Solving linear systems by GMRES with preconditioning, weighting and deflation

FreeFem Days 2023

December 8th, 2023

Nicole Spillane

(CNRS, CMAP, École Polytechnique)

Daniel Szyld

(Temple University, Philadelphia)

Solve $\mathbf{A}\mathbf{x}_* = \mathbf{b}$ for $\mathbf{A} \in \mathbb{K}^{n \times n}$ and $\mathbf{b} \in \mathbb{K}^n$.

Assumption throughout the talk

$\mathbf{A} \in \mathbb{K}^{n \times n}$ is non-singular and $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

A can be

- Hermitian / non-Hermitian,
- indefinite / positive definite,
- of very high order n,
- dense / sparse (*e.g.* finite element discretization of PDE),
- already assembled / with available element matrices (*e.g.*, finite element discretization of PDE) .

Fundamental Questions

- How fast does GMRES converge ?
- How can convergence be accelerated ?

Generalized Minimal Residual Method [Saad and Schultz (1986)]

$\mathbf{x}_i$ (approximate solution at iteration i) characterized by:

$$\mathbf{x}_i = \operatorname{argmin}_{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_i(\mathbf{A}, \mathbf{r}_0)} \{ \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \},$$

$$\text{where} \left\{ \begin{array}{l} \mathcal{K}_i(\mathbf{A}, \mathbf{r}_0) := \operatorname{span}\left\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{i-1}\mathbf{r}_0\right\} \text{ (Krylov subspace)}, \\ \mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0 \text{ (initial residual)}. \end{array} \right.$$

Generalized Minimal Residual Method [Saad and Schultz (1986)]

$\mathbf{x}_i$ (approximate solution at iteration i) characterized by:

$$\mathbf{x}_i = \operatorname{argmin}_{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_i(\mathbf{A}, \mathbf{r}_0)} \{ \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \},$$

where $\begin{cases} \mathcal{K}_i(\mathbf{A}, \mathbf{r}_0) := \operatorname{span}\left\{ \mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{i-1}\mathbf{r}_0 \right\} \text{ (Krylov subspace)}, \\ \mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0 \text{ (initial residual)}. \end{cases}$

Field of value bound aka Elman estimate [Eisenstat, Elman, Schultz (1983)]

$$\frac{\|\mathbf{r}_{i+1}\|}{\|\mathbf{r}_i\|} \leqslant \sqrt{1 - \frac{d(0, W(\mathbf{A}))^2}{\|\mathbf{A}\|^2}},$$

where

- $W(\mathbf{A}) := \left\{ \frac{\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}; \mathbf{u} \in \mathbb{C}^n \right\}$ (Field of values of $\mathbf{A}$),
- $\|\mathbf{A}\|^2 := \max \left\{ \frac{|\langle \mathbf{A}\mathbf{u}, \mathbf{A}\mathbf{u} \rangle|}{\langle \mathbf{u}, \mathbf{u} \rangle}; \mathbf{u} \in \mathbb{K}^n \setminus \{0\} \right\}$.

Long Term Objective

Choose

- **Preconditioner**,
- **Weighted norm**,
- and **Deflation** operators

in order to ensure fast convergence with respect to a **convergence bound** that must also be chosen (or proved).

Preconditioning (by H) and Weighting (by W)

Accelerating GMRES by preconditioning and weighting

- Choose a non-singular preconditioner $\mathbf{H} \in \mathbb{K}^{n \times n}$ and solve

$$\mathbf{H}\mathbf{A}\mathbf{x}_* = \mathbf{H}\mathbf{b} \text{ or } \left(\mathbf{A}\mathbf{H}\mathbf{x}_* = \mathbf{b} \text{ with } \mathbf{x}_* = \mathbf{H}\mathbf{u}_*\right).$$

Accelerating GMRES by preconditioning and weighting

- Choose a non-singular preconditioner $\mathbf{H} \in \mathbb{K}^{n \times n}$ and solve

$$\mathbf{H}\mathbf{A}\mathbf{x}_* = \mathbf{H}\mathbf{b} \text{ or } \left(\mathbf{A}\mathbf{H}\mathbf{x}_* = \mathbf{b} \text{ with } \mathbf{x}_* = \mathbf{H}\mathbf{u}_*\right).$$

- Choose a hpd weight matrix $\mathbf{W} \in \mathbb{K}^{n \times n}$, that induces $\langle \cdot, \cdot \rangle_{\mathbf{W}}$ and $\| \cdot \|_{\mathbf{W}}$ and

$$\text{replace all } \langle \cdot, \cdot \rangle \text{ in GMRES by } \langle \cdot, \cdot \rangle_{\mathbf{W}}.$$

[Cai (1989)] [Cai and Widlund (1992)] [Essai (1998)]

Accelerating GMRES by preconditioning and weighting

- Choose a non-singular preconditioner $\mathbf{H} \in \mathbb{K}^{n \times n}$ and solve

$$\mathbf{H}\mathbf{A}\mathbf{x}_* = \mathbf{H}\mathbf{b} \text{ or } (\mathbf{A}\mathbf{H}\mathbf{x}_* = \mathbf{b} \text{ with } \mathbf{x}_* = \mathbf{H}\mathbf{u}_*)\,.$$

- Choose a hpd weight matrix $\mathbf{W} \in \mathbb{K}^{n \times n}$, that induces $\langle \cdot, \cdot \rangle_{\mathbf{W}}$ and $\| \cdot \|_{\mathbf{W}}$ and

$$\text{replace all } \langle \cdot, \cdot \rangle \text{ in GMRES by } \langle \cdot, \cdot \rangle_{\mathbf{W}}.$$

  [Cai (1989)] [Cai and Widlund (1992)] [Essai (1998)]
- Characterization of the iterates:

  $\mathbf{x}_i = \operatorname{argmin}_{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_i(\mathbf{H}\mathbf{A}, \mathbf{H}\mathbf{r}_0)} \{ \|\mathbf{H}(\mathbf{b} - \mathbf{A}\mathbf{x})\|_{\mathbf{W}} \}$ or $\mathbf{x}_i = \operatorname{argmin}_{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_i(\mathbf{H}\mathbf{A}, \mathbf{H}\mathbf{r}_0)} \{ \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}} \}$,

  where $\mathcal{K}_i := \mathcal{K}_i(\mathbf{H}\mathbf{A}, \mathbf{H}\mathbf{r}_0) := \operatorname{span}\left\{ \mathbf{H}\mathbf{r}_0, \mathbf{H}\mathbf{A}\mathbf{H}\mathbf{r}_0, \dots, (\mathbf{H}\mathbf{A})^{i-1}\mathbf{H}\mathbf{r}_0 \right\}$.

Accelerating GMRES by preconditioning and weighting

- Choose a non-singular preconditioner $\mathbf{H} \in \mathbb{K}^{n\times n}$ and solve

$$\mathbf{H}\mathbf{A}\mathbf{x}_* = \mathbf{H}\mathbf{b} \text{ or } (\mathbf{A}\mathbf{H}\mathbf{x}_* = \mathbf{b} \text{ with } \mathbf{x}_* = \mathbf{H}\mathbf{u}_*)\,.$$

- Choose a hpd weight matrix $\mathbf{W} \in \mathbb{K}^{n\times n}$, that induces $\langle\cdot,\cdot\rangle_{\mathbf{W}}$ and $\|\cdot\|_{\mathbf{W}}$ and

$$\text{replace all } \langle\cdot,\cdot\rangle \text{ in GMRES by } \langle\cdot,\cdot\rangle_{\mathbf{W}}.$$

  [Cai (1989)] [Cai and Widlund (1992)] [Essai (1998)]
- Characterization of the iterates:

$$\mathbf{x}_i = \operatorname{argmin}_{\mathbf{x}\in\mathbf{x}_0+\mathcal{K}_i(\mathbf{H}\mathbf{A},\mathbf{H}\mathbf{r}_0)}\{\|\mathbf{H}(\mathbf{b}-\mathbf{A}\mathbf{x})\|_{\mathbf{W}}\} \text{ or } \mathbf{x}_i = \operatorname{argmin}_{\mathbf{x}\in\mathbf{x}_0+\mathcal{K}_i(\mathbf{H}\mathbf{A},\mathbf{H}\mathbf{r}_0)}\{\|\mathbf{b}-\mathbf{A}\mathbf{x}\|_{\mathbf{W}}\},$$

where $\mathcal{K}_i := \mathcal{K}_i(\mathbf{H}\mathbf{A},\mathbf{H}\mathbf{r}_0) := \operatorname{span}\left\{\mathbf{H}\mathbf{r}_0, \mathbf{H}\mathbf{A}\mathbf{H}\mathbf{r}_0, \dots, (\mathbf{H}\mathbf{A})^{i-1}\mathbf{H}\mathbf{r}_0\right\}$.

Field of Value estimate for preconditioned GMRES weighted by W

$$\frac{\|(\mathbf{H})\mathbf{r}_i\|_{\mathbf{W}}}{\|(\mathbf{H})\mathbf{r}_0\|_{\mathbf{W}}} \leqslant \left[1 - \frac{d(0, W_{\mathbf{W}}(\mathbf{H}\mathbf{A}))^2}{\|\mathbf{H}\mathbf{A}\|_{\mathbf{W}}{}^2}\right]^{i/2};\; W_{\mathbf{W}}(\mathbf{H}\mathbf{A}) := \left\{\frac{\langle\mathbf{H}\mathbf{A}\mathbf{u},\mathbf{u}\rangle_{\mathbf{W}}}{\langle\mathbf{u},\mathbf{u}\rangle_{\mathbf{W}}}; \mathbf{u}\in\mathbb{C}^n\right\}.$$

GMRES as GCR

Right preconditioned GCR in W-inner product

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0;$

$\mathbf{z}_0 = \mathbf{H}\mathbf{r}_0;$

$\mathbf{p}_0 = \mathbf{z}_0;\ \mathbf{q}_0 = \mathbf{A}\mathbf{p}_0;$

for $j = 0, \ldots, \mathit{itmax}$ **do**

$\delta_i = \langle \mathbf{q}_i, \mathbf{q}_i \rangle_{\mathbf{W}}; \quad \gamma_i = \langle \mathbf{q}_i, \mathbf{r}_i \rangle_{\mathbf{W}};$

$\alpha_i = \gamma_i / \delta_i;$

$\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i \mathbf{p}_i\ ;$

$\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i \mathbf{q}_i\ ;$

$\mathbf{z}_{i+1} = \mathbf{H}\mathbf{r}_{i+1};$

for $j = 0,\ 1, \ldots, i;$ ⟵

do

$\Phi_{i,j} = \langle \mathbf{q}_j, \mathbf{A}\mathbf{z}_{i+1} \rangle_{\mathbf{W}},$

$\beta_{i,j} = \Phi_{i,j} / \delta_j;$

end

$$\mathbf{p}_{i+1} = \mathbf{z}_{i+1} - \sum_{j=0}^{i} \beta_{i,j} \mathbf{p}_j;$$

$$\mathbf{q}_{i+1} = \mathbf{A}\mathbf{z}_{i+1} - \sum_{j=0}^{i} \beta_{i,j} \mathbf{q}_j;$$

end

Right preconditioned GCR in W-inner product

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
$\mathbf{z}_0 = \mathbf{H}\mathbf{r}_0$;
$\mathbf{p}_0 = \mathbf{z}_0$; $\mathbf{q}_0 = \mathbf{A}\mathbf{p}_0$;
for $j = 0, \ldots,$ *itmax* **do**
  $\delta_i = \langle \mathbf{q}_i, \mathbf{q}_i \rangle_{\mathbf{W}}$; $\gamma_i = \langle \mathbf{q}_i, \mathbf{r}_i \rangle_{\mathbf{W}}$;
  $\alpha_i = \gamma_i / \delta_i$;
  $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i \mathbf{p}_i$;
  $\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i \mathbf{q}_i$;
  $\mathbf{z}_{i+1} = \mathbf{H}\mathbf{r}_{i+1}$;
  for $j = 0, 1, \ldots, i$; $\leftarrow$
  do
    $\Phi_{i,j} = \langle \mathbf{q}_j, \mathbf{A}\mathbf{z}_{i+1} \rangle_{\mathbf{W}}$,
    $\beta_{i,j} = \Phi_{i,j} / \delta_j$;
  end
  $$\mathbf{p}_{i+1} = \mathbf{z}_{i+1} - \sum_{j=0}^{i} \beta_{i,j} \mathbf{p}_j;$$
  $$\mathbf{q}_{i+1} = \mathbf{A}\mathbf{z}_{i+1} - \sum_{j=0}^{i} \beta_{i,j} \mathbf{q}_j;$$
end

In exact arithmetic, if no breakdown
$\rightarrow$ equivalent to GMRES, *i.e.*,

$$\|\mathbf{r}_i\|_{\mathbf{W}} = \min_{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_i} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}} \quad (*),$$

where $\mathcal{K}_i := \operatorname{span}\left\{ \mathbf{H}\mathbf{r}_0, \mathbf{H}\mathbf{A}\mathbf{H}\mathbf{r}_0, \ldots, (\mathbf{H}\mathbf{A})^{i-1}\mathbf{H}\mathbf{r}_0 \right\}$.

Right preconditioned GCR in W-inner product

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
$\mathbf{z}_0 = \mathbf{H}\mathbf{r}_0$;
$\mathbf{p}_0 = \mathbf{z}_0$; $\mathbf{q}_0 = \mathbf{A}\mathbf{p}_0$;
for $j = 0, \dots,$ *itmax* **do**
  $\delta_i = \langle \mathbf{q}_i, \mathbf{q}_i \rangle_{\mathbf{W}}$; $\gamma_i = \langle \mathbf{q}_i, \mathbf{r}_i \rangle_{\mathbf{W}}$;
  $\alpha_i = \gamma_i / \delta_i$;
  $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i \mathbf{p}_i$;
  $\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i \mathbf{q}_i$;
  $\mathbf{z}_{i+1} = \mathbf{H}\mathbf{r}_{i+1}$;
  for $j = 0, 1, \dots, i$; ← *Gram-Schmidt*
  do
    $\Phi_{i,j} = \langle \mathbf{q}_j, \mathbf{A}\mathbf{z}_{i+1} \rangle_{\mathbf{W}}$,
    $\beta_{i,j} = \Phi_{i,j} / \delta_j$;
  end
  $$\mathbf{p}_{i+1} = \mathbf{z}_{i+1} - \sum_{j=0}^{i} \beta_{i,j} \mathbf{p}_j;$$
  $$\mathbf{q}_{i+1} = \mathbf{A}\mathbf{z}_{i+1} - \sum_{j=0}^{i} \beta_{i,j} \mathbf{q}_j;$$
end

In exact arithmetic, if no breakdown
→ equivalent to GMRES, *i.e.,*

$$\|\mathbf{r}_i\|_{\mathbf{W}} = \min_{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_i} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}} \quad (*),$$

where $\mathcal{K}_i := \operatorname{span}\left\{\mathbf{H}\mathbf{r}_0, \mathbf{H}\mathbf{A}\mathbf{H}\mathbf{r}_0, \dots, (\mathbf{H}\mathbf{A})^{i-1}\mathbf{H}\mathbf{r}_0\right\}$.

Proof

$\{\mathbf{q}_0, \dots, \mathbf{q}_{i-1}\}$ is a **W**-orthonormal basis of $\mathbf{A}\mathcal{K}_i$.

Right preconditioned GCR in W-inner product

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
$\mathbf{z}_0 = \mathbf{H}\mathbf{r}_0$;
$\mathbf{p}_0 = \mathbf{z}_0$; $\mathbf{q}_0 = \mathbf{A}\mathbf{p}_0$;
for $j = 0, \dots, itmax$ **do**
  $\delta_i = \langle \mathbf{q}_i, \mathbf{q}_i \rangle_{\mathbf{W}}$; $\gamma_i = \langle \mathbf{q}_i, \mathbf{r}_i \rangle_{\mathbf{W}}$;
  $\alpha_i = \gamma_i / \delta_i$;
  $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i \mathbf{p}_i$;
  $\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i \mathbf{q}_i$;
  $\mathbf{z}_{i+1} = \mathbf{H}\mathbf{r}_{i+1}$;
  for $j = 0, 1, \dots, i$; $\leftarrow$ *Gram-Schmidt*
  do
    $\Phi_{i,j} = \langle \mathbf{q}_j, \mathbf{A}\mathbf{z}_{i+1} \rangle_{\mathbf{W}}$,
    $\beta_{i,j} = \Phi_{i,j} / \delta_j$;
  end
  $\mathbf{p}_{i+1} = \mathbf{z}_{i+1} - \sum_{j=0}^{i} \beta_{i,j} \mathbf{p}_j$;
  $\mathbf{q}_{i+1} = \mathbf{A}\mathbf{z}_{i+1} - \sum_{j=0}^{i} \beta_{i,j} \mathbf{q}_j$;
end

In exact arithmetic, if no breakdown
$\rightarrow$ equivalent to GMRES, *i.e.*,

$$\|\mathbf{r}_i\|_{\mathbf{W}} = \min_{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_i} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}} \quad (*),$$

where $\mathcal{K}_i := \operatorname{span}\left\{ \mathbf{H}\mathbf{r}_0, \mathbf{HAH}\mathbf{r}_0, \dots, (\mathbf{HA})^{i-1}\mathbf{H}\mathbf{r}_0 \right\}$.

Proof

$\{\mathbf{q}_0, \dots, \mathbf{q}_{i-1}\}$ is a $\mathbf{W}$-orthonormal basis of $\mathbf{A}\mathcal{K}_i$.

$$\begin{aligned}
(*) &\Leftrightarrow \|\mathbf{r}_i\|_{\mathbf{W}} = \min_{\mathbf{r} \in \mathbf{r}_0 + \mathbf{A}\mathcal{K}_i} \|\mathbf{r}\|_{\mathbf{W}}, \\
&\Leftrightarrow \|(\mathbf{r}_0 - \mathbf{r}_i) - \mathbf{r}_0\|_{\mathbf{W}} = \min_{(\mathbf{r}_0 - \mathbf{r}) \in \mathbf{A}\mathcal{K}_i} \|(\mathbf{r}_0 - \mathbf{r}) - \mathbf{r}_0\|_{\mathbf{W}}, \\
&\Leftrightarrow \mathbf{r}_0 - \mathbf{r}_i = \mathbf{W}\text{-orthogonal projection of } \mathbf{r}_0 \text{ onto } \mathbf{A}\mathcal{K}_i
\end{aligned}$$

$$\Leftrightarrow \mathbf{r}_i = \mathbf{r}_0 - \sum_{j=0}^{i-1} \underbrace{\frac{\langle \mathbf{r}_0, \mathbf{q}_j \rangle_{\mathbf{W}}}{\langle \mathbf{q}_j, \mathbf{q}_j \rangle_{\mathbf{W}}}}_{\alpha_j} \mathbf{q}_j.$$

Convergence of Weighted Preconditioned GMRES (1/3)

- From update formula $\mathbf{r}_i = \mathbf{r}_{i+1} + \alpha_i \mathbf{q}_i$ with $\mathbf{r}_{i+1} \perp^{\mathbf{W}} \mathbf{q}_i$:

$$\|\mathbf{r}_i\|_{\mathbf{W}}^2 = \|\mathbf{r}_{i+1}\|_{\mathbf{W}}^2 + |\alpha_i|^2 \|\mathbf{q}_i\|_{\mathbf{W}}^2 = \|\mathbf{r}_{i+1}\|_{\mathbf{W}}^2 + \frac{|\langle \mathbf{q}_i, \mathbf{r}_i \rangle_{\mathbf{W}}|^2}{\|\mathbf{q}_i\|_{\mathbf{W}}^4} \|\mathbf{q}_i\|_{\mathbf{W}}^2.$$

- Equivalently:

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{W}}^2}{\|\mathbf{r}_i\|_{\mathbf{W}}^2} = 1 - \frac{|\langle \mathbf{q}_i, \mathbf{r}_i \rangle_{\mathbf{W}}|^2}{\|\mathbf{q}_i\|_{\mathbf{W}}^2 \|\mathbf{r}_i\|_{\mathbf{W}}^2}.$$

- Taking the **W**-inner product of $\mathbf{q}_i = \mathbf{A}\mathbf{z}_i - \sum_{j=0}^{i-1} \beta_{i,j} \mathbf{q}_j$ by $\mathbf{r}_i$ leads to

$$\langle \mathbf{q}_i, \mathbf{r}_i \rangle_{\mathbf{W}} = \langle \mathbf{A}\mathbf{z}_i, \mathbf{r}_i \rangle_{\mathbf{W}} - \sum_{j=0}^{i-1} \beta_{i,j} \langle \mathbf{q}_j, \mathbf{r}_i \rangle_{\mathbf{W}} = \langle \mathbf{A}\mathbf{z}_i, \mathbf{r}_i \rangle_{\mathbf{W}} = \langle \mathbf{A}\mathbf{H}\mathbf{r}_i, \mathbf{r}_i \rangle_{\mathbf{W}},$$

Convergence of Weighted Preconditioned GMRES (2/3)

- We have

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{W}}^2}{\|\mathbf{r}_i\|_{\mathbf{W}}^2} = 1 - \frac{|\langle \mathbf{AHr}_i, \mathbf{r}_i\rangle_{\mathbf{W}}|^2}{\|\mathbf{q}_i\|_{\mathbf{W}}^2 \|\mathbf{r}_i\|_{\mathbf{W}}^2}.$$

- From the orthogonalisation formula $\mathbf{q}_i = \mathbf{AHr}_i - \sum_{j=0}^{i-1} \beta_{i,j}\mathbf{q}_j$,

$$\|\mathbf{AHr}_i\|_{\mathbf{W}}^2 = \|\mathbf{q}_i + \sum_{j=0}^{i-1} \beta_{i,j}\mathbf{q}_j\|_{\mathbf{W}}^2 = \|\mathbf{q}_i\|_{\mathbf{W}}^2 + \sum_{j=0}^{i-1} |\beta_{i,j}|^2 \|\mathbf{q}_j\|_{\mathbf{W}}^2 \geqslant \|\mathbf{q}_i\|_{\mathbf{W}}^2.$$

- Finally:

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{W}}}{\|\mathbf{r}_i\|_{\mathbf{W}}} \leqslant \left[1 - \frac{|\langle \mathbf{AHr}_i, \mathbf{r}_i\rangle_{\mathbf{W}}|^2}{\|\mathbf{AHr}_i\|_{\mathbf{W}}^2 \|\mathbf{r}_i\|_{\mathbf{W}}^2}\right]^{1/2}.$$

Convergence of Weighted Preconditioned GMRES (2/3)

- We have

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{W}}^2}{\|\mathbf{r}_i\|_{\mathbf{W}}^2} = 1 - \frac{|\langle \mathbf{A}\mathbf{H}\mathbf{r}_i, \mathbf{r}_i\rangle_{\mathbf{W}}|^2}{\|\mathbf{q}_i\|_{\mathbf{W}}^2\|\mathbf{r}_i\|_{\mathbf{W}}^2}.$$

- From the orthogonalisation formula $\mathbf{q}_i = \mathbf{A}\mathbf{H}\mathbf{r}_i - \sum_{j=0}^{i-1} \beta_{i,j}\mathbf{q}_j$,

$$\|\mathbf{A}\mathbf{H}\mathbf{r}_i\|_{\mathbf{W}}^2 = \|\mathbf{q}_i + \sum_{j=0}^{i-1} \beta_{i,j}\mathbf{q}_j\|_{\mathbf{W}}^2 = \|\mathbf{q}_i\|_{\mathbf{W}}^2 + \sum_{j=0}^{i-1} |\beta_{i,j}|^2\|\mathbf{q}_j\|_{\mathbf{W}}^2 \geqslant \|\mathbf{q}_i\|_{\mathbf{W}}^2.$$

- Finally:

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{W}}}{\|\mathbf{r}_i\|_{\mathbf{W}}} \leqslant \left[1 - \frac{|\langle \mathbf{A}\mathbf{H}\mathbf{r}_i, \mathbf{r}_i\rangle_{\mathbf{W}}|^2}{\|\mathbf{A}\mathbf{H}\mathbf{r}_i\|_{\mathbf{W}}^2\|\mathbf{r}_i\|_{\mathbf{W}}^2}\right]^{1/2}.$$

Valid also for **restarted** and **truncated** GMRES (including Minimal Residual).

Convergence of Weighted Preconditioned GMRES (3/3)

Convergence Bound

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{W}}}{\|\mathbf{r}_i\|_{\mathbf{W}}} \leqslant \left[1 - \frac{|\langle \mathbf{AHr}_i, \mathbf{r}_i\rangle_{\mathbf{W}}|^2}{\|\mathbf{AHr}_i\|_{\mathbf{W}}^2 \|\mathbf{r}_i\|_{\mathbf{W}}^2}\right]^{1/2}.$$

- The terms can be grouped to prove the field of value (Elman) bound:

$$\left[\frac{|\langle \mathbf{AHy}, \mathbf{y}\rangle_{\mathbf{W}}|}{\|\mathbf{y}\|_{\mathbf{W}}^2} \times \frac{\|\mathbf{y}\|_{\mathbf{W}}}{\|\mathbf{AHy}\|_{\mathbf{W}}}\right]^2 \geqslant \left[\frac{d(0, W_{\mathbf{W}}(\mathbf{AH}))}{\|\mathbf{AH}\|_{\mathbf{W}}}\right]^2.$$

Hermitian preconditioning for A pd

Notation: Hermitian + Skew-Hermitian splitting of A

Every $\mathbf{A} \in \mathbb{K}^n$ can be split into

$$A = \mathbf{M}(\mathbf{A}) + \mathbf{N}(\mathbf{A}); \text{ with } \mathbf{M}(\mathbf{A}) = \frac{\mathbf{A} + \mathbf{A}^*}{2} \text{ and } \mathbf{N}(\mathbf{A}) = \frac{\mathbf{A} - \mathbf{A}^*}{2}.$$

- $\mathbf{M}(\mathbf{A})$ is Hermitian ($\mathbf{M}(\mathbf{A}) = \mathbf{M}(\mathbf{A})^*$),
- $\mathbf{N}(\mathbf{A})$ is skew-Hermitian ($\mathbf{N}(\mathbf{A}) = -\mathbf{N}(\mathbf{A})^*$).

Remark

$$\langle \mathbf{Ax}, \mathbf{x} \rangle = \underbrace{\langle \mathbf{M}(\mathbf{A})\mathbf{x}, \mathbf{x} \rangle}_{\in \mathbb{R}} + \underbrace{\langle \mathbf{N}(\mathbf{A})\mathbf{x}, \mathbf{x} \rangle}_{\in \mathbb{C}}, \text{ so that } |\langle \mathbf{Ax}, \mathbf{x} \rangle| \geqslant |\langle \mathbf{M}(\mathbf{A})\mathbf{x}, \mathbf{x} \rangle|.$$

Convergence of WP-GMRES with $\mathbf{H}$ spd and $\mathbf{W} = \mathbf{H}$

Two (big) assumptions: preconditioner $\mathbf{H}$ is spd and $\mathbf{W} = \mathbf{H}$.

$$
\begin{aligned}
\inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{AHy}, \mathbf{y}\rangle_{\mathbf{H}}|^2}{\|\mathbf{AHy}\|_{\mathbf{H}}^2 \|\mathbf{y}\|_{\mathbf{H}}^2} &= \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{HAHy}, \mathbf{y}\rangle|^2}{\langle \mathbf{HAHy}, \mathbf{AHy}\rangle \langle \mathbf{Hy}, \mathbf{y}\rangle} \\
&= \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{Ay}, \mathbf{y}\rangle|^2}{\langle \mathbf{HAy}, \mathbf{Ay}\rangle \langle \mathbf{H}^{-1}\mathbf{y}, \mathbf{y}\rangle} \text{ (by } \mathbf{y} \leftarrow \mathbf{Hy}) \\
&\geqslant \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{Ay}, \mathbf{y}\rangle|}{\langle \mathbf{HAy}, \mathbf{Ay}\rangle} \times \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{Ay}, \mathbf{y}\rangle|}{\langle \mathbf{H}^{-1}\mathbf{y}, \mathbf{y}\rangle} \\
&= \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{A}^{-1}\mathbf{y}, \mathbf{y}\rangle|}{\langle \mathbf{Hy}, \mathbf{y}\rangle} \times \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{Ay}, \mathbf{y}\rangle|}{\langle \mathbf{H}^{-1}\mathbf{y}, \mathbf{y}\rangle} \text{ (by } \mathbf{y} \leftarrow \mathbf{A}^{-1}\mathbf{y}) \\
&= \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{M}(\mathbf{A}^{-1})\mathbf{y}, \mathbf{y}\rangle|}{\langle \mathbf{Hy}, \mathbf{y}\rangle} \times \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{M}(\mathbf{A})\mathbf{y}, \mathbf{y}\rangle|}{\langle \mathbf{H}^{-1}\mathbf{y}, \mathbf{y}\rangle} \\
&\geqslant \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{M}(\mathbf{A}^{-1})\mathbf{y}, \mathbf{y}\rangle|}{|\langle \mathbf{M}(\mathbf{A})^{-1}\mathbf{y}, \mathbf{y}\rangle|} \times \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{M}(\mathbf{A})^{-1}\mathbf{y}, \mathbf{y}\rangle|}{\langle \mathbf{Hy}, \mathbf{y}\rangle} \times \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{M}(\mathbf{A})\mathbf{y}, \mathbf{y}\rangle|}{\langle \mathbf{H}^{-1}\mathbf{y}, \mathbf{y}\rangle}.
\end{aligned}
$$

Convergence of WP-GMRES with H spd, W = H, and A pd (1/2)

Additional assumption

$\mathbf{A}(=\mathbf{M}(\mathbf{A})+\mathbf{N}(\mathbf{A}))$ is positive definite *i.e.*, $\mathbf{M}(\mathbf{A})$ is (Hermitian) positive definite.

$$
\begin{aligned}
\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{H}}^2}{\|\mathbf{r}_i\|_{\mathbf{H}}^2} &\leqslant 1 - \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{M}(\mathbf{A}^{-1})\mathbf{y},\mathbf{y}\rangle|}{|\langle \mathbf{M}(\mathbf{A})^{-1}\mathbf{y},\mathbf{y}\rangle|} \times \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{M}(\mathbf{A})^{-1}\mathbf{y},\mathbf{y}\rangle|}{\langle \mathbf{H}\mathbf{y},\mathbf{y}\rangle} \times \inf_{\mathbf{y}\neq 0} \frac{|\langle \mathbf{M}(\mathbf{A})\mathbf{y},\mathbf{y}\rangle|}{\langle \mathbf{H}^{-1}\mathbf{y},\mathbf{y}\rangle} \\
&= 1 - \inf_{\mathbf{y}\neq 0} \frac{\langle \mathbf{M}(\mathbf{A}^{-1})\mathbf{y},\mathbf{y}\rangle}{\langle \mathbf{M}(\mathbf{A})^{-1}\mathbf{y},\mathbf{y}\rangle} \times \inf_{\mathbf{y}\neq 0} \frac{\langle \mathbf{M}(\mathbf{A})^{-1}\mathbf{y},\mathbf{y}\rangle}{\langle \mathbf{H}\mathbf{y},\mathbf{y}\rangle} \times \inf_{\mathbf{y}\neq 0} \frac{\langle \mathbf{M}(\mathbf{A})\mathbf{y},\mathbf{y}\rangle}{\langle \mathbf{H}^{-1}\mathbf{y},\mathbf{y}\rangle} \\
&\leqslant 1 - \inf_{\mathbf{y}\neq 0} \frac{\langle \mathbf{M}(\mathbf{A}^{-1})\mathbf{y},\mathbf{y}\rangle}{\langle \mathbf{M}(\mathbf{A})^{-1}\mathbf{y},\mathbf{y}\rangle} \times \frac{\lambda_{\min}(\mathbf{H}\mathbf{M}(\mathbf{A}))}{\lambda_{\max}(\mathbf{H}\mathbf{M}(\mathbf{A}))} \\
&\leqslant 1 - \frac{1}{(1+\rho(\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A}))) \times \kappa(\mathbf{H}\mathbf{M}(\mathbf{A}))} \quad \text{by [Johnson (1973, 1975)],}
\end{aligned}
$$

where ρ: spectral radius (*i.e.*, norm of eigenvalue of maximal norm).

Convergence of WP-GMRES with H spd, W = H, and A pd (2/2)

Convergence bound

$$\frac{\|\mathbf{r}_i\|_{\mathbf{H}}}{\|\mathbf{r}_0\|_{\mathbf{H}}} \leqslant \left[1 - \frac{1}{\kappa(\mathbf{HM}(\mathbf{A}))(1 + \rho(\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A}))^2)}\right]^{i/2}.$$

If **A** comes from a PDE, **H** can be chosen to be DD+GenEO and

$$\kappa(\mathbf{HM}(\mathbf{A})) \leqslant k_0(1 + k_0\tau^{-1}),$$

where

- k_0: max multiplicity of a mesh element,
- τ is set by the user.

Convergence of WP-GMRES with H spd, W = H, and A pd (2/2)

Convergence bound

$$\frac{\|\mathbf{r}_i\|_{\mathbf{H}}}{\|\mathbf{r}_0\|_{\mathbf{H}}} \leqslant \left[1 - \frac{1}{\kappa(\mathbf{HM}(\mathbf{A}))(1 + \rho(\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A}))^2)}\right]^{i/2}.$$

If **A** comes from a PDE, **H** can be chosen to be DD+GenEO and

$$\kappa(\mathbf{HM}(\mathbf{A})) \leqslant k_0(1 + k_0\tau^{-1}),$$

where

- k_0: max multiplicity of a mesh element,
- τ is set by the user.

The bound does not depend on the number of subdomains !
→Scalability

Numerical results (1/2)

Advection Diffusion Reaction

Strong Formulation

$$\begin{cases} c_0 u + \operatorname{div}(\mathbf{a} u) - \operatorname{div}(\nu \nabla u) & = f, \text{ in } \Omega = [0,1]^2, \\ u & = 0; \text{ on } \partial\Omega. \end{cases}$$

In our numerics

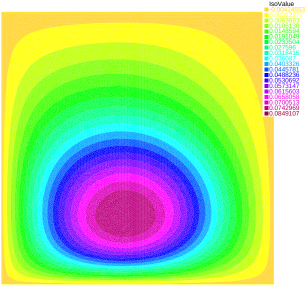
- $f(x, y) = \exp(-10((x - 0.5)^2 + (y - 0.1)^2))$,
- $\mathbf{a} = 2\pi[-(y - 0.1), x - 0.5]$.

Variational Formulation

Find $\mathbf{u} \in H_0^1(\Omega)$ such that:

$$\underbrace{\int_\Omega \left(\left(c_0 + \frac{1}{2} \operatorname{div} \mathbf{a} \right) uv + \nu \nabla u \cdot \nabla v \right)}_{\text{symmetric part}} + \underbrace{\int_\Omega \left(\frac{1}{2} \mathbf{a} \cdot \nabla uv - \frac{1}{2} \mathbf{a} \cdot \nabla vu \right)}_{\text{skew-symmetric part}} = \int_\Omega fv, \quad \forall\, v \in H_0^1(\Omega).$$

Solution for $\nu = c_0 = 1$



The overall bound does not depend on h or number of subdomains.

Lagrange Finite Element discretization of

$$\underbrace{\int_\Omega ((c_0 + \frac{1}{2}\operatorname{div}\mathbf{a})uv + \nu\nabla u \cdot \nabla v)}_{\text{"}\mathbf{M}(\mathbf{A})\text{"}} + \underbrace{\int_\Omega (\frac{1}{2}\mathbf{a}\cdot\nabla uv - \frac{1}{2}\mathbf{a}\cdot\nabla vu)}_{\text{"}\mathbf{N}(\mathbf{A})\text{"}} = \int_\Omega fv.$$

Recall that

$$\frac{\|\mathbf{r}_i\|_\mathbf{H}}{\|\mathbf{r}_0\|_\mathbf{H}} \leqslant \left[1 - \frac{1}{\kappa(\mathbf{HM}(\mathbf{A}))(1+\rho(\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A}))^2)}\right]^{i/2}$$

If H is DD + GenEO:

$\kappa(\mathbf{HM}(\mathbf{A}))$ independent of

- discretization step
- and number of subdomains.

Bound for the spectral radius of $\mathbf{M}(\mathbf{A})^{-1}\,\mathbf{N}(\mathbf{A})$

$$\rho(\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A})) \leqslant \|\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A})\|_{\mathbf{M}(\mathbf{A})} \leqslant \frac{1}{2}\frac{\|\mathbf{a}\|_{L^\infty(\Omega)}}{\sqrt{\inf(\nu)\inf(c_0+\frac{1}{2}\operatorname{div}(\mathbf{a}))}}.$$

The overall bound does not depend on *h* or number of subdomains.

Lagrange Finite Element discretization of

$$\underbrace{\int_\Omega ((c_0 + \frac{1}{2}\operatorname{div}\mathbf{a})uv + \nu\nabla u\cdot\nabla v)}_{\text{"}\mathbf{M(A)}\text{"}} + \underbrace{\int_\Omega (\frac{1}{2}\mathbf{a}\cdot\nabla uv - \frac{1}{2}\mathbf{a}\cdot\nabla vu)}_{\text{"}\mathbf{N(A)}\text{"}} = \int_\Omega fv.$$

Recall that

$$\frac{\|\mathbf{r}_i\|_{\mathbf{H}}}{\|\mathbf{r}_0\|_{\mathbf{H}}} \leqslant \left[1 - \frac{1}{\kappa(\mathbf{HM(A)})(1+\rho(\mathbf{M(A)}^{-1}\mathbf{N(A)})^2)}\right]^{i/2}$$

If H is DD + GenEO:

$\kappa(\mathbf{HM(A)})$ independent of

- discretization step
- and number of subdomains.

Bound for the spectral radius of $\mathbf{M(A)}^{-1}\,\mathbf{N(A)}$

$$\rho(\mathbf{M(A)}^{-1}\mathbf{N(A)}) \leqslant \|\mathbf{M(A)}^{-1}\mathbf{N(A)}\|_{\mathbf{M(A)}} \leqslant \frac{1}{2}\frac{\|\mathbf{a}\|_{L^\infty(\Omega)}}{\sqrt{\inf(\nu)\inf(c_0+\frac{1}{2}\operatorname{div}(\mathbf{a}))}}.$$

Proof uses [Bonazzoli, Claeys, Nataf, Tournier (2021)]

The overall bound does not depend on h or number of subdomains.

Lagrange Finite Element discretization of

$$\underbrace{\int_\Omega ((c_0 + \frac{1}{2}\operatorname{div}\mathbf{a})uv + \nu\nabla u \cdot \nabla v)}_{\text{"}\mathbf{M}(\mathbf{A})\text{"}} + \underbrace{\int_\Omega (\frac{1}{2}\mathbf{a}\cdot\nabla uv - \frac{1}{2}\mathbf{a}\cdot\nabla vu)}_{\text{"}\mathbf{N}(\mathbf{A})\text{"}} = \int_\Omega fv.$$

Recall that

$$\frac{\|\mathbf{r}_i\|_{\mathbf{H}}}{\|\mathbf{r}_0\|_{\mathbf{H}}} \leqslant \left[1 - \frac{1}{\kappa(\mathbf{HM(A)})(1 + \rho(\mathbf{M(A)}^{-1}\mathbf{N(A)})^2)}\right]^{i/2}$$

If H is DD + GenEO:

$\kappa(\mathbf{HM(A)})$ independent of

- discretization step
- and number of subdomains.

Bound for the spectral radius of $\mathbf{M(A)}^{-1}$ $\mathbf{N(A)}$

$$\rho(\mathbf{M(A)}^{-1}\mathbf{N(A)}) \leqslant \|\mathbf{M(A)}^{-1}\mathbf{N(A)}\|_{\mathbf{M(A)}} \leqslant \frac{1}{2}\frac{\|\mathbf{a}\|_{L^\infty(\Omega)}}{\sqrt{\inf(\nu)\inf(c_0 + \frac{1}{2}\operatorname{div}(\mathbf{a}))}}.$$

Proof uses [Bonazzoli, Claeys, Nataf, Tournier (2021)]

Does not depend on h !

Scalability

- Freefem++ with ffddm developed by Tournier, Hecht, Jolivet, Nataf.
- Weighted GMRES with (DD + GenEO) preconditioner of $\mathbf{M}(\mathbf{A})$.

```
-ffddm_schwarz_method asm
-ffddm_geneo_threshold 0.15
-ffddm_schwarz_coarse_correction BNN.
```

- $\int_\Omega ((c_0 + \frac{1}{2}\operatorname{div}\mathbf{a})uv + \nu \nabla u \cdot \nabla v) + \int_\Omega (\frac{1}{2}\mathbf{a} \cdot \nabla uv - \frac{1}{2}\mathbf{a} \cdot \nabla vu) = \int_\Omega fv.$
  - $\mathbf{a} = 2\pi[-(y - 0.1), x - 0.5]$
  - $c_0 = \nu = 1$

Iteration count when number of subdomains and h vary

Number of subdomains	4	8	16	32
$h = 1/200$	19	20	20	20
$h = 1/500$	18	19	19	20

Dependency on h

- Freefem++ with ffddm developed by Tournier, Hecht, Jolivet, Nataf.
- Weighted GMRES with (DD + GenEO) preconditioner of $\mathbf{M}(\mathbf{A})$.
  - Partition into 8 subdomains computed by Metis,
  - Stopping criterion: $\|\mathbf{Hr}_i\| < 10^{-6}$.
- $\displaystyle\int_\Omega ((c_0 + \frac{1}{2}\operatorname{div}\mathbf{a})uv + \nu\nabla u \cdot \nabla v) + \int_\Omega (\frac{1}{2}\mathbf{a} \cdot \nabla uv - \frac{1}{2}\mathbf{a} \cdot \nabla vu) = \int_\Omega fv.$
  - $\mathbf{a} = 2\pi[-(y - 0.1), x - 0.5]$

Iteration count when h and strength of non-symmetry vary

	$1/h$	1000	500	200	100
	global number of dofs	1 002 001	251 001	40 401	10 201
$c_0 = \nu = 0.1$	iteration count	40	42	43	41
$c_0 = \nu = 1$	iteration count	18	19	20	21
$c_0 = \nu = 10$	iteration count	16	17	17	20

Dependency on strength of non-symmetry

- Freefem++ with ffddm developed by Tournier, Hecht, Jolivet, Nataf.
- Weighted GMRES with (DD + GenEO) preconditioner of $\mathbf{M}(\mathbf{A})$.
  - $h = 1/500$,
  - Partition into 8 subdomains computed by Metis,
  - Stopping criterion: $\|\mathbf{H}\mathbf{r}_i\| < 10^{-6}$.
- $\int_\Omega ((c_0 + \frac{1}{2}\operatorname{div}\mathbf{a})uv + \nu\nabla u \cdot \nabla v) + \int_\Omega (\frac{1}{2}\mathbf{a}\cdot\nabla uv - \frac{1}{2}\mathbf{a}\cdot\nabla vu) = \int_\Omega fv.$
  - $\mathbf{a} = 2\pi[-(y-0.1), x-0.5]$

Discretization step when strength of non-symmetry varies

$c_0 = \nu$	0.001	0.01	0.1	1	10	only symmetric part
Iteration count	> 1000	161	42	19	17	17

1 Preconditioning (by $\mathbf{H}$) and Weighting (by $\mathbf{W}$)

2 GMRES as GCR

3 Hermitian preconditioning for $\mathbf{A}$ pd

4 Numerical results (1/2)

5 Spectral Deflation

6 Numerical results (2/2)

Spectral Deflation

Deflation

Following [Tang, Nabben, Vuik, Erlangga (2009)] [García Ramos, Kehl, Nabben (2020)]

- Choose $\mathbf{Y}, \mathbf{Z} \in \mathbb{K}^{n\times m}$ two full rank matrices.
- Let $\mathbf{P}_D = \mathbf{I} - \mathbf{AZ}(\mathbf{Y}^*\mathbf{AZ})^{-1}\mathbf{Y}^*$ and $\mathbf{Q}_D = \mathbf{I} - \mathbf{Z}(\mathbf{Y}^*\mathbf{AZ})^{-1}\mathbf{Y}^*\mathbf{A}$.
  These are projection operators if $\mathbf{Y}^*\mathbf{AZ}$ is non-singular.

$$\begin{aligned}\mathbf{Ax}_* = \mathbf{b} &\Leftrightarrow (\mathbf{P}_D\mathbf{Ax}_* = \mathbf{P}_D\mathbf{b} \text{ and } (\mathbf{I} - \mathbf{P}_D)\mathbf{Ax}_* = (\mathbf{I} - \mathbf{P}_D)\mathbf{b}) \\ &\Leftrightarrow (\mathbf{A}\underbrace{\mathbf{Q}_D\mathbf{x}_*}_{:=\mathbf{y}_*} = \mathbf{P}_D\mathbf{b} \text{ and } \mathbf{A}\underbrace{(\mathbf{I} - \mathbf{Q}_D)\mathbf{x}_*}_{:=\mathbf{z}_*} = (\mathbf{I} - \mathbf{P}_D)\mathbf{b})\end{aligned}$$

- Krylov subspace solver is applied to:

$$\mathbf{P}_D\mathbf{AH}\tilde{\mathbf{u}} = \mathbf{P}_D\mathbf{b} \text{ and } \mathbf{Q}_D\mathbf{H}\tilde{\mathbf{u}} = \mathbf{Q}_D\mathbf{x}_* = \mathbf{y}_*.$$

Requirements [García Ramos, Kehl, Nabben (2020)]

- For the projection operators to be well defined:

$$\ker(\mathbf{Y}^*) \cap \mathrm{range}(\mathbf{AZ}) = \{\mathbf{0}\}, \text{ or equivalently } \mathbf{Y}^*\mathbf{AZ} \text{ is non-singular.}$$

$$\rightarrow \text{OK if } \mathbf{Z} = \mathbf{Y} = \mathbf{AY} \text{ or if } \mathbf{Y} = \mathbf{AZ}.$$

- For the iterations to be well defined when applying $\mathbf{HP}_D\mathbf{AQ}_D$:

$$\ker(\mathbf{Q}_D) \cap \mathrm{range}(\mathbf{HP}_D) = \{\mathbf{0}\} \Leftrightarrow \mathrm{range}(\mathbf{Z}) \cap \ker(\mathbf{Y}^*\mathbf{H}^{-1}) = \{\mathbf{0}\}$$
$$\Leftrightarrow \mathbf{Y}^*\mathbf{H}^{-1}\mathbf{Z} \text{ is non-singular.}$$

$$\rightarrow \text{OK if } \mathbf{Z} = \mathbf{HY} \text{ or } \mathbf{Y} = \mathbf{H}^*\mathbf{Z}.$$

Remark

Both OK if $\mathbf{Y} = \mathbf{AZ} = \mathbf{AHY}$.

Bound for WPD-GMRES

If $\mathbf{Y} = \mathbf{HAZ}$, then $\mathbf{P}_D$ is $\mathbf{H}$-orthogonal and

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{H}}^2}{\|\mathbf{r}_i\|_{\mathbf{H}}^2} \leqslant 1 - \inf_{\mathbf{y}\in\ker(\mathbf{Z}^*\mathbf{A}^*\mathbf{A}^{-1})\setminus\{\mathbf{0}\}} \frac{\langle \mathbf{A}^{-1}\mathbf{y}, \mathbf{y}\rangle}{\langle \mathbf{M}(\mathbf{A})^{-1}\mathbf{y}, \mathbf{y}\rangle} \times \kappa(\mathbf{HM}(\mathbf{A})).$$

From $\mathbf{A}^{-1} = (\mathbf{M}(\mathbf{A}) + \mathbf{N}(\mathbf{A}))^{-1} = (\mathbf{I} + \mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A}))^{-1}\mathbf{M}(\mathbf{A})^{-1}$.

After some work, it all comes down to the spectrum of $\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A})$:

Consider the gevp: find $\lambda_j \in \mathbb{C}$ and $\mathbf{z}^{(j)} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ such that

$$\mathbf{N}(\mathbf{A})\mathbf{z}^{(j)} = \lambda_j \mathbf{M}(\mathbf{A})\mathbf{z}^{(j)}.$$

- Eigenvectors $\mathbf{z}^{(j)}$ form an $\mathbf{M}(\mathbf{A})$-orthonormal basis of $\mathbb{C}^n$,
- Eigenvalues λ_j are either 0 or purely imaginary and in complex conjugate pairs.

Convergence Theorem

Let us assume that **A** is positive definite, **H** is hpd, and **W** = **H**. With

- $\operatorname{span}(\mathbf{Z}) = \operatorname{span}\{\mathbf{z}^{(j)}; |\lambda_j| > \tau\}$ for $\mathbf{N}(\mathbf{A})\mathbf{z}^{(j)} = \lambda_j \mathbf{M}(\mathbf{A})\mathbf{z}^{(j)}$,
- and $\mathbf{Y} = \mathbf{HAZ}$,

the convergence of WPD-GMRES is bounded by

$$\frac{\|\mathbf{r}_{i+1}\|_{\mathbf{H}}^2}{\|\mathbf{r}_i\|_{\mathbf{H}}^2} \leqslant 1 - \frac{1}{\kappa(\mathbf{HM}(\mathbf{A})) \times (1 + \tau^2)}.$$

Numerical results (2/2)

Solution of the Generalized eigenvalue problem

Variational Formulation discretized by $\mathbb{P}_1$ finite elements with 31502 dofs

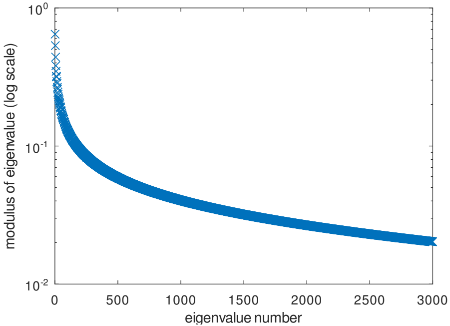
With $\mathbf{a} = 2\pi[-(y-0.1), x-0.5]$, find $\mathbf{u} \in H^1_0(\Omega)$ such that:

$$\underbrace{\int_\Omega \left(\left(1+\frac{1}{2}\operatorname{div}\mathbf{a}\right)uv + \nabla u \cdot \nabla v\right)}_{\mathbf{M}(\mathbf{A})} + \underbrace{\int_\Omega \eta\left(\frac{1}{2}\mathbf{a}\cdot\nabla uv - \frac{1}{2}\mathbf{a}\cdot\nabla vu\right)}_{\mathbf{N}(\mathbf{A})} = \int_\Omega fv, \quad \forall\, v \in H^1_0(\Omega),$$

First 3000 (out of 31502) eigenvalues of :

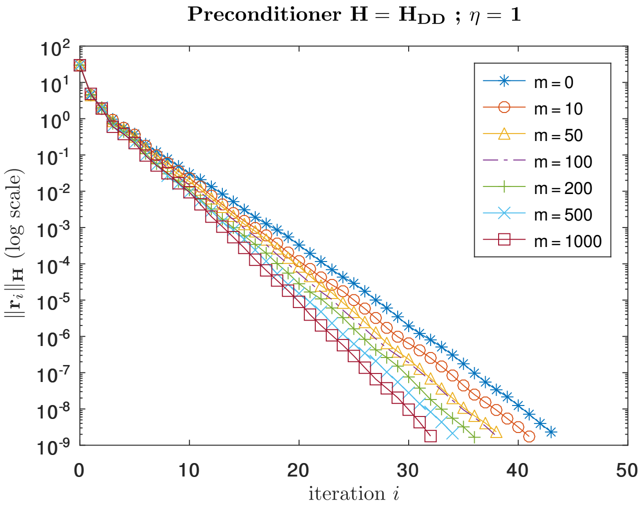
$$\mathbf{N}(\mathbf{A})\mathbf{z}^{(j)} = \lambda_j \mathbf{M}(\mathbf{A})\mathbf{z}^{(j)}$$

in the case $\eta = 1$.



Convergence for a mildly non-symmetric problem ($\eta = 1$)

preconditioned by $\mathbf{H}_{\mathrm{DD}}$ = DD with GenEO

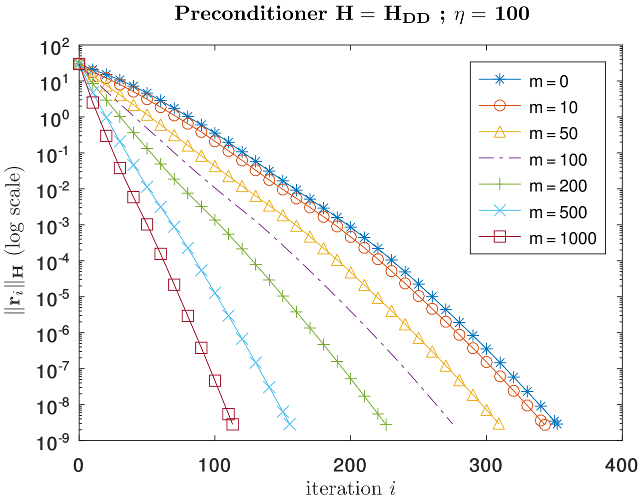


m	iter
0	43
10	41
50	38
100	38
200	36
500	34
1000	32

- $\kappa(\mathbf{H}_{\mathrm{DD}}\mathbf{M}(\mathbf{A})) = 16.241$
- $\rho(\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A})) = 0.646$

Convergence for a non-symmetric problem (η = 100)

preconditioned by $\mathbf{H}_{\mathrm{DD}}$ = DD with GenEO



m	iter
0	352
10	343
50	309
100	275
200	226
500	155
1000	113

- $\kappa(\mathbf{H}_{\mathrm{DD}}\mathbf{M}(\mathbf{A}))$ = 16.241
- $\rho(\mathbf{M}(\mathbf{A})^{-1}\mathbf{N}(\mathbf{A}))$ = 64.6

Comparison of WPD-GMRES and PD-GMRES

We compare

- Weighted GMRES
  i.e, $\mathbf{W} = \mathbf{H}$.
- (Unweighted) GMRES
  i.e, $\mathbf{W} = \mathbf{I}$.

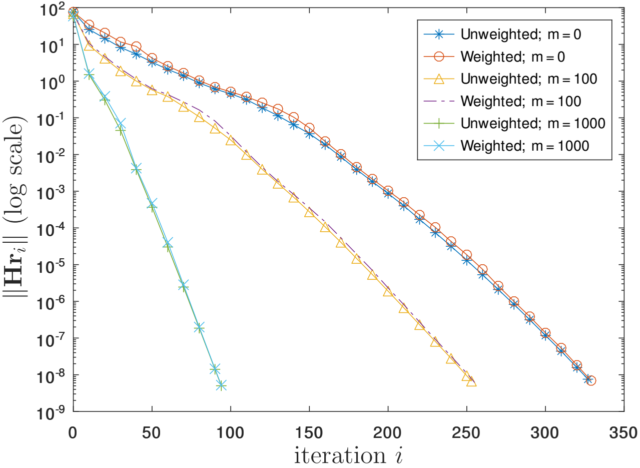
Recall that

the residual is minimized in the **W**-norm.

Stopping criterion

$\|\mathbf{Hr}_i\| / \|\mathbf{Hr}_0\| < 10^{-10}$.

Preconditioner $\mathbf{H} = \mathbf{H}_{\mathrm{DD}}$; $\eta = 100$



Comparison of WPD-GMRES and PD-GMRES

We compare

- Weighted GMRES
  i.e, $\mathbf{W} = \mathbf{H}$.
- (Unweighted) GMRES
  i.e, $\mathbf{W} = \mathbf{I}$.

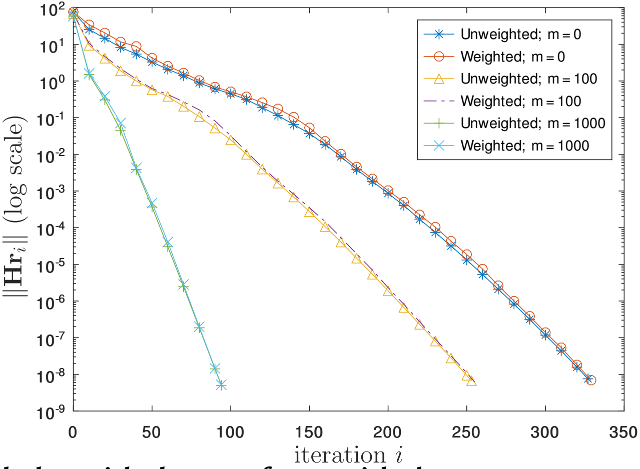
Recall that

the residual is minimized in the $\mathbf{W}$-norm.

Stopping criterion

$\|\mathbf{Hr}_i\| / \|\mathbf{Hr}_0\| < 10^{-10}$.

Preconditioner $\mathbf{H} = \mathbf{H}_{\mathrm{DD}}$; $\eta = 100$



$\rightarrow$ **the weight helps with the proof, not with the convergence.**

Conclusion

Takeaway

Weighted inner product, Preconditioning and Deflation can work together to give a very fast method.

Is FOV the right convergence bound ?

- FOV bound has restriction $0 \notin FOV(\mathbf{A})$.
- FOV bound misses superlinear convergence.
- Other options ?

Introduced a provably scalable algorithm

- for a class of positive definite non-Hermitian problems,
- convergence is h independant for advection-diffusion,
- with a spectral deflation space.