Notes on the Convergence of the Restarted GMRES

Eugene Vecharynski

Julien Langou

Department of Mathematical & Statistical Sciences

University of Colorado Denver

Outline

- Brief overview
- The cycle-convergence of the restarted GMRES for normal matrices
- The cycle-convergence of the restarted GMRES in the general case

Brief overview: GMRES (1 of 2)

GMRES (Generalized Minimal residual method) is a well known Krylov subspace method for solving linear systems of equations with non-Hermitian matrices

$$Ax = b, A \in \mathbb{C}^{n \times n}, b \in \mathbb{C}^{n}. \tag{1}$$

The basic idea of GMRES is to construct approximations x_m to the exact solution of (1) of the form

$$x_m = x_0 + u_m,\ u_m \in \mathcal{K}_m(A, r_0), \tag{2}$$

where $\mathcal{K}_m(A, r_0) = \text{span}\left\{r_0, Ar_0, \ldots, A^{m-1}r_0\right\}$ is the m-dimensional Krylov subspace, x_0 - any initial guess, $r_0 = b - Ax_0$ - the initial residual vector.

Brief overview: GMRES (2 of 2)

At each step m, the approximation x_m to the exact solution is chosen according to the condition that the corresponding *residual vector* r_m *has the smallest 2-norm* over the affine space $r_0 + A\mathcal{K}_m(A, r_0)$. Namely,

$$\|r_m\| = \min_{r \in r_0 + A\mathcal{K}_m(A,r_0)} \|r\| = \min_{u \in \mathcal{K}_m(A,r_0)} \|r_0 - Au\| \,.$$

In other words, the orthogonality condition

$$r_m \perp A\mathcal{K}_m(A, r_0)$$

needs to be satisfied at each GMRES iteration, resulting in the *increasing storage and time complexity of the method at every new step* (r_m needs to be orthogonolized against r_0, r_1, $\ldots r_{m-1}$).

Brief overview: GMRES with restarts or GMRES(m)

The GMRES(m) algorithm is based simply on restarting GMRES every m steps, using the latest iterate as the initial guess for the next GMRES run.

A single run of m GMRES iterations within the described framework is called a *GMRES cycle*. Thus, *GMRES with restarts is a sequence of GMRES cycles*.

In contrast to its restarted counterpart, we refer to the original method as *full* GMRES.

Convergence of full GMRES

A variety of results which characterize the convergence of *full* GMRES is presently available:

- For a normal matrix, the convergence is known to be linear and there exist convergence estimates governed solely by the spectrum of A (H. A. van der Vorst,C. Vuik 1993; V. Simoncini,D. Szyld 2005).
- For a diagonalizable matrix A, some characterizations of the convergence rely on the condition number of the eigenbasis (H. A. van der Vorst,C. Vuik 1993).
- Some estimates rely on the field of values of A (e.g. H .Elman 1982) or pseudospectra (L.N. Trefethen 1990).
- In general, *any nonincreasing convergence curve is possible for GMRES*, moreover the eigenvalues of A alone do not determine the convergence (Greenbaum, Pták, and Strakoš, 1996).

Motivation

While a lot of efforts have been put in the characterization of the convergence of full GMRES, *we have noticed that very few efforts have been made for characterizing the convergence of restarted GMRES.*

Our current research is aimed to *better understand restarted GMRES.*

Sublinear cycle-convergence: several experiments

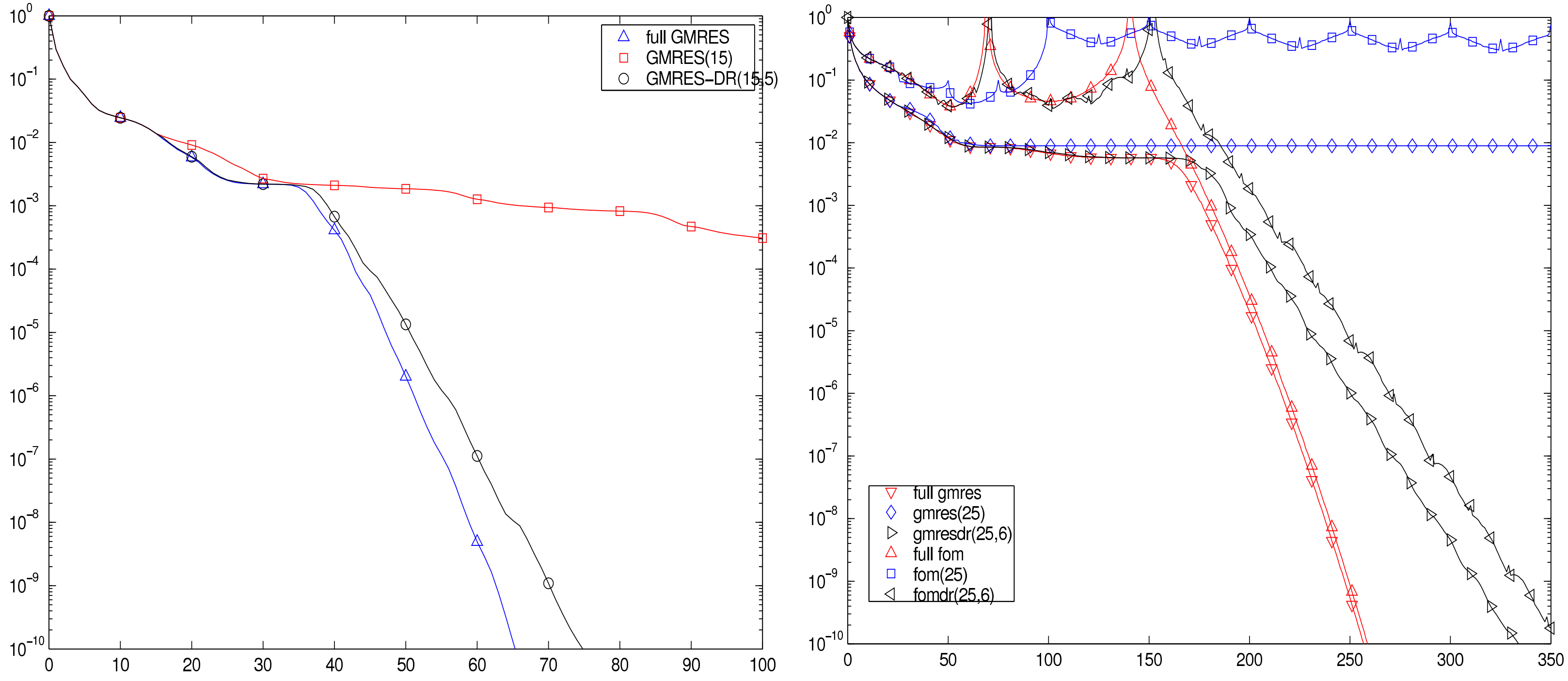

Figure 1: GMRES–DR(15,5), full GMRES and GMRES(15) are run on SAYLR4, a matrix of order 3564 from Matrix Market (**left**); comparison of MINRES solvers (GMRES's) and Galerkin projection solvers (FOM's) on the bidiagonal matrix with entries 0.01, 0.1, 1, 2, . . ., 997, 998 on the main diagonal and 1's on the super diagonal (**right**).

Sublinear cycle-convergence of GMRES(m) for normal matrices

Theorem 1 (The sublinear cycle-convergence of GMRES(m)) *Let $\{r_k\}$ be a sequence of nonzero residual vectors produced by GMRES(m) applied to the system $Ax = b$ with a nonsingular normal matrix $A \in \mathbb{C}^{n\times n}$, $1 \le m \le n-1$. Then*

$$\frac{\|r_k\|}{\|r_{k-1}\|} \le \frac{\|r_{k+1}\|}{\|r_k\|}, \quad k = 1, \ldots, q-1,$$

where q is the total number of GMRES(m) cycles.

In other words, any cycle-convergence curve of a restarted GMRES(m), applied to a system of linear equations with a nonsingular normal matrix A, is nonincreasing and convex (concave up).

The cycle-convergence is sublinear for normal matrices (1 of 2)

Consider 30 cycles of GMRES(20) applied to a *normal* 300×300 matrix A. The RHS vector b is randomly chosen. Spectrum of the matrix A is clustered around $-50 + 5i$.

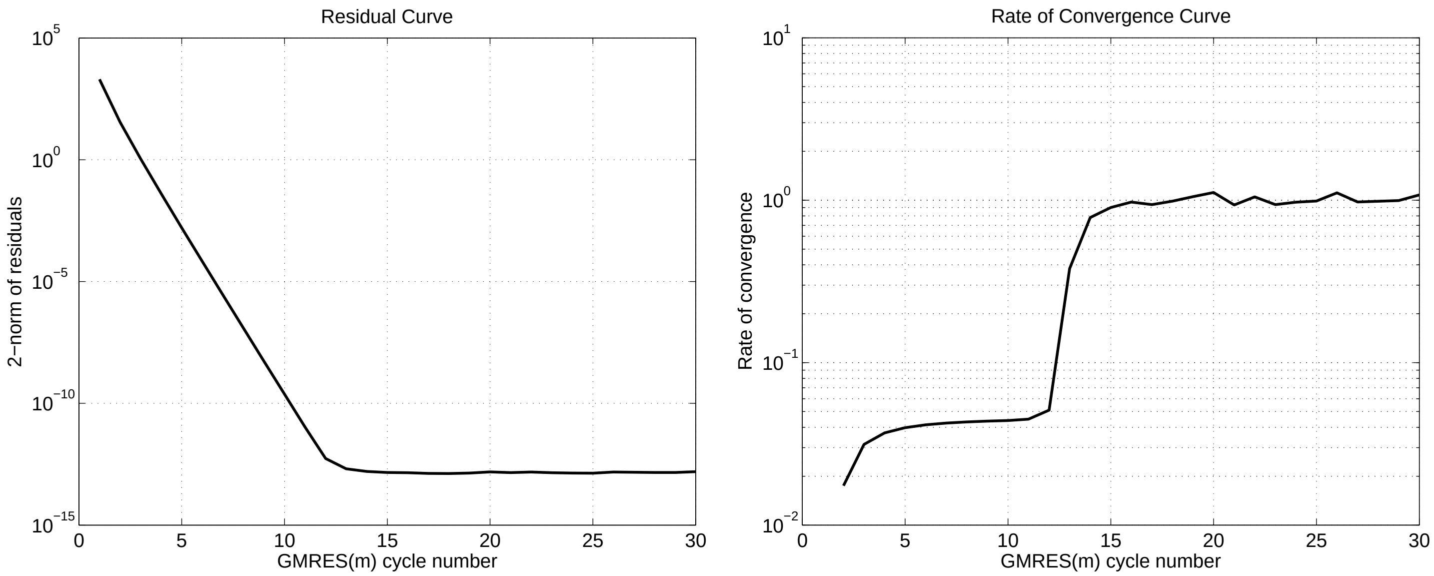

Figure 2: Residual (left) and rate of convergence (right) curves. A is normal, $n = 300$, $m = 20$.

The cycle-convergence is sublinear for normal matrices (2 of 2)

Consider 30 cycles of GMRES(10) applied to a *normal* 500×500 matrix A. The RHS vector b is randomly chosen. Spectrum of the matrix A is $\{k + ki,\ k = 1, \dots, n\}$.

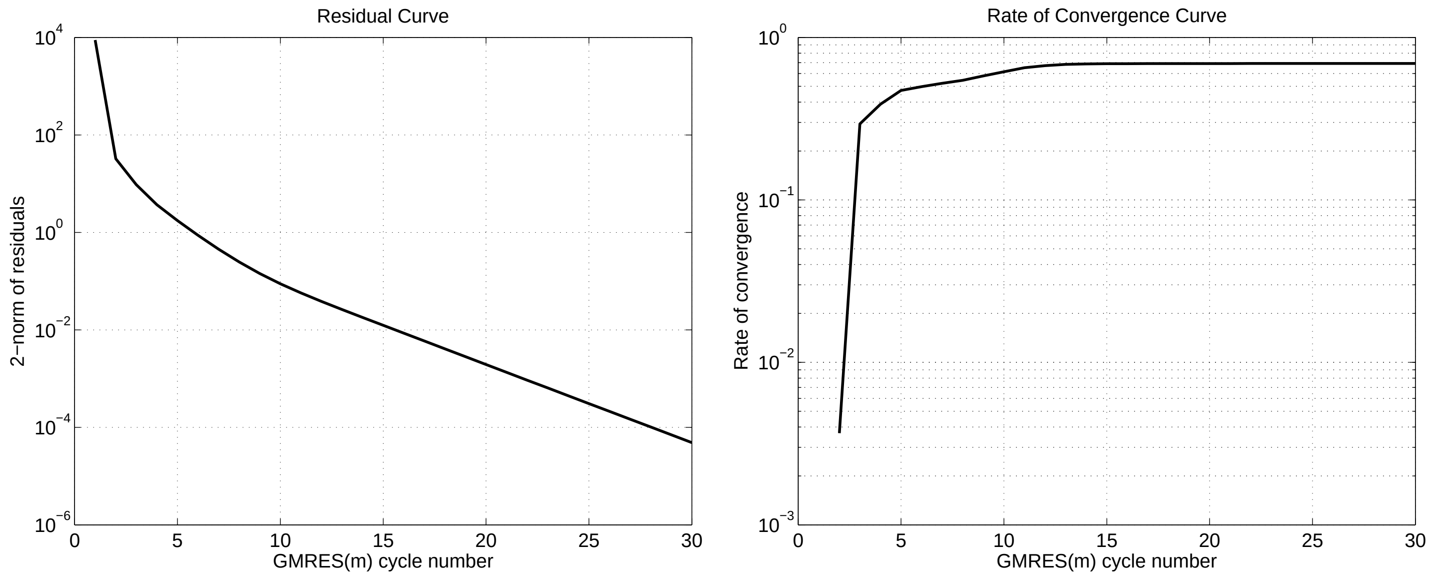

Figure 3: Residual (left) and rate of convergence (right) curves. A is normal, $n = 500$, $m = 10$.

Corollaries

Corollary 1 (The cycle-convergence of GMRES($n-1$)) *Let $A \in \mathbb{C}^{n\times n}$ be a nonsingular normal matrix. Let r_0 be the initial residual vector and r_1 - the residual vector at the end of the first GMRES($n-1$) cycle. Then*

$$\|r_k\| = \|r_1\| \left(\frac{\|r_1\|}{\|r_0\|}\right)^{k-1}, k = 2, 3 \ldots \tag{3}$$

Corollary 2 (The alternating residuals) *When $A \in \mathbb{C}^{n\times n}$ is Hermitian or skew-Hermitian and the restart parameter $m = n-1$, GMRES($n-1$) produces a sequence of residual vectors at the end of each restart cycle such that*

$$r_{k+1} = \alpha_k r_{k-1}, \alpha_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2} \in (0, 1], k = 1, 2, \ldots \tag{4}$$

The cycle-convergence is NOT necessarily sublinear for non-normal matrices

Consider 50 cycles of GMRES(10) applied to a *non-normal* 200×200 matrix A. The RHS vector b is randomly chosen. Spectrum of the matrix A is normally distributed on $[100, 300]$.

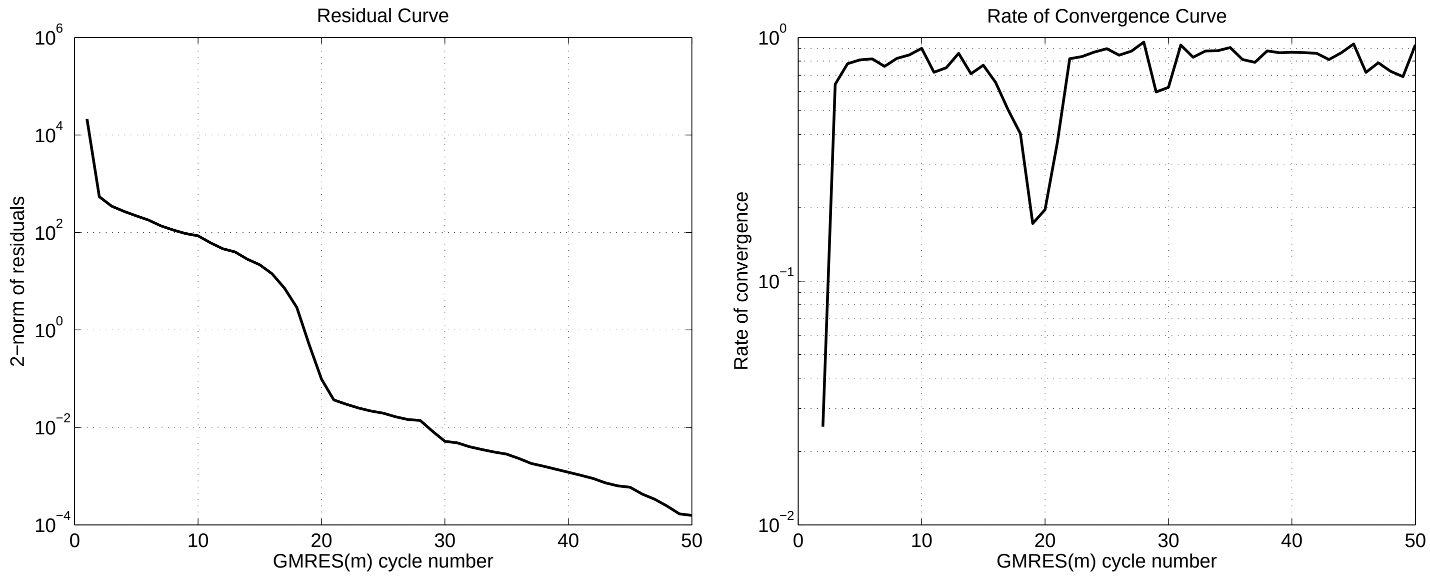

Figure 4: Residual (left) and rate of convergence (right) curves. A is non-normal, $n = 200$, $m = 10$.

The cycle-convergence of GMRES(m): departure from normality

Lemma 1 *Let* $\{r_k\}$ *be a sequence of nonzero residual vectors produced by GMRES(m) applied to the system* $Ax = b$ *with a nonsingular diagonalizable matrix* $A \in \mathbb{C}^{n \times n}$, $A = V \Lambda V^{-1}$, $1 \leq m \leq n-1$. *Then*

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq \frac{\alpha\left(\|r_{k+1}\| + \beta_k\right)}{\|r_k\|}, \quad k = 1, \ldots, q-1, \tag{5}$$

where $\alpha = \frac{1}{\sigma^2_{min}(V)}$, $\beta_k = \|p_k(A)(I - VV^H)r_k\|$, $p_k(z)$ *is the polynomial constructed at the cycle GMRES(A, m,* r_k*), and where* q *is the total number of GMRES(m) cycles. Note that as* $V^H V \longrightarrow I$, $0 < \alpha \longrightarrow 1$ *and* $0 < \beta_k \longrightarrow 0$.

The cycle-convergence of GMRES(m) in the general case (1 of 2)

Theorem 2 (Greenbaum, Pták, and Strakoš, 1996) *Given a nonincreasing positive sequence* $f(0) \geq f(1) \geq \cdots \geq f(n-1) > 0$, *there exists an* n*-by-*n *matrix* A *and a vector* r_0 *with* $\|r_0\| = f(0)$ *such that* $f(k) = \|r_k\|$, $k = 1, \ldots, n-1$, *where* r_k *is the residual at step* k *of the GMRES algorithm applied to the linear system* $Ax = b$, *with initial residual* $r_0 = b - Ax_0$. *Moreover, the matrix* A *can be chosen to have any desired eigenvalues.*

Theorem 3 *Given a matrix order* n, *a restart parameter* m $(m < n)$, *a decreasing positive sequence* $f(0) > f(1) > \ldots > f(q) \geq 0$, *where* $q < n/m$, *there exits an* n*-by-*n *matrix* A *and a vector* r_0 *with* $\|r_0\| = f(0)$ *such that* $\|r_k\| = f(k)$, $k = 1, \ldots, q$, *where* r_k *is the residual at cycle* k *of restarted GMRES with restart parameter* m *applied to the linear system* $Ax = b$, *with initial residual* $r_0 = b - Ax_0$. *Moreover, the matrix* A *can be chosen to have any desired eigenvalues.*

The cycle-convergence of GMRES(m) in the general case (2 of 2)

Several remarks:

- Theorem 3 is to restarted GMRES what Theorem 2 is to full GMRES.
- The proof we provide is constructive and directly inspired from the article of Greenbaum, Pták, and Strakoš, 1996; several specific difficulties ahead in the case of the restarted GMRES.
- We extend the result to the case of stagnating cycle-convergence curves and variable restart parameter.
- We generated two Matlab functions that correspond to Theorem 2 and Theorem 3. Given a matrix size, a restart parameter, a convergence curve and a spectrum, we construct the appropriate matrix and right-hand side. See: `http://www-math.cudenver.edu/~eugenev/edf.software/anycurve/`.

Summary

- The cycle-convergence of the restarted GMRES for normal matrices is sublinear.
- Any admissible cycle-convergence curve is possible for the restarted GMRES at a number of the initial cycles, and eigenvalues alone do not determine the cycle-convergence.

References:

1. E. Vecharynski and J. Langou. *The cycle-convergence of restarted GMRES for normal matrices is sublinear*. SIAM Journal on Scientific Computing, to appear.
(See also: A. Baker, E. Jessup, Tz. Kolev. *A simple strategy for varying the restart parameter in GMRES(m)*. J. of Comp. and Appl. Math., 2009)
2. E. Vecharynski and J. Langou. *Any decreasing cycle-convergence curve is possible for restarted GMRES*. Submitted for publication, 2009.

Thank you!