Affinity を用いた漸近理論

林 正人 [1]　　京都大学 理学研究科 数学教室

概 要

本稿では,Akahira and Takeuchi [1] によって導入された Affinity を用いることによって非正則モデルにも適用できる漸近理論を展開する. その際, 大偏差の評価を用いることにする.

1　はじめに

$\mathcal{S} := \{p_\theta | \theta \in \Theta \subset \mathbb{R}\}$ を確率空間 Ω 上の 1-パラメータ確率分布族とする. 今, ここで $X_1, \ldots, X_n$ を互いに独立に確率分布 p_θ に従う n 個の確率変数とする. このとき n 個のデータに対する未知パラメータ θ の推定量は Ω^n から $\mathbb{R}$ への可測写像 T_n で与えられる. 以下では可測写像の列 $T = \{T_n\}_{n=1}^{\infty}$ を推定量と呼ぶことにする. さらに Ω 上の 2 つの確率分布 p, q の Affinity $I(p\|q)$ を以下で定義する [1].

$$I(p\|q) := -8\log\left(\int_\Omega \sqrt{\frac{\partial p}{\partial r}}\sqrt{\frac{\partial q}{\partial r}} r(\,d\omega)\right). \tag{1}$$

r は p, q 双方が絶対連続となる測度とし, $\frac{\partial p}{\partial r}, \left(\frac{\partial q}{\partial r}\right)$ をそれぞれ $p(q)$ の r に関する Radon-Nikodym の微分とする. 定義式 (1) は r の取り方に依存しないことに注意せよ. 本稿の主な目的は推定量 T の性能を漸近論の立場から評価することにある. 評価の基準としては平均 2 乗誤差及び大偏差型がよく用いられるが, 本稿では大偏差型で評価することにする. 大偏差型の評価では通常相対エントロピーを用いる手法が知られている. しかし, 非正則な場合は相対エントロピーが無限大となり, この方法による大偏差型の評価は破綻している. このような事情から, 大偏差型の評価においても, 非正則なモデルでは超有効性に類する現象が起きることが知られている [2, 3].

以下では Affinity を用いることによって大偏差型の下限を与える. 正則モデルでの大偏差型評価では弱一致性のみで下限が与えれたが, 非正則の場合の Affinity による下限は弱一致性よりも強い一致性条件の下で与えられる. その条件は本稿で導入されるもので, 局所強一致性条件と呼ばれる. さらに, そのような一致性条件が外れると先に述べたような超有効性に類する現象が起ることが確かめられる.

正則なモデルでは, 良く知られているように, 推定の大偏差型評価と検定の双方で相対エントロピーが用いられることを通して, 推定を検定の無限近似と見なすことが可能となる. しかし, 非正則なモデルでは検定の精度を評価する相対エントロピーが無限大になる一方で, 推定論的に意味のある Affinity による下限は有限となる. このようなギャップが生ずるために, 非正則モデルでは, 推定を検定の無限近似と見なすことは出来ない. 従って, 非正則モデルでは検定と推定では根本的に異なった取り扱いを必要とする.

同様な現象は量子状態の統計的推測を考えた時にも起ることが知られている. 量子系で検定を考えたときに, 検定の精度を表す量は (問題設定にもよるが) 相対エントロピーの量子的対応物である量子相対エントロピーで与えられる. これについては文献 [4, 5] が詳しい. 一方で,(1-パラメータ族の) 状態推定を考えたときには, Affinity の量子的対応物である量子 Affinity で大偏差型の下限が与えられる [6]. モデルに量子論的非可換性があるときには量子相対エントロピーの極限と量子 Affinity の極限は一致せず, 一般に, 前者方が後者よりも大きくなる. このような現象は前者が発散しない場合であっても起る. 基本的に量子論的な場合でも以下にのべる非正則モデルの場合とほぼ同様にして上記のことを示すことが出来る.

[1] e-mail address: masahito@kusm.kyoto-u.ac.jp

2　主定理とその証明

推定量 $T, \epsilon > 0$ に対して $\beta(T,\theta,\epsilon)$ を以下で定義する.

$$\beta(T,\theta,\epsilon) := -\lim_{n\to\infty}\frac{1}{n}\log p_\theta^n\left(|T_n-\theta|\geq\epsilon\right).$$

以下では大偏差型評価とはこの量の評価を意味する. 非正則モデルではこの量の評価は困難であるため, $\epsilon\to 0$ での振る舞いを調べることにする. $\beta^+(T,\theta,a),\beta^-(T,\theta,a)$ を以下で定義する.

$$\beta^+(T,\theta,a) := -\lim_{n\to\infty}\frac{1}{n}\log p_\theta^n\left(T_n\geq a\right)$$
$$\beta^-(T,\theta,a) := -\lim_{n\to\infty}\frac{1}{n}\log p_\theta^n\left(T_n\leq a\right).$$

さらに $\beta^+(T,\theta,a),\beta^-(T,\theta,a)$ を下半連続になるように再定義したものを $\overline{\beta}^+(T,\theta,a),\overline{\beta}^-(T,\theta,a)$ と記すことにする. すなわち以下のように定義する.

$$\overline{\beta}^+(T,\theta,a) := \lim_{\epsilon\to+0}\beta^+(T,\theta,a+\epsilon)$$
$$\overline{\beta}^-(T,\theta,a) := \lim_{\epsilon\to-0}\beta^-(T,\theta,a+\epsilon).$$

$\beta^+(T,\theta,a)$ は a について単調増加であり, $\beta^-(T,\theta,a)$ は a について単調減少であることに注意せよ.

定理 1 $\theta' > \theta$ に対して以下の条件が成立するとき,

$$\beta^+(T,\theta,\theta) = \beta^-(T,\theta',\theta') = 0,$$

以下の不等式が成立する.

$$\inf_{\{\theta''|\theta'\geq\theta''\geq\theta\}}\frac{\overline{\beta}^+(T,\theta,\theta'')+\overline{\beta}^-(T,\theta',\theta'')}{2}\leq\frac{1}{8}I(p_\theta\|p_{\theta'}). \tag{2}$$

証明　m を任意の整数とし, 区間 $[\theta,\theta']$ を m 等分する.$i=0,\ldots,m$ に対して, $\theta_i := \theta+\frac{\theta'-\theta}{m}i$ と定義する. さらに任意の $\epsilon>0$ に対して以下の条件を満たすような十分大きな整数 N をとる.

$$\frac{1}{n}p_\theta^n\left(T_n\geq\theta_i\right)\leq\beta^+(T,\theta,\theta_i)+\epsilon$$
$$\frac{1}{n}p_\theta^n\left(T_n\leq\theta_i\right)\leq\beta^-(T,\theta,\theta_i)+\epsilon,\ \text{for }\forall n\geq N, 0\leq\forall i\leq m.$$

affinity の単調性より

$$\begin{aligned}
&-\frac{n}{8}I(p_\theta\|p_{\theta'})\\
&\leq\log\left(\sqrt{p_\theta^n\left(T_n\leq\theta\right)}\sqrt{p_{\theta'}^n\left(T_n\leq\theta\right)}+\sum_{i=1}^m\sqrt{p_\theta^n\left(\theta_{i-1}<T_n\leq\theta_i\right)}\sqrt{p_{\theta'}^n\left(\theta_{i-1}<T_n\leq\theta_i\right)}+\sqrt{p_\theta^n\left(\theta'<T_n\right)}\sqrt{p_{\theta'}^n\left(\theta'<T_n\right)}\right)\\
&\leq\log\left(\sqrt{p_{\theta'}^n\left(T_n\leq\theta\right)}+\sum_{i=1}^m\sqrt{p_\theta^n\left(\theta_{i-1}<T_n\right)}\sqrt{p_{\theta'}^n\left(T_n\leq\theta_i\right)}+\sqrt{p_\theta^n\left(\theta'<T_n\right)}\right)\\
&\leq\log\left(\exp\left(-n\left(\frac{\beta^+(T,\theta,\theta_{i-1})+\beta^-(T,\theta',\theta_i)}{2}+\epsilon\right)\right)+\exp\left(-n\left(\frac{\beta^+(T,\theta,\theta')}{2}+\epsilon\right)\right)+\exp\left(-n\left(\frac{\beta^-(T,\theta',\theta)}{2}+\epsilon\right)\right)\right)\\
&\leq\log\left((m+2)\exp\left(-n\left(\min_{1\leq i\leq m}\frac{\beta^+(T,\theta,\theta_{i-1})+\beta^-(T,\theta',\theta_i)}{2}+\epsilon\right)\right)\right)\\
&=\left(\log(m+2)-n\left(\min_{1\leq i\leq m}\frac{\beta^+(T,\theta,\theta_{i-1})+\beta^-(T,\theta',\theta_i)}{2}+\epsilon\right)\right).
\end{aligned}$$

両辺を n で割った後で $n \to \infty$ を考えると,

$$I(p_\theta \| p_{\theta'}) \geq 8 \left(\min_{1 \leq i \leq m} \frac{\beta^+(T, \theta, \theta_{i-1}) + \beta^-(T, \theta', \theta_i)}{2} + \epsilon \right).$$

$\epsilon > 0$ は任意であるから,

$$I(p_\theta \| p_{\theta'}) \geq 8 \min_{1 \leq i \leq m} \frac{\beta^+(T, \theta, \theta_{i-1}) + \beta^-(T, \theta', \theta_i)}{2}.$$

$m \to \infty$ を考えると, (2) を得る. □

定義 **2** $\theta_0 \in \Theta$ の適当な近傍 U_{δ, θ_0} で

$$\lim_{\epsilon \to 0} \frac{\beta(T, \theta, \epsilon)}{I(p_\theta \| p_{\theta+\epsilon})}. \tag{3}$$

が一様収束するとき推定量 T を θ_0 で局所強一致推定量と呼ぶことにする. さらに *(3)* の量を $\alpha(T, \theta)$ と呼ぶことにする. そして $\overline{\alpha}(T, \theta)$ を以下のように定義する.

$$\overline{\alpha}(T, \theta) := \limsup_{\epsilon \to \pm 0} \alpha(T, \theta + \epsilon).$$

定理 **3** 推定量 T が $\theta_0 \in \Theta$ で局所強一致であるとき以下の不等式が成立する.

$$\overline{\alpha}(T, \theta_0) \leq \begin{cases} \frac{1}{4}, & \text{if } \lim_{\epsilon \to 0} \frac{I(p_\theta \| p_{\theta+\epsilon})}{\epsilon} > 0 : \\ \frac{1}{8} \lim_{\epsilon \to \infty} \frac{I(p_\theta \| p_{\theta+2\epsilon})}{I(p_\theta \| p_{\theta+\epsilon})}, & \text{if } \lim_{\epsilon \to 0} \frac{I(p_\theta \| p_{\theta+\epsilon})}{\epsilon} = 0. \end{cases} \tag{4}$$

さらに

$$I(p_\theta \| p_{\theta_\epsilon}) \cong J_{\theta_0} \epsilon^\gamma, \quad 1 < \gamma \leq 2$$

となるときには (4) の右辺は $\frac{2^\gamma}{8}$ となる.

以下の定理は超有効性がある場合に有効である.

定理 **4** 推定量 T が $\theta_0 \in \Theta$ で局所強一致であるとき以下の不等式が成立する.

$$B(\alpha(T, \theta_0), \overline{\alpha}(T, \theta_0), I(p_\theta \| p_{\theta+\cdot})) \leq \frac{1}{8}$$

ここで $\alpha_1, \alpha_2 > 0$ と $f(\epsilon) \to 0$ as $\epsilon \to 0$ となる関数 $f(\epsilon) > 0$ に対して $B(\alpha_1, \alpha_2, f)$ を以下で定義した.

$$B(\alpha_1, \alpha_2, f) := \lim_{\epsilon \to +0} \frac{1}{f(\epsilon)} \min_{\{\gamma | 0 \leq \gamma \leq \epsilon\}} \frac{\alpha_1 f(\gamma) + \alpha_2 f(\epsilon - \gamma)}{2}. \tag{5}$$

定理 4 は (4) の下のときのみ有効である. (4) の上の場合では (4) より良い評価を与えない.

系 **5**

$$\alpha(T, \theta) = \infty$$

であり, なおかつ T が θ で局所強一致であるときには, 以下の式が成立する.

$$\overline{\alpha}(T, \theta_0) \leq \frac{1}{4}. \tag{6}$$

系 **6** 推定量 T が $\theta_0 \in \Theta$ で局所強一致であり,

$$I(p_\theta \| p_{\theta_\epsilon}) \cong J_{\theta_0} \epsilon^\gamma, \quad 1 < \gamma \leq 2$$

となるときには (5) は以下のように書き直される.

$$\frac{\overline{\alpha}(T,\theta_0)\alpha(T,\theta_0)^{\frac{\gamma}{\gamma-1}} + \alpha(T,\theta_0)\overline{\alpha}(T,\theta_0)^{\frac{\gamma}{\gamma-1}}}{\left(\alpha(T,\theta_0)^{\frac{1}{\gamma-1}} + \overline{\alpha}(T,\theta_0)^{\frac{1}{\gamma-1}}\right)^\gamma} \leq \frac{1}{4}.$$

特に $\gamma = 2$ のときには以下のように書き換えられる.

$$\frac{\overline{\alpha}(T,\theta_0)\alpha(T,\theta_0)^2 + \alpha(T,\theta_0)\overline{\alpha}(T,\theta_0)^2}{\left(\alpha(T,\theta_0) + \overline{\alpha}(T,\theta_0)\right)^2} \leq \frac{1}{4}.$$

3 例への適用

以下では前節の内容を具体例に適用して検討することにする.

例 **1** 一様分布を考える. すなわち密度関数 $f_\theta(x)$ が以下で定義される確率分布族 $\{p_\theta | \theta \in \mathbb{R}\}$ を考える.

$$f_\theta(x) := \begin{cases} 1, & |\theta - x| < \frac{1}{2} : \\ 0, & |\theta - x| \geq \frac{1}{2}. \end{cases}$$

そして $X_1, \ldots, X_n$ を確率分布 p_θ に従う確率変数とする. このとき

$$I(p_\theta \| p_{\theta+\epsilon}) = -8\log(1-\epsilon) \cong 8\epsilon.$$

なので, T が局所強一致推定量ならば, 定理 3 の (4) が成立する. このとき推定量 $M = \{M_n\}$ を以下のように定義する.

$$M_n := \frac{\max_{1\leq k\leq n} X_k + \min_{1\leq l\leq n} X_l}{2}.$$

このとき

$$p_\theta^n \left(|M_n - \theta| \geq \epsilon\right) = (1-2\epsilon)^n$$

となる. 従って

$$\alpha(M,\theta) = \lim_{\epsilon\to 0} \frac{-\log(1-2\epsilon)}{-8\log(1-\epsilon)} = \frac{1}{4}.$$

従って推定量 M は (4) の下限を達成する. 次に推定量 M を変形して M' を以下のように構成する.

$$M_n' := \begin{cases} 0 & \text{全ての } k \text{ について } X_k \in [-1/2, 1/2] \\ M_n & \text{上記以外 .} \end{cases}$$

このとき, 任意の $\epsilon > 0$ について

$$p_\theta^n \left(|M_n' - \theta| \geq \epsilon\right) \begin{cases} = 0, & \text{if } \theta = 0 \\ \leq (1-2\epsilon)^n + (1-|\theta|)^n, & \text{if } 0 < |\theta| < \frac{1}{2} \text{ and } |\theta| \leq \epsilon \\ \geq (1-|\theta|)^n, & \text{if } 0 < |\theta| < \frac{1}{2} \text{ and } |\theta| \leq \epsilon \\ \leq (1-2\epsilon)^n, & \text{if } 0 < |\theta| < \frac{1}{2} \text{ and } |\theta| \geq \epsilon \\ \geq (1-2\epsilon)^n - (1-|\theta|)^n, & \text{if } 0 < |\theta| < \frac{1}{2} \text{ and } |\theta| > 2\epsilon \\ = (1-2\epsilon)^n, & \text{if } |\theta| \geq \frac{1}{2}. \end{cases}$$

任意の $\epsilon > 0$ について

$$\beta(M', \theta, \epsilon) \begin{cases} = \infty, & \text{if } \theta = 0 \\ = -\log(1 - |\theta|), & \text{if } 0 < |\theta| < \frac{1}{2} \text{ and } |\theta| \leq \epsilon \\ \geq -\log(1 - 2\epsilon), & \text{if } 0 < |\theta| < \frac{1}{2} \text{ and } |\theta| \geq \epsilon \\ = -\log(1 - 2\epsilon), & \text{if } 0 < |\theta| < \frac{1}{2} \text{ and } |\theta| > 2\epsilon \\ = -\log(1 - 2\epsilon), & \text{if } |\theta| \geq \frac{1}{2}. \end{cases}$$

この場合は (3) で一様収束していないので定理は適用できないが以下のように計算できる.

$$\alpha(M', \theta) = \begin{cases} \infty & \text{if } \theta = 0 \\ \frac{1}{4} & \text{if } \theta \neq 0. \end{cases}$$

例 **2** $f(x)$ を $[-1, 1]$ でのみ値を持ち以下の条件を満たす密度関数とする.

$$f(x-1) \cong A(x-1)^{a-1} \text{ as } x \to 1$$
$$f(x+1) \cong A(x+1)^{b-1} \text{ as } x \to -1$$

とし $a, b > 2$ とする. そして $f_\theta(x) := f(x - \theta)$ で定義される確率分布 p_θ の族を考える. このとき Akahira and Takeuchi [1] より

$$J_\theta := \lim_{\epsilon \to 0} \frac{I(p_\theta \| p_{\theta+\epsilon})}{\epsilon^2} < \infty.$$

を得る. そして $X_1, \ldots, X_n$ を確率分布 p_θ に従う確率変数とする. 最尤推定量を $\hat{\theta}_{ML} := \{\hat{\theta}_{ML,n}\}$ と書くと以下の不等式を得る.

$$\alpha(\hat{\theta}_{ML}, \theta) = \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{\beta(\hat{\theta}_{ML}, \theta, \epsilon)}{J_\theta \epsilon^2} = \frac{1}{2}. \tag{7}$$

このとき (4) の右辺は $1/2$ なので最尤推定量は下限を達成している. また, 最尤推定量は位置に関して共変であるから, (7) は θ に関して一様収束している. さらに, 次に最尤推定量 $\hat{\theta}_{ML}$ を変形して $\hat{\theta}'_{ML} := \{\hat{\theta}'_{ML,n}\}$ を以下のように構成する.

$$\hat{\theta}'_{ML,n} := \begin{cases} 0 & \text{全ての } k \text{ について } X_k \in [-1, 1] \\ \hat{\theta}_{ML,n} & \text{上記以外 .} \end{cases}$$

このとき先ほどと同じような計算をすると, 以下の式を得る.

$$\alpha(\hat{\theta}'_{ML}, \theta) = \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{\beta(\hat{\theta}'_{ML}, \theta, \epsilon)}{J_\theta \epsilon^2} = \begin{cases} \infty, & \text{if } \theta = 0 \\ \frac{1}{2}, & \text{if } \theta \neq 0. \end{cases} \tag{8}$$

局所強一致であれば系 5 の (6) が成立することを考えると, $\hat{\theta}'_{ML}$ は局所強一致でないことが分る. すなわち (8) は 0 で局所一様収束しないことが分る. 一見 $\alpha(T, \theta)$ の値だけを見ると $\hat{\theta}_{ML}$ よりも $\hat{\theta}'_{ML}$ の方が優れているように思われるが $\hat{\theta}'_{ML}$ の方は上に見たように $\theta = 0$ の近傍で (8) が一様収束しないので必ずしも $\hat{\theta}'_{ML}$ の方が優れているとは言えない. 今後 (8) 型の評価を行うときは一様収束しているかどうかについても吟味する必要がある.

4 結論

本稿では今回新たに導入された局所強一致性という概念と Affinity を用いることによって, [2, 3] で指摘された超有効性を排除することに成功した. 例 1 及び例 2 で構成した推定量 M' 及び $\hat{\theta}'_{ML}$ は推定量と言うよりもむしろ $\theta = 0$ であるか否かを調べる検定量として見るべきである. $\theta = 0$ であるか否かを調べる検定量の rate は相対エントロピーで評価されることが知られている. 正則なモデルでは相対エントロピーの極限として得られる $\alpha(T, \theta)$ の上限と Affinity の極限から得られる $\alpha(T, \theta)$ の上限が一致するために, 検定量と推定量の差は $\epsilon \to 0$ で無視できた. しかし, 非正則なモデルでは上記 2 つの情報量の極限にギャップがあるために, 検定と推定には無視できない差が生じる.

今後非正則なモデルを扱うとき, 推定量に課すべき条件としては, 弱一致性だけではなく今回導入した局所強一致性を課して考える必要が有る.

参考文献

[1] M. Akahira and K. Takeuchi, *Non-Regular Statistical Estimation* LNS 107, Springer, (1995).

[2] I. A. Ibragimov and R. Z. Has'minskii, *Statistical Estimation*, Springer, (1981).

[3] 赤平昌文, "統計的推測と大偏差型確率," 数理科学 2 月号, (1995).

[4] F. Hiai and D. Petz, Commun. Math. Phys. **143**, 99-114 (1991).

[5] M. Hayashi, "Asymptotic Attainment for Quantum Relative Entropy," LANL e-print quant-ph/9704044 (1997).

[6] 林正人, "量子 1-パラメータ族の推定における大偏差型評価" 「情報幾何の新しい展開」(統計数理研究所共同研究会) 講演資料, (1999).