Dual Padé iterations for the matrix pth sector function and pth root, and one topic more

Krystyna Ziętak

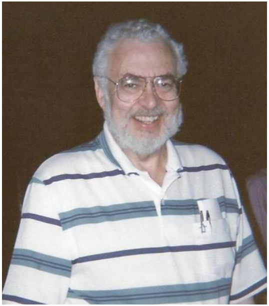
Wrocław School of Information Technology

the second affiliation: Wroclaw University of Technology

Householder Symposium XIX, Spa
June 10, 2014

Outline

- 1993 - Householder Symposium XII, Lake Arrowhead
  invitation by Gene Golub and Tony Chen



- 2014 - Householder Symposium XIX, Spa
  invitation by Ilse Ipsen and Paul Van Dooren

- 1996 - Householder Symposium XIII, Pontresina
  Diane O'Leary, W. Gander, M. Guthnecht

Ilse Ipsen (Pontresina 1996)
Paul Van Dooren (ILAS, Chemnitz 1996)

Part I - dual Padé family of iterations for matrix p-sector function and pth root

Guo (2010) shows some common properties of Newton and Halley iterations for computing matrix pth root.

Why "Newton" and "Halley" have some common properties?

"Halley" is Padé iteration, but "Newton" is not.

It was reason for which I introduce in 2013 a new dual Padé family of iterations which includes "Newton" and "Halley".

Dual Padé iterations have properties observed by Guo for "Newton" and "Halley".

Part II - join work with Joao Cardoso (Coimbra, Portugal)

sub-Stiefel Procrustes problem

Let $\mathbb{S}_{sub}$ denote set of **sub-Stiefel matrices** of order n

sub-Stiefel matrix is obtained by taking off the last row and last column of orthogonal matrix of order $n+1$

We find matrix X_* for which minimum is reached

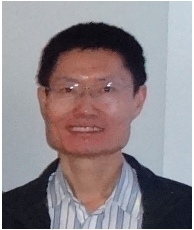
$$\min_{X \in \mathbb{S}_{sub}} \|A - BX\|_F$$

A, B given square matrices

PART I

"*Padé people*" - Kenney and Laub (see their home pages)

inspirations: Guo, Higham and Iannazzo



PART II

inspirations: Elden, Ferreira and Park (see their home pages)

Iterations generated by Padé approximants

- sign function
  Kenney, Laub (1991)
- square root
  Higham (1997)
  Higham, Mackey, Mackey, Tisseur (2004)
- polar decomposition
  Higham, Functions of Matrices,... (2008)
- p-sector function and pth root
  Laszkiewicz, Zię̧tak (2009)
- sign function
  reciprocal Padé iterations
  Greco-Iannazzo-Poloni (2012)
- p-sector function and pth root
  dual Padé iterations, *Zię̧tak* (2014)

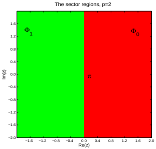
scalar p-sector function

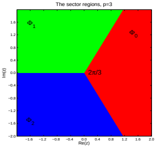
$$\mathrm{sect}_p(z) = \frac{z}{\sqrt[p]{z^p}}$$

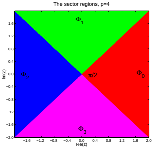
the nearest *pth* root of unity to z

for $p = 2$ sign function

$$\mathrm{sign}(z) = \begin{cases} 1 & \text{if } \mathrm{Re}(z) > 0 \\ -1 & \text{if } \mathrm{Re}(z) < 0 \end{cases}$$







red - principal sector Φ_0

Principal *p*th root and *p*-sector function

matrix principal *p*th root $X = A^{1/p}$

no eigenvalue of A lies on closed negative real axis

$$X^p = A, \qquad \lambda_j(X) \in \Phi_0$$

matrix *p*-sector function

A nonsingular

$$\arg(\lambda_j(A)) \neq (2q+1)\pi/p, \qquad q = 0, 1, \ldots, p-1$$

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

Gauss hypergeometric function

$$g(z) = (1-z)^{-1/p} = {}_2F_1(1/p, 1; 1; z)$$

Padé approximants $P_{km}(z)/Q_{km}(z)$ to $g(z)$

for all integer k, m

$$\frac{P_{km}(z)}{Q_{km}(z)} = \frac{{}_2F_1(-k, \frac{1}{p} - m; -k - m; z)}{{}_2F_1(-m, -\frac{1}{p} - k; -k - m; z)}$$

Gomilko, Greco, KZ (Numer. Lin. Alg. Appl. 2012)
Gomilko, Karp, Lin, KZ (J. Comput. Appl. Math. 2012)
KZ (J. Comput. Appl. Math. 2014)

roots and poles of Padé approximants $P_{km}(z)/Q_{km}(z)$

- If $1 \leq k \leq m$, then all roots of $P_{km}(z)$ lie in $(1, \infty)$
- If $k > m \geq 1$, then m roots lie in $(1, \infty)$, remaining roots have moduli bigger than 1
- all poles have moduli bigger than 1

positivity of coefficients of power series expansions of

$$\frac{P_{km}(z)}{Q_{km}(z)}, \qquad \frac{1}{Q_{km}(z)}$$

$$f_{km}(z) = 1 - (1 - z)\left(\frac{P_{km}(z)}{Q_{km}(z)}\right)^p$$

Iterations for matrix p-sector function

Pade $$X_{j+1} = X_j \frac{P_{km}(I - X_j^p)}{Q_{km}(I - X_j^p)}, \qquad X_0 = A$$

Laszkiewicz, KZ (2009)

for $p = 2$ (sign) *Kenney-Laub (1991)*

Halley $k = m = 1$

dual Pade $$X_{j+1} = X_j \frac{Q_{km}(I - X_j^{-p})}{P_{km}(I - X_j^{-p})}, \qquad X_0 = A$$

KZ (2013)

Halley $k = m = 1$, Newton $k = 0, m = 1$

Schröder $k = 0$, m arbitrary *Cardoso, Loureiro (2011)*

Principal Padé iterations ($k = m$) for p-sector are **structure preserving**.

arbitrary p - Laszkiewicz, KZ (2009)
$p = 2$ Higham, Mackey, Mackey, Tisseur (2004)

After suitable change of variable, (dual) Padé iterations for p-sector can be applied to computing

- pth roots
- square root ($p = 2$)
- polar decomposition

Convergence

Pure matrix iterations (Iannazzo, 2008):
convergence of scalar sequences of eigenvalues → convergence of matrix sequences

Certain regions of convergence for p-sector function

eigenvalues $\lambda_j(A)$ in regions:

Padé

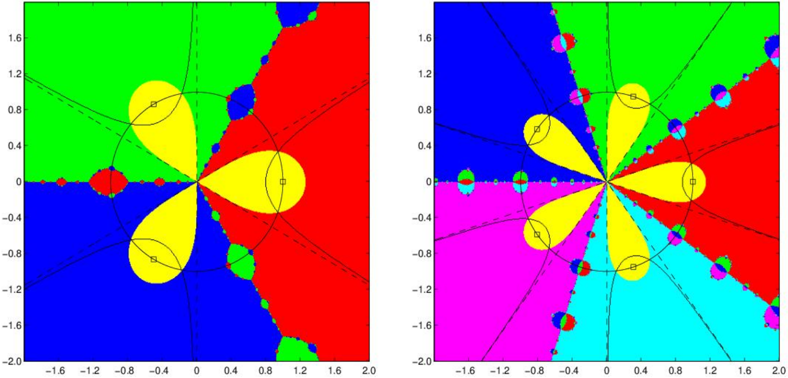
$$\mathbb{L}_p = \{z \in \mathbb{C} : |1 - z^p| < 1\}, \quad X_0 = A$$

"yellow flowers"

dual Padé

$$\mathbb{L}_{-p} = \{z \in \mathbb{C} : |1 - z^{-p}| < 1\}, \quad X_0 = A$$

solid countur



Halley iterations for $p = 3$ and $p = 5$

the unit circle (solid contour), the pth roots of unity (boxes)

$\mathbb{L}_p^{(Pade)}$ for "Padé" (yellow flower)

$\mathbb{L}_{-p}^{(Pade)}$ for "dual Padé" (solid contour)

$p = 2$ sign function,
Padé approximant to $g(z) = (1 - z)^{-1/2}$

Padé for sign (Kenney-Laub, 1991)

$$Y_{j+1} = Y_j \frac{P_{km}(I - Y_j^2)}{Q_{km}(I - Y_j^2)}, \qquad Y_0 = A$$

reciprocal Padé for sign (Greco-Iannazzo-Poloni, 2012)

$$Y_{j+1} = \frac{Q_{km}(I - Y_j^2)}{Y_j P_{km}(I - Y_j^2)}, \qquad Y_0 = A$$

dual Padé for sign (KZ 2014)

$$Y_{j+1} = \frac{Y_j Q_{km}(I - Y_j^{-2})}{P_{km}(I - Y_j^{-2})}, \qquad Y_0 = A$$

Properties of dual Padé family for pth root

residuals for Padé iteration generated by $[k/m]$

$$S_\ell = I - A^{-1}X_\ell^p$$

$$S_{\ell+1} = f_{km}(S_\ell)$$

residuals for dual Padé iteration generated by $[k/m]$

$$R_\ell = I - AX_\ell^{-p}$$

$$R_{\ell+1} = f_{km}(R_\ell)$$

$$f_{km}(z) = 1 - (1-z)\left(\frac{P_{km}(z)}{Q_{km}(z)}\right)^p$$

Guo (2010) applies "dual residuals" to investigation of convergence of Newton and Halley iterations

binomial expansion

$$(1-z)^{1/p} = \sum_{j=0}^{\infty} \beta_j z^j$$

the ℓth iterate Y_ℓ, computed by the dual Padé iteration generated by $[k/m]$ Padé approximant applied to computing $(I-B)^{1/p}$, satisfies

$$Y_\ell = \sum_{j=0}^{\infty} \varphi_{km,j}^{(\ell)} B^j$$

where $\varphi_{km,j}^{(\ell)} = \beta_j$ for $j = 0, \ldots, (k+m+1)^\ell - 1$

Guo (2010) - Newton ($k=0, m=1$) and Halley ($k=m=1$)
KZ (2014) - arbitrary k, m

PART II - sub-Stiefel Procrustes problem

characterization

X is sub-Stiefel iff $\sigma(X) = \{1, \ldots, 1, s\}, \quad 0 \leq s \leq 1$

Procrustes problems (Frobenius norm):

$$\min_{X \in \mathbb{M}} \|A - BX\|_F$$

where $\mathbb{M}$:

- orthogonal matrices (Green, 1952)
- symmetric matrices (Higham, 1988)
- Stiefel matrices (Elden and Park, 1999)
- other types (Andresson and Elfiving 1997)
- sub-Stiefel (join work with Cardoso, 2014)

Motivation for sub-Stiefel matrices:
surface unfolding problem in **computer vision***:*
reconstructing smooth, flexible and isometrically embedded flat surfaces (*Ferreira, Xavier, Costeira* (2009), $n = 2$)

- Properties of sub-Stiefel matrices.
- Necessary conditions for solution of sub-Stiefel Procrustes problem.
- When sub-Stiefel Procrustes problem has orthogonal solution?
- Iterative algorithm - in each iteration one solves some orthogonal Procrustes problem.

Cardoso, KZ, Numerical Linear Alg. Appl.

Let X_* be solution of **sub-Stiefel Procrus.**

$$\min_{X \in \mathbb{S}_{sub}} \|A - BX\|_F = \|A - BX_*\|_F$$

Let

$$Y_* = \begin{bmatrix} X_* & u_* \\ v_*^T & \alpha_* \end{bmatrix}$$

orthogonal for appropriate vectors u_*, v_* and number α_*.
Then Y_* is the solution of the **orth. Procrus.** problem (with extended matrices):

$$\min_{Y \mathrm{orth}} \left\| \begin{bmatrix} A & Bu_* \\ v_*^T & \alpha_* \end{bmatrix} - \begin{bmatrix} B & 0 \\ 0^T & 1 \end{bmatrix} Y \right\|_F =$$

$$\left\| \begin{bmatrix} A & Bu_* \\ v_*^T & \alpha_* \end{bmatrix} - \begin{bmatrix} B & 0 \\ 0^T & 1 \end{bmatrix} Y_* \right\|_F$$

Iterative algorithm for sub-Stiefel Procrustes problem

- Let X_0 be initial approximation of solution X_*.
- The next iterate X_1 is subblock of solution Y_1 of orthogonal Procrustes problem:

$$\min_{Y \text{orth.}} \left\| \begin{bmatrix} A & Bu_0 \\ v_o^T & \alpha_0 \end{bmatrix} - \begin{bmatrix} B & 0 \\ 0^T & 1 \end{bmatrix} Y \right\|_F,$$

where $\alpha_0 = \sigma_{min}(X_0)$ and vectors u_0, v_0 such that

$$\begin{bmatrix} X_0 & u_0 \\ v_0^T & \alpha_0 \end{bmatrix} \quad \text{orthogonal.}$$

compare iterative algorithms
for Stiefel (unbalanced) Procrustes problem:
Ten Berg (1984), Park (1991), Zhang 2006, ...

Summary

- Properties of Padé approximants to $(1-z)^{-1/p}$ have been shown.
- The dual Padé iterations for matrix sector function and matrix pth root have been introduced and investigated.
- Sub-Stiefel Procrustes problem has been formulated and investigated.
- Iterative algorithm for Sub-Stiefel Procrustes problem has been proposed.

References I

- **Cardoso, Loureiro**, On the convergence of Schröder iteration function for the pth roots of complex numbers, *Applied Math. Comput.* (2011).
- **Elden, Park**, A Procrustes problem on the Stiefel manifold, *Numer. Math.* (1999).
- **Greco, Iannazzo, Poloni**, The Padé iterations for the matrix sign function and their reciprocal are optimal, *Lin. Alg. Appl.* (2012).
- **Guo**, On Newton's and Halley's method for the principal pth root of a matrix, *Lin. Alg. Appl.* (2010).
- **Higham, Mackey, Mackey, Tisseur**, Computing the polar decomposition and the matrix sign decomposition in matrix group, *SIAM J. Matrix Anal.* 25 (2004).
- **Kenney, Laub**, Rational iterative methods for the matrix sign function, *SIAM J. Matrix Anal. Appl.* 12 (1991).

References II

- **Cardoso, Ziętak**, On a sub-Stiefel Procrustes problem arising in computer vision, *Numer. Lin. Alg. Appl. submitted*.
- **Gomilko, Greco, Ziętak,** A Padé family of iterations for the matrix sign function and related problems, *Numer. Lin. Alg. Appl.* (2011).
- **Gomilko, Karp, Lin, Ziętak,** Regions of convergence of a Padé family of iterations for the matrix sector function, *J. Comput. Appl. Math.* (2012).
- **Laszkiewicz, Ziętak,** A Padé family of iterations for the matrix sector function and the matrix *p*th root, *Numer. Lin. Alg. Appl.* (2009).
- **Ziętak**, The dual Padé families of iterations for the matrix *p*th root and the matrix *p*-sector function, *J. Comput. Appl. Math.* (2014).