

マルコフ連鎖に関して講義中に紹介できなかった定理とその証明を与える.

1 マルコフ連鎖の性質

表記

- I : マルコフ連鎖の状態空間 (高々可算)
- $i \rightarrow j$: i は j に到達可能
- $T_j = \inf\{n \geq 1 \mid X_n = j\}$: 状態 j への初到達時間 ($\inf \emptyset = \infty$ とする)
- $T_j^{(k)}$: 状態 j に k 回到達した時間.
- $N_j(n)$: 時刻 n までに j を訪れた回数.
- $f(i, j) = P(T_j < \infty \mid X_0 = i)$: i から j への到達確率
- $m(i, j) = E[T_j \mid X_0 = i]$: 平均初到達時間

1.1 周期性

Definition 1 (周期性). マルコフ連鎖の再帰的な状態 i の周期 $d(i)$ とは, $P^{(m)}(i, i) > 0$ を満たす整数 $m \geq 1$ の最大公約数である. $d(i) = 1$ なら状態 i を非周期的とよぶ.

Theorem 2 (同値類における周期の同一性). $i, j \in I$ を再帰的な状態とする. $i \leftrightarrow j$ なら $d(i) = d(j)$.

これより, 周期はクラスの性質であることがわかる.

Proof. $i \leftrightarrow j$ より, ある k, ℓ が存在して, $P^{(k)}(i, j) > 0$ かつ $P^{(\ell)}(j, i) > 0$ である. このとき, $P^{(k+\ell)}(i, i) \geq P^{(k)}(i, j)P^{(\ell)}(j, i) > 0$ であるので, $k + \ell$ は $d(i)$ の倍数である. 一方, $P^{(m)}(j, j) > 0$ である任意の m に対して,

$$P^{(k+\ell+m)}(i, i) \geq P^{(k)}(i, j)P^{(m)}(j, j)P^{(\ell)}(j, i) > 0$$

である. すると, 周期の定義より $k+\ell+m$ は $d(i)$ の倍数である. $k+\ell$ が $d(i)$ の倍数なので, m も $d(i)$ の倍数である. ここで, $d(j)$ はそのような m の最大公約数であるので $d(i)$ は $d(j)$ 以下である. 同様に $d(i) \geq d(j)$ も言えるので, $d(i) = d(j)$. \square

Theorem 3 (周期の性質). $i \leftrightarrow j$ として, それらの共通の周期を d とすると,

$$P^{(\ell)}(i, j) > 0, P^{(m)}(i, j) > 0 \Rightarrow \ell = m \pmod{d}$$

である.

Proof. ある n が存在して, $P^{(n)}(j, i) > 0$ であるので,

$$\begin{aligned} P^{(\ell+n)}(i, i) &\geq P^{(\ell)}(i, j)P^{(n)}(j, i) > 0, \\ P^{(m+n)}(i, i) &\geq P^{(m)}(i, j)P^{(n)}(j, i) > 0, \end{aligned}$$

であるので, $\ell + m$ と $m + n$ はともに d の倍数である. よって, $\ell - m = (\ell + m) - (m + n)$ も d の倍数である. \square

この定理より, 既約かつ閉な集合 $R \subset I$ を次のように分割することができる. すなわち, 状態 i を基準として各 $j \in R$ を $P^{(\ell)}(i, j) > 0$ なる ℓ を周期 d で割った余りに応じて $C_i(\ell \pmod{d})$ なる集合に割り付けるこ

とができる:

$$R = C_i(1) \cup C_i(2) \cup C_i(3) \cup \dots \cup C_i(d).$$

ここで、周期の性質から、各 j の割り当ては $P^{(\ell)}(i, j) > 0$ である ℓ の選び方に依存しないことに注意されたい。

1.2 吸収確率

$T = \{i \in I \mid \exists j \in I \text{ s.t. } i \rightarrow j, j \not\rightarrow i\}$ を非再帰的な同値類とする。再帰的な状態 j に対して、 $C(j)$ を j を含む再帰的な同値類とする: $C(j) = \{i \in I \mid j \rightarrow i\}$.

Lemma 4. j を再帰的な状態であるとする。すると、任意の $i \in C(j)$ に対して、 $f(i, j) = 1$ が成り立つ。

Proof. $j \rightarrow i$ であることより、 j から i へ至る経路 $j = y_1, y_2, \dots, y_K = i$ で、 $y_\ell \neq j$ かつ $y_\ell \neq i$ で

$$P(y_1, y_2)P(y_2, y_3) \dots P(y_{K-1}, y_K) > 0$$

を満たすものが存在する。もし、途中で j を経由することがあれば、そこを y_1 とすれば、必ず $y_\ell \neq j$ ($\forall \ell$) とすることができる。すると、 j から出発して i に至り、それから二度と j に戻らない事象を考えることで

$$0 \geq 1 - f(j, j) \geq P(y_1, y_2)P(y_2, y_3) \dots P(y_{K-1}, y_K)(1 - f(i, j))$$

が得られる (再帰性の定義から $f(j, j) = 1$ であることに注意)。 $P(y_1, y_2)P(y_2, y_3) \dots P(y_{K-1}, y_K) > 0$ なので、 $f(i, j) = 1$ を得る。 \square

Theorem 5. $i \in T$ から j への吸収確率 $a(i) = f(i, j)$ は

$$a(i) = \sum_{k \in C(j)} P(i, k) + \sum_{k \in T} P(i, k)a(k) \quad (i \notin C(j)), \quad (1a)$$

$$a(k) = 1 \quad (\forall k \in C(j)), \quad (1b)$$

を満たす最小の非負解である。

Proof. まず、 $k \in C(j)$ に対しては、Lemma 4 から $a(k) = f(k, j) = 1$ である。また、 $i \notin C(j)$ に対しては、

$$\begin{aligned} a(i) &= P(T_j < \infty \mid X_0 = i) \\ &= \sum_{k \in I} P(T_j < \infty \mid X_1 = k, X_0 = i)P(X_1 = k \mid X_0 = i) \\ &= \sum_{k \in I} P(T_j < \infty \mid X_0 = k)P(X_1 = k \mid X_0 = i) \\ &= \sum_{k \in I} f(k, j)P(i, k) \end{aligned}$$

である。Lemma 4 から全ての $k \in C(j)$ に対して $f(k, j) = 1$ が成り立ち、分解定理から $k \in (T \cup C(j))^c$ に対しては $f(k, j) = 0$ が成り立つ。よって、上の等式は

$$a(i) = \sum_{k \in T} P(i, k)f(k, j) + \sum_{k \in C(j)} P(i, k)$$

と書き換えられる。

次に $a(k)$ が非負解の中で最小であることを示す。他に方程式 (1) を満たす非負解 $b(k)$ ($k \in I$) があるとす。今、 p_{ij} を

この b が任意の $n = 1, 2, \dots$ において、

$$b(k) \geq \begin{cases} P(T_j \leq n | X_0 = k) & (k \neq j), \\ 1 & (k = j), \end{cases} \quad (2)$$

を満たすことが示せれば, $n \rightarrow \infty$ で, 右辺は $P(T_j < \infty | X_0 = k) = f(k, j)$ に単調に収束するので, $b(k) \geq a(k)$ が示される.

これを n に関する帰納法で示す. まず, 式 (1) から $b(k) = 1$ ($k \in C(j)$) なので, $k \in C(j)$ に対しては, 式 (2) が自明に成り立つ. 以下では, $k \notin C(j)$ を仮定する. $n = 0$ ならば, $b(k) \geq P(T_j = 0 | X_0 = k) = 0$ は自明に成り立つ. もし, ある $n \geq 0$ まで式 (2) が成り立っていると仮定すると, b は式 (1) を満たすので,

$$\begin{aligned} b(k) &= \sum_{\ell \in C(j)} P(k, \ell) + \sum_{\ell \in T} P(k, \ell)b(\ell) \\ &= \sum_{\ell \neq j} P(k, \ell)b(\ell) + P(k, j) \\ &\geq \sum_{\ell \neq j} P(k, \ell)P(T_j \leq n | X_0 = \ell) + P(k, j) \\ &= P(T_j \leq n + 1 | X_0 = k) \end{aligned}$$

である. よって, 式 (2) は $n + 1$ においても成り立つ. \square

1.3 極限分布

Lemma 6. $d(i) = 1$ なら, ある m_0 が存在して $P^{(m)}(i, i) > 0$ が全ての $m \geq m_0$ に対して成り立つ.

Proof. $D(i) = \{m \geq 1 \mid P^{(m)}(i, i) > 0\}$ とする. まず, $D(i)$ に連続した二整数が含まれていることを示す. $n_0, n_0 + k \in D(i)$ であるとする. もし $k = 1$ なら, 連続した二整数が含まれることになる. もし, そうでないなら, $D(i)$ の最大公約数が $d(i)$ であることから, ある $n_1 \in D(i)$ が存在して, k は n_1 の約数にはならない. この n_1 を $n_1 = mk + r$ ($0 < r < k$) と表すと, $D(i)$ が和に関して閉じていることから (これは簡単に確認できる), $(m + 1)(n_0 + k) > (m + 1)n_0 + n_1$ はともに $D(i)$ に含まれる. それらの差は,

$$k(m + 1) - n_1 = k - r < k$$

である. これを高々 k 回繰り返すことで, ある十分大きな整数 N で $N, N + 1 \in D(i)$ となるものが見つかる.

すると, $m_0 = N^2$ とすれば $m \geq N^2$ なら $m - N^2 = kN + r$ ($0 \leq r < N$) と書けるが, $m = r + N^2 + kN = r(1 + N) + (N - r + k)N \in D(i)$ がわかる. \square

Theorem 7 (極限分布への収束定理). マルコフ連鎖が既約で非周期的で定常分布 π を持つなら,

$$P^{(n)}(i, j) \rightarrow \pi(j) \quad (\forall i, j \in I).$$

Proof. (カップリング法)

まず, マルコフ連鎖は既約で定常分布を持つので, 再帰的である.

$I^2 = I \times I$ とする. $\bar{P}((i_1, j_1), (i_2, j_2)) = P(i_1, i_2)P(j_1, j_2)$ とする. つまり, 各座標は独立に動くとする. \bar{P} の既約性を示す. P の既約性より, 任意の i_1, i_2, j_1, j_2 に対して, ある m, k が存在して,

$$P^{(m)}(i_1, i_2) > 0, \quad P^{(k)}(j_1, j_2) > 0$$

である. また Lemma 6 より十分大きな ℓ があって, $P^{(k+\ell)}(i_2, i_2) > 0$ かつ $P^{(m+\ell)}(j_2, j_2) > 0$ である. よって,

$$\begin{aligned} \bar{P}^{(k+\ell+m)}((i_1, j_1), (i_2, j_2)) &= P^{(k+\ell+m)}(i_1, i_2)P^{(k+\ell+m)}(j_1, j_2) \\ &\geq P^{(m)}(i_1, i_2)P^{(\ell+k)}(i_2, i_2)P^{(k)}(j_1, j_2)P^{(\ell+m)}(i_2, i_2) > 0 \end{aligned}$$

となる。つまり、 $(i_1, j_1) \rightarrow (i_2, j_2)$ である。 $(i_1, j_1), (i_2, j_2)$ は任意なので、 \bar{P} も既約である。

次に \bar{P} の再帰性を示す。 $\bar{\pi}(i, j) = \pi(i)\pi(j)$ が \bar{P} の定常分布になることはすぐにわかる。 \bar{P} は既約なので、定常分布が存在することから正再帰的である。

(X_n, Y_n) を \bar{P} のマルコフ連鎖とする。 T を $X_n = Y_n$ となる初めての時間とする ($T = \inf\{n \geq 1 \mid X_n = Y_n\}$)。また、 $T_{(x,x)}$ を $(x, x) \in I^2$ に到達する初めての時間とする ($T_{(x,x)} = \inf\{n \geq 1 \mid X_n = Y_n = x\}$)。 \bar{P} は既約かつ再帰的なので $P(T_{(x,x)} < \infty) = 1$ であり、特に $P(T < \infty) = 1$ でもある。

$\{T \leq n\}$ なる条件のもとで、 X_n, Y_n が同分布に従うことを示す。実際に、

$$\begin{aligned} P(X_n = y, T \leq n) &= \sum_{m=1}^n \sum_x P(T = m, X_m = x, X_n = y) \\ &= \sum_{m=1}^n \sum_x P(T = m, X_m = x) P(X_n = y \mid X_m = x) \quad (\because \text{マルコフ性}) \\ &= \sum_{m=1}^n \sum_x P(T = m, Y_m = x) P(Y_n = y \mid Y_m = x) \\ &= P(Y_n = y, T \leq n), \end{aligned}$$

である。よって、

$$P(X_n = y) = P(Y_n = y, T \leq n) + P(X_n = y, T > n) \leq P(Y_n = y) + P(X_n = y, T > n).$$

が成り立ち、同様に $P(Y_n = y) \leq P(X_n = y) + P(Y_n = y, T > n)$ も示せる。よって、

$$\sum_y |P(X_n = y) - P(Y_n = y)| \leq 2P(T > n)$$

を得る。 $X_0 = x$ として、 Y_0 が π に従うとすると、

$$\sum_y |P^{(n)}(x, y) - \pi(y)| \leq 2P(T > n) \rightarrow 0 \quad (\text{as } n \rightarrow \infty).$$

□

1.4 定常分布と大数の強法則

Theorem 8 (マルコフ連鎖の定常分布と正再帰性). マルコフ連鎖が既約であるとする。すると、以下は同値である。

1. ある $j \in I$ が正再帰的である。
2. 全ての $j \in I$ が正再帰的である。
3. 定常分布がただ一つ存在する。

また、その定常分布は

$$\pi(j) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\ell=1}^n P^{(\ell)}(j, j) = \frac{1}{m(j, j)}$$

で与えられる。

この定理より、正再帰性はクラスの性質であることがわかる。

Proof. (2) から (1) は自明。

(1) から (2) を示す。任意の $i \in I$ をとってくる。既約性より、 $i \leftrightarrow j$ なのである m, k が存在して、

$$P^{(m)}(i, j) > 0, \quad P^{(k)}(j, i) > 0$$

とできる。よって、

$$\begin{aligned}
\frac{1}{m(i, i)} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\ell=1}^n P^{(\ell)} P(i, i) \\
&\geq \lim_{n \rightarrow \infty} \frac{1}{n+m+k} \sum_{\ell=1}^n P^{(m)}(i, j) P^{(n)}(j, j) P^{(k)}(j, i) \\
&= P^{(m)}(i, j) P^{(k)}(j, i) \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\ell=1}^n P^{(n)}(j, j) \\
&= P^{(m)}(i, j) P^{(k)}(j, i) \lim_{n \rightarrow \infty} \frac{1}{m(j, j)} > 0
\end{aligned}$$

を得る。よって、 i も正再帰的である。

(2) から (3) を示す。まず、定常分布の存在は Theorem 9 から従う。

次に、定常分布が存在すれば (1) が成り立つことを示し、同時に定常分布の一意性も示す。 $\pi(j) > 0$ なる j に対して、既約性より任意の j' について $j \rightarrow j'$ が成り立つので、ある n が存在して、

$$P^{(n)}(j, j') > 0$$

である。よって、

$$\pi(j') = \sum_i \pi(i) P^{(n)}(i, j') > \pi(j) P^{(n)}(j, j') > 0$$

である。

$j \in I$ を任意の状態とすると、 π が定常分布であることより、Fubini の定理とルベグの収束定理を用いて、

$$\pi(j) = \frac{1}{n} \sum_{m=1}^n \left(\sum_{j'} \pi(j') P^{(m)}(j', j) \right) = \sum_{j'} \pi(j') \left(\frac{1}{n} \sum_{m=1}^n P^{(m)}(j', j) \right) \rightarrow \sum_{j'} \pi(j') \frac{1}{m(j, j)} = \frac{1}{m(j, j)}$$

である。これより、 $\pi(j) = \frac{1}{m(j, j)}$ がわかり、 $\pi(j) > 0$ であることから、 $m(j, j) < \infty$ も示される。よって、 j は正再帰的である。また、上記の議論から任意の定常分布 π に対して $\pi(j) = \frac{1}{m(j, j)}$ が示されているので、定常分布は一意的である。□

Theorem 9 (平均再帰時間と定常分布). 状態 j を正再帰的として、

$$\mu(i) = \sum_{n=0}^{\infty} P(X_n = i, T_j > n | X_0 = j)$$

とする。すると、 $\pi(i) = \mu(i) / \mathbb{E}[T_j | X_0 = j] = \mu(i) / m(j, j)$ は定常分布になる。

Proof. まず、 $\mu(j) = 1$ がすぐわかる。また、 $\sum_i \pi(i) = \sum_{n=0}^{\infty} P(T_j > n) = \mathbb{E}[T_j | X_0 = j] = m(j, j) < \infty$ より、 $\sum_i \pi(i) = 1$ がわかる。よって、 π は確率分布になっている。

$\bar{P}^{(n)}(j, i) = P(X_n = i, T_j > n | X_0 = j)$ とすれば、

$$\sum_i \mu(i) P(i, k) = \sum_{n=0}^{\infty} \sum_i \bar{P}^{(n)}(j, i) P(i, k)$$

である。

(i) ($k \neq j$ の場合)

$$\begin{aligned}
\sum_i \bar{P}^{(n)}(j, i) P(i, k) &= \sum_i P(X_n = i, T_j > n, X_{n+1} = k | X_0 = j) \\
&= P(T_j > n+1, X_{n+1} = k | X_0 = j) = \bar{P}^{(n+1)}(j, k).
\end{aligned}$$

両辺 n で和を取ることで

$$\sum_{n=0}^{\infty} \sum_i \bar{P}^{(n)}(j, i) P(i, k) = \sum_{n=0}^{\infty} \bar{P}^{(n+1)}(j, k) = \mu(k)$$

が示される.

(ii) ($k = j$ の場合) 上記と同様の議論により

$$\begin{aligned} \sum_i \bar{P}^{(n)}(j, i) P(i, k) &= \sum_i P(X_n = i, T_j > n, X_{n+1} = j | X_0 = j) \\ &= P(T_j = n + 1 | X_0 = j) \end{aligned}$$

を得る. よって,

$$\sum_{n=0}^{\infty} \sum_i \bar{P}^{(n)}(j, i) P(i, k) = \sum_{n=0}^{\infty} P(T_j = n + 1 | X_0 = j) = 1 = \mu(i).$$

□

Theorem 10 (大数の強法則). マルコフ連鎖が既約かつ定常分布 π を持つとする^{*1}. $f : I \rightarrow \mathbb{R}$ が $\sum_j \pi(j) |f(j)| < \infty$ を満たすなら,

$$\frac{1}{n} \sum_{k=1}^n f(X_k) \xrightarrow{\text{a.s.}} \sum_{j \in I} \pi(j) f(j).$$

Proof. 既約かつ定常分布が存在するので, マルコフ連鎖は正再帰的である. よって, 定常分布は一意で, 任意の状態 i, j に対して $f(i, j) = 1$ である. ある状態 j を固定して, j を k 回目に訪れた時間を $T_j^{(k)}$ とし, $T_j^{(0)} = 0$ とする. このとき,

$$W_k = f(X_{T_j^{(k-1)}+1}) + f(X_{T_j^{(k-1)}+2}) + \cdots + f(X_{T_j^{(k)}})$$

とすれば, W_k ($k = 2, 3, \dots$) は独立同一分布に従う. なお,

$$Z_k = |f(X_{T_j^{(k-1)}+1})| + |f(X_{T_j^{(k-1)}+2})| + \cdots + |f(X_{T_j^{(k)}})|$$

とすると, $k = 2, 3, \dots$ に対して, Fubini の定理から

$$\begin{aligned} \mathbb{E}[|W_k|] &\leq \mathbb{E}[Z_k] = \sum_{n=1}^{\infty} \sum_{k=1}^n \mathbb{E}[f(X_n) | T_j = n, X_0 = j] P(T_j = n | X_0 = j) \\ &= \sum_{n=1}^{\infty} \sum_{k=1}^n \sum_{i \in I} |f(i)| P(X_k = i | T_j = n, X_0 = j) P(T_j = n | X_0 = j) \\ &= \sum_{i \in I} \sum_{k=1}^{\infty} \sum_{n \geq k} |f(i)| P(X_k = i | T_j = n, X_0 = j) P(T_j = n | X_0 = j) \\ &= \sum_{i \in I} |f(i)| \sum_{n=1}^{\infty} P(X_n = i, T_j \geq n | X_0 = j) \\ &= \sum_{i \in I} |f(i)| \pi(i) m(j, j) < \infty \quad (\because \text{Theorem 9}) \end{aligned}$$

^{*1} 周期性は必要ない.

となり, $E[|W_k|] < \infty$ がわかる. 同様に, 任意の i に対して $f(i, j) = 1$ なので $|W_1| < \infty$ (a.s.) もわかる. 上の議論から, $k \geq 2$ において

$$E[W_k] = E[W_1 | X_0 = j] = \sum_{i \in I} f(i) \pi(i) m(j, j)$$

である. j は再帰的なので, $N_j(n) \rightarrow \infty$ (a.s.) が成り立つので, 大数の強法則より,

$$\frac{1}{n} \sum_{k=1}^{N_j(n)} W_k = \frac{N_j(n) - 1}{n} \frac{1}{N_j(n) - 1} \sum_{k=2}^{N_j(n)} W_k + \frac{W_1}{n} \xrightarrow{\text{a.s.}} \frac{1}{m(j, j)} E[W_1 | X_0 = j] = \sum_j \pi(j) f(j)$$

となる.

一方で,

$$\sum_{k=1}^{N_j(n)} W_k = \sum_{m=1}^{T_j^{(N_j(n))}} f(X_m)$$

に注意する. ここで, $T_j^{(N_j(n))} \leq n < T_j^{(N_j(n)+1)}$ なので, 興味のある量 $\sum_{m=1}^n f(X_m)$ と $\sum_{m=1}^{T_j^{(N_j(n))}} f(X_m)$ の差は高々 W_k 一つ分である. あとは Lemma 11 から所望の結果が得られる. \square

Lemma 11. Z_1, Z_2, \dots が独立同分布で $P(Z_k \geq 0) = 1$ を満たし, $E[|Z_k|] = E[Z_k] < \infty$ であるとき,

$$\max_{1 \leq i \leq n} Z_i/n \rightarrow 0$$

が成り立つ.

Proof. $\epsilon > 0$ とする. $E[|Z_i|] < \infty$ より,

$$\sum_{i=1}^{\infty} P(Z_i/\epsilon > i) \leq \int_0^{\infty} P(Z_i/\epsilon > x) dx = E[Z_i/\epsilon] < \infty$$

である. よって, Borel-Cantelli の定理より有限個の i に対してのみ, $Z_i > \epsilon i$ が成り立つことがわかる. ϵ は任意なので, $Z_i/i \rightarrow 0$ である.

ここで, 任意の K に対して,

$$\max_{1 \leq i \leq n} Z_i/n \leq \max_{1 \leq i \leq K} Z_i/n + \max_{K < i \leq n} Z_i/i$$

が成り立つことに注意する. 任意の $\epsilon > 0$ に対して, K を十分大きな数とすると, 第二項は常に ϵ より小さくなる. また, 任意の固定した K に対しては第一項は 0 に収束する. \square