Rational Krylov methods for the approximation of matrix functions

Stefan Güttel[1]

[1]Institut für Numerische Mathematik und Optimierung
Technische Universität Bergakademie Freiberg

Numerik-Seminar
04. Dezember 2008

Outline

Motivation

We consider the initial-boundary value problem

$$
\begin{aligned}
\partial_t u - \Delta u &= 0 && \text{in } \Omega = (0,1)^3, \quad t > 0,\\
u(x,t) &= 0 && \text{on } \Gamma = \partial\Omega, \quad t > 0,\\
u(x,0) &= u_0(x) && \text{in } \Omega.
\end{aligned}
$$

Seven-point stencil on a uniform grid involving n_1 interior grid points in each Cartesian direction yields IVP

$$
\begin{aligned}
\boldsymbol{u}'(t) &= A\boldsymbol{u}(t), \quad t > 0,\\
\boldsymbol{u}(0) &= \boldsymbol{u}_0,
\end{aligned}
$$

with $n \times n$ matrix A ($n = n_1^3$) and an initial vector $\boldsymbol{u}_0$ consisting of the values $u_0(x)$ at the grid points x, the solution of which is given by

$$\boldsymbol{u}(t) = \exp(tA)\boldsymbol{u}_0.$$

Define Krylov space of order m

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

- There exists an invariance index L such that $\mathscr{K}_{L-1} \subset \mathscr{K}_L = \mathscr{K}_{L+1}$.
- $\mathscr{K}_m \cong \mathscr{P}_{m-1}$ for $m \leq L$.
- The minimal polynomial $\Psi_{A,\boldsymbol{b}}$ of $\boldsymbol{b}$ with respect to A has degree L.
- The exact solution $\exp(tA)\boldsymbol{b}$ is contained in $\mathscr{K}_L$.

Approximation methods can be characterized by an m-dimensional approximation space $\mathscr{V}_m$ and the way how approximations are extracted from $\mathscr{V}_m$.

Method 1 (SLDM)

Space: Approximate $\exp(tA)\boldsymbol{b}$ in $\mathscr{V}_m = \mathscr{K}_m$.

Extraction: Rayleigh-Ritz extraction

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

where V_m is ONB-basis of $\mathscr{K}_m$.

There holds

$$\boldsymbol{f}_m(t) = W_m f(t\Theta_m)\, W_m^* \boldsymbol{b},$$

where the columns of $W_m = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]$ are orthonormal Ritz vectors and $\Theta_m = \operatorname{diag}(\theta_1, \ldots, \theta_m)$ associated Ritz values from $\mathscr{K}_m$.

Let $A = U\Lambda U^*$ be a spectral decomposition for A, $U = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n]$ and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Then

$$\boldsymbol{f}(t) = f(tA)\boldsymbol{b} = U f(t\Lambda)\, U^* \boldsymbol{b}.$$

Error in direction $\boldsymbol{u}_j$:

$$|\boldsymbol{u}_j^*(\boldsymbol{f}(t) - \boldsymbol{f}_m(t))| = |f(t\lambda_j)\,\boldsymbol{u}_j^*\boldsymbol{b} - \boldsymbol{u}_j^*\,W_m f(t\Theta_m)\,W_m^*\boldsymbol{b}|.$$

How close are $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ to some of A's eigenvectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$?
(Enumerate A's eigenvalues/vectors s.t. λ_j is closest to θ_j.)

Approximation quality of eigenvalues/vectors determined by

- spread/gap ratio (van den Eshof/Hochbruck), e.g.,

$$\frac{\lambda_n - \lambda_1}{\lambda_2 - \lambda_1}$$

- constrained equilibrium problem on the spectral interval of A.

Method 2 (shift-invert)

Space: Approximate $\exp(tA)\boldsymbol{b}$ in $\mathscr{V}_m = \mathscr{K}_m((\gamma I - A)^{-1}, \boldsymbol{b})$.

Extraction: Back-transform T_m

$$\boldsymbol{f}_m(t) = V_m f(t\,(\gamma I - T_m^{-1}))\,V_m^* \boldsymbol{b},$$

where V_m is an orthonormal basis of $\mathscr{V}_m$ satisfying

$$(I - \gamma A)^{-1}\,V_m = V_m\,T_m + \boldsymbol{v}_{m+1}\,t_{m+1,m}\,\boldsymbol{e}_m^T.$$

Note: The Rayleigh quotient of A is

$$V_m^* A V_m = (\gamma I - T_m^{-1}) - V_m^* A \boldsymbol{v}_{m+1}\,t_{m+1,m}\,\boldsymbol{e}_m^T\,T_m^{-1}.$$

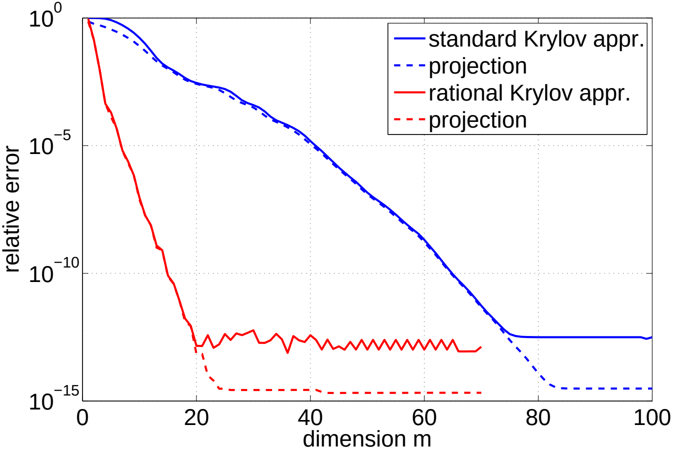


Figure: Heat equation $n = 15^3 = 3,375,\ t = 0.1,\ \gamma = 1.$

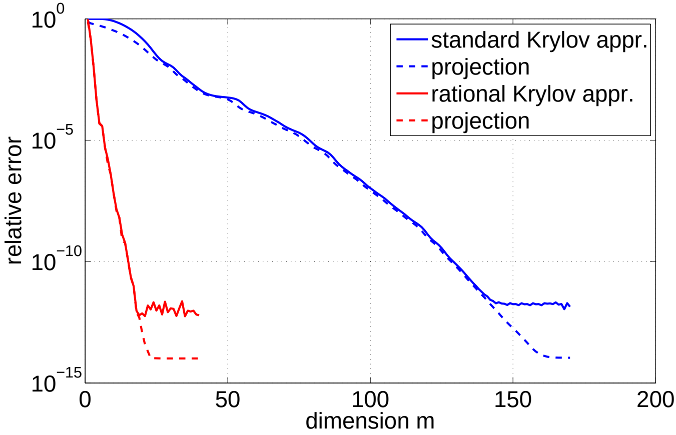


Figure: Heat equation $n = 31^3 = 29,791$, $t = 0.1$, $\gamma = 1$.

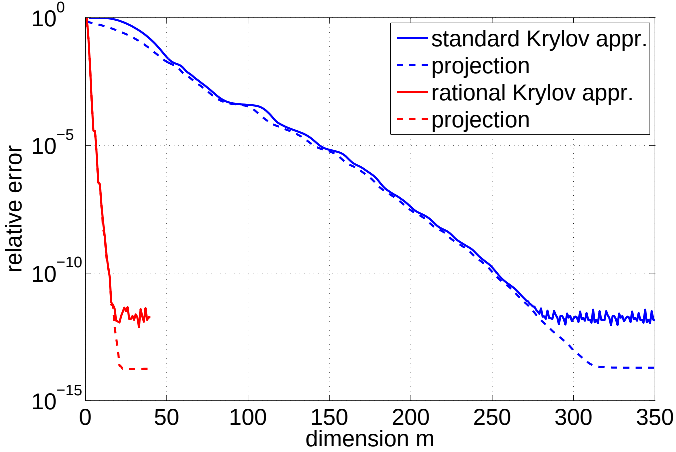


Figure: Heat equation $n = 63^3 = 250,047$, $t = 0.1$, $\gamma = 1$.

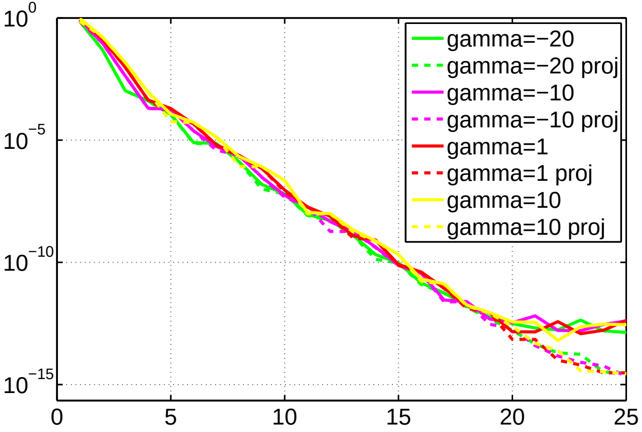


Figure: Dependence on γ, $n = 15^3 = 3,375$, $t = 0.1$.

Shift-invert Lanczos can be interpreted as SLDM for the transformed function $\tilde{f}(z) = f(t(\gamma - 1/z))$ and the spectrally transformed matrix $\tilde{A} = (\gamma I - A)^{-1}$.

This transforms eigenvalues, but not eigenvectors.

The convergence of Ritz values depends on the eigenvalue distribution of A and $\tilde{A}$, respectively.

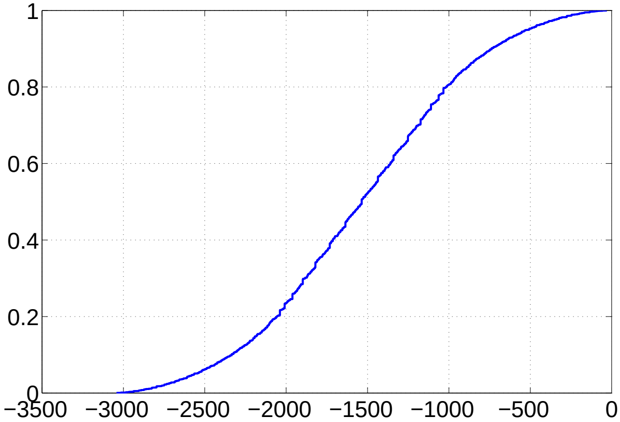


Figure: Eigenvalue distribution of A for $n = 15^3 = 3,375$.

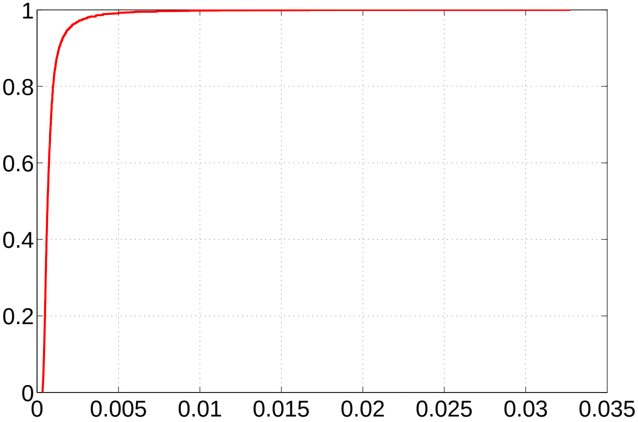


Figure: Eigenvalue distribution of $(\gamma I - A)^{-1}$ for $n = 15^3 = 3,375$, $\gamma = 1$.

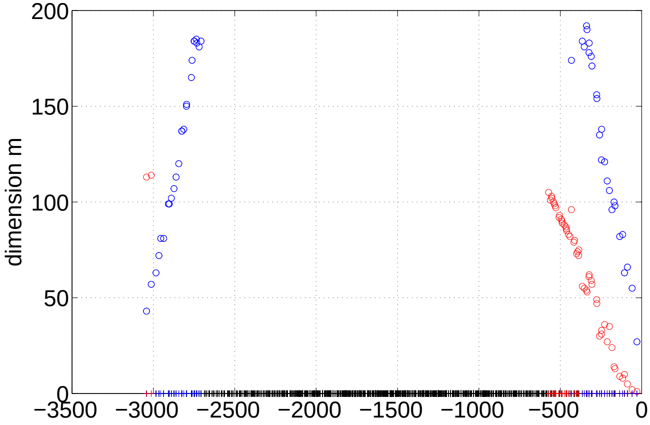


Figure: Ritz values approximating eigenvalues for SLDM and shift-invert.

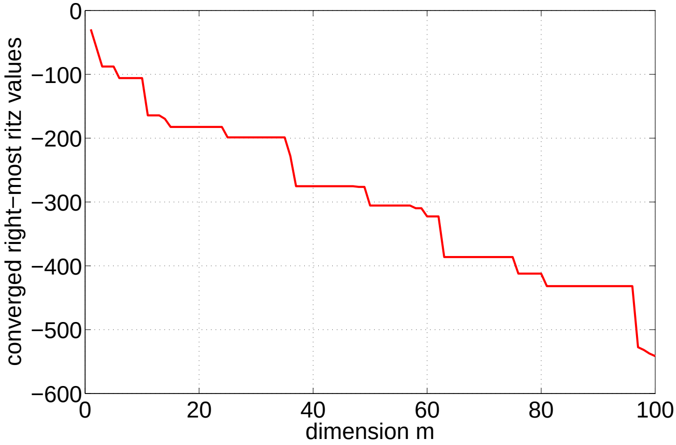


Figure: Ritz values approximating eigenvalues for shift-invert.

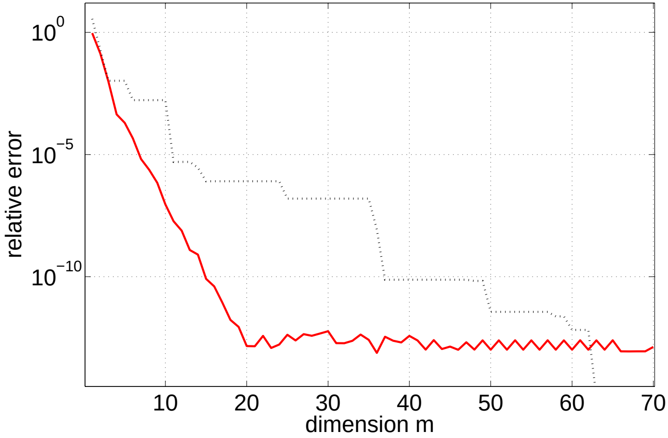


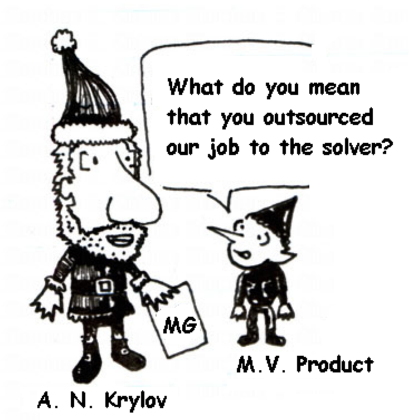
Figure: Simple error bound on the basis of converged right-most Ritz values.

Idea: In standard Krylov methods the matrix function $g(A) = A$ is applied in each step to build up a Krylov space.

In rational Krylov methods the matrix function $g(A) = (\gamma I - A)^{-1}$ is applied in each step to build up a Krylov space.

Use many applications of $g(A)\boldsymbol{x}$ on vectors to approximate the action of a more complicated function $f(A)$ on $\boldsymbol{b}$.

Maybe we could benefit from progress made for the solution of linear systems by displacing iteration work to the linear system solver.


What do you mean that you outsourced our job to the solver?
MG
M.V. Product
A. N. Krylov

Rational Krylov spaces

Definition

Let q_{m-1} be a polynomial of degree $m-1$ which is nonzero at all eigenvalues of A. Then $\mathscr{V}_m = [q_{m-1}(A)]^{-1}\mathscr{K}_m(A, \boldsymbol{b})$ is the *rational Krylov space of order* m *associated with* A, $\boldsymbol{b}$, *and* q_{m-1}.

Basic facts:

- $[q_{m-1}(A)]^{-1}\mathscr{K}_m(A, \boldsymbol{b}) = \mathscr{K}_m(A, [q_{m-1}(A)]^{-1}\boldsymbol{b}) \cong \mathscr{P}_{m-1}/q_{m-1}$.
- $\dim \mathscr{V}_m = \min\{m, L\}$.
- $\mathscr{V}_m \subseteq \mathscr{V}_L = \mathscr{K}_L$.
- $f(A)\boldsymbol{b} \in \mathscr{V}_L$.

For two polynomials p and q define

$$p(z) \odot q(z) := \frac{1}{2\pi i} \int_\Gamma \frac{\Psi_{A,\boldsymbol{b}}(\zeta) - \Psi_{A,\boldsymbol{b}}(z)}{\zeta - z} \frac{p(\zeta)q(\zeta)}{\Psi_{A,\boldsymbol{b}}(\zeta)} \, d\zeta.$$

$p \odot q \in \mathscr{P}_{L-1}$ is the Hermite-interpolating polynomial for pq at the nodes $\Psi_{A,\boldsymbol{b}}$.

Structural result

There holds

$$[q_{m-1}(A)]^{-1} \mathscr{K}_m(A, \boldsymbol{b}) \cong \hat{q} \odot \mathscr{P}_{m-1},$$

where $\hat{q}$ is a polynomial such that $[q_{m-1}(A)]^{-1} = \hat{q}(A)$.

Representation result

For every m-dimensional subspace $\mathscr{V}_m \subseteq \mathscr{K}_L(A, \boldsymbol{b})$ with a Lagrange basis there exists a polynomial $\hat{q} \in \mathscr{P}_{L-m}$ such that

$$\hat{q} \odot \mathscr{P}_{m-1} \cong \mathscr{V}_m.$$

Question: Is every m-dimensional subspace $\mathscr{V}_m \subseteq \mathscr{K}_L(A, \boldsymbol{b})$ a rational Krylov space $[q_{m-1}(A)]^{-1}\mathscr{K}_m(A, \boldsymbol{b})$?

Necessary: $\boldsymbol{b} \in \mathscr{V}_m$.

Answer: No. For $\Psi_{A,\boldsymbol{b}}(z) = z^3$ and $\mathscr{V}_m = \mathrm{span}\{\boldsymbol{b}, A^2\boldsymbol{b}\}$ there exists no $\hat{q}$ such that $\hat{q} \odot \mathscr{P}_{m-1} \cong \mathscr{V}_m$.

Rational Krylov decompositions

For an ordered set of poles $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{m-1}) \subset \overline{\mathbb{C}}$ and scale factors $\boldsymbol{\mu} = (\mu_1, \dots, \mu_{m-1}) \subset \mathbb{C}$, define *continuation matrices*

$$C_j := (\mu_j \sigma_j A - I)(\sigma_j I - A)^{-1} \qquad (j = 1, \dots, m-1)$$

and the rational Krylov space

$$\mathscr{K}_{\boldsymbol{\mu},\boldsymbol{\sigma}}(A, \boldsymbol{b}) := \mathsf{span}\{\boldsymbol{b}, C_1 \boldsymbol{b}, \dots, C_{m-1} \cdots C_1 \boldsymbol{b}\}.$$

We will always assume that none of the poles $\sigma_j \in \overline{\mathbb{C}}$ coincides with an eigenvalue of A and that $\sigma_j \neq \mu_j^{-2}$ for $j = 1, \dots, m-1$ (in particular, we require $\sigma_j \neq 0$ if $\mu_j = \infty$).

An elementary calculation shows that every $\boldsymbol{v} \in \mathscr{K}_{\boldsymbol{\mu},\boldsymbol{\sigma}}$ is of the form

$$\boldsymbol{v} = \frac{\sum_{j=0}^{m-1} \alpha_j (\mu_1\sigma_1 A - I)\cdots(\mu_j\sigma_j A - I)(\sigma_{j+1} I - A)\cdots(\sigma_{m-1} I - A)}{(\sigma_1 I - A)\cdots(\sigma_{m-1} I - A)}\boldsymbol{b},$$

and therefore

$$\boldsymbol{v} \in \frac{1}{(\sigma_1 I - A)\cdots(\sigma_{m-1} I - A)}\mathscr{K}_m(A, \boldsymbol{b}).$$

Hence,

$$\mathscr{K}_{\boldsymbol{\mu},\boldsymbol{\sigma}}(A, \boldsymbol{b}) \subseteq [q_{m-1}(A)]^{-1}\mathscr{K}_m(A, \boldsymbol{b})$$

with $q_{m-1}(z) = (\sigma_1 - z)\cdots(\sigma_{m-1} - z)$.

Equality?

We have $\mathscr{K}_{\boldsymbol{\mu},\boldsymbol{\sigma}} = [q_{m-1}(A)]^{-1}\mathscr{K}_m$ iff the polynomials $p_0, \ldots, p_{m-1}$ defined by their zeros

$$\begin{array}{rcl} p_0 & : & \sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_{m-1} \\ p_1 & : & 1/\mu_1\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_{m-1} \\ p_2 & : & 1/\mu_1\sigma_1, 1/\mu_2\sigma_2, \sigma_3, \ldots, \sigma_{m-1} \\ & \vdots & \\ p_{m-2} & : & 1/\mu_1\sigma_1, 1/\mu_2\sigma_2, 1/\mu_3\sigma_3, \ldots, \sigma_{m-1} \\ p_{m-1} & : & 1/\mu_1\sigma_1, 1/\mu_2\sigma_2, 1/\mu_3\sigma_3, \ldots, 1/\mu_{m-1}\sigma_{m-1}, \end{array}$$

satisfy

$$p_0(1/\mu_1\sigma_1) \neq 0, p_1(1/\mu_2\sigma_2) \neq 0, \ldots, p_{m-2}(1/\mu_{m-1}\sigma_{m-1}) \neq 0,$$

$$\text{and } p_{m-1}(\sigma_{m-1}) \neq 0.$$

The following algorithm constructs an orthonormal basis of $\mathscr{K}_{\boldsymbol{\mu},\boldsymbol{\sigma}}$:

Algorithm 1: RKS algorithm.

Given: $A, \boldsymbol{b}, \boldsymbol{\mu}, \boldsymbol{\sigma}$

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

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

  $\boldsymbol{w} := (\mu_j \sigma_j A - I)(\sigma_j I - A)^{-1} \boldsymbol{v}_j$

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

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

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

  end

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

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

end

Resulting decomposition

$$AV_m \left(H_m D_{m-1} + \begin{bmatrix} M_{m-1}\tilde{S}_{m-1} \\ \mathbf{0} \end{bmatrix} \right) = V_m \left(H_m \tilde{S}_{m-1} + \begin{bmatrix} D_{m-1} \\ \mathbf{0} \end{bmatrix} \right),$$

where

$$\begin{aligned} V_m &:= [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m], \\ H_m &:= [h_{i,j}] \ m \times m-1 \text{ upper Hessenberg matrix}, \\ M_{m-1} &:= \mathrm{diag}(\mu_1, \ldots, \mu_{m-1}). \end{aligned}$$

Let further $S_{m-1} := \mathrm{diag}(\sigma_1, \ldots, \sigma_{m-1})$.
Then $S_{m-1} =: \tilde{S}_{m-1} D_{m-1}^{-1}$ such that all entries in $\tilde{S}_{m-1}$ are finite.

Special cases of Rational Krylov methods:

1. Standard Arnoldi: $\boldsymbol{\mu} = (1, 1, \dots, 1)$, $\boldsymbol{\sigma} = (\infty, \infty, \dots, \infty)$.
2. Extended Krylov: $\boldsymbol{\mu} = (1, 0, 1, 0, \dots)$, $\boldsymbol{\sigma} = (\infty, 0, \infty, 0, \dots)$.
3. Shift-invert: $\boldsymbol{\mu} = (0, 0, \dots, 0)$, $\boldsymbol{\sigma} = (\gamma, \gamma, \dots, \gamma)$.

In cases 1. and 2. the Rayleigh coefficient $V_m^* A V_m$ can be computed without explicit projection. In the general case it is useful to set $\sigma_m = \infty$ and to perform one more step of the RKS algorithm [Beckermann/Reichel]. Then

$$V_m^* A V_m = \left([I_m, \mathbf{0}] H_{m+1} \tilde{S}_m + D_m\right) \left([H_m D_{m-1}, \mathbf{0}] + M_m \tilde{S}_m\right)^{-1}.$$

We can recover a basis of the standard Krylov space $\mathscr{K}_m$ from V_m as $W_m := q_{m-1}(A)\,V_m$. One can show

Lemma

With $\tilde{W}_m := W_m\left([H_m D_{m-1}, \mathbf{0}] + M_m \tilde{S}_m\right)$ *we have*

$$\tilde{W}_m^\dagger A\,\tilde{W}_m = \left([H_m D_{m-1}, \mathbf{0}] + M_m \tilde{S}_m\right)^{-1}\left([I_m, \mathbf{0}] H_{m+1}\tilde{S}_m + D_m\right) + \boldsymbol{c}\boldsymbol{e}_m^T,$$

where $\boldsymbol{c} := \tilde{\sigma}_m h_{m+1,m}(\tilde{W}_m^\dagger \boldsymbol{w}_{m+1})$.

As $\tilde{\sigma}_m \to 0$ we obtain a Rayleigh coefficient for a standard Krylov space and with it the standard Ritz values.

Rayleigh-Ritz approximation

Lemma (Exactness)

Let $[q(A)]^{-1}\mathscr{K}_m(A,\boldsymbol{b})$ *be a rational Krylov space of dimension* m *and let* V_m *be a basis of it. Then for* $K_m := V_m^\dagger A V_m$ *and every rational function* r *of the form* $r = p/q$ *(*$p \in \mathscr{P}_{m-1}$*) there holds* $r(A)\boldsymbol{b} = V_m r(K_m) V_m^\dagger \boldsymbol{b}$*.*

Proof: Since $[q(A)]^{-1}\mathscr{K}_m(A,\boldsymbol{b}) = \mathscr{K}_m(A,[q(A)]^{-1}\boldsymbol{b})$, $p(A)[q(A)]^{-1}\boldsymbol{b} = V_m p(K_m) V_m^\dagger [q(A)]^{-1}\boldsymbol{b}$ for all $p \in \mathscr{P}_{m-1}$ (standard Krylov). Particulary, q has degree $\leq m-1$ and hence $V_m^\dagger \boldsymbol{b} = V_m^\dagger q(A)[q(A)]^{-1}\boldsymbol{b} = V_m^\dagger V_m q(K_m) V_m^\dagger [q(A)]^{-1}\boldsymbol{b} = q(K_m) V_m^\dagger [q(A)]^{-1}\boldsymbol{b}$ proving $[q(K_m)]^{-1} V_m^\dagger \boldsymbol{b} = V_m^\dagger [q(A)]^{-1}\boldsymbol{b}$. Now, $r(A)\boldsymbol{b} = p(A)[q(A)]^{-1}\boldsymbol{b} = V_m p(K_m) V_m^\dagger [q(A)]^{-1}\boldsymbol{b} = V_m p(K_m)[q(K_m)]^{-1} V_m^\dagger \boldsymbol{b}$.

Lemma (Interpolation)

Let $\boldsymbol{c} := [q(A)]^{-1}\boldsymbol{b}$ *and* V_m *be a basis of* $[q(A)]^{-1}\mathscr{K}_m(A,\boldsymbol{b}) = \mathscr{K}_m(A,\boldsymbol{c})$, $K_m := V_m^\dagger A V_m$. *Then*

$$V_m f(K_m)\, V_m^\dagger \boldsymbol{b} = V_m \tilde{f}(K_m)\, V_m^\dagger \boldsymbol{c} = \tilde{p}_{m-1}(A)\boldsymbol{c},$$

where $\tilde{f} := q \cdot f$ *and* $\tilde{p}_{m-1}$ *is a polynomial of degree* $m-1$ *that Hermite-interpolates* $\tilde{f}$ *at* $\Lambda(K_m)$.

Proof: To prove the first equality, note that $\deg q \leq m-1$ and $q(A)\boldsymbol{c} = \boldsymbol{b}$. By Lemma (Exactness) there holds $\boldsymbol{b} = V_m q(K_m)\, V_m^\dagger \boldsymbol{c}$. The second equality results from interpolation properties of standard Krylov methods.

Let w_m denote the characteristic polynomial of K_m and let Γ be a contour containing $\Lambda(K_m)$ in its interior $\mathrm{int}(\Gamma)$. Furthermore f is assumed to be analytic in $\mathrm{int}(\Gamma)$ (and so is $\tilde{f}$). Then the polynomial $\tilde{p}_{m-1}$ can be expressed using Hermite's formula

$$\tilde{p}_{m-1}(z) = \frac{1}{2\pi i} \int_\Gamma \frac{w_m(\zeta) - w_m(z)}{w_m(\zeta)(\zeta - z)} \tilde{f}(\zeta)\, d\zeta,$$

and for the interpolation error

$$\tilde{f}(z) - \tilde{p}_{m-1}(z) = \frac{1}{2\pi i} \int_\Gamma \frac{w_m(z)}{w_m(\zeta)(\zeta - z)} \tilde{f}(\zeta)\, d\zeta.$$

Dividing this equation by $q(z)$ and setting $s_m(z) := q(z)/w_m(z)$ we obtain

$$f(z) - \tilde{p}_{m-1}(z)/q(z) = \frac{1}{2\pi i} \int_\Gamma \frac{s_m(\zeta)}{s_m(z)(\zeta - z)} f(\zeta)\, d\zeta.$$

Using Lemma (Interpolation) we have the following representation for the error of a Rayleigh-Ritz approximation

$$f(A)\boldsymbol{b} - V_m f(K_m)\, V_m^\dagger \boldsymbol{b} = \left(\frac{1}{2\pi i} \int_\Gamma (\zeta I - A)^{-1} s_m(\zeta) f(\zeta)\, d\zeta \right) [s_m(A)]^{-1} \boldsymbol{b}$$

Aim: Choose $q(z)$ such that $s_m(z) = q(z)/w_m(z)$ is as large as possible on $\Lambda(A)$ and as small as possible on Γ.

Lemma (Optimality)

Let A be normal and $[q(A)]^{-1}\mathscr{K}_m(A,\boldsymbol{b})$ be a rational Krylov space of dimension m and let V_m be a basis of it. Let $K_m := V_m^\dagger A V_m$ and $\Omega := \Lambda(A) \cup \Lambda(K_m)$. Then

$$\|f(A)\boldsymbol{b} - V_m f(K_m)\, V_m^\dagger \boldsymbol{b}\| \leq 2\|\boldsymbol{b}\| \min_{p \in \mathscr{P}_{m-1}} \max_{\lambda \in \Omega} |f(\lambda) - p(\lambda)/q(\lambda)|.$$

Aim: Choose poles $q(z)$ such that there exists a rational function $p(z)/q(z)$ which is close to f on Ω.

Orthogonal rational functions

Consider a sequence of orthonormal basis vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ of a rational Krylov space $[q(A)]^{-1}\mathscr{K}_m(A, \boldsymbol{b})$. There exist polynomials $p_1, \ldots, p_m$ of (not necessarily ascending) degree $m-1$ such that $\boldsymbol{v}_j = p_j(A)[q(A)]^{-1}\boldsymbol{b} = r_j(A)\boldsymbol{b}$ ($r_j := p_j/q$, $j = 1, \ldots, m$). Let $A = U\Lambda U^*$ be normal, $U^*U = I$. Then

$$\delta_{j,k} = \langle r_j(A)\boldsymbol{b}, r_k(A)\boldsymbol{b}\rangle_2 = \langle r_j(\Lambda)\,U^*\boldsymbol{b}, r_k(\Lambda)\,U^*\boldsymbol{b}\rangle_2 = \langle r_j(z), r_k(z)\rangle_\mu$$

where

$$\langle r_j(z), r_k(z)\rangle_\mu := \sum_{i=1}^{n} r_j(\lambda_i)\overline{r_k(\lambda_i)}\,|\boldsymbol{e}_i^T U^*\boldsymbol{b}|^2.$$

Hence, the r_j are orthonormal rational functions with respect to the inner product $\langle\cdot,\cdot\rangle_\mu$.

On the other hand, the following interpretation is also possible

$$\delta_{j,k} = \langle r_j(A)\boldsymbol{b}, r_k(A)\boldsymbol{b}\rangle_2 = \left\langle (p_j/q)(\Lambda)\, U^*\boldsymbol{b}, (p_k/q)(\Lambda)\, U^*\boldsymbol{b}\right\rangle_2 = \langle p_j, p_k\rangle_{\mu_q},$$

where

$$\langle p_j(z), p_k(z)\rangle_{\mu_q} := \sum_{i=1}^{n} p_j(\lambda_i)\overline{p_k(\lambda_i)}\, |q(\lambda_i)|^{-2}\, |\boldsymbol{e}_i^T U^*\boldsymbol{b}|^2.$$

Outlook

- Useful results from the theory of orthogonal rational functions available?
- Study convergence of Ritz values of spectrally transformed matrices.
- Error bounds/indicators.
- Implementation issues, shifted MG methods and direct solvers.