

確率分布に関する数値積分の改良: 基底関数の拡張と積分経路の変更

数理情報学専攻 48206212 木村 直人

指導教員 田中 健一郎 准教授

1 概要

値を解析的に計算できない積分を数値的に計算する手法として、数値積分は応用上非常に重要である。本研究では確率分布に関する数値積分に着目し、期待値の数値積分についての Tanaka-Toda の研究 [3] および積分を利用した狭義安定分布の確率密度関数の数値計算についての Saenko の研究 [2] の改良をそれぞれ試みた。

2 基底関数の拡張による数値積分公式の改良

2.1 既存研究

区間 I 上の確率密度関数 $f(x)$ と I 上可積分関数 $g(x)$, 標本点 $x_1, \dots, x_N \in I$ が与えられたとき,

$$E[g(X)] = \int_I g(x)f(x)dx \approx \sum_{n=1}^N p_n g(x_n) \quad (1)$$

という近似が成り立つ重み p_1, \dots, p_N を, 次の制約

$$\sum_{n=1}^N p_n = 1, p_1, \dots, p_N \geq 0, \quad (2)$$

$$\tilde{T}_m = \sum_{n=1}^N p_n T_m(x_n) \quad (m = 1, \dots, M) \quad (3)$$

の下で求めることを考える。ここで $m = 1, \dots, M$ に対し $T_m(x)$ は I 上可積分関数, $\tilde{T}_m := E[T_m(X)]$ である。また, $T_m(x)$, \tilde{T}_m をそれぞれ基底関数, モーメントとよぶことにする。Tanaka-Toda は (2) と (3) を満たす重み p_1, \dots, p_N を, 次の無制約凸最適化問題

$$(D) \min_{\lambda \in \mathbb{R}^M} \left[-\lambda^\top \tilde{T} + \log \left(\sum_{n=1}^N q_n e^{\lambda^\top T(x_n)} \right) \right] \quad (4)$$

の解 $\tilde{\lambda}$ を用いて

$$p_n = \frac{q_n e^{\tilde{\lambda}^\top T(x_n)}}{\sum_{n=1}^N q_n e^{\tilde{\lambda}^\top T(x_n)}} \quad (n = 1, \dots, N) \quad (5)$$

で与えた。ここで $T(x) := (T_1(x), \dots, T_M(x))^\top$, $\tilde{T} := (\tilde{T}_1, \dots, \tilde{T}_M)^\top$ である。

2.2 提案手法

本研究では, 複合台形公式の改良に関する Fornberg の研究 [1] に着想を得て, Tanaka-Toda の数値積分にお

ける基底関数の個数をパラメータを与えて連続化することで精度の改善を試みた。詳細には, $I = [0, 1]$, $f(x) = 1$ とし, 基底関数をパラメータ $z \in [s, t]$ ($0 < s < t$) を用いて $T(x; z) := x^z$ とする。このとき, モーメント $\tilde{T}(z)$ は $\tilde{T}(z) = 1/(z+1)$ で与えられる。さらに, z の関数 $a(z), b(z)$ に対し内積を $\langle a, b \rangle := \int_s^t a(z)b(z)dz$ と定め, 問題 (D) を

$$(D') \min_{\lambda(z)} \left[-\langle \lambda, \tilde{T} \rangle + \log \left(\sum_{n=1}^N q_n e^{\langle \lambda, T(x_n) \rangle} \right) \right] \quad (6)$$

と拡張する。一般の $\lambda(z)$ について (D') を解くのは困難であるため, 本研究では $c_1, \dots, c_M \in \mathbb{R}$ として $\lambda(z) = \sum_{m=1}^M c_m (z+1)^m$ という形式を仮定した。この仮定により問題 (D') は c_1, \dots, c_M に関する無制約凸最適化問題に帰着されるため, 既存研究と同様にして重み p_1, \dots, p_N を計算できる。最適化問題の変数を増やせずに基底関数の個数に対応するパラメータ s, t を選ぶことができる点が既存研究との大きな差異である。

2.3 数値実験

テスト問題 $\int_0^1 (e^x + x^{\frac{1}{2}})dx$ に対し, 複合台形公式, 既存研究, 提案手法のそれぞれで数値積分した際の標本点の個数と絶対誤差との関係を図 1 に示す。このテスト問題においては既存研究と比較して提案手法が高精度であることがわかる。

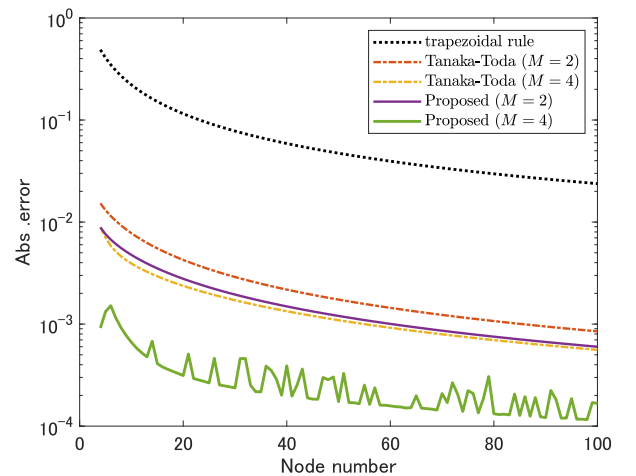


図 1. $\int_0^1 (e^x + x^{\frac{1}{2}})dx$ を複合台形公式, 既存研究および提案手法 ($M = 2, 4$; $(s, t) = (0.6, 2)$) で計算した場合の標本点の個数と絶対誤差との関係

3 積分経路の変更による安定分布の確率密度関数の数値計算

3.1 既存研究

安定分布は次のように定義される: 確率変数 X に対し, X と i.i.d. である確率変数 X_1, X_2 と任意の $a, b > 0$ に対し, ある $c > 0, d \in \mathbb{R}$ が存在して $aX_1 + bX_2 \stackrel{d}{=} cX + d$ であるとき, X が従う確率分布を安定分布といい, 特に $d = 0$ であるとき狭義安定分布という. 安定分布の確率密度関数は一般には閉じた式で表せない一方で, その特性関数は一般に閉じた式で表せる. 狭義安定分布の (規格化された) 特性関数は $0 < \alpha \leq 2, |\theta| \leq \min\{1, 2/\alpha - 1\}$ をパラメータとして

$$\psi(t; \alpha, \theta) := \exp \left[-|t|^\alpha \exp \left(-i \frac{\pi}{2} \alpha \theta \operatorname{sign} t \right) \right] \quad (7)$$

で与えられる. 確率密度関数 $f(x; \alpha, \theta)$ は

$$f(x; \alpha, \theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \psi(t; \alpha, \theta) dt \quad (8)$$

と表されるため, 数値積分により値を求めることができる. しかし, この積分は振動型積分であり, 通常の数値積分では高精度な値を得るのが困難である.

Saenko はこの振動型積分を, 振動項が消失する積分経路を導出することで数値積分が容易である非振動型積分に帰着させる手法を提案した. 詳細には, $0 < \alpha < 1, 1 < \alpha \leq 2, |\theta| \leq \min(1, 2/\alpha - 1), x > 0$ のとき,

$$\frac{\alpha}{\pi x |1 - \alpha|} \int_{-\frac{\pi}{2}\theta}^{\frac{\pi}{2}} A(\varphi; \alpha, \theta, x) e^{-A(\varphi; \alpha, \theta, x)} d\varphi \quad (9)$$

により $f(x; \alpha, \theta)$ の値を求めることができる. ここで, $A(\varphi; \alpha, \theta, x) := \left\{ \frac{\sin[\alpha(\varphi + \frac{\pi}{2}\theta)]}{x \cos \varphi} \right\}^{\frac{\alpha}{1-\alpha}} \frac{\cos[\varphi(1-\alpha) - \frac{\pi}{2}\alpha\theta]}{\cos \varphi}$ である. なお, $\alpha = 1$ または $x = 0$ の場合, $f(x; \alpha, \theta)$ は閉じた式で表され, $x < 0$ の場合, $f(-x; \alpha, \theta) = f(x; \alpha, -\theta)$ より $x > 0$ の場合に帰着される.

3.1.1 提案手法

既存研究はパラメータ α が 1 に近いとき被積分関数が発散してしまい数値積分が非常に困難であるという課題を抱えている. 本研究ではその課題に着目し, そのような発散が起こらない新たな積分経路を導出して特性指数が 1 に近い場合でも数値計算が可能である非振動型積分を与えた. 詳細には, $u = \left\{ \frac{\sin[\alpha(\varphi + \frac{\pi}{2}\theta)]}{x \cos \varphi} \right\}^{\frac{\alpha}{1-\alpha}}$ という変数変換を行うことにより

$$\frac{1}{\pi} \int_0^\infty \frac{u^{\frac{1-\alpha}{\alpha}} C(\varphi; \alpha, \theta) e^{-u C(\varphi; \alpha, \theta)} \cos \varphi}{C(\varphi; \alpha, \theta) - (1 - \alpha) \cos[\alpha(\varphi + \frac{\pi}{2}\theta)]} du \quad (10)$$

で表される新たな非振動型積分を導出した. ここで, $C(\varphi; \alpha, \theta) := \frac{\cos[\varphi(1-\alpha) - \frac{\pi}{2}\alpha\theta]}{\cos \varphi}$ である. また, 積分中の φ については, $u^{\frac{1-\alpha}{\alpha}} = \frac{\sin[\alpha(\varphi + \frac{\pi}{2}\theta)]}{x \cos \varphi}$ の右辺の狭義単調増加性を利用した二分探索法により, 高精度な値を高速に数値計算することができる.

3.2 数値実験

既存研究では計算が非常に困難である $\alpha = 1.000001$ の場合の確率密度関数の値を提案手法を用いて計算した. 計算した確率密度関数を x に関してプロットしたものを図 2 に, 計算した確率密度関数の具体的な数値を表 1 に示す. このように, 提案手法は α が 1 に近い場合であっても計算可能であることがわかる.

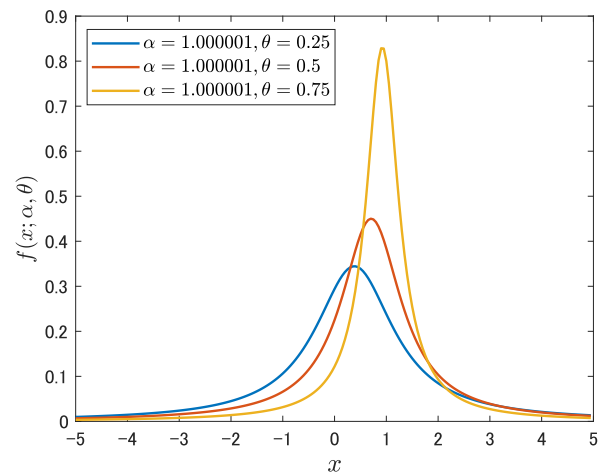


図 2. $f(x; 1.000001, \theta)$ の概形 ($\theta = 0.25, 0.5, 0.75$)

表 1. $f(x; 1.000001, 0.25)$ の値

x	$f(x; 1.000001, 0.25)$
-4	$1.465894635570591 \times 10^{-2}$
-1	$1.063440008212844 \times 10^{-1}$
-0.1	$2.706580507932997 \times 10^{-1}$
0.1	$3.150416686344176 \times 10^{-2}$
1	$2.381924579874976 \times 10^{-1}$
4	$2.109833035729237 \times 10^{-2}$

参考文献

- [1] B. Fornberg: Improving the Accuracy of the Trapezoidal Rule. *SIAM Rev.*, vol. 63, no. 1, pp. 167–180, 2021.
- [2] V. V. Saenko: The Calculation of the Density and Distribution Functions of Strictly Stable Laws. *Math.*, vol. 8, no. 5, 2020.
- [3] K. Tanaka and A. A. Toda: Discretizing Distributions with Exact Moments: Error Estimate and Convergence Analysis. *SIAM J. Numer. Anal.*, vol. 53, pp. 2158–2177, 2015.