

由 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j = \frac{1}{\mu_j}$, 故 $\mu_1 = \frac{7}{5}$, $\mu_2 = \frac{7}{2}$, 且

$$\lim_{n \rightarrow \infty} P^n = \lim_{n \rightarrow \infty} \begin{pmatrix} 3/4 & 1/4 \\ 5/8 & 3/8 \end{pmatrix}^n = \begin{pmatrix} 5/7 & 2/7 \\ 5/7 & 2/7 \end{pmatrix}. \quad \square$$

例 2 设 $\{X_n, n \geq 0\}$ 为艾伦费斯特链, $S = \{0, 1, 2, \dots, 2N\}$, 转移概率为

$$p_{i,i+1} = \frac{2N-i}{2N} \quad (0 \leq i \leq 2N-1),$$

$$p_{ii} = 0 \quad (0 \leq i \leq 2N), \quad p_{i,i-1} = \frac{i}{2N} \quad (1 \leq i \leq 2N).$$

求此链的 π_j 及 μ_j ($i \in S$).

解 由 $\pi = \pi P$, 得

$$\pi_i = \frac{2N-i+1}{2N} \pi_{i-1} + \frac{i+1}{2N} \pi_{i+1}, \quad 1 \leq i \leq 2N-1,$$

$$\pi_0 = \frac{\pi_1}{2N}, \quad \pi_{2N} = \frac{\pi_{2N-1}}{2N}.$$

解此方程组得 $\pi_i = C_{2N}^i \pi_0$ ($1 \leq i \leq 2N$). 又因为 $\sum_{i=0}^{2N} \pi_i = 1$, 因此 $\pi_0 = 2^{-2N}$, 于是有 $\pi_i = C_{2N}^i 2^{-2N}$ ($1 \leq i \leq 2N$). 再由 $\mu_i = \frac{1}{\pi_i}$ 得

$$\mu_i = 2^{2N} \frac{i!(2N-i)!}{(2N)!}, \quad 0 \leq i \leq 2N). \quad \square$$

4. 在马氏链蒙特卡罗随机模型方法中的应用

蒙特卡罗 (Monte Carlo) 随机模拟方法如今广泛应用于生物学、化学、信息科学、军事科学、金融经济、工程管理科学、材料科学、现代物理学和统计学等诸多领域, 特别是马尔科夫链蒙特卡罗 (Markov Chain Monte Carlo, MCMC) 方法更是为现代统计建模提供了耐人寻味的途径.

定义 3.5.3 为要模拟服从给定分布 π 的随机变量, 用生成一个易于实现的不可约遍历链 $X = \{X_n, n \geq 0\}$ 作为随机样本, 使其平稳分布为 π 的方法, 称为马氏链蒙特卡罗方法.

蒙特卡罗方法的一个首要步骤是产生服从给定的概率分布函数 π 的随机变量 (或称为随机样本), 由概率论知识, 熟知下面的结论.

引理 3.5.1 生成随机变量 U , 使其分布满足 $U[0, 1]$, 记为 $U \sim U[0, 1]$, $F(x)$ 是给定的一个分布函数, 记 $F^{-1}(y) = \sup\{x: F(x) < y, y \in (0, 1)\}$ 为 $F(x)$ 的反函数, 则 $X = F^{-1}(U)$ 分布函数为 $F(x)$ (参见文献 [22, 27, 28]).

上述结果为产生服从给定分布的一维随机变量提供了一种方法, 但很多分布都没有解析表达式 (例如正态分布), 利用上述结果实际上是行不通的. 冯·诺伊曼 (J. Von Neumann) 在 1951 年提出了如下的巧妙算法.

命题 3.5.1 假设 $\pi(x)$ 是给定的分布函数, 选取一个易于生成的分布函数 $g(x)$ 和常数 $M > 0$, 满足 $\pi(x) \leq Mg(x), \forall x$. 由下列步骤生成随机变量 Z :

(1) 抽取 $Y \sim g(y), U \sim U[0, 1]$;

(2) 若 $U \leq \pi(Y)/(Mg(Y))$, 取 $Z = Y$; 否则返回步骤 (1). 则 $Z \sim \pi(x)$.

此法尤其适合于高维分布的情形 (证明留作练习, 可参见文献 [22, 27, 28]).

应用举例: 计算积分 $\int h(x)d\pi(x) = Eh(Z)$, 其中 $\pi(x)$ 为 Z 的概率分布函数, 只需生成服从 $\pi(x)$ 的 n 个独立样本 Z_1, \dots, Z_n , 若 $Eh(Z)$ 存在, 则对充分大的 n , 取 $\hat{y}_n = \sum_{k=1}^n h(Z_k)/n$ 作为积分 $\int h(x)d\pi(x) = Eh(Z)$ 的近似值 (由大数定律知 $\hat{y}_n \xrightarrow{P} Eh(Z), n \rightarrow \infty$).

然而, 在许多实际问题中, 生成给定分布 π 的独立样本是不可取的, 而 MCMC 方法的基本技巧是生成一个相关的马氏链 $\{X_n, n \geq 0\}$, 使其平稳分布为给定的 π . 由于当今应用问题中涉及的统计模型越来越复杂, 有些难以用已有的解析方法或数值方法解决, 而 MCMC 恰好提供了一个使复杂问题得以有效分析的统一的数学框架, 从而显示其魅力和威力.

米特罗波利斯 (Metropolis) 等人在 1953 年最早给出了通过生成一马氏链实现从分布 π 中采样 (生成相关的样本) 这一重要基本思想. 随后, 哈斯汀 (Hastings) 将其推广到更一般的形式. 下面仅叙述状态空间 S 为至多可数的情形.

Metropolis-Hastings 算法

命题 3.5.2 设 $\pi = (\pi(i), i \in S)$ 为任意给定的概率分布 ($\pi(i) > 0, i \in S$), 而 $T = (T(i, j), i, j \in S)$ 为任选的易于实现的条件概率转移矩阵, ($T(i, j) > 0, i, j \in S$) (称 T 为参照矩阵). $X = \{X_n, n \geq 0\}$ 由下列步骤生成:

(1) 给定 $X_n (X_n \in S, n \geq 0)$. 由 $T(X_n, \cdot)$ 抽取 Y , 并计算

$$\rho(X_n, Y) = \min\{1, \pi(Y)T(Y, X_n)/(\pi(X_n)T(X_n, Y))\};$$

(2) 抽取 $U \sim U[0, 1]$, 若 $U < \rho(X_n, Y)$, 则令 $X_{n+1} = Y$; 否则舍去 Y , 返回

步骤 (1).

则 $X = \{X_n, n \geq 0\}$ 是不可约遍历马氏链, 平稳分布为 π .

证明梗概: 易证 X 是不可约遍历马氏链, 起一步转移概率 $P_{ij} = T(i, j)\rho(i, j)$ ($i, j \in S$) 有对称性, 即 $\pi(i)P_{ij} = \pi(j)P_{ji} (\forall i, j \in S)$. 事实上

$$\begin{aligned}\pi(i)P_{ij} &= \pi(i)T(i, j)\rho(i, j) = \pi(i)T(i, j) \min\{1, \pi(j)T(j, i)/(\pi(i)T(i, j))\} \\ &= \min\{\pi(i)T(i, j), \pi(j)T(j, i)\} = \pi(j)P_{ji}.\end{aligned}$$

上式两边对 $j \in S$ 求和, 得

$$\pi(i) = \sum_{j \in S} \pi(j)P_{ji}. \quad (*)$$

由 (*) 式即证 $\pi = (\pi(i), i \in S)$ 是 X 的平稳分布.

3.6 离散时间的 Phase-Type 分布及其反问题

本节讨论离散时间的 Phase-Type 分布及其反问题, 先给出它的定义.

定义 3.6.1 设 $\{X_n, n \geq 0\}$ 是马尔可夫链, 状态空间 $\tilde{S} = S \cup S_0$ 有限, $S = \{1, 2, \dots, p\}$ 为瞬时态集, $S_0 = \{0\}$ 为吸收态集. 一步转移概率矩阵

$$\tilde{P} = \begin{pmatrix} P & P_0 \\ 0 & 1 \end{pmatrix},$$

其中, P 为瞬时态集的转移矩阵, $P_0 = (I - P)e$, $e = (1, 1, \dots, 1)^T$ 为 p 维单位列向量, $\tau = \inf\{n: n \geq 0, X_n \in S_0\}$, 称 τ 为从瞬时态集到吸收态集的首达时间, 称 τ 的分布为 Phase-Type 分布(简称 PH 分布).

令 $\pi(0) = (\alpha_0, \alpha)$, 其中 $\alpha = (\alpha_1, \dots, \alpha_p)$, $\alpha_k \geq 0$, $\sum_{k \in \tilde{S}} \alpha_k = 1$. $g_k = P(\tau = k)$, $g_k(i) = P(\tau = k | X_0 = i)$, $g_k = (g_k(i), i \in S)^T$, $g(i, \lambda) = E(\lambda^\tau | X_0 = i)$, $g(\lambda) = (g(i, \lambda), i \in S)$, $g(\lambda) = E(\lambda^\tau) = \sum_{k=0}^{\infty} g_k \lambda^k$.

下面先求 τ 的分布 $\{g_k, k \geq 1\}$, τ 的条件分布向量 $\{g_k, k \geq 1\}$ 及其生成函数. 有如下的定理:

定理 3.6.1 在上述记号下, 有

(1) $g_0 = \alpha_0$,

$$\forall k \geq 1, g_k = \alpha P^{k-1} P_0 = \alpha P^{k-1} (I - P)e; \quad (3.6.1)$$