表現論を用いた指数型分布族の構成法

理化学研究所革新知能統合研究センター　東條 広一
東京大学大学院数理科学研究科　吉野 太郎
Koichi Tojo
RIKEN Center for Advanced Intelligence Project,
Taro Yoshino
Graduate School of Mathematical Science, The University of Tokyo

目次

1 導入

本節では，表現論を用いた指数型分布族の構成法を考える動機とその手法について述べる．用いる概念についての定義は § 1.3 を参照されたい．

1.1 動機

指数型分布族 (Definition 1.6) は情報幾何学において重要な役割を果たす．また共役事前分布族 (Definition 3.1) を許容するためベイズ推定においてもよく用いられる．ベルヌーイ分布族，正規分布族，ガンマ分布族，フォンミーゼス分布族，ウィシャート分布族などは指数型分布族の有名な例である．一方で，定義から指数型分布族は無数にあるが，これら広く使われている分布族は，そのごく一部に過ぎない．それらは個々にはよく調べられているが，我々はそれらを系統的に理解したい．そのためには次のような性質を満たす手法があ

ることが望ましい：

(i) その手法は多くの良く知られた分布族を生成しうる．
(ii) その手法によって得られる分布族は限られており，列挙可能である．

その手法を考える上で次の観察が有用である：

- 広く使われる分布族の多くは標本空間が等質空間 G/H とみなせる，
- その分布族は G 不変 (Definition 1.9) である．

そこで性質 (i), (ii) を満たす手法として [TY18] において表現論を用いて等質空間上に分布族を構成する手法を提案した (Method 1.1)．本稿では，この手法とその性質についてまとめる．ただし，主張の証明はここには記さないため，詳しくは参考文献に提示した論文を参照されたい．また，最後の節では，応用として上半平面上のベイズ推定を可能にするために，上半平面上のポアンカレ分布族 (Example 2.6) に対する共役事前分布族を明示的に提示する．

1.2 提案手法 (G/H-method) とその性質

G を連結成分有限個のリー群, H を G の閉部分群とする, このとき商空間 $X := G/H$ は自然に多様体の構造を持ち, G の等質空間と呼ばれる．X を標本空間とみなす．X 上には非ゼロ相対 G 不変測度 (Definition 1.10) が存在すると仮定する．

Method 1.1 ([TY18], G/H-method)**.** G/H-method は次の二つの入力から X 上の分布族 $\{p_\theta\}_{\theta\in\Theta}$ を生成する手法である．

(i) 有限次元実表現 $\rho\colon G \to GL(V)$,
(ii) H 不変ベクトル $v_0 \in V$.

まず次のベクトル空間を考える．

$$W_0(G,H) := \{\tau\colon G \to \mathbb{R} \mid \tau\text{は連続群準同型写像で}\tau|_H = 0\text{ を満たす }\}$$

この空間は条件 $\tau|_H = 0$ から X 上の連続関数全体の集合の部分集合とみなせる．次に X 上の非ゼロ相対 G 不変測度 $\nu \in \mathcal{R}(X)$ を一つ固定する．ここで $\mathcal{R}(X)$ は X 上のラドン測度全体の集合を表す．各 $\theta := (\xi,\tau) \in V^\vee \oplus W_0(G,H)$ に対し，X 上の測度 $\tilde{p}_\theta$ を次にように構成する．

$$d\tilde{p}_\theta(x) := \exp(-\langle \xi, x\cdot v_0\rangle + \tau(x))d\nu(x).$$

ここで，$V^\vee$ は V の双対空間を表し，$\langle\cdot,\cdot\rangle$ は $V^\vee \times V$ 上のペアリングを表す．さらにこれらの測度 $\tilde{p}_\theta$ を次のように正規化する：

$$\begin{aligned}
\Theta &:= \left\{\theta = (\xi,\tau) \in V^\vee \oplus W_0(G,H) \;\middle|\; \int_X d\tilde{p}_\theta < \infty\right\},\\
c_\theta &:= \int_X d\tilde{p}_\theta \quad (\theta \in \Theta),\\
p_\theta &:= c_\theta^{-1}\tilde{p}_\theta \quad (\theta \in \Theta).
\end{aligned}$$

結果として，X 上の確率測度の族 $\mathcal{P} := \{p_\theta\}_{\theta\in\Theta}$ を得る．

$W_0(G,H) = \{0\}$ のとき Θ を $V^\vee$ の部分集合とみなせる．

Remark 1.2. 得られた分布族 $\mathcal{P}$ は相対 G 不変測度 ν の選び方に依らない．実際，他の X 上の非ゼロ相対 G 不変測度 ν' に対し，次の事実から $c \in \mathbb{R}_{>0}$ と $\tau \in W_0(G, H)$ で $d\nu'(x) = ce^{\tau(x)}d\nu(x)$ となるものが取れる．

Fact 1.3 (See [TY21a, Corollary 4.27] for example)**.** G/H 上には非ゼロ相対 G 不変測度が存在すると仮定し，その一つを取り ν_0 とおく．このとき G/H 上の非ゼロ相対 G 不変測度全体の集合は次で与えられる．

$$\{ce^{\tau}\nu_0 \mid c \in \mathbb{R}_{>0}, \tau \in W_0(G, H)\}.$$

Theorem 1.4 ([TY21a, Theorem 3.2])**.** G/H-method を用いて得られた分布族は X 上の G 不変指数型分布族である．

正規分布族は $G := \mathbb{R}^{\times} \ltimes \mathbb{R}$, $H := \mathbb{R}^{\times}$ として G/H-method を用いて構成することができる．このとき上記定理は，正規分布をスケーリングと平行移動によって変形した分布もまた正規分布であることを主張していることに相当する．

逆に X 上の G 不変指数型分布族は G/H-method で次の意味で構成できる：

Theorem 1.5 ([TY21a, Theorem 3,7])**.** $\mathcal{P}$ を X 上の G 不変指数型分布族とし，(μ, V, T) を $\mathcal{P}$ の最小次元実現とする．$\Theta \subset V^{\vee}$ をパラメータ空間とする．さらに次を仮定する．

(i) G/H 上には非ゼロ相対 G 不変測度が存在する．
(ii) Θ は $V^{\vee}$ の開集合である．

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できる．

1.3 用語

この節では，本稿で用いる概念の定義をまとめておく．

X を多様体とし，$\mathcal{R}(X)$ を X 上のラドン測度全体の集合とする．

Definition 1.6 ([BN70, §5], 指数型分布族)**.** 確率測度からなる部分集合 $\mathcal{P} \subset \mathcal{R}(X)$ が X 上の**指数型分布族**であるとは，次の 4 条件を満たす三つ組 (μ, V, T) が存在することをいう．

(i) $\mu \in \mathcal{R}(X)$,
(ii) V は有限次元実ベクトル空間である，
(iii) $T \colon X \to V$ は連続写像である，
(iv) 任意の $p \in \mathcal{P}$ に対し，ある $\theta \in V^{\vee}$ が存在して次が成り立つ．

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

ここで $c_{\theta} = \int_{x \in X} \exp(-\langle \theta, T(x) \rangle) d\mu(x)$ であり，c_{θ} は**正規化定数**と呼ばれる．このとき三つ組 (μ, V, T) を指数型分布族 $\mathcal{P}$ の**実現**と呼ぶ．実現の中で V の次元が最も小さいものを**最小次元実現**と呼ぶ．

Remark 1.7. 統計の分野では三つ組 (μ, V, T) を表現 (representation) と呼ぶことが多いが，ここでは群の表現と区別するために実現と呼ぶことにする．

[BN70] では指数型分布族の概念は可測空間と可測写像のカテゴリで定義されているが，本稿では，位相空間と連続写像のカテゴリで考える．

以下，G をリー群，H をその閉部分群とし $X := G/H$ を G の等質空間とする．

Definition 1.8. $\mathcal{R}(X)$ への (左)G 作用を以下で定義する.

$$G \times \mathcal{R}(X) \to \mathcal{R}(X),\ (g, \mu) \mapsto g \cdot \mu,$$
$$(g \cdot \mu)(B) := \mu(g^{-1}B) \quad (B \in \mathcal{B}(X)).$$

ここで $\mathcal{B}(X)$ は X 上のボレル集合族を表す. 言い換えれば, $g \cdot \mu$ は写像 $g\cdot\colon X \to X,\ x \mapsto gx$ による μ の押し出し測度である.

Definition 1.9 (G 不変)**.** 部分集合 $\mathcal{P} \subset \mathcal{R}(X)$ が G 不変であるとは, 任意の $g \in G$ と $p \in \mathcal{P}$ に対し, $g \cdot p \in \mathcal{P}$ が成り立つことをいう.

Definition 1.10 (相対 G 不変測度)**.** 測度 $\mu \in \mathcal{R}(X)$ が相対 G 不変であるとは, 正値関数 $\chi\colon G \to \mathbb{R}_{>0}$ で次が成り立つものが存在することをいう:

$$\frac{d(g \cdot \mu)}{d\mu}(x) = \chi(g)^{-1} \quad (g \in G, x \in X).$$

2 具体例とその構成

本節では, 広く用いられている分布族が G/H-method で構成できることを紹介し, そのいくつかを実際に Method 1.1 に基づいて構成する.

2.1 得られる分布族の例

リー群 G, H と G の表現 V, H 不変ベクトル v_0 を表のように選ぶことで G/H-method によって次の表にある分布族を構成することができる. それぞれ詳しい構成方法については [TY18] を参照されたい. ただし, 一般化逆ガウス分布族については [TY19, §3] を参照.

表 1 G/H-method で得られる分布族の例と入力 (G, H, V, v_0)

分布族	標本空間	G	H	V	v_0
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
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	$\mathbb{R}^n$	$GL(n, \mathbb{R}) \ltimes \mathbb{R}^n$	$GL(n, \mathbb{R})$	$\mathrm{Sym}(n+1, \mathbb{R})$	$E_{n+1,n+1}$
ガンマ	$\mathbb{R}_{>0}$	$\mathbb{R}_{>0}$	$\{1\}$	$\mathbb{R}_1$	1
逆ガンマ	$\mathbb{R}_{>0}$	$\mathbb{R}_{>0}$	$\{1\}$	$\mathbb{R}_{-1}$	1
一般化逆ガウス	$\mathbb{R}_{>0}$	$\mathbb{R}_{>0}$	$\{1\}$	$\mathbb{R}_1 \oplus \mathbb{R}_{-1}$	$\frac{1}{2}(e_1 + e_2)$
正規逆ガンマ	$\mathbb{R}_{>0} \ltimes \mathbb{R}$	$\mathbb{R}_{>0} \ltimes \mathbb{R}$	$\{(1, 0)\}$	$\mathbb{R}^4$	$e_3 + e_4$
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
フォンミーゼス–フィッシャー	S^{n-1}	$SO(n)$	$SO(n-1)$	$\mathbb{R}^n$	e_1
フィッシャー–ビンガム	S^{n-1}	$SO(n)$	$SO(n-1)$	$\mathbb{R}^n \oplus \mathrm{Sym}(n, \mathbb{R})$	$e_1 \oplus E_{11}$
ハイパボロイド	H^n	$SO_0(1, n)$	$SO(n)$	$\mathbb{R}^{n+1}$	e_0
ポアンカレ	$\mathcal{H}$	$SL(2, \mathbb{R})$	$SO(2)$	$\mathrm{Sym}(2, \mathbb{R})$	I_2

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

次の節では以下の 4 つの分布族が実際に G/H-method で構成できることを見る．

(i) $\mathbb{R}$ 上の正規分布族,
(ii) $\mathbb{R}$ 上のガンマ分布族,
(iii) 球面 S^{n-1} 上のフォンミーゼス–フィッシャー分布族,
(iv) 上半平面上のポアンカレ分布族.

次の Fact 2.1 は例を構成する際に $W_0(G,H)$ を決定するために用いられる．

Fact 2.1 ([TY21a, Proposition 4.28])**.** G を連結成分有限個のリー群とし，H を G の閉部分群とする．$\mathfrak{g}$, $\mathfrak{h}$ をそれぞれ G, H のリー代数をとする．このとき，次の条件のいずれかが成り立つならば，$W_0(G,H)=\{0\}$ である．

(i) $\mathfrak{g}=\mathfrak{h}+[\mathfrak{g},\mathfrak{g}]$,
(ii) G はコンパクト，
(iii) G は半単純.

2.2 $\mathbb{R}$ 上の正規分布族の構成

Example 2.2 (正規分布族)**.** $G:=\mathbb{R}^\times\ltimes\mathbb{R}$, $H:=\mathbb{R}^\times\subset G$ とおき，$X:=G/H$ を G の等質空間とする．有限次元実表現 $\rho\colon G\to GL(\mathrm{Sym}(2,\mathbb{R}))$ を次で定める．

$$\rho(g)S:=\begin{pmatrix}a & b\\ 0 & 1\end{pmatrix}S\,{}^t\!\begin{pmatrix}a & b\\ 0 & 1\end{pmatrix}\quad (g=(a,b)\in\mathbb{R}^\times\ltimes\mathbb{R}, S\in\mathrm{Sym}(2,\mathbb{R})).$$

$v_0:=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\in\mathrm{Sym}(2,\mathbb{R})$ とおくと v_0 は H 不変ベクトルであり，この ρ, v_0 に対し，G/H-method を適用すると $\mathbb{R}$ 上の正規分布族

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

が得られる．以下，実際に G/H-method を適用する．X は次の同型を通じて $\mathbb{R}$ と同一視されることに注意する．

$$X\to\mathbb{R},\ gH\mapsto b\quad (g=(a,b)\in G).$$

Fact 2.1 (i) より $W_0(G,H)=\{0\}$ となる．以下の $\mathrm{Sym}(2,\mathbb{R})$ 上の内積によって $\mathrm{Sym}(2,\mathbb{R})^\vee$ を $\mathrm{Sym}(2,\mathbb{R})$ と同一視する．

$$\langle x,y\rangle:=\mathrm{trace}(xy)\quad (x,y\in\mathrm{Sym}(2,\mathbb{R})).\tag{2}$$

X 上の非ゼロ相対 G 不変測度としてルベーグ測度 dx を取る．このとき $\theta=\begin{pmatrix}\theta_1 & \theta_2\\ \theta_2 & \theta_3\end{pmatrix}\in\mathrm{Sym}(2,\mathbb{R})$ に対して，X 上の測度 $\tilde{p}_\theta$ は次のように定まる．

$$\begin{aligned}d\tilde{p}_\theta&=\exp(-\mathrm{trace}(\theta\rho(x)v_0)dx\\&=\exp\left(-\mathrm{trace}\left(\begin{pmatrix}\theta_1 & \theta_2\\ \theta_2 & \theta_3\end{pmatrix}\begin{pmatrix}x^2 & x\\ x & 1\end{pmatrix}\right)\right)dx\\&=\exp(-(\theta_1x^2+2\theta_2x+\theta_3))dx.\end{aligned}$$

このときパラメータ空間 Θ と正規化定数は簡単な計算により次で与えられる．

$$\begin{aligned}\Theta &= \left\{\theta = \begin{pmatrix}\theta_1 & \theta_2\\ \theta_2 & \theta_3\end{pmatrix} \in \mathrm{Sym}(2,\mathbb{R}) \;\middle|\; \int_{\mathbb{R}} d\tilde{p}_\theta < \infty\right\}\\ &= \{\theta \in \mathrm{Sym}(2,\mathbb{R}) \mid \theta_1 > 0\},\\ c_\theta &= \sqrt{\frac{\pi}{\theta_1}} \exp\left(\frac{\theta_2^2 - \theta_1\theta_3}{\theta_1}\right).\end{aligned}$$

従って，X 上の確率測度 p_θ は次で与えられる．

$$dp_\theta(x) = c_\theta^{-1} d\tilde{p}_\theta(x) = \sqrt{\frac{\theta_1}{\pi}} \exp\left(-\theta_1\left(x + \frac{\theta_2}{\theta_1}\right)^2\right) dx \quad (\theta \in \Theta).$$

以上より，次のように $\mathbb{R}$ 上の確率測度の族が得られた．

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

さらにパラメータの変換 $\mu = -\frac{\theta_2}{\theta_1}$, $\sigma = \frac{1}{\sqrt{2\theta_1}}$ を行うことで正規分布族の良く知られた形 (1) が得られる．

2.3 $\mathbb{R}_{>0}$ 上のガンマ分布族の構成

Example 2.3 (ガンマ分布族)**.** $G := \mathbb{R}_{>0}$, $H := \{1\}$, $X := G/H = \mathbb{R}_{>0}$ とおく．有限次元実表現 $\rho\colon G \to GL(1,\mathbb{R})$ を次で定める:

$$\rho(x) := x \quad (x \in \mathbb{R}_{>0}).$$

$v_0 := 1 \in \mathbb{R}$ は H 不変ベクトルであり，この ρ, v_0 に対し，G/H-method を適用すると $\mathbb{R}$ 上のガンマ分布族

$$\left\{\frac{1}{\Gamma(k)}\left(\frac{x}{\theta}\right)^k e^{-\frac{x}{\theta}} \frac{dx}{x}\right\}_{(k,\theta)\in\mathbb{R}_{>0}\times\mathbb{R}_{>0}} \tag{3}$$

が得られる．以下, 実際に G/H-method を適用する．$\tau \in W_0(G,H)$ の連続性により $W_0(G,H) = \{\tau\colon \mathbb{R}_{>0} \ni x \mapsto \beta\log x \in \mathbb{R} \mid \beta \in \mathbb{R}\} \simeq \mathbb{R}$ とわかる．標準的な $\mathbb{R}$ 上の内積により $\mathbb{R}^\vee$ を $\mathbb{R}$ と同一視する．X 上の非ゼロ相対 G 不変測度として $\frac{dx}{x}$ を取る．ここで，dx は X 上のルベーグ測度を表す．このとき $\theta = (\alpha,\beta) \in \mathbb{R} \times W_0(G,H)$ に対し，X 上の測度 $\tilde{p}_\theta$ は次で与えられる：

$$d\tilde{p}_\theta(x) = \exp(-\alpha x + \beta\log x)\frac{dx}{x} = \exp(-\alpha x)x^\beta \frac{dx}{x}$$

このとき直接計算によりパラメータ空間 Θ と正規化定数は次のように求められる．

$$\begin{aligned}\Theta &= \left\{\theta = (\alpha,\beta) \in \mathbb{R}\times\mathbb{R} \;\middle|\; \int_{\mathbb{R}_{>0}} d\tilde{p}_\theta < \infty\right\}\\ &= \mathbb{R}_{>0} \times \mathbb{R}_{>0},\\ c_\theta &= \int_{\mathbb{R}_{>0}} d\tilde{p}_\theta = \frac{\Gamma(\beta)}{\alpha^\beta}.\end{aligned}$$

従って，X 上の確率測度 p_θ は次で与えられる．

$$dp_\theta(x) = c_\theta^{-1} d\tilde{p}_\theta = \frac{\alpha^\beta}{\Gamma(\beta)} \exp(-\alpha x)x^\beta \frac{dx}{x} \quad (\theta \in \Theta).$$

以上により，次のように $\mathbb{R}_{>0}$ 上の分布族が得られた．

$$\mathcal{P} = \left\{\frac{\alpha^\beta}{\Gamma(\beta)} \exp(-\alpha x)x^\beta \frac{dx}{x}\right\}_{(\alpha,\beta)\in\mathbb{R}_{>0}\times\mathbb{R}_{>0}}.$$

さらに，パラメータの変換 $\alpha = \frac{1}{\theta}, \beta = k$ を行うことでガンマ分布族の良く知られた形 (3) が得られる．

2.4 球面 S^{n-1} 上のフォンミーゼス–フィッシャー分布族の構成

Example 2.4 (フォンミーゼス–フィッシャー分布族)**.** $G := SO(n)$, $H := SO(n-1)$ とし，$X := G/H$ を G の等質空間とする．ただし，H は G の “ななめ下” に実現されているとする．写像 $G/H \to S^{n-1}$, $gH \mapsto ge_1$ により X と $n-1$ 次元球面を同一視する．$\rho\colon G \to GL(n,\mathbb{R})$ を自然表現とする．$v_0 := e_1 \in \mathbb{R}^n$ とおくと v_0 は H 不変ベクトルであり，この ρ, v_0 に対し，G/H-method を適用すると S^{n-1} 上のフォンミーゼス–フィッシャー分布族

$$\left\{ \frac{1}{(2\pi)^{\frac{n}{2}} \|\theta\|^{1-\frac{n}{2}} I_{\frac{n}{2}-1}(\|\theta\|)} \exp(-{}^t\theta x) dx \right\}_{\theta \in \mathbb{R}^n} \tag{4}$$

が得られる．ここで dx は，S^{n-1} 上の一様測度で $\int_{S^{n-1}} dx = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$ を満たすものであり，$I_m(r)$ は m 次第一種変形ベッセル関数を表す：

$$I_m(r) = \frac{1}{2\pi} \int_0^{2\pi} \cos(mx) e^{r \cos x} dx \quad (m \in \mathbb{Z}, r \in \mathbb{R}).$$

以下，実際に G/H-method を適用する．$G = SO(n)$ はコンパクトであるから Fact 2.1 (ii) から $W_0(G,H) = \{0\}$ である．標準内積を取ることにより $(\mathbb{R}^n)^\vee$ を $\mathbb{R}^n$ と同一視する．X 上の非ゼロ相対 G 不変測度として S^{n-1} 上の一様測度 dx で $\int_{S^{n-1}} dx = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$ となるものを取る．このとき $\theta \in \mathbb{R}^n$ に対し，X 上の測度 $\tilde{p}_\theta$ は次で与えられる．

$$d\tilde{p}_\theta(x) = \exp(-{}^t\theta x) dx$$

$X = S^{n-1}$ はコンパクトであるからパラメータ空間 Θ は $\mathbb{R}^n$ であり，正規化定数 c_θ は [BNBE89, Example 8.3] より次で与えられる．

$$c_\theta = (2\pi)^{\frac{n}{2}} \|\theta\|^{1-\frac{n}{2}} I_{\frac{n}{2}-1}(\|\theta\|).$$

以上により $\tilde{p}_\theta$ を正規化することで，フォンミーゼス–フィッシャー分布族 (4) が得られる．

Remark 2.5. 関連研究として [CW15] があり，コンパクト群のユニタリ表現を用いて球面などのコンパクト等質空間上の指数型分布族を構成する手法および，実際に 2 次元球面上で最尤推定を現実的な計算量で行う方法が提案された．我々の立場からは，[CW15] による分布族構成手法は，G/H-method において G をコンパクト群に限った場合とみなせる．実際，G がコンパクト群であるとき，Fact 2.1 より $W_0(G,H) = \{0\}$ であり，コンパクト群の任意の連続表現はユニタリ化可能であるため，[CW15] で得られる分布族は G/H-method によっても構成できる．

2.5 上半平面上のポアンカレ分布族の構成

Example 2.6 (ポアンカレ分布族)**.** $G := SL(2,\mathbb{R})$, $H := SO(2)$ とし，$X := G/H$ を G の等質空間とする．上半平面 $\mathcal{H}$ は，G の 1 次分数変換による作用を許容する．この作用は推移的でありポアンカレ計量を保つことが知られている．このとき $i \in \mathcal{H}$ の固定部分群は H となり，上半平面は次の対応により X と同一視できる．

$$\begin{aligned} &\mathcal{H} \to X, \\ &z = x + iy \mapsto \begin{pmatrix} \sqrt{y} & \frac{x}{\sqrt{y}} \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix} H. \end{aligned}$$

G の有限次元実表現 $\rho\colon G \to GL(\mathrm{Sym}(2,\mathbb{R}))$ を次で定義する．

$$\rho(g)v := gv^t g \quad (g \in G, v \in \mathrm{Sym}(2,\mathbb{R})).$$

$v_0 := I_2 \in \mathrm{Sym}(2,\mathbb{R})$ は H 不変ベクトルであり，この ρ, v_0 に G/H-method を適用すると上半平面 $\mathcal{H}$ 上の以下のポアンカレ分布族が得られる．

$$\left\{ \frac{De^{2D}}{\pi} \exp\left(-\frac{a(x^2+y^2)+2bx+c}{y} \right) \frac{dxdy}{y^2} \right\}_{\begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathrm{Sym}^+(2,\mathbb{R})}. \tag{5}$$

ここで，$D = \sqrt{ac-b^2}$, $\mathrm{Sym}^+(2,\mathbb{R}) := \{S \in \mathrm{Sym}(2,\mathbb{R}) \mid S \text{ は正値 }\}$ である．

以下，実際に G/H-method を適用する．$SL(2,\mathbb{R})$ は単純であるから，Fact 2.1 (iii) より $W_0(G,H) = \{0\}$ となる．(2) と同様に $\mathrm{Sym}(2,\mathbb{R})$ に内積を入れることで $\mathrm{Sym}(2,\mathbb{R})^\vee$ を $\mathrm{Sym}(2,\mathbb{R})$ と同一視する．X 上の非ゼロ相対 G 不変測度として $\frac{dxdy}{y^2}$ を取る．このとき $\theta = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathrm{Sym}(2,\mathbb{R})$ に対して, X 上の測度 $\tilde{p}_\theta$ は次のように定まる．

$$\begin{aligned} d\tilde{p}_\theta(z) &= \exp(-\operatorname{trace}(\theta\rho(z)I_2)) \frac{dxdy}{y^2} \quad (z = x+iy \in \mathcal{H}) \\ &= \exp\left(-\operatorname{trace}\left(\begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \sqrt{y} & \frac{x}{\sqrt{y}} \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix} I_2 {}^t\!\begin{pmatrix} \sqrt{y} & \frac{x}{\sqrt{y}} \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix} \right) \right) \frac{dxdy}{y^2} \\ &= \exp\left(-\frac{a(x^2+y^2)+2bx+c}{y} \right) \frac{dxdy}{y^2} \end{aligned}$$

このとき [TY21b, Lemma 4] によりパラメータ空間と正規化定数は次のように求められる．

$$\begin{aligned} \Theta &= \left\{ \theta \in \mathrm{Sym}(2,\mathbb{R}) \;\middle|\; \int_{\mathcal{H}} d\tilde{p}_\theta < \infty \right\} \\ &= \mathrm{Sym}^+(2,\mathbb{R}), \\ c_\theta &= \int_{\mathcal{H}} d\tilde{p}_\theta = \frac{\pi}{De^{2D}} \quad (\theta \in \Theta). \end{aligned}$$

ここで $D := \sqrt{ac-b^2}$. 以上より $\tilde{p}_\theta$ を正規化することで，上半平面上の指数型分布族 (5) が得られた．

3 ポアンカレ分布族の共役事前分布族

共役事前分布族 (Definition 3.1) を持つ分布族に対してはベイズ推定を簡単に実行することができる（ベイズ推定については例えば [RS61] を参照されたい）．しかしながら一般には共役事前分布族は存在するとは限らない．一方で，指数型分布族に対してはいつでも共役事前分布族が存在し，それもまた指数型分布族となることが知られている (Fact 3.3)．この節ではポアンカレ分布族 (Example 2.6) の共役事前分布族を一つ正規化定数まで含めて明示的に提示する (Fact 3.4)．

3.1 指数型分布族の共役事前分布族

本節では共役事前分布族の定義を与え，[DY79] による指数型分布族に対する共役事前分布族構成法を復習する．

Definition 3.1 ([RS61], 共役事前分布族)**.** X を多様体とし，$\mathcal{P} = \{p_\theta(x)d\mu(x)\}_{\theta\in\Theta} \subset \mathcal{R}(X)$ を X 上の分布族とする．さらにパラメータ空間 Θ も多様体であると仮定し，$\mathcal{Q} = \{q_c(\theta)d\lambda(\theta)\}_{c\in C} \subset \mathcal{R}(\Theta)$ をハイパー

パラメータ空間を C とする Θ 上の分布族とする．このとき $\mathcal{Q}$ が P に**共役** (あるいは $\mathcal{P}$ の**共役事前分布族**) であるとは，任意のサンプルデータ $x_0 \in X$ と $p_\theta d\mu \in \mathcal{P}$ と $c \in C$ に対して，ある $c' \in C$ が一意的に存在して，$p_\theta(x_0)q_c(\theta)$ が $q_{c'}(\theta)$ と Θ 上の関数として正のスカラー倍を除いて一致することをいう．このとき $q_{c'}(\theta)d\lambda(\theta)$ を事後分布という．

Remark 3.2. 共役事前分布族の概念は well-defined である．実際，$\mathcal{P}$ における μ の取り方に依らない．一方で，与えられた $\mathcal{P}$ に対し，その共役事前分布族は一意ではない．実際，$\mathcal{Q}$ における λ の取り方に依存する．

指数型分布族に対しては次のように共役事前分布族を構成することができる．

Fact 3.3 ([DY79])**.** X を多様体，$\mathcal{P}$ を X 上の正則な指数型分布族 (“正則” の定義は例えば [BN78a, §8.1] を参照) とし，(μ, V, T) を $\mathcal{P}$ の最小次元実現とする．$(t, v) \in \mathbb{R}_{>0} \times V$ でパラメトライズされる Θ 上の関数 $f_{t,v}$ を次で定義する:

$$f_{t,v}(\theta) := \exp(-\langle \theta, v \rangle - t\varphi(\theta)).$$

ここで $\varphi(\theta) := \log c_\theta$ である．Θ 上の測度 $\lambda \in \mathcal{R}(\Theta)$ を一つ取り，部分集合 $C_\lambda \subset \mathbb{R}_{>0} \times V$ を次で定義する．

$$C_\lambda := \left\{ (t, v) \in \mathbb{R}_{>0} \times V \;\middle|\; \int_\Theta f_{t,v} d\lambda < \infty \right\}.$$

さらに C_λ でパラメトライズされる Θ 上の確率測度 $q_{t,v}$ を以下で定める．

$$\begin{aligned} dq_{t,v}(\theta) &:= c_{t,v}^{-1} f_{t,v}(\theta) d\lambda(\theta), \\ c_{t,v} &:= \int_\Theta f_{t,v} d\lambda \quad ((t, v) \in C_\lambda). \end{aligned}$$

このとき，$\mathcal{Q} := \{q_{t,v}\}_{(t,v) \in C_\lambda}$ は Θ 上の指数型分布族であり，$\mathcal{P}$ の共役事前分布族となる．与えられたデータ $x_0 \in X$ に対し，事後分布のハイパーパラメータは次のように更新される:

$$(t, v) \mapsto (t + 1, v + T(x_0)). \tag{6}$$

3.2 ポアンカレ分布族の共役事前分布族

本節では Fact 3.3 による共役事前分布族の構成法をポアンカレ分布族に対して適用することによって得られた分布族を明示的な形で紹介する．詳しい議論については [TY21b] を参照されたい．得られた分布族は $\mathrm{Sym}^+(2, \mathbb{R})$ 上の 4 次元パラメータを持つ指数型分布族となる．これはウィシャート分布族とよく似た形をしているが異なるものである．しかしながら，ともにより大きな 5 次元パラメータを持つ指数型分布族の部分族として実現できる (Proposition 3.7).

Fact 3.4 ([TY21b, Theorem 1])**.** 次の $\Theta := \mathrm{Sym}^+(2, \mathbb{R})$ 上の分布族はポアンカレ分布族の共役事前分布族である．

$$\left\{ c_{t,v}^{-1} \left(\frac{re^{2r}}{\pi} \right)^t \exp(-(\alpha a + 2\beta b + \gamma c)) \frac{da db dc}{2r^2} \right\}_{(t,v) \in C} \quad \left(\begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \Theta \right). \tag{7}$$

ここで $r := \sqrt{ac - b^2}$，$v = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \in \mathrm{Sym}^+(2, \mathbb{R}) \subset V$ である．ハイパーパラメータ空間 C と正規化定数

$c_{t,v}$ は次で与えられる．

$$C := \left\{ \left(t, \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \right) \in \mathbb{R}_{>0} \times \mathrm{Sym}^+(2, \mathbb{R}) \;\middle|\; t < D \right\},$$
$$c_{t,v} := \frac{\Gamma(t)}{2^t \pi^{t-1} D (D-t)^t}.$$

ここで $D := \sqrt{\alpha\gamma - \beta^2}$ である．

Remark 3.5 ([TY21b, Remark 7])**.** 与えられたデータ $x_0 + iy_0 \in \mathcal{H}$ に対し，ハイパーパラメータの更新は (6) より，以下のように与えられる．

$$C \to C, \quad \left(t, \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \right) \mapsto \left(t+1, \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} + \frac{1}{y_0} \begin{pmatrix} x_0^2 + y_0^2 & x_0 \\ x_0 & 1 \end{pmatrix} \right).$$

Remark 3.6. Fact 3.4 で与えた分布族は $\mathrm{Sym}^+(2, \mathbb{R})$ 上の 4 次元パラメータを持つ指数型分布族である．一方，$\mathrm{Sym}^+(2, \mathbb{R})$ 上の有名な分布族としてウィシャート分布族がある．これも 4 次元パラメータを持つが，Fact 3.4 で与えた分布族とは異なり，一方を他方の部分族とみなすことはできない．しかしながら，ともに以下の 5 次元パラメータを持つ $\mathrm{Sym}^+(2, \mathbb{R})$ 上の指数型分布族に含まれる．

Proposition 3.7. 次は $\mathrm{Sym}^+(2, \mathbb{R})$ 上の $SL(2, \mathbb{R}) \times \mathbb{R}_{>0}$ 不変指数型分布族であり，その部分族として (7) とウィシャート分布族を含む．

$$\left\{ \frac{D(t+2D)^\kappa}{\pi \Gamma(\kappa)} e^{-tr} r^\kappa \exp(-(\alpha a + 2\beta b + \gamma c)) \frac{da db dc}{2r^2} \right\}_{(\kappa, t, v) \in C} \quad \left(\begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathrm{Sym}^+(2, \mathbb{R}) \right) \tag{8}$$

ここで $v = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$, $D := \sqrt{\det v} = \sqrt{\alpha\gamma - \beta^2}$, $r := \sqrt{ac - b^2}$ であり，パラメータ空間 C は次で与えられる．

$$C = \{ (\kappa, t, v) \in \mathbb{R}_{>0} \times \mathbb{R} \times \mathrm{Sym}^+(2, \mathbb{R}) \mid t + 2D > 0 \}.$$

実際, $G := SL(2, \mathbb{R}) \times \mathbb{R}_{>0}$, $H := SO(2)$ として [TY21b, Lemma 1] で与えられる写像を通じて $\mathrm{Sym}^+(2, \mathbb{R})$ を $(SL(2, \mathbb{R}) \times \mathbb{R}_{>0}) / SO(2) = (SL(2, \mathbb{R}) / SO(2)) \times \mathbb{R}_{>0} \simeq \mathcal{H} \times \mathbb{R}_{>0}$ と同一視する．

次で定義される有限次元実表現 $\rho \colon G \to GL(\mathrm{Sym}(2, \mathbb{R}) \oplus \mathbb{R})$ を考える．

$$\rho((h, r))(S, a) := (r\, {}^t h^{-1} S\, {}^t h, ra) \quad ((h, r) \in G, (S, a) \in \mathrm{Sym}(2, \mathbb{R}) \oplus \mathbb{R})$$

このとき $v_0 := (I_2, 1)$ は H 不変ベクトルであり ρ, v_0 に対して G/H-method を適用することで $\mathrm{Sym}^+(2, \mathbb{R})$ 上の分布族 (8) が得られる．パラメータ空間や正規化定数の計算は Fact 3.4 と同様にできる．

パラメータ空間において $t = -2\kappa$ という制約を追加すると分布族 (7) が実現できる．一方，$t = 0$ という制約を追加するとウィシャート分布族が実現できる．

参考文献

[BN70] O. Barndorff-Nielsen, *Exponential family: exact theory*, Various Publication Series, No. 19, Matematisk Institut, Aarhus Universitet, Aarhus (1970).

[BN78a] O. E. Barndorff-Nielsen, *Information and exponential families in statistical theory* Chichester: Wiley (1978).

[BN78b] O. E. Barndorff-Nielsen, *Hyperbolic distributions and distribution on hyperbolae*, Scand. J. Statist. **8**, 151–157 (1978).

[BNBE89] O. E. Barndorff-Nielsen, P. Blæsild, P. S. Eriksen, *Decomposition and invariance of measures, and statistical transformation models*, New York: Springer Verlag (1989).

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

[DY79] P. Diaconis, F. Ylvisoker, *Conjugate priors for exponential families*, The Annals of Statistics, Vol.7, No.2, 269–281 (1979).

[RS61] H. Raiffa, R. Schlaifer, *Applied statistical decision theory*, Division of Research, Graduate School of Business Administration, Harvard University (1961).

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

[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).

[TY21a] K. Tojo, T. Yoshino, *Harmonic exponential families on homogeneous spaces*, Info. Geo. **4**, 215–243 (2021), `https://doi.org/10.1007/s41884-020-00033-3`.

[TY21b] K. Tojo, T. Yoshino, *An Exponential Family on the Upper Half Plane and Its Conjugate Prior*, In: Barbaresco F., Nielsen F. (eds) Geometric Structures of Statistical Physics, Information Geometry, and Learning. SPIGL 2020. Springer Proceedings in Mathematics & Statistics, vol 361. Springer, Cham. (2021), `https://doi.org/10.1007/978-3-030-77957-3_4`.