ISSN: 1017-060X (Print)

ISSN: 1735-8515 (Online)

Bulletin of the Iranian Mathematical Society

Vol. 40 (2014), No. 5, pp. 1097–1117

Title:

Theoretical results on the global GMRES method for solving generalized Sylvester matrix equations

Author(s):

F. Panjeh Ali Beik

Published by Iranian Mathematical Society
http://bims.ims.ir

Bull. Iranian Math. Soc.
Vol. 40 (2014), No. 5, pp. 1097–1117
Online ISSN: 1735-8515

THEORETICAL RESULTS ON THE GLOBAL GMRES METHOD FOR SOLVING GENERALIZED SYLVESTER MATRIX EQUATIONS

F. PANJEH ALI BEIK

(Communicated by Nezam Mahdavi-Amiri)

ABSTRACT. The global generalized minimum residual (Gl-GMRES) method is examined for solving the generalized Sylvester matrix equation

$$\sum_{i=1}^{q} A_i X B_i = C.$$

Some new theoretical results are elaborated for the proposed method by employing the Schur complement. These results can be exploited to establish new convergence properties of the Gl-GMRES method for solving general (coupled) linear matrix equations. In addition, the Gl-GMRES method for solving the generalized Sylvester-transpose matrix equation is briefly studied. Finally, some numerical experiments are presented to illustrate the efficiently of the Gl-GMRES method for solving the general linear matrix equations.
Keywords: Linear matrix equation, Krylov subspace, global GMRES, Schur complement.
MSC(2010): Primary: 65F10; Secondary: 15A24.

1. **Introduction**

Consider the generalized Sylvester matrix equation

$$\sum_{i=1}^{q} A_i X B_i = C, \tag{1.1}$$

Article electronically published on October 27, 2014.
Received: 25 July 2011, Accepted: 6 August 2013.

where the matrices $A_i \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{p\times p}$ $(i = 1, 2, \ldots, q)$ and $C \in \mathbb{R}^{n\times p}$ are given and $X \in \mathbb{R}^{n\times p}$ is an unknown matrix to be determined.

The linear matrix equations arise in the solution of large eigenvalue problems and in the boundary value problems. Moreover, they play a central role in the control and communication theory and image restoration; for further details see [1, 2, 11] and the references therein.

Note that the linear matrix equation (1.1) can be reformulated by the following $np \times np$ linear system:

$$\mathcal{A}vec(X) = vec(C), \tag{1.2}$$

where $\mathcal{A} := \sum_{i=1}^{q} (B_i^T \otimes A_i)$. Evidently, (1.1) has a unique solution if and only if the coefficient matrix $\mathcal{A}$ is nonsingular. Throughout this paper, we assume that this condition is satisfied. It is true that the Krylov subspace methods can be used to solve the linear system (1.2). Nevertheless, even for moderate values of n and p, the size of the coefficient matrix $\mathcal{A}$ may become too large and the Krylov subspace methods consume more computer time and memory once the size of the system is large. To overcome these complications and drawbacks, we first consider the following linear operator $\mathcal{M}$ defined as

$$\begin{aligned} \mathcal{M} : \mathbb{R}^{n\times p} &\to \mathbb{R}^{n\times p}, \\ X &\mapsto \mathcal{M}(X) := \sum_{i=1}^{q} A_i X B_i. \end{aligned}$$

Therefore, we can rewrite Equation (1.1) as follows:

$$\mathcal{M}(X) = C. \tag{1.3}$$

Hence, the Gl-GMRES method [6] can be utilized for solving (1.3) which is equivalent to the matrix equation (1.1).

Notations: For a given matrix $X \in \mathbb{R}^{n\times p}$, the notation $vec(X)$ stands for a vector of dimension np obtained by stacking the columns of the matrix X. For an arbitrary square matrix Z, $\det(Z)$ denotes the determinant of Z, $\mathrm{tr}(Z)$ represents the trace of Z and $\lambda_{\min}(Z)$ ($\lambda_{\max}(Z)$) signifies the smallest (largest) eigenvalue of Z. For two given matrices $X \in \mathbb{R}^{n\times p}$ and $Y \in \mathbb{R}^{q\times l}$, the Kronecker product $X \otimes Y$ is the $nq \times pl$ matrix determined by $X \otimes Y = [X_{i,j}Y]$. For two given matrices $Y, Z \in \mathbb{R}^{n\times p}$, the inner product $\langle Y, Z\rangle_F$ is specified by $\langle Y, Z\rangle_F = tr(Y^T Z)$, the associate norm is the well-known Frobenius norm denoted by $\|.\|_F$. Throughout this paper, a set of matrices in $\mathbb{R}^{n\times p}$ is said to be F-orthonormal if it is orthonormal with respect to the scalar product $\langle .,. \rangle_F$.

This paper is organized as follows. In Section 2, we recollect some useful definitions and theorems and review some properties of the Schur complement and the $\diamond$ and Kronecker products. Section 3 is devoted to presenting the Gl-GMRES method for solving the generalized Sylvester matrix equation (1.1). In Section 4, we establish some new theoretical results for the Frobenius norm of the residual matrix obtained by the Gl-GMRES method. In Section 5, the Gl-GMRES method is examined for solving generalized Sylvester-transpose matrix equation by changing the definition of linear operator $\mathcal{M}$. In Section 6, two numerical examples are presented to demonstrate the applicability of the Gl-GMRES method for solving generalized linear matrix equations. Finally, the paper is ended with a brief conclusion in Section 7.

2. **Preliminaries**

In this section, we recall some theorems and concepts which are utilized in the next sections.

Lemma 2.1. (**The Kantorovich inequality**) *Let B be any symmetric positive definite real matrix and $\lambda_{max}, \lambda_{min}$ be its largest and smallest eigenvalues, respectively. Then,*

(a)

$$\frac{\langle Bx, x\rangle_2 \langle B^{-1}x, x\rangle_2}{\langle x, x\rangle_2^2} \leq \frac{(\lambda_{\max} + \lambda_{\min})^2}{4\lambda_{\max}\lambda_{\min}}, \qquad \forall x \neq 0,$$

where $\langle x, x\rangle_2^2 = x^T x$.

(b)

$$\frac{\langle Bx, x\rangle_2 \langle B^{-1}x, x\rangle_2}{\langle x, x\rangle_2^2} \leq \frac{(1 + \chi(B))^2}{4\chi(B)}, \qquad \forall x \neq 0,$$

where $\chi(Z)$ is the condition number of the matrix Z.

Proof. See [8]. □

2.1. ***The $\diamond$ product.***

Definition 2.2. (R. Bouyouli et al. [3]). *Let $A = [A_1, A_2, \ldots, A_p]$ and $B = [B_1, B_2, \ldots, B_\ell]$ be two given matrices of dimensions $n \times ps$ and*

$n \times \ell s$, respectively, where A_i and B_j are $n \times s$ matrices. Then the $p \times \ell$ matrix $A^T \diamond B = [(A^T \diamond B)_{ij}]$ is defined by

$$(A^T \diamond B)_{ij} = \langle A_i, B_j \rangle_F, \qquad i = 1, 2, \ldots, p, \quad j = 1, 2, \ldots, \ell.$$

Proposition 2.3. Let $A, B, C \in \mathbb{R}^{n \times ps}, D \in \mathbb{R}^{n \times n}, L \in \mathbb{R}^{p \times p}$. Then, we have
(1) $(A + B)^T \diamond C = A^T \diamond C + B^T \diamond C$.
(2) $A^T \diamond (B + C) = A^T \diamond B + A^T \diamond C$.
(3) $(A^T \diamond B)^T = B^T \diamond A$.
(4) $(DA)^T \diamond B = A^T \diamond (D^T B)$.
(5) $A^T \diamond (B(L \otimes I_s)) = (A^T \diamond B)L$.
(6) $\left\| A^T \diamond B \right\|_F \leq \|A\|_F \|B\|_F$.

Proof. See [3]. ☐

2.2. ***Some Schur complement identities.*** In the current subsection, we recall the definition of *Schur* complement and present some of its properties; see [7, 9, 10] for more details.

Definition 2.4. *Let M be a matrix partitioned into four blocks as follows:*

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

where the submatrix D is assumed to be square and nonsingular. The Schur complement of D in M is denoted by (M/D) and defined by:

$$(M/D) = A - BD^{-1}C.$$

Proposition 2.5. *Let $A \in \mathbb{R}^{n \times s}, B \in \mathbb{R}^{n \times ks}, C \in \mathbb{R}^{k \times p}, G \in \mathbb{R}^{k \times k}$ and $E \in \mathbb{R}^{n \times s}$. If the matrix G is nonsingular, then*

$$E^T \diamond \left(\begin{pmatrix} A & B \\ C \otimes I_s & G \otimes I_s \end{pmatrix} \Big/ G \otimes I_s \right) = \begin{pmatrix} E^T \diamond A & E^T \diamond B \\ C & G \end{pmatrix} \Big/ G.$$

Proof. See [3]. ☐

Proposition 2.6. *If the matrices M and D are square and nonsingular, then*

$$M^{-1} = \begin{pmatrix} (M/D)^{-1} & -(M/D)^{-1} B D^{-1} \\ -D^{-1} C (M/D)^{-1} & D^{-1} + D^{-1} C (M/D)^{-1} B D^{-1} \end{pmatrix}.$$

Proof. See [10]. ☐

3. **Gl-GMRES for generalized Sylvester matrix equations**

In this section we employ the Gl-GMRES method for solving the generalized Sylvester matrix equation (1.1). Here, we would like to point out that the Gl-GMRES for solving (1.1) has been experimentally examined in [2] and the main contribution of the current work is to analyze the application of the method form a theoretical point of view. To this end, consider the block Krylov subspace

$$\mathcal{K}_m(\mathcal{M}, R_0) = span\{R_0, \mathcal{M}(R_0), \ldots, \mathcal{M}^{m-1}(R_0)\}, \tag{3.1}$$

where $R_0 = C - \mathcal{M}(X_0)$ and $X_0 \in \mathbb{R}^{n\times p}$ is a given initial guess for the solution of Equation (1.1).

Now, the global Arnoldi process is presented which constructs an F-orthonormal basis for the $\mathcal{K}_m(\mathcal{M}, R_0)$.

Algorithm 1. *Global Arnoldi process.*
1. Set $V_1 = R_0/\left\|R_0\right\|_F$.
2. For $j = 1, 2, \ldots, m$ Do
3. $\quad W := \mathcal{M}(V_j)$
4. $\quad$ For $i = 1, 2, \ldots, j$ Do
5. $\quad\quad h_{ij} := \langle W, V_i\rangle_F$
6. $\quad\quad W := W - h_{ij}V_i$
7. $\quad$ End for
8. $\quad h_{j+1,j} := \left\|W\right\|_F$. If $h_{j+1,j} := 0$, then stop.
9. $\quad V_{j+1} := W/h_{j+1,j}$
10. End for

Let $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$ and let $\bar{H}_m = [h_{ij}]_{(m+1)\times m}$ be the Hessenberg matrix whose nonzero entries are computed by Algorithm 1. Suppose that H_m is the $m \times m$ matrix obtained from $\bar{H}_m$ by deleting its last row. It is not difficult to verify that the global Arnoldi process produces an F-orthonormal basis $V_1, V_2, \ldots, V_m$ for the block Krylov subspace $\mathcal{K}_m$, i.e., for $i, j = 1, 2, \ldots, m$, we have

$$\begin{cases} tr(V_i^T V_j) = 0, & (i \neq j) \\ tr(V_i^T V_i) = 1. \end{cases}$$

Or equivalently, $\mathcal{V}_m^T \diamond \mathcal{V}_m = I_m$.

Theorem 3.1. *Let $\mathcal{V}_m$, H_m, and $\bar{H}_m$ be defined as before. Then the following relations hold*

$$[\mathcal{M}(V_1), \mathcal{M}(V_2), \ldots, \mathcal{M}(V_m)] = \sum_{i=1}^{q} A_i \mathcal{V}_m (I_m \otimes B_i)$$

$$= \mathcal{V}_m (H_m \otimes I_p) + h_{m+1,m} V_{m+1} (e_m^T \otimes I_p), \tag{3.2}$$

$$[\mathcal{M}(V_1), \mathcal{M}(V_2), \ldots, \mathcal{M}(V_m)] = \sum_{i=1}^{q} A_i \mathcal{V}_m (I_m \otimes B_i)$$

$$= \mathcal{V}_{m+1} (\bar{H}_m \otimes I_p). \tag{3.3}$$

Proof. The relation (3.2) follows from the fact that

$$\sum_{i=1}^{q} A_i \mathcal{V}_m (I_m \otimes B_i) = \left[\sum_{i=1}^{q} A_i V_1 B_i, \sum_{i=1}^{q} A_i V_2 B_i, \ldots, \sum_{i=1}^{q} A_i V_m B_i \right],$$

and the following equality obtained from Lines 3-6 and 9 of Algorithm 1,

$$\sum_{i=1}^{q} A_i V_j B_i = \sum_{k=1}^{j+1} h_{kj} V_k, \qquad j = 1, 2, \ldots, m.$$

It is not difficult to verify that the relation (3.3) is a reformulation of (3.2). □

For a given initial guess $X_0 \in \mathbb{R}^{n \times p}$, with the corresponding residual $R_0 = C - \mathcal{M}(X_0)$, the Gl-GMRES computes the new approximate solution X_m such that

$$X_m \in X_0 + \mathcal{K}_m(\mathcal{M}, R_0), \tag{3.4}$$

and

$$R_m \perp_F \mathcal{K}_m(\mathcal{M}, \mathcal{M}(R_0)), \tag{3.5}$$

where $R_m = C - \mathcal{M}(X_m)$.
Let $\mathcal{V}_m$ be the F-orthonormal basis for $\mathcal{K}_m(\mathcal{M}, R_0)$ constructed by the global Arnoldi process. From (3.4), it can be verified that

$$X_m = X_0 + \mathcal{V}_m (y_m \otimes I_p), \tag{3.6}$$

where the vector $y_m \in \mathbb{R}^m$ is obtained by imposing the orthogonality condition (3.5). On the other hand, by some straightforward computations, it can be shown that y_m is the solution of the following least-squares problem too

$$\min_{y\in\mathbb{R}^m} \left\| R_0 - [\sum_{i=1}^{q} A_i \mathcal{V}_m (y \otimes I_p) B_i] \right\|_F$$

$$= \min_{y\in\mathbb{R}^m} \left\| R_0 - [\sum_{i=1}^{q} A_i \mathcal{V}_m (I_m \otimes B_i)](y \otimes I_p) \right\|_F . \tag{3.7}$$

Note that in the relation (3.7), we have utilized the following relation

$$(y \otimes I_p) Z = (I_m \otimes Z)(y \otimes I_p).$$

where Z is an arbitrary $p \times p$ real matrix, $y \in \mathbb{R}^m$ and m is a given integer. On the other hand, from (3.3), we have

$$\left\| R_0 - [\sum_{i=1}^{q} A_i \mathcal{V}_m (I_m \otimes B_i)](y \otimes I_p) \right\|_F = \left\| \mathcal{V}_{m+1}((\beta e_1 - \bar{H}_m y) \otimes I_p) \right\|_F .$$

Therefore, we conclude that y_m is the solution of the following least-squares problem

$$\min_{y\in\mathbb{R}^m} \left\| \beta e_1 - \bar{H}_m y \right\|_2 . \tag{3.8}$$

Like the GMRES algorithm [8], in application, the Gl-GMRES algorithm is restarted every m inner iterations, where m is a given fixed integer and the corresponding algorithm is denoted by Gl-GMRES (m) and presented as follows:

Algorithm 2. *Gl-GMRES* (m) *algorithm for* (1.1).
1. Choose X_0, and a tolerance ϵ. Compute $R_0 = C - \mathcal{M}(X_0)$ and $V_1 = R_0$.
2. Construct the orthonormal basis $V_1, V_2, \ldots, V_m$ by Algorithm 1.
3. Determine y_m as the solution of the least-squares problem

$$\min_{y\in\mathbb{R}^m} \left\| \beta e_1 - \bar{H}_m y \right\|_2 .$$

4. Calculate $X_m = X_0 + \mathcal{V}_m(y_m \otimes I_p)$.
5. Compute the residual R_m and $\|R_m\|_F$.
6. If $\frac{\|R_m\|_F}{\|R_0\|_F} < \epsilon$ Stop; else $R_0 := R_m$, $V_1 := R_0$, Go to 2.

The weighted versions of the Krylov subspace methods for solving the matrix equation $AX = B$ are studied in [5]. To improve the speed of convergence of the Gl-GMRES method to solve general (coupled) linear matrix equations, the application of weighted versions of the Gl-GMRES method can also be a subject of interest to be discused from both theoretical and experimental points of view.

3.1. **Complexity considerations.** In the following, a rough estimation of the operation requirements for each restart of Algorithm 2 is determined. In the second line of the algorithm, the global Arnoldi process (Algorithm 1) is performed. Accordingly, we first compute the complexity of Algorithm 1. Suppose that $N_z(A_i)$ and $N_z(B_i)$ are the number of nonzero elements of A_i and B_i for $i = 1, 2, \ldots, q$ respectively.

In Line 3 of Algorithm 1, we must establish by calculation $\mathcal{M}(V_j)$ where $V_j \in \mathbb{R}^{n\times p}$ and

$$\mathcal{M}(V_j) = \sum_{i=1}^{q} A_i V_j B_i, \qquad j = 1, 2, \ldots\ m.$$

For calculating $S_i = A_i V_j B_i$, in practice, we first figure $S_i = A_i V_j$ and then $S_i = S_i B_i$. That is, p matrix-vector and n vector-matrix products are computed to obtain $S_i = A_i V_j B_i$. Hence, the total operations over m steps can be approximated by

$$2mp \sum_{i=1}^{q} N_z(A_i) + 2mn \sum_{i=1}^{q} N_z(B_i).$$

Each one of the Frobenius scalar products requires $2np$ operations. Thence, each of the Gram-Schmidt steps (Lines 5 and 6 of Algorithm 1) needs approximately $4 \times j \times np$ operations which offer m steps to roughly $2m^2np$ operations.

To compute the rest of the required operations in Algorithm 2, we point out that the costs of Lines 4, 5 and 6 are trifling relative to the cost of Line 3. In Line 3 of Algorithm 2, we must solve a least-square problem of size $(m + 1) \times m$ at each restart. To this end, the QR decomposition of $\bar{H}_m$ based on Givens rotations is exploited (see pages 162 and 163 of [8]). It is not difficult to verify that the number of required operations for computing y_m in Line 3 of Algorithm 2 can be estimated by $m^2 + 3(m - 1)(m - 2) + 18m$. Therefore the total costs of

Algorithm 2, at each restart, is roughly

$$2mp\sum_{i=1}^{q} N_z(A_i) + 2mn\sum_{i=1}^{q} N_z(B_i) + 2m^2np + 4m^2.$$

We assume that A_i and B_i $(i = 1, 2, \ldots, q)$ are sparse matrices. Consequently, at each restart, the total cost of Algorithm 2 may be approximated by $\mathcal{O}(m^2np)$.

4. New theoretical results on the Gl-GMRES

Recently, Bouyouli et al. [3, 4] have established some new theoretical results for the global minimal residual (Gl-MR) method to solving the multiple linear system $AX = B$. However different versions of these results can be derived by using the orthonormal basis of the block Krylov subspace constructed by the global Arnoldi process.

Considering the orthonormal basis $\mathcal{V}_m$ of $\mathcal{K}_m(\mathcal{M}, R_0)$. We develop some new theoretical results for the Gl-GMRES method for solving (1.1).

For simplicity, we signify $\mathcal{W}_m$ as follows:

$$\mathcal{W}_m := [\mathcal{M}(V_1), \mathcal{M}(V_2), \ldots, \mathcal{M}(V_m)] = \sum_{i=1}^{q} A_i\mathcal{V}_m(I_m \otimes B_i). \tag{4.1}$$

The orthogonality condition (3.5) implies that

$$\begin{aligned} 0 &= \mathcal{W}_m^T \diamond R_m \\ &= \mathcal{W}_m^T \diamond [R_0 - \sum_{i=1}^{q} A_i\mathcal{V}_m(I_m \otimes B_i)(y \otimes I_p)] \\ &= \mathcal{W}_m^T \diamond R_0 - \mathcal{W}_m^T \diamond (\mathcal{W}_m(y \otimes I_p)). \end{aligned}$$

Consequently, y_m is the solution of the following linear system

$$(\mathcal{W}_m^T \diamond \mathcal{W}_m)y_m = \mathcal{W}_m^T \diamond R_0. \tag{4.2}$$

From Equation (4.2), we infer that y_m exists if and only if $\mathcal{W}_m^T \diamond \mathcal{W}_m$ is a nonsingular matrix. In this paper, we assume that this condition is satisfied. Straightforward computations demonstrate that

$$\begin{aligned} R_m &= R_0 - \mathcal{W}_m[(\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1}(\mathcal{W}_m^T \diamond R_0) \otimes I_p] \\ &= R_0 - \mathcal{W}_m[(\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1} \otimes I_p][(\mathcal{W}_m^T \diamond R_0) \otimes I_p]. \end{aligned}$$

From the definition of the Schur complement, we derive:

(4.3)
$$R_m = \left(\begin{bmatrix} R_0 & \mathcal{W}_m \\ (\mathcal{W}_m^T \diamond R_0) \otimes I_p & (\mathcal{W}_m^T \diamond \mathcal{W}_m) \otimes I_p \end{bmatrix} \Big/ (\mathcal{W}_m^T \diamond \mathcal{W}_m) \otimes I_p \right).$$

Theorem 4.1. *Assume that $\mathcal{W}_m^T \diamond \mathcal{W}_m$ is nonsingular. The residual matrix R_m, obtained by the Gl-GMRES algorithm at step m, satisfies the following relation*

$$\|R_m\|_F^2 = \frac{det[\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1}]}{det[\mathcal{W}_m^T \diamond \mathcal{W}_m]}, \tag{4.4}$$

where $\bar{\mathcal{V}}_{m+1} = [R_0, \mathcal{W}_m]$.

Proof. From the orthogonality condition (3.5), we get

$$R_m^T \diamond R_m = R_0^T \diamond R_m.$$

By using Proposition 2.5 and Equation (4.3), we have

$$\begin{aligned} R_0^T \diamond R_m &= \left(\begin{bmatrix} R_0^T \diamond R_0 & R_0^T \diamond \mathcal{W}_m \\ \mathcal{W}_m^T \diamond R_0 & \mathcal{W}_m^T \diamond \mathcal{W}_m \end{bmatrix} \Big/ \mathcal{W}_m^T \diamond \mathcal{W}_m \right) \\ &= (\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1} / \mathcal{W}_m^T \diamond \mathcal{W}_m). \end{aligned}$$

Or equivalently,

$$R_m^T \diamond R_m = (\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1} / \mathcal{W}_m^T \diamond \mathcal{W}_m). \tag{4.5}$$

(Note that $\|R_m\|_F^2 = R_m^T \diamond R_m$ is a scalar.) $\square$

Theorem 4.2. *At step m, assume that R_m denotes the residual produced by the Gl-GMRES method. Then we have*

$$\|R_m\|_F^2 = \frac{\det\left(\begin{bmatrix} \beta^2 & \beta e_1^T H_m \\ \beta H_m^T e_1 & \bar{H}_m^T \bar{H}_m \end{bmatrix} \right)}{\det[\bar{H}_m^T \bar{H}_m]}, \tag{4.6}$$

where $\beta = \|R_0\|_F$.

Proof. Invoking Equation (3.3), it can be derived that

$$\mathcal{W}_m^T \diamond \mathcal{W}_m = (\mathcal{V}_{m+1}(\bar{H}_m \otimes I_p))^T \diamond (\mathcal{V}_{m+1}(\bar{H}_m \otimes I_p)).$$

Since $\mathcal{V}_{m+1}^T \diamond \mathcal{V}_{m+1} = I$, we deduce that

$$\mathcal{W}_m^T \diamond \mathcal{W}_m = \bar{H}_m^T (\mathcal{V}_{m+1}^T \diamond \mathcal{V}_{m+1}) \bar{H}_m = \bar{H}_m^T \bar{H}_m. \tag{4.7}$$

Using Equation (3.2), we may verify that

$$
\begin{aligned}
R_0^T \diamond \mathcal{W}_m &= R_0^T \diamond [\mathcal{V}_m(H_m \otimes I_p) + h_{m+1,m} V_{m+1}(e_m^T \otimes I_p)] \\
&= (R_0^T \diamond \mathcal{V}_m) H_m + h_{m+1,m}(R_0^T \diamond V_{m+1}) e_m^T. \qquad (4.8)
\end{aligned}
$$

It is known that $R_0^T = \beta V_1$ and $V_1^T \diamond V_i = 0$ for $i \neq 1$. Hence, we can rewrite (4.8) as follows:

$$
R_0^T \diamond \mathcal{W}_m = (R_0^T \diamond \mathcal{V}_m) H_m = \beta e_1^T H_m. \qquad (4.9)
$$

On the other hand,

$$
\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1} = \begin{bmatrix} R_0^T \diamond R_0 & R_0^T \diamond \mathcal{W}_m \\ \mathcal{W}_m^T \diamond R_0 & \mathcal{W}_m^T \diamond \mathcal{W}_m \end{bmatrix}. \qquad (4.10)
$$

By substituting (4.7) and (4.9) in the above relation, the result follows from Theorem 4.1 immediately. □

Theorem 4.3. *Let $\bar{\mathcal{V}}_{m+1} = [R_0, \mathcal{W}_m]$. Moreover, suppose that $\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1}$ and $\mathcal{W}_m^T \diamond \mathcal{W}_m$ are nonsingular matrices. Then, the residual R_m satisfies the following relation*

$$
\|R_m\|_F^2 = \frac{1}{e_1^T (\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1})^{-1} e_1}. \qquad (4.11)
$$

Proof. By the assumption $\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1}$ and $\mathcal{W}_m^T \diamond \mathcal{W}_m$ are nonsingular matrices, therefore the Schur complement $(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1})/(\mathcal{W}_m^T \diamond \mathcal{W}_m)$ is nonzero. Using Proposition 2.6, we achieve to the following relation

$$
e_1^T (\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1})^{-1} e_1 = \frac{1}{(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1} / \mathcal{W}_m^T \diamond \mathcal{W}_m)}.
$$

Now, the result follows immediately from (4.5).

□

Proposition 4.4. *Assume that $Z = [Z_1, Z_2, \ldots, Z_k]$ and $Z_i \in \mathbb{R}^{n \times p}$ for $i = 1, 2, \ldots, k$. Furthermore, suppose that $Z^T \diamond Z$ is a nonsingular matrix. Then $Z^T \diamond Z$ is a symmetric positive definite matrix.*

Proof. Evidently $Z^T \diamond Z$ is a symmetric matrix. Assume that $y \in \mathbb{R}^k$ is an arbitrary nonzero vector. Using Proposition 2.3, we have

$$\begin{aligned} y^T(Z^T \diamond Z)y &= y^T(Z^T \diamond (Z(y \otimes I_p))) \\ &= (Z(y \otimes I_p))^T \diamond (Z(y \otimes I_p)) \\ &= \|Z(y \otimes I_p)\|_F^2 \geq 0. \end{aligned}$$

Now, it is sufficient to show that

$$\|Z(y \otimes I_p)\|_F^2 > 0.$$

To this end, suppose that $\|Z(y \otimes I_p)\|_F^2 = 0$ which implies that $Z(y \otimes I_p) = 0$. Thus, $Z^T \diamond (Z(y \otimes I_p)) = 0$. This is equivalent to say that $(Z^T \diamond Z)y = 0$, which is a contradiction. □

Theorem 4.5. *Under the same assumptions as in Theorem 4.3, the residual R_m satisfies the following relation*

$$\frac{4\chi(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1})}{(1+\chi(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1}))^2} \leq \frac{\|R_m\|_F^2}{\|R_0\|_F^2} \leq 1,$$

where $\chi(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1})$ is the condition number of the matrix $\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1}$.

Proof. From Proposition 4.4, we may conclude that $(\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1}$ is a symmetric positive definite matrix. Hence, the relations (4.5) and (4.10) imply that

$$\begin{aligned} \|R_m\|_F^2 &= (\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1} / \mathcal{W}_m^T \diamond \mathcal{W}_m) \\ &= R_0^T \diamond R_0 - [R_0^T \diamond \mathcal{W}_m](\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1}[R_0^T \diamond \mathcal{W}_m]^T \leq R_0^T \diamond R_0. \end{aligned} \tag{4.12}$$

By Theorem 4.3, Lemma 2.1 and the fact that

$$R_0^T \diamond R_0 = e_1^T(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1})e_1,$$

we may conclude that

$$R_0^T \diamond R_0 \geq \frac{1}{e_1^T(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1})^{-1}e_1} \geq \frac{4\chi(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1})}{(1+\chi(\bar{\mathcal{V}}_{m+1}^T \diamond \bar{\mathcal{V}}_{m+1}))^2} R_0^T \diamond R_0.$$

□

Remark 4.6. *The above theorem shows that the Gl-GMRES is not convergent as long as the matrix $\bar{\mathcal{V}}_m^T \diamond \bar{\mathcal{V}}_m$ is well conditioned.*

Remark 4.7. *From (4.12), it can be easily verified that where $R_0^T \diamond \mathcal{W}_m \neq 0$, at each restart, the norm of the residual matrix decreases strictly. On the other hand, from (4.9), it can be concluded that if $R_0^T \diamond \mathcal{W}_m = 0$ then the matrix H_m is singular. Therefore, it seems that if at each iteration the matrix H_m is a nonsingular matrix the well conditioning of $\bar{\mathcal{V}}_m^T \diamond \bar{\mathcal{V}}_m$ does not cause losing of the convergence.*

Here, we would like to comment that by developing the approach elaborated in Chapter 5 of [8], an upper bound for the norm of the residual matrix can be established. To this end, considering the relation (4.12), it is can be observed that

$$\|R_m\|_F^2 = R_0^T \diamond R_0 - [R_0^T \diamond \mathcal{W}_m](\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1}[R_0^T \diamond \mathcal{W}_m]^T.$$

By (4.9), it is known that $R_0^T \diamond \mathcal{W}_m = \|R_0\|_F\, e^T H$. Assume that $\|R_0\|_F \neq 0$, H_m is nonsingular and $H_{1,m}$ stands for its first row. Under the same assumptions as in Theorem 4.3, straightforward computations show that

$$\begin{aligned}\|R_m\|_F^2 &= \|R_0\|_F^2 \left[1 - \left\langle (\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1} H_{1,m}, H_{1,m}\right\rangle_2\right] \\ &= \|R_0\|_F^2 \left[1 - \frac{\left\langle (\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1} H_{1,m}, H_{1,m}\right\rangle_2}{\left\langle H_{1,m}, H_{1,m}\right\rangle_2}\left\langle H_{1,m}, H_{1,m}\right\rangle_2\right].\end{aligned}$$

Note that Proposition 4.4 implies that $(\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1}$ is a symmetric positive definite matrix, consequently

$$\frac{\left\langle (\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1} H_{1,m}, H_{1,m}\right\rangle_2}{\left\langle H_{1,m}, H_{1,m}\right\rangle_2} \geq \lambda_{\min}\left((\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1}\right).$$

Hence, it can be concluded that

$$\|R_m\|_F^2 \leq \|R_0\|_F^2 \left\{1 - \lambda_{\min}\left((\mathcal{W}_m^T \diamond \mathcal{W}_m)^{-1}\right) \left\|H_{1,m}\right\|_2^2\right\},$$

which is equivalent to say that

$$\|R_m\|_F \leq \|R_0\|_F \left\{1 - \lambda_{\max}\left(\mathcal{W}_m^T \diamond \mathcal{W}_m\right) \left\|H_{1,m}\right\|_2^2\right\}^{\frac{1}{2}}.$$

Let m be a given positive integer. In the rest of this section, suppose that $H_m^{(s)}$ and $\mathcal{V}_m^{(s)} = [V_1^{(s)}, V_2^{(s)}, \ldots, V_m^{(s)}]$ are computed by the global Arnoldi process at sth restart. Corresponding to $\mathcal{V}_m^{(s)}$, at sth restart, we consider $\mathcal{W}_m^{(s)} = [W_1^{(s)}, W_2^{(s)}, \ldots, W_m^{(s)}]$ in which $W_j^{(s)} = \mathcal{M}(V_j^{(s)})$ for $j = 1, 2, \ldots, m$. In what follows, at each restart, we assume that

$(\mathcal{W}_m^{(s)})^T \diamond \mathcal{W}_m^{(s)}$ is a nonsingular matrix. Moreover, suppose that $R_m^{(s)}$ and $R_0^{(s)}$ are the corresponding residual matrices computed in Algorithm 2. Note that $R_0^{(s)} = R_m^{(s-1)}$. From Theorem 4.5, it is known that the sequence $\|R_m^{(s)}\|_F$ should generally converge but not necessarily toward 0. Consider the case that $\|R_m^{(s)}\|_F$ converges to a nonzero constant. It means that for (eventually) large s, the norm of $\|R_m^{(s)}\|_F$ becomes a nonzero constant. In this case, from (4.12), we get

$$(R_0^{(s)})^T \diamond \mathcal{W}_m^{(s)} = 0.$$

Now, (4.9) implies that $H_m^{(s)}$ is a singular matrix.

In the following theorem, we present sufficient conditions under which the sequence $\|R_m^{(s)}\|_F$ converges to zero as $s \to \infty$, i.e., conditions under which Algorithm 2 converges to the exact solution of (1.1).

Theorem 4.8. *Let* $\mathcal{V}_m^{(s)} = [V_1^{(s)}, V_2^{(s)}, \ldots, V_m^{(s)}]$ *and* $\mathcal{W}_m^{(s)} = [W_1^{(s)}, W_2^{(s)}, \ldots, W_m^{(s)}]$ *be defined as before. Assume that at each restart of Algorithm 2,* $(\mathcal{W}_m^{(s)})^T \diamond \mathcal{W}_m^{(s)}$ *is a nonsingular matrix* $(s = 1, 2, \ldots)$. *If the linear operator* $\mathcal{M}$ *satisfies the following condition*

$$\langle Z, \mathcal{M}(Z) \rangle_F \neq 0, \qquad \forall Z \in \mathbb{R}^{n\times p} \quad (Z \neq 0), \tag{4.13}$$

then,

$$\|R_m^{(s)}\|_F \to 0 \qquad as \quad s \to \infty,$$

where $R_m^{(s)} = C - \mathcal{M}(X_m^{(s)})$ *and* $X_m^{(s)}$ *is the approximate solution to* (1.1) *computed by Algorithm 2* (Gl-GMRES (m)).

Proof. From Theorem 4.5, it can be found that $\|R_m^{(s)}\|_F$ is a convergent sequence. Let $\|R_m^{(s)}\|_F \to \ell$ as $s \to \infty$. It is known that $(\mathcal{W}_m^{(s)})^T \diamond \mathcal{W}_m^{(s)}$ is a symmetric positive definite matrix. Assume that $\ell \neq 0$. Hence, we conclude that $R_0^{(s)} \neq 0$. Therefore, using (4.12) and the fact that $V_1^{(s)} = R_0^{(s)}/\beta^{(s)}$, we find that $\beta^{(s)}((V_1^{(s)})^T \diamond \mathcal{W}_m^{(s)}) \to 0$ as $s \to \infty$ where $\beta^{(s)} = \|R_0^{(s)}\|_F$. That is, $(V_1^{(s)})^T \diamond \mathcal{W}_m^{(s)} = 0$ for (eventually) large s. Hence, we have $\left\langle V_1^{(s)}, \mathcal{M}(V_1^{(s)}) \right\rangle_F = 0$ whereas $V_1^{(s)} \neq 0$ which is on the contrary with (4.13). $\square$

Remark 4.9. *Recently, Beik and Salkuyeh* [1] *have introduced the* $\circledcirc$ *product to illustrate how the Gl-FOM and Gl-GMRES methods can be handled for solving the coupled linear matrix equations. The theoretical* *results, presented in this section, can be extended for the Gl-GMRES method for solving the mentioned coupled linear matrix equations in* [1]. *To this end, it is sufficient to replace the $\diamond$ product by the $\circledcirc$ product.*

5. Gl-GMRES for more general linear matrix equation

Recently, Xie et al. [11] have considered the following general linear matrix equation

$$\sum_{i=1}^{p} A_i X B_i + \sum_{i=1}^{q} C_i X^T D_i = F, \tag{5.1}$$

where $A_i \in \mathbb{R}^{r\times m}, B_i \in \mathbb{R}^{n\times s}, C_i \in \mathbb{R}^{r\times n}, D_i \in \mathbb{R}^{m\times s}$ and $F \in \mathbb{R}^{r\times s}$ are given constant matrices, $X \in \mathbb{R}^{m\times n}$ is the unknown matrix to be determined. In the case that (5.1) has a unique solution, a gradient based iterative algorithm with its convergence analysis have been presented for solving (5.1). Given an arbitrary initial approximate solution $X^{(0)}$, the proposed method computes the sequence of approximate solutions $\left\{X^{(k)}\right\}_{k=1}^{\infty}$ of (5.1) by the following recursive formulas

$$\begin{aligned}
X^{(k)} &= \tfrac{1}{p+q}\left[\sum_{j=1}^{p} X_j^{(k)} + \sum_{l=1}^{q} X_{p+l}^{(k)}\right],\\
X_j^{(k)} &= X^{(k-1)} + \mu A_j^T\left[F - \sum_{i=1}^{p} A_i X^{(k-1)} B_i - \sum_{l=1}^{q} C_i (X^{(k-1)})^T D_i\right] B_j^T,\\
X_{p+l}^{(k)} &= X^{(k-1)} + \mu D_l\left[F - \sum_{i=1}^{p} A_i X^{(k-1)} B_i - \sum_{l=1}^{q} C_i (X^{(k-1)})^T D_i\right]^T C_l.
\end{aligned}$$

A conservative choice of the convergence factor μ is $0 < \mu < \mu_0$ where

$$\mu_0 = 2\left\{\sum_{j=1}^{p} \lambda_{\max}\left[A_j A_j^T\right]\lambda_{\max}\left[B_j^T B_j\right] + \sum_{l=1}^{q} \lambda_{\max}\left[C_l C_l^T\right]\lambda_{\max}\left[D_l^T D_l\right]\right\}^{-1}. \tag{5.2}$$

As seen, the presented method is not suitable for large matrices (for further details see [11]).

Consider a special case where the matrices A_i, B_i, C_i and D_i are squares matrices of the same order. In this section, we show that the Gl-GMRES method can be applied for solving (5.1) with a minor change in Algorithms 1 and 2. In Section 6, we examine the numerical example presented in [11]. The reported results reveal that even for small size matrices the Gl-GMRES method surpasses the gradient based iterative in terms of both number of iterations and CPU-times(s). In what follows, we presume that in (5.1) the coefficient matrices A_i, B_i, C_i and D_i are $n \times n$ real matrices.

With a similar to the manner exploited in Section 1, we may rewrite (5.1) by $\hat{\mathcal{M}}(X) = F$ where the linear operator $\hat{\mathcal{M}}$ is specified as

$$\begin{aligned}\hat{\mathcal{M}} : \mathbb{R}^{n\times n} &\to \mathbb{R}^{n\times n},\\ X &\mapsto \hat{\mathcal{M}}(X) = \sum_{i=1}^{p} A_i X B_i + \sum_{i=1}^{q} C_i X^T D_i.\end{aligned}$$

Using the linear operator $\hat{\mathcal{M}}$ instead of $\mathcal{M}$ in the Line 3 of Algorithm 1, we may construct an F-orthonormal basis for $\mathcal{K}_m(\hat{\mathcal{M}}, R_0) = span\{R_0, \hat{\mathcal{M}}(R_0), \ldots, \hat{\mathcal{M}}^{m-1}(R_0)\}$.

In the rest of this section, we suppose that the linear operator $\hat{\mathcal{M}}$ is utilized in Line 3 of Algorithm 1 and m steps of the algorithm have been performed. We set $\mathcal{V}_m = [V_1, V_2, \ldots, V_m]$ and $\hat{\mathcal{V}}_m = [V_1^T, V_2^T, \ldots, V_m^T]$ where the $n \times n$ matrices V_j for $j = 1, 2, \ldots, m$ are computed by Algorithm 1 in which the linear operator $\hat{\mathcal{M}}$ is used instead of $\mathcal{M}$. It is not difficult to establish the following theorem.

Theorem 5.1. *Assume that $\mathcal{V}_m, \hat{\mathcal{V}}_m$ H_m, and $\bar{H}_m$ have the same structures defined as before and obtained from Algorithm 1 where the linear operator $\hat{\mathcal{M}}$ is used in Line 3. Then the following relations hold*

$$\begin{aligned}[\hat{\mathcal{M}}(V_1), \hat{\mathcal{M}}(V_2), \ldots, \hat{\mathcal{M}}(V_m)] &= \sum_{i=1}^{p} A_i \mathcal{V}_m (I_m \otimes B_i) + \sum_{i=1}^{q} C_i \hat{\mathcal{V}}_m (I_m \otimes D_i)\\ &= \mathcal{V}_m (H_m \otimes I_p) + h_{m+1,m} V_{m+1} (e_m^T \otimes I_p), \qquad (5.3)\\ [\hat{\mathcal{M}}(V_1), \hat{\mathcal{M}}(V_2), \ldots, \hat{\mathcal{M}}(V_m)] &= \sum_{i=1}^{p} A_i \mathcal{V}_m (I_m \otimes B_i) + \sum_{i=1}^{q} C_i \hat{\mathcal{V}}_m (I_m \otimes D_i)\\ &= \mathcal{V}_{m+1} (\bar{H}_m \otimes I_p). \qquad (5.4)\end{aligned}$$

Proof. We may prove relations 5.3 and 5.4 with the same strategy employed in the proof of Theorem 3.1. □

Assume that m is a given fixed integer. With Theorem 5.1 and similar discussions in Section 3, it is natural to present the Gl-GMRES(m) for solving (5.1) as follows.

Algorithm 3. *Gl-GMRES* (m) *for solving* (5.1).
1. Choose X_0, and a tolerance ϵ. Compute $R_0 = F - \hat{\mathcal{M}}(X_0)$, and $V_1 = R_0$.
2. Construct the orthonormal basis $V_1, V_2, \ldots, V_m$ by Algorithm 1 (the linear operator $\hat{\mathcal{M}}$ is used instead of $\mathcal{M}$ in the line 3 of Algorithm 1).
3. Determine y_m as the solution of the least-squares problem:

$$\min_{y \in \mathbb{R}^m} \left\| \beta e_1 - \bar{H}_m y \right\|_2 .$$

4. Calculate $X_m = X_0 + \mathcal{V}_m(y_m \otimes I_p)$.
5. Compute the residual $R_m = F - \hat{\mathcal{M}}(X_m)$ and $\|R_m\|_F$.
6. If $\frac{\|R_m\|_F}{\|R_0\|_F} < \epsilon$ Stop; else $R_0 := R_m$, $V_1 := R_0$, Go to 2.

Remark 5.2. *It is not difficult to verify that all of the theoretical results proved in Section 4 stay valid when the Gl-GMRES method is applied to solve* (5.1). *To this end, we only need to change the definition of* $\mathcal{W}_m$ *in Equation* (4.1) *to*

$$\begin{aligned} \mathcal{W}_m &= [\hat{\mathcal{M}}(V_1), \hat{\mathcal{M}}(V_2), \ldots, \hat{\mathcal{M}}(V_m)] \\ &= \sum_{i=1}^{p} A_i \mathcal{V}_m (I_m \otimes B_i) + \sum_{i=1}^{q} C_i \hat{\mathcal{V}}_m (I_m \otimes D_i). \end{aligned}$$

6. **Numerical experiments**

In this section, two numerical examples are presented to illustrate the effectiveness of the Gl-GMRES method for solving (1.1) and (5.1). All the numerical procedures are performed in MATHEMATICA 6.

In the first example, the test was stopped as soon as

$$\text{Err} := \frac{\|R_m\|_F}{\|R_0\|_F} < 10^{-8},$$

and the initial guess was taken to be zero.

Example 6.1. Assume that the $p \times p$ matrix $T_{d,p}$ is defined as

$$T_{d,p} = tridiag\left(-1 + \frac{10}{p+1}, d, -1 + \frac{10}{p+1}\right).$$

TABLE 1. Numerical results for Example 6.1.

n	iters	Err
250	57	0.969446×10^{-8}
500	58	0.807621×10^{-8}
750	57	0.622639×10^{-8}
1000	57	0.801407×10^{-8}

Consider the linear matrix equation $A_1XB_1 + A_2XB_2 = C$ where

$$A_1 = \begin{pmatrix} 4 & -1 & & -1 \\ -1 & 4 & \ddots & \\ & \ddots & \ddots & -1 \\ -1 & & -1 & 4 \end{pmatrix}_{n\times n}, \qquad A_2 = \begin{pmatrix} 8 & -2 & & -2 \\ -2 & 8 & \ddots & \\ & \ddots & \ddots & -2 \\ -2 & & -2 & 8 \end{pmatrix}_{n\times n},$$

$B_1 = T_{2,10}$ and $B_2 = T_{3,10}$. The matrix C is generated such that $X = [X_{ij}]$ is the exact solution of $A_1XB_1 + A_2XB_2 = C$ where the nonzero elements of X are $X_{ii} = 1$, $i = 1, 2, \ldots, min(n, p)$, and $X_{i,i-1} = X_{i-1,i} = -1$ for $i = 2, 3, \ldots, min(n, p)$. The numerical results for different values of n are reported in Table 1 in which "iters" stands for the number of iterations required for the convergence.

Example 6.2. In this example, we compare the application of Gl-GMRES(5) with gradient based iterative method presented in [11]. To this end, the second example in the numerical experiment of [11] is chosen in which

$$A_1XB_1 + A_2XB_2 + C_1X^TD_1 + C_2X^TD_2 = F, \tag{6.1}$$

where

$$A_1 = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 3 & -1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix},$$

$$B_2 = \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix}, \quad C_1 = \begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix}, \quad C_2 = \begin{pmatrix} -1 & 3 \\ -1 & 2 \end{pmatrix},$$

$$D_1 = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 35 & 9 \\ 20 & 7 \end{pmatrix}.$$

Table 2. Numerical results for Example 6.2.

	Gl-GMRES(5)	Gradient based [11]
Iters	10	223
CPU-time(s)	0.047	0.187

As described in [11], it can be verified that the exact solution of (6.1) is

$$X = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}.$$

In this example, we set $X_0 = 10^{-6} I_{2\times 2}$ as the initial guess and the following stopping criterion is used

$$\delta_k = \frac{\|X_k - X\|_F}{\|X\|_F} < 10^{-5},$$

where X_k stands for the kth approximate solution.

In Table 2, the numerical results of the Gl-GMRES(5) algorithm together with the gradient based iterative method [11] have been illustrated. In this table, the corresponding CPU-time (in seconds) for computing the approximate solutions has been also reported. As observed, for this example the numerical results in terms of both number of iterations and CPU-time(s) for the Gl-GMRES(5) algorithm outperforms the gradient based iterative method proposed in [11]. We point out that in the gradient based iterative method the convergence factor μ should be chosen such that $0 < \mu \leq \mu_0$ where μ_0 defined by (5.2). For Example 6.2, $\mu_0 = \frac{1}{121.2}$ as given in [11]. However, the condition (5.2) for μ is a sufficient condition and the gradient based method may converge when $\mu \notin (0, \mu_0]$. In fact, the way of choosing a best convergence factor for gradient based algorithm is still a project to be studied which is a disadvantage for the gradient based algorithm. In [11], the best numerical result was reported for $\mu = \frac{1}{50}$ which does not satisfy (5.2). Hence, in Table 2 for the gradient based method we have set $\mu = \frac{1}{50}$.

7. Conclusion and further works

The applicability of the Gl-GMRES method for solving the general linear matrix equations has been studied. Some new theoretical results for the Gl-GMRES method, to solve the mentioned linear matrix equations, have been established. Numerical experiments have been given to illustrate the good performance of the Gl-GMRES method in comparison with the gradient based iterative method. The theoretical results elaborated in this work may be utilized for presenting links between Gl-GMRES method and its weighted versions for solving the general (coupled) linear matrix equations. In order to improve the speed of the convergence of the Gl-GMRES method for solving the large and sparse (coupled) linear matrix equations, the application of suitable preconditioners can be a subject of interest.

Acknowledgments

The author wishes to thank Dr. Davod Khojasteh Salkuyeh for carefully reading the first version of the paper and for his valuable comments and suggestions. The author also would like to express her sincere gratitude to Dr. Mehran Nomjoo and an anonymous referee for their helpful comments and recommendations which improved the quality of the paper.

References

[1] F. P. A. Beik and D. K. Salkuyeh, On the global Krylov subspace methods for solving general coupled matrix equation, *Comput. Math. Appl.* **62** (2011), no. 12, 4605–4613.

[2] A. Bouhamidi and K. Jbilou, A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications, *Appl. Math. Comput.* **206** (2008), no. 2, 687–694.

[3] R. Bouyouli, K. Jbilou, A. Messaoudi and H. Sadok, On block minimal residual methods, *Appl. Math. Lett.* **20** (2007), no. 3, 284–289.

[4] R. Bouyouli, K. Jbilou, R. Sadaka and H. Sadok, Convergence properties of some block Krylov subspace methods for multiple linear systems, *J. Comput. Appl. Math.* **196** (2006), no. 2, 498–511.

[5] M. Heyouni and A. Essai, Matrix Krylov subspace methods for linear systems with multiple right-hand sides, *Numer. Algorithms,* **40** (2005), no. 2, 137–156.

[6] K. Jbilou, A. Messaoudi and H. Sadok, Global FOM and GMRES algorithms for matrix equations, *Appl. Numer. Math.* **31** (1999), no. 1, 49–63.

[7] K. Jbilou, A. Messaoudi and K. Tabaâ, Some Schur complement identities and applications to matrix extrapolation methods, *Linear Algebra Appl.* **392** (2004) 195–210.

[8] Y. Saad, Iterative Methods for Sparse Linear Systems, Society for Industrial and Applied Mathematics, Philadelphia, 2003.

[9] I. Schur, Potenzreihn im Innern des Einheitskreises, *J. Reine Angew. Math.* **147** (1917) 205–232.

[10] F. Zhang, The Schur Complement and its Application, Springer-Verlag, New York, 2005.

[11] L. Xie, J. Ding and F. Ding, Gradient based iterative solutions for general linear matrix equations, *Comput. Math. Appl.* **58** (2009), no. 7, 1441–1448.

(Fatemeh Panjeh Ali Beik) Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran

E-mail address: `f.beik@vru.ac.ir, f.beik@member.ams.org`