Rational Krylov methods for $f(A)\mathbf{b}$

Michael Eiermann, Oliver G. Ernst, and Stefan Güttel

DWCAA09

September 8th, 2009

Problem

We consider the vector $f(A)\mathbf{b}$, where

- A is a large N-by-N matrix,
- $\mathbf{b}$ is a vector of length N,
- f is a suitable function.

Compute approximation $\mathbf{f}_m \approx f(A)\mathbf{b}$ from a **rational Krylov space.**

What is a rational Krylov space?

Let $\{\xi_1, \xi_2, \dots\} \subseteq \Xi \subset \overline{\mathbb{C}}$ be a given sequence of **poles**.

Define the polynomials

$$q_m(z) = \prod_{\substack{j=1 \\ \xi_j \neq \infty}}^{m} (z - \xi_j) \in \mathcal{P}_m.$$

Assume that $q_m(A)^{-1}$ exists. Then

$$\mathcal{Q}_{m+1}(A, \mathbf{b}) = \mathcal{K}_{m+1}(A, q_m(A)^{-1}\mathbf{b})$$

is the rational Krylov space associated with $(A, \mathbf{b}, q_m)$.

Special cases

- $\Xi = \{\infty\} \Rightarrow$ polynomial Krylov $\mathcal{Q}_{m+1} = \mathcal{K}_{m+1}$

  [Nauts & Wyatt 83] [van der Vorst 87] [Druskin & Knizhnerman 88]

  [Gallopoulos & Saad 92] [Hochbruck & Lubich 97] [Eiermann & Ernst 06]

- $\Xi = \{\xi\} \Rightarrow$ shift-invert Krylov

  [Moret & Novati 04] [van den Eshof & Hochbruck 06]

- $\Xi = \{0, \infty\} \Rightarrow$ extended Krylov

  [Druskin & Knizhnerman 98] [Knizhnerman & Simoncini 08]

- Ξ arbitrary $\Rightarrow$ rational Krylov

  [Ruhe 84] [Beattie 04] [Beckermann & Reichel 08] [Knizhnerman et al 08]

Rational Arnoldi algorithm [Ruhe 84/94]

Input: A, $\mathbf{b}$, $\{\xi_1, \xi_2, \ldots, \xi_m\}$

$\mathbf{v}_1 := \mathbf{b}/\|\mathbf{b}\|$

for $j = 1, 2, \ldots, m$ **do**

$\quad \mathbf{x} := (I - A/\xi_j)^{-1} A\mathbf{v}_j$

$\quad H(1:j, j) := [\mathbf{v}_1, \ldots, \mathbf{v}_j]^* \mathbf{x}$

$\quad \mathbf{x} := \mathbf{x} - [\mathbf{v}_1, \ldots, \mathbf{v}_j] H(1:j, j)$

$\quad H(j+1, j) := \|\mathbf{x}\|$

$\quad \mathbf{v}_{j+1} := \mathbf{x}/H(j+1, j)$

end

Yields decomposition $AV_{m+1}(\underline{I_m} + \underline{H_m} X_m^{-1}) = V_{m+1}\underline{H_m}$.

Rational Krylov decompositions

Theorem (G., 2009): Let a general decomposition

$$AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$$

be given, where V_{m+1} has $m+1$ linearly independent columns, $\underline{K_m} \in \mathbb{C}^{(m+1)\times m}$, $\underline{H_m} \in \mathbb{C}^{(m+1)\times m}$, and $\underline{H_m}$ is of rank m.

Rational Krylov decompositions

Theorem (G., 2009): Let a general decomposition

$$AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$$

be given, where V_{m+1} has $m+1$ linearly independent columns, $\underline{K_m} \in \mathbb{C}^{(m+1)\times m}$, $\underline{H_m} \in \mathbb{C}^{(m+1)\times m}$, and $\underline{H_m}$ is of rank m. Then

1. $\underline{K_m}$ is of rank m.
2. $\mathrm{colspan}(V_{m+1}) = \mathcal{K}_{m+1}(A, \mathbf{q})$ for a vector $\mathbf{q}$.
3. For every vector $\mathbf{b} \in \mathrm{colspan}(V_{m+1})$ there exists a unique polynomial q_m, $\deg(q_m) \leq m$, such that $\mathbf{b} = q_m(A)\mathbf{q}$. Hence, if $q_m(A)$ is invertible, $\mathrm{colspan}(V_{m+1}) = \mathcal{Q}_{m+1}(A, \mathbf{b})$.

Rational Krylov approximations

A special case is the (reduced) decomposition

$$AV_mK_m = V_{m+1}\underline{H_m}.$$

As an approximation to $f(A)\mathbf{b}$ we consider

$$\mathbf{f}_m := V_m f(H_m K_m^{-1}) V_m^\dagger \mathbf{b}.$$

Theorem (Interpolation): There holds

$$\mathbf{f}_m = r_m(A)\mathbf{b} = \frac{p_{m-1}}{q_{m-1}}(A)\mathbf{b},$$

where r_m Hermite-interpolates f at $\Lambda(H_m K_m^{-1})$.

Example

The iteration

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{b}, \\ \beta_j \mathbf{v}_{j+1} &= (I - A/\xi_j)^{-1}(A - \alpha_j I)\mathbf{v}_j, \quad j = 1, \ldots, m, \end{aligned}$$

yields a decomposition $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with
$V_{m+1} = [\mathbf{v}_1, \ldots, \mathbf{v}_{m+1}]$,

$$\underline{K_m} = \left[\begin{array}{cccc} 1 & & & \\ \beta_1/\xi_1 & 1 & & \\ & \beta_2/\xi_2 & \ddots & \\ & & \ddots & 1 \\ \hline & & & \beta_m/\xi_m \end{array}\right] \quad \text{and} \quad \underline{H_m} = \left[\begin{array}{cccc} \alpha_1 & & & \\ \beta_1 & \alpha_2 & & \\ & \beta_2 & \ddots & \\ & & \ddots & \alpha_m \\ \hline & & & \beta_m \end{array}\right].$$

Example

The iteration

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{b}, \\ \beta_j \mathbf{v}_{j+1} &= (I - A/\xi_j)^{-1}(A - \alpha_j I)\mathbf{v}_j, \quad j = 1, \dots, m, \end{aligned}$$

can be used for explicit rational interpolation:

By Theorem (Interpolation) we know that

$$\mathbf{f}_m = V_m f(H_m K_m^{-1})\mathbf{e}_1 = r_m(A)\mathbf{b} = \frac{p_{m-1}}{q_{m-1}}(A)\mathbf{b},$$

where r_m Hermite-interpolates f at $\Lambda(H_m K_m^{-1}) = \{\alpha_1, \dots, \alpha_m\}$.

Example

PAIN method
Preassigned Poles and Interpolation Nodes

The iteration

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{b}, \\ \beta_j \mathbf{v}_{j+1} &= (I - A/\xi_j)^{-1}(A - \alpha_j I)\mathbf{v}_j, \quad j = 1, \ldots, m, \end{aligned}$$

can be used for explicit rational interpolation:

By Theorem (Interpolation) we know that

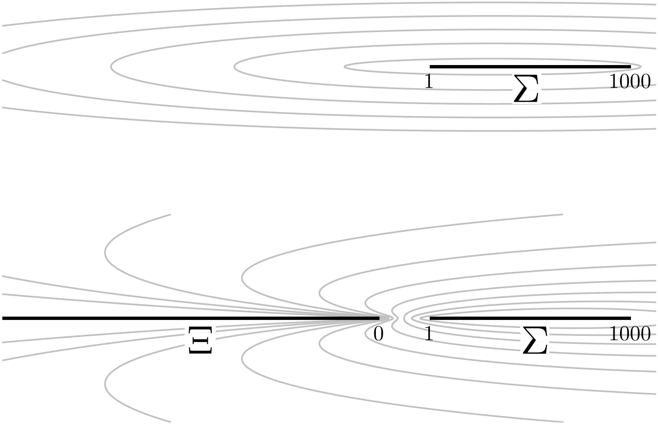
$$\mathbf{f}_m = V_m f(H_m K_m^{-1})\mathbf{e}_1 = r_m(A)\mathbf{b} = \frac{p_{m-1}}{q_{m-1}}(A)\mathbf{b},$$

where r_m Hermite-interpolates f at $\Lambda(H_m K_m^{-1}) = \{\alpha_1, \ldots, \alpha_m\}$.

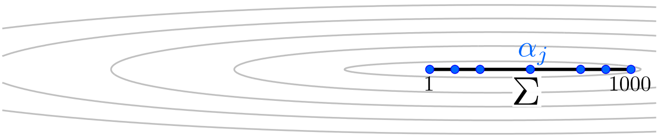
Remarks

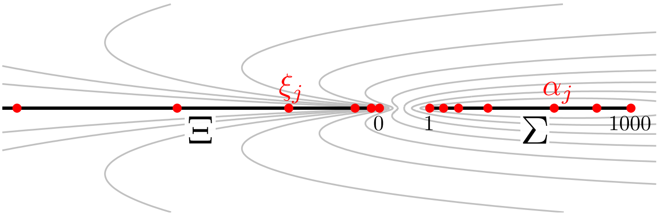
- 2 vectors storage need, 0 inner-products
- If all $\xi_j = \infty \Rightarrow$ polynomial interpolation at $\{\alpha_1, \ldots, \alpha_m\}$
- Polynomial interpolation methods have been considered before [Huisinga et al 99] [Bergamaschi, Caliari & Vianello 04]
  - For $\{\alpha_1, \ldots, \alpha_m\}$ use Leja points, scaled to a set of unit capacity for stability [Reichel 90].
  - No such scaling is necessary with the PAIN method: simply choose β_j such that $\|\mathbf{v}_{j+1}\| = 1, j = 1, \ldots, m$.
- For rational interpolation use **Leja-Bagby points.**

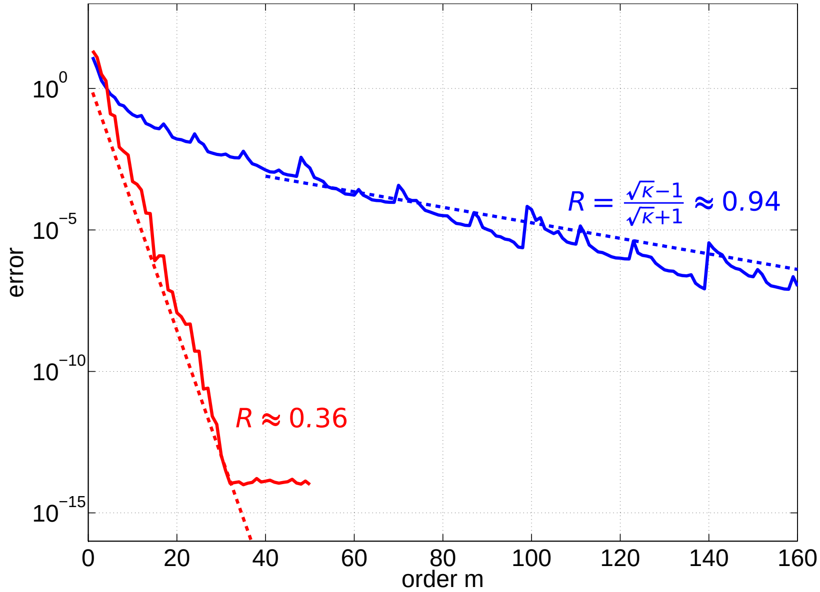
Compute: $f(A)\mathbf{b} = \sqrt{A}\mathbf{b}$, $A = \mathrm{diag}(1, \ldots, 1000)$, $\mathbf{b} = [1, \ldots, 1]^T$.



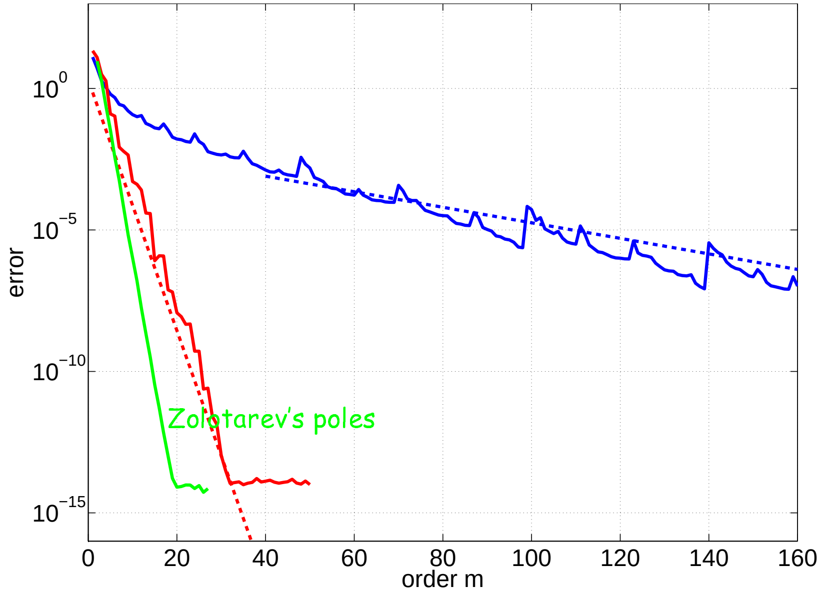
Compute: $f(A)\mathbf{b} = \sqrt{A}\mathbf{b}$, $A = \mathrm{diag}(1, \ldots, 1000)$, $\mathbf{b} = [1, \ldots, 1]^T$.







10^{0}
10^{-5}
10^{-10}
10^{-15}
error
0
20
40
60
80
100
120
140
160
order m
$R = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \approx 0.94$
$R \approx 0.36$



error
order m
Zolotarev's poles
10^{0}
10^{-5}
10^{-10}
10^{-15}
0
20
40
60
80
100
120
140
160

If good (or best) rational approximation r_m^* to f is known explicitly, one can directly evaluate $r_m^*(A)\mathbf{b} \approx f(A)\mathbf{b}$.

[Trefethen et al 06] [Frommer et al 06] [Schmelzer et al 07] [Hale et al 08]

However, using the poles ξ_j of r_m^* and suitable interpolation nodes α_j as inputs for PAIN, we can achieve essentially the same accuracy at the same computational cost.

Moreover, the PAIN method is implicitly based on **exact** interpolation of f and hence robust to perturbations in r_m^*:

$$\limsup_{m\to\infty} \| f(A)\mathbf{b} - \mathbf{f}_m \|^{1/m} \leq R < 1$$

if α_j, ξ_j are equilibrium-distributed on Σ, Ξ.

Rayleigh-Ritz extraction

There is a way to automatically choose near-optimal interpolation points $\{\alpha_1, \ldots, \alpha_m\}$ at iteration m:

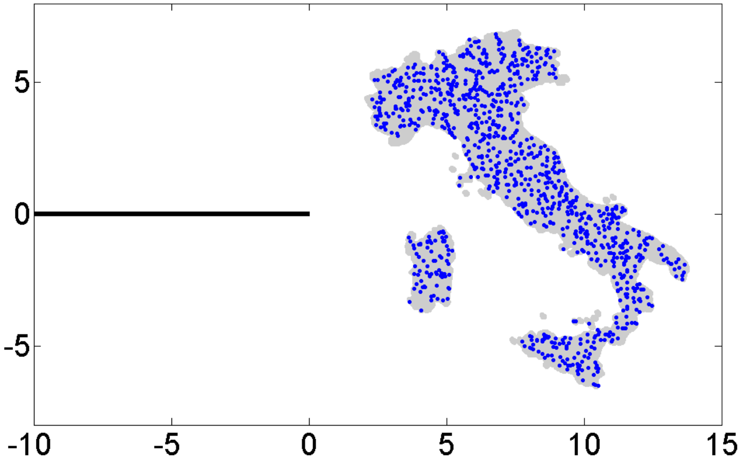
1. Compute orthonormal basis V_m of $\mathcal{Q}_m = q_{m-1}^{-1}\mathcal{K}_m$.
2. "Determine" Rayleigh quotient $A_m = V_m^* A V_m$.
3. Compute $\mathbf{f}_m = V_m f(A_m) V_m^* \mathbf{b}$.

Theorem: $\{\alpha_1, \ldots, \alpha_m\} = \Lambda(A_m)$.

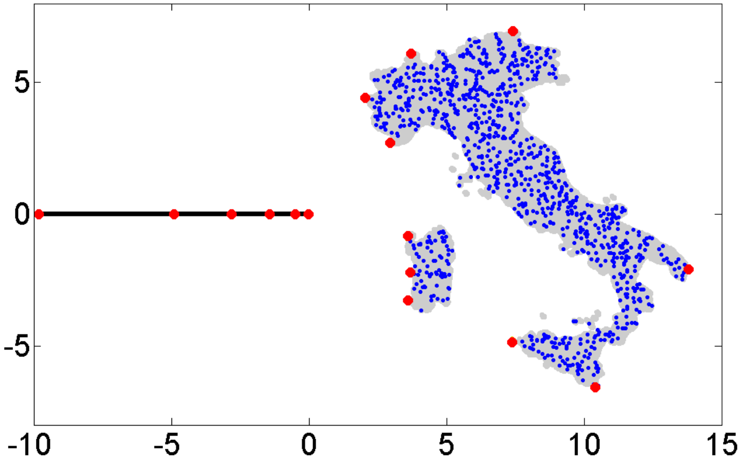
Theorem: $\|f(A)\mathbf{b} - \mathbf{f}_m\| \leq C \min_{p \in \mathcal{P}_{m-1}} \|f - p/q_{m-1}\|_{F(A)}$.

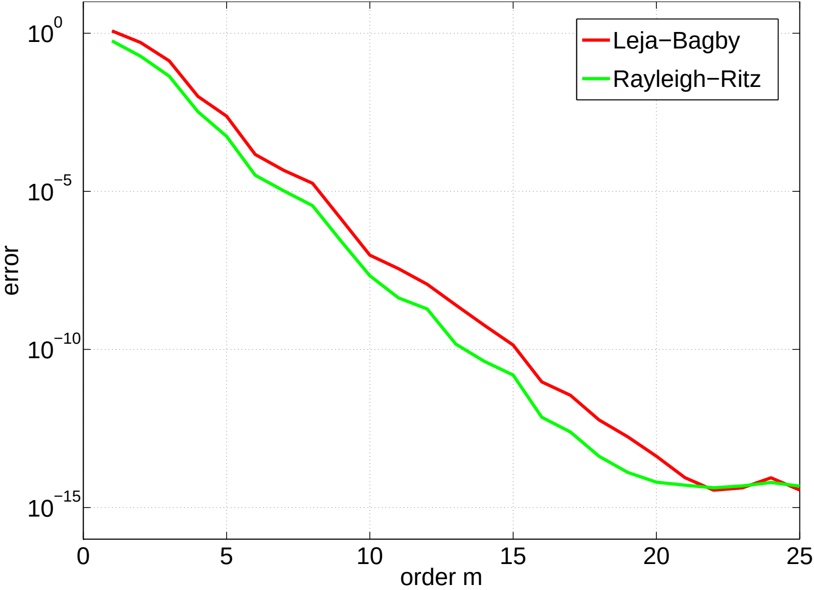
Price: m vectors storage need, $m^2/2$ inner-products.

Compute: $f(A)\mathbf{b} = \log(A)\mathbf{b}$, A normal with 1000 eigenvalues

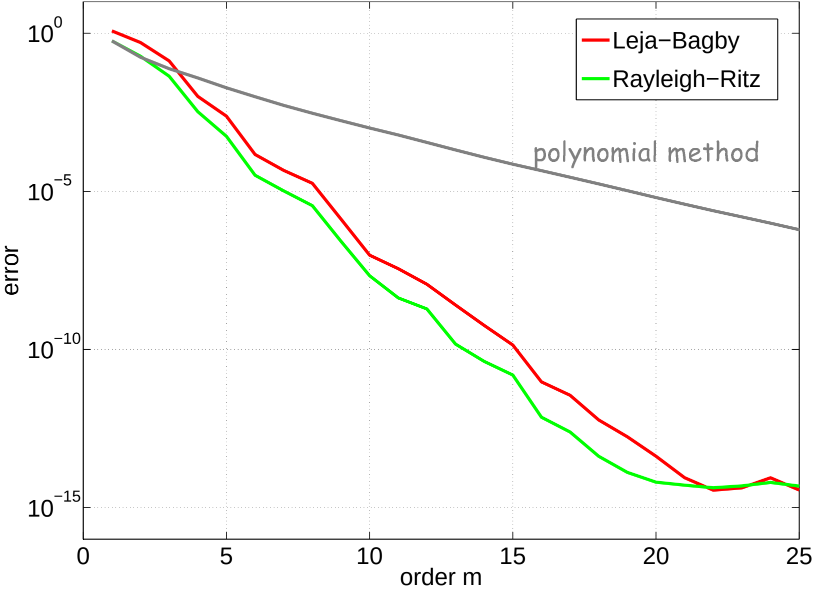


Compute: $f(A)\mathbf{b} = \log(A)\mathbf{b}$, A normal with 1000 eigenvalues





Leja−Bagby
Rayleigh−Ritz
error
order m
10^{0}
10^{-5}
10^{-10}
10^{-15}
0
5
10
15
20
25



Leja−Bagby
Rayleigh−Ritz
polynomial method
error
order m
10^{0}
10^{-5}
10^{-10}
10^{-15}
0
5
10
15
20
25

Parameter-dependent problems

In practice, one often is not interested in $f(A)\mathbf{b}$ but in $f^\tau(A)\mathbf{b}$, for $\tau \in T$ from some **parameter set** T.

Given **a single** rational Krylov decomposition (as before)

$$AV_m K_m = V_{m+1} \underline{H_m}, \quad \mathcal{R}(V_m) = \mathcal{Q}_m = q_{m-1}^{-1} \mathcal{K}_m,$$

we compute **several** approximations

$$\mathbf{f}_m^\tau = V_m f^\tau(H_m K_m^{-1}) V_m^\dagger \mathbf{b} = r_m^\tau(A)\mathbf{b} = \frac{p_{m-1}^\tau}{q_{m-1}}(A)\mathbf{b},$$

where r_m^τ Hermite-interpolates f^τ at $\Lambda(H_m K_m^{-1})$.

Example: Transfer function

$f^\tau(z) = (z - \tau)^{-1}$, spectrum $\Sigma = [0, +\infty)$, parameters $T = i[1, c]$.

Let $\omega_m(z) = (z - \alpha_1) \cdots (z - \alpha_m)$, then

$$r_m^\tau(z) = \frac{1 - \frac{q_{m-1}(\tau)}{q_{m-1}(z)} \frac{\omega_m(z)}{\omega_m(\tau)}}{z - \tau} = \frac{p_{m-1}^\tau}{q_{m-1}}$$

Hermite-interpolates f^τ at $\{\alpha_1, \ldots, \alpha_m\}$. Hence, $\mathbf{f}_m^\tau = r_m^\tau(A)\mathbf{b}$.

Example: Transfer function

$f^\tau(z) = (z-\tau)^{-1}$, spectrum $\Sigma = [0, +\infty)$, parameters $T = i[1, c]$.

Let $\omega_m(z) = (z-\alpha_1)\cdots(z-\alpha_m)$, then

$$r_m^\tau(z) = \frac{1 - \frac{q_{m-1}(\tau)}{q_{m-1}(z)}\frac{\omega_m(z)}{\omega_m(\tau)}}{z-\tau} = \frac{p_{m-1}^\tau}{q_{m-1}}$$

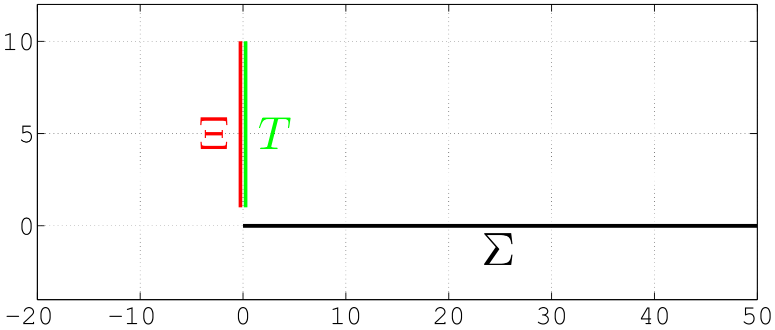
Hermite-interpolates f^τ at $\{\alpha_1, \ldots, \alpha_m\}$. Hence, $\mathbf{f}_m^\tau = r_m^\tau(A)\mathbf{b}$.

The relative error is

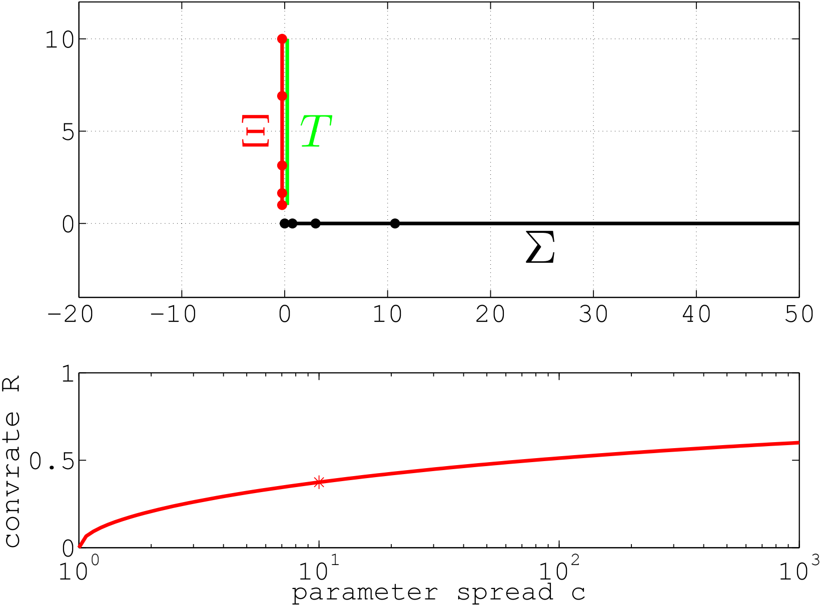
$$[f^\tau(z) - r_m^\tau(z)]/f^\tau(z) = \frac{q_{m-1}(\tau)}{q_{m-1}(z)}\frac{\omega_m(z)}{\omega_m(\tau)}, \quad z \in \Sigma, \tau \in T,$$

and if $\Xi = T$, its minimization is related to the ADI problem.

[Knizhnerman, Druskin & Zaslavsky 08]

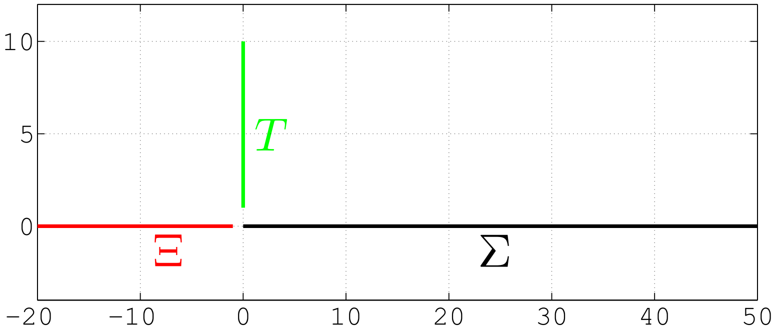


10
5
0
−20
−10
0
10
20
30
40
50
Ξ
T
Σ



10
5
0
−20
−10
0
10
20
30
40
50
Ξ
T
Σ
1
0.5
0
convrate R
10^0
10^1
10^2
10^3
parameter spread c

In practical computations it is more convenient to have real poles, i.e., $\Xi = [-\infty, 0)$. How to select $\{\xi_1, \xi_2, \ldots\} \subset \Xi$?



In practical computations it is more convenient to have real poles, i.e., $\Xi = [-\infty, 0)$. How to select $\{\xi_1, \xi_2, \dots\} \subset \Xi$?

Use standard tool to solve nonstandard approximation problem:

Assume to have a single repeated pole ξ. Then

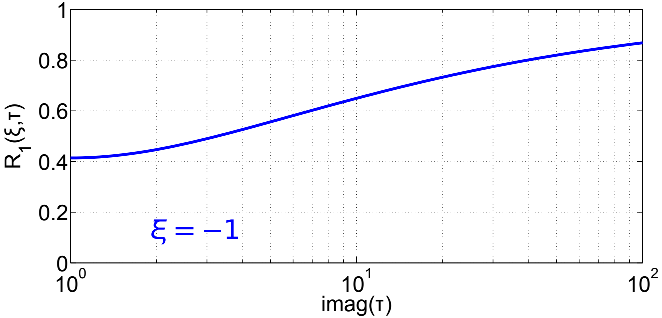
$$r_m^\tau(z) = \frac{p_{m-1}^\tau(z)}{q_{m-1}(z)} = \frac{p_{m-1}^\tau(z)}{(z-\xi)^{m-1}} = \widehat{p}_{m-1}^\tau(\widehat{z}), \quad \widehat{z} = (z-\xi)^{-1},$$

i.e., we have a polynomial problem: among $p \in \mathcal{P}_{m-1}$

$$\text{minimize } \left\| f^\tau(\widehat{z}^{-1} + \xi) - p(\widehat{z}) \right\|_{\widehat{\Sigma}}, \quad \widehat{\Sigma} = \{\widehat{z} : z \in \Sigma\}.$$

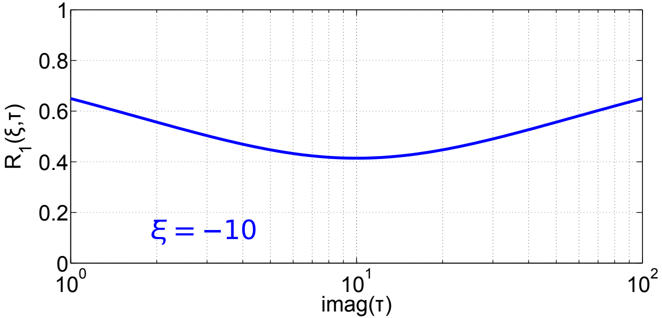
Apply Walsh's theory on polynomial approximation to obtain the asymptotic convergence rate $R_1(\xi, \tau)$. For the transfer function:

$$R_1(\xi,\tau) = \left(1 + \frac{\sqrt{8}d^{3/4} + 4d^{1/2} + \sqrt{8}d^{1/4}}{1+d}\right)^{-1/2}, \quad d = -\tau^2/\xi^2.$$



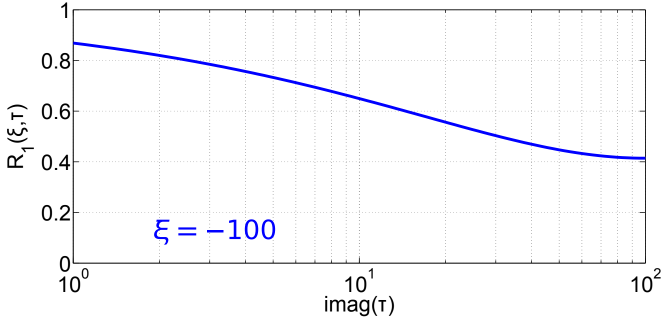
Apply Walsh's theory on polynomial approximation to obtain the asymptotic convergence rate $R_1(\xi, \tau)$. For the transfer function:

$$R_1(\xi,\tau) = \left(1 + \frac{\sqrt{8}d^{3/4} + 4d^{1/2} + \sqrt{8}d^{1/4}}{1+d}\right)^{-1/2}, \quad d = -\tau^2/\xi^2.$$



Apply Walsh's theory on polynomial approximation to obtain the asymptotic convergence rate $R_1(\xi,\tau)$. For the transfer function:

$$R_1(\xi,\tau) = \left(1 + \frac{\sqrt{8}d^{3/4} + 4d^{1/2} + \sqrt{8}d^{1/4}}{1+d}\right)^{-1/2}, \quad d = -\tau^2/\xi^2.$$



Consider p poles $\{\xi_1, \dots, \xi_p\}$ repeated cyclically.

The product form of the error

$$f^\tau(z) - r_m^\tau(z) = \frac{\frac{q_{m-1}(\tau)}{q_{m-1}(z)} \frac{\omega_m(z)}{\omega_m(\tau)}}{z - \tau}$$

allows to conclude that

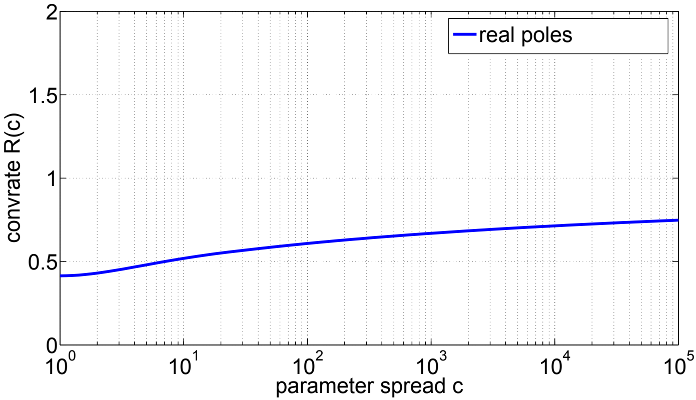
$$R(\{\xi_1, \dots, \xi_p\}, \tau) = \prod_{j=1}^{p} R_1(\xi_j, \tau)^{1/p}$$

is the asymptotic convergence rate for this pole sequence.

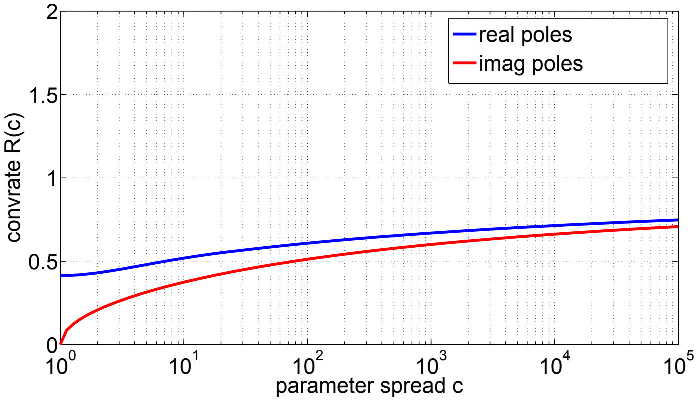
⇒ **Find** $\{\xi_1^*, \dots, \xi_m^*\}$ **minimizing the worst-case rate**

$$\max_{\tau \in T} R(\{\xi_1, \dots, \xi_m\}, \tau).$$

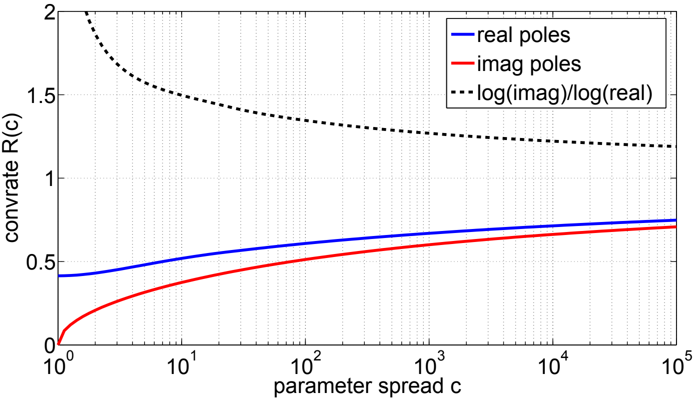
Find optim. poles by nonneg. minimization $\|\mathbf{e} - M\mathbf{x}\|_\infty$. Here is the optimal overall-convergence rate on $T = i[1, c]$ depending on c.



Find optim. poles by nonneg. minimization $\|\mathbf{e} - M\mathbf{x}\|_\infty$. Here is the optimal overall-convergence rate on $T = i[1, c]$ depending on c.



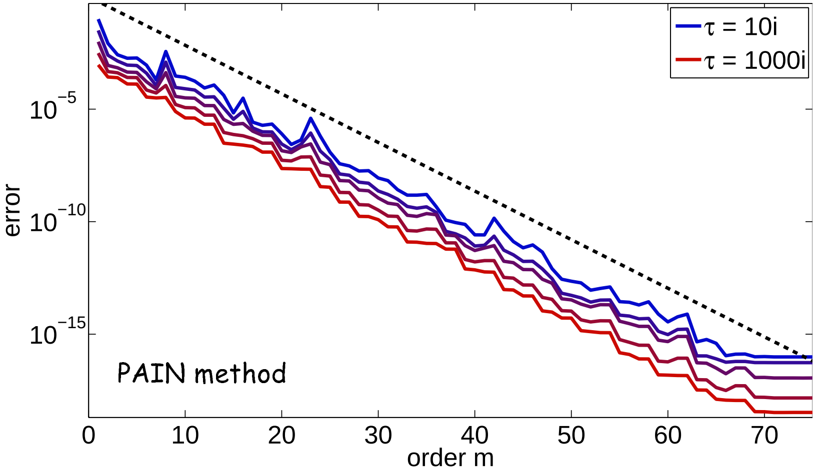
Find optim. poles by nonneg. minimization $\|\mathbf{e} - M\mathbf{x}\|_\infty$. Here is the optimal overall-convergence rate on $T = i[1, c]$ depending on c.



Ex: If solve of complex system is 1.5 as expensive as real solve, use imaginary poles only if $c < 10$!

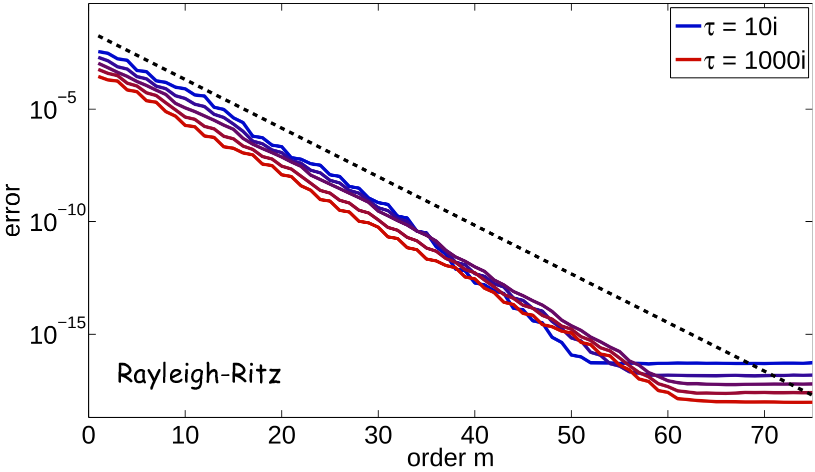
Compute: $f^{\tau}(A)\mathbf{b} = (A - \tau I)^{-1}\mathbf{b}$,

$$A = \mathrm{diag}(0, \ldots, 1e4),\ \tau \in T = i[10, 1000],\ \mathbf{b} = \mathtt{randn}.$$



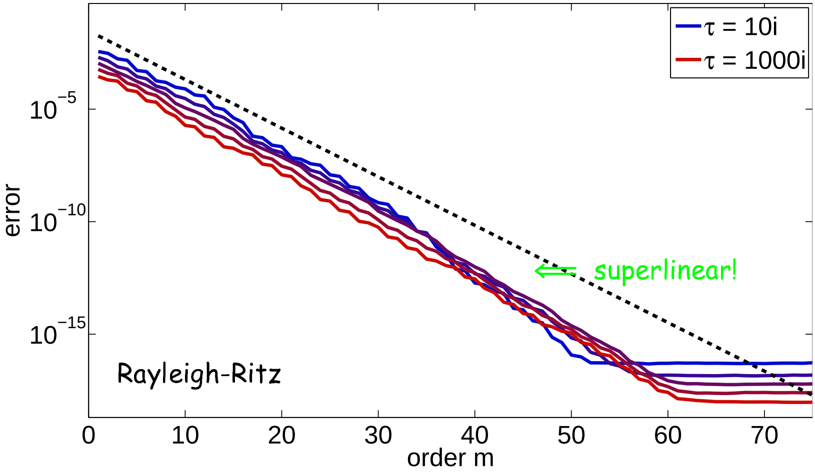
Compute: $f^{\tau}(A)\mathbf{b} = (A - \tau I)^{-1}\mathbf{b}$,

$A = \mathrm{diag}(0, \ldots, 1e4)$, $\tau \in T = i[10, 1000]$, $\mathbf{b} = \mathrm{randn}$.



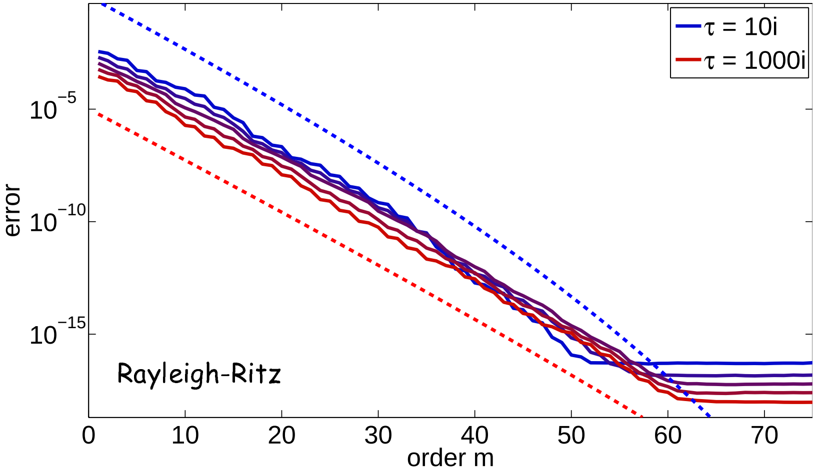
Compute: $f^{\tau}(A)\mathbf{b} = (A - \tau I)^{-1}\mathbf{b}$,

$A = \mathrm{diag}(0, \ldots, 1e4)$, $\tau \in T = i[10, 1000]$, $\mathbf{b} = \mathrm{randn}$.



Compute: $f^\tau(A)\mathbf{b} = (A - \tau I)^{-1}\mathbf{b}$,

$A = \mathrm{diag}(0, \ldots, 1e4)$, $\tau \in T = i[10, 1000]$, $\mathbf{b} = \mathtt{randn}$.



Summary

- Have characterized the general form of a rational Krylov decomposition.
- All existing rational Krylov methods fit into this framework.
- Propose "PAIN" as an efficient and robust rational Krylov method for problems with known spectral properties.
- Have presented simple method for finding constrained pole sequences yielding asymptotically optimal convergence.
- This method may be applied for general f by using Cauchy integral representation.
- Can explain superlinear convergence observed with Rayleigh-Ritz extraction for Hermitian problems using weighted potential theory [Beckermann, G. & Vandebril 09].

See my poster for “rational Ritz values” and “inexact solves”:

Rational Krylov methods and approximation of $f(A)b$

B. Beckermann[‡], M. Eiermann[†], O. G. Ernst[†], Stefan Güttel[†], and R. Vandebril[‡]

[†]TU Bergakademie Freiberg, Germany
Institut für Numerische Mathematik und Optimierung

[‡]Université Lille 1, France
Equipe d'Analyse Numérique et d'Optimisation

Matrix Functions

Given a square matrix A of size $N \times N$, a vector $\boldsymbol{b}$ of length N and a scalar function $f(z)$,

$$f(A)\boldsymbol{b} := p(A)\boldsymbol{b},$$

where $p \in \mathscr{P}_{N-1}$ is a polynomial of degree $\leq N-1$ that Hermite-interpolates f at the eigenvalues of A. In typical applications the matrix A is large and sparse.

Some Applications

- $f(z) = (z - i\omega)^{-1}$: model reduction in the frequency domain,
- $f(z) = \exp(-tz)$: time-integration of linear ODE's, exponential integrators, e.g., in geophysics or chemistry,
- $f(z) = \sqrt{tz}$: simulation of Brownian motion of molecules or sampling from Gaussian Markov random fields,
- $f(z) = \mathrm{sign}(z)$: simulations in quantum chromodynamics.

Rational Krylov Spaces

Definition: Given a sequence of polynomials

$$q_{m-1}(z) = \prod_{\substack{j=1 \\ \xi_j \neq \infty}}^{m-1} (z - \xi_j), \quad m = 1, 2, \ldots,$$

where $\xi_j \in \overline{\mathbb{C}} \setminus \Lambda(A)$. Then the associated rational Krylov spaces of order m are defined as

$$\mathscr{Q}_m(A, \boldsymbol{b}) := q_{m-1}(A)^{-1} \mathscr{K}_m(A, \boldsymbol{b}),$$

where $\mathscr{K}_m(A, \boldsymbol{b}) = \mathrm{span}\{A^0\boldsymbol{b}, A^1\boldsymbol{b}, \ldots, A^{m-1}\boldsymbol{b}\}$.

Rational Ritz values

... are the eigenvalues of the Rayleigh quotient $A_m = Q_m^* A Q_m$, denoted by $\Theta = \{\theta_1, \ldots, \theta_m\}$.

Let A be Hermitian. Then the θ_k's lie in the spectral interval of A and *interlace* the eigenvalues $\Lambda(A) = \{\lambda_1, \ldots, \lambda_N\}$:

(*) In any interval $(\theta_\kappa, \theta_{\kappa+1})$ there is at least one eigenvalue λ_k of A.

Moreover, the rational Ritz values are zeros of orthogonal rational functions and may be characterized as (see, e.g., [2, 3])

(**) The θ_k's are the zeros of the minimizer of $\| p(A) q_{m-1}(A)^{-1} \boldsymbol{b} \|$ among all monic $p \in \mathscr{P}_m^\infty$.

Logarithmic potential theory can explain the *asymptotic* distribution of the rational Ritz values. Therefore we consider

- a sequence of Hermitian matrices $\{A_N\}$, each of size $N \times N$, whose eigenvalue counting measures converge to a Borel probability measure σ in the weak-* sense,
- a sequence of vectors $\{\boldsymbol{b}_N\}$, each of length N,
- a ray sequence of integers $\{m_N\}$ such that

$$m_N / N \to t \in (0,1) \quad \text{as } N \to +\infty,$$

- a sequence of polynomials $\{q_N\}$, each of degree $m_N - 1$, whose zero counting measures converge to a Borel measure ν, $\|\nu\| = t$,
- the sequence $\{\Theta_N\}$ of rational Ritz values of order m_N.

Tools from Potential Theory Associated with a (signed) Borel measure μ_1 is the logarithmic potential

$$U^{\mu_1}(z) := \int \frac{1}{\log|x - z|} \mathrm{d}\mu_1(x).$$

Inexact solves & error estimators

In each iteration of the rational Arnoldi method a linear system of the form $(A - \xi_j I)\boldsymbol{x}_j = \boldsymbol{q}_j$ is solved. If the residuals are collected in a matrix R_m, then (1) becomes

$$AQ_{m+1}\underline{K_m} = Q_{m+1}\underline{H_m} + R_m. \qquad (2)$$

Setting $E_m := -R_m \underline{K_m}^\dagger Q_{m+1}^*$, we observe that we have computed an exact Arnoldi decomposition

$$(A + E_m) Q_{m+1} \underline{K_m} = Q_{m+1} \underline{H_m}$$

for the matrix $A + E_m$. The Rayleigh quotient $\widetilde{A}_m$ computed from the data $\underline{K_m}$ and $\underline{H_m}$ satisfies

$$\begin{aligned} \widetilde{A}_m &= Q_m^*(A + E_m)Q_m \\ &= Q_m^* A Q_m + Q_m^*(-R_m \underline{K_m}^\dagger Q_{m+1}^*) Q_m \\ &= \widehat{A}_m - Q_m^* R_m \underline{K_m}^\dagger \underline{I_m}. \end{aligned}$$

Here, $\widehat{A}_m := Q_m^* A Q_m$ is referred to as the *corrected Rayleigh quotient*, because it is a compression of A instead of $A + E_m$. It can be computed from $\widetilde{A}_m$ without explicit projection, only by additional inner-products $Q_m^* R_m$.

We now decompose the error

$$\| f(A)\boldsymbol{b} - \boldsymbol{f}_m \| \leq \underbrace{\| f(A)\boldsymbol{b} - f(A + E_m)\boldsymbol{b} \|}_{\text{sensitivity error}} + \underbrace{\| f(A + E_m)\boldsymbol{b} - \boldsymbol{f}_m \|}_{\text{approximation error}},$$

and estimate

$$\text{sensitivity error} \approx \| f(\widetilde{A}_m) Q_m^* \boldsymbol{b} - f(\widehat{A}_m) Q_m^* \boldsymbol{b} \|.$$

It is advisable to terminate the rational Arnoldi method if the ap-