Convergence minimal residual methods

- Upper bounds
- Accelerations
- Ritz values and eigenvalues
- Polynomial preconditioners

$\mathcal{P}_k = \{$polynomials of degree $\leq k\}$,
$\mathcal{P}_k^1 = \{p \in \mathcal{P}_k \mid p(0) = 1\}$

$\mathbb{A}$ is **diagonalizable** if
for some non-singular $n \times n$ matrix $\mathbb{X}$

$$\mathbb{A}\mathbb{X} = \Lambda\mathbb{X} \quad \text{with} \quad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n). \qquad (1)$$

Then $\quad \mathcal{C}_E := \min \|\mathbb{X}\|\,\|\mathbb{X}^{-1}\|$,
where min over all $\mathbb{X}$ with (1),
is the **condition number**
of the **eigen problem** of $\mathbb{A}$.

If $\mathbb{A}$ is normal, i.e. $\mathbb{A}^*\mathbb{A} = \mathbb{A}\mathbb{A}^*$, then $\mathcal{C}_E = 1$.

$\sigma(\mathbb{A}) = \{\lambda_i\}$, the **spectrum** of $\mathbb{A}$,
$\quad\quad = $ the set of eigenvalues.

Exact arithmetic

Breakdown, stagnation

If GCR does not break down
and $\quad$ GCR $\mathbf{r}_0$ = GMRES $\mathbf{r}_0$
$\quad \Rightarrow \quad$ GCR $\mathbf{r}_k$ = GMRES $\mathbf{r}_k$

$\|\mathbf{r}_{k+1}\| = \|\mathbf{r}_k\| \qquad$ (**stagnation**)
$\Leftrightarrow \rho_k = (\mathbf{r}_k, \mathbb{A}\mathbf{r}_k) = 0 \quad$ (for GCR)
$\Leftrightarrow \sigma_{k+1} = (\mathbf{c}_{k+1}, \mathbf{c}_{k+1}) = 0$
(breakdown GCR)

$(\mathbb{A}\mathbf{y}, \mathbf{y}) \neq 0 \quad \forall\, \mathbf{y} \neq 0 \Rightarrow \|\mathbf{r}_{k+1}\| < \|\mathbf{r}_k\| \quad \forall k$
no stagnation

Minimal residuals $\qquad$ [GMRES, MINRES, ...]
$\mathbf{r}_k^{\mathsf{MR}} := \mathbf{r}_0 - \mathbb{A}\mathbf{y}_k^{\mathsf{MR}} \quad$ with $\quad \mathbf{y}_k^{\mathsf{MR}} \in \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$
such that
$$\|\mathbf{r}_k^{\mathsf{MR}}\| = \min\{\|\mathbf{r}_0 - \mathbb{A}\mathbf{y}\| \mid \mathbf{y} \in \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)\}$$
or, equivalently,
$$\mathbf{r}_k^{\mathsf{MR}} \perp \mathbb{A}\mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$$

Orthogonal residuals $\qquad$ [FOM, CG, ...]
$\mathbf{r}_k^{\mathsf{OR}} := \mathbf{r}_0 - \mathbb{A}\mathbf{y}_k^{\mathsf{OR}} \quad$ with $\quad \mathbf{y}_k^{\mathsf{OR}} \in \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$
such that
$$\mathbf{r}_k^{\mathsf{OR}} \perp \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$$

[Brown '91]

Theorem. With $s := \|\mathbf{r}_k^{\mathsf{MR}}\| / \|\mathbf{r}_{k-1}^{\mathsf{MR}}\|$,
$$\boxed{\|\mathbf{r}_k^{\mathsf{MR}}\| = \sqrt{1 - s^2}\, \|\mathbf{r}_k^{\mathsf{OR}}\|}$$

Proof.
$$\mathcal{K}_{k-1}(\mathbb{A}; \mathbf{r}_0) \subset \mathbb{A}\,\mathcal{K}_{k-1}(\mathbb{A}; \mathbf{r}_0) \subset \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$$

$\mathbf{r}_{k-1}^{\mathsf{MR}} = \mathbf{r}_0 - \mathbf{w} \perp \mathbb{A}\,\mathcal{K}_{k-1}(\mathbb{A}; \mathbf{r}_0)$
for some $\mathbf{w} \in \mathbb{A}\,\mathcal{K}_{k-1}(\mathbb{A}; \mathbf{r}_0)$

Take $\mathbf{v} \in \mathbb{A}\,\mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$
such that $\mathbf{v} \perp \mathbb{A}\,\mathcal{K}_{k-1}(\mathbb{A}; \mathbf{r}_0)$, $\|\mathbf{v}\| = 1$.
Take $\mathbf{r} \in \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$
such that $\mathbf{r} \perp \mathbb{A}\,\mathcal{K}_{k-1}(\mathbb{A}; \mathbf{r}_0)$, $\mathbf{r} \neq 0$.
Observe that $\mathbf{r} = \mathbf{r}_{k-1}^{\mathsf{MR}}$

Then $\quad \mathbf{r}_k^{\mathsf{MR}} = \mathbf{r}_{k-1}^{\mathsf{MR}} - \alpha\,\mathbf{v} \perp \mathbf{v}$: $\qquad \alpha = (\mathbf{r}_{k-1}^{\mathsf{MR}}, \mathbf{v})$
$\quad \mathbf{r}_k^{\mathsf{OR}} = \mathbf{r}_{k-1}^{\mathsf{MR}} - \beta\,\mathbf{v} \perp \mathbf{r}$: $\beta = \frac{(\mathbf{r},\mathbf{r})}{(\mathbf{r},\mathbf{v})} = \frac{\|\mathbf{r}_{k-1}^{\mathsf{MR}}\|^2}{\alpha}$

Hence $\qquad \|\mathbf{r}_k^{\mathsf{MR}}\|^2 + \alpha^2 = \|\mathbf{r}_{k-1}\|^2$,
$\qquad \beta^2 = \|\mathbf{r}_{k-1}^{\mathsf{MR}}\|^2 + \|\mathbf{r}_k^{\mathsf{OR}}\|^2$

For $\mathbf{r}_0$ and GMRES (or GCR)
the residuals $\mathbf{r}_k$ are characterized by

$\mathbf{r}_k \in \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$ and $p_k \in \mathcal{P}_k^1$ are such that
$$\mathbf{r}_k = p_k(\mathbb{A})\mathbf{r}_0 \quad \text{and} \quad \|\mathbf{r}_k\| = \min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\mathbf{r}_0\|.$$

The following holds

$$\boxed{\begin{aligned} \|\mathbf{r}_k\| &\leq \left(\min_{p \in \mathcal{P}_{k-\ell}^1} \|p(\mathbb{A})\|\right) \|\mathbf{r}_\ell\| \\ &\leq \mathcal{C}_E \left(\min_{p \in \mathcal{P}_k^1} \max_{\lambda \in \sigma(\mathbb{A})} |p(\lambda)|\right) \|\mathbf{r}_0\| \qquad (*) \\ &\leq \mathcal{C}_E \left(\min_{p \in \mathcal{P}_k^1} \max_{\zeta \in \mathcal{G}} |p(\zeta)|\right) \|\mathbf{r}_0\| \qquad (**) \end{aligned}}$$

(*) if $\mathbb{A}$ diagonalizable
(**) if $\sigma(\mathbb{A}) \subset \mathcal{G}$

Minimization problems

Is there a **unique** polynomial p in $\mathcal{P}_k^1$
that solves the minimization problems?

$\min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\mathbf{r}_0\|$	Existence! Uniqueness!	GMRES
$\min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\|$	Existence! Uniqueness??	ideal GMRES
$\min_{p \in \mathcal{P}_k^1} \max_{\zeta \in \mathcal{G}} \|p(\zeta)\|$	Existence! Uniqueness!	

We are approximating the constant pol. 1
by polynomials $\zeta q(\zeta)$, where $q \in \mathcal{P}_{k-1}$,
with respect to some norm on the space of polynomials.

$$\min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\mathbf{r}_0\| \leq \left(\min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\|\right) \|\mathbf{r}_0\|$$

Choice initial residual

Theorem
$\mathbb{A}$ normal $\Rightarrow \forall\, k\ \exists\, \mathbf{r}_0$ such that "=" $\qquad$ [Greenbaum '79]
Counter examples for $\mathbb{A}$ non-normal $\qquad$ ['95]

Numerical experiments
- $\forall k$, $\exists\, \mathbf{r}_0$ such that "$\approx$".
- $\mathbf{r}_k$ and $\mathbf{r}_k'$ minimal residuals
  with initial residual $\mathbf{r}_0$ and $\mathbf{r}_0'$, resp.
  $\Rightarrow$ for some "small" $\ell \in \mathbb{Z} \qquad \|\mathbf{r}_k\| \approx \|\mathbf{r}_{k-\ell}'\|$

(proof for pos. def. case) $\qquad$ [S. & van der Sluis '96]

$\mathbb{A}$ pos. def. $\Rightarrow \min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\| < 1 \Rightarrow \|\mathbf{r}_{k+1}\| < \|\mathbf{r}_k\|$
$(\|\mathbf{I} - \epsilon\,\mathbb{A}\| < 1$ if $0 < \varepsilon \ll 1)$

Suppose $\mathbb{A}$ diagonizable (& $\mathcal{C}_E$ not too large)

Then convergence is measured by

$$\min_{p\in\mathcal{P}_k^1}\ \max_{\lambda\in\sigma(\mathbb{A})} |p(\lambda)|$$

For $\quad k\in\mathbb{N}\quad$ and $\quad\mathcal{G}\subset\mathbb{C}\quad$ we put

$$\boxed{\varepsilon_k(\mathcal{G}) := \min\left\{\max_{\zeta\in\mathcal{G}} |p(\zeta)| \mid p\in\mathcal{P}_k^1\right\}}$$

$\mathbb{A}$ diagonizable $\Rightarrow$

$$\boxed{\sigma(\mathbb{A})\subset\mathcal{G}\ \Rightarrow\ \frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|}\leq\mathcal{C}_E\,\varepsilon_k(\sigma(\mathbb{A}))\leq\mathcal{C}_E\,\varepsilon_k(\mathcal{G})}$$

Of interest if $\mathcal{C}_E$ is not too large.

Suppose $\mathbb{A}$ is normal. Then

$\mathbb{A}\mathbf{x}_i=\lambda_i\mathbf{x}_i$ where $\mathbf{x}_i\perp\mathbf{x}_j$, $\|\mathbf{x}_i\|=1\quad(i\neq j)$.

$$\mathbf{r}_0=\sum_i\mu_i\,\mathbf{x}_i\qquad\text{where}\qquad\mu_i=(\mathbf{r}_0,\mathbf{x}_i)$$

and

$$\|p(\mathbb{A})\mathbf{r}_0\|=\|\sum_i p(\lambda_i)\,\mu_i\,\mathbf{x}_i\|=\sqrt{\sum_i|p(\lambda_i)|^2\,|\mu_i|^2}$$

For some weight function ω we have that

$$\sum_i f(\lambda_i)\,|\mu_i|^2\approx\int f(\lambda)\,\omega(\lambda)\,\mathrm{d}\lambda.$$

Therefore, alternatives

$$\text{I.}\qquad\min_{p\in\mathcal{P}_k^1}\max_{\lambda\in\mathcal{G}}|p(\lambda)|$$

$$\text{II.}\quad\min_{p\in\mathcal{P}_k^1}\sqrt{\int|p(\lambda)|^2\omega(\lambda)\,\mathrm{d}\lambda}$$

$$\|\mathbf{r}_k\|\leq\mathcal{C}_E\ \min_{p\in\mathcal{P}_k^1}\max_{\zeta\in\mathcal{G}}|p(\zeta)|\,\|\mathbf{r}_0\|$$

Example. Suppose $\sigma(\mathbb{A})\subset[\lambda_-,\lambda_+]\subset(0,\infty)$.

Then, with $\mathcal{C}:=\lambda_+/\lambda_-$, [Meinardus-Kaniel]

$$\boxed{\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|}\leq 2\,\mathcal{C}_E\ \exp\left(-\frac{2k}{\sqrt{\mathcal{C}}}\right)}$$

$$k\geq\frac{1}{2}\sqrt{\mathcal{C}}\ \log\frac{2\,\mathcal{C}_E}{\epsilon}\ \Rightarrow\ \frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|}\leq\epsilon$$

On $\mathcal{G}:=[\lambda_-,\lambda_+]$,
the shifted and scaled Chebychev polynomial

$$\widetilde{p}_k(\zeta):=\frac{T_k\left(\frac{\lambda_+-\zeta}{\lambda_+-\lambda_-}-\frac{\zeta-\lambda_-}{\lambda_+-\lambda_-}\right)}{T_k\left(\frac{\lambda_++\lambda_-}{\lambda_+-\lambda_-}\right)}$$

solves

$$\min_{p\in\mathcal{P}_k^1}\ \max_{\zeta\in\mathcal{G}}|p(\zeta)|\leq\frac{1}{T_k\left(\frac{\lambda_++\lambda_-}{\lambda_+-\lambda_-}\right)}$$

Chebychev polynomials

$T_k\in\mathcal{P}_k$ are defined by

$$\begin{cases}T_0(\zeta)=1,\quad T_1(\zeta)=\zeta\\ T_{k+1}(\zeta)=2\,\zeta\,T_k(\zeta)-T_{k-1}(\zeta)\quad k=1,2,\ldots\end{cases}$$

or, equivalently,

$$T_k\left(\frac{1}{2}\left(\zeta+\frac{1}{\zeta}\right)\right)=\frac{1}{2}\left(\zeta^k+\frac{1}{\zeta^k}\right).$$

Properties.

- $|T_k|\leq 1$ on $[-1,+1]$.
- $p\in\mathcal{P}_k$, $|p|\leq 1$ on $[-1,+1]$

  $\Rightarrow\ |T_k|\geq|p|$ on $\mathbb{R}\backslash[-1,+1]$

[Gauss-Seidel, SOR, ...]

Classical methods

with **relaxation parameter** $\alpha_k=\alpha\in\mathbb{C}\ (\forall k)$

```
Choose x_0
k = 0
r_0 = b - A x_0
while  ||r_k|| > tol  do
   solve u_k from  K u = r_k
   c_k = A u_k
   x_{k+1} = x_k + α_k u_k
   r_{k+1} = r_k - α_k c_k
   k = k + 1
end while
```

(locally) **M**inimal **R**esidual: $\alpha_k=(\mathbf{r}_k,\mathbf{c}_k)/(\mathbf{c}_k,\mathbf{c}_k)$

Classical methods II

Then $\quad\mathbf{r}_k=(\mathbf{I}-\alpha\,\mathbb{A})^k\mathbf{r}_0$

$$\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|}\leq\|(\mathbf{I}-\alpha\,\mathbb{A})^k\|\leq\mathcal{C}_E\ \max_{\zeta\in\sigma(\mathbb{A})}|1-\alpha\,\zeta|^k$$

Convergence if $\quad\max_{\zeta\in\sigma(\mathbb{A})}|1-\alpha\,\zeta|<1$

Suppose $\quad\sigma(\mathbb{A})\subset[\lambda_-,\lambda_+].\quad\mathcal{C}=\lambda_+/\lambda_-$.

$\alpha=2/\left(\lambda_++\lambda_-\right)$ minimizes "max ...". Then

$$\max\ldots=\frac{\lambda_+-\lambda_-}{\lambda_++\lambda_-}$$

$$\boxed{\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|}\leq\mathcal{C}_E\left(\frac{\lambda_+-\lambda_-}{\lambda_++\lambda_-}\right)^k\leq\mathcal{C}_E\ \exp\left(-\frac{2k}{\mathcal{C}}\right)}$$

$$k\geq\frac{1}{2}\mathcal{C}\ \log\frac{\mathcal{C}_E}{\epsilon}\ \Rightarrow\ \frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|}\leq\epsilon$$

$$\frac{\eta+\eta^{-1}}{2}=\tau=\frac{\lambda_+-\zeta}{\lambda_+-\lambda_-}-\frac{\zeta-\lambda_-}{\lambda_+-\lambda_-}$$

$$T_k(\tau)=\frac{\eta^k+\eta^{-k}}{2}$$

$$\tau_0:=\frac{\rho_++\rho_+^{-1}}{2}\ \Rightarrow\ |T_k(\tau_0)|=\frac{\rho_+^k+\rho_+^{-k}}{2}$$

$$|\eta|\leq\rho_-\ \Rightarrow\ |T_k(\tau)|\leq\frac{\rho_-^k+\rho_-^{-k}}{2}$$

If $\quad|\eta|\leq\rho_-\quad$ and $\quad\widetilde{p}_k(\zeta):=\dfrac{T_k(\tau)}{T_k(\tau_0)}\quad$ then

$$|\widetilde{p}_k(\zeta)|\leq\frac{\rho_-^k+\rho_-^{-k}}{\rho_+^k+\rho_+^{-k}}$$

Estimates on ellipses

$$\mathcal{E}_\rho:=\left\{\zeta\ \mid\ \frac{|\zeta-\lambda_-|+|\zeta-\lambda_+|}{|\lambda_+-\lambda_-|}\leq\frac{\rho+\rho^{-1}}{2}\right\}$$

$$\boxed{\begin{array}{l}\sigma(\mathbb{A})\subset\mathcal{E}_{\rho_-}\quad\&\quad 0\in\partial(\mathcal{E}_{\rho_+})\quad\Rightarrow\\ \varepsilon_k(\sigma(\mathbb{A}))\leq\dfrac{\rho_-^k+\rho_-^{-k}}{\rho_+^k+\rho_+^{-k}}\leq 2\left(\dfrac{\rho_-}{\rho_+}\right)^k\end{array}}$$

Observe that $\quad\dfrac{\rho_-}{\rho_+}<1\ \Leftrightarrow\ 0\notin\mathcal{E}_{\rho_-}$

Examples.

- $\rho_-=1\ \Rightarrow\mathcal{E}_{\rho_-}=[\lambda_-,\lambda_+]$ & $\quad\dfrac{\rho_-}{\rho_+}\leq\dfrac{\sqrt{\mathcal{C}}-1}{\sqrt{\mathcal{C}}+1}$.
- $\lambda_-=\lambda_+=\mu\quad\&\quad\mathcal{E}_{\rho_-}=\{\zeta\mid|\zeta-\mu|\leq\rho_0\}$

  $\Rightarrow\quad\dfrac{\rho_-}{\rho_+}=\dfrac{\rho_0}{|\mu|}$

Estimating accelerations

$$\boxed{\begin{array}{l}\sigma(\mathbb{A}) \subset \{\lambda_0\} \cup \mathcal{E}_{\rho_-} \ \ \& \ \ 0 \in \partial\mathcal{E}_{\rho_+} \ \Rightarrow \\ \varepsilon_{k+1}(\sigma(\mathbb{A})) \leq 2\left(\frac{\rho_-}{\rho_+}\right)^k \max_{\zeta \in \mathcal{E}_{\rho_-}} \left|1 - \frac{\zeta}{\lambda_0}\right|\end{array}}$$

Proof.
If $\sigma(\mathbb{A}) \subset \{\lambda_0\} \cup \mathcal{E}_{\rho_-}$ and $0 \notin \mathcal{E}_{\rho_-}$ then, with

$$\tilde{p}_{k+1}(\zeta) := \left(1 - \frac{\zeta}{\lambda_0}\right) \frac{T_k(\tau)}{T_k(\tau_0)}$$

we have

$$|\tilde{p}_{k+1}| \leq 2\left(\frac{\rho_-}{\rho_+}\right)^k \max_{\zeta \in \mathcal{E}_{\rho_-}} \left|1 - \frac{\zeta}{\lambda_0}\right| \quad \text{on} \quad \sigma(\mathbb{A})$$

$$\|p(\mathbb{A})\| \leq \mathcal{C}_E \max_{\lambda \in \sigma(\mathbb{A})} |p(\lambda)|$$

For cases where
- $\mathbb{A}$ is not diagonizable or
- $\mathcal{C}_E$ is extremely large:

For $\varepsilon > 0$

$$\begin{aligned}\sigma_\varepsilon(\mathbb{A}) \ &:= \ \left\{\zeta \in \mathbb{C} \mid \|(\zeta\,\mathbf{I} - \mathbb{A})^{-1}\|^{-1} \leq \varepsilon\right\} \\ &= \bigcup_{\|\Delta\| \leq \varepsilon} \sigma(\mathbb{A} + \Delta)\end{aligned}$$

is the set of ε-**pseudo-eigenvalues** of $\mathbb{A}$.

[Trefethen '90]

Theorem. For arc $\mathcal{C}$ enclosing $\sigma_\varepsilon(\mathbb{A})$

$$\boxed{\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} \leq \frac{\text{length}(\mathcal{C})}{2\pi\varepsilon}\, \varepsilon_k(\sigma_\varepsilon(\mathbb{A}))}$$

Proof. $p_k(\mathbb{A})\mathbf{r}_0 = \frac{1}{2\pi i}\oint (\zeta\,\mathbf{I} - \mathbb{A})^{-1}\mathbf{r}_0\, p_k(\zeta)\, \mathrm{d}\zeta.$

Holomorphic functions

Consider a holomorphic function f on some open neighborhood $\mathcal{U} \subset \mathbb{C}$ of $\sigma(\mathbb{A})$.
$\mathcal{C}$ arc enclosing $\sigma(\mathbb{A})$, counterclockwise oriented

smooth simple
closed curve
in $\mathcal{U}$

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

$\sigma(\mathbb{A}) = \{*\}$

Matrix functions

For f holomorphic on some open neighborhood $\mathcal{U}$ of $\sigma(\mathbb{A})$, and arc $\mathcal{C}$ such that ..., define

$$\boxed{f(\mathbb{A}) := \frac{1}{2\pi i}\oint_{\mathcal{C}} (\zeta\,\mathbf{I} - \mathbb{A})^{-1} f(\zeta)\, \mathrm{d}\zeta}$$

$$(f(\mathbb{A})\mathbf{x}, \mathbf{y}) = \frac{1}{2\pi i}\oint_{\mathcal{C}} ((\zeta\,\mathbf{I} - \mathbb{A})^{-1}\mathbf{x}, \mathbf{y}) f(\zeta)\, \mathrm{d}\zeta$$

Properties. Linear in f.

$f(\zeta) = \zeta^j \Rightarrow f(\mathbb{A}) = \mathbb{A}^j.$

$g(\mathbb{A})f(\mathbb{A}) = f(\mathbb{A})g(\mathbb{A}) = fg(\mathbb{A}).$

$(g \circ f)(\mathbb{A}) = g(f(\mathbb{A})).$

$$\|f(\mathbb{A})\| \leq \frac{\text{length}\,\mathcal{C}}{2\pi} \max_{\zeta \in \mathcal{C}} \|(\zeta\,\mathbf{I} - \mathbb{A})^{-1}\| \max_{\zeta \in \mathcal{C}} |f(\zeta)|$$

$\mathbb{A}$ normal $\quad \mathbb{A} = \int \zeta\, \mathrm{d}P_\eta \quad$ spectral decomposition
$f(\mathbb{A}) = \int f(\zeta)\, \mathrm{d}P_\zeta. \quad$ spectral measure

[Greenbaum, Strakos, Ptak '96]

$$\|\mathbf{r}_k\| \leq \mathcal{C}_E \min_{p \in \mathcal{P}_k^1} \max_{\lambda \in \sigma(\mathbb{A})} |p(\lambda)|\, \|\mathbf{r}_0\|$$

Do the eigenvalues tell it all?

Given $\quad \lambda_1, \ldots, \lambda_n \in \mathbb{C}$

Given $\quad \phi_1 \geq \phi_2 \geq \ldots \geq \phi_{n-1}$

Theorem. There is an $n \times n$ irreducible Hessenberg matrix $\mathbf{H}$ and an n-vector $\mathbf{r}_0$ such that for the GMRES residuals $\mathbf{r}_k$, with initial residual $\mathbf{r}_0$, we have

$$\sigma(\mathbf{H}) = \{\lambda_1, \ldots, \lambda_n\} \quad \& \quad \|\mathbf{r}_k\| = \phi_k$$

Any monotonic decreasing graph can be the convergence history of a minimal residual method for a problem with any eigenvalue distribution

[Greenbaum, Strakos '93]

Stagnation for any eigenvalue distribution

$$\mathbf{H} = \begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \ldots & \alpha_{n-1} & \alpha_n \\ 1 & 0 & \ldots & \ldots & 0 & 0 \\ & 1 & \ddots & & & \\ & & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & 0 & 0 \\ 0 & & & & 1 & 0 \end{bmatrix}$$

$\alpha_n \neq 0$

$$\sigma(\mathbf{H}) = \{\lambda \mid \lambda^n - \alpha_1\lambda^{k-1} - \ldots - \alpha_n = 0\}$$

Solve $\ \mathbf{H}\,\tilde{\mathbf{y}} = e_n\ $ with GMRES and initial residual $\ \mathbf{r}_0 := e_n$

Note that $\ \mathbf{H}\,e_n = \alpha_n\, e_1.$

Therefore $\ \mathcal{K}_k(\mathbf{H}; \mathbf{H}\,\mathbf{r}_0) = \text{span}(e_1, \ldots, e_k).$

Hence $\ \mathbf{r}_k = e_n = \mathbf{r}_0 \quad \forall\, k < n.$

Hessenberg matrices tell it all

If $\ \rho\,\mathbf{A}\,\mathbf{r}_0 = \mathbf{c}_0, \mathbf{c}_1, \ldots\ $ orthonormal such that

$$\mathcal{K}_k(\mathbf{A}, \mathbf{A}\,\mathbf{r}_0) = \text{span}(\mathbf{c}_0, \ldots, \mathbf{c}_{k-1}) \quad \forall\, k$$

then, with $\ \mathbf{V} := [\mathbf{c}_0|\ldots|\mathbf{c}_{n-1}],$

$$\boxed{\mathbf{A}\,\mathbf{V} = \mathbf{H}\,\mathbf{V}^*, \quad \mathbf{V}^*\mathbf{V} = \mathbf{I}, \quad \mathbf{H} \text{ Hessenberg}}$$

$$\mathbf{A}\,\mathbf{y} = \mathbf{r}_0^{\mathbf{A}} = \mathbf{r}_0 \Leftrightarrow$$

$$\mathbf{H}\,\mathbf{y}^{\mathbf{H}} = \mathbf{r}_0^{\mathbf{H}} := \mathbf{V}^*\,\mathbf{r}_0^{\mathbf{A}} \quad \& \quad \mathbf{y}^{\mathbf{A}} = \mathbf{V}\,\mathbf{y}^{\mathbf{H}}$$

then $\ \|\mathbf{r}_k^{\mathbf{A}}\| = \|\mathbf{r}_k^{\mathbf{H}}\|, \quad$ since $\quad \mathbf{V}^*\mathbf{V} = \mathbf{I}$

Unitary and pos.def. tell it all

$\mathbf{H}$ Hessenberg, $\mathbf{r}_k^{\mathbf{H}}$ residuals of GMRES applied to $\ \mathbf{H}\,\tilde{\mathbf{y}} = \tilde{\mathbf{r}}_0\ $ with initial residual $\tilde{\mathbf{r}}_0$.
$\mathbf{B}$ Hessenberg, $\mathbf{r}_k^{\mathbf{B}}$ residuals of GMRES applied to $\ \mathbf{B}\,\tilde{\mathbf{y}} = \tilde{\mathbf{r}}_0\ $ with initial residual $\tilde{\mathbf{r}}_0$.

$$\boxed{\text{When is } \ \mathbf{r}_k^{\mathbf{H}} = \mathbf{r}_k^{\mathbf{B}} \ \text{ for all } k?}$$

Theorem. $\quad$ Iff $\quad \mathbf{B} = \mathbf{U}\mathbf{H}$
for some non-singular upper triangular $\mathbf{U}$.

Proof. Since $\ \mathbf{H}\,\tilde{\mathbf{r}}_0 = e_1\ $ we have
$\mathcal{K}_k(\mathbf{H}; \mathbf{H}\,\tilde{\mathbf{r}}_0) = \text{span}(e_1, \ldots, e_k) = \mathcal{K}_k(\mathbf{U}\mathbf{H}; \mathbf{U}\mathbf{H}\,\tilde{\mathbf{r}}_0)$

Applications

- $\mathbf{H} = \mathbf{U}\mathbf{Q}\ $ for some $\ \mathbf{Q}, \quad \mathbf{Q}^*\mathbf{Q} = \mathbf{I}.$
- $(\mathbf{H}\mathbf{y}, \mathbf{y}) \neq 0 \quad \forall\, \mathbf{y} \neq 0$

$$\Rightarrow \ \mathbf{H} = \mathbf{U}\mathbf{L} \ = \underbrace{(\mathbf{U}\mathbf{L}^{-*})}_{\text{upper triang.}}\underbrace{\mathbf{L}^*\mathbf{L}}_{\text{hermitean pos. def.}}$$

[van der Vorst, Vuik '93, vd V, van der Sluis '86, S. '90]

$\mathbf{r}_k$ GMRES residuals,

$\mathbb{A}$ diagonizable $\Rightarrow$ $\boxed{\dfrac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} \leq \mathcal{C}_E\, \varepsilon_k(\sigma(\mathbb{A}))}$

Super-linear convergence I

p_k GMRES pol: $p_k \in \mathcal{P}_k^1$ s.t. $\mathbf{r}_k = p_k(\mathbb{A})\mathbf{r}_0$.

Choose $\underline{\underline{\lambda \in \sigma(\mathbb{A}),\ \ell \in \mathbb{N} \text{ and a zero } \theta \text{ of } p_\ell}}$.

Put $$F := \max_{\zeta \in \sigma(\mathbb{A})} \left| \frac{1 - \zeta/\lambda}{1 - \zeta/\theta} \right|$$

Note that $\theta \approx \lambda \Rightarrow F \approx 1$

$\mathbb{A}$ diagonizable $\Rightarrow$
$$\boxed{\frac{\|\mathbf{r}_{k+\ell}\|}{\|\mathbf{r}_\ell\|} \leq \mathcal{C}_E\, F\, \varepsilon_k(\sigma(\mathbb{A})\backslash\{\lambda\})}$$

Projections

Consider a $\lambda \in \sigma(\mathbb{A})$.

Suppose λ is **semi-simple**, i.e.
$$\text{Kern}(\mathbb{A} - \lambda\mathbf{I}) \cap \text{Im}(\mathbb{A} - \lambda\mathbf{I}) = \{0\}.$$

$$\mathbb{C}^n = \underset{\mathbf{P}_\lambda}{\text{Kern}(\mathbb{A} - \lambda\mathbf{I})} \oplus \underset{\mathbf{I} - \mathbf{P}_\lambda}{\text{Im}(\mathbb{A} - \lambda\mathbf{I})} \quad \text{projections}$$

$$\mathbf{P}_\lambda\, \mathbb{A} = \mathbb{A}\, \mathbf{P}_\lambda, \qquad \mathbb{A} - \lambda\, \mathbf{I} = (\mathbb{A} - \lambda\, \mathbf{I})(\mathbf{I} - \mathbf{P}_\lambda)$$

Suppose λ **simple**, i.e. dim Kern$(\mathbb{A} - \lambda\mathbf{I}) = 1$.

$\mathbf{x}$ and $\mathbf{y}$, $\neq \mathbf{0}$, s.t. $\mathbb{A}\mathbf{x} = \lambda\mathbf{x}$ & $\mathbf{y}^*\mathbb{A} = \lambda\, \mathbf{y}^*$

$$\Rightarrow \quad \mathbf{P}_\lambda = \frac{\mathbf{x}\, \mathbf{y}^*}{\mathbf{y}^*\mathbf{x}}$$

If $\mathbb{A}$ normal then $\mathbf{y} = \mathbf{x}$

$$\mathbf{P}_\lambda = \frac{1}{2\pi i} \oint_{\mathcal{C}} (\zeta\, \mathbf{I} - \mathbb{A})^{-1} f(\zeta)\, \mathrm{d}\zeta \quad \text{with } f \text{ such that}$$
$$f(\lambda) = 1,\ f(\zeta) = 0 \text{ all } \zeta \in \sigma(\mathbb{A}),\ \zeta \neq \lambda$$

$\mathbf{r}_k$ GMRES res. $\quad p_k \in \mathcal{P}_k^1$ s.t. $\mathbf{r}_k = p_k(\mathbb{A})\mathbf{r}_0$.

$\lambda \in \sigma(\mathbb{A}), \quad \ell \in \mathbb{N}, \quad p_\ell(\theta) = 0,$

$$F := \max_{\zeta \in \sigma(\mathbb{A})} \left| \frac{1 - \zeta/\lambda}{1 - \zeta/\theta} \right|$$

Super-linear convergence II

Suppose λ semi-simple.

$$\mathbb{A}' : \text{Im}(\mathbb{A} - \lambda\,\mathbf{I}) \xrightarrow{\mathbb{A}} \text{Im}(\mathbb{A} - \lambda\,\mathbf{I})$$
$$\mathbb{A}' := \mathbb{A}_{|\text{Im}(\mathbb{A} - \lambda\mathbf{I})} = (\mathbf{I} - \mathbf{P}_\lambda)\mathbb{A}_{|\text{Im}(\mathbb{A} - \lambda\mathbf{I})}$$

$$\Rightarrow \sigma(\mathbb{A}') = \sigma(\mathbb{A})\backslash\{\lambda\} \ \& \ \mathcal{C}_E(\mathbb{A}') \leq \mathcal{C}_E(\mathbb{A}) := \mathcal{C}_E$$

<u>**Comparison process**</u>

$\mathbf{r}'_k$ is the kth GMRES residual for $\mathbb{A}'$

with $\mathbf{r}'_0 := (\mathbf{I} - \mathbf{P}_\lambda)\mathbf{r}_\ell$

$$\Rightarrow \qquad \boxed{\|\mathbf{r}_{k+\ell}\| \leq \mathcal{C}_E\, F\, \|\mathbf{r}'_k\|}$$

Proof. $\quad \widetilde{h}(\zeta) := \dfrac{1 - \zeta/\lambda}{1 - \zeta/\theta}, \qquad h := \widetilde{h}\, p_\ell$

Then
$$h \in \mathcal{P}_{k+\ell}^1, \quad \widetilde{h}(\lambda) = 0, \quad F = \max_{\zeta \in \sigma(\mathbb{A})} |\widetilde{h}(\zeta)|.$$

$$\mathbb{B} := \widetilde{h}(\mathbb{A}) = (\mathbf{I} - \mathbb{A}/\theta)^{-1}(\mathbf{I} - \mathbb{A}/\lambda)$$
$$\Rightarrow \mathbb{B} = \mathbb{B}\,(\mathbf{I} - \mathbf{P}_\lambda) = \widetilde{h}(\mathbb{A}')$$

For $\quad q_k \in \mathcal{P}_k^1, \quad$ we have $\quad q_k h \in \mathcal{P}_{k+\ell}^1 \quad$ and
$$\begin{aligned} \|\mathbf{r}_{k+l}\| \leq \|q_k(\mathbb{A})h(\mathbb{A})\mathbf{r}_0\| &= \|\mathbb{B}\, q_k(\mathbb{A})\, p_\ell(\mathbb{A})\, \mathbf{r}_0\| \\ &= \|\mathbb{B}\, q_k(\mathbb{A}')\, (\mathbf{I} - \mathbf{P}_\lambda)\, p_\ell(\mathbb{A})\, \mathbf{r}_0\| \end{aligned}$$

I.
$$\begin{aligned} \|\mathbf{r}_{k+\ell}\| &\leq \|\mathbb{B}\, q_k(\mathbb{A})\|\, \|\mathbf{r}_\ell\| \\ &\leq \mathcal{C}_E \max\nolimits_{\zeta \in \sigma(\mathbb{A})} |\widetilde{h}(\zeta) q_l(\zeta)|\, \|\mathbf{r}_\ell\| \\ &\leq \mathcal{C}_E \max_{\zeta \in \sigma(\mathbb{A})\backslash\{\lambda\}} |\widetilde{h}(\zeta) q_l(\zeta)|\, \|\mathbf{r}_\ell\| \end{aligned}$$

Now, take the minimum over all $q_k \in \mathcal{P}_k^1$.

II. Take q_k such that $\mathbf{r}'_k = q_k(\mathbb{A}')\, \mathbf{r}'_0$:
$$\|\mathbf{r}_{k+\ell}\| \leq \|\mathbb{B}\|\, \|\mathbf{r}'_k\| \leq \mathcal{C}_E \max_{\zeta \in \sigma(\mathbb{A})} |\widetilde{h}(\zeta)|\, \|\mathbf{r}'_k\|$$

Accelerations and Ritz values

$\underline{\underline{\lambda \in \sigma(\mathbb{A}),\ \ell \in \mathbb{N} \text{ and } \theta \text{ Ritz value order } \ell}}$.

Put $$F := \max_{\zeta \in \sigma(\mathbb{A})} \left| \frac{1 - \zeta/\lambda}{1 - \zeta/\theta} \right|$$

$\mathbb{A}$ diagonizable, $s = \|\mathbf{r}_\ell\|/\|\mathbf{r}_{\ell-1}\| \Rightarrow$
$$\boxed{\frac{\|\mathbf{r}_{k+\ell}\|}{\|\mathbf{r}_\ell\|} \leq \mathcal{C}_E \frac{F}{\sqrt{1 - s^2}} \varepsilon_k(\sigma(\mathbb{A})\backslash\{\lambda\})}$$

<u>**Comparison process**</u>. Suppose λ semi-simple.

$\mathbf{r}'_k$ is the kth GMRES residual for $\mathbb{A}'$ with

$\mathbf{r}'_0 := (\mathbf{I} - \mathbf{P}_\lambda)\, \mathbf{r}_\ell^{\mathrm{OR}}$, where $\mathbf{r}_\ell^{\mathrm{OR}}$ FOM residual

$$\Rightarrow \qquad \boxed{\|\mathbf{r}_{k+\ell}\| \leq \mathcal{C}_E\, F\, \|\mathbf{r}'_k\|}$$

Proof. $\widetilde{h}(\zeta) := \frac{1-\zeta/\lambda}{1-\zeta/\theta}, \quad h := \widetilde{h} p_\ell^{\mathrm{OR}}, \quad \|\mathbf{r}_\ell^{\mathrm{OR}}\| \leq \frac{1}{\kappa_\ell}\|\mathbf{r}_\ell\|$

Why approximating eigenvalues?

- If the physical problem requires the computation of eigenvalues.

- For linear problems, approximate eigenvalues provide information on the spectrum of $\mathbf{A}$ that can be used
  - to assess the effect of the preconditioner,
  - to construct an effective preconditioner,

- In the context of PDEs, approximate eigenvalues can be used
  - to analyze stability of integration method,
  - ...

Ritz-values (approximate eigenvalues) can cheaply be computed as a side-product of certain Krylov subspace methods as GCR, GMRES, ..., (Bi-CG, Bi-CGSTAB,...).

Approximate eigenvalues

If $(\theta, \mathbf{u})$ is an **approximate eigenpair**,

i.e. $\theta \in \mathbb{C}$ and $\mathbf{u} \in \mathbb{C}^n$,

then $\mathbf{r} := \mathbf{A}\mathbf{u} - \theta\mathbf{u}$ is the **residual**.

<u>**Theorem.**</u> If $\mathbf{A}$ is diagonalizable [Bauer-Fike]

then there is an eigenvalue λ of $\mathbf{A}$ such that
$$\boxed{|\lambda - \theta| \leq \mathcal{C}_E \frac{\|\mathbf{r}\|}{\|\mathbf{u}\|}}$$

<u>**Theorem**</u>. If $(\mathbf{A} + \Delta_A)\mathbf{u} = \theta\,\mathbf{u}$ then

for some eigenvalue λ with

right eigenvector $\mathbf{x}$ and left eigenvector $\mathbf{y}$,

i.e. $\quad \mathbf{A}\mathbf{x} = \lambda\,\mathbf{x}$ and $\mathbf{y}^*\,\mathbf{A} = \lambda\,\mathbf{y}^*$

Then $$\boxed{|\lambda - \theta| \lesssim \|\Delta_A\| \frac{\|\mathbf{x}\|\, \|\mathbf{y}\|}{(\mathbf{x}, \mathbf{y})}}$$

Petrov values

Consider a (skew) projection Π on a subspace $\mathcal{V}$ of $\mathbb{C}^n$ (typically dim$(\mathcal{V}) \ll n$):

$\Pi(\mathbf{z}) \in \mathcal{V} \quad (\mathbf{z} \in \mathbb{C}^n), \qquad \Pi(\mathbf{w}) = \mathbf{w} \quad (\mathbf{w} \in \mathcal{V})$

$\theta \in \mathbb{C}$ is a **Petrov value** of $\mathbf{A}$ w.r.t. Π

with **Petrov vector** $\mathbf{u} \in \mathbb{C}^n$ if
$$\mathbf{u} \in \mathcal{V} \quad \text{and} \quad \Pi(\mathbf{A}\mathbf{u} - \theta\mathbf{u}) = 0,$$

or, equivalently,
$$(\theta, \mathbf{u}) \text{ is an eigenpair of } \mathbf{A}_\Pi := \Pi\mathbf{A}_{|\mathcal{V}}$$

If the columns of $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$ span $\mathcal{V}$

and H_Π is the matrix of $\mathbf{A}_\Pi$ w.r.t. $\mathbf{v}_1, \ldots, \mathbf{v}_k$

then $(\theta, \mathbf{V}\,\vec{\gamma})$ is a **Petrov pair** $\Leftrightarrow H_\Pi\,\vec{\gamma} = \theta\,\vec{\gamma}$

Practical computation.

Compute H_Π, solve the eigenproblem for H_Π with, for instance, the QR algorithm.

Improved eigenvector approximation

For θ Petrov value, find $\tilde{\mathbf{u}} \in \mathcal{V}$ that minimizes

$$\|\mathbf{A}\tilde{\mathbf{u}} - \theta\,\tilde{\mathbf{u}}\|.$$

$\tilde{\mathbf{u}}$ is the smallest singular vector of $\mathbf{A}_{|\mathcal{V}}$.

The quality of Petrov values

Consider $\quad \mathbf{A}\mathbf{x} = \lambda\,\mathbf{x}$

Is there a Petrov value θ close to λ?

Depends on

- $\angle(\mathbf{x}, \mathcal{V})$
- $\mathcal{C}_E$
- $\|\Pi\|$

If Π is **orthogonal**,

i.e. $\quad \mathbf{z} - \Pi(\mathbf{z}) \perp \mathcal{V} \quad (\mathbf{z} \in \mathbb{C}^n)$,

then $\|\Pi\| = 1$ and the Petrov pair is a **Ritz pair**

Examples

For the circular matrix,

$$\text{for } n = 4: \quad \mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$\mathcal{V} = \text{span}(e_1, \ldots, e_k)$ gives bad approximation for $k < n$.

For any matrix,

$$\Pi(\textstyle\sum_j \mu_j\,\mathbf{v}_j) := \sum_{j<k} \mu_j\,\mathbf{v}_j$$

with $\mathbf{v}_1, \mathbf{v}_2 := \mathbf{A}\mathbf{v}_1, \mathbf{v}_3 := \mathbf{A}\mathbf{v}_2, \ldots$,

gives bad approximations for $k < n$:

$$H_\Pi(e_i) = e_{i+1}$$

Ritz values

Consider a subspace $\mathcal{V}$ of $\mathbb{C}^n$ ($\dim(\mathcal{V}) \ll n$).

Proposition. For a $\theta \in \mathbb{C}$ and a $\mathbf{u} \in \mathcal{V}$ TFAE

(a) $\mathbf{A}\mathbf{u} - \theta\,\mathbf{u} \perp \mathcal{V}$;

(b) $(\theta, \mathbf{u})$ is a Petrov pair w.r.t. Π, where

Π is the **orthogonal** projection on $\mathcal{V}$,

that is, $\Pi(z) \in \mathcal{V}$ and $\mathbf{z} - \Pi(\mathbf{z}) \perp \mathcal{V}$ $(\mathbf{z} \in \mathbb{C}^n)$.

A **Ritz pair** of $\mathbf{A}$ w.r.t. $\mathcal{V}$ is a pair $(\theta, \mathbf{u})$ with one of the above properties.

If the columns of $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$ form an orthonormal basis of $\mathcal{V}$ then

$$\Pi = \mathbf{V}\mathbf{V}^* \quad \text{and} \quad H_\Pi := H_\mathcal{V} := \mathbf{V}^*\mathbf{A}\mathbf{V}$$

$(\theta, \mathbf{u})$ Ritz-pair $\quad \Leftrightarrow \quad H_\mathcal{V}\,\vec{\gamma} = \theta\,\vec{\gamma}$ & $\mathbf{u} = \mathbf{V}\,\vec{\gamma}$

The quality of Ritz values

Consider $\quad \mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x} \quad$ with $\quad \|\mathbf{x}\| = 1$

Write

$\mathbf{x} = \mathbf{u} + \mathbf{z}$ with $\mathbf{u} \in \mathcal{V}$, $\mathbf{z} \perp \mathcal{V}$.

Now, $\quad \sin\phi_x = \text{sinus } \angle\,(\mathbf{x}, \mathcal{V}) = \|\mathbf{z}\|$.

With $\mathbf{A}_\mathcal{V} := \Pi\mathbf{A}_{|\mathcal{V}}$,

$$\begin{aligned} \|(\mathbf{A}_\mathcal{V} - \lambda)\mathbf{u}\| &= \|\mathbf{A}\,\mathbf{x} - \mathbf{A}\,\mathbf{z} - \lambda\,\mathbf{u}\| \\ &= \|(\lambda - \mathbf{A})\mathbf{z}\| \leq 2\|\mathbf{A}\|\,\sin\phi_x. \end{aligned}$$

For certain Ritz value θ w.r.t. $\mathcal{V}$ we have

$$\boxed{|\theta - \lambda| \leq \mathcal{C}_E\,2\,\|\mathbf{A}\|\,\tan\phi_x}$$

Ritz values on Krylov subspaces

Consider Ritz values on $\mathcal{V} = \mathcal{K}_k(\mathbf{A}; \mathbf{v}_1)$.

Also the power method takes approximate eigenpairs from $\mathcal{K}_k(\mathbf{A}; \mathbf{v}_1)$. For this method we know that

$\tan\phi_x \ll 1$ for "non-large" k

and for dominant $\mathbf{x}$.

However, for all $\zeta \in \mathbb{C}$,

$\mathcal{K}_k(\mathbf{A}; \mathbf{v}_1) = \text{span}\big(\mathbf{v}_1, (\mathbf{A} - \zeta)\mathbf{v}_1, \ldots, (\mathbf{A} - \zeta)^{k-1}\mathbf{v}_1\big)$.

Therefore,

$$\boxed{\tan\phi_x \ll 1 \text{ for "non-large" } k \text{ and extremal } \lambda}$$

For normal matrices

Suppose $\mathbf{A}$ is diagonizable:

$\mathbf{A}\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i, \quad \mathbf{x} = \mathbf{x}_1, \mathbf{x}_2, \ldots \quad$ basis

Write $\quad \mathbf{v}_1 = \sum_{i=1}^n \mu_i\,\mathbf{x}_i \quad$ for certain $\mu_i \in \mathbb{C}$.

Then $\quad \mathcal{K}_k(\mathbf{A}; \mathbf{v}_1) = \left\{ \sum_{i=1}^n \mu_i\,p(\lambda_i)\,\mathbf{x}_i \mid p \in \mathcal{P}_k \right\}$

Consider $\quad p \in \mathcal{P}_k, \quad \mu_1\,p(\lambda) = 1 \quad (\lambda = \lambda_1)$.

Then $\quad \sin\angle(\mathbf{x}, \mathcal{K}_k(\mathbf{A}; \mathbf{v}_1)) \leq \|\mathbf{r}\|$

where $\quad \mathbf{r} := \sum_{i=2}^n \mu_i\,p(\lambda_i)\,\mathbf{v}_i$.

Therefore,
an upper bound of $\sin\angle$ is proportional to

$$\boxed{\min\{\max_{\substack{\zeta \in \sigma(\mathbf{A}) \\ \zeta \neq \lambda}} |p(\zeta)| \mid p \in \mathcal{P}_k, \ \mu_1\,p(\lambda) = 1\}}$$

For symmetric matrices

$$\mathbf{A}^* = \mathbf{A}, \quad \lambda = \lambda_1 < \lambda_2 < \ldots < \lambda_n.$$

$$\tau := \tan\angle(\mathbf{x}, \mathbf{v}_1), \qquad \beta := \frac{\lambda_2 - \lambda_1}{\lambda_n - \lambda_1}$$

Gap ratio

[Kaniel-Paige-Saad]

Then, for the smallest Ritz value θ of order k,

$$\boxed{0 \leq \frac{\theta - \lambda}{\lambda_n - \lambda_1} \leq \tau^2 \exp(-[4k + 2]\sqrt{\beta})}$$

Projections: examples

Example I

If the columns of $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$ span $\mathcal{V}$
& the columns of $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_k]$ span $\mathcal{W}$
then $\quad \Pi := \mathbf{V}(\mathbf{W}^*\mathbf{V})^{-1}\mathbf{W}^*$
is a projection on $\mathcal{V}$ orthogonal to $\mathcal{W}$:

$$\Pi(\mathbf{z}) \in \mathcal{V} \ \ \& \ \ \mathbf{z} - \Pi(\mathbf{z}) \perp \mathcal{W} \quad (\mathbf{z} \in \mathbb{C}^n)$$

Then $\quad H_\Pi = (\mathbf{W}^*\mathbf{V})^{-1}\mathbf{W}^*\mathbf{A}\mathbf{V}.$

Example II

If $\mathbf{v}_1, \mathbf{v}_2, \ldots$ span $\mathbb{C}^n$ then

$$\Pi_k(\mathbf{z}) := \sum_{i=1}^{k} \mu_i \mathbf{v}_i \quad \text{for all} \quad \mathbf{z} = \sum_{i=1}^{n} \mu_i \mathbf{v}_i \in \mathbb{C}^n,$$

is a projection onto $\mathcal{V} := \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.
The $k \times k$ left upper block
of the matrix of $\mathbf{A}$ w.r.t. the $\mathbf{v}_j$ is
H_Π, the matrix of $\mathbf{A}_\Pi$ w.r.t. $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Krylov basis

$\mathbf{v}_1, \mathbf{v}_2, \ldots$ Krylov basis, i.e. for all k

$$\mathcal{K}_k(\mathbf{A}; \mathbf{v}_1) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$$

Then

$$\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k, \quad \mathbf{V}_k := [\mathbf{v}_1|\ldots|\mathbf{v}_k]}$$

$\underline{H}_k$ is an $(k+1) \times k$ upper Hessenberg matrix:

$$\underline{H}_k = \begin{bmatrix} \star & \star & \star & & \\ \star & \star & \star & & \\ 0 & \star & \star & & \\ \vdots & \vdots & \ddots & \ddots & \\ & & \ddots & \star & \star \\ & & & 0 & \star \end{bmatrix}$$

H_k is the $k \times k$ upper block of $\underline{H}_k$;
H_k is upper Hessenberg.

Proposition.

$$H_k \text{ irreducible} \ \Leftrightarrow \ \dim \mathcal{K}_k(\mathbf{A}; s_0) = k.$$

$\mathbf{v}_1, \mathbf{v}_2, \ldots$ Krylov basis

$$\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k, \quad \mathbf{V}_k := [\mathbf{v}_1|\ldots|\mathbf{v}_k]}$$

H_k is the $k \times k$ upper block of $\underline{H}_k$.

Projections on Krylov subspaces

Π_k is the projection onto $\mathcal{K}_k(\mathbf{A}; \mathbf{v}_1)$ given by

$$\Pi_k(\mathbf{z}) := \sum_{i=1}^{k} \mu_i \mathbf{v}_i \quad \text{for all } \mathbf{z} = \sum_{i=1}^{n} \mu_i \mathbf{v}_i \in \mathbb{C}^n.$$

Put $\quad \mathbf{A}_k := \Pi_k \mathbf{A}_{|\mathcal{K}_k(\mathbf{A};\mathbf{v}_1)}$

Proposition. H_k is the matrix of $\mathbf{A}_k$ with respect to the basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$.
$(\theta, \mathbf{V}_k\vec{\gamma})$ is a Petrov pair of $\mathbf{A}$ w.r.t. Π
$\Leftrightarrow$ $(\theta, \vec{\gamma})$ is an eigenpair of H_k.

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

Petrov values on Krylov subspaces

Consider a Petrov pair $(\theta, \mathbf{u} := \mathbf{V}_k\vec{\gamma})$:

$$H_k\,\vec{\gamma} = \theta\,\vec{\gamma}.$$

Then, with $\quad \vec{\gamma} = (\gamma_1, \ldots, \gamma_k)^T, \quad \underline{H}_k = (h_{ij}),$

$$\boxed{\begin{aligned} \mathbf{A}\mathbf{u} - \theta\,\mathbf{u} &= \mathbf{V}_{k+1}\,(\underline{H}_k\vec{\gamma} - \theta\,\vec{\gamma}) \\ &= h_{k+1\,k}\,\mathbf{v}_{k+1}\,\gamma_k \end{aligned}}$$

For certain eigenvalue λ of $\mathbf{A}$ we have

$$\boxed{|\theta - \lambda| \leq \mathcal{C}_E\,|h_{k+1\,k}|\,|\gamma_k|\,\frac{\|\mathbf{v}_{k+1}\|}{\|\mathbf{u}\|}}$$

The quantities $h_{k+1\,k}$, γ_k, $\|\mathbf{v}_k\|$ are explicitly known and a lower bound for $\|\mathbf{u}\|$ can (easily) be estimated.

$\mathbf{v}_1, \mathbf{v}_2, \ldots$ Krylov basis

$$\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k, \quad \mathbf{V}_k := [\mathbf{v}_1|\ldots|\mathbf{v}_k]}$$

H_k is the $k \times k$ upper block of $\underline{H}_k$.

Π_k is the projection onto $\mathcal{K}_k(\mathbf{A}; \mathbf{v}_1)$ given by

$$\Pi_k(\mathbf{z}) := \sum_{i=1}^{k} \mu_i \mathbf{v}_i \quad \text{for all} \quad \mathbf{z} = \sum_{i=1}^{n} \mu_i \mathbf{v}_i \in \mathbb{C}^n.$$

Ritz values on Krylov subspaces

Proposition. If $\mathbf{v}_1, \ldots$ are orthogonal then Π_k is orthogonal and $(\theta, \mathbf{V}_k\vec{\gamma})$ is a Ritz pair of $\mathbf{A}$ w.r.t. $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$ $\Leftrightarrow$ $(\theta, \vec{\gamma})$ is an eigenpair of H_k.

Examples. From Arnoldi, GMRES, GCR.

[Arnoldi '54]

Arnoldi for eigenvalue computation

Select a $\mathbf{v}_1$.

For $\quad k = 1, 2, \ldots$

- use Arnoldi's expansion to get
  $\mathbf{A}\,\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k$ with $\mathbf{V}_{k+1}^*\mathbf{V}_{k+1} = I$
- compute all eigenpairs $(\theta, \vec{\gamma})$ of H_k:
  $H_k\,\vec{\gamma} = \theta\,\vec{\gamma} \quad$ with $\quad \|\vec{\gamma}\| = 1,$
- accept $(\theta, \mathbf{u})$, with $\mathbf{u} := \mathbf{V}_k\,\vec{\gamma}$,
  as an approximate eigenpair if
  $|h_{k+1\,k}|\,|e_k^*\vec{\gamma}| < \mathit{tol}$

end

Proposition. If $(\theta, \mathbf{u})$, with $\mathbf{u} := \mathbf{V}_k\,\vec{\gamma}$, is accepted, then, for some eigenvalue λ of $\mathbf{A}$,

$$|\theta - \lambda| \leq \mathcal{C}_E\,\mathit{tol}$$

Matrix representation of GCR

Take $\quad \mathbf{v}_1 = \mathbf{A}\mathbf{r}_0.$

$\mathbf{K}^{-1}\mathbf{R}_{k+1} = \mathbf{U}_{k+1}B_{k+1} \ \Rightarrow \ \mathbf{A}\mathbf{R}_{k+1} = \mathbf{C}_{k+1}B_{k+1}$
where $\quad \mathbf{C}_{k+1} := [\mathbf{c}_0\,|\,\ldots\,|\,\mathbf{c}_k] = \mathbf{A}\mathbf{U}_{k+1}.$
The $\quad \mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{k-1} \quad$ form an orthogonal basis of $\mathcal{K}_k(\mathbf{A}; \mathbf{A}\mathbf{r}_0)$.

$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\,\mathbf{c}_k \ \Rightarrow \ \mathbf{R}_{k+1}\underline{S}_k = \mathbf{C}_k\,D_k,$ where

$$\underline{S}_k = \begin{bmatrix} 1 & 0 & \ldots & & \\ -1 & 1 & \ddots & & \\ 0 & -1 & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ & & & -1 & 1 \\ & & & 0 & -1 \end{bmatrix},$$

$$\& \quad D_k = \operatorname{diag}(\alpha_0, \ldots, \alpha_{k-1}).$$

$\mathbf{K}^{-1}\mathbf{R}_{k+1} = \mathbf{U}_{k+1}B_{k+1} \ \Rightarrow \ \mathbf{A}\mathbf{R}_{k+1} = \mathbf{C}_{k+1}B_{k+1}$
where $\quad \mathbf{C}_{k+1} = [\mathbf{c}_0\,|\,\ldots\,|\,\mathbf{c}_k] = \mathbf{A}\mathbf{U}_{k+1}.$
The $\quad \mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{k-1} \quad$ are orthogonal.

$$\mathbf{R}_{k+1}\,\underline{S}_k = \mathbf{C}_k D_k,$$

GCR and Ritz values

$$\mathbf{A}\,\mathbf{C}_k D_k = \mathbf{C}_{k+1}\,B_{k+1}\,\underline{S}_k \quad \Rightarrow$$

$$\boxed{\mathbf{A}\,\mathbf{C}_k = \mathbf{C}_{k+1}\,\underline{H}_k \text{ where } \underline{H}_k = B_{k+1}\,\underline{S}_{k+1}\,D_k^{-1}}$$

$\underline{H}_k$ is an $(k+1) \times k$ upper Hessenberg matrix.
Its matrix entries can easily be computed from β_{ij}, α_j and σ_j.

With $\Delta_k = \operatorname{diag}(\sigma_0, \ldots, \sigma_{k-1})$, we have that

$$\mathbf{C}_k^T\mathbf{C}_k = \Delta_k.$$

Therefore $\mathbf{C}_k\Delta_k^{-\frac{1}{2}}$ forms an orthonormal basis.

[Morgan '92, Paige Parlett van der vorst '95]

Harmonic Ritz values

Consider a subspace $\mathcal{V}$ of $\mathbb{C}^n$ ($\dim(\mathcal{V}) \ll n$).

Proposition. For a $\theta \in \mathbb{C}$ and a $\mathbf{u} \in \mathcal{V}$ TFAE

(a) $\mathbf{A}\mathbf{u} - \theta\,\mathbf{u} \perp \mathbf{A}\mathcal{V}$;

(b) $(\theta^{-1}, \mathbf{A}\mathbf{u})$ is a Ritz pair of $\mathbf{A}^{-1}$ w.r.t. $\mathbf{A}\mathcal{V}$.

(c) $(\theta, \mathbf{u})$ is a Petrov pair w.r.t. the projection Π on $\mathcal{V}$ that is orthogonal to $\mathcal{W} := \mathbf{A}\mathcal{V}$:

$$\Pi(z) \in \mathcal{V}, \quad \mathbf{z} - \Pi(\mathbf{z}) \perp \mathcal{W} \quad (\mathbf{z} \in \mathbb{C}^n).$$

A **harmonic Ritz pair** of $\mathbf{A}$ w.r.t. $\mathcal{V}$ is a pair $(\theta, \mathbf{u})$ with one of the above properties.

Why harmonic Ritz values?

Consider a Ritz pair $(\theta, \mathbf{u})$ of $\mathbf{A}$ w.r.t. $\mathcal{V}$:

$$\mathbf{A}\mathbf{u} - \theta\,\mathbf{u} \perp \mathcal{V}.$$

Note that $\qquad \theta = (\mathbf{A}\mathbf{u}, \mathbf{u})/(\mathbf{u}, \mathbf{u})$

Suppose $\mathbf{A}$ is normal. Write $\mathbf{u} = \sum \mu_j\,\mathbf{x}_j$.

Then $\qquad \theta = \left(\sum |\mu_j|^2\,\lambda_j\right) / \left(\sum |\mu_j|^2\right)$:

θ is a convex combination of the λ_j.

$$\boxed{\theta \approx \lambda_\ell \;\Rightarrow\; \mathbf{u} \approx \mathbf{x}_\ell?}$$

$\theta \approx \lambda_\ell$ & λ_ℓ extremal $\quad \Rightarrow \quad |\mu_j| \ll |\mu_\ell| \;\; (j \neq \ell)$

$\theta \approx \lambda_\ell$ & λ_ℓ interior $\quad \not\Rightarrow \quad |\mu_j| \ll |\mu_\ell| \;\; (j \neq \ell)$

Note $\lambda_\ell \approx 0$ then λ_ℓ^{-1} extremal eigenvalue $\mathbf{A}^{-1}$.

$\mathbf{v}_1, \mathbf{v}_2, \ldots$ Krylov basis,

$$\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k, \quad \mathbf{V}_k := [\mathbf{v}_1|\ldots|\mathbf{v}_k]}$$

Π_k is the projection onto $\mathcal{K}_k(\mathbf{A}; \mathbf{v}_1)$ given by

$$\Pi_k(\mathbf{z}) := \sum_{i=1}^{k} \mu_i\,\mathbf{v}_i \quad \text{for all} \quad \mathbf{z} = \sum_{i=1}^{n} \mu_i\,\mathbf{v}_i \in \mathbb{C}^n.$$

Harmonic Ritz values on Krylov subspaces

Proposition. If $\mathbf{v}_1, \ldots$ are $\mathbf{A}$**–orthogonal**, that is $\quad \mathbf{v}_i \perp \mathbf{A}\mathbf{v}_j, \quad (\mathbf{v}_j, \mathbf{A}\mathbf{v}_j) \neq 0 \quad (j < i)$, then Π_k is orthogonal to $\mathbf{A}\,\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$ and $(\theta, \mathbf{V}_k\vec{\gamma})$ is a harmonic Ritz pair of $\mathbf{A}$ w.r.t. $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$

$\Leftrightarrow$ $(\theta, \vec{\gamma})$ is an eigenpair of H_k.

Examples. From GCR *(and GMRES)*.

GMRES and GCR residuals

The GCR (and GMRES) residuals $\mathbf{r}_0, \ldots, \mathbf{r}_{k-1}$ form an $\mathbf{A}$-orthogonal basis of $\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$ $\forall k$.

Proposition. $\mathbf{r}_0, \mathbf{r}_1, \ldots$ are $\mathbf{A}$-orthogonal and $\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0) = \mathrm{span}(\mathbf{r}_0, \ldots, \mathbf{r}_{k-1})$ for all k.

Proof. $\mathbf{r}_i \perp \mathrm{span}(\mathbf{c}_0, \ldots, \mathbf{c}_{j-1}) = \mathcal{K}_j(\mathbf{A}; \mathbf{A}\mathbf{r}_0) = \mathbf{A}\mathcal{K}_j(\mathbf{A}; \mathbf{r}_0)$

$\mathcal{K}_j(\mathbf{A}; \mathbf{r}_0) \ni \mathbf{r}_i \quad$ for $\quad i < j$.

and $\qquad \rho_j = (\mathbf{r}_j, \mathbf{c}_j) = (\mathbf{r}_j, \mathbf{A}\mathbf{r}_j - \sum_{i<j} \beta_{ij}\mathbf{c}_i) = (\mathbf{r}_j, \mathbf{A}\mathbf{r}_j)$.

$\rho_j \neq 0 \quad$ all j $\;\Leftrightarrow\;$ no breakdown.

$\mathbf{K}^{-1}\mathbf{R}_{k+1} = \mathbf{U}_{k+1}B_{k+1} \;\Rightarrow\; \mathbf{A}\mathbf{R}_{k+1} = \mathbf{C}_{k+1}B_{k+1}$

where $\quad \mathbf{C}_{k+1} = [\mathbf{c}_0|\ldots|\mathbf{c}_k] = \mathbf{A}\mathbf{U}_{k+1}$

$$\mathbf{R}_{k+1}\,\underline{S}_k = \mathbf{C}_k\,D_k$$

GCR and harmonic Ritz values

$$\mathbf{A}\,\mathbf{R}_k = \mathbf{C}_k\,B_k = \mathbf{R}_{k+1}\,\underline{S}_k\,D_k^{-1}\,B_k$$

$$\boxed{\mathbf{A}\mathbf{R}_k = \mathbf{R}_{k+1}\,\underline{\widetilde{H}}_k \quad \text{where} \quad \underline{\widetilde{H}}_k = \underline{S}_k\,D_k^{-1}\,B_k.}$$

$\underline{\widetilde{H}}_k$ is an $(k+1) \times k$ upper Hessenberg matrix. Its matrix entries can easily be computed from β_{ij}, α_j and σ_j.

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots$ be an orthonormal Krylov basis,

$$\boxed{\mathbf{A}\,\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k, \quad \mathbf{V}_k := [\mathbf{v}_1|\ldots|\mathbf{v}_k]}$$

GMRES and harmonic Ritz values

Theorem. With $\beta := h_{k+1\,k}$, the harmonic Ritz values of $\mathbf{A}$ are the eigenvalues of

$$H_k^{-*}(\underline{H}_k^*\,\underline{H}_k) = H_k + |\beta|^2(H_k^{-*}e_k)e_k^*.$$

Proof. $(\theta, \mathbf{u} := \mathbf{V}_k\,\vec{\gamma})$ is an harmonic Ritz pair

$\Leftrightarrow \quad \mathbf{A}\mathbf{u} - \theta\,\mathbf{u} \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$

$\Leftrightarrow \quad \mathbf{A}\mathbf{V}_k\,\vec{\gamma} - \theta\,\mathbf{V}_k\,\vec{\gamma} \perp \mathbf{A}\,\mathbf{V}_k$

$\Leftrightarrow \quad \mathbf{V}_{k+1}\,\underline{H}_k\,\vec{\gamma} - \theta\mathbf{V}_k\,\vec{\gamma} \perp \mathbf{V}_{k+1}\,\underline{H}_k$

$\Leftrightarrow \quad \underline{H}_k^*\,\underline{H}_k\,\vec{\gamma} - \theta\,H_k^*\,\vec{\gamma} = 0.$

In the last "$\Leftrightarrow$" we used $\underline{H}_k^*\,\mathbf{V}_{k+1}^*\,\mathbf{V}_k = H_k^*$.

Finally, note that $\underline{H}_k = H_k + \beta e_{k+1}e_k^*$. Therefore

$$H_k^{-*}(\underline{H}_k^*\underline{H}_k) = H_k^{-*}(H_k^*H_k + |\beta|^2 e_k e_k^*) = H_k + |\beta|^2 H_k^{-*} e_k e_k^*$$

Projections on Krylov subspaces

$\mathbf{v}_1, \mathbf{v}_2, \ldots$ Krylov basis

$$\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k, \quad \mathbf{V}_k := [\mathbf{v}_1|\ldots|\mathbf{v}_k]}$$

H_k is the $k \times k$ upper block of $\underline{H}_k$.

Π_k is the projection onto $\mathcal{K}_k(\mathbf{A}; \mathbf{v}_1)$ given by

$$\Pi_k(\mathbf{z}) := \sum_{i=1}^{k} \mu_i\,\mathbf{v}_i \quad \text{for all } \mathbf{z} = \sum_{i=1}^{n} \mu_i\,\mathbf{v}_i \in \mathbb{C}^n.$$

Put $\qquad \mathbf{A}_k := \Pi_k\mathbf{A}_{|\mathcal{K}_k(\mathbf{A};\mathbf{v}_1)}$

Proposition. H_k is the matrix of $\mathbf{A}_k$ with respect to the basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

$(\theta, \mathbf{V}_k\vec{\gamma})$ is a Petrov pair of $\mathbf{A}$ w.r.t. Π

$\Leftrightarrow$ $(\theta, \vec{\gamma})$ is an eigenpair of H_k.

Hessenberg matrices and polynomials

Theorem. If $q \in \mathcal{P}_k^1$ then

$\Pi_k(q(\mathbf{A})\mathbf{r}_0) = 0 \;\Leftrightarrow$

$$\{\theta \mid q(\theta) = 0\} = \sigma(H_k) = \sigma(\mathbf{A}_k).$$

Lemma. $\qquad \sigma(H_k) = \{\theta_i\}$. For $q \in \mathcal{P}_k^1$:

(a) $\Pi_k q(\mathbf{A})\mathbf{r}_0 = q(\mathbf{A}_k)\mathbf{v}_1 = \rho\,\mathbf{V}_k\,q(H_k)e_1$

where $\quad \mathbf{r}_0 = \rho\,\mathbf{v}_1$.

(b) $\Pi_k\,(q(\mathbf{A})\mathbf{r}_0) = 0 \;\Leftrightarrow\; q(\zeta) = \prod_{j=1}^{k}\left(1 - \zeta/\theta_j\right)$.

Proof. (b) $\quad \Pi_k(q(\mathbf{A})\mathbf{r}_0) = 0 \;\Leftrightarrow\; q(H_k)e_1 = 0$.

(a) $\quad \Pi_k\mathbf{A}^j\mathbf{r}_0 = \Pi_k\mathbf{A}\,\underbrace{(\mathbf{A}^{j-1}\mathbf{r}_0)}_{\substack{\cap \\ \mathcal{K}_k(\mathbf{A};\mathbf{r}_0)}} = \mathbf{A}_k(\mathbf{A}^{j-1}\mathbf{r}_0) \underset{\text{ind.}}{=} \mathbf{A}_k^j\mathbf{r}_0 \qquad j \leq k$

$$\mathbf{A}_k^j\,\mathbf{V}_k = \mathbf{A}_k^{j-1}\,\mathbf{A}_k\,\mathbf{V}_k = \mathbf{A}_k^{j-1}\,\mathbf{V}_k\,H_k = \mathbf{V}_k\,H_k^j,$$

$$\mathbf{A}_k^j\,\mathbf{r}_0 = \mathbf{A}_k^j\,\mathbf{V}_k\,\rho\,e_1 = \rho\,\mathbf{V}_k\,H_k^j e_1.$$

Irreducible Hessenberg matrices

$$H = \begin{bmatrix} * & * & \cdots & * & * \\ h_2 & * & \cdots & * & * \\ & \ddots & \ddots & \vdots & \vdots \\ & & \ddots & * & * \\ 0 & & & h_k & * \end{bmatrix}$$

$k \times k$ Hessenberg

irreducible iff $h_i \neq 0\ \forall i$

Proposition.

$e_1,\quad He_1,\quad H^2e_1,\quad \ldots,\quad H^{k-1}e_1$ linearly independent

$$\begin{bmatrix}1\\0\\0\\0\\\vdots\\0\end{bmatrix} \begin{bmatrix}*\\h_2\\0\\0\\\vdots\\0\end{bmatrix} \begin{bmatrix}*\\*\\h_2h_3\\0\\\vdots\\0\end{bmatrix} \ldots \begin{bmatrix}*\\*\\*\\*\\\vdots\\h_2\cdots h_k\end{bmatrix}$$

Irreducible Hessenberg matrices

Proposition. $q \in \mathcal{P}_{k-1}$.

$$q(H)e_1 = 0 \ \Leftrightarrow\ q = 0$$

Proof. $0 = q(H)e_1 = \alpha_0 e_1 + \alpha_1 He_1 + \ldots + \alpha_{k-1}H^{k-1}e_1$

$\Rightarrow \alpha_0 = \alpha_1 = \ldots = \alpha_{k-1} = 0 \Rightarrow q = 0$.

Proposition. $\sigma(H) = \{\theta_i\}$. $\quad p \in \mathcal{P}_k^1$.

$$p(\zeta) = \prod_{i=1}^{k}(1 - \zeta/\theta_i) \ \Leftrightarrow\ p(H)e_1 = 0$$

Proof. Cayley-Hamilton: "$\Rightarrow$".

$p, q \in \mathcal{P}_k^1$. $\quad p(H)e_1 = q(H)e_1 = 0 \quad \& \quad q \notin \mathcal{P}_{k-1}$

$$\left.\begin{array}{l} \Rightarrow\ p - \alpha q \in \mathcal{P}_{k-1} \text{ for some } \alpha \in \mathbb{C}. \\ \quad (p(H) - \alpha q(H))e_1 = 0 \end{array}\right\} \begin{array}{l} \Rightarrow p - \alpha q = 0 \\ \Rightarrow 1 = p(0) = \alpha q(0) = \alpha. \end{array}$$

Zeros of FOM polynomials

Orthogonal (=FOM) residual polynomial:

$$p_k^{\mathrm{OR}} \in \mathcal{P}_k^1 \quad \& \quad p_k^{\mathrm{OR}}(\mathbf{A})\mathbf{r}_0 \perp \mathcal{K}_k(\mathbf{A}; \mathbf{r}_0).$$

Theorem. For a $\theta \in \mathbb{C}$ TFAE:

(a) $p_k^{\mathrm{OR}}(\theta) = 0$;

(b) there is a $\mathbf{u} \in \mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$, $\mathbf{u} \neq 0$ for which

$$\mathbf{A}\mathbf{u} - \theta\mathbf{u} \perp \mathcal{K}_k(\mathbf{A}; \mathbf{r}_0).$$

Proof. $\mathbf{v}_1, \mathbf{v}_2 \ldots$ orthonormal system as produced by FOM ($\mathbf{v}_1 = \mathbf{r}_0$) with associated orthogonal projection $\Pi_k \ \Rightarrow\ \Pi_k(p_k^{\mathrm{OR}}(\mathbf{A})\mathbf{r}_0) = 0 \ \Rightarrow$

$$\sigma(\mathbf{A}_k) = \sigma(H_k) = \{\theta \mid p_k^{\mathrm{OR}}(\theta) = 0\}$$

Zeros of GMRES polynomials

Minimal (=GMRES) residual polynomial:

$$p_k^{\mathrm{MR}} \in \mathcal{P}_k^1 \quad \& \quad p_k^{\mathrm{MR}}(\mathbf{A})\mathbf{r}_0 \perp \mathbf{A}\,\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0).$$

Theorem. For a $\theta \in \mathbb{C}$ TFAE:

(a) $p_k^{\mathrm{MR}}(\theta) = 0$;

(b) there is a $\mathbf{u} \in \mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$, $\mathbf{u} \neq 0$ for which

$$\mathbf{A}\mathbf{u} - \theta\mathbf{u} \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0).$$

Remark. The GMRES residuals $\mathbf{r}_0^{\mathrm{MR}}, \ldots, \mathbf{r}_{k-1}^{\mathrm{MR}}$ form an $\mathbf{A}$-orthogonal basis of $\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)\ \ \forall k$.

Proof. Take $\mathbf{v}_i = \mathbf{r}_i^{\mathrm{MR}}\ \ \forall i$, with associated projection Π_k. Then Π_k is orthogonal to $\mathbf{A}\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$

$\Rightarrow \Pi_k(p_k^{\mathrm{MR}}(\mathbf{A})\mathbf{r}_0) = 0 \ \Rightarrow\ p_k(H_k)e_1 = 0 \ \Rightarrow$

$$\sigma(\mathbf{A}_k) = \sigma(H_k) = \{\theta \mid p_k^G(\theta) = 0\}$$

Polynomial preconditioning

Suppose, $\quad \|\tilde{p}(\mathbb{A})\| \approx \min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\|$,

$$\tilde{p}(\zeta) = 1 - \alpha_1\zeta - \ldots - \alpha_k\zeta^k = 1 - \zeta\tilde{q}(\zeta)$$

for some k and some explicitly known $\tilde{p} \in \mathcal{P}_k^1$

To be used

- to approximate the solution of $\mathbf{A}\,\mathbf{x} = \tilde{\mathbf{b}}$:
With $\tilde{\mathbf{r}}_0 = \tilde{\mathbf{b}} - \mathbf{A}\,\tilde{\mathbf{x}}_0$, $\ \tilde{\mathbf{x}}_k = \tilde{\mathbf{x}}_0 + \mathbf{K}^{-1}\tilde{q}(\mathbb{A})\tilde{\mathbf{r}}_0$
corresponds to $\ \tilde{\mathbf{r}}_k = \tilde{\mathbf{b}} - \mathbf{A}\,\tilde{\mathbf{x}}_k$
- as a preconditioner:
$\mathbf{K}^{-1}\tilde{q}(\mathbb{A})\tilde{\mathbf{b}}$ approx. the solution of $\mathbf{A}\,\mathbf{x} = \tilde{\mathbf{b}}$.
i.e., replace "Solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u} = \mathbf{r}$" by

$$\boxed{\text{Solve } \mathbf{u} \text{ from } \mathbf{K}\,\mathbf{u} = \tilde{q}(\mathbb{A})\mathbf{r}}$$

Advantage. No DOTs are needed.

Drawback. Sensitive to choice $\tilde{p}$.

$$\tilde{p}(\zeta) = 1 - \alpha_1\zeta - \ldots - \alpha_k\zeta^k = 1 - \zeta\tilde{q}(\zeta)$$

and $\qquad \|\tilde{p}(\mathbb{A})\| \approx \min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\|$

$\tilde{\mathbf{x}}_k := \tilde{\mathbf{x}}_0 + \mathbf{K}^{-1}\tilde{q}(\mathbb{A})\tilde{\mathbf{r}}_0$ approx. solution $\mathbf{A}\mathbf{x} = \tilde{\mathbf{b}}$

Horner's scheme

$\mathbf{w} = \alpha_k\,\tilde{\mathbf{r}}_0$
for $\quad i = 1, \ldots, k-1 \quad$ do
  solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u} = \mathbf{w}$
  $\mathbf{w} \leftarrow \mathbf{A}\mathbf{u} + \alpha_{k-i}\,\tilde{\mathbf{r}}_0$
end for
solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u} = \mathbf{w}$
$\tilde{\mathbf{x}}_k = \tilde{\mathbf{x}}_0 + \mathbf{u}$
$\tilde{\mathbf{r}}_k = \tilde{\mathbf{r}}_0 - \mathbf{A}\,\mathbf{u}$

$$\tilde{p}(\zeta) = 1 - \alpha_1\zeta - \ldots - \alpha_k\zeta^k = 1 - \zeta\tilde{q}(\zeta)$$

and $\qquad \|\tilde{p}(\mathbb{A})\| \approx \min_{p \in \mathcal{P}_k^1} \|p(\mathbb{A})\|$

$\tilde{\mathbf{x}}_k := \tilde{\mathbf{x}}_0 + \mathbf{K}^{-1}\tilde{q}(\mathbb{A})\tilde{\mathbf{r}}_0$ approx. solution $\mathbf{A}\mathbf{x} = \tilde{\mathbf{b}}$

$$\tilde{p}(\zeta) = \prod_{j=1}^{k}\left(1 - \frac{\zeta}{\theta_j}\right).$$

Smolariski & Saylor scheme

for $\quad i = 1, \ldots, k \quad$ do
  solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u} = \tilde{\mathbf{r}}_{i-1}$
  $\mathbf{c} \leftarrow \mathbf{A}\mathbf{u}$
  $\tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_{i-1} + (1/\theta_i)\,\mathbf{u}$
  $\tilde{\mathbf{r}}_i = \tilde{\mathbf{r}}_{i-1} - (1/\theta_i)\,\mathbf{c}$
end for

Leja ordering

For stability,

order zero's $\tilde{p}$ according to **Leja ordering**:

$$|\theta_1| = \max_{1 \le j \le k} |\theta_j|,$$

$$|\theta_j| \prod_{i=1}^{j-1} |\theta_j - \theta_i| \le \max_{j \le \ell \le k} \left(|\theta_\ell| \prod_{i=1}^{j-1} |\theta_\ell - \theta_i|\right)$$

Real arithmetic

For real problems,

use polynomial factors of degree 2:

$\theta_2 = \overline{\theta}_1 \ \Rightarrow$

$$\left(1 - \frac{\zeta}{\theta_1}\right)\left(1 - \frac{\zeta}{\theta_2}\right) = 1 - 2\,\mathrm{Re}\left(\frac{1}{\theta_1}\right)\zeta + \frac{\zeta^2}{|\theta_1|^2}$$

Using (harmonic) Ritz values

Harmonic Ritz values are zero's of p_k, where p_k is some kth GMRES polynomial.
Apply Leja, etc..

I. Use (harmonic) Ritz values to estimate a polygonial area $\mathcal{G}$ in $\mathbb{C}$ that contains $\sigma(\mathbb{A})$.

A. Estimate an ellipse $\mathcal{E}$ ($0 \notin \mathcal{E}$) that contains $\mathcal{G}$. Take for $\widetilde{p}$ the shifted and scaled Chebychev polynomial. [Manteuffel '79]

B. Choose a weight function and compute $\widetilde{p}$ as the polynomial in $\mathcal{P}_k^1$ that minimizes on $\partial\mathcal{G}$ with respect to the weighted 2-norm. [van Gijzen '93]

II. (If $\mathbb{A}$ is positive definite.) Use the Ritz values values to estimate a weight function ω for which $\int f\omega \approx \sum f(\lambda_i)|\mu_i|^2$. Compute $\widetilde{p}$ as the polynomial in $\mathcal{P}_k^1$ (a Jacobi-type of polynomial) that minimizes with respect to the weighted 2-norm. [Fischer '93]

Using the GCR polynomial

Consider the GCR (=GMRES) polynomial p_k (with respect to some $\mathbf{r}_0$), that is
$p_k \in \mathcal{P}_k^1$ & $\mathbf{r}_k = p_k(\mathbb{A})\mathbf{r}_0$ the kth GCR residual

I. The harmonic Ritz values are the zero's of p_k and can cheaply be computed from GCR. (Apply Leja, etc..) [Nachtigal Reichel Trefethen '92]

II. Mimic GCR:

Store the coefficients β_{ij}, α_j, $\vec{\gamma}_k$ as computed in k steps GCR with initial residual $\mathbf{r}_0$.

Now apply k steps of "(efficient) GCR" with the new initial approximation, but using the **old** GCR coefficients. [Sleijpen van der Vorst Fokkema '94]

Drawback. Depends on the initial residual.

$\tilde{\mathbf{x}}_k := \tilde{\mathbf{x}}_0 + \mathbf{K}^{-1}q(\mathbb{A})\tilde{\mathbf{r}}_0$ approx. sol. $\mathbf{A}\mathbf{x} = \widetilde{\mathbf{b}}$
$\tilde{\mathbf{r}}_k = \widetilde{\mathbf{b}} - \mathbf{A}\tilde{\mathbf{x}}_k = p_k(\mathbb{A})\tilde{\mathbf{r}}_0$ where $p_k(\zeta) = 1 - \zeta q(\zeta)$

Mimic of efficient GCR

```
Choose x̃_0
r̃_0 = b̃ − A x̃_0
for ℓ = 0,...,k−1 do
    solve u from  K u = r̃_ℓ
    c ← Au
    for i = 0,...,ℓ−1 do
        c ← c − β_iℓ c̃_i
    end for
    c̃_ℓ = c,
    r̃_ℓ+1 = r̃_ℓ − α_ℓ c̃_ℓ
end for
solve y from  K y = R̃_k γ⃗_k
x̃_k = x̃_0 + y
```