Algorithms for matrix sector function

Krystyna Ziętak

(the joint work with Beata Laszkiewicz)

Institute of Mathematics and Computer Science
Wrocław University of Technology

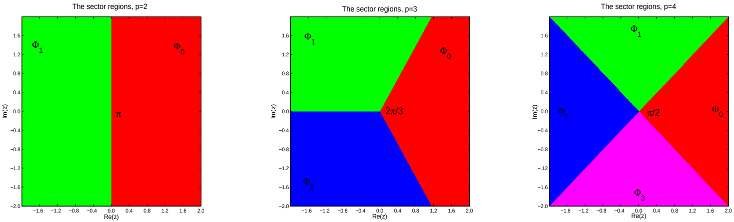
Novi Sad, April 28 - 30, 2008

Outline

The sector regions

$$\Phi_k = \left\{ z \in \mathbb{C} : \frac{2k\pi}{p} - \frac{\pi}{p} < \arg(z) < \frac{2k\pi}{p} + \frac{\pi}{p} \right\}$$

$k = 0, \ldots, p-1$



The scalar p-sector function

- $s_p(\lambda)$ is the pth root of unity which lies in the same sector Φ_k in which λ is.

Representation

$$s_p(\lambda) = \frac{\lambda}{\sqrt[p]{\lambda^p}}$$

- $\sqrt[p]{a}$ principal pth root of $a \notin \mathbb{R}^-$, $\sqrt[p]{a}$ lies in Φ_0
- $s_p(\lambda)$ is not defined for the pth roots of nonpositive real numbers.

Principal matrix pth root

Let nonsingular complex matrix A have no negative eigenvalue. There is a unique pth root of A:

$$X = A^{1/p}$$

all of whose eigenvalues lie in the region Φ_0.

$$X^p = A, \qquad \arg \lambda_j(X) \in \left(-\frac{\pi}{p}, \frac{\pi}{p}\right)$$

N.J. Higham, *Functions of Matrices: Theory and Computation*, SIAM 2008

The scalar p-sector function

Let $\lambda = |\lambda| e^{i\varphi} \in \mathbb{C}, \quad \lambda \neq 0,$

$$\varphi = \arg(\lambda) \neq \frac{2k\pi}{p} + \frac{\pi}{p}, \quad k \in \{0, 1, \ldots, p-1\}$$

Then

$$s_p(\lambda) = e^{i2\pi q/p}$$

where $q \in \{0, 1, \ldots, p-1\}$ such that

$$\frac{2q\pi}{p} - \frac{\pi}{p} < \varphi < \frac{2q\pi}{p} + \frac{\pi}{p}$$

Matrix sector function of $A \in \mathbb{C}^{n\times n}$

$$\mathrm{sect}_p(A) = A\left(\sqrt[p]{A^p}\right)^{-1}$$

$$\lambda_j(A) \neq 0, \quad arg(\lambda_j) \neq 2\pi(q+\tfrac{1}{2})/p$$

$$q \in \{0, \ldots, p-1\}$$

Matrix sector function is some *pth* root of identity I.

Matrix sector function

$$\mathrm{sect}_p(A) = Z \mathrm{diag}\left(s_p(\lambda_j) I_{r_j}\right) Z^{-1}$$

$$A = Z \mathrm{diag}\left(J_1, J_2, \ldots, J_m\right) Z^{-1},$$

Jordan canonical form

Jordan block $J_k(\lambda)$

Matrix sign function

Let

$$A = Z\mathrm{diag}(J_1, J_2)Z^{-1},$$

eigenvalues of J_1 lie in the open left half-plane, those of J_2 in open right half-plane. Then

$$\mathrm{sign}(A) = Z\mathrm{diag}(-I_1, I_2)Z^{-1}$$

Algorithms for matrix sign function: Schur method, Newton's method, Pade family of iterations,...

Fréchet derivative and condition numbers of matrix function

Let $F = F(X)$ be a matrix function. The Fréchet derivative of F at X in the direction E is a linear mapping such that

$$F(X+E) - F(X) = L(X,E) = o(||E||).$$

Absolute and relative condition numbers of $F(X)$

$$\mathrm{cond}_{\mathrm{abs}}(F,X) = \lim_{\varepsilon\to 0} \sup_{||E||\le\varepsilon} \frac{||F(X+E)-F(X)||}{\varepsilon} = ||L(X)||$$

$$\mathrm{cond}_{\mathrm{rel}}(F,X) = \frac{||L(X)||\ ||X||}{||F(X)||}$$

Fréchet derivative of matrix sign function

Matrix sign decomposition - Higham

$$A = SN, \quad S = \mathrm{sign}(A), \quad N = (A^2)^{1/2}$$
$$S^2 = I, \qquad S^{-1} = S$$

$$S + \Delta_S = \mathrm{sign}(A + \Delta_A)$$

$L = L(A, \Delta_A)$ Fréchet derivative of matrix sign function of A in direction Δ_A

$$\Delta_S - L = o(||\Delta_A||)$$

Kenney-Laub

L satisfies $\quad NL + LN = \Delta_A - S\Delta_A S.$

Fréchet derivative of matrix sector function

$$\mathrm{sect}_p(A) + \Delta_S = \mathrm{sect}_p(A + \Delta_A)$$

Matrix sector decomposition $A = SN$,

$S = \mathrm{sect}_p(A), \quad N = (A^p)^{1/p}, \qquad S^{-1} = S^{p-1}$

The Fréchet derivative $L = L(A, \Delta_A)$ of matrix sector function is the unique solution of

$$NL + \sum_{k=0}^{p-2} S^k L S^{-k} N = \Delta_A - S^{-1}\Delta_A S$$

Real Schur algorithm for $f(A)$

$A \in \mathbb{R}^{n\times n}$, $A = QRQ^T$ real Schur decomposition

R upper quasi-triangular and block, Q orthogonal

Matrix function f of R has the same block structure as R

Parlett 1976

Main blocks of R are 1×1 or 2×2.

Recurrence relations between blocks of R and $f(R)$ lead to real Schur algorithm for f.

Schur algorithm for matrix *p*th root

Schur algorithms

Higham 1987 - square root
Smith 2003 - *p*th root

Stability of Schur algorithm - Smith 2003

Let $A = QRQ^T$ be real Schur decomposition, $U = (R)^{1/p}$.

$$\beta(U) = \frac{||U||_F^p}{||R||_F} \geq 1$$

Schur algorithm for *p*th root stable provided $\beta(U)$ is sufficiently small.

Algorithms for matrix sector function

$$\mathrm{sect}_p(A) = A(A^p)^{-1/p}$$

$$\mathrm{sect}_p(A) = A\exp(-\log(A^p)/p)$$

MATLAB: `expm`, `logm`

- real Schur algorithm
- Newton's iterations
- Halley's method

Real Schur algorithm for sector

$$A = QRQ^T \qquad \text{real Schur decomposition}$$

$$U = \mathrm{sect}_p(R), \qquad \mathrm{sect}_p(A) = QUQ^T.$$

$$RU = UR, \qquad U^p = I$$

Recurrence relations between blocks of R and U and some Sylvester equations for the blocks lead to real Schur algorithm for sector.

Remark. If A has multiple complex eigenvalues in the sectors different from $\Phi_{p/2}$ (if p even) and Φ_0 then real Schur algorithm does not work.

Newton's method for sector

Shieh, Tsay, Wang, 1984

$$X_0 = A$$

$$X_{k+1} = \left((p-1)X_k^p + I\right) pX_k^{1-p}$$

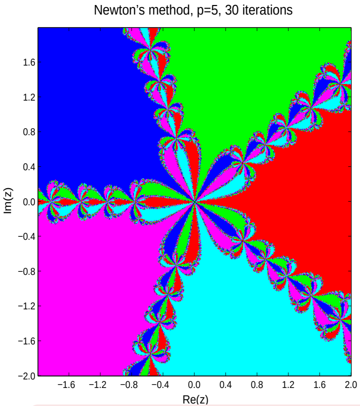
Newton's method is applied to the scalar equation

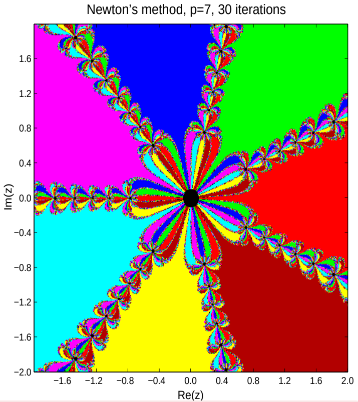
$$x^p - 1 = 0; \quad x_0 = \lambda_j(A)$$

Convergence regions for matrix sector function follow from the results of Higham and Iannazzo for matrix pth roots.

Regions of convergence of Newton for sector

determined experimentally





ω_j pth root of unity

color: $|x_{30} - \omega_j| < 10^{-5}$

Convergence of Newton for sector

If all eigenvalues of A lie in

$$\bigcup_{k=0}^{p-1}(\mathbb{B}_k \cup \mathbb{C}_k \cup \mathbb{R}_k^+)$$

$$\mathbb{B}_k = \left\{ z \in \mathbb{C} : \ |z| \geq 1, \ \frac{2k\pi}{p} - \frac{\pi}{2p} < \arg(z) < \frac{2k\pi}{p} + \frac{\pi}{2p} \right\}$$

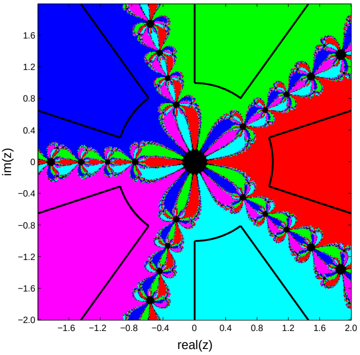
$$\mathbb{C}_k = \left\{ z \in \mathbb{C} : \frac{1}{2^{1/p}} \leq |z| \leq 1, \frac{2k\pi}{p} - \frac{\pi}{2p} < \arg(z) < \frac{2k\pi}{p} + \frac{\pi}{2p} \right\}$$

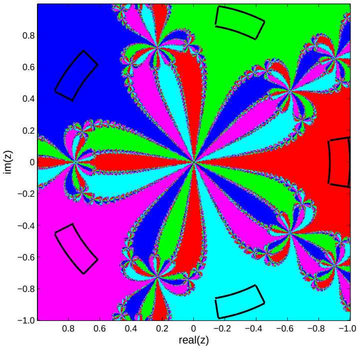
$$\mathbb{R}_k^+ = \left\{ z : \mathbb{C} : \ \mathrm{Re}\ z > 0 \text{ and } \frac{2k\pi}{p} - \frac{\pi}{2p} < \arg(z) < \frac{2k\pi}{p} + \frac{\pi}{2p} \right\}$$

then Newton is convergent

Convergence regions of Newton

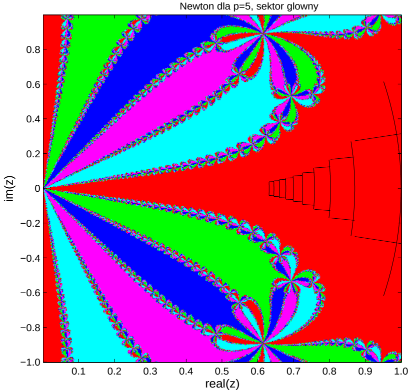
Region $\mathbb{B}_k$ and Region $\mathbb{C}_k$

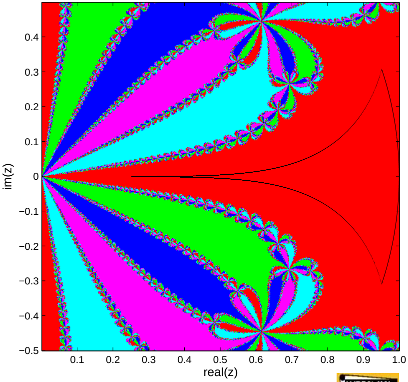




Convergence regions of Newton

Additional regions





Halley's method for sector

Bakkaloğlu, Koç, 1995

$$X_0 = A$$

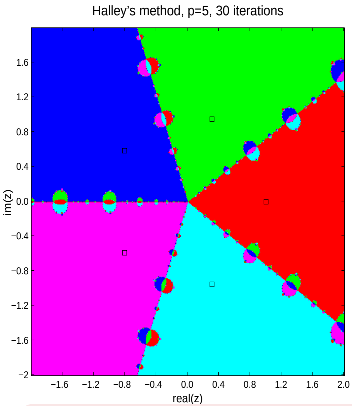
$$X_{k+1} = X_k \left[(p-1)X_k^p + (p+1)I\right] \times \left[(p+1)X_k^p - (p-1)I\right]^{-1}$$

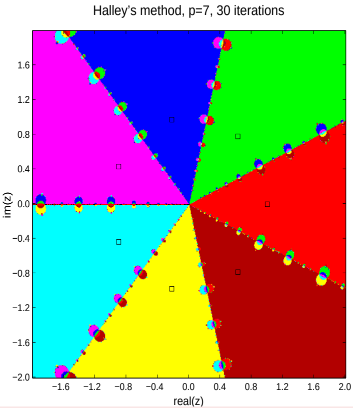
Halley's method is applied to the scalar equation

$$x^p - 1 = 0; \quad x_0 = \lambda_j(A)$$

Regions of convergence of Halley for sector

determined experimentally





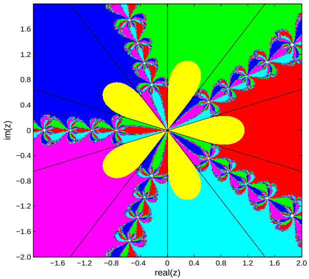
ω_j pth root of unity

color: $|z_{30} - \omega_j| < 10^{-5}$

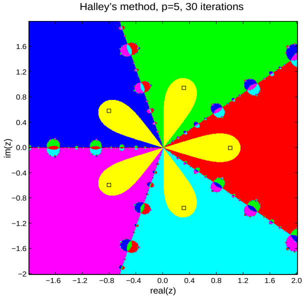
Halley for sector

"yellow flowers" - conjecture for Pade

Newton and $\mathbb{B}_p^{\mathrm{hall}}$



Halley and Pade



If all eigenvalues of A lie in

$$\mathbb{B}_p^{\mathrm{hall}} = \bigcup_{k=0}^{p-1} \left\{ z \in \mathbb{C} : \frac{2k\pi}{p} - \frac{\pi}{2p} < \arg(z) < \frac{2k\pi}{p} + \frac{\pi}{2p} \right\}$$

then Halley is convergent to sector.

Stability of Newton's and Halley's methods for matrix sector function

- Matrix sector function is idempotent, i.e. $\mathrm{sect}_p(\mathrm{sect}_p(A)) = \mathrm{sect}_p(A)$.
- From the theorem of Higham we deduce that Newton's and Halley's iterations are stable, i.e. Fréchet derivatives of the functions, generating iterations, have bounded powers.

Stability of Newton's and Halley's methods for matrix sector function

- Matrix sector function is idempotent, i.e. $\mathrm{sect}_p(\mathrm{sect}_p(A)) = \mathrm{sect}_p(A)$.
- From the theorem of Higham we deduce that Newton's and Halley's iterations are stable, i.e. Fréchet derivatives of the functions, generating iterations, have bounded powers.

Fréchet derivative

Let $A \in \mathbb{C}^{n\times n}$ be such that $\mathrm{sect}_p(A)$ exists and the Newton iterates X_k are convergent to $\mathrm{sect}_p(A)$. Let

$$Y_{k+1} = \frac{1}{p}\left((p-1)Y_k - X_k^{1-p}\left(\sum_{j=0}^{p-2} X_k^{p-2-j} Y_k X_k^j\right) X_k^{1-p}\right),$$

$$Y_0 = \Delta_A, \qquad X_0 = A.$$

Then the sequence Y_k tends to the Fréchet derivative $L(A, \Delta_A)$ of $\mathrm{sect}_p(A)$: $\lim_{k\to\infty} Y_k = L(A, \Delta_A)$.

Matrix sign ($p = 2$) Kenney-Laub

$$Y_{k+1} = \tfrac{1}{2}(Y_k - X_k^{-1} Y_k X_k^{-1})$$

Implementation

Newton

$$X_{k+1} = \left[(p-1)X_k^p + I\right]\left(X_k^{-1}\right)^{p-1}$$

Halley

$$X_{k+1} = X_k\left[(p-1)X_k^p + (p+1)I\right] \times \left[(p+1)X_k^p - (p-1)I\right]^{-1}$$

Example 1- test matrix

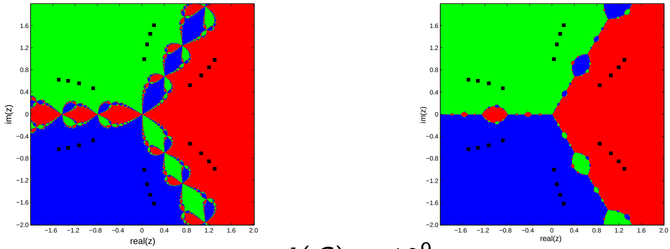
$$A \in \mathbb{C}^{n\times n}, \qquad Y = A^{1/p}$$

$$C = \begin{bmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & \ddots & I \\ A & & & & 0 \end{bmatrix} \in \mathbb{C}^{pn\times pn}.$$

$$\mathrm{sect}_p(C) = \begin{bmatrix} 0 & Y^{-1} & & 0 \\ \vdots & 0 & \ddots & \\ 0 & \ddots & \ddots & Y^{-1} \\ AY^{-1} & 0 & \cdots & 0 \end{bmatrix}.$$

eigenvalues of $A \in \mathbb{R}^{8\times 8}$: $\frac{-k^2}{10} \pm ik, \qquad k = 1, 2, 3, 4$

black boxes - eigenvalues of C for $p = 3$, convergence regions



$$\mathrm{cond}(C) \approx 10^9$$

C has 4 groups of eigen. with $2p$ eigenvalues
with the same module in each group

for $p = 3$: $\beta(U) \approx \mathbf{10^{16}}$

for $p = 6$: $\beta(U) \approx \mathbf{10^{35}}$

$U = \mathrm{sect}(R)$, R quasi-triang. from Schur decomp. of C

Table: Results for C

$n = 24, \quad p = 3, \quad \|\hat{X}\| = 1.71 \times 10^6, \quad iter_{\text{Newt}} = 8, \quad iter_{\text{Hall}} = 5$

$alg.$	$\|\hat{X}^p - I\|$	$\|C\hat{X} - \hat{X}C\|$	$\frac{\|C\hat{X}-\hat{X}C\|}{\|\hat{X}\|\|C\|}$
Newt	$1.32e-09$	$1.50e-09$	$1.94e-18$
Hall	$1.88e-09$	$3.12e-09$	$4.03e-18$
r − Sch	**2.76e − 06**	$8.91e-08$	$1.15e-16$

$n = 48, \quad p = 6, \quad \|\hat{X}\| = 8.76 \times 10^5, \quad iter_{\text{Newt}} = 9, \quad iter_{\text{Hall}} = 5$

$alg.$	$\|\hat{X}^p - I\|$	$\|C\hat{X} - \hat{X}C\|$	$\frac{\|C\hat{X}-\hat{X}C\|}{\|\hat{X}\|\|C\|}$
Newt	$5.07e-09$	$3.21e-09$	$8.10e-18$
Hall	$4.00e-09$	$3.57e-09$	$9.03e-18$
r − Sch	**8.81e − 04**	$5.81e-08$	$1.47e-16$

for $p = 6$ $\qquad max_j|\lambda_j^{\text{schur}} - \lambda_j^A| \approx 10^{-10}$

Example 2

$$A = D + T, \qquad D = \mathrm{diag}(\lambda_j), \quad \text{complex}$$

T triangular real, $\quad n = 40$

Table: Results for A

$$p = 5, \quad \|\hat{X}\| = 1.1, \quad iter_{\mathrm{Newt}} = 28, \quad iter_{\mathrm{Hall}} = 16$$

alg.	$\|\hat{X}^p - I\|$	$\|A\hat{X} - \hat{X}A\|$	$\frac{\|A\hat{X}-\hat{X}A\|}{\|\hat{X}\|\|A\|}$
Newt	$6.40e-16$	$5.57e-15$	$4.13e-17$
Hall	$1.45e-15$	$\mathbf{1.65e-11}$	$1.22e-13$

Example 3

A as in the previous example, $n = 10$

$$p = 4, \quad \|\hat{X}\| = 1.01, \quad iter_{\mathrm{Newt}} = 22, \quad iter_{\mathrm{Hall}} = 13$$

$alg.$	$\|\hat{X}^p - I\|$	$\frac{\|\hat{X}^p - I\|}{\|\hat{X}\|^p}$	$\|A\hat{X} - \hat{X}A\|$	$\frac{\|A\hat{X} - \hat{X}A\|}{\|\hat{X}\|\|A\|}$
Newt	$2.68e-18$	$2.54e-18$	$1.47e-15$	$1.52e-17$
Hall	$4.44e-16$	$4.22e-16$	$4.32e-15$	$4.46e-17$
c − Sch	$2.68e-18$	$2.55e-18$	$3.57e-16$	$3.69e-18$

slow convergence of Newton

$$p = 10, \quad \|\hat{X}\| = 1.02, \quad \mathbf{iter}_{\mathrm{Newt}} = \mathbf{51}, \quad iter_{\mathrm{Hall}} = 28$$

$alg.$	$\|\hat{X}^p - I\|$	$\frac{\|\hat{X}^p - I\|}{\|\hat{X}\|^p}$	$\|A\hat{X} - \hat{X}A\|$	$\frac{\|A\hat{X} - \hat{X}A\|}{\|\hat{X}\|\|A\|}$
Newt	$1.32e-15$	$1.08e-15$	$1.75e-15$	$1.47e-17$
Hall	$1.94e-15$	$1.59e-15$	$\mathbf{3.29e-08}$	$2.76e-10$
c − Sch	$1.28e-15$	$1.05e-15$	$4.11e-16$	$3.45e-18$

Summary

- Real Schur algorithm for the matrix sector function was proposed.
- Some convergence regions of Newton's and Halley's iterations were given.
- Conditioning and stability of the algorithms were discussed.
- Numerical experiments were presented.
  - the commutativity condition was not well satisfied by Halley in some cases,
  - accuracy of Schur algorithm for A with multiple eigenvalues was bad.

Other results in PhD of Beata Laszkiewicz, in preparation.

Summary

- Real Schur algorithm for the matrix sector function was proposed.
- Some convergence regions of Newton's and Halley's iterations were given.
- Conditioning and stability of the algorithms were discussed.
- Numerical experiments were presented.
  - the commutativity condition was not well satisfied by Halley in some cases,
  - accuracy of Schur algorithm for A with multiple eigenvalues was bad.

Other results in PhD of Beata Laszkiewicz, in preparation.

Summary

- Real Schur algorithm for the matrix sector function was proposed.
- Some convergence regions of Newton's and Halley's iterations were given.
- Conditioning and stability of the algorithms were discussed.
- Numerical experiments were presented.
  - the commutativity condition was not well satisfied by Halley in some cases,
  - accuracy of Schur algorithm for A with multiple eigenvalues was bad.

Other results in PhD of Beata Laszkiewicz, in preparation.

Summary

- Real Schur algorithm for the matrix sector function was proposed.
- Some convergence regions of Newton's and Halley's iterations were given.
- Conditioning and stability of the algorithms were discussed.
- Numerical experiments were presented.
  - the commutativity condition was not well satisfied by Halley in some cases,
  - accuracy of Schur algorithm for A with multiple eigenvalues was bad.

Other results in PhD of Beata Laszkiewicz, in preparation.

Thank you for your attention!

References

- N.J. Higham, *Functions of a matrix: Theory and Computation*, SIAM 2008.
- B. Iannazzo, *Numerical solution of certain nonlinear matrix equations*, PhD, Pisa 2007.
- C. Kenney, A. Laub, *Rational iterative methods for the matrix sign function*, SIAM J. MAtrix Anal. Appl. 12 (2): 273 — 291, 1991.
- Ç.K. Koç, B. Bakkaloğlu, *Halley's method for the matrix sector function* IEEE Trans. on Automatic Control 40 (5): 994 — 948, 1995.
- L.S. Shieh, Y.T. Tsay, C.T. Wang, *Matrix sector function and their applications to system theory*, IEE Proceedings 131 (5): 171 — 181, 1984.