Convergence of a restarted and augmented GMRES method

Jan Zítko

Charles University, Faculty of Mathematics and Physics, Department of Numerical Mathematics, Sokolovská 83, 186 75 Praha 8, Czech Republic

`zitko@karlin.mff.cuni.cz`

Abstract

The bahaviour of the estimate for the quotient $\|r_{m+k}\|/\|r_0\|$ is investigated for GMRES(m,k) method. Sufficient conditions for convergence are formulated. All estimates are independent of the choice of an initial approximation.

Let us consider the GMRES method for solving the non-Hermitian and non-singular system

$$Ax = b, \quad A \in \mathbb{C}^{n \times n}, \quad x, b \in \mathbb{C}^n. \qquad (1)$$

In what follows, let x_0 be an initial approximation, $r_0 = b - Ax_0 \neq 0$ the associated residual, $v = r_0/\|r_0\|$ and $V_j = [Av, A^2v, \ldots, A^jv]$.

We will consider the GMRES(m,k) method, what means that to $\mathcal{K}_m(A, r_0)$ a space $\mathcal{Y}_k$ of dimension k is added.

If we carry out the GMRES(m,k) process, we practically perform the GMRES algorithm, but we manipulate with the space $\mathcal{K}_m(A, r_0) + \mathcal{Y}_k$ instead of $\mathcal{K}_m(A, r_0)$. (For more details see Morgan, Zítko.)

Put $s = m + k$ and let $\mathcal{Y}_k = span\{Y_k\}$, where $Y_k \in \mathbb{C}^{n \times k}$.

Assumption 1. Let all considered Krylov and augmented Krylov subspaces have a maximal dimension.

1. The first restarted run of GMRES(m,k). Conclusions for restarted GMRES.

In the first restart we usually put $y_j = A^{m+j-1}r_0$ for $j = 1, 2, \ldots, k$. Hence the estimate for $\|r_s\|^2/\|r_0\|^2$ is equivalent with the estimate for GMRES(s). The approximation $x_s \in x_0 + \mathcal{K}_s(A, r_0)$, $r_s = b - Ax_s$. We have $r_s = \|r_0\|v - q_s(A)v$, where $r_s \perp$ span$\{V_s\}$ and q_s is the polynomial of degree at most s such that $q_s(0) = 0$. (We will write $q_s \in \mathcal{P}_s^0$.)

It is $\|r(\alpha)\| \geq \|r_s\|$, where $r(\alpha) = \|r_0\|v - \alpha q(A)v$, $q \in \mathcal{P}_s^0$ and $\alpha \in \mathbb{C}$ and the same inequality holds for the vector r, where $r = \arg_{r(\alpha)} \min_{\alpha \in \mathbb{C}} \|r(\alpha)\|$. We obtain for the quotient $\|r_s\|^2/\|r_0\|^2$:

$$\|r_s\|^2/\|r_0\|^2 = \left(1 - \frac{|v^H q_s(A)v|^2}{\|q_s(A)v\|^2}\right) \qquad (2)$$
$$\leq \left(1 - \max_{q \in \mathcal{P}_s^0} \min_{w \in S_n} \frac{|w^H q(A)w|^2}{\|q(A)\|^2}\right)$$
$$= \left(1 - \max_{q \in \mathcal{P}_s^0} \min_{w \in S_n} \frac{|w^H H_q w|^2 + |w^H S_q w|^2}{\|(q(A))\|^2}\right).$$

The matrices H_q and $\mathrm{i}S_q$ denote the Hermitian and skew-Hermitian part of the matrix $q(A)$ respectively.

Remark 1 {for GMRES(s)}. As s can lie in the set $\{1, 2, \ldots, n-1\}$ we have the folowing result formulated for GMRES in real case by [Grcar].

Proposition 1. Let $s \in \{1, 2, \ldots, n-1\}$. If a polynomial q of degree s with $q(0) = 0$ exists such that

$$\min_{w \in S_n} |w^H H_q w| > 0, \text{ or } \min_{w \in S_n} |w^H S_q w| > 0.$$

Then GMRES(j) is convergent for all $j \geq s$, i.e., the iterations converge to the unique solution of **(1)**.

Remark 2. It holds for arbitrary $x \in S_n$ and $q \in \mathcal{P}_s^0$:

$$x^H q^2(A)x = x^H(H_q + \mathrm{i}S_q)^2 x = \|H_q x\|^2 - \|S_q x\|^2 + \mathrm{i}2Re(x^H H_q S_q x).$$

Now we apply the original results of [Szyld and Simoncini] for the polynomial $q(A)$ in the complex case and we obtain the following result. Let H_q be nonsingular and $\|S_q H_q^{-1}\| < 1$ or S_q be nonsingular and $\|H_q S_q^{-1}\| < 1$. Then $Re(x^H q^2(A)x) > 0$ for all $x \in S_n$ and hence GMRES(j) is convergent for all $j \geq 2s$.

2. Augmented GMRES method

Let $y \in \mathbb{C}^n$ is added to $\mathcal{K}_m(A, r_0)$. The iteration x_{m+1} is constructed in $x_0 + \mathcal{K}_m(A, r_0) + span\{y\}$. In this case $s = m + 1$. Let the corresponding residual vector r_{m+1} be written in the form

$$r_{m+1} = \|r_0\|v - (q_m(A)v + \beta_{q_m}Ay), \qquad (3)$$

where the coefficients of the polynomial $q_m \in \mathcal{P}_m^0$ and $\beta_{q_m} \in \mathbb{C}$ are determined such that $r_{m+1} \perp span\{V_m, Ay\}$. Especially, $r_{m+1} \perp (q_m(A)v + \beta_{q_m}Ay)$. Hence for arbitrary polynomial $q \in \mathcal{P}_m^0$ and $\beta \in \mathbb{C}$ we have $\|r_{m+1}\| \leq \|r\|$ where $r = \|r_0\|v - q(A)v - \beta Ay = (\underbrace{\|r_0\|I - q(A)}_{p(A)})v - A\underbrace{\hat{y}}_{\beta y}$ and $p(0) = \|r_0\|$. The last relations yield the following theoretical theorem.

Proposition 2. *Let $m \in \{1, 2, \ldots, n-1\}$ and p be a polynomial of degree at most m, $p(0) = \|r_0\|$. If the vector $\hat{y} \in \mathbb{C}^n$ solves the equation*

$$A\hat{y} = p(A)v. \qquad (4)$$

Then $r_{m+1} = r = 0$.

Similar formulation is given by [Saad.] But to solve the equation **(4)** is the problem equivalent with the original one. Let $\mathcal{Y}_k \subset \mathbb{C}^n$ be added to $\mathcal{K}_m(A, r_0)$. We have, roughly speaking, two opportinities in practice:

The space $\mathcal{Y}_k$ is chosen arbitrary, or

The space $\mathcal{Y}_k$ is an invariant subspace.

Bacause the estimates are formally identical, we will suppose that $\mathcal{Y}_k$ is an invariant subspace. All estimates will be considered after the second restart.

3. Let an invariant subspace $\mathcal{Y}_k = span\{Y_k\}$ be added to $\mathcal{K}_m(A, r_0)$ in all restarted runs.

After the first restart we use the fact that that the new initial residual $r_0 \perp span\{Y_k\}$, i.e. $r_0^H Y_k = v^H Y_k = 0$ and $r_s \perp span\{V_m, AY_k\}$. Because $span\{AY_k\} = span\{Y_k\}$, we can take the matrix $P = U(U^H U)^{-1}U^H$, where $U = [q(A)v, Y_k]$ for the orthogonal projector for the space $span\{q(A)v, AY_k\}$. We have

$$\|r_s\|^2 = \|r_0\|^2 - r_0^H P r_0$$
$$= (1 - (U^H v)^H (U^H U)^{-1} U^H v)\|r_0\|^2.$$

However $U^H v =$

$$\begin{bmatrix} (q(A)v)^H v \\ Y_k^H v \end{bmatrix} = \begin{bmatrix} (q(A)v)^H v \\ 0 \text{ (dim } k) \end{bmatrix} = ((q(A)v)^H v)e_1$$

and hence $\|r_s\|^2/\|r_0\|^2$

$$\leq 1 - |v^H q(A)v|^2 e_1^T (U^H U)^{-1} e_1 \leq 1 - |v^H q(A)v|^2/\|q(A)\|^2$$

where $v \perp span\{Y_k\}$. Let $\mathcal{L}_{n-k} = \{w \in \mathbb{C}^n | w^H Y_k = 0\}$, $S_{n-k} = \{S_n \cap \mathcal{L}_{n-k}\}$.

Theorem 1. *Let the invariant subspace $span\{Y_k\}$ be added to the corresponding Krylov subspace for all restarted runs. Then for all restarts, excepts for the first one, we have the estimate*

$$\frac{\|r_s\|^2}{\|r_0\|^2} \leq 1 - \max_{q \in \mathcal{P}_m^0} \min_{w \in S_{n-k}} \frac{|w^H q(A)w|^2}{\|q(A)\|^2} \qquad (5)$$

Remark 3. The same estimate is obtained if arbitrary vectors are added to the Krylov subspace in all restarts. In this case we consider minimum for $w \in S_n, w^H AY_k = 0$ in the last formula.

Theorem 2. *Let $m, k, s \in \{1, 2, \ldots, n-1\}$, $s = m + k < n$. Let the invariant subspace $span\{Y_k\}$ be added to the corresponding Krylov subspace for all restarted runs. Let a polynomial $q \in \mathcal{P}_m^0$ exists such, that the system of equations*

$$w^H q(A)w = 0 \qquad (6)$$
$$w^H Y_k = 0 \qquad (7)$$

has only the solution $w = 0$ in C^n (or equivalently does not have any solutin on S_n.) Then GMRES(m,k) is convergent.

Conclusion. Let the equation **(6)** has a nontrivial solution, i.e., $0 \in W(q(A))$- field of values of the matrix $q(A)$.
Let $\mathcal{M} = \{W(q(A)) \cap S_n | w^H q(A)w = 0\}$. The condition **(7)** offers to find Y_k such that $u^H Y_k \neq 0 \forall u \in \mathcal{M}$ and therefore make do the quotient in **(5)** less than 1.

4. Let a nearly invariant subspace be added to $\mathcal{K}_m(A, r_0)$ in all restarted runs.

Let $\Theta(X, Y)$ denotes the gap between subspaces X and Y. Let the space $\mathcal{Y}_k$ be added to $\mathcal{K}_m(A, r_0)$ in some restart and the residual r_s be obtained after one restarted run. Let us write $r_0^+ = r_s$ and let the space $\mathcal{Y}^+$ be added to $\mathcal{K}_(A, r_0)$ in the next restart with the residual vector r_s^+. We have $v^+ = r_{0+}/\|r_{0+}\|$, $v^+ \perp A\mathcal{Y}_k$. Let $\mathcal{Y}$ be an invariant subspace. Let $\mathcal{L}_{n-k} = \mathcal{Y}^\perp$ and

$\widehat{S}_{n-k} = \{v \in \mathbb{C}^n, \|v\| = 1 | \sin\angle(v, \mathcal{L}_{n-k}) \leq \varepsilon\}$

Theorem 3. Let $\varepsilon \in (0, 1)$, $\Theta(A\mathcal{Y}_k, \mathcal{Y}) \leq \varepsilon$ and $\Theta(A\mathcal{Y}_k^+, \mathcal{Y}) \leq \varepsilon$. Then

$$\frac{\|r_{s+}\|^2}{\|r_{0+}\|^2} \leq 1 - \max_{q \in \mathcal{P}_m^0} \min_{w \in \widehat{S}_{n-k}} \frac{|w^H q(A)w|^2}{\|q(A)\|^2}$$

5. Conclusions.

Restarting slows down the convergence and trouble may be caused by the presence of the eigenvalues nearest to zero. These are bad because it is impossible to have a polynomial p of degree n such that $p(0) = 1$ and $|p(z)| < 1$ on any Jordan curve around the origin (see [Greenbaum, p55]). Usually, an eigenspace corresponding to the smallest eigenvalues is taken for $\mathcal{Y}_k$ and the corresponding algorithms give good results ([Morgan, Saad]). If GMRES(m) is convergent and if arbitrary vectors are add to $\mathcal{K}_m(A, r_0)$, then the numerical results are not in many cases essentially better, i.e. we do not improve the convergence. In this theoretical paper, an arbitrary invariant subspace is considered and the question ”to find some sufficient condition for convergence” is transformed to the question whether the intersection of fields of values and sets of the form **(7)** contains zero or not. The field of values are considered for vectors in spaces perpendicular to a chosen invariant space.

The open question 1. Given a polynomial q. To estimate all solutions of the equation $w^H q(A)w = 0$ on S_{n-k} and analogously for the equations $w^H H_q w = 0; w^H S_q w = 0;$

The open question 2. To find min$\{|z| : z \in \mathcal{F}(\mathcal{A})\}$ and analogously for H_q and S_q.

The open question 3. To construct, for special matrices and polynomials, the behaviour of the integer function $f(j) = 1 - \max_{q_j \in \mathcal{P}_j^0} \min_{w \in S_{n-k}} \frac{|w^H qj(A)w|^2}{\|q_j(A)\|^2}$, $j \in [1, \mathrm{restart}]$ and compare $f(j)$ with the behaviour of the sequence $\{\frac{\|r_j\|^2}{\|r_0\|^2}\}_{j=1}^{\mathrm{restart}}$. (This would be the answer on the question how descriptive are these bounds.)

The open question 4. How to find very fast an inexact approximate solution of **(4)?**

References

[1] R. B. Morgan: A restarted GMRES method augmented with eigenvectors. *SIAM Journal on Matrix Analysis and Aplication 1995; 16(4): 1154–1171*.

[2] J. Zítko: Convergence conditions for a restarted GMRES method augmented with eigenspaces. *Numerical Linear algebra with Application 2005; 12: 373–390*.

[3] J. F. Grcar: Operator coefficient methods for linear equations. A restarted GMRES method augmented with eigenvectors. *Technical Report SAND89-8691, Sandia National Laboratories, November 1989. Available at* `http://seesar.lbl.gov/ccse/Publications/sepp/ocm/SAND89-8691.pdf`.

[4] V. Simoncini, D. B. Szyld: New conditions for non-stagnation of minimal residual methods *Report 07-4-17, April 2007. Available at* `http://www.math.temple.edu/ szyld`.

[5] Y. Saad: Analysis of augmented Krylov subspace methods. *SIAM Journal on Matrix Analysis and Aplication 1997; 18(2): 435-449*.

[6] A. Greenbaum: Iterative methods for solving linear systems. *SIAM, Philadelphia 1997.*