マルコフ連鎖の基本

@litharge3141

2020 年 5 月 10 日

1　マルコフ連鎖の基本

有限ないし高々可算の状態空間を持ち，かつ離散的という最もシンプルな場合を通して，マルコフ過程の概念を整理する．大数の強法則を証明したことがある程度の知識を仮定する．本文全体を通して I を高々可算集合とし，その σ-代数として I の部分集合全体 2^I を取る．また，自然数の全体 $\mathbb{N}$ は 0 を含むものとする．

1.1　マルコフ連鎖の定義

時間変化する確率変数は，各時刻で I 上の確率分布を与える．それが時間無限大でどうなるかとか，そういう問題を考えたい．I は高々可算だから，I 上の分布は次のように言い換えられる．

Definition 1.1. 写像 $\nu\colon I \to [0,1]$ が確率ベクトルであるとは，$\sum_{i\in I} \nu(i) = 1$ が成立することをいう．ν を $(\nu_i)_{i\in I}$ ともかき，$\nu(i)$ を単に ν_i とかく．写像 $A\colon I \times I \to [0,1]$ が確率行列であるとは，任意の $i \in I$ に対して，$\sum_{j\in I} A(i,j) = 1$ が成立することをいう．A を $(A_{ij})_{i,j\in I}$ ともかき，$A(i,j)$ を単に A_{ij} とかく．

確率ベクトル ν に対して $P\colon 2^I \to [0,1]$ を $P(E) \coloneqq \sum_{i\in E} \nu_i$ により定めると，P は I 上の確率測度となる．逆に I 上の確率測度 P に対して $\nu_i \coloneqq P(\{i\})$ と定めると，ν は確率ベクトルになる．したがって，I 上の確率分布を定めることは，確率ベクトルを定めることと同じである．

Example 1.1. $i \in I$ に対して，確率ベクトル δ_i を

$$\delta_i(j) = \begin{cases} 1 & (j = i) \\ 0 & (j \neq i) \end{cases}$$

によって定めることができる．この記号 δ_i は後で用いる．

Theorem 1.1. 確率ベクトル ν と確率行列 A の積 $\nu A\colon I \to [0,1]$ を $j \in I$ に対し $\nu A(j) \coloneqq \sum_{i\in I} \nu_i A_{ij}$ によって定めることができ，νA は確率ベクトルになる．確率行列 A, B の積 $AB\colon I \times I \to [0,1]$ を $AB(i,j) = \sum_{k\in I} A_{ik} B_{kj}$ によって定めることができ，AB は確率行列になる．

Proof. 非負項の二重級数はいつでも和をとる順序を交換できるので，定理が従う．□

確率行列は確率ベクトルの変換を定めるが，これは分布の変換を定めているのと同じことである．

Example 1.2 (破産問題)**.** $I = \mathbb{Z}$ とし，X_n を n 回目の試行の後の所持金とする．確率 $1/2$ で $X_{n+1} = X_n + 1$ とし，確率 $1/2$ で $X_{n+1} = X_n - 1$ とするような賭けを考える．最初の所持金を i とすると，初期分布は δ_i で

与えられる．この試行の確率行列は

$$A_{ij} = \begin{cases} 1/2 & (j = i \pm 1) \\ 0 & (\text{それ以外}) \end{cases}$$

によって与えられる．

初期分布と時間によらない分布の変換が与えられているとき，それにしたがって発展するような分布を持つ確率変数列のことをマルコフ連鎖という．素直に数式で表現すると次のようになる．

Definition 1.2 (Markov 連鎖っぽいなにか)**.** 確率ベクトル ν と確率行列 A が与えられたとする．$(\Omega, \mathcal{F}, P)$ を確率空間として，Ω から I への確率変数列 $(X_n)_{n=0}^{\infty}$ が遷移行列 A，初期分布 ν をもつマルコフ連鎖であるとは，任意の $E \subset I$ に対して $P(X_0 \in E) = \sum_{i \in E} \nu_i$ が成立し，さらに任意の $n \in \mathbb{N}$ と任意の $E \subset I$ に対して $P(X_{n+1} \in E) = \sum_{i \in E} \sum_{j \in I} P(X_n = j) A_{ji}$ が成立することをいう．

この定義は任意の $i \in I$ に対して $P(X_{n+1} = i) = \sum_{j \in I} P(X_n = j) A_{ji}$ が成立すること，と書き換えてもよい．この定義は次の定義を採用すればそれから導かれる．

Definition 1.3 (Markov 連鎖)**.** 確率ベクトル ν と確率行列 A が与えられたとする．$(\Omega, \mathcal{F}, P)$ を確率空間として，Ω から I への確率変数列 $(X_n)_{n=0}^{\infty}$ が遷移行列 A，初期分布 ν をもつマルコフ連鎖であるとは，任意の $n \in \mathbb{N}$ と任意の $i_0, \ldots, i_n \in I$ に対して，$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \nu_{i_0} A_{i_0 i_1} A_{i_1 i_2} \cdots A_{i_{n-1} i_n}$ が成立することをいう．マルコフ連鎖を初期分布と遷移行列，確率測度との組にして $((X_n)_{n=0}^{\infty}, A, \nu, P)$ と書き表すこともある．

この定義は任意の $i, j \in I$ に対して $P(X_0 = i) = \nu_i$ および $P(X_{n+1} = i \mid X_n = j) = A_{ji}$ が成立すること，と書き換えてもよい．っぽいなにかのほうだと後で証明が回らなくなるようなので，このノートではこちらの定義を採用する．実は同値だったとか，具体的にどこがまずいのかとか分かったら追記する．マルコフ連鎖が常に存在するかは自明ではないので，存在を示す．

Theorem 1.2. 確率ベクトル ν と確率行列 A が与えられたとする．このとき，ある確率空間 $(\Omega, \mathcal{F}, P)$ と Ω から I への確率変数列 $(X_n)_{n=0}^{\infty}$ が存在して，遷移行列 A，初期分布 ν をもつマルコフ連鎖となる．

Proof. コルモゴロフの拡張定理を使うために，便宜的に $\mathbb{R}^n$ 上の確率測度を構成する．I は高々可算集合，したがって単射 $\phi\colon I \to \mathbb{N}$ が存在する．$n = 1, 2, \ldots$ に対して写像 $\mu_n\colon \mathcal{B}(\mathbb{R}^n) \to \mathbb{R}$ を $\mu_n(E) \coloneqq \sum_{(\phi(i_0), \ldots, \phi(i_{n-1})) \in E} \nu_{i_0} A_{i_0 i_1} A_{i_1 i_2} \cdots A_{i_{n-2} i_{n-1}}$ と定めると，これは $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上の確率測度になる．さらに拡張定理の仮定である整合条件 $\mu_{n+1}(E \times \mathbb{R}) = \mu_n(E)$ が満たされることも示せるので，コルモゴロフの拡張定理から $\mathbb{R}^{\mathbb{N}}$ 上の確率測度 μ で任意の $n = 1, 2, \ldots$ と任意の $A \in \mathcal{B}(\mathbb{R}^n)$ に対して $\mu(A \times \mathbb{R}^{\mathbb{N}}) = \mu_n(A)$ を満たすものが一意的に存在する．$n = 1, 2, \ldots$ に対して $Z_n\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ を n 成分への射影 $Z_n((x_1, x_2, \ldots)) = x_n$ によって定めて，$n \in \mathbb{N}$ に対して $Y_n \coloneqq Z_{n+1}$ とする．$(Y_n)_{n=0}^{\infty}$ はほとんどいたるところ $\phi(I)$ に値を取る確率変数列で，任意の $n \in \mathbb{N}$ と任意の $i_0, \ldots, i_n \in I$ に対して，$P(Y_0 = \phi(i_0), Y_1 = \phi(i_1), \ldots, Y_n = \phi(i_n)) = \nu_{i_0} A_{i_0 i_1} A_{i_1 i_2} \cdots A_{i_{n-1} i_n}$ を満たす．零集合上の値を修正して I 値にすれば，求めるマルコフ連鎖 $(X_n)_{n=0}^{\infty}$ が得られる． □

しばしば現れてきた $\{\omega \mid X_0(\omega) = i_0, \ldots, X_n(\omega) = i_n\}$ のような形の事象全体で作られる，自然な増大情報系という概念を導入しておくと便利である．

Definition 1.4. $(\Omega, \mathcal{F}, P)$ を確率空間として，Ω から I への確率変数列 $X = (X_n)_{n=0}^{\infty}$ が与えられたとする．$n \in \mathbb{N}$ に対して $\mathcal{F}$ の部分 σ-代数 $\mathcal{F}_n$ を $\mathcal{F}_n \coloneqq \sigma(X_0, \dots, X_n)$ と定める．$(\mathcal{F}_n)_{n=0}^{\infty}$ を X に関する自然な増大情報系という．

I は高々可算集合なので，$F \subset I^n$ に対して $E = (X_0, \dots, X_n)^{-1}(F)$ という形でかける E は $E = \bigcup_{(i_0,\dots,i_n)\in F}\{X_0 = i_0, \dots, X_n = i_n\}$ という交わらない可算和で書き直せる．このような元全体で生成されるのが $\mathcal{F}_n$ である．各時間 n において，X に関連する事象で確率を計算し得るものは全て含む．そのような中で最小のもの，というのが自然という言葉の意味である．

マルコフ連鎖は時刻 $n+1$ での確率分布が時刻 n での分布にのみ依存するという性質を持つ．これをマルコフ性という．それを示そう．

Theorem 1.3 (マルコフ性 1)**.** $(\Omega, \mathcal{F}, P)$ を確率空間として，Ω から I への確率変数列 $(X_n)_{n=0}^{\infty}$ は遷移行列 A，初期分布 ν をもつマルコフ連鎖であるとする．このとき，$P(X_n = i_n, \dots, X_0 = i_0) > 0$ が成り立つような任意の $n \in \mathbb{N}$ と任意の $i_0, \dots, i_{n+1} \in I$ に対して，$P(X_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) = A_{i_n i_{n+1}}$ が成立する．

Proof. $n \in \mathbb{N}$ と $i_0, \dots, i_{n+1} \in I$ が任意に与えられたとすると，$P(X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0)P(X_n = i_n, \dots, X_0 = i_0) = P(X_{n+1}, X_n = i_n, \dots, X_0 = i_0)$ が成り立つ．$(X_n)_{n=0}^{\infty}$ がマルコフ連鎖であるから，$P(X_n = i_n, \dots, X_0 = i_0) = \nu_{i_0} A_{i_0 i_1} \cdots A_{i_{n-1} i_n}$ および $P(X_{n+1}, X_n = i_n, \dots, X_0 = i_0) = \nu_{i_0} A_{i_0 i_1} \cdots A_{i_n i_{n+1}}$ となる．これを代入して計算をすれば定理を得る． □

マルコフ性は別の定式化をすることもできる．次の性質も成り立つ．

Theorem 1.4 (マルコフ性 2)**.** $(\Omega, \mathcal{F}, P)$ を確率空間として，Ω から I への確率変数列 $(X_n)_{n=0}^{\infty}$ は遷移行列 A，初期分布 ν をもつマルコフ連鎖であるとする．$P(X_m = i) > 0$ が成り立つような任意の $i \in I$ と任意の $m \in \mathbb{N}$ に対して，$Y_n \coloneqq X_{n+m}$ として $(Y_n)_{n=0}^{\infty}$ を定めると，$((Y_n)_{n=0}^{\infty}, A, \delta_i, P(\cdot \mid X_m = i))$ はマルコフ連鎖となる．

Proof. $i \in \mathbb{N}$ と $i_m, i_{m+1}, \dots, i_{m+n} \in I$ が任意に与えられたとする．

$$
\begin{aligned}
&P(Y_0 = i_m, \dots, Y_n = i_{m+n} \mid X_m = i) \\
&= P(X_m = i, X_m = i_m, \dots, X_{m+n} = i_{m+n})/P(X_m = i) \\
&= \frac{\sum_{i_0,\dots,i_{m-1}\in I} P(X_0 = i_0, \dots, X_{m-1} = i_{m-1}, X_m = i, X_m = i_m, \dots, X_{m+n} = i_{m+n})}{\sum_{j_0,\dots,j_{m-1}\in I} P(X_0 = j_0, \dots, X_{m-1} = j_{m-1}, X_m = i)} \\
&= \frac{\sum_{i_0,\dots,i_{m-1}\in I} \delta_i(i_m)\nu_{i_0} A_{i_0 i_1} \cdots A_{i_{m+n-1} i_{m+n}}}{\sum_{i_0,\dots,i_{m-1}\in I} \nu_{j_0} A_{j_0 j_1} \cdots A_{j_{m-1} i}} \\
&= \delta_i(i_m) A_{i_m i_{m+1}} \cdots A_{i_{m+n-1} i_{m+n}} \frac{\sum_{i_0,\dots,i_{m-1}\in I} \nu_{i_0} A_{i_0 i_1} \cdots A_{i_{m-1} i}}{\sum_{i_0,\dots,i_{m-1}\in I} \nu_{j_0} A_{j_0 j_1} \cdots A_{j_{m-1} i}} \\
&= \delta_i(i_m) A_{i_m i_{m+1}} \cdots A_{i_{m+n-1} i_{m+n}}
\end{aligned}
$$

したがって，示された． □

Theorem1.4 を使うとより強い次の結果を示すことができる．

Theorem 1.5 ($\mathcal{F}_m$ との独立性)**.** $(\Omega, \mathcal{F}, P)$ を確率空間として，Ω から I への確率変数列 $(X_n)_{n=0}^{\infty}$ は遷移行

列 A，初期分布 ν をもつマルコフ連鎖であるとする．$P(X_m = i) > 0$ が成り立つような任意の $i \in I$ と任意の $m \in \mathbb{N}$ に対して，$Y_n \coloneqq X_{n+m}$ として $(Y_n)_{n=0}^{\infty}$ を定める．このとき，確率空間 $(\Omega, \mathcal{F}, P(\cdot \mid X_m = i))$ の下で $(Y_n)_{n=0}^{\infty}$ と $\mathcal{F}_m$ は独立である．

Proof. $n \in \mathbb{N}$ と $j_0, j_1, \ldots, j_m, i_m, i_{m+1}, \ldots, i_{m+n} \in I$ が任意に与えられたとする．$E \coloneqq \{X_0 = j_0, \ldots, X_m = j_m\}$ とおく．$E \in \mathcal{F}_m$ は定義から明らかである．この E に対して

$$\begin{aligned} &P((Y_0 = i_m, \ldots, Y_n = i_{m+n}) \cap E \mid X_m = i) \\ &= P(Y_0 = i_m, \ldots, Y_n = i_{m+n} \mid X_m = i) P(E \mid X_m = i) \end{aligned} \tag{1}$$

を示す．$(X_n)_{n=0}^{\infty}$ のマルコフ性から

$$\begin{aligned} &P((Y_0 = i_m, \ldots, Y_n = i_{m+n}) \cap E \mid X_m = i) \\ &= \delta_i(i_m) P(X_0 = j_0, \ldots, X_m = j_m, X_m = i_m, \ldots, X_{m+n} = i_{m+n}) / P(X_m = i) \\ &= \delta_{j_m i_m} \delta_i(i_m) \nu_{j_0} A_{j_0 j_1} \cdots A_{i_{m+n-1} i_{m+n}} / P(X_m = i) \end{aligned}$$

および

$$\begin{aligned} P(E \mid X_m = i) &= \delta_i(j_m) P(X_0 = j_0, \ldots, X_{m-1} = j_{m-1}, X_m = j_m) / P(X_m = i) \\ &= \delta_i(j_m) \nu_{j_0} A_{j_0 j_1} \cdots A_{j_{m-1} j_m} / P(X_m = i) \end{aligned}$$

が得られる．また Theorem1.4 により，

$$P(Y_0 = i_m, \ldots, Y_n = i_{m+n} \mid X_m = i) = \delta_i(i_m) A_{i_m i_{m+1}} \cdots A_{i_{m+n-1} i_{m+n}}$$

となる．これらの式から (1) が得られた．一般の $E \in \mathcal{F}_m$ に対して証明するため，上の結果を利用する．

$$\begin{aligned} \mathcal{C} \coloneqq \{E \mid & \forall n \in \mathbb{N},\ \forall i_m, \ldots, i_{m+n} \in I, \\ & P((Y_0 = i_m, \ldots, Y_n = i_{m+n}) \cap E \mid X_m = i) \\ & = P(Y_0 = i_m, \ldots, Y_n = i_{m+n} \mid X_m = i) P(E \mid X_m = i)\} \end{aligned}$$

とおく．$E = \{X_0 = j_0, \ldots, X_m = j_m\}$ の形で表される E の全体は $\mathcal{C}$ に含まれるから，$\mathcal{C}$ が σ-代数であることを示せばよいが，それは $P(\cdot \mid X_m = i)$ が確率測度であることから直ちにしたがう．よって示された． $\square$

1.2 到達確率と差分作用素

マルコフ性の応用として到達確率と差分作用素を扱う．

Definition 1.5. $(\Omega, \mathcal{F}, P)$ を確率空間として，Ω から I への確率変数列 $(X_n)_{n=0}^{\infty}$ は遷移行列 A，初期分布 ν をもつマルコフ連鎖であるとする．$E \subset I$ に対して，到達時刻 $\tau_E \colon \Omega \to \mathbb{N} \cup \{\infty\}$ を

$$\tau_E(\omega) \coloneqq \inf\{n \in \mathbb{N} \mid X_n(\omega) \in E\}$$

によって定める．ただし，$\inf \emptyset \coloneqq \infty$ とする．

Theorem 1.6. $(\Omega, \mathcal{F}, P)$ を確率空間として，Ω から I への確率変数列 $(X_n)_{n=0}^{\infty}$ は遷移行列 A，初期分布 ν をもつマルコフ連鎖であるとする．任意の $E \subset I$ に対して，到達時刻 τ_E は $\mathbb{N} \cup \{\infty\}$ の部分集合全体を σ-代数として可測であり，さらに任意の $n \in \mathbb{N}$ に対して $\{\tau_E = n\} \in \mathcal{F}_n$ となる．

Proof. $\mathbb{N} \cup \{\infty\}$ は可算集合だから，τ_E による一点集合の引き戻しが可測であることを示せば十分である．$\{\tau_E = \infty\} = \bigcap_{n\in\mathbb{N}} X_n^{-1}(E^c)$ は明らかに可測集合である．$n \in \mathbb{N}$ が任意に与えられたとして，$\{\tau_E = n\} \in \mathcal{F}_n$ を示す．E および E^c は高々可算集合だから，値で場合分けをして

$$\{\tau_E = n\} = \bigcup_{i_0 \in E^c, \dots, i_{n-1} \in E^c, i_n \in E} \{X_0 = i_0, \dots, X_n = i_n\}$$

となる．したがって $\{\tau_E = n\} \in \mathcal{F}_n$ である． □

Definition 1.6. 確率行列 A と $E \subset I$ が与えられたとする．マルコフ連鎖の E への到達確率 $e\colon I \to \mathbb{R}$ を，$i \in I$ に対して，初期分布 δ_i，遷移行列 A を持つ任意のマルコフ連鎖 (X, A, δ_i, P) を用いて $e(i) \coloneqq P(\tau_E < \infty \mid X_0 = i)$ によって定める．$e(i)$ を e_i ともかく．

与えられているのは確率行列 A と初期分布 δ_i のみだから，Theorem1.2 から，適切な確率空間を設定すれば少なくとも一つはマルコフ連鎖が存在する．マルコフ連鎖の取り方によらないことを確かめる．

Theorem 1.7. この定義は *well-defined*，すなわちマルコフ連鎖の取り方には依存しない．

Proof. $i \in I$ が任意に与えられたとする．(X, A, δ_i, P_0) および (Y, A, δ_i, P_1) をマルコフ連鎖とする．X に対する到達時刻を τ_E^0 とかき，到達確率を e_0 とする．Y に対する到達確率を τ_E^1 とかき，到達確率を e_1 とする．

$$\begin{aligned}
e_0(i) &= P_0(\tau_E^0 < \infty \mid X_0 = i) \\
&= \sum_{n=0}^{\infty} P_0(\tau_E^0 = n \mid X_0 = i) \\
&= \sum_{n=0}^{\infty} \sum_{i_0 \notin E, \dots, i_{n-1} \notin E, i_n \in E} P_0(X_0 = i_0, \dots, X_n = i_n \mid X_0 = i) \\
&= \sum_{n=0}^{\infty} \sum_{i_0 \notin E, \dots, i_{n-1} \notin E, i_n \in E} \delta_i(i_0) A_{i_0 i_1} \cdots A_{i_{n-1} i_n} \\
&= \sum_{n=0}^{\infty} \sum_{i_0 \notin E, \dots, i_{n-1} \notin E, i_n \in E} P_1(Y_0 = i_0, \dots, Y_n = i_n \mid Y_0 = i) \\
&= \sum_{n=0}^{\infty} P_1(\tau_E^1 = n \mid Y_0 = i) \\
&= P_1(\tau_E^1 < \infty \mid Y_0 = i) \\
&= e_1(i)
\end{aligned}$$

となるから，示された． □

たとえ最初に $P(X_0 = i) = \nu_i = 0$ であったとしても，Theorem1.4 から，$P(X_m = i) = \nu_i > 0$ となる $m \in \mathbb{N}$ が存在すれば，$(Y_n)_{n=0}^{\infty} = (X_{n+m})_{n=0}^{\infty}$ を初期分布が δ_i となるように取り直すことができる．この定義はこの取り直しを念頭に置いている．

Definition 1.7 (差分作用素)**.** 確率行列 A が与えられたとする．任意の $i \in I$ に対し $\sum_{j\in I} A_{ij} f(j)$ が絶対収束する $f\colon I \to \mathbb{R}$ に対し，実数値関数 $\mathcal{L}f$ を対応させる差分作用素 $\mathcal{L}$ を $\mathcal{L}f(i) \coloneqq \sum_{j\in I} A_{ij} f(j) - f(i)$ によって定める．

以前に定義した確率ベクトルと確率行列との積は横ベクトルとみてのものであり，この定義に登場するものとは微妙に異なることに注意する．次の定理が目標である．

Theorem 1.8 (境界値問題)**.** 確率行列 A と $E \subset I$ が与えられたとする．このとき，到達確率 e は

$$\begin{cases} \mathcal{L}e(i) = 0, & i \notin E \\ e(i) = 1, & i \in E \end{cases}$$

を満たす最小の非負実数値関数である．

Proof. まず e が方程式を満たすことを示す．$i \in E$ のとき $e(i) = 1$ は明らかだから，$i \notin E$ とする．以下の式中に出てくる X は，初期分布が δ_i や δ_{i_1} で，遷移行列 A を持つようなマルコフ連鎖である．

$$\begin{aligned}
e(i) &= P(\tau_E < \infty \mid X_0 = i) \\
&= \sum_{n=1}^{\infty} P(\tau_E = n \mid X_0 = i) \quad (\because i \notin E) \\
&= \sum_{n=1}^{\infty} \sum_{i_0 \notin E, \ldots, i_{n-1} \notin E, i_n \in E} P(X_0 = i_0, \ldots, X_n = i_n \mid X_0 = i) \\
&= \sum_{n=1}^{\infty} \sum_{i_0 \notin E, \ldots, i_{n-1} \notin E, i_n \in E} \delta_i(i_0) A_{i_0 i_1} \cdots A_{i_{n-1} i_n} \quad (\because \text{Theorem1.4}) \\
&= \sum_{n=2}^{\infty} \sum_{i_1 \notin E, \ldots, i_{n-1} \notin E, i_n \in E} A_{i i_1} A_{i_1 i_2} \cdots A_{i_{n-1} i_n} + \sum_{i_1 \in E} A_{i i_1} \\
&= \sum_{n=3}^{\infty} \sum_{i_1 \notin E} A_{i i_1} \sum_{i_2 \notin E, \ldots, i_{n-1} \notin E, i_n \in E} \delta_{i_1}(i_1) A_{i_1 i_2} \cdots A_{i_{n-1} i_n} \\
&\quad + \sum_{i_1 \in E} A_{i i_1} + \sum_{i_1 \notin E, i_2 \in E} A_{i i_1} A_{i_1 i_2} \\
&= \sum_{n=3}^{\infty} \sum_{i_1 \notin E} A_{i i_1} \sum_{i_2 \notin E, \ldots, i_{n-1} \notin E, i_n \in E} P(X_0 = i_1, \ldots, X_{n-1} = i_n \mid X_0 = i_1) \quad (\because \text{Theorem1.4}) \\
&\quad + \sum_{i_1 \in E} A_{i i_1} P(\tau_E < \infty \mid X_0 = i_1) + \sum_{i_1 \notin E} A_{i i_1} P(\tau_E = 1 \mid X_0 = i_1) \\
&= \sum_{n=3}^{\infty} \sum_{i_1 \notin E} A_{i i_1} P(\tau_E = n-1 \mid X_0 = i_1) \\
&\quad + \sum_{i_1 \in E} A_{i i_1} P(\tau_E < \infty \mid X_0 = i_1) + \sum_{i_1 \notin E} A_{i i_1} P(\tau_E = 1 \mid X_0 = i_1) \\
&= \sum_{i_1 \notin E} A_{i i_1} \sum_{n=2}^{\infty} P(\tau_E = n-1 \mid X_0 = i_1) + \sum_{i_1 \in E} A_{i i_1} P(\tau_E < \infty \mid X_0 = i_1) \\
&= \sum_{i_1 \notin E} A_{i i_1} P(\tau_E < \infty \mid X_0 = i_1) + \sum_{i_1 \in E} A_{i i_1} P(\tau_E < \infty \mid X_0 = i_1) \\
&= \sum_{j \in I} A_{ij} P(\tau_E < \infty \mid X_0 = j) \\
&= \sum_{j \in I} A_{ij} e(j)
\end{aligned}$$

により示された．最小性を示す．$f\colon I \to \mathbb{R}$ を非負実数値関数とし，方程式を満たすとする．$e \leq f$ を示せばよい．$i \in E$ の時は明らかだから，$i \notin E$ とする．任意の $n \in \mathbb{N}$ に対して

$$f(i) \geq \sum_{k=0}^{n} P(\tau_E = k \mid X_0 = i)$$

を示す．帰納法で証明する．$n = 0$ の時は $i \notin E$ より $P(\tau_E = 0 \mid X_0 = i) = 0 \leq f(i)$ となる．n までの成立を仮定する．

$$\begin{aligned}
f(i) &= \sum_{j \in I} A_{ij} f(j) \\
&\geq \sum_{j \in I} A_{ij} \sum_{k=0}^{n} P(\tau_E = k \mid X_0 = j) \\
&= \sum_{j \notin E} A_{ij} \sum_{k=1}^{n} P(\tau_E = k \mid X_0 = j) \\
&\quad + \sum_{j \in E} A_{ij} \sum_{k=0}^{n} P(\tau_E = k \mid X_0 = j) \\
&= \sum_{j \notin E} \sum_{k=1}^{n} A_{ij} \sum_{i_0 \notin E, \ldots, i_{k-1} \notin E, i_k \in E} P(X_0 = i_0, \ldots, X_k = i_k \mid X_0 = j) \\
&\quad + \sum_{j \in E} A_{ij} \\
&= \sum_{k=1}^{n} \sum_{j \notin E} \sum_{i_0 \notin E, \ldots, i_{k-1} \notin E, i_k \in E} A_{ij} \delta_j(i_0) A_{i_0 i_1} \cdots A_{i_{k-1} i_k} \\
&\quad + \sum_{j \in E} P(X_0 = i, X_1 = j \mid X_0 = i) \\
&= \sum_{k=2}^{n} \sum_{j \notin E, i_1 \notin E, \ldots, i_{k-1} \notin E, i_k \in E} A_{ij} A_{j i_1} \cdots A_{i_{k-1} i_k} \\
&\quad + \sum_{j \notin E, i_1 \in E} A_{ij} A_{j i_1} + P(\tau_E = 1 \mid X_0 = i) \\
&= \sum_{k=2}^{n} \sum_{j \notin E, i_1 \notin E, \ldots, i_{k-1} \notin E, i_k \in E} P(X_0 = i, X_1 = j, X_2 = i_1, \ldots, X_{k+1} = i_k \mid X_0 = i) \\
&\quad + \sum_{j \notin E, i_1 \in E} P(X_0 = i, X_1 = j, X_2 = i_1 \mid X_0 = i) + P(\tau_E = 1 \mid X_0 = i) \\
&= \sum_{k=2}^{n} P(\tau_E = k+1 \mid X_0 = i) + P(\tau_E = 2 \mid X_0 = i) + P(\tau_E = 1 \mid X_0 = i) \\
&= \sum_{k=0}^{n} P(\tau_E = k \mid X_0 = i) \quad (\because i \notin E)
\end{aligned}$$

により，示された．$n \to \infty$ として，

$$f(i) \geq \sum_{n=0}^{\infty} P(\tau_E = n \mid X_0 = i) = e(i)$$

となり，示された． □

この定理により，様々な到達確率を求めることができる．例として，破産問題を扱う．

Example 1.3. $I = \mathbb{Z}$ とし，X_n を n 回目の試行の後の所持金とする．$p + q = 1$ となる非負の実数 p, q に対して，確率 p で $X_{n+1} = X_n + 1$ とし，確率 q で $X_{n+1} = X_n - 1$ とするような賭けを考える．最初の所持金を $i \in \mathbb{Z}$ とすると，初期分布は δ_i で与えられる．この試行の確率行列は

$$A_{ij} = \begin{cases} p & (j = i + 1) \\ q & (j = i - 1) \\ 0 & (\text{それ以外}) \end{cases}$$

によって与えられる．所持金が 0 になった時破産したということにして，破産する確率を求めたい．$E \subset \mathbb{Z}$ を $E = \{m \in \mathbb{Z} \mid m \leq 0\}$ として，i から出発して破産する確率は E への到達確率 $e(i)$ で与えられる．Theorem1.8 により e は

$$\begin{cases} \mathcal{L}e(i) = 0, & i \notin E \\ e(i) = 1, & i \in E \end{cases}$$

を満たす最小の非負実数値関数である．第一式から $i > 0$ に対して漸化式 $pe(i+1) + qe(i-1) = e(i)$ を得る．$e(i+1) - e(i) = (q/p)(e(i) - e(i-1))$ と変形するなどして，

$$e(i) = \begin{cases} C_0 + C_1 i & (p = q) \\ C_0 + C_1 (q/p)^i & (p \neq q) \end{cases}$$

として解くことができる．ここで C_0, C_1 は定数で，非負実数解として最小になるように定める．$p \neq q$ のときを考える．$e(0) = 1$ より $C_0 + C_1 = 1$ であることに注意する．$p < q$ ならば，$0 \leq e(i) \leq 1$ より $C_1 = 0$ となり，$e(i) = 1$ となる．したがってどこから出発しても必ず破産する．$p > q$ ならば $i \to \infty$ を考えると $e(i) \to C_0 \geq 0$ となる．よって最小解となるためには $C_0 = 0$ が必要で，$e(i) = (q/p)^i$ となる．$p = q$ のときは $0 \leq e(i) \leq 1$ より $C_1 = 0$ となり，$e(i) = 1$ となる．したがって必ず破産する．たとえ公平な賭けとしてもこの場合は必ず破産することが分かった．

1.3 エルゴード性

この節では，I は有限集合であることを約束する．これまで見てきた通り確率行列 A さえ指定すれば A を遷移行列とするマルコフ連鎖のほとんどのことはわかるし，実際到達確率は遷移行列を用いた方程式で特徴づけられた．この節では最初に与えた分布が時間無限大の極限においてどうなるのかを考える．

最初に与えた分布を確率ベクトル v と同一視すると，v の時間発展は vA^n によってあらわされる．もしこの極限が存在するならば，$\pi = \lim_{n\to\infty} vA^n$ と置くと $\pi A = \lim_{n\to\infty} vA^{n+1} = \pi$ となる．A がエルゴード性という条件を満たせばこの極限が存在する．このことと，極限が存在する場合に大数の法則が成立することを証明する．

確率行列の積は確率行列だから，A^n も確率行列である．そこで A^n の i, j 成分を A^n_{ij} と書くことにする．

Definition 1.8. A を確率行列とする．A がエルゴード的とは，ある $n_0 \in \mathbb{N}$ が存在して，任意の $i, j \in I$ に対して $A^{n_0}_{ij} > 0$ が成立することをいう．A が既約であるとは，任意の $i, j \in I$ に対してある $n_1 = n_1(i, j) \in \mathbb{N}$ が存在して $A^{n_1}_{ij} > 0$ が成立することをいう．$i \in I$ が A について非周期的であるとは，ある $n_2 = n_2(i) \in \mathbb{N}$ が存在して，任意の $n \geq n_2$ に対して $A^n_{ii} > 0$ が成立することをいう．

エルゴード的な確率行列 A を特徴づける前に，簡単な補題を示す．

Lemma 1.9. A, B を確率行列とする．任意の $i, j \in I$ に対して $B_{ij} > 0$ となるとき，任意の $i, j \in I$ に対して $(AB)_{ij} > 0$ となる．

Proof. $i, j \in I$ が任意に与えられたとする．$\sum_{k\in I} A_{ik} = 1$ だから，ある $k_0 \in I$ が存在して $A_{ik_0} > 0$ となる．したがって $(AB)_{ij} = \sum_{k\in I} A_{ik}B_{kj} \geq A_{ik_0}B_{k_0j} > 0$ となり，示された． □

確率行列 A のエルゴード性を特徴づける定理を述べる．I が有限集合でないと成り立たないことに注意する．

Theorem 1.10. A を確率行列とすると，次は同値．

(1) A はエルゴード的である．
(2) A は既約，かつ全ての $i \in I$ が A について非周期的である．
(3) A は既約，かつある $i_0 \in I$ が存在して A について非周期的である．

Proof. (1) $\Rightarrow$ (2) を示す．A はエルゴード的だから，ある $n_0 \in \mathbb{N}$ が存在して，任意の $i, j \in I$ に対して $A^{n_0}_{ij} > 0$ となる．Lemma1.9 を $B = A^{n_0}$ として用いれば，$n \geq n_0$ のとき，任意の $i, j \in I$ に対して $A^n_{ij} > 0$ である．このことから (2) は直ちに従う．(2) $\Rightarrow$ (3) は明らかである．(3) $\Rightarrow$ (1) を示す．$i_0 \in I$ は A について非周期的だから，$n_2 = n_2(i_0) \in \mathbb{N}$ が存在して $n \geq n_2$ ならば $A^n_{i_0i_0} > 0$ である．A は既約だから，任意の $i, j \in I$ に対し $n_1(i, i_0), n_1(i_0, j) \in \mathbb{N}$ が存在して $A^{n_1(i,i_0)}_{ii_0} > 0$ および $A^{n_2(i_0j)}_{i_0j} > 0$ となる．そこで $n_0 \coloneqq \max_{i,j\in I}\{n_1(i, i_0) + n_2(i_0) + n_1(i_0, j)\}$ と定める．Lemma1.9 を繰り返し用いると，任意の $i, j \in I$ に対して，$n \coloneqq n_0 - n_1(i, i_0) - n_1(i_0, j) \geq n_2(i_0)$ とすると $A^{n_0}_{ij} \geq A^{n_1(i,i_0)}_{ii_0} A^n_{i_0i_0} A^{n_1(i_0,j)}_{i_0j} > 0$ が成立することが分かる．したがって示された． □

本節冒頭に書いた π を不変分布といい，次のように定める．

Definition 1.9. 確率ベクトル π が確率行列 A の不変分布であるとは，$\pi A = \pi$ が成立することをいう．

エルゴード的な確率行列に対しては初期分布によらずに一意的に存在することを証明する．さらに収束の速さまで分かる．いくつか準備をする．不変分布の構成では次の有名な不動点定理の証明と本質的に同じことをする．直接使うわけではないので，証明は省略する．

Theorem 1.11 (Banach の不動点定理)**.** $(X, \|\cdot\|)$ を *Banach* 空間とそのノルムの組とし，線形写像 $T\colon X \to X$ が縮小条件

$$0 < \exists\gamma < 1, \quad \forall x, y \in X, \quad \|T(x - y)\| \leq \gamma\|x - y\|$$

を満たすと仮定する．このとき，$T(x) = x$ を満たす $x \in X$ がただ一つ存在して，任意の $x_0 \in X$ に対して $x_n \coloneqq T^n(x_0)$ で定めた点列は x にノルム収束する．ここで，T^n は T の n 回合成である．

T として確率行列を右からかける写像を取れば良い．あとは Banach 空間 X と縮小条件を満たす定数を取らなければならない．X は次のようにとる．

Definition 1.10. I 上の確率ベクトルの全体のなす集合を $\mathcal{P}(I)$ とかき，その上の実数値関数 $\|\cdot\| : \mathcal{P}(I) \to \mathbb{R}$ を $\nu \in \mathcal{P}(I)$ に対して $\|\nu\| \coloneqq \frac{1}{2}\sum_{i\in I} |\nu_i|$ によって定める．

Theorem 1.12. 以下が成立する．

(1) $(\mathcal{P}(I), \|\cdot\|)$ は *Banach* 空間である．

(2) 任意の $\nu, \mu \in \mathcal{P}(I)$ に対して $\|\nu - \mu\| = \sum_{i\in I}(\nu_i - \mu_i)^+$ が成立する．ここで，実数 a に対して $a^+ := \max\{a, 0\}$ とする．

(3) 任意の $\nu, \mu \in \mathcal{P}(I)$ に対して $\|\nu - \mu\| \leq 1$ が成立する．

Proof. $\mathbb{R}^{|I|}$ の部分空間と考えれば Banach 空間であることは明らかである．$\nu, \mu \in \mathcal{P}(I)$ が任意に与えられたとする．実数 a に対して $a^- := \min\{a, 0\}$ とする．

$$\begin{aligned}
2\|\nu - \mu\| &= \sum_{i\in I}(\nu_i - \mu_i)^+ - \sum_{i\in I}(\nu_i - \mu_i)^- \\
&= \sum_{i\in I, \nu_i \geq \mu_i}(\nu_i - \mu_i) - \sum_{i\in I, \nu_i \leq \mu_i}(\nu_i - \mu_i) \\
&= \sum_{i\in I, \nu_i \geq \mu_i}(\nu_i - \mu_i) - \left(\sum_{i\in I}(\nu_i - \mu_i) - \sum_{i\in I, \nu_i \geq \mu_i}(\nu_i - \mu_i)\right) \\
&= 2\sum_{i\in I, \nu_i \geq \mu_i}(\nu_i - \mu_i) \\
&= 2\sum_{i\in I}(\nu_i - \mu_i)^+
\end{aligned}$$

により，(2) が従う．$\sum_{i\in I}|\nu_i - \mu_i| \leq \sum_{i\in I}\nu_i + \mu_i = 2$ より，(3) も従う． □

縮小条件の定数は次のようにしてとる．

Theorem 1.13. 確率行列 A が与えられたとする．ある $0 < \gamma < 1$ が存在して任意の $i, j \in I$ に対して $A_{ij} \geq \gamma$ となるとき，

$$\forall \nu, \mu \in \mathcal{P}(I), \quad \|(\nu - \mu)A\| \leq (1 - \gamma)\|\nu - \mu\|$$

が成立する．

Proof. $\nu, \mu \in \mathcal{P}(I)$ が任意に与えられたとする．Thoerem1.12 により，$\|(\nu - \mu)A\| = \sum_{j\in I}\left((\nu - \mu)A\right)_j^+ = \sum_{j\in I}\left(\sum_{i\in I}(\nu_i - \mu_i)A_{ij}\right)^+$ となる．任意の $j \in I$ に対して $\sum_{i\in I}(\nu_i - \mu_i)A_{ij} > 0$ とすると，$\sum_{j\in I}\sum_{i\in I}(\nu_i - \mu_i)A_{ij} = \sum_{i\in I}(\nu_i - \mu_i) = 0$ となることに反する．したがってある $j_0 \in I$ が存在して $\sum_{i\in I}(\nu_i - \mu_i)A_{ij_0} = 0$

となる．そこで

$$
\begin{aligned}
\sum_{j\in I}\left(\sum_{i\in I}(\nu_i-\mu_i)A_{ij}\right)^+ &= \sum_{j\neq j_0}\left(\sum_{i\in I}(\nu_i-\mu_i)A_{ij}\right)^+ \\
&\le \sum_{j\neq j_0}\sum_{i\in I}(\nu_i-\mu_i)^+A_{ij} \\
&= \sum_{i\in I}(\nu_i-\mu_i)^+\sum_{j\neq j_0}A_{ij} \\
&= \sum_{i\in I}(\nu_i-\mu_i)^+(1-A_{ij_0}) \\
&\le \sum_{i\in I}(\nu_i-\mu_i)^+(1-\gamma) \\
&= (1-\gamma)\left\|\nu-\mu\right\|
\end{aligned}
$$

として変形できるから，示された． □

以上の準備の下で，次の定理が証明できる．

Theorem 1.14. 確率行列 A がエルゴード的であるとする．このとき，A に対する不変分布 π がただ一つ存在して，次が成立する．

(1) 任意の $i,j\in I$ に対して，$\lim_{n\to\infty}A^n_{ij}=\pi_j$ が成立する．

(2) ある $C>0$ とある $0<\lambda<1$ が存在して，任意の $n\in\mathbb{N}$ と任意の $i,j\in I$ に対して $\left|A^n_{ij}-\pi_j\right|\le C\lambda^n$ が成立する．

Proof. 不変分布の一意存在から証明する．A はエルゴード的だから，ある $n_0\in\mathbb{N}$ が存在して任意の $i,j\in I$ に対して $A^{n_0}_{ij}>0$ となる．そこで $\gamma\coloneqq\min_{i,j\in I}A^{n_0}_{ij}$ とおくと $0<\gamma<1$ であり，$A^{n_0}_{ij}\ge\gamma$ となる．$\nu\in\mathcal{P}(I)$ が任意に与えられたとして，$\nu_n\coloneqq\nu A^n$ と定める．$(\nu_n)_{n=0}^\infty$ が Cauchy 列であることを証明する．$\varepsilon>0$ が任意に与えられたとする．$(1-\gamma)^{m_0}<\varepsilon$ となるように $m_0\in\mathbb{N}$ をとる．$N\coloneqq n_0m_0$ とする．Theorem1.12 と Theorem1.13 から，$n,m\ge N$ に対して，$\left\|\nu_n-\nu_m\right\|=\left\|\nu A^n-\nu A^m\right\|=\left\|(\nu A^{n-N}-\nu A^{m-N})A^N\right\|\le(1-\gamma)^{m_0}\left\|\nu A^{n-N}-\nu A^{m-N}\right\|<\varepsilon$ となり，示された．$\mathcal{P}(I)$ の完備性から極限 π が存在する．$\pi A=\lim_{n\to\infty}\nu A^{n+1}=\pi$ より π は不変分布．$\tilde{\pi}$ も不変分布とすると，$\left\|\pi-\tilde{\pi}\right\|=\left\|(\pi-\tilde{\pi})A^{n_0}\right\|\le(1-\gamma)\left\|\pi-\tilde{\pi}\right\|$ となるから，$\left\|\pi-\tilde{\pi}\right\|=0$ となり，不変分布は一意である．$i\in I$ に対して $\nu=\delta_i$ とすると $\lim_{n\to\infty}(\nu A^n)_j=\lim_{n\to\infty}A^n_{ij}=\pi_j$ となり，(1) が示された．(2) を証明する．$i,j\in I$ と $n\in\mathbb{N}$ が任意に与えられたとする．ノルムの定義から $\left|A^n_{ij}-\pi_j\right|\le2\left\|\delta_iA^n-\pi\right\|$ が成立する．任意の $k\in\mathbb{N}$ と n を n_0 で割った商 $m\in\mathbb{N}$ に対して，$\left\|\delta_iA^n-\delta_iA^{n+k}\right\|\le(1-\gamma)^m$ が成立するから，$k\to\infty$ として $\left\|\delta_iA^n-\pi\right\|\le(1-\gamma)^m$ を得る．$m\ge(n-n_0)/n_0$ だから，$\left\|\delta_iA^n-\pi\right\|\le(1-\gamma)^{(n-n_0)/n_0}$ となり，$\lambda\coloneqq(1-\gamma)^{n_0}$ として $\left|A^n_{ij}-\pi_j\right|\le2\left\|\delta_iA^n-\pi\right\|\le2\lambda^{n-n_0}$ となる．$C=2/\lambda^{n_0}$ として (2) が示された． □

以上でエルゴード的な遷移行列をもつマルコフ連鎖は初期分布によらない一定の分布に収束すること，その収束は $\mathcal{P}(I)$ のノルムで指数的な速さであることが分かった．この応用として大数の弱法則が成立することを証明する．

Definition 1.11. $\nu\in\mathcal{P}(I)$ と $f\colon I\to\mathbb{R}$ に対して，$E^\nu[f]\coloneqq\sum_{i\in I}\nu_if(i)$ と定義する．これは ν の定める I 上の確率測度による f の積分と一致する．

Theorem 1.15. 確率行列 A はエルゴード的とし，π を A に対する不変分布とする．確率ベクトル ν が任意に与えられたとする．任意のマルコフ連鎖 (X, A, ν, P) と任意の $f\colon I \to \mathbb{R}$ に対して，$\sum_{k=1}^{n} f(X_k)/n$ は $E^{\pi}[f]$ に確率収束する．

Proof. (X, A, ν, P) をマルコフ連鎖として，確率空間を $(\Omega, \mathcal{F}, P)$ と書くことにする．$f - E^{\pi}[f]$ を考えることで $E^{\pi}[f] = 0$ としてよい．$\varepsilon > 0$ が任意に与えられたとして，$P\left(\left|\sum_{k=0}^{n} f(X_k)/n\right| > \varepsilon\right)$ が 0 に収束することを証明すればよい．

$$
\begin{aligned}
P\left(\left|\sum_{k=1}^{n} f(X_k)/n\right| > \varepsilon\right) &= P\left(\left|\sum_{k=1}^{n} f(X_k)\right| > n\varepsilon\right) \\
&\leq \frac{1}{n^2\varepsilon^2} E\left[\left(\sum_{k=1}^{n} f(X_k)\right)^2\right] \quad (\because \text{Chebyshev の不等式}) \\
&= \frac{1}{n^2\varepsilon^2} \sum_{k=1}^{n} E\left[f(X_k)^2\right] + \frac{2}{n^2\varepsilon^2} \sum_{1 \leq k_1 < k_2 \leq n} E\left[f(X_{k_1}) f(X_{k_2})\right]
\end{aligned}
$$

となる．第一項は

$$
\frac{1}{n^2\varepsilon^2} \sum_{k=1}^{n} E\left[f(X_k)^2\right] \leq \frac{1}{n^2\varepsilon^2} n \left\|f\right\|_{\infty}^2 = \frac{\left\|f\right\|_{\infty}^2}{n\varepsilon}
$$

と評価でき，$n \to \infty$ で 0 に収束する．第二項を評価する．$k_1 < k_2$ に対して

$$
\begin{aligned}
E\left[f(X_{k_1}) f(X_{k_2})\right] &= \sum_{i,j \in I} P(X_{k_1} = i) P(X_{k_2} = j \mid X_{k_1} = i) f(i) f(j) \\
&= \sum_{i_0, i, j \in I} \nu_{i_0} A_{i_0 i}^{k_1} A_{ij}^{k_2 - k_1} f(i) f(j) \quad (\because \text{Theorem1.4}) \\
&= \sum_{i_0, i, j \in I} \nu_{i_0} A_{i_0 i}^{k_1} \left(A_{ij}^{k_2 - k_1} - \pi_j\right) f(i) f(j) \quad (\because E^{\pi}[f] = 0)
\end{aligned}
$$

が成立する．ここで Thoerem1.14 よりある $C > 0$ と $0 < \lambda < 1$ が存在して

$$
\left|A_{ij}^{k_2 - k_1} - \pi_j\right| \leq C \lambda^{k_2 - k_1}
$$

となるから，

$$
\begin{aligned}
\left|E\left[f(X_{k_1}) f(X_{k_2})\right]\right| &\leq \sum_{i_0, i, j \in I} \nu_{i_0} A_{i_0 i}^{k_1} \left|A_{ij}^{k_2 - k_1} - \pi_j\right| \left|f(i) f(j)\right| \\
&\leq \left\|f\right\|_{\infty}^2 \left|I\right| C \lambda^{k_2 - k_1}
\end{aligned}
$$

を得る．したがって，第二項の絶対値は

$$
\begin{aligned}
\left|\frac{2}{n^2\varepsilon^2} \sum_{1 \leq k_1 < k_2 \leq n} E\left[f(X_{k_1}) f(X_{k_2})\right]\right| &\leq \frac{2}{n^2\varepsilon^2} \sum_{1 \leq k_1 < k_2 \leq n} \left|E\left[f(X_{k_1}) f(X_{k_2})\right]\right| \\
&\leq \frac{2}{n^2\varepsilon^2} \sum_{1 \leq k_1 < k_2 \leq n} \left\|f\right\|_{\infty}^2 \left|I\right| C \lambda^{k_2 - k_1}
\end{aligned}
$$

となる．ここで

$$\begin{aligned}\sum_{1\le k_1<k_2\le n}\lambda^{k_2-k_1} &= \sum_{k_1=1}^{n-1}\lambda^{-k_1}\sum_{k_2=k_1+1}^{n}\lambda^{k_2}\\ &= \sum_{k_1=1}^{n-1}\lambda^{-k_1}\frac{\lambda^{k_1+1}(1-\lambda^{n-k_1})}{1-\lambda}\\ &\le \frac{n\lambda}{1-\lambda}\end{aligned}$$

だから，

$$\begin{aligned}\frac{2}{n^2\varepsilon^2}\sum_{1\le k_1<k_2\le n}\|f\|_\infty^2\,|I|\,C\lambda^{k_2-k_1} &\le \frac{2}{n^2\varepsilon^2}\|f\|_\infty^2\,|I|\,C\frac{n\lambda}{1-\lambda}\\ &= \frac{2\|f\|_\infty^2\,|I|\,C\lambda}{n(1-\lambda)}\end{aligned}$$

となり，第二項も $n\to\infty$ で 0 に収束する．以上により，示された． □

弱法則の証明では非対角部分の相関が問題になるが，k_2-k_1（対角から離れたところ）が指数的に減衰するので評価が上手くいった．このように時間が離れたところの相関が小さくなることを，過去の情報を忘れると表現したりする．

1.4 あとがき

以上のノートの内容は「確率論」（舟木直久，朝倉出版）と「Essentials of Stochastic Processes」（Richard Durrett，Springer）を参考にした．マルチンゲール問題や強マルコフ性，ランダムウォークは扱わなかったので，それらは上記の参考書で補ってほしい．SDE の解の強マルコフ性がよくわからなかったので簡単な場合で様子をつかもうとして書き始めたのだが，強マルコフ性を使わずに到達時刻や差分方程式までやれてしまった…