7 Discrete Time Markov chains

7.1 Definitions and Basic Properties

In this section, we discuss the discrete time Markov chains. These are special stochastic processes satisfying the property that at a given time, the future evolution depends only on the state at the *current time* instead of the past of the whole process.

We take E to be a finite or countable set equipped with the σ-algebra $\mathcal{P}(E)$.

Definition 7.1.1 : A function $Q : E \times E \longrightarrow \mathbb{R}$ is called *transition matrix* (轉移矩陣) or *stochastic matrix* (隨機矩陣) on E if it satisfies

(1) $0 \leqslant Q(x, y) \leqslant 1$ for all $x, y \in E$;

(2) $\sum_{y \in E} Q(x, y) = 1$ for all $x \in E$.

Remark 7.1.2 : The notion of transition matrix on E is equivalent to the notion of transition probability from E to E. Indeed, given a transition matrix Q on E, we may define $\nu : E \times \mathcal{P}(E) \longrightarrow \mathbb{R}$ by

$$\nu(x, A) = \sum_{y \in A} Q(x, y), \quad \forall x \in E, A \subseteq E.$$

Then, one can easily check that ν is a transition probability from E to E. Conversely, given a transition probability ν, we can obtain a transition matrix on E by letting $Q(x, y) = \nu(x, \{y\})$.

Definition 7.1.3 : We define $Q_n = Q^n$ for all positive integer $n \geqslant 1$. In other words, $Q_1 = Q$ and the following recurrence relation holds for all $n \geqslant 1$,

$$Q_{n+1}(x, y) = \sum_{z \in E} Q_n(x, z) Q(z, y). \tag{7.1}$$

This equality is called *Chapman-Kolmogorov equality* (Chapman-Kolmogorov 等式). We can check that Q_n is still a transition matrix on E. In order to extend the validity of to the case $n = 0$, we may define $Q_0(x, y) = \mathbb{1}_{x=y}$, which is a diagonal transition matrix.

第七章 離散時間馬可夫鏈

第一節 定義及基本性質

在此章節中，我們要討論的是離散時間的馬可夫鏈。這是個滿足下列性質的隨機過程：固定時間點時，未來的演變只取決於此刻的狀態，而不是過去所有的狀態。

我們考慮有限或可數空間 E，以及相對應的 σ 代數 $\mathcal{P}(E)$。

定義 7.1.1 ： 若 $Q : E \times E \longrightarrow \mathbb{R}$ 滿足

(1) 對於所有 $x, y \in E$，$0 \leqslant Q(x, y) \leqslant 1$；

(2) 對於所有 $x \in E$，我們有 $\sum_{y \in E} Q(x, y) = 1$。

則稱之為在 E 上的<u>轉移矩陣</u> (transition matrix) 或<u>隨機矩陣</u> (stochastic matrix) 。

註解 7.1.2 ： 在 E 上轉移矩陣的概念與由 E 至 E 的轉移機率等價。給定轉移矩陣 Q，我們可以將 $\nu : E \times \mathcal{P}(E) \longrightarrow \mathbb{R}$ 定義做

$$\nu(x, A) = \sum_{y \in A} Q(x, y), \quad \forall x \in E, A \subseteq E,$$

則我們可以驗證 ν 的確是個由 E 至 E 的轉移機率；反之，給定轉移機率 ν，我們可以由 $Q(x, y) = \nu(x, \{y\})$ 得到在 E 上的轉移矩陣 Q。

定義 7.1.3 ： 對於所有正整數 $n \geqslant 1$，我們定義 $Q_n = Q^n$；也就是說，$Q_1 = Q$ 且對於所有 $n \geqslant 1$，我們有下列遞迴關係：

$$Q_{n+1}(x, y) = \sum_{z \in E} Q_n(x, z) Q(z, y). \tag{7.1}$$

此等式稱作<u>*Chapman-Kolmogorov* 等式</u> (Chapman-Kolmogorov equality) 。我們可以檢查 Q_n 仍然是在 E 上的轉移矩陣。此外，我們可以定義 $Q_0(x, y) = \mathbb{1}_{x=y}$，這是個對角的轉移矩陣，這可以讓我們把推廣到 $n = 0$ 的情況。

Definition 7.1.4 : For any function $f : E \longrightarrow \mathbb{R}_{\geqslant 0}$, we define $Qf : E \longrightarrow \mathbb{R}_{\geqslant 0}$ to be

$$Qf(x) = \sum_{y \in E} Q(x,y)f(y), \quad \forall x \in E. \tag{7.2}$$

Remark 7.1.5 :

(1) The summation in (7.2) is always well defined, being a countable summation of non-negative terms.

(2) If we see f as a column vector, then the way the function Qf is defined is exactly through a matrix multiplication.

Definition 7.1.6 : Let Q be a transition matrix on E and $(X_n)_{n \geqslant 0}$ be a stochastic process with values in E. We say that $(X_n)_{n \geqslant 0}$ is a *Markov chain* (馬可夫鏈) with transition matrix Q if for all $n \geqslant 0$, the conditional distribution of X_{n+1} knowing $(X_0, X_1, \dots, X_n)$ is $Q(X_n, \cdot)$. Since E is a discrete space, this condition is equivalent to the follows,

$$\mathbb{P}(X_{n+1} = y \mid X_0 = x_0, \dots, X_n = x_n) = Q(x_n, y),$$

for all $x_0, x_1, \dots, x_n, y \in E$ with $\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) > 0$.

Remark 7.1.7 :

(1) In general, the conditional distribution of X_{n+1} knowing $(X_0, X_1, \dots, X_n)$ does not only depend on X_n, but on $X_0, X_1, \dots, X_n$. In the case of a Markov chain, the fact that the above conditional distribution only depends on X_n is called *Markov property* (馬可夫性質). This should be understood as follows: knowing the whole past $(X_0, X_1, \dots, X_n)$ does not give more information than knowing only the current state X_n.

(2) The conditional distribution $Q(x, \cdot)$ mentioned above does not depend on the time n. In this case, we say that the Markov chain is *homogeneous* (均勻). We may also consider a transition matrix that evolve with time n.

Proposition 7.1.8 : *Given a stochastic process $(X_n)_{n \geqslant 0}$ with values in E. Then, $(X_n)_{n \geqslant 0}$ is a Markov chain with transition matrix Q if and only if for all $n \geqslant 0$ and $x_0, x_1, \dots, x_n \in E$, the following holds,*

$$\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_0 = x_0) Q(x_0, x_1) \dots Q(x_{n-1}, x_n). \tag{7.3}$$

Besides, when $\mathbb{P}(X_0 = x_0) > 0$, we have

$$\mathbb{P}(X_n = x_n \mid X_0 = x_0) = Q_n(x_0, x_n). \tag{7.4}$$

定義 7.1.4 ： 對於任意函數 $f : E \longrightarrow \mathbb{R}_{\geqslant 0}$，我們定義 $Qf : E \longrightarrow \mathbb{R}_{\geqslant 0}$ 為

$$Qf(x) = \sum_{y \in E} Q(x,y)f(y), \quad \forall x \in E. \tag{7.2}$$

註解 7.1.5 ：

(1) (7.2)中的和永遠是定義良好的，因為他是可數多個非負項的和。

(2) 如果我們將 f 視為一個行向量，則 Qf 可以看作透過矩陣乘法所定義出來的值。

定義 7.1.6 ： 令 Q 為在 E 上的轉移矩陣，$(X_n)_{n \geqslant 0}$ 為值域為 E 的隨機過程。若對於所有 $n \geqslant 0$，X_{n+1} 在已知 $(X_0, X_1, \dots, X_n)$ 的條件分佈為 $Q(X_n, \cdot)$，則我們說 $(X_n)_{n \geqslant 0}$ 是個轉移矩陣為 Q 的馬可夫鏈 (Markov chain) 。由於 E 是個離散空間，此條件與下列等價：對於所有 $x_0, x_1, \dots, x_n, y \in E$，若 $\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) > 0$，則我們也會有

$$\mathbb{P}(X_{n+1} = y \mid X_0 = x_0, \dots, X_n = x_n) = Q(x_n, y).$$

註解 7.1.7 ：

(1) 一般來說，X_{n+1} 對於 $(X_0, X_1, \dots, X_n)$ 的條件分佈會不會只取決於 X_n，而是會取決於所有的 $X_0, X_1, \dots, X_n$。在馬可夫鏈的情況中，此條件分佈只取決於 X_n，這樣的性質稱作馬可夫性質 (Markov property)。此概念應該被理解為：當我們知道完整的過去 $(X_0, X_1, \dots, X_n)$ 時，並不會比只知道此刻時間的狀態 X_n 所給出的資訊還多。

(2) 上述所描述的條件分佈 $Q(x, \cdot)$ 並不取決於時間 n，這種情況下，我們討論的是個均勻 (homogeneous) 的馬可夫鏈；我們當然也可以考慮會隨著時間 n 改變的轉移矩陣 Q。

命題 7.1.8 ： 給定值域為 E 的隨機過程 $(X_n)_{n \geqslant 0}$。若且唯若對於所有 $n \geqslant 0$ 及 $x_0, x_1, \dots, x_n \in E$，下列等式成立：

$$\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_0 = x_0) Q(x_0, x_1) \dots Q(x_{n-1}, x_n), \tag{7.3}$$

則 $(X_n)_{n \geqslant 0}$ 是個轉移矩陣為 Q 的馬可夫鏈。此外，當 $\mathbb{P}(X_0 = x_0) > 0$ 時，我們有

$$\mathbb{P}(X_n = x_n \mid X_0 = x_0) = Q_n(x_0, x_n). \tag{7.4}$$

Proof : Let $(X_n)_{n\geqslant 0}$ be a Markov chain with transition matrix Q. Then, we have

$$\begin{aligned}&\mathbb{P}(X_0 = x_0, \ldots, X_{n+1} = x_{n+1})\\ &= \mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) \times \mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n)\\ &= \mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) \times Q(x_n, x_{n+1}).\end{aligned}$$

Hence by induction, we obtain Eq. (7.3). Conversely, if Eq. (7.3) holds, we can check

$$\mathbb{P}(X_{n+1} = y \mid X_0 = x_0, \ldots, X_n = x_n) = \frac{\mathbb{P}(X_{n+1} = y, X_0 = x_0, \ldots, X_n = x_n)}{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n)} = Q(x_n, y).$$

Finally, the marginal distribution Eq. (7.4) can be obtained from the following,

$$Q_n(x_0, x_n) = \sum_{x_1,\ldots,x_{n-1}\in E} Q(x_0, x_1)\ldots Q(x_{n-1}, x_n).$$

□

Remark 7.1.9 : We know from Eq. (7.3) that the initial condition, or the distribution of X_0, along with the transition matrix Q, determines the distribution of the whole Markov chain $(X_n)_{n\geqslant 0}$.

The following proposition contains different properties of a Markov chain.

Proposition 7.1.10 : *Let $(X_n)_{n\geqslant 0}$ be a Markov chain with transition matrix Q.*

(1) For all $n \geqslant 0$ and non-negative measurable function $f : E \longrightarrow \mathbb{R}_{\geqslant 0}$, we have

$$\mathbb{E}[f(X_{n+1}) \mid X_0, X_1, \ldots, X_n] = \mathbb{E}[f(X_{n+1}) \mid X_n] = Qf(X_n).$$

More generally speaking, for any finite subset $\{i_1, \ldots, i_k\}$ of $\{0, \ldots, n-1\}$, we have

$$\mathbb{E}[f(X_{n+1}) \mid X_{i_1}, \ldots, X_{i_k}, X_n] = Qf(X_n).$$

(2) For any $n \geqslant 0$, $p \geqslant 1$ and $y_1, \ldots y_p \in E$, we have

$$\mathbb{P}(X_{n+1} = y_1, \ldots, X_{n+p} = y_p \mid X_0, \ldots, X_n) = Q(X_n, y_1)Q(y_1, y_2)\ldots Q(y_{p-1}, y_p), \tag{7.5}$$

and in consequence,

$$\mathbb{P}(X_{n+p} = y_p \mid X_n) = Q_p(X_n, y_p). \tag{7.6}$$

If we set $Y_p = X_{n+p}$ for all $p \geqslant 0$, then $(Y_p)_{p\geqslant 0}$ is still a Markov chain with transition matrix Q.

證明： 令 $(X_n)_{n\geqslant 0}$ 是個轉移矩陣為 Q 的馬可夫鏈，我們有：

$$\begin{aligned}&\mathbb{P}(X_0 = x_0, \ldots, X_{n+1} = x_{n+1})\\ &= \mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) \times \mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n)\\ &= \mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) \times Q(x_n, x_{n+1}).\end{aligned}$$

因此根據遞迴，我們可以得到式 (7.3)。反之，若式 (7.3) 成立，則我們可以驗證

$$\mathbb{P}(X_{n+1} = y \mid X_0 = x_0, \ldots, X_n = x_n) = \frac{\mathbb{P}(X_{n+1} = y, X_0 = x_0, \ldots, X_n = x_n)}{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n)} = Q(x_n, y).$$

最後，式 (7.4) 中的邊緣分佈可由下列式子得到：

$$Q_n(x_0, x_n) = \sum_{x_1,\ldots,x_{n-1}\in E} Q(x_0, x_1)\ldots Q(x_{n-1}, x_n).$$

□

註解 7.1.9 ： 從式 (7.3) 可以得到，初始分佈也就是 X_0 的分佈以及轉移矩陣 Q 共同決定了整個馬可夫鏈 $(X_n)_{n\geqslant 0}$ 的分佈。

下列命題把馬可夫鏈的各種性質集結起來。

命題 7.1.10 ： 令 $(X_n)_{n\geqslant 0}$ 為轉移矩陣為 Q 的馬可夫鏈。

(1) 對於所有 $n \geqslant 0$ 以及非負可測函數 $f : E \longrightarrow \mathbb{R}_{\geqslant 0}$，我們有

$$\mathbb{E}[f(X_{n+1}) \mid X_0, X_1, \ldots, X_n] = \mathbb{E}[f(X_{n+1}) \mid X_n] = Qf(X_n).$$

更一般來說，對於 $\{0, \ldots, n-1\}$ 的任意有限子集合 $\{i_1, \ldots, i_k\}$，我們有

$$\mathbb{E}[f(X_{n+1}) \mid X_{i_1}, \ldots, X_{i_k}, X_n] = Qf(X_n).$$

(2) 對任意整數 $n \geqslant 0$ 及 $p \geqslant 1$ 以及 $y_1, \ldots y_p \in E$，我們有

$$\mathbb{P}(X_{n+1} = y_1, \ldots, X_{n+p} = y_p \mid X_0, \ldots, X_n) = Q(X_n, y_1)Q(y_1, y_2)\ldots Q(y_{p-1}, y_p), \tag{7.5}$$

因此

$$\mathbb{P}(X_{n+p} = y_p \mid X_n) = Q_p(X_n, y_p). \tag{7.6}$$

若對於所有 $p \geqslant 0$，設 $Y_p = X_{n+p}$，則 $(Y_p)_{p\geqslant 0}$ 仍是轉移矩陣為 Q 的馬可夫鏈。

Proof :

(1) From the definition, we have

$$\mathbb{E}[f(X_{n+1}) \,|\, X_0, \ldots, X_n] = \sum_{y \in E} Q(X_n, y) f(y) = Qf(X_n).$$

Moreover, if $\{i_1, \ldots, i_k\}$ is a finite subset of $\{0, \ldots, n-1\}$, we have

$$\begin{aligned}\mathbb{E}[f(X_{n+1}) \,|\, X_{i_1}, \ldots, X_{i_k}, X_n] &= \mathbb{E}[\mathbb{E}[f(X_{n+1}) \,|\, X_1, \ldots, X_n] \,|\, X_{i_1}, \ldots, X_{i_k}, X_n] \\ &= \mathbb{E}[Qf(X_n) \,|\, X_{i_1}, \ldots, X_{i_k}, X_n] \\ &= Qf(X_n).\end{aligned}$$

(2) Eq. (7.5) is a direct application of Eq. (7.3). Similarly to the proof of Eq. (7.4), Eq. (7.6) can be obtained by summing all the possible values taken by $y_1, \ldots, y_{p-1}$. To conclude, we obtain the following from Eq. (7.5),

$$\mathbb{P}(Y_0 = y_0, \ldots, Y_n = y_n) = \mathbb{P}(X_n = y_0) Q(y_0, y_1) \ldots Q(y_{p-1}, y_p),$$

and use Proposition 7.1.8.

□

Example 7.1.11 : Below are examples of Markov chains.

(1) (Independent random variables) If $(X_n)_{n \geqslant 0}$ is an i.i.d. sequence of random variables with values in E, let us denote by μ the distribution of any term, then $(X_n)_{n \geqslant 0}$ is a Markov chain with transition matrix

$$Q(x, y) = \mu(y), \quad \forall x, y \in E.$$

(2) (Random walk on $\mathbb{Z}^d$) Let $\eta, \xi_1, \ldots, \xi_n$ be independent random variables with values in $\mathbb{Z}^d$. Assume that $(\xi_n)_{n \geqslant 1}$ is an i.i.d. sequence with distribution μ. For all $n \geqslant 0$, let

$$X_n = \eta + \xi_1 + \cdots + \xi_n.$$

Then, $(X_n)_{n \geqslant 0}$ is a Markov chain with transition matrix

$$Q(x, y) = \mu(y - x), \quad \forall x, y \in E.$$

Using the independence of ξ_{n+1} with $(X_0, \ldots, X_n)$, we have

$$\begin{aligned}&\mathbb{P}(X_{n+1} = y \,|\, X_0 = x_0, \ldots, X_n = x_n) \\ &= \mathbb{P}(\xi_{n+1} = y - x_n \,|\, X_0 = x_0, \ldots, X_n = x_n) \\ &= \mathbb{P}(\xi_{n+1} = y - x_n) = \mu(y - x_n).\end{aligned}$$

Let $(e_1, \ldots, e_d)$ be the canonical basis (正則基底) of $\mathbb{R}^d$. The Markov chain $(X_n)_{n \geqslant 0}$ is called *symmetric simple random walk* (對稱簡單隨機漫步) on $\mathbb{Z}^d$ if μ satisfies

$$\mu(e_i) = \mu(-e_i) = \frac{1}{2d}, \quad \forall i \in \{1, \ldots, d\}.$$

證明：

(1) 根據定義，我們有

$$\mathbb{E}[f(X_{n+1}) \,|\, X_0, \ldots, X_n] = \sum_{y \in E} Q(X_n, y) f(y) = Qf(X_n).$$

此外，若 $\{i_1, \ldots, i_k\}$ 是個 $\{0, \ldots, n-1\}$ 的有限子集合，我們有

$$\begin{aligned}\mathbb{E}[f(X_{n+1}) \,|\, X_{i_1}, \ldots, X_{i_k}, X_n] &= \mathbb{E}[\mathbb{E}[f(X_{n+1}) \,|\, X_1, \ldots, X_n] \,|\, X_{i_1}, \ldots, X_{i_k}, X_n] \\ &= \mathbb{E}[Qf(X_n) \,|\, X_{i_1}, \ldots, X_{i_k}, X_n] \\ &= Qf(X_n).\end{aligned}$$

(2) 式 (7.5) 是式 (7.3) 的直接應用，且如同式 (7.4) 的證明，式 (7.6) 可由對所有可能的 $y_1, \ldots, y_{p-1}$ 求和而得到。最後，根據式 (7.5)，我們可以得到

$$\mathbb{P}(Y_0 = y_0, \ldots, Y_n = y_n) = \mathbb{P}(X_n = y_0) Q(y_0, y_1) \ldots Q(y_{p-1}, y_p),$$

並利用命題 7.1.8 總結。

□

範例 7.1.11 ： 以下是馬可夫鏈的例子。

(1) 【獨立隨機變數】若 $(X_n)_{n \geqslant 0}$ 是個值域為 E 的 i.i.d. 隨機變數序列，將任意項的分佈記作 μ，則 $(X_n)_{n \geqslant 0}$ 是個轉移矩陣為

$$Q(x, y) = \mu(y), \quad \forall x, y \in E$$

的馬可夫鏈。

(2) 【$\mathbb{Z}^d$ 上的隨機漫步】令 $\eta, \xi_1, \ldots, \xi_n$ 為取值在 $\mathbb{Z}^d$ 上的獨立隨機變數，並假設 $(\xi_n)_{n \geqslant 1}$ 為分佈為 μ 的 i.i.d. 序列。對於所有 $n \geqslant 0$，設

$$X_n = \eta + \xi_1 + \cdots + \xi_n,$$

則 $(X_n)_{n \geqslant 0}$ 是個轉移矩陣為

$$Q(x, y) = \mu(y - x), \quad \forall x, y \in E$$

的馬可夫鏈。利用 ξ_{n+1} 與 $(X_0, \ldots, X_n)$ 的獨立性，我們有

$$\begin{aligned}&\mathbb{P}(X_{n+1} = y \,|\, X_0 = x_0, \ldots, X_n = x_n) \\ &= \mathbb{P}(\xi_{n+1} = y - x_n \,|\, X_0 = x_0, \ldots, X_n = x_n) \\ &= \mathbb{P}(\xi_{n+1} = y - x_n) = \mu(y - x_n).\end{aligned}$$

(3) (Simple random walk on graphs) Let E be any set, $\mathcal{P}_2(E)$ be the set consisting of all the pairwise disjoint elements from E and F be a subset of $\mathcal{P}_2(E)$. For all $x \in E$, set

$$F_x = \{y \in E : \{x, y\} \in F\}.$$

We assume that the set F_x is not empty and is finite for all $x \in E$. Let us define Q to be the transition matrix on E,

$$\forall x, y \in E, \quad Q(x, y) = \begin{cases} |F_x|^{-1} & \text{if } \{x, y\} \in F, \\ 0 & \text{otherwise.} \end{cases}$$

A Markov chain with transition matrix Q is called *simple random walk* (簡單隨機漫步) on the graph (E, F).

(4) (Branching process) We recall the branching process discussed in Example 6.3.8 and Exercise 6.24.

Let μ be a probability distribution on non-negative integers. Assume that μ is integrable and denote its expectation $m < \infty$. We exclude the special cases where $\mu = \delta_0$ or δ_1. Consider an i.i.d. sequence $(\xi_{n,j})_{n,j \geqslant 0}$ of random variables with distribution μ. Fix an integer $\ell \geqslant 1$, and define another sequence $(X_n)_{n \geqslant 0}$ of random variables by induction,

$$X_0 = \ell,$$
$$X_{n+1} = \sum_{j=1}^{X_n} \xi_{n,j}, \quad \forall n \geqslant 0.$$

The random process $(X_n)_{n \geqslant 0}$ can be understood as the evolution of a parthenogenetic population where μ is the distribution of the offspring of each family member, and the population at the n-th generation is given by X_n. In this case, $(X_n)_{n \geqslant 0}$ is a Markov chain on $E = \mathbb{N}_0$ with transition matrix

$$Q(x, y) = \mu^{*x}(y), \quad \forall x, y \in \mathbb{N}_0, \tag{7.7}$$

where μ^{*x} is the x-fold convolution (捲積) of μ with itself, which can also be interpreted as the distribution of the sum of x i.i.d. random variables with distribution μ. Eq. (7.7) is a consequence of the below computation,

$$\mathbb{P}(X_{n+1} = y \,|\, X_0 = x_0, \ldots, X_n = x_n)$$
$$\mathbb{P}\Big(\sum_{j=1}^{x_n} \xi_{n,j} = y \,\Big|\, X_0 = x_0, \ldots, X_n = x_n\Big)$$
$$\mathbb{P}\Big(\sum_{j=1}^{x_n} \xi_{n,j} = y\Big) = \mu^{*x_n}(y),$$

where in the second equality, we use the property that $(\xi_{n,j})_{j \geqslant 1}$ and $X_0, \ldots, X_n$ are independent.

令 $(e_1, \ldots, e_d)$ 為 $\mathbb{R}^d$ 上的正則基底 (canonical basis)，當 μ 滿足

$$\mu(e_i) = \mu(-e_i) = \frac{1}{2d}, \quad \forall i \in \{1, \ldots, d\},$$

我們將馬可夫鏈 $(X_n)_{n \geqslant 0}$ 稱做在 $\mathbb{Z}^d$ 上的對稱簡單隨機漫步 (symmetric simple random walk)。

(3) 【圖上的簡單隨機漫步】令 E 為任意集合，$\mathcal{P}_2(E)$ 為由所有 E 中兩個不相等的元素構成的集合，F 為 $\mathcal{P}_2(E)$ 的子集合。對於所有 $x \in E$，我們設

$$F_x = \{y \in E : \{x, y\} \in F\}.$$

假設對於所有 $x \in E$，集合 F_x 非空且有限，我們定義 E 上的轉移矩陣 Q 為

$$\forall x, y \in E, \quad Q(x, y) = \begin{cases} |F_x|^{-1} & \text{若 } \{x, y\} \in F, \\ 0 & \text{其他情況.} \end{cases}$$

我們稱轉移矩陣為 Q 的馬可夫鏈為在圖 (E, F) 上的簡單隨機漫步 (simple random walk)。

(4) 【分支過程】我們回顧在範例 6.3.8 以及習題 6.24 中討論過的分支過程。

令 μ 為在非負整數上的機率分佈，假設 μ 可積，並將其期望值記作 $m < \infty$。我們不考慮 $\mu = \delta_0$ 或 δ_1 的特殊情況。考慮 i.i.d. 隨機變數序列 $(\xi_{n,j})_{n,j \geqslant 0}$ 且各項分佈皆由 μ 所給定。固定正整數 $\ell \geqslant 1$，並以遞迴方式來定義隨機變數序列 $(X_n)_{n \geqslant 0}$：

$$X_0 = \ell,$$
$$X_{n+1} = \sum_{j=1}^{X_n} \xi_{n,j}, \quad \forall n \geqslant 0.$$

我們可以將 $(X_n)_{n \geqslant 0}$ 視為描述一個（單性生殖）家族人口數量的隨機過程：μ 為家族成員中，每個成員後代數量的機率分佈，則第 n 代的人口數量為 X_n。在這樣的情況下，$(X_n)_{n \geqslant 0}$ 是個在 $E = \mathbb{N}_0$ 上的馬可夫鏈，其轉移矩陣寫作

$$Q(x, y) = \mu^{*x}(y), \quad \forall x, y \in \mathbb{N}_0, \tag{7.7}$$

其中 μ^{*x} 是 μ 與自己取 x 次的捲積 (convolution)，也可以理解做 x 個 i.i.d. 有著與 μ 相同分佈的隨機變數和。式 (7.7) 是下列計算的結果

$$\mathbb{P}(X_{n+1} = y \,|\, X_0 = x_0, \ldots, X_n = x_n)$$
$$\mathbb{P}\Big(\sum_{j=1}^{x_n} \xi_{n,j} = y \,\Big|\, X_0 = x_0, \ldots, X_n = x_n\Big)$$
$$\mathbb{P}\Big(\sum_{j=1}^{x_n} \xi_{n,j} = y\Big) = \mu^{*x_n}(y),$$

7.2 Canonical Markov Chain

In Definition 7.1.6, we gave the property that a Markov chain needs to satisfy. In the below proposition, we explain, given a transition matrix, how to construct a corresponding Markov chain. Later in Theorem 7.2.3, we will see the uniqueness of the distribution of such a Markov chain. Thus, we can say *the* Markov chain instead of *a* Markov chain when the transition matrix is given.

Proposition 7.2.1 : *Let Q be a transition matrix on E. We can find a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ such that for any $x \in E$, we can construct a Markov chain $(X_n^x)_{n\geqslant 0}$ with transition matrix Q and initial state $X_0^x = x$.*

Proof : Take $\Omega' = [0,1)$, $\mathcal{F}' = \mathcal{B}(\Omega')$ and $\mathbb{P}'$ to be the Lebesgue measure. For any real number $\omega \in [0,1)$, we have the dyadic expansion,

$$\omega = \sum_{n=0}^{\infty} \varepsilon_n(\omega) 2^{-n-1}, \quad \varepsilon_n(\omega) \in \{0,1\}.$$

The above provides us with an i.i.d. sequence $(\varepsilon_n)_{n\geqslant 0}$ of random variables such that $\mathbb{P}(\varepsilon_n = 1) = \mathbb{P}(\varepsilon_n = 0) = \frac{1}{2}$. If $\varphi : \mathbb{N}_0 \times \mathbb{N}_0 \longrightarrow \mathbb{N}_0$ is an injective function, define $\eta_{i,j} = \varepsilon_{\varphi(i,j)}$ for all $i, j \in \mathbb{N}_0$, then $(\eta_{i,j})_{i,j\geqslant 0}$ is still an i.i.d. sequence of random variables. Let

$$\forall i \geqslant 0, \quad U_i = \sum_{j\geqslant 0} \eta_{i,j} 2^{-j-1}.$$

Then, $(U_i)_{i\geqslant 0}$ is an i.i.d. sequence of random variables with the uniform distribution on $[0,1]$.

Since E is countable, we can enumerate the elements by $y_1, y_2, \ldots$. Given $x \in E$, let $X_0^x = x$ and we define the value of X_1^x using the uniform random variable U_0,

$$X_1^x = y_k \quad \text{if} \quad \sum_{1\leqslant j\leqslant k-1} Q(x, y_j) < U_0 \leqslant \sum_{1\leqslant j\leqslant k} Q(x, y_j).$$

We can easily check that we do have $\mathbb{P}(X_1^x = y) = Q(x,y)$ for all $y \in E$. Then, similarly to the definition of X_1^x, we define the Markov chain $(X_n^x)_{n\geqslant 0}$ by induction: for all $n \geqslant 0$, define

$$X_{n+1}^x = y_k \quad \text{if} \quad \sum_{1\leqslant j\leqslant k-1} Q(X_n^x, y_j) < U_n \leqslant \sum_{1\leqslant j\leqslant k} Q(X_n^x, y_j).$$

其中在第二個等式中，我們使用 $(\xi_{n,j})_{j\geqslant 1}$ 與 $X_0, \ldots, X_n$ 獨立的性質。

第二節 正則馬可夫鏈

在定義 7.1.6 中，我們給出馬可夫鏈需要滿足的性質，在下列命題中，則告訴我們在給定轉移矩陣時，如何構造出相對應的馬可夫鏈。稍後在定理 7.2.3 中，我們會看到這樣的馬可夫鏈，其分佈是唯一的。因此在後續章節中，當我們有轉移矩陣時，我們可以以定冠詞 the 來描述（正則）馬可夫鏈。

命題 7.2.1 ： 令 Q 為在 E 上的轉移矩陣。我們可以找到機率空間 $(\Omega', \mathcal{F}', \mathbb{P}')$ 使得對於任意 $x \in E$，我們可以構造轉移機率為 Q，起始位置為 $X_0^x = x$ 的馬可夫鏈 $(X_n^x)_{n\geqslant 0}$。

證明： 取 $\Omega' = [0,1)$，$\mathcal{F}' = \mathcal{B}(\Omega')$ 以及 $\mathbb{P}'$ 為勒貝格測度。對於任意實數 $\omega \in [0,1)$，我們有二元展開式：

$$\omega = \sum_{n=0}^{\infty} \varepsilon_n(\omega) 2^{-n-1}, \quad \varepsilon_n(\omega) \in \{0,1\}.$$

由此我們得到一個 i.i.d. 隨機變數序列 $(\varepsilon_n)_{n\geqslant 0}$ 使得 $\mathbb{P}(\varepsilon_n = 1) = \mathbb{P}(\varepsilon_n = 0) = \frac{1}{2}$。若 $\varphi : \mathbb{N}_0 \times \mathbb{N}_0 \longrightarrow \mathbb{N}_0$ 是個單射函數，則對於 $i, j \in \mathbb{N}_0$，定義 $\eta_{i,j} = \varepsilon_{\varphi(i,j)}$，則 $(\eta_{i,j})_{i,j\geqslant 0}$ 仍是個 i.i.d. 隨機變數序列。令

$$\forall i \geqslant 0, \quad U_i = \sum_{j\geqslant 0} \eta_{i,j} 2^{-j-1},$$

則 $(U_i)_{i\geqslant 0}$ 是個 i.i.d. 隨機變數序列，且皆為在 $[0,1]$ 之上的均勻分佈。

由於 E 可數，我們將裡面的元素記作 $y_1, y_2, \ldots$。給定 $x \in E$，令 $X_0^x = x$，接著根據均勻隨機變數 U_0 的值來決定 X_1^x 的值：

$$X_1^x = y_k \quad 若 \quad \sum_{1\leqslant j\leqslant k-1} Q(x, y_j) < U_0 \leqslant \sum_{1\leqslant j\leqslant k} Q(x, y_j).$$

我們可以輕易驗證，對於所有 $y \in E$，我們的確有 $\mathbb{P}(X_1^x = y) = Q(x,y)$。我們可以以遞迴的方式來定義整個馬可夫鏈 $(X_n^x)_{n\geqslant 0}$：對於所有 $n \geqslant 0$，定義

$$X_{n+1}^x = y_k \quad 若 \quad \sum_{1\leqslant j\leqslant k-1} Q(X_n^x, y_j) < U_n \leqslant \sum_{1\leqslant j\leqslant k} Q(X_n^x, y_j).$$

Then, using the independence of (U_i), we reach at

$$\begin{aligned}
&\mathbb{P}(X_{n+1}^x = y_k \,|\, X_0^x = x_0, \dots, X_n^x = x_n) \\
&= \mathbb{P}\left(\sum_{1 \leqslant j \leqslant k-1} Q(x_n, y_j) < U_n \leqslant \sum_{1 \leqslant j \leqslant k} Q(x_n, y_j) \,\middle|\, X_0^x = x_0, \dots, X_n^x = x_n \right) \\
&= \mathbb{P}\left(\sum_{1 \leqslant j \leqslant k-1} Q(x_n, y_j) < U_n \leqslant \sum_{1 \leqslant j \leqslant k} Q(x_n, y_j) \right) \\
&= Q(x_n, y_k).
\end{aligned}$$

As a consequence, $(X_n^x)_{n \geqslant 0}$ is a Markov chain with transition matrix Q. □

In what follows, we will choose a canonical space on which the Markov chain is defined. If the state space is E, the sample space is $\Omega = E^{\mathbb{N}_0}$. An element $\omega \in \Omega$ in the sample space is written as $\omega = (\omega_0, \omega_1, \dots)$, which is a sequence in E. For all $n \geqslant 0$, we can define the coordinate function (座標函數) X_n to be

$$X_n : \begin{array}{ccc} \Omega = E^{\mathbb{N}_0} & \to & E \\ \omega = (\omega_0, \omega_1, \dots) & \mapsto & \omega_n \end{array}.$$

On Ω, we consider $\mathcal{F}$ to be the smallest σ-algebra that makes all the coordinate functions X_n measurable, that is

$$\mathcal{F} := \sigma(X_n^{-1}(x) : \forall n \geqslant 0, \forall x \in E).$$

This σ-algebra can also be generated by the following cylindrical events (圓柱事件)

$$\forall n \geqslant 0, \quad \forall x_0, \dots, x_n \in E, \quad C = \{\omega \in \Omega : \omega_0 = x_0, \dots, \omega_n = x_n\}.$$

In the rest of this section, we will define a probability measure on Ω so that we can talk about the uniqueness of the Markov chain in the sense of distribution.

Proposition 7.2.1 above gives the existence of the Markov chain once the transition matrix and the initial state are given. The uniqueness, however, does not hold in general. As discussed in Section 2.1.2, the random variable itself is not unique, what is unique is its *distribution*.

Lemma 7.2.2 : *Let $(G, \mathcal{G})$ be a measurable space and $\psi : (G, \mathcal{G}) \longrightarrow (\Omega, \mathcal{F})$ be a function. Then, the two following properties are equivalent.*

(1) $X_n \circ \psi$ is measurable for all $n \geqslant 0$.

(2) ψ is measurable.

接著利用 (U_i) 的獨立性，我們得到

$$\begin{aligned}
&\mathbb{P}(X_{n+1}^x = y_k \,|\, X_0^x = x_0, \dots, X_n^x = x_n) \\
&= \mathbb{P}\left(\sum_{1 \leqslant j \leqslant k-1} Q(x_n, y_j) < U_n \leqslant \sum_{1 \leqslant j \leqslant k} Q(x_n, y_j) \,\middle|\, X_0^x = x_0, \dots, X_n^x = x_n \right) \\
&= \mathbb{P}\left(\sum_{1 \leqslant j \leqslant k-1} Q(x_n, y_j) < U_n \leqslant \sum_{1 \leqslant j \leqslant k} Q(x_n, y_j) \right) \\
&= Q(x_n, y_k).
\end{aligned}$$

因此 $(X_n^x)_{n \geqslant 0}$ 是個轉移矩陣由 Q 給定的馬可夫鏈。 □

在接續的章節中，我們要選定慣例的機率空間，並在上面定義馬可夫鏈：若狀態空間為 E，樣本空間寫作 $\Omega = E^{\mathbb{N}_0}$。在樣本空間中的元素 $\omega \in \Omega$ 可以寫作 $\omega = (\omega_0, \omega_1, \dots)$，這是個在 E 中的序列。對於所有 $n \geqslant 0$，我們可以定義座標函數 (coordinate function) X_n 如下：

$$X_n : \begin{array}{ccc} \Omega = E^{\mathbb{N}_0} & \to & E \\ \omega = (\omega_0, \omega_1, \dots) & \mapsto & \omega_n \end{array}.$$

在 Ω 上，我們考慮 $\mathcal{F}$ 為使所有座標函數 X_n 皆可測的最小的 σ 代數，也就是說

$$\mathcal{F} := \sigma(X_n^{-1}(x) : \forall n \geqslant 0, \forall x \in E).$$

這個 σ 代數也可以被以下「圓柱事件 (cylindrical events)」所生成：

$$\forall n \geqslant 0, \quad \forall x_0, \dots, x_n \in E, \quad C = \{\omega \in \Omega : \omega_0 = x_0, \dots, \omega_n = x_n\}.$$

在此章節接續的部份，我們要在 Ω 上定義機率測度，進而得到馬可夫鏈的唯一性。

上述命題 7.2.1 給出了在給定轉移矩陣以及起始位置的情況下，馬可夫鏈的存在性，但一般來說，我們不會有唯一性，因為如同在第 2.1.2 小節討論過的，隨機變數本身沒有唯一性，有唯一性的是他的分佈。

引理 7.2.2 ： 令 $(G, \mathcal{G})$ 為可測空間，以及函數 $\psi : (G, \mathcal{G}) \longrightarrow (\Omega, \mathcal{F})$。則下列兩性質等價：

(1) 對於所有 $n \geqslant 0$，$X_n \circ \psi$ 為可測函數。

(2) ψ 為可測函數。

Proof : We only need to prove that (1) ⇒ (2). Let

$$\mathcal{F}' = \{A \in \mathcal{F} : \psi^{-1}(A) \in \mathcal{G}\}.$$

It is not hard to check that $\mathcal{F}'$ is indeed a σ-algebra on Ω. For any fixed $y \in E$, since E is discrete, $\{y\}$ is measurable. By (1), we know that $\psi^{-1}(X_n^{-1}(\{y\})) = (X_n \circ \psi)^{-1}(\{y\}) \in \mathcal{G}$, so $X_n^{-1}(\{y\}) \in \mathcal{F}'$. Therefore, $\mathcal{F}'$ makes all the coordinate functions X_n measurable, giving the equality $\mathcal{F}' = \mathcal{F}$. □

$$\begin{tikzcd}
(G,\mathcal{G}) \arrow[dr, dashed, "\psi"] \arrow[dd, dashed, "X_n\circ\psi"'] & \\
& (\Omega,\mathcal{F}) \arrow[dl, "X_n"] \\
(E,\mathcal{P}(E)) &
\end{tikzcd}
\qquad
\begin{tikzcd}
(\Omega',\mathcal{F}') \arrow[r, "\mathbb{P}'"] \arrow[d, "X_n\circ\Psi"'] \arrow[dr, "\Psi"] & {[0,1]} \\
(E,\mathcal{P}(E)) & (\Omega=E^{\otimes\mathbb{N}_0},\mathcal{F}) \arrow[l, "X_n"] \arrow[u, "\mathbb{P}_x"']
\end{tikzcd}$$

Figure 7.1: **Left**: The coordinate function X_n is measurable for all $n \geqslant 0$. Lemma 7.2.2 states that ψ is measurable if and only if $X_n \circ \psi$ is measurable for all $n \geqslant 0$. **Right**: For the construction of the canonical Markov chain in Theorem 7.2.3, we push forward the measure defined on $(\Omega', \mathcal{F}')$ to the canonical space $(\Omega, \mathcal{F})$.

Theorem 7.2.3 (Canonical Markov chain) **:** *Let Q be a transition matrix on E. Then, for all $x \in E$, there exists a unique probability measure $\mathbb{P}_x$ on $\Omega = E^{\mathbb{N}_0}$ such that the stochastic process defined by the coordinate functions $(X_n)_{n\geqslant 0}$ is a Markov chain with transition matrix Q and $\mathbb{P}_x(X_0 = x) = 1$.*

Proof : Let $x \in E$. Proposition 7.2.1 gives the existence of a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ on which we can construct a Markov chain $(X_n^x)_{n\geqslant 0}$ with transition matrix Q and $X_0^x = x$. From Lemma 7.2.2, the below function is measurable,

$$\begin{aligned} \Psi : \quad \Omega' &\longrightarrow \Omega = E^{\otimes \mathbb{N}_0} \\ \omega' &\longrightarrow (X_n^x(\omega'))_{n\geqslant 0} \end{aligned}.$$

Let us define $\mathbb{P}_x$ to be the image measure of $\mathbb{P}'$ under Ψ, that is $\mathbb{P}_x = \mathbb{P}' \circ \Psi^{-1}$, then we have

$$\mathbb{P}_x(X_0 = x) = \mathbb{P}' \circ \Psi^{-1}(X_0 = x) = \mathbb{P}'(X_0^x = x) = 1.$$

證明：我們只需要證明 (1) ⇒ (2)。令

$$\mathcal{F}' = \{A \in \mathcal{F} : \psi^{-1}(A) \in \mathcal{G}\}.$$

我們不難檢查 $\mathcal{F}'$ 是個在 Ω 上的 σ 代數。對於任意固定的 $y \in E$ 來說，由於 E 是離散的，$\{y\}$ 是個可測集合。根據 (1)，我們知道 $\psi^{-1}(X_n^{-1}(\{y\})) = (X_n \circ \psi)^{-1}(\{y\}) \in \mathcal{G}$，所以 $X_n^{-1}(\{y\}) \in \mathcal{F}'$。因此，$\mathcal{F}'$ 會讓所有的座標函數 X_n 可測，所以給出等式 $\mathcal{F}' = \mathcal{F}$。 □

$$\begin{tikzcd}
(G,\mathcal{G}) \arrow[dr, dashed, "\psi"] \arrow[dd, dashed, "X_n\circ\psi"'] & \\
& (\Omega,\mathcal{F}) \arrow[dl, "X_n"] \\
(E,\mathcal{P}(E)) &
\end{tikzcd}
\qquad
\begin{tikzcd}
(\Omega',\mathcal{F}') \arrow[r, "\mathbb{P}'"] \arrow[d, "X_n\circ\Psi"'] \arrow[dr, "\Psi"] & {[0,1]} \\
(E,\mathcal{P}(E)) & (\Omega=E^{\otimes\mathbb{N}_0},\mathcal{F}) \arrow[l, "X_n"] \arrow[u, "\mathbb{P}_x"']
\end{tikzcd}$$

圖 7.1: **左圖**：對於所有 $n \geqslant 0$，座標函數 X_n 皆是可測的。引理 7.2.2 的敘述告訴我們，若且唯若 ψ 可測，則對於所有 $n \geqslant 0$，合成函數 $X_n \circ \psi$ 皆是可測的。**右圖**：定理 7.2.3 中構造正則馬可夫鏈，我們將在 $(\Omega', \mathcal{F}')$ 機率空間上的測度，推至正則空間 $(\Omega, \mathcal{F})$ 上。

定理 7.2.3 【正則馬可夫鏈】：令 Q 為 E 上的轉移矩陣，則對於所有 $x \in E$，存在唯一在 $\Omega = E^{\mathbb{N}_0}$ 上的機率測度 $\mathbb{P}_x$ 使得在 $\mathbb{P}_x$ 之下，座標函數定義出來的隨機過程 $(X_n)_{n\geqslant 0}$，是個轉移矩陣為 Q 的馬可夫鏈，且滿足 $\mathbb{P}_x(X_0 = x) = 1$。

證明：令 $x \in E$，根據命題 7.2.1 ，存在機率空間 $(\Omega', \mathcal{F}', \mathbb{P}')$ 使得我們可以在上面構造轉移矩陣為 Q 的馬可夫鏈 $(X_n^x)_{n\geqslant 0}$，且滿足 $X_0^x = x$。根據引理 7.2.2 ，下述函數可測：

$$\begin{aligned} \Psi : \quad \Omega' &\longrightarrow \Omega = E^{\otimes \mathbb{N}_0} \\ \omega' &\longrightarrow (X_n^x(\omega'))_{n\geqslant 0} \end{aligned}.$$

我們將 $\mathbb{P}_x$ 定義為 $\mathbb{P}'$ 在 Ψ 之下的影像測度：$\mathbb{P}_x = \mathbb{P}' \circ \Psi^{-1}$，則我們有

$$\mathbb{P}_x(X_0 = x) = \mathbb{P}' \circ \Psi^{-1}(X_0 = x) = \mathbb{P}'(X_0^x = x) = 1.$$

Then, for all $x_0, x_1, \dots, x_n \in E$, we also have

$$\begin{aligned}\mathbb{P}_x(X_0 = x_0, \dots, X_n = x_n) &= \mathbb{P}'(X_0^x = x_0, \dots, X_n^x = x_n) \\ &= \mathbb{P}'(X_0^x = x_0)Q(x_0, x_1)\dots Q(x_{n-1}, x_n) \\ &= \mathbb{P}_x(X_0 = x_0)Q(x_0, x_1)\dots Q(x_{n-1}, x_n).\end{aligned}$$

Hence, from Proposition 7.1.8, under $\mathbb{P}_x$, $(X_n)_{n\geqslant 0}$ is a Markov chain with transition matrix Q and initial position x.

Then we discuss uniqueness. If $\mathbb{P}_x$ and $\mathbb{P}'_x$ both satisfy the properties required by the theorem, then they take the same value on all cylindrical events. Since the set of cylindrical events is closed under finite intersections and generate the σ-algebra $\mathcal{F}$, the monotone class lemma tells us that $\mathbb{P}_x = \mathbb{P}'_x$ (see Corollary 1.1.19). □

Remark 7.2.4 :

(1) From Proposition 7.1.8, we know that we have, for all $n \geqslant 0$ and $x, y \in E$, that

$$\mathbb{P}_x(X_n = y) = Q_n(x, y).$$

(2) If μ is a probability measure on E, denote

$$\mathbb{P}_\mu = \sum_{x\in E} \mu(x)\mathbb{P}_x,$$

which is a probability measure on Ω. We can check that under $\mathbb{P}_\mu$, the process $(X_n)_{n\geqslant 0}$ of coordinate functions is a Markov chain with transition matrix Q and X_0 has distribution μ.

(3) On the probability space $(\Omega', \mathcal{F}', \mathbb{P}')$, if $(X'_n)_{n\geqslant 0}$ is a Markov chain with transition matrix Q and initial distribution μ, then for any measurable set $B \subseteq \Omega = E^{\mathbb{N}_0}$, we have

$$\mathbb{P}'((X'_n)_{n\geqslant 0} \in B) = \mathbb{P}_\mu(B).$$

In the proof of Theorem 7.2.3, we know that the above holds for all cylindrical events, hence for all the measurable events from the monotone class lemma. This equality tells us that any statement that holds for the canonical Markov chain will also be valid for any Markov chain with the same transition matrix, and the same initial condition.

Next, we will discuss the *Markov properties* (馬可夫性質), and some more notations need to be introduced.

Definition 7.2.5 : First, we define the translation operator (平移算子). For all non-negative integer $k \geqslant 0$, define $\theta_k : \Omega \longrightarrow \Omega$ to be

$$\theta_k((\omega_n)_{n\geqslant 0}) = (\omega_{k+n})_{n\geqslant 0}.$$

From Lemma 7.2.2, we know that these operators are measurable.

此外，對於所有 $x_0, x_1, \dots, x_n \in E$，我們也有

$$\begin{aligned}\mathbb{P}_x(X_0 = x_0, \dots, X_n = x_n) &= \mathbb{P}'(X_0^x = x_0, \dots, X_n^x = x_n) \\ &= \mathbb{P}'(X_0^x = x_0)Q(x_0, x_1)\dots Q(x_{n-1}, x_n) \\ &= \mathbb{P}_x(X_0 = x_0)Q(x_0, x_1)\dots Q(x_{n-1}, x_n).\end{aligned}$$

因此根據命題 7.1.8 ，在 $\mathbb{P}_x$ 之下，$(X_n)_{n\geqslant 0}$ 是個轉移矩陣為 Q 且起始位置為 x 的馬可夫鏈。

我們來討論唯一性：若 $\mathbb{P}_x$ 及 $\mathbb{P}'_x$ 皆滿足此定理敘述中所要求的性質，則他們在圓柱事件上取值相同。由於圓柱事件構成的集合在有限交集下是封閉的，且能夠生成 σ 代數 $\mathcal{F}$，單調類引理告訴我們 $\mathbb{P}_x = \mathbb{P}'_x$（參考系理 1.1.19）。 □

註解 7.2.4 ：

(1) 根據命題 7.1.8 ，我們知道對於所有 $n \geqslant 0$ 以及 $x, y \in E$，我們有

$$\mathbb{P}_x(X_n = y) = Q_n(x, y).$$

(2) 若 μ 是個在 E 上的機率測度，我們記

$$\mathbb{P}_\mu = \sum_{x\in E} \mu(x)\mathbb{P}_x,$$

這會是個在 Ω 上的機率測度。我們可以檢驗在 $\mathbb{P}_\mu$ 之下，座標函數 $(X_n)_{n\geqslant 0}$ 是個轉移矩陣為 Q 的馬可夫鏈，且 X_0 與 μ 有著相同的分佈。

(3) 在機率空間 $(\Omega', \mathcal{F}', \mathbb{P}')$ 上，若 $(X'_n)_{n\geqslant 0}$ 是個轉移矩陣為 Q 且初始分佈為 μ 的馬可夫鏈，則對於任意可測集合 $B \subseteq \Omega = E^{\mathbb{N}_0}$，我們有

$$\mathbb{P}'((X'_n)_{n\geqslant 0} \in B) = \mathbb{P}_\mu(B).$$

在定理 7.2.3 的證明中，我們知道上式對於所有圓柱事件皆成立，因此同樣根據單調類引理，對於所有可測集合皆成立。此等式告訴我們，任何對正則馬可夫鏈為真的敘述，對任意有著相同轉移矩陣、以及相同初始條件的馬可夫鏈也會為真。

接著我們要討論馬可夫性質 (Markov properties) ，在此之前我們還需要引進一些新的記號。

定義 7.2.5 ： 首先，我們定義平移算子 (translation operator)。對於所有非負整數 $k \geqslant 0$，定義 $\theta_k : \Omega \longrightarrow \Omega$ 為

$$\theta_k((\omega_n)_{n\geqslant 0}) = (\omega_{k+n})_{n\geqslant 0}.$$

從引理 7.2.2 我們知道這些算子是可測的。

Also, let us write $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ for all $n \geqslant 0$, meaning that $(\mathcal{F}_n)_{n\geqslant 0}$ is the canonical filtration (正則濾鏈) of $(X_n)_{n\geqslant 0}$. Besides, we write $\mathbb{E}_x$ for the expectation under the probability measure $\mathbb{P}_x$.

Theorem 7.2.6 (Simple Markov property) : *Let G be a non-negative measurable function on Ω. Fix $n \geqslant 0$. Then, for any fixed $x \in E$, we have*

$$\mathbb{E}_x[F \cdot G \circ \theta_n] = \mathbb{E}_x[F\ \mathbb{E}_{X_n}[G]], \tag{7.8}$$

for every non-negative $\mathcal{F}_n$-measurable function F. Equivalently, for any fixed $x \in E$, the above can be reformulated as follows,

$$\mathbb{E}_x[G \circ \theta_n \,|\, \mathcal{F}_n] = \mathbb{E}_{X_n}[G], \tag{7.9}$$

which can be interpreted as: the conditional distribution of $\theta_n(\omega)$ knowing $(X_0, \ldots, X_n)$ is $\mathbb{P}_{X_n}$.

Proof : The second part of the theorem is equivalent to the first part, so we only need to show the first part. To prove Eq. (7.8), we only need to check it for all $x_0, x_1, \ldots, x_n \in E$ and F of the form

$$F = \mathbb{1}_{X_0=x_0,\ldots,X_n=x_n}.$$

Similarly, for $p \geqslant 0$ and $y_0, \ldots, y_p \in E$, consider

$$G = \mathbb{1}_{X_0=y_0,\ldots,X_p=y_p}. \tag{7.10}$$

Given $y \in E$, we have

$$\mathbb{E}_y[G] = \mathbb{1}_{y_0=y} Q(y_0, y_1) \cdots Q(y_{p-1}, y_p),$$

and

$$\begin{aligned}\mathbb{E}_x[F \cdot G \circ \theta_n] &= \mathbb{P}_x(X_0 = x_0, \ldots, X_n = x_n, X_n = y_0, X_{n+1} = y_1, \ldots, X_{n+p} = y_p) \\ &= \mathbb{1}_{x_0=x} Q(x_0, x_1) \cdots Q(x_{n-1}, x_n) \mathbb{1}_{y_0=x_n} Q(y_0, y_1) \cdots Q(y_{p-1}, y_p),\end{aligned}$$

so Eq. (7.8) is true for G of the form Eq. (7.10). The monotone class lemma then allows to conclude that Eq. (7.8) holds for all $G = \mathbb{1}_A$ with $A \in \mathcal{F}$. □

The above theorem is a general form of the simple Markov property, meaning that knowing the past $(X_0, \ldots, X_n)$, the distribution $\theta_n(\omega)$ of the Markov chain in the future time only depends on the current state X_n. Below we will talk about the *strong Markov property* (強馬可夫性質) , in which case the "current time" needs not to be a deterministic time n but is given by a random variable T. Later in Section 7.3, we will see some applications of the strong Markov property.

Theorem 7.2.7 (Strong Markog property) : *Let T be a stopping time with respect to the filtration $(\mathcal{F}_n)_{n\geqslant 0}$ and G be a non-negative measurable function defined on Ω. Then, for any fixed $x \in E$, we*

對於所有 $n \geqslant 0$，我們定義 $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$，也就是說 $(\mathcal{F}_n)_{n\geqslant 0}$ 是 $(X_n)_{n\geqslant 0}$ 的正則濾鏈 (canonical filtration)。另外，我們定義 $\mathbb{E}_x$ 為在機率測度 $\mathbb{P}_x$ 之下的期望值。

定理 7.2.6 【簡單馬可夫性質】：令 G 為在 Ω 上的非負可測函數。固定 $n \geqslant 0$。則對於所有 $x \in E$，以及任意非負 $\mathcal{F}_n$ 可測函數 F，我們有

$$\mathbb{E}_x[F \cdot G \circ \theta_n] = \mathbb{E}_x[F\ \mathbb{E}_{X_n}[G]]. \tag{7.8}$$

以等價的方式描述，對於任意的 $x \in E$ 來說，上述也可以重新寫做

$$\mathbb{E}_x[G \circ \theta_n \,|\, \mathcal{F}_n] = \mathbb{E}_{X_n}[G], \tag{7.9}$$

這可以解讀為：$\theta_n(\omega)$ 在已知 $(X_0, \ldots, X_n)$ 的條件分佈為 $\mathbb{P}_{X_n}$。

證明：定理第二部份的敘述與第一部份等價，因此只需要證明第一部份。要證明式 (7.8)，只需要對於下列形式的 F 進行驗證：對於所有 $x_0, x_1, \ldots, x_n \in E$，考慮

$$F = \mathbb{1}_{X_0=x_0,\ldots,X_n=x_n}.$$

同樣的，對於 $p \geqslant 0$，$y_0, \ldots, y_p \in E$，考慮

$$G = \mathbb{1}_{X_0=y_0,\ldots,X_p=y_p}. \tag{7.10}$$

給定 $y \in E$，我們有

$$\mathbb{E}_y[G] = \mathbb{1}_{y_0=y} Q(y_0, y_1) \cdots Q(y_{p-1}, y_p),$$

以及

$$\begin{aligned}\mathbb{E}_x[F \cdot G \circ \theta_n] &= \mathbb{P}_x(X_0 = x_0, \ldots, X_n = x_n, X_n = y_0, X_{n+1} = y_1, \ldots, X_{n+p} = y_p) \\ &= \mathbb{1}_{x_0=x} Q(x_0, x_1) \cdots Q(x_{n-1}, x_n) \mathbb{1}_{y_0=x_n} Q(y_0, y_1) \cdots Q(y_{p-1}, y_p),\end{aligned}$$

因此式 (7.8) 對式 (7.10) 中形式的 G 為真。透過單調類引理，我們知道式 (7.8) 對於所有的 $G = \mathbb{1}_A$，且 $A \in \mathcal{F}$ 也為真，故得證。□

上述定理是簡單馬可夫性質的推廣形式：在知道過去 $(X_0, \ldots, X_n)$ 的情況下，馬可夫鏈在未來時間的分佈 $\theta_n(\omega)$ 只取決於當下的狀態 X_n。下面我們要探討的是強馬可夫性質 (strong Markov property)，也就是說所謂的「當下」，不一定會是個固定時間 n，而會是個由隨機變數 T 所給定的時間。在後面第 7.3 節中，我們會看到強馬可夫性質的應用。

定理 7.2.7 【強馬可夫性質】：令 T 對於濾鏈 $(\mathcal{F}_n)_{n\geqslant 0}$ 的停止時間，F 及 G 為兩個定義在 Ω

have

$$\mathbb{E}_x[\mathbb{1}_{T<\infty} F \cdot G \circ \theta_T] = \mathbb{E}_x[\mathbb{1}_{T<\infty} F \ \mathbb{E}_{X_T}[G]], \tag{7.11}$$

for every non-negative $\mathcal{F}_T$-measurable function F. The above can also be reformulated as below for any fixed $x \in E$,

$$\mathbb{E}_x[\mathbb{1}_{T<\infty} G \circ \theta_T \mid \mathcal{F}_T] = \mathbb{1}_{T<\infty} \mathbb{E}_{X_T}[G]. \tag{7.12}$$

Remark 7.2.8 : We can note that the random variable X_T is defined on a $\mathcal{F}_T$-measurable set $\{T < \infty\}$ and is $\mathcal{F}_T$-measurable. The expectation $\mathbb{E}_{X_T}[G]$ is a random variable which is also defined on $\{T < \infty\}$. It can be seen as the composition between the functions $\omega \mapsto X_T(\omega)$ and $x \mapsto \mathbb{E}_x[G]$.

Proof : Given $n \geqslant 0$, since $\mathbb{1}_{T=n} F$ is $\mathcal{F}_n$-measurable, we can apply the simple Markov property stated in Theorem 7.2.6, giving

$$\mathbb{E}_x[\mathbb{1}_{T=n} F \cdot G \circ \theta_T] = \mathbb{E}_x[\mathbb{1}_{T=n} F \cdot G \circ \theta_n] = \mathbb{E}_x[\mathbb{1}_{T=n} F \ \mathbb{E}_{X_n}[G]].$$

Then, we take the summation over $n \in \mathbb{N}_0$ and obtain Eq. (7.11). □

Corollary 7.2.9 : *Let T be a stopping time satisfying $\mathbb{P}_x(T < \infty) = 1$. Suppose that there exists $y \in E$ such that $\mathbb{P}_x(X_T = y) = 1$, then under the probability measure $\mathbb{P}_x$, $\theta_T(\omega)$ is independent of $\mathcal{F}_T$ and has the same distribution as $\mathbb{P}_y$.*

Proof : It is a direct application of Theorem 7.2.7,

$$\mathbb{E}_x[F \cdot G(\theta_T(\omega))] = \mathbb{E}_x[F \ \mathbb{E}_{X_T}[G]] = \mathbb{E}_x[F \ \mathbb{E}_y[G]] = \mathbb{E}_x[F]\,\mathbb{E}_y[G].$$

□

上的非負可測函數。那麼對於所有 $x \in E$，以及非負 $\mathcal{F}_T$ 可測函數 F 來說，我們有

$$\mathbb{E}_x[\mathbb{1}_{T<\infty} F \cdot G \circ \theta_T] = \mathbb{E}_x[\mathbb{1}_{T<\infty} F \ \mathbb{E}_{X_T}[G]]. \tag{7.11}$$

對於任意固定的 $x \in E$ 來說，上述也可以重新寫做

$$\mathbb{E}_x[\mathbb{1}_{T<\infty} G \circ \theta_T \mid \mathcal{F}_T] = \mathbb{1}_{T<\infty} \mathbb{E}_{X_T}[G]. \tag{7.12}$$

註解 7.2.8 ： 我們注意到，這裡的隨機變數 X_T 是定義在 $\mathcal{F}_T$ 可測的集合 $\{T < \infty\}$ 之上，且是 $\mathcal{F}_T$ 可測的。期望值 $\mathbb{E}_{X_T}[G]$ 是個隨機變數，且也是定義在 $\{T < \infty\}$ 之上，可以視為函數 $\omega \mapsto X_T(\omega)$ 與 $x \mapsto \mathbb{E}_x[G]$ 的合成函數。

證明：給定整數 $n \geqslant 0$，由於 $\mathbb{1}_{T=n} F$ 是 $\mathcal{F}_n$ 可測的，我們可以將定理 7.2.6 中的簡單馬可夫性質用此函數上，進而得到

$$\mathbb{E}_x[\mathbb{1}_{T=n} F \cdot G \circ \theta_T] = \mathbb{E}_x[\mathbb{1}_{T=n} F \cdot G \circ \theta_n] = \mathbb{E}_x[\mathbb{1}_{T=n} F \ \mathbb{E}_{X_n}[G]].$$

接著，我們將上式對 $n \in \mathbb{N}_0$ 取合，便可以得到式 (7.11)。 □

系理 7.2.9 ： 令 T 是個滿足 $\mathbb{P}_x(T < \infty) = 1$ 的停止時間，假設存在 $y \in E$ 使得 $\mathbb{P}_x(X_T = y) = 1$，則在機率測度 $\mathbb{P}_x$ 之下，$\theta_T(\omega)$ 與 $\mathcal{F}_T$ 獨立，且有著與 $\mathbb{P}_y$ 相同的分佈。

證明：這是定理 7.2.7 的直接應用：

$$\mathbb{E}_x[F \cdot G(\theta_T(\omega))] = \mathbb{E}_x[F \ \mathbb{E}_{X_T}[G]] = \mathbb{E}_x[F \ \mathbb{E}_y[G]] = \mathbb{E}_x[F]\,\mathbb{E}_y[G].$$

□