A Convergent Restarted GMRES Method
For Large Linear Systems

Minghua Xu\textsuperscript{1,2}, Jinxi Zhao\textsuperscript{1}, Jiancheng Wu\textsuperscript{2} and Hongjun Fan\textsuperscript{1}
\textsuperscript{1}.State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, P.R.China
\textsuperscript{2}.Department of Information Science, Jiangsu Polytechnic University, Changzhou, 213016, P.R.China

October 3, 2003

Abstract The GMRES method is popular for solving nonsymmetric linear
equations. It is generally used with restarting to reduce storage and orthogonal-
ization costs. However, it is possible to show that the restarted GMRES method
may not converge, i.e., it may be stationary. To remedy this difficulty, a new
convergent restarted GMRES method is discussed in this paper.

Key words GMRES, Krylov subspace, iterative methods, nonsymmetric sys-
tems.

AMS(MOS) subject classifications 65F10

1. Introduction

The restarted GMRES algorithm GMRES(m)\cite{Saad1} proposed by Saad and Schultz
is one of the most popular iterative methods for solving large linear systems of
equations

\[ Ax = b, A \in \mathbb{R}^{n \times n}, x, b \in \mathbb{R}^{n}, \]

with a sparse, nonsymmetric, and nonsingular matrix $A$. It is known that when $A$
is positive real, the restarted GMRES method will produce a sequence of ap-
proximates $x_k$ that converge to the exact solution. However, when $A$ is not
positive real, this method often slows down convergence and stagnates. The
analysis and implementation of the restarted GMRES algorithm continue to re-
ceive considerable attention \cite{Saad2,Saad3,Saad4,Saad5,Saad6,Saad7}. For example, Y.Saad suggested a flexible
inner-outer preconditioned GMRES method FGMRES(m)\cite{Saad2}. R.B.Morgan gave
a restarted GMRES method augmented with eigenvectors \cite{Saad3}, and Cao Zhihao et.
al. presented a convergent restarted GMRES algorithm based on the algorithm
FGMRES(m)\cite{Saad4}. We will now briefly review the algorithm GMRES in this sec-
tion. A new restarted GMRES method and its analysis will be given in section
2, section 3 gives the examples and comparisons, and conclusions are given in
section 4. The restarted GMRES can be briefly described as follows.

Algorithm 1: GMRES(m) for systems (1.1)
1. Start: Choose \( x_0 \) and compute \( r_0 = b - Ax_0 \) and \( \beta = \|r_0\|, v_1 = r_0/\beta \).

2. Iterate: For \( j = 1, 2, \cdots, m \) do:
   \[
   h_{i,j} = (Av_j, v_i), \quad i = 1, 2, \cdots, j,
   \]
   \[
   \bar{v}_{j+1} = Av_j - \sum_{i=1}^{j} h_{i,j} v_i,
   \]
   \[
   h_{j+1,j} = \|\bar{v}_{j+1}\|,
   \]
   \[
   v_{j+1} = \bar{v}_{j+1}/h_{j+1,j}.
   \]

3. Form the approximate solution:
   \[
   x_m = x_0 + V_m y_m, \text{ where } y_m \text{ minimizes } \|\beta e_1 - H_m y\|, y \in R^m. \text{ Here } H_m \text{ is the } (m+1) \times m \text{ matrix whose only nonzero entries are the elements } h_{i,j} \text{ defined in step 2. } V_m = [v_1, v_2, \cdots, v_m] \text{ and the vector } e_1 \text{ is the first column of the } (m+1) \times (m+1) \text{ identity matrix.}
   \]

4. Restart:
   Compute \( r_m = b - Ax_m \), if satisfied then stop else compute \( x_0 := x_m, r_0 := r_m, \beta := \|r_0\|, v_1 := r_0/\beta \) and go to 2.

If \( A \) is not positive real, then \( r_0 \perp \text{span}\{Ar_0, A^2r_0, \cdots, A^mr_0\} \) may happen. In this situation the restarted GMRES method is stationary. To avoid this disadvantage, we introduce and analyze a new convergent restarted GMRES method. Conveniently, we use the term CGMRES(m) to denote the method.

2. CGMRES(m)

The linear systems associated with (1.1) can be taken as the following form

\[
\begin{bmatrix}
I & A \\
-A^T & 0
\end{bmatrix}
\begin{bmatrix}
u^* \\
x
\end{bmatrix} =
\begin{bmatrix}
f \\
g
\end{bmatrix},
\]

where \( I \in R^{n\times n} \) is the identity matrix, while \( u^* \in R^n \) is a given vector and \( f = u^* + b, g = -A^T u^* \in R^n \). Since \( A \) is nonsingular, thus the system (2.1) has an unique solution \( z^* = \begin{bmatrix} u^* \\ x^* \end{bmatrix} \). Let \( z_0 \) is the initial approximate solution of (2.1), \( B = \begin{bmatrix}
I & A \\
-A^T & 0
\end{bmatrix} \), \( \bar{r}_0 = \begin{bmatrix} f \\ g \end{bmatrix} - Bz_0 \). Solving the systems (2.1) with GMRES(m), where \( m \geq 2 \), we have the following results:

**Proposition 2.1** Denoting by \( \beta_i, i = 1, 2, \cdots, n \), the eigenvalues of \( A^T A \) and supposing

\[
\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n \geq 1/4,
\]

then we have that the eigenvalues of matrix \( B \) have positive real part.

**Proof** We have

\[
|\lambda I_1 - B| = \begin{vmatrix}
(\lambda - 1)I & -A \\
A^T & \lambda I
\end{vmatrix} ,
\]

(2.3)
where $I_1 \in R^{2n \times 2n}$ is a identity matrix. If $\lambda = 1$ is the eigenvalue of matrix $B$, by (2.3), we get

$$\begin{vmatrix} 0 & -A \\ A^T & I \end{vmatrix} = 0.$$  

This is a contradiction with that matrix $A$ is nonsingular. Hence, we have $\lambda \neq 1$. According to (2.3) and $\lambda \neq 1$, we have

$$|\lambda I_1 - B| = \left| (\lambda - 1)I \begin{array}{c} -A \\ \lambda I \end{array} \right| = \left| (\lambda - 1)\lambda I + (\lambda - 1)^{-1}A^TA \right|.$$  

If $\{\lambda_{i,j}|i = 1, 2, \cdots, n; j = 1, 2\}$ denote the eigenvalues of $B$, then $\lambda_{i,j}, j = 1, 2$, can be given by solving the following equation

$$(\lambda_{i,j} - 1)\lambda_{i,j} + \beta_i = 0, i = 1, 2, \cdots, n, j = 1, 2. \quad (2.4)$$  

Solving (2.4) we obtain

$$\lambda_{i,1} = \frac{1 + \sqrt{1 - 4\beta_i}}{2} \quad (2.5)$$  

and

$$\lambda_{i,2} = \frac{1 - \sqrt{1 - 4\beta_i}}{2}, \quad (2.6)$$  

$i = 1, 2, \cdots, n$. Using (2.2) yields the desired result.

According to [1] and Proposition 2.1, we know that using the restarted GMRES method to solve (2.1) will produce a sequence of approximations which converges to the exact solution of (2.1) when $\beta_i \geq 1/4, i = 1, 2, \cdots, n$. If the conditions $\beta_i \geq 1/4, i = 1, 2, \cdots, n$ do not hold, we can use

**Proposition 2.2** Assume that $y_k, k = 1, 2, \cdots, m$ minimizes $\|\beta e_1 - H_k y\|$, $y \in R^k$, $H_k$ is the $(k + 1) \times k$ matrix whose nonzero entries are the elements $h_{i,j}$ defined by GMRES(m) for (2.1), $z_k = z_0 + V_k y_k$ is the approximate solution of (2.1), where $V_k = [v_1, v_2, \cdots, v_k]$, $v_i$ is the Arnoldi vector generated by GMRES(m) for (2.1), $i = 1, 2, \cdots, k$, $k = 1, 2, \cdots, m (m \geq 2)$. Suppose that $z_k = \begin{bmatrix} u_k \\ x_k \end{bmatrix} \in R^{2n}$, $u_k, x_k \in R^n$ and residual $\bar{r}_k = \begin{bmatrix} f \\ g \end{bmatrix} - B z_k$ then the following results hold:

(1) $\|\bar{r}_m\|_2 < \|\bar{r}_0\|_2$ and $z_k$ tends to the exact solution $z^* = \begin{bmatrix} u^* \\ x^* \end{bmatrix}$ of (2.1).

(2) $x^*$ is the exact solution of (1.1).

**Proof of (1)** The residual vector of the approximate solution $z_k$ can be written as

$$\bar{r}_k = \begin{bmatrix} f \\ g \end{bmatrix} - B z_k = \begin{bmatrix} u^* + b - u_k - A x_k \\ g + A^T u_k \end{bmatrix} = \begin{bmatrix} \bar{r}_{k1} \\ \bar{r}_{k2} \end{bmatrix}, \quad k = 0, 1, \cdots, m.
Suppose further that \( \bar{v}_1 = \bar{r}_0/\|\bar{r}_0\|_2 = \begin{bmatrix} \bar{v}_{1,1} \\ \bar{v}_{1,2} \end{bmatrix} \neq 0 \). We find

\[
\bar{v}_1^T B \bar{v}_1 = \bar{v}_{1,1}^T \bar{v}_{1,1} + \bar{v}_{1,1}^T A \bar{v}_{1,2} - \bar{v}_{1,2}^T A^T \bar{v}_{1,1} = \bar{v}_{1,1}^T \bar{v}_{1,1} \geq 0
\]

and if \( \bar{v}_{1,1} = 0 \), we have

\[
\bar{v}_1^T B^2 \bar{v}_1 = \begin{bmatrix} -\bar{v}_{1,2}^T A^T, 0 \end{bmatrix} \begin{bmatrix} A \bar{v}_{1,2} \\ 0 \end{bmatrix} = -\bar{v}_{1,2}^T A^T A \bar{v}_{1,2} \leq 0
\]

Let \( K_m = \text{span}\{\bar{r}_0, B\bar{r}_0, \ldots, B^{m-1}\bar{r}_0\} \). We have

\[
\|\bar{r}_m\|_2 = \min_{z \in K_m} \left\| \begin{bmatrix} f \\ g \end{bmatrix} - B[z_0 + z] \right\|_2
\]

\[
= \min_{z \in K_m} \|\bar{r}_0 - Bz\|_2
\]

\[
= \min_{y \in R^m} \|\beta \bar{v}_1 - BV_m y\|_2,
\]

where \( z = V_m y \).

Using \( BV_m = V_{m+1}H_m \), we get

\[
\min_{y \in R^m} \|\beta \bar{v}_1 - BV_m y\|_2 = \min_{y \in R^m} \|\beta e_1 - H_m y\|_2,
\]

and (2.10), we obtain

\[
\|\bar{r}_m\|_2 = \min_{z \in K_m} \|\bar{r}_0 - Bz\|_2 \leq \|\bar{r}_0 - \frac{\beta \bar{v}_1^T B \bar{v}_1}{\|B \bar{v}_1\|_2^2} B \bar{v}_1\|_2.
\]

Let

\[
c_1 = \frac{\beta \bar{v}_1^T B \bar{v}_1}{\|B \bar{v}_1\|_2^2}, R_1 = \bar{r}_0 - c_1 B \bar{v}_1.
\]

According to \( \bar{v}_{1,1} \neq 0 \) and (2.7), we have

\[
c_1 > 0, R_1^T B \bar{v}_1 = 0
\]

and

\[
\|\bar{r}_0\|_2 = \|\bar{r}_0 - c_1 B \bar{v}_1 + c_1 B \bar{v}_1\|_2
\]

\[
= \sqrt{\|R_1\|_2^2 + c_1^2 \|B \bar{v}_1\|_2^2} > \|R_1\|_2.
\]

By (2.11) and (2.12), we can get

\[
\|\bar{r}_m\|_2 \leq \|R_1\|_2 < \|\bar{r}_0\|_2.
\]

In similar way, if \( \bar{v}_{1,1} = 0 \) then \( \bar{v}_{1,2} \neq 0 \) we have

\[
\|\bar{r}_m\|_2 = \min_{z \in K_m} \|\bar{r}_0 - Bz\|_2 \leq \|\bar{r}_0 - \frac{\beta \bar{v}_1^T B^2 \bar{v}_1}{\|B^2 \bar{v}_1\|_2^2} B^2 \bar{v}_1\|_2,
\]

\[
\|\bar{r}_m\|_2 \leq \|R_1\|_2 < \|\bar{r}_0\|_2.
\]
where $m \geq 2$.

Let

$$c_2 = \frac{\beta \bar{v}_1^T B^2 \bar{v}_1}{\|B^2 \bar{v}_1\|_2^2}, R_2 = \bar{r}_0 - c_2 B^2 \bar{v}_1.$$ 

According to $\bar{v}_{1,2} \neq 0$ and (2.8), we have

$$c_2 < 0, R_2^T B^2 \bar{v}_1 = 0$$

and

$$\|\bar{r}_0\|_2^2 = \|\bar{r}_0 - c_2 B^2 \bar{v}_1 + c_2 B^2 \bar{v}_1\|_2^2 = \sqrt{\|R_2\|_2^2 + c_2^2 \|B^2 \bar{v}_1\|_2^2} > \|R_2\|_2.$$

Thus, we can get

$$\|\bar{r}_m\|_2 \leq \|R_2\|_2 < \|\bar{r}_0\|_2.$$ (2.15)

Applying (2.13) and (2.16), we know if $\bar{r}_0 \neq 0$ then $\|\bar{r}_m\|_2 < \|\bar{r}_0\|_2$, the result (1) holds.

**Proof of (2)** Since $z^* = \begin{bmatrix} u^* \\ x^* \end{bmatrix}$ satisfies (2.1), we can find

$$-A^T u^* = g$$ (2.11)

and

$$Ax^* = (f - u^*) = b$$ (2.12)

By (2.12) we have thus obtained the result (2).

According to Propositions 2.1 and 2.2, using algorithm GMRES(m) to solve system (2.1), we can obtain and approximate solution of (1.1). Now the CGMRES(m) algorithm can be briefly described as follows:

**Algorithm 2: CGMRES(m) for systems (1.1)**

1. Start: Choose $z_0$ and compute $r_0 = \begin{bmatrix} f \\ g \end{bmatrix} - Bz_0$ and $\beta = \|r_0\|, v_1 = r_0/\beta$.

2. Iterate: For $j = 1, 2, \cdots, m$ do:

   $$h_{i,j} = (Bv_j, v_i), \quad i = 1, 2, \cdots, j,$$

   $$\bar{v}_{j+1} = Bv_j - \sum_{i=1}^j h_{i,j} v_i,$$

   $$h_{j+1,j} = \|\bar{v}_{j+1}\|,$$

   $$v_{j+1} = \bar{v}_{j+1}/h_{j+1,j}.$$

3. Form the approximate solution of (2.1):

   $$z_m = z_0 + V_m y_m, \text{ where } V_m = [v_1, v_2, \cdots, v_m] \text{ and } y_m \text{ minimizes } \|\beta e_1 - H_m y\|, \text{ for } y \in R^m. \text{ Here } H_m \text{ is the } (m+1) \text{ by } m \text{ matrix whose only nonzero entries are the elements } h_{i,j} \text{ defined in step 2, and } e_1 \text{ is the first column of the } (m+1) \times (m+1) \text{ identity matrix.}$$
4. Form the approximate solution of (1.1):
\[ x_m = \begin{bmatrix} \mathbf{z}^{(n+1)}, \mathbf{z}^{(n+2)}, \cdots, \mathbf{z}^{(2n)} \end{bmatrix}^T, \]
where \( \mathbf{z}_m = \begin{bmatrix} \mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \cdots, \mathbf{z}^{(2n)} \end{bmatrix}^T \).

5. Restart:
\[
\text{Compute } r_m = b - Ax_m, \text{ if satisfied, then stop, else compute } r_m = \begin{bmatrix} f \\ g \end{bmatrix} - Bz_m, \text{ set } z_0 := z_m, r_0 := r_m, \beta := \|r_0\|, v_1 := r_0/\beta, \text{ and go to 2.}
\]

In comparing algorithm CGMRES(m) with GMRES(m), where \( m \geq 2 \), it is clear that CGMRES(m) has all the advantages of algorithm GMRES(m) and is a convergent algorithm, but it needs more storage than is required by GMRES(m), and costs nearly as much as by GMRES(m) in each inner loop.

3. Numerical experiments
In this section we report a few numerical experiments comparing the performances of CGMRES(m) with GMRES(m).

Example 1. Consider \( A = \text{Toeplitz}([1, -3.5, 1, 1, 1]) \in \mathbb{R}^{200 \times 200} \), where the diagonal element of the matrix underlined. The matrix \( A \) has extreme singular value \( \sigma_{200} = 4.1375 \times 10^{-11} \) and \( \sigma_1 = 5.4955 \). Let \( b = A[2, 2, \cdots, 2]^T, x_0 = [0, 0, \cdots, 0]^T \in \mathbb{R}^{200}, m = 10, \) and \( z_0 = [0, 0, \cdots, 0]^T \in \mathbb{R}^{400} \). The logarithm of the norm of relative residual is given by \( \log_{10}(\|b - Ax_m\|/\|b\|) \). Figure 1 exhibits the convergence histories of GMRES(m) and CGMRES(m) against the number of matrix-vector products.

Example 2. Consider \( A = \text{Toeplitz}([1, 0, 0, 1, 1, 1]) \in \mathbb{R}^{200 \times 200} \), and \( b = A[2, 2, \cdots, 2]^T \in \mathbb{R}^{200}, x_0, m, \) and \( z_0 \) as in Example 1. Figure 2 exhibits the convergence histories of GMRES(m) and CGMRES(m) against the number of matrix-vector products.
4. Conclusion
Algorithm CGMRES(m) is useful especially when GMRES(m) is stationary. It can avoid stagnation arising from algorithm GMRES(m). However, we can’t draw the conclusion that the convergence rate of CGMRES(m) is faster than GMRES(m) when GMRES(m) is convergent.

References