

Spectral behaviour of GMRES applied to singular systems

Laurent Smoch*

June 11, 2004

Abstract

The purpose of this paper is to develop a spectral analysis of the Hessenberg matrix obtained by the GMRES algorithm used for solving a linear system with a singular matrix. We prove that the singularity of the Hessenberg matrix depends on the nature of A and some others criteria like the zero eigenvalue multiplicity and the projection of the initial residual on particular subspaces. We also introduce some new results about the distinct kinds of breakdown which may occur in the algorithm when the system is singular.

Key words: GMRES, Krylov subspace, singular system, Hessenberg matrix, geometric multiplicities, breakdowns

1 Introduction

The Generalized Minimal Residual algorithm (GMRES) developed by Saad and Schultz in [9] is an effective method for solving large systems of linear equations

$$Ax = b \tag{1}$$

and its behavior is really well understood when the systems are nonsingular. We study in this paper the theoretical behavior of GMRES applied to singular systems in exact arithmetic. We use for this some special tools like the minimal polynomial associated to A and a particular vector, the eigen subspaces and the geometric multiplicities. Several authors have already tried to analyse the effects of a singular matrix on a Krylov subspace method and we give below a list of their works. We remind that Brown and Walker have introduced in [3] conditions concerning the singular matrix A under which the GMRES iterates converge safely to a solution of (1) and remark that these theoretical results are still true for any mathematically equivalent method. They also discuss about the distinct kinds of breakdown appearing in the GMRES algorithm by considering the Krylov subspaces used for seeking the final solution. Ipsen and Meyer [5] also deal within the general framework of the Krylov subspace methods when A is a singular matrix and explain why the singular systems are different. Several results about homogeneous systems

*Laboratoire de Mathématiques Pures et Appliquées, Université du Littoral, zone universitaire de la Mi-voix, bâtiment H. Poincaré, 50 rue F. Buisson, BP 699, F-62228 Calais Cedex, France. Email: smoch@lmpa.univ-littoral.fr

have also been developed by Freund and Hochbruck in [4] implying the Drazin inverse solution of a system (Wang, Wei and Qiao have widely developed this special inverse in [12]). Freund and Hochbruck discuss about linear systems arising in Markov chain modeling and properties under which the QMR and TFQMR converge. More recently, Wei and Wu have extended in [13] these results obtained with an index of A equal to 1 by considering more general parameters. Sidi and Kluzner have developed in [10] a Bi-CG type Krylov subspace method suitable for the general case in which A is not necessarily real symmetric, its index is arbitrary and its spectrum is not necessarily real. At last, Reichel and Ye introduced in [7] an extension of GMRES called BFGMRES (for Breakdown Free GMRES) that overcomes breakdown for consistent and inconsistent linear systems of equations with a singular matrix. These works prove that some alternatives exist for solving particular singular systems but we will not pursue this subject here any further. All these authors have proved the important role played by the initial guess used in the algorithm but nothing in the papers treats with accuracy the spectral nature of the Hessenberg matrix produced by the GMRES. In exact arithmetic, if m is the termination step of the algorithm, the nonsingularity of H_m is a necessary and sufficient condition to obtain a solution, our purposes in this paper are first to establish some new results concerning the spectral properties of the Hessenberg matrix and next, to state some links between the results of the previous papers and those appearing in this work.

The remainder of this paper is organized as follows. In section 2, we briefly describe the Full Orthogonalization Method (FOM) developed in [2], the GMRES algorithm and the tools which will be used in section 3 to present some new results about the singularity of H_m . Some further details are brought in section 4 about the distinct kinds of breakdown appearing in GMRES, which are precisely assigned to the nature of the matrix H_m . In the fifth section, we will describe some numerical examples which confirm our theory, and finally, in the last section, we will give some concluding remarks.

2 Preliminaries

We start in this section with a brief description of the Arnoldi and GMRES algorithms [8, 9] and we present some basic results of linear algebra which will be very useful in the following. No restriction is put on the matrix A which is not necessarily hermitian or Hermitian positive semi. Nevertheless, we put a restriction on the linear system (1) which has to be consistent to admit a solution.

2.1 Presentation of Arnoldi and GMRES algorithms

We are interested in solving the linear system (1) where $A \in \mathbb{R}^{n \times n}$ is a singular matrix and $x, b \in \mathbb{R}^n$. If x_0 is an initial guess of (1), letting $x = x_0 + z$, we obtain the equivalent system

$$Az = r_0$$

where $r_0 = b - Ax_0$ is the initial residual and $z \in \mathbb{R}^n$. Arnoldi's method and GMRES both find an approximate solution

$$x_l = x_0 + z_l \text{ with } z_l \in \mathcal{K}_l,$$

where $\mathcal{K}_l \equiv \mathcal{K}_l(A, r_0)$ is the Krylov subspace determined by A and r_0 , defined by

$$\mathcal{K}_l \equiv \text{span}\{r_0, Ar_0, \dots, A^{l-1}r_0\}.$$

The approximate solution is such that

1. $(b - Ax_l) \perp \mathcal{K}_l$ for Arnoldi's method and
2. $\|b - Ax_l\|_2 = \min_{x \in x_0 + \mathcal{K}_l} \|b - Ax\|_2$ for GMRES.

Here $\|\cdot\|_2$ denotes the Euclidean norm on \mathbb{R}^n and orthogonality is meant in Euclidean sense, we also note that (\cdot, \cdot) denotes the euclidean inner product.

The following algorithm is a version of the Full Orthogonalization Method (GMRES) algorithm. First, Arnoldi's method is used for the construction of an orthonormal basis $\{v_1, v_2, \dots, v_l\}$ for \mathcal{K}_l where the first vector v_1 is obtained by normalizing r_0 and next, the algorithm builds an approximate solution x_l which satisfies 1. (or 2. for GMRES).

1. Arnoldi process.
 - (a) *Start.* Choose an initial guess x_0 and a tolerance ϵ , form $r_0 = b - Ax_0 \in \mathbb{R}^n$ and $v_1 = r_0 / \|r_0\|_2$.
 - (b) *Iterate.* For $j = 1, 2, \dots$, do
 - $v_{j+1} \equiv Av_j$,
 - for $i = 1, 2, \dots, j$ do
 - $h_{i,j} \equiv v_{j+1}^t v_i$, $v_{j+1} = v_{j+1} - h_{i,j} v_i$,
 - $h_{j+1,j} \equiv \|v_{j+1}\|_2$, $v_{j+1} = v_{j+1} / h_{j+1,j}$.
 - Compute the residual norm $\rho_j = \|b - Ax_j\|_2$ of the solution x_j that would be obtained if we stopped at this step.
 - If $\rho_j \leq \epsilon$ set $l = j$ and go to 2.
2. Form the approximate solution.

Arnoldi. Define H_l the $l \times l$ (Hessenberg) matrix whose nonzero entries are the coefficients $h_{i,j}$, $1 \leq i \leq \min\{j+1, l\}$, $1 \leq j \leq l$ and define $V_l = [v_1, v_2, \dots, v_l]$, compute $x_l^A = x_0 + z_l^A$ where $z_l^A = \beta V_l H_l^{-1} e_1$ with $\beta = \|r_0\|_2$ and $e_1 \in \mathbb{R}^l$ is the first vector of the canonical basis of \mathbb{R}^l , if H_l is nonsingular.

GMRES. Define \overline{H}_l the $(l+1) \times l$ (Hessenberg) matrix whose nonzero entries are the coefficients $h_{i,j}$, $1 \leq i \leq j+1$, $1 \leq j \leq l$ and define $V_l = [v_1, v_2, \dots, v_l]$,

- (a) find the vector y_l^{GM} which verifies $\forall y \in \mathbb{R}^l$, $J(y_l^{GM}) \leq J(y)$ with $J(y) = \|\beta e_1 - \overline{H}_l y\|_2$ where $e_1 = (1, 0, \dots, 0)^t \in \mathbb{R}^{l+1}$.
- (b) Compute $x_l^{GM} = x_0 + V_l y_l^{GM}$.

From the Arnoldi process, it follows that

$$AV_l = V_l H_l + h_{l+1,l} v_{l+1} e_l^t \quad (2)$$

where $e_l = (0, \dots, 0, 1)^t \in \mathbb{R}^l$.

This relation can be rewritten as

$$AV_l = V_{l+1}\overline{H}_l. \quad (3)$$

In contrast with the Full Orthogonalization Method algorithm, the approximation x_l^{GM} obtained with GMRES is defined for all l when A is nonsingular. We wonder if the situation is still true when the matrix A is singular.

2.2 Basic theory

We only consider in this section the case where A is singular. Let $Sp(A) = \{\lambda_1, \lambda_2, \dots, \lambda_p\}$ be the set of distinct eigenvalues of the matrix A , with $p \geq 2$ to ensure that at least one eigenvalue is nonzero. We note that all these eigenvalues are not counted with their algebraic multiplicity $\alpha_1, \alpha_2, \dots, \alpha_p$ respectively. In the sequel, we assume that $\lambda_1 = 0$ and $\alpha_1 \geq 1$. We recall now some classical definitions in matrix algebra.

The polynomial

$$P_A(\lambda) \equiv \lambda^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2} \dots (\lambda - \lambda_p)^{\alpha_p} \quad (4)$$

is called the characteristic polynomial of A and we get by means of Cayley-Hamilton theorem that $P_A(A) = 0$. The annihilating polynomial of A of minimal degree such that its highest order coefficient is one and is called the minimal polynomial of A . It can be written as described below

$$Q_A(\lambda) \equiv \lambda^{\nu_1}(\lambda - \lambda_2)^{\nu_2} \dots (\lambda - \lambda_p)^{\nu_p} \quad (5)$$

where the ν_i 's are respectively called geometric multiplicities of the λ_i 's, $\forall i \in \{1, 2, \dots, p\}$. It also satisfies $Q_A(A) = 0$. Using the previous notations, we point out that the following decomposition holds

$$\mathbb{R}^n = \bigoplus_{i=1}^p N_{\lambda_i}, \quad (6)$$

where $N_{\lambda_i} \equiv Ker[(A - \lambda_i I)^{\nu_i}] \forall i \in \{1, 2, \dots, p\}$.

Finally, the normalized polynomial $\Pi_A(\lambda)$ of minimal degree m such that $\Pi_A(A)r_0 = 0$ is called the minimal polynomial associated to A and r_0 . It can be written as follows

$$\Pi_A(\lambda) = \lambda^{\mu_1}(\lambda - \lambda_2)^{\mu_2} \dots (\lambda - \lambda_p)^{\mu_p} \quad (7)$$

with $0 \leq \mu_i \leq \nu_i \leq \alpha_i, \forall i \in \{1, 2, \dots, p\}$.

We also note that if the matrix A is diagonalizable, $\nu_i = 1, \forall i \in \{1, 2, \dots, p\}$ and $N_{\lambda_i} = E_{\lambda_i} \equiv Ker(A - \lambda_i I), \forall i \in \{1, 2, \dots, p\}$.

In exact arithmetic, if m is the termination step of GMRES, the degree of Π_A equals to m and $\Pi_A(A)r_0 = 0$ simultaneously. In this case, $h_{m+1,m} = 0$ and we get by using (2) the relation

$$AV_m = V_m H_m \quad (8)$$

We give the following

Lemma 2.1

$$V_m^t A^k V_m = H_m^k, \quad \forall k \in \mathbb{N} \quad (9)$$

Proof. This lemma is proved by induction. The result (9) is obvious for $k = 1$ by using the relation (8). We assume next that the relation is true at the rank k . Then, $V_m^t A^{k+1} V_m = V_m^t A^k A V_m = V_m^t A^k V_m H_m = H_m^k H_m = H_m^{k+1}$. ■

Then, we give the

Proposition 2.2

$$\forall A \in \mathbb{R}^{n \times n}, \quad \Pi_A(A)r_0 = 0 \Rightarrow \Pi_A(H_m) = 0 \quad (10)$$

Proof. As seen before, when $h_{m+1,m} = 0$, the degree of Π_A equals to m . The polynomial may be written as

$$\Pi_A(\lambda) = a_0 + a_1 \lambda + \dots + a_m \lambda^m = \sum_{i=0}^m a_i \lambda^i.$$

We also know that $\Pi_A(A)r_0 = 0$ and $r_0 = \|r_0\|V_m e_1$. Then,

$$\sum_{i=0}^m a_i A^i V_m e_1 = 0 \Rightarrow \sum_{i=0}^m a_i V_m^t A^i V_m e_1 = 0.$$

By using (9) from the previous lemma, we obtain

$$\sum_{i=0}^m a_i H_m^i e_1 = 0 \Leftrightarrow \Pi_A(H_m)e_1 = 0.$$

We multiply this equality on the left by H_m^j for $j = 0, 1, \dots, m-1$ and

$$H_m^j \Pi_A(H_m)e_1 = 0 \Leftrightarrow \Pi_A(H_m)H_m^j e_1 = 0 \text{ for } j = 0, 1, \dots, m-1.$$

Since $h_{j+1,j} \neq 0$ for $j = 1, 2, \dots, m-1$, the space $\text{Span}\{e_1, H_m e_1, \dots, H_m^{m-1} e_1\}$ equals to \mathbb{R}^m . This yields $\Pi_A(H_m) = 0$. ■

A similar result was stated by Van der Vorst and Vuik in [11] for the FOM method.

We also remind the definition of the index of A . First, for any matrix $M \in \mathbb{R}^{l \times k}$, we denote by $\mathcal{R}(M) \equiv \{My | y \in \mathbb{R}^k\}$ the range of M . Therefore, the lowest nonnegative integer ν such that $\mathcal{R}(A^\nu) = \mathcal{R}(A^{\nu+1})$ is called the index of A and is denoted by $\text{ind}(A)$. Namely, the index is the size of the largest Jordan block corresponding to the zero eigenvalue of A . We get

$$\mathbb{R}^n = \mathcal{R}(A^\nu) \oplus \text{Ker}(A^\nu). \quad (11)$$

It is straightforward by using the properties of the direct sum and (6) that $\nu = \nu_1$, i.e. the geometric multiplicity of the zero eigenvalue is also called the index of A .

It is interesting now to state the connections between the previous multiplicities and the projections of r_0 on the subspaces defined above. Since $r_0 \in \mathbb{R}^n$, by using (6), the initial residual can be written

$$r_0 = u_1 + u_2 + \dots + u_p \quad (12)$$

with $u_i \in N_{\lambda_i}, \forall i \in \{1, 2, \dots, p\}$. Obviously some changes may occur when we consider particular cases of A . The following proposition holds

Proposition 2.3

Let $A \in \mathbb{R}^{n \times n}$ with $\text{ind}(A) = 1$ and $b \in \mathcal{R}(A)$ then, the initial residual r_0 may be written as

$$r_0 = u_2 + u_3 + \dots + u_p.$$

Proof. Suppose that $r_0 = u_1 + u_2 + \dots + u_p$. Since $\nu_1 = 1, u_1 \in E_{\lambda_1} = N_{\lambda_1} = \text{Ker}(A)$. We know that $r_0 = b - Ax_0$ then, if $b \in \mathcal{R}(A), r_0 \in \mathcal{R}(A)$. We have

$$(A - \lambda_i I)^{\nu_i} = \sum_{k=0}^{\nu_i} C_k^{\nu_i} A^k (-\lambda_i I)^{\nu_i - k}, \quad \forall i \in \{2, 3, \dots, p\}$$

but $(A - \lambda_i I)^{\nu_i} u_i = 0, \forall i \in \{2, 3, \dots, p\}$ then

$$\sum_{k=0}^{\nu_i} C_k^{\nu_i} A^k (-\lambda_i I)^{\nu_i - k} u_i = 0, \quad \forall i \in \{2, 3, \dots, p\}.$$

Thus

$$-(-\lambda_i I)^{\nu_i} u_i = A \sum_{k=1}^{\nu_i} C_k^{\nu_i} A^{k-1} (-\lambda_i I)^{\nu_i - k} u_i, \quad \forall i \in \{2, 3, \dots, p\}$$

and then, $u_i \in \mathcal{R}(A), \forall i \in \{2, 3, \dots, p\}$. By letting

$$r_0 = u_1 + w \text{ where } w = u_2 + u_3 + \dots + u_p,$$

we have $u_1 = r_0 - w \in \mathcal{R}(A)$, and then, $u_1 \in \mathcal{R}(A) \cap \text{Ker}(A) = \{0\}$. Therefore, $r_0 = u_2 + u_3 + \dots + u_p$. ■

The previous result means that if the geometric multiplicity associated to the zero eigenvalue equals 1, the decomposition of the initial residual on $\text{Ker}(A^{\nu_1})$ is zero. Therefore, by using (11), we get $r_0 \in \mathcal{R}(A^{\nu_1})$.

Remark

The previous proposition is not true in the general case. We take as an example a singular matrix A such that $\nu_1 > 1$ by letting

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

A is defective and $Sp(A) = \{0, 1\}$. Let $\lambda_1 = 0$ and $\lambda_2 = 1$ with $\alpha_1 = 3$ and $\alpha_2 = 1$. We also take $b = Ay = (4, 2, 1, 0)^t$ where $y = (1, 1, 1, 1)^t$ and $x_0 = (0, 0, 0, 0)^t$. We assume that m is the termination step in the algorithm in exact arithmetic. Then, in this case, $h_{m+1, m} = 0$ if and only if $m = 3$. We get $r_0 = (4, 2, 1, 0)^t$ and the minimal polynomial of A for r_0 is

$$\Pi_A(A)r_0 = A^2(A - I)r_0 = 0.$$

If $r_0 = u_2$ then, the minimal polynomial of A and r_0 would have the form $\Pi_A(A)r_0 = (A - I)r_0$, but as $(A - I)r_0 \neq 0$, $r_0 = u_1 + u_2$.

The condition $\nu_1 = 1$ is satisfied by some particular matrices as shown below.

Corollary 2.4

Consider the matrix $A \in \mathbb{R}^{n \times n}$ and the right-hand side $b \in \mathcal{R}(A)$. If A is diagonalizable and the algebraic multiplicity of $\lambda_1 = 0$ is such that $\alpha_1 \geq 1$ then, $r_0 = u_2 + u_3 + \dots + u_p$.

Proof. We have seen already that the diagonalizability of A yields the fact that $\nu_i = 1$, $\forall i \in \{1, 2, \dots, p\}$. It is obvious that the assumptions which are used are the same as in the previous proposition, which concludes the proof. ■

Next, we introduce $\mathbb{K}[\lambda]$ the space of the polynomials whose coefficients are in \mathbb{K} where \mathbb{K} is \mathbb{R} or \mathbb{C} . Then, the following result holds

Lemma 2.5

Let P_1, P_2, \dots, P_r be some polynomials $\in \mathbb{K}[\lambda]$ such that P_i and P_j are prime among themselves, $\forall i, j \in \{1, 2, \dots, r\}$ and $P = P_1 \times P_2 \times \dots \times P_r$, then

$$Ker(P(A)) = \bigoplus_{j=1}^r Ker(P_j(A)) \text{ holds.} \quad (13)$$

Proof. This well known result can be proved by induction and Bezout's identity [6]. ■

We know that Π_A can be written as (7) where $\mu_1 + \mu_2 + \dots + \mu_p = m$ therefore, this suggests the following proposition

Proposition 2.6

Assume that $\Pi_A(A)r_0 = 0$. Then,

$$u_i \in Ker[(A - \lambda_i I)^{\mu_i}], \forall i \in \{1, 2, \dots, p\}.$$

Proof. Let $r_0 = u_1 + u_2 + \dots + u_p$ where $u_i \in Ker[(A - \lambda_i I)^{\mu_i}]$, $\forall i \in \{1, 2, \dots, p\}$.

Then, $\Pi_A(A)r_0 = 0$ implies $(A - \lambda_1 I)^{\mu_1} (A - \lambda_2 I)^{\mu_2} \dots (A - \lambda_p I)^{\mu_p} r_0 = 0$. By premultiplying $\Pi_A(A)$ by $(A - \lambda_i I)^{\nu_i - \mu_i}$, $\forall i \in \{2, 3, \dots, p\}$, we have

$$(A - \lambda_1 I)^{\mu_1} (A - \lambda_2 I)^{\nu_2} \dots (A - \lambda_p I)^{\nu_p} r_0 = 0.$$

and this yields $(A - \lambda_1 I)^{\mu_1} (A - \lambda_2 I)^{\nu_2} \dots (A - \lambda_p I)^{\nu_p} u_1 = 0$ since $u_i \in N_{\lambda_i}$, $\forall i \in \{2, 3, \dots, p\}$.

Thus, $u_1 \in Ker[(A - \lambda_1 I)^{\mu_1} (A - \lambda_2 I)^{\nu_2} \dots (A - \lambda_p I)^{\nu_p}]$.

But $(A - \lambda_i I)$ and $(A - \lambda_j I)$ are prime among themselves for $i \neq j$, $\forall i, j \in \{1, 2, \dots, p\}$ then, by using Lemma 2.5, we have

$$u_1 \in Ker[(A - \lambda_1 I)^{\mu_1}] \oplus Ker[(A - \lambda_2 I)^{\nu_2}] \oplus \dots \oplus Ker[(A - \lambda_p I)^{\nu_p}].$$

Since $u_1 \notin \text{Ker}[(A - \lambda_j I)^{\nu_j}]$, $\forall j \in \{2, 3, \dots, p\}$, $u_1 \in \text{Ker}[(A - \lambda_1 I)^{\mu_1}]$.

By using the same technique for the vectors u_2, u_3, \dots, u_p , we obtain

$$u_i \in \text{Ker}[(A - \lambda_i I)^{\mu_i}], \quad \forall i \in \{1, 2, \dots, p\}.$$

■

This proposition simply means that we may specify the subspaces where the u_i 's lie since $0 \leq \mu_i \leq \nu_i$, $\forall i \in \{1, 2, \dots, p\}$, this provides us the decomposition of the initial residual. According to this proposition, we can prove the following result

Proposition 2.7 *We assume that $A \in \mathbb{R}^{n \times n}$ is singular. Then,*

$$r_0 = \sum_{i=1, \mu_i \neq 0} u_i$$

where u_i is the decomposition of r_0 on the subspace $\text{Ker}[(A - \lambda_i I)^{\mu_i}]$ and μ_i is the multiplicity of λ_i in the minimal polynomial Π_A of A and r_0 .

Proof. We assume first that $r_0 = \sum_{\substack{i=1 \\ i \neq i_0}}^p u_i$ for $i_0 \in \{1, 2, \dots, p\}$. Then, by using the previous proposition, we obtain $u_i \in \text{Ker}[(A - \lambda_i I)^{\mu_i}]$, $\forall i \in \{1, \dots, p\}$ such that $i \neq i_0$ and by using Lemma 2.5 once again, we obtain

$$r_0 \in \text{Ker}[(A - \lambda_1 I)^{\mu_1} \dots (A - \lambda_{i_0-1} I)^{\mu_{i_0-1}} (A - \lambda_{i_0+1} I)^{\mu_{i_0+1}} \dots (A - \lambda_p I)^{\mu_p}]$$

or under a compact form $\prod_{\substack{i=1 \\ i \neq i_0}}^p (A - \lambda_i I)^{\mu_i} r_0 = 0$. Due to the uniqueness of the minimal polynomial of A for r_0 , we obtain $\mu_{i_0} = 0$.

We prove now the other implication which seems obvious from Proposition 2.6. Indeed,
 $\mu_i = 0 \Rightarrow u_i \in \text{Ker} I \Leftrightarrow u_i = 0$.

■

Both the previous propositions are easy to state, and they were proved noting the way the u_i 's and the μ_i 's took place in the minimal polynomial of A and r_0 . We would like to state some links between these results and the singularity of the Hessenberg matrix obtained during the Arnoldi's process. For this, we use Proposition 2.2 which states in exact arithmetic that if $\Pi_A(A)r_0 = 0$ then, $\Pi_A(H_m) = 0$. It means that the minimal polynomial of A and r_0 is also the characteristic polynomial of H_m . Then, if $\Pi_A(0) = 0$, it is equivalent to the fact that H_m admits at least one zero eigenvalue.

Finally, the nonsingularity of H_m depends on the multiplicity μ_1 by using (7) or u_1 due to Proposition 2.7. This proves the essential role played by the initial guess x_0 in GMRES. In the next section, we introduce some new results about GMRES applied to singular systems in exact arithmetic. It deals first with the success of the determination of a solution of (1) and next, the termination step of the algorithm when A is diagonalizable.

3 Theoretical results

In this section, we wish first to state some links between the previous results and those obtained by Ipsen and Meyer in [5]. We give the following

Proposition 3.1

In exact arithmetic, if $\nu_1 \neq 0$,

$$r_0 \in R(A^{\nu_1}) \text{ if and only if } H_m \text{ is nonsingular.}$$

Proof. Assume first that H_m is nonsingular, we have seen previously that it is equivalent to $\mu_1 = 0$. By using Proposition 2.7, we get $r_0 = u_2 + u_3 + \dots + u_p$. Therefore, $\Pi_A(A)r_0 = 0$ yields

$$(A - \lambda_2 I)^{\nu_2} (A - \lambda_3 I)^{\nu_3} \dots (A - \lambda_p I)^{\nu_p} r_0 = 0.$$

By multiplying the the left-hand side of this equality by $(A - \lambda_i I)^{\nu_i - \mu_i}$, $\forall i \in \{2, 3, \dots, p\}$ and by using Lemma 2.5, we have

$$r_0 \in Ker[(A - \lambda_2 I)^{\nu_2} (A - \lambda_3 I)^{\nu_3} \dots (A - \lambda_p I)^{\nu_p}] = \bigoplus_{i=2}^p Ker[(A - \lambda_i I)^{\nu_i}].$$

From the equalities (6) and (11), it is clear that $r_0 \in R(A^{\nu_1})$ since the subspaces that we use are complementary.

Assume now that $r_0 \in R(A^{\nu_1})$. Then,

$$r_0 \in \bigoplus_{i=2}^p Ker[(A - \lambda_i I)^{\nu_i}].$$

Hence, $(A - \lambda_2 I)^{\nu_2} (A - \lambda_3 I)^{\nu_3} \dots (A - \lambda_p I)^{\nu_p} r_0 = 0$. But we can use the fact that $(A - \lambda_2 I)^{\nu_2} (A - \lambda_3 I)^{\nu_3} \dots (A - \lambda_p I)^{\nu_p} \neq 0$ since $\nu_1 \neq 0$. Therefore, $\mu_1 = 0$ by using the fact that $\Pi_A(\lambda)$ is a factor of $(x - \lambda_2)^{\nu_2} (x - \lambda_3)^{\nu_3} \dots (x - \lambda_p)^{\nu_p}$ due to its minimality. We get finally that H_m is nonsingular. ■

This result according to Proposition 3.1 will be very useful in justifying the existence of a solution of (1).

Remark 3.1

The previous proposition states clearly that the choice of x_0 may bring some alternatives about the nonsingularity of the Hessenberg matrix H_m . The initial guess x_0 must be chosen such that

$$b - Ax_0 \in R(A^{\nu_1}).$$

Then, if $b = A(b_1 + A(b_2 + \dots))$ and $x_0 = x_{0,1} + A(x_{0,2} + A(x_{0,3} + \dots))$, $r_0 \in Im(A^{\nu_1})$ if and only if

$$Ab_1 + A^2 b_2 + \dots + A^{\nu_1 - 1} b_{\nu_1 - 1} = Ax_{0,1} + A^2 x_{0,2} + \dots + A^{\nu_1 - 1} x_{0, \nu_1 - 1}.$$

It seems impossible to obtain such a decomposition of x_0 . However, it gives a way to prevent the singularity of H_m .

Let us discuss now about the singularity of H_m . The following proposition holds

Proposition 3.2

Let $A \in \mathbb{R}^{n \times n}$ be a singular matrix such that $Ax = b$ with $b \in \mathcal{R}(A)$. Let A be defective, and $x_0 \in \mathbb{R}^n$ an initial guess such that $r_0 = b - Ax_0 = u_1 + u_2 + \dots + u_p$ with $u_i \in \text{Ker}[(A - \lambda_i I)^{\mu_i}]$, $\forall i \in \{1, 2, \dots, p\}$. Then, in exact arithmetic

- if $\nu_1 = 1$, the Hessenberg matrix H_m produced by GMRES at the termination step m is nonsingular and a solution of (1) is found by GMRES,
- if $\nu_1 > 1$, then
 - if u_1 is zero, the Hessenberg matrix H_m produced by GMRES is nonsingular and a solution of (1) is found by GMRES,
 - if u_1 is nonzero, the Hessenberg matrix H_m produced by the GMRES is singular and GMRES does not find any solution of (1).

Proof. These assertions are very easy to prove. If $\nu_1 = 1$, we have seen that $u_1 = 0$ by using Proposition 2.3. It is equivalent to $\mu_1 = 0$ by Proposition 2.7 and it yields $\Pi_A(0) \neq 0$, i.e. H_m is nonsingular. Since $h_{m+1,m} = 0$, we have $\|\beta e_1^{(m+1)} - \overline{H}_m y\|_2 = \|\beta e_1^{(m)} - H_m y\|_2$ in exact arithmetic at the termination step m , when the solution of (1) is found and is given by $x_m = x_0 + V_m H_m^{-1} \beta e_1^{(m)}$.

If $\nu_1 > 1$, Proposition 2.3 does not apply. Two cases may appear :

- r_0 is such that its component on $\text{Ker}(A^{\mu_1})$ is zero. Then, by using Proposition (2.7), $\mu_1 = 0$ holds, which is equivalent to the fact that H_m is nonsingular and the same conclusions as the previous case hold.
- r_0 is such that its component on $\text{Ker}(A^{\mu_1})$ is not zero. Then, $\mu_1 \neq 0$. Indeed if $\mu_1 = 0$, by using Proposition (2.7), $r_0 = u_2 + \dots + u_p$ which contradicts our hypothesis. This yields H_m is singular because in this case, $\Pi_A(0) = 0$ and GMRES does not find a solution. ■

By using the previous result and Proposition 3.1, we get the same conclusions as obtained in theorem 3 by Ipsen and Meyer [5], i.e. a square linear system $Ax = b$ has a Krylov solution (a solution which lies in the Krylov subspace $K_m(A, r_0)$) if and only if $r_0 \in \mathcal{R}(A^{\nu_1})$.

We consider now the particular case of diagonalizable matrices. We know from Corollary 2.4 that u_1 is zero. Therefore, $\Pi_A(A)r_0 = 0$ yields

$$(A - \lambda_2 I)^{\mu_2} (A - \lambda_3 I)^{\mu_3} \dots (A - \lambda_p I)^{\mu_p} (u_2 + u_3 + \dots + u_p) = 0, \quad (14)$$

with $\mu_2 + \mu_3 + \dots + \mu_p = m$ and $0 \leq \mu_i \leq 1$, $\forall i \in \{2, 3, \dots, p\}$. We remind that Saad and Schultz [9] have shown that the number of iterations required by GMRES to find the exact solution is

equal to the degree of Π_A . Let see what happens when A is singular and diagonalizable.

We consider the following cases :

- We assume that $r_0 = \sum_{i=2}^p u_i$. Then, the equality (14) holds if and only if $\mu_i = 1, \forall i \in \{2, 3, \dots, p\}$, which yields $m = \mu_2 + \mu_3 + \dots + \mu_p = p - 1$.
- Next, we assume that $l - 1$ vectors in the decomposition of r_0 are zero with $2 \leq l \leq p - 1$ and $u_{i_j} \in \{u_2, u_3, \dots, u_p\}, \forall j \in \{1, 2, \dots, p - 1\}$. Then, it is evident that

$$r_0 = u_{i_l} + u_{i_{l+1}} + \dots + u_{i_{p-1}}.$$

By using Proposition (2.7) and the uniqueness of the minimal polynomial of A for r_0 , we get $(A - \lambda_{i_l} I)^{\mu_{i_l}} (A - \lambda_{i_{l+1}} I)^{\mu_{i_{l+1}}} \dots (A - \lambda_{i_{p-1}} I)^{\mu_{i_{p-1}}} r_0 = 0$ which yields

$$\mu_{i_l} + \mu_{i_{l+1}} + \dots + \mu_{i_{p-1}} = m.$$

But $\mu_{i_j} = 1, \forall j \in \{l, l + 1, \dots, p - 1\}$. Thus, $\mu_{i_l} + \mu_{i_{l+1}} + \dots + \mu_{i_{p-1}} = p - l = m$.

Therefore, when A is diagonalizable, $0 < p - l \leq m \leq p - 1$. Finally we get the

Proposition 3.3

Let A be a $n \times n$ singular matrix such that $Ax = b$. We assume that A is diagonalizable, $b \in \mathcal{R}(A)$, $x_0 \in \mathbb{R}^n$ an initial guess. Then, the GMRES applied to (1) produces the solution in $m = p - 1$ steps if $r_0 = u_2 + \dots + u_p$ and $0 < m = p - l$ steps in the case where l vectors are zero in the decomposition of r_0 .

The same result is obtained in [5] and in [1] by Axelsson (under a more generally formulation). We also note that no particular relation on p is found when A is nondiagonalizable. Indeed, nothing is known about the μ_i 's in this case.

4 The distinct kinds of breakdown

In the following, we introduce some results from Brown and Walker [3].

We consider $A(\mathcal{K}_l)$ the l -th Krylov subspace defined by

$$A(\mathcal{K}_l) \equiv \text{span}\{Ar_0, A^2r_0, \dots, A^l r_0\}. \quad (15)$$

We note that $\dim A(\mathcal{K}_l) \leq \dim(\mathcal{K}_l) \leq l$ for each $l \in \mathbb{N}$. We shall say that GMRES does not break down at the l -th step if $\dim A(\mathcal{K}_l) = l$. In this case, $\dim A(\mathcal{K}_l) = \dim(\mathcal{K}_l)$ and therefore, the problem

$$\min_{x \in x_0 + \mathcal{K}_l} \|b - A(x_0 + z)\|_2 = \min_{x \in x_0 + \mathcal{K}_l} \|r_0 - Az\|_2 \quad (16)$$

has a unique solution.

As it has been proved in [3], two distinct kinds of breakdown can occur in the algorithm,

1. when there is a rank deficiency of the least-squares problem (16), which occurs when we have the relation $\dim A(\mathcal{K}_l) < \dim(\mathcal{K}_l)$,
2. when there is a degeneracy of the Krylov subspace \mathcal{K}_l , which occurs when $\dim(\mathcal{K}_l) < l$.

Thus, we remind the following theorem from [3]

Theorem 4.1

We assume that we apply GMRES to the system (1). Then, at some step, either

1. *GMRES breaks down through rank deficiency of the least-squares problem (16) without determining a solution*
or
2. *GMRES determines a solution without breakdown and then, breaks down at the next step through degeneracy of the Krylov subspace.*

We also note that if A is nonsingular, GMRES does not break down until the solution of (1) has been found.

Unfortunately these breakdowns are not precisely related to the singularity of the Hessenberg matrix H_m . Our aim in this section is to state some connections between the results of the previous chapter and those obtained by Brown and Walker, in order to understand why a Krylov solution may be found or not and what kind of breakdown appears in this case.

We assume now that until the l -th step

$$\dim A(\mathcal{K}_l) = \dim(\mathcal{K}_l) = l, \quad l > 1, \tag{17}$$

i.e. no breakdown has occurred.

The equality (17) yields the fact that the minimal polynomial of A for r_0 has a greater degree than $l - 1$.

Let us assume now that at the $(l + 1) - th$ step,

$$\dim(\mathcal{K}_{l+1}) < l + 1. \tag{18}$$

Therefore, $\dim(\mathcal{K}_{l+1}) = l$ and $A^l r_0 \in \mathcal{K}_l$, i.e. $\exists \alpha_1, \alpha_2, \dots, \alpha_l \in \mathbb{R}$ such that

$$A^l r_0 = \alpha_1 r_0 + \alpha_2 A r_0 + \dots + \alpha_l A^{l-1} r_0.$$

Hence, $R_A(A)r_0 = 0$ by letting $R_A(\lambda) \equiv \alpha_1 + \alpha_2 \lambda + \dots + \alpha_l \lambda^{l-1} - \lambda^l$. In this case, $R_A = \Pi_A$ is the minimal polynomial of A for r_0 by uniqueness and it admits a degree $m = l$. By supposing that $\Pi_A(0) = 0$, we get $\alpha_1 = 0$ and $\dim A(\mathcal{K}_l) < l$, which contradicts (17). Thus, H_m is nonsingular.

Let us assume now that H_m is nonsingular. Then, the minimal polynomial of A for r_0 is such that

$$\Pi_A(A)r_0 = (a_0 I + a_1 A + \dots + a_m A^m)r_0 = 0$$

with $a_m \neq 0$ since the degree of Π_A equals to m and $a_0 \neq 0$ by assumption. Thus, $A^m r_0 = (\alpha_0 I + \alpha_1 A + \dots + \alpha_{m-1} A^{m-1}) r_0$ with $\alpha_i = -a_i/a_m$ for $i \in \{0, 1, \dots, m-1\}$. Thus, if $\dim(\mathcal{K}_m) = m$, $\dim(\mathcal{K}_{m+1}) < m+1$. Since $\alpha_0 = -a_0/a_m \neq 0$, $\dim A(\mathcal{K}_m) = \dim(\mathcal{K}_m)$. Hence, we obtain the

Proposition 4.2

We consider the system (1) with $A \in \mathbb{R}^{n \times n}$ a singular matrix, $x, b \in \mathbb{R}^n$ and $b \in \mathcal{R}(A)$. We apply GMRES to (1) which stops at the m -th step. Then, by supposing that $\dim A(\mathcal{K}_l) = \dim(\mathcal{K}_l) = l$, $\forall l < m$ and $\dim(\mathcal{K}_m) = m$,

$$H_m \text{ is nonsingular if and only if } \begin{cases} \dim A(\mathcal{K}_m) = \dim(\mathcal{K}_m) \\ \dim(\mathcal{K}_{m+1}) < m+1 \end{cases} .$$

By using the same techniques as previously, it is quite easy to prove

Proposition 4.3

We consider the system (1) with $A \in \mathbb{R}^{n \times n}$ a singular matrix, $x, b \in \mathbb{R}^n$ and $b \in \mathcal{R}(A)$. We apply GMRES to (1) which stops at the m -th step. Then, by supposing that $\dim A(\mathcal{K}_l) = \dim(\mathcal{K}_l) = l$, $\forall l < m$ and $\dim(\mathcal{K}_m) = m$,

$$H_m \text{ is singular if and only if } \dim A(\mathcal{K}_m) < \dim(\mathcal{K}_m).$$

Therefore, in exact arithmetic, the singularity or the nonsingularity of H_m is equivalent to a particular breakdown. By using the results of chapter 3, we can state some links between the multiplicity μ_1 or the vector u_1 as described above.

Theorem 4.4

We consider the system (1) with $A \in \mathbb{R}^{n \times n}$ singular, $x, b \in \mathbb{R}^n$ and $b \in \mathcal{R}(A)$. We apply GMRES to (1) which stops at the m -th step. Let r_0 be the initial residual such that $r_0 = u_1 + u_2 + \dots + u_p$ with $u_i \in \text{Ker}[(A - \lambda_i I)^{\mu_i}]$, $\forall i \in \{1, 2, \dots, p\}$. Then

1. if u_1 is zero, GMRES determines a solution of (1) without breakdown and breaks down at the next step through degeneracy of the Krylov subspace,
2. if u_1 is nonzero, GMRES breaks down through rank deficiency of the least-squares problem (16) without determining any solution.

Proof. We have seen that $r_0 = u_2 + u_3 + \dots + u_p$ implies $\mu_1 = 0$ by Proposition 2.7. Therefore, H_m is nonsingular and it is equivalent by using Proposition (4.2) to $\begin{cases} \dim A(\mathcal{K}_m) = \dim(\mathcal{K}_m) \\ \dim(\mathcal{K}_{m+1}) < m+1 \end{cases}$. If u_1 is nonzero, it yields $\mu_1 \neq 0$ from Proposition 2.7. Thus, H_m is singular and $\dim A(\mathcal{K}_m) < \dim(\mathcal{K}_m)$ by using Proposition (4.3). ■

These results allow us to understand why the nonsingularity of H_m plays an essential role in GMRES in exact arithmetic. All the criteria like the determination of a Krylov solution or the kind of breakdown which occurs in the algorithm depend on H_m and the analysis of this matrix finds its convenience here. We introduce in the following section some examples, of small size to ensure the exact arithmetic, that illustrate all the previous discussions.

5 Numerical examples

In our numerical experiments, we take as right hand member $b = Ay \in \mathbb{R}^n$, where $y = (1, 1, \dots, 1)^t$ and the initial guess $x_0 = (0, 0, \dots, 0)^t$ is such that $r_0 = b - Ax_0 = b \in \mathcal{R}(A)$.

- *Experiment 4.1.* We choose first some diagonalizable matrices and $\alpha_1 \geq 1$. We consider $p \geq 2$ by taking A such that

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ -1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 1 & 0 \\ \vdots & \ddots & -1 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix}.$$

We assume that n is odd, $n = 21$, in which case A is singular whose spectrum contains one zero eigenvalue ($\alpha_1 = 1$).

By applying GMRES, we know help to the previous section that H_m is nonsingular. We get $h_{m+1,m} = 0$ if $m = 20$, this yields $r_0 = u_2 + \dots + u_p$ where $p = 21 = m + 1$. We also note that $Sp(H_m) = \{\lambda_2, \lambda_3, \dots, \lambda_p\} = Sp(A) \setminus \{0\}$.

By taking n even such that $n = 22$ we get $h_{m+1,m} = 0$ for $m = 10$. H_m is nonsingular and $Sp(H_m) \subset Sp(A) \setminus \{0\}$. We know that $p = 22$ and $p - l = m$ by using Proposition 3.3. Thus, we get $l = 12$ where l is the number of vectors which are zero in the decomposition of r_0 .

We can also consider the case $\nu_1 = 1$ and $\alpha_1 > 1$ with the matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ -1 & \ddots & \ddots & \ddots & & & \vdots \\ 0 & \ddots & \ddots & 1 & \ddots & & \vdots \\ \vdots & \ddots & -1 & 0 & 0 & \ddots & \vdots \\ \vdots & & \ddots & 0 & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 & 0 \end{pmatrix}.$$

We take $n = 25$ and $\alpha_1 = 5$. Then, $h_{m+1,m} = 0$ if $m = 20$ and H_m is nonsingular. This yields $r_0 = u_2 + \dots + u_p$ where $p = 21$. We also get $Sp(H_m) = Sp(A) \setminus \{0\}$.

If we take $n = 26$, the matrix H_m is nonsingular and $h_{m+1,m} = 0$ if and only if $m = 10$, $Sp(H_m) \subset Sp(A) \setminus \{0\}$ and $l \geq 1$ ($l = 12$) as in both the examples.

It is proved theoretically in the both examples that a Krylov solution is obtained and we checked it numerically using Matlab on SunSparc station.

- *Experiment 4.2.* Now we look at defective matrices, we consider $\nu_1 = 1$ by means of the matrix

$$A = \begin{pmatrix} 0 & 1 & \dots & 1 \\ 0 & 1 & \dots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

We take $n = 20$. Then, H_m is nonsingular with $h_{m+1,m} = 0$ if $m = 19$. This yields u_1 is zero. We also consider the same matrix with random values on the diagonal, i.e. $A(i, i) = \text{rand}(1, 1)$, $\forall i \in \{2, 3, \dots, n\}$. Therefore, H_m is still nonsingular but $h_{m+1,m} = 0$ for a different m . A Krylov solution is determined by GMRES at the 19-th step and GMRES breaks down at the next step through degeneracy of the Krylov subspace.

We consider now the defective case and $\nu_1 > 1$. We work with the following matrix

$$A = \begin{pmatrix} 1 & \dots & \dots & \dots & \dots & \dots & 1 \\ 0 & \ddots & & & & & \vdots \\ \vdots & \ddots & 1 & \dots & \dots & \dots & 1 \\ \vdots & & \ddots & 0 & 1 & \dots & 1 \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 \end{pmatrix}.$$

We take $n = 25$ and $\nu_1 = k = 5$, i.e. the last k diagonal elements are zero. Then, we get $h_{m+1,m} = 0$ for $m = 24$ and H_m is singular. Consequently u_1 is nonzero, and $\Pi_A(A)r_0 = A^4(A - I)^{20}r_0 = 0$. Hence, we remark numerically that no solution is found by GMRES.

By taking the same matrix with the first $k = 5$ diagonal elements equal to zero and the last diagonal elements equal to 1, we obtain that $h_{m+1,m} = 0$ if $m = 20$. Then, H_m is nonsingular and u_1 is zero. The minimal polynomial of A for r_0 can be written under the form $\Pi_A(A)r_0 = (A - I)^{20}r_0 = 0$ and this time, a Krylov solution is determined by GMRES. The vector u_1 is nonzero and according to Theorem 4.4, GMRES breaks down through rank deficiency of the least-squares problem (16).

These two examples show that when A is a defective matrix and $\nu_1 > 1$, the Hessenberg matrix H_m may be either singular or not, we have seen that it depends on the definition of the vector u_1 . Since $r_0 = b - Ax_0$, the nonsingularity of H_m depends on the choice of the initial guess x_0 as it has been shown in the Remark 3.1.

6 Conclusion

In exact arithmetic, the determination of a Krylov solution depends on the nonsingularity of the Hessenberg matrix produced by the GMRES algorithm. An analysis of the nature

of H_m was given in this paper help to the notion of the minimal polynomial of A for the initial residual r_0 . The results which are also obtained connect the discussion to breakdowns occurring in the GMRES when the system is singular and consistent. Nevertheless, these results can not be used in practical applications since the method is generally terminated prematurely and $h_{l+1,l} = 0$ is not checked. It is clear that GMRES does not run to completion and it would be interesting to obtain similar results in finite precision by using this time the matrix \overline{H}_l .

Acknowledgements. The author would like to thank the referees and especially Ken Hayami for their valuable suggestions and for encouraging the revision of the presentation..

References

- [1] O. AXELSSON. *A generalized conjugate gradient, least-square method*, Numer. Math., 51 (1987), pp. 209-227.
- [2] P. N. BROWN. *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Stat. Comput., 12 (1991), pp. 58-78.
- [3] P. N. BROWN, H. F. WALKER. *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. App., 18 (1997), pp. 37-51.
- [4] R. W. FREUND, M. HOCHBRUCK. *On the use of two QMR algorithms for solving singular systems and applications in Markov chain modeling*, Num. Lin. Algebra, 1 (1994), pp. 403-420.
- [5] I. C. F. IPSEN, C. D. MEYER. *The idea behind Krylov methods*, Amer. Math. Monthly, 105 (1998), pp. 889-899.
- [6] S. LANG. *Algèbre linéaire*, 2 (1976), pp. 294-295, InterEditions Paris.
- [7] L. REICHEL, Q. YE. *Breakdown-free GMRES for singular systems*, To appear (2004).
- [8] Y. SAAD. *Krylov subspace methods for solving large unsymmetric linear systems*, Math. Comp., 37 (1981), pp. 105-26.
- [9] Y. SAAD, M. H. SCHULTZ. *GMRES : a generalized residual method for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856-869.
- [10] A. SIDI, V. KLIZNER. *A Bi-CG type iterative method for Drazin-inverse solution of singular inconsistent nonsymmetric linear systems of arbitrary index*, The Electronic Journal of Linear Algebra, 6 (2000), pp. 72-94.
- [11] H. A. VAN DER VORST, C. VUIK. *The superlinear convergence behaviour of GMRES*, J. Comp. App. Math., 48 (1993), pp. 327-341.

- [12] G. WANG, Y. WEI, S. QIAO. *Generalized Inverses : Theory and Computations*, Science Press, Beijing (2004).
- [13] Y. WEI, H. WU. *Convergence properties of Krylov subspace methods for singular linear systems with arbitrary index*, J. Comp. App. Math., 114 (2000), pp. 305-318.