2021 年度　数理科学特別講義 C, G
大阪府立大学
8 月 27 日 第 3 コマ

今野良彦

日本女子大学 理学部 数物科学科

2021 年 8 月 27 日

1

内容

(1) M 推定量とは

(2) 有限次元の母数に対するM推定量の一致性

(3) 有限次元の母数に対するM推定量の漸近正規性

M推定量とは

$(\Omega, \mathcal{F}, \mathbb{P})$：確率空間

$X : \Omega \to \mathbb{R}$：確率変数

注意 $(\mathcal{X}, d_{\mathcal{X}})$を確率空間とし、$X : \Omega \to \mathcal{X}$ を確率要素としても形式的な議論はほとんど変わらない。

$P = \mathbb{P} \circ X^{-1}$ とする。すなわち、$P(B) = \mathbb{P}(X \in B)$ $({}^{\forall} B \in \mathcal{B}(\mathbb{R}))$.

$X_1, \ldots, X_n$：X の独立複製

(Θ, d)：距離空間（ここでは、$\Theta \subset \mathbb{R}^p$ とは限らない）

$\|\cdot\|_d = \sqrt{d(\theta, \theta)}$ $(\theta \in \Theta)$

ある $\theta \in \Theta$ に対して

$\rho_\theta : \mathbb{R} \to [0, \infty)$

は損失関数（可測）とする。

すべての $\theta \in \Theta$ に対して、

$$P|\gamma_\theta| = \int_{\mathbb{R}} |\gamma_\theta(x)| \, dP(x) < \infty$$

とする。

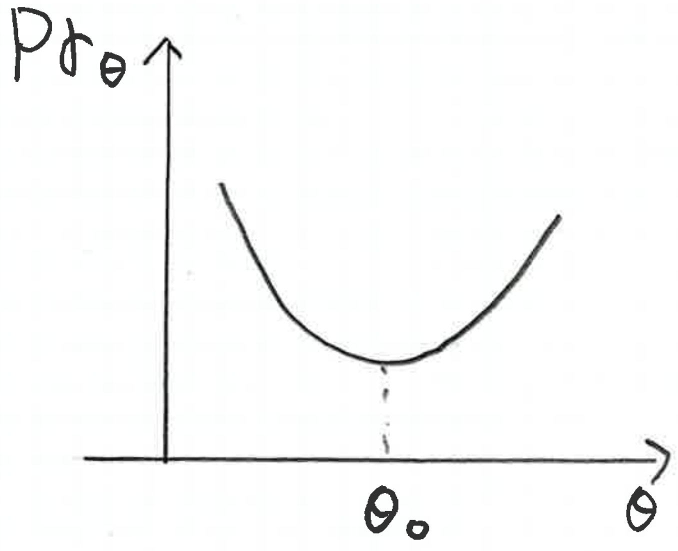
未知の母数を

$$\theta_0 := \underset{\theta \in \Theta}{\operatorname{argmin}} \, P\gamma_\theta$$

とする。すなわち、θ_0 は関数

$$\Theta \ni \theta \longmapsto P\gamma_\theta$$

の最小値を与える点である。



$\Theta = \mathbb{R}$

仮定

θ_0 は一意的に存在する。

θ_0のM推定量

$$\hat{\theta}_n := \underset{\theta\in\Theta}{\operatorname{argmin}}\, P_n \gamma_\theta = \underset{\theta\in\Theta}{\operatorname{argmin}} \left\{ \frac{1}{n}\sum_{i=1}^{n} \gamma_\theta(X_i) \right\}.$$

仮定

$\hat{\theta}_n$は存在する。一意性は保証せず。

例 15.1 $\Theta = \mathbb{R}$とする。

(1a) $\gamma_\theta(x) = (x-\theta)^2$.

(1b) $\gamma_\theta(x) = |x-\theta|$.

(2) パラメトリックモデルの最尤推定量

p_θ : 確率密度函数 $\theta \in \Theta = \mathbb{R}$.

$\gamma_\theta = -\log p_\theta$ 正確には $\gamma_\theta = -(\log p_\theta)\mathbb{I}\{p_\theta > 0\}$.

M推定量の一致性

記号　$\theta \in \Theta$ に対して

$$R(\theta) := P r_\theta = \int_{\mathbb{R}} r_\theta(x)\, dP(x).$$

$$R_n(\theta) := P_n r_\theta = \frac{1}{n} \sum_{i=1}^{n} r_\theta(X_i).$$

定義 15.1　θ_0 はうまく分離されているとは

$\forall \eta > 0$ に対して，

$$\inf \{ R(\theta) ;\ d(\theta, \theta_0) > \eta \} > R(\theta_0).$$

定理 15.1　函数族 $\{\gamma_\theta ; \theta \in \Theta\}$ は GC 族と仮定する。
すなわち、

$$\sup_{\theta \in \Theta} |R(\theta) - R_n(\theta)| \xrightarrow{P} 0 \quad (n \to \infty) \qquad (a)$$

と仮定している。このとき、

$$R(\hat{\theta}_n) \xrightarrow{P} R(\theta_0) \qquad (n \to \infty).$$

さらに、θ_0 がうまく分離されていれば、

$$\hat{\theta}_n \xrightarrow{P} \theta_0 \quad (n \to \infty).$$

注意　$\hat{\theta}_n \xrightarrow{P} \theta_0 \ (n \to \infty) \iff d(\hat{\theta}_n, \theta_0) \xrightarrow{P} 0 \ (n \to \infty)$

証明

$$0 \leqq R(\hat{\theta}_n) - R(\theta_0) \qquad (\because\ \theta_0 \text{の定義})$$

$$= R(\hat{\theta}_n) - R(\theta_0) - \{R_n(\hat{\theta}_n) - R_n(\theta_0)\}$$

$$+ \{R_n(\hat{\theta}_n) - R_n(\theta_0)\}$$

$$\leqq R(\hat{\theta}_n) - R_n(\hat{\theta}_n) + R_n(\theta_0) - R(\theta_0)$$

$$(\because\ R_n(\hat{\theta}_n) - R_n(\theta_0) \leqq 0)$$

$$\leqq |R(\hat{\theta}_n) - R_n(\hat{\theta}_n)| + |R(\theta_0) - R_n(\theta_0)|$$

$$\leqq 2 \sup_{\theta \in \Theta} |R(\theta) - R_n(\theta)| \xrightarrow{P} 0 \quad (n \to \infty).$$

よって、 $|R(\hat{\theta}_n) - R(\theta_0)| \xrightarrow{P} 0 \quad (n \to \infty).$

No. 7

次に，θ_0 はうまく分離されているとする．

このとき，$^\forall \varepsilon > 0$ に対して，$\exists \eta > 0$ があって，$^\forall \theta \in \Theta$ に対して，

$$d(\theta, \theta_0) > \eta \Longrightarrow R(\theta) - R(\theta_0) > \varepsilon .$$

これの対偶を取れば；

$$R(\theta) - R(\theta_0) \leqq \varepsilon \Longrightarrow d(\theta, \theta_0) \leqq \eta$$

となるので，

$$\mathbb{P}\left(d(\hat{\theta}_n, \theta_0) \leqq \eta \right) \geq \mathbb{P}\left(R(\hat{\theta}_n) - R(\theta_0) \leqq \varepsilon \right)$$

$$\longrightarrow 1 \quad (n \to \infty).$$

よって

$$\mathbb{P}(d(\hat{\theta}_n, \theta_0) > \eta) \longrightarrow 0 \quad (n \longrightarrow \infty).$$

注意 15.1 定理15.1の条件は、ほぼ Θ がコンパクトであることを要求している。

以下の仮定を考えてみる。

(1) Θ はノルム空間(ノルムを $\|\cdot\|_d$ とかく)の凸集合。

$0 \leq {}^\forall \alpha \leq 1$ と ${}^\forall \theta_1, \theta_2 \in \Theta$ に対して

$$\alpha \theta_1 + (1-\alpha)\theta_2 \in \Theta$$

(2) 函数 $\Theta \ni \theta \longmapsto \delta_\theta \in [0, \infty)$ は狭義凸。

すなわち、$0 < {}^\forall \alpha < 1$, ${}^\forall \theta_1, \theta_2 \in \Theta$ に対して

$$\delta_{\alpha\theta_1 + (1-\alpha)\theta_2}(x) < \alpha \delta_{\theta_1}(x) + (1-\alpha)\delta_{\theta_2}(x) \quad (x \in \mathbb{R})$$

この仮定より、$\Theta \ni \theta \longmapsto \delta_\theta \in [0, \infty)$ は連続。

(3) ある $\varepsilon > 0$ があって、・・

$$\Theta_\varepsilon := \{ \theta \in \Theta ;\ \| \theta - \theta_0 \|_d \leq \varepsilon \}$$

(4) $G_\varepsilon(x) := \sup_{\theta \in \Theta} | \gamma_\theta(x) |$ としたとき、

$$P G_\varepsilon = \int_{\mathbb{R}} G_\varepsilon(x)\, dP(x) < \infty .$$

補題 15.2 仮定 (1) ~ (4) のもと

$$\hat{\theta}_n \xrightarrow{P} \theta_0 \quad (n \to \infty) .$$

<u>証明</u>：(3) と (4) より，

$$\sup_{\theta \in \Theta_\varepsilon} |R_n(\theta) - R(\theta)| \xrightarrow{P} 0 \quad (n \to \infty)$$

となることがわかる。

いま，

$$\tilde{\theta}_n := \alpha \hat{\theta}_n + (1-\alpha)\theta_0$$

$$\alpha := \frac{\varepsilon}{\varepsilon + \|\hat{\theta}_n - \theta_0\|_d}$$

とおく。

すると

$$\|\hat{\theta} - \theta_0\| \leq \varepsilon \Longrightarrow \hat{\theta} \in \Theta_\varepsilon$$

となる。すると、

$$0 \leq R_n(\tilde{\theta}_n) - R(\theta_0)$$

$$\leq 2 \sup_{\theta \in \Theta_\varepsilon} |R_n(\theta) - R(\theta)|$$

（定理15.1の証明中の議論と同じようにする。）

を示すことができる。さらに、写数 $\theta \mapsto \sigma_\theta$ の凸狭義性より分離性が保証されるので、

$$\tilde{\theta}_n \xrightarrow{P} \theta_0 \quad (n \to \infty)$$

さらに、

$\|\hat{\theta}_n - \theta_0\|_d = \alpha \|\tilde{\theta}_n - \theta_0\|_d$ より、$\hat{\theta}_n \xrightarrow{P} \theta_0 \ (n \to \infty)$。 □

例 15.2： $\Theta = \mathbb{R}$ とし、$r \geq 1$ に対し、 No.12

$$\gamma_\theta(x) = |x - \theta|^r$$

とおく。さらに、

$$P\gamma_\theta = \int_{\mathbb{R}} |x-\theta|^r \, dP(x) < \infty$$

と仮定する。

$$\theta_0 = \underset{\theta \in \Theta}{\operatorname{argmin}} \; P\gamma_\theta$$

は一意的に存在すると仮定する。
すると、$\forall \varepsilon > 0$ に対して、

$$\sup_{|\theta - \theta_0| \leq \varepsilon} |\gamma_\theta(x)| \leq 2^{r-1} \{ |x-\theta_0|^r + \varepsilon^r \} =: G_\varepsilon(x)$$

すると

$$P G_\varepsilon = 2^{r-1}\{\mathbb{E}[|X-\theta_0|^r] + \varepsilon^r\} < \infty$$

となるので,

$$G_\varepsilon \in L^1(P).$$

よって, 補題 15.2 より,

$$\hat{\theta}_n \xrightarrow{P} \theta_0 \quad (n \to \infty).$$

M推定量の漸近正規性

以下では、$p \in \mathbb{N}$ とし、

$$\Theta \subset \mathbb{R}^p, \quad \theta_0 \text{ は } \Theta \text{ の内点}$$

とする。Euclid ノルムを $|\cdot|$ と書く。

定義15.2 θ_0 の推定量 $\{\hat{\theta}_n\}_{n\geq 1}$ は漸近線形であるとは、以下が成立することである。

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \sqrt{n}\, P_n \ell + o_p(1)$$

ただし、

$$\ell^T(x) = (\ell_1(x), \ldots, \ell_p(x)) \quad (x \in \mathbb{R})$$

$$P\ell = (P\ell_1, \ldots, P\ell_p) = (0, \ldots, 0).$$

条件(d)

$$\mathbb{E}\left[X^2 \, \mathbb{I}\{|X-\theta| \leq k\}\right] < \infty$$

でよい。

補題15.3より、

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{L} N\left(0, \frac{J}{V^2}\right)$$

$$V = 2\{F(k+\theta_0) - F(-k+\theta_0)\}$$

$$J = \mathbb{E}[\psi^2_{\theta_0}(X)]$$

である。

仮定

(a) ある $\varepsilon>0$ が存在して, $|\theta-\theta_0|<\varepsilon$ なる θ において, 関数 $\theta \mapsto r_\theta(x)$ は微分可能で

$$\psi_\theta(x) := \frac{\partial}{\partial\theta} r_\theta(x) = \left(\frac{\partial}{\partial\theta_1} r_\theta(x), \ldots, \frac{\partial}{\partial\theta_p} \partial r_\theta(x)\right)^T$$

となる。上のことは $x \in \mathbb{R}$ に対して成立である。

ほとんどいたるところで

(b)

$$\lim_{\theta\to\theta_0} P(\psi_\theta-\psi_{\theta_0}) = V(\theta-\theta_0) + o(1)|\theta-\theta_0|.$$

ただし, V は $p\times p$ の正定値行列である。

(c) $\mathcal{G} := \{\psi_\theta ; |\theta - \theta_0| < \varepsilon\}$ とおいたとき

$\mathcal{G}$ はドンスカー族. $\Rightarrow$ $\psi_{\theta_0} \in L^2(P)$ を仮定することになる.

(d)

$$\lim_{\theta \to \theta_0} \sqrt{P|\psi_\theta - \psi_{\theta_0}|^2} = 0.$$

補題 15.3　条件 (a)〜(d) のもと, $\hat{\theta}_n$ は漸近線形で.

$$\ell(x) = -V^{-1}\psi_{\theta_0}(x).$$

したがって.

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{L} N_p(0, V^{-1}JV^{-1})$$

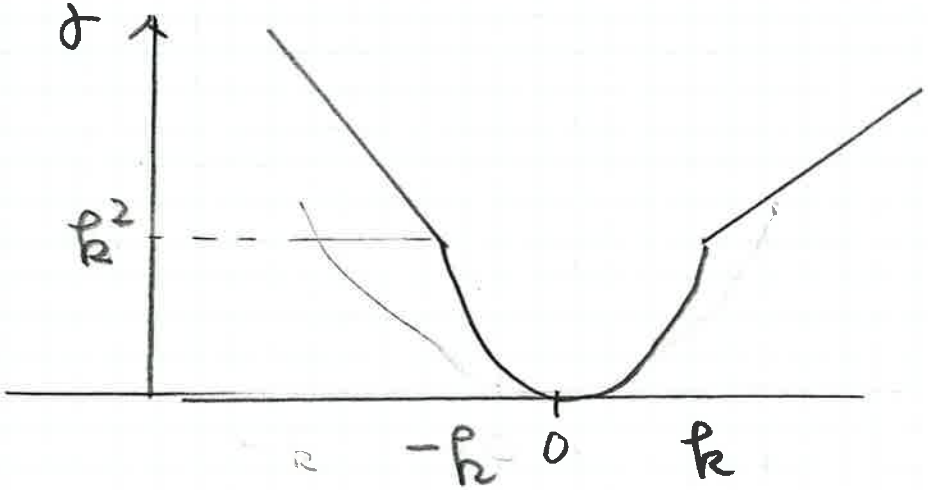
$$J := P[\psi_{\theta_0}\psi_{\theta_0}^T] : p \times p.$$

例 15.4 $\Theta = \mathbb{R}$とする。 $0 < k < \infty$を固定し、

$$\gamma_\theta(x) = \gamma(x-\theta)$$

$$\gamma(x) := x^2 \mathbb{I}\{|x| \le k\} + (2k|x| - k^2)\mathbb{I}\{|x| > k\}$$

$(x \in \mathbb{R})$



$$\hat{\theta}_n = \operatorname*{argmin}_{\theta \in \Theta} P_n r_\theta$$

を Huber 型推定量という.

条件 (a)

$$\psi_\theta(x) = \begin{cases} 2k & (x-\theta < -k) \\ -2(x-\theta) & (-k \leq x-\theta \leq k) \\ -2k & (x-\theta > k) \end{cases}$$

条件 (b)

$$V = 2\{F(k+\theta_0) - F(-k+\theta_0)\}.$$

条件 (c) $\{\psi_\theta ; \theta \in \Theta\}$ は VC 族 かつ

$$\sup_{x \in \mathbb{R}} |\psi_\theta(x)| \leq 2k.$$

例 15.5　$f(x) = \frac{d}{dx}F(x)$ とおく。

No. 21

θ_0 は中央値

$\hat{\theta}_n$ は中央値推定量

$$\sqrt{n}\,(\hat{\theta}_n - \theta_0) \xrightarrow{L} N\left(0,\ \frac{1}{4\{f(\theta_0)\}^2}\right)$$