

第 16 讲 马尔可夫链

定义 1.1

设 $\{X_n\}$ 是一个在 I 中取值的随机序列, 如果对任意正整数 n 和 $i, j, i_0, i_1, \dots, i_{n-1} \in I$, 有

$$\begin{aligned} & P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\ &= P(X_{n+1} = j | X_n = i) \\ &= P(X_1 = j | X_0 = i), \quad i, j \in I, \end{aligned}$$

则称 $\{X_n\}$ 为时齐的马尔可夫链, 简称为马氏链. 这时称

$$p_{ij} = P(X_1 = j | X_0 = i), \quad i, j \in I$$

为马氏链 $\{X_n\}$ 的转移概率, 称矩阵

$$P = (p_{ij}) = (p_{ij})_{i,j \in I}$$

为马氏链 $\{X_n\}$ 的一步转移概率矩阵, 简称为转移矩阵.

由于 $\{X_1 = j\}(j \in I)$ 是完备事件组, 所以有

$$\sum_{j \in I} p_{ij} = \sum_{j \in I} P(X_1 = j | X_0 = i) = P\left(\bigcup_{j \in I} \{X_1 = j\} | X_0 = i\right) = 1.$$

于是转移矩阵 P 的各行之和等于 1, 这样的矩阵也被称为随机矩阵.

对于马氏链的直观理解是：已知现在的状态，将来与过去相独立. 人们习惯地称这种性质为 马氏性.

定理 2.1

对于事件 A, B, C , 当 $P(AB) > 0$, 条件

$$P(C|BA) = P(C|B)$$

与条件

$$P(AC|B) = P(A|B)P(C|B)$$

等价.

证明.

引入条件概率 $P_B(\cdot) = P(\cdot|B)$, 我们知上面两个等式分别等价于 $P_B(C|A) = P_B(C)$ 和 $P_B(AC) = P_B(A)P_B(C)$. 这两个公式都表示对于概率 P_B 来说, A 和 C 独立, 所以它们等价. \square

根据上述定理, $P(C|BA) = P(C|B)$ 等价于已知 B 发生时, C 和 A 独立. 于是, $\{X_n\}$ 有马氏性的充分必要条件是对于任意 $n \geq 1$, 已知现在 $B = \{X_n = i\}$, 将来 $C = \{X_{n+1} = j\}$ 与过去 $A = \{X_{n-1} = i_{n-1}, \dots, X_0 = i_0\}$ 独立.

例 2.2: 两状态马尔可夫链

状态空间为 $\{0, 1\}$, 转移概率为 $p_{01} = 1 - p_{00} = p$, $p_{10} = 1 - p_{11} = q$. 转移矩阵可以表示如下:

$$\begin{matrix} & \begin{matrix} 0 & 1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \end{matrix}$$

该马尔可夫链可以用来作为一个描述电话状态的简单模型, 其中 $X_n = 0$ 表明电话在时刻 n 空闲, $X_n = 1$ 表明电话在时刻 n 繁忙. 我们假设在每个时间间隔内有一个电话打进的概率为 p (为了方便起见, 假定在任意一个特定时间的的时间间隔内至多有一个电话打进). 当电话繁忙时, 来到的呼叫无法进入系统. 我们还假设前一段时间间隔内忙的电话在下一时间间隔内空闲的概率为 q .

例 2.3

设想一个质点在直线的整数点上作简单随机游动：质点一旦到达某状态后，下次向右移动一步的概率是 p ，向左移动一步的概率是 $q = 1 - p$ ， $pq > 0$ 。现在用 X_0 表示质点的初始状态，用 X_n 表示质点在时刻 n 的状态，则 $\{X_n\}$ 是马氏链，并且

$$\begin{cases} p_{i,i-1} = P(X_{n+1} = i - 1 | X_n = i) = q, \\ p_{i,i+1} = P(X_{n+1} = i + 1 | X_n = i) = p. \end{cases}$$

例 2.4: 两端是吸收壁的简单随机游动

设质点在状态 $\{1, 2, \dots, n-1\}$ 中按上例的规律做简单随机游动, 但是质点一旦到达状态 n 或状态 0 后将永远停留在 n 或 0 . 用 X_0 表示质点的初始状态, 用 X_n 表示质点在时刻 n 的状态, 则 $\{X_n\}$ 是马氏链, 并且

$$p_{ij} = \begin{cases} q, & 1 \leq i \leq n-1, j = i-1, \\ p, & 1 \leq i \leq n-1, j = i+1, \\ 1, & (i, j) = (0, 0) \text{ 或 } (i, j) = (n, n), \\ 0, & \text{其他.} \end{cases}$$

相应的转移概率矩阵是

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ q & 0 & p & 0 & 0 & \cdots & 0 \\ 0 & q & 0 & p & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

例 2.5: 两端是反射壁的简单随机游动

设质点在状态 $\{1, 2, \dots, n-1\}$ 中按上例的规律做简单随机游动, 但是质点到达状态 n 后下一步一定返回 $n-1$, 到达状态 0 后下一步一定返回状态 1 . 用 X_0 表示质点的初始状态, 用 X_n 表示质点在时刻 n 的状态, 则 $\{X_n\}$ 是马氏链, 并且

$$p_{ij} = \begin{cases} q, & 1 \leq i \leq n-1, j = i-1, \\ p, & 1 \leq i \leq n-1, j = i+1, \\ 1, & (i, j) = (0, 1) \text{ 或 } (i, j) = (n, n-1), \\ 0, & \text{其他.} \end{cases}$$

相应的转移概率矩阵是

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

下面的定理提供了一个非常有用的获得马尔可夫链的方法，并可用于检验一随机过程是否为马尔可夫链.

定理 2.6

设随机过程 $\{X_n : n \geq 0\}$ 满足:

- (1) $X_n = f(X_{n-1}, \xi_n) (n \geq 1)$, 其中 $f : S \times S \rightarrow S$, 且 ξ_n 取值在 S 上,
- (2) $\{\xi_n : n \geq 1\}$ 为独立同分布随机变量, 且 X_0 与 $\{\xi_n : n \geq 1\}$ 也相互独立,

则 $\{X_n, n \geq 0\}$ 是马尔可夫链, 而且其一步转移概率为

$$p_{ij} = P(f(i, \xi_1) = j).$$

证明.

设 $n \geq 1$, 注意到 ξ_{n+1} 与 X_0, X_1, \dots, X_n 相互独立 (这是因为 X_i 是 X_0, ξ_1, \dots, ξ_i 的函数), 有

$$\begin{aligned} P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) &= P(f(X_n, \xi_{n+1}) = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) \\ &= P(f(i_n, \xi_{n+1}) = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) \\ &= P(f(i_n, \xi_{n+1}) = i_{n+1}). \end{aligned}$$

同样地,

$$P(X_{n+1} = i_{n+1} | X_n = i_n) = P(f(X_n, \xi_{n+1}) = i_{n+1}).$$

因此

$$P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) = P(X_{n+1} = i_{n+1} | X_n = i_n).$$

从而 $\{X_n, n \geq 0\}$ 是马尔可夫链, 且其转移概率为 $p_{ij} = P(f(i, \xi_1) = j)$. □

为了方便地研究马氏链的性质, 有必要回忆条件概率的基本公式. 设 $P(AB) > 0$, 则

$$P_A(C|B) = P(C|BA).$$

如果 $P(A) > 0$, 事件 B_k 互不相容使得 $C \subset \bigcup_{k=1}^{\infty} B_k$, 则有全概率公式

$$P(C|A) = \sum_{k=1}^{\infty} P(B_k|A)P(C|B_kA).$$

定理 2.7

设 I 是马氏链 $\{X_n\}$ 的状态空间, $A, A_j \subset I$, 则有

- (1) 已知 $X_n = i$ 的条件下, 将来 $\{X_m; m \geq n + 1\}$ 与过去 $\{X_j; j \leq n - 1\}$ 独立;
- (2) $P(X_{n+k} = j | X_n = i) = P(X_k = j | X_0 = i)$;
- (3) $P(X_{n+k} = j | X_n = i, X_{n-1} \in A_{n-1}, \dots, X_0 \in A_0) = P(X_k = j | X_0 = i)$;
- (4) $P(X_{n+k} \in A | X_n = i, X_{n-1} \in A_{n-1}, \dots, X_0 \in A_0) = P(X_k \in A | X_0 = i)$.

证明.

对于任何 $m > n \geq 1$ 和 I 中的 i_0, i_1, \dots, i_m , 引入 $B_k = \{X_k = i_k\}$, $k = 0, 1, \dots, m$, 则有

$$\begin{aligned} A &:= \{X_m = i_m, X_{m-1} = i_{m-1}, \dots, X_{n+1} = i_{n+1}\} \\ &= B_m B_{m-1} \cdots B_{n+1}, \end{aligned}$$

$$\begin{aligned} C &:= \{X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \dots, X_0 = i_0\} \\ &= B_{n-1} B_{n-2} \cdots B_0, \end{aligned}$$

下面证明已知现在 $B = \{X_n = i_n\} = B_n$ 后, 将来 A 与过去 C 独立. 用乘法公式和马氏性得到

$$\begin{aligned} P(ABC) &= P(B_0 B_1 \cdots B_m) \\ &= P(B_0) P(B_1 | B_0) P(B_2 | B_1) \cdots P(B_m | B_{m-1}) \\ &= P(B_m | B_{m-1}) P(B_{m-1} | B_{m-2}) \cdots P(B_1 | B_0) P(B_0). \end{aligned}$$

证明.

同理有

$$P(BC) = P(B_0 B_1 \cdots B_n) = P(B_n | B_{n-1}) P(B_{n-1} | B_{n-2}) \cdots P(B_1 | B_0) P(B_0).$$

对 $k > n$ 注意使用 $P(B_{k+1} | B_k) = P(B_{k+1} | B_k B_{k-1} \cdots B_n)$ 就得到

$$\begin{aligned} P(A|BC) &= \frac{P(ABC)}{P(BC)} \\ &= P(B_n | B_{m-1} \cdots B_n) P(B_{m-1} | B_{m-2} \cdots B_n) \cdots P(B_{n+1} | B_n) \\ &= P(B_m B_{m-1} \cdots B_{n+1} | B_n) = P(A|B). \end{aligned}$$

从而 (1) 成立. □

证明.

下面我们用归纳法来证明 (2). 由马氏链的定义我们知对 $k = 1$ 以及任意 n 结论成立. 设结论对 $k - 1$ 以及任意 n 成立, 用 (1) 中的结论以及全概率公式我们有

$$\begin{aligned} P(X_{n+k} = j | X_n = i) &= \sum_{l \in I} P(X_{n+k} = j | X_{n+k-1} = l, X_n = i) P(X_{n+k-1} = l | X_n = i) \\ &= \sum_{l \in I} P(X_k = j | X_{k-1} = l, X_0 = i) P(X_{k-1} = l | X_0 = i) \\ &= P(X_k = j | X_0 = i). \end{aligned}$$

由 (1) 和 (2) 我们可得 (3) 和 (4). □

对于马氏链 $\{X_n\}$, 我们知道 $P(X_{k+n} = j | X_n = i)$ 和 n 无关, 所以对 $i, j \in I$ 定义

$$P_{ij}^{(0)} = P(X_0 = j | X_0 = i) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

$$P_{ij}^{(k)} = P(X_{n+k} = j | X_n = i) = P(X_k = j | X_0 = i),$$

并且称 $p_{ij}^{(k)}$ 为 $\{X_n\}$ 的 k 步转移概率, 称矩阵

$$p^{(k)} = (p_{ij}^{(k)})$$

为 $\{X_n\}$ 的 k 步转移概率矩阵, 简称为 k 步转移矩阵. 这时,

$$P^{(1)} = P, \quad P^{(0)} = \text{单位阵}.$$

下面的柯尔莫哥洛夫-切普曼 (Kolmogorov-Chapman) 方程 (简称 K-C 方程) 告诉我们从一部转移概率矩阵 P 计算 k 步转移概率矩阵 $P^{(k)}$ 的方法.

定理 3.1: (K-C 方程)

对任意 $m, n \geq 0$, 有

$$\begin{cases} p_{ij}^{(n+m)} = \sum_{k \in I} p_{ik}^{(n)} p_{kj}^{(m)}, \\ P^{(n+m)} = P^{n+m}. \end{cases} \quad (3.1)$$

证明.

由定理 ?? 我们有

$$\begin{aligned}
 p_{ij}^{(n+m)} &= P(X_{n+m} = j | X_0 = i) \\
 &= \sum_{k \in I} P(X_n = k | X_0 = i) P(X_{n+m} = j | X_n = k, X_0 = i) \\
 &= \sum_{k \in I} p_{ik}^{(n)} p_{kj}^{(m)}.
 \end{aligned}$$

写成矩阵的形式, 我们得到 $P^{(n+m)} = P^{(n)}P^{(m)}$. 由 n, m 的任意性得到

$$P^{(n+m)} = P P^{(n+m-1)} = P^2 P^{(n+m-2)} = \dots = P^{n+m}.$$

□

推论 3.2

对于正整数 $n, m, k, n_1, n_2, \dots, n_k$ 和状态 i, j, l 总有

$$(1) p_{ij}^{(n+m)} \geq p_{il}^{(n)} p_{lj}^{(m)};$$

$$(2) p_{ii}^{(n+k+m)} \geq p_{ij}^{(n)} p_{jl}^{(k)} p_{li}^{(m)};$$

$$(3) p_{ii}^{(n_1+n_2+\dots+n_k)} \geq p_{ii}^{(n_1)} p_{ii}^{(n_2)} \dots p_{ii}^{(n_k)};$$

$$(4) p_{ii}^{(nk)} \geq (p_{ii}^{(n)})^k.$$

证明.

利用 K-C 方程我们有

$$p_{ij}^{(n+m)} = \sum_{s \in I} p_{is}^{(n)} p_{sj}^{(m)} \geq p_{il}^{(n)} p_{lj}^{(m)},$$

从而 (1) 得证. 两次利用 (1) 我们有

$$p_{ii}^{(n+k+m)} \geq p_{il}^{(n+k)} p_{li}^{(m)} \geq p_{ij}^{(n)} p_{jl}^{(k)} p_{li}^{(m)},$$

于是 (2) 成立. 在 (1) 中取 $j = i$ 并应用归纳法我们可以得到 (3). 在 (3) 中取 $n_i = n$ 我们就得到 (4). □

例 3.3

考虑两状态的马尔可夫链所描述的电话状态模型, 假定电话在时刻 0 时空闲. 设 $p = \frac{1}{4}$, $q = \frac{1}{6}$, 令 $n = 6$, 则

$$\mathbf{P}^6 = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{6} & \frac{5}{6} \end{pmatrix}^6 = \begin{pmatrix} 0.424 & 0.576 \\ 0.384 & 0.616 \end{pmatrix}.$$

如果电话在 0 时刻空闲, 想要知道在此条件下, 电话在时刻 6 繁忙的概率, 我们可以计算

$$p_{01}^{(6)} = 0.576$$

即可.

设 $\{X_n\}$ 的一步转移概率矩阵是 P , X_0 有概率分布

$$\pi_j = P(X_0 = j), j \in I.$$

称 X_0 的分布列 (认为 $I = \{1, 2, \dots\}$)

$$\boldsymbol{\pi}^{(0)} = [\pi_1, \pi_2, \dots]$$

为 $\{X_n\}$ 的初始分布, 它表明了质点在初始状态的分布情况. 再引入 X_n 的概率分布

$$\pi_j^{(n)} = P(X_n = j), j \in I$$

和

$$\boldsymbol{\pi}^{(n)} = [\pi_1^{(n)}, \pi_2^{(n)}, \dots],$$

则 $\boldsymbol{\pi}^{(n)}$ 表明质点在 n 时刻所处状态的分布.

定理 4.1

设马氏链 $\{X_n\}$ 有初始分布 $\pi^{(0)} = [\pi_1, \pi_2, \dots]$ 和转移概率矩阵 $P = (p_{ij})$.

(1) 对于 $n_0 < n_1 < \dots < n_m$, 有

$$\begin{aligned} & P(X_{n_0} = i_0, X_{n_1} = i_1, \dots, X_{n_m} = i_m) \\ &= \pi_{i_0}^{(n_0)} p_{i_0 i_1}^{(n_1 - n_0)} p_{i_1 i_2}^{(n_2 - n_1)} \dots p_{i_{m-1} i_m}^{(n_m - n_{m-1})}; \end{aligned}$$

(2) 对 $n \geq 1$, 有

$$\begin{cases} \pi_j^{(n)} = \sum_{i \in I} \pi_i p_{ij}^{(n)}, \\ \boldsymbol{\pi}^{(n)} = \boldsymbol{\pi}^{(k)} P^{(n-k)}, \quad 0 \leq k \leq n. \end{cases}$$

证明.

(1) 利用概率的乘法公式我们有

$$\begin{aligned} & P(X_{n_0} = i_0, X_{n_1} = i_1, \dots, X_{n_m} = i_m) \\ &= P(X_{n_0} = i_0)P(X_{n_1} = i_1 | X_{n_0} = i_0) \cdots P(X_{n_m} = i_m | X_{n_{m-1}} = i_{m-1}) \\ &= \pi_{i_0}^{n_0} p_{i_0 i_1}^{(n_1 - n_0)} p_{i_1 i_2}^{(n_2 - n_1)} \cdots p_{i_{m-1} i_m}^{(n_m - n_{m-1})}. \end{aligned}$$

(2) 用全概率公式我们可以得到, 当 $n \geq 1$ 时,

$$\begin{aligned} \pi_j^{(n)} &= P(X_n = j) \\ &= \sum_{i \in I} P(X_0 = i)P(X_n = j | X_0 = i) \\ &= \sum_{i \in I} \pi_i p_{ij}^{(n)}. \end{aligned}$$

写成矩阵的形式就得到

$$\begin{aligned} \boldsymbol{\pi}^{(n)} &= \boldsymbol{\pi}^{(0)} P^{(n)} = [\boldsymbol{\pi}^{(0)} P^{(1)}] P^{(n-1)} \\ &= \boldsymbol{\pi}^{(1)} P^{(n-1)} = \dots = \boldsymbol{\pi}^{(k)} P^{(n-k)}. \end{aligned}$$

□

极限 $\lim_{n \rightarrow \infty} p_{ij}^{(n)}$ 表示质点从 X_0 出发后, 在充分远的将来处于状态 j 的概率.

考虑直线上的简单随机游动. 如果 $p > q$, 无论质点从哪里出发, 随着时间的推移向右越跑越远. 直观上应该有

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0.$$

再考虑两端带有吸收壁的简单随机游动. 状态 $0, n$ 和 $\{1, \dots, n-1\}$ 中的状态有明显的不同. 直观上质点不可能永远在 $\{1, \dots, n-1\}$ 中转移, 所以有

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0, \quad j = 1, 2, \dots, n-1.$$