Algorithms for matrix sector function

Krystyna Ziętak

(the joint work with Beata Laszkiewicz)

Institute of Mathematics and Computer Science
Wrocław University of Technology

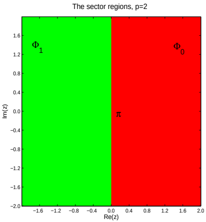
Manchester, May 15 - 16, 2008

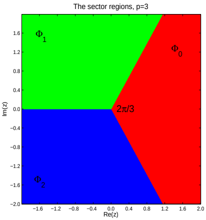
Outline

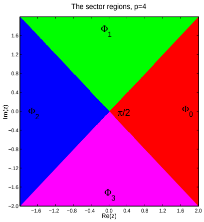
The sector regions

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

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







The scalar p-sector function

- $s_p(\lambda)$ is the nearest *pth* root of unity

  (which lies in the same sector Φ_k in which λ is).

- $s_p(\lambda)$ is not defined for the *pth* roots of nonpositive real numbers.

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

Principal matrix *p*th 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 sector Φ_0.

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

- $A \in \mathbb{C}^{n \times n}$ nonsingular
- $arg(\lambda_j) \neq 2\pi(q + \frac{1}{2})/p$

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

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

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

Matrix sector function is a specific *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 $J_k(\lambda)$

$p = 2$ matrix sign function

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
- Padé family of iterations

Real Schur algorithm for sector

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

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

Parlett 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.

Smith 2003 - any primary matrix pth root

Newton's method for sector

Shieh, Tsay, Wang, 1984

$$X_0 = A$$

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

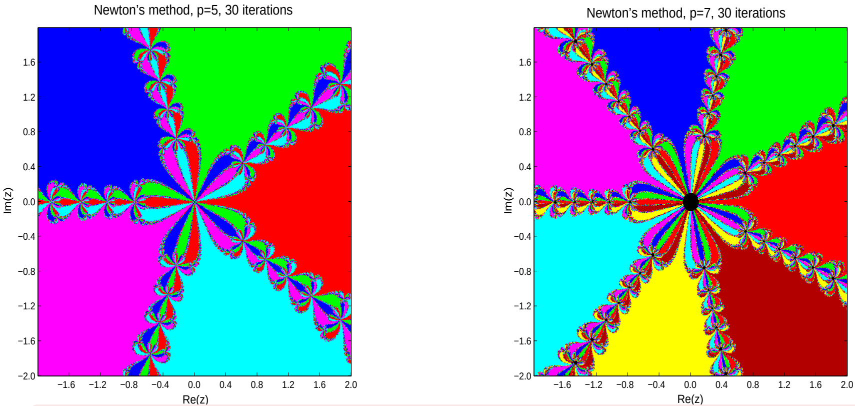
Newton's method is applied to the scalar equation

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

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

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

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

$$\mathbb{C}_k = \left\{ z \in \mathbb{C} : \frac{1}{2^{1/p}} \leq |z| \leq 1, \frac{2k\pi}{p} - \frac{\pi}{4p} < \arg(z) < \frac{2k\pi}{p} + \frac{\pi}{4p} \right\}$$

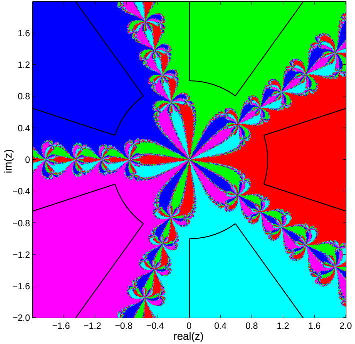
$$\mathbb{R}_k^+ = \left\{ z \in \mathbb{C} : \ z = r\epsilon_k, r \in \mathbb{R}^+ \right\}$$

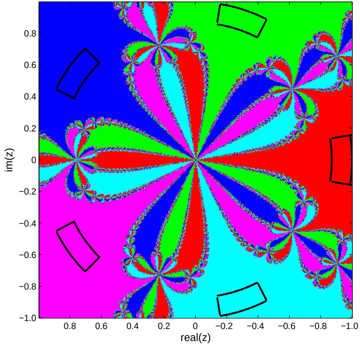
then Newton is convergent.

Iannazzo *p*th roots

Convergence regions of Newton

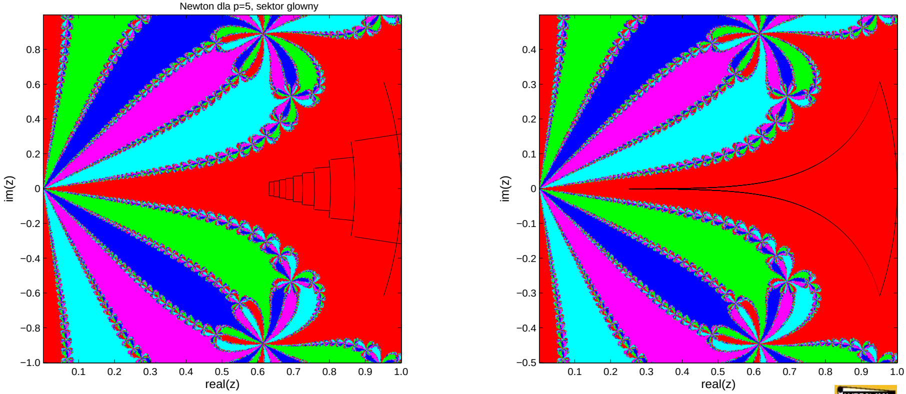
Region $\mathbb{B}_k$ and Region $\mathbb{C}_k$





Convergence regions of Newton

Additional regions



Halley's method for sector

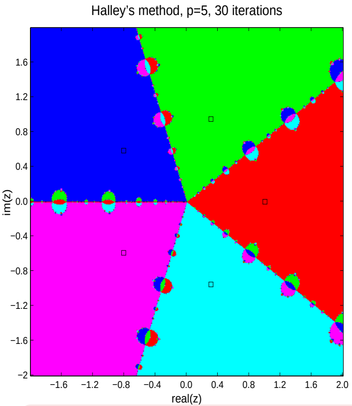
Bakkaloğlu, Koç, 1995

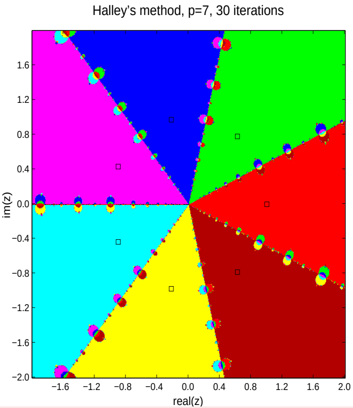
$$X_0 = A$$

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

Regions of convergence of Halley for sector

determined experimentally



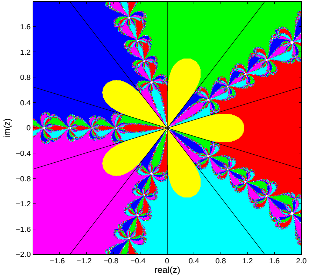


ω_j pth root of unity

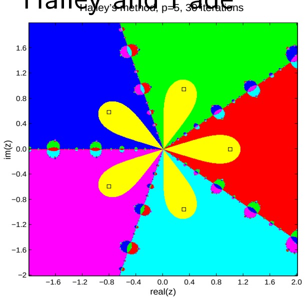
color: $|z_{30} - \omega_j| < 10^{-5}$

Halley for sector

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

From the theorems of Higham and Iannazzo 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 Schur method for pth roots

Smith 2003
Let $\hat{U}$ be computed upper triangular pth roots of R from Schur decomposition of A. Then

$$\hat{U}^p = R + E, \qquad |E| \leq c\ p\ n\ u|\hat{U}|^p$$

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

Schur method for pth root is stable provided $\beta(U)$ is sufficiently small

The Padé family iterations for sector function

$$s_p(\lambda) = \frac{\lambda}{\sqrt[p]{\lambda^p}} = \frac{\lambda}{\sqrt[p]{1-z}}$$

$$z = 1 - \lambda^p$$

$$x_{i+1} = x_i \frac{P_{km}(1 - x_i^p)}{Q_{km}(1 - x_i^p)}$$

$$x_0 = \lambda_j$$

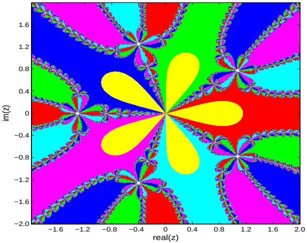
P_{km}/Q_{km} - $[k/m]$ Padé approximant to $(1 - z)^{-1/p}$

$p = 2$ sign function

Kenney, Laub, 1991

The convergence region of Padé iterations $[m-1/m]$

$[0/1]$



$[1/2]$

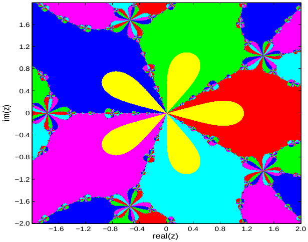
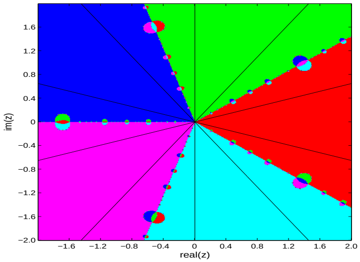
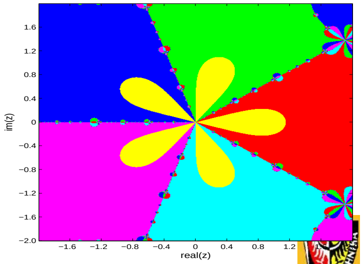


Figure: The convergence region of Padé iterations

[4/4]



[3/4]



Padé $[k/m]$ for sector

"yellow flowers" - known for $p = 2$ (Kenney-Laub)

First observation - Padé for sector

For $k \geq m - 1$, if

$$x_{n+1} = x_n \frac{P_{km}(1 - x_n^p)}{Q_k m(1 - x_n^p)}$$

$$|1 - x_0^p| < 1$$

then

$$|1 - x_n^p| \leq |1 - x_0^p|^{(k+m+1)^n}$$

$$\lim_{n \to \infty} x_n \to s_p(x_0)$$

Principal Padé for sector

Principal Padé iteration for matrix sector function preserve structure (automorphism group)!!!

Second observation - principal Padé for sector

If all eigenvalues of A lie in $\mathbb{B}_p^{\mathrm{hall}}$ then principal Pade $[m/m]$ iterations are convergent to sector.

"Yellow flowers"' lie in this region.

Implementation

Newton

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

Halley 1

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

Halley 2

$$X_{k+1} = \frac{p-1}{p+1}X_k + \frac{4p}{p+1}X_k\left[(p+1)X_k^p + (p-1)I\right]^{-1}$$

Example 1

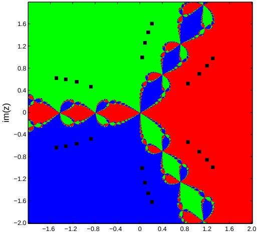
$$A \in \mathbb{C}^{n\times n}, \qquad Y = A^{1/p}$$

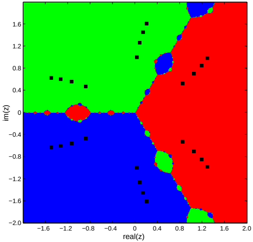
$$C = \begin{bmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & \ddots & I \\ A & & & & 0 \end{bmatrix} \in \mathbb{C}^{pn\times pn}.$$

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

A in real Schur form, $\mathrm{cond}(A) = 1.4e + 16$
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





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

Table: Results for C, $\mathrm{cond}(C) \approx 10^9$

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

alg.	CPU	$\|\hat{X}^p - I\|$	$\|C\hat{X} - \hat{X}C\|$	$\frac{\|C\hat{X}-\hat{X}C\|}{\|\hat{X}\|\|C\|}$
Newt	$1.14e-02$	$1.32e-09$	$1.50e-09$	$1.94e-18$
Hall 1	$8.68e-03$	$1.88e-09$	$3.12e-09$	$4.03e-18$
Hall 2		$2.34e-09$	**$6.67e-06$**	$8.62e-15$
r − Sch	$6.17e-02$	**$2.76e-06$**	$8.91e-08$	$1.15e-16$
p − root		$6.02e-09$	$5.37e-09$	$6.94e-18$

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

alg.	CPU	$\|\hat{X}^p - I\|$	$\|C\hat{X} - \hat{X}C\|$	$\frac{\|C\hat{X}-\hat{X}C\|}{\|\hat{X}\|\|C\|}$
Newt	$3.74e-02$	$5.07e-09$	$3.21e-09$	$8.10e-18$
Hall 1	$2.63e-02$	$4.00e-09$	$3.57e-09$	$9.03e-18$
Hall 2		$2.08e-09$	**$6.27e-07$**	$1.58e-15$
r − Sch	$4.98e-01$	**$8.81e-04$**	$5.81e-08$	$1.47e-16$

Accuracy of computed Schur decomposition of C for $p = 6$, $n = 48$, $\mathrm{cond}(C) \approx 10^9$

$$\max_j |\lambda_j^{\mathrm{schur}} - \lambda_j^{A}| \approx 1.21e - 10$$

$$\max_j |\lambda_j^{\mathrm{schur}} - \lambda_j^{eig}| \approx 1.12e - 10$$

$$\max_j |\lambda_j^{A} - \lambda_j^{eig}| \approx 3.14e - 11$$

- $\lambda_j^{\mathrm{schur}}$ - eigenvalues of C computed directly from diagonal blocks of R
- λ_j^{eig} - eigenvalues of C computed by `eig`
- λ_j^{A} - eigenvalues of C computed as pth roots of eigenvalues of A

$$\beta = \frac{||\mathrm{sect}_p(R)||_F^p}{||R||_F} = 1.25e + 35$$

R triangular from Schur decomposition of C, Matlab 6.5

Example 2

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

A triangular

T triangular real, $\quad n = 40 \quad \mathrm{cond}(A) = 9.81$

Table: Results for A

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

alg.	*CPU*	$\|\hat{X}^p - I\|$	$\|A\hat{X} - \hat{X}A\|$	$\frac{\|A\hat{X} - \hat{X}A\|}{\|\hat{X}\|\|A\|}$
Newt	$3.12e-01$	$6.40e-16$	$5.57e-15$	$4.13e-17$
Hall 1	$2.51e-01$	$1.45e-15$	$\mathbf{1.65e-11}$	$1.22e-13$
Hall 2		$8.26e-16$	$\mathbf{4.23e-15}$	$3.13e-17$
c − Sch	$9.83e-02$	$1.1e-15$	$2.13e-15$	$1.59e-17$

Example 3

slow convergence of Newton

$n = 10, \quad A$ as in Example 2, complex triangular

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

$alg.$	CPU	$\|\hat{X}^p - I\|$	$\|A\hat{X} - \hat{X}A\|$	$\frac{\|A\hat{X}-\hat{X}A\|}{\|\hat{X}\|\|A\|}$
Newt	$3.59e-02$	$1.32e-15$	$1.75e-15$	$1.47e-17$
Hall 1	$3.03e-02$	$1.94e-15$	$\mathbf{3.29e-08}$	$2.76e-10$
Hall 2		$8.90e-16$	$1.53e-15$	$1.29e-17$
c − Sch	$1.00e-02$	$1.28e-15$	$4.11e-16$	$3.45e-18$

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 $A = SN$,

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

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

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

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

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

Summary

- Real Schur algorithm for the matrix sector function was proposed.
- Some convergence regions of Newton's and Halley's iterations were proven.
- Padé family for the matrix sector function was introduced.
- 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 sometimes not good because of inaccuracy in computed by `MATLAB` Schur decomposition and ill conditioning.

Other results in PhD of Beata Laszkiewicz.

Summary

- Real Schur algorithm for the matrix sector function was proposed.
- Some convergence regions of Newton's and Halley's iterations were proven.
- Padé family for the matrix sector function was introduced.
- 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 sometimes not good because of inaccuracy in computed by MATLAB Schur decomposition and ill conditioning.

Other results in PhD of Beata Laszkiewicz.

Summary

- Real Schur algorithm for the matrix sector function was proposed.
- Some convergence regions of Newton's and Halley's iterations were proven.
- Padé family for the matrix sector function was introduced.
- 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 sometimes not good because of inaccuracy in computed by MATLAB Schur decomposition and ill conditioning.

Other results in PhD of Beata Laszkiewicz.

Summary

- Real Schur algorithm for the matrix sector function was proposed.
- Some convergence regions of Newton's and Halley's iterations were proven.
- Padé family for the matrix sector function was introduced.
- 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 sometimes not good because of inaccuracy in computed by `MATLAB` Schur decomposition and ill conditioning.

Other results in PhD of Beata Laszkiewicz.

Thank you for your attention!

References

- 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.