Algorithms for the ADI optimum parameters and Zolotarev's coefficients - numerical experiments

Beata Laszkiewicz, Krystyna Ziętak

Wrocław University of Technology
Institute of Mathematics and Computer Science

Birmingham, September 13–15, 2007

Zolotarev and ADI minimax problems

- Complete elliptic integral and Jacobi elliptic functions
- Algorithms for ADI coefficients

$$r_{j,m}, \qquad j = 1, \ldots, m$$

$$m \quad \text{arbitrary or} \quad m = 2^p$$

- Numerical experiments

Why such a talk?

- The paper of Kennedy on fast evaluation of Zolotarev coefficients is cited in: **N.J. Higham, Functions of matrices: theory and computation** (2008, book to appear).
- Computing Zolotarev's coeeficients leads to the same algorithms as computing the ADI coefficients.

In this way I return to my PhD (1972).

Why such a talk?

- The paper of Kennedy on fast evaluation of Zolotarev coefficients is cited in: **N.J. Higham, Functions of matrices: theory and computation** (2008, book to appear).
- Computing Zolotarev's coeeficients leads to the same algorithms as computing the ADI coefficients.

In this way I return to my PhD (1972).

ADI minimax problem

$$\text{Let } k' \in \mathbb{R}, \qquad 0 < k' < 1.$$

Find parameters $r_{j,m}$ which minimize

$$\max_{k' \le x \le 1} |f_m(x)|,$$

where

$$f_m(x) = \prod_{j=1}^{m} \frac{x - r_{j,m}}{x + r_{j,m}}.$$

Zolotarev problem

The best uniform rational relative approximation to $1/\sqrt{x}$

Let $a, b \in \mathbb{R}$, $0 < a < b$, $c = \frac{b}{a}$.

Find rational function $g \in \mathcal{R}_{m-1,m}$ which minimizes

$$\max_{x \in [1, c^2]} |1 - \sqrt{x} g(x)|$$

Uniform rational approximation to sign

$$\max_{x\in[-b,-a]\cup[a,b]} |\mathbf{sign}(x) - s(x)|.$$

over $s(x) \in \mathbb{R}_{2m-1,2m}$. Then

$$s(x) = tg\left(t^2\right), \qquad t = x/a$$

where $\mathbf{g} \in \mathcal{R}_{m-1,m}$ **solves Zolotarev problem** $(c = \frac{b}{a})$.

Zolotarev and ADI coefficients are known analytically

Jacobi elliptic functions

$$\mathrm{sn}(\xi K(k); k), \qquad \mathrm{dn}(\xi K(k); k)$$

Complete elliptic integral

$$K(k) = \int_0^1 \frac{dt}{\sqrt{(1 - t^2)(1 - k^2 t^2)}}$$

Jacobi elliptic functions: sn, dn

- $\mathrm{sn}(u;k)$ is defined implicitly by the following relation:

$$u = \int_0^{\mathrm{sn}} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}$$

- $$\mathrm{dn}^2(u;k) + k^2\mathrm{sn}^2(u;k) = 1$$

Jacobi elliptic functions: sn, dn

- $\mathrm{sn}(u;k)$ is defined implicitly by the following relation:

$$u = \int_0^{\mathrm{sn}} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}$$

- $$\mathrm{dn}^2(u;k) + k^2\mathrm{sn}^2(u;k) = 1$$

ADI coefficients $r_{j,m}$

ordered coefficients

$$r_{j,m} = \mathrm{dn}\Big(\frac{2m-2j+1}{2m}K(k);k\Big), \qquad j = 1,2,\dots,m$$

ADI - alternating set

$$u_{j,m} = \mathrm{dn}\Big(\frac{m-j}{m}K(k);k\Big), \qquad j = 0,1,\dots,m$$

where $k = \sqrt{1-k'^2}$

Zolotarev coefficients $c_{j,m}$

$$c_{j,m} = \frac{s^2}{1 - s^2}$$

$$s = \mathrm{sn}\left(\frac{j}{2m}K(k); k\right)$$

where $k = (1 - 1/c^2)^{1/2}, j = 1, \ldots, 2m - 1$

Zolotarev - alternating set

$$v_{j,m} = \left[\mathrm{dn}\left(\frac{j}{2m}K(k); k\right)\right]^{-2}, \quad j = 0, 1, \ldots, 2m$$

Zolotarev and ADI coefficients

$$\mathrm{sn}(\xi K(k); k), \qquad \mathrm{dn}(\xi K(k); k)$$

$$0 \le \xi \le 1, \qquad 0 \le k \le 1$$

$$\textbf{ADI}: \quad [k', 1], \quad k = (1 - k'^2)^{1/2}$$

$$\textbf{Zolotarev}: \quad [1, c^2], \quad k = (1 - k'^2)^{1/2}, \quad k' = 1/c$$

longer intervals $\rightarrow$ k closer to 1,

i.e. k' close to zero

Properties of $\mathrm{dn}(u;k)$ and $K(k)$

$$k' = \sqrt{1-k^2} \le \mathrm{dn}(\xi K(k); k) \le 1$$

$$0 < \xi < 1$$

If $k \to 1$ then

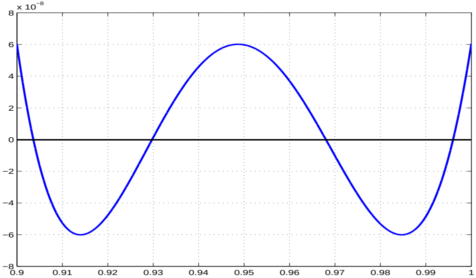
$$K(k) \to \infty, \qquad \mathrm{dn}(\xi K(k); k) \to 0$$

If $k \to 0$ then

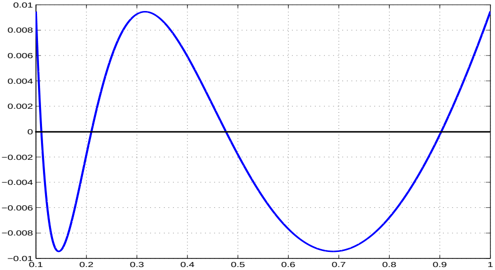
$$K(k) \to \frac{\pi}{2}, \qquad \mathrm{dn}(\xi K(k); k) \to 1$$

ADI rational function

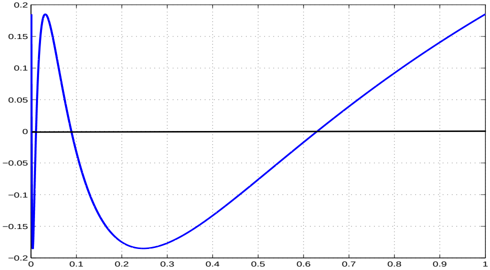
$$[k', 1]$$



$$m = 4, \quad k' = 0.9$$


0.01
0.008
0.006
0.004
0.002
0
−0.002
−0.004
−0.006
−0.008
−0.01
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

$$m = 4, \quad k' = 0.1$$



$$m = 4, \quad k' = 10^{-3}$$

- MATLAB - ELLIPKE
- MAPLE - 20 digits

Arithmetic-geometric mean **AGM**

$$a_0 = 1, \qquad b_0 = \sqrt{1-k^2}$$

$$a_{n+1} = \frac{1}{2}(a_n + b_n), \qquad b_{n+1} = \sqrt{a_n b_n}$$

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \frac{\pi}{2K(k)}$$

Conditioning of $K(k)$

$$cond = \left| \frac{k}{K(k)} \times \frac{d}{dk} K(k) \right| = \left| \frac{E(k)}{k'^2 K(k)} - 1 \right|$$

- if $k \to 0$ then cond $\to 0$
- if $k \to 1$ then cond $\to \infty$

$$K(1) = \infty, \qquad K(0) = \frac{\pi}{2}$$

Computing $K(k)$

$K(k)$ for k' given

$$[k', 1], \qquad k = (1 - k'^2)^{1/2}$$

ELLIPKE (q) - (MATLAB)

$$q = k^2 = (1 - k')(1 + k')$$

$$b_0 = (1 - q)^{1/2}$$

AGM (k) - (Kennedy version)

$$k = [(1 - k')(1 + k')]^{1/2}$$

$$b_0 = [(1 - k)(1 + k)]^{1/2}$$

$K(k)$ for k' given

k'	AGM	ELLIPKE
0.9	1.65461 66675 2253	1.65461 66675 2253
0.01	5.99158 93405 0711	5.99158 93405 0705
0.001	8.29405 14635 3450	8.29405 14635 4555
0.000001	15.20176 04707 7238	15.20176 04707 7226

$K(k)$ for k' given

ELLIPKE: $q = (1 - k')(1 + k')$

MAPLE: $k = [(1 - k')(1 + k')]^{1/2}$

k'	MAPLE	ELLIPKE
0.99	1.57869 97420 3901	1.57869 97420 3901
0.9	1.65461 66675 2252	1.65461 66675 2253
0.1	3.69563 73629 8987	3.69563 73629 8987
10^{-2}	5.99158 93405 0699	5.99158 93405 0705
10^{-4}	10.59663 47570 8766	10.59663 47545 7528
10^{-6}	15.20180 49190 8781	15.20176 04707 7226
10^{-8}	19.80697 51050 7225	19.40812 10556 7847
10^{-9}	22.10956 01980 6630	∞

MATLAB - ELLIPJ (see Abramowitz, Stegun)

$$a_0 = 1, \quad b_0 = \sqrt{1-k^2},$$

$$a_{n+1} = \frac{1}{2}(a_n + b_n), \quad b_{n+1} = \sqrt{a_n b_n},$$

$$c_{n+1} = \frac{1}{2}(a_n - b_n)$$

$$\varphi_N = 2^{N-1} a_N u$$

$$\sin(2\varphi_{n-1} - \varphi_n) = \frac{c_n}{a_n} \sin\varphi_n, \qquad n = N, N-1, \ldots, 1$$

$$\mathrm{sn}(u;k) = \sin\varphi_0$$

- **Kennedy (2005)**

$$\mathrm{dn}(\xi K(k); k)$$

- **KenI** - Kennedy
  without computing explicitly **K(k)**
- **KenII** - Kennedy
  with **K(k)** by **MAPLE**

Gauss transformation

$$\mathrm{sn}\left((k+1)u; \frac{2\sqrt{k}}{1+k}\right) = \frac{(k+1)\mathrm{sn}(u;k)}{1+k\,\mathrm{sn}^2(u;k)}$$

Find $\hat{k}, \hat{u}$ that
$\mathrm{sn}(u;k)$ is Gauss transformation of $\mathrm{sn}(\hat{u};\hat{k})$.

$$\hat{k} = \frac{(1-\sqrt{1-k^2})^2}{k^2} < k,$$

$$\hat{u} = \frac{u}{\hat{k}+1}$$

We repeat: $s_n = \mathrm{sn}(u_n; k_n)$,

$$k_{n+1} = \frac{k_n^2}{(1+\sqrt{1-k_n^2})^2}, \qquad u_{n+1} = \frac{u_n}{k_{n+1}+1}.$$

Then $k_n \longrightarrow 0$, $\qquad u_n \longrightarrow u_\infty \frac{u\pi}{2K(k)}$,

$$s_n \longrightarrow \mathrm{sn}(u_\infty; 0) = \sin(u_\infty)$$

$$\mathrm{sn}(u; 0) = \sin(u)$$

Kennedy 2005

AGM: $a_0 = 1, \quad b_0 = \sqrt{1-k^2}$

$$a_{n+1} = \frac{1}{2}(a_n + b_n), \quad b_{n+1} = \sqrt{a_n b_n}$$

$n = 0, 1, \ldots$ until $b_{n+1} \leq b_n$

Gauss transformation - back process

$$s_N = \sin\left(a_N u\right)$$

$$s_n = \frac{2a_n s_{n+1}}{(a_n + b_n) + (a_n - b_n)s_{n+1}^2}$$

$n = N-1, \ldots, 1,$

$$\mathrm{sn}(u; k) \approx s_1$$

Kennedy - back process for $u = \xi K(k)$

$$s_N = \sin(a_N u), \quad a_N \approx a_\infty = \frac{\pi}{2K(k)}$$

$$u = \xi K(k)$$

$$s_N \approx \sin(a_\infty \xi K(k)) = \sin(\xi \frac{\pi}{2})$$

KenI: $s_N = \sin(\xi \frac{\pi}{2})$

KenII: $s_N = \sin(\xi a_N K(k))$,
where $K(k)$ by **MAPLE**

Algorithms for ADI coefficients

$r_{j,m} = \mathrm{dn}(\xi_j K(k); k)$, **k′ given**

m arbitrary

- MATLAB

  ELLIPJ: $\mathbf{dn}(u; k)$
  ELLIPKE: $K(k)$

- Kennedy and AGM: $\mathbf{sn}(u; k)$

$r_{j,m}$ for $m = 2^p$

- WR - Wachspress (1963)
- RJ - Ziętak (PhD 1972)

$$\mathrm{sn}^2(u;k) + \mathrm{cn}^2(u;k) = 1$$

$$\mathrm{dn}^2(u;k) + k^2\mathrm{sn}^2(u;k) = 1$$

$$\mathrm{dn}^2(u;k) = \frac{\mathrm{dn}(2u;k) + k^2\mathrm{cn}(2u;k) + k'^2}{1 + \mathrm{dn}(2u;k)}$$

RJ algorithm $m = 2^p$

$$r_{1,1} = \mathrm{dn}(\frac{1}{2}K(k); k) = \sqrt{k'}$$

$$r_{1,1} \Longrightarrow r_{1,2}, r_{2,2} \Longrightarrow r_{1,4}, r_{2,4}, r_{3,4}, r_{4,4} \Longrightarrow \ldots$$

$$r_{j,n} \quad \text{for} \quad n = 1, 2, 4, 8, 16, \ldots, 2^p$$

Let $s = r_{j,n}$. Then

$$r_{n+j,2n} = \left(\frac{s + k'^2 + \sqrt{(1 - k'^2)(s^2 - k'^2)}}{1 + s} \right)^{1/2}$$

$$r_{j,m} = \mathrm{dn}\Big(\frac{2m-2j+1}{2m}K(k); k\Big)$$

$$j = 1, 2, \ldots, m$$

$$k' \le r_{1,m} < r_{2,m} < \cdots < r_{m,m} \le 1$$

$$r_{j,2n} = \frac{k'}{r_{2n-j+1,2n}} \qquad j = 1, \ldots, n$$

version a: all $r_{j,m}$

version b: half of $r_{j,m}$

Wachspress algorithm $m = 2^p$

Let $a_0 = k', \quad b_0 = 1$

$$a_{i+1} = \sqrt{a_i b_i}, \quad b_{i+1} = \frac{a_i + b_i}{2}$$

$s_{1,1} = a_{p+1} = \sqrt{a_p b_p} \Longrightarrow r_{jm} = s_{jm}$

$n = 2^{p-i}$

$$s = s_{j,n} \quad \Longrightarrow \quad s_{j+n,2n} = s + \sqrt{s^2 - a_i^2}$$

$$s_{n-j,2n} = \frac{a_{i-1} b_{i-1}}{s_{n+j,2n}}$$

Comparison: RJ and WR

RJ: $r_{j,2n}$ **and** $r_{j+n,2n}$ **are roots of trinomial**

$$(1+r)y^2 - 2(k'^2+r)y + k'^2(1+r) = 0, \qquad (r = r_{jn}).$$

WR: Let $n = 2^{p-i}$**.**

Then $s_{n+j,2n}$ and $s_{n-j+1,2n}$ are roots of trinomial

$$y^2 - 2sy + a_i^2 = 0, \qquad (s = s_{jn}).$$

For $i = p$ we have: $\Delta = 4(a_{p+1}^2 - a_p^2) \to 0 \quad \text{when } p \to \infty$

Rounding errror analysis of **WR** and **RJ** (Ziętak 1974)
In **JR** damping of errors can occur ($r_{jn} > \sqrt{k'}$).
In **WR** propagation factors of errors can be large.

Numerical experiments for ADI rational function f_m $\quad (m = 2^p)$

Theoretically at **alternating points** $u_{j,m}$ we have:

$$\mathbf{f_m(u_{j,m}) = \pm||f_m||_\infty}$$

In floating point arithmetic we compute:

$$\max = \max_j |f_m(u_{j,m})|, \qquad \min = \min_j |f_m(u_{j,m})|$$

where $u_{j,m} = \mathrm{dn}\left(\frac{m-j}{m}K(k); k\right)$.

For all algorithms we compare

$$\frac{\max - \min}{\max}$$

Results for $\mathbf{k}' = \mathbf{10^{-6}}$ and $\mathbf{m} = \mathbf{4}$

alg	max	$max - min$	$\frac{max-min}{max}$
$Kenl_a$	$5.399\ 1574\ 881e-01$	$6.3e-07$	$1.8e-06$
$Kenl_b$	$5.399\ 2282\ 663e-01$	$7.5e-06$	$1.4e-05$
$KenII_a$	$5.399\ 1728\ 154e-01$	$1.5e-05$	$2.8e-05$
$\mathbf{KenII_b}$	$5.399\ 1728\ 155e-01$	$1.4e-14$	$\mathbf{2.7e-14}$
$matlab_a$	$5.399\ 1574\ 882e-01$	$6.4e-07$	$1.2e-06$
$matlab_b$	$5.399\ 2282\ 663e-01$	$7.5e-06$	$1.4e-05$
WR	$5.399\ 1728\ 155e-01$	$2.2e-16$	$\mathbf{4.1e-16}$
RJ	$5.399\ 1728\ 155e-01$	$2.2e-16$	$\mathbf{4.1e-16}$

Results for $\mathbf{k}' = \mathbf{10^{-6}}$ and $\mathbf{m = 4}$

alg	max	$max - min$	$\frac{max - min}{max}$
$Kenl_a$	$5.399\,1574\,881e-01$	$6.3e-07$	$1.8e-06$
	$5.399\,3143\,136e-01$	$1.6e-05$	$3.0e-05$
$matlab_a$	$5.399\,1574\,882e-01$	$6.4e-07$	$1.2e-06$
	$5.399\,3143\,136e-01$	$1.6e-05$	$3.0e-05$
RJ	$5.399\,1728\,155e-01$	$2.2e-16$	$\mathbf{4.1e-16}$

Results for $\mathbf{k}' = \mathbf{10}^{-3}$ and $\mathbf{m} = \mathbf{4}$

alg	max	$max - min$	$\frac{max-min}{max}$
$KenI_a$	$1.850880205\ 709\ 998e-01$	$2.4e-11$	$1.3e-10$
$KenI_b$	$1.850880205\ 771\ 299e-01$	$1.3e-11$	$7.2e-11$
$KenII_a$	$1.850880205\ 703\ 919e-01$	$4.3e-11$	$2.3e-10$
$\mathbf{KenII_b}$	$1.850880205\ 677\ 388e-01$	$5.8e-15$	$\mathbf{3.1e-14}$
$matlab_a$	$1.850880205\ 717\ 716e-01$	$1.1e-11$	$6.1e-11$
$matlab_b$	$1.850880205\ 758\ 433e-01$	$1.1e-11$	$6.2e-11$
$\mathbf{WR}$	$1.850880205\ 677\ 349e-01$	$8.3e-17$	$\mathbf{4.5e-16}$
$\mathbf{RJ}$	$1.850880205\ 677\ 348e-01$	$8.3e-17$	$\mathbf{4.5e-16}$

Results for $\mathbf{k}' = 10^{-1}$ and $\mathbf{m} = 4$

alg	*max*	*max − min*	$\frac{max-min}{max}$
$KenI_a$	9.978243 2409 184e − 10	1.2e − 21	1.2e − 12
$KenI_b$	9.978243 2409 097e − 10	3.4e − 23	3.4e − 14
$KenII_a$	9.978243 2409 186e − 10	1.2e − 21	1.2e − 12
$\mathbf{KenII_b}$	9.978243 2409 097e − 10	8.7e − 24	**8.7e − 15**
$matlab_a$	9.978243 2409 121e − 10	4.3e − 22	4.3e − 13
$matlab_b$	9.978243 2409 099e − 10	2.8e − 23	2.8e − 14
WR	9.978243 3898 946e − 10	3.0e − 17	**3.0e − 08**
RJ	9.978243 2409 097e − 10	2.3e − 23	2.3e − 14

Results for $\mathbf{k'} = \mathbf{0.99}$ and $\mathbf{m} = \mathbf{4}$

alg	*max*	*max − min*	$\frac{max-min}{max}$
$Kenl_a$	$4.9818\ 2756808\ 1315e-12$	$1.0e-24$	$2.0e-13$
$Kenl_b$	$4.9818\ 2756808\ 1770e-12$	$1.5e-24$	$3.0e-13$
$KenII_a$	$4.9818\ 2756808\ 0871e-12$	$1.0e-24$	$2.0e-13$
$KenII_b$	$4.9818\ 2756808\ 0871e-12$	$1.0e-24$	$2.0e-13$
$matlab_a$	$4.9818\ 2756808\ 1315e-12$	$1.0e-24$	$2.0e-13$
$matlab_b$	$4.9818\ 2756808\ 1770e-12$	$1.5e-24$	$3.0e-13$
WR	$4.9818\ 4655115\ 7935e-12$	$3.8e-17$	$\mathbf{7.6e-06}$
RJ	$4.9818\ 2756808\ 1870e-12$	$2.1e-24$	$4.2e-13$

$$k' = 0.001, \quad m = 16, \quad max \approx 1.46 \times 10^{-4}$$

$$err = \frac{max - min}{max}$$

1. **MATLAB**: $err = 2.5e - 09$
2. **Kennedy$_{\mathbf{a}}$**: $err = 1.7e - 09$
3. **WR** : $err = 1.4e - 14$
4. **RJ** : $err = 3.3e - 15$

$$k' = 0.1, \quad m = 16, \quad max \approx 9.97 \times 10^{-10}$$

$$err = \frac{max - min}{max}$$

1. **MATLAB**: $err = 4.3e - 13$
2. **Kennedy$_a$**: $err = 1.2e - 12$
3. **WR** : $err = 3.0e - 08$
4. **RJ** : $err = 2.3e - 14$

$$k' = 0.99, \quad m = 16, \quad max \approx 7.69 \times 10^{-47}$$

$$err = \frac{max - min}{max}$$

1. **MATLAB**: $err = 2.6e - 12$
2. **Kennedy$_{\mathbf{a}}$**: $err = 3.0e - 12$
3. **WR** : $err = 1$!!!
4. **RJ** : $err = 3.4e - 12$

Errors in Kennedy, ELLIPJ...

- rounding errors in whole process
- approximation error of Jacobi elliptic functions
- error in argument $u = \xi K(k)$

Conclusions

KenII with $s_N = \sin\left(\xi a_N K(k)\right)$

is better than **KenI** with $s_N = \sin\left(\xi \frac{\pi}{2}\right)$

where $K(k)$ by **MAPLE**

(**not** by MATLAB).

The accuracy of ELLIPJ and ELLIPKE is not good for small k'.

Conclusions

version b with computing $r_{j,m} = \dfrac{k'}{r_{m-j+1,m}}$ is better than **version a**

The accuracy of **RJ** is very good in all cases.

The accuracy of **WR** is not enough for k' close to 1 (small interval) or large m.

References

1. **Kennedy**, Fast evaluation of Zolotarev coefficients, in: *Lecture Notes in Computational Science and Engineering*, vol. 47, Springer 2005.
2. **Wachspress**, Extended applications of alternating direction model problem theory, *J. SIAM 11 (1963)*.
3. **Ziętak**, Construction and features of the optimum rational function used in the ADI method, *Applicationes Mathematicae 14 (1974)*.
4. **Ziętak**, Investigation of the algorithms determining the optimum rational function of the ADI-method, *Applicationes Mathematicae 14 (1974)*.

Wrocław University of Technology

Institute of Mathematics and Computer Science