Rational minimax approximation via adaptive barycentric representations

Silviu Filip, CAIRN team, Univ Rennes, Inria, CNRS, IRISA
joint work with Bernhard Beckermann, Yuji Nakatsukasa
and Lloyd N. Trefethen

January 22, 2018

Rational functions

Why are they important?

$\rightarrow$ powerful **approximations** near singularities or on unbounded domains

Rational functions

Why are they important?

$\rightarrow$ powerful **approximations** near singularities or on unbounded domains

Some applications:

- elementary + special functions
- recursive filter design
- matrix exponentials & stiff PDEs
- optimal control problems
- ...

Rational minimax approximation

Input: $f \in \mathcal{C}([a,b])$, target type $(m,n) \in \mathbb{N}^2$

Output: $r^* \in \mathcal{R}_{m,n} = \left\{ \frac{p}{q}, p \in \mathbb{R}_m[x], q \in \mathbb{R}_n[x] \right\}$ s.t.

$$\|f - r^*\|_\infty \text{ is } \textbf{minimal}.$$

$\rightarrow$ denote this **minimax** error with $E_{m,n}(f)$

Rational minimax approximation

Input: $f \in \mathcal{C}([a,b])$, target type $(m,n) \in \mathbb{N}^2$

Output: $r^* \in \mathcal{R}_{m,n} = \left\{ \frac{p}{q}, p \in \mathbb{R}_m[x], q \in \mathbb{R}_n[x] \right\}$ s.t.

$\|f - r^*\|_\infty$ is **minimal**.

$\rightarrow$ denote this **minimax** error with $E_{m,n}(f)$

$\rightarrow$ theoretical results:

Rational minimax approximation

Input: $f \in \mathcal{C}([a, b])$, target type $(m, n) \in \mathbb{N}^2$

Output: $r^* \in \mathcal{R}_{m,n} = \left\{ \frac{p}{q}, p \in \mathbb{R}_m[x], q \in \mathbb{R}_n[x] \right\}$ s.t.

$\|f - r^*\|_\infty$ is **minimal**.

→ denote this **minimax** error with $E_{m,n}(f)$

→ theoretical results:

- existence & unicity of r^* [de la Vallée Poussin, Walsh]

Rational minimax approximation

Input: $f \in \mathcal{C}([a,b])$, target type $(m,n) \in \mathbb{N}^2$

Output: $r^* \in \mathcal{R}_{m,n} = \left\{ \frac{p}{q}, p \in \mathbb{R}_m[x], q \in \mathbb{R}_n[x] \right\}$ s.t.

$$\|f - r^*\|_\infty \text{ is } \textbf{minimal}.$$

→ denote this **minimax** error with $E_{m,n}(f)$

→ theoretical results:

- existence & unicity of r^* [de la Vallée Poussin, Walsh]
- presence of the *defect*:

  → if $r^* = p^*/q^*$ in irreducible form, then its defect is

$$d = \min\{m - \deg p^*, n - \deg q^*\}$$

Rational minimax approximation

Input: $f \in \mathcal{C}([a,b])$, target type $(m,n) \in \mathbb{N}^2$

Output: $r^* \in \mathcal{R}_{m,n} = \left\{ \frac{p}{q}, p \in \mathbb{R}_m[x], q \in \mathbb{R}_n[x] \right\}$ s.t.

$$\|f - r^*\|_\infty \text{ is } \textbf{minimal}.$$

$\rightarrow$ denote this **minimax** error with $E_{m,n}(f)$

$\rightarrow$ theoretical results:

- existence & unicity of r^* [de la Vallée Poussin, Walsh]
- presence of the *defect*:

  $\rightarrow$ if $r^* = p^*/q^*$ in irreducible form, then its defect is

$$d = \min\{m - \deg p^*, n - \deg q^*\}$$

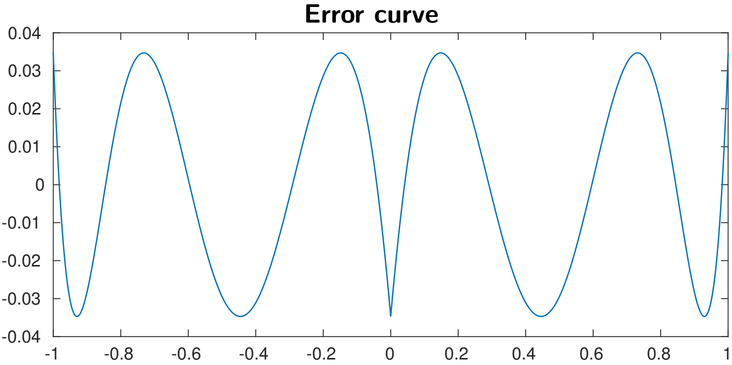
- **Alternation Theorem** [Achieser 1930]:
  $\rightarrow f - r^*$ equioscillates at least $m + n + 2 - d$ times

A classic example: $f(x) = |x|, x \in [-1, 1]$

A classic example: $f(x) = |x|, x \in [-1, 1]$

$\rightarrow$ consider best uniform approximations:

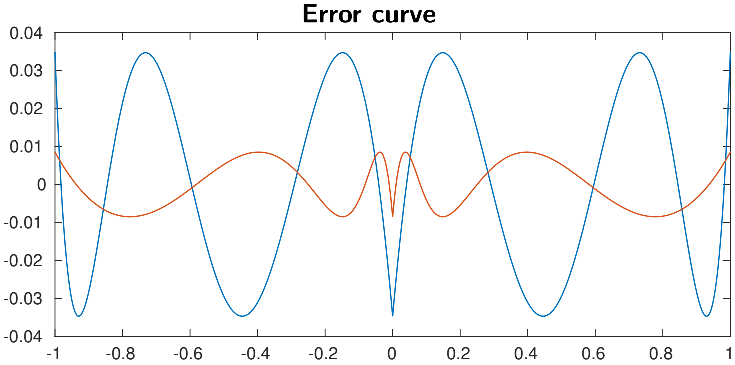
- degree 8 polynomial



A classic example: $f(x) = |x|, x \in [-1, 1]$

→ consider best uniform approximations:

- degree 8 polynomial
- type $(4, 4)$ rational function



A classic example: $f(x) = |x|, x \in [-1, 1]$

$\rightarrow$ asymptotic behavior

$E_{n,0}(f) \sim \beta/n,$	$\beta = 0.2801...$	[Varga & Carpenter 1985]
$E_{n,n}(f) \sim 8e^{-\sqrt{n}},$		[Newman 1964, Stahl 1993]

A classic example: $f(x) = |x|, x \in [-1, 1]$

$\rightarrow$ asymptotic behavior

$$E_{n,0}(f) \sim \beta/n, \qquad \beta = 0.2801... \quad \text{[Varga \& Carpenter 1985]}$$
$$E_{n,n}(f) \sim 8e^{-\sqrt{n}}, \qquad \text{[Newman 1964, Stahl 1993]}$$

$\rightarrow$ rational minimax approximations can be difficult to compute
e.g. [Varga, Ruttan & Carpenter 1991] conjecture Stahl's result using 200-digit arithmetic for $n \leqslant 80$

A classic example: $f(x) = |x|, x \in [-1, 1]$

$\to$ asymptotic behavior

$$E_{n,0}(f) \sim \beta/n, \qquad \beta = 0.2801... \quad \text{[Varga \& Carpenter 1985]}$$
$$E_{n,n}(f) \sim 8e^{-\sqrt{n}}, \qquad \text{[Newman 1964, Stahl 1993]}$$

$\to$ rational minimax approximations can be difficult to compute
e.g. [Varga, Ruttan & Carpenter 1991] conjecture Stahl's result using 200-digit arithmetic for $n \leqslant 80$

$\to$ **codes** (the Remez algorithm):

- Maple: `numapprox[minimax]`
- Mathematica: `MinimaxApproximation`

A classic example: $f(x) = |x|, x \in [-1, 1]$

→ asymptotic behavior

$E_{n,0}(f) \sim \beta/n,$	$\beta = 0.2801...$	[Varga & Carpenter 1985]
$E_{n,n}(f) \sim 8e^{-\sqrt{n}},$		[Newman 1964, Stahl 1993]

→ rational minimax approximations can be difficult to compute
e.g. [Varga, Ruttan & Carpenter 1991] conjecture Stahl's result using 200-digit arithmetic for $n \leqslant 80$

→ **codes** (the Remez algorithm):

- Maple: `numapprox[minimax]`
- Mathematica: `MinimaxApproximation`
- Chebfun (Matlab): `minimax`

Barycentric representations

$\rightarrow$ many different ways of representing rational functions

Barycentric representations

$\rightarrow$ many different ways of representing rational functions

$\rightarrow$ barycentric form for type (n, n) rational functions

$$r(z) = \frac{N(z)}{D(z)} = \sum_{k=0}^{n} \frac{\alpha_k}{z - t_k} \Bigg/ \sum_{k=0}^{n} \frac{\beta_k}{z - t_k}$$

Notation:

- $\{\alpha_k\}, \{\beta_k\}$ barycentric coefficients
- $\{t_k\}$ support points

Barycentric representations

Why use adaptive barycentric formulas?

$\rightarrow$ problem dependent $\{t_k\}$ $\Rightarrow$ well conditioned representation

Barycentric representations

Why use adaptive barycentric formulas?

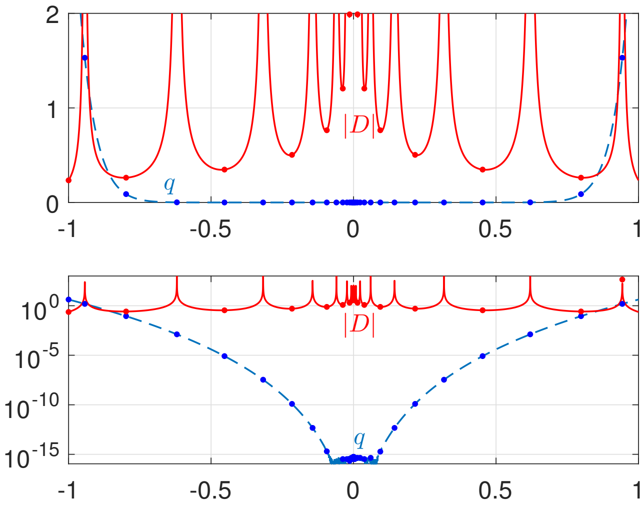
$\rightarrow$ problem dependent $\{t_k\} \Rightarrow$ well conditioned representation

Example:

$\rightarrow$ the adaptive Antoulas-Anderson (AAA) algorithm [Nakatsukasa, Sète & Trefethen 2018]: greedy least squares approximation

Example: $f(x) = |x|, x \in [-1, 1]$, type $(20, 20)$

p/q vs N/D



The rational Remez algorithm

The rational Remez algorithm

Some assumptions:

- no defect $(d = 0) \rightarrow$ required
- diagonal case $m = n \rightarrow$ for convenience

The rational Remez algorithm

Some assumptions:

- no defect ($d = 0$) $\rightarrow$ required
- diagonal case $m = n$ $\rightarrow$ for convenience

Step 1: choose a reference set $a \leqslant x_0 < \cdots < x_{2n+1} \leqslant b$

The rational Remez algorithm

Some assumptions:

- no defect ($d = 0$) $\rightarrow$ required
- diagonal case $m = n$ $\rightarrow$ for convenience

Step 1: choose a reference set $a \leqslant x_0 < \cdots < x_{2n+1} \leqslant b$

$\rightarrow$ iterate the following steps until convergence:

The rational Remez algorithm

Some assumptions:

- no defect ($d = 0$) $\rightarrow$ required
- diagonal case $m = n$ $\rightarrow$ for convenience

Step 1: choose a reference set $a \leqslant x_0 < \cdots < x_{2n+1} \leqslant b$

$\rightarrow$ iterate the following steps until convergence:

Step 2: find $r \in \mathcal{R}_{n,n}$ and $\lambda \in \mathbb{R}$ s.t.

$$f(x_k) - r(x_k) = (-1)^{k+1}\lambda, \qquad k = 0, \ldots, 2n+1$$

The rational Remez algorithm

Some assumptions:

- no defect ($d = 0$) $\rightarrow$ required
- diagonal case $m = n$ $\rightarrow$ for convenience

Step 1: choose a reference set $a \leqslant x_0 < \cdots < x_{2n+1} \leqslant b$

$\rightarrow$ iterate the following steps until convergence:

Step 2: find $r \in \mathcal{R}_{n,n}$ and $\lambda \in \mathbb{R}$ s.t.

$$f(x_k) - r(x_k) = (-1)^{k+1}\lambda, \qquad k = 0, \ldots, 2n+1$$

Step 3: among local extrema of $f - r$, take $2n + 2$ new points

$$a \leqslant x_0' < \cdots < x_{2n+1}' \leqslant b,$$

$f - r$ alternates in sign + at least one global extrema over $[a, b]$ and

$$|f(x_k') - r(x_k')| \geqslant |\lambda|, \qquad k = 0, \ldots, 2n+1$$

The rational Remez algorithm

Convergence:

→ usually quadratic [Curtis & Osborne 1966]
→ guaranteed only if starting reference set is close enough to optimal

The rational Remez algorithm

Convergence:

$\rightarrow$ usually quadratic [Curtis & Osborne 1966]
$\rightarrow$ guaranteed only if starting reference set is close enough to optimal

What can go wrong?

$\rightarrow$ no pole-free solution in **Step 2**

Step 1: initial reference set

→ **need** suff. good initial guess for $\{x_k\}$

Step 1: initial reference set

$\rightarrow$ **need** suff. good initial guess for $\{x_k\}$

Our strategy:

- use Carathéodory-Fejér (CF) approximation [Trefethen & Gutknecht 1983]

Step 1: initial reference set

$\rightarrow$ **need** suff. good initial guess for $\{x_k\}$

Our strategy:

- use Carathéodory-Fejér (CF) approximation [Trefethen & Gutknecht 1983]
- AAA-Lawson approx. (adaptively re-weighted least squares AAA variant)

Step 1: initial reference set

$\rightarrow$ **need** suff. good initial guess for $\{x_k\}$

Our strategy:

- use Carathéodory-Fejér (CF) approximation [Trefethen & Gutknecht 1983]
- AAA-Lawson approx. (adaptively re-weighted least squares AAA variant)
- extrapolation from lower degree approx. ($(2,2),(3,3),(4,4),\ldots$)

Step 2: find r

$\rightarrow$ find $r = N/D \in \mathcal{R}_{n,n}$ s.t.

$$N(x_k) = D(x_k)(f(x_k) - (-1)^{k+1}\lambda), \qquad k = 0, \dots, 2n+1$$

Step 2: find r

→ find $r = N/D \in \mathcal{R}_{n,n}$ s.t.

$$N(x_k) = D(x_k)(f(x_k) - (-1)^{k+1}\lambda), \qquad k = 0, \dots, 2n+1$$

→ matrix form

$$C\alpha = \left(\begin{bmatrix} f(x_0) & & & \\ & f(x_1) & & \\ & & \ddots & \\ & & & f(x_{2n+1}) \end{bmatrix} - \lambda \begin{bmatrix} -1 & & & \\ & 1 & & \\ & & -1 & \\ & & & \ddots \end{bmatrix} \right) C\beta,$$

$C \in \mathbb{R}^{(2n+2)\times(n+1)}$ Cauchy matrix, $C_{k,j} = 1/(x_k - t_j)$

Step 2: find r

$\to$ find $r = N/D \in \mathcal{R}_{n,n}$ s.t.

$$N(x_k) = D(x_k)(f(x_k) - (-1)^{k+1}\lambda), \qquad k = 0, \dots, 2n+1$$

$\to$ matrix form

$$C\alpha = \left(\begin{bmatrix} f(x_0) & & & \\ & f(x_1) & & \\ & & \ddots & \\ & & & f(x_{2n+1}) \end{bmatrix} - \lambda \begin{bmatrix} -1 & & & \\ & 1 & & \\ & & -1 & \\ & & & \ddots \end{bmatrix} \right) C\beta,$$

$C \in \mathbb{R}^{(2n+2)\times(n+1)}$ Cauchy matrix, $C_{k,j} = 1/(x_k - t_j)$

$\to$ generalized eigenvalue problem

$$\begin{bmatrix} C & -FC \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \lambda \begin{bmatrix} 0 & -SC \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$F = \mathrm{diag}(f(x_k)), S = \mathrm{diag}((-1)^{k+1})$

Step 2: find r

$\rightarrow$ generalized eigenvalue problem

$$\begin{bmatrix} C & -FC \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \lambda \begin{bmatrix} 0 & -SC \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Step 2: find r

$\rightarrow$ generalized eigenvalue problem

$$\begin{bmatrix} C & -FC \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \lambda \begin{bmatrix} 0 & -SC \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$\rightarrow$ can transform it into a symmetric eigenvalue problem

$$Q_1^T (SF) Q_1 R\beta = \lambda R\beta,$$

where $\omega_x(x) = \prod_{k=0}^{2n+1}(x - x_k), \quad \omega_t(x) = \prod_{j=0}^{n}(x - t_j),$

$$\Delta = \text{diag}\left(\frac{\omega_t(x_0)^2}{\omega_x'(x_0)}, \ldots, \frac{\omega_t(x_{2n+1})^2}{\omega_x'(x_{2n+1})}\right)$$

and $|\Delta|^{1/2} C = Q_1 R$

Step 2: find r

$\rightarrow$ generalized eigenvalue problem

$$\begin{bmatrix} C & -FC \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \lambda \begin{bmatrix} 0 & -SC \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$\rightarrow$ can transform it into a symmetric eigenvalue problem

$$Q_1^T (SF) Q_1 R\beta = \lambda R\beta,$$

where $\omega_x(x) = \prod_{k=0}^{2n+1}(x - x_k), \quad \omega_t(x) = \prod_{j=0}^{n}(x - t_j),$

$$\Delta = \text{diag}\left(\frac{\omega_t(x_0)^2}{\omega_x'(x_0)}, \ldots, \frac{\omega_t(x_{2n+1})^2}{\omega_x'(x_{2n+1})}\right)$$

and $|\Delta|^{1/2} C = Q_1 R$
$\rightarrow$ well conditioned eigenvalue computation

Step 2: choice of the $\{t_k\}$

Step 2: choice of the $\{t_k\}$

$\rightarrow$ because we perform the QR factorization of $|\Delta|^{1/2}C$, take $\{t_k\}$ to minimize

$$\min_{\Gamma} \kappa_2(|\Delta|^{1/2}C\Gamma),$$

Γ diagonal scaling matrix

Step 2: choice of the $\{t_k\}$

$\rightarrow$ because we perform the QR factorization of $|\Delta|^{1/2}C$, take $\{t_k\}$ to minimize

$$\min_{\Gamma} \kappa_2(|\Delta|^{1/2}C\Gamma),$$

Γ diagonal scaling matrix

$\rightarrow$ we show that this happens (with optimum 1) for

$$t_k = x_{2k+1}, \qquad k = 0, \ldots, n$$

Step 3: next reference set

Approach:

Step 3: next reference set

Approach:

$\rightarrow$ decompose $[a, b]$ into nondegenerate intervals

$$[a, x_0], [x_0, x_1], \ldots, [x_{2n}, x_{2n+1}], [x_{2n+1}, b]$$

Step 3: next reference set

Approach:

$\rightarrow$ decompose $[a, b]$ into nondegenerate intervals

$$[a, x_0], [x_0, x_1], \ldots, [x_{2n}, x_{2n+1}], [x_{2n+1}, b]$$

$\rightarrow$ detect singularities of f + further decomp. of $[a, b]$

$\rightarrow$ use `splitting on` [Pachón, Platte & Trefethen 2010]

Step 3: next reference set

Approach:

$\rightarrow$ decompose $[a, b]$ into nondegenerate intervals

$$[a, x_0], [x_0, x_1], \ldots, [x_{2n}, x_{2n+1}], [x_{2n+1}, b]$$

$\rightarrow$ detect singularities of f + further decomp. of $[a, b]$

$\rightarrow$ use `splitting on` [Pachón, Platte & Trefethen 2010]

$\rightarrow$ Chebyshev interpolants of $e(x) = f(x) - r(x)$ on each subinterval

$\rightarrow$ colleague matrix root finding [Specht, Good]

Examples

DEMO

Conclusion

$\rightarrow$ robust rational Remez algorithm (available now in Chebfun):

- adaptive barycentric representation $\rightarrow$ eigenvalue problem with good stability
- good choice of the initial reference points (CF, AAA-Lawson, extrapolation)
- colleague matrix root finding

Conclusion

$\rightarrow$ robust rational Remez algorithm (available now in Chebfun):

- adaptive barycentric representation $\rightarrow$ eigenvalue problem with good stability
- good choice of the initial reference points (CF, AAA-Lawson, extrapolation)
- colleague matrix root finding

$\rightarrow$ what I did not talk about:

- what to do in degenerate $d > 0$ cases
- how do we handle $m \neq n$ problem instances

Conclusion

→ robust rational Remez algorithm (available now in Chebfun):

- adaptive barycentric representation → eigenvalue problem with good stability
- good choice of the initial reference points (CF, AAA-Lawson, extrapolation)
- colleague matrix root finding

→ what I did not talk about:

- what to do in degenerate $d > 0$ cases
- how do we handle $m \neq n$ problem instances

→ the details:
B. Beckermann, S.-I. Filip, Y. Nakatsukasa, L. N. Trefethen, *Rational minimax approximation via adaptive barycentric representations*, arXiv:1705.10132, under minor revision for *SIAM Journal on Scientific Computing*

Conclusion

$\rightarrow$ directions for future work:

- handle weighted approximation problems (*e.g.*, relative error)
- extensions to:
  - the complex case
  - bivariate approximation

Conclusion

→ directions for future work:

- handle weighted approximation problems (*e.g.*, relative error)
- extensions to:
  - the complex case
  - bivariate approximation

Thank you!