1999 年情報論的学習理論ワークショップ
1999 Workshop on Information-Based Induction Sciences (IBIS'99)
Izu, Japan, August 26-27, 1999.

適応的に測定を選択できる場合での大偏差型評価

Large deviation evaluation in adaptive measurement

林 正人*
Masahito Hayashi

Abstract: This manuscript is based on the joint research with Dr. Matsumoto. We deal with a large deviation theory of adaptive experimental design. First, we propose a framework for the general discussion of asymptotic estimation and detection theory of adaptive experimental design. Then, two analogues of Bahadur inequality, and their attainability are proved under some regularity conditions.

キーワード: 量子推定, 大偏差型評価, 適応的測定選択, 推定と検定のギャップ, 実験計画, 能動学習

1 まえがき

本稿では, データの取り方 (測定) を適応的に選択できる場合に, パラメータ推定の大偏差型評価を試みる. データの取り方を適応的に選択できるという条件で未知パラメータを推定するという問題設定は極めて自然である. そのため, 数学的にはこれと同値である, 異なった名称で呼ばれている分野が複数存在する. 例えば, 実験計画や本研究集会のテーマと関連の深い能動学習などが挙げられる. 実験計画では情報行列を最適化するという戦略が取られる. 能動的学習では, システムの入力と出力との間に成立する条件付き確率を推定する問題を扱う. この場合, 適切に入力分布を選択することが問題となる.

これと同種な問題は量子力学に従う系で未知状態を (量子相関を用いない条件の下で) 推定するという問題を考えたときにも表れる. この場合についても, 単にデータから未知状態を推定する推定量を構成するのではなく, データの抽出過程である, 量子測定の選択の仕方が問題となる. このような理論は量子推定理論と呼ばれ 1960 年代後半に Helstrom により始められた [1]. 量子推定理論では未知パラメータを推定する際の誤差評価の方法として, 平均 2 乗型だけでなく大偏差型の評価も試みられている [2].

本稿では, 量子推定理論の表記にならい未知状態がパラメータ $\theta = (\theta^1, \ldots, \theta^d) \in \mathbb{R}^d$ でパラメトライズされた状態族 $\mathcal{S} := \{\rho_\theta | \theta \in \Theta \subset \mathbb{R}^d\}$ に含まれていることが既知の上で, 未知パラメータ θ を推定する問題を扱う. このような視点からの, 平均 2 乗誤差に注目した漸近理論は幾つかなされている [3, 4]. 本稿では, 大偏差型の評価で漸近理論を展開する. なぜなら大偏差を用いることによって, この問題設定の下での推定と検定の本質的相違が初めて明らかになるからである. なお, この相違を考える上で未知パラメータが複数の場合について考えることは本質的に不可欠である. なぜなら,1 次元の場合にはモデルがある種の正則を満たしている限りそのような相違が表れないからである.

ここでもし, 状態族 $\mathcal{S}$ とともに, 測定 M もあらかじめ指定されていたとすると, 確率分布 $P_\theta(\, d\omega) = P(\, d\omega|\rho_\theta, M)$ に従うデータ $\omega \in \Omega$ を見て θ を推定するという単なる確率分布のパラメータ推定問題になる. 本稿で扱う問題ではあらかじめ指定されるのは条件付き確率 $P(\, d\omega|\rho_\theta, M)$ (以下, 場合によっては $P(\rho_\theta, M)$ と記す) のみであり, データを取る操作である測定 M はこれまでに得られたデータをもとに適応的に最適なものを選択することが可能となる.

2 問題の定式化

2.1 全般的な定義及び例

以下, 数式を用いてもう少し細かい議論を行う. 測定 M の集合を $\mathcal{M}$ で表すことにする. そして測定値の集合を Ω で表すことにするが, このとき, 異なる測定 $M_1, M_2 \in \mathcal{M}$ のサポートは disjoint になっているとす

*京都大学大学院 理学研究科 数学教室, 〒 606-8502 京都市左京区北白川追分町 tel. 075-753-3695, e-mail masahito@kusm.kyoto-u.ac.jp,
Department of Mathematics, Faculty of Science, Kyoto University, Kyoto, 606-8502, Japan

る. すなわち,

$$\left(\bigcup_{\theta\in\Theta}\mathrm{Supp}P(\rho_\theta,M_1)\right)\bigcap\left(\bigcup_{\theta\in\Theta}\mathrm{Supp}P(\rho_\theta,M_2)\right)=\emptyset. \tag{1}$$

とする. (1) が満たされていない場合は改めて $\Omega\times\mathcal{M}$ を Ω と取り直すと満たされる. この問題設定の下で我々が可能な戦略は単に単独の測定 $M\in\mathcal{M}$ を用いるだけでなく, 確率的に測定 $M\in\mathcal{M}$ を用いることも含まれる. 従って, 我々が単一のサンプルに対して可能な選択は $\mathcal{M}$ 上の確率分布 μ で記述される. 以下 $\mathcal{M}$ 上の確率分布全体の集合を $P(\mathcal{M})$ で表し, 単に測定と呼ぶときは $P(\mathcal{M})$ の要素を指し, $\mathcal{M}$ の要素を指すときは基本測定と呼ぶことにする.

例 **1** 粒子が二次元空間中に拘束されて存在し, その座標は $\theta=(\theta^1,\theta^2)$ を中心として共分散行列 $\sigma^2 I_2$ のガウス分布にしたがって揺らいでいるとする. 今, 同じ二次元空間中に拘束されたカメラの映像を通じて $\theta=(\theta^1,\theta^2)$ を推定もしくは検定する問題を考える. このときカメラの映像は *1* 次元に制限されているとする. 状態 ρ_θ は, 粒子の座標が従う確率分布に対応する. ある方向 $x\in S^1$ (S^1 は単位ベクトルの集合すなわち単位円) から粒子を撮影するという測定を M_x と書くと, 得られるデータは粒子の位置を x 方向に射影で得られる. すなわち $\mathrm{Supp}(P(\rho_\theta,M_x))=\{tx|t\in\mathbb{R}\}\subset\mathbb{R}^2$ となる. そしてその分布は $P(\rho_\theta,M_x)=P_{\theta,x}\circ P_x^{-1}$ で与えられる. なお P_x は $\mathbb{R}^2$ から $\mathbb{R}$ への射影で $y\mapsto(x,y)$ で与えられる. また $\mathbb{R}$ 上の確率分布 $P_{\theta,x}$ は以下で与えられる.

$$P_{\theta,x}(B):=\frac{1}{\sqrt{2\pi\sigma^2}}\int_B\exp\left(-\frac{(y-(\theta,x))^2}{2\sigma^2}\right)\,dy.$$

2.2 推定量及び検定量の定義

適応的な測定による未知パラメータ θ の推定 (もしくは検定) は以下のように定式化される. n 個の未知状態 ρ_θ が与えられとの仮定の下では, 未知パラメータの推定 (または検定) には $k-1$ 番目までの測定値 $\omega_1,\ldots,\omega_{k-1}$ から, k 番目の測定 ($\mathcal{M}$ 上の確率分布) $\mu_k(\omega_1,\ldots,\omega_{k-1})$ を与える関数の系列 $\mathcal{E}_n:=\{\mu_k\}_{k=1}^n$ と n 個の測定値 $\omega_1,\ldots,\omega_n$ から推定値 $T_n(\omega_1,\ldots,\omega_n)$ を与える関数 T_n が必要である. 以下表記の簡略化のために $(\omega_1,\ldots,\omega_k)\in\Omega^{k-1}$ を $\boldsymbol{\omega}_k$ で記すことにする. また同一の測定 μ を n 回行う測定の系列を $\mu\times n$ で表すことにする. したがってこれらの組 $(\mathcal{E}_n,T_n)$ が最適化の対象となり, これを n 個のサンプルに対する推定量と呼ぶ. そして, 大偏差型評価ではその推定量の系列 $\mathcal{E}:=\{(\mathcal{E}_n,T_n)\}_{n=1}^\infty$ の裾確率の rate の極限を評価することになる.

また本稿では帰無仮説が $\theta\in\Theta_1$ で対立仮説が $\theta=\theta_0$ である場合の検定を考える. この場合, 検定量は先に与えた k 番目の測定を与える関数の系列 $\mathcal{E}_n:=\{\mu_k\}_{k=1}^n$ と n 個の測定値から推定値から帰無仮説が正しいか否かを判断する 2 値関数 T_n の組 $(\mathcal{E}_n,T_n)$ で記述される.

3 相対エントロピー,Fisher情報行列

真のパラメータが θ であるときに, 先に与えた k 番目の測定を与える測定の系列 $\mathcal{E}_n:=\{\mu_k\}_{k=1}^n$ にしたがって測定を行ってデータを取ったとき得られる確率分布 $P(\rho_\theta,\mathcal{E}_n)$ は以下で与えられる.

$$P(\,d\boldsymbol{\omega}_n|\rho_\theta,\mathcal{E}_n):=$$
$$P(\,d\omega_1|\rho_\theta,\mu_1)P\,(\,d\omega_2|\rho_\theta,\mu_2(\omega_1))\cdots P\,(\,d\omega_n|\rho_\theta,\mu_n(\boldsymbol{\omega}_{n-1}))\,.$$

なお, ここで $\mu\in P(\mathcal{M})$ に対して Ω 上の確率分布 $P(\rho_\theta,\mu)$ を以下で定義した.

$$P(\,d\omega|\rho_\theta,\mu):=\int_{\mathcal{M}}P(\,d\omega|\rho_\theta,M)\mu(\,dM).$$

以下, 相対エントロピー $D(p\|q):=\int\log\frac{p}{q}p(\,d\omega)$ に関するいくつかの式を導入する. このとき条件(1) を用いると, 以下の式を示すことができる. なお以下では表記の簡略化のため $D(P(\rho_\theta,\mathcal{E}_n)\|P(\rho_{\theta'},\mathcal{E}_n))$,$D(P(\rho_\theta,\mu)\|P(\rho_{\theta'},\mu))$, $D(P(\rho_\theta,M)\|P(\rho_{\theta'},M))$ をそれぞれ, $D^{\mathcal{E}_n}(\theta\|\theta')$, $D^{\mu}(\theta\|\theta')$, $D^{M}(\theta\|\theta')$ と書くことにする. (2) ついてはいきなり一般的に証明しようとすると手のつけようがないが, $n=2$ の場合について考えて, 後で一般化を考えるとよい.

$$D^{\mathcal{E}_n}(\,\theta\,\|\theta')=D^{\mu_1}(\theta\|\theta')$$
$$+\sum_{k=2}^n\int_{\Omega^{k-1}}D^{\mu_k(\boldsymbol{\omega}_{k-1})}(\,\theta\|\theta')\,P(\,d\boldsymbol{\omega}_{k-1}|\rho_\theta,\mathcal{E}_{k-1}) \tag{2}$$
$$D^{\mu}(\,\theta\|\theta')=\int_{\mathcal{M}}D^{M}(\theta\|\theta')\mu(\,dM). \tag{3}$$

ここで $\mathcal{M}$ 上の確率分布

$$\mu_n(\theta):=\frac{1}{n}\left(\sum_{k=2}^n P(\theta,\mathcal{E}_{k-1})\circ\mu_k^{-1}+\mu_1\right) \tag{4}$$

を用いると(2) は以下の式に書き換えられる.

$$D^{\mathcal{E}_n}(\rho_\theta\|\theta')=nD^{\mu_n(\theta)}(\theta\|\theta').$$

なお(4) の右辺の $\sum$ 内は $P\mathcal{M}$ 上での確率分布を表しているいるが,(4) ではそれを $\mathcal{M}$ 上の確率分布と見なして $\mu_n(\theta)$ を定義している.

さらに確率分布族 $\{P(\rho_\theta, M)|\theta \in \Theta\}, \{P(\rho_\theta, \mu)|\theta \in \Theta\}$ の $\theta \in \Theta$ での Fisher 情報行列をそれぞれ $J^M(\theta), J^\mu(\theta)$ と書くことにする. すると(3) の極限を考えることにより以下の式を得る.

$$J^\mu(\theta) = \int_{\mathcal{M}} J^M(\theta)\mu(\, dM).$$

以下の節でこれらの式を用いて推定量及び検定量の漸近的評価を試みる.

4 検定量の漸近的評価

以下では検定量の系列 $\mathcal{D} := \{(\mathcal{E}_n, T_n)\}_{n=1}^\infty$ について以下の評価を考える. まず検定量の系列 $\mathcal{D}$ に対して第 1 種の誤り ($\theta \neq \theta_0$ であるのに $\theta = \theta_0$ と判断する) 確率が漸近的に 0 に近づくとの仮定をおく. すなわち以下を仮定する.

(I) $\lim_{n\to\infty} P(T_n = \theta_0|\rho_\theta, \mathcal{E}_n) = 0, \quad \forall\theta \in \Theta_1.$

上記の条件を満たす検定量の系列 $\mathcal{D}$ に対して第 2 種の誤り確率 ($\theta = \theta_0$ であるのに $\theta \neq \theta_0$ と判断する) 確率の rate $R(\mathcal{D})$(以下に定義する) の評価を考える.

$$R(\mathcal{D}) := -\lim_{n\to\infty} \frac{1}{n} \log P(T_n \neq \theta_0|\rho_{\theta_0}, \mathcal{E}_n). \qquad (5)$$

定理 1 条件 **(I)** を満たす検定量の系列について以下の不等式が成立する.

$$R(\mathcal{D}) \leq \inf_{\theta\in\Theta_1} \sup_{M\in\mathcal{M}} D^M(\theta\|\theta_0). \qquad (6)$$

証明 p_n, q_n を以下のように定義する.

$$p_n := P(T_n \neq \theta_0|\rho_{\theta_0}, \mathcal{E}_n), \quad q_n(\theta) := P(T_n \neq \theta_0|\rho_\theta, \mathcal{E}_n).$$

条件 **(I)** より $\lim_{n\to\infty} q_n(\theta) = 1$ となる. h を 2 値エントロピーとすると相対エントロピーの単調性より以下の評価式を得る.

$$nD^{\mu_n(\theta)}(\theta\|\theta_0) = D^{\mathcal{E}_n}(\theta\|\theta_0)$$
$$\geq -q_n(\theta)\log p_n - (1-q_n(\theta))\log(1-p_n) - h(q_n(\theta)).$$

この評価式を変形し極限を取ると次の評価式が得られる.

$$\limsup_{n\to\infty} -\frac{1}{n}\log p_n \leq \limsup_{n\to\infty} D^{\mu_n(\theta)}(\theta\|\theta_0). \qquad (7)$$

従って $\theta \in \Theta_1$ 任意より(6) を得る. □

なお, Stein の補題の適応的測定版として, 第 1 種の誤り確率が ϵ 以下という条件の下でも第 2 種の誤り確率の rate が(6) の右辺で押さえられるのではないかとの予想が考えられる. 少なくともモデルがある正則性条件を満たす場合は示すことはできる. 例えば $\mathcal{M}$ が有限集合であればよい. 仮にその正則性条件を満たさない場合でも ((6) の右辺) $\times(1-\epsilon)$ で押さえられることは容易に示せる.

定理 2 モデルが正則性条件 (2.1)(2.2) を満たすとする. 数列 $\{a_n\}$ を $a_n \to \infty, a_n/n \to 0$ を満たす数列とする. このとき, 以下に定義する検定量の系列 $\{\mathcal{D}_{\theta_0}(a_n)\}$ は条件 **(I)** を満たし,(6) で等号を達成する.

$\{\mathcal{D}_{\theta_0}(a_n)\}$ の構成: a_n を自然数列としたときに検定量の系列 $\{\mathcal{D}_{\theta_0}(a_n)\}$ は以下のように構成される. n 回の測定のうち最初の a_n 回は正則性条件 (2.1) を満たす測定 $\mu_0 \in P(\mathcal{M})$ を a_n 回行い, 確率分布族 $\{P(\rho_\theta, \mu_0 \times a_n)|\theta \in \Theta_1\}$ に対して最尤推定 $\check{\theta}_n$ を得る. 次に正則性条件 (2.2) を満たす基本測定 M_θ を残り $n - a_n$ 回行い, 帰無仮説 $\theta = \check{\theta}_n$ 対立仮説 $\theta = \theta_0$ となる場合の仮説検定で Stein の補題の意味で最適な検定を行い, その結果を最終的な検定結果とする.

正則性条件 (2.1): 以下を満たす測定 $\mu_0 \in P(\mathcal{M})$ が存在する.

$$D^\mu(\theta\|\theta') > 0, \quad \forall\theta, \theta' \in \Theta_1.$$

正則性条件 (2.2): 任意の $\theta \in \Theta_1$ に対して以下を満たす基本測定 M_θ が存在する.

$$D^{M_\theta}(\theta\|\theta_0) \geq \inf_{\theta\in\Theta_1} \sup_{M\in\mathcal{M}} D^M(\theta\|\theta_0).$$

正則性条件 (2.3): 帰無仮説 $P(\rho_\theta, M_\theta)$ 対立仮説 $P(\rho_{\theta_0}, M_\theta)$ となる仮説検定で Stein の補題の意味で最適な検定を行ったときに, 第 2 種の誤り確率の rate が $D^{M_\theta)}(\theta\|\theta_0)$ に収束する速さが $\theta \in \Theta_1$ に関して一様であると仮定する.

証明 正則性条件 (2.3) より検定量の系列 $\{\mathcal{D}_{\theta_0}(a_n)\}$ が(6) の等号を達成していることは容易に示すことが出来る. 条件 **(I)** についても正則性条件 (2.1) を考慮に入れると満足していることが分る. □

5 推定量の漸近的評価

5.1 rate の上からの評価

次に推定量の列 $\mathcal{E} := \{(\mathcal{E}_n, T_n)\}_{n=1}^\infty$ に対する漸近的な評価を行う. 推定量の列 $\mathcal{E} := \{(\mathcal{E}_n, T_n)\}_{n=1}^\infty$ が以下の条件 **(II)** を満たすとき弱一致であるという.

(II) 任意の $\theta \in \Theta$ に対して以下の等式が成立する.

$$\lim_{n\to\infty} P(\, \|T_n - \theta\| > \epsilon|\, \rho_\theta, \mathcal{E}_n) = 0.$$

そして, 弱一致性を満たす推定量の系列 $\mathcal{E} := \{(\mathcal{E}_n, T_n)\}_{n=1}^\infty$ に対して以下の量が収束すると仮定して以下の量を考

える.

$$\beta(\mathcal{E},\theta,\epsilon) := -\lim_{n\to\infty}\frac{1}{n}\log P\left(\|T_n-\theta\| > \epsilon \middle| \rho_\theta, \mathcal{E}_n\right).$$

そして $\beta(\mathcal{E},\theta,\epsilon)$ の $\epsilon\to 0$ の極限として以下の量を導入する.

$$\alpha(\mathcal{E},\theta) := \lim_{\epsilon\to 0}\frac{\beta(\mathcal{E},\theta,\epsilon)}{\epsilon^2}. \tag{8}$$

定理 3 極限の順序交換等に関する以下の式が成立していることとする.

$$\begin{aligned}
&\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\sup_{M\in\mathcal{M}} D^M(\theta_0+\epsilon\xi\|\theta_0)\\
&= \sup_{M\in\mathcal{M}}\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}D^M(\theta_0+\epsilon\xi\|\theta_0)\\
&= \frac{1}{2}\sup_{M\in\mathcal{M}}\sum_{i,j} J^M_{i,j}(\theta_0)\xi^i\xi^j \quad \forall\xi\in\mathbb{R}^d, \forall M\in\mathcal{M}. \quad (9)
\end{aligned}$$

ξ は $\mathbb{R}^d$ の要素を表すので以下では $\xi\in\mathbb{R}^d$ は省略する. このとき, 弱一致性を満たす推定の系列 $\mathcal{E} := \{(\mathcal{E}_n, T_n)\}_{n=1}^\infty$ に対して以下の各点 $Bahadur$ 型の評価式が成立する:

$$\alpha(\mathcal{E},\theta_0) \le B^w(\theta_0). \tag{10}$$

ここで各点 $Bahadur$ 型の上限 $B^w(\theta)$ は以下で定義した.

$$B^w(\theta) := \inf\left\{ \sup_{M\in\mathcal{M}}\sum_{i,j}\frac{J^M_{i,j}(\theta)\xi^i\xi^j}{2} \middle| \|\xi\|=1 \right\}. \tag{11}$$

証明 $\Theta_1 := \{\theta\in\Theta | \|\theta-\theta_0\| > \epsilon\}$ とおいて, 定理 1 を適用すると,

$$\beta(\mathcal{E},\theta_0,\epsilon) \le \inf\left\{ \sup_{M\in\mathcal{M}} D^M(\theta\|\theta_0) \middle| \theta:\ \|\theta-\theta_0\| > \epsilon \right\}.$$

従って $\|\xi\| > 1$ に対して以下の式を得る.

$$\frac{\beta(\mathcal{E},\theta_0,\epsilon)}{\epsilon^2} \le \frac{1}{\epsilon^2}\sup_{M\in\mathcal{M}} D^M(\theta_0+\epsilon\xi\|\theta_0).$$

ここで条件(9) を用いて $\epsilon\to 0$ を考えると以下の式を得る.

$$\alpha(\theta_0) \le \frac{1}{2}\sup_{M\in\mathcal{M}}\sum_{i,j} J^M_{i,j}(\theta_0)\xi^i\xi^j.$$

また以下の式を示すことができる.

$$\begin{aligned}
&\inf\left\{ \sup_{M\in\mathcal{M}}\sum_{i,j} J^M_{i,j}(\theta_0)\xi^i\xi^j \middle| \|\xi\| > 1 \right\}\\
&= \inf\left\{ \sup_{M\in\mathcal{M}}\sum_{i,j} J^M_{i,j}(\theta_0)\xi^i\xi^j \middle| \|\xi\| = 1 \right\}.
\end{aligned}$$

したがって(10) を得る. □

ここでさらに, 推定量の系列 $\mathcal{E} := \{(\mathcal{E}_n, T_n)\}_{n=1}^\infty$ に対して弱一致性に加えて以下の条件を考える.

(III) (8) の右辺は θ_0 の近傍で局所一様収束する.

(IV) $\alpha(\mathcal{E},\theta_0) = \lim_{\theta\to\theta_0}\alpha(\mathcal{E},\theta)$.

以下の定理が成立する.

定理 4 任意の測定の列 μ_n に対して極限の順序交換等に関する以下の式が成立していることとする.

$$\begin{aligned}
&\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\liminf_{n\to\infty} D^{\mu_n}(\theta_0\|\theta_0+\epsilon\xi)\\
&= \liminf_{n\to\infty}\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}D^{\mu_n}(\theta_0\|\theta_0+\epsilon\xi)\\
&= \frac{1}{2}\liminf_{n\to\infty}\sum_{i,j} J^{\mu_n}_{i,j}(\theta_0)\xi^i\xi^j \quad \forall\xi\in\mathbb{R}^d. \qquad (12)
\end{aligned}$$

さらに集合 $\{J^M | M\in\mathcal{M}\}$ は有界であるとする. このとき, 推定量の系列 $\mathcal{E}$ が弱一致性 **(II)** に加えて,**(III)**,**(IV)** を満たすとき以下の一様 $Bahadur$ 型の不等式が成立する.

$$\alpha(\mathcal{E},\theta_0) \le B^s(\theta_0). \tag{13}$$

ここで一様 $Bahadur$ 型の下限 $B^s(\theta)$ は以下で定義した.

$$B^s(\theta) := \sup_{\mu\in P(\mathcal{M})}\inf\left\{ \sum_{i,j}\frac{J^\mu_{i,j}(\theta)\xi^i\xi^j}{2} \middle| \|\xi\|=1 \right\}. \tag{14}$$

ここで $P(\mathcal{M})$ は $\mathcal{M}$ 上の確率分布のなす集合を表す.

証明 $\theta,\theta'\in\Theta, \|\theta-\theta'\| > \epsilon$ に対して, (7) と同様にして,

$$\beta(\mathcal{E},\theta,\epsilon) \le \liminf_{n\to\infty} D^{\mu_n(\theta')}(\theta'\|\theta) \tag{15}$$

が示せる. ここで(15) の右辺が(7) と異なって $\liminf$ でよいのは左辺が収束値になっているからである. したがって $\|\xi\| > 1$ となる任意の $\xi\in\mathbb{R}^d$ について以下の式を得る.

$$\frac{\beta(\mathcal{E},\theta_0+\epsilon\xi,\epsilon)}{\epsilon^2} \le \frac{1}{\epsilon^2}\liminf_{n\to\infty} D^{\mu_n(\theta_0)}(\theta_0\|\theta_0+\epsilon\xi).$$

ここで条件 **(III)**,**(IV)** 及び(12) を用いて $\epsilon\to 0$ の極限を用いると次式を得る.

$$\alpha(\mathcal{E},\theta_0) \le \frac{1}{2}\liminf_{n\to\infty}\sum_{i,j} J^{\mu_n(\theta_0)}_{i,j}(\theta_0)\xi^i\xi^j.$$

ここで $\{J^M | M\in\mathcal{M}\}$ は有界性より, $\{J^\mu | \mu\in P(\mathcal{M})\}$ の有界性が得られる. 有界性を用いると以下の式が成立する.

$$\inf\left\{ \frac{1}{2}\liminf_{n\to\infty}\sum_{i,j} J^{\mu_n(\theta_0)}_{i,j}(\theta_0)\xi^i\xi^j \middle| \|\xi\| > 1 \right\} \le B^s(\theta_0)$$

したがって(13) が成立する. □

一般的に $B^w(\theta) \geq B^s(\theta)$ が成り立つことは容易に分るが, どのような場合に両者が一致するのだろうか. しかしながら, その必要十分条件は見つかっていない. 少なくとも $\overline{\{J^M(\theta)|M \in \mathcal{M}\}}$ が非負定値の意味で順序を入れたときに最大元を持つ場合には両者が一致する. したがって自明な場合として $d = 1$ のときは両者は一致する. ここで §2 の例 1 で扱った例について考える. $x = (\cos\phi, \sin\phi) \in S^1$ とすると,

$$J^{M_x}(\theta) = \frac{1}{2\sigma^2}\begin{pmatrix} \cos^2\phi & \cos\phi\sin\phi \\ \cos\phi\sin\phi & \sin^2\phi \end{pmatrix}$$

したがって $\{J^\mu(\theta)|\mu \in P(\mathcal{M})\}$ は 2 次元の実対称行列でトレースが $\frac{1}{2\sigma^2}$ の非負定値行列全体になる. この場合 $B^s(\theta) = \frac{1}{2\sigma^2}, B^s(\theta) = \frac{1}{4\sigma^2}$ となり一致しない.

5.2 上限を達成する推定量の系列

次に条件 **(II)** のみを満たす推定量で各点 Bahadur 型の上限 $B^w(\theta_0)$ を達成する推定量の構成する. 一般にこのような推定量は条件 **(III)**,**(IV)** を満たさないことを考えると, 一点でのみ非常によい rate を実現し, 他の点ではかろうじて弱一致性 **(II)** を満たしているものであると考えられる. その後で条件 **(III)**,**(IV)** を満たし, 一様 Bahadur 型の上限 $B^s(\theta_0)$ を達成する推定量の構成する.

定理 5 モデルが正則性条件 (5.1)(5.2) を満たすとする. このとき, 数列 $\{a_n\}$ を $a_n \to \infty, a_n/n \to 0$ を満たす数列とする. $\theta_0 \in \Theta$ に対して, 以下のように推定量の系列 $\{\mathcal{E}_{\theta_0}(a_n)\}$ を構成すると, 弱一致性 **(II)** を満たし θ_0 において(10) の等号を達成する.

$\{\mathcal{E}_{\theta_0}(a_n)\}$ の構成: b_n を自然数列としたとき推定量の系列 $\{\mathcal{E}_{\theta_0}(a_n)\}$ は次のように定義される. n 回の測定のうち最初の a_n 回は正則性条件 (5.1) を満たす測定 M_0 を施し, 確率分布族 $\{P(\rho_\theta, M_0 \times a_n)|\theta \in \Theta\}$ の最尤推定量 $\check{\theta}_n$ を得る. ここで $\hat{\xi}_n := (\check{\theta}_n - \theta_0)/\|\check{\theta}_n - \theta_0\|$ とおく. 次に正則性条件 (5.2) を満たす測定 $M_{\theta_0,\hat{\xi}_n}$ を残り $n-a_n$ 回行い, 確率分布族 $\{P(\rho_{\theta_0+t\hat{\xi}_n}, M_{\theta_0,\hat{\xi}_n} \times (n-a_n))|t \in \mathbb{R}\}$ の最尤推定量 $\hat{t}_{n-a_n}$ を得る. そして $\hat{\theta}_n := \theta_0 + \hat{t}_{n-a_n}\hat{\xi}_n$ を最終的な推定値とする.

正則性条件 (5.1): 以下の条件を満たす測定 μ_0 が存在する. 確率分布族 $\{P(\rho_\theta, \mu_0 \times n)|\theta \in \Theta\}$ の最尤推定量が弱一致性を満たす.

正則性条件 (5.2): 以下の条件を満たす基本測定 $M_{\theta_0,\xi}$ が存在する. 任意の十分小さな $\delta > 0$ に対して, 確率分布族 $\{P(\rho_{\theta_0+t\xi}, M_{\theta_0,\xi} \times n)|t \in \mathbb{R}\}$ の最尤推定量 $\hat{t}_n$ が $t = 0$ のとき漸近的に Bahadur の意味での上限を達成する. すなわち

$$-\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{\epsilon^2 n}\log P(|\hat{t}_n| > \epsilon|\rho_\theta, M_{\theta_0,\xi} \times n) = \frac{1}{2}\sup_{M\in\mathcal{M}}\sum_{i,j} J^M_{i,j}(\theta)\xi^i\xi^j. \quad (16)$$

なおこのとき $\|\xi\| = 1$ なる $\xi \in \mathbb{R}^d$ に関して一様に収束するものとする. さらに, 状態族 $\{\rho_{\theta_0+t\xi}|t \in \mathbb{R}\}$ から真のパラメータ θ までのノルムの意味で距離が δ より小さい点では上記, 最尤推定量 $\hat{t}_n$ から得られる推定値 $\theta_0 + \hat{t}_n\xi$ が真値から 2δ より大きくずれる確率が $n \to \infty$ で 0 になる. このとき 0 に近づく速さは $\|\xi\| = 1$ となる $\xi \in \mathbb{R}^d$ に関して一様であるとする.

定理 6 モデルが正則性条件 (6.1) ~ (6.4) を満たすとする. 数列 $\{a_n\}$ を $a_n \to \infty, a_n/n \to c$ を満たす数列とする任意の $\theta \in \Theta$ で以下に定義する推定量の系列 $\{\mathcal{E}(a_n)\}$ について以下の式を得る.

$$\alpha(\mathcal{E}(a_n), \theta) = (1-c)B^s(\theta). \quad (17)$$

なお $a_n/n \to 0$ となる場合では(17) は成立しない.

$\{\mathcal{E}(a_n)\}$ の構成: a_n を自然数列としたとき推定量の系列 $\{\mathcal{E}_\theta(a_n)\}$ は次のように定義される. n 回の測定のうち最初の a_n 回は正則性条件 (3) を満たす適当な測定 μ_0 を施し, 確率分布族 $\{P(\rho_\theta, \mu_0 \times a_n)|\theta \in \Theta\}$ の最尤推定量 $\check{\theta}_n$ を得る. 次に以下の条件を満たす測定 $\mu_{\check{\theta}_n}$ を残り $n-a_n$ 回行い, 確率分布族 $\{P(\rho_\theta, \mu_{\check{\theta}_n} \times (n-a_n))|\theta \in \Theta\}$ の最尤推定量 $\hat{\theta}_{n-a_n}$ を最終的な推定値とする.

正則性条件 (6.1): 以下の条件を満たす測定 μ_0 が存在する: 確率分布族 $\{P(\rho_\theta, \mu_0 \times n)|\theta \in \Theta\}$ の最尤推定量 $\hat{\theta}_n$ が $\forall\theta \in \Theta$ について以下の式を満たす:

$$\alpha(\{(\hat{\theta}_n, \mu_0 \times n)\}, \theta) \geq C(\theta). \quad (18)$$

なお $C(\theta)$ は $\theta \in \Theta$ ごとに決まる定数である.

正則性条件 (6.2): 任意の $\theta \in \Theta$ に対して以下の条件を満たす測定 μ_θ が存在する: 任意の $\epsilon' > 0$ に対して以下の条件を満たす $\delta > 0$ が取れる: $\|\theta - \theta'\| > \delta$ となる θ' について以下の式が成立する:

$$-\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{\epsilon^2 n}\log P(\|\hat{\theta}_n - \theta'\| > \epsilon|\rho_{\theta'}, \mu_\theta) \leq B^s(\theta) - \epsilon'. \quad (19)$$

なお, $\hat{\theta}_n$ は確率分布族 $\{P(\rho_\theta, \mu_\theta \times n)|\theta \in \Theta\}$ の最尤推定量とする.

正則性条件 (6.3): (19) の左辺の収束の速さであるが, $\theta \in \Theta$ に関して局所一様に収束する.

正則性条件 (6.4): $B^s(\theta)$ は θ に関して連続である.

確率分布族のパラメータ推定の大偏差型評価の場合, (10) 及び(13) に対応する評価は 1 次元の場合については Bahadur によりなされている [5]. この議論は多次元の確率分布族の場合にも容易に拡張できるがどうゆうわけか Bahadur の論文にはその記述が見られない. 一般に多次元の確率分布族のパラメータ推定では多次元でも両者の右辺は一致する. 逆に定理 6 のような任意の $\theta \in \Theta$ での上限の達成は 1 次元の確率分布族の場合に Fu により尤度関数に関する正則性条件の下で示されている [6]. 一方, 多次元の上限達成に関する研究は見当たらない. 少なくとも問題となる確率分布族が曲指数型分布族となる場合では(10) の上限を達成する推定量は構成できる. したがって(14) の右辺の上限を達成する測定 μ_θ を行ったときに得られる確率分布族 $\{P(\rho_\theta, \mu_\theta)\}$ が曲指数型分布族である場合には正則性条件 (6.1)(6.2)(6.4) は満たされる. 同様に(11) の右辺の上限を達成する測定 $\mu_{\theta_0,\xi}$ を行ったときに得られる確率分布族 $\{P(\rho_\theta, \mu_{\theta_0,\xi})\}$ が曲指数型分布族である場合には正則性条件 (5.1) 及び (5.2) の前半は満たされる. なお例 1 で扱った例ではこれらの正則性条件は全て満たされている.

6 結論

本稿では推定量に条件 **(II)** のみを課したときの rate の上限 $B^w(\theta)$ とそれに付け加えて条件 **(III)**,**(IV)** まで課したときの上限 $B^s(\theta)$ の 2 つの上限を導いた. $B^w(\theta)$ は §4 で扱った仮説検定の極限として導かれるものである. そして定理 5 で扱った $B^w(\theta)$ を達成する推定量の系列は定理 2 で扱った(6) の上限を達成する検定量の系列のある種の極限になっている. しかし一点でのみよい rate を実現するだけでは実際に推定量としては有効に機能しない. したがって §5.1 で導入した条件 **(III)**,**(IV)** を満たす必要があり, 上限 $B^s(\theta)$ を本来の推定量の rate の上限として捉えるべきである. ただ正則な確率分布族の場合には 2 つの上限が一致するので推定は検定のある種の極限として捉えることができた. しかし, 本稿で扱った測定を適応的に選択できる場合では推定は検定の極限として捉えることはできず, 推定問題と §4 で扱った検定問題との間には大きなギャップがある. ただし, このような問題設定の下でも未知パラメータが 1 つの場合にはこのようなギャップが生まれない. このようなギャップを見るには多次元であることが本質的である.

なお, 量子相関を許さない量子推定に本稿の議論を適用するためには量子的に定義された測定集合 $\mathcal{M}$ から, 一様 Bahadur 型の上限 $B^s(\theta_0)$ や各点 Bahadur 型の上限 $B^w(\theta_0)$ を求めることが問題になる. $B^w(\theta_0)$ の方は比較的容易に求まるが, $B^s(\theta_0)$ については部分的にしか求まっていない. これの計算には, 非可換的な議論が必要になり, 極めて困難である.

なお, これと類似のギャップは 1 パラメータであっても, 非正則モデルの推定や量子相関を許した場合での量子推定には現れる [7, 8].

謝辞

本稿は松本啓史氏との共同研究に基づいている. 氏には本研究を進める上で貴重なアドバイスを頂いた. この場を借りて感謝の意を表したい.

参考文献

[1] C. W. Helstrom, *Quantum detection and Estimation Theory* ,Academic Press, (1976).

[2] 長岡浩司, “Kullbacl Divergence と Fisher 情報量について -古典系から量子系へ- ,” 情報理論とその応用学会ジョイントミニワークショップ講演資料, pp. 63 (1992).

[3] 松本啓史, “データのとり方を適応的に変えれる場合の一次漸近論の構想,” 情報理論的学習ワークショップ予稿集 (1998).

[4] 林正人, 松本啓史, “測定自由度をもつ統計モデルと量子力学,” 数理解析研究所講究録 **1055**, 96-110, (1998).

[5] R. R. Bahadur, *Some Limit Theorems in Statistics* Regional Conference Series in Applied Mathematics, No 4, SIAM, Philadelphia (1971).

[6] J. C. Fu, On a theorem of Bahadur on the rate of convergence of point estimator, Ann. stat., **1**, 745-749 (1973).

[7] 林正人, “Affinity を用いた漸近論,” 数理解析研究所講究録 **1101**, 32-37, (1999).

[8] 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).