On the convergence of the residual inverse iteration

Daniel Kressner

Chair of Numerical Algorithms and HPC
MATHICSE / SMA / SB
EPF Lausanne
http://anchp.epfl.ch

$N(\lambda)$2014, Manchester

Joint work with C. Effenberger.

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Nonlinear eigenvalue problems

Mathematical formulation.
For analytic $T : \Omega \to \mathbb{C}^{n \times n}$, $\Omega \subseteq \mathbb{C}$, find eigenvalues $\lambda \in \mathbb{C}$:

$$T(\lambda)\, x = 0, \quad x \neq 0 \tag{NLEVP}$$

Several exciting recent algorithmic developments:

- nonlinear Arnoldi [Voss'2004]
- rational linearization [Su/Bai'2011]
- block Newton [K.'2009]
- infinite Arnoldi [Jarlebring/Meerbergen/Michiels'2010/2012]
- contour integral method [Asakura/Sakurai/Tadano/Ikegami/Kimura'2009, Beyn'2012, ...]
- interpolation based approaches [Effenberger/K.'2012, Van Beeumen/Meerbergen/Michiels'2013, K./Roman'2013, Güttel/Van Beeumen/Meerbergen/Michiels'2013]
- Jacobi-Davidson with eigenvalue locking [Effenberger'2013]
- ...

Nonlinear eigenvalue problems

Mathematical formulation.
For analytic $T : \Omega \to \mathbb{C}^{n\times n}$, $\Omega \subseteq \mathbb{C}$, find eigenvalues $\lambda \in \mathbb{C}$:

$$T(\lambda)\, x = 0, \quad x \neq 0 \tag{NLEVP}$$

Several exciting recent algorithmic developments:

- nonlinear Arnoldi [Voss'2004]
- rational linearization [Su/Bai'2011]
- block Newton [K.'2009]
- infinite Arnoldi [Jarlebring/Meerbergen/Michiels'2010/2012]
- contour integral method [Asakura/Sakurai/Tadano/Ikegami/Kimura'2009, Beyn'2012, ...]
- interpolation based approaches [Effenberger/K.'2012, Van Beeumen/Meerbergen/Michiels'2013, K./Roman'2013, Güttel/Van Beeumen/Meerbergen/Michiels'2013]
- Jacobi-Davidson with eigenvalue locking [Effenberger'2013]
- ...

Not the topic of this talk!

Residual inverse iteration

Plain Newton applied to nonlinear system of equations

$$T(\lambda)x = 0, \quad w^* x = 1$$

yields:

$$\begin{aligned} y_{j+1} &= T(\lambda_j)^{-1} T'(\lambda_j) v_j \\ \lambda_{j+1} &= \lambda_j - 1/w^* y_{j+1} \\ v_{j+1} &= y_{j+1}/w^* y_{j+1} \end{aligned}$$

$\rightsquigarrow$ Need derivative and need to refactorize $T(\lambda_j)$ in every step.

Residual inverse iteration

Quasi-Newton applied to nonlinear system of equations

$$T(\lambda)x = 0, \quad w^* x = 1$$

yields:

$$\begin{aligned}
y_{j+1} &= T(\lambda_j)^{-1} \frac{1}{\lambda_{j+1} - \lambda_j} (T(\lambda_{j+1}) - T(\lambda_j)) v_j \\
\lambda_{j+1} &= \lambda_j - 1/w^* y_{j+1} \\
v_{j+1} &= y_{j+1}/w^* y_{j+1}
\end{aligned}$$

Residual inverse iteration

Quasi-Newton applied to nonlinear system of equations

$$T(\lambda)x = 0, \quad w^* x = 1$$

yields:

$$v_{j+1} = v_j - T(\lambda_j)^{-1} T(\lambda_{j+1}) v_j$$

Residual inverse iteration

Quasi-Newton applied to nonlinear system of equations

$$T(\lambda)x = 0, \quad w^* x = 1$$

yields:

$$v_{j+1} = v_j - T(\lambda_j)^{-1} T(\lambda_{j+1}) v_j$$

$\rightsquigarrow$ Ups... how do we get λ_{j+1}?

Residual inverse iteration

Quasi-Newton applied to nonlinear system of equations

$$T(\lambda)x = 0, \quad w^* x = 1$$

yields:

$$v_{j+1} = v_j - T(\lambda_j)^{-1} T(\lambda_{j+1}) v_j$$

$\rightsquigarrow$ Ups... how do we get λ_{j+1}?

- In the linear Hermitian case: Rayleigh quotient solves $v_j^* A v_j - \lambda v_j^* v_j = 0$.
- In the nonlinear Hermitian case: λ_{j+1} solves $v_j^* T(\lambda) v_j = 0$.

Residual inverse iteration

for $j = 0, 1, 2, \ldots$

Compute solution λ_{j+1} of $v_j^* T(\lambda) v_j = 0$.

$v_{j+1} = v_j - T(\lambda_j)^{-1} T(\lambda_{j+1}) v_j$

Normalize v_{j+1}.

end for

$\rightsquigarrow$ Still need to refactorize $T(\lambda_j)$ in every step!

Residual inverse iteration

Choose fixed σ sufficiently close to eigenvalue of interest.

for $j = 0, 1, 2, \ldots$

Compute solution λ_{j+1} of $v_j^* T(\lambda) v_j = 0$.

$v_{j+1} = v_j - T(\sigma)^{-1} T(\lambda_{j+1}) v_j$

Normalize v_{j+1}.

end for

Finally...
... this is RESINVIT (residual inverse iteration) by Neumaier [1985].

Residual inverse iteration

Choose fixed σ sufficiently close to eigenvalue of interest.

for $j = 0, 1, 2, \ldots$

Compute solution λ_{j+1} of $v_j^* T(\lambda) v_j = 0$.

$v_{j+1} = v_j - T(\sigma)^{-1} T(\lambda_{j+1}) v_j$

Normalize v_{j+1}.

end for

- For linear $T(\lambda) = \lambda I - A$: Odd way of writing inverse iteration.

  When using preconditioner $P \approx -T(\sigma) \rightsquigarrow$ PINVIT.

Preconditioner for nonlinear eigenvalue problems?

Model problem

Vibrating string with elastically attached mass [Solov'ëv'2006]:

$$-u''(x) = \lambda u(x), \quad u(0) = 0, \quad u'(1) + \phi(\lambda)u(1) = 0, \quad \phi(\lambda) = \frac{\lambda\eta m}{\lambda - \eta}.$$

Discretization with pw linear FEs $\rightsquigarrow$

$$\big(K(\lambda) - \lambda M\big)v = 0$$

with $K(\lambda) = K_0 + \phi(\lambda)e_n e_n^{\mathsf{T}}$ and

$$K_0 = \frac{1}{h}\begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & 2 & -1 \\ & & -1 & 1 \end{bmatrix}, \quad M = \frac{h}{6}\begin{bmatrix} 4 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & 4 & 1 \\ & & 1 & 2 \end{bmatrix}.$$

Model problem

- Aim to compute smallest eigenvalue $\lambda_1 = 4.482024295\ldots$.
- Apply RESINVIT with preconditioner for $T(\sigma) =$ one W-cycle of multigrid with Jacobi smoother.
- Stop when residual $\leq 10^{-6}$.

	#iterations		
h	$\sigma = 0$	$\sigma = 2$	$\sigma = 4$
2^{-5}	12	8	5
2^{-6}	12	8	5
2^{-7}	12	8	5
2^{-8}	12	7	4
2^{-9}	12	7	4
2^{-10}	12	7	4
2^{-11}	11	7	4
2^{-12}	11	7	4
2^{-13}	11	6	4
2^{-14}	10	6	4
2^{-15}	10	6	4
2^{-16}	10	6	4
2^{-17}	9	6	3
2^{-18}	9	5	3

Goal of this talk

Derive convergence result that allows to conclude

mesh-independent asymptotic convergence rates

when using a spectrally equivalent preconditioner for $T(\sigma)$.

[Solov'ëv'2006] has established such a result for "monotonic" nonlinear eigenvalue problems

$$T(\lambda) = A(\lambda) - \lambda B(\lambda)$$

with spd $A(\lambda), B(\lambda)$.

Convergence result by Neumaier

- Consider simple eigenvalue λ_1 with eigenvector x_1 normalized s.t. $w^* x_1 = 1$.
- Assume that σ is sufficiently close to λ_1.
- Exact preconditioner $T(\sigma)^{-1}$.

Then RESINVIT converges with

$$\frac{\|v_{j+1} - x_1\|_2}{\|v_j - x_1\|_2} = O(|\sigma - \lambda_1|), \qquad |\lambda_{j+1} - \lambda_1| = O(\|v_{j+1} - x_1\|_2).$$

Convergence result by Neumaier

- Consider simple eigenvalue λ_1 with eigenvector x_1 normalized s.t. $w^* x_1 = 1$.
- Assume that σ is sufficiently close to λ_1.
- Assume that T is Hermitian.
- Exact preconditioner $T(\sigma)^{-1}$.

Then RESINVIT converges with

$$\frac{\|v_{j+1} - x_1\|_2}{\|v_j - x_1\|_2} = O(|\sigma - \lambda_1|), \qquad |\lambda_{j+1} - \lambda_1| = O(\|v_{j+1} - x_1\|_2^2).$$

Convergence result by Neumaier

- Consider simple eigenvalue λ_1 with eigenvector x_1 normalized s.t. $w^* x_1 = 1$.
- Assume that σ is sufficiently close to λ_1.
- Assume that T is Hermitian.
- Exact preconditioner $T(\sigma)^{-1}$.

Then RESINVIT converges with

$$\frac{\|v_{j+1} - x_1\|_2}{\|v_j - x_1\|_2} = O(|\sigma - \lambda_1|), \qquad |\lambda_{j+1} - \lambda_1| = O(\|v_{j+1} - x_1\|_2^2).$$

But: Not helpful if constant in $O(|\sigma - \lambda_1|)$ deteriorates as $h \to 0$.

Convergence rate

Detailed analysis of convergence rate by [Jarlebring/Michiels'2011]:

$$\text{asympt convergence rate} = |\sigma - \lambda_1|\rho(B_1) + o(|\sigma - \lambda_1|)$$

with

$$\begin{aligned} B_0 &= I - \frac{T'(\lambda_1)x_1x_1^*}{x_1^*T'(\lambda_1)x_1} \\ B_1 &= B_0\left(-T'(\lambda_1)T(\lambda_1)^+B_0 - \frac{T''(\lambda_1)x_1x_1^*}{x_1^*T'(\lambda_1)x_1}\right)\left(I + \frac{T'(\lambda_1)x_1x_1^*}{x_1^*T'(\lambda_1)x_1}\right)^{-1} \end{aligned}$$

Inexact variant of RESINVIT analyzed by [Szyld/Xue'2013].

PINVIT (Preconditioned inverse iteration)

for $j = 0, 1, 2, \ldots$

$$v_{j+1} = v_j + P^{-1}\big((v_j^* A v_j I - A) v_j\big)$$

Normalize v_{j+1}.

end for

Elegant convergence analyses by [D'yakonov/Orekhov'1980], [Knyazev'1998], [Neymeyr'2001], [Knyazev/Neymeyr'2003], ...:

- Setting: symmetric positive definite A and aim at smallest eigenvalue λ_1.
- Neymeyr's analysis relies on mini-dimensional analysis for Rayleigh quotient.
- Allow to conclude mesh-independent convergence rates.

PINVIT (Preconditioned inverse iteration)

for $j = 0, 1, 2, \ldots$

$$v_{j+1} = v_j + P^{-1}\big((v_j^* A v_j I - A) v_j\big)$$

Normalize v_{j+1}.

end for

Elegant convergence analyses by [D'yakonov/Orekhov'1980], [Knyazev'1998], [Neymeyr'2001], [Knyazev/Neymeyr'2003], ...:

- Setting: symmetric positive definite A and aim at smallest eigenvalue λ_1.
- Neymeyr's analysis relies on mini-dimensional analysis for Rayleigh quotient.
- Allow to conclude mesh-independent convergence rates.

Difficult to extend this analysis to
general nonlinear eigenvalue problems.

Instead: Extend poor men's approach by interpreting PINVIT as perturbed inverse iteration.

Rayleigh functionals

Back to nonlinear Hermitian $T(\lambda)$.

Assume that scalar nonlinear equation

$$x^* T(\lambda) x = 0$$

admits unique solution $\lambda \in \Omega$ for every x in open set $D \in \mathbb{C}^n$.

Defines Rayleigh functional

$$\rho : D \to \Omega, \quad x \mapsto \lambda.$$

Additional assumption:

$$x^* T'(\rho(x)) x > 0 \quad \forall x \in D.$$

- Allows to carry over variational principles [Hadeler'1968], [Voss/Werner'1982], ...
- For example:

$$\lambda_1 = \inf_{x \in D} \rho(x).$$

$\rightsquigarrow -T(\sigma)$ is spd (on D) for any $\sigma < \lambda_1$.

Rayleigh functional for model problem

$$T(\lambda) = \lambda M - K_0 - \phi(\lambda) e_n e_n^{\mathsf{T}}.$$

with spd M, K_0 and $\phi(\lambda) = \frac{\lambda \eta m}{\lambda - \eta}$.

- $x^* T(\lambda) x = 0$ has exactly one solution in $]\eta, \infty]$ for every $x \in \mathbb{C}^n \setminus \{0\}$.

$\rightsquigarrow$ $\exists$ Rayleigh functional for $\Omega =]\eta, \infty]$ and $D = \mathbb{C}^n \setminus \{0\}$.

One step of residual inverse iteration

RESINVIT with Rayleigh functional ρ:

$$v^+ = v + P^{-1} T\big(\rho(v)\big) v$$

for some spd preconditioner P.

Ignore scaling: will work with angles.

What is the right geometry?

Geometry induced by P

RESINVIT with Rayleigh functional ρ:

$$v^+ = v + P^{-1}T(\rho(v))v$$

for some spd preconditioner P.

- Geometry induced by P:

$$\langle v, w\rangle_P := v^* P w,$$

  and induced norm $\|v\|_P := \sqrt{v^* P v}$.
- Angle $\phi_P(v, x_1)$ between v and x_1:

$$\cos\phi_P(v, x_1) := \frac{\mathrm{Re}\langle v, x_1\rangle_P}{\|v\|_P\|w\|_P}$$

- Given P-orthogonal decomposition

$$v = v_1 + v_\perp, \quad v_1 \in \mathrm{span}\{x_1\}, \quad \langle v_1, v_\perp\rangle_P = 0,$$

  $\rightsquigarrow$

$$\tan\phi_P(v, x_1) = \|v_\perp\|_P / \|v_1\|_P.$$

Convergence result

Theorem [Effenberger/K.'2013].

$$\tan \phi_P(v^+, x_1) \le \gamma \cdot \tan \phi_P(v, x_1) + O(\varepsilon^2)$$

with $\varepsilon = \tan \phi_P(v, x_1)$ and convergence factor

$$\gamma = \max_{\substack{z \neq 0 \\ \langle z, x_1 \rangle_P = 0}} \left| 1 + \frac{z^* T(\lambda_1) z}{z^* P z} \right|.$$

Proof based on expansion $T(\rho(v)) = T(\lambda_1) + O(\varepsilon^2)$.

- Linear case $T(\lambda) = \lambda I - A$ with exact preconditioner $P = A \rightsquigarrow$

$$\gamma = 1 + \frac{\lambda_1 - \lambda_2}{\lambda_2} = \frac{\lambda_1}{\lambda_2}$$

- Sufficient condition for convergence: $\exists 0 \le \tilde{\gamma} < 1$ such that

$$(1 - \tilde{\gamma}) z^* P z \le -z^* T(\lambda_1) z \le (1 + \tilde{\gamma}) z^* P z \quad \forall z \perp_P x_1.$$

  Satisfied, e.g., by multigrid for $-T(\sigma)$ with
  σ sufficiently close to λ_1.

Summary

RESINVIT for Hermitian $T(\lambda)$ with preconditioner $P \approx -T(\sigma)$ has linear convergence rate

$$\gamma = \max_{\substack{z \neq 0 \\ \langle z, x_1 \rangle_P = 0}} \left| 1 + \frac{z^* T(\lambda_1) z}{z^* P z} \right|.$$

- For $P = -T(\sigma)$: $\gamma < 1$ if λ_1 simple eigenvalue and $\sigma \approx \lambda_1$
- Aim:
  Show that γ stays away from 1 as $h \to 0$ for discretized PDE.

Function space setting

- Hilbert spaces V and H with dense and compact embedding $V \hookrightarrow H$ (think of H^1 and L^2)
- Variational formulation of NLEVP:

$$\text{Find } (u, \lambda) \in V \times \Omega \text{ such that}$$
$$a(u, v, \lambda) = 0 \quad \forall v \in V.$$

  with bounded bilinear form $a(\cdot, \cdot, \lambda)$ depending holomorphically on λ.
- Fix λ and consider linear eigenvalue problem associated with $\bar{a}(\cdot, \cdot, \lambda) := -a(\cdot, \cdot, \lambda)$:

$$\text{Find } (u, \mu) \in V \times \mathcal{D} \text{ such that}$$
$$\bar{a}(u, v, \lambda) = \mu \langle u, v \rangle_H \quad \forall v \in V.$$

  Assume

$$\exists c, s \quad \text{such that} \quad \bar{a}(v, v, \lambda) \geq c\|v\|_V^2 - s\|v\|_H^2 \quad \forall v \in V.$$

- Modification of setting considered by [Voss'2013]: Existence of Rayleigh functional $\rightsquigarrow$ variational principles for eigenvalues.

Function space setting for model problem

For model problem $V = H^1(]0,1[), H = L^2(]0,1[)$, and

$$a(u, v, \lambda) = \int_\Omega \nabla u \cdot \nabla v \; dx + \varphi(\lambda) u(1) v(1) - \lambda \int_\Omega uv \; dx.$$

$\leadsto$ all assumptions are satisfied.

Further assumptions

- $0 = \mu_1(\lambda_1) < \mu_2(\lambda_1) \leq \mu_3(\lambda_1) \leq \cdots$
- a bounded on $V \times V$
- $\frac{\partial}{\partial \lambda} a$ bounded on $H \times H$
- $\frac{\partial}{\partial \lambda} a$ Lipschitz continuous wrt λ in neighbourhood of λ_1

Attaining mesh-independent convergence rate

Outline of proof:

1. For eigenpair (λ_1, u_1) choose $\sigma < \lambda_1$ sufficiently close to λ_1
$\rightsquigarrow$ (shifted) $\bar{a}(\cdot, \cdot, \sigma)$ is elliptic on V.
2. Assumptions imply existence of $0 \leq \delta < 1$ such that

$$(1-\delta)\bar{a}(v, v, \sigma) \leq \bar{a}(v, v, \lambda_1) \leq (1+\delta)\bar{a}(v, v, \sigma), \qquad (1)$$

holds for all $v \neq 0$ with $a(u_1, v, \sigma) = 0$.
3. Show that (1) holds asymptotically ($h \to 0$) for

$$(1-\delta)\bar{a}_h(v, v, \sigma) \leq \bar{a}_h(v, v, \lambda_{1,h}) \leq (1+\delta)\bar{a}_h(v, v, \sigma), \qquad (2)$$

with $a_h(\sigma), a_h(\lambda_{1,h})$ obtained by Galerkin projection on $V_h \subset V$.
Main problem: $\lambda_{1,h}$ is the eigenvalue of the discrete problem!
4. Show that (1) remains true when $a_h(\sigma)$ is replaced by spectrally equivalent preconditioner P.

Establishes h-independent convergence rate δ.

Conclusions

- Proposed and analyzed preconditioned variant of RESINVIT for nonlinear eigenvalue problems admitting a Rayleigh functional.
- Surrogate for understanding convergence of more advanced algorithms, e.g., nonlinear Arnoldi.
- Analysis requires preconditioner of $T(\sigma)$ with σ sufficiently close to λ_1. In practice: Not a severe limitation.
- Ongoing work (with A. Miedlar): Adaptive FEM.

References:

C. Effenberger, D. K. On the convergence of the residual inverse iteration for nonlinear eigenvalue problems admitting a Rayleigh functional. Preprint available from `http://anchp.epfl.ch`.

C. Effenberger. Robust solution methods for nonlinear eigenvalue problems. PhD thesis, EPFL, 2013.