Chapter 4

连续时间马氏过程

4.1 泊松（Poisson）过程

作为最简单的例子，我们考虑定义在 $\mathbb{Z}^d$ 上的随机过程，并且令其各个时间的变化方式是互相独立，且不随时间改变的。也就是说，

$$\mathbb{P}(X(t_1)-X(t_0)=\vec{j}_1,\ldots,X(t_n)-X(t_{n-1})=\vec{j}_n)=\prod_{m=1}^{n}\mathbb{P}(X(t_m-t_{m-1})=\vec{j}_m).$$

这里 $\vec{j}_m\in\mathbb{Z}^d$。不失一般性，我们可以假定 $X_0=\vec{0}$。

4.1.1 简单泊松过程

考虑定义在 $\mathbb{N}$ 上的随机过程 $\{N(t):t\geq 0\}$ 如下：首先，初始条件为 $N(0)=0$，然后从 0 时刻开始，这个随机过程待在 0 点总时间 E_1，之后向右走到 1，再花时间 E_2 待在 1 点，⋯⋯，且所有 E_n 各自独立，都满足标准指数分布 $\mathbb{P}(E_n>t)=e^{-t}$。也就是说，令 $\{E_n:n\geq 1\}$ 为相互独立，标准指数分布的随机变量，且

$$J_n=\begin{cases}0, & n=0,\\ \sum_{m=1}^{n}E_m, & n\geq 1,\end{cases}$$

则由

$$N(t)=\max\{n\geq 0:J_n\leq t\}$$

定义的随机过程 $\{N(t):t\geq 1\}$ 就是简单泊松过程。

关于简单泊松过程，我们首先有：所有 E_n 都大于 0，且 $\sum_{m=1}^{\infty} E_m = \infty$ 的概率为 1。所以，在概率 1 下，$N(t)$ 的路径（也就是从 $t \in [0,\infty)$ 到 $N(t) \in \mathbb{N}$ 的映射）是一个分片常数值，右连续函数，并且当它在不连续点跳跃时，总是增加 1：对于所有 $t \in (0,+\infty)$，$N(t) - N(t-) \in \{0,1\}$)（这里 $N(t-)$ 指左极限 $\lim_{s \nearrow t} N(s)$)。

然后，我们有，$\{N(t) : t \geq 0\}$ 在各个时间段的变化规律是各自独立，且不随时间改变的。也就是说，对于 $s,t < (0,+\infty)$，$N(s+t) - N(s)$ 独立于在时间 s 之前 $N(\tau)$ $(0 \leq \tau \leq s)$ 的历史，或者更精确地说，随机变量 $N(t) - N(s)$ 独立于由 $\{N(\tau) : \tau \in [0,s]\}$ 生成的 σ-代数（中的任意事件），且其分布等同于 $N(t)$：

$$\mathbb{P}(N(s+t) - N(s) \geq n \mid N(\tau), \tau \in [0,s]) = \mathbb{P}(N(t) \geq n), \quad n \in \mathbb{N}. \tag{4.1}$$

要证明此结论，只要对任意 $s,t \in [0,\infty)$ 和任意 $A \in \sigma(\{N(\tau) : \tau \in [0,s]\})$，验证对任意 $n \geq 0$，$\mathbb{P}(\{N(s+t) - N(s) \geq n\} \cap A) = \mathbb{P}(N(t) \geq n)\mathbb{P}(A)$。为此，我们表示 A 为

$$A = \bigcup_{m=0}^{\infty} A_m, \quad A_m = A \cap \{N(s) = m\}.$$

注意到 $A_m \in \sigma(\{N(\tau) : \tau \in [0,s]\})$，我们只要证明所有 A_m 都与 $N(s+t) - N(s)$ 独立即可。进一步，我们可以表示 A_m 为

$$A_m = \{J_{m+1} > s\} \cap B, \quad B \in \sigma(\{E_1, \ldots, E_m\}),$$

且在 B 上 $J_m \leq s$。于是，利用 E_{m+1} 服从指数分布的性质以及 E_{m+1} 与 B 相互独立，我们得到

$$\mathbb{P}(A_m) = \mathbb{P}(\{E_{m+1} > s - J_m\} \cap B) \overset{\text{对 } E_{m+1} \text{ 积分}}{\underset{\text{是 } J_m \text{ 的函数}}{=}} \mathbb{E}[e^{-(s-J_m)}, B].$$

另一方面，因为 $E_{m+1}, E_{m+2}, \ldots$ 与 B 相互独立，

$$\begin{aligned}
&\mathbb{P}(\{N(s+t) - N(s) \geq n\} \cap A_m) \\
&= \mathbb{P}(\{J_{m+n} \leq s+t\} \cap \{J_{m+1} > s\} \cap B) \\
&= \mathbb{P}(\{E_{m+1} + \cdots + E_{m+n} \leq s+t-J_m\} \cap \{E_{m+1} > s - J_m\} \cap B) \\
&= \underbrace{\mathbb{E}\left[\underbrace{\mathbb{P}(E_{m+1} + \cdots + E_{m+n} \leq s+t-J_m, \& E_{m+1} > s - J_m \mid J_m)}_{\text{先对 } E_{m+1}, \ldots, E_{m+n} \text{ 积分，结果是 } J_m \text{ 的函数}}, B\right]}_{\text{这个期望在 } \sigma(\{E_1, \ldots, E_m\}) \text{ 上计算}}.
\end{aligned}$$

由于 E_{m+1} 服从指数分布，满足一种“无记忆”性质，也就是对于 $a \geq 0$ 和任意 b，$\mathbb{P}(E_{m+1} > a+b \mid E_{m+1} > a) = \mathbb{P}(E_{m+1} > b)$，或者说 $\mathbb{P}(E_{m+1} > a+b, E_{m+1} > a) = \mathbb{P}(E_{m+1} > a+b)\mathbb{P}(E_{m+1} > a+b, \& E_{m+1} > a) = \mathbb{P}(E_{m+1} > a)\mathbb{P}(E_{m+1} > b)$，我们得到

$$
\begin{aligned}
&\mathbb{P}_{E_{m+1}}(E_{m+1} \leq t - (E_{m+2} + \cdots + E_{m+n}) + (s - J_m), \\
&\qquad \& E_{m+1} > s - J_m \mid J_m) \\
&= \mathbb{E}[\underbrace{\mathbb{P}(E_{m+1} \leq t - (E_{m+2} + \cdots + E_{m+n}))}_{\text{和 } J_m \text{ 相互独立}} \mathbb{P}_{E_{m+1}}(E_{m+1} > s - J_m) \mid J_m] \\
&= \mathbb{P}_{E_{m+1}}(E_{m+1} \leq t - (E_{m+2} + \cdots + E_{m+n}))\mathbb{P}_{E_{m+1}}(E_{m+1} > s - J_m \mid J_m).
\end{aligned}
$$

并且简化上面公式为

$$
\begin{aligned}
&\mathbb{P}(\{N(s+t) - N(s) \geq n\} \cap A_m) \\
&= \mathbb{E}\left[\mathbb{P}(E_{m+1} + \cdots + E_{m+n} \leq t)\mathbb{P}(E_{m+1} > s - J_m \mid J_m), B\right] \\
&= \mathbb{P}(E_{m+1} + \cdots + E_{m+n} \leq t)\mathbb{E}[\mathbb{P}(E_{m+1} > s - J_m \mid J_m) \cap B) \\
&= \mathbb{P}(N(t) \leq n)\mathbb{P}\{E_{m+1} > s - J_m\} \cap B) \\
&= \mathbb{P}(N(t) \leq n)\mathbb{P}(A_m).
\end{aligned}
$$

至此，我们证明(4.1)。

至于 $\mathbb{P}(N(t) \leq n)$，因为它是由几个相互独立的指数分布随机变量定义的，通过计算定积分，我们可以得到 $\mathbb{P}(N(t) \leq n) = t^n/n!$，也就是说，$N(t)$ 是一个依均值 t 泊松分布的随机变量。

有上述结果，我们有

$$
\mathbb{P}(N(s+t) - N(s) = n \mid N(\tau), \tau \in [0,s]) = e^{-t}\frac{t^n}{n!},
$$
$$
\mathbb{P}(N(s+t) = n \mid N(\tau), \tau \in [0,s]) = e^{-t}\frac{t^{n-N(s)}}{(n-N(s))!}1_{[0,n]}(N(s)).
$$

4.1.2 $\mathbb{Z}^d$ 上的复合泊松过程

令 μ 为 $\mathbb{Z}^d$ 上的一个概率分布向量，且 $\mu_{\vec{0}} = 0$。则此分布加上一个跳跃速率 $R > 0$ 定义了一个初始条件为 $\vec{X}(0) = \vec{0}$ 的复合泊松过程 $\{\vec{X}(t) : t \geq 0\}$。这个随机过程开始于 $\vec{0}$，在 0 上停留一段随机时间，且这个随机时间依均值为 R^{-1} 的指数分布，。然后这个随机过程立即进行第一次跳跃，以概率 $\mu_{\vec{k}}$ 跳跃到 $\vec{k}$ 上。每次跳跃后，都等待依均值 R^{-1} 的指数分布的时间，然后跳

跃，以概率 $\mu_{\vec{k}}$ 使位置发生 $\vec{k}$ 改变……上一节介绍的简单泊松过程，就是 $d=1$，$\mu_1=1$ 且 $R=1$ 的特例。

复合泊松过程的具体构造如下：令 $\{\vec{B}_n\}$ 为一列相互独立的，服从分布 μ 的 $\mathbb{Z}^d$ 值随机变量。令 $\vec{X}_0=\vec{0}$，而对于 $n\geq 1$，令 $X_n=\sum_{m=1}^n \vec{B}_m$。接着，定义 $\{\vec{X}(t):t\geq 0\}$ 为 $\vec{X}(t)=\vec{X}_{N(R(t))}$，这里 $\{N(t):t\geq 0\}$ 是和 $\{\vec{B}_m\}$ 相互独立的简单泊松过程。

显然，$\vec{X}(0)=\vec{0}$，且作为从 $[0,\infty)$ 到 $\mathbb{Z}^d$ 的映射，$\vec{X}(t)$ 是分段常数值，右连续的。因为我们一开始假设 $\vec{B}_m$ 都不为 $\vec{0}$，所以在时间段 $(s,t]$ 上，复合泊松过程跳跃次数等于 $N(Rt)-N(Rs)$。同理可知，$\vec{B}_n=\vec{X}_n-\vec{X}_{n-1}$ 是 $\vec{X}(t)$ 的第 n 次跳跃的值。于是，如果记 $J_0=0$ 且当 $n\geq 1$ 时 J_n 为 $\vec{X}(t)$ 第 n 次跳跃的时间，则 $N(Rt)=n$ 当且仅当 $J_n\leq Rt<J_{n+1}$，且这时 $\vec{X}(J_n)-\vec{X}(J_{n-1})=\vec{X}_n-\vec{X}_{n-1}=\vec{B}_n$。换句话说，如果记 $\{E_n:n\geq 1\}$ 为构造 $\{N(t):t\geq 0\}$ 时所用的指数分布随机变量，则

$$J_n-J_{n-1}=\frac{E_n}{R},\quad \vec{X}(t)-\vec{X}(t-)=\begin{cases}\vec{0}, & t\in(J_{n-1},J_n),\\ \vec{B}_n, & t=J_n.\end{cases}$$

很明显，我们这里构造的复合泊松过程满足一开始的描述。

下面，我们证明复合泊松过程的变化规律是不随时间改变，且不同时间相互独立的：

$$\mathbb{P}(\vec{X}(s+t)-\vec{X}(s)=\vec{k}\mid \vec{X}(\tau),\tau\in[0,s])=\mathbb{P}(\vec{X}(t)=\vec{k}),\quad \vec{k}\in\mathbb{Z}^d.$$

和简单泊松过程中类似结论证明相似，我们只需要证明，对任意事件 $A\in\sigma(\{\vec{X}(\tau):\tau\in[0,s]\})$，有

$$\begin{aligned}\mathbb{P}(\{\vec{X}(s+t)-\vec{X}(s)=\vec{k}\}\cap A)&=\mathbb{P}(\vec{X}(s+t)-\vec{X}(s)=\vec{k})\mathbb{P}(A)\\&=\mathbb{P}(\vec{X}(t)=\vec{k})\mathbb{P}(A).\end{aligned}\tag{4.2}$$

不失一般性，我们假定这个 A 满足 $N(Rs)=m$（也就是相当于简单泊松过程的证明中的 A_m）。这样，A 独立于 $\sigma(\{\vec{X}_{m+n}-\vec{X}_m:n\geq 0\})$（这是由于随机游动 $\{\vec{X}_n\}$ 的马氏性）和 $\sigma(\{N(R(s+t))-N(Rs)\})$（这是由于之前证

明的简单泊松过程的性质)，而这两个 σ-代数互相独立。于是

$$
\begin{aligned}
&\mathbb{P}(\{\vec{X}(s+t)-\vec{X}(s)=\vec{k}\}\cap A)\\
&=\sum_{n=0}^{\infty}\mathbb{P}(\{\vec{X}(s+t)-\vec{X}(s)=\vec{k}\& N(R(s+t))-N(Rs)=n\}\cap A)\\
&=\sum_{n=0}^{\infty}\mathbb{P}(\{\vec{X}_{m+n}-\vec{X}_m=\vec{k}\& N(R(s+t))-N(Rs)=n\}\cap A)\\
&=\sum_{n=0}^{\infty}\mathbb{P}(\vec{X}_n=\vec{k})\mathbb{P}(N(R(s+t))-N(Rs)=n)\mathbb{P}(A)\\
&=\sum_{n=0}^{\infty}\mathbb{P}(\vec{X}_n=\vec{k})\mathbb{P}(N(Rt)=n)\mathbb{P}(A)\\
&=\sum_{n=0}^{\infty}\mathbb{P}(\vec{X}_n=\vec{k}\& N(Rt)=n)\mathbb{P}(A)=\mathbb{P}(\vec{X}(t)=\vec{k})\mathbb{P}(A).
\end{aligned}
$$

就证明了(4.2)。

然后我们计算 $\vec{X}(t)$ 的分布。我们已经记 $\vec{B}_1$（也就是任意 $\vec{B}_n$）的分布为 μ。现在，我们引入 μ^{*n}，来标记 $\vec{X}_n=\vec{B}_1+\vec{B}_2+\cdots+\vec{B}_n$ 的分布。(同时，令 μ^{*0} 为 $\vec{0}$ 处单点分布：$(\mu^{*0})_{\vec{k}}=\delta_{\vec{0},\vec{k}}$)。显然，

$$
(\mu^{*n})_{\vec{k}}=\sum_{\vec{j}\in\mathbb{Z}^d}(\mu^{*(n-1)})_{\vec{k}-\vec{j}}(\mu)_{\vec{j}},\quad n\geq 1.
$$

所以，

$$
\mathbb{P}(\vec{X}(t)=\vec{k})=\sum_{n=0}^{\infty}\mathbb{P}(\vec{X}_n=\vec{k}\& N(Rt)=n)=e^{-Rt}\sum_{n=0}^{\infty}\frac{(Rt)^n}{n!}(\mu^{*n})_{\vec{k}}.
$$

并且，通过(4.2)，我们得到（这里假定 $A\in\sigma(\{\vec{X}(\tau):\tau\in[0,s]\})$ 而且不假定 $N(Rs)$ 的值）

$$
\begin{aligned}
&\mathbb{P}(\{\vec{X}(s+t)=\vec{k}\}\cap A)\\
&=\sum_{\vec{j}\in\mathbb{Z}^d}\mathbb{P}(\{\vec{X}(s+t)=\vec{k}\}\cap A\cap\{\vec{X}(s)=\vec{j}\})\\
&=\sum_{\vec{j}\in\mathbb{Z}^d}\mathbb{P}(\{\vec{X}(s+t)-\vec{X}(s)=\vec{k}-\vec{j}\}\cap A\cap\{\vec{X}(s)=\vec{j}\})\\
&=\sum_{j\in\mathbb{Z}^d}(\mathbf{P}(t))_{\vec{j}\vec{k}}\mathbb{P}(A\cap\{\vec{X}(s)=\vec{j})\\
&=\mathbb{E}[(\mathbf{P}(t))_{\vec{X}(s),\vec{k}},A],
\end{aligned}
$$

这里

$$(\mathbf{P}(t))_{\vec{j},\vec{k}} = e^{-Rt}\sum_{m=0}^{\infty}\frac{(Rt)^m}{m!}(\mu^{*m})_{\vec{k}-\vec{j}}.$$

换句话说，我们证明了

$$\mathbb{P}(\vec{X}(s+t)=\vec{k} \mid \vec{X}(\sigma),\sigma\in[0,s]) = (\mathbf{P}(t))_{\vec{X}(s),\vec{k}},$$

复合泊松过程 $\{\vec{X}(t):t\geq 0\}$ 是一个连续时间马氏过程并且其转移概率为 $\mathbf{P}(t)$。

这里的 $\{\mathbf{P}(t):t\geq 0\}$ 是一个半群：它满足查普曼-科尔莫哥罗夫（*Chapman-Kolmogorov*）方程

$$\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t), \quad s,t\in[0,\infty).$$

这个可以直接验证，也可以首先观察到 $(\mathbf{P}(t))_{\vec{j}\vec{k}} = (\mathbf{P}(t))_{\vec{0},\vec{k}-\vec{j}}$，然后用 $\vec{X}(0)=\vec{0}$,

$$\begin{aligned}
(\mathbf{P}(s+t))_{\vec{0},\vec{k}} &= \mathbb{P}(\vec{X}(s+t)=\vec{k}) \\
&= \sum_{\vec{j}\in\mathbb{Z}^d}\mathbb{P}(\vec{X}(s+t)=\vec{k}\,\&\,\vec{X}(s)=\vec{j}) \\
&= \sum_{\vec{j}\in\mathbb{Z}^d}\mathbb{P}(\vec{X}(s+t)=\vec{k} \mid \vec{X}(s)=\vec{j})\mathbb{P}(\vec{X}(s)=\vec{j}) \\
&= \sum_{\vec{j}\in\mathbb{Z}^d}(\mathbf{P}(t)_{\vec{j},\vec{k}}(\mathbf{P}(t)_{\vec{0},\vec{j}} = (\mathbf{P}(s)\mathbf{P}(t))_{\vec{0},\vec{k}}.
\end{aligned}$$

4.2 有界速率的马氏过程

令 $\mathbb{S}$ 为一个可数状态空间, $\mathbf{P}$ 一个概率转移矩阵且对所有 $i\in\mathbb{S}$, $(\mathbf{P})_{ii}=0$。又令 $\mathfrak{R}=\{R_i:i\in\mathbb{S}\}\subseteq[0,\infty)$ 为一族速率。这样，我们定义 $\mathbb{S}$ 上的连续时间马氏过程，使其速率由 $\mathfrak{R}$ 给定而转移概率由 $\mathbf{P}$ 给定，且满足如下要求：

1. 作为从 $[0,\infty)$ 到 $\mathbb{S}$ 的映射，$X(t)$ 是分段常数值，且右连续的。
2. 令 $J_0=0$ 且对任意 $n\geq 1$，J_n 为 $X(t)$ 的第 n 此跳跃。$J_{n-1}<\infty$ 这个条件下，

   $$\mathbb{P}(J_n > J_{n-1}+t\,\&\,X(J_n)=j \mid X(\tau),\tau\in[0,J_n)) = e^{-tR_{X(J_{n-1}}}(\mathbf{P})_{X(J_{n-1}),j}.$$

在本章，我们假定 $\sup\mathfrak{R} < \infty$，也就是有界速率。

下面我们构造满足要求的马氏过程。先考虑我们一个技术性的非退化假设：对所有 $i \in \mathbb{S}$，$R_i > 0$。这个假设的直观意义是，马氏链到了任何状态，都可以继续走下去。（反例如作业里出现的 Galton-Watson 过程，如果此过程到了人口为 0 这个状态，就永远人口为零，走不出去了。）我们可以得到对任意 $n \geq 0$，$\mathbb{P}(J_n < \infty) = 1$。我们先令 $X_n = X(J_n)$ 以及 $E_n = (J_n - J_{n-1})R_{X_{n-1}}$（$n = 1, 2, \dots$），然后注意到，由第二项要求，

$$\mathbb{P}(E_n > t \& X_n = j \mid \{E_1, \dots, E_{n-1} \cup \{X_0, \dots, X_{n-1}\}) = e^{-t}(\mathbf{P})_{X_{n-1},j}. \tag{4.3}$$

这样，$\{X_n : n \geq 0\}$ 就是以 $\mathbf{P}$ 为概率转移矩阵的马氏链，且其初始分布等于 $X(0)$ 的分布。$\{E_n : n \geq 1\}$ 是一列互相独立的标准指数分布随机变量。最后，$\{X_n : n \geq 0\}$（生成的 σ-代数）和 $\{E_n : n \geq 1\}$（生成的 σ-代数）是相互独立的。于是，$\{X_n : n \geq 0\}$ 和 $\{E_n : n \geq 1\}$ 有惟一确定的联合分布。

反过来，如果我们有 $\{X_n : n \geq 0\}$ 和 $\{E_n : n \geq 1\}$，也可以确定 $\{X(t) : t \geq 0\}$。考虑 $(e_1, \dots, e_n, \dots) \in (0, \infty)^{\mathbb{Z}_+}$ 和 $(j_0, \dots, j_n, \dots) \in \mathbb{S}^{\mathbb{N}}$。令

$$\Phi^{(\mathfrak{R},\mathbf{P})}(t; (e_1, \dots, e_n, \dots), (j_0, \dots, j_n, \dots)) = j_n, \quad \text{如果 } \xi_n \leq t < \xi_n,$$

这里 $\xi_0 = 0$, 对 $n \geq 1$ 则 $\xi_n = \sum_{m=1}^{n} R_{j_{m-1}}^{-1} e_m$。则如果 $\{X_n : n \geq 0\}$ 和 $\{E_n : n \geq 1\}$ 是由 $\{X(t) : t \geq 0\}$ 定义的, 就有 $X(t) = \Phi^{(\mathfrak{R},\mathbf{P})}(t; (E_1, \dots, E_n, \dots), (X_0, \dots, X_n, \dots))$, 只要 $0 \leq t < \sum_{m=1}^{\infty} R_{j_{m-1}}^{-1} E_m$。如果我们假定了 $\sup\mathfrak{R} < \infty$，则以概率 1 我们有 $0 \leq t < \sum_{m=1}^{\infty} R_{j_{m-1}}^{-1} E_m$。所以，在速率有界条件下，$\{X(t) : t \geq 0\}$ 与 $\{X_n : n \geq 0\} \cup \{E_n : n \geq 1\}$ 一一对应。

下面我们考虑退化情形，也就是某些 R_i 等于 0。令 $\mathbb{S}_0 = \{i : R_i = 0\}$。定义 $\bar{\mathfrak{R}} = \{\bar{R}_i : i \in \mathbb{S}\}$ 为

$$\bar{R}_i = \begin{cases} R_i, & i \notin \mathbb{S}_0, \\ 1, & i \in \mathbb{S}_0. \end{cases}$$

这个 $\bar{\mathfrak{R}}$ 与 $\mathbf{P}$ 惟一确定了一个非退化的马氏过程 $\{\bar{X}(t) : t \geq 0\}$。令 $\xi = \inf\{t \geq 0 : \bar{X}(t) \in \mathbb{S}_0\}$。我们可以证明，$\{X(t) : t \geq 0\}$ 的分布等同于 $\{\bar{X}(t \wedge \zeta) : t \geq 0\}$。具体来说，就是定义一个 $\{\bar{X}(t)\}$，使得它在 $t \wedge \zeta$ 时间前显然地等于 $\{X(t)\}$，然后再证明 $\{\bar{X}(t)\}$ 是一个如上所述的非退化马氏过程。

对每个 $i \in \mathbb{S}_0$，令 $\{X_n^{(i)} : n \geq 0\}$ 为 $\mathbb{S}$ 值的随机变量，且令 $\{\bar{E}_n : n \geq 1\}$ 为 $(0, \infty)$ 值随机变量，并要求

1. $\sigma(\{X_n^{(i)} : n \geq 0 \& i \in \mathbb{S}\})$，$\sigma(\{\bar{E}_n : n \geq 1\})$ 和 $\sigma(\{X(t) : t \geq 0\})$ 相互独立。

2. 对于每个 $i \in \mathbb{S}_0$，$\{X_n^{(i)} : n \geq 0\}$ 是个以 $\mathbf{P}$ 为转移概率矩阵的马氏链，且其初始概率分布为 $X_0^{(i)} = i$。

3. $\{\bar{E}_n : n \geq 1\}$ 为相互独立同分布的单位指数分布随机变量。

对于 $i \in \mathbb{S}$，可以定义以 $\bar{X}^{(i)}(0) = i$ 为初始概率分布的连续时间马氏过程

$$\bar{X}^{(i)}(t) = \Phi^{(\bar{\mathfrak{R}}, \mathbf{P})}(t; (\bar{E}_1, \ldots, \bar{E}_n, \ldots), (X_0^{(i)}, \ldots, X_n^{(i)}, \ldots)).$$

于是，我们取的 $\{\bar{X}(t)\}$ 就是把 $\{X(t)\}$ 和 $\{\bar{X}^{(i)}(t)\}$ 接起来：

$$\bar{X}(t) = \begin{cases} X(t), & t < \zeta, \\ \bar{X}^{(i)}(t - \zeta), & t \geq \zeta \text{且} X(\zeta) = i \in \mathbb{S}_0. \end{cases}$$

下面我们所要做的，就是证明这个 $\{\bar{X}(t)\}$ 确实是 $\bar{\mathfrak{R}}$ 和 $\mathbf{P}$ 定义的马氏过程。

令 $\bar{J}_0 = 0$ 而 $\bar{J}_m$ 为 $\bar{X}(t)$ 第 m 次跳跃的时间。我们需要对任意 n 和任意 $A \in \sigma(\{\bar{X}(\tau) : \tau \in [0, \bar{J}_n)\}$，

$$\mathbb{P}(\{\bar{J}_n > \bar{J}_{n-1} + t \& \bar{X}(\bar{J}_n) = j\} \cap A) = \mathbb{E}[e^{-t\bar{R}_{\bar{X}(J_{n-1})}}(\mathbf{P})_{\bar{X}(\bar{J}_{n-1}), j}, A]. \quad (4.4)$$

不失一般性，我们可以假定在对 $m = 0, \ldots, n-1$，A 上 $\bar{X}(\bar{J}_m) = j_m$。(因为一般的 A 都是这样的特殊形式的集合的不相交并集)。如果所有的 $j_0, \ldots, j_{n-1}$ 都不属于 $\mathbb{S}_0$，也就是所有的 $R_{j_m} > 0$，我们有 $A \in \sigma(\{X(\sigma) : \sigma \in [0, J_n)\})$，而且对 $\tau \in [0, J_n)$，都有 $\bar{X}(t) = X(t)$。所以这种情况下，(4.4)等价于对(4.3)。不然的话，令 $m = \min\{k < n : j_k \in \mathbb{S}_0 \text{ 且 } j_m = i\}$，则

$$\bar{J}_k = \begin{cases} J_k, & k = 0, \ldots, m, \\ \bar{J}_{k-m}^{(i)} + J_m, & k = m+1, \ldots, n, \end{cases}$$

并且 $A \in \sigma(\{X_0, X_1, \ldots, X_m, X_1^{(i)}, \ldots, X_{n-m-1}^{(i)}, E_1, \ldots, E_m, E_1^{(i)}, \ldots, E_{n-m-1}^{(i)}\})$。应用独立性关系，可以看到 $A = B \cap C$，其中 $B \in \sigma(\{X_0, X_1, \ldots, X_m, E_1, \ldots, E_m\}$，

$C \in \sigma(\{X_1^{(i)}, \ldots, X_{n-m-1}^{(i)}, E_1^{(i)}, \ldots, E_{n-m-1}^{(i)}\})$，且 B, C 相互独立，$\mathbb{P}(A) = \mathbb{P}(B)\mathbb{P}(C)$。于是

$$\begin{aligned}
&\mathbb{P}(\{\bar{J}_n > \bar{J}_{n-1} + t \& \bar{X}(\bar{J}_n) = j\} \cap A) \\
&= \mathbb{P}(\{\bar{J}_n > \bar{J}_{n-1} + t \& \bar{X}(\bar{J}_n) = j\} \cap C \cap B) \\
&= \mathbb{P}(\{\bar{J}_{n-m}^{(i)} > \bar{J}_{n-m-1}^{(i)} + t \& \bar{X}_{n-m}^{(i)} = j\} \cap C)\mathbb{P}(B) \\
&= \underbrace{\mathbb{P}(\{\bar{J}_{n-m}^{(i)} > \bar{J}_{n-m-1}^{(i)} + t \& \bar{X}_{n-m}^{(i)} = j\} \mid C)}_{\text{这个计算完全在 } \{\bar{X}^{(i)}(t)\} \text{ 马氏过程上}} \mathbb{P}(C)\mathbb{P}(B) \\
&= \exp(-t\bar{R}_{j_{n-1}})(\mathbf{P})_{j_{n-1},j}\mathbb{P}(C)\mathbb{P}(B) \exp(-t\bar{R}_{j_{n-1}})(\mathbf{P})_{j_{n-1},j}\mathbb{P}(A) \\
&= \mathbb{E}[\exp(-t\bar{R}_{\bar{X}(J_{n-1})})(\mathbf{P})_{\bar{X}(J_{n-1}),j}, A].
\end{aligned}$$

由上面的构造，我们知道，只要 $\mathfrak{R}$ 是有界速率，连续时间马氏过程总可以用由 $\mathbf{P}$ 定义的马氏过程（$X_0 = X(0)$）和与之相互独立的一列服从标准指数分布的 E_i 构造。如果速率中有些 $R_i = 0$，则不见得所有的 E_i 和 X_i 都用得上。

4.2.1 马氏性

在速率有界的假定下,我们证明,如果令 $(\mathbf{P}(t))_{ij} = \mathbb{P}(X(t) = j \mid X(0) = i)$，则

$$\mathbb{P}(X(s+t) = j \mid X(\tau), \tau \in [0, s]) = (\mathbf{P}(t))_{X(s),j}. \tag{4.5}$$

我们只要证明，对于任意事件 $A \in \sigma(\{X(\tau) : \tau \in [0, s]\})$，

$$\mathbb{P}(\{X(s+t) = j\} \cap A) = (\mathbf{P}(t))_{X(s),j}\mathbb{P}(A). \tag{4.6}$$

不失一般性，我们可以假定在 A 上，$X(s) = i$。(不然的话，按 $X(s)$ 的值，把 A 分为可数多互不相交的子集。）则我们只需要证明

$$\mathbb{P}(\{X(s+t) = j\} \cap A) = (\mathbf{P}(t))_{ij}\mathbb{P}(A). \tag{4.7}$$

然后，因为

$$A = \bigcup_{m=0}^{\infty} A_m, \quad A_m = A \cap \{N(s) = m\},$$

我们又可以把(4.7)中的 A 换成 A_m。接着，又有

$$\begin{aligned}
&A_m = \{E_{m+1} > R_i(s - J_m)\} \cap B_m, \\
&\qquad B_m \in \sigma(\{E_1, \ldots, E_m\} \cup \{X_0, \ldots, X_m\}) \text{ 且} B_m \subseteq \{J_m \le s\}.
\end{aligned}$$

为此，我们先验证如果 $s \in [\xi_m, \xi_{m+1})$，则

$$\begin{aligned}&\Phi^{\mathfrak{R},\mathbf{P}}(s+t;(e_1,\dots,e_n,\dots),(j_0,\dots,j_n,\dots)) = \\ &\Phi^{\mathfrak{R},\mathbf{P}}(t;(e_{m+1}-R_{j_m}(s-\xi_m),e_{m+2},\dots,e_{m+n},\dots),(j_m,\dots,j_{m+n},\dots)).\end{aligned}$$

最后，因为

$$\begin{aligned}&\mathbb{P}(\{X(s+t)=j\}\cap A_m) \\ &= \mathbb{P}(\{X(s+t)=j\&E_{m+1}>R_i(s-J_m)\}\cap B_m) \\ &= \mathbb{P}(\{\Phi^{\mathfrak{R},\mathbf{P}}(s+t;(E_1,\dots,E_n,\dots);(X_0,\dots,X_n,\dots))=j\} \\ &\qquad\qquad \cap\{E_{m+1}>R_i(s-J_m)\}\cap B_m) \\ &= \mathbb{P}(\{\Phi^{\mathfrak{R},\mathbf{P}}(t;(E_{m+1}-R_i(s-J_m),E_{m+2},\dots,E_{m+n},\dots); \\ &\qquad\qquad (i,X_{m+1},\dots,X_{m+n},\dots))=j\}\cap\{E_{m+1}>R_i(s-J_m)\}\cap B_m) \\ &= \mathbb{P}(X(t)=j \mid X(0)=i)\mathbb{P}(\{E_{m+1}>R_i(s-J_m)\}\cap B_m) \\ &= (\mathbf{P}(t))_{ij}\mathbb{P}(A_m).\end{aligned}$$

这样我们证明了马氏性。

作为一个特例，我们得到 $\{\mathbf{P}(t):t\ge 0\}$ 的半群性质：

$$\begin{aligned}&(\mathbf{P}(s+t))_{ij} \\ &= \sum_{k\in\mathbb{S}}\mathbb{P}(X(s+t)=j\&X(s)=k \mid X(0)=i) \\ &= \sum_{k\in\mathbb{S}}\mathbb{P}(X(s+t)=j \mid X(s)=k,X(0)=i)\mathbb{P}(X(s)=k \mid X(0)=i) \\ &= \sum_{k\in\mathbb{S}}(\mathbf{P}(t))_{k,j}(\mathbf{P}(s))_{i,k} = (\mathbf{P}(s)\mathbf{P}(t))_{ij}.\end{aligned}$$

最后，我们由上面结果可以得到：连续时间马氏过程的分布，完全由其初始分布以及转移矩阵半群 $\{\mathbf{P}(t):t\ge 0\}$ 惟一决定。精确地说，假定 $\{X(t):t\ge 0\}$ 是一族 $\mathbb{S}$ 值的随机变量，满足(4.5)，且 $X(0)$ 的分布给定为 μ。则对任意 $n\ge 1$，$0=t_0<t_1<\cdots<t_n$ 以及 $j_0,\dots,j_n\in\mathbb{S}$，都有

$$\mathbb{P}(X(t_0)=j_0,\dots,X(t_n)=j_n)=\mu_{j_0}(\mathbf{P}(t_1-t_0))_{j_0j_1}\cdots(\mathbf{P}(t_n-t_{n-1}))_{j_{n-1}j_n}.$$

当 $n=1$ 时，此结论就是(4.5)的直接推论。当 $n>1$ 时，令 $s=t_{n-1}$，$t=t_n-t_{n-1}$ 且 $A=\{X(t_0)=j_0,\dots,X(t_{n-1})=j_{n-1}\}\in\sigma(\{X(\tau):\tau\in[0,s]\})$。应用(4.6)，我们得到

$$\mathbb{P}(X(t_0)=j_0,\dots,X(t_n)=j_n)=\mathbb{P}(X(s+t)=j_n\cap A)=(\mathbf{P}(t))_{j_{n-1}j_n}\mathbb{P}(A),$$

于是应用数学归纳法，得到结论。

4.2.2 Q 矩阵和科尔莫哥罗夫后向方程

从 $\mathfrak{R}$ 和 $\mathbf{P}$ 出发，怎样构造 $\{\mathbf{P}(t) : t \geq 0\}$?

我们希望表示 $\mathbf{P}(t)$ 为 $e^{t\mathbf{Q}}$，这样，$\mathbf{Q}$ 应该像 $\mathbf{P}(t)$ 在零点处对导数。

为此，我们要证明对于任意 $t > 0$

$$(\mathbf{P}(t))_{ij} = \delta_{ij} e^{-tR_i} + R_i \int_0^t e^{-\tau R_i} (\mathbf{P}\mathbf{P}(t-\tau))_{ij} d\tau. \tag{4.8}$$

如果 $R_i = 0$，显然 $(\mathbf{P}(t))_{ij} = \delta_{ij}$，而等式右边也是 δ_{ij}，所以等式(4.8)得证。下面，我们证明等式(4.8)在 $R_i > 0$ 时也成立。

因为

$$(\mathbf{P}(t))_{ij} = \delta_{ij} \mathbb{P}(E_1 > tR_1 \mid X(0) = i) + \mathbb{P}(E_1 \leq tR_i \& X(t) = j \mid X(0) = i),$$

其中 $\mathbb{P}(E_1 > tR_1 \mid X(0) = i) = e^{-tR_i}$，而

$$\begin{aligned}
&\mathbb{P}(E_1 \leq tR_i \& X(t) = j \mid X(0) = i) \\
&= \mathbb{E}[\mathbb{P}(X(t) = j \mid E_1), E_1 \leq tR_i \mid X(0) = i] \\
&= \mathbb{E}\left[\underbrace{(\mathbf{P}(t - R_i^{-1} E_1)_{X_2, j}}_{\text{马氏性(4.5)，其中 } s = R_i^{-1} E_1}, E_1 \leq tR_i \middle| X(0) = i \right] \\
&= \mathbb{E}\left[\mathbb{E}\left[(\mathbf{P}(t - R_i^{-1} E_1)_{X_2, j} \middle| X(0) = i \right], E_1 \leq tR_i \right], \\
&\quad \text{（利用 } E_1 \text{ 与 } X(0) = X_1 \text{ 相互独立）} \\
&= \mathbb{E}\left[\sum_{k \in \mathbb{S}} (\mathbf{P}(t - R_i^{-1} E_1)_{kj} (\mathbf{P})_{ik}, E_1 \leq tR_i \right] \\
&= \mathbb{E}\left[(\mathbf{P}\mathbf{P}(t - R_i^{-1} E_1))_{ij}, E_1 \leq tR_i \right] \\
&= \int_0^t e^{-\tau R_i} (\mathbf{P}\mathbf{P}(t-\tau))_{ij} d\tau.
\end{aligned}$$

结合上面两部分，就证明了(4.8)。

在(4.8)两边对 t 求导（练习），我们得到科尔莫哥洛夫向后方程

$$\frac{d}{dt}(\mathbf{P}(t))_{ij} = -R_i (\mathbf{P}(t))_{ij} + R_i (\mathbf{P}\mathbf{P}(t))_{ij},$$

或者，引入所谓的 “**Q**-矩阵”：

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{Q}\mathbf{P}(t), \quad \mathbf{Q} = \mathrm{diag}(R_i)(\mathbf{P} - I). \tag{4.9}$$

下面，我们记 $\mathbf{R} := \mathrm{diag}(R_i)$。

4.2.3 科尔莫哥洛夫向前方程

回忆3.2.1节引入的算子范数 $\|\cdot\|_{\infty,\infty}$。因为我们考虑的马氏过程速率有界，所以 $\|\mathbf{Q}\|_{\infty,\infty} \le \sup_{i\in\mathbb{S}} R_i < \infty$。由(4.9)，通过对两边分别积分，得

$$\begin{aligned}
\mathbf{P}(t) &= I + \int_0^t \mathbf{Q}\mathbf{P}(s)ds \\
&= I + \int_0^t \mathbf{Q}\left(I + \int_0^s \mathbf{Q}\mathbf{P}(\tau)d\tau\right)ds \\
&= I + t\mathbf{Q} + \int_0^t (t-\tau)\mathbf{Q}^2\mathbf{P}(\tau)d\tau.
\end{aligned}$$

由此，对于任意 $t > 0$，

$$\begin{aligned}
\|\mathbf{P}(t) - I - t\mathbf{Q}\|_{\infty,\infty} &= \left\|\int_0^t (t-\tau)\mathbf{Q}^2\mathbf{P}(\tau)d\tau\right\|_{\infty,\infty} \\
&\le \int_0^t (t-\tau)\|\mathbf{Q}\|_{\infty,\infty}^2 \cdot \underbrace{\|\mathbf{P}(\tau)\|_{\infty,\infty}}_{\le 1} d\tau \qquad (4.10) \\
&\le \frac{t^2}{2}\|\mathbf{Q}\|_{\infty,\infty}^2.
\end{aligned}$$

再用 $\mathbf{P}(t)$ 的半群性质，得

$$\left\|\frac{\mathbf{P}(t+h) - \mathbf{P}(t)}{h} - \mathbf{P}(t)\mathbf{Q}\right\|_{\infty,\infty} = \frac{1}{h}\|\mathbf{P}(t)\| \cdot \|\mathbf{P}(h) - I - h\mathbf{Q}\| \le \frac{h}{2}\|\mathbf{Q}\|_{\infty,\infty}^2. \tag{4.11}$$

取 $h \to 0$ 的极限，我们得到微分公式，也就是所谓的**科尔莫哥洛夫向前方程**

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{Q}, \quad \mathbf{P}(0) = I.$$

4.2.4 求解科尔莫哥洛夫方程

很容易猜出来

$$\mathbf{P}(t) = e^{t\mathbf{Q}} := \sum_{m=0}^{\infty} \frac{t^m}{m!}\mathbf{Q}^m. \tag{4.12}$$

至于严格证明，首先，我们可以看到右边的级数在 $\|\cdot\|_{\infty,\infty}$ 下收敛，其次可以验证 $e^{(s+t)\mathbf{Q}} = e^{s\mathbf{Q}}e^{t\mathbf{Q}}$，也就是说，$\{e^{t\mathbf{Q}} : t \geq 0\}$ 也构成一个半群。（其实，$e^{t\mathbf{Q}}$ 对 $t < 0$ 时也有良好定义，并且 $\{e^{t\mathbf{Q}} : t \in \mathbb{R}\}$ 是一个单参数群。）由因为和(4.10)类似，我们有

$$\|e^{t\mathbf{Q}} - I - t\mathbf{Q}\|_{\infty,\infty} \leq \frac{t^2}{2} e^{t\|\mathbf{Q}\|_{\infty,\infty}} \|\mathbf{Q}\|^2_{\infty,\infty}, \tag{4.13}$$

所以，和(4.11)类似，

$$\left\| \frac{e^{(t+h)\mathbf{Q}} - e^{t\mathbf{Q}}}{h} - e^{t\mathbf{Q}}\mathbf{Q} \right\|_{\infty,\infty} \leq \frac{h}{2} e^{t\|\mathbf{Q}\|_{\infty,\infty}} \|\mathbf{Q}\|^2_{\infty,\infty},$$

也就意味着微分公式 $\frac{d}{dt}e^{t\mathbf{Q}} = e^{t\mathbf{Q}}\mathbf{Q}$。最后，我们考虑 $e^{(t-\tau)\mathbf{Q}}\mathbf{P}(\tau)$ ($\tau \in [0,t]$)。如果我们能证明它是（算子空间的）常数，则 $\mathbf{P}(t)$ ($\tau = t$ 时) 等于 $e^{t\mathbf{Q}}$ ($\tau = 0$ 时)。为此，我们只要验证

$$\begin{aligned} \frac{d}{d\tau}\left(e^{(t-\tau)\mathbf{Q}}\mathbf{P}(\tau)\right) &= \left(\frac{d}{d\tau}e^{(t-\tau)\mathbf{Q}}\right)\mathbf{P}(\tau) + e^{(t-\tau)\mathbf{Q}}\frac{d}{d\tau}\left(\mathbf{P}(\tau)\right) \\ &= \left(-e^{(t-\tau)\mathbf{Q}}\mathbf{Q}\right)\mathbf{P}(\tau) + e^{(t-\tau)\mathbf{Q}} \underbrace{\left(\mathbf{Q}\mathbf{P}(\tau)\right)}_{\text{应用向后方程}} \\ &= 0 \end{aligned} \tag{4.14}$$

即完成证明。

4.2.5 马氏过程的无穷小描述

考虑一种由 $\mathfrak{R}$ 与 $\mathbf{P}$ 描述的随机过程（不预先假定马氏性），使得：

- 如果在某时间 t 此过程在状态 $i \in \mathbb{S}$，（不管之前的状态），则在很短时间 h 内，转移到别的状态的概率大约为 hR_i，
- 如果在某时间 t 此过程在状态 $i \in \mathbb{S}$，且在很短时间 h 内转移到别的状态上去，则在此条件下，（不管之前的状态），转移到状态 $j \neq i$ 的概率大约是 $(\mathbf{P})_{ij}$。

或者用更精确的数学描述：令 $\epsilon : \mathbb{R}_+ \to \mathbb{R}_+$ 为一个函数，满足 $\lim_{h\to 0}\epsilon(h) = 0$。上面两个条件表示为

$$\left|\mathbb{P}(X(t+h) \neq X(t) \mid X(\tau), \tau \in [0,t]) - hR_{X(t)}\right| \leq h\epsilon(h),$$
$$\left|\mathbb{P}(X(t+h) = j \mid X(\tau), \tau \in [0,t], \& X(t+h) \neq X(t)) - (\mathbf{P})_{X(t),j}\right| \leq \epsilon(h).$$

这两个不等式可以写成一个：对任意 $j \neq X(t)$：

$$\left|\mathbb{P}(X(t+h)=j \mid X(\tau), \tau \in [0,t]) - h(\mathbf{Q})_{X(t),j}\right| \leq h\epsilon(h)(R_{X(t)} + \epsilon(h)).$$

如果我们进一步要求 $\sup_{i\in\mathbb{S}} R_i < \infty$，也就是速率有界，则上面的公式等价于对任意 $j \in \mathbb{S}$，存在一个 $\epsilon' : \mathbb{R}_+ \to \mathbb{R}_+$ 且 $\lim_{h\to 0} \epsilon'(h) = 0$，使得

$$\left|\mathbb{P}(X+h)=j \mid X(\tau), \tau \in [0,t]) - \delta_{X(t),j} - h(\mathbf{Q})_{X(t),j}\right| \leq h\epsilon'(h). \quad (4.15)$$

下面我们证明，这个随机过程就是由 $\mathbf{P}(t) = e^{t\mathbf{Q}}$ 定义的马氏过程。

令 $s > 0$ 且 $A \in \sigma(\{X(\tau) : \tau \in [0,s]\})$。对每个 $t \geq 0$，定义 $\mathbb{S}$ 上的行向量 $\mu(t)$，使得对任意 $j \in \mathbb{S}$，$(\mu(t))_j = \mathbb{P}(\{X(s+t)=j\} \cap A)$。于是，又可以写成 $(\mu(t))_j = \mathbb{E}[\delta_{X(s+t),j}, A]$。对于 $h \geq 0$，我们有

$$\begin{aligned}
(\mu(t+h))_j &= \mathbb{E}[\delta_{X(s+t+h),j}, A] \\
&= \mathbb{E}\left[\underbrace{\mathbb{E}[\delta_{X(s+t+h),j} \mid X(\tau), \tau \in [0,s+t])}_{\sigma(\{X(\tau), \tau \in [0,s+t]\}) \text{ 上的随机变量}}, A\right] \\
&= \mathbb{E}[\mathbb{P}(X(s+t+h)=j \mid X(\tau), \tau \in [0,s+t]), A]
\end{aligned}$$

由(4.15)，我们有

$$\begin{aligned}
&|(\mu(t+h))_j - (\mu(t))_j - h(\mu(t)\mathbf{Q})_j| \\
&= |(\mu(t+h))_j - (\mu(t))_j - h\mathbb{E}[(\mathbf{Q})_{X(s+t),j}, A]| \\
&= |\mathbb{E}[\mathbb{P}(X(s+t+h)=j \mid X(\tau), \tau \in [0,s+t]), A] \\
&\quad - \mathbb{E}[\delta_{X(s+t),j}, A] - h\mathbb{E}[(\mathbf{Q})_{X(s+t),j}, A]| \\
&= |\mathbb{E}[\underbrace{\mathbb{P}(X(s+t+h)=j \mid X(\tau), \tau \in [0,s+t]) - \delta_{X(s+t),j} - h(\mathbf{Q})_{X(s+t),j}}_{\leq h\epsilon'(h)}, A]| \\
&\leq h\epsilon'(h).
\end{aligned}$$

所以，通过取 $h \to 0$ 的极限，对任意 $t \geq 0$，我们有 $\frac{d}{dt}\mu(t) = \mu(t)\mathbf{Q}$。于是，类似(4.14)，我们得到

$$\frac{d}{d\tau}\mu(\tau)e^{(t-\tau)\mathbf{Q}} = 0, \quad \tau \in (0,t).$$

于是，我们得到

$$\mathbb{P}(\{X(s+t)=j\} \cap A) = \mu(t) = \mu(0)e^{t\mathbf{Q}} = \mu(0)\mathbf{P}(t) = \mathbb{E}[(\mathbf{P}(t))_{X(s),j}, A],$$

等价于(4.5)。

4.3 遍历性质

4.3.1 状态的分类

仿照对于离散时间马氏链，我们先考虑对状态的分类。我们已经有了 $\mathbf{Q}=\mathbf{R}(\mathbf{P}-I)$。再定义 $\mathbf{P}^{\mathfrak{R}}$

$$(\mathbf{P}^{\mathfrak{R}})_{ij}=\begin{cases}(\mathbf{P})_{ij}, & R_i>0,\\ \delta_{ij}, & R_i=0.\end{cases}$$

显然 $\mathbf{P}^{\mathfrak{R}}$ 由 $\mathbf{Q}$ 完全决定，且 $\mathbf{Q}=\mathbf{R}(\mathbf{P}^{\mathfrak{R}}-I)$。根据(4.12)，我们有：如果

(1) 对某个 $t>0$ 使得 $(\mathbf{P}(t))_{ij}>0$，则

(2) 存在 $n\geq 0$ 使得 $(\mathbf{Q}^n)_{ij}>0$，也就得到

(1') 对与所有 $t>0$，$(\mathbf{P}(t))_{ij}>0$。接着可以验证，

(3) 存在 $n\geq 0$ 使得 $(\mathbf{Q}^n)_{ij}>0$ 等价于状态 i 相对于转移状态矩阵 $\mathbf{P}^{\mathfrak{R}}$ 可到达状态 j。

如果上面相互等价的条件满足，则我们称状态 i “**Q**-可到达” 状态 j，记为 $i\overset{\mathbf{Q}}{\to} j$。如果 $i\overset{\mathbf{Q}}{\to} i$ 且 $j\overset{\mathbf{Q}}{\to} i$，则称 i 与 j “**Q**-互通”，记为 $i\overset{\mathbf{Q}}{\leftrightarrow} j$。如果所有状态两两 **Q**-户通，则称 $\mathbb{S}$ 为 “**Q**-不可约” 的。

接着我们定义状态的常返性。如果 $R_i=0$（或者说 $i\in\mathbb{S}_0$），则马氏过程到了状态 i 就不会离开，所以此状态应该被认识为常返。在 $R_i>0$ 情况下，我们说 i 是常返态，或者说**Q**-常返态，如果 $\mathbb{P}(\sigma_i<\infty\mid X(0)=i)=1$，这里 $\sigma_j=\inf\{t\geq J_1: X(t)=i\}$（$J_1$ 是第一次跳跃时间）。不然的话，我们说 i 是**Q**-瞬时态。

对任意状态 $i\in\mathbb{S}$, **Q**-常返/瞬时与 $\mathbf{P}^{\mathfrak{R}}$ 这个状态转移矩阵定义的常返/瞬时态等价。如果 $R_i=0$，则显然 i 在两种定义下都常返，不然的话，i 是 **Q**-瞬时，当且仅当 **Q** 决定的连续时间马氏过程满足

$$\begin{aligned}&\mathbb{P}(\text{第一次跳跃没跳到 } \mathbb{S}_0 \text{ 且第二次跳跃跳到 } i\mid X_0=i)\\ &+\mathbb{P}(\text{前两次跳跃没跳到 } \mathbb{S}_0\cup\{i\} \text{ 且第三次跳跃跳到 } i\mid X_0=i)\\ &+\mathbb{P}(\text{前三次跳跃没跳到 } \mathbb{S}_0\cup\{i\} \text{ 且第四次跳跃跳到 } i\mid X_0=i)\\ &+\dots\end{aligned}$$

的和为 1。但是，上面求和中的各个概率，分别等于 $\mathbf{P}^{\mathfrak{R}}$ 决定的马氏链中的概率

$$
\begin{aligned}
&\mathbb{P}(X_1 \notin \mathbb{S}_0 \& X_2 = i \mid X_0 = i)\\
&\mathbb{P}(X_1, X_2 \notin \mathbb{S}_0 \cup \{i\} \& X_3 = i \mid X_0 = i)\\
&\mathbb{P}(X_1, X_2, X_3 \notin \mathbb{S}_0 \cup \{i\} \& X_4 = i \mid X_0 = i)\\
&\dots,
\end{aligned}
$$

而上面的概率求和等于 1 即为 $\mathbf{P}^{\mathfrak{R}}$ 决定的马氏链中 i 为常返的定义。所以我们得到结论。这个结论也告诉我们，$\mathbf{Q}$-常返/瞬时是（由 $\mathbf{Q}$ 决定的）互通类性质。

和离散时间马氏链中的结果定理2.3.2类似，对 $i \in \mathbb{S} \setminus \mathbb{S}_0$，我们也可以定义

$$
\sigma_i^{(m)} = \begin{cases} 0, & m = 0, \\ \sigma_i, & m = 1, \\ \infty, & m > 1 且 \sigma^{(m-1)} = \infty, \\ \inf\{t \geq J_{l+1} : X(t) = j\}, & m > 1 且 \sigma^{(m-1)} = J_l < \infty. \end{cases}
$$

（对于 $i \in \mathbb{S}_0$，可以直接令 $\sigma_i^{(0)} = 0$，$\sigma_i = \sigma_i^{(1)} = \sigma_i^{(2)} = \cdots = \infty$。）并且利用马氏性，证明

$$
\mathbb{P}(\sigma_i^{(m)} < \infty \mid X_0 = i) = \mathbb{P}(\sigma_i < \infty \mid X_0 = i)^m,
$$

以及

$$
\begin{aligned}
&\mathbb{E}\left[\int_{\sigma_i^{(m)}}^{\sigma_i^{(m+1)}} 1_{\{i\}}(X(t))dt, \sigma_i^{(m)} < \infty \middle| X_0 = i\right]\\
&= \mathbb{E}\left[\int_0^{\sigma_i} 1_{\{i\}}(X(t))dt \middle| X_0 = i\right] \mathbb{P}(\sigma_i^{(m)} < \infty \mid X(0) = i)\\
&= \mathbb{E}[J_1 \mid X_0 = i]\mathbb{P}(\sigma_i^{(m)} < \infty \mid X(0) = i)\\
&= \frac{\mathbb{P}(\sigma_i < \infty \mid X(0) = i)^m}{R_i},
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^{\infty} 1_{\{j\}}(X(t))dt \middle| X(0) = i\right] =\\
&\qquad \mathbb{E}\left[\int_0^{\infty} 1_{\{j\}}(X(t))dt \middle| X(0) = j\right] \mathbb{P}(\sigma_j < \infty \mid X(0) = i),
\end{aligned}
$$

并且因此

$$\mathbb{E}\left[\int_0^\infty 1_{\{j\}}(X(t))dt \middle| X(0)=i\right] = \frac{1}{R_j}\left(\delta_{ij} + \frac{\mathbb{P}(\sigma_j < \infty \mid X_0 = i)}{\mathbb{P}(\sigma_j = \infty \mid X_0 = j)}\right).$$

由此可得，

$$\mathbb{E}\left[\int_0^\infty 1_{\{i\}}(X(t))dt \middle| X(0)=i\right] = \infty$$

当且仅当

$$\mathbb{P}\left(\int_0^\infty 1_{\{i\}}(X(t))dt = \infty \middle| X(0)=i\right) = 1,$$

而

$$\mathbb{E}\left[\int_0^\infty 1_{\{i\}}(X(t))dt \middle| X(0)=i\right] < \infty$$

当且仅当

$$\mathbb{P}\left(\int_0^\infty 1_{\{i\}}(X(t))dt < \infty \middle| X(0)=i\right) = 1.$$

对于常返性，我们还有以下结论：

定理 4.3.1. *对于任意状态 $i \in \mathbb{S}$，以下结论等价：*

1. *i 为 $\mathbf{Q}$-常返/瞬时。*

2. *存在一个 $t \in (0,\infty)$ 使得 i 关于转移概率矩阵 $\mathbf{P}(t)$ 常返/瞬时。*

3. *对任意 $t \in (0,\infty)$ 使得 i 关于转移概率矩阵 $\mathbf{P}(t)$ 常返/瞬时。*

证明. 因为

$$\mathbb{E}\left[\int_0^\infty 1_{\{i\}}(X(t))dt \middle| X(0)=i\right] = \int_0^\infty (\mathbf{P}(t))_{ii}dt,$$

i 是 $\mathbf{Q}$-常返/瞬时的，或者 $\mathbf{P}(t)$-常返/瞬时的，当且仅当

$$\int_0^\infty (\mathbf{P}(t))_{ii}dt \quad \text{或者} \quad \sum_{n=0}^\infty (\mathbf{P}(t)^n)_{ii}$$

等于 ∞ 或者小于 ∞。利用 $(\mathbf{P}(h))_{ii} \geq \mathbb{P}(X(h)=i, \sigma_i > h \mid X(0)=i) = \mathbb{P}(J_1 > h \mid X(0)=i) = e^{-hR_i}$，我们有

$$(\mathbf{P}(t))_{ii} \geq (\mathbf{P}(t-s))_{ii}(\mathbf{P}(s))_{ii} \geq e^{-(t-s)R_i}(\mathbf{P}(s))_{ii}, \quad 0 \leq s < t.$$

于是对任意 $t > 0$，$n \in \mathbb{N}$ 和 $\tau \in [nt,(n+1)t)$，

$$\begin{aligned}(\mathbf{P}(t)^{n+1})_{ii} = (\mathbf{P}((n+1)t))_{ii} \geq e^{-tR_i}(\mathbf{P}(\tau))_{ii},\\ e^{-tR_i}(\mathbf{P}(t)^n)_{ii} = e^{-tR_i}(\mathbf{P}(nt))_{ii} \leq (\mathbf{P}(\tau))_{ii}.\end{aligned} \tag{4.16}$$

所以，对任意 $t>0$，

$$te^{-tR_i}\sum_{n=0}^{\infty}(\mathbf{P}(t)^n)_{ii}\le\int_0^{\infty}(\mathbf{P}(\tau))_{ii}d\tau\le te^{tR_i}\sum_{n=0}^{\infty}(\mathbf{P}(t)^{n+1})_{ii}.$$

至此我们证明结论。 □

4.3.2 平稳测度和极限定理

定理 4.3.2. *(a)* 对每个状态 $j\in\mathbb{S}$，$\lim_{t\to\infty}(\mathbf{P}(t))_{jj}$，存在。如果我们记此极限为 $\hat{\pi}_{jj}$，则对任意状态 $i\neq j$，

$$\lim_{t\to\infty}(\mathbf{P}(t))_{ij}=\mathbb{P}(\sigma_j<\infty\mid X(0)=i)\hat{\pi}_{jj}.$$

我们记上述极限为 $\hat{\pi}_{ij}$。

(b) 如果 $\hat{\pi}_{jj}>0$，则对任意与 $j\mathbf{Q}$-互通的状态 i，都有 $\hat{\pi}_{ii}>0$。

(c) 如果 $\hat{\pi}_{jj}>0$，令 $C=\{i\in\mathbb{S}:i\overset{\mathbf{Q}}{\leftrightarrow}j\}$，$\hat{\pi}^C$ 为行向量，各分量为 $(\hat{\pi}^C)_i=1_C(i)\hat{\pi}_i$。则对任意 $s>0$，$\hat{\pi}^C$ 是 $\mathrm{Stat}(\mathbf{P}(s))$ 中唯一的满足如 $k\notin C$，则 k 状态对应分量为 0 的概率向量/分布。

(d) 如果 $\mu\in\mathrm{Stat}(\mathbf{P}(s))$，则

$$\mu_j=\left(\sum_{i\overset{\mathbf{Q}}{\leftrightarrow}j}\mu_i\right)\hat{\pi}_{jj}.$$

证明这个定理之前，我们首先建立与(3.3)平行的更新方程：

$$(\mathbf{P}(t))_{ij}=e^{-tR_i}\delta_{ij}+\mathbb{E}[(\mathbf{P}(t-\sigma_j))_{jj},\sigma_j\le t\mid X(0)=i].\tag{4.17}$$

这个更新方程的左边可以写成

$$\mathbb{P}(X(t)=j,\sigma_j>t\mid X(0)=i)+\sum_{m=1}^{\infty}\mathbb{P}(X(t)=j,\sigma_j=J_m\le t\mid X(0)=i).$$

其中第一项对应右边的第一项，而后面对 m 的求和，用 $X(t)$ 的构造，其各项可以写为

$$\begin{aligned}&\mathbb{P}(\Phi^{\mathfrak{R},\mathbf{P}}(t-J_m;(E_{m+1},\dots,E_{m+n},\dots),(j,X_{m+1},\dots,X_{m+n},\dots))=j,\\&\qquad\sigma_j=J_m\le t\mid X_0=i)=\mathbb{E}[(\mathbf{P}(t-J_m))_{jj},\sigma_j=J_m\le t\mid X(0)=i].\end{aligned}$$

把所有的 $m\ge1$ 项都加起来，就得到(4.17)右边第二项。

定理*4.3.2*的证明. 第(a)部分的证明可以从离散时间马氏链的相关结果推出。

更新方程(4.17)的一个推论是：$(\mathbf{P}(s))_{ii} \geq e^{-sR_i} > 0$。所以，以 $\mathbf{P}(s)$ 为转移概率矩阵的马氏链在任意状态上都是非周期的。于是用(3.10)，我们得到

$$\lim_{n\to\infty}(\mathbf{P}(s)^n)_{ii} = \lim_{n\to\infty}(\mathbf{P}(ns))_{ii} = \pi(s)_{ii}$$

存在。

这些 $\pi(s)_{ii}$ 有可能依赖 s。但是容易看到，对任意 $m \in \mathbb{Z}_+$，$\pi(m^{-1})_{ii} = \pi(1)_{ii}$。下面我们证明，$\lim_{t\to\infty}(\mathbf{P}(t))_{ii}$ 存在且等于 $\pi(1)_{ii}$。利用和(4.16)类似的推导，对于任意 $t \in (n/m, (n+1)/m)$，都有

$$(\mathbf{P}(t))_{ii} \geq e^{-R_i/m}(\mathbf{P}(\frac{n}{m}))_{ii}, \quad (\mathbf{P}(t))_{ii} \leq e^{R_i/m}(\mathbf{P}(\frac{n+1}{m}))_{ii}.$$

因此，

$$\begin{aligned}
\liminf_{t\to\infty}((\mathbf{P}(t))_{ii} &\geq e^{-\frac{R_i}{m}} \lim_{n\to\infty}(\mathbf{P}(\frac{n}{m}))_{ii} = e^{-\frac{R_i}{m}}\pi(1)_{ii}, \\
\limsup_{t\to\infty}((\mathbf{P}(t))_{ii} &\leq e^{\frac{R_i}{m}} \lim_{n\to\infty}(\mathbf{P}(\frac{n}{m}))_{i+1,i+1} = e^{\frac{R_i}{m}}\pi(1)_{ii}.
\end{aligned}$$

令 $m \to \infty$，我们得到结论。所以，$\hat{\pi}_{ii}$ 就是 $\pi(1)_{ii}$。

对于 $i \neq j$，利用更新方程(4.17)即可证得：

$$\begin{aligned}
\lim_{t\to\infty}(\mathbf{P}(t))_{ij} &= \lim_{t\to\infty}\mathbb{E}[(\mathbf{P}(t-\sigma_j))_{jj}, \sigma_j \leq t \mid X(0) = i] \\
&\overset{\text{一下假定 } X(0)=j \text{ 条件}}{=} \lim_{t\to\infty}\int_0^\infty (\mathbf{P}(t-\sigma_j))_{jj} 1_{\sigma_j \leq t} d\sigma_j \\
&\underset{g(\sigma_j)=1}{\overset{\text{Lebesgue 控制收敛}}{=}} \int_0^\infty \lim_{t\to\infty}(\mathbf{P}(t-\sigma_j))_{jj} 1_{\sigma_j \leq t} d\sigma_j \\
&= \int_0^\infty \hat{\pi}_{jj} d\sigma_j = \hat{\pi}_{jj}\mathbb{P}(\sigma_j < \infty).
\end{aligned}$$

要证第(b)部分，我们注意到 $\hat{\pi}_{jj} = \pi(1)_{jj} > 0$，意味着 j 是以 $\mathbf{P}(1)$ 为转移概率矩阵的马氏链的正常返态。如果 $i \overset{\mathbf{Q}}{\leftrightarrow} j$，则对于 $\mathbf{P}(1)$ 也有 $i \leftrightarrow j$，于是 j 也是其正常返态，$\pi(1)_{ii} = \hat{\pi}_{ii} > 0$。

同样的思路可以证明第(c)和(d)部分。我们注意到，$C = \{i : i \overset{\mathbf{Q}}{\leftrightarrow} j\}$ 等价于由 $\mathbf{P}(s)$ 定义的 $\{i : i \leftrightarrow j\}$。对于第(c)部分，如果定义 $\pi(s)^C$ 为一个概率向量，且 $(\pi(s)^C)_i = 1_C(i)(\pi(s))_{ii}$，则显然 $\pi(s)^C$ 为唯一的 $\mathrm{Stat}(\mathbf{P}(s))$ 中满足要求的概率向量/分布。又因为 $\hat{\pi}^C = \pi(s)^C$，我们得到结论。对于第(d)部分，以 $\mathbf{P}(s)$ 为转移概率矩阵的马氏链相应的结果便可。 □

和离散马氏链类似，对于一个常返态 i，如果 $\hat{\pi}_{ii} > 0$，我们称 i 为正常返态，反之为零常返态，并且正常返性是 $\mathbf{Q}$-互通等价类性质。

接下来，我们有和定理3.2.2平行的平均遍历定理：

推论 4.3.1. 假设 j 是 $\mathbf{Q}$-常返的状态，$C = \{i \in \mathbb{S} : i \overset{\mathbf{Q}}{\leftrightarrow} j\}$，且马氏过程从 C 里面开始，也就是 $\mathbb{P}(X(0) \in C) = 1$，则

$$\lim_{t\to\infty} \mathbb{E}\left[\left(\frac{1}{T}\int_0^T 1_{\{j\}}(X(t))dt - \hat{\pi}_{jj}\right)^2\right] = 0.$$

证明. 此定理的证明与定理3.2.2的证明平行。首先，我们同样假设 $S = C$，把对于一般初始条件的证明归结为初始条件为 $X(0)$ 分布为 π^C 这个特殊的初始条件。

接着，定义函数 $f : \mathbb{S} \to \mathbb{R}$，使得 $f(i) = \delta_j(i) + \hat{\pi}_{jj}$。同样，我们用列向量 $\vec{f}$ 表示 f，且 $f_i = f(i)$。类似(3.9)，我们只需要证明在 $X(t)$ 的分布都为 $\hat{\pi}^C$ 这个条件下，

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{T}\int_0^T f(X(t))dt\right)^2\right] \\
&= \frac{2}{T^2}\int_0^T \left(\int_0^t \underbrace{\mathbb{E}[f(X(s))f(X(t))]}_{=\alpha(t):=\sum_{i\in C}(\hat{\pi}^C)_i f_i(\mathbf{P}(t)f)_i} ds\right) dt \\
&= \frac{2}{T^2}\int_0^T \left(\int_0^t \alpha(t-s)ds\right) dt \\
&= \frac{2}{T}\int_0^T \left(1 - \frac{t}{T}\right)\alpha(t)dt = 2\int_0^1 (1-s)\alpha(Ts)ds.
\end{aligned}$$

注意到对于任意 i，随着 $t \to \infty$，$((\mathbf{P}(t)f)_i = (\mathbf{P}(t))_{ij} - \pi_{jj} \to 0$。所以利用勒贝格控制收敛定理（控制函数 $\hat{(\pi)}_i^C$），我们知道 $\alpha(t) \to 0$。接着，又因为 $\alpha(t) \in [0,1]$，再用勒贝格控制收敛定理在 $u\int_0^1 (1-s)\alpha(Ts)ds$ 上（控制函数 $1-s$），得到最后的收敛结论。 □

4.3.3 $(\hat{\pi}^C)_{ij}$ 的另一重意义

回忆在离散时间情况下，我们得到了了 π_{ij} 用 $\mathbb{E}[\rho_j \mid X_0 = j]$ 和 $\mathbb{P}(\rho_j < \infty \mid X_0 = i$ 的表示。我们的工具是 $\mathbf{R}(s)$（$\mathbf{P}$ 的阿贝尔求和）和更新方程。在连续时间情况下，类似的推导仍然成立。和 $(\mathbf{R}(s))_{ij}$ 相对应的，是（注意，α 对应 $1-s$）

$$\mathbf{L}(\alpha)_{ij} = \alpha\mathbb{E}\left[\int_0^\infty e^{-\alpha t}1_{\{j\}}(X(t))dt \middle| X(0) = i\right] = \alpha\int_0^\infty e^{-\alpha t}(\mathbf{P}(t))_{ij}dt.$$

因为 $t\to\infty$ 时 $(\mathbf{P}(t))_{ij}\to\hat{\pi}_{ij}$，我们有 $\lim_{\alpha\searrow 0}\mathbf{L}(\alpha)_{ij}=\hat{\pi}_{ij}$。同时，由更新方程(4.17)，我们得到（假设条件 $X(0)=i$）

$$
\begin{aligned}
\mathbf{L}(\alpha)_{ii} &= \alpha\int_0^\infty e^{-\alpha t}\left(e^{-tR_i}+\mathbb{E}[(\mathbf{P}(t-\sigma_i))_{ii},\sigma_i\le t\mid X(0)=i]\right)dt\\
&= \alpha\left(\frac{1}{\alpha+R_i}+\int_0^\infty e^{-\alpha t}\int_0^t(\mathbf{P}(t-s))_{ii}d\sigma_i(s)dt\right)\\
&= \alpha\left(\frac{1}{\alpha+R_i}+\int_0^\infty\int_s^\infty(\mathbf{P}(t-s))_{ii}e^{-\alpha t}dtd\sigma_i(s)\right)\\
&= \alpha\left(\frac{1}{\alpha+R_i}+\int_0^\infty e^{-\alpha s}\int_s^\infty(\mathbf{P}(t-s))_{ii}e^{-\alpha(t-s)}dtd\sigma_i(s)\right)\\
&= \alpha\left(\frac{1}{\alpha+R_i}+\int_0^\infty e^{-\alpha s}\underbrace{\int_0^\infty(\mathbf{P}(t))_{ii}e^{-\alpha t}dt}_{\alpha^{-1}\mathbf{L}(\alpha)_{ii}}\,d\sigma_i(s)\right)\\
&= \alpha\left(\frac{1}{\alpha+R_i}+\alpha^{-1}\mathbf{L}(\alpha)_{ii}\underbrace{\int_0^\infty e^{-\alpha s}d\sigma_i(s)}_{\mathbb{E}[e^{-\alpha\sigma_i}\mid X(0)=i]}\right)\\
&= \frac{\alpha}{\alpha+R_i}+\mathbb{E}[e^{-\alpha\sigma_i}\mid X(0)=i]\mathbf{L}(\alpha)_{ii}.
\end{aligned}
$$

当 $R_i=0$ 时，$\mathbb{P}(\sigma_i=\infty\mid X(0)=i)=1$，所以对所有 $\alpha\in(0,1)$，我们有 $\mathbb{E}[e^{-\alpha\sigma_i}\mid X_0=i]=0$。当 $R_i>0$ 时，

$$
\begin{aligned}
\lim_{\alpha\searrow 1}\alpha^{-1}(1-\mathbb{E}[e^{-\alpha\sigma_i}\mid X(0)=i]) &= \lim_{\alpha\searrow 1}\mathbb{E}[\alpha^{-1}(1-e^{-\alpha\sigma_i})\mid X(0)=i]\\
&\overset{\text{Lebesgue 控制收敛}}{\underset{g(x)=\alpha^{-1}}{=}}\mathbb{E}\left[\lim_{\alpha\searrow 0}\alpha^{-1}(1-e^{-\alpha\sigma_i})\middle|X(0)=i\right]\\
&= \mathbb{E}[\sigma_i\mid X(0)=i].
\end{aligned}
$$

综上可得

$$
\hat{\pi}_{ii}=\begin{cases}1, & R_i=0,\\ \frac{1}{R_i\mathbb{E}[\sigma_i\mid X(0)=i]}, & R_i>0.\end{cases}
$$

再利用定理4.3.2的(a)部分，得到

$$
\hat{\pi}_{ij}=\begin{cases}\delta_{ij}+\mathbb{P}(\sigma_j<\infty\mid X(0)=i), & R_i=0,\\ \frac{\mathbb{P}(\sigma_j<\infty\mid X(0)=i)}{R_j\mathbb{E}[\sigma_j\mid X(0)=j]}, & R_j>0.\end{cases}
$$

最后，我们得到，i 是正常返态的充要条件是：$R_i=0$ 或 $\mathbb{E}[\sigma_i\mid X(0)=i]<\infty$。