Chapter 3

马氏链的遍历理论

3.1 状态的分类

之前我们已经定义了从一个状态“可到达”另一个：$i \to j$，意思是从状态 i 经过若干步，马氏链可以走到状态 j，或者等价地，对于某个 $n \geq 0$，$(\mathbf{P}^n)_{ij} > 0$。状态间的可到达性满足直观的性质（证明作为习题）：

$$i \to j, \quad j \to k \implies i \to k.$$

如果 $i \to j$ 并且 $j \to i$，我们说 i,j 这两个状态*互通*，记作 $i \leftrightarrow j$。不难发现，互通是一种等价关系，因为（1）对任意状态 i，$i \leftrightarrow i$，（2）如果 $i \leftrightarrow j$，则 $j \leftrightarrow i$，以及（3）由 $i \leftrightarrow j$，$j \leftrightarrow k$ 可得 $i \leftrightarrow k$。于是，根据这个等价关系，可以把状态空间 $\mathbb{S}$ 划分成互不相交的等价类：$[i] = \{j \in \mathbb{S} : j \leftrightarrow i\}$。如果整个状态空间就是一个等价类（也就是说，所有状态都两两互通），则我们称这个马氏链为*不可约的*。

3.1.1 状态分类：常返和瞬时

我们在这一节证明，如果一个状态 i 是常返或瞬时的，则所有于它互通的状态也是。也就是说，常返/瞬时状态是定义在等价类上的。

定理 3.1.1. *如果 i 是常返态，$j \neq i$ 是另一个状态，则*

1. $i \to j$ 当且仅当 $\mathbb{P}(\rho_j < \rho_i \mid X_0 = i) > 0$。

2. 如果 $i \to j$，则

$$\mathbb{P}(\rho_j < \infty \mid X_0 = i) = \mathbb{P}(\rho_i < \infty \mid X_0 = j) = \mathbb{P}(\rho_j < \infty \mid X_0 = j) = 1,$$

尤其是 $i \leftrightarrow j$ 且 j 是常返的。

证明. 我们首先证明第1部分。由 $\mathbb{P}(\rho_j < \rho_i \mid X_0 = i) > 0$ 推出 $i \to j$ 是显然的。下面我们证明如果 $\mathbb{P}(\rho_i > \rho_j \mid X_0 = i) = 0$，则 $i \not\to j$，思路大体是：如果 i 返回前一定走不到 j，则第二次，第三次，——返回前都到不了 j。我们使用之前引进的 $F_{N,i}(k_0, \ldots, k_N)$ 记号。定义

$$G_n(k_0, \ldots, k_n) = \underbrace{[F_{n-1,i}(k_0, \ldots, k_{n-1}) - F_{n,i}(k_0, \ldots, k_n)]F_{n,j}(k_0, \ldots, k_n)}_{\text{当 } X_{k_0}, \ldots, X_{k_n} \neq j\text{，} X_{k_0}, \ldots, X_{k_{n-1}} \neq i \text{ 且 } X_{k_n} = i \text{ 时，才为 } 1\text{。}}$$

我们跟之前一样定义 $\{\rho_j^{(m)} : m \geq 0\}$，于是

$$\begin{aligned}
&\mathbb{P}(\rho_i^{(m+1)} < \rho_j \mid X_0 = i) \\
&= \sum_{l=1}^{\infty} \mathbb{P}(\rho_i^{(m)} = l \& \rho_i^{(m+1)} < \rho_j \mid X_0 = i) \\
&= \sum_{l=1}^{\infty} \sum_{n=1}^{\infty} \mathbb{P}(\rho_i^{(m)} = l, \rho_i^{(m+1)} = l + n < \rho_j \mid X_0 = i) \\
&= \sum_{l,n=1}^{\infty} \mathbb{E}[G_n(X_l, \ldots, X_{l+n}), \rho_i^{(m)} = l < \rho_j \mid X_0 = i] \\
&= \sum_{l,n=1}^{\infty} \mathbb{E}[G_n(X_l, \ldots, X_{l+n}) \mid \rho_i^{(m)} = l < \rho_j, X_0 = i] \\
&\qquad \times \mathbb{P}(\rho_i^{m)} = l < \rho_j \mid X_0 = i) \\
&= \sum_{l,n=1}^{\infty} \mathbb{E}[G_n(X_0, \ldots, X_n) \mid X_0 = i] \mathbb{P}(\rho_i^{m)} = l < \rho_j \mid X_0 = i) \\
&= \sum_{l,n=1}^{\infty} \mathbb{P}(\rho_i = n < \rho_j \mid X_0 = i) \mathbb{P}(\rho_i^{(m)} = l < \rho_j \mid X_0 = 0) \\
&= \mathbb{P}(\rho_i < \rho_j \mid X_0 = i) \mathbb{P}(\rho_i^{(m)} < \rho_j \mid X_0 = i).
\end{aligned}$$

所以，对于任意 $j \neq i$，我们有

$$\mathbb{P}(\rho_i^{(m)} < \rho_j \mid X_0 = i) = \mathbb{P}(\rho_i < \rho_j \mid X_0 = i)^m.$$

因为我们假设了 $\mathbb{P}(\rho_j < \rho_i \mid X_0 = i) = 0$，对于所有的 $m \geq 1$，$\mathbb{P}(\rho_i^{(m)} < \rho_j \mid X_0 = i) = 1$。因为 $\rho^{(m)} \geq m$，我们得到 $\mathbb{P}(\rho_j = \infty \mid X_0 = i) = 1$，也就是 $i \not\to j$。

接着考虑定理的第2部分。如果 $i \to j$，

$$
\begin{aligned}
&\mathbb{P}(\rho_j < \rho_i < \infty \mid X_0 = i) \\
=& \lim_{n\to\infty} \sum_{m=1}^{\infty} \mathbb{P}(\rho_j = m < \rho_i \le m+n \mid X_0 = i) \\
=& \lim_{n\to\infty} \sum_{m=1}^{\infty} \mathbb{E}[\underbrace{1 - F_{n,i}(X_m, \ldots, X_{m+n})}_{X_m,\ldots,X_{m+n}\text{ 中没有 } i}, \rho_j = m < \rho_i \mid X_0 = i] \\
=& \lim_{n\to\infty} \sum_{m=1}^{\infty} \mathbb{P}(\rho_j = m < \rho_i \mid X_0 = i)\mathbb{E}[1 - F_{n,i}(X_0, \ldots, X_n) \mid X_0 = j] \\
=& \sum_{m=1}^{\infty} \mathbb{P}(\rho_j = m < \rho_i \mid X_0 = i)\mathbb{P}(\rho_i < \infty \mid X_0 = j) \\
=& \mathbb{P}(\rho_j < \rho_i \mid X_0 = i)\mathbb{P}(\rho_i < \infty \mid X_0 = j).
\end{aligned}
$$

利用已知的 $\mathbb{P}(\rho_i < \infty \mid X_0 = i) = 1$，可得

$$
\underbrace{\mathbb{P}(\rho_j < \rho_i \mid X_0 = i)}_{>0} = \underbrace{\mathbb{P}(\rho_j < \rho_i \mid X_0 = i)}_{>0}\mathbb{P}(\rho_i < \infty \mid X_0 = j),
$$

所以 $\mathbb{P}(\rho_i < \infty \mid X_0 = j) = 1$。由此可知，$j \to i$，也就是 $i \leftrightarrow j$。

类似地，

$$
\begin{aligned}
&\mathbb{P}(\rho_j < \infty \mid X_0 = i) \\
=& \underbrace{\mathbb{P}(\rho_j < \rho_i \mid X_0 = i)}_{\rho_i\text{几乎处处}<\infty} + \mathbb{P}(\rho_i < \rho_j < \infty \mid X_0 = i) \\
=& \underbrace{\mathbb{P}(\rho_j < \rho_i \mid X_0 = i) + \mathbb{P}(\rho_i < \rho_j \mid X_0 = i)}_{\text{和}=1}\mathbb{P}(\rho_j < \infty \mid X_0 = i).
\end{aligned}
$$

可以推导出

$$
\mathbb{P}(\rho_j < \infty \mid X_0 = i)\underbrace{\mathbb{P}(\rho_j < \rho_i \mid X_0 = i)}_{>0} = \mathbb{P}(\rho_j < \rho_i \mid X_0 = i).
$$

所以 $i \to j$ 可以推出 $\mathbb{P}(\rho_j < \infty \mid X_0 = i) = 1$。

最后，我们应用

$$
\mathbb{P}(\rho_i < \rho_j < \infty \mid X_0 = j) = \mathbb{P}(\rho_j < \infty \mid X_0 = i)\mathbb{P}(\rho_i < \rho_j \mid X_0 = j)
$$

和（因为 $i \leftrightarrow j$）$\mathbb{P}(\rho_i < \infty \mid X_0 = j) = 1$，$\mathbb{P}(\rho_i = \rho_j \mid X_0 = j) \le \mathbb{P}(\rho_i = \infty \mid X_0 = j) = 0$，

$$\begin{aligned} & \mathbb{P}(\rho_j < \infty \mid X_0 = j) \\ \overset{\rho_i \text{几乎处处} < \infty}{=} & \mathbb{P}(\rho_j < \rho_i \mid X_0 = j) + \mathbb{P}(\rho_i < \rho_j < \infty \mid X_0 = j) \\ \overset{\mathbb{P}(\rho_i = \rho_j \mid X_0 = j) = 0}{=} & \mathbb{P}(\rho_j < \rho_i \mid X_0 = j) + \mathbb{P}(\rho_j < \infty \mid X_0 = i)\mathbb{P}(\rho_i < \rho_j \mid X_0 = j) \\ = & 1. \end{aligned}$$

□

推论 3.1.1. 如果 $i \leftrightarrow j$，则 j 是常返/瞬时，当且仅当 i 是常返/瞬时。如果 $i \leftrightarrow j$ 且 i 是常返的，$\mathbb{P}(\rho_j < \infty \mid X_0 = i) = 1$，如果 $i \not\leftrightarrow j$（即 $i \not\to j$ 或 $i \not\leftarrow j$）且 i 是常返的，则 $\mathbb{P}(\rho_j < \infty \mid X_0 = i) = 0$，也就是说，对于所有 $n \ge 0$，$(\mathbf{P}^n)_{ij} = 0$，或者说 $i \not\to j$。

因为常返/瞬时性只与状态所在的互通类有关，所以，如果一个马氏连是可约的，则它的所有状态或者全部常返，或者全部瞬时。这时，我们就称这个马氏连为常返的或者瞬时的。

3.1.2 常返性/瞬时性的判别法则

定义状态空间 $\mathbb{S}$ 上的一个函数 $u : \mathbb{S} \to \mathbb{R}$，在 $\mathbb{S} \subseteq \mathbb{Z}_+$ 这个假定下，我们记 $u_i = u(i)$，并且用列向量 $\vec{u} = (u_i)$ 来表示 u 这个函数。

首先我们介绍一个技术性结果：

定理 3.1.2. 如果 u 是一个 $\mathbb{S}$ 上的非负函数，且对所有 $i \in \mathbb{S}$，$(\mathbf{P}\vec{u})_i \le u_i$。则如果对某个 $j \in \mathbb{S}$，$(\mathbf{P}\vec{u})_j < u_j$，则 j 是瞬时的。

证明. 令 $\vec{f} = \vec{u} - \mathbf{P}\vec{u}$。则 $f_j > 0$。我们有，对于所有 $n \ge 1$，

$$\begin{aligned} u_j & \ge u_j - (\mathbf{P}^n\vec{u})_j \\ & = \sum_{m=0}^{n-1} (\mathbf{P}^m\vec{u})_j - (\mathbf{P}^{m+1}\vec{u})_j = (\mathbf{P}^m(\vec{u} - \mathbf{P}\vec{u}))_j \\ & = \sum_{m=0}^{n-1} (\mathbf{P}^m\vec{f})_j \\ & \underbrace{\ge}_{\text{利用 } f_i \text{ 的非负性}} \sum_{m=0}^{n-1} (\mathbf{P}^m)_{jj} f_j. \end{aligned}$$

于是，$\sum_{m=0}^{\infty}(\mathbf{P}^m)_{jj} \le u_j/f_j < \infty$。注意到 $\sum_{m=0}^{\infty}(\mathbf{P}^m)_{jj} = \mathbb{E}[T_j \mid T_0 = j]$，由已知结论(2.11)，状态 j 是瞬时的。 □

下面我们介绍一个更有概率意义的结果。首先我们介绍一个杜布停时（Doob stoping time）定理的特例：

引理 3.1.1. 假设 $u:\mathbb{S}\to\mathbb{R}$ 是下有界函数(即存在 $C > -\infty$ 使对所有 $i\in\mathbb{S}$，$u_i > C$)，Γ 是 $\mathbb{S}$ 的非空子集，并且定义停时 $\rho_\Gamma = \inf\{n\ge 1: X_n\in\Gamma\}$。如果对所有 $i\notin\Gamma$，$(\mathbf{P}\vec{u})_i \le u_i$，则对所有 $n\ge 0$ 和所有 $i\in\mathbb{S}\setminus\Gamma$，

$$\mathbb{E}[u(X_{n\wedge\rho_\Gamma}) \mid X_0 = i] \le u_i.$$

另外，如果假设中的不等式改为 $(\mathbf{P}\vec{u})_i \le u_i$，则结论变为 $\mathbb{E}[u(X_{n\wedge\rho_\Gamma} \mid X_0 = i] \le u_i$。

证明此引理之前，我们先介绍一下：引理中不等式（等式）满足的情况下，$\{u(X_n)\}$ 是一列下鞅（鞅）。

证明. 我们只证明不等式情况，因为等式情况同理。令 $A_n = \{\rho_\Gamma > n\}$ 为一个事件，则 A_n 可以用 $X_0,\dots,X_n$ 表示（或者说，A_n 在 $X_0,\dots,X_n$ 生成的 σ-代数中）。对任意 $i\notin\Gamma$，

$$\begin{aligned}
&\mathbb{E}[u(X_{(n+1)\wedge\rho_\Gamma}) \mid X_0 = i] \\
&= \mathbb{E}[\underbrace{u(X_{n\wedge\rho_\Gamma})}_{\text{在 } A_n^c \text{ 上，} (n+1)\wedge\rho_\Gamma = n\wedge\rho_\Gamma}, A_n^c \mid X_0 = i] \\
&\quad + \sum_{k\notin\Gamma}\mathbb{E}[u(X_{n+1}, A_n\cap\{X_n = k\} \mid X_0 = i] \\
&= \mathbb{E}[u(X_{n\wedge\rho_\Gamma}), A_n^c \mid X_0 = i] \\
&\quad + \sum_{k\notin\Gamma}\underbrace{(\mathbf{P}\vec{u})_k}_{\text{当 } X_n = k \text{ 时，} \mathbb{E}[X_{n+1}] \text{ 的值}}\mathbb{P}(A_n\cap\{X_n = k\} \mid X_0 = i) \\
&\le \mathbb{E}[u(X_{n\wedge\rho_\Gamma}), A_n^c \mid X_0 = i] + \sum_{k\notin\Gamma}\underbrace{u_k\mathbb{P}(A_n\cap\{X_n = k\} \mid X_0 = i)}_{\substack{=\mathbb{E}[u(X_n), A_n\cap\{X_n=k\}\mid X_0=i] \\ =\mathbb{E}[u(X_{n\wedge\rho_\Gamma}), A_n\cap\{X_n=k\}\mid X_0=i]}} \\
&= \mathbb{E}[u(X_{n\wedge\rho_\Gamma}) \mid X_0 = i].
\end{aligned}$$

□

定理 3.1.3. 令 $j\in\mathbb{S}$ 且令 $C = \{i : i\leftrightarrow j\}$。

1. 假设 j 是常返的。如果 $u:\mathbb{R}_+$ 是有界函数，且对任意 $i\in C\setminus\{j\}$，

(i) $u_i=(\mathbf{P}\vec{u})_i$ 或者

(ii) $u_j\ge u_i\ge(\mathbf{P}\vec{u})_i$，

则 u 在 C 上为常数。

2. 假设 j 是瞬时的。则对任意 $i\in\mathbb{S}\setminus\{j\}$，由

$$u(i)=\begin{cases}1, & i=j,\\ \mathbb{P}(\rho_j<\infty\mid X_0=i), & i\neq j\end{cases}$$

是方程 $u(i)=(\mathbf{P}\vec{u})_i=\sum_{k\in\mathbb{S}}(\mathbf{P})_{ik}u(k)$ 的有界非常数解。

证明. 首先我们证明第1部分。我们应用引理3.1.1，并令 $\Gamma=\{j\}\cup(\mathbb{S}\setminus C)$。根据推论3.1.1，由 $X_0=i\in C$ 出发的马氏链到达 $\mathbb{S}\setminus C$ 的概率为 0。如果条件(i)成立，则对于 $i\in C\setminus\{j\}$ 和任意 $n>0$，

$$u(i)=u(j)\mathbb{P}(\rho_j\le n\mid X_0=i)+\mathbb{E}[u(X_n),\rho_j>n\mid X_0=i].$$

因为随着 $n\to\infty$，$\mathbb{P}(\rho_j\le n\mid X_0=i)\to 1$，并且因为 u 是有界的，上式取 $n\to\infty$ 极限后得到 $u(i)=u(j)$。如果条件(ii)成立，则有

$$u(j)\ge u(i)\ge u(j)\mathbb{P}(\rho_j\le n\mid X_0=i)+\mathbb{E}[u(X_0),\rho_j>0\mid X_0=i],$$

取 $n\to\infty$ 极限后得到 $u(j)\ge u(i)\ge u(j)$，也得到 $u(i)=u(j)$。

然后我们证明2部分。因为 j 是瞬时的，我们有 $\mathbf{P}_{jj}<1$。如果 u 是由定理中表达式给出的，则

$$\begin{aligned}1>\mathbb{P}(\rho_j<\infty\mid X_0=j)&=(\mathbf{P})_{jj}+\underbrace{\sum_{i\neq j}(\mathbf{P})_{j,i}u(i)}_{=\sum_{i\neq j}\mathbb{P}(X_1=i,\rho_j<\infty\mid X_0=j)}\\ &\ge(\mathbf{P})_{jj}+(1-\mathbf{P}_{jj})\inf_{i\neq j}u(i).\end{aligned}$$

于是可以看出，$\inf_{i\neq j}u(i)<1$，所以 $u(i)$ 作为 i 的函数，不是平凡的。接着可以验证

$$u(i)=\mathbb{P}(\rho_j<\infty\mid X_0=i)=\mathbf{P}_{ij}+\underbrace{\sum_{k\neq j}\mathbf{P}_{ik}\mathbb{P}(\rho_j<\infty\mid X_0=k)}_{=\sum_{k\neq j}\mathbb{P}(X_1=k,\rho_j<\infty\mid X_0=i)}=(\mathbf{P}\vec{u})_i.$$

□

定理 3.1.4. 令 $\{B_m : m \geq 0\}$ 为 $\mathbb{S}$ 的一列子集，且 $B_0 \subseteq B_1 \subseteq \cdots \subseteq B_n \subseteq \cdots$。假设存在 $j \in B_0$，使得对所有 $m \geq 0$，

$$\mathbb{P}(\exists n \in \mathbb{N}\ X_n \notin B_m \mid X_0 = j) = 1.$$

假设存在一个非负函数 u，满足对所有 $(\mathbf{P}\vec{u})_i \leq u_i$ 且随着 $m \to \infty$ 我们有 $a_m := \inf_{i \notin B_m} u_i \to \infty$，则 j 为常返态。

证明. 对任意 $m \geq 0$，令 $\Gamma_m = \{j\} \cup B_m^c$，$\tau_m = \inf\{n \geq 1 : X_n \notin B_m\}$。(我们注意到，在 $X_0 = j$ 条件下，τ_m 几乎处处有限。) 然后令 $\rho_{\Gamma_m} = \inf\{X_n \in \Gamma_m\} = \rho_j \wedge \tau_m$。由引理3.1.1，对任意 n, m，

$$u_j \geq \mathbb{E}[u(X_{n\wedge\rho_{\Gamma_m}}) \mid X_0 = j] \geq a_m \underbrace{\mathbb{P}(\tau_m \leq n \wedge \rho_j \mid X_0 = j)}_{\substack{\text{走到 } B_m^c \text{ 停住，而不是走到 } j \text{ 停住，} \\ \text{或者一直走到 } n \text{ 停不下来的概率}}} .$$

取 $n \to \infty$ 的极限，我们得到对任意 m，$u_j \geq a_m \mathbb{P}(\tau_m \leq \rho_j \mid X_0 = j)$。于是，再取 $m \to \infty$ 的极限，我们得到 $\lim_{m\to\infty} \mathbb{P}(\tau_m \leq \rho_j \mid X_0 = j) = 0$。最后，因为

$$\mathbb{P}(\rho_j < \infty \mid X_0 = j) \geq \mathbb{P}(\rho_j < \tau_m \mid X_0 = j) = 1 - \underbrace{\mathbb{P}(\tau_m \leq \rho_j \mid X_0 = j)}_{\text{随着 } m \text{ 增大 } \nearrow 1}$$

可得 j 为常返态。 □

推论 3.1.2. 假设 $\mathbf{P}$ 是不可约的。令 $\{F_m : m \geq 0\}$ 为 $\mathbb{S}$ 中一列非空、有限子集，且 $F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots$。如果 $j \in F_0$ 且存在 $\mathbb{S}$ 上的非负函数 u，使得对不等于 j 的状态 i，都有 $(\mathbf{P}\vec{u})_i \leq u_i$ 以及 $\inf_{i \notin F_m} u_i \to \infty$，则 j 是常返的。

证明. 根据上面定理3.1.4，我们只需要证明对于任意 m，$\mathbb{P}(\exists n \in \mathbb{N}\ X_n \notin F_m \mid X_0 = j) = 1$。为此，令 $\tau_m = \inf\{n \geq 1 : X_n \notin F_m\}$。因为此马氏链是不可约的，从 j 总能到达 F_m^c，对任意 m 和 $i \in \mathbb{S}$，$\mathbb{P}(\tau_m < \infty \mid X_0 = i) > 0$。因为 F_m 是个有限集合，存在 $\theta_m \in (0,1)$ 和 $N_m \geq 1$ 使得对于任意 $i \in F_m$，$\mathbb{P}(\tau_m > N_m \mid X_0 = i) \leq \theta_m$。这个意味着

$$\begin{aligned}
&\mathbb{P}(\tau_m > (l+1)N_m \mid X_0 = j) \\
&= \sum_{i \in F_m} \mathbb{P}(\tau_m > (l+1)N_m \& X_{lN_m} = i \mid X_0 = j) \\
&= \sum_{i \in F_m} \mathbb{P}(\tau_m > N_m \mid X_0 = i) \underbrace{\mathbb{P}(\tau_m > lN_m \& X_{lN_m} = i \mid X_0 = j)}_{\text{前 } lN_m \text{ 步}} \\
&\leq \theta_m \mathbb{P}(\tau_m > lN_m \mid X_0 = j).
\end{aligned}$$

于是，$\mathbb{P}(\tau_m > lN_m \mid X_0 = j) \le \theta_m^l$，也就意味着 $\mathbb{P}(\tau_m = \infty \mid X_0 = j) = 0$，与之前的假设矛盾。 □

本节给出的结果只是一些零散的例子。他们的共同特征是，找到一个 $\mathbb{S}$ 上的函数 u，使得某种意义下 $\mathbf{P}\vec{u} = \vec{u}$ 或者 $(\mathbf{P}\vec{u})_i \le u_i$，也就是说，$\{u(X_n)\}$ 构成鞅或者上鞅。本节中结果的证明在学了鞅论之后可以简化。

3.1.3 周期性

对任意一个状态 $i \in \mathbb{S}$，令

$$S(i) = \{n \ge 0 : (\mathbf{P}^n)_{ii} > 0\} \quad \text{以及} \quad d(i) = \gcd(S(i)).$$

$d(i)$ 是状态 i 的周期。如果 $d(i) = 1$，这个状态是非周期的。(周期可以为 ∞，当且仅当 $S(i) = \{0\}$ 时。)

周期是互通类的性质：如果 $i \leftrightarrow j$，则 $d(i) = d(j)$。对此我们只要证明 $d(i) \le d(j)$ 同时 $d(j) \le d(i)$。假设 $d(i) = k$，且 $\min\{n \ge 1 : (\mathbf{P}^n)_{i,j} > 0\} = a$，$\min\{n \ge 1 : (\mathbf{P}^n)_{j,i} > 0\} = b$。则 $(\mathbf{P}^{a+b})_{i,i} \ge (\mathbf{P}^a)_{i,j}(\mathbf{P}^b)_{j,i} > 0$，所以 $a + b \in S(i)$。于是 $k \mid (a + b)$。对每个 $m \in S(i)$，我们有 $(\mathbf{P}^{b+m+a})_{j,j} \ge (\mathbf{P}^b)_{j,i}(\mathbf{P}^m)_{i,i}(\mathbf{P}^n)_{i,j} > 0$，所以 $(a+b)+m \in S(j)$。我们有 $\gcd((a+b)+S(i)) = \gcd(S(i)) = k$，而 $(a+b)+S(i)$ 是 $S(j)$ 的一个子集，所以 $\gcd(S(j)) \le k = \gcd(S(i))$。一个不等式得证。另一个不等式的证明完全平行，这里略去。

一个比较难的结果是：如果状态 i 的周期有限，则对所有充分大的 $n \in \mathbb{Z}^+$，$(\mathbf{P}^{nd(i)})_{ii} > 0$。特别地，如果 i 是非周期的，则对所有充分大的 n，有 $(\mathbf{P}^n)_{ii} > 0$。这个结果依赖于一个数论定理：

定理 3.1.5. 给定 $\emptyset \ne S \subseteq \mathbb{Z}^+$，则 $\gcd(S) \le \min(S)$，且等式成立当且仅当 $\gcd(S) \in S$。即使等式不成立，也可以找到有限个 $s_1, \ldots, s_n \in S$ 以及 $a_1, \ldots, a_n \in \mathbb{Z}$ 使得 $\gcd(S) = a_1 s_1 + \cdots + a_n s_n$。最后，如果 S 满足以下条件：如果 $a, b \in S$，则 $a + b \in S$，那么存在一个 $M \in \mathbb{Z}^+$，使得 $\{s \in S : s \ge M \gcd(S)\} = \{m \gcd(S) : m \ge M\}$。

因为此定理为纯粹数论命题，证明从略。因为对任意状态 i，我们有：如果 $a, b \in S(i)$，则 $(\mathbf{P}^a)_{ii} > 0$，$(\mathbf{P}^b)_{ii} > 0$，于是 $(\mathbf{P}^{a+b})_i i \ge (\mathbf{P}^a)_{ii}(\mathbf{P}^b)_{ii} > 0$，也即 $a + b \in S(i)$。所以可以应用上述定理，证明之前陈述的结论。

推论 3.1.3. 假设 $\mathbf{P}$ 是一个有限状态空间 $\mathbb{S}$ 上的转移概率矩阵。如果有一个状态 $j_0 \in \mathbb{S}$ 是非周期的，且其他所有状态 $i \in \mathbb{S}$ 都满足 $i \to j_0$，则存在

$M \in \mathbb{Z}^+$ 和 $\epsilon > 0$，使得对所有 i，$(\mathbf{P}^M)_{i,j_0} \geq \epsilon$。这样，对任意初始分布 μ，

$$\|\vec{\mu}\mathbf{P}^n - \pi\|_1 \leq 2(1-\epsilon)^{\lfloor \frac{n}{M} \rfloor}.$$

证明. 因为 j_0 是非周期的，存在一个 $M_0 \in \mathbb{Z}^+$ 使得对于所有的 $n \geq M$，$(\mathbf{P}^n)_{ii} > 0$。有因为每个 i 都可到达 j_0，所以存在 $m(i)$ 使得 $(\mathbf{P}^{m(i)})_{i,j_0} > 0$。于是，可以看出来如果取 $M = M_0 + \max_{i \neq j} m(i)$，则对所有 $i \in \mathbb{S}$，$(\mathbf{P})_{i,j_0} > 0$。这样的话，取 $\epsilon = \min_{i \in \mathbb{S}} (\mathbf{P})_{i,j_0}$，就可以证明结论（最后的不等式需要(2.4)）。 □

3.2 没有德布林条件的遍历论

3.2.1 矩阵的范数与收敛性

我们以前使用了两个向量范数 $\|\cdot\|_1$ 与 $\|\cdot\|_\infty$，并且前一个只用于行向量（用来代表分布），后一个只用于列向量（用来代表 $\mathbb{S}$ 上的函数）。下面我们介绍一种（有限或无限维）矩阵的算子范数：

$$\|M\|_{\infty,\infty} = \sup\{\|Mf\|_\infty : \|f\|_\infty \leq 1\}.$$

如果在 $\mathbb{S} \subseteq \{1, 2, \ldots\}$ 上，M 表示为 $M_{i,j}$，则可以看到

$$\begin{aligned}
\|M\|_{\infty,\infty} &= \sup_{\|f\|_\infty \leq 1} \|Mf\| = \sup_{\|f\|_\infty \leq 1} \sup_{i \in \mathbb{S}} \left| \sum_{j \in \mathbb{S}} M_{ij} f_j \right| \\
&\leq \sup_{\|f\|_\infty \leq 1} \sup_{i \in \mathbb{S}} \|M_{ij}\|_1 \|f_j\|_\infty = \sup_{i \in \mathbb{S}} \|M_{ij}\|_1 \\
&= \sup_{i \in \mathbb{S}} \sum_{j \in \mathbb{S}} |M_{ij}|.
\end{aligned}$$

另一方面，可以证明 $\|M\|_{\infty,\infty} \geq \sum_{j \in \mathbb{S}} |M_{ij}|$（作为习题），所以我们有

$$\|M\|_{\infty,\infty} = \sup_{i \in \mathbb{S}} \sum_{j \in \mathbb{S}} |M_{ij}|.$$

最后，任意一个抽象的，把 $\ell^\infty(\mathbb{S})$ 映射到自身上的有界算子 M，都可以表示为 M_{ij} 形式，且

$$M_{ij} = M(e_j) \text{ 的第 } i \text{ 个分量}, \quad e_j = (0, \ldots, 0, \underbrace{1}_{\text{第 } j \text{ 个分量}}, 0, \ldots). \tag{3.1}$$

所以,把 $\ell^\infty(\mathbb{S})$ 映射到自身上的有界算子的线性空间,就是满足 $\sup_{i\in\mathbb{S}}\sum_{j\in\mathbb{S}}|M_{ij}|<\infty$ 的矩阵空间。

对于算子范数，我们有从反函分析得来的性质，比如这个有界算子空间的完备性，以及 $\|MM'\|_{\infty,\infty}\le\|M\|_{\infty,\infty}\|M'\|_{\infty,\infty}$。

3.2.2 阿贝尔收敛性

我们考虑过 $(\mathbf{P}^n)_{ij}$ 在 $n\to\infty$ 时的收敛性，也考虑过比它更弱的 $\mathbf{A}_n=\frac{1}{n}\sum_{m=0}^{n-1}(\mathbf{P}^m)_{ij}$ 的收敛性。（$\mathbf{A}_n$ 收敛，也叫做 $(\mathbf{P}^n)_{ij}$ 的切萨罗（Césaro）收敛。）下面我们介绍一种更弱的收敛性：阿贝尔（Abel）收敛。

如果一个有界数列 $\{x_n\}_0^\infty$ 满足

$$\lim_{s\nearrow 1}(1-s)\sum_{n=1}^{\infty}s^n x_n=x,$$

我们称 $\{x_n\}_0^\infty$ 阿贝尔收敛到极限 x。容易证明，如果 $\{x_n\}_1^\infty$ 收敛到极限 x，则其阿贝尔极限也为 x，反之，阿贝尔极限存在，数列极限，甚至切萨罗极限，不见得存在。(举例作为习题。)

对于 $\mathbb{S}$ 上的概率转移矩阵 $\mathbf{P}$，我们定义

$$\mathbf{R}(s)=(1-s)\sum_{n=0}^{\infty}s^n\mathbf{P}^n,\quad s\in[0,1).$$

对于 $s\in[0,1)$，$\mathbf{R}(s)$ 存在且唯一。因为 $\|\mathbf{P}\|_{\infty,\infty}=1$，所以 $\sum_{n=0}^{\infty}s^n\mathbf{P}^n$ 显然在 $\|\cdot\|_{\infty,\infty}$ 有界算子空间中收敛。于是，由(3.1)，$\mathbf{R}(s)_{ij}$ 也存在且唯一。

本节将要证明

$$\lim_{s\nearrow 1}\mathbf{R}(s)_{ij}=\pi_{ij}:=\begin{cases}1/\mathbb{E}[\rho_j\mid X_0=j], & i=j,\\ \mathbb{P}(\rho_j<\infty\mid X_0=i)/\mathbb{E}[\rho_j\mid X_0=j], & i\ne j.\end{cases}\tag{3.2}$$

证明的关键步骤是更新方程：对于 $n\ge 1$，

$$(\mathbf{P}^n)_{ij}=\sum_{m=1}^{n}\mathbb{P}(\rho_j=m\mid X_0=i)(\mathbf{P}^{n-m})_{jj}.$$

这个方程的证明如下：

$$\begin{aligned}(\mathbf{P}^n)_{ij} &= \sum_{m=1}^{n} \mathbb{P}(X_n = j \& \rho_j = m \mid X_0 = i) \\ &= \sum_{m=1}^{n} \mathbb{P}(X_n = j \mid \rho_j = m, X_0 = i)\mathbb{P}(\rho_j = m \mid X_0 = i) \\ &= \sum_{m=1}^{n} \mathbb{P}(X_{n-m} = j \mid X_0 = j)\mathbb{P}(\rho_j = m \mid X_0 = j).\end{aligned}$$

定义

$$f(s)_{ij} = \sum_{m=1}^{\infty} s^m \mathbb{P}(\rho_j = m \mid X_0 = i) = \mathbb{E}[s^{\rho_j} \mid X_0 = i].$$

于是

$$\begin{aligned}\mathbf{R}(s)_{ij} &= (1-s)\delta_{i,j} + (1-s)\sum_{n=1}^{\infty} s^n \sum_{m=1}^{n} \mathbb{P}(\rho_j = m \mid X_0 = i)(\mathbf{P}^{n-m})_{jj} \\ &= (1-s)\delta_{i,j} + (1-s)\sum_{m=1}^{\infty} s^n \mathbb{P}(\rho_j = m \mid X_0 = i) \sum_{n=m}^{\infty} s^{n-m}(\mathbf{P}^{n-m})_{jj} \\ &= (1-s)\delta_{i,j} + f(s)_{ij}\mathbf{R}(s)_{jj}.\end{aligned}$$

当 $i = j$，得到

$$\mathbf{R}(s)_{jj} = (1-s) + f(s)_{jj}\mathbf{R}(s)_{jj} \iff \mathbf{R}(s) = \frac{1-s}{1-f(s)_{jj}},$$

当 $i \neq j$，得到

$$\mathbf{R}(s)_{ij} = f(s)_{ij}\mathbf{R}(s)_{jj}.$$

如果 j 是瞬时的，则 $\mathbb{E}[\rho_j \mid X_0 = j] = \infty$，并且

$$\lim_{s\nearrow 1} \mathbf{R}(s)_{jj} = \frac{\lim_{s\nearrow 1}(1-s)}{\lim_{s\nearrow 1}(1-f(s)_{jj})} = \frac{0}{\mathbb{P}(\rho_j = \infty \mid X_0 = j)} = 0, \tag{3.3}$$

于是我们证明(3.2)，并且所有 $\pi_{ij} = 0$。如果 j 是常返的，利用单调收敛定理，以及当 $s \nearrow 1$ 则 $(1-s^m)/(1-s) \nearrow m$ 的事实，我们得到当 $s \nearrow 1$ 时，

$$\begin{aligned}\lim_{s\nearrow 1} \frac{1-f(s)_{jj}}{1-s} &= \lim_{s\nearrow 1} \sum_{m=1}^{\infty} \frac{1-s^m}{1-s}\mathbb{P}(\rho_j = m \mid X_0 = i) \\ &\overset{\text{这里用单调收敛定理}}{=} \sum_{m=1}^{\infty} m\mathbb{P}(\rho_j = m \mid X_0 = i) = \mathbb{E}[\rho_j \mid X_0 = j],\end{aligned} \tag{3.4}$$

$$\lim_{s\nearrow 1} f(s)_{jj} = \lim_{s\nearrow 1}\sum_{m=1}^{\infty} s^m \mathbb{P}(\rho_j = m \mid X_0 = i)$$
$$= \sum_{m=1}^{\infty} \mathbb{P}(\rho_j = m \mid X_0 = i) = \mathbb{P}(\rho_j < \infty \mid X_0 = i).$$

所以 $\lim_{s\nearrow 1}\mathbf{R}(s)_{jj} = \mathbb{E}[\rho_j \mid X_0 = j]^{-1}$ 以及如果 $i \neq j$ 则 $\lim_{s\nearrow 1}\mathbf{R}(s)_{ij} = \mathbb{P}(\rho_j < \infty \mid X_0 = i)\mathbb{E}[\rho_j \mid X_0 = j]^{-1}$。

3.2.3 平稳概率分布的结构

如果一个（$\mathbb{S}$ 上的）概率分布向量 $\vec{\mu}$ 满足 $\mu = \mu\mathbf{P}$，则称为$\mathbf{P}$-平稳，记为 $\vec{\mu} \in \mathrm{Stat}(\mathbf{P})$。显然，如果 $\vec{\mu} \in \mathrm{Stat}(\mathbf{P})$，则 $\vec{\mu} = \vec{\mu}\mathbf{R}(s)$（$s \in [0,1)$）。于是，根据(3.2)，（令 $s \in [0,1)$）

$$\mu_j = (\vec{\mu}\mathbf{R}(s))_j = \sum_{i\in\mathbb{S}} \mu_i \mathbf{R}(s)_{ij} = \lim_{s\nearrow 1} \mu_i \mathbf{R}(s)_{ij} = \sum_{i\in\mathbb{S}} \mu_i \pi_{ij}.$$

一方面，如果 j 是瞬时的，则 $\pi_{ij} = 0$。另一方面，如果 j 是常返的，则当 $i \leftrightarrow j$ 时，因为由推论3.1.1，$\mathbb{P}(\rho_j < \infty \mid X_0 = i) = 1$，我们根据(3.2)有 $\pi_{ij} = \pi_{jj}$；当 $i \not\leftrightarrow j$ 时，同样理由我们有 $\mathbb{P}(\rho_j < \infty \mid X_0 = i) = 0$ 和 $\pi_{ij} = 0$。在两种情况下，对于 $\mathrm{Stat}(\mathbf{P})$ 中的分布向量 $\vec{\mu}$，我们都有

$$\mu_j = \left(\sum_{i\in\mathbb{S}\text{且}i\leftrightarrow j} \mu_i\right)\pi_{jj}. \tag{3.5}$$

下面，我们要证明，如果对某个状态 j，定义 $C = C_j$ 以及 $\mathbb{R}^{\mathbb{S}}$ 上的行向量 π^C 如下：

$$C = \{i : i \leftrightarrow j\}, \quad (\pi^C)_i = \begin{cases} 0, & i \notin C, \\ \pi_{ii}, & i \in C. \end{cases} \tag{3.6}$$

则只要 $\pi_{jj} > 0$，π^C 就是 $\mathbf{P}$-平稳的。

要证明上述结果，我们先注意到，$\pi_{jj} > 0$ 时，j 常返，于是所有的状态 $i \in C$ 都常返并且根据定理3.1.1，$\mathbb{P}(\rho_i < \infty \mid X_0 = j) = 1$。而对于 $i \notin C$，则同样由定理3.1.1，$j \not\to i$（而不是 $j \to i$ 且 $i \not\to j$），所以对于 $s \in (0,1)$，我们有 $(\mathbf{R})_{ji} = 0$。所以综合上文，对任意 i，

$$(\pi^C)_i = \lim_{s\nearrow 1} \mathbf{R}(s)_{j,i}. \tag{3.7}$$

根据法都（Fatou）引理[1]，

$$\sum_{i\in\mathbb{S}}(\pi^C)_i = \sum_{i\in\mathbb{S}} \liminf_{s\nearrow 1} \mathbf{R}(s)_{ji} \le \liminf_{s\nearrow 1} \sum_{i\in\mathbb{S}} \mathbf{R}(s)_{ji} = \liminf_{s\nearrow 1} 1 = 1.$$

同样，对于任意 i，

$$(\pi^C\mathbf{P})_i = \sum_{k\in C} \pi_{kk}(\mathbf{P})_{ki} \le \liminf_{s\nearrow 1} \sum_{k\in C} (\mathbf{R}(s))_{jk}(\mathbf{P})_{ki} = \liminf_{s\nearrow 1} s(\mathbf{R}(s))_{ji} = (\pi^C)_i.$$

我们只要证明，在上面两个不等式中，实际上成立的是等式。假设第二个不是，则

$$\begin{aligned}\sum_{k\in\mathbb{S}}(\pi^C)_k &= \sum_{k\in\mathbb{S}}(\pi^C)_k\left(\sum_{i\in\mathbb{S}}(\mathbf{P})_{ki}\right)\\ &\overset{\text{富比尼（Fubini）定理}}{=} \sum_{i\in\mathbb{S}}\left(\sum_{k\in\mathbb{S}}(\pi^C)_k(\mathbf{P})_{ki}\right) < \sum_{i}(\pi^C)_i.\end{aligned}$$

导出矛盾。这样，我们有了 $\pi^C = \pi^C\mathbf{P}$，也就有了 $\pi^C = \pi^C\mathbf{R}(s)$（$s\in[0,1)$）。所以

$$\lim_{s\nearrow 1}\sum_{i\in\mathbb{S}}(\pi^C)_i(\mathbf{R}(s))_{ij} = \lim_{s\nearrow 1}(\pi^C)_j = (\pi^C)_j.$$

另一方面，上面的极限也等价于（利用勒贝格（Lebesgue）控制收敛定理，以及 $(\mathbf{R}(s))_{ij}\le 1$）

$$\begin{aligned}\lim_{s\nearrow 1}\sum_{i\in C}(\pi^C)_i(\mathbf{R}(s))_{ij} &= \sum_{i\in C}(\pi^C)_i \underbrace{\lim_{s\nearrow 1}(\mathbf{R}(s))_{ij}}_{=\mathbb{P}(\rho_j<\infty|X_0=i)\pi_{jj}}\\ &= \sum_{i\in C}(\pi^C)_i(\pi^C)_j = \left(\sum_{i\in C}(\pi^C)_i\right)(\pi^C)_j.\end{aligned}$$

这样，我们有 $\sum_{i\in C}(\pi^C)_i = 1$。于是完成了证明。

下面，我们引入两个概念：首先是凸性这个概念。令 A 为某线性空间的一个子集，如果对于任意 $a, a'\in A$ 以及 $\theta\in[0,1]$，都有 $(1-\theta)a+\theta a'\in A$，则我们称 A 为凸集。如果 $b\in A$ 且对任意 $\theta\in(0,1)$ 和 $a,a'\in A$，$b=(1-\theta)a+\theta a'$ 只在 $a=a'=b$ 时成立，则我们称 b 为 A 的一个极点。对一个凸集来说，知道了它所有的极点，也就唯一确定了它。其次是正常返这个概念。如果在某个马氏链里，状态 $j\in\mathbb{S}$ 满足 $\mathbb{E}[\rho_j\mid X_0=j]<\infty$，则我们称 j 为正常返态。如果状态 $i\in$ 常返但是 $\mathbb{E}[\rho_i\mid X_0=i]=\infty$，则我们称 i 为零常返态。

[1]法都引理是关于积分的结果。这里，我们把在 $\mathbb{S}$ 上求和理解为积分。

定理 3.2.1. *1.* Stat($\mathbf{P}$) 是 $\mathbb{R}^{\mathbb{S}}$ 的凸子集。

2. 对任意 $\mu \in \mathrm{Stat}(\mathbf{P})$，等式(3.5)成立。

3. Stat($\mathbf{P}$) 非空的充分必要条件是，至少有一个状态是正常返状态。

4. μ 为 Stat($\mathbf{P}$) 的极点的充分必要条件为，存在某个正常返态 j 以及由(3.6)定义的 C 以及 π^C 并且 $\pi^C = \mu$。

5. 如果 j 是个瞬时态，则对任意 $\mu \in \mathrm{Stat}(\mathbf{P})$，$\mu_j = 0$。如果 j 是个常返态且 $\mu_j = 0$（或 $\mu_j > 0$），则对于所有与 j 互通的状态 i，都有 $\mu_i = 0$（或 $\mu_i > 0$）。

证明. 第1部分，Stat($\mathbf{P}$) 的凸性是显然的。

第2部分前面已经证明。

第3部分，如果 $\mu \in \mathrm{Stat}(\mathbf{P})$ 且 $\mu_j > 0$，则根据(3.5)，$\pi_{jj} > 0$，也就是说 j 是正常返态。相反地，如果 j 是正常返态，则由(3.6)定义的 π^C 是非平凡平稳分布。

下面证明第4部分。假设 μ 是非零 $\mathbf{P}$-平稳概率分布向量，且对任意常返态 j 以及由 j 定义的 $C = C_j$，都有 $\mu \neq \pi^C$。我们首先看到，令 $u_i \neq 0$ 的 i 不可能包含在一个互通类中。不然，所有这样的 i 都在某个 $C = C_j$ 中，于是 $\sum_{i\in\mathbb{S}且i\leftrightarrow j} \mu_i = 1$。这样的话，根据(3.5)，这个 μ 就等于 π^C，跟假设矛盾。

所以，存在不互通的两个状态 j 和 j'，使得 $\mu_j > 0$ 且 $\mu_{j'} > 0$。这时，根据(3.5)，我们可以把 μ 写成

$$\mu = \theta\pi^C + (1-\theta)\nu, \quad C = C_j = \{i : i \leftrightarrow j\}且\theta = \sum_{i\in C} \mu_i \in (0,1).$$

这里当 $i \in C$ 时 $\nu_i = 0$，否则 $\nu_i = (1-\theta)^{-1}\mu_i$。显然，$\nu \in \mathrm{Stat}(\mathbf{P})$，且 $\nu \neq \pi^C$。所以 μ 不是 Stat($\mathbf{P}$) 的极点。如上我们证明 Stat($\mathbf{P}$) 的极点都是 π^C 形式的。

反之，如果某个 π^C 不是极点，而是可以写成 $(1-\theta)\mu + \theta\nu$ 形式，（这里 $\mu, \nu \in \mathrm{Stat}(\mathbf{P})$），则对所有 $i \notin C$，我们有 $\mu_i = 0$。由之前“不等于 π^C 的 μ 一定不能只对属于 C 的 i 有 $\mu_i > 0$”的论证，我们知道 $\mu = \pi^C$。这样的话，ν 也等于 π^C，我们证明 π^C 是个极点。

下面证明第5部分。首先我们说明，当 j 为一个瞬时态或零常返态时，对任意 $\mu \in \mathrm{Stat}(\mathbf{P})$，$\mu_j = 0$。这是因为 $\pi_{jj} = 0$（瞬时态情况见(3.3)，零常返态情况见(3.4)），我们根据(3.5)有 $\mu_j = 0$。

如果 j 是一个正常返态，则由(3.6)至少 $(\pi^C)_j > 0$。对于任意 $\mu \in \mathrm{Stat}(\mathbf{P})$，如果 $i \to j$，也就是存在 $m > 0$ 使得 $(\mathbf{P}^m)_{ij} > 0$，则 $\mu_i > 0$ 可以推出 $\mu_j = (\mu\mathbf{P}^m)_j \ge \mu_i(\mathbf{P}^m)_{ij} > 0$。所以，在 $C = C_j = \{i : i \leftrightarrow j\}$ 上，μ_i 或者都等于 0，或者都大于 0。 □

由定理3.2.1的证明看，我们知道正常返性是关于状态等价类的函数：如果 $i \leftrightarrow j$，则两者都为正常返态，或者两者都不是。

3.2.4 由阿贝尔求和到切萨罗求和

我们刚才得到的最重要的结论是(3.7)。但是，(3.7)的左边是 $(\mathbf{P}^n)_{j,i}$ 关于 n 的阿贝尔极限，没有什么概率意义。我们准备把它替换成 $(\mathbf{P}^n)_{j,i}$ 的切萨罗极限 $\lim_{n\to\infty}(\mathbf{A}_n)_{j,i} = \lim_{n\to\infty} n^{-1}\sum_{m=0}^{n-1}(\mathbf{P}^m)_{j,i} = \pi_{j,i}$。这个极限有概率意义（按时间平均）。

我们需要一个技术引理：如果 $\{a_m\}_0^\infty$ 是一列实数且 $0 \le a_m \le 1$，且 $A_n = n^{-1}\sum_{m=0}^{n-1} a_m$ 是此数列前 n 项的平均，则

$$|A_n - A_{n-m}| \le \frac{m}{n}, \quad 0 \le m < n.$$

（证明留作练习。）

引理 3.2.1. 对所有的 $i, j \in \mathbb{S}$，$\limsup_{n\to\infty}(\mathbf{A}_n)_{ij} \le e\pi_{ij}$。对任意 $j \in \mathbb{S}$ 以及任意 $\{n_1 < n_2 < \cdots\} \subseteq \mathbb{N}$，如果 $\lim_{l\to\infty}(\mathbf{A}_{n_l})_{jj} = \alpha$，则对另外 $i \in \mathbb{S}$

$$\lim_{l\to\infty}(\mathbf{A}_{n_l})_{ij} = \mathbb{P}(\rho_j < \infty \mid X_0 = i)\alpha.$$

证明. 第一部分，可以用如下不等式去估计 $\limsup_{n\to\infty}(\mathbf{A}_n)_{ij}$:

$$(\mathbf{A}_n)_{ij} \le \frac{1}{n}(1-\frac{1}{n})^{-n}\sum_{m=0}^{n-1}(1-\frac{1}{n})^m(\mathbf{P}^m)_{ij} \le (1-\frac{1}{n})^{-n}(\mathbf{R}(1-\frac{1}{n}))_{ij}.$$

然后用已经证得的阿贝尔极限公式，在不等式两边取上极限，就得到了结论。

第二部分，我们利用

$$(\mathbf{A}_n)_{ij} = \sum_{m=1}^{n-1}\mathbb{P}(\rho_j = m \mid X_0 = i)(1-\frac{m}{n})(\mathbf{A}_{n-m})_{jj},$$

得到（应用上面的不等式结果 $(\mathbf{A}_{n-m})_{jj} \le \frac{m}{n} + (\mathbf{A}_n)_{jj}$）

$$\begin{aligned}
&|(\mathbf{A}_n)_{ij} - \mathbb{P}(\rho_j < n \mid X_0 = i)\alpha| \\
&= \left|\sum_{m=1}^{n-1} \mathbb{P}(\rho_j = m \mid X_0 = i)\left[(1-\frac{m}{n})(\mathbf{A}_{n-m})_{jj} - \alpha\right]\right| \\
&\le \sum_{m=1}^{n-1} \mathbb{P}(\rho_j = m \mid X_0 = i)\left|((\mathbf{A}_n)_{jj} - \alpha) + (\frac{m}{n} - \frac{m^2}{n^2} - \frac{m}{n}(\mathbf{A}_n)_{jj})\right| \\
&\le \left(2\sum_{m=1}^{n-1} \frac{m}{n}\mathbb{P}(\rho_j = m \mid X_0 = i)\right) + |(\mathbf{A}_n)_{jj} - \alpha|.
\end{aligned}$$

因为右边的两项都随着 $n \to \infty$ 趋近 0（第一项需要一点证明），我们得到结论。 $\square$

现在我们可以证明

$$\lim_{n\to\infty} (\mathbf{A}_n)_{ij} = \pi_{ij}.$$

如果 $\pi_{jj} = 0$,则 $\pi_{ij} = 0$,用上面引理的第一部分,我们就知道 $\lim_{n\to\infty}(\mathbf{A}_n)_{ij} = 0$，证明完毕。如果 $\pi_{jj} > 0$，根据定理3.2.1，j 是个正常返态，并且，如果 $C = C_j = \{i : i \leftrightarrow j\}$，则 $\pi^C \in \mathrm{Stat}(\mathbf{P})$。于是，根据(3.6)，对于任意 n，$\pi_{jj} = (\pi^C)_j = \sum_{i\in C}(\pi^C)_i(\mathbf{A}_n)_{ij}$。令 $\alpha \in [0,1]$ 为 $\{(\mathbf{A}_n)_{jj} : n \in \mathbb{N}\}$ 的一个聚点，且 $\{n_l\} \subseteq \mathbb{N}$ 为 $\mathbb{N}$ 的子序列使 $\lim_{l\to\infty}(\mathbf{A}_{n_l})_{jj} = \alpha$。于是，根据上面引理的第二部分，如果 $i \in C$，则

$$\lim_{l\to\infty} (\mathbf{A}_{n_l})_{ij} = \mathbb{P}(\rho_j < \infty \mid X_0 = i)\alpha = \alpha.$$

（这里我们利用了推论3.1.1）。这样，我们就有（如果 C 是无穷集，这里需要勒贝格控制收敛定理）

$$\pi_{jj} = \lim_{l\to\infty} \sum_{i\in C} (\pi^C)_i(\mathbf{A}_{n_l})_{ij} = \sum_{i\in C}(\pi^C)_i\alpha = \alpha.$$

于是，π_{jj} 是 $\{(\mathbf{A}_n)_{jj} : n \in \mathbb{N}\}$ 唯一的聚点，也就是 $\lim_{n\to\infty}(\mathbf{A}_n)_{jj} = \pi_{jj}$。最后，应用上个引理的第二部分，我们得到 $\lim_{n\to\infty}(\mathbf{A}_n)_{ij} = \mathbb{P}(\rho_j < \infty \mid X_0 = i)\lim_{n\to\infty}(\mathbf{A}_n)_{jj} = \mathbb{P}(\rho_j < \infty \mid X_0 = i)\pi_{jj} = \pi_{ij}$。（最后一步用了(3.2)。）

3.2.5 平均遍历定理

定理 3.2.2. 令 C 为一个由正常返态组成的联通类。如果马氏链从 C 里面开始，也就是 $\mathbb{P}(X_0 \in C) = 1$，则

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\frac{1}{n}\sum_{m=0}^{n-1} 1_{\{j\}}(X_m) - \pi_{jj}\right)^2\right] = 0.$$

证明. 因为这个马氏链走出 C 的概率为零，所以不失一般性，我们假设 $\mathbb{S} = C$。（如果有非零概率从 $i \in C$ 走到 $j \notin C$，则 $i \to j$ 且 $j \not\to i$。由定理3.1.1，这与 i 的常返性矛盾。）于是 π^C 是唯一的 **P**-平稳的概率分布向量。这里，我们记 $\pi = \pi^C$。下面，我们认为 $\mathbb{S} = \{0, 1, 2, \ldots\}$ 且 $j = 0$。

首先，我们把各个初始值拆开考虑：如果 $\mu_i = \mathbb{P}(X_0 = i)$，则

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{m=0}^{n-1} 1_{\{0\}}(X_m) - \pi_{00}\right)^2\right] = \sum_{i\in\mathbb{S}} \mu_i \mathbb{E}\left[\left(\frac{1}{n}\sum_{m=0}^{n-1} 1_{\{0\}}(X_m) - \pi_{00}\right)^2 \middle| X_0 = i\right],$$

所以我们只需要证明对任意 $i \in \mathbb{S}$,

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\frac{1}{n}\sum_{m=0}^{n-1} 1_{\{0\}}(X_m) - \pi_{00}\right)^2 \middle| X_0 = i\right] = 0.$$

反过来，如果初始分布是唯一的平稳概率分布，也就是 $\mathbb{P}(X_0 = i) = \pi_i$，则只要证明

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\frac{1}{n}\sum_{m=0}^{n-1} 1_{\{0\}}(X_m) - \pi_{00}\right)^2 \middle| X_0 = Y\right] = 0,$$

这里 Y 是一个分布为 $\mathbb{P}(Y = i) = \pi_i$ 的随机变量。

令 $\vec{f}$ 为一个列向量，其分量为 $f_i = \delta_{0,i} - \pi_{00}$。这样，

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{n}\sum_{m=0}^{n-1} 1_{\{0\}}(X_m) - \pi_{00}\right)^2\right] \\
&= \frac{1}{n^2}\sum_{m=0}^{n-1}\mathbb{E}(1_{\{0\}}(X_m) - \pi_{00})^2 \\
&\quad + \frac{2}{n^2}\sum_{m=0}^{n-2}\mathbb{E}\left[(1_{\{0\}}(X_m) - \pi_{00})\left(\sum_{l=m+1}^{n-1}(1_{\{0\}}(X_l) - \pi_{jj})\right)\right] \\
&\le \frac{2}{n^2}\mathbb{E}\left[\sum_{m=0}^{n-1}(1_{\{0\}}(X_m) - \pi_{00})\left(\sum_{l=m}^{n-1}(1_{\{0\}}(X_l) - \pi_{jj})\right)\right] \\
&= \frac{2}{n^2}\sum_{m=0}^{n-1}\mathbb{E}\left[f_{X_m}\sum_{l=m}^{n-1} f_{X_l}\right] \\
&= \frac{2}{n^2}\sum_{m=0}^{n-1}(n-m)\mathbb{E}[f_{X_m}(\mathbf{A}_{n-m}\vec{f})_{X_m}].
\end{aligned}$$

既然我们假定 X_0 服从平稳分布 π，则所有 X_m 都服从同样的平稳分布 π，从而

$$\mathbb{E}[f_{x_m}(\mathbf{A}_{n-m}\vec{f})_{X_m}] = \sum_{i\in\mathbb{S}}\pi_i f_i(\mathbf{A}_{n-m}\vec{f})_i.$$

所以，我们可以把之前的不等式写为

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{m=0}^{n-1} 1_{\{0\}}(X_m) - \pi_{00}\right)^2\right] \le \frac{2}{n^2}\sum_{k=1}^{n} k\sum_{i\in\mathbb{S}}\pi_i f_i(\mathbf{A}_k\vec{f})_i.$$

最后，因为

$$\sum_{i\in\mathbb{S}}\pi_i f_i(\mathbf{A}_k\vec{f})_i = \sum_{i\in\mathbb{S}}\pi_i f_i(\underbrace{(\mathbf{A}_k)_{i0} - \pi_{00}}_{\text{当 } k\to\infty \text{ 时趋近 } 0}),$$

当 $k \to \infty$ 时趋近于 0，结论证毕。 □

3.2.6 非周期情况下的更强结论

在本节，我们证明以下结论：

$$\lim_{n\to\infty}(\mathbf{P}^n)_{ij} = \pi_{ij} \quad \text{如果 } j \text{ 是瞬时态或者非周期态.} \tag{3.8}$$

如果 j 是瞬时态，结论比较易得：首先，由(3.2)，我们得到对任意 i，$\pi_{ij}=0$。然后，对 $n>0$，应用

$$(\mathbf{P}^n)_{ij}=\sum_{m=1}^{n}\mathbb{P}(\rho_j^{(m)}=n\mid X_0=i)=\sum_{m=1}^{\infty}\mathbb{P}(\rho_j^{(m)}=n\mid X_0=i),\tag{3.9}$$

我们有

$$\sum_{n=0}^{\infty}(\mathbf{P}^n)_{ij}=\delta_{ij}+\sum_{m=1}^{\infty}\mathbb{P}(\rho_j^{(m)}<\infty\mid X_0=i)<\infty,$$

于是 $\lim_{n\to\infty}(\mathbf{P}^n)_{ij}=0$，等式得证。下面，我们在非周期情况证明等式。

首先，我们注意到，如果 j 是个非周期态，则存在一个充分大的 $N\in\mathbb{N}$，使得对于所有 $n\geq N$，

$$\max_{1\leq m\leq n}\mathbb{P}(\rho_j^{(m)}=n\mid X_0=j)>0.\tag{3.10}$$

这个结论同样可以用(3.9)证明。

下面我们证明一个比较难的技术引理：

引理 3.2.2. 令 j 为一个非周期的常返态。记 $\alpha_j^-=\liminf_{n\to\infty}(\mathbf{P}^n)_{jj}$，$\alpha_j^+=\limsup_{n\to\infty}(\mathbf{P}^n)_{jj}$。这样的话，就存在 $\mathbb{N}$ 的子序列 $\{n_l^-:l\geq 1\}$，$\{n_l^+:l\geq 1\}$，使得对于任意 $r\in\mathbb{Z}_+$，

$$\alpha_j^{\pm}=\lim_{l\to\infty}(\mathbf{P}^{n_l^{\pm}-r})_{jj}.$$

证明. 我们只证明 $+$ 情况，因为 $-$ 情况类似。

首先，我们取 $\{n_l\}$ 为一个任意的，使 $\lim_{l\to\infty}(\mathbf{P}^{n_l})_{jj}=\alpha_j^+$ 成立的子序列。在此情况下，我们证明引理对充分大的 r 成立。根据(3.10)，如果 r 大于等于某个 N，则有依赖于 r 的 $m\in\{1,\ldots,r\}$ 使 $\mathbb{P}(\rho_j^{(m)}=r\mid X_0=j)=\delta>0$。然后，我们取 $M\in\mathbb{Z}_+$，且对于任意大于 $M+r$ 的 n_l，分解

$$\begin{aligned}(\mathbf{P}^{n_l})_{jj}&=\mathbb{P}(X_{n_l}=j\&\rho_j^{(m)}=r\mid X_0=j)+\mathbb{P}(X_{n_l}=j\&\rho_j^{(m)}\neq r\mid X_0=j)\\&=\delta(\mathbf{P}^{n_l-r})_{jj}+\mathbb{P}(X_{n_l}=j\&\rho_j^{(m)}\geq\rho_j^{(m)}\neq r\mid X_0=j)\\&\quad+\mathbb{P}(X_{n_l}=j\&\rho_j^{(m)}>n_l-M\mid X_0=j)\end{aligned}\tag{3.11}$$

对等式右边的三项，我们有

$$\begin{aligned}
&\mathbb{P}(X_{n_l} = j \& \rho_j^{(m)} \geq \rho_j^{(m)} \neq r \mid X_0 = j) \\
=& \sum_{\substack{k \in \{m, m+1, \ldots, n_l - M\} \\ k \neq r}} \mathbb{P}(\rho_j^{(m)} = k \mid X_0 = j)(\mathbf{P}^{n_l - k})_{jj} \\
\leq& \left[\sum_{\substack{k \in \{m, m+1, \ldots, n_l - M\} \\ k \neq r}} \mathbb{P}(\rho_j^{(m)} = k \mid X_0 = j) \right] \sup_{n \geq M} (\mathbf{P}^n)_{jj} \\
\leq& (1 - \delta) \sup_{n \geq M} (\mathbf{P}^n)_{jj},
\end{aligned}$$

$$\mathbb{P}(X_{n_l} = j \& \rho_j^{(m)} > n_l - M \mid X_0 = j) \leq \mathbb{P}(\rho_j^{(m)} > n_l - M \mid X_0 = j).$$

因为当 $l \to \infty$ 时，$\mathbb{P}(\rho_j^{(m)} > n_l - M \mid X_0 = j) \to \mathbb{P}(\rho_j^{(m)} = \infty \mid X_0 = j) = 0$，我们有，在等式(3.11)取 $l \to \infty$ 的下极限，则

$$\alpha_j^+ \leq \liminf_{l \to \infty} \delta (\mathbf{P}^{n_l - r})_{jj} + (1 - \delta) \sup_{n \geq M} (\mathbf{P}^n)_{jj}.$$

又因为当 $M \to \infty$ 时，$\sup_{n \geq M} (\mathbf{P}^n)_{jj} \to \alpha_j^+$，我们可以得到 $\liminf_{l \to \infty} (\mathbf{P}^{n_l - r})_{jj} \geq \alpha_j^+$。最后，因为 $\limsup_{l \to \infty} (\mathbf{P}^{n_l - r})_{jj} \leq \alpha_j^+$，我们得到

$$\lim_{l \to \infty} (\mathbf{P}^{n_l - r})_{jj} = \alpha_j^+, \quad r \geq N.$$

如果我们吧一开始的收敛于 α_j^+ 的子序列取为 $n_l^+ = n_l - N$（忽略前几个未负数的项），则定理所需要求完全满足。 □

引理 3.2.3. *如果 j 是非周期的常返态，则 $\alpha_j^+ := \limsup_{n \to \infty} (\mathbf{P}^n)_{jj} \leq \pi_{jj}$。并且对于任何与 j 互通的状态 i，如果 $\{n_l^\pm\}$ 是如引理3.2.2定义的子序列，则 $\lim_{l \to \infty} (\mathbf{P}^{n_l^\pm})_{ij} = \alpha_+^\pm$。*

证明. 首先证明第二部分。因为（对 $i \neq j$）

$$(\mathbf{P}^{n_l^\pm})_{ij} = \sum_{r=1}^{n_l^\pm} \mathbb{P}(\rho_j = r \mid X_0 = i)(\mathbf{P}^{n_l^\pm - r})_{jj},$$

令 $l \to \infty$，应用勒贝格控制收敛定理（控制函数 $1 \geq (\mathbf{P}^{n_l^\pm - r})_{jj}$），得到

$$\lim_{l \to \infty} (\mathbf{P}^{n_l^\pm})_{ij} = \sum_{r=1}^{n_l^\pm} \mathbb{P}(\rho_j = r \mid X_0 = i)\alpha_j^\pm = \alpha_j^\pm.$$

然后证明第二部分。因为 $\pi_{jj} = 1/\mathbb{E}[\rho_j \mid X_0 = j]$，我们只要证明 $\alpha_j^+ \mathbb{E}[\rho_j \mid X_0 = j] \le 1$，也即

$$\alpha_j^+ \sum_{r=1}^{\infty} \mathbb{P}(\rho_j \ge r \mid X_0 = j) \le 1,$$

或者，证明上面不等式把 ∞ 换成任意的 N，都成立。又因为

$$\alpha_j^+ \sum_{r=1}^{N} \mathbb{P}(\rho_j \ge r \mid X_0 = j) = \lim_{l\to\infty} \sum_{r=1}^{N} \mathbb{P}(\rho_j \ge r \mid X_0 = j)(\mathbf{P}^{n_l^+ - r})_{jj},$$

我们只需要有对所有的 $n \ge N \ge 1$，

$$\sum_{r=1}^{N} \mathbb{P}(\rho_j \ge r \mid X_0 = j)(\mathbf{P}^{n-r})_{jj} \le 1, \tag{3.12}$$

就完成了证明。为了证明此不等式，我们考虑比 $\sum_{r=1}^{N} \mathbb{P}(\rho_j \ge r \mid X_0 = j)(\mathbf{P}^{n-r})_{jj}$ 更大一些的 $\sum_{r=1}^{n} \mathbb{P}(\rho_j \ge r \mid X_0 = j)(\mathbf{P}^{n-r})_{jj}$，并且有

$$\begin{aligned}
&\sum_{r=1}^{n} \mathbb{P}(\rho_j \ge r \mid X_0 = j)(\mathbf{P}^{n-r})_{jj} \\
&= \sum_{r=1}^{n} \left(\sum_{k=0}^{r} (\mathbf{P}^{n-k})_{jj} \right) \mathbb{P}(\rho_j = r \mid X_0 = 0) + \mathbb{P}(\rho_j > n \mid X_0) \\
&= \mathbb{P}(\text{这个马氏链最后去了 } \{n+1, n+2, \dots\}) = 1.
\end{aligned}$$

所以 $\sum_{r=1}^{N} \mathbb{P}(\rho_j \ge r \mid X_0 = j)(\mathbf{P}^{n-r})_{jj} \le 1$。 □

最后我们证明(3.8)的非周期常返态情况。首先，如果 $\pi_{jj} = 0$（也就是说 j 是零常返态），根据引理3.2.3，我们有 $\lim_{n\to\infty}(\mathbf{P}^n)_{jj} = 0$，且对 $i \ne j$，利用同样的引理以及 $\alpha^{\pm} = 0$ 和 $\{n_l^+\} = \{n_-^-\} = \mathbb{N}$，我们得到

$$\lim_{n\to\infty} (\mathbf{P}^n)_{ij} = 0 = \pi_{ij}.$$

另一方面，如果 j 是正常返态，令 $C = \{i : i \leftrightarrow j\}$，并定义 π^C 如(3.6)。我们有 $\pi^C \in \mathrm{Stat}(\mathbf{P})$。因为 $\sum_{i\in C} (\pi^C)_i (\mathbf{P}^{n_l^{\pm}})_{ij} = \pi_i^C = \pi_{jj}$，我们取 $l \to \infty$ 的极限，有

$$\pi_{ij} = \lim_{l\to\infty} \sum_{i\in C} (\pi^C)_i (\mathbf{P}^{n_l^{\pm}})_{ij} = \alpha_j^{\pm} \sum_{i\in C} (\pi^C)_i = \alpha_j^{\pm}.$$

于是 $\alpha_j^+ = \alpha_j^- = \alpha_j$，所以 $\lim_{n\to\infty}(\mathbf{P}^n)_{jj} = \pi_{jj}$。最后，对于 $i \neq j$，

$$\begin{aligned}\lim_{n\to\infty}(\mathbf{P}^n)_{ij} &= \lim_{n\to\infty}\sum_{r=1}^{n}\mathbb{P}(\rho_j = r \mid X_0 = i)\underbrace{(\mathbf{P}^{n-r})_{jj}}_{\text{极限为 } \pi_{jj}} \\ &= \sum_{r=1}^{\infty}\mathbb{P}(\rho_j = r \mid X_0 = i)\pi_{jj} = \mathbb{P}(\rho_j < \infty \mid X_0 = i)\pi_{jj} = \pi_{ij}.\end{aligned}$$