# 正則条件下における最尤推定量の一致性

渡辺澄夫
東京工業大学

# 問題

実数に値をとる確率変数 X が確率密度関数 $p(x|w_0)$ を持つとする。ここで $w_0$ はある実数。

確率変数 $X_1, X_2, \cdots, X_n$ は独立で X と同じ確率密度を持つとする。実数 w の関数(対数尤度関数を (−n) で割ったもの)

$$L(w) = -(1/n) \sum_{i=1}^{n} \log p(X_i|w)$$

を最小にする w を $w^*$ と書いて最尤推定量という。
どのような条件のもとで

「$\forall \varepsilon > 0, \quad \lim_{n\to\infty} P(|w^* - w_0| > \varepsilon) = 0$」が成り立つか。

この命題が成り立つとき $w^*$ は $w_0$ に確率収束するといい、
「最尤推定量は一致性を持つ」という。

# 注意

◎　関数 $L(w)$ の平均値は

$$E[L(w)] = -\int p(x|w_0) \log p(x|w)\, dx$$

$$= \int p(x|w_0) \log \{ p(x|w_0)/p(x|w) \}\, dx$$

$$- \int p(x|w_0) \log p(x|w_0)\, dx$$

であるから $E[L(w)]$は $w=w_0$ で最小値をとる。

◎　大数の法則から $n \to \infty$のとき、各点 $w$ ごとに

$L(w) \to E[L(w)]$ が成り立つ。

◎　$w_0=0$, $E[L(0)]=0$ とおいても一般性を失わない。

# 定理

条件　$x, w \in \mathbf{R}$, $f : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$

X は確率変数（密度関数 $q(x)$を持つとする）。以下を仮定。

(1) $F(w) = E[\ f(X,w)\ ]$ は $F(0)=0$ で $F(|w|)$ は $|w|$ の増加関数。

次の条件を満たす $M>0$ が存在する。

(2) $g(x) \equiv \inf_{|w|>M} f(x,w)$ とおくと $E[g(X)] = C > 0$

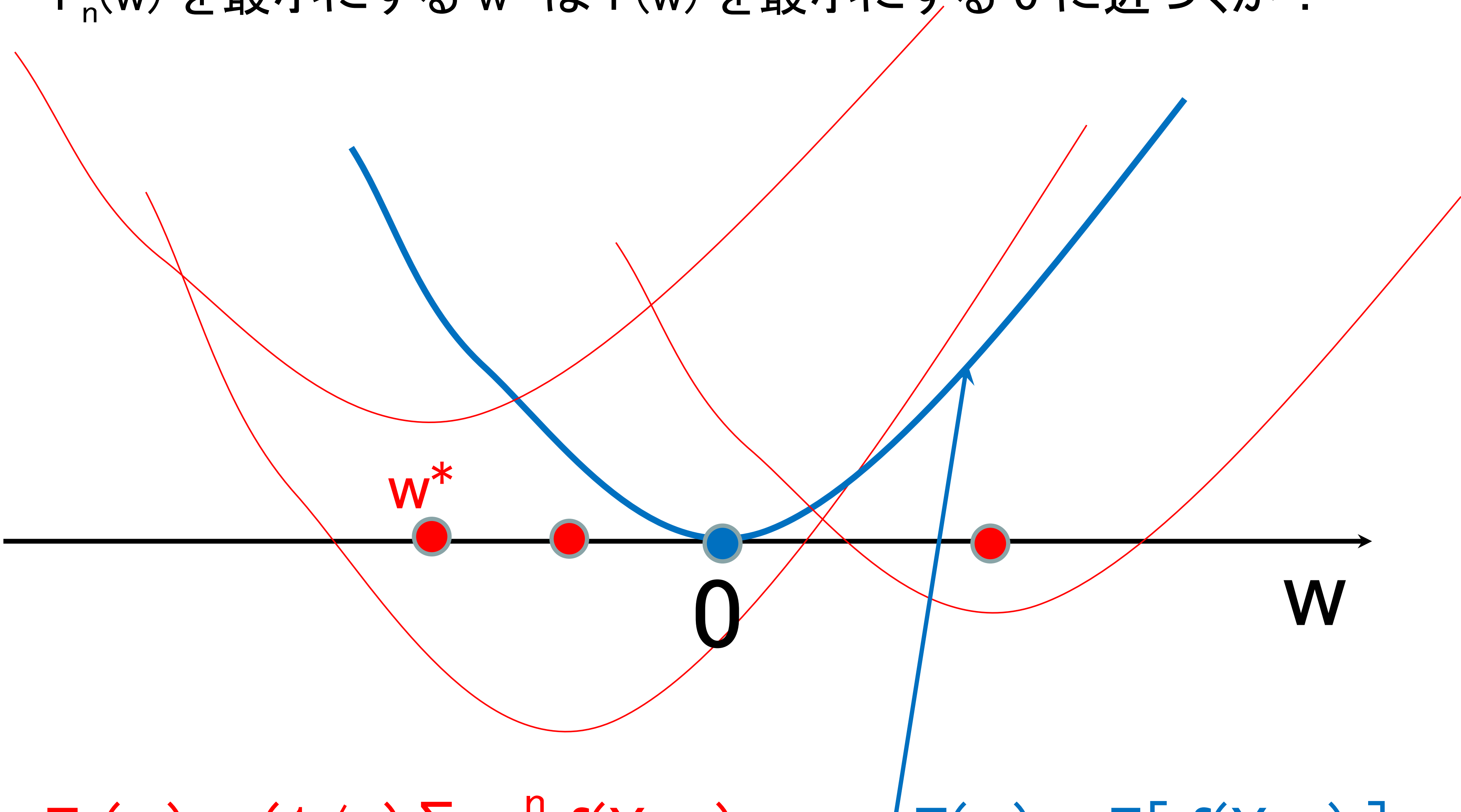
(3) $h(x) \equiv \sup_{|w|<M} \left| (\partial f / \partial w)(x,w) \right|$ とおくと $E[h(X)]=D < \infty$

定理　$X_1, X_2, \cdots, X_n$ は X と同じ分布に従う独立な確率変数

$\Rightarrow$ $F_n(w) = (1/n) \sum_{i=1}^{n} f(X_i,w)$ を最小にする $w^*$ は 0 に確率収束

# 問題の図示

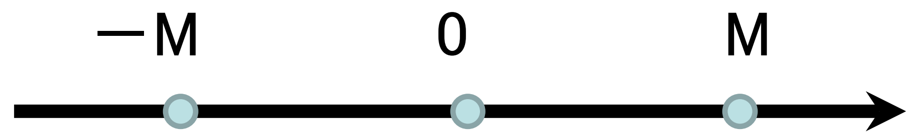
$F_n(w)$ を最小にする $w^*$ は $F(w)$ を最小にする 0 に近づくか？




$F_n(w) = (1/n)\,\Sigma_{i=1}^{n}\, f(X_i,w)$

$F(w) = E[\, f(X,w)\, ]$

例




例. $f(x,w) = w^2 - 2xw$, $X \sim N(0,1^2)$ とする。条件を確認する。

（1） $F(w) = w^2$ だから $F(0)=0$ かつ $F(|w|)$ は $|w|$ の増加関数

（2） $|x| > M$ のとき $\inf_{|w|>M} f(x,w) \geqq -|x|^2$

$|x| \leqq M$ のとき $\inf_{|w|>M} f(x,w) \geqq |w|^2 - 2|x||w| \geqq M^2 - 2|x|M$

$$E[g(X)] \geqq -\int_{|x|>M} |x|^2 q(x)dx + \int_{|x|\leqq M} (M^2 - 2|x|M)\, q(x)dx$$

$$\geqq M^2 \int_{|x|\leqq M} q(x)dx - 2ME[|X|] - E[|X|^2]$$

M を十分大きくとればこれを正の値にできる。

（3） $|w|<M$ のとき

$|(\partial f / \partial w)| = 2|w-x|$ だから $E|(\partial f / \partial w)| \leqq 2M + 2E[|x|] < \infty$

# 例　つづき

例.　$f(x,w) = w^2 - 2xw$,　$X \sim N(0,1^2)$ の場合は

$F_n(w) = (1/n) \sum_{i=1}^n f(X_i,w)$ を最小にする $w^*$ は具体的に計算できる。

実際　$F_n(w) = w^2 - 2w\{(1/n)\sum_{i=1}^n X_i\}$　だから

$w^* = \{(1/n)\sum_{i=1}^n X_i\}$. 各 $X_i$ は $N(0,1^2)$ に従うから $E[X_i]=0$.

従って 大数の法則から $X^* \to 0$ が成り立つ。

これは概収束したがって確率収束である。

定理は、このように具体的でない場合でも確率収束が成り立つということを主張している。

# 基本的な補題

補題1.　$P(B_n) \to 1$, $P(A_n \cap B_n) \to 0$ ならば $P(A_n) \to 0$

（証明）　$P(A_n) = P(A_n \cap B_n) + P(A_n \cap B_n^C)$

$$\leqq P(A_n \cap B_n) + P(B_n^C) \to 0$$

補題2.「$X_n \to 0$, $Y_n \to A > 0$（共に確率収束）ならば $P(X_n \geqq Y_n) \to 0$」

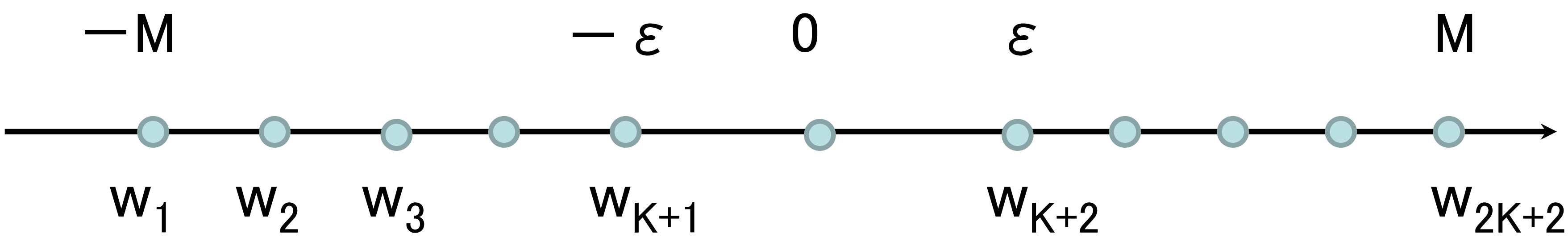
（証明）

$P(|X_n| < A/3) \to 1$,　$P(|Y_n - A| < A/3) \to 1$

$P(X_n < A/3) \to 1$,　$P(Y_n > 2A/3) \to 1$

$P(X_n \geqq Y_n) \geqq P(X_n < A/3 \cap Y_n > 2A/3) \to 1$

# 証明の流れ




ε >0 を任意に取って固定。M は条件で与えられている。

ε とM に対して K を十分大きくとって固定。

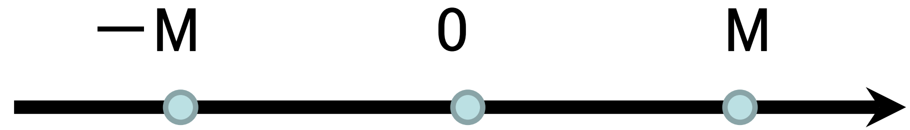
まず w* がMより外にある確率は零に近づくを示す。

☆　$P(|w^*| \geqq M) \to 0$

次に j＝K+1以外で　$|w^*| \in [w_j, w_{j+1}]$ となる確率が零に近づく

☆☆　$P(\ |w^*| \in [w_j, w_{j+1}]\ ) \to 0$　　を示す。

# 証明(1)




☆ $P(|w^*| \geqq M) \to 0$ : $w^*$ がMより外の確率は零に近づくを示す。

まず 条件(2) と大数の法則から

$$\inf_{|w|>M} F_n(w) \geqq (1/n) \Sigma_{i=1}^{n} \inf_{|w|>M} f(X_i, w)$$

$$= (1/n) \Sigma_{i=1}^{n} g(X_i) \equiv Y_n \to C > 0 \quad \text{（確率収束）}$$

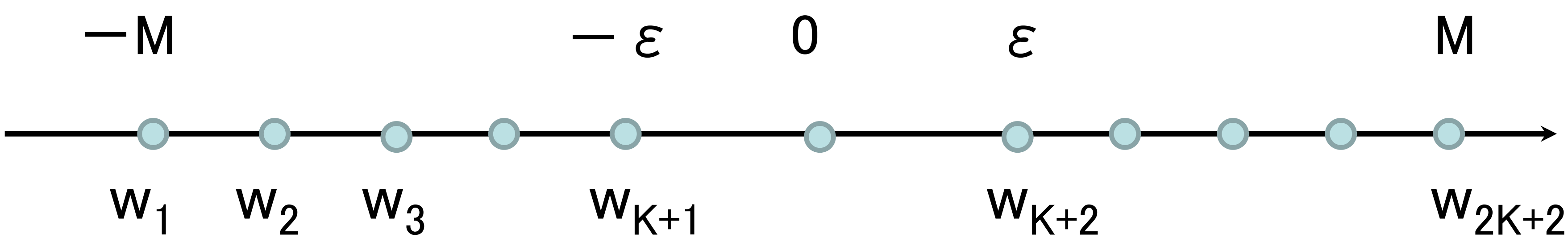
次に 大数の法則から $F_n(0) \to F(0) = 0$（確率収束）

$F_n(w)$ は $w^*$ で最小になるから

$|w^*| \geqq M$ ならば $Y_n \leqq F_n(0)$ が成り立つので補題2から

$P(|w^*| \geqq M) \leqq P(F_n(0) \geqq Y_n) \to 0$ ☆が示せた。

# 証明(2)




記号の決め方

$\varepsilon>0$ を任意に(好きなだけ小さく)決定。

$\delta = \min_k \ \{ F(-\varepsilon),F(\varepsilon) \}$ と定める ($\delta>0$)。

M を十分大きくとる。$K > 3MD/\delta$ となるように K を定める。

$[-M, -\varepsilon]$ と $[\varepsilon, M]$ を K 等分。各区間の長さは $(M-\varepsilon)/K$

条件(1)の単調の仮定より $F(w_j) \geqq \delta$ が成り立つ。

# 証明（３）

j＝K+1以外で $|w^*| \in [w_j, w_{j+1}]$ となる確率が零に近づくこと

☆☆ $P( |w^*| \in [w_j, w_{j+1}] ) \to 0$ を示す。

平均値の定理からある $w^+$ $(w_j < w^+ < w_{j+1})$ が存在して

$$\sup_{w \in [w_j, w_{j+1}]} | F_n(w) - F_n(w_j) | = \sup_{w \in [w_j, w_{j+1}]} |w - w_j| \, |F_n'(w^+)|$$

$j \neq K+1$ とする。上式の値をΔと書くと条件(3)から

$$\Delta \leqq \sup_{w \in [w_j, w_{j+1}]} |w - w_j| \, (1/n) \Sigma_{i=1}^{n} h(X_i)$$

$$\leqq (M/K) \, (1/n) \Sigma_{i=1}^{n} h(X_i)$$

$$\leqq (\delta/3D) \, (1/n) \Sigma_{i=1}^{n} h(X_i) \to \delta/3 \text{（確率収束）}$$

従って $P( \Delta \leqq \delta/2 ) \to 1$

# 証明（4）

前ページより　$P(\ \sup_{w\in[w_j,w_{j+1}]} | F_n(w)-F_n(w_j) | \leqq \delta/2\ ) \to 1$

$Z_n \equiv \inf_{w\in[w_j,w_{j+1}]} F_n(w)$ とおくと　$P(\ Z_n \geqq F_n(w_j) - \delta/2\ ) \to 1$

一方、各　$\{w_j\}$ において $F_n(w_j) \to F(w_j) \geqq \delta > 0$　（確率収束）

$P(\ Z_n \geqq F_n(w_j) - \delta/2\ ) \to 1$ および

$P(F_n(0) \geqq Z_n \cap Z_n \geqq F_n(w_j) - \delta/2) \leqq P(F_n(0) \geqq F_n(w_j) - \delta/2\ ) \to 0$

より補題1より $P(F_n(0) \geqq Z_n\ ) \to 0$

$w^* \in [w_j,w_{j+1}]$ ならば $Z_n \leqq F_n(0)$ が成り立つので

$P(\ w^* \in [w_j,w_{j+1}]\ ) \leqq P(F_n(0) \geqq Z_n) \to 0$　　☆☆が示せた。

# 証明(5)

以上で示したことをもう一度書くと

☆ $P(|w^*| \geqq M) \to 0$

☆☆ $j = K+1$以外で $P(w^* \in [w_j, w_{j+1}]) \to 0$

従って $P(w^* \in (-\varepsilon, \varepsilon)) \to 1$ が得られた.

(注) 定理が成り立つための十分条件は

もう少し緩めることができます。