8

Ergodic Theorem

We have seen in Corollary 7.5.3 that a recurrent and irreducible Markov chain has ergodicity, in this chapter, we will going to discuss the ergodic theory in a more general setting. This can be understood as the strong law of large numbers discussed in Theorem 4.3.1 where we extend the result to sequences of non i.i.d. random variables.

8.1 Measure-Preserving Transformations and Properties

Definition 8.1.1 : A random process $(X_n)_{n\geqslant 0}$ is a *stationary sequence* (平穩序列) or a *stationary process* (平穩過程) if for all non-negative integer $k \geqslant 0$, the random processes $(X_n)_{n\geqslant 0}$ and $(X_{n+k})_{n\geqslant 0}$ have the same distribution.

Example 8.1.2 : Below are a few examples of stationary processes.

(1) Any i.i.d. sequence $(X_n)_{n\geqslant 0}$ of random variables is a stationary process.

(2) Let $(X_n)_{n\geqslant 0}$ be a Markov chain with transition matrix Q. If it possesses a stationary probability measure π, then when $X_0 \sim \pi$, the random process $(X_n)_{n\geqslant 0}$ is stationary, see Remark 7.4.3.

Then, we introduce the notion of *measure-preserving transformations*, generalizing the notion of stationary processes.

Definition 8.1.3 : Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A measurable function $\varphi : \Omega \longrightarrow \Omega$ is said to be a *measure-preserving transformation* (測度守恆變換) on $(\Omega, \mathcal{F}, \mathbb{P})$ if $\mathbb{P}(\varphi^{-1}(A)) = \mathbb{P}(A)$ for all $A \in \mathcal{F}$. In this case, we also say that $\mathbb{P}$ is an *invariant measure* (不變測度) for φ.

More generally speaking, we do not need to assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space in Definition 8.1.3, it is enough to have a σ-finite measured space. Since our class is about probability theory, we will only focus on the case of probability spaces in what follows.

Example 8.1.4 : Below are a few examples of measure-preserving measures.

(1) Consider the measured space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda^d)$, then for any $y \in \mathbb{R}$, the translation function

$$\theta_y : \quad x \quad \mapsto \quad x + y$$

is a measure-preserving transformation.

遍歷定理

第八章

我們在系理 7.5.3 中看到，一個重現且不可約的馬可夫鏈，有遍歷性的性質；在這個章節中，我們將要探討更一般的遍歷定理。這也可以被理解做我們在定理 4.3.1 中看過得強大數法則，在非 i.i.d. 隨機變數序列情況下的推廣。

第一節 測度守恆變換及性質

定義 8.1.1 ：給定隨機過程 $(X_n)_{n\geqslant 0}$。若對於所有非負整數 $k \geqslant 0$，隨機過程 $(X_n)_{n\geqslant 0}$ 及 $(X_{n+k})_{n\geqslant 0}$ 有相同的分佈，則我們說 $(X_n)_{n\geqslant 0}$ 是個平穩序列 (stationary sequence) 或是平穩過程 (stationary process) 。

範例 8.1.2 ：下列是幾個平穩過程的例子。

(1) 任意 i.i.d. 隨機變數序列 $(X_n)_{n\geqslant 0}$ 皆是平穩過程。

(2) 令 $(X_n)_{n\geqslant 0}$ 是個轉移矩陣為 Q 的馬可夫鏈，若存在平穩機率測度 π，則當 $X_0 \sim \pi$ 時，$(X_n)_{n\geqslant 0}$ 是個平穩過程，參見註解 7.4.3 。

接著我們定義測度守恆變換的概念，將平穩過程的概念推廣至更一般的情況。

定義 8.1.3 ：給定 σ 機率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 以及可測函數 $\varphi : \Omega \longrightarrow \Omega$。如果對於所有 $A \in \mathcal{F}$，我們有 $\mathbb{P}(\varphi^{-1}(A)) = \mathbb{P}(A)$，則我們說 φ 是個在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的測度守恆變換 (measure-preserving transformation) 。在此情況下，我們也說 $\mathbb{P}$ 是 φ 的不變測度 (invariant measure) 。

更一般來說，在定義 8.1.3 中，我們不需要 $(\Omega, \mathcal{F}, \mathbb{P})$ 是個機率空間的假設，σ 有限的測度空間即足夠；但由於這是機率論課程，後續我們只會討論機率空間的情況。

範例 8.1.4 ：以下是測度守恆變換的例子。

(1) 考慮測度空間 $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda^d)$，則對於任意 $y \in \mathbb{R}$，平移函數

$$\theta_y : \quad x \quad \mapsto \quad x + y$$

(2) Consider the unit circle $\mathbb{S}^1 = \{x \in \mathbb{C} : |x| = 1\}$ and the uniform probability distribution μ defined above. For any $\beta \in \mathbb{R}$, we can define the rotation operator

$$\theta_\beta : e^{2\pi\,\mathrm{i}\,\alpha} \mapsto e^{2\pi\,\mathrm{i}(\alpha+\beta)}.$$

which is a measure-preserving transformation on $(\mathbb{S}^1, \mathcal{B}(\mathbb{S}^1), \mu)$.

(3) Consider the probability space $([0,1], \mathcal{B}([0,1]), \mathbb{P})$, where $\mathbb{P}$ is the Lebesgue measure on $[0,1]$. Then,

$$\theta(x) := \begin{cases} 2x & \text{if } x \in [0, \frac{1}{2}), \\ 2-2x & \text{if } x \in [\frac{1}{2}, 1], \end{cases}$$

is a measure-preserving transformation.

The below proposition shows that, the notion of a stationary sequence defined previously is indeed a special case of invariant measure with respect to a properly chosen measure-preserving transformation.

Proposition 8.1.5 : *Let $E = \mathbb{R}$ or $\mathbb{R}^d$ be the state space, $\Omega = E^{\mathbb{N}_0}$ be the sample space, $\mathcal{F} = \mathcal{B}(E)^{\otimes \mathbb{N}_0}$ be the smallest σ-algebra making all the coordinate functions measurable, $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$. If $\omega \sim \mathbb{P}$ is a stationary process, then the shift operator* (推移算子) *$\theta := \theta_1$ defined as $\theta((\omega_n)_{n\geqslant 0}) = (\omega_{n+1})_{n\geqslant 0}$ is a measure-preserving transformation on $(\Omega, \mathcal{F}, \mathbb{P})$.*

Proof : First, let us check that θ is measurable. For any cylinder event A in $\mathcal{F}$, we have $\theta^{-1}(A) = E \times A$ which is also a cylinder event. At the same time, we know that the cylinder events generate $\mathcal{F}$, so θ is measurable. Then, due to the stationarity of $\mathbb{P}$, we have $\mathbb{P}(A) = \mathbb{P}(\theta^{-1}(A))$ for all cylinder event A, so the measures $\mathbb{P}$ and $\mathbb{P} \circ \theta^{-1}$ take the same value on the collection of cylinder events which generates $\mathcal{F}$. We are done with the proof. □

Proposition 8.1.6 : *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. If φ is a measure-preserving transformation on $(\Omega, \mathcal{F}, \mathbb{P})$, then for any random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (G, \mathcal{G})$ taking values in a measurable space $(G, \mathcal{G})$, we can define $X_n(\omega) = X(\varphi^n(\omega))$ for any non-negative integer $n \geqslant 0$, and we have that $(X_n)_{n\geqslant 0}$ is a stationary process.*

Remark 8.1.7 : In Proposition 8.1.6, if we choose $(\Omega, \mathcal{F}) = (E^{\mathbb{N}_0}, \mathcal{B}(E)^{\otimes \mathbb{N}_0})$ and $(G, \mathcal{G}) = (E, \mathcal{B}(E))$ and assume that the shift operator θ is measure-preserving, then by taking $X = X_0$, we get the converse of Proposition 8.1.5. This tells us that a stationary sequence is indeed a special case of invariant measure with respect to a properly chosen measure-preserving transformation.

是個測度守恆變換。

(2) 考慮單位圓 $\mathbb{S}^1 = \{x \in \mathbb{C} : |x| = 1\}$ 以及定義在上面的均勻分佈 μ。對於任意 $\beta \in \mathbb{R}$，我們可以定義旋轉算子

$$\theta_\beta : e^{2\pi\,\mathrm{i}\,\alpha} \mapsto e^{2\pi\,\mathrm{i}(\alpha+\beta)}.$$

這會是個在 $(\mathbb{S}^1, \mathcal{B}(\mathbb{S}^1), \mu)$ 上的測度守恆變換。

(3) 考慮機率空間 $([0,1], \mathcal{B}([0,1]), \mathbb{P})$，其中 $\mathbb{P}$ 為在 $[0,1]$ 上的勒貝格測度，則

$$\theta(x) := \begin{cases} 2x & \text{若 } x \in [0, \frac{1}{2}), \\ 2-2x & \text{若 } x \in [\frac{1}{2}, 1], \end{cases}$$

是個測度守恆變換。

下面的命題告訴我們，稍早定義的平穩序列的確只是由某特定的測度守恆轉換所對應到的不變測度的特例而已。

命題 8.1.5 ： 令 $E = \mathbb{R}$ 或 $\mathbb{R}^d$ 為狀態空間，$\Omega = E^{\mathbb{N}_0}$ 為樣本空間，$\mathcal{F} = \mathcal{B}(E)^{\otimes \mathbb{N}_0}$ 為使所有座標函數可測的最小 σ 代數，$\mathbb{P}$ 為在 $(\Omega, \mathcal{F})$ 上的機率測度。若 $\omega \sim \mathbb{P}$ 是個平穩過程，則推移算子 (shift operator) $\theta := \theta_1$，定義做 $\theta((\omega_n)_{n\geqslant 0}) = (\omega_{n+1})_{n\geqslant 0}$，是個在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的測度守恆變換。

證明： 首先，我們來檢查 θ 是個可測函數。對於任意在 $\mathcal{F}$ 中的圓柱事件 A，我們有 $\theta^{-1}(A) = E \times A$，這也會是個圓柱事件。同時，我們知道圓柱事件可以生成 $\mathcal{F}$，因此 θ 為可測函數。接著，由於 $\mathbb{P}$ 定義的是個平穩過程，對於所有圓柱事件 A，我們會有 $\mathbb{P}(A) = \mathbb{P}(\theta^{-1}(A))$，因此測度 $\mathbb{P}$ 及 $\mathbb{P} \circ \theta^{-1}$ 在圓柱事件構成的集合上取值相同，且圓柱事件生成 $\mathcal{F}$，得證。 □

命題 8.1.6 ： 令 $(\Omega, \mathcal{F}, \mathbb{P})$ 為機率空間。若 φ 是個在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的測度守恆變換，則對於任意取值在可測空間 $(G, \mathcal{G})$ 中的隨機變數 $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (G, \mathcal{G})$，我們可以對所有非負整數 $n \geqslant 0$ 定義 $X_n(\omega) = X(\varphi^n(\omega))$，則 $(X_n)_{n\geqslant 0}$ 是個平穩過程。

註解 8.1.7 ： 在命題 8.1.6 中，若取 $(\Omega, \mathcal{F}) = (E^{\mathbb{N}_0}, \mathcal{B}(E)^{\otimes \mathbb{N}_0})$ 及 $(G, \mathcal{G}) = (E, \mathcal{B}(E))$，且假設推移算子 θ 是個測度守恆變換，當我們取 $X = X_0$ 的座標函數時，我們得到命題 8.1.5 的逆命題。由此我們可以得知，平穩序列的確只是在特定測度守恆轉換下的不變測度而已。

Proof : Given $n \geqslant 0$ and let $B \in \mathcal{G}^{\otimes(n+1)}$ and $A = \{\omega \in \Omega : (X_0(\omega), \dots, X_n(\omega)) \in B\}$. Then,

$$\begin{aligned}\mathbb{P}((X_k(\omega), \dots, X_{k+n}(\omega)) \in B) &= \mathbb{P}(\omega : \varphi^k(\omega) \in A)\\ &= \mathbb{P}(\omega \in A)\\ &= \mathbb{P}((X_0(\omega), \dots, X_n(\omega)) \in B).\end{aligned}$$

$\square$

Definition 8.1.8 : Let φ be a measure-preserving transformation on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The *invariant set* (不變集合) of φ is defined as

$$\mathcal{I} = \mathcal{I}_\varphi := \{A \in \mathcal{F} : \varphi^{-1}(A) = A\}. \tag{8.1}$$

We can check that $\mathcal{I}_\varphi$ is a σ-algebra, so it is also called *invariant σ-algebra* (不變 σ 代數) of φ. Additionally, if $\mathcal{I}_\varphi$ is a trivial σ-algebra, we say that φ is *ergodic* (遍歷性) .

Question 8.1.9: Check that the invariant set $\mathcal{I}$ defined in Definition 8.1.8 is indeed a σ-algebra.

Remark 8.1.10 : If φ is not ergodic, then we can find $A \in \mathcal{I}$ such that $\varphi(A) = A$, $\varphi(A^c) = A^c$ and $\mathbb{P}(A) > 0$, $\mathbb{P}(A^c) > 0$. In other words, φ is reducible.

Question 8.1.11: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which we consider a measure-preserving transformation φ.

(1) A random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is $\mathcal{I}$-measurable if and only if $X \circ \varphi = X$.

(2) φ is ergodic if and only if all the $\mathcal{I}$-measurable random variables $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ are $\mathbb{P}$-a.s. constant.

For a given sequence $(X_n)_{n\geqslant 0}$ of random variables, we recall the *asymptotic σ-algebra* (漸進 σ 代數) defined in Definition 4.2.1, which we also call *tail σ-algebra* (尾端 σ 代數) , detnoed $\mathcal{T}$,

$$\mathcal{T} = \bigcap_{n=0}^{\infty} \mathcal{F}^n, \qquad \mathcal{F}^n = \sigma(X_n, X_{n+1}, \dots), \quad \forall n \geqslant 0.$$

Proposition 8.1.12 : *Let $E = \mathbb{R}$ or $\mathbb{R}^d$ be the state space and $\Omega = E^{\mathbb{N}_0}$ be the sample space. Take $\mathcal{F} = \mathcal{B}(E)^{\otimes\mathbb{N}_0}$ and assume that the shift operator θ is measure-preserving. Then, $\mathcal{I} \subseteq \mathcal{T}$.*

Remark 8.1.13 : Consider the case where $(X_n)_{n\geqslant 0}$ is an i.i.d. sequence. By Kolmogorov's 0-1 law (Theorem 4.2.2), we know that $\mathcal{T}$ is a trivial σ-algebra, so $\mathcal{I}$ is also trivial. In other words, for any i.i.d. sequence of random variables, the shift operator is ergodic.

證明：給定 $n \geqslant 0$ 並令 $B \in \mathcal{G}^{\otimes(n+1)}$ 及 $A = \{\omega \in \Omega : (X_0(\omega), \dots, X_n(\omega)) \in B\}$，則

$$\begin{aligned}\mathbb{P}((X_k(\omega), \dots, X_{k+n}(\omega)) \in B) &= \mathbb{P}(\omega : \varphi^k(\omega) \in A)\\ &= \mathbb{P}(\omega \in A)\\ &= \mathbb{P}((X_0(\omega), \dots, X_n(\omega)) \in B).\end{aligned}$$

$\square$

定義 8.1.8 ：令 φ 為在機率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的測度守恆變換，我們定義 φ 的不變集合 (invariant set) 為

$$\mathcal{I} = \mathcal{I}_\varphi := \{A \in \mathcal{F} : \varphi^{-1}(A) = A\}. \tag{8.1}$$

我們可以驗證 $\mathcal{I}_\varphi$ 是個 σ 代數，因此也可以稱他為 φ 的不變 σ 代數 (invariant σ-algebra) 。此外，若 $\mathcal{I}_\varphi$ 為平凡 σ 代數，則我們說 φ 有遍歷性 (ergodic) 。

問題 8.1.9：驗證定義 8.1.8 中定義的不變集合 $\mathcal{I}$ 的確是個 σ 代數。

註解 8.1.10 ：若 φ 不具有遍歷性，則我們可以找到 $A \in \mathcal{I}$ 使得 $\varphi(A) = A$，$\varphi(A^c) = A^c$ 及 $\mathbb{P}(A) > 0$，$\mathbb{P}(A^c) > 0$；換句話說，φ 是可約的。

問題 8.1.11：令 $(\Omega, \mathcal{F}, \mathbb{P})$ 為機率空間，φ 是個在他上面的測度守恆函數。

(1) 給定隨機變數 $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$。若且唯若 $X \circ \varphi = X$，則 X 會對 $\mathcal{I}$ 可測。

(2) 若且唯若所有 $\mathcal{I}$ 可測的隨機變數 $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 皆 $\mathbb{P}$-a.s. 為常數，則 φ 有遍歷性。

給定隨機變數序列 $(X_n)_{n\geqslant 0}$，我們回顧在定義 4.2.1 中定義過漸進 σ 代數 (asymptotic σ-algebra) ，這裡我們也稱作尾端 σ 代數 (tail σ-algebra) ，記作 $\mathcal{T}$：

$$\mathcal{T} = \bigcap_{n=0}^{\infty} \mathcal{F}^n, \qquad \mathcal{F}^n = \sigma(X_n, X_{n+1}, \dots), \quad \forall n \geqslant 0.$$

命題 8.1.12 ：令 $E = \mathbb{R}$ 或 $\mathbb{R}^d$ 為狀態空間，考慮樣本空間 $\Omega = E^{\mathbb{N}_0}$ 及 $\mathcal{F} = \mathcal{B}(E)^{\otimes\mathbb{N}_0}$，並假設推移運算子 θ 是個測度守恆變換，則 $\mathcal{I} \subseteq \mathcal{T}$。

註解 8.1.13 ：考慮 $(X_n)_{n\geqslant 0}$ 為 i.i.d. 序列的狀況，從 Kolmogorov 零一律（定理 4.2.2 ），我們得知 $\mathcal{T}$ 為平凡 σ 代數，因此 $\mathcal{I}$ 也是個平凡 σ 代數，也就是說，任意 i.i.d. 隨機變數序列的推移算子皆具有遍歷性。

Proof : Given an invariant set $A \in \mathcal{I}$, we have

$$A = \theta^{-1}(A) = \{\omega : \theta(\omega) \in A\} \in \mathcal{F}^1.$$

If we keep iterating, we obtain $A \in \mathcal{F}^n$ for all $n \geqslant 0$, that is $A \in \mathcal{T}$. □

Question 8.1.14: Please find an example with $\mathcal{I} \subsetneq \mathcal{T}$.

Definition 8.1.15 : We keep the notations in Proposition 8.1.5. If $\omega \sim \mathbb{P}$ is a stationary process and the shift operator θ is ergodic, we say that $\omega = (\omega_n)_{n \geqslant 0}$ is an *ergodic process* (遍歷過程) .

Remark 8.1.16 : We note that, according to Definition 8.1.15, for a random process to be called ergodic, it needs first to be stationary. If we come back to the setting of Markov chains, Remark 7.4.3 says that if a Markov chain evolves from the initial state given by its stationary probability measure (assuming existence), then we do have a stationary process. However, for a recurrent and irreducible Markov chain, even without the assumption of stationarity, for any given initial state $\mathbb{P}$ and bounded (or non-negative) function $f : E \longrightarrow \mathbb{R}$, the ergodic theorem (Corollary 7.5.3)

$$\frac{1}{n}\sum_{k=0}^{n} f(X_k) \longrightarrow \mu(f), \qquad \mathbb{P}\text{-a.s.}$$

holds and sometimes we also say such a process is *ergodic*.

Now, let us discuss the ergodicity of Markov chains in the sense of Definition 8.1.15.

Proposition 8.1.17 : *Given a Markov chain $(X_n)_{n \geqslant 0}$ defined on a countable state space E and assume that it has an invariant probability measure π such that $\pi(x) > 0$ for all $x \in E$. Then, the Markov chain $(X_n)_{n \geqslant 0}$ is irreducible if and only if it is ergodic.*

Proof : First we note that, by Example 8.1.2 (2) and Proposition 8.1.5, we know that the shift operator θ is a measure-preserving transformation for $\mathbb{P}_\pi$.

We can show that all the states are recurrent by adapting the proof of Proposition 7.4.15. Moreover, Theorem 7.3.6 says that $E = \sqcup R_i$ where R_i's are disjoint irreducible sets. Besides, we also know that if $X_0 \in R_i$, then $X_n \in R_i$ for all $n \geqslant 1$, implying

$$\{\omega : X_0(\omega) \in R_i\} \in \mathcal{I}, \qquad \forall i.$$

In consequence, this tells us that if a Markov chain is not irreducible, then the shift operator θ is not ergodic.

Conversely, consider $A \in \mathcal{I}_\theta$, then for any iteration of the shift operator $\theta_n = \theta^n$, we have $\mathbb{1}_A \circ \theta_n = \mathbb{1}_A$. Hence, if we define $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ for $n \geqslant 0$, the simple Markov property gives

$$\mathbb{E}_\pi[\mathbb{1}_A \,|\, \mathcal{F}_n] = \mathbb{E}_\pi[\mathbb{1}_A \circ \theta_n \,|\, \mathcal{F}_n] = \mathbb{E}_{X_n}[\mathbb{1}_A] = h(X_n),$$

證明：給定不變集合 $A \in \mathcal{I}$，則我們有

$$A = \theta^{-1}(A) = \{\omega : \theta(\omega) \in A\} \in \mathcal{F}^1.$$

若繼續迭代下去，我們可以得到，對於所有 $n \geqslant 0$，$A \in \mathcal{F}^n$，也就是說 $A \in \mathcal{T}$。 □

問題 8.1.14：請找出滿足 $\mathcal{I} \subsetneq \mathcal{T}$ 的例子。

定義 8.1.15 ： 我們沿用命題 8.1.5 中的記號。若 $\omega \sim \mathbb{P}$ 是個平穩過程且推移算子 θ 具有遍歷性，我們將 $\omega = (\omega_n)_{n \geqslant 0}$ 稱為遍歷過程 (ergodic process) 。

註解 8.1.16 ： 我們注意到，根據定義 8.1.15 ，一個隨機過程要被稱作遍歷過程，首先必須要先是個平穩過程。回到馬可夫鏈的框架，根據註解 7.4.3 ，如果馬可夫鏈的初始狀態由平穩機率測度所給定（假設存在性），則我們的確有個平穩過程；但對於一個重現且不可約的馬可夫鏈，即使我們沒有平穩性質，若對於任意初始狀態 $\mathbb{P}$，有界函數（或非負函數） $f : E \longrightarrow \mathbb{R}$，遍歷定理（系理 7.5.3 ）

$$\frac{1}{n}\sum_{k=0}^{n} f(X_k) \longrightarrow \mu(f), \qquad \mathbb{P}\text{-a.s.}$$

成立，我們偶爾仍會說這樣的隨機過程是遍歷過程。

現在，我們用定義 8.1.15 來探討馬可夫鏈的遍歷性。

命題 8.1.17 ： 給定一個在可數狀態集合 E 上的馬可夫鏈 $(X_n)_{n \geqslant 0}$，且存在不變機率測度 π，使得對於所有 $x \in E$，我們有 $\pi(x) > 0$。那麼若且唯若馬可夫鏈 $(X_n)_{n \geqslant 0}$ 具有遍歷性，則他是不可約的。

證明：首先注意到，藉由範例 8.1.2 (2) 及命題 8.1.5 我們知道，推移算子 θ 對 $\mathbb{P}_\pi$ 來說，是個測度守恆變換。

修改命題 7.4.15 的證明，我們可以得知所有的狀態皆是重現狀態，且定理 7.3.6 告訴我們，我們有 $E = \sqcup R_i$，其中 R_i 為互斥不可約的集合。此外，我們也知道若 $X_0 \in R_i$，則對於所有 $n \geqslant 1$，我們也會有 $X_n \in R_i$，所以說

$$\{\omega : X_0(\omega) \in R_i\} \in \mathcal{I}, \qquad \forall i.$$

因此，這告訴我們，若馬可夫鏈不可約，則推移算子 θ 不具有遍歷性。

反之，考慮 $A \in \mathcal{I}_\theta$，則對任意推移算子的迭代 $\theta_n = \theta^n$，我們會有 $\mathbb{1}_A \circ \theta_n = \mathbb{1}_A$。因此，若

where $h(x) = \mathbb{E}_x[\mathbb{1}_A]$. Since $A \in \mathcal{I}$ and the left side of the above formula is an uniformly integrable martingale, Theorem 6.5.6 implies that it converges almost surely to $\mathbb{E}_\pi[\mathbb{1}_A \mid \mathcal{F}_\infty] = \mathbb{1}_A$. Besides, since $(X_n)_{n \geqslant 0}$ is irreducible and recurrent, for any $y \in E$, the right side of the above formula can take the value $h(y)$ infinitely many times, so $h \equiv 0$ or $h \equiv 1$, that is $\mathbb{P}_\pi(A) = 0$ or 1. □

Below is a useful criterion to check whether a measure-preserving transformation is ergodic.

Definition 8.1.18 : Let φ be a measure-preserving transformation on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that φ is *mixing* (混合性) if for all $F, G \in \mathcal{F}$,

$$\mathbb{P}(F \cap \varphi^{-n}(G)) \xrightarrow[n \to \infty]{} \mathbb{P}(F)\ \mathbb{P}(G),$$

Lemma 8.1.19 : *Let φ be a measure-preserving transformation on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If φ is mixing, then φ is also ergodic.*

Proof : Let $F \in \mathcal{I}$. Then, we have

$$\mathbb{P}(F) = \mathbb{P}(F \cap \varphi^{-n}(F)) \longrightarrow \mathbb{P}(F)^2.$$

This gives $\mathbb{P}(F) = \mathbb{P}(F)^2$, that is $\mathbb{P}(F) \in \{0, 1\}$. Hence, φ is ergodic. □

Exercise 8.6 is an example where we use Lemma 8.1.19 to check that a measure-preserving transformation is ergodic. Exercise 8.14 also gives another useful lemma to check the ergodicity.

8.2 Birkhoff's Ergodic Theorem

In this section, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we consider a measure-preserving transformation φ. Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (G, \mathcal{G})$ be a random variable and define the random process $(X_n)_{n \geqslant 0}$ with $X_n(\omega) = X(\varphi^n(\omega))$.

Theorem 8.2.1 (Birkhoff's ergodic theorem) **:** *Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ be an integrable random variable and φ be a measure-preserving transformation on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the following convergence holds $\mathbb{P}$-a.s. and in L^1,*

$$\frac{1}{n} \sum_{k=0}^{n-1} X_k = \frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k \xrightarrow[n \to \infty]{} \mathbb{E}[X \mid \mathcal{I}].$$

對於所有 $n \geqslant 0$，定義 $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$，簡單馬可夫性質給出：

$$\mathbb{E}_\pi[\mathbb{1}_A \mid \mathcal{F}_n] = \mathbb{E}_\pi[\mathbb{1}_A \circ \theta_n \mid \mathcal{F}_n] = \mathbb{E}_{X_n}[\mathbb{1}_A] = h(X_n),$$

其中 $h(x) = \mathbb{E}_x[\mathbb{1}_A]$。由於 $A \in \mathcal{I}$ 且上式左方為均勻可積的鞅，根據定理 6.5.6 ，我們知道他會 a.s. 收斂至 $\mathbb{E}_\pi[\mathbb{1}_A \mid \mathcal{F}_\infty] = \mathbb{1}_A$。此外，由於 $(X_n)_{n \geqslant 0}$ 是不可約且重現的，對於任意 $y \in E$，上式右方會無限次取值 $h(y)$，因此 $h \equiv 0$ 或 $h \equiv 1$，也就是說 $\mathbb{P}_\pi(A) = 0$ 或 1。 □

下面是可以拿來檢查一個測度守恆變換是否具有遍歷性的好用條件。

定義 8.1.18 ： 令 φ 為在機率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的測度守恆變換若對於所有 $F, G \in \mathcal{F}$，我們有

$$\mathbb{P}(F \cap \varphi^{-n}(G)) \xrightarrow[n \to \infty]{} \mathbb{P}(F)\ \mathbb{P}(G).$$

則我們說 φ 有混合性 (mixing) 。

引理 8.1.19 ： 令 φ 為在機率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的測度守恆變換。若 φ 具有混合性，則 φ 也具有遍歷性。

證明： 令 $F \in \mathcal{I}$，則我們有

$$\mathbb{P}(F) = \mathbb{P}(F \cap \varphi^{-n}(F)) \longrightarrow \mathbb{P}(F)^2.$$

所以 $\mathbb{P}(F) = \mathbb{P}(F)^2$，也就是說 $\mathbb{P}(F) \in \{0, 1\}$，因此 φ 對於 φ 有遍歷性。 □

習題 8.6 是個利用引理 8.1.19 來檢查測度守恆變換是否具有遍歷性的例子。習題 8.14 也會給出其他檢查遍歷性的有用引理。

第二節 Birkhoff 遍歷定理

接下來在此章節中，我們固定機率空間 $(\Omega, \mathcal{F}, \mathbb{P})$，取 φ 為在此空間上的測度守恆變換，考慮隨機變數 $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (G, \mathcal{G})$ 並定義隨機過程 $(X_n)_{n \geqslant 0}$，其中 $X_n(\omega) = X(\varphi^n(\omega))$。

定理 8.2.1 【Birkhoff 遍歷定理】 **：** 令 $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ 為可積隨機變數，φ 為在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的測度守恆變換，則下列收斂會 $\mathbb{P}$-a.s. 成立，也會在 L^1 中成立：

$$\frac{1}{n} \sum_{k=0}^{n-1} X_k = \frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k \xrightarrow[n \to \infty]{} \mathbb{E}[X \mid \mathcal{I}].$$

Lemma 8.2.2 (Hopf's maximal ergodic lemma) **:** *If we define*

$$\begin{aligned} S_0 = 0, \quad S_n &= \sum_{k=0}^{n-1} X \circ \varphi^k, &\quad \forall n \geqslant 1, \\ M_n &= \max_{0 \leqslant k \leqslant n} S_k, &\quad \forall n \geqslant 1, \end{aligned}$$

then for all $n \geqslant 1$*, we have*

$$\mathbb{E}[X \mathbb{1}_{M_n > 0}] \geqslant 0.$$

Proof : Given a positive integer $n \geqslant 1$. First we note that

$$S_{k+1} = X + S_k \circ \varphi, \qquad \forall k \geqslant 0.$$

Thus, we have, for $1 \leqslant k \leqslant n$,

$$S_k = X + S_{k-1} \circ \varphi \leqslant X + M_n \circ \varphi.$$

On the event $\{M_n > 0\}$, we have $M_n = \max_{1 \leqslant k \leqslant n} S_k$ (note that the index k starts from 1 instead of 0). We can multiply the above formula by $\mathbb{1}_{M_n > 0}$ and take $\max$ on the left side for $1 \leqslant k \leqslant n$ to obtain

$$M_n \mathbb{1}_{M_n > 0} \leqslant (X + M_n \circ \varphi) \mathbb{1}_{M_n > 0}.$$

Finally, we conclude by writing

$$\begin{aligned} \mathbb{E}[X \mathbb{1}_{M_n > 0}] &\geqslant \mathbb{E}[M_n \mathbb{1}_{M_n > 0}] - \mathbb{E}[(M_n \circ \varphi) \mathbb{1}_{M_n > 0}] \\ &\geqslant \mathbb{E}[M_n] - \mathbb{E}[M_n \circ \varphi] = 0, \end{aligned}$$

where in the inequality in the second line, we use the fact that $M_n \equiv 0$ and $M_n \circ \varphi \geqslant 0$ on $\{M_n > 0\}^c$, and in the last equality, we use the assumption that φ is a measure-preserving transformation. ☐

Now we are ready to prove the ergodic theorem in Theorem 8.2.1.

Proof : We can assume $\mathbb{E}[X \,|\, \mathcal{I}] = 0$ since without loss of generality, we can replace X with $X - \mathbb{E}[X \,|\, \mathcal{I}]$. Now we want to show

$$\frac{1}{n} \sum_{k=0}^{n-1} X_k \xrightarrow[n \to \infty]{} 0. \tag{8.2}$$

Given $\varepsilon > 0$ and let

$$L_\varepsilon = \left\{ \limsup_{n \to \infty} \frac{S_n}{n} > \varepsilon \right\}.$$

To prove Eq. (8.2), it is sufficient to show $\mathbb{P}(L_\varepsilon) = 0$ then replace X with $-X$ by symmetry to conclude, since ε is arbitrarily small.

引理 8.2.2 【Hopf 極大遍歷引理】： 若我們定義

$$\begin{aligned} S_0 = 0, \quad S_n &= \sum_{k=0}^{n-1} X \circ \varphi^k, &\quad \forall n \geqslant 1, \\ M_n &= \max_{0 \leqslant k \leqslant n} S_k, &\quad \forall n \geqslant 1, \end{aligned}$$

則對於所有 $n \geqslant 1$，我們有

$$\mathbb{E}[X \mathbb{1}_{M_n > 0}] \geqslant 0.$$

證明：給定正整數 $n \geqslant 1$。首先我們注意到，

$$S_{k+1} = X + S_k \circ \varphi, \qquad \forall k \geqslant 0.$$

因此所有對於 $1 \leqslant k \leqslant n$，我們有

$$S_k = X + S_{k-1} \circ \varphi \leqslant X + M_n \circ \varphi.$$

在事件 $\{M_n > 0\}$ 上，我們有 $M_n = \max_{1 \leqslant k \leqslant n} S_k$（注意到這裡下標 k 從 1 開始而不是從 0 開始），因此將上式乘上指標函數 $\mathbb{1}_{M_n > 0}$ 並在左側對 $1 \leqslant k \leqslant n$ 取 $\max$，我們得到

$$M_n \mathbb{1}_{M_n > 0} \leqslant (X + M_n \circ \varphi) \mathbb{1}_{M_n > 0}.$$

最後，我們總結

$$\begin{aligned} \mathbb{E}[X \mathbb{1}_{M_n > 0}] &\geqslant \mathbb{E}[M_n \mathbb{1}_{M_n > 0}] - \mathbb{E}[(M_n \circ \varphi) \mathbb{1}_{M_n > 0}] \\ &\geqslant \mathbb{E}[M_n] - \mathbb{E}[M_n \circ \varphi] = 0, \end{aligned}$$

其中在第二行的不等式中，我們用了在 $\{M_n > 0\}^c$ 上，$M_n \equiv 0$ 及 $M_n \circ \varphi \geqslant 0$ 的性質；在最後一個等式，我們使用了 φ 是個測度守恆變換的假設。□

接著我們可以證明定理 8.2.1 中的遍歷定理。

證明：由於我們可以將 X 用 $X - \mathbb{E}[X \,|\, \mathcal{I}]$ 代替，因此不失一般性，我們可以假設 $\mathbb{E}[X \,|\, \mathcal{I}] = 0$ 並且想要證明

$$\frac{1}{n} \sum_{k=0}^{n-1} X_k \xrightarrow[n \to \infty]{} 0. \tag{8.2}$$

給定 $\varepsilon > 0$，令

$$L_\varepsilon = \left\{ \limsup_{n \to \infty} \frac{S_n}{n} > \varepsilon \right\}.$$

若要證明式 (8.2)，我們只需要證明 $\mathbb{P}(L_\varepsilon) = 0$，並且根據對稱性，將 X 用 $-X$ 取代，並使用 ε 可以無窮小而總結。

Define the following notations

$$\forall k \geqslant 0, \qquad X_k^\varepsilon = (X_k - \varepsilon)\mathbb{1}_{L_\varepsilon},$$
$$\forall n \geqslant 0, \qquad S_n^\varepsilon = \sum_{k=0}^{n-1} X_k^\varepsilon, \qquad M_n^\varepsilon = \max_{0 \leqslant k \leqslant n} S_n^\varepsilon.$$

We note that $L_\varepsilon = \varphi^{-1}(L_\varepsilon)$, hence $L_\varepsilon \in \mathcal{I}$ and $(X_k^\varepsilon)_{k \geqslant 0}$ is a stationarity process. Using Lemma 8.2.2, we obtain

$$\forall n \geqslant 0, \qquad \mathbb{E}[X_0^\varepsilon \mathbb{1}_{M_n^\varepsilon > 0}] \geqslant 0.$$

Taking the limit $n \to \infty$ in the above formula, we get

$$\mathbb{E}[X_0^\varepsilon \mathbb{1}_{\sup_{n \geqslant 0} S_n^\varepsilon > 0}] \geqslant 0.$$

But at the same time, we have

$$\Big\{\sup_{n \geqslant 0} S_n^\varepsilon > 0\Big\} = \Big\{\sup_{n \geqslant 0} \frac{S_n}{n} > \varepsilon\Big\} \cap L_\varepsilon = L_\varepsilon,$$

so

$$0 \leqslant \mathbb{E}[(X_0 - \varepsilon)\mathbb{1}_{L_\varepsilon}] = \mathbb{E}[\mathbb{1}_{L_\varepsilon}\,\mathbb{E}[X \,|\, \mathcal{I}]] - \varepsilon\,\mathbb{P}(L_\varepsilon) = -\varepsilon\,\mathbb{P}(L_\varepsilon).$$

This tells us that $\mathbb{P}(L_\varepsilon) = 0$, giving the a.s. convergence in the ergodic theorem.

Next we discuss the L^1 convergence. We may show that $(\frac{S_n}{n})_{n \geqslant 1}$ is uniformly integrable to deduce the L^1 convergence. But we will use the following more straightforward method.

Given $M > 0$ and define

$$X' = X\mathbb{1}_{|X| \leqslant M} \quad \text{and} \quad X'' = X - X'.$$

Using the a.s. ergodic theorem above, we have

$$\frac{1}{n}\sum_{k=0}^{n-1} X' \circ \varphi^k \xrightarrow[n \to \infty]{} \mathbb{E}[X' \,|\, \mathcal{I}], \qquad \text{a.s.}$$

Since X' is bounded, from the dominated convergence theorem, the above convergence also holds in L^1. Then we note that

$$\mathbb{E}\left[\left|\Big(\frac{1}{n}\sum_{k=0}^{n-1} X'' \circ \varphi^k\Big) - \mathbb{E}[X'' \,|\, \mathcal{I}]\right|\right] \leqslant 2\,\mathbb{E}[|X''|] \xrightarrow[M \to \infty]{} 0.$$

Therefore, for any $\varepsilon > 0$, we can choose a large enough M such that $\mathbb{E}[|X''|] < \varepsilon$, then a large enough n such that

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{k=0}^{n-1} X' \circ \varphi^k - \mathbb{E}[X' \,|\, \mathcal{I}]\right|\right] \leqslant \varepsilon,$$

to deduce the L^1 convergence of the ergodic theorem. □

我們定義下列記號

$$\forall k \geqslant 0, \qquad X_k^\varepsilon = (X_k - \varepsilon)\mathbb{1}_{L_\varepsilon},$$
$$\forall n \geqslant 0, \qquad S_n^\varepsilon = \sum_{k=0}^{n-1} X_k^\varepsilon, \qquad M_n^\varepsilon = \max_{0 \leqslant k \leqslant n} S_n^\varepsilon.$$

我們注意到，$L_\varepsilon = \varphi^{-1}(L_\varepsilon)$，因此 $L_\varepsilon \in \mathcal{I}$，且 $(X_k^\varepsilon)_{k \geqslant 0}$ 是個平穩過程。我們可以利用引理 8.2.2 而得到

$$\forall n \geqslant 0, \qquad \mathbb{E}[X_0^\varepsilon \mathbb{1}_{M_n^\varepsilon > 0}] \geqslant 0.$$

在上式中，我們可以對 $n \to \infty$ 取極限，進而得到

$$\mathbb{E}[X_0^\varepsilon \mathbb{1}_{\sup_{n \geqslant 0} S_n^\varepsilon > 0}] \geqslant 0.$$

但同時，我們有

$$\Big\{\sup_{n \geqslant 0} S_n^\varepsilon > 0\Big\} = \Big\{\sup_{n \geqslant 0} \frac{S_n}{n} > \varepsilon\Big\} \cap L_\varepsilon = L_\varepsilon,$$

因此

$$0 \leqslant \mathbb{E}[(X_0 - \varepsilon)\mathbb{1}_{L_\varepsilon}] = \mathbb{E}[\mathbb{1}_{L_\varepsilon}\,\mathbb{E}[X \,|\, \mathcal{I}]] - \varepsilon\,\mathbb{P}(L_\varepsilon) = -\varepsilon\,\mathbb{P}(L_\varepsilon).$$

這告訴我們 $\mathbb{P}(L_\varepsilon) = 0$，因此得證 a.s. 收斂的遍歷定理。

接著我們來討論 L^1 中的收斂。我們可以證明 $(\frac{S_n}{n})_{n \geqslant 1}$ 為均勻可積，從而直接得到 L^1 收斂，但也可以使用下列比較直接的方法。

給定 $M > 0$，我們定義

$$X' = X\mathbb{1}_{|X| \leqslant M} \quad \text{且} \quad X'' = X - X'.$$

利用前面證明過的 a.s. 遍歷定理，我們會有

$$\frac{1}{n}\sum_{k=0}^{n-1} X' \circ \varphi^k \xrightarrow[n \to \infty]{} \mathbb{E}[X' \,|\, \mathcal{I}], \qquad \text{a.s.}$$

由於 X' 有界，根據勒貝格控制收斂定理，上述收斂也會在 L^1 中成立。接著我們注意到：

$$\mathbb{E}\left[\left|\Big(\frac{1}{n}\sum_{k=0}^{n-1} X'' \circ \varphi^k\Big) - \mathbb{E}[X'' \,|\, \mathcal{I}]\right|\right] \leqslant 2\,\mathbb{E}[|X''|] \xrightarrow[M \to \infty]{} 0.$$

因此對於任意 $\varepsilon > 0$，我們可以選擇夠大的 M 使得 $\mathbb{E}[|X''|] < \varepsilon$，再選擇夠大的 n，使得

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{k=0}^{n-1} X' \circ \varphi^k - \mathbb{E}[X' \,|\, \mathcal{I}]\right|\right] \leqslant \varepsilon,$$

進而得到 L^1 收斂的遍歷定理。□

Question 8.2.3: We keep the notations in Theorem 8.2.1 and assume additionally that $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$. Prove the L^p convergence of the ergodic theorem.

Example 8.2.4 : Below are a few applications of Theorem 8.2.1.

(1) If $(X_k)_{k\geqslant 0}$ is an i.i.d. sequence of random variables in L^1, Remark 8.1.13 says that $\mathcal{I}$ is a trivial σ-algebra, giving the following a.s. convergence and L^1 convergence,

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow[n\to\infty]{} \mathbb{E}[X_0].$$

(2) Consider an irreducible Markov chain $(X_n)_{n\geqslant 0}$ on a countable state space E with invariant probability measure μ. In Proposition 8.1.17, we have seen that the shift operator θ is ergodic with respect to $\mathbb{P}_\mu$. Let $f : E \longrightarrow \mathbb{R}$ with $\mu(|f|) < \infty$, then we have the following $\mathbb{P}_\mu$-a.s. convergence and convergence in L^1,

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \xrightarrow[n\to\infty]{} \mathbb{E}_\mu[f \mid \mathcal{I}_\theta] = \mu(f).$$

(3) We carry on the discussion of the rotation operator on the unit circle from Example 8.1.4 (2). According to Exercise 8.12, when θ is irrational, θ_β is ergodic. Thus, given $A \in \mathcal{B}(\mathbb{S}^1)$, we have

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{\theta_\beta^k(\omega)\in A} \xrightarrow[n\to\infty]{} \mu(A),$$

where μ is the uniform probability measure on $\mathbb{S}^1$. This is Weyl's equidistribution theorem (等分佈定理).

8.3 Recurrence of Stationary Sequences

In this section, we discuss the recurrence of stationary sequences using the ergodic theorem seen in Section 8.2.

Below, we want to consider random processes indexed by time steps in $\mathbb{N}$ with values in $\mathbb{R}^d$. We define the sample space to be $\Omega = (\mathbb{R}^d)^{\mathbb{N}}$ and the usual σ-algebra $\mathcal{F} := \mathcal{B}(\mathbb{R}^d)^{\otimes\mathbb{N}}$, which is the smallest σ-algebra making all the coordinate functions measurable. Let $\mathbb{P}$ be a probability measure such that the shift operator θ is measure-preserving on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Take $\omega \sim \mathbb{P}$, that is the sequence $(X_k(\omega))_{k\geqslant 1}$ given by the coordinate functions is stationary with values in $\mathbb{R}^d$. For all positive integer $n \geqslant 1$, let

$$\begin{aligned} S_n &= X_1 + \cdots + X_n, \\ R_n &= |\{S_1, \cdots, S_n\}|. \end{aligned}$$

If we interprete $(X_k)_{k\geqslant 1}$ as increments of a random walk, then S_n represents its position at time n, R_n the number of sites visited by the walk up to time n.

We have the following result.

問題 8.2.3：我們沿用定理 8.2.1 中的記號，若假設 $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$，試證 L^p 收斂的遍歷定理。

範例 8.2.4 ： 下列是幾個定理 8.2.1 的應用。

(1) 若 $(X_k)_{k\geqslant 0}$ 是個 i.i.d. 且在 L^1 中的隨機變數序列，由註解 8.1.13 得知，$\mathcal{I}$ 為平凡 σ 代數，因此下列會 a.s. 收斂及在 L^1 中收斂：

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow[n\to\infty]{} \mathbb{E}[X_0].$$

(2) 考慮在可數空間 E 上的不可約馬可夫鏈 $(X_n)_{n\geqslant 0}$，並假設他有不變機率測度 μ。我們在命題 8.1.17 中，證明了推移算子 θ 對於 $\mathbb{P}_\mu$ 具有遍歷性。令 $f : E \longrightarrow \mathbb{R}$ 使得 $\mu(|f|) < \infty$，則我們有下列 $\mathbb{P}_\mu$-a.s. 及 L^1 收斂：

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \xrightarrow[n\to\infty]{} \mathbb{E}_\mu[f \mid \mathcal{I}_\theta] = \mu(f).$$

(3) 我們繼續討論範例 8.1.4 (2) 中考慮的單位圓上的旋轉算子。根據習題 8.12 ，我們知道在 θ 為無理數時，θ_β 具有遍歷性。因此，給定 $A \in \mathcal{B}(\mathbb{S}^1)$，我們會有

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{\theta_\beta^k(\omega)\in A} \xrightarrow[n\to\infty]{} \mu(A),$$

其中 μ 是在 $\mathbb{S}^1$ 上的均勻機率測度。這是 Weyl 的等分佈定理 (equidistribution theorem)。

第三節 平穩序列的重現性

在此章節中，我們要用第 8.2 節中的遍歷定理來探討平穩序列的重現性。

以下我們要考慮的是時間在 $\mathbb{N}$ 上，取值在 $\mathbb{R}^d$ 中的隨機過程。我們定義樣本空間 $\Omega = (\mathbb{R}^d)^{\mathbb{N}}$，定義 σ 代數 $\mathcal{F} := \mathcal{B}(\mathbb{R}^d)^{\otimes\mathbb{N}}$ 為使其座標函數可測的最小 σ 代數，及機率測度 $\mathbb{P}$ 使得推移算子 θ 會是機率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的測度守恆變換。

考慮 $\omega \sim \mathbb{P}$，換句話說，座標函數給定的序列 $(X_k(\omega))_{k\geqslant 1}$ 為取值在 $\mathbb{R}^d$ 中的平穩序列。對於所有正整數 $n \geqslant 1$，令

$$\begin{aligned} S_n &= X_1 + \cdots + X_n, \\ R_n &= |\{S_1, \cdots, S_n\}|. \end{aligned}$$

如果把 $(X_k)_{k\geqslant 1}$ 看作隨機漫步的步差增量，則 S_n 為隨機漫步在時間 n 的位置，R_n 為隨機漫步至時間 n 為止，拜訪過的位置數量。

我們有下列結果。

Theorem 8.3.1 : *Let*

$$A = \{S_k \neq 0,\ \forall k \geqslant 1\}.$$

Then we have the following convergence,

$$\frac{R_n}{n} \xrightarrow{a.s.} \mathbb{P}(A \,|\, \mathcal{I}).$$

Proof : First, we note that R_n can be represented as $R_n = |B_n|$ with

$$B_n = \{m \geqslant 1 : S_{m+k} \neq S_m,\ \forall 1 \leqslant k \leqslant n-m\}, \qquad \forall n \geqslant 1.$$

Since the following equivalence holds for all integers $m \geqslant 1$,

$$\theta^m(\omega) \in A \Leftrightarrow \omega \in \{S_{m+k} - S_m \neq 0,\ \forall k \geqslant 1\},$$

we have

$$R_n \geqslant \sum_{m=1}^{n} \mathbb{1}_A(\theta^m(\omega)), \qquad \forall n \geqslant 1.$$

From Birkhoff's ergodic theorem (Theorem 8.2.1), we have

$$\liminf_{n\to\infty} \frac{R_n}{n} \geqslant \mathbb{E}[\mathbb{1}_A \,|\, \mathcal{I}], \qquad \text{a.s.}$$

Next, we consider, for a given integer $k \geqslant 1$,

$$A_k = \{S_j \neq 0,\ \forall 1 \leqslant j \leqslant k\}.$$

Similar to the above discussion, we have

$$R_n \leqslant k + \sum_{m=1}^{n-k} \mathbb{1}_{A_k}(\theta^m(\omega)).$$

We use Birkhoff's ergodic theorem again to obtain

$$\limsup_{n\to\infty} \frac{R_n}{n} \leqslant \mathbb{E}[\mathbb{1}_{A_k} \,|\, \mathcal{I}], \qquad \text{a.s.}$$

Due to the monotonicity of the sequence $(A_k)_{k\geqslant 1}$ and the fact that A_k decreases to A when $k \to \infty$, we conclude using the monotone convergence theorem. □

The below theorem extends the results on random walks with i.i.d. increments (Example 7.3.9 and Exercise 7.9) to random walks with increments given by a stationary sequence.

Theorem 8.3.2 : *Consider a stationary sequence $(X_k)_{k\geqslant 1}$ in L^1 with values in $\mathbb{Z}$. Then, we have the following properties.*

(1) *If $\mathbb{E}[X_1 \,|\, \mathcal{I}] = 0$, then $\mathbb{P}(A) = 0$.*

定理 8.3.1 ： 令

$$A = \{S_k \neq 0,\ \forall k \geqslant 1\}.$$

則我們有下列收斂：

$$\frac{R_n}{n} \xrightarrow{\text{a.s.}} \mathbb{P}(A \,|\, \mathcal{I}).$$

證明：首先我們注意到，R_n 可以寫作 $R_n = |B_n|$ 其中 B_n 定義做：

$$B_n = \{m \geqslant 1 : S_{m+k} \neq S_m,\ \forall 1 \leqslant k \leqslant n-m\}, \qquad \forall n \geqslant 1.$$

由於對於所有整數 $m \geqslant 1$，我們有下列等價關係：

$$\theta^m(\omega) \in A \Leftrightarrow \omega \in \{S_{m+k} - S_m \neq 0,\ \forall k \geqslant 1\},$$

因此我們有

$$R_n \geqslant \sum_{m=1}^{n} \mathbb{1}_A(\theta^m(\omega)), \qquad \forall n \geqslant 1.$$

根據 Birkhoff 遍歷定理（定理 8.2.1），我們得到

$$\liminf_{n\to\infty} \frac{R_n}{n} \geqslant \mathbb{E}[\mathbb{1}_A \,|\, \mathcal{I}], \qquad \text{a.s.}$$

接著，給定整數 $k \geqslant 1$，我們考慮

$$A_k = \{S_j \neq 0,\ \forall 1 \leqslant j \leqslant k\}.$$

與上面的討論相似，我們可以得到

$$R_n \leqslant k + \sum_{m=1}^{n-k} \mathbb{1}_{A_k}(\theta^m(\omega)).$$

再次根據 Birkhoff 遍歷定理，我們有

$$\limsup_{n\to\infty} \frac{R_n}{n} \leqslant \mathbb{E}[\mathbb{1}_{A_k} \,|\, \mathcal{I}], \qquad \text{a.s.}$$

由於序列 $(A_k)_{k\geqslant 1}$ 有單調性，且當 $k \to \infty$ 時會遞減至 A，我們以單調收斂定理總結。 □

下列定理將 i.i.d. 步差的隨機漫步的結果（範例 7.3.9 及習題 7.9）推廣至平穩序列的步差。

定理 8.3.2 ： 假設 $(X_k)_{k\geqslant 1}$ 取值在 $\mathbb{Z}$ 中的 L^1 平穩序列，則我們有：

(1) 若 $\mathbb{E}[X_1 \,|\, \mathcal{I}] = 0$，則 $\mathbb{P}(A) = 0$。

(2) If $\mathbb{P}(A) = 0$, then $\mathbb{P}(S_n = 0\ i.o.) = 1$.

This means that if the conditional expectation of the increments of a random walk is zero with respect to the invariant σ-algebra, then it is recurrent.

Proof :

(1) If $\mathbb{E}[X_1 \,|\, \mathcal{I}] = 0$, then the ergodic theorem says that $\frac{S_n}{n}$ converges to 0 almost surely. Thus, we have

$$\limsup_{n\to\infty}\left(\max_{1\leqslant k\leqslant n}\frac{|S_k|}{n}\right) = \limsup_{n\to\infty}\left(\max_{K\leqslant k\leqslant n}\frac{|S_k|}{n}\right) \leqslant \max_{k\geqslant K}\frac{|S_k|}{k} \xrightarrow[K\to\infty]{} 0,$$

that is

$$\lim_{n\to\infty}\left(\max_{1\leqslant k\leqslant n}\frac{|S_k|}{n}\right) = 0.$$

Since $R_n \leqslant 1 + 2\max_{1\leqslant k\leqslant n}|S_k|$, we get $\frac{R_n}{n} \longrightarrow 0$ and Theorem 8.3.1 implies $\mathbb{P}(A) = 0$.

(2) We note that $A^c = \{\exists n \geqslant 1 : S_n = 0\}$, which means that if $\mathbb{P}(A) = 0$, then $(S_n)_{n\geqslant 1}$ visits 0 at least once with probability 1. According to the first visit time of $(S_n)_{n\geqslant 1}$ to 0, we decompose A^c into a disjoint union of events $A^c = \sqcup_{k\geqslant 1} F_k$ with

$$F_k = \{S_i \neq 0 \text{ for all } 1 \leqslant i < k \text{ and } S_k = 0\}, \qquad \forall k \geqslant 1.$$

Since we want to look at the probability that $(S_n)_{n\geqslant 1}$ visits 0 at least two times, we define the following events,

$$G_{j,k} = \theta^j(F_k) = \{S_{j+i} - S_j \neq 0 \text{ for all } 1 \leqslant i < k \text{ and } S_{j+k} - S_j = 0\}, \qquad \forall j,k \geqslant 1.$$

Due to the stationarity of $(X_k)_{k\geqslant 1}$, we know that $\mathbb{P}(G_{j,k}) = \mathbb{P}(F_k)$ for all $j,k \geqslant 1$ and when j is fixed, $(G_{j,k})_{k\geqslant 1}$ is a collection of disjoint sets with total measure 1. In other words, we have

$$\sum_{j,k\geqslant 1}\mathbb{P}(F_j \cap G_{j,k}) = \sum_{j\geqslant 1}\mathbb{P}(F_j) = 1.$$

The event $F_j \cap G_{j,k}$ represents $S_j = 0$ and $S_{j+k} = 0$, hence we have shown that $(S_n)_{n\geqslant 1}$ visits 0 at least twice with probability 1. We can repeat this technique to show that for any $m \geqslant 1$, $(S_n)_{n\geqslant 1}$ visits 0 at least m times with probability 1. □

To close this section, we extend the result of Corollary 7.4.13 to stationary sequences.

Theorem 8.3.3 : *Consider a stationary sequence $(X_n)_{n\geqslant 0}$ in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a measurable set $A \in \mathcal{F}$ and define the stopping times $T_0 = 0$ and*

$$T_{n+1} = \inf\{k > T_n : X_k \in A\}, \qquad \forall n \geqslant 0.$$

(2) 若 $\mathbb{P}(A) = 0$，則 $\mathbb{P}(S_n = 0\ i.o.) = 1$。

也就是說，若隨機漫步的步差對於不變 σ 代數的條件期望值 $\mathbb{E}[X_1 \,|\, \mathcal{I}]$ 為 0，則他是重現的。

證明：

(1) 若 $\mathbb{E}[X_1 \,|\, \mathcal{I}] = 0$，則遍歷定理告訴我們 $\frac{S_n}{n}$ 會 a.s. 收斂至 0。因此，我們有

$$\limsup_{n\to\infty}\left(\max_{1\leqslant k\leqslant n}\frac{|S_k|}{n}\right) = \limsup_{n\to\infty}\left(\max_{K\leqslant k\leqslant n}\frac{|S_k|}{n}\right) \leqslant \max_{k\geqslant K}\frac{|S_k|}{k} \xrightarrow[K\to\infty]{} 0,$$

也就是說

$$\lim_{n\to\infty}\left(\max_{1\leqslant k\leqslant n}\frac{|S_k|}{n}\right) = 0.$$

由於 $R_n \leqslant 1 + 2\max_{1\leqslant k\leqslant n}|S_k|$，我們會得到 $\frac{R_n}{n} \longrightarrow 0$，因此定理 8.3.1 給出 $\mathbb{P}(A) = 0$。

(2) 首先我們有 $A^c = \{\exists n \geqslant 1 : S_n = 0\}$，也就是若 $\mathbb{P}(A) = 0$，則 $(S_n)_{n\geqslant 1}$ 拜訪 0 至少一次的機率為 1。我們根據 $(S_n)_{n\geqslant 1}$ 第一次拜訪 0 的時間將 A^c 拆做互斥事件的聯集 $A^c = \sqcup_{k\geqslant 1} F_k$，其中

$$F_k = \{S_i \neq 0 \text{ 對於所有 } 1 \leqslant i < k \text{ 且 } S_k = 0\}, \qquad \forall k \geqslant 1.$$

接著，我們想要探討 $(S_n)_{n\geqslant 1}$ 拜訪 0 至少兩次的機率，因此我們定義下列事件

$$G_{j,k} = \theta^j(F_k) = \{S_{j+i} - S_j \neq 0 \text{ 對於所有 } 1 \leqslant i < k \text{ 且 } S_{j+k} - S_j = 0\}, \qquad \forall j,k \geqslant 1.$$

根據 $(X_k)_{k\geqslant 1}$ 的平穩性，我們知道對於所有 $j,k \geqslant 1$ 會有 $\mathbb{P}(G_{j,k}) = \mathbb{P}(F_k)$ 且當 j 固定時，$(G_{j,k})_{k\geqslant 1}$ 為互斥集合且總測度為 1。換句話說，我們有

$$\sum_{j,k\geqslant 1}\mathbb{P}(F_j \cap G_{j,k}) = \sum_{j\geqslant 1}\mathbb{P}(F_j) = 1.$$

由於事件 $F_j \cap G_{j,k}$ 代表 $S_j = 0$ 及 $S_{j+k} = 0$，因此我們證明了，$(S_n)_{n\geqslant 1}$ 拜訪 0 至少兩次的機率為 1。我們重複此技巧，可以證明對於任意 $m \geqslant 1$，$(S_n)_{n\geqslant 1}$ 拜訪 0 至少 m 次的機率為 1。 □

在結束此章節前，我們將系理 7.4.13 的結果推廣至平穩序列上。

定理 8.3.3 ： 考慮在機率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的平穩序列 $(X_n)_{n\geqslant 0}$，可測集合 $A \in \mathcal{F}$，並定義停止時間 $T_0 = 0$ 及

$$T_{n+1} = \inf\{k > T_n : X_k \in A\}, \qquad \forall n \geqslant 0.$$

If we have $T_1 < \infty$ $\mathbb{P}$-a.s., then under the conditional probability $\mathbb{P}(\cdot \mid X_0 \in A)$, the sequence of random variables $(\tau_n := T_{n+1} - T_n)_{n\geqslant 0}$ is stationary and

$$\mathbb{E}[\tau_0 \mid X_0 \in A] = \frac{1}{\mathbb{P}(X_0 \in A)}. \tag{8.3}$$

Remark 8.3.4 : If $(X_n)_{n\geqslant 0}$ is a positive recurrent and irreducible Markov chain with the unique invariant probability measure μ, then under $\mathbb{P}_\mu$, the process $(X_n)_{n\geqslant 0}$ is stationary. For any $x \in E$, if we take $A = \{x\}$, then Eq. (8.3) simplifies to Eq. (7.19). Thus, Theorem 8.3.3 can be seen as an extension of Corollary 7.4.13, where the assumption of the Markov chain is removed and the starting point needs not be a fixed point almost surely.

Proof : Since we need to take conditional probabilities and conditional expectations with respect to $X_0 \in A$, we can use the following trick to simplify the computations. By Exercise 8.2, we can extend $(X_n)_{n\geqslant 0}$ to a stationary sequence $(X_n)_{n\in\mathbb{Z}}$ indexed by $\mathbb{Z}$. Then, define the following disjoint events

$$C_k = \{X_{-1} \notin A, \ldots, X_{-(k-1)} \notin A, X_{-k} \in A\}, \qquad \forall k \geqslant 1.$$

We have

$$\Big(\bigsqcup_{k=1}^{K} C_k\Big)^c = \{X_k \notin A \text{ for all } -K \leqslant k \leqslant -1\}.$$

By stationarity, the above event on the right side has the same probability as $\{X_k \notin A \text{ for all } 1 \leqslant k \leqslant K\}$, taking $K \to \infty$ and using the assumption of the theorem, we have

$$\mathbb{P}\Big(\bigsqcup_{k=1}^{\infty} C_k\Big) = 1. \tag{8.4}$$

Now, we are ready to show that under the conditional probability $\mathbb{P}(\cdot \mid X_0 \in A)$, the sequence $(\tau_n)_{n\geqslant 0}$ is stationary. Given a positive integer j, non-negative integers $n_1 < n_2 < \cdots < n_j$ and positive integers $m_1, \cdots, m_j$, we have

$$\begin{aligned}\mathbb{P}(\tau_{n_1+1} = m_1, \ldots, \tau_{n_j+1} = m_j, X_0 \in A) &= \sum_{k\geqslant 1} \mathbb{P}(\tau_0 = k, \tau_{n_1+1} = m_1, \ldots, \tau_{n_j+1} = m_j, X_0 \in A)\\ &= \sum_{k\geqslant 1} \mathbb{P}(C_k, X_0 \in A, \tau_{n_1} = m_1, \ldots, \tau_{n_j} = m_j),\\ &= \mathbb{P}(X_0 \in A, \tau_{n_1} = m_1, \ldots, \tau_{n_j} = m_j),\end{aligned}$$

where we use the assumption of stationarity in the second equality, and Eq. (8.4) in the last equality to conclude.

若 $\mathbb{P}$-a.s. 我們有 $T_1 < \infty$，則在條件機率 $\mathbb{P}(\cdot \mid X_0 \in A)$ 之下，隨機變數序列 $(\tau_n := T_{n+1} - T_n)_{n\geqslant 0}$ 會是個平穩序列且

$$\mathbb{E}[\tau_0 \mid X_0 \in A] = \frac{1}{\mathbb{P}(X_0 \in A)}. \tag{8.3}$$

註解 8.3.4 ： 若 $(X_n)_{n\geqslant 0}$ 為正重現且不可約的馬可夫鏈，將其唯一不變機率測度記作 μ，則在 $\mathbb{P}_\mu$ 之下，$(X_n)_{n\geqslant 0}$ 是個平穩序列。對於任意 $x \in E$，我們取 $A = \{x\}$，則式 (8.3) 化簡為式 (7.19)。因此定理 8.3.3 可以視為系理 7.4.13 的推廣，我們移除馬可夫鏈的假設，並且出發點未必是 a.s. 為固定點。

證明：由於必須要對 $X_0 \in A$ 做條件機率及條件期望值，我們可以使用下列技巧，來簡化這樣的計算。利用習題 8.2 ，我們可以將 $(X_n)_{n\geqslant 0}$ 拓延成下標在 $\mathbb{Z}$ 上的平穩序列 $(X_n)_{n\in\mathbb{Z}}$。接著，定義下列互斥事件

$$C_k = \{X_{-1} \notin A, \ldots, X_{-(k-1)} \notin A, X_{-k} \in A\}, \qquad \forall k \geqslant 1.$$

我們有

$$\Big(\bigsqcup_{k=1}^{K} C_k\Big)^c = \{X_k \notin A \text{ 對於所有 } -K \leqslant k \leqslant -1\}.$$

使用平穩序列的性質，上式右側事件與事件 $\{X_k \notin A \text{ 對於所有 } 1 \leqslant k \leqslant K\}$ 有相同的機率，取 $K \to \infty$，根據定理假設，我們有

$$\mathbb{P}\Big(\bigsqcup_{k=1}^{\infty} C_k\Big) = 1. \tag{8.4}$$

我們來證明在條件機率 $\mathbb{P}(\cdot \mid X_0 \in A)$ 之下 $(\tau_n)_{n\geqslant 0}$ 的平穩性。給定正整數 j、非負整數 $n_1 < n_2 < \cdots < n_j$ 及正整數 $m_1, \cdots, m_j$，我們有

$$\begin{aligned}\mathbb{P}(\tau_{n_1+1} = m_1, \ldots, \tau_{n_j+1} = m_j, X_0 \in A) &= \sum_{k\geqslant 1} \mathbb{P}(\tau_0 = k, \tau_{n_1+1} = m_1, \ldots, \tau_{n_j+1} = m_j, X_0 \in A)\\ &= \sum_{k\geqslant 1} \mathbb{P}(C_k, X_0 \in A, \tau_{n_1} = m_1, \ldots, \tau_{n_j} = m_j),\\ &= \mathbb{P}(X_0 \in A, \tau_{n_1} = m_1, \ldots, \tau_{n_j} = m_j),\end{aligned}$$

其中在第二個等式中，我們使用了平穩序列的假設；在第三個等式中，使用了式 (8.4) 總結。

Finally, let us prove Eq. (8.3),

$$\begin{aligned}\mathbb{E}[\tau_0 \mid X_0 \in A] &= \sum_{k\geqslant 1}\mathbb{P}(\tau_0 \geqslant k \mid X_0 \in A)\\ &= \mathbb{P}(X_0 \in A)^{-1}\sum_{k\geqslant 1}\mathbb{P}(\tau_0 \geqslant k, X_0 \in A)\\ &= \mathbb{P}(X_0 \in A)^{-1}\sum_{k\geqslant 1}\mathbb{P}(C_k) = \mathbb{P}(X_0 \in A)^{-1},\end{aligned}$$

where in the third equality, we use the property of a stationary sequence,

$$\begin{aligned}\mathbb{P}(\tau_0 \geqslant k, X_0 \in A) &= \mathbb{P}(X_0 \in A, X_1 \notin A, \dots, X_{k-1} \notin A)\\ &= \mathbb{P}(\theta^{-k}(X_0 \in A, X_1 \notin A, \dots, X_{k-1} \notin A)) = \mathbb{P}(C_k).\end{aligned}$$

□

Question 8.3.5: We keep the notations from Theorem 8.3.3. Let $B \in \mathcal{F}$ satisfying $A \cap B = \varnothing$. Show

$$\mathbb{E}\left[\sum_{k=0}^{T_1-1}\mathbb{1}_{X_k\in B}\,\middle|\, X_0 \in A\right] = \frac{\mathbb{P}(X_0 \in B)}{\mathbb{P}(X_0 \in A)}.$$

How to compare this to the case of a Markov chain in Theorem 7.4.10?

8.4 Subadditive Ergodic Theorem

In this chapter, we will discuss a generalization of Birkhoff's ergodic theorem. In Theorem 8.2.1, the assumption on the stationarity of the sequence of random variables was necessary to obtain the ergodic theorem. However, in real-world problems, this assumption is still too restrictive and hence unrealistic. Later, we will see in Theorem 8.4.3 that under appropriate conditions, even without stationarity, we are still able to get a similar result as Birkhoff's ergodic theorem.

A real sequence $(x_n)_{n\geqslant 1}$ is said to be *subadditive* (劣加性) if it satisfies

$$x_{n+m} \leqslant x_n + x_m, \qquad \forall n, m \geqslant 1.$$

Lemma 8.4.1 (Fekete 引理) **:** *The following limit exists for any subadditive sequence* $(x_n)_{n\geqslant 1}$,

$$\lim_{n\to\infty}\frac{x_n}{n} = \inf_{n\geqslant 1}\frac{x_n}{n} \in \mathbb{R}\cup\{-\infty\}.$$

Proof : First note that we have

$$\liminf_{n\to\infty}\frac{x_n}{n} \geqslant \inf_{n\geqslant 1}\frac{x_n}{n}.$$

最後，我們來證明式 (8.3)：

$$\begin{aligned}\mathbb{E}[\tau_0 \mid X_0 \in A] &= \sum_{k\geqslant 1}\mathbb{P}(\tau_0 \geqslant k \mid X_0 \in A)\\ &= \mathbb{P}(X_0 \in A)^{-1}\sum_{k\geqslant 1}\mathbb{P}(\tau_0 \geqslant k, X_0 \in A)\\ &= \mathbb{P}(X_0 \in A)^{-1}\sum_{k\geqslant 1}\mathbb{P}(C_k) = \mathbb{P}(X_0 \in A)^{-1},\end{aligned}$$

其中在第三個不等式中，我們使用了平穩序列的性質：

$$\begin{aligned}\mathbb{P}(\tau_0 \geqslant k, X_0 \in A) &= \mathbb{P}(X_0 \in A, X_1 \notin A, \dots, X_{k-1} \notin A)\\ &= \mathbb{P}(\theta^{-k}(X_0 \in A, X_1 \notin A, \dots, X_{k-1} \notin A)) = \mathbb{P}(C_k).\end{aligned}$$

□

問題 8.3.5：我們沿用定理 8.3.3 中的記號。令 $B \in \mathcal{F}$ 滿足 $A \cap B = \varnothing$，證明

$$\mathbb{E}\left[\sum_{k=0}^{T_1-1}\mathbb{1}_{X_k\in B}\,\middle|\, X_0 \in A\right] = \frac{\mathbb{P}(X_0 \in B)}{\mathbb{P}(X_0 \in A)}.$$

如何與定理 7.4.10 中馬可夫鏈的情況做比較呢？

第四節 劣加性遍歷定理

在此章節中，我們要探討的是 Birkhoff 遍歷定理的加強版。在定理 8.2.1 中，我們對隨機變數序列必須要有平穩性的假設，才有辦法得到遍歷定理；但在很多現實世界的問題中，這樣的假設還是太強，因此顯得不切實際。在定理 8.4.3 中我們會看到，在某些好的假設之下，即使沒有平穩性，我們仍然可以得到類似 Birkhoff 遍歷定理的敘述。

若實數序列 $(x_n)_{n\geqslant 1}$ 滿足

$$x_{n+m} \leqslant x_n + x_m, \qquad \forall n, m \geqslant 1,$$

則我們說他具有劣加性 (subadditive)。

引理 8.4.1 【Fekete 引理】： 對於任意劣加性數列 $(x_n)_{n\geqslant 1}$ 來說，下列極限存在

$$\lim_{n\to\infty}\frac{x_n}{n} = \inf_{n\geqslant 1}\frac{x_n}{n} \in \mathbb{R}\cup\{-\infty\}.$$

證明：首先注意到，我們有

$$\liminf_{n\to\infty}\frac{x_n}{n} \geqslant \inf_{n\geqslant 1}\frac{x_n}{n}.$$

Given $\varepsilon > 0$, take $N \geqslant 1$ such that

$$\frac{x_N}{N} \leqslant \inf_{n\geqslant 1} \frac{x_n}{n} + \varepsilon.$$

For any $n \geqslant N$, we can write $n = kN + r$ with $k \in \mathbb{N}$ and $0 \leqslant r \leqslant N-1$. By subadditivity, we have

$$\frac{x_n}{n} \leqslant \frac{kx_N}{n} + \frac{x_r}{n}.$$

Taking $\limsup$ in the above formula, we get

$$\limsup_{n\to\infty} \frac{x_n}{n} \leqslant \frac{x_N}{N} \leqslant \inf_{n\geqslant 1} \frac{x_n}{n} + \varepsilon.$$

Since $\varepsilon > 0$ can be arbitrarily small, the proof is complete. □

Example 8.4.2 (self-avoiding walk) **:** We are given a path $(x_k)_{0\leqslant k\leqslant n}$ of length $n \geqslant 1$ on $\mathbb{Z}^d$. It is called a *self-avoiding walk* (自迴避路徑) of length n if

- we have $x_k \sim x_{k+1}$ for all $0 \leqslant k \leqslant n-1$;
- we have $x_j \neq x_k$ for all $0 \leqslant j < k \leqslant n$.

Let us denote by $\mathrm{SAW}(n)$ the set of all the self-avoiding walks starting from 0 of length n and denote its cardinal by $c_n = |\mathrm{SAW}(n)|$. The sequence $(\ln c_n)_{n\geqslant 1}$ is subadditive and one can apply Lemma 8.4.1 to obtain

$$\gamma := \lim_{n\to\infty} \frac{\ln c_n}{n} = \inf_{n\geqslant 1} \frac{\ln c_n}{n} \in \mathbb{R} \cup \{-\infty\}.$$

The constant $\mu = e^{\gamma}$ is called *connective constant* (連通常數).

On any periodic lattice in any dimension, we can use Lemma 8.4.1 to show the existence of the connective constant but it is highly non trivial to compute its exact value. The only exact computation available at the current moment is on the planar hexagonal lattice with value $\mu = \sqrt{2+\sqrt{2}}$. [1]

The below theorem is the stochastic version of Fekete's lemma.

Theorem 8.4.3 (Subadditive ergodic theorem) **:** *Let $(X_{m,n})_{n>m\geqslant 0}$ be a sequence of random variables satisfying*

(a) $X_{0,n} \leqslant X_{0,m} + X_{m,n}$ for all $n > m \geqslant 1$.

(b) The sequence $(X_{nk,(n+1)k})_{n\geqslant 0}$ is a stationary sequence for all integers $k \geqslant 1$.

(c) The distribution of the sequence $(X_{m,m+k})_{k\geqslant 1}$ does not depend on $m \geqslant 0$.

---

[1] Duminil-Copin, Hugo and Stanislav Smirnov (2012). "The connective constant of the honeycomb lattice equals $\sqrt{2+\sqrt{2}}$". *Annals of Mathematics*, 1653-1665.

給定 $\varepsilon > 0$，取 $N \geqslant 1$ 使得

$$\frac{x_N}{N} \leqslant \inf_{n\geqslant 1} \frac{x_n}{n} + \varepsilon.$$

對於任意 $n \geqslant N$，我們可以將他寫作 $n = kN + r$，其中 $k \in \mathbb{N}$ 及 $0 \leqslant r \leqslant N-1$，使用劣加性，我們可以得到

$$\frac{x_n}{n} \leqslant \frac{kx_N}{n} + \frac{x_r}{n}.$$

對上式取 $\limsup$，我們得到

$$\limsup_{n\to\infty} \frac{x_n}{n} \leqslant \frac{x_N}{N} \leqslant \inf_{n\geqslant 1} \frac{x_n}{n} + \varepsilon.$$

由於 $\varepsilon > 0$ 可以無限小，得證。 □

範例 8.4.2 【自迴避漫步】**：** 在 $\mathbb{Z}^d$ 上，給定長度為 $n \geqslant 1$ 的路徑 $(x_k)_{0\leqslant k\leqslant n}$，若滿足

- 對於所有 $0 \leqslant k \leqslant n-1$，我們有 $x_k \sim x_{k+1}$；
- 對於所有 $0 \leqslant j < k \leqslant n$，我們有 $x_j \neq x_k$，

則我們說 $(x_k)_{0\leqslant k\leqslant n}$ 是個長度為 n 的自迴避路徑 (self-avoiding walk)。我們將從原點 0 出發，長度為 n 的自迴避路徑構成的集合記作 $\mathrm{SAW}(n)$，並將其元素個數記作 $c_n = |\mathrm{SAW}(n)|$。則數列 $(\ln c_n)_{n\geqslant 1}$ 具有劣加性並且可以使用引理 8.4.1 得到

$$\gamma := \lim_{n\to\infty} \frac{\ln c_n}{n} = \inf_{n\geqslant 1} \frac{\ln c_n}{n} \in \mathbb{R} \cup \{-\infty\}.$$

常數 $\mu = e^{\gamma}$ 稱作連通常數 (connective constant)。

在任意維度的週期性網格上，我們都可以利用引理 8.4.1 來證明連通常數存在，但計算他的值卻不是件容易的事情。目前關於連通常數唯一的確切計算是在二維的六角網格網格上，其值為 $\mu = \sqrt{2+\sqrt{2}}$。[1]

我們下面接著要探討的定理是隨機變數版本的 Fekete 引理。

定理 8.4.3 【劣加性遍歷定理】**：** 令 $(X_{m,n})_{n>m\geqslant 0}$ 為滿足下列性質的實隨機變數序列：

(a) 對於所有整數 $n > m \geqslant 1$，我們有 $X_{0,n} \leqslant X_{0,m} + X_{m,n}$。

(b) 對於所有整數 $k \geqslant 1$，序列 $(X_{nk,(n+1)k})_{n\geqslant 0}$ 是個平穩序列。

(c) 序列 $(X_{m,m+k})_{k\geqslant 1}$ 的分佈與 $m \geqslant 0$ 無關。

---

[1] Duminil-Copin, Hugo and Stanislav Smirnov (2012). "The connective constant of the honeycomb lattice equals $\sqrt{2+\sqrt{2}}$". *Annals of Mathematics*, 1653-1665.

(d) $X_{0,n} \in L^1$ for all $n \geqslant 1$.

$$\inf_{n\geqslant 1} \frac{\mathbb{E}[X_{0,n}]}{n} > -\infty.$$

Then, the following statements holds.

(1) The convergence holds almost surely and in L^1,

$$X_\infty := \lim_{n\to\infty} \frac{X_{0,n}}{n}.$$

(2) The following equality holds,

$$\mathbb{E}[X_\infty] = \lim_{n\to\infty} \frac{\mathbb{E}[X_{0,n}]}{n} = \inf_{n\geqslant 1} \frac{\mathbb{E}[X_{0,n}]}{n}.$$

(3) If the sequences in the assumption (b) are ergodic for all $k \geqslant 0$, then $X_\infty = \mathbb{E}[X_\infty]$ a.s.

Remark 8.4.4 :

(1) This theorem was first proved in 1973 by Kingman, then improved in 1985 by Liggett. The statement here is the version from Liggett. The difference between two versions lies in the assumption (a), Kingman required the stronger assumption below,

$$X_{l,n} \leqslant X_{l,m} + X_{m,n}, \qquad \forall 0 \leqslant l \leqslant m \leqslant n.$$

(2) If we use the notations from Theorem 8.2.1 and define

$$X_{m,n} = \sum_{k=m}^{n-1} X \circ \varphi^k, \qquad \forall n > m \geqslant 0,$$

then the subadditive ergodic theorem gives the below almost sure convergence and convergence in L^1,

$$X_\infty = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k,$$

where $\mathbb{E}[X_\infty] = \mathbb{E}[X]$. Additionally, for all $A \in \mathcal{I}_\varphi$, we have

$$\mathbb{E}[X_\infty \mathbb{1}_A] = \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k \mathbb{1}_A\right] = \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n} \sum_{k=0}^{n-1} X \mathbb{1}_{\varphi^{-k}(A)}\right] = \mathbb{E}[X \mathbb{1}_A],$$

giving that $X_\infty = \mathbb{E}[X \mid \mathcal{I}_\varphi]$, which is the result from Theorem 8.2.1.

Before proving Theorem 8.4.3, we still need the following Definition 8.4.5 and Lemma 8.4.6.

(d) 對於所有 $n \geqslant 1$，我們有 $X_{0,n} \in L^1$ 且

$$\inf_{n\geqslant 1} \frac{\mathbb{E}[X_{0,n}]}{n} > -\infty.$$

則我們有：

(1) 下列會 a.s. 收斂且在 L^1 中收斂：

$$X_\infty := \lim_{n\to\infty} \frac{X_{0,n}}{n}.$$

(2) 下列等式會成立：

$$\mathbb{E}[X_\infty] = \lim_{n\to\infty} \frac{\mathbb{E}[X_{0,n}]}{n} = \inf_{n\geqslant 1} \frac{\mathbb{E}[X_{0,n}]}{n}.$$

(3) 若對於所有 $k \geqslant 0$，假設 (b) 中的序列皆具有遍歷性，則 $X_\infty = \mathbb{E}[X_\infty]$ a.s.。

註解 8.4.4 ：

(1) 此定理最早於 1973 年由 Kingman 證明出來，接著於 1985 年由 Liggett 改良，這裡的敘述是 Liggett 的版本。Kingman 的假設差別在於 (a)，他要求的假設較強，如下：

$$X_{l,n} \leqslant X_{l,m} + X_{m,n}, \qquad \forall 0 \leqslant l \leqslant m \leqslant n.$$

(2) 若我們回到定理 8.2.1 的記號，並且定義

$$X_{m,n} = \sum_{k=m}^{n-1} X \circ \varphi^k, \qquad \forall n > m \geqslant 0,$$

則劣加性遍歷定理會給我們下列的 a.s. 收斂及在 L^1 中的收斂：

$$X_\infty = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k,$$

其中 $\mathbb{E}[X_\infty] = \mathbb{E}[X]$。此外，對於所有的 $A \in \mathcal{I}_\varphi$，我們會有

$$\mathbb{E}[X_\infty \mathbb{1}_A] = \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k \mathbb{1}_A\right] = \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n} \sum_{k=0}^{n-1} X \mathbb{1}_{\varphi^{-k}(A)}\right] = \mathbb{E}[X \mathbb{1}_A],$$

所以我們也可以推得 $X_\infty = \mathbb{E}[X \mid \mathcal{I}_\varphi]$，也就是定理 8.2.1 給出的結果。

在證明定理 8.4.3 之前，我們還需要下列定義 8.4.5 及引理 8.4.6 。

Definition 8.4.5 : Let X and Y be two random variables with values in $\mathbb{R}^d$. We say that X *stochastically dominates* (隨機優勢) Y if

$$\mathbb{E}[f(X)] \geqslant \mathbb{E}[f(Y)],$$

for all non-decreasing measurable functions $f : \mathbb{R}^d \longrightarrow \mathbb{R}_{\geqslant 0}$, where we consider the partial order $\leqslant$ on $\mathbb{R}^d$ to be defined as

$$(x_1, \dots, x_d) \leqslant (y_1, \dots, y_d) \quad \Leftrightarrow \quad x_i \leqslant y_i \quad \forall i = 1, \dots, d.$$

We may also write $X \succeq Y$.

If we make use of the notion of coupling introduced in Definition 7.5.9, then we can find random variables X' and Y' defined on the same probability space satisfying $X \overset{(d)}{=} X'$, $Y \overset{(d)}{=} Y'$ and $X \geqslant Y$ a.s.

Lemma 8.4.6 : *Given a sequence $(X_n)_{n \geqslant 1}$ of real-valued random variables and assume that the sequence $(X_n^-)_{n \geqslant 1}$ of negative parts is uniformly integrable. Then, Fatou's lemma still holds,*

$$\mathbb{E}\left[\liminf_{n \to \infty} X_n\right] \leqslant \liminf_{n \to \infty} \mathbb{E}[X_n].$$

Proof : Let $\varepsilon > 0$ and $a > 0$ such that

$$\mathbb{E}\left[(X_n^-) \mathbb{1}_{X_n^- \geqslant a}\right] < \varepsilon.$$

First, we apply Fatou's lemma to $(X_n + a)^+$,

$$\mathbb{E}\left[\liminf_{n \to \infty} (X_n + a)\right] \leqslant \mathbb{E}\left[\liminf_{n \to \infty} (X_n + a)^+\right] \leqslant \liminf_{n \to \infty} \mathbb{E}[(X_n + a)^+].$$

Besides, we have, for all $n \geqslant 1$, that

$$(X_n + a)^+ = (X_n + a) + (X_n + a)^- \leqslant (X_n + a) + X_n^- \mathbb{1}_{X_n^- \geqslant a}.$$

We combine the above three inequalities together to obtain

$$\mathbb{E}\left[\liminf_{n \to \infty} (X_n + a)\right] \leqslant \liminf_{n \to \infty} \mathbb{E}[(X_n + a) + X_n^- \mathbb{1}_{X_n^- \geqslant a}] = \liminf_{n \to \infty} \mathbb{E}[(X_n + a)] + \varepsilon.$$

We substract a from both sides and take $\varepsilon \to 0$ to complete the proof. □

Now we are ready to prove Theorem 8.4.3.

Proof : We define the following notations,

$$\forall n \geqslant 1, \quad a_n = \mathbb{E}[X_{0,n}], \qquad \overline{X}_\infty = \limsup_{n \to \infty} \frac{X_{0,n}}{n}, \qquad \underline{X}_\infty = \liminf_{n \to \infty} \frac{X_{0,n}}{n},$$

and decompose the prove into four steps.

定義 8.4.5 ： 令 X 及 Y 為兩個取值在 $\mathbb{R}^d$ 中的隨機變數。若對於所有非遞減的可測函數 $f : \mathbb{R}^d \longrightarrow \mathbb{R}_{\geqslant 0}$，我們有

$$\mathbb{E}[f(X)] \geqslant \mathbb{E}[f(Y)],$$

其中我們考慮在 $\mathbb{R}^d$ 上的偏序 $\leqslant$ 定義做：

$$(x_1, \dots, x_d) \leqslant (y_1, \dots, y_d) \quad \Leftrightarrow \quad x_i \leqslant y_i \quad \forall i = 1, \dots, d,$$

則我們說 X 對於 Y 有隨機優勢 (stochastic dominance) ，也會記作 $X \succeq Y$。

若我們將在定義 7.5.9 中探討的耦合概念帶進來討論，若且唯若 X 對 Y 有隨機優勢，則我們可以找到定義在同一個機率空間上的隨機變數 X' 及 Y'，滿足 $X \overset{(d)}{=} X'$ 及 $Y \overset{(d)}{=} Y'$ 且 $X' \geqslant Y'$ a.s.。

引理 8.4.6 ： 給定實數隨機變數序列 $(X_n)_{n \geqslant 1}$ 並假設由其負數部份構成的序列 $(X_n^-)_{n \geqslant 1}$ 為均勻可積，則我們仍然有 Fatou 引理：

$$\mathbb{E}\left[\liminf_{n \to \infty} X_n\right] \leqslant \liminf_{n \to \infty} \mathbb{E}[X_n].$$

證明：令 $\varepsilon > 0$ 及 $a > 0$ 使得

$$\mathbb{E}\left[(X_n^-) \mathbb{1}_{X_n^- \geqslant a}\right] < \varepsilon.$$

首先，我們可以對 $(X_n + a)^+$ 使用 Fatou 引理：

$$\mathbb{E}\left[\liminf_{n \to \infty} (X_n + a)\right] \leqslant \mathbb{E}\left[\liminf_{n \to \infty} (X_n + a)^+\right] \leqslant \liminf_{n \to \infty} \mathbb{E}[(X_n + a)^+].$$

此外，對於所有 $n \geqslant 1$，我們有

$$(X_n + a)^+ = (X_n + a) + (X_n + a)^- \leqslant (X_n + a) + X_n^- \mathbb{1}_{X_n^- \geqslant a}.$$

我們將上面三個不等式合併而得到

$$\mathbb{E}\left[\liminf_{n \to \infty} (X_n + a)\right] \leqslant \liminf_{n \to \infty} \mathbb{E}[(X_n + a) + X_n^- \mathbb{1}_{X_n^- \geqslant a}] = \liminf_{n \to \infty} \mathbb{E}[(X_n + a)] + \varepsilon.$$

將左右兩邊各扣除 a 並取 $\varepsilon \to 0$，得證。 □

現在讓我們來證明定理 8.4.3 。

證明：我們定義下列記號

$$\forall n \geqslant 1, \quad a_n = \mathbb{E}[X_{0,n}], \qquad \overline{X}_\infty = \limsup_{n \to \infty} \frac{X_{0,n}}{n}, \qquad \underline{X}_\infty = \liminf_{n \to \infty} \frac{X_{0,n}}{n},$$

並將此證明拆解為四步驟：

(i) Use Fekete's lemma to show that a_n/n a.s. converges and denote this limit by γ.

(ii) Show that $\mathbb{E}[\overline{X}_\infty] \leqslant \gamma$.

(iii) Show that $\mathbb{E}[\underline{X}_\infty] \geqslant \gamma$ and hence from (ii) we obtain $X_\infty := \underline{X}_\infty = \overline{X}_\infty$ a.s. and $\mathbb{E}[X_\infty] = \gamma$.

(iv) Show that $X_{0,n}/n$ converges in L^1.

We start with the prove of (i). From assumptions (a) and (c), we have

$$a_{n+m} \leqslant a_n + a_m, \qquad \forall n, m \geqslant 1.$$

So Fekete's subadditive lemma (Lemma 8.4.1) gives

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geqslant 1} \frac{a_n}{n} =: \gamma. \tag{8.5}$$

Then, let us prove (ii). Given a positive integer $m \geqslant 1$, use the assumption (b) and Birkhoff's ergodic theorem (Theorem 8.2.1), the following converges a.s. and in L^1,

$$\frac{1}{k}\sum_{j=0}^{k-1} X_{jm,(j+1)m} \xrightarrow[n\to\infty]{} \mathbb{E}[X_{0,m} \,|\, \mathcal{I}_m] =: A_m,$$

where $\mathcal{I}_m$ is the invariant σ-algebra for the m-th iterated shift operator θ^m. We make use of the assumption (a) to obtain

$$\limsup_{k\to\infty} \frac{X_{0,km}}{km} \leqslant \frac{A_m}{m}. \tag{8.6}$$

However, the quantity we want to look at is $\overline{X}_\infty = \limsup \frac{X_{0,n}}{n}$, so we need to estimate the corresponding missing terms in Eq. (8.6). Any positive integer $n \geqslant 1$ can be written as $n = km + l$ with $1 \leqslant l \leqslant m$, so we have

$$X_{0,n} \leqslant X_{0,km} + X_{km,km+l}. \tag{8.7}$$

If l is fixed, the assumption (c) guarantees that $X_{km,km+l}$ has the same distribution as $X_{0,l} \in L^1$ and we have, for all $\varepsilon > 0$,

$$\sum_{k\geqslant 1} \mathbb{P}(|X_{km,km+l}| > k\varepsilon) = \sum_{k\geqslant 1} \mathbb{P}(|X_{0,l}| > k\varepsilon) = \frac{1}{\varepsilon}\,\mathbb{E}\left[|X_{o,l}|\right] < \infty.$$

By the Borel-Cantelli lemma, this implies that almost surely, we have

$$\lim_{k\to\infty} \frac{X_{km,km+l}}{k} = 0, \quad \forall m \geqslant 1, 1 \leqslant \ell \leqslant m.$$

Therefore, along with Eq. (8.6) and Eq. (8.7), we find

$$\overline{X}_\infty \leqslant \frac{A_m}{m}. \tag{8.8}$$

This also implies

$$\mathbb{E}[\overline{X}_\infty] \leqslant \frac{\mathbb{E}[A_m]}{m} = \frac{a_m}{m}.$$

We can take inf in m in the right side of the above formula to conclude $\mathbb{E}[\overline{X}_\infty] \leqslant \gamma$.

(i) 利用 Fekete 引理來證明 a_n/n 會 a.s. 收斂，並將其極限記作 γ。

(ii) 證明 $\mathbb{E}[\overline{X}_\infty] \leqslant \gamma$。

(iii) 證明 $\mathbb{E}[\underline{X}_\infty] \geqslant \gamma$，因此從 (ii)，我們得到 $X_\infty := \underline{X}_\infty = \overline{X}_\infty$ a.s. 且 $\mathbb{E}[X_\infty] = \gamma$。

(iv) 證明 $X_{0,n}/n$ 會在 L^1 中收斂。

首先來證明 (i)。根據假設 (a) 及 (c)，我們有

$$a_{n+m} \leqslant a_n + a_m, \qquad \forall n, m \geqslant 1.$$

因此根據 Fekete 的劣加性引理（引理 8.4.1），我們有

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geqslant 1} \frac{a_n}{n} =: \gamma. \tag{8.5}$$

接著證明 (ii)。給定正整數 $m \geqslant 1$，利用 (b) 的假設及 Birkhoff 遍歷定理（定理 8.2.1），下列式子會 a.s. 收斂且在 L^1 中收斂：

$$\frac{1}{k}\sum_{j=0}^{k-1} X_{jm,(j+1)m} \xrightarrow[n\to\infty]{} \mathbb{E}[X_{0,m} \,|\, \mathcal{I}_m] =: A_m,$$

其中 $\mathcal{I}_m$ 是 m 次推移算子 θ^m 的不變 σ 代數。再次使用假設 (a) 我們得到

$$\limsup_{k\to\infty} \frac{X_{0,km}}{km} \leqslant \frac{A_m}{m}. \tag{8.6}$$

但由於我們想要計算的是 $\overline{X}_\infty = \limsup \frac{X_{0,n}}{n}$，我們必須去對式 (8.6) 中漏掉的項做估計。我們可以把任意正整數 $n \geqslant 1$ 寫作 $n = km + l$，其中 $1 \leqslant l \leqslant m$，因此我們有

$$X_{0,n} \leqslant X_{0,km} + X_{km,km+l}. \tag{8.7}$$

若我們固定 l，假設 (c) 告訴我們，$X_{km,km+l}$ 的分佈與 $X_{0,l} \in L^1$ 相同，且對於所有 $\varepsilon > 0$，我們有

$$\sum_{k\geqslant 1} \mathbb{P}(|X_{km,km+l}| > k\varepsilon) = \sum_{k\geqslant 1} \mathbb{P}(|X_{0,l}| > k\varepsilon) = \frac{1}{\varepsilon}\,\mathbb{E}\left[|X_{o,l}|\right] < \infty.$$

所以從 Borel-Cantelli 引理我們幾乎必然會有

$$\lim_{k\to\infty} \frac{X_{km,km+l}}{k} = 0, \quad \forall m \geqslant 1, 1 \leqslant \ell \leqslant m.$$

因此，上式再加上式 (8.6) 及式 (8.7)，我們可以推得

$$\overline{X}_\infty \leqslant \frac{A_m}{m}. \tag{8.8}$$

這也給出了下列關係

$$\mathbb{E}[\overline{X}_\infty] \leqslant \frac{\mathbb{E}[A_m]}{m} = \frac{a_m}{m}.$$

We can note that, if in the assumption (b), the sequence $(X_{nk,(n+1)k})_{n\geqslant 0}$ is ergodic for all $k\geqslant 1$, then we can deduce directly $\overline{X}_\infty\leqslant\gamma$ a.s.

Next, let us prove (iii). If we can construct a stationary sequence $(Y_k)_{k\geqslant 1}$ satisfying

$$\begin{cases}\mathbb{E}[Y_k]\geqslant\gamma, & \forall k\geqslant 1,\\ (X_{0,k})_{1\leqslant k\leqslant n}\succeq(Y_1+\dots Y_k)_{1\leqslant k\leqslant n}, \quad \text{a.s.}, & \forall n\geqslant 1,\end{cases}\tag{8.9}$$

then we can easily get (iii).

Given a positive integer $p\geqslant 1$, let U_p be a random variable with uniform distribution on $\{1,\dots,p\}$. Let us assume that $(U_p)_{p\geqslant 1}$ is an independent sequence and that it is independent of $(X_{m,n})_{n>m\geqslant 0}$ as well. For any positive integers $k,p\geqslant 1$, we define

$$Y_k^{(p)}=X_{0,k+U_p}-X_{0,k+U_p-1},$$

then we have

$$\mathbb{E}[Y_k^{(p)}]=\frac{1}{p}\,\mathbb{E}[X_{0,k+p}-X_{0,k}]=\frac{1}{p}(a_{k+p}-a_k).$$

Then, the assumptions (a) and (c) give that

$$\mathbb{E}[(Y_k^{(p)})^+]=\frac{1}{p}\sum_{l=1}^{p}\mathbb{E}[(X_{0,k+l}-X_{0,k+l-1})^+]\leqslant\frac{1}{p}\sum_{l=1}^{p}\mathbb{E}[X^+_{k+l-1,k+l}]=\mathbb{E}[X^+_{0,1}].$$

In consequence, when k is fixed, we obtain

$$\begin{aligned}\sup_{p\geqslant 1}\mathbb{E}[|Y_k^{(p)}|]&=\sup_{p\geqslant 1}\big(2\,\mathbb{E}[(Y_k^{(p)})^+]-\mathbb{E}[Y_k^{(p)}]\big)\\&\leqslant 2\,\mathbb{E}[X^+_{0,1}]-\inf_{p\geqslant 1}\big(\frac{1}{p}(a_{k+p}-a_k)\big)<\infty,\end{aligned}$$

and also the following using Eq. (8.5),

$$\lim_{p\to\infty}\mathbb{E}[Y_k^{(p)}]=\gamma.$$

By Exercise 1.26, we know that an L^1-bounded sequence of random variables has an almost surely converging subsequence, moreover, the a.s. convergence implies the convergence in distribution, so Cantor's diagonal argument (對角論證法) provides us with a subsequence $(p_i)_{i\geqslant 1}$ such that the sequence $(Y_k^{(p_i)})_{k\geqslant 1}$ of random variables converges in distribution with limit denoted by $(Y_k)_{k\geqslant 1}$. This is equivalent to saying that for any $k\geqslant 1$ and a bounded measurable function f on $\mathbb{R}^k$, the following convergence holds,

$$\mathbb{E}[f(Y_1,\dots,Y_k)]=\lim_{i\to\infty}\frac{1}{p_i}\sum_{l=1}^{p_i}\mathbb{E}[f(X_{0,l+1}-X_{0,l},\dots,X_{0,l+k}-X_{0,l+k-1})].$$

From this formula we can also conclude that $(Y_k)_{k\geqslant 1}$ is stationary.

Now let us check that the sequence $(Y_k)_{k\geqslant 1}$ constructed above satisfies Eq. (8.9). First, the assumptions (a) and (c) give us

$$Y_1^{(p)}=X_{0,U_p+1}-X_{0,U_p}\leqslant X_{U_p,U_p+1}\overset{(d)}{=}X_{0,1},$$

我們在上式右方中對 m 取 $\inf$ 可以得到 $\mathbb{E}[\overline{X}_\infty]\leqslant\gamma$。

我們注意到，若在假設 (b) 中，對於所有 $k\geqslant 1$，序列 $(X_{nk,(n+1)k})_{n\geqslant 0}$ 具有遍歷性，則我們可以直接推得 $\overline{X}_\infty\leqslant\gamma$ a.s.。

再來證明 (iii)。若我們可以構造一個平穩序列 $(Y_k)_{k\geqslant 1}$ 滿足

$$\begin{cases}\mathbb{E}[Y_k]\geqslant\gamma, & \forall k\geqslant 1,\\ (X_{0,k})_{1\leqslant k\leqslant n}\succeq(Y_1+\dots Y_k)_{1\leqslant k\leqslant n}, \quad \text{a.s.}, & \forall n\geqslant 1,\end{cases}\tag{8.9}$$

則我們可以輕易得到 (iii)。

給定正整數 $p\geqslant 1$，令 U_p 為在 $\{1,\dots,p\}$ 上有均勻分佈的隨機變數。我們假設 $(U_p)_{p\geqslant 1}$ 為獨立隨機變數序列，而且也和 $(X_{m,n})_{n>m\geqslant 0}$ 相互獨立。對於任意正整數 $k,p\geqslant 1$，我們定義

$$Y_k^{(p)}=X_{0,k+U_p}-X_{0,k+U_p-1},$$

則我們有

$$\mathbb{E}[Y_k^{(p)}]=\frac{1}{p}\,\mathbb{E}[X_{0,k+p}-X_{0,k}]=\frac{1}{p}(a_{k+p}-a_k).$$

接著根據假設 (a) 及 (c)，我們有

$$\mathbb{E}[(Y_k^{(p)})^+]=\frac{1}{p}\sum_{l=1}^{p}\mathbb{E}[(X_{0,k+l}-X_{0,k+l-1})^+]\leqslant\frac{1}{p}\sum_{l=1}^{p}\mathbb{E}[X^+_{k+l-1,k+l}]=\mathbb{E}[X^+_{0,1}].$$

因此，在固定 k 的情況下，我們得到

$$\begin{aligned}\sup_{p\geqslant 1}\mathbb{E}[|Y_k^{(p)}|]&=\sup_{p\geqslant 1}\big(2\,\mathbb{E}[(Y_k^{(p)})^+]-\mathbb{E}[Y_k^{(p)}]\big)\\&\leqslant 2\,\mathbb{E}[X^+_{0,1}]-\inf_{p\geqslant 1}\big(\frac{1}{p}(a_{k+p}-a_k)\big)<\infty,\end{aligned}$$

而且再利用式 (8.5)，我們有

$$\lim_{p\to\infty}\mathbb{E}[Y_k^{(p)}]=\gamma.$$

根據習題 1.26 ，我們知道一個在 L^1 中有界的隨機變數序列會有一個 a.s. 收斂的子序列，且 a.s. 收斂蘊含分佈收斂，所以 Cantor 的對角論證法 (diagonal argument) 告訴我們，可以找到子序列 $(p_i)_{i\geqslant 1}$ 使得隨機變數序列 $(Y_k^{(p_i)})_{k\geqslant 1}$ 會分佈收斂，我們將其極限記作 $(Y_k)_{k\geqslant 1}$。這意味著對於任意 $k\geqslant 1$ 及在 $\mathbb{R}^k$ 上有界的可測函數 f，我們會有下列收斂

$$\mathbb{E}[f(Y_1,\dots,Y_k)]=\lim_{i\to\infty}\frac{1}{p_i}\sum_{l=1}^{p_i}\mathbb{E}[f(X_{0,l+1}-X_{0,l},\dots,X_{0,l+k}-X_{0,l+k-1})].$$

由上式我們也可以看出來，$(Y_k)_{k\geqslant 1}$ 是個平穩序列。

現在我們來檢查，我們上面構造出來的序列 $(Y_k)_{k\geqslant 1}$ 滿足式 (8.9)。首先，根據假設 (a) 及 (c)，我們有

$$Y_1^{(p)}=X_{0,U_p+1}-X_{0,U_p}\leqslant X_{U_p,U_p+1}\overset{(d)}{=}X_{0,1},$$

so the sequence $((Y_1^{(p)})^+)_{p\geqslant 1}$ is uniformly integrable. From Fatou's lemma (Lemma 8.4.6), we have

$$\mathbb{E}[Y_1] = \mathbb{E}\left[\lim_{i\to\infty} Y_1^{(p_i)}\right] \geqslant \limsup_{i\to\infty} \mathbb{E}[Y_1^{(p_i)}] = \gamma,$$

Again using Birkhoff's ergodic theorem (Theorem 8.2.1), we conclude that the limit

$$Y := \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} Y_k, \quad \text{a.s.}$$

exists and $\mathbb{E}[Y] = \mathbb{E}[Y_1] \geqslant \gamma$.

In the end, we still need to check that the sequence $(X_{0,1}, X_{0,2}, \dots, X_{0,n})$ stochastically dominates $(Y_1, Y_1 + Y_2, \dots, Y_1 + \cdots + Y_n)$. Given $n \geqslant 1$ and a non-decreasing and non-negative measurable function f on $\mathbb{R}^n$, we have

$$\begin{aligned}
&\mathbb{E}[f(Y_1, Y_1 + Y_2, \dots, Y_1 + \cdots + Y_n)] \\
&\quad = \lim_{i\to\infty} \frac{1}{p_i}\sum_{l=1}^{p_i} \mathbb{E}[f(X_{0,l+1} - X_{0,l}, X_{0,l+2} - X_{0,l}, \dots, X_{0,l+n} - X_{0,l})] \\
&\quad \leqslant \lim_{i\to\infty} \frac{1}{p_i}\sum_{l=1}^{p_i} \mathbb{E}[f(X_{l,l+1}, X_{l,l+2}, \dots, X_{l,l+n})] \\
&\quad = \mathbb{E}[f(X_{0,1}, X_{0,2}, \dots, X_{0,n})].
\end{aligned}$$

Finally, let us prove (iv). From the assumption (a), we have

$$(X_{0,n} - nX_\infty)^+ \leqslant \sum_{k=0}^{n-1} (X_{k,k+1} - X_\infty)^+,$$

and since the assumption (c) says that the terms in the sequence $((X_{k,k+1} - X_\infty)^+)_{k\geqslant 0}$ of random variables are identically distributed, $((\frac{X_{0,n}}{n} - X_\infty)^+)_{n\geqslant 0}$ is an uniformly integrable sequence. Using the fact that X_∞ is the a.s. limit of $\frac{X_{0,n}}{n}$, we obtain

$$\mathbb{E}\left[\left(\frac{X_{0,n}}{n} - X_\infty\right)^+\right] \xrightarrow[n\to\infty]{} 0.$$

Since we also have $\mathbb{E}[\frac{X_{0,n}}{n}] \to \mathbb{E}[X_\infty]$, using the identity $|x| = 2x^+ - x$, we find

$$\mathbb{E}\left[\left|\frac{X_{0,n}}{n} - X_\infty\right|\right] \xrightarrow[n\to\infty]{} 0.$$

□

因此序列 $((Y_1^{(p)})^+)_{p\geqslant 1}$ 是均勻可積的。根據 Fatou 引理（引理 8.4.6），我們有

$$\mathbb{E}[Y_1] = \mathbb{E}\left[\lim_{i\to\infty} Y_1^{(p_i)}\right] \geqslant \limsup_{i\to\infty} \mathbb{E}[Y_1^{(p_i)}] = \gamma,$$

再根據 Birkhoff 遍歷定理（定理 8.2.1），我們得到

$$Y := \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} Y_k, \quad \text{a.s.}$$

存在且 $\mathbb{E}[Y] = \mathbb{E}[Y_1] \geqslant \gamma$。

最後，我們來證明隨機變數序列 $(X_{0,1}, X_{0,2}, \dots, X_{0,n})$ 對於 $(Y_1, Y_1 + Y_2, \dots, Y_1 + \cdots + Y_n)$ 具有隨機優勢。給定 $n \geqslant 1$ 以及在 $\mathbb{R}^n$ 上的非遞減非負可測函數 f，我們有

$$\begin{aligned}
&\mathbb{E}[f(Y_1, Y_1 + Y_2, \dots, Y_1 + \cdots + Y_n)] \\
&\quad = \lim_{i\to\infty} \frac{1}{p_i}\sum_{l=1}^{p_i} \mathbb{E}[f(X_{0,l+1} - X_{0,l}, X_{0,l+2} - X_{0,l}, \dots, X_{0,l+n} - X_{0,l})] \\
&\quad \leqslant \lim_{i\to\infty} \frac{1}{p_i}\sum_{l=1}^{p_i} \mathbb{E}[f(X_{l,l+1}, X_{l,l+2}, \dots, X_{l,l+n})] \\
&\quad = \mathbb{E}[f(X_{0,1}, X_{0,2}, \dots, X_{0,n})].
\end{aligned}$$

最後我們來證明 (iv)。根據假設 (a)，我們有

$$(X_{0,n} - nX_\infty)^+ \leqslant \sum_{k=0}^{n-1} (X_{k,k+1} - X_\infty)^+,$$

其中根據假設 (c)，隨機變數序列 $((X_{k,k+1} - X_\infty)^+)_{k\geqslant 0}$ 中的各項皆有著相同的分佈，因此 $((\frac{X_{0,n}}{n} - X_\infty)^+)_{n\geqslant 0}$ 是個均勻可積的序列。因此利用 X_∞ 為 $\frac{X_{0,n}}{n}$ 的 a.s. 極限，我們得到

$$\mathbb{E}\left[\left(\frac{X_{0,n}}{n} - X_\infty\right)^+\right] \xrightarrow[n\to\infty]{} 0.$$

由於我們也有 $\mathbb{E}[\frac{X_{0,n}}{n}] \to \mathbb{E}[X_\infty]$，使用關係式 $|x| = 2x^+ - x$，我們得到

$$\mathbb{E}\left[\left|\frac{X_{0,n}}{n} - X_\infty\right|\right] \xrightarrow[n\to\infty]{} 0.$$

□