Stefan Güttel

Rational approximation III+IV: Computational methods and applications

MANCHESTER 1824

My Woudschoten experience in 2009

Overview

I will focus on some practical aspects of rational approximation, and discuss applications to problems in numerical linear algebra and scientific computing.

Main interest will be on rational Krylov methods but I will also mention algorithms for scalar rational approximation in the second hour.

Rational Krylov

Rational Krylov = combination of rational approximation and numerical linear algebra.

Natural extension of scalar rational approximation to matrices and vectors.

Many computational methods based on **rational Krylov spaces** (even if only implicitly).

First introduced by Axel Ruhe (LAA 1984):

Rational Krylov Sequence Methods for Eigenvalue Computation

Algorithms to solve large sparse eigenvalue problems are considered. A new class of algorithms which is based on rational functions of the matrix is described. The Lanczos method, the Arnoldi method, the spectral transformation Lanczos method, and Rayleigh quotient iteration all are special cases, but there are also new algorithms which correspond to rational functions with several poles. In the simplest case a basis of a rational Krylov subspace is found in which the matrix eigenvalue problem is formulated as a linear matrix pencil with a pair of Hessenberg matrices.

From scalars to vectors and matrices

Given a scalar rational function

$$r(z) = \frac{p(z)}{q(z)},$$

a matrix $A \in \mathbb{C}^{N \times N}$ and a vector $\boldsymbol{b} \in \mathbb{C}^N$. Assume that $q(A)$ is nonsingular, then

$$r(A)\boldsymbol{b} = q(A)^{-1} p(A)\boldsymbol{b}.$$

It is useful to consider the linear span of such vectors for $p \in \mathscr{P}_{m-1}$.

Definition (rational Krylov space)

Given $q_{m-1} \in \mathscr{P}_{m-1}$ such that $q_{m-1}(A)$ is nonsingular, we define

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

Note that q_{m-1} is implicit in our notation of $\mathscr{Q}_m(A, \boldsymbol{b})$.

Rational Krylov space: properties

$$\mathcal{Q}_m(A, \boldsymbol{b}) := q_{m-1}(A)^{-1} \operatorname{span}\{\boldsymbol{b}, A\boldsymbol{b}, \ldots, A^{m-1}\boldsymbol{b}\}$$

- $\mathcal{Q}_m(A, \boldsymbol{b}) \subseteq \mathbb{C}^N$
- there exists *invariance index* $L \leq N$ for $(A, \boldsymbol{b})$ such that $\dim(\mathcal{Q}_m) = \min\{m, L\}$
- $\mathcal{Q}_m(A, \boldsymbol{b}) = \mathcal{K}_m(A, q_{m-1}(A)^{-1}\boldsymbol{b}) =$ polynomial Krylov space for $q_{m-1}(A)^{-1}\boldsymbol{b}$

We can generally take q_{m-1} to be of the form

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

for a *sequence of poles* $\xi_1, \xi_2, \ldots \in \overline{\mathbb{C}}$. Then $\mathcal{Q}_1 \subset \mathcal{Q}_2 \subset \cdots \subset \mathcal{Q}_L = \mathcal{Q}_{L+1} = \cdots$

Rational Krylov spaces: examples

- Polynomial: if all $\xi_j = \infty$, then $\mathscr{Q}_m(A, \boldsymbol{b}) = \mathscr{K}_m(A, \boldsymbol{b})$
- Shift-invert: if $\xi_j = \sigma$ fixed, then $\mathscr{Q}_m(A, \boldsymbol{b}) = \mathscr{K}_m((A - \sigma I)^{-1}, \boldsymbol{b})$
- Extended Krylov: if $\xi_j \in \{\infty, 0\}$ alternating, i.e.,

$$\mathscr{Q}_m(A, \boldsymbol{b}) = \operatorname{span}\{\boldsymbol{b}, A\boldsymbol{b}, A^{-1}\boldsymbol{b}, A^2\boldsymbol{b}, A^{-2}\boldsymbol{b}, \ldots\}$$

- Partial fractions: if all ξ_j distinct and finite, then

$$\mathscr{Q}_m(A, \boldsymbol{b}) = \operatorname{span}\{\boldsymbol{b}, (A - \xi_1 I)^{-1}\boldsymbol{b}, (A - \xi_2 I)^{-1}\boldsymbol{b}, \ldots, (A - \xi_{m-1} I)^{-1}\boldsymbol{b}\}$$

Back to scalar rational approximation

Theorem

Given a pair $(A, \boldsymbol{b})$ and a pole sequence $\xi_1, \ldots, \xi_{m-1}$. Assume that $\dim(\mathcal{Q}_m(A, \boldsymbol{b})) = m$. Then every vector $\boldsymbol{v} \in \mathcal{Q}_m(A, \boldsymbol{b})$ is in one-to-one correspondence with a rational function

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

such that

$$\boldsymbol{v} = r_m(A)\boldsymbol{b}, \quad p_{m-1} \in \mathscr{P}_{m-1}, \quad q_{m-1} = \prod_{\substack{j=1 \\ \xi_j \neq \infty}}^{m-1} (z - \xi_j).$$

Example: rational approximation with prescribed denominator

Aim: Find ratfun r_m with negative poles s.t. $r_m(x) \approx f(x) = \exp(-x)$ for $x \in [0, 1000]$

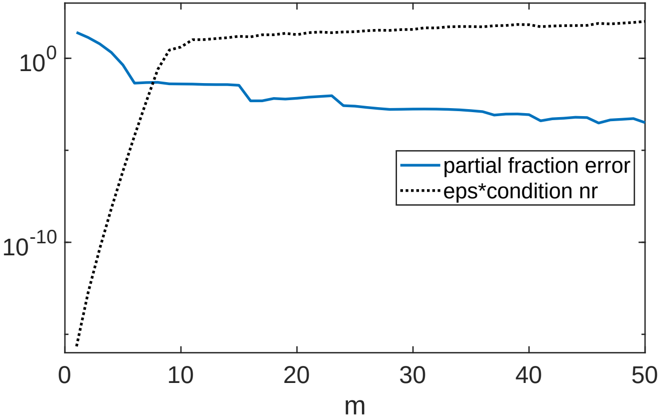
Attempt 1:

```
N = 1e4; x = logspace(-3,3,N).';
f = exp(-x);
xi = linspace(-10,-1,50);
C = 1./(x - xi);

for j = 1:length(xi)
    coeffs = C(:,1:j)\f;
    err(j) = norm(f - C(:,1:j)*coeffs);
    cnd(j) = cond(C(:,1:j));
end

semilogy(err), hold on, semilogy(cnd,'k:')
```

The basis matters



Partial fraction basis results in Cauchy matrix $C(i,j) = 1/(x_i - \xi_j)$ with exponentially growing condition number (Beckermann & Townsend 2019). Growth factor known in terms of $\mathrm{cap}(\Sigma, \Xi)$.

Aim: Construct *orthonormal* rational Krylov basis, one vector at a time

Input: matrix $A \in \mathbb{C}^{N \times N}$, vector $\boldsymbol{b} \in \mathbb{C}^{N}$, pole sequence $\xi_1, \dots, \xi_m$

$\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$

for $j = 1 : m$

$\quad \boldsymbol{w} = (A - \xi_j I)^{-1} \boldsymbol{v}_j$

$\quad$ for $i = 1 : j$

$\quad\quad h_{i,j} = \boldsymbol{v}_i^* \boldsymbol{w}$

$\quad\quad \boldsymbol{w} = \boldsymbol{w} - \boldsymbol{v}_i h_{i,j}$

$\quad h_{j+1,j} = \|\boldsymbol{w}\|$

$\quad \boldsymbol{v}_{j+1} = \boldsymbol{w}/h_{j+1,j}$

Output: Orthonormal $V_{m+1} = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_{m+1}]$, coefficients $\underline{H_m} = [h_{i,j}]$

Rational approximation with prescribed poles: revisited

Aim: Find ratfun r_m with negative poles s. t. $r_m(x) \approx f(x) = \exp(-x)$ for $x \in [0, 1000]$

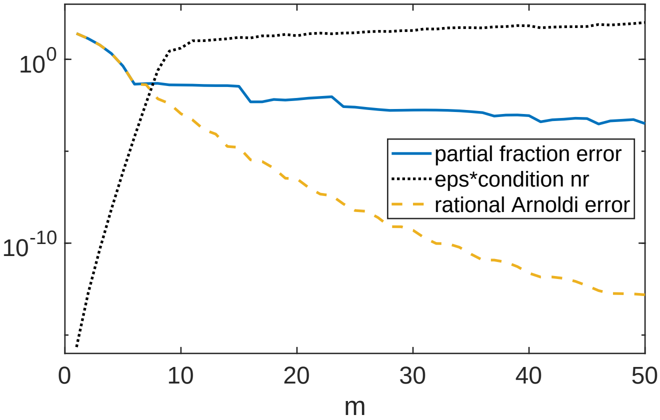
Attempt 2:

```
N = 1e4; x = logspace(-3,3,N).';
A = diag(x); b = ones(N,1);      % define (A,b)
f = exp(-x);                     % function to approximate
xi = linspace(-10,-1,50);
[V,H] = ratarnoldi(A,b,xi);

for j = 1:length(xi)
    coeffs = V(:,1:j)'*f;
    ratfun = V(:,1:j)*coeffs;  % = r(A)b
    err(j) = norm(f - ratfun);
end

semilogy(err)
```

Rational approximation with prescribed poles: revisited



Arnoldi process fixes stability issue, but how do we evaluate $r_m(z)$?

Given $(A, \boldsymbol{b})$ and poles ξ_j, orthonormal rational Krylov basis V_{m+1}, and coefficients

$$\texttt{coeffs} = V_{m+1}^* \boldsymbol{f}.$$

Then $V_{m+1} \cdot \texttt{coeffs} = r_m(A)\boldsymbol{b}$ with a unique rational function $r_m(z)$.

Idea: Rerun Arnoldi for arbitrary $(\widehat{A}, \widehat{\boldsymbol{b}})$ using the previous coefficients $\underline{H_m}$ and poles ξ_j. This will return $\widehat{V}_m$. Then form $r_m(\widehat{A})\widehat{\boldsymbol{b}} = \widehat{V}_{m+1} \cdot \texttt{coeffs}$.

Special case: scalar evaluation with $\widehat{A} = [z]$, $\widehat{\boldsymbol{b}} = [1]$.

(Basis of the `feval` method in RKToolbox. Demo of that later.)

Rational Krylov for matrix functions

Given a pair $(A, \boldsymbol{b})$ and a scalar function f, we now consider the approximation of $f(A)\boldsymbol{b}$, where $f(A)$ is a **matrix function.** (Higham 2008)

Definition (matrix function)

Given scalar function $f(z)$ and *diagonalizable* matrix $A = XDX^{-1}$, $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$. Then $f(A) := Xf(D)X^{-1}$, where $f(D) := \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_N))$.

In many applications, A is large and sparse and we cannot compute $f(A)$ explicitly.

Solution: Use Arnoldi process to compute orthonormal rational Krylov basis V_m for $(A, \boldsymbol{b})$ and form the *(rational) Arnoldi approximation*

$$\boldsymbol{f}_m = V_m f(A_m) V_m^* \boldsymbol{b}, \quad A_m = V_m^* A V_m.$$

If $m \ll N$, computational cost dominated by m solves of linear systems $(A - \xi_i I)\boldsymbol{w}_i = \boldsymbol{v}_i$.

Arnoldi approximation to $f(A)\boldsymbol{b}$

$$\boldsymbol{f}_m = V_m f(A_m) V_m^* \boldsymbol{b}, \quad A_m = V_m^* A V_m$$

Theorem (interpolation)

We have $\boldsymbol{f}_m = r_m(A)\boldsymbol{b}$, where $r_m = p_{m-1}/q_{m-1}$ interpolates f at the eigenvalues of A_m (the rational Ritz values), and $q_{m-1}(z) = \prod_{j=1}^{m-1}(z - \xi_j)$.

Theorem (near-optimality)

Assume f is analytic in a neighborhood of the numerical range $\mathbb{W}(A) := \{\boldsymbol{v}^* A \boldsymbol{v} : \|\boldsymbol{v}\|_2 = 1\}$. Then

$$\|f(A)\boldsymbol{b} - \boldsymbol{f}_m\|_2 \leq 5\|\boldsymbol{b}\|_2 \min_{p \in \mathscr{P}_{m-1}} \left\| f(z) - \frac{p(z)}{q_{m-1}(z)} \right\|_{\mathbb{W}(A)}.$$

Ericsson '90, Saad '92, Druskin-Knizhnerman '98, Beckermann-Reichel '09, G. '10, Crouzeix-Palencia '17, ...

Pole optimization

Near-optimality $\Longrightarrow$ focus on finding “good” poles ξ_j to make $\|f(A)\boldsymbol{b} - \boldsymbol{f}_m\|_2$ small.

Essentially three approaches.

- **Analytic approach:** Assuming knowledge of $\mathbb{W}(A)$, construct scalar rational approximant (e.g., using Zolotarev functions, Faber transform) and use its poles as ξ_j parameters.
- **Scalar numerical:** Use a method like Remez or AAA to compute best or near-best rational approximant $r_m \approx f$ on (a discretized version of) $\mathbb{W}(A)$. Use its poles.
- **Adaptive rational Krylov:** Methods that choose their own pole parameters (greedy rational Krylov, IRKA, RKFIT).

RKToolbox demo

Demonstrate rational Arnoldi approximation of $A^{-1/2}\boldsymbol{b}$ using RKToolbox.

We will use Zolotarev approximant $r_m(z) \approx z^{-1/2}$ on spectral interval E.

This approximant is explicitly known in terms of elliptic functions.

Rel. appoximation error decreases like $\exp(-2/\mathrm{cap}(E,F))^m$, where $F = (-\infty, 0]$.

We evaluate $r_m(A)\boldsymbol{b}$ and check the abs. error $\|A^{-1/2}\boldsymbol{b} - r_m(A)\boldsymbol{b}\|_2$ as a function of m.

```
%% install RKToolbox
unzip('http://guettel.com/rktoolbox/rktoolbox.zip');
cd('rktoolbox'); addpath(fullfile(cd)); savepath

%% construct and evaluate RKFUN object
r = rkfun('invsqrt',5,1e4);    % degree 5 Zolotarev
r(8)                           % scalar evaluation
r(A,b)                         % compute r(A)*b
```

RKToolbox demo

Poles of Zolotarev approximant $r_m(z) \approx z^{-1/2}$ are **not nested.**

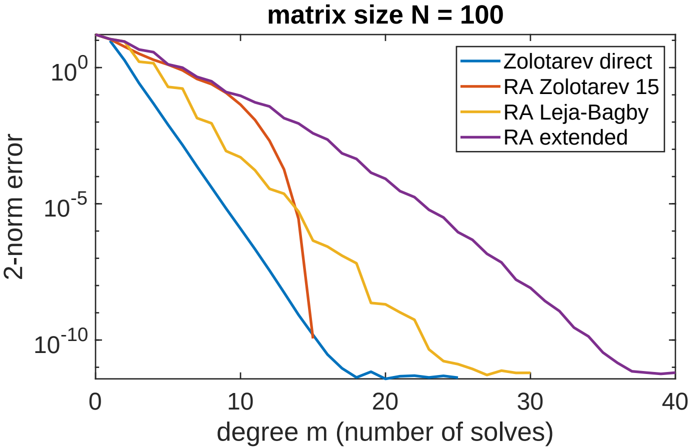
I.e., all m poles change when m increases.

In practice, we prefer nested pole sequences as we can simply add another pole (basis vector) to the rational Krylov space if our Arnoldi approximation $\boldsymbol{f}_m$ isn't good enough.

For the case of $z^{-1/2}$ on interval $E = [\lambda_{\min}, \lambda_{\max}] > 0$, we have some natural choices:

- **Leja–Bagby:** nested discretization of equilibrium measure on (E, F)
  $\Longrightarrow$ may require about twice as large m as Zolotarev
- **repeated single pole:** $\xi_j = -\sqrt{\lambda_{\min}\lambda_{\max}}$
  $\Longrightarrow$ expected linear convergence at rate $\frac{\sqrt[4]{\kappa}-1}{\sqrt[4]{\kappa}+1}$, $\kappa = \lambda_{\max}/\lambda_{\min}$
- **extended Krylov:** $\xi_j \in \{0, \infty\}$ alternating
  $\Longrightarrow$ same $\frac{\sqrt[4]{\kappa}-1}{\sqrt[4]{\kappa}+1}$ rate but only half as many solves

Convergence of rational Arnoldi



Approximating $A^{-1/2}\boldsymbol{b}$ using direct evaluation $r_m(A)\boldsymbol{b}$ of Zolotarev approximant vs rational Arnoldi approximant with different pole sequences

A closer look at the error behavior

For symmetric $A = U\Lambda U^*$, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$, we have the following

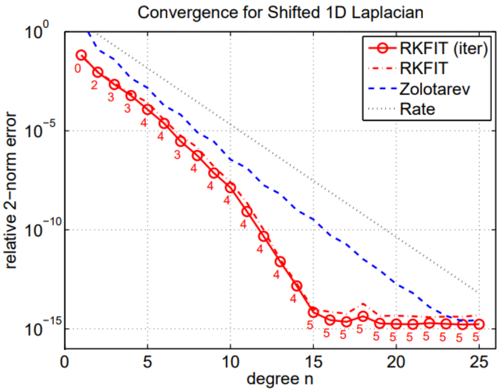
$$\begin{aligned} \|f(A)\boldsymbol{b} - r_m(A)\boldsymbol{b}\|_2^2 &= \sum_{i=1}^{N} |f(\lambda_i) - r_m(\lambda_i)|^2 \cdot |\boldsymbol{u}_i^* \boldsymbol{b}|^2 \\ &\le \|\boldsymbol{b}\|_2^2 \cdot \max_{x \in [\lambda_{\min}, \lambda_{\max}]} |f(x) - r_m(x)|^2. \end{aligned}$$

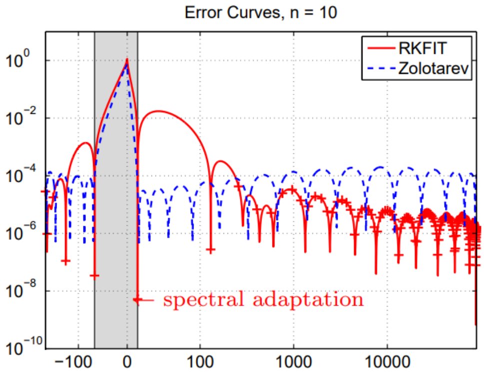
Uniform best approximant r_m would minimize blue term, but the red error we're actually interested in could be smaller with another r_m.

Recall that the r_m underlying the **rational Arnoldi approximant** is a rational function with the chosen poles ξ_j that **interpolates** f **at the Ritz values** $\Lambda(V_m^* A V_m)$.

The function r_m is therefore not necessarily a good uniform approximant; it actually adapts to the discrete spectrum of A, leading to superlinear convergence! This is called **spectral adaptation.**

Spectral adaptation





Left: Superlinear convergence of rational Krylov approximant $r_n(A)\boldsymbol{b}$ for $f(A)\boldsymbol{b}$ with $f(z) = \sqrt{z + (hz/2)^2}$ for an indefinite shifted 1D Laplacian A, compared to uniform Zolotarev approximant on negative and positive spectral subinterval. Right: Error curves $|f(z) - r_{10}(z)|$.

Rational least squares fitting

RKFIT

Aims to compute a rational function r_m such that

$$\|F\boldsymbol{b} - r_m(A)\boldsymbol{b}\|_2^2 \rightsquigarrow \min, \quad F = f(A),$$

with the minimum taken over all rational functions $r_m(z) = \frac{p_m(z)}{q_m(z)}$.

This is a nonlinear nonconvex approximation problem.

We know that $r_m(A)\boldsymbol{b}$ is an element of some rational Krylov space

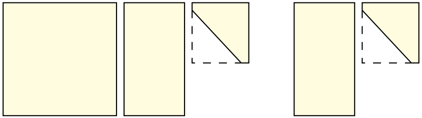
$$\mathscr{Q}_{m+1}(A, \boldsymbol{b}) = \mathrm{span}\{\boldsymbol{b}, (A - \xi_1 I)^{-1}\boldsymbol{b}, \dots, (A - \xi_m I)^{-1}\boldsymbol{b}\},$$

but what are good pole parameters $\xi_1, \dots, \xi_m$?

Rational least squares approximation

For every rational Krylov space $\mathscr{Q}_{m+1}(A, \boldsymbol{b})$ there exists an orthonormal basis V_{m+1} that satisfies a rational Arnoldi decomposition:

$$A \quad V_{m+1} \; \underline{K_m} \; = V_{m+1} \; \underline{H_m}$$



where

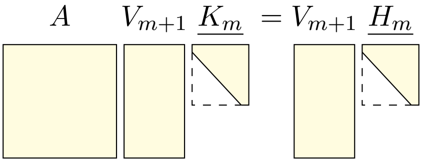
- $\underline{H_m}, \underline{K_m}$ are $(m+1) \times m$ upper-Hessenberg matrices
- subdiagonal quotients $h_{j+1,j}/k_{j+1,j}$ are the poles ξ_j
- first column of V_{m+1} is $\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2$

Can show: One-to-one correspondence between $\boldsymbol{v}_1$ and $\xi_1, \ldots, \xi_m$.

Rational implicit Q theorem[1]

Given rational Arnoldi decomposition $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with $\xi_j := h_{j+1,j}/k_{j+1,j}$. Then the orthonormal basis V_{m+1} and $(\underline{H_m}, \underline{K_m})$ are essentially uniquely determined by $\boldsymbol{v}_1$ and the poles $\xi_1, \ldots, \xi_m$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{m+1}:



[1] Berljafa & G. 2015, Camps-Meerbergen-Vandebril 2019, ...

Rational implicit Q theorem[1]

Given rational Arnoldi decomposition $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$

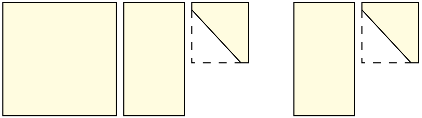
with $\xi_j := h_{j+1,j}/k_{j+1,j}$. Then the orthonormal basis V_{m+1} and

$(\underline{H_m}, \underline{K_m})$ are essentially uniquely determined by $\boldsymbol{v}_1$ and the poles $\xi_1, \ldots, \xi_m$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{m+1}:

insert unitary

$$A \quad V_{m+1} \underbrace{}^{P_{m+1}P_{m+1}^*} \underline{K_m} = V_{m+1} \underbrace{}^{P_{m+1}P_{m+1}^*} \underline{H_m}$$



[1]Berljafa & G. 2015, Camps-Meerbergen-Vandebril 2019, ...

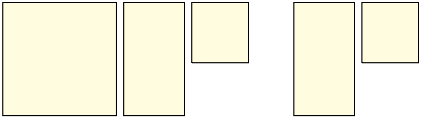
Rational implicit Q theorem[1]

Given rational Arnoldi decomposition $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with $\xi_j := h_{j+1,j}/k_{j+1,j}$. Then the orthonormal basis V_{m+1} and $(\underline{H_m}, \underline{K_m})$ are essentially uniquely determined by $\boldsymbol{v}_1$ and the poles $\xi_1, \ldots, \xi_m$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{m+1}:

rotate basis to $\widetilde{V}_{m+1} = V_{m+1}P_{m+1}$

$$A \quad \widetilde{V}_{m+1} \, \underline{\widetilde{K}_m} = \widetilde{V}_{m+1} \, \underline{\widetilde{H}_m}$$



[1] Berljafa & G. 2015, Camps-Meerbergen-Vandebril 2019, . . .

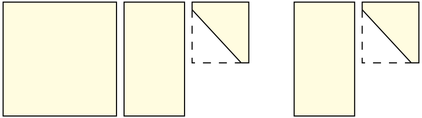
Rational implicit Q theorem[1]

Given rational Arnoldi decomposition $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with $\xi_j := h_{j+1,j}/k_{j+1,j}$. Then the orthonormal basis V_{m+1} and $(\underline{H_m}, \underline{K_m})$ are essentially uniquely determined by $\boldsymbol{v}_1$ and the poles $\xi_1, \ldots, \xi_m$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{m+1}:

QZ transform on lower $m \times m$ part of $(\underline{\widehat{H}_m}, \underline{\widehat{K}_m})$

$$A \quad \widehat{V}_{m+1}\,\underline{\widehat{K}_m} = \widehat{V}_{m+1}\,\underline{\widehat{H}_m}$$



[1]Berljafa & G. 2015, Camps-Meerbergen-Vandebril 2019, ...

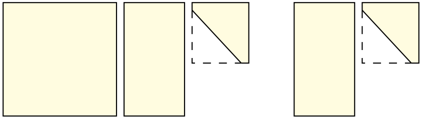
Rational implicit Q theorem[1]

Given rational Arnoldi decomposition $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with $\xi_j := h_{j+1,j}/k_{j+1,j}$. Then the orthonormal basis V_{m+1} and $(\underline{H_m}, \underline{K_m})$ are essentially uniquely determined by $\boldsymbol{v}_1$ and the poles $\xi_1, \ldots, \xi_m$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{m+1}:

QZ transform on lower $m \times m$ part of $(\underline{\widehat{H}_m}, \underline{\widehat{K}_m})$

$$A \quad \widehat{V}_{m+1} \; \underline{\widehat{K}_m} \; = \; \widehat{V}_{m+1} \; \underline{\widehat{H}_m}$$



Read off new poles $\widehat{\xi}_j := \widehat{h}_{j+1,j}/\widehat{k}_{j+1,j}$

[1] Berljafa & G. 2015, Camps-Meerbergen-Vandebril 2019, ...

Rational least squares fitting with RKFIT

These pole transformations can be used to approximately solve

$$\|F\boldsymbol{b} - r_m(A)\boldsymbol{b}\|_2^2 \rightsquigarrow \min, \quad F = f(A).$$

Idea: Given a set of poles $\xi_1, \ldots, \xi_m$ and the corresponding Krylov basis V_{m+1}, find a unit vector $\boldsymbol{v} \in \mathrm{span}(V_{m+1})$ so that $F\boldsymbol{v}$ is well approximated by some $r_m(A)\boldsymbol{v} \in \mathrm{span}(V_{m+1})$.

Solution:

$$\operatorname{arg\,min}_{\boldsymbol{c}\in\mathbb{C}^{m+1}, \|\boldsymbol{c}\|=1} \|(I - V_{m+1}V_{m+1}^*)FV_{m+1}\boldsymbol{c}\|_2$$

Bring $\boldsymbol{v}$ to the first column of V_{m+1} and read off new poles of $r_m(A)\boldsymbol{v}$. Iterate.

Theory: Can show that $f \in \mathscr{P}_m/q_m$ will be identified in one iteration. (Berljafa-G. 2017)

Application: Compressing layered waveguides

Druskin et al. 2022

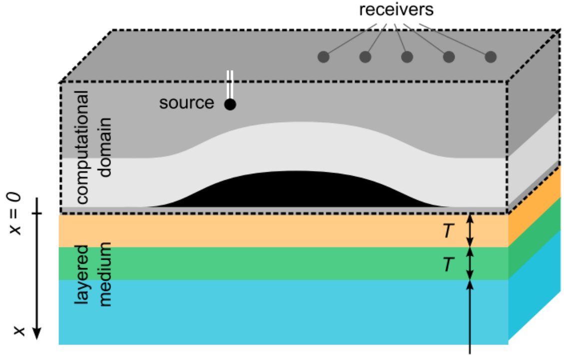
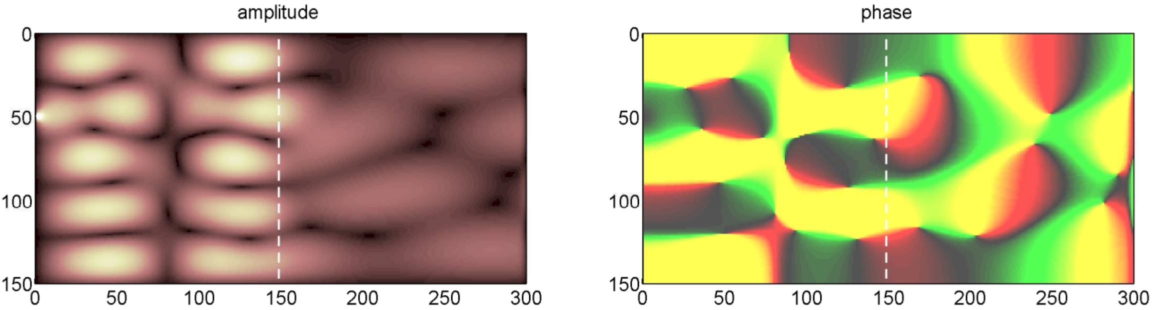


Fig. 5 Typical setup of a seismic exploration of the Earth's subsurface. It is of practical interest to compress the layered medium in $x \geq 0$ into a single PML with a small number of grid points

Application: Compressing layered waveguides

Semi-discretization of indefinite Helmholtz problem on $[0, +\infty)$:

$$\boldsymbol{u}''(x) = [\, A + c(x) I \,] \boldsymbol{u}(x), \quad \boldsymbol{u}(0) = \boldsymbol{b}, \quad \boldsymbol{u}(x) \text{ bounded as } x \to \infty.$$



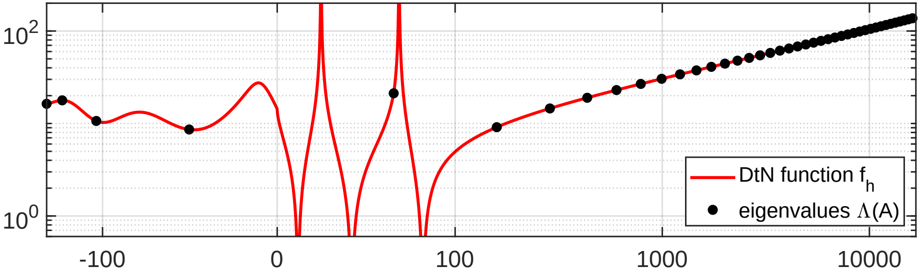
Discretization with 300×150 grid points + PML on the right, $h = 1/150$, effective wavenumber on the left half is $k = 16$, on the right half is $k = 9$

Application: Compressing layered waveguides

Semi-discretization of indef. Helmholtz problem on $[0, +\infty)$:

$$\boldsymbol{u}''(x) = [\,A + c(x)I\,]\,\boldsymbol{u}(x), \quad \boldsymbol{u}(0) = \boldsymbol{b}, \quad \boldsymbol{u}(x) \text{ bounded as } x \to \infty.$$

Can write DtN map as $\boxed{\boldsymbol{u}'(0) = f(A)\boldsymbol{b}}$ BUT $f(z)$ highly irregular near $\Lambda(A)$:



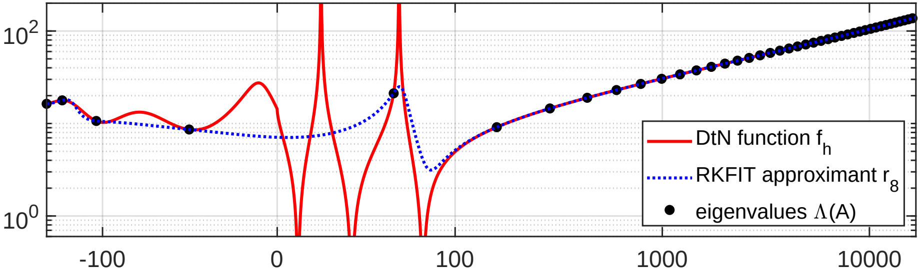
There is no way $r_m \approx f$ uniformly on A's spectral interval, but ...

Application: Compressing layered waveguides

Semi-discretization of indef. Helmholtz problem on $[0, +\infty)$:

$$\boldsymbol{u}''(x) = [\,A + c(x)I\,]\boldsymbol{u}(x), \quad \boldsymbol{u}(0) = \boldsymbol{b}, \quad \boldsymbol{u}(x) \text{ bounded as } x \to \infty.$$

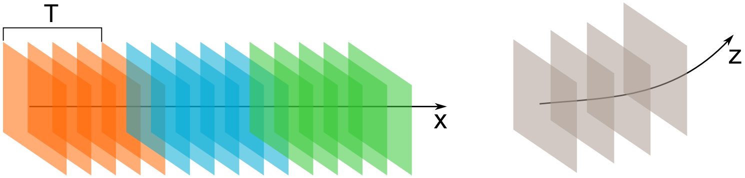
Can write DtN map as $\boxed{\boldsymbol{u}'(0) = f(A)\boldsymbol{b}}$ BUT $f(z)$ highly irregular near $\Lambda(A)$:



There is no way $r_m \approx f$ uniformly on A's spectral interval, but we can still compute RKFIT approximant due to spectral adaptation.

Application: Compressing layered waveguides

Rational approximant r_m of DtN map can be converted into equivalent finite difference representation acting as perfectly matched layer (PML):



As consequence of spectral adaptation, fewer than 2 grid points per wavelength required:

	$T = 0.25$	$T = 0.5$	$T = 1$	$T = 2$
Nyquist minimum N	8.75	17.5	35	70
SEM minium $\frac{\pi}{2}$N	13.7	27.5	55.0	110.0
RKFIT-FD	8	10	16	19

Adaptive rational Krylov methods

There is large body of literature on methods that choose their pole parameters; e.g.:

- **greedy parameter selection for model order reduction**
  Grimme, 1997; Frangos & Jaimoukha, **Proc. EEC** 2007; Druskin & Simoncini, **Syst. Control. Lett.** 2011; Benner, Gugercin, Willcox, **SIAM Rev.** 2015; Frie & Eberhard **Multibody Sys. Dyn.** 2023,...
- **greedy pole selection for matrix functions**
  Druskin, Lieberman, Zaslavsky, **SIAM J. Sci. Comp.** 2010; G. & Knizhnerman, **BIT Numer. Math.** 2013
- **automatic approximation of nonlinear eigenproblems**
  Lietaert, Pérez, Vandereycken, Meerbergen, **IMA J. Numer. Anal.** 2021
- **iterated poles for $\mathscr{H}_2$-optimal model reduction (IRKA)**
  Gugercin, Antoulas, Beattie, **SIAM J. Matrix Anal. Appl.** 2008; Flagg, Beattie, Gugercin, **Syst. Control. Lett.** 2012; Borghi & Breiten, **Adv. Comp. Math.** 2024; Aumann Werner, **Adv. Comp. Math.** 2024
- **nonlinear rational least squares fitting (vector fitting, RKFIT)**
  Gustavsen & Semlyen, **IEEE Trans. Power Del.** 1999, Berljafa & G., **SIAM J. Sci. Comp.** 2017

Algorithms for scalar rational approximation

Many interesting problems require merely scalar-valued rational approximation $r_m(z) \approx f(z)$.

If approximation is around a single-point $z = z_0$ in the complex plane, **Padé approximation** is a popular approach (matching derivatives $f^{(i)}(0) \approx r_m^{(i)}(z_0)$, $i = 0, 1, \ldots$). One can also construct **multi-point Padé approximants.**

If the approximation is on a real interval and f is continuous, there is a unique best rational approximant. Good implementation of **Remez algorithm** in Chebfun.

For scalar approximation on discrete sets, the **AAA algorithm** is a very powerful tool.

Stands for adaptive Antoulas–Anderson and uses barycentric interpolatory representation

$$r_m(z) = \frac{\displaystyle\sum_{i=0}^{m} \frac{w_i f_i}{z - z_i}}{\displaystyle\sum_{i=0}^{m} \frac{w_i}{z - z_i}} = \frac{n_m(z)}{d_m(z)}$$

with distinct support points z_i and nonzero weights w_i. As $z \to z_i$ we have $r_m(z) \to f_i$.

Key idea: Greedily add interpolation point z_{m+1} from a discrete set Σ such that

$$|f(z_{m+1}) - r_m(z_{m+1})| = \max_{z \in \Sigma} |f(z) - r_m(z)|.$$

Then compute new weights $w_0, w_1, \ldots, w_{m+1}$ of r_{m+1} by solving

$$\|f(\Sigma) d_{m+1}(\Sigma) - n_{m+1}(\Sigma)\|_2 \to \min_{\{w_i\}} .$$

Variants of AAA

Over the last five years, many variations of AAA have been proposed, e.g., AAA-Lawson, fastAAA, set-valued AAA, block-AAA, parametric-AAA, sketch-AAA, . . .

Has been very popular in particular in the model order reduction community.

Can also be applied to solve problems of the following form:

Nonlinear eigenvalue problem (NEP)

Given a holomorphic matrix-valued function $F(z) : \mathbb{C} \to \mathbb{C}^{N\times N}$, find points $\lambda \in \mathbb{C}$ (the eigenvalues) such that $F(\lambda)$ is singular.

NEPs arise in many applications incl. structural mechanics, delay-differential equations, ROMs with nonlinear parameter dependencies, etc. They can be large-scale.

Nonlinear eigenvalue problems (NEPs)

There are many techniques to solve NEPs numerically, such as Newton-based methods, contour-integral methods, but also methods based on rational approximation

$$R_m(z) \approx F(z) \qquad \text{for } z \in \Sigma \text{ compact.}$$

Then the NEP is solved for R_m in place of F. This is "easy" because R_m can be linearized.

It is crucial that $R_m \approx F$ is a **uniformly accurate rational approximation.**
For example, we may want to impose that

$$\|F - R_m\|_\Sigma := \max_{z \in \Sigma} \|F(z) - R_m(z)\|_2 \leq \varepsilon,$$

because then the eigenvalues of R_m in Σ can be guaranteed to be approximations to some of the eigenvalues of F. This can be seen as follows.

Uniform rational approximation $\Longrightarrow$ good eigenvalue approximation

Assume that $(\lambda, \boldsymbol{v})$ with $\lambda \in \Sigma$ and $\|\boldsymbol{v}\|_2 = 1$ is an eigenpair of R_m, i.e., $R_m(\lambda)\boldsymbol{v} = 0$. Then from

$$\|F(\lambda)\boldsymbol{v}\|_2 = \|\big(F(\lambda) - R_m(\lambda)\big)\boldsymbol{v}\|_2 \leq \|F(\lambda) - R_m(\lambda)\|_2 \leq \varepsilon$$

we find that $(\lambda, \boldsymbol{v})$ has a bounded residual for the original NEP $F(\lambda)\boldsymbol{v} = 0$.

Conversely, if $\mu \in \Sigma$ is *not* an eigenvalue of F, i.e., $F(\mu)$ is nonsingular, then a sufficiently accurate approximant R_m is also nonsingular at μ:

Assume that $\|F(\mu) - R_m(\mu)\|_2 < \|F(\mu)^{-1}\|^{-1}$, then

$$\|I - F(\mu)^{-1}R_m(\mu)\|_2 \leq \|F(\mu)^{-1}\|_2 \|F(\mu) - R_m(\mu)\|_2 < 1.$$

Hence all eigenvalues of $I - F(\mu)^{-1}R_m(\mu)$ are strictly smaller in modulus than 1.
As a consequence, $F(\mu)^{-1}R_m(\mu)$ and hence $R_m(\mu)$ are nonsingular.

Ideally, R_m does not have any eigenvalues in Σ which are in the resolvent set of F.
In this case we say that R_m is free of *spurious* eigenvalues on Σ.

Solving NEPs via rational approximation methods

Assume split form $F(z) = f^{(1)}(z)A_1 + \cdots + f^{(\ell)}(z)A_\ell$ with fixed matrices A_i.

Key steps:

1. Uniformly approximate scalar functions $f^{(i)}$ by $r_m^{(i)}$ all sharing the same denominator
2. Linearize the rational NEP $R_m(z) = r_m^{(1)}(z)A_1 + \cdots + r_m^{(\ell)}(z)A_\ell$
3. Solve generalized linear eigenvalue problem $A_{Nm}\boldsymbol{x} = \lambda B_{Nm}\boldsymbol{x}$

For Step 1, use the set-valued AAA method (Lietaert-Meerbergen-Pérez-Vandereycken 2022) or a randomized probing approach (G.-Kressner-Vandereycken 2024).

For Step 3, we can use a rational Krylov method as originally proposed by Ruhe in 1984.

See example in RKToolbox
`http://guettel.com/rktoolbox/examples/html/example_nlep.html`

Summary

- Rational Krylov $=$ rational approximation $+$ numerical linear algebra
- Rational Arnoldi process to compute orthonormal rational Krylov bases
- Arnoldi approximation of $f(A)\boldsymbol{b}$, and some theory (interpolation, near-optimality)
- Approaches to pole optimization: analytical, scalar numerical, greedy/iterated Krylov
- RKFIT for rational least squares approximation $r_m(A)\boldsymbol{b} \approx F\boldsymbol{b}$
- Spectral adaptivity and application to PMLs for variable coefficient media
- AAA algorithm and application to solving nonlinear eigenvalue problems

Bibliography

- Beckermann, B., & Reichel, L. (2009). Error estimates and evaluation of matrix functions via the Faber transform. SIAM Journal on Numerical Analysis, 47(5), 3849–3883.
- Beckermann, B., & Townsend, A. (2017). On the singular values of matrices with displacement structure. SIAM Journal on Matrix Analysis and Applications, 38(4), 1227–1248.
- Berljafa, M., & Güttel, S. (2015). Generalized rational Krylov decompositions with an application to rational approximation. SIAM Journal on Matrix Analysis and Applications, 36(2), 894-916.
- Berljafa, M., & Güttel, S. (2017). The RKFIT algorithm for nonlinear rational approximation. SIAM Journal on Scientific Computing, 39(5), A2049–A2071.
- Camps, D., Meerbergen, K., & Vandebril, R. (2019). An implicit filter for rational Krylov using core transformations. Linear Algebra and its Applications, 561, 113–140.

  Druskin, V., Güttel, S., & Knizhnerman, L. (2022). Model order reduction of layered waveguides via rational Krylov fitting. BIT Numerical Mathematics, 62(4), 1551–1572.
- Druskin, V., & Knizhnerman, L. (1998). Extended Krylov subspaces: approximation of the matrix square root and related functions. SIAM Journal on Matrix Analysis and Applications, 19(3), 755–771.
- Druskin, V., & Simoncini, V. (2011). Adaptive rational Krylov subspaces for large-scale dynamical systems. Systems & Control Letters, 60(8), 546–560.
- Elsworth, S., & Güttel, S. (2020). The block rational Arnoldi method. SIAM Journal on Matrix Analysis and Applications, 41(2), 365-388.

Bibliography

- Gugercin, S., Antoulas, A. C., & Beattie, C. (2008). $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. SIAM Journal on Matrix Analysis and Applications, 30(2), 609–638.
- Gustavsen, B., & Semlyen, A. (1999). Rational approximation of frequency domain responses by vector fitting. IEEE Transactions on Power Delivery, 14(3), 1052–1061.
- Güttel, S. (2013). Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection. GAMM-Mitteilungen, 36(1), 8–31.
- Güttel, S., Kressner, D., Vandereycken, B. (2024). Randomized sketching of nonlinear eigenvalue problems. SIAM Journal on Scientific Computing, 46(5), A3022–A3043.
- Lietaert, P., Meerbergen, K., Pérez, J., & Vandereycken, B. (2022). Automatic rational approximation and linearization of nonlinear eigenvalue problems. IMA Journal of Numerical Analysis, 42(2), 1087–1115.
- Nakatsukasa, Y., Sète, O., & Trefethen, L. N. (2018). The AAA algorithm for rational approximation. SIAM Journal on Scientific Computing, 40(3), A1494–A1522.
- Ruhe, A. (1984). Rational Krylov sequence methods for eigenvalue computation. Linear Algebra and its Applications, 58, 391–405.
- Ruhe, A. (1994). Rational Krylov algorithms for nonsymmetric eigenvalue problems. II. Matrix pairs. Linear algebra and its Applications, 197, 283–295.
- Saad, Y. (1992). Analysis of some Krylov subspace approximations to the matrix exponential operator. SIAM Journal on Numerical Analysis, 29(1), 209–228.