Argonne National Laboratory, May 1997

expanding efficiently effective search subspaces for Jacobi-Davidson type of methods

Gerard Sleijpen
Diederik Fokkema
Henk van der Vorst

Universiteit Utrecht
Department of Mathematics

http://www.math.ruu.nl/people/sleijpen

$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{B}\,\mathbf{x}$

JDQZ

Select $\alpha, \beta \in \mathbb{C}$. Consider $\mathbf{C} := \alpha\,\mathbf{A} + \beta\,\mathbf{B}$.

- Solve $\mathbf{W}^*\,\mathbf{A}\,\mathbf{V}\,y - \theta\,\mathbf{W}^*\,\mathbf{B}\,\mathbf{V}\,y = 0$.
  $\mathbf{u} = \mathbf{V}\,y$, $\mathbf{r} := \mathbf{A}\,\mathbf{u} - \theta\,\mathbf{B}\,\mathbf{u}$, $\mathbf{w} = \mathbf{C}\,\mathbf{u}$.
- Solve (approximately) the Cor. Eq.
  $$\left(\mathbf{I} - \frac{\mathbf{w}\,\mathbf{w}^*}{\mathbf{w}^*\mathbf{w}}\right)(\mathbf{A} - \theta\,\mathbf{B})\left(\mathbf{I} - \frac{\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{z} = -\mathbf{r}.$$
- Expand search – and test subspace:
  $\mathbf{V} = \text{modGS}([\mathbf{V}\,|\,\mathbf{z}])$, $\mathbf{W} = \text{modGS}([\mathbf{W}\,|\,\mathbf{C}\,\mathbf{z}])$

Th. *Exact solve* $\Rightarrow$ *quadratic convergence.*

Program

- Incorporating preconditioners
  Example: MRILU
- Selecting the mixture $\alpha\,\mathbf{A} + \beta\,\mathbf{B}$.

Solve approximately

$$\left(\mathbf{I} - \frac{\mathbf{w}\,\mathbf{w}^*}{\mathbf{w}^*\mathbf{w}}\right)(\mathbf{A} - \theta\mathbf{B})\left(\mathbf{I} - \frac{\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{z} = -\mathbf{r}.$$

Any preconditioned it. method can be used.

Preconditioning: $\mathbf{M} \approx \mathbf{A} - \theta\,\mathbf{B}$

Preconditioning equation:

$$\left(\mathbf{I} - \frac{\mathbf{w}\,\mathbf{w}^*}{\mathbf{w}^*\mathbf{w}}\right)\mathbf{M}\left(\mathbf{I} - \frac{\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{t} = \mathbf{s}$$

Solution:

$$\mathbf{M}\,\mathbf{t} = \mathbf{s} + \mathbf{w}\,\alpha \quad \text{such that} \quad \mathbf{t} \perp \mathbf{u},$$

or

$$\mathbf{t} = \mathbf{M}^{-1}\mathbf{s} + \mathbf{M}^{-1}\mathbf{w}\,\alpha \quad \& \quad \alpha = (\mathbf{u}^*\mathbf{M}^{-1}\mathbf{w})^{-1}\mathbf{u}^*\mathbf{M}^{-1}\mathbf{s}.$$

Solve approximately

$$\left(\mathbf{I}-\frac{\mathbf{w}\,\mathbf{w}^*}{\mathbf{w}^*\mathbf{w}}\right)(\mathbf{A}-\theta\mathbf{B})\left(\mathbf{I}-\frac{\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{z}=-\mathbf{r}.$$

$\mathbf{M}\approx\mathbf{A}-\theta\,\mathbf{B}$, $\quad\left(\mathbf{I}-\frac{\mathbf{w}\,\mathbf{w}^*}{\mathbf{w}^*\mathbf{w}}\right)\mathbf{M}\left(\mathbf{I}-\frac{\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{t}=\mathbf{s}.$

Explicit left preconditioning:

- The correction equation:

$$\left(\mathbf{I}-\frac{\mathbf{v}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{v}}\right)\mathbf{M}^{-1}(\mathbf{A}-\theta\,\mathbf{B})\left(\mathbf{I}-\frac{\mathbf{v}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{v}}\right)\mathbf{z}=-\mathbf{r}'.$$

where

$$\mathbf{M}\,\mathbf{v}=\mathbf{w}=\mathbf{B}\,\mathbf{u}\quad\&\quad\mathbf{r}'=\left(\mathbf{I}-\frac{\mathbf{v}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{v}}\right)\mathbf{r}.$$

$\mathbf{M}\approx\mathbf{A}-\theta\,\mathbf{I}$, $\quad\left(\mathbf{I}-\frac{\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{M}\left(\mathbf{I}-\frac{\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{z}=-\mathbf{r}$

What if $\mathbf{M}\tilde{\mathbf{x}}=\mathbf{0}$ & $\tilde{\mathbf{x}}\approx\mathbf{x}$?

Compute $\mathbf{y}$ such that $\mathbf{y}^*\mathbf{M}=\mathbf{0}$.

Find $\mathbf{p}$ such that

- $\mathbf{M}:\mathbf{p}^\perp\to\mathbf{y}^\perp$ well-cond., $\quad\mathbf{p}^*\tilde{\mathbf{x}}\gg 0$, &
- solving $\mathbf{t}\perp\mathbf{p}$ from $\mathbf{M}\mathbf{t}=\mathbf{s}\perp\mathbf{y}$ is easy.

Solve $\mathbf{t}\perp\mathbf{p}$ such that

$$\mathbf{M}\mathbf{t}=-\mathbf{r}+\alpha\,\mathbf{u}\perp\mathbf{y}$$

Then $\qquad\mathbf{z}=\mathbf{t}+\beta\,\tilde{\mathbf{x}}\perp\mathbf{u}$

$\mathbf{M}\tilde{\mathbf{x}}=\mathbf{0}$, $\qquad\mathbf{y}^*\mathbf{M}=\mathbf{0}$,

Find $\mathbf{p}$ such that

- $\mathbf{M}:\mathbf{p}^\perp\to\mathbf{y}^\perp$ well-cond. &
- solving $\mathbf{t}\perp\mathbf{p}$ from $\mathbf{M}\mathbf{t}=\mathbf{s}\perp\mathbf{y}$ is easy.

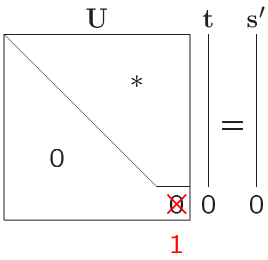
Example (*Complete LU decomposition*)

$\mathbf{M}=\mathbf{L}\mathbf{U}$ with $e_n^*\mathbf{U}=0$ & $\mathrm{diag}(\mathbf{L})=\mathbf{I}$.

Then $\mathbf{y}^*\mathbf{L}=e_n^*$

$\mathbf{U}\mathbf{t}=\mathbf{s}':=\mathbf{L}^{-1}\mathbf{s}$

$$\begin{aligned}e_n^*\,\mathbf{s}'&=e_n^*\,\mathbf{L}^{-1}\mathbf{s}\\&=\mathbf{y}^*\mathbf{s}=0\end{aligned}$$



Take $\mathbf{p}=e_n$.

Put $\mathbf{U}_{nn}=1$ and solve $\mathbf{U}\mathbf{t}=\mathbf{L}^{-1}\mathbf{s}$

$\mathbf{A}\mathbf{x}=\mathbf{b}$

[Botta & Wubs 97]

MRILU

Reorder & partition

$$\mathbf{A}^{(0)}=\mathbf{A}=\begin{bmatrix}\mathbf{A}_1&\mathbf{E}\\\mathbf{F}&\mathbf{A}_2\end{bmatrix}$$

such that $\quad\mathbf{A}_1\approx\mathbf{D}:=\mathrm{diag}(\mathbf{A}_1)$

Take $\quad\mathbf{M}^{(0)}=\mathbf{M}=\begin{bmatrix}\mathbf{D}&\mathbf{E}\\\mathbf{F}&\mathbf{A}_2\end{bmatrix}$

Use Schur complement $\widetilde{\mathbf{A}}_2=\mathbf{A}_2-\mathbf{F}\mathbf{D}^{-1}\mathbf{E}$

for solving $\quad\mathbf{M}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}b\\c\end{bmatrix}$:

- solve (approximately) $\widetilde{\mathbf{A}}_2\,y=c-\mathbf{G}\mathbf{D}^{-1}b$
- solve exactly $\qquad\mathbf{D}\,x=b-\mathbf{E}\,y$

Can be repeated with
$\mathbf{A}$ replaced by $\mathbf{A}^{(1)}=\widetilde{\mathbf{A}}_2-\mathrm{small}_\varepsilon(\widetilde{\mathbf{A}}_2)$, etc.

$(\mathbf{A}-\lambda\mathbf{B})\mathbf{x}=\mathbf{0}$, $\mathbf{B}$ diagonal

MRILU

[Wubs 97]

Reorder & partition

$$\mathbf{A}^{(0)}=\mathbf{A}=\begin{bmatrix}\mathbf{A}_1 & \mathbf{E}\\ \mathbf{F} & \mathbf{A}_2\end{bmatrix},\qquad \mathbf{B}^{(0)}=\mathbf{B}=\begin{bmatrix}\mathbf{B}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{B}_2\end{bmatrix}$$

such that $\mathbf{A}_1\approx\mathbf{D}:=\mathrm{diag}(\mathbf{A}_1)$

Take $\mathbf{M}^{(0)}=\mathbf{M}=\begin{bmatrix}\mathbf{D}-\lambda\mathbf{B}_1 & \mathbf{E}\\ \mathbf{F} & \mathbf{A}_2-\lambda\mathbf{B}_2\end{bmatrix}$

Use Schur complement

$$\widetilde{\mathbf{M}}_2=\mathbf{A}_2-\lambda\mathbf{B}_2-\mathbf{F}(\mathbf{D}-\lambda\mathbf{B}_1)^{-1}\mathbf{E}$$

to solve (appr) λ, x & y from $\mathbf{M}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$:

- ∘ solve (appr.) λ & y from $\widetilde{\mathbf{M}}_2\,y=0$
- ∘ solve exactly x from $(\mathbf{D}-\lambda\mathbf{B}_2)\,x=-\mathbf{E}\,y$

$\widetilde{\mathbf{M}}_2=\mathbf{A}_2-\lambda\mathbf{B}_2-\mathbf{F}(\mathbf{D}-\lambda\mathbf{B}_1)^{-1}\mathbf{E}$

∘ solve (appr.) λ & y from $\widetilde{\mathbf{M}}_2\,y=0$

$$\begin{aligned}\widetilde{\mathbf{M}}_2 &= \mathbf{A}_2-\lambda\mathbf{B}_2-\mathbf{F}(\mathbf{D}-\lambda\mathbf{B}_1)^{-1}\mathbf{E}\\ &\qquad\qquad \textit{Neumann series}\\ &\approx \mathbf{A}_2-\lambda\mathbf{B}_2-\mathbf{F}(\mathbf{D}^{-1}+\lambda\mathbf{D}^{-1}\mathbf{B}_1\mathbf{D}^{-1})\mathbf{E}\\ &= (\mathbf{A}_2-\mathbf{F}\mathbf{D}^{-1}\mathbf{E})-\lambda(\mathbf{B}_2-\mathbf{F}\mathbf{D}^{-1}\mathbf{B}_1\mathbf{D}^{-1}\mathbf{E})\\ &= \qquad \widetilde{\mathbf{A}}_2 \qquad - \qquad \lambda \qquad \tilde{\mathbf{B}}_2\end{aligned}$$

Use $\widetilde{\widetilde{\mathbf{M}}}_2=\widetilde{\mathbf{A}}_2-\lambda\tilde{\mathbf{B}}_2$

∘ solve (approximately) $(\widetilde{\mathbf{A}}_2-\lambda\tilde{\mathbf{B}}_2)y=0$.

Exact solve: *generalized eigenproblem*

Approximate solve: repeat process with

$\mathbf{B}^{(1)}$ = diagonalized $\tilde{\mathbf{B}}_2$ by lumping *(smooth eigenv.)*

$\mathbf{A}^{(1)}=\widetilde{\mathbf{A}}_2-\mathrm{small}_\varepsilon(\widetilde{\mathbf{A}}_2)$

etc.

With exact solve after ℓ steps we have

- ∘ an approximate eigenvalue $\tilde{\lambda}$
- ∘ an approximate eigenvector $\tilde{x}$
- ∘ an $\mathbf{M}=\mathbf{M}^{(\ell)}$ in factorized form

with the properties

- $\mathbf{M}\tilde{\mathbf{x}}=\mathbf{0}$
- $\mathbf{y}$ with $\mathbf{y}^*\mathbf{M}=\mathbf{0}$ is easy to compute
- $\mathbf{M}\mathbf{t}=\mathbf{s}$ for $\mathbf{s}\perp\mathbf{y}$

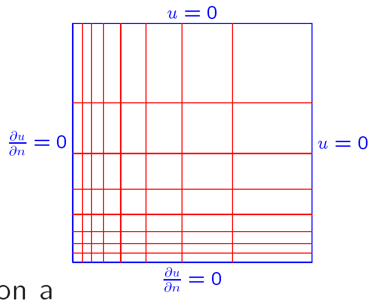
can efficiently be solved:
use complete LU-decomposition for $\widetilde{\mathbf{A}}_2^{(\ell)}-\tilde{\lambda}\tilde{\mathbf{B}}_2^{(\ell)}$

Property "$\mathbf{s}\perp\mathbf{y}$" is propagated to lower blocks

- Factorization can be used for other approximate eigenpairs:

compute other solutions $(\lambda, y^{(\ell)})$ of

$$(\widetilde{\mathbf{A}}_2^{(\ell)}-\lambda\tilde{\mathbf{B}}_2^{(\ell)})y^{(\ell)}=0$$



Numerical Example

2d Laplace operator on a streched $N\times N$ regular grid.

Dirichlet on east & north boundary
Neumann on west and south boundary

Preconditioned for 3 smallest eigenvalues
$\epsilon=10^{-8}$
(3 eigenvalues detected at machine precision)

N	Cond($\mathbf{A}$)	fill in	levels	# JD steps
7	$2.4\,10^5$	5.6	3	7
15	$7.4\,10^6$	7.3	8	10
31	$7.3\,10^7$	7.9	12	12
63	$6.1\,10^8$	8.4	17	13

$\mathbf{w} = \alpha\,\mathbf{A}\,\mathbf{u} + \beta\,\mathbf{B}\,\mathbf{u}$

Choosing the mixture of $\mathbf{A}$ and $\mathbf{B}$

First approach:
aim for stable approximate "left Schur vector"

Inspiration:

$\mathbf{B} = \mathbf{I}$: choose $\quad \mathbf{w} = \mathbf{u} = 0\,\mathbf{A}\,\mathbf{u} + 1\,\mathbf{B}\,\mathbf{u}$.

$\mathbf{A} = \mathbf{I}$: choose $\quad \mathbf{w} = \mathbf{u} = 1\,\mathbf{A}\,\mathbf{u} + 0\,\mathbf{B}\,\mathbf{u}$.

$\mathbf{w} = 1\,\mathbf{A}\,\mathbf{u} + 0\,\mathbf{B}\,\mathbf{u}$ is "instable" choice

if $\mathbf{B} = \mathbf{I}$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \approx \mathbf{0}$

Select α, β such that

— $|\alpha|^2 + \beta^2 = 1$, $\beta \geq 0$, and

— $\|\alpha\,\mathbf{A}\,\mathbf{u} + \beta\,\mathbf{B}\,\mathbf{u}\|$ approximately maximal:

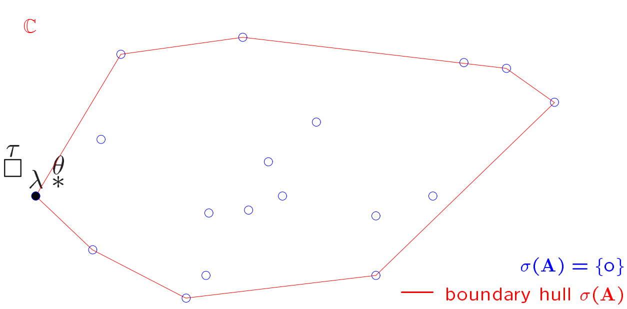
$$\mathbf{C} = \frac{\bar{\tau}}{\sqrt{1+|\tau|^2}}\mathbf{A} + \frac{1}{\sqrt{1+|\tau|^2}}\mathbf{B}$$

Selecting Ritz pairs

Suppose $\mathbf{A}$ normal & $\mathbf{B} = \mathbf{I}$.

$$\theta = \sum_i |\gamma_i|^2 \lambda_i \quad \text{with} \quad \mathbf{u} = \sum_i \gamma_i \mathbf{x}_i$$

Suppose τ outside hull spectrum:



If $\theta \approx \tau$ then wanted component large & selecting extremal Ritz value is safe.

What if τ is in the interior of the hull of the spectrum?

Second approach: avoid misselection.

$$(\mathbf{A} - \lambda\,\mathbf{I})\mathbf{x} = \mathbf{0} \quad \Leftrightarrow \quad (\mathbf{A} - \tau\,\mathbf{I})^{-1}\mathbf{x} = \frac{1}{\lambda - \tau}\mathbf{x}:$$

if $\lambda \approx \tau$ then $1/(\lambda - \tau)$ extremal for $(\mathbf{A} - \tau\,\mathbf{I})^{-1}$.

Avoid inversion: use Harmonic Ritz values

Galerkin for $(\mathbf{A} - \tau\,\mathbf{I})^{-1}$, with

search space = test space = span$((\mathbf{A} - \tau\,\mathbf{I})\mathbf{V})$,

is equivalently to Petrov for $\mathbf{A}$, with

search space $\mathbf{V}$ & test space $\mathbf{W} = (\mathbf{A} - \tau\,\mathbf{I})\mathbf{V}$.

Harmonic Ritz value := this Petrov value

For generalized eigenproblems, take

$$\text{span}(\mathbf{W}) = \text{span}((\mathbf{A} - \tau\,\mathbf{B})\mathbf{V}),$$

$$\mathbf{C} = \frac{1}{\sqrt{1+|\tau|^2}}\mathbf{A} - \frac{\tau}{\sqrt{1+|\tau|^2}}\mathbf{B}.$$

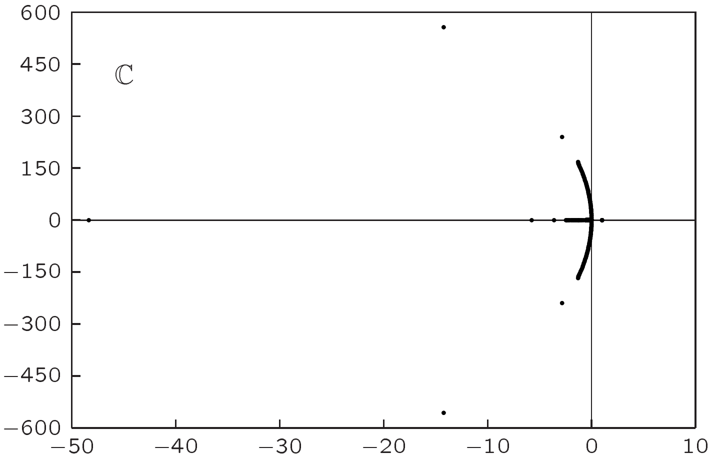
The problem

$$\boxed{\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{B}\,\mathbf{x}}$$

For MHD-problems:

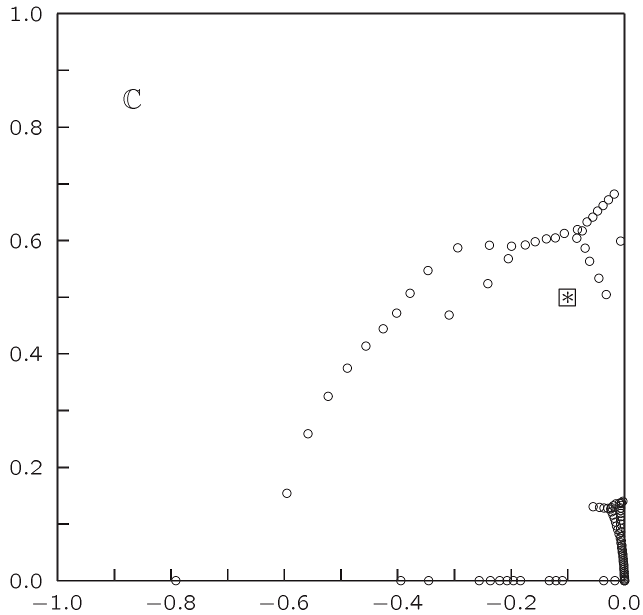
- $\mathbf{A}$ is non Hermitian
- $\mathbf{B}$ Hermitian, positive definite
- $\mathbf{A}$, $\mathbf{B}$ block tridiagonal — 100 - 2000 blocks of size 80 -320
- interior eigenvalues
- very large order — 8 000 - 640 000
- inversion unpractical

Spectrum MHD test problem



Entire spectrum MHD test problem ($n = 416$).
Correct eigenvalues (.) are shown.

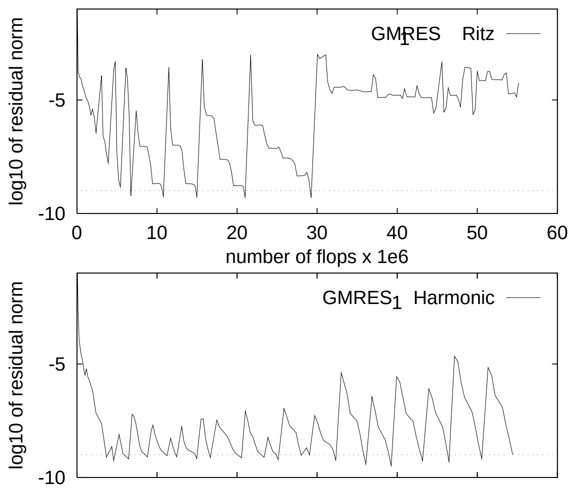
Alfvén spectrum



MHD test problem ($n = 416$).
Correct eigenvalues ($\circ$), target ($\boxed{*}$).

MHD416

Searching for the Alfvén branch



ILU(2) for $\mathbf{A} - \tau\mathbf{B}$ is used as preconditioner.

max dim($\mathbf{V}$)=15
min dim($\mathbf{V}$)=10

More details

JD method:

Sleijpen & van der Vorst
SIMAX, 17 (2), March 1996 pp. 401-425

JD for generalized problems:

Sleijpen, et al
BIT, 36:3 (1996), pp. 595–633

JDQR and JDQZ:

Fokkema, et al
Preprint 941, UU, 1996
To appear in SIAM J. Sc. Comput

http://www.math.ruu.nl/people/sleijpen