7　確率過程と条件付き期待値

後半のテーマは，時間的に変化するランダムな現象を表す確率モデルである．本節の授業では∗のついた小節を省きます．

以後，時間は非負の整数値を取って進むとする．時刻 n でランダムな現象が表す状態を X_n により表す．状態はランダムな現象の実現値（標本を選んだときに得られる値）である．この値の集合を S により表すと，X_n は Ω から S への関数である．このとき，$\{X_n; n=0,1,\ldots\}$ を状態空間 S をもつ離散時間確率過程（または単に確率過程）と呼ぶ．$\mathbb{Z}_+$ を非負の整数全体の集合とすると，各 $\omega\in\Omega$ に対する確率過程の標本 $\{X_n(\omega); n=0,1,\ldots\}$ は，$\mathbb{Z}_+$ から S への関数である．この関数を標本関数または標本路 (sample path) と呼ぶ．本節では確率過程を観測して得られる情報を表すためにフィルトレーションと条件付き期待値を定義する．

7.1　抽象化の目的

X を Ω から状態空間 S への関数とする．$B\subset S$ に対して $\{X\in B\}$ の確率を測る，すなわち，$\{X\in B\}$ を事象とするために，B を S 上の σ-集合体の要素とすると都合が良い．なぜならば，この S 上の σ-集合体を $\mathscr{H}$ とするとき，

$$\sigma(X)\equiv\{\{X\in B\}; B\in\mathscr{H}\}$$

もまた σ-集合体となるからである．この $\mathscr{H}$ を作るためには，S の要素間に何らかの関係があるとよい．そこで，S は距離空間であるとする．すなわち，$x,y\in S$ に対して次の条件を満たす $\rho(x,y)$ が定義されているとする．

(a) $\rho(x,y)\ge 0,\quad \rho(x,y)=\rho(y,x)$,

(b) $\rho(x,y)=0$ が成り立つのは $x=y$ のときに限る，

(c) $\forall x,y,z\in S$ に対して $\rho(x,z)\le\rho(x,y)+\rho(y,z)$.

このとき，$\rho(x,y)$ を x,y の距離，(S,ρ) を距離空間と呼ぶ．また，$\varepsilon>0$ と $x\in S$ に対して，$V_\varepsilon(x)\equiv\{y\in S;\rho(x,y)<\varepsilon\}$ を x の ε-近傍と呼ぶ．一般に，$\mathscr{O}$ が以下の 3 条件：

(a') $S\in\mathscr{O}$

(b') 有限個の $i=1,2,\ldots,n$ に対して $O_i\in\mathscr{O}$ ならば $\cap_{i=1}^n O_i\in\mathscr{O}$,

(c') 任意の集合 A に対して，$\alpha\in A$ ならば $O_\alpha\in\mathscr{O}$ であるとき，$\cup_{\alpha\in A}O_\alpha\in\mathscr{O}$.

を満たすとき開集合と呼び，$(S,\mathscr{O})$ を位相空間と呼ぶ．このとき，$C\subset S$ に対して，$C^c\equiv S\setminus C$ が開集合ならば C を閉集合と呼ぶ．距離空間 (S,ρ) において，S の部分集合 O が，$\forall x\in O$ に対して $\exists\varepsilon>0$ があり $V_\varepsilon(x)\subset O$ ならば，O を開集合とするとき，位相空間である．

S のすべての開集合を含む S 上の最小の σ-集合体を S 上のボレル集合体と呼び $\mathscr{B}(S)$ と表す．これまで，実数の集合や実ベクトル空間 $\mathbb{R}^d$ を距離空間へ，更に位相空間へ抽象化してきた．この抽象化は概念の本質的な部分を取り出すことが目的である．抽象化は具体的なイメージがもちにくいので難しいと思われがちであるが，問題を簡単にする．抽象化は思考の節約であり数理的思考に欠かせない．

演習問題 7.1 (a) S が可算集合であるとき，距離を定義して距離空間を作りボレル集合体 $\mathscr{B}(S)$ を求めよ．
(b) $S = \mathbb{R}$ であるとき，$x, y \in \mathbb{R}$ に対して $\rho(x, y) = |x - y|$ と定義する．ρ は距離であり，ボレル集合体 $\mathscr{B}(\mathbb{R})$ がこれまで用いてきた $\sigma(\{(a, b]; a, b \in \mathbb{R}, a < b\})$ に等しいことを示せ．

時間が連続的である場合には，時間を実数により表し，連続時間確率過程と呼ぶ．以後特に断らない限り，離散時間確率過程のみを論じ，単に確率過程と呼ぶ．

7.2 確率過程における情報とは？

状態空間 S をもつ確率過程 $\{X_n; n \geq 0\}$ においては時刻の経過と共に観測結果として得られる情報量が増える．このときの情報量とは何であろうか？確率モデルである確率空間 $(\Omega, \mathscr{F}, \mathbb{P})$ に立ち戻って考えてみよう．このとき，$\forall A \in \mathscr{F}$ に対して確率 $\mathbb{P}(A)$ が与えられている．これは σ-集合体 $\mathscr{F}$ が情報をもっている Ω の部分集合の集まりであることを表している．従って，$\mathscr{F}$ はこのモデルにおける情報の集まりと解釈することもできる．

ここで注意すべきことは，$\mathbb{P}(A)$ が実際にすべての $A \in \mathscr{F}$ に対して計算できるわけではないことである．多くの確率モデルでは確率測度 $\mathbb{P}$ の存在について論じても具体的に求めることは容易ではない．比較的簡単な場合が，実数値を取る確率変数 X についてその分布 F が既知の場合である．このときには，$S = \mathbb{R}$ であり，任意の $B \in \mathscr{B}(\mathbb{R})$ に対して，

$$\mathbb{P}(X \in B) = \int_{-\infty}^{+\infty} 1(x \in B) F(dx)$$

により，事象 $\{X \in B\} \equiv \{\omega \in \Omega; X(\omega) \in B\}$ の確率が（原理的に）計算できる．そこで，このような事象の集まりを

$$\sigma(X) = \{\{X \in B\}; B \in \mathscr{B}(\mathbb{R})\}$$

と表し，X を観測したときに得られる情報と解釈する．確率変数の定義より $\sigma(X) \subset \mathscr{F}$ であり，$\sigma(X)$ は Ω 上の σ-集合体である．

演習問題 7.2 $\sigma(X)$ が Ω 上の σ-集合体であることを証明せよ．

複数の確率変数についても同様な σ-集合体を作る．同時に複数の確率変数から作られる事象も含む必要がある．そこで，n 個の確率変数 $X_1, X_2, \ldots, X_n$ に対して，

$$\begin{aligned}\sigma(X_1, X_2, \ldots, X_n) &= \sigma(\cup_{i=1}^n \sigma(X_i)) \\ &\equiv \cup_{i=1}^n \sigma(X_i) \text{ を含む } \Omega \text{ 上の最小の } \sigma\text{-集合体}\end{aligned}$$

により定義した σ-集合体を $X_1, X_2, \ldots, X_n$ に関する情報の全体とみなす．

演習問題 7.3 コインを投げ，表ならば 1 裏ならば 0 により表す．$\Omega = \{(i,j); i,j = 0,1\}$ を標本空間とする確率空間 $(\Omega, \mathscr{F}, \mathbb{P})$ によりコインを 2 回投げることを表す．この確率モデルにおいて，コインを 1 回投げた場合の情報を表す σ-集合体 $\mathscr{G}$ を求めよ．

7.3　情報のフィルトレーション

状態空間 S をもつ確率過程 $\{X_n; n \geq 0\}$ に対して，時刻 n まで観測したときの情報を σ-集合体 $\mathscr{F}_n^X$ により表す．すなわち，

$$\mathscr{F}_n^X = \sigma(X_0, X_1, \ldots, X_n), \qquad n \geq 0, \tag{7.1}$$

である．$\mathscr{B}(S^n)$ を S^n の開集合から生成した S 上の σ-集合体とすると，

$$\mathscr{F}_n^X = \{\{(X_0, X_1, \ldots, X_n) \in \boldsymbol{B}\}; \boldsymbol{B} \in \mathscr{B}(S^n)\}$$

と表すこともできる．

定義より，$\mathscr{F}_n^X \subset \mathscr{F}$ であり，$\mathscr{F}_n^X$ は n について非減少である．一般にこの特性を持つ σ-集合体の列を定義しておこう．

定義 7.1 (フィルトレーション) 各 $n \geq 0$ に対して，$\mathscr{F}$ の部分集合 $\mathscr{F}_n$ が

(a) $\mathscr{F}_n$ は Ω 上の σ-集合体である，

(b) $\mathscr{F}_n \subset \mathscr{F}_{n+1}$，

を満たすとき，$\{\mathscr{F}_n; n = 0, 1, \ldots\}$ をフィルトレーションと呼び，$\mathbb{F}$ と表す．また，状態空間 S をもつ確率過程 $\{X_n; n = 0, 1, \ldots\}$ が，フィルトレーション $\mathbb{F}$ に対して，

(c) 任意の $n \geq$ に対して $\mathscr{F}_n^X \subset \mathscr{F}_n$，

を満たすならば，$\mathbb{F}$ に適合した確率過程であるという．

注 7.1 フィルトレーションという呼び名は *filtration* の音訳である．しかし，日本語で表現するとろ過機構となるが，ろ過は日常的に余り使われていないのでフィルトレーションとした．

確率過程 $\{X_n; n \geq 0\}$ に対する $\mathbb{F}^X \equiv \{\mathscr{F}_n^X; n = 0, 1, \ldots\}$ もフィルトレーションである．ではなぜより一般的なフィルトレーション $\mathbb{F}$ を定義したのだろうか？これは対象としている確率過程の外にもランダムな影響を及ぼす要因ある場合に備えるためである．言葉を換えると柔軟性のあるモデル化を行い応用範囲を拡げるためである．

例 7.1 (ランダムウォーク) (a) 表が出る確率が p であるコイン投げの n 回目の結果を確率変数 U_n により表す．ここに，$0 < p < 1$，

$$U_n = \begin{cases} 1, & n \text{ 回目の結果が表}, \\ -1, & n \text{ 回目の結果が裏}, \end{cases}$$

とする．このとき，$U_1, U_2, \ldots$ は，独立で同一の分布，

$$\mathbb{P}(U_n = 1) = p, \qquad \mathbb{P}(U_n = -1) = 1 - p,$$

に従うとする．次に，$a \in \mathbb{Z}$ に対して $X_0 = a$，

$$X_n = X_0 + U_1 + U_2 + \ldots + U_n, \qquad n \geq 1, \tag{7.2}$$

により確率過程 $\{X_n; n = 0, 1, \ldots\}$ を定義する．その状態空間 S は整数全体の集合 $\mathbb{Z}$ に等しい．この確率過程を初期値 a をもつ単純ランダムウォークと呼ぶ．

(b) 単純ランダムウォークにおいて，U_n の分布を $\mathbb{R}$ の n に依らない一般の分布 F に置き換えた場合を考えてみよう．このとき，確率過程 $\{X_n; n = 0, 1, \ldots\}$ の状態空間 S は $\mathbb{R}$（またはその部分集合）に等しい．$X_n - X_0$ は独立で同一の分布に従う確率変数の和であるから，X_n の時刻 n が大きいときの漸近特性を大数の法則と中心極限定理を使って調べることができる．(7.2) は次のように表すこともできる．

$$X_n = X_{n-1} + U_n, \qquad n \geq 1. \tag{7.3}$$

この場合に，X_n を数直線上の位置であると見なすと，U_n は X_{n-1} から X_n へ移動する変化量である．これは，物体が数直線上をランダムに動く様を表しているので，ランダムウォークと呼ばれる．数学的に見たランダムウォークの特徴は，U_n が X_{n-1} と独立であり，X_n が X_{n-1} と U_n によって決まることである．更に，実数 a に対して，$X_0 = a$ と定義し，(7.3) により，$n \geq 1$ に対して X_n を定義するとき，$\{X_n; n \geq 0\}$ を初期値が a であるランダムウォークと呼ぶ．ここで，$U_0 = X_0$ とおいて 2 つのフィルトレーション

$$\mathbb{F}^U \equiv \{\sigma(U_0, U_1, \ldots, U_n); n \geq 0\}, \qquad \mathbb{F}^X \equiv \{\sigma(X_0, X_1, \ldots, X_n); n \geq 0\},$$

をを定義すると，(7.3) より $\mathbb{F}^U = \mathbb{F}^X$ である．

(c) 非負の整数 a を初期値とするランダムウォーク $\{X_n; n \geq 0\}$ に対して，原点に壁を置き，ランダムウォークが原点で負の方向に動こうとするならば原点に止まるとする．すなわち，(7.3) を

$$Y_n = \max(0, Y_{n-1} + U_n), \qquad n \geq 1. \tag{7.4}$$

に置き換える．この $\{Y_n; n \geq 0\}$ を原点に柔らかい反射壁をもつランダムウォークと呼ぶ．このランダムウォークは，どんな $n \geq 0$ に対しても $Y_n \geq 0$ である．しかし，Y_n は依然として Y_{n-1} と U_n により決まる．しかし，フィルトレーションには注意が必要である．すなわち，$\mathbb{F}^Y \equiv \{\sigma(Y_0, Y_1, \ldots, Y_n); n \geq 0\}$ とすると，$\mathbb{F}^Y \subset \mathbb{F}^X$ であるが，一般に $\mathbb{F}^Y$ は $\mathbb{F}^X$ に一致しない． ∎

次に双六遊びを考えてみよう．ただし，簡単のために一人で双六を行うとする．

例 7.2 (双六) 1 から M までの番号がついた M 個の場所があり，X_n を n 回目の場所とする．X_{n-1} から X_n への変化は，X_{n-1} とサイコロを投げた結果により決まる．n 回目のサイコロの結果を U_n により表すと，n に依存した確定的な関数 h_n があり，$n+1$

$$X_n = h_n(X_{n-1}, U_n), \qquad n = 1, 2, \ldots,$$

と表すことができる．この場合の状態空間は $S = \{1, 2, \ldots, M\}$ である． ∎

7.4 条件付き期待値

本節では，確率変数 X や確率過程 $\{X_n; n \geq 0\}$ の状態は実数値を取る，すなわち，$S = \mathbb{R}$ とする．確率過程 $\{X_n; n \geq 0\}$ においては時間の経過と共に観測情報が増える．そこで，各時刻 n において，$X_0, X_1, \ldots, X_n$ の値を知っているという条件の下で X_{n+1} の分布や期待値を求めることが重要になってくる．テスト関数を用いて分布は期待値の特別な場合であると考えると，この問題は，σ-集合体 $\mathscr{F}_n^X$ の下で X_{n+1} またはその関数の期待値を求める問題に帰着する．そこで，問題を一般化するために可測性を一般化する．

定義 7.2 確率空間 $(\Omega, \mathscr{F}, \mathbb{P})$ において，$\mathscr{G} \subset \mathscr{F}$ かつ Ω 上の σ-集合体ならば，$\mathscr{G}$ を部分 σ-集合体と呼ぶ．この $\mathscr{G}$ と確率変数 X に対して，

$$\forall B \in \mathscr{B}(\mathbb{R}) \text{ に対して，} \quad \{X \in B\} \in \mathscr{G}, \tag{7.5}$$

ならば，X は $\mathscr{G}$-可測であるという．

$$\forall B \in \mathscr{B}(\mathbb{R}) \text{ と } \forall A \in \mathscr{G} \text{ に対して，} \quad \mathbb{P}(\{X \in B\} \cap A) = \mathbb{P}(X \in B)\mathbb{P}(A) \tag{7.6}$$

ならば，X と $\mathscr{G}$ は独立であるという．

補題 7.1 確率変数 X, Y が $\mathscr{G}$-可測ならば，定数 a, b に対して $aX + bY$ も $\mathscr{G}$-可測である．

証明 aX と $X + Y$ が $\mathscr{G}$-可測であることを示せばよい．$a = 0$ ならば $aX = 0$ であるから，aX は $\mathscr{G}$-可測である．$a \neq 0$ のとき，$\forall x \in \mathbb{R}$ に対して，

$$\{aX \leq x\} = \begin{cases} \{X \leq x/a\}, & a > 0, \\ \{X \geq x/a\}, & a < 0, \end{cases}$$

であるから，$\{aX \leq x\} \in \mathscr{G}$ となり，aX は $\mathscr{G}$-可測である．次に，$X + Y$ が $\mathscr{G}$-可測であることを示す．このためには，$\forall x \in \mathbb{R}$ に対して，$\{X + Y \leq x\} \in \mathscr{G}$ を示せば十分である．各整数 $n \geq$ に対して，

$$Y_n = \sum_{i=-\infty}^{+\infty} i2^{-n} 1(i2^{-n} < Y \leq (i+1)2^{-n})$$

とおくと，

$$\{X+Y_n \le x\} = \bigcup_{i=-\infty}^{+\infty} \{X + i2^{-n} \le x\} \cap \{i2^{-n} < Y \le (i+1)2^{-n}\} \in \mathscr{G}$$

であるから，$X+Y_n$ は $\mathscr{G}$-可測である．一方，$X+Y_n \le X+Y \le X+Y_n+2^{-n}$ であるから，

$$\lim_{n\to\infty}(X+Y_n) = X+Y$$

が成り立つ．従って，$X+Y_n$ が $\mathscr{G}$-可測であることから，$X+Y$ も $\mathscr{G}$-可測である．　∎

演習問題 7.4 $\mathscr{G}$-可測な確率変数列 $\{X_n; n\ge 1\}$ に対して，

(a) $\inf_{\ell\ge 1}\sup_{n\ge \ell} X_n$ が $\mathscr{G}$-可測であることを証明せよ．

(b) $\Omega_0 = \{\omega\in\Omega; \lim_{n\to\infty} X_n(\omega)$ が存在し有限 $\}$ とおく．$\Omega_0 \in \mathscr{G}$ を証明せよ．

Ω 上の部分 σ-集合体 $\mathscr{G}$ が与えられたとき，確率変数 X の条件付き期待値はどのように定義すればよいだろうか？1 つの考え方は，各 $A\in\mathscr{G}$ に対して，条件付き期待値 $\mathbb{E}(X|A) \equiv \mathbb{E}(X1_A)/\mathbb{P}(A)$ を対応させる方法である．しかし，$\mathbb{P}(A)=0$ のとき，この方法は使えない．そこで，$A\in\mathscr{G}$ ではなく，各 $\omega\in\Omega$ に対して条件付き期待値を定義する．すなわち，条件付き期待値を確率変数により表す次の定義である．

定義 7.3 確率空間 $(\Omega,\mathscr{F},\mathbb{P})$ において，X を有限な期待値 $\mathbb{E}(X)$ をもつ確率変数，$\mathscr{G}$ を Ω 上の部分 σ-集合体とするとき，確率変数 Z が

(i) $\mathscr{G}$-可測であり，

(ii) 任意の $A\in\mathscr{G}$ に対して，$\mathbb{E}(Z1_A) = \mathbb{E}(X1_A)$,

を満たすならば，Z を $\mathbb{E}(X|\mathscr{G})$ と表し，X の $\mathscr{G}$ の下での条件付き期待値と呼ぶ．条件付き期待値は確率変数であるから $\omega\in\Omega$ の関数であり，$\mathbb{E}(X|\mathscr{G})(\omega)$ と書くこともある．また，$A\in\mathscr{F}$ に対して，

$$\mathbb{P}(A|\mathscr{G}) \equiv \mathbb{E}(1_A|\mathscr{G}),$$

により，$\mathbb{P}(A|\mathscr{G})$ を定義し，事象 A の $\mathscr{G}$ の下での条件付き確率と呼ぶ.

これらの定義は条件 (i) と (ii) を満たす Z が存在し唯 1 つのときのみ有効である．唯 1 つに関しては少し条件を緩める必要がある．初めに唯一つであることを確認しておこう．

補題 7.2 定義 7.3 の条件 (i) と (ii) を満たす確率変数 Z は確率 1 でただ一つである．すなわち，確率変数 Z, Z' が件 (i) と (ii) を満たすならば，$\mathbb{P}(Z=Z')=1$ である．

証明　$\mathbb{P}(Z = Z') = 1$ は $\mathbb{P}(Z \neq Z') = 0$ と同値であるから後者を示す．$A = \{\omega \in \Omega; Z > Z'\}$ とおくと，条件 (i) と補題 7.1 より，$A = \{Z - Z' \leq 0\} \in \mathscr{G}$ である．一方，条件 (ii) より，

$$\mathbb{E}(Z1_A) = \mathbb{E}(X1_A), \qquad \mathbb{E}(Z'1_A) = \mathbb{E}(X1_A),$$

であるから，$\mathbb{E}((Z-Z')1_A) = 0$ である．$\mathbb{P}(A) > 0$ と仮定すると，$\mathbb{E}((Z-Z')1_A) > 0$ であり，矛盾である．従って，$\mathbb{P}(Z > Z') = \mathbb{P}(A) = 0$ である．Z と Z' を入れ換えれば，$\mathbb{P}(Z' > Z) = 0$ である．従って，$\mathbb{P}(Z \neq Z') = 0$ である．■

この補題から (i) と (ii) を満たす Z があれば確率 1 で条件付き期待値が得られたことになる．このような Z を簡単な場合で考えてみよう．

例 7.3 (条件付き期待値の具体例)　$A \neq \varnothing$ を満たす $A \in \mathscr{F}$ に対して，$\mathscr{G} = \{\varnothing, A, A^c, \Omega\}$ であり，確率変数 X が有限な期待値 $\mathbb{E}(X)$ をもつとする．ここで，確率変数 Z が $\mathscr{G}$-可測であるから，$Z(\omega)$ は $\omega \in A$ に対して一定の値を取る，すなわち，Z は A 上で一定の値を取ると推測される．同様に，A^c 上でも一定の値を取ると推測する．これらの推測は簡単に証明できるが，ここでは Z を定義するための方針として使うだけなので証明は必要ない．以上のことをふまえ，$\mathbb{P}(C) > 0$ である $C \in \mathscr{F}$ に対して $\mathbb{E}(X|C) = \mathbb{E}(X1_C)/\mathbb{P}(C)$ とするとき，

$$x_1 = \begin{cases} \mathbb{E}(X|A), & \mathbb{P}(A) > 0, \\ 0, & \mathbb{P}(A) = 0, \end{cases} \qquad x_2 = \begin{cases} \mathbb{E}(X|A^c), & \mathbb{P}(A^c) > 0, \\ 0, & \mathbb{P}(A^c) = 0, \end{cases}$$

により，定数 x_1, x_2 を定義し，

$$Z(\omega) = x_1 1_A(\omega) + x_2 1_{A^c}(\omega)$$

により Ω から $\mathbb{R}$ への関数 Z を定義する．このとき，$1_A, 1_{A^c}$ は $\mathscr{G}$-可測であるから，補題 7.1 により Z は $\mathscr{G}$-可測であり，

$$\mathbb{E}(Z1_A) = x_1\mathbb{P}(A) = \mathbb{E}(X1_A), \qquad \mathbb{E}(Z1_{A^c}) = x_2\mathbb{P}(A^c) = \mathbb{E}(X1_{A^c}),$$

が成り立つ．よって，$Z = \mathbb{E}(X|\mathscr{G})$ である．特に，$\mathbb{P}(A) = 0$ の場合には，任意の定数 a に対して $x_1 = a$ としても条件 (i) と (ii) が成り立つ．この場合の Z を Z_a と表すと，$A \neq \varnothing$ より $Z \neq Z_a$ であるが，Z_a は (i) と (ii) を満たすので $Z_a = \mathbb{E}(X|\mathscr{G})$ である．従って，(i) と (ii) を満たす Z は ω の関数としては唯一つではない．しかし，$\mathbb{P}(Z = Z_a) = \mathbb{P}(A^c) = 1 - \mathbb{P}(A) = 1$，すなわち，確率 1 で Z_a は Z と一致する．■

この例は，条件付き期待値を具体的に作る方法を示唆している．条件付き期待値の定義は抽象的であるが，具体性をもった概念ともいえる．Z の存在について次の結果がある．

補題 7.3　定義 7.3 の条件 (i) と (ii) を満たす確率変数 Z が存在する．

補題 7.3 の証明は難しいので，$\mathcal{G}$ が可算個の集合から生成される場合についてのみ証明する．初めに，次の条件：

$$i \neq j \Rightarrow A_i \cap A_j = \emptyset, \qquad \cup_{i=1}^{\infty} A_i = \Omega, \tag{7.7}$$

を満たす $i = 1, 2, \ldots$ に対する $A_i \in \mathcal{F}$ に対して，Ω の部分集合の集まり $\mathcal{G}$ を

$$\mathcal{G} = \{\cup_{i \in B} A_i; B \subset \mathbb{N}\} \tag{7.8}$$

により定義する．ここに，$\mathbb{N} = \{1, 2, \ldots\}$ である．この $\mathcal{G}$ は Ω 上の σ-集合体である．

演習問題 7.5 (7.8) の $\mathcal{G}$ が Ω 上の σ-集合体であることを証明せよ．

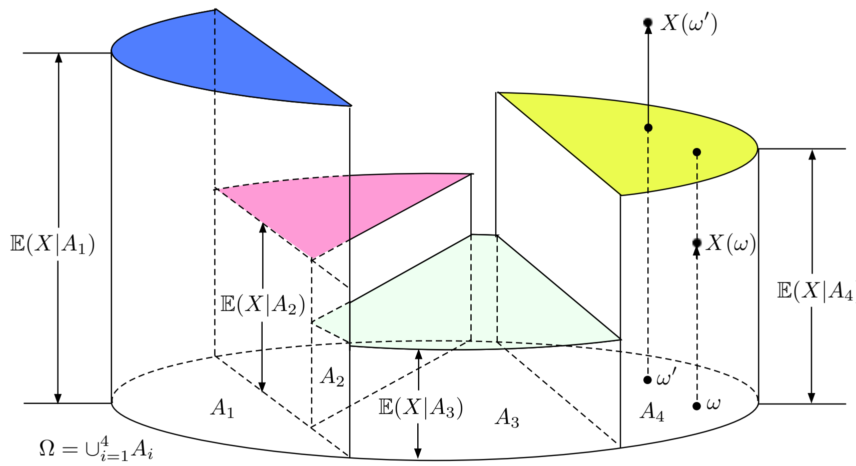


図 7.1: $\mathcal{G} = \{\cup_{i \in B} A_i; B \subset \{1, 2, 3, 4\}\}$ の場合の条件付き期待値 $Z \equiv \mathbb{E}(X|\mathcal{G})$ のイメージ

(7.8) の $\mathcal{G}$ に対する補題 7.3 の証明　確率変数 X が有限な期待値 $\mathbb{E}(X)$ をもつとする．例 7.3 と同様にして，条件付き期待値の定義の条件 (i) と (ii) を満たす確率変数 Z を具体的に作り存在を証明する．このために，

$$x_i = \begin{cases} \mathbb{E}(X|A_i), & \mathbb{P}(A_i) > 0, \\ 0, & \mathbb{P}(A_i) = 0, \end{cases} \qquad i = 1, 2, \ldots,$$

とおき，$Z(\omega) = \sum_{i=1}^{\infty} x_i 1_{A_i}(\omega)$ により Ω から $\mathbb{R}$ への関数を定義する（図 7.1 参照）．各 1_{A_i} は $\mathcal{G}$-可測であるから，Z は $\mathcal{G}$-可測である．また，$A \in \mathcal{G}$ はある $B \subset \mathbb{N}$ に対して $A = \cup_{i \in B} A_i$ と表すことができるので，

$$\begin{aligned} \mathbb{E}(Z1_A) &= \mathbb{E}\left(\sum_{k \in B} Z1_{A_k}\right) = \sum_{k \in B} \mathbb{E}\left(Z1_{A_k}\right) \\ &= \sum_{k \in B} x_k \mathbb{P}(A_k) \\ &= \sum_{k \in B} \mathbb{E}(X1_{A_k}) = \mathbb{E}(X \sum_{k \in B} 1_{A_k}) = \mathbb{E}(X1_A) \end{aligned}$$

が成り立つ．従って，Z は条件 (i) と (ii) を満たす確率変数である． ∎

$\mathcal{G}$ が一般の場合には，$\mathcal{G}$ を可算個の分割から生成された σ-集合体の極限として表し，具体的に計算できる条件付き期待値の極限として $\mathcal{G}$ の下での条件付き期待値の存在を証明することができる．しかし，この極限操作は難しいので証明を省く．

演習問題 7.6 確率変数 X は有限な期待値をもつとする．
(a) X が $\mathcal{G}$-可測ならば，$\mathbb{E}(X|\mathcal{G}) = X$ が確率 1 で成り立つことを示せ．
(b) X と $\mathcal{G}$ が独立ならば，$\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$ が確率 1 で成り立つことを示せ．

条件付き期待値は期待値同様に非負性と線型性をもつ．すなわち，定数 a, b，有限な期待値をもつ確率変数 X, Y と部分 σ-集合体 $\mathcal{G}$ に対して，

$$\mathbb{P}(X \geq 0) = 1 \text{ ならば } \mathbb{E}(X|\mathcal{G}) \geq 0, \tag{7.9}$$

$$\mathbb{E}(aX + bY|\mathcal{G}) = a\mathbb{E}(X|\mathcal{G}) + b\mathbb{E}(Y|\mathcal{G}), \tag{7.10}$$

が確率 1 で成り立つ．証明は条件付き期待値の定義に戻って確かめればよい．この際に役立つ結果を証明しておこう．

補題 7.4 有限な期待値をもつ確率変数 X, Y と部分 σ-集合体 $\mathcal{G}$ に対して，$\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(Y|\mathcal{G})$ が確率 1 で成り立つことは，次の条件が成り立つこと同値である．

$$\mathbb{E}(X 1_A) = \mathbb{E}(Y 1_A), \qquad \forall A \in \mathcal{G}. \tag{7.11}$$

証明 条件付き期待値の定義から (7.11) の必要性は明らかである．十分性を示すために，(7.11) が成り立つとする．このとき，条件付き期待値の定義から

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})1_A) = \mathbb{E}(X 1_A) = \mathbb{E}(Y 1_A) \qquad \forall A \in \mathcal{G}.$$

である．$\mathbb{E}(X|\mathcal{G})$ は $\mathcal{G}$-可測であり，Y の $\mathcal{G}$ の下での条件付き期待値 $\mathbb{E}(Y|\mathcal{G})$ は確率 1 で唯 1 つ決まることから，$\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(Y|\mathcal{G})$ が確率 1 で成り立つ． ∎

この結果を使うと (7.10) を確かめるためには

$$\mathbb{E}(\mathbb{E}(aX + bY|\mathcal{G})1_A) = \mathbb{E}((a\mathbb{E}(X|\mathcal{G}) + b\mathbb{E}(Y|\mathcal{G}))1_A), \qquad \forall A \in \mathcal{G}, \tag{7.12}$$

を示せばよい．条件付き期待値の定義に戻れば，この式の左辺は

$$\begin{aligned}\mathbb{E}((aX + bY)1_A) &= a\mathbb{E}(X 1_A) + b\mathbb{E}(Y 1_A) \\ &= a\mathbb{E}(\mathbb{E}(X|\mathcal{G})1_A) + b\mathbb{E}(\mathbb{E}(Y|\mathcal{G}))1_A)\end{aligned}$$

であるから，(7.12) を示すことができた．

7.5＊　最小 2 乗推定と条件付き期待値

条件付き期待値 $\mathbb{E}(X|\mathscr{G})$ には別の顔がある．これは，条件付き期待値を最小 2 乗法から求めるという考えである．与えられた情報は $\mathscr{G}$ に対して，確率変数 X を最小の誤差で推定する問題を考える．この推定値を確率変数 Y により表すと Y は $\mathscr{G}$-可測であり，推定の 2 乗誤差は $\mathbb{E}((Y-X)^2)$ である．この誤差を最小とする Y が最適な推定量と考えられる．このために，$\mathbb{E}(X^2)<\infty$ を仮定する．このとき，

$$
\begin{aligned}
\mathbb{E}((Y-X)^2) &= \mathbb{E}((Y-\mathbb{E}(X|\mathscr{G})+\mathbb{E}(X|\mathscr{G})-X)^2) \\
&= \mathbb{E}((Y-\mathbb{E}(X|\mathscr{G}))^2)+\mathbb{E}((\mathbb{E}(X|\mathscr{G})-X))^2)+2\mathbb{E}((Y-\mathbb{E}(X|\mathscr{G}))(\mathbb{E}(X|\mathscr{G})-X))
\end{aligned}
$$

である．ここで，Y が $\mathscr{G}$-可測であることから，$Y-\mathbb{E}(X|\mathscr{G})$ も $\mathscr{G}$-可測であるので，

$$
\begin{aligned}
\mathbb{E}((Y-\mathbb{E}(X|\mathscr{G}))(\mathbb{E}(X|\mathscr{G})-X)) &= \mathbb{E}(\mathbb{E}[(Y-\mathbb{E}(X|\mathscr{G}))(\mathbb{E}(X|\mathscr{G})-X)|\mathscr{G}]) \\
&= \mathbb{E}((Y-\mathbb{E}(X|\mathscr{G}))\mathbb{E}[(\mathbb{E}(X|\mathscr{G})-X)|\mathscr{G}]) \\
&= \mathbb{E}((Y-\mathbb{E}(X|\mathscr{G}))[\mathbb{E}(X|\mathscr{G})-\mathbb{E}(X|\mathscr{G})]) = 0
\end{aligned}
$$

である．従って，

$$
\mathbb{E}((Y-X)^2) = \mathbb{E}((Y-\mathbb{E}(X|\mathscr{G}))^2)+\mathbb{E}((\mathbb{E}(X|\mathscr{G})-X))^2).
$$

この式は，$Y=\mathbb{E}(X|\mathscr{G})$ のとき最小の値を取る．従って，条件付き期待値 $\mathbb{E}(X|\mathscr{G})$ とは X の $\mathscr{G}$ の下での最適な推定値である．$\mathbb{E}(X^2)<\infty$ の場合には，この事実を使って条件付き期待値 $\mathbb{E}(X|\mathscr{G})$ の存在を証明する方法がある．

以上のことから，$\mathbb{E}(X|\mathscr{G})$ は $\mathscr{G}$-可測の Y の中で 2 乗誤差 $\mathbb{E}((Y-X)^2)$ を最小とする確率変数である．従って，確率変数 X を

$$
X = \mathbb{E}(X|\mathscr{G}) + X - \mathbb{E}(X|\mathscr{G})
$$

と表すと，X が $\mathscr{G}$ の情報で予測できる部分 $\mathbb{E}(X|\mathscr{G})$ と予測できない部分 $X-\mathbb{E}(X|\mathscr{G})$ に分解できる．この予測できない部分は $\mathscr{G}$ の情報からは得ることができないランダムな要素であると見なすことができ，白色ノイズと呼ばれることもある．

8 マルチンゲール

実数値を状態に取る確率過程を $\{X_n; n \ge 0\}$，この確率過程の観測から得られる情報を含むフィルトレーションを $\mathbb{F} = \{\mathscr{F}_n; n \ge 0\}$ により表す．すなわち，$\{X_n; n \ge 0\}$ は $\mathbb{F}$ に適合しているとする．この状況において，各時刻における状態 X_n を $\mathscr{F}_n$ から予測できる項，すなわち，$\mathscr{F}_{n-1}$-可測である項と，予測できない項に分解したい．どのように分解すればよいか，この分解の役割が何かを考えることが本節の主題である．なお授業では∗のついた小節を省きます．

8.1 フィルトレーションと条件付き期待値

これまで条件付期待値を σ-集合体を与えた下で定義してきた．この定義を使って確率変数を条件とする条件付期待値を定義しておこう．

定義 8.1 確率変数 X, Y に対して，X の期待値が有限であるとき，$\mathbb{E}(X|\sigma(Y))$ を Y を与えた下での X の条件付き期待値と呼び，$\mathbb{E}(X|Y)$ と表す．ここに，$\sigma(Y) = \{\{Y \in B\}; B \in \mathscr{B}(\mathbb{R})\}$ とする．Y が複数の確率変数 $Y_1, Y_2, \dots, Y_n$ の列であるときには，同様に，

$$\mathbb{E}(X|Y_1, Y_2, \dots, Y_n) \equiv \mathbb{E}(X|\sigma(Y_1, Y_2, \dots, Y_n))$$

を，$Y_1, Y_2, \dots, Y_n$ を与えた下での X の条件付き期待値とする．

補題 8.1 $\mathbb{E}(X|Y)$ は Y の関数である，すなわち，Y の値で $\mathbb{E}(X|Y)$ の値が決まる．

証明 $\{Y = y\}$ 上で $\mathbb{E}(X|Y)$ が異なる値 x_1, x_2 を取ると仮定する．$\{\mathbb{E}(X|Y) = x_1, Y = y\} \in \sigma(Y)$ と $\{\mathbb{E}(X|Y) = x_2, Y = y\} \in \sigma(Y)$ より，ある $B_1, B_2 \in \sigma(Y)$ に対して，$\{\mathbb{E}(X|Y) = x_1, Y = y\} = \{Y \in B_1\}$ と $\{\mathbb{E}(X|Y) = x_2, Y = y\} = \{Y \in B_2\}$ が成り立つ．これらの式から $y \in B_1 \cap B_2$ であるが，$x_1 \ne x_2$ であるから，$B_1 \cap B_2 = \emptyset$ となり矛盾である．従って，$\{Y = y\}$ 上で $\mathbb{E}(X|Y)$ が異なる値を取ることはできない． ∎

条件付期待値をより小さな σ-集合体の下で条件付期待値を取ると単純にこの σ-集合体の下で条件付期待値に一致するという特性がある．すなわち，情報が少ないほど条件付き期待値は平均化される．この特性を証明しよう．

補題 8.2 X が有限な期待値をもつ確率変数であり，$\mathscr{G}_1, \mathscr{G}_2$ が $\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{F}$ を満たす Ω 上の σ-集合体ならば，

$$\mathbb{E}(\mathbb{E}(X|\mathscr{G}_2)|\mathscr{G}_1) = \mathbb{E}(X|\mathscr{G}_1) \tag{8.1}$$

が成り立つ．両辺の期待値を取れば，

$$\mathbb{E}(\mathbb{E}(X|\mathscr{G}_2)) = \mathbb{E}(X) \tag{8.2}$$

であり，特に，$\mathscr{G}_0 = \{\emptyset, \Omega\}$ とすると，$\mathbb{E}(X|\mathscr{G}_0) = \mathbb{E}(X)$ である．

証明　$\mathbb{E}(X|\mathscr{G}_2)$ の $\mathscr{G}_1$ の下での条件付き期待値の定義より，

$$\mathbb{E}(\mathbb{E}(\mathbb{E}(X|\mathscr{G}_2)|\mathscr{G}_1)1_A) = \mathbb{E}(\mathbb{E}(X|\mathscr{G}_2)1_A), \qquad A \in \mathscr{G}_1, \tag{8.3}$$

である．$\mathscr{G}_1 \subset \mathscr{G}_2$ より，$A \in \mathscr{G}_1$ ならば $A \in \mathscr{G}_2$ であるから，(8.3) の右辺は $\mathbb{E}(X1_A)$ に等しいので

$$\mathbb{E}(\mathbb{E}(\mathbb{E}(X|\mathscr{G}_2)|\mathscr{G}_1)1_A) = \mathbb{E}(X1_A) = \mathbb{E}(\mathbb{E}(X|\mathscr{G}_1)1_A), \qquad A \in \mathscr{G}_1,$$

が成り立つ．一方，$\mathbb{E}(\mathbb{E}(X|\mathscr{G}_2)|\mathscr{G}_1)$ は $\mathscr{G}_1$-可測であるから，条件付き確率が確率 1 で唯一つであることより，(8.1) が得られた．∎

(8.2) を使うと，条件付き期待値から期待値を計算できる．例えば，Y の分布を F とすると，$\mathbb{E}(X|Y)$ は Y の関数であるから，$Y=y$ のとき $\mathbb{E}(X|Y)$ を $\mathbb{E}(X|Y=y)$ と表すと，

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) = \int_{-\infty}^{+\infty} \mathbb{E}(X|Y=y)F(dy). \tag{8.4}$$

特に，Y が有限個の値 $y_1, y_2, \ldots, y_m$ のみを正の確率で取る確率変数であるならば，

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) = \sum_{i=1}^{m} \mathbb{E}(X|Y=y_i)\mathbb{P}(Y=y_i) \tag{8.5}$$

が得られる．これらの式は直接確認することもできるが，(8.2) がもつ意味を表していると考えることができる．なお，$\mathbb{E}(X|Y)$ を $\mathbb{E}(X|Y=y)$ は $\mathbb{P}(Y=y)=0$ の場合にも定義される．従って，$\mathbb{E}(X|Y=y)$ は $\mathbb{E}(X1(Y=y))/\mathbb{P}(Y=y)$ と異なることがある．

演習問題 8.1 (a) g を $[0,1]$ から $\mathbb{R}$ への連続関数とする．確率変数 Y が $[0,1]$ 上の一様分布に従うとき，(8.4) を用いて $\mathbb{E}(g(Y)) = \int_0^1 g(y)dy$ を証明せよ．
(b) $\mathbb{P}(Y=y) > 0$ ならば，$\mathbb{E}(X|Y=y) = \mathbb{E}(X1(Y=y))/\mathbb{P}(Y=y)$ を証明せよ．
(c) (b) の結果を用いて (8.5) を直接証明せよ．

8.2　マルチンゲールとセミマルチンゲール

ここで，実数値を状態に取る確率過程 $\{X_n; n \geq 0\}$ から予測できないランダムな要素を取り出す問題に戻る．$\mathbb{F} = \{\mathscr{F}_n; n \geq 0\}$ をフィルトレーションとし，$\{X_n; n \geq 0\}$ は $\mathbb{F}$ に適合しているとする．このとき，7.5 節の結果から $\mathbb{E}(X_n|\mathscr{F}_{n-1})$ は $\mathscr{F}_{n-1}$ の下での X_n の最適な推定値であるから，

$$X_n = \mathbb{E}(X_n|\mathscr{F}_{n-1}) + X_n - \mathbb{E}(X_n|\mathscr{F}_{n-1}), \qquad n \geq 0, \tag{8.6}$$

と分解し，予測できない部分 $X_n - \mathbb{E}(X_n|\mathscr{F}_{n-1})$ を取り出すことが考えられる．一方，予測できない部分を累積した量を M_n により表すと

$$M_n = M_0 + \sum_{\ell=1}^{n}(X_\ell - \mathbb{E}(X_\ell|\mathscr{F}_{\ell-1})), \qquad n \geq 0, \tag{8.7}$$

である．ここに，M_0 は $\mathscr{F}_0$-可測とする．$\{M_n; n\geq 0\}$ は純粋にランダムな要素を表す確率過程と見ることができる．ここで，

$$\begin{aligned}\mathbb{E}(M_n|\mathscr{F}_{n-1}) &= M_0 + \sum_{\ell=1}^{n-1}(X_\ell - \mathbb{E}(X_\ell|\mathscr{F}_{\ell-1})) + \mathbb{E}((X_n - \mathbb{E}(X_n|\mathscr{F}_{n-1}))|\mathscr{F}_{n-1})) \\ &= M_{n-1} - (\mathbb{E}(X_n|\mathscr{F}_{n-1}) - \mathbb{E}(\mathbb{E}(X_n|\mathscr{F}_{n-1})|\mathscr{F}_{n-1})) = M_{n-1}\end{aligned}$$

が成り立つことに注目し，次の定義を行う．

定義 8.2 実数値を値取る確率過程 $\{M_n; n\geq 0\}$ が $\mathbb{F}$ に適合し，すべての $n\geq 0$ に対して $\mathbb{E}(M_n)$ が有限であり，

$$\mathbb{E}(M_n|\mathscr{F}_{n-1}) = M_{n-1}, \qquad n\geq 0, \tag{8.8}$$

を満たすならば，$\{M_n; n\geq 0\}$ を $\mathbb{F}$-マルチンゲール (martingale) と呼ぶ．条件 (8.8) において，等号を $\geq$ に置き換えた式，すなわち，$\mathbb{E}(M_n|\mathscr{F}_{n-1}) \geq M_{n-1}$ が成り立つならば，劣マルチンゲール (submartingale)，$\leq$ に置き換えた式，すなわち，$\mathbb{E}(M_n|\mathscr{F}_{n-1}) \leq M_{n-1}$ が成り立つならば，優マルチンゲール (supermartingale) と呼ぶ．

演習問題 8.2 $U_1, U_2, \ldots$ は独立で同一の分布に従い有限な期待値をもつ確率変数の列とする．$X_0 = 0$，$n\geq 1$ に対して $X_n = \sum_{\ell=1}^n (U_\ell - \mathbb{E}(U_\ell))$ により確率過程 $X_\bullet = \{X_n; n\geq 0\}$ を定義する．$\mathscr{F}_n^X = \sigma(X_0, \ldots, X_n)$ とおき $\mathbb{F}^X = \{\mathscr{F}_n^X; n\geq 0\}$ とするとき，$X_\bullet$ が $\mathbb{F}^X$-マルチンゲールであることを示せ．

定義 8.3 確率過程 $\{Y_n; n\geq 0\}$ は，Y_0 が $\mathscr{F}_0$-可測であり，各 $n\geq 1$ に対して Y_n が $\mathscr{F}_{n-1}$ 可測であるならば，$\mathbb{F}$-予測可能であるという．

X_n を Y_n を用いて分解することを考えてみよう．一般にこのような分解は 1 つとは限らない．しかし，予測できない部分にマルチンゲールを使うと確率 1 で唯 1 つ決まる．これが次の結果である．

補題 8.3 $\{X_n; n\geq 0\}$ を $\mathbb{F}$ に適合した実数値確率過程とし，すべての $n\geq 0$ に対して $\mathbb{E}(X_n)$ が有限であるならば，

$$X_n = Y_n + M_n, \qquad n\geq 0, \tag{8.9}$$

を満たす $\mathbb{F}$-予測可能な確率過程 $\{Y_n; n\geq 0\}$ と $\mathbb{F}$-マルチンゲール $\{M_n; n\geq 0\}$ が存在し，$M_0 = 0$ が与えられたとき，これらの確率過程は確率 1 で唯一つ定まる．

証明 これまで述べてきたことから，$M_0 = 0$ とし，(8.7) により $\{M_n; n\geq 0\}$ を定義すれば，$\mathbb{F}$-マルチンゲールである．$n\geq 0$ に対して $Y_n = X_n - M_n$ により確率過程 $\{Y_n; n\geq 0\}$ を定義する．(8.7) より，

$$Y_n = X_n - \sum_{\ell=1}^{n}(X_\ell - \mathbb{E}(X_\ell|\mathscr{F}_{\ell-1})) = \mathbb{E}(X_n|\mathscr{F}_{n-1}) - \sum_{\ell=1}^{n-1}(X_\ell - \mathbb{E}(X_\ell|\mathscr{F}_{\ell-1}))$$

であるから，Y_n は $\mathscr{F}_{n-1}$-可測である．従って，$\{Y_n; n \geq 0\}$ は $\mathbb{F}$-予測可能な確率過程であり，条件を満たす 2 つの確率過程が得られた．確率 1 で唯一つであることを示すために，$X_n = Y'_n + M'_n$ をみたす $\mathbb{F}$-予測可能な $\{Y'_n; n \geq 0\}$ と $\mathbb{F}$-マルチンゲール $\{M'_n; n \geq 0\}$ があるとする．このとき，$Y_n + M_n = X_n = Y'_n + M'_n$ であるから，

$$Y_n - Y'_n = M'_n - M_n, \qquad n \geq 0, \tag{8.10}$$

が成り立つ．両辺の $\mathscr{F}_{n-1}$ の下での条件付き期待値を取ると，Y_n, Y'_n が $\mathscr{F}_{n-1}$ 可測であることと M_n, M'_n が $\mathbb{F}$-マルチンゲールであることから，

$$Y_n - Y'_n = \mathbb{E}(M'_n - M_n | \mathscr{F}_{n-1}) = M'_{n-1} - M_{n-1}, \qquad n \geq 1,$$

が成り立つ．従って，$M'_n - M_n = M'_{n-1} - M_{n-1}$．この式を繰り返し使うと

$$Y_n - Y'_n = M_n - M'_n = M'_0 - M_0 = 0$$

が成り立つので，$Y_n = Y'_n$，$M'_n = M_n$ となり，確率 1 での唯一性が言えた．　■

定義 8.4 実数値を取る確率過程 $\{X_n; n \geq 0\}$ に対して，(8.9) を満たす $\mathbb{F}$-予測可能な確率過程 $\{Y_n; n \geq 0\}$ と $M_0 = 0$ を満たす $\mathbb{F}$-マルチンゲール $\{M_n; n \geq 0\}$ が存在するならば，$\{X_n; n \geq 0\}$ を $\mathbb{F}$-セミマルチンゲール (semimartingale) と呼ぶ．

補題 8.3 より，$\{X_n; n \geq 0\}$ が $\mathbb{F}$ に適合し，各 $n \geq 0$ に対して有限な期待値 $\mathbb{E}(X_n)$ をもつならば，$\mathbb{F}$-セミマルチンゲールである．

8.3 停止時刻と随意選択 (Optional sampling) 定理

確率過程の応用においては，観測結果に基づいて値が決まる時刻がある．この時刻を τ とする．τ は確率変数であるが $+\infty$ も値に取るとする．これは有限の時間内に決まらない場合を含めたいためである．このような τ を定義しよう．

定義 8.5 確率変数 τ が $\mathbb{N} \cup \{\infty\}$ に値を取り，条件：

$$\{\tau \leq n\} \in \mathscr{F}_n, \qquad n \geq 0, \tag{8.11}$$

を満たすならば $\mathbb{F}$-停止時刻 (stopping time) と呼ぶ．この条件は時刻 n までの情報で $\tau \leq n$ がであるかないかを判断できることを表している．

演習問題 8.3 τ を $\mathbb{N} \cup \{\infty\}$ に値を取る確率変数とする．$\forall n \in \mathbb{N}$ に対して $\{\tau > n\} \in \mathscr{F}_n$ ならば，τ が $\mathbb{F}$-停止時刻であることを示せ．

定義 8.6 $\mathbb{F}$-停止時刻 τ に対して，

$$\mathscr{F}_\tau \equiv \{A \in \mathscr{F}; \forall n \in [0, \infty], A \cap \{\tau \leq n\} \in \mathscr{F}_n\} \tag{8.12}$$

を $\mathbb{F}$-停止 σ 集合体と呼ぶ．ここに，$\mathscr{F}_\infty = \bigcup_{n=0}^{\infty} \mathscr{F}_n$ とする．

演習問題 8.4 (a) $\mathscr{F}_\infty$ が $\mathscr{F}$ の部分 σ-集合体であることを示せ．
(b) τ が $\mathbb{F}$-停止時刻のとき，$\mathscr{F}_\tau$ が $\mathscr{F}$ の部分 σ-集合体であることを示せ．

補題 8.4 τ, η が $\mathbb{F}$-停止時刻ならば，$\tau \wedge \eta$，$\tau \vee \eta$，$\tau + \eta$ も $\mathbb{F}$-停止時刻である．ここに，$a \wedge b = \min(a, b)$，$a \vee b = \max(a, b)$ とする．特に，整数の定数 $k \geq 1$ は $\mathbb{F}$-停止時刻であるから，$\tau \wedge k$，$\tau \vee k$，$\tau + k$ も $\mathbb{F}$-停止時刻である．

証明 $n \geq 0$ に対して $\{\tau \leq n\} \in \mathscr{F}_n$，$\{\eta \leq n\} \in \mathscr{F}_n$ であるから，

$$\{\tau \wedge \eta \leq n\} = \{\tau \leq n\} \cup \{\eta \wedge n\} \in \mathscr{F}_n$$

より，$\tau \wedge \eta$ は $\mathbb{F}$-停止時刻である．$\tau \vee \eta$，$\tau + \eta$ も同様に確かめられる． ∎

$k \geq 0$ に対して $\tau + k$ は $\mathbb{F}$-停止時刻であるが，

$$\{\tau - k \leq n\} = \{\tau \leq n + k\}$$

より，$\tau - k$ は必ずしも $\mathbb{F}$-停止時刻ではない．同様に，$\mathbb{F}$-停止時刻 τ, η に対して，$\tau - \eta$ は必ずしも $\mathbb{F}$-停止時刻ではない．

停止時刻の代表的な例はフィルトレーション $\mathbb{F}$ に適合した確率過程 $\{X_n; n \geq 0\}$ が状態の集合 $B \in \mathscr{B}(S)$ に到達する時刻である．この時刻を τ_B とすると，

$$\tau_B = \inf\{n \geq 1; X_n \in B\}$$

であり，τ_B が $\mathbb{F}$-停止時刻であることを確認できる．ここに，すべての $n \geq 0$ に対して $X_n \notin B$ ならば，$\tau_B = \infty$ である．例えば，$\{X_n; n \geq 0\}$ がランダムウォークで，$\mathbb{E}(X_{n+1} - X_n) > 0$ ならば，$X_0 = 0$ のとき，$B = (-\infty, 0]$ に対して，$\tau_B = \infty$ も起こりうる．このように，停止時刻は ∞ となる可能性を常に考慮する必要がある．

定理 8.1 (随意選択 (optional sampling) 定理) $\{M_n; n \geq 0\}$ が $\mathbb{F}$-マルチンゲールであり，$\mathbb{F}$-停止時刻 τ が有界，すなわち，ある非負の整数 m があり，確率 1 で $\tau \leq m$ ならば，

$$\mathbb{E}(M_\tau | \mathscr{F}_0) = M_0, \qquad n \geq 0, \tag{8.13}$$

が確率 1 で成り立つ．

証明 $\tau = 0$ のときは明らかに成り立つので，$1 \leq \tau \leq m$ とする．このとき，

$$\begin{aligned} M_\tau &= \sum_{\ell=0}^{\tau} M_\ell - \sum_{\ell=0}^{\tau-1} M_\ell = \sum_{\ell=0}^{m} M_\ell 1(\tau \geq \ell) - \sum_{\ell=0}^{m-1} M_\ell 1(\tau \geq \ell + 1)) \\ &= \sum_{\ell=1}^{m} (M_\ell - M_{\ell-1}) 1(\tau \geq \ell) + M_0 \end{aligned} \tag{8.14}$$

です．ここで，$\{\tau \geq \ell\} = \Omega \setminus \{\tau \leq \ell - 1\} \in \mathscr{F}_{\ell-1}$ であるから，$\{M_n; n \geq 0\}$ が $\mathbb{F}$-マルチンゲールであることより，$\ell \geq 1$ に対して，

$$\begin{aligned}\mathbb{E}\left((M_\ell - M_{\ell-1})1(\tau \geq \ell)|\mathscr{F}_0\right) &= \mathbb{E}\left(\mathbb{E}[(M_\ell - M_{\ell-1})1(\tau \geq \ell)|\mathscr{F}_{\ell-1}]|\mathscr{F}_0\right)\\ &= \mathbb{E}\left((\mathbb{E}[M_\ell|\mathscr{F}_{\ell-1}] - M_{\ell-1})1(\tau \geq \ell)|\mathscr{F}_0\right) = 0\end{aligned}$$

が成り立つ．従って，$\mathscr{F}_0$ の下で (8.14) の両辺の条件付き期待値を取ると (8.13) が得られる．∎

演習問題 8.5 (a) $\{M_n; n \geq 0\}$ を $\mathbb{F}$-マルチンゲール，τ を $\mathbb{F}$-停止時刻とする．$n \geq 0$ に対して $X_n = M_{\tau\wedge n}$ とするならば，$\{X_n; n \geq 0\}$ は $\mathbb{F}$-マルチンゲールであることを示せ．
(b) (a) を用いて定理 8.1 を証明せよ．

例 8.1 確率変数列 $U_1, U_2, \ldots$ が独立で同一の分布に従い，$\mathbb{P}(U_n = 1) = \mathbb{P}(U_n = -1) = \frac{1}{2}$ を満たすとする．このとき，$X_0 = 0$, $n \geq 1$ に対して $X_n = U_1 + U_2 + \ldots + U_n$ により，ランダムウォーク $\{X_n; n \geq 0\}$ を定義する．この X_n は偏りのないコインを n 回投げたときの表が出た回数から裏が出た回数を引いた結果と見ることもできる．$\mathscr{F}_n = \sigma(\{X_\ell; 0 \leq \ell \leq n\})$ とし，$\mathbb{F} = \{\mathscr{F}_n; n \geq 0\}$ によりフィルトレーション $\mathbb{F}$ を定義する．U_n が $\mathscr{F}_{n-1}$ と独立であることから，

$$\mathbb{E}(X_n|\mathscr{F}_{n-1}) = \mathbb{E}(X_{n-1} + U_n|\mathscr{F}_{n-1}) = X_{n-1} + \mathbb{E}(U_n|\mathscr{F}_{n-1}) = X_{n-1}, \qquad n \geq 1,$$

が成り立つ．また，$\mathbb{E}(|X_n|) < \infty$ であるから，$\{X_n; n \geq 0\}$ は $\mathbb{F}$-マルチンゲールである．このマルチンゲールに対して，

$$\tau = \inf\{n \geq 1; X_n \geq 2X_0\}$$

により τ を定義すれば，τ は停止時刻であり，τ は確率 1 で有限であることが知られている．X_n をコイン投げによる賭の時刻 n での所持金とすると，$X(0) \geq 1$ ならば，τ は所持金が 2 倍となる時刻である．$X(0) = a \geq 1$ を最初の所持金とすると，$\mathbb{E}(X_\tau) = 2a$ である．τ は有界ではないが，定理 8.1 が成り立つとすると

$$\mathbb{E}(X_\tau) = \mathbb{E}(X_0) = a$$

となり，$\mathbb{E}(X_\tau) = 2a$ と矛盾する．従って，この場合定理 8.1 は成り立たない．∎

上記の例から，定理 8.1 において τ が有界であるという条件は一般に外すことができない．これは応用において大きな制約である．そこで，停止時刻 τ を非負の整数 n を用いて $\tau \wedge n \equiv \min(\tau, n)$ で置き換えることを考える．このとき，$\tau \wedge n$ は有界であり，補題 8.4 より $\mathbb{F}$-停止時刻である．従って，定理 8.1 において τ を $\tau \wedge n$ に置き換えると，

$$\mathbb{E}(M_{\tau\wedge n}|\mathscr{F}_0) = M_0, \qquad n \geq 0, \tag{8.15}$$

が確率 1 で成り立つ．

一方，$\mathbb{P}(\tau < \infty) = 1$ ならば，確率 1 で $\lim_{n\to\infty} M_{\tau\wedge n} = M_\tau$ であるから，定理 8.1 を次のように拡張することができる．

系 8.1 $\{M_n; n \ge 0\}$ を $\mathbb{F}$-マルチンゲール，τ を確率 1 で有限な $\mathbb{F}$-停止時刻とする．ある有限な期待値をもつ確率変数 Y に対して $|M_n| \le Y$ がすべての $0 \le n \le \tau$ に対して成り立つならば，(8.13) が成り立つ．

証明 (8.15) が確率 1 で成り立つことは，

$$\mathbb{E}(M_{\tau\wedge n}1_A) = \mathbb{E}(\mathbb{E}(M_{\tau\wedge n}|\mathscr{F}_0)1_A) = \mathbb{E}(M_0 1_A), \qquad \forall A \in \mathscr{F}_0, n \ge 0,$$

が成り立つことに等しい．仮定より $|M_{\tau\wedge n}1_A| \le Y$ であるから，有界収束定理により

$$\lim_{n\to\infty} \mathbb{E}(M_{\tau\wedge n}1_A) = \mathbb{E}(\lim_{n\to\infty} M_{\tau\wedge n}1_A) = \mathbb{E}(M_\tau 1_A)$$

であるから，$\mathbb{E}(M_\tau 1_A) = \mathbb{E}(M_0 1_A)$ が成り立つ．従って，(8.13) が確率 1 で成り立つ． ∎

系 8.1 において，停止時刻 τ が確率 1 で有限という条件は取り除くことができる．しかし，その証明には M_n の極限が存在することが必要である．この証明には，いくつかの準備が必要なので節を小節に分けて行う．その前に，系 8.1 が役立つ例を示そう．

例 8.2 確率過程 $\{X_n; n \ge 0\}$ を例 8.1 で定義したコイン投げから作られたランダムウォーク，$\mathbb{F}$ をフィルトレーションとする．$a \le 0 < b$ を満たす整数 a, b に対して，$n \ge 1$ に対して X_n が a に到達する前に b に到達する確率を求めたい．この問題は $-a$ の資金をもってコインを投げ，表ならば 1 ドルを得て裏ならば 1 ドルを失う賭を行うとき，資金が尽きる前に $b - a$ ドルに到達する確率を求めることに等しい．この問題を定理 8.1 を適用して解くために，$k = a, b$ に対して τ_k を

$$\tau_k = \inf\{n \ge 1; X_n = k\},$$

により定義すると，τ_a, τ_b は $\mathbb{F}$-停止時刻である．従って，$\tau \equiv \tau_a \wedge \tau_b$ も $\mathbb{F}$-停止時刻である．更に，$\mathbb{E}(X_n) = 0$，$V(X_n) = 2$ であり，$\frac{a}{2\sqrt{n}} \vee \frac{b}{2\sqrt{n}} \le |\varepsilon|$ を満たす $\varepsilon > 0$ に対して，

$$\begin{aligned}\mathbb{P}(\tau > n) &= \mathbb{P}(a < X_\ell < b, \forall \ell[0,n]) \le \mathbb{P}(a < X_n < b) \\ &= \mathbb{P}\left(\frac{a}{2\sqrt{n}} < \frac{X_n}{2\sqrt{n}} < \frac{b}{2\sqrt{n}}\right) \le \mathbb{P}\left(-\varepsilon < \frac{X_n}{2\sqrt{n}} < \varepsilon\right)\end{aligned}$$

が成り立つ．従って，中心極限定理により，

$$\lim_{n\to\infty} \mathbb{P}(\tau > n) \le \lim_{n\to\infty} \mathbb{P}\left(-\varepsilon < \frac{X_n}{2\sqrt{n}} < \varepsilon\right) = \int_{-\varepsilon}^{\varepsilon} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \le 2\varepsilon \downarrow 0$$

である．一方，$\{\tau \le n\} \uparrow \cup_{n=1}^{\infty} \{\tau \le n\} = \{\tau < \infty\}$ であるから，

$$\mathbb{P}(\tau < \infty) = \lim_{n\to\infty} \mathbb{P}(\tau \le n) = 1 - \lim_{n\to\infty} \mathbb{P}(\tau > n) = 1.$$

また，$n \le \tau$ のとき $|X_n| \le |a| \vee b$ であるから，マルチンゲール $\{X_n; n \ge 0\}$ と $\mathbb{F}$-停止時刻 τ に対して系 8.1 を適用することができる．従って，$X_\tau = a1(\tau_a < \tau_b) + b1(\tau_b < \tau_a)$ より，

$$a\mathbb{P}(\tau_a < \tau_b) + b\mathbb{P}(\tau_b < \tau_a) = \mathbb{E}(X_0) = 0.$$

一方，$\mathbb{P}(\tau_a < \tau_b) + \mathbb{P}(\tau_b < \tau_a) = \mathbb{P}(\tau < \infty) = 1$ であるから，

$$\mathbb{P}(\tau_b < \tau_a) = \frac{-a}{b - a}$$

は，X_n が a に到達する前に b に到達する確率である． ∎

8.4＊ 凸関数と劣マルチンゲール

$\{M_n; n \geq 1\}$ が $\mathbb{F}$-マルチンゲールならば，

$$\mathbb{E}(|M_n||\mathscr{F}_{n-1}) \geq |\mathbb{E}(M_n|\mathscr{F}_{n-1})| = |M_{n-1}|$$

が確率 1 で成り立つので，$\{|M_n|; n \geq 1\}$ は $\mathbb{F}$-劣マルチンゲールである．絶対値をより一般的な関数 g で置き換えたとき，$\{g(M_n); n \geq 1\}$ は $\mathbb{F}$-劣マルチンゲールなるような関数を定義しよう．

定義 8.7 f を $\mathbb{R}$ の $\mathbb{R}$ から $\mathbb{R}$ への関数とする．$p + q = 1$ を満たす任意の $p, q > 0$ に対して，以下の条件を満たす f を凸関数と呼ぶ．

$$f(px + qy) \leq pf(x) + qf(y), \qquad \forall x, y \in \mathbb{R}. \tag{8.16}$$

次の結果は凸関数のグラフがグラフ上の各点における接線上にあることを示している．

補題 8.5 f が凸関数であるとは，$\forall x \in \mathbb{R}$ に対してある $c_x \in \mathbb{R}$ があり次の条件が成り立つことに等しい．

$$c_x(y - x) \leq f(y) - f(x), \qquad \forall y \in \mathbb{R}. \tag{8.17}$$

証明 f が凸関数であるとする．初めに $x < y$ とする．このとき，$p + q = 1$ を満たす $p, q > 0$ に対して，$w = y - p(y - x)$ とおくと

$$p = \frac{y - w}{y - x}, \qquad q = \frac{w - x}{y - x}, \qquad px + qy = w,$$

であるから，(8.16) は，$x < w < y$ を満たす w に対して

$$f(w) \leq \frac{y - w}{y - x}f(x) + \frac{w - x}{y - x}f(y), \tag{8.18}$$

が成り立つことに等しい．この式は

$$(y - x)f(w) \leq (y - x - (w - x))f(x) + (w - x)f(y),$$

に等しいので，(8.16) は，$x < w < y$ を満たす w に対する

$$\frac{f(w) - f(x)}{w - x} \leq \frac{f(y) - f(x)}{y - x}. \tag{8.19}$$

と同値である．ここで，$g(w) = \dfrac{f(w) - f(x)}{w - x}$ とおくと，$g(w)$ は $w > x$ に対して非減少である．一方，$z < x < w$ を満たす z, x, w に対して，(8.19) と同様に

$$\frac{f(x) - f(z)}{x - z} \leq \frac{f(w) - f(x)}{w - x} = g(w). \tag{8.20}$$

が成り立つので，$g(w)$ は非減少であり，$w \downarrow x$ のとき下に有界であるから，$c_x = \lim_{w\downarrow x} g(w)$ が存在し有限である．よって，(8.19) から (8.17) が得られた．次に $x > y$ とする．この場合は (8.16) が x と y を入れ換えても成り立つことから，$y < x < w$ を満たす y, z, w と (8.20) において $z = y$ を代入した式から，

$$\frac{f(y) - f(x)}{y - x} = \frac{f(x) - f(y)}{x - y} \le \frac{f(w) - f(x)}{w - x} = g(w).$$

$y - x < 0$ であるから，この式より

$$(y - x)g(w) \le f(y) - f(x)$$

従って，$w \downarrow x$ とすれば，(8.17) が得られる．

逆に (8.17) が成り立つとする．(8.17) において x に $w \equiv px + qy$ を代入すると，

$$c_w(y - w) \le f(y) - f(px + qy) \tag{8.21}$$

であり，(8.17) の x, y を入れ換えた式 $c_y(x - y) \le f(x) - f(y)$ において，y に $w \equiv px + qy$ を代入すると，

$$c_w(x - w) \le f(x) - f(px + qy) \tag{8.22}$$

が得られる．(8.21) に q をかけ (8.22) に p をかけ，各辺を加えると，左辺の和は

$$qc_w(y - w) + pc_w(x - w) = c_w(q(p(y - x) + p(q(x - y))) = 0$$

であるから，$p + q = 1$ より (8.16) が得られる． ∎

確率変数を変数とする凸関数の期待値について一般的に成り立つ不等式がある．

補題 8.6 (Jensen) 確率変数 X と $\mathbb{R}$ 上の凸関数 f に対して, $\mathbb{E}(X)$ と $\mathbb{E}(f(X))$ が有限ならば

$$f(\mathbb{E}(X)) \le \mathbb{E}(f(X)). \tag{8.23}$$

証明 補題 8.5 において $x = \mathbb{E}(X)$ とすれば,

$$f(X) - f(\mathbb{E}(X)) \ge c_{\mathbb{E}(X)}(X - \mathbb{E}(X))$$

である．この式の両辺の期待値を取ると

$$\mathbb{E}(f(X) - f(\mathbb{E}(X))) \ge c_{\mathbb{E}(X)}\mathbb{E}(X - \mathbb{E}(X)) = 0.$$ ∎

次の結果はマルチンゲールの凸関数による写像が劣マルチンゲールとなることを示している．

補題 8.7 f を $\mathbb{R}$ 上の凸関数とする．$X_\bullet \equiv \{X_n; n \geq 0\}$ が $\mathbb{F}$-マルチンゲールであり，すべての $n \geq 0$ に対して $\mathbb{E}(X_n)$ と $\mathbb{E}(f(X_n))$ が有限ならば，$\{f(X_n); n \geq 0\}$ は $\mathbb{F}$-劣マルチンゲールである．

証明 $Y_n = \mathbb{E}(X_n|\mathscr{F}_{n-1})$ とおくと補題 8.5 より

$$f(X_n) - f(Y_n) \geq c_{Y_n}(X_n - Y_n)$$

$\mathscr{F}_{n-1}$ の下でこの式の条件付き期待値を取ると，Y_n が $\mathscr{F}_{n-1}$ であることから，

$$\mathbb{E}(f(X_n)|\mathscr{F}_{n-1}) - f(\mathbb{E}(X_n|\mathscr{F}_{n-1})) \geq c_{Y_n}\mathbb{E}(X_n - Y_n|\mathscr{F}_{n-1}) = 0.$$

一方，$X_\bullet$ が $\mathbb{F}$-マルチンゲールであることから，$\mathbb{E}(X_n|\mathscr{F}_{n-1}) = X_{n-1}$ である．よって，上記の式から

$$\mathbb{E}(f(X_n)|\mathscr{F}_{n-1}) \geq f(X_{n-1}),$$

が得られ，$\{f(X_n)\}$ は $\mathbb{F}$-劣マルチンゲールである．■

補題 8.7 において，$X_\bullet$ を $\mathbb{F}$-劣マルチンゲールに置き換えることは一般にできないが，f に条件を追加することにより可能となる．

補題 8.8 $X_\bullet$ が $\mathbb{F}$-劣マルチンゲールであり，f が非減少な凸関数ですべての $n \geq 0$ に対して $\mathbb{E}(|f(X_n)|) < \infty$ を満たすならば，$\{f(X_n); n \geq 0\}$ は $\mathbb{F}$-劣マルチンゲールである．特に，各 $a \in \mathbb{R}$ に対して，$\{a + (X_n - a)^+; n \geq 0\}$ は $\mathbb{F}$-劣マルチンゲールである．

証明 補題 8.7 の証明から，

$$\mathbb{E}(f(X_n)|\mathscr{F}_{n-1}) - f(\mathbb{E}(X_n|\mathscr{F}_{n-1})) \geq 0,$$

が得られる．ここで，$X_\bullet$ が $\mathbb{F}$-劣マルチンゲールであることから，$\mathbb{E}(X_n|\mathscr{F}_{n-1}) \geq X_{n-1}$ であるから，f が非減少であることにより，$\mathbb{E}(f(X_n)|\mathscr{F}_{n-1}) \geq f(X_{n-1})$ が成り立つ．この不等式と上記の不等式より，$\{f(X_n)\}$ は $\mathbb{F}$-劣マルチンゲールである．■

8.5∗　マルチンゲールの極限と随意選択定理の拡張

マルチンゲール M_n の期待値は一定であるが，純粋に（偏りのない）ランダムな量の累積である．しかし，正の部分 $M_n^+ \equiv M_n \vee 0$ の期待値が有界ならば，$n \to \infty$ のとき M_n がある確率変数 X_∞ に概収束する．これは直感的に理解しがたい結果である．この結果を証明することが本節の目的である．初めに，予測可能な確率過程のマルチンゲールによる離散的な積分がマルチンゲールと成ることを示す．

補題 8.9 (離散的積分) $X_\bullet$ が $\mathbb{F}$-マルチンゲール（または，劣（優）マルチンゲール）ならば，有界で $\mathbb{F}$-予測可能な確率過程 $H_\bullet \equiv \{H_n; n \geq 0\}$ に対して，

$$(H \cdot X)_0 = 0, \qquad (H \cdot X)_n = \sum_{\ell=1}^{n} H_\ell (X_\ell - X_{\ell-1}), \quad n \geq 1.$$

により $(H \cdot X)_\bullet \equiv \{H \cdot X)_n; n \geq 0\}$ を定義すると，$(H \cdot X)_\bullet$ は $\mathbb{F}$-マルチンゲール（$\forall n \geq 0$ に対して $H_n \geq 0$ のとき $\mathbb{F}$-劣（優）マルチンゲール）である．

証明 H_n が有界であるから，$\mathbb{E}((H \cdot X)_n)$ は有限である．$n \geq 1$ に対して，$(H \cdot X)_n = (H \cdot X)_{n-1} + H_n(X_n - X_{n-1})$ であり，H_n は $\mathscr{F}_{n-1}$-可測であるから，$X_\bullet$ が $\mathbb{F}$-マルチンゲールであることより，

$$\mathbb{E}((H \cdot X)_n | \mathscr{F}_{n-1}) = (H \cdot X)_{n-1} + H_n \mathbb{E}\left((X_n - X_{n-1}) | \mathscr{F}_{n-1}\right) = (H \cdot X)_{n-1}. \tag{8.24}$$

従って，$(H \cdot X)_\bullet$ は $\mathbb{F}$-マルチンゲールである．$X_\bullet$ が $\mathbb{F}$-劣マルチンゲールであり，$\forall n \geq 1$ に対して $H_n \geq 0$ ならば，

$$H_n \mathbb{E}\left((X_n - X_{n-1}) | \mathscr{F}_{n-1}\right) = H_n \left(\mathbb{E}(X_n | \mathscr{F}_{n-1}) - X_{n-1}\right) \geq 0$$

であるから，(8.24) の 2 番目の等号を "$\geq$" に置き換えた式が成り立つので，$(H \cdot X)_\bullet$ は $\mathbb{F}$-劣マルチンゲールである．優マルチンゲールの場合も同様に証明できる．■

マルチンゲール X_n の概収束を調べるために，X_n の値がどの程度ばらつくかを区間 (a, b)（$a < b$ とする）を下から上に横切る回数により評価する．$a < b$ である任意の a, b に対して，$n \to \infty$ のとき横切る回数が 0 に近づけば X_n が収束することが期待される．

$a < b$ を満たす定数 a, b に対して，$N_0 = -1$，$k \geq 1$ に対して

$$N_{2k-1} = \inf\{\ell > N_{2k-2}; X_\ell \leq a\}, \quad N_{2k} = \inf\{\ell > N_{2k-1}; X_\ell \geq b\},$$

により，上から a 以下になる回数 N_{2k-1} と下から b 以上になる回数 N_{2n} を定義する．このとき，$U_n = \sup\{k \geq 1; N_{2k} \leq n\}$ とおく．U_n は時刻 n までに区間 (a, b) を下から上へと横切る回数である．

補題 8.10 $X_\bullet$ が $\mathbb{F}$-劣マルチンゲールならば，$a < b$ である任意の a, b に対して，

$$(b - a)\mathbb{E}(U_n) \leq \mathbb{E}((X_n - a)^+) - \mathbb{E}((X_0 - a)^+), \qquad n \geq 0, \tag{8.25}$$

ここに，$x \in \mathbb{R}$ に対して $x^+ = \max(0, x)$ とする．

証明 $H_n = 1(\exists k \geq 1, N_{2k-1} < n \leq N_{2k})$，$Y_n = a + (X_n - a)^+$ とおく．$X_\bullet$ が $\mathbb{F}$-劣マルチンゲールであり，$f(x) = a + (x - a)^+$ は非減少な凸関数であるから，補題 8.8 より $Y_\bullet$ は $\mathbb{F}$-劣マルチンゲールである．定義から $H_\bullet$ は予測可能であり，$0 \leq H_n \leq 1$ であるから，補題 8.9 より $(H \cdot Y)_\bullet$ も $\mathbb{F}$-劣マルチンゲールである．$(H \cdot Y)_n$ は時刻 n までに区間 (a, b) を下から上へと横切る回数に $(b - a)$ をかけた数であるから，

$$(b - a)U_n \leq (H \cdot Y)_n$$

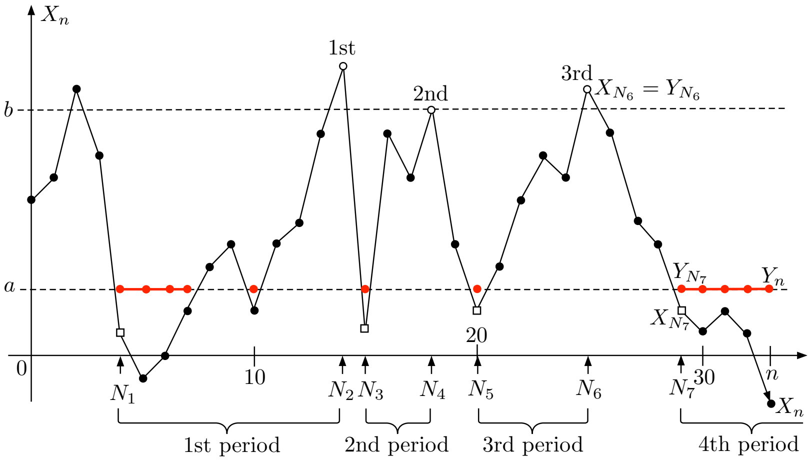


図 8.1: $X_\bullet$ が区間 (a,b) を下から上へと横切る回数が 3 の場合

が成り立つ．$K_\ell = 1 - H_\ell$ とおくと

$$Y_n - Y_0 = (H \cdot Y)_n + (K \cdot Y)_n.$$

であり，補題 8.9 より $(K \cdot Y)_\bullet$ も $\mathbb{F}$-劣マルチンゲールである．従って， $\mathbb{E}((K \cdot Y)_n) \geq \mathbb{E}((K \cdot Y)_0) = 0$ であるから，$(b-a)\mathbb{E}(U_n) \leq \mathbb{E}((H \cdot Y)_n) \leq \mathbb{E}(Y_n - Y_0)$. を得る．すなわち，(8.25) が証明された． ∎

定理 8.2 $X_\bullet$ が $\mathbb{F}$-劣マルチンゲールであり，$\sup_{n\geq 0}\mathbb{E}(X_n^+) < \infty$ ならば，$\mathbb{E}(|X_\infty|) < \infty$ を満たすある確率変数 X_∞ があり，$n \to \infty$ のとき $X_n \overset{a.s}{\to} X_\infty$ が成り立つ．

証明 $(X_n - a)^+ \leq |X_n| + |a|$ であるから，補題 8.10 より

$$\mathbb{E}(U_n) \leq \frac{1}{b-a}(|a| + \mathbb{E}(X_n^+)), \qquad n \geq 0$$

が成り立つ．U_n は n について非減少であるから，ある U へ収束する．この U は区間 (a,b) を下から上へ通過する総回数である．仮定より $\sup_{n\geq 0}\mathbb{E}(X_n^+) < \infty$ であるから，$\mathbb{E}(U) < \infty$ である．従って，U は確率 1 で有限である．すなわち，ある時間以降区間 (a,b) を下から上へ通過することはない．ここで，事象 $A_{a,b}$ を

$$A_{a,b} = \{\liminf_{n\to\infty} X_n < a < b < \limsup_{n\to\infty} X_n\}$$

により定義すると，$\omega \in A_{a,b}$ ならば，無限に多くの n に対して $X_n(\omega) \leq a$ であり，無限に多くの m に対して $X_m(\omega) \geq b$ であるから，$\mathbb{P}(A_{a,b}) = 0$ である．従って，Q を有理数を全て集めた集合とすると，Q が可算であることから，$\mathbb{P}(\bigcap_{r,q\in Q} A_{r,q}) = 0$．従って，

$$\mathbb{P}\Big(\liminf_{n\to\infty} X_n \geq \limsup_{n\to\infty} X_n\Big) = 1 - \mathbb{P}\Big(\liminf_{n\to\infty} X_n < \limsup_{n\to\infty} X_n\Big) = 1$$

一方，$\liminf_{n\to\infty} X_n \le \limsup_{n\to\infty} X_n$ は常に成り立つので，$X_\infty \equiv \lim_{n\to\infty} X_n$ の存在が示された．残された証明は $\mathbb{E}(|X_\infty|) < \infty$ である．この条件は，$X_\infty^- = (-X_\infty)^+$ とすると，$\mathbb{E}(X_\infty^+) < \infty$ かつ $\mathbb{E}(X_\infty^-) < \infty$ に等しい．$\lim_{n\to\infty} X_n^+ = X_\infty^+$ であるから，Fatou の補題より，

$$\mathbb{E}(X_\infty^+) = \mathbb{E}\left(\liminf_{n\to\infty} X_n^+\right) \le \liminf_{n\to\infty} \mathbb{E}(X_n^+) < \infty.$$

である．$X_\bullet$ は $\mathbb{F}$-劣マルチンゲールであるから，$\mathbb{E}(X_0) \le \mathbb{E}(X_n)$ である．従って，

$$\mathbb{E}(X_n^-) = \mathbb{E}(X_n^+) - \mathbb{E}(X_n) \le \mathbb{E}(X_n^+) - \mathbb{E}(X_0).$$

よって，再び Fatou の補題より，

$$\mathbb{E}(X_\infty^-) = \mathbb{E}\left(\liminf_{n\to\infty} X_n^-\right) \le \liminf_{n\to\infty} \mathbb{E}(X_n^-) \le \liminf_{n\to\infty} \mathbb{E}(X_n^+) - \mathbb{E}(X_0) < \infty.$$

従って，$\mathbb{E}(|X_\infty|) < \infty$ が得られ，証明が完成した．■

定理 8.3 $\{M_n; n \ge 0\}$ を $\mathbb{F}$-マルチンゲール，τ を $\mathbb{F}$-停止時刻とする．このとき，ある有限な期待値をもつ Y に対して $|M_n| \le Y$ がすべての $0 \le n \le \tau$ に対して成り立つならば，$\mathbb{E}(|M_\infty|) < \infty$ を満たす $\mathscr{F}_\infty$-可測な確率変数 M_∞ が存在し，$\{M_n; n \in [0,\infty]\}$ は $\mathbb{F}$-マルチンゲールであり，

$$\mathbb{E}(M_\tau|\mathscr{F}_0) = M_0, \qquad n \in [0,\infty], \tag{8.26}$$

が確率 1 で成り立つ．ここに，$\mathscr{F}_\infty = \bigcup_{n=0}^\infty \mathscr{F}_n$ であることを再記する．

証明 仮定より $\sup_{n\ge 0} \mathbb{E}(M_n^+) \le \mathbb{E}(Y) < \infty$ であるから，定理 8.2 より，$M_n \overset{a.s.}{\to} M_\infty$ かつ $\mathbb{E}(M_\infty) < \infty$ を満たす $\mathscr{F}_\infty$-可測な確率変数 M_∞ が存在する．従って，$\{M_n; n \in [0,\infty)\}$ が $\mathbb{F}$-マルチンゲールであることから，$\{M_n; n \in [0,\infty]\}$ が $\mathbb{F}$-マルチンゲールであることを示すためには，$\mathbb{E}(M_\infty|\mathscr{F}_n) = M_n$ が確率 1 で成り立つことを示せばよい．マルチンゲール性より，$k \ge n$ ならば $\mathbb{E}(M_k|\mathscr{F}_n) = M_n$ であるから，

$$\mathbb{E}(M_n 1_A) = \mathbb{E}(\mathbb{E}(M_k|\mathscr{F}_n)1_A) = \mathbb{E}(M_k 1_A), \qquad A \in \mathscr{F}_n,$$

である．この式で $k \to \infty$ とすれば，有界収束定理と $M_k \overset{a.s.}{\to} M_\infty$ より $\mathbb{E}(M_n 1_A) = \mathbb{E}(M_\infty 1_A)$ である．従って，$\mathbb{E}(M_\infty|\mathscr{F}_n) = M_n$ が得られる．(8.26) の証明は系 8.1 の証明と同様に τ を $\tau \wedge n$ に置き換え，$n \to \infty$ とすればよいので詳細は省く．■

条件 $|M_n| \le Y$ がすべての $0 \le n \le \tau$ に対して成り立つという条件は一様積分の条件：

$$\lim_{a\to\infty} \sup_{n\ge 0} \mathbb{E}(|M_n| 1(|M_n| > a) = 0$$

に置き換えることができる．証明はマルチンゲールに関する本を参照してほしい．

9 マルコフ連鎖と推移確率

状態空間 S が可算集合である確率過程 $\{X_n; n = 0, 1, \ldots\}$ の確率法則は，各 $n \geq 0$ に対する結合分布

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n), \qquad i_j \in S,$$

により決まる．この結合分布の計算が簡単にできると解析的に扱いやすい．これから学ぶマルコフ連鎖は，マルコフ性と呼ばれる条件を仮定し，結合分布の計算を簡単にする．なお，マルコフ（Markov）は最初にこのような確率過程を研究した人の名前である．マルコフ連鎖やその時間連続型であるマルコフ過程は，解析的に扱いやすく，状態空間の取り方により広く応用できるため，確率過程の基本モデルとして広く使われている．

9.1 マルコフ連鎖の定義

結合分布の計算において，$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_n) \neq 0$ のとき，

$$\begin{aligned}&\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_{n+1} = i_{n+1}) \\ &\quad = \mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)\mathbb{P}(X_n = i_{n+1} | X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)\end{aligned}$$

であるから，すべての $n \geq 1$ に対するすべての $i, j, i_0, i_1, \ldots, i_{n-1} \in S$ に対して

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j | X_n = i), \quad n \geq 0, \quad (9.1)$$

が成り立つならば，結合分布が帰納的に計算できる．この条件は，時刻 n までの観測結果が得られた下での X_{n+1} の条件付き確率が，X_n の値のみを観測した下での条件付き確率に等しいことを表している．

定義 9.1 離散的な状態空間 S をもつ確率過程 $\{X_n; n = 0, 1, \ldots\}$ が，条件 (9.1) を満たすならば，マルコフ連鎖といい，(9.1) をマルコフ性と呼ぶ．更に，(9.1) の右辺が n に依らないならば，

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i), \qquad i, j \in S, \qquad (9.2)$$

とおくことができる．このとき，p_{ij} を状態 i から j への推移確率と呼び，マルコフ連鎖は定常な推移確率をもつという．

以下では，$\{X_n; n = 0, 1, \ldots\}$ を状態空間 S と定常な推移確率をもつマルコフ連鎖とする．このとき，各 $n \geq 1$ に対して $X_0, X_1, \ldots, X_n$ の結合分布を X_0 の分布と p_{ij} から計算することができる．

補題 9.1 定常な推移確率をもつマルコフ連鎖 $\{X_n\}$ において，任意の $n \geq 0$ と任意の $i_0, i_1, \ldots, i_n \in S$ に対して

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = a_{i_0} p_{i_0 i_1} \times \cdots \times p_{i_{n-1} i_n} \qquad (9.3)$$

が成り立つ．ここに，$a_i = \mathbb{P}(X_0 = i)$，$i \in S$ とする．

証明 条件付き確率の定義より，

$$\begin{aligned}&\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)\\&\quad = \mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1})\mathbb{P}(X_n = i_n | X_0 = i_0, \ldots, X_{n-1} = i_{n-1})\end{aligned}$$

である．右辺にマルコフ性の条件 (9.2) を適用すると，

$$\begin{aligned}&\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)\\&\quad = \mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1})\mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1})\\&\quad = \mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1})p_{i_{n-1}i_n}\end{aligned}$$

である．$\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1})$ に対して同様な操作を繰り返すと，最終的に (9.3) が得られる． ∎

補題 9.1 により，マルコフ連鎖（正確には，マルコフ連鎖の確率分布）は，状態空間 S，推移確率 $\{p_{ij}; i, j \in S\}$ と初期分布 $\{a_i; i \in S\}$ により決まる．初期分布は任意に与えることができるので，マルコフ連鎖が表す確率モデルは，状態空間 S と推移確率 $\{p_{ij}; i, j \in S\}$ により決まると言うことができる．

補題 9.2 定常な推移確率をもつマルコフ連鎖 $\{X_n\}$ において，$\ell = 0, 1, \ldots, n-1$ である $B_\ell \subset S$ に対して $A = \cap_{\ell=1}^{n-1}\{X_\ell \in B_\ell\}$ とするならば，任意の $i, j \in S$ に対して，

$$\mathbb{P}(X_{n+1} = j | A, X_n = i) = \mathbb{P}(X_{n+1} = j | X_n = i), \quad n \geq 0, \tag{9.4}$$

が成り立つ．

証明 (9.3) から

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i, X_{n+1} = j) = a_{i_0}p_{i_0i_1} \times \cdots \times p_{i_{n-2}i_{n-1}}p_{i_{n-1},i}p_{i,j}$$

である．この式の両辺を $\ell = 0, 1, \ldots, n-1$ に対して $i_\ell \in B_\ell$ を満たす i_ℓ について和を取ると

$$\mathbb{P}(A, X_n = i, X_{n+1} = j) = \mathbb{P}(A, X_n = i)p_{i,j}$$

が成り立つ．この式の両辺を $\mathbb{P}(A, X_n = i)$ で割れば (9.4) が得られる． ∎

状態空間と推移確率の与え方には，次の 3 つの方法がある．

i) S と p_{ij} を式で表す．

ii) p_{ij} を (i, j) 要素とする $S \times S$ 行列により表す．この行列を推移確率行列と呼ぶ．

iii) S の要素を頂点とし，$i, j \in S$ に対して $p_{ij} > 0$ ならば，頂点 i から頂点 j への辺があり重み p_{ij} もつグラフを描く．このグラフを推移図と呼ぶ．

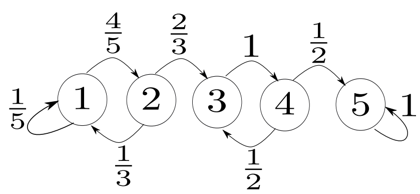
例 9.1 (状態空間と推移確率) マルコフ連鎖 $\{X_n\}$ の状態空間が $S=\{1,2,3,4,5\}$ により与えられ，推移確率 p_{ij} が以下のように与えられている．

$$p_{11}=\frac{1}{5},\quad p_{12}=\frac{4}{5},\quad p_{21}=\frac{1}{3},\quad p_{23}=\frac{2}{3},\quad p_{34}=1,\quad p_{43}=\frac{1}{2},\quad p_{45}=\frac{1}{2},\quad p_{55}=1,$$

ここに，上記以外の p_{ij} はすべて 0 に等しい．このマルコフ連鎖の推移行列を P とすると，

$$P=\begin{pmatrix} \frac{1}{5} & \frac{4}{5} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

であり，推移図は下記のように表すことができる．



∎

演習問題 9.1 状態空間が $S=\{1,2,3,4,5,6\}$ であるマルコフ連鎖において推移確率 p_{ij} が，$p+q=1$ を満たす $p,q>0$ に対して以下のように与えられている．

$$p_{12}=p_{23}=p_{34}=p_{45}=p_{56}=p,\quad p_{65}=p_{54}=p_{43}=p_{32}=p_{21}=q,\quad p_{11}=q,p_{66}=p.$$

このマルコフ連鎖の推移確率行列を求め，推移図を描け．

9.2 推移確率の計算

マルコフ連鎖 $\{X_n\}$ によるモデル化の目的の 1 つは，現在の状態から将来の状態を推測することである．このために，$X_0=i$ が与えられたという条件の下で X_n の確率分布を計算する．この計算は行列の積の計算と同じように行うことができる．このために推移確率行列を P により表す．

定義 9.2 $n\geq 0$ に対して，

$$p_{ij}^{(n)}=\mathbb{P}(X_n=j|X_0=i),\qquad i,j\in S,$$

により，$p_{ij}^{(n)}$ を定義し，n 次の推移確率と呼ぶ．また，$p_{ij}^{(n)}$ を (i,j) 要素とする $S\times S$ 行列を $P^{(n)}$ により表し，n 次の推移確率行列と呼ぶ．この定義から，I を $S\times S$ 単位行列とすると，$P^{(0)}=I$ であり，P の定義から $P^{(1)}=P$ である．

2 つの $S \times S$ 行列 $A = \{a_{ij}\}$ と $B = \{b_{ij}\}$ に対して，積 AB を行列の積と同様に，

$$AB \text{ の } (i,j) \text{ 要素 } = \sum_{k \in S} a_{ik} b_{kj}$$

により定義する．S は無限個の要素を持つ集合の場合もあるので，右辺の和が確定する場合に限り，この定義は有効である．

補題 9.3 $n \geq 1$ に対して，$P^{(n)} = P^n$ が成り立つ．ここに，P^n は P を n 乗した行列である．

証明 (9.2) において，$i_0 = i$, $i_n = j$ とおき，$i_1, i_2, \ldots, i_{n-1} \in S$ について和を取れば，

$$\mathbb{P}(X_0 = i, X_n = j) = \mathbb{P}(X_0 = i) \times P^n \text{ の } (i,j) \text{ 要素}$$

である．この式の両辺を $\mathbb{P}(X_0 = i)$ で割れば，左辺は $P^{(n)}$ の (i,j) 要素となり，証明できた． ∎

この補題より，$m, n \geq 0$ に対して，

$$P^{(m+n)} = P^{m+n} = P^m P^n = P^{(m)} P^{(n)}$$

であるから次の系が得られる．

系 9.1 (Chapman-Kormogorov の公式) $\ell, m \geq 0$ に対して，

$$p_{ij}^{(\ell+m)} = \sum_{k \in S} p_{ik}^{(\ell)} p_{kj}^{(m)}, \qquad i, j \in S, \tag{9.5}$$

が成り立つ．

例 9.2 (推移確率の計算) 例 9.1 のマルコフ連鎖に対して，$P^{(n)}$ を $n = 2, 3$ に対して計算すると

$$P^{(2)} = P^2 = \begin{pmatrix} \frac{23}{75} & \frac{4}{25} & \frac{8}{15} & 0 & 0 \\ \frac{1}{15} & \frac{4}{15} & 0 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad P^{(3)} = P^3 = \begin{pmatrix} \frac{23}{75} & \frac{4}{25} & \frac{8}{15} & 0 & 0 \\ \frac{43}{375} & \frac{92}{375} & \frac{8}{75} & \frac{8}{15} & 0 \\ \frac{23}{225} & \frac{4}{75} & \frac{31}{90} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

となる．これより，最初の 3 行 3 列の部分行列が複雑であることが見て取れる．しかし，次の節で述べる理論により，$n \to \infty$ のとき $P^{(n)}$ は 5 列目がすべて 1 の要素を持つ行列に収束することを証明できる． ∎

$P^{(n)}$ を解析的に求めるために行列母関数を使う方法がある．推移確率行列 P に対して，

$$U(z) = \sum_{n=0}^{\infty} z^n P^{(n)}, \qquad |z| < 1,$$

により，z を変数とする行列 $U(z)$ を定義し，行列母関数と呼ぶ．この $U(z)$ を解析的に求めることができれば，$U(z)$ を z について多項式展開し，z^n の係数として $P^{(n)}$ が得られる．$P^{(n)} = PP^{(n-1)}$ であるから，この式の両辺に z^n をかけて $n = 1, 2, \dots$ について加えると，

$$U(z) - I = zPU(z)$$

である．従って，$|z| < 1$ ならば，$\sum_{n=0}^{\infty} z^n P^n$ が収束することから，

$$U(z) = (I - zP)^{-1} \tag{9.6}$$

が得られる．S が有限集合，特に，状態数が 2 または 3 の場合には計算が可能である．しかし，状態数が増えると，$U(z)$ の計算は一般に困難である．

演習問題 9.2 演習問題 9.1 のマルコフ連鎖に対して $p_{24}^{(4)}$ を求めよ．

演習問題 9.3 $\{X_n; n \geq 0\}$ が状態空間 S と推移行列 P をもつマルコフ連鎖であるとき，$Y_n = X_{2n}$ により確率過程 $\{Y_n; n \geq 0\}$ を定義する．この確率過程がマルコフ連鎖であることを示し，その推移行列を求めよ．

9.3　マルコフ連鎖の応用例

マルコフ連鎖は各種のシステムを表す確率モデルとして広く使われている．その中からいくつかを紹介する.

本節の確率過程 $X_\bullet \equiv \{X_n; n \geq 0\}$ はすべて，標準的なフィルトレーションを使う．すなわち，

$$\mathscr{F}_n = \sigma(X_0, X_1, \dots, X_n)$$

とし，$\mathbb{F} = \{\mathscr{F}_n; n \geq 0\}$ によりフィルトレーションを定義する．

9.3.1　在庫管理と反射壁をもつランダムウォーク

小売店がある商品について，在庫量が s を下回ったとき $S\,(> s)$ になるまで在庫量を増やすように発注を行う方法があり，$s-S$ 政策と呼ばれている．1 日ごとに発注をおこなうことができるとする．1 日の販売数はランダムであり，各日ごとの販売数は独立であるとする．1 日目から始め，n 日目の需要数を A_n とするとき，販売終了後に発注を行い翌日までに納入されるとする．なお，需要数が在庫数より多いときには，在庫が空になるとする．1 日目 の最初の在庫量を a とする．ここに，$s \leq a \leq S$ とする．このとき，$X_0 = a$,

$$X_n = \begin{cases} X_{n-1} - A_n, & X_{n-1} - A_n \geq s, \\ S, & X_{n-1} - A_n < s, \end{cases} \qquad n \geq 1, \tag{9.7}$$

により確率過程 $X_\bullet \equiv \{X_n; n \geq 0\}$ を定義すれば，X_n は n 日目の販売開始時の在庫量である．日々の需要が独立で同一の分布に従うと仮定する．すなわち，各 $n \geq 1$ に対して A_n は $\mathscr{F}_{n-1}$ と独立であり，n に依らない同一の分布に従うと仮定する．この分布を $\{a_k; k \geq 0\}$ により表す．すなわち，$a_k = \mathbb{P}(A_n = k)$ である．

$X_\bullet$ は状態空間が $S = \{s, s+1, \ldots, S\}$ のマルコフ連鎖である．なぜならば，

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} a_{i-j}, & s \leq j \leq i, \\ \sum_{k=i-s+1}^{\infty} a_k, & j = S, \end{cases}$$

であるから，$X_\bullet$ は定常な推移をもち，A_{n+1} は $X_0, X_1, \ldots, X_n$ と独立であるから，

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = j | X_0 = a, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i) \\
&\quad = \mathbb{P}(j = i - A_{n+1} \geq s | X_0 = a, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i) \\
&\qquad + \mathbb{P}(j = S, i - A_{n+1} < s | X_0 = a, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i) \\
&\quad = 1(s \leq j \leq i) a_{i-j} + 1(j = S) \sum_{k=i-s+1}^{\infty} a_k \\
&\quad = \mathbb{P}(X_{n+1} = j | X_n = i)
\end{aligned}$$

より，マルコフ性が成り立つからである．

在庫問題では最適な s と S を求めることが問題である．最適を求めるためにの評価量として売上高から需要に応じることができなかったことに依る損失を差し引いた利益を使う．売れた商品 1 つ当たりの利益を a 円，売れ残った商品の在庫管理費を 1 つ当たり h 円とする．なお，応じることができなかった需要よる損失もあるが考慮しないことにする．このとき，n 日目の利益を R_n とすると，

$$\begin{aligned}
R_n &= a \min(A_n, X_{n-1}) - h(X_{n-1} - A_n)^+ \\
&= a(X_{n-1} + \min(0, A_n - X_{n-1})) - h(X_{n-1} - A_n)^+ \\
&= a(X_{n-1} - \max(0, X_{n-1} - A_n)) - h(X_{n-1} - A_n)^+ \\
&= aX_{n-1} - (a+h)(X_{n-1} - A_n)^+, \qquad n \geq 1,
\end{aligned}$$

である．ここに，$x^+ = \max(0, x)$ とする．少し複雑な式であるが，R_n の期待値により期待利益を計算できる．この場合には，s, S に関係なく $\mathbb{E}(R_n | X_{n-1} = x)$ を最小にする x の値を求める単純な在庫問題が考えられる．

演習問題 9.4 この在庫問題において，$\mathbb{E}(R_n | X_{n-1} = x)$ を最小とする x を求めよ．

一方，発注を行った場合に経費 K 円がかかるとすると，1 日当たりの総経費 T_n は

$$\begin{aligned}
T_n &= a \min(A_n, X_{n-1}) - h(X_{n-1} - A_n)^+ - 1(X_{n-1} - A_n < s)K \\
&= a(X_{n-1} + \min(A_n - X_{n-1}, 0)) - h(X_{n-1} - A_n)^+ - 1(X_{n-1} - A_n < s)K \\
&= aX_{n-1} - (a+h)(X_{n-1} - A_n)^+ - 1(X_{n-1} - A_n < s)K, \qquad n \geq 1,
\end{aligned}$$

である．この場合には時間平均

$$\overline{T} \equiv \lim_{n\to\infty} \frac{1}{n}\sum_{\ell=1}^{n} \mathbb{E}(T_\ell | X_0 = a)$$

などを評価量として計算し，最小とする s, S を求めることが問題である．この式の計算は複雑そうに見えるが，$n \to \infty$ のときの X_n の極限分布が存在し，この分布を使って $\overline{T}$ を求めることができる，

9.3.2　出生死滅過程

生物の個体数は出生と死滅により変化する．各時刻で個体数が i のとき，個体数が 1 増える確率を $p(i)$，1 減る確率を $q(i)$，変化しない確率を $r(i)$ とし，$p(i) + q(i) = 1$ とする．ただし，個体数が 0 のときは，$q(0)$ は個体数が変わらない確率とする．このとき，時刻 n での個体数を X_n とし，$X_\bullet$ は推移確率が

$$\mathbb{P}(X_n = j | X_{n-1} = i) = \begin{cases} p(i), & j = i + 1, \\ q(i), & j = (i-1)^+, \end{cases}$$

により与えられたマルコフ連鎖とする．その状態空間は $S = \{0, 1, \ldots\}$ である．このマルコフ連鎖を出生死滅過程と呼ぶ．例えば，正の整数 a に対して，

$$p(i) = \frac{a}{i+a}, \qquad q(i) = \frac{i}{i+a}, \qquad i \geq 0,$$

とすると，$0 \leq i \leq a$ ならば個体数が増加し，$i > a$ ならば減少する傾向があることを表している．この場合，時間の経過による個体数の分布は X_n の分布を計算することにより求めることができる．この場合にも，$n \to \infty$ のとき X_n の分布がある分布に近づくことを証明することができる．

9.3.3　分枝過程

生物の個体数の変化をより詳細に調べるモデルがある．このモデルにおいては，個々の個体は各時刻で与えられた分布に従う数の子供を産み死滅する．この分布を $\{p(k); k \geq 0\}$ とし，$X_0 = 1$，

$$X_{n+1} = \sum_{k=1}^{X_n} A_{n,k}, \qquad n \geq 0, \tag{9.8}$$

により確率過程 $X_\bullet \equiv \{X_n; n \geq 0\}$ を定義する．X_n は時刻 n における総個体数を表している．ここに，$\{A_{n,k}; k \geq 1\}$ は独立な確率変数列であり，$\mathscr{F}_n$ と独立であるとする．n, k に依らない同一の分布 $\{p(k); k \geq 0\}$ に従うと仮定する．このとき，

$$\begin{aligned}\mathbb{P}(X_{n+1} = j | X_0 = 1, X_1 = i_1, \ldots, X_n = i) &= \mathbb{P}\left(\sum_{k=1}^{X_n} A_{n,k} = j \,\middle|\, X_n = i\right) \\ &= \mathbb{P}\left(\sum_{k=1}^{i} A_{n,k} = j\right) = \mathbb{P}(X_{n+1} = j | X_n = i)\end{aligned}$$

が成り立つ．従って，$X_\bullet$ は定常な推移をもつマルコフ連鎖である．このマルコフ連鎖を分枝過程と呼ぶ．

分枝過程では，個体が消滅する確率や消滅するまでの時間の分布を求めることが問題である．$\lambda = \mathbb{E}(A_{n,k})$ とおき，(9.8) の期待値を取ると

$$\begin{aligned}\mathbb{E}(X_{n+1}) &= \mathbb{E}\Big(\sum_{k=1}^{X_n} A_{n,k}\Big) = \mathbb{E}\Big(\sum_{i=1}^{\infty}\sum_{k=1}^{i} A_{n,k}1(X_n = i)\Big) \\ &= \sum_{i=1}^{\infty} i\mathbb{E}(A_{n,k})\mathbb{P}(X_n = i) = \lambda\mathbb{E}(X_n)\end{aligned}$$

が成り立つ．従って，$X_0 = 1$ より，$\mathbb{E}(X_n) = \lambda^n$ である．これより，$\lambda < 1$ のときに限り，$n \to \infty$ のとき $\mathbb{E}(X_n)$ は 0 へ収束する．X_n は非負であるから，これより，X_n は 0 へ確率収束する．これは，1 個体当たりの平均出生数，すなわち，出生率 λ が 1 より小さいならば個体全体がいつか消滅することを表している．

$\mathscr{F}_0 = \{\varnothing, \Omega\}$，$n \geq 1$ に対して $\mathscr{F}_n = \sigma(A_{1,k}, A_{2,k}, \ldots, A_{n-1,k}; k \geq 1)$ とし，$\mathbb{F} \equiv \{\mathscr{F}_n; n \geq 0\}$ をフィルトレーションとするならば，$X_\bullet$ は $\mathbb{F}$ に適合した確率過程である．このとき，X_n が $\mathscr{F}_n$-可測であり，$\mathscr{F}_n$ と $A_{n,k}$ が独立であることから，$\lambda = \mathbb{E}(A_{n,k}) < 1$ ならば，

$$\mathbb{E}(X_{n+1}|\mathscr{F}_n) = \sum_{k=1}^{X_n} \mathbb{E}(A_{n,k}|\mathscr{F}_n) = \lambda X_n \leq X_n$$

が確率 1 で成り立つ．すなわち，$X_\bullet$ は $\mathbb{F}$-優マルチンゲールである．よって，$\{-X_n; n \geq 0\}$ は $\mathbb{F}$-劣マルチンゲールとなり，$X_n \geq 0$ より $\mathbb{E}((-X_n)^+) = 0$ であるから，定理 8.2 より，$n \to \infty$ のとき X_n はある X_∞ へ概収束する．このとき，X_n は X_∞ へ確率収束もする．一方，X_n は 0 へ確率収束するので，$X_\infty = 0$ である．従って，X_n は 0 へ概収束する．

以上をまとめると，1 個体から生まれる平均の子供数が 1 より小ならば，確率 1 で全ての個体が消滅するという結果が得られた．

9.3.4　待ち行列過程

サービスを受けるための待ち行列は日常よく見かける．1 つのサービス窓口があり 1 つの待ち行列に並び先着順にサービスを受けるシステムにおける待ち行列長の変化を確率過程で表してみよう．簡単のために書く離散的な時刻 n において到着する客数を A_n 各時刻で 1 人のサービスを行うとする．時刻 n でのサービス中も含めた客数を X_n とすると，

$$X_{n+1} = \max(0, X_n + A_n - 1), \qquad n \geq 0 \tag{9.9}$$

と表すことができる．ここに，A_n は $\mathscr{F}_n$ と独立であり，n に依らない同一の分布に従うとする．このとき，$X_\bullet$ が $\mathbb{F}$-マルコフ連鎖であることを確かめることができる．

(9.9) において，$X_0 = 0$ として，繰り返し代入すると，

$$
\begin{aligned}
X_{n+1} &= \max(0, \max(0, X_{n-1} + A_{n-1} - 1) + A_n - 1) \\
&= \max(0, A_n - 1, X_{n-1} + A_{n-1} + A_n - 2) \\
&\ \ \vdots \\
&= \max(0, A_n - 1, A_n + A_{n-1} - 2, \ldots, A_n + A_{n-1} + \ldots, A_1 - n)
\end{aligned}
$$

が得られる．仮定から，$A_1, A_2, \ldots, A_n$ は独立で同一の分布に従う．これより．$A_n, A_n + A_{n-1}, \ldots, A_n + A_{n-1} + \ldots, A_1$ の同時分布は $A_1, A_1 + A_2, \ldots, A_1 + A_2 + \ldots + A_n$ の同時分布に等しい．従って，$\ell \geq 1$ に対して $S_\ell = A_1 + \ldots A_\ell - \ell$ とおき，等しい分布を持つことを $\overset{d}{\sim}$ により表すと，

$$X_{n+1} \overset{d}{=} \max(0, S_1, S_2, \ldots, S_n)$$

である．$\lambda = \mathbb{E}(A_1)$ とおくと，$\mathbb{E}(S_1) = \lambda - 1$ であり，S_n は独立で同一の分布に従う確率変数の列 $\{A_n - 1; n \geq 1\}$ であるから，大数の強法則により，

$$\lim_{n\to\infty} \frac{1}{n} S_n = \lambda - 1$$

が確率 1 で成り立つ．

従って，$\lambda < 1$，すなわち，平均到着数がサービス能力より小さいときには，$\lambda - 1 < 0$ より，確率 1 で

$$\lim_{n\to\infty} S_n = -\infty$$

が成り立つ．これより，$Y_n = \max(0, S_1, S_2, \ldots, S_n)$ とおくと，Y_n は非減少であるから，有限な確率変数 $Y_\infty \equiv \max(0, S_1, S_2, \ldots)$ へ収束する．

以上をまとめると，$\lambda < 1$ ならば，待ち行列の長さ X_n の極限分布は

$$\lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \lim_{n\to\infty} \mathbb{P}(Y_n \leq x) = \mathbb{P}(Y_\infty \leq x), \qquad x \geq 0,$$

により与えられる．一方，$\lambda > 1$ ならば，$Y_n \to \infty$ であるから，

$$\lim_{n\to\infty} \mathbb{P}(X_n > x) = 1, \qquad x \geq 0,$$

である．すなわち，待ち行列長は無限に大きくなる．

演習問題 9.5 (9.9) により定義された待ち行列長過程 $\{X_n; n \geq 0\}$ が $\mathbb{F}$-マルコフ連鎖であることを証明せよ．

10　マルコフ連鎖とフィルトレーション

これまで状態空間 S をもつマルコフ連鎖 $\{X_n; n \ge 0\}$ を $X_0, X_1, \dots, X_n$ の値を与えた下での X_{n+1} の条件付き確率を用いて定義してきた（(9.1) と (9.2) 参照）．この定義をフィルトレーションを使って拡張すると共に，マルコフ性の条件 (9.1) において時刻 n を停止時刻に置き換えらることを示す．

10.1　フィルトレーションとテスト関数

$X_\bullet \equiv \{X_n; n \ge 0\}$ を離散的な状態空間 S をもつ確率過程とする．フィルトレーション $\mathbb{F}^X \equiv \{\mathscr{F}_n^X; n \ge 0\}$ を

$$\mathscr{F}_n^X = \sigma(X_0, X_1, \dots, X_n), \qquad n \ge 0,$$

により定義する．多くの場合このフィルトレーションで十分であるが，より多くの情報が観測できる場合にはよりおおきなフィルトレーションを使う．このフィルトレーションを $\mathbb{F} \equiv \{\mathscr{F}_n; n \ge 0\}$ より表す．ここに，$\mathscr{F}_n^X \subset \mathscr{F}_n$ とする．すなわち，$X_\bullet$ は $\mathbb{F}$ に適合する．

フィルトレーションにより観測結果を記述できるが，状態は一般に数ではない．このため状態の変化を計算をしたり比較するためには都合がよくない．そこで，状態を数に対応させることを考える．この対応を S から $\mathbb{R}$ への関数を用いて表す．この関数をテスト関数と呼ぶ．S が数の集合のときにはこのような対応は必要ないが，状態の評価値を考える際にはテスト関数が役立つ．

一般に 1 つのテスト関数では状態の情報を十分に表すことができない．そこで，テスト関数の集合を使う．この集合として，S から $\mathbb{R}$ への有界な関数すべての集合 $K_b(S)$ を使う．$i \in S$ に対して $f(x) = 1(x = i)$ とすれば，$f \in K_b(S)$ である．このとき，$\mathbb{E}(f(X_n)) = \mathbb{P}(X_n = i)$ であるから，X_n の分布はこのようなテスト関数を使って求められる．従って，すべての $f \in K_b(S)$ を使えば X_n の分布を求めることができる．

有界でないテスト関数を使うこともある．例えば，$S = \{a_1, a_2, \dots\}$ であるとき，$f(a_n) = n$ はそのようなテスト関数である．各テスト関数 f に対して，$\{f(X_n); n \ge 0\}$ も $\mathbb{F}$ に適合した確率過程である．この確率過程は実数を値に取るので，マルチンゲールやセミマルチンゲールを使うことができる．テスト関数は確率過程 $X_\bullet$ を解析する上で重要な道具である．

10.2　マルコフ連鎖のフィルトレーションによる再定義

フィルトレーション $\mathbb{F}$ とテスト関数の集合 $K_b(S)$ を使ってマルコフ連鎖を再定義することを考える．まず，条件 (9.1) は

$$\mathbb{E}(f(X_{n+1})|\mathscr{F}_n^X) = \mathbb{E}(f(X_{n+1})|X_n), \qquad f \in K_b(S), \tag{10.1}$$

が確率 1 で成り立つことに置き換えることができる．なぜならば，$\mathscr{F}_n$ は互いに背反な事象：

$$A_{n,i_1,i_2,\dots,i_{n-1},i} \equiv \{X_\ell = i_\ell, \ell = 0, 1, \dots, n,\}, \qquad i_0, i_1, \dots, i_{n-1}, i \in S,$$

から生成された σ-集合体であるからである．このとき，$f(x) = 1(x = j)$ とすれば，$\mathbb{P}(A_{n,i_1,i_2,\ldots,i_n}) > 0$ ならば，事象 $A_{n,i_1,i_2,\ldots,i_n}$ 上における $\mathbb{E}(f(X_{n+1})|\mathcal{F}_n^X)$ の値は (9.1) の左辺に等しい．また，$\{X_n = i\}$ 上における $\mathbb{E}(f(X_{n+1})|X_n)$ の値は (9.1) の右辺に等しい．$\{X_n = i\} \supset A_{n,i_1,i_2,\ldots,i_{n-1},i}$ であるから，(9.1) は (10.1) が $A_{n,i_1,i_2,\ldots,i_{n-1},i}$ 上で成り立つことに等しい．従って，(9.1) と (10.1) は同値な条件式である．

この条件の比較において，$\mathbb{P}(A_{n,i_1,i_2,\ldots,i_n}) > 0$ を仮定したが，これは (9.1) の条件に必要なためである．(9.1) では明記していないが，この仮定は暗黙のうちに使われている．これに対して，(10.1) ではこのような仮定は必要ない．これは，確率変数や σ-集合体の下での条件付き期待値の大きな利点である．

同様に，定常な推移をもつ条件 (9.2) は

$$\mathbb{E}(f(X_{n+1})|X_n) = \sum_{i\in S} p_{X_n,i} f(i) \tag{10.2}$$

を満たす推移確率 $\{p_{i,j}; i,j \in S\}$ が存在すると言い換えることができる．

これらの条件をまとめると，状態空間 S をもつ確率過程 $\{X_n; n \geq 0\}$ が定常な推移をもつマルコフ過程とは，$p_{i,j} \geq 0$ と $\sum_{j\in S} p_{i,j} = 1$ を満たすある $\{p_{i,j}; i,j \in S\}$ があり，

$$\mathbb{E}(f(X_{n+1})|\mathcal{F}_n^X) = \sum_{i\in S} p_{X_n,i} f(i), \qquad f \in K_b(S), \tag{10.3}$$

が成り立つことを言う．

マルコフ連鎖を観測したとき，時刻 n で得られる情報が $\mathcal{F}_n^X$ より多いこともある．この場合を含めるためマルコフ連鎖を次のように拡張する．

定義 10.1 確率過程 $\{X_n; n \geq 0\}$ がフィルトレーション $\mathbb{F} \equiv \{\mathcal{F}_n; n \geq 0\}$ に適合し，すなわち，すべての $n \geq 0$ に対して $\mathcal{F}_n^X \subset \mathcal{F}_n$ であり，$p_{i,j} \geq 0$ と $\sum_{j\in S} p_{i,j} = 1$ を満たすある $\{p_{i,j}; i,j \in S\}$ に対して，

$$\mathbb{E}(f(X_{n+1})|\mathcal{F}_n) = \sum_{i\in S} p_{X_n,i} f(i), \qquad f \in K_b(S), \tag{10.4}$$

が成り立つならば，$X_\bullet$ を定常な推移をもつ $\mathbb{F}$-マルコフ連鎖といい，$S \times S$ 行列 $P \equiv \{p_{i,j}; i,j \in S\}$ を推移確率行列と呼ぶ．

演習問題 10.1 定常な推移をもつ $\mathbb{F}$-マルコフ連鎖 $X_\bullet$ について，(a) $\ell \geq 0, m \geq 1$ に対して，

$$\mathbb{E}(\mathbb{E}(f(X_{\ell+m})|\mathcal{F}_\ell) = \mathbb{E}(f(X_{\ell+m})|X_\ell), \qquad f \in K_b(S),$$

を証明せよ．(b) この結果を用いて Chapman-Kormogorov の公式 (9.5) を証明せよ．

マルコフ連鎖において，ある停止時刻後は同じ状態に止まるように状態推移を変えたい場合がある．例えば，賭をマルコフ連鎖で表すとき，破産して賭を途中で止める場合に適用できる．この場合には定常な推移をもつ条件は必ずしも満たされない．そこで，定義 10.1 を以下のように拡張する．

定義 10.2 確率過程 $\{X_n; n \geq 0\}$ がフィルトレーション $\mathbb{F} \equiv \{\mathscr{F}_n; n \geq 0\}$ に適合し，

$$\mathbb{E}(f(X_{n+1})|\mathscr{F}_n) = \mathbb{E}(f(X_{n+1})|X_n), \qquad f \in K_b(S), \tag{10.5}$$

が成り立つならば，$X_\bullet$ を $\mathbb{F}$-マルコフ連鎖と呼ぶ．

定義 10.1 との違いは，定義 10.2 においては定常な推移をもつことを仮定していないことである．

10.3　強マルコフ性

$X_\bullet \equiv \{X_n; n \geq 0\}$ が定常な推移をもつ $\mathbb{F}$-マルコフ連鎖であるとき，時刻 1 以上で初めて状態 $i \in S$ に到達する時刻を τ_i とする．すなわち，

$$\tau_i = \inf\{n \geq 1; X_n = i\} \tag{10.6}$$

により，τ_i を定義する．ここに，すべての $n \geq 1$ に対して $X_n \neq i$ ならば，$\tau_i = \infty$ とする．この τ_i は $\mathbb{F}$-停止時刻である．

演習問題 10.2 τ_i は $\mathbb{F}$-停止時刻であることを証明せよ．

この τ_i が有限であるとき，マルコフ性

$$\mathbb{E}(f(X_{\tau_i+1})|\mathscr{F}_{\tau_i}) = \sum_{j \in S} p_{X_\tau,j} f(j), \qquad f \in K_b(S), \tag{10.7}$$

は成り立つであろうか？ここに，

$$\mathscr{F}_{\tau_i} = \{A \in \mathscr{F}; \forall n \geq 0, A \cap \{\tau_i \leq n\} \in \mathscr{F}_n\}$$

であることを思い出してほしい（定義 8.6 参照）．(10.7) が成り立つと，τ_i 時間以降もマルコフとなるので好都合である．そこで，次の定義を行う．

定義 10.3 $X_\bullet \equiv \{X_n; n \geq 0\}$ が定常な推移確率 $\{p_{i,j}; i,j \in S\}$ をもつ $\mathbb{F}$-マルコフ連鎖であるとき，任意の有限な $\mathbb{F}$-停止時刻 τ に対して，

$$\mathbb{E}(f(X_{\tau+1})|\mathscr{F}_\tau) = \sum_{j \in S} p_{X_\tau,j} f(j), \qquad f \in K_b(S), \tag{10.8}$$

が成り立つならば，$X_\bullet$ を定常な推移をもつ $\mathbb{F}$-強マルコフ連鎖（または $\mathbb{F}$-強マルコフ過程）と呼ぶ．

演習問題 10.3 $X_\bullet$ が定常な推移をもつ $\mathbb{F}$-強マルコフ連鎖であるとき，任意の $k \geq 1$ と $f \in K_b(S)$ に対して，$\mathbb{E}(f(X_{\tau+k})|\mathscr{F}_\tau) = \sum_{j \in S} p^{(k)}_{X_\tau,j} f(j)$ が成り立つことを証明せよ．

補題 10.1 $X_\bullet$ が定常な推移をもつ $\mathbb{F}$-マルコフ連鎖ならば，定常な推移をもつ $\mathbb{F}$-強マルコフ連鎖である．

証明 τ を任意の有限な $\mathbb{F}$-停止時刻とする．$\mathbb{P}(\tau < \infty) = 1$ であるから，任意の $n \geq 0$ に対して，$\{\tau = n\}$ 上で (10.8) が成り立つことを示せばよい．これは，

$$\mathbb{E}(f(X_{\tau+1})1_{A\cap\{\tau=n\}}) = \mathbb{E}\left(\sum_{j\in S} p_{X_\tau,j} f(j) 1_{A\cap\{\tau=n\}}\right), \qquad A \in \mathscr{F}_\tau, f \in K_b(S),$$

すなわち，

$$\mathbb{E}(f(X_{n+1})1_{A\cap\{\tau=n\}}) = \mathbb{E}\left(\sum_{j\in S} p_{X_n,j} f(j) 1_{A\cap\{\tau=n\}}\right), \qquad A \in \mathscr{F}_\tau, f \in K_b(S), \tag{10.9}$$

を示すことと同値である．この式の右辺は $X_\bullet$ が $\mathbb{F}$-マルコフ連鎖であることから，

$$\begin{aligned}\sum_{j\in S} \mathbb{E}(p_{X_n,j} 1_{A\cap\{\tau=n\}}) f(j) &= \mathbb{E}(\mathbb{E}(f(X_{n+1})|X_n)1_{A\cap\{\tau=n\}}) \\ &= \mathbb{E}(\mathbb{E}(f(X_{n+1})|\mathscr{F}_n)1_{A\cap\{\tau=n\}})\end{aligned}$$

に等しい．$A \cap \{\tau = n\} \in \mathscr{F}_n$ であるから，最後の項は

$$\mathbb{E}(\mathbb{E}(f(X_{n+1})|\mathscr{F}_n)1_{A\cap\{\tau=n\}}) = \mathbb{E}(\mathbb{E}(f(X_{n+1})1_{A\cap\{\tau=n\}}|\mathscr{F}_n))$$

となり，(10.8) の左辺に一致する． ∎

補題 10.1 と同様に次のことが成り立つ．

系 10.1 $\mathbb{F}$-マルコフ連鎖は $\mathbb{F}$-強マルコフ連鎖である．

補題 10.2 $\mathbb{F}$-マルコフ連鎖 $X_\bullet$ と $\mathbb{F}$-停止時刻 τ に対して，$Y_n = X_{\tau\wedge n}$ により確率過程 $Y_\bullet \equiv \{Y_n; n \geq 0\}$ を定義すれば，$Y_\bullet$ は $\mathbb{F}$-マルコフ連鎖である．

証明 $Y_\bullet$ が $\mathbb{F}$-マルコフ連鎖であることを示すためには，任意の $f \in K_b(S)$ に対して，

$$\mathbb{E}(f(X_{\tau\wedge(n+1)})|\mathscr{F}_n) = \mathbb{E}(f(X_{\tau\wedge(n+1)})|X_{\tau\wedge n}), \qquad , \tag{10.10}$$

を証明すればよい．これは，条件付き期待値の定義より，

$$\mathbb{E}(f(X_{\tau\wedge(n+1)})1_A) = \mathbb{E}(\mathbb{E}(f(X_{\tau\wedge(n+1)})|X_{\tau\wedge n})1_A), \qquad A \in \mathscr{F}_n, \tag{10.11}$$

を示すことに等しい．ここに，

$$\sigma(X_{\tau\wedge n}) = \sigma(\cup_{\ell=0}^n(\{\tau = \ell\} \cap \sigma(X_\ell)) \cup (\{\tau > n\} \cap \sigma(X_n))) \subset \mathscr{F}_n$$

より，$\{\tau \le n\}, \{\tau > n\} \in \sigma(X_{\tau\wedge n})$ であるから，(10.11) の右辺は，

$$\begin{aligned}
&\mathbb{E}(\mathbb{E}(f(X_{\tau\wedge(n+1)})|X_{\tau\wedge n})1_A)\\
&\quad= \mathbb{E}(\mathbb{E}(f(X_{\tau\wedge(n+1)})(1(\tau \le n) + 1(\tau > n))|X_{\tau\wedge n})1_A)\\
&\quad= \mathbb{E}(\mathbb{E}(f(X_{\tau\wedge n})1(\tau \le n) + f(X_{(\tau\wedge n)+1})1(\tau > n))|X_{\tau\wedge n})1_A)\\
&\quad= \mathbb{E}(f(X_{\tau\wedge n})1(\tau \le n)1_A) + \mathbb{E}(f(X_{(\tau\wedge n)+1})|X_{\tau\wedge n})1(\tau > n)1_A)
\end{aligned} \tag{10.12}$$

である．ここに，最後の等式では $\mathbb{E}(f(X_{\tau\wedge n})1(\tau \le n)|X_{\tau\wedge n}) = f(X_{\tau\wedge n})1(\tau \le n)$ を用いた．(10.12) の最後の右辺の 2 番目の項に強マルコフ性を適用すると右辺は

$$\mathbb{E}(f(X_{\tau\wedge n})1(\tau \le n)1_A) + \mathbb{E}(f(X_{(\tau\wedge n)+1})|\mathscr{F}_{\tau\wedge n})1_{\{\tau>n\}\cap A})$$

と書くことができるが，$\{\tau > n\} \cap A \in \mathscr{F}_{\tau\wedge n}$ であるから，条件付き期待値の定義より，

$$\begin{aligned}
&\mathbb{E}(f(X_{\tau\wedge n})1(\tau \le n)1_A) + \mathbb{E}(f(X_{(\tau\wedge n)+1}))1_{\{\tau>n\}\cap A})\\
&\quad= \mathbb{E}(f(X_{\tau\wedge(n+1)})1(\tau \le n)1_A) + \mathbb{E}(f(X_{\tau\wedge(n+1)}))1(\tau > n)1_A)\\
&\quad= \mathbb{E}(f(X_{\tau\wedge(n+1)})1_A)
\end{aligned}$$

となり，(10.11) の左辺に一致する．　■

この補題の $Y_\bullet$ は τ-停止 $\mathbb{F}$-マルコフ連鎖と呼ばれる．一般に $Y_\bullet$ は定常な推移をもたない．例えば，正の整数 k に対して，$\tau \equiv k$ とすると，τ はどんなフィルトレーションに対しても停止時刻であるが，$Y_n = X_{\tau\wedge n} = X_{k\wedge n}$ であるから，$n \ge k$ に対して $Y_n = X_k$ であり，時刻 k より前と後では推移確率が異なる．なお，この τ-停止 $\mathbb{F}$-マルコフ連鎖 $Y_\bullet$ のフィルトレーション $\mathbb{F}$ を $\mathbb{F}^\tau \equiv \{\mathscr{F}_{\tau\wedge n}; n \ge 0\}$ によって定義したフィルトレーション $\mathbb{F}^\tau$ に置き換えることもできる（演習問題 10.5 参照).

以上のことから，マルコフ連鎖はマルチンゲールと同様に停止時刻により停止しても同じ種類の確率過程である．このような確率過程を停止時刻による局所化（localization）に対して安定であるという．

10.4　マルコフ連鎖とマルチンゲール

マルコフ連鎖をテスト関数を使ってセミマルチンゲールとして表し，予測可能な部分と純粋にランダムな部分に分けることを考えてみよう．

$X_\bullet \equiv \{X_n; n \ge 0\}$ は定常な推移確率 $\{p_{i,j}; i, j \in S\}$ をもつ $\mathbb{F}$-マルコフ連鎖であるとする．テスト関数 $f \in K_b(S)$ に対して，実数値の確率過程 $\{f(X_n); n \ge 0\}$ にセミマルチンゲール分解を適用する．(8.7) と同様に

$$M_n = M_0 + \sum_{\ell=1}^{n}(f(X_\ell) - \mathbb{E}(f(X_\ell)|\mathscr{F}_{\ell-1})), \qquad n \ge 0,$$

により M_n を定義すれば $M(\cdot) \equiv \{M_n; n \geq 0\}$ は $\mathbb{F}$-マルチンゲールである．$M_0 = 0$ とすると，予測可能な部分 Y_n は

$$
\begin{aligned}
Y_n &= f(X_n) - M_n \\
&= \mathbb{E}(f(X_n)|\mathscr{F}_{n-1}) - \sum_{\ell=1}^{n-1}(f(X_\ell) - \mathbb{E}(f(X_\ell)|\mathscr{F}_{\ell-1})) \\
&= \sum_{\ell=1}^{n}(\mathbb{E}(f(X_\ell)|X_{\ell-1}) - f(X_{\ell-1})) + f(X_0) \\
&= f(X_0) + \sum_{\ell=0}^{n-1}\sum_{j\in S} p_{X_\ell j}(f(j) - f(X_\ell))
\end{aligned}
$$

である．従って，セミマルチンゲール表現

$$
f(X_n) = f(X_0) + \sum_{\ell=0}^{n-1}\sum_{j\in S} p_{X_\ell j}(f(j) - f(X_\ell)) + M_n \tag{10.13}
$$

が得られた．この式における予測可能な部分は時刻 $n-1$ までの状態変化量の推移確率による重みつきの累積である．(10.13) は標本関数の式を変形したものであるから，標本関数を推移確率移を用いて表した式であるという見方もできる．

$K_b(S)$ から $K_b(S)$ への関数 $\mathscr{A}$ を

$$
\mathscr{A} f(i) = \sum_{j\in S} p_{ij}(f(j) - f(i)), \qquad i \in S, f \in K_b(S), \tag{10.14}
$$

により定義し生成作用素と呼ぶ．この関数は線形関数であるから，生成作用素は線型作用素である．このとき，(10.13) は

$$
f(X_n) = f(X_0) + \sum_{\ell=0}^{n-1} \mathscr{A} f(X_\ell) + M_n \tag{10.15}
$$

と表すことができる．この式から，$\mathscr{A} f(X_\ell)$ は時刻 ℓ での可測によって得られる情報量であると解釈することができる．

演習問題 10.4 $\mathscr{A}$ を推移確率行列 P をもつマルコフ連鎖の生成作用素とする．(10.14) を第 i 要素が $f(i)$ であるベクトル $\boldsymbol{f}$ に対する行列 $\mathscr{A}$ の定義式とみなすとき，行列 $\mathscr{A}$ を P を用いて表し，$\mathscr{A}\mathbf{1} = \mathbf{0}$ が成り立つことを示せ．ここに，$\mathbf{1}$ はすべての要素が 1 である S-次元ベクトルとする．

演習問題 10.5 補題 10.2 の $Y_\bullet$ は $\mathbb{F}^\tau$-マルコフ連鎖であることを証明せよ．ここに，$\mathbb{F}^\tau \equiv \{\mathscr{F}_{\tau\wedge n}; n \geq 0\}$ とする．

11　状態の分類

$X_\bullet \equiv \{X_n; n \geq 0\}$ を定常な推移をもつマルコフ連鎖とする．その状態空間 S や次数 n が大きくなると推移確率 $p_{ij}^{(n)}$ の計算が難しい．そこで，$n \to \infty$ のとき一定の値に収束することが証明できれば，$p_{ij}^{(n)}$ を近似的に求めることが可能となる．このために，各 $i, j \in S$ に対して

1) $n \to \infty$ のとき $p_{ij}^{(n)}$ の極限が存在するのはどのようなときか？

2) 1) の極限の値を求めることができるか？

という問題がある．本節では，この問題に答えるために次の 3 つの観点から状態を分類する．

① 状態 i から状態 j が到達可能か？

② $p_{ii}^{(n)} > 0$ となる n に周期はあるか？

③ 状態 i から出発したとき状態 i へ確率 1 で戻るのか，戻るならば，戻るまでの平均時間は有限か？

11.1　到達可能性

定義 11.1 $i, j \in S$ に対して，ある $n \geq 1$ があり，$p_{ij}^{(n)} > 0$ ならば，j は i から到達可能であるといい，$i \Rightarrow j$ と表す．$i \Rightarrow j$ かつ $j \Rightarrow i$ ならば，i, i は互いに到達可能であるといい，$i \Leftrightarrow j$ と表す．更に，$C \subset S$ に対して，

(i) 任意の $i, j \in C$ に対して，$i \Leftrightarrow j$ ならば，C は既約であるという．特に，S が既約ならば，マルコフ連鎖が既約であるという．

(ii) 任意の $i \in C$ と任意の $j \in S \setminus C$ に対して，$p_{ij} = 0$ ならば，C を閉集合という．特に，$i \in S$ に対して，$\{i\}$ が閉集合であるとき，i を吸収状態と呼ぶ．

この定義より，$i \Rightarrow j$ かつ $j \Rightarrow k$ ならば，ある $\ell, m \geq 1$ があり，$p_{ij}^{(\ell)} > 0$ かつ $p_{jk}^{(m)} > 0$ である．従って，系 9.1 より，

$$p_{ik}^{(\ell+m)} = \sum_{j' \in S} p_{ij'}^{(\ell)} p_{j'k}^{(m)} \geq p_{ij}^{(\ell)} p_{jk}^{(m)} > 0$$

であるから，$i \Rightarrow k$ である．また，$i \in C$ に対して $i \Rightarrow i$ ならば，$i \Leftrightarrow i$ である．特に空でない C が既約な閉集合ならば，$i \in C$ に対して $i \Rightarrow i$ であるから，既約な閉集合上で $\Rightarrow$ は同値関係を表す．既約と閉集合により状態空間 S を次のように分割することができる．

補題 11.1 マルコフ連鎖の状態空間 S は互いに排反である既約な閉集合と既約な閉集合を部分集合として含まない状態の集合に分割することができる．

証明　各 $i \in S$ に対して，$C(i) = \{j \in S; i \Leftrightarrow j\}$ により状態の集合を定義する．$i, j \in S$ に対して，$C(i) \cap C(j) \neq \varnothing$ ならば，上記で述べた同値関係より $C(i) = C(j)$ である．各 $i \in S$ に対して，

$$T = S \setminus (\cup_{i \in S} C(i))$$

とおく．このとき，S は T と空集合でない異なる $C(i)$, $i \in S$ に分割される．∎

この結果より，$p_{ij}^{(n)}$ を計算するためには，既約な閉集合または既約な閉集合を部分集合としてもたない集合 T について調べれば十分であることがわかる．更に，$j \in T$ ならば，いつか j へ戻ることができない状態に到達するので，$n \to \infty$ のとき $p_{jj}^{(n)}$ は 0 へ収束することが予想され，$p_{ij}^{(n)}$ も 0 へ収束することが予想される．従って，既約な閉集合について，$p_{ij}^{(n)}$ を考えれば十分である．

例 11.1　例 9.1 のマルコフ連鎖の状態を補題 11.1 に従って分類すると，

$$C(5) = \{5\}, \qquad T = \{1, 2, 3, 4\}$$

である．なお，$\{1, 2\}$ と $\{3, 4\}$ は既約であるが閉集合ではない．∎

演習問題 11.1　図 11.1 を推移図とするマルコフ連鎖の状態を補題 11.1 に従って分類せよ．

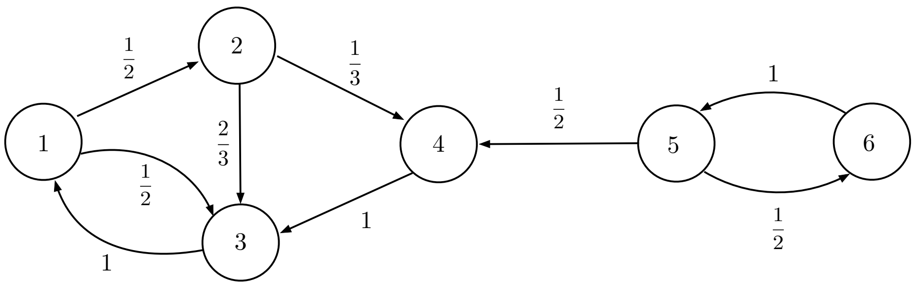


図 11.1: 演習問題 11.1 の推移図

②と③の問題を考える際に状態 j に初めて到達するまでの時間が役立つ．この時間を，

$$\tau_j = \inf\{n \geq 1; X_n = j\}, \qquad i \in S,$$

により定義する．ここに，すべての $n \geq 1$ に対して $X_n \neq j$ ならば，$\tau_j = \infty$ とする．状態 $i, j \in S$ に対して，$X_0 = i$ の下での τ_j の分布を

$$f_{ij}^{(n)} = \mathbb{P}(\tau_j = n | X_0 = i), \qquad n \geq 1, \tag{11.1}$$

により表す．次の補題は $f_{ij}^{(n)}$ と $p_{ij}^{(n)}$ の関係を表す重要な結果である．

補題 11.2 任意の $n \geq 0$ と $i,j \in S$ に対して次の式が成り立つ．

$$p_{ij}^{(n)} = \begin{cases} 1(i=j), & n=0, \\ \sum_{\ell=1}^{n} f_{ij}^{(\ell)} p_{jj}^{(n-\ell)}, & n \geq 1. \end{cases} \tag{11.2}$$

証明 $n=0$ のとき，$p_{ij}^{(0)} = 1(i=j)$ であるから (11.2) が成り立つ．$n \geq 1$ のときは，τ_j の定義より，$A_{1,\ell-1} = \{X_m \neq j, m=1,2,\ldots,\ell-1\}$ とおくと，

$$\{\tau_j = \ell\} = A_{1,\ell-1} \cap \{X_\ell = j\}$$

であるから，

$$\begin{aligned}
\mathbb{P}(X_n = j | X_0 = i) &= \sum_{\ell=1}^{n} \mathbb{P}(\tau_j = \ell, X_n = j | X_0 = i) \\
&= \sum_{\ell=1}^{n} \mathbb{P}(\tau_j = \ell | X_0 = i)\mathbb{P}(X_n = j | \tau_j = \ell, X_0 = i) \\
&= \sum_{\ell=1}^{n} f_{ij}^{(n)} \mathbb{P}(X_n = j | A_{1,\ell-1}, X_\ell = j) \\
&= \sum_{\ell=1}^{n} f_{ij}^{(n)} \mathbb{P}(X_n = j | X_\ell = j) = \sum_{\ell=1}^{n} f_{ij}^{\ell} p_{jj}^{(n-\ell)}
\end{aligned}$$

となり，(11.2) が得られる．∎

11.2 周期

定義 11.2 $i \in S$ に対して，

$$d(i) = \{n \geq 1; p_{ii}^{(n)} > 0\} \text{ の最大公約数}$$

により $d(i)$ を定義し，状態 i の周期と呼ぶ．なお，$p_{ii}^{(n)} > 0$ を満たす $n \geq 1$ が存在しない場合には，$d(i) = \infty$ とする．

この定義より，$i \Leftrightarrow i$ ならば，ある ℓ に対して $p_{ii}^{(\ell)} > 0$ である．これより，$1 \leq d(i) \leq \ell$ である．従って，$d(i) < \infty$ である．

補題 11.3 $i,j \in S$ に対して，$i \Leftrightarrow j$ ならば，$d(i) = d(j) < \infty$ である．

証明 $i \Leftrightarrow j$ より，$i \Leftrightarrow i$，$j \Leftrightarrow j$ であるから，$d(i), d(j) < \infty$ である．従って，$i = j$ ならば，補題が成り立つのは自明である．$i \neq j$ と仮定する．$i \Leftrightarrow j$ より，ある $\ell, m \geq 1$ で，$p_{ij}^{(\ell)} > 0$ かつ $p_{ji}^{(m)} > 0$ となるものが存在する．このとき，$p_{jj}^{(n)} > 0$ を満たす任意の $n \geq 0$ に対して，

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

であるから，$\ell+n+m$ は $d(i)$ の倍数である．特に，$n=0$ のときは，$p_{ii}^{(0)}=1$ であるから，$\ell+m$ も $d(i)$ の倍数である．従って，$p_{jj}^{(n)}>0$ を満たす任意の $n\geq 1$ は $d(i)$ の倍数である．これは，$d(j)$ が $d(i)$ の倍数であることを示している．i と j の役割を入れ換えれば逆も成り立つので補題が証明された． ∎

周期を求めるためには $p_{jj}^{(n)}>0$ であるすべての n を求める必要があるが，状態空間 S が大きい場合には必ずしも容易ではない．そこで，$p_{jj}^{(n)}>0$ の換わりに $f_{jj}^{(n)}>0$ を使う方法がある．

補題 11.4 $\{n\geq 1; f_{jj}^{(n)}>0\}$ の最大公約数を $f(j)$ とするならば，$f(j)$ は j の周期 $d(j)$ に等しい．

証明 $f_{jj}^{(n)}>0$ ならば $p_{jj}^{(n)}>0$ であるから，$f(j)$ は $d(j)$ の倍数である．逆に，$p_{jj}^{(n)}>0$ を満たす任意の $n\geq 1$ に対して，補題 11.2 よりある正の整数 $\ell_1\leq n$ があり，

$$f_{jj}^{(\ell_1)}p_{jj}^{(n-\ell_1)}>0$$

が成り立つ．$n-\ell_1>0$ ならば，同様に，ある正の整数 $\ell_2\leq n$ に対して，$f_{jj}^{(\ell_2)}p_{jj}^{(n-\ell_1-\ell_2)}>0$ である．このような操作を有限回繰り返すと，ある正の整数 k に対して，$n=\ell_1+\ell_2+\ldots+\ell_k$ かつ

$$f_{jj}^{(\ell_1)}f_{jj}^{(\ell_2)}\times\cdots\times f_{jj}^{(\ell_k)}>0$$

を満たす正の整数 $\ell_1,\ell_2,\ldots,\ell_k$ が存在する．従って，n は $f(j)$ で割り切れる．よって，$f(j)=d(j)$ である． ∎

例 11.2 例 9.1 のマルコフ連鎖の状態の周期は，$1\Leftrightarrow 2$，$3\Leftrightarrow 4$ であるから，

$$d(1)=d(2)=1,\qquad d(3)=d(4)=2,\qquad d(5)=1$$

である． ∎

演習問題 11.2 演習問題 11.1 の図 11.1 を推移図とするマルコフ連鎖について，各状態の周期を求めよ．

11.3 再帰性

③の問題を考えるために次の記号を定義する．$i,j\in S$ に対して，

$$f_{ij}^{(*)}=\mathbb{P}(\tau_j<\infty|X_0=i),$$
$$\mu_{ij}=\begin{cases}\mathbb{E}(\tau_j|X_0=i), & f_{ij}^{(*)}=1,\\ \infty & f_{ij}^{(*)}<1.\end{cases}$$

ここに，$f_{ij}^{(*)}=1$ の場合でも，$\mathbb{E}(\tau_j|X_0=i)=\infty$ となる場合がある．

演習問題 11.3 $f_{ij}^{(*)} = \sum_{\ell=1}^{\infty} f_{ij}^{(\ell)}$ を証明せよ．

定義 11.3 (a) $f_{ii}^{(*)} < 1$ ならば，i を一時的状態，
(b) i が一時的でない，すなわち，$f_{ii}^{(*)} = 1$ ならば，i を再帰的状態，
(c1) $f_{ii}^{(*)} = 1$ かつ $\mu_{ii} < \infty$ ならば，i を正再帰的状態，
(c2) $f_{ii}^{(*)} = 1$ かつ $\mu_{ii} = \infty$ ならば，i を零再帰的状態と呼ぶ．

補題 11.2 を使って，$f_{ij}^{(n)}$ を $p_{ij}^{(n)}$ と $p_{jj}^{(n)}$ から求めるために次の積率母関数を定義する．

$$F_{ij}(z) = \sum_{n=1}^{\infty} f_{ij}^{(n)} z^n, \qquad U_{ij}(z) = \sum_{n=0}^{\infty} p_{ij}^{(n)} z^n, \qquad i, j \in S.$$

補題 11.5 $i, j \in S$ と $0 < z < 1$ である z に対して，次の式が成り立つ．

$$U_{ij}(z) - 1(i = j) = F_{ij}(z) U_{jj}(z), \tag{11.3}$$

$$U_{ii}(z) = \frac{1}{1 - F_{ii}(z)}, \qquad |z| < 1, \tag{11.4}$$

$$\lim_{z \uparrow 1} F_{ii}(z) = f_{ii}^{(*)}, \tag{11.5}$$

$$\lim_{z \uparrow 1} \frac{1 - F_{ii}(z)}{1 - z} = \mu_{ii}, \tag{11.6}$$

が成り立つ．

証明 (11.2) の両辺に z^n をかけて $n = 1, 2, \ldots$ について加えると，これらの積率母関数より，

$$\begin{aligned} U_{ij}(z) - 1(i = j) &= \sum_{n=1}^{\infty} z^n \sum_{\ell=1}^{n} f_{ij}^{(\ell)} p_{jj}^{(n-\ell)} \\ &= \sum_{\ell=1}^{\infty} \sum_{n=\ell}^{n} z^\ell f_{ij}^{(\ell)} z^{n-\ell} p_{jj}^{(n-\ell)} = F_{ij}(z) U_{jj}(z). \end{aligned}$$

よって，(11.3) が示せた．この式で $j = i$ とすると，$|z| < 1$ に対して $|1 - F_{ii}(z)| > 0$ であるから，(11.4) を得る．(11.5) は τ_i の定義より明らかである．i が一時的ならば，$\mu_{ii} = \infty$ であり，$F_{ii}(1) = f_{ii}^{(*)} < 1$ であるから (11.6) を得る．i が再帰的ならば，$F_{ii}(1) = f_{ii}^{(*)} = 1$ であるから，

$$\frac{1 - F_{ii}(z)}{1 - z} = \frac{1}{1 - z} \sum_{\ell=1}^{\infty} (1 - z^\ell) f_{ii}^{(\ell)} = \sum_{\ell=1}^{\infty} \sum_{n=0}^{\ell-1} z^n f_{ii}^{(\ell)}.$$

$z \uparrow 1$ のとき，$\sum_{n=0}^{\ell-1} z^n \uparrow \ell$ であるから，有界収束定理より (11.6) を得る．∎

定理 11.1 状態を一時的，正再帰的，零再帰的に分類するとき，異なる $i, j \in S$ に対して，$i \Leftrightarrow j$ ならば，i と j は共に同じ種類に分類される．

証明 補題 11.3 の証明と同様に，$i \Leftrightarrow j$ より，ある $\ell, m \geq 1$ で，$p_{ij}^{(\ell)} > 0$ かつ $p_{ji}^{(m)} > 0$ となるものが存在するので，任意の $n \geq 0$ に対して，

$$p_{ii}^{(\ell+n+m)} \geq p_{ij}^{(\ell)} p_{jj}^{(n)} p_{ji}^{(m)}$$

が成り立つ．この式の両辺に $z^{\ell+n+m}$ をかけて，$n = 0, 1, \ldots$ について加えると

$$U_{ii}(z) \geq \sum_{n=0}^{\infty} z^{\ell+n+m} p_{ii}^{(\ell+n+m)} \geq p_{ij}^{(\ell)} p_{ji}^{(m)} z^{\ell+m} U_{jj}(z).$$

この式に (11.4) を適用すると，

$$(1 - F_{jj}(z)) \geq \sum_{n=0}^{\infty} z^{\ell+n+m} p_{ii}^{(\ell+n+m)} \geq p_{ij}^{(\ell)} p_{ji}^{(m)} z^{\ell+m} (1 - F_{ii}(z)) \geq 0. \tag{11.7}$$

よって，j が再帰的ならば，$F_{jj}(1) = 1$ であり，この式から $F_{ii}(1) = 1$ となるので，i も再帰的である．i, j が再帰的のとき，(11.7) の両辺を $1 - z$ で割り，$z \uparrow 1$ とすれば，(11.6) より，$\mathbb{E}(\tau_j | X_0 = j) < \infty$ ならば，$\mathbb{E}(\tau_i | X_0 = i) < \infty$ である．従って，j が正再帰的ならば，i も正再帰的である．i と j を入れ換えれば逆も言えるので定理が証明された．∎

状態が一時的であることは，推移図から直感的にわかることが多い．しかし，直感的なことを式で確認することも大事である．このために，次の式が役立つ．$i \in S$ に対して，

$$f_{ii}^{(*)} = p_{ii} + \sum_{j \in S \setminus \{i\}} p_{ij} f_{ji}^{(*)} \tag{11.8}$$

が成り立つ．この式を示すために，状態 i から初めて i へ戻る事象を時刻 1 での状態で分類し確率を求める．

$$\begin{aligned}\mathbb{P}(\tau_i < \infty | X_0 = i) &= \sum_{j \in S} \mathbb{P}(\tau_i < \infty | X_1 = j) p_{ij} \\ &= \mathbb{P}(\tau_i < \infty | X_1 = i) p_{ii} + \sum_{j \neq i} \mathbb{P}(\tau_i < \infty | X_1 = j) p_{ij}\end{aligned}$$

$\mathbb{P}(\tau_i < \infty | X_1 = i) = 1$，$\mathbb{P}(\tau_i < \infty | X_0 = i) = f_{ii}$ であり，$j \neq i$ のとき，

$$\mathbb{P}(\tau_i < \infty | X_1 = j) = \mathbb{P}(\tau_i < \infty | X_0 = j) = f_{ji}$$

であるから，(11.8) が得られる．

例 11.3 例 9.1 のマルコフ連鎖について再帰性に関する分類を行う．例 11.2 と同様に，$1 \Leftrightarrow 2$，$3 \Leftrightarrow 4$ であるから，状態空間を $\{1, 2\}$，$\{3, 4\}$，$\{5\}$ に分け，答えればよい．状態 5 は明らかに正再帰的である．状態 4 については，(11.8) より，

$$f_{44}^{(*)} = p_{43} f_{34}^{(*)} + p_{45} \times 0 \leq p_{43} = \frac{1}{2} < 1$$

であるから，状態 $3, 4$ は一時的である．同様に，状態 2 について，

$$f_{22}^{(*)} = p_{21} f_{12}^{(*)} + p_{23} \times 0 \leq p_{21} = \frac{1}{3} < 1$$

であるから，状態 $1, 2$ も一時的である．∎

演習問題 11.4 S をマルコフ連鎖の状態空間とする．C, D を $C \cap D = \varnothing$ かつ $C \cup D = S$ を満たす空でない S の部分集合とする，すなわち，S を空でない部分集合 C, D に分割する．C が既約な集合，D が閉集合であり，ある C の状態からある D の状態へ到達可能ならば，C の状態はすべて一時的であることを証明せよ．

12　極限分布と定常分布

定常な推移をもつマルコフ連鎖 $X_\bullet \equiv \{X_n; n \geq 0\}$ に対して，$p_{ij}^{(n)}$ が $n \to \infty$ としたとき収束するか，収束するならば，その極限値はどのように求められるかについて答えたい．このために，これから使う次の結果を証明しておこう．

補題 12.1 (Fatou の補題) 各 $n = 1, 2, \ldots$ に対して f_n を可算集合 S 上の非負値関数とするとき，

$$\sum_{i \in S} \liminf_{n \to \infty} f_n(i) \leq \liminf_{n \to \infty} \sum_{i \in S} f_n(i) \tag{12.1}$$

が成り立つ．

証明 S は可算集合であるから，$\ell \to \infty$ のとき，$S_\ell \uparrow S$ となる有限集合 S_ℓ の列，$S_1, S_2, \ldots$ が存在する．S_ℓ は有限集合であり，$f_n(i) \geq 0$ であるから，

$$\sum_{i \in S_\ell} \liminf_{n \to \infty} f_n(i) = \liminf_{n \to \infty} \sum_{i \in S_\ell} f_n(i) \leq \liminf_{n \to \infty} \sum_{i \in S} f_n(i).$$

この式の左辺で $\ell \to \infty$ とすれば (12.1) が得られる．　∎

補題 12.2 (離散型有界収束定理) ν を可算集合 S 上の非負値関数とする．各 $n = 1, 2, \ldots$ に対して f_n を S 上の実数値関数とするとき，

$$\lim_{n \to \infty} f_n(i) = g(i), \qquad i \in S, \tag{12.2}$$

を満たす S 上の関数 g があり，ある S 上の非負値関数 h に対して，

$$|f_n(i)| \leq h(i), \qquad \forall n \geq 1, i \in S, \tag{12.3}$$

$$\sum_{i \in S} \nu(i) h(i) < \infty, \tag{12.4}$$

が満たされるならば，次の式が成り立つ．

$$\lim_{n \to \infty} \sum_{i \in S} \nu(i) f_n(i) = \sum_{i \in S} \nu(i) g(i). \tag{12.5}$$

証明 S_ℓ を $\ell \to \infty$ のとき $S_\ell \uparrow S$ となる有限集合の列とする．このとき，(12.2) より，

$$\lim_{n \to \infty} \sum_{i \in S_\ell} \nu(i) f_n(i) = \sum_{i \in S_\ell} \nu(i) g(i) \tag{12.6}$$

が成り立つ．(12.4) より，任意の $\epsilon > 0$ に対して，十分大きな ℓ_0 を選ぶと，任意の $\ell \geq \ell_0$ に対して，

$$\sum_{i \in S \setminus S_\ell} \nu(i) h(i) < \epsilon$$

とすることができる．従って，(12.3) より，

$$\sum_{i\in S\setminus S_\ell}\nu(i)|f_n(i)| \le \sum_{i\in S\setminus S_\ell}\nu(i)h(i) < \epsilon, \qquad \ell \ge \ell_0,$$

であり，

$$\sum_{i\in S_\ell}\nu(i)f_n(i) - \epsilon \le \sum_{i\in S}\nu(i)f_n(i) \le \sum_{i\in S_\ell}\nu(i)f_n(i) + \epsilon, \qquad \ell \ge \ell_0,$$

が成り立つ．この式において，$n \to \infty$ として (12.6) を適用すると，

$$\sum_{i\in S_\ell}\nu(i)g(i) - \epsilon \le \liminf_{n\to\infty}\sum_{i\in S}\nu(i)f_n(i) \le \limsup_{n\to\infty}\sum_{i\in S}\nu(i)f_n(i) \le \sum_{i\in S_\ell}\nu(i)g(i) + \epsilon$$

よって，$\ell \to \infty$ とした後に，$\epsilon \downarrow 0$ とすれば，(12.5) が得られる． ∎

12.1 一時的な場合の極限

状態が一時的な場合には $p_{ij}^{(n)}$ の極限は比較的簡単である．

定理 12.1 j が一時的状態ならば，任意の $i \in S$ に対して，

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0 \tag{12.7}$$

である．

証明 初めに，$i = j$ の場合について証明する．j が一時的であるから，$F_{jj}(1) < 1$．従って，(11.4) より，

$$\sum_{n=0}^{\infty} p_{jj}^{(n)} = \lim_{z\uparrow 1} U_{jj}(z) = \frac{1}{1 - F_{jj}(1)} < \infty.$$

$\lim_{n\to\infty}\sum_{\ell=0}^{n} p_{jj}^{(\ell)} = \sum_{\ell=0}^{\infty} p_{jj}^{(\ell)} < \infty$ であるから，

$$\lim_{n\to\infty} p_{jj}^{(n)} = \lim_{n\to\infty}\left(\sum_{\ell=0}^{n} p_{jj}^{(\ell)} - \sum_{\ell=0}^{n-1} p_{jj}^{(\ell)}\right) = \lim_{n\to\infty}\sum_{\ell=0}^{n} p_{jj}^{(\ell)} - \lim_{n\to\infty}\sum_{\ell=0}^{n-1} p_{jj}^{(\ell)} = 0 \tag{12.8}$$

である．従って，$i = j$ に対して (12.7) が得られた．$i \ne j$ のときは，補題 11.2 より

$$p_{ij}^{(n)} = \sum_{\ell=1}^{\infty} f_{ij}^{(\ell)} 1(\ell \le n) p_{jj}^{(n-\ell)} \tag{12.9}$$

である．$1(\ell \le n)p_{jj}^{(n-\ell)} \le 1$ であるから，この式で $n \to \infty$ とすると，(12.8) と有界収束定理（補題 12.2）より (12.7) を得る． ∎

演習問題 12.1 定常な推移をもつマルコフ連鎖において，状態空間 S が有限集合ならば再帰的な状態が存在することを証明せよ．

12.2　再生定理

推移確率 $p_{ij}^{(n)}$ の極限を j が再帰的な場合に求めるために，確率分布 $\{f_n; n \geq 0\}$ に対して，

$$u_n = \begin{cases} 1, & n = 0 \\ \sum_{\ell=1}^{n} f_\ell u_{n-\ell}, & n \geq 1, \end{cases} \tag{12.10}$$

により数列 $\{u_n; n \geq 0\}$ を帰納的に定義する．

例 12.1 定常な推移をもつマルコフ連鎖において，状態 $j \in S$ が再帰的である場合，j へ初めて戻るまでの時間を τ_j とする．このとき，$f_\ell = f_{jj}^{(\ell)} \equiv \mathbb{P}(\tau_j = \ell | X_0 = j)$ とおくと，$\{f_\ell; \ell \geq 0\}$ は確率分布である．また，$p_{jj}^{(0)} = 1$ であり，補題 11.2 より，

$$p_{jj}^{(n)} = \sum_{\ell=1}^{n} f_\ell p_{jj}^{(n-\ell)}, \qquad n \geq 1$$

が成り立つので，$u_n = p_{jj}^{(n)}$ とすれば，(12.10) が成り立つ．このとき，u_n は時刻 n で j に戻っている確率を表している．これを状態 j の再生と見ることができることから，(12.10) は再生方程式と呼ばれている．

$p_{ij}^{(n)}$ の極限を求めるための最も基本的結果を述べよう．

定理 12.2 (再生定理) 確率分布 $\{f_n; n \geq 0\}$ が非周期的，すなわち，$\{n \geq 1; f_n > 0\}$ の最大公約数が 1 ならば，再生方程式 (12.10) の解 $\{u_n; n \geq 0\}$ の極限は

$$\lim_{n\to\infty} u_n = 1/\mu \tag{12.11}$$

により与えられる．ここに，$\mu = \sum_{\ell=1}^{\infty} \ell f_\ell$ であり，$\mu = \infty$ ならば，$1/\mu = 0$ とする．

この定理の証明に次の補題を使う．

補題 12.3 正の整数 $a_1, a_2, \ldots, a_m$ の最大公約数が 1 ならば，ある正の整数 n_0 があり，任意の $n \geq n_0$ に対して，

$$n = k_1 a_1 + k_2 a_2 + \ldots + k_m a_m$$

を満たす非負の整数 $k_1, k_2, \ldots, k_m$ が存在する．

証明　一般性を失うことなく，$0 < a_1 < a_2 < \ldots < a_m$ と仮定する．初めに，

$$1 = h_1 a_1 + h_2 a_2 + \ldots + h_m a_m \tag{12.12}$$

を満たす整数 $h_1, h_2, \ldots, h_m$ が存在するならば，補題が得られることを確認する．

$$a_0 = |h_1| a_1 + |h_2| a_2 + \ldots + |h_m| a_m$$

とおく．$n \ge a_0 a_1$ に対して，n を a_1 で割った商を b_0 余りを c_0 とすると

$$\begin{aligned} n &= a_1 b_0 + c_0 = a_1 b_0 + c_0(h_1 a_1 + h_2 a_2 + \ldots + h_m a_m) \\ &= (b_0 - a_0) a_1 + (a_1 |h_1| + c_0 h_1) a_1 + (a_1 |h_2| + c_0 h_2) a_2 + \ldots + (a_1 |h_m| + c_0 h_m) a_m \end{aligned}$$

である．$b_0 \ge a_0$，$0 \le c_0 < a_1$ であるから，任意の $n \ge n_0 \equiv a_0 a_1$ に対して，$k_1 = b_0 - a_0 + a_1|h_1| + c_0 h_1$ と $\ell = 2, 3, \ldots, m$ に対する $k_\ell = (a_1|h_\ell| + c_0 h_\ell)$ は条件を満たす非負の整数である．

(12.12) を帰納法で証明しよう．$m = 1$ のときは，$a_1 = 1$ であるから $h_1 = 1$ とすれば成り立つ．帰納法の証明では $m = 2$ の場合を使うので，$m = 2$ について証明する．a_2 を a_1 で割った商を b_1，余りを c_1 とすると，

$$a_2 = b_1 a_1 + c_1, \qquad 0 \le c_1 \le a_1 - 1,$$

である．a_2, a_1 の約数は a_1, c_1 の約数でありその逆も言えるので，a_2, a_1 の最大公約数が 1 であることから，a_1, c_1 の最大公約数も 1 である．特に，$c_1 = 0$ ならば，$a_1 = 1$ であり，(12.12) が成り立つ．$c_1 \ne 0$ ならば，$1 \le c_1 \le a_1 - 1$ である．このとき，a_1 を c_1 で割った商を b_2，余りを c_2 とすると，

$$a_1 = b_2 c_1 + c_2, \qquad 0 \le c_2 \le c_1 - 1 \le a_1 - 2,$$

である．先の場合と同様にして，c_1, c_2 の最大公約数は 1 である．よって，$c_2 = 0$ ならば，$c_1 = 1$ である．$c_1 \ne 1$ ならばさらに同様な操作を繰り返すと，$c_1 > c_2 > \ldots > c_\ell = 1$ となる $\ell \le a_1$ が存在する．$c_1, c_2, \ldots, c_\ell$ は a_1 と a_2 の線型和で表すことができ，$c_\ell = 1$ であるから，$m = 2$ に対して (12.12) が示せた．

$m = i \ge 2$ で成り立つとする．このとき，$a_1, a_2, \ldots, a_i$ の最大公約数が 1 ならば，$m = i + 1$ でも成り立つ．$a_1, a_2, \ldots, a_i$ の最大公約数が 1 より大きな数 j であるとする．このとき，$\ell = 1, 2, \ldots, i$ に対して $b_\ell = a_\ell / j$ とおくと，$b_1, b_2, \ldots, b_i$ の最大公約数は 1 であるから，

$$j = j(k'_1 b_1 + k'_2 b_2 + \ldots + k'_i b_i) = k'_1 a_1 + k'_2 a_2 + \ldots + k'_i a_i \tag{12.13}$$

を満たす $k'_1, k'_2, \ldots, k'_i$ が存在する．一方，j と a_{i+1} の最大公約数は 1 であるから，帰納法の仮定より，$1 = d_1 j + d_2 a_{i+1}$ を満たす整数 d_1, d_2 がある．この式の式に (12.13) の j を代入すれば求める表現が $m = i + 1$ に対して得られる． ∎

定理 12.2 の証明 u_n の定義式 (12.10) より，帰納的に $u_n \le 1$ が成り立つ．従って，

$$\overline{u} = \limsup_{n \to \infty} u_n \le 1$$

が存在する．よって，$\{u_n; \ge 0\}$ の部分列で $k \to \infty$ のとき $u_{n(k)} \to \overline{u}$ となるものが存在する．$f_i > 0$ を満たす i に対して $k \to \infty$ のとき $u_{n(k)-i} \to \overline{u}$ を示そう．$r_j = \sum_{n=j+1}^{\infty} f_n$ とおく．$\sum_{n=0}^{\infty} f_n = 1 < \infty$ であるから，$\lim_{j \to \infty} r_j = 0$ である．従って，任意の $\varepsilon > 0$ に対して十分大きな j_0 を選ぶと $r_{j_0} < \varepsilon$ である．十分大きな k に対して，$|u_{n(k)} - \overline{u}| < \varepsilon$ であることから，

$$\overline{u} - \varepsilon \le u_{n(k)} = \sum_{j=1}^{n(k)} f_j u_{n(k)-j} \le r_{j_0} + \sum_{j=1}^{j_0} f_j u_{n(k)-j} \le \varepsilon + f_i u_{n(k)-i} + (1 - f_i)(\overline{u} + \varepsilon)$$

である．従って，$f_i\overline{u} \le f_i u_{n(k)-i} + 3\varepsilon$ となるので，

$$\overline{u} \le \liminf_{k\to\infty} u_{n(k)-i} \le \limsup_{k\to\infty} u_{n(k)-i} \le \overline{u}$$

が成り立つので，$\lim_{k\to\infty} u_{n(k)-i} = \overline{u}$ である．同様な論証を繰り返すと，任意の正の整数 m に対して，$\lim_{k\to\infty} u_{n(k)-mi} = \overline{u}$ である．

初めに $f_1 > 0$ の場合について証明する．この場合には $i = 1$ であるから，

$$\lim_{k\to\infty} u_{n(k)-m} = \overline{u}, \qquad m \ge 0, \tag{12.14}$$

である．一方，$f_0 = 0$ と (12.10) より，

$$\sum_{\ell=0}^{n} r_{n-\ell}u_\ell = \sum_{\ell=0}^{n}(r_{n-\ell-1} - f_{n-\ell})u_\ell = \sum_{\ell=0}^{n} r_{n-\ell-1}u_\ell - u_n = \sum_{\ell=0}^{n-1} r_{n-\ell-1}u_\ell = \ldots = r_0u_0 = 1,$$

であるから，$\sum_{\ell=0}^{n} r_\ell u_{n-\ell} = 1$ である．この式より，各 $\ell_0 \ge 1$ に対して $n \to \infty$ とすると，

$$1 \ge \sum_{\ell=0}^{\ell_0} r_\ell u_{n-\ell} \to \sum_{\ell=0}^{\ell_0} r_\ell \overline{u}$$

$\sum_{\ell=0}^{\infty} r_\ell = \sum_{\ell=1}^{\infty} \ell f_\ell = \mu$ であるから，$\mu = \infty$ ならば，$\overline{u} = 0$，$\mu < \infty$ ならば，$\overline{u} \le 1/\mu$ である．$\mu < \infty$ の場合に同様な論証を $\underline{u} = \liminf_{n\to\infty} u_n$ に対して行うと，

$$1 \le \sum_{\ell=0}^{\ell_0} r_\ell u_{n-\ell} + \sum_{\ell=\ell_0+1}^{\infty} r_\ell \to \sum_{\ell=0}^{\ell_0} r_\ell \underline{u} + \sum_{\ell=\ell_0+1}^{\infty} r_\ell.$$

$\mu < \infty$ より，$\ell_0 \to \infty$ のとき $\sum_{\ell=\ell_0+1}^{\infty} r_\ell \to 0$ であるから，$\underline{u} \ge 1/\mu$ が得られる．従って，$i = 1$ のとき，(12.11) が証明された．

残された $f_1 = 0$ の場合について証明する．この場合は $\{i \ge 1; f_i > 0\}$ の最大公約数が 1 である．この集合の要素を小さいものから大きいものへ $a_1, a_2, \ldots$ と並べると，ある $m \ge 1$ に対して $a_1, a_2, \ldots, a_j$ の最大公約数は 1 に等しい．なぜならば，どんな j に対しても最大公約数が 2 以上ならば $\{i \ge 1; f_i > 0\}$ の最大公約数の 2 以上となるからである．これらの $a_1, a_2, \ldots, a_j$ に対して，先に示した，任意の正の整数 m に対して，$\lim_{k\to\infty} u_{n(k)-mi} = \overline{u}$ を導いた方法を適用すれば，任意の正の整数 $m_1, m_2, \ldots, m_j$ に対して，

$$\lim_{k\to\infty} u_{n(k)-(m_1a_1+m_2a_2+\ldots+m_ja_j)} = \overline{u}$$

を導くことができる．補題 12.3 より，ある $m_0 \ge 1$ があり，

$$\lim_{k\to\infty} u_{n(k)-m} = \overline{u}, \qquad \forall m \ge m_0,$$

が得られる．この結果は (12.14) と $m \ge m_0$ の制限がある点が異なるが，$n(k)$ を $n(k) + m_0$ で置き換えれば同じ結果と見なすことができる．従って，(12.14) に基づいた証明から，$\overline{u} = 1/\mu$，$\mu < \infty$ ならば $\underline{u} = 1/\mu$ であり，(12.11) が証明された．∎

12.3　再帰的な場合の極限

状態 $j \in S$ が再帰的な場合に，再生定理を使って $p_{ij}^{(n)}$ の極限を求めよう．以下の結果では j へ初めて到達するまでの時間 τ_j の期待値を使う．なお，$f_{ij}^{*} = \mathbb{P}(\tau_j < \infty | X_0 = i)$ である．

定理 12.3 状態 $j \in S$ が非周期的かつ再帰的ならば，

$$\lim_{n\to\infty} p_{jj}^{(n)} = \frac{1}{\mu_{jj}}, \tag{12.15}$$

$$\lim_{n\to\infty} p_{ij}^{(n)} = f_{ij}^{(*)} \frac{1}{\mu_{jj}}, \qquad i \in S, \tag{12.16}$$

が成り立つ．ここに，$\mu_{jj} = \mathbb{E}(\tau_j | X_0 = j)$ とし，$\mu_{jj} = \infty$ ならば，$1/\mu_{jj} = 0$ とする．

証明 例 12.1 で述べたように，$f_\ell = f_{jj}^{(\ell)}$，$u_n = p_{jj}^{(n)}$ とすれば，(12.10) を満たす．一方，j は非周期的であるから，補題 11.4 より $\{\ell \geq 1; f_{jj}^{(\ell)} > 0\}$ の最大公約数は 1 である．従って，定理 12.2 より (12.15) を得る．(12.16) の証明には，補題 11.2 の

$$p_{ij}^{(n)} = \sum_{\ell=1}^{n} f_{ij}^{(\ell)} p_{jj}^{(n)}, \qquad n \geq 0,$$

を使う．この式で，$n \to \infty$ とすると，(12.15) と有界収束定理より，

$$\lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{\ell=1}^{\infty} f_{ij}^{(\ell)} p_{jj}^{(n)} 1(\ell \leq n) = \sum_{\ell=1}^{\infty} f_{ij}^{(\ell)} \lim_{n\to\infty} p_{jj}^{(n)} 1(\ell \leq n) = f_{ij}^{(*)} \frac{1}{\mu_{jj}}$$

となり，(12.16) が得られた．　■

状態 j の周期を d とするとき，$Y_n = X_{nd}$ とおくと，$X_0 = j$ の下で，$Y_\bullet \equiv \{Y_n, n = 0, 1, \ldots\}$ はマルコフ連鎖であり，その推移確率は $\{p_{ij}^{(d)}; i, j \in S\}$ である．このマルコフ連鎖において j は非周期的である．$Y_\bullet$ が初めて j に戻る時間を $\tau_j^{(d)}$ とすると，$i = j$ に対する定理 12.3 の (12.15) から，j が再帰的ならば，

$$\lim_{n\to\infty} p_{jj}^{(nd)} = \frac{1}{\mathbb{E}(\tau_j^{(d)} | X_0 = j)} \tag{12.17}$$

である．$X_0 = Y_0 = j$ の下で，$\tau_i = d\,\tau_i^{(d)}$ であるから，定理 12.3 から次の結果が得られる．

系 12.1 状態 $j \in S$ が再帰的ならば，j の周期を d とするとき，

$$\lim_{n\to\infty} p_{ij}^{(nd)} = f_{ij}^{(*)} \frac{d}{\mu_{jj}} \tag{12.18}$$

が成り立つ．ただし，j が零再帰的の場合は右辺は 0 に等しいとする．

例 12.2 (出生死滅過程) 9.3.2 節で述べた出生死滅過程の状態空間は $S = \{0, 1, \ldots\}$ である．簡単な場合として，推移確率が $p + q = 1$ である正の数 p, q に対して，$p_{00} = q$，$i \geq 0$ に対して，$p_{i(i+1)} = p$, $p_{(i+1)i} = q$，その他の $i, j \in S$ に対して，$p_{ij} = 0$ であるとする．このモデルの推移図は下記のようになる．

この図より，既約で非周期的であることがわかる．従って，定理 12.3 により，$n \to \infty$ のとき $p_{ij}^{(n)}$ の極限が存在する．しかし，極限の値を求めることは難しそうである．　■

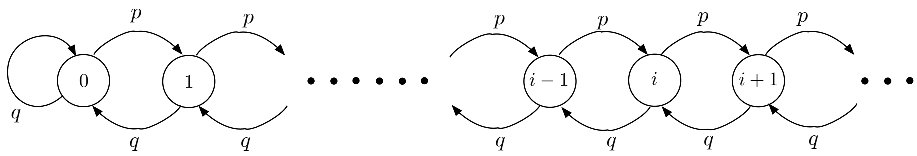


図 12.1: 出生死滅過程の推移図

12.4　定常測度と定常分布

定義 12.1 すべての $i \in S$ に対して $\pi(i) \geq 0$ のとき $\{\pi(i); i \in S\}$ を S 上の測度と呼ぶ．S 上の測度 $\{\pi(i); i \in S\}$ が，

$$\pi(j) = \sum_{i \in S} \pi(i) p_{ij}, \qquad j \in S, \tag{12.19}$$

を満たすとき，$\{\pi(i); i \in S\}$ を定常測度，(12.19) を定常方程式と呼ぶ．更に，

$$\sum_{i \in S} \pi(i) = 1 \tag{12.20}$$

を満たすならば，$\{\pi(i); i \in S\}$ を定常分布と呼ぶ．

(12.19) が成り立つならば，両辺に p_{jk} をかけて j について加えると

$$\sum_{j \in S} \pi(j) p_{jk} = \sum_{j \in S} \sum_{i \in S} \pi(i) p_{ij} p_{ik} = \sum_{i \in S} \pi(i) \sum_{j \in S} p_{ij} p_{jk} = \sum_{i \in S} \pi(i) p_{ik}^{(2)}$$

であるから，(12.19) を $i = j, j = k$ に対して適用すると，

$$\pi(k) = \sum_{i \in S} \pi(i) p_{ik}^{(2)}$$

が得られる．同様な計算を繰り返すことにより，任意の正の整数 n に対して，

$$\pi(j) = \sum_{i \in S} \pi(i) p_{ij}^{(n)}, \qquad j \in S, \tag{12.21}$$

が成り立つ．これより，定常測度を初期測度に選ぶとどんな時刻でも測度が変わらない．これが，定常測度と呼ぶ理由である．

定理 12.4 状態空間 S と推移確率 $\{p_{ij}\}$ をもつマルコフ連鎖 $\{X_n\}$ が既約かつ非周期的であるとする．このとき，次の 2 つの条件は同値である．

(i) 定常分布が存在する．

(ii) 全ての状態が正再帰的である．

これらのいずれかが成り立つならば，定常分布は唯 1 つであり，定常分布を π とすると，

$$\lim_{n \to \infty} p_{ij}^{(n)} = \pi(j) = \frac{1}{\mu_{jj}}. \qquad i, j \in S, \tag{12.22}$$

が成り立つ．

注 12.1 この定理において非周期性の仮定は除くことができる（注 12.2 参照）．ただし，定常分布の一意性は成り立たなくなる．例えば，周期を d とすると，推移確率 $\{p_{ij}^{(d)}; i.j \in S\}$ をもつマルコフ連鎖は非周期的であるから，定常分布 $\pi^{(d)}$ をもつ．このとき，

$$\pi(j) = \sum_{\ell=0}^{d-1} \sum_{k \in S} \pi^{(d)}(k) p_{kj}^{(\ell)}, \qquad j \in S,$$

により π を定義すると

$$\begin{aligned}\sum_{j' \in S} \pi(j') p_{j',j} &= \sum_{j' \in S} \sum_{\ell=0}^{d-1} \sum_{k \in S} \pi^{(d)}(k) p_{kj'}^{(\ell)} p_{j',j} = \sum_{\ell=0}^{d-1} \sum_{k \in S} \pi^{(d)}(k) p_{kj}^{(\ell+1)} \\ &= \sum_{\ell=1}^{d-1} \sum_{k \in S} \pi^{(d)}(k) p_{kj}^{(\ell)} + \sum_{k \in S} \pi^{(d)}(k) p_{kj}^{(d)} \\ &= \sum_{\ell=1}^{d-1} \sum_{k \in S} \pi^{(d)}(k) p_{kj}^{(\ell)} + \pi^{(d)}(j) = \pi(j)\end{aligned}$$

であるから，π は確かに定常分布である．一方，π' を

$$\pi'(j) = \sum_{\ell=1}^{d} \sum_{k \in S} \pi^{(d)}(k) p_{kj}^{(\ell)}, \qquad j \in S,$$

により定義すれば，π' も定常分布である．一般に $\pi \neq \pi'$ であるから，定常分布の一意性は一般に成り立たない．

証明 (i) が成り立つとし，定常分布を π により表す．(ii) が成り立たないと仮定する．これより，一時的または零再帰的状態が存在する．既約の仮定から，全ての状態が一時的または零再帰的である．従って，定理 12.1 と定理 12.3 より

$$\lim_{n \to \infty} p_{ij}^{(n)} = 0$$

である．これを (12.21) に適用すると，有界収束定理（補題 12.2）より，任意の $j \in S$ に対して，

$$\pi(j) = \lim_{n \to \infty} \sum_{i \in S} \pi(i) p_{ij}^{(n)} = \sum_{i \in S} \pi(i) \lim_{n \to \infty} p_{ij}^{(n)} = 0$$

となり，π が確率分布であることに矛盾する．従って，(i) ならば (ii) である．
次に，(ii) を仮定する．定理 12.3 より，

$$\lim_{n \to \infty} p_{ij}^{(n)} = \frac{1}{\mu_{jj}} \tag{12.23}$$

が成り立つ．$\pi(j) = 1/\mu_{jj}$ とおく．ここで，定理 12.3 と Fatou の補題 12.1 より，

$$\sum_{j \in S} \frac{1}{\mu_{jj}} = \sum_{j \in S} \lim_{n \to \infty} p_{ij}^{(n)} = \sum_{j \in S} \liminf_{n \to \infty} p_{ij}^{(n)} \leq \liminf_{n \to \infty} \sum_{j \in S} p_{ij}^{(n)} = 1$$

であるから，$\sum_{j\in S}\frac{1}{\mu_{jj}} \le 1 < \infty$ である．一方，

$$p_{ik}^{(n+1)} = \sum_{j\in S} p_{ij}^{(n)} p_{ij}$$

であるから，$n\to\infty$ とすると，

$$\pi(k) = \liminf_{n\to\infty}\sum_{j\in S} p_{ij}^{(n)} p_{jk} \ge \sum_{j\in S}\liminf_{n\to\infty} p_{ij}^{(n)} p_{jk} = \sum_{j\in S}\pi(j)p_{jk} \tag{12.24}$$

両辺を $k\in S$ について加えると，

$$\sum_{k\in S}\pi(k) \ge \sum_{k\in S}\sum_{j\in S}\pi(j)p_{jk} = \sum_{j\in S}\pi(j)$$

である．よって，$\sum_{k\in S}\pi(k) = \sum_{j\in S}\frac{1}{\mu_{jj}} < \infty$ より，不等式 (12.24) は等式となり，定常方程式 (12.19) が成り立つ．(12.21) と有界収束定理より，

$$\pi(j) = \lim_{n\to\infty}\sum_{i\in S}\pi(i)p_{ij}^{(n)} = \sum_{i\in S}\pi(i)\lim_{n\to\infty} p_{ij}^{(n)} = \sum_{i\in S}\pi(i)\frac{1}{\mu_{jj}} = \pi(j)\sum_{i\in S}\pi(i) \tag{12.25}$$

が成り立つので，$\sum_{i\in S}\pi(i) = 1$．従って，π は確率分布であり，(12.24) は等号で成り立つので，π は定常分布である．すなわち，(i) が成り立つ．最後に，(i) または (ii) が成り立つとき，定常分布がただ一つであることを証明する．$\nu \equiv \{\nu(i); i\in S\}$ を任意の定常分布とするとき，(12.21) より，

$$\nu(j) = \sum_{i\in S}\nu(i)p_{ij}^{(n)}, \qquad j\in S,$$

が成り立つ．この式で $n\to\infty$ とすれば，(12.23) が成り立つので，有界収束定理より，

$$\nu(j) = \lim_{n\to\infty}\sum_{i\in S}\nu(i)p_{ij}^{(n)} = \sum_{i\in S}\nu(i)\frac{1}{\mu_{jj}} = \frac{1}{\mu_{jj}}$$

である．従って，定常分布は唯 1 つである．(12.22) は (12.25) において既に証明されている．∎

例 12.3 (出生死滅過程の続き) 例 12.2 で述べた出生死滅過程の定常方程式 (12.21) をたてると，

$$\begin{aligned}
\nu(0) &= q\nu(0) + q\nu(1),\\
\nu(i) &= p\nu(i-1) + q\nu(i+1), \qquad i\ge 1,
\end{aligned}$$

である．第 2 式より，

$$q\nu(i) - p\nu(i-1) = q\nu(i+1) - p\nu(i)$$

である．この式を $i = 1, 2, \ldots, j-1$ まで，繰り返し使うと，

$$\begin{aligned} q\nu(1) - p\nu(0) &= q\nu(2) - p\nu(1) \\ &= q\nu(3) - p\nu(2) \\ &\quad \cdots \\ &= q\nu(j) - p\nu(j-1) \end{aligned}$$

が得られる．定常方程式の第 1 式より，$q\nu(1) - p\nu(0) = 0$ であるから，$\rho = \frac{p}{q}$ とおくと，

$$\nu(j) = \rho\nu(j-1) = \ldots = \rho^j \nu(0).$$

となる．従って，ν は定常測度である．ν が定常分布となることは，条件 $\rho < 1$ が成り立つときのみである．このとき，定常分布 π は

$$\pi(j) = (1-\rho)\rho^j, \qquad j = 0, 1, 2, \ldots.$$

により与えられる．∎

演習問題 12.2 定常な推移をもつマルコフ連鎖の状態空間 S が有限であるとき，
(a) 既約かつ非周期的ならば，定常分布が存在し唯一つであることを証明せよ．
(b) 一般に少なくとも 1 つの正再帰的状態が存在することを証明せよ．

演習問題 12.3 演習問題 11.1 の図 11.1 を推移図とするマルコフ連鎖について，
(a) 定常分布を求めよ．
(b) 一時的状態の集合，零再帰的状態の集合，正再帰的状態の集合を求めよ．

12.5 定常測度の存在

定理 12.4 により正再帰的な場合に限り定常分布が存在することが分かったが，零再帰的な場合にはどのようなことが言えるだろうか？定常分布に換わるものがあるだろうか？この疑問に答えるのが次の結果である．

定理 12.5 状態空間 S と推移確率 $\{p_{ij}\}$ をもつマルコフ連鎖 $\{X_n\}$ が既約かつ再帰的であるとする．このとき，任意に選んだ $k \in S$ に対して，S から $\mathbb{R}_+$ への関数 ν を

$$\nu(j) = \mathbb{E}\Big(\sum_{\ell=1}^{\infty} 1(X_\ell = j, \ell \le \tau_k) \Big| X_0 = k \Big), \qquad j \in S, \tag{12.26}$$

により定義すれば，ν は S 上の定常測度である．すなわち，$0 \le \nu(j) < \infty$ であり，(12.19)π を ν で置き換えた式：

$$\nu(j) = \sum_{i \in S} \nu(i) p_{ij}, \qquad j \in S, \tag{12.27}$$

を満たす．特に正再帰的ならば，$\pi(j) = \frac{1}{\mu_{kk}}\nu(j)$ とおくと，$\{\pi(j); j \in S\}$ は定常分布である．

注 12.2 この定理は非周期性を仮定していない．従って，正再帰的ならば，周期的な場合にも定常分布が存在することが再確認される（注 12.1 参照）．

証明 定義より $\nu(k) = 1$ であり，既約であることから，任意の $j \in S$ に対してある $\ell \geq 1$ があり，$p_{kj}^{(\ell)} > 0$ であるから，(12.27) が $j = k$ に対して成り立つならばすべての $\nu(i)$ は有限である．$\{\ell \leq \tau_k\} \in \mathscr{F}_{\ell-1}$ であるから，

$$\begin{aligned}
\nu(j) &= \mathbb{E}\Big(\sum_{\ell=1}^{\infty} 1(X_\ell = j, \ell \leq \tau_k)\Big| X_0 = k\Big) \\
&= p_{kj} + \sum_{\ell=2}^{\infty} \sum_{i \in S \setminus \{k\}} \mathbb{P}(X_{\ell-1} = i, X_\ell = j, \ell \leq \tau_k | X_0 = k) \\
&= p_{kj} + \sum_{\ell=2}^{\infty} \sum_{i \in S \setminus \{k\}} \mathbb{P}(X_{\ell-1} = i, \ell \leq \tau_k | X_0 = k)\mathbb{P}(X_\ell = j | X_0 = k, X_{\ell-1} = i, \ell \leq \tau_k) \\
&= p_{kj} + \sum_{\ell=1}^{\infty} \sum_{i \in S \setminus \{k\}} \mathbb{P}(X_\ell = i, \ell + 1 \leq \tau_k | X_0 = k) p_{ij}
\end{aligned}$$

である．ここで，$\{\ell + 1 \leq \tau_k\} = \{\ell \leq \tau_k\} \setminus \{\ell = \tau_k\}$ と

$$\sum_{i \in S \setminus \{k\}} \mathbb{P}(X_\ell = i, \ell = \tau_k | X_0 = k) p_{ij} = 0$$

より，

$$\begin{aligned}
\nu(j) &= p_{kj} + \sum_{\ell=1}^{\infty} \sum_{i \in S \setminus \{k\}} \mathbb{P}(X_\ell = i, \ell \leq \tau_k | X_0 = k) p_{ij} \\
&= \sum_{\ell=1}^{\infty} \sum_{i \in S} \mathbb{P}(X_\ell = i, \ell \leq \tau_k | X_0 = k) p_{ij} = \sum_{i \in S} \nu(i) p_{ij}
\end{aligned}$$

であり，(12.27) が得られた．　∎

ここまで来ると，定常測度は再帰的な場合にのみ存在するのかという疑問が出てくる．答えは否定的である．すなわち，一時的な場合には定常測度がある例とない例を作ることができる．

12.6 定常分布の意味

定常分布は応用上広く使われている．これは定常分布が安定なシステムを長時間にわたり稼働したときの平均的な特性を表すことによる．次の結果はこれを表すものである．

定理 12.6 定常な推移をもつ既約かつ非周期的なマルコフ連鎖 $\{X_n; n \geq 0\}$ が定常分布をもつならば，定常分布を $\pi \equiv \{\pi(i); i \in S\}$ とするとき任意の S から $\mathbb{R}_+$ への関数 f に対して，

$$\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=1}^{n} f(X_\ell) = \sum_{i \in S} f(i)\pi(i) \tag{12.28}$$

が確率 1 で成り立つ．X を分布 π に従う S に値を取る確率変数とするならば，この式の右辺は $\mathbb{E}(f(X))$ と表すことができる．なお，(12.28) の右辺は発散してもよい．

注 12.3 (12.28) は時間平均（左辺）が空間平均（右辺）に等しいことを表している．一般にこの 2 つの平均が一致することをエルゴード的であるという．

注 12.4 この定理では $f(i) \geq 0$ を仮定しているが，期待値の定義における場合と同様に，非負であることは本質的な仮定（絶対に必要な仮定）ではない．本質的な仮定は，∞ も含めた期待値の存在である．一般の場合に証明は期待値の定義のときと同様に行う．すなわち，$f^+(i) = \max(0, f(i))$ と $f^-(i) = \max(0, -f(i))$ を定義し，$f(i) = f^+(i) - f^-(i)$ と表す．f^+, f^- は非負値であるから，それぞれに上記の定理の結果が適用できる．これらの結果を組み合わせれば (12.28) が一般の f に対して得られる．

証明 各 $i \in S$ に対して，$X_0 = i$ のとき (12.28) を証明すればよい．定常分布をもつことから，定理 12.4 よりすべての状態が正再帰的である．従って，$\tau_i(n)$ を $n \geq 1$ 回目に状態 i を訪れる時刻とすると，確率 1 で有限であり，マルコフ性より，$T_i(n) \equiv \tau_i(n) - \tau_i(n-1)$ は $n = 1, 2, \ldots$ に対して独立で同一の分布に従う確率変数列である．ここに，$\tau_i(0) = 0$ とする．更に，正再帰的であることから，$\mu_{ii} = \mathbb{E}(T_i(n)|X_0 = i) < \infty$ である．次に，

$$Y_n = \sum_{\ell=\tau_i(n-1)+1}^{\tau_i(n)} f(X_\ell), \qquad n = 1, 2, \ldots,$$

とおく．マルコフ性により $n = 1, 2, \ldots$ に対して独立で同一の分布に従う．以下では，$\mathbb{E}(Y_1) < \infty$ の場合と $\mathbb{E}(Y_1) = \infty$ の場合に分けて証明する．初めに $\mathbb{E}(Y_1) < \infty$ を仮定する．この場合には，大数の強法則により，

$$\lim_{m\to\infty} \frac{1}{n} \sum_{\ell=1}^{m} Y_\ell = \mathbb{E}(Y_1) \tag{12.29}$$

が確率 1 で成り立つ．任意の $n \geq 1$ に対して，$\tau_i(n_0 - 1) \leq n < \tau_i(n_0)$ を満たす n_0 があり，$n \to \infty$ のとき $n_0 \to \infty$ である．更に，大数の強法則より，

$$\lim_{n\to\infty} \frac{n_0}{n} = \lim_{n_0\to\infty} \frac{n_0}{\tau_i(n_0)} = \frac{1}{\mu_{ii}} \tag{12.30}$$

が確率 1 で成り立つ．一方，

$$\frac{n_0 - 1}{n} \frac{1}{n_0 - 1} \sum_{m=1}^{n_0-1} Y_m \leq \frac{1}{n} \sum_{\ell=1}^{n} f(X_\ell) \leq \frac{n_0}{n} \frac{1}{n_0} \sum_{m=1}^{n_0} Y_m$$

であるから，この式の各辺を $n \to \infty$ とすれば，(12.29) と (12.30) より

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\ell=1}^{n} f(X_\ell) = \frac{1}{\mu_{ii}} \mathbb{E}(Y_1) \tag{12.31}$$

が確率 1 で成り立つ．一方，

$$
\begin{aligned}
\frac{1}{\mu_{ii}}\mathbb{E}(Y_1) &= \frac{1}{\mu_{ii}}\mathbb{E}\left(\sum_{\ell=1}^{\tau_i} f(X_\ell)\middle| X_0 = i\right) \\
&= \frac{1}{\mu_{ii}}\mathbb{E}\left(\sum_{\ell=1}^{\infty} f(X_\ell)1(\ell \le \tau_i)\middle| X_0 = i\right) \\
&= \frac{1}{\mu_{ii}}\mathbb{E}\left(\sum_{\ell=1}^{\infty}\sum_{j\in S} f(X_\ell)1(X_\ell = j, \ell \le \tau_i)\middle| X_0 = i\right) \\
&= \frac{1}{\mu_{ii}}\sum_{j\in S} f(j)\mathbb{E}\left(\sum_{\ell=1}^{\infty} 1(X_\ell = j, \ell \le \tau_i)\middle| X_0 = i\right) \\
&= \sum_{j\in S} f(j)\pi(j) \qquad (12.32)
\end{aligned}
$$

である．ここに最後の等式は，定理 **??**と 12.4 より得られる．従って，(12.31) より (12.28) が確率 1 で成り立つ．なお，このとき，$\sum_{j\in S} f(j)\pi(j) < \infty$ である．最後に，$\mathbb{E}(Y_1) = \infty$ の場合を証明する．(12.32) より，$\sum_{j\in S} f(j)\pi(j) = \infty$ である．この場合には，正の整数 k に対して関数 $f^{(k)}$ を $f^{(k)}(i) = \min(k, f(i))$ により定義し，関数 f を関数 $f^{(k)}$ と入れ換える．このとき，

$$
\sum_{j\in S} f^{(k)}(j)\pi(j) \le k < \infty
$$

であるから，$\sum_{j\in S} f(j)\pi(j)$ が有限の場合の結果が適用でき，

$$
\lim_{n\to\infty}\frac{1}{n}\sum_{\ell=1}^{n} f^{(k)}(X_\ell) = \sum_{j\in S} f^{(k)}(j)\pi(j)
$$

が確率 1 で成り立つ．従って，$f(i) \ge f^{(k)}(i)$ より

$$
\liminf_{n\to\infty}\frac{1}{n}\sum_{\ell=1}^{n} f(X_\ell) \ge \liminf_{n\to\infty}\frac{1}{n}\sum_{\ell=1}^{n} f^{(k)}(X_\ell) = \sum_{j\in S} f^{(k)}(j)\pi(j)
$$

が得られる．この式の右辺は $k\to\infty$ のとき単調収束定理により $\sum_{j\in S} f(j)\pi(j)$ に収束するので発散する．従って，この場合も (12.28) が確率 1 で成り立つ．　∎

12.7　定常分布の存在条件

既約で非周期的マルコフ連鎖の状態空間 S が有限ならば，演習問題 12.2 より定常分布が唯 1 つ存在し，全ての状態が正再帰的である．一般に，状態空間が有限である限り，定常分布の存在は有限個の連立方程式である定常方程式を解くことにより容易に確かめることができる．しかし，$|S| = \infty$ の場合は簡単ではない．定常方程式を解くことは一般にできないので，何か別の方法が必要である．この場合に定常分布の存在を確かめる方法を考えてみよう．

マルコフ連鎖 $\{X_n; n \geq 0\}$ は状態空間 S と定常な推移確率 $\{p_{ij}; i, j \in S\}$ をもち，既約かつ非周期的であるとする．定理 12.6 により，定常分布は各状態が正再帰的な場合のみに存在する．従って，ある $i \in S$ に対して $\mu_{ii} \equiv \mathbb{E}(\tau_i | X_0 = i) < \infty$ が証明できれば良い．これは状態 i から出発したマルコフ連鎖が確率 1 で十分に速く i へ戻って来れば良いと考えられる．これを，以下の 2 条件を満たす S から $\mathbb{R}_+$ への関数 f を使って調べる．

(i) ある $\delta > 0$ に対して，$A_\delta \equiv \{i \in S; f(i) \leq \delta\}$ が有限集合である．

(ii) $i \in S$ に対して $f(i)$ は状態 i から原点への距離を表す．

この関数 f に対して，$n \to \infty$ のとき $f(X_n)$ が δ 以内になれば $X_n \in A_\delta$ となるので，$i \in A_\delta$ に対して $\mu_{ii} < \infty$ が期待できる．一般に，条件 (i), (ii) を満たす関数を Lyapunov 関数と呼ぶ．

定理 12.7 状態空間 S と推移確率 $\{p_{ij} : i, j \in S\}$ をもつマルコフ連鎖 $\{X_n; n \geq 0\}$ が既約かつ非周期的であるとする．このとき，上記の条件 (i) を満たす S から $\mathbb{R}_+$ への関数 f に対して，すべての $n \geq 1$ に対して $\mathbb{E}(f(X_n))$ が有限であり，ある $\varepsilon > 0$ があり，

$$\forall n \geq 0, X_n \notin A_\delta \text{ ならば，} \quad \mathbb{E}(f(X_{n+1}) | X_n) - f(X_n) \leq -\varepsilon \tag{12.33}$$

が確率 1 で成り立つならば，マルコフ連鎖は正再帰的である．

証明 このマルコフ連鎖が適合するフィルトレーションを $\mathbb{F}$ とし，$\sigma_\delta = \inf\{n \geq 1; f(X_n) \leq \delta\}$ とおく．σ_δ は $\mathbb{F}$-停止時刻である．初めに，任意の $i \in A_\delta$ に対して $\mathbb{E}(\sigma_\delta | X_0 = i) < \infty$ を証明しよう．任意の $n \geq 1$ に対して，$\{n \leq \sigma_\delta\} = \{\sigma_\delta \leq n-1\}^c \in \mathscr{F}_{n-1}$ であるから，マルコフ性と (12.33) を用いると，

$$\begin{aligned}
&\mathbb{E}(f(X_n) 1(n \leq \sigma_\delta) | \mathscr{F}_{n-1}) - f(X_{n-1}) 1(n-1 \leq \sigma_\delta) \\
&\quad = \mathbb{E}(f(X_n) | \mathscr{F}_{n-1}) 1(n \leq \sigma_\delta) - f(X_{n-1}) 1(n-1 \leq \sigma_\delta) \\
&\quad = \mathbb{E}(f(X_n) | X_{n-1}) 1(n \leq \sigma_\delta) - f(X_{n-1}) 1(n-1 \leq \sigma_\delta) \\
&\quad \leq (f(X_{n-1}) - \varepsilon) 1(n \leq \sigma_\delta) - f(X_{n-1}) 1(n-1 \leq \sigma_\delta) \\
&\quad = -\varepsilon 1(n \leq \sigma_\delta) - f(X_{n-1}) 1(n-1 = \sigma_\delta) \leq -\varepsilon 1(n \leq \sigma_\delta)
\end{aligned}$$

が得られる．この式の両辺を $X_0 = i$ の下で条件付期待値を取り，$\ell \geq 1$ に対して $n = 1, 2, \ldots, \ell$ まで加えると，

$$\mathbb{E}(f(X_n) 1(\ell \leq \sigma_\delta) | X_0 = i) - \mathbb{E}(f(X_0) | X_0 = i) \leq -\varepsilon \sum_{n=1}^{\ell} \mathbb{E}(1(n \leq \sigma_\delta) | X_0 = i)$$

である．左辺の第 1 項は非負であり，$\sum_{n=1}^{\ell} \mathbb{E}(1(n \leq \sigma_\delta) | X_0 = i) = \mathbb{E}(\ell \wedge \sigma_\delta | X_0 = i)$ であるから，

$$\varepsilon \mathbb{E}(\ell \wedge \sigma_\delta | X_0 = i) \leq \mathbb{E}(f(X_0) | X_0 = i) = f(i) \leq \delta$$

が成り立つ．この式で $\ell \to \infty$ とすれば，$\mathbb{E}(\sigma_\delta | X_0 = i) \le \delta/\varepsilon < \infty$ が任意の $i \in A_\delta$ 対して成り立つ．従って，A_δ が有限集合であることから，

$$\max_{i \in A_\delta} (\mathbb{E}(\sigma_\delta | X_0 = i) < \infty \tag{12.34}$$

が得られる．この結果を用いて，ある $j \in A_\delta$ に対して，$\mathbb{E}(\tau_j | X_0 = j) < \infty$ であることを示そう．σ_δ は $\mathbb{F}$-停止時刻であるから，$\sigma_\delta(0) = 0$，$\sigma_\delta(n)$ を n 回目に A_δ へ戻った時刻とし $Y_n = X_{\sigma_\delta(n)}$ とおくと，強マルコフ性により $\{Y_n; n \ge 0\}$ もマルコフ連鎖であり，その状態空間 A_δ は有限である．従って，正再帰的状態が少なくとも 1 つある（演習問題 12.2 参照)．この状態を j とし，$\tau_j^Y = \inf\{n \ge 1; Y_n = j\}$ とおく．正再帰性より $\mathbb{E}(\tau_j^Y | Y_0 = j) < \infty$ である．このとき，$\{Y_n; n \ge 0\}$ が適合しマルコフ性の条件を満たすフィルトレーションを $\mathbb{F}^Y \equiv \{\mathscr{F}_n^Y; n \ge 0\}$ とすれば，$\{n \le \tau_j^Y$ は $\mathscr{F}_{n-1}^Y$ 可測であるから，

$$\begin{aligned}
\mathbb{E}(\tau_j | X_0 = j) &= \mathbb{E}\left(\sum_{n=1}^{\tau_j^Y} \sigma_\delta(n) \middle| Y_0 = j \right) \\
&= \mathbb{E}\left(\sum_{n=1}^{\infty} 1(n \le \tau_j^Y) \mathbb{E}(\sigma_\delta(n) | Y_{n-1}) \middle| Y_0 = j \right) \\
&\le \max_{i \in A_\delta} (\mathbb{E}(\sigma_\delta(n) | Y_{n-1} = i)) \mathbb{E}\left(\sum_{n=1}^{\infty} 1(n \le \tau_j^Y) \middle| Y_0 = j \right) \\
&\le \max_{i \in A_\delta} (\mathbb{E}(\sigma_\delta | X_0 = i)) \mathbb{E}\left(\tau_j^Y \middle| Y_0 = j \right)
\end{aligned}$$

となる．この式の最後の右辺は (12.34) より有限であるから，$\mu_{jj} \equiv \mathbb{E}(\tau_j | X_0 = j) < \infty$ が示せた．　■