

# 行列対数関数のための二重指数関数型公式の収束率について

立岡文理, 曾我部知広, 剣持智哉, 張紹良

名古屋大学 大学院工学研究科 応用物理学専攻

## On the convergence rate of the double exponential formula for the computation of the matrix logarithm

Fuminori Tatsuoka, Tomohiro Sogabe, Tomoya Kemmochi, Shao-Liang Zhang

Department of Applied Physics, Graduate School of Engineering, Nagoya University

E-mail: f-tatsuoka@na.nuap.nagoya-u.ac.jp

### 1 はじめに

正方行列  $A \in \mathbb{C}^{n \times n}$  に対して, 方程式

$$\exp(X) \left( := I + X + \frac{1}{2!}X^2 + \frac{1}{3!}X^3 + \cdots \right) = A$$

を満たす行列  $X \in \mathbb{C}^{n \times n}$  を  $A$  の対数と呼ぶ.  $A$  が正則ならば対数は無限個存在することが知られており [12, p. 269], 特に対数の全ての固有値が  $\{z \in \mathbb{C}: |\operatorname{Im}(z)| < \pi\}$  に属するとき, その対数を主対数と呼び  $\log(A)$  と表記する. なお,  $A$  の全ての固有値が  $\{z \in \mathbb{C}: z \notin (-\infty, 0]\}$  に属するならば,  $\log(A)$  は一意に存在する [12, Thm. 1.31]. 本稿では, これ以降  $\log(A)$  が一意に存在するとし, また単に対数と述べたときでも主対数を指すものとする. 行列対数関数の応用としては機械学習 [2, 6, 11] や量子情報科学 [9] などがある.

行列対数関数の計算手法としては Inverse scaling and squaring アルゴリズム [1, 7] や反復法 [3], 数値積分に基づく手法 (これ以降数値積分法と略す) [4, 14] があるが, 本稿では以下の 2 点の長所を持つ数値積分法に着目する. 1 点目は, 数値積分法は行列対数関数-ベクトル積  $\log(A)\mathbf{b}$  ( $\mathbf{b} \in \mathbb{C}^n$ ) の計算に対応できる点である. 応用上, 大規模な  $A$  に対して  $\log(A)$  そのものではなく  $\log(A)\mathbf{b}$  のみを必要とするときがある. このような場合, Krylov 部分空間などの低次元空間への射影と  $\log(A)$  の計算手法とを組み合わせる方法 ([12, pp. 301–306] 等) が考えられるが, 数値積分法は  $\log(A)\mathbf{b}$  の計算をシフト線形方程式の計算に帰着させるため, 線形方程式さえ計算できれば簡単に対応できる. 2 点目はアルゴリズムの階層的並列性により並列計算に向くことである.

---

本稿は研究集会 (諸科学分野を結ぶ基礎学問としての数値解析学) における講演「数値積分に基づく行列対数関数の計算について」に対する講義録である.

数値積分を行うための  $\log(A)$  の積分表示として、実軸上の積分表示

$$\begin{aligned}\log(A) &= (A - I) \int_0^1 [u(A - I) + I]^{-1} du \\ &= (A - I) \int_{-1}^1 [(1+t)A + (1-t)I]^{-1} dt \quad \left(u(t) = \frac{t+1}{2}\right)\end{aligned}\quad (1)$$

が知られており ([12, Thm. 11.1] 等), 本稿では積分表示 (1) を採用する<sup>\*1</sup>. 積分表示 (1) に対する求積法としては Gauss–Legendre (GL) 求積や二重指数関数型 (DE) 公式が考えられる.

$A$  が正定値エルミート行列の場合は GL 求積の収束率が解析されており [8]<sup>\*2</sup>, 実験的には  $A$  の条件数が小さいときは GL 求積が速く収束し, 大きいときは DE 公式が速く収束する [14]. しかし, DE 公式の収束率は具体的には明らかになっていない.

そこで, 本稿では正定値エルミート行列  $A$  に対する DE 公式の収束率を解析し, GL 求積と比較して計算の際の求積法の選び方の指針を述べる. ここでは, 収束率の解析を文献 [8] の解析手法に沿って, すなわち,  $\log(A)$  に対する DE 公式の誤差を  $A$  の固有値の対数に対する DE 公式の誤差に帰着させて解析を行う.

本稿では次節において正定値エルミート行列の対数に対する DE 公式の収束率を求め, GL 求積との比較を行う. 3 節で数値例を示し, 4 節でまとめと今後の展望を述べる.

**表記について.** これ以降,  $\mathcal{D}_d$  を複素平面上の幅  $2d$  ( $d > 0$ ) の帯状領域  $\{z \in \mathbb{C}: |\operatorname{Im}(z)| < d\}$  とする.

## 2 収束率の解析

本節では, 正定値エルミート行列  $A$  の対数に対する DE 公式の誤差および収束率を求める. まず, DE 変換  $t(x) = \tanh(\frac{\pi}{2} \sinh(x))$  を積分表示 (1) に適用すると

$$\log(A) = \int_{-\infty}^{\infty} F_{\text{DE}}(x) dx$$

となる. ただし,

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

<sup>\*1</sup> 積分表示 (1) 以外にも Cauchy の積分公式に基づく積分表示が知られているが, 計算の際には  $A$  の固有値分布を考慮して積分経路を設定する必要がある. 文献 [10] では  $A$  の固有値が実軸付近にあるときの積分経路の決め方が提案されている.

<sup>\*2</sup> 文献 [8] では,  $A$  が正定値エルミート行列であるときの, 積分  $\int_{-1}^1 (t-1)^{\alpha-1} (t+1)^{-\alpha} [(1+t)A + (1-t)I]^{-1} dt$  ( $\alpha \in (0, 1)$ ) に対する重み  $(t-1)^{\alpha-1} (t+1)^{-\alpha}$  の Gauss–Jacobi 求積が解析されている. 重みを除いた被積分関数が共通であるため, 積分表示 (1) に対する GL 求積の収束率は [8] で解析された収束率と等しい.

である． $A$  が正定値エルミート行列であるとき， $\log(A)$  に対する DE 公式の誤差は固有値の対数に対する DE 公式の誤差に帰着する．すなわち，

$$\left\| \int_{-\infty}^{\infty} F_{\text{DE}}(x) dx - \sum_{k=-\infty}^{\infty} h F_{\text{DE}}(kh) \right\|_2 = \max_{\lambda \in \Lambda(A)} \left| \int_{-\infty}^{\infty} f_{\text{DE}}(x, \lambda) dx - \sum_{k=-\infty}^{\infty} h f_{\text{DE}}(kh, \lambda) \right|$$

である．ただし， $\Lambda(A)$  は  $A$  の固有値の集合であり，

$$f_{\text{DE}}(x, \lambda) = \frac{\pi}{2} \frac{\cosh(x)}{\cosh^2(\frac{\pi}{2} \sinh(x))} \frac{\lambda - 1}{(1 + t(x))\lambda + (1 - t(x))}$$

である．そこで，まず 2.1 小節でスカラー対数関数に対する DE 公式の誤差を求め，2.2 小節で 2.1 小節の結果から  $\log(A)$  に対する DE 公式の収束率を求める．そして，2.3 小節で DE 公式と Gauss 求積を比較する．

## 2.1 スカラー対数関数に対する DE 公式の誤差

無限区間上の解析関数に対する台形則の誤差に関して以下の定理が知られている．

**定理 1** ([15, Thm. 5.1]).  $d > 0$  に対して帯状領域  $\mathcal{D}_d$  で解析的な関数  $g(z)$  が  $g(z) \rightarrow 0$  ( $|\operatorname{Re}(z)| \rightarrow \infty$ ) を満たすとする．さらに，ある定数  $c > 0$  が存在して，任意の  $y \in (-d, d)$  に対して

$$\int_{-\infty}^{\infty} |g(x + iy)| dx \leq c$$

であるとする．このとき，任意の  $h > 0$  に対して，無限区間の台形則  $\sum_{k=-\infty}^{\infty} hg(kh)$  の値が存在し，

$$\left| \int_{-\infty}^{\infty} g(x) dx - \sum_{k=-\infty}^{\infty} g(kh) \right| \leq \frac{2c}{\exp(2\pi d/h) - 1} \quad (2)$$

が成り立つ．

そこで，本小節では定理 1 を用いて，スカラー対数関数に対する DE 公式の誤差を求める．

定理 1 を用いるためには，被積分関数が解析的であるような帯状領域を知る必要がある．そこで，以下の命題に  $f_{\text{DE}}(z, \lambda)$  が解析的である帯状領域を示す．

**命題 2.**  $\lambda > 0$  に対して

$$d_0(\lambda) = \arcsin \left( \sqrt{\frac{(\log \lambda)^2 + 2\pi^2 - \sqrt{[(\log \lambda)^2 + 2\pi^2]^2 - 4\pi^4}}{2\pi^2}} \right) \quad (3)$$

とする．このとき， $f_{\text{DE}}(z, \lambda)$  は  $\mathcal{D}_{d_0(\lambda)}$  で解析的である．

**証明.**  $f_{\text{DE}}(z, \lambda)$  を構成する関数のうち， $1/\cosh(\frac{\pi}{2} \sinh(z))$  は  $\mathcal{D}_{\pi/2}$  で解析的であり [13]， $\cosh(z)$  は  $\mathbb{C}$  で解析的であるため，残りの  $1/(1+t(z))\lambda + (1-t(z))$  について考える．ここでは  $(1+t(z))\lambda +$

$(1 - t(z))$  の最も実軸に近い零点を求め、その零点の虚部の絶対値が  $d_0(\lambda)$  ( $\leq \pi/2$ ) となることから命題 2 を示す.

$(1 + t(z))\lambda + (1 - t(z)) = 0$  を変形すると

$$\exp(\pi \sinh z) = -\frac{1}{\lambda}$$

が得られる. そこで,  $z$  の実部と虚部を  $x, y$  とすると右辺と左辺はそれぞれ

$$\begin{aligned} \exp(\pi \sinh z) &= \exp(\pi \sinh x \cos y + 2i\gamma \cosh x \sin y), \\ -\frac{1}{\lambda} &= \exp(-\log \lambda + ik\pi) \quad (k = \pm 1, \pm 3, \pm 5, \dots) \end{aligned}$$

となり, 上述の 2 式の右辺を比較して

$$\begin{cases} \pi \sinh x \cos y = -\log \lambda, \\ \pi \cosh x \sin y = k\pi \end{cases}$$

となる. 両辺を 2 乗すると

$$\begin{cases} \pi^2 \sinh^2 x \cos^2 y = (\log \lambda)^2, \\ \pi^2 \cosh^2 x \sin^2 y = k^2 \pi^2. \end{cases} \quad (4)$$

が得られる. 命題 2 を示すためには零点の虚部が分かれば良いので式 (4) を  $y$  について解く. そこで,  $\cosh^2 x = 1 + \sinh^2 x$  および  $\cos^2 y = 1 - \sin^2 y$  を式 (4) に代入すると

$$\left( \pi^2 + \frac{(\log \lambda)^2}{1 - \sin^2 y} \right) \sin^2 y = k^2 \pi^2. \quad (5)$$

となる.

ここで, 方程式 (5) の左辺を  $s(y)$  とおき, 方程式  $s(y) = k^2 \pi^2$  に対する絶対値最小の解を考える. なお,  $s(y)$  は偶関数なので一般性を失わず  $y \geq 0$  とする. まず, 方程式 (5) は任意の  $k$  について  $(0, \pi/2)$  の中に解を 1 つ必ず持つ. これは  $s(y)$  が  $(0, \pi/2)$  において単調増加し,  $s(0) = 0$  かつ  $\lim_{y \rightarrow \pi/2} s(y) = \infty$  となるためである. 加えて,  $s(y)$  の単調増加性から, 方程式 (5) はあらゆる  $k$  の中で  $k = \pm 1$  のときに絶対値最小の解を持つ.

以上より, 方程式  $s(y) = \pi^2$  を  $y$  について解くと, 関数  $(1 + t(z))\lambda + (1 - t(z))$  の最も実軸に近い零点の虚部が得られ, 具体的には式 (2) の  $d_0(\lambda)$  ( $\leq \pi/2$ ) となる. 従って  $f_{\text{DE}}(z, \lambda)$  において最も実軸に近い特異点の虚部の絶対値は  $d_0(\lambda)$  であり,  $f_{\text{DE}}(z, \lambda)$  は  $\mathcal{D}_{d_0(\lambda)}$  において解析的である.  $\square$

なお,  $\delta$  を十分小さな正の数とし,  $x + iy \in \mathcal{D}_{d_0(\lambda) - \delta}$  とすると,  $|f_{\text{DE}}(x + iy, \lambda)|$  は  $|x| \rightarrow 0$  で二重指数関数的に減衰する. 従って  $\int_{-\infty}^{\infty} |f_{\text{DE}}(x + iy, \lambda)| dx$  は  $y$  について一様有界であり, 定理 1 を適用できる. 以上より,  $f_{\text{DE}}$  に対する DE 公式の誤差は十分小さな  $h$  において

$$\left| \int_{-\infty}^{\infty} f_{\text{DE}}(x, \lambda) dx - \sum_{k=-\infty}^{\infty} h f_{\text{DE}}(kh, \lambda) \right| \leq \mathcal{O} \left( \exp \left( -\frac{2\pi d_0(\lambda)}{h} \right) \right) \quad (6)$$

である.

## 2.2 行列対数関数に対する DE 公式の収束率

第 2.1 小節ではスカラー対数関数  $\log(\lambda)$  に対する DE 公式の誤差を解析した．本小節では前節の結果を用いて  $\log(A)$  に対する DE 公式の収束率を求める．

$\log(A)$  に対する DE 公式の誤差を求めるにあたって，収束の律速となる固有値を考える必要がある．そこで， $\log(\lambda)$  に対する DE 公式の収束の速さが  $\lambda$  によってどのように変化するかを考える． $d_0(\lambda)$  は  $\lambda < 1$  で単調減少し  $\lambda > 1$  で単調増加する．従って， $\lambda$  が 1 から離れる，すなわち  $\log(|\lambda|)$  が大きいと DE 公式の収束が遅くなる．また  $d_0(\lambda) = d_0(1/\lambda)$  が成り立つことから， $\log(\lambda)$  と  $\log(1/\lambda)$  に対する DE 公式は同じ速さで収束する．なお， $d_0(\lambda)$  の単調増加・単調減少性と  $d_0(\lambda) = d_0(1/\lambda)$  であることは， $d_0(\lambda)$  をプロットした図 1 から確認できる．

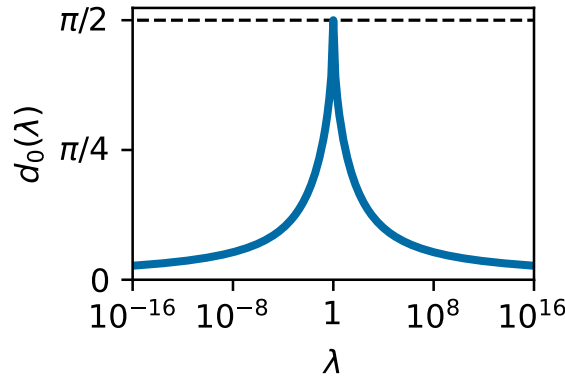


図 1  $d_0(\lambda)$  (式 (3)) のグラフ

上述より， $h$  が十分小さいとき  $\log(A)$  に対する DE 公式の誤差は最大固有値  $\lambda_{\max}$  と最小固有値  $\lambda_{\min}$  のみに依存する．すなわち，

$$\left\| \log(A) - \sum_{k=-\infty}^{\infty} h F_{\text{DE}}(kh) \right\|_2 \leq \max_{\lambda \in \{\lambda_{\max}, \lambda_{\min}\}} \left| \log(\lambda) - \sum_{k=-\infty}^{\infty} h f_{\text{DE}}(kh, \lambda) \right|$$

である．さらに， $\lambda_{\max} \lambda_{\min} = 1$ ，すなわち， $|\log(\lambda_{\max})| = |\log(\lambda_{\min})|$  を仮定すると， $d_0(\lambda_{\max}) = d_0(1/\lambda_{\max}) = d_0(\lambda_{\min})$  であることから  $\log(\lambda_{\max})$  と  $\log(\lambda_{\min})$  に対する DE 公式の収束率は等しくなり，(6) より

$$\left\| \log(A) - \sum_{k=-\infty}^{\infty} h F_{\text{DE}}(ih) \right\|_2 \leq \mathcal{O} \left( \exp \left( -\frac{2\pi d_0(\lambda_{\max})}{h} \right) \right) = \mathcal{O} \left( \exp \left( -\frac{2\pi d_0(\sqrt{\kappa(A)})}{h} \right) \right)$$

となる．ここで  $\kappa(A) = \|A\|_2 \|A^{-1}\|_2$  である．なお， $\log(A) = \log(pA) - \log(p)I$  ( $p > 0$ ) が成り立つので， $A$  を  $pA$  ( $p = 1/\sqrt{\lambda_{\max} \lambda_{\min}}$ ) と定数倍することを考えれば仮定  $\lambda_{\max} \lambda_{\min} = 1$  は一般性を失わない．

さらに，DE 公式を計算するための有限の積分区間が  $[l, r]$  と得られているとする．なお，このような区間は文献 [14] における Algorithm 1 の第 2 ステップから第 11 ステップを用いると得られ

る\*3. このとき、台形則の誤差は積分点数  $m = (r - l)/h + 1$  を用いて

$$\left\| \log(A) - \sum_{k=1}^{m-2} h F_{\text{DE}}(l + kh) - \frac{h}{2} (F_{\text{DE}}(l) + F_{\text{DE}}(r)) \right\|_2 \leq \mathcal{O} \left( \exp \left( -\frac{2\pi d_0(\sqrt{\kappa(A)})}{r-l} m \right) \right).$$

と表される. 以上より, 本研究の主結果として,  $\log(A)$  に対する DE 公式の収束率

$$\mathcal{O} \left( \exp \left( -\frac{2\pi d_0(\sqrt{\kappa(A)})}{r-l} \right) \right) \quad (7)$$

を得る.

## 2.3 DE 公式と GL 求積との比較

本小節では DE 公式と GL 求積の収束の速さを比較する.  $A$  の最大固有値と最小固有値の積が 1 のとき, どちらの求積法も積分点数に対して指数的に誤差が減少し, その誤差を  $\mathcal{O}(\exp(-\phi m))$  と表せる. ここで,  $\phi$  は求積法と  $\kappa(A)$  のみに依存する定数であり,  $\phi$  が大きいほど収束が速い. DE 公式では式 (7) より

$$\phi = \frac{2\pi d_0(\sqrt{\kappa(A)})}{r-l},$$

GL 求積では

$$\phi = 2 \log \left( \frac{\kappa(A)^{1/4} + 1}{\kappa(A)^{1/4} - 1} \right)$$

である [8]. そこで, 図 2 に DE 公式と GL 求積における収束の速さ  $\phi$  を  $\kappa(A)$  に対してプロットしたものを示す. なお, DE 公式の有限な積分区間  $[l, r]$  は積分区間決定のアルゴリズム [14, Alg. 1] を用いて, 積分区間の打ち切り誤差が  $\varepsilon = 2^{-53} \approx 1.1 \times 10^{-16}$  以下になるように得た\*4. 図 2 より,  $\kappa(A) \gtrsim 2.7 \times 10^3$  のとき DE 公式が速く収束し,  $\kappa(A) \lesssim 2.7 \times 10^3$  のとき GL 求積が速く収束することがわかる. この結果から, 正定値エルミート行列の対数を数値積分法で計算するときには, 条件数が大きいときは DE 公式, 小さいときは GL 求積を選べば良いと分かる.

\*3 積分区間決定のアルゴリズム [14, Alg. 1] は DE 変換として  $t(x) = \tanh(\sinh(x))$  を用いることを前提としているが, アルゴリズムの第 10 ステップおよび第 11 ステップにおいて,  $l = \text{asinh}(2 \tanh(2a - 1)/\pi)$ ,  $r = \text{asinh}(2 \tanh(2b - 1)/\pi)$  と計算することで, 本稿で用いた変数変換  $t(x) = \tanh(\frac{\pi}{2} \sinh(x))$  に対応できる.

\*4 積分区間決定のアルゴリズム [14, Alg. 1] は  $\|A^{-1}\|_2$ ,  $\|A - I\|_2$ , 及び  $A$  のスペクトル半径を必要とするが,  $A$  が正定値エルミート行列かつ  $\lambda_{\max} \lambda_{\min} = 1$  のときは  $\kappa(A)$  からその値を求めることができる. したがって本稿では積分区間  $[l, r]$  は  $\kappa(A)$  と  $\varepsilon$  のみに依存する.

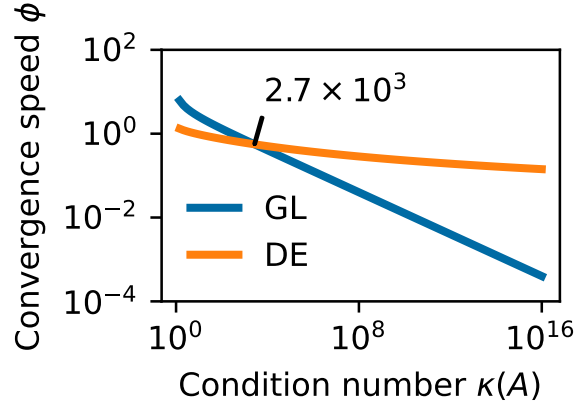


図2 DE 公式と GL 求積の収束の速さ.  $\phi$  が大きいほど収束が速い.

表1 数値例で用いたテスト行列

行列	$n$	$\kappa(A)$
nos4 [5]	100	$1.6 \times 10^3$
bcsstk04 [5]	132	$2.3 \times 10^6$
lund_b [5]	147	$3.0 \times 10^4$

### 3 数値例

本節では表1に示す行列を用いて、DE 公式の収束率の見積りの妥当性を確かめる. なお、プログラミングは Julia を用いて行い、特に明記しない限り倍精度演算を用いた. 任意精度演算は Julia の BigFloat 型を用いて実装し、GL 求積の積分点と重みの計算には QuadGK.jl<sup>\*5</sup>を用いた. なお、計算に用いたソースコードと実行結果は <https://github.com/f-ttok/article-rims2019> にアップロードした. 計算手順は以下のとおりである.

1.  $\tilde{A} = A/\sqrt{\lambda_{\max}\lambda_{\min}}$  とスケーリングし、 $\log(\tilde{A})$  を任意精度演算向きの Inverse Scaling and Squaring アルゴリズム [8, Alg. 5.1] によって計算した.
2.  $\log(\tilde{A})$  を倍精度で DE 公式によって計算した.
3.  $\log(\tilde{\lambda}_{\max})$  ( $\tilde{\lambda}_{\max} = \lambda_{\max}/\sqrt{\lambda_{\max}\lambda_{\min}}$ ) を DE 公式によって計算した. ただし  $\log(\tilde{A})$  を DE 公式で計算したときと同じ積分区間を用いた.
4. 参考のため、 $\log(\tilde{A})$  を GL 求積を用いて計算した.

結果の誤差履歴を図3に示す.

図3より以下の3点が観察できた:

- $\lambda_{\max}\lambda_{\min} = 1$  の条件下では  $\log(\tilde{\lambda}_{\max})$  と  $\log(\tilde{A})$  の収束履歴はほぼ一致した.

<sup>\*5</sup> <https://github.com/JuliaMath/QuadGK.jl>

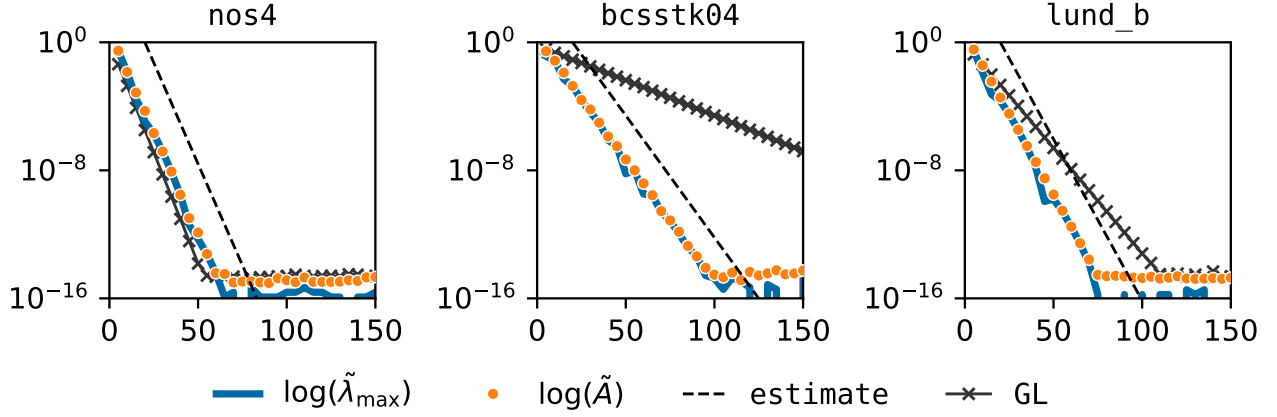


図3 テスト行列に対する DE 公式の誤差履歴.

- DE 公式の誤差履歴の傾きは、収束率の見積りから得た線の傾きとほぼ一致した。
- GL 求積と DE 公式を比較すると、nos4 では同程度の速さで、ほかの 2 つのテスト行列では DE 公式が速く収束した。これはテスト行列の条件数を考えると妥当な結果である。

以上より本稿の議論の妥当性を確認できた。

## 4 まとめ

本稿では、正定値エルミート行列  $A$  において、 $\log(A)$  を計算するための DE 公式の収束率を  $A$  の条件数をもって見積もった。この見積もりを用いて GL 求積と比較すると、 $A$  の条件数が多いときは DE 公式が速く収束し、小さいときは GL 求積が速く収束することが分かった。条件数の分水嶺の数値はおおよそ  $2.7 \times 10^3$  である。また、数値例を通して議論の妥当性を確認した。

今後の課題として、正定値エルミート行列でない一般の行列に対する GL 求積と DE 公式の解析が挙げられる。また、並列化を行った上で大規模な実問題へ数値積分法を適用し、 $\log(A)b$  の計算における性能評価を行うことも挙げられる。

## 謝辞

DE 公式の誤差解析についてコメントをいただいた東京大学の田中健一郎先生に感謝申し上げます。本研究は日本学術振興会科研費 18J22501, 18H05392 の助成を受けた。

## 参考文献

- [1] A. H. Al-Mohy and N. J. Higham. Improved inverse scaling and squaring algorithms for the matrix logarithm. *SIAM J. Sci. Comput.*, 34(4):C153–C169, 2012.
- [2] M. Anitescu, J. Chen, and L. Wang. A matrix-free approach for solving the parametric



- Gaussian process maximum likelihood problem. *SIAM J. Sci. Comput.*, 34(1):A240–A262, 2012.
- [3] J. R. Cardoso and R. Ralha. Matrix arithmetic-geometric mean and the computation of the logarithm. *SIAM J. Matrix Anal. Appl.*, 37(2):719–743, 2016.
  - [4] P. I. Davies and N. J. Higham. Computing  $f(A)b$  for matrix functions  $f$ . In *QCD and numerical analysis III*, volume 47 of *Lect. Notes Comput. Sci. Eng.*, page 15–24. Springer, Berlin, 2005.
  - [5] T. A. Davis and Y. Hu. The University of Florida sparse matrix collection. *ACM Trans. Math. Software*, 38(1):Art. 1, 25 pp., 2011.
  - [6] K. Dong, D. Eriksson, H. Nickisch, D. Bindel, and A. G. Wilson. Scalable log determinants for Gaussian process kernel learning. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 6327–6337. Curran Associates, Inc., 2017.
  - [7] M. Fasi and N. J. Higham. Multiprecision algorithms for computing the matrix logarithm. *SIAM J. Matrix Anal. Appl.*, 39(1):472–491, 2018.
  - [8] M. Fasi and B. Iannazzo. Computing the weighted geometric mean of two large-scale matrices and its inverse times a vector. *SIAM J. Matrix Anal. Appl.*, 39(1):178–203, 2018.
  - [9] H. Fawzi, J. Saunderson, and P. A. Parrilo. Semidefinite approximations of the matrix logarithm. *Found. Comput. Math.*, 19(2):259–296, 2019.
  - [10] N. Hale, N. J. Higham, and L. N. Trefethen. Computing  $\mathbf{A}^\alpha$ ,  $\log(\mathbf{A})$ , and related matrix functions by contour integrals. *SIAM J. Numer. Anal.*, 46(5):2505–2523, 2008.
  - [11] I. Han, D. Malioutov, and J. Shin. Large-scale log-determinant computation through stochastic Chebyshev expansions. In F. Bach and D. Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 908–917, Lille, France, 2015. PMLR.
  - [12] N. J. Higham. *Functions of Matrices: Theory and Computation*. SIAM, Philadelphia, PA, 2008.
  - [13] H. Takahasi and M. Mori. Double exponential formulas for numerical integration. *Publ. Res. Inst. Math. Sci.*, 9:721–741, 1974.
  - [14] F. Tatsuoka, T. Sogabe, Y. Miyatake, and S.-L. Zhang. Algorithms for the computation of the matrix logarithm based on the double exponential formula. *J. Comput. Appl. Math.*, 373(1):112396, 2019.
  - [15] L. N. Trefethen and J. A. C. Weideman. The exponentially convergent trapezoidal rule. *SIAM Rev.*, 56(3):385–458, 2014.