International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:10, No:9, 2016

The BGMRES Method for Generalized Sylvester Matrix Equation $AXB - X = C$ and Preconditioning

Azita Tajaddini, Ramleh Shamsi

Abstract—In this paper, we present the block generalized minimal residual (BGMRES) method in order to solve the generalized Sylvester matrix equation. However, this method may not be converged in some problems. We construct a polynomial preconditioner based on BGMRES which shows why polynomial preconditioner is superior to some block solvers. Finally, numerical experiments report the effectiveness of this method.

Keywords—Linear matrix equation, Block GMRES, matrix Krylov subspace, polynomial preconditioner.

I. Introduction

CONSIDER the generalized linear sylvester equation

$$AXB - X = C, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{s\times s}$, $C \in \mathbb{R}^{n\times s}$, and X is the unknown matrix in $\mathbb{R}^{n\times s}$. The matrix equation (1) plays an important role in control and communication theory; see [3], [4]. Also, the discrete -time time -invariant linear systems

$$x_{k+1} = Ax_k + Bu_k,$$
$$y_k = cx_k,$$

$k = 0, 1, 2, ...$ and x_0 is given. These systems are associated, for instance, with the discrete -time Lyapunov equation

$$AXA^T - X = -BB^T.$$

It arises naturally in a wide variety of control applications such as stability analysis [17] model order reduction [12], [15], [18] and Newton's method for discrete algebraic Riccati equations [2]. The discrete-time Lyapunov equation is a special case of matrix equation (1).

The analytical solution of the matrix equation (1) has been considered by many authors; see [6]. They have proposed the Bartels-Stewart algorithm for solving matrix equation (1). The direct methods for solving the matrix equation (1) such as those proposed in [6], [7] are attractive if the matrices are of small size. These methods are based on the Schur's decomposition, by which the original equation is transformed into a form that is easily solved by a forward substitution. In recent years, several iterative methods based on the Arnoldi process have been proposed to solve the matrix equation (1). For example, in [1], Bao et al. proposed an iterative method which allows us to construct an orthonormal basis of certain Krylov subspace and simultaneously reduce the order of the generalized Sylvester equation (1). By the same token, in [3], Bouhamidi et al. proposed the global generalized minimum residual (GL - GMRES) to solve the following general linear matrix equations:

$$\sum_{j=1}^{p} A_j X B_j = C,$$

and then they applied ILU (incomplete LU factorization) and SSOR preconditioning in order to solve the Sylvester equation AX+XB=C. Here, we use ILU and SSOR preconditioning the same as the methods that [3] were used.

In this paper, we generalize the GMRES method to obtain a matrix iterative method to solve the matrix equation (1). Also, we present a polynomial preconditioner. Then, we compare these methods with BGMRES (ILU), BGMRES (SSOR) and NSCG [19], squared Smith (SM) and restarted Krylov squared Smith (RKSS) [10], [16] methods. Finally, we show that PBGMRES is more effective than the other methods.

Let $X = [x_1, x_2, \ldots, x_s]$, where $x_i, i = 1, 2, ..., s$ is ith column of X. We define a linear operator

$$vec : \mathbb{R}^{n\times s} \longrightarrow \mathbb{R}^{ns}$$
$$X \longmapsto [x_1^T, x_2^T, \ldots, x_s^T]^T \tag{2}$$

Hence, the linear matrix equation (1) can be written as the following $ns \times ns$ linear system

$$\mathcal{A}x = c, \tag{3}$$

where $\mathcal{A} = (B^T \otimes A - I_{ns}), x = vec(X), c = vec(C)$ and $\otimes$ denotes the Kronecker product, see [4]. This product satisfies the properties

$$(A \otimes B)(C \otimes D) = (AC \otimes BD),\ (A \otimes B)^T = A^T \otimes B^T.$$

Equation (3) has a unique solution if and only if the matrix $\mathcal{A}$ is nonsingular.

Throughout this paper, we use the following notations: Let $X, Y \in \mathbb{R}^{n\times p}$, the Frobenius inner product is defined $< X, Y >_F = tr(X^T Y)$, where $tr(.)$ denotes the trace and X^T the transpose of the matrix X. The associated norm is the well-known Frobenius norm denotes by $\|.\|_F$. A system of matrices of $\mathbb{R}^{n\times p}$ is said to be F-orthogonal, if it is orthogonal with respect to the scalar product $< .,. >_F$, that means $tr(X^T Y) = 0$. The Kronecker product of $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$ is defined by $A \otimes B := [a_{ij}B] \in \mathbb{R}^{mp\times nq}$. We

A. Tajaddini is with the Department of Applied Mathematics, Faculty of Mathematics & Computer Sciences, Shahid Bahonar University of Kerman, Kerman, Iran (e-mail: atajadini@uk.ac.ir).
R. Shamsi is with Anonymous University.

use the notation $*$ for the following product [8], [9], let $\mathcal{V}_m$ denote the $n \times ms$ block matrix $\mathcal{V}_m = [V_1, \dots, V_m]$, where $V_i \in \mathbb{R}^{n\times s}$ for $i = 1, \dots, m$. If $D_m = [d_1, \dots, d_m] \in \mathbb{R}^{m\times m}$ and $\alpha = [\alpha_1, \dots, \alpha_m]^T \in \mathbb{R}^m$, then we have

$$\mathcal{V}_m * \alpha = \sum_{i=1}^{m} \alpha_i V_i,$$
$$\mathcal{V}_m * D_m = [\mathcal{V}_m * d_1, \dots, \mathcal{V}_m * d_m].$$

If $H_m \in \mathbb{R}^{m\times m}$ and $\alpha, \beta \in \mathbb{R}^m$, then the matrix product $*$ satisfies the following properties

$$\mathcal{V}_m * (\alpha + \beta) = (\mathcal{V}_m * \alpha) + (\mathcal{V}_m * \beta),$$
$$(\mathcal{V}_m * H_m) * \alpha = \mathcal{V}_m * (H_m \alpha),$$
$$(\mathcal{V}_m * \alpha)^T = \mathcal{V}_m^T * \alpha.$$

The remainder of this paper is organized as follows: In Section II, a description of the block generalized minimal residual (BGMRES) is given. In Section III, we show how to apply polynomial preconditioner in order to solve the generalized Sylveter matrix equation $AXB - X = C$. Section IV is devoted to some numerical experiments. Finally, conclusion is given in Section V.

II. The Block GMRES Method

In this section, we define the mth generalized matrix Krylov subspace and recall the modified global Arnoldi process, for more details, see [11], [13].

Definition 1. *Let V be any $n\times s$ matrix. Then, the generalized matrix Krylov subspace is associated to (A, V, B) and an integer m is defined as*

$$\mathcal{GK}_m(A, V, B) = span\{V, AVB, \dots, A^{m-1}VB^{m-1}\}.$$

The modified global Arnoldi process allows us to construct an F-orthonormal basis for the generalized matrix Krylov $\mathcal{GK}_m(A, V, B)$, see [11], [13].

Algorithm 1 The Modified Global Arnoldi Algorithm

Require: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{s\times s}$, $m \in \mathbb{N}$ and the nonzeros matrix $V \in \mathbb{R}^{n\times s}$.
Ensure: The block vectors $V_1, V_2, ..., V_{m+1}$ and the semi upper hessenberg matrix $\bar{H}_m = (h_{ij})$.

```
set V_1 = V/||V||_F;
for j = 1,...,m do
    W = AV_j;
    W = WB;
    for i = 1,...,j do
        h_ij = <V_i, W>_F = tr(W^T V_i);
        W = W - h_ij V_i;
    end for
    Compute h_{j+1,j} = ||W||_F; if h_{j+1,j} = 0, stop
    Compute V_{j+1} = W/h_{j+1,j};
end for
```

Since the modified Arnoldi algorithm involves the Gram-Schmidt process, algorithm 1 builds an F-orthonormal basis $\mathcal{V}_m = [V_1, V_2, \dots, V_m]$, $V_i \in \mathbb{R}^{n\times s}$ for the generalized Krylov subspace $\mathcal{GK}_m(A, V, B)$ and a semi upper Hessenberg matrix $\bar{H}_m \in \mathbb{R}^{m+1\times m}$. The following theorem can be easily proved.

Theorem 1. *Let $\mathcal{V}_m, \bar{H}_m$ and H_m be as defined above. The global Arnoldi process satisfies the following*

1) $A\mathcal{V}_m(I_m \otimes B) = \mathcal{V}_m * H_m + E_{m+1}$, *where* $E_{m+1} = h_{m+1,m}[O_{n\times s}, \dots, O_{n\times s}, V_{m+1}]$.
2) $A\mathcal{V}_m(I_m \otimes B) = \mathcal{V}_{m+1} * \bar{H}_m$.
3) *For any $(m+1)\times s$ matrix G, we have* $\|\mathcal{V}_{m+1} * G\|_F = \|G\|_2$.

Theorem 2. *Let $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{s\times s}$, $C \in \mathbb{R}^{n\times s}$. Let $X_0 \in \mathbb{R}^{n\times s}$ be an initial guess and R_0 is its corresponding residual. Then*

$$\mathcal{K}_m(B^T \otimes A, r_0) = \mathcal{K}_m(\mathcal{A}, r_0),$$

where $\mathcal{A} = (B^T \otimes A - I_{ns})$, $r_0 = vec(R_0)$ and $\mathcal{K}_m(\mathcal{A}, r_0)$, $\mathcal{K}_m(B^T \otimes A, r_0)$ are the classic Krylov subspaces.

Remark 1. *Suppose that*

$$\mathcal{K}_m(B^T\otimes A, r_0) = span\{r_0, (B^T\otimes A)r_0, \dots, (B^T\otimes A)^{m-1}r_0\},$$

where $r_0 = vec(R_0)$. The map

$$T : \mathcal{GK}_m(A, R_0, B) \longrightarrow \mathcal{K}_m(B^T \otimes A, r_0)$$

given by $Z \longrightarrow vec(Z)$ is an isomorphism. Hence, by theorem 2 and above discussion, the two subspace $\mathcal{GK}_m(A, R_0, B)$ and $\mathcal{K}_m(\mathcal{A}, r_0)$ are isomorph, i.e.

$$\mathcal{GK}_m(A, R_0, B) \simeq \mathcal{K}_m(\mathcal{A}, r_0).$$

Therefore, we can conclude that

$$A\mathcal{GK}_m(A, R_0, B) \simeq \mathcal{AK}_m(\mathcal{A}, r_0), \quad (4)$$

where

$$A\mathcal{GK}_m(A, R_0, B)B = span\{AR_0B, A^2R_0B^2, \dots, A^mR_0B^m\},$$

and

$$\mathcal{AK}_m(\mathcal{A}, r_0) = span\{\mathcal{A}r_0, \mathcal{A}^2r_0, \dots, \mathcal{A}^m r_0\}.$$

Let $X_0 \in \mathbb{R}^{n\times s}$ be an initial guess and the corresponding residual is
$R_0 = C - AX_0B + X_0$. The block GMRES algorithm, at the mth step, constructs the approximation solution X_m to the solution of (1) such that

$$X_m = X_0 + Z_m \qquad s.t. \qquad Z_m \in \mathcal{GK}_m(A, R_0, B), \quad (5)$$

with F-orthogonality relation

$$R_m = C - AX_mB + X_m \perp_F A\mathcal{GK}_m(A, R_0, B)B. \quad (6)$$

In theorem 3, we show that F-orthogonality (6) is equivalent to minimization problem (7).

Theorem 3. *Let A and B be two arbitrary matrices. Let X_0 be an initial guess and R_0 is its corresponding residual. Then a matrix X_m is the result of an oblique projection method onto $\mathcal{GK}_m(A, R_0, B)$ and F-orthogonal to $A\mathcal{GK}_m(A, R_0, B)B$ if and only if it minimizes the F-norm of the residual matrix*

$R = C - AXB + X$ over
$X \in X_0 + \mathcal{GK}_m(A, R_0, B)$, i.e, if and only if

$$||R_m||_F = \min_{X \in X_0 + \mathcal{GK}_m(A,R_0,B)} ||R||_F,$$

where $R_m = C - AX_mB + X_m$.

Therefore, the F-orthogonality (6) is equivalent to minimizing the Frobenius norm of the residual

$$\|R_m\|_F = \min_{Z \in \mathcal{GK}_m(A,R_0,B)} \|C - A(X_0 + Z)B + X_0 + Z\|_F \tag{7}$$

i.e. X_m solves the minimization problem (7). Consider the F-orthogonal basis $\mathcal{V}_m$, constructed with algorithm 1. Application of m steps of the algorithm 1 to the matrices A and B with the nonzero residual matrix R_0 yields theorem 1. The least squares problem (7) can be reformulated as:

$$\begin{aligned}
&\|C - A(X_0 + Z)B + X_0 + Z\|_F \\
&= \|C - A(X_0 + \mathcal{V}_m * y_m)B + X_0 + \mathcal{V}_m * y_m\|_F \\
&= \|R_0 - A(\mathcal{V}_m * y_m)B + \mathcal{V}_m * y_m\|_F \\
&= \|R_0 - (A\mathcal{V}_m(I_m \otimes B)) * y_m + \mathcal{V}_m * y_m\|_F.
\end{aligned}$$

By $*$ product properties and theorem 1, we have

$$\begin{aligned}
&\|R_0 - (A\mathcal{V}_m(I_m \otimes B) - \mathcal{V}_m) * y_m\|_F \\
&= \|\|R_0\|_F V_1 - (\mathcal{V}_{m+1} * (\bar{H}_m - \begin{pmatrix} I_m \\ 0_{s\times m} \end{pmatrix})) * y_m\|_F \\
&= \|\mathcal{V}_{m+1} * (\|R_0\|_F e_1 - \begin{pmatrix} H_m - I_m \\ h_{m+1,m}e_m^T \end{pmatrix} y_m)\|_F \\
&= \|\|R_0\|_F e_1 - \begin{pmatrix} H_m - I_m \\ h_{m+1,m}e_m^T \end{pmatrix} y_m\|_2.
\end{aligned}$$

The above results are summarized in the following theorem.

Theorem 4. *At step m, the approximation solution X_m produced by the block GMRES method is given by $X_m = X_0 + \mathcal{V}_m * y_m$ where y_m is the solution of the following small least square problem*

$$y_m = arg \min_{y \in \mathbb{R}^m} \|\|R_0\|_F e_1 - \begin{pmatrix} H_m - I_m \\ h_{m+1,m}e_m^T \end{pmatrix} y\|_2, \tag{8}$$

where e_1 is the first unit vector of $\mathbb{R}^{m+1}$.

The minimization (8) can be solved by the QR factorization of $\begin{pmatrix} H_m - I_m \\ h_{m+1,m}e_m^T \end{pmatrix}$ with Givens rotations.

By algorithm 1 , the $\begin{pmatrix} H_m - I_m \\ h_{m+1,m}e_m^T \end{pmatrix}$ is an unreduced semi upper Hessenberg matrix and $rank(\begin{pmatrix} H_m - I_m \\ h_{m+1,m}e_m^T \end{pmatrix}) = m$. Then, there is an orthogonal matrix $Q_m \in \mathbb{R}^{(m+1)\times(m+1)}$ and an invertible upper triangular matrix $\bar{R}_m \in \mathbb{R}^{(m)\times(m)}$ such that

$$\begin{pmatrix} H_m - I_m \\ h_{m+1,m}e_m^T \end{pmatrix} = Q_m \begin{pmatrix} \bar{R}_m \\ 0 \end{pmatrix} \tag{9}$$

With substitution (9) into (8), we obtain

$$\|R_m\|_F = \min_{y \in \mathbb{R}^m} \|\|R_0\|_F (Q_m^T e_1) - \begin{pmatrix} \bar{R}_m \\ 0 \end{pmatrix} y\|_2$$

for obtaining y_m, we solve the following upper triangular system:

$$\bar{R}_m y = \|R_0\|_F \begin{pmatrix} I_m, 0 \end{pmatrix} (Q_m^T e_1).$$

Then, we get the BGMRES iterative solution to (1),

$$X_m = X_0 + \mathcal{V}_m * y_m.$$

It can be easily shown that the residual matrix form is as:

Theorem 5. *The residual matrix at step m, $R_m = C - AX_mB + X_m$ produced by the block GMRES for the linear matrix equation satisfies the following properties*

$$R_m = \gamma_{m+1}\mathcal{V}_{m+1} * \big(Q_m e_{m+1}\big),$$

and

$$\|R_m\|_F = |\gamma_{m+1}|,$$

where γ_{m+1} is the last component of the vector $g_m = \|R_0\|_F Q_m^T e_1$ and
$e_{m+1} = (0, \ldots, 0, 1)^T \in \mathbb{R}^{m+1}$.

Finally, the previous results can summarized in the following algorithm.

Algorithm 2 The block GMRES algorithm (BGMRES)

Require: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{s\times s}$, an initial guess matrix $X_0 \in \mathbb{R}^{n\times s}$, $m \in \mathbb{N}$ and ε.
Ensure: The solution X_m.
Compute $R_0 = C - AX_0B + X_0, \beta = \|R_0\|_F, V_1 = \dfrac{R_0}{\beta}$;
Construct the F-orthonormal basis $V_1, V_2, ..., V_{m+1}$ and the semi upper Hessenberg matrix $\bar{H}_m$ by Algorithm 1;
Solve the least squares problem

$$y_m = arg \min_{y \in \mathbb{R}^m} \|\|R_0\|_F (Q_m^T e_1) - \begin{pmatrix} \bar{R}_m \\ 0 \end{pmatrix} y\|_2$$

Compute: $X_m = X_0 + \mathcal{V}_m * y_m$;
Compute the residual $R_m = C - AX_mB + X_m$ and $\|R_m\|_F$ by using Theorem 1;
if $\|R_m\|_F < \varepsilon$ **then**
  Stop;
end if
  set $X_0 = X_m$ Go to 1;

In the next section, we construct the polynomial preconditioned BGMRES based on BGMRES.

III. Polynomial Preconditioning

Consider the generalized Sylvester matrix equation (1). In Section II, we will show that this equation can be solved by the block GMRES method, however this method may slow down the convergence. In order to accelerate the convergence, we construct a polynomial preconditioned BGMRES based on BGMRES. Let $X_0 \in \mathbb{R}^{n\times s}$ be an initial guess and $R_0 = C - AX_0B + X_0$ is its corresponding residual. Define the $n \times ms$ block matrix $\mathbb{K}_m$ as:

$$\mathbb{K}_m = \big[R_0, AR_0B, \ldots, A^{m-1}R_0B^{m-1}\big]. \tag{10}$$

Assume that Algorithm 1 does not stop before the mth step. Then by relation 1 from the theorem 1, we get

$$V_{m+1} = h_{m+1,m}^{-1}\big(AV_mB - \mathcal{V}_m * H_{.,m}\big), \tag{11}$$

where

$$H_{.,m} = (h_{1m}, \ldots, h_{mm})^T.$$

Since the space generated by the matrix $\mathbb{K}_m$ is the same as the space spanned by the matrix $\mathcal{V}_m$. Therefore, we have

$$\mathcal{V}_m = \mathbb{K}_m * S_m, \tag{12}$$

where S_m is an upper triangular matrix

$$S_m = \begin{bmatrix} s_{11} & \ldots & s_{1m} \\ & \ddots & \vdots \\ 0 & & s_{mm} \end{bmatrix}.$$

By Kronecker and * product properties, we have

$$\begin{aligned} \mathcal{V}_m * H_{.,m} &= (\mathbb{K}_m * S_m) * H_{.,m} \\ &= \mathbb{K}_m * (S_m H_{.,m}) \\ &= \big[\mathbb{K}_m, A^m R_0 B^m\big] * \begin{pmatrix} S_m H_{.,m} \\ 0 \end{pmatrix} \\ &= \mathbb{K}_{m+1} * \begin{pmatrix} S_m H_{.,m} \\ 0 \end{pmatrix}. \end{aligned} \tag{13}$$

Since $A\mathcal{V}_m(I_m \otimes B) = (A\mathbb{K}_m(I_m \otimes B)) * S_m$, then

$$\begin{aligned} AV_mB &= (A\mathbb{K}_m(I_m \otimes B)) * S_{.,m} \\ &= \big(AR_0B, A^2R_0B^2, \ldots, A^mR_0B^m\big) * S_{.,m} \\ &= \mathbb{K}_{m+1} * \begin{pmatrix} 0 \\ S_{.,m} \end{pmatrix}. \end{aligned} \tag{14}$$

By substitution (13) and (14) into (11), we obtain

$$V_{m+1} = \mathbb{K}_{m+1} * (h_{m+1,m}^{-1}[\begin{pmatrix} 0 \\ S_{.,m} \end{pmatrix} - \begin{pmatrix} S_m H_{.,m} \\ 0 \end{pmatrix}]) \tag{15}$$

Since $\mathcal{V}_{m+1} = \mathbb{K}_{m+1} * S_{m+1}$, then

$$V_{m+1} = \mathbb{K}_{m+1} * S_{.,m+1}. \tag{16}$$

From relations (15), (16) and linearly independent columns of the $\mathbb{K}_{m+1}$, we conclude

$$S_{.,m+1} = h_{m+1,m}^{-1}\begin{pmatrix} 0 \\ S_{.,m} \end{pmatrix} - h_{m+1,m}^{-1}\begin{pmatrix} S_m H_{.,m} \\ 0 \end{pmatrix}. \tag{17}$$

In the sequel, consider the approximation solution X_m that by using BGMRES method is obtained

$$X_m = X_0 + \mathcal{V}_m * y_m.$$

Since $\mathcal{V}_m * y_m = K_m * (S_m y_m)$ and

$$\begin{aligned} R_m &= R_0 - AX_mB + X_m \\ &= R_0 - A(\mathcal{V}_m * y_m)B + \mathcal{V}_m * y_m \\ &= R_0 - (A\mathbb{K}_m(I_m \otimes B)) * S_m y_m + \mathbb{K}_m * S_m y_m \\ &= R_0 - (A\mathbb{K}_m(I_m \otimes B) - \mathbb{K}_m) * S_m y_m. \end{aligned} \tag{18}$$

Then by applying vec(.) operator on equation (18), we get

$$vec(R_m) = \tag{19}$$

$$\begin{aligned} &(I_{ns} - (B^T \otimes A - I_{ns}) \sum_{j=0}^{m-1} a_j((B^T)^j \otimes A^j))vec(R_0) \\ &= P(B^T; A)vec(R_0) \end{aligned} \tag{20}$$

where

$$(a_0, \ldots, a_{m-1})^T = S_m y_m, \tag{21}$$

and the residual polynomial is as:

$$P_m(x,y) = 1 - (xy-1)Q_{m-1}(x,y), \tag{22}$$

where

$$Q_{m-1}(x,y) = \sum_{j=0}^{m-1} a_j x^j y^j. \tag{23}$$

From (22), we have $Q_{m-1}(x,y) = \frac{1-P_m(x,y)}{(xy-1)}$ and $(B^T \otimes A - I_{ns})^{-1} \approx Q_{m-1}(B^T; A)$. Therefore, we can apply $Q_{m-1}(x,y)$ as the preconditioner polynomial. We must solve the linear equation system:

$$Q_{m-1}(B^T; A)(B^T \otimes A - I_{ns})vec(X) = Q_{m-1}(B^T; A)vec(C), \tag{24}$$

by using BGMRES. The algorithm BGMRES, at the kth step, constructs the approximate solution X_k to the solution (24) such that

$$X_k = X_0 + Z_k \qquad s.t. \qquad Z_k \in \mathcal{GK}_k(A, R_0, B) \tag{25}$$

and with F-orthogonality relation

$$\begin{aligned} R_k &= \\ &\sum_{j=0}^{m-1} a_j A^j C B^j - \sum_{j=0}^{m-1} a_j(A^{j+1}X_kB^{j+1} - A^jX_kB^j) \\ &\perp_F A\mathcal{GK}_k(A, R_0, B)B. \end{aligned}$$

The above F-orthogonal relation is equivalent to the following relation:

$$\|R_k\|_F = min_{Z \in \mathcal{GK}_k(A,R_0,B)} \|R\|_F, \tag{26}$$

where

$$R = \sum_{j=0}^{m-1} a_j(A^jCB^j - (A^{j+1}(X_0+Z)B^{j+1} - A^j(X_0+Z)B^j)).$$

Consider F-orthonormal basis $\mathcal{V}_k$, which is constructed by using algorithm 1. After kth step of algorithm 1 to the matrices A and B with the nonzero residual matrix R_0, we can rewrite the relation (26) as follows:

$$\begin{aligned}
&\| \sum_{j=0}^{m-1} a_j A^j C B^j - \\
&\sum_{j=0}^{m-1} a_j (A^{j+1}(X_0+Z)B^{j+1} - A^j(X_0+Z)B^j)\|_F \\
&= \| \sum_{j=0}^{m-1} a_j A^j C B^j - \\
&\sum_{j=0}^{m-1} a_j (A^{j+1}(X_0+\mathcal{V}_k * y_k)B^{j+1} - A^j(X_0+\mathcal{V}_k * y_k)B^j)\|_F \\
&= \|R_0 - \sum_{j=0}^{m-1} a_j (A^{j+1}(\mathcal{V}_k * y_k)B^{j+1} - A^j(\mathcal{V}_k * y_k)B^j)\|_F \\
&= \|R_0 - \sum_{j=0}^{m-1} a_j (A^{j+1}\mathcal{V}_k(I_k \otimes B^{j+1}) - \\
&A^j\mathcal{V}_k(I_k \otimes B^j)) * y_k\|_F.
\end{aligned}$$

By * product properties and theorem 1, we get

$$\begin{aligned}
&\|R_0 - (-a_0\mathcal{V}_k + (a_0-a_1)A\mathcal{V}_k(I_k \otimes B) + \\
&(a_1-a_2)A^2\mathcal{V}_k(I_k \otimes B^2) + ... + \\
&(a_{m-2}-a_{m-1})A^{m-1}\mathcal{V}_k(I_k \otimes B^{m-1}) + \\
&a_{m-1}A^m\mathcal{V}_k(I_k \otimes B^m)) * y_k\|_F \\
&= \|R_0 - (-a_0\mathcal{V}_k + (a_0-a_1)\mathcal{V}_{k+1}(\bar{H}_k \otimes I_s) \\
&+ (a_1-a_2)\mathcal{V}_{k+2}(\bar{H}_{k+1}\bar{H}_k \otimes I_s) + ... + \\
&(a_{m-2}-a_{m-1})\mathcal{V}_{k+m-1}(\bar{H}_{k+m-2}...\bar{H}_k \otimes I_s) \\
&+ a_{m-1}\mathcal{V}_{k+m}(\bar{H}_{k+m-1}...\bar{H}_k \otimes I_s)) * y_k\|_F
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
&= \|\mathcal{V}_{m+k}(\|R_0\|_F(e_1 \otimes I_s) - (-a_0(\begin{pmatrix} I_k \\ 0_{m,k} \end{pmatrix} \otimes I_s) + \\
&(a_0-a_1)(\begin{pmatrix} \bar{H}_k \\ 0_{m-1,k} \end{pmatrix} \otimes I_s) \\
&+ (a_1-a_2)(\begin{pmatrix} \bar{H}_{k+1}\bar{H}_k \\ 0_{m-2,k} \end{pmatrix} \otimes I_s) + \cdots + \\
&(a_{m-2}-a_{m-1})(\begin{pmatrix} \bar{H}_{k+m-2}...\bar{H}_k \\ 0_{1,k} \end{pmatrix} \otimes I_s) \\
&+ a_{m-1}(\bar{H}_{k+m-1}...\bar{H}_k \otimes I_s)) * y_k\|_F \\
&= \|\|\|R_0\|_F e_1 - (-a_0\begin{pmatrix} I_k \\ 0_m \end{pmatrix} + \sum_{j=0}^{m-2}(a_j - a_{j+1})\begin{pmatrix} \bar{H}_{k+j}...\bar{H}_k \\ 0_{m-1-j,k} \end{pmatrix} \\
&+ a_{m-1}(\bar{H}_{k+m-1}...\bar{H}_k)y_k\|_2.
\end{aligned}$$

The above results, we can summarize in the following theorem:

Theorem 6. *After kth step of the algorithm BGMRES, the approximation solution* X_k *of* (24) *is given by* $X_k = X_0 + \mathcal{V}_k * y_k$, *where*

$$\begin{aligned}
y_k = argmin_{y\in\mathbb{R}^k}\|\|\|R_0\|_F e_1 - (-a_0\begin{pmatrix} I_k \\ 0_m \end{pmatrix} + \\
\sum_{j=0}^{m-2}(a_j - a_{j+1})\begin{pmatrix} \bar{H}_{k+j}...\bar{H}_k \\ 0_{m-1-j,k} \end{pmatrix} \qquad (27) \\
+ a_{m-1}\bar{H}_{k+m-1}...\bar{H}_k)y_k\|_2, \\
(28)
\end{aligned}$$

with $e_1 \in \mathbb{R}^{k+m}$.

The minimization problem (28) can be solved by the QR factorization the matrix

$$\begin{aligned}
&-a_0\begin{pmatrix} I_k \\ 0_m \end{pmatrix} + \sum_{j=0}^{m-2}(a_j - a_{j+1})\begin{pmatrix} \bar{H}_{k+j}...\bar{H}_k \\ 0_{m-1-j,k} \end{pmatrix} \\
&+ a_{m-1}\bar{H}_{k+m-1}...\bar{H}_k.
\end{aligned}$$

This matrix is transformed to an upper triangular matrix and then by solving the upper triangular system, we can obtain the vector y_k.
Using the above results, the polynomial preconditioner, i.e. PBGMRES algorithm based on the BGMRES algorithm is summarized in algorithm 3:

IV. Numerical Examples

In this section, we present some numerical examples to illustrate the potential of the new algorithm with polynomial preconditioner for the solution of the generalized linear Sylvester equation (1). In the following examples, we mainly evaluate and compare the performance of the new method against block GMRES with ILU and SSOR preconditioner [3], NSCG [19], squared Smith and restarted Krylov squared Smith [10], [16]. We use Matlab 2014a on a PC- Pentium(R), CPU 2.66GHz, 4.00 GB of RAM. We use the zero initial vector and stopping criterion $\|R_k\|_F < 1e-9$ for all the methods. In examples $4, 5, C = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}(1, \cdots, 1)$ for all the solvers. In Tables I and II, we note that Iter, Res.norm and cputime denote iteration number, residual norm and cputime, respectively.

Example 1. *In this example, we have tested the BGMRES, BGMRES with polynomial preconditioning, squared Smith (SM) and restarted Krylov squared Smith (RKSS) methods with* $m = 5, k = 10$ *on selected numerical example from [10], [16]. The convergence behavior of these methods are shown in Table* I .

From Table I, *we can see that the PBGMRES is faster than the other methods.*

Example 2. *([16]). We consider the continuous- time Lyapunov equation* $TX + XT^T = -EE^T$, *where* T *is the matrix* $TUB1000^3$ *of order n=1000 representing the Jacobian of a tabular rector model, and E is a one column vector such*

Algorithm 3 The PBGMRES algorithm

Require: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{s\times s}$, $X_0 \in \mathbb{R}^{n\times s}$, k, ε and degree m of $Q_{m-1}(x, y)$.

Ensure: X_k

Compute polynomial preconditioner $Q_{m-1}(x, y)$.
for $l = 1, \cdots$, until convergence **do**
  Compute$R_0 = C - AX_0B + X_0, \beta = \|R_0\|_F$, $V_1 = \frac{R_0}{\beta}, S_1 = (\frac{1}{\beta})$.
  for $j = 1, ..., m$ **do**
    $W = AV_j$;
    $W = WB$;
    for $i = 1, ..., j$ **do**
      $h_{ij} =< W, V_i >_F$;
      $W = W - h_{ij}V_i$;
    end for
    $h_{j+1,j} = \|W\|_F; if\ h_{j+1,j} = 0$ stop;
    $V_{j+1} = \frac{W}{h_{j+1,j}}$;
    set $H_{.,j} = \begin{bmatrix} h_{1j} \\ \vdots \\ h_{j,j} \end{bmatrix}, S_j = \begin{bmatrix} s_{11} & \cdots & s_{1j} \\ & \ddots & \vdots \\ & & s_{jj} \end{bmatrix}$;
    Compute $\begin{bmatrix} s_{1,j+1} \\ \vdots \\ s_{j+1,j+1} \end{bmatrix} =$

$$h_{j+1,j}^{-1} \begin{bmatrix} 0 \\ S_{.,j} \end{bmatrix} - h_{j+1,j}^{-1} \begin{bmatrix} S_j H_{.,j} \\ 0 \end{bmatrix};$$

  end for
  Solve the least squares problem

$$y_m = arg \min_{y\in\mathbb{R}^m} \| \|R_0\|_F e_1 - \begin{pmatrix} H_m - I_m \\ h_{m+1,m}e_m^T \end{pmatrix} y\|_2.$$

  Compute:$X_m = X_0 + \mathcal{V}_m * y_m$;
  Compute:$R_m = C - AX_mB + X_m$; if $||R_m||_F < \epsilon$, *stop*
  Compute polynomial preconditioner $Q_{m-1}(x, y)$:

$$S_m y_m = (a_0, \cdots, a_{m-1})^T,$$

$$Q_{m-1}(x, y) = \sum_{j=0}^{m-1} a_j x^j y^j.$$

  Compute the solution of

$$Q_{m-1}(B^T; A)(B^T \otimes A - I_{ns})vec(X) = Q_{m-1}(B^T; A)vec(C);$$

  by solving the least square problem (28) and $X_k = X_m + \mathcal{V}_k * y_k$;
  Set $X_0 = X_k$
end for

that $E(K) = 1/k, k = 1, 2, \cdots, n$. The Lyapunov equation is transformed to an equation $AXA^T - X = -BB^T$ with

$$A = (1 - T)^{-1}(1 + T), \qquad B = \sqrt{2}(1 - T)^{-1}(1 + T)E$$

see, e.g. [5]. Now, we compare BGMRES, PBGMRES, and

TABLE I
CONVERGENCE RESULTS FOR BGMRES, PBGMRES, SM, RKSS, WITH $m = 5, n = 500, \alpha = 0.4$.

	BGMRES	NSCG	PBGMRES	SM	RKKS
Iter	Res.norm	Res.norm	Res.norm	Res.norm	Res.norm
1	1.4142	1.4142	1.4142	0.1280	0.1280
2	1.9626e- 05	0.0121	2.2823e- 16	0.0275	0.0275
3	1.7228e- 05	0.0485	5.2886e- 18	0.0208e- 01	0.0208e- 01
4	6.0121e- 09	0.0350	1.2740e- 18	0.0239e- 03	0.0239e- 03
5	4.0623e- 11	0.1088	2.6130e- 19	0.0722e- 07	0.0834e- 04
6	4.9543e- 12	0.0350	1.7657e- 19	0.0168e- 14	0.0295e- 04

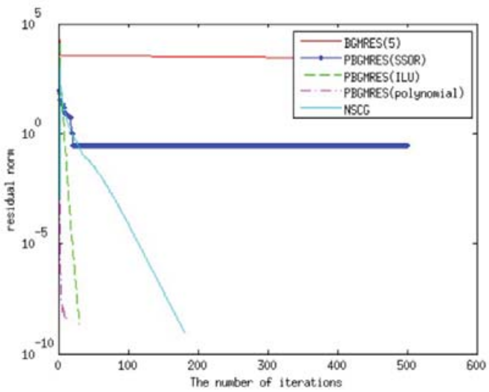
RKSS methods with $m = k = 10$ for solving equation $AXA^T - X = -BB^T$. The results are summarized in Table II.

TABLE II
CONVERGENCE RESULTS FOR BGMRES, PBGMRES, RKSS WITH $m = k = 10$.

Method	Res.norm	Iter.
PBGMRES(10)	0.00763e- 09	9
BGMRES(10)	0.0247	450
RKSS	0.0997e- 09	443

Example 3. *The purpose of this example is to illustrate the numerical behavior of BGMRES and BGMRES with ILU, SSOR and polynomial preconditioning and $m = 5$. The matrix $A \in \mathbb{R}^{n\times n}$ is a bidiagonal matrix with entries $2, 2, 3, 4, ..., 64$ on the main diagonal, and super diagonal entries 1. The matrix B is the same as A. Also, the right hand side of the generalized linear Sylvester equation $AXB - X = C$ is such that $X = 1$ is the exact solution. The numerical computations are carried out with $m = 6, k = 10$. The convergence curves plotted in Fig. 1. From Fig. 1, we can see that BGMRES with polynomial preconditioning is faster than the other methods.*

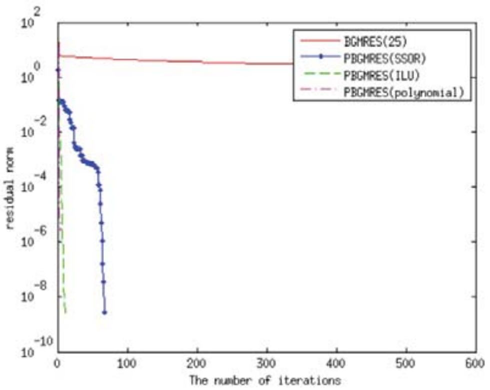
Figure 1 The convergence results of BGMRES with SSOR, ILU, polynomial BGMRES and NSCG methods with $m = 5$



Example 4. *In this example, we use the matrices $A = tridiag(1+d, 4, 1-d), B = A$ with $d = 5$ and $n = s = 64$ and $k = m = 25$. We evaluate the performance of the four block*

solvers. In Fig. 2, we show that BGMRES with polynomial preconditioning is faster convergence than the other methods.

Figure 2 The convergence result of BGMRES with SSOR, ILU, polynomial preconditioning and BGMRES methods



Example 5. *The matrices are the same as in Example 4, but d=8, m=k=10. We compare the convergence behavior of BGMRES, PBGMRES, PBGMRES (ILU, SSOR). The results are summarized in Table III.*

TABLE III
CONVERGENCE RESULTS FOR BGMRES, PBGMRES, PBGMRES (ILU, SSOR)

Method	Iter.	Res.norm	Cputime(seconds)
BGMRES	5000	0.8503	1.438e+03
PBGMRES(polynomial)	7	3.8132e- 10	14.235
PBGMRES(ILU)	19	5.6427e- 10	2.5800
PBGMRES(SSOR)	88	1.1340e- 10	1.789e+3
NSCG	218	NaN	-

V. CONCLUSION

We have derived a polynomial preconditioning block GMRES method for the generalized Sylvester matrix equation. It is observed by examples that PBGMRES is faster than some other block solvers.

APPENDIX A
PROOF OF THEOREM 2

Since

$$\mathcal{A}r_0 = (B^T \otimes A)r_0 - r_0.$$

Thus $\mathcal{A}r_0$ is a combination of $(B^T \otimes A)r_0$ and r_0 and

$$span\{r_0, \mathcal{A}r_0\} = span\{r_0, (B^T \otimes A)r_0\}.$$

Next, we consider $\mathcal{A}^2 r_0$,

$$\mathcal{A}^2 r_0 = (B^T \otimes A)^2 r_0 - 2(B^T \otimes A)r_0 + r_0.$$

Therefore,

$$span\{r_0, \mathcal{A}r_0, \mathcal{A}^2 r_0\} = span\{r_0, (B^T \otimes A)r_0, (B^T \otimes A)^2 r_0\}.$$

Continuing this, the two subspaces are the same. This establishes the claim.

APPENDIX B
PROOF OF THEOREM 3

The proof of theorem 3 proceeds as:
$\min_{X \in X_0 + \mathcal{GK}_m(A,R_0,B)} ||R||_F$

$$= \min_{X \in X_0 + \mathcal{GK}_m(A,R_0,B)} ||C - AXB + X||_F.$$

Let X^* be the exact solution of the matrix equation (1), therefore

$$= \min_{X \in X_0 + \mathcal{GK}_m(A,R_0,B)} ||A(X^* - X)B - (X^* - X)||_F$$

since $X \in X_0 + \mathcal{GK}_m(A, R_0, B)$, there exists a $Z \in \mathcal{GK}_m(A, R_0, B)$ such that $X = X_0 + Z$ hence

$$= \min_{Z \in \mathcal{GK}_m(A,R_0,B)} ||A(Z^* - Z)B - (Z^* - Z)||_F$$

by (4), we have

$$= \min_{Y \in A\mathcal{GK}_m(A,R_0,B)B} ||Y^* - Y||_F,$$

where $Y = AZB - Z$. By (corollary 1.39, [14]),

$$\min_{Y \in A\mathcal{GK}_m(A,R_0,B)B} ||Y^* - Y||_F = ||Y^* - Y_m||_F$$

if and only if

$$\begin{cases} Y_m \in A\mathcal{GK}_m(A, R_0, B)B, \\ Y^* - Y_m \perp_F A\mathcal{GK}_m(A, R_0, B)B. \end{cases}$$

Since $Y^* - Y_m = R_m$, then $R_m = C - AX_mB + X_m \perp_F A\mathcal{GK}_m(A, R_0, B)B$.

REFERENCES

[1] L. Bao, Y. Lin and Y. Wei, "Krylov subspace methods for the generalized Sylvester equation", *Appl. Mathem. Comput.* vol. 175, 2006, pp. 557–573.
[2] D. S. Bernstein and W. M. Haddad, "LQG control with a $H_{\inf}$ performance bound: a Riccati equation approach", *IEEE Trans. Automat. Control*, vol. AC-34, 1989, pp. 293-305.
[3] A. Bouhamidi and K. Jbilou, "A note on the numerical approximate solutions for generalized sylvester matrix equations with applications", *Appl. Math. Comput.*, vol. 206, 2008, pp. 687–694.
[4] J. W. Demmel, Applied numerical linear algebra; *SIAM*, 1997.
[5] Z. Gajic and M. T. J. Qureshi, Lyapunov Matrix Equation in System Stability and Control; *Dover, Mineola, NY*, 2008.
[6] G. H. Golub, S. Nash and C. F. Vanloan , "A Hessenberg- Schur method for problem $AX + XB = C$", *IEEE Trans. Automat. Contr.*, vol. 24, 1979, pp. 909–913.
[7] G. H. Golub, S. Nash and C. F. Vanloan, Matrix Computations, *Jhons Hoplins U. P., Baltimore, 3th edn*, 1996.
[8] K. Jbilou, A. Messaoudi and A. Sadok, "Global FOM and GMRES algorithms for matrix equation", *Appl. Numer. Math.*, vol. 31, 1999, pp. 49–63.
[9] D. Khojasteh Salkuyeh and F. Toutounian, "New approaches for solving large sparse Sylvester equations", *Appl. Math. Comput.*, vol. 173, no. 1, 2006, pp. 9–18.
[10] T. Li, P. Chang-Yi Weng,E. King- Wash Chu and W. W. Lin, "Large-Scale Stein and Lyapunov equations, Smith methods and applications", *Numer. Algor*, vol. 63, 2013, pp. 727–752.
[11] M. Mohseni Moghadam, A. Rivaz, A. Tajaddini and F. Saberi Mouvahed, "Convergence analysis of the global FOM and GMRES methods for solving matrix equations $AXB = C$ with spd coefficients", *Bull. Iranian Math. Soc*, vol. 41, no. 4, 2015, pp. 981–1001.
[12] B. C. Moore, "Principal component analysis in linear systems: controllability, observability, and model reduction", *IEEE Trans. Automat. Contr.*, vol. AC-26, 1981, pp. 17-32.
[13] F. Panjeh Ali Beik, "Note to the global GMRES for solving the matrix equation $AXB = C$", *Int. J. Eng. Nat. Sci.*, vol. 5, no. 2, 2003, pp. 101–105.

[14] Y. Saad, Iterative Methods for Sparse Linear Systems, *SIAM, Philadelphia*, 2003.
[15] J. Saak, Efficient Numerical Solution of Large Scale Algebraic Matrix Equations in PDE Control and Model Order Reduction, *Dr. Rer. Nat. dissertation, Chemnitz University of Technology, Germany*, 2009.
[16] M. Sadkane, "A low rank Krylov squared Smith method for large-scale discrete-time Lyapunov equations", *Linear Algebra Appl.*, vol. 436, 2012, pp. 2807–2827.
[17] J. L. Salle and Lefschetz, Stability by Lyapunovs Direct Method with Applications, *Academic press*, 1961.
[18] M. G. Sanfonov and R. Y. Chiang, A Schur method for balanced model reduction, *Proc. Amer. Control Conf., Atlanta, GA*, 1988.
[19] K. Yi-Fen and M. Chang- Feng, preconditioned nested splitting conjugate gradient iterative method for the large sparse generalized Sylvester equation, *Computers and Mathematics with Applications*, vol. 68, no. 10, 2014, pp. 1409–1420.