数理手法VI

第4回

演習問題の解答

演習問題

$(Z_n)_{n\in\mathbb{N}}$: *i.i.d. r.v.s*, $Z_1 \in L^4$, $\mu = E[Z_1]$, $v = E[(Z_1-\mu)^2]$ とする.

$$X_{m,n} = \frac{Z_m - \mu}{\sqrt{nv}}, \quad \mathcal{F}_{m,n} = \sigma(Z_1, \cdots, Z_m)$$

とおくと，定理 *1.6* の *1.* と *2.* が成立.

演習問題の解答

（解答例）
$\sigma(Z_m)$ と $\mathcal{F}_{m-1,n}$ は独立なので,

$$\sum_{m=1}^{n} E[X_{m,n}^2|\mathcal{F}_{m-1,n}] = \sum_{m=1}^{n} E\left[\left(\frac{Z_m-\mu}{\sqrt{nv}}\right)^2\right] = \frac{n \cdot E[(Z_1-\mu)^2]}{nv} = 1.$$

よって 1. が成り立つ.

また,

$$\sum_{m=1}^{n} E[X_{m,n}^4] = \frac{\sum_{m=1}^{n} E[(Z_m-\mu)^4]}{n^2v^2} = \frac{E[(Z_1-\mu)^4]}{nv^2} \to 0 \quad (n\to\infty)$$

より 2. が成立.

AR 過程の最尤推定

$0 < \theta_0 < 1$, $(Z_n)_{n \in \mathbb{N}}$: i.i.d. 確率変数列で $Z_n \sim N(0,1)$ とする.
確率変数列 $(X_n)_{n \in \mathbb{Z}_+}$ を

$$X_0 = 0, \quad X_n = \theta_0 X_{n-1} + Z_n \ (n \in \mathbb{N})$$

と定める.

最尤推定量 $\hat{\theta}_n$ に対して

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, 1 - \theta_0^2)$$

を示した.

3．信頼区間の計算

漸近正規性を使うと，推定量の存在する範囲を表す「信頼区間」を近似的に計算することができる．

$Z \sim N(0,1)$ の時，$P(-1.96 \le Z \le 1.96) \approx 0.95$ となることが知られている．

$V \sim N(0, 1-\theta_0^2)$ の時，$\frac{V}{\sqrt{1-\theta_0^2}} \sim N(0,1)$　なので，

$$P(-1.96 \times \sqrt{1-\theta_0^2} \le V \le 1.96 \times \sqrt{1-\theta_0^2}) \approx 0.95.$$

3．信頼区間の計算

漸近正規性より，n が十分大きい時,

$$P\left(|\hat{\theta}_n - \theta_0| \le \frac{1.96 \times \sqrt{1-\theta_0^2}}{\sqrt{n}}\right)$$
$$= P(\sqrt{n}|\hat{\theta}_n - \theta_0| \le 1.96 \times \sqrt{1-\theta_0^2}) \approx 0.95.$$

よって推定誤差は $1/\sqrt{n}$ に比例して小さくなることがわかる.

3．信頼区間の計算

また，

$$
\begin{aligned}
0.95 &\approx P\left(-1.96 \times \sqrt{1-\theta_0^2} \le \sqrt{n}(\hat{\theta}_n - \theta_0) \le 1.96 \times \sqrt{1-\theta_0^2}\right) \\
&= P\left(\hat{\theta}_n - \frac{1.96 \times \sqrt{1-\theta_0^2}}{\sqrt{n}} \le \theta_0 \le \hat{\theta}_n + \frac{1.96 \times \sqrt{1-\theta_0^2}}{\sqrt{n}}\right) \\
&\approx P\left(\hat{\theta}_n - \frac{1.96 \times \sqrt{1-\hat{\theta}_n^2}}{\sqrt{n}} \le \theta_0 \le \hat{\theta}_n + \frac{1.96 \times \sqrt{1-\hat{\theta}_n^2}}{\sqrt{n}}\right).
\end{aligned}
$$

3．信頼区間の計算

よって n が大きい時未知パラメータ θ_0 はおおよそ 95 %の確率で，

$$\hat{\theta}_n - \frac{1.96 \times \sqrt{1-\hat{\theta}_n^2}}{\sqrt{n}} \quad \text{から} \quad \hat{\theta}_n + \frac{1.96 \times \sqrt{1-\hat{\theta}_n^2}}{\sqrt{n}}$$

の範囲に存在することがわかる．（95 %信頼区間）

第 2 章 マルコフ連鎖

2.1 マルコフ連鎖の定義

d 次元ボレル集合族

$d \geq 1$ として，d 次元ユークリッド空間 $\mathbb{R}^d$ 上の

$$\{(-\infty, a_1] \times \cdots \times (-\infty, a_d]; a_1, \cdots, a_d \in \mathbb{R}\}$$

が生成する σ-加法族を $\mathcal{B}(\mathbb{R}^d)$ と書き，d 次元ボレル集合族と呼ぶ.
d 次元ボレル集合族の元を (d 次元) ボレル集合と呼ぶ.
開集合や閉集合，その可算和集合などが d 次元ボレル集合族に含まれる.

可測関数 $X : \Omega \to \mathbb{R}^d$ を $\mathbb{R}^d$-値確率変数（もしくは d 次元確率変数）といい，$\sigma(X)$ が同様に定義される.

$X = x$ の下での条件付期待値

$X, Y \in L^1$ とする．マルコフ連鎖では「$X = x$ の下での Y の条件付期待値（確率）」を考える．

条件付期待値 $E[Y|\sigma(X)]$ は数理手法 IV で定義していて，これは $A \in \mathcal{B}(\mathbb{R})$ に対して，$X^{-1}(A)$ 上で

$$E[E[Y|\sigma(X)], X^{-1}(A)] = E[Y, X^{-1}(A)]$$

となるような確率変数であった．

特に $A = \{x\}$ として，$P(X = x) > 0$ の時は，$X^{-1}(\{x\})$ 上，$E[Y|\sigma(X)]$ は定数となり，

$$E[Y|\sigma(X)] \times P(X = x) = E[Y, \{X = x\}].$$

$X = x$ の下での条件付期待値

よって

$$E[Y|\sigma(X)](\omega) = \frac{E[Y, \{X = x\}]}{P(X = x)}. \quad (\omega \in X^{-1}(\{x\}).$$

($X^{-1}(\{x\})$ 上，$E[Y|\sigma(X)]$ は定数となるのは，数理手法 IV の命題 2.16 の証明と同様に示せる）

上式右辺の形は（高校数学で学ぶような）$X = x$ の下での Y の条件付期待値の形になっている．よって，$X^{-1}(\{x\})$ 上の $E[Y|\sigma(X)]$ の値として，$X = x$ の下での Y の条件付期待値を定義すればよいように思える．

- このような議論を $P(X = x) = 0$ の場合にも拡張したい．

$\sigma(X)$-可測関数の性質

命題 2.1
X を $\mathbb{R}^d$-値確率変数として，T を $\sigma(X)$-可測確率変数とする．この時，可測関数 $g : \mathbb{R}^d \to \mathbb{R}$ が存在して，$T = g(X)$.

- $\sigma(X)$-可測な関数は X の関数の形で書ける．

Proof. まず，ある $B \in \sigma(X)$ に対して，$T = 1_B$ となる場合を考える．この時，ある $A \in \mathcal{B}(\mathbb{R}^d)$ に対して，$B = X^{-1}(A)$ となるので，

$$T(\omega) = 1_{X^{-1}(A)}(\omega) = 1_A(X(\omega)).$$

よって $g(x) = 1_A(x)$ に対して，$T = g(X)$ と書けるので OK.

$\sigma(X)$-可測関数の性質

次に T が $\sigma(X)$-可測単関数の時も同様にある g に対して $T = g(X)$ と書ける．T が一般の $\sigma(X)$-可測関数の時は，$\sigma(X)$-可測な単関数の列 $(T_n)_{n=1}^{\infty}$ があって，

$$T_n(\omega) \to T(\omega) \quad \text{as} \quad n \to \infty \quad (\omega \in \Omega)$$

となる．この時，可測関数列 $g_n : \mathbb{R}^d \to \mathbb{R}$ があって，$T_n(\omega) = g_n(X(\omega))$ となる．よって

$$\lim_{n\to\infty} g_n(X(\omega)) = \lim_{n\to\infty} T_n(\omega) = T(\omega).$$

$\sigma(X)$-可測関数の性質

$E := \{x|$ 有限な $\lim_{n\to\infty} g_n(x)$ が存在する $\}$ とおくと，
$X(\omega) \in E$ $(\omega \in \Omega)$. よって

$$g(x) = \begin{cases} \lim_{n\to\infty} g_n(x) & \text{if } x \in E \\ 0 & \text{if } x \in E^c \end{cases}$$

とおくと， $T(\omega) = g(X(\omega)) \quad (\omega \in \Omega)$. $\square$

条件付期待値・条件付確率の定義

定義 2.2

1. X を $\mathbb{R}^d$-値確率変数，$Y \in L^1$ に対し，$E[Y|\sigma(X)] = E[Y|X]$ と書く．命題 *2.1* より，
$$E[Y|X] = g(X)$$
となる可測関数 $g : \mathbb{R}^d \to \mathbb{R}$ が存在する．
この時 $E[Y|X = x] := g(x)$ と定義する．

2. $\mathbb{R}^d$-値確率変数 X，$A \in \mathcal{F}$ に対して，
$$P(A|X = x) = E[1_A|X = x], \quad P(A|X) = E[1_A|X]$$
と定め，条件付確率と呼ぶ．

条件付期待値の一意性

g, g' が上の定義の条件を満たす時，条件付期待値の a.s. での一意性より

$$g(X) = g'(X) \quad \text{a.s.}$$

よって X の分布を μ_X と書くと，

$$\mu_X(\{x | g(x) = g'(x)\}) = P(\{\omega | g(X(\omega)) = g'(X(\omega))\}) = 1.$$

つまり，$E[Y|X = x]$ は μ_X-a.s. に一意に定まる．

例

例 2.3

Y_1, Y_2: 独立な確率変数，$P(Y_i = \pm 1) = 1/2$ $(i = 1, 2)$ とする．
$X_1 = Y_1$, $X_2 = Y_1 + Y_2$ とする．この時 $E[X_2|X_1] = X_1$ なので $g(x) = x$ として，$E[X_2|X_1 = x] = x$ となる．$X_1 = \pm 1$ の下で X_2 の条件付期待値は ± 1 ということ．

マルコフ連鎖の定義

定義 2.4

写像 $p : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0,1]$ が推移確率であるとは，

1. 各 $x \in \mathbb{R}$ に対して，$p(x,\cdot)$ は $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度
2. 各 $A \in \mathcal{B}(\mathbb{R})$ に対して，$p(\cdot, A)$ は可測

をみたす時をいう．

推移確率はマルコフ・カーネルとも呼ばれる．

マルコフ連鎖の定義

定義 2.5

確率過程 $(X_n)_{n\in\mathbb{Z}_+}$ がマルコフ連鎖であるとは，推移確率 p があって，任意の $k \in \mathbb{Z}_+$ と $A \in \mathcal{B}(\mathbb{R})$ に対して，

$$P(X_{k+1} \in A | (X_0, \cdots, X_k) = (x_0, \cdots, x_k)) = p(x_k, A) \quad \mu_{(X_0,\cdots,X_k)} - a.s. \tag{1}$$

（つまり $\mu_{(X_0,\cdots,X_k)}(\{(x_0,\cdots,x_k) | (1)\text{ が成立 }\}) = 1$）

をみたす時をいう．

※ (1) は $P(X_{k+1} \in A | X_0, \cdots, X_k) = p(X_k, A) \quad \text{a.s.}$ と同値．
※ X_{k+1} の状態が $X_0, \cdots, X_{k-1}$ の値に関わらず直前の X_k の値だけで決まるということ．

X_0 の分布 μ_{X_0} を初期分布という．

条件付期待値の性質

命題 2.6

X: $\mathbb{R}^k$-値確率変数，Y: $\mathbb{R}^l$-値確率変数，X と Y は独立，
$\varphi : \mathbb{R}^k \times \mathbb{R}^l \to \mathbb{R}$ を可測関数で $E[|\varphi(X,Y)|] < \infty$ とする．この時，
$g(x) := E[\varphi(x,Y)]$ は可測関数で

$$E[\varphi(X,Y)|X] = g(X) \quad \text{a.s.}$$

証明は Durrett 5th ed. の Example 4.1.7 を見よ．

マルコフ連鎖の例

例 2.7

$(Z_n)_{n\in\mathbb{N}}$: *i.i.d.* 確率変数列，$X_0 = 0$, $X_n = \sum_{k=1}^{n} Z_k$ $(n \in \mathbb{N})$ とする．この時 $(X_n)_{n\in\mathbb{Z}_+}$ はマルコフ連鎖.

※ $(X_n)_{n\in\mathbb{Z}_+}$ がマルチンゲールになるためには $E[Z_1] = 0$ が必要だったが，マルコフ連鎖では必要ない.（L^1 である必要もない.）

演習問題

命題 *2.6* を使って上の $(X_n)_{n\in\mathbb{Z}_+}$ がマルコフ連鎖になることを証明せよ.

マルコフ連鎖の性質

次にマルコフ連鎖の性質をいくつか調べる．

命題 2.8

$(X_n)_{n\in\mathbb{Z}_+}$ をマルコフ連鎖とすると，任意の $k\in\mathbb{Z}_+$ と可測関数 $f:\mathbb{R}\to\mathbb{R}$ に対して，$f(X_{k+1})\in L^1$ なら，$x_k\mapsto\int f(x)p(x_k,dx)$ は可測で，

$$E[f(X_{k+1})|X_0,\cdots,X_k]=\int f(x)p(X_k,dx)\quad\text{a.s.}$$

マルコフ連鎖の性質

Proof. $A \in \mathcal{B}(\mathbb{R})$ に対し，$f = 1_A$ なら

$$x_k \mapsto \int f(x)p(x_k, dx) = p(x_k, A)$$

は推移確率の定義より可測で，マルコフ連鎖の定義より，

$$E[f(X_{k+1})|X_0, \cdots, X_k] = p(X_k, A) = \int f(x)p(X_k, dx) \quad \text{a.s.}$$

より OK.

マルコフ連鎖の性質

（条件付）期待値の線形性より f が単関数の場合も成立し，単調収束定理等より f が一般の可測関数の場合も成立．

（$f(X_{k+1}) \in L^1$ で f が非負の時

$$\int f(x)p(X_k, dx) = E[f(X_{k+1})|X_0, \cdots, X_k] < \infty \quad \text{a.s.}$$

より f が一般でも $\int f(x)p(X_k, dx)$ が定義できる.） □