

マルコフ連鎖と定常分布

長谷部高広

マルコフ連鎖は確率過程論の中で学びやすい対象でありながらも多くの応用を持つ。筆者は学生の頃に統計物理学のモデルでマルコフ連鎖を使ったことがある。意外なところではトランプを何回シャッフルすれば十分に混ざるのかといった研究もある（6節で紹介する）。本稿ではマルコフ連鎖において不動点の役割を果たす定常分布という概念に焦点を当てて、特に定常分布の分類と定常分布への収束を中心に解説する。マルコフ連鎖を全般的にきちんと学習したい読者には文献^{3, 6, 7, 10, 12, 13}を挙げておく。他にも文献⁴は豊富な例を含んでおり学習に役立つだろう。

1. 高校数学で現れる確率過程の例

高校数学の中で既に色々な確率過程が現れている。読者は次のような問題に見覚えがあるかもしれない。図1のような正四面体の頂点 a, b, c, d の上を点 X が以下のルールにより移動していくとする：(α) X は時刻0で頂点 a にいる、(β) X が時刻 n においてある頂点にいるとき、時刻 $n+1$ では他の三つの頂点のいずれかにそれぞれ確率 $\frac{1}{3}$ で移動する。基本的な問題として、時刻 n で頂点 a, b, c, d にいる確率をそれぞれ求める問題がある。この例では、いわゆる確率漸化式を立ててこれらの確率を求めることができる。ただし、大学以降の数学では確率漸化式にはあまり頼らない。というのも、確率漸化式の代わりに行列を用いる方が様々な面で便利だからだ。このことは本稿を読み進めてもらうといくぶんか理解できるだろう。

実は、上で考えた確率過程 X は本稿の主題となるマルコフ連鎖に近いが、まだマルコフ連鎖とは言えない。理由の一つとして X の「作り方」が明確でないことが挙げられる。詳しくは2節で論じる。

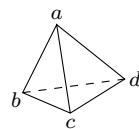


図1 正四面体

2. マルコフ連鎖

1節の例をもう少し一般化・数式化してみる。 X が取りうる値の集合（状態空間）を S とする。1節の例だと $S = \{a, b, c, d\}$ である。本稿ではごく一部を除いて S は有限集合とする。一般的に S の要素を状態と呼ぶが、本稿では点あるいは頂点と呼ぶこともある。

まず、非負の実数 $\mu(i), p(i, j) \geq 0$ ($i, j \in S$) で以下の条件を満たすものを用意する：

$$\sum_{i \in S} \mu(i) = 1, \quad \sum_{j \in S} p(i, j) = 1 \quad (i \in S).$$

$\mu(i), p(i, j)$ たちをそれぞれまとめて $\mu = (\mu(i))_{i \in S}$, $P = (p(i, j))_{i, j \in S}$ と表し、 μ を分布、 P を推移確率という。点 X が以下のルールで S 上を動くとする：全ての $i, j \in S$ と整数 $n \geq 0$ について、

- 時刻0で X が i にいる確率は $\mu(i)$ である、すなわち

$$\mathbb{P}[X_0 = i] = \mu(i), \quad (1)$$

- 時刻 n で X が i にいるとき、時刻 $n+1$ で X が j に移動する確率は $p(i, j)$ である、すなわち

$$\mathbb{P}[X_{n+1} = j | X_n = i] = p(i, j). \quad (2)$$

ここで X_n は時刻 n における X の位置を表し、 $\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$ は条件付き確率を表している。た

だし、条件付き確率は分母が 0 になると意味をなさないため、正確には $\mathbb{P}[X_n = i] > 0$ を満たす i と n に対してのみ条件 (2) を考える。

条件 (1), (2) はそれぞれ 1 節の $(\alpha), (\beta)$ の一般化になっている。1 節の例では、初期分布は $\mu(a) = 1, \mu(b) = \mu(c) = \mu(d) = 0$ で与えられる。このような場合、確率過程 X は a から出発すると言う。また推移確率は

$$p(i, j) = \begin{cases} 0, & i = j, \\ \frac{1}{3}, & i \neq j \end{cases} \quad (3)$$

となっている。

実は、条件 (1), (2) だけでは色々な基本的確率が計算できない。例えば、確率 $\mathbb{P}[X_0 = i, X_1 = j, X_2 = k]$ ^{*1)} の値は μ, P のみでは決まらない。そこで、(2) より強い次の条件を考える：全ての $i_0, i_1, \dots, i_{n+1} \in \mathcal{S}, n \geq 0$ に対して^{*2)}

$$\mathbb{P}[X_{n+1} = i_{n+1} | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0] = p(i_n, i_{n+1}). \quad (4)$$

条件 (1) と (4) を満たす X のことを初期分布 μ と推移確率 P を持つマルコフ連鎖と呼ぶ。推移確率が時刻 n に依存するマルコフ連鎖も応用上重要なことがあるが、簡単のため今後は n に依存しない場合（時間的一様という）のみを考える。なお、マルコフ連鎖はよくランダムウォークと呼ばれることがある。どのような場合にそう呼ばれるのか明確なルールはなさそうだが、筆者の見たところ、次の定義が「ランダムウォーク」という用語の実際の使われ方をうまく表現している：状態空間にグラフ構造や群構造など何らかの数学的構造があるときに、その構造に適したマルコフ連鎖をランダムウォークと呼ぶ。^{14, 序文)}

条件 (4) から条件 (2) が導かれることは以下のようにして分かる。簡単のため $n = 1$ の場合を考える。条件 (4) は $j = i_2, i = i_1$ と置くと

$$\begin{aligned} \mathbb{P}[X_2 = j, X_1 = i, X_0 = i_0] \\ = p(i, j) \mathbb{P}[X_1 = i, X_0 = i_0] \end{aligned}$$

となる。両辺で i_0 について総和を取って $\mathbb{P}[X_1 = i]$ で

*1) コンマによる区切りは「かつ」と読む。つまりこれは「 $X_0 = i$ かつ $X_1 = j$ かつ $X_2 = k$ となる確率」を意味する。

*2) 正確には、条件 (2) の場合と同様に、ここでも分母が 0 にならない場合のみを考える。つまり $\mathbb{P}[X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0] > 0$ を満たすような $i_0, i_1, \dots, i_n \in \mathcal{S}$ のみを考える。

割ると (2) が得られる。

条件 (2) と (4) を合わせると、等式

$$\begin{aligned} \mathbb{P}[X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_0 = i_0] \\ = \mathbb{P}[X_{n+1} = i_{n+1} | X_n = i_n] \end{aligned} \quad (5)$$

が得られる（マルコフ性）。これはマルコフ連鎖が将来どのように変化するかについての確率が現時刻 n での状態のみで決まり、過去の経過に依存しないという性質を表している。等式 (5) のみでマルコフ連鎖の定義としている文献もあるが、これだと推移確率 $p(i, j)$ を導入しにくいので、ここではそういった定義は採用しないことにした。条件付き確率は分母が 0 になると意味をなさないという点がやっかいである。

マルコフ連鎖に対しては、基本的に $X_0, X_1, \dots$ を用いて表される事象の確率は全て μ と P で決定される。このため、マルコフ連鎖はさまざまな確率計算がしやすい確率過程となっている。例えば $\mathbb{P}[X_0 = i, X_1 = j, X_2 = k]$ は条件付き確率の定義により $\mathbb{P}[X_2 = k | X_1 = j, X_0 = i]$, $\mathbb{P}[X_1 = j | X_0 = i]$, $\mathbb{P}[X_0 = i]$ の積と等しいことが分かるから、その値は $\mu(i)p(i, j)p(j, k)$ と計算される。同様にして、全ての $i_0, i_1, \dots, i_n \in \mathcal{S}, n \geq 0$ に対して

$$\begin{aligned} \mathbb{P}[X_0 = i_0, X_1 = i_1, \dots, X_n = i_n] \\ = \mu(i_0)p(i_0, i_1)p(i_1, i_2) \cdots p(i_{n-1}, i_n) \end{aligned} \quad (6)$$

が成り立つ。ただし $n = 0$ のときの条件 (6) は条件 (1) を意味すると解釈する。実は条件 (6) は (1) かつ (4) と同値になる。文献³⁾などでは条件 (6) をマルコフ連鎖の定義としている。この定義では条件付き確率が現れないため、分母の確率が 0 にならないかどうかを気にする必要はない。

実際にマルコフ連鎖を「作る」にはどうしたらよいだろうか？ 人が物理的な地点から地点へ次々に移動するとしよう。各地点 i にくじ引き L_i を用意しておく。このくじ引きでは確率 $p(i, j)$ で j という結果が出るとする。^{*3)} また初期分布 μ についても同様に、確率 $\mu(i)$ で i が出るようにくじ引き L を用意する。以上の準備をしたところで、ある人物 X が以下のルールで移動する：まずくじ引き L を実施し、出た結果が i であれば地点 i に移動し、そこを時刻 0 での出発点とす

*3) 確率論では無理数の確率も許容するのが一般的だ。しかし、 $p(i, j)$ が無理数の場合はくじ引きを作るのは困難であろう。その点が気になる読者は $p(i, j)$ は有理数と考えておくとうい。

る ($X_0 = i$)。続いて地点 i でくじ引き L_i を行い、出た結果が j であれば j に移動し ($X_1 = j$)、また移動先の j でくじ引き L_j を行うという具合で進んでいけば、初期分布 μ と推移確率 P を持つマルコフ連鎖になる (特に条件 (4) が成り立つ)。実は、この「作り方」で大切なのは異なる時刻におけるくじ引き結果が「独立」なことである。要するに、過去の経緯と独立に次の行き先を確率的に決めていくのがマルコフ連鎖というわけだ。条件 (2) や 1 節の条件 (β) ではこの独立性を十分に表現できていない。以上の「作り方」は文献⁵⁾の “random mapping representation” や文献¹²⁾の式 (2.1.3) とおおよそ同じ考え方に基づいている。

3. 行列としての推移確率

有限状態空間 $\mathcal{S}$ 上の初期分布 μ 、推移確率 P を持つマルコフ連鎖 X に対して $\mathbb{P}[X_n = i]$ を計算しよう。状態に番号をつけておくとう便利なので、以降では $\mathcal{S} = \{1, 2, \dots, s\}$ としよう (s は自然数)。このように仮定しても一般性を失わないことは明らかだろう。^{*4)} まず連立の確率漸化式

$$\begin{aligned}\mathbb{P}[X_{n+1} = j] &= \sum_{i=1}^s \mathbb{P}[X_{n+1} = j | X_n = i] \mathbb{P}[X_n = i] \\ &= \sum_{i=1}^s \mathbb{P}[X_n = i] p(i, j)\end{aligned}\quad (7)$$

が成り立つ ($1 \leq j \leq s, n = 0, 1, 2, \dots$)。これを行列を用いて書き換えてみる。時刻 n における確率を並べた横ベクトル (X_n の分布という) を

$$\mu_n = (\mathbb{P}[X_n = 1], \mathbb{P}[X_n = 2], \dots, \mathbb{P}[X_n = s])$$

と表す。また推移確率 $P = (p(i, j))_{i,j=1}^s$ を (i, j) 成分が $p(i, j)$ であるような s 次正方行列とみなすことにする (推移確率行列)。すると (7) は

$$\mu_{n+1} = \mu_n P, \quad n = 0, 1, 2, \dots \quad (8)$$

^{*4)} $\mathcal{S}$ の番号付けの仕方によって推移確率行列の形 (行列表示) は変わる。番号付けを変えることは線形代数学で言うところの基底の変換になる。番号付けをすることで対称性を壊してしまって、かえって行列の形が複雑になってしまう恐れもある。よって番号付けをしない方が便利の場合もあり、このときは、 P を行列ではなく $C(\mathcal{S})$ における線型写像として扱うことができる。ただし $C(\mathcal{S})$ は $\mathcal{S}$ 上の実数値関数全体のなすベクトル空間である。

と書き換えることができて見通しが良い。ただし $\mu_0 = \mu$ は初期分布とする。この関係式を繰り返し用いると、 $\mu_n = \mu P^n$ が得られる。従って、 $\mathbb{P}[X_n = i]$ の計算は推移確率行列 P の n 乗を求める問題に帰着される。行列のベキ乗は対角化や Jordan 標準形によって計算することが可能である。

ところで、導出を見れば分かるように、確率漸化式 (7) が成り立つためには X はマルコフ連鎖である必要はなく、条件 (1), (2) さえあれば良い。従って、大学入試問題などで条件 (1), (2) のみを考えているような場合は、 X はマルコフ連鎖とまでは言えないが、 μ_n についての確率漸化式を解く上では特に問題はないと言える。

4. 定常分布

確率 $\mathbb{P}[X_n = i]$ の計算は推移確率行列のベキ乗の計算に帰着されるが、状態空間の要素の個数が増えてくると容易ではない。そこで、より易しい (?) 問題として、極限值

$$\lim_{n \rightarrow \infty} \mathbb{P}[X_n = i]$$

が計算できないかどうかを考えてみる。まず、仮にこの極限值が全ての $i \in \mathcal{S}$ に対して存在すると仮定してみると、横ベクトルとしての極限值

$$\pi := \lim_{n \rightarrow \infty} \mu_n \quad (9)$$

が存在することになる。等式 (8) の極限を取ると

$$\pi = \pi P \quad (10)$$

という等式が得られる。^{*5)}つまり、 π は P を横ベクトルに右からかける操作 (写像) の不動点になっている。

ここでいったん μ_n の極限ということは忘れて、等式 (10) を満たす任意の分布 π を定常分布と呼ぶ。等式 (10) は s 次の連立一次方程式であるため、それほど大きくない s の場合には手計算で解くことも十分可能である。1 節の例、つまり (3) で与えられる P について、 $a = 1, b = 2, c = 3, d = 4$ と考えて $\pi = \pi P$ を

^{*5)} 数学では写像としての正方行列や長方形行列を縦ベクトルに左からかける記法が圧倒的に主流である。しかしマルコフ連鎖の分野では、(8) や (10) のように、横ベクトルとしての分布に正方行列を右からかける記法がよく使われる。 P や π の代わりにそれらの転置を考えることにすれば「主流の記法」に変換できる。他にも、圏論では $f(x)$ を xf と書くことがある。

解いてみると、定常分布は $\pi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ のただ一つだと分かる。ちなみに、「定常」とは「変わらない」という意味で、定常分布は「不変分布」と呼ばれることもある。なぜこのように呼ばれるかという、もしマルコフ連鎖の初期分布が定常分布であれば、任意の時刻 n で X_n の分布は初期分布と等しい、つまり時間が経っても分布が変化しないからである。

以降では、定常分布に関する次の三つの基本問題について見ていく。

- (Q1) 定常分布はどれだけ存在するか？
- (Q2) 極限値 (9) は存在するか？ もし極限値が存在すれば定常分布になることは既に見た通り。
- (Q3) 極限値 (9) が存在するとき、収束の速さはどの程度か？

5. 定常分布はどれだけ存在するか

以降では、 P^n の第 (i, j) 成分を $p_n(i, j)$ と表すことにする ($n \geq 0$)。これは i から出発したマルコフ連鎖が時刻 n で j にいる確率と等しい。このことは、 μ を第 i 成分が 1 で他は全て 0 の横ベクトルとした場合に横ベクトル $\mu_n = \mu P^n$ の第 j 成分が $p_n(i, j)$ と一致することから分かる。

まず定常分布は存在するかという問題について、結論から言うと以下のことが知られている。

定理 1 有限サイズの推移確率行列は定常分布を持つ。

複数の証明があるが、筆者の知る限りどれも難しい。まずコンパクト凸集合に対する Brouwer の不動点定理を使う方法がある。¹⁵⁾ 具体的には、確率分布の集合 $\Delta = \{(\pi_1, \pi_2, \dots, \pi_s) : \pi_i \geq 0, \sum_{i=1}^s \pi_i = 1\}$ は $\mathbb{R}^s$ のコンパクト凸部分集合となっていて、さらに写像 $\pi \mapsto \pi P$ は Δ 上の連続写像なので、Brouwer の不動点定理により不動点、つまり定常分布が存在する。この証明では Brouwer の不動点定理自体の証明が難しい。この節の最後で別の証明方法を紹介する。

参考までに、無限状態空間でも無限サイズの推移確率行列を考えることができるが、この場合は定常分布が存在しないことも頻繁に起こる。例えば、整数点上の単純ランダムウォーク（隣り合う整数のいずれかに $\frac{1}{2}$ の確率で移動するマルコフ連鎖）は定常分布を持たない。

次に、定常分布はいくつあるか調べてみよう。複数

の定常分布を持つような例は簡単に作ることができる。以下の推移確率行列を考えてみる。

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

P は定常分布 $\pi_1 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$, $\pi_2 = (0, 0, \frac{1}{2}, \frac{1}{2})$ を持つことが確認できる。この例では状態 1, 2 と状態 3, 4 は断絶されており行き来が全くできないので、二つの定常分布があるのは自然に了解できる。さらに、これらの凸結合 $t\pi_1 + (1-t)\pi_2$ ($0 \leq t \leq 1$) も定常分布になる。これで全ての定常分布が尽くされることも容易に分かる。

定常分布の一意性に関する有名な定理の一つを紹介する。有限状態空間 S と推移確率 P が与えられたとする。 S の要素 i, j について、 i から j に到達可能であるとは、ある $n \geq 0$ が存在して $p_n(i, j) > 0$ となることを言う。図 2 のような状態遷移図を利用すると分かりやすい。 i から j に到達可能とは、状態遷移図において i からうまく矢印に沿って進んでいけば j まで行けるという意味である。例えば図 2 では $4 \rightarrow 5 \rightarrow 3$ と進めるから、状態 4 から状態 3 へは到達可能である。状態 3 から状態 1 にはどう矢印をたどっても決して行けないので、到達不可能である。

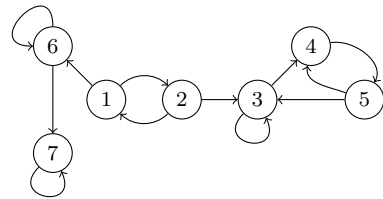


図 2 i から j に向かう矢印は $p(i, j) > 0$ を意味する。逆に i から j に向かう矢印が存在しないことは $p(i, j) = 0$ を意味する。

任意の $i, j \in S$ について i から j に到達可能なとき、 P は既約であると言う。要するに、どの状態から出発しても（時間はかかるかもしれないが）他のどの状態にも行くことができるという状況である。1 節の例では P は既約である。逆に既約でないとは、何か二つの状態 i, j が存在して i から j へ到達不可能だという状況である。図 2 の場合、 P は既約ではない。

定理 2 有限サイズの推移確率行列 P が既約ならば、

定常分布がただ一つ存在する。

定理 2 は Perron-Frobenius の定理と呼ばれる結果の一部として知られている。^{9, Theorem 1.5)} さらに別の証明として文献^{5,7)}にあるような確率論的方法もある。既約な場合には、到達時刻に関する確率または期待値を利用して定常分布を表わすことができて、文献^{5,7)}にある証明はその表示を利用している。既に定理 1 で見たように、定常分布の存在だけを証明するには既約性の仮定は不要ではあるが、既約性を仮定すると上記のように Brouwer の不動点定理を用いずに二通りの方法で証明することが可能になる。

では、既約でない場合に定常分布はいくつあるのか考えてみる。まず、既約でなくても定常分布が一つだけ存在する例がある。 $S = \{1, 2\}$, $p(1, 1) = 0, p(1, 2) = 1, p(2, 1) = 0, p(2, 2) = 1$ とする。対応するマルコフ連鎖は状態 1 から出発したら必ず時刻 1 で状態 2 に移動して、それ以降は状態 2 に留まる。また状態 2 から出発したら、以降は全く動かない。とてもつまらないマルコフ連鎖である。状態 2 から出発したら状態 1 に決して行けないから、この P は既約ではない。しかし定常分布は $\pi = (0, 1)$ の一つしかないことが簡単な計算から分かる。余談だが、このような最も簡単な例から出発して徐々に難しい例を調べていくことは研究において基本的なやり方である。

実は、一般の既約とは限らない状況でも定常分布を全て決定することは可能である。そのためには S を同値類に分解する。まず S の要素 i, j が P に関して同値であるとは、 i から j に到達可能かつ j から i に到達可能なことをいう。いわば、 i と j はお互いに行き来できる。互いに同値な状態どうしを全て集めて S の中の一つのグループ(部分集合)とし、これを同値類と呼ぶ。 S が既約というのは、同値類が一つしかないことを意味する。図 2 の場合なら同値類は $\{1, 2\}, \{3, 4, 5\}, \{6\}, \{7\}$ の全部で 4 グループある。

同値類たちをさらに二種類に区別する。同値類 C が開いているというのは、マルコフ連鎖が C の中から外に出られることを言う(つまりある状態 $i \in C$ からある状態 $j \in S \setminus C$ へ到達可能である)。そうでない場合には閉じていると呼ぶ。これは、マルコフ連鎖が C の中から外に決して出られないことを意味する。図 2 で言うと、同値類 $\{1, 2\}$ と $\{6\}$ は開いており、同値類 $\{3, 4, 5\}$ と $\{7\}$ は閉じている。特にいったん状態 7 に到達すると、以降動かない。このような状態を吸収

状態と呼ぶ。

閉じた同値類を用いると、有限状態空間における定常分布を全て決定することができる。まず、有限状態空間では閉じた同値類が少なくとも一つ存在する(背理法で考えると良い)。全ての閉じた同値類を $R_1, R_2, \dots, R_l$ と表すことにしよう。任意の i を固定して R_i に着目すると、 R_i 内の状態から出発したマルコフ連鎖はずっと R_i 内に留まるので、 R_i 自体を状態空間とみなしてその中だけで推移確率(行列)を考えることが可能である。 R_i の中ではどの状態どうしも行き来できるから、 R_i の中での推移確率は既約になり、定理 2 により定常分布をただ一つ持つ。この R_i 上の定常分布を R_i の外では 0 として S 全体に拡張したものを π_i と表すと、これは S 上の定常分布になることが比較的容易な考察から分かる。

定理 3 上述の $\pi_1, \pi_2, \dots, \pi_l$ たちの凸結合

$$\sum_{i=1}^l t_i \pi_i \quad \left(t_i \geq 0, \sum_{i=1}^l t_i = 1 \right)$$

は全て S 上の定常分布であり、しかもこれで S 上の全ての定常分布が尽くされる。特に、定常分布がただ一つになるための必要十分条件は、閉じた同値類の個数が 1 であることである。

π_i たちの凸結合が定常分布になることは容易に分かるが、それで全ての定常分布が尽くされることを示すには込み入った議論が必要になる。参考文献を挙げると、結果の述べ方はかなり異なるが実質的に^{12, Theorem 4.1.10)}が定理 3 に相当する。なおこの文献では無限状態空間も扱っている。以上で問題 (Q1) が解決した。

問題 (Q2) に移る前に、ここで Brouwer の不動点定理を使わない定理 1 の二通りの別証明について述べる。どちらも定理 2 を用いる。まず定理 3 を議論している途中で定理 2 を用いて定常分布を構成しているので、それが既に定理 1 の別証明になっている。ちなみに、推移確率行列が既約というのは閉じた同値類の個数が 1 かつ開いた同値類の個数が 0 ということなので、定理 3 は定理 2 も包含している。ただし本稿では $\pi_1, \pi_2, \dots, \pi_l$ の構成が定理 2 に基づいているので、定理 3 によって定理 2 の別証明ができるわけではない。

また、同値類ではなく近似の観点から定理 1 を証明することもできるので紹介する。補助的に全ての成分が $1/s$ である正方行列 J を用意して、 P と J の凸結合

$$P_t = (1 - t)P + tJ$$

を定める ($0 < t \leq 1$). すると P_t も推移確率行列となり, しかも全ての成分が正なので既約になることがすぐに分かる. よって定理 2 により P_t は定常分布 π_t を持つ. あとはいわゆるコンパクト性の議論によって, π_t の部分列をうまく選んで $t \rightarrow 0^+$ の極限を取れば収束し, その極限が P の定常分布になる.

6. 定常分布への収束

問題 (Q2) の極限值 (9) の存在について述べるための準備として, 次の集合の最大公約数を P に関する i の周期と呼ぶ:

$$\mathcal{T}(i) = \{n \in \mathbb{N} : p_n(i, i) > 0\}.$$

P が既約のときは, いずれの状態も同じ周期を持つことが分かる. このとき, この共通の値を単に P の周期と呼ぶ. P の周期が 1 のとき, P は非周期的と言う.

周期とは何か分かりづらいかもしれない. 例で計算してみよう. 1 節の例, つまり (3) で与えられる P は既約であり, $i = a, b, c, d$ のいずれについても

$$\mathcal{T}(i) = \{2, 3, 4, 5, 6, \dots\}$$

となる. すなわち, 頂点 i から出発して 1 ステップで i に戻ることはできないが, 2 ステップ, 3 ステップ, ... のいずれでも i に戻るができる. よって, i の周期は 1 であり, P は非周期的である. 別の例として, 頂点 1, 2, 3, 4 を持つ正方形上で隣り合う二つの頂点のいずれかに等確率 $\frac{1}{2}$ で移動していくマルコフ連鎖 (図 3) を考える. この推移確率は既約である. 頂点 i から出発して奇数ステップで i に戻って来ることはできないから, n が奇数だと $p_n(i, i) = 0$ となる. 逆に, n が偶数だと $p_n(i, i) > 0$ となることも分かる. よって, $i = 1, 2, 3, 4$ のいずれについても

$$\mathcal{T}(i) = \{2, 4, 6, 8, \dots\}$$

となり, P の周期は 2 になる.

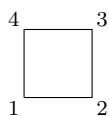


図 3 正方形

既約性+非周期性は次の分かりやすい特徴づけを持つ. 証明は例えば文献¹³⁾にある.

命題 4 有限サイズの推移確率行列 P が既約かつ非周期的になるための必要十分条件は, ある自然数 n が存在して P^n の全ての成分が正になることである.

既約性と非周期性はよくマルコフ連鎖の理論で仮定される. その利点は問題 (Q2) について以下のように解答が得られることである. 証明は文献^{3, 5, 13)}などを参照のこと.

定理 5 有限サイズの推移確率行列 P が既約かつ非周期的ならば, 任意の初期分布 μ について, $\lim_{n \rightarrow \infty} \mu P^n$ が存在する.

この定理に関する注意点を二つ述べる. まず極限 $\lim_{n \rightarrow \infty} \mu P^n$ は定常分布になることが分かっていて, さらに既約性より P の定常分布はただ一つだから, 極限 $\lim_{n \rightarrow \infty} \mu P^n$ は μ に依存しない. そして初期分布 μ は任意に選べることから, 行列の全ての成分が収束するという意味で $\lim_{n \rightarrow \infty} P^n$ が存在し, その極限行列は定常分布を s 個縦に並べた正方行列になることが分かる.

このように初期点に写像を反復して適用していくとその写像の不動点に収束するという議論は, マルコフ連鎖以外の分野でもよく現れる. このような議論では, 写像の「縮小性」と呼ばれる性質を利用することが多い. 定理 5 の場合も P はある種の「縮小性」を持っており, それを用いた証明が文献³⁾にある.

次に問題 (Q3) の定常分布への収束の速さについて, 一つの結果を紹介する. まず定理 1 より P は定常分布を持つから P は固有値 1 を持つ (行列を転置しても固有値は変わらないことに注意する). このことは定理 1 を使わなくても簡単に示すことができる: 全ての成分が 1 であるような縦ベクトルを $\mathbf{1}$ とすると, P の各行の和が 1 になることから $P\mathbf{1} = \mathbf{1}$ が成り立つので, これより P は固有値 1 を持つと分かる. 実は, その他の 1 でない固有値 (複素数になりうる) の中で, 絶対値が最も大きいものが収束の速さを調べる上で大きな役割を果たす. これを第二固有値と呼ぶ.

定理 6 既約かつ非周期的な有限サイズの推移確率行列 P を考える. また, π を P の定常分布とする. このとき以下が成り立つ.

- (i) 固有値 1 の多重度は 1 である.
- (ii) 1 でない P の固有値の絶対値はいずれも 1 より小さい.

- (iii) 1 でない P の固有値の絶対値の中で最大のものを r とする. ある定数 $C > 0$ と $g \geq 0$ が存在して, 任意の初期分布 μ と n に対して,

$$\|\mu P^n - \pi\| \leq C n^g r^n$$

が成り立つ. ただし $\|v\|$ は横ベクトル v の長さを表す.

(i), (ii) は Perron-Frobenius の定理の一部になっている.^{9, Theorem 1.1)} (iii) は定理 5 の精密化になっている. P の Jordan 標準形を利用して証明することができる. $n^g r^n$ は Jordan 細胞を n 乗すると現れる項で, 定数 g は最大の Jordan 細胞のサイズから 1 を引いた数として取ることができる. もし P が対角化可能であれば, $g = 0$ で良い. 詳しくは文献⁸⁾を参照のこと.

定常分布への収束の速さに関しては **cutoff 現象**が興味深い. ある時刻を境にマルコフ連鎖の分布 μ_n が定常分布へ急激に近づくという現象である. 特に状態空間のサイズが大きい場合によく見られる. この現象が見られる有名な例として, トランプのシャッフルを何回くらい行えばよく混ざるのかという問題がある. まずカードの可能な並び順を全て列挙した集合を状態空間とみなす. 従って, 状態空間は $52!$ 個の要素を持つ (52 ではなく $52!$, およそ 10^{68} 個). 一回シャッフルするという操作をこの巨大な状態空間の中の状態が確率的に別の状態へ推移すると考えると, シャッフルを繰り返す操作をマルコフ連鎖としてとらえることができる. 推移確率はまさにシャッフルの仕方によって決まる. 通常, 定常分布は一様分布になる. 一様分布とはカードの並び方 $52!$ 通りのうちいずれも等しい確率で現れることを意味する. カードの並び方が一様分布のときにカードが最もよく混ざっていると考えるのが合理的である. そこで, マルコフ連鎖の時間を進めて (= シャッフルを繰り返して) いき, その分布が一様分布に十分に近づいたらカードがよく混ざったと考えることにする. シャッフルのやり方もいろいろあるが, riffle shuffle というやり方の場合は, 7 回目あたりで急激によく混ざることが Diaconis たちの研究で分かっている. 詳しくは文献^{1, 2, 5, 11)}などを参照のこと.

7. 展望

定常分布は線型写像の不動点なので, 不動点の問題としては易しく見えるかもしれない. 実際, 定理 3 で

不動点の構造はかなり分かっている. しかし, 具体的なマルコフ連鎖は状態空間が巨大だったり複雑になりやすく, 定常分布に関する解析は容易でないことがある. この辺りは文献⁵⁾に詳しく書かれている.

本稿では触れなかったが, 時間と状態空間が連続の場合にもマルコフ連鎖の類似が考えられる. これはマルコフ過程と呼ばれており, 統計物理学に由来する Brown 運動 (Wiener 過程とも呼ばれる) が代表例になる. 連続の場合は確率積分, 確率微分方程式といった「微積分」の手法が確立されており, これらを利用することにより深い計算が可能になる. マルコフ連鎖をある程度理解できた読者はぜひこちらにもチャレンジして欲しい. 連続とは言っても, マルコフ過程のシミュレーションを行う場合などは時間と状態空間をともに離散化してマルコフ連鎖で近似するのが基本になるため, マルコフ連鎖もやはり大切な役割を果たす.

最後に, 原稿段階で洞彰人氏, 宮尾忠宏氏, 坂井哲氏からいただいた貴重なご意見が本稿の改善に役立ったことを記しておく.

参考文献

- 1) P. Diaconis, The cutoff phenomenon in finite Markov chains, Proc. Natl. Acad. Sci. USA 93 (1996), 1659–1664.
- 2) P. Diaconis and J. Fulman, *The Mathematics of Shuffling Cards*, Amer. Math. Soc., 2023.
- 3) 舟木直久, 確率論, 朝倉書店, 2004.
- 4) S. カーリン, 確率過程講義 (訳 佐藤健一・佐藤由身子), 産業図書, 1974.
- 5) D. A. Levin and Y. Peres, *Markov Chains and Mixing Times* (2nd ed.), Amer. Math. Soc., 2017.
- 6) J. R. Norris, *Markov Chains*, Cambridge University Press, 1998.
- 7) 尾畑伸明, 確率モデル要論, 牧野書店, 2012.
- 8) J. S. Rosenthal, Convergence rates for Markov chains, SIAM Review 37, Issue 3 (1995), 387–405.
- 9) E. Seneta, *Non-negative Matrices and Markov Chains* (2nd ed.), Springer, 1980.
- 10) R. B. シナジ, マルコフ連鎖から格子確率モデルへ (訳 今野紀雄・林俊一), 丸善出版, 2012.
- 11) 白井朋之, マルコフ連鎖と混合時間, 大学院 GP 数学レクチャーノートシリーズ GP-TML03, 2008.
- 12) D. W. Stroock, *An Introduction to Markov Processes* (2nd ed.), Springer, 2014.
- 13) 竹居正登, 入門 確率過程, 森北出版, 2020.
- 14) W. Woess, *Random Walks on Infinite Graphs and Groups*, Cambridge University Press, 2000.
- 15) <https://math.stackexchange.com/questions/119891/finite-state-markov-chain-stationary-distribution>