Error Estimation and Evaluation of Matrix Functions

Bernd Beckermann

Carl Jagels

Miroslav Pranić

Lothar Reichel

UC3M, Nov. 16, 2010

Outline:

- Approximation of functions of matrices: $f(A)\boldsymbol{v}$ with A large and sparse.
- Polynomial approximation:
  - Reduction to small problem by Arnoldi.
  - Error bounds via the Faber transform.
- Rational approximation:
  - Reduction to small problem by rational Arnoldi.
  - A symmetric, one or several distinct poles: Derivation of short recursion formulas.
  - Computed examples.

Approximation of functions of matrices

$A \in \mathbb{R}^{n\times n}$ large, sparse or structured.

f nonlinear, $\boldsymbol{v}$ unit vector.

Approximate

$$\boldsymbol{w} := f(A)\boldsymbol{v}.$$

Examples:

$$f(t) = \exp(t), \qquad f(t) = \sqrt{t}, \qquad f(t) = \ln(t).$$

If A small, then several approaches are possible, including the use of the spectral factorization

$$A = S\Lambda S^{-1}, \qquad \Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$$

and

$$f(A)\boldsymbol{v} = Sf(\Lambda)S^{-1}\boldsymbol{v}, \quad f(\lambda) = \mathrm{diag}[f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)].$$

Using other factorizations (Schur, Cholesky), squaring and scaling also options.

Reference for small problems: N. J. Higham, Functions of Matrices, SIAM, 2008.

If A large, then first reduce to small matrix.

Polynomial Approximation

m steps of the Arnoldi process with initial vector $\boldsymbol{v}$ gives

$$AV_m = V_m H_m + \boldsymbol{g}_m \boldsymbol{e}_m^T,$$

where

$$\begin{aligned} V_m &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m] \in \mathbb{R}^{n \times m}, \quad \boldsymbol{v}_1 = \boldsymbol{v}, \quad V_m^T V_m = I, \\ H_m &= V_m^T A V_m \quad \text{Hessenberg}, \qquad V_m^T \boldsymbol{g}_m = \boldsymbol{0}, \\ \boldsymbol{e}_m &= [0, \ldots, 0, 1]^T \in \mathbb{R}^m. \end{aligned}$$

Define the Krylov subspace:

$$\mathbf{K}^m(A, \boldsymbol{v}) = \mathrm{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{m-1}\boldsymbol{v}\}.$$

Then

$$\mathrm{range}(V_m) = \mathbf{K}^m(A, \boldsymbol{v}), \quad V_m \boldsymbol{e}_j = p_{j-1}(A)\boldsymbol{v}, \quad p_{j-1} \in \mathbf{P}_{j-1}.$$

Approximate $\boldsymbol{w} := f(A)\boldsymbol{v}$ by

$$\boldsymbol{w}_m := V_m f(H_m)\boldsymbol{e}_1.$$

This is a polynomial approximant:

$$V_m f(H_m)\boldsymbol{e}_1 = p(A)\boldsymbol{v}, \qquad p \in \mathbf{P}_{m-1}.$$

For any polynomial $p \in \mathbf{P}_{m-1}$,

$$p(A)\boldsymbol{v} = p(A)V_m\boldsymbol{e}_1 = V_m p(H_m)\boldsymbol{e}_1.$$

Therefore

$$\begin{aligned} \|f(A)\boldsymbol{v} - V_m f(H_m)\boldsymbol{e}_1\| &= \|(f-p)(A)\boldsymbol{v} - V_m(f-p)(H_m)\boldsymbol{e}_1\| \\ &\leq \|(f-p)(A)\| + \|(f-p)(H_m)\|. \end{aligned}$$

How can we bound the right-hand side?

Crouzeix 2006: There is a universal constant $2 \le C \le 11.5$, such that for any $A \in \mathbb{C}^{n\times n}$ and any function f analytic in the field of values

$$\mathbb{W}(A) = \{\boldsymbol{y}^* A \boldsymbol{y} : \boldsymbol{y} \in \mathbb{C}^n,\ \|\boldsymbol{y}\| = 1\},$$

there holds

$$\|f(A)\| \le C \|f\|_{L_\infty(\mathbb{W}(A))}.$$

Corollary: Let f be analytic in $\mathbb{W}(A)$. Then

$$\|f(A)\boldsymbol{v} - V_m f(H_m)\boldsymbol{e}_1\| \le 23 \min_{p\in\mathbf{P}_{m-1}} \|f - p\|_{L_\infty(\mathbb{W}(A))}.$$

Application of the Faber transform yields a sharper bound.

Error bounds via the Faber transform

Let $\mathbb{E}$ be a convex compact set symmetric with respect to the real axis containing the field of values. Let $\mathbb{E}^c = \mathbb{C} \backslash \mathbb{E}$.

Example: When $A \in \mathbb{R}^{n \times n}$ is symmetric, let $\mathbb{E}$ be the a real interval containing $\lambda(A)$.

Example: When $A \in \mathbb{R}^{n \times n}$ is normal, let $\mathbb{E}$ be the convex hull of $\lambda(A)$.

The **Faber transform** Φ maps the polynomial

$$p(w) = a_0 w^0 + a_1 w^1 + \ldots + a_m w^m, \qquad a_j \in \mathbb{C},$$

to the polynomial

$$\Phi(p)(z) = a_0 f_0(z) + a_1 f_1(z) + \ldots + a_m f_m(z),$$

where f_j is the Faber polynomial of degree j for $\mathbb{E}$.

Example: Let $\mathbb{E} = [-1, 1]$. The Faber polynomials are scaled Chebyshev polynomials of the first kind.

Example: Let $\mathbb{E} = \{z : |z - c| \leq r\}$. Then $f_m(z) = (z - c)^m / r^m$.

Let $\mathbb{D}$ closed unit disc in $\mathbb{C}$ and $\phi : \mathbb{E}^c \to \mathbb{D}^c$ unique conformal mapping with

$$\phi(\infty) = \infty, \qquad \phi'(\infty) > 0, \qquad \psi = \phi^{-1}.$$

The Faber transform Φ is a bijection between functions F analytic in $\mathbb{D}$ and functions f analytic in $\mathbb{E}$:

$$z \in \mathrm{Int}(\mathbb{E}) : \quad f(z) \quad = \quad \Phi(F)(z) = \frac{1}{2\pi i} \int_{\partial \mathbb{E}} \frac{F(\phi(\zeta))}{\zeta - z} d\zeta,$$

$$w \in \mathrm{Int}(\mathbb{D}) : \quad F(w) \quad = \quad \Phi^{-1}(f)(w) = \frac{1}{2\pi i} \int_{\partial \mathbb{D}} \frac{f(\psi(u))}{u - w} du.$$

Example: The Faber polynomials are given by

$$f_m(z) = \Phi(P)(z) \quad \text{for} \quad P(w) = w^m, \qquad m = 0, 1, 2, \dots \ ,$$

i.e., $f_m \in \mathbf{P}_m$ is the polynomial part of $\phi(z)^m$.

Theorem:

$$f = \Phi(F) \quad \Longrightarrow \quad \frac{1}{\|\Phi^{-1}\|}\eta_m(F,\mathbb{D}) \le \eta_m(f,\mathbb{E}) \le \|\Phi\|\eta_m(F,\mathbb{D}),$$

where

$$\|\Phi\| \le 2$$

and

$$\begin{aligned} \eta_m(f,\mathbb{E}) &= \min_{p\in\mathbf{P}_m} \|f-p\|_{L_\infty(\mathbb{E})}, \\ \eta_m(F,\mathbb{D}) &= \min_{P\in\mathbf{P}_m} \|F-P\|_{L_\infty(\mathbb{D})}. \end{aligned}$$

Theorem: Let $\mathbb{E}$ be convex and $\mathbb{W}(A) \subset \mathbb{E}$. Define

$$\Phi_+(F)(z) = \Phi(F)(z) + F(0)$$

for F analytic in $\mathbb{D}$. Then

$$\|\Phi_+\| \leq 2.$$

Moreover, for F analytic in $\mathbb{D}$,

$$\|\Phi_+(F)(A)\| \leq 2\|F\|_{L_\infty(\mathbb{D})}.$$

"Proof:" Use the representation

$$\Phi_+(F)(A) = \Phi(F)(A) + F(0)I = \frac{1}{\pi}\int_{\partial\mathbb{D}} F(w)K(w)\,|dw|,$$

where

$$\begin{aligned} K(w) \quad &:= \quad \frac{1}{2i}\Bigg(\psi'(w)(\psi(w)I - A)^{-1}\frac{dw}{|dw|} \\ &\qquad - \overline{\psi'(w)}(\overline{\psi(w)}I - A^*)^{-1}\frac{\overline{dw}}{|dw|}\Bigg) \end{aligned}$$

and A^* is the conjugate transpose of A.

Corollary: Let $\mathbb{E}$ be convex and $\mathbb{W}(A) \subset \mathbb{E}$. Let

$$f \in \mathcal{A}(\mathbb{E}), \qquad F := \Phi^{-1}(f), \qquad P \in \mathbf{P}_m,$$

and

$$p(z) := \Phi_+(P)(z) - F(0).$$

Then

$$||f(A) - p(A)|| \leq 2\,||F - P||_{L_\infty(\mathbb{D})}.$$

Theorem: Let $\mathbb{W}(A) \subset \mathbb{E}$ and let $f = \Phi(F)$ be analytic in $\mathbb{E}$. Then

$$\|f(A)\boldsymbol{v} - V_{m+1}f(H_{m+1})\boldsymbol{e}_1\| \leq 4\eta_m(F, \mathbb{D}).$$

The Arnoldi process gives accurate results if F can be approximated well by a polynomial of fairly low degree on $\mathbb{D}$.

Related results shown by Druskin, Knizhnerman, Hochbruck, Lubich,...

Rational Arnoldi

Determine an orthonormal basis $\{\boldsymbol{v}_j\}_{j=1}^{m+1}$ of the rational Krylov subspace

$$(q(A))^{-1}\text{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^m\boldsymbol{v}\}$$

where

$$q(z) := (z - z_1)(z - z_2)\cdots(z - z_m).$$

Let $z_0 \in \mathbb{C}$, $z_0 \neq z_j$, $j \geq 1$, and let $\boldsymbol{v}_1 = \boldsymbol{v}$.
For $j = 1, 2, \ldots$, determine $\boldsymbol{v}_{j+1}$ by orthonormalizing

$$(z_j - z_0)(z_j I - A)^{-1}(A - z_0 I)\boldsymbol{v}_j$$

against available basis vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_j$.

This defines the coefficients $h_{k,j}$ in

$$\begin{aligned} h_{j+1,j}\boldsymbol{v}_{j+1} &= (z_j - z_0)(z_j I - A)^{-1}(A - z_0 I)\boldsymbol{v}_j \\ &\quad + h_{1,j}\boldsymbol{v}_1 + h_{2,j}\boldsymbol{v}_2 + \ldots + h_{j,j}\boldsymbol{v}_j, \quad j = 1, 2, \ldots . \end{aligned}$$

In matrix notation, with $z_{m+1} = \infty$,

$$(A - z_0 I)V_{m+1}(I + H_{m+1}D_{m+1}) = V_{m+1}H_{m+1} + h_{m+2,m+1}\boldsymbol{v}_{m+2}\boldsymbol{e}_{m+1}^T,$$

where

$$D_{m+1} = \mathrm{diag}[(z_1 - z_0)^{-1}, (z_2 - z_0)^{-1}, \ldots, (z_{m+1} - z_0)^{-1}].$$

Projected matrix:

$$A_{m+1} := V_{m+1}^T A V_{m+1} = z_0 I + H_{m+1}(I + H_{m+1} D_{m+1})^{-1}.$$

Simplifies to the (standard) Arnoldi projection

$$A_{m+1} = H_{m+1}$$

used for polynomial approximation when

$$z_0 = 0, \qquad z_j = \infty, \quad j = 1, 2, \dots \; .$$

Theorem: Let $\mathbb{W}(A) \subset \mathbb{E}$ and let $f = \Phi(F)$ be analytic in $\mathbb{E}$. Let $z_1, z_2, \ldots, z_m \notin \mathbb{E}$ and $z_{m+1} = \infty$. Then

$$\|f(A)\boldsymbol{v} - V_{m+1} f(H_{m+1})\boldsymbol{e}_1\| \leq 4\eta_m^Q(F, \mathbb{D}),$$

where

$$Q(w) = (w - w_1)(w - w_2)\cdots(w - w_m), \quad w_j = \Phi^{-1}(z_j)$$

and

$$\eta_m^Q(F, \mathbb{D}) = \min_{P \in \mathbf{P}_m} \|F - \frac{P}{Q}\|_{L_\infty(\mathbb{D})}$$

Approximation of Markov functions

Let $d\mu$ be a positive measure with support in $[\alpha, \beta]$, $-\infty \leq \alpha < \beta < \infty$. Then

$$f(z) = \int_{\alpha}^{\beta} \frac{d\mu(x)}{z - x}$$

is a Markov function. Note: f analytic in $\bar{\mathbb{C}} \setminus [\alpha, \beta]$.

Examples:

$$\begin{aligned} f(z) &= \frac{\log(1+z)}{z}, \\ f(z) &= z^{-\gamma}, \qquad 0 < \gamma < 1, \qquad z \in \mathbb{C} \setminus \mathbb{R}_-, \end{aligned}$$

are Markov functions.

Define the polynomial

$$\widetilde{q}(w) = \prod_{j=1}^{m} (w - w_j), \qquad 1 < |w_j| \le \infty,$$

with real or complex conjugate zeroes.

Introduce the Blaschke product

$$B(w) = \frac{w^m \widetilde{q}(1/w)}{\widetilde{q}(w)} = \prod_{j=1}^{m} \frac{1 - w_j w}{w - w_j}.$$

Theorem: Let $\mathbb{E}$ be compact, convex, and symmetric with respect to the real axis. Let f be a Markov function with

$$-\infty \leq \alpha < \beta < \gamma = \min\{\mathrm{Re}(z) : z \in \mathbb{E}\}.$$

Then $\widetilde{f} = \mathcal{F}^{-1}(f)$ is a Markov function,

$$\widetilde{f}(w) = \int_{\alpha}^{\beta} \frac{\phi'(x) d\mu(x)}{w - \phi(x)} =: \int \frac{d\widetilde{\mu}(x)}{w - x}.$$

Theorem (cont'd): Let $R = P/\widetilde{q}$ with $P \in \mathbf{P}_{m-1}$ be a rational interpolant of $\widetilde{f}$ with prescribed poles w_j at the reflected points $1/\overline{w}_j$, $j = 1, 2, \ldots, m$. Define

$$\widetilde{r}(w) = R(w) + B(w)\left(\frac{\widetilde{f}(1) - R(1)}{2B(1)} + \frac{\widetilde{f}(-1) - R(-1)}{2B(-1)}\right).$$

Then $\widetilde{r} \in \mathbf{P}_m/\widetilde{q}$ and

$$\eta_m^{\widetilde{q}}(\mathcal{F}^{-1}(f), \mathbb{D}) \leq \|\widetilde{f} - \widetilde{r}\|_{L_\infty(\mathbb{D})} \leq \frac{\|f\|_{L_\infty(\mathbb{E})}}{|\phi(\beta)|} \max_{y \in \phi([\alpha,\beta])} \frac{1}{|B(y)|}.$$

A bound for rational approximants of f on $\mathbb{E}$ is given by

$$\eta_m^q(f, \mathbb{E}) \leq 2\, \eta_m^{\widetilde{q}}(\mathcal{F}^{-1}(f), \mathbb{D}).$$

Rational Lanczos with a fixed pole

Inspired by Druskin and Knizhnerman (SIMAX, '98).

Let A be symmetric and nonsingular. Determine orthonormal bases for the rational Krylov subspaces

$$\mathbf{K}^{\ell,m}(A,\boldsymbol{v}) := \operatorname{span}\{A^{-\ell+1}\boldsymbol{v},\ldots,A^{-1}\boldsymbol{v},\boldsymbol{v},A\boldsymbol{v},\ldots,A^{m-1}\boldsymbol{v}\}.$$

We consider $m = i\ell$ for $i = 1, 2, 3, \ldots$

A Lanczos-like method for orthogonalizing $\mathbf{K}^{1,2}(A, \boldsymbol{v}), \mathbf{K}^{2,3}(A, \boldsymbol{v}), \ldots$ for A SPD $(i = 1)$.

Let

$$\{\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_{-1}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{m-1}, \boldsymbol{v}_{-m+1}, \boldsymbol{v}_m\}$$

be an ON basis for $\mathbf{K}^{m,m+1}(A, \boldsymbol{v})$.

Represent the $\boldsymbol{v}_j$ in terms of monic orthogonal Laurent polynomials:

$$\phi_j(x) := \begin{cases} x^j + \displaystyle\sum_{k=-j+1}^{j-1} c_{j,k} x^k, & j = 0, 1, \ldots, m, \\ x^j + \displaystyle\sum_{k=j+1}^{-j} c_{j,k} x^k, & j = -1, -2, \ldots, -m+1. \end{cases}$$

with

$$\begin{aligned} \boldsymbol{w}_j &= \phi_j(A)\boldsymbol{v}, \\ \boldsymbol{v}_j &= \boldsymbol{w}_j / ||\boldsymbol{w}_j||. \end{aligned}$$

Orthogonality with respect to the inner product

$$(p, q) := (p(A)\boldsymbol{v})^T (q(A)\boldsymbol{v})$$

By symmetry of A:

$$(xp, q) = (p, xq).$$

Theorem (Njåstad and Thron, '83): The orthogonal Laurent polynomials ϕ_j satisfy short recursion relations.

Survey: Jones and Njåstad, JCAM, '91

Recent work: Díaz-Mendoza, Gonzáles-Vera, Jiménez Paiz, and Njåstad.

Simplest recursions when A SPD $\implies$ trailing coefficient of every ϕ_j is nonvanishing.

Algorithm: Compute ON basis $\{\boldsymbol{v}_k\}_{k=-m+1}^{m}$ of $\mathbf{K}^{m,m+1}(A,\boldsymbol{v})$.

$\delta_0 := ||\boldsymbol{v}||$; $\boldsymbol{v}_0 := \boldsymbol{v}/\delta_0$;

$\boldsymbol{u} := A\boldsymbol{v}_0$; $\alpha_0 := \boldsymbol{v}_0^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_0\boldsymbol{v}_0$;

$\delta_1 := ||\boldsymbol{u}||$; $\boldsymbol{v}_1 := \boldsymbol{u}/\delta_1$;

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

$\boldsymbol{w} := A^{-1}\boldsymbol{v}_k$;

$\beta_{-k+1} := \boldsymbol{v}_{-k+1}^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - \beta_{-k+1}\boldsymbol{v}_{-k+1}$;

$\beta_k := \boldsymbol{v}_k^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - \beta_k\boldsymbol{v}_k$;

$\delta_{-k} := ||\boldsymbol{w}||$; $\boldsymbol{v}_{-k} := \boldsymbol{w}/\delta_{-k}$;

$\boldsymbol{u} := A\boldsymbol{v}_{-k}$;

$\alpha_{-k} := \boldsymbol{v}_{-k}^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{-k}\boldsymbol{v}_{-k}$;

$\alpha_k := \boldsymbol{v}_k^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_k\boldsymbol{v}_k$;

$\delta_{k+1} := ||\boldsymbol{u}||$; $\boldsymbol{v}_{k+1} := \boldsymbol{u}/\delta_{k+1}$;

end

Let

$$V_{2m-1} = [\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_{-1}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{m-1}, \boldsymbol{v}_{-m+1}].$$

From the recursion formulas:

$$H_{2m-1} = V_{2m-1}^T A V_{2m-1}, \qquad A V_{2m-1} = V_{2m} \widetilde{H}_{2m-1},$$

$$G_{2m} = V_{2m}^T A^{-1} V_{2m}, \qquad A^{-1} V_{2m} = V_{2m+1} \widetilde{G}_{2m}.$$

Odd numbered columns of H_{2m-1} have at most 3 nontrivial elements and even numbered columns have at most 5.

Example:

$$
H_8 = \begin{bmatrix}
* & * & & & & & & \\
* & * & * & * & & & & \\
 & * & * & * & & & & \\
 & * & * & * & * & * & & \\
 & & & * & * & * & & \\
 & & & * & * & * & * & * \\
 & & & & & * & * & * \\
 & & & & & * & * & *
\end{bmatrix}.
$$

Example:

$$G_8 = \begin{bmatrix} * & * & * & & & & & \\ * & * & * & & & & & \\ * & * & * & * & * & & & \\ & & * & * & * & & & \\ & & * & * & * & * & * & \\ & & & & * & * & * & \\ & & & & * & * & * & * \\ & & & & & & * & * \end{bmatrix}.$$

Both G_{2m} and H_{2m} are pentadiagonal.

Moreover,

$$H_{2m}G_{2m} = I + e_{2m}u_{2m}^T,$$

where only the last two entries of u_{2m} may be nonvanishing.

A Lanczos-like method for orthogonalizing $\mathbf{K}^{1,2}(A,\boldsymbol{v}),\mathbf{K}^{2,3}(A,\boldsymbol{v}),\ldots$ for A indefinite $(i=1)$.

The trailing coefficient of the ϕ_j may vanish $\Longrightarrow$ new derivation of recursion formulas. There may be 5-term recursions.

Example:

$$H_7 = \begin{bmatrix} * & * & & & & & \\ * & * & * & * & & & \\ & * & * & * & & & \\ & * & * & * & * & * & \\ & & & * & & & \\ & & & * & & * & * \\ & & & & & * & * \end{bmatrix}$$

A Lanczos-like method for orthogonalizing $\mathbf{K}^{1,2}(A,\boldsymbol{v}),\mathbf{K}^{2,4}(A,\boldsymbol{v}),\ldots$ for A SPD.

Orthogonal Laurent polynomials:

$$\phi_j(x) := \begin{cases} x^j + \displaystyle\sum_{k=-\lfloor (j-1)/2 \rfloor}^{j-1} c_{j,k} x^k, & j = 0,1,\ldots,m, \\ x^j + \displaystyle\sum_{k=j+1}^{-2j} c_{j,k} x^k, & j = -1,-2,\ldots,-m+1. \end{cases}$$

with

$$\boldsymbol{v}_j = \frac{\phi_j(A)\boldsymbol{v}}{\|\phi_j(A)\boldsymbol{v}\|}, \qquad j = 0,1,2,-1,3,4,-2,5,\ldots\ .$$

Then

$$\{\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_{-1}, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{-m+1}, \boldsymbol{v}_{2m-1}\}$$

an orthonormal basis for $\mathbf{K}^{m,2m}(A, \boldsymbol{v})$.

Example: The matrix H_{10} is pentadiagonal:

$$
\begin{bmatrix}
+ & + & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
+ & + & + & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & + & + & + & + & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & + & + & + & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & + & + & + & + & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & + & + & + & + & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & + & + & + & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & + & + & + & + & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & + & + & + \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & + & +
\end{bmatrix}.
$$

A Lanczos-like method for orthogonalizing $\mathbf{K}^{1,i}(A,\boldsymbol{v}), \mathbf{K}^{2,2i}(A,\boldsymbol{v}), \ldots$ for A SPD, $i \geq 2$.

We want to determine orthonormal basis of

$$\boldsymbol{v}, A\boldsymbol{v}, A^2\boldsymbol{v}, \ldots A^i\boldsymbol{v}, A^{-1}\boldsymbol{v}, A^{i+1}\boldsymbol{v}, \ldots, A^{2i}\boldsymbol{v}, A^{-2}\boldsymbol{v}, A^{2i+1}\boldsymbol{v}, \ldots .$$

Associated orthogonal Laurent polynomials

$$\phi_0, \phi_1, \ldots \phi_i, \phi_{-1}, \phi_{i+1}, \ldots, \phi_{2i}, \phi_{-2}, \ldots$$

of the form

$$\phi_j(x) := \begin{cases} x^j + \displaystyle\sum_{k=-\lfloor (j-1)/i \rfloor}^{j-1} c_{j,k} x^k, & j = 1, 2, 3, \dots\ , \\ x^j + \displaystyle\sum_{k=j+1}^{-ij} c_{j,k} x^k, & j = -1, -2, -3, \dots\ . \end{cases}$$

with $\phi_0(x) := 1$.

Example: Let $m = 2$ and $i = 3$. Then H_8 of the form

$$\begin{bmatrix} x & x & 0 & 0 & 0 & 0 & 0 & 0 \\ x & x & x & 0 & 0 & 0 & 0 & 0 \\ 0 & x & x & x & 0 & 0 & 0 & 0 \\ 0 & 0 & x & x & x & x & 0 & 0 \\ 0 & 0 & 0 & x & x & x & 0 & 0 \\ 0 & 0 & 0 & x & x & x & x & 0 \\ 0 & 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & x \end{bmatrix}$$

Computed examples.

$A \in \mathbb{R}^{1000\times 1000}, \quad \boldsymbol{v} \in \mathbb{R}^{1000}$ random.

Tabulate error in approximations obtained by 42 steps of standard Lanczos or rational Lanczos.

Example 1. $A = n^2\,[-1 \;\; 2 \quad -1]$ tridiagonal, SPD;
$n = 1000$.

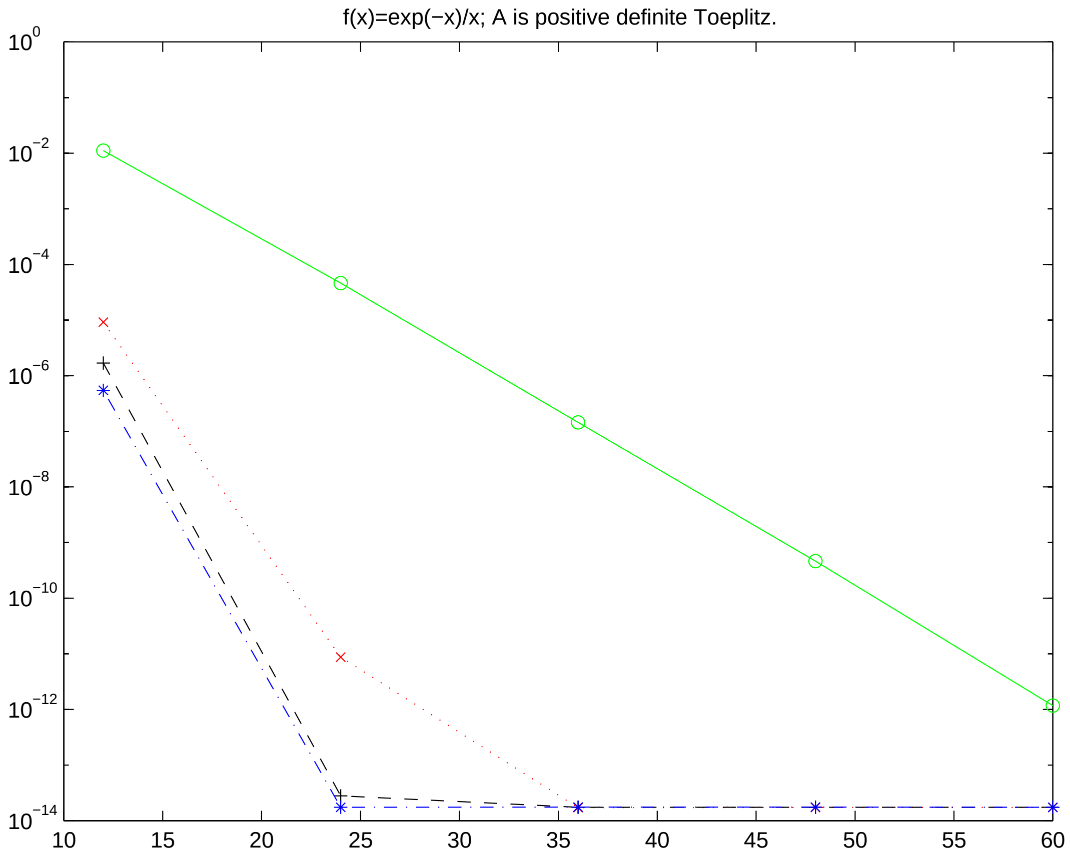
$f(x)$	Lan. (42)	Rat. Lan. (21,22)	Rat. Lan. (14,29)
$\exp(x)$	$2.3 \cdot 10^{-6}$	$3.4 \cdot 10^{-15}$	$3.8 \cdot 10^{-15}$
$\sqrt{x}$	$1.3 \cdot 10^{0}$	$2.1 \cdot 10^{-2}$	$3.6 \cdot 10^{-2}$
$\exp(-\sqrt{x})$	$1.0 \cdot 10^{-3}$	$2.5 \cdot 10^{-13}$	$2.6 \cdot 10^{-13}$
$\ln(x)$	$1.8 \cdot 10^{-1}$	$3.4 \cdot 10^{-4}$	$7.1 \cdot 10^{-4}$
$\exp(x)/x$	$2.4 \cdot 10^{-7}$	$3.5 \cdot 10^{-16}$	$3.9 \cdot 10^{-16}$

Example 2. $A = [a_{j,k}]$, $a_{j,k} = 1/(1 + |j - k|)$, Toeplitz, SPD, 1000×1000.

$f(x)$	Lan. (42)	Rat. Lan. (21,22)	Rat. Lan. (14,29)
$\exp(x)$	$8.2 \cdot 10^{-15}$	$8.2 \cdot 10^{-15}$	$8.1 \cdot 10^{-15}$
$\sqrt{x}$	$1.9 \cdot 10^{-11}$	$1.0 \cdot 10^{-14}$	$1.0 \cdot 10^{-14}$
$\exp(-\sqrt{x})$	$1.9 \cdot 10^{-11}$	$6.9 \cdot 10^{-15}$	$7.0 \cdot 10^{-15}$
$\ln(x)$	$3.0 \cdot 10^{-10}$	$1.4 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$
$\exp(x)/x$	$7.6 \cdot 10^{-9}$	$1.6 \cdot 10^{-14}$	$1.5 \cdot 10^{-14}$

Example 2 (cont'd)

$f(x) = \exp(x)/x$. Approximation errors from top to bottom: Lanczos, rational Lanczos for $i = 1, 2, 3$.

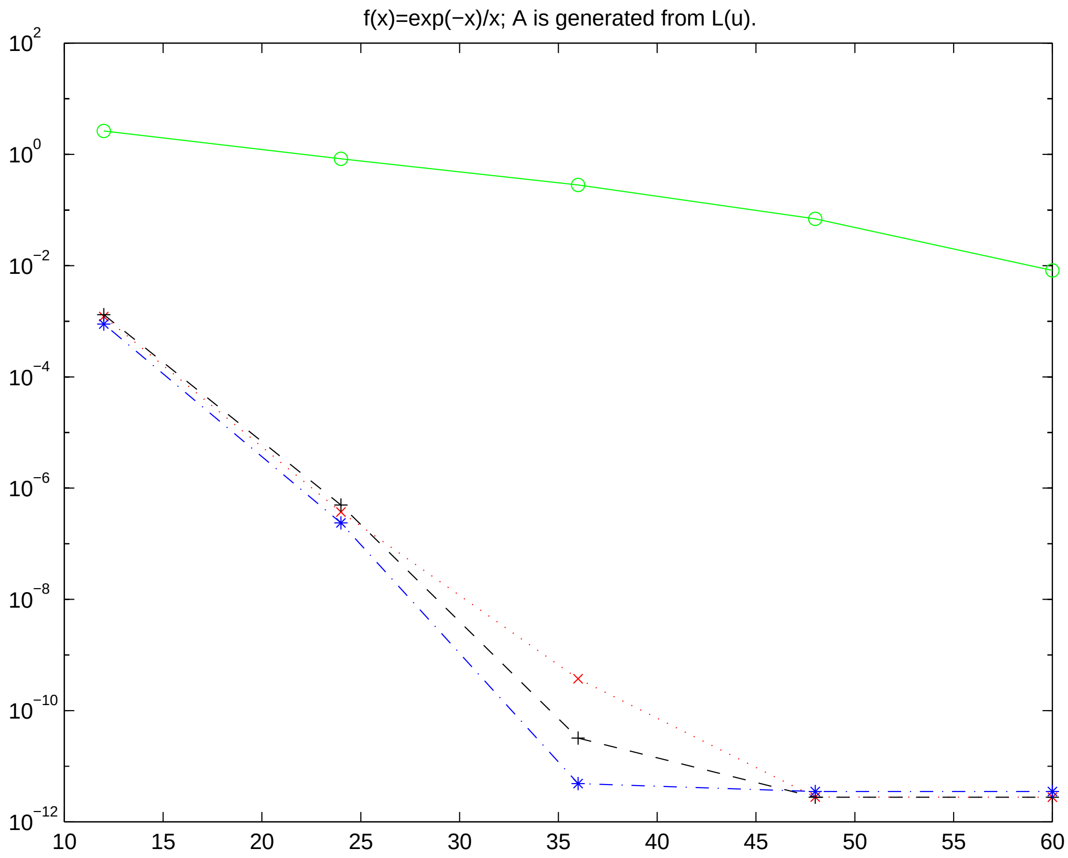


Example 3. Matrix $A \in \mathbb{R}^{1600\times 1600}$ stems from the discretization of the differential operator $L(u) = \frac{1}{10}u_{xx} - 100u_{yy}$ on the unit square.

$f(x)$	Lan. (42)	Rat. Lan. (21,22)	Rat. Lan. (14,29)
$1/\sqrt{x}$	$1.4 \cdot 10^{-2}$	$5.6 \cdot 10^{-13}$	$2.7 \cdot 10^{-12}$

Example 3 (cont'd)

$f(x) = \exp(x)/x$. Approximation errors from top to bottom: Lanczos, rational Lanczos for $i = 1, 2, 3$.



Orthogonal rational functions with several fixed poles

$d\mu$: nonnegative measure on (part of) the real axis

$$(f, g) = \int_a^b f(x)g(x)d\mu : \qquad \text{inner product}$$

$\mathbf{P}$: space of all polynomials with real coefficients

$$\mathbf{Q} = \text{span}\left\{ \frac{1}{(x-\alpha_k)^s} \;:\; s \in \mathbb{N}, \;\; \alpha_k \notin [a, b] \cup \{\infty\} \right\} :$$

space of rational functions with real or complex conjugate poles α_k.

Assume poles ordered so that

$$\operatorname{Im}(\alpha_j) > 0 \qquad \Rightarrow \qquad \operatorname{Im}(\alpha_{j+1}) = -\operatorname{Im}(\alpha_j)$$

Replace for all $s = 1, 2, \dots$,

$$\frac{1}{(x-\alpha_j)^s} \quad \text{and} \quad \frac{1}{(x-\alpha_{j+1})^s}$$

by

$$\frac{1}{(x^2+p_j x+q_j)^s} \quad \text{and} \quad \frac{x}{(x^2+p_j x+q_j)^s},$$

where

$$x^2+p_j x+q_j = (x-\alpha_j)(x-\alpha_{j+1}), \qquad p_j, q_j \in \mathbb{R}.$$

Define linear space

$$\mathbf{P} + \mathbf{Q} \quad = \quad \text{span}\{1,\ x^s,\ \frac{1}{(x-\alpha_k)^s},\ \frac{1}{(x^2+p_jx+q_j)^s},$$

$$\frac{x}{(x^2+p_jx+q_j)^s} \ : \quad s \in \mathbb{N},$$

$$\alpha_k \in \mathbb{R}\backslash[a,b], \alpha_j \in \mathbb{C}\backslash\mathbb{R},\ |\alpha_k|, |\alpha_j| < \infty\}.$$

Let

$$\Psi = \{\psi_0, \psi_1, \psi_2, \dots\ \}$$

denote an elementary basis for $\mathbf{P} + \mathbf{Q}$:
$\psi_0(x) = 1$ and $\psi_\ell(x)$, $\ell = 1, 2, \dots$, one of the functions

$$x^s, \ \frac{1}{(x-\alpha_k)^s}, \ \frac{1}{(x^2+p_jx+q_j)^s}, \ \frac{x}{(x^2+p_jx+q_j)^s}$$

for some positive integers k, j, and s.

Gram-Schmidt process applied to the basis Ψ yields basis of orthonormal rational functions

$$\Phi = \{\phi_0, \phi_1, \phi_2, \dots\ \}.$$

The recursion relations for the ϕ_j depend on the ordering of the basis functions ψ_j of Ψ.

We say the ordering of Ψ is natural if for all integers $s \geq 1$, all real poles α_k, and all pairs $\{p_j, q_j\}$,

- $x^{s-1} \prec x^s$,
- $\dfrac{1}{(x-\alpha_k)^s} \prec \dfrac{1}{(x-\alpha_k)^{s+1}}$,

- $$\frac{1}{(x^2 + p_j x + q_j)^s} \prec \frac{x}{(x^2 + p_j x + q_j)^s} \prec \frac{1}{(x^2 + p_j x + q_j)^{s+1}},$$
- $$\frac{x}{(x^2 + p_j x + q_j)^s} \prec \frac{1}{(x^2 + p_j x + q_j)^{s+1}} \prec \frac{x}{(x^2 + p_j x + q_j)^{s+1}}.$$

Theorem: Let the basis $\Psi = \{\psi_0, \psi_1, \psi_2, \dots\ \}$ be naturally ordered. Let

- every sequence of m_1 consecutive basis functions $\psi_k, \psi_{k+1}, \dots, \psi_{k+m_1-1}$ contain at least one power x^ℓ,
- there be at most m_2 basis functions between every pair of functions

$$\left\{ \frac{1}{(x^2 + p_j x + q_j)^s}, \frac{x}{(x^2 + p_j x + q_j)^s} \right\}, \quad s = 1, 2, \dots$$

Then the orthonormal Laurent polynomials

$$\phi_0, \phi_1, \phi_2, \dots$$

satisfy a $(2m+1)$-term recurrence relation of the form

$$x\phi_k(x) = \sum_{i=-m}^{m} c_{k,k+i}\phi_{k+i}(x), \qquad k = 0, 1, 2, \dots\ ,$$

with $m = \max\{m_1, m_2 + 1\}$. Here $c_{k,k+i}$ and ϕ_{k+i} with $k + i < 0$ are zero.

Note:

- If $\mathbf{Q} \neq \emptyset$, then we may order the basis Ψ to get the smallest possible value of m, which is 2. This gives a 5-term recursion formula.
- If $\mathbf{Q} = \emptyset$, then $m = 1$ and we obtain the 3-term recusion formula for orthogonal polynomials.

Theorem: Let the basis $\Psi = \{\psi_0, \psi_1, \psi_2, \dots\ \}$ be naturally ordered. Let

- every sequence of m_1 consecutive basis functions $\psi_k, \psi_{k+1}, \dots, \psi_{k+m_1-1}$ contain at least one power $(x - \alpha_\ell)^{-t}$,
- there be at most m_2 basis functions between every pair of functions

$$\left\{ \frac{1}{(x^2 + p_j x + q_j)^s}, \frac{x}{(x^2 + p_j x + q_j)^s} \right\}, \quad s = 1, 2, \dots$$

Then the orthonormal Laurent polynomials

$$\phi_0, \phi_1, \phi_2, \ldots$$

satisfy a $(2m+1)$-term recurrence relation of the form

$$\frac{1}{x-\alpha_\ell}\phi_k(x) = \sum_{i=-m}^{m} c_{k,k+i}^{(\ell)} \phi_{k+i}(x), \quad k = 0, 1, 2, \ldots \ ,$$

with $m = \max\{m_1, m_2 + 1\}$. Here $c_{k,k+i}$ and ϕ_{k+i} with $k + i < 0$ are zero.

Note: Let

$$\mathbf{P}+\mathbf{Q} = \text{span}\{1, x, \ldots, x^{\ell}, x^{-1}, x^{\ell+1}, \ldots, x^{2\ell}, x^{-2}, x^{2\ell+1}, \ldots\}$$

This defines a basis Ψ with $m_1 = \ell + 1$ and $m_2 = 0$. Therefore,

$$\frac{\phi_k(x)}{x}$$

satisfies a recursion formula with $2\ell + 3$ terms. Note that

$$x\phi_k(x)$$

satisfies a 5-term recursion.

Short recursion formulas for

$$\frac{1}{x^2 + p_j x + q_j}\phi_k(x), \quad \frac{x}{x^2 + p_j x + q_j}\phi_k(x)$$

also can be established.

Extensions (work in progress):

- Application to the evaluation of matrix functions $f(A)b$ for nonsymmetric matrices.
- New derivation of rational Gauss quadrature rules.

Muchas Gracias