

Key concepts:

- 连续时间 *Markov* 链;
- 转移速率矩阵/生成元/ Q 矩阵;
- *Kolmogorov* 向前-向后方程。

9.1 连续时间 Markov 链的基本定义

先给出连续时间 Markov 链的定义

Definition 9.1 (连续时间 Markov 链) 设 E 是一个可数集合, 称状态空间为 E 的连续时间随机过程 $\{X(t)\}$ 为一个连续时间 *Markov* 链 (*Continuous-time Markov Chain, CTMC*), 如果对于任意 $t, s \geq 0$, $s > s_k \geq s_{k-1} \geq \cdots \geq s_1 > 0, k \in \mathbb{N}$, 任意状态 $i_1, i_2, \dots, i_k, i, j \in E$, 有 *Markov* 性:

$$P(X(t+s) = j | X(s) = i, X(s_k) = i_k, \dots, X(s_1) = i_1) = P(X(t+s) = j | X(s) = i) \quad (9.1)$$

成立。如果

$$P((X(t+s) = j | X(s) = i) = P((X(t) = j | X(0) = i), \quad \forall s \geq 0$$

则该 *Markov* 链为齐次的。记 $P_{ij}(t)$ 为从状态 i 出发经过 t 时间间隔后转移到状态 j 的转移概率, 即

$$P_{ij}(t) = P(X(t) = j | X(0) = i)$$

以后讨论中如无特别说明, 所讨论的 *Markov* 链是齐次的。

Definition 9.2 (转移概率矩阵) 称由转移概率组成的矩阵

$$P(t) = \{P_{ij}(t)\}_{i,j \in E} = \begin{pmatrix} P_{00}(t) & P_{01}(t) & P_{02}(t) & \cdots \\ P_{10}(t) & P_{11}(t) & P_{12}(t) & \cdots \\ P_{20}(t) & P_{21}(t) & P_{22}(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

转移概率矩阵，其中，矩阵中各元素满足

$$P_{ij}(t) \geq 0, \quad i, j \in E, \quad \sum_{j \in E} P_{ij}(t) = 1$$

Proposition 9.3 (Chapman–Kolmogorov 方程)

对任意 $i, j \in E, t, s \geq 0$,

$$P_{ij}(s+t) = \sum_{k \in E} P_{ik}(s)P_{kj}(t)$$

写成矩阵形式为

$$P(s+t) = P(s)P(t), \quad P(0) = I \quad (9.2)$$

Proof:

$$\begin{aligned} P(X(s+t) = j | X(0) = i) &= \sum_{k \in E} P(X(s+t) = j, X(s) = k | X(0) = i) \\ &= \sum_{k \in E} P(X(s+t) = j | X(s) = k, X(0) = i) P(X(s) = k | X(0) = i) \\ &= \sum_{k \in E} P_{kj}(t) P_{ik}(s) \quad (\text{Markov 性}) \end{aligned}$$

■

注1. 我们需要对 $P(t)$ 做一个连续性假设

$$\lim_{\Delta t \downarrow 0} P(\Delta t) = I$$

即 $P(\Delta t)$ 的每一个元素都在原点处连续。满足这个条件的转移概率通常称为标准的(standard)转移概率。转移概率的标准性可以保证其不仅在原点，而且在任意的 t 处都是连续的，即

$$\lim_{\Delta t \downarrow 0} P(t + \Delta t) = P(t)$$

标准性的物理概念也非常明显。当转移所需要的时间趋于 0 时，不同状态间转移发生的概率也随之减小，链以接近于 1 的概率停留在原状态。这是由于对于一般物理系统，转移需要消耗能量，而在瞬间内积聚状态改变所需的能量需要无穷大的功率，这一般情况下是无法达到的。所以今后如不加说明，所研究的转移概率都是标准的。

注2. 式(9.2)被称为 $\{P(t)\}_{t \geq 0}$ 的半群(Semigroup)性质，我们称 $\{P(t)\}_{t \geq 0}$ 为一个Markov半群。

9.2 转移速率矩阵与Kolmogorov向前-向后方程

离散时间时齐Markov的 n 步转移矩阵

$$P^{(n)} = P^n$$

是由一个“最小的” P 所生成的，但是连续时间和离散时间 Markov 链不同的是，时间 s, t 不是非负整数，所以找不到“最小”的时间单位 Δ ，使得对于所有的 t 都有

$$P(t) = P^{t/\Delta}(\Delta)$$

这正是连续时间随机过程的复杂之处。因而需要找到一个能起到相当于离散 Markov 链中一步转移矩阵 P 所起“作用”的量，并利用它来计算 $P(t)$ 。

从CK方程

$$P(s+t) = P(s)P(t), \quad P(0) = I$$

来看，这是一个无限维的矩阵值函数方程，我们先形式地考虑一个函数方程

$$f(s+t) = f(s)f(t), \quad f(0) = 1$$

这个方程在 $t = 0$ 处连续的通解为

$$f(t) = e^{\alpha t}$$

其中 $\alpha = f'(0)$ ，由此，我们猜测 $P(t)$ 可以表示为

$$P(t) = \exp(P'(0)t)$$

下面我们先定义转移速率矩阵。

Definition 9.4 (转移速率矩阵) 转移矩阵 $P(t)$ 对应的转移速率矩阵 $Q = (q_{ij})$ 定义为

$$Q = \lim_{\Delta t \rightarrow 0} \frac{P(\Delta t) - I}{\Delta t} = \frac{d}{dt} P(t)|_{t=0} = P'(0)$$

分量表示为

$$q_{ii} = \lim_{\Delta t \rightarrow 0} \frac{P_{ii}(\Delta t) - 1}{\Delta t} =: -q_i, \quad q_i \geq 0 \quad (9.3)$$

$$q_{ij} = \lim_{\Delta t \rightarrow 0} \frac{P_{ij}(\Delta t)}{\Delta t}, \quad j \neq i \quad (9.4)$$

称 q_{ij} 为从状态 i 到状态 j 的转移速率 (transition rate)。

注1. 转移速率矩阵 Q 也被称为 $P(t)$ 的生成元(generator)，也简称为 Q 矩阵。

注2. 转移速率矩阵 Q 的良好性：

Proposition 9.5 设转移矩阵 $P(t)$ 是标准的，即满足连续性条件 $\lim_{\Delta t \downarrow 0} P(\Delta t) = I$ ，那么

(1)

$$q_{ii} = \lim_{\Delta t \downarrow 0} \frac{P_{ii}(\Delta t) - 1}{\Delta t} =: -q_i$$

存在，但可能为 ∞ ；

(2)

$$q_{ij} = \lim_{\Delta t \downarrow 0} \frac{P_{ij}(\Delta t)}{\Delta t}, \quad j \neq i$$

存在且有限。

(3)

$$\sum_{j \neq i} q_{ij} \leq q_i$$

证明参考《应用随机过程》，陈大岳、章复熹，北京大学出版社，2023，命题2.6.3.

总结命题9.5中的性质，可以给出下面的定义

Definition 9.6 (Q 矩阵) $Q = (q_{ij})$ 称为一个 Q 矩阵, 如果

$$(1) \quad q_{ii} = -q_i \leq 0, \quad q_i \text{ 可以取 } +\infty$$

$$(2) \quad 0 \leq q_{ij} < +\infty, \quad j \neq i$$

$$(3) \quad \sum_{j \neq i} q_{ij} \leq q_i$$

进一步, 若 $\sum_{j \neq i} q_{ij} = q_i < +\infty$, 即 $\sum_{j \in E} q_{ij} = 0$, 则称 Q 矩阵是保守的 (*conservation*), 此后我们都考虑保守的 Q 矩阵。

对于给定 Q 矩阵, 若存在 *Markov* 链满足

$$q_{ii} = \lim_{\Delta t \downarrow 0} \frac{P_{ii}(\Delta t) - 1}{\Delta t} =: -q_i, \quad q_i \geq 0$$

$$q_{ij} = \lim_{\Delta t \downarrow 0} \frac{P_{ij}(\Delta t)}{\Delta t}, \quad j \neq i$$

则称此 *Markov* 链为以 Q 为密度的连续时间 *Markov* 链。

对于离散时间的情形, 我们通过一步转移矩阵 P 计算 n 步转移矩阵, 对于连续时间, 计算转移矩阵 $P(t)$ 会复杂很多, 下面我们考虑状态空间 E 有限时如何计算 $P(t)$.

给定矩阵 Q , 定义

$$e^{Qt} \triangleq \sum_{k=0}^{\infty} \frac{Q^k}{k!} t^k$$

Proposition 9.7 对于有限状态连续时间 *Markov* 链, 则它的转移速率矩阵 Q 一定是保守的, 而且转移矩阵 $\{P(t)\}$ 满足 *Kolmogorov* 向前方程

$$\frac{d}{dt} P(t) = P(t)Q, \quad P(0) = I \quad (9.5)$$

以及 *Kolmogorov* 向后方程

$$\frac{d}{dt} P(t) = QP(t), \quad P(0) = I \quad (9.6)$$

即 $P(t) = e^{Qt}$.

Proof: 由定义,

$$q_i = \lim_{\Delta t \rightarrow 0} \frac{1 - P_{ii}(\Delta t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \sum_{j \neq i} \frac{P_{ij}(\Delta t)}{\Delta t} \stackrel{*}{=} \sum_{j \neq i} \lim_{\Delta t \rightarrow 0} \frac{P_{ij}(\Delta t)}{\Delta t} = \sum_{j \neq i} q_{ij}$$

*成立是因为状态空间有限, 所以 Q 是保守的。由CK方程

$$\begin{aligned} \frac{1}{\Delta t} (P_{ij}(t + \Delta t) - P_{ij}(t)) &= \frac{1}{\Delta t} \left(\sum_{k=1}^N P_{ik}(t) P_{kj}(\Delta t) - P_{ij}(t) \right) \\ &= \sum_{k=1}^N P_{ik}(t) \frac{P_{kj}(\Delta t) - \delta_{kj}}{\Delta t} \\ \frac{1}{\Delta t} (P_{ij}(t + \Delta t) - P_{ij}(t)) &= \frac{1}{\Delta t} \left(\sum_{k=1}^N P_{ik}(\Delta t) P_{kj}(t) - P_{ij}(t) \right) \\ &= \sum_{k=1}^N \frac{P_{ik}(\Delta t) - \delta_{ik}}{\Delta t} P_{kj}(t) \end{aligned}$$

令 $\Delta t \downarrow 0$, 即得

$$\frac{d}{dt} P(t) = P(t)Q = QP(t)$$

由于状态空间有限, 所以 Q 是有界算子, 那么方程解是唯一的, 即 $P(t) = e^{Qt}$. ■

注1. Kolmogorov向前方程又称Fokker-Planck方程

注2. 对于状态空间是可数的形式, Kolmogorov向前-向后方程需要一些条件才能成立。具体来说, 设 $\{P(t)\}$ 是一个Markov半群, Q 为其对应的转移速率矩阵, 若 Q 是保守的, 则Kolmogorov向后方程成立; 若 $\sup_{j \in E} q_j < \infty$, 则Kolmogorov向前方程成立。

注3. 矩阵乘积运算是不可交换的, 所以Kolmogorov向前-向后方程是两个方程, 只是由于时齐性看起来有了 $\frac{d}{dt} P(t) = P(t)Q = QP(t)$, 向后方程本质上应该写为

$$\begin{aligned} &\frac{1}{\Delta t} (P_{ij}(t) - P_{ij}(t - \Delta t)) \\ &= \frac{1}{\Delta t} \left(\sum_{k=1}^N P_{ik}(\Delta t) P_{kj}(t - \Delta t) - P_{ij}(t - \Delta t) \right) \\ &= \sum_{k=1}^N \frac{P_{ik}(\Delta t) - \delta_{ik}}{\Delta t} P_{kj}(t - \Delta t) \end{aligned}$$

有限的所以可以交换极限

$$P_{ij}(t) = \sum_{k=1}^N P_{ik}(t) P_{kj}(0)$$

$$P_{ij}(t) = \sum_{k=1}^N P_{ik}(0) P_{kj}(t)$$

反证法

思考：若连续时间Markov链是非时齐的，即转移速率矩阵依赖于时间 $Q(t)$ ，那么转移概率 $P(s, t)$ 的Kolmogorov向前-向后方程是什么？

9.3 例子

$$\begin{aligned}\frac{d}{dt} P(s, t) &= P(s, t) Q(t) \\ \frac{d}{ds} P(s, t) &= -Q(s) P(s, t)\end{aligned}$$

Example 9.8 (Poisson过程) 回顾Poisson过程的定义：若计数过程 $\{N(t), t \geq 0\}$ 满足：

(1) $N(0) = 0$;

(2) 具有独立增量；

(3) 在长度为 t 的任意区间中的事件数服从以 λt 为均值的Poisson分布，即，对于任意 $s, t \geq 0$ ，有

$$P\{N(t+s) - N(s) = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, \dots$$

则称为具有速率 $\lambda (\lambda > 0)$ 的Poisson过程。Poisson过程 $\{N(t), t \geq 0\}$ 是状态空间为 $\{0, 1, 2, \dots\}$ ，转移概率为

$$P_{ij}(t) = P(N(t+s) = j | N(s) = i) = \begin{cases} \frac{(\lambda t)^{j-i}}{(j-i)!} \exp(-\lambda t), & j \geq i \\ 0, & j < i \end{cases}$$

的连续时间Markov链，其中 $P_{ij}(0) = \delta_{ij}$

Proof: 由于条件概率

$$\begin{aligned}P(N(t_{n+1}) = j | N(t_n) = i, N(t_{n-1}) = i_{n-1}, \dots, N(t_0) = i_0) \\ &= P(N(t_{n+1}) - N(t_n) = j - i | N(t_n) = i, \dots, N(t_0) = i_0) \\ &= P(N(t_{n+1}) - N(t_n) = j - i) \quad (\text{独立增量}) \\ &= P(N(t_{n+1} - t_n) = j - i). \quad (\text{平稳增量})\end{aligned}$$

只依赖于 $i, j, t_{n+1} - t_n$ ，所以Poisson过程 $\{N(t), t \geq 0\}$ 是一个连续时间Markov链，转移概率为

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

所以转移矩阵为

$$P(t) = \begin{pmatrix} 1 & \lambda t & \frac{(\lambda t)^2}{2!} & \cdots & \frac{(\lambda t)^k}{k!} & \cdots \\ 0 & 1 & \lambda t & \cdots & \frac{(\lambda t)^{k-1}}{(k-1)!} & \cdots \\ 0 & 0 & 1 & \cdots & \frac{(\lambda t)^{k-2}}{(k-2)!} & \cdots \\ & & & \ddots & \vdots & \\ & & & & 1 & \cdots \\ & & & & & \ddots \end{pmatrix} \exp(-\lambda t)$$

下面我们check Poisson过程的 Q 矩阵,

$$q_{ij} \triangleq \lim_{\Delta t \downarrow 0} \frac{P_{ij}(\Delta t) - \delta_{ij}}{\Delta t} = \begin{cases} \lambda, & j = i + 1 \\ -\lambda, & j = i \\ 0, & \text{其他} \end{cases}$$

$$Q = (-\lambda) \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

用归纳法容易验证

$$Q^k = (-\lambda)^k \begin{pmatrix} C_k^0 & C_k^1(-1)^1 & C_k^2(-1)^2 & C_k^3(-1)^3 & \cdots \\ 0 & C_k^0 & C_k^1(-1)^1 & C_k^2(-1)^2 & \cdots \\ 0 & 0 & C_k^0 & C_k^1(-1)^1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix},$$

其中对于 $j < 0$ 或 $j > k$, 规定 $C_k^j = 0$. 定义

$$q_{ij}^{(k)} = (-\lambda)^k C_k^{j-i} (-1)^{j-i}, k \geq 1, i, j \in I,$$

利用 $P_{ij}(0) = \delta_{ij}$ 得到

$$Q^k = \left(q_{ij}^{(k)} \right), Q^0 = I,$$

并且对 $j \geq i$, 有

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{t^k}{k!} q_{ij}^{(k)} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda^{j-i} C_k^{j-i} (-\lambda)^{k-(j-i)} \\ &= \sum_{k=j-i}^{\infty} \frac{(\lambda t)^{j-i}}{(j-i)! [k-(j-i)]!} (-\lambda t)^{k-(j-i)} \\ &= \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t} = P_{ij}(t) \end{aligned}$$

写成矩阵的形式就得到

$$P(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} (q_{ij}^{(k)}) = \sum_{k=0}^{\infty} \frac{1}{k!} (tQ)^k = e^{tQ}.$$

■

Poisson过程的例子是我们先写出转移矩阵 $P(t)$, 然后再通过求导得到 Q 矩阵, 然而现实中, 往往是先得到描述转移概率变化率的 Q 矩阵, 然后通过Kolmogorov向前向后方程求解转移矩阵 $P(t)$ 。下面我们看一个例子。

Example 9.9 (机器维修问题) 设某个机器的正常工作时间服从参数为 λ 的指数分布, 一旦机器损坏则立刻进行修理, 修理时间服从参数为 μ 的指数分布。如果在起始时刻 $t = 0$ 机器处于正常工作状态, 计算在时刻 T 机器处于工作状态的概率?

就是去算 9

解. 这是一个两状态的 Markov 链。设正常工作状态为 0, 修理状态为 1, 则状态空间为 $\{0, 1\}$, 在 Δt 时间内, 机器从正常状态转入损坏修理状态的概率为

$$P_{01}(\Delta t) = 1 - \exp(-\lambda \Delta t) = \lambda \Delta t + o(\Delta t)$$

且有 $P_{00}(\Delta t) = 1 - P_{01}(\Delta t)$ 。另一方面, 在 Δt 时间内, 机器从修理状态转入正常状态的概率为

$$P_{10}(\Delta t) = 1 - \exp(-\mu \Delta t) = \mu \Delta t + o(\Delta t)$$

且有 $P_{11}(\Delta t) = 1 - P_{10}(\Delta t)$ 。所以 Q 矩阵为

$$Q = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$$

求 Q 的特征值为

$$\lambda_1 = 0, \quad \lambda_2 = -(\lambda + \mu)$$

相应的特征向量为

$$(1, 1)^T, \quad \left(-1, \frac{\mu}{\lambda}\right)^T$$

求解Kolmogorov方程, 可得

$$P(t) = \exp(Qt)$$

$$\begin{aligned} &= \begin{pmatrix} 1 & -1 \\ 1 & \frac{\mu}{\lambda} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \exp(-(\lambda + \mu)t) \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & \frac{\mu}{\lambda} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \exp(-(\lambda + \mu)t) & \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} \exp(-(\lambda + \mu)t) \\ \frac{\mu}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} \exp(-(\lambda + \mu)t) & \frac{\lambda}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \exp(-(\lambda + \mu)t) \end{pmatrix} \end{aligned}$$

由于机器在时刻 0 正常工作, 所以初始概率为 $(1, 0)^T$, 因此在时刻 T 仍然正常工作的概率为

$$P_0(T) = P_0(0)P_{00}(T) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \exp(-(\lambda + \mu)T)$$

9.4 Q 矩阵的概率意义



连续时间 Markov 链的行为可以从它的两个特征“停留”和“跳变”来考察, 这两个特征都和 Q 矩阵有着密切关系, 本节我们给出 Q 矩阵的元素 q_{ij} 的概率意义。

Theorem 9.10 设 $X(t)$ 为一个连续时间 Markov 过程, 定义 $\tau \triangleq \inf\{t \geq 0 : X(t) \neq i, X(0) = i\}$ 表示 $X(t)$ 在状态 i 的停留时间, 则

$$(1) P(\tau > t | X(0) = i) = e^{-q_i t}, t \geq 0;$$

$$(2) \text{ 当 } j \neq i \text{ 时, } P(X(\tau) = j, \tau \leq t | X(0) = i) = \frac{q_{ij}}{q_i} (1 - e^{-q_i t});$$

$$(3) \text{ 当 } j \neq i \text{ 时, } P(X(\tau) = j | X(0) = i) = \frac{q_{ij}}{q_i}$$

$= \sum_j q_{ij}$ \rightarrow 直观理解: 一个事件发生的概率

Proof: (1) 注意到

$$P(\tau > t | X(0) = i) = P(X(u) = i, u \in [0, t] | X(0) = i)$$

停留多久

跳变到哪

我们下面计算上式右端。注意到集合列 $B_n \triangleq \{ \frac{jt}{2^n}, 1 \leq j \leq 2^n \}$ 单调上升(越加越细), 事件列

$$A_n \triangleq \{ X(\frac{jt}{2^n}) = i, 1 \leq j \leq 2^n \} = \{ X(u) = i, u \in B_n \}$$

单调下降: $A_1 \supset A_2 \supset \dots$, 因为 $B = \bigcup_{n=1}^{\infty} B_n$ 在 $[0, t]$ 中稠密, $X(t)$ 的轨道又是阶梯形状和右连续的, 所以有

$$\{X(u) = i, u \in [0, t]\} = \bigcap_{j=1}^{\infty} A_n \quad a.s..$$

由 **概率的连续性**

$$\begin{aligned} & P(X(u) = i, u \in [0, t] \mid X(0) = i) \\ &= \lim_{n \rightarrow \infty} P(A_n \mid X(0) = i) \\ &= \lim_{n \rightarrow \infty} P\left(X\left(\frac{jt}{2^n}\right) = i, 1 \leq j \leq 2^n \mid X(0) = i\right) \\ &= \lim_{n \rightarrow \infty} \left[P_{ii}\left(\frac{t}{2^n}\right) \right]^{2^n} \\ &= \lim_{n \rightarrow \infty} \left[P_{ii}\left(\frac{t}{n}\right) \right]^n \quad \text{展开} \\ &= \lim_{n \rightarrow \infty} \left[P_{ii}(0) + P'_{ii}(0) \frac{t}{n} + o\left(\frac{t}{n}\right) \right]^n \\ &= \lim_{n \rightarrow \infty} \left[1 - q_i \frac{t}{n} + o\left(\frac{t}{n}\right) \right]^n \\ &= e^{-q_i t}. \end{aligned}$$

(2) 重新定义

$$B_n \triangleq \{ jt/2^n \mid 1 \leq j \leq 2^n - 1 \}$$

$$A_n \triangleq \{ X(u) = i, u \in B_n \}$$

A_n 单调下降, 使得

$$\{X(u) = i, u \in [0, t)\} = \bigcap_{j=1}^{\infty} A_n.$$

对于 $t, \Delta t > 0$ 和 $j \neq i$, 从 Taylor 展开公式

$$P(X(t) = j \mid X(t - \Delta t) = i) = P_{ij}(\Delta t) = q_{ij} \Delta t + o(\Delta t)$$

知道

$$\lim_{\Delta t \rightarrow 0} P(X(t) = j \mid X(t - \Delta t) = i) = q_{ij} \Delta t.$$

取 $\Delta t = \frac{t}{2^n}$, 得到

$$\begin{aligned}
 & P(X(\tau) = j, \tau = t | X(0) = i) \\
 &= P(\{X(u) = i, u \in [0, t]\}, X(t) = j | X(0) = i) \\
 &= \lim_{n \rightarrow \infty} P(A_n, X(t) = j | X(0) = i) \\
 &= \lim_{n \rightarrow \infty} P(A_n | X(0) = i) P(X(t) = j | A_n, X(0) = i) \\
 &= \lim_{n \rightarrow \infty} [P_{ii}(\frac{t}{2^n})]^{2^n-1} P(X(t) = j | X(t - \Delta t) = i) \\
 &= \lim_{n \rightarrow \infty} (1 - q_i \frac{t}{2^n} + o(\frac{t}{2^n}))^{2^n-1} (q_{ij} \Delta t + o(\Delta t)) \\
 &= e^{-q_i t} q_{ij} dt.
 \end{aligned}$$

两边对于 $t \in [0, s]$ 积分得到

$$P(X(\tau) = j, \tau \leq s | X(0) = i) = \int_0^s q_{ij} e^{-q_i t} dt = \frac{q_{ij}}{q_i} (1 - e^{-q_i s})$$

(3) 在(2)中令 $t \rightarrow \infty$ 即得。 ■

注1. 当 $q_i = 0$ 时,

$$P(\tau > t | X(0) = i) = e^{-q_i t} = 1$$

这表明 $X(t)$ 不会离开状态 i , 称状态 i 为吸收状态(absorbing/traps state)。当 $q_i = +\infty$ 时,

$$P(\tau > t | X(0) = i) = 0$$

这表明 $X(t)$ 在状态 i 几乎不会停留, 称状态 i 为瞬时状态(instantaneous state)。当 $0 < q_i < +\infty$ 时, 称状态 i 为逗留状态(stable state)。我们只考虑逗留的Markov链。

注2. 由定理9.10的结论, 从状态 i 出发在状态 i 的停留时间服从参数 q_i 的指数分布, 基于此我们可以给出连续时间Markov链行为的具体描述。

定义

$$k_{ij} \triangleq \begin{cases} q_{ij}/q_i, & \text{当 } q_i > 0, j \neq i, \\ 0, & \text{当 } q_i > 0, j = i, \\ \delta_{ij}, & \text{当 } q_i = 0, \end{cases}$$

则 $K = (k_{ij})$ 的各行之和为1. 定义

$$\begin{aligned}\tau_0 &= 0, \\ \tau_1 &= \inf\{t > 0 | X(t) \neq X(0)\}, \\ \tau_2 &= \inf\{t > \tau_1 | X(t) \neq X(\tau_1)\}, \\ &\dots \\ \tau_n &= \inf\{t > \tau_{n-1} | X(t) \neq X(\tau_{n-1})\}, \\ &\dots\end{aligned}$$

有趣

则 τ_i 是马氏链 $\{X(t)\}$ 的第 i 次转移时刻。 $T_i = \tau_{i+1} - \tau_i$ 是第 i 次转移后的停留时间。从定理9.10得到Markov链的以下结果：

(1) 连续时间 \rightarrow 离散时间：

$X_n = X(\tau_n)$ ($n = 0, 1, \dots$) 是以 $K = (k_{ij})$ 为一步转移概率矩阵的离散时间Markov链；

沿着 $\{X(\tau_n)\}$ 的给定轨迹 $i_0 \rightarrow i_1 \rightarrow i_2 \rightarrow \dots$ ，Markov链在各状态的依次停留时间 T_0, T_1, T_2, \dots 相互独立， T_j 服从参数为 q_{i_j} 的指数分布， $j = 0, 1, 2, \dots$ ；

(2) 离散时间 \rightarrow 连续时间：

设 $\{Y_n\}$ 是离散时间Markov链，以 $K = (k_{ij})$ 为转移概率矩阵。对每个 $i \in I$ ，假设 $\{Y_n\}$ 每次到达 i 后，在 i 的停留时间是相互独立的随机变量，服从共同的参数为 q_i 指数分布，停留结束时以概率 k_{ij} 转移到状态 j ($j \neq i$)，进一步假设 $\{Y_n\}$ 在不同状态的停留时间相互独立，则用 $X(t)$ 表示 t 时 $\{Y_n\}$ 的状态时， $\{X(t)\}$ 是连续时间的Markov链，有转移速率矩阵 Q 。

在 (1) 中，称离散时间Markov链 $\{X_n\}$ 为 $\{X(t)\}$ 的嵌入链(embedded chain)或跳跃链(jump chain)。同理在 (2) 中也称离散时间马氏链 $\{Y_n\}$ 为 $\{X(t)\}$ 的嵌入链。嵌入链的转移概率矩阵 K 由转移速率矩阵 Q 决定。