International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:5, No:2, 2011

Extending Global Full Orthogonalization method for Solving the Matrix Equation $AXB = F$

Fatemeh Panjeh Ali Beik

Abstract—In the present work, we propose a new method for solving the matrix equation $AXB = F$. The new method can be considered as a generalized form of the well-known global full orthogonalization method (Gl-FOM) for solving multiple linear systems. Hence, the method will be called extended Gl-FOM (EGl-FOM). For implementing EGl-FOM, generalized forms of block Krylov subspace and global Arnoldi process are presented. Finally, some numerical experiments are given to illustrate the efficiency of our new method.

Keywords—Matrix equations, Iterative methods, Block Krylov subspace methods.

I. Introduction

CONSIDER the multiple linear system

$$AX = C,$$

where $A \in \mathbb{R}^{n\times n}$ is a large and sparse nonsingular matrix, *C* and *X* are $n \times s$ rectangular real matrices.

For nonsymmetric problems, recently, some *block Krylov subspace methods* have been developed; see [1, 3, 5, 7-9, 12] and the references therein. The global full orthogonalization method (Gl-FOM) and its weighted version, for solving the multiple linear system $AX = C$, are projection methods on the *block Krylov subspace*

$$\mathcal{K}_m(A,V) = span\{V, AV, AV, ..., A^{m-1}V\},$$

where $V \in \mathbb{R}^{n\times s}$ is given.

Now, consider the following matrix equation

$$AXB = F, \tag{1}$$

where $A \in \mathbb{R}^{p\times n}$, is a full column-rank matrix and $B \in \mathbb{R}^{s\times q}$ is a full row-rank matrix.

Recently, there has been an increased interest in solving matrix equations; for more details see [2,4] and their references. In [2], Ding et al. proposed an iterative method for solving the matrix equation (1) by extending the well-known *Jacobi* and *Gauss-Seidel* methods.

Lemma I.1. *If A is a full column-rank matrix and B is a full row-rank matrix* ($p \geq n, s \leq q$)*, then in the sense of least-squares,* (1) *has the unique solution*

$$X = (A^TA)^{-1}A^TFB^T(BB^T)^{-1}.$$

Proof. See[2]. □

Fatemeh Panjeh Ali Beik is with the Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran e-mail: f.beik@vru.ac.ir.

Theorem I.2. *If the conditions of Lemma 1.1 hold, the gradient based iterative algorithm of* (1)*,*

$$X(k) = X(k-1) + \mu A^T[F - AX(k-1)B]B^T,$$
$$0 < \mu < \frac{2}{\lambda_{\max}[AA^T]\lambda_{\max}[BB^T]} \quad \text{or} \quad \mu \leq \frac{2}{||A||^2\,||B||^2},$$

yields $X(k) \to X$**.**

proof. See[2]. □

It is obvious that finding a proper μ by the conditions described in Theorem 1.2, is too expensive. On the other hand, in application, the value of μ approximated by Theorem 1.2 may become too small, hence the algorithm may become divergent.

It is known that the global full orthogonalization method (Gl-FOM) is suitable for solving multiple linear systems with large coefficient matrix. Hence, we are interested to present a new iterative method, by extending Gl-FOM, for solving the matrix equation (1). To this end, we need to generalize the definition of the block Krylov subspace. On the other hand, it is obvious that each system of the form (1) can be reformed as $(A^TA)X(BB^T) = A^TFB^T$**.** Hence, without loss of generality, we will consider the following matrix equation

$$AXB = F, \tag{2}$$

where $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{s\times s}$ are *nonsingular* matrices and $X, F \in \mathbb{R}^{n\times s}$**.**

For two matrices *Y* and *Z* in $\mathbb{R}^{n\times s}$, we define the inner product $< Y, Z >_F = tr(Y^TZ)$, where $tr(Y^TZ)$ denotes the trace of the matrix Y^TZ**.** The associated norm is the well-known *Frobenius* norm denoted by $||.||_F$ **.** A system of vectors (matrices) of $\mathbb{R}^{n\times s}$ is said to be *F-orthonormal* if it is orthonormal with respect to $< ., . >_F$.

For a matrix $V \in \mathbb{R}^{n\times s}$, we denote by $vec(V)$ the vector of $\mathbb{R}^{ns}$ defined by

$$vec(V) = \left[v(.,1)^T, \quad v(.,2)^T, ..., v(.,s)^T\right]^T$$

where $v(.,j), \quad j = 1, ..., s$, is the *j*-th column of *V*.

For given matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{k\times l}$, 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, $A\otimes B = [a_{ij}B]$**.**(for more details see [6])

Definition I.3. (R. Bouyouli et al.[1]). Let $A = [A_1, A_2, ..., A_p]$ and $B = [B_1, B_2, ..., B_\ell]$ be 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 $A^T\diamondsuit B$ is defined by $[A^T\diamondsuit B]_{i,j} =< A_1, B_1 >_F, \;\; i = 1, 2, ..., p, \;\; j = 1, 2..., \ell.$

The outline of the paper is organized as follows. In Section 2, a generalized form of the global Arnoldi process is presented for implementing our new method. A new method called the extended global full orthogonalization method (EGl-FOM), for solving the matrix equation (2), is proposed in Section 3. To illustrate the efficiency of our new method in comparison with the method presented in [2], some numerical experiments are presented in Section 4. Finally, the paper is ended with a brief conclusion in Section 5.

II. Generalized global Arnoldi process

We can easily see that the matrix equation (2) is equivalent to the following linear system of equations

$$(B^T \otimes A)vec(X) = vec(F).$$

However, the size of the linear equations $(B^T \otimes A)vec(X) = vec(F)$ is too large and the block Krylov subspace methods consume more computer time and memory once the size of the system is large. To overcome these complications and drawbacks, by extending global full orthogonalization method (Gl-FOM), we propose an extended global full orthogonalization method (EGl-FOM) for solving the matrix equation (2). To this end, we need to generalize the definition of the block Krylov subspace in the following.

Definition II.1. Suppose that $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{s\times s}$, and $V \in \mathbb{R}^{n\times s}$, we define the *generalized block Krylov subspace* as follows

$\mathcal{GK}_m \equiv \mathcal{GK}_m(A,V,B) \equiv$

$$span\{V, AVB, A^2VB^2, ..., A^{m-1}VB^{m-1}\}. \quad (3)$$

In the following, a generalized form of the global Arnoldi process is presented for constructing an F-orthonormal basis for the $\mathcal{GK}_m$.

Algorithm II.2. *(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. $W = AV_j$
4. $W = WB$
5. For $i = 1, 2, ..., j$ Do:
6. $h_{ij} =< W, V_i >_F$
7. $W = W - h_{ij}V_i$
8. EndDo
9. $h_{j+1,j} = \|W\|_F$. If $h_{j+1,j} = 0$, then stop.
10. $V_{j+1} = W/h_{j+1,j}$
11. EndDo.

Denote by $\mathcal{V}_m$, the $n \times ms$ matrix with columns $V_1, V_2, ..., V_m$, $\overline{H}_m$, the $(m+1)\times m$ Hessenberg matrix whose nonzero entries $h_{ij}, i = 1, 2, ..., m+1, j = 1, ...m$, are defined by Algorithm 2.2, and by H_m, the matrix obtained from $\overline{H}_m$ by deleting its last row.

It is obvious that, the generalized global Arnoldi process constructs an F-orthonormal basis $V_1, V_2, ..., V_m$ for the matrix block Krylov $\mathcal{GK}_m(A,V,B)$, i.e., the matrices $V_1, V_2, ..., V_m$ satisfy the following conditions

$$tr(V_i^T V_j) = 0, \quad tr(V_i^T V_i) = 1, \text{ for } i \neq j, \quad i,j = 1, 2, ..., m. \quad (4)$$

Theorem II.3. *Let $\mathcal{V}_m$, H_m, and $\overline{H}_m$ be defined as before. Then the following relations hold*

$$A\mathcal{V}_m(I_m \otimes B) = \mathcal{V}_m(H_m \otimes I_s) + [0_{n\times s}, ..., 0_{n\times s}, h_{m+1,m}V_{m+1}], \quad (5)$$

$$A\mathcal{V}_m(I_m \otimes B) = \mathcal{V}_m(H_m \otimes I_s) + h_{m+1,m}V_{m+1}(e_m^T \otimes I_s) = \mathcal{V}_{m+1}(\overline{H}_m \otimes I_s), \quad (6)$$

where $e_m^T = \begin{bmatrix} 0, & ..., & 0, & 1 \end{bmatrix}_{1\times m}$.

Proof. From lines 3, 4 and 7 of Algorithm 2.2, we deduce that

$$AV_jB = \sum_{i=1}^{j+1} h_{ij}V_i, \qquad j = 1, 2, ..., i, \quad i = 1, 2, ..., m.$$

The above relation is equivalent to (5). The relation (6) is a reformulation of (5). □

III. Extended global full orthogonalization method (EGl-FOM)

In this section, we present our new method, EGl-FOM for solving the matrix equation (2).

Given an initial guess X_0, with the corresponding residual $R_0 = F - AX_0B$, the EGl-FOM constructs the new approximate solution X_m to the solution of (2) such that

$$X_m \in X_0 + \mathcal{GK}_m(A, R_0, B), \quad (7)$$

and,

$$R_m = F - AX_mB \perp_F \mathcal{GK}_m(A, R_0, B). \quad (8)$$

Consider the F-orthonormal basis $\mathcal{V}_m$, constructed with generalized global Arnoldi process. From the relation (7), we deduce that

$$X_m = X_0 + \mathcal{V}_m(y_m \otimes I_s), \quad (9)$$

where the vector $y_m \in \mathbb{R}^m$ is obtained by imposing the orthogonality condition (8). By substituting (9) in R_m, we get

$$R_m = F - AX_mB = F - A(X_0 + \mathcal{V}_m(y_m \otimes I_s))B$$
$$= R_0 - A\mathcal{V}_m(I_m \otimes B)(y_m \otimes I_s).$$

By imposing the orthogonality condition we conclude that

$0 = \mathcal{V}_m^T \diamondsuit R_m$

$= \mathcal{V}_m^T \diamondsuit (\mathcal{V}_m((\beta e_1 - H_m y_m) \otimes I_s))$

$$- h_{m+1,m}\mathcal{V}_m^T \diamondsuit (V_{m+1}(e_m^T y_m \otimes I_s))$$

$$= (\mathcal{V}_m^T \diamondsuit \mathcal{V}_m)(\beta e_1 - H_m y_m) - h_{m+1,m}(\mathcal{V}_m^T \diamondsuit V_{m+1})(e_m^T y_m).$$

On the other hand $\mathcal{V}_m^T \diamondsuit \mathcal{V}_m = I$ and $\mathcal{V}_m^T \diamondsuit V_{m+1} = 0$. Therefore, the vector $y_m \in \mathbb{R}^m$ is the solution of the following linear system

$$H_m y_m = \beta e_1.$$

Now, we propose the EGl-FOM algorithm for solving the matrix equation (2) as follows.

Algorithm III.1. *(EGl-FOM)*
1. Choose X_0, a tolerance ε, compute $R_0 = B - AX_0$ and set $V = R_0$.
2. For $m = 1,\ 2,\ 3, ...$ Do:

3. Construct the F-orthonormal basis $V_1, V_2, ..., V_m$ by Algorithm 2.2.
4. Find y_m as the solution of $H_m y_m = \beta e_1$.
5. Compute the approximate solution $X_m = X_0 + \mathcal{V}_m(y_m \otimes I_s)$, and

$$R_m = F - AX_mB.$$

6. If $||R_m||_F < \varepsilon$, then stop.
7. Set $X_0 = X_m$, $R_0 = R_m$, $V = R_0$, and go to 3.
8. EndDo.

The EGl-FOM 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 matrices. To cure the storage problem, encountered also in Gl-FOM, the value of m is limited by storage constraint and by avoiding rounding errors; for more details, see [11]. Hence, Algorithm 3.1 can be restarted after m iterations. The corresponding algorithm is called the restarted EGl-FOM (m).

Lemma III.2. *Suppose that $X_m \in X_0 + \mathcal{GK}_m(A, R_0, B)$, is the approximate solution computed by the EGl-FOM algorithm with the corresponding residual matrix R_m. Then*

$$R_m = -h_{m+1,m} V_{m+1}(e_m^T y_m \otimes I_s).$$

Therefore, $||R_m||_F = h_{m+1,m}\left|y_m^{(m)}\right|$ where $y_m^{(m)}$is the last component of the vector y_m.

Proof. It is easy to see that, $R_m = R_0 - A\mathcal{V}_m(y_m \otimes I_s)B$. By some easy computation, we get $R_m = R_0 - \mathcal{V}_m(H_m y_m \otimes I_s) - h_{m+1,m} V_{m+1}(e_m^T y_m \otimes I_s)$, therefore, $R_m = \beta V_1 - \mathcal{V}_m(\beta e_1 \otimes I_s) - h_{m+1,m} V_{m+1}(e_m^T y_m \otimes I_s)$. On the other hand, it is clear that $\mathcal{V}_m(\beta e_1 \otimes I_s) = \beta V_1$. Hence, the result follows immediately. □

Lemma III.3. *Suppose that $X_m \in X_0 + \mathcal{GK}_m(A, R_0, B)$, is the approximate solution computed by the EGl-FOM algorithm with the corresponding residual matrix R_m. Then*

$$||R_m||_F = ||R_0||_F \frac{\left|\prod_{i=2}^{m+1} h_{i,i-1}\right|}{\det(H_m)}$$

and $\frac{||R_m||_F}{||R_{m-1}||_F} = \frac{h_{m+1,m}}{h_{m,m}} \frac{\det(H_{m-1})}{\det(H_m)}$.

Proof. From the previous lemma, we have $||R_m||_F = h_{m+1,m}\left|y_m^{(m)}\right|$. On the other hand y_m is the solution of the linear system $H_m y_m = \beta e_1$. Hence, the results follow by computing $y_m^{(m)}$ by the Cramer rule. □

IV. Numerical Experiments

In this section, we give an example to illustrate the results presented in this paper. Also, we will compare our new method with the method given in [2]. For simplicity we called the method, proposed in [2], as Ding's Method. The experiments were performed by Mathematica 6. The initial guess X_0 was chosen such that $X_0 = 0$ and the tests were stopped as soon as

$$||R_m||_F = ||F - AX_mB||_F \leq 0.5 \times 10^{-6}.$$

Example IV.1. Let the matrix A be a $n \times n$ matrix defined as follows:

$$\begin{pmatrix}
1 & 0.21 & 1.2 & 0 & 0.13 & 1.42 & & & & \\
0.45 & 2 & 0.21 & 1.2 & 0 & 0.13 & 1.42 & & & \\
0 & 0.45 & 3 & 0.21 & 1.2 & 0 & 0.13 & & & \\
0.12 & 0 & 0.45 & 4 & 0.21 & & & & & \\
0.11 & 0.12 & 0 & 0.45 & \ddots & \ddots & & & & 1.42 \\
 & 0.11 & 0.12 & 0 & \ddots & \ddots & \ddots & & & 0.13 \\
 & & & & & \ddots & \ddots & \ddots & & 0 \\
 & & & & & & \ddots & \ddots & \ddots & 1.2 \\
 & & & & & & & \ddots & \ddots & 0.21 \\
 & & & & & 0.11 & 0.12 & 0 & 0.45 & n
\end{pmatrix}_{n\times n}.$$

The $s \times s$ matrix B is defined by $B = tridiag\{-1, 4, -1\}$. The $n \times s$, matrix F is chosen such that $X = [X_{ij}]_{n\times s}$ be the solution of the matrix equation (2), where nonzero elements of X are $X_{ii} = 1, i = 1, 2, ..., min(n, s)$.

The results of performing EGl-FOM(2) in terms of both number of restarts and CPU consuming time in seconds are prented in the following table.

n	s	iteration	CPU- time(s)
200	30	292	17.16
400	50	459	403.794

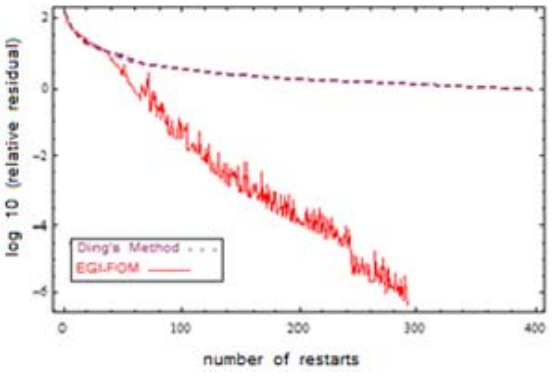

Fig. 1: n = 200, s = 30

The numerical comparison results for Ding's Method and EGl-FOM are given in the form of Figures 1 and 2. The relation between the number of restarts as the *x*-axis and the relative residual's logarithm (in base 10) as the *y*-axis. As it is seen the Ding's Method is not convergent for this example, although we have chosen a proper μ which satisfies in the conditions of Theorem 1.2.

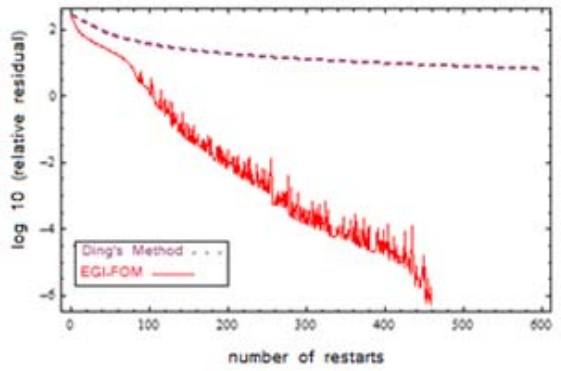

Fig. 2: n = 400, s = 50

V. Conclusion and further works

We introduced a generalized forms for the block Krylov subspace and global Arnoldi process. Then, an extended global full orthogonalization method (EGl-FOM) was presented for solving the matrix equation $AXB = F$. The weighted version of the EGl-FOM can be presented in a similar way which have proposed by the authors in [8]. Extending global generalized minimum residual (Gl-GMRES) method [5] and its weighted version [9] for solving the matrix equation $AXB = F$ with its convergent properties is under investigation. In a similar way discussed in [10], the EGl-FOM can be applied for solving the shifted matrix equation $(A - \sigma I)X(B - \sigma I) = F$.

References

[1] R. Bouyouli, K. Jbilou, A. Messaoudi and H. Sadok, On block minimal residual methods, Appl. Math. Lett, 20 (2007), 284–289.
[2] F. Ding, P. X. Liu and J. Ding, Iterative solutions of the generalized Sylvester matrix equation by using hierarchical identification principle, Appl. Math. Comput, 197 (2008), 41–50.
[3] A. El Guennouni, K. Jbilou and H. Sadok, A block version of BiCGSTAB for linear systems with multiple right-hand sides, Trans. Numer. Anal, 16 (2003), 129–142.
[4] G. X. Huang, F. Yin, K. Guo, An iterative method for the skew-symmetric solution and the optimal approximate solution of the matrix equation *AXB =C*, J. Comput. Appl. Math, 212 (2008), 231–244.
[5] K. Jbilou, A. Messaoudi and H. Sadok, Global FOM and GMRES algorithms for matrix equations, Appl. Numer. Math, 31 (1999), 43–49.
[6] P. Lancaster. Theory of Matrix. Academic Press, New York, 1969.
[7] Y.-Q. Lin, Implicitly restarted global FOM and GMRES for nonsymmetric matrix equations and Sylvester equations, Appl. Math. Comput, 167 (2005), 1004–1025.
[8] M. Mohseni Moghadam and F. Panjeh Ali Beik, A new weighted global full orthogonalization method for solving nonsymmetric linear systems with multiple right-hand sides, Int. Elect. J. Pure Appl. Math, 2(2)(2010), 47–67.
[9] M. Mohseni Moghadam and F. Panjeh Ali Beik, On a new weighted global GMRES for solving nonsymmetric linear system with multiple right-hand sides, Int. Journal of contemp. Math. Sciences, 5 (2010), 2237–2255.
[10] M. Mohseni Moghadam and F. Panjeh Ali Beik, A new weighted global full orthogonalization method for shifted linear system with multiple right-hand sides, Int. Math. Forum, 5 (2010), 2857–2874.
[11] Y. Saad, Iterative Methods for Sparse Linear System, PWS Press, New York, 1995.
[12] M. Sadkane, Block Arnoldi and Davidson methods for unsymmetrical large eigenvalue problems, Numer. Math, 64 (1993), 687–706.