Available online at www.sciencedirect.com
ScienceDirect
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
Journal of Computational and Applied Mathematics 219 (2008) 350–358
www.elsevier.com/locate/cam

New convergence results on the global GMRES method for diagonalizable matrices

M. Bellalij[a], K. Jbilou[b], H. Sadok[b,*]

[a] Université de Valenciennes, Le Mont Houy, 59313 Valenciennes Cedex, France
[b] Université du Littoral, zone universitaire de la Mi-voix, batiment H. Poincarré, 50 rue F. Buisson, BP 699, F-62280 Calais Cedex, France

Received 11 November 2006; received in revised form 6 July 2007

Dedicated to Claude Brezinski

Abstract

In the present paper, we give some new convergence results of the global GMRES method for multiple linear systems. In the case where the coefficient matrix A is diagonalizable, we derive new upper bounds for the Frobenius norm of the residual. We also consider the case of normal matrices and we propose new expressions for the norm of the residual.

MSC: 65F10

Keywords: Convergence; Global GMRES; Iterative methods; Matrix Krylov subspace; Multiple linear systems

1. Introduction

We consider the following multiple linear system:

$$AX = B, \tag{1}$$

where A is an $n \times n$ real large and sparse matrix, B and X are $n \times s$ rectangular matrices with $s \ll n$.

For nonsymmetric problems, some block Krylov subspace methods have been developed these last years; see [3,9,12,14] and the references therein.

In [9], we introduced a global approach for solving (1) and derived the global Arnoldi and the global GMRES methods. They are generalizations of the global MR method proposed by Saad [12, p. 300] for approximating the inverse of a matrix. These global methods are also effective, as compared to block Krylov subspace methods, when applied for solving large and sparse low rank right-hand sides Lyapunov and Sylvester matrix [10,11,16]. Other applications of the global Arnoldi and global Lanczos methods in control theory, model reduction and quadratic matrix equations are given in [5,6,8,17].

In the present paper we give some new convergence results for the global GMRES method when applied to the multiple linear system (1); numerical tests and comparison with other block methods are given in [9,11]. Our approach

* Corresponding author.
E-mail addresses: Mohammed.Bellalij@univ-valenciennes.fr (M. Bellalij), jbilou@lmpa.univ-littoral.fr (K. Jbilou), Sadok@lmpa.univ-littoral.fr (H. Sadok).

0377-0427/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2007.09.016

ignores the way global GMRES is implemented and uses the fact that the Frobenius norm of the residual is minimized at each iteration.

The paper is organized as follows. In Section 2, we recall some properties of the $\diamond$ matrix product introduced in [4]. In Section 3, we review the global GMRES method for multiple linear systems. Section 4 is devoted to the convergence analysis of the global GMRES method. We give new convergence results for the case A diagonalizable and also the particular case of normal matrices.

We use the following notation. For two matrices Y and Z in $\mathbb{R}^{n\times s}$, we define the inner product $\langle Y, Z\rangle_F = \mathrm{trace}(Y^T Z)$, the associated norm is the Frobenius norm denoted by $\|.\|_F$. The 2-norm of a matrix X is denoted by $\|X\|_2$. A system of vectors (matrices) of $\mathbb{R}^{n\times s}$ is said to be F-orthonormal if it is orthonormal with respect to the scalar product $\langle ., .\rangle_F$. The Kronecker product of the matrices C and D is given by $C \otimes D = [c_{i,j}D]$. Finally, if X is an $n \times s$ matrix, the $x = \mathrm{vec}(X)$ is the ns vector obtained by stacking the s columns of the matrix X.

2. Definitions and properties

2.1. The $\diamond$ product

In the following we recall the product denoted by $\diamond$ and defined as follows [4]:

Definition 1. Let $A = [A_1, A_2, \ldots, A_p]$ and $B = [B_1, B_2, \ldots, B_l]$ be matrices of dimension $n \times ps$ and $n \times ls$, respectively, where A_i and B_j ($i = 1, \ldots, p$; $j = 1, \ldots, l$) are $n \times s$ matrices. Then the $p \times l$ matrix $A^T \diamond B$ is defined by

$$A^T \diamond B = \begin{pmatrix} \langle A_1, B_1\rangle_F & \langle A_1, B_2\rangle_F & \ldots & \langle A_1, B_l\rangle_F \\ \langle A_2, B_1\rangle_F & \langle A_2, B_2\rangle_F & \ldots & \langle A_2, B_l\rangle_F \\ \vdots & \vdots & \vdots & \vdots \\ \langle A_p, B_1\rangle_F & \langle A_p, B_2\rangle_F & \ldots & \langle A_p, B_l\rangle_F \end{pmatrix}.$$

Remarks. (1) If $s = 1$ then $A^T \diamond B = A^T B$.

(2) If $s = 1$, $p = 1$ and $l = 1$, then setting $A = u \in \mathbb{R}^n$ and $B = v \in \mathbb{R}^n$, we have $A^T \diamond B = u^T v \in \mathbb{R}$.

(3) The matrix $A = [A_1, A_2, \ldots, A_p]$ is F-orthonormal if and only if $A^T \diamond A = I_p$.

(4) If $X \in \mathbb{R}^{n\times s}$, then $X^T \diamond X = \|X\|_F^2$.

It is not difficult to show the following properties satisfied by the product $\diamond$.

Proposition 1. *Let $A, B, C \in \mathbb{R}^{n\times ps}$, $D \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{p\times p}$ and $\alpha \in \mathbb{R}$. 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. $(\alpha A)^T \diamond C = \alpha(A^T \diamond C)$.
4. $(A^T \diamond B)^T = B^T \diamond A$.
5. $(DA)^T \diamond B = A^T \diamond (D^T B)$.
6. $A^T \diamond (B(L \otimes I_s)) = (A^T \diamond B)L$.
7. $\|A^T \diamond B\|_F \leqslant \|A\|_F \|B\|_F$.

2.2. GMRES-type methods for multiple linear systems

The multiple linear system (1) could be solved by applying the classical one right-hand-side GMRES [13] method to the s linear systems separately. Starting from an initial block $X_0 = [X_0^{(1)}, X_0^{(2)}, \ldots, X_0^{(s)}]$ and the corresponding residual $R_0 = [R_0^{(1)}, \ldots, R_0^{(s)}]$, with $R_0^{(i)} = B^{(i)} - AX_0^{(i)}$, the kth approximation $X_k^{(i)}$, $i = 1, \ldots, s$ is such that $X_k^{(i)} - X_0^{(i)} \in K_k(A, R_0^{(i)})$ where $K_k(A, R_0^{(i)})$ is the ith classical Krylov subspace $K_k(A, R_0^{(i)}) = \mathrm{span}\{R_0^{(i)}, AR_0^{(i)}, \ldots, A^{k-1}R_0^{(i)}\}$. Another way of solving (1) is to consider all the s-second right-hand sides $B^{(i)}$, $i = 1, \ldots, s$, at the same time and this leads to the block GMRES [18]; see also [4,7,14] and the global GMRES [9] methods. For the block GMRES method, we consider the block Krylov subspace defined by

$$\mathbb{K}_k(A, R_0) = \mathrm{Range}([R_0, AR_0, \ldots, A^{k-1}R_0]) \subset \mathbb{R}^n. \tag{2}$$

Note that the block Krylov subspace $\mathbb{K}_k(A, R_0)$ is a sum of s classical Krylov subspaces

$$\mathbb{K}_k(A, R_0) = \sum_{i=1}^{s} K_k(A, R_0^{(i)}).$$

For the global GMRES method, we consider the matrix Krylov subspace $\mathbf{K}_k(A, R_0)$ defined as the subset of $\mathbb{R}^{n\times s}$ generated by the matrices $R_0, AR_0, \ldots, A^{k-1}R_0$.

For the three methods, the approximation X_k can be defined as follows:

$$X_k - X_0 = \sum_{i=1}^{k} A^{i-1}R_0\Omega_i, \tag{3}$$

where Ω_i, $i = 1, \ldots, k$ is an $s \times s$ coefficient matrix defined by

- $\Omega_i = \alpha_i I_s$, where α_i is a scalar coefficient defined from the orthogonality relation $R_k \perp_{\mathrm{F}} \mathbf{K}_k(A, AR_0)$ for the global GMRES where $R_k = B - AX_k$ is the residual corresponding to the approximate solution X_k.
- Ω_i is a full $s\times s$ matrix defined from $R_k^{(i)} \perp \mathbb{K}_k(A, AR_0)$, $i = 1, \ldots, s$; for the block GMRES. Note that the residual R_k can be given as

  $$R_k = \mathbb{P}_k(A) \circ R_0 = \sum_{i=0}^{k} A^i R_0 \Omega_i,$$

  where $\Omega_0 = I_s$ and $\mathbb{P}_k$ is the matrix-valued polynomial defined by $\mathbb{P}_k(t) = \sum_{i=1}^{k} t^i \Omega_i$. We note that a breakdown occurs in block GMRES, at step k, if the matrix $[R_0, AR_0, \ldots, A^k R_0]$ is rank deficient.
- $\Omega_i = \mathrm{diag}(\omega_1^{(i)}, \ldots, \omega_s^{(i)})$ is a diagonal matrix calculated from the orthogonality relations $R_k^{(i)} \perp K_k(A, AR_0^{(i)})$, $i = 1, \ldots, s$ for the classical GMRES.

When applying GMRES to the s right-hand side linear systems separately, it is well known [15] that

$$\|R_k^{(i)}\|_2^2 = \frac{1}{e_1^{\mathrm{T}}(K_{i,k+1}^{\mathrm{T}} K_{i,k+1})^{-1} e_1}, \quad i = 1, \ldots, s, \tag{4}$$

where $K_{i,k+1}$, $i = 1, \ldots, s$ is the Krylov matrix defined by $K_{i,k+1} = [R_0^{(i)}, AR_0^{(i)}, \ldots, A^k R_0^{(i)}]$ and $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^{k+1}$.

The relation (4) is important, it is the key for developing important convergence results for GMRES [2]. For the block GMRES method, we have the following expression of the norm of the residual [3]:

$$\|R_k\|^2 = \frac{\det\left(\begin{bmatrix} r_0^{\mathrm{T}} r_0 & r_0^{\mathrm{T}} B_k \\ B_k^{\mathrm{T}} r_0 & B_k^{\mathrm{T}} B_k \end{bmatrix}\right)}{\det(B_k^{\mathrm{T}} B_k)} \tag{5}$$

with $r_0 = \mathrm{vec}(R_0)$ and $B_k = I_s \otimes W_k$ where $W_k = [AR_0, \ldots, A^k R_0]$. Therefore, from the expression (5), it can be shown that

$$\|R_k\|^2 = \frac{1}{e_1^{\mathrm{T}}(\widetilde{K}_{k+1}^{\mathrm{T}} \widetilde{K}_{k+1})^{-1} e_1}, \tag{6}$$

where $\widetilde{K}_{k+1} = [r_0, B_k]$ and e_1 is the first vector of the canonical basis of $\mathbb{R}^{ks^2+1}$. Note that when $s = 1$, the block GMRES becomes the classical GMRES and the relation (6) reduces to (4). New convergence results based on the expression (6) are in progress for block GMRES method.

The convergence of a Krylov subspace method could be slow and usually, these methods are used with preconditioning techniques. With left and right preconditioners, we transform the original linear system to the new equivalent one

$$M_1^{-1}AM_2^{-1}Y = M_1^{-1}B \quad \text{and} \quad X = M_2^{-1}Y. \tag{7}$$

A good preconditioner is such that it should be easy to compute the solution of the linear systems $M_1Z_1 = W_1$, $M_2Z_2 = W_2$ for given W_1, W_2 and the matrix $M_1^{-1}AM_2^{-1}$ of the linear system (7) must be "close" to the identity. One of the widely used left-right preconditioner for sparse matrices is the incomplete LU factorization: ILU(τ) where τ is some dropping tolerance [12]. In [9], we applied the global GMRES method with the incomplete *LU* factorization to some numerical examples.

In what follows, we will be interested only in the global GMRES method. Our aim is to analyze the convergence of this method for some problems.

3. The global GMRES method

Let $\mathbf{K}_k(A, V) = \text{span}\{V, AV, \ldots, A^{k-1}V\}$ denotes the matrix Krylov subspace of $\mathbb{R}^{n\times s}$ spanned by the matrices $V, AV, \ldots, A^{k-1}V$ where V is an $n \times s$ matrix. Note that $Z \in \mathbf{K}_k(A, V)$ means that

$$\begin{aligned} Z &= \sum_{i=1}^{k} \alpha_i A^{i-1}V, \quad \alpha_i \in \mathbb{R}, \; i = 1, \ldots, k, \\ &= p_k(A)V, \end{aligned}$$

where p_k is the scalar polynomial defined by $p_k(t) = \sum_{i=1}^{k} \alpha_i t^{i-1}$.

Now consider the block linear system of Eq. (1) and let X_0 be an initial $n \times s$ matrix with the corresponding residual $R_0 = B - AX_0$.

The global GMRES method constructs, at step *k*, the approximation X_k satisfying the following two relations:

$$X_k - X_0 \in \mathbf{K}_k(A, R_0) \quad \text{and} \quad R_k \perp_{\mathrm{F}} \mathbf{K}_k(A, AR_0),$$

where the notation $\perp_{\mathrm{F}}$ means the orthogonality with respect to the scalar product $\langle Y, Z\rangle_{\mathrm{F}}$. From these two relations, we obtain

$$X_k = X_0 + \mathscr{K}_k(\alpha \otimes I_s)$$

and

$$R_k = R_0 - \mathscr{W}_k(\alpha \otimes I_s),$$

where $\mathscr{K}_k = [R_0, AR_0, \ldots, A^{k-1}R_0]$, $\mathscr{W}_k = A\mathscr{K}_k$ and α is such that

$$(\mathscr{W}_k^{\mathrm{T}} \diamond \mathscr{W}_k)\alpha = \mathscr{W}_k^{\mathrm{T}} \diamond R_0.$$

If $\mathscr{P}_k$ denotes the F-orthogonal projector onto the matrix Krylov subspace $\mathbf{K}_k(A, AR_0)$, then the residual R_k can be expressed as $R_k = R_0 - \mathscr{P}_k R_0$. As we are dealing with an orthogonal projection method onto the Krylov subspace $\mathbf{K}_k(A, AR_0)$, we have the minimization property

$$\|R_k\|_{\mathrm{F}} = \min_{Z \in \mathbf{K}_k(A, R_0)} \|R_0 - AZ\|_{\mathrm{F}}. \tag{8}$$

The problem (8) is solved by applying the global Arnoldi process [9], to get an F-orthonormal basis $\{V_1, \ldots, V_k\}$ of the matrix Krylov subspace $\mathbf{K}_k(A, R_0)$. Note that if $\mathscr{V}_k$ is the $n \times ks$ matrix $\mathscr{V}_k = [V_1, \ldots, V_k]$, then $\mathscr{V}_k^{\mathrm{T}} \diamond \mathscr{V}_k = I_k$. The minimization problem (8) is then equivalent to a small least-squares $(k+1) \times k$ problem. In this paper, we are interested in the convergence analysis of the global GMRES method and ignore how the method is implemented.

4. Convergence analysis of the global GMRES method

In this section, we give new convergence results for the global GMRES method. We will consider the case where A is diagonalizable and the case of normal matrices.

We first review some comparisons between the global GMRES for solving the multiple linear system (1) and the standard GMRES method [13] applied to each single linear system $AX^{(i)} = B^{(i)}$; $i = 1, \ldots, s$ where $B^{(i)}$ is the ith column of the $n \times s$ matrix B.

Theorem 1 (*Bouyouli et al. [4]*). *Let $K_{i,k+1}$, $i = 1, \ldots, s$ be the Krylov matrix defined by*

$$K_{i,k+1} = [R_0^{(i)}, AR_0^{(i)}, \ldots, A^k R_0^{(i)}] \quad \textit{with } R_0^{(i)} = B^{(i)} - AX^{(i)}, \quad i = 1, \ldots, s.$$

Then

$$\mathscr{K}_{k+1}^{\mathrm{T}} \diamond \mathscr{K}_{k+1} = \sum_{i=1}^{s} K_{i,k+1}^{\mathrm{T}} K_{i,k+1}, \tag{9}$$

where $\mathscr{K}_{k+1} = [R_0, AR_0, \ldots, A^k R_0]$.

It has been proved in [4] that when applying the global GMRES method to the multiple linear system (1), we obtain

$$\|R_k\|_{\mathrm{F}}^2 = \frac{1}{e_1^{\mathrm{T}} (\mathscr{K}_{k+1}^{\mathrm{T}} \diamond \mathscr{K}_{k+1})^{-1} e_1}, \tag{10}$$

where e_1 is the first unit vector of $\mathbb{R}^{k+1}$. The relation (10) is the key for developing new convergence results. We first give a result about the rate of the convergence of the global GMRES for general matrices.

Theorem 2 (*Bouyouli et al. [4]*). *Let R_k be the kth residual obtained by the global GMRES. Then*

$$\frac{\|R_k\|_{\mathrm{F}}}{\|R_{k-1}\|_{\mathrm{F}}} = \sqrt{1 - c_k^2},$$

with

$$c_k^2 = \frac{\det(\mathscr{K}_k^{\mathrm{T}} \diamond \mathscr{W}_k)^2}{\det(\mathscr{K}_k^{\mathrm{T}} \diamond \mathscr{K}_k) \det(\mathscr{W}_k^{\mathrm{T}} \diamond \mathscr{W}_k)},$$

and $\mathscr{W}_k = A\mathscr{K}_k$.

We consider now the case where the symmetric part A_{S} of the matrix A is positive definite.

Theorem 3. *If the symmetric part of the matrix A is positive definite, then*

$$\frac{\|R_k\|_{\mathrm{F}}}{\|R_{k-1}\|_{\mathrm{F}}} \leqslant \sqrt{1 - \frac{(\lambda_{\min}(A_{\mathrm{S}}))^2}{\|A\|_2^2}},$$

where $\lambda_{\min}(A_{\mathrm{S}})$ is the smallest eigenvalue of the matrix A_{S}.

The proof is easily obtained by using Theorem 2 and the first part of Theorem 6 of [4].

We shall consider the case where the matrix A is diagonalizable, and the particular case where A is a normal matrix.

4.1. Diagonalizable matrices

Let $A = XDX^{-1}$ where D is the diagonal matrix whose elements are the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the matrix A. In [4], the following upper bound for the norm of the residual R_k was given:

$$\|R_k\|_{\mathrm{F}} \leqslant \kappa_{\mathrm{F}}(X)\|R_0\|_{\mathrm{F}}\sqrt{n} \min_{p\in\mathscr{P}_k,\, p(0)=1}\left(\max_{i=1,\ldots,n}|p(\lambda_i)|\right), \tag{11}$$

where $\kappa_{\mathrm{F}}(X) = \|X\|_{\mathrm{F}}\|X^{-1}\|_{\mathrm{F}}$ and $\mathscr{P}_k$ is the set of polynomials of degree less or equal than k. We will give a more sharp upper bound for the norm of the residual R_k.

We first give the following lemma to be used later.

Lemma 1 (*Bellalij and Sadok [2]*). *Let E and F be two matrices of $\mathbb{C}^{n,k+1}$ $(k \leqslant n-1)$ and $\mathbb{C}^{n,n}$, respectively, and let $\sigma_1(F) \geqslant \sigma_2(F) \geqslant \cdots \geqslant \sigma_n(F) > 0$ be the singular values of F. If the matrix E is of full rank, then*

$$\frac{\sigma_n(F)^2}{e_1^{\mathrm{T}}(E^{\mathrm{H}}E)^{-1}e_1} \leqslant \frac{1}{e_1^{\mathrm{T}}(E^{\mathrm{H}}(F^{\mathrm{H}}F)E)^{-1}e_1} \leqslant \frac{\sigma_1(F)^2}{e_1^{\mathrm{T}}(E^{\mathrm{H}}E)^{-1}e_1},$$

where E^{H} is the conjugate transpose of the matrix E and e_1 is the first unit vector of $\mathbb{R}^{k+1}$.

Theorem 4. *Let the initial residual R_0 be decomposed as $R_0 = X\beta$ where β is an $n \times s$ matrix whose columns are denoted by $\beta^{(1)}, \ldots, \beta^{(s)}$. Let $R_k = B - AX_k$ be the kth residual obtained by applying global GMRES to* (1). *Then we have*

$$\|R_k\|_{\mathrm{F}}^2 \leqslant \frac{\|X\|_2^2}{e_1^{\mathrm{T}}(V_{k+1}^{\mathrm{H}}\tilde{D}V_{k+1})^{-1}e_1}, \tag{12}$$

where $\tilde{D}$ is the diagonal matrix defined by

$$\tilde{D} = \mathrm{diag}\left(\sum_{i=1}^{s}|\beta_1^{(i)}|, \ldots, \sum_{i=1}^{s}|\beta_n^{(i)}|\right)$$

and

$$V_{k+1} = \begin{pmatrix} 1 & \lambda_1 & \ldots & \lambda_1^k \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \ldots & \lambda_n^k \end{pmatrix}. \tag{13}$$

The coefficients $\beta_1^{(i)}, \ldots, \beta_n^{(i)}$ are the components of the vector $\beta^{(i)}$, $i = 1, \ldots, s$.

Proof. The ith matrix Krylov $K_{i,k+1}$ can be written as

$$K_{i,k+1} = [X\beta^{(i)}, XD\beta^{(i)}, \ldots, XD^k\beta^{(i)}], \quad i = 1, \ldots, s$$

which is decomposed as

$$K_{i,k+1} = XD_{\beta^{(i)}}V_{k+1}, \tag{14}$$

where $D_{\beta^{(i)}}$ is the diagonal matrix defined as

$$D_{\beta^{(i)}} = \begin{pmatrix} \beta_1^{(i)} & 0 & \ldots & 0 \\ 0 & \beta_2^{(i)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \beta_n^{(i)} \end{pmatrix}.$$

Using the relation (14) and the fact that $K_{i,k+1}$ is a real matrix, it follows that

$$K_{i,k+1}^{\mathrm{T}} K_{i,k+1} = K_{i,k+1}^{\mathrm{H}} K_{i,k+1} = V_{k+1}^{\mathrm{H}} D_{\beta^{(i)}}^{\mathrm{H}} X^{\mathrm{H}} X D_{\beta^{(i)}} V_{k+1}.$$

Invoking the relation (9), we obtain

$$\mathscr{K}_{k+1}^{\mathrm{T}} \diamond \mathscr{K}_{k+1} = \sum_{i=1}^{s} V_{k+1}^{\mathrm{H}} D_{\beta^{(i)}}^{\mathrm{H}} X^{\mathrm{H}} X D_{\beta^{(i)}} V_{k+1}. \tag{15}$$

On the other hand using Lemma 1, we get

$$e_1^{\mathrm{T}} (V_{k+1}^{\mathrm{H}} D_{\beta^{(i)}}^{\mathrm{H}} X^{\mathrm{H}} X D_{\beta^{(i)}} V_{k+1})^{-1} e_1 \leqslant \|X\|_2^{-2} e_1^{\mathrm{T}} (V_{k+1}^{\mathrm{H}} D_{\beta^{(i)}}^{\mathrm{H}} D_{\beta^{(i)}} V_{k+1})^{-1} e_1 \tag{16}$$

and then from the relations (15) and (16) and the fact that $\tilde{D} = \sum_{i=1}^{s} D_{\beta^{(i)}}^{\mathrm{H}} D_{\beta^{(i)}}$, we get

$$e_1^{\mathrm{T}} (\mathscr{K}_{k+1}^{\mathrm{T}} \diamond \mathscr{K}_{k+1})^{-1} e_1 \geqslant \|X\|_2^{-2} e_1^{\mathrm{T}} (V_{k+1}^{\mathrm{H}} \tilde{D} V_{k+1})^{-1} e_1. \tag{17}$$

Finally, the relations (10) and (17) imply

$$\|R_k\|_{\mathrm{F}}^2 \leqslant \frac{\|X\|_2^2}{e_1^{\mathrm{T}} (V_{k+1}^{\mathrm{H}} \tilde{D} V_{k+1})^{-1} e_1}$$

which shows the result. □

Let us set $\gamma_i = \sum_{l=1}^{s} |\beta_i^{(l)}|^2 / \|\beta\|_{\mathrm{F}}^2$, $i = 1, \ldots, n$ and let D_γ be the diagonal matrix defined by

$$D_\gamma = \begin{pmatrix} \gamma_1 & 0 & \ldots & 0 \\ 0 & \gamma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \gamma_n \end{pmatrix}. \tag{18}$$

Note that $\gamma_i \geqslant 0$; $i = 1, \ldots, n$, and $\sum_{i=1}^{n} \gamma_i = 1$.

In the following theorem, we give an upper bound for the norm of the residual.

Theorem 5. *Let $\mathscr{P}_k$ be the set of polynomials of degree less or equal than k, and let $\kappa_2(X) = \|X\|_2 \|X^{-1}\|_2$. Then we have the following results*:

$$\frac{\|R_k\|_{\mathrm{F}}^2}{\|\beta\|_{\mathrm{F}}^2} \leqslant \frac{\|X\|_2^2}{e_1^{\mathrm{T}} (V_{k+1}^{\mathrm{H}} D_\gamma V_{k+1})^{-1} e_1} \tag{19}$$

and

$$\frac{\|R_k\|_{\mathrm{F}}}{\|R_0\|_{\mathrm{F}}} \leqslant \kappa_2(X) \min_{p \in \mathscr{P}_k;\, p(0)=1} \left(\max_{\lambda \in \mathrm{Sp}(A)} |p(\lambda)| \right), \tag{20}$$

where $\mathrm{Sp}(A)$ *the set of eigenvalues of the matrix* A.

Proof. The relation (19) is obtained directly from (12) and the definition of D_γ. To show (20), we first use the fact that $\beta = X^{-1} R_0$ and then $\|\beta\|_{\mathrm{F}} = \|X^{-1} R_0\|_{\mathrm{F}} \leqslant \|X^{-1}\|_2 \|R_0\|_{\mathrm{F}}$. Therefore, replacing in (19), we get

$$\frac{\|R_k\|_{\mathrm{F}}^2}{\|R_0\|_{\mathrm{F}}^2} \leqslant \frac{\|X\|_2^2 \|X^{-1}\|_2^2}{e_1^{\mathrm{T}} (V_{k+1}^{\mathrm{H}} D_\gamma V_{k+1})^{-1} e_1}. \tag{21}$$

Finally, we invoke the following result already proved in [1]:

$$\frac{1}{e_1^{\mathrm{T}}(V_{k+1}^{\mathrm{H}} D_\gamma V_{k+1})^{-1} e_1} \leqslant \min_{p\in\mathscr{P}_k;\, p(0)=1} \left(\max_{\lambda\in\mathrm{Sp}(A)} |p(\lambda)| \right). \tag{22}$$

Hence, replacing (22) in (21), the result (20) follows. □

Note that since $\kappa_{\mathrm{F}}(X)$ is larger than $\kappa_2(X)$, the upper bound given by (20) is smaller than the one given by (11). In the case $s = 1$, the results of Theorem 3 reduce to those given in [2] for the classical GMRES.

4.2. Normal matrices

In this subsection, we assume that the matrix A is normal and we set $A = XDX^{\mathrm{H}}$ with $X^{\mathrm{H}}X = I$. We will give an expression, not only an upper bound, for the norm of the residual R_k.

Theorem 6. *Let $A = XDX^{\mathrm{T}}$ with $X^{\mathrm{T}}X = I$ and $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Consider the eigen-decomposition of the initial residual $R_0 = X\beta$. Then we have*

$$\frac{\|R_k\|_{\mathrm{F}}^2}{\|R_0\|_{\mathrm{F}}^2} = \frac{1}{e_1^{\mathrm{T}}(V_{k+1}^{\mathrm{H}} D_\gamma V_{k+1})^{-1} e_1} \tag{23}$$

and

$$\frac{\|R_k\|_{\mathrm{F}}^2}{\|R_0\|_{\mathrm{F}}^2} \leqslant \max_{\gamma\geqslant 0,\, \sum_{i=1}^n \gamma_i = 1} \left(\frac{1}{e_1^{\mathrm{T}}(V_{k+1}^{\mathrm{H}} D_\gamma V_{k+1})^{-1} e_1} \right), \tag{24}$$

where $\gamma = (\gamma_1, \ldots, \gamma_n)^{\mathrm{T}}$, V_{k+1} and D_γ are defined by (13) *and* (18), *respectively.*

Proof. Since $X^{\mathrm{H}}X = I$, the relation (15) becomes

$$\mathscr{K}_k^{\mathrm{T}} \diamond \mathscr{K}_k = \sum_{i=1}^{s} V_k^{\mathrm{H}} D_{\beta^{(i)}}^{\mathrm{H}} D_{\beta^{(i)}} V_k \tag{25}$$

which is expressed as

$$\mathscr{K}_k^{\mathrm{T}} \diamond \mathscr{K}_k = V_k^{\mathrm{H}} \tilde{D} V_k, \tag{26}$$

therefore

$$e_1^{\mathrm{T}}(\mathscr{K}_k^{\mathrm{T}} \diamond \mathscr{K}_k)^{-1} e_1 = e_1^{\mathrm{T}}(V_k^{\mathrm{T}} \tilde{D} V_k)^{-1} e_1. \tag{27}$$

Invoking (10) and the fact that $\|R_0\|_{\mathrm{F}} = \|\beta\|_{\mathrm{F}}$, we get

$$\frac{\|R_k\|_{\mathrm{F}}^2}{\|R_0\|_{\mathrm{F}}^2} = \frac{1}{e_1^{\mathrm{T}}(V_{k+1}^{\mathrm{H}} D_\gamma V_{k+1})^{-1} e_1}$$

which shows the result.

The relation (24) is directly derived from (23). □

We can state now the following main result.

Theorem 7. *Let A be a normal matrix and let R_k be the obtained residual, at step k, when applying the global GMRES method to the multiple linear system* (1). *Then we have*

$$\frac{\|R_k\|_{\mathrm{F}}}{\|R_0\|_{\mathrm{F}}} \leqslant \min_{p\in\mathscr{P}_k;\, p(0)=1} \left(\max_{\lambda\in\mathrm{Sp}(A)} |p(\lambda)| \right),$$

where $\mathscr{P}_k$ is the set of polynomials of degree $\leqslant k$ and $\mathrm{Sp}(A)$ is the set of eigenvalues of the matrix A.

Proof. The proof is derived from the relations (22) and (23). □

5. Conclusion

We developed in this paper new convergence results for the global GMRES method when applied to multiple linear systems. We considered the case when the matrix A is diagonalizable and the case of normal matrices. We showed that these new results are better than the ones given already in [9]. For $s = 1$, our results reduce to those obtained for the classical GMRES method.

Acknowledgment

We would like to thank the referees for their valuable comments and suggestions.

References

[1] M. Bellalij, Y. Saad, H. Sadok, Approximation and convex programming approaches for solving a Minmax problem, submitted for publication.
[2] M. Bellalij, H. Sadok, New tool for studying the GMRES convergence in exact arithmetic, submitted for publication.
[3] R. Bouyouli, K. Jbilou, A. Messaoudi, H. Sadok, On block minimal residual methods, Appl. Math. Lett. 20 (2007) 284–289.
[4] R. Bouyouli, K. Jbilou, R. Sadaka, H. Sadok, Convergence properties of some block Krylov subspace methods for multiple linear systems, J. Comput. Appl. Math. 196 (2006) 498–511.
[5] C.C. Chu, M.H. Lai, W.S. Feng, MIMO interconnects order reductions by using the multiple point adaptive-order rational global Arnoldi algorithm, IEICE Trans. Elect. E89-C (2006) 792–808.
[6] C.C. Chu, M.H. Lai, W.S. Feng, The multiple point global Lanczos method for mutiple-inputs multiple-outputs interconnect order reduction, IEICE Trans. Elect. E89-A (2006) 2706–2716.
[7] M.H. Gutknecht, Block Krylov space methods for linear systems with multiple right-hand sides: an introduction, February 2006; in: Y. Kaneda, H. Kawamura, M. Sasai (Eds.), Frontiers of Computational Science—Proceedings of the International Symposium on Frontiers of Computational Science 2005, Springer, Berlin, Heidelberg, March 2007.
[8] M. Heyouni, K. Jbilou, Matrix Krylov subspace methods for large scale model reduction problems, Appl. Math. Comput. 181 (2006) 1215–1228.
[9] K. Jbilou, A. Messaoudi, H. Sadok, Global FOM and GMRES algorithms for matrix equations, Appl. Numer. Math. 31 (1999) 49–63.
[10] K. Jbilou, A.J. Riquet, Projection methods for large Lyapunov matrix equations, Linear Algebra Appl. 31 (2006) 49–63.
[11] Y.Q. Lin, Implicitly restarted global FOM and GMRES for nonsymmetric matrix equations and Sylvester equations, Appl. Math. Comput. 167-2 (2005) 1004–1025.
[12] Y. Saad, Iteratives Methods for Sparse Linear System, PWS press, New york, 1996.
[13] Y. Saad, M. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Statist. Comput. 7 (1986) 856–869.
[14] M. Sadkane, M. Robbé, Exact and inexact breakdowns in the block GMRES method, Linear Algebra Appl. 419 (2006) 265–285.
[15] H. Sadok, Analysis of the convergence of the minimal and the orthogonal residual methods, Numer. Algorithms 40 (2005) 101–115.
[16] D.K. Salkuyeh, F. Toutounian, New approaches for solving large Sylvester equations, Appl. Math. Comput. 173 (2006) 9–18.
[17] N. Wagner, A Krylov based method for quadratic matrix equations, in: K. Willner, M. Hanss (Eds.), Angewandte und experimentelle Mechanik—Ein Querschnitt, Der Andere Verlag, Tönning, Lübeck und Marburg, 2006, pp. 283–303.
[18] P. Vital, Etude de quelques méthodes de résolution de problèmes linéaires de grande taille sur multiprocesseur, Ph.D., Thesis, Université de Rennes, France, 1990.