Nested Krylov methods for shifted linear systems

M. Baumann[*,†] and M. B. van Gijzen[†]

[*]Email: M.M.Baumann@tudelft.nl
[†]Delft Institute of Applied Mathematics
Delft University of Technology
Delft, The Netherlands

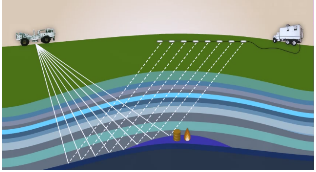
University of Wuppertal - September 25, 2014

Motivation

Full-waveform inversion

PDE-constrained optimization:

$$\min_{\rho(\mathbf{x}),c_p(\mathbf{x}),c_s(\mathbf{x})} \left\| \mathbf{u}_{sim} - \mathbf{u}_{meas} \right\|,$$



where in our application:

- $\mathbf{u}_{sim}$ is the (numerical) solution of the elastic wave equation,
- $\mathbf{u}_{meas}$ is obtained from measurements,
- $\rho(\mathbf{x})$ is the density of the earth layers we are interested in.

The modelling is done in frequency-domain...

Frequency-domain approach:

The time-harmonic elastic wave equation

For **many** (angular) frequencies ω_k, we solve

$$-\omega_k^2 \rho(\mathbf{x})\hat{\mathbf{u}} - \nabla \cdot \sigma(\hat{\mathbf{u}}, c_p, c_s) = \hat{\mathbf{s}}, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^{2,3},$$

together with absorbing or reflecting boundary conditions.

Inverse (discrete) Fourier transform:

$$\mathbf{u}(\mathbf{x}, t) = \sum_k \hat{\mathbf{u}}(\mathbf{x}, \omega_k) e^{i\omega_k t}$$

Motivation

Shifted linear systems

The **discretized** time-harmonic elastic wave equation is quadratic in ω_k:

$$(K + i\omega_k C - \omega_k^2 M)\underline{\hat{\mathbf{u}}} = \underline{\hat{\mathbf{s}}},$$

which can be re-arranged as,

$$\left[\begin{pmatrix} iM^{-1}C & M^{-1}K \\ I & 0 \end{pmatrix} - \omega_k \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\right] \begin{pmatrix} \omega_k \underline{\hat{\mathbf{u}}} \\ \underline{\hat{\mathbf{u}}} \end{pmatrix} = \begin{pmatrix} M^{-1}\underline{\hat{\mathbf{s}}} \\ 0 \end{pmatrix}.$$

The latter is of the form:

$$(A - \omega_k I)\mathbf{x}_k = \mathbf{b}, \quad k = 1, \ldots, N.$$

$\hookrightarrow$ movie

Motivation

Shifted linear systems

The **discretized** time-harmonic elastic wave equation is quadratic in ω_k:

$$(K + i\omega_k C - \omega_k^2 M)\underline{\hat{\mathbf{u}}} = \underline{\hat{\mathbf{s}}},$$

which can be re-arranged as,

$$\left[\begin{pmatrix} iM^{-1}C & M^{-1}K \\ I & 0 \end{pmatrix} - \omega_k \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\right] \begin{pmatrix} \omega_k \underline{\hat{\mathbf{u}}} \\ \underline{\hat{\mathbf{u}}} \end{pmatrix} = \begin{pmatrix} M^{-1}\underline{\hat{\mathbf{s}}} \\ 0 \end{pmatrix}.$$

The latter is of the form:

$$(A - \omega_k I)\mathbf{x}_k = \mathbf{b}, \quad k = 1, ..., N.$$

$\hookrightarrow$ movie

TU Delft

Outline

What's a shifted linear system?

Definition

Shifted linear systems are of the form

$$(A - \omega I)\mathbf{x}^{(\omega)} = \mathbf{b},$$

where $\omega \in \mathbb{C}$ is the *shift*.

For the simultaneous solution, **Krylov methods** are well-suited because of the *shift-invariance* property:

$$\mathcal{K}_m(A, \mathbf{b}) \equiv \text{span}\{\mathbf{b}, A\mathbf{b}, ..., A^{m-1}\mathbf{b}\} = \mathcal{K}_m(A - \omega I, \mathbf{b}).$$

"Proof" (shift-invariance)

For $m = 2$: $\mathcal{K}_2(A, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b}\}$

$$\mathcal{K}_2(A - \omega I, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b} - \omega\mathbf{b}\} = \text{span}\{\mathbf{b}, A\mathbf{b}\}$$

What's a shifted linear system?

Definition

Shifted linear systems are of the form

$$(A - \omega I)\mathbf{x}^{(\omega)} = \mathbf{b},$$

where $\omega \in \mathbb{C}$ is the *shift*.

For the simultaneous solution, **Krylov methods** are well-suited because of the *shift-invariance* property:

$$\mathcal{K}_m(A, \mathbf{b}) \equiv \text{span}\{\mathbf{b}, A\mathbf{b}, ..., A^{m-1}\mathbf{b}\} = \mathcal{K}_m(A - \omega I, \mathbf{b}).$$

"Proof" (shift-invariance)

For $m = 2$: $\mathcal{K}_2(A, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b}\}$

$$\mathcal{K}_2(A - \omega I, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b} - \omega\mathbf{b}\} = \text{span}\{\mathbf{b}, A\mathbf{b}\}$$

What's a shifted linear system?

Definition

Shifted linear systems are of the form

$$(A - \omega I)\mathbf{x}^{(\omega)} = \mathbf{b},$$

where $\omega \in \mathbb{C}$ is the *shift*.

For the simultaneous solution, **Krylov methods** are well-suited because of the *shift-invariance* property:

$$\mathcal{K}_m(A, \mathbf{b}) \equiv \text{span}\{\mathbf{b}, A\mathbf{b}, ..., A^{m-1}\mathbf{b}\} = \mathcal{K}_m(A - \omega I, \mathbf{b}).$$

"Proof" (shift-invariance)

For $m = 2$: $\mathcal{K}_2(A, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b}\}$

$$\mathcal{K}_2(A - \omega I, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b} - \omega\mathbf{b}\} = \text{span}\{\mathbf{b}, A\mathbf{b}\}$$

TUDelft

For example, multi-shift GMRES

After m steps of Arnoldi, we have,

$$AV_m = V_{m+1}\underline{H}_m,$$

and the approximate solution yields:

$$\mathbf{x}_m \approx V_m\mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{H}_m\mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1 \right\|.$$

For shifted systems, we get

$$(A - \omega I)V_m = V_{m+1}(\underline{H}_m - \omega\underline{I}_m),$$

and, therefore,

$$\mathbf{x}_m^{(\omega)} \approx V_m\mathbf{y}_m^{(\omega)}, \quad \text{where } \mathbf{y}_m^{(\omega)} = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{H}_m^{(\omega)}\mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1 \right\|.$$

TUDelft

For example, multi-shift GMRES

After m steps of Arnoldi, we have,

$$AV_m = V_{m+1}\underline{\mathrm{H}}_m,$$

and the approximate solution yields:

$$\mathbf{x}_m \approx V_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{\mathrm{H}}_m \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1 \right\|.$$

For shifted systems, we get

$$(A - \omega I)V_m = V_{m+1}(\underline{\mathrm{H}}_m - \omega \underline{\mathrm{I}}_m),$$

and, therefore,

$$\mathbf{x}_m^{(\omega)} \approx V_m \mathbf{y}_m^{(\omega)}, \quad \text{where } \mathbf{y}_m^{(\omega)} = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{\mathrm{H}}_m^{(\omega)} \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1 \right\|.$$

TU Delft

For example, multi-shift GMRES

After m steps of Arnoldi, we have,

$$AV_m = V_{m+1}\underline{\mathrm{H}}_m,$$

and the approximate solution yields:

$$\mathbf{x}_m \approx V_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{\mathrm{H}}_m \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1 \right\|.$$

For shifted systems, we get

$$(A - \omega I)V_m = V_{m+1}(\underline{\mathrm{H}}_m - \omega \underline{\mathrm{I}}_m),$$

and, therefore,

$$\mathbf{x}_m^{(\omega)} \approx V_m \mathbf{y}_m^{(\omega)}, \quad \text{where } \mathbf{y}_m^{(\omega)} = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{\mathrm{H}}_m^{(\omega)} \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1 \right\|.$$

Preconditioning is a problem

Main disadvantage:

Preconditioners are in general not easy to apply. For

$$(A - \omega I)\mathcal{P}_\omega^{-1}\mathbf{y}^{(\omega)} = \mathbf{b}, \quad \mathcal{P}_\omega \mathbf{x}^{(\omega)} = \mathbf{y}^{(\omega)}$$

it does not hold:

$$\mathcal{K}_m(A\mathcal{P}^{-1}, \mathbf{b}) \neq \mathcal{K}_m(A\mathcal{P}_\omega^{-1} - \omega \mathcal{P}_\omega^{-1}, \mathbf{b}).$$

However, there are ways...

Reference

B. Jegerlehner, *Krylov space solvers for shifted linear systems.* Published online `arXiv:hep-lat/9612014`, 1996.

Preconditioning is a problem

Main disadvantage:

Preconditioners are in general not easy to apply. For

$$(A - \omega I)\mathcal{P}_\omega^{-1}\mathbf{y}^{(\omega)} = \mathbf{b}, \quad \mathcal{P}_\omega \mathbf{x}^{(\omega)} = \mathbf{y}^{(\omega)}$$

it does not hold:

$$\mathcal{K}_m(A\mathcal{P}^{-1}, \mathbf{b}) \neq \mathcal{K}_m(A\mathcal{P}_\omega^{-1} - \omega\mathcal{P}_\omega^{-1}, \mathbf{b}).$$

However, there are ways...

Reference

B. Jegerlehner, *Krylov space solvers for shifted linear systems.* Published online `arXiv:hep-lat/9612014`, 1996.

Preconditioning is a problem

... or has been a problem ?

Short historical overview:

2002 Shift-and-invert preconditioner:

$$\mathcal{P} = (A - \tau I), \quad \tau \approx \{\omega_1, ..., \omega_N\}$$

2007 Many shift-and-invert preconditioners:

$$\mathcal{P}_j = (A - \tau_j I)$$

2013 Polynomial preconditioners:

$$p_n(A) \approx A^{-1}, \quad p_n^{\omega}(A) \approx (A - \omega I)^{-1}$$

2014 Nested Krylov methods

TUDelft

Shift-and-invert preconditioner

Choose $\mathcal{P} \equiv (A - \tau I)$. Then,

$$\begin{aligned}
(A - \omega I)\mathcal{P}_\omega^{-1} &= A\mathcal{P}^{-1} - \eta_\omega I \\
&= A(A - \tau I)^{-1} - \eta_\omega I \\
&= [A - \eta_\omega (A - \tau I)]\,(A - \tau I)^{-1} \\
&= \left[A + \frac{\eta_\omega \tau}{1 - \eta_\omega} I\right](1 - \eta_\omega)(A - \tau I)^{-1}.
\end{aligned}$$

From the fit $\omega = -\frac{\eta_\omega \tau}{1-\eta_\omega}$, we conclude

$$\eta_\omega = \frac{\omega}{\omega - \tau}, \quad \mathcal{P}_\omega = \frac{1}{1 - \eta_\omega}\mathcal{P} = \frac{\tau - \omega}{\tau}\mathcal{P}.$$

TUDelft

Polynomial preconditioners - Theory

Suppose we have found,

$$p_n(A) \equiv \sum_{i=0}^{n} \alpha_i A^i \approx A^{-1}$$

Question: Can we find $p_n^\omega(A) \equiv \sum_{i=0}^{n} \alpha_i^\omega A^i$ such that

$$(A - \omega I) p_n^\omega(A) = A p_n(A) - \eta_\omega I \quad ?$$

Reference

M. I. Ahmad, D. B. Szyld, and M. B. van Gijzen, *Preconditioned multishift BiCG for $\mathcal{H}_2$-optimal model reduction.* Technical report, 2013.

Polynomial preconditioners - Theory

Suppose we have found,

$$p_n(A) \equiv \sum_{i=0}^{n} \alpha_i A^i \approx A^{-1}$$

Question: Can we find $p_n^\omega(A) \equiv \sum_{i=0}^{n} \alpha_i^\omega A^i$ such that

$$(A - \omega I)p_n^\omega(A) = Ap_n(A) - \eta_\omega I \quad ?$$

Reference

M. I. Ahmad, D. B. Szyld, and M. B. van Gijzen, *Preconditioned multishift BiCG for $\mathcal{H}_2$-optimal model reduction.* Technical report, 2013.

TU Delft

Polynomial preconditioners - Theory

From

$$(A - \omega I) p_n^{\omega}(A) = A p_n(A) - \eta_{\omega} I$$

we get

$$\sum_{i=0}^{n} \alpha_i^{\omega} A^{i+1} - \sum_{i=0}^{n} \omega \alpha_i^{\omega} A^{i} - \sum_{i=0}^{n} \alpha_i A^{i+1} + \eta_{\omega} I = 0$$

The latter can be solved to:

$$\begin{aligned} \alpha_n^{\omega} &= \alpha_n \\ \alpha_{i-1}^{\omega} &= \alpha_{i-1} + \omega \alpha_i^{\omega}, \quad \text{for } i = n, \ldots, 1 \\ \eta_{\omega} &= \omega \alpha_0^{\omega} \end{aligned}$$

Polynomial preconditioners - In practice

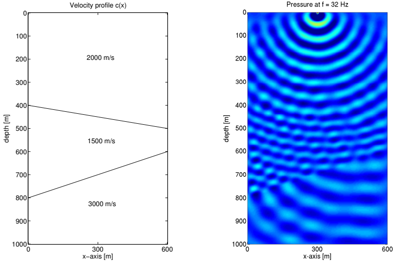
What goes wrong in practice?

- Use Chebyshev polynomials for $p_n(A) \approx A^{-1}$.
  - Based on ellipse that surrounds the spectrum of A,
  - Does not work for indefinite matrix A.
- Instead, approximate the shift-and-invert preconditioner $p_n(A) \approx (A - \tau I)^{-1}$, i.e. *shift the spectrum*.
- For Helmholtz, this resembles an approximate shifted Laplace preconditioner.

Polynomial preconditioners - Example (1/2)

Acoustic wave propagation (Helmholtz equation):

$$-\Delta p - \left(\frac{2\pi f_k}{c(\mathbf{x})}\right) p = s$$
$$s = \delta(x_1 - 300, x_2)$$



We prescribe absorbing boundary conditions (Sommerfeld conditions).

Polynomial preconditioners - Example (2/2)

We consider 5 different meshes:

Grid	$h[m]$	gridpoints	$f_k[Hz]$	No prec.	Exact prec.	PP $n=3$
1	100	77	1	179	77	87
2	100/2	273	1,2	457	156	194
3	100/4	1025	1,2,4	1160	319	457
4	100/8	3969	1,2,4,8	4087	788	1115
5	100/16	15617	1,2,4,8,16	9994	1830	2429

- MS-QMRIDR(8), Seed: $\tau = (1 - 8i)2\pi f_{max}$
- Note: we have to use a shift with a large imaginary part to obtain a converging Chebyshev polynomial.

[Ongoing Research]

Nested multi-shift Krylov methods

Methodology:

- We have learned: Polynomial preconditioners exist
- Question: Can we use a Krylov polynomial?

Nested multi-shift Krylov methods:

- Use an inner multi-shift Krylov method as preconditioner.
- For inner method, require collinear residuals [$\mathbf{r}_j^{(\omega)} = \gamma \mathbf{r}_j$]. This is the case for:
  - multi-shift GMRES [1998]
  - multi-shift FOM [2003]
  - multi-shift BiCG [2003]
  - multi-shift IDR(s) [new!]
- Using γ, we can preserve the shift-invariance in the outer Krylov iteration.

Nested multi-shift Krylov methods

Methodology:

- We have learned: Polynomial preconditioners exist
- Question: Can we use a Krylov polynomial?

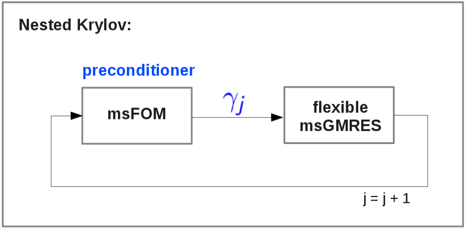
Nested multi-shift Krylov methods:

- Use an inner multi-shift Krylov method as preconditioner.
- For inner method, require collinear residuals $[\mathbf{r}_j^{(\omega)} = \gamma \mathbf{r}_j]$. This is the case for:
  - multi-shift GMRES [1998]
  - multi-shift FOM [2003]
  - multi-shift BiCG [2003]
  - multi-shift IDR(s) [new!]
- Using γ, we can preserve the shift-invariance in the outer Krylov iteration.

TUDelft

Nested multi-shift Krylov methods

Overview of one possible combination:



Multi-shift FOM as inner method

Classical result: In FOM, the residuals are

$$\mathbf{r}_j = \mathbf{b} - A\mathbf{x}_j = \dots = -h_{j+1,j}\mathbf{e}_j^T\mathbf{y}_j\mathbf{v}_{j+1}.$$

Thus, for the shifted residuals it holds:

$$\mathbf{r}_j^{(\omega)} = \mathbf{b} - (A - \omega I)\mathbf{x}_j^{(\omega)} = \dots = -h_{j+1,j}^{(\omega)}\mathbf{e}_j^T\mathbf{y}_j^{(\omega)}\mathbf{v}_{j+1},$$

which gives $\gamma = y_j^{(\omega)}/y_j$.

Reference

V. Simoncini, *Restarted full orthogonalization method for shifted linear systems.* BIT Numerical Mathematics, 43 (2003).

Flexible multi-shift GMRES as outer method

Use flexible GMRES in the outer loop,

$$(A - \omega I)\widehat{V}_m = V_{m+1}\underline{\mathrm{H}}_m^{(\omega)},$$

where one column yields

$$(A - \omega I)\underbrace{\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j}_{\text{inner loop}} = V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}, \quad 1 \leq j \leq m.$$

The "inner loop" is the truncated solution of $(A - \omega I)$ with right-hand side $\mathbf{v}_j$ using `msFOM`.

Flexible multi-shift GMRES as outer method

Use flexible GMRES in the outer loop,

$$(A - \omega I)\widehat{V}_m = V_{m+1}\underline{\mathrm{H}}_m^{(\omega)},$$

where one column yields

$$(A - \omega I)\underbrace{\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j}_{\text{inner loop}} = V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}, \quad 1 \leq j \leq m.$$

The "inner loop" is the truncated solution of $(A - \omega I)$ with right-hand side $\mathbf{v}_j$ using `msFOM`.

Flexible multi-shift GMRES

The inner residuals are:

$$\mathbf{r}_j^{(\omega)} = \mathbf{v}_j - (A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j,$$
$$\mathbf{r}_j = \mathbf{v}_j - A\mathcal{P}_j^{-1}\mathbf{v}_j,$$

Imposing $\mathbf{r}_j^{(\omega)} = \gamma \mathbf{r}_j$ yields:

$$(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j = \gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j \qquad (*)$$

Note that the right-hand side in $(*)$ is a preconditioned shifted system!

Flexible multi-shift GMRES

The inner residuals are:

$$\mathbf{r}_j^{(\omega)} = \mathbf{v}_j - (A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j,$$
$$\mathbf{r}_j = \mathbf{v}_j - A\mathcal{P}_j^{-1}\mathbf{v}_j,$$

Imposing $\mathbf{r}_j^{(\omega)} = \gamma \mathbf{r}_j$ yields:

$$(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j = \gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j \qquad (*)$$

Note that the right-hand side in $(*)$ is a preconditioned shifted system!

Flexible multi-shift GMRES

Altogether,

$$
\begin{aligned}
(A-\omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma-1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}(\gamma-1)\,\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma-1)\,\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}
\end{aligned}
$$

which yields:

$$\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

Flexible multi-shift GMRES

Altogether,

$$
\begin{aligned}
(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}(\gamma - 1)\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma - 1)\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}
\end{aligned}
$$

which yields:

$$
\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,
$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

Flexible multi-shift GMRES

Altogether,

$$
\begin{aligned}
(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}(\gamma - 1)\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma - 1)\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}
\end{aligned}
$$

which yields:

$$
\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,
$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

Flexible multi-shift GMRES

Altogether,

$$
\begin{aligned}
(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}\,(\gamma - 1)\,\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma - 1)\,\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}
\end{aligned}
$$

which yields:

$$
\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,
$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

Flexible multi-shift GMRES

Altogether,

$$\begin{aligned}(A-\omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}\\ \gamma A\mathcal{P}_j^{-1}\mathbf{v}_j-(\gamma-1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}\\ \gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}(\gamma-1)\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}\\ V_{m+1}\left(\gamma\underline{\mathbf{h}}_j-(\gamma-1)\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}\end{aligned}$$

which yields:

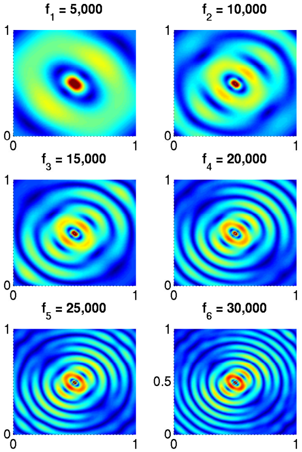
$$\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

A first example - The setting

Test case from literature:

- $\Omega = [0, 1] \times [0, 1]$
- $h = 0.01$ implying
  $n = 10.201$ grid points
- system size:
  $4n = 40.804$
- $N = 6$ frequencies
- point source at center

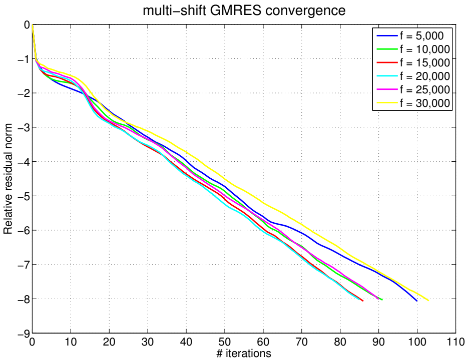


Reference

- T. Airaksinen, A. Pennanen, and J. Toivanen, *A damping preconditioner for time-harmonic wave equations in fluid and elastic material.* Journal of Computational Physics, 2009.

A first example - Convergence behavior (1/2)

Preconditioned **multi-shift GMRES**:

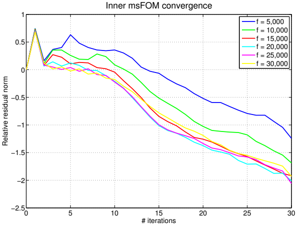


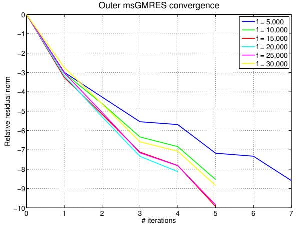
We observe:

- simultaneous solve
- CPU time: 17.71s

A first example - Convergence behavior (2/2)

Preconditioned **nested FOM-FGMRES**:





We observe:

- 30 inner iterations
- truncate when inner residual norm ~ 0.1
- very few outer iterations
- CPU time: 9.62s

TUDelft

A first example - More nested methods

We were running the same setting with different (nested) multi-shift Krylov methods:

	multi-shift Krylov methods			
	msGMRES	rest_msGMRES	QMRIDR(4)	msIDR(4)
# inner iterations	-	20	-	-
# outer iterations	103	7	136	134
seed shift τ	0.7-0.7i	0.7-0.7i	0.7-0.7i	0.7-0.7i
CPU time	17.71s	6.13s	22.35s	22.58s
	nested multi-shift Krylov methods			
	FOM-FGMRES	IDR(4)-FGMRES	FOM-FQMRIDR(4)	IDR(4)-FQMRIDR(4)
# inner iterations	30	25	30	30
# outer iterations	7	9	5	15
seed shift τ	0.7-0.7i	0.7-0.7i	0.7-0.7i	0.7-0.7i
CPU time	9.62s	32.99s	8.14s	58.36s

Summary

✓ Nested Krylov methods for $A\mathbf{x} = \mathbf{b}$ are widely used
$\hookrightarrow$ extension to shifted linear systems is possible

✓ Multiple combinations of inner-outer methods possible, e.g. `FOM-FGMRES`, `IDR-FQMRIDR`, ...

✓ The shift-and-invert preconditioner (or the polynomial preconditioner) can be applied on top

✗ Future work: recylcing, deflation, ...

Thank you for your attention!

Further reading:

M. Baumann and M. B. van Gijzen. *Nested Krylov methods for shifted linear systems.* DIAM technical report 14-01, 2014.

Further coding:

https://bitbucket.org/ManuelMBaumann/nestedkrylov

IDR(s) in a nutshell

IDR(s) is a Krylov subspace method that enforces the residuals $\mathbf{r}_n$ to be,

$$\mathcal{G}_{j+1} \ni \mathbf{r}_{n+1} = (I - \mu_{j+1} A)\mathbf{v}_n, \quad \text{with } \mathbf{v}_n \in \mathcal{G}_j \cap \mathcal{P}^{\perp},$$

where $\mu_{j+1} \in \mathbb{C} \setminus \{0\}$ and $\mathcal{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_s]$ are chosen freely.

The IDR theorem states:

1. $\mathcal{G}_{j+1} \subset \mathcal{G}_j$ for all $j \geq 0$,
2. $\mathcal{G}_j = \{\mathbf{0}\}$ for some $j \leq N$.

Reference

- P. Sonneveld, M. B. van Gijzen, *IDR(s): A family of simple and fast algorithms for solving large nonsymmetric systems of linear equations.* SIAM J. Sci. Comput., 2008.

IDR(s) in a nutshell

IDR(s) is a Krylov subspace method that enforces the residuals $\mathbf{r}_n$ to be,

$$\mathcal{G}_{j+1} \ni \mathbf{r}_{n+1} = (I - \mu_{j+1}A)\mathbf{v}_n, \quad \text{with } \mathbf{v}_n \in \mathcal{G}_j \cap \mathcal{P}^\perp,$$

where $\mu_{j+1} \in \mathbb{C} \setminus \{0\}$ and $\mathcal{P} = [\mathbf{p}_1, ..., \mathbf{p}_s]$ are chosen freely.

The IDR theorem states:

1. $\mathcal{G}_{j+1} \subset \mathcal{G}_j$ for all $j \geq 0$,
2. $\mathcal{G}_j = \{\mathbf{0}\}$ for some $j \leq N$.

Reference

- P. Sonneveld, M. B. van Gijzen, *IDR(s): A family of simple and fast algorithms for solving large nonsymmetric systems of linear equations.* SIAM J. Sci. Comput., 2008.

TU Delft

IDR(s) with collinear residuals

Notation: Let's call $\widehat{A} \equiv (A - \omega I)$ and $\hat{\mathbf{r}}_n, \hat{\mu}_{j+1}, \hat{\mathbf{v}}_n, \ldots$

Assuming, we have $\hat{\mathbf{v}}_n = \alpha_n \mathbf{v}_n$ (not trivial!), then we want:

$$
\begin{aligned}
\gamma_{n+1}\mathbf{r}_{n+1} &= \hat{\mathbf{r}}_{n+1} \\
\gamma_{n+1}\left(I - \mu_{j+1}A\right)\mathbf{v}_n &= \left(I - \hat{\mu}_{j+1}\widehat{A}\right)\hat{\mathbf{v}}_n \\
\gamma_{n+1}\left(I - \mu_{j+1}A\right)\mathbf{v}_n &= \left(I - \hat{\mu}_{j+1}(A - \omega I)\right)\alpha_n\mathbf{v}_n \\
\gamma_{n+1}\mathbf{v}_n - \gamma_{n+1}\mu_{j+1}A\mathbf{v}_n &= (\alpha_n + \alpha_n\hat{\mu}_{j+1}\omega)\mathbf{v}_n - \alpha_n\hat{\mu}_{j+1}A\mathbf{v}_n
\end{aligned}
$$

Choose $\hat{\mu}_{j+1}$ and γ_{n+1} such that the two terms match:

$$
\hat{\mu}_{j+1} = \frac{\mu_{j+1}}{1 - \omega\mu_{j+1}}, \quad \gamma_{n+1} = \frac{\alpha_n}{1 - \omega\mu_{j+1}}.
$$

TUDelft

IDR(s) with collinear residuals

Notation: Let's call $\widehat{A} \equiv (A - \omega I)$ and $\hat{\mathbf{r}}_n, \hat{\mu}_{j+1}, \hat{\mathbf{v}}_n, \ldots$

Assuming, we have $\hat{\mathbf{v}}_n = \alpha_n \mathbf{v}_n$ (not trivial!), then we want:

$$\begin{aligned}
\gamma_{n+1}\mathbf{r}_{n+1} &= \hat{\mathbf{r}}_{n+1} \\
\gamma_{n+1}\left(I - \mu_{j+1}A\right)\mathbf{v}_n &= \left(I - \hat{\mu}_{j+1}\widehat{A}\right)\hat{\mathbf{v}}_n \\
\gamma_{n+1}\left(I - \mu_{j+1}A\right)\mathbf{v}_n &= \left(I - \hat{\mu}_{j+1}(A - \omega I)\right)\alpha_n\mathbf{v}_n \\
\gamma_{n+1}\mathbf{v}_n - \gamma_{n+1}\mu_{j+1}A\mathbf{v}_n &= (\alpha_n + \alpha_n\hat{\mu}_{j+1}\omega)\mathbf{v}_n - \alpha_n\hat{\mu}_{j+1}A\mathbf{v}_n
\end{aligned}$$

Choose $\hat{\mu}_{j+1}$ and γ_{n+1} such that the two terms match:

$$\hat{\mu}_{j+1} = \frac{\mu_{j+1}}{1 - \omega\mu_{j+1}}, \quad \gamma_{n+1} = \frac{\alpha_n}{1 - \omega\mu_{j+1}}.$$