表現論を用いた指数型分布族の構成法

東條 広一

理化学研究所革新知能統合研究センター汎用基盤技術研究グループ 数理科学チーム,

慶應義塾大学理工学部数理科学科

2022 年 4 月 20 日 (水)

本発表は吉野太郎氏 (東京大学大学院数理科学研究科) との共同研究に基づく.

自己紹介

ポスドクです．

氏名：東條 広一
（とうじょう　こういち）

所属：理研 AIP 数理科学チーム

（勤務地：慶應義塾大学矢上キャンパス）

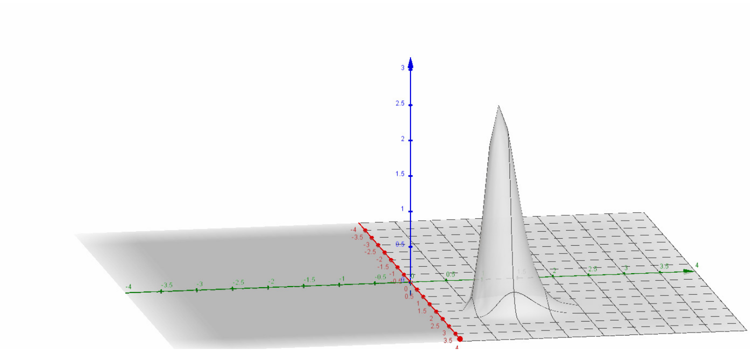
専門分野：数学のリー群論，表現論
空間の “対称性” を記述する学問

趣味：水泳，歌うこと，ダジャレ, 将棋（マイブーム）

デモ：上半平面上の指数型分布族

目標

良い分布族を系統的に扱う枠組みを与え，実際に良い分布族を提供し，その応用手法を与えたい．



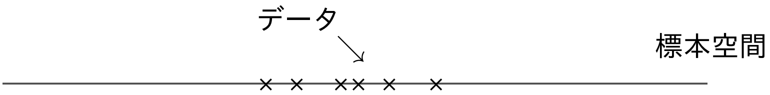
上半平面

Contents

背景 (機械学習分野)

問題

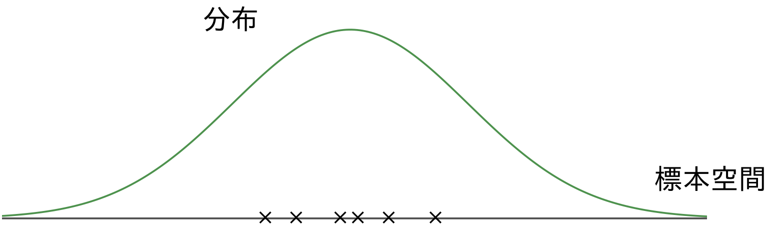
データがある分布に従って生成されているとする．与えられたデータからこの分布を推定せよ．



背景 (機械学習分野)

問題

データがある分布に従って生成されているとする．与えられたデータからこの分布を推定せよ．



分布族と学習

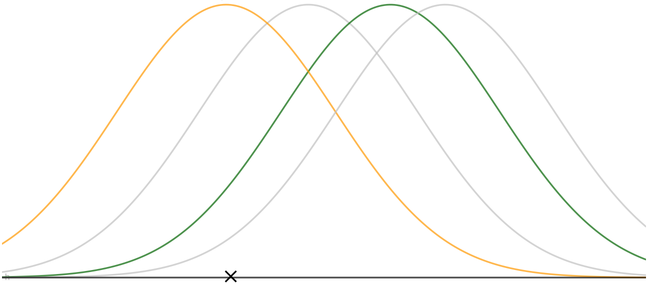
主要な手法の一つ：分布族を用いてデータから学習し近似する

分布族＝パラメトライズされた分布の族

学習＝分布族において最適なパラメータを探す行為

あるパラメータにおける分布 ↓

真の分布 ↓

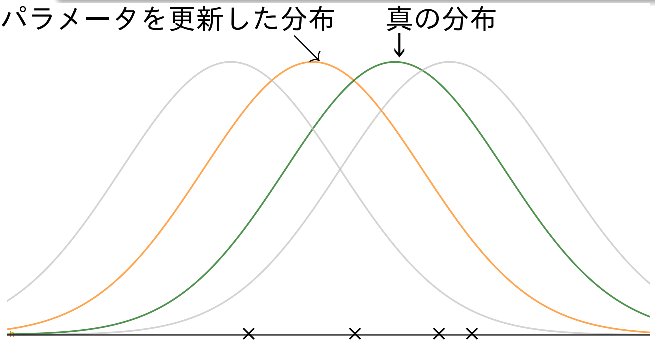


分布族と学習

主要な手法の一つ：分布族を用いてデータから学習し近似する

分布族＝パラメトライズされた分布の族

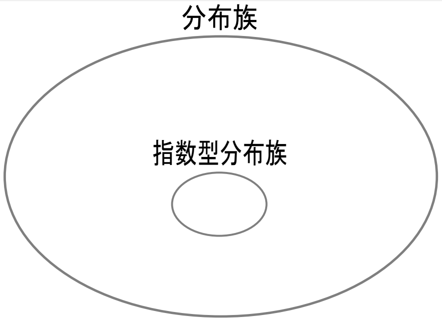
学習＝分布族において最適なパラメータを探す行為



重要な分布族：指数型分布族

指数型分布族

- 情報幾何の分野で重要な研究対象
- ベイズ推定で有用
- よく使われる分布族の多くを含む



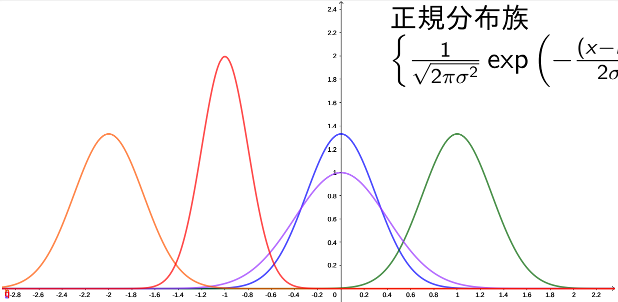
指数型分布族の例

例 1.1.

- $\mathbb{R}$ 上の正規分布族
- $\{\pm 1\}$ 上のベルヌーイ分布族
- $\mathbb{R}_{>0}$ 上のガンマ分布族
- $\mathrm{Sym}^{+}(n, \mathbb{R})$ 上のウィシャート分布族
- S^1 上のフォンミーゼス分布族

正規分布族

$$\left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2} \right) dx \right\}_{(\sigma, m) \in \mathbb{R}_{>0} \times \mathbb{R}}$$



指数型分布族の定義

X: 多様体, $\mathcal{R}(X)$: X 上の (ラドン) 測度全体の集合.

定義 1.2 (指数型分布族).

$\emptyset \neq \mathcal{P} \subset \mathcal{R}(X)$ が，X 上の指数型分布族 $\overset{\text{def}}{\Longleftrightarrow}$ ${}^{\exists}(\mu, V, T)$ s.t.

1. $\mu \in \mathcal{R}(X)$,
2. V は $\mathbb{R}$ 上の有限次元ベクトル空間,
3. $T : X \to V$, $x \mapsto T(x)$ は連続写像,
4. ${}^{\forall}p \in \mathcal{P}$, ${}^{\exists}\theta \in V^{\vee}$ s.t.

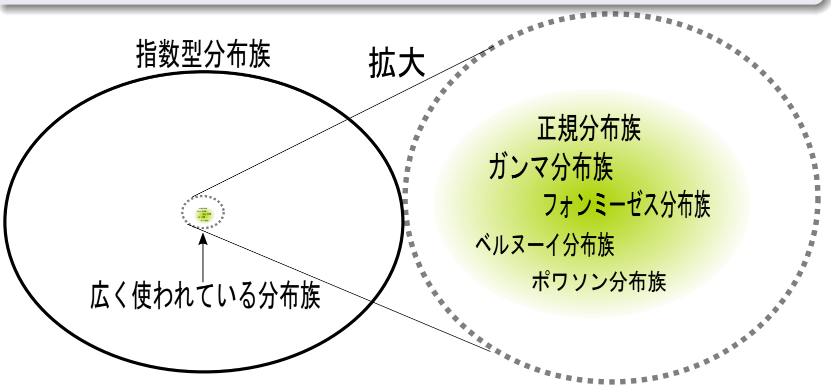
$$dp(x) = c_\theta^{-1} \exp(-\langle \theta, T(x) \rangle) d\mu(x).$$

ここで $c_\theta = \int_{x \in X} \exp(-\langle \theta, T(x) \rangle) d\mu(x)$ (正規化定数).

背景

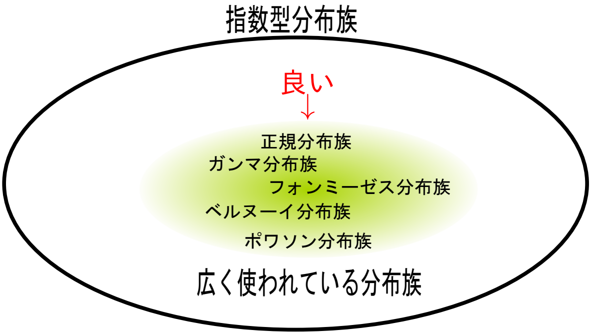
注意 1.3.

- 定義上，無数の指数型分布族がある．
- 広く使われている分布族はごく一部である．



動機・目的

“良い” 指数型分布族があると考えられる.
“良い” 指数型分布族を系統的に理解・応用する枠組みが欲しい.



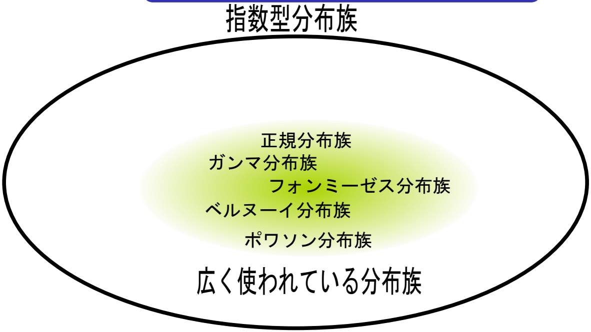
観察 1.4.

有用な指数型分布族は標本空間と同じ対称性を持っている.

- 標本空間：等質空間 G/H
- 分布族：誘導される G の作用で不変

$\rightsquigarrow$ アイデア: **表現論**を用いる

対称性のある空間=等質空間



観察 1.4.

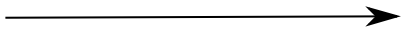
有用な指数型分布族は標本空間と同じ対称性を持っている.

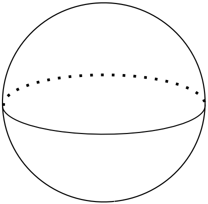
- 標本空間：等質空間 G/H
- 分布族：誘導される G の作用で不変

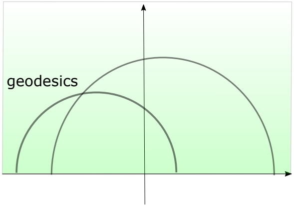
$\rightsquigarrow$ アイデア: 表現論を用いる

対称性のある空間=等質空間

$\mathbb{R} \simeq (\mathbb{R}^\times \ltimes \mathbb{R})/\mathbb{R}^\times$ $\quad S^2 \simeq SO(3)/SO(2)$ $\quad \mathcal{H} \simeq SL(2,\mathbb{R})/SO(2)$







本手法 (*G* / *H*-method)

表現論を用いて指数型分布族を構成する手法を提案した.

- よく知られている分布族の多くを生成する.
- この手法によって得られる分布族は列挙可能である.
- 対称性を持つ空間に系統的に分布族を提供できる.

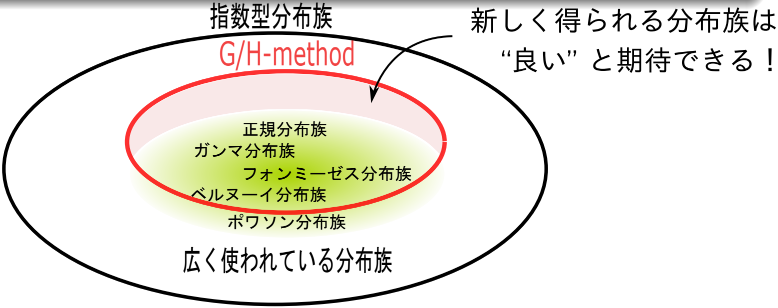
吉野太郎氏との共同研究:

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

本手法 (G/H-method)

表現論を用いて指数型分布族を構成する手法を提案した.

- よく知られている分布族の多くを生成する.
- この手法によって得られる分布族は列挙可能である.
- 対称性を持つ空間に系統的に分布族を提供できる.



吉野太郎氏との共同研究:

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

G/H-method により得られる例

表: 例とその入力 (G, H, V, v_0)

分布族	標本空間 X	G	H	V	v_0
正規	$\mathbb{R}$	$\mathbb{R}^\times \ltimes \mathbb{R}$	$\mathbb{R}^\times$	$\mathrm{Sym}(2,\mathbb{R})$	E_{22}
多変量正規	$\mathbb{R}^n$	$GL(n,\mathbb{R}) \ltimes \mathbb{R}^n$	$GL(n,\mathbb{R})$	$\mathrm{Sym}(n+1,\mathbb{R})$	$E_{n+1,n+1}$
ベルヌーイ	$\{\pm 1\}$	$\{\pm 1\}$	$\{1\}$	$\mathbb{R}_{sgn}$	1
カテゴリカル	$\{1,\cdots,n\}$	$\mathfrak{S}_n$	$\mathfrak{S}_{n-1}$	W	w
ガンマ	$\mathbb{R}_{>0}$	$\mathbb{R}_{>0}$	$\{1\}$	$\mathbb{R}$	1
逆ガンマ	$\mathbb{R}_{>0}$	$\mathbb{R}_{>0}$	$\{1\}$	$\mathbb{R}_{-1}$	1
ウィシャート	$\mathrm{Sym}^+(n,\mathbb{R})$	$GL(n,\mathbb{R})$	$O(n)$	$\mathrm{Sym}(n,\mathbb{R})$	I_n
フォンミーゼス	S^1	$SO(2)$	$\{I_2\}$	$\mathbb{R}^2$	e_1

Here $W = \{(x_1,\cdots,x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 0\}$,
$w = (-(n-1), 1, \cdots, 1) \in W$.

G/H-method: 概要

G/H-method (T–, Yoshino, 2018)

次の入力から等質空間 $X := G/H$ 上の分布族を出力する手法

- 有限次元実表現 $\rho : G \to GL(V)$,
- 非ゼロ H 不変ベクトル $v_0 \in V^H$.

手続き:

1. X 上の非ゼロ相対 G 不変測度全体の集合 $\Omega(G, H)$ を考える.
2. $V^\vee \times \Omega(G, H)$ でパラメトライズされる X 上の測度の族を上の入力に応じて作る.
3. それらを正規化して分布族を得る.

G/H-method: 標本空間

以下，正規分布族を例にとって G/H-method による分布族の構成法を述べる．

表現論:

設定

G: リー群,
H: G の閉部分群,
$X := G/H$: 等質空間.

X 上の確率測度の族
$\mathcal{P} := \{p_\theta\}_{\theta \in \Theta}$ を構成する

正規分布族:

設定

$G = \mathbb{R}^\times \ltimes \mathbb{R}$
(スケーリングと平行移動に対応),
$H = \mathbb{R}^\times$,
$X = G/H \simeq \mathbb{R}$.

G/H と $\mathbb{R}$ を以下で同一視する

$$G/H \ni (t, x)H \mapsto x \in \mathbb{R}$$

G/H-method: 入力

表現論:

入力

1. V: 有限次元実線型空間
2. 表現 $\rho: G \to GL(V)$
3. H 不変ベクトル $v_0 \in V^H$

正規分布族:

入力

1. $V := \mathrm{Sym}(2, \mathbb{R})$
2. $\rho: \mathbb{R}^{\times} \ltimes \mathbb{R} \to GL(\mathrm{Sym}(2, \mathbb{R}))$
   $$\rho(t, x)S := \begin{pmatrix} t & x \\ & 1 \end{pmatrix} S \begin{pmatrix} t & \\ x & 1 \end{pmatrix}$$
3. $v_0 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in V^H$

G/H-method: 相対 G 不変測度

定義 2.1.

測度 $\mu \in \mathcal{R}(X)$ が相対 G 不変
$\overset{\text{def}}{\Longleftrightarrow}$ ${}^{\exists}\chi : G \to \mathbb{R}_{>0}$ 連続写像 s.t. $d\mu(gx) = \chi(g)d\mu(x)$.

表現論:
$\Omega(G, H) := \{\mu \in \mathcal{R}(X) \mid$ μは非ゼロ相対 G 不変 $\}$ とおく.

正規分布族:
$\Omega(G, H) = \{\lambda\, dx \mid \lambda > 0\}$. ここで dx はルベーグ測度 (これは相対 G 不変).

G/H-method: X 上の測度 $\tilde{p}_\theta$

表現論:
$(\xi, \mu) \in V^\vee \times \Omega(G, H)$ でパラメトライズされる X 上の測度を次で定める.

$$
\begin{aligned}
d\tilde{p}_\theta(x) &= d\tilde{p}_{\xi,\mu}(x) \\
&:= \exp(-\langle \xi, x v_0 \rangle) d\mu(x) \\
&\quad (x \in X)
\end{aligned}
$$

正規分布族:
V 上の内積
$(S_1, S_2) := \mathrm{trace}(S_1 S_2)$ によって
$V^\vee$ と $V = \mathrm{Sym}(2, \mathbb{R})$ を同一視.
$d\tilde{p}_\theta(x) =$

$$
\exp\left(-\mathrm{trace}\left(\begin{pmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_3 \end{pmatrix} \rho(t, x) v_0\right)\right) \lambda dx
$$

$= \exp(-(\theta_1 x^2 + 2\theta_2 x + \theta_3)) \lambda dx$

G/H-method: 測度 $\tilde{p}_\theta$ の正規化

表現論:
$\Theta := \{\theta = (\xi, \mu) \in V^\vee \times \Omega(G, H) \mid \int_X d\tilde{p}_\theta < \infty\}$,
$p_\theta := c_\theta^{-1} \tilde{p}_\theta$,
$c_\theta := \int_X d\tilde{p}_\theta$ (正規化定数).
確率測度の族 $\{p_\theta\}_{\theta \in \Theta}$ を得る.

正規分布族:
$\Theta = \{\theta = (\begin{pmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_3 \end{pmatrix}, \lambda) \in \mathrm{Sym}(2, \mathbb{R}) \times \mathbb{R}_{>0} \mid \theta_1 > 0\}$,
$dp_\theta(x) = \sqrt{\frac{\theta_1}{\pi}} \exp(-\theta_1 (x + \frac{\theta_2}{\theta_1})^2) dx$,
$c_\theta = \lambda \sqrt{\frac{\pi}{\theta_1}} \exp(\frac{\theta_2^2 - \theta_1 \theta_3}{\theta_1})$.

G/H-method: 正規分布族

以下の $\mathbb{R}$ 上の確率測度の族を得る．

$$\left\{ \sqrt{\frac{\theta_1}{\pi}} \exp\left(-\theta_1 \left(x + \frac{\theta_2}{\theta_1} \right)^2 \right) dx \right\}_{(\theta_1, \theta_2) \in \mathbb{R}_{>0} \times \mathbb{R}}.$$

パラメータの変数変換

$$m = -\frac{\theta_2}{\theta_1}, \quad \sigma = \frac{1}{\sqrt{2\theta_1}},$$

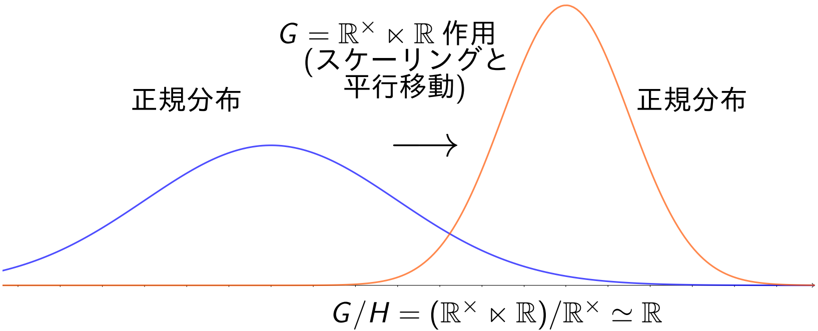
によって，以下の正規分布族を得る．

$$\left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2} \right) dx \right\}_{(\sigma, m) \in \mathbb{R}_{>0} \times \mathbb{R}}.$$

(性質) 得られる分布族は G 不変指数型分布族

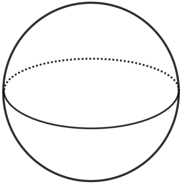
命題 2.2.

本手法で得られる G/H 上の分布族は G 不変指数型分布族である.



等質空間 G/H の持つ構造・対称性を反映した分布族が得られる.

系統的に良い分布族が得られる例 1: 球面



$G = SO(n+1)$, $H = SO(n)$,
$X := G/H \simeq S^n$: n 次元球面
低次元表現の場合: 有名分布族が復元可

- $\rho : SO(n+1) \to GL(\mathbb{R}^{n+1})$ 自然表現
  $\rightsquigarrow$ フォンミーゼス・フィシャー分布族 (発見的)
- $\rho : SO(n+1) \to GL(\mathbb{R}^{n+1} \oplus \mathrm{Sym}(n+1, \mathbb{R}))$
  $\rho(g)(v, S) = (gv, gS{}^tg)$ $((v, S) \in \mathbb{R}^{n+1} \times \mathrm{Sym}(n+1, \mathbb{R}))$.
  $\rightsquigarrow$ フィッシャー・ビンガム分布族 (発見的)

より高次元の表現の場合も，現時点ではあまり有名ではないが，本手法により分布族が構成できる．T. S. Cohen, M. Welling による研究 [CW15] で同じ分布族が現れる．

関連研究: [CW15]

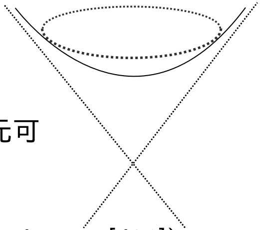
[CW15] T. S. Cohen, M. Welling, *Harmonic exponential families on manifolds*, In Proceedings of the 32nd International Conference on Machine Learning(ICML), volume 37 (2015), 1757–1765

1. コンパクト群の等質空間，主に $S^1 \simeq SO(2)$, $S^2 \simeq SO(3)/SO(2)$ 上の "調和指数型分布族" を提案．
2. S^1, S^2 上の高速フーリエ変換を用いた勾配法で最尤法を実現．
3. 地球を球面 S^2 とみなし，主要な地震の分布問題に応用．

本手法は [CW15] の自然な拡張とみなせる

[CW15] で提案された分布族は，G/H-method で G をコンパクト群とした特別な場合として得られる．

系統的に良い分布族が得られる例 2: 双曲空間

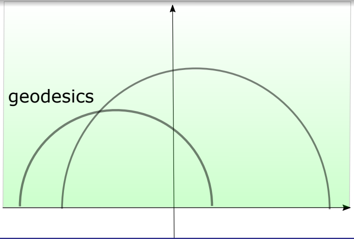


$G = SO_0(n,1)$, $H = SO(n)$,
$X := G/H \simeq H^n$: n 次元双曲空間

- 低次元表現の場合: 有名分布族が復元可
  $\rho : G \to GL(\mathbb{R}^{n+1})$ 自然表現
  $\rightsquigarrow$ ハイパボロイド分布族
  (O. E. Barndorff-Nielsen [BN78], J. L. Jensen [J81])
- より高次元の表現の場合も，本手法により新しい分布族が得られる

系統的に良い分布族が得られる例 3: 上半平面

$G = SL(2,\mathbb{R}),\ H = SO(2),$
$X := G/H \simeq \mathcal{H}$:上半平面,
低次元表現の場合:
$\rho : SL(2,\mathbb{R}) \to GL(\mathrm{Sym}(2,\mathbb{R})),$
$\rho(g)S = gS{}^tg \quad (S \in \mathrm{Sym}(2,\mathbb{R})).$



Poincaré 分布族 (冒頭のデモで示した分布族)

$$\left\{ \frac{De^{2D}}{\pi} \exp\left(- \frac{a(x^2+y^2)+2bx+c}{y} \right) \frac{dxdy}{y^2} \right\}_{\begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathrm{Sym}^+(2,\mathbb{R})}$$

ここで $D = \sqrt{ac-b^2}$.

より高次元表現の場合も，本手法により新しい分布族が得られる
これは双曲空間の特別な場合である.

$\mathbb{R}$ 上のスケーリング・平行移動不変分布族

$\mathbb{R}$ 上の正規分布族は $\mathbb{R}^\times \ltimes \mathbb{R}$(スケーリングと平行移動) 不変.

Question 2.3.

$\mathbb{R}$ 上の $\mathbb{R}^\times \ltimes \mathbb{R}$ 不変な指数型分布族は他にあるか？

$\mathbb{R}$ 上のスケーリング・平行移動不変分布族

$\mathbb{R}$ 上の正規分布族は $\mathbb{R}^\times \ltimes \mathbb{R}$(スケーリングと平行移動) 不変.

Question 2.3.

$\mathbb{R}$ 上の $\mathbb{R}^\times \ltimes \mathbb{R}$ 不変な指数型分布族は他にあるか？ $\rightsquigarrow$ ある.

多項式指数型分布族 (polynomial exponential family)

以下で与えられる指数型分布族は $\mathbb{R}$ 上の多項式指数型分布族と呼ばれ, $\mathbb{R}^\times \ltimes \mathbb{R}$ 不変である.

$$\left\{ c_\theta^{-1} \exp\left(-\sum_{i=1}^{2n} \theta_i x^i \right) dx \right\}_{\theta \in \Theta}$$

Question 2.4.

多項式指数型分布族は G/H-method で得られるか？

$\mathbb{R}$ 上のスケーリング・平行移動不変分布族

$\mathbb{R}$ 上の正規分布族は $\mathbb{R}^\times \ltimes \mathbb{R}$(スケーリングと平行移動) 不変.

Question 2.3.

$\mathbb{R}$ 上の $\mathbb{R}^\times \ltimes \mathbb{R}$ 不変な指数型分布族は他にあるか？ $\rightsquigarrow$ ある.

多項式指数型分布族 (polynomial exponential family)

以下で与えられる指数型分布族は $\mathbb{R}$ 上の多項式指数型分布族と呼ばれ，$\mathbb{R}^\times \ltimes \mathbb{R}$ 不変である.

$$\left\{ c_\theta^{-1} \exp\left(-\sum_{i=1}^{2n} \theta_i x^i \right) dx \right\}_{\theta \in \Theta}$$

Question 2.4.

多項式指数型分布族は G/H-method で得られるか？ $\rightsquigarrow$ YES!

問の一般化

Question 2.5.

G/H 上の G 不変指数型分布族は G/H-method で得られるか？

$\rightsquigarrow$ 弱い仮定の下で YES!.
$\rightsquigarrow$ ラフに言えば

$$\{G/H \text{ 上の } G \text{ 不変指数型分布族}\}$$
$$\text{“}=\text{”}\{G/H\text{-method によって得られる } G/H \text{ 上の分布族}\}$$

問に対する答え

$\mathcal{P} := \{p_\theta\}_{\theta\in\Theta}$ を G/H 上の G 不変指数型分布族とする．ここで Θ はパラメータ空間である．

定理 2.6.

次の 2 点を仮定する．

1. G/H は非ゼロ相対 G 不変測度を持つ，
2. Θ は開集合．

このとき $\mathcal{P}$ は G/H-*method* で得られる分布族の部分族として実現される．

より詳しくは，[TY20] を参照．

G 不変指数型分布族の分類問題

G/H: 球面や双曲空間などの有用な等質空間

問題 3.1.

G/H 上の G 不変指数型分布族を分類せよ.

前頁の定理より，上記問題は以下の Question に帰着される.

Question 3.2.

G/H-method を用いて得られる G/H 上の分布族を分類せよ.

分類の現状

現状の結果

G/H が以下の場合に分布族の分類ができた．

- $\mathbb{R} \simeq \mathbb{R}_{>0} \ltimes \mathbb{R}/\mathbb{R}_{>0}$,
- S^1, S^2.

$G/H = (\mathbb{R}_{>0} \ltimes \mathbb{R})/\mathbb{R}_{>0} \simeq \mathbb{R}$ のとき

G/H-method によって得られる分布族は，多項式指数型分布族のみ.

現在，双曲空間を含む既約リーマン対称空間のクラスで分類問題に取り組み中．

分類結果の例 (informal)

$G/H = SO(2) \simeq S^1$ のとき

G/H-method を用いて得られる分布族は以下の形：

$$\{(\text{正規化定数}) \exp\left(-\sum_{k=1}^{n} \theta_{k,1} \cos kx + \theta_{k,2} \sin kx\right) d\mu(x)\}$$

ここで μ は S^1 上の $SO(2)$ 不変な測度 (一様測度).

$G/H = SO(3)/SO(2) \simeq S^2$ のとき

G/H-method を用いて得られる分布族は以下の形：

$$\{(\text{正規化定数}) \exp(\text{球面調和関数たちの線形和}) d\mu\}$$

ここで μ は S^2 上の $SO(3)$ 不変な測度 (一様測度).

参考文献

[BN78] O. E. Barndorff-Nielsen, *Hyperbolic distributions and distribution on hyperbolae*, Scan. J. Stat. **8** (1978), 151–157.

[CW15] T. S. Cohen, M. Welling, *Harmonic exponential families on manifolds*, In Proc. of the 32nd Int. Conf. on Machine Learning(ICML), volume 37 (2015), 1757–1765

[J81] J. L. Jensen, *On the hyperboloid distribution*, Scand. J. Statist. **8** (1981), 193–206.

[TY18] K. Tojo, T. Yoshino, *A method to construct exponential families by representation theory*, arXiv:1811.01394v3.

[TY19] K. Tojo, T. Yoshino, *On a method to construct exponential families by representation theory*, Geometric Science of Information. GSI2019, Lecture Notes in Computer Science, vol 11712, 147–156 (2019).

[TY20] K. Tojo, T. Yoshino, *Harmonic exponential families on homogeneous spaces*, Info. Geo. (2020). https://doi.org/10.1007/s41884-020-00033-3