

平滑化相互作用エネルギーに関する種々の考察及びサンプリングへの適用

数理情報学専攻 48226230 福元 啓悟

指導教員 田中 健一郎 准教授

1 はじめに

真の確率分布を有限個の点による離散分布によって近似する問題を考える。この問題は、真の分布と離散分布の差を何らかの関数によって数値化することによって、目的関数の最小化問題として定式化することができる。

この問題に対して、[1] は平滑化相互作用エネルギー (MIE) の最小化によるアプローチを提案した。MIE は測地的凸性の「必要条件」となる不等式の成立など、良い性質を持つことが示されている。しかし、この不等式は極限でしか成り立たないため、実際に最適化を行う際には測地的凸性の恩恵を受けることができないという課題がある。

2 既存研究

$\mathcal{X} \subset \mathbb{R}^n$ を標本点の属する空間、 μ^* を真の確率分布、 $p \in C^1(\mathcal{X}), 0 < p(x) < \infty$ を真の分布の密度関数とする。

2.1 MIE の定義

定義 1 ([1, Def. 3.1]). $\phi_\epsilon : \mathbb{R}^n \rightarrow \mathbb{R}$ について、 $\{\phi_\epsilon\}_{\epsilon>0}$ は以下の 3 条件を満たすとき軟化子の族と呼ばれる。

1. $\forall \epsilon > 0, \forall x \in \mathbb{R}^n, \phi_\epsilon(x) \geq 0$ and $\phi_\epsilon(x) = \phi_\epsilon(-x)$.
2. $\forall \epsilon > 0, \|\phi_\epsilon\|_1 = 1$ and $\sup_{x \in \mathbb{R}^n} \phi_\epsilon(x) < \infty$.
3. $\forall \delta > 0, \forall p \in \{1, \infty\}, \lim_{\epsilon \rightarrow 0} \|\mathbb{1}_{\mathbb{R}^n \setminus B_\delta(0)} \phi_\epsilon\|_p = 0$.

定義 2 ([1, Def. 3.2]). $\{\phi_\epsilon\}_{\epsilon>0}$ を軟化子の族とする。このとき、 $\epsilon > 0$ に対して、対称カーネル $W_\epsilon : \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, \infty]$ を以下のように定義する、

$$W_\epsilon(x, y) = \begin{cases} \phi_\epsilon(x - y)(p(x)p(y))^{-\frac{1}{2}} & \text{if } x, y \in \mathcal{X}, \\ \infty & \text{otherwise.} \end{cases} \quad (1)$$

この W_ϵ を用いて、 $\mathcal{E}_\epsilon : \mathcal{P}(\mathcal{X}) \rightarrow [0, \infty]$ を以下の通り定義する、

$$\mathcal{E}_\epsilon(\mu) = \int \int W_\epsilon(x, y) d\mu(x) d\mu(y). \quad (2)$$

この \mathcal{E}_ϵ を MIE と呼ぶ。

MIE は、 $\epsilon \rightarrow 0$ の極限にて χ^2 -divergence に収束することが知られている [1, Thm. 3.3].

2.1.1 測地的凸性の「必要条件」

定理 1 ([1, Prop. 3.8.]). $\forall \xi \in C_c^\infty(\mathcal{X}; \mathbb{R}^{*n})$ に対して

$$\mu_t = (I + t\xi)_{\#} \mu^* \quad (3)$$

とすると、いくつかの仮定のもとで

$$\lim_{\epsilon \rightarrow 0} \frac{d^2}{dt^2} \mathcal{E}_\epsilon(\mu_t) \Big|_{t=0} \geq 0 \quad (4)$$

が成り立つ。

定理 1 は「 $\epsilon \rightarrow 0$ の極限かつ真の分布」という限定的な場合のみにおいて意味のあるものである。したがって、そのまま実際のアルゴリズムに活用することは非常に難しい。

2.2 離散 MIE

上で定義した MIE は空間方向に連続な一般の確率分布について定義されているため、実際に数値実験を行うためには以下で定義される離散 MIE を用いる。

定義 3. サンプル点 $\omega_N = \{x_1, x_2, \dots, x_N\} \subset \mathcal{X}$ について、離散 MIE は以下の式で定義される：

$$E_\epsilon(\omega_N) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \phi_\epsilon(x_i - x_j) (p(x_i)p(x_j))^{-\frac{1}{2}}. \quad (5)$$

[1] ではこの離散 MIE を用いて数値実験にて有効性を示しているものの、離散 MIE に最小化アルゴリズムを適用したときにどのような条件で減少列となるかの理論解析は全く行っていない。

3 本研究 1 : MIE に対する正則化項の付加

本研究では MIE に正則化項を付加することで測地的凸性の必要条件である 2 階微分の不等式が成り立つ範囲を広げ、実際の MIE によるサンプリングの性能を向上させることを目指した。

定義 4. $\{\phi_\epsilon\}_{\epsilon>0}$ を軟化子の族とする。この時 MIE および、 $\mathcal{R}(\mu) : \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}_{\geq 0}$ について、以下の式を考える：

$$\mathcal{E}_\epsilon^{\mathcal{R}}(\mu) = \int \int W_\epsilon(x, y) d\mu(x) d\mu(y) + \mathcal{R}(\mu). \quad (6)$$

このとき、(6) 式で定義される $\mathcal{E}_\epsilon^{\mathcal{R}} : \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}$ を、 \mathcal{R} を正則化項とする RMIE と言い、 \mathcal{R} を正則化項と呼ぶ。

定義 5. single mean の正則化項 $\mathcal{R}_{\text{single}}$ は以下のよう
に定義される:

$$\mathcal{R}_{\text{single}}(\mu; \sigma, m^*) = \frac{\sigma}{2} \int \|x - m^*\|_2^2 d\mu(x). \quad (7)$$

ただし,

$$\sigma \in \mathbb{R}_{\geq 0}, m^* \in \mathbb{R}^n \quad (8)$$

である.

定理 2. $\|\xi(\cdot)\sqrt{p(\cdot)}\|_2 = 1$ を満たす $\forall \xi \in C_c^\infty(\mathcal{X}; \mathcal{X})$
に対して

$$\mu_t = (I + t\xi)_{\#}\mu^* \quad (9)$$

とすると, $\mathcal{R}_{\text{single}}$ を正則化項とする RMIE $\mathcal{E}_\epsilon^{\mathcal{R}_{\text{single}}}$ に
ついて, 一定の仮定のもとで

$$\begin{aligned} \frac{d^2}{dt^2} \mathcal{E}_\epsilon^{\mathcal{R}_{\text{single}}}(\mu_t) &\geq \sigma - C_1(\epsilon)n^2 \|\xi\|_*^2 - C_2(\epsilon)n^3 t \|\xi\|_*^3 \\ &\quad - C_3(\epsilon)n^4 t^2 \|\xi\|_*^4 \end{aligned} \quad (10)$$

が成り立つ. ただし, $\|\cdot\|_*$ はノルムである.

本定理より, MIE では極限でしか成り立たなかった
測地的凸性の必要条件となる不等式が, 正則化項の付加
によって必ずしも極限をとらなくても成り立つように
なった.

4 本研究 2 : MIE/RMIE に関する理論評価

本研究では, 実験に使用する離散 MIE/離散 RMIE に
ついて理論解析を行い, 最急降下法の適用時に一定の
ステップサイズのもとで減少列となることを示した. 本
稿では結果の一部を抜粋する.

定義 6. サンプル点 $\omega_N = \{x_1, x_2, \dots, x_N\} \subset \mathcal{X}$ が
与えられたとする. この時, 離散 MIE E_ϵ および関数
 $R: \mathcal{X}^N \rightarrow \mathbb{R}_{\geq 0}$ について, 以下の式を考える:

$$E_\epsilon^R(\omega_N) = \frac{1}{N^2} \sum_{i,j=1}^N \phi_\epsilon(x_i - x_j) (p(x)p(y))^{-\frac{1}{2}} + R(\omega_N). \quad (11)$$

この時, (11) 式で定義される $E_\epsilon^R: \mathcal{X}^N \rightarrow \mathbb{R}$ を, R を正
則化項とする RMIE と言い, R を正則化項と呼ぶ.

定理 3. R_{single} ($\mathcal{R}_{\text{single}}$ の離散版) をもとにした離散
RMIE について, 一定の仮定のもとで

$$\begin{aligned} &E_\epsilon^{R_{\text{single}}}(\omega_N^{m+1}) - E_\epsilon^{R_{\text{single}}}(\omega_N^m) \\ &\leq -\gamma \left(1 - \left(2L + \frac{1}{2}\sigma\right)\gamma\right) \|\nabla E_\epsilon^{R_{\text{single}}}(\omega_N^m)\|_2^2 \end{aligned} \quad (12)$$

が成り立つ. ただし, $L > 0$ は定数である.

5 数値実験

5.1 アルゴリズム

離散 MIE については, 単純に E_ϵ に対して Adam を
適用して最小化を行った.

一方で離散 RMIE については, 下記アルゴリズムに
基づいて最小化を行った.

Algorithm 1 離散 RMIE の最小化アルゴリズム

Require: $\sigma_0 > 0, 0 < \sigma_1 < 1, \omega_N = \{x_i\}_{i=1}^N, N_1 \in \mathbb{N}, N_2 \in \mathbb{N}, R(\omega_N; \sigma)$: 正則化項

- 1: **for** $i = 0 \sim N_1 - 1$ **do**
- 2: $\sigma' \leftarrow \sigma_1^i$
- 3: **for** $j = 0 \sim N_2 - 1$ **do**
- 4: $e_1 \leftarrow E_\epsilon(\omega_N)$
- 5: $e_2 \leftarrow R(\omega_N; 1)$
- 6: $\sigma \leftarrow \frac{e_1}{e_2} \sigma_0 \cdot \sigma'$
- 7: $E_\epsilon^R(\cdot) = E_\epsilon(\cdot) + R(\cdot; \sigma)$ に Adam を適用して
 ω_N を更新
- 8: **end for**
- 9: **end for**

Ensure: ω_N

5.2 実験結果

$r \leq \sin^2 \theta + 0.2$ 上の一様分布を真の分布として数値
実験を行った結果, 下記の通りとなった.

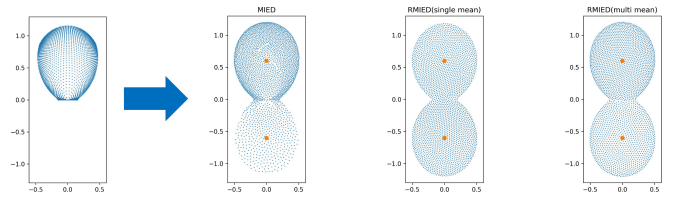


図 1. $r \leq \sin^2 \theta + 0.2$ 上の一様分布を真の
分布として MIED, RMIED(single mean),
RMIED(multi mean) を適用した際のサンプ
ル点の分布.

図 1 の通り, 正則化項を付加することでより真の分布
に近い形に収束するという結果となった.

参考文献

- [1] Lingxiao Li, Qiang Liu, Anna Korba, Mikhail Yurochkin, and Justin Solomon. Sampling with mollified interaction energy descent. In *The Eleventh International Conference on Learning Representations*, 2023.