Convergence Results for GMRES(m)

John C. Strikwerda
Suzan C. Stodder

Technical Report #1280

August 1995

Convergence Results for GMRES(m)

John C. Strikwerda*

and

Suzan C. Stodder*

August 1995

Abstract. We present a necessary and sufficient condition for the convergence of the GMRES(m) method. We present examples illustrating situations in which GMRES(m) stalls, i.e., fails to converge. Several theorems are proved regarding the convergence of GMRES(m) for special classes of matrices.

Keywords: GMRES, convergence estimates.

AMS(MOS) classifications: 65F10

1. Introduction.

The GMRES method was introduced by Saad and Schultz [8] as an iterative method to solve linear systems of equations. The advantage of the GMRES method over methods such as the conjugate gradient method and many other methods is that it can be applied to systems for which the matrix of coefficients is not symmetric or positive definite.

The GMRES method, as a Krylov space method, is a process that starts with a vector x generating a sequence of vectors and then chooses the update $y - x$ from the span of this sequence so as to reduce the residual. In practice, to limit the storage used, GMRES is replaced by GMRES(m) in which at most m vectors are generated to span the Krylov space. The GMRES(m) method consists of regenerating m vectors and choosing the update within the span of these vectors.

However, GMRES(m) has the disadvantage that it can stall. That is, for some initial vectors the best update is the zero vector. In this paper we state the precise conditions

* Department of Computer Sciences, University of Wisconsin - Madison, 1210 W. Dayton St., Madison, WI 53706.

that determine the stalling of GMRES(m). We also present a convergence estimate for GMRES(m) showing a geometric rate of convergence.

By convergence of GMRES(m) we mean that repeated application of the GMRES(m) algorithm produces a sequence of vectors that converge to the solution of the linear system and that this convergence holds for all initial vectors.

Several variants of GMRES(m) have been presented with the intent of reducing the likelihood of stalling. Among these are the FGMRES(m) method of [7], the DQGMRES method of [9], the GMBACK algorithm of [5], the GMRESR(m) method of [10], and the method of Joubert [4]. The methods of this paper should useful in analyzing the convergence of these methods. Another useful way of considering GMRES(m) from a theoretical point of view is presented in [3].

We begin in section 2 with a basic description of the GMRES(m) method. In section 3 we present and prove the basic theorem on the stalling of GMRES(m). Several examples illustrating stalling are presented in section 4 and several general converge conditions are presented in 5. A convergence estimate is given in section 6.

2. Description of GMRES(m).

We now present the description of the GMRES algorithm to solve $Ax = b$. This first presentation of the algorithm will be for ease of understanding.

We consider the linear system $Ax = b$ where A is a nonsingular $N \times N$ matrix with real entries. For some applications we can consider A to have complex entries, in which case we use the Hermitian inner product on $\mathbb{C}^N$.

We start with a description of the GMRES(m) method. There are several computational short cuts that we ignore in this description.

The algorithm starts with an initial iterate x^0. The residual $r^0 = b - Ax^0$ and the unit vector $v^0 = r^0/\|r^0\|$ are computed.

Iterate: For $j = 0, \ldots, m-1$

1. Compute Av^j.
2. Compute $h_{i,j} = (Av^j, v^i)$, for $i = 0, \ldots, j$.
3. Compute

$$\hat{v}^{j+1} = Av^j - \sum_{i=0}^{j} h_{i,j} v^i \tag{2.1}$$

and set $h_{j+1,j} = \|\hat{v}^{j+1}\|$.

If $h_{j+1,j} = 0$ or is sufficiently small, go to step Q with $q = j+1$

4. Compute $v^{j+1} = \hat{v}^{j+1}/h_{j+1,j}$.

End of loop on j.

Set $q = m$.

Stopping.

Q. Construct the new solution x^1 given by

$$x^1 = x^0 + \sum_{i=0}^{q-1} \alpha_i v^i \tag{2.2}$$

so that $\|b - Ax^1\|$ is minimal.

The success of GMRES(m) as a computational procedure depends on having an efficient means of solving the least squares problem (2.2). The solution of the least squares problem depends on the Arnoldi relations and is discussed in [8].

We begin by deriving the Arnoldi relation. From the equation (2.1), we obtain

$$Av^j = h_{j+1,j}v^{j+1} + \sum_{i=0}^{j} h_{i,j}v^i = \sum_{i=0}^{j+1} h_{i,j}v^i \,. \tag{2.3}$$

We form the matrices V_k whose columns are the vectors $v^0, \ldots, v^{k-1}$. Note that the V_k are orthogonal matrices. (When the index on a matrix or vector is related to the number of columns or components, as with V_k, we adopt the convention used in the language C, for index k, the indexing on the columns or components starts at 0 and runs up to $k-1$.)

The relation (2.3) can then be written as

$$Av^j = V_{j+2}h_j$$

where h_j is the vector with components $h_{0,j}, h_{1,j}, \ldots, h_{j+1,j}$. If we form the $(k+1) \times k$ matrix

$$H_k = \begin{pmatrix} h_{0,0} & h_{0,1} & \ldots & h_{0,k-1} \\ h_{1,0} & h_{1,1} & \ldots & h_{1,k-1} \\ 0 & h_{2,1} & \ldots & h_{2,k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & h_{k,k-1} \end{pmatrix}$$

then we obtain the Arnoldi relation

$$AV_k = V_{k+1}H_k \ . \tag{2.4}$$

The matrices H_k have upper Hessenberg form, that is, all elements are zero below the first subdiagonal.

We use the Arnoldi relation (2.4) to modify the minimization problem. Let a^m be the vector with components $\alpha_0, \alpha_1, \ldots, \alpha_{m-1}$. Then from (2.2) we have

$$\begin{aligned} b - Ax^1 &= b - A\left(x^0 + \sum_{i=0}^{m} \alpha_i v^i\right) \\ &= b - A\left(x^0 + V_m a^m\right) \\ &= b - Ax^0 - AV_m a^m \\ &= r^0 - AV_m a^m \ . \end{aligned}$$

Let $\beta = ||r^0||$ and let e^0 be the vector all of whose components are 0 except for the first component, which is 1. (The dimension of e^0 need not be specified.) Then, from the definition of v^0, we have $r^0 = \beta v^0 = \beta V_\ell e^0$. So we have by the Arnoldi relation

$$\begin{aligned} b - Ax^1 &= r^0 - AV_m a^m = \beta V_{m+1} e^0 - V_{m+1} H_m a^m \\ &= V_{m+1}\left(\beta e^0 - H_m a^m\right) \end{aligned}$$

and so

$$||b - Ax^1|| = ||V_{m+1}\left(\beta e^0 - H_m a^m\right)|| = ||\beta e^0 - H_m a^m||$$

since V_{m+1} is an orthogonal matrix.

This shows that the minimization problem is reduced from a problem with the $N \times N$ matrix A to one involving only the $(m+1) \times m$ matrix H_m. This makes the problem far more tractable. The system to solve is

$$H_m a^m \approx \beta e^0 \tag{2.5}$$

where the meaning of $\approx$ is that the two sides are as close as possible in the sense of least squares.

3. Stalling of GMRES(m).

The stalling of GMRES(m) occurs when the least squares solution of (2.5) is the solution $a^m = 0$. To see when this occurs we look at the structure of the linear system (2.5). We now look at the particular case of $m = 3$. The system is

$$\begin{pmatrix} h_{0,0} & h_{0,1} & h_{0,2} \\ h_{1,0} & h_{1,1} & h_{1,2} \\ 0 & h_{2,1} & h_{2,2} \\ 0 & 0 & h_{3,2} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} \approx \begin{pmatrix} \beta \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

A quick inspection shows that if the first row of the matrix H_m contains only zero entries, then the optimal solution is the zero solution. Moreover, if any element of the first row, say $h_{0,k}$ is nonzero, then nonzero values α_k can be chosen so that the the least squares solution is less than β. An examination of when the first row of the H matrices can be identically zero leads us to the following theorem.

Theorem 3.1. *A necessary and sufficient condition for GMRES(m) to converge, that is, not stall, is that the set of vectors*

$$\mathcal{V}_m = \{v : (v, A^j v) = 0 \quad \textit{for} \quad 1 \leq j \leq m\}$$

contains only the vector $\vec{0}$.

Proof.

If the initial vector v^0 is in $\mathcal{V}_m$, then it is easy to see that the first row of the matrix H_m is zero. Indeed, $h_{0,0} = (v^0, Av^0) = 0$, and so $v^1 = Av^0/||Av^0||$. We have also

$$h_{0,1} = (v^0, Av^1) = (v^0, A^2v^0)/||Av^0|| = 0.$$

Each successive v^j for j larger than 1 is a linear combination of the vectors $A^k v^0$ for k strictly greater than 0 and k at most j. Thus all inner products (v^0, Av^j) are zero.

If the first row of H_m is zero, then the least squares problem has the solution $a^m = 0$. Indeed, the optimal solution is the solution to

$$H_m^T H_m a^m = \beta \, H_M^T e^0 = \vec{0}$$

which has the solution $a^m = \vec{0}$.

Conversely, if $\mathcal{V}_m = \{\vec{0}\}$, then for each vector v^0 there is some element $h_{0,j}$ that is nonzero. By choosing $\alpha_j = \varepsilon$ and all other $\alpha_k = 0$ we have, for ε to be sufficiently small and with the sign of $h_{0,j}$,

$$||H_m a^m - \beta e^0||^2 = (\varepsilon h_{0,j} - \beta)^2 + O(\varepsilon)^2 < \beta^2$$

(recall that β is positive). Thus the least squares solution results in a solution with residual norm less than β and GMRES(m) does not stall. ∎

In [1], it is observed that the stalling of GMRES(m) is related to the singularity of the matrix H'_m formed from the first m rows of H_m. As Theorem 3.1 shows, stalling is equivalent to this matrix having the first row be identically zero. If H'_m is singular, but the first row of H'_m is not identically zero, then the algorithm will not stall, and in fact it converges in that step.

The condition that $\mathcal{V}_m$ contain only the zero vector can be viewed as a generalization of the matrix A being positive definite. Indeed, $\mathcal{V}_1 = \{\vec{0}\}$ is equivalent to A being either positive definite or negative definite.

4. Examples.

In this section we present two examples of matrices of dimension N with $N > m$ for which GMRES(m) fails to converge. This first example has been presented by Brown in [1].

Example 1.

We consider an orthonormal set of p vectors v^k for $k = 0, \ldots, p-1$. The matrix P is the permutation matrix corresponding to $v^0 \to v^1 \to v^2 \to \ldots \to v^{p-1} \to v^0$. Define P to be the identity on the orthogonal complement of the span of the v^j. Because $(v^0, v^j) = (v^0, P^j v^0)$ for $j = 1, \ldots, p-1$ we see that GMRES(m) will fail to converge with initial residual v^0 for matrix P if m is less than p.

Notice that the minimal polynomial of P is $\lambda^p - 1$.

Example 2.

For the second example we consider $(N+1)\times(N+1)$ matrices of the form

$$\begin{pmatrix} 1 & -\gamma & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -\gamma & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & -\gamma & & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & & 1 & -\gamma & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & -\gamma \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix} \tag{4.1}$$

The condition $(x, A^k x) = 0$ for $1 \le k \le m \le N$ with $x = (x_0, x_1, \dots, x_N)$ is then

$$\sum_{\ell=0}^{k} (-\gamma)^\ell \binom{k}{\ell} \sum_{j=0}^{N-\ell} x_j x_{j+\ell} = 0.$$

By taking x to be a unit vector, and by comparing successive conditions, we see that these conditions are equivalent to

$$\sum_{j=0}^{N-\ell} x_j x_{j+\ell} = \gamma^{-\ell} \quad \text{for} \quad 0 \le \ell \le m. \tag{4.2}$$

We now show that the system (4.2) has nontrivial solutions for γ sufficiently large. Thus for γ sufficiently large, GMRES(m) will not converge for the matrix (4.1).

Theorem 4.1. *There is a constant $\Gamma(m)$, depending only on m such that for $\gamma \ge \Gamma(m)$ and for $m \le N$ the system of equations (4.2) has solutions.*

Proof.

We consider the nonlinear system $F(x(\varepsilon)) = b(\varepsilon)$ where F and $b(\varepsilon)$ are defined by

$$F_0(x) = \frac{1}{2}\sum_{j=0}^{N} x_j^2 = \frac{1}{2}$$

$$F_\ell(x) = \sum_{j=0}^{N-\ell} x_j x_{j+\ell} = \varepsilon^\ell \quad \text{for } \ell > 0$$

For $\varepsilon = 0$ we have the solution $x_0(0) = 1$ and $x_\ell(0) = 0$ for $\ell > 0$. If we set $x_j = 0$ for $m < j \leq N$ we have $m+1$ equations in the remaining $m+1$ unknowns. The Jacobian of F at this x is the identity matrix. Thus there is a solution x for all vectors b near to $b(0)$, see e.g., Buck [2] sect. 5.7. In particular, for ε sufficiently small there is a solution $x(\varepsilon)$ corresponding to $b(\varepsilon)$. By taking $\gamma = \varepsilon^{-1}$ there is a solution to the system (4.2). ∎

This second example is interesting because it shows that the stalling phenomena is not dependent just on the eigenvalues, but on the structure of the eigenspace.

We think that these two examples illustrate the two basic reasons that GMRES(m) stalls. Either the eigenvalues are too scattered around the origin or the matrix is too far from being a normal matrix or both of these conditions hold together. These ideas are illustrated by the theorems in the next section. The paper [6] by Nachtigal, Reddy, and Trefethen contains several test matrices for GMRES and other methods. Those that cause difficulty for GMRES(m) have the basic form of the above examples.

5. Special Cases.

In this section we present several special classes of matrices for which GMRES(m) converges. These results are all applications of Theorem 3.1.

Theorem 5.1. *If A is a symmetric matrix or if A is a skew-symmetric matrix, then GMRES(2) converges.*

Proof.

We have for all nonzero vectors v

$$(v, A^2 v) = \pm(v, A^T A v) = \pm\|Av\|^2 \neq 0$$

where the plus sign corresponds to the symmetric case, and the minus sign to the skew-symmetric case. ∎

We also state a theorem related to the first example in section 4. For the next two theorems we can allow A to be a matrix with complex entries.

Theorem 5.2. *If A is a normal matrix and $p(\cdot)$ is a polynomial of degree m that maps the spectrum of A into the right half plane, then GMRES(m) converges for A.*

Proof.

Let $p(\mu) = \sum_{j=0}^{m} a_j \mu^j$ and let v be a nonzero vector decomposed as $\sum_{\alpha=1}^{N} v_\alpha$ where v_α is an eigenvector of A with eigenvalue λ_α. Then

$$(v, p(A)v) = \sum_{j=0}^{m} a_j (v, A^j v) = \sum_{\alpha=1}^{N} p(\lambda_\alpha)\|v_\alpha\|^2$$

So

$$\mathrm{Re} \sum_{j=0}^{m} a_j (v, A^j v) = \mathrm{Re} \sum_{\alpha=1}^{N} p(\lambda_\alpha)\|v_\alpha\|^2 > 0.$$

Hence, for each nonzero vector v some of the inner products $(v, A^j v)$ must be nonzero and so GMRES(m) will converge. ∎

Theorem 5.1 follows from Theorem 5.2 using the polynomial μ^2 for the symmetric case and $-\mu^2$ for the skew case. Using the special polynomial μ^m we have the following result.

Theorem 5.3. *Let A be a normal matrix. If there is an integer m and a real number δ such that each eigenvalue of A is in one of the sets*

$$\mathcal{I}_k = \left\{ z = re^{i\theta} : |\theta - \frac{2k\pi}{m} - \delta| < \frac{\pi}{2m} \right\} \quad \textit{for} \quad k = 0, 1, \ldots, m-1,$$

then GMRES(m) converges for all initial vectors.

Proof.

If λ_j is in $\mathcal{I}_k$, then λ_j^m is in the set

$$\left\{ z = re^{i\theta} : |\theta - 2k\pi - m\delta| < \frac{\pi}{2} \right\} = \left\{ z = re^{i\theta} : |\theta - m\delta| < \frac{\pi}{2} \right\}.$$

Thus the real part of $e^{-im\delta}\lambda_j^m$ has positive real part and so by Theorem 5.2 with the polynomial $e^{-im\delta}\mu^m$ GMRES(m) will converge for A. ∎

For general matrices, i.e., not necessarily normal, we have the next results.

Theorem 5.4. *A sufficient condition for GMRES(m) to converge, that is, not stall, is that there is a polynomial $g(\cdot)$ of degree m with $g(0) = 0$ such that $g(A)$ is positive definite.*

Proof.

If $g(t) = \sum_{j=1}^{m} g_j t^j$ is a polynomial such that $g(A)$ is positive definite, then for each nonzero vector v

$$0 < (v, g(A)v) = \sum_{j=1}^{m} g_j (v, A^j v)$$

and thus some of the inner products $(v, A^j v)$ must be nonzero, and by Theorem 3.1, GMRES(m) converges. ∎

Theorem 5.5. *If the minimal polynomial of A has degree m, then GMRES(m) converges.*

Proof.

Let $p(t) = \sum_{j=0}^{m} a_j t^j$ be the minimal polynomial of A. Since A is nonsingular, a_0 is not zero. So for any nonzero vector x

$$0 \neq a_0 \|x\|^2 = -\sum_{j=1}^{m} a_j (x, A^j x) .$$

Thus for each vector x some of the inner products $(x, A^j x)$ must be nonzero. Thus GMRES(m) converges. ∎

Theorem 5.6. *If GMRES(m) converges for A, then GMRES(m) converges for A^T.*

Proof. The proof follows from Theorem 3.1 and the observation that $(v, A^j v) = (v, (A^T)^j v)$. ∎

6. A Convergence Estimate.

We now prove a convergence estimate for GMRES(m).

Theorem 6.1. *If r^k is the residual after k steps of GMRES(m), then*

$$\|r^k\|^2 \leq (1-\rho_m)^k \|r^0\|^2 \tag{6.1}$$

where

$$\rho_m = \rho_m(A) = \min_{\|v\|=1} \left(\frac{\sum_{j=0}^{m-1} (v^0, Av^j)^2}{\sum_{j=0}^{m-1} \|Av^j\|^2} \right)$$

and the vectors v^j are the unit vectors generated in the GMRES procedure starting with $v^0 = v$.

Notice that $\rho_m = 0$ is equivalent to the necessary and sufficient condition for stalling presented in Theorem 3.1.

Proof.

The least squares solution to (2.5) has a smaller residual than does the solution to

$$H_m h_0 \alpha \approx \beta e^0 \tag{6.2}$$

where h_0 is the column vector with entries $h_{0,j}$. The solution to (2.5) with $v = \alpha h_0$ is the solution of

$$\min_{\alpha} \left[(\alpha P - \beta)^2 + \alpha^2 Q^2 \right]$$

where

$$P = \sum_{j=0}^{k-1} h_{0,j}^2$$

and

$$Q^2 = \sum_{i=1}^{k} \left(\sum_{j=i-1}^{k-1} h_{i,j} h_{0,j} \right)^2 .$$

At the minimum

$$\alpha = \beta \frac{P}{P^2 + Q^2}$$

and

$$\|r^1\|^2 \le \beta^2 \frac{Q^2}{P^2+Q^2} \, .$$

We now estimate the expression $Q^2/(P^2+Q^2) = (Q^2/P^2)/(1+Q^2/P^2)$.

$$\frac{Q^2}{P^2} = \frac{\sum_{i=1}^k \left(\sum_{j=i-1}^{k-1} h_{i,j} h_{0,j}\right)^2}{(\sum_{j=0}^{k-1} h_{0,j}^2)^2}$$

$$\le \frac{\sum_{i=1}^k \left(\sum_{j=i-1}^{k-1} h_{i,j}^2 \sum_{j=i-1}^{k-1} h_{0,j}^2\right)}{(\sum_{j=0}^{k-1} h_{0,j}^2)^2} \le \frac{\sum_{i=1}^k \sum_{j=i-1}^{k-1} h_{i,j}^2}{(\sum_{j=0}^{k-1} h_{0,j}^2)}$$

by the Cauchy-Schwartz inequality. Continuing by interchanging summations

$$\frac{Q^2}{P^2} \le \frac{\sum_{i=0}^k \sum_{j=i-1}^{k-1} h_{i,j}^2}{\sum_{j=0}^{k-1} h_{0,j}^2} = \frac{\sum_{j=0}^{k-1} \sum_{i=0}^{j+1} h_{i,j}^2}{\sum_{j=0}^{k-1} h_{0,j}^2}$$

$$= \frac{\sum_{j=0}^{k-1} \left(h_{j+1,j}^2 + \sum_{i=1}^{j} h_{i,j}^2\right)}{\sum_{j=0}^{k-1} h_{0,j}^2} \, .$$

Using the relation, from (2.1), see also [8],

$$h_{j+1,j}^2 = \|Av^j\|^2 - \sum_{i=0}^{j} h_{i,j}^2$$

we have

$$\frac{Q^2}{P^2} \le \frac{\sum_{j=0}^{k-1} \|Av^j\|^2 - \sum_{j=0}^{k-1} h_{0,j}^2}{\sum_{j=0}^{k-1} h_{0,j}^2} = \frac{\sum_{j=0}^{k-1} \|Av^j\|^2}{\sum_{j=0}^{k-1} h_{0,j}^2} - 1 \, .$$

Thus, using $h_{0,j} = (v^0, Av^j)$,

$$\frac{Q^2}{P^2+Q^2} = 1 - \frac{P^2}{P^2+Q^2} = 1 - \frac{1}{1+Q^2/P^2} \le 1 - \frac{\sum_{j=0}^{k-1} (v^0, Av^j)^2}{\sum_{j=0}^{k-1} \|Av^j\|^2}$$

and by taking the maximum of the right-hand side of this expression, we have (6.1). This proves the theorem. ∎

The convergence estimate of Theorem 6.1 is not very useful since the value of ρ_m is not easily obtainable in most cases. However, it does show that if GMRES(m) does converge, it does so with a geometric rate of convergence. Also, it relates the necessary and sufficient condition of Theorem 3.1 with the rate of convergence.

7. GMRES(m) and Preconditioning.

Frequently, GMRES(m) is used with preconditioning, that is, a matrix B is used to modify the original linear system to either

$$(B^{-1}A)x = B^{-1}b,$$

which is left preconditioning or

$$(AB^{-1})Bx = b,$$

which is right preconditioning. The matrix B is called the preconditioning matrix and it is chosen so that the matrices $B^{-1}A$ or AB^{-1} have nice properties with respect to the iterative method. Preconditioning is a common practice to increase the speed of iterative procedures such as GMRES(m) and other conjugate-gradient-like methods.

The condition that the left preconditioned system will stall is that there is a nonzero solution v to

$$(v, (B^{-1}A)^j v) = 0 \quad \text{for} \quad 1 \le j \le m, \tag{7.1}$$

and the condition that the right preconditioned system will stall is that there is a nonzero solution v to

$$(v, (AB^{-1})^j v) = 0 \quad \text{for} \quad 1 \le j \le m. \tag{7.2}$$

By using left and right preconditioning on alternate GMRES(m) steps, the condition that the iterative process stall is that there must be a nonzero vector v satisfying both (7.1) and (7.2). Because this involves $2m$ conditions, we see that the alternating of the left and right preconditioning may be less likely to stall than is either left or right preconditioning by itself. Of course, in the trivial case with B the identity matrix there is no advantage to this strategy. However, in real computation in which the matrix B is a reasonable preconditioner it is quite likely that the alternating of the preconditioning will prevent stalling in many situations.

The matrices $B^{-1}A$ and AB^{-1} have the same eigenvalues, however, the relations (7.1) and (7.2) will, in general, be independent conditions. With the left/right preconditioning the estimate (6.1) is replaced by

$$\|r^{2k}\|^2 \le \left([1 - \rho_m(B^{-1}A)][1 - \rho_m(AB^{-1})]\right)^k \|r^0\|^2 .$$

However, even if both $\rho_m(B^{-1}A)$ and $\rho_m(AB^{-1})$ are zero, alternating the preconditioning will still converge if there are no nontrivial vectors satisfying both (7.1) and (7.2).

The flexible GMRES algorithm (FGMRES) of Saad [7] is different in approach to that suggested here in that it changes the preconditioner within the GMRES(m) steps. It should be possible to extend the results of this paper to the FGMRES algorithm and other suggested variants of GMRES(m).

8. Conclusions.

We have given a necessary and sufficient condition for the GMRES(m) method to converge. A convergence estimate has also been established. We have applied the convergence condition to obtain sufficient conditions for convergence for some special classes of matrices. Further research is being done on extending the methods of this paper to some of the variants of GMRES.

References

[1] P. N. Brown, *A Theoretical Comparison of the Arnoldi and GMRES Algorithms*, SIAM J. Sci. Comput., 12 (1991), pp. 58-78.

[2] R. C. Buck, *Advanced Calculus*, McGraw-Hill, New York, 1965.

[3] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as Matrix Approximation Problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359-368.

[4] W. Joubert, *A Robust GMRES-Based Adaptive Polynomial Preconditioning Algorithm for Nonsymmetric Linear Systems*, SIAM J. Sci. Comput., 15 (1994), pp. 427-439.

[5] E. M. Kasenally, *GMBACK: A Generalized Minimum Backward Error Algorithm for Nonsymmetric Linear Systems*, SIAM J. Sci. Comput., 16 (1995), pp. 698-719.

[6] N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, *How Fast Are Nonsymmetric Matrix Iterations?*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 778-795.

[7] Y. Saad, *A Flexible Inner-Outer Preconditioned GMRES Algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461-469.

[8] Y. Saad and M.H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856-869.

[9] Y. Saad and K. Wu, *DQGMRES: A Direct Quasi-Minimal Residue Algorithm Based on Incomplete Orthogonalization*, Univ. Minn. Tech. Report (1995).

[10] H.A. van der Vorst and C. Vuik, *GMRESR: a family of nested GMRES methods*, Tech. Report 91-80, Delft University of Technology (1991).