

第五章 连续时间马氏链

随机过程及其在金融中的应用

中国人民大学出版社

本章内容

① 定义和例子

- 连续时间马氏链的概念
- 连续时间马氏链的性质
- 转移速率
- 连续时间马氏链的例子

② 转移概率

- 柯尔莫哥洛夫向后方程

- 转移概率的求解

③ 极限行为

- 平稳分布
- 细致平衡条件

④ 嵌入链

- 嵌入链的概念及特征
- 利用嵌入链计算平稳分

布

连续时间马氏链的概念

对于任意状态 i, j, i_0, \dots, i_n 以及任意时间

$0 \leq s_0 < s_1 < \dots < s_n < s$, 有:

$$\begin{aligned}\mathbb{P}(X_{t+s} = j | X_s = i, X_{s_n} = i_n, \dots, X_{s_0} = i_0) &= \mathbb{P}(X_{t+s} = j | X_s = i) \\ &= \mathbb{P}(X_t = j | X_0 = i)\end{aligned}$$

满足上式的就是连续时间马氏链。

从时刻 s 的状态 i 到时刻 $(t+s)$ 的状态 j 的概率, 只依赖于时间间隔 t , 而与起始时间 s 无关。

注意:

在离散时间马氏链中, 时间和状态均是离散的。而在连续时间马氏链中, 时间是连续的, 状态是离散的。

连续时间马氏链的概念 (cont.)

为了便于区分, 离散时间马氏链经过 n 步, 从状态 i 转移到状态 j 的概率记作 $p^n(i, j)$ (注意: n 是以上标形式体现), 即:

$$p^n(i, j) = \mathbb{P}(X_{t+n} = j | X_t = i) = \mathbb{P}(X_n = j | X_0 = i), \quad n, t \in \mathbb{Z}$$

而在连续时间马氏链中, 在 $h > 0$ 时间段, 从状态 i 转移到状态 j 的概率记作 $p_h(i, j)$ (注意: h 是以下标形式体现), 即:

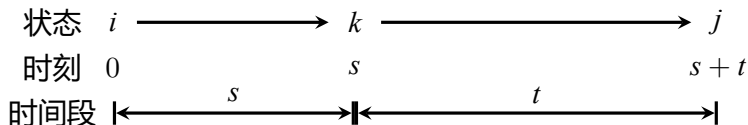
$$p_h(i, j) = \mathbb{P}(X_{t+h} = j | X_t = i) = \mathbb{P}(X_h = j | X_0 = i), \quad h, t \in \mathbb{R}^+$$

此处的连续时间马氏链仍然假设其具有时间齐次性。

连续时间马氏链的性质

连续时间马氏链的 C-K 方程：

$$\sum_k p_s(i, k) p_t(k, j) = p_{s+t}(i, j)$$



与离散时间马氏链类似，由连续时间马氏链的 C-K 方程也可以得到如下不等式：

$$p_{s+t}(i, j) \geq p_s(i, k) \cdot p_t(k, j), \quad \forall k$$

连续时间马氏链的相关定理

- 假设 T_i 是连续时间马氏链停留在状态 i 的时长, 则 T_i 服从指数分布。
- 对于连续时间马氏链, 若对于某个 $t > 0$, 有 $p_t(i, j) > 0$, 则对任意 $s > 0$, 均有 $p_{t+s}(i, j) > 0$ 。

连续时间马氏链中的所有状态均是非周期的, 因此无须考虑周期性。

转移速率

当 $h \rightarrow 0$ 时, 引入转移速率 $q(i, j)$, 即:

$$q(i, j) = \lim_{h \rightarrow 0} \frac{p_h(i, j)}{h} = \lim_{h \rightarrow 0} \frac{\mathbb{P}(X_{t+h} = j | X_t = i)}{h}$$

其中, $q(i, j)$ 表示从状态 i 跳到 j 的转移速率。

在连续时间马氏链中, 除了要考虑某一时刻马氏链处于什么状态以外, 还要关心它在离开这个状态之前会停留多长时间。前面已经证明了这个停留时间具备无记忆性, 服从指数分布。

通常在连续时间下, 通过转移速率来描述系统更为简单。

转移速率 $q(i, j)$ 的性质

- ① $q(i, i) \leq 0, \quad i = 1, 2, \dots, N$
- ② $q(i, j) \geq 0, \quad i \neq j, \quad i, j = 1, 2, \dots, N$
- ③ $\sum_{j=1}^N q(i, j) = 0, \quad i = 1, 2, \dots, N$

由于状态间的转移是发生在状态空间内，因此总的转移速率应当为零。这就如同一个封闭容器中水的总量不变，若水从 A 处流向 B 处的流速为 x ，则意味着由 B 处流向 A 处的流速为 $-x$ ，流速之和仍然为零。

如果转移速率 $q(i, j)$ 对任意状态 j 均满足 $q(i, j) = 0$ ，则称状态 i 是吸收态。此时的转移速率均为零，对应到转移速率矩阵上，体现为状态 i 一行的元素取值均为零，说明马氏链将永远停留在状态 i 处。

转移速率矩阵

$$\mathbf{Q} = \begin{bmatrix} q(1,1) & q(1,2) & \cdots & q(1,N) \\ q(2,1) & q(2,2) & \cdots & q(2,N) \\ \vdots & \vdots & & \vdots \\ q(N,1) & q(N,2) & \cdots & q(N,N) \end{bmatrix}$$

需要注意的是，由于转移速率矩阵每行元素之和等于零，因此该矩阵对角线上的元素 $q(i,i)$ 的绝对值应当等于该行其他元素之和，即：

$$|q(i,i)| = \sum_{k \neq i} q(i,k), \quad \forall i$$

记 $q_i = |q(i,i)| = -q(i,i)$ ，这里的 q_i 就是离开状态 i 的速率。

举例 1: 泊松过程

$N(t)$ 表示速率为 λ 的泊松过程到时刻 t 为止的到达数。对于任意时间段 h , 到达数由 n 增加到 $(n+1)$ 的概率为:

$$\begin{aligned} p_h(n, n+1) &= \frac{(\lambda h)^1}{1!} e^{-\lambda h} = \lambda h e^{-\lambda h} \\ &= \lambda h \left[1 - \lambda h + \frac{1}{2}(\lambda h)^2 - \frac{1}{3!}(\lambda h)^3 + \cdots \right] \end{aligned}$$

因此:

$$\begin{aligned} q(n, n+1) &= \lim_{h \rightarrow 0} \frac{p_h(n, n+1)}{h} \\ &= \lim_{h \rightarrow 0} \lambda \left[1 - \lambda h + \frac{1}{2}(\lambda h)^2 - \frac{1}{3!}(\lambda h)^3 + \cdots \right] = \lambda \end{aligned}$$

泊松过程 (cont.)

在时间 h 内至少经历两步转移的概率为:

$$\begin{aligned}\mathbb{P}[N(t+h) \geq n+2 | N(t) = n] &= 1 - [p_h(n, n) + p_h(n, n+1)] \\&= 1 - (e^{-\lambda h} + \lambda h e^{-\lambda h}) = 1 - (1 + \lambda h)e^{-\lambda h} \\&= 1 - (1 + \lambda h) \left[1 - \lambda h + \frac{1}{2}(\lambda h)^2 - \frac{1}{3!}(\lambda h)^3 + \dots \right]\end{aligned}$$

因此:

$$\begin{aligned}\frac{\mathbb{P}[N(t+h) \geq n+2 | N(t) = n]}{h} &= \frac{1}{2}\lambda^2 h + \frac{1}{3}\lambda^3 h^2 + \dots \\ \lim_{h \rightarrow 0} \frac{\mathbb{P}[N(t+h) \geq n+2 | N(t) = n]}{h} &= o(h)\end{aligned}$$

从而:

$$q(n, n+k) = 0, \quad k \geq 2$$

泊松过程 (cont.)

因此, 在到达发生的速率为 λ 的泊松过程中, 到达数 $N(t)$ 由 n 增加到 $(n+1)$ 的速率为 λ , 其余情形下超过两步的转移速率为 0, 即:

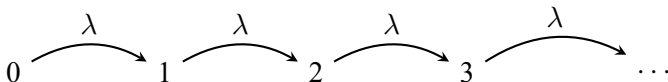
$$\begin{cases} q(n, n+1) = \lambda, \\ q(n, n+k) = 0, \quad k \geq 2 \end{cases} \quad \forall n \geq 0$$

相应的转移速率矩阵 \mathbf{Q} 如下:

$$\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \cdots \\ 0 & 0 & 0 & -\lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

泊松过程的转移速率图

$$\begin{cases} q(n, n+1) = \lambda, \\ q(n, n+k) = 0, \quad k \geq 2 \end{cases} \quad \forall n \geq 0$$



说明:

泊松过程属于计数过程, 因此 $m > n$ 时, $q(m, n) = 0$ 。

举例 2: 生灭过程

在状态空间为 $\{0, 1, 2, \dots, N\}$ 的生灭过程 (birth-and-death process) 中, 出生率为 λ_n , 死亡率为 μ_n , 因此有:

$$q(n, n+1) = \lambda_n, \quad n = 0, 1, 2, \dots, N-1$$

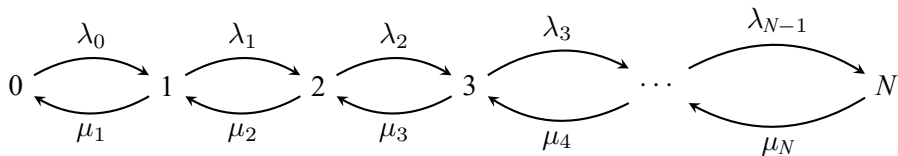
$$q(n, n-1) = \mu_n, \quad n = 1, 2, \dots, N$$

相应的转移速率矩阵 \mathbf{Q} 如下:

$$\mathbf{Q} = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots & 0 & 0 \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \mu_{N-1} & -(\lambda_{N-1} + \mu_{N-1}) & \lambda_{N-1} \\ 0 & 0 & 0 & 0 & \cdots & \mu_N & -\mu_N \end{bmatrix}$$

生灭过程的转移速率图

$$\begin{cases} q(n, n+1) = \lambda_n, & n = 0, 1, 2, \dots, N-1 \\ q(n, n-1) = \mu_n, & n = 1, 2, \dots, N \end{cases}$$



举例 3: 纯生过程

当 $\mu = 0$ 时, 生灭过程称为纯生过程 (pure-birth process)。

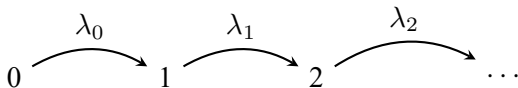
$$q(n, n+1) = \lambda_n, \quad n = 0, 1, 2, \dots$$

$$q(n, n-1) = 0, \quad n = 1, 2, \dots$$

相应的转移速率矩阵 \mathbf{Q} 如下:

$$\mathbf{Q} = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ 0 & -\lambda_1 & \lambda_1 & 0 & \cdots \\ 0 & 0 & -\lambda_2 & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

纯生过程的转移速率图



注意：

泊松过程可看作转移速率不变 ($\lambda_n = \lambda$) 的纯生过程。

举例 4: $M/M/s$ 排队系统

顾客按速率 λ 到达, 柜台有 s 个服务窗口, 对每个顾客的服务时间相互独立且服从速率为 μ 的指数分布, 因此:

$$q(n, n+1) = \lambda, \quad n = 0, 1, 2, \dots, s-1$$

$$q(n, n-1) = \begin{cases} n\mu, & 0 \leq n \leq s \\ s\mu, & n \geq s \end{cases}$$

即, 当顾客数不多于服务窗口数 ($n \leq s$) 时, 他们都在接受服务, 离开的速率为 $n\mu$; 当顾客数不少于服务窗口数 ($n \geq s$) 时, 所有服务窗口都在工作, 离开的速率为 $s\mu$ 。

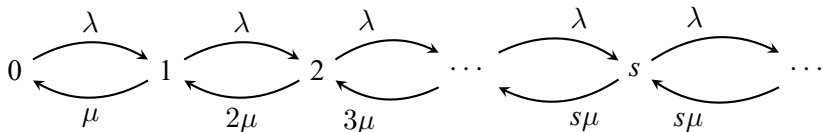
举例 4: $M/M/s$ 排队系统 (cont.)

离开速率的这一特点与指数分布的特征有关。指数分布具有如下性质：

对于独立同分布的随机变量 $\tau_1, \tau_2, \dots, \tau_n \sim \mathcal{E}(\lambda)$ ，下式成立：

$$\tau = \min(\tau_1, \tau_2, \dots, \tau_n) \sim \mathcal{E}(n\lambda)$$

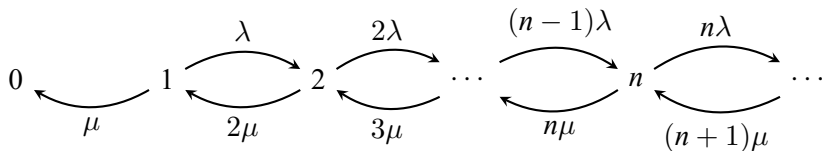
因此，在有 n 位顾客正在接受服务的情况下，最先结束服务离开的时间就是 τ ，相应的离开速率即为 $n\lambda$ 。



举例 5: 分支过程

假设每个个体死亡的速率为 μ ; 生育一个新个体的速率为 λ , 则有:

$$q(n, n+1) = n\lambda, \quad q(n, n-1) = n\mu$$

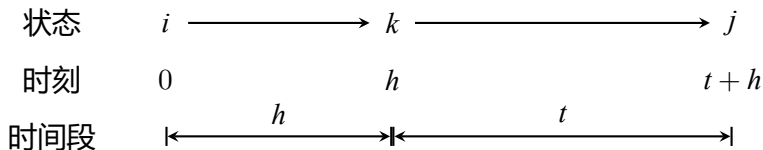


说明:

当 $\mu = 0$ 时, 该过程也称为尤尔过程 (Yule process)。

柯尔莫哥洛夫向后方程

将 $[0, t+h]$ 拆分成 $[0, h]$ 和 $[h, t+h]$



$$\begin{aligned}
 p_{t+h}(i, j) - p_t(i, j) &= \left[\sum_k p_h(i, k) p_t(k, j) \right] - p_t(i, j) \\
 &= \left[\sum_{k \neq i} p_h(i, k) p_t(k, j) \right] + [p_h(i, i) - 1] p_t(i, j)
 \end{aligned}$$

对上式两端同时除以 h , 并令 $h \rightarrow 0$ 。

$$\lim_{h \rightarrow 0} \frac{p_{t+h}(i, j) - p_t(i, j)}{h} = \lim_{h \rightarrow 0} \left[\sum_{k \neq i} \frac{p_h(i, k)}{h} p_t(k, j) \right] + \left[\lim_{h \rightarrow 0} \frac{p_h(i, i) - 1}{h} \right] p_t(i, j)$$

$$p'_t(i, j) = \sum_{k \neq i} q(i, k) p_t(k, j) + \left[\lim_{h \rightarrow 0} \frac{p_h(i, i) - 1}{h} \right] p_t(i, j)$$

注意:

$$\lim_{h \rightarrow 0} \frac{p_h(i, i) - 1}{h} = - \lim_{h \rightarrow 0} \sum_{k \neq i} \frac{p_h(i, k)}{h} = - \sum_{k \neq i} q(i, k) = -\lambda_i$$

因此:

$$\begin{aligned} p'_t(i, j) &= \sum_{k \neq i} q(i, k) p_t(k, j) - \left[\lim_{h \rightarrow 0} \frac{1 - p_h(i, i)}{h} \right] p_t(i, j) \\ &= \sum_{k \neq i} q(i, k) p_t(k, j) - \lambda_i p_t(i, j) \\ &= \sum_k Q(i, k) p_t(k, j) \end{aligned}$$

$$p'_t(i, j) = \sum_k Q(i, k) p_t(k, j)$$

其中:

$$Q(i, k) = \begin{cases} q(i, k), & i \neq k \\ -\lambda_i, & i = k \end{cases}$$

对应的矩阵形式如下:

$$\begin{bmatrix} p'_t(1, 1) & \cdots & p'_t(1, n) \\ p'_t(2, 1) & \cdots & p'_t(2, n) \\ \vdots & & \vdots \\ p'_t(n, 1) & \cdots & p'_t(n, n) \end{bmatrix} = \begin{bmatrix} Q(1, 1) & \cdots & Q(1, n) \\ Q(2, 1) & \cdots & Q(2, n) \\ \vdots & & \vdots \\ Q(n, 1) & \cdots & Q(n, n) \end{bmatrix} \begin{bmatrix} p_t(1, 1) & \cdots & p_t(1, n) \\ p_t(2, 1) & \cdots & p_t(2, n) \\ \vdots & & \vdots \\ p_t(n, 1) & \cdots & p_t(n, n) \end{bmatrix}$$

使用矩阵符号, 上式可简写为:

$$\frac{d\mathbf{P}_t}{dt} = \mathbf{Q}\mathbf{P}_t$$

该方程称为柯尔莫哥洛夫向后方程 (Kolmogorov backward equation), 其中 \mathbf{Q} 称作转移速率矩阵。在这里将 $[0, t+h]$ 拆分成 $[0, h]$ 和 $[h, t+h]$ 。

举例 6: 泊松过程的柯尔莫哥洛夫向后方程

假设泊松过程的速率为 λ , 因此:

$$q(i, i+1) = \lambda, \quad q(i, i) = -\lambda$$

根据 $p'_t(i, j) = \sum_k Q(i, k)p_t(k, j)$, 可得:

$$\begin{aligned} p'_t(i, j) &= q(i, i+1)p_t(i+1, j) + q(i, i)p_t(i, j) \\ &= \lambda p_t(i+1, j) - \lambda p_t(i, j) \end{aligned}$$

假设生灭过程的出生速率为 λ_n , 死亡速率为 μ_n , 因此:

$$q(n, n+1) = \lambda_n, \quad q(n, n-1) = \mu_n, \quad q(n, n) = -(\lambda_n + \mu_n)$$

泊松过程的柯尔莫哥洛夫向后方程 (cont.)

根据 $p'_t(n, m) = \sum_k Q(n, k)p_t(k, m)$, 可得:

$$\begin{aligned} p'_t(n, m) &= q(n, n+1)p_t(n+1, m) + q(n, n-1)p_t(n-1, m) + q(n, n)p_t(n, m) \\ &= \lambda_n p_t(n+1, m) + \mu_n p_t(n-1, m) - (\lambda_n + \mu_n)p_t(n, m) \end{aligned}$$

柯尔莫哥洛夫向前方程

如果将 $[0, t + h]$ 拆分成 $[0, t]$ 和 $[t, t + h]$, 那么得到的方程称为柯尔莫哥洛夫向前方程 (Kolmogorov forward equation), 使用矩阵符号可以简写为:

$$\frac{d\mathbf{P}_t}{dt} = \mathbf{P}_t \mathbf{Q}$$

转移概率的求解

根据微分方程的知识, 以下方程的通解为 $P(t) = Ce^{at}$:

$$\frac{dP(t)}{dt} = aP(t)$$

在多维情形下也有类似的结论。

由于:

$$\frac{d\mathbf{P}_t}{dt} = \mathbf{Q}\mathbf{P}_t, \quad \frac{d\mathbf{P}_t}{dt} = \mathbf{P}_t\mathbf{Q}$$

因此:

$$\mathbf{P}_t = e^{\mathbf{Q}t}$$

其中, \mathbf{Q} 是转移速率矩阵; \mathbf{P}_t 是转移概率矩阵。

举例

考虑两状态 $\{0, 1\}$ 的马氏链，其转移速率矩阵如下：

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$$

要计算转移概率矩阵 \mathbf{P}_t ，就要求得 $e^{\mathbf{Q}t}$ 。

举例 (cont.)

将矩阵 \mathbf{Q} 对角化: $\mathbf{Q} \rightarrow \mathbf{\Lambda}$, 相应地:

$$\mathbf{e}^{\mathbf{Q}t} \rightarrow \mathbf{e}^{\mathbf{\Lambda}t} = \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t})$$

其中, $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 。将矩阵 \mathbf{Q} 对角化, 其特征值为 0 和 -3 , 相应的矩阵可写为:

$$\mathbf{Q} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$$

其中:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}, \quad \mathbf{\Lambda} = \begin{bmatrix} 0 & 0 \\ 0 & -3 \end{bmatrix}, \quad \mathbf{X}^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

举例 (cont.)

相应地：

$$\begin{aligned}\mathbf{P}_t &= \mathbf{e}^{\mathbf{Q}t} = \mathbf{X}\mathbf{e}^{\mathbf{\Lambda}t}\mathbf{X}^{-1} \\ &= \mathbf{X} \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{e}^{-3t} \end{bmatrix} \mathbf{X}^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} + \mathbf{e}^{-3t} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{bmatrix}\end{aligned}$$

最终可得转移概率矩阵：

$$\begin{aligned}\mathbf{P}_t &= \mathbf{e}^{\mathbf{Q}t} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} + \mathbf{e}^{-3t} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{bmatrix} \\ &= \begin{bmatrix} \frac{2}{3} + \frac{1}{3}\mathbf{e}^{-3t} & \frac{1}{3} - \frac{1}{3}\mathbf{e}^{-3t} \\ \frac{2}{3} - \frac{2}{3}\mathbf{e}^{-3t} & \frac{1}{3} + \frac{2}{3}\mathbf{e}^{-3t} \end{bmatrix}\end{aligned}$$

举例 (cont.)

$$\mathbf{P}_t = \begin{bmatrix} \frac{2}{3} + \frac{1}{3}e^{-3t} & \frac{1}{3} - \frac{1}{3}e^{-3t} \\ \frac{2}{3} - \frac{2}{3}e^{-3t} & \frac{1}{3} + \frac{2}{3}e^{-3t} \end{bmatrix}$$

当 $t \rightarrow \infty$ 时,

$$\lim_{t \rightarrow \infty} \mathbf{P}_t = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

该连续时间马氏链的平稳分布为

$$\pi_1 = \frac{2}{3}, \quad \pi_2 = \frac{1}{3}$$

不可约马氏链

定义

如果对任意状态 i 和 j , 都有可能从 i 经过有限步转移到 j , 则称马氏链是不可约的, 即, 存在状态序列 $k_0 = i, k_1, \dots, k_n = j$, 使得 $q(k_{m-1}, k_m) > 0$ ($1 \leq m \leq n$)。

定理

- 如果一个连续时间马氏链 X_t 不可约, 且 $t > 0$, 则 $p_t(i, j) > 0$ 。
- 如果一个连续时间马氏链 X_t 不可约, 且具有平稳分布 π , 则:

$$\lim_{t \rightarrow \infty} p_t(i, j) = \pi(j)$$

平稳分布

连续时间马氏链具有非周期性，并且停留时间服从指数分布，因此连续时间马氏链的极限行为比离散时间马氏链更为简单。

在离散时间下，平稳分布是 $\pi \mathbf{P} = \pi$ 的一个解；而在连续时间下，对 $\forall t > 0$ ，都有 $\pi \mathbf{P}_t = \pi$ 成立。但是在实践中，由于 \mathbf{P}_t 不容易计算，因此 $\pi \mathbf{P}_t = \pi$ 成立的条件难以验证。

定理

当且仅当 $\pi \mathbf{Q} = 0$ 时， π 是一个平稳分布。

平稳分布 (cont.)

简单证明

由概率平稳性的定义可知： $\pi \mathbf{P}_t = \pi$ ，根据柯尔莫哥洛夫向后方程，有：

$$\frac{d\mathbf{P}_t}{dt} = \mathbf{P}_t \mathbf{Q}$$

方程两端同时左乘平稳概率 π ，可得：

$$\begin{aligned}\frac{d(\pi \mathbf{P}_t)}{dt} &= (\pi \mathbf{P}_t) \mathbf{Q} \\ \frac{d\pi}{dt} &= \pi \mathbf{Q}\end{aligned}$$

由于平稳概率 π 与时间 t 无关，因此可得：

$$\pi \mathbf{Q} = \mathbf{0}$$

举例 7: 天气链

某地的天气有三种状态，分别为晴、雾、雨。已知晴天的持续时间服从均值为 3 天的指数分布；雾天的持续时间服从均值为 4 天的指数分布；雨天的持续时间服从均值为 1 天的指数分布。由此得到的转移速率矩阵如下：

$$\mathbf{Q} = \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} & 0 \\ 0 & -\frac{1}{4} & \frac{1}{4} \\ 1 & 0 & -1 \end{bmatrix}$$

求每种天气所占的比例。

思路：

根据 $\pi \mathbf{Q} = \mathbf{0}$ 计算。

天气链 (cont.)

$$\begin{cases} -\frac{1}{3}\pi_1 + \pi_3 = 0 \\ \frac{1}{3}\pi_1 - \frac{1}{4}\pi_2 = 0 \\ \frac{1}{4}\pi_2 - \pi_3 = 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \Rightarrow \begin{cases} \pi_1 = 0.375 \\ \pi_2 = 0.5 \\ \pi_3 = 0.125 \end{cases}$$

因此，晴天、雾天和雨天所占的比例分别为 37.5%、50% 和 12.5%。

使用软件求解

需要将方程组中的第三个方程删去，可得：

$$\begin{cases} -\frac{1}{3}\pi_1 + \pi_3 = 0 \\ \frac{1}{3}\pi_1 - \frac{1}{4}\pi_2 = 0 \\ \frac{1}{4}\pi_2 - \pi_3 = 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \Rightarrow \begin{cases} -\frac{1}{3}\pi_1 + \pi_3 = 0 \\ \frac{1}{3}\pi_1 - \frac{1}{4}\pi_2 = 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

删减后的方程组的矩阵-向量形式如下：

$$\underbrace{\begin{bmatrix} \pi_1 & \pi_2 & \pi_3 \end{bmatrix}}_{\boldsymbol{\pi}} \underbrace{\begin{bmatrix} -\frac{1}{3} & \frac{1}{3} & 1 \\ 0 & -\frac{1}{4} & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}}_{\mathbf{b}} \Rightarrow \boldsymbol{\pi} \mathbf{A} = \mathbf{b}$$

使用软件求解 (cont.)

对矩阵-向量形式的方程进行运算, 可得:

$$\pi = \mathbf{bA}^{-1}$$

将数值代入计算, 可得:

$$\pi = \mathbf{bA}^{-1} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} & 1 \\ 0 & -\frac{1}{4} & 1 \\ 1 & 0 & 1 \end{bmatrix}^{-1}$$

π 的结果刚好就是对应的 \mathbf{A}^{-1} 的最后一行。

$$\pi = \begin{bmatrix} 0.375 & 0.5 & 0.125 \end{bmatrix}$$

细致平衡条件

连续时间马氏链，若对 $\forall j \neq k$ ，均有：

$$\pi(k)q(k,j) = \pi(j)q(j,k)$$

则称 π 满足细致平衡条件。

对比：

离散时间马氏链的细致平衡条件：

$$\pi(k)p(k,j) = \pi(j)p(j,k)$$

细致平衡条件 (cont.)

若连续时间马氏链满足细致平衡条件, 即:

$$\pi(k)q(k,j) = \pi(j)q(j,k), \quad \forall j \neq k$$

则 π 是一个平稳分布。

可以根据此定理, 利用细致平衡条件, 进而求得平稳分布 π , 从而不必通过 $\pi Q = 0$ 来求平稳分布。

举例 8：生灭链

考虑状态空间为 $S = \{0, 1, \dots, N\}$ 的生灭链，其转移速率如下：

$$\begin{cases} q(n, n+1) = \lambda_n, & n < N \\ q(n, n-1) = \mu_n, & n > 0 \end{cases}$$

求其平稳分布。

生灭链 (cont.)

根据细致平衡条件, 可得:

$$\pi(n)q(n, n+1) = \pi(n+1)q(n+1, n)$$

$$\pi(n)\lambda_n = \pi(n+1)\mu_{n+1}$$

$$\frac{\pi(n+1)}{\pi(n)} = \frac{\lambda_n}{\mu_{n+1}}$$

因此:

$$\frac{\pi(n+1)}{\pi(n)} \cdot \frac{\pi(n)}{\pi(n-1)} \cdots \frac{\pi(1)}{\pi(0)} = \frac{\lambda_n}{\mu_{n+1}} \cdot \frac{\lambda_{n-1}}{\mu_n} \cdots \frac{\lambda_0}{\mu_1}$$

$$\pi(n+1) = \frac{\lambda_n \cdot \lambda_{n-1} \cdots \lambda_0}{\mu_{n+1} \cdot \mu_n \cdots \mu_1} \pi(0)$$

从而:

$$\pi(n) = \frac{\lambda_{n-1} \cdot \lambda_{n-2} \cdots \lambda_0}{\mu_n \cdot \mu_{n-1} \cdots \mu_1} \cdot \pi(0), \quad 0 < n < N$$

其中, $\pi(n)$ 是生灭链中状态 n 的平稳概率。

举例 9: $M/M/\infty$ 排队系统

状态 $S = \{0, 1, \dots\}$, 其转移速率如下:

$$q(n, n+1) = \lambda, \quad q(n, n-1) = n\mu$$

求其平稳分布。

思路:

$M/M/s$ 排队系统 ($s \rightarrow \infty$) 可看作生灭过程的特殊形式。

$$\mu_n = n\mu, \quad \lambda_0 = \lambda_1 = \dots = \lambda$$

$M/M/\infty$ 排队系统 (cont.)

根据生灭链的公式可得：

$$\pi(n) = \frac{\lambda_{n-1} \cdot \lambda_{n-2} \cdots \lambda_0}{\mu_n \cdot \mu_{n-1} \cdots \mu_1} \pi(0) = \frac{\lambda^n}{n! \mu^n} \pi(0) = \frac{(\lambda/\mu)^n}{n!} \pi(0)$$

进一步地，由于概率要满足完备性，因此：

$$\sum_{n=0}^{\infty} \pi(n) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n! \mu^n} \pi(0) = \pi(0) \cdot \sum_{n=0}^{\infty} \frac{(\lambda/\mu)^n}{n!} = 1$$

因此：

$$\pi(0) = e^{-\lambda/\mu}$$

最终可得：

$$\pi(n) = \frac{(\lambda/\mu)^n}{n!} \cdot e^{-\lambda/\mu}$$

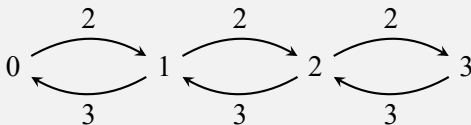
由此可见， $M/M/\infty$ 排队系统的平稳分布 π 服从均值为 λ/μ 的泊松分布。

举例 10: 理发店问题

理发店一名理发师理发的速率为 3 (这里以每小时的顾客数为单位, 即每位顾客的理发时间服从均值为 20 分钟的指数分布), 假设顾客按照一个速率为 2 的泊松过程到达, 但是, 如果顾客到达时等候室的两把椅子都已坐满, 那么他将离开。

求其平稳分布。

思路: 转移速率图



理发店问题 (cont.)

该问题中的状态空间 $S = \{0, 1, 2, 3\}$, 其对应的转移速率如下:

$$\begin{cases} q(n, n+1) = 2, & n = 0, 1, 2 \\ q(n, n-1) = 3, & n = 1, 2, 3 \end{cases}$$

根据细致平衡条件 $\pi(n)q(n, n+1) = \pi(n+1)q(n+1, n)$, 有:

$$\begin{cases} \pi(0)q(0, 1) = \pi(1)q(1, 0) \\ \pi(1)q(1, 2) = \pi(2)q(2, 1) \\ \pi(2)q(2, 3) = \pi(3)q(3, 2) \end{cases} \Rightarrow \begin{cases} 2\pi(0) = 3\pi(1) \\ 2\pi(1) = 3\pi(2) \\ 2\pi(2) = 3\pi(3) \end{cases}$$

结合 $\pi(0) + \pi(1) + \pi(2) + \pi(3) = 1$, 可得:

$$\pi(0) = \frac{27}{65}, \quad \pi(1) = \frac{18}{65}, \quad \pi(2) = \frac{12}{65}, \quad \pi(3) = \frac{8}{65}$$

理发店问题 (cont.)

若采用 $\pi Q = 0$ 求解, 过程如下:

$$Q = \begin{bmatrix} -2 & 2 & 0 & 0 \\ 3 & -5 & 2 & 0 \\ 0 & 3 & -5 & 2 \\ 0 & 0 & 3 & -3 \end{bmatrix}$$

从而得到:

$$\begin{cases} -2\pi_0 + 3\pi_1 = 0 \\ 2\pi_0 - 5\pi_1 + 3\pi_2 = 0 \\ 2\pi_1 - 5\pi_2 + 3\pi_3 = 0 \\ 2\pi_2 - 3\pi_3 = 0 \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \Rightarrow \begin{cases} \pi_0 = \frac{27}{65} \\ \pi_1 = \frac{18}{65} \\ \pi_2 = \frac{12}{65} \\ \pi_3 = \frac{8}{65} \end{cases}$$

嵌入链

嵌入链 (embedded chain) 是对连续时间马氏链的另一种描述方式。

定义

对于转移速率矩阵为 Q 的连续时间马氏链而言, 满足如下条件的 $r(i, j)$ 所构成的就是对应的嵌入链。

$$r(i, j) = \begin{cases} \frac{q(i, j)}{\sum_{i \neq j} q(i, j)} = \frac{q(i, j)}{q(i)}, & i \neq j \\ 0, & i = j \end{cases}$$

其中, $q(i, j)$ 是转移速率矩阵 Q 的对应元素, $q(i)$ 是离开状态 i 的速率, 并且 $q(i) = -q(i, i)$ 。

理发店问题回顾

状态空间 $S = \{0, 1, 2, 3\}$ 对应的转移速率矩阵 \mathbf{Q} 如下:

$$\mathbf{Q} = \begin{bmatrix} -2 & 2 & 0 & 0 \\ 3 & -5 & 2 & 0 \\ 0 & 3 & -5 & 2 \\ 0 & 0 & 3 & -3 \end{bmatrix}$$

根据嵌入链的定义, 可以得到对应的嵌入链转移概率矩阵 \mathbf{R} 如下:

$$\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{3}{5} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{3}{5} & 0 & \frac{2}{5} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

理发店问题 (cont.)

$$\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{3}{5} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{3}{5} & 0 & \frac{2}{5} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

这个矩阵反映了由状态 i 转移到可能的状态 j 的概率。该矩阵的主对角线元素均为零，说明嵌入链的转移概率不考虑状态在原地不变的情形。

注意：

这里转移概率的取值是假设转移服从速率为 $q(i,j)$ 的指数分布，并利用指数分布的性质得到的。

理发店问题 (cont.)

$$\mathbf{Q} = \begin{bmatrix} -2 & 2 & 0 & 0 \\ 3 & -5 & 2 & 0 \\ 0 & 3 & -5 & 2 \\ 0 & 0 & 3 & -3 \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{3}{5} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{3}{5} & 0 & \frac{2}{5} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

从状态 1 转移到状态 0 的速率为 3, 转移到状态 2 的速率为 2, 因此从状态 1 首先转移到状态 0 的概率为:

$$\mathbb{P}(X_\tau = 0 | X_0 = 1) = \mathbb{P}[T_1 = \min(T_1, T_2)] = \frac{3}{2+3} = \frac{3}{5}$$

类似地, 从状态 1 首先转移到状态 2 的概率为:

$$\mathbb{P}(X_\tau = 2 | X_0 = 1) = \mathbb{P}[T_2 = \min(T_1, T_2)] = \frac{2}{2+3} = \frac{2}{5}$$

泊松过程的嵌入链

对于泊松过程，其转移速率满足 $q(i, i+1) = \lambda$ ，并且 $q(i, i) = -\lambda$ ，因此：

$$r(i, i) = 0, \quad r(i, i+1) = 1$$

相应的转移速率矩阵与嵌入链的转移概率矩阵分别如下：

$$\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & 0 & 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & 0 & 0 & -\lambda & \lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

有吸收态的嵌入链

$$\mathbf{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 10 & -20 & 10 & 0 & 0 \\ 0 & 10 & -30 & 20 & 0 \\ 0 & 0 & 10 & -40 & 30 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

对于有吸收态 i 的连续时间马氏链, $q(i, i) = 0$ 。同时, 与之对应的嵌入链转移概率 $r(i, i) = 1$ 。

上面的转移速率矩阵 \mathbf{Q} 有两个吸收态 1 和 5。于是对应的嵌入链转移概率矩阵 \mathbf{R} 如下:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 2/3 & 0 \\ 0 & 0 & 1/4 & 0 & 3/4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

嵌入链转移概率的特点

嵌入链的转移概率与连续时间马氏链的转移概率的不同之处在于，前者并未考虑时间因素，只关心状态与状态之间的转移概率。

嵌入链

根据 $\pi Q = 0$ 可得:

$$\sum_i \pi(i)q(i,j) = 0 \quad \Rightarrow \quad \sum_{i \neq j} \pi(i)q(i,j) = \pi(j)q(j)$$

定义 $\psi(i) = \pi(i)q(i)$, 将其分别代入上式的左右两侧, 可得:

$$\begin{cases} \sum_{i \neq j} \pi(i)q(i,j) = \sum_{i \neq j} \psi(i) \frac{q(i,j)}{q(i)} = \sum_{i \neq j} \psi(i)r(i,j) \\ \pi(j)q(j) = \psi(j) \end{cases}$$

因此:

$$\psi(j) = \sum_{i \neq j} \psi(i)r(i,j)$$

其中, $r(i,j)$ 是嵌入链对应的转移概率, 反映了连续时间马氏链状态之间转移的概率。

矩阵-向量形式

$$\psi(j) = \sum_{i \neq j} \psi(i) r(i, j)$$

上式的矩阵-向量形式如下：

$$\psi \mathbf{R} = \psi$$

其中， ψ 是嵌入链的平稳分布。

注意：

可根据该性质，通过嵌入链 \mathbf{R} 计算连续时间马氏链的平稳分布 π 。

通过嵌入链计算平稳分布的具体步骤

- 基于嵌入链 \mathbf{R} , 利用 $\psi\mathbf{R} = \psi$ 计算出嵌入链的平稳分布 ψ :

$$\psi = [\psi(1) \quad \psi(2) \quad \cdots \quad \psi(n)]$$

- 将得到的嵌入链平稳分布中的各元素分别除以对应的转移速率矩阵主对角线元素的绝对值, 即 $\psi(i)/q(i)$, 从而得到向量 $\tilde{\pi}$, 即:

$$\tilde{\pi} = \left[\frac{\psi(1)}{q(1)} \quad \frac{\psi(2)}{q(2)} \quad \cdots \quad \frac{\psi(n)}{q(n)} \right]$$

- 最后对向量 $\tilde{\pi}$ 进行归一化 (normalize) 操作, 得到连续时间马氏链的平稳分布 π , 即:

$$\pi = \frac{1}{A} \tilde{\pi} = \frac{1}{A} \left[\frac{\psi(1)}{q(1)} \quad \frac{\psi(2)}{q(2)} \quad \cdots \quad \frac{\psi(n)}{q(n)} \right]$$

其中, A 是归一化因子, 并且 $A = \sum_{k=1}^n \frac{\psi(k)}{q(k)}$.

$\pi(i)$ 与 $\psi(i)$ 的关系

连续时间马氏链的平稳分布概率 $\pi(i)$ 衡量的是，长期看来过程花费在状态 i 上的时间占总时间的比重；而嵌入链的平稳分布概率 $\psi(i)$ 衡量的则是，长期看来转移至状态 i 的次数占总的转移次数的比重。

简言之， π 关注花费在各状态上的时间所占的比重； ψ 关注各状态转移的次数所占的比重。

举例：

考虑一个三状态连续时间马氏链，其转移速率矩阵如下：

$$\mathbf{Q} = \begin{bmatrix} -2 & 1 & 1 \\ 2 & -4 & 2 \\ 4 & 4 & -8 \end{bmatrix}$$

求该链的平稳分布 π 。

解法一

使用 $\pi Q = 0$ 来求解，可以列出如下方程组：

$$\begin{cases} -2\pi(1) + 2\pi(2) + 4\pi(3) = 0 \\ \pi(1) - 4\pi(2) + 4\pi(3) = 0 \\ \pi(1) + \pi(2) + \pi(3) = 1 \end{cases} \Rightarrow \begin{cases} \pi(1) = \frac{4}{7} \\ \pi(2) = \frac{2}{7} \\ \pi(3) = \frac{1}{7} \end{cases}$$

解法二

由转移速率矩阵 \mathbf{Q} , 可得到对应的嵌入链 \mathbf{R} 。

$$\mathbf{R} = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix}$$

由于该链是一个双随机链, 因此:

$$\psi = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

从而:

$$\tilde{\pi} = \begin{bmatrix} \frac{\psi(1)}{q(1)} & \frac{\psi(2)}{q(2)} & \frac{\psi(3)}{q(3)} \end{bmatrix} = \begin{bmatrix} \frac{1/3}{2} & \frac{1/3}{4} & \frac{1/3}{8} \end{bmatrix} = \begin{bmatrix} \frac{1}{6} & \frac{1}{12} & \frac{1}{24} \end{bmatrix}$$

解法二 (cont.)

相应的归一化因子为：

$$A = \frac{1}{6} + \frac{1}{12} + \frac{1}{24} = \frac{7}{24}$$

因此：

$$\pi = \frac{1}{A} \tilde{\pi} = \frac{24}{7} \begin{bmatrix} \frac{1}{6} & \frac{1}{12} & \frac{1}{24} \end{bmatrix} = \begin{bmatrix} \frac{4}{7} & \frac{2}{7} & \frac{1}{7} \end{bmatrix}$$

不难看出，两种解法得到的结果完全相同。