Computation of the matrix logarithm using the double exponential formula

F. Tatsuoka*, T. Sogabe*, Y. Miyatake†, S.-L. Zhang*

*Nagoya Univ. (Japan), †Osaka Univ. (Japan)

NASCA 2018 (July 5, 2018) @ Kalamata, Greece

Outline

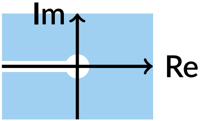
- **Matrix logarithm $\log(A)$**
  - Definition
  - Applications
  - Computational methods ◀ Numerical quadrature
- **Gauss–Legendre quadrature for $\log(A)$**
- **The DE formula for $\log(A)$ (our research)**
- **Numerical experiments**
- **Conclusion**

Outline

- **Matrix logarithm $\log(A)$**
  - Definition
  - Applications
  - Computational methods ◀ Numerical quadrature
- **Gauss–Legendre quadrature for $\log(A)$**
- **The DE formula for $\log(A)$ (our research)**
- **Numerical experiments**
- **Conclusion**

Definition of log(A)

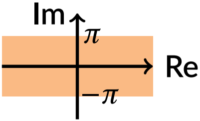
Assumption. $A \in \mathbb{C}^{n\times n}$ does not have real eigenvalues less than or equal to 0.



Definition (Principal matrix logarithm).

$\log(A)$ is a matrix $X \in \mathbb{C}^{n\times n}$ s.t.

- $\exp(X) = A \quad \left(\exp(X) := I + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots\right)$
- All eigenvalues of X lie in $\{z \in \mathbb{C} : |\mathrm{Im}(z)| < \pi\}$



Properties. (See, e.g., [Higham, 2008])

- For any A, $\log(A)$ uniquely exists.
- $\log(A) = (A - I)\int_0^1 [t(A-I)+I]^{-1}\,\mathrm{d}t$

Applications & Computational methods

Applications

- Quantum communication (Holevo information)
- Lattice QCD (noisy Monte Carlo algorithm)
- Medical imaging (Image registration)

Computational methods

- Inverse scaling and squaring (e.g., [Al-Mohy, Higham, 2012])
  - Appropriate for dense A
- **Numerical quadrature**
  - Appropriate for large and sparse A

Outline

- **Matrix logarithm log(A)**
  - Definition
  - Applications
  - Computational methods ◀ Numerical quadrature
- **Gauss–Legendre quadrature for log(A)**
- **The DE formula for log(A) (our research)**
- **Numerical experiments**
- **Conclusion**

Gauss–Legendre quadrature for log(A)

$$\log(A) = (A - I)\int_0^1 \left[t(A-I)+I\right]^{-1} \mathrm{d}t$$

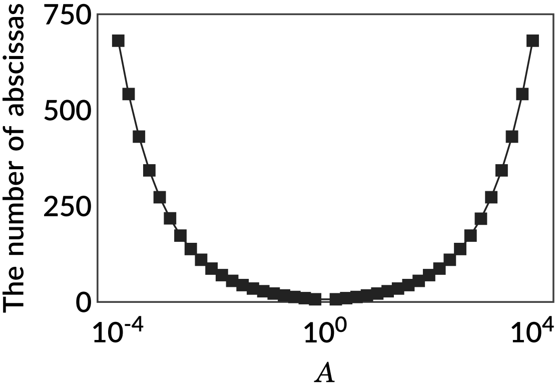
When $A \approx I$, Gauss–Legendre quadrature may be the best choice.

- Gauss–Legendre quadrature for $\log(A)$ coincides with the Padé approximation if the spectrum radius of $A - I$ is less than 1 [Dieci et al., 1996].

Problem: When $A \not\approx I$, the convergence of Gauss–Legendre quadrature become slow (even if A is a scalar).

Numerical example

For $A \in \mathbb{R}^{1\times 1}$, we counted the minimum number of abscissas (integral nodes) of Gauss–Legendre quadrature for $\log(A)$ when the relative error is less than 10^{-12}.

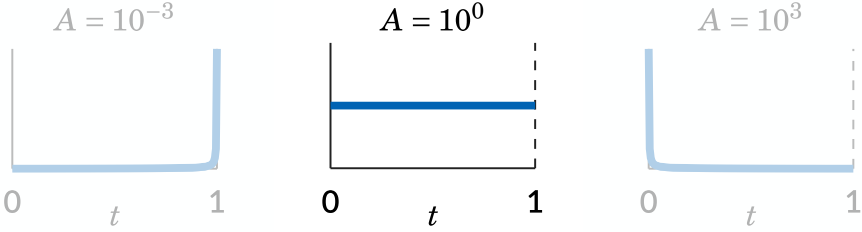


The convergence is slow when $A \approx 0$ or $A \gg 1$.

Motivation

Slow convergence may be due to the near-singularity (in function sense) of the integrand.

Shapes of the integrand $F(t) = [t(A - I) + I]^{-1}$



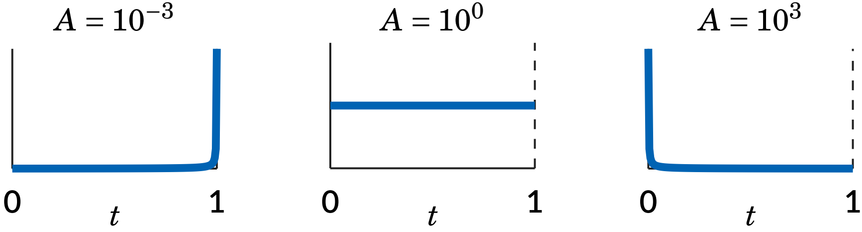
Motivation.

Another approach that can deal with near-singularity is needed for efficient computation.

Motivation

Slow convergence may be due to the near-singularity (in function sense) of the integrand.

Shapes of the integrand $F(t) = [t(A - I) + I]^{-1}$

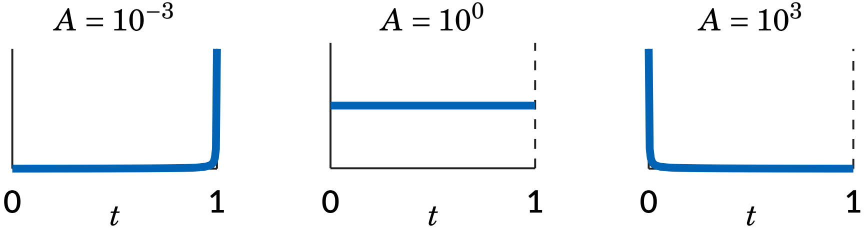


Motivation.
Another approach that can deal with near-singularity is needed for efficient computation.

Motivation

Slow convergence may be due to the near-singularity (in function sense) of the integrand.

Shapes of the integrand $F(t) = [t(A - I) + I]^{-1}$



Motivation.
Another approach that can deal with near-singularity is needed for efficient computation.

Outline

- **Matrix logarithm log(A)**
  - Definition
  - Applications
  - Computational methods ◀ Numerical quadrature
- **Gauss–Legendre quadrature for log(A)**
- **The DE formula for log(A) (our research)**
- **Numerical experiments**
- **Conclusion**

Our research

Idea. **The double exponential (DE) formula** [Takahashi, Mori, 1974]

- works well even if the integrand is singular at the endpoint.
- needs **parameter tuning** for practical computation.

Contribution.

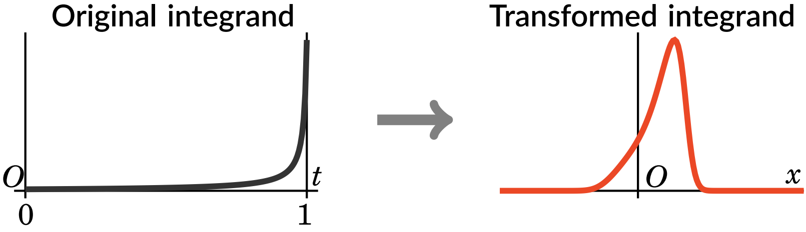
Proposing an algorithm **that determines the parameters theoretically** for practical use.

Procedure of the DE formula

i. Variable transformation by the specific function.

- The transformed interval is $(-\infty, \infty)$.
- The transformed integrand decays double exponentially.

Example.



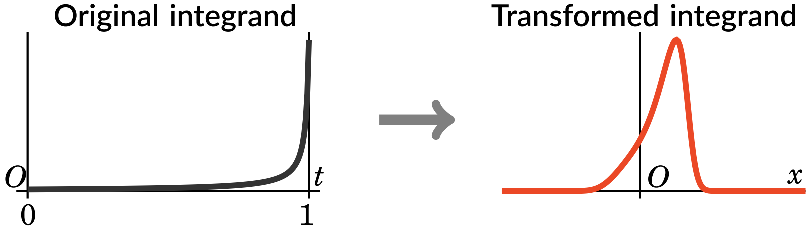
ii. Truncating the interval into suitable finite interval $[l, r]$.

iii. Approximation using trapezoidal rule on $[l, r]$

Procedure of the DE formula

i. Variable transformation by the specific function.
- The transformed interval is $(-\infty, \infty)$.
- The transformed integrand decays double exponentially.

Example.



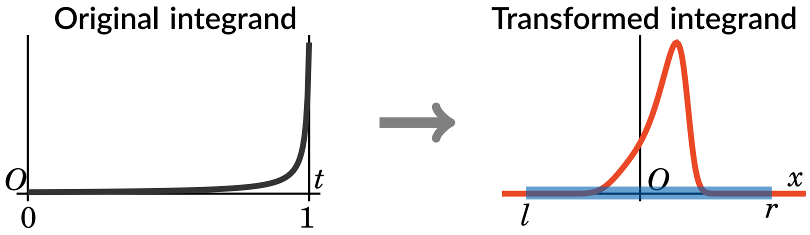
ii. Truncating the interval into suitable finite interval $[l, r]$.

iii. Approximation using trapezoidal rule on $[l, r]$

Procedure of the DE formula

i. Variable transformation by the specific function.
- The transformed interval is $(-\infty, \infty)$.
- The transformed integrand **decays double exponentially**.

Example.



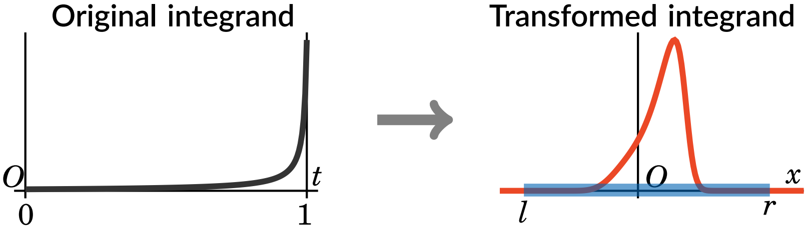
ii. **Truncating the interval into suitable finite interval $[l, r]$.**

iii. Approximation using trapezoidal rule on $[l, r]$

Procedure of the DE formula

i. Variable transformation by the specific function.
- The transformed interval is $(-\infty, \infty)$.
- The transformed integrand decays double exponentially.

Example.



ii. Truncating the interval into suitable finite interval $[l, r]$.

iii. Approximation using trapezoidal rule on $[l, r]$

Applying the DE formula to log(A)

i. Applying the DE transformation:

$$\log(A) = (A-I)\int_0^1 \left[t(A-I)+I\right]^{-1} \mathrm{d}t$$

$$\downarrow\ t(x) = \frac{\tanh(\sinh(x))+1}{2}$$

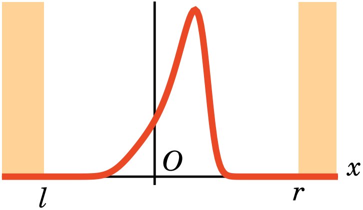
$$= (A-I)\int_{-\infty}^{\infty} F_{\mathrm{DE}}(x)\,\mathrm{d}x$$

$$\left(F_{\mathrm{DE}}(x) := \frac{\cosh(x)}{2\cosh^2(\sinh(x))}\left[t(x)(A-I)+I\right]^{-1}\right)$$

ii. A way of truncating the interval based on truncation error estimation.

Idea for the step ii

$$\underbrace{\int_{-\infty}^{\infty} F_{\mathrm{DE}}(x)\,\mathrm{d}x - \int_{l}^{r} F_{\mathrm{DE}}(x)\,\mathrm{d}x}_{\text{Truncation error}} = \int_{-\infty}^{l} F_{\mathrm{DE}}(x)\,\mathrm{d}x + \int_{r}^{\infty} F_{\mathrm{DE}}(x)\,\mathrm{d}x$$



Error estimation

Representing upper bound of the norm of the RHS using l and r.

Setting a finite interval

Setting l and r such that the norm of the RHS is smaller than a given tolerance.

Result of error estimation

Lemma.

An upper bound of relative truncation error of the DE formula for $\log(A)$ can be represented as

$$\frac{\left\|\log(A) - (A-I)\int_l^r F_{\mathrm{DE}}(x)\,\mathrm{d}x\right\|}{\|\log(A)\|} \leq \boxed{c_l}\left[\tanh(\sinh(l)) + 1\right] + \boxed{c_r}\left[1 - \tanh(\sinh(r))\right],$$

where $\|\cdot\|$ is a consistent matrix norm, and c_l, c_r are specific positive constants depending on A, if l, r satisfy

$\boxed{\text{some conditions}}$.

Lemma.

An upper bound of relative truncation error of the DE formula for $\log(A)$ can be represented as

$$\frac{\left\|\log(A)-(A-I)\int_l^r F_{\mathrm{DE}}(x)\,\mathrm{d}x\right\|}{\|\log(A)\|} \le \underbrace{\frac{3\|A-I\|}{4\left|\log\left(\rho(A)\right)\right|}}_{c_l}\left[\tanh(\sinh(l))+1\right]+\underbrace{\frac{3\|A-I\|\|A^{-1}\|}{4\left|\log\left(\rho(A)\right)\right|}}_{c_r}\left[1-\tanh(\sinh(r))\right]$$

($\rho(A)$ is spectral radius)

where $\|\cdot\|$ is a consistent matrix norm, if l, r satisfy

$$\tanh(\sinh(l)) \le \frac{1}{\|A-I\|}-1, \quad \tanh(\sinh(r)) \ge \frac{2\|A^{-1}\|-1}{2\|A^{-1}\|+1}.$$

Setting the interval

Proposition

For a given tolerance $\varepsilon > 0$, we can set a finite interval such that the truncation error is less than or equal to ε:

$$\frac{\left\| \log(A) - (A - I)\int_l^r F_{\mathrm{DE}}(x)\,\mathrm{d}x \right\|}{\|\log(A)\|} \leq \varepsilon,$$

by choosing l, r as

$$l = l(\varepsilon) := \min\left\{ \mathrm{arsinh}(\mathrm{artanh}(\tfrac{1}{2c_l}\varepsilon) - 1),\ \ l_{\max} \right\},$$
$$r = r(\varepsilon) := \max\left\{ \mathrm{arsinh}(1 - \mathrm{artanh}(\tfrac{1}{2c_r}\varepsilon)),\ \ r_{\min} \right\},$$

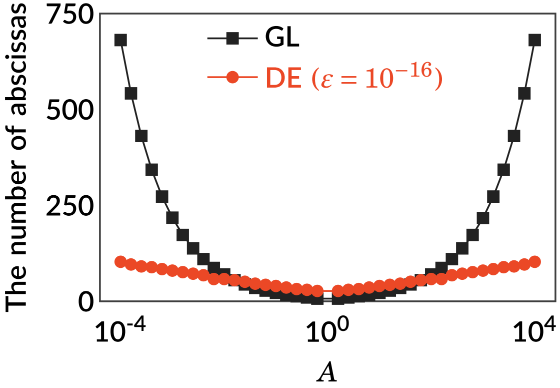
where c_l, c_r, $l_{\max}$, $r_{\min}$ are some specific constants depending on A.

The algorithm is ready!

Outline

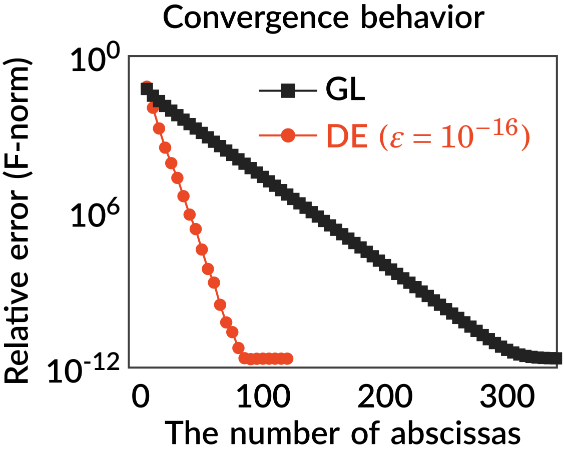
Example 1 (Scalar)

For $A \in \mathbb{R}^{1\times 1}$, we counted the number of minimum abscissas when the relative error is less than 10^{-12}.



The DE formula converged faster than Gauss–Legendre quadrature when $A \approx 0$ or $A \gg 1$

Example 2 (SPD Matrix)



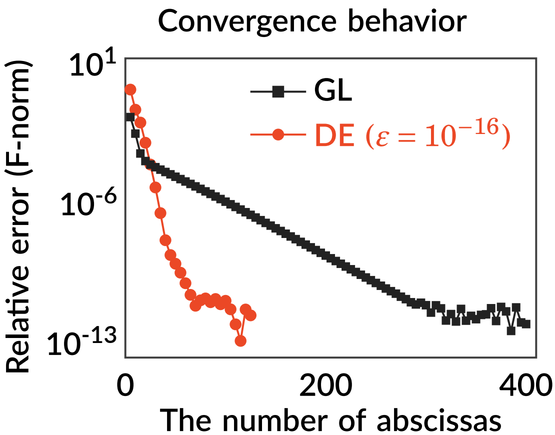
Test matrix A (bcsskt03)

- SPD
- size 112
- $\kappa_2(A)$ 6.7×10^6

- The "exact" solution was computed using arbitrary precision computation.

The convergence of the DE formula is faster than that of Gauss–Legendre quadrature.

Example 3 (Unsymmetric Matrix)



Test matrix A
(Flank matrix)

- Unsymmetric
- size 10
- $\kappa_2(A)$ 2.85×10^7

- The "exact" solution was computed using arbitrary precision computation.

The convergence of the DE formula is faster than that of Gauss–Legendre quadrature.

- **Matrix logarithm log(A)**
  - Definition
  - Applications
  - Computational methods ◀ Numerical quadrature
- **Gauss–Legendre quadrature for log(A)**
- **The DE formula for log(A) (our research)**
- **Numerical experiments**
- **Conclusion**

Conclusion

Summary.

- When $A \not\approx I$, the DE formula seems to be promising for numerical integration of $\log(A)$.
- For practical use, we proposed an algorithm based on the truncation error estimate.
- From results of examples, the DE formula seems to be efficient when $A \not\approx I$.

Future work.

- Error analysis (convergence speed) of the DE formula for $\log(A)$.

Reference

- A. H. Al-Mohy, N. J. Higham, Improved inverse scaling and squaring algorithms for the matrix logarithm, SIAM J. Sci. Comput. 34 (2012), pp. C153–C169.
- L. Dieci, B. Morini, A. Papini, Computational techniques for real logarithms of matrices, SIAM J. Matrix Anal. Appl. 17 (1996), pp. 570–593.
- N. J. Higham, Functions of Matrices: Theory and computation. SIAM, Philadelphia, 2008.
- H. Takahashi, M. Mori, Double exponential formulas for numerical integration, Publ. RIMS Kyoto Univ. 9 (1974), pp. 721–741