自己整合量子トモグラフィ MLE の一致性の証明

大阪大学・基礎工学研究科　田中冬彦
Fuyuhiko Tanaka
Graduate School of Engineering Science
Osaka University

Abstract

RIMS 研究会の参加者と議論した結果を参考に自己整合量子トモグラフィにおける最尤推定量や最小二乗推定量, 正則化推定量の一致性の証明を統一的に行う. 特に, 量子トモグラフィで前提となる強い条件があるため, M 推定の一致性の証明で用いる仮定や一様大数の法則を前提にしない方法で示す.

1　損失関数のクラスと一致性を示す準備

1.1　量子トモグラフィの概略と本稿で示すこと

まず, 記法と問題設定について説明する. ここで考えている, 量子トモグラフィでは M 種類の実験を n 回ずつ繰り返し, それぞれは, 二項分布に従う確率変数 $X^1,\dots,X^M$ として観測される. 後の議論のため, それぞれの平均をとった

$$\vec{f}_n = \begin{pmatrix} X^1/n, \\ \vdots \\ X^M/n \end{pmatrix} \in [0,1]^{\times M} \tag{1}$$

を $\mathbf{R}^M$ 値確率変数として考える. ここで, $\vec{f}_n$ の分布は未知パラメータ $\vec{\theta} = (\theta^1,\dots,\theta^M)^\top \in [0,1]^{\times M}$ によって完全に指定される. ただし, 以下の議論では二項分布という性質は使わない. $[0,1]^{\times M}$ に値をとる確率変数のみ仮定すればよい.

一方, 密度行列や測定, ゲートを指定する物理的なパラメータの動く範囲 $\mathcal{S} \subset \mathbf{R}^k$ はコンパクトであり, s を一つ指定すれば Born の公式から $\vec{\theta}$ が一意に定まる. $\vec{\theta}(s)$ は密度行列や測定, ゲートを s でパラメータ表示して, ゲート列を指定すれば, 明示的に s の関数として書くことができ, 通常は無限回微分可能な関数になる. （一般形はかなり煩雑になるので, 省略する. ）ただし, 以下の議論では $\vec{\theta}(s)$ は $\mathcal{S} \subset \mathbf{R}^k$ 上の連続関数であることのみ

を仮定して話を進める.

$\vec{\theta}(s)$ と観測される量, $\vec{f_n}$ は以下の式で結ばれる.

$$E_{s_0}[\vec{f_n}] = \vec{\theta}(s_0) \tag{2}$$

ここで, s_0 は真値をあらわすものとする. 上のモーメント条件は量子力学的な期待値に関する式を数理統計の用語で表現したものである.

ここで注意すべきことは, $\vec{\theta}(s)$ は一般に単射ではないことである. そこで, $[s_0] := \{s \in \mathcal{S} : \vec{\theta}(s_0) = \vec{\theta}(s)\}$ と定義すると, これらは, 無限に観測データを手に入れても真値 s_0 と区別がつかないパラメータの集合をあらわす. 同値関係が定義できるため, 同値類になっている (**ゲージ同値類**という.)

やや単純化した例としては $M = 1$ 種類の実験として

$$X \sim Bin\left(n, \frac{1+s_1 s_2}{2}\right),$$
$$\mathcal{S} := [-1,1] \times [-1,1]$$

が挙げられる. この時, $f_n = X/n$, $E[f_n] = \theta(s) = \frac{1+s_1 s_2}{2}$ である. $s_0 = (s_{0,1}, s_{0,2})$ が $\mathcal{S}$ の内点の場合は, どんなに n を増やしてもパラメータ s_1, s_2 は一意に定まらない[*1].

これが, 自己整合量子トモグラフィにおける**ゲージ不定性**とよばれるものである. そこで, このような場合に, 以下のような $\mathcal{S}$ 上の距離関数 d, 既知のパラメータ点 s_{tar}, 正則化パラメータ $\lambda_n (> 0)$ を用いて目的関数

$$F_n(s) := \left\| \vec{f_n} - \vec{\theta}(s) \right\|_2^2 + \lambda_n d(s, s_{tar}) \tag{3}$$

を最小化することでゲージ不定性を取り除いた推定値 ($s_{n,est}$ とあらわす) を求める. ここでは見やすいように, ややシンボリックに $d(s, s_{tar})$ と記載したが, 実際には s, s_{tar} と $1:1$ に対応する行列たちの何らかのノルムを考える.

[*1] より正確には $-1 < s_1 s_2 < 1$ では無数の解をもち, $s_1 s_2 = 1 \Leftrightarrow (s_1, s_2) = (1,1), (-1,-1)$, $s_1 s_2 = -1 \Leftrightarrow (s_1, s_2) = (-1,1), (1,-1)$ である.

なお, 数理統計では, ある関数を最小化することで得られる推定量を**最小コントラスト推定量** [3] とよぶ. 例えば, 尤度関数にマイナスをかけた関数の最小コントラスト推定量は最尤推定量であり, 正則化した最尤推定量も同様である.

上の最小化は, ラフに述べると $n \to \infty$ とした際に

$$\min_{s \in \mathcal{S}} d(s, s_{tar}) \qquad \text{subj. to } \vec{\theta}(s_0) = \vec{\theta}(s)$$

という最適化問題を解くことに相当する. s は $[s_0]$ の元ならなんでもよいのだが, その中で一番, s_{tar} に近いものを真値と考える.

本原稿では, 目的関数 (3) の右辺の第一項がもっと一般の形の場合も含めて, 推定量の列 $\{s_{n,est}\}$ が強一致性をもつことを示す. より具体的には

$$\lim_{n \to \infty} d(s_{n,est}, [s_0]) = 0, \qquad \text{a.s.}$$

を示す.

Remark 1 目的関数 (3) の右辺の第一項は一般に

$$F(\vec{f}_n) = \frac{1}{n} \sum_{j=0}^{n} G(X_j)$$

のような独立同一な確率変数 X_j の入った関数の平均の形に書けるとは**限らない**. もし, かける場合には M 推定の一般論 *(*特に一様大数の法則*)* が使える.

Remark 2 最小コントラスト推定量が強一致性をもつためのもっとも一般的な条件は吉田 *[3]* 定理 *4.4* で与えられている. ただし, そこでは損失関数をひとつ固定しており, かつ, 連続性や下半連続性, 極限の存在も不要にする代償として煩雑な条件が与えられている. 本稿では, 超伝導量子ビット系の量子トモグラフィやゲージ不定性をもつという背景 *(*詳しくは, 例えば杉山 *[2]* を参照せよ.*)* の下, 自然な条件を設定し, どういうクラスの損失関数であれば強一致性が保障されるかを明らかにする.

Remark 3 超伝導量子ビット系以外の物理系での量子トモグラフィに適用する上で, 緩められる仮定は以下のとおりである.

(a) 式 (1) では $\vec{f_n}$ の各成分の値域を $[0,1]$ に制限する必要はない.

(b) $\vec{f_n}$ に関するモーメント条件, 式 (2) の仮定も不要である. ($\vec{f_n} \to \vec{\theta}(s_0)$, a.s. を示すことができればよい.)

本稿では, 数理統計の専門以外の異なる分野の研究者にも読みやすいようにかなり行間を埋めて細かく証明を書いている. また, 数理統計の理論研究としての学術的な新規性をことさら主張するつもりはない. 特に, 本稿で示す定理の前提となる幾つかの条件ははるかに緩めることができる. 例えば有限次元ユークリッド空間の仮定などは不要であるが証明が無駄に抽象的になることを避けた.

1.2 損失関数の導入

目的関数 (3) の右辺の第一項をより一般的な形で考えるため, 以下のような関数 L を導入する.

$$\begin{aligned} L : [0,1]^{\times M} \times [0,1]^{\times M} &\longrightarrow [0,+\infty], \\ (\vec{\eta},\vec{\theta}) &\longmapsto L(\vec{\eta},\vec{\theta}) \end{aligned}$$

通常と同様にここでは L を**損失関数**とよぶことにする.

このとき, 損失関数 L に関して, 片側一様条件は以下で与えられる.

片側一様条件

各 $n = 1, 2, \ldots,$ に対し確率変数ベクトル $\vec{f_n}$ は二項分布 $Bin(n; \vec{\theta}(s_0))$ から得られるとする. この時, 確率変数ベクトルの列 $\{\vec{f_n}\}$ が L に関して, 片側一様条件を満たすとは以下を言う.
$\forall \epsilon > 0,$

$$\liminf_n \left\{ \inf_{s:\ d(s,[s_0]) \geq \epsilon} L\left(\vec{f_n}, \vec{\theta}(s)\right) \right\} \geq \inf_{s:\ d(s,[s_0]) \geq \epsilon} L\left(\vec{\theta}(s_0), \vec{\theta}(s)\right) \ \text{a.s.}$$

これから考えたいことは (正則化項はいったん忘れて) 損失関数について $L(\vec{f_n}, \vec{\theta}(s))$ の最小化で得られる s の推定量が, $n \to \infty$ で $L(\vec{\theta}(s_0), \vec{\theta}(s))$ の最小化解に近づいていくか

ということである．s の関数 $L(\vec{f_n},\vec{\theta}(s))$ が $L(\vec{\theta}(s_0),\vec{\theta}(s))$ に一様収束していれば後の証明はかなり平易になる．上の条件で符号を逆にした条件も成り立てば一様収束と同値であることがいえる．しかし, 尤度関数を扱うような場合, 片側一様条件しか成立しない．そこで, 本稿でも上の片側一様条件のみを示して一致性を示す．仮定が一様収束より緩くなるため証明は煩雑になる．

さらに

$$\vec{\eta}=\vec{\theta}\Leftrightarrow L(\vec{\eta},\vec{\theta})=0 \tag{4}$$

を仮定する（この条件は緩められる．）と識別性条件は以下のように与えられる．識別性条件は損失関数自体の性質として一致性の証明に使われる．

識別性条件

損失関数 L が以下を満たすとき, 識別性条件を満たすという．$\forall\epsilon>0,$

$$\inf_{s:d(s,[s_0])\geq\epsilon} L(\vec{\theta}(s_0),\vec{\theta}(s))>0$$

この条件の意味を理解するためにはたとえば $L(s,t)=g(s-t),\ g(x)=x^2e^{-x}$ という関数の振る舞いを考えてみるといいだろう．t を固定した時, $L(s-t)$ が最小になるのは $s=t$ だが, $s-t\to\infty$ で $L(s-t)\to 0$ となってしまう．

ここまでは, 相対エントロピーのような非対称なタイプも想定しているので, 対称性 $(L(\vec{\eta},\vec{\theta})=L(\vec{\theta},\vec{\eta}))$ や三角不等式は仮定しない．

2 章でみるように, 片側一様条件と識別性条件を仮定すると, 次の最小コントラスト推定量

$$s_{est,n}=\mathrm{argmin}\, L\left(\vec{f_n},\vec{\theta}(s)\right)$$

や正則化項 $R(s)(\geq 0)$, $\lambda_n>0$, $(\lambda_n\to 0)$ を加えて最小化する

$$s_{est,n}=\mathrm{argmin}\left\{L\left(\vec{f_n},\vec{\theta}(s)\right)+\lambda_n R(s)\right\}$$

の強一致性が示せる．なお，一般には min が存在するとは限らないが，量子トモグラフィの文脈では目的関数は s の連続関数であり動く範囲も通常，有界閉集合であるため min が存在する．(詳しくは 2 章を参照.)

上で与えた識別性条件は損失関数 L について連続性や下半連続性すら仮定していないし，$\vec{f_n}$ が平均の形であるとか独立観測の仮定も不要である．一方で，既に述べたように，量子トモグラフィの設定では自然な条件がいくつか入ってくる．また，実際には，高次元パラメータについて非線形制約条件の下の最適化を扱うため，損失関数や正則化項の検討も重要である．

そこで，本原稿では，どのようなクラスの損失関数であれば，上に掲げた片側一様条件や識別性条件を満たすのか明らかにしたい．これは純粋な数理統計的研究とは異なる観点である．緩めすぎても意味がないので，ここでは以下の二つのタイプの損失関数を挙げておく．

(i) 連続な距離関数による損失

(ii) クロス型損失と拡張クロス型損失

1.2.1 (i) 連続な距離関数による損失

(i) はユークリッドノルム $\left\|\vec{\theta}-\vec{\eta}\right\|$ に代表されるように以下の性質を満たす損失関数である．

(a) $L(\vec{\theta},\vec{\eta})=L(\vec{\eta},\vec{\theta})$ であり，三角不等式を満たす (つまり，通常の意味での距離関数)

(b) $L(\vec{\theta},\vec{\eta})$ は片方の引数を固定したときに，もう片方の引数について連続.

なお，後で見るように連続な距離関数の損失では，今回の枠組みでは一様収束を示せる．そのため，本稿のような方法を経由せず，かなり直接的に強一致性を示せる．本稿では KL-divergence のように (i) に含まれない損失で一致性を担保できる損失関数のクラスを

調べるのが主眼である.

1.2.2 (ii) クロス型損失と拡張クロス型損失

(ii) は確率変数の部分とパラメータの部分が積の形に分離したタイプの和である.

Definition 1 クロス型の損失関数を以下で定義する.

$$L(\vec{\eta},\vec{\theta})=\sum_{j=1}^{a}g_j(\vec{\eta})h_j(\vec{\theta}),\ \vec{\eta},\vec{\theta}\in[0,1]^{\times M}$$

ここで $g_1,\dots,g_a$ は非負の連続関数, $h_1,\dots,h_a$ は非負の下半連続関数 *(lsc)* とする. また, $0\cdot\infty=0$ と約束する.

一般に上の形のままだと (4) を満たさない. そこで以下のようにして (4) を満たすように修正した損失関数を使う.

Definition 2 クロス型の損失関数 L に対し, $M(\vec{\eta})=-L(\vec{\eta},\vec{\eta})$ は $\vec{\eta}$ の連続関数とする. このとき以下で定義される損失関数を拡張クロス型損失関数とよぶ.

$$\widetilde{L}(\vec{\eta},\vec{\theta})=L(\vec{\eta},\vec{\theta})+M(\vec{\eta}),\qquad \vec{\eta},\vec{\theta}\in[0,1]^{\times M} \tag{5}$$

識別性条件は拡張クロス型として (4) を仮定して示す. 片側一様条件はクロス型の方が示しやすいので, 最初にクロス型, そのあと拡張クロス型で示す.

例 1: 多項分布 $\overrightarrow{W}=(W_1,\dots,W_t)\sim MN(n,q_1,\dots,q_t)$

最尤推定量は（M 推定量の枠組みでとらえることもできるが）次のように KL-divergence の最小化と考えることができる.

$$\begin{aligned}\widetilde{L}(\overrightarrow{W}/n,\vec{q})&=\sum_{j=1}^{t}\frac{W_j}{n}\log\frac{W_j}{n}+\sum_{j=1}^{t}\frac{W_j}{n}\log\frac{1}{q_j}\\&=M(\overrightarrow{W})+\sum_{j=1}^{t}g_j(\overrightarrow{W}/n)h_j(\vec{q})\end{aligned}$$

ここで, $g_j(\overrightarrow{W}/n) = W_j/n$, $h_j(\vec{q}) = \log 1/q_j$ でありそれぞれ非負の連続関数と下半連続関数だから

$$L(\overrightarrow{W}/n, \vec{q}) := \sum_{j=1}^{t} g_j(\overrightarrow{W}/n) h_j(\vec{q})$$

はクロス型損失関数である．また, $M(\overrightarrow{W}) = \sum_{j=1}^{t} \frac{W_j}{n} \log \frac{W_j}{n}$ はパラメータ $q_1, \ldots, q_t$ を含まないから最小化推定を考える際には寄与しないが

$$M(\overrightarrow{W}) = -L(\overrightarrow{W}/n, \overrightarrow{W}/n)$$

の形をしている．また, M 自体は連続関数になっている.[*2] この意味で KL-divergence は拡張クロス型損失関数 (5) である．

例 2：量子トモグラフィの設定 (m 種類の実験) $X^j \sim Bin(n, \theta_j)$; $j = 1, \ldots, m$

$\vec{f}_n = (X^1, \ldots, X^m)/n$ として最尤推定量は次のような KL-divergence の最小化に帰着する．

$$\begin{aligned} L(\vec{f}_n, \vec{\theta}) &= \sum_{j=1}^{m} f_n^j \log f_n^j + \sum_{j=1}^{m} f_n^j \log \frac{1}{\theta_j} \\ &\qquad + \sum_{j=1}^{m} (1 - f_n^j) \log(1 - f_n^j) + \sum_{j=1}^{m} (1 - f_n^j) \log \frac{1}{1 - \theta_j} \\ &= M(\vec{f}_n) + \sum_{j=1}^{2m} g_j(\vec{f}_n) h_j(\vec{\theta}) \end{aligned}$$

ここで

$$g_j(\vec{f}_n) = \begin{cases} f_n^j, & j = 1, \ldots, m, \\ 1 - f_n^j, & j = m+1, \ldots, 2m, \end{cases}$$

$$h_j(\vec{\theta}) = \begin{cases} \theta_j, & j = 1, \ldots, m, \\ 1 - \theta_j, & j = m+1, \ldots, 2m, \end{cases}$$

[*2] $f(x, y) = x \log y$ は $0 \le x \le 1$ を固定すると $0 \le y \le 1$ に関して lsc であるが $h(x) = x \log x$ は $[0, 1]$ 上連続関数である．

および $M(\vec{f_n}) = \sum_{j=1}^{m} \{f_n^j \log f_n^j + (1 - f_n^j) \log(1 - f_n^j)\}$ とおいた．例 1 と同様に拡張クロス型損失関数 (5) の形になっている．

上の例からわかるように $L(\vec{\eta}, \vec{\theta})$ において第一引数 $\vec{\eta}$ には確率変数, 第二引数 $\vec{\theta}$ にはパラメータが入り, それぞれが分離されている点が後の証明で重要な役割を果たす．

1.3　連続な距離関数と拡張クロス型損失が識別性条件を満たすこと

前節で導入した二つのタイプの損失関数ともに, 第一引数を固定した場合に第二引数について lsc である．このことと $\mathcal{S}$ のコンパクト性から識別性条件は容易に示せる．

まず, 証明に必要な条件をまとめておくと以下のようになる．ここではクロス型は考えない．

(i) $L \geq 0$

(ii) $L(\vec{\theta}(s_0), \vec{\theta}(s))$ は s に関して lsc.

(iii) $L(\vec{\theta}(s_0), \vec{\theta}(s)) = 0 \Leftrightarrow s \in [s_0]$

二つのタイプの損失関数とも上の三つの条件を満たしている．まず, 定義から上の条件 (i), (iii) を満たすことは明らかであろう．また, 条件 (ii) は以下の Lemma からわかる．(証明は省略．)

Lemma 1 *$\vec{\theta} : \mathcal{S} \longrightarrow [0,1]^{\times M}$ が連続, $h : [0,1]^{\times M} \longrightarrow \mathbf{R}$ が lsc とすると, 合成関数 $h(\vec{\theta}(s))$ も lsc である.*

さて, 以上を踏まえると識別性条件が容易に示せる．

Lemma 2 *$\vec{\theta}(s)$ を $\mathcal{S}$ 上の連続関数とせよ．また, $L(\vec{\eta}, \vec{\theta})$ を前節で定義した (i) 連続な距*

離関数, *(ii)* 拡張クロス型, いずれかの損失関数とせよ. このとき, 任意の $\epsilon > 0$ について

$$\inf_{s:d(s,[s_0])\geq\epsilon} L(\vec{\theta}(s_0), \vec{\theta}(s)) > 0$$

となる.

proof

まず, inf の範囲は $\mathcal{S}$ の閉部分集合だからコンパクト. $L(\vec{\theta}(s_0), \vec{\theta}(s))$ は s の関数として lsc なので, 最小値を達成する点 s_* が存在. 仮に $L(\vec{\theta}(s_0), \vec{\theta}(s_*)) = 0$ とすると $d(s_*, [s_0]) > 0$ という条件に反する. したがって, $L(\vec{\theta}(s_0), \vec{\theta}(s_*)) > 0$. *Q.E.D.*

Remark 4 L, $\vec{\theta}(s)$ に対する条件をはずして, 上の識別性条件そのものを仮定にすることもできる. *(*吉田 *[3],* 条件 *[C2], p.171).*

1.4 連続な距離関数が片側一様条件を満たすこと

本節と次節で片側一様条件をより強い以下の形で示す.

片側一様条件 (強いバージョン)

任意のコンパクト部分集合 $K \subset \mathcal{S}$ について

$$\liminf_n \left\{ \inf_{s\in K} L\left(\vec{f}_n, \vec{\theta}(s)\right) \right\} \geq \inf_{s\in K} L\left(\vec{\theta}(s_0), \vec{\theta}(s)\right) \text{ a.s.}$$

証明は (i) の連続な距離関数の場合, $\vec{f}_n \to \vec{\theta}(s_0)$ a.s. と三角不等式から明らかである. 実際,

$$L\left(\vec{\theta}(s_0), \vec{\theta}(s)\right) \leq L\left(\vec{\theta}(s_0), \vec{f}_n\right) + L\left(\vec{f}_n, \vec{\theta}(s)\right)$$

に注意すると

$$\inf_K L\left(\vec{\theta}(s_0), \vec{\theta}(s)\right) \leq L\left(\vec{\theta}(s_0), \vec{f}_n\right) + \inf_K L\left(\vec{f}_n, \vec{\theta}(s)\right)$$

ここで, $\liminf$ をとると, a_n が収束する場合, $\liminf(a_n+b_n)=\lim a_n+\liminf b_n$ とかけることに注意すれば

$$\begin{aligned}\inf_K L\left(\vec{\theta}(s_0),\vec{\theta}(s)\right) &\le \lim_n L\left(\vec{\theta}(s_0),\vec{f}_n\right)+\liminf_n\left\{\inf_K L\left(\vec{f}_n,\vec{\theta}(s)\right)\right\}\\ &=\liminf_n\left\{\inf_K L\left(\vec{f}_n,\vec{\theta}(s)\right)\right\}\end{aligned}$$

となる.

Remark 5 連続な距離関数の損失のみを取り扱って一致性を示すのであれば, 片側一様条件を示さず, 一様収束を示せばよい. つまり, $g_n(s)=L\left(\vec{f}_n,\vec{\theta}(s)\right)$ 及び, $g(s)=L\left(\vec{\theta}(s_0),\vec{\theta}(s)\right)$ とおくと三角不等式により

$$|g_n(s)-g(s)|\le L\left(\vec{\theta}(s_0),\vec{f}_n\right),\qquad \forall s\in\mathcal{S}$$

となる. これより

$$\inf_{s\in\mathcal{S}}|g_n(s)-g(s)|\le L\left(\vec{\theta}(s_0),\vec{f}_n\right)\to 0$$

から $\{g_n\}$ の g への一様収束が示せる. 上の証明から明らかなように $\mathcal{S}$ のコンパクト性は不要である.

次に (ii) のクロス型損失関数の場合であるが, こちらが非自明であり K のコンパクト性が本質的である. KL-divergence も含む広い場合に適用するため, 節を分けて幾つかの補題と合わせて証明する.

1.5 クロス型損失が片側一様条件を満たすこと

以下では, まず, クロス型損失の場合に, 片側一様条件を満たすことを示す. その後, KL-divergence などを含む拡張クロス型損失について, 片側一様条件を満たすことを示す. 後者は, 前者の結果を利用して容易に得られる.

まず, クロス型損失関数 $L(\vec{\eta},\vec{\theta})$ では第二引数 $\vec{\theta}$ について下半連続に条件を緩めていることに注意して, 損失関数の極限の形を調べる.

また, 本節では元々の $\vec{\theta}(s)$ が期待値パラメータである (式 (2)) という仮定（及び大数の強法則）を緩め

$$\vec{f}_n \to \vec{\theta}(s_0) \qquad \text{a.s.}$$

を仮定として話を進める. したがって, 観測の独立性なども不要であり M 推定量の一致性とは異なる議論になる.

Lemma 3 確率変数列 $\{\vec{f}_n\}$ が $\vec{\theta}(s_0)$ に as 収束しているとせよ. クロス型損失関数 $L(\vec{\eta}, \vec{\theta})$ に対し s を任意に固定するとき以下が成立.

$$\lim_n L(\vec{f}_n, \vec{\theta}(s)) = L(\vec{\theta}(s_0), \vec{\theta}(s)) \tag{6}$$

ここで, 両辺が ∞ も含めている.

proof

今, 仮定より L は

$$L(\vec{\eta}, \vec{\theta}) = \sum_{j=1}^{a} g_j(\vec{\eta}) h_j(\vec{\theta}), \qquad \vec{\eta}, \vec{\theta} \in [0,1]^{\times M}$$

とかける.

(i) $L(\vec{\theta}(s_0), \vec{\theta}(s)) < \infty$ の場合:

この時は $\vec{f}_n \to \vec{\theta}(s_0)$ および $g_j(\vec{\eta}), j = 1, \ldots, a$ の連続性から (6) 式は明らかである.

(ii) $L(\vec{\theta}(s_0), \vec{\theta}(s)) = \infty$ の場合:

この時は, 少なくとも一つの j について

$$g_j(\vec{\theta}(s_0)) > 0,\ h_j(\vec{\theta}(s)) = \infty$$

になっている. g_j の連続性から $\vec{\theta}(s_0)$ の近傍でも同様であり, とくに, ある番号 N が存在して $n \geq N \Rightarrow g_j(\vec{f}_n) > 0$. したがって, 各関数の非負性から $n \geq N$ のとき

$$L(\vec{f}_n, \vec{\theta}(s)) \geq g_j(\vec{f}_n) h_j(\vec{\theta}(s)) = \infty$$

より両辺 ∞ となって (6) 式が示される. *Q.E.D.*

Remark 6 M 推定量では極限値の議論ができない. s を任意に固定するとき

$$\liminf_n \frac{1}{n}\sum_{j=1}^{n} g(X_j, \vec{\theta}(s)) \geq g(E[X_j], \vec{\theta}(s)) \tag{7}$$

という不等式のみ示せる. ただし, 推定量を得る際に s に関する最小化のみを考えるため, 上のような評価で十分である.

以下がこの原稿の主要結果である.

Theorem 1 確率変数列 $\{\vec{f}_n\}$ が $\vec{\theta}(s_0)$ に as 収束しているとせよ. このとき, クロス型の損失関数 L, および任意のコンパクト部分集合 K に対し, 片側一様条件（強いバージョン）が成立する. すなわち,

$$\liminf_n \inf_{s\in K} L(\vec{f}_n, \vec{\theta}(s)) \geq L_0 \qquad \text{a.s.}$$

ここで $L_0 := \inf_{s\in K} L(\vec{\theta}(s_0), \vec{\theta}(s))$.

式 (7) における独立変数の和 $\frac{1}{n}\sum$ の部分が $L(\vec{f}_n, \vec{\theta}(s))$ のように変わり, さらに $\inf_K$ をとっていることに注意する.

さて, Theorem 1 の証明のためにいくつか準備する. まず, lsc の一般的な性質として下側でみると連続関数と同じ性質が成り立つ. (証明は省略.)

Lemma 4 距離空間 (X, d) 上で $f(x)$ が lsc のとき, $\rho > 0$ を任意にとると

$$\lim_{\rho\to 0} \inf_{d(x,x')\leq\rho} f(x') = f(x)$$

が成立.

次に任意の $\rho > 0$ について

$$\varphi(\vec{\eta}, s, \rho) := \inf_{s' :\ d(s,s') \le \rho} L(\vec{\eta}, \vec{\theta}(s'))$$

を導入する．（lower envelope などともいう．）

上の Lemma 4 を用いると以下が示される．

$$L(\vec{\eta}, \vec{\theta}(s)) \ge \varphi(\vec{\eta}, s, \rho), \tag{8}$$

$$\lim_{\rho \downarrow 0} \varphi(\vec{\eta}, s, \rho) = L(\vec{\eta}, \vec{\theta}(s)), \tag{9}$$

$$\varphi(\vec{\eta}, s, \rho) \ge \sum_{j=1}^{a} g_j(\vec{\eta}) \inf_{d(s,s') \le \rho} h_j(\vec{\theta}(s')) \tag{10}$$

最後の式 (10) はクロス型損失の仮定，$g_j(\vec{\eta}) \ge 0$ であることと $\inf_s\{a(s) + b(s)\} \ge \inf_s a(s) + \inf_s b(s)$ から従う．

また, Theorem 1 の証明は長いので, 例外処理を先に示しておく．

Lemma 5 *Theorem 1* の条件下で $L_0 = \infty$ の場合,

$$\liminf_{n} \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) = \infty \qquad \text{a.s.}$$

が成立．

proof

仮定より $L(\vec{\theta}(s_0), \vec{\theta}(s)) = \infty,\ \forall s \in K$ であるから, 各 s ごとに少なくとも一つの j があり $\xi_j := g_j(\vec{\theta}(s_0)) > 0$ および $h_j(\vec{\theta}(s)) = \infty$. そこで J をそのような添え字の全体とする．（$J \subset \{1, \ldots, a\}$ で有限集合に注意する．）このとき,

$$\sum_{j \in J} h_j(\vec{\theta}(s)) = \infty,\ \forall s \in K$$

が成立する．

一方で十分, 大きい N をとれば

$$n \ge N \Rightarrow g_j(\vec{f}_n) > \xi_j/2,\ \forall j \in J$$

が成立する. このとき, 任意の s について

$$
\begin{aligned}
L(\vec{f}_n, \vec{\theta}(s)) &\geq \sum_{j \in J} g_j(\vec{f}_n) h_j(\vec{\theta}(s)) \\
&\geq \sum_{j \in J} \frac{\xi_j}{2} h_j(\vec{\theta}(s)) \\
&\geq \frac{1}{2} \min_{j \in J} \xi_j \sum_{j \in J} h_j(\vec{\theta}(s)) = \infty
\end{aligned}
$$

つまり,

$$
n \geq N \Rightarrow \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) = \infty
$$

より題意が示された.

Q.E.D.

以上の準備を踏まえて Theorem 1 を証明する.

proof

$L_0 = \infty$ の場合は Lemma 5 で示したので, $L_0 < \infty$ の場合のみ示せばよい.

まず, $\sup_{s \in K} L(\vec{\theta}(s_0), \vec{\theta}(s)) < \infty$ として証明する. $\sup_{s \in K} L(\vec{\theta}(s_0), \vec{\theta}(s)) = \infty$ の場合は, 最後に説明する.

さて, s_0, s を固定すると, 任意の $\rho > 0$ について (8), (10) に注意して

$$
\begin{aligned}
L(\vec{\theta}(s_0), \vec{\theta}(s)) &\geq \varphi(\vec{\theta}(s_0), s, \rho) \\
&\geq \sum_{k=1}^{a} g_k(\vec{\theta}(s_0)) \inf_{d(s,s') \leq \rho} h_k(\vec{\theta}(s'))
\end{aligned}
$$

を得る. (9) より $\rho \to 0$ とすると最後の項は $L(\vec{\theta}(s_0), \vec{\theta}(s))(< \infty)$ に（単調増加で）収束する.

以下では $\epsilon > 0$ を任意に固定する. すると, 各点 $s \in K$ において, $\rho_s > 0$ が存在し

$$
\sum_{k=1}^{a} g_k(\vec{\theta}(s_0)) \inf_{d(s,s') \leq \rho_s} h_k(\vec{\theta}(s')) \geq L(\vec{\theta}(s_0), \vec{\theta}(s)) - \epsilon \tag{11}
$$

とできる *3. (ρ_s は L の第一引数 $\vec{\theta}(s_0)$ に依存してもよい. ）

さて, ここで次のような開被覆 V_s をとる.

$$V_s := \{s' \in K : \ d(s,s') < \rho_s\}$$

すると, K のコンパクト性から高々有限個の $s_1, \ldots, s_\alpha$ について

$$K \subset \bigcup_{l=1}^{\alpha} V_l$$

が成立. ここで, $V_l := V_{s_l}$ とした. また, $\rho_l := \rho_{s_l}$ と略記する.

さて, 今, 任意の s をとると, 必ず, ある番号 j があって $s \in V_j$, つまり, $d(s,s_j) \leq \rho_j$ を満たしているので,

$$\begin{aligned} L(\vec{f}_n, \vec{\theta}(s)) &\geq \inf_{s' \in V_j} L(\vec{f}_n, \vec{\theta}(s')) \\ &= \varphi(\vec{f}_n, s_j, \rho_j) \end{aligned} \tag{12}$$

ここで最後の式は $j = 1, \ldots, \alpha$ の最小値でさらに下からおさえられる.

以上から任意の s について

$$L(\vec{f}_n, \vec{\theta}(s)) \geq \min_{l=1,\ldots,\alpha} \varphi(\vec{f}_n, s_l, \rho_l)$$

が成立する. さらに左辺で inf をとると

$$\inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) \geq \min_{l=1,\ldots,\alpha} \varphi(\vec{f}_n, s_l, \rho_l)$$

この不等式は任意の n, $\vec{f}_n$ で成立するので下極限をとると

$$\begin{aligned} \liminf_n \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) &\geq \liminf_n \left\{ \min_{l=1,\ldots,\alpha} \varphi(\vec{f}_n, s_l, \rho_l) \right\} \\ &= \min_{l=1,\ldots,\alpha} \liminf_n \varphi(\vec{f}_n, s_l, \rho_l) \\ &\geq \min_{l=1,\ldots,\alpha} \liminf_n \sum_{k=1}^{a} g_k(\vec{f}_n) \inf_{d(s_l,s') \leq \rho_l} h_k(\vec{\theta}(s')) \end{aligned}$$

最後の不等式は (10) を用いた.

---

*3 $\sup_{s \in K} L(\vec{\theta}(s_0), \vec{\theta}(s)) < \infty$ の仮定はここで使っている.

ここで仮定から $g_k(\vec{f}_n) \to g_k(\vec{\theta}(s_0))$ a.s. だから

$$
\begin{aligned}
\liminf_n \inf_{s\in K} L(\vec{f}_n, \vec{\theta}(s)) &\geq \min_{l=1,\ldots,\alpha} \left\{ \sum_{k=1}^{a} g_k(\vec{\theta}(s_0)) \inf_{d(s_l,s')\leq \rho_l} h_k(\vec{\theta}(s')) \right\} \\
&\geq \min_{l=1,\ldots,\alpha} \{L(\vec{\theta}(s_0), \vec{\theta}(s_l)) - \epsilon\} \\
&= \min_{l=1,\ldots,\alpha} \{L(\vec{\theta}(s_0), \vec{\theta}(s_l))\} - \epsilon \\
&\geq \inf_{s\in K} L(\vec{\theta}(s_0), \vec{\theta}(s)) - \epsilon \\
&= L_0 - \epsilon
\end{aligned}
$$

以上より ϵ が任意だったから

$$
\liminf_n \inf_{s\in K} L(\vec{f}_n, \vec{\theta}(s)) \geq L_0
$$

が示された．

最後に $\sup_{s\in K} L(\vec{\theta}(s_0), \vec{\theta}(s)) = \infty$ の場合を考える．$L_0 < \infty$ より，正の定数 $\Lambda > L_0$ を用いて

$$
h_j^{\Lambda}(\vec{\theta}) = h_j(\vec{\theta}) \wedge \Lambda
$$

と定義し，

$$
L^{\Lambda}(\vec{\eta}, \vec{\theta}) = \sum_{j=1}^{a} g_j(\vec{\eta}) h_j^{\Lambda}(\vec{\theta})
$$

とおく．h_j^{Λ} は上に有界な lsc であり，コンパクト集合 $\mathcal{S}$ 上で g_j が有界なこととあわせると L^{Λ} は有界なクロス型損失になっている．また，

$$
\begin{aligned}
&L(\vec{\eta}, \vec{\theta}) \geq L^{\Lambda}(\vec{\eta}, \vec{\theta}), \qquad \forall \vec{\eta}, \forall \vec{\theta}, \\
&\inf_{s\in K} L^{\Lambda}(\vec{\theta}(s_0), \vec{\theta}(s)) = L_0, \\
&\sup_{s\in K} L^{\Lambda}(\vec{\theta}(s_0), \vec{\theta}(s)) < \infty
\end{aligned}
$$

が成り立つ．以上を用いると L^{Λ} について上の証明を用いることができて

$$
\begin{aligned}
\liminf_n \inf_{s\in K} L(\vec{f}_n, \vec{\theta}(s)) &\geq \liminf_n \inf_{s\in K} L^{\Lambda}(\vec{f}_n, \vec{\theta}(s)) \\
&\geq \inf_{s\in K} L^{\Lambda}(\vec{\theta}(s_0), \vec{\theta}(s)) \\
&\geq L_0
\end{aligned}
$$

を得る.

Q.E.D.

上の結果からただちに次のことがいえる.

Theorem 2 確率変数列 $\{\vec{f}_n\}$ が $\vec{\theta}(s_0)$ に *as* 収束しているとせよ. このとき, クロス型の損失関数に対し, 極限と最小化の順序交換が可能である. つまり

$$\lim_n \left\{ \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) \right\} = L_0 = \inf_{s \in K} \left\{ \lim_n L(\vec{f}_n, \vec{\theta}(s)) \right\} \quad \text{a.s.}$$

が成立.

Proof 任意に s を固定するとき, inf の定義から

$$L(\vec{f}_n, \vec{\theta}(s)) \geq \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s))$$

となる. 両辺で上極限を考えると, 左辺は *Lemma 3* から極限が存在するので

$$L(\vec{\theta}(s_0), \vec{\theta}(s)) = \lim_n L(\vec{f}_n, \vec{\theta}(s)) \geq \limsup_n \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s))$$

$\limsup, \liminf$ の不等式と *Theorem 1* を合わせると

$$\begin{aligned} L(\vec{\theta}(s_0), \vec{\theta}(s)) &\geq \limsup_n \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) \\ &\geq \liminf_n \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) \\ &\geq L_0 \end{aligned}$$

となる. 最後に s について inf をとると

$$\begin{aligned} L_0 = \inf_{s \in K} L(\vec{\theta}(s_0), \vec{\theta}(s)) &\geq \limsup_n \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) \\ &\geq \liminf_n \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) \\ &\geq L_0 \end{aligned}$$

によりすべて等号成立.

Q.E.D.

Remark 7 ここで用いているのは, 確率変数列 $\{\vec{f_n}\}$ が $\vec{\theta}(s_0)$ に as 収束するという仮定と損失関数の形の制限のみであり,

$$\frac{1}{n}\sum_j g(X_j, \vec{\theta}(s))$$

のような和の形は仮定していない点が重要である. そのため, 一様大数の法則は使わずに示している点で M 推定量の一致性の証明とは異なる.

Remark 8 不等式 (12) において ρ_j は $\vec{\theta}(s_0)$ の点でとっているため

$$L(\vec{\theta}(s_0), \vec{\theta}(s)) \geq \varphi(\vec{\theta}(s_0), s_j, \rho_j) \geq L(\vec{\theta}(s_0), \vec{\theta}(s)) - \epsilon,$$

は成立するが

$$L(\vec{f_n}, \vec{\theta}(s)) \geq \varphi(\vec{f_n}, s_j, \rho_j) \geq L(\vec{f_n}, \vec{\theta}(s)) - \epsilon,$$

は保証していない.

1.6　拡張クロス型損失関数の場合

クロス型損失 L に $\vec{\eta}$ の連続関数 $M(\vec{\eta})$ を付け加えた場合も同様の結果が成立する. そこで, 拡張クロス型損失関数 $\widetilde{L}$ を

$$\widetilde{L}(\vec{\eta}, \vec{\theta}) = L(\vec{\eta}, \vec{\theta}) + M(\vec{\eta}) \tag{13}$$

とおいて以下を示そう.

Collorary 1 確率変数列 $\{\vec{f_n}\}$ が $\vec{\theta}(s_0)$ に as 収束しているとせよ. このとき, 拡張クロス型損失関数 (13) に対し片側一様条件（強いバージョン）が成立する. すなわち,

$$\liminf_n \inf_{s\in K} \widetilde{L}(\vec{f_n}, \vec{\theta}(s)) \geq \widetilde{L}_0 \tag{14}$$

ここで $\widetilde{L}_0 := \inf_{s\in K} \widetilde{L}(\vec{\theta}(s_0), \vec{\theta}(s))$.

proof

まず,

$$\begin{aligned}\widetilde{L}_0 &= \{\inf_{s\in K} L(\vec{\theta}(s_0),\vec{\theta}(s))\} + M(\vec{\theta}(s_0))\\ &= L_0 + M(\vec{\theta}(s_0))\end{aligned}$$

とかけることに注意する．さて，示すべき不等式の左辺は

$$\begin{aligned}\liminf_n \inf_{s\in K} \widetilde{L}(\vec{f}_n,\vec{\theta}(s)) &= \liminf_n \left\{\inf_{s\in K} L(\vec{f}_n,\vec{\theta}(s)) + M(\vec{f}_n)\right\}\\ &= \liminf_n \inf_{s\in K} L(\vec{f}_n,\vec{\theta}(s)) + M(\lim_n \vec{f}_n) \qquad (*)\\ &= \liminf_n \inf_{s\in K} L(\vec{f}_n,\vec{\theta}(s)) + M(\vec{\theta}(s_0))\end{aligned}$$

となるから，*Theorem 1* より (14) が示される．

Q.E.D.

Remark 9 $M(\vec{\eta})$ は負もとりえる関数であることに注意する．もし非負なら $g_{a+1}(\vec{\eta}) = M(\vec{\eta})$, $h_{a+1}(\vec{\theta}) = 1$ としてクロス型に含まれる．

Remark 10 一般に二つの実数列 $\{a_n\},\{b_n\}$ では $\liminf(a_n + b_n) \geq \liminf a_n + \liminf b_n$ であり，等号が成立するとは限らない[*4] しかし，片方が収束列の場合は等号が成立する．上の変形では $*$ の部分で，$M(\vec{\eta})$ の連続性から $M(\vec{f}_n)$ は収束列になっているため等号で結んでいる．

以下も *Theorem 2* を利用して容易に示せるので証明は省略する．

Collorary 2 確率変数列 $\{\vec{f}_n\}$ が $\vec{\theta}(s_0)$ に *as* 収束しているとせよ．このとき，拡張クロス型損失関数 (13) に対し，極限と最小化の順序交換が可能である．つまり

$$\lim_n \left\{\inf_{s\in K} \widetilde{L}(\vec{f}_n,\vec{\theta}(s))\right\} = \widetilde{L}_0 = \inf_{s\in K}\left\{\lim_n \widetilde{L}(\vec{f}_n,\vec{\theta}(s))\right\} \quad \text{a.s.}$$

[*4] 例えば $a_n = (-1)^n, b_n = (-1)^{n+1}$ としてみよ．

2　一致性の証明

先の損失関数 $L(\vec{\eta}, \vec{\theta})$ が拡張クロス型の場合に最小コントラスト推定量の一致性を証明する．*Theorem 2* が示せていることと，損失に関する識別性条件件があるため，後は，通常の M 推定量の一致性の証明などと同様にできる．ここでは *Ferguson [1]* で紹介している証明 *(Theorem17, p.114-115.)* と同様に示す．

Theorem 3 $\vec{\theta}(s)$ をコンパクト集合 $\mathcal{S}$ から $[0,1]^M$ への連続関数とし，確率変数列 $\{\vec{f_n}\}$ が $\vec{\theta}(s_0)$ に *as* 収束しているとする．また，$L(\vec{\eta}, \vec{\theta})$ を (4) を満たす拡張クロス型の損失関数とせよ．このとき，$\inf_{s \in \mathcal{S}} L(\vec{f_n}, \vec{\theta}(s))$ において inf を達成する s の一つを $s_{n,est}$ と書くと以下が成立．

$$\lim_n d(s_{n,est}, [s_0]) = 0 \qquad \text{a.s.}$$

つまり，最小コントラスト推定量の列 $\{s_{n,est}\}$ は強一致性をもつ．

proof

任意の $\epsilon > 0$ をひとつ固定する．以下，確率変数列の収束はすべて *almost sure* であるが，煩雑になるので省略する．

まず，拡張クロス型損失は識別性条件 *(Lemma 2)* を満たすから，この ϵ に対して

$$\eta_\epsilon := \inf_{s : d(s,[s_0]) \geq \epsilon} L(\vec{\theta}(s_0), \vec{\theta}(s)) > 0$$

が成立する．

次に，$s_{n,est}$ の定義と *Theorem 2* により，$\lim_n L(\vec{f_n}, \vec{\theta}(s_{n,est})) = 0$ である．従って，十分大きい N_1 をとると

$$n \geq N_1 \Rightarrow 0 \leq L(\vec{f_n}, \vec{\theta}(s_{n,est})) < \eta_\epsilon / 2 \tag{15}$$

が成立．

さて，$K = \{s \in \mathcal{S} :\ d(s, [s_0]) \geq \epsilon\}$ とおくと，K はコンパクト集合 $\mathcal{S}$ の閉部分集合だから，やはりコンパクト集合である．

従って，*Theorem 2* が使え，

$$\liminf_n \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) \geq \eta_\epsilon$$

が成立している．そこで，適当な N_2 をとれば

$$n \geq N_2 \Rightarrow \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) > \eta_\epsilon / 2 \tag{16}$$

が成立．

そこで $N = \max\{N_1, N_2\}$ ととれば，(15)，(16) により

$$n \geq N \Rightarrow \inf_{s \in K} L(\vec{f}_n, \vec{\theta}(s)) > L(\vec{f}_n, \vec{\theta}(s_{n,est}))$$

これは $n \geq N \Rightarrow d(s_{n,est}, [s_0]) < \epsilon$ を示している．つまり，

$$\lim_{n \to \infty} d(s_{n,est}, [s_0]) = 0$$

が示された．

Q.E.D.

損失関数 L に $\lambda_n > 0,\ \lambda_n \to 0$ として s の連続関数 $R(s)(\geq 0)$ を加えた場合を考えても同様に一致性が示せる．そこで，

$$s_{n,reg} = \operatorname*{argmin}_{s \in \mathcal{S}} \left\{ L(\vec{f}_n, \vec{\theta}(s)) + \lambda_n R(s) \right\}$$

とおくと以下が成立．

Theorem 4 $\vec{\theta}(s)$ をコンパクト集合 $\mathcal{S}$ から $[0,1]^M$ への連続関数とし，確率変数列 $\{\vec{f}_n\}$ が $\vec{\theta}(s_0)$ に as 収束しているとする．また，$L(\vec{\eta}, \vec{\theta})$ を (4) を満たす拡張クロス型の損失関数とせよ．このとき

$$\lim_n d(s_{n,reg}, [s_0]) = 0 \qquad \text{a.s.}$$

が成り立つ．つまり，正則化最小コントラスト推定量の列 $\{s_{n,reg}\}$ は強一致性をもつ．

proof

任意の $\epsilon > 0$ をひとつ固定する．

まず，*Theorem 3* と同じように

$$\eta_\epsilon := \inf_{s:d(s,[s_0])\geq\epsilon} L(\vec{\theta}(s_0), \vec{\theta}(s)) > 0$$

が成立する．

次に

$$F_n(s) = L(\vec{f}_n, \vec{\theta}(s)) + \lambda_n R(s)$$

とおく．$s_{n,reg}$ は $F_n(s)$ の最小化元であり，$R(s_0) < \infty$ であることから

$$\begin{aligned}
0 \leq F_n(s_{n,reg}) &= L(\vec{f}_n, \vec{\theta}(s_{n,reg})) + \lambda_n R(s_{n,reg}) \\
&\leq L(\vec{f}_n, \vec{\theta}(s_0)) + \lambda_n R(s_0) \\
&\to 0, \qquad as\ n \to \infty
\end{aligned}$$

により $\lim_n F_n(s_{n,reg}) = 0$ である．従って，十分大きい N_1 をとると

$$n \geq N_1 \Rightarrow 0 \leq F_n(s_{n,reg}) < \eta_\epsilon/2 \tag{17}$$

が成立．

さて，$K = \{s \in \mathcal{S} :\ d(s,[s_0]) \geq \epsilon\}$ とおくと，*Theorem 3* と同じように適当な N_2 をとれば

$$n \geq N_2 \Rightarrow \inf_{s\in K} L(\vec{f}_n, \vec{\theta}(s)) > \eta_\epsilon/2 \tag{18}$$

が成立．

そこで $N = \max\{N_1, N_2\}$ ととれば，(17)，(18) により $n \geq N$ で

$$\begin{aligned}
\inf_{s\in K} F_n(s) &\geq \inf_{s\in K} L(\vec{f}_n, \vec{\theta}(s)) \\
&> \eta_\epsilon/2 \\
&> F_n(s_{n,reg})
\end{aligned}$$

を得る．これは $n \geq N \Rightarrow d(s_{n,reg}, [s_0]) < \epsilon$ を示している．つまり，

$$\lim_{n\to\infty} d(s_{n,reg}, [s_0]) = 0$$

が示された.

Q.E.D.

3　おわりに

MLE は *KL-divergence* の最小化であり, 拡張クロス型損失関数に関する最小コントラスト推定量になっている. 通常は, 独立観測を仮定して M 推定量の強一致性として証明する. しかし, 上のようにして, *MLE* や正則化 *MLE* の推定量の強一致性を示すこともできる.

本原稿の結果を用いれば, 実験での観測結果に応じて適応的に測定を変える場合や, *KL-divergence* を拡張したような場合でも拡張クロス型損失関数であれば強一致性が示せるだろう. 確認すべき条件は実験データから得られるベクトル (確率変数) $\vec{f}_n$ に対し, $\vec{f}_n \to \vec{\theta}(s_0)$ a.s. のみである.

REFERENCES

[1] *T. Ferguson: A Course in Large Sample Theory. CRC Press, New York, 1996.*

[2] 杉山 太香典: 大規模量子計算機の実用化における統計的課題, 数理解析研究所講究録 *2018 (2017)*, 京都大学, *pp.1-17.*

[3] 吉田 朋広: 講座 数学の考え方〈*21*〉数理統計学. 朝倉書店, *2006.*