

最適領域の推定に関する

統計的構造

統計数理研究所

野田 一雄

多賀 保志

§ 1. 最適領域の決定問題

(X, Ω) を標本空間とし、その上の確率測度のある族 P_1 が与えられているとする。

X 上の Ω -可測な実関数のベクトル $\phi = (\phi_1, \phi_2, \dots, \phi_l)$;

$$(1.1) \quad \sum_{j=1}^l \phi_j(x) = 1, \quad \phi_j(x) \geq 0 \quad (j=1, 2, \dots, l)$$

を X の l -分割とよぶ。これらの全体を Ω で表わすことに

しよう。そのうち制約条件 Ω を満たすものの全体を $\Omega(\Omega)$

で Ω で表わして、これが空でないことを仮定しておく。また

制約条件 Ω を置かないときは、 Ω を直接の対象と考えればよい。

いま、 $\forall p \in P_1$ について可積分であるような Ω -可測な実関数 $g_i(x)$ をとり、 p に関して (g_i) と ϕ の内積の関数となるようなある目的関数 $v(\phi, p)$ を考える。 $\Omega(\Omega)$ の中に

この目的関数 $v(\phi, p)$ の最大値もしくは最小値を実現するものが存在するとき、これをその最適分割とよび、その一つを ϕ_p で表わすことにする。ここでは ϕ_p は $v(\phi, p)$ の最大値を実現するものとしよう。

このとき、ある $\phi \in \mathfrak{A}(\mathcal{C})$ を選んだときの損失を

$$(1.2) \quad L(\phi, p) = v(\phi_p, p) - v(\phi, p) \geq 0$$

とする。

一般に $p \in \mathcal{P}_1$ は未知であるから、 ϕ を決定する際に事前情報 $\mathfrak{B}$ をとることを考える。 $\mathfrak{B}$ が値をとる標本空間を $(\mathcal{S}, \mathfrak{B})$ とし、その上の確率測度の族 $\mathcal{Q}_1 = \{Q_p : p \in \mathcal{P}_1\}$ が与えられているとする。つまり、 X が $p \in \mathcal{P}_1$ に従うとき、 $\mathfrak{B}$ はこれとは独立に $Q_p \in \mathcal{Q}_1$ に従うと考える。

かくして、写像 $\hat{\phi} : \mathcal{S} \rightarrow \mathfrak{A}(\mathcal{C})$ を決定関数と考えるのであるが、これにはある種の可測性を仮定することによって、 $\mathcal{X} \times \mathcal{S}$ 上の $\mathcal{Q} \times \mathfrak{B}$ -可測な実関数のベクトル $\hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_l)$;

$$(1.3) \quad \sum_{j=1}^l \hat{\phi}_j(x, s) = 1, \quad \hat{\phi}_j(x, s) \geq 0 \quad (j=1, 2, \dots, l)$$

をもつて表わすことが出来る ([3])。もちろん $\mathfrak{A}(\mathcal{C})$ は、

この決定の行動空間とみなされるわけである。

そこで、この場合、危険関数を

$$(1.4) \quad r(\hat{\phi}, p) = \mathbb{E}_{Q_p} [L(\hat{\phi}, p)]$$

$$\hat{\phi} \in \mathfrak{A}(\mathcal{C}), \quad p \in \mathcal{P}_1$$

と考える。(1.3) を満足する $\hat{\phi}$ の全体を $\hat{\Phi}$, そのうち制約条件 $\mathcal{C}$ を満足する部分空間を $\hat{\Phi}(\mathcal{C})$ で表す。直接には, $\hat{\Phi}(\mathcal{C})$ が決定関数の空間となるわけである。

この場合, ϕ_p のミニマックス推定量 $\hat{\phi}$ は,

$$(1.5) \quad \inf_{\hat{\phi} \in \hat{\Phi}(\mathcal{C})} \sup_{p \in \mathcal{P}_1} r(\hat{\phi}, p)$$

を実現するもので, これが当面の目標となるであろう。

このミニマックス $\hat{\phi}$ を求めるために, Bayes 推定量との関係を論じるには, 次のような先験測度 ξ のある集合 Ξ_1 を考えていく。すなわち, $\mathcal{P}_1$ の部分集合から作られる σ -代数 $\mathcal{C}_1$ を適当にとって, $(\mathcal{P}_1, \mathcal{C}_1)$ 上の確率測度 ξ のある族 Ξ_1 を選ぶ。

この場合, Bayes 危険は,

$$(1.6) \quad r(\hat{\phi}, \xi) = E_{\xi} [r(\hat{\phi}, p)] \\ \hat{\phi} \in \hat{\Phi}(\mathcal{C}), \xi \in \Xi_1$$

となり, その $\inf$ を実現する $\hat{\phi} \in \hat{\Phi}(\mathcal{C})$ を ϕ_p の Bayes 推定量とよび $\hat{\phi}_{\xi}$ で表すことにしよう。

一般には, 先験確率測度 ξ は, $(\mathcal{P}_1 \times \mathcal{Q}_1, \mathcal{C}_1 \times \mathcal{D}_1)$ の上の測度と考えるべきであるが, この場合 p と Q_p との独立性を仮定していいので, 上記のように表現してよい。

さて, 以上の設定の下で, 次のようなことが問題となる。

くる。

[1°] 一般的な条件の下で, 最適分割 ϕ_p の存在およびその explicit form を求めること。

[2°] 一般的な条件の下で, Bayes 推定量 $\hat{\phi}_{\xi}$ の存在およびその explicit form を求めること。

[3°] 一般的な条件の下で, ミニマックス推定量の存在およびその explicit form を求めること。

特に適当に $\{\xi_i : i=1, 2, \dots\} \subset \Xi_1$ を選出することによ

$$(1.7) \quad \sup_{p \in \mathcal{P}_1} r(\tilde{\phi}, p) = \lim_{i \rightarrow \infty} r(\hat{\phi}_{\xi_i}, \xi_i)$$

なる関係を求めること。

[4°] 有限個ないしは可算個の $\mathcal{P}_0 = \{p_i\} \subset \mathcal{P}_1$ を適当にとることにより, 族 $\mathcal{P}_0$ についての Bayes 推定量ならびにミニマックス推定量が $\hat{\phi}_{\xi}$, $\tilde{\phi}$ の近似になつていふようになること。

特に $\mathcal{P}_1$ に適当な距離を導入することにより, その ε -被覆を上述の問題に適用すること。

[5°] $\tilde{\phi}$ は (1.7) の関係から許容性をもつことになるが, 更に $\hat{\phi}_{\xi}$, $\tilde{\phi}$ の consistency, すなわち, $\mathcal{S}$ のサンプルサイズ n を増大させるとき, 危険関数が 0 に収束せしめるような性質をもつことを示すこと。

以上の問題が [2], [3] において、次のような定式化の下で議論された。

ν を $(X, \mathcal{O})$ 上のある与えられた σ -有限な測度とする。
 $f_i(x)$ を $\mathcal{O}$ -可測で、 ν および $\forall p \in \mathcal{P}_1$ について可積分な実関数、 $g_{ij}(x)$ を $\forall p \in \mathcal{P}_1$ について可積分な実関数とする。

このとき、制約条件を、 C_{ij} をある定数として、

$$(1.8) \quad \tau_{ij}(\phi) = \int_X f_i(x) \phi_j(x) \nu(dx) \leq C_{ij} \quad (\alpha = C_{ij})$$

$$(i=1, 2, \dots, m; j=1, 2, \dots, l)$$

によって定める。

次に目的関数 v は、まず

$$(1.9) \quad \psi_{i'j}(\phi, p) = \int_X g_{i'}(x) \phi_j(x) p(dx)$$

$$(i'=1, 2, \dots, k; j=1, 2, \dots, l)$$

と置く。 $k \times l$ -実数値行列空間のある凸部分集合 Z をとることにより、 $\psi_{ij}(\phi, p)$ の値がすべてこの $\phi \in \hat{\mathcal{C}}(\mathcal{C})$, $p \in \mathcal{P}_1$ について Z の中に実現されることを仮定しておく。このとき Z 上の実関数 $h(z)$ をとり、

$$(1.10) \quad v(\phi, p) = h_0[\psi_{ij}(\phi, p)]$$

と定める。

かくして、 $\mathcal{O}$, $\mathcal{P}$, および $\mathcal{P}_1$, $\mathcal{Q}_1$ および h に一般的な条件を置くことによって、[1°] ~ [5°] の議論が展開される。

さて、ここでは、[5°] に関して、[3] でとられた方法に

依らず、ミニマックス $\hat{\phi}$ の収束の速さを求めることを考えよう。

§2. ミニマックス推定量 $\hat{\phi}_n$ の収束の速さ

事前情報 $\mathcal{S}_n$ のサンプル・サイズを n とする。すなわち、
 $(\mathcal{S}_n, \mathcal{B}, \mathcal{Q}_1) = (\mathcal{X}^n, \mathcal{Q}_1^n, \mathcal{P}_1^n)$ なる積空間において考
 える。

また、簡便のため、ここでは、制約条件を

$$(2.1) \quad \int_{\mathcal{X}} f(x) \phi(x) \nu(dx) \leq C \quad (C = C)$$

とし、目的関数を

$$(2.2) \quad v(\phi, p) = \int_{\mathcal{X}} g(x) \phi(x) p(dx)$$

と置く。ここでは、 ϕ は、 $0 \leq \phi(x) \leq 1$ なる $\mathcal{Q}_1$ -可測関数の全体とする。

また損失は、

$$(2.3) \quad L(\phi, p) = \int_{\mathcal{X}} g(\phi_p - \phi) p(dx)$$

となり、危険関数は、

$$(2.4) \quad r(\hat{\phi}, p) = \int_{\mathcal{X}^n} \int_{\mathcal{X}} g(\phi_p(x) - \hat{\phi}(x, s)) p(dx) p^n(ds)$$

となる。

さて、 $\mathcal{S}_n$ をもとにした経験分布を $\hat{p}_n(\cdot | s)$ で表わすことに
 する。

まず幾つかの lemma を用意しよう。

Lemma 2.1 (Dudley [1] の Proposition の拡張)

$g(x)$ を $\forall p \in \mathcal{P}_1$ について自乗可積分であるとす。そして

$$(2.5) \quad G_p(A; g) = \int_A g(x) p(dx), \quad A \in \mathcal{O}, \quad p \in \mathcal{P}_1.$$

と記すことにしよう。

このとき, $A \in \mathcal{O}$ の任意の有限個の分割

$$\{A_j : j=1, 2, \dots, m\}$$

$$A = \bigcup_{j=1}^m A_j, \quad A_i \cap A_j = \emptyset \quad \text{for } i \neq j$$

について,

$$(2.6) \quad E_{p^n} \sum_{j=1}^m \left| \int_{A_j} g(x) \{ \hat{p}_n(dx|s) - p(dx) \} \right| \\ \leq \left[\frac{m}{n} G_p(A; g^2) \right]^{\frac{1}{2}}$$

さて, $(X, \mathcal{O})$ 上の確率測度で, $f(x)$ を可積分, $g(x)$ を自乗可積分とするようなものの全体を $\mathcal{P}_2$ ($\supset \mathcal{P}_1$) としよう。

このとき, $\mathcal{P}_2$ に次のような距離 ρ を導入する。

$$(2.7) \quad \rho(p_1, p_2) = \sup_{g \in \mathcal{F}} \left| \int_X g(x) \phi(x) (p_1(dx) - p_2(dx)) \right|, \\ p_1, p_2 \in \mathcal{P}_2.$$

ここで, $E_{p^n} \rho(\hat{p}_n, p)$ を評価したいのであるが, そのために次のような準備をする。

X を separable metric space とし, $\mathcal{O}$ をその Borel-field としておく。

このとき, $\forall \varepsilon > 0$ に対して, 直至 2ε 以下なる被覆 $\{U_j\}$

をとり、これが高々可算集合であるようにあることが出来る。

しかも、 $\eta(\varepsilon) > 0$ に対し、

$$(2.8) \quad G_p(W(\varepsilon, p); |g|) \leq \eta(\varepsilon)$$

なる $W(\varepsilon, p) \in \Omega$ を除けば、 $x \sim W(\varepsilon, p)$ は有限個の $\{U_i\}$ で
おおわれっているわけである。そこでその minimal number
を $N(p, \varepsilon, \eta(\varepsilon))$ で表わすことにしよう。

Lemma 2.2

$k > 2$ なる実数に対し、 $0 < K(p) < \infty$ があって、任意の
 $0 < \varepsilon \leq 1$ について

$$(2.9) \quad N(p, \varepsilon, \varepsilon^{k/(k-2)}) \leq K(p) \varepsilon^{-k}$$

が満たされているものとする。

このとき $0 < M(k, K(p), G_p(x; g^2)) < \infty$ があって

$$(2.10) \quad E_{p_n} P(\hat{P}_n(\cdot | s), p) \leq M n^{-\frac{1}{k}}, \quad \text{for } p \in \mathcal{P}_2$$

が成立する。

次に、 $s \in S_n = \mathcal{X}^n$ を固定するとき、 $p_0 \in \mathcal{P}_1$ と適当に選
んで

$$(2.11) \quad \mathcal{P}_0(s) \equiv \{p: p \in \mathcal{P}_1, P(\hat{P}_n(\cdot | s), p) \leq P(\hat{P}_n(\cdot | s), p_0)\}$$

を定める。このとき、 $\mathcal{P}_0(s)$ は $\mathcal{P}_1$ の有界集合となるが、
これが全有界であるように p_0 がとれることを仮定しよう。

この仮定のちとで、 $\mathcal{P}_0(s)$ について、 $\forall \varepsilon > 0$ に対し、直交 2ε

以下の有限個の被覆 $\{\mathcal{P}_{\varepsilon,i} : i=1,2,\dots,m(\varepsilon,s)\}$ をとる。各 $\mathcal{P}_{\varepsilon,i}$ から代表元 $p_{\varepsilon,i} \in \mathcal{P}_{\varepsilon,i}$ を選んで固定する。このうち、 $\rho(p_{\varepsilon,i}, \tilde{p}_n(\cdot|s))$ の最小値を実現するものを $\rightarrow$ 固定して $\tilde{p}_n(\cdot|s)$ と記すことにしよう。この $\tilde{p}_n(\cdot|s)$ については $\mathcal{B}$ -可測性をいっておく必要があるが、ここでは触れないでおく。

Lemma 2.3

Lemma 2.2 の仮定を仮定し、 $\mathcal{P}_0(s)$ が全有界であるように $\mathcal{P}_0 \in \mathcal{P}_1$ がとれることを仮定すれば

$$(2.12) \quad E_{p_n} \rho(\tilde{p}_n(\cdot|s), p) \leq 4M(k, p) n^{-\frac{1}{k}} \\ \text{for } p \in \mathcal{P}_1.$$

が成立する。ただし、 $k, M(k, p)$ は Lemma 2.2 で与えられるものである。

次に、制約条件 (2.1) のもとで、 $s \in \mathcal{S}_n$ を固定するとき

$$(2.13) \quad v(\hat{\phi}, p) = \int_{\mathcal{X}} g(x) \hat{\phi}(x, s) \tilde{p}_n(dx|s)$$

の $\sup$ を実現する $\hat{\phi} \in \hat{\mathcal{C}}(\mathcal{C})$ を $\hat{\phi}_{\tilde{p}_n}(x, s)$ と記すことにしよう。

Lemma 2.4

$s \in \mathcal{S}_n$ を固定して考える。 $\nu \ll P_\nu$ とするとき、 $(\mathcal{X}, \mathcal{O}_2)$ 上の確率測度 $\tilde{p}_n^*$ を

$$(2.14) \quad \tilde{p}_n^*(dx|s) = \frac{1}{2} [P_\nu(dx) + \tilde{p}_n(dx|s)].$$

でもって定義する。このとき、 $\hat{p}_n$, ν の $\hat{p}_n^*$ に関する一般化した密度を $\frac{d\hat{p}_n}{d\hat{p}_n^*}(x, s)$, $\frac{d\nu}{d\hat{p}_n^*}(x, s)$ とし、これらが $\mathcal{X}$ の $\mathcal{B}$ -可測であることを仮定しておく。

このとき、(2.13) の $\sup$ を実現する $\hat{\phi}_{\hat{p}_n}(x, s) \in \hat{\Phi}$ が存在し、またこれは、 $\hat{p}_n^*$ -測度 0 を除外して、

$$(2.15) \quad \hat{\phi}_{\hat{p}_n}(x, s) = \begin{cases} 1 & \text{for } g(x) \frac{d\hat{p}_n}{d\hat{p}_n^*}(x, s) - \hat{\lambda}_n(s) f(x) \\ & \times \frac{d\nu}{d\hat{p}_n^*}(x, s) > 0 \\ 0 & \text{for } \quad \quad \quad < 0 \end{cases}$$

となることが必要十分である。ただし、 $\hat{\lambda}_n(s)$ は、 $\mathcal{B}$ -可測関数で (2.15) を満たす $\hat{\phi}_{\hat{p}_n}(x, s)$ が制約条件 (2.1) を満足するように定められる。

Lemma 2.5

Lemma 2.3, 2.4 の仮定のもとで

$$(2.16) \quad r(\hat{\phi}_{\hat{p}_n}, p) \leq 7M(k, p)n^{-\frac{1}{k}}, \quad p \in \mathcal{P}_1$$

が成立する。

以上の準備のもとに、ミニマックス推定量 $\hat{p}_n$ の収束の速さは次のように示される。

Theorem 2.6

次のような仮定を置く。

1) Lemma 2.2 の仮定。特に $K(p)$, $G_p(x; g^2)$ が $p \in \mathcal{P}_1$ について有界であること。したがって

$$(2.17) \quad M(k, p) \leq \bar{M}(k) < \infty \quad \text{for all } p \in \mathcal{P}_1$$

なる $\bar{M}(k)$ の存在を仮定しておく。

2) 距離 ρ によって $\mathcal{P}_0(s)$ が全有界となるように $\rho \in \mathcal{P}_1$ が選べる。

3) Lemma 2.4 の $\frac{d\tilde{P}_n}{dP_n^*}$, $\frac{d\nu}{dP_n^*}$ の $\mathcal{O}_2 \times \mathcal{B}$ -可測性を仮定する。

以上の仮定のもとでミニマックス推定量 $\hat{\phi}_n(x, s)$ について

$$(2.18) \quad r(\hat{\phi}_n, p) \leq 4\bar{M}(k)n^{-\frac{1}{k}} \quad \text{for all } p \in \mathcal{P}_1.$$

が成立する。

(証明). 以上の lemma により.

$$(2.19) \quad \sup_{p \in \mathcal{P}_1} r(\tilde{\phi}_n, p) = \inf_{\hat{\phi} \in \hat{\Phi}(\mathcal{C})} \sup_{p \in \mathcal{P}_1} r(\hat{\phi}_n, p)$$

$$\leq \sup_{p \in \mathcal{P}_1} r(\hat{\phi}_{\tilde{P}_n}^*, p) \leq 4\bar{M}(k)n^{-\frac{1}{k}}$$

したがって (2.18) の成立を見る。

Proposition 2.7 (Dudley の proposition の拡張)

$\mathcal{X}$ を p 次元ユークリッド空間, $\mathcal{O}$ をその Borel-field とする。

このとき,

$$(2.20) \quad d = pk / (k-p)(k-2), \quad k > p, \quad k > 2.$$

なる d について

$$(2.21) \quad G_p(x; |x|^d |g(x)|) < \infty$$

が成立すれば Lemma 2.2 の仮定が満足される。

例1

<1> (x, σ) をユークリッド空間とし、その上の確率測度の族 $\mathcal{P}_1 = \{P(\mu, \sigma^2) : -k_1 \leq \mu \leq k_2, \delta \leq \sigma \leq k_3\}$ を $P(\mu, \sigma^2)$ が正規分布 $N(\mu, \sigma^2)$ であるようにとっておく。ただし、 k_1, k_2, k_3, δ は適当な正数である。

このとき、Theorem 2.6 の条件 2), 3) が満足されることは明白であろう。

条件 1) については、Proposition 2.7 の適用を考えればよい。 $g(x)$ が簡単に x などであれば Proposition 2.7 の仮定を満足していることは見やすい。

$G_p(x; g^2)$, $L(p)$ は、例えば $k=3$ としおくと、それぞれ $\mu^2 + \sigma^2$, $\int_{\mathcal{P}_1} |x|^{3/2} P(\mu, \sigma^2)(dx)$ となるが、いずれも $-k_1 \leq \mu \leq k_2, \delta \leq \sigma \leq k_3$ で有界(連続)である。したがって Theorem 2.6 が、この場合 成立する。

$g(x)$ が連続関数の場合は、適当な区間で多項式近似を考えればよい。

<2>

特にいま, C を正数として, 制約条件を

$$(2.22) \quad \int_{\mathcal{X}} \phi(x) \nu(dx) \leq C$$

とし, 目的関数を

$$(2.23) \quad \int_{\mathcal{X}} x \phi(x) p_{(\mu, 1^2)}(dx), \quad -K_1 \leq \mu \leq K_2$$

としよう。ただし, ν はルベック測度である。

このとき, 最適領域 ϕ_p は,

$$(2.24) \quad \phi_p(x) = \begin{cases} 1 & \text{for } x p_{(\mu, 1^2)}(x) - \lambda_\mu > 0, \\ 0 & \text{for } \quad \quad \quad < 0 \end{cases}$$

で与えられる。ただし λ_μ は (2.22) を満足するような正の数である。また $p_{(\mu, 1^2)}$ は $P_{(\mu, 1^2)}$ の密度

$$p_{(\mu, 1^2)} = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2}\right]$$

である。

このミニマックス推定量 $\tilde{\phi}_n$ は,

$$(2.25) \quad \tilde{\phi}_n(x, s) = \begin{cases} 1 & \text{for } x p_n(x, s) - \tilde{\lambda}_n(s) > 0, \\ 0 & \text{for } \quad \quad \quad < 0 \end{cases}$$

$$p_n(x, s) = \frac{\sqrt{n}}{\sqrt{2\pi} \sqrt{n+1}} \exp\left[-\frac{(x-\bar{x})^2}{2 \cdot \frac{n+1}{n}}\right],$$

$$\bar{x} = \sum_{i=1}^n x_i / n, \quad s = (x_1, x_2, \dots, x_n)$$

として与えられる。ただし $\tilde{\lambda}_n(s)$ は 正值可測関数で, 制約条件 (2.22) を満足するように定められる。

文 献

- [1] Dudley, R.M.: "The speed of mean Glivenko - Canteli convergence", A.M.S., Vol. 40, No. 1, 40~50, (1969)
- [2] Noda, K & Taga, Y.; "Bayes and minimax estimation methods for the optimum decomposition of a sample space based on prior information", Review of I.S.I., Vol. 38: 3, 324~350 (1970)
- [3] Noda, K & Taga, Y.; "Minimax estimation methods for the optimum decomposition of a sample space based on prior information"
(A.I.S.M. Vol. 23. No. 1 に発表予定 (1971)).