The Matrix Logarithm: from Theory to Computation

Nick Higham
School of Mathematics
The University of Manchester

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

6th European Congress of Mathematics, July 2012

Matrix Logarithm

Definition

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

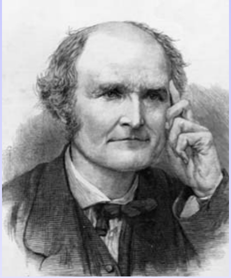
- Implicit definition.
- Properties, classification?

Outline

Cayley and Sylvester

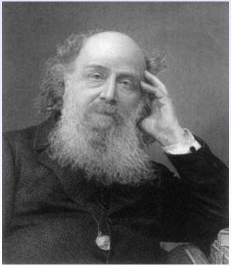
Matrix algebra developed by **Arthur Cayley**, FRS (1821–1895) in *Memoir on the Theory of Matrices (1858)*.

- Cayley considered matrix square roots.



Term **"matrix"** coined in 1850 by **James Joseph Sylvester**, FRS (1814–1897).

- Gave (1883) first definition of $f(A)$ for general f.



Multiplicity of Definitions

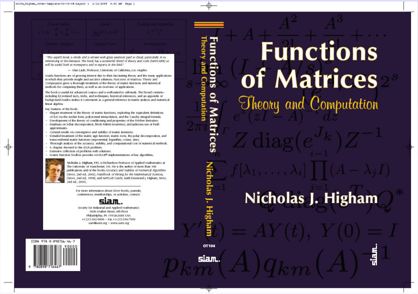
There have been proposed in the literature since 1880 ***eight distinct definitions*** *of a matric function, by Weyr, Sylvester and Buchheim, Giorgi, Cartan, Fantappiè, Cipolla, Schwerdtfeger and Richter.*

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

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 (1955)



Jordan Canonical Form

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

Definition

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

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

Primary and Nonprimary Logarithms

$A = \operatorname{diag}(1, 1, e, e)$.

$$\textbf{Primary}: \quad \log(A) = \operatorname{diag}(0, 0, 1, 1).$$

$$\textbf{Nonprimary}: \quad \log(A) = \operatorname{diag}(0, 2\pi i, 1, 1).$$

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 $\lambda(A)$.

Mercator's Series

By integrating $(1+t)^{-1} = 1 - t + t^2 - t^3 + \cdots$ between 0 and x we obtain Mercator's series (1668),

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots, \quad |x| < 1.$$

For $A \in \mathbb{C}^{n\times n}$,

$$\log(I+A) = A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} + \cdots, \quad \rho(A) < 1.$$

Composite Functions

Theorem

$$f(t) = g(h(t)) \;\Rightarrow\; f(A) = g(h(A)).$$

Corollary

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

Composite Functions

Theorem

$f(t) = g(h(t)) \;\Rightarrow\; f(A) = g(h(A)).$

Corollary

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

What about $\log(\exp(A)) = A$?

Matrix unwinding number

$$\mathcal{U}(A) = \frac{A - \log(\exp(A))}{2\pi i}.$$

Outline

1. Definition and Properties

2. Applications

3. Theory

4. Computing the Matrix Logarithm and its Fréchet derivative

Toolbox of Matrix Functions

$$\frac{d^2y}{dt^2} + Ay = 0, \qquad y(0) = y_0, \quad y'(0) = y'_0$$

has solution

$$y(t) = \cos(\sqrt{A}\,t)y_0 + \left(\sqrt{A}\right)^{-1} \sin(\sqrt{A}\,t)y'_0.$$

Toolbox of Matrix Functions

$$\frac{d^2y}{dt^2} + Ay = 0, \qquad y(0) = y_0, \quad y'(0) = y'_0$$

has solution

$$y(t) = \cos(\sqrt{A}\,t)y_0 + \left(\sqrt{A}\right)^{-1}\sin(\sqrt{A}\,t)y'_0.$$

But

$$\begin{bmatrix} y' \\ y \end{bmatrix} = \exp\left(\begin{bmatrix} 0 & -tA \\ tI_n & 0 \end{bmatrix}\right)\begin{bmatrix} y'_0 \\ y_0 \end{bmatrix}.$$

Toolbox of Matrix Functions

$$\frac{d^2y}{dt^2} + Ay = 0, \qquad y(0) = y_0, \quad y'(0) = y_0'$$

has solution

$$y(t) = \cos(\sqrt{A}\,t)y_0 + \left(\sqrt{A}\right)^{-1}\sin(\sqrt{A}\,t)y_0'.$$

But

$$\begin{bmatrix} y' \\ y \end{bmatrix} = \exp\left(\begin{bmatrix} 0 & -tA \\ tI_n & 0 \end{bmatrix}\right)\begin{bmatrix} y_0' \\ y_0 \end{bmatrix}.$$

- In software want to be able evaluate interesting f at matrix args as well as scalar args.
- MATLAB has **`expm`**, **`logm`**, **`sqrtm`**, **`funm`**.

Application: Control Theory

Convert **continuous-time system**

$$\frac{dx}{dt} = Fx(t) + Gu(t),$$
$$y = Hx(t) + Ju(t),$$

to **discrete-time state-space system**

$$x_{k+1} = Ax_k + Bu_k,$$
$$y_k = Hx_k + Ju_k.$$

Have

$$A = e^{F\tau}, \qquad B = \left(\int_0^\tau e^{Ft}dt\right)G,$$

where τ is the sampling period.
MATLAB Control System Toolbox: `c2d` and `d2c`.

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**:

$$S^T J S = J = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix}.$$

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

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

Markov Models

- Time-homogeneous continuous-time Markov process with transition probability matrix $P(t) \in \mathbb{R}^{n\times n}$.
- **Transition intensity matrix** Q: $q_{ij} \geq 0$ $(i \neq j)$, $\sum_{j=1}^{n} q_{ij} = 0$, $\boxed{P(t) = e^{Qt}.}$

For *discrete-time* Markov processes:

Embeddability problem

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

Markov Models (1)—Example

With $x = -e^{-2\sqrt{3}\pi} \approx -1.9 \times 10^{-5}$,

$$P = \frac{1}{3}\begin{bmatrix} 1+2x & 1-x & 1-x \\ 1-x & 1+2x & 1-x \\ 1-x & 1-x & 1+2x \end{bmatrix}.$$

- P diagonalizable, $\Lambda(P) = \{1, x, x\}$.
- Every primary log complex (can't have complex conjugate ei'vals).
- Yet a generator is the non-primary log

$$Q = 2\sqrt{3}\pi \begin{bmatrix} -2/3 & 1/2 & 1/6 \\ 1/6 & -2/3 & 1/2 \\ 1/2 & 1/6 & -2/3 \end{bmatrix}.$$

Markov Models (2)

- Suppose $P \equiv P(1)$ has a generator $Q = \log P$.
  Then $P(t)$ at other times is $P(t) = \exp(Qt)$.
  E.g., if P transition matrix for 1 year,
  $P(1/12) = e^{\frac{1}{12}\log P} \equiv P^{1/12}$ is matrix for 1 month.
- **Problem**: $\log P$, $P^{1/k}$ may have wrong sign patterns $\Rightarrow$ "regularize".

HIV to Aids Transition

- Estimated 6-month transition matrix.
- Four AIDS-free states and 1 AIDS state.
- 2077 observations (Charitos et al., 2008).

$$P = \begin{bmatrix} 0.8149 & 0.0738 & 0.0586 & 0.0407 & 0.0120 \\ 0.5622 & 0.1752 & 0.1314 & 0.1169 & 0.0143 \\ 0.3606 & 0.1860 & 0.1521 & 0.2198 & 0.0815 \\ 0.1676 & 0.0636 & 0.1444 & 0.4652 & 0.1592 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Want to estimate the **1-month transition matrix**.

$$\Lambda(P) = \{1, 0.9644, 0.4980, 0.1493, -0.0043\}.$$

N. J. Higham and L. Lin.
On *p*th roots of stochastic matrices, LAA, 2011.

Outline

Logs of $A = I_3$

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

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

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

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

All Solutions of $e^X = A$

Theorem (Gantmacher, 1959)

$A \in \mathbb{C}^{n\times n}$ 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 $A = I_3$ (again)

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

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

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

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

Square Roots of Rotations

$$G(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

$G(\theta/2)$ is the natural square root of $G(\theta)$.

For $\theta = \pi$,

$$G(\pi) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad G(\pi/2) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

$G(\pi/2)$ is a **nonprimary** square root.

Principal Logarithm 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

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

Principal Logarithm 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

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

Principal pth root

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

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

Outline

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 *Numeri continue Medij inter Denarium & Unitatem.*	E *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	10011,24941,39987,98758,85395,51805	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

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

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

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

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

Inverse Scaling and Squaring Method

Take $B = C$ in previous theorem:

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

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

Inverse Scaling and Squaring Method

Take $B = C$ in previous theorem:

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

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

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

Inverse Scaling and Squaring Method

Take $B = C$ in previous theorem:

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

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

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

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

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

Choice of Parameters s, m

Must have $\|I - A^{1/2^s}\| < 1$.

- Larger Padé degree m means smaller s.

Let $h_{2m+1}(X) = e^{r_m(X)} - X - I$.
Assume $\rho(r_m(X)) < \pi$, so $\log(e^{r_m(X)}) = r_m(X)$. Then

$$r_m(X) = \log(I + X + h_{2m+1}(X)) =: \log(I + X + \Delta X),$$

where

$$h_{2m+1}(X) = \sum_{k=2m+1}^{\infty} c_k X^k.$$

Bounding the Backward Error

Want to bound norm of $h_{2m+1}(X) = \sum_{k=2m+1}^{\infty} c_k X^k$.

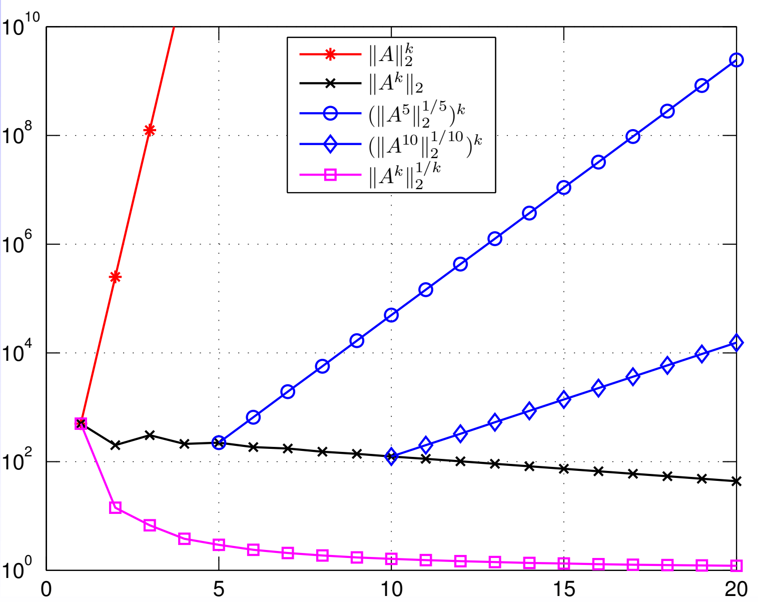
- Non-normality implies $\rho(A) \ll \|A\|$.
- Note that

$$\rho(A) \le \|A^k\|^{1/k} \le \|A\|, \qquad k = 1 : \infty.$$

  and $\lim_{k\to\infty} \|A^k\|^{1/k} = \rho(A)$.

- Use $\|A^k\|^{1/k}$ instead of $\|A\|$ in the truncation bounds.

$$A = \begin{bmatrix} 0.9 & 500 \\ 0 & -0.5 \end{bmatrix}.$$



Algorithm of Al-Mohy & H (2011)

- Truncation bounds use $\|A^k\|^{1/k}$ rather than $\|A\|$, leading to major benefits in speed and accuracy.
  Matrix norms not such a blunt tool!
- Use *estimates* of $\|A^k\|$ (alg of H & Tisseur (2000)).
- Choose s and m to achieve double precision backward error at minimal cost.
- Initial Schur decomposition: $A = QTQ^*$.
- Directly and accurately compute certain elements of $T^{1/2^s} - I$ and $\log(T)$. Use

$$a^{1/2^s} - 1 = \frac{a-1}{\prod_{i=1}^{s}(1 + a^{1/2^i})}.$$

Frechét Derivative of Logarithm

$$f(A+E) - f(A) - L(A,E) = o(\|E\|).$$

- Integral formula

$$L(A,E) = \int_0^1 \left(t(A-I)+I\right)^{-1} E \left(t(A-I)+I\right)^{-1} dt.$$

- Method based on

$$f\left(\begin{bmatrix} X & E \\ 0 & X \end{bmatrix}\right) = \begin{bmatrix} f(X) & L(X,E) \\ 0 & f(X) \end{bmatrix}.$$

- Kenney & Laub (1998): Kronecker–Sylvester alg, Padé of $\tanh(x)/x$. Requires complex arithmetic.

Algorithm of Al-Mohy, H & Relton (2012)

Fréchet differentiate the ISS algorithm!

$E_0 = E$
for $i = 1\colon s$
    Compute $A^{1/2^i}$.
    Solve the Sylvester eqn $A^{1/2^i} E_i + E_i A^{1/2^i} = E_{i-1}$.
end
$\log(A) \approx 2^s r_m(A^{1/2^s} - I)$
$L_{\log}(A, E) \approx 2^s L_{r_m}(A^{1/2^s} - I, E_s)$

Backward Error Result

$$r_m(X) = \log(I + X + \Delta X),$$
$$L_{r_m}(X, E) = L_{\log}(I + X + \Delta X, E + \Delta E).$$

Conclusions & Future Directions

- Log appears in a growing number of applications.
- Have good algorithms for $\log(A)$, $L_{\log}(A)$ and estimating the condition number.
- If A is real can work entirely in real arithmetic.

- Conditioning of $f(A)$.
- Non-primary functions.
- Functions of structured matrices.

References I

A. H. Al-Mohy and N. J. Higham.
A new scaling and squaring algorithm for the matrix exponential.
SIAM J. Matrix Anal. Appl., 31(3):970–989, 2009.

A. H. Al-Mohy and N. J. Higham.
Improved inverse scaling and squaring algorithms for the matrix logarithm.
MIMS EPrint 2011.83, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Oct. 2011.
18 pp.
Revised April 2012. To appear in SIAM J. Sci. Comput.

References II

T. Charitos, P. R. de Waal, and L. C. van der Gaag.
Computing short-interval transition matrices of a discrete-time Markov chain from partially observed data.
Statistics in Medicine, 27:905–921, 2008.

W. F. Harris.
The average eye.
Opthal. Physiol. Opt., 24:580–585, 2005.

N. J. Higham.
The Matrix Function Toolbox.
http://www.ma.man.ac.uk/~higham/mftoolbox.

References III

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

N. J. Higham.
Functions of Matrices: Theory and Computation.
Society for Industrial and Applied Mathematics,
Philadelphia, PA, USA, 2008.
ISBN 978-0-898716-46-7.
xx+425 pp.

N. J. Higham and A. H. Al-Mohy.
Computing matrix functions.
Acta Numerica, 19:159–208, 2010.

References IV

N. J. Higham and L. Lin.
On *p*th roots of stochastic matrices.
Linear Algebra Appl., 435(3):448–463, 2011.

C. S. Kenney and A. J. Laub.
Condition estimates for matrix functions.
SIAM J. Matrix Anal. Appl., 10(2):191–209, 1989.

C. S. Kenney and A. J. Laub.
A Schur–Fréchet algorithm for computing the logarithm and exponential of a matrix.
SIAM J. Matrix Anal. Appl., 19(3):640–663, 1998.