

研究終了報告書

「等質空間上の調和指数型分布族とその応用」

研究期間：2019年10月～2022年3月

研究者：東條 広一

1. 研究のねらい

本研究の目的は、等質空間上の指数型分布族を表現論によって統一的に扱う枠組みを与えること、および球面や双曲空間などの重要な空間上に良い指数型分布族を与え応用可能にすることである。

研究の背景：機械学習や統計学の分野で、標本空間の性質を反映した訓練可能な分布族は理論上も応用上も重要である。特に、近年機械学習の分野で階層構造を持つデータの解析に上半平面、双曲空間が有用であることが示されており、その上の良い分布族を与えることが期待されている。分布族の中でも指数型分布族は、情報幾何の分野で重要な研究対象であり、共役事前分布を持つことからベイズ推定の分野でも重宝される。指数型分布族は定義上、無数に考えられるが、広く用いられているのはそのごく一部に過ぎない。それらは指数型分布族の中でも“良い”分布族と考えられ、個別にはよく調べられているが、“良い”指数型分布族のみを統一的に扱う枠組みは未整備である。そこで、これら“良い”指数型分布族は、標本空間が対称性を持つ空間(=等質空間)上の対称性を持つ分布族であることに注目し、リー群の表現論を用いて等質空間上の指数型分布族を構成する手法を[TY18]において提案した。本手法を用いて目的にアプローチしたい。

[TY18] K. Tojo, T. Yoshino, A method to construct exponential families by representation theory, arXiv:1811.01394, preprint, 2018

2. 研究成果

(1) 概要

A. 提案手法の持つ性質について

研究目的を達成するにあたり導入した、表現論を用いて等質空間上に指数型分布族を構成する手法が、実際に対称性を持つ分布族を統一的に扱う枠組みを与えるのに必要な性質(詳細参照)を持つことを示した。これにより、対称性を持つ分布族を表現論の手法を用いて解析することが可能となった。

B. 上半平面上の指数型分布族とその共役事前分布族

提案手法を用いて得られた上半平面上の指数型分布族に対して、共役事前分布族を明示的に与えた。これにより、上半平面上でのこの分布族・共役事前分布族を用いたベイズ推定が可能となった。

(2) 詳細

A 提案手法の持つ性質について

G をリー群, H を G の閉部分群とする. このとき G の等質空間 G/H 上の分布全体の集合にはリー群 G の作用が自然に定まる. このとき, “良い” 指数型分布族は G/H 上の G 不変な指数型分布族と考えられる. 例えば, 正規分布族の場合, スケーリングと平行移動によって正規分布を変形してもまた正規分布となっていることを意味する. 提案手法は G/H 上の G 不変指数型分布族を与える. 逆に以下の問が自然に考えられる:

問: G/H 上の G 不変指数型分布族は提案手法を用いて得られるか?

これに対し, 弱い仮定のもとで肯定的な答えを与えた(研究成果リスト[1]). これにより, ラフには以下のように考えてよい.

{ G/H 上の G 不変指数型分布族} “=” {提案手法を用いて得られる G/H 上の分布族}

B 上半平面上の指数型分布族とその共役事前分布族

パラメライズされた分布族に対し, その共役事前分布族があると, ベイズ推定が簡単に実行できる. 実際に得られたデータに応じて, 共役事前分布族のハイパーパラメータを更新していけばよいためである. 指数型分布族には共役事前分布族が存在することが知られている. しかしながら, 与えられた指数型分布族に対し, 正規化定数まで含めて共役事前分布族を明示的に与えることは一般には難しい.

上半平面には $SL(2, \mathbb{R})$ の 1 次分数変換による自然な作用が定まっている. 本研究では, 上半平面上で最も基本的だと考えられる $SL(2, \mathbb{R})$ 不変な指数型分布族とその共役事前分布族を正規化定数まで含めて明示的に与えた(研究成果リスト[2]). さらにパラメータ空間に定まる $SL(2, \mathbb{R})$ 作用から誘導される作用に関して, 共役事前分布族も $SL(2, \mathbb{R})$ 不変であることを示した.

3. 今後の展開

機械学習への応用という観点から, 既約リーマン対称空間の中でも双曲空間は, 重要な標本空間であると考えられる. そこで, 双曲空間上に提案手法を用いて得られる分布族を用いた最尤推定手法を構築したい. 球面上の対称性を持つ分布族を用いた最尤推定手法を提案した先行研究として, [CW15]がある. 実際, 分布族の対称性を用いて計算を高速化(高速フーリエ変換)することで, 現実的な計算量でパラメータ推定を可能としている. 提案手法は[CW15]で提案されている分布族の構成法の一般化とみなせるため, [CW15]をお手本として, 類いの議論をすることで, 双曲空間における最尤推定手法を与えることが今後の目標の一つである.

[CW15] T. S. Cohen, M. Welling, Harmonic exponential families on manifolds, In Proceedings of the 32nd International Conference on Machine Learning (ICML), vol. 37 (2015), 1757–1765.

4. 自己評価

「今後の展開」で述べた双曲空間上での最尤推定手法の構築は、当初の研究計画に含まれていたが、あまり進められなかった。球面や双曲空間を含む既約リーマン対称空間と呼ばれるのクラスに対して提案手法を適用する際に手法が許容する入力ペアの分類に取り組み、その論文執筆に予想以上に時間がかかっているのが原因の一つと考えられる。応用するにはその論文の内容が必要ではあるが、双曲空間に特化した最尤推定手法についてであれば、ACT-X の他の研究者と議論しやすい側面もあると考えられ、機械学習分野におけるインパクトも強いと想像できるため、より積極的にそちらを平行に進めることができればよりよかったかもしれない。予定していた研究集会や対面での議論がコロナ禍においてキャンセルとなり、十分に予算を執行できず、他の研究者と議論する機会が減ってしまったことも、応用研究への推進力に対して、負に働いてしまったと考えられる。

一方で、当初の研究計画にはなかった、等質空間上の q 指数型分布族という未開拓な分野があることに気づき、実際に成果が一部上がりつつある。これにより、指数型分布族では扱えない現象を扱う方法の一つを与えることができる。しかしながら、この分野においてまだまだすべきことは多く残されているので、今後数学的なバックグラウンドを持った研究者が参入することが期待される。

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数: ●件

1. Koichi Tojo, Taro Yoshino, Harmonic exponential families on homogeneous spaces, Information Geometry, 2021, 4, 215-243..
等質空間 G/H 上に表現論を用いて分布族を構成する手法を導入し、その手法によって得られる分布族が G 不変指数型分布族であること、および G 不変指数型分布族はこの手法によって得られる分布族の部分族として実現できることを示した。
2. Koichi Tojo, Taro Yoshino, An exponential family on the upper half plane and its conjugateprior, Springer Proceedings in Mathematics & Statistics, 2021, vol 361, 84-95..
上半平面上に提案手法を用いて指数型分布族を具体的に構成し、さらにその共役事前分布族を明示的な形で与えた。

(2) 特許出願

研究期間全出願件数: 0 件(特許公開前のものも含む)

1	発 明 者	〇〇 〇〇
	発 明 の 名 称	〇〇〇〇
	出 願 人	〇〇大学

	出 願 日	201x/xx/xx
	出 願 番 号	Xxxxxxx
	概 要	100 字程度
2	発 明 者	〇〇 〇〇
	発 明 の 名 称	〇〇〇〇
	出 願 人	〇〇大学
	出 願 日	201x/xx/xx
	出 願 番 号	Xxxxxxx
	概 要	100 字程度

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

国際カンファレンスである Joint Structures and Common Foundations of Statistical Physics, Information Geometry and Inference for Learning (SPIGL'20)において keynote speaker として講演した。