FDU 随机过程导论 4. Markov 链

本文根据王老师课堂笔记整理而成，并参考了以下教材:

- 随机过程 (苏中根) 第 4 章
- Introduction to Probability Models: Applied Stochastic Processes (S. Ross)
- 应用随机过程概率论模型导论 (S. Ross) 龚光鲁译 第 4 章

欢迎批评指正!

4.1 Introduction

假设随机过程 $\{X_n : n \in \mathbb{N}\}$ 只有有限个或可数个可能取值.
除非特殊指定，我们默认其状态空间为 $\mathcal{E} = \mathbb{N} = \{0, 1, 2, \dots\}$

如果 $X_n = i$，则我们称该过程在时刻 n 为状态 i.
假设只要过程在状态 i 就有一个**固定的转移概率** p_{ij} 使它在下一个时刻为状态 j.
也就是说， **将来的状态** X_{n+1} **只依赖于现在的状态** X_n，而独立于过去的状态 $X_0, X_1, \dots, X_{n-1}$，
即对于一切状态 $i_0, i_1, \dots, i_{n-1}, i, j$ 和 $n \geq 0$ 都有:

$$\mathrm{P}\{X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0\} = \mathrm{P}\{X_{n+1} = j | X_n = i\} \overset{\Delta}{=} p_{ij}$$

那么我们称随机过程 $\{X_n : n \in \mathbb{N}\}$ 为 **Markov 链**.

基于概率的非负性和归一性，我们有 $\begin{cases} p_{ij} \geq 0 \ \ (\forall\, i, j \in \mathbb{N}) \\ \sum\limits_{j=0}^{\infty} p_{ij} = 1 \ \ (\forall\, i \in \mathbb{N}) \end{cases}$

我们定义**一步转移概率矩阵**为 $P = [p_{ij}]$

下面是一些具体的例子:

- **(通信系统, S. Ross 例 4.2):**
  考虑一个传送数字 0 和 1 的通信系统，
  记 X_n 为第 n 个阶段内传送的信号，
  设每个阶段信号保持不变的概率为 p，发生改变的概率为 $1-p$.
  则 $\{X_n : n \in \mathbb{N}\}$ 是一个拥有两个状态的 Markov 链，
  具有一步状态转移矩阵 $\mathrm{P} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$
- **(天气预报: 将一个过程转变为 Markov 链, S. Ross 例 4.4):**
  假设明天是否下雨依赖于昨天和今天是否下雨.
  如果昨天、今天都不下雨，则明天下雨的概率为 0.2;
  如果昨天不下雨、今天下雨，则明天下雨的概率为 0.5;
  如果昨天下雨、今天不下雨，则明天下雨的概率为 0.4;
  如果昨天、今天都下雨，则明天下雨的概率为 0.7;
  将这个问题建模为一个 Markov 链需要一点小技巧 (**两天为一个单位，可以交叠**):
  我们假定过程有 4 个状态:
  - 状态 0: 昨天、今天都不下雨 (分别以 0.2 和 0.8 的概率转移到状态 1 和状态 0)
  - 状态 1: 昨天不下雨、今天下雨 (分别以 0.5 和 0.5 的概率转移到状态 3 和状态 2)
  - 状态 2: 昨天下雨、今天不下雨 (分别以 0.4 和 0.6 的概率转移到状态 1 和状态 0)
  - 状态 3: 昨天、今天都下雨 (分别以 0.7 和 0.3 的概率转移到状态 3 和状态 2)

这样我们就定义了一个拥有 4 个状态的 Markov 链,
具有一步转移概率矩阵 $P = \begin{bmatrix} 0.8 & 0.2 & & \\ & & 0.5 & 0.5 \\ 0.6 & 0.4 & & \\ & & 0.3 & 0.7 \end{bmatrix}$

- **(简单随机游动, simple random walk, S. Ross 例 4.5):**
  一个人在直线上行走，他在每一个时间点以概率 p 向右走一步，或以概率 $1-p$ 向左走一步.
  其中 $p \in (0,1)$
  这个随机过程的状态空间为 $i = 0, \pm 1, \pm 2, \cdots$
  任意状态 i 的转移概率为 $\begin{cases} p_{i,i+1} = p \\ p_{i,i-1} = 1-p \end{cases}$
- **(赌博模型, S. Ross 例 4.6):**
  一个赌徒，在每局中赢 1 元的概率为 p，输 1 元的概率为 $1-p$.
  假设他在破产或财富到达 N 元时离开.
  此时赌徒的财富是一个具有吸收壁 (状态 0 和 N) 的有限状态的简单随机游动,
  具有转移概率 $\begin{cases} p_{i,i+1} = p = 1 - p_{i,i-1} \quad (i = 1, 2, \ldots, N-1) \\ p_{00} = p_{NN} = 1 \end{cases}$
  状态 0 和 N 称为**吸收态**，因为一旦进入此状态，它们就不再离开.
- **(保险, S. Ross 例 4.7):**
  参保人被赋予一个正整数值的状态，而年保险金是该状态的函数.
  在每一年中，若参保人无理赔要求，则其状态降低，保险金降低;
  否则，状态增加，保险金增加;
  以 $s_i(k)$ 记一个在上一年处于状态 i 且该年有 k 次理赔的参保人下一年的状态.
  假设该参保人年理赔次数是参数为 $\lambda > 0$ 的 Poisson 随机变量,
  那么此参保人下一年的状态构成一个 Markov 链,
  具有转移概率 $p_{ij} = \sum\limits_{k:s_i(k)=j} \mathrm{P}\{\text{Poisson}(\lambda) = k\} = \sum\limits_{k:s_i(k)=j} e^{-\lambda} \frac{\lambda^k}{k!} \quad (\forall\, j \geq 0)$

4.2 Chapman–Kolmogorov 方程

我们已经定义了一步转移概率 $p_{ij} = \mathrm{P}\{X_{k+1} = j | X_k = i\} \quad (\forall\, k \geq 0, i, j \geq 0)$
现在我们定义 $n \geq 0$ 步转移概率 $p_{ij}^n = \mathrm{P}\{X_{k+n} = j | X_k = i\} \quad (\forall\, k \geq 0, i, j \geq 0)$
Chapman–Kolmogorov 方程提供了 n 步转移概率的计算方法:
对于任意 $\begin{cases} n, m \geq 0 \\ i, j \geq 0 \end{cases}$ 都有 $p_{ij}^{n+m} = \sum\limits_{k=0}^{\infty} p_{ik}^n p_{kj}^m$

- **证明:**
  $$\begin{aligned} p_{ij}^{n+m} &= \mathrm{P}\{X_{n+m} = j | X_0 = i\} \\ &= \sum_{k=0}^{\infty} \mathrm{P}\{X_{n+m} = j, X_n = k | X_0 = i\} \\ &= \sum_{k=0}^{\infty} \mathrm{P}\{X_{n+m} = j | X_n = k, X_0 = i\} \mathrm{P}\{X_n = k | X_0 = i\} \\ &= \sum_{k=0}^{\infty} p_{kj}^m p_{ik}^n \quad (\text{单纯地调换顺序}) \\ &= \sum_{k=0}^{\infty} p_{ik}^n p_{kj}^m \end{aligned}$$

记 n 步转移概率矩阵 $P^{(n)}=[p_{ij}^n]$,
那么 **Chapman–Kolmogorov 方程**表明 $P^{(n+m)}=P^n\cdot P^{(m)}$ (矩阵乘法)
特别地, $P^{(n)}=P^{(n-1)}\cdot P=P^{(n-2)}\cdot P\cdot P=\cdots=P^n$
即 n 步转移概率矩阵 $P^{(n)}$ 可以由 P 自乘 n 次得到.

- **延续 (天气预报: 将一个过程转变为 Markov 链, S. Ross 例 4.4):**
  在 4.1 节的 S. Ross 例 4.4 中我们构建了如下的 Marcov 链:
  - 状态 0: 昨天、今天都不下雨 (分别以 0.2 和 0.8 的概率转移到状态 1 和状态 0)
  - 状态 1: 昨天不下雨、今天下雨 (分别以 0.5 和 0.5 的概率转移到状态 3 和状态 2)
  - 状态 2: 昨天下雨、今天不下雨 (分别以 0.4 和 0.6 的概率转移到状态 1 和状态 0)
  - 状态 3: 昨天、今天都下雨 (分别以 0.7 和 0.3 的概率转移到状态 3 和状态 2)

  这样我们就定义了一个拥有 4 个状态的 Markov 链,
  具有一步转移概率矩阵 $P=\begin{bmatrix}0.8&0.2&&\\&&0.5&0.5\\0.6&0.4&&\\&&0.3&0.7\end{bmatrix}$
  现在已知周一、周二都下雨, 问周四下雨的概率是多少?

  Solution:
  我们计算两步概率转移矩阵:
  $$P^{(2)}=P^2=\begin{bmatrix}0.8&0.2&&\\&&0.5&0.5\\0.6&0.4&&\\&&0.3&0.7\end{bmatrix}\begin{bmatrix}0.8&0.2&&\\&&0.5&0.5\\0.6&0.4&&\\&&0.3&0.7\end{bmatrix}=\begin{bmatrix}0.64&0.16&0.1&0.1\\0.3&0.2&0.15&0.35\\0.48&0.12&0.2&0.2\\0.18&0.12&0.21&0.49\end{bmatrix}$$

  已知周一、周二都下雨 (状态 3),
  则周四下雨对应于状态 1 (周三不下雨、周四下雨) 或状态 3 (周三下雨、周四下雨),
  因此概率为 $P_{31}^{(2)}+P_{33}^{(2)}=0.12+0.49=0.61$

假设初始状态 X_0 的概率分布为 $p_i^{(0)}\equiv\mathrm{P}\{X_0=i\}\ (\forall\, i\geq 0)$, 满足 $\sum_{i=0}^{\infty}p_i^{(0)}=1$
一切无条件概率都可以利用对初始状态 X_0 取条件来计算
即对于任意给定的 $j\in\mathbb{N}$, 我们都有:
$$\begin{aligned}\mathrm{P}\{X_n=j\}&=\sum_{i=0}^{\infty}\mathrm{P}\{X_n=j|X_0=i\}\mathrm{P}\{X_0=i\}\\&=\sum_{i=0}^{\infty}P_{ij}^{(n)}p_i^{(0)}\end{aligned}$$
因此状态 X_n 的概率分布仅由初始概率 $\{p_i^{(0)}\}_{i=0}^{\infty}$ 和 n 步转移概率矩阵 $P^{(n)}$ 确定.

考虑 $(X_0,X_1,\dots,X_m)$ 的联合分布:
$$\begin{aligned}\mathrm{P}\{X_0=i_0,X_1=i_1,\dots,X_m=i_m\}&=\mathrm{P}\{X_0=i_0\}\cdot\mathrm{P}\{X_1=i_1,\dots,X_m=i_m|X_0=i_0\}\\&=p_{i_0}^{(0)}P_{i_0,i_1}\dots P_{i_{m-1},i_m}\end{aligned}$$

考虑 (X_1,X_3,X_6) 的分布:
$$\begin{aligned}\mathrm{P}\{X_1=i_1,X_3=i_3,X_6=i_6\}&=\mathrm{P}\{X_1=i_1\}\cdot\mathrm{P}\{X_3=i_3,X_6=i_6|X_1=i_1\}\\&=(\sum_{i=0}^{\infty}\mathrm{P}\{X_1=i_1|X_0=i\}\mathrm{P}\{X_0=i\})\cdot P_{i_1,i_3}^{(2)}P_{i_3,i_6}^{(3)}\\&=(\sum_{i=0}^{\infty}P_{i,i_1}p_i^{(0)})\cdot P_{i_1,i_3}^{(2)}P_{i_3,i_6}^{(3)}\end{aligned}$$

4.2.1 曾经命中概率

考虑一个具有转移概率 P_{ij} 的 Markov 链 $\{X_n : n \in \mathbb{N}\}$.
以 $S \subset \mathbb{N} = \{0, 1, 2, \dots\}$ 记某个状态集合,
我们想求此 Markov 链在时刻 m 前曾经进入过 S 中任意状态的概率.
也就是说，给定初始状态 $X_0 = i \notin S$,
我们想求 $\mathrm{P}\{X_k \in S \text{ for some } k = 1, \dots, m | X_0 = i \notin S\}$ 的值.

为确定上述概率，我们定义一个 Markov 链 $\{W_n : n \in \mathbb{N}\}$
其状态空间包含所有不属于 S 的状态,
外加一个附加状态 s_0 (一旦 Markov 链 $\{W_n : n \in \mathbb{N}\}$ 进入状态 s_0 就永远保持在其中)

定义 Markov 链 $\{X_n : n \in \mathbb{N}\}$ 首次进入 S 中的某个状态的时间为:
$N = \min\{n : X_n \in S\}$
(若对于一切 n 都有 $X_n \notin S$，则记 $N = \infty$)

现在定义 $W_n = \begin{cases} X_n & \text{if } n < N \\ s_0 & \text{if } n \geq N \end{cases}$
其转移概率为 $\begin{cases} Q_{i,j} = P_{i,j} & \text{for all } i, j \notin S \\ Q_{i,s_0} = \sum\limits_{j \in S} P_{i,j} & \text{for all } i \notin S \\ Q_{s_0,s_0} = 1 \end{cases}$
因此 $\{X_n : n \in \mathbb{N}\}$ 在时刻 m 前进入 S 中的状态,
当且仅当 $\{W_n : n \in \mathbb{N}\}$ 在时刻 m 的状态是 s_0
由此我们可知:

$$\begin{aligned} \mathrm{P}\{X_k \in S \text{ for some } k = 1, \dots, m | X_0 = i\} &= \mathrm{P}\{W_m = s_0 | X_0 = i\} \\ &= \mathrm{P}\{W_m = s_0 | W_0 = i\} \\ &= Q_{i,A}^{(m)} \end{aligned}$$

也就是说，所求概率等于新 Markov 链 $\{W_n : n \in \mathbb{N}\}$ 的一个 m 步转移概率.

(S. Ross 例 4.12)
在一系列独立抛掷一个公平硬币的试验中，以 N 记首次出现连续 3 次正面所需的抛掷次数.

- (1) 求 $\mathrm{P}\{N \leq 8\}$
  Solution:
  我们定义具有状态 $0, 1, 2, 3$ 的 Markov 链 $\{X_n : n \in \mathbb{N}\}$
  - 状态 $i < 3$ 表示目前处于相继正面的一个 i 连贯
  - 状态 $i = 3$ 代表一个相继正面的 3 连贯已经出现

  $\{X_n : n \in \mathbb{N}\}$ 的转移概率矩阵为 $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & & \\ \frac{1}{2} & & \frac{1}{2} & \\ \frac{1}{2} & & & \frac{1}{2} \\ & & & 1 \end{bmatrix}$

  $P \to P^2 \to P^4 \to P^8$ 得到 8 步转移概率矩阵 $P^{(8)} = \dfrac{1}{256} \begin{bmatrix} 81 & 44 & 24 & 107 \\ 68 & 37 & 20 & 131 \\ 44 & 24 & 13 & 175 \\ & & & 256 \end{bmatrix}$

  因此 $\mathrm{P}\{N \leq 8\} = \mathrm{P}\{X_8 = 3\} = P_{0,3}^{(8)} = \frac{107}{256} \approx 0.4180$

- (2) 求 $\mathrm{P}\{N=8\}$

  Solution:

  $$\begin{aligned}\mathrm{P}\{N=8\} &= \mathrm{P}\{N\le 8\}-\mathrm{P}\{N\le 7\}\\ &= \mathrm{P}\{X_8=3\}-\mathrm{P}\{X_7=3\}\\ &= P_{0,3}^{(8)}-P_{0,3}^{(7)}\end{aligned}$$

  我们计算得 7 步转移概率矩阵 $P^{(7)}=\dfrac{1}{128}\begin{bmatrix}44 & 24 & 13 & 47\\ 37 & 20 & 11 & 60\\ 24 & 13 & 7 & 84\\ & & & 128\end{bmatrix}$

  于是有:

  $$\begin{aligned}\mathrm{P}\{N=8\} &= P_{0,3}^{(8)}-P_{0,3}^{(7)}\\ &= \frac{107}{256}-\frac{47}{128}\\ &= \frac{13}{256}\\ &\approx 0.0508\end{aligned}$$

  Another Solution:

  我们定义具有状态 $0,1,2,3,4$ 的 Markov 链 $\{Y_n:n\in\mathbb{N}\}$

  - 状态 $i<3$ 表示目前处于相继正面的一个 i 连贯
  - 状态 $i=3$ 代表一个相继正面的 3 连贯刚刚出现
  - 状态 $i=4$ 代表在过去出现了相继正面的 3 连贯

  $\{Y_n:n\in\mathbb{N}\}$ 的转移概率矩阵为 $Q=\begin{bmatrix}\frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & & \frac{1}{2} & & \\ \frac{1}{2} & & & \frac{1}{2} & \\ & & & & 1\\ & & & & 1\end{bmatrix}$

  $Q\to Q^2\to Q^4\to Q^8$ 得到 8 步转移概率矩阵 $Q^{(8)}=\dfrac{1}{256}\begin{bmatrix}81 & 44 & 24 & 13 & 94\\ 68 & 37 & 20 & 11 & 120\\ 44 & 24 & 13 & 7 & 168\\ & & & & 256\\ & & & & 256\end{bmatrix}$

  因此 $\mathrm{P}\{N=8\}=\mathrm{P}\{Y_8=3\}=Q_{0,3}^{(8)}=\frac{13}{256}\approx 0.0508$

4.2.2 未命中概率

考虑一个具有转移概率 P_{ij} 的 Markov 链 $\{X_n:n\in\mathbb{N}\}$.
以 $S\subset\mathbb{N}=\{0,1,2,\dots\}$ 记某个状态集合,
我们想求此 Markov 链在时刻 m 进入状态 $j\notin S$ 而且从时刻 1 开始未曾进入 S 中任意状态的概率.
也就是说，给定初始状态 $X_0=i$ 和目标状态 $X_m=j\notin S$,
我们想求 $\mathrm{P}\{X_m=j,X_k\notin S \text{ for all } k=1,\dots,m-1|X_0=i\}$ 的值.

参考 4.2.1 节的内容，我们定义新的 Markov 链 $\{Y_n : n \in \mathbb{N}\}$

$$W_n = \begin{cases} X_n & \text{if } n < N \\ s_0 & \text{if } n \geq N \end{cases}$$

其中 $N = \min\{n : X_n \in S\}$ 代表 $\{X_n : n \in \mathbb{N}\}$ 首次进入 S 中的某个状态的时间.

$\{Y_n : n \in \mathbb{N}\}$ 的转移概率为 $\begin{cases} Q_{i,j} = P_{i,j} & \text{for all } i, j \notin S \\ Q_{i,s_0} = \sum\limits_{j \in S} P_{i,j} & \text{for all } i \notin S \\ Q_{s_0,s_0} = 1 \end{cases}$

对于 $i \notin S, j \notin S$ 的情况:

$$\begin{aligned} \mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1, \ldots, m-1 | X_0 = i\} &= \mathrm{P}\{W_m = j | X_0 = i\} \\ &= \mathrm{P}\{W_m = j | W_0 = i\} \\ &= Q_{ij}^{(m)} \end{aligned}$$

对于 $i \notin S, j \notin S$ 的情况，我们也可以计算 X_m 的条件概率:

$$\begin{aligned} &\mathrm{P}\{X_m = j | X_0 = i, X_k \notin S \text{ for all } k = 1, \ldots, m\} \\ &= \frac{\mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1, \ldots, m | X_0 = i\}}{\mathrm{P}\{X_k \notin S \text{ for all } k = 1, \ldots, m | X_0 = i\}} \\ &= \frac{\mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1, \ldots, m | X_0 = i\}}{\sum\limits_{r \notin S} \mathrm{P}\{X_m = r, X_k \notin S \text{ for all } k = 1, \ldots, m | X_0 = i\}} \\ &= \frac{Q_{i,j}^{(m)}}{\sum\limits_{r \notin S} Q_{j,r}^{(m)}} \end{aligned}$$

对于 $i \in S, j \notin S$ 的情况，我们可以确定概率 (对时刻 1 的状态取条件):

$$\begin{aligned} &\mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1, \ldots, m-1 | X_0 = i\} \\ &= \sum_{r \notin S} \mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1, \ldots, m-1 | X_1 = r, X_0 = i\} \cdot \mathrm{P}\{X_1 = r | X_0 = i\} \\ &= \sum_{r \notin S} \mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 2, \ldots, m-1 | X_1 = r\} \cdot P_{i,r} \\ &= \sum_{r \notin S} \mathrm{P}\{X_{m-1} = j, X_k \notin S \text{ for all } k = 1, \ldots, m-2 | X_0 = r\} \cdot P_{i,r} \\ &= \sum_{r \notin S} Q_{r,j}^{(m-1)} P_{i,r} \end{aligned}$$

最后一步代入的是 $\mathrm{P}\{X_{m-1} = j, X_k \notin S \text{ for all } k = 1, \ldots, m-2 | X_0 = r\} = Q_{r,j}^{(m-1)}$

(它是起始未命中 $(r \notin S)$，目标时刻也未命中 $(j \notin S)$ 的情况)

(S. Ross 例 4.13)

考虑一个状态为 $1, 2, 3, 4, 5$ 的 Markov 链 $\{X_n : n \in \mathbb{N}\}$.

我们想要计算 $\mathrm{P}\{X_4 = 2, X_3 \leq 2, X_2 \leq 2, X_1 \leq 2 | X_0 = 1\}$ 的值，

即要计算 $\mathrm{P}\{X_4 = 2, X_k \notin \{3, 4, 5\} \text{ for all } k = 1, 2, 3 | X_0 = 1\}$ 的值，

即从状态 1 出发，在时刻 4 进入状态 2，并且从未进入集合 $S = \{3, 4, 5\}$ 的概率.

为计算此概率，我们只需要转移概率 $P_{11}, P_{12}, P_{21}, P_{22}$.

假定 $\begin{cases} P_{11} = 0.3, \ P_{12} = 0.3 \\ P_{21} = 0.1, \ P_{22} = 0.2 \end{cases}$

我们构造新的 Markov 链 $\{Y_n : n \in \mathbb{N}\}$，其状态为 $1, 2, 3$

我们定义状态 3 为吸收态，$\{Y_n : n \in \mathbb{N}\}$ 一旦进入状态 3 就永远保持在其中.

$\{Y_n : n \in \mathbb{N}\}$ 的转移概率矩阵为:

$$Q = \begin{bmatrix} 0.3 & 0.3 & 0.4 \\ 0.1 & 0.2 & 0.7 \\ & & 1 \end{bmatrix}$$

$Q \to Q^2 \to Q^4$ 得到 $Q^{(4)} = \begin{bmatrix} 0.0219 & 0.0285 & 0.9496 \\ 0.0095 & 0.0124 & 0.9781 \\ & & 1 \end{bmatrix}$

因此我们有:

$$\begin{aligned} \mathrm{P}\{X_4 = 2, X_3 \le 2, X_2 \le 2, X_1 \le 2 | X_0 = 1\} &= \mathrm{P}\{X_4 = 2, X_k \notin \{3,4,5\} \text{ for all } k = 1,2,3 | X_0 = 1\} \\ &= \mathrm{P}\{W_4 = 2 | X_0 = 1\} \\ &= \mathrm{P}\{W_4 = 2 | W_0 = 1\} \\ &= Q_{1,2}^{(4)} \\ &= 0.0285 \end{aligned}$$

4.2.3 命中概率

考虑一个具有转移概率 P_{ij} 的 Markov 链 $\{X_n : n \in \mathbb{N}\}$.
以 $S \subset \mathbb{N} = \{0, 1, 2, \dots\}$ 记某个状态集合，
我们想求此 Markov 链在时刻 m 进入状态 $j \in S$，但之前从没有进入过 S 中任意状态的概率.
也就是说，给定初始状态 $X_0 = i$ 和目标状态 $X_m = j \in S$，
我们想求 $\mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1, \dots, m-1 | X_0 = i\}$ 的值.

参考 4.2.1 节的内容，我们定义新的 Markov 链 $\{Y_n : n \in \mathbb{N}\}$

$$W_n = \begin{cases} X_n & \text{if } n < N \\ s_0 & \text{if } n \ge N \end{cases}$$

其中 $N = \min\{n : X_n \in S\}$ 代表 $\{X_n : n \in \mathbb{N}\}$ 首次进入 S 中的某个状态的时间.

$\{Y_n : n \in \mathbb{N}\}$ 的转移概率为 $\begin{cases} Q_{i,j} = P_{i,j} & \text{for all } i, j \notin S \\ Q_{i,s_0} = \sum\limits_{j \in S} P_{i,j} & \text{for all } i \notin S \\ Q_{s_0,s_0} = 1 \end{cases}$

对于 $i \notin S, j \in S$ 的情况，我们可以确定概率 (对时刻 $m-1$ 的状态取条件):

$$\begin{aligned} &\mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1, \dots, m-1 | X_0 = i\} \\ &= \sum_{r \notin S} \mathrm{P}\{X_m = j, X_{m-1} = r, X_k \notin S \text{ for all } k = 1, \dots, m-2 | X_0 = i\} \\ &= \sum_{r \notin S} \mathrm{P}\{X_m = j | X_{m-1} = r, X_k \notin S \text{ for all } k = 1, \dots, m-2, X_0 = i\} \\ &\quad \times \mathrm{P}\{X_{m-1} = r, X_k \notin S \text{ for all } k = 1, \dots, m-2 | X_0 = i\} \\ &= \sum_{r \notin S} \mathrm{P}\{X_m = j | X_{m-1} = r\} \cdot \mathrm{P}\{X_{m-1} = r, X_k \notin S \text{ for all } k = 1, \dots, m-2 | X_0 = i\} \\ &= \sum_{r \notin S} P_{r,j} \cdot Q_{i,r}^{(m-1)} \end{aligned}$$

对于 $i \in S, j \in S$ 的情况，我们可以确定概率 (对时刻 1 的状态取条件):

$$
\begin{aligned}
&\mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1,\dots,m-1 | X_0 = i\} \\
&= \sum_{r \notin S} \mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 1,\dots,m-1 | X_1 = r, X_0 = i\} \cdot \mathrm{P}\{X_1 = r | X_0 = i\} \\
&= \sum_{r \notin S} \mathrm{P}\{X_m = j, X_k \notin S \text{ for all } k = 2,\dots,m-1 | X_1 = r\} \cdot P_{i,r} \\
&= \sum_{r \notin S} \mathrm{P}\{X_{m-1} = j, X_k \notin S \text{ for all } k = 1,\dots,m-2 | X_0 = r\} \cdot P_{i,r} \\
&= \sum_{r \notin S} (\sum_{s \notin S} P_{s,j} Q_{r,s}^{(m-1)}) P_{i,r}
\end{aligned}
$$

最后一步代入的是 $\mathrm{P}\{X_{m-1} = j, X_k \notin S \text{ for all } k = 1,\dots,m-2 | X_0 = r\} = \sum_{s \notin S} P_{s,j} Q_{r,s}^{(m-1)}$

(它是起始未命中 $(r \notin S)$，目标时刻命中 $(j \in S)$ 的情况)

4.3 状态的分类

4.3.1 状态类

若存在某个 $n \geq 0$ 使得 $P_{ij}^{(n)} > 0$，则我们称状态 j 是**状态 i 可达的** (accessible from state i).
我们可以断言:
从状态 i 开始的 Markov 过程最终可能到达状态 j，当且仅当状态 j 是状态 i 可达的.

- 右侧条件的充分性是显然的，下面我们用反证法证明必要性:
  (反证法) 假设状态 j 不是状态 i 可达的，那么我们有:

$$
\begin{aligned}
\mathrm{P}\{\text{最终进入状态 } j \,|\text{初始状态为状态 } i\} &= \mathrm{P}\{\bigcup_{n=0}^{\infty}\{X_n = j\} | X_0 = i\} \\
&\leq \sum_{n=0}^{\infty} \mathrm{P}\{X_n = j | X_0 = i\} \\
&= \sum_{n=0}^{\infty} \mathrm{P}_{ij}^{(n)} \quad (\forall\, n \in \mathbb{N},\ P_{ij}^{(n)} = 0) \\
&= 0
\end{aligned}
$$

互相可达的两个状态 i, j 称为**互通的** (communicate)，记为 $i \leftrightarrow j$
特殊地，任意状态 i 和它自身都是互通的，因为 $P_{ii}^{(0)} = \mathrm{P}\{X_0 = i | X_0 = i\} = 1$

互通关系 $\leftrightarrow$ 是一个**等价关系**，它满足:

- 自反性: 对于任意 $i \geq 0$ 都有 $i \leftrightarrow i$
- 对称性: 若 $i \leftrightarrow j$，则 $j \leftrightarrow i$
- 传递性: 若 $\begin{cases} i \leftrightarrow j \\ j \leftrightarrow k \end{cases}$，则 $i \leftrightarrow k$

我们称两个互通的状态位于同一个**状态类** (state class) 中.
任意两个状态类或者相同，或者不相交.
也就是说，互通关系将状态空间分为许多分离的状态类.
若一个 Markov 链的状态空间只有一个状态类，即所有状态都是互通的，
则我们称其为**不可约的** (irreducible).

4.3.2 周期

对于任意状态 i，我们记其周期为 $d_i = \gcd\{n \geq 1 : P_{ii}^{(n)} > 0\}$
(其中 $\gcd$ 代表最大公因子)
假设 $d_i = d$，那么存在 $m \in \mathbb{N}$ 使得对于任意 $n \geq m$ 都有 $P_{ii}^{(nd)} > 0$ 成立.

若 $d_i = 1$，则我们称状态 i 是**非周期的**;
若 $d_i > 2$，则我们称状态 i 是**周期的**;

定理 4.3.1: (互通的状态具有相同的周期, 苏中根 定理 4.1)
若状态 i 和状态 j 是互通的 (即 $i \leftrightarrow j$)，则有 $d_i = d_j$ 成立.

- **Proof:**
  注意到 d_i 和 d_j 都是自然数，因此只要证明 d_i 和 d_j 相互整除即可.
  不失一般性，假设 $i \neq j$.
  根据 $i \leftrightarrow j$ 可知，存在 $k_1, k_2 \geq 1$ 使得 $\begin{cases} P_{ji}^{(k_1)} > 0 \\ P_{ij}^{(k_2)} > 0 \end{cases}$ 成立.
  于是我们有 $P_{jj}^{(k_1+k_2)} \geq P_{ji}^{(k_1)} P_{ij}^{(k_2)} > 0$ 成立.
  (第一个不等号 $\geq$ 是因为右侧路径只是左侧路径中的一部分)
  因此 $d_j | (k_1 + k_2)$
  对于任意 n 使得 $P_{ii}^{(n)} > 0$，我们都有 $P_{jj}^{(k_1+n+k_2)} \geq P_{ji}^{(k_1)} P_{ii}^{(n)} P_{ij}^{(k_2)} > 0$ 成立.
  因此 $d_j | (k_1 + n + k_2)$
  结合 $\begin{cases} d_j | (k_1 + k_2) \\ d_j | (k_1 + n + k_2) \text{ for all } n \text{ such that } P_{ii}^{(n)} > 0 \\ d_i = \gcd\{n \geq 1 : P_{ii}^{(n)} > 0\} \end{cases}$ 可知 $d_j | d_i$
  类似地，我们可以证明 $d_i | d_j$
  这样我们就有 $d_i = d_j$，命题得证.

4.3.3 状态的分类

给定状态 j，定义 $T_j = \inf\{n \geq 1 : X_n = j\}$
我们约定 $\inf\{\emptyset\} = \infty$，即若不存在 $n \geq 1$ 使得 $X_n = j$，则令 $T_j = \infty$.

- 当 $X_0 = i \neq j$ 时，T_j 代表从状态 i 出发的 Markov 链首次到达状态 j 的时刻;
- 当 $X_0 = j$ 时，T_j 代表从状态 j 出发的 Markov 链首次返回状态 j 的时刻;

常返态、瞬时态的第一种等价定义:
对于任意状态 i，
我们定义从状态 i 开始的过程在时刻 $n \geq 1$ 首次返回状态 i 的概率为:
$$\begin{aligned} f_{ii}^{(n)} &= \mathrm{P}\{T_i = n | X_0 = i\} \\ &= \mathrm{P}\{X_n = i, X_k \neq i \text{ for all } k = 1, \ldots, n-1 | X_0 = j\} \end{aligned}$$

我们记 $f_{ii} = \mathrm{P}\{T_i < \infty | X_0 = i\}$ 为从状态 i 开始的过程迟早返回状态 i 的概率.
我们可以将 f_i 表示为:

$$\begin{aligned} f_{ii} &= \mathrm{P}\{T_i < \infty | X_0 = i\} \\ &= \sum_{n=1}^{\infty} \mathrm{P}\{T_i = n | X_0 = i\} \\ &= \sum_{n=1}^{\infty} f_{ii}^{(n)} \end{aligned}$$

- 若 $0 \leq f_{ii} < 1$，则我们称状态 i 为**瞬时态** (transient);
- 若 $f_{ii} = 1$，则我们称状态 i 为**常返态** (recurrent);
  对于常返态 i，我们定义**平均常返时间**为 $\tau_i = \mathrm{E}[T_i | X_0 = i]$
  我们可以将 τ_i 表示为:

  $$\begin{aligned} \tau_i &= \mathrm{E}[T_i | X_0 = i] \\ &= \sum_{n=1}^{\infty} n \mathrm{P}\{T_i = n | X_0 = i\} \\ &= \sum_{n=1}^{\infty} n f_{ii}^{(n)} \end{aligned}$$

  - 若 $\tau_i < \infty$，则我们称状态 i 为**正常返态**
  - 若 $\tau_i = \infty$，则我们称状态 i 为**零常返态**

$f_{ii} = \mathrm{P}\{T_i < \infty | X_0 = i\}$ 为从状态 i 开始的过程迟早返回状态 i 的概率.
我们容易知道:

- 如果状态 i 是**常返态**，那么从状态 i 开始的过程将一再返回状态 i (事实上是无穷多次)
- 如果状态 i 是**瞬时态**，
  那么过程每次进入状态 i 都有一个正的概率 $1 - f_{ii}$ 不再进入这个状态.
  所以，从状态 i 开始的过程恰好在状态 i 停留 $n \geq 1$ 个时间周期的概率为 $(f_{ii})^{n-1}(1 - f_{ii})$
  换句话说，从状态 i 开始的过程在状态 i 停留的时间周期数 N_i 服从几何分布 $\mathrm{Geo}(1 - f_{ii})$
  其期望为有限值 $\mathrm{E}[N_i] = \frac{1}{1-f_{ii}}$

我们可以直接得到**常返态、瞬时态的第二种等价定义:**
定理 4.3.2: (常返态、瞬时态的第二种等价定义, 苏中根 定理 4.8)

- 状态 i 是常返态 (recurrent)，如果 $\mathrm{P}\{N_i = \infty | X_0 = i\} = 1$
- 状态 i 是瞬时态 (transient)，如果 $\mathrm{P}\{N_i = \infty | X_0 = i\} = 0$

它与第一种定义是等价的，这是因为:

$$\begin{aligned} \mathrm{P}\{N_i \geq n | X_0 = i\} &= \mathrm{P}\{T_i < \infty | X_0 = i\} \cdot \mathrm{P}\{N_i \geq n-1 | X_0 = i\} \\ &= (\mathrm{P}\{T_i < \infty | X_0 = i\})^2 \cdot \mathrm{P}\{N_i \geq n-2 | X_0 = i\} \\ &= \cdots \\ &= (\mathrm{P}\{T_i < \infty | X_0 = i\})^n \\ &= (f_{ii})^n \end{aligned}$$

因此我们有:

$$\begin{aligned} \mathrm{P}\{N_i = \infty | X_0 = i\} &= \lim_{n \to \infty} \mathrm{P}\{N_i = \infty | X_0 = i\} \\ &= \lim_{n \to \infty} (f_i)^n \\ &= \begin{cases} 1, & \text{if } f_{ii} = 1 \\ 0, & \text{if } 0 \leq f_{ii} < 1 \end{cases} \end{aligned}$$

> **常返态、瞬时态的第一种等价定义:**

> - 若 $0 \le f_{ii} < 1$，则我们称状态 i 为**瞬时态** (transient);
> - 若 $f_{ii} = 1$，则我们称状态 i 为**常返态** (recurrent);

进一步，状态 i 是**常返态的充要条件**是:
从状态 i 开始的过程在状态 i 停留的时间周期数 N_i 的期望 $\mathrm{E}[N_i] = \infty$

定义 $I_n = \begin{cases} 1 & \text{if } X_n = i \\ 0 & \text{if } X_n \ne i \end{cases}$
则我们有 $N_i = \sum\limits_{n=0}^{\infty} I_n$，表示从状态 i 开始的过程在状态 i 停留的时间周期数.
再根据:

$$\begin{aligned} \mathrm{E}[N_i|X_0 = i] &= \mathrm{E}[\sum_{n=0}^{\infty} I_n|X_0 = I] \\ &= \sum_{n=0}^{\infty} \mathrm{E}[I_n|X_0 = i] \\ &= \sum_{n=0}^{\infty} \mathrm{P}\{X_n = i|X_0 = i\} \\ &= \sum_{n=0}^{\infty} P_{ii}^{(n)} \end{aligned}$$

我们就证明了如下的命题:

定理 4.3.3: (常返态、瞬时态的第三种等价定义, S. Ross 命题 4.1)

- 状态 i 是常返态 (recurrent)，如果 $\sum\limits_{n=0}^{\infty} P_{ii}^{(n)} = \infty$
- 状态 i 是瞬时态 (transient)，如果 $\sum\limits_{n=0}^{\infty} P_{ii}^{(n)} < \infty$

证明定理 4.3.3 的推理过程显然更加重要，
它表明了**一个瞬时态只能被访问有限次**.
因此一个有限状态空间的 Markov 链中，所有的状态不可能都是瞬时态.
(反证法) 假设有限状态空间为 $\{0, 1, \dots, M\}$ 且所有状态都是瞬时态，
则存在有限时间 $T_0, T_1, \dots, T_M$，使得时间 T_i 后，状态 i 将不再被访问 $(i = 0, 1, \dots, M)$.
那么经过有限时间 $T = \max\{T_0, T_1, \dots, T_M\}$ 后该 Markov 过程将无状态可访，产生矛盾.
因此对于有限状态空间的 Markov 链，它至少有一个常返态.

作为定理 4.3.3 的直接推论，我们可知**常返性是一个类性质** (recurrence is a class property)
推论 4.3.4: (常返性是一个类性质, S. Ross 推论 4.2, 苏中根 定理 4.2)
若状态 i 是常返态，而状态 j 与状态 i 互通，则状态 j 也是常返态.

- **(S. Ross 命题 4.5)**
  若状态 i 是正常返态，而状态 j 与状态 i 互通，则状态 j 也是正常返态.
  若状态 i 是零常返态，而状态 j 与状态 i 互通，则状态 j 也是零常返态.
- 我们也容易知道: **瞬时性也是一个类性质**
  也就是说，若状态 i 是瞬时态，而状态 j 与状态 i 互通，则状态 j 也是瞬时态.
  (假如状态 j 为常返态，则根据推论 4.3.4 可知状态 i 是常返态，这与我们假设矛盾)
- 进一步，我们知道:
  状态空间有限且不可约的 Markov 链的所有状态都是常返态.
  (因为它只有一个状态类，且至少有一个常返态 (因为不能全为瞬时态)，故所有状态均为常返态)

更具体来说，**状态空间有限且不可约的 Markov 链的所有状态都是正常返态.**
因为不能全为零常返态，
否则根据 4.4.1 节的**推论 4.4.2** 可知所有的极限概率 $\lim\limits_{n\to\infty} p_{ij}^{(n)}$ 都是 0，
而这是不可能的，因为这个 Markov 链状态有限.

- **Proof:**
  (我们这里只证明常返性是一个类性质，正常返、零常返的结论后面会证明)
  由于状态 j 与状态 i 互通，故存在 $k_1, k_2 \in \mathbb{N}$ 使得 $\begin{cases} P_{ij}^{(k_1)} > 0 \\ P_{ji}^{(k_2)} > 0 \end{cases}$
  注意到: 对于任意整数 n 都有 $P_{jj}^{(k_2+n+k_1)} \geq P_{ji}^{(k_2)} P_{ii}^{(n)} P_{ij}^{(k_1)}$
  这是显然的，因为右侧实现的转移路径只是左侧实现的转移路径中的一部分.
  对 n 求和，我们得到:
  $$\begin{aligned}\sum_{n=0}^{\infty} P_{jj}^{(k_2+n+k_1)} &\geq \sum_{n=0}^{\infty} P_{ji}^{(k_2)} P_{ii}^{(n)} P_{ij}^{(k_1)} \\ &= P_{ji}^{(k_2)} P_{ij}^{(k_1)} \sum_{n=0}^{\infty} P_{ii}^{(n)} \quad (P_{ji}^{(k_2)} P_{ij}^{(k_1)} > 0) \\ &= \infty\end{aligned}$$
  因此我们有:
  $$\begin{aligned}\sum_{n=0}^{\infty} P_{jj}^{(n)} &= \sum_{n=0}^{k_1+k_2-1} P_{jj}^{(n)} + \sum_{n=k_1+k_2}^{\infty} P_{jj}^{(n)} \\ &= \sum_{n=0}^{k_1+k_2-1} P_{jj}^{(n)} + \sum_{n=0}^{\infty} P_{jj}^{(k_2+n+k_1)} \\ &= \infty\end{aligned}$$
  根据**定理 4.3.3** 给出的常返态的等价定义，我们知道状态 j 也是常返态.
  命题得证.

定理 4.3.5: (不同常返态的等价类是互不相交的闭集, 苏中根 定理 4.3)
对于常返态 i，其等价类 $C_i = \{j \in \mathbb{N} : j \leftrightarrow i\}$ 是闭集，
即不存在 $\begin{cases} j \in C_i \\ k \notin C_i \end{cases}$ 使得 $j \to k$ 成立.

定理 4.3.6: (状态空间的分解, 苏中根 定理 4.4)
对于一般的状态空间 $\mathcal{E}$，它可以分解为:
$\mathcal{E} = C_1 + C_2 + \cdots + C_M + \mathcal{N}$
其中 $\{C_i\}$ 表示一系列互不相交的常返态等价类 (M 可以取 ∞)，
而 $\mathcal{N}$ 是瞬时态的全体构成的集合.

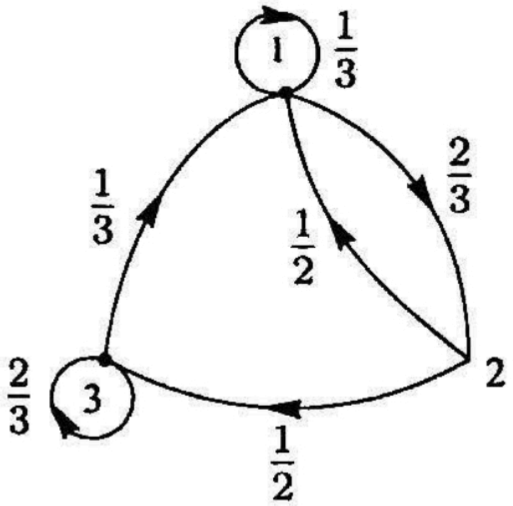
4.3.4 示例

(S. Ross 例 4.17)
考虑状态为 $0, 1, 2, 3, 4$ 的 Markov 链的概率转移矩阵:
$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ & & \frac{1}{2} & \frac{1}{2} & \\ & & \frac{1}{2} & \frac{1}{2} & \\ \frac{1}{4} & \frac{1}{4} & & & \frac{1}{2} \end{bmatrix}$$
这个 Markov 链由三个类构成: $\{0, 1\}$, $\{2, 3\}$ 和 $\{4\}$

前两个类的状态都是常返态，而第三个类的状态是瞬时态 (其返回次数服从几何分布 $\mathrm{Geo}(\frac{1}{2})$)

(苏中根 例 4.18)
考虑状态为 $1, 2, 3$ 的 Markov 链的概率转移矩阵:
$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \\ \frac{1}{2} & & \frac{1}{2} \\ \frac{1}{3} & & \frac{2}{3} \end{bmatrix}$$
这个 Markov 链具有有限状态空间.
容易验证所有状态都是互通的，因而状态空间是不可约的.
根据推论 4.3.4 的结论，**状态空间有限且不可约的 Markov 链的所有状态都是正常返态.**

容易看出状态 1 的周期为 1，因而所有状态的周期都是 1.



下面考虑通过计算验证状态 1 的常返性:
回顾本节的开头，定义状态 1 的首次返回时间为 $T_1 = \inf\{n \geq 1 : X_n = 1\}$
我们不难得到以下概率:

- $f_{11}^{(1)} = \mathrm{P}\{T_1 = 1 | X_0 = 1\} = \frac{1}{3}$
- $f_{11}^{(2)} = \mathrm{P}\{T_1 = 2 | X_0 = 1\} = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$
- $f_{11}^{(3)} = \mathrm{P}\{T_1 = 3 | X_0 = 1\} = \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{9}$
- 继续下去，对于任意 $n \geq 4$，我们有:
  $f_{11}^{(n)} = \mathrm{P}\{T_1 = n | X_0 = 1\} = \frac{2}{3} \cdot \frac{1}{2} \cdot (\frac{2}{3})^{n-3} \cdot \frac{1}{3} = \frac{1}{9}(\frac{2}{3})^{n-3}$
  (实际上 $n = 3$ 也满足这个通式)

因此我们有:

$$
\begin{aligned}
f_1 &= \mathrm{P}\{T_1 < \infty | X_0 = 1\} \\
&= \sum_{n=1}^{\infty} f_{11}^{(n)} \\
&= f_{11}^{(1)} + f_{11}^{(2)} + \sum_{n=3}^{\infty} f_{11}^{(n)} \\
&= \frac{1}{3} + \frac{1}{3} + \sum_{n=3}^{\infty} \frac{1}{9} (\frac{2}{3})^{n-3} \\
&= \frac{2}{3} + \frac{1}{9} \frac{1-0}{1-\frac{2}{3}} \\
&= 1
\end{aligned}
$$

所以状态 1 是常返的.

考虑计算状态 1 的平均常返时间 τ_1:

$$
\begin{aligned}
\tau_1 &= \mathrm{E}[T_1 | X_0 = 1] \\
&= \sum_{n=1}^{\infty} n \mathrm{P}\{T_1 = n | X_0 = 1\} \\
&= \sum_{n=1}^{\infty} n f_{11}^{(n)} \\
&= 1 \cdot f_{11}^{(1)} + 2 \cdot f_{11}^{(2)} + \sum_{n=3}^{\infty} n f_{11}^{(n)} \\
&= 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} + \sum_{n=3}^{\infty} n \cdot \frac{1}{9} (\frac{2}{3})^{n-3} \\
&= 1 + \frac{1}{9} \sum_{n=3}^{\infty} n (\frac{2}{3})^{n-3}
\end{aligned}
$$

我们记 $\mathrm{sum} = \sum_{n=3}^{\infty} n(\frac{2}{3})^{n-3}$

容易知道 $\frac{2}{3}\mathrm{sum} = \frac{2}{3}\sum_{n=3}^{\infty} n(\frac{2}{3})^{n-3} = \sum_{n=3}^{\infty}(n-1)(\frac{2}{3})^{n-3} - 2$

两式相减得到 $\frac{1}{3}\mathrm{sum} = \sum_{n=3}^{\infty}(\frac{2}{3})^{n-3} + 2 = \frac{1-0}{1-\frac{2}{3}} + 2 = 5$

因此 $\mathrm{sum} = 15$

于是我们知道:

$$
\begin{aligned}
\tau_1 &= 1 + \frac{1}{9} \sum_{n=3}^{\infty} n (\frac{2}{3})^{n-3} \\
&= 1 + \frac{1}{9} \mathrm{sum} \\
&= 1 + \frac{15}{9} \\
&= \frac{8}{3} < \infty
\end{aligned}
$$

因此状态 1 的**平均常返时间** τ_1 是一个有限值，故状态 1 是**正常返态**.

由于所有状态都是互通的，故所有状态都是**正常返态**.

(王勤文 补充示例)

考虑具有状态 $1, 2$ 的 Markov 过程，其概率转移矩阵为:

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

其中 $0 < p < q < 1$

根据推论 4.3.4 的结论，**状态空间有限且不可约的 Markov 链的所有状态都是正常返态.**

下面考虑通过计算验证状态 1 的常返性:

回顾本节的开头，定义状态 1 的首次返回时间为 $T_1 = \inf\{n \geq 1 : X_n = 1\}$

我们不难得到以下概率:

- $f_{11}^{(1)} = \mathrm{P}\{T_1 = 1 | X_0 = 1\} = (1-p)$
- $f_{11}^{(2)} = \mathrm{P}\{T_1 = 2 | X_0 = 1\} = pq$
- 继续下去，对于任意 $n \geq 3$，我们有:
  $f_{11}^{(n)} = \mathrm{P}\{T_1 = n | X_0 = 1\} = p(1-q)^{n-2}q$
  (实际上 $n = 2$ 也满足这个通式)

因此我们有:

$$\begin{aligned} f_1 &= \mathrm{P}\{T_1 < \infty | X_0 = 1\} \\ &= \sum_{n=1}^{\infty} f_{11}^{(n)} \\ &= f_{11}^{(1)} + \sum_{n=2}^{\infty} f_{11}^{(n)} \\ &= (1-p) + \sum_{n=2}^{\infty} pq(1-q)^{n-2} \\ &= 1 - p + pq \cdot \frac{1-0}{1-(1-q)} \\ &= 1 \end{aligned}$$

所以状态 1 是常返的.

考虑计算状态 1 的平均常返时间 τ_1:

$$\begin{aligned} \tau_1 &= \mathrm{E}[T_1 | X_0 = 1] \\ &= \sum_{n=1}^{\infty} n\mathrm{P}\{T_1 = n | X_0 = 1\} \\ &= \sum_{n=1}^{\infty} n f_{11}^{(n)} \\ &= 1 \cdot f_{11}^{(1)} + \sum_{n=2}^{\infty} n f_{11}^{(n)} \\ &= 1 \cdot (1-p) + \sum_{n=2}^{\infty} n \cdot pq(1-q)^{n-2} \\ &= 1 - p + pq \sum_{n=2}^{\infty} n(1-q)^{n-2} \end{aligned}$$

我们记 $\mathrm{sum} = \sum_{n=2}^{\infty} n(1-q)^{n-2}$

容易知道 $(1-q)\mathrm{sum} = (1-q)\sum_{n=2}^{\infty} n(1-q)^{n-2} = \sum_{n=2}^{\infty} (n-1)(1-q)^{n-2} - 1$

两式相减得到 $q \cdot \mathrm{sum} = \sum_{n=2}^{\infty} (1-q)^{n-2} + 1 = \frac{1-0}{1-(1-q)} + 1 = \frac{1}{q} + 1$

因此 $\mathrm{sum} = \frac{1}{q}(\frac{1}{q} + 1)$

于是我们知道:

$$\begin{aligned}\tau_1 &= 1-p+pq\sum_{n=2}^{\infty}n(1-q)^{n-2} \\ &= 1-p+pq\cdot\text{sum} \\ &= 1-p+pq\cdot\frac{1}{q}(\frac{1}{q}+1) \\ &= \frac{p+q}{q}<\infty\end{aligned}$$

因此状态 1 的**平均常返时间** τ_1 是一个有限值，故状态 1 是**正常返态**.
由于所有状态都是互通的，故所有状态都是**正常返态**.

(简单随机游动, S. Ross 例 4.18 & 苏中根 例 4.20)
一个人在直线上行走，他在每一个时间点以概率 p 向右走一步，或以概率 $1-p$ 向左走一步.
其中 $p\in(0,1)$
这个随机过程的状态空间为 $i=0,\pm1,\pm2,\cdots$
任意状态 i 的转移概率为 $\begin{cases}p_{i,i+1}=p \\ p_{i,i-1}=1-p\end{cases}$

因为其所有状态都是互通的，由推论 4.3.3 可知它们要么都是常返态，要么都是瞬时态.
所以我们考察状态 0，尝试确定 $\sum\limits_{n=0}^{\infty}P_{00}^{(n)}$ 是有限还是无穷大.

- 奇数次行走后肯定不能回到原点 0，因此 $P_{00}^{(2n-1)}=0\ \ (\forall\ n=1,2,\dots)$
- 偶数次行走最终回到原点 0 当且仅当我们一半步数向右走，一半步数向左走.
  这是一个二项概率: $P_{00}^{(2n)}=\binom{2n}{n}p^n(1-p)^n=\frac{(2n)!}{n!n!}(p(1-p))^n\ \ (\forall\ n=0,1,\dots)$

Stirling 公式表明: $n!\sim\sqrt{2\pi n}(\frac{n}{e})^n$
(其中 $a_n\sim b_n$ 代表 $\lim\limits_{n\to\infty}\frac{a_n}{b_n}=1$，易知 $\sum\limits_{n=0}^{\infty}a_n<\infty$ 当且仅当 $\sum\limits_{n=0}^{\infty}b_n<\infty$)
因此 $P_{00}^{(2n)}\sim\frac{\sqrt{2\pi\cdot2n}(\frac{2n}{e})^{2n}}{\sqrt{2\pi n}(\frac{n}{e})^n\sqrt{2\pi n}(\frac{n}{e})^n}(p(1-p))^n=\frac{(4p(1-p))^n}{\sqrt{\pi n}}$
我们容易验证 $\sum\limits_{n=0}^{\infty}\frac{(4p(1-p))^n}{\sqrt{\pi n}}<\infty$ 当且仅当 $4p(1-p)<1$ (这个不等式解得 $p\in(0,\frac{1}{2})\cup(\frac{1}{2},1)$)
因此 $\sum\limits_{n=0}^{\infty}P_{00}^{(n)}=0+\sum\limits_{n=0}^{\infty}P_{00}^{(2n)}<\infty$ 当且仅当 $p\in(0,\frac{1}{2})\cup(\frac{1}{2},1)$

综上所述:
当 $p=\frac{1}{2}$ 时 (对称随机游动)，这个 Markov 链是**(零)常返的**;
当 $p\neq\frac{1}{2}$ 时 (非对称随机游动)，这个 Markov 链是**瞬时的**.

- 若 $p\in(\frac{1}{2},1)$，则 $X_n\to\infty\ (n\to\infty)$
  若 $p\in(-1,\frac{1}{2})$，则 $X_n\to-\infty\ (n\to\infty)$
  这直观上说明:
  随着时间推移，非对称随机游动 X_n 离任何一个状态都将越来越远，
  不会 "经常性地" 返回某一状态.

二维情况下，我们也可以证明**只有对称的二维随机游动是常返的**.
(对于二维情况，我们定义每次转移都以 $\frac{1}{4}$ 的概率向左、右、上、下四个方向之一移动一步)
直观来说:
一维对称情况 $P_{00}^{(2n)}$ 为 $O(\frac{1}{\sqrt{n}})$ 阶，二维对称情况 $P_{00}^{(2n)}$ 为 $O(\frac{1}{n})$ 阶，
因而求和 $\sum\limits_{n=0}^{\infty}P_{00}^{(n)}=0+\sum\limits_{n=0}^{\infty}P_{00}^{(2n)}$ 都是发散的，说明是常返的.

但是对于更高维的情况 (例如三维情况)，即使是对称的随机游动也是瞬时的.
三维对称情况 $P_{00}^{(2n)}$ 为 $O(\frac{1}{n\sqrt{n}})$ 阶,
因而求和 $\sum\limits_{n=0}^{\infty} P_{00}^{(n)} = 0 + \sum\limits_{n=0}^{\infty} P_{00}^{(2n)}$ 是收敛的，说明是瞬时的.
(更多细节请参考 S. Ross 例 4.18 以及苏中根 例 4.20)

4.4 平稳概率 & 极限概率

4.4.1 极限概率

定理 4.4.1: (苏中根 定理 4.6)

- 若状态 i 是瞬时态，则有 $\lim\limits_{n\to\infty} P_{ii}^{(n)} = 0$ 成立.
  (这一点可由**定理 4.3.3 第三种等价定义**的 $\sum\limits_{n=0}^{\infty} P_{ii}^{(n)} < \infty$ 直接得到)
- 假设状态 i 是常返态，记其平均常返时间为 $\tau_i = \mathrm{E}[T_i|X_0 = i]$
  - 若状态 i 是正常返态 $(\tau_i < \infty)$，周期为 d_i (非周期的情况我们就当 $d_i = 1$ 好了)
    则有 $\lim\limits_{n\to\infty} P_{ii}^{(n)} = \frac{d_i}{\tau_i}$ 成立.
  - 若状态 i 是零常返态 $(\tau_i = \infty)$，则有 $\lim\limits_{n\to\infty} P_{ii}^{(n)} = 0$ 成立.

推论 4.4.2: (苏中根 推论 4.1)

- 若状态 j 是瞬时态或零常返态,
  则对于任意状态 i 都有 $\lim\limits_{n\to\infty} P_{ij}^{(n)} = 0$ 成立.
- 定义 $f_{ij} = \mathrm{P}\{T_j < \infty|X_0 = i\}$ 为从状态 i 开始的过程迟早到达状态 j 的概率,
  其中 $T_j = \inf\{n \geq 1 : X_n = j\}$
  (也即 $f_{ij} = \mathrm{P}\{X_n = j \text{ for some } n > 0|X_0 = i\}$)

  对于常返态 i，我们定义**平均常返时间**为 $\tau_i = \mathrm{E}[T_i|X_0 = i]$
  若状态 j 是非周期 $(d_i = 1)$ 正常返状态 $(\tau_i < \infty)$,
  则对于任意状态 i 都有 $\lim\limits_{n\to\infty} P_{ij}^{(n)} = \frac{f_{ij}}{\tau_j}$ 成立.
  - **(S. Ross 命题 4.3)**
    进一步，我们可以断言:
    若 i, j 常返且互通，则 $f_{ij} = 1$
    于是我们有 $\lim\limits_{n\to\infty} P_{ij}^{(n)} = \frac{1}{\tau_j}$ **(这给出了极限概率与平均常返时间的关系)**

    (S. Ross 命题 4.4)
    因此若 Markov 链是不可约且常返的,
    则对于任意初始状态 i 都有 $\lim\limits_{n\to\infty} P_{ij}^{(n)} = \frac{1}{\tau_j}$ 成立.

(苏中根 例 4.22)

考虑 2 状态 Markov 链，转移概率矩阵为 $P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$ (其中 $0 < p, q < 1$)
我们将其特征分解为 $P = Q\Lambda Q^{-1}$:

- 求特征值:
$$\begin{aligned}\det(\lambda I_2 - P) &= (\lambda - 1 + p)(\lambda - 1 + q) - pq \\ &= \lambda^2 + (p+q-2)\lambda + 1 - p - q + pq - pq \\ &= \lambda^2 + (p+q-2)\lambda + 1 - p - q \\ &= (\lambda - 1)(\lambda + p + q - 1)\end{aligned}$$
令上式 $=0$ 解得 $\lambda_1 = 1$ 和 $\lambda_2 = 1 - p - q$
因此 $\Lambda = \begin{bmatrix} 1 & \\ & 1-p-q \end{bmatrix}$
- 求特征向量:
$(\lambda_1 I_2 - P)x = \begin{bmatrix} p & -p \\ -q & q \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$ 可解出一个特征向量 $x^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$(\lambda_2 I_2 - P)x = \begin{bmatrix} -q & -p \\ -q & -p \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$ 可解出一个特征向量 $x^{(2)} = \begin{bmatrix} p \\ -q \end{bmatrix}$
因此 $Q = \begin{bmatrix} 1 & p \\ 1 & -q \end{bmatrix}$
而其逆矩阵 $Q^{-1} = \dfrac{1}{\det(Q)}\begin{bmatrix} -q & -p \\ -1 & 1 \end{bmatrix} = -\dfrac{1}{p+q}\begin{bmatrix} -q & -p \\ -1 & 1 \end{bmatrix}$

> 考虑 2×2 可逆矩阵 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 的求逆:
> 其逆矩阵为 $A^{-1} = \dfrac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

这样我们就得到了特征分解 $P = Q\Lambda Q^{-1}$
那么我们有:
$$\begin{aligned}\lim_{n\to\infty} P^n &= \lim_{n\to\infty}(Q\Lambda Q^{-1})^n \\ &= \lim_{n\to\infty} Q\Lambda^n Q^{-1} \\ &= \lim_{n\to\infty}\{\begin{bmatrix} 1 & p \\ 1 & -q \end{bmatrix}\begin{bmatrix} 1^n & \\ & (1-p-q)^n \end{bmatrix}\cdot(-\frac{1}{p+q})\begin{bmatrix} -q & -p \\ -1 & 1 \end{bmatrix}\} \\ &= -\frac{1}{p+q}\begin{bmatrix} 1 & p \\ 1 & -q \end{bmatrix}\begin{bmatrix} 1 & \\ & 0 \end{bmatrix}\begin{bmatrix} -q & -p \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{bmatrix}\end{aligned}$$
因此该 Markov 链一定存在极限分布 $\mu = (\frac{q}{p+q}, \frac{p}{p+q})$
表明无论初始状态 X_0 是 1 还是 2，它们很久以后进入状态 1 和状态 2 的概率服从上述极限分布.

(苏中根 例 4.24 并非所有 Markov 链都有极限分布)
考虑 3 状态 Markov 链，转移状态矩阵为 $P = \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}$
根据状态转换图和一些简单的计算，我们知道 $P^n = \begin{cases} P & \text{if } n = 3k+1 \\ P^2 & \text{if } n = 3k+2 \ (\forall\ k \geq 0) \\ P^3 & \text{if } n = 3k+3 \end{cases}$
而显然 P, P^2, P^3 并不两两相等,
也就是说序列 $\{P^n\}$ 具有 3 个收敛到不同极限的子列,
因此序列 $\{P^n\}$ 不收敛，表明极限分布不存在.

(苏中根 例 4.25 极限分布也可能依赖于初始状态)

考虑 4 状态 Markov 链，转移概率矩阵为 $P = \begin{bmatrix} 0.4 & 0.6 & & \\ 0.6 & 0.4 & & \\ & & 0.5 & 0.5 \\ & & 0.5 & 0.5 \end{bmatrix}$

注意这不是一个不可约 Markov 链.
通过计算可得 (实际上可以根据**苏中根 例 4.22** 的结论):

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \frac{0.6}{0.6+0.6} & \frac{0.6}{0.6+0.6} & & \\ \frac{0.6}{0.6+0.6} & \frac{0.6}{0.6+0.6} & & \\ & & \frac{0.5}{0.5+0.5} & \frac{0.5}{0.5+0.5} \\ & & \frac{0.5}{0.5+0.5} & \frac{0.5}{0.5+0.5} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & & \\ 0.5 & 0.5 & & \\ & & 0.5 & 0.5 \\ & & 0.5 & 0.5 \end{bmatrix}$$

与之前的例子不同，该 Markov 链的极限分布依赖于初始分布.
任意给定 $0 \le \alpha \le 1$

- 从初始分布 $(\alpha, 1-\alpha, 0, 0)$ 出发的 Markov 链的极限分布为 $(0.5, 0.5, 0, 0)$
- 从初始分布 $(0, 0, \alpha, 1-\alpha)$ 出发的 Markov 链的极限分布为 $(0, 0, 0.5, 0.5)$

任意给定 $0 \le \alpha_1, \alpha_2, \alpha_3, \alpha_4 \le 1$ 满足 $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$

- 从初始分布 $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ 出发的 Markov 链的极限分布为 $(\frac{\alpha_1+\alpha_2}{2}, \frac{\alpha_1+\alpha_2}{2}, \frac{\alpha_3+\alpha_4}{2}, \frac{\alpha_3+\alpha_4}{2})$

4.4.2 平稳概率

对于常返态 i，**平均常返时间** $\tau_i = \mathrm{E}[T_i | X_0 = i] < \infty$
其中 $T_i = \inf\{n \ge 1 : X_n = i\}$

我们定义向状态 i 转移的平稳概率为 $\pi_i = \frac{1}{\tau_i}$，
于是向状态 j 转移的平稳概率为 $\pi_j = \sum_i \pi_i P_{ij}$

事实上，我们可以证明以下的重要定理:
定理 4.4.3: (S. Ross 定理 4.1)
考虑一个不可约的 Markov 链.

- 若此 Markov 链是正常返的，
  则平稳分布 π 是线性方程组 $\begin{cases} \pi_j = \sum_i \pi_i P_{ij} & (\forall\ j \text{ by column}) \\ \sum_j \pi_j = 1 & \\ \pi_j \ge 0 & (\forall\ j) \end{cases}$ 的唯一解.
  其中 $\pi_j = \sum_i \pi_i P_{ij}\ \ (\forall\ j)$ 可以紧凑地写成 $\pi = \pi P$ (这里视 π 为行向量)
  - **(苏中根 定理 4.9)**
    考虑一个非周期不可约的 Markov 链.
    则当且仅当该 Markov 链正常返时，存在唯一的平稳分布.
    并且对于任意状态 $i \in \mathcal{E}$ 都有 $\pi_i = \frac{1}{\tau_i}$

    > 定义 $T_i = \inf\{n \ge 1 : X_n = i\}$
    > 对于常返态 i，我们定义**平均常返时间**为 $\tau_i = \mathrm{E}[T_i | X_0 = i]$

    (苏中根 推论 4.2)
    此时平稳分布 π 即为极限分布 μ.

一般情况下，极限分布与平稳分布之间并没有必然联系.
例如对于周期 Markov 链，上述推论不一定成立.

> 回顾 4.1 节的**苏中根 例 4.24**:
> 考虑 3 状态 Markov 链，转移状态矩阵为 $P = \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}$
> 我们证明过它没有极限分布.
> 但是它的平稳分布存在且唯一: $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$

- 若上述线性方程组无解，
  则此 Markov 链是瞬时的或零常返的，且 $\pi_j = 0 \ \ (\forall\, j)$.
  (根据 4.3.3 节的推论 4.3.4 可知，此 Markov 链的状态空间一定是无限的)

  (苏中根 注 4.4)
  我们认为，不可约瞬时或零常返 Markov 链没有平稳分布.

  - **瞬时**的例子: 直线上**非对称**简单随机游动没有平稳分布
  - **零常返**的例子: 直线上**对称**简单随机游动没有平稳分布

π_i 之所以称为**平稳概率**，是因为:
若初始状态按平稳分布选取，则在任意时刻 n 的状态分布仍为平稳分布.
即若 $p_0(i) = \mathrm{P} = \{X_0 = i\} = \pi_i \ \ (\forall\, i)$,
则我们有 $p_n(i) = \mathrm{P}\{X_n = i\} = \pi_j \ \ (\forall\, n, j)$ 成立.
上述性质由方程 $\pi P = \pi$ 保证.

存在平稳分布 π 的 Markov 链称为**平稳 Markov 链**

- 随机过程 $\mathbf{X} = \{X_n : n \geq 0\}$ 是**强平稳过程**,
  当且仅当对于任意 $\begin{cases} n \geq 0 \\ m \geq 1 \end{cases}$ 都有 $(X_m, X_{m+1}, \ldots, X_{m+n}) \overset{d}{=} (X_0, X_1, \ldots, X_n)$ 成立.
- 考虑 Markov 过程:
  我们记初始分布为 p_0，记时刻 n 的分布为 $p_n = p_0 P^{(n)} = p_0 P^n$
  显然，此 Markov 链为**强平稳过程**，当且仅当 $p_0 = p_1 = p_2 = \cdots$
  这个条件等价于 $p_0 = p_0 P = p_0 P^2 = \cdots$
  即等价于 $p_0 = p_0 P$

(苏中根 例 4.22 续)
考虑 2 状态 Markov 链，转移概率矩阵为 $P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$ (其中 $0 < p, q < 1$)
我们解方程 $\begin{cases} (1-p)\pi_1 + q\pi_2 = \pi_1 \\ p\pi_1 + (1-q)\pi_2 = \pi_2 \\ \pi_1 + \pi_2 = 1 \end{cases}$ 得到 $\begin{cases} \pi_1 = \frac{q}{p+q} \\ \pi_2 = \frac{p}{p+q} \end{cases}$ 恰与极限分布 $\begin{cases} \mu_1 = \frac{q}{p+q} \\ \mu_2 = \frac{p}{p+q} \end{cases}$ 相同.

(苏中根 例 4.23-例 4.26)
考虑 3 状态 Markov 链，转移概率矩阵为 $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{8} & \frac{3}{8} \\ 1 & & \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$

- 求转移概率矩阵的**谱分解** $P = Q\Lambda Q^{-1}$:

  $\det(\lambda I_3 - P) = (\lambda-1)(\lambda^2 + \frac{\lambda}{4} + \frac{5}{32}) = 0$ 解得 $\begin{cases}\lambda_1 = 1\\ \lambda_2 = \frac{1}{8}(-1+3i)\\ \lambda_3 = \frac{1}{8}(-1-3i)\end{cases}$

  对应的特征向量可以是 $\begin{cases}x^{(1)} = (1,1,1)^T\\ x^{(2)} = (-7-9i, -16+24i, 26)^T\\ x^{(3)} = (-7+9i, -16-24i, 26)^T\end{cases}$

  变换矩阵 $Q = \begin{bmatrix}1 & -7-9i & -7+9i\\ 1 & -16+24i & -16-24i\\ 1 & 26 & 26\end{bmatrix}$

  其逆矩阵为: **(伴随矩阵求逆法)**

  $$\begin{aligned}Q^{-1} &= \frac{1}{\det(Q)}\mathrm{adj}(Q)\\ &= \frac{1}{2340i}\begin{bmatrix}1248i & 468i & 624i\\ -42-24i & 33-9i & 9+33i\\ 42-24i & -33-9i & -9+33i\end{bmatrix}\\ &= \frac{1}{780}\begin{bmatrix}416 & 156 & 208\\ -8+14i & -3-11i & 11-3i\\ -8-14i & -3-11i & 11-3i\end{bmatrix}\end{aligned}$$

  其中**伴随矩阵** $\mathrm{adj}(Q)$ 每个元素是 Q^T 对应位置的代数余子式.

- 计算**极限分布**:

  $$\Lambda^n = \begin{bmatrix}1^n & & \\ & (\frac{-1+3i}{8})^n & \\ & & (\frac{-1-3i}{8})^n\end{bmatrix} \to \begin{bmatrix}1 & & \\ & 0 & \\ & & 0\end{bmatrix} (n\to\infty)$$

  因此我们有:

  $$\begin{aligned}\lim_{n\to\infty} P^n &= \begin{bmatrix}1 & -7-9i & -7+9i\\ 1 & -16+24i & -16-24i\\ 1 & 26 & 26\end{bmatrix}\begin{bmatrix}1 & & \\ & 0 & \\ & & 0\end{bmatrix}\frac{1}{780}\begin{bmatrix}416 & 156 & 208\\ -8+14i & -3-11i & 11-3i\\ -8-14i & -3-11i & 11-3i\end{bmatrix}\\ &= \begin{bmatrix}\frac{8}{15} & \frac{1}{5} & \frac{4}{15}\\ \frac{8}{15} & \frac{1}{5} & \frac{4}{15}\\ \frac{8}{15} & \frac{1}{5} & \frac{4}{15}\end{bmatrix}\end{aligned}$$

  因此极限分布为 $\mu = (\frac{8}{15}, \frac{1}{5}, \frac{4}{15})$

- 计算**平稳分布**:

  $\begin{cases}\pi = \pi P\\ 1_3^T\pi = 1\end{cases} \Rightarrow \begin{cases}\pi_1 = \frac{1}{2}\pi_1 + 1\pi_2 + \frac{1}{4}\pi_3\\ \pi_2 = \frac{1}{8}\pi_1 + 0\pi_2 + \frac{1}{2}\pi_3\\ \pi_3 = \frac{3}{8}\pi_1 + 0\pi_2 + \frac{1}{4}\pi_3\\ \pi_1+\pi_2+\pi_3 = 1\end{cases}$ (回顾 $P = \begin{bmatrix}\frac{1}{2} & \frac{1}{8} & \frac{3}{8}\\ 1 & & \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4}\end{bmatrix}$) 解得 $\begin{cases}\pi_1 = \frac{8}{15}\\ \pi_2 = \frac{1}{5}\\ \pi_3 = \frac{4}{15}\end{cases}$

  我们发现该 Markov 链的平稳分布 $\pi = (\frac{8}{15}, \frac{1}{5}, \frac{4}{15})$ 与极限分布 $\mu = (\frac{8}{15}, \frac{1}{5}, \frac{4}{15})$ 相同.

- 根据**定理 4.4.3** 可知，状态 $1,2,3$ 的**平均常返时间**分别为 $\begin{cases}\tau_1 = \frac{1}{\pi_1} = \frac{15}{8}\\ \tau_2 = \frac{1}{\pi_2} = 5\\ \tau_3 = \frac{1}{\pi_3} = \frac{15}{4}\end{cases}$

(苏中根 例 4.25 续-4.27, 平稳分布可能不唯一)

考虑 4 状态 Markov 链，转移概率矩阵为 $P = \begin{bmatrix}0.4 & 0.6 & & \\ 0.6 & 0.4 & & \\ & & 0.5 & 0.5\\ & & 0.5 & 0.5\end{bmatrix}$

注意这不是一个不可约 Markov 链.

$$\begin{cases}\pi=\pi P\\ 1_4^T\pi=1\end{cases}\Rightarrow\begin{cases}\pi_1=0.4\pi_1+0.6\pi_2\\ \pi_2=0.4\pi_1+0.6\pi_2\\ \pi_3=0.5\pi_3+0.5\pi_4\\ \pi_4=0.5\pi_3+0.5\pi_4\\ \pi_1+\pi_2+\pi_3+\pi_4=1\end{cases}\quad \text{解得}\begin{cases}\pi_1=\pi_2=\alpha\\ \pi_3=\pi_4=\frac{1}{2}-\alpha\end{cases}(\forall\ 0\le\alpha\le\frac{1}{2})$$

其平稳分布不唯一 (极限分布实际上也不唯一，而且与平稳分布构成双射)

(苏中根 例 4.28)

设 $\mathbf{X}=\{X_n:n\ge 0\}$ 是 Markov 链，状态空间为 $\mathcal{E}=\mathbb{N}_+$

$$\text{转移概率矩阵为 } P=\begin{bmatrix}\frac{1}{2}&\frac{1}{2}&&&&\\ \frac{3}{4}&&\frac{1}{4}&&&\\ &\frac{7}{8}&&\frac{1}{8}&&\\ &&\frac{15}{16}&&\ddots&\\ &&&\ddots&&\end{bmatrix}$$

$$\text{为计算平稳分布，我们求解方程组}\begin{cases}\pi_1=\frac{1}{2}\pi_1+\frac{3}{4}\pi_2\\ \pi_2=\frac{1}{2}\pi_1+\frac{7}{8}\pi_3\\ \pi_3=\frac{1}{4}\pi_2+\frac{15}{16}\pi_4\\ \cdots\end{cases}\text{解得}\begin{cases}\pi_2=\frac{2}{3}\pi_1\\ \pi_3=\frac{4}{3\cdot 7}\pi_1\\ \pi_4=\frac{8}{3\cdot 7\cdot 15}\pi_1\\ \cdots\end{cases}$$

通式为 $\pi_k=\frac{2^{k-1}}{\prod_{i=1}^k(2^i-1)}\pi_1=\frac{2^{k-1}}{1\cdot 3\cdot 7\cdot 15\ldots(2^k-1)}\pi_1$

代入 $\pi_1+\pi_2+\cdots=1$ 就得到 $\pi_1=[\sum\limits_{k=1}^{\infty}\frac{2^{k-1}}{\prod_{i=1}^k(2^i-1)}]^{-1}$

其他项可以根据 π_1 得到.