Chentsov の定理とその周辺 (I)*

藤原彰夫

12 September 2016

1　はじめに

サイズ n の根元事象系を標準的に

$$\Omega_n = \{1, 2, \ldots, n\}$$

でラベルづける．そして Ω_n 上の確率分布全体の集合を

$$\mathcal{P}(\Omega_n) := \left\{ p : \Omega_n \to \mathbb{R}_{++} \;\middle|\; \sum_{\omega \in \Omega_n} p(\omega) = 1 \right\}$$

と表す．ここに $\mathbb{R}_{++} := \{x \in \mathbb{R} \,|\, x > 0\}$ である．$\mathcal{P}(\Omega_n)$ の各元である確率分布を n 次元数ベクトル

$$(p(1), p(2), \ldots, p(n))$$

と同一視すれば，集合 $\mathcal{P}(\Omega_n)$ は $\mathbb{R}^n$ 内の（開）単体とみなせる．これは <u>確率単体</u> と呼ばれる $n-1$ 次元多様体であり，以下では $\mathcal{S}_{n-1}$ で表す．

Chentsov の定理（informal version）

$\mathcal{S}_{n-1}$ 上の計量 g，アファイン接続 ∇ で，“ある種の不変性” を満たすものは

$$\begin{array}{ccl} g & \to & \text{Fisher 計量（の定数倍）} \\ \nabla & \to & \alpha\text{-接続（}\forall \alpha \in \mathbb{R}\text{）} \end{array}$$

に限られる．

1.1　“ある種の不変性” とは？

例えば 3 つの根元事象 $\{a, b, c\}$ を持つ確率的情報源があったとしよう．これを $\Omega_3 = \{1, 2, 3\}$ 上の確率分布で記述する場合，各事象 a, b, c をラベル $1, 2, 3$ のどれに割り当てるかは全く本質ではない．従って，対応する確率分布空間の幾何構造もラベルの取り替えに関して不変であるべきである．さらに，本当は $\{a, b, c\}$ 上の確率事象なのに，データを送信する機械の故障により，b と c がどちらも $*$ という記号に置き換わってしまう状況を想像しよう．受信者にとってみれば，最初から $\{a, *\}$ だったのか，それとも背後に $\{a, b, c\}$ があったのかは識別できない．こうした事情は，統計学における “十分統計量” の概念を通じ，確率分布空間の幾何構造に決定的な影響を与えることになる．

* ミニワークショップ：統計多様体の幾何学とその周辺 (8)（@北大数学）

1.2　十分統計量

定義

$\mathcal{X}$，$\mathcal{Y}$ を有限集合とし，$\{P_\theta \,|\, \theta \in \Theta\}$ を $\mathcal{X}$ 上のパラメトリックモデルとする．写像

$$F : \mathcal{X} \to \mathcal{Y}$$

がパラメタ θ の 十分統計量 であるとは，$F(X)$ が与えられた時の条件つき分布

$$P_\theta(X = x \,|\, F(X) = y) := \frac{P_\theta(X = x \wedge F(X) = y)}{P_\theta(F(X) = y)} \qquad (x \in \mathcal{X},\ y \in \operatorname{Ran} F)$$

が θ に依存しないことである．

十分統計量に関しては,「十分性」の名前の由来ともなっている次の事実が基本的である.

定理 1.1 (分解定理). F が十分統計量であるための必要十分条件は

$$P_\theta(X = x) = g_\theta(F(x)) \cdot h(x)$$

と分解できることである.

Proof. （必要性）$F(x) = y$ なる任意の $x \in \mathcal{X}$，$y \in \mathcal{Y}$ に対し，

$$\begin{aligned} P_\theta(X = x) &= P_\theta(X = x \wedge F(X) = y) \\ &= \underbrace{P_\theta(X = x \,|\, F(X) = y)}_{h(x)} \cdot \underbrace{P_\theta(F(X) = y)}_{g_\theta(y)} \end{aligned}$$

（十分性）$\forall y \in \operatorname{Ran} F$ に対し

$$P_\theta(F(X) = y) = \sum_{x : F(x) = y} P_\theta(X = x) = \sum_{x : F(x) = y} g_\theta(F(x)) \cdot h(x) = g_\theta(y) \sum_{x : F(x) = y} h(x)$$

だから，$F(x) = y$ なる任意の $x \in \mathcal{X}$，$y \in \mathcal{Y}$ に対し，

$$\begin{aligned} P_\theta(X = x \,|\, F(X) = y) &= \frac{P_\theta(X = x \wedge F(X) = y)}{P_\theta(F(X) = y)} = \frac{P_\theta(X = x)}{P_\theta(F(X) = y)} \\ &= \frac{g_\theta(y) \cdot h(x)}{g_\theta(y) \displaystyle\sum_{x : F(x) = y} h(x)} = \frac{h(x)}{\displaystyle\sum_{x : F(x) = y} h(x)} \end{aligned}$$

これは確かに θ に依存しない. □

例 1.2. 【指数型分布族】

$$p_\theta(x) = \exp\left[C(x) + \theta^i F_i(x) - \psi(\theta)\right] = \underbrace{\exp\left[C(x)\right]}_{h(x)} \times \underbrace{\exp\left[\theta^i F_i(x) - \psi(\theta)\right]}_{g_\theta(F_1(x), \ldots, F_d(x))}$$

例 1.3. 【壊れたデータ送信機】

$\mathcal{X} = \{a, b, c\}$，$\mathcal{Y} = \{a, *\}$ とするとき，壊れた送信機は

$$F(a) = a, \qquad F(b) = F(c) = *$$

で表される．$\mathcal{X}$ を Ω_3 で，$\mathcal{Y}$ を Ω_2 でラベルづけると，$F : \Omega_3 \to \Omega_2$ は次のようになる．

$$F(\omega) = \begin{cases} 1, & (\omega = 1) \\ 2, & (\omega = 2, 3) \end{cases}$$

さて，粗視化 F で情報が落ちないような（すなわち F を十分統計量とする）Ω_3 上の１次元パラメトリックモデルがあったとしたら，それは Ω_2 上の確率分布全体の集合 $\mathcal{S}_{2-1}$ と統計学的には同一視できる．

では，F を十分統計量とする Ω_3 上のモデルはどのようなものがあるだろうか？例 1.2 を参考に，指数型分布族の形でそれを実現してみよう．例えば

$$C(\omega) = \begin{cases} 0, & (\omega = 1) \\ \log Q^2, & (\omega = 2) \\ \log Q^3, & (\omega = 3) \end{cases}$$

としてみる．ここに Q^2 と Q^3 は，$Q^2 + Q^3 = 1$ を満たす正の実数とする．すると

$$p_\theta(\omega) = \exp\left[C(\omega) + \theta F(\omega) - \psi(\theta)\right] = \left(\frac{1}{1+e^\theta}, \ \frac{e^\theta}{1+e^\theta} Q^2, \ \frac{e^\theta}{1+e^\theta} Q^3\right)$$

一方，Ω_2 上の確率分布は

$$q_\theta = \left(\frac{1}{1+e^\theta}, \ \frac{e^\theta}{1+e^\theta}\right)$$

こうして，任意定数 Q^2，Q^3 を選ぶごとに，Ω_2 上の確率分布全体の集合 $\mathcal{S}_{2-1} = \{q_\theta\}_\theta$ と統計学的に同等な Ω_3 上の１次元パラメトリックモデル p_θ が作られる．

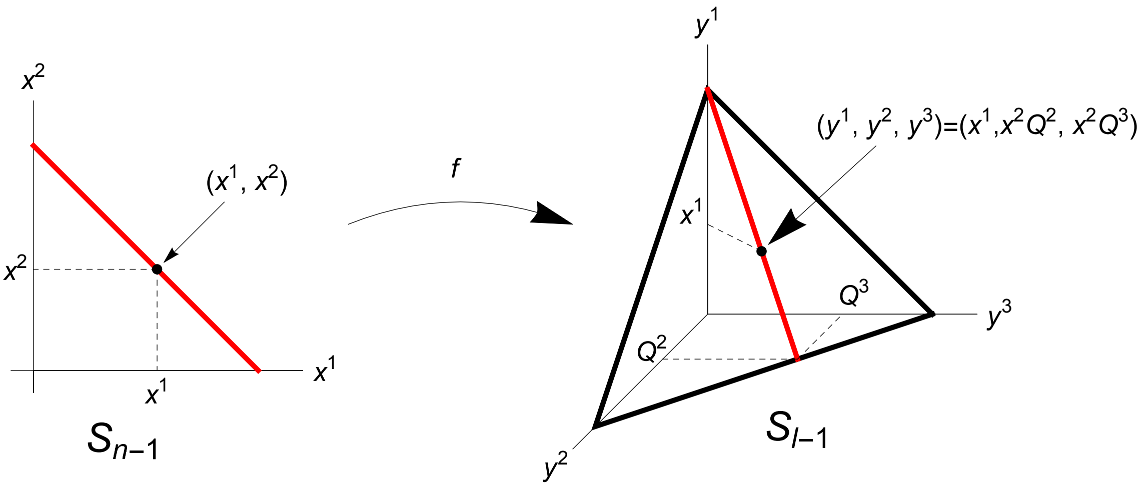

Figure 1: Markov 埋め込み f による $\mathcal{S}_{n-1}$ の $\mathcal{S}_{\ell-1}$ への埋め込み．ここでは，例 1.3 に合わせて $n = 2$，$\ell = 3$ とし，$C_{(1)} = \{1\}$，$C_{(2)} = \{2, 3\}$，$Q_{(1)} = (1, 0, 0)$，$Q_{(2)} = (0, Q^2, Q^3)$ としている．

1.3 Markov 埋め込み

上記着想を一般化すると，次の概念に自然に到達する．

定義

n, ℓ は $2 \leq n \leq \ell$ を満たす自然数とする．以下のように構成される写像

$$f : \mathcal{S}_{n-1} \longrightarrow \mathcal{S}_{\ell-1}$$

を Markov 埋め込み という．

(i) Ω_ℓ を，空でなく互いに交わらない部分集合の族 $\{C_{(1)}, \ldots, C_{(n)}\}$ に分割する．

(ii) 各 j（$1 \leq j \leq n$）に対し，$C_{(j)}$ に台を持つ Ω_ℓ 上の確率分布

$$Q_{(j)} = \left(Q^1_{(j)}, Q^2_{(j)}, \ldots, Q^\ell_{(j)}\right)$$

を付随させる．ここに $Q^k_{(j)}$ は，$k \in C_{(j)}$ ならば正の値であり，$k \notin C_{(j)}$ ならばゼロとする．

(iii) $(y^1, \ldots, y^\ell) = f(x^1, \ldots, x^n)$ を次で定義する．

$$y^k := \sum_{j=1}^{n} x^j Q^k_{(j)}.$$

なお，$n = \ell$ の時はラベルの入れ替えに相当するので，上で想定した不変性が課せられる状況はすべて Markov 埋め込みで表現される．

要請

$\mathcal{S}_{n-1}$ の幾何は，$\mathcal{S}_{\ell-1}$ に埋め込まれた部分多様体 $f(\mathcal{S}_{n-1})$ の幾何と同等であるべきである．

ここまでの話の流れからも分かるように，上記要請は $\mathcal{S}_{n-1}$ が確率分布からなる空間であるということから当然満たされるべき自然な要請である．ところが驚くべきことに，この要請だけから $\mathcal{S}_{n-1}$ の幾何構造がほとんど決定されてしまうのである．これが Chentsov の定理である．

2 Chentsov の定理

2.1 $(0, 2)$ 型テンソル

定理 2.1 (Chentsov)**.** $\mathcal{S}_{n-1}$ 上の $(0, 2)$ 型テンソル場 $g^{[n]}$ からなる列 $\{g^{[n]} \mid n = 2, 3, \ldots\}$ であって，任意の *Markov* 埋め込み f に関する不変性

$$g^{[n]}_p(X, Y) = g^{[\ell]}_{f(p)}(f_* X, f_* Y) \tag{1}$$

を満たすものは，定数倍を除いて

$$g^{[n]}_p(X, Y) = \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega)) \tag{2}$$

に限られる．

Proof. Markov 埋め込みは自然に $f : \mathbb{R}^n_{++} \to \mathbb{R}^\ell_{++}$ に拡張される．そこで，不変性の要請 (1) を満たす $\mathbb{R}^n_{++}$ の $(0, 2)$ 型テンソル場の列に対する同様の性質を導き，それを $\mathcal{S}_{n-1}$ に制限することで定

理を証明しよう．証明を４つのステップに分けて示す．以下では，$\mathbb{R}^n_{++} = \{p = (x^1, \ldots, x^n) \,|\, x^i > 0\}$ の大域的座標系として $(x^1, \ldots, x^n)$ 自身を用いる．また

$$\mathbb{R}^n_{++} = \bigsqcup_{\alpha > 0} \mathcal{S}^{(\alpha)}_{n-1}$$

と分解しておく．ここに

$$\mathcal{S}^{(\alpha)}_{n-1} := \left\{ (x^1, \ldots, x^n) \in \mathbb{R}^n_{++} \;\middle|\; \sum_{i=1}^n x^i = \alpha \right\}.$$

【ステップ１】$\mathcal{S}^{(\alpha)}_{n-1}$ の重心

$$p_0 = \left(\frac{\alpha}{n}, \ldots, \frac{\alpha}{n} \right)$$

で考える．Markov 埋め込みによる不変性 (1) の特殊ケースとして，$\ell = n$ の場合，すなわち事象のラベルづけに関する不変性を考えると，

$$g^{[n]}_{p_0} \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^i} \right)$$

は i によらないはずであり，$i \neq j$ に対し，

$$g^{[n]}_{p_0} \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right)$$

は i, j によらないはずである．従って，α のみに依存する滑らかな関数の列 $A^{[n]}(\alpha)$，$B^{[n]}(\alpha)$ が存在して

$$g^{[n]}_{p_0} \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) = \delta_{ij} A^{[n]}(\alpha) + B^{[n]}(\alpha)$$

【ステップ２】ある自然数 N が存在して $\ell = Nn$ となっている状況で

$$f(x^1, \ldots, x^n) = (\underbrace{\frac{x^1}{N}, \ldots, \frac{x^1}{N}}_{N}, \;\ldots\ldots,\; \underbrace{\frac{x^n}{N}, \ldots, \frac{x^n}{N}}_{N}) =: (y^{1_1}, \ldots, y^{1_N}, \;\ldots\ldots,\; y^{n_1}, \ldots, y^{n_N})$$

という Markov 埋め込みを考える．このとき

$$f_* : \frac{\partial}{\partial x^i} \;\longmapsto\; \frac{1}{N} \sum_{r=1}^N \frac{\partial}{\partial y^{i_r}}$$

であり，$\mathcal{S}^{(\alpha)}_{n-1}$ の重心 p_0 の像 $f(p_0)$ は $\mathcal{S}^{(\alpha)}_{\ell-1}$ の重心となる．よって不変性 (1) より

$$\begin{aligned}
\delta_{ij} A^{[n]}(\alpha) + B^{[n]}(\alpha) &= g^{[n]}_{p_0} \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \\
&= g^{[\ell]}_{f(p_0)} \left(f_* \frac{\partial}{\partial x^i}, f_* \frac{\partial}{\partial x^j} \right) \\
&= g^{[\ell]}_{f(p_0)} \left(\frac{1}{N} \sum_{r=1}^k \frac{\partial}{\partial y^{i_r}}, \frac{1}{N} \sum_{s=1}^N \frac{\partial}{\partial y^{j_s}} \right) \\
&= \frac{1}{N^2} \sum_{r,s=1}^N g^{[\ell]}_{f(p_0)} \left(\frac{\partial}{\partial y^{i_r}}, \frac{\partial}{\partial y^{j_s}} \right) \\
&= \frac{1}{N^2} \sum_{r,s=1}^N \left(\delta_{i_r, j_s} A^{[\ell]}(\alpha) + B^{[\ell]}(\alpha) \right) \\
&= \frac{\delta_{ij}}{N} A^{[\ell]}(\alpha) + B^{[\ell]}(\alpha)
\end{aligned}$$

これより
$$\frac{A^{[n]}(\alpha)}{n} = \frac{A^{[\ell]}(\alpha)}{\ell}, \qquad B^{[n]}(\alpha) = B^{[\ell]}(\alpha)$$
を得る．従ってある関数 $\lambda(\alpha)$，$\mu(\alpha)$ が存在して
$$A^{[n]}(\alpha) = \lambda(\alpha)n, \qquad B^{[n]}(\alpha) = \mu(\alpha)$$
となる．

【ステップ 3】$\alpha \in \mathbb{Q} \cap \mathbb{R}_{++}$ および $\mathcal{S}_{n-1}^{(\alpha)}$ 上の有理点 p を任意にとり，それを共通の分母を持つ分数で
$$p = \left(\frac{m_1}{\ell}, \ldots, \frac{m_n}{\ell}\right), \qquad (\ell, m_1, \ldots, m_n \in \mathbb{N})$$
と表しておく．そして
$$f(x^1, \ldots, x^n) = (\underbrace{\frac{x^1}{m_1}, \ldots, \frac{x^1}{m_1}}_{m_1}, \ \ldots\ldots, \ \underbrace{\frac{x^n}{m_n}, \ldots, \frac{x^n}{m_n}}_{m_n}) =: (y^{1_1}, \ldots, y^{1_{m_1}}, \ \ldots\ldots, \ y^{n_1}, \ldots, y^{n_{m_n}})$$
という Markov 埋め込みを考えると，$f(p)$ は $\mathcal{S}_{\ell-1}^{(\alpha)}$ の重心だから，不変性 (1) より
$$\begin{aligned}
g_p^{[n]}\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) &= g_{f(p)}^{[\ell]}\left(f_* \frac{\partial}{\partial x^i}, f_* \frac{\partial}{\partial x^j}\right) \\
&= g_{f(p)}^{[\ell]}\left(\frac{1}{m_i}\sum_{r=1}^{m_i}\frac{\partial}{\partial y^{i_r}}, \frac{1}{m_j}\sum_{s=1}^{m_j}\frac{\partial}{\partial y^{j_s}}\right) \\
&= \frac{1}{m_i m_j}\sum_{r=1}^{m_i}\sum_{s=1}^{m_j} g_{f(p)}^{[\ell]}\left(\frac{\partial}{\partial y^{i_r}}, \frac{\partial}{\partial y^{j_s}}\right) \\
&= \frac{1}{m_i m_j}\sum_{r=1}^{m_i}\sum_{s=1}^{m_j}\left(\delta_{i_r, j_s} A^{[\ell]}(\alpha) + B^{[\ell]}(\alpha)\right) \\
&= \left(\frac{1}{m_i m_j}\sum_{r=1}^{m_i}\sum_{s=1}^{m_j}\delta_{ij}\delta_{rs}A^{[\ell]}(\alpha)\right) + B^{[\ell]}(\alpha) \\
&= \frac{\delta_{ij}}{(m_i)^2} m_i A^{[\ell]}(\alpha) + B^{[\ell]}(\alpha) \\
&= \frac{\delta_{ij}}{m_i}\lambda(\alpha)\ell + \mu(\alpha) \\
&= \lambda(\alpha)\frac{\delta_{ij}}{p(i)} + \mu(\alpha)
\end{aligned}$$
ここで最後から 2 番目の等号ではステップ 2 の結果を用いた．こうして不変性の要請 (1) を満たす $\mathbb{R}_{++}^n$ の計量は，$\mathcal{S}_{n-1}^{(\alpha)}$ 上のすべての有理点 p 上で
$$g_p^{[n]}\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \lambda(\alpha)\frac{\delta_{ij}}{p(i)} + \mu(\alpha)$$
と書かれることが分かった．さらにテンソル場の連続性により，結局すべての点 $p \in \mathbb{R}_{++}^n$ において上式が成立することが結論される．

【ステップ 4】上記結果を $\mathcal{S}_{n-1}$ へ制限する．まず，$\mathcal{S}_{n-1}$ に座標系を導入しよう．第 i 番目の事象が確率 1 で起こる確率分布を δ_i と記す．すなわち
$$\delta_i(\omega) = \begin{cases} 1, & \omega = i \\ 0, & \omega \neq i \end{cases}$$

である．すると，任意の点 $p \in \mathcal{S}_{n-1}$ に対し

$$p(\omega) = \sum_{a=1}^{n-1} \xi^a \delta_a(\omega) + \left(1 - \sum_{a=1}^{n-1} \xi^a\right) \delta_n(\omega) \tag{3}$$

を満たす正の実数の組 $(\xi^1, \ldots, \xi^{n-1})$ がただ一つ定まる．そこで，(ξ^a) を $\mathcal{S}_{n-1}$ の座標系にとることにする．つまり

$$p = (x^1, \ldots, x^{n-1}, x^n) = \left(\xi^1, \ldots, \xi^{n-1}, 1 - \sum_{a=1}^{n-1} \xi^a\right)$$

である．よって

$$\frac{\partial}{\partial \xi^a} = \frac{\partial x^i}{\partial \xi^a} \frac{\partial}{\partial x^i} = \frac{\partial}{\partial x^a} - \frac{\partial}{\partial x^n}$$

だから，$\lambda := \lambda(1)$，$\mu := \mu(1)$ とすると

$$\begin{aligned} g_p^{[n]}\left(\frac{\partial}{\partial \xi^a}, \frac{\partial}{\partial \xi^b}\right) &= g_p^{[n]}\left(\frac{\partial}{\partial x^a} - \frac{\partial}{\partial x^n}, \frac{\partial}{\partial x^b} - \frac{\partial}{\partial x^n}\right) \\ &= \lambda \left\{\frac{\delta_{ab}}{p(a)} + \frac{1}{p(n)}\right\} + \mu \left\{(1-1)(1-1)\right\} \\ &= \lambda \sum_{\omega=1}^{n} \frac{(\delta_a(\omega) - \delta_n(\omega))(\delta_b(\omega) - \delta_n(\omega))}{p(\omega)} \end{aligned}$$

ここで (3) より

$$\delta_a(\omega) - \delta_n(\omega) = \frac{\partial}{\partial \xi^a} p(\omega)$$

だから，上式はさらに変形できて

$$\begin{aligned} g_p^{[n]}\left(\frac{\partial}{\partial \xi^a}, \frac{\partial}{\partial \xi^b}\right) &= \lambda \sum_{\omega=1}^{n} \frac{\left(\frac{\partial}{\partial \xi^a} p(\omega)\right)\left(\frac{\partial}{\partial \xi^b} p(\omega)\right)}{p(\omega)} \\ &= \lambda \sum_{\omega=1}^{n} p(\omega) \left(\frac{\partial}{\partial \xi^a} \log p(\omega)\right)\left(\frac{\partial}{\partial \xi^b} \log p(\omega)\right) \end{aligned}$$

以上で定理は証明された．□

2.2 $(0,3)$ 型テンソル

定理 2.2 (Chentsov)**.** $\mathcal{S}_{n-1}$ 上の $(0,3)$ 型テンソル場 $S^{[n]}$ からなる列 $\{S^{[n]} \mid n = 2, 3, \ldots\}$ であって，任意の *Markov* 埋め込み f に関する不変性

$$S_p^{[n]}(X, Y, Z) = S_{f(p)}^{[\ell]}(f_* X, f_* Y, f_* Z) \tag{4}$$

を満たすものは，定数倍を除いて

$$S_p^{[n]}(X, Y, Z) = \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega))(Z \log p(\omega)) \tag{5}$$

に限られる．

Proof. 定理 2.1 の証明と同様，不変性の要請 (4) を満たす $\mathbb{R}^n_{++}$ のテンソル場を特徴づけ，それを $\mathcal{S}_{n-1}$ に制限することにより証明する．

【ステップ１】$\mathcal{S}^{(\alpha)}_{n-1}$ の重心

$$p_0 = \left(\frac{\alpha}{n}, \dots, \frac{\alpha}{n}\right)$$

で考える．Markov 埋め込みによる不変性 (4) の特殊ケースとして，$\ell = n$ の場合，すなわち事象の並べ替えに関する不変性を考えよう．

$$S^{[n]}_{p_0}\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right)$$

のインデックス i, j, k の組み合わせを，すべてが一致する組み合わせ，二つが一致して一つが異なる組み合わせ，すべてが異なる組み合わせ，の３通りに分けると，不変性の要請から，それぞれの組み合わせのところでは値が一致するので，ある滑らかな関数の列 $A^{[n]}(\alpha)$，$B^{[n]}(\alpha)$，$C^{[n]}(\alpha)$，$D^{[n]}(\alpha)$，$E^{[n]}(\alpha)$ が存在して

$$S^{[n]}_{p_0}\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right) = \delta_{ijk} A^{[n]}(\alpha) + \Big(\delta_{ij} B^{[n]}(\alpha) + \delta_{jk} C^{[n]}(\alpha) + \delta_{ki} D^{[n]}(\alpha)\Big) + E^{[n]}(\alpha)$$

となる．ここに δ_{ijk} は $i = j = k$ のときのみ 1 で，その他は 0 という記号である．

【ステップ２】ある自然数 N が存在して $\ell = Nn$ となっている状況で

$$f(x^1, \dots, x^n) = (\underbrace{\frac{x^1}{N}, \dots, \frac{x^1}{N}}_{N},\ \dots\dots,\ \underbrace{\frac{x^n}{N}, \dots, \frac{x^n}{N}}_{N}) =: (y^{1_1}, \dots, y^{1_N},\ \dots\dots,\ y^{n_1}, \dots, y^{n_N})$$

という Markov 埋め込みを考える．このとき，$\mathcal{S}^{(\alpha)}_{n-1}$ の重心 p_0 の像 $f(p_0)$ は $\mathcal{S}^{(\alpha)}_{\ell-1}$ の重心となるので，不変性 (4) より

$$\begin{aligned}
&\delta_{ijk} A^{[n]} + \Big(\delta_{ij} B^{[n]} + \delta_{jk} C^{[n]} + \delta_{ki} D^{[n]}\Big) + E^{[n]} \\
&\qquad = S^{[n]}_{p_0}\left(\frac{\partial}{\partial x^i},\ \frac{\partial}{\partial x^j},\ \frac{\partial}{\partial x^k}\right) \\
&\qquad = S^{[\ell]}_{f(p_0)}\left(f_*\,\frac{\partial}{\partial x^i},\ f_*\,\frac{\partial}{\partial x^j},\ f_*\,\frac{\partial}{\partial x^k}\right) \\
&\qquad = S^{[\ell]}_{f(p_0)}\left(\frac{1}{N}\sum_{r=1}^{N}\frac{\partial}{\partial y^{i_r}},\ \frac{1}{N}\sum_{s=1}^{N}\frac{\partial}{\partial y^{j_s}},\ \frac{1}{N}\sum_{t=1}^{N}\frac{\partial}{\partial y^{k_t}}\right) \\
&\qquad = \frac{1}{N^3}\left\{N\,\delta_{ijk} A^{[\ell]} + N^2\left(\delta_{ij} B^{[\ell]} + \delta_{jk} C^{[\ell]} + \delta_{ki} D^{[\ell]}\right) + N^3 E^{[\ell]}\right\} \\
&\qquad = \frac{1}{N^2}\delta_{ijk} A^{[\ell]} + \frac{1}{N}\left(\delta_{ij} B^{[\ell]} + \delta_{jk} C^{[\ell]} + \delta_{ki} D^{[\ell]}\right) + E^{[\ell]}
\end{aligned}$$

これより

$$\frac{A^{[n]}}{n^2} = \frac{A^{[\ell]}}{\ell^2}, \quad \frac{B^{[n]}}{n} = \frac{B^{[\ell]}}{\ell}, \quad \frac{C^{[n]}}{n} = \frac{C^{[\ell]}}{\ell}, \quad \frac{D^{[n]}}{n} = \frac{D^{[\ell]}}{\ell}, \quad E^{[n]} = E^{[\ell]}$$

を得る．従って関数 $\lambda(\alpha)$，$\mu_B(\alpha)$，$\mu_C(\alpha)$，$\mu_D(\alpha)$，$\nu(\alpha)$ が存在して

$$\begin{aligned}
&A^{[n]} = \lambda(\alpha)\, n^2, \\
&B^{[n]} = \mu_B(\alpha)\, n, \quad C^{[n]} = \mu_C(\alpha)\, n, \quad D^{[n]} = \mu_D(\alpha)\, n, \\
&E^{[n]} = \nu(\alpha)
\end{aligned}$$

【ステップ３】$\alpha \in \mathbb{Q} \cap \mathbb{R}_{++}$ および $\mathcal{S}^{(\alpha)}_{n-1}$ 上の有理点 p を任意にとり，それを共通の分母を持つ分数で

$$p = \left(\frac{m_1}{\ell}, \dots, \frac{m_n}{\ell}\right), \qquad (\ell, m_1, \dots, m_n \in \mathbb{N})$$

と表しておく．そして

$$f(x^1,\ldots,x^n)=(\underbrace{\frac{x^1}{m_1},\ldots,\frac{x^1}{m_1}}_{m_1},\ \ldots\ldots,\ \underbrace{\frac{x^n}{m_n},\ldots,\frac{x^n}{m_n}}_{m_n})=:(y^{1_1},\ldots,y^{1_{m_1}},\ \ldots\ldots,\ y^{n_1},\ldots,y^{n_{m_n}})$$

という Markov 埋め込みを考えると，$f(p)$ は $\mathcal{S}^{(\alpha)}_{\ell-1}$ の重心だから，不変性 (4) より

$$\begin{aligned}
S^{[n]}_p\left(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j},\frac{\partial}{\partial x^k}\right) &= S^{[\ell]}_{f(p)}\left(f_*\frac{\partial}{\partial x^i},f_*\frac{\partial}{\partial x^j},f_*\frac{\partial}{\partial x^k}\right)\\
&= S^{[\ell]}_{f(p)}\left(\frac{1}{m_i}\sum_{r=1}^{m_i}\frac{\partial}{\partial y^{i_r}},\frac{1}{m_j}\sum_{s=1}^{m_j}\frac{\partial}{\partial y^{j_s}},\frac{1}{m_k}\sum_{t=1}^{m_k}\frac{\partial}{\partial y^{k_t}}\right)\\
&= \frac{\delta_{ijk}}{(m_i)^3}\,m_i\,A^{[\ell]}+\frac{\delta_{ij}}{m_i^2m_k}m_im_kB^{[\ell]}+\frac{\delta_{jk}}{m_j^2m_i}m_jm_iC^{[\ell]}\\
&\qquad +\frac{\delta_{ki}}{m_k^2m_j}m_km_jD^{[\ell]}+\frac{1}{m_im_jm_k}m_im_jm_kE^{[\ell]}\\
&= \lambda(\alpha)\,\frac{\delta_{ijk}}{p(i)^2}+\mu_B(\alpha)\frac{\delta_{ij}}{p(i)}+\mu_C(\alpha)\frac{\delta_{jk}}{p(j)}+\mu_D(\alpha)\frac{\delta_{ki}}{p(k)}+\nu(\alpha)
\end{aligned}$$

これを連続拡張すれば，結局すべての点 $p\in\mathbb{R}^n_{++}$ において上式が成立することが結論される．

【ステップ４】上記結果を $\mathcal{S}_{n-1}$ へ制限する．(3) で定まる $\mathcal{S}_{n-1}$ の座標系 (ξ^a) を用い，$\lambda:=\lambda(1)$, $\mu_\bullet:=\mu_\bullet(1)$，$\nu:=\nu(1)$ とおけば，

$$\begin{aligned}
&S^{[n]}_p\left(\frac{\partial}{\partial\xi^a},\frac{\partial}{\partial\xi^b},\frac{\partial}{\partial\xi^c}\right)=S^{[n]}_p\left(\frac{\partial}{\partial x^a}-\frac{\partial}{\partial x^n},\frac{\partial}{\partial x^b}-\frac{\partial}{\partial x^n},\frac{\partial}{\partial x^c}-\frac{\partial}{\partial x^n}\right)\\
&=\lambda\left(\frac{\delta_{abc}}{p(a)^2}-\frac{1}{p(n)^2}\right)+\mu_B\left\{\underbrace{\left(\frac{\delta_{ab}}{p(a)}+\frac{1}{p(n)}\right)}_{\partial/\partial x^c\text{ に由来}}-\underbrace{\left(\frac{\delta_{ab}}{p(a)}+\frac{1}{p(n)}\right)}_{\partial/\partial x^n\text{ に由来}}\right\}\\
&\quad+\mu_C\left\{\underbrace{\left(\frac{\delta_{bc}}{p(b)}+\frac{1}{p(n)}\right)}_{\partial/\partial x^a\text{ に由来}}-\underbrace{\left(\frac{\delta_{bc}}{p(b)}+\frac{1}{p(n)}\right)}_{\partial/\partial x^n\text{ に由来}}\right\}+\mu_D\left\{\underbrace{\left(\frac{\delta_{ca}}{p(c)}+\frac{1}{p(n)}\right)}_{\partial/\partial x^b\text{ に由来}}-\underbrace{\left(\frac{\delta_{ca}}{p(c)}+\frac{1}{p(n)}\right)}_{\partial/\partial x^n\text{ に由来}}\right\}\\
&\quad+\nu\left\{(1-1)(1-1)(1-1)\right\}\\
&=\lambda\sum_{\omega=1}^{n}\frac{(\delta_a(\omega)-\delta_n(\omega))\,(\delta_b(\omega)-\delta_n(\omega))\,(\delta_c(\omega)-\delta_n(\omega))}{p(\omega)^2}\\
&=\lambda\sum_{\omega=1}^{n}\frac{\left(\frac{\partial}{\partial\xi^a}p(\omega)\right)\left(\frac{\partial}{\partial\xi^b}p(\omega)\right)\left(\frac{\partial}{\partial\xi^c}p(\omega)\right)}{p(\omega)^2}\\
&=\lambda\sum_{\omega=1}^{n}p(\omega)\left(\frac{\partial}{\partial\xi^a}\log p(\omega)\right)\left(\frac{\partial}{\partial\xi^b}\log p(\omega)\right)\left(\frac{\partial}{\partial\xi^c}\log p(\omega)\right)
\end{aligned}$$

以上で定理は証明された． □

2.3 Fisher 計量と α-接続

Markov 埋め込み $f:\mathcal{S}_{n-1}\to\mathcal{S}_{\ell-1}$ の下での $\mathcal{S}_{n-1}$ 上の計量 $g^{[n]}$ に対する不変性の要請は

$$g^{[n]}_p(X,Y)=g^{[\ell]}_{f(p)}(f_*X,f_*Y)\tag{6}$$

であり，$\mathcal{S}_{n-1}$ 上のアファイン接続 $\nabla^{[n]}$ に対する不変性の要請は

$$g_p^{[n]}(\nabla_X^{[n]}Y, Z) = g_{f(p)}^{[\ell]}(\nabla_{f_*X}^{[\ell]} f_*Y, f_*Z) \tag{7}$$

と書ける．このうち，要請 (6) については定理 2.1 で解決済み．そしてこの計量に付随する Levi-Civita 接続を $\overline{\nabla}^{[n]}$ と書くと，(6) より $\overline{\nabla}^{[n]}$ は自動的に (7) を満たす．すなわち，

$$\begin{aligned}
&g_{f(p)}^{[\ell]}(\overline{\nabla}_{f_*\partial_i}^{[\ell]} f_*\partial_j, f_*\partial_k) \\
&\qquad = \frac{1}{2}\left\{(f_*\partial_i)g_{f(p)}^{[\ell]}(f_*\partial_j, f_*\partial_k) + (f_*\partial_j)g_{f(p)}^{[\ell]}(f_*\partial_k, f_*\partial_i) - (f_*\partial_k)g_{f(p)}^{[\ell]}(f_*\partial_i, f_*\partial_j)\right\} \\
&\qquad = \frac{1}{2}\left\{\partial_i g_p^{[n]}(\partial_j, \partial_k) + \partial_j g_p^{[n]}(\partial_k, \partial_i) - \partial_k g_p^{[n]}(\partial_i, \partial_j)\right\} \\
&\qquad = g_p^{[n]}(\overline{\nabla}_{\partial_i}^{[n]}\partial_j, \partial_k)
\end{aligned}$$

従って，Levi-Civita 接続 $\overline{\nabla}^{[n]}$ は $\mathcal{S}_{n-1}$ 上に許容される接続の一つである．

さて，$\mathcal{S}_{n-1}$ 上に許容される接続はこれだけではない．多様体上に共変微分を一つ（例えば $\overline{\nabla}^{[n]}$）を固定すると，他の共変微分 $\nabla^{[n]}$ との差は $(1,2)$ 型テンソル場と 1 対 1 に対応する．従ってそれを計量で射影して考えれば，

$$g_p^{[n]}(\nabla_X^{[n]}Y, Z) - g_p^{[n]}(\overline{\nabla}_X^{[n]}Y, Z)$$

と $(0,3)$ 型テンソル場は 1 対 1 に対応することになる．そして $\nabla^{[n]}$ に不変性 (7) を要請することは，上記 $(0,3)$ 型テンソル場に不変性を要請することに他ならない．しかも不変性を満たす $(0,3)$ 型テンソル場は定数倍を除いて (5) に定まることが定理 2.2 で証明されている．そこで，この定数倍因子を改めて $-\dfrac{\alpha}{2}$ と置くと，実数 α を任意に与えるごとに

$$g_p^{[n]}(\nabla_X^{[n]}Y, Z) - g_p^{[n]}(\overline{\nabla}_X^{[n]}Y, Z) = -\frac{\alpha}{2}\, S_p^{[n]}(X, Y, Z)$$

によって不変性 (7) を満たす接続 $\nabla^{[n]}$ が一つ定まることが分かる．以上をまとめて，次の結論を得る．

定理 2.3 (Chentsov)**.** *Markov* 埋め込みの下での不変性 *(6)* を満たす $\mathcal{S}_{n-1}$ 上の *Riemann* 計量は，定数倍を除いて

$$g_p(X, Y) := \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega)) \tag{8}$$

に限られる．一方，不変性 *(7)* を満たすアファイン接続 $\nabla^{(\alpha)}$ は，関係

$$g_p(\nabla_X^{(\alpha)}Y, Z) := g_p(\overline{\nabla}_X Y, Z) - \frac{\alpha}{2}\, S_p(X, Y, Z) \tag{9}$$

により，実数 α と 1 対 1 に対応する．ここに $\overline{\nabla}$ は計量 *(8)* に付随する *Levi-Civita* 接続，S_p は

$$S_p(X, Y, Z) := \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega))(Z \log p(\omega))$$

で定義される $(0,3)$ 型対称テンソル場である．

定義

(8) で定まる計量 g を $\mathcal{S}_{n-1}$ の Fisher 計量といい，各実数 α に対し (9) で定まる接続 $\nabla^{(\alpha)}$ を $\mathcal{S}_{n-1}$ の α-接続という．