STAT0008: Stochastic Processes

Lecture 10 - Invariant Distribution,Limit Behavior and Reversibility

Lecturer: Weichen Zhao Spring 2025

Key concepts:

- 不变分布;
- 遍历性;
- 可逆性;
- *Glauber Dynamics.*

上一节课我们建立了研究连续时间Markov链的基础，这一节讨论极限行为，不变分布，可逆性等Markov链的性质。

10.1 状态分类

Definition 10.1 (互通性)
(1) 设$i, j \in E$是*Markov*链中的两个状态，如果$\exists t \geq 0$，使得

$$P_{ij}(t) > 0$$

就称状态i可达状态j，记为$i \to j$;

(2) 如果$i \to j$，且$j \to i$，则称i和j互通，记为$i \leftrightarrow j$;

注. 互通是一个等价关系。

Definition 10.2 (不可约性) 如果对任意i, j, 都存在t, 使得$P_{ij}(t) > 0$, 那么称连续时间*Markov*链 $X(t)$是不可约的，也称Q不可约。

连续时间Markov链的不可约性等价于其嵌入链的不可约性，具体我们有以下命题。

Proposition 10.3 设状态$i \neq j \in E$，则下面几条等价：

(1) $i \to j$;

(2) 对于嵌入链，$i \to j$;

(3) 存在$i = i_0, i_1, i_2, \cdots, i_n = j$，使得$q_{i_0 i_1}, \cdots, q_{i_{n-1} i_n} > 0$;

(4) 对所有$t > 0$，$P_{ij}(t) > 0$;

(5) 存在$t > 0$，使得$P_{ij}(t) > 0$.

注. (4)和(5)等价说明：对任意给定的i, j，$P_{ij}(t)$，$\forall t > 0$要么恒为0，要么恒为正，这意味着，连续时间Markov链都是非周期的。

Definition 10.4 (常返性) 对状态$i \in E$，如果

$$P(\forall t > 0, \exists s > t, \text{s.t. } X(s) = i \mid X(0) = i) = 1$$

则称状态i是常返的，否则称i为暂留的，或者非常返的。

连续时间Markov链的常返性依然可以通过其嵌入链的常返性刻画。定义

$$\sigma_i \triangleq \inf\{t \geqslant S_1 : X(t) = i\}$$

为连续时间Markov链$X(t)$首次“跳到”状态i的时刻，其中S_1是$X(t)$的首次起跳时刻。

用$\sigma_i^{(k)}$表示$X(t)$第k次跳到i的时间，它一定是某个跳跃时刻，记为S_{N_k}。具体地，令$\sigma_i^{(1)} = \sigma_i, S_{N_1} = \sigma_i^{(1)}$,

$$\sigma_i^{(k+1)} = \inf\{t \geqslant S_{N_k+1} : X(t) = i\}, \quad S_{N_{k+1}} = \sigma_i^{(k+1)}.$$

具体我们有以下命题。

Proposition 10.5 下面几条等价：

(1) i是常返态，即

$$P(\forall t>0, \exists s>t, \text{s.t. } X(s)=i \mid X(0)=i)=1$$

(2) $q_i=0$ *(*吸收态*)* 或从状态i出发迟早回到i的概率

$$f_{ii} \triangleq P(\sigma_i<\infty|X(0)=i)=1;$$

(3) $q_i=0$ *(*吸收态*)* 或对任何自然数$k\geqslant 1$,

$$P(\sigma_i^{(k)}<\infty \mid X(0)=i)=1;$$

(4) $G_{ii} \triangleq \int_0^\infty P_{ii}(t)dt=\infty$;

(5) i 是嵌入链的常返态。

注. 实际上，我们只考虑$0<q_i<+\infty$的逗留Markov链。

10.2 不变分布

Definition 10.6 (不变测度) 设$\pi=\{\pi_i, i\in E\}$为E上的测度，若π满足不变方程

$$\pi=\pi P(t), \quad \forall t>0$$

则称π为$P(t)$的不变测度。进一步，若$\sum_{i\in E}\pi_i=1$，则称π为不变分布。

和离散时间一样，我们也希望能够通过一步转移矩阵给出不变分布的定义。

Theorem 10.7 $\pi=\{\pi_i, i\in E\}$是不变测度的充分必要条件为：

$$\pi Q=0$$

Proof: 只证有限状态情形。

$$
\begin{aligned}
&\pi Q = 0 \\
&\Leftrightarrow \sum_{n=1}^{N} \frac{t^n}{n!}\pi Q^n = 0 \quad \Leftrightarrow \sum_{n=0}^{N} \frac{t^n}{n!}\pi Q^n = \pi \quad \Leftrightarrow \pi \exp(Qt) = \pi \\
&\Leftrightarrow \pi P(t) = \pi
\end{aligned}
$$

■

如同连续时间Markov链与其嵌入链的不可约性以及常返性上的关系，两者不变分布的关系可以表述为下面命题。

Proposition 10.8 *设$X(t)$为一连续时间Markov链，转移矩阵为$P(t)$，嵌入链为X_n，则*

(1) 设π为$X(t)$的不变分布，那么

$$\mu = (\mu_i)_{i\in E}, \quad \mu_i \triangleq \frac{\pi_i q_i}{\sum\limits_{j\in E} \pi_j q_j}$$

为嵌入链X_n的不变分布；

(2) 设μ为X_n的不变分布，且$\sum_{i\in E} \frac{\mu_i}{q_i} < \infty$，那么

$$\pi = (\pi_i)_{i\in E}, \quad \pi_i \triangleq \frac{\mu_i}{q_i \sum\limits_{j\in E} \frac{\mu_j}{q_j}}$$

为连续时间Markov链$X(t)$的不变分布。

容易自行验证。

10.3 极限行为

类似于离散时间Markov链，我们感兴趣转移概率的极限行为。

Proposition 10.9 (访问频率) 设Q不可约常返，那么对任意状态i,

$$P_i\left(\lim_{t\to\infty}\frac{1}{t}\int_0^t 1_{\{X_s=i\}}\mathrm{d}s=\frac{1}{q_i\mathbb{E}[\sigma_j]}\right)=1.$$

其中$P_i(\cdot)=P(\cdot\mid X(0)=i)$表示在$X(0)=i$条件下的条件分布。

证明参考：陈&章，《应用随机过程》，命题2.4.2

Definition 10.10 (正常返) 设状态i是常返的，如果平均返回时间

$$\mu_{ii}\triangleq\mathbb{E}[\sigma_i]<\infty$$

则称状态i正常返；否则称状态i零常返。

注. 虽然连续时间Markov链与其嵌入链的不可约性以及常返性一致，但是两者的正常返性不一定一致，反例参考：陈&章，《应用随机过程》，例2.4.11

类似于离散时间Markov链，下面定理给出不变分布的存在唯一性。

Theorem 10.11 设Q不可约，则下面几条等价：

(1) 所有状态正常返*;*

(2) 存在正常返态*;*

(3) 不变分布存在。

特别地，若π是不变分布，则

$$\pi_i=\frac{1}{q_i\mathbb{E}[\sigma_i]}$$

证明参考：陈&章，《应用随机过程》，定理2.4.8

类似于离散时间Markov链，有(平均)遍历定理:

Theorem 10.12 (遍历定理) 设$\{X(t)\}$是不可约常返的$Markov$链，π是不变分布，f是E上的函数，满足 $\sum_{i\in E}\pi_i|f(i)|<\infty$，则

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f(X(s))ds=\sum_{i\in E}\pi_i f(i)$$

由于连续时间情形不涉及周期性，那么自然有强遍历性的相关结果

Theorem 10.13 (强遍历定理) 设$\{X(t)\}$是不可约常返的$Markov$链，π是不变分布，那么对任意状态i,j,

$$\lim_{t\to\infty}P_{ij}(t)=\pi_j$$

注1. 如果连续时间Markov链$X(t)$不可约，但不变分布不存在，那么

$$\lim_{t\to\infty}P_{ij}(t)=0$$

注2. 如果连续时间Markov链$X(t)$不可约正常返，我们称$X(t)$是遍历的。(离散时间要求不可约非周期正常返)

10.4 可逆性

Definition 10.14 (时间逆转过程)
设$\{X(t)\}$是一个连续时间$Markov$链，对任意固定时间T，称连续时间随机过程$\{Y(t)\triangleq X(T-t),0\le t\le T\}$为$\{X(t)\}$的时间逆转过程。

注. 更严格来说，为了保证逆转过程的轨道右连续，应该定义$Y(t)\triangleq \lim\limits_{s\to(T-t)^-}X(s)$.

Proposition 10.15 设转移速率矩阵$Q=(q_{ij})$不可约，具有不变分布π，给定时间$0<T<\infty$，设$\{X(t)\}$是密度为Q的连续时间$Markov$链，初分布为π，则它的逆转过程$\{Y(t),0\le t\le T\}$为一个以

$$\tilde{q}_{ij}\triangleq\frac{\pi_j}{\pi_i}q_{ji}$$

为转移速率矩阵的连续时间$Markov$链。

证明参考 Markov chains, J. R. Norris, Cambridge university press, Theorem 3.7.1.

Definition 10.16 (可逆过程) 设π为E上的一个概率分布，Q为不可约连续时间$Markov$链$\{X(t)\}$的转移速率矩阵，如果Q满足

$$\pi_i q_{ij} = \pi_j q_{ji}, \quad \forall i, j \in E \tag{10.1}$$

则称$\{X(t)\}$是可逆的。(10.1)式称为细致平衡条件，分布π称为Q的一个可逆分布。

Proposition 10.17 可逆分布是不变分布。

10.5 Ising模型与Glauber动力学

Ising模型是由物理学家 Ernst Ising 于1925 年提出的统计物理模型，用来研究铁磁性物质磁性性质的模型，本小节我们简单介绍有限网格(lattice)上的最近邻相互作用(nearest neighbour interaction) Ising 模型。

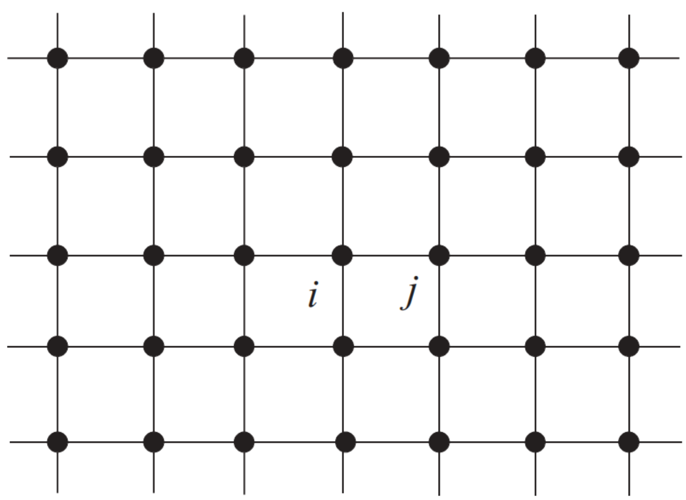

图 10.1: 二维网格以及其上的最近邻粒子i, j

记$T := \{1, 2, \ldots, N\}$是一个有限集合，定义$G \triangleq T^d$为每边为N个格点的d维格点阵列。设每个格点(site)上有一个粒子，可以处于+1或−1两种状态之一，称为该粒子的自旋(spin).

那么T^d个粒子组成的粒子系统中每个粒子赋予一种自旋，就组成了粒子系统的一种构型/组态(configuration)，全体构型组成的集合记为S，称为构型空间，它是一个有限集合

$$S := \{+1, -1\}^G = \{\sigma : \sigma = (\sigma(i)), i \in T^d, \sigma(i) = +1\text{或} -1\}$$

在构型空间S上可以定义如下的能量函数/哈密顿量(energy function/Hamiltonian)：

$$H(\sigma) = -J \sum_{(ij)\in B} \sigma(i)\sigma(j) - h \sum_i \sigma(i).$$

其中B是键(bond)的集合，J是相互作用强度，h是外场强度，$-J\sum_{(ij)\in B}\sigma(i)\sigma(j)$表示相邻格点对上自旋相互作用对能量的贡献，$-h\sum_i \sigma(i)$表示外场带来的势能。

记σ^i为构型σ只在格点$i \in T^d$处翻转自旋符号后得到的构型，即

$$\sigma^i(j) := \begin{cases} \sigma(j) & \text{如果 } j \neq i, \\ -\sigma(i) & \text{如果 } j = i \end{cases}$$

下面定义描述Ising模型演化的Glauber动力学

Definition 10.18 (Glauber动力学)
Ising模型的Glauber动力学(Glauber Dynamics)是定义在构型空间S上的连续时间Markov链，具有如下的转移速率矩阵

$$q_{\sigma\eta} := \begin{cases} c(i,\sigma) & \text{如果 } \eta = \sigma^i, \\ -\sum_i c(i,\sigma) & \text{如果 } \eta = \sigma \\ 0 & \text{否则} \end{cases}$$

其中转移速率$c(i,\sigma)$有以下几种取法：

(1) $c(i,\sigma) = \min\{e^{-\beta(H(\sigma^i)-H(\sigma))}, 1\}$

(2) $c(i,\sigma) = \frac{1}{1+e^{\beta(H(\sigma^i)-H(\sigma))}} = \frac{e^{-\beta H(\sigma^i)}}{e^{-\beta H(\sigma)}+e^{-\beta H(\sigma^i)}}$.

(3) $c(i,\sigma) = \frac{1}{2}\left[1 + e^{-\beta(H(\sigma^i)-H(\sigma))}\right]$.

其中(1)称为Metropolis dynamics，(2)称为heat-bath dynamics.

由于构型空间S是有限的，$|S| = 2^{N^d}$*，所以其转移矩阵为*$P(t) = e^{tQ}$.

定义Boltzmann-Gibbs分布

$$P(\sigma) = \frac{\mathrm{e}^{-\beta H(\sigma)}}{Z} \tag{10.2}$$

其中$\beta = \frac{1}{T}$是逆温度参数(inverse temperature)，Z是配分函数(partition function)

$$Z = \sum_{\sigma} e^{-\beta H(\sigma)}.$$

容易通过细致平衡条件验证Boltzmann-Gibbs分布(10.2)是Glauber动力学的可逆分布，所以有限格点上的Ising模型收敛到热力学平衡时，每个构型服从Boltzmann-Gibbs分布，统计物理中一般通过对Boltzmann-Gibbs分布计算热力学平均来刻画物理量。