第 25 讲 Kolmogorov 方程

回忆 K-C 方程是

$$p_{ij}(t+s) = \sum_{k\in I} p_{ik}(t)p_{kj}(s),\ \ i,j\in I.$$

让我们约定把右边称为前，把左边称为后. 在上式两边对于后面的 t 在 $t=0$ 处求导数，形式上得到

$$p'_{ij}(s) = \sum_{k\in I} q_{ik}p_{kj}(s).$$

写成矩阵的形式就得到 $P'(t) = QP(t)$. 如果对于 s 在 $s=0$ 处求导，形式上得到

$$p'_{ij}(t) = \sum_{k\in I} p_{ik}(t)q_{kj}.$$

写成矩阵的形式就得到 $P'(s) = P(s)Q$.

定理 1.1

设 Q 是连续时间马氏链 $\{X(t)\}$ 的转移速率矩阵，$q_i = |q_{ii}|$，则有

(1) $P'(t) = QP(t)$，或等价地写成

$$p'_{ij}(t) = \sum_{k \in I} q_{ik} p_{kj}(t), \quad i, j \in I;$$

(2) 当 $q = \sup\{q_i | i \in I\} < \infty$ 时，有 $P'(t) = P(t)Q$，或等价地写成

$$p'_{ij}(t) = \sum_{k \in I} p_{ik}(t) q_{kj}, \quad i, j \in I.$$

在 K-C 方程中, 质点先从 i 经过时间 t 到 k, 再从 k 经过时间 s 到 j. 对于后面的时间 t 求导得到向后方程, 对前面的时间 s 求导得到向前方程, 这是向后, 向前的含义.

设右连续函数 $g(t)$ 满足

$$g(t+s)=g(t)g(s)>0,\ \ s,t>0,$$

则 $g(t)$ 是指数函数，即

$$g(t)=\sum_{k=0}^{\infty}\frac{1}{k!}(ct)^k=e^{ct},\ \ g'(t)=cg(t),\ \ c=g'(0).$$

于是得到

$$g(t)=\sum_{k=0}^{\infty}\frac{1}{k!}(g'(0)t)^k=\exp(g'(0)t).$$

对于方阵 A，当 $\sum_{k=0}^{\infty}A^k/k!$ 的每个元收敛时，定义

$$e^A=\exp(A)=\sum_{k=0}^{\infty}\frac{1}{k!}A^k,\ \ A^0=\text{单位矩阵}.$$

对于马氏链的转移概率矩阵 $P(t)$ 和转移速率矩阵 Q，由 K-C 方程我们得

$$P(t+s) = P(t)P(s), \ \ s,t \geq 0,$$

所以我们也期望有某个矩阵 C，使得

$$P(t) = \sum_{k=0}^{\infty} \frac{1}{k!}(Ct)^k \equiv \exp(Ct),$$

其中 $(Ct)^0$ 表示单位阵. 因为 $P'(0) = Q$，所以猜测应当有 $C = Q$ 和

$$P(t) = \sum_{k=0}^{\infty} \frac{1}{k!}(Qt)^k \equiv \exp(Qt). \tag{2.1}$$

如果马氏链 $\{X(t)\}$ 的状态空间 I 是有限集合，则称 $\{X(t)\}$ 是有限状态马氏链. 设状态空间 I 有 n 个状态，取

$$q = \max_{i,j \in I} |q_{ij}| = \max_{i \in I} q_i,$$

则矩阵 $(Qt)^k$ 的元素绝对值都小于 $n^{k-1}(qt)^k$. 于是

$$\exp(Qt) \equiv \sum_{k=0}^{\infty} \frac{1}{k!}(Qt)^k$$

的每个元收敛，关于 t 逐项求导得到

$$\begin{aligned}(\exp(Qt))' &= (\sum_{k=0}^{\infty} \frac{1}{k!}(Qt)^k)' \\ &= \sum_{k=1}^{\infty} \frac{1}{(k-1)!}(Qt)^{k-1}Q \\ &= \exp(Qt)Q, \ t \in (-\infty, \infty).\end{aligned}$$

定理 2.1

有限状态连续时间马氏链转移概率矩阵 $P(t)$ 由转移概率矩阵 Q 唯一决定，并且有

$$P(t) = \sum_{k=0}^{\infty} \frac{1}{k!}(Qt)^k \equiv \exp(Qt), t \geq 0.$$

证明.

由方程 (2.1) 我们知 $P(t)=\exp(Qt)$ 满足 $P'(t)=\exp(Qt)Q=P(t)Q$，所以是向前方程的解. 下面证明它是向前方程的唯一解.
如果 $P(t)$ 满足向前方程 $P'(t)=P(t)Q$，且 $P(0)$ 是单位阵，我们证明 $P(t)=\exp(Qt)$. 利用公式

$$(A(t)B(t)' = A'(t)B(t) + A(t)B'(t)$$

和向前方程我们有

$$\begin{aligned}(P(t)\exp(-Qt))' &= P'(t)\exp(-Qt) - P(t)Q\exp(-Qt)\\ &= (P'(t)-P(t)Q)\exp(-Qt) = 0.\end{aligned}$$

这说明 $P(t)\exp(-Qt)=C$ 是常数矩阵. 再由 $P(0)=$ 单位矩阵 我们有

$$单位阵 = P(t)\exp(-Qt).$$

证明.

两边右乘 $\exp(Qt)$，得到

$$\begin{aligned}
\exp(Qt) &= P(t)\exp(-Qt)\exp(Qt) \\
&= P(t)\sum_{k=0}^{\infty}\frac{1}{k!}(-Qt)^k\sum_{j=0}^{\infty}\frac{1}{j!}(Qt)^j \\
&= P(t)\sum_{j=0}^{\infty}\sum_{k=0}^{j}\frac{1}{k!(j-k)!}t^k(-t)^{j-k}Q^j \\
&= P(t)\sum_{j=0}^{\infty}\frac{1}{j!}\Big(\sum_{k=0}^{j}C_j^k t^k(-t)^{j-k}\Big)Q^j \\
&= P(t)\sum_{j=0}^{\infty}j!(t-t)^jQ^j \\
&= P(t).
\end{aligned}$$

这就证明了向前方程有唯一解

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

Q 唯一决定了 $P(t)$. $\square$

例 2.2

给定连续时间马氏链的转移速率矩阵

$$Q = P'(0) = \frac{1}{5}\begin{pmatrix} -9 & 3 & 6 \\ 6 & -12 & 6 \\ 6 & 3 & -9 \end{pmatrix}$$

计算转移概率矩阵.

证明.

我们先将 Q 写成若当标准型

$$Q = MDM^{-1},$$

其中 D 为 Q 的若当标准型，M 为可逆矩阵，用线性代数的方法我们可以求得：

$$D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \; M = \begin{pmatrix} 0 & 1 & -1 \\ 2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}, \; M^{-1} = \frac{1}{5}\begin{pmatrix} -1 & 2 & -1 \\ 2 & 1 & 2 \\ -3 & 1 & 2 \end{pmatrix}.$$

这样得到 $Q^k = MD^kM^{-1}$，于是得到

$$\begin{aligned} P(t) &= \sum_{k=0}^{\infty} \frac{1}{k!}(Qt)^k = M \sum_{k=0}^{\infty} \frac{1}{k!}(Qt)^k M^{-1} \\ &= M \begin{pmatrix} e^{-3t} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-3t} \end{pmatrix} M^{-1} \\ &= \frac{1}{5}\begin{pmatrix} 2+3e^{-3t} & 1-e^{-3t} & 2-2e^{-3t} \\ 2-2e^{-3t} & 1+4e^{-3t} & 2-2e^{-3t} \\ 2-2e^{-3t} & 1-e^{-3t} & 2+3e^{-3t} \end{pmatrix}. \end{aligned}$$

□

定理 2.3

马氏链的转移概率 $p_{ij}(t)$ 满足柯尔莫哥洛夫向前和向后方程，而且是向后或向前方程的唯一解.

上面定理的唯一性是说, 如果有另外的非负函数 $g_{ij}(t)$ 满足相同的向后方程(或向前方程), 并且满足 K-C 方程
$g_{ij}(s+t) = \sum_{k\in I} g_{ik}(s)g_{kj}(t)\ (i,j\in I,\ s,t\geq 0)$ 和

$$\sum_{j\in I} g_{ij}(t) \leq 1,\ g_{ij}(0) = \lim_{t\downarrow 0} g_{ij}(t) = \delta_{ij},\ i,j\in I,\ t\geq 0,$$

则 $g_{ij}(t) = p_{ij}(t)$.

例 2.4

马氏链只有 $1, 0$ 两个状态, 在状态 1 和 0 的停留时间分别是来自总体 $\mathcal{E}(\lambda)$, $\mathcal{E}(\mu)$ 的随机变量.

(1) 写出嵌入链的转移概率矩阵 K 和 Q 矩阵;

(2) 写出 p_{00} 和 p_{11} 满足的向前方程;

(3) 计算 $p_{00}(t)$, $p_{11}(t)$.

解.

(1) 对于连续时间马尔可夫链，从状态 i 离开的速率 q_i 等于其停留时间指数分布的参数. 在状态 1 的停留时间服从 $\mathcal{E}(\lambda)$，所以 $q_1 = \lambda$。这意味着对角线元素 $q_{11} = -q_1 = -\lambda$。在状态 0 的停留时间服从 $\mathcal{E}(\mu)$，所以 $q_0 = \mu$。这意味着对角线元素 $q_{00} = -q_0 = -\mu$。因为总共有两个状态，离开状态 1 只能转移到状态 0，所以转移速率 $q_{10} = \lambda$；同理，离开状态 0 只能转移到状态 1，所以 $q_{01} = \mu$.

因此，Q 矩阵为:

$$Q = \begin{pmatrix} q_{00} & q_{01} \\ q_{10} & q_{11} \end{pmatrix} = \begin{pmatrix} -\mu & \mu \\ \lambda & -\lambda \end{pmatrix}$$

系统只要离开当前状态，就必须跳跃到另一个状态，不存在原地跳跃 $(k_{00} = 0, k_{11} = 0)$。因此，嵌入链的转移概率矩阵 K 为:

$$K = \begin{pmatrix} k_{00} & k_{01} \\ k_{10} & k_{11} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

解.

(2) Kolmogorov 向前微分方程的形式为 $\frac{\mathrm{d}}{\mathrm{d}t}P(t) = P(t)Q$.

1. 对于 $p_{00}(t)$ 的向前方程（对应 $P(t)Q$ 的第 1 行第 1 列）:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{00}(t) = p_{00}(t)q_{00} + p_{01}(t)q_{10}$$

代入 $q_{00} = -\mu$，$q_{10} = \lambda$，并利用全概率公式 $p_{01}(t) = 1 - p_{00}(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{00}(t) = -\mu p_{00}(t) + \lambda(1 - p_{00}(t))$$

简化得:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{00}(t) = \lambda - (\lambda + \mu)p_{00}(t).$$

2. 对于 $p_{11}(t)$ 的向前方程（对应 $P(t)Q$ 的第 2 行第 2 列）:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{11}(t) = p_{10}(t)q_{01} + p_{11}(t)q_{11}$$

代入 $q_{01} = \mu$，$q_{11} = -\lambda$，并利用全概率公式 $p_{10}(t) = 1 - p_{11}(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{11}(t) = \mu(1 - p_{11}(t)) - \lambda p_{11}(t)$$

解.

(3) 利用一阶常系数线性微分方程的求解方法，结合各自的初始条件进行计算。

1. 求解 $p_{00}(t)$ 微分方程为：

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{00}(t) + (\lambda+\mu)p_{00}(t) = \lambda$$

其通解形式为：$p_{00}(t) = \frac{\lambda}{\lambda+\mu} + C_1 e^{-(\lambda+\mu)t}$，代入初始条件 $p_{00}(0) = 1$：

$$1 = \frac{\lambda}{\lambda+\mu} + C_1 \implies C_1 = \frac{\mu}{\lambda+\mu}$$

所以：

$$p_{00}(t) = \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}e^{-(\lambda+\mu)t}.$$

解.

2. 求解 $p_{11}(t)$ 微分方程为:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{11}(t)+(\lambda+\mu)p_{11}(t)=\mu$$

其通解形式为:

$$p_{11}(t)=\frac{\mu}{\lambda+\mu}+C_2e^{-(\lambda+\mu)t}$$

代入初始条件 $p_{11}(0)=1$:

$$1=\frac{\mu}{\lambda+\mu}+C_2 \implies C_2=\frac{\lambda}{\lambda+\mu}$$

所以:

$$p_{11}(t)=\frac{\mu}{\lambda+\mu}+\frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}.$$

□