

マルコフ連鎖入門 第1回

永幡幸生
新潟大学

nagahata@eng.niigata-u.ac.jp

2021 第1学期

Markov 連鎖

Markov 連鎖は確率過程の中でも取り扱いやすく、現実問題への応用という観点から考えると、物理現象との相性が良い。

物理現象と呼ばれると少し違和感を感じるかもしれないが、コロナ関係で「8割削減」というキーワードが出てきたが、この数字は「SIR モデル」と呼ばれる感染症の分野で基本的に使われる Markov 連鎖を用いて導き出した値である。

取り扱いやすいと言ってもやはり奥が深く、いろいろな取り扱ひ方があるが、一つの方法として線形代数をよく知っていると分かってくることがある。

応用の観点から Markov 連鎖の族を考えそのスケーリング (時空間に関するスケール極限) を考えたい。このスケーリングをどう決めればよいのかということの一つの指標がスペクトルギャップもしくははスペクトルギャップのオーダーである。また逆にこのオーダーの評価があるとその後の解析ができるようになる例がある。この講義ではなるべく線形代数の観点からこれらおよび関連する話をする。

この講義では条件付確率は高校の数 I の範囲の割り算で定義するもの、すなわち

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

で与えられるもので十分であるが、 $P(B) = 0$ になるところでも (ダミーで) 定義されているものと思う。但し基本的には

$P(B) = 0$ になるような時には $P(A|B)$ が単独で表れることはなく必ず $P(B)P(A|B)$ のように 0 になる部分が入っており、表記を簡単にするために導入するだけである。

表記を簡単にするため $P(A, B) = P(A \cap B)$ 等と $\cap$ を省略して書くことにする。

問題

$P(B, C) > 0$ のとき $P(A, B|C) = P(A|B, C)P(B|C)$ を示せ。

Markov 連鎖

Markov 過程は確率過程の特別なものである。

確率過程とは時間パラメータの入った確率変数であり

$(X_n)_{n \geq 0}$ を確率過程とした時 $X_n \in S$ となる S のことを状態空間と呼ぶ。

通常確率過程 $(X_n)_{n \geq 0}$ に対して X_n の確率法則はそれ以前の状態 $X_0, X_1, \dots, X_{n-1}$ に依存する。

(離散時間)Markov 過程とは次の Markov 性を満たす確率過程のことである

$P(X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_1 = x_1, X_0 = x_0) > 0$ であれば

$$\begin{aligned} &P(X_n = x_n, | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_1 = x_1, X_0 = x_0) \\ &= P(X_n = x_n | X_{n-1} = x_{n-1}) \end{aligned}$$

すなわち X_n の確率法則が 1 歩前の X_{n-1} の状態だけで決定されるものである。

Markov 連鎖

用語として **Markov 過程** と **Markov 連鎖** とあるが本講義では使い分けをしないが以下の設定では **Markov 連鎖** のみを取り扱っている。数学辞典によると **Markov 過程** は状態空間を (d-次元) 実数空間もしくは局所コンパクト Hausdorff 空間で第2可算公理を満たすものとし、**Markov 連鎖** はこのうち状態空間が高々可算集合のものである。

本講義では基本的には状態空間は有限集合を考えるが、可算集合の場合でも成り立つ主張が多くある。同時に特に線形代数、行列の話で有限集合、有限サイズの行列に限った話も出てくる。状態空間が可算集合の場合にも成り立つ主張を状態空間が有限集合の場合に適用すると自明な結果しか得られないことが多々ある。

Markov 連鎖

(離散時間)Markov 過程は X_n の確率法則が 1 歩前の X_{n-1} の状態だけで決定されるものであるが次のように考えるとランダムウォークの自然な拡張と思える。

ランダムウォーカーはコイン (サイコロ) を投げて出た面 (目) によって進む方向を決めて行く。特にこのコイン (サイコロ) はランダムウォーカーの持ち物でありいつも同じものを使っている。

一方で Markov 過程ではウォーカーは自分自身のコイン (サイコロ) をもっておらず、そのたどり着いた場所に備え付けのコイン (サイコロ) を使っている。そのためコイン (サイコロ) 達が場所によって違うことを考える必要がある。

Markov 連鎖

(人間の考える) 多くの確率過程は (状態空間を大きく取れば) Markov 過程になる。一方でその Markov 過程の一部分だけを取り出して確率過程と捉えると Markov 過程にならない (Markov 性を持たない) 例が多くある。

例 倍賭け法のもうけ

まず 2 次元の確率過程 $(X_n, Y_n)_{n \geq 0}$ として勝つか負けるかだけが確率的に決まることに次のような戦略で賭けを行なった時の時刻 n でのもうけを X_n , 次にいくら賭けるかを Y_n で表すものとする。

● 勝てば掛け金を 2 倍に、負ければ 1 にリセットする。

問題を単純にするため $X_0 = 0$, $Y_0 = 1$ とし、毎時独立で表が出る確率が $0 < p < 1$ のコインを投げ、表が出た時に勝ちとする。

$$(X_{n+1}, Y_{n+1}) = \begin{cases} (X_n + Y_n, 2Y_n) & \text{表が出た場合：確率 } p \\ (X_n - Y_n, 1) & \text{裏が出た場合：確率 } 1 - p \end{cases}$$

と解釈する。

この設定では当然 $(X_n, Y_n)_{n \geq 0}$ は Markov 過程である。

さらに $(Y_n)_{n \geq 0}$ だけを見ても、 Y_{n+1} は Y_n の 2 倍になるか (確率 p)、1 になるか (確率 $1 - p$) だけなので Markov 過程である。

Markov 連鎖

一方コインが「表表裏」「裏裏表」と出た場合を考えると
 $(X_n, Y_n)_{n=0,1,2,3}$ は

	(X_0, Y_0)		(X_1, Y_1)		(X_2, Y_2)		(X_3, Y_3)
表表裏	$(0, 1)$	$\rightarrow$	$(1, 2)$	$\rightarrow$	$(3, 4)$	$\rightarrow$	$(-1, 1)$
裏裏表	$(0, 1)$	$\rightarrow$	$(-1, 1)$	$\rightarrow$	$(-2, 1)$	$\rightarrow$	$(-1, 2)$

と推移するので共に $X_3 = -1$ である。

しかし

$$P(X_4 = k | X_3 = -1, X_2 = 3, X_1 = 1, X_0 = 0) = \begin{cases} p & k = 0 \\ 1 - p & k = -2 \\ 0 & \text{その他} \end{cases}$$

$$P(X_4 = k | X_3 = -1, X_2 = -2, X_1 = -1, X_0 = 0) = \begin{cases} p & k = 1 \\ 1 - p & k = -3 \\ 0 & \text{その他} \end{cases}$$

であるため $(X_n)_{n \geq 0}$ だけみて Markov 過程になるのであれば

$$\begin{aligned} & P(X_4 = k | X_3 = -1, X_2 = 3, X_1 = 1, X_0 = 0) \\ &= P(X_4 = k | X_3 = -1) \\ &= P(X_4 = k | X_3 = -1, X_2 = -2, X_1 = -1, X_0 = 0) \end{aligned}$$

であるはずであるが明らかに成立していない。

例 Markov 過程とは思えないが Markov 過程に帰着できる例
(ランダムな) 力学モデルの離散化として $V_n = X_n - X_{n-1}$ を速度
(の対応物) ととらえ、 X_{n+1} の確率法則は X_n, V_{n-1} で与えられる
ようなモデルを考えるがこれは X_{n+1} の確率法則は X_n, X_{n-1} で与
えられるようなモデルと考え直すことができる。
例として

$$\begin{aligned} X_{n+1} &= X_n + f(X_n, V_n) + \text{ランダム項} \\ &= X_n + f(X_n, X_n - X_{n-1}) + \text{ランダム項} \end{aligned}$$

を考えるが、ランダム項は X_n, X_{n-1} の値には依存してもよい。
このとき X_n は Markov 性を満たさないが
 $Y_n := (X_n, V_n) = (X_n, X_n - X_{n-1})$ とおけば $(Y_n)_n$ が Markov 性を
持つことを確認することは容易である。

Markov 連鎖は物理現象との相性が良いため、確率的な挙動を示す現象に対して対応する Markov 連鎖モデルを考え、その解析を行うことが多い。

取り扱いやすいとは言ったが、実際のモデルを考えた場合は複雑なものになる場合が多く、数学として抽象的なことは分かるが、実際の値が分からないことは多く、そのモデルの物理量の値、もしくはその近似値を得るためにはコンピュータを用いて、モンテカルロシミュレーションを行うことが多い。

次にあげる命題は、モデルが与えられた場合には必ずモンテカルロシミュレーションを行うことができると主張していて、さらにその証明自身がプログラムを与えていると思えるものである。

定義 (a random mapping representation)

推移確率行列 P の関数 f と確率変数 Z による表現とは、関数 $f : S \times \Lambda \rightarrow S$ と Λ を状態空間とする確率変数 Z (その確率を Q とする) を用いて

$$Q(\{f(x, Z) = y\}) = P_{x,y}$$

を満たすこととする。

命題 (a random mapping representation)

状態空間 S を有限集合とすると、推移確率行列 P はある関数 f と、ある確率変数 Z による表現が可能である。

先にも挙げたが、実際にコンピュータシミュレーションを行なう上ではこの命題は重要であり、この命題 (の証明) に従ってプログラムを書いていると言っても過言ではない。

証明でも用いるが $\Lambda = [0, 1]$ として Z を $[0, 1]$ 上の一様乱数として実現するのが有効。

証明

$S = \{s_1, s_2, \dots, s_n\}$ として $\Lambda = [0, 1]$ とし Z を $\Lambda = [0, 1]$ 上の一様分布に従う確率変数とする。

$\{F_{j,k}; 1 \leq j \leq n, 0 \leq k \leq n\}$ を $F_{j,0} = 0, F_{j,k} = \sum_{i=1}^k P_{s_j, s_i}$ ととり $F_{j,k-1} < z \leq F_{j,k}$ のとき $f(s_j, z) := s_i$ とする。但し $f(s, 0) = s_1$ とする。

このとき明らかに

$$Q(\{f(x, Z) = y\}) = P_{x,y}$$

を満たす。



参考文献

少なくとも以下の教科書などは参考になっている。但し、ずいぶん前に読んで勉強した教科書の内容をそれと認識せずに参考になっていることもありえるのでその点に関してはご容赦頂きたい。

Markov 過程の基本的なところでは [Ko] の教科書の該当部分を多く参考になっている。

参考文献

- [AF] Aldous D.; Fill J.A.
Reversible Markov Chains and Random Walks on Graphs,
(未刊の教科書 現在 (2021/1) も 「Aldous Fill Markov chain」 で検索をすれば pdf、html バージョンを発見できる)
- [Ka] N. Kahale, Nabil; A semidefinite bound for mixing rates of Markov chains. Proceedings of the Workshop on Randomized Algorithms and Computation (Berkeley, CA, 1995). Random Structures Algorithms 11 (1997), no. 4, 299313.
- [Ko] 小谷真一, 測度と確率, 岩波書店
- [LPW] Levin D.A.; Peres Y.; Wilmer E.L.
Markov Chains and Mixing Times, AMS
- [Sai] 斎藤正彦, 線形代数入門, 東京大学出版会

参考文献

[Sal] Saloff-Coste L.

Lectures on Finite Markov Chains, Lecture on Probability 1665

[Se] Seneta E.

Non-negative Matrices and Markov Chains, Springer

[V] Varga R.S.

Matrix Iterative Analysis, Springer