Australian Journal of Basic and Applied Sciences, 5(8): 1128-1134, 2011
ISSN 1991-8178

Global Generalized Minimum Residual method for Solving Sylvester Equation AX+XB=C

[1]F. Panjeh Ali Beik, [2]M. Mohseni Mohgadam

[1]Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.
[2]Center of Excellence of Linear Algebra and Optimzation, Shahid BahonarUniversity of Kerman, Kerman, Iran.

Abstract: In the present paper, by extending the well-known global generalized minimum residual (GI-GMRES) method, we present new method for solving the Sylvester matrix equation AX+XB=C. the new method is based on the extending of the block krylov subspace and global Arnoldi process and will be called the Extended Gi-GMRES (EGI-GMRES). Theoretical results are given to show that the new method is convergent. Finally, some numerical examples are presented to illustrate the efficiency of our new method.

Key words: Gi-GMERS; Matrix equations; Block Krylov subspace.

INTRODUCTION

Consider the Sylvester matrix equation

$$AX + XB = C, \tag{1}$$

where $A \in \Re^{n\times n}$, $B \in \Re^{s\times s}$ are nonsingular matrices and $C \in \Re^{n\times s}$ is a given rectangular real matrix. Sylvester matrix equations play a key role and have numerous applications in control and communication theory, model reduction problems. Moreover, the solution of the Sylvester equations is needed in the block diagonalization of a matrix by a similarity transformation, in decoupling techniques for ordinary and partial differential equations, in filtering and image restoration; for more details refer to Heyouni, M., 2010 and the references therein.

A necessary and sufficient condition for (1.1) to have a unique solution is that

$$\lambda_i(A) + \mu_j(B) \neq 0,$$

for all $i = 1,\dots,n$, $j = 1,\dots s,$, where $\lambda_k(Z)$ is the *k*-th eigenvalue of the matrix. Throughout of this paper, we assume that the previous condition is satisfied.

In the following, we recall some definitions and theorems which are useful in the next sections.

We first recall the definition of Schur complements (Schur, I., 1917), then some of their properties are given.

Definition 1:

Let M be a matrix partitioned into four blocks:

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

where the submatrix D is assumed to be square and nonsingular. The Schur complement of D in M, denoted by (M / D), is defined by $(M / D) = A - BD^{-1}C.$

For two matrices $A, B \in \Re^{m\times n}$ we define the inner product $<A, B>_F = trace(B^T A).$ The associate norm is the well-known Frobenius norm denoted by $\|.\|_F$. For a matrix $A \in \Re^{m\times n}$ the so-called stretching function is defined as $vec(A) = (a_1^T\ a_2^T\ \dots a_n^T)^T$, where a_i is the *i*-th column of $A.$ Let $\mathcal{V} = [V_1, \dots, V_m]_{n\times ms}$, where V_i is an $n \times s$ matrix for $i = 1, 2, \dots, m.$ Corresponding to the $\mathcal{V}$, we associate a $ns \times m$ matrix $\mathcal{V}^v$ defined as $\mathcal{V}^v = [vec(V_1), vec(V_2), \dots, vec(V_m)].$ For the given

Corresponding Author: F. Panjeh Ali Beik, Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.
E-mail: f.beik@vru.ac.ir; panjeh_ali_beik@yahoo.com

matrices $A \in \Re^{n\times m}$ and $B \in \Re^{k\times l}$, the so-called Kronecker product of the matrices *A* and *B*, denoted by $A \otimes B$, is defined by the following $nk \times ml$ matrix, $A \otimes B = \left[a_{ij} B \right]$. Some properties of this product are given as follows (Lancaster, P., 1969) :

(1) $(A \otimes B)(E \otimes F) = (AE \otimes BF)$.

(2) If A and B are nonsingular matrices of dimension $n \times n$ and $p \times p$, respectively, then

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

(3) $vec(ABC) = (C^T \otimes A)\ vec(B)$.

(4) $vec(A)^T vec(B) = tr(A^T B)$.

(5) $\|A\|_F^2 = \|vec(A)\|_2$.

Definition 2:

(Bouyouli, R., Jbilou, K., 2007)

Let $A = [A_1, A_2, ..., A_p]$ and $B = [B_1, B_2, ..., B_\ell]$ be two 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 $A^T \lozenge B$ is defined by

$$[A^T \lozenge B]_{ij} = \langle A_i, B_j \rangle_F, \quad i = 1, 2, ..., p \quad j = 1, 2, ..., \ell.$$

Remarks:

(1) The matrix $A = [A_1, A_2, ..., A_p]$ is F-orthonormal iff $A^T \lozenge A = I_p$.

(2) If $X \in \Re^{n\times s}$, then $X^T \lozenge X = \|X\|_F^2$.

(3) Let $A, B \in \Re^{n\times ps}$, $L \in \Re^{p\times p}$ then $A^T \lozenge (B(L \otimes I_s)) = (A^T \lozenge B)L$.

Proposition 3:

(Bouyouli, R., Jbilou, K., 2007)

Let $A \in \Re^{n\times ps}, B \in \Re^{n\times ks}, C \in \Re^{k\times p}, G \in \Re^{k\times k}$ and $E \in \Re^{n\times s}$. If the matrix G is nonsingular matrix then

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

Proposition 4:

(Zhang, F., 2005, *p. 165*)

If the matrices M and D are square and nonsingular, then

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

The rest of this paper is organized as follows. In Section 2, we introduce the EGl-GMRES method for solving the Sylvester matrix equation (1). Section 3 is devoted to presenting some new theoretical results in order to prove the convergence of EGl-GMRES method. In order to illustrate the efficiency of the EGl-GMRES method, some numerical examples are given in Section 4. Finally, the paper is ended with a brief conclusion in Section 5.

Extended Gl-Gmres (EGL-Gmres):

In this section by extending the Gl-GMRES method (Jbilou, k., Messaoudi, A., and Sadok, H., 1999), we propose a new method for solving the Sylvester matrix equation (1). To this end, first, we need to extend the definition of the block Krylov subspace in the following.

Definition 5:

Suppose that $A \in \Re^{n\times n}, B \in \Re^{s\times s}$, and $V \in \Re^{n\times s}$ we define the extended block Krylov subspace as follows:

$\mathcal{EK}_m(A,V,B) = span\{W_1, W_2, ..., W_m\}$,
where $W_1 = V$ and $W_j = AW_{j-1} + W_{j-1}B$ for $j = 1, 2, ..., m$.

Now, the generalized global Arnoldi process is presented which constructs an *F*-orthonormal basis $V_1, V_2, ..., V_m$ for the extended block Krylov subspace $\mathcal{EK}_m(A,V,B)$.

Algorithm 1:

(Generalized global Arnoldi process)

1. Choose an $n \times s$ matrix *V*. Set $\beta = \|V\|_F$, $V_1 = V/\beta$,
2. For $j = 1, 2, ..., m$ Do:
3. $W = AV_j + V_j B$
4. For $i = 1, 2, ..., j$ Do:
5. $h_{i,j} = \langle W, V_i \rangle_F$
6. $W = W - h_{ij} V_i$
7. EndDo
8. $h_{j+1,j} = \|W\|_F$. If $h_{j+1,j} = 0$ Stop
9. $V_{j+1} = W / h_{j+1,j}$
10. EndDo.

Denote by $\mathcal{V}_m$, an $n \times ms$ block matrix defined by the columns $V_1, V_2, ..., V_m$, by $\overline{H}_m$, the $(m+1) \times m$ Hessenberg matrix whose nonzero entries h_{ij}, $i = 1, 2, ..., m+1$, $j = 1, ..., m$, are defined by Algorithm 1, and by H_m, the matrix obtained from $\overline{H}_m$ by deleting its last row. It is not difficult to see that $\left(\mathcal{V}_m^{\nu}\right)^T \mathcal{V}_m^{\nu} = I$.

Theorem 6:

Let $\mathcal{V}_m$, H_m and $\overline{H}_m$, be defined as before. Then the following relation holds

$$A\mathcal{V}_m + \mathcal{V}_m(I_m \otimes B) = \mathcal{V}_m(H_m \otimes I_s) + h_{m+1,m} V_{m+1}(e_m^T \otimes I_s) = \mathcal{V}_{m+1}(\overline{H}_m \otimes I_s) \ . \tag{2}$$

Proof:

From lines 3, 5 and 9 in Algorithm 1, we deduce that

$$AV_j + V_j B = \sum_{i=1}^{j+1} h_{ij} V_i \ , \quad j = 1, 2, ..., m.$$

By the definition of the Kronecker product, the result follows immediately.

Now, we present our new method, EGl-GMRES method, for solving Sylvester matrix equation (1). In EGl-GMRES, for a given initial a approximation X_0, the new approximation X_m computed such that:
$X_m \in X_0 + \mathcal{EK}_m(A, R_0, B)$.

Or Equivalently:

$$X_m = X_0 + \mathcal{V}_m(y_m \otimes I_s), \tag{3}$$

and the vector $y_m \in \Re^m$ is obtained by imposing the following orthogonality condition:

$$R_m = C - AX_m - X_m B \perp_F A\mathcal{EK}_m(A, R_0, B) + \mathcal{EK}_m(A, R_0, B)(I_m \otimes B). \tag{4}$$

Note that in a special case when $B = 0_{s \times s}$, the EGl-GMRES method reduces to the Gl-GMRES method (Jbilou, k., Messaoudi, A., and Sadok, H., 1999).

Straightforward computations show that

$$R_m = C - AX_m - X_m B = C - AX_0 - X_0 B - A\mathcal{V}_m(y_m \otimes I_s) - \mathcal{V}_m(y_m \otimes I_s)B$$
$$= R_0 - A\mathcal{V}_m(y_m \otimes I_s) - \mathcal{V}_m(I_m \otimes B)(y_m \otimes I_s)$$
$$= R_0 - [A\mathcal{V}_m + \mathcal{V}_m(I_m \otimes B)](y_m \otimes I_s).$$

On the other hand, in a similar way which is employed (by M. Mohseni Moghadam and F. Panjeh Ali Beik, 2010), it can be proved that the approximate solution X_m, computed by imposing the orthogonality condition (4), is the solution of the following least-square problem:

$$\min_{X \in X_0 + \mathcal{E}\mathcal{K}_m} \| C - AX - X B \|_F . \tag{5}$$

From the relation (2) and some easy computations, it is not difficult to see that:

$$\| C - AX - X B \|_F = \| R_0 - [A\mathcal{V}_m + \mathcal{V}_m(I_m \otimes B)](y \otimes I_s) \|_F$$
$$= \| \mathcal{V}_{m+1}(\beta e_1 \otimes I_s) - [A\mathcal{V}_m + \mathcal{V}_m(I_m \otimes B)](y \otimes I_s) \|_F$$
$$= \| \mathcal{V}_{m+1}(\beta e_1 \otimes I_s) - \mathcal{V}_{m+1}(\overline{H}_m \otimes I_s)(y \otimes I_s) \|_F = \| \mathcal{V}_{m+1}(\beta e_1 \otimes I_s) - \mathcal{V}_{m+1}(\overline{H}_m y \otimes I_s) \|_F$$
$$= \| \mathcal{V}_{m+1}[(\beta e_1 - \overline{H}_m y) \otimes I_s] \|_F = \| \mathcal{V}^{v}_{m+1}(\beta e_1 - \overline{H}_m y) \|_2$$
$$= \| \beta e_1 - \overline{H}_m y \|_2 .$$

Therefore, the vector y_m is the solution of the following least square problem

$$\min_{y \in \mathbb{R}^m} \| \beta e_1 - \overline{H}_m y \|_2 . \tag{6}$$

Note that the EGl-GMRES algorithm requires the storage of $\mathcal{V}_m$. That is, in order to save the vector $\mathcal{V}_m$, we need an *m* dimensional vectors space whose entries are $n \times s$ matrices. To cure the storage problem, encountered also in GMRES, the value of *m* is limited by storage constraint and by avoiding rounding errors. Hence, this algorithm can be restarted after *m* iterations. The corresponding algorithm is called the restarted EGl-GMRES (*m*), see Saad, Y., 1995

Algorithm 2:

(EGl-GMRES (*m*))

1. Choose X_0, *m* and a tolerance ε, Compute $R_0 = C - AX_0 - X_0 B$, Set $V = R_0$.
2. Construct the F-orthonormal basis $V_1, V_2, \ldots, V_m$ by the generalized global Arnoldi process.
3. Find y_m as the solution of $\min_{y \in \mathbb{R}^m} \| \beta e_1 - \overline{H}_m y \|_2$.
4. Compute the approximate solution $X_m = X_0 + \mathcal{V}_m(y_m \otimes I_s)$, and $R_m = C - AX_m - X_m B$.
5. If $\|R_m\|_F < \varepsilon$ Stop.
6. Set $X_0 = X_m$, $R_0 = R_m$, $V = R_0$, and go to 2.

Convegence Properties:

In this section, we will establish some new expressions for the norms of the residual matrices corresponding to the approximations obtained by the EGl-GMRES. Then, it is shown that the EGl-GMRES method is convergent. From orthogonality condition (4), we deduce that

$$(A\mathcal{V}_m + \mathcal{V}_m(I_m \otimes B))^T \lozenge R_m = 0.$$

For simplicity, we set $\mathcal{W}_m = A\mathcal{V}_m + \mathcal{V}_m(I_m \otimes B)$. The above relation and the properties of the $\lozenge$ product imply that

$$\mathcal{W}_m^T \lozenge R_0 = (\mathcal{W}_m^T \lozenge \mathcal{W}_m) y_m . \tag{7}$$

Theorem 7:

Suppose that $\mathcal{W}_m^T \lozenge \mathcal{W}_m$, is a nonsingular matrix. Then the residual matrix R_m can be expressed by the following Schur complements

$$R_m = \begin{pmatrix} R_0 & \mathcal{W}_m \\ (\mathcal{W}_m^T \lozenge R_0)\otimes I_s & (\mathcal{W}_m^T \lozenge \mathcal{W}_m)\otimes I_s \end{pmatrix} \Big/ ((\mathcal{W}_m^T \lozenge \mathcal{W}_m)\otimes I_s), \tag{8}$$

where $\mathcal{W}_m = A\mathcal{V}_m + \mathcal{V}_m(I_m \otimes B)$.

Proof:

It is not difficult to see that

$R_m = R_0 - \mathcal{W}_m(y_m \otimes I_s)$.

From (7), we have

$R_m = R_0 - \mathcal{W}_m\left[[(\mathcal{W}_m^T \lozenge \mathcal{W}_m)^{-1}(\mathcal{W}_m^T \lozenge R_0)]\otimes I_s\right]$.

By using the properties of the Kronecker product, we get

$$R_m = R_0 - \mathcal{W}_m\left[[(\mathcal{W}_m^T \lozenge \mathcal{W}_m)^{-1}\otimes I_s][(\mathcal{W}_m^T \lozenge R_0)\otimes I_s]\right]$$

Hence, we can easily conclude the result from the definition of the Schur complement.

Theorem 8:

Suppose that R_m is the residual matrix obtained by the EGl-GMRES method to (1). Let $\mathcal{W}_m^T \lozenge \mathcal{W}_m$ be a nonsingular matrix. Then

$$\| R_m \|_F^2 = \left((\overline{\mathcal{V}}_{m+1}^T \lozenge \overline{\mathcal{V}}_{m+1}) \big/ (\mathcal{W}_m^T \lozenge \mathcal{W}_m) \right), \tag{9}$$

where $\mathcal{W}_m = A\mathcal{V}_m + \mathcal{V}_m(I_m \otimes B)$ and $\overline{\mathcal{V}}_{m+1} = [R_0, \mathcal{W}_m]$.

Proof:

Note that $R_m = R_0 - \mathcal{W}_m(y_m \otimes I_s)$. Invoking the orthogonality condition (8), we deduce :

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

On the other hand, it is well-known that $\| R_m \|_F^2 = R_m^T \lozenge R_m$, therefore

$$\| R_m \|_F^2 = R_m^T \lozenge R_m = R_0^T \lozenge R_m.$$

Using Proposition 3 and the relation (8), we have:

$$R_0^T \lozenge R_m = \begin{pmatrix} R_0^T \lozenge R_0 & R_0^T \lozenge \mathcal{W}_m \\ \mathcal{W}_m^T \lozenge R_0 & \mathcal{W}_m^T \lozenge \mathcal{W}_m \end{pmatrix} \Big/ (\mathcal{W}_m^T \lozenge \mathcal{W}_m) = (\overline{\mathcal{V}}_{m+1}^T \lozenge \overline{\mathcal{V}}_{m+1}) \big/ (\mathcal{W}_m^T \lozenge \mathcal{W}_m).$$

Theorem 9:

Let $\overline{\mathcal{V}}_{m+1} = [R_0, \mathcal{W}_m]$. Assume that $\overline{\mathcal{V}}_{m+1}^T \lozenge \overline{\mathcal{V}}_{m+1}$ and $\mathcal{W}_m^T \lozenge \mathcal{W}_m$ are nonsingular matrices. Then, the residual R_m satisfies in the following relation:

$$\| R_m \|_F^2 = 1/(e_1^T (\overline{\mathcal{V}}_{m+1}^T \lozenge \overline{\mathcal{V}}_{m+1})^{-1} e_1).$$

Proof:

By the assumption the matrices $\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1}$ and $\mathcal{W}_m^T \Diamond \mathcal{W}_m$ are nonsingular matrices, therefore the Schur complement $(\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1}) / (\mathcal{W}_m^T \Diamond \mathcal{W}_m)$ is nonzero. Hence, Proposition 4 implies that

$$e_1^T (\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1})^{-1} e_1 = \left[(\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1}) / (\mathcal{W}_m^T \Diamond \mathcal{W}_m) \right]^{-1},$$

The result follows from Theorem 8 immediately.

The following theorem shows that the EGl-GMRES method is convergent for solving Sylvester matrix equation (1).

Theorem 10:

The residual R_m satisfies in the following relation:

$$\frac{4\chi(\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1})}{(1+\chi(\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1}))^2} \le \frac{\|R_m\|_F^2}{\|R_0\|_F^2} \le 1,$$

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

Proof:

It is not difficult to see that $R_0^T \Diamond R_0 = e_1^T (\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1})\ e_1$. So, by using *Kantorovich* inequality (Saad, Y., 1995, Chapter 5), we derive:

$$R_0^T \Diamond R_0 \ge 1 \Big/ (e_1^T (\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1})^{-1} e_1) \ge \left[4\chi(\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1}) \Big/ (1+\chi(\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1}))^2 \right] R_0^T \Diamond R_0.$$

By using Theorem 9, we rewrite the above relation

$$\|R_0\|_F^2 \ge \|R_m\|_F^2 \ge \left[4\chi(\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1}) \Big/ (1+\chi(\overline{\mathcal{V}}_{m+1}^{T} \Diamond \overline{\mathcal{V}}_{m+1}))^2 \right] \|R_0\|_F^2 .$$

Numerical Experiments:

The results reported in this section were run using Mathematica 6. For all the experiments the initial guess X_0 was taken to be the zero $n \times s$ matrix. The test were stopped as soon as $\|R_m\|_F / \|R_0\|_F \le 10^{-5}$. The matrix C in (1) is generated such that $X = [X_{ij}]_{n \times s}$ is the solution of the Sylvester equation $AX + XB = C$, where nonzero elements of X are $X_{ii} = 1$ for $i = 1, 2, \ldots, \min(n, s)$.

Example 1:

Let the matrices A, B are $n \times n$ and $s \times s$, respectively, defined as

$$A = tridiag\left\{-1-\frac{10}{n+1}, 2, -1+\frac{10}{n+1}\right\}, \quad B = tridiag\left\{-1-\frac{10}{s+1}, 2, -1+\frac{10}{s+1}\right\}.$$

The results of performing EGl-GMRES (5) are presented in Table 1 for different values of *n,s*.

Table 1:

n	*S*	*iteration*	$\|R_m\|_F$
1000	*10*	*20*	0.550590×10^{-6}
2000	*10*	*33*	0.311107×10^{-6}
3000	*10*	*35*	0.443554×10^{-6}

Example 2:

Consider the Sylvester equation $AX + XB = C$ where A, B both taken from Harwell-Boeing Collocation. In fact, we choose NOS6 (675×675) and NOS5 (468×468) from the set LANPRO.
The result of applying EGl-GMRES(5) are given in the following table.

Table 2:

A	*B*	*Iteration*	$\|R_m\|_F$
NOS5	*Example 1*	*35*	0.443554×10^{-6}
NOS5	*NOS5*	*92*	0.116622×10^{-6}
NOS6	*NOS5*	*12*	0.311388×10^{-6}

Conclusion:

In this paper, we present an iterative projection method for solving Sylvester matrix equations. We also derived some new convergent theorems for our new method. Numerical examples show that our method is interesting.

REFERENCES

Bouyouli, R., K. Jbilou, 2007. A. Messaoudi and H. Sadok, On block minimal residual methods, Appl. Math. Lett., 20: 284-289.

El Guennouni, A., K. Jbilou and A.J. Riquet, 2002. Block Krylov subspace methods for solving large Sylvester equations, Numer. Alg., 29: 75-96.

Heyouni, M., 2010. Extended Arnoldi methods for low-rank Sylvester matrix equations, Appl. Numer. Math., 60: 1171-1182.

Jbilou, K., A. Messaoudi and H. Sadok, 1999. Global FOM and GMRES algorithms for matrix equations, Appl. Numer. Math., 31: 43-49.

Lancaster, P., 1969. Theory of Matrix, Academic Press, New York.

Lin, Y.Q., 2005. Implicitly restarted global FOM and GMRES for nonsymmetric matrix equations and Sylvester equations, Appl. Math. Comput., 167: 1004-1025.

Mohseni Moghadam, M., F. Panjeh Ali Beik, On a new weighted global GMRES for solving nonsymmetric multiple linear systems, Int. J. Contemp. Math. Sciences, 5(45): 2237-2255.

Saad, Y., 1995. Iterative Methods for Spars Linear System, PWS Press, New York.

Schur, I., 1917. Potenzreihn im Innern des Einheitskreises, J. Reine Angew. Math., 147: 205-232.

Zhang, F., 2005. The Schur complement and its application, New York.