Computation of generalized matrix functions with rational Krylov methods

Angelo A. Casulli[†] Igor Simunec[†]

[†]Scuola Normale Superiore di Pisa

Due Giorni di Algebra Lineare Numerica e Applicazioni

Napoli, 14-15 febbraio 2022

Outline

1. Background
    - Generalized matrix functions
    - Rational Krylov subspaces
2. Computation of GMFs
    - Golub-Kahan bidiagonalization
    - Rational Krylov methods
    - Short term recurrence
    - Error bounds
3. Numerical results

Outline

1. Background

- Generalized matrix functions
- Rational Krylov subspaces

2. Computation of GMFs

- Golub-Kahan bidiagonalization
- Rational Krylov methods
- Short term recurrence
- Error bounds

3. Numerical results

Standard matrix functions

Let A be a diagonalizable $n \times n$ matrix, with spectral decomposition

$$A = V \Lambda V^{-1}, \qquad \text{where } \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

If $f : \mathbb{C} \to \mathbb{C}$ is a function defined on the spectrum of A, the **matrix function** $f(A)$ is defined as

$$f(A) = V f(\Lambda) V^{-1}, \qquad \text{where } f(\Lambda) = \operatorname{diag}(f(\lambda_1), \dots, f(\lambda_n)).$$

For a general matrix A, a matrix function can be defined via the Jordan canonical form.

Generalized matrix functions

Generalized matrix functions (GMFs) are an extension of standard matrix functions to the rectangular case, defined using the SVD instead of a spectral decomposition [Hawkins–Ben-Israel, 1973].

Given an $m \times n$ matrix A with SVD $A = U\Sigma V^T$ and a function $f : \mathbb{R}^+ \to \mathbb{R}$, a **generalized matrix function** of A, denoted by $f^\diamond(A)$, is defined as

$$f^\diamond(A) = Uf^\diamond(\Sigma)V^T \in \mathbb{R}^{m\times n},$$

where $f^\diamond(\Sigma)$ is diagonal with entries

$$f^\diamond(\Sigma)_{ii} = \begin{cases} f(\sigma_i) & \text{if } \sigma_i > 0, \\ 0 & \text{otherwise.} \end{cases}$$

This definition does not depend on the value of f at $z = 0$, so we can always assume that $f(0) = 0$ and f is **odd**.

Examples

- For $f(z) = z$, $f^\diamond(A) = A$.
- For $f(z) = z^3$, $f^\diamond(A) = AA^TA$.
- For $f(z) = z^{-1}$, $f^\diamond(A)^T$ is the Moore-Penrose pseudoinverse A^+.
- If f is odd and we define $\mathcal{A} = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$, then

$$f(\mathcal{A}) = \begin{bmatrix} 0 & f^\diamond(A) \\ f^\diamond(A^T) & 0 \end{bmatrix}.$$

Rational Krylov subspaces

Expressions of the form $f(A)\boldsymbol{b}$ can be efficiently computed using polynomial and rational Krylov subspaces.

Given a sequence of poles $\{\xi_j\}_{j\geq 1} \subset (\mathbb{C} \cup \infty) \setminus \Lambda(A)$, the associated **rational Krylov subspace** is

$$\mathcal{Q}_k(A, \boldsymbol{b}) = q_{k-1}(A)^{-1}\mathcal{K}_k(A, \boldsymbol{b}),$$

where $q_{k-1}(z) = \prod_{j=1}^{k-1}(z - \xi_j)$ and $\mathcal{K}_k(A, \boldsymbol{b})$ is the **Krylov subspace**

$$\mathcal{K}_k(A, \boldsymbol{b}) = \operatorname{span}\{\boldsymbol{b}, A\boldsymbol{b}, \ldots, A^{k-1}\boldsymbol{b}\}.$$

An orthonormal basis V_k of $\mathcal{Q}_k(A, \boldsymbol{b})$ can be computed with the **rational Arnoldi algorithm**, which requires the solution of $k-1$ shifted linear systems with the matrix A.

Outline

1. Background
   - Generalized matrix functions
   - Rational Krylov subspaces

2. Computation of GMFs
   - Golub-Kahan bidiagonalization
   - Rational Krylov methods
   - Short term recurrence
   - Error bounds

3. Numerical results

Computation of $f^\diamond(A)\boldsymbol{b}$ via GK bidiagonalization

In [Arrigo-Benzi-Fenu, 2016] the following method is proposed for the efficient computation of $f^\diamond(A)\boldsymbol{b}$.

- Perform k steps of Golub-Kahan bidiagonalization on A. This produces a $k \times k$ bidiagonal matrix B_k and matrices P_k, Q_k with orthonormal columns, such that $B_k = P_k^T A Q_k$.
- Approximate $f^\diamond(A)\boldsymbol{b}$ with

$$\boldsymbol{y}_k = P_k f^\diamond(B_k) Q_k^T \boldsymbol{b} = P_k f^\diamond(B_k) \boldsymbol{e}_1 \|\boldsymbol{b}\|_2.$$

Computation of $f^\diamond(A)\boldsymbol{b}$ via GK bidiagonalization

In [Arrigo-Benzi-Fenu, 2016] the following method is proposed for the efficient computation of $f^\diamond(A)\boldsymbol{b}$.

- Perform k steps of Golub-Kahan bidiagonalization on A. This produces a $k \times k$ bidiagonal matrix B_k and matrices P_k, Q_k with orthonormal columns, such that $B_k = P_k^T A Q_k$.
- Approximate $f^\diamond(A)\boldsymbol{b}$ with

$$\boldsymbol{y}_k = P_k f^\diamond(B_k) Q_k^T \boldsymbol{b} = P_k f^\diamond(B_k) \boldsymbol{e}_1 \|\boldsymbol{b}\|_2.$$

Another efficient method based on Chebyshev interpolation is proposed in [Aurentz-Austin-Benzi-Kalantzis, 2019].

Both methods have very good performance for analytic functions such as $\sin(z)$, $\sinh(z)$.

Computation of $f^\diamond(A)\boldsymbol{b}$ via rational Krylov

The matrices P_k and Q_k constructed in the Golub-Kahan bidiagonalization process are orthonormal bases of the **polynomial Krylov subspaces**

$$\operatorname{span} Q_k = \mathcal{K}_k(A^T A, \boldsymbol{b}) \quad \text{and} \quad \operatorname{span} P_k = \mathcal{K}_k(AA^T, A\boldsymbol{b}).$$

In this talk we present a new class of methods, obtained by replacing $\mathcal{K}_k(A^T A, \boldsymbol{b})$ and $\mathcal{K}_k(AA^T, A\boldsymbol{b})$ with the corresponding **rational Krylov subspaces**.

Computation of $f^\diamond(A)\boldsymbol{b}$ via rational Krylov

The matrices P_k and Q_k constructed in the Golub-Kahan bidiagonalization process are orthonormal bases of the **polynomial Krylov subspaces**

$$\operatorname{span} Q_k = \mathcal{K}_k(A^T A, \boldsymbol{b}) \quad \text{and} \quad \operatorname{span} P_k = \mathcal{K}_k(AA^T, A\boldsymbol{b}).$$

In this talk we present a new class of methods, obtained by replacing $\mathcal{K}_k(A^T A, \boldsymbol{b})$ and $\mathcal{K}_k(AA^T, A\boldsymbol{b})$ with the corresponding **rational Krylov subspaces**.

Let Q_k and P_k be matrices with orthonormal columns such that

$$\operatorname{span} Q_k = \mathcal{Q}_k(A^T A, \boldsymbol{b}) \quad \text{and} \quad \operatorname{span} P_k = \mathcal{Q}_k(AA^T, A\boldsymbol{b}).$$

Defining $B_k = P_k^T A Q_k$, we can approximate $f^\diamond(A)\boldsymbol{b}$ with

$$\boldsymbol{y}_k = P_k f^\diamond(B_k) Q_k^T \boldsymbol{b} = P_k f^\diamond(B_k) \boldsymbol{e}_1 \|\boldsymbol{b}\|_2.$$

Computational remarks

- A basis Q_k of $\mathcal{Q}_k(A^T A, \boldsymbol{b})$ can be computed by solving $k-1$ shifted linear systems with $A^T A$.
- Since $\mathcal{Q}_k(AA^T, A\boldsymbol{b}) = A\mathcal{Q}_k(A^T A, \boldsymbol{b})$, we can obtain P_k and B_k with a thin QR decomposition of AQ_k.

Computational remarks

- A basis Q_k of $\mathcal{Q}_k(A^TA, \boldsymbol{b})$ can be computed by solving $k-1$ shifted linear systems with A^TA.
- Since $\mathcal{Q}_k(AA^T, A\boldsymbol{b}) = A\mathcal{Q}_k(A^TA, \boldsymbol{b})$, we can obtain P_k and B_k with a thin QR decomposition of AQ_k.
- In the Golub-Kahan bidiagonalization, B_k is **bidiagonal**, so the next columns of P_k, Q_k and B_k can be computed with a short term recurrence.

Computational remarks

- A basis Q_k of $\mathcal{Q}_k(A^T A, \boldsymbol{b})$ can be computed by solving $k-1$ shifted linear systems with $A^T A$.
- Since $\mathcal{Q}_k(AA^T, A\boldsymbol{b}) = A\mathcal{Q}_k(A^T A, \boldsymbol{b})$, we can obtain P_k and B_k with a thin QR decomposition of AQ_k.
- In the Golub-Kahan bidiagonalization, B_k is **bidiagonal**, so the next columns of P_k, Q_k and B_k can be computed with a short term recurrence.
- In the rational case, B_k is upper triangular but **not** bidiagonal.
- However, we can still compute the columns of P_k and B_k with a **short recurrence** by exploiting the **quasiseparable** structure of B_k, i.e. that all the blocks in its strictly upper triangular part have rank 1.

Short term recurrence

Let

$$B_k = \begin{bmatrix} d_1 & \beta_1 & \gamma_1 & \dots & * \\ & \ddots & \ddots & \ddots & \vdots \\ & & d_{k-2} & \beta_{k-2} & \gamma_{k-2} \\ & & & d_{k-1} & \beta_{k-1} \\ & & & & d_k \end{bmatrix}.$$

We have

$$A\boldsymbol{q}_k = AQ_k\boldsymbol{e}_k = P_kB_k\boldsymbol{e}_k = d_k\boldsymbol{p}_k + \boldsymbol{x}_k, \quad \text{where } \boldsymbol{x}_k = [\,P_{k-1}\ 0\,]B_k\boldsymbol{e}_k.$$

Short term recurrence

Let

$$B_k = \begin{bmatrix} d_1 & \beta_1 & \gamma_1 & \dots & * \\ & \ddots & \ddots & \ddots & \vdots \\ & & d_{k-2} & \beta_{k-2} & \gamma_{k-2} \\ & & & d_{k-1} & \beta_{k-1} \\ & & & & d_k \end{bmatrix}.$$

We have

$$A\boldsymbol{q}_k = AQ_k\boldsymbol{e}_k = P_kB_k\boldsymbol{e}_k = d_k\boldsymbol{p}_k + \boldsymbol{x}_k, \quad \text{where } \boldsymbol{x}_k = [\,P_{k-1}\;0\,]B_k\boldsymbol{e}_k.$$

Using the fact that the submatrices of B_k in the strictly upper triangular part have rank at most 1, we can compute $\boldsymbol{x}_k$ with the recursive relation

$$\boldsymbol{x}_k = \frac{\gamma_{k-2}}{\beta_{k-2}}\boldsymbol{x}_{k-1} + \beta_{k-1}\boldsymbol{p}_{k-1}.$$

Algorithm

Algorithm 1: Short recurrence rational Krylov approximation of $f^\diamond(A)\boldsymbol{b}$

Input: $A \in \mathbb{R}^{n\times n}, \boldsymbol{b} \in \mathbb{R}^n, f, \{\xi_1, \dots, \xi_{k-1}\}$

Output: $\boldsymbol{y}_k \in \mathcal{Q}_k(AA^T, A\boldsymbol{b})$ s.t. $\boldsymbol{y}_k \approx f^\diamond(A)\boldsymbol{b}$

$\boldsymbol{q}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2$

$\boldsymbol{w}_1 = (I - A^TA/\xi_1)^{-1}A^TA\boldsymbol{q}_1$ `// can use other choices`

Compute $\boldsymbol{q}_2$ by orthogonalizing $\boldsymbol{w}_1$ against $\boldsymbol{q}_1$

Compute the QR decomposition $[\boldsymbol{p}_1, \boldsymbol{p}_2]\begin{bmatrix} d_1 & \beta_1 \\ 0 & d_2 \end{bmatrix} = [\boldsymbol{q}_1, \boldsymbol{q}_2]$

Define $B_2 = \begin{bmatrix} d_1 & \beta_1 \\ 0 & d_2 \end{bmatrix}$ and $\boldsymbol{x}_2 = \beta_1\boldsymbol{p}_1$

for $j = 2, \dots, k-1$ **do**

- $\boldsymbol{w}_j = (I - A^TA/\xi_j)^{-1}A^TA\boldsymbol{q}_j$ `// can use other choices`
- Compute $\boldsymbol{q}_{j+1}$ by orthogonalizing $\boldsymbol{w}_j$ against $[\boldsymbol{q}_1, \dots, \boldsymbol{q}_j]$
- Compute $\boldsymbol{p}_{j+1}, d_{j+1}, \beta_j, \gamma_{j-1}$ with short recurrence
- $B_{j+1} = \begin{bmatrix} B_j & \boldsymbol{s}_{j+1} \\ 0 & d_{j+1} \end{bmatrix}$, where $s_{j+1} = \begin{bmatrix} \frac{\gamma_{j-1}}{\beta_{j-1}}(B_j)_{1:j-1,j} \\ \beta_j \end{bmatrix}$

$P_k = [\boldsymbol{p}_1, \dots, \boldsymbol{p}_k]$

Compute $f^\diamond(B_k)$, e.g. via an SVD of B_k

$\boldsymbol{y}_k = P_k f^\diamond(B_k)\boldsymbol{e}_1\|\boldsymbol{b}\|_2$

Error bound

Let $A \in \mathbb{R}^{m\times n}$, and let σ_1, σ_n be the first and n-th singular values of A, where we use the notation $\sigma_n := 0$ when $n > m$.

Theorem

Let $\boldsymbol{y}_k \in \mathcal{Q}_k(AA^T, A\boldsymbol{b})$ be the approximation to $f^\diamond(A)\boldsymbol{b}$ obtained after k steps of a rational Krylov method, with $q_{k-1}(z) = \prod_{j=1}^{k-1}(z - \xi_j)$ as denominator polynomial. Then

$$\|f^\diamond(A)\boldsymbol{b} - \boldsymbol{y}_k\|_2 \le 2\|\boldsymbol{b}\|_2 \min_{p\in\mathcal{P}_{k-1}} \|f(z) - q_{k-1}(z^2)^{-1}p(z^2)z\|_{\infty,[\sigma_n,\sigma_1]}.$$

The bound can be also reformulated as

$$\|f^\diamond(A)\boldsymbol{b} - \boldsymbol{y}_k\|_2 \le 2\sigma_1\|\boldsymbol{b}\|_2 \min_{p\in\mathcal{P}_{k-1}} \left\| \frac{f(\sqrt{z})}{\sqrt{z}} - \frac{p(z)}{q_{k-1}(z)} \right\|_{\infty,[\sigma_n^2,\sigma_1^2]}.$$

How to deal with $\sigma_n = 0$

When $A \in \mathbb{R}^{m\times n}$ with $n > m$, we have $\sigma_n = 0$ and the matrix B_k can have arbitrarily small singular values even if $\sigma_m > 0$.

Example. Consider $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} \epsilon & 1 \end{bmatrix}^T$, for small $\epsilon > 0$. Then we have $Q_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2 = \frac{1}{\sqrt{1+\epsilon^2}}\boldsymbol{b}$ and $P_1 = A\boldsymbol{b}/\|A\boldsymbol{b}\|_2 = 1$. So we get $B_1 = P_1^T A Q_1 = \frac{\epsilon}{\sqrt{1+\epsilon^2}}$, which can be arbitrarily close to zero.

How to deal with $\sigma_n = 0$

When $A \in \mathbb{R}^{m\times n}$ with $n > m$, we have $\sigma_n = 0$ and the matrix B_k can have arbitrarily small singular values even if $\sigma_m > 0$.

Example. Consider $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} \epsilon & 1 \end{bmatrix}^T$, for small $\epsilon > 0$. Then we have $Q_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2 = \frac{1}{\sqrt{1+\epsilon^2}}\boldsymbol{b}$ and $P_1 = A\boldsymbol{b}/\|A\boldsymbol{b}\|_2 = 1$. So we get $B_1 = P_1^T A Q_1 = \frac{\epsilon}{\sqrt{1+\epsilon^2}}$, which can be arbitrarily close to zero.

We can overcome this difficulty with the identity

$$f^\diamond(A)\boldsymbol{b} = (A^+)^T f^\diamond(A^T)A\boldsymbol{b},$$

by first computing $\boldsymbol{w} = f^\diamond(A^T)A\boldsymbol{b}$ with a rational Krylov method on A^T and then recovering $f^\diamond(A)\boldsymbol{b}$ as the solution of the least squares problem

$$f^\diamond(A)\boldsymbol{b} = (A^+)^T\boldsymbol{w} = \arg\min_{\boldsymbol{y}} \|A^T\boldsymbol{y} - \boldsymbol{w}\|_2.$$

The projected matrix now has singular values in the interval $[\sigma_m, \sigma_1]$.

Outline

Numerical results

We performed experiments on random matrices with prescribed singular values to investigate the sharpness of the error bounds.

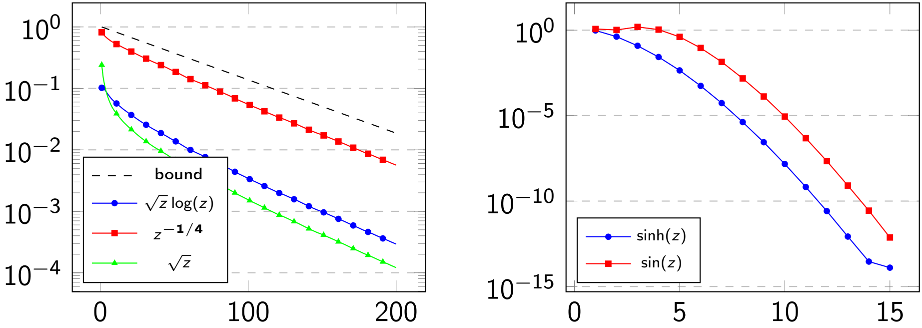
We compare the following methods:

- **polynomial** Krylov method;
- **extended** Krylov method, with alternating poles at 0 and ∞;
- **Shift-and-Invert** Krylov method, with repeated pole $\xi = -\sigma_{\min}\sigma_{\max}$;
- rational Krylov method with **asymptotically optimal poles** for Laplace-Stieltjes functions, from [Massei-Robol, 2020].

The poles for the fourth method were chosen according to the bound

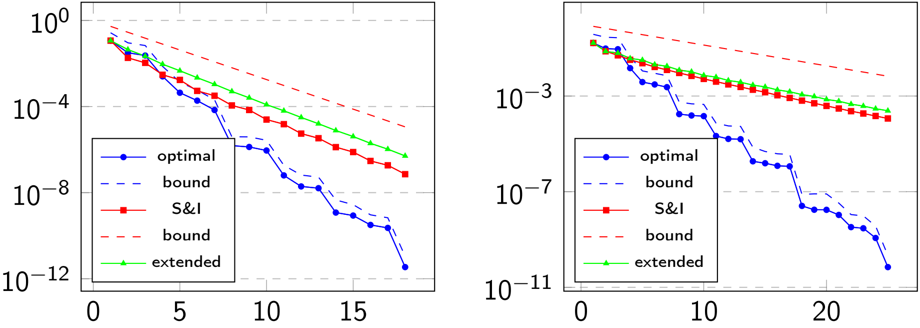
$$\|f^\diamond(A)\boldsymbol{b} - \boldsymbol{y}_k\|_2 \le 2\sigma_1\|\boldsymbol{b}\|_2 \min_{p\in\mathcal{P}_{k-1}} \left\| \frac{f(\sqrt{z})}{\sqrt{z}} - \frac{p(z)}{q_{k-1}(z)} \right\|_{\infty,[\sigma_n^2,\sigma_1^2]}.$$

Convergence – Polynomial



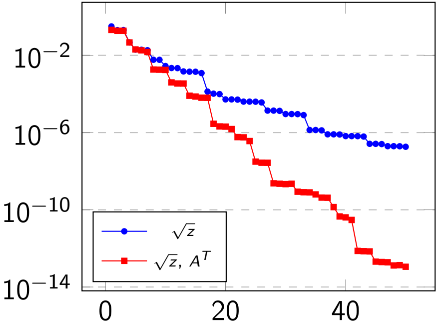
Convergence of the polynomial Krylov method, for a 2000×2000 matrix whose singular values are the Chebyshev points of the second kind for the interval $[10^{-1}, 10]$.

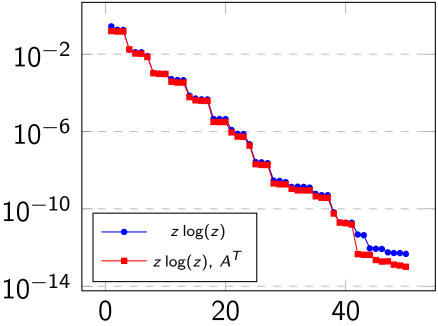
Convergence – Rational



Convergence of rational Krylov methods for $f^{\diamond}(A)\boldsymbol{b}$, where A is a 2000×2000 matrix with logspaced singular values in the interval $[1, 10]$ (left) or $[10^{-1}, 10]$ (right), and $f(z) = \sqrt{z}\log(1 + \sqrt{z})$.

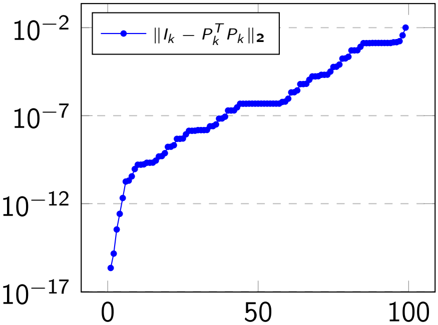
Convergence – Rectangular

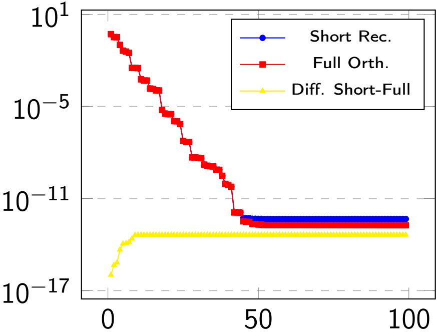




Convergence of the asymptotically optimal rational Krylov method for $f^{\diamond}(A)\boldsymbol{b}$, where A is a rectangular 1000×1500 matrix whose singular values are Chebyshev points of the second kind in the interval $[10^{-2}, 10]$.

Loss of orthogonality





Effects of the loss of orthogonality in the rational Golub-Kahan algorithm for the approximation of $f^{\diamond}(A)\boldsymbol{b}$, where $f(z) = \sqrt{z}$ and A is a 2000×2000 matrix with logspaced singular values in the interval $[10^{-1}, 10^{2}]$.

Conclusions

- We have proposed a class of rational Krylov methods for the computation of $f^{\diamond}(A)\boldsymbol{b}$.

Conclusions

- We have proposed a class of rational Krylov methods for the computation of $f^{\diamond}(A)\boldsymbol{b}$.
- The projected matrix B_k and the Krylov basis P_k can be computed with a short recurrence by exploiting the quasiseparable structure of B_k.

Conclusions

- We have proposed a class of rational Krylov methods for the computation of $f^{\diamond}(A)\boldsymbol{b}$.
- The projected matrix B_k and the Krylov basis P_k can be computed with a short recurrence by exploiting the quasiseparable structure of B_k.
- We have proved error bounds that relate the convergence rate of these methods to rational approximation of f on $[\sigma_n, \sigma_1]$.

Conclusions

- We have proposed a class of rational Krylov methods for the computation of $f^{\diamond}(A)\boldsymbol{b}$.
- The projected matrix B_k and the Krylov basis P_k can be computed with a short recurrence by exploiting the quasiseparable structure of B_k.
- We have proved error bounds that relate the convergence rate of these methods to rational approximation of f on $[\sigma_n, \sigma_1]$.
- Our numerical experiments indicate that the bounds can accurately predict convergence, and show that the rational Krylov methods converge faster than the polynomial ones when f has low regularity.

Conclusions

- We have proposed a class of rational Krylov methods for the computation of $f^{\diamond}(A)\boldsymbol{b}$.
- The projected matrix B_k and the Krylov basis P_k can be computed with a short recurrence by exploiting the quasiseparable structure of B_k.
- We have proved error bounds that relate the convergence rate of these methods to rational approximation of f on $[\sigma_n, \sigma_1]$.
- Our numerical experiments indicate that the bounds can accurately predict convergence, and show that the rational Krylov methods converge faster than the polynomial ones when f has low regularity.

Thank you for your attention!

Our preprint on arXiv:

A. A. Casulli, I. Simunec, *Computation of generalized matrix functions with rational Krylov methods*, arXiv:2107.12074.

References:

J. B. Hawkins, A. Ben-Israel, *On generalized matrix functions*, Linear and Multilinear Algebra, 1973.

F. Arrigo, M. Benzi, C. Fenu, *Computation of generalized matrix functions*, SIMAX, 2016.

J. L. Aurentz, A. P. Austin, M. Benzi, V. Kalantzis, *Stable computation of generalized matrix functions via polynomial interpolation*, SIMAX, 2019.

S. Massei, L. Robol, *Rational Krylov for Stieltjes matrix functions: convergence and pole selection*, BIT, 2021.

Backup slides

Practical polynomial error bound

Assume $\sigma_1 > 0$, let $1 < \rho \leq \dfrac{\sigma_1 + \sigma_n}{\sigma_1 - \sigma_n}$, and denote by E_ρ the ellipse with vertices at $\frac{1}{2}(\sigma_n^2 + \sigma_1^2) \pm \frac{1}{4}(\rho + \frac{1}{\rho})(\sigma_n^2 - \sigma_1^2)$ and foci at σ_n^2 and σ_1^2.

Theorem (Bernstein)

Let the function g be analytic in the interior of the ellipse E_ρ, and assume that $\max_{z \in E_\rho} |g(z)| \leq M$. Then

$$\min_{p \in \mathcal{P}_k} \|g(z) - p(z)\|_{\infty,[\sigma_n^2, \sigma_1^2]} \leq \frac{2M}{\rho - 1} \rho^{-k}.$$

If the function $f(\sqrt{z}/\sqrt{z})$ is analytic in the interior of E_ρ, we get

$$\|f^\diamond(A)\boldsymbol{b} - \boldsymbol{y}_k\|_2 \leq 4M\sigma_1 \|\boldsymbol{b}\|_2 \frac{\rho}{\rho - 1} \rho^{-k},$$

where $M = \max_{z \in E_\rho} \left| f(\sqrt{z}/\sqrt{z}) \right|$ and $1 < \rho \leq \dfrac{\sigma_1 + \sigma_n}{\sigma_1 - \sigma_n}$.

Shift-and-Invert bound

Let $\boldsymbol{y}_k$ be the approximation to $f^\diamond(A)\boldsymbol{b}$ from the Shift-and-Invert Krylov method with the single pole $\xi = -\sigma_{\min}\sigma_{\max}$. Then

$$\|f^\diamond(A)\boldsymbol{b} - \boldsymbol{y}_k\|_2 \leq 2\|\boldsymbol{b}\|_2 M \sqrt{\frac{\sigma_{\max}}{\sigma_{\min}}} \exp\left(-2k\sqrt{\frac{\sigma_{\min}}{\sigma_{\max}}}\right),$$

where

$$M = \|h(z)\|_{\infty,[0,-\xi^{-1}]}, \qquad h(z) = \frac{f(\sqrt{z^{-1}+\xi})}{\sqrt{z^{-1}+\xi}}.$$

The function h has the property that

$$\frac{f(\sqrt{z})}{\sqrt{z}} = h((z-\xi)^{-1}).$$

Proof sketch of polynomial error bound

If $f = p_{2k-1}$ is an odd polynomial of degree $\leq 2k-1$, then it is not hard to see that $\boldsymbol{y}_k = f^\diamond(A)\boldsymbol{b}$.

For a general f, take an approximating odd polynomial $p_{2k-1} \in \mathcal{P}_{2k-1}$. Then, defining $h_k = f - p_{2k-1}$, we have

$$\|f^\diamond(A)\boldsymbol{b} - \boldsymbol{y}_k\|_2 \leq \|h_k^\diamond(A)\boldsymbol{b}\|_2 + \|P_k h_k^\diamond(B_k) Q_k^T \boldsymbol{b}\|_2.$$

Since we have

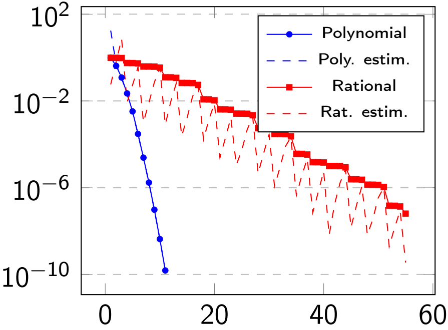
$$\|h_k^\diamond(A)\|_2 \leq \|h_k\|_{\infty,[\sigma_{min\{m,n\}},\sigma_1]},$$
$$\|h_k^\diamond(B_k)\|_2 \leq \|h_k\|_{\infty,[\sigma_n,\sigma_1]},$$

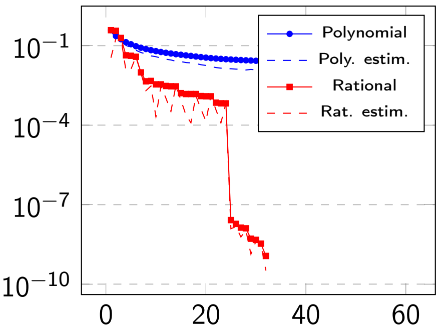
we obtain

$$\|f^\diamond(A)\boldsymbol{b} - \boldsymbol{y}_k\|_2 \leq 2\|\boldsymbol{b}\|_2\|h_k\|_{\infty,[\sigma_n,\sigma_1]}.$$

The statement follows by minimizing over the polynomial p_{2k-1}.

Polynomial vs Rational – Convergence





Comparison between polynomial and rational Krylov for $f^{\diamond}(A)\boldsymbol{b}$, where A is the 8490×8490 adjacency matrix of the directed graph `p2p-Gnutella30` and $\boldsymbol{b}$ is the vector of all ones. Left: $f_1(z) = \sinh(z)$. Right: $f_2(z) = z^{1/3}$.

Polynomial vs Rational – Execution time

function	polynomial			rational		
	k	t_k	E_k	k	t_k	E_k
$\sinh(z)$	11	0.0103	1.54e-10	55	10.0469	6.41e-08
$z^{1/3}$	2000	41.9933	2.48e-03	32	5.6455	1.16e-09

Number of iterations k, execution time t_k in seconds required to achieve tolerance $\texttt{tol} = 10^{-9}$, and actual error E_k at iteration k. The execution times are for the short recurrence implementations, obtained as an average over 10 runs.