MCS.T312「マルコフ解析」講義資料

マルコフ過程入門

中野 張

2022 年 6 月 9 日

目次

記法

- $\mathbb{N} = \{1, 2, \ldots\}$.
- $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$.
- $\mathbb{R}^d$: d 次元ユークリッド空間.
- $\mathbb{Z}^d = \{(x^1, \ldots, x^d) : x^i \in \mathbb{Z},\ \ 1 \leq i \leq d\}$.
- $\mathbb{R}^{m \times d}$: 実 $m \times d$ 行列全体.
- $\mathbb{C}$: 複素数全体.
- $|x|$: $x \in \mathbb{R}^d$ のとき，x のユークリッドノルム.
- $|S|$: S が集合のとき，S の要素数.
- $x_+ = \max\{x, 0\}$, $x \in \mathbb{R}$.
- $\mathbb{E}[X]$: 確率変数 X の期待値.
- $\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$: 確率変数 X の分散.
- I_m: $\mathbb{R}^{m \times m}$ の単位行列.
- 証明*: 初読の際は読み飛ばしてもよい証明の見出し.

第 1 章

離散時間マルコフ過程

一般に，**確率過程**（stochastic process）とは，確率変数の族 $\{X_\lambda\}_{\lambda\in\Lambda}$ であり，パラメータ集合 Λ が時間などによって意味付けられるもののことをいう．特に，$\Lambda=\mathbb{N}\cup\{0\}$ のとき，$\{X_n\}_{n=0}^\infty$ を離散時間確率過程，$\Lambda=[0,\infty)$ のとき，$\{X_t\}_{t\geq 0}$ を連続時間確率過程と呼ぶ．それぞれ有限時間までに制限して考えることも多い．

本章では，離散時間確率過程の内，離散時間マルコフ過程の代表例である離散時間マルコフ連鎖を扱う．以下，現れる確率変数，確率過程は全て確率空間 $(\Omega,\mathcal{F},\mathbb{P})$ 上のものとする．なお，確率論の基礎事項については付録 A を参照のこと．

1.1　マルコフ性と離散時間マルコフ連鎖

S を高々可算集合（i.e., S は有限集合あるいは可算無限集合）とする．$i,j\in S$ に対して $p_{ij}\in[0,1]$ が対応していて

$$\sum_{j\in S}p_{ij}=1,\quad i\in S$$

をみたすとき，$P=\{p_{ij}\}_{i,j\in S}$ を**推移確率**（transition probability）という．また，$\mathbb{R}^{|S|\times|S|}$ の行列と同一視するとき，P を**推移確率行列**（transition probability matrix）ともいう．

定義 **1.1.1.** S 値確率過程 $\{X_n\}_{n=0}^\infty$ が，任意の $n=0,1,\ldots$ と任意の $j\in S$, および $\mathbb{P}(X_0=i_0,\ldots,X_{n-1}=i_{n-1},X_n=i)>0$ なる任意の $i_0,\ldots,i_{n-1},i\in S$ に対して

$$\mathbb{P}(X_{n+1}=j\,|\,X_0=i_0,\ldots,X_{n-1}=i_{n-1},X_n=i)=p_{ij}\tag{1.1.1}$$

を満たすとき，推移確率 P をもつ**マルコフ連鎖**（Markov chain）という．

- 条件 (1.1.1) より

  $$\mathbb{P}(X_{n+1}=j\,|\,X_0=i_0,\ldots,X_{n-1}=i_{n-1},X_n=i)=\mathbb{P}(X_{n+1}=j\,|\,X_n=i)$$

  である．これは $X_{n+1}=j$ の条件付確率は直前の X_n の値にのみ依存し，それ以前の履歴とは無関係に決まることを意味している．このような性質のことを**マルコフ性**という．

- S の元を状態と呼ぶ.
- マルコフ連鎖の値をとる集合 S を**状態空間**あるいは**状態集合**という.
- $\{X_n\}$ がマルコフ連鎖のとき，$\mathbb{P}(X_{n+1}=j\,|\,X_n=i)$ は n に依存しない．この意味で $\{X_n\}$ を**一様**（time-homogeneous）マルコフ連鎖ともいう.

例 **1.1.2** (独立試行)**.** S 値確率過程 $\{X_n\}$ が独立同分布（Independent, Identically Distributed, IID）確率変数列で,

$$\mathbb{P}(X_0=i)=\mu_i,\quad i\in S$$

のとき,

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

となるので, $\{X_n\}$ はマルコフ連鎖である．さらに, $|S|=m<\infty$ のとき, S を $\{1,\dots,m\}$ と同一視すれば，推移確率 P は行列として

$$P=\begin{array}{c}\mathbf{1}\\ \vdots\\ \mathbf{m}\end{array}\begin{pmatrix}\mu_1&\cdots&\mu_m\\ \vdots&\ddots&\vdots\\ \mu_1&\cdots&\mu_m\end{pmatrix}$$

と表される．ここで，太字の数は状態集合 $S=\{1,\dots,m\}$ の元を指し示している．すなわち，上の行列表示より，$p_{11}=\mu_1$, $p_{32}=\mu_2$ 等と読める.

例 **1.1.3** (ランダム・ウォーク)**.** $\{Z_n\}_{n=1}^{\infty}$ を $\mathbb{Z}^d$ 値 IID とし，これを基に別の確率過程 $\{X_n\}_{n=0}^{\infty}$ を

$$X_0=0,\quad X_n=\sum_{k=1}^{n}Z_k$$

により定義する．このとき，Z_{n+1} と $\{X_0,X_1,\dots,X_n\}$ は独立なので,

$$\begin{aligned}&\mathbb{P}(X_{n+1}=i_{n+1}\,|\,X_0=0,X_1=i_1,\dots,X_n=i_n)\\&=\mathbb{P}(Z_{n+1}+i_n=i_{n+1}\,|\,X_0=0,X_1=i_1,\dots,X_n=i_n)=\mathbb{P}(Z_{n+1}=i_{n+1}-i_n)\\&=\mathbb{P}(X_{n+1}=i_{n+1}\,|\,X_n=i_n).\end{aligned}$$

よって $\{X_n\}$ はマルコフ連鎖である．このように定義される $\{X_n\}$ をランダム・ウォーク（random walk），あるいは**乱歩**，**酔歩**という．特に，Z_1 が原点から長さ 1 となる $2d$ 個のベクトルに等確率に分布しているとき，$\{X_n\}$ を**単純ランダム・ウォーク**（simple random walk）という.

例 **1.1.4** (ギャンブラーの破産問題)**.** 確率 $p=0.4$ で 100 円の勝ち，確率 $q=0.6$ で 100 円の負けとなる賭けを繰り返し行うゲームを考えよう．このゲームは，プレイヤーの持ち金がゼロになるか，$100\times N$ 円になった時点で終了になるものとする．今，n 回後の持ち

金を X_n とするとき，$0 < X_n < 100N$ ならば

$$X_{n+1} = \begin{cases} X_n + 100 & (\text{確率 } 0.4), \\ X_n - 100 & (\text{確率 } 0.6) \end{cases}$$

である．もし $X_n = 0$ ならゲームは終了しているが，形式的に $X_{n+1} = 0$ とおくことにする．同様に，$X_n = 100N$ のときも $X_{n+1} = 100N$ とおくことにする．このとき，$X_0 = 100N_0$ ($N_0 \in \{1, \dots, N-1\}$) とすれば，$\{X_n\}$ は状態集合 $\{0, 100, \dots, 100N\}$ のマルコフ連鎖である．

例 **1.1.5** (在庫管理)**.** ある薬局では，その日の終わりに店内の医薬品 A の数が 0 個か 1 個になれば，A の合計が 6 個になるように発注する．A を購入する客は 1 日に 0, 1, 2, 3, 4 人のどれかであり，それぞれ確率 0.2, 0.3, 0.2, 0.2, 0.1 で来店すると仮定する．また，客は 1 人 1 個ずつ購入し，新しい医薬品は翌日，店が開く前に届くものとする．n 日の閉店時に店内にある医薬品 A の数を X_n とするとき，$\{X_n\}$ はマルコフ連鎖であり，推移確率行列は次のようになる:

$$\begin{array}{c} \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \end{array} \begin{pmatrix} 0 & 0 & 0.1 & 0.2 & 0.2 & 0.3 & 0.2 \\ 0 & 0 & 0.1 & 0.2 & 0.2 & 0.3 & 0.2 \\ 0.5 & 0.3 & 0.2 & 0 & 0 & 0 & 0 \\ 0.3 & 0.2 & 0.3 & 0.2 & 0 & 0 & 0 \\ 0.1 & 0.2 & 0.2 & 0.3 & 0.2 & 0 & 0 \\ 0 & 0.1 & 0.2 & 0.2 & 0.3 & 0.2 & 0 \\ 0 & 0 & 0.1 & 0.2 & 0.2 & 0.3 & 0.2 \end{pmatrix}$$

また，$\{X_n\}$ は次の漸化式により表される:

$$X_{n+1} = \begin{cases} (X_n - D_{n+1})_+, & X_n > 1, \\ (6 - D_{n+1})_+, & X_n \leq 1. \end{cases}$$

ただし，D_n は n 日目の客の需要，すなわち n 日目における客数×各々の購入希望数（この例では 1）を表す．

例 **1.1.6** ($M/G/1$ 待ち行列)**.** X_n をある一つの窓口に並ぶ客の時刻 n での人数とし，Z_n を時刻 n で新たに行列に加わった人数とする．窓口は単位時間で先着順に 1 人ずつサービスを提供し，$\{Z_n\}$ は $\mathbb{N} \cup \{0\}$ 値 IID 列とする．このとき，$X_n = (X_{n-1} - 1)_+ + Z_n$ であり，$\{X_n\}$ は状態集合 $\mathbb{N} \cup \{0\}$，推移確率

$$p_{0j} = \mathbb{P}(Z_1 = j), \;\; p_{i,i-1+j} = \mathbb{P}(Z_1 = j), \quad i \geq 1, \; j \geq 0$$

をもつマルコフ連鎖となる．

例 **1.1.7** (ハイパーリンク解析)**.** Web ページをランダムに閲覧するユーザを考えよう．まず，ユーザが閲覧する可能性のある Web ページの集合を番号付けて $S = \{1, \dots, N\}$ により表す．$i \in S$ に対し，Web ページ i のリンク先の Web ページの集合を $F(i)$ により

表し，$F(i) \neq \emptyset$ を仮定する．これはすなわち，リンク先のないページは存在しないという仮定である．n_i を Web ページ i の持つリンク先ページの総数とし，

$$p_{ij} = \begin{cases} \dfrac{1}{n_i} & (i \text{ から } j \text{ へのリンクが存在する場合}) \\ 0 & (i \text{ から } j \text{ へのリンクが存在しない場合}) \end{cases}$$

と定義すると，$\sum_{j\in S} p_{ij} = \sum_{j\in F(i)}(1/n_i) = n_i/n_i = 1$ であるから $P = \{p_{ij}\}_{i,j\in S}$ は推移確率を定義する．

ユーザは単位時間で 1 つの Web ページを閲覧し，リンク先 Web ページへ等確率で遷移すると仮定する．時刻 n にユーザが閲覧している Web ページを X_n により表すことにすると，P の定義より

$$\mathbb{P}(X_{n+1} = j \,|\, X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j \,|\, X_n = i) = p_{ij}$$

であり，$\{X_n\}$ は推移確率 P を持つマルコフ連鎖になる．

例 **1.1.8** (債券格付けモデル)**.** 国や企業，地方公共団体等が発行する債券の格付けについて考えよう．債券の購入者にとって，格付けの変化を定量的に把握することは信用リスク管理の一つの手段となる．ある債券の格付けの時系列 $\{X_n\}$ にマルコフ性を仮定し，格付けの集合 $S = \{AAA, AA, A, BBB, BB, B, CCC, CC, C, D\}$ を考える．ここで D は何らかの意味での債務不履行（デフォルト）を表す．このとき，$\{X_n\}$ は S を状態空間とするマルコフ連鎖となる．

	AAA	AA	A	BBB	BB	B	CCC-C	D
AAA	91.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0
AA	0.8	94.2	5.0	0.0	0.0	0.0	0.0	0.0
A	0.0	1.8	94.5	3.5	0.1	0.0	0.0	0.1
BBB	0.0	0.0	3.8	93.4	2.6	0.0	0.0	0.1
BB	0.0	0.0	0.2	8.2	86.4	2.6	0.1	2.5
B	0.0	0.0	0.0	0.8	9.8	77.3	0.8	11.4
CCC-C	0.0	0.0	0.0	0.0	0.0	4.4	88.9	6.7

表 1.1.1　日本企業発行社債の 1 年あたりの平均格付け変動率（%, 1978–2014）．格付投資情報センター発行レポート（2015.6.30）より．CCC-C は CCC, CC, C の合算．

マルコフ連鎖の定義から直ちに従う事実を紹介しておこう．

命題 **1.1.9.** $\{X_n\}$ をマルコフ連鎖とし，$n \geq 0, m \geq 1$ とする．このとき，$i_n, \ldots, i_{n+m} \in S$ に対し，

$$\mathbb{P}(X_n = i_n, \ldots, X_{n+m} = i_{n+m}) = \mathbb{P}(X_n = i_n) p_{i_n i_{n+1}} \cdots p_{i_{n+m-1} i_{n+m}}.$$

証明．まず，$\mathbb{P}(X_n = i_n) = 0$ ならば示すべき等式は両辺 0 で成立する．よって $\mathbb{P}(X_n = i_n) > 0$ を仮定する．$k = \max\{\ell \geq 0 : \mathbb{P}(X_n = i_n, \ldots, X_{n+\ell} = i_{n+\ell}) > 0\}$ とおくとき，

マルコフ過程の定義より，

$$
\begin{aligned}
&\mathbb{P}(X_n = i_n, \ldots, X_{n+k} = i_{n+k}) \\
&= \mathbb{P}(X_{n+k} = i_{n+k} \,|\, X_{n+k-1} = i_{n+k-1}, \ldots, X_n = i_n)\mathbb{P}(X_{n+k-1} = i_{n+k-1}, \ldots, X_n = i_n) \\
&= p_{i_{n+k-1}i_{n+k}}\mathbb{P}(X_{n+k-1} = i_{n+k-1}, \ldots, X_n = i_n).
\end{aligned}
$$

この議論を繰り返すことにより，

$$
\begin{aligned}
\mathbb{P}(X_n = i_n, \ldots, X_{n+k} = i_{n+k}) &= p_{i_{n+k-1}i_{n+k}} \cdots p_{i_{n+1}i_{n+2}}\mathbb{P}(X_{n+1} = i_{n+1}, X_n = i_n) \\
&= p_{i_{n+k-1}i_{n+k}} \cdots p_{i_n i_{n+1}}\mathbb{P}(X_n = i_n)
\end{aligned}
$$

を得る． □

また，マルコフ性の以下の意味での拡張が成り立つ．

命題 1.1.10. マルコフ連鎖 $\{X_n\}$ と $m, n \in \mathbb{N}$ を考える．$\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) > 0$ なる任意の $i_0, \ldots, i_n \in S$ と任意の $j_1, \ldots, j_m \in S$ に対し，

$$
\begin{aligned}
&\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m \,|\, X_0 = i_0, \ldots, X_n = i_n) \\
&= \mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m \,|\, X_n = i_n).
\end{aligned}
$$

証明．命題 1.1.9 を繰り返し用いて，

$$
\begin{aligned}
&\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m \,|\, X_0 = i_0, \ldots, X_n = i_n) \\
&= \frac{\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n, X_{n+1} = j_1, \ldots, X_{n+m} = j_m)}{\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n)} \\
&= \frac{\mathbb{P}(X_0 = i_0)p_{i_0 i_1} \cdots p_{i_{n-1}i_n}p_{i_n j_1} \cdots p_{j_{m-1}j_m}}{\mathbb{P}(X_0 = i_0)p_{i_0 i_1} \cdots p_{i_{n-1}i_n}} \\
&= p_{i_n j_1} \cdots p_{j_{m-1}j_m} \\
&= \frac{\mathbb{P}(X_n = i_n, X_{n+1} = j_1, \ldots, X_{n+m} = j_m)}{\mathbb{P}(X_n = i_n)} \\
&= \mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m \,|\, X_n = i_n).
\end{aligned}
$$

ゆえに命題が従う． □

本節の最後に，任意に与えた初期分布と推移確率をもつマルコフ連鎖を明示的に構成する問題を考える．この部分は初読の際は読み飛ばしてもよい．

簡単のため，$S = \mathbb{N} \cup \{0\}$ とする．初期分布を $\mu = \{\mu_i\}_{i=0}^{\infty}$ で，推移確率を $P = \{p_{ij}\}_{i,j \in S}$ により表す．

まず，$\{U_n\}_{n=0}^{\infty}$ を $[0, 1]$ 上の一様分布を共通の分布とする IID 列とする．

$$
X_0 = \sum_{k=0}^{\infty} k 1_{(\sum_{i=0}^{k-1}\mu_i, \sum_{i=0}^{k}\mu_i]}(U_0).
$$

ただし $\sum_{i=0}^{-1}\mu_i = 0$ と解釈する．このとき，$X_0 = k \Leftrightarrow \sum_{i=0}^{k-1}\mu_i < U_0 \leq \sum_{i=0}^{k}\mu_i$ より，

$$
\mathbb{P}(X_0 = k) = \mu_k, \quad k \in S.
$$

すなわち，U_0 のシミュレーションにより μ に従う確率変数 X_0 をシミュレーションできる．

次に，関数 $f : S \times [0,1] \to S$ を

$$f(i,u) = \sum_{k=0}^{\infty} k 1_{(\sum_{j=0}^{k-1} p_{ij}, \sum_{j=0}^{k} p_{ij}]}(u)$$

により定義すると，$f(i,u) = k \Leftrightarrow \sum_{j=0}^{k-1} p_{ij} < u \le \sum_{j=0}^{k} p_{ij}$ である．この f を用いて，確率過程 $\{X_n\}$ を

$$X_{n+1} = f(X_n, U_{n+1}), \quad n \ge 0$$

により定義する．このとき，次が成り立つ．

定理 1.1. $\{X_n\}_{n=0}^{\infty}$ は初期分布 μ と推移確率 P をもつマルコフ連鎖である．

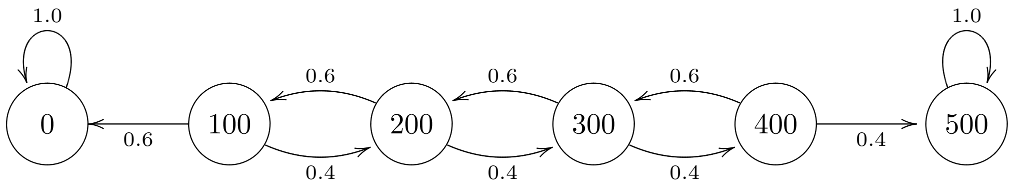
証明. $n \ge 0$ を固定し，$i_0, \ldots, i_{n-1}, i, j \in S$ を勝手にとる．このとき，f の定義と U_{n+1} が一様分布であること，および $X_0, \ldots, X_n$ が U_{n+1} と独立であることより，条件付けの事象の確率が正との仮定の下で，

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) \\
&= \mathbb{P}(f(X_n, U_{n+1}) = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) \\
&= \mathbb{P}(f(i, U_{n+1}) = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) \\
&= \mathbb{P}(f(i, U_{n+1}) = j) = \mathbb{P}\left(\sum_{k=0}^{j-1} p_{ik} < U_{n+1} \le \sum_{k=0}^{j} p_{ik} \right) = p_{ij}
\end{aligned}$$

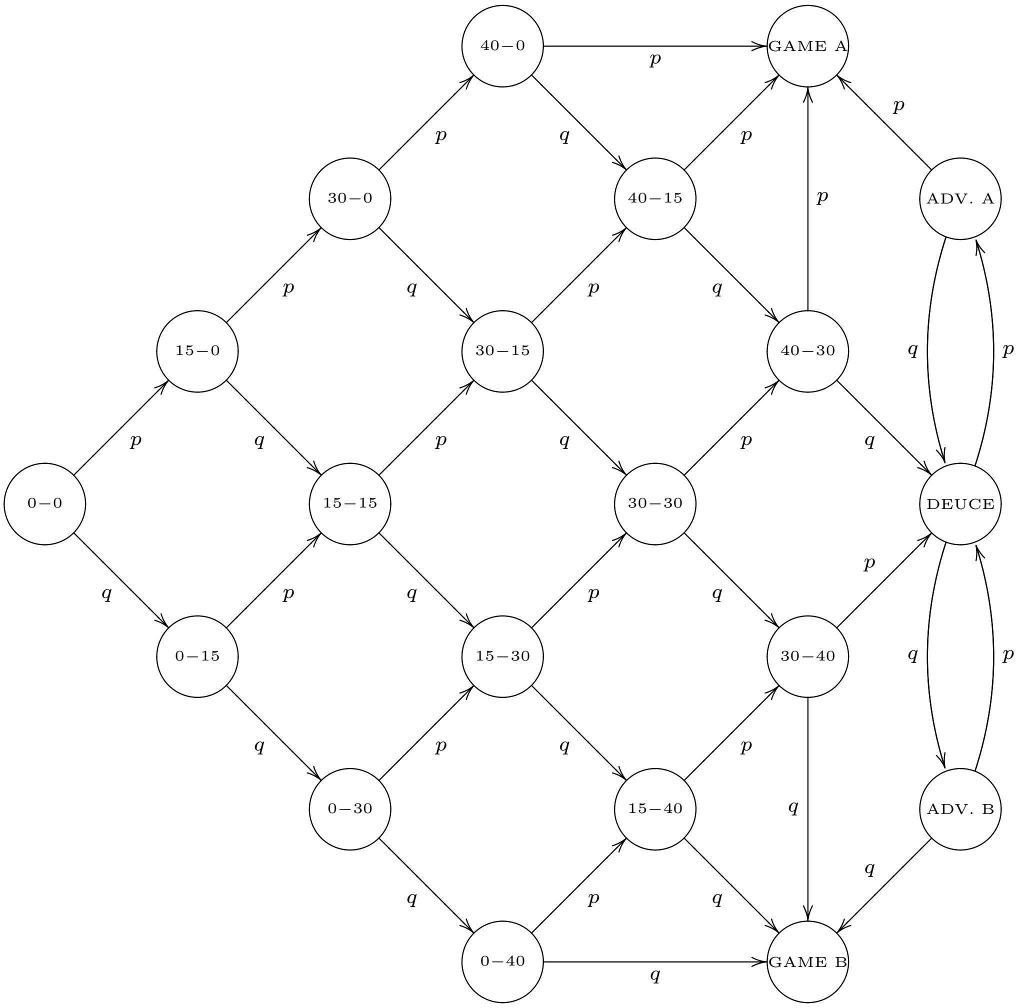
を得る．右辺が $i_0, \ldots, i_{n-1}$ に依存しないことからマルコフ性が従う． □

1.2 推移図と状態の確率分布

例 1.1.4 を思い出そう．$N = 5$ とするとき，このマルコフ連鎖の推移の様子は次の図のように，状態を頂点，状態推移を有向辺とする有向グラフによって表すことができる．ここで，辺に付記されている数値は推移確率である．このような図をマルコフ連鎖の**推移図**と呼ぶ．



もう一つ推移図の例を挙げよう．テニスのスコアの推移は，（タイブレークが関係ないときは）下の推移図を持つマルコフ連鎖によって記述することができる．ここで，p, q はそれぞれ，プレイヤー A，プレイヤー B がポイントを取る確率であり，$q = 1 - p$ である．



次に，$P=(p_{ij})$ を推移確率とする．$P^2=PP$ の (i,j) 成分を $p_{ij}^{(2)}$ により表すとき，

$$p_{ij}^{(2)}=\sum_{k\in S}p_{ik}p_{kj}. \tag{1.2.1}$$

ここで，推移確率の定義より，$\sum_k p_{ik}p_{kj}\leq\sum_k p_{ik}=1$. ゆえに，$S$ が無限集合であっても (1.2.1) の右辺は収束する．

同様に，$P^3=PP^2=P^2P$ の (i,j) 成分を $p_{ij}^{(3)}$ により表すとき，

$$p_{ij}^{(3)}=\sum_{k\in S}p_{ik}^{(2)}p_{kj}=\sum_{k\in S}p_{ik}p_{kj}^{(2)}$$

であり，より一般に，$P^n=P^{n-1}P=PP^{n-1}$ の (i,j) 成分を $p_{ij}^{(n)}$ により表すとき，

$$p_{ij}^{(n)}=\sum_{k\in S}p_{ik}^{(n-1)}p_{kj}=\sum_{k\in S}p_{ik}p_{kj}^{(n-1)} \tag{1.2.2}$$

である．また，$P^0 = (p_{ij}^{(0)})$ を

$$p_{ij}^{(0)} = \begin{cases} 1 & (i = j), \\ 0 & (i \neq j) \end{cases}$$

により定義する．

次の命題でみるように，n ステップ後の状態の確率を計算するためには推移確率行列の n 乗を計算すればよい．

命題 1.2.1. $n, m \in \mathbb{N} \cup \{0\}$, $i, j \in S$ とする．$\mathbb{P}(X_m = i) > 0$ ならば

$$p_{ij}^{(n)} = \mathbb{P}(X_{n+m} = j \,|\, X_m = i). \tag{1.2.3}$$

証明. $n = 0, 1$ のとき (1.2.3) が成立するのは明らかである．勝手な $n \in \mathbb{N}$ に対して (1.2.3) の成立を仮定すると，$\mathbb{P}(X_m = i) > 0$ ならば,

$$\begin{aligned} &\mathbb{P}(X_{n+m+1} = j \,|\, X_m = i) \\ &= \sum_k \mathbb{P}(X_{n+m+1} = j, X_{n+m} = k, X_m = i)/\mathbb{P}(X_m = i) \\ &= \sum_k \mathbb{P}(X_{n+m+1} = j \,|\, X_{n+m} = k, X_m = i)\mathbb{P}(X_{n+m} = k, X_m = i)/\mathbb{P}(X_m = i). \end{aligned}$$

ここで $k \in S$ についての和は $\mathbb{P}(X_{n+m} = k, X_m = i) > 0$ を満たすものに関してとる．この表現に注目し,

$$\begin{aligned} &\mathbb{P}(X_{n+m+1} = j \,|\, X_{n+m} = k, X_m = i) \\ &= \sum_{i_1, \ldots, i_{n-1} \in S} \frac{\mathbb{P}(X_{n+m+1} = j, X_{n+m} = k, X_{n+m-1} = i_{n-1}, \ldots, X_{m+1} = i_1, X_m = i)}{\mathbb{P}(X_{n+m} = k, X_m = i)} \\ &= \sum_{i_1, \ldots, i_{n-1} \in S} \frac{p_{kj}\mathbb{P}(X_{n+m} = k, X_{n+m-1} = i_{n-1}, \ldots, X_{m+1} = i_1, X_m = i)}{\mathbb{P}(X_{n+m} = k, X_m = i)} \\ &= p_{kj} \end{aligned}$$

であることと帰納法の仮定より,

$$\mathbb{P}(X_{n+m+1} = j \,|\, X_m = i) = \sum_k p_{kj} p_{ik}^{(n)}. \tag{1.2.4}$$

ここで，$k \in S$ は $\mathbb{P}(X_{n+m} = k, X_m = i) > 0$ を満たすものに限定されていたが，この条件は $p_{ik}^{(n)} > 0$ と同値なので，(1.2.4) の右辺の和は S の全ての元に渡ってとるとしてよい．よって，(1.2.2) より

$$\mathbb{P}(X_{n+m+1} = j \,|\, X_m = i) = p_{ij}^{(n+1)}.$$

ゆえに命題の主張が示された． □

等式 $P^{n+m} = P^n P^m$ を成分で書き下すと

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)}, \quad i, j \in S \tag{1.2.5}$$

となる．これはチャップマン・コルモゴロフ方程式（Chapman-Kolmogorov equation）と呼ばれている．

1.3 状態の分類: 連結性

$i, j \in S$ に対して，

- $p_{ij}^{(n)} > 0$ を満たすような $n \geq 0$ が存在するとき，状態 j は状態 i から**到達可能** (accessible) であるといい，$i \to j$ と書く．
- 状態 i と状態 j が互いに到達可能のとき，i と j は**連結** (communicate) しているといい，$i \leftrightarrow j$ と書く．

例 **1.3.1.** ギャンブラーの破産問題（例 1.1.4）の場合，明らかに $0 \to 0$, $100N \to 100N$ である．さらに，各 $n \in \{1, \dots, N-1\}$ に対して，$100n \to 0$, $100n \to 100N$ である．また，$m, n \in \{1, \dots, N-1\}$, $m \neq n$, のとき，$100m$ と $100n$ は連結している．

補題 **1.3.2.** S の元の間の関係 $\leftrightarrow$ は同値関係である．すなわち以下が成り立つ．

(1) $i \leftrightarrow i$.

(2) $i \leftrightarrow j$ ならば $j \leftrightarrow i$.

(3) $i \leftrightarrow j$ かつ $j \leftrightarrow k$ ならば $i \leftrightarrow k$.

証明. (1) は $p_{ii}^{(0)} = 1$ より従う．(2) は自明である．(3) を示すために $i \leftrightarrow j$, $j \leftrightarrow k$ を仮定する．このとき，$i \to j$, $j \to k$ より，ある $n, m \geq 0$ が存在して $p_{ij}^{(n)} > 0$, $p_{jk}^{(m)} > 0$ である．ここで，チャップマン・コルモゴロフ方程式 (1.2.5) より $p_{ik}^{(n+m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0$ であるから $i \to k$ が従う．同様の議論により，$k \to j$, $j \to i$ から $k \to i$ が得られる． □

S の部分集合 C に対して，

- 任意の $i \in C$, $j \notin C$ に対して $p_{ij} = 0$ が成立つとき，C は**閉**（closed）であるという．
- 一点集合 $\{i\}$ が閉であるとき，i を**吸収状態**（absorbing state）という．
- $i \in S$ が吸収状態であるための必要十分条件は $p_{ii} = 1$ である．
- 任意の $i, j \in C$ が $i \leftrightarrow j$ を満たすとき，C は**既約**（irreducible）であるという．
- S 自身が既約のとき，そのマルコフ連鎖を既約という．

S は既約性に関する同値類によって分割できる．まず適当な $i_0 \in S$ を選び，i_0 と連結する全ての状態を集めた集合を C_0 とおく．次に，勝手な $i_1 \in S \setminus C_0$ をとり，i_1 と連結する全ての状態を集めた集合を C_1 とする．この手続きを状態が選び尽くされるまで続けることにより，

$$S = \bigcup_i C_i, \quad C_i \cap C_j = \emptyset \ \ (i \neq j)$$

を満たす同値類 $C_0, C_1, \dots$ が得られる．

例 **1.3.3.** ギャンブラーの破産問題（例 1.1.4）で例えば $N = 3$ の場合，状態集合は $S = \{0, 100, 200, 300\}$，推移確率は

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 \\ 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

となる．このとき，既約性に関する同値類は $\{0\}$, $\{100, 200\}$, $\{300\}$ であり，0 と 300 が吸収状態である．

例 **1.3.4.** 推移確率

$$P = \begin{array}{c} \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \end{array} \begin{pmatrix} 0.2 & 0.8 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0.8 & 0.2 \end{pmatrix}$$

を持つマルコフ連鎖の場合，$\{0, 1\}$ と $\{2, 3\}$ が既約かつ閉じた集合である．

1.4 周期性

$i \in S$ に対して，

- $I(i) = \{n \in \mathbb{N} \cup \{0\} : p_{ii}^{(n)} > 0\}$ とおく．
- $I(i)$ の最大公約数を $d(i)$ と書き，i の**周期**という．
- $d(i) = 1$ のとき，i は**非周期的** (aperiodic) であるという．
- $p_{ii} > 0$ なら $i \in S$ は非周期的である．

例 **1.4.1.** 1 次元単純ランダム・ウォーク（例 1.1.3）を考える．このとき，$S = \mathbb{Z}$ であり，$p_{i,i+1} = p_{i,i-1} = 0.5$ である．n が偶数なら $n \in I(0)$ であり，n が奇数なら $n \notin I(0)$ であるから，状態 0 の周期は 2 であることが分かる．

周期については以下の性質が基本的である．

命題 **1.4.2.** $i, j \in S$ ついて，$i \leftrightarrow j$ ならば $d(i) = d(j)$ である．すなわち，i と j は同じ周期をもつ．

証明. 仮定より，ある $k, \ell \geq 0$ が存在し，$p_{ij}^{(k)} > 0$, $p_{ji}^{(\ell)} > 0$ である．$n \in I(i)$ に対し，チャップマン・コルモゴロフ方程式より，

$$p_{jj}^{(n+k+\ell)} \geq p_{ji}^{(\ell)} p_{ii}^{(n)} p_{ij}^{(k)} > 0.$$

よって，$n + k + \ell \in I(j)$．この事実と $0 \in I(i)$ より，$k + \ell \equiv 0$, $n + k + \ell \equiv 0 \pmod{d(j)}$ が従う．これより $n \equiv 0 \pmod{d(j)}$ であるから，結局 $d(j)$ は $I(i)$ の公約数である．従って $d(i) \geq d(j)$ となる．

i と j の役割を入れ替えて同じ議論をすることで $d(j) \geq d(i)$ も得られる． □

命題 1.4.3. 任意の $i \in S$ に対して，ある $n_i \in \mathbb{N}$ が存在し，すべての $m \geq n_i$ に対して $md(i) \in I(i)$ である．

証明．$\{n/d(i) : n \in I(i)\}$ の最大公約数が 1 であるから，ある $n_1, \ldots, n_k \in I(i)$ が存在して，$\{n_1/d(i), \ldots, n_k/d(i)\}$ が互いに素となる．よって $\ell_1, \ldots, \ell_k \in \mathbb{Z}$ が存在して

$$\ell_1 n_1 + \cdots + \ell_k n_k = d(i)$$

となる．このとき，

$$N := |\ell_1| n_1 + \cdots + |\ell_k| n_k \in I(i), \quad N + d(i) = (|\ell_1| + \ell_1) n_1 + \cdots + (|\ell_k| + \ell_k) n_k \in I(i)$$

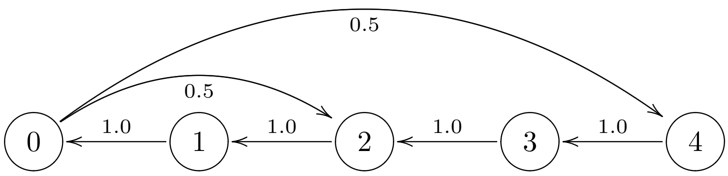
である．実際，一般に $n, m \in I(i)$ なら $n + m \in I(i)$ となるからである．よって，$n \geq N^2$ とし，$n - N^2 = tN + r \ (0 \leq r < N)$ と表すと

$$nd(i) = (r + N^2 + tN)d(i) = r(d(i) + N) + (Nd(i) + td(i) - r)N \in I(i)$$

である．ゆえに $n_i = N^2$ とすることで命題の主張が得られる． □

- i が非周期的のとき，命題 1.4.3 より，十分大きな全ての m について $p_{ii}^{(m)} > 0$ となる．従って，例えば何らかの方法で $p_{ii}^{(2)} > 0$ かつ $p_{ii}^{(3)} > 0$ であることが分かれば，十分大きな全てのステップ数 m について $p_{ii}^{(m)} > 0$ であることも分かる．

例 1.4.4. 推移図



を持つマルコフ連鎖の場合，3 ステップまたは 5 ステップで状態 0 に戻ってくるので，

$$I(0) = \{0, 3, 5, 6, 8, 9, 10, 11, 12, \ldots\}$$

であり，状態 0 の周期は 1 となる．

1.5 再帰性

この節では，ある状態から出発したマルコフ連鎖が再びその状態に戻ってくるか，という問題について議論する．そのために，各状態から出発するマルコフ連鎖を定義しておこう．任意の $i \in S$ に対して $\mathbb{P}(X_0 = i) > 0$ を満たし，推移確率 $P = \{p_{ij}\}$ を持つマルコフ連鎖 $\{X_n\}$ を考える．これに対し，$\{X_0 = i\}$ に関する条件付き確率を $\mathbb{P}_i$ とおく．すなわち，$\mathbb{P}_i(A) = \mathbb{P}(A \,|\, X_0 = i)$, $A \in \mathcal{F}$．このとき，$\mathbb{P}_i(X_0 = i) = 1$ であり，

$\mathbb{P}_i(X_1 = i_1, \ldots, X_n = i_n) > 0$ を満たす $i_1, \ldots, i_n \in S$ と $j \in S$ について

$$\mathbb{P}_i(X_{n+1} = j \,|\, X_1 = i_1, \ldots, X_n = i_n) = \frac{\mathbb{P}_i(X_1 = i_1, \ldots, X_n = i_n, X_{n+1} = j)}{\mathbb{P}_i(X_1 = i_1, \ldots, X_n = i_n)}$$
$$= \frac{\mathbb{P}(X_0 = i, X_1 = i_1, \ldots, X_n = i_n, X_{n+1} = j)}{\mathbb{P}(X_0 = i, X_1 = i_1, \ldots, X_n = i)} = \mathbb{P}(X_{n+1} = j \,|\, X_0 = i, \ldots, X_n = i_n)$$
$$= p_{i_n j}.$$

すなわち，各 $i \in S$ に対して，$\{X_n\}$ は $\mathbb{P}_i$ の下で推移確率 P を持ち，確率 1 で i から出発するマルコフ連鎖である．

$i \in S$ に対し，

- $\tau_i = \inf\{n \in \mathbb{N} : X_n = i\}$ とおく．ただし $\inf \emptyset = \infty$ とする．（τ_i は確率変数であることに注意.）
- τ_i を i への**到達時刻**（hitting time）と呼ぶ.
- $\mathbb{P}_i(\tau_i < \infty) = 1$ が成立つとき，i は**再帰的**（recurrent）であるという.
- $\mathbb{P}_i(\tau_i < \infty) < 1$ が成立つとき，i は**非再帰的**（transient）であるという.
- S の全ての元が再帰的であるとき，マルコフ連鎖が再帰的であるという.

状態の再帰性について調べるために

$$f_{ij}^{(n)} = \mathbb{P}_i(\tau_j = n), \quad f_{ij} = \mathbb{P}_i(\tau_j < \infty), \quad i, j \in S$$

を考える.

命題 **1.5.1.** $i, j \in S$ について，

$$p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}, \quad n \in \mathbb{N}.$$

証明. $p_{jj}^{(0)} = 1$ より $n = 1$ のときは明らか．$n \geq 2$ のとき，$\{X_n = j\} \subset \{\tau_j \leq n\}$ より，

$$p_{ij}^{(n)} = \mathbb{P}_i(X_n = j) = \mathbb{P}_i(X_n = j, \tau_j \leq n) = \sum_{k=1}^{n} \mathbb{P}_i(X_n = j, \tau_j = k).$$

今，$k \leq n-1$ に対して，

$$\mathbb{P}_i(\tau_j = k, X_n = j) = \sum_{\substack{i_1, \ldots, i_{k-1} \neq j \\ j_1, \ldots, j_{n-k-1} \in S}} \mathbb{P}_i(X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = j, X_{k+1} = j_1, \ldots, X_{n-1} = j_{n-k-1}, X_n = j)$$

である．ここで $i_1, \ldots, i_{k-1}$ に関する和は $\mathbb{P}_i(X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = j) > 0$

を満たすものについてとる．それぞれの確率は

$$
\begin{aligned}
&\mathbb{P}_i(X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = j, X_{k+1} = j_1, \ldots, X_{n-1} = j_{n-k-1}, X_n = j) \\
&= \mathbb{P}_i(X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = j) \\
&\quad \times \mathbb{P}_i(X_{k+1} = j_1, \ldots, X_{n-1} = j_{n-k-1}, X_n = j \mid X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = j) \\
&= \mathbb{P}_i(X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = j)\mathbb{P}_j(X_1 = j_1, \ldots, X_{n-k-1} = j_{n-k-1}, X_{n-k} = j) \\
&= \mathbb{P}_i(X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, \tau_j = k)\mathbb{P}_j(X_1 = j_1, \ldots, X_{n-k-1} = j_{n-k-1}, X_{n-k} = j)
\end{aligned}
$$

と表せるので，$i_1, \ldots, i_{k-1}, j_1, \ldots, j_{n-k-1}$ について和をとることにより，

$$
\mathbb{P}_i(\tau_j = k, X_n = j) = \mathbb{P}_i(\tau_j = k)\mathbb{P}_j(X_{n-k} = j) = f_{ij}^{(k)} p_{jj}^{(n-k)}
$$

を得る．また，

$$
\mathbb{P}_i(X_n = j, \tau_j = n) = \mathbb{P}_i(\tau_j = n) = f_{ij}^{(n)} = f_{ij}^{(n)} p_{jj}^{(0)}
$$

であるから，以上を併せて命題の主張が得られた． □

命題 **1.5.2.** 状態 $i \in S$ が再帰的であるための必要十分条件は $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$ である．

証明．命題 1.5.1 より，$N > 1$ に対して，

$$
\begin{aligned}
\sum_{n=0}^{N} p_{ii}^{(n)} &= 1 + \sum_{n=1}^{N} p_{ii}^{(n)} = 1 + \sum_{n=1}^{N}\sum_{k=1}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)} \\
&= 1 + \sum_{k=1}^{N}\sum_{n=k}^{N} p_{ii}^{(n-k)} f_{ii}^{(k)} = 1 + \sum_{k=1}^{N} f_{ii}^{(k)} \sum_{n=0}^{N-k} p_{ii}^{(n)}.
\end{aligned}
\tag{1.5.1}
$$

ここで $f_{ii} = 1$ を仮定する．もし $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$ ならば，(1.5.1) で $N \to \infty$ とすると，$\sum_{k=1}^{\infty} f_{ii}^{(k)} = f_{ii}$ より

$$
\sum_{n=0}^{\infty} p_{ii}^{(n)} = 1 + \sum_{n=0}^{\infty} p_{ii}^{(n)}.
$$

これより $0 = 1$ が導かれてしまうので，$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$ でなくてはならない．

逆に，$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$ を仮定する．(1.5.1) より，

$$
\sum_{n=0}^{N} p_{ii}^{(n)} \left(1 - \sum_{k=1}^{N} f_{ii}^{(k)}\right) = 1 - \sum_{k=1}^{N} f_{ii}^{(k)} \sum_{n=N-k+1}^{N} p_{ii}^{(n)}
$$

であるから，

$$
\sum_{n=0}^{N} p_{ii}^{(n)} (1 - f_{ii}) \le \sum_{n=0}^{N} p_{ii}^{(n)} \left(1 - \sum_{k=1}^{N} f_{ii}^{(k)}\right) \le 1.
$$

よって $N \to \infty$ として $\infty \times (1 - f_{ii}) \le 1$ となり，これが成り立つのは $f_{ii} = 1$ のときに限る． □

例 **1.5.3.** 1 次元単純ランダムウォークの場合，奇数ステップで状態 0 に戻ってくることはないので，$p_{00}^{(2n+1)} = 0$ である．偶数ステップについては，

$$p_{00}^{(2n)} = \mathbb{P}(n \text{ 回右に進み，} n \text{ 回左に進む}) = C_n^{2n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n.$$

ただし，$C_n^m = m!/((m-n)!n!)$．ここでスターリングの公式

$$n! \sim \sqrt{2\pi} e^{-n} n^{n+1/2}, \quad n \to \infty$$

（数列 $\{a_n\}$, $\{b_n\}$ に対して，$a_n \sim b_n$, $n \to \infty$, は $\lim_{n\to\infty}(a_n/b_n) = 1$ を意味することを思い出そう）より，

$$C_n^{2n} = \frac{(2n)!}{(n!)^2} \sim \frac{4^n}{\sqrt{\pi n}}, \quad n \to \infty.$$

よって，

$$p_{00}^{(2n)} \sim \frac{1}{\sqrt{\pi n}}, \quad n \to \infty$$

であるから $\sum_{n=0}^{\infty} p_{00}^{(n)} = \infty$ を得る．ゆえに命題 1.5.2 より状態 0 は再帰的である．

次に，状態 j に訪れる回数を N_j で表すと，

$$N_j = \sum_{n=1}^{\infty} 1_{\{X_n = j\}}$$

であり，これは

$$\mathbb{E}_i[N_j] = \sum_{n=1}^{\infty} \mathbb{P}_i(X_n = j) = \sum_{n=1}^{\infty} p_{ij}^{(n)}$$

を満たす．ここで $\mathbb{E}_i$ は $\mathbb{P}_i$ の下での期待値を表し，最初の等号の導出には命題 A.3.2(1) を用いた．命題 1.5.2 より，i が再帰的であるための必要十分条件は $\mathbb{E}_i N_i = +\infty$ となる．

期待値だけでなく，N_j の分布そのものを明示的に求めることもできる．

定理 **1.5.4.** $i, j \in S$ に対して，

$$\mathbb{P}_i(N_j = k) = \begin{cases} 1 - f_{ij} & \text{if } k = 0, \\ f_{ij} f_{jj}^{k-1}(1 - f_{jj}) & \text{if } k \geq 1. \end{cases}$$

証明*. まず，$\mathbb{P}_i(N_j \geq 1) = \mathbb{P}_i(\tau_j < \infty) = f_{ij}$ より $\mathbb{P}_i(N_j = 0) = 1 - \mathbb{P}_i(N_j \geq 1) = 1 - f_{ij}$ である．

次に，j に 2 回目に到達する時刻

$$\tau_j(2) = \inf\{n > \tau_j : X_n = j\}$$

を考えると，

$$\begin{aligned} \mathbb{P}_i(N_j \geq 2) = \mathbb{P}_i(\tau_j(2) < \infty) &= \sum_{n\geq 1} \mathbb{P}_i(\tau_j(2) = \tau_j + n) \\ &= \sum_{n,m\geq 1} \mathbb{P}_i(\tau_j = m, \tau_j(2) = m + n) \end{aligned}$$

となる．ここで，m は $\mathbb{P}_i(\tau_j = m) > 0$ を満たす範囲でとれば十分である．このとき，命題 1.1.10 を適用し，かつ $\mathbb{P}_i(X_m = j) \geq \mathbb{P}_i(\tau_j = m) > 0$ に注意すれば

$$\begin{aligned}
&\mathbb{P}_i(\tau_j = m, \tau_j(2) = m+n) \\
&= \sum_{\substack{i_1,\ldots,i_{m-1}\neq j \\ j_1,\ldots,j_{n-1}\neq j}} \mathbb{P}_i(X_p = i_p\,(1 \leq p \leq m-1), X_m = j, X_{m+q} = j_q\,(1 \leq q \leq n-1), X_{m+n} = j) \\
&= \sum_{\substack{i_1,\ldots,i_{m-1}\neq j \\ j_1,\ldots,j_{n-1}\neq j}} \mathbb{P}_i(X_1 = i_1, \ldots, X_{m-1} = i_{m-1}, X_m = j) \\
&\qquad\qquad \times \mathbb{P}_i(X_{m+1} = j_1, \ldots, X_{m+n-1} = j_{n-1}, X_{m+n} = j \,|\, X_m = j)
\end{aligned}$$

を得る．さらに，命題 1.1.9 と $\mathbb{P}_j(X_0 = j) = 1$ より

$$\begin{aligned}
&\mathbb{P}_i(X_{m+1} = j_1, \ldots, X_{m+n-1} = j_{n-1}, X_{m+n} = j \,|\, X_m = j) \\
&= p_{jj_1} \cdots p_{j_{n-1}j} \\
&= \mathbb{P}_j(X_1 = j_1, \ldots, X_{n-1} = j_{n-1}, X_n = j).
\end{aligned}$$

よって，

$$\begin{aligned}
\mathbb{P}_i(\tau_j = m, \tau_j(2) = m+n) &= \sum_{i_1,\ldots,i_{m-1}\neq j} \mathbb{P}_i(X_1 = i_1, \ldots, X_{m-1} = i_{m-1}, X_m = j) \\
&\quad \times \sum_{j_1,\ldots,j_{n-1}\neq j} \mathbb{P}_j(X_1 = j_1, \ldots, X_{n-1} = j_{n-1}, X_n = j) \\
&= \mathbb{P}_i(\tau_j = m)\mathbb{P}_j(\tau_j = n).
\end{aligned}$$

ゆえに

$$\mathbb{P}_i(N_j \geq 2) = \mathbb{P}_i(\tau_j < \infty)\mathbb{P}_j(\tau_j < \infty) = f_{ij}f_{jj}.$$

一般の $k \geq 2$ に対しては，k 回目に到達する時刻を導入し，（煩雑だが）上と同様の議論を行うことにより，

$$\mathbb{P}_i(N_j \geq k) = \mathbb{P}_i(\tau_j < \infty)\mathbb{P}_j(\tau_j < \infty)^{k-1} = f_{ij}f_{jj}^{k-1} \tag{1.5.2}$$

が得られるので，これより

$$\mathbb{P}_i(N_j = k) = \mathbb{P}_i(N_j \geq k) - \mathbb{P}_i(N_j \geq k+1) = f_{ij}f_{jj}^{k-1}(1 - f_{jj})$$

が従う． □

系 **1.5.5.**

(1) 状態 $j \in S$ が再帰的ならば，

$$\mathbb{P}_i(N_j = \infty) = f_{ij}.$$

従って，$\mathbb{P}_j(N_j = \infty) = 1$ である．

(2) j が非再帰的ならば，

$$\mathbb{E}_i[N_j] = \frac{f_{ij}}{1 - f_{jj}} < \infty.$$

従って，$\mathbb{P}_i(N_j < \infty) = 1$ である．

証明. j が再帰的とすると，定義より $f_{jj} = \mathbb{P}_j(\tau_j < \infty) = 1$. よって (1.5.2) と命題 A.1.10 より

$$\mathbb{P}_i(N_j = \infty) = \lim_{k\to\infty} \mathbb{P}_i(N_j \geq k) = f_{ij}.$$

j が非再帰的のとき，$f_{jj} < 1$ であり，(1.5.2) と命題 A.3.2(2) より

$$\mathbb{E}_i[N_j] = \sum_{n=0}^{\infty} \mathbb{P}_i(N_j > n) = \sum_{n=0}^{\infty} f_{ij} f_{jj}^m = \frac{f_{ij}}{1 - f_{jj}} < \infty.$$

□

次に，A を S の部分集合とし，マルコフ連鎖が A に到達する確率を求めよう．すなわち，

$$\tau_A = \inf\{n \in \mathbb{N} : X_n \in A\}$$

とおき，

$$f(i) := \mathbb{P}_i(\tau_A < \infty), \quad i \in S$$

を求める問題を考える．

まず，

$$f(i) = \sum_{j\in S} \mathbb{P}_i(X_1 = j, \tau_A < \infty) = \sum_{j\in A} p_{ij} + \sum_{j\in S\setminus A} \mathbb{P}_i(X_1 = j, \tau_A < \infty)$$

である．ここで，$X_1 = j$ を出発点として考えると，マルコフ性より直感的には $\mathbb{P}_i(X_1 = j, \tau_A < \infty) = p_{ij}\mathbb{P}_j(\tau_A < \infty)$ である．よって，これを認めれば，

$$f(i) = \mathbb{P}_i(\tau_A = 1) + \sum_{j\in S\setminus A} p_{ij} f(j), \quad i \in S \setminus A$$

となり，f を $S \setminus A$ に制限したもの $f|_{S\setminus A}$ に関する方程式と見なせる．ゆえに，この方程式を解くことで $i \in S \setminus A$ から出発した場合の A への到達確率を求めることができる．このことを正当化しよう．

定理 **1.5.6.** $S \setminus A$ 上の関数 $g = f|_{S\setminus A}$ は方程式

$$g(i) = \mathbb{P}_i(\tau_A = 1) + \sum_{j\in S\setminus A} p_{ij} g(j), \quad i \in S \setminus A \tag{1.5.3}$$

の非負の解の中で最小の解である．

証明. 任意の $j \in S \setminus A$ をとり固定する．このとき，$\mathbb{P}_i(X_1 = j, \tau_A = 1) = 0$ であり，

$$\begin{aligned}\mathbb{P}_i(X_1 = j, \tau_A = 2) &= \mathbb{P}_i(X_1 = j, X_2 \in A) = \sum_{j_2\in A} p_{ij} p_{jj_2} = p_{ij}\mathbb{P}_j(X_1 \in A) \\ &= p_{ij}\mathbb{P}_j(\tau_A = 1)\end{aligned}$$

である．$n \geq 3$ に対しては，

$$
\begin{aligned}
\mathbb{P}_i(X_1 = j, \tau_A = n) &= \mathbb{P}_i(X_1 = j, X_2 \notin A, \dots, X_{n-1} \notin A, X_n \in A) \\
&= \sum_{\substack{j_2,\dots,j_{n-1} \in S\setminus A \\ j_n \in A}} \mathbb{P}_i(X_1 = j, X_2 = j_2, \dots, X_{n-1} = j_{n-1}, X_n = j_n) \\
&= \sum_{\substack{j_2,\dots,j_{n-1} \in S\setminus A \\ j_n \in A}} p_{ij} p_{jj_2} \cdots p_{j_{n-1}j_n} \\
&= p_{ij} \sum_{\substack{j_2,\dots,j_{n-1} \in S\setminus A \\ j_n \in A}} \mathbb{P}_j(X_1 = j_2, \dots, X_{n-2} = j_{n-1}, X_{n-1} = j_n) \\
&= p_{ij} \mathbb{P}_j(\tau_A = n-1).
\end{aligned}
$$

まとめると，

$$
\mathbb{P}_i(X_1 = j, \tau_A = n) = p_{ij} \mathbb{P}_j(\tau_A = n-1), \quad n \geq 2. \tag{1.5.4}
$$

従って，

$$
\begin{aligned}
\mathbb{P}_i(X_1 = j, \tau_A < \infty) = \sum_{n=2}^{\infty} \mathbb{P}_i(X_1 = j, \tau_A = n) &= p_{ij} \sum_{n=2}^{\infty} \mathbb{P}_j(\tau_A = n-1) \\
&= p_{ij} \mathbb{P}_j(\tau_A < \infty) = p_{ij} f(j).
\end{aligned}
$$

ゆえに $f|_{S\setminus A}$ は (1.5.3) の解である．

次に，$S\setminus A$ 上の非負関数 g が (1.5.3) の解であるとすると，$g(i) \geq \mathbb{P}_i(\tau_A = 1), i \in S\setminus A$ である．これを (1.5.3) に代入すれば，(1.5.4) より

$$
\begin{aligned}
g(i) &\geq \mathbb{P}_i(\tau_A = 1) + \sum_{j \in S\setminus A} p_{ij} \mathbb{P}_j(\tau_A = 1) = \mathbb{P}_i(\tau_A = 1) + \sum_{j \in S\setminus A} \mathbb{P}_i(X_1 = j, \tau_A = 2) \\
&= \mathbb{P}_i(\tau_A = 1) + \mathbb{P}_i(X_1 \notin A, \tau_A = 2) = \mathbb{P}_i(\tau_A = 1) + \mathbb{P}_i(\tau_A = 2).
\end{aligned}
$$

これを再び (1.5.3) に代入し，

$$
g(i) \geq \mathbb{P}_i(\tau_A = 1) + \mathbb{P}_i(\tau_A = 2) + \mathbb{P}_i(\tau_A = 3).
$$

この手続きを繰り返すことにより，

$$
g(i) \geq \sum_{n=1}^{\infty} \mathbb{P}_i(\tau_A = n) = \mathbb{P}_i(\tau_A < \infty) = f(i), \quad i \in S \setminus A
$$

を得る．ゆえに $f|_{S\setminus A}$ の最小性が従う． □

例 **1.5.7.** 推移確率 P が

$$
P = \begin{array}{c} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \end{array}
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0.1 & 0.2 & 0.2 & 0.2 & 0.2 & 0.1 \\
0.1 & 0.1 & 0.3 & 0.3 & 0.1 & 0.1 \\
0 & 0 & 0.4 & 0.3 & 0.2 & 0.1 \\
0.1 & 0.2 & 0.3 & 0.4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
$$

により与えられるとする．このとき，直感的には明らかに，マルコフ連鎖が状態 $2,3,4,5$ のいずれから出発しても，やがては吸収状態 1 と 6 のどちらかに必ず到達する．このことを定理 1.5.6 を使って計算で確かめてみよう．

まず，計算をし易くするため，状態を並び替えて行列

$$\tilde{P}=\begin{array}{c}\mathbf{2}\\\mathbf{3}\\\mathbf{4}\\\mathbf{5}\\\mathbf{1}\\\mathbf{6}\end{array}\begin{pmatrix}0.2&0.2&0.2&0.2&0.1&0.1\\0.1&0.3&0.3&0.1&0.1&0.1\\0&0.4&0.3&0.2&0&0.1\\0.2&0.3&0.4&0&0.1&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{pmatrix}$$

を考える．$A=\{1,6\}$ とすると，

$$R:=\begin{pmatrix}\mathbb{P}_2(X_1\in A)\\\mathbb{P}_3(X_1\in A)\\\mathbb{P}_4(X_1\in A)\\\mathbb{P}_5(X_1\in A)\end{pmatrix}=\begin{pmatrix}0.2\\0.2\\0.1\\0.1\end{pmatrix}$$

であり，

$$F=\begin{pmatrix}g(2)\\g(3)\\g(4)\\g(5)\end{pmatrix},\quad Q=\begin{pmatrix}0.2&0.2&0.2&0.2\\0.1&0.3&0.3&0.1\\0&0.4&0.3&0.2\\0.2&0.3&0.4&0\end{pmatrix}$$

とおくことにより，方程式 (1.5.3) は

$$F=R+QF \tag{1.5.5}$$

と表される．$\mathbb{R}^{4\times4}$ の単位行列を I_4 と書くとき，I_4-Q が正則ならば (1.5.5) は一意解 $F=(I_4-Q)^{-1}R$ を持ち，定理 1.5.6 より，それが求める到達確率を与える．実際に計算してみると

$$(I_4-Q)^{-1}=\begin{pmatrix}1.71523179&1.9205298&1.82781457&0.90066225\\0.58278146&3.04635762&1.93377483&0.80794702\\0.54304636&2.38410596&3.16556291&0.98013245\\0.73509934&2.25165563&2.21192053&1.81456954\end{pmatrix}$$

となり，各 $i=2,3,4,5$ に対して（わずかに誤差はあるが）$f(i)=1.0$ となる．

系 **1.5.8.** A を S の部分集合，$h(i)$ を $i\in A$ から出発し A から出ない確率とすると，$h(i)$, $i\in A$, は

$$h(i)=\sum_{j\in A}p_{ij}h(j),\quad h(i)\le 1, \tag{1.5.6}$$

の最大解である．

例 **1.5.9.** 例 1.1.4 において，$N\ge 3$ として，$A:=\{100,\dots,100(N-1)\}$ の状態から出発し A に留まり続ける確率，すなわち，プラスの持ち金から始めてゲームが永遠に続く確率を求めてみよう．

方程式 (1.5.6) は

$$\begin{cases} h(100m) = 0.6h(100(m-1)) + 0.4h(100(m+1)), \quad m \in \{2, \ldots, N-2\}, \\ h(100) = 0.4h(200), \quad h(100(N-1)) = 0.6h(100(N-2)) \end{cases}$$

となる．これより，$h(200) = (10/19)h(300)$ が得られ，同様にして，任意の $m \in \{1, \ldots, N-2\}$ に対して $h(100m) = q_m h(100(m+1))$ を満たすような $q_m \in (0,1)$ が存在することが分かる．よって，$h(100(N-1)) = 0.6q_{N-2}h(100(N-1))$ であり，従って $h(100(N-1)) = 0$. ゆえに任意の $m \in \{1, \ldots, N-2\}$ に対して $h(100m) = 0$ となる．すなわち，ギャンブラーは破産するか望む収益 $100N$ を得るかのどちらかである．

1.6 定常分布

この節では，ある初期分布から出発したマルコフ連鎖がその推移確率によって最終的にどのような分布に落ち着くのか，という問題を考えよう．そのために，まずは確率測度の定義を思い出しておこう．$\pi = \{\pi_i\}_{i \in S}$ が S 上の確率測度であるとは，π が S 上の非負関数であり $\sum_{i \in S} \pi_i = 1$ を満たすときにいうのであった．

定義 **1.6.1.** S 上の確率測度 $\pi = \{\pi_i\}_{i \in S}$ が推移確率 $P = \{p_{ij}\}_{i,j \in S}$ を持つマルコフ連鎖の**定常分布**（stationary distribution）あるいは**不変分布**（invariant distribution）であるとは，

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}, \quad j \in S \tag{1.6.1}$$

が成り立つときにいう．

- P を $\mathbb{R}^{|S| \times |S|}$ の（無限）行列と，π を $\mathbb{R}^{|S|}$ の（無限）行ベクトルと同一視するとき，(1.6.1) は

  $$\pi = \pi P \tag{1.6.2}$$

  と表せる．
- (1.6.2) より任意の $m \in \mathbb{N}$ に対して

  $$\pi = \pi P^2 = \cdots = \pi P^m.$$

  よって π が定常分布であるということは推移行列を何回作用させても変わらないようなベクトルであることを意味する．

例 **1.6.2.** $\alpha, \beta \in (0,1)$ とする．推移確率

$$P = \mathbf{\frac{1}{2}} \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

を持つマルコフ連鎖の場合，条件 (1.6.1) は

$$\pi_1 = \pi_1(1-\alpha) + \pi_2\beta$$
$$\pi_2 = \pi_1\alpha + \pi_2(1-\beta)$$

となり，これより $\pi_1\alpha = \pi_2\beta$ を得る．これと確率の条件 $\pi_1 + \pi_2 = 1$ より定常分布は

$$\pi_1 = \frac{\beta}{\alpha+\beta}, \quad \pi_2 = \frac{\alpha}{\alpha+\beta}$$

により与えられる．

例 **1.6.3.** 1 次元単純ランダムウォークの場合，条件 (1.6.1) は

$$\pi_i = \frac{1}{2}\pi_{i+1} + \frac{1}{2}\pi_{i-1}, \quad i \in \mathbb{Z}$$

となる．これより，

$$\pi_i = \pi_0 + (\pi_1 - \pi_0)i, \quad i \in \mathbb{Z}$$

を得る．他方，$\{\pi_i\}$ が確率測度になるためには $\pi_i \in [0,1]$ でなければならない．よって $\pi_1 - \pi_0 = 0$ となり，$\pi_i = \pi_0$, $i \in \mathbb{Z}$, が従う．しかし，π_0 の値にかかわらず，これは $\mathbb{Z}$ 上の確率測度を定義しない．結局，1 次元単純ランダムウォークには定常分布は存在しない．

定常分布の存在と一意性について議論するため，もう少し広いクラスのベクトルで定常性を考えることからはじめよう．

定義 **1.6.4.** S 上の非負実数値関数 $\nu = \{\nu_i\}_{i\in S}$ が推移確率 $P = \{p_{ij}\}_{i,j\in S}$ を持つマルコフ連鎖の**不変測度**（invariant measure）であるとは，

$$\nu_j = \sum_{i\in S} \nu_i p_{ij}, \quad j \in S \tag{1.6.3}$$

が成り立つときにいう．

例 **1.6.5.** 1 次元ランダムウォークを考える．すなわち，$S = \mathbb{Z}$ で，$p_{i,i+1} = p$, $p_{i,i-1} = 1-p$, $i \in S$ のマルコフ連鎖を考える．ただし，$p \in (0,1)$ とする．このとき，$\nu_i = 1$, $i \in S$ は不変測度である．$p \neq 1/2$ の場合，これに加えて，$\nu_i = (p/(1-p))^i$, $i \in S$ も不変測度である．

固定された $i \in S$ に対し，

$$\nu_j = \mathbb{E}_i \sum_{n=0}^{\tau_i - 1} 1_{\{X_n = j\}} \tag{1.6.4}$$

とおく．ν_j は i に初めて戻ってくる前に j に滞在した回数の期待値である．従って特に，$\nu_i = 1$ である．

定理 **1.6.6.** $i \in S$ が再帰的のとき，(1.6.4) で定義される $\nu = \{\nu_j\}_{j\in S}$ は不変測度である．

証明*. 仮定より，$\mathbb{P}_i(\tau_i < \infty) = 1$ である．$j \neq i$ とするとき，$\{X_0 = j\} = \{X_n =$

$j, \tau_i = n\} = \emptyset$ より，

$$
\begin{aligned}
\nu_j &= \mathbb{E}_i \sum_{n=1}^{\infty} 1_{\{X_n = j, \tau_i > n\}} = \mathbb{E}_i \sum_{n=1}^{\infty} 1_{\{X_n = j, \tau_i \geq n\}} \\
&= p_{ij} + \sum_{n=2}^{\infty} \mathbb{P}_i(X_n = j, \tau_i \geq n) = p_{ij} + \sum_{k \neq i} \sum_{n=2}^{\infty} \mathbb{P}_i(X_n = j, \tau_i \geq n, X_{n-1} = k).
\end{aligned}
$$

$n = 2$, $k \neq i$ に対して，

$$
\begin{aligned}
\mathbb{P}_i(X_2 = j, \tau_i \geq 2, X_1 = k) &= \mathbb{P}_i(X_n = j, X_1 = k) \\
&= p_{jk} \mathbb{P}_i(X_1 = k) = p_{kj} \mathbb{P}_i(X_1 = k, \tau_i \geq 2).
\end{aligned}
$$

$n \geq 3$, $k \neq i$ に対して，$\{\tau_i \geq n, X_{n-1} = k\} = \{X_1 \neq i, \ldots, X_{n-2} \neq i, X_{n-1} = k\}$ より，

$$
\begin{aligned}
&\mathbb{P}_i(X_n = j, \tau_i \geq n, X_{n-1} = k) \\
&= \sum_{i_1, \ldots, i_{n-2} \neq i} \mathbb{P}_i(X_1 = i_1, \ldots, X_{n-2} = i_{n-2}, X_{n-1} = k, X_n = j) \\
&= \sum_{i_1, \ldots, i_{n-2} \neq i} p_{kj} \mathbb{P}_i(X_1 = i_1, \ldots, X_{n-2} = i_{n-2}, X_{n-1} = k) \\
&= p_{kj} \mathbb{P}_i(X_{n-1} = k, \tau_i \geq n).
\end{aligned}
$$

従って，

$$
\begin{aligned}
\nu_j &= p_{ij} + \sum_{k \neq i} \sum_{n=2}^{\infty} p_{kj} \mathbb{P}_i(X_{n-1} = k, \tau_i \geq n) \\
&= \nu_i p_{ij} + \sum_{k \neq i} p_{kj} \sum_{n=1}^{\infty} \mathbb{P}_i(X_n = k, \tau_i \geq n + 1) \\
&= \nu_i p_{ij} + \sum_{k \neq i} p_{kj} \mathbb{E}_i \sum_{n=1}^{\tau_i - 1} 1_{\{X_n = k\}}
\end{aligned}
$$

を得る．ここでさらに $k \neq i$ であることを使うと，

$$
\nu_j = \nu_i p_{ij} + \sum_{k \neq i} p_{kj} \mathbb{E}_i \sum_{n=0}^{\tau_i - 1} 1_{\{X_n = k\}} = \nu_i p_{ij} + \sum_{k \neq i} p_{kj} \nu_k = \sum_{k \in S} \nu_k p_{kj} \tag{1.6.5}
$$

となる．

他方，$k \neq i$ と $\mathbb{P}_i(X_n = k, \tau_i > n) > 0$ を満たす $n \geq 1$ に対して，上と同様の議論により

$$
p_{ki} = \mathbb{P}_i(X_{n+1} = i \,|\, X_n = k) = \mathbb{P}_i(X_{n+1} = i \,|\, X_n = k, \tau_i > n)
$$

が従うので，

$$
\mathbb{P}_i(X_n = k, \tau_i > n) p_{ki} = \mathbb{P}_i(X_{n+1} = i, X_n = k, \tau_i > n) = \mathbb{P}_i(X_n = k, \tau_i = n + 1).
$$

この等式は $\mathbb{P}_i(X_n = k, \tau_i > n) = 0$ のときも成立する．よって，

$$\begin{aligned}\sum_{k\neq i} \nu_k p_{ki} &= \sum_{k\neq i}\sum_{n=1}^{\infty} \mathbb{P}_i(X_n = k, \tau_i > n)p_{ki} = \sum_{k\neq i}\sum_{n=1}^{\infty} \mathbb{P}_i(X_n = k, \tau_i = n+1)\\ &= \sum_{n=1}^{\infty}\sum_{k\neq i} \mathbb{P}_i(X_n = k, \tau_i = n+1) = \sum_{n=1}^{\infty} \mathbb{P}_i(\tau_i = n+1).\end{aligned}$$

ここで，一般に非負数列 $\{a_{m,n}\}_{m,n\geq 1}$ に対して $\sum_{m\geq 1}\sum_{n\geq 1} a_{m,n} = \sum_{n\geq 1}\sum_{m\geq 1} a_{m,n}$ であることを用いた．さらに，$\nu_i p_{ii} = \mathbb{P}_i(X_1 = i) = \mathbb{P}_i(\tau_i = 1)$ であるから，結局，

$$\sum_{k\in S} \nu_k p_{ki} = \sum_{n=0}^{\infty} \mathbb{P}_i(\tau_i = n+1) = \mathbb{P}_i(\tau_i < \infty) = 1 = \nu_i$$

が従う．これと (1.6.5) より，ν は (1.6.3) を満たす．

最後に $\nu_j < \infty$ を示そう．もし $j \to i$ ならば，ある $m \geq 1$ が存在して，$p_{ji}^{(m)} > 0$. 他方，$\nu = \nu P$ より，$\nu = \nu P^m$. よって，

$$1 = \sum_{k\in S} v_k p_{ki}^{(m)} \geq \nu_j p_{ji}^{(m)}.$$

両辺 $p_{ji}^{(m)}$ で割ることにより $\nu_j < \infty$ を得る．

次に $j \nrightarrow i$ とする．このとき，$n < m$ に対して命題 1.1.9 より

$$\begin{aligned}&\mathbb{P}_i(X_n = j, \tau_i = m)\\ &= \mathbb{P}_i(X_n = j, X_1 \neq i, \cdots, X_{n-1} \neq i, X_{n+1} \neq i, \cdots, X_{m-1} \neq i, X_m = i)\\ &= \sum_{\substack{j_1,\ldots,j_{m-1}\neq i\\ j_n = j}} \mathbb{P}_i(X_1 = j_1, \ldots, X_{m-1} = j_{m-1}, X_m = i)\\ &= \sum_{\substack{j_1,\ldots,j_{m-1}\neq i\\ j_n = j}} p_{ij_1}\cdots p_{j_{m-1}i}\\ &= \sum_{\substack{j_1,\ldots,j_{m-1}\neq i\\ j_n = j}} \mathbb{P}_i(X_1 = j_1, \ldots, X_n = j_n)\mathbb{P}_j(X_1 = j_{n+1}, \ldots, X_{m-n-1} = j_{m-1}, X_{n-m} = i)\\ &= \mathbb{P}_i(\tau_i > n, X_n = j)\mathbb{P}_j(\tau_i = m-n)\end{aligned}$$

であるから，$j \nrightarrow i$ より $\mathbb{P}_i(X_n = j, \tau_i = m) = 0$ となる．よって，

$$\nu_j = \sum_{m=1}^{\infty} \mathbb{E}_i \sum_{n=0}^{m-1} 1_{\{X_n = j, \tau_i = m\}} = \sum_{m=1}^{\infty}\sum_{n=0}^{m-1} \mathbb{P}_i(X_n = j, \tau_i = m) = 0.$$

従って特に $\nu_j < \infty$ を得る． □

- $i \in S$ が $\mathbb{E}_i[\tau_i] < \infty$ を満たすとき，i を**正再帰的**（positive recurrent）であるという．
- 正再帰的なら再帰的である．
- $i \in S$ が再帰的ではあるが正再帰的でないとき，**零再帰的**（null recurrent）という．

定理 **1.6.7.** マルコフ連鎖が既約で再帰的のとき，不変測度は定数倍を除いて一意的に定まり，すべての点で測度の値は正である．さらに，マルコフ連鎖が既約で正再帰的のとき，定常分布 $\pi = \{\pi_j\}_{j \in S}$ が一意的に存在し，

$$\pi_j = \frac{1}{\mathbb{E}_j[\tau_j]}, \quad j \in S$$

により与えられる．

証明*. $i \in S$ を固定し，(1.6.4) により定義される ν を考える．任意の $j \in S$ について，既約性より $i \to j$ なので，ある m が存在して $p_{ij}^{(m)} > 0$. 前の定理より ν は不変測度であるから $\nu = \nu P^m$ となり，これより

$$\nu_j = \sum_{k \in S} \nu_k p_{kj}^{(m)} \geq p_{ij}^{(m)} > 0.$$

ゆえに ν は正値である．

$\mu = \{\mu_j\}_{j \in S}$ を任意の不変測度とする．上と同じ議論により，$\mu_j > 0$, $j \in S$ が従う．μ を μ_i で割ってできる測度を考え，これも記号を流用して μ により表すことにする．すなわち，μ は $\mu_i = 1$, $\mu_j > 0$, $j \in S$ を満たす不変測度である．

今，

$$\mu_j = p_{ij} + \sum_{k \neq i} \mu_k p_{kj}, \quad j \in S \tag{1.6.6}$$

より，$\mu_j \geq p_{ij}$, $j \in S$ である．これを (1.6.6) に代入すれば

$$\mu_j \geq p_{ij} + \sum_{k \neq i} p_{ik} p_{kj}.$$

これを再び (1.6.6) に代入すると，

$$\begin{aligned}
\mu_j &\geq p_{ij} + \sum_{k_2 \neq i} \left(p_{ik_2} + \sum_{k_1 \neq i} p_{ik_1} p_{k_1 k_2} \right) p_{k_2 j} \\
&= p_{ij} + \sum_{k_2 \neq i} p_{ik_2} p_{k_2 j} + \sum_{k_1, k_2 \neq i} p_{ik_1} p_{k_1 k_2} p_{k_2 j} \\
&= \mathbb{P}_i(X_1 = j) + \mathbb{P}_i(X_1 \neq i, X_2 = j) + \mathbb{P}_i(X_1 \neq i, X_2 \neq i, X_3 = j) \\
&= \mathbb{P}_i(X_1 = j) + \mathbb{P}_i(\tau_i > 1, X_2 = j) + \mathbb{P}_i(\tau_i > 2, X_3 = j).
\end{aligned}$$

これを繰り返すことにより，任意の $m \geq 2$ に対して

$$\mu_j \geq \sum_{n=0}^{m} \mathbb{P}_i(\tau_i > n, X_{n+1} = j).$$

他方，$j \neq i$ のとき，

$$\nu_j = \sum_{n=1}^{\infty} \mathbb{P}_i(X_n = j, \tau_i \geq n) = \sum_{n=0}^{\infty} \mathbb{P}_i(X_{n+1} = j, \tau_i \geq n+1)$$

であるから，$\mu_j \geq \nu_j$ が従う．$\mu_i = \nu_i = 1$ であるから，結局，すべての $j \in S$ に対して $\mu_j \geq \nu_j$ である．よって

$$1 = \mu_i = \sum_{j\in S} \mu_j p_{ji}^{(n)} \geq \sum_{j\in S} \nu_j p_{ji}^{(n)} = \nu_i = 1.$$

ゆえに $\sum_{j\in S}(\mu_j - \nu_j)p_{ji}^{(n)} = 0$ であり，従って，任意の $j \in S$ に対して $(\mu_j - \nu_j)p_{ji}^{(n)} = 0$. ここで既約性より $p_{ji}^{(n)} > 0$ なる n が存在するから，（各 j に対し予めこのような n で上の議論を進めることで）$\mu_j = \nu_j$ が得られる．以上より一意性が従う．

最後に，マルコフ連鎖が既約かつ正再帰的であるとき，

$$\sum_{j\in S} \nu_j = \sum_{j\in S} \mathbb{E}_i \sum_{n=0}^{\tau_i - 1} 1_{\{X_n = j\}} = \mathbb{E}_i \sum_{n=0}^{\tau_i - 1} \sum_{j\in S} 1_{\{X_n = j\}} = \mathbb{E}_i \sum_{n=0}^{\tau_i - 1} 1 = \mathbb{E}_i \tau_i < \infty.$$

よって $\pi_j = \nu_j / \mathbb{E}_i \tau_i$, $j \in S$ は確率分布を定義し，従って定常分布である．ところで，$\pi_i = 1/\mathbb{E}_i \tau_i$ となり，i は証明の最初に任意に固定したものだったから，結局，どの i を用いても π は

$$\pi_i = \frac{1}{\mathbb{E}_i \tau_i}$$

を満たすことになり，これは最後の主張を意味する． □

上の定理では既約なマルコフ連鎖が正再帰的のとき，定常分布が一意に存在することが分かった．実は逆も成り立つ．

定理 **1.6.8.** 既約なマルコフ連鎖に対して，定常分布が一意に存在することとすべての状態が正再帰的であることは同値である．

証明は割愛する．例えば [15, 命題 11.19] を参照のこと．

1.7 極限定理

本節では，マルコフ連鎖の状態集合 S は有限集合（i.e., $|S| < \infty$）と仮定する．

- ある $n \in \mathbb{N}$ が存在して，すべての $i, j \in S$ に対して $p_{ij}^{(n)} > 0$ のとき，推移確率 P あるいはそれを持つマルコフ連鎖は**エルゴード的**（ergodic）という．
- すなわち，エルゴード的とは，ある（十分大きな）ステップ後には，いかなる状態にも遷移し得ることを意味する．
- 推移確率 P が**正行列**（positive matrix）を定義する場合，i.e., P の各成分が正となる場合，マルコフ連鎖はエルゴード的である．
- 吸収状態を持つマルコフ連鎖はエルゴード的ではない．

命題 **1.7.1.** 次の条件 (1)–(3) は同値である．

(1) マルコフ連鎖はエルゴード的である．

(2) マルコフ連鎖は既約で，すべての状態は非周期的である．

(3) マルコフ連鎖は既約で，非周期的な状態が存在する．

証明．(1) $\Rightarrow$ (2) を示す．エルゴード性の定義よりマルコフ連鎖が既約であることは明らかである．また，定義より，ある n_0 が存在し，任意の $i, j \in S$ に対して $p_{ij}^{(n_0)} > 0$ である．このとき，P は推移確率であるから，任意の $i \in S$ に対して $\sum_{k\in S} p_{ik} = 1$ より，$p_{ik} > 0$ を満たす k が存在する．よって，チャップマン・コルモゴロフ方程式より

$$p_{ij}^{(n_0+1)} = \sum_{k'\in S} p_{ik'}p_{k'j}^{(n_0)} \geq p_{ik}p_{kj}^{(n_0)} > 0.$$

この議論により，結局，任意の $n \geq n_0$ に対して $p_{ij}^{(n)} > 0$, $i, j \in S$ であることが分かる．これはすなわち，任意の i について $I(i)$ の最大公約数が 1 となることを意味し，従って，すべての状態は非周期的である．

(2) $\Rightarrow$ (3) は明らか．(3) $\Rightarrow$ (1) を示そう．既約であることから，任意の i, j に対して $n_1 = n_1(i, j) \in \mathbb{N}$ が存在して $p_{ij}^{(n_1)} > 0$ を満たす．また，$i_0 \in S$ が非周期的とすると，命題 1.4.3 より，ある $n_2 = n_2(i_0) \in \mathbb{N}$ が存在し，

$$p_{i_0i_0}^{(n)} > 0, \quad n \geq n_2. \tag{1.7.1}$$

今，

$$n_0 := \max_{i,j\in S}\{n_1(i, i_0) + n_2(i_0) + n_1(i_0, j)\}$$

とおく．このとき，各 i, j に対して $n := n_0 - n_1(i, i_0) - n_1(i_0, j) \geq n_2(i_0)$ を考えると，再びチャップマン・コルモゴロフ方程式と (1.7.1) より

$$p_{ij}^{(n_0)} \geq p_{ii_0}^{(n_1(i,i_0))}p_{i_0i_0}^{(n)}p_{i_0j}^{(n_1(i_0,j))} > 0.$$

n_0 は i, j に依存しないので，これにより P はエルゴード的であることが分かる． □

次の結果はマルコフ連鎖のエルゴード定理と呼ばれている．

定理 **1.7.2.** マルコフ連鎖がエルゴード的であるとき，定常分布 $\pi = \{\pi_i\}_{i\in S}$ が一意に存在し，次を満たす: ある定数 $C > 0$ と $\gamma \in (0, 1)$, $n_0 \in \mathbb{N}$ が存在し，

$$|p_{ij}^{(n)} - \pi_j| \leq C\gamma^n, \quad i, j \in S, \ \ n \geq n_0. \tag{1.7.2}$$

従って，このとき特に，

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j, \quad i, j \in S$$

が成り立つ．

- 定理 1.7.2 より，マルコフ連鎖がエルゴード的ならば，定常分布は推移確率行列のべき乗により近似計算できる．すなわち，$i \in S$ を任意に選んで固定し，δ_i を第 i 成分が 1 で残りの成分が 0 となるような $\mathbb{R}^{|S|}$ の行ベクトルとすると，十分大きな n に対し，

$$\pi \approx \delta_i P^n.$$

定理 1.7.2 の証明のための準備をしよう．まず，S 上の確率測度全体を $\mathcal{P}$ とおく．このとき，$\mathcal{P}$ の確率測度は $\mathbb{R}^{|S|}$ の元とみなせる．すなわち，

$$\mathcal{P} = \left\{ \{\nu_i\}_{i\in S} : \sum_{i\in S} \nu_i = 1,\ \ \nu_i \geq 0,\ \ i \in S \right\} \subset \mathbb{R}^{|S|}$$

である．さらに，

$$d(\nu, \mu) := \frac{1}{2} \sum_{i\in S} |\nu_i - \mu_i|, \quad \nu = \{\nu_i\}_{i\in S},\ \ \mu = \{\mu_i\}_{i\in S}$$

は $\mathbb{R}^{|S|}$ の距離を定義するので，結局，$(\mathcal{P}, d)$ は完備距離空間である．

補題 **1.7.3.** 以下が成り立つ．

(1) 任意の $\nu \in \mathcal{P}$ と任意の $m \in \mathbb{N}$ に対して $\nu P^m \in \mathcal{P}$ である．

(2) P がエルゴード的のとき，ある $n_0 \in \mathbb{N}$ と $\lambda \in (0, 1)$ が存在して，

$$d(\nu P^{n_0}, \mu P^{n_0}) \leq \lambda d(\nu, \mu), \quad \nu, \mu \in \mathcal{P}.$$

証明. (1). まずは（今更だが），P^m が推移確率であることを示そう．$m = 1$ に対しては明らかで，$m \geq 2$ に対して P^{m-1} が推移確率と仮定すると，チャップマン・コルモゴロフ方程式より，

$$\sum_{j\in S} p_{ij}^{(m)} = \sum_{j\in S}\sum_{k\in S} p_{ik} p_{kj}^{(m-1)} = \sum_{k\in S} p_{ik} \sum_{j\in S} p_{kj}^{(m-1)} = \sum_{k} p_{ik} = 1.$$

よって，帰納法より任意の m について P^m は推移確率である．このことより，$\nu \in \mathcal{P}$ に対して，

$$\sum_{j\in S} (\nu P^m)_j = \sum_{j\in S}\sum_{i\in S} \nu_i p_{ij}^{(m)} = \sum_{i\in S} \nu_i \sum_{j\in S} p_{ij}^{(m)} = \sum_{i\in S} \nu_i = 1$$

となり，$\nu P^m \in \mathcal{P}$ が従う．

(2). $\nu, \mu \in \mathcal{P}$ とする．エルゴード性より，ある $n_0 \in \mathbb{N}$ が存在して $p_{ij}^{(n_0)} > 0$, $i, j \in S$ である．S は有限集合だったから，$\lambda := 1 - \min_{i,j\in S} p_{ij}^{(n_0)} \in (0, 1)$ となる．

今，

$$\sum_{j\in S} \left| \sum_{i\in S} (\nu_i - \mu_i) p_{ij}^{(n_0)} \right| = \sum_{j\in S} \left(\sum_{i\in S} (\nu_i - \mu_i) p_{ij}^{(n_0)} \right)_+ + \sum_{j\in S} \left(\sum_{i\in S} (\mu_i - \nu_i) p_{ij}^{(n_0)} \right)_+$$

である．ところが，$\nu P^{n_0}, \mu P^{n_0} \in \mathcal{P}$ より，

$$0 = \sum_{j\in S}\sum_{i\in S} (\nu_i - \mu_i) p_{ij}^{(n_0)} = \sum_{j\in S} \left(\sum_{i\in S} (\nu_i - \mu_i) p_{ij}^{(n_0)} \right)_+ - \sum_{j\in S} \left(\sum_{i\in S} (\mu_i - \nu_i) p_{ij}^{(n_0)} \right)_+ . \tag{1.7.3}$$

よって，

$$d(\nu P^{n_0}, \mu P^{n_0}) = \sum_{j\in S} \left(\sum_{i\in S} (\nu_i - \mu_i) p_{ij}^{(n_0)} \right)_+ . \tag{1.7.4}$$

もし任意の $j \in S$ について $\sum_i (\nu_i - \mu_i) p_{ij}^{(n_0)} > 0$ ならば,

$$\sum_{j \in S} \sum_{i \in S} (\nu_i - \mu_i) p_{ij}^{(n_0)} = \sum_{j \in S} \left(\sum_{i \in S} (\nu_i - \mu_i) p_{ij}^{(n_0)} \right)_+ > 0$$

となり, (1.7.3) に矛盾する．故に, $j_0 \in S$ が存在して, $\left(\sum_i (\nu_i - \mu_i) p_{ij_0}^{(n_0)} \right)_+ = 0$. よって

$$\begin{aligned} d(\nu P^{n_0}, \mu P^{n_0}) &= \sum_{j \neq j_0} \left(\sum_{i \in S} (\nu_i - \mu_i) p_{ij}^{(n_0)} \right)_+ \leq \sum_{j \neq j_0} \sum_i (\nu_i - \mu_i)_+ p_{ij}^{(n_0)} \\ &= \sum_i (\nu_i - \mu_i)_+ \sum_{j \neq j_0} p_{ij}^{(n_0)} = \sum_i (\nu_i - \mu_i)_+ (1 - p_{ij_0}^{(n_0)}) \\ &\leq \lambda \sum_i (\nu_i - \mu_i)_+ = \lambda d(\nu, \mu). \end{aligned}$$

ここで，最後の等号は (1.7.4) と同様にして得られる． □

定理 1.7.2 の証明. $\nu_0 \in \mathcal{P}$ を $(\nu_0)_i = 1$, $(\nu_0)_j = 0$, $j \neq i$ により定義し，$\nu_n = \nu_0 P^n$ により ν_n を定義する．このとき，補題 1.7.3(1) より $\nu_n \in \mathcal{P}$ であり，$(\nu_n)_j = p_{ij}^{(n)}$ と表される．仮定と補題 1.7.3(2) より，$n_0 \in \mathbb{N}$ と $\lambda \in (0,1)$ が存在し，

$$\begin{aligned} d(\nu_n, \nu_{n+k}) &= d\left((\nu_0 P^{n-n_0}) P^{n_0}, (\nu_0 P^{n+k-n_0}) P^{n_0} \right) \leq \lambda d\left(\nu_0 P^{n-n_0}, \nu_0 P^{n+k-n_0} \right) \\ &\leq \lambda^2 d\left(\nu_0 P^{n-2n_0}, \nu_0 P^{n+k-2n_0} \right) \leq \cdots \leq \lambda^m d\left(\nu_0 P^{n-mn_0}, \nu_0 P^{n+k-mn_0} \right). \end{aligned}$$

ここで，m は n/n_0 を越えない最大の整数である．他方，任意の $\nu, \mu \in \mathcal{P}$ に対して $d(\nu, \mu) \leq (1/2) \sum_i (\nu_i + \mu_i) = 1$ であるから，結局，$n \geq n_0$ と $k \geq 0$ に対して，

$$d(\nu_n, \nu_{n+k}) \leq \lambda^m \tag{1.7.5}$$

となる．$n \to \infty$ のとき $m \to \infty$ であるから，(1.7.5) は $\{\nu_n\}_{n=1}^\infty$ が $(\mathcal{P}, d)$ のコーシー列であることを意味する．従って完備性より $\{\nu_n\}$ の極限 π が存在する．d の定義より，この収束は各状態で成立する．すなわち，$\lim_{n \to \infty} (\nu_n)_j = \pi_j$, $j \in S$. よって各状態の極限を考えることにより，

$$\pi P = \lim_{n \to \infty} \nu_n P = \lim_{n \to \infty} \nu_{n+1} = \pi.$$

すなわち，π は不変分布である．もし他の不変分布 $\tilde{\pi}$ が存在すれば,

$$d(\pi, \tilde{\pi}) = d\left(\pi P^{n_0}, \tilde{\pi} P^{n_0} \right) \leq \lambda d(\pi, \tilde{\pi})$$

より $\pi = \tilde{\pi}$ が従う．よって不変分布は一意である．

さらに，(1.7.5) で $k \to \infty$ とし $m > -1 + n/n_0$ に注意すると

$$d(\nu_n, \pi) \leq \lambda^{-1} (\lambda^{1/n_0})^n.$$

従って，$C = 2\lambda^{-1}$, $\gamma = \lambda^{1/n_0}$ とすることで (1.7.3) を得る． □

定常分布を計算する別の方法を示そう．以下，$I_{|S|}$ を $\mathbb{R}^{|S|\times|S|}$ の単位行列，$\mathbf{1}_{|S|}$ を成分がすべて 1 の $|S|\times|S|$ 行列とする．

定理 1.7.4. P がエルゴード的のとき，定常分布 π は

$$\pi = (1,\dots,1)(I_{|S|} - P + \mathbf{1}_{|S|})^{-1}$$

により与えられる．

証明. 定常分布の定義より，$\pi(I_{|S|} - P) = 0$, $\pi\mathbf{1}_{|S|} = (1,\dots,1)$ である．よって，

$$\pi(I_{|S|} - P + \mathbf{1}_{|S|}) = (1,\dots,1).$$

従って，後は $I_{|S|} - P + \mathbf{1}_{|S|}$ が正則であることを示せば定理の結論が得られる．

$x \in \mathbb{R}^{|S|}$ が $(I_{|S|} - P + \mathbf{1}_{|S|})x = 0$ を満たすとする．このとき，

$$0 = \pi(I_{|S|} - P + \mathbf{1}_{|S|})x = \pi\mathbf{1}_{|S|}x = (1,\dots,1)x$$

であり，従って，$\mathbf{1}_{|S|}x = 0$. このことと x の満たす条件より $(I_{|S|} - P)x = 0$. すなわち，$x = Px$. よって任意の $k \in \mathbb{N}$ に対して $x = P^k x$ であり，ゆえに $(1/n)\sum_{k=1}^{n} P^k x = x$. ここで定理 1.7.2 を用いると

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P^k = \Pi.$$

ただし，Π の (i,j) 成分は π_j により与えられる（すなわち，第 j 列は全て π_j）．これより，$x = \Pi x$ を得る．これを成分で表すと $x_i = \sum_{j\in S}\pi_j x_j$, $i \in S$, となり，左辺が i に依存しないことから，適当な定数 c を用いて，$x_i = c$, $i \in S$, と書ける．ところが，

$$0 = (1,\dots,1)x = c|S|$$

であるから，$c = 0$. すなわち $x = 0$ を得る．これより行列 $(I_{|S|} - P + \mathbf{1}_{|S|})$ が正則であることが従う． □

例 **1.7.5** (PageRank アルゴリズム)**.** 例 1.1.7 を思い出そう．検索サイト Google で使われている PageRank アルゴリズムでは，先の例で考えたランダムに Web ページを閲覧するユーザ（これをランダムサーファーと呼ぶ）が滞在する割合によって Web ページの重要度を評価している．より正確に言えば，ランダムサーファーが定義するマルコフ連鎖の定常分布 $\{\pi_i\}_{i\in S}$ を求め，π_i たちの値の大小によって，Web ページの集合 S のランク付けを行っている．

定常分布を求める問題についてもう少し考えてみよう．ここでは，より現実的に，リンク先のないページも存在し得る場合を考える．このとき，例 1.1.7 と同様に $P = \{p_{i,j}\}_{i,j\in S}$ を定義すると，成分がすべて 0 となる行も存在し得るため，P は一般には推移確率にはならない．この問題を解決するため，ランダムサーファーの挙動に以下のような仮定をおく:

- 現在滞在している Web ページにリンク先が存在しない場合，ランダムに選んだ Web ページに遷移する．

- 現在滞在している Web ページにリンク先が存在する場合，次の挙動 (1) または (2) をランダムに選択する．
  (1) 現在滞在している Web ページ内のリンクをランダムに選択し遷移する．
  (2) 現在滞在している Web ページ内のリンクと関係なく，ランダムに選んだページに遷移する．

Web ページ間を遷移する確率は一様で，挙動 (1) を選択する確率を $d \in (0,1)$ とすると，ランダムサーファーが Web ページ i から j へ遷移する確率 g_{ij} は

$$g_{ij} = \begin{cases} \dfrac{d}{n_i} + \dfrac{1-d}{N} & (i \text{ にリンク先が存在し，} j \in F(i) \text{ の場合}) \\ \dfrac{1-d}{N} & (i \text{ にリンク先が存在し，} j \notin F(i) \text{ の場合}) \\ \dfrac{1}{N} & (i \text{ にリンク先が存在しない場合}) \end{cases}$$

となり，$G = \{g_{ij}\}_{i,j \in S}$ は推移確率を定義する．さらに，G は正行列であるからエルゴード的であり，従って定常分布が一意に存在する．この定常分布の算出には定理 1.7.2 のべき乗法や定理 1.7.4 の連立方程式を解く方法などを用いることができる．

マルコフ連鎖がエルゴード的であるとき，独立同分布列の場合と同様に，大数の弱法則が成り立つ．

定理 **1.7.6.** マルコフ連鎖はエルゴード的であるとする．このとき，任意の $f : S \to \mathbb{R}$ に対して，確率収束の意味で，

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) = \sum_{i \in S} f(i)\pi_i.$$

証明．定理の主張は定数を加えても同値であるため，はじめから $\sum_{i\in S} f(i)\pi_i = 0$ を仮定しても一般性を失わない．定理の仮定より，定理 1.7.2 における C, γ, n_0 が存在する．このとき，命題 A.3.2(3)（チェビシェフの不等式）より，$n > n_0$ と固定された $\varepsilon > 0$ に対して，

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} f(X_k)\right| > \varepsilon\right) \le \frac{1}{\varepsilon^2 n^2}\mathbb{E}\left(\sum_{k=1}^{n} f(X_k)\right)^2$$
$$= \frac{1}{\varepsilon^2 n^2}\sum_{k=1}^{n}\mathbb{E}[f(X_k)^2] + \frac{2}{\varepsilon^2 n^2}\sum_{\substack{1\le k<m\le n \\ m-k\le n_0}}\mathbb{E}[f(X_k)f(X_m)] + \frac{2}{\varepsilon^2 n^2}\sum_{\substack{1\le k<m\le n \\ m-k> n_0}}\mathbb{E}[f(X_k)f(X_m)].$$

右辺の第 1,2,3 項をそれぞれ I_1, I_2, I_3 とおく．$\|f\|_\infty := \max_{i\in S}|f(i)|$ と表すとき，

$$I_1 \le \frac{1}{\varepsilon^2 n^2}\sum_{k=1}^{n}\|f\|_\infty^2 = \frac{\|f\|_\infty^2}{\varepsilon^2 n} \to 0, \quad n \to \infty.$$

同様に，

$$I_2 \le \frac{2\|f\|_\infty^2}{\varepsilon^2 n^2} \sum_{\substack{1\le k<m\le n\\ m-k\le n_0}} = \frac{2\|f\|_\infty^2}{\varepsilon^2 n^2} \sum_{k=1}^{n-1} \sum_{m=k+1}^{k+n_0} = \frac{2\|f\|_\infty^2 (n-1)n_0}{\varepsilon^2 n^2} \to 0, \quad n\to\infty.$$

I_3 を評価するため，$\sum_{j\in S} f(j)\pi_j = 0$ に注意すれば

$$\begin{aligned}
\mathbb{E}[f(X_k)f(X_m)] &= \sum_{i,j\in S} f(i)f(j)\mathbb{P}(X_k=i, X_m=j) \\
&= \sum_{i_0,i,j\in S} f(i)f(j)\mathbb{P}(X_0=i_0, X_k=i, X_m=j) \\
&= \sum_{i_0,i,j\in S} f(i)f(j)\mathbb{P}(X_0=i_0) p_{i_0 i}^{(k)} p_{ij}^{(m-k)} \\
&= \sum_{i_0,i,j\in S} f(i)f(j)\mathbb{P}(X_0=i_0) p_{i_0 i}^{(k)} (p_{ij}^{(m-k)} - \pi_j)
\end{aligned}$$

を得る．よって，定理 1.7.2 より，$m-k>n_0$ のとき，

$$\begin{aligned}
|\mathbb{E}[f(X_k)f(X_m)]| &\le \|f\|_\infty^2 \sum_{i_0,i,j\in S} \mathbb{P}(X_0=i_0) p_{i_0 i}^{(k)} |p_{ij}^{(m-k)} - \pi_j| \\
&\le C\|f\|_\infty^2 |S| \gamma^{m-k}.
\end{aligned}$$

他方，

$$\sum_{\substack{1\le k<m\le n\\ m-k>n_0}} \gamma^{m-k} = \sum_{k=1}^{n-1} \sum_{m=k+n_0+1}^{n} \gamma^{m-k} \le \sum_{k=1}^{n-1} \sum_{m=1}^{\infty} \gamma^m = \frac{\gamma(n-1)}{1-\gamma}$$

であるから，

$$I_3 \le \frac{2C\|f\|_\infty^2 |S| \gamma(n-1)}{\varepsilon^2 n^2 (1-\gamma)} \to 0, \quad n\to\infty.$$

ゆえに定理の主張が従う． □

- エルゴード性の仮定の下，実は大数の強法則も成立する．証明については（測度論の準備が必要だが）例えば [15] の第 11 章を見よ．

1.8 マルコフ連鎖モンテカルロ法

X を確率変数，f を X の状態集合上の関数として，$f(X)$ の期待値 $\mathbb{E}[f(X)]$ を効率的に計算する問題を考えよう．この問題はベイズ統計やファイナンスなど，多くの応用において重要である．X が密度 $\rho(x)$ を持つ場合，

$$\mathbb{E}[f(X)] = \int f(x)\rho(x)dx$$

であり，右辺の積分が厳密に計算できない場合，諸々の数値積分法に頼ることになる．X が離散分布に従う場合，

$$\mathbb{E}[f(X)] = \sum_x f(x)\mathbb{P}(X=x) \tag{1.8.1}$$

であり，右辺の計算は単純な足し算になるが，$X(\Omega) = \{X(\omega) : \omega \in \Omega\}$ の要素数が非常に多いときは，直接的な和計算は非効率的かもしれない．

モンテカルロ法（Monte Carlo method）は，確率変数の分布が連続的か離散的かに依らない期待値の近似計算法である．$\{X_n\}$ が X と同じ分布を持つ IID 列であるとき，（適当な可積分性の下で）大数の法則より，

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) = \mathbb{E}[f(X)]$$

が確率 1 で成立する．よって，n を十分大きくとり，X の従う分布から乱数 $x_1, \ldots, x_n$ を独立に発生させることで，

$$\frac{1}{n}\sum_{k=1}^{n} f(x_k) \approx \mathbb{E}[f(X)] \tag{1.8.2}$$

とみなせる．モンテカルロ法とは右辺の期待値を左辺の和で近似する方法のことである．このとき，

$$\mathbb{E}\left|\frac{1}{n}\sum_{k=1}^{n} f(X_k) - \mathbb{E}[f(X)]\right|^2 = \frac{1}{n}\mathbb{V}[f(X)]$$

より，(1.8.2) の近似誤差は平均的には $O(1/\sqrt{n})$ である．従って，モンテカルロ法は，$X(\Omega)$ の次元に依存しない期待値の近似法であり，特に X が離散的な場合，$\sqrt{n}$ が十分大きく，かつ $|X(\Omega)| > n$ が満たされる場合には，(1.8.2) の近似は有用になり得る．

モンテカルロ法を適用するためには，上述の通り，X の分布から乱数を発生させる必要がある．このためには，X の分布の形が分かっているだけでは不十分で，計算機上で擬似乱数を構成することが可能でなくてはならない．この問題について考えるため，本章では離散状態の確率変数を扱っているので，X が有限集合 S 上の確率変数で離散分布 $\{\pi_i\}_{i\in S}$ に従う場合に限定して話を進めよう．まず，π のシミュレーション法を紹介する．

例 **1.8.1.** $S = \{1, \ldots, N\}$ と仮定する．U を $[0,1]$ 上の一様分布に従う確率変数とし，

$$X = \sum_{\ell=1}^{N} \ell 1_{\{F_{\ell-1} \le U < F_\ell\}} \tag{1.8.3}$$

により確率変数 X を定義する．ここで，$F_\ell = \sum_{i=1}^{\ell} \pi_i$, $\ell = 1, \ldots, N$, $F_0 = 0$ である．このとき，$\mathbb{P}(X = \ell) = \mathbb{P}(F_{\ell-1} \le U < F_\ell) = F_\ell - F_{\ell-1} = \pi_\ell$ である．従って，$[0,1]$ 上の一様分布から独立に n 個の乱数 $u_1, \ldots, u_n$ を発生させ，x_k を (1.8.3) において U を u_k に置き換えたものとして定義することで π のシミュレーションができることになる．

例 1.8.1 の方法では，$\{x_1, \ldots, x_n\}$ を得るために，例えば，全ての u_i について最近傍探索により対応する ℓ を見つける必要がある．$\{F_\ell\}$ の計算に $O(N)$ を要するので，二分探索を使うとすれば，これには $O(N + n\log N)$ の計算量を必要とする．期待値の算出まで含めて $O(n + N + n\log N)$ の計算量であるから，理論的には，単純な足し算による期待値計算を選ばない理由は見つけにくい．

以上のような理由により，π からの乱数生成が容易な場合以外では，モンテカルロ法による期待値計算は非効率的であることが多い．前置きが長くなってしまったが，本節では，それに代わる方法として，**マルコフ連鎖モンテカルロ法**（Markov chain Monte Carlo method, MCMC 法）を紹介する．前節の定理 1.7.6 を思い出そう．これはマルコフ連鎖に対する大数の法則を記述したものであった．MCMC 法とは，π が適当なエルゴード的マルコフ連鎖 $\{X_n\}$ の定常分布とみなせるように $\{X_n\}$ を構成する方法の一群のことを指す．もしこのような $\{X_n\}$ がシミュレートできるなら，モンテカルロ法と同様に

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \approx \mathbb{E}[f(X)]$$

が成り立つ．

以下，各 $i \in S$ に対して $\pi_i > 0$ を満たすと仮定する．問題は π を定常分布として持つようなマルコフ連鎖を構成することである．推移確率 $P = \{p_{ij}\}_{i,j\in S}$ を持つ $\{X_n\}$ で**詳細釣り合い条件**（detailed balance condition）

$$\pi_i p_{ij} = \pi_j p_{ji}, \quad i, j \in S,$$

を満たすものを考える．このとき，$\pi P = \pi$ が成り立つので，π は $\{X_n\}$ の定常分布である．**メトロポリス・ヘイスティングス法**（Metropolis–Hastings method）では，

$$\begin{aligned} p_{ij} &= q_{ij}\alpha_{ij}, \quad i \neq j, \\ p_{ii} &= 1 - \sum_{j\neq i} p_{ij} \end{aligned} \tag{1.8.4}$$

を満たす P を構成する．ここで，$Q = \{q_{ij}\}$ は別の推移確率で，α_{ij} は，$q_{ij}, q_{ji} > 0$ なる $i, j \in S$ に対しては

$$\alpha_{ij} = \frac{s_{ij}}{1 + \dfrac{\pi_i q_{ij}}{\pi_j q_{ji}}}$$

により定義される．$q_{ji} = 0$, $q_{ij} > 0$ のときは $\alpha_{ij} = 0$, $q_{ij} = q_{ji} = 0$ のときは $\alpha_{ij} = 1$ とおく．$\{s_{ij}\}$ は対称行列で $\alpha_{ij} \in [0, 1]$ を満たすように選ばれる．(1.8.4) を満たす推移確率 P が詳細釣り合い条件を満たすことは容易に確かめることができる．

(1.8.4) を満たす推移確率 P を持つマルコフ連鎖 $\{X_n\}$ は例えば次のようにして構成される．

(i) $i_0 \in S$ を任意に取り，$X_0 = i_0$ とおく．

(ii) 各 $n \geq 0$ と $i \in S$ に対して，$\mathbb{P}(Y_{n,i} = j \,|\, X_n = i) = q_{ij}$, $j \in S$, を満たす確率変数 $Y_{n,i}$ をとる．

(iii) 各 $n \geq 0$ と $i, j \in S$ に対して，$\{0, 1\}$-値確率変数 $Z_{n,i,j}$ で，

$$\mathbb{P}(Z_{n,i,j} = 1 \,|\, Y_{n,i} = j, X_n = i) = \alpha_{ij} = 1 - \mathbb{P}(Z_{n,i,j} = 0 \,|\, X_n = i, Y_{n,i} = j)$$

を満たすものを選ぶ．

(iv) $\{X_n\}$ を

$$X_{n+1} = \tilde{Y}_n 1_{\{\tilde{Z}_n=1\}} + X_n 1_{\{\tilde{Z}_n=0\}}, \quad n \geq 0,$$

により定める．ここで，$\tilde{Y}_n = Y_{n,X_n}$, $\tilde{Z}_n = Z_{n,\tilde{Y}_n}$.

実際，$i \neq j$ に対して，

$$\begin{aligned}
\mathbb{P}(X_{n+1} = j \,|\, X_n = i) &= \mathbb{P}(Y_{n,i} = j, Z_{n,i,j} = 1 \,|\, X_n = i) \\
&= \mathbb{P}(Y_{n,i} = j \,|\, X_n = i)\mathbb{P}(Z_{n,i,j} = 1 \,|\, X_n = i, Y_{n,i} = j) = q_{ij}\alpha_{ij} \\
&= p_{ij}
\end{aligned}$$

であり，さらに，

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = i \,|\, X_n = i) \\
&= \mathbb{P}(Y_{n,i} = i \,|\, X_n = i) + \mathbb{P}(Y_{n,i} \neq i, Z_{n,i} = 0 \,|\, X_n = i) \\
&= \mathbb{P}(Y_{n,i} = j \,|\, X_n = i) + \sum_{j \neq i} \mathbb{P}(Y_{n,i} = j \,|\, X_n = i)\mathbb{P}(Z_{n,i,j} = 0 \,|\, X_n = i, Y_{n,i} = j) \\
&= q_{ii} + \sum_{i \neq j} q_{ij}(1 - \alpha_{ij}) = 1 - \sum_{i \neq j} p_{ij} = p_{ii}
\end{aligned}$$

である．従って，$\{X_n\}$ のシミュレーションは以下の手順により実行できることになる．

(i) $i_0 \in S$ を任意にとり，$X_0 = i_0$ と定める．

(ii) 各 $n \geq 0$ に対して，$X_n = i \in S$ のとき，確率分布 $\{q_{ij}\}_{j \in S}$ の（擬似）乱数を生成する．

(ii-a) この乱数の値が $j \in S$ のとき，確率 α_{ij} で $X_{n+1} = j$ とし，確率 $1 - \alpha_{ij}$ で $X_{n+1} = i$ とする．

すなわち，上の手順において，$X_n = i$ で分布 $\{q_{ij}\}_{j \in S}$ から発生させた乱数が j のとき，確率 α_{ij} で j を次のステップの値として**採択**し，確率 $1 - \alpha_{ij}$ で**棄却**することになる．この意味で Q は**提案分布**と呼ばれる．Q の選び方には任意性があるが，乱数生成が容易なものであり，かつ後で述べるようにエルゴード的である必要がある．具体的には，ある種のランダムウォークが使われることが多い．例えば $S = \{1, \ldots, N\}$ の場合には，$q_{11} > 0$, $q_{12} = 1 - q_{11}$, $q_{i,i+1} > 0$, $q_{i,i-1} = 1 - q_{i,i+1}$, $i = 2, \ldots, N-1$, $q_{NN} = 1 - q_{N,N-1}$, を満たすものである．

例 **1.8.2.** 行列 Q が対称で，

$$s_{ij} = \begin{cases} 1 + \dfrac{\pi_i}{\pi_j} & \dfrac{\pi_j}{\pi_i} \geq 1, \\[2ex] 1 + \dfrac{\pi_j}{\pi_i} & \dfrac{\pi_j}{\pi_i} < 1, \end{cases}$$

とするとき，

$$\alpha_{ij} = \min\left(1, \frac{\pi_i}{\pi_j}\right)$$

となる．このような構成法は**メトロポリス法**と呼ばれる．

例 **1.8.3.** 行列 Q が対称で $s_{ij}=1$ とすると，

$$\alpha_{ij}=\frac{\pi_j}{\pi_i+\pi_j}$$

となる．このような構成法は**バーカー法**（Barker's method）と呼ばれる．

上のような手続きにより構成した $\{X_n\}$ の分布が確かに π に収束しているかどうか，そして収束している場合はその収束速度，を知ることはモンテカルロ法の適用において重要なことである．これについては定理 1.7.2 が教えてくれる．

命題 **1.8.4.** 推移確率 Q がエルゴード的であると仮定する．さらに，各 $i,j\in S$ に対して $\alpha_{ij}>0$ を仮定する．P を (1.8.4) を満たす S 上の推移確率とする．このとき，P を推移確率とする S 上のマルコフ連鎖 $\{X_n\}$ はエルゴード的であり，以下が成り立つ: ある定数 $C>0$ と $\gamma\in(0,1)$, $n_0\in\mathbb{N}$ が存在し，$\mathbb{P}(X_0=i)>0$ なる任意の $i\in S$ と任意の $j\in S$ に対し，

$$|\mathbb{P}(X_n=j\,|\,X_0=i)-\pi_j|\le C\gamma^n,\ \ n\ge n_0.$$

証明. 仮定より，ある $n_0\in\mathbb{N}$ が存在して，任意の $i,j\in S$ について $q_{ij}^{(n)}>0$. さらに，$\varepsilon:=\min_{i,j\in S}\alpha_{ij}>0$. これより

$$p_{ij}^{(2)}=\sum_{k\in S}p_{ik}p_{kj}=\sum_{k\in S}\alpha_{ik}q_{ik}\alpha_{kj}q_{kj}\ge\varepsilon^2q_{ij}^{(2)}.$$

よって，帰納法を使うことで，任意の $n\ge1$ に対して $p_{ij}^{(n)}\ge\varepsilon^nq_{ij}^{(n)}$ が成り立つことが示せる．$n=n_0$ として P のエルゴード性が従う．残りの主張は定理 1.7.2 より従う． □

実用上は命題 1.8.4 における定数 C, γ, n_0 を具体的に知りたいところである．定理 1.7.2 の証明中では，n_0 はエルゴード性の条件に現れるステップ数，すなわち，各 $i,j\in S$ に対して $p_{ij}^{(n_0)}>0$ をなるような n_0 である．命題 1.8.4 の証明から，n_0 としては Q のエルゴード性に関するステップ数を採用すればよい．さらに，$\lambda=1-\min_{i,j}p_{ij}^{(n_0)}$ として $C=2/\lambda$, $\gamma=\lambda^{1/n_0}$ である．しかしこれは必ずしも精密な評価ではなく，実際の収束の速さを知るには保守的過ぎるかもしれない．P が詳細釣り合い条件を満たす場合は，定理 1.7.2 より精密な評価が知られている．これについては例えば [2] の第 6 章で論じられている．

MCMC 法の別の例を示そう．ここでは $S=\{j=(j_1,\dots,j_d):j_i=1,\dots,N,\ i=1,\dots,d\}$ と表される場合を考える．まず，X が π に従う確率変数としていたことを思い出そう．今の場合では，X は d 次元の確率変数で $X=(X(1),\dots,X(d))$ と表される．$i,j\in S$, $k=1,\dots,d$ に対し，

$$\pi(j_k\,|\,i_{-k})=\mathbb{P}(X(k)=j_k\,|\,X(\ell)=i_\ell,\,\ell\neq k)$$

とおく．各 $k=1,\dots,d$ に対し $P(k)=\{p_{ij}(k)\}$ を次のように定義する．$i_m=j_m$, $m\neq k$, を満たすような $i,j\in S$ に対して，

$$p_{ij}(k)=\pi(j_k\,|\,i_{-k})$$

とし，それ以外の $i, j \in S$ に対しては $p_{ij}(k) = 0$ と定める．このとき，各 $k = 1, \ldots, d$, $i, j \in S$ に対して，

$$\begin{aligned}
\sum_{i \in S} \pi_i p_{ij}(k) &= \sum_{i \in S} \mathbb{P}(X = i) \mathbb{P}(X(k) = j_k \,|\, X(\ell) = i_\ell,\, \ell \neq k) \\
&= \sum_{i \in S} \mathbb{P}(X(k) = i_k, X(\ell) = i_\ell,\, \ell \neq k) \mathbb{P}(X(k) = j_k \,|\, X(\ell) = i_\ell,\, \ell \neq k) \\
&= \sum_{i \in S} \mathbb{P}(X = j) \mathbb{P}(X(k) = i_k \,|\, X(\ell) = j_\ell,\, \ell \neq k) \\
&= \mathbb{P}(X = j) \frac{\sum_{i \in S} \mathbb{P}(X(k) = i_k, X(\ell) = j_\ell,\, \ell \neq k)}{\mathbb{P}(X(\ell) = j_\ell,\, \ell \neq k)} \\
&= \pi_j.
\end{aligned}$$

ゆえに，推移確率

$$P := P(1) \cdots P(d) \tag{1.8.5}$$

は $\pi P = \pi$ を満たす．すなわち，P を推移確率とするようなマルコフ連鎖 $\{X_n\}$ は π を定常分布に持つ．この $\{X_n\}$ についての MCMC 法は**ギブス・サンプリング**（Gibbs sampling）と呼ばれている．

ギブス・サンプリングは例えば以下の手順により実行できる．

(i) $i_0 \in S$ を任意にとり，$X_0 = i_0$ とおく．

(ii) 各 $n \geq 0$ と $i \in S$ に対して，$X_n = i$ のとき，まず $j = i$ とおく．

  (ii-a) k を 1 から d まで順に動かし次を実行する: $\{1, \ldots, N\}$ 上の確率分布 $\pi(\cdot \,|\, j_{-k})$ からランダムに ι を発生させる．j の第 k 成分を ι に更新する．

  (ii-b) $X_{n+1} = j$ と定める．

ギブス・サンプリングはすなわち，多次元分布の乱数生成を比較的計算コストの低い 1 次元の場合に帰着させる方法である．特定の分布のクラス，例えばギブス分布の場合には，これは上手く機能する．$\pi(\cdot \,|\, j_{-k})$ のシミュレーションにメトロポリス・ヘイスティングス法を適用することもある．命題 1.8.4 と同様に，以下が得られる．

命題 **1.8.5.** P を (1.8.5) により定義される推移確率とする．P を推移確率とする S 上のマルコフ連鎖 $\{X_n\}$ はエルゴード的であり，次が成り立つ: ある定数 $C > 0$ と $\gamma \in (0, 1)$, $n_0 \in \mathbb{N}$ が存在し，$\mathbb{P}(X_0 = i) > 0$ なる任意の $i \in S$ と任意の $j \in S$ に対し，

$$|\mathbb{P}(X_n = j \,|\, X_0 = i) - \pi_j| \leq C\gamma^n, \;\; n \geq n_0.$$

証明. 任意の $i \in S$ に対して $\pi_i > 0$ と仮定しているので，$i_\ell = j_\ell$, $\ell \neq k$ に対しては $p_{ij}(k) > 0$ である．このことと $P = \{p_{ij}\}$ の構成の仕方より，P は既約であり，全ての i に対して $p_{ii} > 0$ である．よって全ての状態が非周期的である．従って，命題 1.7.1 より $\{X_n\}$ はエルゴード的である．あとは定理 1.7.2 を適用すればよい． □

1.9 マルコフ連鎖の制御

例 1.1.5 を思い出そう．この例では，一日の終わりに医薬品 A が 1 個以下になっていたら合計 6 個になるよう発注するのだった．これを仮に，一日の終わりに 2 個以下のとき合計 6 個となるよう発注することにすると，n 日の閉店時にある A の数 X_n は

$$X_{n+1} = (X_n + \alpha_n(X_n) - D_{n+1})_+$$

により表される．ここで，$\alpha_n(i)$ は n 日の終わりの在庫数 i に応じて決定される発注方針であり，

$$\alpha_n(i) = \begin{cases} 6 - i & (i \leq 2), \\ 0 & (i > 2) \end{cases}$$

と定義される．このように書き表してみると，発注方針ごとに異なるマルコフ連鎖が得られることが分かるだろう．様々な発注方針の中で，どれが平均的に効率的な在庫管理を実現するか，という問題に対する基本的な手法を紹介するのが本節の目的である．

ここでは，

- A, D: 高々可算集合.
- 任意の $n = 0, \ldots, N-1$, $a \in A_n(i)$ と $w \in D$ に対して $f_n(i, a, w) \in S$.
- $\{W_n\}_{n=1}^N$: D 値独立確率変数列.

という条件を満たす集合 A, D, 集合列 $\{A_n(i)\}$, 関数列 $\{f_n\}$, 確率変数列 $\{W_n\}$ を考え，さらに

$$\mathcal{A} = \left\{ \alpha = \{\alpha_n\}_{n=0}^{N-1} \;\middle|\; \alpha_n : S \to A,\ \alpha_n(i) \in A_n(i),\ i \in S,\ n = 0, \ldots, N-1 \right\}$$

とおく．各 $\alpha \in \mathcal{A}$ に対して

$$\begin{cases} X_{n+1} = f_n(X_n, \alpha_n(X_n), W_{n+1}), \quad n = 0, \ldots, N-1, \\ X_0 = i \end{cases}$$

により確率過程 $X_n \equiv X_n^{i,\alpha}$, $n = 0, \ldots, N$ を定義すると，これは時刻 0 で状態 i から出発する S 値マルコフ連鎖となる．一般に，$\alpha \in \mathcal{A}$ を制御則（control policy），$\{X_n\}$ を**被制御マルコフ連鎖**（controlled Markov chain）と呼ぶ.

被制御過程 $\{X_n\}$ の各時刻のパフォーマンスを関数 g_n で評価し，その時刻 N までの平均量を最小化する問題

$$V^*(i) := \inf_{\alpha \in \mathcal{A}} \mathbb{E}\left[g_N(X_N^{i,\alpha}) + \sum_{n=0}^{N-1} g_n(X_n^{i,\alpha}, \alpha_n(X_n^{i,\alpha}), W_{n+1}) \right] \tag{1.9.1}$$

を考えよう．ここで，g_n たちは全て有界な実数値関数とする．(1.9.1) を**確率制御問題**（stochastic control problem）という．

問題 (1.9.1) は関数列 $\{\alpha_n(\cdot)\}_{n=0}^{N-1}$ の最適化問題であるが，確率現象の時間発展があるため通常の多変数の最適化問題とは多少異なる．このような問題を解くための方法としては**動的計画法**（dynamic programming）が基本的である．やや天下り的ではあるが，以下で定義される関数

$$\begin{cases} V_N(i) = g_N(i), \\ V_n(i) = \inf_{a \in A_n(i)} \mathbb{E}\left[V_{n+1}(f_n(i,a,W_{n+1})) + g_n(i,a,W_{n+1})\right], \quad i \in S, \ \ n = 0, \ldots, N-1. \end{cases}$$

を考えよう．これは終端時刻から時間に関して後退的に定義される漸化式である．

このとき，次を得る:

定理 1.9.1 (動的計画法の原理)**.** 任意の $i \in S$ に対して

$$V_0(i) = V^*(i)$$

が成り立つ．さらに，各 $V_n(i)$ の定義式における inf を達成する $a^*_{n,i} \in A_n(i)$ が存在するとき，$\alpha^*_n(i) = a^*_{n,i}$ により $\alpha^*_n : S \to A$, $n = 0, \ldots, N-1$ を定義すると，これが問題 (1.9.1) の最適解となる．

- $V_n(i)$, $(n,i) \in \{0, \ldots, N\} \times S$ を**値関数**（value function）という．

証明の前に例を挙げておこう．

例 1.9.2. 例 1.1.5 の枠組みで，定理 1.9.1 を使って最適な在庫管理方針を求めてみよう．この場合，$S = A = \{0,1,2,3,4,5,6\}$, $A_n(i) = \{a \in A : i + a \leq 6\}$, $D = \{0,1,2,3,4\}$ であり，不確実性の素になる確率変数列 $\{W_n\}$ は客の需要を表す確率変数列 $\{D_n\}$ により与えられる．分布は

$$\mathbb{P}(D_n = 0) = 0.2, \quad \mathbb{P}(D_n = 1) = 0.3, \quad \mathbb{P}(D_n = 2) = 0.2,$$
$$\mathbb{P}(D_n = 3) = 0.2, \quad \mathbb{P}(D_n = 4) = 0.1$$

である．$N = 3$ とし，医薬品 A の最初の在庫は無いものとしよう．すなわち $X_0 = 0$ とする．ここでは，在庫のコストは保有と発注にかかるコストの 2 つにより与えられると考える．時刻 3 以降の在庫の状態については考えないので，この時点のコストはゼロとする．それまでの時刻においては，在庫が足りなくて購入できない客が存在する時にはその個数分の損失が，発注には 1 個あたり 0.5 の費用がそれぞれ発生すると仮定する．よって，

$$g_3(i) = 0, \quad g_n(i,a,w) = (-(i+a-w))_+ + a/2, \quad i \in S, \ \ a \in A, \ \ w \in D, \ \ n = 0,1,2$$

とし，これより $V_3(i) = 0$, $i \in S$ である．このとき，V_2 の定義は

$$V_2(i) = \min_{a \in A_2(i)} \sum_{w=0}^{4} \{(-(i+a-w))_+ a/2\} \, \mathbb{P}(D_3 = w).$$

であり，各状態ごとに計算すると

$$
\begin{aligned}
V_2(0) &= \min_{a\le 6}[(-(a-1))_+ \times 0.3 + (-(a-2))_+ \times 0.2 + (-(a-3))_+ \times 0.2 \\
&\qquad + (-(a-4))_+ \times 0.1 + 0.5a] = 1.4, \quad a_2^*(0) = 1 \text{ or } 2, \\
V_2(1) &= \min_{a\le 5}[(-(a-1))_+ \times 0.2 + (-(a-2))_+ \times 0.2 + (-(a-3))_+ \times 0.1 \\
&\qquad + 0.5a] = 0.9, \quad a_2^*(1) = 0 \text{ or } 1, \\
V_2(2) &= \min_{a\le 4}[(-(a-1))_+ \times 0.2 + (-(a-2))_+ \times 0.1 + 0.5a] = 0.4, \quad a_2^*(2) = 0, \\
V_2(3) &= \min_{a\le 3}[(-(a-1))_+ \times 0.1 + 0.5a] = 0.1, \quad a_2^*(2) = 0, \\
V_2(4) &= V_2(5) = V_2(6) = 0, \quad a_2^*(4) = a_2^*(5) = a_2^*(6) = 0.
\end{aligned}
$$

となる．これを用いると $n=1$ の値関数と最適な a は

$$
\begin{aligned}
V_1(0) = \min_{a\le 6} \big[& V_2(a) \times 0.2 + (V_2((a-1)_+) + (-(a-1))_+) \times 0.3 \\
& + (V_2((a-2)_+) + (-(a-2))_+) \times 0.2 + (V_2((a-3)_+) + (-(a-3))_+) \times 0.2 \\
& + (V_2((a-4)_+) + (-(a-4))_+) \times 0.1 + 0.5a \big] = 2.34, \quad a_1^*(0) = 3, \\
V_1(1) = \min_{a\le 5} \big[& V_2(a+1) \times 0.2 + V_2(a) \times 0.3 + (V_2((a-1)_+) + (-(a-1))_+) \times 0.2 \\
& + (V_2((a-2)_+) + (-(a-2))_+) \times 0.2 + (V_2((a-3)_+) + (-(a-3))_+) \times 0.1 \\
& + 0.5a \big] = 1.84, \quad a_1^*(1) = 2, \\
V_1(2) = \min_{a\le 4} \big[& V_2(a+2) \times 0.2 + V_2(a+1) \times 0.3 + V_2(a) \times 0.2 \\
& + (V_2((a-1)_+) + (-(a-1))_+) \times 0.2 + (V_2((a-2)_+) + (-(a-2))_+) \times 0.1 \\
& + 0.5a \big] = 1.34, \quad a_1^*(2) = 1, \\
V_1(3) = \min_{a\le 3} \big[& V_2(a+3) \times 0.2 + V_2(a+2) \times 0.3 + V_2(a+1) \times 0.2 + V_2(a) \times 0.2 \\
& + (V_2((a-1)_+) + (-(a-1))_+) \times 0.1 + 0.5a \big] = 0.84, \quad a_1^*(3) = 0, \\
V_1(4) = 0.43, \quad & V_1(5) = 0.19, \quad V_1(6) = 0.06, \quad a_1^*(4) = a_1^*(5) = a_1^*(6) = 0.
\end{aligned}
$$

となり，これより

$$
\begin{aligned}
V_0(0) = \min_{a\le 6} \big[& V_1(a) \times 0.2 + (V_1((a-1)_+) + (-(a-1))_+) \times 0.3 \\
& + (V_1((a-2)_+) + (-(a-2))_+) \times 0.2 + (V_1((a-3)_+) + (-(a-3))_+) \times 0.2 \\
& + (V_1((a-4)_+) + (-(a-4))_+) \times 0.1 + 0.5a \big] = 3.208, \quad a_0^*(0) = 4
\end{aligned}
$$

となることが確かめられる．すなわち，最初にまず 4 個発注し，時刻 1 の終わりには，次の日の開店時に在庫が 3 個になるように発注する．そして，時刻 2 の終わりには，次の日の開店時に在庫が 1 個か 2 個になるように発注するのが最適ということになる．

さて，定理 1.9.1 の証明のための準備をしよう．まず，途中の時刻から考えた制御則の集合

$$
\mathcal{A}_n = \left\{ \alpha = \{\alpha_k\}_{k=n}^{N-1} \;\middle|\; \alpha_k : S \to A,\ \alpha_k(i) \in A_k(i),\ i \in S,\ k = n, \ldots, N-1 \right\}
$$

を導入する．$\alpha \in \mathcal{A}_n$ に対し，時刻 n で i から出発する被制御マルコフ連鎖 $\{X_k^{n,i,\alpha}\}_{k=n}^N$ を

$$\begin{cases} X_k^{n,i,\alpha} = f_k(X_k^{n,i,\alpha}, \alpha_k(X_k^{n,i,\alpha}), W_{k+1}), \quad k = n+1, \dots, N-1, \\ X_n^{n,i,\alpha} = i. \end{cases}$$

により定義する．このとき，$\alpha \in \mathcal{A}$ に対して $X_n^{0,i,\alpha} = X_n^{i,\alpha}$ である．

補題 **1.9.3.** 任意の $n = 1, \dots, N-1$, $\alpha \in \mathcal{A}_n$, $k = n, \dots, N$ に対して，

$$X_k^{n,j,\alpha}\big|_{j=X_n^{n-1,i,\alpha}} = X_k^{n-1,i,\alpha}, \quad i \in S. \tag{1.9.2}$$

証明. k に関する帰納法で証明する．$X_n^{n,j,\alpha} = j$ であるから，$X_n^{n,j,\alpha}\big|_{j=X_n^{n-1,i,\alpha}} = X_n^{n-1,i,\alpha}$ となり，$k = n$ のとき (1.9.2) は成立する．次に一般の k について (1.9.2) が成り立つとする．このとき，$X_{k+1}^{n,j,\alpha} = f_k(X_k^{n,j,\alpha}, \alpha_k(X_k^{n,j,\alpha}), W_{k+1})$ と帰納法の仮定より

$$\begin{aligned} X_{k+1}^{n,j,\alpha}\big|_{j=X_n^{n-1,i,\alpha}} &= f_k(X_k^{n,j,\alpha}, \alpha_k(X_k^{n,j,\alpha}), W_{k+1})\big|_{j=X_n^{n-1,i,\alpha}} \\ &= f_k(X_k^{n-1,i,\alpha}, \alpha_k(X_k^{n-1,i,\alpha}), W_{k+1}) \\ &= X_{k+1}^{n-1,i,\alpha}. \end{aligned}$$

よって $k+1$ に対しても (1.9.2) が成り立つ． □

定理 1.9.1 の証明. 時刻 n を出発点とする確率制御問題

$$\begin{cases} V_n^*(i) = \inf_{\alpha \in \mathcal{A}_n(i)} \mathbb{E}\left[g_N(X_N^{n,i,\alpha}) + \sum_{k=n}^{N-1} g_k(X_k^{n,i,\alpha}, \alpha_k(X_k^{n,i,\alpha}), W_{k+1}) \right], \quad n = 0, \dots, N-1, \\ V_N^*(i) = g_N(i) \end{cases}$$

を考える．はじめに，各 $n = 0, \dots, N$ に対して

$$V_n(i) \le V_n^*(i), \quad i \in S \tag{1.9.3}$$

が成り立つことを後ろ向きの帰納法で示そう．$n = N$ のとき (1.9.3) が成立するのは明らか．一般の n に対して (1.9.3) を仮定する．このとき，任意の $\alpha \in \mathcal{A}_n$ に対して，

$$V_n(j) \le \mathbb{E}\left[g_N(X_N^{n,j,\alpha}) + \sum_{k=n}^{N-1} g_k(X_k^{n,j,\alpha}, \alpha_k(X_k^{n,j,\alpha}), W_{k+1}) \right], \quad j \in S.$$

よって，任意の $\alpha \in \mathcal{A}_{n-1}$ について，

$$\begin{aligned} V_{n-1}(i) &\le \mathbb{E}\left[g_{n-1}(i, \alpha_{n-1}(i), W_n) + V_n(f_{n-1}(i, \alpha_{n-1}(i), W_n)\right] \\ &\le \mathbb{E}[g_{n-1}(i, \alpha_{n-1}(i), W_n)] \\ &\quad + \mathbb{E}\left[\mathbb{E}\left[g_N(X_N^{n,j,\alpha}) + \sum_{k=n}^{N-1} g_k(X_k^{n,j,\alpha}, \alpha_k(X_k^{n,j,\alpha}), W_{k+1}) \right] \Bigg|_{j=f_{n-1}(i,\alpha_{n-1}(i),W_n)} \right]. \end{aligned}$$

ここで，$\{W_k\}$ が独立であることと $X_n^{n-1,i,\alpha} = f_{n-1}(i, \alpha_{n-1}(i), W_n)$ であることから命題 A.3.2(4) と補題 1.9.3 を適用し，

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[g_N(X_N^{n,j,\alpha}) + \sum_{k=n}^{N-1} g_k(X_k^{n,j,\alpha}, \alpha_k(X_k^{n,j,\alpha}), W_{k+1})\right]\Bigg|_{j=f_{n-1}(i,\alpha_{n-1}(i),W_n)}\right] \\
&= \mathbb{E}\left[g_N(X_N^{n-1,i,\alpha}) + \sum_{k=n}^{N-1} g_k(X_k^{n-1,i,\alpha}, \alpha_k(X_k^{n-1,i,\alpha}), W_{k+1})\right]
\end{aligned}$$

を得る．よって，

$$V_n(i) \leq \mathbb{E}\left[g_N(X_N^{n-1,i,\alpha}) + \sum_{k=n-1}^{N-1} g_k(X_k^{n-1,i,\alpha}, \alpha_k(X_k^{n-1,i,\alpha}), W_{k+1})\right].$$

$\alpha \in \mathcal{A}_n$ は任意だったから，これについて inf をとることで $V_n(i) \leq V_n^*(i)$, $i \in S$ が従う．

逆の不等号を示すために，$\varepsilon > 0$ を任意にとる．これと任意の n, i に対し，$\tilde{a}_{n,i} \in A_n(i)$ が存在して，

$$V_n(i) \geq \mathbb{E}[g_n(i, \tilde{a}_{n,i}, W_n) + V_{n+1}(f_n(i, \tilde{a}_{n,i}, W_n))] - \varepsilon. \tag{1.9.4}$$

この $\tilde{a}_{n,i}$ たちを用いて，$\tilde{\alpha}_n : S \to A$ を $\tilde{\alpha}_n(j) = \tilde{a}_{n,j}$ により定義すると $\tilde{\alpha} := \{\tilde{\alpha}_k\} \in \mathcal{A}$ である．

次に，任意の $n = 0, \ldots, N-1$ に対して

$$V_n(i) \geq -(N-n)\varepsilon + \mathbb{E}\left[g_N(X_N^{n,i,\tilde{\alpha}}) + \sum_{k=n}^{N-1} g_k(X_k^{n,i,\tilde{\alpha}}, \tilde{\alpha}_k(X_k^{n,i,\tilde{\alpha}}), W_{k+1})\right], \quad i \in S \tag{1.9.5}$$

が成り立つことを後ろ向きの帰納法で示す．$V_N = g_N$ より $n = N-1$ のときは明らかである．一般の n に対して (1.9.5) が成り立つとき，命題 A.3.2(4) と補題 1.9.3 を再度用いて

$$\begin{aligned}
&\mathbb{E}[V_n(f_{n-1}(i, \tilde{\alpha}_{n-1}(i), W_n)] \\
&\geq -(N-n)\varepsilon + \mathbb{E}\left[\mathbb{E}\left[g_N(X_N^{n,j,\tilde{\alpha}}) + \sum_{k=n}^{N-1} g_k(X_k^{n,j,\tilde{\alpha}}, \tilde{\alpha}_k(X_k^{n,j,\tilde{\alpha}}), W_{k+1})\right]\Bigg|_{j=f_{n-1}(i,\tilde{\alpha}_{n-1}(i),W_n)}\right] \\
&= -(N-n)\varepsilon + \mathbb{E}\left[g_N(X_N^{n-1,i,\tilde{\alpha}}) + \sum_{k=n}^{N-1} g_k(X_k^{n-1,i,\tilde{\alpha}}, \tilde{\alpha}_k(X_k^{n-1,i,\tilde{\alpha}}), W_{k+1})\right]
\end{aligned}$$

を得る．これと (1.9.4) より $n-1$ に対しても (1.9.5) が成り立つ．従って，任意の $n = 0, \ldots, N-1$ に対して (1.9.5) が成り立ち，それゆえ，

$$V_n(i) \geq -(N-n)\varepsilon + V_n^*(i), \quad i \in S.$$

よって $\varepsilon \to 0$ として $V_n(i) \geq V_n^*(i)$, $i \in S$ となる．

最後に，最適制御則についての主張を示す．α^* を定理内で記述されたものとすると，

$$V_n(i) = \mathbb{E}[g_n(i, \alpha_n^*(i), W_n) + V_{n+1}(f_n(i, \alpha_n^*(i), W_n))].$$

これを基に上と同様の議論をすることにより，

$$V_n(i) = \mathbb{E}\left[g_N(X_N^{n,i,\alpha^*}) + \sum_{k=n}^{N-1} g_k(X_k^{n,i,\alpha^*}, \alpha_k^*(X_k^{n,i,\alpha^*}), W_{k+1})\right], \quad i \in S$$

が得られる．ところが既に示したように，$V_n = V_n^*$ であるから，$n = 0$ として，

$$V_0^*(i) = \mathbb{E}\left[g_N(X_N^{i,\alpha^*}) + \sum_{k=0}^{N-1} g_k(X_k^{i,\alpha^*}, \alpha_k^*(X_k^{i,\alpha^*}), W_{k+1})\right], \quad i \in S.$$

これは α^* の最適性を意味している． □

第 2 章

連続時間マルコフ過程

応用上現れる問題を連続的な時間の流れで定式化することにより，解析的な扱いが可能になり，計算上有用なことが多くある．本章では，マルコフ性をもつ連続時間確率過程とそれらによって定義される確率過程を扱う．以下，現れる確率変数，確率過程は全て確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上のものとする．

2.1　ポアソン過程

ある特定の期間における災害などのイベント発生件数はポアソン分布を使ってモデル化されることが多い．期間を固定せず，継続的に件数を計測する場合に適した確率過程としてはポアソン過程が一般的である．

定義 **2.1.1.** $\lambda > 0$ とする．連続時間確率過程 $\{N_t\}_{t\geq 0}$ は以下の (1)-(5) を満たすとき，**強度**（intensity）λ の**一様ポアソン過程**（homogenous Poisson process）という．あるいは単に強度 λ のポアソン過程と呼ぶ.

(1) $N_0 = 0$.

(2) $t \mapsto N_t$ は右連続かつ左極限をもつ.

(3) 独立増分性: 任意の $n \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots < t_n$ に対して，$N_{t_i} - N_{t_{i-1}}$, $i = 1, \ldots, n$, は独立である.

(4) 定常増分性: 任意の $t, s \geq 0$ に対して，$N_{t+s} - N_t$ は N_s と同分布である.

(5) 各 $t \geq 0$ に対して N_t はパラメータ λt のポアソン分布に従う.

$\{N_t\}_{t\geq 0}$ を強度 λ のポアソン過程とするとき，上の定義より $E[N_t] = \lambda t$ である．従って，ポアソン過程をイベントの発生件数とみなすとき，強度 λ は単位時間当たりの期待発生件数と解釈できる．また，Taylor の定理より，$1 - e^{-h} = h + o(h)$ $(h \to 0)$ であるから，

$$\mathbb{P}(N_{t+h} - N_t = 1) = e^{-\lambda h}\lambda h = \lambda h + o(h) \quad (h \to 0)$$

が成り立つ．すなわち，微小な区間幅 h の間で N_t がジャンプする確率はおおよそ h に比

例し，その比例定数が λ というわけである．さらに，

$$\begin{aligned}\mathbb{P}(N_{t+h}-N_t\geq 2)&=1-\mathbb{P}(N_h=0)-\mathbb{P}(N_h=1)=1-e^{-\lambda h}-\lambda he^{-\lambda h}\\&=\lambda h+o(h)-\lambda he^{-\lambda h}=\lambda h(1-e^{-\lambda h})+o(h)\\&=o(h)\qquad(h\to 0).\end{aligned}$$

また，独立増分性と定常増分性を用いることで $\{N_t\}$ の有限次元分布を容易に計算することが出来る．実際，$t_1<t_2<\cdots<t_n$ に対して，

$$\begin{aligned}&(N_{t_1},N_{t_2},\ldots,N_{t_n})=\\&\left(N_{t_1},N_{t_2}-N_{t_1}+N_{t_1},\ldots,N_{t_1}+\sum_{i=2}^{n}(N_{t_i}-N_{t_{i-1}})\right)\end{aligned}$$

とみれば

$$\begin{aligned}&\mathbb{P}(N_{t_1}=k_1,N_{t_2}=k_1+k_2,\ldots,N_{t_n}=k_1+\cdots+k_n)\\&=\mathbb{P}(N_{t_1}-N_0=k_1,N_{t_2}-N_{t_1}=k_2,\ldots,N_{t_n}-N_{t_{n-1}}=k_n)\\&=\mathbb{P}(N_{t_1}-N_0=k_1)\mathbb{P}(N_{t_2}-N_{t_1}=k_2)\cdots\mathbb{P}(N_{t_n}-N_{t_{n-1}}=k_n)\\&=\mathbb{P}(N_{t_1}=k_1)\mathbb{P}(N_{t_2-t_1}=k_2)\cdots\mathbb{P}(N_{t_n-t_{n-1}}=k_n)\\&=e^{-\lambda t_1}\frac{(\lambda t_1)^{k_1}}{k_1!}e^{-\lambda(t_2-t_1)}\frac{(\lambda(t_2-t_1))^{k_2}}{k_2!}\cdots e^{-\lambda(t_n-t_{n-1})}\frac{(\lambda(t_n-t_{n-1}))^{k_n}}{k_n!}\\&=e^{-\lambda t_n}\frac{(\lambda t_1)^{k_1}}{k_1!}\frac{(\lambda(t_2-t_1))^{k_2}}{k_2!}\cdots\frac{(\lambda(t_n-t_{n-1}))^{k_n}}{k_n!}.\end{aligned}$$

確率変数列 $\{W_n\}_{n=1}^{\infty}$ に対し $T_0=0$, $T_n=W_1+\cdots+W_n$, $n\in\mathbb{N}$ とおき，連続時間確率過程

$$N_t:=\max\{n\in\mathbb{N}\cup\{0\}:T_n\leq t\},\quad t\geq 0 \tag{2.1.1}$$

を考える．

定理 **2.1.2.** 確率変数列 $\{W_n\}_{n=1}^{\infty}$ は独立同分布で，各 W_n がパラメータ λ の指数分布に従うとき，(2.1.1) で定義される $\{N_t\}_{t\geq 0}$ は強度 λ のポアソン過程となる．

証明．$\{W_n\}$ に対する仮定と帰納法により

$$\mathbb{P}(T_n\leq x)=1-e^{-\lambda x}\sum_{k=0}^{n-1}\frac{(\lambda x)^k}{k!},\quad x\geq 0 \tag{2.1.2}$$

を確かめることができる．よって，各 $n\in\mathbb{N}\cup\{0\}$ に対して $\{N_t=n\}=\{T_n\leq t<T_{n+1}\}$ であることより

$$\mathbb{P}(N_t=n)=\mathbb{P}(T_n\leq t)-\mathbb{P}(T_{n+1}\leq t)=e^{-\lambda t}\frac{(\lambda t)^n}{n!} \tag{2.1.3}$$

となる．これより

$$\begin{aligned}&\mathbb{P}(N_t=n,N_{t+s}=n)=\mathbb{P}(T_n\leq t,\ t+s<T_n+W_{n+1})\\&=\int_0^t\int_{t+s-x}^{\infty}\lambda^2e^{-\lambda(x+y)}\frac{(\lambda x)^{n-1}}{(n-1)!}dydx=e^{-\lambda(t+s)}\frac{(\lambda t)^n}{n!}\end{aligned}$$

である．加えて，

$$
\begin{aligned}
&\mathbb{P}(N_t = n, N_{t+s} = n+1) \\
&= \mathbb{P}(T_n \le t < T_n + W_{n+1}, T_n + W_{n+1} \le t+s < T_n + W_{n+1} + W_{n+2}) \\
&= \int_0^t \int_{t-x}^{t+s-x} \int_{t+s-x-y}^{\infty} \lambda^3 e^{-\lambda(x+y+z)} \frac{(\lambda x)^{n-1}}{(n-1)!} dzdydx = e^{-\lambda(t+s)}(\lambda s)\frac{(\lambda t)^n}{n!}
\end{aligned}
$$

であり，$m \ge 2$ のとき，$Z := W_{n+2} + \cdots + W_{n+m}$ は T_{m-1} と同分布であるから，(2.1.2) を用いて，

$$
\begin{aligned}
&\mathbb{P}(N_t = n, N_{t+s} = n+m) \\
&= \mathbb{P}(T_n \le t < T_n + W_{n+1}, T_n + W_{n+1} + Z \le t+s < T_n + W_{n+1} + Z + W_{n+m+1}) \\
&= \int_0^t \int_{t-x}^{t+s-x} \int_0^{t+s-x-y} \int_{t+s-x-y-z}^{\infty} \lambda^4 e^{-\lambda(x+y+z+u)} \frac{(\lambda z)^{m-2}}{(m-2)!} \frac{(\lambda x)^{n-1}}{(n-1)!} dudzdydx \\
&= e^{-\lambda(t+s)} \frac{(\lambda s)^m}{m!} \frac{(\lambda t)^n}{n!}
\end{aligned}
$$

を得る．よって，(2.1.3) を考慮すれば

$$
\begin{aligned}
\mathbb{P}(N_{t+s} - N_t = m) &= \sum_{n=0}^{\infty} \mathbb{P}(N_{t+s} - N_t = m, N_t = n) = \sum_{n=0}^{\infty} \mathbb{P}(N_{t+s} = n+m, N_t = n) \\
&= \sum_{n=0}^{\infty} e^{-\lambda(t+s)} \frac{(\lambda s)^m}{m!} \frac{(\lambda t)^n}{n!} = e^{-\lambda s} \frac{(\lambda s)^m}{m!} = \mathbb{P}(N_s = m)
\end{aligned}
$$

となる．すなわち，$\{N_t\}$ は定常増分性を持つ．さらに，

$$
\mathbb{P}(N_t = n, N_{t+s} - N_t = m) = \mathbb{P}(N_{t+s} = n+m, N_t = n) = \mathbb{P}(N_t = n)\mathbb{P}(N_{t+s} - N_t = m)
$$

が従う．ゆえに N_t と $N_{t+s} - N_t$ は独立である．この議論を（煩雑だが）一般化することで，$\{N_t\}$ の独立増分性も示すことができる． □

定理 **2.1.3.** $\{N_t\}_{t\ge 0}$ をポアソン過程とする．このとき，任意の $n \in \mathbb{N}, 0 \le t_1 < \cdots < t_n$ なる $t_1, \ldots, t_n \ge 0$, $k_1 \le \cdots \le k_n$ なる $k_1, \ldots, k_n \in \mathbb{N} \cup \{0\}$ に対して，

$$
\mathbb{P}(N_{t_n} = k_n \,|\, N_{t_1} = k_1, \ldots, N_{t_{n-1}} = k_{n-1}) = \mathbb{P}(N_{t_n} = k_n \,|\, N_{t_{n-1}} = k_{n-1}).
$$

- 定理 2.1.3 はポアソン過程のマルコフ性を意味している．実際，後で見るように，ポアソン過程は連続時間マルコフ連鎖の特別な例となっている．
- 定理 2.1.3 はポアソン過程の有限次元分布を用いて容易に確かめることができる．よって証明は省略する．

ある程度長い時間区間では，強度も時間が経つにつれて変化すると考える方が現実的である．このことを考慮して，ポアソン過程の定義も一般化しておこう．

定義 **2.1.4** (非一様ポアソン過程). $\lambda(t)$, $t \geq 0$ を非負の区分的連続関数とする．連続時間確率過程 $\{N_t\}_{t\geq 0}$ は以下の (1)-(4) を満たすとき，強度関数（intensity function）λ の**非一様ポアソン過程**（non-homogenous Poisson process）という．

(1) $N_0 = 0$.
(2) $t \mapsto N_t$ は右連続かつ左極限をもつ．
(3) $\{N_t\}$ は独立増分性をもつ．
(4) 任意の $0 \leq s < t$ に対して，$N_t - N_s$ はパラメータ $\int_s^t \lambda(u)du$ のポアソン分布に従う．

$\{N_t\}_{t\geq 0}$ を強度関数 λ のポアソン過程とするとき，一様の場合と同じ議論により，$h \to 0$ のとき

$$\begin{aligned}\mathbb{P}(N_{t+h} - N_t = 1) &= e^{-\int_t^{t+h}\lambda(u)du}\int_t^{t+h}\lambda(u)du = \int_t^{t+h}\lambda(u)du + o(h)\\ &= \lambda(t)h + o(h)\end{aligned}$$

であり，

$$\mathbb{P}(N_{t+h} - N_t \geq 2) = o(h)$$

となる．

- 非一様の場合も含み，ポアソン過程をデンマークの火災発生件数へ応用した例が [3] に挙げられている．

2.2 複合ポアソン過程

$\{N_t\}$ を強度 λ の一様ポアソン過程，$\{X_n\}_{n=1}^{\infty}$ を独立同分布かつ $\{N_t\}$ と独立となるような確率変数列で，各 X_n は正値の確率変数とする．このとき，次で定義される確率過程 $\{S_t\}$ を考える:

$$S_t = \sum_{n=1}^{N_t} X_n, \quad t \geq 0 \tag{2.2.1}$$

ただし $\sum_{n=1}^{0} = 0$ とする．$\{S_t\}_{t\geq 0}$ は**複合ポアソン過程**（compound Poisson process）と呼ばれる．

- N_t を時刻 t までに起こった災害件数，X_n を n 番目に起きた災害による損害額とすると，S_t は時刻 t までに起こった災害による累計損害額となる．
- [3] において，やはりデンマークの火災データを用いて損害額分布 X_n の推定が行われている．

S_t の期待値，分散はそれぞれ

$$\mathbb{E}[S_t] = \mathbb{E}[X_1]\mathbb{E}[N_t] = \lambda t\mathbb{E}[X_1], \quad \mathbb{V}[S_t] = \mathbb{E}[N_t]\mathbb{V}[X_1] + \mathbb{E}[X_1]^2\mathbb{V}[N_t] = \lambda t\mathbb{E}[X_1^2]$$

により与えられる．また S_t の特性関数 $\phi_{S_t}(u)$ は

$$\begin{aligned}\varphi_{S_t}(u) &= \sum_{n=0}^{\infty} \mathbb{E}[e^{\sqrt{-1}u\sum_{k=0}^{n} X_k} 1_{N_t=n}] = \sum_{n=0}^{\infty} \varphi(u)^n \mathbb{P}(N_t = n) \\ &= \mathbb{E}[\varphi_{X_1}(u)^{N_t}] = \mathbb{E}[e^{N_t \log \varphi_{X_1}(u)}], \quad u \in \mathbb{R}\end{aligned}$$

によって与えられる．ここで，N_t は $\mathbb{N} \cup \{0\}$-値であるから上式の右辺は log の分枝の取り方によらず決まる．さらに，N_t はパラメータ λt のポアソン分布に従うので，

$$\mathbb{E}[e^{hN_t}] = \sum_{k=0}^{\infty} e^{hk} e^{-\lambda t} \frac{(\lambda t)^k}{k!} = e^{-\lambda t(1-e^h)}, \quad h \in \mathbb{C}$$

となる．従って，

$$\varphi_{S_t}(u) = e^{-\lambda t(1-\varphi_{X_1}(u))}.$$

一般に，$\mathbb{R}$ 上の関数 F が右連続かつ非減少で，$F(-\infty) = \lim_{x\to-\infty} F(x) = 0$ および $F(+\infty) = \lim_{x\to\infty} F(x) = 1$ を満たすとき，これを**分布関数**（distribution function）という．確率変数 X に対して $P(X \leq x)$, $x \in \mathbb{R}$, は上の性質を満たしており，X の分布関数と呼ぶ．これ以後，本節の終わりまで，X の分布関数を F_X と書くことにする．

一般に，確率変数 X, Y が独立のとき，

$$F_{X+Y}(x) = \int_{-\infty}^{\infty} F_X(x-y)dF_Y(y) = (F_X * F_Y)(x), \quad x \in \mathbb{R}$$

と表される．ただし，分布関数 F, G に対して

$$F * G(x) = \int_{-\infty}^{\infty} F(x-y)dG(y), \quad x \in \mathbb{R}.$$

これは分布関数 F と G の**畳み込み**（convolution）と呼ばれ，再び分布関数になる．今，

$$S_0 = 0, \quad S_n = X_1 + \cdots + X_n, \quad n \in \mathbb{N}$$

とおくとき，

$$F_{S_t}(x) = \mathbb{E}[\mathbb{P}(S_t \leq x \,|\, N_t)] = \sum_{n=0}^{\infty} \mathbb{P}(S_n \leq x)\mathbb{P}(N_t = n)$$

となる．従って，分布関数 F に対して，

$$F^{*0}(x) = \begin{cases} 1 & (x \geq 0), \\ 0 & (x < 0), \end{cases}$$
$$F^{*n}(x) = F * F^{*(n-1)}(x), \quad n \in \mathbb{N},\ x \in \mathbb{R}$$

と定義すると，

$$F_{S_t}(x) = \sum_{n=0}^{\infty} \mathbb{P}(N_t = n)F_{S_n}(x) = \sum_{n=0}^{\infty} \mathbb{P}(N_t = n)F_{X_1}^{*n}(x) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} F_{X_1}^{*n}(x)$$

と表すことができる．

例 **2.2.1.** X_1 の分布がパラメータ δ の指数分布である場合を考えよう．このとき，$F_{X_1}^{*n}$ は

$$F_{X_1}^{*n}(x) = \frac{\delta^n}{(n-1)!}\int_0^x e^{-\delta y} y^{n-1} dy, \quad x \in [0,\infty), \ n \in \mathbb{N}$$

により与えられる（読者はこれを確かめよ）．従って，$F_{S_t}(0) = e^{-\lambda}$ であり，また F_{S_t} は $(0,\infty)$ で微分可能であって，S_t のそこでの密度関数 $F'_{S_t}(x)$, $x \in (0,\infty)$, は

$$F'_{S_t}(x) = e^{-(\lambda+x\delta)}\sum_{n=1}^{\infty}\frac{1}{n!(n-1)!}\frac{1}{x}(x\lambda\delta)^n = e^{-(\lambda+x\delta)}\sqrt{\lambda\delta/x}\,I_1(2\sqrt{\lambda x\delta})$$

と書ける．ただし $I_\nu(z)$ は以下で定義される修正 Bessel 関数である:

$$I_\nu(z) = \sum_{n=0}^{\infty}\frac{(z/2)^{2n+\nu}}{n!\Gamma(\nu+n+1)}, \quad z,\nu \in \mathbb{R}.$$

2.3 連続時間マルコフ連鎖

ここでは連続時間のマルコフ連鎖について考えよう．第 1 章と同様に，S を高々可算集合とし，これをマルコフ連鎖の状態空間とする．

定義 **2.3.1.** $i,j \in S$ に対して $p_{ij} : [0,\infty) \to [0,1]$ が存在して以下の (1)–(3) をみたすとき，$P(t) = \{p_{ij}(t)\}_{i,j\in S}$, $t \geq 0$ を**推移確率関数**（transition probability function）という．

(1) 各 $i \in S$ と $t \geq 0$ に対して，

$$\sum_{j\in S} p_{ij}(t) = 1.$$

(2) 半群性:

$$P(s+t) = P(s)P(t), \quad s,t \geq 0.$$

(3) $t \mapsto P(t)$ は 0 で連続である．i.e.,

$$\lim_{h\searrow 0} P(h) = P(0) = I_{|S|}.$$

- 離散時間の場合と同様に，$\mathbb{R}^{|S|\times|S|}$ の行列と同一視するとき，$P(t)$ を推移確率行列と呼ぶ．
- 性質 (2) を成分で書き表すと，

$$p_{ij}(s+t) = \sum_{k\in S} p_{ik}(s)p_{kj}(t), \quad i,j \in S, \ \ s,t \geq 0.$$

となる．これ（あるいは性質 (2) 自体）をチャップマン・コルモゴロフ方程式と呼ぶ．

- δ_{ij} をディラックのデルタとする．すなわち，$i, j \in S$ に対して，

$$\delta_{ij} = \begin{cases} 1 & (i = j), \\ 0 & (i \neq j). \end{cases}$$

このとき，定義 2.3.1 の (3) は

$$\lim_{t \searrow 0} p_{ij}(t) = \delta_{ij}, \quad i, j \in S$$

と書き表される．

定義 **2.3.2.** S 値連続時間確率過程 $\{X_t\}_{t \geq 0}$ が，任意の $n \geq 1$, $h > 0$, $0 \leq t_0 < t_1 < \cdots < t_{n-1} < t_n$ と任意の $j \in S$, および $\mathbb{P}(X_{t_0} = i_0, \ldots, X_{t_{n-1}} = i_{n-1}, X_{t_n} = i) > 0$ なる任意の $i_0, \ldots, i_{n-1}, i \in S$ に対して

$$\mathbb{P}(X_{t_n+h} = j \,|\, X_{t_0} = i_0, \ldots, X_{t_{n-1}} = i_{n-1}, X_{t_n} = i) = p_{ij}(h) \tag{2.3.1}$$

を満たすとき，推移確率 $P(\cdot)$ をもつ**連続時間マルコフ連鎖**（continuous-time Markov chain）という．

- 離散時間の場合と同様に，条件 (2.3.1) より

$$\mathbb{P}(X_{t_n+h} = j \,|\, X_{t_0} = i_0, \ldots, X_{t_{n-1}} = i_{n-1}, X_{t_n} = i) = \mathbb{P}(X_{t_n+h} = j \,|\, X_{t_n} = i)$$

である．

例 **2.3.3.** 強度 λ の一様ポアソン過程 $\{N_t\}$ が状態空間 $S = \mathbb{N} \cup \{0\}$ 上の連続時間マルコフ連鎖であることを確かめてみよう．定義より，

$$\begin{aligned} \mathbb{P}(N_{t+h} = i + n \,|\, N_t = i) &= \frac{\mathbb{P}(N_{t+h} - N_t = n, N_t = i)}{\mathbb{P}(N_t = i)} \\ &= \mathbb{P}(N_{t+h} - N_t = n) = \mathbb{P}(N_h = n) = e^{-\lambda h} \frac{(\lambda h)^n}{n!}. \end{aligned}$$

よって，

$$p_{ij}(t) = \begin{cases} e^{-\lambda t} \dfrac{(\lambda t)^{j-i}}{(j-i)!} & (j \geq i), \\ 0 & (j < i) \end{cases}$$

と定義すれば，定理 2.1.3 より，$\{N_t\}$ が (2.3.1) を満たすことが分かる．また，この $\{p_{ij}(\cdot)\}_{i,j \in S}$ が定義 2.3.1 の (1) と (3) を満たすことは明らかなので，(2) を直接計算により確かめる．$j < i$ のとき，任意の $k \in S$ に対して $p_{ik}(t) p_{kj}(s) = 0$. よって

$$\sum_{k \in S} p_{ik}(t) p_{kj}(s) = 0 = p_{ij}(t + s).$$

$j \ge i$ のとき，

$$
\begin{aligned}
\sum_{k\in S} p_{ik}(t)p_{kj}(s) &= e^{-\lambda(t+s)} \sum_{k=i}^{j} \frac{(\lambda t)^{k-i}}{(k-i)!}\frac{(\lambda s)^{j-k}}{(j-k)!} \\
&= \frac{e^{-\lambda(t+s)}}{(j-i)!}\sum_{\ell=0}^{j-i}\frac{(j-i)!}{\ell!(j-i-\ell)!}(\lambda t)^{\ell}(\lambda s)^{j-i-\ell} = e^{-\lambda(t+s)}\frac{(\lambda(t+s))^{j-i}}{(j-i)!} \\
&= p_{ij}(t+s).
\end{aligned}
$$

次は命題 1.1.9 の連続時間版である．

命題 2.3.4. $\{X_t\}$ を連続時間マルコフ連鎖とし，$n \ge 0$, $m \ge 1$, $0 \le t_0 \le \cdots \le t_{n+m}$ とする．このとき，$i_n, \ldots, i_{n+m} \in S$ に対し，

$$
\begin{aligned}
&\mathbb{P}(X_{t_n} = i_n, \ldots, X_{t_{n+m}} = i_{n+m}) \\
&= \mathbb{P}(X_{t_n} = i_n)p_{i_n i_{n+1}}(t_{n+1}-t_n)\cdots p_{i_{n+m-1}i_{n+m}}(t_{n+m}-t_{n+m-1}).
\end{aligned}
$$

証明は命題 1.1.9 のそれとほとんど同じなので省略する．

推移確率関数の主要な性質を述べる．

補題 2.3.5. $P(\cdot) = \{p_{ij}(\cdot)\}$ が推移確率関数ならば，各 $p_{ij}(\cdot)$ は $[0,\infty)$ 上連続である．

証明. $t, h \ge 0$ に対して，チャップマン・コルモゴロフ方程式と $p_{kj}(t) \le 1$ より，

$$
\begin{aligned}
|p_{ij}(t+h) - p_{ij}(t)| &= \left|\sum_{k\ne i} p_{ik}(h)p_{kj}(t) + (p_{ii}(h)-1)p_{ij}(t)\right| \le \sum_{k\ne i} p_{ik}(h) + 1 - p_{ii}(h) \\
&\le 2(1-p_{ii}(h)).
\end{aligned}
$$

推移確率関数の性質 (3) より，$h \searrow 0$ のとき，右辺は 0 に収束する．

同様に，

$$
\begin{aligned}
|p_{ij}(t-h) - p_{ij}(t)| &= \left|(1-p_{ii}(h))p_{ij}(t-h) + \sum_{k\ne i} p_{ik}(h)p_{kj}(t-h)\right| \\
&\le 2(1-p_{ii}(h)) \to 0, \quad h \searrow 0.
\end{aligned}
$$

以上より p_{ij} は各 $t > 0$ で連続である． □

補題 2.3.6. $f : (0,\infty) \to \mathbb{R}$ は $\lim_{t\searrow 0} f(t) = 0$ となる非負関数で，劣加法的とする．すなわち，

$$
f(t+s) \le f(t) + f(s), \quad s, t \in (0,\infty)
$$

を満たすとする．このとき，

$$
q = \sup_{t>0} \frac{f(t)}{t} \in [0,\infty]
$$

とおくと，

$$
\lim_{h\searrow 0} \frac{f(h)}{h} = q.
$$

証明. q の定義より，任意の $a \in [0,q)$ に対して，ある $t_0 > 0$ が存在し，$f(t_0)/t_0 \geq a$. 勝手な $t > 0$ をとり，$n \in \mathbb{N}$ を t_0/t の整数部分とすると，$nt \leq t_0 < nt + t$ であるから $t_0 = nt + \delta$ なる $\delta \in [0,t)$ が存在する．ここで劣加法性を（繰り返し）用いて

$$f(t_0) \leq f(nt) + f(\delta) \leq nf(t) + f(\delta)$$

が得られ，従って

$$a \leq \frac{f(t_0)}{t_0} \leq \frac{nt}{t_0} \cdot \frac{f(t)}{t} + \frac{f(\delta)}{t_0}.$$

$t_0 - t < nt \leq t_0$ を考慮すると，$t \searrow 0$ のとき，$nt \to t_0$ であり，$\delta \to 0$. これらと f に対する仮定より，

$$a \leq \liminf_{t \searrow 0} \frac{f(t)}{t} \leq \limsup_{t \searrow 0} \frac{f(t)}{t} \leq q.$$

$a < q$ は任意だったので，補題の主張が従う． □

定理 **2.3.7.** $P(\cdot) = \{p_{ij}(\cdot)\}_{i,j \in S}$ を推移確率関数とする．このとき，各 $i, j \in S$ に対し，極限

$$q_{ij} := \lim_{h \searrow 0} \frac{p_{ij}(h) - \delta_{ij}}{h} \tag{2.3.2}$$

が存在する．特に，$i \neq j$ のとき，この極限値は有限である．すなわち，$q_{ij} < \infty$. さらに，S が有限集合ならば，$q_{ii} > -\infty$.

証明*. $\lim_{h \searrow 0} p_{ii}(h) = 1$ より，ある $\varepsilon > 0$ が存在し，$p_{ii}(h) > 0$, $h \in [0, \varepsilon]$. 任意の $t \geq 0$ に対し，$t/n \leq \varepsilon$ となるように十分大きい $n \in \mathbb{N}$ をとれば，$p_{ii}(t) \geq (p_{ii}(t/n))^n > 0$. よって，関数

$$f(t) := -\log p_{ii}(t), \quad t > 0$$

は非負実数値で，$\lim_{t \searrow 0} f(t) = 0$ を満たす．さらに，$p_{ii}(t+s) \geq p_{ii}(t) p_{ii}(s)$ より，f は劣加法的である．これより補題 2.3.6 が適用でき，

$$\lim_{h \searrow 0} \frac{f(h)}{h} = \sup_{t > 0} \frac{f(t)}{t} \in [0, \infty].$$

ゆえに，

$$\lim_{h \searrow 0} \frac{1 - p_{ii}(h)}{h} = \lim_{h \searrow 0} \frac{1 - e^{-f(h)}}{f(h)} \frac{f(h)}{h} = \sup_{t > 0} \frac{f(t)}{t}.$$

すなわち，q_{ii} は $-\infty$ に発散するかもしれないが，確定する．

次に，$i \neq j$ の場合に (2.3.2) が成り立つことを示そう．$h > 0$ を任意にとって固定し，$\{X_n\}_{n=0}^{\infty}$ を推移確率 $P(h)$ を持つ離散時間マルコフ連鎖で，各 $i \in S$ に対して $\mathbb{P}(X_0 = i) > 0$ をみたすものとする．

$$p_{ii}^{j0}(h) = 1, \quad p_{ii}^{j1}(h) = p_{ii}(h),$$
$$p_{ii}^{jm}(h) = \mathbb{P}(X_m = i, X_\ell \neq j, \ell = 1, \ldots, m-1 \mid X_0 = i), \quad m \geq 2$$

とおく．このとき，$i \neq j$ より $n \geq 3$ に対して

$$\begin{aligned}
p_{ij}(nh) &= \mathbb{P}(X_n = j \,|\, X_0 = i) \\
&= \mathbb{P}(X_n = j, X_1 = j \,|\, X_0 = i) \\
&\quad + \sum_{m=2}^{n} \mathbb{P}(X_n = j, X_m = j, X_\ell \neq j, \ell = 1, \ldots, m-1 \,|\, X_0 = i) \\
&= \mathbb{P}(X_n = j, X_1 = j \,|\, X_0 = i) \\
&\quad + \sum_{m=2}^{n} \sum_{\substack{j_\ell \neq j \\ \ell = 1, \ldots, m-1}} \mathbb{P}(X_n = j, X_m = j, X_\ell = j_\ell, \ell = 1, \ldots, m-1 \,|\, X_0 = i).
\end{aligned}$$

最後の等式の右辺に命題 1.1.9 を適用し，

$$\begin{aligned}
p_{ij}(nh) &\geq p_{ij}(h)p_{jj}((n-1)h) + p_{ii}(h)p_{ij}(h)p_{jj}((n-2)h) \\
&\quad + \sum_{m=3}^{n} \sum_{\substack{j_\ell \neq j \\ \ell = 1, \ldots, m-2}} p_{ij_1}(h)p_{j_1 j_2}(h) \cdots p_{j_{m-2} i}(h)p_{ij}(h)p_{jj}((n-m)h).
\end{aligned}$$

ここで $p_{ii}^{jm}(h)$ の定義を思い出し

$$\begin{aligned}
&\sum_{\substack{j_\ell \neq j \\ \ell = 1, \ldots, m-2}} p_{ij_1}(h)p_{j_1 j_2}(h) \cdots p_{j_{m-2} i}(h) \\
&= \mathbb{P}(X_{m-1} = i, X_\ell \neq j, \ell = 1, \ldots, m-2 \,|\, X_0 = i) = p_{ii}^{j, m-1}(h).
\end{aligned}$$

よって

$$p_{ij}(nh) \geq \sum_{m=0}^{n-1} p_{ii}^{jm}(h)p_{ij}(h)p_{jj}((n-m-1)h). \tag{2.3.3}$$

さらに，

$$f_{ij}^1(h) = p_{ij}(h), \quad f_{ij}^m(h) = \mathbb{P}(X_m = j, X_\ell \neq j, \ell = 1, \ldots, m-1 \,|\, X_0 = i), \quad m \geq 2$$

とおく．このとき，$m \geq 3$ に対して，

$$\begin{aligned}
p_{ii}(mh) &= \mathbb{P}(X_m = i, X_\ell \neq j, \ell = 1, \ldots, m-1 \,|\, X_0 = i) + \mathbb{P}(X_m = i, X_1 = j \,|\, X_0 = i) \\
&\quad + \sum_{\nu=2}^{m-1} \mathbb{P}(X_m = i, X_\nu = j, X_\ell \neq j, \ell = 1, \ldots, \nu-1 \,|\, X_0 = i).
\end{aligned}$$

これを基に上と同様の議論を行って，$m = 2$ の場合も含めて，

$$p_{ii}(mh) = p_{ii}^{jm}(h) + \sum_{\nu=1}^{m-1} f_{ij}^{\nu}(h)p_{jj}((m-\nu)h)$$

を得る．これと $\sum_{\nu=1}^{m-1} f_{ij}^{\nu}(h) \leq 1$ より，

$$p_{ii}(mh) \leq p_{ii}^{jm}(h) + \max_{1 \leq \nu < m} p_{jj}((m-\nu)h).$$

ゆえに,

$$p_{ii}^{jm}(h) \geq p_{ii}(mh) - \max_{1\leq \nu < m} p_{jj}((m-\nu)h), \quad m \geq 2. \tag{2.3.4}$$

他方,推移確率の定義より,十分小さい $\varepsilon > 0$ に対し,$\delta > 0$ が存在し,

$$\begin{aligned} \sup_{0\leq h\leq \delta} p_{ij}(h) < \varepsilon, &\quad \sup_{0\leq h\leq \delta} p_{ji}(h) < \varepsilon, \\ \inf_{0\leq h\leq \delta} p_{ii}(h) > 1-\varepsilon, &\quad \inf_{0\leq h\leq \delta} p_{jj}(h) > 1-\varepsilon. \end{aligned} \tag{2.3.5}$$

従って,(2.3.3)–(2.3.5) より,$nh < \delta$ のとき,

$$\begin{aligned} p_{ij}(nh) &\geq p_{ij}(h)p_{jj}((n-1)h) + p_{ii}(h)p_{ij}(h)p_{jj}((n-2)h) \\ &\quad + \sum_{m=2}^{n-1} p_{ii}^{jm}(h)p_{ij}(h)p_{jj}((n-m-1)h) \\ &\geq (1-\varepsilon)p_{ij}(h) + (1-\varepsilon)^2 p_{ij}(h) + \sum_{m=2}^{n-1}(1-2\varepsilon)(1-\varepsilon)p_{ij}(h). \end{aligned}$$

ところが $\min\{1-\varepsilon, (1-\varepsilon)^2, (1-2\varepsilon)(1-\varepsilon)\} \geq 1-3\varepsilon$ であるから結局,$p_{ij}(nh) \geq (1-3\varepsilon)np_{ij}(h)$. すなわち $nh < \delta$ のとき,

$$\frac{p_{ij}(h)}{h} \leq \frac{1}{1-3\varepsilon}\frac{p_{ij}(nh)}{nh}. \tag{2.3.6}$$

$t, h \in (0,\delta)$ とし,n を t/h の整数部分とする.このとき,$\lim_{h\searrow 0} nh = t$ であり,補題 2.3.5 より $s \mapsto p_{ij}(s)/s$ は t で連続だから,$\lim_{h\searrow 0} p_{ij}(nh)/(nh) = p_{ij}(t)/t$. さらに,$nh \leq t < \delta$ であるから (2.3.6) を適用し

$$\limsup_{h\searrow 0}\frac{p_{ij}(h)}{h} \leq \frac{1}{1-3\varepsilon}\frac{p_{ij}(t)}{t} < \infty.$$

ここで $t \searrow 0$ として

$$\limsup_{h\searrow 0}\frac{p_{ij}(h)}{h} \leq \frac{1}{1-3\varepsilon}\liminf_{t\searrow 0}\frac{p_{ij}(t)}{t} < \infty.$$

最後に $\varepsilon \searrow 0$ として

$$\limsup_{h\searrow 0}\frac{p_{ij}(h)}{h} \leq \liminf_{t\searrow 0}\frac{p_{ij}(t)}{t} < \infty$$

を得る.これは (2.3.2) の成立を意味する.

S を有限集合とすると,極限と和の交換ができるので,

$$q_{ii} = \lim_{h\searrow 0}\frac{p_{ii}(h)-1}{h} = -\lim_{h\searrow 0}\sum_{j\neq i}\frac{p_{ij}(h)}{h} = -\sum_{j\neq i}\lim_{h\searrow 0}\frac{p_{ij}(h)}{h} = -\sum_{j\neq i} q_{ij} > -\infty.$$

以上で定理の主張が全て示された. □

- 各 q_{ij} を(i から j への)**推移率**(transition rate),$Q = \{q_{ij}\}_{i,j\in S}$ を**推移率行列**(transition rate matrix)という.

- 推移率行列のことを Q 行列と呼ぶこともある．
- (2.3.2) を行列表記すれば，

$$Q = \lim_{h \searrow 0} \frac{P(h) - P(0)}{h}.$$

- 各 q_{ij} は

$$q_{ij} \geq 0, \quad i \neq j, \ \ i, j \in S \tag{2.3.7}$$

  を満たしていることに注意しよう．

例 **2.3.8.** 例 2.3.3 より，ポアソン過程に対応する推移確率関数は

$$p_{ij}(t) = \begin{cases} e^{-\lambda t} \dfrac{(\lambda t)^{j-i}}{(j-i)!} & (j \geq i), \\ 0 & (j < i) \end{cases}$$

である．これより，推移率行列 $Q = \{q_{ij}\}$ は

$$q_{ij} = \begin{cases} -\lambda & (j = i) \\ \lambda & (j = i+1) \\ 0 & (j \neq i, i+1). \end{cases}$$

により与えられる．

推移確率関数の定義より $\sum_{j \in S} p_{ij}(h) = 1$ であるから，

$$\frac{1 - p_{ii}(h)}{h} = \sum_{j \neq i} \frac{p_{ij}(h)}{h}.$$

ここで両辺 $h \searrow 0$ とするとき，もし右辺の和と極限の交換ができるなら（これは特に S が有限集合のとき可能である），$-q_{ii} = \sum_{j \neq i} q_{ij}$ となる．よって

$$q_{ii} > -\infty, \quad i \in S \tag{2.3.8}$$

ならば

$$\sum_{j \in S} q_{ij} = 0, \quad i \in S \tag{2.3.9}$$

が得られる．

逆に，(2.3.7)–(2.3.9) を満たす二重数列 $Q = \{q_{ij}\}_{i,j \in S}$ が与えられたとき，それを推移率行列とするようなマルコフ連鎖を構成することもできる．次の定理の（完全な）証明と具体的な構成法については例えば [4] の第 5 章や [2] の第 9 章を見よ．

定理 **2.3.9.** 二重数列 $Q = \{q_{ij}\}_{i,j \in S}$ が (2.3.7)–(2.3.9) を満たすとする．このとき，Q を推移率行列とする連続時間マルコフ連鎖が存在する．

次に，S を有限集合と仮定して，推移率行列と推移確率関数の関係をより詳細に調べる．$P = \{p_{ij}\}_{i,j\in S}$ を推移確率関数，Q を対応する推移率行列とする．このとき，チャップマン・コルモゴロフ方程式より

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h} = \sum_{k\in S}\frac{p_{ik}(h)-\delta_{ik}}{h}p_{kj}(t),$$
$$\frac{p_{ij}(t-h)-p_{ij}(t)}{-h} = \sum_{k\in S}\frac{p_{ik}(h)-\delta_{ik}}{h}p_{kj}(t-h)$$

を得る．これらと定理 2.3.7，および S が有限集合であることより，$t\mapsto p_{ij}(t)$ は $t>0$ で微分可能であり，$\{p_{ij}(t)\}_{i,j\in S}$ は微分方程式の系

$$\frac{dp_{ij}(t)}{dt} = \sum_{k\in S} q_{ik}p_{kj}(t), \quad i,j\in S$$

を満たす．これを行列で書けば

$$\frac{dP(t)}{dt} = QP(t), \quad t>0. \tag{2.3.10}$$

初期条件は $P(0)=I_{|S|}$ である．(2.3.10) を**コルモゴロフの後退方程式**（Kolmogorov backward equation）という．

また，チャップマン・コルモゴロフ方程式の使い方を変えると，

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h} = \sum_{k\in S}p_{ik}(t)\frac{p_{kj}(h)-\delta_{kj}}{h}$$

を得るので，両辺 $h\searrow 0$ とすると，上と同様に，微分方程式の系

$$\frac{dp_{ij}(t)}{dt} = \sum_{k\in S} p_{ik}(t)q_{kj}, \quad i,j\in S$$

が導かれる．これを行列で書き表すと

$$\frac{dP(t)}{dt} = P(t)Q, \quad t>0. \tag{2.3.11}$$

初期条件はやはり $P(0)=I_{|S|}$ である．(2.3.11) を**コルモゴロフの前進方程式**（Kolmogorov forward equation）という．

- (2.3.10) を導く際にはチャップマン・コルモゴロフ方程式を

  $$p_{ij}(t+h) = \sum_{k\in S}p_{ik}(h)p_{kj}(t)$$

  という形で用いた．$t_0>0$ をとり，h を十分小さいとすると，上の等式は

  $$\mathbb{P}(X_{t_0+t}=j\,|\,X_{t_0-h}=i) = \sum_{k\in S}\mathbb{P}(X_{t_0+t}=j\,|\,X_{t_0}=k)\mathbb{P}(X_{t_0}=k\,|\,X_{t_0-h}=i)$$

  と表される．すなわち，これは，推移確率を考える区間を $[t_0,t_0+t]$ から過去に遡って $[t_0-h,t_0+t]$ に延長し，$[t_0-h,t_0]$ の間の変化を捉えたものである．この意味で (2.3.10) は「後退」方程式と呼ばれる．

- 同様に，(2.3.11) を導く際に用いたチャップマン・コルモゴロフ方程式は

  $$\mathbb{P}(X_{t_0+t+h}=j \mid X_{t_0}=i)=\sum_{k\in S}\mathbb{P}(X_{t_0+t+h}=j \mid X_{t_0+t}=k)\mathbb{P}(X_{t_0+t}=k \mid X_{t_0}=i)$$

  と表される．今度は推移確率を考える区間を $[t_0, t_0+t]$ から時間に関して前向きに $[t_0, t_0+t+h]$ へ延ばした形であり，この意味で (2.3.11) は「前進」方程式と呼ばれる．

(2.3.10)（あるいは (2.3.11)）は行列の線形常微分方程式であり，これの解は行列の指数関数により与えられることが知られている．一般に，$A\in\mathbb{R}^{|S|\times|S|}$ と $s_0>0$ に対して，級数 $\sum_{n=0}^{\infty}(sA)^n/(n!)$ は $s\in[-s_0,s_0]$ 上一様収束する．従って，行列値関数

$$\exp(sA):=\sum_{n=0}^{\infty}\frac{(sA)^n}{n!},\quad s\in\mathbb{R}$$

は定義可能であり，$\mathbb{R}$ 上連続である．さらに，各成分で微分を考えるとき，$s\mapsto\exp(sA)$ は $\mathbb{R}$ 上微分可能で，

$$\frac{d}{ds}\exp(sA)=A\exp(sA)=\exp(sA)A$$

である（微分方程式の教科書を参照のこと）．

定理 **2.3.10.** S を有限集合とする．このとき, 任意の推移確率関数 $P(\cdot)$ は方程式 (2.3.10) を満たし，

$$P(t)=\exp(tQ),\quad t\geq 0 \tag{2.3.12}$$

により与えられる．ここで Q は $P(\cdot)$ に付随する推移率行列である．

逆に，(2.3.7) と (2.3.9) を満たす $Q=\{q_{ij}\}_{i,j\in S}\in\mathbb{R}^{|S|\times|S|}$ に対し，(2.3.12) により定義される $P(\cdot)$ は Q を推移率行列とする推移確率関数である．

証明．常微分方程式の一般論より, (2.3.12) により定義される $P(\cdot)$ が初期条件 $P(0)=I_{|S|}$ を持つ (2.3.10) の一意解である．よって定理の前半の主張は直ちに従う．

次に，$Q=\{q_{ij}\}_{i,j\in S}\in\mathbb{R}^{|S|\times|S|}$ が (2.3.7) と (2.3.9) を満たすと仮定する．もし $Q=0$ なら $P(t)=I_{|S|}$ であり，これが推移確率関数であることは明らかである．よって $Q\neq 0$ とする．(2.3.9) より $Q(1,\ldots,1)^{\mathsf{T}}=0$. ここで，$x^{\mathsf{T}}$ はベクトル x の転置を表す．従って，各 n に対して $Q^n(1,\ldots,1)^{\mathsf{T}}=0$. これより $\exp(tQ)(1,\ldots,1)^{\mathsf{T}}=(1,\ldots,1)^{\mathsf{T}}$. これは定義 2.3.1(1) の成立を意味する．

$\exp(tQ)$ の各成分が非負であることを示すため，$a=\max\{-q_{ii}: i\in S\}$ を考える．このとき，(2.3.7) と (2.3.9) より $a\geq 0$ である．仮に $a=0$ とすると，$0=-q_{ii}=\sum_{j\neq i}q_{ij}$ となるため，各 i,j について $q_{ij}=0$, すなわち $Q=0$ となり最初の仮定に矛盾する．よって $a>0$ でなければばならない．これと a の定義より $\tilde{P}:=a^{-1}Q+I_{|S|}$ は非負行列である．Q を $\tilde{P}$ で表すことにより

$$\exp(tQ)=\exp(at(\tilde{P}-I_{|S|}))=\sum_{n=0}^{\infty}\frac{(at)^n(\tilde{P})^n}{n!}e^{-at}. \tag{2.3.13}$$

ここで，一般に $A, B \in \mathbb{R}^{|S|\times|S|}$ が $AB = BA$ を満たすなら

$$\exp(A+B) = \exp(A)\exp(B) \tag{2.3.14}$$

が成り立つこと，および $\exp(-atI_{|S|}) = e^{-at}I_{|S|}$ を用いた．(2.3.13) より $\exp(tQ)$ は非負行列である．

また，(2.3.14) より定義 2.3.1(2) は直ちに従い，先に述べた指数関数の性質より定義 2.3.1(3) が満たされることも明らかである． □

S を一般の場合に戻して漸近的結果に話を移そう．

定義 **2.3.11.** $\{X_t\}_{t\geq 0}$ を推移確率関数 $P(\cdot) = \{p_{ij}(\cdot)\}_{i,j\in S}$ を持つマルコフ連鎖とする．

(1) S 上の確率測度 $\pi = \{\pi_i\}_{i\in S}$ は

$$\pi_j = \sum_{i\in S} \pi_i p_{ij}(t), \quad t > 0$$

を満たすとき，$\{X_t\}_{t\geq 0}$ の定常分布（stationary distribution）であるという．

(2) $\{X_t\}_{t\geq 0}$ は $i \neq j$ なる任意の $i, j \in S$ と任意の $t > 0$ に対して $p_{ij}(t) > 0$ を満たすとき，既約（irreducible）であるという．

$\{X_t\}_{t\geq 0}$ を推移確率関数 $P(\cdot) = \{p_{ij}(\cdot)\}_{i,j\in S}$ を持つマルコフ連鎖とし，Q をその推移率行列とする．このとき以下の 2 つが成り立つ（読者はこれらを確かめよ）:

- S が有限集合ならば，π が $\{X_t\}$ の定常分布であることと $\pi Q = 0$ が成り立つことは同値である．
- $\{X_t\}$ が既約であることと，$i \neq j$ なる任意の $i, j \in S$ に対して $i_1, \ldots, i_n \in S$ $(i_k \neq i_\ell)$ が存在して $q_{ii_1}q_{i_1i_2}\cdots q_{i_nj} > 0$ が成り立つことは同値である．

定理 **2.3.12.** S を有限集合とする．$\{X_t\}_{t\geq 0}$ を推移確率関数 $P(\cdot) = \{p_{ij}(\cdot)\}$ を持つ既約なマルコフ連鎖と仮定する．このとき，定常分布 π が一意的に存在し，

$$\lim_{t\to\infty} p_{ij}(t) = \pi_j, \quad i, j \in S.$$

証明. 定常分布 π の存在については証明略（[12] の系 6.3.1 参照）．それ以外の主張を示すために，まず，(2.3.13) より $p_{ii}(t) \geq e^{-at} > 0$ となることを注意しておく．よって，既約性より任意の $i, j \in S$, $t > 0$ に対して $p_{ij}(t) > 0$ である．ゆえに $h > 0$ を任意にとるとき，離散時間マルコフ連鎖 $\{X_{nh}\}_{n\geq 0}$ はエルゴード的であり，定常分布 π^h が一意に存在する．π 自身も $\{X_{nh}\}_{n\geq 0}$ の定常分布であるから，$\pi = \pi^h$ である．すなわち，π^h は h に依存しない．また，これより連続時間マルコフ連鎖の定常分布の一意性も従う．

次に，$P(t)$ は連続なので，$[0,1]$ 上で一様連続である．ゆえに，任意の $\varepsilon > 0$ に対して，ある $\delta \in (0,1)$ が存在し，$|s-u| \leq \delta$ なる $s, u \in [0,1]$ に対して $|p_{ij}(s) - p_{ij}(u)| \leq \varepsilon$ が各 $i, j \in S$ について成立する．よって特に，十分大きい任意の $t > 0$ に対して n を t/δ の整数部分とすると，$0 \leq t - n\delta < \delta$ となるので，$|p_{ij}(t-n\delta) - \delta_{ij}| \leq \varepsilon$. 他方，定理 1.7.2

より，ある n_0 が存在して，$N \geq n_0$ ならば $|p_{ij}(N\delta) - \pi_j| \leq \varepsilon$. 従って，$t > 0$ が十分大きければ $n \geq n_0$ とできるので，$|p_{ij}(n\delta) - \pi_j| \leq \varepsilon$ を得る．ゆえに，十分大きな $t > 0$ と各 $i, j \in S$ に対して

$$\begin{aligned}|p_{ij}(t) - \pi_j| &\leq |p_{ij}(t) - p_{ij}(n\delta)| + |p_{ij}(n\delta) - \pi_j| \\ &\leq \sum_{k \in S} |p_{ik}(t - n\delta) - \delta_{ik}| p_{kj}(n\delta) + |p_{ij}(n\delta) - \pi_j| \leq (|S| + 1)\varepsilon.\end{aligned}$$

これより定理の主張が従う． □

2.4 出生死滅過程

出生（しゅっしょう）死滅過程は生物の個体数の増減について研究するために考えだされたモデルであり，確率過程としては連続時間マルコフ連鎖である．個体数に制限を課す場合と課さない場合があり，前者は有限状態のマルコフ連鎖となるため前節の全ての結果が適用できるが，後者については個別の取扱いが（少なくとも本稿の枠組みでは）必要になる部分もある．本節で得られた結果を次節で扱う待ち行列の問題に応用する．

前節で述べたように，推移率行列を与える形で連続時間マルコフ連鎖を定義することができる（定理 2.3.9 参照）．出生死滅過程はこの方法で定義される．

定義 2.4.1. $S = \mathbb{N} \cup \{0\}$ とし，$\{\lambda_i\}_{i=0}^{\infty}$, $\{\mu_i\}_{i=0}^{\infty}$ を非負の実数列で $\mu_0 = 0$ を満たすものとする．$Q = \{q_{ij}\}_{i,j \in S}$ を

$$q_{ij} = \begin{cases} \lambda_i & (j = i + 1), \\ -(\lambda_i + \mu_i) & (j = i), \\ \mu_i & (j = i - 1), \\ 0 & (j \notin \{i - 1, i, i + 1\}) \end{cases}$$

により定義する．このとき，Q を推移率行列とする連続時間マルコフ連鎖を**出生死滅過程**（birth and death process）という．

- 推移率の定義より，

  $$\mathbb{P}(X_{t+h} = i + 1 \,|\, X_t = i) = \lambda_i h + o(h), \quad h \searrow 0$$

  となる．ここで，$o(h)$, $h \searrow 0$ は $\lim_{h \searrow 0} f(h)/h = 0$ を満たす関数 f を汎用的に表すことを思い出そう．よって，X_t を時刻 t での個体数とすると，λ_i は個体数が i のときの**出生率**（birth rate）を意味する．
- 同様に，

  $$\mathbb{P}(X_{t+h} = i - 1 \,|\, X_t = i) = \mu_i h + o(h), \quad h \searrow 0$$

  であるから，μ_i は個体数が i のときの**死滅率**（death rate）を表している．
- $\mu_0 = 0$ としているのは，個体数ゼロから死滅する現象を排除するためである．

命題 2.4.2. 出生死滅過程の推移確率関数 $P(\cdot)$ はコルモゴロフの後退方程式 (2.3.10) を満たす．

*証明**. $i,j\in S$, $t>0$ を固定する．前節と同様に，

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h}=\frac{p_{ii}(h)-1}{h}p_{ij}(t)+\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t) \tag{2.4.1}$$

と表し，右辺の極限について調べる．任意の $N\in\mathbb{N}$ に対して

$$\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\geq\sum_{\substack{k=0\\k\neq i}}^{N}\frac{p_{ik}(h)}{h}p_{kj}(t) \tag{2.4.2}$$

であり，右辺は有限和であるから $h\searrow 0$ のとき $\sum_{\substack{k=0\\k\neq i}}^{N}q_{ik}p_{kj}(t)$ に収束する．従って，(2.4.2) において $h\searrow 0$ としてから $N\to\infty$ とすることで，

$$\liminf_{h\searrow 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\geq\sum_{k\neq i}q_{ik}p_{kj}(t). \tag{2.4.3}$$

次に，$N>i$ に対して，

$$\begin{aligned}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)&\leq\sum_{\substack{k=0\\k\neq i}}^{N}\frac{p_{ik}(h)}{h}p_{kj}(t)+\sum_{k>N}\frac{p_{ik}(h)}{h}\\&=\sum_{\substack{k=0\\k\neq i}}^{N}\frac{p_{ik}(h)}{h}p_{kj}(t)+\frac{1-p_{ii}(h)}{h}-\sum_{\substack{k=0\\k\neq i}}^{N}\frac{p_{ik}(h)}{h}\end{aligned}$$

であるから，

$$\limsup_{h\searrow 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\leq\sum_{\substack{k=0\\k\neq i}}^{N}q_{ik}p_{kj}(t)-q_{ii}-\sum_{\substack{k=0\\k\neq i}}^{N}q_{ik}.$$

両辺 $N\to\infty$ とすれば，$\sum_k q_{ik}=0$ より，

$$\limsup_{h\searrow 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\leq\sum_{k\neq i}q_{ik}p_{kj}(t). \tag{2.4.4}$$

(2.4.3) と (2.4.4) より (2.4.1) の右辺第二項の極限が存在し，

$$\lim_{h\searrow 0}\frac{p_{ij}(t+h)-p_{ij}(t)}{h}=\sum_{k\in S}q_{ik}p_{kj}(t). \tag{2.4.5}$$

同じ議論により，

$$\lim_{h\searrow 0}\frac{p_{ij}(t-h)-p_{ij}(t)}{-h}=\sum_{k\in S}q_{ik}p_{kj}(t)$$

を示せるので，これと (2.4.5) より命題の主張が従う． □

状態空間が有限のときは，π が定常分布であることは $\pi Q = 0$ と同値であった．同様のことが出生死滅過程についても言える．

定理 **2.4.3.** $\sup_{i\in S}(\lambda_i+\mu_i)<\infty$ を仮定する．このとき，S 上の確率測度 π が $P(\cdot)$ の定常分布であるための必要十分条件は $\pi Q = 0$ が成り立つことである．

証明*. 最初に十分性を示す．命題 2.4.2 より，任意の $t>0$ に対して，

$$p_{ij}(t)=\delta_{ij}+\int_0^t\sum_{k\in S_i}q_{ik}p_{kj}(s)ds,\quad i,j\in S.$$

ただし $S_0=\{0,1\}$, $S_i=\{i-1,i,i+1\}$, $i\geq 1$. 今，$k\in S_i$ のとき $|q_{ik}p_{kj}(s)|\leq\lambda_i+\mu_i$ であり，$\sum_{i\in S}\pi_i(\lambda_i+\mu_i)<\infty$ を仮定したので，積分と和の交換ができき，

$$\begin{aligned}\sum_{i\in S}\pi_i p_{ij}(t)&=\sum_{i\in S}\pi_i\delta_{ij}+\sum_{i\in S}\pi_i\sum_{k\in S_i}\int_0^t q_{ik}p_{kj}(s)ds\\&=\pi_j+\int_0^t\sum_{i\in S}\sum_{k\in S_i}\pi_i q_{ik}p_{kj}(s)ds.\end{aligned}$$

ここで，$\sum_{i\in S}\sum_{k\in S_i}\pi_i q_{ik}p_{kj}(s)=\sum_{k\in S}\sum_{i\in S}\pi_i q_{ik}p_{kj}(s)$ であるから（これは絶対収束性から従うが，直接計算によっても確かめられる）．定理の仮定より，この等式の右辺は 0 であるから，結局 π は定常分布である．

逆に，π が $P(\cdot)$ の定常分布と仮定する．このとき，任意の $h>0$ と $j\in S$ に対して，

$$0=\sum_{i\in S}\pi_i\frac{p_{ij}(h)-\delta_{ij}}{h}. \tag{2.4.6}$$

今，命題 2.4.2 より，

$$\frac{p_{ij}(h)-\delta_{ij}}{h}=\frac{1}{h}\int_0^h\sum_{k\in S}q_{ik}p_{kj}(s)ds.$$

この等式と $q_{ii}\leq 0$, $q_{ik}\geq 0$, $i\neq k$ より

$$q_{ii}\leq\frac{1}{h}\int_0^h q_{ii}p_{ij}(s)ds\leq\frac{p_{ij}(h)-\delta_{ij}}{h}\leq\frac{1}{h}\int_0^h\sum_{k\neq i}q_{ik}ds=-q_{ii}.$$

すなわち，

$$\left|\frac{p_{ij}(h)-\delta_{ij}}{h}\right|\leq\lambda_i+\mu_i.$$

よって，定理の仮定より，確率空間 $(S,2^S,\pi)$ において優収束定理（定理 A.4.4）が適用できるので，(2.4.6) において $h\searrow 0$ とすれば，

$$0=\lim_{h\searrow 0}\sum_{i\in S}\pi_i\frac{p_{ij}(h)-\delta_{ij}}{h}=\sum_{i\in S}\pi_i\lim_{h\searrow 0}\frac{p_{ij}(h)-\delta_{ij}}{h}=\sum_{i\in S}\pi_i q_{ij}.$$

ゆえに $\pi Q = 0$ が従う． □

方程式 $\pi Q = 0$ は

$$\begin{cases} \lambda_0\pi_0 - \mu_1\pi_1 = 0, \\ \lambda_{i-1}\pi_{i-1} - (\lambda_i + \mu_i)\pi_i + \mu_{i+1}\pi_{i+1} = 0, \quad i \geq 1 \end{cases}$$

と書き表される．π_1 から帰納的に解いていくことにより，この方程式の解は

$$\pi_i = \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i}\pi_0, \quad i \geq 1 \tag{2.4.7}$$

により一意的に与えられることが分かる．よって，

$$\sum_{i=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i} < \infty \tag{2.4.8}$$

のとき，

$$\pi_0 = \left(1 + \sum_{i=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i}\right)^{-1} \tag{2.4.9}$$

と定義すれば，この $\{\pi_i\}_{i\in S}$ は確率分布になり，従って $P(\cdot)$ の定常分布である．

例 **2.4.4.** $\lambda_i \equiv \lambda > 0$, $\mu_i \equiv \mu > 0$ の場合を考える．このとき，$\rho := \lambda/\mu < 1$ ならば $1/\pi_0 = \sum_{i=0}^{\infty} \rho^i = 1/(1-\rho)$ であるから，定常分布は

$$\pi_i = (1-\rho)\rho^i, \quad i \in S$$

により与えられる．すなわち定常分布はパラメータ ρ の幾何分布である．

命題 **2.4.5.** 任意の $i \in S$ に対して $\lambda_i, \mu_{i+1} > 0$ ならば $P(\cdot)$ は既約である．

証明*. 命題 2.4.2 より，

$$\frac{dp_{ij}(t)}{dt} = -(\lambda_i + \mu_i)p_{ij}(t) + \mu_i p_{i-1,j}(t) + \lambda_i p_{i+1,j}(t), \quad i, j \in S.$$

これを，$p_{i-1,j}$, $p_{i+1,j}$ を既知として $p_{i,j}$ に関して解けば，

$$p_{ij}(t) = e^{-(\lambda_i+\mu_i)t}\delta_{ij} + \int_0^t e^{-(\lambda_i+\mu_i)(t-s)}(\mu_i p_{i-1,j}(s) + \lambda_i p_{i+1,j}(s))ds. \tag{2.4.10}$$

となる．ただし，$i = 0$ のときは右辺の第二項は現れない．これより，任意の $i \in S$, $t > 0$ に対して $p_{ii}(t) \geq e^{-(\lambda_i+\mu_i)t} > 0$. これと (2.4.10) を用いて，

$$p_{i,i+1}(t) \geq \lambda_i \int_0^t e^{-(\lambda_i+\mu_i)(t-s)} p_{i+1,i+1}(s)ds > 0,$$
$$p_{i,i-1}(t) \geq \mu_i \int_0^t e^{-(\lambda_i+\mu_i)(t-s)} p_{i-1,i-1}(s)ds > 0$$

を得る．今，任意の $m \in \mathbb{N}$ に対して

$$p_{i,i+m}(t) > 0, \quad p_{j,j-m}(t) > 0, \quad i \geq 0, \ j \geq m, \ t > 0 \tag{2.4.11}$$

が成り立つことを帰納法により示そう．これにより命題の証明は完了する．$m=1$ の場合に成り立つことは上で確認したので，一般の m に対して (2.4.11) の成立を仮定する．このとき，$m=1$ のときと同様に，(2.4.10) より，

$$
\begin{aligned}
p_{i,i+m+1}(t) &\geq \lambda_i \int_0^t e^{-(\lambda_i+\mu_i)(t-s)} p_{i+1,i+1+m}(s)ds > 0, \quad i \geq 0,\ t > 0, \\
p_{i,i-m-1}(t) &\geq \mu_i \int_0^t e^{-(\lambda_i+\mu_i)(t-s)} p_{i-1,i-1-m}(s)ds > 0, \quad i \geq m+1,\ t > 0
\end{aligned}
$$

となる．よって $m+1$ に対しても (2.4.11) が成り立つ． □

ある番号から先の出生率と死滅率をゼロにすることで個体数が一定の範囲内に収まるような出生死滅過程を定義できる．途中の出生率と死滅率が常に正のとき，命題の証明の議論により，この出生死滅過程が既約であることが確かめられる．このとき，状態空間が有限集合とみなせるので，前節の定理 2.3.12 より次の結果が得られる:

定理 **2.4.6.** $N \in \mathbb{N}$ とし，

$$
\begin{aligned}
\lambda_i > 0, \quad i = 0, \ldots, N, \quad \lambda_i = 0, \quad i \geq N, \\
\mu_i > 0, \quad i = 1, \ldots, N, \quad \mu_i = 0, \quad i \geq N+1
\end{aligned}
$$

を仮定する．このとき，$P(\cdot)$ の定常分布を $\{\pi_i\}_{i\in S}$ とすると，

$$
\lim_{t\to\infty} p_{ij}(t) = \pi_j, \quad i, j \in \{0, 1, \ldots, N\}.
$$

2.5 待ち行列

待ち行列（queue, waiting line）とは，あるサービスを受けるために並んでいる人や物のことをいう．例えば，映画館や美術館のチケットブースでチケットの購入のために並んでいる人がそれに当たるであろう．**待ち行列理論**（queueing theory）では，待ち行列を数理モデル化し，待ち現象の解析やサービス提供のためのシステム設計・評価を行う．

待ち行列は，電話回線，通信，交通渋滞など人間以外の待ち現象をも指すが，本稿では，上述のようなチケット購入のために並ぶ人を想定して説明しよう．この場合，人はチケットを手に入れるために所定の場所に行く．このことを抽象化し，チケットを手に入れるということを「サービスを受ける」と呼び，サービスが提供される場所を「窓口」，窓口に向かう人を「客」，客が待ち行列に加わることを「到着」と呼ぶ．もし待ち人数がゼロであれば直接窓口に向かうことになるが，これも到着とみなす．

待ち行列モデルは

(1) 到着時間間隔分布: 客が到着する時間間隔の確率分布，

(2) サービス時間分布: 客 1 人に要するサービス時間の確率分布，

(3) 窓口の数: 同時にサービスを受けることのできる客数，

(4) 待ち行列の容量: 窓口の数と待ち行列に実際に並ぶことのできる人数の和，

(5) サービス方法: サービスの提供が先着順（FCFS: First Come, First Served）なのか後着順（LCFS: Last Come, First Served）なのかなど，

により表現する．一般に，これらは**ケンドールの記号**を用いて示される．ケンドールの記号は

$$A/B/c/K/Z$$

という形で待ち行列モデルを特定する．ここで A は到着時間間隔分布，B はサービス時間分布，c は窓口の数，K は待ち行列の容量，Z はサービス方法を表す．例えば，$M/M/1/\infty/$FCFS は客はポアソン過程に従って到着し，サービス時間は指数分布，単一窓口，待ち行列の容量は無限で先着順にサービスされるようなモデルである．これを $M/M/1$ と略記することも多い．

以後，待ち行列モデルは M/M 型として，「待ち行列の長さと一人当たりのサービス時間は独立である」ことを仮定する．すなわち，たくさん並んでいるからサービス時間を短縮する，というようなことは考慮せず話を進める．さらに，

- $A(i)$: 系内客数 i のときの到着時間間隔 = 状態 i から $i+1$ へ遷移するまでの時間
- $D(i)$: 系内客数 i のときのサービス時間 = 状態 i から $i-1$ へ遷移するまでの時間

を考える．ここで系内客数とは窓口にいる客と待ち行列に並んでいる客の合計である．このとき，各 $A(i), D(i)$ は指数分布に従い各々独立と仮定する．そのパラメータをそれぞれ λ_i, μ_i とする．さらに，時刻 t における系内客数を X_t により表すと，

$$\begin{aligned}
\mathbb{P}(X_{t+h} = i+1 \,|\, X_t = i) &= \mathbb{P}(A(i) \le h, A(i) < D(i)) + o(h) \\
&= \int\int_{\{x \le h, 0 < x < y\}} \lambda_i \mu_i e^{-\lambda_i x} e^{-\mu_i y} dxdy + o(h) \\
&= \int_0^h \int_x^\infty \lambda_i \mu_i e^{-\lambda_i x} e^{-\mu_i y} dydx + o(h) \\
&= \frac{\lambda_i}{\lambda_i + \mu_i}(1 - e^{-(\lambda_i + \mu_i)h}) + o(h) \\
&= \lambda_i h + o(h) \quad (h \searrow 0).
\end{aligned}$$

ここで $1 - e^{-x} = x + o(x)$ $(x \to 0)$ を用いた．同様にして，

$$\begin{aligned}
&\mathbb{P}(X_{t+h} = i-1 \,|\, X_t = i) = \mu_i h + o(h), \\
&\mathbb{P}(X_{t+h} = i \,|\, X_t = i) = 1 - (\lambda_i + \mu_i)h + o(h) \quad (h \searrow 0).
\end{aligned}$$

従って，$\{X_t\}$ がマルコフ性を持てばこれは出生死滅過程である．待ち行列理論では，出生率 λ_i を到着率，死滅率 μ_i をサービス率と呼ぶ．

- ここでは，待ち行列モデルが M/M 型であることに加えて系内客数にマルコフ性を考慮したが，実は M/M 型であれば必然的に $\{X_t\}$ はマルコフ連鎖となることが知られている．これが指数分布を用いる理由でもある．詳細については例えば [12] の第 6 章を参照のこと．

前節で示したように，出生死滅過程は (2.4.8) を満たすとき定常分布 $\{\pi_i\}_{i\in S}$ を持ち，それは (2.4.7) と (2.4.9) により与えられる．分布が $\{\pi_i\}$ に従う確率変数を X とすると，これは X_t の $t\to\infty$ における定常状態を表す確率変数とみなせる．

$M/M/1$ モデル

$\lambda>0$ と $\mu>0$ をとり，到着率とサービス率がそれぞれ

$$\lambda_i=\lambda,\quad \mu_{i+1}=\mu,\quad i\in S$$

により与えられる場合を考える．このとき，到着率は客数に依存せず $S=\{0,1,\cdots\}$ 上一定であるから，このことにより待ち行列の容量が無限であることが表現できる．また，一般的に考えれば，窓口全体の処理能力は窓口の数とそこにいる客数に応じて決まるものであり，今は系内客数が一人減るための率を一定と仮定していることから，窓口は一つとみなせる．すなわち，この場合の出生死滅過程は $M/M/1$ モデルを表現している．

例 2.4.4 より，$\lambda<\mu$ のとき，定常分布，すなわち X の分布はパラメータ $\rho=\lambda/\mu$ の幾何分布である．今，$\mathbb{E}[A(i)]=1/\lambda, \mathbb{E}[D(i)]=1/\mu$ であるから，

$$\rho=\frac{\text{平均サービス時間}}{\text{平均到着時間間隔}}$$

と表される．また，

$$\mathbb{P}(X\geq 1)=\rho,\quad \mathbb{P}(X=0)=1-\rho$$

である．言うなれば，将来的には，ある客が訪れた時に窓口が塞がっている確率が ρ であり，空いている確率が $1-\rho$ である．この意味で ρ を利用率という．

次に，定常状態における

- 平均系内客数 L
- 平均滞在時間 W

を考える．$L=\mathbb{E}[X]$ であるから，

$$L=\frac{\rho}{1-\rho}.$$

よって，平均系内客数は利用率が 1 に近づくにつれて（急激に）増加する．

γ を定常状態において客が訪れた時の待ち時間とすると，平均滞在時間は平均待ち時間と平均サービス時間の和であるから $W=\mathbb{E}[\gamma]+1/\mu$ である．他方，定常状態で j 人にサービスする時間の合計を γ_j とすると（$\gamma_0=0$ とおく），$\gamma=\sum_{j=0}^{\infty}1_{\{X=j\}}\gamma_j$ と表される．$j\geq 1$ のとき，γ_j はパラメータ μ の指数分布に従う j 個の独立な確率変数の和であ

り，上で課した待ち行列の長さとサービス時間の独立性と (2.1.2) から，

$$
\begin{aligned}
\mathbb{P}(\gamma \le t) &= \sum_{j=0}^{\infty} \mathbb{P}(\gamma_j \le t, X = j) = \pi_0 + \sum_{j=1}^{\infty} \pi_j \mathbb{P}(\gamma_j \le t) \\
&= \pi_0 + \sum_{j=1}^{\infty} \pi_j \int_0^t \frac{\mu(\mu x)^{j-1}}{(j-1)!} e^{-\mu x} dx
\end{aligned}
$$

となる．π_j の具体形を代入し，積分計算を進めることにより

$$
\begin{aligned}
\mathbb{P}(\gamma \le t) &= 1 - \rho + \sum_{j=1}^{\infty} (1-\rho)\rho^j \int_0^t \frac{\mu(\mu x)^{j-1}}{(j-1)!} e^{-\mu x} dx \\
&= 1 - \rho + (1-\rho)\mu\rho \int_0^t \sum_{j=1}^{\infty} \frac{(\mu\rho x)^{j-1}}{(j-1)!} e^{-\mu x} dx \\
&= 1 - \rho + (1-\rho)\mu\rho \int_0^t e^{-\mu(1-\rho)x} dx \\
&= 1 - \rho e^{-\mu(1-\rho)t}
\end{aligned}
$$

を得る．従って，

$$
W = \mathbb{E}[\gamma] + \frac{1}{\mu} = \frac{1}{\mu(1-\rho)}.
$$

よって平均滞在時間も利用率が 1 に近づくにつれて急激に増加する．また，これより

$$
L = \lambda W. \tag{2.5.1}
$$

関係式 (2.5.1) をリトルの公式という．

$M/M/1/N$ モデル

$N \ge 2$, $\lambda > 0$, $\mu > 0$ とし，到着率とサービス率がそれぞれ

$$
\lambda_i = \begin{cases} \lambda & (0 \le i \le N-1), \\ 0 & (i \ge N), \end{cases} \qquad \mu_i = \begin{cases} \mu & (1 \le i \le N), \\ 0 & (i \ge N+1) \end{cases}
$$

により与えられる場合を考えよう．この場合，系内客数が N になれば新たな到着は発生せず，客の出発のみが起こると考えることができる．よって，この出生死滅過程は $M/M/1/N$ モデルを表現している．定常分布 $\{\pi_i\}$ は，各 $i = 0, 1, \ldots, N$ について

$$
\pi_i = \begin{cases} \dfrac{(1-\rho)\rho^i}{1-\rho^{N+1}} & (\rho \ne 1), \\ \dfrac{1}{N+1} & (\rho = 1) \end{cases}
$$

により与えられる．よって平均系内客数 $L = \mathbb{E}[X]$ は

$$
L = \begin{cases} \dfrac{\rho}{1-\rho} - \dfrac{(N+1)\rho^{N+1}}{1-\rho^{N+1}} & (\rho \ne 1), \\ \dfrac{N+1}{2} & (\rho = 1) \end{cases}
$$

となる．

- $M/M/1/N$ モデルにおいても，リトルの公式 $L = \lambda W$ が成り立つことが知られている．さらに言えば，リトルの公式はモデルの仮定がほとんど無い状況でも成り立つ．詳しくは [12] の第 2 章を参照のこと．

生産管理への応用

待ち行列理論の応用として，ある製品の生産管理について考えよう．問題を単純化し，顧客による製品の注文と，生産後の在庫，製品の受渡しの 3 点のみを考える．**受注生産方式**では，顧客の注文を受けてから先着順に生産する．この場合，製品の注文が待ち行列モデルにおける客の到着に当たり，製品の受渡しは窓口でのサービスとみなせる．受渡しを待っている注文の数を系内客数とし，この待ち行列を $M/M/1$ モデルと仮定すると，先の結果が適用でき，

$$\text{注文後，すぐに生産されない確率} = \mathbb{P}(X \geq 1) = \rho,$$
$$\text{注文後，製品を受け取るまでの平均時間} = \text{平均滞在時間} = \frac{1}{\mu(1-\rho)}$$

と見積もることができる．

見込み生産方式では，製品の需要を想定してあらかじめ生産しておき，注文に応じて在庫から出荷する．この方法では，もちろん，顧客の待ち時間は短くなるが，製品の保管コストとのトレードオフが発生する．**かんばん方式**は，この点を考慮した生産管理法である．在庫の各製品には「かんばん」と呼ばれる記録書が付与されており，納品書として機能する．製品の受渡し時には，かんばんは外さる．外されたかんばんは，製造部門にまわされ，新しい製品の生産指示書として役割を変える．この方式により，在庫が減ったときのみ製造が行われ，かんばんの数によって在庫数がコントロールされることになる．

かんばん方式を待ち行列モデルで解析してみよう．もし注文時に在庫があれば，製品の受渡し後，かんばんの到着によって生産が開始される．在庫が無いときには，受注生産と同様に生産を待つことになる．よって，「生産指示かんばん」の数と在庫が無いときの注文数の総和，すなわち生産待ちの注文数を待ち行列モデルにおける系内客数とみなそう．受注生産の場合と同様に，生産待ち注文の到着はポアソン過程に従うと仮定する．さらに，窓口のサービスをやはり製品の受渡しとみなすことで，結局この待ち行列を $M/M/1$ モデルと仮定することができる．今，定常状態の生産待ち注文数を X，かんばんの数を z とすると，

- 在庫数 $= (z - X)_+ =: I$,
- 在庫切れ時の注文数 $= (X - z)_+ =: B$,
- 生産指示かんばんの数 $= \min\{X, z\} =: C$

となる．これより，

$$\text{平均在庫数} = \mathbb{E}[I] = \sum_{i=0}^{z}(z-i)_+(1-\rho)\rho^i = (1-\rho)\rho^z\sum_{n=1}^{z} n\rho^{-n} = z - \frac{\rho(1-\rho^z)}{1-\rho}$$

となり，かんばん数と在庫数は平均的には概ね同数であることが分かる．さらに，

$$\text{在庫切れの確率} = \mathbb{P}(I=0) = \mathbb{P}(X \geq z) = \rho^z.$$

すなわち，かんばん数を増やすことで在庫も増えるので在庫切れになる確率は小さくなる．しかし，利用率 ρ がそれ程高くない時には，かんばん数を一定以上増やしても在庫切れ確率に大きな影響を与えない．実際，$\rho = 0.1$ の時には $z = 4$ 程度で在庫切れ確率は十分小さい．

もし在庫切れの時には注文がキャンセルされてしまう場合には，かんばん方式は $M/M/1/z$ モデルとなる．この場合，系内客数は生産指示かんばん数に一致し，これが全体のかんばん数に等しいときには新たな到着が発生しないからである．従って，

$$\text{注文がキャンセルされる確率} = \mathbb{P}(X=z) = \begin{cases} \dfrac{(1-\rho)\rho^z}{1-\rho^{z+1}} & (\rho \neq 1), \\ \dfrac{1}{z+1} & (\rho = 1) \end{cases}$$

となる．

2.6 ブラウン運動

ブラウン運動とは，1827 年に英国の植物学者ロバート・ブラウンが発見した，水に浮かんだ花粉中の微粒子の不規則運動を数理モデル化したものである．現在では，このような現象の解釈からは離れ，様々な数理解析の分野において用いられている．例としては，確率微分方程式に代表されるような，ランダムさを持つ動的システムにおける不確実性の源泉としての役割が挙げられる．

初めに，$S_n = \sum_{k=1}^n X_k$ $(n \geq 0)$ を 0 から出発する 1 次元対称ランダム・ウォークとする．小さい時間区間で変動させるため，$\sqrt{n}$ で標準化した上で

$$W_0^{(n)} := \frac{1}{\sqrt{n}}S_0 = 0,\ \ W_{1/n}^{(n)} := \frac{1}{\sqrt{n}}S_1,\ \ W_{2/n}^{(n)} := \frac{1}{\sqrt{n}}S_2, \cdots$$

とおき，これらの線形補間を $W_t^{(n)}$ とする．すなわち，

$$W_t^{(n)} = \frac{1}{\sqrt{n}}\left[S_{\lfloor nt\rfloor} + X_{\lfloor nt\rfloor+1}(nt - \lfloor nt\rfloor)\right], \quad t \geq 0.$$

$W_t^{(n)}$ の何らかの意味での極限 $n \to \infty$ で得られる関数について考えよう．まず，$\{W_t^{(n)}\}$ の有限次元分布については以下のことが分かる．

命題 **2.6.1.** $0 = t_0 \le t_1 \le t_2 \le \cdots \le t_m$ とするとき，$(m+1)$ 次元確率変数 $(W_{t_0}^{(n)}, W_{t_1}^{(n)}, \ldots, W_{t_m}^{(n)})$ は以下の性質をもつ $(m+1)$ 次元確率変数 $(W_{t_0}, W_{t_1}, \ldots, W_{t_m})$ に分布収束する:

(1) $W_{t_0} = 0$ a.s.

(2) $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_m} - W_{t_{m-1}}$ は独立.

(3) 各 $W_{t_k} - W_{t_{k-1}}$ は平均 0 分散 $t_k - t_{k-1}$ の正規分布に従う.

証明. $m = 2$ の場合を考える．一般の m の場合も煩雑だが同様に示せる．$s = t_1$, $t = t_2$ とおく.

$$\left| W_t^{(n)} - \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor} \right| \le \frac{1}{\sqrt{n}}$$

であるから，明らかに

$$\left| (W_s^{(n)}, W_t^{(n)}) - \frac{1}{\sqrt{n}} (S_{\lfloor sn \rfloor}, S_{\lfloor tn \rfloor}) \right| \to 0 \quad \text{a.s.}$$

従って，分布収束の意味で

$$\frac{1}{\sqrt{n}} \left(\sum_{j=1}^{\lfloor sn \rfloor} X_j, \sum_{j=1}^{\lfloor tn \rfloor} X_j \right) \to (W_s, W_t) \tag{2.6.1}$$

が成り立つことを示せばよい．i を虚数単位とし，$\alpha, \beta \in \mathbb{R}$ とするとき，$\{\xi_j\}$ の独立同分布性より

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(i\alpha \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} X_j + i\beta \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor tn \rfloor} X_j \right)\right] \\
&= \mathbb{E}\left[\exp\left(i(\alpha - \beta) \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} X_j + i\beta \frac{1}{\sqrt{n}} \sum_{j=\lfloor sn \rfloor + 1}^{\lfloor tn \rfloor} X_j \right)\right] \\
&= \mathbb{E}\left[\exp\left(i(\alpha - \beta) \sqrt{\frac{\lfloor sn \rfloor}{n}} \frac{1}{\sqrt{\lfloor sn \rfloor}} \sum_{j=1}^{\lfloor sn \rfloor} X_j \right)\right] \\
&\qquad \times \mathbb{E}\left[\exp\left(i\beta \sqrt{\frac{\lfloor tn \rfloor - \lfloor sn \rfloor}{n}} \frac{1}{\sqrt{\lfloor tn \rfloor - \lfloor sn \rfloor}} \sum_{j=1}^{\lfloor tn \rfloor - \lfloor sn \rfloor} X_j \right)\right].
\end{aligned}$$

今，$(sn-1)/n \le \lfloor sn \rfloor / n \le s$ であるから $\lfloor sn \rfloor / n \to s$．また，中心極限定理より $\frac{1}{\sqrt{\lfloor sn \rfloor}} \sum_{j=1}^{\lfloor sn \rfloor} X_j$ は標準正規分布に従う確率変数に分布収束するから，

$$\begin{aligned}
\mathbb{E}\left[\exp\left(i\alpha \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} X_j + i\beta \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor tn \rfloor} X_j \right)\right] &\to \mathbb{E}[e^{i(\alpha-\beta)W_s}] \mathbb{E}[e^{i\beta(W_{t-s})}] \\
&= \mathbb{E}[e^{i\alpha W_s + i\beta W_t}].
\end{aligned}$$

よって特性関数の収束により (2.6.1) が従う. □

ゆえに，命題 2.6.1 の (1)-(3) をみたす確率過程 $\{W_t\}$ は $\{W_t^{(n)}\}$ の確率過程としてのある種の極限とみなせる．$\{W_t\}$ のような確率過程をブラウン運動と呼び，不規則挙動の数学的に理想化されたモデルとして用いる．

定義 **2.6.2.** （実数値）確率過程 $\{W_t\}_{t\geq 0}$ は以下の条件を満たすとき，**ブラウン運動**（Brownian motion）であるという:

(1) $W_0 = 0$ a.s.
(2) $t \mapsto W_t$ は確率 1 で連続．
(3) 独立増分性: $0 = t_0 < t_1 < \cdots < t_m$ に対し，確率変数列 $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_m} - W_{t_{m-1}}$ は独立．
(4) 定常増分性: $s \leq t$ に対し，$W_t - W_s$ は，平均 0 分散 $t - s$ の正規分布に従う．

ブラウン運動の解析を進める前に，命題 2.6.1 では，ブラウン運動の存在は保証できていないことに注意してほしい．もし存在するとすれば有限次元分布が一致するということを主張しているのみである．しかし，ここではブラウン運動の存在問題に立ち入らず，結果のみを述べる．

定理 **2.6.3.** 標準ブラウン運動は適当な確率空間上に存在する．

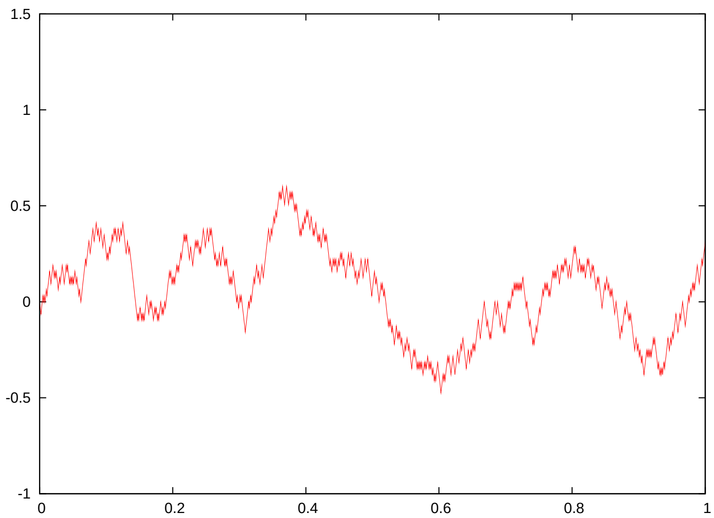


図 2.6.1　ブラウン運動の標本路．

- ブラウン運動を Wiener 過程ともいう．
- d 次元確率過程 $W_t = (W_t^1, \ldots, W_t^d)$, $t \geq 0$, は各成分がブラウン運動で互いに独立のとき，d 次元ブラウン運動とよぶ．

次に，ブラウン運動の標本路の不規則性についてみてみよう．

定理 **2.6.4.** ブラウン運動の標本路は各点について確率 1 で微分不可能である．すなわち，$\{W_t\}$ をブラウン運動とするとき，任意の $s \geq 0$ に対して $\mathbb{P}(\{\omega \in \Omega : t \mapsto W_t(\omega)$ は s で微分不可能 $\}) = 1$.

証明．まず $s \geq 0$ を固定し $A_s = \{\omega : s \geq 0$ で $t \mapsto W_t(\omega)$ は微分可能 $\}$ とおく．よって $\omega \in A_s$ とすると，極限 $\lim_{h \searrow 0}(W_{s+h}(\omega) - W_s(\omega))/h$ が存在して有限値である．特に，$|W_{s+h}(\omega) - W_s(\omega)|/h \leq \delta$ $(0 < \forall h < h_0)$ を満たす $\delta > 0$ と $h_0 > 0$ が存在する．よって $\sup_{n\geq 1} n|W_{s+1/n}(\omega) - W_s(\omega)| < \infty$ であり，これより，ある $N \geq 1$ が存在して，各 $n \geq 1$ に対して $n|W_{s+1/n}(\omega) - W_s(\omega)| \leq N$. これは

$$A_s \subset \cup_{N\geq 1} \cap_{n\geq 1} \{n|W_{s+1/n} - W_s| \leq N\}$$

を意味する．事象の単調性より

$$\mathbb{P}(A_s) \leq \lim_{N\to\infty} \inf_{n\geq 1} \mathbb{P}(n|W_{s+1/n} - W_s| \leq N)$$

を得る．今，$W_{s+1/n} - W_s \sim N(0, 1/n)$ であるから，$\xi \sim N(0,1)$ とするとき，

$$\inf_{n\geq 1} \mathbb{P}(n|W_{s+1/n} - W_s| \leq N) = \inf_{n\geq 1} \mathbb{P}(n\sqrt{1/n}|\xi| \leq N) = \inf_{n\geq 1} \mathbb{P}(|\xi| \leq Nn^{-1/2}) = 0.$$

従って，$\mathbb{P}(A_s) = 0$ である． □

- この命題よりブラウン運動の標本路は非常に粗く変動していることが分かる．
- 実は，ブラウン運動の標本路は確率 1 で各点について微分不可能となることも示せる．証明については例えば [16, 定理 1.2.2] を参照のこと．

別の尺度を用いてブラウン運動の標本路の不規則性を示そう．各 $t > 0$ に対し

$$V_W([0,t]) := \sup_{k\geq 0} \sup_{\pi} \sum_{i=0}^{k} |W_{t_{i+1}} - W_{t_i}|$$

と定義する．これは区間 $[0,t]$ 上の $\{W_t\}$ の**全変動**（total variation）と呼ばれている．ここで 2 つ目の sup は k 個の分点をもつ $[0,t]$ の分割 $\pi : 0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = t$ 全体にわたってとる．これはすなわち，ある地点から別の地点までの道における高低差の合計を表す．全変動が大きいほど，アップダウンのきつい道路であるといえる．

定理 **2.6.5.** ブラウン運動の標本路の全変動は確率 1 で無限大である．すなわち，$\mathbb{P}(\forall t > 0,\ V_W([0,t]) = \infty)$.

証明*．まず，$[0,t]$ 上の任意の分割 π に対して

$$\mathbb{E}\sum_{t_i\in\pi}(W_{t_{i+1}} - W_{t_i})^2 = \sum_{t_i\in\pi}(t_{i+1} - t_i) = t$$

であることに注意する．$Z_i = (W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)$ とおき，$\xi \sim N(0,1)$ とすると，Z_i たちは互いに独立で，それぞれは $(\xi^2 - 1)(t_{i+1} - t_i)$ と同分布である．よって

$$\mathbb{E}\left[\left(\sum_{t_i\in\pi}(W_{t_{i+1}} - W_{t_i})^2 - t\right)^2\right] = \mathbb{E}\sum_{t_i\in\pi} Z_i^2 = \mathbb{E}[(\xi^2 - 1)^2]\sum_{t_i\in\pi}(t_{i+1} - t_i)^2.$$

分割列 π_n を $\Delta_n := \sup_{t_i\in\pi_n}|t_{i+1}-t_i| \to 0$ を満たすようにとれば，上式の右辺は $t\mathbb{E}[(\xi^2-1)^2]\Delta_n$ で押さえられる．従って，確率収束の意味で，

$$Q_n := \sum_{t_i\in\pi_n}(W_{t_{i+1}}-W_{t_i})^2 \to t,\ \ n\to\infty.$$

よって概収束する部分列 Q_{n_k} が存在する．ゆえに，仮に $\mathbb{P}(V_W([0,t])<\infty)>0$ とすると，ブラウン運動の標本路の連続性より $\sup_{t_i\in\pi_{n_k}}|W_{t_{i+1}}-W_{t_i}|\to 0$ であるから

$$t \le \lim_{k\to\infty}\left(\sup_{t_i\in\pi_{n_k}}|W_{t_{i+1}}-W_{t_i}|\right)\sum_{t_i\in\pi_{n_k}}|W_{t_{i+1}}-W_{t_i}| = 0$$

なる事象の確率も正となる．しかしもちろんこれは $t>0$ と矛盾する．よって，任意の $t>0$ に対して $\mathbb{P}(V_W([0,t])=\infty)$．さらに，任意の $t>0$ に対し $s<t$ なる有理数 s を考えれば $V_W([0,s])\le V_W([0,t])$ を得るので，

$$1=\mathbb{P}(\forall s\in\mathbb{Q}\cap(0,\infty),\ V_W([0,s])=\infty)\le\mathbb{P}(\forall t>0,\ V_W([0,t])=\infty).$$

ゆえに定理の結論が従う．　□

- 上の定理の証明より，任意の分割 $0=t_0<t_1<\cdots<t_n<t_{n+1}=t$ に対して $\Delta_n=\sup|t_{i+1}-t_i|\to 0$ のとき，確率収束の意味で

$$\langle W\rangle_t := \lim\sum_{i=0}^{n}(W_{t_{i+1}}-W_{t_i})^2 = t$$

  が成り立つ．
- $\langle W\rangle_t,\ t\ge 0$ を $\{W_t\}$ の**二次変分**（quadratic variation）という．すなわち，ブラウン運動の一次の変分は発散するが，二次のレベルでは有限になるのである．この性質が確率積分の定義に生かされる．

次にブラウン運動のマルコフ性について考える．マルコフ連鎖の場合と異なり，ブラウン運動は確率密度を持つので（後述），マルコフ性の定義を修正する必要がある．しかしここでは厳密な定義をすることは避け，直感的に考えてみよう．マルコフ連鎖の場合，マルコフ性とは，粗っぽく言えば，過去の履歴による条件付き確率がその直前の条件にのみ依存することであった．この意味で，ブラウン運動もマルコフ性を持つといえる．実際，$t\ge s$ に対して

$$W_t = W_s + W_t - W_s$$

と表されるので，W_t-W_s が時刻 s 以前の値の情報と独立であれば，W_t の予測のためには s 時点の情報のみが有効であり，それより前の履歴は関係ないことになる．

定理 **2.6.6** (マルコフ性)**.** $0\le s_1<s_2<\cdots<s_m<s<t$ とするとき，W_t-W_s は $W_{s_1},\dots,W_{s_m},W_s$ と独立である．

証明. $\xi_1, \ldots, \xi_{m+1} \in \mathbb{R}$ に対して，

$$\xi_{m+1} W_s + \sum_{j=1}^{m} \xi_j W_{s_j} = \sum_{j=1}^{m+1} \sum_{k=j}^{m+1} \xi_k (W_{s_j} - W_{s_{j-1}}).$$

ただし $s_{m+1} = s$, $s_0 = 0$. これとブラウン運動の独立増分性より，i を虚数単位，$\xi_{m+2} \in \mathbb{R}$ として

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(i\xi_{m+2}(W_t - W_s) + \xi_{m+1} W_s + \sum_{j=1}^{m} \xi_j W_{s_j}\right)\right] \\
&= \mathbb{E}[e^{i\xi_{m+2}(W_t - W_s)}] \prod_{j=1}^{m+1} \mathbb{E}\left[\exp\left(i \sum_{k=j}^{m+1} \xi_k (W_{s_j} - W_{s_{j-1}})\right)\right] \\
&= \mathbb{E}[e^{i\xi_{m+2}(W_t - W_s)}] \mathbb{E}\left[\exp\left(\sum_{j=1}^{m+1} \sum_{k=j}^{m+1} \xi_k (W_{s_j} - W_{s_{j-1}})\right)\right] \\
&= \mathbb{E}[e^{i\xi_{m+2}(W_t - W_s)}] \mathbb{E}\left[\exp\left(\xi_{m+1} W_s + \sum_{j=1}^{m} \xi_j W_{s_j}\right)\right].
\end{aligned}$$

従って定理 A.3.5 より結論を得る. □

正規分布に従う確率変数が一意でないのと同様に，ブラウン運動も一意ではなく，例えば次が成り立つ:

命題 2.6.7. $\{W_t\}_{t\geq 0}$ をブラウン運動とするとき以下の (1)–(3) の確率過程は全てブラウン運動である．ただし $s > 0$, $c \neq 0$ とする．

(1) $\{-W_t\}_{t\geq 0}$.
(2) $\{W_{t+s} - W_s\}_{t\geq 0}$.
(3) $\{cW_{t/(c^2)}\}_{t\geq 0}$.

証明は読者の演習問題とする．

次に，ブラウン運動の分布の性質についてみていこう．$s \geq 0$, $x \in \mathbb{R}$ に対し

$$W_t^{s,x} := x + W_t - W_s, \quad t \geq s$$

は，言わば時刻 s で x から出発するブラウン運動である．このとき，$W_t^{s,x}$ の確率密度関数

$$p(t, y \,|\, s, x) := \frac{\partial}{\partial y} \mathbb{P}(W_t^{s,x} \leq y), \quad t \geq s, \ \ y \in \mathbb{R}$$

を (s,x) から (t,y) への**推移密度**（transition density）と呼ぶ．これは 2 階放物型偏微分方程式

$$\frac{\partial}{\partial t} p = \frac{1}{2} \frac{\partial^2}{\partial y^2} p, \tag{2.6.2}$$

$$\frac{\partial}{\partial s} p = \frac{1}{2} \frac{\partial^2}{\partial x^2} p, \tag{2.6.3}$$

を満たす．(2.6.2) をコルモゴロフの前進方程式，(2.6.3) をコルモゴロフの後退方程式という．

$p_t(x,y) = p(t,y\,|\,0,x)$ とおき，$f \in C_b^2(\mathbb{R})$ に対して

$$P_t f(x) := \begin{cases} \displaystyle\int_{-\infty}^{\infty} p_t(x,y) f(y) dy & (t > 0), \\ f(x) & (t = 0) \end{cases}$$

を考える．ここで，$C_b^2(\mathbb{R})$ は $\mathbb{R}$ 上 2 回偏微分可能で各偏導関数が有界連続となる関数全体である．このとき，

$$\mathbb{E}[f(W_{s+t}^{s,x})] = \mathbb{E}[f(x + W_t)] = P_t f(x)$$

であり，チャップマン・コルモゴロフ方程式

$$P_{t+s} = P_t P_s = P_s P_t, \quad t, s \geq 0$$

が成り立つ．これより，

$$\frac{d}{dt} P_t = \lim_{h \searrow 0} \frac{1}{h}(P_{t+h} - P_t) = P_t \mathcal{A} = \mathcal{A} P_t$$

が導かれる．ここで，

$$\mathcal{A} := \lim_{h \searrow 0} \frac{1}{h}(P_h - I)$$

であり，I は $C_b^2(\mathbb{R})$ 上の恒等作用素とし，極限は作用させた関数の各点で考えるものとする（厳密な議論は省略．ここでは形式的に理解すれば十分である）．$\mathcal{A}$ をブラウン運動の生成作用素（infinitesimal generator）という．テイラーの定理を適用すると，各 $\sqrt{h}y$ に依存する $\theta \in (0,1)$ を用いて

$$\begin{aligned} \lim_{h \searrow 0} \frac{1}{h}(P_h f - f)(x) &= \lim_{h \searrow 0} \int_{-\infty}^{\infty} \frac{f(x + \sqrt{h}y) - f(x)}{h} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy \\ &= \lim_{h \searrow 0} \int_{-\infty}^{\infty} \frac{1}{h}[\sqrt{h}y f'(x) + \frac{1}{2} h y^2 f''(x + \theta\sqrt{h}y)] \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy \end{aligned}$$

と表される．第一項の積分は部分積分によりゼロ，第二項は積分と極限の順序を入れ替えることにより $f''(x)/2$ となる．よって，

$$\mathcal{A} f(x) = \frac{1}{2} \frac{d^2}{dx^2} f(x), \quad f \in C_b^2(\mathbb{R}).$$

すなわち，ブラウン運動の生成作用素は 2 階の（偏）微分作用素である．

最後に，ブラウン運動の数値シミュレーションの方法に言及しておこう．例えば $0 = t_0 < t_1 < \cdots < t_n = t$ に対して $W_{t_1}, \ldots, W_{t_n}$ の標本路をシミュレーションしたい場合を考えよう．$W_{t_j} - W_{t_{j-1}}$ が平均 0 分散 $t_j - t_{j-1}$ の正規分布に従い，各々独立であることを考慮し，標準正規乱数列 $\{\xi_j\}_{j=1}^n$ を独立に生成し，漸化式

$$\tilde{W}_j = \tilde{W}_{j-1} + \sqrt{t_j - t_{j-1}} \xi_j, \quad j = 1, \ldots, n$$

により $\{\tilde{W}_j\}_{j=1}^n$ を定義すると，各 $\tilde{W}_j$ を W_{t_j} の擬似的な標本とみなすことができる．

さらなる学習のために

- マルコフ連鎖: [2], [6], [4].
- 待ち行列: [13], [11], [12].
- 確率制御理論: [1], [16].
- ブラウン運動，確率微分方程式: [5], [14], [16].
- 確率論，確率過程論: [15], [7], [8], [9], [10].

付録 A

測度論的確率論

A.1　確率空間，確率測度

定義 **A.1.1.** Ω を任意の集合とする．Ω の集合族 $\mathcal{F}$ が Ω 上の σ-**集合体**（または σ-加法族）であるとは以下の条件が満たされるときにいう．

(1) $\emptyset \in \mathcal{F}$.

(2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$．ただし $A^c = \Omega \setminus A$.

(3) $A_1, A_2, \dots \in \mathcal{A} \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

- $A \in \mathcal{F}$ を（$\mathcal{F}$ に関して）可測な集合，あるいは**事象**と呼ぶ．
- 組 $(\Omega, \mathcal{F})$ を**可測空間**という．

例 **A.1.2.** 任意の集合 Ω に対して，Ω の部分集合全体 $\mathcal{F} = 2^{\Omega} := \{A : A \subset \Omega\}$ は σ-集合体である．

命題 **A.1.3.** $(\Omega, \mathcal{F})$ を可測空間とし，$A_i \in \mathcal{F}$, $i = 1, 2, \dots$ とする．このとき，以下の集合はすべて $\mathcal{F}$ に関して可測な集合である．

$$\bigcup_{i=1}^{n} A_i, \quad \bigcap_{i=1}^{n} A_i, \quad \bigcap_{i=1}^{\infty} A_i, \quad \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i, \quad \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i.$$

注意 **A.1.4.** 確率論では基本的に，Ω の部分集合を不確実に起きる現象の集まりとみなし，その確率を測るなど考察の対象にする．このとき σ-集合体 $\mathcal{F}$ は考察可能な現象の適当なクラスと思えばよい．例えば，考察可能な現象の集まり A と B があり（i.e.,$A, B \in \mathcal{F}$），A と B が同時に起きる現象 $A \cap B$ や A は起きるが B は起きない現象 $A \cap B^c$ を考えたい場合に，これらの集合も考察可能でなければならない（i.e, $A \cap B, A \cap B^c \in \mathcal{F}$）．事象の集まりが σ-集合体であることはこのための一つの要請である．つまり，上の命題でみたように σ-集合体は様々な集合演算に関して閉じているため，複雑な事象の組み合わせを考察対象にできる良い定義域になっているのである．しかし，何故そんなことを考えなければならないのだろうか？ 常に Ω の部分集合全体だけをその定義域として採用していてはダメなのだろうか？ ここでは理由を述べないが、それではしばしば不都合が起こる

のである．一般に，問題を解くときには設定に応じて σ-集合体を適切に選ぶ必要がある．とはいえ，ある程度慣れてしまえば σ-集合体について強く意識することもなくなるであろう．

Ω の部分集合族 $\mathcal{G}$ に対して，

$$\sigma(\mathcal{G}) := \cap\{\mathcal{H} : \Omega \text{ 上の } \sigma\text{-集合体 s.t. } \mathcal{G} \subset \mathcal{H}\}$$

とおく．これは $\mathcal{G}$ を含む最小の σ-集合体である．

例 **A.1.5.** $A \in \mathcal{F}$ とし，$\mathcal{G} = \{A\}$ の場合は，$\sigma(\mathcal{G})$（これを $\sigma(A)$ とも書く）$= \{\emptyset, A, A^c, \Omega\}$ である．

Ω をある位相空間，$\mathcal{G}$ を Ω の開集合全体とするとき，$\sigma(\mathcal{G})$ を Ω 上の Borel σ-集合体といい，$\mathcal{B}(\Omega) = \sigma(\mathcal{G})$ と書く．よく使うのは，$\Omega = \mathbb{R}^n, [a,b]$ の場合等である．また，$\mathcal{B}([a,b])$ を $\mathcal{B}[a,b]$ などと略記することもある．

定義 **A.1.6.** $\mathbb{P} : \mathcal{F} \to [0,1]$ が可測空間 $(\Omega, \mathcal{F})$ 上の**確率測度**であるとは，以下が満たされるときにいう．

(1) $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$.

(2) $A_1, A_2, \dots \in \mathcal{F}$ かつ $A_i \cap A_j = \emptyset$ $(i \neq j)$ ならば

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

- 3 つ組み $(\Omega, \mathcal{F}, \mathbb{P})$ を**確率空間**という．
- $\mathbb{P}(A) =$「事象 A が起こる確率」
- $\mathbb{P}(A) = 1$ のとき，「事象 A は確率 1 で起こる」，「事象 A はほとんど確実に（almost surely, a.s.）起こる」などという．

例 **A.1.7.** $(\Omega, \mathcal{F})$ を可測空間とする．$\omega_0 \in \mathcal{F}$ を取って固定し，$\mathbb{P} : \mathcal{F} \to [0,1]$ を

$$\mathbb{P}(A) = \begin{cases} 1, & \text{if } \omega_0 \in A, \\ 0, & \text{if } \omega_0 \notin A \end{cases}$$

により定義する．このとき $\mathbb{P}$ は $(\Omega, \mathcal{F})$ 上の確率測度である．この場合，$\mathbb{P}$ は ω_0 に集積したディラック測度という．

例 **A.1.8.** Ω を有限集合（i.e., $\#\Omega < \infty$）とし，$\mathcal{F}$ を Ω の部分集合全体とする．各 $\omega \in \Omega$ に対してその確率 $p_\omega \in [0,1]$ を与え，$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$ と定義する．この手続きによって $(\Omega, \mathcal{F})$ 上の任意の確率測度が構成できる．

例 **A.1.9** (ルベーグ測度)**.** $((0,1], \mathcal{B}((0,1]))$ 上の確率測度 μ で，

$$\mu((a,b]) = b - a, \quad 0 \leq a \leq b \leq 1$$

を満たすものの存在が知られている．すなわち区間の長さを測度とするものである．これを $((0,1],\mathcal{B}(0,1])$ 上の**ルベーグ測度**という．$\{0\}$ の測度をゼロとすることで，$([0,1],\mathcal{B}[0,1])$ 上の測度とみることもある．

また，$(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 上の非負測度（i.e., 定義 A.1.6 の (2) を満たす非負値集合関数）ν で

$$\nu((a,b]) = b-a, \quad -\infty \le a \le b \le +\infty$$

を満たすものが存在する．これを $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 上のルベーグ測度という．

さらに，この測度の $[\alpha,\beta]$ への制限を $([\alpha,\beta],\mathcal{B}[\alpha,\beta])$ 上のルベーグ測度という．先に述べた $((0,1],\mathcal{B}(0,1])$ 上のルベーグ測度は，$\mathbb{R}$ 上のルベーグ測度の $(0,1]$ への制限とみなしてもよい．

命題 **A.1.10.** $\mathbb{P}$ を $(\Omega,\mathcal{F})$ 上の確率測度とするとき，以下が成り立つ:

(1) $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, が単調増加（i.e., $A_1 \subset A_2 \subset \cdots$）のとき，

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

(2) $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, が単調減少（i.e., $A_1 \supset A_2 \supset \cdots$）のとき，

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

以下，この章の終わりまで確率空間 $(\Omega,\mathcal{F},\mathbb{P})$ を考える．

- $A, B \in \mathcal{F}$, $\mathbb{P}(B) > 0$ とする．

$$\mathbb{P}(A \,|\, B) := \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$$

  を事象 B が与えられたときの事象 A の**条件付き確率**（conditional probability）という．
- $\mathbb{P}(A \,|\, B)$ は B が起こるという前提で A の起こる確率をみるものであり，$\mathbb{P}(B)$ で割ることにより規格化している．

命題 **A.1.11.** $\mathbb{P}(B) > 0$ なる $B \in \mathcal{F}$ が与えられたときの条件付き確率の族 $\mathbb{P}(A \,|\, B)$, $A \in \mathcal{F}$, は $(\Omega,\mathcal{F})$ 上の確率測度である．

- 事象 $A, B \in \mathcal{F}$ が独立（independent）であるとは，

$$\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

  を成り立つときにいう．
- $\mathbb{P}(B) > 0$ のとき，独立性は $\mathbb{P}(A \,|\, B) = \mathbb{P}(A)$ と同値であり，これは「B が起こるという前提であっても A の確率は変わらない」ことを意味している．

- 事象 $A_i \in \mathcal{F}$, $i = 1, \ldots, n$, が独立であるとは，任意の $\ell \in \{1, \ldots, n\}$ と互いに異なる任意の $i_1, \ldots, i_\ell \in \{1, \ldots, n\}$ に対して

$$\mathbb{P}\left(\bigcap_{k=1}^{\ell} A_{i_k}\right) = \prod_{k=1}^{\ell} \mathbb{P}(A_{i_k})$$

が成り立つときにいう．

A.2　確率変数

確率空間とはいわば不確実性の「泉」とそれを測る道具の組であるが，個々の不確実性の素に対する実現値を記述するのが確率変数である．

定義 **A.2.1.** $X : \Omega \to \mathbb{R} \cup \{\pm\infty\}$ は

$$\{\omega \in \Omega : X(\omega) > a\} \in \mathcal{F}, \quad a \in \mathbb{R}$$

を満たすとき，（$\mathcal{F}$-可測）**確率変数**であるという．

- 事象 $\{\omega \in \Omega : X(\omega) \in B\}$ は通常，$\{X \in B\}$ などと略記される．

例 **A.2.2.** 次の夏の平均気温の集合を $\Omega = \{$ 低い, 平年並, 高い $\}$ とする．このとき，あるビール会社の夏期の売上高 X は Ω 上の確率変数と考えることができる．

- X が確率変数ならば，任意の区間 I に対して $\{X \in I\} \in \mathcal{F}$.
- Ω の部分集合 A に対して

$$1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

で定義された関数 1_A を集合 A の定義関数あるいは**指示関数**という．
- 1_A が確率変数 $\iff A \in \mathcal{F}$.
- 確率変数の列 $\{X_k\}_{k=1}^n$ が独立であるとは，任意の $a_1, \ldots, a_n \in \mathbb{R}$ に対して

$$\mathbb{P}(X_1 \le a_1, \ldots, X_n \le a_n) = \mathbb{P}(X_1 \le a_1) \cdots \mathbb{P}(X_n \le a_n)$$

が成り立つときにいう．
- 確率変数の（無限）列 $\{X_n\}_{n=1}^\infty$ が独立であるとは，任意の相異なる $n_1, \ldots, n_k \in \mathbb{N}$ に対して $\{X_{n_i}\}_{i=1}^k$ が独立であるときにいう．
- 確率変数のベクトル $X = (X_1, \ldots, X_m)$ と $Y = (Y_1, \ldots, Y_n)$ が独立であるとは，任意の $a_1, \ldots, a_m, b_1, \ldots, b_n \in \mathbb{R}$ に対して

$$\begin{aligned} &\mathbb{P}(X_1 \le a_1, \ldots, X_m \le a_m, Y_1 \le b_1, \ldots, Y_n \le b_n) \\ &= \mathbb{P}(X_1 \le a_1, \ldots, X_m \le a_m)\mathbb{P}(Y_1 \le b_1, \ldots, Y_n \le b_n) \end{aligned}$$

が満たされるときにいう．

例 **A.2.3.** $p \in [0,1]$ とする．$\{0,1,\dots,n\}$-値確率変数 X が

$$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

を満たすとき，X はパラメータ (n,p) の二項分布に従うという．

例 **A.2.4.** $\lambda > 0$ とする．0 以上の自然数の集合 $\mathbb{N} \cup \{0\}$ に値をとる確率変数 X が

$$\mathbb{P}(X=k) = e^{-\lambda} \frac{\lambda^k}{k}, \quad k \in \mathbb{N} \cup \{0\}$$

を満たすとき，X はパラメータ λ の**ポアソン分布**に従うという．

- X が実数値のとき，右連続単調非減少関数

$$F_X(x) := \mathbb{P}(X \in (-\infty, x]) = \mathbb{P}(X \leq x), \quad x \in \mathbb{R}$$

  を X の分布関数という．
- 実数値確率変数 X について，$\mathbb{P}(X \in S) = 1$ を満たす区間 S を X の**台**（support）と呼ぶ．
- X を実数値確率変数とし，S をその台とする．$\int_S f(x)dx = 1$ なる S 上の非負連続関数 f が

$$\mathbb{P}(a < X \leq b) = \int_a^b f(x)dx, \quad -\infty \leq a < b \leq +\infty$$

  を満たすと仮定する．ただし S の外側では f をゼロと定義しておく．この f を X の**密度関数**（density function）と呼ぶ．また，X の分布は密度 f をもつという．
- 密度関数はより弱い形で定義可能だが，応用上は上記の定義で十分である．

例 **A.2.5.** $m \in \mathbb{R}, \sigma > 0$ とする．確率変数 X が $\mathbb{R}$ を台に持ち，さらに密度関数

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma}}, \quad x \in \mathbb{R}$$

をもつとき，X は平均 m, 分散 σ の**正規（ガウス）分布**に従うといい，$X \sim N(m,\sigma)$ と書く．$m=0, \sigma=1$ のときは X は標準正規分布に従うという．

例 **A.2.6.** $\lambda > 0$ とする．確率変数 X の台が $(0,\infty)$ であり，密度が

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0$$

で与えられるとき，X はパラメータ λ の指数分布に従うという．

A.3 期待値

X を確率変数とする．

(1) $X(\Omega)=\{X(\omega):\omega\in\Omega\}$ が有限集合のとき，X の**期待値**（expectation）$\mathbb{E}[X]$ を

$$\mathbb{E}[X]=\sum_{x\in X(\Omega)}x\mathbb{P}(X=x)$$

により定義する．このとき，任意の実数値関数 f について

$$\mathbb{E}[f(X)]=\sum_{x\in X(\Omega)}f(x)\mathbb{P}(X=x).$$

(2) X が密度関数 p をもつとき，その期待値 $\mathbb{E}[X]$ を

$$\mathbb{E}[X]=\int_S xp(x)dx$$

により定義する．ここで，S は X の台とし，右辺は（リーマン）積分可能とする．このとき，任意の有界連続関数 f に対して，

$$\mathbb{E}[f(X)]=\int_S f(x)p(x)dx.$$

(3) 一般の非負確率変数 X に対する期待値は

$$\mathbb{E}[X]=\lim_{n\to\infty}\sum_{k=0}^{n2^n-1}\frac{k}{2^n}\mathbb{P}\left(\frac{k}{2^n}\le X<\frac{k+1}{2^n}\right)+n\mathbb{P}(X\ge n)\in[0,\infty]$$

により定義される．

(4)（非負とは限らない）X に対して定義される確率変数

$$X^+(\omega):=\max\{X(\omega),0\},\quad X^-(\omega):=-\min\{X(\omega),0\},\quad \omega\in\Omega$$

はともに非負確率変数であるから上のように期待値が定義される．よって $\mathbb{E}[X^+]$, $\mathbb{E}[X^-]$ のどちらかが有限値のとき，X の期待値 $\mathbb{E}[X]$ を

$$\mathbb{E}[X]=\mathbb{E}[X^+]-\mathbb{E}[X^-]$$

により定義する．

(5) X,Y を実数値確率変数，$i=\sqrt{-1}$ を虚数単位とするとき，$Z=X+iY$ は複素数値確率変数となる．複素数値確率変数 Z の期待値は

$$\mathbb{E}[Z]=\mathbb{E}[X]+i\mathbb{E}[Y]$$

と定義される．例えば，（実）確率変数 X と $t\in\mathbb{R}$ に対して，

$$\mathbb{E}[e^{itX}]=\mathbb{E}[\cos(tX)]+i\mathbb{E}[\sin(tX)].$$

- 一般的な期待値の定義 (3) および (4) は離散的な場合の定義 (1) と連続的な場合の定義 (2) と整合的である．

- $|X| = X^+ + X^-$ より $\mathbb{E}[X]$ が有限値であることと $\mathbb{E}[|X|] < \infty$ であることは同値である．
- 確率変数 X の分散 $\mathbb{E}[(X - \mathbb{E}[X])^2]$ を $\mathbb{V}(X)$ で表す．

命題 **A.3.1.** X, Y を実確率変数とし，これらの期待値が定義されるとする．$a, b \in \mathbb{R}$ に対して以下が成り立つ．

(1) $X = Y$ a.s. $\Longrightarrow \mathbb{E}[X] = \mathbb{E}[Y]$.
(2) $X \leq Y$ a.s. $\Longrightarrow \mathbb{E}[X] \leq \mathbb{E}[Y]$.
(3) $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$（これは右辺が $\infty - \infty$ でない限りで成り立つ）．
(4) $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.
(5) $\mathbb{E}[|X|] < \infty \Longrightarrow |X| < \infty$ a.s.
(6) $X \geq 0$ a.s., $\mathbb{E}[X] = 0 \Longrightarrow X = 0$ a.s.
(7) $X \geq Y$ a.s., $\mathbb{E}[X] = \mathbb{E}[Y] \Longrightarrow X = Y$ a.s.

第 1 章以降で使う期待値の性質のいくつかを次の命題でまとめておこう．

命題 **A.3.2.** 次の (1)–(4) が成り立つ．

(1) $a_n \geq 0$, $A_n \in \mathcal{F}$, $n \in \mathbb{N} \cup \{0\}$ に対し，

$$\mathbb{E}\left[\sum_{n=0}^{\infty} a_n 1_{A_n}\right] = \sum_{n=0}^{\infty} a_n \mathbb{P}(A_n).$$

(2) $\mathbb{N} \cup \{0\}$-値確率変数 X に対し，

$$\mathbb{E}[X] = \sum_{n=0}^{\infty} \mathbb{P}(X > n).$$

(3) (チェビシェフの不等式). $\varepsilon > 0$, $p \geq 1$ と $\mathbb{E}[|X|^p] < \infty$ を満たす実確率変数 X に対して

$$\mathbb{P}(|X| > \varepsilon) \leq \frac{1}{\varepsilon^p}\mathbb{E}[|X|^p].$$

(4) A と B を高々可算集合とし，X と Y はそれぞれ A と B に値をとる確率変数とする．X と Y が独立のとき，任意の有界関数 $f : A \times B \to \mathbb{R}$ に対して，

$$\mathbb{E}[f(X, Y)] = \mathbb{E}\left[\mathbb{E}[f(x, Y)]|_{x=X}\right].$$

- 命題 A.3.2(4) より，$f(X, Y)$ の期待値を計算するためには，まず X を定数とみなして $f(X, Y)$ の期待値を Y に関してとり，その後で X に関して期待値をとればよいことになる．
- 命題 A.3.2(4) はより一般の確率変数に対して成り立つが，本稿では上記のような命題で十分である．

一般に，d 次元確率変数 $X=(X_1,\ldots,X_d)$ に対して

$$\varphi_X(t)=\mathbb{E}\left[e^{i\sum_{k=1}^d t_kX_k}\right],\quad t=(t_1,\ldots,t_d)\in\mathbb{R}^d$$

を X の特性関数という．ただし $i=\sqrt{-1}$. 特性関数により確率変数の分布は完全に決定される．

命題 **A.3.3.** X,Y を実数値確率変数とする．任意の $t\in\mathbb{R}^d$ について $\varphi_X(t)=\varphi_Y(t)$ ならば $F_X=F_Y$ である．

独立性を特性関数で判定することもできる．

定理 **A.3.4.** 確率変数列 $\{X_k\}_{k=1}^n$ が独立であるための必要十分条件は,

$$\varphi_X(t)=\prod_{k=1}^n\varphi_{X_k}(t_k),\quad t_1,\ldots,t_n\in\mathbb{R}.$$

ここで $X=(X_1,\ldots,X_n)$, $t=(t_1,\ldots,t_n)$.

定理 **A.3.5.** 確率変数のベクトル $X=(X_1,\ldots,X_m)$ と $Y=(Y_1,\ldots,Y_n)$ が独立であるための必要十分条件は，任意の $\xi_1,\ldots,\xi_m,\eta_1,\ldots,\eta_n\in\mathbb{R}$ に対し

$$\mathbb{E}\left[\exp\left(i\sum_{k=1}^m\xi_kX_k+i\sum_{j=1}^n\eta_jY_j\right)\right]=\mathbb{E}\left[\exp\left(i\sum_{k=1}^m\xi_kX_k\right)\right]\mathbb{E}\left[\exp\left(i\sum_{j=1}^n\eta_jY_j\right)\right]$$

が成り立つことである．

A.4 確率変数の収束

実確率変数列 $\{X_n\}_{n=1}^\infty$ と実確率変数 X が与えられているとする．

- $X_n(\omega)$ が $X(\omega)$ に収束するような $\omega\in\Omega$ の集合が確率 1 であるとき，すなわち，

  $$\mathbb{P}(X=\lim_{n\to\infty}X_n)=1$$

  のとき，$\{X_n\}$ は X に概収束するといい，$\lim_{n\to\infty}X_n=X$, a.s. と書く．また，$\{X_n\}$ は X にほとんど確実に収束するということも多い．
- 任意の $\varepsilon>0$ に対して

  $$\lim_{n\to\infty}\mathbb{P}(|X_n-X|>\varepsilon)=0$$

  が成り立つとき，$\{X_n\}$ は X に確率収束するという．
- $p\in\mathbb{N}$ に対して

  $$\lim_{n\to\infty}\mathbb{E}[|X_n-X|^p]=0$$

  が成り立つとき，$\{X_n\}$ は X に p 次平均収束する，あるいは L^p-収束するという．
- $\mathbb{R}$ 上の任意の有界連続関数 f に対して

  $$\lim_{n\to\infty}\mathbb{E}[f(X_n)]=\mathbb{E}[f(X)]$$

が成り立つとき，$\{X_n\}$ は X に法則収束する，あるいは分布収束するという．

定理 **A.4.1.** 次の (1)–(4) が成り立つ．

(1) $\{X_n\}$ が X に概収束するならば確率収束する．
(2) $\{X_n\}$ が X に p 次平均収束するならば確率収束する．
(3) $\{X_n\}$ が X に確率収束するならば法則収束する．
(4) $\{X_n\}$ が X に確率収束するとき，概収束するような部分列が存在する．

極限と期待値の交換可能性について，以下の 3 つの事実が基本的である．

定理 **A.4.2** (単調収束定理)**.** $\{X_n\}$ は $0 \le X_1 \le X_2 \le \cdots$ a.s. を満たす確率変数列とする．このとき，

$$\mathbb{E}[X_n] \nearrow \mathbb{E}[X] \quad (n \to \infty).$$

補題 **A.4.3** (Fatou（ファトゥ）の補題)**.** $\{X_n\}$ を非負確率変数列とするとき，

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n\right] \le \liminf_{n\to\infty} \mathbb{E}[X_n].$$

定理 **A.4.4** (優収束定理)**.** 確率変数列 $\{X_n\}$ と確率変数 X が次の条件を満たすとする．

(1) $X_n \to X$ a.s.
(2) ある可積分確率変数 Y が存在して，任意の n に対して $|X_n| \le Y$ a.s.

このとき，

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

A.5 極限定理

コイン投げの繰り返しにおいて，各回に表が出たら $+1$, 裏が出たら -1 として値を加減していき，n 回目までの合計を S_n により表す．表と裏の出る確率が等しいとき，n を大きくすれば平均値 S_n/n が 0 に近づくことが予想できるであろう．このことを正当化した定理が大数の法則であり，収束の様子まで記述したものが中心極限定理である．

定理 **A.5.1** (大数の弱法則)**.** 確率変数列 $\{X_n\}_{n=1}^{\infty}$ が次の (1) と (2) を満たすと仮定する:

(1) 任意の $i, j \in \mathbb{N}$, $i \ne j$, に対して，X_i と X_j は独立．
(2) $\sup_{n\in\mathbb{N}} \mathbb{V}(X_n) < \infty$.

このとき，$\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])/n$ は 0 に確率収束する．すなわち，

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])\right| > \varepsilon\right) = 0.$$

$\{X_n\}$ が確率変数列として独立であり（第 A.2 参照），各 X_n が同じ分布を持つとき（こ

のとき $\{X_n\}$ は**独立同分布**であるという)，より強い収束が得られる．

定理 A.5.2 (大数の強法則)**.** $\{X_n\}_{n=1}^{\infty}$ を独立同分布の確率変数列で $\mathbb{E}|X_1| < \infty$ を満たすと仮定する．このとき，$\sum_{i=1}^{n} X_i/n$ は共通の期待値に概収束する．すなわち，

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mathbb{E}[X_1], \quad \text{a.s.}$$

n ではなく $\sqrt{n}$ のスケールで考えることで，収束の様子が見えてくる．

定理 A.5.3 (中心極限定理)**.** $\{X_n\}_{n=1}^{\infty}$ を独立同分布確率変数列で，$m = \mathbb{E}[X_1]$, $v = \mathbb{V}(X_1)$ とおく．このとき，

$$Z_n = \frac{1}{\sqrt{nv}} \sum_{i=1}^{n} (X_i - m), \quad n \in \mathbb{N}$$

は平均 0, 分散 1 の正規分布に従う確率変数に分布収束する．

参考文献

[1] D. Bertsekas. *Dynamic programming and optimal control*, Vol. I. Athena Scientific, Belmmont, 3rd edition, 2005.

[2] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues.* Springer, New York, 1999.

[3] T. Mikosch. *Non-life insurance mathematics.* Springer-Verlag, Berlin, 2 edition, 2009.

[4] S. I. Resnick. *Adventures in stochastic processes.* Birkhäuser, Boston, 2002.

[5] B. エクセンダール（著）, 谷口説男（訳）. 確率微分方程式. 丸善出版, 2012.

[6] R. デュレット. 確率過程の基礎. シュプリンガー・フェアラーク東京, 2005. (今野紀雄他訳，原著英語版 1999).

[7] W. フェラー（著）, 河田龍夫（監訳）. 確率論とその応用 I 上. 紀伊国屋書店, 1960.

[8] W. フェラー（著）, 河田龍夫（監訳）. 確率論とその応用 I 下. 紀伊国屋書店, 1961.

[9] W. フェラー（著）, 河田龍夫（監訳）. 確率論とその応用 II 上. 紀伊国屋書店, 1969.

[10] W. フェラー（著）, 河田龍夫（監訳）. 確率論とその応用 II 上. 紀伊国屋書店, 1970.

[11] 塩田茂雄, 河西憲一, 豊泉洋, 会田雅樹（著）, 川島幸之助（監修）. 待ち行列理論の基礎と応用. 共立出版, 2014.

[12] 宮沢政清. 待ち行列の数理とその応用. 牧野書店, 2013.

[13] 高橋幸雄, 森村英典. 混雑と待ち. 朝倉書店, 2001.

[14] 舟木直久. 確率論. 朝倉書店, 2004.

[15] 小谷眞一. 測度と確率. 岩波書店, 2006.

[16] 長井英生. 確率微分方程式. 共立出版, 2003.