

マルコフ連鎖まとめ

@skbtkey

概要

この文書は卒研セミナーでのマルコフ連鎖のセミナーについて、毎回その要点をまとめるものである。気を付けなければいけないと思ったことは詳しく書く。また僕自身にとって簡単であるようなことは省略する。なお用いる教科書は Paul G. Hoel, Sidney C. Port, Charles J. Stone 著の Introduction to Stochastic Processes.

1 マルコフ鎖の定義と簡単な性質, いくつかの例

*1

1.1 マルコフ性の定義

\mathcal{J} は有限個の状態の集合とする。 \mathcal{J} のことを状態空間とよぶ。 離散時間 $n = 0, 1, 2, \dots$ が与えられていて X_n はそれぞれの時刻 n における状態を表すものとする。 ($n \geq 0$). この状態を数値に対応させることで X_n はある確率空間上の確率変数とみなす。

定義 1.1 (マルコフ性). 確率空間上 (確率測度は \mathbb{P}) で定義された確率変数列 $(X_n)_{n \in \mathbb{N}}$ がマルコフ性を満たすとは、次を満たすことである。

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \quad (1)$$

マルコフ性の式を言葉で説明してみよう。 離散時間が与えられていてそれぞれの時刻で様々な状態を取るというモデルがある。 計測開始時点から現在までの状態が分かっているとする。 このとき、次の計測時点での状態が x_{n+1} であるような確率は、現在の状態のみに依存する。 現在の状態さえわかっているならば過去の状態は一つ次の未来がどんな状態であ

*1 Markov Chain の訳は本によってさまざま。 マルコフ系列だったりマルコフ過程だったりします。 ここでは読んでいる教科書の用語を直訳することを意識してマルコフ鎖, マルコフ連鎖とします。

るのかという確率には影響を与えない。

定義 1.2. 条件付き確率 $\mathbb{P}(X_{n+1} = y | X_n = x)$ のことを鎖の遷移確率 (transition probabilities of the chain) という。

この文書では変化しない遷移確率 (stationary transition probability) を持つマルコフ鎖を考える。すなわち、 $\mathbb{P}(X_{n+1} = y | X_n = x)$ が n に依存しないようなマルコフ鎖のみを考える。 n に依存しないとは、もっと具体的に言うと、

$$\mathbb{P}(X_1 = y | X_0 = x) = \mathbb{P}(X_2 = y | X_1 = x) = \cdots = \mathbb{P}(X_{n+1} = y | X_n = x)$$

が成り立つことを意味する。

定義 1.3. (マルコフ鎖の定義) 確率変数列 $(X_n)_{n \in \mathbb{N}}$ がマルコフ鎖をなすとは、これらの確率変数たちがマルコフ性を持ち、変化しない遷移確率をもつときにいう。

1.2 二つの状態をもつマルコフ鎖

表 1 機械の状態と遷移の確率の様子

nth day	遷移確率	(n + 1)th day
正常	\xrightarrow{q}	故障
	$\xrightarrow{1-q}$	正常
故障	\xrightarrow{p}	正常
	$\xrightarrow{1-p}$	故障

ある機械は 1 日の始まりに正常に動くか故障しているかチェックされる。表のように n 日目の状態と $n + 1$ 日目の状態の遷移の確率が定義されている。 $0 \leq p, q \leq 1$ である。機械が正常である時を 1, 故障しているときを 0 とする。 $(X_n)_{n \in \mathbb{N}}$ を n 日目の機械の状態を表すとする。このようにすると $(X_n)_{n \in \mathbb{N}}$ はマルコフ性を持つ確率変数であると考えられる (あくまでそんな気がする!)。

$$\mathbb{P}(X_{n+1} = 1 | X_n = 0) = p$$

$$\mathbb{P}(X_{n+1} = 0 | X_n = 1) = q$$

$$\mathbb{P}(X_0 = 0) =: \pi_0(0)$$

とおくと、 n 日目の状態が $x_n (= 0 \text{ or } 1)$ である確率を計算することができる。

$$\begin{aligned} \mathbb{P}(X_{n+1} = 0) &= \mathbb{P}(X_n = 0 \text{ and } X_{n+1} = 0) + \mathbb{P}(X_n = 1 \text{ and } X_{n+1} = 0) \\ &= \mathbb{P}(X_n = 0)\mathbb{P}(X_{n+1} = 0|X_n = 0) + \mathbb{P}(X_n = 1)\mathbb{P}(X_{n+1} = 0|X_n = 1) \\ &= \dots \\ &= (1 - p - q)\mathbb{P}(X_n = 0) + q \end{aligned}$$

というふうに漸化式が得られるので、一般項を求める計算をすると、

$$\mathbb{P}(X_n = 0) = (1 - p - q)^n \pi_0(0) + q \sum_{j=0}^{n-1} (1 - p - q)^j \quad (2)$$

総和を計算すると、

$$\mathbb{P}(X_n = 0) = \frac{q}{p + q} + (1 - p - q)^n \left(\pi_0(0) - \frac{q}{p + q} \right) \quad (3)$$

が得られる。したがって、

$$\mathbb{P}(X_n = 1) = \frac{p}{p + q} + (1 - q - p)^n \left(\pi_0(1) - \frac{p}{q + p} \right) \quad (4)$$

ここで、 $p = q = 1$ でも $p = q = 0$ でもないとする。すなわち、 $|1 - p - q| < 1$ であるから、 $n \rightarrow \infty$ とすると、

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0) = \frac{q}{p + q}$$

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 1) = \frac{p}{p + q}$$

この機械の例では、先ほども述べたように、その日の機械の状態を指示している確率変数 $(X_n)_{n \in \mathbb{N}}$ が本当にマルコフ性を持っているかどうかはわからない。実際、今までの計算でマルコフ性を用いてはいない。ここではこれらの確率変数たちがマルコフ性を持つと仮定して話を進める。すると、次のような計算が可能になる。 $n = 2, x_0, x_1, x_2$ は $1, 0$ のいずれかであるとき、

$$\begin{aligned} &\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2) \\ &= \mathbb{P}(X_0 = x_0 \text{ and } X_1 = x_1)\mathbb{P}(X_2 = x_2|X_0 = x_0 \text{ and } X_1 = x_1) \\ &= \mathbb{P}(X_0 = x_0)\mathbb{P}(X_0 = x_0|X_1 = x_1)\mathbb{P}(X_2 = x_2|X_1 = x_1) \end{aligned}$$

マルコフ性が無ければこの計算の 2 行目から 3 行目への変形はできない。それが意味するのは、 $\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2)$ は p, q, π_0 を用いて表すことができないということである。

1.3 遷移関数と初期分布

$(X_n)_{n \in \mathbb{N}}, n \geq 0$ を状態空間 \mathcal{J} を持つマルコフ鎖とする。関数 $P(x, y) (x, y \in \mathcal{J})$ を次で定める。

$$P(x, y) = \mathbb{P}(X_1 = x | X_0 = y) [x, y \in \mathcal{J}] \quad (5)$$

これを鎖の遷移関数と呼ぶ。当然,

$$P(x, y) \geq 0 \quad (6)$$

$$\sum_y P(x, y) = 1 \quad (7)$$

である。この文書ではマルコフ鎖は変化しない遷移確率をもつとしていたので、すなわち,

$$\mathbb{P}(X_{n+1} = x | X_n = y) = P(x, y) [n \geq 1] \quad (8)$$

が成立する。これと、マルコフ性より,

$$\mathbb{P}(X_{n+1} = x | X_0 = x_0, \dots, X_n = y) = P(x, y) \quad (9)$$

が成り立つ。

また、関数 π_0 を次で定義する。

$$\pi_0(x) = \mathbb{P}(X_0 = x) [x \in \mathcal{J}] \quad (10)$$

次が成り立つことは明らかである。

$$\pi_0(x) \geq 0 \quad (11)$$

$$\sum_x \pi_0(x) = 1 \quad (12)$$

これらの関数を用いるとマルコフ性を特徴づけすることができる。先ほども書いたように,

$$\begin{aligned} & \mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2) \\ &= \mathbb{P}(X_0 = x_0) \mathbb{P}(X_1 = x_1 | X_0 = x_0) \mathbb{P}(X_2 = x_2 | X_1 = x_1) \end{aligned}$$

であるから,

$$\begin{aligned}\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2) \\ = \pi_0(x_0)P(x_1, x_0)P(x_2, x_1)\end{aligned}$$

これを帰納的に計算することで, 次を得る.

$$\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) = \pi_0(x_0)P(x_1, x_0)P(x_2, x_1) \cdots P(x_n, x_{n-1}) \quad (13)$$

ここで定義をまとめておく.

定義 1.4 (遷移関数と初期分布). 関数 $P(x, y) [x, y \in \mathcal{J}]$ が鎖の遷移関数であるとは, (6), (7) を満たすことである. 関数 $\pi_0(x) [x \in \mathcal{J}]$ が初期分布であるとは, (11), (12) を満たすことである.

コロモゴロフの拡張定理より, 遷移関数と初期分布が与えられると, ある確率空間とその空間の上で定義された確率変数列が式 (13) を満たすようなものが存在することを示すことができる.*²

*² 後で書く