# On the convergence curves that can be generated by restarted GMRES

Jurjen Duintjer Tebbens

joint work with

Gérard Meurant

Institute of Computer Science,
Academy of Sciences of the Czech Republic

June 12, XIX-th. Householder Symposium, Spa, Belgium, 2014

# Outline

# Introduction

- GMRES(m) is among the most widely used Krylov methods for solving linear systems with large, sparse and nonsymmetric matrices.

# Introduction

- GMRES(m) is among the most widely used Krylov methods for solving linear systems with large, sparse and nonsymmetric matrices.
- Many results on convergence, some inherited from the theory for full GMRES, some focussing on restart (e.g. non-stagnation conditions).

# Introduction

- GMRES(m) is among the most widely used Krylov methods for solving linear systems with large, sparse and nonsymmetric matrices.
- Many results on convergence, some inherited from the theory for full GMRES, some focussing on restart (e.g. non-stagnation conditions).

In this talk we investigate the behavior that GMRES(m) may possibly exhibit. More precisely:

# Introduction

- GMRES(m) is among the most widely used Krylov methods for solving linear systems with large, sparse and nonsymmetric matrices.
- Many results on convergence, some inherited from the theory for full GMRES, some focussing on restart (e.g. non-stagnation conditions).

In this talk we investigate the behavior that GMRES(m) may possibly exhibit. More precisely:

- What convergence curves can GMRES(m) generate ?

# Introduction

- GMRES(m) is among the most widely used Krylov methods for solving linear systems with large, sparse and nonsymmetric matrices.
- Many results on convergence, some inherited from the theory for full GMRES, some focussing on restart (e.g. non-stagnation conditions).

In this talk we investigate the behavior that GMRES(m) may possibly exhibit. More precisely:

- What convergence curves can GMRES(m) generate ?
- What convergence curves can GMRES(m) generate with what spectrum?

# Introduction

- GMRES(m) is among the most widely used Krylov methods for solving linear systems with large, sparse and nonsymmetric matrices.
- Many results on convergence, some inherited from the theory for full GMRES, some focussing on restart (e.g. non-stagnation conditions).

In this talk we investigate the behavior that GMRES(m) may possibly exhibit. More precisely:

- What convergence curves can GMRES(m) generate ?
- What convergence curves can GMRES(m) generate with what spectrum?
- What Ritz values can GMRES(m) generate ?

# Introduction

- GMRES(m) is among the most widely used Krylov methods for solving linear systems with large, sparse and nonsymmetric matrices.
- Many results on convergence, some inherited from the theory for full GMRES, some focussing on restart (e.g. non-stagnation conditions).

In this talk we investigate the behavior that GMRES(m) may possibly exhibit. More precisely:

- What convergence curves can GMRES(m) generate ?
- What convergence curves can GMRES(m) generate with what spectrum?
- What Ritz values can GMRES(m) generate ?
- What convergence curves can GMRES(m) generate with what Ritz values?

# Introduction

Similar questions were answered for full GMRES:

# Introduction

Similar questions were answered for full GMRES:

- *Any nonincreasing convergence curve is possible for GMRES* [Greenbaum & Pták & Strakoš, 1996] showed additionally that this is possible with any nonzero spectrum

# Introduction

Similar questions were answered for full GMRES:

- *Any nonincreasing convergence curve is possible for GMRES* [Greenbaum & Pták & Strakoš, 1996] showed additionally that this is possible with any nonzero spectrum
- A parametrization was given of all matrices with a prescribed spectrum and right-hand sides generating a prescribed convergence curve [Arioli & Pták & Strakoš, 1998]

# Introduction

Similar questions were answered for full GMRES:

- *Any nonincreasing convergence curve is possible for GMRES* [Greenbaum & Pták & Strakoš, 1996] showed additionally that this is possible with any nonzero spectrum
- A parametrization was given of all matrices with a prescribed spectrum and right-hand sides generating a prescribed convergence curve [Arioli & Pták & Strakoš, 1998]
- *Any Ritz-value behavior is possible for Arnoldi and for GMRES* [DT & Meurant, 2012] showed that any Ritz values can be prescribed in Arnoldi processes

# Introduction

Similar questions were answered for full GMRES:

- *Any nonincreasing convergence curve is possible for GMRES* [Greenbaum & Pták & Strakoš, 1996] showed additionally that this is possible with any nonzero spectrum
- A parametrization was given of all matrices with a prescribed spectrum and right-hand sides generating a prescribed convergence curve [Arioli & Pták & Strakoš, 1998]
- *Any Ritz-value behavior is possible for Arnoldi and for GMRES* [DT & Meurant, 2012] showed that any Ritz values can be prescribed in Arnoldi processes
- A parametrization to see this was given in [DT & Meurant, 2013]:

# The parametrization for full GMRES

To prescribe Ritz values and residual norms in GMRES [DT & Meurant 2014]:

# The parametrization for full GMRES

To prescribe Ritz values and residual norms in GMRES [DT & Meurant 2014]:

- Choose a unitary matrix $V$ and put $b = Ve_1$ and

$$A = VHV^*, \qquad H \text{ upper Hessenberg.}$$

# The parametrization for full GMRES

To prescribe Ritz values and residual norms in GMRES [DT & Meurant 2014]:

- Choose a unitary matrix $V$ and put $b = Ve_1$ and

$$A = VHV^*, \qquad H \text{ upper Hessenberg.}$$

- To force the desired eigenvalues, $H$ will be of the form

$$H = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

where $C$ is the companion matrix for the prescribed spectrum.

# The parametrization for full GMRES

To prescribe Ritz values and residual norms in GMRES [DT & Meurant 2014]:

- Choose a unitary matrix $V$ and put $b = Ve_1$ and

$$A = VHV^*, \qquad H \text{ upper Hessenberg.}$$

- To force the desired eigenvalues, $H$ will be of the form

$$H = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

  where $C$ is the companion matrix for the prescribed spectrum.
- Let

$$U = \begin{bmatrix} & g^T & \\ 0 & T & \end{bmatrix}$$

# The parametrization for full GMRES

- To force the residual norm sequence $f(0) \geq \cdots \geq f(n-1) > 0$, the first row $g^T$ of $U$ has entries

$$g_1 = \frac{1}{f(0)}, \qquad g_k = \sqrt{\frac{1}{f(k-1)^2} - \frac{1}{f(k-2)^2}}, \qquad k = 2, \ldots, n.$$

# The parametrization for full GMRES

- To force the residual norm sequence $f(0) \geq \cdots \geq f(n-1) > 0$, the first row $g^T$ of $U$ has entries

$$g_1 = \frac{1}{f(0)}, \qquad g_k = \sqrt{\frac{1}{f(k-1)^2} - \frac{1}{f(k-2)^2}}, \qquad k = 2, \ldots, n.$$

- To force at iteration $k$ the Ritz values $\rho_1^{(k)} \ldots \rho_k^{(k)}$, $k = 1, \ldots, n-1$, the remaining entries of $U$ must satisfy

$$\prod_{i=1}^{k} (\lambda - \rho_i^{(k)}) = g_{k+1} + \sum_{i=1}^{k} t_{i,k} \lambda^i.$$

# The parametrization for full GMRES

- To force the residual norm sequence $f(0) \geq \cdots \geq f(n-1) > 0$, the first row $g^T$ of $U$ has entries

$$g_1 = \frac{1}{f(0)}, \qquad g_k = \sqrt{\frac{1}{f(k-1)^2} - \frac{1}{f(k-2)^2}}, \qquad k = 2, \ldots, n.$$

- To force at iteration $k$ the Ritz values $\rho_1^{(k)} \ldots \rho_k^{(k)}$, $k = 1, \ldots, n-1$, the remaining entries of $U$ must satisfy

$$\prod_{i=1}^{k} (\lambda - \rho_i^{(k)}) = g_{k+1} + \sum_{i=1}^{k} t_{i,k} \lambda^i.$$

Prescribing residual norms in restarted GMRES was considered in the paper [Vecharinsky & Langou 2011].

# Results for restarted GMRES

The paper assumes a special situation in GMRES(m):

1. During every restart cycle, all residual norms stagnate except for the very last iteration inside the cycle.
2. In this very last iteration it is assumed that the residual norm is strictly decreasing.

# Results for restarted GMRES

The paper assumes a special situation in GMRES(m):

1. During every restart cycle, all residual norms stagnate except for the very last iteration inside the cycle.
2. In this very last iteration it is assumed that the residual norm is strictly decreasing.

**Theorem 1** [Vecharinsky & Langou 2011]. *Let $k$ positive decreasing numbers*

$$f(0) > f(1) > \cdots > f(k-1) > 0,$$

*and $n$ complex nonzero numbers $\lambda_1, \ldots, \lambda_n$ be given. With the assumptions 1. and 2. above, let the very last residual at the end of the $j$th cycle be denoted by $\bar{r}_j$. If $km < n$, then:*

- There exists a matrix $A$ of order $n$ with a right hand side such that GMRES(m) generates residual norms at the end of cycles satisfying

$$\|\bar{r}_i\| = f(i), \qquad i = 0, 1, \ldots, k.$$

- The matrix $A$ has the eigenvalues $\lambda_1, \ldots, \lambda_n$.

# Residual norms inside cycles

- To prescribe residual norms inside restart cycles we will construct linear systems generating Hessenberg matrices with prescribed entries:

# Residual norms inside cycles

- To prescribe residual norms inside restart cycles we will construct linear systems generating Hessenberg matrices with prescribed entries:
- To generate in the $k$th cycle the residual norms

$$f^{(k)}(0) \geq \cdots \geq f^{(k)}(m),$$

we choose the $k$th Hessenberg matrix of dimension $(m+1) \times m$ is to be of the form

$$\hat{H}_m^{(k)} = \begin{bmatrix} g_1^{(k)} & \dots & g_{m+1}^{(k)} \\ 0 & T_m^{(k)} & \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I_m \end{bmatrix} \begin{bmatrix} g_1^{(k)} & \dots & g_m^{(k)} \\ 0 & T_{m-1}^{(k)} & \end{bmatrix},$$

where

$$g_1^{(k)} = 1/(f^{(k)}(0))^2, \quad g_i^{(k)} = \sqrt{\frac{1}{(f(i-1)^{(k)})^2} - \frac{1}{(f(i-2)^{(k)})^2}}.$$

# Outline

# Comparison of the first two cycles

The constructed matrices and right-hand sides will, as before, have the form

$$A = VHV^*, \quad H \text{ upper Hessenberg}, \qquad b = Ve_1.$$

# Comparison of the first two cycles

The constructed matrices and right-hand sides will, as before, have the form

$$A = VHV^*, \quad H \text{ upper Hessenberg}, \qquad b = Ve_1.$$

GMRES and restarted GMRES being unitarily invariant, we concentrate on the choice of $H$.

# Comparison of the first two cycles

The constructed matrices and right-hand sides will, as before, have the form

$$A = VHV^*, \quad H \text{ upper Hessenberg}, \qquad b = Ve_1.$$

GMRES and restarted GMRES being unitarily invariant, we concentrate on the choice of $H$.
After $m$ iterations of the initial cycle we wish to have the Arnoldi decomposition

$$AV_m^{(1)} \;\; = \;\; V_{m+1}^{(1)} \hat{H}_m^{(1)}, \quad V_{m+1}^{(1)^*} V_{m+1}^{(1)} = I_{m+1}.$$

# Comparison of the first two cycles

The constructed matrices and right-hand sides will, as before, have the form

$$A = VHV^*, \quad H \text{ upper Hessenberg}, \qquad b = Ve_1.$$

GMRES and restarted GMRES being unitarily invariant, we concentrate on the choice of $H$.

After $m$ iterations of the initial cycle we wish to have the Arnoldi decomposition

$$AV_m^{(1)} \;=\; V_{m+1}^{(1)}\hat{H}_m^{(1)}, \quad V_{m+1}^{(1)^*}V_{m+1}^{(1)} = I_{m+1}.$$

To generate $\hat{H}_m^{(1)}$ we need

$$H \begin{bmatrix} I_m \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{H}_m^{(1)} \\ 0 \end{bmatrix}.$$

# Comparison of the first two cycles

The constructed matrices and right-hand sides will, as before, have the form

$$A = VHV^*, \quad H \text{ upper Hessenberg}, \qquad b = Ve_1.$$

GMRES and restarted GMRES being unitarily invariant, we concentrate on the choice of $H$.

After $m$ iterations of the initial cycle we wish to have the Arnoldi decomposition

$$AV_m^{(1)} \;\; = \;\; V_{m+1}^{(1)} \hat{H}_m^{(1)}, \quad {V_{m+1}^{(1)}}^* V_{m+1}^{(1)} = I_{m+1}.$$

To generate $\hat{H}_m^{(1)}$ we need

$$H \begin{bmatrix} I_m \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{H}_m^{(1)} \\ 0 \end{bmatrix}.$$

The $m$ iterations of the second cycle should give the Arnoldi decomposition

$$AV_m^{(2)} = V_{m+1}^{(2)} \hat{H}_m^{(2)}, \quad {V_{m+1}^{(2)}}^* V_{m+1}^{(2)} = I_{m+1}.$$

# Comparison of the first two cycles

If $r_m^{(1)}$ is the residual vector at the end of the first cycle, then

$$V_{m+1}^{(2)} e_1 = \frac{r_m^{(1)}}{\|r_m^{(1)}\|} \equiv V_{m+1}^{(1)} z^{(1)}.$$

# Comparison of the first two cycles

If $r_m^{(1)}$ is the residual vector at the end of the first cycle, then

$$V_{m+1}^{(2)} e_1 = \frac{r_m^{(1)}}{\|r_m^{(1)}\|} \equiv V_{m+1}^{(1)} z^{(1)}.$$

**Lemma 1.** *If the initial cycle generates residual norms*

$$\|r_0^{(1)}\| \geq \|r_1^{(1)}\| \geq \cdots \geq \|r_m^{(1)}\|$$

*and we define*

$$g_0^{(1)} = 1/\|r_0^{(1)}\|^2, \quad g_i^{(1)} = \sqrt{\frac{1}{\|r_{i-1}^{(1)}\|^2} - \frac{1}{\|r_{i-2}^{(1)}\|^2}}, \quad g^{(1)} \equiv [g_0^{(1)}, \ldots, g_m^{(1)}]^T,$$

*then* $z^{(1)} = g^{(1)}$ *and* $g^{(1)} \perp \hat{H}_m^{(1)} e_i, \quad i = 1, \ldots, m.$

# Comparison of the first two cycles

If $r_m^{(1)}$ is the residual vector at the end of the first cycle, then

$$V_{m+1}^{(2)} e_1 = \frac{r_m^{(1)}}{\|r_m^{(1)}\|} \equiv V_{m+1}^{(1)} z^{(1)}.$$

**Lemma 1.** *If the initial cycle generates residual norms*

$$\|r_0^{(1)}\| \geq \|r_1^{(1)}\| \geq \cdots \geq \|r_m^{(1)}\|$$

*and we define*

$$g_0^{(1)} = 1/\|r_0^{(1)}\|^2, \quad g_i^{(1)} = \sqrt{\frac{1}{\|r_{i-1}^{(1)}\|^2} - \frac{1}{\|r_{i-2}^{(1)}\|^2}}, \quad g^{(1)} \equiv [g_0^{(1)}, \ldots, g_m^{(1)}]^T,$$

*then* $z^{(1)} = g^{(1)}$ *and* $g^{(1)} \perp \hat{H}_m^{(1)} e_i, \quad i = 1, \ldots, m.$

Here is a relation between the small Hessenberg matrix $\hat{H}_m^{(2)}$ of the second cycle and the large Hessenberg matrix $H$ in the matrix $A = VHV^*$ which we construct:

# Comparison of the first two cycles

**Lemma 2.** *The matrix* $\hat{H}_m^{(2)}$ *is the Hessenberg matrix generated by* $m$ *iterations of the Arnoldi process with input* $H$ *and* $\left[\begin{array}{cc} g^{(1)T} & 0 \end{array}\right]^T$*:*

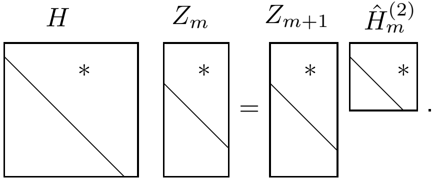
$$HZ_m = Z_{m+1}\hat{H}_m^{(2)}, \quad Z_{m+1}e_1 = \left[\begin{array}{c} g^{(1)} \\ 0 \end{array}\right], \quad Z_{m+1}^* Z_{m+1} = I_{m+1}. \quad (1)$$

# Comparison of the first two cycles

**Lemma 2.** *The matrix* $\hat{H}_m^{(2)}$ *is the Hessenberg matrix generated by* $m$ *iterations of the Arnoldi process with input* $H$ *and* $\begin{bmatrix} g^{(1)T} & 0 \end{bmatrix}^T$*:*

$$HZ_m = Z_{m+1}\hat{H}_m^{(2)}, \quad Z_{m+1}e_1 = \begin{bmatrix} g^{(1)} \\ 0 \end{bmatrix}, \quad Z_{m+1}^* Z_{m+1} = I_{m+1}. \quad (1)$$

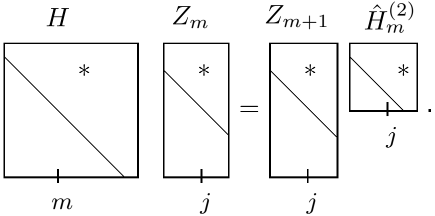
Can we construct the columns $m+1, m+2, \ldots, 2m$ of $H$ such that (1) is satisfied with a prescribed Hessenberg matrix $\hat{H}_m^{(2)}$ an "inverse" Arnoldi problem? This will depend on the number of trailing zeroes in $g^{(1)}$ because the Arnoldi decomposition has the structure



# Stagnation at the end of cycles

Because $g_i^{(1)} = \sqrt{\frac{1}{\|r_{i-1}^{(1)}\|^2} - \frac{1}{\|r_{i-2}^{(1)}\|^2}}$, with $j$ stagnation steps at the end of the first restart cycle, the trailing $j$ entries of $g^{(1)}$, are zero.

Then the Arnoldi decomposition $HZ_m = Z_{m+1}\hat{H}_m^{(2)}$ looks like




The first $j$ columns of the Hessenberg matrix of the second cycle $\hat{H}_m^{(2)}$ are fully determined by $\hat{H}_m^{(1)}$ and $g^{(1)}$ - they cannot be prescribed.

# Stagnation at the end of cycles

**Lemma 3** *If the last $j$ residual norms stagnate in the initial cycle, i.e.*

$$\|r_0^{(1)}\| \geq \|r_1^{(1)}\| \geq \cdots \geq \|r_{m-j-1}^{(1)}\| > \|r_{m-j}^{(1)}\| = \cdots = \|r_m^{(1)}\|$$

*then the first row of $\hat{H}_m^{(2)}$ is zero on positions one till $j$.* Proof: For $i < j$,

$$\begin{aligned} e_1^T \begin{bmatrix} \hat{H}_m^{(2)} \\ 0 \end{bmatrix} e_i &= e_1^T \begin{bmatrix} Z_{m+1}^* H Z_m \\ 0 \end{bmatrix} e_i \\ &= \begin{bmatrix} {g^{(1)}}^* & 0 \end{bmatrix} H Z_m e_i = {g^{(1)}}^* \hat{H}_m^{(1)} \begin{bmatrix} Z_m e_i \\ 0 \end{bmatrix} = 0. \end{aligned}$$

# Stagnation at the end of cycles

**Lemma 3** *If the last $j$ residual norms stagnate in the initial cycle, i.e.*

$$\|r_0^{(1)}\| \geq \|r_1^{(1)}\| \geq \cdots \geq \|r_{m-j-1}^{(1)}\| > \|r_{m-j}^{(1)}\| = \cdots = \|r_m^{(1)}\|$$

*then the first row of $\hat{H}_m^{(2)}$ is zero on positions one till $j$.* Proof: For $i < j$,

$$\begin{aligned} e_1^T \begin{bmatrix} \hat{H}_m^{(2)} \\ 0 \end{bmatrix} e_i &= e_1^T \begin{bmatrix} Z_{m+1}^* H Z_m \\ 0 \end{bmatrix} e_i \\ &= \begin{bmatrix} {g^{(1)}}^* & 0 \end{bmatrix} H Z_m e_i = {g^{(1)}}^* \hat{H}_m^{(1)} \begin{bmatrix} Z_m e_i \\ 0 \end{bmatrix} = 0. \end{aligned}$$

**Corollary** *If the last $j$ residual norms stagnate in the initial cycle, i.e.*

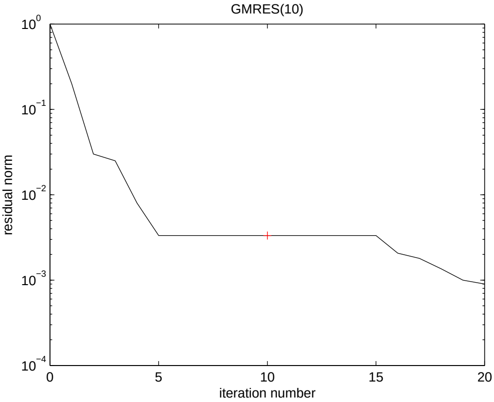
$$\|r_0^{(1)}\| \geq \|r_1^{(1)}\| \geq \cdots \geq \|r_{m-j-1}^{(1)}\| > \|r_{m-j}^{(1)}\| = \cdots = \|r_m^{(1)}\|$$

*then the first $j$ residual norms stagnate in the second cycle,*

$$\|r_0^{(2)}\| = \|r_1^{(2)}\| = \cdots = \|r_j^{(2)}\|.$$

# Stagnation at the end of cycles

Hence stagnation in one cycle is literally mirrored in the next cycle and we cannot prescribe any residual norm history!



# Prescribed Hessenberg matrices in GMRES(m)

If we assume no restart cycles stagnate in their last iterations, we can try to solve the "inverse" Arnoldi problem:

$$HZ_m = Z_{m+1}\hat{H}_m^{(2)}, \quad Z_{m+1}e_1 = \begin{bmatrix} g^{(1)} \\ 0 \end{bmatrix}, \quad Z_{m+1}^* Z_{m+1} = I_{m+1}.$$

# Prescribed Hessenberg matrices in GMRES(m)

If we assume no restart cycles stagnate in their last iterations, we can try to solve the "inverse" Arnoldi problem:

$$HZ_m = Z_{m+1}\hat{H}_m^{(2)}, \quad Z_{m+1}e_1 = \begin{bmatrix} g^{(1)} \\ 0 \end{bmatrix}, \quad Z_{m+1}^* Z_{m+1} = I_{m+1}.$$

This can be done in many ways, depending on the choice of columns of $Z_{m+1}$ (the Arnoldi vectors). Let us consider the most straightforward, "canonical" choice

$$Z_{m+1} \equiv \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \\ \vdots & 0 \end{bmatrix}.$$

# Prescribed Hessenberg matrices in GMRES(m)

If we assume no restart cycles stagnate in their last iterations, we can try to solve the "inverse" Arnoldi problem:

$$HZ_m = Z_{m+1}\hat{H}_m^{(2)}, \quad Z_{m+1}e_1 = \left[\begin{array}{c} g^{(1)} \\ 0 \end{array}\right], \quad Z_{m+1}^* Z_{m+1} = I_{m+1}.$$

This can be done in many ways, depending on the choice of columns of $Z_{m+1}$ (the Arnoldi vectors). Let us consider the most straightforward, "canonical" choice

$$Z_{m+1} \equiv \left[\begin{array}{cc} g^{(1)} & 0 \\ 0 & I_m \\ \vdots & 0 \end{array}\right].$$

Equating the subsequent columns in $HZ_m = Z_{m+1}\hat{H}_m^{(2)}$ gives explicit values for columns $m+1$ till $2m$ of $H$ and leads to:

# Prescribed Hessenberg matrices in GMRES(m)

**Theorem 2** **[DT & Meurant 2014?]** *Let $A \in \mathbb{C}^{n \times n}$ and $b \in \mathbb{C}^n$ and let for $km < n$,*

$$\hat{H}_m^{(1)}, \ldots, \hat{H}_m^{(k)} \in \mathbb{C}^{(m+1) \times m}$$

*be $k$ given unreduced upper Hessenberg matrices with positive subdiagonal and nonsingular leading $m \times m$ principal submatrix. Then the $k$th cycle of GMRES(m) applied to $A$ and $b$ generates the Hessenberg matrix*

$$\hat{H}_m^{(k)}$$

*if*

$$A = VHV^*, \qquad b = Ve_1,$$

*where $V$ is unitary, $H$ is upper Hessenberg and the columns $(k-1)m + 1$ till $km$ corresponding to the $k$th cycle are of the form:*

# Prescribed Hessenberg matrices in GMRES(m)

$$H\left[e_{(k-1)m+1},\ldots,e_{km}\right]=\begin{bmatrix} & \gamma^{(1)}g^{(1)}\hat{h}^T_{(k)} \\ & \vdots \\ \vdots & \\ & \gamma^{(k-1)}g^{(k-1)}\hat{h}^T_{(k)} \\ t^{(k)} & g^{(k)}\hat{h}^T_{(k)} \\ 0 & \begin{bmatrix} 0 & I_m \end{bmatrix}\hat{H}^{(k)}_m\begin{bmatrix} 0 \\ I_{m-1}\end{bmatrix} \\ 0 & 0 \end{bmatrix},\quad \textit{where}$$

$$\hat{h}^T_{(k)}=e_1^T\hat{H}^{(k)}_m\begin{bmatrix} 0 \\ I_{m-1}\end{bmatrix},\quad \gamma^{(\ell)}=\left(\prod_{i=\ell+1}^{k-1} g_1^{(i-1)}\right)\quad \text{and}$$

$$t^{(k)}=\frac{1}{\zeta^{(k-1)}_{m+1}}\left[(h^{(k)}_{1,1}z^{(k-1)}-\hat{H}^{(k-1)}_m[\zeta^{(k-1)}_1,\ldots,\zeta^{(k-1)}_m]^T),h^{(k)}_{2,1}\right]^T.$$

# Prescribed Hessenberg matrices in GMRES(m)

Thus we know how to generate, by the right choice of columns of $H$, arbitrary Hessenberg matrices during *all* restarts. Therefore we may prescribe not only GMRES residual norms *inside* cycles, but also Ritz values and possibly other values (singular values [Ernst & Eiermann 2001], harmonic Ritz values, etc. ...).

# Prescribed Hessenberg matrices in GMRES(m)

Thus we know how to generate, by the right choice of columns of $H$, arbitrary Hessenberg matrices during *all* restarts. Therefore we may prescribe not only GMRES residual norms *inside* cycles, but also Ritz values and possibly other values (singular values [Ernst & Eiermann 2001], harmonic Ritz values, etc. ...).

**Remark:** Prescribing $k$ restarts under the condition $km < n$ means that in the parametrization

$$A = VHV^*, \qquad b = \|b\|Ve_1,$$

we put conditions on the first $km < n$ columns of $H$ only. The last column can be chosen arbitrarily. It can be checked (see, e.g., [Parlett & Strang 2008], that any nonzero spectrum of $A$ is possible with an appropriate choice of the last column.

# Outline

# Other constructions

- We constructed matrices and right-hand sides yielding a sequence of prescribed Hessenberg matrices if the GMRES process is restarted.
- The construction found the entries of $H$ using the "canonical" choice

$$Z_{m+1} \equiv \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \\ \vdots & 0 \end{bmatrix}$$

of $Z_{m+1}$ in $HZ_m = Z_{m+1}\hat{H}_m^{(2)}$.

# Other constructions

- We constructed matrices and right-hand sides yielding a sequence of prescribed Hessenberg matrices if the GMRES process is restarted.
- The construction found the entries of $H$ using the "canonical" choice

$$Z_{m+1} \equiv \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \\ \vdots & 0 \end{bmatrix}$$

of $Z_{m+1}$ in $HZ_m = Z_{m+1}\hat{H}_m^{(2)}$.

- Other choices lead to different parametrizations, e.g.

$$\tilde{H} = \begin{bmatrix} I_{m+1} & \hat{H}_m^{(1)} S_0 & 0 \\ 0 & S_1 & 0 \\ 0 & 0 & I_{n-2m-1} \end{bmatrix}^{-1} H \begin{bmatrix} I_{m+1} & \hat{H}_m^{(1)} S_0 & 0 \\ 0 & S_1 & 0 \\ 0 & 0 & I_{n-2m-1} \end{bmatrix}$$

where $S_1$ is nonsingular upper triangular such that $S_1^* S_1 - I_m$ is positive definite and $S_0$ satisfies $(\hat{H}_m^{(1)} S_0)^* \hat{H}_m^{(1)} S_0 = S_1^* S_1 - I_m$.

# Relation with full GMRES

- All constructions give the same behavior of GMRES(m) (it generates the same small Hessenberg matrices)
- but obviously not the same behavior if full GMRES is applied

# Relation with full GMRES

- All constructions give the same behavior of GMRES(m) (it generates the same small Hessenberg matrices)
- but obviously not the same behavior if full GMRES is applied

**Theorem 3** **[DT & Meurant 2014?]** *Consider the linear system constructed in Theorem 2 using the "canonical" basis for $Z_{m+1}$. The residual norms generated when GMRES($m$) is applied to this system are the same as when full GMRES is applied to this system.*

# Relation with full GMRES

- All constructions give the same behavior of GMRES(m) (it generates the same small Hessenberg matrices)
- but obviously not the same behavior if full GMRES is applied

**Theorem 3** **[DT & Meurant 2014?]** *Consider the linear system constructed in Theorem 2 using the "canonical" basis for $Z_{m+1}$. The residual norms generated when GMRES($m$) is applied to this system are the same as when full GMRES is applied to this system.*

This means the linear systems in Theorem 2 represent the optimal situation for GMRES(m), where restarting is as efficient as running full GMRES.

# Relation with full GMRES

- All constructions give the same behavior of GMRES(m) (it generates the same small Hessenberg matrices)
- but obviously not the same behavior if full GMRES is applied

**Theorem 3** **[DT & Meurant 2014?]** *Consider the linear system constructed in Theorem 2 using the "canonical" basis for $Z_{m+1}$. The residual norms generated when GMRES($m$) is applied to this system are the same as when full GMRES is applied to this system.*

This means the linear systems in Theorem 2 represent the optimal situation for GMRES(m), where restarting is as efficient as running full GMRES.
Equivalently, we constructed systems where full GMRES can be computed with short recurrences ...

# Relation with full GMRES

The linear systems in Theorem 2 also enable to construct situations where GMRES(m) converges faster than GMRES(m+i) for some $i > 0$:

# Relation with full GMRES

The linear systems in Theorem 2 also enable to construct situations where GMRES(m) converges faster than GMRES(m+i) for some $i > 0$:

With GMRES(m) exhibiting the optimal convergence speed of full GMRES, no GMRES(m+i) process can do better!

# Relation with full GMRES

The linear systems in Theorem 2 also enable to construct situations where GMRES(m) converges faster than GMRES(m+i) for some $i > 0$:

With GMRES(m) exhibiting the optimal convergence speed of full GMRES, no GMRES(m+i) process can do better!

Example: We construct a linear system with

$$A \in \mathbb{R}^{16\times 16}, \quad b \in \mathbb{R}^{16}$$

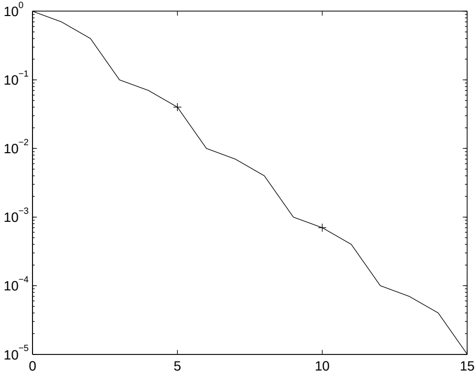
such that GMRES(5) generates the prescribed residual norms

$$\left[\|r_0^{(1)}\|, \|r_1^{(1)}\|, \dots, \|r_5^{(1)}\|\right] = [1, 0.7, 0.4, 0.1, 0.07, 0.04]$$

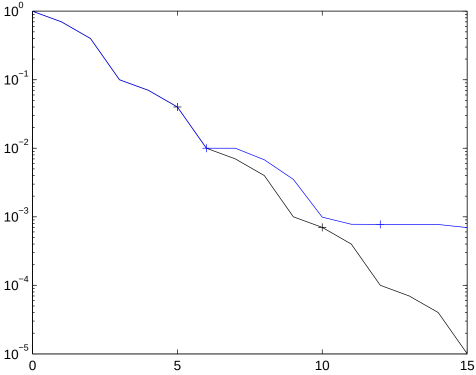
$$\left[\|r_0^{(2)}\|, \|r_1^{(2)}\|, \dots, \|r_5^{(2)}\|\right] = [0.04, 0.01, 0.007, 0.004, 0.001]$$

$$\left[\|r_0^{(3)}\|, \|r_1^{(3)}\|, \dots, \|r_5^{(3)}\|\right] = [0.001, 7 \cdot 10^{-4}, 4 \cdot 10^{-4}, 10^{-4}, 7 \cdot 10^{-5}]$$

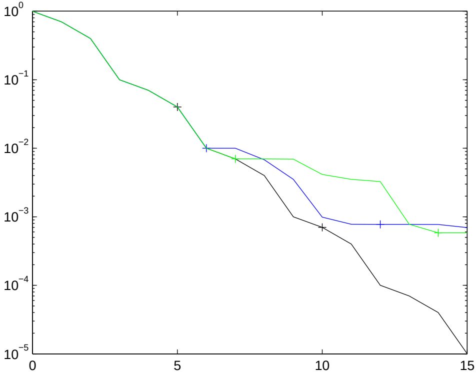
# Example



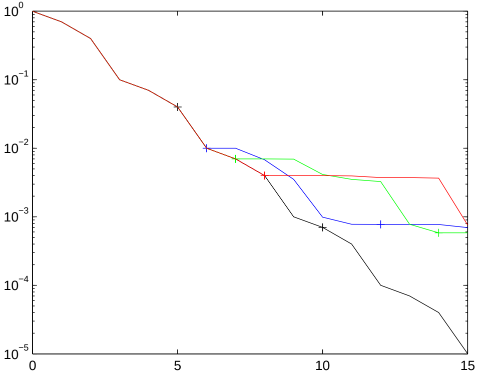
# Example



# Example



# Example



# Outline

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.
- When we prescribe stagnation at the end of a cycle, we must also prescribe it at the beginning of the next cycle. The same holds for zero Ritz values at the end of a cycle.

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.
- When we prescribe stagnation at the end of a cycle, we must also prescribe it at the beginning of the next cycle. The same holds for zero Ritz values at the end of a cycle.
- We constructed short-recurrence but full GMRES processes

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.
- When we prescribe stagnation at the end of a cycle, we must also prescribe it at the beginning of the next cycle. The same holds for zero Ritz values at the end of a cycle.
- We constructed short-recurrence but full GMRES processes
- Questions for future work include:

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.
- When we prescribe stagnation at the end of a cycle, we must also prescribe it at the beginning of the next cycle. The same holds for zero Ritz values at the end of a cycle.
- We constructed short-recurrence but full GMRES processes
- Questions for future work include:
  - Can systems be constructed generating prescribed residual norms for both GMRES(m) and GMRES(m+i)?

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.
- When we prescribe stagnation at the end of a cycle, we must also prescribe it at the beginning of the next cycle. The same holds for zero Ritz values at the end of a cycle.
- We constructed short-recurrence but full GMRES processes
- Questions for future work include:
  - Can systems be constructed generating prescribed residual norms for both GMRES(m) and GMRES(m+i)?
  - What is the relation between our constructions and components of $b$ in the eigenspaces ?

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.
- When we prescribe stagnation at the end of a cycle, we must also prescribe it at the beginning of the next cycle. The same holds for zero Ritz values at the end of a cycle.
- We constructed short-recurrence but full GMRES processes
- Questions for future work include:
  - Can systems be constructed generating prescribed residual norms for both GMRES(m) and GMRES(m+i)?
  - What is the relation between our constructions and components of $b$ in the eigenspaces ?
  - Are there consequences for restarted Arnoldi ?

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.
- When we prescribe stagnation at the end of a cycle, we must also prescribe it at the beginning of the next cycle. The same holds for zero Ritz values at the end of a cycle.
- We constructed short-recurrence but full GMRES processes
- Questions for future work include:
  - Can systems be constructed generating prescribed residual norms for both GMRES(m) and GMRES(m+i)?
  - What is the relation between our constructions and components of $b$ in the eigenspaces ?
  - Are there consequences for restarted Arnoldi ?
  - Can the sufficiency of shorter restart parameters be detected during the GMRES(m) process ?

# Conclusions and future work

- Any prescribed non-stagnating residual norms and nonzero Ritz values are possible for the first $n$ iterations of restarted GMRES, with any spectrum of $A$.
- When we prescribe stagnation at the end of a cycle, we must also prescribe it at the beginning of the next cycle. The same holds for zero Ritz values at the end of a cycle.
- We constructed short-recurrence but full GMRES processes
- Questions for future work include:
  - Can systems be constructed generating prescribed residual norms for both GMRES(m) and GMRES(m+i)?
  - What is the relation between our constructions and components of $b$ in the eigenspaces ?
  - Are there consequences for restarted Arnoldi ?
  - Can the sufficiency of shorter restart parameters be detected during the GMRES(m) process ?
  - What can be said for GMRES(m) after iteration number $n$ ?

**Thank you for your attention!**

# Related papers


- A. Greenbaum and Z. Strakoš, [Matrices that generate the same Krylov residual spaces, IMA Vol. Math. Appl., 60 (1994), pp. 95–118.]
- A. Greenbaum, V. Pták and Z. Strakoš, [Any nonincreasing convergence curve is possible for GMRES, SIMAX, 17 (1996), pp. 465–469.]
- M. Arioli, V. Pták and Z. Strakoš, [Krylov sequences of maximal length and convergence of GMRES, BIT, 38 (1996), pp. 636–643.]
- E. Vecharinsky and J. Langou, [ Any admissible cycle-convergence behavior is possible for restarted GMRES at its initial cycles, 18 (2011), Num. Lin. Alg. Appl., pp. 499?-511. ]
- J. Duintjer Tebbens and G. Meurant, [Any Ritz value behavior is possible for Arnoldi and for GMRES, SIMAX, 33 (2012), pp. 958–978.]
- J. Duintjer Tebbens and G. Meurant, [Prescribing the behavior of early terminating GMRES and Arnoldi iterations, Numer. Algorithms, 65 (2014), pp. 69–90. ]
- J. Duintjer Tebbens, G. Meurant, H. Sadok and Z. Strakoš, [ On Investigating GMRES Convergence Using Unitary Matrices , 450 (2014), Lin. Alg. Appl., pp. 83–107. ]