マルコフ連鎖入門 第2回

永幡幸生
新潟大学
nagahata@eng.niigata-u.ac.jp

2021 第1学期

Markov 連鎖

Markov 過程のうち特に

$$P(X_{n+1} = y | X_n = x) = P(X_1 = y | X_0 = x), \quad \forall n$$

を満たすものを時間的に一様な Markov 過程と呼ぶ。
以下では特に断らない限り、時間的に一様な Markov 過程を取り扱うことにする。
時間的に一様な Markov 過程において

$$P_{x,y} := P(X_1 = y | X_0 = x)$$

とおき (巨大なサイズの) 行列を考える。
この行列を推移確率もしくは推移確率行列と呼ぶ。

行列による表記？

$P(X_1 = y | X_0 = x)$ だけを考えたが一般に $P(X_{n+m} = y | X_m = x)$ はどうなるのか？$P_{x,y}$ もしくはこの行列を使って書けるのか？

Markov 連鎖

次のような計算をする。

$$
\begin{aligned}
&P(X_{n+m} = y | X_m = x) \\
&= P(X_{n+m} = y, \bigcup_{y_{n+m-1}} X_{n+m-1} = y_{n+m-1} | X_m = x) \\
&= \sum_{y_{n+m-1}} P(X_{n+m} = y, X_{n+m-1} = y_{n+m-1} | X_m = x) \\
&= \sum_{y_{n+m-1}} P(X_{n+m} = y | X_{n+m-1} = y_{n+m-1}) \\
&\qquad \times P(X_{n-1} = y_{n+m-1} | X_m = x) \\
&= \sum_{y_{n+m-1}} P(X_{n+m-1} = y_{n+m-1} | X_m = x) P_{y_{n+m-1},y}
\end{aligned}
$$

Markov 連鎖

以下帰納的に計算することで

$$
\begin{aligned}
&P(X_{n+m} = y | X_m = x) \\
&= \sum_{y_{n+m-1}} P(X_{n+m-1} = y_{n+m-1} | X_m = x) P_{y_{n+m-1}, y} \\
&\vdots \\
&= \sum_{y_{m+1}} \sum_{y_{m+2}} \cdots \sum_{y_{n+m-1}} P_{x, y_{m+1}} P_{y_{m+1}, y_{m+2}} \cdots P_{y_{n+m-1}, y} \\
&= (P^n)_{x,y}
\end{aligned}
$$

と行列 P の n 乗に帰着できる。

Markov 連鎖

推移確率は $P_{x,y} = P(X_1 = y | X_0 = x)$ で与えたので

$$
\begin{aligned}
0 \leq P_{x,y} \leq 1, &\quad \forall x, y \\
\sum_y P_{x,y} = 1, &\quad \forall x
\end{aligned}
$$

を満たしている。
逆にこの 2 つの条件を満たしていれば対応する Markov 過程が存在する。

Markov 連鎖

一般に $P(X_{n+m}=y|X_m=x)$ が推移確率行列 P の n 乗で書けるので、行列の計算、特に固有値、固有ベクトル、対角化が重要になる。

$$Q^{-1}PQ=\Lambda, \qquad P=Q\Lambda Q^{-1}$$

$$\Lambda=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$$

と書けたとき、全ての固有値が実数だと取り扱いやすい。
この十分条件として P が対称行列であることがあげられるが、これの拡張として

Markov 連鎖

定義 (π に対して対称)

$\pi = (\pi_1, \pi_2, \ldots, \pi_n) > 0$ に対して P が対称であるとは全ての x, y に対して以下を満たすこと

$$\pi_x P_{x,y} = \pi_y P_{y,x}$$

$$\Pi = \begin{pmatrix} \pi_1 & & \\ & \ddots & \\ & & \pi_n \end{pmatrix}, \quad \sqrt{\Pi} = \begin{pmatrix} \sqrt{\pi_1} & & \\ & \ddots & \\ & & \sqrt{\pi_n} \end{pmatrix}$$

等とおくとき

$$S = \sqrt{\Pi} P \sqrt{\Pi}^{-1}$$

とおくと

Markov 連鎖

$$S = \sqrt{\Pi} P \sqrt{\Pi}^{-1} = \sqrt{\Pi}^{-1} (\Pi P) \sqrt{\Pi}^{-1}$$

と書けるが右辺に出てくる 3 つの行列はすべて対称行列であるため S も対称行列になる。
S が実対称行列なので全ての固有値が実数で対角化できるので Λ を対角行列として $S = R \Lambda R^{-1}$ とすると

$$P = \sqrt{\Pi}^{-1} S \sqrt{\Pi} = (\sqrt{\Pi}^{-1} R) \Lambda (R^{-1} \sqrt{\Pi}) = (\sqrt{\Pi}^{-1} R) \Lambda (\sqrt{\Pi}^{-1} R)^{-1}$$

と書けるため P と S は同じ固有値を持っていることが分かる。

Markov 連鎖

逆に全ての固有値が実数でも、どのような π に対しても対称にならない例も作れる。

例

$$\begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}$$

の固有値は $1, 0, -1/2$ になる。

一方 $P_{x,y}, P_{y,x}$ のうち一方が 0 もう一方が 0 でない場合にはどのような $\pi > 0$ を持ってきても対称にすることは不可能。

$$\frac{1}{12}\begin{pmatrix} 2 & 8 & 2 \\ 5 & 2 & 5 \\ 5 & 2 & 5 \end{pmatrix}$$

の固有値も $1, 0, -1/4$ になり、計算するとやはりどのような $\pi > 0$ を持ってきても対称にすることは不可能。(後で解説)

Markov 連鎖

同値類 $\leftrightarrow$、既約

定義 ($\to$、$\leftrightarrow$)

推移確率行列 P が与えられているとき、関係 $\to$、$\leftrightarrow$ を次のように定義する；
(1) $x \to x$.
(2) $\exists n,\ x = x_0, x_1, x_2, \ldots, x_n = y$ s.t. $P_{x_{i-1},x_i} > 0,\quad 1 \le \forall i \le n$ ならば $x \to y$
(3) $x \to y$ かつ $y \to x$ ならば $x \leftrightarrow y$.

この定義では推移確率行列を用いたが一般に行列で定義してよい。

問題

(2) の条件は $\exists n$ s.t. $P(X_n = y | X_0 = x) > 0$ と同値であることを示せ。

Markov 連鎖

補題 (同値関係 $\leftrightarrow$)

関係 $\leftrightarrow$ は同値関係である。すなわち
(1) $x \leftrightarrow x$
(2) $x \leftrightarrow y$ ならば $y \leftrightarrow x$
(3) $x \leftrightarrow y$ かつ $y \leftrightarrow z$ ならば $x \leftrightarrow z$

問題

ほぼ自明ではあるが、この補題を証明せよ。

定義 (既約)

$\forall x, y (\in S)$ に対して $x \leftrightarrow y$ ならば状態空間 S は既約である。
また P が既約であるとも呼ぶ。

既約でない場合は $\leftrightarrow$ で分類される同値類毎に考えればよいので
以後既約の場合だけを考えることにする。

Markov 連鎖

周期

定義 (周期)

x 毎に $N(x) := \{n; (P^n)_{x,x} > 0\}$ とおき、$N(x)$ の最大公約数を d_x と書き、これを周期と呼ぶ。

命題 (周期)

$x \leftrightarrow y$ ならば $d_x = d_y$

Markov 連鎖

証明

$x \leftrightarrow y$ なので $x \to y$ かつ $y \to x$ なので

$$\exists n_1, n_2 \text{ s.t. } P(X_{n_1} = y | X_0 = x) > 0,\ P(X_{n_2} = x | X_0 = y) > 0$$

また $n \in N(x)$ とすると ($n = 0$ も可能であることに注意する)

$$P(X_n = x | X_0 = x) > 0.$$

従って

$$\begin{aligned} & P(X_{n_1+n+n_2} = y | X_0 = y) \\ & \geq P(X_{n_1+n+n_2} = y, X_{n+n_2} = x, X_{n_2} = x | X_0 = y) \\ & = P(X_{n_1} = x | X_0 = y) P(X_{n_2} = x | X_0 = y) P(X_{n_2} = x | X_0 = y) > 0 \end{aligned}$$

すなわち $n_1 + n + n_2 \in N(y)$ である。

Markov 連鎖

$n=0$ でも成立しているので $n_1+n_2 \in N(y)$ なので

$$\exists k \text{ s.t. } n_1+n_2 = kd_y$$

また $n \in N(x)$ に対して同様に $n_1+n+n_2 \in N(y)$ なので

$$\exists l \text{ s.t. } n_1+n+n_2 = ld_y$$

より

$$n = (l-k)d_y$$

と書ける。従って $d_x \geq d_y$ である。
今の議論は x, y を入れ替えても構わないので $d_y \geq d_x$ でもある。 $\square$

Markov 連鎖

補題

$\exists n_x$ s.t. $\forall k \geq n_x$, $kd_x \in N(x)$

問題

この補題を証明せよ。実質的には $d_x = 1$ の場合だけを考えればよい。

注意：$d_x = 1$ で $\#S = n$ とすると $n_x \leq n$ のような気がするが

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

とおくと $n_3 = 5 > 4 = n$ になる。

Markov 連鎖

系

状態空間 S は既約とする。このとき全ての点の周期は等しくなるので d とする。状態空間の番号付けを変更して推移確率行列を書き直す必要があるが
$\exists n$ s.t.

$$P^n = \begin{pmatrix} P_1 & & & \\ & P_2 & & \\ & & \ddots & \\ & & & P_d \end{pmatrix}$$

とブロック行列に分解できて $P_i > 0$ である。更に各 P_i, $1 \leq i \leq d$ は推移確率行列になる。

Markov 連鎖

注意：P_i のサイズが i によって異なっても構わない。例えば次は $d = 2,\ n = 2$ になるケース

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}, \quad P^2 = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Markov 連鎖

定常測度、不変測度、定常分布、不変分布

定義 (定常測度、定常分布)

S が有限集合 ($\#S = n$ とする) の場合は $\mu = (\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n})$、$S$ が可算集合の場合は $\mu = (\mu_{x_1}, \mu_{x_2}, \dots,)$ が $\forall x$ に対して $\mu_x \geq$ かつ $\mu \neq 0$ とする。
$\mu P = \mu$ を満たすものを定常測度、不変測度と呼び特に $\sum_x \mu_x = 1$ を満たすものを定常分布、定常確率測度、不変分布、不変確率測度と呼ぶ。

注意：定常測度を定数倍したものは定常測度になる。特に S が有限集合の場合は $\sum_x \mu_x$ で割って規格化して不変分布に変換することができる。

Markov 連鎖

P のサイズが大きい時には不変測度 μ を探すのは困難になるが以下の事実が使える時がある。

問題

以下を示せ
$\pi > 0$ に対して P が対称であれば π は P の定常測度になる。
($\pi \geq 0$ でも (対称の定義を拡張した上で) この主張は成立する。)

Markov 連鎖

$$\frac{1}{12}\begin{pmatrix} 2 & 8 & 2 \\ 5 & 2 & 5 \\ 5 & 2 & 5 \end{pmatrix}$$

の固有値も $1, 0, -1/4$ になり、計算するとやはりどのような $\pi > 0$ を持ってきても対称にすることは不可能。の解説
これを示すには

$$\pi_x P_{x,y} = \pi_y P_{y,x}, \quad x, y = a, b, c,$$

という方程式系を満たす自明でない π を探すことなのでこの方程式系に対応する行列を考えそのランクを考えるのが正統的。
一方で P を対称にする π があったとするとその π は不変測度であるが、不変測度は固有値 1 に対する左固有ベクトルでなければいけない。この例では固有値を先に計算していて固有値 1 は単純特性根なので固有ベクトル $\pi = (1, 1, 1)$ に対して対称になっているかを確かめればよい。

Markov 連鎖

測度 $\mu > 0$ に対して (重み付き) 内積 $\langle \cdot, \cdot \rangle_\mu$ を

$$\langle x, y \rangle_\mu = \sum_i x_i y_i \mu_i$$

とする。通常の内積は $\mathbf{1} = (1, 1, \ldots, 1)$ に対応する、すなわち
$x \cdot y = \langle x, y \rangle_{\mathbf{1}}$ と思える。
全ての x, y に対して $\langle x, Ay \rangle_\mu = \langle y, A^* x \rangle_\mu$ を満たす $A^* = A^*_\mu$ を
A の μ に対する共役行列と呼ぶことにする。

問題

$(A^*_\mu)_{i,j} = \mu_i^{-1} A_{j,i} \mu_j$ を示せ。
また A が μ 対称ならば $(A^*_\mu) = A$ を示せ。

この問題の前半部分から次のことが分かる。

命題 (P^* と定常測度)

P を確率行列とする。P^*_μ が再び確率行列になるのは μ が P の定常測度の時のみ。

証明

先の問題から $(P^*_\mu)_{i,j} = \mu_i^{-1} P_{j,i} \mu_j$ と書けているので $(P^*_\mu)_{i,j} > 0$ は分かる。

$\forall i,\ \sum_j (P^*_\mu)_{i,j} = 1$ となるのが μ が定常測度の時だけであることをいえばよい。

$$\sum_j (P^*_\mu)_{i,j} = \sum_j \mu_i^{-1} P_{j,i} \mu_j = \mu_i^{-1} \sum_j \mu_j P_{j,i}$$

であることに注意すれば μ が定常測度の場合は明らかに成り立っており、定常測度でない場合は $\sum_j \mu_j P_{j,i} \neq \mu_i$ となる i が少なくとも 1 つあるので成り立たない。 □