On control strategy for stagnated GMRES(m)

Juan Carlos Cabral[1 2], Pedro Torres[1] and Christian E. Schaerer[1,3]
[1] LCCA, National University of Asuncion, Paraguay.
P.O.Box: 2111 SL. Email: jccabral19@gmail.com, pjtorres@pol.una.py, cschaer@pol.una.py
[3] CIMA, Dr. Cesar López Moreira 693, Paraguay. P.O.Box: 1766. Email: chris.schaerer@cima.org.py

Abstract

In this work, GMRES(m) method is formulated as a dynamic discrete system with feedback control and a condition for the convergence of the method is discussed. Using the proposed formulation, deflation and augmentation techniques are introduced in the context of GMRES(m) for avoiding the stagnation in the convergence process. The resulting method is tested on matrices based on real problems showing that the method overcome the stagnation problem.

Keywords: Krylov subspace, restarted GMRES, control strategy, augmentation, deflation.

1. Introduction

The GMRES is a popular iterative method for solving the large nonsymmetric system of linear equations

$$Ax = b; \qquad A \in \mathbb{C}^{N\times N}; \qquad x, b \in \mathbb{C}^{N\times 1}. \tag{1}$$

To limit the computational cost and storage requirements, the method is usually restarted after a fixed number of iterations, in our case we take the fixed value m. The resulted method is called restarted GMRES method, denoted by GMRES(m).

A drawback with the GMRES(m) algorithm is that its rate of convergence can deteriorate and even stagnate [3, 6]. The full GMRES algorithm is guaranteed to converge in at most N steps, but this would be impractical if there were many steps required for convergence [12]. In the first part of this work we present a control formulation for the restarted GMRES and in the

[2]Corresponding author: jccabral19@gmail.com

second we use a technique, developed in the context of the acceleration for convergence of full GMRES, but now in the context of restarted GMRES.

This work is organized as follows. In §2, we formulate the GMRES(m) as a dynamical system with feedback control and some consideration about the convergence of the method are discussed. In §3, it is presented the control formulation of the deflated augmented method for GMRES(m) and numerical results are presented at §4. The conclusion are presented at §5 showing that we can improve the convergence using deflated augmented technique and recover the global optimality lost due to restarting.

2. GMRES(m) as dynamical systems with feedback control. GMRES(m) approximate the solution to system (1) at kth iteration from a residual $r^{(k-1)}$ at the iteration $(k-1)$ which is used to construct a Krylov subspace of dimension m. At the kth iteration it is obtained the vector $x^{(k)}$ that solves the least squares problem

$$\min_{x^{(k)} \in K(A,v)} \| b - Ax^{(k)} \|_2 \tag{2}$$

over the Krylov subspace $\mathcal{K}_m(A, v^{k-1}) = span\{v^{k-1}, Av^{k-1}, \dots A^{m-1}v^{k-1}\}$. To solve this problem, an orthonormal basis for the Krylov subspace is developed through the Arnoldi process. The first m steps of this procedure can be collected into the relationship

$$AV_m^{k-1} \quad = \quad V_{m+1}^{k-1}\tilde{H}_m^{k-1} \tag{3}$$

$$AV_m^{k-1} \quad = \quad V_m^{k-1}H_m^{k-1} + h_{m+1,m}^{k-1}v_m^{k-1}e_m^T \tag{4}$$

where $V_m^{k-1} \in \mathbb{C}^{N\times m}$ and $V_{m+1}^{k-1} := [V_m^{k-1}v_{m+1}^{k-1}] \in \mathbb{C}^{N\times(m+1)}$ have orthonormal columns and $\tilde{H}_m^{k-1} \in \mathbb{C}^{(m+1)\times m}$ is upper Hessenberg with upper $m \times m$ block H_m^{k-1} and $(m+1, m)$ entry $h_{m+1,m}^{k-1}$. If the Arnoldi process starts with $v^{k-1} = r^{(k-1)}/ \| r^{(k-1)} \|_2$, then $\mathcal{K}_m(A, v^{k-1}) = span(V_m^{k-1})$.
The Petrov - Galerkin condition on GMRES(m) at the kth iteration is given by $r^{(k)} \perp A\mathcal{K}_m^{(k-1)}(A, r^{(k-1)})$, which in terms of expression (3), it can be considered as

$$(r^{(k)}, AV_m^{(k-1)}) = 0. \tag{5}$$

In addition, considering that

$$x^{(k)} \in x^{(k-1)} + \mathcal{K}_m^{(k-1)}(A, r^{(k-1)}) \tag{6}$$

then it is obtained

$$x^{(k)} = x^{(k-1)} + V_m^{(k-1)} y^{(k-1)}. \tag{7}$$

Observing that

$$\begin{array}{rcl} 0 & = & V_m^{(k-1)*} A^* r^{(k)} = V_m^{(k-1)*} A^* (b - Ax^{(k)}) \\ 0 & = & V_m^{(k-1)*} A^* r^{(k-1)} - V_m^{(k-1)*} A^* A V_m^{(k-1)} y^{(k-1)} \end{array}$$

then we have

$$y^{(k-1)} = (V_m^{(k-1)*} A^* A V_m^{(k-1)})^{-1} V_m^{(k-1)*} A^* r^{(k-1)} \tag{8}$$

or equivalently for $x^{(k)}$,

$$x^{(k)} = x^{(k-1)} + V_m^{(k-1)} (V_m^{(k-1)*} A^* A V_m^{(k-1)})^{-1} V_m^{(k-1)*} A^* r^{(k-1)} \tag{9}$$

which in terms of the residual the iterative process is

$$r^{(k)} = (I - A V_m^{(k-1)} (V_m^{(k-1)*} A^* A V_m^{(k-1)})^{-1} V_m^{(k-1)*} A^*) r^{(k-1)}. \tag{10}$$

Control formulation. In terms of control theory, Iterative methods for linear systems can be represented as [2]:

$$\left\{ \begin{array}{lcl} x^{(k)} & = & x^{(k-1)} + u^{(k-1)}, \\ r^{(k-1)} & = & b - Ax^{(k-1)}, \\ u^{(k-1)} & = & f(r^{(k-1)}). \end{array} \right. \tag{11}$$

In words, it is desired to choose the control signal $u^{(k-1)}$ in order to zero the steady-state error by choice of the control law $f(\cdot)$, that is, $\lim\limits_{k\to\infty} r^{(k)} = 0$, or, equivalently, $\lim\limits_{k\to\infty} Ax^{(k)} = b$. Notice that the control law $f(\cdot)$ has to be chosen such that the asymptotic stability of the closed- loop system (11) implies the convergence of the residue $r^{(k)}$ to 0.

Comparing both formulations (and (11)) it is observed that GMRES(m) can be expressed as:

$$x^{(k)} = x^{(k-1)} + K^{(k-1)} r^{(k-1)} \tag{12}$$

where: matrix $K^{(k-1)}$ corresponds to a gain feedback gain matrix, $u^{(k-1)} := K^{(k-1)} r^{(k-1)}$ is the control signal and k is the iteration counter.
Analogously to equation (10), using the expression (12) it is obtained:

$$r^{(k)} = (I - AK^{(k-1)}) r^{(k-1)} \tag{13}$$

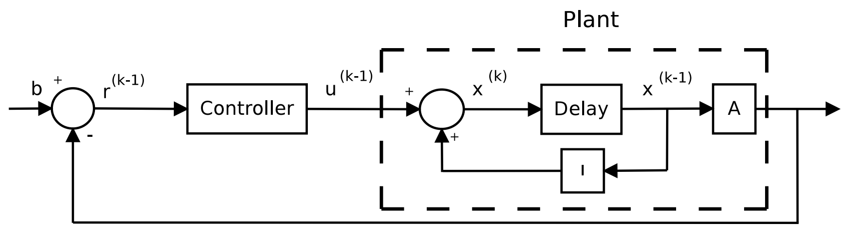


Figure 1: Block diagram of linear system $Ax = b$..

where by comparison matrix $K^{(k-1)}$ has the form

$$K^{(k-1)} = V_m^{(k-1)}(V_m^{(k-1)*}A^*AV_m^{(k-1)})^{-1}V_m^{(k-1)*}A^* \tag{14}$$

where to emphasize the dependence of matrix $K^{(k-1)}$ of the restarting parameter m, it is adopted the notation $K_m^{(k-1)}$.

To obtain the convergence of the corresponding iteration method, it is necessary to assure the Schur stability of matrix S, where

$$S = I - AK_m^{(k-1)}$$

and where the eigenvalues of this matrix are given by the solution of the characteristic equation

$$\det(I\lambda - I + AK_m^{(k-1)}) = 0.$$

Observe that the choice of matrix $K_m^{(k-1)}$ affects directly to the convergence of the method. A crucial problem here is to find a rule for obtaining a matrix $K_m^{(k-1)}$ for obtaining an adequate for a vast class of matrices A. More specifically, in practice the GMRES(m) can have its rate of convergence deteriorated or even its convergence stagnated if the matrix $K_m^{(k-1)}$ is not adequately chosen.

3. Modifying Krylov subspace using deflated and augmented method.

In this section, we use in the context of GMRES(m) two approaches used for the acceleration of the convergence of the full GMRES. The techniques are denoted as deflation and augmentation. In the deflation technique the system

(1) is multiplied with a suitably chosen projection with the objective to eliminate components that slow down convergence, while in the augmentation technique, the search space of the Krylov subspace method $\mathcal{K}_m(A, v^{k-1})$ is enlarged by a suitably chosen subspace. A typical goal is to add information about the problem to the search space that is slowly revealed in the Krylov subspace itself [10].

Let us for the moment consider an arbitrary possibly singular $\hat{A} \in \mathbb{C}^{N\times N}$ and a arbitrary vector $\hat{v} \in \mathbb{C}^N$, such that the Krylov subspace $\mathcal{K}_m(\hat{A}, \hat{v})$ has dimension m. We focus from now an augmented Krylov subspace of the form

$$\mathcal{S}_m = \mathcal{K}_m^{(k-1)}(\hat{A}, \hat{v}^{(k-1)}) + \mathcal{U} \tag{15}$$

$$x^{(k)} \in x^{(k-1)} + \mathcal{S}_m \tag{16}$$

so that the corresponding residual is

$$r^{(k)} \perp A\mathcal{S}_m. \tag{17}$$

We suppose that $\mathcal{U}$ has dimension d, $0 < d < N$, and denote by $U \in \mathbb{C}^{N\times d}$ a matrix whose columns form a basis of $\mathcal{U}$, and by $V_m^{(k-1)} \in \mathbb{C}^{N\times m}$ one whose columns form a basis of $\mathcal{K}_m^{(k-1)}(\hat{A}, \hat{v}^{(k-1)})$, so that (16) can be written as

$$x^{(k)} = x^{(k-1)} + V_m^{(k-1)} y^{(k-1)} + U u^{(k-1)} \tag{18}$$

for some vectors $y^{(k-1)} \in \mathbb{C}^m$ and $u^{(k-1)} \in \mathbb{C}^d$.

To satisfy (17), the residual $r^{(k)} = b - Ax^{(k)} = r^{(k-1)} - AV_m^{(k-1)} y^{(k-1)} - AUu^{(k-1)}$ must be orthogonal to both $A\mathcal{K}_m^{(k-1)}(\hat{A}, \hat{v}^{(k-1)}$ and $A\mathcal{U}$, hence it must satisfy the pair of orthogonality conditions

$$r^{(k)} \perp A\mathcal{K}_m^{(k-1)}(\hat{A}, \hat{v}^{(k-1)}) \qquad \textit{and} \qquad r^{(k)} \perp A\mathcal{U}. \tag{19}$$

Using the first condition of (19) it is determined $y^{(k-1)}$ (see expression (8)). The second condition of (19) can be written as

$$0 = U^* A^* r^{(k)} = U^* A^* (r^{(k-1)} - AK^{(k-1)} r^{(k-1)} - AUu^{(k-1)}) =$$

$$= U^* A^* (I - AK^{(k-1)}) r^{(k-1)} - U^* A^* AUu^{(k-1)}$$

where

$$E_A := U^* A^* AU \in \mathbb{C}^{d\times d}. \tag{20}$$

We assume that E_A is nonsingular, then the second orthogonality gives

$$u^{(k-1)} = E_A^{-1} U^* A^* (I - AK^{(k-1)}) r^{(k-1)}. \tag{21}$$

Substituting this into (18) and defining $M_A := U E_A^{-1} U^*$ we obtaing

$$\begin{aligned} x^{(k)} &= x^{(k-1)} + K^{(k-1)} r^{(k-1)} + U E_A^{-1} U^* A^* (r^{(k-1)} - AK^{(k-1)} r^{(k-1)}) \\ &= (I - M_A A^* A)(x^{(k-1)} + K^{(k-1)} r^{(k-1)}) + M_A A^* b \end{aligned} \tag{22}$$

and

$$\begin{aligned} r^{(k)} &= r^{(k-1)} - AK^{(k-1)} r^{(k-1)} - AM_A A^* (r^{(k-1)} - AK^{(k-1)} r^{(k-1)}) \\ &= (I - AM_A A^*)(r^{(k-1)} - AK^{(k-1)} r^{(k-1)}) \end{aligned} \tag{23}$$

To simplify the notation we define the $N \times N$ matrices: $P_A := I - AM_A A^*$ and $Q_A := I - M_A A^* A$. Hence equations (22) and (23) take the form

$$x^{(k)} = Q_A (x^{(k-1)} + K^{(k-1)} r^{(k-1)}) + M_A A^* b \tag{24}$$

and

$$r^{(k)} = P_A (r^{(k-1)} - AK^{(k-1)} r^{(k-1)}). \tag{25}$$

Let $A, \in \mathbb{C}^{N \times N}$ and $U \in \mathbb{C}^{N \times d}$ be such that $E_A = U^* A^* A U$ is nonsingular (which implies that rank U=d). Then the matrices M_A, P_A and Q_A are well defined and the following statements hold [8]:

1. P_A is the orthogonal projection onto $(A\mathcal{U})^\perp$ along $A\mathcal{U}$.
2. Q_A is the projection onto $(A^* A\mathcal{U})^\perp$ along $\mathcal{U}$.
3. $P_A A = A Q_A$.

In practice, $\hat{A}$ and $\hat{v}$ should be somehow related to A, however. One specific choice is suggested in [8]

$$\hat{A} := P_A A, \qquad \hat{v} := P_A (b - Ax^{(k)}) \qquad \textit{and} \qquad \hat{b} := P_A b.$$

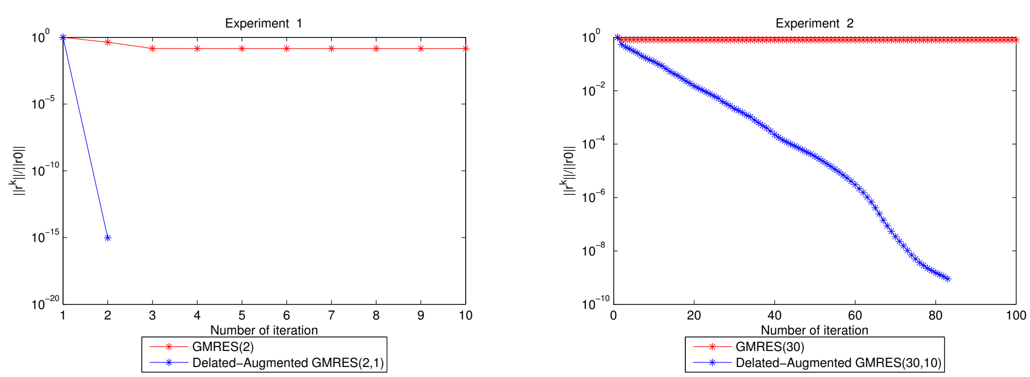
In this way our proposal of Deflated-Augmented GMRES(m, d) have the Arnoldi matrix $V_m^{(k-1)}$ which is a $N \times m$ matrix where its columns form an orthogonal basis of the Krylov Subspace $\mathcal{K}_m^{(k-1)}(\hat{A}, \hat{v}^{(k-1)})$ and the matrix U is formed by the d-eigenvectors corresponding to the smallest eigenvalues.

Table 1: Matrices information used in numerical experiments.

Matrix	N	nnz	Cond	λ_1	λ_2	λ_3
Sherman5	3312	20793	3,90E+05	4.6925E-02	1.25457E-01	4.0266E-02
Morgan2	1000	998001	4,72E+06	1E-02	2E-02	3E-02
Cavity10	2597	76171	4,46E+06	4.305E-06	4.3058E-04	4.3056E-04

4. Numerical experiments.

In this section, we present some examples about the numerical behavior of GMRES(m) and Deflated-Augmentation GMRES(m, d) introduced above. We emphasize in the phenomenon a of stagnation based on matrices from Matrix Market Collection [9] and [8]. Algorithms initial configurations are: the initial solution is $x_0 = 0$, the stopping criterion is $\frac{\|r^{(k)}\|}{\|r^{(0)}\|} \leq 10^{-9}$ and the maximum number of iterations is 100. The number of restarts in experiments 2, 3 and 4 is $m = 30$, which it is similar to the values used in [7, 11]. These matrices are outlines in Table 1, where N is the size of A, nnz is the number of nonzero elements, $Cond$ is the condition number and λ_i, $i = 1, 2, 3$, refer to the estimation of the three smallest eigenvalues, respectively.

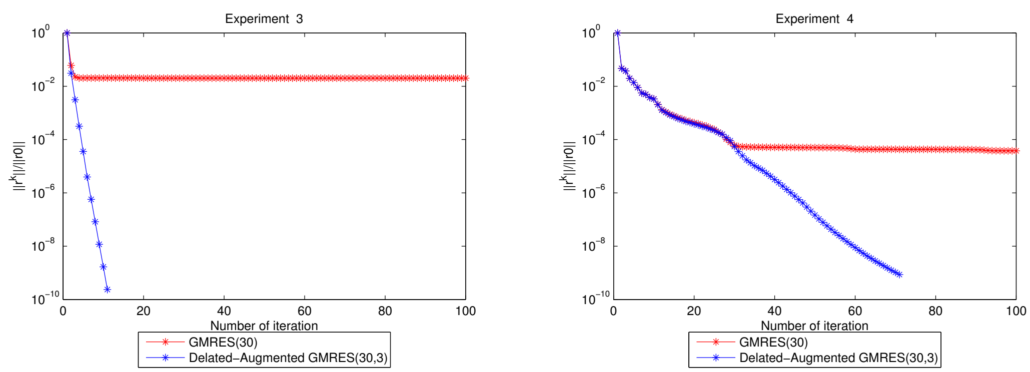


(a) (b)

Figure 2: The convergence curves of relative residual norm. (a) Embree matrix extracted from [3] and (b) Sherman 5 problem extracted from [9].

Experiment 1. Simple example showing stagnation problem in the convergence of GMRES(2). We choose the square matrix presented in [3], with $N = 3$ and eigenvalues $\{1, 2, 3\}$ for which restarted GMRES also produces an extreme behavior. This problem is well know that GMRES(1) converges but GMRES(2) does not. The residual behavior is shown at Figure 2 (a). Observe that although the restarting parameter is relative large in comparison with the dimension of the matrix, this does not assure an improvement in the convergence of the method, however, the deflation - augmentation part, improves drastically its convergence.

Experiment 2. Benchmark problem from computational fluid dynamics. In this case we consider a benchmark problem: *Sherman* 5 [9]. This is a difficult real non-symmetric matrix with smallest eigenvalues (very close to zero). It is observed that GMRES stagnates at just at beginning of the iteration process, however the deflated - augmented GMRES(30,10) maintains the rate of convergence almost linearly up to the pre-specified tolerance.



(a) (b)

Figure 3: The convergence curves of relative residual norm. *Morgan* 2 extracted from [10] and (b) *Cavity10* extracted from [9].

Experiment 3. For this, we consider the matrix named as *Morgan* 2 extracted from [10]. This matrix has some very small eigenvalues. The convergence results is reported in Figure 3 (a). It is observed that the deflated-augmented GMRES(30, 3) exhibit a constant rate of convergence, whereas GMRES(30) present a stagnation after the second iteration ($k > 2$).

Experiment 4. The matrix is *Cavity10* extracted from [9]. This problem has a real non-symmetric matrix with a relative small eigenvalue in magnitude, arising from finite element modeling [9]. It can be observed at Figure 3 (b) that the overall process of GMRES(30) maintain the rate of convergence up to a certain tolerance, however below of this tolerance, the rate of convergence deteriorates and the method stagnates. By its turns, deflated - augmented GMRES(30,3) maintains the rate of convergence up to the pre-specified tolerance.

5. Conclusions.

In this paper we present the feedback control formulation for restarted GMRES and deflated - augmented restarted GMRES. Future research are orientated to understand from this perspective the robustness of the proposal, since using only the restarted parameter, a priori convergence of the restarted GMRES can not be assured in terms of stagnation. Therefore, a deflation - augmentation is used successfully. This technique introduces another question about how large must to be the parameter d of the deflated - augmented part and how sensible (robust) is the method to its variation.

Acknowledgments: JCC acknowledges the financial support given by CONACyT through scholarship Prociencia-2015 and CES thanks PRONII-CONACyT. This work was partially supported by 14-INV-186 CABIBESKRY - PROCIENCIA Project

References

[1] Bhaya, A., and Kaszkurewicz, E. (2003, December). Iterative methods as dynamical systems with feedback control. In 42nd IEEE Conf. on Decision and Control (pp. 2347-2380).

[2] Bhaya, A., and Kaszkurewicz, E. (2003, December). Control Perspective on Numerical Algorithms and Matrix Problems, Advances in Design and Control, Vol.DC10 SIAM: Philadelphia, 2006.

[3] M. Embreé, "The tortoise and the hare restart GMRES", SIAM J. on Num. Analysis, 2003.

[4] Schaerer, C. and Kaszkurewicz, E., 2001, The shooting method for the solution of ordinary differential equations: a Control-Theoretical Perspective, "Internat. J. Systems Science", Vol. 32, No. 8, pp. 10471053.

[5] W. D. Joubert and T. A. Manteuffel. Iterative methods for nonsymmetric linear systems. In Iterative Methods for Large Linear Systems, D. R. Kincaid and L. J. Hayes, editors, pages 149-171. Academic Press, Boston MA, 1990.

[6] W. Joubert , On the Convergence Behavior of the Restarted GMRES Algorithm for Solving Nonsym- metric Linear Systems , Numerical Linear Algebra with Applications, Volume 1, Issue 4, 1994.

[7] Cuevas, R., Schaerer, C.E., Bhaya, A., *A proportional-derivative control strategy for varying the restart parameter in GMRES(m)*, Anais do XXXIII CNMAC, V.3, pp. 1000-1001, 2010.

[8] Gaul, A., Gutknecht, M. H., Liesen, J., Nabben, R. A framework for deflated and augmented Krylov subspace methods. SIAM Journal on Matrix Analysis and Applications 34.2 (2013): 495-518.

[9] T. Davis and Y. Hu, *The University of Florida Matrix Collection. Univ. of Florida and ATT Research.* Sitio web: http://www.cise.ufl.edu/research/sparse/matrices/.

[10] R. B. Morgan, *A restarted GMRES methods augmented with eigenvectors*, SIAM J. MATRIX ANAL. APPL. Vol. 16, No. 4, pp. 1154-1171, October 1995.

[11] Qiang Niu, and Linzhang Lu, *Restarted GMRES method Augmented with the Combination of Harmonic Ritz Vectors and Error Approximations*, International Journal of Mathematical and Computational Sciences 6, 2012.

[12] Yousef Saad, *Iterative Methods for Sparse Linear Systems*,SIAM, (2003),