilde{X}(0)\big\|^2-\big\{\frac{2(1-\omega)\mu\omega}{p}-\big[\frac{(1-\omega)\mu\omega}{p}\big]^2\pi\big\}\sum_{i=0}^{k-1}\big\|Z(i)\big\|^2.
\end{aligned}
$$

So

$$0<\big\{\frac{2(1-\omega)\mu\omega}{p}-\big[\frac{(1-\omega)\mu\omega}{p}\big]^2\pi\big\}\sum_{i=0}^{k-1}\big\|Z(i)\big\|^2\le\big\|\widetilde{X}(0)\big\|^2,$$

or,

$$0<\big\{\frac{2(1-\omega)\mu\omega}{p}-\big[\frac{(1-\omega)\mu\omega}{p}\big]^2\pi\big\}\sum_{i=0}^{\infty}\big\|Z(i)\big\|^2\le\big\|\widetilde{X}(0)\big\|^2.$$

If the parameter μ is chosen to satisfy the following

$$0<\mu<\frac{2p}{\pi(1-\omega)\omega}.$$

One has

$$\sum_{i=0}^{\infty}\big\|Z(i)\big\|^2<\infty.$$

Then

$$\big\|Z(i)\big\|^2\to 0,\ Z(i)\to 0,$$

$$\sum_{l=1}^{p}A_l\widetilde{X}(k)B_l+\sum_{l=1}^{p}C_l\overline{\widetilde{X}(k)}D_l\to 0.$$

Thus

$$\widetilde{X}(k)\to 0\quad\text{as}\quad k\to\infty.$$

This completes the proof of Theorem 4.1. □

Corollary 4.2 *If the equation in (19) has a unique solution X_*, then for any initial value $X(0)$, the iterative solution $X(k)$ given in (26) converges to X_*, i.e., $\lim_{k\to\infty} X(k) = X_*$, if*

$$0 < \mu < \frac{1}{\omega(1-\omega)\big(\sum\limits_{i=1}^{p} \|C_i\|_2^2\|D_i\|_2^2 + \sum\limits_{j=1}^{p} \|B_j\|_2^2\|A_j\|_2^2\big)}.$$

Proof For two arbitrary matrices,

$$\|A \otimes B\| = \|A\|\|B\|.$$

Then,

$$\begin{aligned}
&\Big\| \sum_{i=1}^{p} \big((D_i)_\sigma P_n\big) \otimes \big((C_i^T)_\sigma P_m\big) + \sum_{j=1}^{p} \big(\overline{B}_j\big)_\sigma \otimes \big(A_j^H\big)_\sigma \Big\|_2^2 \\
&\le \Big(\Big\| \sum_{i=1}^{p} \big((D_i)_\sigma P_n\big) \otimes \big((C_i^T)_\sigma P_m\big)\Big\|_2 + \Big\| \sum_{j=1}^{p} \big(\overline{B}_j\big)_\sigma \otimes \big(A_j^H\big)_\sigma \Big\|_2\Big)^2 \\
&= \Big(\Big\| \sum_{i=1}^{p} (D_i)_\sigma P_n\Big\|_2 \Big\| \sum_{i=1}^{p} (C_i^T)_\sigma P_m\Big\|_2 + \Big\| \sum_{j=1}^{p} \big(\overline{B}_j\big)_\sigma\Big\|_2 \Big\| \sum_{j=1}^{p} \big(A_j^H\big)_\sigma \Big\|_2\Big)^2 \\
&= \Big(\sum_{i=1}^{p} \|C_i\|_2\|D_i\|_2 + \sum_{j=1}^{p} \|B_j\|_2\|A_j\|_2\Big)^2 \\
&\le 2\Big(\sum_{i=1}^{p} \|C_i\|_2\|D_i\|_2\Big)^2 + 2\Big(\sum_{j=1}^{p} \|B_j\|_2\|A_j\|_2\Big)^2 \\
&\le 2p\Big(\sum_{i=1}^{p} \|C_i\|_2^2\|D_i\|_2^2 + \sum_{j=1}^{p} \|B_j\|_2^2\|A_j\|_2^2\Big).
\end{aligned}$$

So

$$\begin{aligned}
&\frac{1}{\omega(1-\omega)\big(\sum\limits_{i=1}^{p} \|C_i\|_2^2\|D_i\|_2^2 + \sum\limits_{j=1}^{p} \|B_j\|_2^2\|A_j\|_2^2\big)} \\
&\quad \le \frac{2p}{\omega(1-\omega)\big\| \sum\limits_{i=1}^{p} \big((D_i)_\sigma P_n\big) \otimes \big((C_i^T)_\sigma P_m\big) + \sum\limits_{j=1}^{p} \big(\overline{B}_j\big)_\sigma \otimes \big(A_j^H\big)_\sigma \big\|_2^2}.
\end{aligned}$$

This completes the proof of Corollary 4.2. □

5. Numerical examples

This section gives two examples to testify the iterative algorithm.

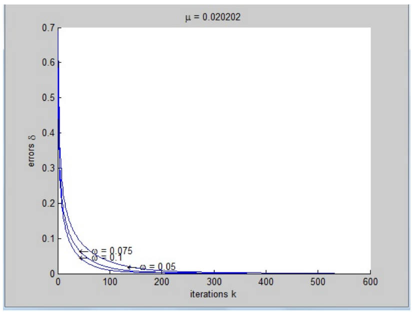

Figure 1 The convergence performance of algorithm (14) for different ω when $\mu = 0.02$

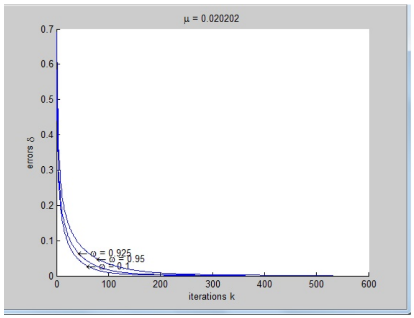

Figure 2 The convergence performance of algorithm (14) for different ω when $\mu = 0.02$

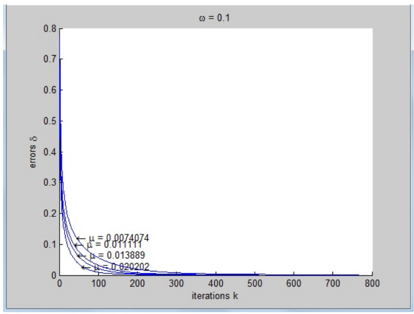

Figure 3 The convergence performance of algorithm (14) for different μ when $\omega = 0.1$

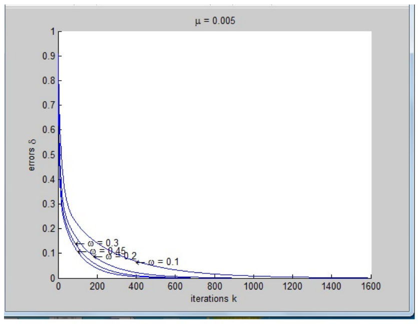

Figure 4 The convergence performance of algorithm (24)–(26) for different ω when $\mu = 0.005$

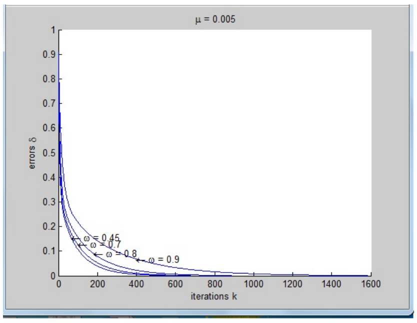

Figure 5 The convergence performance of algorithm (24)–(26) for different ω when $\mu = 0.005$

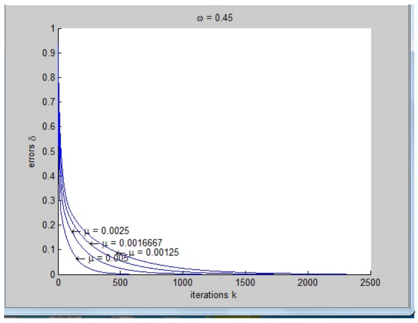

Figure 6 The convergence performance of algorithm (24)–(26) for different μ when $\omega = 0.45$

Example 5.1 Suppose the equation $AXB + C\overline{X}D = F$ with the matrices

$$A = \begin{bmatrix} 1+2i & 2-i \\ 1-i & 2+3i \end{bmatrix}, \quad B = \begin{bmatrix} 2-4i & i \\ -1+3i & 2 \end{bmatrix}, \quad C = \begin{bmatrix} -1-i & -3i \\ 0 & 1+2i \end{bmatrix}$$

$$D = \begin{bmatrix} -2 & 1-i \\ 1+i & -1-i \end{bmatrix}, \quad F = \begin{bmatrix} 21+11i & -9+7i \\ 52-22i & -18+i \end{bmatrix}$$

and exact solution $X = \left[\begin{smallmatrix} 1+2i & -i \\ 2+i & -1+i \end{smallmatrix}\right]$. We apply algorithm in (11)–(13) with the initial matrix

$X(0) = 10^{-6} I_{2\times 2}$, and define the relative error $\sigma = \frac{\|X(k)-X\|}{\|X\|}$, we have the following figures and the tables.

According to Theorem 3.1, we have $0 < \mu < \frac{1}{\omega(1-\omega)(\|B\|_2^2\|A\|_2^2+\|D\|_2^2\|C\|_2^2)}$. $\|A\|_2^2 = 18.0902$, $\|B\|_2^2 = 31.9309$, $\|C\|_2^2 = 15.3485$, $\|D\|_2^2 = 9.5826$, so we have $0 < \mu < \frac{1}{720.8\omega(1-\omega)}$. Through a large number of experiments, we know when $\omega = 0.1$ and $\mu = 0.02$ are selected, it can bring about the faster convergent rate and smaller relative error, which are shown in Figures 1–3. The matrices used in this example are taken from [10]. For $\omega = 0.1$, $\mu = 0.02$, we give the following Tables 1 and 2, where δ_1 denotes the relative error of the algorithm in [10].

k	x_{11}	x_{12}	x_{21}	x_{22}
5	$1.0950 - 1.7129i$	$-0.0370 + 1.1602i$	$1.6925 - 1.0521i$	$-1.2797 - 0.7333i$
15	$1.0679 - 1.8095i$	$0.0476 + 1.1447i$	$1.8243 - 0.8629i$	$-1.2935 - 0.9498i$
25	$1.0822 - 1.8852i$	$0.0838 + 1.0974i$	$1.8865 - 0.9326i$	$-1.1638 - 0.9554i$
35	$1.0727 - 1.9216i$	$0.0812 + 1.0680i$	$1.9264 - 0.9605i$	$-1.0983 - 0.9617i$
45	$1.0582 - 1.9428i$	$0.0680 + 1.0505i$	$1.9514 - 0.9745i$	$-1.0623 - 0.9701i$
55	$1.0446 - 1.9563i$	$0.0537 + 1.0393i$	$1.9677 - 0.9826i$	$-1.0408 + 0.9777i$
65	$1.0335 - 1.9656i$	$0.0413 + 1.0316i$	$1.9787 - 1.9787i$	$-1.0270 - 0.9837i$
75	$1.0248 - 1.9724i$	$0.0313 + 1.0259i$	$1.9862 - 0.9913i$	$-1.0178 + 0.9883i$
85	$1.0182 - 1.9774i$	$0.0237 + 1.0217i$	$1.9914 - 0.9937i$	$-1.0114 + 0.9917i$
95	$1.0134 - 1.9812i$	$0.0179 + 1.0183i$	$1.9951 - 0.9955i$	$-1.0070 + 0.9942i$
105	$1.0098 - 1.9842i$	$0.0136 + 1.0157i$	$1.9976 - 0.9968i$	$-1.0040 + 0.9960i$

Table 1 The numerical solution of algorithm (14)

k	δ	δ_1
5	0.2673	0.2645
15	0.1306	0.1388
25	0.0841	0.0907
35	0.0592	0.0682
45	0.0432	0.0512
55	0.0320	0.0391
65	0.0239	0.0301
75	0.0180	0.0233
85	0.0138	0.0182
95	0.0107	0.0143
105	0.0084	0.0114

Table 2 The comparison of relative error of algorithm (14) and algorithm of Wu in [10]

The unique solution of this equation is $X = \begin{bmatrix} 1+2i & -i \\ 2+i & -1+i \end{bmatrix}$. Through comparison, we find though we do not determine the optimal value of ω and μ, the algorithm (14) is better than the algorithm of Wu in [10].

Example 5.2 Suppose the equation $A_1XB_1 + C_1\overline{X}D_1 + C_2\overline{X}D_2 = F$ with the exact solution $X = \begin{bmatrix} -1+i & 1 \\ 1+2i & 1-i \end{bmatrix}$, and the matrices

$$A_1 = \begin{bmatrix} -1-i & 2-i \\ 1-i & -1+2i \end{bmatrix},\ B_1 = \begin{bmatrix} 1-2i & i \\ -1+3i & -1+3i \end{bmatrix},\ C_1 = \begin{bmatrix} 2+3i & -1-i \\ 1 & 2+2i \end{bmatrix},$$

$$D_1 = \begin{bmatrix} 2 & 3-i \\ 1+3i & 1+2i \end{bmatrix},\ C_2 = \begin{bmatrix} -1 & 3+i \\ -2+i & 2i \end{bmatrix},\ D_2 = \begin{bmatrix} 2 & -2+3i \\ -1+10i & 4i \end{bmatrix},$$

$$F = \begin{bmatrix} -22-8i & -21+31i \\ -22-20i & -40-10i \end{bmatrix}.$$

According to Theorem 4.1, we have

$$0 < \mu < \frac{1}{\omega(1-\omega)\big(\sum\limits_{i=1}^{p} \|C_i\|_2^2\|D_i\|_2^2 + \sum\limits_{j=1}^{p} \|B_j\|_2^2\|A_j\|_2^2\big)}.$$

For Example 5.2, we have $0 < \mu < \frac{1}{2724.4\omega(1-\omega)}$. Through a large number of experiments, we know when $\omega = 0.45$, $\mu = 0.005$ are selected, it can bring about the faster convergent rate and smaller relative error, which are Shown in Figures 4–6. The matrices used in this example are taken from [10]. For $\omega = 0.45$, $\mu = 0.005$ we give the following Tables 3 and 4, where δ_1 denotes the relative error of the algorithm in [10]. The unique solution of this equation is $X = \begin{bmatrix} -1+i & 1 \\ 1+2i & 1-i \end{bmatrix}$. Through comparison, we find though we do not determine the optimal value of ω and μ, the algorithm (26) is better than the algorithm of Wu in [10].

k	x_{11}	x_{12}	x_{21}	x_{22}
10	$-0.6608-0.2506i$	$0.4724+0.0220i$	$1.0844-1.2476i$	$0.7246+1.2776i$
20	$-0.7066-0.5834i$	$0.4528+0.0321i$	$1.1682-1.5824i$	$0.8931+1.3493i$
30	$-0.7626-0.7237i$	$0.4768+0.0554i$	$1.1564-1.7197i$	$0.9487+1.3360i$
40	$-0.8159-0.7880i$	$0.5142+0.0652i$	$1.1267-1.7915i$	$0.9711+1.3045i$
50	$-0.8600-0.8230i$	$0.5541+0.0671i$	$1.0969-1.8367i$	$0.9821+1.2712i$
60	$-0.8948-0.8462i$	$0.5927+0.0655i$	$1.0717-1.8686i$	$0.9884+1.2405i$
70	$-0.9217-0.8639i$	$0.6287+0.0624i$	$1.0517-1.8927i$	$0.9926+1.2135i$
80	$-0.9424-0.8786i$	$0.6620+0.0587i$	$1.0363-1.9115i$	$0.9955+1.1898i$
90	$-0.9581-0.8912i$	$0.6925+0.0548i$	$1.0246-1.9264i$	$0.9976+1.1691i$
100	$-0.9701-0.9023i$	$0.7205+0.0509i$	$1.0159-1.9384i$	$0.9991+1.1510i$
110	$-0.9791-0.9121i$	$0.7460+0.0471i$	$1.0094-1.9481i$	$1.0001+1.1351i$
120	$-0.9859-0.9208i$	$0.7692+0.0434i$	$1.0047-1.9559i$	$1.0008+1.1210i$
130	$-0.9909-0.9286i$	$0.7904+0.0399i$	$1.0012-1.9624i$	$1.0013+1.1087i$
140	$-0.9946-0.9356i$	$0.8097+0.0366i$	$0.9988-1.9677i$	$1.0016+1.0977i$
150	$-0.9974-0.9418i$	$0.8272+0.0335i$	$0.9971-1.9721i$	$1.0018+1.0879i$

Table 3 The numerical solution of algorithm (14) and algorithm of Wu in [10]

k	δ	δ_1
10	0.4014	0.446844
20	0.2825	0.326441
30	0.2364	0.273501
40	0.2069	0.239945
50	0.1835	0.213944
60	0.1638	0.192289
70	0.1468	0.173764
80	0.1319	0.157687
90	0.1188	0.143585
100	0.1072	0.131105
110	0.0968	0.119974
120	0.0875	0.109981
130	0.0792	0.100961
140	0.0717	0.092782
150	0.0649	0.085338

Table 4 The comparison of relative error of algorithm (14) and algorithm of Wu in [10]

6. Conclusions

This paper gives an iterative algorithm for solving a class of Sylvester-conjugate matrix equations by using the hierarchical identification principle and introducing a relaxation parameter. The analysis of theorems and numerical examples illustrate the performance of algorithm. Obviously, the relaxation parameter influences the convergence of the iterative algorithm. It is a difficult task to give an insightful conclusion about the determination of the optimal value, which needs further investigation.

References

[1] Feng DING, P. X. LIU, Jie DING. *Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle.* Appl. Math. Comput., 2008, **197**(1): 41–50.
[2] Feng DING, Tongwen CHEN. *Gradient based iterative algorithms for solving a class of matrix equations.* IEEE Trans. Automat. Control, 2005, **50**(8): 1216–1221.
[3] Feng DING, Tongwen CHEN. *Iterative least-squares solutions of coupled Sylvester matrix equations.* Systems Control Lett., 2005, **54**(2): 95–107.
[4] Feng DING, Tongwen CHEN. *On iterative solutions of general coupled matrix equations.* SIAM J. Control Optim., 2006, **44**(6): 2269–2284.
[5] Qiang NIU, Xiang WANG, Linzhang LIU. *A relaxed gradient based algorithm for solving Sylvester equations.* Asian J. Control, 2011, **13**(3): 461–464.
[6] Caiqin SONG, Guoliang CHEN. *An efficient algorithm for solving extended Sylvester-conjugate transpose matrix equations.* Arab J. Math. Sci., 2011, **17**(2): 115–134.
[7] Caiqin SONG, Guoliang CHEN, Linlin ZHAO. *Iterative solutions to coupled Sylvester-transpose matrix equations.* Appl. Math. Model., 2011, **35**(10): 4675–4683.

[8] Aiguo WU, Lingling LV, Guangren DUAN. *Iterative algorithms for solving a class of complex conjugate and transpose matrix equations*. Appl. Math. Comput., 2011, **217**(21): 8343–8353.

[9] Aiguo WU, Gang FENG, Guangren DUAN, et al. *Iterative solutions to coupled Sylvester-conjugate matrix equations*. Comput. Math. Appl., 2010, **60**(1): 54–66.

[10] Aiguo WU, Xianlin ZENG, Guangren DUAN, et al. *Iterative solutions to the extended Sylvester-conjugate matrix equations*. Appl. Math. Comput., 2010, **217**(1): 130–142.

[11] Aiguo WU, Gang FENG, Guangren DUAN, et al. *Iterative solutions to coupled Sylvester-conjugate matrix equations*. Comput. Math. Appl., 2010, **60**(1): 54–66.

[12] Li XIE, Yanjun LIU, Huizhong YANG. *Gradient based and least squares based iterative algorithms for matrix equations $AXB + CXTD = F$*. Appl. Math. Comput., 2010, **217**(5): 2191–2199.

[13] Li XIE, Jie DING, Feng DING. *Gradient based iterative solutions for general linear matrix equations*. Comput. Math. Appl., 2009, **58**(7): 1441–1448.

[14] Xiang WANG, Lin DAI, Dan LIAO. *A modified gradient based algorithm for solving Sylvester equations*. Appl. Math. Comput., 2012, **218**(9): 5620–5628.