

Key concepts:

- 可逆分布;
- 细致平衡条件。

Markov性的一个等价表述是：在已知“现在”的条件下，“过去”和“未来”是独立的，这表明了“过去”和“未来”在时间上的对称性。本节讨论在时间上具有对称性的，即可逆Markov链。

此外，我们已经知道Markov链的长时间行为与不变分布紧密联系，但是直接通过不变方程求解不变分布是困难的，对于可逆Markov链，本节还会给出一个更容易计算不变分布的条件。

7.1 可逆Markov链

Definition 7.1 (时间逆转) 设 $\{X_n\}$ 是一个Markov链，对任意固定时间 N ， $\{Y_n \triangleq X_{N-n} : 0 \leq n \leq N\}$ 称为 $\{X_n\}$ 的时间逆转过程(*Time Reversal*)。

Proposition 7.2 时间逆转过程 $\{Y_n, 0 \leq n \leq N\}$ 是一个Markov链，其一步转移概率为

$$\tilde{P}_{ij} \triangleq P(Y_{n+1} = j | Y_n = i) = \frac{\mu_j^{(N-n-1)}}{\mu_i^{(N-n)}} P_{ji},$$

其中 $\mu^{(n)}$ 是 X_n 的分布。特别地，如果 $\{X_n\}$ 初分布是不变分布，记为 π ，则 $\{Y_n\}$ 是时齐的，一步转移概率为

$$\tilde{P}_{ij} = \frac{\pi_j}{\pi_i} P_{ji}.$$

Proof: 对任意 n , 状态 $i_{N-n}, i_{N-n+1}, \dots, i_N$,

$$\begin{aligned}
& P(Y_n = i_{N-n} | Y_{n-1} = i_{N-n+1}, \dots, Y_0 = i_N) \\
&= P(X_{N-n} = i_{N-n} | X_{N-n+1} = i_{N-n+1}, \dots, X_N = i_N) \\
&= \frac{P(X_{N-n} = i_{N-n}, X_{N-n+1} = i_{N-n+1}, \dots, X_N = i_N)}{P(X_{N-n+1} = i_{N-n+1}, \dots, X_N = i_N)} \\
&= \frac{P(X_{N-n} = i_{N-n})P_{i_{N-n}, i_{N-n+1}}P_{i_{N-n+1}, i_{N-n+2}} \cdots P_{i_{N-1}, i_N}}{P(X_{N-n+1} = i_{N-n+1})P_{i_{N-n+1}, i_{N-n+2}} \cdots P_{i_{N-1}, i_N}} \\
&= \frac{P(X_{N-n} = i_{N-n})P_{i_{N-n}, i_{N-n+1}}}{P(X_{N-n+1} = i_{N-n+1})} \quad (*) \\
&= P(X_{N-n} = i_{N-n} | X_{N-n+1} = i_{N-n+1}) \\
&= P(Y_n = i_{N-n} | Y_{n-1} = i_{N-n+1})
\end{aligned}$$

所以 $\{Y_n\}$ 是一个Markov链, 从推导中的(*)式可以看出一步转移概率为

$$\tilde{P}_{ij} \triangleq P(Y_{n+1} = j | Y_n = i) = \frac{\mu_j^{(N-n-1)}}{\mu_i^{(N-n)}} P_{ji}.$$

■

下面我们定义一类性质良好的Markov链, 它的时间逆转链的转移矩阵依然是 P 。

Definition 7.3 (可逆Markov链) 设 π 为 E 上的一个概率分布, P 为不可约Markov链 $\{X_n\}$ 的转移矩阵, 如果 P 满足

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad \forall i, j \in E \quad (7.1)$$

则称Markov链 $\{X_n\}$ 是可逆的 (*reversible*)。 (7.1)式称为细致平衡条件 (*detailed balance*), 分布 π 称为 P 的一个可逆分布 (*reversible distribution*)。

注1. 一般的如果 E 上测度 ν 满足细致平衡(7.1), 则称 ν 为 P 的一个配称测度 (*Symmetric Measure*), 称 P 是可配称的。

注2. 对于可逆Markov链 $\{X_n\}$, 对任意 k 都有: $(X_k, X_{k-1}, \dots, X_0)$ 和 (X_0, X_1, \dots, X_k) 同分布。

注3. 对于有限状态空间上的可逆Markov链, 转移矩阵 P 的所有特征值为实数。

证明只需要考虑实对称矩阵

$$Q_{ij} \triangleq \frac{\sqrt{\pi_i}}{\sqrt{\pi_j}} P_{ij}$$

那么 Q 的所有特征值为实数，那么

$$P = \text{diag}(\sqrt{\pi_1}, \dots, \sqrt{\pi_n})^{-1} Q \text{diag}(\sqrt{\pi_1}, \dots, \sqrt{\pi_n})$$

的所有特征值为实数。

Proposition 7.4 可逆分布是不变分布。

Proof: 由细致平衡条件

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad \forall i, j \in E$$

对 j 求和有

$$\sum_j \pi_i P_{ij} = \sum_j \pi_j P_{ji}, \quad \forall i \in E$$

由于 $\sum_j P_{ij} = 1$ ，所以

$$\pi_i = \sum_j \pi_j P_{ji}, \quad \forall i \in E$$

即 π 为不变分布。 ■

注1. 该命题的逆命题不一定成立，给一个简单的反例：

考虑状态空间为 $\{1, 2, 3\}$ 的Markov链，转移矩阵为

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

该链具有不变分布 $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ ，然而

$$\pi_1 P_{12} - \pi_2 P_{21} = \frac{1}{6} \neq 0$$

所以该链的不变分布不是可逆分布。

注2. 可逆分布一个良好的性质是它限制在状态空间 E 的子集 $A \in E$ 上依然是可逆分布。令矩阵 \hat{P} 为

$$\hat{P}_{ij} \triangleq \begin{cases} P_{ij}, & \forall j \in A, j \neq i, \\ P_{ii} + \sum_{j \notin A} P_{ij}. \end{cases}$$

则 \hat{P} 是 A 上的转移矩阵。若 π 是 P 的可逆分布，则 \hat{P} 的可逆分布为 $\pi|_A$ 。但是， P 的不变分布 μ 却未必能保证 $\mu|_A$ 是 \hat{P} 的不变测度。

注3. 可逆分布另一个好处是容易计算，看一个简单的例子

Example 7.5 (有限图上的随机游走) 设 $\{X_n\}$ 是连通图 $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ 上的随机游走，其中 $|\mathcal{V}| < \infty$ 。记节点 i 的度为 d_i ，令 $\mu_i = d_i$ ，那么

$$\mu_i p_{ij} = \mu_j p_{ji} = \begin{cases} 1, & \text{若 } i, j \text{ 是邻居,} \\ 0, & \text{否则.} \end{cases}$$

因此细致平衡条件成立，即 μ 是配称测度，又 $|\mathcal{V}| < \infty$ ，做归一化，令

$$\pi \triangleq \frac{d_i}{\sum_j d_j}$$

则 π 是可逆分布，自然 π 也是不变分布。

7.2 Metropolis 链

给定一个不可约转移矩阵 P ，我们之前研究了它的不变分布的相关问题，现在我们考虑它的反问题：

给定状态空间 E 上的概率分布 π ，如何构造转移矩阵 P ，使得 P 的唯一平稳分布为 π ？

本节我们讨论一个重要的可逆Markov链——Metropolis 链，初步回答这个问题。

考虑有限状态空间 $E = \{1, 2, \dots, n\}$ ，设 π 为 E 上任意概率分布，令 Q 是其上的一个不可约对称转移矩阵，即 $Q_{ij} = Q_{ji}$ ，那么由细致平衡条件

$$\nu_i Q_{ij} = \nu_j Q_{ji}$$

可知 E 上的均匀分布 ν 为 Q 的可逆分布。我们现在考虑调整转移概率 Q_{ij} ，使它的不变分布从均匀分布调整到目标分布 π 。定义

$$P_{ij} \triangleq \begin{cases} Q_{ij}a(i, j) & \text{if } j \neq i, \\ 1 - \sum_{k:k \neq i} Q_{ik}a(i, k) & \text{if } j = i. \end{cases}$$

其中 $a(i, j)$ 为接受概率，即我们有 $a(i, j)$ 的概率接受以 Q_{ij} 这个概率从状态 i 转移到 j ，有 $1 - a(i, j)$ 的概率从 i 转移到 j 概率为0。

我们希望通过调整接受概率 $a(i, j)$ 来使得 P 的不变分布为目标分布 π 。我们知道满足细致平衡条件

$$\pi_i Q_{ij} a(i, j) = \pi_j Q_{ji} a(j, i), \quad \forall i \neq j$$

的 P 一定以 π 为不变分布，由于 $Q_{ij} = Q_{ji}$ ，所以只需要满足

$$\pi_i a(i, j) = \pi_j a(j, i), \quad \forall i \neq j$$

注意到 $a(i, j) \leq 1$ ，那么

$$\pi_i a(i, j) \leq \pi_i, \quad \pi_i a(i, j) = \pi_j a(j, i) \leq \pi_j$$

即 $\pi_i a(i, j) \leq \min\{\pi_i, \pi_j\}$ ，我们希望接受概率高一些，不然构造的Markov链转移效率会十分低下，于是取

$$\pi_i a(i, j) = \min\{\pi_i, \pi_j\}$$

即接受概率

$$a(i, j) = \min\{1, \pi_j/\pi_i\}$$

至此，我们构造了一个以给定分布 π 为不变分布的Markov链，称为Metropolis 链。

Definition 7.6 (Metropolis 链) 对于一个概率分布 π 和对称转移矩阵 Q ，对应的Metropolis 链定义为以

$$P_{ij} \triangleq \begin{cases} Q_{ij} \min\{1, \pi_j/\pi_i\} & \text{if } j \neq i, \\ 1 - \sum_{k:k \neq i} Q_{ik} \min\{1, \pi_k/\pi_i\} & \text{if } j = i. \end{cases}$$

为转移概率的Markov链。

注. 给定目标分布 π ，我们可以通过Metropolis 链的转移得到 π 的样本，这种方法称为 Markov Chain Monte Carlo (MCMC) 方法。除了Metropolis 链，还有有很多种方法可以构造MCMC，我们会在之后的课程中介绍。