片側切断指数型分布族の情報幾何

大阪大学・基礎工学研究科　吉岡 正記
Masaki Yoshioka
Graduate School of Engineering Science, Osaka University

大阪大学・全学教育推進機構　田中 冬彦
Fuyuhiko Tanaka
Center for Education in Liberal Arts and Sciences, Osaka University

1　はじめに

統計理論と統計モデルの幾何学的性質には深い関係があることが情報幾何学によって明かされてきた [5]. 例えば, 漸近理論において高次の誤差項が曲率に依存することなどが良く知られている [4]. また, Bayes 統計学においても, 幾何学的性質に基づいた無情報事前分布の提案等が行われている [18]. これらの結果を支える幾何学的構造は, Fisher 情報行列から定義される Riemann 計量 (Fisher 計量) と, アファイン接続の 1 パラメータ族 (α 接続) である.

しかし, 上記の議論は正則条件を満たすような統計モデルを前提としており, そうでない統計モデルの取り扱いが課題の一つとなっている. よく知られる非正則モデルとして, 片側切断指数型分布族 (oTEF) と呼ばれる統計モデルが挙げられる [6, 1]. このモデルの分布は指数型分布族に近い密度関数を持つ一方, その密度のサポートはパラメータに依存しており, 正則条件は満たさない. 情報幾何学における oTEF の扱いは少なく, Pareto 分布族を対象として, 形式的な幾何構造を入れたものが多かった [14, 17, 16]. そのため, 幾何構造と統計理論との関係性は不明瞭であった. 最近, Yoshioka and Tanaka [19] は oTEF に幾何構造を入れ, $\alpha = 1$ での α 平行事前分布の存在を示した。しかし, α 接続の定義は形式的なままであった.

本稿では, oTEF に幾何学的構造を入れた下で曲率に関する 2 つの性質を導く．一部に Yoshioka and Tanaka [19] の概説を含む．幾何学的構造を入れる際には, 正則モデルの Fisher 計量・α 接続を拡張し, oTEF 上の Riemann 計量とアファイン接続を定める．Riemann 計量は oTEF の最尤推定量の漸近的振る舞いをもとに定議する．一方, アファイン接続は, 等積接続という観点から α 接続を拡張したもの (β 接続) を用いる．この幾何学的構造の下で, oTEF が $\beta = 1$ でアファイン座標系を持つこと, 及びサブモデルに制限した際に $\beta = 0$ でスカラー定曲率になることを示す.

2　準備: 正則モデルの情報幾何

非正則モデルを扱うための前準備として, この節では正則モデルの情報幾何で用いられる幾何学的構造を紹介する．詳細は Amari [4] や Amari and Nagaoka [5] 等を参照せよ.

その後, Bayes 統計学への応用として, α 平行事前分布 [18] を紹介する.

2.1　統計モデル多様体

n 個の実数 $\theta = (\theta^1, \dots, \theta^n)$ でパラメータ付けされた, 標本空間 χ 上の確率分布 P_θ の族

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\}$$

を考える．ここで, Θ は $\mathbb{R}^n$ の (連結な) 開集合である．このような確率分布族 $\mathcal{P}$ を, n 次元パラメトリック統計モデルと呼び, n 次元多様体として扱う．また, 分布 P_θ は (ルベーグ測度に関する) 確率密度関数 $p(x, \theta)$ を持つとし, その対数を $l(x, \theta) = \log p(x, \theta)$ と表記する.

情報幾何学では, 統計モデルに対して次の正則条件を仮定する.

定義 2.1 (正則条件 [4]).

パラメトリック統計モデル $\mathcal{P}$ に課される以下の条件を正則条件と呼ぶ.

1. 確率密度関数 (x, θ) は, x の関数として θ によらないサポートをもつ.
2. 以降で扱う関数に対して, 偏微分 $\frac{\partial}{\partial \theta^i}$ とルベーグ測度による積分は交換可能である.
3. 任意の θ において,

$$\frac{\partial}{\partial \theta^i} l(x, \theta) \quad (i = 1, 2, \dots, n)$$

は x の関数として線型独立である.

4. $\dfrac{\partial}{\partial\theta^i} l(x,\theta)$ の任意の次数のモーメントが存在する.

■

正則条件を満たす統計モデルを正則モデルと呼ぶ. 以降, この節では正則モデルのみ取り扱う.

次に, 統計モデルで重要な幾何量である Fisher 計量を紹介する. Fisher 計量とは, Fisher 情報行列をもとに定義される Riemann 計量である.

定義 2.2 (Fisher 計量 [4]).

Fisher 計量とは, 第 i, j 成分 $(i, j = 1, \dots, n)$ を

$$g^F_{ij}(\theta) = \mathbf{E}_\theta\left[\partial_i l(X,\theta)\partial_j l(X,\theta)\right] \quad (\theta \in \Theta,\ X \sim P_\theta)$$

とする正定値行列 $\left(g^F_{ij}\right)$ によって定まる Riemann 計量である. ■

接ベクトル $A = A^i\partial_i, B = B^i\partial_i \in T_\theta\mathcal{P}$ の Fisher 計量は

$$\langle A, B\rangle = \mathbf{E}_\theta\left[Al(X,\theta)Bl(X,\theta)\right]$$

となる.

また, 情報幾何学では, α 接続と呼ばれるアファイン接続の族を用いる.

定義 2.3 (α 接続 [4]).

統計モデル $\mathcal{P}$ の α 接続とは, 接続係数

$$\overset{(\alpha)}{\Gamma}{}_{ij,k}(\theta) = \mathbf{E}\left[\partial_i\partial_j l\,\partial_k l\right] + \frac{1-\alpha}{2}\mathbf{E}\left[\partial_i l\,\partial_j l\,\partial_k l\right]$$
$$(\alpha \in \mathbb{R},\ i, j, k = 1, \dots, n)$$

によって定まるアファイン接続の, パラメータ $\alpha \in \mathbb{R}$ に関する族である. ここで, $l = l(X,\theta)$ と省略した. ■

α 接続に対応する共変微分を $\overset{(\alpha)}{\nabla}$ と書く.

α 接続係数は, Fisher 計量に基づく Levi-Civita 接続係数 Γ^g と 3 次テンソル $T_{ijk} = \mathbf{E}\left[\partial_i l\partial_j l\partial_k l\right]$ を用いて

$$\overset{(\alpha)}{\Gamma}{}_{ij,k}(\theta) = \Gamma^g_{ij,k}(\theta) - \frac{\alpha}{2}T_{ijk}(\theta)$$

と表せる. この表現により, 0 接続, つまり $\alpha = 0$ の α 接続は Levi-Civita 接続に対応していると分かる. なお, Levi-Civita 接続とは Riemman 計量から定められるアファイン接

続であり, 接続係数は

$$\Gamma^{g}_{ij,k} = \partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}$$

と表現できる. [13].

2.2 体積要素と α 平行事前分布

情報幾何学の Bayes 統計学への貢献の一つに, α 平行事前分布という無情報事前分布がある.

パラメータ空間 $\Theta \subset \mathbb{R}^n$ 上の事前分布 π は, 情報幾何学において, 統計モデル多様体上の体積要素と同一視される. 例えば, Jeffreys 事前分布 [12] $\pi_J \propto \sqrt{\det(g^F)}$ に対応する体積要素 ω_J は, 局所座標系を用いて

$$\omega_J = \sqrt{\det(g^F)} d\theta^1 \wedge \cdots \wedge d\theta^n$$

と表現される. そして, 体積要素 ω_J は 0 平行という幾何的特徴をもっている.

Takeuchi and Amari [18] は 0 平行な体積要素を α 平行な体積要素へと拡張し, α 平行事前分布を提案した. Jeffreys 事前分布と同様に, α 平行事前分布もパラメータ変換に対する不変性をもつ.

この節では, 正則モデルにおける α 平行事前分布とその存在性に関わる定理を簡単に紹介する.

まず, α 平行事前分布を定義するために必要な微分幾何学の概念を述べる.

定義 2.4 (等積接続 [21]).
$\mathcal{P}$ を n 次元多様体とし, ∇ を $\mathcal{P}$ 上の捩れの無いアファイン接続とする.
$\mathcal{P}$ のいたるところで

$$\nabla \omega = 0$$

となる体積要素 ω が存在するとき, ∇ を (局所) 等積接続 (equiaffine connection) という (または等積的という). また, そのような ω を ∇ について平行な体積要素という. ■

Riemann 多様体の場合, Levi-Civita 接続が等積接続の一種となる.

アファイン接続 ∇ が 等積接続 となるための必要十分条件は, 曲率によって記述される. ここで, ∇ による Riemann 曲率テンソルを

$$R_{ijk}{}^{l} = \partial_i \Gamma_{jk}{}^{l} - \partial_j \Gamma_{ik}{}^{l} + \Gamma_{im}{}^{l} \Gamma_{jk}{}^{m} - \Gamma_{jm}{}^{l} \Gamma_{ik}{}^{m} \quad (i, j, k, l = 1, \dots, n)$$

と定める．また, Ricci 曲率テンソルを Ricci_{ij} と表記する．

アファイン接続を持つ一般の多様体に対して, 次の命題が成立する.

命題 2.5 (野水 and 佐々木 [21]).

$\mathcal{P}$ を多様体とし, ∇ を $\mathcal{P}$ 上の捩れの無いアファイン接続とする.

このとき, 次の三つは同値である.

1. ∇ は (局所) 等積接続 である
2. $\mathcal{P}$ 上で $R_{ijk}{}^{k} = 0 \quad (i, j = 1, \dots, n)$ が成立する.
3. $\mathcal{P}$ 上で $\mathrm{Ricci}_{ij} = \mathrm{Ricci}_{ji} \quad (i, j = 1, \dots, n)$ が成立する.

■

さて, 統計モデル $\mathcal{P}$ の場合に話を戻し, α 平行事前分布を定義する.

定義 2.6 (α 平行事前分布 [18]).

統計モデル $\mathcal{P}$ 上に, α 接続に対して平行な体積要素 $\omega^{(\alpha)}$ が存在すると仮定する.

このとき, $\omega^{(\alpha)}$ の θ 座標系による表現

$$\omega^{(\alpha)} = \pi^{(\alpha)} d\theta^1 \wedge \cdots \wedge d\theta^n$$

に表れる関数 $\pi \in C^{\infty}(\mathcal{P})$ を α 平行事前分布という. ■

ここで, $\omega^{(\alpha)}$ は α 平行な体積要素とも呼ばれる.

正則な統計モデルに命題 2.5 を適用すると, 次のような結果が得られる.

命題 2.7 (Takeuchi and Amari [18]).

統計的モデル $\mathcal{P}$ について,

$$\partial_i T_{jk}{}^{k} - \partial_j T_{ik}{}^{k} = 0 \quad (i, j = 1, \dots, n)$$

ならば, 任意の $\alpha \in \mathbb{R}$ に対して α 接続は等積的である. そうでなければ, 0 接続のみ等積的である. ■

定義 2.8 (統計的等積 [18]).

統計モデル $\mathcal{P}$ について, $\mathcal{P}$ 上の各点で

$$\partial_i T_{jk}{}^{k} - \partial_j T_{ik}{}^{k} = 0 \quad (i, j = 1, \dots, n)$$

が成立するとき, $\mathcal{P}$ は統計的等積 (statisticaly equiaffine) であるという. ■

すなわち, $\mathcal{P}$ が統計的等積とは, 任意の α について α 平行事前分布が存在することである. 統計的等積なモデルの例として, 指数型分布族が挙げられる [18].

3　片側切断指数型分布族とその幾何学的構造

この節では, 片側切断指数型分布族 (oTEF) について簡単に解説した後, oTEF 上に Riemann 計量とアファイン接続を定義する.

3.1　片側切断指数型分布族

oTEF とは, 次のような確率分布族である.

定義 3.1 (片側切断指数型分布族 (One-sided Truncated Exponential Family) [6]).

パラメータ θ, γ をもつ確率分布族 $\mathcal{P} = \left\{P_{\theta,\gamma} : \theta \in \Theta, \gamma \in I\right\}$ を考える. なお, パラメータ空間 Θ を $\mathbb{R}^n$ の (連結) 開集合とし, $I = (I_1, I_2)$ を開区間とする. そして, それぞれの確率分布 $P_{\theta,\gamma}$ が確率密度関数

$$p(x, \theta, \gamma) = \exp\left\{\sum_{i=1}^{n} \theta^i F_i(x) + C(x) - \psi(\theta, \gamma)\right\} \cdot \mathbb{1}_{[\gamma, I_2)}(x) \quad (x \in I) \quad (3.1)$$

を持つとき, $\mathcal{P}$ を片側切断指数型分布族とよぶ. ここで, $C \in C^\infty(I), F_i \in C^\infty(I)\,(i = 1, \dots, n), \psi \in C^\infty(\Theta \times I)$ とした. ■

θ を自然パラメータ, γ を切断パラメータと呼ぶ. 以降, $\partial_i = \partial/\partial_i$, $\partial_\gamma = \partial/\partial_\gamma$ と略記する. また, 添え字 $i, j, k, \dots$ の範囲が $1, \dots, n$, 添え字 $a, b, c, \dots$ の範囲が $1, \dots, n, \gamma$ となるように書き分ける.

oTEF は正則条件を満たさないような統計モデルである. 実際, 確率密度関数 $p(x, \theta, \gamma)$ の x に関するサポートは区間 $[\gamma, I_2]$ であり, パラメータ γ に依存している.

更に, パラメータ γ による偏微分と x に関する積分の順序交換ができないという点でも正則条件を満たしていない. 例えば, 対数尤度の γ 偏微分の期待値は 0 ではなく,

$$\mathbf{E}\left[\partial_\gamma l(X, \theta, \gamma)\right] = -\partial_\gamma \psi(\theta, \gamma)$$

となる. ただし, $x = \gamma$ で $\partial_\gamma l$ は存在しないため, 積分範囲を (γ, I_2) として計算した. 同様に,

$$\mathbf{E}\left[\partial_i \partial_\gamma l(X, \theta, \gamma)\right] = -\partial_\gamma \partial_i \psi(\theta, \gamma) \quad (i = 1, \dots, n)$$
$$\mathbf{E}\left[\partial_\gamma \partial_\gamma l(X, \theta, \gamma)\right] = -\partial_\gamma \partial_\gamma \psi(\theta, \gamma)$$

となるため, Fisher 情報量に関する等式 [7]

$$-\mathbf{E}\left[\partial_a\partial_\gamma l(X,\theta,\gamma)\right]=\mathbf{E}\left[\partial_a l(X,\theta,\gamma)\partial_\gamma l(X,\theta,\gamma)\right]\quad (a=1,\ldots,n,\gamma)$$

も一般に成り立たない．これらの計算は, 対数尤度の γ 偏微分

$$\partial_\gamma l(x,\theta,\gamma)=-\partial_\gamma\psi(\theta,\gamma)\quad (x>\gamma)$$

が x に関して定数となることによって簡単に分かる.

oTEF は非正則モデルだが, パラメータ γ を固定した場合には指数型分布族となり, 正則モデルとなる．このことは確率密度関数の形 (3.1) から確かめられる．つまり, 正則条件に抵触しているのはパラメータ γ に関する部分のみである.

3.2 Riemann 計量

この節では, 最尤推定量の漸近的性質を用いて oTEF の Riemann 計量を定義する.

定義 3.2 (oTEF の Riemann 計量 [19]).
$\mathcal{P}=\left\{P_{\theta,\gamma}:\theta\in\theta,\gamma\in I\right\}$ を oTEF とする．oTEF の Rimann 計量を,

$$\begin{aligned}g_{ij}&=\mathbf{E}\left[\partial_i l(x,\theta,\gamma)\partial_j l(x,\theta,\gamma)\right]\\ g_{i\gamma}&=0\\ g_{\gamma\gamma}&=\left\{\partial_\gamma\psi(\theta,\gamma)\right\}^2\end{aligned}$$

から定まる Riemann 計量とする．ここで, $i=1,\ldots,n$ とした. ■

なお, g_{ij} は, 指数型分布族と同じように

$$\begin{aligned}g_{ij}&=-\mathbf{E}\left[\partial_i\partial_j l(x,\theta,\gamma)\right]\\ &=\partial_i\partial_j\psi(\theta,\gamma)\end{aligned}$$

と関数 ψ の二次導関数のみで表せる.

次に, 定義 3.2 に至った手順について解説する.

■Riemann 計量の定義の仕方　まず, oTEF $\mathcal{P}$ を二つの多様体

$$\begin{aligned}\mathcal{E}_\gamma&=\left\{P_{\theta,\gamma}:\theta\in\theta\right\}\\ \mathcal{F}_\theta&=\left\{P_{\theta,\gamma}:\gamma\in I\right\}\end{aligned}$$

に分けて考える．$\mathcal{E}_\gamma$ は oTEF の切断パラメータ γ を固定して得られる n 次元指数型分布族であり, 正則モデルとなる．一方, $\mathcal{F}_\theta$ は自然パラメータ γ を固定して得られる 1 次元

非正則モデルである．このとき, 接空間は

$$\begin{aligned} T_{\theta,\gamma}\mathcal{P} &= \mathrm{Span}\{\partial_1, \dots, \partial_n\} + \mathrm{Span}\{\partial_\gamma\} \\ &= T_{\theta,\gamma}\mathcal{E}_\gamma + T_{\theta,\gamma}\mathcal{F}_\theta \end{aligned}$$

と分解される．すると, 正則モデルの接空間 $T_{\theta,\gamma}\mathcal{E}_\gamma$ には, Riemann 計量として Fisher 計量が自然に入る．よって,

$$g_{ij} = \mathbf{E}\left[\partial_i l(X,\theta,\gamma)\partial_j l(X,\theta,\gamma)\right] \quad (i,j=1,\dots,n)$$

と定めればよい．残るは, 接ベクトル ∂_i と ∂_γ の内積 $g_{i\gamma} \quad (i=1,\dots,n)$ 及び $\mathcal{F}_\theta$ の Riemann 計量 $g_{\gamma\gamma}$ である．

ここで, Fisher 計量の正則モデルにおける統計的意味を振り返る．$\mathcal{P}_0$ を, パラメータ $\theta \in \mathbb{R}^n$ を持つような正則モデルとする．Fisher 計量とは, Fisher 情報行列から定義される Riemann 計量である．そして, 最尤推定量 $\hat{\theta}$ の分散を展開すると

$$\mathbf{V}\left[\hat{\theta}\right] = \frac{1}{N}\left(g_{ij}^F\right)^{-1}$$

となり, 1 次の係数と $\left(g_{ij}^F\right)^{-1}$ が対応する [4].

oTEF では逆の手順を取り, 最尤推定量の分散の係数から Riemann 計量を定める．oTEF の最尤推定量 $\hat{\theta}, \hat{\gamma}$ の分散は,

$$\begin{aligned} \mathbf{V}\left[\hat{\theta}\right] &= \frac{1}{N}\left(g_{ij}\right)^{-1} + \mathrm{O}\left(\frac{1}{N^2}\right), \\ \mathbf{V}\left[\hat{\gamma}\right] &= \frac{1}{N^2\left\{\partial_\gamma\psi(\theta,\gamma)\right\}^2} + \mathrm{O}\left(\frac{1}{N^3}\right), \\ \mathbf{Cov}\left[\hat{\theta},\hat{\gamma}\right] &= \mathrm{O}\left(\frac{1}{N^3}\right) \end{aligned}$$

と表せる [3]．ここで, N はサンプル数を表す．すると, Fisher 情報行列と同じように, $\left\{\partial_\gamma\psi(\theta,\gamma)\right\}^2$ は第一項の係数の逆数として現れる．このことから, $\mathcal{F}_\theta$ の Riemann 計量を

$$g_{\gamma\gamma} = \left\{\partial_\gamma\psi(\theta,\gamma)\right\}^2$$

と定義する．

また, $\mathbf{V}\left[\hat{\theta}\right]$ や $\mathbf{V}\left[\hat{\gamma}\right]$ と比較して $\mathbf{Cov}\left[\hat{\theta},\hat{\gamma}\right]$ のオーダーが小さくなっていることに注目する．更に, $\hat{\gamma}$ に関して, θ 既知か否かは分散の一次の項に影響しない [3]．これは θ の推定でも同様である．これらの事実を鑑みて, 接ベクトル ∂_i と ∂_γ は直交しているとし,

$$g_{i\gamma} = 0 \quad (i=1,\dots,n)$$

と定める.

以上により, 定義 3.2 の Riemann 計量が得られた.

さて, 我々の Riemann 計量と, 先行研究 (Rylov [16],Li et al. [14]) で用いられていた Riemann 計量を比較する. 先行研究では, Fisher 計量をもとに形式的に定義された Riemann 計量を用いていた. 実は, その Riemann 計量は定義 3.2 と同じものである. つまり

$$g_{ab} = \mathbf{E}\left[\partial_a l(x,\theta,\gamma)\partial_b l(x,\theta,\gamma)\right] \quad (a,b=1,\dots,n,\gamma) \tag{3.2}$$

が成立する.

3.3 アファイン接続

この節では, oTEF 上のアファイン接続について検討する. 目的は, 正則モデルで用いられてきた α 接続を oTEF へと拡張することである.

しかし, oTEF への α 接続の拡張は, 複数通り考えられる. 例えば, Rylov [16] と Sun et al. [17] は Pareto 分布族に α 接続を与えて議論しているが, それら二つの α 接続は異なるものになっている.

ここでは, 等積接続という観点から oTEF のアファイン接続を考える. 以前より, 指数型分布族は統計的等積という事実が知られている [18]. つまり, 指数型分布族において, 任意の $\alpha \in \mathbb{R}$ に対し, α 接続は等積接続である. そこで, この性質を保存するような片側切断指数型分布族のアファイン接続を考える.

まず, Sun et al. [17] が Pareto 分布族を扱った時と同様の方法で, oTEF の α 接続を形式的に定める.

定義 3.3 (oTEF の α 接続 [19]).

$\alpha \in \mathbb{R}$ に対し, oTEF 上の α 接続を, 接続係数

$$\overset{(\alpha)}{\Gamma}_{ab,c}(\theta,\gamma) = \mathbf{E}\left[\partial_a\partial_b l\,\partial_c l\right] + \frac{1-\alpha}{2}\mathbf{E}\left[\partial_a l\,\partial_b l\,\partial_c l\right] \quad (a,b,c=1,\dots,n,\gamma)$$

によって定義する. ここで, $l = l(X,\theta,\gamma)$ と省略した. ■

上記の定義は, 正則モデルにおける α 接続の式に oTEF の確率密度関数を代入することによって得られる. ただし, 正則条件を満たさないため, Levi-Civita 接続を含んでいない.

α 接続係数は, 添え字 $i, j, k = 1, \dots, n$ に対して,

$$
\begin{aligned}
\overset{(\alpha)}{\Gamma}_{ij,k} &= \frac{1-\alpha}{2}\partial_i\partial_j\partial_k\psi, \\
\overset{(\alpha)}{\Gamma}_{ij,\gamma} &= \frac{1+\alpha}{2}\partial_i\partial_j\psi\partial_\gamma\psi, \\
\overset{(\alpha)}{\Gamma}_{i\gamma,j} &= -\frac{1-\alpha}{2}\partial_i\partial_j\psi\partial_\gamma\psi, \\
\overset{(\alpha)}{\Gamma}_{i\gamma,\gamma} &= \partial_i\partial_\gamma\psi\partial_\gamma\psi, \\
\overset{(\alpha)}{\Gamma}_{\gamma\gamma,i} &= 0, \\
\overset{(\alpha)}{\Gamma}_{\gamma\gamma,\gamma} &= \partial_\gamma\partial_\gamma\psi\partial_\gamma\psi - \frac{1-\alpha}{2}\left(\partial_\gamma\psi\right)^3,
\end{aligned}
$$

のように関数 $\psi(\theta, \gamma)$ を用いて表せる.

上記の α 接続について, 等積接続になるかどうかを調べたところ, 次のような結果が得られた.

定理 3.4 (Yoshioka and Tanaka [19]).

$\mathcal{P}$ を oTEF とし, 定義 3.2 の Riemann 計量と定義 3.3 の α 接続を微分幾何学的構造として入れる.

この幾何構造のもと, $\alpha = 1$ のとき, α 平行な体積要素が存在する. そして, 1 平行な体積要素の密度を $\pi^{(1)}$ とすると,

$$
\pi^{(1)}(\theta, \gamma) \propto -\partial_\gamma\psi(\theta, \gamma)
$$

と表せる. ■

定理 3.4 より, oTEF の等積接続として, 1 接続が得られた. 他に, 定義 3.2 の Riemann 計量に関する Levi-Civita 接続も等積的である. これらを用いて, 等積接続の 1 パラメータ族を構成する.

まず, 一般のアファイン接続に対して次の補題が成立する.

補題 3.5.

$\mathcal{P}$ を n 次元多様体, ∇^0, ∇^1 を $\mathcal{P}$ 上の等積接続とする. $\beta \in \mathbb{R}$ に対し, $\mathcal{P}$ 上のアファイン接続 $\overset{(\beta)}{\nabla}$ を

$$
\overset{(\beta)}{\nabla} = (1-\beta)\nabla^1 + \beta\nabla^0
$$

と定義する.

このとき, 任意の $\beta \in \mathbb{R}$ に対し, $\overset{(\beta)}{\nabla}$ は等積接続である. ■

証明. $\mathcal{P}$ の座標系 $\theta = (\theta^1, \dots, \theta^n) \in \Theta$ をとる．θ 座標系の下で，アファイン接続 ∇^0, ∇^1 の接続係数をそれぞれ $\left\{ \overset{0}{\Gamma}{}_{ij}{}^{k} \right\}, \left\{ \overset{1}{\Gamma}{}_{ij}{}^{k} \right\}$ とかく．また，$\overset{(\beta)}{\nabla}$ に対する曲率テンソルを $\overset{(\beta)}{R}$ と表記する．

このとき，$\overset{(\beta)}{\nabla}$ の接続係数は

$$\overset{(\beta)}{\Gamma}{}_{ij}{}^{k} = (1-\beta)\overset{0}{\Gamma}{}_{ij}{}^{k} + \beta\overset{1}{\Gamma}{}_{ij}{}^{k} \quad (i, j, k = 1, \dots, n)$$

となる．

この接続において，命題 2.5 の条件

$$\overset{(\beta)}{R}{}_{ijk}{}^{k} = 0 \quad (i, j = 1, \dots, n)$$

の成立を確かめればよい．

左辺を計算すると，

$$\begin{aligned}
\overset{(\beta)}{R}{}_{ijk}{}^{k} &= \partial_i \overset{(\beta)}{\Gamma}{}_{jk}{}^{k} - \partial_j \overset{(\beta)}{\Gamma}{}_{ik}{}^{k} + \overset{(\beta)}{\Gamma}{}_{jk}{}^{l} \overset{(\beta)}{\Gamma}{}_{il}{}^{k} - \overset{(\beta)}{\Gamma}{}_{ik}{}^{l} \overset{(\beta)}{\Gamma}{}_{jl}{}^{k} \\
&= \partial_i \overset{(\beta)}{\Gamma}{}_{jk}{}^{k} - \partial_j \overset{(\beta)}{\Gamma}{}_{ik}{}^{k} \\
&= (1-\beta)\left(\partial_i \overset{0}{\Gamma}{}_{jk}{}^{k} - \partial_j \overset{0}{\Gamma}{}_{ik}{}^{k} \right) + \beta \left(\partial_i \overset{1}{\Gamma}{}_{jk}{}^{k} - \partial_j \overset{1}{\Gamma}{}_{ik}{}^{k} \right)
\end{aligned}$$

となる．

ここで，∇^0, ∇^1 が等積接続であることから，

$$\begin{aligned}
\partial_i \overset{0}{\Gamma}{}_{jk}{}^{k} - \partial_j \overset{0}{\Gamma}{}_{ik}{}^{k} = 0 \\
\partial_i \overset{1}{\Gamma}{}_{jk}{}^{k} - \partial_j \overset{1}{\Gamma}{}_{ik}{}^{k} = 0
\end{aligned}$$

が任意の $i, j = 1, \dots, n$ で成立する．

よって，アファイン接続 $\overset{(\beta)}{\nabla}$ において

$$\overset{(\beta)}{R}{}_{ijk}{}^{k} = 0 \quad (i, j = 1, \dots, n)$$

が成立する．したがって，命題 2.5 より，$\overset{(\beta)}{\nabla}$ は等積接続である． ■

上記の補題に基づき，1 接続と Levi-Civita 接続をパラメータで結んだアファイン接続の族を定義する．

定義 3.6 (β 接続).

$\beta \in \mathbb{R}$ に対し, 接続係数

$$\overset{(\beta)}{\Gamma}_{ab,c}(\theta,\gamma) = \beta \mathbf{E}\left[\partial_a \partial_b l \, \partial_c l\right] + (1-\beta)\Gamma^g_{ab,c} \quad (a,b,c = 1,\dots,n,\gamma)$$

によって定義されるアファイン接続を β 接続と呼ぶ. ただし, Γ^g は Levi-Civita 接続の接続係数を表す. ■

正則モデルにおいて, β 接続は $\beta = \alpha$ の α 接続に一致する. よって, β 接続とは正則モデルの α を定義 3.3 とは異なる形で拡張したものである.

β 接続係数は, 関数 ψ を用いて

$$\begin{aligned}
\overset{(\beta)}{\Gamma}_{ij,k} &= \frac{1-\beta}{2}\partial_i\partial_j\partial_k\psi(\theta,\gamma),\\
\overset{(\beta)}{\Gamma}_{ij,\gamma} &= \beta\partial_i\partial_j\psi\partial_\gamma\psi - \frac{1-\beta}{2}\partial_i\partial_j\partial_\gamma\psi,\\
\overset{(\beta)}{\Gamma}_{i\gamma,j} &= \frac{1-\beta}{2}\partial_i\partial_j\partial_\gamma\psi,\\
\overset{(\beta)}{\Gamma}_{i\gamma,\gamma} &= \partial_i\partial_\gamma\psi\partial_\gamma\psi,\\
\overset{(\beta)}{\Gamma}_{\gamma\gamma,i} &= -(1-\beta)\,\partial_{i\gamma}\psi\partial_\gamma\psi,\\
\overset{(\beta)}{\Gamma}_{\gamma\gamma,\gamma} &= \beta\partial_\gamma\partial_\gamma\psi\partial_\gamma\psi + (1-\beta)\,\partial_\gamma\partial_\gamma\psi\partial_\gamma\psi
\end{aligned}$$

と表せる.

上記の β 接続に関して β 平行事前分布の存在を調べると, 次のような結果が得られる. なお, β 平行事前分布とは β 接続に関して平行な体積要素に対応する事前分布のことである.

定理 3.7 (β 平行事前分布).

$\mathcal{P}$ を oTEF とし, 定義 3.2 の Riemann 計量と定義 3.6 の β 接続を幾何構造として入れる.

このとき, 任意の $\beta \in \mathbb{R}$ において β 接続は等積的であり, β 平行事前分布が存在する. そして, $\pi^{(\beta)}$ を β 平行事前分布とすると,

$$\pi^{(\beta)} \propto \left\{\det\left(g_{ij}\right)\right\}^{\frac{1-\beta}{2}}\left(-\partial_\gamma\psi\right)$$

と表せる. ■

証明. 任意に $\beta \in \mathbb{R}$ を一つ固定する.

表 1　α 接続の比較

	α 接続 (指数型分布族)	α 接続 (oTEF)	β 接続 (oTEF)
等積接続	$\alpha \in \mathbb{R}$	$\alpha = 1$	$\beta \in \mathbb{R}$
Levi-Civita 接続	$\alpha = 0$	含まない	$\beta = 0$

補題 3.5 と β 接続の定め方より, β 接続は等積接続である.

β 平行事前分布の表示を求める.

β 接続係数の和を取ると,

$$\overset{(\beta)}{\Gamma}{}_{ia}{}^{a} = \frac{1-\beta}{2}\partial_i \log\left(\det(g_{jk})\right) + \partial_i \log\left(-\partial_\gamma \psi\right),$$
$$\overset{(\beta)}{\Gamma}{}_{\gamma a}{}^{a} = \frac{1-\beta}{2}\partial_\gamma \log\left(\det(g_{jk})\right) + \partial_\gamma \log\left(-\partial_\gamma \psi\right).$$

となることが分かる.

よって, α 平行事前分布と同様に, Takeuchi and Amari [18] の命題 1 より,

$$\pi^{(\beta)} \propto \left\{\det\left(g_{ij}\right)\right\}^{\frac{1-\beta}{2}}\left(-\partial_\gamma \psi\right)$$

が得られる. ■

β 平行事前分布は, Jeffreys 事前分布と定理 3.4 の 1 平行事前分布を含むような事前分布の族である. $\beta = 0$ のときに Jeffreys 事前分布, $\beta = 1$ のときに定理 3.4 の 1 平行事前分布を表す.

定理 3.7 により, β 接続は指数型分布族における α 平行事前分布の存在性を保存していることが分かる. よって, 等積的という観点からみた場合, 定義 3.3 の α 接続よりも定義 3.6 の β 接続の方がふさわしいといえる. これらの接続の比較を表 1 に表した.

また, 上記の 1 平行事前分布は, ある Reference prior [9] と一致する. Reference prior とは, Bernardo [8] によって提唱された, 情報理論の観点から導き出される無情報事前分布である. 具体的には, 事後分布と事前分布の KL ダイバージェンスを最大にするような事前分布として定義され, nuisance パラメータがない場合には Jeffreys 事前分布に一致する [8]. しかし, nuisance パラメータが存在する場合にはそうとは限らない. Ghosh and Mukerjee [10] らは, 適切な罰則項の下で汎関数の最大化を考えることにより, nuisance パラメータが存在する場合の Reference prior に新たな定式化を与えた. 更に, Ghosal [9] はその Reference prior を密度のサポートがパラメータに依存するような非正則モデルに拡張した. γ を興味のあるパラメータ, θ を nuisance パラメータとして oTEF に適用した

とき, Ghosal [9] の Reference prior は

$$\pi_{\mathrm{Ghosal}}(\theta, \gamma) \propto -\partial_\gamma \psi$$

となる. よって, 定理 3.4 の 1 平行事前分布に一致する.

4　片側切断指数型分布族の曲率

3 節で oTEF に幾何構造が導入されたことで, 幾何的性質の議論が可能となった. この節では, oTEF の曲率に関する性質を二つ紹介する. 一つは, oTEF が $\beta = 1$ でアファイン座標系をもつというものであり, もう一つは oTEF のあるサブモデルが Levi-Civita 接続 ($\beta = 0$) に対してスカラー定曲率になるというものである.

4.1　片側切断指数型分布族の 1 アファイン座標系

アファイン接続をもつ多様体には,

$$\Gamma_{ij}{}^{k} = 0 \quad (i, j, k = 1, \dots, n)$$

となる特別な座標系が存在することがあり, アファイン座標系と呼ばれている [20]. この座標系はユークリッド空間におけるカーテシアン座標系に対応する. 統計モデルの場合, 指数型分布族の自然パラメータは 1 接続に対するアファイン座標系となる [4].

同様に, oTEF にもアファイン座標系が存在する.

定理 4.1.

$\mathcal{P}$ を oTEF とし, β 接続を用いる. また, $\xi = (\theta^1, \dots, \theta^n, \psi(\theta, \gamma)) \in \Theta \times \psi(\Theta \times I)$ を座標系として入れる.

このとき, $\beta = 1$ で ξ はアファイン座標系となる. つまり,

$$\overset{(1)}{\Gamma}{}_{st}{}^{u}(\xi) = 0 \quad (s, t, u = 1, \dots, n, n+1)$$

が成立する. ■

証明. 二つの座標系を

$$y = (\theta^1, \dots, \theta^n, \gamma)$$
$$\xi = (\theta^1, \dots, \theta^n, \psi)$$

と表記する.

藤原 [20] より,

$$\overset{(1)}{\Gamma}{}_{st}{}^{u}(\xi) = 0 \quad (s, t, u = 1, \dots, n, n+1)$$

となるための必要十分条件は

$$\left(\frac{\partial \xi^u}{\partial y^c}\right) \overset{(1)}{\Gamma}{}_{ab}{}^{c}(y) = \frac{\partial^2 \xi^u}{\partial y^a \partial y^b} \quad (a, b, u = 1, \dots, n, n+1) \tag{4.1}$$

で与えられる. この条件の成立を確認すればよい.

まず, 式 (4.1) の右辺を計算する. y から ξ への座標変換における Jacobi 行列は,

$$\left(\frac{\partial \xi^s}{\partial y^a}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \partial_1 \psi & \cdots & \partial_n \psi & \partial_\gamma \psi \end{pmatrix}$$

であるから, $a, b = 1, \dots, n, \gamma$ に対して

$$\frac{\partial^2 \xi^u}{\partial y^a \partial y^b} = \begin{cases} 0 & (u = 1, \dots, n) \\ \partial_a \partial_b \psi & (u = n+1) \end{cases}$$

となる.

次に, 式 (4.1) の左辺を計算する. 添え字 $a, b = 1, \dots, n, \gamma$, $i = 1, \dots, n$ に対して, 1 接続係数の反変表現は,

$$\begin{aligned} \overset{(1)}{\Gamma}{}_{ab}{}^{i}(y) &= 0 \\ \overset{(1)}{\Gamma}{}_{ab}{}^{\gamma}(y) &= \frac{\partial_a \partial_b \psi}{\partial_\gamma \psi} \end{aligned}$$

となる. よって,

$$\begin{aligned} \left(\frac{\partial \xi^u}{\partial y^c}\right) \overset{(1)}{\Gamma}{}_{ab}{}^{c}(y) &= \left(\frac{\partial \xi^u}{\partial \gamma}\right) \overset{(1)}{\Gamma}{}_{ab}{}^{\gamma}(y) \\ &= \begin{cases} 0 & (u = 1, \dots, n) \\ \partial_a \partial_b \psi & (u = n+1) \end{cases} \end{aligned}$$

である.

したがって, 式 (4.1) が成立するため, $\xi = (\theta^1, \dots, \theta^n, \psi)$ はアファイン座標系である. ■

多様体に関して, アファイン座標系を持つことと平坦であることは同値であるため [20], 次の事実が導かれる.

命題 4.2.

$\mathcal{P}$ を oTEF とし, β 接続をアファイン接続として入れる.

このとき, $\mathcal{P}$ は $\beta = 1$ で平坦である. ■

4.2 定曲率なサブモデル

この節では, oTEF のあるサブモデルにおいて, Levi-Civita 接続に関するスカラー曲率が定数になることを示す. なお, Levi-Civita 接続 は $\beta = 0$ の β 接続に対応する.

上記の結果は, Pareto 分布族が定曲率であることの拡張といえる. 先行研究では, oTEF の一種である Pareto 分布族について, その Riemann 幾何学的性質が調べられてきた. 具体的には, Pareto 分布族は Levi-Civita 接続に関して定曲率空間であること [16] や Pareto 分布族と Poincarè 上半平面は等長であること [14] などが知られている. なお, 式 (3.2) より, これらの研究で採用されている Riemann 計量は定義 3.2 と同じものである. 実は, Pareto 分布族の高次元版を考えることにより, 類似した幾何的性質を導くことができる.

まず, 先行研究を高次元に拡張するために, Pareto 分布族の n 次元版にあたる統計モデルを考える.

定義 4.3 (γ 共通サブモデル).

γ 共通サブモデルとは, 確率密度関数

$$q(x,\theta,\gamma) = \exp\left\{ -\sum_{i=1}^{n} \theta^i \left\{F_i(x_i) - F_i(\gamma)\right\} + \sum_{i=1}^{n} \log \theta^i \right\} \cdot \mathbb{1}_{I^n(\gamma)}(x)$$

$$\left(x = (x_1, \dots, x_n) \in x \in \prod I^n \right)$$

を持つような $\mathbb{R}^n$ 上の確率分布 $Q_{\theta,\gamma}$ の集まり $\mathcal{Q} = \left\{ Q_{\theta,\gamma} : \theta \in \Theta, \gamma \in I \right\}$ である. ここで, $F_i \in C^1(\mathbb{R})$ を単調増加関数, $I = F_i(\mathbb{R}), I(\gamma) = I \cap [\gamma, \infty)$ とし, パラメータ空間を $\Theta = \mathbb{R}^n_{>0}, \gamma \in I$ とした. ■

γ 共通サブモデルは, 切断指数分布を関数 F_i で変換して同時分布をとったのちに, γ を共通とすることによって得られる. 確率変数 $X_i\,(i = 1, \dots, n)$ は確率密度関数

$$p_E(x,\theta,\gamma) = \exp\left\{-\theta(x-\gamma)\right\} \cdot \mathbb{1}_{[\gamma,\infty)}(x)$$

を持つような分布に従うものとする．この分布は切断指数分布や 2 パラメータ指数分布と呼ばれ, oTEF に含まれている．X_i を変換した $F_i(X_i)$ でパラメータ γ を取り直すと, その確率密度関数は

$$p_i(x,\theta,\gamma) = F_i'(x)\exp\{-\theta(F_i(x)-F_i(\gamma))\}\cdot \mathbb{1}_{[\gamma,\infty)}(x)$$

となる．ここで, F_i' は F_i の導関数を意味する．そして, $\{F_1(X_1),\ldots,F_n(X_n)\}$ は共通の切断パラメータ γ を持つと仮定する．すると, $\{F_1(X_1),\ldots,F_n(X_n)\}$ の分布は $Q_{\theta,\gamma}$ となる．

γ 共通サブモデルの代表例には, γ 共通の Pareto 分布族 (Rohatgi and Saleh [15]) や γ 共通の切断指数分布族 (Ghosh and Razmpour [11]) などが挙げられる．

では, γ 共通サブモデルの曲率を調べるために, Riemann 計量テンソルと Levi-Civita 接続係数を計算する．Riemann 計量を計算すると,

$$\begin{aligned} g_{ij} &= \frac{1}{(\theta^i)^2}\delta_{ij} \\ g_{i\gamma} &= 0 \\ g_{\gamma\gamma} &= \{G(\theta,\gamma)\}^2 \end{aligned}$$

となる．そして, Levi-Civita 接続係数は

$$\begin{aligned} \Gamma_{ij,k} &= -\frac{1}{(\theta^i)^3}\delta_{ijk} \quad (i,j,k=1,\ldots,n), \\ \Gamma_{ij,\gamma} &= 0, \\ \Gamma_{i\gamma,j} &= 0, \\ \Gamma_{i\gamma,\gamma} &= F_i'(\gamma)G(\theta,\gamma), \\ \Gamma_{\gamma\gamma,i} &= -F_i'(\gamma)G(\theta,\gamma), \\ \Gamma_{\gamma\gamma,\gamma} &= G(\theta,\gamma)\,\partial_\gamma G(\theta,\gamma) \end{aligned}$$

となる．ただし, $G(\theta,\gamma)=\sum\theta^i F_i'(\gamma)$ とした．なお, $F_i'(\gamma)$ は導関数 $dF_i/d\gamma(\gamma)$ を表す．

上記の計算結果から, γ 共通サブモデルの曲率テンソル

$$\begin{aligned} R_{i\gamma j\gamma} &= \frac{F_i'(\gamma)}{\theta^i}G(\theta,\gamma)\delta_{ij}, \\ R_{ijab} &= R_{abij} = 0 \end{aligned}$$

が得られる．なお, 曲率テンソルの他の成分は 0 または等式

$$R_{abcd} = -R_{bacd} = -R_{abdc} \quad (a,b,c,d=1,\ldots,n,\gamma)$$

により求められる [13].

そして, 曲率テンソルから γ 共通サブモデルのスカラー曲率を計算すると, 以下のように定数となる.

定理 4.4 (Yoshioka and Tanaka [19]).

γ 共通サブモデルのスカラー曲率は

$$\mathcal{R} = -2$$

である. ■

この定理は, Pareto 分布族が定曲率という結果を, $n+1$ 次元に拡張したものである. ただし, 主張はスカラー曲率定数に弱められている. 一般の Riemann 多様体において, スカラー曲率が定数となることは定曲率を意味しない. しかし, $1+1$ 次元の場合, つまり 2 次元多様体に限っては定曲率とスカラー曲率が定数になることは同値である. この事実は先行研究における Pareto 分布族が定曲率 -2 という結果に該当する.

では, γ 共通サブモデルが定曲率になるのはどんな場合だろうか. 実は, $n=1$ でのみ定曲率になることが分かっている.

命題 4.5.

γ 共通サブモデルは $n=1$ でのみ定曲率. ■

証明. $n=1$ で定曲率 -2 となることはすでに述べた. $n \geq 2$ では定曲率にならないことを示す.

$\mathcal{Q}$ を $n \geq 2$ の γ 共通サブモデルとし, 定曲率 $K \in \mathbb{R}$ であると仮定する.

このとき, 定曲率の同値条件 [13]

$$R_{abcd} = K\left(g_{ad}g_{bc} - g_{ac}g_{bd}\right)$$

より, モデル $\mathcal{Q}$ の曲率テンソルは

$$\begin{aligned} R_{i\gamma i\gamma} - K\left(g_{i\gamma}g_{\gamma i} - g_{ii}g_{\gamma\gamma}\right) &= \frac{F_i'(\gamma)}{\theta^i}G\left(\theta,\gamma\right) + \frac{K}{\left(\theta^i\right)^2}\left\{G\left(\theta,\gamma\right)\right\}^2 \\ &= 0 \end{aligned}$$

を満たす. この式を変形すると,

$$F_i'(\gamma) + K\sum_{j=1}^{n}\frac{\theta^j}{\theta^i}F_j'(\gamma) = 0$$

が得られる．

よって，$F_i'(\gamma) > 0$ より，$\theta^j \,/\, \theta^i$ はすべて定数である．

しかし，これは (θ, γ) が $n+1$ 次元の座標系であることに矛盾する．したがって，Ω は定曲率でない． ■

5　まとめと今後の展望

本稿では，非正則モデルの一つである片側切断指数型分布族に幾何構造を入れ，いくつかの幾何的性質を導いた．幾何構造として，最尤推定量の漸近的性質に基づいて定義されたと，等積になるような α 接続の拡張を用いた．そのもとで，また，あるサブモデルにおいてスカラー曲率が一定となることを示し，Pareto 分布族が定曲率という定理を拡張した．

今後の研究の方向性として，β 接続と高次漸近理論との関係を詳しく調べることが挙げられる．例えば，最尤推定量の漸近分散において，切断パラメータの有無が高次の項に影響すると知られており [1]，幾何的性質による表現が期待される．また，切断指数型分布族から切断分布族 [2] への理論の拡張も考えられる．

6　謝辞

本研究は JSPS 科研費 JP23K11006 及び JST 次世代研究者挑戦的研究プログラム JPMJSP2138 の助成を受けたものである．

参考文献

[1] Akahira, M. (2017). *Statistical Estimation for Truncated Exponential Families.* SpringerBriefs in Statistics. Springer, Singapore. xi+122.

[2] Akahira, M. (2021). “Maximum Likelihood Estimation for a One-Sided Truncated Family of Distributions”. In: *Jpn J Stat Data Sci* **4**, pp. 317–344.

[3] Akahira, M. and Ohyauchi, N. (2017). “Second-Order Asymptotic Loss of the MLE of a Truncation Parameter for a Truncated Exponential Family of Distributions”. In: *Comm. Statist. Theory Methods* **46**, pp. 6085–6097.

[4] Amari, S.-i. (1985). *Differential-Geometrical Methods in Statistics.* **28**. Lecture Notes in Statistics. Berlin: Springer-Verlag.

[5] Amari, S.-i. and Nagaoka, H. (2000). *Methods of Information Geometry*. **191**. Translations of Mathematical Monographs. Providence, RI: American Mathematical Society. x+206.

[6] Bar-Lev, S. K. (1984). "Large Sample Properties of the Mle and Mcle for the Natural Parameter of a Truncated Exponential Family". In: *Ann Inst Stat Math* **36**, pp. 217–222.

[7] Bartlett, M. S. (1953). "Approximate Confidence Intervals". In: *Biometrika* **40**, pp. 12–19.

[8] Bernardo, J. M. (1979). "Reference Posterior Distributions for Bayesian Inference". In: *Journal of the Royal Statistical Society: Series B (Methodological)* **41**, pp. 113–128.

[9] Ghosal, S. (1997). "Reference Priors in Multiparameter Nonregular Cases". In: *Test* **6**, pp. 159–186.

[10] Ghosh, J. K. and Mukerjee, R. (1992). "Non-Informative Priors". In: *Bayesian Statistics, 4 (Peñíscola, 1991)*. Oxford Univ. Press, New York, pp. 195–210.

[11] Ghosh, M. and Razmpour, A. (1984). "Estimation of the Common Location Parameter of Several Exponentials". In: *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* **46**, pp. 383–394.

[12] Jeffreys, H. (1961). *Theory of Probability*. Third. Oxford: Clarendon Press. ix+447.

[13] Kobayashi, S. and Nomizu, K. (1963). *Foundations of Differential Geometry. Vol I*. New York-London: Interscience Publishers (a division of John Wiley & Sons, Inc.) xi+329.

[14] Li, M. et al. (2022). "Fisher–Rao Geometry and Jeffreys Prior for Pareto Distribution". In: *Communications in Statistics - Theory and Methods* **51**, pp. 1895–1910.

[15] Rohatgi, V. K. and Saleh, A. K. M. E. (1987). "Estimation of the Common Scale Parameter of Two Pareto Distributions in Censored Samples". In: *Naval Research Logistics (NRL)* **34**, pp. 235–238.

[16] Rylov, A. (2018). "Constant Curvature Connections On Statistical Models". In: *Information Geometry and Its Applications*. Ed. by N. Ay et al. **252**. Cham: Springer International Publishing, pp. 349–361.

[17] Sun, F. et al. (2021). "The Bayesian Inference of Pareto Models Based on Information Geometry". In: *Entropy* **23**, p. 45.

[18] Takeuchi, J. and Amari, S. (2005). "Alpha-Parallel Prior and Its Properties". In: *IEEE Transactions on Information Theory* **51**, pp. 1011–1023.

[19] Yoshioka, M. and Tanaka, F. (in press). "Information-Geometric Approach for One-Sided Truncated Exponential Family". In: *Entropy*.

[20] 藤原 彰夫 (2021).『情報幾何学の基礎: 情報の内的構造を捉える新たな地平』. Tokyo: 共立出版.

[21] 野水 克己・佐々木 武 (1994).『アファイン微分幾何学: アファインはめ込みの幾何』. Tokyo: 裳華房.