Majorization

大上由人

2024 年 5 月 9 日

1　古典的エントロピー及びダイバージェンス

1.1　古典的状態及び系

必要な量を定義する。

Def: 状態を表す確率分布

古典的系における状態は確率分布

$$p = (p_1, p_2, \cdots, p_d)^\top \tag{1.1}$$

で表される。ここで、$p_i \geq 0$ かつ $\sum_{i=1}^{d} p_i = 1$ である。また、d 次元の確率分布全体の集合を、$\mathcal{P}_d$ と表す。
また、その集合に属する一様分布を、

$$u = \left(\frac{1}{d}, \frac{1}{d}, \cdots, \frac{1}{d}\right)^\top \tag{1.2}$$

と表す。
また、異なる確率分布の積を、

$$p \otimes q \in \mathcal{P}_{dd'} \quad p \in \mathcal{P}_d, q \in \mathcal{P}_{d'} \tag{1.3}$$

と表し、とくに、同じ確率分布の累乗を、

$$p^{\otimes n} \in \mathcal{P}_{d^n} \quad p \in \mathcal{P}_d \tag{1.4}$$

と表す。

Def:Supp

確率分布 $p = p_{i_i} \in \mathcal{P}_d$ に対して、p の台を、

$$\mathrm{spp}(p) = \{i \in [d] | p_i > 0\} \subset \{1, 2, \cdots, d\} \tag{1.5}$$

と表す。また、

$$\mathrm{rank}(p) = |\mathrm{spp}(p)| \tag{1.6}$$

を、p のランクという。とくに、$\mathrm{rank}(p) = d$ のとき、p はフルランクであるという。

要するに、確率が 0 でないようなインデックスの集合を台と呼び、その要素数をランクと呼ぶ。

Def: 確率遷移行列

古典的な確率分布の時間発展は、確率遷移行列 T を用いて以下のように表される。

$$p'_i = \sum_{j=1} T_{ij} p_j \tag{1.7}$$

Prop: 確率遷移行列の性質

確率遷移行列 T は以下の性質を持つ。

$$\sum_{i=1}^{d} T_{ij} = 1 \tag{1.8}$$

Prf

略 (確率の規格化を利用する。) □

Def: 二重確率遷移行列

確率遷移行列 T が、

$$\sum_{j=1} T_{ij} = 1 \tag{1.9}$$

をみたすとき、二重確率遷移行列という。

Prop: 二重確率遷移行列の特徴づけ

以下の二つは同値である。

1. T は二重確率遷移行列である。
2. 一様分布 u は T に対して不変である。すなわち、$u = Tu$ である。

Prf

$$
\begin{aligned}
p'_i &= \sum_{j=1} T_{ij} u_j &(1.10)\\
&= \frac{1}{d} \sum_{j=1}^{d} T_{ij} &(1.11)\\
&= \frac{1}{d} \cdot d &(1.12)\\
&= 1 &(1.13)
\end{aligned}
$$

であることからわかる。

Def: トレース距離

二つの確率分布 p, q のトレース距離は、

$$D(p,q) = \frac{1}{2} \sum_{i=1}^{d} |p_i - q_i| \tag{1.14}$$

で定義される。

Prop: トレース距離の性質

トレース距離は、T に対して非増加である。すなわち、

$$D(p,q) \geq D(Tp, Tq) \tag{1.15}$$

が成り立つ。

Prf

後により一般の証明をするため、ここでは省略する。

1.2　シャノンエントロピー及び KL ダイバージェンス

Def: シャノンエントロピー

確率分布 $p \in \mathcal{P}_d$ のシャノンエントロピーは、

$$S_1(p) = -\sum_{i=1}^{d} p_i \log p_i \tag{1.16}$$

で定義される。

Def:KL ダイバージェンス

二つの確率分布 $p, q \in \mathcal{P}_d$ の KL ダイバージェンスは、

$$S_1(p||q) = \sum_{i=1}^{d} p_i \log \frac{p_i}{q_i} \tag{1.17}$$

で定義される。ただし、$\mathrm{supp}(p) \subset \mathrm{supp}(q)$ でないときは、$S_1(p||q) = \infty$ とする。

このとき、エントロピーと KL ダイバージェンスの関係がわかる。

Prop: エントロピーと KL ダイバージェンスの関係

任意の $p, q \in \mathcal{P}_d$ に対して、

$$S_1(p) = ln(d) - S_1(p||u) \tag{1.18}$$

Prf

$$\begin{aligned}
S_1(p||u) &= \sum_{i=1}^{d} p_i \log \frac{p_i}{\frac{1}{d}} && (1.19)\\
&= \sum_{i=1}^{d} p_i \log dp_i && (1.20)\\
&= \sum_{i=1}^{d} p_i \log d + \sum_{i=1}^{d} p_i \log p_i && (1.21)\\
&= \log d - S_1(p) && (1.22)
\end{aligned}$$

であることからわかる。 □

これより、$S_1(p) \leq \log d$ であることがわかる。

このとき、KL ダイバージェンスのテイラー展開は以下のようになる。

$$S_1(p||p - \Delta p) = \frac{1}{2} \sum_i \frac{(\Delta p_i)^2}{p_i} + O(\Delta p^3) \tag{1.23}$$

$$\tag{1.24}$$

ここで、$\sum_i \Delta p_i = 0$ を用いている。

Prop:KL ダイバージェンスの単調性

KL ダイバージェンスは、$p' = Tp$ および $q' = Tq$ に対して、

$$S_1(p||q) \geq S_1(p'||q') \tag{1.25}$$

が成り立つ。

Prf

後に一般に示す。

注意されたいこととして、KL ダイバージェンスの単調性の逆はいえない。すなわち、単調性を満たすが、$p' = Tp$ および $q' = Tq$ を満たすような T が存在しない場合がある。

次に、二重確率遷移行列について考える。このとき、KL ダイバージェンスの単調性と、

$$S_1(p) \leq S_1(Tp) \tag{1.26}$$

が成り立つ。
すなわち、二重確率遷移行列による時間発展は、エントロピーを増加させる。

Def: 相互情報量

二つの確率分布 $p, q \in \mathcal{P}_d$ の相互情報量は、

$$I_1(p_{AB})_{A:B} = S_1(p_A) + S_1(p_B) - S_1(p_{AB}) = S_1(p_{AB}||p_A \otimes p_B) \geq 0 \tag{1.27}$$

で定義される。

この量は、A と B の相関を表す量である。

Prop: 相互情報量の性質

任意の $p, q \in \mathcal{P}_d$ に対して、

$$I_1(p_{AB})_{A:B} = 0 \Leftrightarrow p_{AB} = p_A \otimes p_B \tag{1.28}$$

が成り立つ。また、KL ダイバージェンスの単調性から、

$$I_1(p_{AB})_{A:B} \geq I_1(T_A \otimes T_B p_{AB})_{A:B} \tag{1.29}$$

が成り立つ。ただし、$T_A \otimes T_B$ は、各 A,B に独立に作用する確率遷移行列である。

Prf

略 (過去のゼミ資料を参考せよ) $\square$

1.3 Rényi エントロピー及びダイバージェンス

シャノンエントロピーを包含する概念として、Rényi エントロピーがある。

Def:Rényi エントロピー

確率分布 $p \in \mathcal{P}_d$ の Rényi エントロピーは、$0 \leq \alpha \leq \infty$、$p \in \mathcal{P}_d$ に対して、

$$S_\alpha(p) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{d} p_i^\alpha\right) \tag{1.30}$$

で定義される。

また、ダイバージェンスについても、Rényi ダイバージェンスがある。

Def:Rényi ダイバージェンス

二つの確率分布 $p, q \in \mathcal{P}_d$ の Rényi ダイバージェンスは、$0 \leq \alpha \leq \infty$、$p \in \mathcal{P}_d$ に対して、

$$S_\alpha(p||q) = \frac{1}{\alpha-1} \log\left(\sum_{i=1}^{d} p_i^\alpha q_i^{1-\alpha}\right) \tag{1.31}$$

で定義される。ただし、$\mathrm{supp}(p) \subset \mathrm{supp}(q)$ でないときは、$S_\alpha(p||q) = \infty$ とする。

これらの量が、たしかにシャノンエントロピーと KL ダイバージェンスを包含していることを示す。

Prop:Rényi エントロピーとシャノンエントロピーの関係

Rényi-1 エントロピーは、シャノンエントロピーに一致する。すなわち、

$$S_1(p) = S_\alpha(p)|_{\alpha=1} \tag{1.32}$$

が成り立つ。

Prf

$$\begin{aligned}
S_\alpha(p)|_{\alpha=1} &= -lim_{\alpha\to 1} \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{d} p_i^\alpha\right) && (1.33)\\
&= -\frac{d}{d\alpha} \log\left(\sum_{i=1}^{d} p_i^\alpha\right)|_{\alpha=1} && (1.34)\\
&= -\frac{\sum_{i=1}^{d} p_i \log p_i}{\sum_{i=1}^{d} p_i} && (1.35)\\
&= -\sum_{i=1}^{d} p_i \log p_i && (1.36)\\
&= S_1(p) && (1.37)
\end{aligned}$$

であることからわかる。 □

Prop:Rényi ダイバージェンスとダイバージェンスの関係

Rényi-1 ダイバージェンスは、KL ダイバージェンスに一致する。すなわち、

$$S_\alpha(p||q)|_{\alpha=1} = S_1(p||q) \tag{1.38}$$

が成り立つ。

Prf

$$\begin{aligned}
S_\alpha(p||q)|_{\alpha=1} &= lim_{\alpha\to 1}\frac{1}{\alpha-1}\log\left(\sum_{i=1}^d p_i^\alpha q_i^{1-\alpha}\right) && (1.39)\\
&= \frac{d}{d\alpha}\log\left(\sum_{i=1}^d p_i^\alpha q_i^{1-\alpha}\right)|_{\alpha=1} && (1.40)\\
&= \frac{\sum_{i=1}^d q_i(p_i/q_i)^\alpha \log(p_i/q_i)}{\sum_{i=1}^d p_i^\alpha q_i^{1-\alpha}}|_{\alpha=1} && (1.41)\\
&= \sum_{i=1}^d p_i \log\frac{p_i}{q_i} && (1.42)\\
&= S_1(p||q) && (1.43)
\end{aligned}$$

であることからわかる。 □

また、$\alpha = 0, \infty$ の場合は重要らしいので、以下で確認する。

Prop:Rényi エントロピー/ダイバージェンスの極限

Rényi-0 エントロピーは、

$$S_0(p) = \log(\mathrm{rank}(p)) \tag{1.44}$$

で定義される。また、Rényi-∞ エントロピーは、

$$S_\infty(p) = -(\log \max_i p_i) \tag{1.45}$$

で定義される。また、Rényi-0 ダイバージェンスは、

$$S_0(p||q) = -\log\left(\sum_{i:p_i>0} q_i\right) \tag{1.46}$$

で定義される。また、Rényi-∞ ダイバージェンスは、

$$S_\infty(p||q) = \log\left(\max_i \frac{p_i}{q_i}\right) \tag{1.47}$$

で定義される。

Prf

Prop:Rényi エントロピーと KL ダイバージェンスの関係

任意の $p, q \in \mathcal{P}_d$ に対して、

$$S_\alpha(p) = \frac{1}{1-\alpha}\log d - S_\alpha(p||u) \tag{1.48}$$

が成り立つ。

以下、Rényi ダイバージェンスについての性質を示す。

Prop:Rényi ダイバージェンスの非負性

任意の $p, q \in \mathcal{P}_d$ に対して、

$$S_\alpha(p||q) \geq 0 \tag{1.49}$$

が成り立つ。また、$0 < \alpha \leq \infty$ に対して、

$$S_\alpha(p||q) = 0 \Leftrightarrow p = q \tag{1.50}$$

であり、また、$\alpha = 0$ のとき、

$$S_0(p||q) = 0 \Leftrightarrow \mathrm{supp}(p) \subset \mathrm{supp}(q) \tag{1.51}$$

が成り立つ。

Prop:Rényi ダイバージェンスの単調性

任意の $p' = Tp, q' = Tq \in \mathcal{P}_d$ に対して、

$$S_\alpha(p||q) \geq S_\alpha(p'||q') \tag{1.52}$$

が成り立つ。また、二重確率遷移行列に対して、

$$S_\alpha(p) \leq S_\alpha(Tp) \tag{1.53}$$

が成り立つ。

Prop:Rényi ダイバージェンスの単調性 (2)

$\alpha \leq \alpha'$ に対して、

$$S_\alpha(p||q) \leq S_{\alpha'}(p||q) \tag{1.54}$$

が成り立ち、また、

$$S_\alpha(p) \geq S_{\alpha'}(p) \tag{1.55}$$

が成り立つ。

Lem:

f を下に凸な関数であるとし、$p, q, p', q' \in \mathbb{R}^d$ がすべて正であるとする。もし、$p' = Tp, q' = Tq$ であるとき、

$$\sum_{i=1}^{d} q_i' f\left(\frac{p_i'}{q_i'}\right) \leq \sum_{i=1}^{d} q_i f\left(\frac{p_i}{q_i}\right) \tag{1.56}$$

が成り立つ。

Prf

Jensen の不等式より、

$$\sum_{j=1}^{d} q_j' f\left(\frac{p_j'}{q_j'}\right) \leq \sum_{j=1}^{d}\sum_{i=1}^{d} \frac{T_{ji}q_i}{q_j'} f\left(\frac{p_i}{q_i}\right) = \sum_{i=1}^{d} q_i f\left(\frac{p_i}{q_i}\right) \tag{1.57}$$

である。 □

ここで、$p, q \in \mathcal{P}_d$ として、$f(x) = xlogx$ とすると、

$$S_1(p||q) \geq S_1(p'||q') \tag{1.58}$$

が成り立つ。これは、KL ダイバージェンスの単調性を示している。

Prf:(Rényi ダイバージェンスの非負性)

$f_\alpha(x) = x^\alpha$ とすると、$1 < \alpha \leq \infty$ に対して、$f_\alpha(x)$ は下に凸な関数である。

Jensen の不等式より

$$\sum_{i=1}^{d} q_i f(\frac{p_i}{q_i}) \geq f(\sum_{i=1}^{d} q_i \frac{p_i}{q_i}) = f(1) = 0 \tag{1.59}$$

である。両辺対数をとることにより、

$$\log\left(\sum_{i=1}^{d} q_i \frac{p_i}{q_i}\right) \geq 0 \tag{1.60}$$

である。これを両辺 $\frac{1}{\alpha-1}$ かけることにより、

$$\frac{1}{\alpha-1}\log\left(\sum_{i=1}^{d} q_i \frac{p_i}{q_i}\right) \geq 0 \tag{1.61}$$

である。したがって、

$$S_\alpha(p||q) \geq 0 \tag{1.62}$$

である。また、$0 < \alpha < 1$ のときは、上の Jensen の不等式で不等号が逆になり、$\frac{1}{\alpha-1}$ をかけるときにもう一度符号が逆になることに注意して、同様に示すことができる。
$\alpha = 0, 1, \infty$ の場合は、それぞれの定義から自明である。

□

Prf:(Rényi ダイバージェンスの単調性)
$1 < \alpha < \infty$ のとき、Lem で、$f(x) = x^\alpha$ として示した不等式を用いると、

$$\sum_{i=1}^{d} q'_i \frac{p_i'^\alpha}{q_i'^\alpha} \leq \sum_{i=1}^{d} q_i \frac{p_i^\alpha}{q_i^\alpha} \tag{1.63}$$

である。これの両辺対数をとることにより、

$$\log\left(\sum_{i=1}^{d} p_i'^\alpha q_i^{1-\alpha}\right) \leq \log\left(\sum_{i=1}^{d} p_i^\alpha q_i^{1-\alpha}\right) \tag{1.64}$$

である。したがって、

$$\frac{1}{\alpha-1}\log\left(\sum_{i=1}^{d} p_i'^\alpha q_i^{1-\alpha}\right) \leq \frac{1}{\alpha-1}\log\left(\sum_{i=1}^{d} p_i^\alpha q_i^{1-\alpha}\right) \tag{1.65}$$

である。したがって、

$$S_\alpha(p||q) \geq S_\alpha(p'||q') \tag{1.66}$$

である。
$0 < \alpha < 1$ のときは、同様に示すことができる。
$\alpha = 0, 1, \infty$ の場合は、それぞれの定義から自明である。

$\square$

Prf:(Rényi ダイバージェンスの単調性 (2))
$\alpha \leq \alpha'$ に対して、$f(x) = x^{\frac{\alpha-1}{\alpha'-1}}$ とすると、この関数は、$1 < \alpha < \alpha' < \infty$ に対して下に凸な関数であり、$0 < \alpha < \alpha' < 1$ および $0 < \alpha < 1 < \alpha'$ に対して上に凸な関数である。
したがって、Jensen の不等式より、

$$S_\alpha(p||q) = \frac{1}{\alpha-1}\log\left(\sum_{i=1} p_i \left(\frac{p_i}{q_i}\right)^{\alpha-1}\right) \tag{1.67}$$

$$= \frac{1}{\alpha-1}\log\left(\sum_{i=1} p_i \left(\frac{p_i}{q_i}\right)^{(alpha'-1)(\frac{\alpha-1}{\alpha'-1})}\right) \tag{1.68}$$

$$\leq \frac{1}{\alpha'-1}\log\left(\sum_{i=1} p_i \left(\frac{p_i}{q_i}\right)^{\alpha'-1}\right) \tag{1.69}$$

$$= S_{\alpha'}(p||q) \tag{1.70}$$

である。
また、$\alpha = 0, 1, \infty$ の場合は、それぞれの定義から自明である。

$\square$

さらに、f-ダイバージェンスという概念がある。

Def:f-ダイバージェンス

$f(0,\infty) \to \mathbb{R}$ を下に凸な関数とし、$x = 1$ で $f(x)$ が狭義凸かつ $f(1) = 0$ であるとする。このとき、$p, q \in \mathcal{P}_d$ に対して、

$$D_f(p||q) = \sum_{i=1}^{d} q_i f\left(\frac{p_i}{q_i}\right) \tag{1.71}$$

で定義される。

KL ダイバージェンスは、$f(x) = x\log x$ のときの f-ダイバージェンスである。また、Rényi ダイバージェンスは、$f(x) = x^\alpha$ として log をとって $\frac{1}{\alpha-1}$ をかけたものである。

Prop:f-ダイバージェンスの非負性

任意の $p,q \in \mathcal{P}_d$ に対して、

$$D_f(p||q) \geq 0 \tag{1.72}$$

が成り立つ。また、

$$D_f(p||q) = 0 \Leftrightarrow p = q \tag{1.73}$$

が成り立つ。

Prf

Jensen の不等式より、

$$\sum_{i=1}^{d} q_i f\left(\frac{p_i}{q_i}\right) \geq f\left(\sum_{i=1}^{d} q_i \frac{p_i}{q_i}\right) = f(1) = 0 \tag{1.74}$$

である。ことからわかる。 □

Prop:f-ダイバージェンスの単調性

任意の $p' = Tp, q' = Tq \in \mathcal{P}_d$ に対して、

$$D_f(p||q) \geq D_f(p'||q') \tag{1.75}$$

が成り立つ。

Prf

よくよく見ると、これは Lem で示した不等式と同じである。 □

1.4 フィッシャー情報量

以下、我々は、なめらかにパラメータ化された確率分布 $p(\theta)$ について考える。ただし、θ の取りうる領域は、$\mathbb{R}^m$ の開部分集合である。

Def: フィッシャー情報量

$p(\theta) \in \mathcal{P}_d$ がフルランクであるとし、$\theta \in \mathbb{R}^m$ をパラメータとする。このとき、フィッシャー情報量は $m \times m$ 行列で、

$$J_{p(\theta),kl} = \sum_{i=1}^{d} p_i(\theta)\partial_k[\log p_i(\theta)]\partial_l[\log p_i(\theta)] = \sum_{i=1}^{d} \frac{\partial_k p_i(\theta)\partial_l p_i(\theta)}{p_i(\theta)} \tag{1.76}$$

で定義される。

フィッシャー情報量は、f-ダイバージェンスの極限として得られる。

Prop: フィッシャー情報量の単調性

任意の確率遷移行列 T に対して、

$$J_{p(\theta)} \geq J_{Tp(\theta)} \tag{1.77}$$

が成り立つ。

Prf

$p' = Tp$ とし、$c = (c^1, ..., c^m) \in \mathbb{R}^m$ として、$\partial = \sum_k c^k \partial_k$ とする。このとき、

$$c^\top J_{p(\theta)} c = \sum_i p_i \left(\frac{\partial p_i}{p_i}\right)^2 \tag{1.78}$$

$$c^\top J_{p'(\theta)} c = \sum_i p'_i \left(\frac{\partial p'_i}{p'_i}\right)^2 \tag{1.79}$$

$$\tag{1.80}$$

である。したがって、Lemma で $f = x^2$ として示した不等式より、

$$c^\top J_{p(\theta)} c \geq c^\top J_{p'(\theta)} c \tag{1.81}$$

である。したがって、

$$J_{p(\theta)} \geq J_{Tp(\theta)} \tag{1.82}$$

である。 □

フィッシャー情報量の操作的な意味付けとして、Cramér-Rao の不等式がある。

Thm:Cramér-Rao の不等式

あるパラメータ θ に対する不偏推定量 θ_{est} に対して、不偏条件 $\sum_{i=1}^d p_i(\theta)\theta_{est}(i) = \theta$ が成り立つとする。このとき、正確さは共分散行列

$$Cov_\theta^{kl}(\theta_{est}) = \sum_{i=1} p_i(\theta)(\theta_{est}^k(i) - \theta^k)(\theta_{est}^l(i) - \theta^l) \tag{1.83}$$

により表現される。このとき、

$$\mathrm{Cov}_\theta^{kl}(\theta_{est}) \geq (J_{p(\theta)})_{kl}^{-1} \tag{1.84}$$

が成り立つ。

すなわち、フィッシャー情報量は、θ の不偏推定量が、フィッシャー情報量によって制限されることを示している。

　例として、指数型分布族とよばれる確率分布の集合族を考える。簡単のために、パラメータを θ のみとし、

$$p_i(\theta) = h_i \exp(\theta T_i - A(\theta)) \tag{1.85}$$

であるとする。ただし、$A(\theta)$ は、θ のなめらかな関数である。このとき、簡単な計算により、

$$\sum_{i=1} T_i p_i(\theta) = A'(\theta) \quad \sum_{i=1} T_i^2 p_i(\theta) = A''(\theta) + A'(\theta)^2 \tag{1.86}$$

である。したがって

$$J_{p(\theta)} = A''(\theta) \tag{1.87}$$

である。
熱力学の文脈では、$p_i(\theta)$ をギブス分布、T_i をエネルギーとして、$-\theta$ を逆温度として解釈することができる。そして、$\theta^{-1}A(\theta)$ は、自由エネルギーである。すなわち、

$$p_i(\theta) = \frac{1}{Z(\theta)} \exp(-\theta E_i) \tag{1.88}$$

である。
また、(1.89) は、T_i が、$A'(\theta)$ の不偏推定量であることを示している。それに対応するフィッシャー情報量は、$J_{p(\theta)} = A''(\theta)^{-1}$ である。というのも、$\frac{\mathrm{d}}{\mathrm{d}\theta'} = (A''(\theta))^{-1} \frac{\mathrm{d}}{\mathrm{d}\theta}$ であるからである。

また、情報幾何の文脈では、フィッシャー情報量は、確率分布空間の計量として考えられる。以下では、monotone 計量という概念を導入する。

Def:Monotone 計量

$G_p : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ が、フルランクの $p \in \mathcal{P}_d$ に対して、以下を満たすとき、G_p は Monotone 計量であるという。

- G_p は双線形である。
- $G_p(a,a) \geq 0$ であり、$G_p(a,a) = 0 \Leftrightarrow a = 0$ である。
- $p \mapsto G_p(a,a)$ は任意の a に対して連続である。
- 任意の T, a, p に対して、$G_p(a,a) \geq G_{Tp}(Ta, Ta)$ が成り立つ。

このとき、特にフィッシャー情報計量は、

$$G_p^F(a,b) = \sum_{i=1}^{d} \frac{a_i b_i}{p_i} \tag{1.89}$$

で定義される。ただし、$a = (a_1, ..., a_d), b = (b_1, ..., b_d)$ である。
この計量は、フィッシャー情報量行列と、

$$J_{p(\theta),kl} = G_{p(\theta)}^F(\partial_k p(\theta), \partial_l p(\theta)) \tag{1.90}$$

との関係がある。

Prop: フィッシャー情報量計量の単調性

任意の確率遷移行列 T に対して、

$$G_p^F(a,a) \geq G_{Tp}^F(Ta,Ta) \tag{1.91}$$

が成り立つ。

Prf

逆に、任意の monotone 計量は、フィッシャー情報量計量を用いて、以下のように表現できる。

Thm:Chentsov の定理

任意の monotone 計量 G_p は、フィッシャー情報量計量を用いて、

$$G_p(a,b) = kG_p^F(a,b) + k'(\sum_{i=1}^{d} a_i)(\sum_{i=1}^{d} b_i) \tag{1.92}$$

と表現できる。ただし、$k, k' \geq 0$ である。

Prf

略