Computers and Mathematics with Applications 58 (2009) 1441–1448

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Gradient based iterative solutions for general linear matrix equations[☆]

Li Xie[*], Jie Ding, Feng Ding[1]

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

ARTICLE INFO

Article history:
Received 8 October 2008
Received in revised form 3 June 2009
Accepted 17 June 2009

Keywords:
Lyapunov matrix equations
Sylvester matrix equations
Iterations
Least-squares
Estimation

ABSTRACT

In this paper, we present a gradient based iterative algorithm for solving general linear matrix equations by extending the Jacobi iteration and by applying the hierarchical identification principle. Convergence analysis indicates that the iterative solutions always converge fast to the exact solutions for any initial values and small condition numbers of the associated matrices. Two numerical examples are provided to show that the proposed algorithm is effective.

1. Introduction

Matrix equations are often encountered in systems and control, 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, however, which involve the inversion of the associated large matrix and result in increasing computation and excessive computer memory. By extending the Jacobi iteration [1], the gradient based and least-squares based iterative methods proposed in [2,3] can be used to solve the general matrix equation:

$$\sum_{i=1}^{p} \boldsymbol{A}_i \boldsymbol{X} \boldsymbol{B}_i = \boldsymbol{F},$$

which includes the Sylvester matrix equation as a special form. But the method there is not suitable for solving general linear matrix equations in (1) in the next section, which include the Lyapunov equations, Sylvester equations as the special cases, e.g., [4].

Iterative approaches for solving matrix equations and recursive identification for parameter estimation have received much attention, e.g., [5–10]. For example, Dehghan and Hajarian studied the iterative algorithm for the reflexive solutions of the generalized coupled Sylvester matrix equations [11]; Mukaidani et al. gave a numerical algorithm for finding solution of cross-coupled algebraic Riccati equations [12]; Zhou et al. studied the explicit solutions to generalized Sylvester matrix equations [13,14]. Also, Kilicman et al. presented the vector least-squares solutions for coupled singular matrix equations [5]; Ding and Chen presented a gradient based and a least-squares based iterative algorithms for generalized Sylvester matrix equations and general coupled matrix equations by introducing the star ($\star$) product of matrices [15,16]. Finally, Al Zhour et al. discussed some new connections between matrix products for partitioned and non-partitioned matrices, including the star product [17] and the solutions of other matrix equations can be found in [18–20].

☆ This work was supported by the Scientific Research Foundation for the Returned Overseas Chinese Scholars (State Education Ministry).

* Corresponding author.
E-mail addresses: xieli2412@126.com (L. Xie), djvsqxb@163.com (J. Ding), fding@jiangnan.edu.cn (F. Ding).

1 Tel.: +86 510 88637783; fax: +86 510 85910633.

0898-1221/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2009.06.047

This paper decomposes the system in (1) into several subsystems by applying the hierarchical identification principle [15,21,22], regards the unknown matrix $\boldsymbol{X}$ as the system parameter matrix, and presents a gradient based iterative algorithm for solving the matrix equations in (1).

The rest of the paper is organized as follows. Section 2 derives iterative algorithm for solving the general matrix equations in (1) and studies convergence properties of the algorithm. Section 3 provides two examples to illustrate the effectiveness of the proposed algorithm. Finally, we offer some concluding remarks in Section 4.

2. The exact and iterative solutions of general matrix equations

In this section, we apply the hierarchical identification to solve the following general linear matrix equations:

$$\sum_{i=1}^{p}\boldsymbol{A}_i\boldsymbol{X}\boldsymbol{B}_i+\sum_{i=1}^{q}\boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}\boldsymbol{D}_i=\boldsymbol{F}, \tag{1}$$

where $\boldsymbol{A}_i\in\mathbb{R}^{r\times m}$, $\boldsymbol{B}_i\in\mathbb{R}^{n\times s}$, $\boldsymbol{C}_i\in\mathbb{R}^{r\times n}$, $\boldsymbol{D}_i\in\mathbb{R}^{m\times s}$ and $\boldsymbol{F}=[\boldsymbol{f}_1,\boldsymbol{f}_2,\ldots,\boldsymbol{f}_s]\in\mathbb{R}^{r\times s}$ are given constant matrices, $\boldsymbol{X}\in\mathbb{R}^{m\times n}$ is the unknown matrix to be solved.

Let us introduce some notations first. The symbol $\boldsymbol{I}$ or $\boldsymbol{I}_n$ stands for an identity matrix of appropriate sizes or size $n\times n$. For two matrices $\boldsymbol{M}$ and $\boldsymbol{N}$, $\boldsymbol{M}\otimes\boldsymbol{N}$ is their Kronecker product (called direct product); for an $m\times n$ matrix

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

col$[\boldsymbol{X}]$ is an mn-dimensional vector formed by the columns of $\boldsymbol{X}$, i.e.,

$$\mathrm{col}[\boldsymbol{X}]=\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \vdots\\ \boldsymbol{x}_n\end{bmatrix}\in\mathbb{R}^{mn}.$$

According to the above definitions, the unique solution of the equation $\boldsymbol{AX}+\boldsymbol{XB}=\boldsymbol{F}$ can be expressed as

$$\mathrm{col}[\boldsymbol{X}]=(\boldsymbol{I}\otimes\boldsymbol{A}+\boldsymbol{B}^{\mathrm{T}}\otimes\boldsymbol{I})^{-1}\mathrm{col}[\boldsymbol{F}],$$

if $\boldsymbol{I}\otimes\boldsymbol{A}+\boldsymbol{B}^{\mathrm{T}}\otimes\boldsymbol{I}$ is invertible. Referring to Al Zhour and Kilicman's work [17], let $\boldsymbol{P}_{mn}\in\mathbb{R}^{mn\times mn}$ be a square $mn\times mn$ matrix partitioned into $m\times n$ submatrices such that ijth submatrix has a 1 in its jith position and zeros elsewhere, i.e.,

$$\boldsymbol{P}_{mn}=\sum_{i=1}^{m}\sum_{j=1}^{n}\boldsymbol{E}_{ij}\otimes\boldsymbol{E}_{ij}^{\mathrm{T}},$$

where $\boldsymbol{E}_{ij}=\boldsymbol{e}_i\boldsymbol{e}_j^{\mathrm{T}}$ called an elementary matrix of order $m\times n$, and $\boldsymbol{e}_i$ ($\boldsymbol{e}_j$) is a column vector with a unity in the ith (jth) position and zeros elsewhere of order $m\times 1$ ($n\times 1$). Using this definition, we have

$$\boldsymbol{P}_{mn}\mathrm{col}[\boldsymbol{X}^{\mathrm{T}}]=\mathrm{col}[\boldsymbol{X}],\qquad \boldsymbol{P}_{mn}\boldsymbol{P}_{nm}=\boldsymbol{I}_{mn},\qquad \boldsymbol{P}_{mn}^{\mathrm{T}}=\boldsymbol{P}_{mn}^{-1}=\boldsymbol{P}_{nm}.$$

Thus the solution of equation $\boldsymbol{AX}+\boldsymbol{X}^{\mathrm{T}}\boldsymbol{B}=\boldsymbol{F}$ can be expressed as

$$\mathrm{col}[\boldsymbol{X}]=(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W})^{-1}\boldsymbol{W}^{\mathrm{T}}\mathrm{col}[\boldsymbol{F}],$$

where $\boldsymbol{W}=\boldsymbol{I}\otimes\boldsymbol{A}+(\boldsymbol{B}^{\mathrm{T}}\otimes\boldsymbol{I})\boldsymbol{P}_{nm}$.

The following studies contain the exact and iterative solutions of the general matrix equation in (1).

2.1. The exact solution

Lemma 1. *Let*

$$\boldsymbol{S}:=\sum_{i=1}^{p}\boldsymbol{B}_i^{\mathrm{T}}\otimes\boldsymbol{A}_i+\sum_{i=1}^{q}(\boldsymbol{D}_i^{\mathrm{T}}\otimes\boldsymbol{C}_i)\boldsymbol{P}_{nm},$$

then Eq. (1) has a unique solution if and only if $\mathrm{rank}\{\boldsymbol{S},\mathrm{col}[\boldsymbol{F}]\}=\mathrm{rank}[\boldsymbol{S}]=mn$ *(i.e., $\boldsymbol{S}$ has a full column rank). In this case, the unique solution is given by*

$$\mathrm{col}[\boldsymbol{X}]=(\boldsymbol{S}^{\mathrm{T}}\boldsymbol{S})^{-1}\boldsymbol{S}^{\mathrm{T}}\mathrm{col}[\boldsymbol{F}], \tag{2}$$

and the corresponding homogeneous matrix equation in (1) with $\boldsymbol{F}=\boldsymbol{0}$ has a unique solution $\boldsymbol{X}=\boldsymbol{0}$.

The results of Lemma 1 is obvious, and the proof of which is omitted here. □

2.2. The iterative solution

Eq. (2) can give the solution of (1) but it requires excessive computer memory because of computing the inversion of the large matrix $\boldsymbol{S}^{\mathrm{T}}\boldsymbol{S}$ of size $(mn)\times(mn)$ as the dimension of $\boldsymbol{X}$ increases. This motivates us to study the iterative algorithm to solve (1). The following uses the hierarchical identification principle, regards the unknown matrix $\boldsymbol{X}$ as the parameter matrix of the system in (1), decomposes (1) into $p+q$ subsystems, and then presents a gradient based iterative method to obtain the iterative solution of the parameter matrix of each subsystem.

Define the following matrices:

$$\boldsymbol{Q}_j := \boldsymbol{F} - \sum_{i=1,i\neq j}^{p} \boldsymbol{A}_i\boldsymbol{X}\boldsymbol{B}_i - \sum_{i=1}^{q} \boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}\boldsymbol{D}_i, \quad j=1,2,\ldots,p. \tag{3}$$

$$\boldsymbol{Q}_{p+l} := \boldsymbol{F} - \sum_{i=1}^{p} \boldsymbol{A}_i\boldsymbol{X}\boldsymbol{B}_i - \sum_{i=1,i\neq l}^{q} \boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}\boldsymbol{D}_i, \quad l=1,2,\ldots,q. \tag{4}$$

Then from (1), we obtain $p+q$ fictitious subsystems

Subsystem j: $\quad \boldsymbol{A}_j\boldsymbol{X}\boldsymbol{B}_j = \boldsymbol{Q}_j, \quad j=1,2,\ldots,p.$
Subsystem $p+l$: $\quad \boldsymbol{C}_l\boldsymbol{X}^{\mathrm{T}}\boldsymbol{D}_l = \boldsymbol{Q}_{p+l}, \quad l=1,2,\ldots,q.$

Let $\boldsymbol{X}_i(k)$ be the estimate or iterative solution of $\boldsymbol{X}$ at iteration k, associated with ith subsystem. Applying the gradient search method [2] or Corollary 3 in [16] to Subsystem i, $i=1,2,\ldots,p+q$, we can obtain the iterative algorithms:

$$\boldsymbol{X}_j(k) = \boldsymbol{X}_j(k-1) + \mu\boldsymbol{A}_j^{\mathrm{T}}[\boldsymbol{Q}_j - \boldsymbol{A}_j\boldsymbol{X}_j(k-1)\boldsymbol{B}_j]\boldsymbol{B}_j^{\mathrm{T}}, \quad j=1,2,\ldots,p. \tag{5}$$

$$\boldsymbol{X}_{p+l}(k) = \boldsymbol{X}_{p+l}(k-1) + \mu\boldsymbol{D}_l[\boldsymbol{Q}_{p+l}^{\mathrm{T}} - \boldsymbol{D}_l^{\mathrm{T}}\boldsymbol{X}_{p+l}(k-1)\boldsymbol{C}_l^{\mathrm{T}}]\boldsymbol{C}_l, \quad l=1,2,\ldots,q. \tag{6}$$

The convergence factor $\mu > 0$ will be given later. Substituting (3) and (4) into (5) and (6) gives

$$\boldsymbol{X}_j(k) = \boldsymbol{X}_j(k-1) + \mu\boldsymbol{A}_j^{\mathrm{T}}\left[\boldsymbol{F} - \sum_{i=1,i\neq j}^{p} \boldsymbol{A}_i\boldsymbol{X}\boldsymbol{B}_i - \sum_{i=1}^{q} \boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}\boldsymbol{D}_i - \boldsymbol{A}_j\boldsymbol{X}_j(k-1)\boldsymbol{B}_j\right]\boldsymbol{B}_j^{\mathrm{T}}, \tag{7}$$

$$\boldsymbol{X}_{p+l}(k) = \boldsymbol{X}_{p+l}(k-1) + \mu\boldsymbol{D}_l\left[\boldsymbol{F} - \sum_{i=1}^{p} \boldsymbol{A}_i\boldsymbol{X}\boldsymbol{B}_i - \sum_{i=1,i\neq l}^{q} \boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}\boldsymbol{D}_i - \boldsymbol{C}_l\boldsymbol{X}_{p+l}^{\mathrm{T}}(k-1)\boldsymbol{D}_l\right]^{\mathrm{T}}\boldsymbol{C}_l. \tag{8}$$

Here, a difficulty arises in that the expressions on the right-hand sides of (7) and (8) contain the unknown matrix $\boldsymbol{X}$; so it is impossible to realize the algorithm. The solution is based on the hierarchical identification principle [15,16,21,22]: the unknown variable $\boldsymbol{X}$ in (7) and (8) is replaced by its estimate $\boldsymbol{X}_j(k-1)$ and $\boldsymbol{X}_{p+l}(k-1)$ at time $(k-1)$. Hence, we have

$$\boldsymbol{X}_j(k) = \boldsymbol{X}_j(k-1) + \mu\boldsymbol{A}_j^{\mathrm{T}}\left[\boldsymbol{F} - \sum_{i=1}^{p} \boldsymbol{A}_i\boldsymbol{X}_j(k-1)\boldsymbol{B}_i - \sum_{i=1}^{q} \boldsymbol{C}_i\boldsymbol{X}_j^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right]\boldsymbol{B}_j^{\mathrm{T}}, \tag{9}$$

$$\boldsymbol{X}_{p+l}(k) = \boldsymbol{X}_{p+l}(k-1) + \mu\boldsymbol{D}_l\left[\boldsymbol{F} - \sum_{i=1}^{p} \boldsymbol{A}_i\boldsymbol{X}_{p+l}(k-1)\boldsymbol{B}_i - \sum_{i=1}^{q} \boldsymbol{C}_i\boldsymbol{X}_{p+l}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right]^{\mathrm{T}}\boldsymbol{C}_l. \tag{10}$$

In fact, we need only an iterative solution $\boldsymbol{X}(k)$ rather than $p+q$ solutions $\boldsymbol{X}_i(k)$: $i=1,2,\ldots,p+q$. Taking the average of the $p+q$ solutions as the iterative solution $\boldsymbol{X}(k)$ of $\boldsymbol{X}$, we obtain a gradient based iterative (GI) algorithm for the general matrix equation in (1):

$$\boldsymbol{X}(k) = \frac{1}{p+q}\left[\sum_{j=1}^{p} \boldsymbol{X}_j(k) + \sum_{l=1}^{q} \boldsymbol{X}_{p+l}(k)\right], \tag{11}$$

$$\boldsymbol{X}_j(k) = \boldsymbol{X}(k-1) + \mu\boldsymbol{A}_j^{\mathrm{T}}\left[\boldsymbol{F} - \sum_{i=1}^{p} \boldsymbol{A}_i\boldsymbol{X}(k-1)\boldsymbol{B}_i - \sum_{i=1}^{q} \boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right]\boldsymbol{B}_j^{\mathrm{T}}, \tag{12}$$

$$\boldsymbol{X}_{p+l}(k) = \boldsymbol{X}(k-1) + \mu\boldsymbol{D}_l\left[\boldsymbol{F} - \sum_{i=1}^{p} \boldsymbol{A}_i\boldsymbol{X}(k-1)\boldsymbol{B}_i - \sum_{i=1}^{q} \boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right]^{\mathrm{T}}\boldsymbol{C}_l. \tag{13}$$

A conservative choice of the convergence factor μ is

$$0 < \mu < 2\left\{\sum_{j=1}^{p} \lambda_{\max}[\boldsymbol{A}_j\boldsymbol{A}_j^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{B}_j^{\mathrm{T}}\boldsymbol{B}_j] + \sum_{l=1}^{q} \lambda_{\max}[\boldsymbol{C}_l\boldsymbol{C}_l^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{D}_l^{\mathrm{T}}\boldsymbol{D}_l]\right\}^{-1} =: \mu_0. \tag{14}$$

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

Theorem 1. *If Equation in (1) has a unique solution $\boldsymbol{X}$, then the iterative solution $\boldsymbol{X}(k)$ given by the algorithm in (11)–(14) converges to $\boldsymbol{X}$, i.e., $\lim_{k\to\infty}\boldsymbol{X}(k) = \boldsymbol{X}$; or, the error $\boldsymbol{X}(k) - \boldsymbol{X}$ converges to zero for any initial value $\boldsymbol{X}(0)$.*

Referring to the methods in [2,3,15,16], we prove this theorem.

Proof. Define the estimation error matrices:

$$\tilde{\boldsymbol{X}}_i(k) := \boldsymbol{X}_i(k) - \boldsymbol{X},$$
$$\tilde{\boldsymbol{X}}(k) := \boldsymbol{X}(k) - \boldsymbol{X} = \frac{1}{p+q}\left[\sum_{j=1}^{p}\tilde{\boldsymbol{X}}_j(k) + \sum_{l=1}^{q}\tilde{\boldsymbol{X}}_{p+l}(k)\right], \tag{15}$$

and

$$\boldsymbol{\xi}_i(k) := \boldsymbol{A}_i\tilde{\boldsymbol{X}}(k-1)\boldsymbol{B}_i, \qquad \boldsymbol{\eta}_i(k) := \boldsymbol{D}_i^{\mathrm{T}}\tilde{\boldsymbol{X}}(k-1)\boldsymbol{C}_i^{\mathrm{T}}. \tag{16}$$

Using (1), (12)–(13) and (16), it is easy to get

$$\begin{aligned}
\tilde{\boldsymbol{X}}_j(k) &= \boldsymbol{X}_j(k) - \boldsymbol{X} \\
&= \boldsymbol{X}(k-1) - \boldsymbol{X} + \mu\boldsymbol{A}_j^{\mathrm{T}}\left[\boldsymbol{F} - \sum_{i=1}^{p}\boldsymbol{A}_i\boldsymbol{X}(k-1)\boldsymbol{B}_i - \sum_{i=1}^{q}\boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right]\boldsymbol{B}_j^{\mathrm{T}} \\
&= \tilde{\boldsymbol{X}}(k-1) - \mu\boldsymbol{A}_j^{\mathrm{T}}\left[\sum_{i=1}^{p}\boldsymbol{A}_i(\boldsymbol{X}(k-1)-\boldsymbol{X})\boldsymbol{B}_i + \sum_{i=1}^{q}\boldsymbol{C}_i(\boldsymbol{X}^{\mathrm{T}}(k-1)-\boldsymbol{X}^{\mathrm{T}})\boldsymbol{D}_i\right]\boldsymbol{B}_j^{\mathrm{T}} \\
&= \tilde{\boldsymbol{X}}(k-1) - \mu\boldsymbol{A}_j^{\mathrm{T}}\left[\sum_{i=1}^{p}\boldsymbol{A}_i\tilde{\boldsymbol{X}}(k-1)\boldsymbol{B}_i + \sum_{i=1}^{q}\boldsymbol{C}_i\tilde{\boldsymbol{X}}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right]\boldsymbol{B}_j^{\mathrm{T}} \\
&= \tilde{\boldsymbol{X}}(k-1) - \mu\boldsymbol{A}_j^{\mathrm{T}}\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]\boldsymbol{B}_j^{\mathrm{T}}.
\end{aligned} \tag{17}$$

Similarly,

$$\tilde{\boldsymbol{X}}_{p+l}(k) = \tilde{\boldsymbol{X}}(k-1) - \mu\boldsymbol{D}_l\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]^{\mathrm{T}}\boldsymbol{C}_l. \tag{18}$$

Taking the norm of both sides of the above equations and using the formula: $\mathrm{tr}[\boldsymbol{AB}] = \mathrm{tr}[\boldsymbol{BA}]$ and $\mathrm{tr}[\boldsymbol{A}^{\mathrm{T}}] = \mathrm{tr}[\boldsymbol{A}]$ give

$$\begin{aligned}
\|\tilde{\boldsymbol{X}}_j(k)\|^2 &= \mathrm{tr}[\tilde{\boldsymbol{X}}_j^{\mathrm{T}}(k)\tilde{\boldsymbol{X}}_j(k)] \\
&= \left\|\tilde{\boldsymbol{X}}(k-1) - \mu\boldsymbol{A}_j^{\mathrm{T}}\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]\boldsymbol{B}_j^{\mathrm{T}}\right\|^2 \\
&= \|\tilde{\boldsymbol{X}}(k-1)\|^2 - \mu\,\mathrm{tr}\left\{\tilde{\boldsymbol{X}}^{\mathrm{T}}(k-1)\boldsymbol{A}_j^{\mathrm{T}}\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]\boldsymbol{B}_j^{\mathrm{T}}\right\} \\
&\quad - \mu\,\mathrm{tr}\left\{\boldsymbol{B}_j\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i^{\mathrm{T}}(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i(k)\right]\boldsymbol{A}_j\tilde{\boldsymbol{X}}(k-1)\right\} + \mu^2\left\|\boldsymbol{A}_j^{\mathrm{T}}\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]\boldsymbol{B}_j^{\mathrm{T}}\right\|^2 \\
&\le \|\tilde{\boldsymbol{X}}(k-1)\|^2 - 2\mu\,\mathrm{tr}\left\{\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]\boldsymbol{\xi}_j^{\mathrm{T}}(k)\right\} \\
&\quad + \mu^2\lambda_{\max}[\boldsymbol{A}_j\boldsymbol{A}_j^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{B}_j^{\mathrm{T}}\boldsymbol{B}_j]\left\|\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right\|^2.
\end{aligned} \tag{19}$$

Similarly,

$$\begin{aligned}
\|\tilde{\boldsymbol{X}}_{p+l}(k)\|^2 &\le \|\tilde{\boldsymbol{X}}(k-1)\|^2 - 2\mu\,\mathrm{tr}\left\{\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]\boldsymbol{\eta}_l(k)\right\} \\
&\quad + \mu^2\lambda_{\max}[\boldsymbol{C}_l\boldsymbol{C}_l^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{D}_l^{\mathrm{T}}\boldsymbol{D}_l]\left\|\sum_{i=1}^{p}\boldsymbol{\xi}_i(k) + \sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right\|^2.
\end{aligned} \tag{20}$$

Hence, using (19) and (20) and from (15), we have

$$\begin{aligned}
\|\tilde{\boldsymbol{X}}(k)\|^2 &= \frac{1}{(p+q)^2}\left\|\sum_{j=1}^{p}\tilde{\boldsymbol{X}}_j(k)+\sum_{l=1}^{q}\tilde{\boldsymbol{X}}_{p+l}(k)\right\|^2 \\
&\le \frac{1}{p+q}\left(\sum_{j=1}^{p}\|\tilde{\boldsymbol{X}}_j(k)\|^2+\sum_{l=1}^{q}\|\tilde{\boldsymbol{X}}_{p+l}(k)\|^2\right) \\
&\le \frac{1}{p+q}\Bigg\{p\|\tilde{\boldsymbol{X}}(k-1)\|^2-2\mu\,\mathrm{tr}\left\{\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k)+\sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]\left[\sum_{j=1}^{p}\boldsymbol{\xi}_j^{\mathrm{T}}(k)\right]\right\} \\
&\quad +\mu^2\left(\sum_{j=1}^{p}\lambda_{\max}[\boldsymbol{A}_j\boldsymbol{A}_j^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{B}_j^{\mathrm{T}}\boldsymbol{B}_j]\right)\left\|\sum_{i=1}^{p}\boldsymbol{\xi}_i(k)+\sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right\|^2 \\
&\quad +q\|\tilde{\boldsymbol{X}}(k-1)\|^2-2\mu\,\mathrm{tr}\left\{\left[\sum_{i=1}^{p}\boldsymbol{\xi}_i(k)+\sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right]\sum_{l=1}^{q}\boldsymbol{\eta}_l(k)\right\} \\
&\quad +\mu^2\left(\sum_{l=1}^{q}\lambda_{\max}[\boldsymbol{C}_l\boldsymbol{C}_l^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{D}_l^{\mathrm{T}}\boldsymbol{D}_l]\right)\left\|\sum_{i=1}^{p}\boldsymbol{\xi}_i(k)+\sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right\|^2\Bigg\} \\
&= \|\tilde{\boldsymbol{X}}(k-1)\|^2-\frac{1}{p+q}\Bigg\{2\mu-\mu^2\Bigg(\sum_{j=1}^{p}\lambda_{\max}[\boldsymbol{A}_j\boldsymbol{A}_j^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{B}_j^{\mathrm{T}}\boldsymbol{B}_j] \\
&\quad +\sum_{l=1}^{q}\lambda_{\max}[\boldsymbol{C}_l\boldsymbol{C}_l^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{D}_l^{\mathrm{T}}\boldsymbol{D}_l]\Bigg)\Bigg\}\left\|\sum_{i=1}^{p}\boldsymbol{\xi}_i(k)+\sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right\|^2 \\
&\le \|\tilde{\boldsymbol{X}}(0)\|^2-\frac{\mu}{p+q}\Bigg\{2-\mu\Bigg(\sum_{j=1}^{p}\lambda_{\max}[\boldsymbol{A}_j\boldsymbol{A}_j^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{B}_j^{\mathrm{T}}\boldsymbol{B}_j] \\
&\quad +\sum_{l=1}^{q}\lambda_{\max}[\boldsymbol{C}_l\boldsymbol{C}_l^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{D}_l^{\mathrm{T}}\boldsymbol{D}_l]\Bigg)\Bigg\}\left(\sum_{j=1}^{k}\left\|\sum_{i=1}^{p}\boldsymbol{\xi}_i(j)+\sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(j)\right\|^2\right).
\end{aligned}$$

If the convergence factor μ is chosen to satisfy

$$0<\mu<2\left\{\sum_{j=1}^{p}\lambda_{\max}[\boldsymbol{A}_j\boldsymbol{A}_j^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{B}_j^{\mathrm{T}}\boldsymbol{B}_j]+\sum_{l=1}^{q}\lambda_{\max}[\boldsymbol{C}_l\boldsymbol{C}_l^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{D}_l^{\mathrm{T}}\boldsymbol{D}_l]\right\}^{-1},$$

then we have

$$\sum_{j=1}^{k}\left\|\sum_{i=1}^{p}\boldsymbol{\xi}_i(j)+\sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(j)\right\|^2<\infty.$$

For the necessary condition of the series convergence, when $k\to\infty$, we have

$$\left\|\sum_{i=1}^{p}\boldsymbol{\xi}_i(k)+\sum_{i=1}^{q}\boldsymbol{\eta}_i^{\mathrm{T}}(k)\right\|^2\to 0,$$

or

$$\left\|\sum_{i=1}^{p}\boldsymbol{A}_i\tilde{\boldsymbol{X}}(k-1)\boldsymbol{B}_i+\sum_{i=1}^{q}\boldsymbol{C}_i\tilde{\boldsymbol{X}}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right\|^2\to 0.$$

According to Lemma 1, we can get $\tilde{\boldsymbol{X}}(k-1)\to\boldsymbol{0}$ as $k\to\infty$. This proves Theorem 1. □

Next, we show that the convergence rate of the gradient based iterative algorithm in (11)–(14) depends on the condition number of the associated matrix $\boldsymbol{\Phi}$ below, like the iterative algorithm of the equation $\boldsymbol{Ax}=\boldsymbol{b}$ [1,2]. From (15)–(18), we have

$$\tilde{\boldsymbol{X}}(k)=\tilde{\boldsymbol{X}}(k-1)-\frac{\mu}{p+q}\sum_{j=1}^{p}\sum_{i=1}^{p}\boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{A}_i\tilde{\boldsymbol{X}}(k-1)\boldsymbol{B}_i\boldsymbol{B}_j^{\mathrm{T}}-\frac{\mu}{p+q}\sum_{l=1}^{q}\sum_{i=1}^{q}\boldsymbol{D}_l\boldsymbol{D}_i^{\mathrm{T}}\tilde{\boldsymbol{X}}(k-1)\boldsymbol{C}_i^{\mathrm{T}}\boldsymbol{C}_l$$

$$- \frac{\mu}{p+q}\sum_{j=1}^{p}\sum_{i=1}^{q}\boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{C}_i\tilde{\boldsymbol{X}}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\boldsymbol{B}_j^{\mathrm{T}} - \frac{\mu}{p+q}\sum_{l=1}^{q}\sum_{i=1}^{p}\boldsymbol{D}_l\boldsymbol{B}_i^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}(k-1)\boldsymbol{A}_i^{\mathrm{T}}\boldsymbol{C}_l,$$

which can be equivalently expressed as

$$\mathrm{col}[\tilde{\boldsymbol{X}}(k)] = \left(\boldsymbol{I}_{mn} - \frac{\mu}{p+q}\boldsymbol{\Phi}\right)\mathrm{col}[\tilde{\boldsymbol{X}}(k-1)], \tag{21}$$

where

$$\boldsymbol{\Phi} := \sum_{j=1}^{p}\sum_{i=1}^{p}\boldsymbol{B}_j\boldsymbol{B}_i^{\mathrm{T}} \otimes \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{A}_i + \sum_{l=1}^{q}\sum_{i=1}^{q}\boldsymbol{C}_l^{\mathrm{T}}\boldsymbol{C}_i \otimes \boldsymbol{D}_l\boldsymbol{D}_i^{\mathrm{T}} + \sum_{j=1}^{p}\sum_{i=1}^{q}(\boldsymbol{B}_j\boldsymbol{D}_i^{\mathrm{T}} \otimes \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{C}_i)\boldsymbol{P}_{nm} + \sum_{l=1}^{q}\sum_{i=1}^{p}(\boldsymbol{C}_l^{\mathrm{T}}\boldsymbol{A}_i \otimes \boldsymbol{D}_l\boldsymbol{B}_i^{\mathrm{T}})\boldsymbol{P}_{nm}. \tag{22}$$

From (21), we can see that the closer the eigenvalues of $\frac{\mu}{p+q}\boldsymbol{\Phi}$ are to 1, the closer the eigenvalues of $\boldsymbol{I}_{mn} - \frac{\mu}{p+q}\boldsymbol{\Phi}$ tend to be zero, and hence, the faster the error $\mathrm{col}[\tilde{\boldsymbol{X}}(k)]$ or $\tilde{\boldsymbol{X}}(k)$ converges to zero. In other words, the gradient based iterative algorithm in (11)–(14) has a fast convergence rate for small condition numbers of $\boldsymbol{\Phi}$.

Similarly, by means of the hierarchical identification principle and referring to [3,15], we can obtain the least-squares based iterative (LSI) algorithm of (1):

$$\boldsymbol{X}_j(k) = \boldsymbol{X}(k-1) + \mu(\boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{A}_j)^{-1}\boldsymbol{A}_j^{\mathrm{T}}\left[\boldsymbol{F} - \sum_{i=1}^{p}\boldsymbol{A}_i\boldsymbol{X}(k-1)\boldsymbol{B}_i - \sum_{i=1}^{q}\boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right]\boldsymbol{B}_j^{\mathrm{T}}(\boldsymbol{B}_j\boldsymbol{B}_j^{\mathrm{T}})^{-1}, \quad j = 1, 2, \ldots, p, \tag{23}$$

$$\boldsymbol{X}_{p+l}(k) = \boldsymbol{X}(k-1) + \mu(\boldsymbol{D}_l\boldsymbol{D}_l^{\mathrm{T}})^{-1}\boldsymbol{D}_l\left[\boldsymbol{F} - \sum_{i=1}^{p}\boldsymbol{A}_i\boldsymbol{X}(k-1)\boldsymbol{B}_i - \sum_{i=1}^{q}\boldsymbol{C}_i\boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{D}_i\right]^{\mathrm{T}}\boldsymbol{C}_l(\boldsymbol{C}_l^{\mathrm{T}}\boldsymbol{C}_l)^{-1}, \quad l = 1, 2, \ldots, q, \tag{24}$$

$$\boldsymbol{X}(k) = \frac{1}{p+q}\left[\sum_{j=1}^{p}\boldsymbol{X}_j(k) + \sum_{l=1}^{q}\boldsymbol{X}_{p+l}(k)\right], \quad 0 < \mu < 2(p+q). \tag{25}$$

For Eq. (1), when $p = q = 1$ and $\boldsymbol{A}_1 = \boldsymbol{A} \in \mathbb{R}^{n\times m}$, $\boldsymbol{B}_1 = \boldsymbol{C}_1 = \boldsymbol{I}_n$, $\boldsymbol{D}_1 = \boldsymbol{B} \in \mathbb{R}^{m\times n}$, we obtain a special case of the form:

$$\boldsymbol{AX} + \boldsymbol{X}^{\mathrm{T}}\boldsymbol{B} = \boldsymbol{F}, \tag{26}$$

the gradient based iterative (GI) algorithm of (26) is as follows:

$$\boldsymbol{X}(k) = \frac{\boldsymbol{X}_1(k) + \boldsymbol{X}_2(k)}{2}, \tag{27}$$

$$\boldsymbol{X}_1(k) = \boldsymbol{X}(k-1) + \mu\boldsymbol{A}^{\mathrm{T}}[\boldsymbol{F} - \boldsymbol{AX}(k-1) - \boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{B}], \tag{28}$$

$$\boldsymbol{X}_2(k) = \boldsymbol{X}(k-1) + \mu\boldsymbol{B}[\boldsymbol{F} - \boldsymbol{AX}(k-1) - \boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{B}]^{\mathrm{T}}, \tag{29}$$

$$0 < \mu < \mu_0 := \frac{2}{\lambda_{\max}[\boldsymbol{AA}^{\mathrm{T}}] + \lambda_{\max}[\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}]}. \tag{30}$$

Similarly, referring to [3,15], we can obtain the least-squares based iterative (LSI) algorithm of (26) as follows:

$$\boldsymbol{X}_1(k) = \boldsymbol{X}(k-1) + \mu(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}[\boldsymbol{F} - \boldsymbol{AX}(k-1) - \boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{B}], \tag{31}$$

$$\boldsymbol{X}_2(k) = \boldsymbol{X}(k-1) + \mu(\boldsymbol{BB}^{\mathrm{T}})^{-1}\boldsymbol{B}[\boldsymbol{F} - \boldsymbol{AX}(k-1) - \boldsymbol{X}^{\mathrm{T}}(k-1)\boldsymbol{B}]^{\mathrm{T}}, \tag{32}$$

$$\boldsymbol{X}(k) = \frac{\boldsymbol{X}_1(k) + \boldsymbol{X}_2(k)}{2}, \quad 0 < \mu < 4. \tag{33}$$

3. Examples

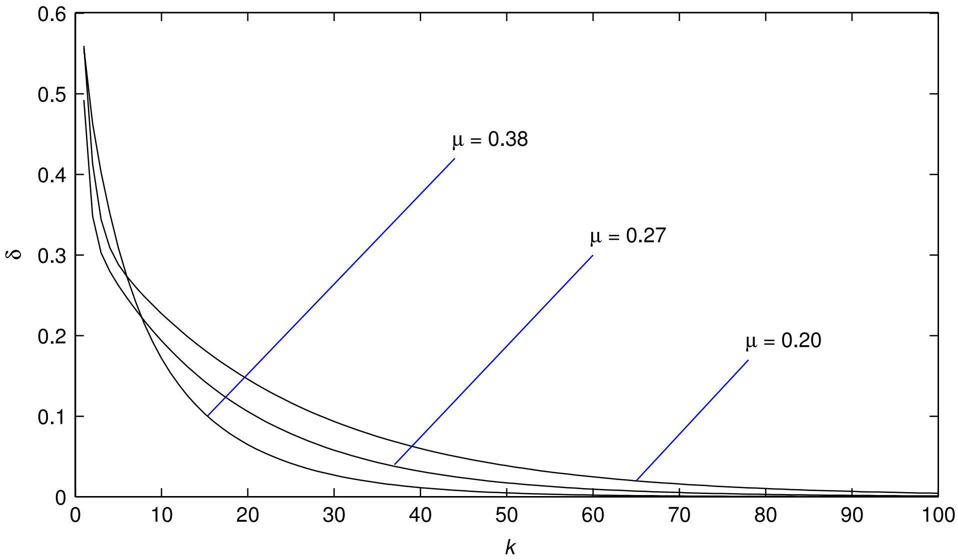
This section gives two examples to illustrate the performances of the proposed algorithms.

Example 1. Suppose that $\boldsymbol{AX} + \boldsymbol{X}^{\mathrm{T}}\boldsymbol{B} = \boldsymbol{F}$, where

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \qquad \boldsymbol{F} = \begin{bmatrix} 8 & 8 \\ 5 & 2 \end{bmatrix}.$$

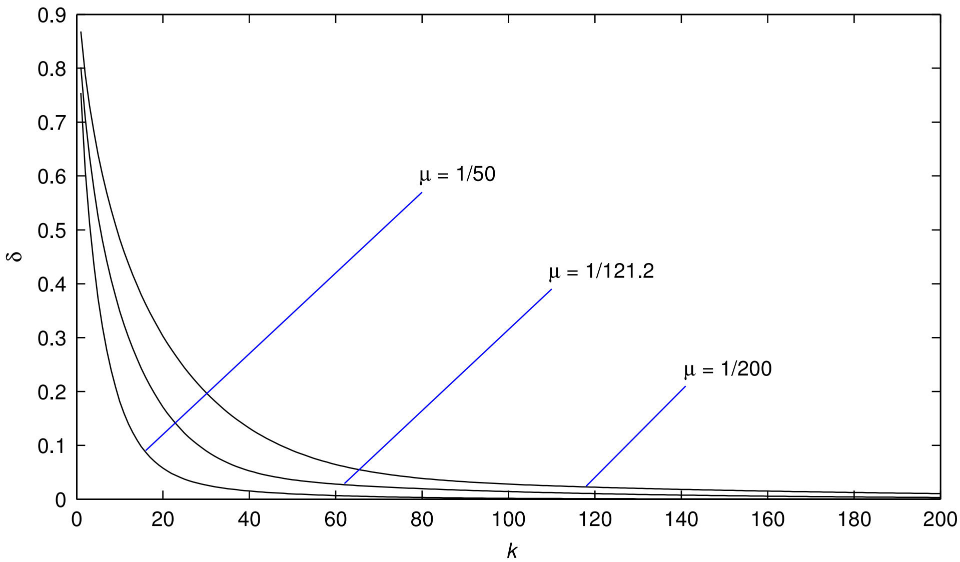
From (2), we can obtain the solution of this matrix equation, which is

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$



$$\mu_0 = 2\{\lambda_{\max}[\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}] + \lambda_{\max}[\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}]\}^{-1} = 0.27$$

Fig. 1. The errors δ versus k of Example 1.



$$\mu_0 = 2\left\{\sum_{j=1}^{2}\lambda_{\max}[\boldsymbol{A}_j\boldsymbol{A}_j^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{B}_j^{\mathrm{T}}\boldsymbol{B}_j] + \sum_{l=1}^{2}\lambda_{\max}[\boldsymbol{C}_l\boldsymbol{C}_l^{\mathrm{T}}]\lambda_{\max}[\boldsymbol{D}_l^{\mathrm{T}}\boldsymbol{D}_l]\right\}^{-1} = \frac{1}{121.2}$$

Fig. 2. The errors δ versus k of Example 2.

Take $\boldsymbol{X}(0) = 10^{-6}\mathbf{1}_{2\times 2}$. Applying the GI algorithm in (27)–(30) to compute $\boldsymbol{X}(k)$, the iterative errors $\delta := \|\boldsymbol{X}(k) - \boldsymbol{X}\|/\|\boldsymbol{X}\|$ versus k are shown in Fig. 1.

From Fig. 1, it is clear that the errors δ become smaller and go to zero as k increases. The effect of changing the convergence factor μ is illustrated in Fig. 1. We can see that for $\mu = 0.20$, 0.27 and 0.38, the larger the convergence factor μ, the faster the convergence rate. However, if we keep enlarging μ, the algorithm will diverge. How to choose a best convergence factor is still a project to be studied.

Example 2. Suppose that $\boldsymbol{A}_1\boldsymbol{X}\boldsymbol{B}_1 + \boldsymbol{A}_2\boldsymbol{X}\boldsymbol{B}_2 + \boldsymbol{C}_1\boldsymbol{X}^{\mathrm{T}}\boldsymbol{D}_1 + \boldsymbol{C}_2\boldsymbol{X}^{\mathrm{T}}\boldsymbol{D}_2 = \boldsymbol{F}$, where

$$\boldsymbol{A}_1 = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix}, \quad \boldsymbol{B}_1 = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}, \quad \boldsymbol{B}_2 = \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix},$$

$$\boldsymbol{C}_1 = \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}, \quad \boldsymbol{C}_2 = \begin{bmatrix} -1 & 3 \\ -1 & 2 \end{bmatrix}, \quad \boldsymbol{D}_1 = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{D}_2 = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \quad \boldsymbol{F} = \begin{bmatrix} 35 & 9 \\ 20 & 7 \end{bmatrix}.$$

From (2), the solution is

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}.$$

Taking $\boldsymbol{X}(0) = 10^{-6}\mathbf{1}_{2\times 2}$, we apply the algorithm in (11)–(14) to compute $\boldsymbol{X}(k)$. The errors $\delta := \|\boldsymbol{X}(k) - \boldsymbol{X}\|/\|\boldsymbol{X}\|$ versus k are shown in Fig. 2.

From Fig. 2, it is clear that the errors δ become smaller and converge to zero as k increases. The effect of changing the convergence factor μ is illustrated in Fig. 2 with $\mu = 1/200$, $1/121.2$ and $1/50$, and a larger μ leads to a faster convergence rate.

4. Conclusions

The gradient based iterative algorithms for solving general matrix equations are studied by using the hierarchical identification principle. We prove that the iterative solutions given by the proposed algorithms converge fast to their true solutions for any initial values and small condition numbers. We test the proposed algorithm using MATLAB and the results verify our theoretical findings. The algorithm is proposed for linear general matrix equations; extending the adopted idea to study iterative solutions for nonlinear matrix equations requires further research.

References

[1] G.H. Golub, C.F. Van Loan, Matrix Computations, 3rd ed., Johns Hopkins Univ. Press, Baltimore, MD, 1996.
[2] 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.
[3] 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.
[4] Z.L. Tian, C.Q. Gu, A numerical algorithm for Lyapunov equations, Applied Mathematics and Computation 202 (1) (2008) 44–53.
[5] A. Kilicman, Z. Al Zhour, Vector least-squares solutions for coupled singular matrix equations, Journal of Computational and Applied Mathematics 206 (2) (2007) 1051–1069.
[6] F. Ding, L. Qiu, T. Chen, Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems, Automatica 45 (2) (2009) 324–332.
[7] F. Ding, T. Chen, Performance analysis of multi-innovation gradient type identification methods, Automatica 43 (1) (2007) 1–14.
[8] F. Ding, P.X. Liu, G. Liu, Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises, Signal Processing 89 (10) (2009) 1883–1890.
[9] 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 (4) (2008) 966–975.
[10] F. Ding, H.Z. Yang, F. Liu, Performance analysis of stochastic gradient algorithms under weak conditions, Science in China Series F–Information Sciences 51 (9) (2008) 1269–1280.
[11] 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.
[12] 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 (2) (2008) 682–685.
[13] B. Zhou, G.R. Duan, Solutions to generalized Sylvester matrix equation by Schur decomposition, International Journal of Systems Science 38 (5) (2007) 369–375.
[14] B. Zhou, G.R. Duan, On the generalized Sylvester mapping and matrix equations, Systems & Control Letters 57 (3) (2008) 200–208.
[15] F. Ding, T. Chen, Iterative least squares solutions of coupled Sylvester matrix equations, Systems & Control Letters 54 (2) (2005) 95–107.
[16] F. Ding, T. Chen, On iterative solutions of general coupled matrix equations, SIAM Journal on Control and Optimization 44 (6) (2005) 2269–2284.
[17] Z. Al Zhour, A. Kilicman, Some new connections between matrix products for partitioned and non-partitioned matrices, Computers & Mathematics with Applications 54 (6) (2007) 763–784.
[18] B. Zhou, G.R. Duan, An explicit solution to the matrix equation $AX - XF = BY$, Linear Algebra and its Applications 402 (1–3) (2005) 345–366.
[19] B. Zhou, G.R. Duan, A new solution to the generalized Sylvester matrix equation $AV - EVF = BW$, Systems & Control Letters 55 (3) (2006) 193–198.
[20] B. Zhou, G.R. Duan, Parametric solutions to the generalized Sylvester matrix equation $AX - XF = BY$ and the regulator equation $AX - XF = BY + R$, Asian Journal of Control 9 (4) (2007) 475–483.
[21] F. Ding, T. Chen, Hierarchical gradient based identification of multivariable discrete-time systems, Automatica 41 (2) (2005) 315–325.
[22] F. Ding, T. Chen, Hierarchical least squares identification methods for multivariable systems, IEEE Transactions on Automatic Control 50 (3) (2005) 397–402.