随机过程 B
第 3 章 Markov 过程 (下)

解扬洋, 殷哲

E-mails: xieyclio@ustc.edu.cn(解扬洋); yinzhe@ustc.edu.cn(殷哲)

本章提纲

1. Markov 链的定义和例子
2. Markov 链状态分类
    - 常返态与瞬过态
3. Markov 链极限定理与平稳分布
    - 极限定理
    - 平稳分布
4. 分支过程
5. 连续时间 Markov 链
    - 连续时间 Markov 链定义
    - 纯生过程

常返态与瞬过态

定义 n 步首达概率

$$f_{ij}^{(n)} = \mathbb{P}(X_n = j, X_k \neq j,\ k = 1, \cdots, n-1 | X_0 = i),\ \forall j$$

- 表示从状态 i 出发，在第 n 步转移时，首次到达状态 j 的概率
- 约定 $f_{ij}^{(0)} = 0,\ \forall j$
  - ▷ 一步都没有走，就无法首次到达，注意到 $j = i$ 时，也有 $f_{ii}^{(0)} = 0$

定义 可达概率

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)},\ \forall j$$

- 表示从 i 出发，最终能够转入状态 j 的概率
- 当 $i \neq j$ 时，$i \to j$ 当且仅当 $f_{ij} > 0$，如何说明？
- 当 $i = j$ 时，$P_{ii}^{(0)} = 1 > 0$，故有 $i \to i$，但 f_{ii} 可能等于 0

常返态与瞬过态 (续)

定义 3.5 常返态与瞬过态

如果 $f_{ii} = 1$, 则我们称状态 i 是**常返的** (recurrent).
一个状态如果不是常返的, 也即 $f_{ii} < 1$, 则称其为**瞬过的** (transient).

- 常返指的是, 由一个状态出发, 经有限步转移, **最终能回到初始状态的概率为 1**
- 瞬过指的是, 最终返回初始状态的概率小于 1
- 运用 n **步转移概率**, 可以判断状态是常返的还是瞬过的

定理 3.2 常返与瞬过的充要条件

状态 i 是常返态的充要条件是

$$\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty.$$

状态 i 是瞬过态的充要条件是

$$\sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty.$$

- 证明如下

定理 3.2 证明

- 先考虑瞬过的充要条件: $f_{ii} < 1 \;\Leftrightarrow\; \sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty$
- 定义随机变量 K 为, 过程返回状态 i 的总次数
  - ▷ 先证 $f_{ii} < 1 \;\Leftrightarrow\; \mathbb{E}(K|X_0 = i) < \infty$, 再证 $\mathbb{E}(K|X_0 = i) = \sum_{n=1}^{\infty} P_{ii}^{(n)}$
  - ▷ 也即瞬过相当于返回初始状态的期望次数有限
- 定义随机变量 N 为, 过程首次返回状态 i 时经过的步数

$$
\begin{aligned}
& \mathbb{P}(K \geq 1|X_0 = i) \\
= & \sum_{n=1}^{\infty} \mathbb{P}(K \geq 1, N = n|X_0 = i) \quad \text{(按照首次返回时的步数拆分事件)} \\
= & \sum_{n=1}^{\infty} \mathbb{P}(N = n|X_0 = i) \quad \text{(等价事件)} \\
= & \sum_{n=1}^{\infty} f_{ii}^{(n)} = f_{ii} \quad (f_{ii}^{(n)} \text{ 定义})
\end{aligned}
$$

常返态与瞬过态 (续)

- 类似地，可以推导出 $\mathbb{P}(K \geq k | X_0 = i)$

$$
\begin{aligned}
& \mathbb{P}(K \geq k | X_0 = i) \\
= & \sum_{n=1}^{\infty} \mathbb{P}(K \geq k, N = n | X_0 = i) \quad (\text{互斥事件拆分}) \\
= & \sum_{n=1}^{\infty} \mathbb{P}(K \geq k | N = n, X_0 = i) \mathbb{P}(N = n | X_0 = i) \quad (\text{条件概率定义}) \\
= & \sum_{n=1}^{\infty} \mathbb{P}(K \geq k | X_n = i, X_{n-1} \neq i, \cdots, X_1 \neq i, X_0 = i) \mathbb{P}(N = n | X_0 = i) \\
= & \sum_{n=1}^{\infty} \color{red}{\mathbb{P}(K \geq k-1 | X_0 = i)} \mathbb{P}(N = n | X_0 = i) \quad (\text{将 } n \text{ 时刻看作新的 } 0 \text{ 时刻}) \\
= & \mathbb{P}(K \geq k-1 | X_0 = i) \sum_{n=1}^{\infty} \mathbb{P}(N = n | X_0 = i) = \mathbb{P}(K \geq k-1 | X_0 = i) \sum_{n=1}^{\infty} f_{ii}^{(n)} \\
= & \mathbb{P}(K \geq k-1 | X_0 = i) f_{ii} \quad (f_{ii} \text{ 定义}) \\
= & \mathbb{P}(K \geq k-2 | X_0 = i) f_{ii}^2 = \cdots = f_{ii}^k \quad (\text{递推})
\end{aligned}
$$

常返态与瞬过态 (续)

已证 $\mathbb{P}(K \geq k|X_0 = i) = f_{ii}^k$

$$
\begin{aligned}
& \mathbb{E}(K|X_0 = i) \\
= & \sum_{k=1}^{\infty} k\mathbb{P}(K = k|X_0 = i) \\
= & \sum_{k=1}^{\infty}\sum_{l=1}^{k} \mathbb{P}(K = k|X_0 = i) \\
= & \sum_{l=1}^{\infty}\sum_{k=l}^{\infty} \mathbb{P}(K = k|X_0 = i) \qquad \text{(交换求和顺序)} \\
= & \sum_{l=1}^{\infty} \mathbb{P}(K \geq l|X_0 = i) \\
= & \sum_{l=1}^{\infty} f_{ii}^l = \frac{f_{ii}}{1 - f_{ii}}
\end{aligned}
$$

- 因此, 当 $f_{ii} < 1$ 时, $\mathbb{E}(K|X_0 = i) < \infty$

已证 $\mathbb{E}(K|X_0=i)<\infty$, 下面证 $\sum_{n=1}^{\infty} P_{ii}^{(n)} = \mathbb{E}(K|X_0=i)$

- 引入示性变量, 记录 n 时刻的状态是否为 i

$$\mathbf{I}_n = \begin{cases} 1, & 若 X_n = i, \\ 0, & 其它情形, \end{cases}$$

- K 的含义是过程返回状态 i 的次数, 因此有

$$K = \sum_{n=1}^{\infty} \mathbf{I}_n$$

- 从而有

$$\begin{aligned} \sum_{n=1}^{\infty} P_{ii}^{(n)} &= \sum_{n=1}^{\infty} \mathbb{P}(X_n = i|X_0 = i) \\ &= \sum_{n=1}^{\infty} \mathbb{E}(\mathbf{I}_n|X_0 = i) \quad (\text{示性变量的期望即是事件发生概率}) \\ &= \mathbb{E}\left(\sum_{n=1}^{\infty} \mathbf{I}_n|X_0 = i\right) \\ &= \mathbb{E}(K|X_0 = i) < \infty \end{aligned}$$

- 上述推导都是双向的，所以有

$$f_{ii} < 1 \Leftrightarrow \sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty$$

- 若状态 i 是常返的，也即 $f_{ii} = 1$
- 类似地有

$$\begin{aligned} & f_{ii} = 1 \\ \Leftrightarrow \quad & \mathbb{E}(K|X_0 = i) = \frac{f_{ii}}{1 - f_{ii}} = \infty \\ \Leftrightarrow \quad & \sum_{n=1}^{\infty} P_{ii}^{(n)} = \mathbb{E}(K|X_0 = i) = \infty \end{aligned}$$

常返态与瞬过态 (续)

推论 3.2

如果状态 i 是常返的, 且 $i \leftrightarrow j$, 则 j 也是常返的.

- 证明见下页

推论 3.2'

如果状态 i 是瞬过的, 且 $i \leftrightarrow j$, 则 j 也是瞬过的.

- 推论 3.2', 可以利用推论 3.2 及反证法证明
  - ▷ 假设 j 不是瞬过的, 则 j 是常返的
  - ▷ 由于 $j \leftrightarrow i$, 利用推论 3.2, 可知 i 也是常返的, 与题设矛盾
  - ▷ 故反证假设不成立, j 是瞬过的

推论 3.2 证明: 与常返态互达的状态也是常返的

- 由 $i \leftrightarrow j$, 可知 $\exists m, n$, 使得 $P_{ji}^{(m)} > 0$ 且 $P_{ij}^{(n)} > 0$
- 利用 Chapman-Kolmogorov 方程, 对任意 s, 有

$$P_{jj}^{(m+s+n)} \geq P_{ji}^{(m)} P_{ii}^{(s)} P_{ij}^{(n)}$$

- 由于状态 i 是常返的, 根据定理 3.2 的充要条件, 有 $\sum_{s=1}^{\infty} P_{ii}^{(s)} = \infty$, 从而

$$\sum_{s=1}^{\infty} P_{jj}^{(s)} \geq \sum_{s=m+n+1}^{\infty} P_{jj}^{(s)} = \sum_{s=1}^{\infty} P_{jj}^{(m+s+n)} \qquad (\text{替换求和变量 } s)$$

$$\geq \sum_{s=1}^{\infty} P_{ji}^{(m)} P_{ii}^{(s)} P_{ij}^{(n)} = P_{ji}^{(m)} P_{ij}^{(n)} \sum_{s=1}^{\infty} P_{ii}^{(s)} = \infty$$

- 再利用定理 3.2 的充要性, 可知 j 也是常返的
- 推论 3.2 表明: 同一等价类中的状态, 要么都是常返的, 要么都是瞬过的

常返态与瞬过态 (续)

例 3.8 随机游动的状态分类

- 考虑整数点上的随机游动: 右移一格的概率为 p, 左移为 $q=1-p$
- 从原点 (状态 0) 出发, 只有转移偶数步才能回到原点, 也即

$$P_{00}^{(2n+1)}=0,\ \ n=1,2,\cdots$$

$$P_{00}^{(2n)}=C_{2n}^{n}p^{n}q^{n},\ \ n=1,2,\cdots$$

- 利用 Stirling 公式进行阶乘的近似

$$\lim_{n\to\infty}\frac{n!}{\sqrt{2\pi n}(\frac{n}{e})^{n}}=1$$

有

$$C_{2n}^{n}=\frac{(2n)!}{n!n!}\approx\frac{\sqrt{2\pi(2n)}(\frac{2n}{e})^{2n}}{\left(\sqrt{2\pi n}(\frac{n}{e})^{n}\right)^{2}}=\frac{2^{2n}}{\sqrt{\pi n}}$$

从而

$$P_{00}^{(2n)}\approx\frac{2^{2n}}{\sqrt{\pi n}}(pq)^{n}=\frac{(4pq)^{n}}{\sqrt{\pi n}}$$

常返态与瞬过态 (续)

$P_{00}^{(2n)} \approx \frac{(4pq)^n}{\sqrt{\pi n}}$

- 已知

$$pq = p(1-p) \leq \frac{1}{4}$$

  且等号仅在 $p = q = \frac{1}{2}$ 时成立

- 因此，当 $p = \frac{1}{2}$ 时，有

$$\begin{aligned} \sum_{n=1}^{\infty} P_{00}^{(n)} &= \sum_{n=1}^{\infty} P_{00}^{(2n)} \\ &\approx \sum_{n=1}^{\infty} \frac{(4pq)^n}{\sqrt{\pi n}} \\ &= \sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}} = \infty \end{aligned}$$

- 故而，状态 0 是常返态
- 又有 $0 \leftrightarrow i$, $\forall i$, 所以对称随机游动的所有状态均是常返的

- 当 $p \neq \frac{1}{2}$ 时, $pq < \frac{1}{4}$

$$\begin{aligned}\sum_{n=1}^{\infty} P_{00}^{(n)} &\approx \sum_{n=1}^{\infty} \frac{(4pq)^n}{\sqrt{\pi n}} \\ &\leq \sum_{n=1}^{\infty} (4pq)^n \\ &= \frac{4pq}{1-4pq} < \infty\end{aligned}$$

- 因此, 非对称随机游动的所有状态均是瞬过的

- 以上是一维随机游动的结果
  - ▷ 二维对称随机游动的所有状态也是常返的
  - ▷ 三维以上对称随机游动的所有状态是瞬过的

常返态与瞬过态 (续)

定义 常返时

对于常返状态 i, 定义 T_i 为过程首次返回状态 i 的时刻, 称为常返时.

- T_i 是一个随机量: $\mathbb{P}(T_i = n | X_0 = i) = f_{ii}^{(n)}$
- i 为常返态时, 有 $\mathbb{P}(T_i < \infty | X_0 = i) = f_{ii} = 1$; 否则 $\mathbb{P}(T_i < \infty | X_0 = i) < 1$, T_i 在实数值域上不满足概率的归一性
- 常返时的期望 (也称平均常返时) 为

$$\mu_i = \mathbb{E}(T_i | X_0 = i) = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

定义 3.6 正常返与零常返

对于常返状态 i, 若 $\mu_i = \infty$, 则称其为零常返的, 若 $\mu_i < \infty$, 则称其为正常返的.

- 如果 Markov 链的状态有限, 则常返状态都是正常返的 (习题 3.15)
- 零常返状态, 只会出现在无穷多状态的 Markov 链中
  - ▷ 如对称随机徘徊, 利用后续的极限定理证明

习题 3:

- 3.1 节作业: 3, 4, 6
- 3.2 节作业: 11, 12, 16

- 随着时间的推移，Markov 链最终会进入怎样的状态呢？
- 本小节将考虑 n 步转移概率矩阵的极限：

$$\lim_{n\to\infty} \mathbf{P}^{(n)}$$

例 3.9

- 考虑只有 $0, 1$ 两个状态的 Markov 链 X_n，其转移概率矩阵为

$$\mathbf{P} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \quad 0 < \alpha, \beta < 1$$

研究 $\mathbf{P}^{(n)}$ 在 $n \to \infty$ 的情况。

例 3.9 (续)

- 根据 C-K 方程，有 $\mathbf{P}^{(n)} = \mathbf{P}^n$
- 可以通过特征分解的方法求 $\mathbf{P}^n$
- 一步概率转移矩阵的特征值为

$$|\lambda\mathbf{I} - \mathbf{P}| = \begin{vmatrix} \lambda - 1 + \alpha & -\alpha \\ -\beta & \lambda - 1 + \beta \end{vmatrix}$$

$$= \lambda^2 + (\alpha + \beta - 2)\lambda - (\alpha + \beta - 1) = 0$$

$$\Rightarrow \quad \lambda_1 = 1, \ \lambda_2 = 1 - \alpha - \beta$$

- 特征向量满足 $(\lambda_i\mathbf{I} - \mathbf{P})\mathbf{e}_i = \mathbf{0}$, 解得

$$\mathbf{e}_1 = [1, \ 1]^T, \ \mathbf{e}_2 = [\alpha, \ -\beta]^T$$

例 3.9 (续)

- 令

$$\mathbf{M} = [\mathbf{e}_1,\ \mathbf{e}_2] = \begin{pmatrix} 1 & \alpha \\ 1 & -\beta \end{pmatrix}$$

$$\mathbf{\Lambda} = \begin{pmatrix} 1 & 0 \\ 0 & 1-\alpha-\beta \end{pmatrix}$$

$$\mathbf{M}^{-1} = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{1}{\alpha+\beta} & -\frac{1}{\alpha+\beta} \end{pmatrix}$$

- 矩阵 $\mathbf{P}$ 可以对角化为

$$\mathbf{P} = \mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1}$$

例 3.9 (续)

- 利用对角化结果，求 n 步转移概率矩阵

$$\begin{aligned}\mathbf{P}^n &= (\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1})^n = \mathbf{M}\mathbf{\Lambda}^n\mathbf{M}^{-1} \\ &= \begin{pmatrix} \frac{\beta+\alpha(1-\alpha-\beta)^n}{\alpha+\beta} & \frac{\alpha-\alpha(1-\alpha-\beta)^n}{\alpha+\beta} \\ \frac{\beta-\beta(1-\alpha-\beta)^n}{\alpha+\beta} & \frac{\alpha+\beta(1-\alpha-\beta)^n}{\alpha+\beta} \end{pmatrix}\end{aligned}$$

- 由 $\alpha, \beta \in (0,1)$ 可知，$|1-\alpha-\beta| < 1$，也即 $\lim_{n\to\infty}(1-\alpha-\beta)^n = 0$
- 因此

$$\lim_{n\to\infty} \mathbf{P}^{(n)} = \lim_{n\to\infty} \mathbf{P}^n = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix}$$

例 3.9 (续)

- 现考虑状态 0 和 1 的平均常返时
  状态 0 的 n 步首达概率为

$$f_{00}^{(1)} = 1-\alpha, \quad f_{00}^{(n)} = \alpha\beta(1-\beta)^{n-2}, \ n = 2, 3, \cdots$$

因此, 状态 0 的平均常返时为

$$\mu_0 = \mathbb{E}(T_0 | X_0 = 0) = \sum_{n=1}^{\infty} n f_{00}^{(n)} = \frac{\alpha+\beta}{\beta}$$

同理得, 状态 1 的平均常返时 $\mu_1 = \frac{\alpha+\beta}{\alpha}$

例 3.9 (续)

- 因此

$$\lim_{n\to\infty} \mathbf{P}^{(n)} = \lim_{n\to\infty} \mathbf{P}^{n} = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix} = \begin{pmatrix} \frac{1}{\mu_0} & \frac{1}{\mu_1} \\ \frac{1}{\mu_0} & \frac{1}{\mu_1} \end{pmatrix}$$

- 记 $\mathbf{Q} = \lim_{n\to\infty} \mathbf{P}^{(n)}$, $\pi_0 = \frac{\beta}{\alpha+\beta}$, $\pi_1 = \frac{\alpha}{\alpha+\beta}$, 观察可得
  - ▷ 过程在进入稳态 (极限) 后处于状态 i 的概率, 总是过程返回该状态平均常返时间的倒数: $\pi_i = \frac{1}{\mu_i}$
  - ▷ 极限转移概率与初始状态无关: 对 $\forall i$, $q_{ij} = \pi_j$, $\{\pi_0, \pi_1\}$ 是本例中 Markov 链的平稳分布

上例中的结果是普遍成立的

定理 3.3 Markov 链基本极限定理

(a) 若状态 i 是瞬过或零常返的，则

$$\lim_{n\to\infty} P_{ii}^{(n)} = 0$$

(b) 若状态 i 是周期为 d 的常返状态，则

$$\lim_{n\to\infty} P_{ii}^{(nd)} = \frac{d}{\mu_i}$$

(c) 若状态 i 是非周期的常返状态 $(d(i) = 1)$，则

$$\lim_{n\to\infty} P_{ii}^{(n)} = \frac{1}{\mu_i}$$

- 我们仅证明 (c)，性质 (a) 和 (b) 的证明过程类似

定理 3.3 证明

- 由习题 3.9 (作业): 从状态 i 出发经过 n 步到达状态 j, 等价于在第 k 步 $(k=1,2,\cdots,n)$ 首达 j, 然后经 $n-k$ 步返回 j

$$P_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} P_{jj}^{(n-k)}$$

特别地，如果 $j=i$，有

$$P_{ii}^{(n)} = \sum_{k=1}^{n} f_{ii}^{(k)} P_{ii}^{(n-k)}$$

- 接下来, 利用形式矩母函数 (离散型随机变量的 $\mathbb{E}e^{tX}$) 来证明
- 记 $\{P_{ii}^{(n)},\ n \geq 0\}$ 和 $\{f_{ii}^{(n)},\ n \geq 0\}$ 形式上的矩母函数为

$$P_i(t) = \sum_{n=0}^{\infty} e^{tn} P_{ii}^{(n)},\quad F_i(t) = \sum_{n=0}^{\infty} e^{tn} f_{ii}^{(n)},\quad t < 0$$

定理 3.3 证明 (续)

- 由 $P_i(t)$定义 有 (注意 $P_{ii}^{(0)} = 1$, $f_{ii}^{(0)} = 0$)

$$
\begin{aligned}
P_i(t) &= P_{ii}^{(0)} + \sum_{n=1}^{\infty} e^{tn} P_{ii}^{(n)} \\
&= 1 + \sum_{n=1}^{\infty} e^{tn} \left(\sum_{k=1}^{n} f_{ii}^{(k)} P_{ii}^{(n-k)} \right) = 1 + \sum_{n=1}^{\infty} \sum_{k=1}^{n} e^{tn} f_{ii}^{(k)} P_{ii}^{(n-k)} \\
&= 1 + \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} e^{tn} f_{ii}^{(k)} P_{ii}^{(n-k)} \qquad \text{(交换求和顺序)} \\
&= 1 + \sum_{k=1}^{\infty} e^{tk} f_{ii}^{(k)} \left(\sum_{n=k}^{\infty} e^{t(n-k)} P_{ii}^{(n-k)} \right) \\
&= 1 + \sum_{k=0}^{\infty} e^{tk} f_{ii}^{(k)} \left(\sum_{n=0}^{\infty} e^{tn} P_{ii}^{(n)} \right) \\
&= 1 + F_i(t) P_i(t)
\end{aligned}
$$

定理 3.3 证明 (续)

- 进一步可得 $(1-e^t)P_i(t)=\frac{1-e^t}{1-F_i(t)}$ 运用 L'Hospital 法则对右边求极限可得

$$\begin{aligned}\lim_{t\to 0^-}\frac{1-e^t}{1-F_i(t)} &= \lim_{t\to 0^-}\frac{\frac{d}{dt}(1-e^t)}{\frac{d}{dt}\left(1-\sum\limits_{n=0}^{\infty}e^{tn}f_{ii}^{(n)}\right)} \\ &= \lim_{t\to 0^-}\frac{-e^t}{-\left(\sum\limits_{n=0}^{\infty}ne^{tn}f_{ii}^{(n)}\right)} \\ &= \frac{1}{\sum\limits_{n=0}^{\infty}nf_{ii}^{(n)}}=\frac{1}{\mu_i}\end{aligned}$$

- 接下来需证 $\lim\limits_{t\to 0^-}(1-e^t)P_i(t)=\lim\limits_{n\to\infty}P_{ii}^{(n)}$

定理 3.3 证明 (续)

$$
\begin{aligned}
\lim_{t\to 0^-}(1-e^t)P_i(t) &= \lim_{t\to 0^-}\frac{P_i(t)}{\frac{1}{1-e^t}} = \lim_{t\to 0^-}\frac{\sum\limits_{n=0}^{\infty} e^{tn}P_{ii}^{(n)}}{\sum\limits_{n=0}^{\infty} e^{tn}} \\
\text{(一致收敛性)} &= \lim_{t\to 0^-}\lim_{k\to\infty}\frac{\sum\limits_{n=0}^{k} e^{tn}P_{ii}^{(n)}}{\sum\limits_{n=0}^{k} e^{tn}} = \lim_{k\to\infty}\lim_{t\to 0^-}\frac{\sum\limits_{n=0}^{k} e^{tn}P_{ii}^{(n)}}{\sum\limits_{n=0}^{k} e^{tn}} \\
\text{(Stolz 定理)} &= \lim_{k\to\infty}\frac{\sum\limits_{n=0}^{k} P_{ii}^{(n)}}{K+1} = \lim_{k\to\infty}P_{ii}^{(k)} = \lim_{n\to\infty}P_{ii}^{(n)}
\end{aligned}
$$

因此，$\lim\limits_{k\to\infty}P_{ii}^{(k)} = \frac{1}{\mu_i}$，定理 3.3 (c) 得证

Stolz 定理

设数列 $\{a_k\}$ 和 $\{b_k\}$ 满足:

(1) $\{b_k\}$ 严格单调递增且 $\lim\limits_{k\to\infty} b_k = +\infty$

(2) $\lim\limits_{k\to\infty} \frac{a_{k+1}-a_k}{b_{k+1}-b_k} = L$

则有 $\lim\limits_{k\to\infty} \frac{a_k}{b_k} = L$

- 也即, 在 k 足够大时, 数列增量的比例就是数列的比例

  令

$$a_k = \sum_{n=0}^{k} P_{ii}^{(n)}, \quad b_k = \sum_{n=0}^{k} 1 = k+1$$

则

$$\lim_{k\to\infty} \frac{\sum\limits_{n=0}^{k} P_{ii}^{(n)}}{K+1} = \lim_{k\to\infty} P_{ii}^{(k)}$$

极限定理 (续)

- 一个正常返 ($f_{ii} = 1$, $\mu_i < \infty$) 非周期 ($d(i) = 1$) 的状态是也称作是**遍历的** (ergodic)
- 定理 3.3 (c) 等价于: 对于遍历状态 i, 有 $\lim_{n\to\infty} P_{ii}^{(n)} = \frac{1}{\mu_i}$, 其中 $\mu_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$ 是常返时的期望

推论 3.3

如果状态 i 是遍历的, 则对所有 $j \to i$, 有

$$\lim_{n\to\infty} P_{ji}^{(n)} = \lim_{n\to\infty} P_{ii}^{(n)} = \frac{1}{\mu_i}$$

证明思路: 利用 $\sum_{n=1}^{\infty} f_{ji}^{(n)} = 1$, $\lim_{n\to\infty} P_{ii}^{(n)} = \frac{1}{\mu_i}$ 及

$$P_{ji}^{(n)} = \sum_{k=1}^{n} f_{ji}^{(k)} P_{ii}^{(n-k)}$$

具体步骤如下

推论 3.3 证明:

对 $1 \le N < n$, 有

$$\sum_{k=1}^{N} f_{ji}^{(k)} P_{ii}^{(n-k)} \le P_{ji}^{(n)} \le \sum_{k=1}^{N} f_{ji}^{(k)} P_{ii}^{(n-k)} + \sum_{k=N+1}^{n} f_{ji}^{(k)}$$

固定 N, 令 $n \to \infty$, 依据 $\lim\limits_{n\to\infty} P_{ii}^{(n)} = \frac{1}{\mu_i}$, 有

$$\frac{1}{\mu_i}\sum_{k=1}^{N} f_{ji}^{(k)} \le \lim_{n\to\infty} P_{ji}^{(n)} \le \frac{1}{\mu_i}\sum_{k=1}^{N} f_{ji}^{(k)} + \sum_{k=N+1}^{\infty} f_{ji}^{(k)}$$

令 $N \to \infty$, 依据 $\sum\limits_{n=1}^{\infty} f_{ji}^{(n)} = 1$, 有

$$\frac{1}{\mu_i} \le \lim_{n\to\infty} P_{ji}^{(n)} \le \frac{1}{\mu_i}$$

推论 3.3 得证

平稳分布

- 通常来说，直接求 $P_{ij}^{(n)}$ 的极限或求 $\mu_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$ 并不简单
- 因此，我们引入平稳分布的定义，来处理 Markov 链的极限

定义 3.7 平稳分布

设 Markov 链的转移概率矩阵为 $\mathbf{P} = (P_{ij})$. 一个概率分布 $\{\pi_j,\ j \geq 0\}$ 如果满足 $\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}$，则称其为 Markov 链的平稳分布.

- 若过程初始时处于平稳分布，也即 $P(X_0 = j) = \pi_j$，则有

$$P(X_1 = j) = \sum_{i=0}^{\infty} P(X_1 = j | X_0 = i) P(X_0 = i) = \sum_{i=0}^{\infty} \pi_i P_{ij} = \pi_j$$

- 运用归纳法，得 $P(X_n = j) = \pi_j,\ \forall n \geq 1$
- 也即所有 X_n 是同分布的

平稳分布 (续)

- 由 Markov 性及 X_n 与 X_0 同分布, 有

$$\begin{aligned}
& P(X_n = i_0, X_{n+1} = i_1, \cdots, X_{n+k} = i_k) \\
= \; & P(X_n = i_0) P(X_{n+1} = i_1 | X_n = i_0) P(X_{n+2} = i_2 | X_{n+1} = i_1, X_n = i_0) \\
& \cdots P(X_{n+k} = i_k | X_{n+k-1} = i_{n+k-1}, \cdots, X_n = i_0) \\
= \; & P(X_n = i_0) P_{i_0,i_1} P_{i_1,i_2} \cdots P_{i_{k-1},i_k} \\
= \; & P(X_0 = i_0) P_{i_0,i_1} P_{i_1,i_2} \cdots P_{i_{k-1},i_k} \\
= \; & P(X_0 = i_0, X_1 = i_1, \cdots, X_k = i_k)
\end{aligned}$$

- 这意味着, 对任意的 k, $X_n, X_{n+1}, \cdots, X_{n+k}$ 的联合分布, 不依赖于 n
- 过程的有限维联合分布不依赖于分布的初始时间 n, 只依赖于分布的维度 k
- 也即随机过程是平稳的, 平稳分布 π 由此得名

平稳分布 (续)

定理 3.4 平稳分布与极限分布

(a) 若一个不可约 (所有状态互达) 的 Markov 链中，所有状态都是遍历 (非周期正常返) 的，则对任意状态 i, j，

$$\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j,$$

其中 $\pi = \{\pi_j,\ j \geq 0\}$ 是一个严格为正的平稳分布，也即

$$\sum_j \pi_j = 1, \quad \pi_j > 0, \quad \sum_i \pi_i P_{ij} = \pi_j.$$

(b) 反之，若一个不可约 Markov 链只存在一个平稳分布，且所有状态都是遍历的，则该平稳分布就是此马氏链的极限分布.

- (a) 不可约遍历 Markov 链的极限分布是严格为正的平稳分布
- (b) 不可约遍历 Markov 链若只有一个平稳分布，则其必是极限分布

定理 3.4 证明

- 在不可约 Markov 链中，由定理 3.3 (c) 及推论 3.3，对遍历状态 j，有

$$\lim_{n\to\infty} P_{ij}^{(n)} = \frac{1}{\mu_j}, \quad \forall i$$

- 状态 j 是遍历的 (非周期正常返)，故有 $\mu_j < \infty$，因而有 $\frac{1}{\mu_j} > 0$
- 记 $\pi_j = \frac{1}{\mu_j}$，取 $0 \le M < n$，由 C-K 方程有

$$P_{ij}^{(n)} = \sum_{k=0}^{\infty} P_{ik}^{(n-1)} P_{kj} \ge \sum_{k=0}^{M} P_{ik}^{(n-1)} P_{kj}$$

先固定 M，令 $n \to \infty$，再令 $M \to \infty$ 有

$$\pi_j \ge \sum_{k=0}^{\infty} \pi_k P_{kj}$$

定理 3.4 证明 (续)

- 下面用反证法证明 $\pi_j = \sum_{k=0}^{\infty} \pi_k P_{kj}$ 假设对某个 j, $\pi_j > \sum_{k=0}^{\infty} \pi_k P_{kj}$ 成立, 将所有不等式求和, 有

$$
\begin{aligned}
\sum_{j=0}^{\infty} \pi_j &> \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \pi_k P_{kj} \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \pi_k P_{kj} = \sum_{k=0}^{\infty} \pi_k \sum_{j=0}^{\infty} P_{kj} \quad (\text{交换求和号}) \\
&= \sum_{k=0}^{\infty} \pi_k \cdot 1 = \sum_{k=0}^{\infty} \pi_k \quad (\text{一步转移矩阵每行和为 } 1)
\end{aligned}
$$

矛盾, 假设不成立, 故

$$
\pi_j = \sum_{k=0}^{\infty} \pi_k P_{kj}
$$

平稳分布 (续)

定理 3.4　证明 (续)

- 不断将 π_j 代入等式右边, 得 $\{\pi_j\}$ 在 n 步转移下也是平稳的

$$\begin{aligned}\pi_j &= \sum_{k=0}^{\infty}\pi_k P_{kj} = \sum_{k=0}^{\infty}\left(\sum_{i=0}^{\infty}\pi_i P_{ik}\right)P_{kj} \\ &= \sum_{i=0}^{\infty}\pi_i\sum_{k=0}^{\infty}P_{ik}P_{kj} \\ &= \sum_{i=0}^{\infty}\pi_i P_{ij}^{(2)} = \cdots = \sum_{i=0}^{\infty}\pi_i P_{ij}^{(n)}\end{aligned}$$

- 接下来证明 $\sum\limits_{j=0}^{\infty}\pi_j = 1$

定理 3.4 证明 (续)

取 $0 \le M < n$, 利用 $\sum_{j=0}^{\infty} P_{ij}^{(n)} = 1$, 有

$$\sum_{j=0}^{M} P_{ij}^{(n)} \le \sum_{j=0}^{\infty} P_{ij}^{(n)} = 1$$

固定 M, 令 $n \to \infty$, 有

$$\sum_{j=0}^{M} \pi_j \le 1$$

令 $M \to \infty$, 有

$$\sum_{j=0}^{\infty} \pi_j \le 1$$

定理 3.4 证明 (续)

取 $0 \le M < n$, 由 $\pi_j = \sum_{k=0}^{\infty} \pi_k P_{kj}^{(n)}$, 可得

$$\pi_j \le \sum_{k=0}^{M} \pi_k P_{kj}^{(n)} + \sum_{k=M+1}^{\infty} \pi_k$$

令 $n \to \infty$, 有

$$\pi_j \le \sum_{k=0}^{M} \pi_k \pi_j + \sum_{k=M+1}^{\infty} \pi_k$$

令 $M \to \infty$, 有

$$\pi_j \le \pi_j \sum_{k=0}^{\infty} \pi_k$$

由于 $\pi_j = \frac{1}{\mu_j} > 0$ (正常返性质), 故有 $\sum_{k=0}^{\infty} \pi_k \ge 1$

平稳分布 (续)

- 综上，证得
$$\sum_j \pi_j = 1, \quad \pi_j > 0, \quad \sum_i \pi_i P_{ij} = \pi_j$$
- 因而 $\{\pi_j,\ j \geq 0\}$ 是一个平稳分布

- 定理 3.4 (b)，若所有状态都是遍历的不可约 Markov 链只存在一个平稳分布，则该平稳分布就是极限分布 的证明：
  - ▷ 利用 (a) 的结论，不可约遍历 Markov 链的极限分布一定存在，且必定是平稳分布
  - ▷ 由于 (b) 的题设中只存在一个平稳分布，故该平稳分布必是极限分布

例 3.9 再讨论

- 两状态 Markov 链, 1 步转移概率矩阵为

$$\mathbf{P} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \quad 0 < \alpha, \beta < 1$$

现利用定理 3.4, 通过平稳分布, 求解 $\lim_{n\to\infty} \mathbf{P}^{(n)}$

- 由定理 3.4, $\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j$, 且 $\{\pi_0, \ \pi_1\}$ 满足

$$\begin{aligned} \pi_0 + \pi_1 &= 1 \\ \pi_0 P_{00} + \pi_1 P_{10} &= \pi_0 \\ \pi_0 P_{01} + \pi_1 P_{11} &= \pi_1 \end{aligned}$$

- 只有两个方程是线性无关的, 代入 P_{ij} 的值, 得

$$\begin{aligned} \pi_0 + \pi_1 &= 1 \\ (1-\alpha)\pi_0 + \beta\pi_1 &= \pi_0 \end{aligned}$$

- 解得 $\pi_0 = \frac{\beta}{\alpha+\beta}$, $\pi_1 = \frac{\alpha}{\alpha+\beta}$, 与之前的计算结果一致

平稳分布 (续)

- 对于平稳分布 $\{\pi_j\}$, 有两种常见的看法
  - ▷ 一种看法是, 作为 $P_{ij}^{(n)}$ 的极限分布, 告诉我们无论初始状态 i 为何, 经过足够长时间后, 过程处于状态 j 的概率就是 π_j
  - ▷ 另一种看法是, π_j 代表了在长期运行过程中, 访问状态 j 的频率 (访问状态 j 次数占总访问次数的比例)

定义

$$I_n = \begin{cases} 1, & \text{若} X_n = j, \\ 0, & \text{其它情形} \end{cases}$$

则 m 步转移中, 访问 j 的次数所占比例为 $\frac{1}{m}\sum\limits_{n=0}^{m-1} I_n$ 若初始状态为 i, 该比例的条件期望为

$$\mathbb{E}\left\{\frac{1}{m}\sum_{n=0}^{m-1} I_n | X_0 = i\right\} = \frac{1}{m}\sum_{n=0}^{m-1} P(X_n = j | X_0 = i) = \frac{1}{m}\sum_{n=0}^{m-1} P_{ij}^{(n)}$$

当 $m \to \infty$, 利用 Stolz 定理, 有

$$\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m-1} P_{ij}^{(n)} = \lim_{m\to\infty} P_{ij}^{(m)} = \pi_j = \frac{1}{\mu_j}$$

第 3 章 (下) 习题一

- 3.3 节作业: 17, 18, 19

分支过程

- 分支过程 (Branching process) 是 Markov 链的一个重要应用, 可以描述生物遗传、原子核的连锁反应等

- 定义分支过程 $\{X_n, \ n \geq 0\}$
  - X_n 为第 n 代个体数, 第 $n+1$ 代由第 n 代的后代组成
  - X_0 为第 0 代的先祖数目, 一般假定为 1
  - 每个个体的繁衍能力相同且相互独立, 均服从分布 $\{p_k\}, k = 0, 1, 2 \cdots$
  - 记 Z_i 为每一代中第 i 个个体所繁衍的后代数, 即 $P(Z_i = k) = p_k$, 则

$$X_{n+1} = \sum_{i=1}^{X_n} Z_i$$

  - 记 $\mathbb{E}Z_i = \mu$, $\mathrm{Var}Z_i = \sigma^2$

分支过程 (续)

- 分支过程作为一个 Markov 链的转移概率

$$
\begin{aligned}
P_{ij} &= P(X_{n+1} = j | X_n = i) \\
&= P\left(\sum_{k=1}^{X_n} Z_k = j | X_n = i\right) = P\left(\sum_{k=1}^{i} Z_k = j\right)
\end{aligned}
$$

- 分支过程的数字特征

  利用第 1 章例 1.12 中独立随机变量和的结果, 可得

$$\mathbb{E}X_{n+1} = \mathbb{E}X_n\mathbb{E}Z_i = \mu\mathbb{E}X_n$$

$$\mathrm{Var}X_{n+1} = \mathbb{E}X_n\mathrm{Var}Z_i + \mathrm{Var}X_n(\mathbb{E}Z_i)^2 = \sigma^2\mathbb{E}X_n + \mu^2\mathrm{Var}X_n$$

  利用递推关系及 $\mathbb{E}X_1 = \mathbb{E}Z_1 = \mu$, 得到

$$\mathbb{E}X_{n+1} = \mu\mathbb{E}X_n = \mu^2\mathbb{E}X_{n-1} = \cdots = \mu^n\mathbb{E}X_1 = \mu^{n+1}$$

$$\begin{aligned}
\mathrm{Var}X_{n+1} &= \sigma^2\mu^n + \mu^2\mathrm{Var}X_n = \sigma^2\mu^n + \mu^2[\sigma^2\mu^{n-1} + \mu^2\mathrm{Var}X_{n-1}] \\
&= \sigma^2(\mu^n + \mu^{n+1}) + \mu^4[\sigma^2\mu^{n-2} + \mu^2\mathrm{Var}X_{n-2}] \\
&= \cdots = \sigma^2\mu^n[1 + \mu + \cdots + \mu^n] \\
&= \begin{cases} \sigma^2\mu^n\frac{1-\mu^{n+1}}{1-\mu}, & \mu \neq 1, \\ (n+1)\sigma^2, & \mu = 1. \end{cases}
\end{aligned}$$

分支过程 (续)

- 群体消亡概率
  从数字特征看出，个体繁衍后代的平均数量 μ 对群体繁衍至关重要
  - 当 $\mu < 1$ 时，随时间推移，$\mathbb{E}X_n$ 和 $\mathrm{Var}X_n$ 都趋于 0
    利用 Chebyshev 不等式，对于任意 ε，有

$$P(|X_n - \mathbb{E}X_n| > \varepsilon) \leq \frac{\mathrm{Var}X_n}{\varepsilon^2}$$

    群体数量最终依概率收敛到 0，也即群体必然消亡
  - 当 $\mu \geq 1$ 时，群体仍会消亡吗？
    群体消亡即在某个时刻，过程进入了吸收态 0
    过程最终进入吸收状态的概率即为消亡概率

- 利用概率生成函数研究消亡概率

定理 3.5

对分支过程 X_n, 若 $p_0 > 0$ 且 $p_0 + p_1 < 1$, 则有

(a) 群体消亡概率 π, 是方程 $\phi(s) = s$ 的最小正解, 其中 $\phi(s) = \sum_{j=0}^{\infty} p_j s^j$

(b) $\pi = 1$ 当且仅当 $\mu \leq 1$

- 证明 (a):
记第一代个体数量 X_1 的概率生成函数为 $\phi(s)$, 因 $X_1 = Z_1$, 有

$$\phi(s) = \mathbb{E}s^{X_1} = \mathbb{E}s^{Z_1} = \sum_{k=0}^{\infty} p_k s^k$$

分支过程 (续)

- 记 π_n 为由单个先祖开始的家族, 在第 n 代前**消亡的概率**
- 当 $n \to \infty$ 时, π_n **的极限**即为整个群体的消亡概率, 已知

$$\pi_n = P(X_n = 0) = \sum_{k=0}^{\infty} s^k P(X_n = k)\Big|_{s=0} = \phi_n(0)$$

其中 $\phi_n(s)$ 为 X_n 的生成函数, 注意 $\phi_1(s) = \phi(s)$

- 利用 Z_i 独立同分布及条件概率**两步走**, 有

$$\begin{aligned}
\phi_{n+1}(s) &= \mathbb{E}s^{X_{n+1}} = \mathbb{E}s^{\sum_{i=1}^{X_n} Z_i} \\
&= \mathbb{E}[\mathbb{E}(s^{\sum_{i=1}^{X_n} Z_i} | X_n)] \qquad (\text{对 } X_n \text{ 取条件}) \\
&= \mathbb{E}[\prod_{i=1}^{X_n} \mathbb{E}s^{Z_i} | X_n] = \mathbb{E}[\prod_{i=1}^{X_n} \phi(s) | X_n] \\
&= \mathbb{E}[\phi(s)]^{X_n} = \phi_n\left(\phi(s)\right)
\end{aligned}$$

分支过程 (续)

- 当 $p_0 = 1$ 时，家族不能发端；当 $p_0 = 0$ 时，家族不会消亡
  因此，假定 $0 < p_0 < 1$，来考察 π_n 极限的存在性
- 对 $s \in (0,1)$，利用函数项级数的一致收敛性，可交换无穷求和运算与求导运算，有

$$\phi'(s) = \frac{d}{ds}\left(\sum_{k=0}^{\infty} p_k s^k\right) = \sum_{k=1}^{\infty} p_k k s^{k-1} > 0 \quad (\text{因存在} k \geq 1 \text{满足} p_k > 0)$$

  因此，$\phi(s)$ 是 $(0,1)$ 上的非负单调递增函数，由 $\phi_{n+1}(s) = \phi_n(\phi(s))$ 可推出 $\phi_n(s)$ 也是 $(0,1)$ 上的非负单增函数
- 又由于 $\phi(0) = p_0 > 0$，故有

$$\pi_{n+1} = \phi_{n+1}(0) = \phi_n(\phi(0)) = \phi_n(p_0) > \phi_n(0) = \pi_n, \quad \forall n$$

- π_n 随 n 单增且有界，故极限存在，记

$$\lim_{n\to\infty} \pi_n = \pi$$

- π 就是群体消亡的概率

分支过程 (续)

- 由 $\phi_{n+1}(s) = \phi_n(\phi(s))$ 可推出 $\phi_{n+1}(s) = \phi(\phi_n(s))$
- 由 $\phi_{n+1}(s) = \phi(\phi_n(s))$ 及 $\phi_n(0) = \pi_n$，可知

$$\pi_{n+1} = \phi_{n+1}(0) = \phi(\phi_n(0)) = \phi(\pi_n)$$

- 令 $n \to \infty$，可得

$$\pi = \phi(\pi)$$

- 因此，可以通过求解方程 $\phi(s) = s$，得到消亡概率 π
- 还需证明其是最小正解，即假定 $\tilde{\pi}$ 是满足方程 $\phi(s) = s$ 的任意解，有 $\tilde{\pi} \geq \pi$ 成立
- 利用数学归纳法证 $\tilde{\pi} \geq \pi_n = P(X_n = 0)$，验证 $n = 1$ 成立，有

$$\tilde{\pi} = \phi(\tilde{\pi}) = \sum_{k=0}^{\infty} \tilde{\pi}^k p_k \geq \tilde{\pi}^0 p_0 = p_0 = P(X_1 = 0)$$

分支过程 (续)

- 假设 $\tilde{\pi} \geq P(X_n = 0)$ 成立, 现验证 $\tilde{\pi} \geq P(X_{n+1} = 0)$ 成立, 有

$$
\begin{aligned}
P(X_{n+1}=0) &= \sum_{j=0}^{\infty} P(X_{n+1}=0|X_1=j)P(X_1=j) \\
&= \sum_{j=0}^{\infty} [P(X_n=0|X_0=1)]^j p_j \quad (j\text{个个体独立繁衍}) \\
&\leq \sum_{j=0}^{\infty} \tilde{\pi}^j p_j \leq \sum_{j=0}^{\infty} \tilde{\pi} p_j = \tilde{\pi}
\end{aligned}
$$

- 取 $n \to \infty$, 有

$$
\tilde{\pi} \geq \lim_{n\to\infty} P(X_n = 0) = \pi
$$

- 从而, π 是方程 $\phi(s) = s$ 的最小正解, 定理 3.5 (a) 得证

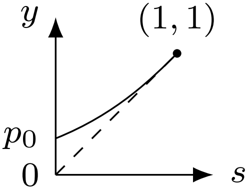
分支过程 (续)

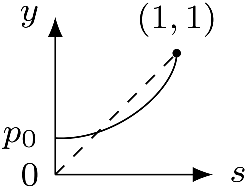
- 证明 (b): $\pi = 1$ 当且仅当 $\mu \le 1$
  相当于考察 $y = \phi(s)$ 和 $y = s$ 在 $(1,1)$ 处是否相交, 及是否最小正解
- 注意到 $\phi(s)$ 在 $(0,1)$ 上是一个严格凸函数, 也即

$$\phi''(s) = \frac{d^2}{ds^2}\left(\sum_{k=0}^{\infty} p_k s^k\right) = \sum_{k=2}^{\infty} p_k k(k-1)s^{k-2} > 0 \quad (\text{由} p_0 + p_1 < 1)$$

- 此外, 有 $\phi(0) = p_0$, $\phi(1) = \sum_{j=0}^{\infty} p_j = 1$
- 故严格凸函数 $y = \phi(s)$ 与直线 $y = s$ 必相交于 $(1,1)$ 点

- 严格凸函数和直线最多只有两个交点
- 故两者在 $s \in (0,1)$ 上至多只有一个交点, 取决于 $\phi'(1)$
  - ▷ 当 $\phi'(1) \le 1$ 时 (左图), 两者在 $s \in (0,1)$ 上不相交
  - ▷ 当 $\phi'(1) > 1$ 时 (右图), 两者在 $s \in (0,1)$ 上一定相交一次





- 注意到

$$\phi'(1) = \frac{d}{ds}\left(\sum_{k=0}^{\infty} p_k s^k\right) = \sum_{k=1}^{\infty} k p_k = \mathbb{E}Z_i = \mu$$

- 因此, $\mu \le 1$ 时, $\phi(s) = s$ 的最小解为 1, 也即 $\pi = 1$ (左图), 种群必然消亡; $\mu > 1$ 时, $0 < \pi < 1$ (右图), (b) 得证.
- 未消亡时, 种群数量会增长到无穷 (由于 $\mathbb{E}X_n = \mu^n$)

例 3.10 Poisson 分布下的分支过程

- 假定生物群体中的每个个体繁衍下一代的数量服从参数为 λ 的 Poisson 分布，也即

$$p_k = P(Z_i = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \ k = 0, 1, 2, \cdots$$

- 由于 i 个独立的参数为 λ 的 Poisson 随机变量和，是参数为 λi 的 Poisson 随机变量，故分支过程 X_n 的转移概率为

$$P_{ij} = P(X_{n+1} = j | X_n = i) = \frac{(\lambda i)^j}{j!}e^{-\lambda i}$$

- X_1 的生成函数为

$$\phi(s) = \sum_{k=0}^{\infty} p_k s^k = e^{\lambda(s-1)} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} e^{-\lambda s} = e^{\lambda(s-1)}$$

- 从而 $\phi'(1) = \lambda e^{\lambda(s-1)}|_{s=1} = \lambda$，或者直接利用期望 $\phi'(1) = \mathbb{E}Z_1 = \lambda$
  - ▷ 由定理 3.5，当 $\lambda \le 1$ 时，消亡概率 $\pi = 1$，群体必然消亡
  - ▷ 当 $\lambda > 1$ 时，消亡概率 $\pi \in (p_0, \ 1)$，例如当 $\lambda = 2$ 时，$\pi \approx 0.2$

第 3 章 (下) 习题二

- 3.4 节作业: 20, 21, 22

- 当时间指标参数从离散的 $n = 0, 1, 2, \cdots$，变为连续的实数 $t \geq 0$ 时，便是连续时间 Markov 链

定义 3.8 连续时间 Markov 链

若对所有非负实数 $s, t \geq 0$ 和任意非负整数状态 i, j 及状态函数 $x(u)$, $0 \leq u \leq s$，随机过程 $X(t)$, $t \geq 0$ 满足

$$
\begin{aligned}
& P(X(t+s) = j | X(s) = i,\ X(u) = x(u),\ 0 \leq u < s) \\
= \quad & P(X(t+s) = j | X(s) = i),
\end{aligned}
$$

则称过程为连续时间 Markov 链.

- 定义所表示的，即是连续时间下的 Markov 性
- 进一步，若 $P(X(t+s) = j | X(s) = i)$ 与 s 无关，则称过程为有平稳转移概率的连续时间 Markov 链
  - ▷ 平稳转移概率可简记为 $P_{ij}(t) = P(X(t+s) = j | X(s) = i)$
  - ▷ 本节后续讨论默认为平稳 Markov 链

命题 3.5

连续时间 Markov 链的**转移概率函数** $P_{ij}(t)$ 和**初始分布** $p_i = P(X(0) = i)$, 完全确定了过程的所有有限维联合分布

证明

- 利用条件概率, 对任何 $0 = t_0 < t_1 < \cdots < t_n$ 和状态 i_k, $0 \le k \le n$, n 为任意非负整数, 有

$$
\begin{aligned}
& P(X(t_n) = i_n, X(t_{n-1}) = i_{n-1}, \cdots, X(t_0) = i_0) \\
= \; & P(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}, \cdots, X(t_0) = i_0) \\
& \cdot P(X(t_{n-1}) = i_{n-1} | X(t_{n-2}) = i_{n-2}, \cdots, X(t_0) = i_0) \\
& \cdots P(X(t_1) = i_1 | X(t_0) = i_0) \cdot P(X(t_0) = i_0) \\
= \; & P_{i_{n-1}, i_n}(t_n - t_{n-1}) P_{i_{n-2}, i_{n-1}}(t_{n-1} - t_{n-2}) \cdots P_{i_0, i_1}(t_1) p_{i_0}
\end{aligned}
$$

- 得证

- 那么，一组对 $t \geq 0$ 定义的函数 $P_{ij}(t)$，在满足何种条件时，才能充当 Markov 链的转移概率呢？
- 首先，应满足概率的非负性和正则性：

$$P_{ij}(t) \geq 0, \quad \sum_j P_{ij}(t) = 1$$

- 其次，还应满足连续形式的 Chapman-Kolomogorov 方程

$$\begin{aligned} P_{ij}(t+s) &= \sum_k P(X(t+s)=j, X(s)=k|X(0)=i) \\ &= \sum_k P(X(t+s)=j|X(s)=k, X(0)=i)P(X(s)=k|X(0)=i) \\ &= \sum_k P_{ik}(s)P_{kj}(t) \end{aligned}$$

- 此外，通常还假设过程在有限时间内不会转移无穷多次，也即

$$\lim_{t\downarrow 0} P_{ii}(t) = 1$$

- 在连续时间 Markov 链中，我们不仅关心过程会转移到哪个状态，转移的机会有多大，还关心过程在某个状态逗留的时间
- 记 τ_i 为过程在转移之前，逗留于状态 i 的概率

$$\{\tau_i > t\} = \{X(s) = i,\ 0 < s \le t | X(0) = i\}$$

- 由 Markov 性和平稳性可知，逗留时间的分布具有无记忆性，也即

$$\begin{aligned}
& P(\tau_i > s + t | \tau_i > s) \\
= & \frac{P(\tau_i > s + t)}{P(\tau_i > s)} \qquad \text{(条件概率定义)} \\
= & \frac{P(X(u) = i,\ 0 < u \le s + t | X(0) = i)}{P(X(u) = i,\ 0 < u \le s | X(0) = i)} \qquad (\tau_i \text{ 定义}) \\
= & \frac{P(X(u) = i,\ 0 < u \le s + t,\ X(0) = i)}{P(X(u) = i,\ 0 < u \le s,\ X(0) = i)} \\
= & P(X(u) = i,\ s < u \le s + t | X(u) = i,\ 0 \le u \le s) \\
= & P(X(u) = i,\ s < u \le s + t | X(s) = i) \qquad \text{(Markov 性)} \\
= & P(X(u) = i,\ 0 < u \le t | X(0) = i) \qquad \text{(平稳性)} \\
= & P(\tau_i > t)
\end{aligned}$$

- 由无记忆性 $P(\tau_i > s+t | \tau_i > s) = P(\tau_i > t)$ 可推出

$$P(\tau_i > s+t) = P(\tau_i > s)P(\tau_i > t)$$

- 记 τ_i 的 CDF 为 $F_i(t)$, pdf 为 $f_i(t)$, $\bar{F}_i(t) = 1 - F_i(t) = P(\tau_i > t)$, 有

$$\bar{F}_i(s+t) = \bar{F}_i(s)\bar{F}_i(t)$$

- 从而有

$$\begin{aligned} \frac{d}{dt}\bar{F}_i(t) &= \lim_{s\to 0}\frac{\bar{F}_i(s+t) - \bar{F}_i(t)}{s} = \lim_{s\to 0}\frac{\bar{F}_i(s)\bar{F}_i(t) - \bar{F}_i(t)}{s} \\ &= \lim_{s\to 0}\frac{\bar{F}_i(s) - 1}{s}\bar{F}_i(t) = -f_i(0)\bar{F}_i(t) \end{aligned}$$

- 求解一阶常微分方程, 得 $\bar{F}_i(t) = e^{-f_i(0)t}$
- 也即, τ_i 服从指数分布, 记其参数为 $\nu_i = f_i(0)$

- 从而，可给出基于逗留时间的连续时间 (平稳)Markov 链等价定义

连续时间 Markov 链等价定义

连续时间 Markov 链是具有以下两条性质的随机过程:

(a) 过程从某一状态 i 转移到其它状态前，在 i 逗留的时间，服从参数为 ν_i 的指数分布

(b) 过程离开状态 i，在下一次转移时转入状态 j 的概率为 P_{ij}，这里 $\sum_{j\neq i} P_{ij} = 1$

- 注意到 ν_i 依赖于状态 i，也即不同状态下的逗留时间分布可以不一样，但都属于指数分布类型
- ν_i 代表过程离开状态 i 的速率
- $\nu_i = 0$ 时，逗留时间 τ_i 的 CDF 为 $P(\tau_i \leq t) = 1 - e^{-\nu_i t} = 0, \ \forall t$
  - 离开状态 i 的速率是 0，状态 i 是吸收态
- P_{ij} 代表过程发生转移时 (条件)，转入状态 j 的概率，$\sum_{j\neq i} P_{ij} = 1$
  - 注意到其与 $P_{ij}(t) = P(X(t+s) = j|X(s) = i)$ 不同，$\sum_j P_{ij}(t) = 1$
- 记 $q_{ij} = \nu_i P_{ij}$，$j \neq i$，代表了过程从状态 i 到 j 的转移速率

例 3.11

- 考虑 Poisson 过程 $\{X(t),\ t \geq 0\}$
- 在时间 $[t, t+\tau]$ 内, 到达的粒子数服从以 $\lambda\tau$ 为参数的 Poisson 分布
- 由增量独立性, 对任意 $0 \leq t_1 < t_2 < \cdots < t_n < t$, 有

$$\begin{aligned}
& P(X(t+\tau) = k+j | X(t) = j, X(t_n) = j_n, \cdots, X(t_1) = j_1) \\
= \quad & P(X(t+\tau) - X(t) = k | X(t) = j, X(t_n) = j_n, \cdots, X(t_1) = j_1) \\
= \quad & P(X(t+\tau) - X(t) = k) \\
= \quad & P(X(t+\tau) = k+j | X(t) = j)
\end{aligned}$$

- Markov 性得证, 故 $X(t)$ 是有平稳转移概率的连续时间 Markov 链
- 其转移概率为, t 时段内到达 $j-i$ 个粒子的概率, 也即

$$P_{ij}(t) = P(X(t) = j | X(0) = i) = \begin{cases} 0, & j < i, \\ \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t}, & j \geq i. \end{cases}$$

- 根据第 2 章对 Poisson 过程的讨论
  - ▷ 当 $j \neq i+1$ 时, $P_{ij} = 0$, 因此, $q_{ij} = 0$
  - ▷ 当 $j = i+1$ 时, $P_{i,i+1} = 1$, $v_i = \lambda$, 因此, $q_{ij} = \lambda$

纯生过程

- 考虑一个种群在保证良好环境、充足食物、无死亡、无外迁的理想条件下的生长模型
- 也可用来描述传染病病例增长
- 种群增加 (新生) 一个个体的概率
  - 群体当前大小 (及过程所处状态) 是相关的
  - 与经过时间长短成比例
- 记 $X(t)$ 为 $(0,t]$ 时间段内新生个体数

定义 纯生过程

当随机过程 $X(t)$ 满足以下四条假设时, 称其为纯生过程.

(i) $X(0)=0$

(ii) $P(X(t+h)-X(t)=1|X(t)=k)=\lambda_k h+o(h)$

(iii) $P(X(t+h)-X(t)=0|X(t)=k)=1-\lambda_k h+o(h)$

(iv) $P(X(t+h)-X(t)<0|X(t)=k)=0$

- 纯生过程是 Poisson 过程的自然推广, 过程关于时间是齐次的, 区别在于到达率 λ_k 依赖于系统当前状态

纯生过程 (续)

- 将从状态 0 开始，经过 t 时间，到达状态 n 的转移概率简记为

$$P_n(t) = P_{0n}(t) = P(X(t) = n | X(0) = 0)$$

- 由全概率公式 (两步走)，得

$$\begin{aligned} P_n(t+h) &= \sum_{k=0}^{\infty} P_k(t) P(X(t+h) = n | X(t) = k) \\ &= \sum_{k=0}^{\infty} P_k(t) P(X(t+h) - X(t) = n-k | X(t) = k) \end{aligned}$$

- 基于上式，利用假设 (ii), (iii), (iv)，对 $n \geq 1$ 有

$$P_n(t+h) = \underbrace{P_n(t)[1 - \lambda_n h + o(h)]}_{t \text{ 时刻为状态 } n} + \underbrace{P_{n-1}(t)[\lambda_{n-1} h + o(h)]}_{t \text{ 时刻为状态 } n-1} + \underbrace{o(h)}_{t \text{ 时刻为其它状态}},$$

- 当 $n = 0$ 时，有

$$P_0(t+h) = \underbrace{P_0(t)[1 - \lambda_0 h + o(h)]}_{t \text{ 时刻为状态 } n} + \underbrace{o(h)}_{t \text{ 时刻为其它状态}}$$

- 移项整理后, 两边同除 h 并取 $h \to 0$, 得

$$
\begin{aligned}
P_0'(t) &= -\lambda_0 P_0(t) \\
P_n'(t) &= -\lambda_n P_n(t) + \lambda_{n-1} P_{n-1}(t), \ n \geq 1
\end{aligned}
$$

- 由第一行的一阶齐次常微分方程, 可解出

$$
P_0(t) = e^{-\lambda_0 t}, \ t \geq 0
$$

- 将该解代入第二行的一阶非齐次常微分方程, 运用归纳法, 可以解出

$$
P_n(t) = \left(\prod_{i=0}^{n-1} \lambda_i \right) \left[\sum_{i=0}^{n} c_{i,n} e^{-\lambda_i t} \right],
$$

其中 $c_{i,n} = \left[\prod_{0 \leq k \leq n, k \neq i} (\lambda_k - \lambda_i) \right]^{-1}$

纯生过程 (续)

- 记纯生过程在某一状态 i 逗留的时间为 T_i
  - ▷ 也即，两个相邻事件发生的间隔时间
- 用 $W_k = \sum_{i=0}^{k-1} T_i$ 表示第 k 个生命的诞生时刻，则有

$$P_n(t) = P(X(t) = n | X(0) = 0) = P(W_n \le t < W_{n+1})$$

- 回忆 Poisson 过程，事件发生的间隔时间服从参数为 λ 的指数分布
  - ▷ 可类比证明，纯生过程事件发生的间隔 $T_i \sim \text{Exp}(\lambda_i)$，且 T_i 相互独立
- $\mathbb{E}T_i = \frac{1}{\lambda_i}$ 即表示过程在状态 i 的平均逗留时间
  - ▷ 纯生过程的状态是不断增大的，一旦离开状态 i，就不再返回
  - ▷ 因此，过程增长到状态 i 的平均时间为

$$\mathbb{E}W_i = \sum_{n=0}^{i-1} \mathbb{E}T_n = \sum_{n=0}^{i-1} \frac{1}{\lambda_n}$$

纯生过程 (续)

- 纯生过程满足 $\sum_{n=0}^{\infty}\frac{1}{\lambda_n}=\infty$, 分析过程如下
- 概率的正则性 (t 时刻过程必然处于某一状态) 要求

$$\sum_{n=0}^{\infty}P_n(t)=1$$

- 如果过程在有限时刻 t 不能增长到 ∞ 状态, 也即

$$P(X(t)=\infty)=1-\sum_{n=0}^{\infty}P_n(t)=0$$

- 这意味着过程状态增长到 ∞ 的时间为有限时间 t 的概率为 0, 也即

$$\sum_{n=0}^{\infty}\frac{1}{\lambda_n}=\infty$$

- 如果过程状态增长到 ∞ 的时间为有限时间 t 的概率为 0, 则过程在有限时刻 t 为 ∞ 的概率为 0, 一定为某个 n, 概率正则性得到满足

例 3.12 Yule 过程

- Yule 过程是一类特殊的纯生过程，在物理学和生物学中有广泛应用
- 假定在长度为 h 的小时间段内，群体中每个成员都有概率 $\beta h+o(h)$ 产生一个新个体
- 假定成员产生新个体的行为是相互独立的，则有

$$\begin{aligned}
& P(X(t+h)-X(t)=0\,|X(t)=n) \\
= \; & C_n^0\,[1-\beta h+o(h)]^n \\
= \; & 1-n\beta h+o(h)
\end{aligned}$$

$$\begin{aligned}
& P(X(t+h)-X(t)=1\,|X(t)=n) \\
= \; & C_n^1\,[\beta h+o(h)]\,[1-\beta h+o(h)]^{n-1} \\
= \; & n\beta h+o(h)
\end{aligned}$$

$$\begin{aligned}
& P(X(t+h)-X(t)\geq 2\,|X(t)=n) \\
= \; & 1-P(X(t+h)-X(t)=0|X(t)=n) \\
& -P(X(t+h)-X(t)=1|X(t)=n) \\
= \; & o(h)
\end{aligned}$$

纯生过程 (续)

$$P(X(t+h)-X(t)=1|X(t)=n)=n\beta h+o(h)$$

- 注意到 Yule 过程满足标准纯生过程的假设 (ii), (iii), (iv)
- 但假设 (i) 要求 $X(0)=0$, 而Yule 过程中 $X(0)=1$
- 令 $Y(t)=X(t)-1$, 此时 $Y(t)$ 满足标准纯生过程的四个假设, 且

$$P(Y(t+h)-Y(t)=1|Y(t)=n-1)=n\beta h+o(h),\ n\geq 1$$

- 这意味着 $\lambda_n=(n+1)\beta$
- 回忆标准纯生过程在 t 时刻种群数量为 n 的概率:

$$P_n(t)=P(Y(t)=n|Y(0)=0)=\left(\prod_{i=0}^{n-1}\lambda_i\right)\left[\sum_{i=0}^{n}c_{i,n}e^{-\lambda_i t}\right],$$

其中 $c_{i,n}=\left[\prod_{0\leq k\leq n,k\neq i}(\lambda_k-\lambda_i)\right]^{-1}$

纯生过程 (续)

- 将 $\lambda_n = (n+1)\beta$ 代入, 得

$$c_{i,n} = \left[\prod_{0\le k\le n, k\ne i} (\lambda_k - \lambda_i)\right]^{-1} = \left[\prod_{0\le k\le n, k\ne i} [k-1-(i-1)]\beta\right]^{-1} = \frac{(-1)^i}{\beta^n i!(n-i)!}$$

- 从而有

$$\begin{aligned} P_n(t) &= \left(\prod_{i=0}^{n-1} \lambda_i\right)\left[\sum_{i=0}^{n} c_{i,n} e^{-\lambda_i t}\right] \\ &= \left(\prod_{i=0}^{n-1} \beta(i+1)\right)\left[\sum_{i=0}^{n} \frac{(-1)^i}{\beta^n i!(n-i)!} e^{-\beta(i+1)t}\right] \\ &= \sum_{i=0}^{n} \frac{n!}{i!(n-i)!}(-1)^i e^{-\beta(i+1)t} \\ &= e^{-\beta t}\sum_{i=0}^{n} C_n^i (-e^{-\beta t})^i \cdot 1^{n-i} = e^{-\beta t}(1-e^{-\beta t})^n \end{aligned}$$

- 由于 $X(t) = Y(t) + 1$, 故

$$P(X(t) = n | X(0) = 1) = P(Y(t) = n-1 | Y(0) = 0) = e^{-\beta t}(1-e^{-\beta t})^{n-1}$$

第 3 章 (下) 习题三

- 3.5 节作业: 23, 24, 26

第 3 章习题汇总

- 3.1 节作业: 3, 4, 6
- 3.2 节作业: 11, 12, 16
- 3.3 节作业: 17, 18, 19
- 3.4 节作业: 20, 21, 22
- 3.5 节作业: 23, 24, 26