



Necessary and sufficient conditions for GMRES complete and partial stagnation

G rard Meurant

30 rue du sergent Bauchat, 75012 Paris, France

ARTICLE INFO

Article history:

Available online 17 June 2013

Keywords:

GMRES

Complete stagnation

Partial stagnation

ABSTRACT

In this paper we give necessary and sufficient conditions for the complete or partial stagnation of the GMRES iterative method for solving real linear systems. Our results rely on a paper by Arioli, Pt k and Strako  (1998), characterizing the matrices having a prescribed convergence curve for the residual norms. We show that we have complete stagnation if and only if the matrix A is orthonormally similar to an upper or lower Hessenberg matrix having a particular first row or column or a particular last row or column. Partial stagnation is characterized by a particular pattern of the matrix Q in the QR factorization of the upper Hessenberg matrix generated by the Arnoldi process.

  2013 IMACS. Published by Elsevier B.V. All rights reserved.

1. Introduction

The Generalized Minimum RESidual iterative method (GMRES) was introduced by Saad and Schultz [13] to solve linear systems $Ax = b$, where A is a nonsingular real matrix of order n . It is a Krylov method based on the Arnoldi orthogonalization process using an orthogonal basis of a Krylov space. The initial residual is denoted as $r^0 = b - Ax^0$ where x^0 is the starting vector. The Krylov subspace of order k based on A and r^0 denoted as $\mathcal{K}_k(r^0, A)$ is $\text{span}\{r^0, Ar^0, \dots, A^{k-1}r^0\}$. The approximate solution x^k at iteration k is sought as $x^k \in x^0 + \mathcal{K}_k(r^0, A)$, such that the norm of the residual vector $r^k = b - Ax^k$ is minimized.

In this paper we study the problem of stagnation of the GMRES algorithm for real matrices and right-hand sides. Throughout the paper we assume that the matrix A is nonderogatory; hence the degree of the minimal polynomial of A is equal to n . We define partial stagnation for a given matrix A and right-hand side b as having

$$\|r^k\| < \|r^{k-1}\|, \quad k = 1, \dots, m, \quad \|r^k\| = \|r^{k-1}\|, \quad k = m + 1, \dots, m + p - 1,$$

and

$$\|r^k\| < \|r^{k-1}\|, \quad k = m + p, \dots, n.$$

Hence the norms of the residual stay the same for p iterations starting from $k = m$. Complete stagnation corresponds to $m = 0$ and $p = n$. Thus

$$\|r^k\| = \|r^0\|, \quad k = 1, \dots, n - 1, \quad \text{and} \quad \|r^n\| = 0.$$

Without loss of generality we will use $x^0 = 0$; therefore $r^0 = b$ and we will assume that $\|b\| = 1$.

E-mail address: gerard.meurant@gmail.com.

Complete stagnation of GMRES has been studied by Zavorin, O’Leary and Elman in [17] and Zavorin in [16]; see also Simoncini and Szyld [15] and Simoncini [14] who studied conditions for non-stagnation, Liesen and Tichý [10] for the study of worst-case GMRES for normal matrices and Arioli [1]. A study of necessary and sufficient conditions for complete stagnation when $n \leq 4$ is given in [11]. Up to our knowledge a theoretical study of partial stagnation does not exist in the literature. Note that the definition of partial stagnation in [17] is different from ours. Their definition corresponds to $m = 0$; that is, stagnation at the beginning of GMRES iterations. We will study the stagnation phenomenon using results established for the convergence curve of the residual norms. In [9] Greenbaum and Strakoš proved that any convergence curve for the residual norm that can be generated with GMRES can be obtained with a matrix having prescribed eigenvalues. Greenbaum, Pták and Strakoš [8] showed later that any nonincreasing sequence of residual norms can be given by GMRES. To complement these results, Arioli, Pták and Strakoš [2] gave a complete parametrization of all pairs $\{A, b\}$ generating a prescribed residual norm convergence curve. More details about the matrices involved in this parametrization are given in [12]. More recently Duintjer Tebbens and Meurant [5] showed how to construct a class of matrices and right-hand sides having prescribed Ritz values at each iteration in addition to prescribed residual norms. This is extended to GMRES early termination in [6].

Of course, having complete or partial stagnation is a very special case of prescribed residual norm convergence curve. In this paper we will build upon Arioli, Pták and Strakoš parametrization to obtain necessary and sufficient conditions for complete or partial stagnation. In the course of our study we will see that the orthogonal matrix Q in the QR factorization of the upper Hessenberg matrix H generated by the Arnoldi process has a very special structure in case of stagnation.

The contents of the paper are as follows. Section 2 recall the results from Arioli, Pták and Strakoš [2] that we need for our purposes. Section 3 proves general results that are useful for studying stagnation. In Section 4 we prove an essential result for the orthogonal factor of a QR factorization of the Hessenberg matrix generated by the Arnoldi process. This result is useful to study complete and partial stagnation. In Section 5 we characterize the matrices and right-hand sides leading to complete stagnation. In Section 6 we study partial stagnation. Finally we give some conclusions.

Throughout the paper e^j will denote the j th column of the identity matrix of appropriate order.

2. The Arioli, Pták and Strakoš parametrization

We recall the following results that were proved in [2, Theorem 2.1 and Corollary 2.4].

Theorem 1. Assume we are given n positive numbers

$$f(0) \geq f(1) \geq \dots \geq f(n-1) > 0$$

and n complex numbers $\lambda_1, \dots, \lambda_n$ all different from 0. Let A be a nonderogatory matrix of order n and b an n -dimensional vector. The following assertions are equivalent:

1. The spectrum of A is $\{\lambda_1, \dots, \lambda_n\}$ and GMRES applied to A and b yields residuals r^j , $j = 0, \dots, n-1$, such that

$$\|r^j\| = f(j), \quad j = 0, \dots, n-1.$$

2. The matrix A is of the form $A = WYCY^{-1}W^*$ and $b = Wh$, where W is a unitary matrix, Y is given by

$$Y = \begin{pmatrix} h & R \\ & 0 \end{pmatrix},$$

R being any nonsingular upper triangular matrix of order $n-1$, h a vector such that

$$h = (\eta_1, \dots, \eta_n)^T, \quad \eta_j = (f(j-1)^2 - f(j)^2)^{1/2},$$

and C is the companion matrix corresponding to the polynomial q ,

$$q(z) = (z - \lambda_1) \cdots (z - \lambda_n) = z^n + \sum_{j=0}^{n-1} \alpha_j z^j.$$

In reference to the three authors of the paper [2] we will call the parametrization, $A = WYCY^{-1}W^*$, the APS parametrization. Here we are only interested in real matrices A and right-hand sides b , so we can replace the $*$ above, that means conjugate transpose, by the transpose. In this parametrization, the prescribed residual norm convergence curve is implicitly contained in the vector h which is the first column of Y . The squares of its elements are the differences of the squares of the residual norms at successive iterations. Whatever the values of W orthogonal, R upper triangular and C a companion matrix are, we will obtain with all these matrices A the same convergence curve for GMRES as long as the right-hand side is $b = Wh$.

When we have partial stagnation some of the residual norms are the same. Hence we have $\eta_j = 0$, $j = m + 1, \dots, m + p - 1$. The vector h has therefore $p - 1$ consecutive zero components. In case of complete stagnation all the residual norms are the same except for the last one. In this case we have $\eta_j = 0$, $j = 1, \dots, n - 1$. The vector h has only the last component which is nonzero. Moreover, since $h = W^*b$ and W is orthogonal, we have that $\|h\| = \|b\| = 1$ from our assumption on the norm of b and therefore $h = e^n$.

3. Some properties of the GMRES algorithm

In this section we prove some general properties for the matrices that are involved in the GMRES algorithm using the APS parametrization. The following results do not assume stagnation. Let K be the Krylov matrix,

$$K = (b \quad Ab \quad A^2b \quad \dots \quad A^{n-1}b),$$

whose columns are the natural basis vectors of the Krylov space \mathcal{K}_n . In the sequel we use the notation of [Theorem 1](#). The first interesting property is that $K = WY$; this was proven in [\[2\]](#). Let V be the orthogonal matrix whose columns are the mutually orthogonal Arnoldi basis vectors. Then let $K = VU$ be the QR factorization of K where the matrix U is upper triangular. It is proved in [\[12, Proposition 3.1\]](#) that the matrix U has positive diagonal elements. The Arnoldi process gradually computes an upper Hessenberg matrix H . At the end of the iterations we have $AV = VH$.

Theorem 2. *The matrix U is the upper triangular Cholesky factor of the matrix $Y^T Y$. The Hessenberg matrix H of the Arnoldi process is given by*

$$H = UCU^{-1}.$$

A QR factorization of the matrix H is $H = Q\mathcal{R}$ where $Q = V^T W$ is upper Hessenberg orthogonal and \mathcal{R} is upper triangular. Moreover, we have

$$Q = UY^{-1} = U^{-T}Y^T, \quad \mathcal{R} = YCU^{-1}.$$

The matrices Q and \mathcal{R} are also related to the APS parametrization of [Theorem 1](#) by $\mathcal{R}Q = YCY^{-1}$.

Proof. We have

$$K^T K = Y^T W^T W Y = Y^T Y = U^T V^T V U = U^T U.$$

The matrix U being upper triangular with positive diagonal elements, U^T is the Cholesky factor of $Y^T Y$. Since $K = WY = VU$ we have $Y = W^T V U$ and

$$\begin{aligned} H &= V^T A V = V^T W Y C Y^{-1} W^T V \\ &= (W^T V)^T W^T V U C U^{-1} (W^T V)^T W^T V \\ &= U C U^{-1}. \end{aligned}$$

This can also be proven using the well-known relation $AK = KC$. The columns of the matrix W in the APS parametrization give a basis of the Krylov space $AK_n(A, b)$ and we have $AK = W\tilde{R}$ where the matrix \tilde{R} is upper triangular. We have also that $K = VU$. Hence

$$AK = AVU = VHU = W\tilde{R}.$$

But, since $K = WY$, we have $AK = AWY = (WYCY^{-1}W^T)WY = WYC$. Therefore, $\tilde{R} = YC$. The previous equality gives

$$H = V^T W \tilde{R} U^{-1} = Q\mathcal{R},$$

where $Q = V^T W$ is orthogonal and $\mathcal{R} = YCU^{-1}$ is upper triangular. This is a QR factorization of H . It gives the relation between both bases since $W = VQ$. Using the other relation linking V and W (that is, $WY = VU$) we obtain

$$Q = UY^{-1} = U^{-T}Y^T.$$

This relation implies that $Y = Q^T U$, which is a QR factorization of Y . The orthogonal factor Q^T is just the transpose of that of H .

We remark that since $H = Q\mathcal{R} = V^T AV$, we have a factorization of the matrix A as

$$A = VQ\mathcal{R}V^T = VQ(\mathcal{R}Q)Q^T V^T = W\mathcal{H}W^T.$$

The matrix

$$\mathcal{H} = \mathcal{R}Q = YCY^{-1}$$

is also upper Hessenberg. This is an RQ factorization of \mathcal{H} . Since

$$H = UCU^{-1} = QYCY^{-1}Q^T,$$

we obtain the relation between H and \mathcal{H} , $H = Q\mathcal{H}Q^T$. The matrix Q transforms the upper Hessenberg matrix \mathcal{H} to the upper Hessenberg matrix H . \square

4. What is the matrix Q ?

In this section we characterize the orthogonal matrix Q that was introduced in [Theorem 2](#) as a factor of the decomposition $H = Q\mathcal{R}$. We would like to relate the entries of Q to the η_j 's defined in [Theorem 1](#) that describe GMRES convergence. The matrix Q is given by $Q = UY^{-1}$. Therefore we first compute U and the inverse of Y .

Lemma 3. Using the notation of [Theorem 1](#) and assuming $\|b\| = 1$, let \hat{h} be the vector of the first $n - 1$ components of h . For the inverse of the matrix Y , we have

$$Y^{-1} = \begin{pmatrix} 0 & \cdots & 0 & 1/\eta_n \\ & R^{-1} & & -R^{-1}\hat{h}/\eta_n \end{pmatrix}. \quad (1)$$

Let \hat{L} be the lower triangular Cholesky factor of $R^T R - R^T \hat{h} \hat{h}^T R$. Then

$$U = \begin{pmatrix} 1 & \hat{h}^T R \\ 0 & \hat{L}^T \\ \vdots & \\ 0 & \end{pmatrix}. \quad (2)$$

Proof. One can easily check that the matrix in (1) is the inverse of the matrix Y . The matrix U^T is the lower triangular Cholesky factor of the symmetric positive definite matrix $Y^T Y$. From the definition of Y and $\|h\| = \|W^* b\| = \|b\| = 1$, we have

$$Y^T Y = \begin{pmatrix} 1 & \hat{h}^T R \\ R^T \hat{h} & R^T R \end{pmatrix}.$$

It is easy to see that the Cholesky factor L of the matrix $Y^T Y$ is

$$L = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ R^T \hat{h} & \hat{L} & & \end{pmatrix}$$

where $\hat{L}\hat{L}^T = R^T R - R^T \hat{h} \hat{h}^T R$ is the Cholesky factorization of the positive definite matrix in the right-hand side. Since $U = L^T$ this proves the result for the matrix U . \square

Since we know the matrices U and Y^{-1} from [Lemma 3](#), we can immediately write down Q as

$$Q = \begin{pmatrix} \hat{h}^T & \frac{1}{\eta_n} - \frac{\|\hat{h}\|^2}{\eta_n} \\ \hat{L}^T R^{-1} & -\frac{\hat{L}^T R^{-1} \hat{h}}{\eta_n} \end{pmatrix}. \quad (3)$$

However this expression can be considerably simplified as we see in the following theorem.

Theorem 4. Let \check{L} be the lower triangular Cholesky factor of the matrix $I - \hat{h} \hat{h}^T$ and \mathcal{S} be a diagonal matrix whose diagonal entries are ± 1 according to the signs of the diagonal entries of R . Then

$$Q = \begin{pmatrix} \hat{h}^T & \frac{\eta_n}{\eta_n} \\ \mathcal{S} \check{L}^T & -\frac{\mathcal{S} \check{L}^T \hat{h}}{\eta_n} \end{pmatrix}. \quad (4)$$

Moreover, the entries of the upper triangular matrix \check{L}^T appearing in (4) are for $j > i$,

$$\begin{aligned} (\check{L}^T)_{i,j} &= -\frac{\eta_i \eta_j}{\sqrt{\eta_{i+1}^2 + \cdots + \eta_n^2} \sqrt{\eta_i^2 + \cdots + \eta_n^2}} \\ &= -\left(\frac{1}{\|r^i\|^2} - \frac{1}{\|r^{i-1}\|^2} \right)^{1/2} (\|r^{j-1}\|^2 - \|r^j\|^2)^{1/2}, \end{aligned}$$

and the diagonal entries are

$$(\check{L}^T)_{i,i} = \frac{\sqrt{\eta_{i+1}^2 + \dots + \eta_n^2}}{\sqrt{\eta_i^2 + \dots + \eta_n^2}} = \frac{\|r^i\|}{\|r^{i-1}\|}.$$

Proof. We first remark that since $\|h\|^2 = \|\hat{h}\|^2 + \eta_n^2 = 1$,

$$\frac{1}{\eta_n} - \frac{\|\hat{h}\|^2}{\eta_n} = \eta_n.$$

From Lemma 3 we have $\hat{L}\hat{L}^T = R^T R - R^T \hat{h}\hat{h}^T R$. This implies

$$\hat{L}^T R^{-1} = \hat{L}^{-1} R^T (I - \hat{h}\hat{h}^T).$$

Hence, we have

$$(R^{-T} \hat{L})(\hat{L}^T R^{-1}) = I - \hat{h}\hat{h}^T = \check{L}\check{L}^T.$$

However we cannot identify $\hat{L}^T R^{-1}$, which appears in Q , with \check{L}^T since this last matrix is the upper triangular Cholesky factor of $I - \hat{h}\hat{h}^T$ and has positive diagonal entries whence the entries of the upper triangular matrix $\hat{L}^T R^{-1}$ may be negative. The matrices \hat{L}^T and R^{-1} being both upper triangular, we have

$$(\hat{L}^T R^{-1})_{i,i} = \frac{(\hat{L}^T)_{i,i}}{(R)_{i,i}}.$$

Since $(\hat{L}^T)_{i,i} > 0$, the diagonal entries $(\hat{L}^T R^{-1})_{i,i}$ have the sign of $(R)_{i,i}$. Let S be a diagonal matrix whose diagonal entry $(S)_{i,i}$ is plus or minus one according to the sign of $(R)_{i,i}$. Then we have $\hat{L}^T R^{-1} = S\check{L}^T$. This together with (3) proves the result.

Moreover, since \check{L} is the lower triangular Cholesky factor of a rank-one modification of the identity matrix, its entries can be computed. This can be done using the results of [7], where the authors considered a factorization $\mathcal{L}\mathcal{D}\mathcal{L}^T$ with \mathcal{L} lower triangular and \mathcal{D} diagonal. The matrix \mathcal{L} has a very special structure with ones on the diagonal and

$$(\mathcal{L})_{i,j} = \eta_i \gamma_j, \quad i = 2, \dots, n, \quad j = 1, \dots, n-1.$$

The values γ_j and the diagonal elements d_j of the diagonal matrix \mathcal{D} are computed by recurrences

$$\delta_1 = -1 = -\frac{1}{\|r^0\|^2},$$

and for $j = 1, \dots, n$,

$$d_j = 1 + \delta_j(\eta_j)^2, \quad \gamma_j = \delta_j \frac{\eta_j}{d_j}, \quad \delta_{j+1} = \frac{\delta_j}{d_j}.$$

It is easy to see that

$$d_j = \frac{\eta_{j+1}^2 + \dots + \eta_n^2}{\eta_j^2 + \dots + \eta_n^2} = \frac{\|r^j\|^2}{\|r^{j-1}\|^2}, \quad \delta_j = -\frac{1}{\eta_j^2 + \dots + \eta_n^2} = -\frac{1}{\|r^{j-1}\|^2},$$

and

$$\gamma_j = -\frac{\eta_j}{\eta_{j+1}^2 + \dots + \eta_n^2} = -\frac{\eta_j}{\|r^j\|^2}.$$

We have to take the square roots of the elements d_j to obtain the Cholesky factor $\check{L} = \mathcal{L}\sqrt{\mathcal{D}}$. The multiplying factor for the column j below the diagonal element is

$$\gamma_j \sqrt{d_j} = -\frac{\eta_j}{\sqrt{\eta_{j+1}^2 + \dots + \eta_n^2} \sqrt{\eta_j^2 + \dots + \eta_n^2}} = -\frac{\eta_j}{\|r^j\| \|r^{j-1}\|},$$

and therefore $(\check{L})_{i,j} = \eta_i \gamma_j \sqrt{d_j}$. By transposing this we obtain the entries we need for the matrix Q . The diagonal element of the column j is

$$(\check{L})_{j,j} = \sqrt{d_j} = \frac{\sqrt{\eta_{j+1}^2 + \dots + \eta_n^2}}{\sqrt{\eta_j^2 + \dots + \eta_n^2}} = \frac{\|r^j\|}{\|r^{j-1}\|}. \quad \square$$

From [Theorem 4](#) we see that the first row of the upper Hessenberg orthogonal matrix Q is h^T . Therefore the GMRES residual norm convergence (described by h) can be read from the first row of Q , the orthogonal factor of a QR factorization of H . Note that, in this factorization, the signs are such that the entries of the first row of Q are positive.

The matrix \check{L} depends only on the components η_j of h . So, up to the signs (which depend on the signs of the diagonal entries of R), the same property is true for Q . The results of [Theorem 4](#) give a complete description of the matrix Q in terms of the components η_j and therefore in terms of the GMRES residual norms.

In practical implementations of GMRES, the upper Hessenberg matrix H is reduced to upper triangular form by Givens rotations. An expression of the matrix Q using the sines and cosines of the rotations was given by P.N. Brown, see [\[3\]](#) and [\[4\]](#).

Note that one can also consider the matrix $\mathcal{R} = YCU^{-1}$. Unfortunately, the expression of \mathcal{R} is not as nice as the one for Q and since it is not relevant for the study of stagnation, we do not consider it furthermore.

5. Complete stagnation

Let us see what we obtain from the relations of the previous section when we have complete stagnation that is, h the first column of Y is e^n . This implies that the vector \hat{h} of the first $n-1$ components is identically zero and $\check{L} = I$. Using [Theorem 4](#) we immediately obtain the following result.

Theorem 5. *In case of complete stagnation, the orthogonal matrix Q in $H = Q\mathcal{R}$ is*

$$Q = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ & S & & 0 \end{pmatrix}, \quad (5)$$

where S is a diagonal matrix with ± 1 as diagonal entries according to the signs of the diagonal entries of R .

Note that [Theorem 5](#) implies that $H = Q\mathcal{R}$ is an upper Hessenberg matrix with a first row proportional to $(e^n)^T$; that is, all the elements of the first row are zero, except the last one which is equal to $-\alpha_0(\check{L}^{-T}e^{n-1})_{n-1}$.

Note also that since $h = e^n$, we have

$$b^T K = b^T WY = (e^n)^T Y = (e^1)^T.$$

Therefore, as it is well known (see [\[17\]](#)), we have $b^T A^j b = 0$, $j = 2, \dots, n-1$. This is a necessary and sufficient condition to have complete stagnation. However, with [Theorem 5](#), we are now able to characterize in a different way the matrices and right-hand sides for which we have complete stagnation.

Theorem 6. *We have complete stagnation in GMRES if and only if the nonderogatory real matrix A can be written as $A = W\mathcal{U}PW^T$, where W is orthogonal, \mathcal{U} is upper triangular and nonsingular and*

$$P = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ & I & & 0 \end{pmatrix} \quad (6)$$

is a permutation matrix. A right-hand side giving stagnation is $b = We^n$.

Proof. Assume that we have complete stagnation. From [Theorems 1 and 2](#)

$$\begin{aligned} A &= WYCY^{-1}W^T \\ &= W\mathcal{R}QW^T. \end{aligned}$$

The matrix Q has the structure given in [\(5\)](#). Such a matrix can be written as $Q = \check{S}P$ with P defined in [\(6\)](#) and \check{S} is a diagonal matrix with $(\check{S})_{1,1} = 1$ and $(\check{S})_{j,j} = (S)_{j-1,j-1}$, $j = 2, \dots, n$. The sign matrix \check{S} can be absorbed in the upper triangular matrix by defining $\mathcal{U} = \mathcal{R}\check{S}$. Moreover $b = Wh = We^n$ in the APS parametrization of [Theorem 1](#).

Conversely, let us assume that $A = W\mathcal{U}PW^T$ and $b = We^n$. From [\[8, Section 2, p. 466\]](#), we just need to show that the columns of W are a basis of AK_n . We proceed by induction. Since $W^T b = e^n$ and $Pe^n = e^1$ we have

$$Ab = W\mathcal{U}PW^T b = W\mathcal{U}e^1.$$

But $\mathcal{U}e^1$ is a vector proportional to e^1 . Hence Ab is proportional to w^1 , the first column of W . More generally, assume that $A^{j-1}b = Wq$ where q is a vector with only the first $j-1$ components nonzero. Then,

$$A^j b = A(A^{j-1}b) = W\mathcal{U}Pq.$$

The j first components of the vector Pq are 0, q_1, \dots, q_{j-1} . Multiplying the vector Pq by the nonsingular upper triangular matrix \mathcal{U} , we obtain a vector with the first j components nonzero showing that $A^j b$ is in the span of the j first columns of W . \square

The result of [Theorem 6](#) could have also been obtained from Theorem 2.2 in [\[2, p. 639\]](#). In case of complete stagnation the matrix denoted \hat{H} in [\[2\]](#) is the same as P . We note that $\mathcal{U}P$ is an upper Hessenberg matrix whose last column is proportional to e^1 .

Other equivalent characterizations can be obtained as shown in the following theorem.

Theorem 7. *We have complete stagnation in GMRES if and only if the nonderogatory real matrix A and the real right-hand side b can be written as in one of the four cases:*

- (1) $A = W\mathcal{U}PW^T, \quad b = We^n,$
- (2) $A = WP^T\mathcal{L}W^T, \quad b = We^n,$
- (3) $A = V\mathcal{P}\mathcal{U}V^T, \quad b = Ve^1,$
- (4) $A = V\mathcal{L}P^TV^T, \quad b = Ve^1,$

where W and V are orthogonal matrices, \mathcal{U} is a nonsingular upper triangular matrix, \mathcal{L} is a nonsingular lower triangular matrix and P is the permutation matrix defined in [\(6\)](#).

Proof. Case (1) has been proved in [Theorem 6](#). Let us consider case (2). If we have complete stagnation for A and b , then we have also complete stagnation for A^T and b . This was shown in [\[17\]](#) and follows from

$$b^T A^j b = b^T (A^j)^T b = b^T (A^T)^j b = 0, \quad j = 1, \dots, n-1.$$

The transpose of A is written as $WP^T\mathcal{U}^TW^T$. The matrix P^T is such that $P^Te^1 = e^n$ and $\mathcal{L} = \mathcal{U}^T$ is lower triangular.

Another characterization of the matrices leading to complete stagnation is obtained by using the orthogonal matrix V in the Arnoldi process. From $A = W\mathcal{R}QW^T$ and $W = VQ$, we have $A = VQ\mathcal{R}V^T$ as well as $b = We^n = Ve^1$. The matrix Q can be written as $Q = P\check{S}$ where P is defined in [\(6\)](#) and \check{S} is a diagonal matrix with $(\check{S})_{j,j} = (S)_{j,j}, j = 1, \dots, n-1$, and $(\check{S})_{n,n} = 1$. The sign matrix \check{S} can be absorbed in the upper triangular matrix by defining $\mathcal{U} = \check{S}\mathcal{R}$. It yields $A = V\mathcal{P}\mathcal{U}V^T$ and proves case (3). Case (4) is proved by considering as before the transpose matrix $A^T = V\mathcal{U}^TP^TV^T$ with $\mathcal{L} = \mathcal{U}^T$. \square

Corollary 8. *We have complete stagnation in GMRES if and only if the nonderogatory real matrix A is $A = VBV^T$ with V orthogonal, the matrix B being one of the four matrices below and the real right-hand sides b are chosen properly,*

- (1) B is an upper Hessenberg matrix with the last column proportional to e_1 and $b = Ve_n$,
- (2) B is a lower Hessenberg matrix with the last row proportional to e_1^T and $b = Ve_n$,
- (3) B is an upper Hessenberg matrix with the first row proportional to e_n^T and $b = Ve_1$,
- (4) B is a lower Hessenberg matrix with the first column proportional to e_n and $b = Ve_1$.

Proof. These results are readily obtained from [Theorem 7](#). For instance in case (1) of [Theorem 7](#), multiplying the upper triangular matrix \mathcal{U} to the right by the matrix P in [\(6\)](#) gives an upper Hessenberg matrix whose last column is zero except for the first entry. \square

6. Partial stagnation

Assume that we have partial stagnation as defined in Section 1. Then, we can characterize the orthogonal matrix Q in the QR factorization of H .

Theorem 9. *In case of partial stagnation of GMRES the columns $j = m+1, \dots, m+p-1$, of Q are zero except for the subdiagonal entries $(j+1, j)$ which are ± 1 . The rows $i = m+2, \dots, m+p$, are zero for columns i to n .*

Proof. We apply [Theorem 4](#). We have seen that $Q_{i,j}$ is proportional to $\eta_{i+1}\eta_j$ for $1 < i \leq j < n$. Therefore, the columns $m+1$ to $m+p-1$ are zero except for the subdiagonal entries which are ± 1 depending on the sign of the corresponding

diagonal entry of R . The rows i from $m + 2$ to $m + p$ are zero for columns i to $n - 1$. The entries in the last column for these rows are zero because of the structure of \tilde{L}^T and \hat{h} . \square

Then we can characterize the matrices and right-hand sides leading to partial stagnation.

Theorem 10. *We have partial stagnation in GMRES for iteration $m \geq 0$ to iteration $m + p - 1$ if and only if the nonderogatory real matrix A and the real right-hand side b can be written as*

$$A = W\mathcal{R}QW^T, \quad b = WQ^Te^1,$$

where W is orthogonal, \mathcal{R} is upper triangular, Q is orthogonal with the sparsity structure defined in Theorem 9. We have also the same partial stagnation if and only if $A = VQ\mathcal{R}V^T$ and $b = Ve^1$ with V orthogonal and Q and \mathcal{R} as before.

Proof. Assume that we have partial stagnation as defined above. From Theorems 1 and 2,

$$A = WYCY^{-1}W^T = W\mathcal{R}QW^T.$$

We have seen from Theorems 4 and 9 that Q has the required structure and values. Moreover $b = Wh = WQ^Te^1$ in the APS parametrization. The proof of the converse is essentially the same as in Theorem 6. The proof of the other assertion follows from the relation between W and V in the APS parametrization. \square

Of course, one can easily extend this theorem to the case where we have several sequences of stagnation separated with phases of strict decrease of the residual norm. We also remark that partial stagnation does not lead to orthogonality of b with respect to columns of K (that is, $b^T A^j b = 0$) like in the case of complete stagnation except when partial stagnation occurs at the beginning of the iterations that is, $m = 0$.

7. Conclusions

In this paper we have given necessary and sufficient conditions for the complete and partial stagnation of GMRES. Extension to complex matrices is not too difficult. Unfortunately, given a matrix A , these conditions are not easy to check in practice. However, it is hoped that our results could lead to a better understanding of GMRES behavior. Theorems 6 and 10 can also be used to generate examples with complete or partial stagnation.

Acknowledgements

This paper was started in 2010 during a visit to the Nečas Center of Charles University in Prague supported by the grant Jindřich Nečas Center for mathematical modeling, project LC06052 financed by MSM. The author thanks particularly Zdeněk Strakoš and Miroslav Rozložník for their kind hospitality. Detailed remarks and comments from the referees lead to improvements in the presentation.

References

- [1] M. Arioli, An analysis of GMRES worst case convergence, 2010, manuscript.
- [2] M. Arioli, V. Pták, Z. Strakoš, Krylov sequences of maximal length and convergence of GMRES, BIT Numer. Math. 38 (4) (1998) 636–643.
- [3] P. Brown, A theoretical comparison of the Arnoldi and GMRES algorithms, SIAM J. Sci. Statist. Comput. 12 (1) (1991) 58–78.
- [4] P. Brown, A. Hindmarsh, Reduced storage methods in stiff ODE systems, J. Appl. Math. Comput. 31 (1989) 40–91.
- [5] J. Duintjer Tebbens, G. Meurant, Any Ritz value behavior is possible for Arnoldi and for GMRES, SIAM J. Matrix Anal. Appl. 33 (3) (2012) 958–978.
- [6] J. Duintjer Tebbens, G. Meurant, Prescribing the behavior of early terminating GMRES and Arnoldi iterations, Numerical Algorithms (2013), in press, <http://dx.doi.org/10.1007/s11075-013-9695-x>.
- [7] P. Gill, G. Golub, W. Murray, M. Saunders, Methods for modifying matrix factorizations, Math. Comp. 28 (126) (1974) 505–535.
- [8] A. Greenbaum, V. Pták, Z. Strakoš, Any nonincreasing convergence curve is possible for GMRES, SIAM J. Matrix Anal. Appl. 17 (3) (1996) 465–469.
- [9] A. Greenbaum, Z. Strakoš, Matrices that generate the same Krylov residual spaces, in: Recent Advances in Iterative Methods, in: IMA Vol. Math. Appl., vol. 60, Springer, New York, 1994, pp. 95–118.
- [10] J. Liesen, P. Tichý, The worst-case GMRES for normal matrices, BIT Numer. Math. 44 (2004) 79–98.
- [11] G. Meurant, The complete stagnation of GMRES for $n \leq 4$, Electron. Trans. Numer. Anal. 39 (2012) 75–101.
- [12] G. Meurant, GMRES and the Arioli, Pták and Strakoš parametrization, BIT Numer. Math. 52 (3) (2012) 687–702.
- [13] Y. Saad, M. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Statist. Comput. 7 (3) (1986) 856–869.
- [14] V. Simoncini, On a non-stagnation condition for GMRES and application to saddle point matrices, Electron. Trans. Numer. Anal. 37 (2010) 202–213.
- [15] V. Simoncini, D. Szyld, New conditions for non-stagnation of minimal residual methods, Numer. Math. 109 (3) (2008) 477–487.
- [16] I. Zavorin, Analysis of GMRES convergence by spectral factorization of the Krylov matrix, PhD thesis, University of Maryland, 2001.
- [17] I. Zavorin, D. O’Leary, H. Elman, Complete stagnation of GMRES, Linear Algebra Appl. 367 (2003) 165–183.