

Iterative solutions to matrix equations of the form $A_iXB_i = F_i$ [☆]

Jie Ding^{*}, Yanjun Liu, Feng Ding

School of Communication and Control Engineering, Jiangnan University, Wuxi 214122, PR China

ARTICLE INFO

Article history:

Received 23 October 2009

Received in revised form 18 March 2010

Accepted 18 March 2010

Keywords:

Matrix equations

Identification

Parameter estimation

Iterative algorithms

ABSTRACT

This paper is concerned with the numerical solutions to the linear matrix equations $A_1XB_1 = F_1$ and $A_2XB_2 = F_2$; two iterative algorithms are presented to obtain the solutions. For any initial value, it is proved that the iterative solutions obtained by the proposed algorithms converge to their true values. Finally, simulation examples are given to verify the proposed convergence theorems.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Research on solutions to matrix equations has been very plentiful over the past decades; for example, Zheng et al. studied the least squares solution to matrix equations by means of singular value decomposition [1]; Xie et al. applied the hierarchical principle to solve the general linear matrix equations $\sum_{i=1}^p A_iXB_i + \sum_{i=1}^q C_iX^TD_i = F$ [2]; Zhou et al. proposed a gradient based iterative algorithm to find the unique solution of the general coupled Sylvester matrix equations by the weighted least squares and the gradient search principle [3]; Ding et al. demonstrated hierarchical principle based iterative algorithms in which the general coupled Sylvester matrix equations were treated as the unknown variables or parameters in control systems [4]. Many other matrix equations have been solved in [5–10].

This paper considers a pair of linear matrix equations

$$A_1XB_1 = F_1, \quad A_2XB_2 = F_2, \quad (1)$$

which play an important role in linear system theory and have been actively studied since the 1960s. For instance, Navarra et al. proposed a representation of the generalized solution to (1) [11]; by using the generalized singular value decomposition and the canonical correlation decomposition, Liao and Lei derived a least squares solution with the minimum norm [12]; Liu employed the matrix rank method to provide the necessary and sufficient condition for solving (1) [13]; Yuan et al. proposed a least squares Hermitian solution with the least norm over the skew field of a quaternion [14]; however, difficulties in study of quaternion matrices existed.

Iterative algorithms are not only widely applied in system identification [15–17], but have also been developed for solving linear matrix equations in (1) [18,19]. For example, in [18], Dehghan and Hajarian presented an algorithm to obtain the generalized centro-symmetric solution of (1) with the condition that the equations in (1) are consistent over the generalized centro-symmetric X . In [19], Cai and Chen derived the least Frobenius solutions of (1) iteratively over the bisymmetric matrices. Inspired by the iterative algorithms in [20–22] which demonstrated iterative methods for Sylvester

[☆] This work was supported by the National Natural Science Foundation of China.

^{*} Corresponding author.

E-mail addresses: djvsqxb@163.com (J. Ding), yanjunliu_1983@126.com (Y. Liu), fding@jiangnan.edu.cn (F. Ding).

matrix equations and general coupled matrix equations, we propose a gradient based iterative algorithm and a least squares iterative algorithm to solve the matrix equations in (1), in which we regard the unknown matrix \mathbf{X} as the system parameters to be identified. For any initial value $\mathbf{X}(0)$, the iterative solution $\mathbf{X}(k)$ is proved to converge to the unique solution \mathbf{X} . Compared with [20,21], the main contribution of this paper lies in the fact that the proposed iterative algorithms deal with the coupling matrix equation (1) in which both equations contain the same unknown \mathbf{X} , while [20,21] considered multivariable matrix equations. The objectives of this paper are as follows.

- To find the constraints for unique solution to the matrix equation (1).
- To develop a stochastic gradient algorithm and a least squares algorithm to iteratively generate the approximate solutions.
- To expand the proposed algorithms to generalized matrix equations of the form $\mathbf{A}_i \mathbf{X} \mathbf{B}_i = \mathbf{F}_i$.

The paper is organized as follows. In Section 2, several important lemmas are introduced. In Section 3, the expressions of the iterative solutions to the matrix equations are given. In Section 4, an example is included to verify the convergence of the algorithms. Some brief concluding remarks are given in Section 5.

2. Main preliminaries

Let us introduce some notations. For two matrices \mathbf{M} and \mathbf{N} , $\mathbf{M} \otimes \mathbf{N}$ stands for their Kronecker product. For two matrices \mathbf{X} and \mathbf{Y} with

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}, \quad \mathbf{x}_i \in \mathbb{R}^m,$$

$$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_s] \in \mathbb{R}^{r \times s}, \quad \mathbf{y}_i \in \mathbb{R}^r,$$

$\text{col}[\mathbf{X}]$ is an mn -dimensional vector formed by columns of \mathbf{X} :

$$\text{col}[\mathbf{X}] = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \in \mathbb{R}^{mn},$$

and

$$\text{col}[\mathbf{X}, \mathbf{Y}] = \begin{bmatrix} \text{col}[\mathbf{X}] \\ \text{col}[\mathbf{Y}] \end{bmatrix} \in \mathbb{R}^{mn+rs}.$$

For matrices $\mathbf{C}_i \in \mathbb{R}^{r_i \times s_i}$, $i = 1, 2, \dots, q$, a cell array \mathbf{C} is introduced here which contains matrices of different dimensions in one variable described by

$$\mathbf{C} = \{\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_q\},$$

and the matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$ can be viewed as a special cell array with the same row or column. In order to more succinctly illustrate the iterative algorithms to be proposed later, we introduce the block-matrix inner product – the star (\star) product [21] for short. Let

$$\mathbf{X} := \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_p \end{bmatrix} \in \mathbb{R}^{(mp) \times n}, \quad \mathbf{Y} := \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_p \end{bmatrix} \in \mathbb{R}^{(np) \times m}, \quad \mathbf{X}_i, \mathbf{Y}_i^T \in \mathbb{R}^{m \times n}, \quad i = 1, 2, \dots, p;$$

then the block-matrix star product \star is defined as

$$\mathbf{X} \star \mathbf{Y} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_p \end{bmatrix} \star \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_p \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \mathbf{Y}_1 \\ \mathbf{X}_2 \mathbf{Y}_2 \\ \vdots \\ \mathbf{X}_p \mathbf{Y}_p \end{bmatrix}.$$

Taking into account the dimension compatibility, the star product is superior to matrix multiplication. Note that $\mathbf{A} \mathbf{B} \star \mathbf{C} = \mathbf{A}(\mathbf{B} \star \mathbf{C}) \neq (\mathbf{A} \mathbf{B}) \star \mathbf{C}$. There are some lemmas for better understanding the algorithms proposed later. Consider the matrix equation

$$\mathbf{A} \mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{n \times n}, \quad \mathbf{b} \in \mathbb{R}^n. \tag{2}$$

Here, $\mathbf{A} = [a_{ij}]$ is a full-rank matrix with non-zero diagonal elements, and $\mathbf{x} \in \mathbb{R}^n$ is an unknown vector to be solved. It is well known that, for the full column-rank non-square matrix \mathbf{A} , we have the least squares solution $\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$. Let $\mathbf{x}(k)$ represent the iterative solution to matrix equation (2); it is not difficult to get the iterative solution by the following gradient based iterative algorithm and the least squares algorithm in [20,21].

Lemma 1 ([4,20,21]). For matrix Eq. (2), assume that \mathbf{A} is a full-rank matrix with non-zero diagonal elements; the solution to (2) given by the following gradient based iterative algorithm,

$$\begin{cases} \mathbf{x}(k) = \mathbf{x}(k-1) + \mu \mathbf{A}^T [\mathbf{b} - \mathbf{A}\mathbf{x}(k-1)], \\ 0 < \mu \leq \frac{2}{\lambda_{\max}[\mathbf{A}^T \mathbf{A}]} \quad \text{or} \quad 0 < \mu \leq \frac{2}{\|\mathbf{A}\|^2}, \end{cases}$$

yields $\lim_{k \rightarrow \infty} \mathbf{x}(k) = \mathbf{x}$. Here, $\|\mathbf{A}\|^2 = \text{tr}[\mathbf{A}\mathbf{A}^T]$, and k is the number of iterations.

Lemma 2 ([4,20,21]). If \mathbf{A} is a full-rank matrix with non-zero diagonal elements, then the following least squares based iterative algorithm leads to $\lim_{k \rightarrow \infty} \mathbf{x}(k) = \mathbf{x}$:

$$\mathbf{x}(k) = \mathbf{x}(k-1) + \mu (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T [\mathbf{b} - \mathbf{A}\mathbf{x}(k-1)], \quad 0 < \mu < 2.$$

Next, we give the iterative algorithms for the matrix equation

$$\mathbf{A}\mathbf{X}\mathbf{B} = \mathbf{F}, \tag{3}$$

where $\mathbf{A} \in \mathbb{R}^{p \times m}$, $\mathbf{B} \in \mathbb{R}^{n \times q}$ and $\mathbf{F} \in \mathbb{R}^{p \times q}$ are given constant matrices, and $\mathbf{X} \in \mathbb{R}^{m \times n}$ is the unknown matrix to be solved.

Lemma 3 ([4]). If \mathbf{A} is a full column-rank matrix and \mathbf{B} is a full row-rank matrix ($p \geq m, n \leq q$), the gradient based iterative algorithm for matrix equation (3),

$$\begin{cases} \mathbf{X}(k) = \mathbf{X}(k-1) + \mu \mathbf{A}^T [\mathbf{F} - \mathbf{A}\mathbf{X}(k-1)\mathbf{B}]\mathbf{B}^T, \\ 0 < \mu < \frac{2}{\lambda_{\max}[\mathbf{A}\mathbf{A}^T]\lambda_{\max}[\mathbf{B}^T\mathbf{B}]}, \end{cases}$$

yields $\mathbf{X}(k) \rightarrow \mathbf{X}$.

Lemma 4 ([4]). If \mathbf{A} is a full column-rank matrix and \mathbf{B} is a full row-rank matrix ($p \geq m, n \leq q$), then the least squares based iterative algorithm for matrix equation (3),

$$\mathbf{X}(k) = \mathbf{X}(k-1) + \mu (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T [\mathbf{F} - \mathbf{A}\mathbf{X}(k-1)\mathbf{B}]\mathbf{B}^T (\mathbf{B}\mathbf{B}^T)^{-1}, \quad 0 < \mu < 2,$$

yields $\mathbf{X}(k) \rightarrow \mathbf{X}$.

3. Iterative algorithms

In order to well represent the solutions to the matrix equations in (1), we rewrite them as follows:

$$\begin{cases} \mathbf{A}_1 \mathbf{X} \mathbf{B}_1 = \mathbf{F}_1, \\ \mathbf{A}_2 \mathbf{X} \mathbf{B}_2 = \mathbf{F}_2, \end{cases} \tag{4}$$

where $\mathbf{A}_i \in \mathbb{R}^{p_i \times m}$, $\mathbf{B}_i \in \mathbb{R}^{n \times q_i}$ and $\mathbf{F}_i \in \mathbb{R}^{p_i \times q_i}$ are given constant matrices ($i = 1, 2$), and $\mathbf{X} \in \mathbb{R}^{m \times n}$ is the unknown matrix to be solved. Define

$$\mathbf{S} := \begin{bmatrix} \mathbf{B}_1^T \otimes \mathbf{A}_1 \\ \mathbf{B}_2^T \otimes \mathbf{A}_2 \end{bmatrix} \in \mathbb{R}^{(p_1 q_1 + p_2 q_2) \times (mn)}.$$

Lemma 5 ([4]). Matrix equation (4) has a unique solution if and only if $\text{rank}\{\mathbf{S}, \text{col}[\mathbf{F}_1, \mathbf{F}_2]\} = \text{rank}\{\mathbf{S}\} = mn$; in this case, the unique solution is given by

$$\text{col}[\mathbf{X}] = [\mathbf{S}^T \mathbf{S}]^{-1} \mathbf{S}^T \text{col}[\mathbf{F}_1, \mathbf{F}_2], \tag{5}$$

and the corresponding homogeneous matrix equations $\mathbf{A}_1 \mathbf{X} \mathbf{B}_1 = \mathbf{0}, \mathbf{A}_2 \mathbf{X} \mathbf{B}_2 = \mathbf{0}$ have a unique solution: $\mathbf{X} = \mathbf{0}$.

Define

$$\mathbf{G} := \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}, \quad \mathbf{H} := [\mathbf{B}_1, \mathbf{B}_2].$$

According to Lemmas 1 and 3 and referring to [4], if \mathbf{A}_i is a non-square $p_i \times m$ full column-rank matrix and \mathbf{B}_i is a non-square $n \times q_i$ full row-rank matrix, then we have the gradient based iterative (GI) algorithm described in the following:

$$\mathbf{X}(k) = \mathbf{X}(k-1) + \mu \mathbf{G}^T \begin{Bmatrix} \mathbf{F}_1 - \mathbf{A}_1 \mathbf{X}(k-1) \mathbf{B}_1 \\ \mathbf{F}_2 - \mathbf{A}_2 \mathbf{X}(k-1) \mathbf{B}_2 \end{Bmatrix} \star \mathbf{H}^T, \tag{6}$$

$$0 < \mu < \frac{2}{\lambda_{\max}[\mathbf{G}\mathbf{G}^T]\lambda_{\max}[\mathbf{H}^T\mathbf{H}]} =: \mu_0. \tag{7}$$

To initialize the GI algorithm, we take $\mathbf{X}(0) = \mathbf{0}$ or some small real matrix, e.g., $\mathbf{X}(0) = 10^{-6}\mathbf{1}_{m \times n}$, with $\mathbf{1}_{m \times n}$ being an $m \times n$ matrix whose elements are 1.

Theorem 1. *If the matrix equations in (4) have a unique solution \mathbf{X} , then the iterative solution $\mathbf{X}(k)$ given by the algorithm in (6) converges to \mathbf{X} , i.e., $\lim_{k \rightarrow \infty} \mathbf{X}(k) = \mathbf{X}$, or the error $\mathbf{X}(k) - \mathbf{X}$ converges to zero for any initial value $\mathbf{X}(0)$.*

Proof. Define the error matrix

$$\tilde{\mathbf{X}}(k) := \mathbf{X}(k) - \mathbf{X}.$$

Using (4) and (6), we have

$$\begin{aligned} \tilde{\mathbf{X}}(k) &= \tilde{\mathbf{X}}(k-1) + \mu \mathbf{G}^T \begin{Bmatrix} \mathbf{A}_1 \mathbf{X} \mathbf{B}_1 - \mathbf{A}_1 \mathbf{X}(k-1) \mathbf{B}_1 \\ \mathbf{A}_2 \mathbf{X} \mathbf{B}_2 - \mathbf{A}_2 \mathbf{X}(k-1) \mathbf{B}_2 \end{Bmatrix} \star \mathbf{H}^T \\ &= \tilde{\mathbf{X}}(k-1) - \mu \mathbf{G}^T \begin{Bmatrix} \mathbf{A}_1 \tilde{\mathbf{X}}(k-1) \mathbf{B}_1 \\ \mathbf{A}_2 \tilde{\mathbf{X}}(k-1) \mathbf{B}_2 \end{Bmatrix} \star \mathbf{H}^T \\ &= \tilde{\mathbf{X}}(k-1) - \mu \mathbf{G}^T \mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H} \mathbf{H}^T. \end{aligned}$$

Using the formula $\text{tr}[\mathbf{AB}] = \text{tr}[\mathbf{BA}]$ and $\text{tr}[\mathbf{A}^T] = \text{tr}[\mathbf{A}]$ gives

$$\begin{aligned} \|\tilde{\mathbf{X}}(k)\|^2 &= \text{tr}[\tilde{\mathbf{X}}^T(k) \tilde{\mathbf{X}}(k)] \\ &= \|\tilde{\mathbf{X}}(k-1)\|^2 - 2\mu \text{tr}[\tilde{\mathbf{X}}^T(k-1) \mathbf{G}^T \mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H} \mathbf{H}^T] + \mu^2 \|\mathbf{G}^T \mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H} \mathbf{H}^T\|^2 \\ &= \|\tilde{\mathbf{X}}(k-1)\|^2 - 2\mu \text{tr}[\mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H} \mathbf{H}^T \tilde{\mathbf{X}}^T(k-1) \mathbf{G}^T] + \mu^2 \|\mathbf{G}^T \mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H} \mathbf{H}^T\|^2 \\ &= \|\tilde{\mathbf{X}}(k-1)\|^2 - 2\mu \|\mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H}\|^2 + \mu^2 \|\mathbf{G}^T \mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H} \mathbf{H}^T\|^2 \\ &\leq \|\tilde{\mathbf{X}}(k-1)\|^2 - 2\mu \|\mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H}\|^2 + \mu^2 \lambda_{\max}[\mathbf{G} \mathbf{G}^T] \lambda_{\max}[\mathbf{H}^T \mathbf{H}] \|\mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H}\|^2 \\ &= \|\tilde{\mathbf{X}}(k-1)\|^2 - \mu \{2 - \mu \lambda_{\max}[\mathbf{G} \mathbf{G}^T] \lambda_{\max}[\mathbf{H}^T \mathbf{H}]\} \|\mathbf{G} \tilde{\mathbf{X}}(k-1) \mathbf{H}\|^2 \\ &\leq \|\tilde{\mathbf{X}}(0)\|^2 - \mu \{2 - \mu \lambda_{\max}[\mathbf{G} \mathbf{G}^T] \lambda_{\max}[\mathbf{H}^T \mathbf{H}]\} \sum_{i=1}^k \|\mathbf{G} \tilde{\mathbf{X}}(i-1) \mathbf{H}\|^2. \end{aligned}$$

Using (7) gives

$$\sum_{k=1}^{\infty} \|\mathbf{G} \tilde{\mathbf{X}}(k) \mathbf{H}\|^2 < \infty.$$

It follows that

$$\mathbf{G} \tilde{\mathbf{X}}(k) \mathbf{H} \rightarrow 0, \quad \text{as } k \rightarrow \infty.$$

According to Lemma 5, we have $\tilde{\mathbf{X}} \rightarrow 0$ as $k \rightarrow \infty$. \square

Also, according to Lemmas 2 and 4 and referring to [4], if \mathbf{A}_i is a non-square $p_i \times m$ full column-rank matrix and \mathbf{B}_i is a non-square $n \times q_i$ full row-rank matrix, the least squares based iterative algorithm (LSI) is described as follows:

$$\mathbf{X}(k) = \mathbf{X}(k-1) + \mu (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \begin{Bmatrix} \mathbf{F}_1 - \mathbf{A}_1 \mathbf{X}(k-1) \mathbf{B}_1 \\ \mathbf{F}_2 - \mathbf{A}_2 \mathbf{X}(k-1) \mathbf{B}_2 \end{Bmatrix} \star \mathbf{H}^T (\mathbf{H} \mathbf{H}^T)^{-1}, \quad 0 < \mu < 2. \tag{8}$$

Theorem 2. *If the matrix equations in (4) have a unique solution \mathbf{X} , then the iterative solution $\mathbf{X}(k)$ given by the algorithm in (8) converges to \mathbf{X} , i.e., $\lim_{k \rightarrow \infty} \mathbf{X}(k) = \mathbf{X}$, or the error $\mathbf{X}(k) - \mathbf{X}$ converges to zero for any initial value $\mathbf{X}(0)$.*

The proof can be obtained in a similar way to that above, and is omitted here. The proposed algorithms can be also applied to the generalized matrix equations:

$$\begin{cases} \mathbf{A}_1 \mathbf{X} \mathbf{B}_1 = \mathbf{F}_1, \\ \mathbf{A}_2 \mathbf{X} \mathbf{B}_2 = \mathbf{F}_2, \\ \vdots \\ \mathbf{A}_p \mathbf{X} \mathbf{B}_p = \mathbf{F}_p. \end{cases} \tag{9}$$

Define

$$\mathbf{G}_p := \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_p \end{bmatrix}, \quad \mathbf{H}_p := [\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_p];$$

Table 1

The iterative solutions of the GI algorithm with $\mu = 1/52.63$.

k	x_{11}	x_{12}	x_{21}	x_{22}	δ (%)
1	-0.944120	-5.521457	1.444380	3.050525	41.498997
2	-0.130592	-4.723496	1.253092	1.162763	25.947768
5	0.603962	-5.058627	1.908965	1.919378	8.171622
10	0.932533	-5.181802	1.979618	1.682126	1.268198
20	0.998122	-5.199455	1.999776	1.699733	0.033617
25	0.999688	-5.199899	1.999996	1.700036	0.005581
Solution	1.000000	-5.200000	2.000000	1.700000	

the GI based solution can be expressed as

$$\mathbf{X}(k) = \mathbf{X}(k - 1) + \mu \mathbf{G}_p^T \begin{Bmatrix} \mathbf{F}_1 - \mathbf{A}_1 \mathbf{X}(k - 1) \mathbf{B}_1 \\ \mathbf{F}_2 - \mathbf{A}_2 \mathbf{X}(k - 1) \mathbf{B}_2 \\ \vdots \\ \mathbf{F}_p - \mathbf{A}_p \mathbf{X}(k - 1) \mathbf{B}_p \end{Bmatrix} \star \mathbf{H}_p^T, \quad 0 < \mu \leq 2 \left(\sum_{i=1}^p \|\mathbf{A}_i\|^2 \|\mathbf{B}_i\|^2 \right)^{-1}.$$

Similarly, one can easily give the LSI solution to the matrix equations in (9):

$$\mathbf{X}(k) = \mathbf{X}(k - 1) + \mu (\mathbf{G}_p^T \mathbf{G}_p)^{-1} \mathbf{G}_p^T \begin{Bmatrix} \mathbf{F}_1 - \mathbf{A}_1 \mathbf{X}(k - 1) \mathbf{B}_1 \\ \mathbf{F}_2 - \mathbf{A}_2 \mathbf{X}(k - 1) \mathbf{B}_2 \\ \vdots \\ \mathbf{F}_p - \mathbf{A}_p \mathbf{X}(k - 1) \mathbf{B}_p \end{Bmatrix} \star \mathbf{H}_p^T (\mathbf{H}_p \mathbf{H}_p^T)^{-1}, \quad 0 < \mu \leq 2.$$

4. Numerical examples

Example 1. Consider the following coupled matrix equations:

$$\mathbf{A}_1 \mathbf{X} \mathbf{B}_1 = \mathbf{F}_1, \quad \mathbf{A}_2 \mathbf{X} \mathbf{B}_2 = \mathbf{F}_2$$

with

$$\mathbf{A}_1 = \begin{bmatrix} 1.00 & 1.00 \\ 2.00 & -1.00 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 1.00 & 3.00 \\ -2.00 & 1.00 \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} 1.00 & -1.00 \\ 2.00 & 0.80 \end{bmatrix},$$

$$\mathbf{B}_2 = \begin{bmatrix} 1.00 & 1.00 \\ 2.50 & -1.00 \end{bmatrix}, \quad \mathbf{F}_1 = \begin{bmatrix} -4.00 & -5.80 \\ -24.20 & -9.68 \end{bmatrix}, \quad \mathbf{F}_2 = \begin{bmatrix} 6.75 & 7.10 \\ 30.25 & -12.10 \end{bmatrix}.$$

Then the solution \mathbf{X} from (5) is

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1.00 & -5.20 \\ 2.00 & 1.70 \end{bmatrix}.$$

Taking $\mathbf{X}(0) = 10^{-6} \mathbf{1}_{2 \times 2}$, we apply the GI algorithm in (6) to compute $\mathbf{X}(k)$. The iterative solutions $\mathbf{X}(k)$ are shown in Table 1, where $\delta := \|\mathbf{X}(k) - \mathbf{X}\|/\|\mathbf{X}\|$ is the relative error. The errors δ versus k with different convergence factors are shown in Fig. 1.

As shown in Table 1 and Fig. 1, the convergence factor μ_0 in (6) may be not the best but the conservative one, when $\mu = 1/52.63$ or larger, the better convergence will be obtained. δ becomes smaller and smaller and goes to zero within several iterations. This indicates that the gradient based iterative algorithm is effective.

Taking $\mu = 0.88$ and $\mu = 1.34$, respectively, and applying the LSI algorithm in (8), the errors versus k are shown in Table 2 and Fig. 2.

From Tables 1–2 and Figs. 1–2, the iterative solutions $\mathbf{X}(k)$ obtained by using the LSI algorithm converge faster than those obtained by using the GI algorithm in Example 1.

Example 2. Consider the following coupled matrix equations:

$$\mathbf{A}_1 \mathbf{X} \mathbf{B}_1 = \mathbf{F}_1, \quad \mathbf{A}_2 \mathbf{X} \mathbf{B}_2 = \mathbf{F}_2$$

with

$$\mathbf{A}_1 = \begin{bmatrix} 1.00 & -0.50 \\ 0.50 & 1.00 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 1.00 & 1.00 \\ -2.00 & 1.00 \\ 1.00 & 1.10 \end{bmatrix},$$

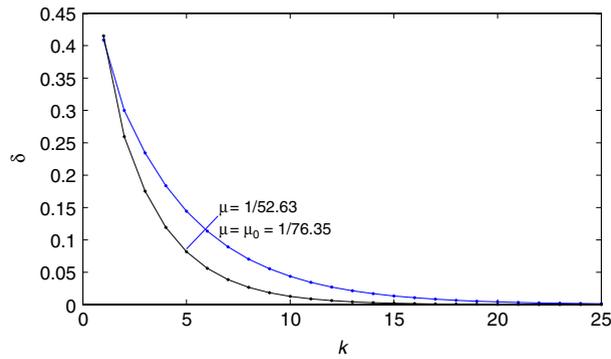


Fig. 1. The relative errors δ versus k of the GI algorithm.

Table 2

The iterative solutions of the GI algorithm with $\mu = 1.34$.

k	x_{11}	x_{12}	x_{21}	x_{22}	δ (%)
1	0.644924	-3.421538	1.304346	1.211674	33.888222
2	0.875295	-4.594682	1.761699	1.553962	11.476654
5	0.994780	-5.176481	1.990840	1.695354	0.443127
10	0.999975	-5.199897	1.999962	1.699981	0.001936
20	1.000000	-5.200000	2.000000	1.700000	0.000000
25	1.000000	-5.200000	2.000000	1.700000	0.000000
Solution	1.000000	-5.200000	2.000000	1.700000	0.000000

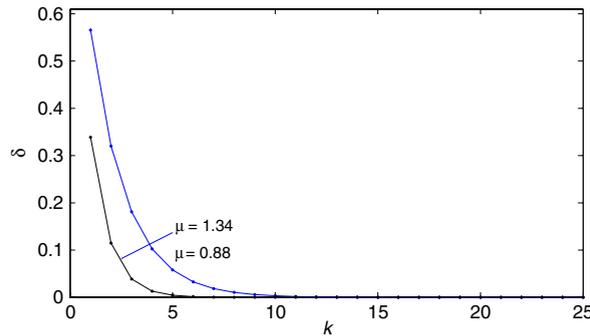


Fig. 2. The relative errors δ versus k of the LSI algorithm.

$$\mathbf{B}_1 = \begin{bmatrix} 1.00 & 3.00 & 0.80 & 2.00 \\ 1.00 & 1.00 & -3.00 & 1.20 \\ -1.10 & -2.10 & 5.00 & 1.00 \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} 1.00 & 2.60 & 0.80 \\ 2.50 & -1.10 & 1.00 \\ 1.00 & -1.50 & 2.00 \end{bmatrix},$$

$$\mathbf{F}_1 = \begin{bmatrix} -8.175 & -13.925 & 23.50 & -4.63 \\ 3.925 & 13.675 & 11.25 & 12.01 \end{bmatrix}, \quad \mathbf{F}_2 = \begin{bmatrix} 4.35 & 13.57 & 8.30 \\ 15.75 & 8.02 & -1.30 \\ 5.165 & 14.742 & 8.81 \end{bmatrix}.$$

Then the solution \mathbf{X} from (5) is

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix} = \begin{bmatrix} 1.00 & -3.00 & 2.70 \\ 5.00 & 1.30 & -0.10 \end{bmatrix}.$$

Taking $\mathbf{X}(0) = 10^{-6} \mathbf{1}_{2 \times 3}$, we apply the algorithms in (6) and (8) to compute $\mathbf{X}(k)$, respectively. The iterative solutions $\mathbf{X}(k)$ are shown in Tables 3–4 and Figs. 3–4.

In Example 2, a larger μ makes the LSI algorithm work as effectively as the GI algorithm. As shown in Figs. 1–4, both the proposed algorithms are effective, but as shown in the algorithms (6) and (8), the computational burden of the GI algorithm is much less than that of the LSI algorithm.

Table 3

The iterative solutions of the GI algorithm with $\mu = 1/66.67$.

k	x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}	δ (%)
1	0.363978	-2.216651	2.516167	3.534082	1.303469	-0.568099	27.882781
2	0.920848	-2.874524	2.472307	4.494598	1.423831	-0.263796	9.189076
5	0.997283	-2.994392	2.705240	4.970215	1.324223	-0.106211	0.599080
10	1.000030	-3.000051	2.699929	4.999561	1.300473	-0.100077	0.009894
20	1.000000	-3.000000	2.700000	5.000000	1.300000	-0.100000	0.000003
25	1.000000	-3.000000	2.700000	5.000000	1.300000	-0.100000	0.000000
Solution	1.000000	-3.000000	2.700000	5.000000	1.300000	-0.100000	

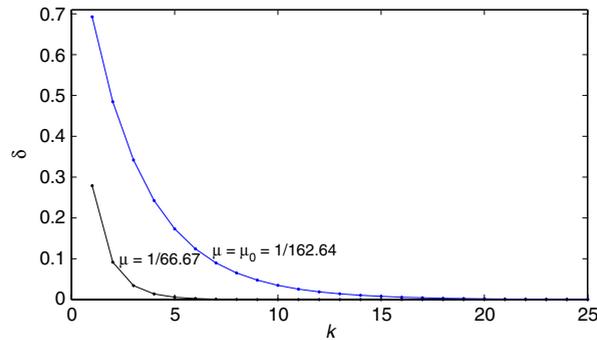


Fig. 3. The relative errors δ versus k of the GI algorithm.

Table 4

The iterative solutions of the LSI algorithm with $\mu = 1.8$.

k	x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}	δ (%)
1	0.583109	-1.546975	0.685887	3.955392	0.939774	0.119684	41.595160
2	0.724009	-1.890218	1.499962	4.793526	1.154572	0.009876	25.335799
5	0.936698	-2.717858	2.366530	4.998992	1.294205	-0.093802	6.656006
10	0.992841	-2.967804	2.662443	4.999820	1.300806	-0.100940	0.753861
20	0.999908	-2.999587	2.699518	4.999997	1.300015	-0.100017	0.009676
25	0.999990	-2.999953	2.699945	5.000000	1.300002	-0.100002	0.001096
Solution	1.000000	-3.000000	2.700000	5.000000	1.300000	-0.100000	

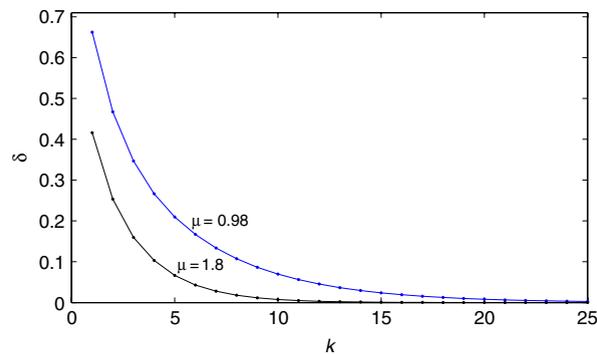


Fig. 4. The relative errors δ versus k of the LSI algorithm.

5. Conclusions

In this paper, iterative algorithms are established to solve the matrix equations $A_1XB_1 = F_1$ and $A_2XB_2 = F_2$. The algorithms can generate an iterative solution which converges to the true solution. But how to choose the convergence factor or step size μ to ensure its properties is worth further study.

References

[1] B. Zheng, L. Ye, D.S. Cvetkovic-Ilic, The congruence class of the solutions of some matrix equations, Computers & Mathematics with Applications 57 (4) (2009) 540–549.
 [2] L. Xie, J. Ding, F. Ding, Gradient based iterative solutions for general linear matrix equations, Computers & Mathematics with Applications 58 (7) (2009) 1441–1448.

- [3] B. Zhou, Z.Y. Li, G.R. Duan, Y. Wang, Weighted least squares solutions to general coupled Sylvester matrix equations, *Journal of Computational and Applied Mathematics* 224 (2) (2009) 759–776.
- [4] 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 (1) (2008) 41–50.
- [5] 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 (2) (2008) 571–588.
- [6] M. Dehghan, M. Hajarian, Finite iterative algorithms for the reflexive and anti-reflexive solutions of the matrix equation $A_1X_1B_1 + A_2X_2B_2 = C$, *Mathematical and Computer Modelling* 49 (9–10) (2009) 1937–1959.
- [7] J.Q. Qiu, W.H. Jiang, Y. Shi, D.M. Xi, Positive solutions for nonlinear n th-order m -point boundary value problem with the first derivative, *International Journal of Innovative Computing, Information and Control* 5 (8) (2009) 2405–2414.
- [8] X.Z. Zhong, T. Zhang, Y. Shi, Oscillation and nonoscillation of neutral difference equation with positive and negative coefficients, *International Journal of Innovative Computing, Information and Control* 5 (5) (2009) 1329–1342.
- [9] C. Yu, Existence and uniqueness of the solution for FM-BEM based on GMRES(m) algorithm, *ICIC Express Letters* 2 (1) (2008) 89–93.
- [10] J. Liu, C. Yu, Y. Chen, Xia Li, Computational formulations for the fundamental solution and kernel functions of elasto-plastic FM-BEM in spherical coordinate system, *ICIC Express Letters* 2 (2) (2008) 207–212.
- [11] A. Navarra, P.L. Odell, D.M. Young, A representation of the general common solution to the matrix equations $A_1XB_1 = C_1$ and $A_2XB_2 = C_2$ with applications, *Computers & Mathematics with Applications* 41 (7–8) (2001) 929–935.
- [12] A. Liao, Y. Lei, Least-squares solution with the minimum-norm for the matrix equation $(AXB, GXH) = (C, D)$, *Computers & Mathematics with Applications* 50 (3–4) (2005) 539–549.
- [13] Y.H. Liu, Ranks of least squares solutions of the matrix equation $AXB = C$, *Computers & Mathematics with Applications* 55 (6) (2008) 1270–1278.
- [14] S. Yuan, A. Liao, Y. Lei, Least squares Hermitian solution of the matrix equation $(AXB, CXD) = (E, F)$ with the least norm over the skew field of quaternions, *Mathematical and Computer Modelling* 48 (1–2) (2008) 91–100.
- [15] D.Q. Wang, F. Ding, Extended stochastic gradient identification algorithms for Hammerstein–Wiener ARMAX systems, *Computers & Mathematics with Applications* 56 (12) (2008) 3157–3164.
- [16] J. Ding, F. Ding, The residual based extended least squares identification method for dual-rate systems, *Computers & Mathematics with Applications* 56 (6) (2008) 1479–1487.
- [17] L.L. Han, F. Ding, Multi-innovation stochastic gradient algorithms for multi-input multi-output systems, *Digital Signal Processing* 19 (4) (2009) 545–554.
- [18] M. Dehghan, M. Hajarian, An iterative algorithm for solving a pair of matrix equations $AYB = E$, $CYD = F$ over generalized centro-symmetric matrices, *Computers & Mathematics with Applications* 56 (12) (2008) 3246–3260.
- [19] J. Cai, G. Chen, An iterative algorithm for the least squares bisymmetric solutions of the matrix equations $A_1XB_1 = C_1$ and $A_2XB_2 = C_2$, *Mathematical and Computer Modelling* 50 (7–8) (2009) 1237–1244.
- [20] F. Ding, T. Chen, Iterative least squares solutions of coupled Sylvester matrix equations, *Systems & Control Letters* 54 (2) (2005) 95–107.
- [21] F. Ding, T. Chen, On iterative solutions of general coupled matrix equations, *SIAM Journal on Control and Optimization* 44 (6) (2006) 2269–2284.
- [22] F. Ding, T. Chen, Gradient based iterative algorithms for solving a class of matrix equations, *IEEE Transactions on Automatic Control* 50 (8) (2005) 1216–1221.