March 26, 2018

Lecture 6 – Optimal Iterative Methods

Let $\mathbf{A}$ be an $n \times n$ matrix. If our interest is solving a linear system $\mathbf{Ax} = \mathbf{b}$, we assume that $\mathbf{A}$ is non-singular. For eigenvalue computations, $\mathbf{A}$ can be singular.

A GMRES, FOM and Arnoldi's method

Select a non-trivial initial vector $\mathbf{r}_0$.
The kth step Arnoldi's approach leads to the **Arnoldi relation**

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k$$

of order k, where $\mathbf{V}_k = [\,\mathbf{v}_1, \ldots, \mathbf{v}_k\,]$ is an $n \times k$ orthonormal matrix of **Arnoldi vectors** $\mathbf{v}_i$ with $\mathbf{r}_0 = \|\mathbf{r}_0\|_2\,\mathbf{v}_1$ and $\underline{H}_k$ is a $(k+1) \times k$ upper Hessenberg matrix. H_k will denote the $k \times k$ upper block of $\underline{H}_k$. Note that $\mathrm{span}(\mathbf{V}_k) = \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$. An Arnoldi relation is also called **Arnoldi factorisation** or **Arnoldi decomposition**.

Both the **Generalised Minimal Residual** method (**GMRES**) and the **Full Orthogonalisation Method** (**FOM**) for solving the linear system $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ are based on the Arnoldi relation. They both take $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$ and in both methods $\mathbf{x}_k$ is obtained as $\mathbf{x}_k = \mathbf{x}_0 + \mathbf{V}_k\,\vec{y}_k$ for some vector $\vec{y}_k \in \mathbb{C}^k$. In particular, $\mathbf{x}_k - \mathbf{x}_0 = \mathbf{V}_k\,\vec{y}_k \in \mathcal{K}(\mathbf{A}, \mathbf{r}_0)$. The methods differ in the way $\vec{y}_k$ is computed. Note that with an approximation $\mathbf{x}_k$ of this form, we have that

$$\begin{aligned} \|\mathbf{r}_k\|_2 &= \|\mathbf{b} - \mathbf{A}\mathbf{x}_k\|_2 = \|\mathbf{r}_0 - \mathbf{A}\mathbf{V}_k\,\vec{y}_k\|_2 \\ &= \|\mathbf{V}_{k+1}\,(\|\mathbf{r}_0\|_2\,e_1 - \underline{H}_k\,\vec{y}_k)\,\|_2 = \|\,\|\mathbf{r}_0\|_2\,e_1 - \underline{H}_k\,\vec{y}_k\|_2. \end{aligned}$$

The following three properties are equivalent: they characterise GMRES.

1) $\|\mathbf{r}_k\|_2 = \min\{\|\mathbf{r}_0 - \mathbf{A}\widetilde{\mathbf{x}}\|_2 \,\big|\, \widetilde{\mathbf{x}} \in \mathrm{span}(\mathbf{V}_k)\} = \min\{\|\mathbf{r}_0 - \mathbf{A}\widetilde{\mathbf{x}}\|_2 \,\big|\, \widetilde{\mathbf{x}} \in \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)\}$.
2) $\mathbf{r}_k \perp \mathbf{A}\mathbf{V}_k$ or, equivalently, $\mathbf{r}_k \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0) = \mathcal{K}_k(\mathbf{A}, \mathbf{A}\mathbf{r}_0)$.
3) $\vec{y}_k = \mathrm{argmin}\{\|\,\|\mathbf{r}_0\|_2\,e_1 - \underline{H}_k\,\vec{y}\,\|_2 \,\big|\, \vec{y} \in \mathbb{C}^k\}$.

GMRES obtains $\vec{y}_k$ as the least square solution of $\underline{H}_k\,\vec{y}_k = \|\mathbf{r}_0\|_2\,e_1$.
The following two equivalent properties characterise FOM.

1) $\mathbf{r}_k \perp \mathbf{V}_k$ or, equivalently, $\mathbf{r}_k \perp \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.
2) $H_k\,\vec{y}_k = \|\mathbf{r}_0\|_2\,e_1$.

FOM uses the solution $\vec{y}_k$ of $H_k\,\vec{y}_k = \|\mathbf{r}_0\|_2\,e_1$.
For ease of notation, assume $\mathbf{x}_0 = \mathbf{0}$. Then $\mathbf{x} = \mathbf{V}_k\,\vec{y}_k \in \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$. The FOM orthogonality condition is equivalent to requiring that the

$$\mathbf{x}_k \in \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0) \quad \text{satisfies} \quad \mathbf{v}^*\mathbf{A}\mathbf{x}_k = \mathbf{v}^*\mathbf{b} \quad \text{for all} \quad \mathbf{v} \in \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0). \tag{6.1}$$

This is known as a **Galerkin** condition: the search subspace ($\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$ in this case) is equal to the **test subspace**. In GMRES the approximate solution from the search subspace $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$ is tested against $\mathbf{A}\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$ (see 2) of the properties that characterise GMRES). This 'skew' way of testing where search subspace differs from the test subspace leads to a so-called **Petrov–Galerkin** condition.
Note that the residuals $\mathbf{r}_k$ can be expressed as $\mathbf{r}_k = p_k(\mathbf{A})\mathbf{r}_0$ for some **residual polynomial** p_k, i.e., a polynomial p_k of degree k that takes the value 1 at 0: $p_k(0) = 1$. The FOM residual polynomial differs from the GMRES residual polynomial. These polynomials play a role in theoretical discussion on convergence. They are never (explicitly) computed.

In **Arnoldi's method** for solving the eigenvalue problem $\mathbf{Ax} = \lambda\mathbf{x}$, the initial vector $\mathbf{r}_0$ is usually selected randomly (unless eigenvalue approximations are computed as a side product of GMRES. Then $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$). Approximate eigenpairs $(\vartheta_k, \mathbf{u}_k)$ are obtained by solving the eigenvalue problem $H_k\,\vec{y}_k = \vartheta_k\vec{y}_k$, and taking $\mathbf{u}_k = \mathbf{V}_k\vec{y}_k$, or, equivalently, $\mathbf{u}_k \in \mathrm{span}(\mathbf{V}_k)$ such

that the Galerkin condition $\mathbf{A}\mathbf{u}_k - \vartheta_k\,\mathbf{u}_k \perp \mathbf{V}_k$ is satisfied (cf., Exercise 6.4). The eigenvalue ϑ_k of H_k is a so-called **Ritz value** (of order k), the associated vector $\mathbf{u}_k$ is a **Ritz vector**, $(\vartheta_k, \mathbf{u}_k)$ is a **Ritz pair**. Note that this method leads to k Ritz pairs, i.e., neglecting scalings (or, assuming $\|\vec{y}_k\|_2 = 1$), to k approximate eigenpairs in step k. Arnoldi's method extract Ritz pairs from the Krylov search subspace as approximate eigenpairs. The variant, where on $\mathbf{u}_k \in \mathrm{span}(\mathbf{V}_k)$ the Petrov–Galerkin condition $\mathbf{A}\mathbf{u}_k - \vartheta_k\mathbf{u}_k \perp \mathbf{A}\mathbf{V}_k$ is imposed leads to harmonic Ritz pairs (see also Exercise 6.4).

Both Ritz values and harmonic Ritz values provid approximate eigenvalues, which is useful in itself, but, as such they are also of importance in the analysis of the convergence of FOM and GMRES, see the discussion in the paragraph before Exercise 6.8,

If $\underline{H}_k = (h_{ij})$ is unreduced, i.e., $h_{i+1,i} \neq 0$ for all i, then $\underline{H}_k$ has a left kernel vector $\vec{\gamma}_{k+1} \in \mathbb{C}^{k+1}$ that is unique up to scaling (see Exercise 6.1). It is convenient to scale $\vec{\gamma}_{k+1}$ to have first coordinate equal to 1:

$$\vec{\gamma}_{k+1} = (1, \gamma_2, \ldots, \gamma_{k+1})^{\mathrm{T}} \quad \text{such that} \quad \vec{\gamma}_{k+1}^{*}\,\underline{H}_k = \vec{0}_k^{*}. \tag{6.2}$$

Here, $\vec{0}_k$ is the k-vector of zeros. This vector $\vec{\gamma}_{k+1}$ plays an important role in the convergence analysis of GMRES, FOM and Arnoldi as we will learn below.

Exercise 6.1. **Hessenberg matrices and residuals.** Let $\mathbf{H} = (h_{ij})$ be an $n \times n$ unreduced upper Hessenberg matrix. $\underline{H}_k$ is the $k+1$ by k upper left block of $\mathbf{H}$. Let $\vec{\gamma}_{k+1}$ be as in (6.2).
(a) Show that, for each k, there exists a $\vec{\gamma}_k$ with the required properties and that $\vec{\gamma}_{k+1}$ can be obtained by extending $\vec{\gamma}_k$ by one coordinate:

$$\vec{\gamma}_{k+1} = (1, \gamma_2, \ldots, \gamma_k, \gamma_{k+1})^{\mathrm{T}} = (\vec{\gamma}_k^{\mathrm{T}}, \gamma_{k+1})^{\mathrm{T}}.$$

Express $\|\vec{\gamma}_{k+1}\|_2^2$ as an update of $\|\vec{\gamma}_k\|_2^2$.

Let y_k^{GMRES} be the minimal residual solution of $\underline{H}_k y = e_1$ (i.e. $\|e_1 - \underline{H}_k y_k^{\mathrm{GMRES}}\|_2$ is minimal) and let y_k^{FOM} denote the solution of $H_k y = e_1$.
(b) Prove that $e_1 - \underline{H}_k y_k^{\mathrm{GMRES}} \perp \mathcal{R}(\underline{H}_k) = \mathcal{N}(\underline{H}_k{}^{*})^{\perp}$ to show that, for some scalar τ, (see also Exercise 0.4(c))

$$e_1 - \underline{H}_k y_k^{\mathrm{GMRES}} = \tau\vec{\gamma}_{k+1}.$$

Show also that, for some scalar $\widetilde{\tau}$,

$$e_1 - \underline{H}_k y_k^{\mathrm{FOM}} = \widetilde{\tau} e_{k+1}.$$

Prove that

$$\rho_k^{\mathrm{G}} \equiv \|e_1 - \underline{H}_k y_k^{\mathrm{GMRES}}\|_2 = \frac{1}{\|\vec{\gamma}_{k+1}\|_2}, \qquad \rho_k^{\mathrm{F}} \equiv \|e_1 - \underline{H}_k y_k^{\mathrm{FOM}}\|_2 = \frac{1}{|\gamma_{k+1}|}. \tag{6.3}$$

Note that the norms of the GMRES residuals and FOM residuals can be computed without computing the solutions (nor the residuals).
(c) Show that $\rho_k^{\mathrm{G}} \leq \rho_{k-1}^{\mathrm{G}}$, $\rho_k^{\mathrm{G}} < \rho_k^{\mathrm{F}}$ and

$$\rho_k^{\mathrm{F}} = \frac{1}{\sqrt{1-\delta^2}}\,\rho_k^{\mathrm{G}}, \quad \text{with} \quad \delta \equiv \frac{\rho_k^{\mathrm{G}}}{\rho_{k-1}^{\mathrm{G}}}.$$

Explain why near stagnation in GMRES corresponds to a high 'bump' in the convergence history of FOM, while the norm of the FOM and GMRES residuals almost coincide at steps where the norm of the GMRES residual decreases significantly (Hint: conclude that $\rho_k^{\mathrm{F}} \approx \rho_k^{\mathrm{G}}$ if $\rho_k^{\mathrm{G}} \ll \rho_{k-1}^{\mathrm{G}}$, while ρ_k^{F} is huge if $\rho_k^{\mathrm{G}} \approx \rho_{k-1}^{\mathrm{G}}$. If $\rho_k^{\mathrm{G}} = \rho_{k-1}^{\mathrm{G}}$, then $\rho_k^{\mathrm{F}} = \infty$).

Consider the Arnoldi relation $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k$. Assume H_n is unreduced.
(d) Describe the relation for $k = n$. Explain how the above results are applicable to GMRES and FOM for this general matrix $\mathbf{A}$. Deduce Alg. 6.1 (pay attention to the case where $\|\mathbf{r}_0\|_2 \neq 1$).

GMRES (AND FOM)
$\rho_0 = \|\mathbf{b}\|_2$, $\mathbf{v} = \mathbf{b}/\rho_0$,
$\mathbf{V} = [\mathbf{v}]$, $k = 1$, $\underline{H} = [\,]$,
$\rho = \rho_0$, $\vec{\gamma} = 1$
`while` $\rho > tol$ `do`
%% *Update the Arnoldi expansion*
$\mathbf{w} = \mathbf{A}\mathbf{v}$
$[\mathbf{v}, \vec{h}\,] = \mathtt{Orth}(\mathbf{V}, \mathbf{w})$
$\mathbf{V} \leftarrow [\,\mathbf{V}, \mathbf{v}\,]$, $k \leftarrow k+1$
$\underline{H} \leftarrow [\,\underline{H}\,; \vec{0}^*_{k-2}]$, $\underline{H} \leftarrow [\,\underline{H}\,, \vec{h}\,]$
%% *Update $\underline{H}$'s left singular vector*
$\gamma = [\vec{\gamma}; 0]^* \vec{h}$, $\gamma \leftarrow \frac{\gamma}{e_k^* \vec{h}}$, $\vec{\gamma} \leftarrow [\vec{\gamma}\,; -\bar{\gamma}]$.
%% *Compute the residual norm*
$\rho = \frac{\rho_0}{\|\vec{\gamma}\|_2}$ %% *or* $\rho = \frac{\rho_0}{|\gamma|}$ *for FOM*
`end while`
%% *Solve the projected problem*
`Solve` $\underline{H}\,\vec{y} = \rho_0\, e_1$ `for` $\vec{y}$ `in Least Square sense`
%% *or for FOM, solve* $H\vec{y} = \rho_0 e_1$ *for* $\vec{y}$
%% *Lift the pre-solution to* $\mathbb{C}^n$
$\vec{y} \leftarrow (\vec{y}; 0)$, $\mathbf{x} = \mathbf{V}\vec{y}$

ALGORITHM 6.1. GMRES (and FOM) for solving $\mathbf{A}\mathbf{x} = \mathbf{b}$ for $\mathbf{x}$ with residual accuracy *tol*. $\mathbf{A}$ is a general square matrix.
Here, we used MATLAB's notation for defining extensions of matrices: $[\vec{y}; 0]$ indicates that the vector $\vec{y}$ is extended with a 0, $[\,\underline{H}\,; \vec{0}^*_{k-2}\,]$ is $\underline{H}$ extended with a row of length $k-2$ of zeros, $[\underline{H}, \vec{h}]$ extends $\underline{H}$ with a column. Note that k is the number of columns of $\mathbf{V}$. If $k = 2$, then $\vec{0}^*_{k-2}$ is empty. 'Orth' is as defined in ALG. 3.1. H is the square upper block of $\underline{H}$. GMRES solves $\underline{H}\,\vec{y} = \rho_0\, e_1$ in the least square sense. The backslash operator in MATLAB, $\underline{H}\backslash(\rho_0 e_1)$, returns the least square solution.

(e) Prove that the following four statements are equivalent:
1) GMRES stagnates in step k, i.e., $\|\mathbf{r}^{\mathrm{GMRES}}_{k-1}\|_2 = \|\mathbf{r}^{\mathrm{GMRES}}_{k}\|_2$
2) $\gamma_{k+1} = 0$ (the size of the FOM residual is ∞)
3) H_k is singular
4) 0 is a kth order Ritz value.
Does the FOM process break down if the norm of its residual is ∞?

The previous exercise showed that the left kernel vector $\vec{\gamma}_{k+1}$ of the Hessenberg matrix $\underline{H}_k$ is usefull to access the accuracy of the FOM and GMRES processes. The next exercise generalises this result to the case where $\mathbf{r}_0$ is not necessarily a multiple of $\mathbf{v}_1$, a situation that may be useful for restarts. This exercise also suggests an easy way to compute y_k^{GMRES}.

Exercise 6.2. Consider the Arnoldi relation $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k$.
Assume $\underline{H}_k$ is unreduced and let $\vec{\gamma}_{k+1}$ be as in (6.2). Assume $\mathbf{b} \in \mathrm{span}(\mathbf{V}_{k+1})$. Let $\vec{\beta}$ be the vector of coordinates of $\mathbf{b}$ with respect to the Arnoldi vectors, i.e., the columns of $\mathbf{V}_{k+1}$: $\vec{\beta} \equiv \mathbf{V}^*_{k+1}\,\mathbf{b}$ and $\mathbf{b} = \mathbf{V}_{k+1}\,\vec{\beta}$. Note that in the standard approach $\vec{\beta} = \|\mathbf{b}\|_2\, e_1$.

(a) Prove that the solution $\mathbf{x}$ of $\mathbf{A}\mathbf{x} = \mathbf{b}$ belongs to $\mathrm{span}(\mathbf{V}_k)$ if and only if $\vec{\beta} \perp \vec{\gamma}_{k+1}$.

ARNOLDI (AND RITZ PAIRS)
Select a $\mathbf{b} \in \mathbb{C}^n$, $\mathbf{b} \neq \mathbf{0}$.
$\rho_0 = \|\mathbf{b}\|_2$, $\mathbf{v} = \mathbf{b}/\rho_0$
$\mathbf{V} = [\mathbf{v}]$, $k = 1$, $\underline{H} = [\,]$
$\rho = \infty$, $\vec{\gamma} = 1$
while $\rho > tol$ do
  %% *Update the Arnoldi expansion*
  $\mathbf{w} = \mathbf{A}\mathbf{v}$
  $[\mathbf{v}, \vec{h}\,] = \mathtt{Orth}(\mathbf{V}, \mathbf{w})$
  $\mathbf{V} \leftarrow [\,\mathbf{V}, \mathbf{v}\,]$, $k \leftarrow k + 1$
  $\underline{H} \leftarrow [\,\underline{H}\,; \vec{0}^*_{k-2}]$, $\underline{H} \leftarrow [\,\underline{H}\,, \vec{h}\,]$
  %% *Solve the projected problem*
  Solve $H\,\vec{y} = \vartheta\,\vec{y}$ with $\|\vec{y}\|_2 = 1$
  Select a pre Ritz pair, say $(\vartheta, \vec{y})$
  %% *Compute the residual norm*
  $\rho = |(e_k^*\vec{h})\,(e_{k-1}^*\vec{y})|$
end while
  %% *Lift the pre-solution to* $\mathbb{C}^n$
$\lambda = \vartheta$, $\vec{y} \leftarrow (\vec{y}; 0)$, $\mathbf{x} = \mathbf{V}\vec{y}$

ALGORITHM 6.2. Arnoldi's method, using Ritz pairs for computing an eigenpair $(\lambda, \mathbf{x})$ with residual accuracy *tol* of a general square matrix $\mathbf{A}$. For notations, see the caption of ALG. 6.1.

(b) Put

$$\vec{\beta}_{k+1} \equiv \vec{\beta} - \vec{\rho}_{k+1} \quad \text{with} \quad \vec{\rho}_{k+1} \equiv \frac{\vec{\gamma}^*_{k+1}\vec{\beta}}{\vec{\gamma}^*_{k+1}\vec{\gamma}_{k+1}}\vec{\gamma}_{k+1}.$$

Show that for an $\mathbf{x}_k = \mathbf{V}_k\vec{y}_k \in \mathrm{span}(\mathbf{V}_k)$ the following statements are equivalent:

1) $\mathbf{x}_k$ is the GMRES solution, i.e., $\mathbf{x}_k = \mathrm{argmin}\{\|\mathbf{b} - \mathbf{A}\widetilde{\mathbf{x}}\|_2 \,\big|\, \widetilde{\mathbf{x}} \in \mathrm{span}(\mathbf{V}_k)\}$,

2) $\vec{y}_k$ exactly satisfies $\underline{H}_k\,\vec{y}_k = \vec{\beta}_{k+1}$,

3) $\mathbf{r}_k \equiv \mathbf{b} - \mathbf{A}\mathbf{x}_k = \mathbf{V}_{k+1}\,\vec{\rho}_{k+1}$.

(c) Let $\mathbf{x}_k = \mathbf{V}_k\vec{y}_k \in \mathrm{span}(\mathbf{V}_k)$ be the GMRES solution. Show that $\vec{y}_k$ is the solution of the upper triangular $k \times k$ system $R\vec{y}_k = \vec{\beta}'$, where R is the lower $k \times k$ block of $\underline{H}_k$ and $\vec{\beta}'$ is lower k-vector of $\vec{\beta}_{k+1}$.
Note that $\|\mathbf{r}_k\|_2 = |\vec{\gamma}^*_{k+1}\vec{\beta}|/\|\vec{\gamma}_{k+1}\|_2$ which is $\|\mathbf{b}\|_2$ times the cosine of the angle between $\vec{\beta}$ and $\vec{\gamma}_{k+1}$. Observe that this is in line with the fact that $\|\mathbf{r}_k\|_2$ is $\|\mathbf{b}\|_2$ times the sine of the angle between $\mathbf{b}$ and span $(\mathbf{A}\mathbf{V}_k)$ (why do we have this fact?).

Exercise 6.3. **Unreduced Hessenberg matrices and GMRES.** Suppose Arnoldi's process in the GMRES (FOM) or Arnoldi context yields $h_{i+1,i} = 0$ for some i. Here, $\underline{H}_k = (h_{ij})$ is the Hessenberg matrix for this process. Discuss the consequences for GMRES (and FOM).

Exercise 6.4. **Arnoldi, Ritz values and harmonic Ritz values.** Consider the Arnoldi relation $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k$. Assume $\underline{H}_k$ is unreduced and let $\vec{\gamma}_{k+1}$ be as in (6.2). Put $\eta \equiv h_{k+1,k}$. The n-vector $\mathbf{u}$ and k-vector $\vec{y}$ are related by $\mathbf{u} = \mathbf{V}_k\,\vec{y}$, ϑ is a scalar, $\vec{y}$ is normalised: $\|\vec{y}\,\|_2 = 1$. $\mathbf{r}$ is the **residual**: $\mathbf{r} \equiv \vartheta\,\mathbf{u} - \mathbf{A}\mathbf{u}$.

(a) Show that $\|\mathbf{u}\|_2 = 1$.

(b) Prove that the following two properties are equivalent

1) $\vartheta\,\mathbf{u} - \mathbf{A}\mathbf{u} \perp \mathbf{V}_k$.

2) $H_k\,\vec{y} = \vartheta\,\vec{y}$, where H_k is the $k \times k$ upper block of $\underline{H}_k$.

Then ϑ is an **Ritz value** with **Ritz vector u** and **pre Ritz vector** $\vec{y}$ (of order k).

(c) Show that, for Ritz pairs (i.e., one of the properties of (b) holds), we have that

$$\|\mathbf{r}\|_2 = |\eta|\,|e_k^*\vec{y}\,| :$$

the residual norm can be computed in k-dimensional space. Actually, only the last coordinate of (the normalised) $\vec{y}$ is needed and the right bottom element of $\underline{H}_k$.

(d) Deduce ALG. 6.2. Note that the problem $H_k\,\vec{y} = \vartheta\,\vec{y}$ is a low dimensional one. It can efficiently be solved with the QR-algorithm (MATLAB's `eig`) yielding k pre Ritz pairs $(\vartheta, \vec{y})$. If the eigenvalue λ of $\mathbf{A}$ with, say, smallest real part[1] is the one that is of interest, then, among all pre Ritz pairs of order k, the ϑ with smallest real part will be an appropriate choice (an appropriate selection strategy).

(e) Show that, for Ritz pairs, we also have that

$$\|\mathbf{r}\|_2 = |\vartheta|\,\frac{|\vec{\gamma}_{k+1}^*\vec{y}|}{|\gamma_{k+1}|} = |\vartheta|\,\cos\angle(\vec{\gamma}_{k+1}, \vec{y})\,\frac{\|\vec{\gamma}_{k+1}\|_2}{|\gamma_{k+1}|}.$$

Note that the last fraction is the ratio $\rho_k^{\mathrm{F}}/\rho_k^{\mathrm{G}}$ of the kth FOM residual norm ρ_k^{F} and the GMRES residual norm ρ_k^{G} as discussed in Exercise 6.1, see (6.3).

(f) Prove that the following five properties are equivalent (we assume $\vec{y}$ to be extended with one zero if dimensions have to be matched).

1) $\vartheta\,\mathbf{u} - \mathbf{A}\mathbf{u} \perp \mathbf{A}\mathbf{V}_k$.

2) $\underline{H}_k\,\vec{y} - \vartheta\,\vec{y} \perp \underline{H}_k$.

3) $\underline{H}_k\,\vec{y} - \vartheta\,\vec{y} = \widetilde{\beta}\,\vec{\gamma}_{k+1}$ for some scalar $\widetilde{\beta}$.

4) $(\underline{H}_k - \beta\,\vec{\gamma}_{k+1}e_k^*)\vec{y} = \vartheta\,\vec{y}$, where $\beta \equiv \eta/\gamma_{k+1}$.

5) $(I_k + \alpha\,\vec{\gamma}_k\vec{\gamma}_k^*)H_k\vec{y}_k = \vartheta\vec{y}$, where $\alpha \equiv |\gamma_{k+1}|^{-2}$.

Note that the bottom row of the matrix $\underline{H}_k - \beta\,\vec{\gamma}_{k+1}e_k^*$ in 4) is a zero vector: this matrix essentially is square, i.e., $k \times k$. To prove 5), note that $\vec{0}^* = \vec{\gamma}_{k+1}^*\,\underline{H}_k = \vec{\gamma}_k^*\,H_k + \eta\,\bar{\gamma}_{k+1}\,e_k^*$.

Then ϑ is called an **harmonic Ritz value** with **harmonic Ritz vector u** and **pre harmonic Ritz vector**: $\vec{y}$.

(g) Show that, for harmonic Ritz pairs, we have

$$\|\mathbf{r}\|_2 = |\eta|\,|e_k^*\vec{y}\,|\frac{\|\vec{\gamma}_{k+1}\|_2}{|\gamma_{k+1}|}.$$

(h) Show that, for harmonic Ritz pairs, we also have

$$\|\mathbf{r}\|_2 = |\vartheta|\,\frac{|\vec{\gamma}_{k+1}^*\vec{y}\,|}{\|\vec{\gamma}_{k+1}\|_2} = |\vartheta|\,\cos\angle(\vec{\gamma}_{k+1}, \vec{y}).$$

(i) Let ϑ be the harmonic Ritz value of (f). Deduce from property 1) that $\vartheta = \frac{\|\mathbf{A}\mathbf{u}\|_2^2}{\mathbf{u}^*\mathbf{A}^*\mathbf{u}}$. Take $\mathbf{y} \equiv \mathbf{A}\mathbf{u}$ and conclude that

$$\frac{1}{\vartheta} = \frac{\mathbf{y}^*\mathbf{A}^{-1}\mathbf{y}}{\mathbf{y}^*\mathbf{y}} \le \|\mathbf{A}^{-1}\|_2 \tag{6.4}$$

in case $\mathbf{A}$ is non-singular: all harmonic Ritz values 'stay away from zero' if $\mathbf{A}$ is non-singular.

For GMRES and harmonic Ritz vectors, projecting the problems onto Krylov subspaces leads to non square systems. The following exercise shows that the projected problems can also be formulated as square systems.

[1] $\mathrm{Re}(\lambda) \le \mathrm{Re}(\lambda')$ for all eigenvalues λ' of $\mathbf{A}$

Exercise 6.5. Let $\underline{H}_k = (h_{i,j})$ be an unreduced $(k+1)\times k$ upper Hessenberg matrix and let $\vec{\gamma}_{k+1}$ be as in (6.2). Put $\vec{\gamma} \equiv \vec{\gamma}_k/\gamma_{k+1}$. As before (cf., Exercise 6.4), H_k is the $k\times k$ upper block of $\underline{H}_k$, $\eta \equiv h_{k+1,k}$, and $\beta \equiv \eta/\gamma_{k+1}$

(a) Show that the bottom row of $\underline{H}_k - \beta\,\vec{\gamma}_{k+1}\,e_k^*$ consists of zeros only. Note that this matrix arises by subtracting a multiple of $\vec{\gamma}_{k+1}$ from the last column of $\underline{H}_k$, with multiple selected such that the right-bottom element of this 'adapted' matrix equals 0.

(b) Show that the $k\times k$ upper block of $\underline{H}_k - \beta\vec{\gamma}_{k+1}e_k^*$ equals $(I_k + \vec{\gamma}\,\vec{\gamma}^*)H_k$.

(c) Prove the following equivalences for $\vec{y}\in\mathbb{C}^k$, $\vartheta\in\mathbb{C}$,

$$(I_k + \vec{\gamma}\,\vec{\gamma}^*)H_k\,\vec{y} = e_1 \qquad \Leftrightarrow \qquad \underline{H}_k\,\vec{y} - e_1 \perp \underline{H}_k,$$

$$(I_k + \vec{\gamma}\,\vec{\gamma}^*)H_k\,\vec{y} = \vartheta\vec{y} \qquad \Leftrightarrow \qquad \underline{H}_k\,\vec{y} - \vartheta\vec{y} \perp \underline{H}_k$$

(extend $\vec{y}$ with a 0 if required for matching dimensions).

(d) Let $\alpha > 0$ be such that $\alpha^2\|\vec{\gamma}\|_2^2 + 2\alpha = 1$. Prove that $(I_k + \alpha\,\vec{\gamma}\,\vec{\gamma}^*)^2 = I_k + \vec{\gamma}\,\vec{\gamma}^*$ and conclude that ϑ is an eigenvalue of $(I_k + \vec{\gamma}\,\vec{\gamma}^*)H_k$ if and only if it is an eigenvalue of

$$(I_k + \alpha\,\vec{\gamma}\,\vec{\gamma}^*)H_k(I_k + \alpha\,\vec{\gamma}\,\vec{\gamma}^*).$$

Note that this matrix is Hermitean if H_k is Hermitean (which is the case if $\mathbf{A}$ is Hermitean).

In Lecture 13.A, we will learn that harmonic Ritz values are usefull objects in eigenvalue computations. Here, in this Lecture 6, they have been introduced since they relate to GMRES, as Ritz values relate to FOM.

The kth FOM residual $\mathbf{r}_k^{\mathrm{FOM}}$ equals $p_k(\mathbf{A})\mathbf{r}_0$ for some residual polynomial p_k of degree k: this polynomial is called the kth FOM residual polymial or kth **FOM polynomial**. The following theorem characterises this polynomial in terms of the kth order Ritz values: the Ritz values are the zeros of the FOM polynomial. In Exercise 6.8, we will see that the harmonic Ritz values are the zeros of the GMRES polynomial.

Theorem 6.1 *Let p be a polynomial of degree k such that $p(0) = 1$.*
The following two properties are equivalent.

1) *p is the kth FOM polynomial* *($p(\mathbf{A})\mathbf{v}_1 \perp \mathbf{V}_k$).*

2) *$p^{(j)}(\vartheta) = 0$ for all eigenvalues ϑ of H_k and all $j < \mu(\vartheta)$* *($p(H_k) = 0$)*

Here $\mu(\vartheta)$ is the multiplicity of the eigenvalue ϑ of H_k.

The next exercise suggests a proof of this theorem. Exercise 6.7 provides an easier proof for the simpler situation where all eigenvalues of H_k are simple.

Exercise 6.6. **Ritz values and zeros of FOM polynomials (Proof of Th. 6.1).**

(a) Show that $\mathbf{r}_k \perp \mathbf{V}_k$: $\mathbf{r}_k$ is the kth FOM residual. Show that there are a polynomial q of degree $k-1$ and a polynomial p_k of degree k such that $p_k(\zeta) = 1 - \zeta q(\zeta)$ $(\zeta\in\mathbb{C})$ and $\mathbf{r}_k = (\mathbf{I} - \mathbf{A}q(\mathbf{A}))\mathbf{r}_0 = p_k(\mathbf{A})\mathbf{r}_0 \perp \mathbf{V}_k$: p_k is the kth **FOM polynomial**.

(b) Show that, $\mathbf{A}\mathbf{V}_k e_j = \mathbf{V}_k H_k e_j$ for all $j < k$, $\mathbf{A}^j\mathbf{V}_k e_1 = \mathbf{V}_k H_k^j e_1$ for all $j < k$,
and $\mathbf{A}^k\mathbf{V}_k e_1 = \mathbf{V}_{k+1}\,\underline{H}_k\,H_k^{k-1}\,e_1$, $\mathbf{V}_k^*\mathbf{A}^j\mathbf{V}_k\,e_1 = H_k^j\,e_1$ for all $j \le k$.
The problem of finding a linear operator $\mathbf{A}_k$ on $\mathcal{K}_k(\mathbf{A},\mathbf{v}_1)$ for which

$$\mathbf{A}_k^j\mathbf{v}_1 = \mathbf{A}^j\mathbf{v}_1 \quad (j<k), \quad \text{and} \quad \mathbf{A}_k^k\mathbf{v}_1 = \mathbf{V}_k\mathbf{V}_k^*\mathbf{A}^k\mathbf{v}_1$$

is known as the **Vorobyev moment problem for Arnoldi**.
Solve this problem (in terms if H_k and $\mathbf{V}_k$).

(c) Show that $\mathbf{V}_k^* p(\mathbf{A})\mathbf{V}_k\,e_1 = p(H_k)e_1$ for any polynomial p of degree k.

(d) Prove Theorem 6.1. (Hint. Use also Theorem 5.9.)

Exercise 6.7. **An alternative proof of Theorem 6.1.** To avoid technical details, we assume that $\dim(\mathcal{K}_{k+1}(\mathbf{A},\mathbf{r}_0)) = k+1$ and that all eigenvalues $\vartheta_1,\ldots,\vartheta_k$ of H_k are simple.

Let p_k be the FOM residual polynomial in step k, i.e., $p_k(0) = 1$, p_k is of degree k and $\mathbf{r}_k = \mathbf{r}_k^{\mathrm{FOM}} = p_k(\mathbf{A})\mathbf{r}_0$.

(a) If $\vartheta = \vartheta_j$ is a zero of p_k, then $p_k(\lambda) = (\lambda - \vartheta)q(\lambda)$ for some polynomial q of degree $k-1$.

Put $\mathbf{u} \equiv q(\mathbf{A})\mathbf{r}_0$.

(b) Prove that $\mathbf{A}\mathbf{u} - \vartheta\mathbf{u} = \mathbf{r}_k \perp \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.
Conclude that $(\vartheta, \mathbf{u})$ is a Ritz pair (for $\mathbf{A}$, w.r.t., $\mathcal{K}_{k+1}(\mathbf{A}, \mathbf{r}_0)$).

(c) Use a dimension argument to show that this approach leads to all Ritz pairs (neglecting scalings of the Ritz vectors). In particular, if $(\vartheta, \mathbf{u})$ is a Ritz pair, then ϑ is a zero of p_k and

$$p_k(\lambda) = \prod_{j=1}^{k} \left(1 - \frac{1}{\vartheta_j}\,\lambda\right) \qquad (\lambda \in \mathbb{C}). \tag{6.5}$$

Note that this polynomial does not exists if a ϑ_j is equal to zero: at such a point, the FOM process cannot produce an approximate solution $\mathbf{x}_k^{\mathrm{FOM}}$ or residual $\mathbf{r}_k^{\mathrm{FOM}}$. However, the Arnoldi process can continue and in the subsequent steps FOM may be able to produce approximations and resiuduals. The convergence curve of the FOM process (the curve of $\|\mathbf{r}_k\|_2$) exhibits a (huge) peak in step k if a Ritz value in this step is (very) near zero.

The formula $p_k(\lambda) = (\lambda - \vartheta_j)q(\lambda)$ in Exercise 6.7(a) is useful for several reasons:
1) As we learnt form the above exercise, it shows that zeros of the FOM polynomial are Ritz values. And, conversely, that Ritz values are zeros of the FOM polynomial. For GMRES, we have a similar result for harmonic Ritz values rather than Ritz values (see Exercise 6.8).
2) If a Ritz value ϑ_j is close to an eigenvalue, then $\mathbf{r}_k = (\mathbf{A} - \vartheta_j\mathbf{I})\mathbf{u}$ (see Exercise 6.7(b)) shows that the component of the associated eigenvector is deflated from the residual (cf., Exercise 4.3(a)). This implies that from this step k on this eigenvector does not play a role anymore in the convergence: the spectrum that effectively determines convergence is the original spectrum from which the 'detected' eigenvalue has been removed. Here, we also used the fact that once an eigenvector is (almost) in the Krylov subspace, this eigenvector will also be (almost) in the Krylov subspaces of higher order. This explains the **super linear convergence** of FOM (and of CG, the positive definite version of FOM). Similar arguments (using harmonic Ritz values) can be used to explain the super linear convergence of other optimal Krylov methods as GMRES, GCR and CR.
3) Repeated factorisation and using the fact that p_k is a residual polynomial, i.e., $p_k(0) = 1$, leads to (6.5). This formula shows that $p_k(\lambda)$, and therefore $\mathbf{r}_k = p_k(\mathbf{A})\mathbf{r}_0$ will be huge if a Ritz value ϑ_j is almost zero. Then we have a peak in the convergence history of FOM. Harmonic Ritz values will not be close to zero: in absolute value, they are at least the smallest singular value of $\mathbf{A}$ (see (6.4)). This is in line with the fact that GMRES converges monotonically.

Exercise 6.8. **Harmonic Ritz values and zeros of GMRES polynomials.** Let $\mathbf{r}_k$ be the GMRES residual: $\mathbf{r}_k \perp \mathbf{A}\mathbf{V}_k$. There is a polynomial p_k of exact degree k such that $p_k(\mathbf{A})\mathbf{r}_0 = \mathbf{r}_k$ and $p_k(0) = 1$: p_k is the kth **GMRES polynomial**.

(a) Adapt the arguments in Exercise 6.7 to prove that harmonic Ritz values are the zeros of the GMRES polynomial.

(b) Prove that the property in (a) characterises GMRES polynomial (as is the case for FOM polynomials in Theorem 6.1).

Exercise 6.9. **Stability of Arnoldi and GMRES.** Consider an $n \times \ell$ matrix $\mathbf{X}$, with QR-decomposition (in economical form) $\mathbf{X} = \mathbf{Q}R$, with $\mathbf{Q}$ an $n \times \ell$ orthonormal matrix and R an $\ell \times \ell$ upper triangular matrix.

Let $\mathbf{A}$ be $n \times n$, with Arnoldi relation

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k.$$

(a) Prove that $\mathbf{V}_{k+1}$ is the 'Q-factor', in the QR-decomposition of

$$\mathbf{X} \equiv [\mathbf{v}_1, \mathbf{A}\mathbf{v}_1, \mathbf{A}\mathbf{v}_2, \ldots, \mathbf{A}\mathbf{v}_k] = [\mathbf{v}_1, \mathbf{A}\mathbf{V}_k].$$

Relate the 'R-factor' to $\mathbf{H}_k$ (cf., (e) of Exercise 3.19).

The computed version $\widehat{\mathbf{Q}}$ of $\mathbf{Q}$ will not be exactly orthonormal. The '**loss of orthogonality**'

$$\nu \equiv \|\widehat{\mathbf{Q}}^* \widehat{\mathbf{Q}} - I\|_2,$$

depends on the orthogonalisation strategy. In floating point arithmetic, with relative machine precision $\mathbf{u}$, we have (you do not have to prove this, see the discussion in Section E of Lecture 3)

- classical Gram–Schmidt: $\nu \leq ck^2\mathbf{u}\mathcal{C}_2^2(\mathbf{X})$;
- modified Gram–Schmidt: $\nu \leq ck^2\mathbf{u}\mathcal{C}_2(\mathbf{X})$;
- repeated Gram–Schmidt: depends on repetition strategy;
- with Householder: $\nu \leq ck^2\mathbf{u}$.

Here c is some modest constant (as $c = 4$).

(b) Prove that the loss of orthogonality in GMRES with modified Gram–Schmidt is approximately

$$\nu \approx \frac{2\mathbf{u}}{\|\mathbf{r}_k^{\mathrm{GMRES}}\|_2}.$$

(Hint: see (c) of Exercise 3.8).

(c) What orthogonalisation strategy should we use in GMRES?

(d) What orthogonalisation strategy should we use in Arnoldi (for computing several eigenpairs)?

B (Lack of) Convergence of GMRES

The following theorem gives an upper bound on the norm of the GMRES residual in terms the spectrum of $\mathbf{A}$ in case $\mathbf{A}$ is diagonizable. The bound contains the condition number $\mathcal{C}_E$ of the eigenvector basis. A similar result holds for other methods that compute residuals that are minimal with respect to some norm, methods as MINRES, SYMMLQ and CG, to be discussed in the next lecture. Moreover, since GMRES finds approximate solutions in the Krylov subspace with smallest residual 2-norm, no other Krylov subspace method (including the ones the be discussed in the next two lectures) will exhibit fast convergence for a specific problem $\mathbf{A}\mathbf{x} = \mathbf{b}$ if GMRES fails to converge quickly (in terms of the number of Matrix-Vector multiplications) for this problem.

Theorem 6.2 *If* $\mathbf{A}$ *is diagonizable and* $\mathbf{r}_k = \mathbf{r}_k^{\mathrm{GMRES}}$ *is the* k*th GMRES residual then*

$$\|\mathbf{r}_k\|_2 \leq \mathcal{C}_E\, \nu_k(\Lambda(\mathbf{A}))\, \|\mathbf{r}_0\|_2, \quad \textit{where} \quad \nu_k(\mathcal{G}) \equiv \min_{p \in \mathcal{P}_k^0} \max\{|p(\zeta)| \,\big|\, \zeta \in \mathcal{G}\} \quad (\mathcal{G} \subset \mathbb{C}), \tag{6.6}$$

and $\mathcal{C}_E$ *is the (smallest) condition number a a basis of eigenvectors.*[2]

Exercise 6.10. Prove Theorem 6.2.

MINRES is the Hermitian variant of GMRES, SYMMLQ computes approximation with minimal errors for Hermitian matrices, and CG minimises the error in the $\mathbf{A}$-norm if $\mathbf{A}$ is positive definite. These methods have been designed for Hermitian matrices where $\mathcal{C}_E = 1$. For these methods the distribution of the eigenvalues gives excellent information on the convergence:

[2] $\mathcal{C}_E \equiv \|\mathbf{X}\|_2\, \|\mathbf{X}^{-1}\|_2$ with $\mathbf{A}\mathbf{X} = \mathbf{X}\Lambda$ and $\mathcal{C}_E$ smallest. Note that the value of $\mathcal{C}_E$ is affected by the scaling of the eigenvectors. Moreover, in case of a semi-simple eigenvalue with multiplicity > 1, the associated eigenvectors could be selected to be skew.

the size of the kth residual can be bounded by the size of the residual polynomial of degree k on the spectrum of the matrix.[3]

In the non-normal case, the conditioning of the eigenvectors may be very large. If the matrix is not diagonizable (i.e., if the Jordan form contains non-trivial Jordan blocks), then the estimate is even not applicable. The following exercise gives an alternative that also covers this case.

Exercise 6.11. **Cauchy's integral formula.** Let f be a complex-valued function that is analytic on some simply connected[4] open subset $\mathcal{G}$. Let Γ be a closed smooth curve (a contour) in $\mathcal{G}$ that encircles λ once counterclockwise. From **Cauchy's integral formula** in complex function theory we know that

$$f(\lambda) = \frac{1}{2\pi i} \oint_\Gamma \frac{f(\zeta)}{\zeta - \lambda} \, \mathrm{d}\zeta.$$

This representation of f in terms of a **contour integral** allows us to turn f into a matrix-valued function for $n \times n$ matrices with spectrum in the area in $\mathcal{G}$ with boundary Γ:

$$f(\mathbf{A}) = \frac{1}{2\pi i} \oint_\Gamma (\zeta \mathbf{I} - \mathbf{A})^{-1} f(\zeta) \, \mathrm{d}\zeta \tag{6.7}$$

Note: it is essential that *all* eigenvalues of $\mathbf{A}$ are the same side of Γ.
The function $\zeta \rightsquigarrow (\zeta \mathbf{I} - \mathbf{A})^{-1}$ is the **resolvent** of $\mathbf{A}$.

(a) For polynomials p, we already have definition of $p(\mathbf{A})$. Prove that this definition is consistent with the one by (6.7) in case $\mathbf{A}$ is diagonizable. (Hint: diagonalise $\mathbf{A}$ and apply the counter integral expression matrix entry wise.)

(b) Prove that

$$\|\mathbf{r}_k^{\mathrm{GMRES}}\|_2 \leq \ell(\Gamma) \max_{\zeta \in \Gamma} \|(\zeta \mathbf{I} - \mathbf{A})^{-1}\|_2 \; \nu_k(\Gamma) \, \|\mathbf{r}_0\|_2, \tag{6.8}$$

where $\ell(\Gamma)$ is the length of the curve Γ.

Formula (6.7) is often useful, but leads to statements as (6.8) with the same disadvantage as (6.6): they do not only involve eigenvalue information (and the size of $\mathbf{r}_0$) but also quantities that are hard to access and that are extremely large for some non-normal matrices. These quantities can completely dominate the bound. Then, these formulas do not allow to extract any information at all on the convergence from the distribution of the eigenvalue. The following theorem shows that that is not just because of an inaccurate bound. The GMRES residuals decrease monotonically, $\|\mathbf{r}_{k+1}^{\mathrm{GMTRES}}\|_2 \leq \|\mathbf{r}_k^{\mathrm{GMTRES}}\|_2$ for all k, and $\|\mathbf{r}_n^{\mathrm{GMTRES}}\|_2 = 0$, but, except for these restrictions, any convergence curve is possible for any distribution of eigenvalues.

The matrices in this theorem are **far from normal**, that is, $\|\mathbf{A}\mathbf{A}^* - \mathbf{A}^*\mathbf{A}\|_2$ is large (relative to $\|\mathbf{A}\|_2^2$), while that is usually not the case for matrices that we encounter in practice. In practice, it appears that the distribution of the eigenvalues usually gives good information on the convergence of methods as GMRES.

Theorem 6.3 *Let $(\rho_0, \rho_1, \rho_2, \ldots, \rho_{n-1})$ be a decreasing sequence of n positive real numbers, that is $\rho_0 \geq \rho_1 \geq \ldots \geq \rho_{n-1} \geq \rho_n = 0$. For an $n \times n$ matrix $\mathbf{A}$ consider the following statement:*

$$\begin{aligned} &\textit{there is a right-hand side vector } \mathbf{b} \textit{ such that, with initial guess } \mathbf{x}_0 = \mathbf{0}, \\ &\textit{for the GMRES residuals } \mathbf{r}_k \textit{ we have that } \|\mathbf{r}_k\|_2 = \rho_k \quad (k = 0, \ldots, n). \end{aligned} \tag{6.9}$$

1) *For any sequence $\lambda_1, \ldots, \lambda_n$ of n non-trivial complex numbers, there is a matrix $\mathbf{A}$ such that the λ_j are the eigenvalues of $\mathbf{A}$ counted according to multiplicity and* (6.9) *holds.*

[3] By $\nu_k \equiv \nu_k(\Lambda(\mathbf{A}))$. The size of the kth residual also depends on the initial residual $\mathbf{r}_0$, but this dependence is usually rather weak, i.e., the estimate $\|\mathbf{r}_k\| \leq \nu_k \|\mathbf{r}_0\|$ tends to be sharp.

[4] between any pair of points in $\mathcal{G}$ there is a smooth curve that connects these points, and any closed curve in $\mathcal{G}$ can be continuously contracted in $\mathcal{G}$ to a point: $\mathcal{G}$ does not have holes.

Before we discuss the proof, let us consider two simple examples: the matrix in the proof of the theorem is essentially a modification of a combination of these two examples.

Let $\mathbf{C}$ be the circular matrix defined by $\mathbf{C}\mathbf{e}_k = \mathbf{e}_{k+1}$, $\mathbf{C}\mathbf{e}_n = \mathbf{e}_1$ and $\mathbf{S}$ is the shift matrix: by $\mathbf{S}\mathbf{e}_k = \mathbf{e}_{k+1}$, $\mathbf{S}\mathbf{e}_n = \mathbf{0}$, with, for both matrices $k = 1, \ldots, n-1$. We take $\mathbf{b} = \mathbf{e}_1$. We consider two cases: a) $\mathbf{A} \equiv \mathbf{C}$ with $\mathbf{x} = \mathbf{e}_1$ and b) $\mathbf{A} \equiv \mathbf{I} - \mathbf{S}$ with $\mathbf{x} = \mathbf{1}$, the n-vector of all ones. Note that in both cases, with $\mathbf{x}_0 \equiv \mathbf{0}$ we have that $\mathbf{b} = \mathbf{r}_0 = \mathbf{e}_1$ and $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$ is spanned by $\mathbf{e}_1, \ldots, \mathbf{e}_k$. Therefore, in both cases $\|\mathbf{x} - \widetilde{\mathbf{x}}_k\|_2 \geq 1$ for all $\widetilde{\mathbf{x}}_k \in \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$. Actually, in case a) we have stagnation ($\|\mathbf{r}_k^{\mathrm{GMRES}}\|_2 = 1$) until step n ($\|\mathbf{r}_n^{\mathrm{GMRES}}\|_2 = 0$), while in case b) we have a slow decrease ($\|\mathbf{r}_k^{\mathrm{GMRES}}\|_2 = 1/\sqrt{k}$. Use that $\vec{\gamma}_{k+1} \equiv \vec{1}$ is the left kernel vector of the $(k+1) \times k$ left upper block $\underline{H}_k$ of $\mathbf{A}$). In example a), the eigenvectors form a well-conditioned basis (othonormal), but the eigenvalues cluster around 0 (the are uniformly distributed on the unit circle). In example b), the eigenvalues cluster away from 0 (are all equal to 1), but there is no basis of eigenvectors (you could state that $\mathcal{C}_E = \infty$).

Exercise 6.12. ***Proof of Theorem*** 6.3***.*** Let (λ_j) and (ρ_j) be as in Theorem 6.3. Eigenvalues are counted according to multiplicity.

(a) Show that there is a sequence $\gamma_1, \ldots, \gamma_{k-1}$ of scalars such that, with $\vec{\gamma}_k \equiv (1, \gamma_1, \ldots, \gamma_k)^{\mathrm{T}}$ all $k < n$ we have that $\rho_k = \frac{1}{\|\vec{\gamma}_k\|_2}$. Note that $\rho_{k+1} = \rho_k \quad \Leftrightarrow \quad \gamma_{k+1} = 0$.

The matrix $\mathbf{A} = \mathbf{H}$ that we will construct here will be Hessenberg and the right-hand side vector $\mathbf{b}$ will be the first standard basis vector $\mathbf{e}_1$. Moreover, with $\underline{H}_k$ the left upper $(k+1) \times k$ block of $\mathbf{H}$, the constructed $\mathbf{H}$ will be such that

$$\mathbf{H} \text{ is unreduced, } \vec{\gamma}_k^* \underline{H}_k = \vec{0}_k^* \text{ for all } k < n \text{ and } \lambda_1, \ldots, \lambda_n \text{ are the eigenvalues of } \mathbf{H}. \tag{6.10}$$

(b) Use Exercise 6.1 to show that the existence of an Hessenberg $\mathbf{H}$ satisfying (6.10) proves Theorem 6.3.

(c) First, consider the case where the sequence of (ρ_k) is strictly decreasing.
Let $\mathbf{S}$ be the $n \times n$ shift matrix: $\mathbf{S}\mathbf{e}_k = \mathbf{e}_{k+1}$ $(k = 1, \ldots, n-1)$. Show that (6.10) holds for

$$\mathbf{H} \equiv \Gamma^{-1}(\mathbf{I} - \mathbf{S})\Lambda\Gamma, \quad \text{where} \quad \Gamma \equiv \mathrm{diag}\,(1, \gamma_1, \ldots, \gamma_{n-1}), \quad \Lambda \equiv \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n). \tag{6.11}$$

(Hint: Note that $\vec{\gamma}_{n-1}^* = \mathbf{1}^*\Gamma$.)

(d) With $p(\lambda) \equiv \lambda^n - (\alpha_{n-1}\lambda^{n-1} + \ldots + \alpha_0)$ $(\lambda \in \mathbb{C})$, show that the characteristic polynomial of $\mathbf{H}$ equals p, where

$$\mathbf{H} \equiv \begin{bmatrix} 0 & 0 & \ldots & 0 & \alpha_0 \\ 1 & 0 & \ddots & \vdots & \alpha_1 \\ & \ddots & \ddots & & \vdots \\ & & 1 & 0 & \alpha_{n-2} \\ & & & 1 & \alpha_{n-1} \end{bmatrix}. \tag{6.12}$$

Check the following. Note that $\mathbf{H}$ is a (type of) companion matrix.[5] The (α_j) can be any sequence of n complex numbers, implying that $\mathbf{H}$ can have any eigenvalue distribution.
$\alpha_0 \neq 0 \quad \Leftrightarrow$ all eigenvalues are non-zero. With $\mathbf{b} = \mathbf{e}_1$, we have that

$$\mathcal{K}_k(\mathbf{H}, \mathbf{e}_1) = \mathrm{span}\,(\mathbf{e}_1, \ldots, \mathbf{e}_k) \quad \text{and} \quad \mathbf{H}\mathcal{K}_k(\mathbf{H}, \mathbf{e}_1) = \mathrm{span}\,(\mathbf{e}_2, \ldots, \mathbf{e}_{k+1}) \perp \mathbf{e}_1 \qquad (k < n).$$

Hence, $\|\mathbf{e}_1 - \mathbf{H}\mathbf{x}_k\|_2^2 \geq \|\mathbf{e}_1\|_2^2 = 1$ for all $\mathbf{x}_k \in \mathcal{K}_k(\mathbf{H}, \mathbf{e}_1)$ and all $k < n$.
Conclude that GMRES can stagnates from the first steps to the last but one ($\rho_0 = \ldots = \rho_{n-1}$), regardless the distribution of the eigenvalues.

(e) A matrix $\mathbf{H}$ for a general decreasing sequence of ρ_k is obtained as a combination of the matrices in (c) and (d).

[5] In the standard form of a companion matrix, the coefficients are on the first row, rather than the last column. Actually, if $\mathbf{P}$ is the permutation that reverse the ordering, then $(\mathbf{P}\widetilde{\mathbf{H}}\mathbf{P})^{\mathrm{T}}$ is in standard form.

$$\mathbf{H} \equiv \begin{bmatrix} H_1 & 0 & \dots & \\ F_1 & H_2 & \ddots & \\ 0 & F_2 & H_3 & \ddots \\ & \ddots & \ddots & \ddots \end{bmatrix}, \tag{6.13}$$

where the H_j are square matrices as in (6.12) (with a possibly different sequence of α_j in the last column), and the F_j are matrices with all entries equal to 0 except for the right top entry which is equal to $-\alpha_0^{(j)}$, with $\alpha_0^{(j)}$ the right top entry of H_j. The dimension of the H_j block matches the length of the required stagnation phase: if, in MATLAB notation, $H_j = \mathbf{H}(J, J)$ with $J = [k+1, \dots, k+\ell]$, then $\rho_{k+1} = \dots = \rho_{k+\ell} > \rho_{k+\ell+1}$; H_j is $\ell \times \ell$. In particular, if there is no stagnation at a step corresponding to H_j, ($\ell = 1$), then H_j is 1×1 and $H_j = (\alpha_0^{(j)})$.
Show that the eigenvalues of $\mathbf{H}$ equal the eigenvalues of the H_j and any distribution can be obtained by an appropriate selection of the last columns of the H_j.
Let Γ the diagonal matrix with kth diagonal entries equal to γ_{k-1} if $\gamma_{k-1} \neq 0$ and equal to 1 if $\gamma_{k-1} = 0$. Prove that $\Gamma^{-1}\mathbf{H}\Gamma$ is unreduced and $\vec{\gamma}_k^* \underline{H}_k = \vec{0}^*$ for all $k < n$.

One may hope that the convergence curve of GMRES has special properties for special matrices as Hermitian, or unitary. This is not the case, as is stated in the next proposition.

Proposition 6.4 2) *There is a unitary matrix* $\mathbf{A}$ *such that* (6.9) *holds.*
If the decrease of the ρ_i *is strict, i.e.,* $\rho_k > \rho_{k+1}$ *for all* $k = 0, \dots, n-1$, *then*
3) (6.9) *holds for some lower triangular matrix* $\mathbf{A}$,
4) *as well as for some positive definite matrix* $\mathbf{A}$.

Recall that, for normal matrices, the eigenvalues do determine the convergence (cf., Theorem 6.2). In view of the above proposition, we have to conclude that the eigenvalues of these 'nice' matrices can have a nasty distribution.

Exercise 6.13. Consider the situation of Proposition 6.4. Prove the following claims.

(a) Prove 1) of Proposition 6.4. (Hint: consider the QR-decomposition of the matrix in (6.13))

(b) If the decrease of the ρ_k is strict, then there is a lower triangular matrix $\mathbf{A}$ and a vector $\mathbf{b}$ for which (6.9) holds. (Hint: use (6.11).)

(c) If $\mathbf{A}$ is Hermitian, then we can not have stagnation in two consecutive steps, that is, if $\|\mathbf{r}_k\|_2 = \|\mathbf{r}_{k+1}\|_2$, then $\|\mathbf{r}_{k+1}\|_2 > \|\mathbf{r}_{k+2}\|_2$. (Hint: see Exercise 5.28(d)).

(d) If the sequence is such that $\rho_k = \rho_{k+1} \quad \Rightarrow \quad \rho_{k+1} > \rho_{k+2}$, then there is an Hermitian matrix $\mathbf{A}$ and a vector $\mathbf{b}$ for which (6.9) is correct.

(e) If $\mathbf{A}$ is positive definite, then GMRES does not stagnate: $\|\mathbf{r}_k\|_2 > \|\mathbf{r}_{k+1}\|_2$ for all k.

(f) If the decrease of the ρ_k is strict, then there is a positive definite matrix $\mathbf{A}$ and a vector $\mathbf{b}$ for which (6.9) holds.

Arbitrary slow convergence may occur even with eigenvalues clustering away from 0 (cf. Theorem 6.3) and also if eigenvectors form a well conditioned basis (cf. Proposition 6.4). Nevertheless, we may have quick convergence even in a situation where both the eigenvalues clustered around 0 and the eigenvectors are ill-conditioned.

Exercise 6.14. For modest positive integer p, let n be a (large) multiple of $2p$. Let C be the $p \times p$ circular matrix and S the $p \times p$ shift matrix. Let $\mathbf{A}$ be the $n \times n$ block diagonal matrix with ith diagonal blocks A_i of size $p \times p$, with $A_{2i} = C$ and $A_{2i+1} = 2I - S$.

(a) Determine the eigenvalues of $\mathbf{A}$ and $\mathcal{C}_E$.

(b) Show that $\|\mathbf{r}_p^{\text{GMRES}}\|_2 = 0$ (regardless $\mathbf{b}$ and $\mathbf{x}_0$).