MANCHESTER 1824

The University of Manchester

Numerical Analysis of the Matrix Logarithm

Nick Higham
School of Mathematics
The University of Manchester

higham@ma.man.ac.uk
http://www.ma.man.ac.uk/~higham/

Computational Methods with Applications
Harrachov 2007

Outline

1. Definition of $\log(A)$
2. Applications
3. Theory
4. Numerical methods

Matrix Logarithm

A logarithm of $A \in \mathbb{C}^{n\times n}$ is any matrix X such that $e^X = A$.

- Existence.
- Representation, classification.
- Computation.
- Conditioning.

First, approach via **theory of matrix functions. . .**

Multiplicity of Definitions

There have been proposed in the literature since 1880
eight distinct definitions *of a matric function,*
by Weyr, Sylvester and Buchheim,
Giorgi, Cartan, Fantappiè, Cipolla,
Schwerdtfeger and Richter.

— **R. F. Rinehart,**
The Equivalence of Definitions of a Matric Function,
Amer. Math. Monthly (1955)

Jordan Canonical Form

$$Z^{-1}AZ = J = \mathrm{diag}(J_1, \dots, J_p), \qquad \underbrace{J_k}_{m_k \times m_k} = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix}$$

Definition

$$f(A) = Zf(J)Z^{-1} = Z\mathrm{diag}(f(J_k))Z^{-1},$$

$$f(J_k) = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \dots & \dfrac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix}.$$

Interpolation

Definition (Sylvester, 1883; Buchheim, 1886)

Distinct e'vals $\lambda_1, \dots, \lambda_s$, n_i = max size of Jordan blocks for λ_i. Then $f(A) = p(A)$, where p is unique Hermite interpolating poly of degree $< \sum_{i=1}^{s} n_i$ satisfying

$$p^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \qquad j = 0\colon n_i - 1, \quad i = 1\colon s.$$

Cauchy Integral Theorem

Definition

$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - A)^{-1}\, dz,$$

where f is analytic on and inside a closed contour Γ that encloses $\lambda(A)$.

Equivalence of Definitions

Theorem

*The three definitions are **equivalent**, modulo analyticity assumption for Cauchy.*

Composite Functions

Theorem

$f(t) = g(h(t)) \;\Rightarrow\; f(A) = g(h(A))$, *provided latter matrix defined.*

Corollary

$\exp(\log(A)) = A$ when $\log(A)$ is defined.

Outline

1. Definition of $\log(A)$
2. Applications
3. Theory
4. Numerical methods

Application: Markov Models

Time-homogeneous continuous-time Markov process with transition probability matrix $P(t) \in \mathbb{R}^{n\times n}$. Transition intensity matrix Q related to P by

$$P(t) = e^{Qt}.$$

Elements of Q satisfy

$$q_{ij} \geq 0, \quad i \neq j, \qquad \sum_{j=1}^{n} q_{ij} = 0.$$

Embeddability problem

When does a given **stochastic** P have a real logarithm Q that is an **intensity matrix**?

The Average Eye

First order character of optical system characterized by transference matrix $T = \begin{bmatrix} S & \delta \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{5\times 5}$, where $S \in \mathbb{R}^{4\times 4}$ is symplectic: $S^T J S = J$, where $J = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix}$.

Average $m^{-1}\sum_{i=1}^{m} T_i$ is not a transference matrix.

Harris (2005) proposes the average $\exp(m^{-1}\sum_{i=1}^{m} \log(T_i))$.

The Average Eye

First order character of optical system characterized by transference matrix $T = \begin{bmatrix} S & \delta \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{5\times 5}$, where $S \in \mathbb{R}^{4\times 4}$ is symplectic: $S^T J S = J$, where $J = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix}$.

Average $m^{-1}\sum_{i=1}^{m} T_i$ is not a transference matrix.

Harris (2005) proposes the average $\exp(m^{-1}\sum_{i=1}^{m}\log(T_i))$.

For Hermitian pos def A and B, Arsigny et al. (2007) define the log-Euclidean mean

$$E(A,B) = \exp(\tfrac{1}{2}(\log(A) + \log(B))).$$

Outline

Logs of $A = I_3$

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 2\pi - 1 & 1 \\ -2\pi & 0 & 0 \\ -2\pi & 0 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 & 2\pi & 1 \\ -2\pi & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$e^B = e^C = e^D = I_3$.

$\Lambda(C) = \Lambda(D) = \{0, 2\pi i, -2\pi i\}$.

Principal Log and pth Root

Let $A \in \mathbb{C}^{n\times n}$ have no eigenvalues on $\mathbb{R}^-$.

Principal log

$X = \log(A)$ denotes unique X such that

- $e^X = A$.
- $-\pi < \operatorname{Im}\big(\lambda(X)\big) < \pi$.

For **next 2 slides only**, allow $\operatorname{Im}\big(\lambda(X)\big) = \pi$.

Principal pth root

For integer $p > 0$, $X = A^{1/p}$ is unique X such that

- $X^p = A$.
- $-\pi/p < \arg(\lambda(X)) < \pi/p$.

All Solutions of $e^X = A$

Theorem (Gantmacher)

$A \in \mathbb{C}^{n\times n}$ nonsing with Jordan canonical form $Z^{-1}AZ = J = \mathrm{diag}(J_1, J_2, \dots, J_p)$. All solutions to $e^X = A$ are given by

$$X = Z\,U\,\mathrm{diag}(L_1^{(j_1)}, L_2^{(j_2)}, \dots, L_p^{(j_p)})\,U^{-1} Z^{-1},$$

where

$$L_k^{(j_k)} = \log(J_k(\lambda_k)) + 2\,j_k\,\pi i\, I_{m_k},$$

$j_k \in \mathbb{Z}$ arbitrary, and U an arbitrary nonsing matrix that commutes with J.

All Solutions of $e^X = A$: Classified

Theorem

$A \in \mathbb{C}^{n\times n}$ nonsing: p Jordan blocks, s distinct ei'vals.
$e^X = A$ has a countable infinity of solutions that are **primary functions** of A:

$$X_j = Z\mathrm{diag}(L_1^{(j_1)}, L_2^{(j_2)}, \dots, L_p^{(j_p)})Z^{-1},$$

where $\lambda_i = \lambda_k$ implies $j_i = j_k$. If $s < p$ then $e^X = A$ has **non-primary solutions**

$$X_j(U) = Z\,U\,\mathrm{diag}(L_1^{(j_1)}, L_2^{(j_2)}, \dots, L_p^{(j_p)})\,U^{-1}Z^{-1},$$

where $j_k \in \mathbb{Z}$ arbitrary, U arbitrary nonsing with $UJ = JU$, and for each j $\exists\, i$ and k s.t. $\lambda_i = \lambda_k$ while $j_i \neq j_k$.

Logs of $A = I_3$

$$C = \begin{bmatrix} 0 & 2\pi - 1 & 1 \\ -2\pi & 0 & 0 \\ -2\pi & 0 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 & 2\pi & 1 \\ -2\pi & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$e^0 = e^C = e^D = I_3$. $\Lambda(C) = \Lambda(D) = \{0, 2\pi i, -2\pi i\}$.

$$U = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{bmatrix}, \quad \alpha \in \mathbb{C},$$

$$X = U \mathrm{diag}(2\pi i, -2\pi i, 0) U^{-1} = 2\pi i \begin{bmatrix} 1 & -2\alpha & 2\alpha^2 \\ 0 & 1 & -\alpha \\ 0 & 0 & 1 \end{bmatrix}.$$

Two Facts on Commuting Matrices

Theorem

If $A, B \in \mathbb{C}^{n\times n}$ commute then $\exists$ a unitary $U \in \mathbb{C}^{n\times n}$ such that U^*AU and U^*BU are both upper triangular.

Two Facts on Commuting Matrices

Theorem

If $A, B \in \mathbb{C}^{n\times n}$ commute then $\exists$ a unitary $U \in \mathbb{C}^{n\times n}$ such that U^*AU and U^*BU are both upper triangular.

Theorem

For $A, B \in \mathbb{C}^{n\times n}$, $e^{(A+B)t} = e^{At}e^{Bt}$ for all t if and only if $AB = BA$.

When Does $\log(BC) = \log(B) + \log(C)$?

Theorem

Let $B, C \in \mathbb{C}^{n\times n}$ commute and have no ei'vals on $\mathbb{R}^-$. If for every ei'val λ_j of B and the corr. ei'val μ_j of C, $|\arg\lambda_j + \arg\mu_j| < \pi$, then $\log(BC) = \log(B) + \log(C)$.

When Does $\log(BC) = \log(B) + \log(C)$?

Theorem

Let $B, C \in \mathbb{C}^{n\times n}$ commute and have no ei'vals on $\mathbb{R}^-$. If for every ei'val λ_j of B and the corr. ei'val μ_j of C, $|\arg\lambda_j + \arg\mu_j| < \pi$, then $\log(BC) = \log(B) + \log(C)$.

Proof. $\log(B)$ and $\log(C)$ commute, since B and C do. Therefore

$$e^{\log(B)+\log(C)} = e^{\log(B)}e^{\log(C)} = BC.$$

Thus $\log(B) + \log(C)$ is *some* logarithm of BC. Then

$$\operatorname{Im}(\log\lambda_j + \log\mu_j) = \arg\lambda_j + \arg\mu_j \in (-\pi, \pi),$$

so $\log(B) + \log(C)$ is the *principal* logarithm of BC. $\square$

Outline

1. Definition of $\log(A)$
2. Applications
3. Theory
4. Numerical methods

Henry Briggs (1561–1630)

- **Arithmetica Logarithmica** (1624)
- Logarithms to base 10 of 1–20,000 and 90,000–100,000 to **14 decimal places**.

Henry Briggs (1561–1630)

- **Arithmetica Logarithmica** (1624)
- Logarithms to base 10 of 1–20,000 and 90,000–100,000 to **14 decimal places**.

Briggs must be viewed as one of the great figures in numerical analysis.

—Herman H. Goldstine,
A History of Numerical Analysis (1977)

ARITHMETICA

LOGARITHMICA

SIVE

LOGARITHMORVM CHILIADES TRIGINTA, PRO

numeris naturali ſerie creſcentibus ab vnitate ad 20,000: et a 90,000 ad 100,000. Quorum ope multa perficiuntur Arithmetica problemata et Geometrica.

HOS NVMEROS PRIMVS INVENIT CLARISSIMVS VIR IOHANNES NEPERVS Baro Merchiſtonij: eos autem ex eiuſdem ſententia mutavit, eorumque ortum et vſum illuſtravit HENRICVS BRIGGIVS, in celeberrima Academia Oxonienſi Geometriæ profeſſor SAVILIANVS.

DEVS NOBIS VSVRAM VITÆ DEDIT ET INGENII, TANQVAM PECVNIÆ, NVLLA PRÆSTITVTA DIE.

LONDINI,
Excudebat GVLIELMVS IONES. 1624.

	D	E
	Numeri continue Medij inter Denarium & Unitatem.	Logarithmi Rationales.
	10	1,000
1	31622,77660,16837,93319,98893,54	0,50
2	17782,79410,03892,28011,97304,13	0,25
3	13335,21432,16332,40256,65389,308	0,125
4	11547,81984,68945,81796,61918,213	0,0625
5	10746,07828,32131,74972,13817,6538	0,03125
6	10366,32928,43769,79972,90627,3131	0,01562,5
7	10181,51721,71818,18414,73723,8144	0,00781,25
8	10090,35044,84144,74377,59005,1391	0,00390,625
9	10045,07364,25446,25156,64670,6113	0,00195,3125
10	10022,51148,29291,29154,65611,7367	0,00097,65625
11	10011,24941,39987,98757,85395,52805	0,00048,82812,5
12	10005,62312,60220,86366,18495,91839	0,00024,41406,25
13	10002,81116,78778,01323,99249,64325	0,00012,20703,125
14	10001,40548,51694,72581,62767,32715	0,00006,10351,5625
15	10000,70271,78941,14355,38811,70845	0,00003,05175,78125
16	10000,35135,27746,18565,08581,37077	0,00001,52587,89062,5
17	10000,17567,48442,26738,33846,78274	0,00000,76293,94531,25
18	10000,08783,70363,46121,46574,07431	0,00000,38146,97265,625
19	10000,04391,84217,31672,36281,88083	0,00000,19073,48632,8125
20	10000,02195,91867,55542,03317,07719	0,00000,09536,74316,40625
21	10000,01097,95873,50204,09754,72940	0,00000,04768,37158,20312,5
22	10000,00548,97921,68211,14626,60250,4	0,00000,02384,18579,10156,25
23	10000,00274,48957,07382,95091,25449,9	0,00000,01192,09289,55078,125
24	10000,00137,24477,59510,83282,69572,5	0,00000,00596,04644,77539,0625
25	10000,00068,62238,56210,25737,18748,2	0,00000,00298,02322,38769,53125
26	10000,00034,31119,22218,83912,75020,8	0,00000,00149,01161,19384,76562,5
27	10000,00017,15559,59637,84719,93879,1	0,00000,00074,50580,59692,38281,25
28	10000,00008,57779,79451,03051,17588,8	0,00000,00037,25290,29846,19140,625
29	10000,00004,28889,89633,54198,42901,3	0,00000,00018,62645,14923,09570,3125
30	10000,00002,14444,94793,77767,42970,4	0,00000,00009,31322,57461,54785,15625
31	10000,00001,07222,47391,14050,76926,8	0,00000,00004,65661,28730,77392,57812,5
32	10000,00000,53611,23694,13317,14831,4	0,00000,00002,32830,64365,38696,28906,25
33	10000,00000,26805,61846,70731,51508,7	0,00000,00001,16415,32182,69348,14453,125
34	10000,00000,13402,80923,26383,99277,7	0,00000,00000,58207,66091,34674,07226,5625
35	10000,00000,06701,40461,60946,55519,6	0,00000,00000,29103,83045,67337,03613,28125
36	10000,00000,03350,70230,79911,91730,0	0,00000,00000,14551,91522,83668,51806,64062,5
37	10000,00000,01675,35115,39815,61857,6	0,00000,00000,07275,95761,41834,25903,32031,25
38	10000,00000,00837,67557,69872,72426,9	0,00000,00000,03637,97880,70917,12951,66015,625
39	10000,00000,00418,83778,84927,59087,9	0,00000,00000,01818,98940,35458,56475,83007,8125
40	10000,00000,00209,41889,42461,60262,5	0,00000,00000,00909,49470,17729,28237,91503,90625
41	10000,00000,00104,70944,71230,25311,0	0,00000,00000,00454,74735,08864,64118,95751,95312

Briggs' Log Method (1617)

$$\log(ab) = \log a + \log b \quad \Rightarrow \quad \log a = 2 \log a^{1/2}.$$

Use repeatedly:

$$\log a = 2^k \log a^{1/2^k}.$$

Write $a^{1/2^k} = 1 + x$ and note $\log(1 + x) \approx x$. Briggs worked to base 10 and used

$$\log_{10} a \approx 2^k \cdot \log_{10} e \cdot (a^{1/2^k} - 1).$$

Matrix Logarithm

Take $B = C$ in previous theorem:

$$\log A = \log\big(A^{1/2} \cdot A^{1/2}\big) = 2 \log A^{1/2},$$

since $\arg \lambda(A^{1/2}) \in (-\pi/2, \pi/2)$.

Matrix Logarithm

Take $B = C$ in previous theorem:

$$\log A = \log\left(A^{1/2} \cdot A^{1/2}\right) = 2 \log A^{1/2},$$

since $\arg \lambda(A^{1/2}) \in (-\pi/2, \pi/2)$.

Use Briggs' idea: $\log A = 2^k \log\left(A^{1/2^k}\right).$

Matrix Logarithm

Take $B = C$ in previous theorem:

$$\log A = \log\left(A^{1/2} \cdot A^{1/2}\right) = 2 \log A^{1/2},$$

since $\arg \lambda(A^{1/2}) \in (-\pi/2, \pi/2)$.

Use Briggs' idea: $\log A = 2^k \log\left(A^{1/2^k}\right).$

Kenney & Laub's (1989) **inverse scaling and squaring** method:

- Bring A close to I by repeated square roots.
- Approximate $\log A^{1/2^k}$ using an $[m/m]$ Padé approximant $r_m(x) \approx \log(1 + x)$.
- Rescale to find $\log(A)$.

Options

Apply ISS

- To original A: Cheng, H, Kenney & Laub (2001). Requires square roots of full matrices.
- To triangular Schur factor.
- To diagonal blocks within the Schur–Parlett method.

Options

Apply ISS

- To original A: Cheng, H, Kenney & Laub (2001). Requires square roots of full matrices.
- To triangular Schur factor.
- To diagonal blocks within the Schur–Parlett method.

★ Use fixed Padé degree m.

★ Let m vary optimally with $\|A\|$.

Options

Apply ISS

- To original A: Cheng, H, Kenney & Laub (2001). Requires square roots of full matrices.
- To triangular Schur factor.
- To diagonal blocks within the Schur–Parlett method.

★ Use fixed Padé degree m.

★ Let m vary optimally with $\|A\|$.

▶ MATLAB's `logm` ($m = 8$).

Options

Apply ISS

- To original A: Cheng, H, Kenney & Laub (2001). Requires square roots of full matrices.
- To triangular Schur factor.
- To diagonal blocks within the Schur–Parlett method.

★ Use fixed Padé degree m.

★ Let m vary optimally with $\|A\|$.

▶ MATLAB's `logm` ($m = 8$).

▶ Improved `logm`.

Padé Approximants

$r_{km} = p_{km}/q_{km}$ is a $[k/m]$ Padé approximant of f if p_{km} and q_{km} are polys of degree at most k and m and

$$f(x) - r_{km}(x) = O\big(x^{k+m+1}\big).$$

For $f(x) = \log(1+x)$,

$$r_{11}(x) = \frac{2x}{2+x},$$

$$r_{22}(x) = \frac{6x+3x^2}{6+6x+x^2},$$

$$r_{33}(x) = \frac{60x+60x^2+11x^3}{60+90x+36x^2+3x^3}.$$

Padé Approximants

$r_{km} = p_{km}/q_{km}$ is a $[k/m]$ Padé approximant of f if p_{km} and q_{km} are polys of degree at most k and m and

$$f(x) - r_{km}(x) = O\big(x^{k+m+1}\big).$$

Theorem (Kenney & Laub, 1989)

For $\|X\| < 1$,

$$\|r_{mm}(X) - \log(I + X)\| \le |r_{mm}(-\|X\|) - \log(1 - \|X\|)|.$$

Algorithmic Ingredients

$$\log A = 2^k \log\left(A^{1/2^k}\right) \approx 2^k r_m\left(A^{1/2^k} - I\right).$$

- For given $A^{1/2^k}$, error bound determines min m s.t. r_m suff. accurate.
- Choose k and $m = m(k)$ to minimize overall cost.
- Since $\left(I - A^{1/2^{k+1}}\right)\left(I + A^{1/2^{k+1}}\right) = I - A^{1/2^k}$,

$$\|I - A^{1/2^{k+1}}\| \approx \frac{1}{2}\|I - A^{1/2^k}\|.$$

- Evaluate the **partial fraction** form

$$r_m(x) = \sum_{j=1}^{m} \frac{\alpha_j^{(m)} x}{1 + \beta_j^{(m)} x},$$

where $\alpha_j^{(m)}$ weights and $\beta_j^{(m)}$ Gauss–Legendre nodes.

Schur–Parlett Algorithm

H & Davies (2003), **funm**:

- Compute **Schur decomposition** $A = QTQ^*$.
- Re-order T to block triangular form in which eigenvalues within a block are "**close**" and those of separate blocks are "**well separated**".
- Evaluate $F_{ii} = f(T_{ii})$.
- Solve the **Sylvester equations**

$$T_{ii}F_{ij} - F_{ij}T_{jj} = F_{ii}T_{ij} - T_{ij}F_{jj} + \sum_{k=i+1}^{j-1}(F_{ik}T_{kj} - T_{ik}F_{kj}).$$

- Undo the unitary transformations.

Schur–Parlett Algorithm

H & Davies (2003), **funm**:

- Compute **Schur decomposition** $A = QTQ^*$.
- Re-order T to block triangular form in which eigenvalues within a block are "**close**" and those of separate blocks are "**well separated**".
- Evaluate $F_{ii} = \log(T_{ii})$.
- Solve the **Sylvester equations**

$$T_{ii}F_{ij} - F_{ij}T_{jj} = F_{ii}T_{ij} - T_{ij}F_{jj} + \sum_{k=i+1}^{j-1}(F_{ik}T_{kj} - T_{ik}F_{kj}).$$

- Undo the unitary transformations.

Function of 2×2 Block

$$f\left(\begin{bmatrix} \lambda_1 & t_{12} \\ 0 & \lambda_2 \end{bmatrix}\right) = \begin{bmatrix} f(\lambda_1) & t_{12}\dfrac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1} \\ 0 & f(\lambda_2) \end{bmatrix}.$$

- **Inaccurate** if $\lambda_1 \approx \lambda_2$.
- Need a better way to compute the divided difference $f[\lambda_2, \lambda_1]$.

Log of 2×2 Block

$$\begin{aligned}\log\lambda_2 - \log\lambda_1 &= \log\left(\frac{\lambda_2}{\lambda_1}\right) + 2\pi i\,\mathcal{U}(\log\lambda_2 - \log\lambda_1)\\ &= \log\left(\frac{1+z}{1-z}\right) + 2\pi i\,\mathcal{U}(\log\lambda_2 - \log\lambda_1),\end{aligned}$$

where U = unwinding number, $z = (\lambda_2 - \lambda_1)/(\lambda_2 + \lambda_1)$.

$$\mathrm{atanh}(z) := \frac{1}{2}\log\left(\frac{1+z}{1-z}\right),$$

Log of 2×2 Block

$$\log \lambda_2 - \log \lambda_1 = \log\left(\frac{\lambda_2}{\lambda_1}\right) + 2\pi i\,\mathcal{U}(\log \lambda_2 - \log \lambda_1)$$
$$= \log\left(\frac{1+z}{1-z}\right) + 2\pi i\,\mathcal{U}(\log \lambda_2 - \log \lambda_1),$$

where $\mathcal{U}$ = unwinding number, $z = (\lambda_2 - \lambda_1)/(\lambda_2 + \lambda_1)$.

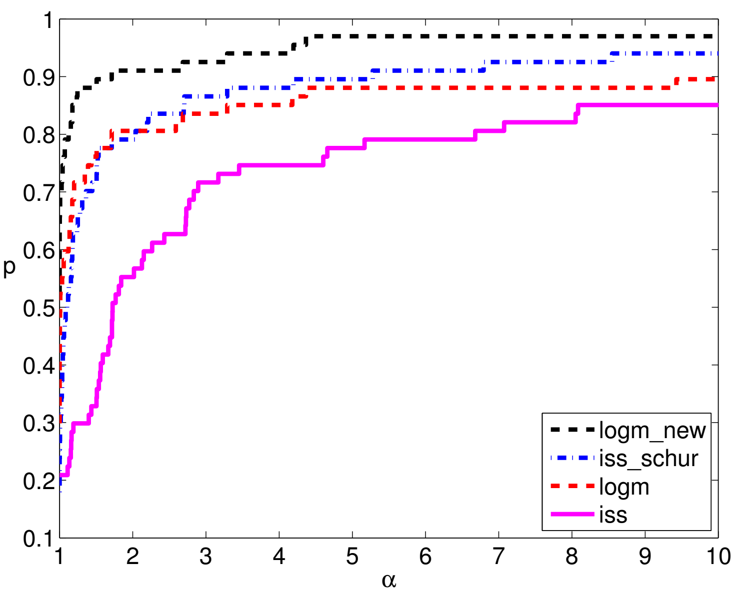
$$\mathrm{atanh}(z) := \frac{1}{2}\log\left(\frac{1+z}{1-z}\right),$$

$$f_{12} = t_{12}\,\frac{2\,\mathrm{atanh}(z) + 2\pi i\,\mathcal{U}(\log \lambda_2 - \log \lambda_1)}{\lambda_2 - \lambda_1}.$$

Numerical Experiment

- 67 test matrices, dimension 2–10.
- Evaluated $\|\widehat{X} - \log(A)\|_F / \|\log(A)\|_F$.
- Notation:
  - `logm`: MATLAB 7.4 (R2007a).
  - `logm_new`: New version of `logm`.
  - `iss_schur`: Schur decomp then ISS.
  - `iss`: ISS on full A.
- $\mathrm{cond}(A) = \lim_{\epsilon\to 0} \max_{\|E\|_2 \le \epsilon \|A\|_2} \dfrac{\|\log(A+E) - \log(A)\|_2}{\epsilon \|\log(A)\|_2}.$

Performance Profile



$\log(A)b$

Hale, H & Trefethen, **Computing A^{α}, $\log(A)$ and Related Matrix Functions by Contour Integrals, 2007.**

- New methods for $f(A)b$ where f has singularities in $(-\infty, 0]$ and A is a matrix with ei'vals on or near $(0, \infty)$.
- Contour integrals + conformal map + repeated trapezium rule $\Rightarrow$ **geometric convergence**.

In Conclusion

- Matrix logarithm and square root are archetypal examples of multivalued matrix functions **(LambertW: Corless, Ding, H & Jeffrey, 2007)**.
- Able to **classify** all logs.
- Non-primary logs of interest, but little is known.
- Improvements to inverse scaling and squaring alg and to `logm`.
- Exploiting structure?

References I

- R. J. Bradford, R. M. Corless, J. H. Davenport, D. J. Jeffrey, and S. M. Watt.
  Reasoning about the elementary functions of complex analysis.
  Annals of Mathematics and Artificial Intelligence, 36:303–318, 2002.

- S. H. Cheng, N. J. Higham, C. S. Kenney, and A. J. Laub.
  Approximating the logarithm of a matrix to specified accuracy.
  SIAM J. Matrix Anal. Appl., 22(4):1112–1125, 2001.

References II

- R. M. Corless, H. Ding, N. J. Higham, and D. J. Jeffrey.
  The solution of $S\exp(S) = A$ is not always the Lambert W function of A.
  In *ISSAC '07: Proceedings of the 2007 International Symposium on Symbolic and Algebraic Computation*, pages 116–121, New York, 2007. ACM Press.

- R. M. Corless and D. J. Jeffrey.
  The unwinding number.
  ACM SIGSAM Bulletin, 30(2):28–35, June 1996.

References III

E. D. Dolan and J. J. Moré.
Benchmarking optimization software with performance profiles.
Math. Programming, 91:201–213, 2002.

F. R. Gantmacher.
The Theory of Matrices, volume one.
Chelsea, New York, 1959.

References IV

N. Hale, N. J. Higham, and L. N. Trefethen.
Computing A^α, $\log(A)$ and related matrix functions by contour integrals.
MIMS EPrint 2007.xxx, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Aug. 2007.

N. J. Higham.
Evaluating Padé approximants of the matrix logarithm.
SIAM J. Matrix Anal. Appl., 22(4):1126–1135, 2001.

References V

N. J. Higham.
Functions of Matrices: Theory and Computation.
2008.
Book to appear.