

目录

第一章 独立随机变量	3
1.1 投球入箱问题 (balls-into-bins)	3
1.2 集中不等式 (concentration inequalities)	4
1.3 独立模型与集中不等式的一些应用	11
第二章 有限状态马尔可夫链	21
2.1 离散马尔可夫链的基本概念	21
2.2 马尔可夫链基本定理	26
2.3 时间反演, 可逆链与 Metropolis-Hastings 算法	29
2.4 耦合 (coupling)	34
2.5 马尔可夫链基本定理的证明	37
2.6 马尔可夫链的收敛速率与混合时间 (mixing time)	41
第三章 可数无限状态马尔可夫链	45
3.1 信息流 (filtration)、停时 (stopping time) 与强马尔可夫性 (strong Markov property)	46
3.2 从有限到无限	49
3.3 首达时间的相关计算	53
3.4 状态的分类	60
3.5 马尔可夫链基本定理 (可数无穷状态)	63

第一章 独立随机变量

随机过程是研究一系列随机变量 X_1, X_2, \dots , 或者一族随机变量 $\{X(t)\}_{t \geq 0}$ 的学问, 其重点会放在随机变量之间的关系上. 最基本的情况便是这些随机变量都是互相独立的. 在概率论课上, 我们学习过经典的大数定律和中心极限定理, 这便是独立随机变量序列的经典结果. 我们在这一节, 介绍一些独立随机变量序列的非渐近结果, 它们不是我们这本书的重点, 但我们会通过这些讨论引入几个重要的概率模型和数学工具, 是未来研究更复杂的随机过程的基础.

1.1 投球入箱问题 (balls-into-bins)

投球入箱是一个简单的随机过程, 或者说是随机试验: 将 m 个球均匀随机地投入 n 个箱子中. 我们用 $[m]$ 来编号所有的球, 用 $[n]$ 来编号所有的箱子, 并用 X_i 来表示第 i 个球落入的箱子的编号. 那么, X_1, X_2, \dots, X_m 就是取值为 $[n]$ 的独立随机变量. 假设每一个 X_i 都是均匀的随机变量是最简单的情况. 对于这个过程, 我们可以提出许多有趣的问题. 我们这儿介绍俩经典的.

生日悖论 (birthday paradox)

生日悖论指的是在一个有 30 个同学左右的班级中某些同学很可能共享相同生日这一反直觉现象. 将箱子视为日期、球视为学生, 两个学生生日相同的事件可以建模为某个箱子包含多于一个球的情况.

注意到每个球的投掷是独立的. 在已投掷 $k-1$ 个球且无碰撞的条件下, 投掷第 k 个球后仍无碰撞的概率为 $\frac{n-k+1}{n}$. 因此,

$$\mathbb{P}[\text{无相同生日}] = \prod_{k=1}^m \frac{n-k+1}{n} = \prod_{k=1}^{m-1} \left(1 - \frac{k}{n}\right).$$

概率论课上遇到的大数定律和中心极限定理都是讨论一系列随机变量 $X_1, X_2, \dots, X_n, \dots$ 当 $n \rightarrow \infty$ 时的渐近性质. 我们这里说的非渐近性质指的是对于固定 (但足够大) 的 n , 序列 X_1, \dots, X_n 的性质.

对于自然数 $n \in \mathbb{N}$, 我们用 $[n]$ 表示集合 $\{1, 2, \dots, n\}$.

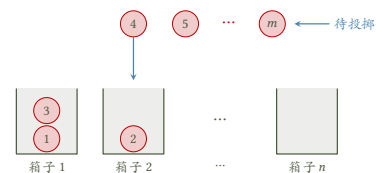


图 1.1: 投球入箱

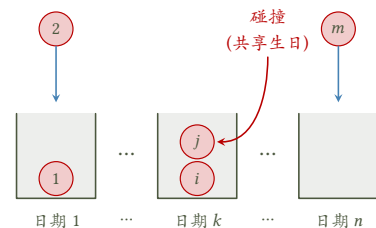


图 1.2: 生日悖论

由 $1 + x \leq e^x$, 上面的概率有上界

$$\exp\left(-\frac{\sum_{k=1}^{m-1} k}{n}\right) = \exp\left(-\frac{m(m-1)}{2n}\right).$$

于是, 可以看到当 m 是 \sqrt{n} 的一个足够大的常数倍时, 该概率可以任意接近 0.

$1 + x \leq e^x$ 是一个常用的不等式, 它来自于对于 e^x 的线性近似, 并且对于任何 $x \in \mathbb{R}$ 均成立. 我们有的时候也想用另一方向的不等式, 一个实用的并且方便记忆的选择是:

$$\forall k > 0, 1 + 1/k \geq e^{1/(k+1)};$$

$$\forall k > 1, 1 - 1/k \geq e^{-1/(k-1)}.$$

奖券收集 (coupon collector) 问题

奖券收集问题是如下问题: 如果某手机游戏开卡包机制是每一次均匀随机的开出 n 种不同的卡片之一, 总共需要开多少包才能收集全部 n 种卡片呢? 用投球入箱的语言表述, 我们用箱子来表示每一种卡片, 用每一个球来表示开启的每一包卡牌, 那么即需要投入多少球才能没有空箱子.

需要购买的卡包数是一个随机变量, 我们用 Y 来表示. 我们首先利用期望的线性性可计算 Y 的期望. 设 Y_i 表示在已持有 $i-1$ 种卡片时, 收集第 i 种卡片所需的开包次数. 那么显然有 $Y = \sum_{i=1}^n Y_i$. 由期望线性性:

$$\mathbf{E}[Y] = \sum_{i=1}^n \mathbf{E}[Y_i].$$

显然 $Y_i \sim \text{Geom}\left(\frac{n-i+1}{n}\right)$, 因此 $\mathbf{E}[Y_i] = \frac{n}{n-i+1}$. 从而

$$\mathbf{E}[Y] = \sum_{i=1}^n \frac{n}{n-i+1} = n \cdot H_n.$$

对于几何分布 $X \sim \text{Geom}(p)$, 有 $\mathbf{E}[X] = p^{-1}$.

这儿 $H_n = \sum_{i=1}^n \frac{1}{i}$ 为调和级数的前 n 项, 满足 $H_n \sim \log n + \gamma$, 其中常数 $\gamma = 0.577\dots$ 被称为欧拉常数.

上述使用期望的线性性进行计算的技巧是非常常用的. 此外, 奖券收集问题的结论也值得记住, 在我们后续课程中会多次的遇到它.

1.2 集中不等式 (concentration inequalities)

在奖券收集问题里, 从实际的角度来说, 我们知道平均 $n \cdot H_n$ 包可以集齐全卡是不够的, 因为实际开包的过程可能和所谓的“期望值”相去甚远. 我们需要知道关于随机变量 Y 更进一步的信息, 比如, 如果我想保证以至少 99% 的概率收集全所有卡片, 我需要开多少包? 这就涉及到随机变量的集中不等式.

马尔可夫不等式 (Markov's inequality)

马尔可夫不等式可能是最简单集中不等式, 它对于任意非负的随机变量均成立.

定理 1.1 马尔可夫不等式.

对任意非负随机变量 $X \geq 0$ 及 $a > 0$,

$$\mathbb{P}[X \geq a] \leq \frac{\mathbf{E}[X]}{a}.$$

证明. 由于 X 非负, 由全期望公式, 有 $\mathbf{E}[X] \geq a \cdot \mathbb{P}[X \geq a] + 0 \cdot \mathbb{P}[X < a]$. 这便等价于我们想要证明的. \square

全期望公式指的是对于样本空间的一个可数划分 $\{A_i\}$, 其中每一个 A_i 是一个事件, 有 $\mathbf{E}[X] = \sum_i \mathbf{E}[X | A_i] \cdot \mathbb{P}[A_i]$.

示例 1.2 奖券收集的集中性.

回忆 Y 为所需球数. 应用马尔可夫不等式, 对 $c > 0$ 有

$$\mathbb{P}[Y \geq c] \leq \frac{\mathbf{E}[Y]}{c} = \frac{n \cdot H_n}{c}.$$

因此, 需要开超过 $100 \cdot n \cdot H_n$ 包的概率小于 0.01.

取 $n = 100$: $H_{100} \approx 5.187$, $\mathbf{E}[Y] \approx 519$. 马尔可夫不等式要求 $c \geq 100 \times 519 \approx 51900$ 包才能保证 99% 的概率集齐.

切比雪夫不等式 (Chebyshev's inequality)

马尔可夫不等式的成立仅仅要求随机变量 X 是非负的, 但实际上, 如果我们考虑 X 和它的期望的偏差程度, 即 $\mathbb{P}[|X - \mathbf{E}[X]| \geq a]$, 那么马尔可夫不等式总能够给出一个上界, 即 $\mathbb{P}[|X - \mathbf{E}[X]| \geq a] \leq \frac{\mathbf{E}[|X - \mathbf{E}[X]|^2]}{a^2}$. 因此, 马尔可夫不等式的成立可以看成是无条件的, 但也说明了它在很多随机变量上并不紧, 因为它没有利用足够多的随机变量的信息.

我们用一个简单的技巧即可改进它. 对于一个 (在我们关心的范围内的) 递增的函数 $f: \mathbb{R} \rightarrow \mathbb{R}$, 我们有 $\mathbb{P}[X \geq a] = \mathbb{P}[f(X) \geq f(a)]$. 我们便可以对后者使用马尔可夫不等式.

定理 1.3 切比雪夫不等式.

对任意具有有限期望 $\mathbf{E}[X]$ 的随机变量及 $a > 0$, 有

$$\mathbb{P}[|X - \mathbf{E}[X]| \geq a] \leq \frac{\mathbf{Var}[X]}{a^2}.$$

证明. 令 $Y = |X - \mathbf{E}[X]|$, 显然 $Y \geq 0$. 因此

$$\mathbb{P}[|X - \mathbf{E}[X]| \geq a] = \mathbb{P}[Y \geq a] = \mathbb{P}[Y^2 \geq a^2].$$

利用马尔可夫不等式, 我们立即得到

$$\mathbb{P}[Y^2 \geq a^2] \leq \frac{\mathbf{E}[Y^2]}{a^2} = \frac{\mathbf{Var}[X]}{a^2}.$$

□

示例 1.4 再探奖券收集.

将切比雪夫不等式应用于奖券收集问题. 沿用之前记号, 有

$$P[Y \geq nH_n + t] \leq P[|Y - \mathbf{E}[Y]| \geq t] \leq \frac{\mathbf{Var}[Y]}{t^2}.$$

我们之前使用 Y_i 表示已有 $i-1$ 种卡时获得一种新卡所需的开包次数. 对不同的 i, j , Y_i 与 Y_j 独立. 因此

$$\mathbf{Var}[Y] = \mathbf{Var}\left[\sum_{i=1}^n Y_i\right] = \sum_{i=1}^n \mathbf{Var}[Y_i].$$

对 $i \in [n]$, $Y_i \sim \text{Geom}\left(\frac{n-i+1}{n}\right)$, 故

$$\mathbf{Var}[Y_i] = \frac{1 - \frac{n-i+1}{n}}{\left(\frac{n-i+1}{n}\right)^2} = \frac{(i-1) \cdot n}{(n-i+1)^2} \leq \frac{n^2}{(n-i+1)^2}.$$

因此, 我们只需要控制住 $\sum_{i=1}^n \frac{1}{(n-i+1)^2} = \sum_{i=1}^n \frac{1}{i^2}$. 注意到

$$\sum_{i=1}^n \frac{1}{i^2} \leq 1 + \int_1^{\infty} \frac{dx}{x^2} = 2$$

成立. 因此 $\mathbf{Var}[X] \leq 2n^2$ 且 $P[X \geq nH_n + t] \leq \frac{2n^2}{t^2}$. 切比雪夫不等式告诉我们, 需要抽取超过 $\sqrt{200n} + nH_n$ 包的概率小于 0.01. 相比马尔可夫不等式 (示例 1.2), 这距离事实要接近很多.

我们发现, 在切比雪夫不等式的证明中, 我们通过控制 $\mathbf{E}[f(|X - \mathbf{E}[X]|)]$, 其中 $f(x) = x^2$, 引入了方差. 一个很自然的问题是, 如果我们选择别的 $f(x)$, 能否得到关于“尾分布” $P[|X - \mathbf{E}[X]| \geq t]$ 的更好的不等式? 切比雪夫不等式仅仅用到了随机变量二阶矩的信息.

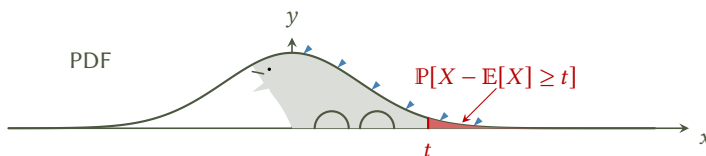


图 1.3: “尾分布”

我们首先要来看一下, 在什么情况下, 切比雪夫不等式对于尾分布得到的界并不紧. 我

对于两两独立的随机变量, 方差和求和可以交换.

在这个问题上, 通过切比雪夫不等式得到的界比马尔可夫不等式紧得多——为达到相同置信度, 马尔可夫不等式说明需选择 $t = \Theta(n \log n)$. 仍取 $n = 100$: $\sqrt{200} \times 100 + 519 \approx 1933$ 包即可保证 99% 集齐——对比示例 1.2 中马尔可夫不等式给出的约 51900 包, 改善了约 27 倍.

们考察熟悉的高斯分布 $X \sim \mathcal{N}(0, 1)$. 使用切比雪夫不等式, 我们得到

$$\mathbb{P}[X \geq t] \leq \frac{\text{Var}[X]}{t^2} = \frac{1}{t^2}.$$

而事实上, 我们可以使用高斯分布的概率密度函数直接计算出这个尾分布概率:

$$\mathbb{P}[X \geq t] = \int_t^{\infty} \phi(x) dx = \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-\frac{x^2}{2}} dx.$$

标准高斯分布的概率密度函数 $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

注意到

$$\int_t^{\infty} e^{-\frac{x^2}{2}} dx \leq \int_t^{\infty} \frac{x}{t} \cdot e^{-\frac{x^2}{2}} dx = t^{-1} e^{-t^2/2},$$

我们有 $\mathbb{P}[X \geq t] \leq \frac{1}{\sqrt{2\pi t}} e^{-\frac{t^2}{2}}$. 这说明, 实际上, 高斯的尾分布是关于 t^2 指数下降的, 对于高斯分布来说, 切比雪夫不等式得到的界离真相还差距很远.

如果一个分布它长得有点接近高斯, 那我们期待能够得到比切比雪夫不等式更加紧的集中不等式. 什么样的分布会接近高斯呢? 在概率论课中, 我们学习过中心极限定理, 我们知道, 在一定条件下, 很多独立的随机变量之和的分布会收敛到高斯分布. 我们接下来两节就讨论这样的随机变量的集中不等式.

切尔诺夫界 (Chernoff bound)

我们接着来考虑一系列独立随机变量的和. 中心极限定理说, 独立随机变量的和的分布 (在一定条件下) 会收敛到高斯分布, 因此, 我们接下来的一系列结论可以看成是非渐近版本的中心极限定理, 也就是当随机变量的个数 n 是固定的, 而不一定趋向于无穷时候的中心极限定理. 最简单的, 就是当每个随机变量是伯努利变量的时候, 也对应了棣莫弗-拉普拉斯 (De Moivre-Laplace) 中心极限定理. 我们把这个结论称作切尔诺夫界.

棣莫弗-拉普拉斯中心极限定理说的是对于一系列 X_1, X_2, \dots , 如果每一个 $X_i \sim \text{Ber}(p)$, 并且是相互独立的, 那么

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{D} Y \sim \mathcal{N}(0, 1).$$

当然, 我们的切尔诺夫界允许每一个 X_i 具有不同的均值 p_i .

定理 1.5 切尔诺夫界.

设 X_1, \dots, X_n 为独立随机变量, 其中每个 $X_i \sim \text{Ber}(p_i)$ ($i = 1, 2, \dots, n$). 令 $X = \sum_{i=1}^n X_i$, 并记 $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$, 则有

- 对于任意 $\varepsilon > 0$, $\mathbb{P}[X \geq (1 + \varepsilon)\mu] \leq \left(\frac{e^\varepsilon}{(1+\varepsilon)^{1+\varepsilon}}\right)^\mu$.
- 对于 $0 < \varepsilon < 1$, $\mathbb{P}[X \leq (1 - \varepsilon)\mu] \leq \left(\frac{e^{-\varepsilon}}{(1-\varepsilon)^{1-\varepsilon}}\right)^\mu$.

证明. 我们这儿仅证明第一条 (即“上尾界 (upper tail bound)”), 对于下尾界的证明类似. 对任意 $\beta > 0$, 有

$$\mathbb{P}[X \geq (1 + \varepsilon)\mu] = \mathbb{P}\left[e^{\beta X} \geq e^{\beta(1+\varepsilon)\mu}\right] \leq \frac{\mathbb{E}\left[e^{\beta X}\right]}{e^{\beta(1+\varepsilon)\mu}}.$$

因此需要估计矩生成函数 $\mathbf{E}[e^{\beta X}]$. 由于 $X = \sum_{i=1}^n X_i$ 是独立伯努利变量之和, 可得

$$\mathbf{E}[e^{\beta X}] = \mathbf{E}\left[e^{\beta \sum_{i=1}^n X_i}\right] = \mathbf{E}\left[\prod_{i=1}^n e^{\beta X_i}\right] = \prod_{i=1}^n \mathbf{E}[e^{\beta X_i}].$$

对 $X_i \sim \text{Ber}(p_i)$, 直接计算得:

$$\mathbf{E}[e^{\beta X_i}] = p_i e^{\beta} + (1 - p_i) = 1 + (e^{\beta} - 1)p_i \leq \exp((e^{\beta} - 1)p_i)$$

因此,

$$\mathbf{E}[e^{\beta X}] \leq \prod_{i=1}^n \exp((e^{\beta} - 1)p_i) = \exp((e^{\beta} - 1)\mu).$$

从而

$$\mathbb{P}[X \geq (1 + \varepsilon)\mu] \leq \left(\frac{\exp(e^{\beta} - 1)}{\exp(\beta(1 + \varepsilon))}\right)^{\mu}.$$

此式对任意 $\beta > 0$ 成立. 选择 β 使分数最小化, 令导数为零:

$$\frac{d}{d\beta} \left(\frac{\exp(e^{\beta} - 1)}{\exp(\beta(1 + \varepsilon))}\right) = \exp(e^{\beta} - 1 - \beta - \beta\varepsilon) \cdot (e^{\beta} - 1 - \varepsilon) = 0.$$

解得 $\beta = \log(1 + \varepsilon)$, 代入得:

$$\mathbb{P}[X \geq (1 + \varepsilon)\mu] \leq \left(\frac{e^{\varepsilon}}{(1 + \varepsilon)^{1 + \varepsilon}}\right)^{\mu}.$$

□

注意到在上述证明中, 我们依旧类似切比雪夫不等式的证明, 先使用 $\mathbb{P}[X \geq a] = \mathbb{P}[f(X) \geq f(a)]$, 然后使用马尔可夫不等式. 我们这儿选取的 $f(x) = e^{\beta x}$ 是指数函数, 因此 $\mathbf{E}[f(X)] = \mathbf{E}[e^{\beta X}]$ 是随机变量 X 的矩生成函数. 在一定条件下, 矩生成函数唯一确定了随机变量. 它可以看成是 X 的所有的 k 阶矩 $\mathbf{E}[X^k]$ 的加权和, 而权重由参数 β 决定. 事实上, 我们可以选择一个合适的 k , 取 $f(x) = x^k$, 然后得到类似的结果 [Tao12]. 但是从证明中可以看到, 选用矩生成函数在计算上十分方便.

事实上, [定理 1.5](#) 里面给出来的界形式上用起来并不是很方便, 也不是很好看出为什么它具有类似于高斯尾分布的下降速度. 我们经常使用下面这个推论里给出的界.

推论 1.6.

对任意 $0 < \varepsilon < 1$, 有更易用 (但稍弱) 的形式:

- $\mathbb{P}[X \geq (1 + \varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2}{3}\mu\right).$
- $\mathbb{P}[X \leq (1 - \varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2}{2}\mu\right).$

如果 X 和 Y 独立, 那么 $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$.

这里又用了 $1 + x \leq e^x$.

在本讲义中我们用 \log 表示自然对数.

证明. 仅证上尾. 需验证当 $0 < \varepsilon < 1$ 时:

$$\frac{e^\varepsilon}{(1+\varepsilon)^{1+\varepsilon}} \leq \exp\left(-\frac{\varepsilon^2}{3}\right).$$

取对数等价于:

$$\varepsilon - (1+\varepsilon)\log(1+\varepsilon) \leq -\frac{\varepsilon^2}{3}.$$

令 $f(\varepsilon) = \varepsilon - (1+\varepsilon)\log(1+\varepsilon) + \frac{\varepsilon^2}{3}$, 其导数满足:

$$f'(\varepsilon) = -\log(1+\varepsilon) + \frac{2}{3}\varepsilon, \quad f''(\varepsilon) = -\frac{1}{1+\varepsilon} + \frac{2}{3}.$$

分析凹凸性可知 $f(\varepsilon) \leq 0$, 故原式成立. □

示例 1.7 公平硬币检验.

现有一枚硬币, 已知其要么为公平硬币 (正面概率 0.5), 要么为有偏硬币 (正面概率 $0.5 + \varepsilon$). 通过抛掷 T 次后计算正面比例 \hat{p} 进行判断是哪种: 若 $\hat{p} \leq 0.5 + \varepsilon/2$ 则判为公平, 否则判为有偏. 现在请问需要投掷多少次硬币, 能够保证以至少 $1 - \delta$ 的概率正确判断.

我们不妨假设硬币本身是公平的, 即 $X_i \sim \text{Ber}(1/2)$. 硬币本身是有偏的场合可以类似说明. 那么, 总正面数 $X = \sum_{i=1}^T X_i$, 则误判概率为:

$$\mathbb{P}[\hat{p} > 0.5 + \varepsilon/2] = \mathbb{P}[X > (1+\varepsilon)\mathbf{E}[X]]$$

由切尔诺夫界, 当

$$\exp\left(-\frac{\varepsilon^2}{3} \cdot \frac{T}{2}\right) \leq \delta$$

时, 解得 $T \geq \frac{6}{\varepsilon^2} \log \frac{1}{\delta}$. 事实上, 在忽略常数倍的意义下, 这个界是最优的.

我们将在作业里证明 $\Omega\left(\frac{1}{\varepsilon^2} \log \frac{1}{\delta}\right)$ 次是必要的.

霍夫丁不等式 (Hoeffding's inequality)

切尔诺夫界的一个恼人限制是要求每个 X_i 必须是伯努利随机变量. 实际上, 正如同中心极限定理一样, 我们可以大大放松这个要求. 正确刻画这类具有类高斯尾分布的概念叫做次高斯分布 (*sub-Gaussian distribution*). 我们在本章习题中会讨论这类分布, 现在我们先介绍一个特殊情况, 也是在本门课中最常用的情况, 即有界独立变量的霍夫丁 (Hoeffding) 不等式.

定理 1.8 霍夫丁不等式.

设 X_1, \dots, X_n 为独立随机变量, 其中每个 X_i 以概率 1 落在区间 $[a_i, b_i]$ 内 ($a_i \leq b_i$). 令 $X = \sum_{i=1}^n X_i$ 则对所有 $t \geq 0$ 有

$$P[|X - \mathbf{E}[X]| \geq t] \leq 2 \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

成立.

类似切尔诺夫界的证明, 建立此类集中不等式的关键在于对矩生成函数进行有效上界估计. 因此, 下面的霍夫丁引理将成为证明的主要技术工具.

引理 1.9 霍夫丁引理.

设 X 为满足 $\mathbf{E}[X] = 0$ 且 $X \in [a, b]$ 的随机变量, 则对任意 $\alpha \in \mathbb{R}$ 有

$$\mathbf{E}[e^{\alpha X}] \leq \exp\left(\frac{\alpha^2(b-a)^2}{8}\right)$$

霍夫丁引理的证明.

我们首先寻找一个线性函数作为 $e^{\alpha x}$ 的上界, 以便利用期望的线性性来估计 $\mathbf{E}[e^{\alpha X}]$. 根据指数函数的凸性, 对于任何 $x \in [a, b]$, 有

$$e^{\alpha x} \leq \frac{e^{\alpha b} - e^{\alpha a}}{b - a}(x - a) + e^{\alpha a}.$$

因此, 根据期望的线性性,

$$\begin{aligned} \mathbf{E}[e^{\alpha X}] &\leq \frac{e^{\alpha b} - e^{\alpha a}}{b - a}(\mathbf{E}[X] - a) + e^{\alpha a} \\ \triangleright \mathbf{E}[X] = 0 &= \frac{-a}{b - a}e^{\alpha b} + \frac{b}{b - a}e^{\alpha a} \\ &= e^{\alpha a} \left(\frac{b}{b - a} - \frac{a}{b - a}e^{\alpha(b-a)} \right) \\ \triangleright \theta = -\frac{a}{b-a}, t = \alpha(b-a) &= e^{-\theta t} (1 - \theta + \theta e^t) \\ &=: e^{g(t)}. \end{aligned}$$

其中 $g(t) = -\theta t + \log(1 - \theta + \theta e^t)$. 根据泰勒定理, 对任意实数 t 存在 $\delta \in (0, t)$ 使得

$$g(t) = g(0) + t g'(0) + \frac{1}{2} g''(\delta) t^2.$$

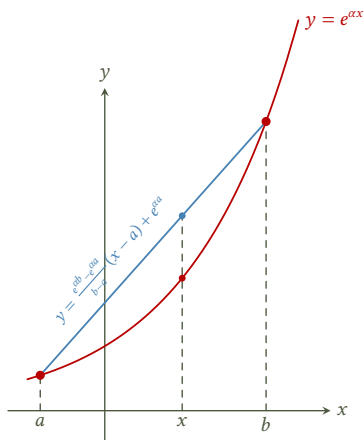


图 1.4: 凸函数 $e^{\alpha x}$ 的线性上界

计算各阶导数:

$$\begin{aligned} g(0) &= 0; \\ g'(0) &= -\theta + \frac{\theta e^t}{1 - \theta + \theta e^t} \Big|_{t=0} = 0; \\ g''(t) &= \frac{(1 - \theta)\theta e^t}{(1 - \theta + \theta e^t)^2} \leq \frac{1}{4}. \end{aligned}$$

因此 $g(t) \leq \frac{1}{8}t^2 = \frac{1}{8}\alpha^2(b-a)^2$. 即得 $\mathbf{E}[e^{\alpha X}] \leq \exp\left(\frac{\alpha^2(b-a)^2}{8}\right)$. □

借助霍夫丁引理, 霍夫丁不等式便很容易证明了.

霍夫丁不等式的证明. 不妨设 $\mathbf{E}[X_i] = 0$ (否则用 $X_i - \mathbf{E}[X_i]$ 代替). 由对称性只需证明 $\mathbb{P}[X \geq t] \leq \exp\left(-\frac{2t^2}{\sum(b_i - a_i)^2}\right)$. 考虑

$$\mathbb{P}[X \geq t] \leq \frac{\mathbf{E}[e^{\alpha X}]}{e^{\alpha t}} = \prod_{i=1}^n \frac{\mathbf{E}[e^{\alpha X_i}]}{e^{\alpha t/n}}.$$

应用霍夫丁引理并取 $\alpha = \frac{4t}{\sum(b_i - a_i)^2}$ 即得结论. □

1.3 独立模型与集中不等式的一些应用

最大负载 (max load) 问题

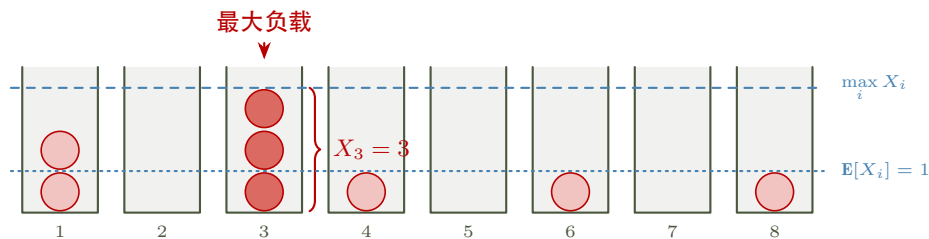


图 1.5: $n = 8$ 时的最大负载问题示意图

投球入箱问题可以被用来建模哈希表中的负载分布. 我们考虑有 m 个数据, 它们被随机地哈希到 n 个桶中. 我们关心的是最大负载, 即每个桶中数据的数量的最大值. 为了方便讨论, 我们接下来假设 $m = n$.

用投球入箱的语言来表述这个问题, 我们有 n 个球被随机地投到 n 个箱子中. 对于每个 $i \in [n]$, 我们定义随机变量 X_i 表示第 i 个桶中的球的数量. 我们要计算的是 $\max_{i \in [n]} X_i$ 的值.

在这之前, 我们首先计算一下 $\mathbf{E}[X_i]$, 即每一个箱子中平均的球的个数. 由于每一个箱子都是对称的, 我们只需要计算 $\mathbf{E}[X_1]$. 我们使用期望的线性性, 有 $\mathbf{E}[X_1] = \sum_{j=1}^n \mathbf{E}[I_j]$, 其中 I_j 是一个指示变量, 表示第 j 个球是否被投到第 1 个箱子中. 由于每个球被随机地投到 n 个箱子中, 因此 $\mathbf{P}[I_j = 1] = \frac{1}{n}$, 从而 $\mathbf{E}[I_j] = \frac{1}{n}$. 因此 $\mathbf{E}[X_1] = n \cdot \frac{1}{n} = 1$. n 个球投入 n 个箱子中, 每个箱子平均有 1 个球, 这很合理. 但是, 我们马上可以看到, 最大负载的典型取值就不再是一个常数了, 而是 $\Theta\left(\frac{\log n}{\log \log n}\right)$.

我们现在先证明最大负载以高概率为 $O\left(\frac{\log n}{\log \log n}\right)$. 我们在后面的章节会证明其为 $\Omega\left(\frac{\log n}{\log \log n}\right)$.

设 $X = \max_{i \in [n]} X_i$, 我们需要说明存在常数 $c > 0$, 使得 $\mathbf{P}\left[X \geq \frac{c \log n}{\log \log n}\right] = o(1)$. 令 $k = \frac{c \log n}{\log \log n}$. 使用联合界 (union bound), 我们有:

$$\mathbf{P}[X \geq k] = \mathbf{P}[\exists i \in [n], X_i \geq k] \leq \sum_{i=1}^n \mathbf{P}[X_i \geq k] = n \cdot \mathbf{P}[X_1 \geq k].$$

注意到 $[X_1 \geq k]$ 这个事件意味着在 n 次投球中, 至少有 k 次投入了第 1 个箱子中. 对于 $S \subseteq [n]$, 我们用 A_S 表示事件 [对于每一个 $j \in S$, 第 j 个球投到了第 1 个箱子中], 因此我们再次使用联合界以及独立性得到:

$$\mathbf{P}[X_1 \geq k] = \mathbf{P}\left[\exists S \subseteq \binom{[n]}{k}, A_S\right] \leq \binom{n}{k} \cdot n^{-k} \leq \left(\frac{e}{k}\right)^k.$$

注意到

$$k \log k = \frac{c \log n}{\log \log n} (\log \log n - \log \log \log n + \log c) \sim c \log n.$$

取 $c = 6$, 我们对于充分大的 n , 有

$$\log n + k - k \log k < -\log n.$$

于是, $\mathbf{P}[X \geq k] \leq n \cdot \left(\frac{e}{k}\right)^k < \frac{1}{n} = o(1)$.

流模型计数

假设我们有一台内存极其有限的路由器 (它背后是一个大型网站), 它可以监控访问网站的所有设备 ID. 面对海量的访问数据, 它需要完成一些计算任务, 比如

- 总共有多少个设备访问了?
- 有多少个不同的设备访问了?
- 哪个设备出现的频率最高?

$\Theta\left(\frac{\log n}{\log \log n}\right)$ 是一个看起来有点奇怪的界, 但重要的是, 它是一个当 n 趋向于无穷的时候趋向无穷的界.

联合界 (union bound) 又称布尔不等式 (Boole's inequality), 指的是对于任意事件 A_1, A_2, \dots, A_n , 有 $\mathbf{P}\left[\bigcup_{i=1}^n A_i\right] \leq \sum_{i=1}^n \mathbf{P}[A_i]$.

$\binom{[n]}{k} := \{S \subseteq [n] : |S| = k\}$.

如果我们能存下访问数据本身，那么解决这些计算问题的复杂度并不高。但是由于数据量巨大，我们无法将所有访问记录都存储下来进行离线处理，只能使用有限的内存来处理这些数据流。这便催生了流模型（streaming model）这个计算模型。在流模型中，输入是一个序列 $\sigma = (a_1, a_2, \dots, a_m)$ ，其中每个 $a_i \in [n]$ 。顾名思义，“流”意味着数据是逐个到达的。

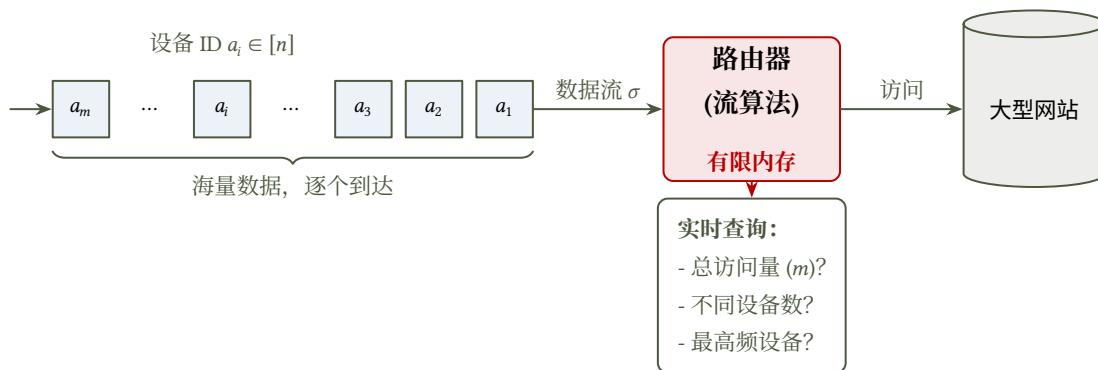


图 1.6: 流模型示意图

我们现在聚焦于最基本的问题：数据流的长度是多少（即 m 的值）？最简单的想法是维护一个计数器 k 。每当一个元素到达时，执行 $k \leftarrow k + 1$ 。不难看出，这需要 $\lceil \log_2 m \rceil$ 比特的内存。

我们能否设计一个更聪明的算法，仅使用 $o(\log m)$ 的内存？实际上 $\lceil \log_2 m \rceil - 1$ 比特的内存也是不够的。我们可以使用鸽巢原理来证明这件事情：

任何状态数少于 m 的确定性算法，必定在处理某两个不同长度 $i \neq j$ 的前缀时陷入相同的内部状态，从而无法区分出它们。

我们将在练习中请读者严格的证明这个结论。

这个事实说明，如果希望精确的统计 m ，那么上面简单维护计数器的方法就是最优的了。但是，如果我们允许近似，情况就很不一样。严格的说，对于给定的误差参数 $\varepsilon > 0$ ，我们希望算法计算出一个估计值 \hat{m} ，使得以高概率有

$$1 - \varepsilon \leq \frac{\hat{m}}{m} \leq 1 + \varepsilon$$

成立。如果把算法的要求放松成这样，我们能否得到更省内存的算法呢？

Morris 算法 (Morris' algorithm) Robert Morris 在 1978 年提出了一个简单但是巧妙的随机算法。

算法：莫里斯计数算法

输入：数据流序列 $\sigma = \langle a_1, a_2, \dots, a_m \rangle$ ，其中 $a_i \in [n]$ 。

输出：序列长度 m 的估计值。

1. 初始化计数器 $X \leftarrow 0$ ；
2. 当每个元素到达时：以 2^{-X} 的概率执行 $X \leftarrow X + 1$ ；
3. 返回估计值 $\hat{m} = 2^X - 1$ 。

一个简单的观察是， X 大概是 \hat{m} 的对数级别大小。因此，如果 $\hat{m} \approx m$ ，那么 X 就是一个 $\log m$ 级别大小的数，从而总体占用 $O(\log \log m)$ 比特的内存就可以实现它。这相当于我们之前的精确算法是指数级的提升。我们接下来就是要逐渐在概率的意义下证明 $\hat{m} \approx m$ 这件事，并且精确的给出实现算法所需要占用的内存。

命题 1.10 莫里斯算法的无偏性。

莫里斯算法的输出 \hat{m} 是 m 的无偏估计，即 $\mathbf{E}[\hat{m}] = m$ 。

证明。 我们对 m 进行归纳。当 $m = 1$ 时，经过首次处理必定有 $X = 1$ ，故 $\mathbf{E}[\hat{m}] = 2^1 - 1 = 1$ ，基础情况成立。假设该结论对于 $k < m$ 均成立，这等价于对于 $k < m$ ， $\mathbf{E}[2^{X_k}] = k + 1$ ，其中 X_k 表示处理完第 k 个输入后计数器 X 的值。利用 LOTUS 与归纳假设，有：

$$\begin{aligned}
 \mathbf{E}[\hat{m}] &= \mathbf{E}[2^{X_m}] - 1 \\
 \triangleright \text{LOTUS} \quad &= \sum_{i=0}^m \mathbf{P}[X_m = i] \cdot 2^i - 1 \\
 \triangleright \text{算法定义} \quad &= \sum_{i=0}^m \left(\mathbf{P}[X_{m-1} = i] \cdot (1 - 2^{-i}) + \mathbf{P}[X_{m-1} = i - 1] \cdot 2^{-(i-1)} \right) \cdot 2^i - 1 \\
 &= \sum_{i=0}^{m-1} \mathbf{P}[X_{m-1} = i] \cdot (2^i + 1) - 1 \\
 &= \mathbf{E}[2^{X_{m-1}}].
 \end{aligned}$$

使用归纳假设，我们有 $\mathbf{E}[2^{X_{m-1}}] = m - 1 + 1 = m$ 。 □

至此，我们证明了其无偏性，也就是说，算法输出的结果，在平均的意义下和正确的结果是一致的。然而，一个实用的随机算法必须额外满足集中性，即输出以高概率集中在期望附近，即对于小量 $\varepsilon, \delta > 0$ ，我们希望满足

$$\mathbf{P}[|\hat{m} - m| \geq \varepsilon m] \leq \delta.$$

LOTUS (Law of the unconscious statistician) 指的是对于任意随机变量 X 和 (可测) 函数 f ，有 $\mathbf{E}[f(X)] = \sum_x f(x) \cdot \mathbf{P}[X = x]$ 。

这是我们之前学习过的集中不等式的形态，我们可以使用，比如切比雪夫不等式来分析莫里斯算法的集中性。为了应用切比雪夫不等式，我们需要计算 \hat{m} 的方差，这需要先计算二阶矩。

引理 1.11.

对于莫里斯算法，有 $\mathbf{E}[(2^{X_m})^2] = \frac{3}{2}m^2 + \frac{3}{2}m + 1$ 。

证明. 我们同样对 m 使用归纳法。当 $m = 1$ 时， $\mathbf{E}[(2^{X_1})^2] = 4$ ，符合公式。假设该结论对于所有 $k < m$ 成立：

$$\begin{aligned} \mathbf{E}[(2^{X_m})^2] &= \sum_{i=0}^m \mathbf{P}[X_m = i] \cdot 2^{2i} \\ \triangleright \text{算法定义} \quad &= \sum_{i=0}^m \left(\mathbf{P}[X_{m-1} = i] \cdot (1 - 2^{-i}) + \mathbf{P}[X_{m-1} = i - 1] \cdot 2^{-(i-1)} \right) \cdot 2^{2i} \\ &= \sum_{i=0}^{m-1} \mathbf{P}[X_{m-1} = i] \cdot (2^{2i} + 3 \cdot 2^i) \\ &= \mathbf{E}[(2^{X_{m-1}})^2] + 3\mathbf{E}[2^{X_{m-1}}]. \end{aligned}$$

使用归纳假设以及之前证明的无偏性，我们有：

$$\mathbf{E}[(2^{X_m})^2] = \left(\frac{3}{2}(m-1)^2 + \frac{3}{2}(m-1) + 1 \right) + 3m = \frac{3}{2}m^2 + \frac{3}{2}m + 1.$$

□

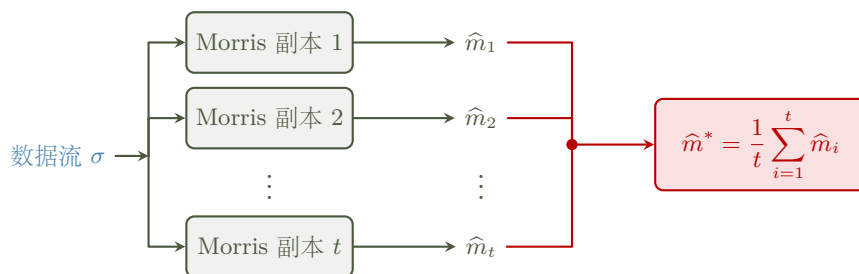
借助这个引理，我们可以直接计算方差：

$$\mathbf{Var}[\hat{m}] = \mathbf{E}[\hat{m}^2] - (\mathbf{E}[\hat{m}])^2 = \mathbf{E}[(2^{X_m} - 1)^2] - m^2 \leq \frac{m^2}{2}.$$

应用切比雪夫不等式 (定理 1.3)，对于任意 $\varepsilon > 0$ ，有：

$$\mathbf{P}[|\hat{m} - m| \geq \varepsilon m] \leq \frac{\mathbf{Var}[\hat{m}]}{\varepsilon^2 m^2} \leq \frac{1}{2\varepsilon^2}. \quad (1.1)$$

仔细观察式 (1.1)，可以看到这个集中不等式的失败概率上限是 $\frac{1}{2\varepsilon^2}$ ，这对于 $\varepsilon < \frac{1}{\sqrt{2}}$ 就已经大于 1 了，从而变得毫无意义。因此，这是一个很差的界。究其原因，在于 \hat{m} 事实上就是一个方差很大的随机变量，我们直接使用它作为算法的输出本身就是很不准确。为了让算法真正有效，我们必须引入系统性的方法来放大算法的集中度，或者说减小算法输出的方差。

图 1.7: 均值技巧: 独立并行运行 t 次, 取均值

均值技巧 (averaging trick) 一个非常直观的让算法变得更稳定的方式是运行多个独立的副本, 然后取它们的平均值作为最终输出. 这件事情本质上是减小了输出的方差, 我们称之为均值技巧. 现在, 我们来严格的分析它.

我们设计一个新算法: 独立且并行地运行莫里斯算法 t 次, 记输出分别为 $\hat{m}_1, \dots, \hat{m}_t$. 最终输出取均值:

$$\hat{m}^* := \frac{1}{t} \sum_{i=1}^t \hat{m}_i.$$

由独立性, $\mathbf{Var}[\hat{m}^*] = \frac{1}{t} \mathbf{Var}[\hat{m}_1]$. 再次应用切比雪夫不等式可得:

$$\mathbb{P}[|\hat{m}^* - m| \geq \epsilon m] \leq \frac{1}{t \cdot 2\epsilon^2}. \quad (1.2)$$

为了将失败概率控制在 δ 以下, 我们只需取 $t \geq \frac{1}{2\epsilon^2\delta}$. 这个新算法消耗了 $O\left(\frac{\log \log m}{\epsilon^2 \delta}\right)$ 比特的内存. 可以看到, 算法的性能 (内存消耗) 体现了对于失败率 δ 和精度 ϵ 的权衡.

式 (1.2) 的界是很不错的, 其导出来的算法对于失败率 δ 的依赖是关于 $1/\delta$ 的线性函数. 回忆我们在 **示例 1.7** 中, 对于公平硬币检验问题, 我们可以使用切尔诺夫界得到一个掷次数关于错误率 δ 的 $\log(1/\delta)$ 级别的依赖. 虽然硬币检测问题和我们这里的问题表面上看起来很不一样, 但实际上它们具有类似的结构. 我们有一个通用的技巧把对于失败率的依赖改进到 $\log(1/\delta)$ 级别, 使用的工具依然是切尔诺夫界.

中位数技巧 (median-of-means trick) 我们在均值技巧中固定 $t = \frac{3}{2\epsilon^2}$. 由前文分析, 此时每次运行的失败概率 $\delta \leq 1/3$. 我们将这个均值算法作为一个整体, 独立并行地运行 s 次, 得到估计集合 $\{\hat{m}_1^*, \dots, \hat{m}_s^*\}$. 对于任意 $i \in [s]$, 均满足 $\mathbb{P}[|\hat{m}_i^* - m| \geq \epsilon m] \leq 1/3$. 最终, 我们输出这 s 个结果的中位数 \hat{M} .

对于每个 $i \in [s]$, 定义指示变量 $Y_i = 1[|\hat{m}_i^* - m| < \epsilon m]$. 令 $Y = \sum_{i=1}^s Y_i$, 显然 $\mathbf{E}[Y] \geq \frac{2}{3}s$.

如果最终的中位数 \hat{M} 是一个“坏”估计, 必然意味着有一半以上的子算法给出了坏估

在这里的讨论中, 我们说算法“成功”是指输出满足 $|\hat{m}^* - m| < \epsilon m$ 的事件, 否则叫做算法“失败”.

严格来说, 每个副本的计数器 X_i 占用的比特数是随机的: 存储 X_i 需要 $\lceil \log_2(X_i + 1) \rceil$ 比特, 只有当 $X_i = O(\log m)$ 时才是 $O(\log \log m)$ 比特. 但 X_i 的尾概率随 X_i 指数衰减, 对 t 个副本取联合界后, 总内存以高概率不超过所述的量. 这一额外的失败概率可以并入整体的 δ 中. 我们这里不深究这个细节.

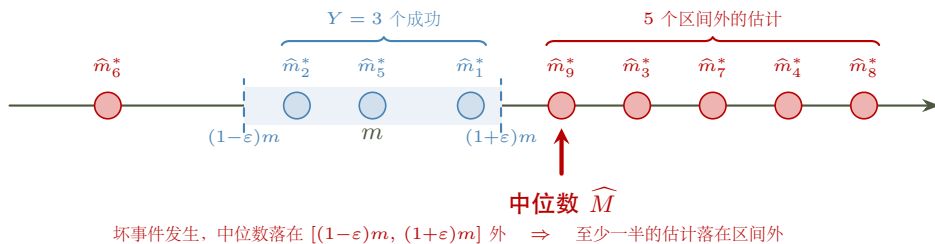


图 1.8: 中位数技巧的“坏事件”: 过半数副本失败时, 中位数落在误差区间外

计. 换言之, 成功的次数 $Y \leq s/2$. 利用推论 1.6 中的下尾切尔诺夫界 (取 $\epsilon' = 1/4$):

$$P[Y \leq s/2] \leq P[Y \leq (1 - 1/4)E[Y]] \leq \exp\left(-\frac{(1/4)^2}{2} \cdot \frac{2}{3}s\right) = \exp(-s/48).$$

为了让整体失败概率不超过 δ , 我们只需令 $\exp(-s/48) \leq \delta$. 取 $s = 48 \log \frac{1}{\delta}$ 即可.

总结起来, 结合均值与中位数技巧的最终算法, 对于任何 $\epsilon, \delta \in (0, 1)$, 以至少 $1 - \delta$ 的概率输出一个满足 $|\widehat{M} - m| < \epsilon m$ 的估计值, 并且使用的总内存为 $O\left(\frac{1}{\epsilon^2} \log \frac{1}{\delta} \log \log m\right)$ 比特.

比较重要的是, 我们使用的均值技巧和中位数技巧, 本身与莫里斯算法没有什么关系. 它们是一种通用的提高随机算法性能的方式.

同前文的讨论, 这里的内存上界也是以高概率成立的: 需要对所有 $s \cdot t$ 个 Morris 副本的计数器取联合界.

多臂老虎机问题 (multi-armed bandit)

在本节中, 我们讨论在线优化中的一个经典模型: 多臂老虎机问题 (multi-armed bandit problem). 我们通过这个问题展示一下集中不等式在算法设计和分析中的作用. 假设有一个 n 臂老虎机, 每一个臂拉一下会给出一个位于 $[0, 1]$ 范围内的随机奖励值. 为了方便, 我们假设第 i 个臂返回的奖励服从分布 $\text{Ber}(\mu_i)$. 我们进一步假设 $\mu_1 > \mu_2 \geq \dots \geq \mu_n$ 成立. 我们现在的可以拉动老虎机 T 轮 (每一轮选一个臂), 目标是使奖励期望最大化. 如果我们已知 μ_1, \dots, μ_k , 最优策略是始终选择臂 1, 此时期望奖励为 $T\mu_1$. 然而, 当我们不知道每个臂的奖励分布及顺序时, 就需要设计一种策略来首先探索老虎机.

记 a_t 为第 t 轮拉动的臂, 因此第 t 轮的奖励为 $X_t \sim \text{Ber}(\mu_{a_t})$. 一种衡量策略的好坏的方式是使用懊悔 (regret). 懊悔定义为总轮次 T 中始终选择最佳臂 1 所能获得的期望奖励与策略实际期望奖励之间的差值, 即未总是选择最佳臂所带来的代价:

$$R(T) := T\mu_1 - \mathbf{E} \left[\sum_{t=1}^T X_t \right] \geq 0.$$

在上述表达式中, 期望 $\mathbf{E}[\cdot]$ 的随机性通常来源于两个方面: 奖励分布 $\text{Ber}(\mu_i)$ 的随机性以及策略本身的随机性. 对于每个 $i \in [n]$, 我们定义 $\Delta_i := \mu_1 - \mu_i \geq 0$, 表示第 i 臂的奖励均值与最优臂的奖励均值之间的差距. 如果我们的策略是平均的拉每一个臂, 那么懊悔为

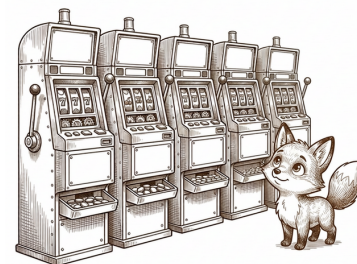


图 1.9: 多臂老虎机问题

$R(T) = \frac{\sum_{i=1}^n \Delta_i}{n} \cdot T$. 我们可以选择 Δ_i 使得这是一个关于 T 线性增长的, 因此, 不是一个好的策略. 接下来, 我们介绍一种名为“探索然后确定”(explore-then-commit, ETC) 的算法, 该算法可以实现次线性懊悔.

探索然后确定 (explore-then-commit, ETC) 算法 为了使懊悔最小化, 策略应尽快识别出最佳臂. 最直接的方法是给每个臂尝试一定次数, 然后选择奖励的**经验均值** (empirical mean) 最大的臂. ETC 算法便是实现了这一思想: 每个臂 i 被拉动 L 次 (总共进行 nL 次探索), 并计算 $\hat{\mu}_i$ (在 L 次尝试中获得的平均奖励). 之后, 始终选择 $\hat{\mu}_i$ 最大的臂. 其懊悔可以写为:

$$\begin{aligned} R(T) &= L \sum_{i=1}^n \Delta_i + \sum_{i=2}^n \Delta_i \cdot \sum_{t=nL+1}^T \mathbb{P} \left[\hat{\mu}_i > \max_{j \neq i} \hat{\mu}_j \right] \\ &= L \sum_{i=1}^n \Delta_i + \sum_{i=2}^n \Delta_i \cdot (T - nL) \mathbb{P} \left[\hat{\mu}_i > \max_{j \neq i} \hat{\mu}_j \right]. \end{aligned}$$

当 $i \neq 1$ 时:

$$\mathbb{P} \left[\hat{\mu}_i > \max_{j \neq i} \hat{\mu}_j \right] \leq \mathbb{P} [\hat{\mu}_i > \hat{\mu}_1].$$

我们用集中不等式来界定上述概率. 为此, 令 X_j 为探索阶段中第 j 次拉动臂 i 时的奖励, Y_j 为探索阶段中第 j 次拉动臂 1 时的奖励. 令 $Z_j = X_j - Y_j \in [-1, 1]$, 则 $\mathbf{E}[Z_j] = -\Delta_i \leq 0$. 令 $Z = \sum_{j=1}^L Z_j$, 则 $\mathbf{E}[Z] = -L\Delta_i$.

根据霍夫丁不等式:

$$\mathbb{P} [\hat{\mu}_i > \hat{\mu}_1] = \mathbb{P} [Z > 0] = \mathbb{P} [Z - \mathbf{E}[Z] > L\Delta_i] \leq \exp \left(-\frac{2(L\Delta_i)^2}{\sum_{j=1}^L 2^2} \right) = \exp \left(-\frac{L\Delta_i^2}{2} \right).$$

因此有:

$$\begin{aligned} R(T) &\leq L \sum_{i=1}^n \Delta_i + (T - nL) \sum_{i=2}^n \Delta_i \exp \left(-\frac{L\Delta_i^2}{2} \right) \\ \triangleright \Delta_i \leq 1 &\leq \sum_{i=1}^n \left(L\Delta_i + T\Delta_i \exp \left(-\frac{L\Delta_i^2}{2} \right) \right). \end{aligned}$$

接下来定义:

$$g(L, \Delta_i) := L + T\Delta_i \exp \left(-\frac{L\Delta_i^2}{2} \right).$$

我们希望找到 L 使得 $R(T)$ 的上界最小化, 即 $\min_L \max_{\Delta_i} R(T)$. 首先计算 $\max_{\Delta_i} g(L, \Delta_i)$:

$$\frac{\partial}{\partial \Delta_i} g(L, \Delta_i) = T(1 - L\Delta_i^2) \exp \left(-\frac{L\Delta_i^2}{2} \right).$$

当 $0 \leq \Delta_i < \frac{1}{\sqrt{L}}$ 时, $\frac{\partial}{\partial \Delta_i} g(L, \Delta_i) > 0$; 当 $1 \geq \Delta_i > \frac{1}{\sqrt{L}}$ 时, $\frac{\partial}{\partial \Delta_i} g(L, \Delta_i) < 0$.

因此, 对于所有 $L > 1$:

$$g(L, \Delta_i) \leq g\left(L, \frac{1}{\sqrt{L}}\right) = L + \frac{Te^{-1/2}}{\sqrt{L}}. \quad (1.3)$$

通过令 $L = \Theta(T^{\frac{2}{3}})$, 最终可以得到

$$R(T) \leq \sum_{i=1}^n \left(L + \frac{Te^{-1/2}}{\sqrt{L}} \right) = \Theta(nT^{\frac{2}{3}}).$$

更好的懊悔 ETC 算法实现了 $\Theta(T^{\frac{2}{3}})$ 的次线性懊悔, 这是不错的结果, 但仍然不是最优的. 其主要缺点在于, 在探索阶段, 该算法对所有臂一视同仁, 无论已经获得的奖励如何, 每个臂都固定被尝试 L 次. 我们可以更加精妙的应用集中不等式, 构造一个称之为上置信界 (*upper confidence bound, UCB*) 的算法. 它可以达到 $\tilde{O}(T^{\frac{1}{2}})$ ¹ 的懊悔, 我们将在本章习题中讨论这个算法.

¹ $\tilde{O}(\cdot)$ 忽略了关于 T 的 \log 项

第二章 有限状态马尔可夫链

大家一定对于“洗牌”过程非常熟悉. 我们在玩扑克牌之前, 总是会把牌洗匀, 以确保每个人摸到的牌是随机的. 常用的洗牌方式有好几种, 比如

- 鸽尾洗牌法 (*dovetail shuffle*) 把牌堆大致分成两半, 然后交替地从两半中抽牌放回去.
- 顶牌随机插入法 (*top-to-random shuffle*) 每次把顶牌取出, 随机插入到牌堆中某个位置.
- 随机置顶法 (*random-to-top shuffle*) 每次随机选一张牌, 放到牌堆顶端.

生活中的经验告诉我们, 当洗的次数足够多之后, 牌堆的顺序会变得足够随机, 以至于每一种排列出现的概率都差不多, 即所谓的“洗匀了”. 但我们不禁要问, 这件事情为什么是对的呢? 更进一步的, 我们要洗多少次牌才能够让牌序足够均匀? 这些洗牌的策略有没有效率上的优劣之分? 为了精确回答的这个问题, 我们需要对洗牌的过程进行数学建模, 而这里用到的模型正是马尔可夫链 (*Markov chain*).

2.1 离散马尔可夫链的基本概念

我们可以仔细地观察洗牌的过程. 假设总共有 n 张牌, 我们用 $1, \dots, n$ 来给它们编号. 我们可以用 X_t 来表示进行了 t 次洗牌后的牌堆状态. 也就是, X_t 是 n 张牌的一个排列. 而洗牌的每一次操作, 就是从 X_t 到 X_{t+1} 的一个随机变换. 不同的洗牌方式实际上对应了这个随机变换的不同. 但这些随机变换有一个共同的性质: X_{t+1} 的分布只依赖于 X_t , 而和 $X_{t-1}, X_{t-2}, \dots, X_0$ 无关. 也就是说, 给定当前的牌堆状态, 下一次洗牌后牌堆状态的分布就已经确定了, 而不需要知道之前牌堆的状态. 我们把这种性质抽象出来, 称为马尔可夫性 (*Markov property*). 满足马尔可夫性质的随机变量序列, 就被称为马尔可夫链.

鸽尾洗牌法

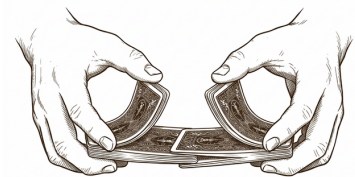


图 2.1: 鸽尾洗牌法又称对切法 (*riffle shuffle*), 是一种在数学上得到了广泛研究的洗牌方式.

在我们这门课程中, 我们总是假设这些随机变量定义在一个合适的概率空间 (Ω, \mathcal{F}, P) 上. 我们说定义在 Ω 上的随机变量 X 是离散的, 当且仅当 $\{X(\omega) : \omega \in \Omega\}$ 是一个可数集合.

定义 2.1 离散马尔可夫链.

假设有一列离散随机变量 $X_0, X_1, \dots, X_t, \dots$, 其中每一个 X_t 的取值都来自于可数集合 S . 如果 $\forall t \geq 1$, 随机变量 X_t 的分布只依赖于 X_{t-1} , 即 $\forall a_0, a_1, \dots, a_t \in S$,

$$P[X_t = a_t | X_{t-1} = a_{t-1}, \dots, X_1 = a_1, X_0 = a_0] = P[X_t = a_t | X_{t-1} = a_{t-1}].$$

那么, 我们称 $\{X_t\}$ 为离散马尔可夫链.

在这个定义里, “离散”这个修饰词是为了强调我们讨论的马尔可夫链的状态空间 S 是一个可数集合. 此外, 我们的时间参数 t 也是离散的 ($t = 0, 1, 2, \dots$). 如果更进一步 S 是一个有限大小的集合, 我们就称这个马尔可夫链为**有限状态马尔可夫链**.

从这个定义可以看出, 我们前面提到的不同的洗牌方式, 实际上对应了概率转移 $P[X_t = a_t | X_{t-1} = a_{t-1}]$ 的不同选择. 我们本章主要关心 S 是有限集合的场合, 因此可以不失一般性的假设 $S = [N]$, 其中 $N \in \mathbb{N}$. 于是, 对于每一个 $t \geq 1$, 我们可以用一个 $N \times N$ 的矩阵 $P^{(t)} = (p_{ij}^{(t)})$ 来表示从时间 $t-1$ 到时间 t 的转移概率, 其中

$$p_{ij}^{(t)} = P[X_t = j | X_{t-1} = i].$$

在我们洗牌的例子里, 由于每一次洗牌的方式都是一样的, 因此 $P^{(t)}$ 实际上和 t 无关, 我们可以把它简记为 P . 满足这个性质的马尔可夫链, 称为时间齐次(*time-homogeneous*) 马尔可夫链. 矩阵 P 常常被称为马尔可夫链的转移矩阵, 它唯一刻画了时间齐次马尔可夫链每一步的行为.

有的时候为了更形象的表明“转移概率”, 我们会用 $P(i \rightarrow j)$ 来表示 p_{ij} .

在转移图里, 当我们说存在一条有向边 (i, j) 时, 一般默认了 $p_{ij} > 0$.

我们可以等价的用一个有 N 个顶点的加权有向图来表示 P , 其中边 (i, j) 的权即为 p_{ij} . 我们称它为转移图.

示例 2.2 有限状态随机游走.

图 2.2 为转移矩阵

$$P = (p_{ij}) = \begin{bmatrix} 1/2 & 3/8 & 1/8 \\ 1/3 & 0 & 2/3 \\ 1/4 & 3/4 & 0 \end{bmatrix}$$

对应的转移图.

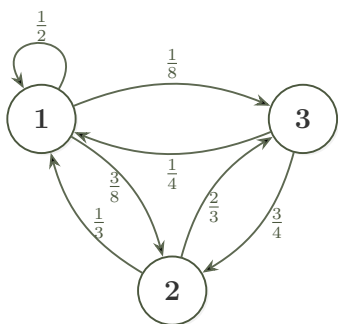


图 2.2: 转移矩阵 P 对应的转移图

我们接下来看一些马尔可夫链的例子.

示例 2.3 \mathbb{Z} 上的随机游走.

考虑 \mathbb{Z} 上的随机游走. 初始位置为 0, 每轮抛掷一枚公平硬币决定移动方向: 以 50% 的概率向左移动, 以 50% 的概率向右移动. 如果用 X_t 表示时间 t 时的位置, 则 $X_0 = 0$, 对于每个 $t > 0$, 有 $X_t = X_{t-1} + 1$ 的概率为 50%, 以及 $X_t = X_{t-1} - 1$ 的概率为 50%. 这是一个马尔可夫链, 因为时间 t 时的位置只依赖于时间 $t-1$ 时的位置. 该链的状态空

间是可数无限集 \mathbb{Z} .

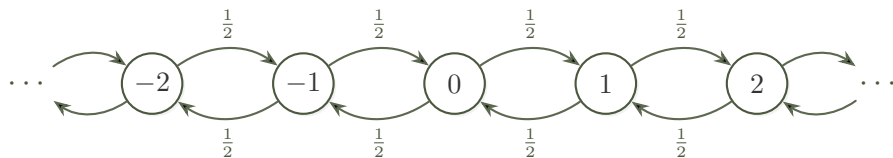


图 2.3: \mathbb{Z} 上的简单随机游走的转移图

示例 2.4 $\{\pm 1\}^d$ 上的随机游走.

考虑 d 维超立方体 (hypercube) 上的随机游走. 状态空间为 $V = \{-1, +1\}^d$, 包含 2^d 个顶点, 每个状态表示为一个长度为 d 的向量 $x = (x_1, \dots, x_d)$, 其中 $x_i \in \{-1, +1\}$. 如果在时间 $t-1$ 时的位置为 X_{t-1} , 则时间 t 的位置 X_t 通过以下方式确定: 从集合 $\{1, \dots, d\}$ 中均匀随机地选取一个坐标索引 k , 并将 X_{t-1} 的第 k 个分量符号翻转 (即乘以 -1), 其余分量保持不变. 这等价于从当前顶点的 d 个邻居中均匀随机地选择一个作为下一步的位置. 这是一个有限状态空间上的马尔可夫链.

示例 2.5 网页浏览.

考虑一个在网上随机浏览的网页的人. 假设状态空间 V 由所有网页组成, 这是一个巨大的有限集. 该随机游走的转移规则由两部分混合而成: 设当前位于页面 i , 且页面 i 有 k_i 个出链. 在每一步, 浏览者以概率 α 从当前页面的 k_i 个出链中均匀随机地选择一个链接跳转; 以概率 $1 - \alpha$ 从整个状态空间 V 中均匀随机地 (或根据特定的个性化分布) 选择一个页面跳转. 这种对浏览网页方式的建模方式被称为随机冲浪者 (random surfer) 模型. 它是搜索引擎 PageRank 算法量化网页重要性的基础.

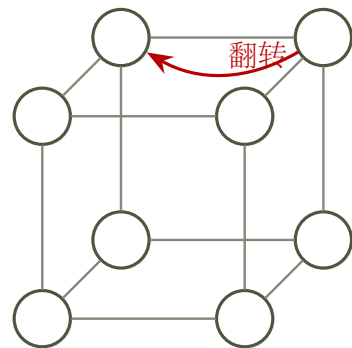


图 2.4: $d = 3$ 时的超立方体图, 其中红色箭头示意一步“翻转”操作

既然使用马尔可夫链可以用来描述这么多具体的问题, 那么, 我们自然也希望能够用它来分析问题的性质. 比如说, 我告诉你一副牌初始是按照顺序排列的, 那我用鸽尾洗牌法洗了三回之后, 牌堆的顺序是什么样子的呢? 显然, 这个顺序是随机的, 那么它的分布是什么? 根据我们刚才的讨论, 对于时间齐次的马尔可夫链, 转移矩阵 P 完全的描述了马尔可夫链每一步的行为. 因此, 我们可以使用线性代数的工具通过研究矩阵 P 来分析马尔可夫链的性质. 我们接下来讨论一些基本的结论.

对概率论熟悉的读者会发现我们这里的“分布” μ 和概率论里面的“分布”有一点不一样，后者是定义在所有事件上的函数。我们这儿的分布实际上是离散场合的概率质量函数 (probability mass function)，这是一个被普遍接受的“记号滥用”。

使用线性代数

一个贯穿本书始终的记号是我们会使用 μ_t 来表示马尔可夫链里 X_t 的分布。也就是说，

$$\forall i \in [N], \mu_t(i) = \mathbb{P}[X_t = i].$$

因此， μ 是一个定义在 $[N]$ 上的，取值为 $[0, 1]$ 的函数，并且满足 $\sum_{i \in [N]} \mu(i) = 1$ 。我们有的时候也会把它等价的看成是 $[0, 1]^{[N]}$ 里的一个列向量，即 $\mu = (\mu(1), \mu(2), \dots, \mu(N))^T$ 。

根据全概率公式，有 $\mu_{t+1}(j) = \sum_i \mu_t(i) \cdot p_{ij}$ 成立。把它写成矩阵的形式，便是

$$\mu_t^T P = \mu_{t+1}^T.$$

这便给了我们一个从任何 μ_s ($s \geq 0$) 出发计算 μ_{s+t} 的公式

$$\mu_{s+t}^T = \mu_s^T P^t. \quad (2.1)$$

从式 (2.1) 也可以看出来，矩阵 P^t 直观上可以解释成马尔可夫链走 t 步的转移矩阵。使用矩阵乘法的结合律，我们显然有等式

$$\forall s, t \geq 0: \quad P^{s+t} = P^s P^t. \quad (2.2)$$

这个等式在随机过程的文献里被称为查普曼-科尔莫戈罗夫 (Chapman-Kolmogorov) 方程。

由此可见，每个时间点的分布完全由转移矩阵 P 和初始分布 μ_0 决定。

简称为 C-K 方程。我们未来会看到它以多种形态出现。

马尔可夫链的平稳分布 (stationary distribution)

在洗牌的例子里，我们关注的最基本的问题是，经过多次洗牌后，牌堆的顺序是否会接近于均匀随机？这就引出了马尔可夫链的一个重要概念：平稳分布。对于洗牌来说，答案是肯定的，其原因就在于均匀分布是洗牌这个马尔可夫链所对应的唯一的平稳分布，并且从任何一个分布出发，执行足够多的操作之后，均会收敛到它。这样一个看似简单的过程，实际上就是所谓的马尔可夫链蒙特卡罗方法 (Markov chain Monte Carlo, MCMC) 的核心思想，我们在本书中会详细的讨论这个方法。

便让我们从平稳分布的定义开始。

MCMC 方法将贯穿本书始终，敬请期待。

定义 2.6 平稳分布.

考虑转移矩阵为 P 的马尔可夫链。如果一个分布 π 在马尔可夫链中随时间推移保持不变，即满足

$$\pi^T P = \pi^T,$$

那么称 π 为 P 的平稳分布。

示例 2.7 平稳分布的例子.

我们逐一讨论上一节中各个例子的平稳分布.

1. 有限状态随机游走. 回忆转移矩阵

$$P = \begin{bmatrix} 1/2 & 3/8 & 1/8 \\ 1/3 & 0 & 2/3 \\ 1/4 & 3/4 & 0 \end{bmatrix}.$$

我们要求解 $\pi^\top P = \pi^\top$, 结合 $\sum_i \pi(i) = 1$. 取后面两个方程:

$$\begin{aligned} \frac{3}{8}\pi(1) + \frac{3}{4}\pi(3) &= \pi(2), \\ \frac{1}{8}\pi(1) + \frac{2}{3}\pi(2) &= \pi(3). \end{aligned}$$

解方程可得 $\pi(1) : \pi(2) : \pi(3) = 16 : 15 : 12$, 即

$$\pi = \left(\frac{16}{43}, \frac{15}{43}, \frac{12}{43} \right)^\top.$$

2. \mathbb{Z} 上的随机游走. 如果 π 是平稳分布, 则对每个 $i \in \mathbb{Z}$,

$$\pi(i) = \frac{1}{2}\pi(i-1) + \frac{1}{2}\pi(i+1),$$

即 $\pi(i+1) - \pi(i) = \pi(i) - \pi(i-1)$. 这意味着差分 $\pi(i+1) - \pi(i)$ 是常数, 也就是 π 是关于 i 的一次函数: $\pi(i) = a + bi$. 但 π 必须非负且可求和 ($\sum_{i \in \mathbb{Z}} \pi(i) = 1$), 所以只能 $b = 0$, 即 π 是常值函数. 然而常值函数在 \mathbb{Z} 的求和是无穷大. 因此, \mathbb{Z} 上的简单随机游走不存在平稳分布.

3. $\{\pm 1\}^d$ 上的随机游走. 超立方体图是 d -正则图 (每个顶点恰好有 d 个邻居), 因此由对称性可以猜测均匀分布 $\pi(x) = \frac{1}{2^d}$ 是平稳分布. 直接验证: 每个状态 x 恰好有 d 个邻居, 从每个邻居出发走到 x 的概率为 $\frac{1}{d}$, 因此

$$[\pi^\top P](x) = d \cdot \frac{1}{2^d} \cdot \frac{1}{d} = \frac{1}{2^d} = \pi(x).$$

4. 网页浏览 (随机冲浪者模型). 在这个模型中, 浏览者每一步都有 $1-\alpha$ 的概率等概率地跳转到任意页面 (包括自身), 因此从任何页面出发一步内都有正概率到达任何其他页面. 如果平稳分布存在且唯一, 那么它就是著名的 *PageRank* 向量: Google 搜索引擎正是通过计算这个向量来量化每个网页的“重要性”. 由于状态空间 (即所有网页的集合) 非常巨大, 直接求解线性方程 $\pi^\top P = \pi^\top$ 是不可行的. 我们后面将看到, 通过不断模拟该马尔可夫链来逼近 π 是一种可行的做法——这正是 MCMC 方法的核心思想.

一般来说, $\pi^\top P = \pi^\top$ 的 N 个方程中总有一个是冗余的 (为什么?), 所以我们需要利用归一化条件 $\sum_i \pi(i) = 1$ 来确定 π .

这个现象的根源在于 \mathbb{Z} 是无穷状态空间. 我们将在下面证明, 有限状态马尔可夫链一定存在平稳分布.

我们刚才看到, 有的马尔可夫链有平稳分布 (如有限状态随机游走和超立方体上的随机游走), 有的则没有 (如 \mathbb{Z} 上的随机游走). 让我们回到洗牌的例子来思考更多的问题. 为什么我们平时常用的那些洗牌手法是有效的? 为什么洗了足够次之后我们相信牌的顺序就会足够随机? 假设我们想出老千, 希望洗完牌后牌的顺序接近我们想要的某种顺序, 应该怎么办? 这些问题都可以归结于马尔可夫链平稳分布的性质. 我们用数学的语言把它们抽象出来, 便是以下几个问题:

- (1) 每个马尔可夫链是否都有平稳分布?
- (2) 如果一个马尔可夫链有平稳分布, 它是否唯一?
- (3) 如果链有唯一的平稳分布, 是否对于任意初始分布 μ_0 , μ_t 总是收敛到它?
- (4) 如果 μ_t 总是收敛到平稳分布, 收敛的速率是多少?
- (5) 如何设计一个马尔可夫链, 使得它的平稳分布是某个指定的分布?

在本书中, 我们将尝试回答这些问题.

2.2 马尔可夫链基本定理

马尔可夫链基本定理回答了上面的前三个问题, 即马尔可夫链平稳分布的存在性、唯一性以及收敛性. 它有很多种证明的方法, 有意思的是, 大部分证明它的方法都可以在一定程度上来回答第四个问题, 即收敛的速率.

平稳分布的存在性

这里要求 π 非负即可, 因为我们总是可以把它除掉 $\|\pi\|_1$ 之后变成一个分布.

这是由于对于任何 i , $\sum_{j \in [N]} P(i \rightarrow j) = 1$.

我们将证明, 对于每个有限的马尔可夫链 P , 总存在某个 π 使得 $\pi^\top P = \pi^\top$. 注意, 这等价于“1 是 P^\top 的一个特征值, 并且它有一个非负特征向量 ($P^\top \pi = \pi$)”. 我们首先注意到, 矩阵 P 满足 $P\mathbf{1} = \mathbf{1}$, 即 1 是 P 的特征值, 而 P 与 P^\top 有同样的特征多项式, 因此, 1 也是 P^\top 的特征值. 于是, 我们可以找到一个向量 \mathbf{v} , 满足 $P^\top \mathbf{v} = \mathbf{v}$. 但是, 这个 \mathbf{v} 并不一定是非负的, 我们现在说明, 如果我们定义 $\pi(i) = |\mathbf{v}(i)|$, 那么 π 是 P^\top 对应于 1 的一个特征向量.

对于任何 $i \in [N]$, 我们可以验证,

$$\pi(i) = |\mathbf{v}(i)| = \left| \sum_{j \in [N]} \mathbf{v}(j) \cdot P(j \rightarrow i) \right| \leq \sum_{j \in [N]} |\mathbf{v}(j)| \cdot P(j \rightarrow i) = \sum_{j \in [N]} \pi(j) \cdot P(j \rightarrow i).$$

我们只需要说明, 对于每一个 i , 上面式子中的 \leq 都必须取到等号即可. 我们用反证法先假设某一个 i 没有取到等号, 那么对所有 $\mathbf{v}(i)$ 求和, 有

$$\begin{aligned} \left| \sum_{i \in [N]} |\mathbf{v}(i)| \right| &< \left| \sum_{i \in [N]} \sum_{j \in [N]} |\mathbf{v}(j)| \cdot P(j \rightarrow i) \right| \\ &= \left| \sum_{j \in [N]} |\mathbf{v}(j)| \cdot \sum_{i \in [N]} P(j \rightarrow i) \right| \\ \triangleright \sum_{i \in [N]} P(j \rightarrow i) &= 1 \quad = \left| \sum_{j \in [N]} |\mathbf{v}(j)| \right|. \end{aligned}$$

但这显然是矛盾的, 因为上述式子的头和尾是同一个东西. 因此, π 是 P^\top 的一个非负的特征向量. 于是, $\frac{\pi}{\|\pi\|_1}$ 是马尔可夫链 P 的一个平稳分布.

唯一性与收敛性

我们接着来讨论平稳分布的唯一性, 以及在已知平稳分布唯一的场合下, 是否能够从任意分布出发收敛到它. 当然了, 我们现在还没有严格定义分布的“收敛”的概念, 在这一节, 我们不妨把它想象成分布向量在每一个维度上的收敛, 也就是说, 对于定义在 $[N]$ 上的分布 μ_t , 我们说它收敛到分布 π , 如果对于每一个 $i \in [N]$, 都有

$$\lim_{t \rightarrow \infty} \mu_t(i) = \pi(i).$$

我们从非平凡的最简单的马尔可夫链, 也就是只有两个状态的马尔可夫链出发来研究这个问题. 所有的两个状态的马尔可夫链都可以被两个参数 $p, q \in [0, 1]$ 来描述, 其状态转移图如图 2.5 所示.

很明显, 该马尔可夫链的转移矩阵为 $P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$. 并且, 容易验证 $\pi = \left(\frac{q}{p+q}, \frac{p}{p+q} \right)^\top$ 是 P 的一个平稳分布. 我们将验证是否从任意初始分布 μ_0 出发, 分布 μ_t 总是收敛到 π . 在我们的例子里, 由于分布只有两个维度, 并且两维之和等于 1, 因此我们只需检查第一维是否收敛, 即

$$\lim_{t \rightarrow \infty} |\mu_0^\top P^t(1) - \pi(1)| \rightarrow 0$$

是否成立. 现在我们定义 $\Delta_t := |\mu_t(1) - \pi(1)|$, 并研究其如何随着 t 变化. 根据式 (2.1), $\mu_t^\top = \mu_{t-1}^\top P$, 因此我们有

$$\Delta_t = \left| \mu_{t-1}^\top \cdot P(1) - \pi(1) \right| = \left| (1-p) \cdot \mu_{t-1}(1) + q \cdot (1 - \mu_{t-1}(1)) - \frac{q}{p+q} \right|.$$

整理可得 $\Delta_t = |1-p-q| \cdot \Delta_{t-1}$. 因此可以看出, 除非 $p=q=0$ 或者 $p=q=1$, $\Delta_t \rightarrow 0$ 总成立.

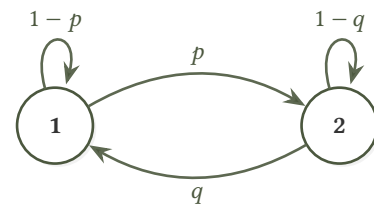


图 2.5: 两个状态的马尔可夫链的转移图

事实上,这两种情况分别对应了两种不收敛的原因,我们分别来讨论.

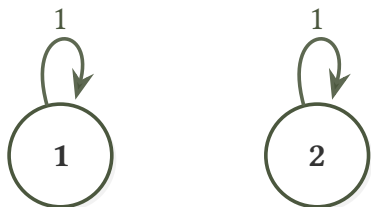


图 2.6: $p = q = 0$ 时候的转移图

给定 $[N]$ 上的两个分布 μ_1 和 μ_2 , 我们说 μ 是它们的混合 (mixture) 当且仅当存在某个 $\alpha \in [0, 1]$ 使得 $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$. 换句话说, 我们可以以如下方式从 μ 中采样一个样本: 以概率 α 从 μ_1 中采样, 以概率 $1 - \alpha$ 从 μ_2 中采样 x .

值得注意的是, 不可约并不是具有唯一平稳分布的必要条件. 考虑 $p = 1, q = 0$ 的情况, 这个马尔可夫链具有唯一的平稳分布 $(0, 1)^T$. 你能根据这个例子想到平稳分布唯一的充分必要条件是什么? 可以回忆一下算法课学过的对于有向图的强连通分量的有向无环图分解.

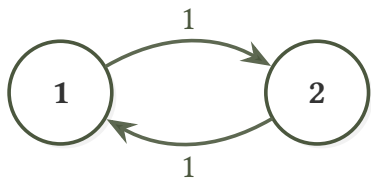


图 2.7: $p = q = 1$ 时候的转移图

这儿对于一个集合 $S \subseteq \mathbb{N}$, $\gcd(S)$ 表示的是 S 中数的最大公约数.

$p = q = 0$ 的情况

首先考虑 $p = q = 0$ 的情况. 此时马尔可夫链如图 2.6 所示.

这个马尔可夫链的转移图是不连通的, 因此可以划分为两个不相交的子图, 每个子图仍然是一个马尔可夫链, 并且各自都有自己的平稳分布. 稍微思考一下可以发现, 这些子分布的任何混合 (mixture) 都是整个马尔可夫链的平稳分布. 因此, 在这种情况下平稳分布不是唯一的. 我们想把这种平稳分布不唯一的原因进行推广, 便促使我们给出如下定义:

定义 2.8 可约与不可约.

如果一个有限马尔可夫链的转移图是强连通的, 我们称该马尔可夫链是不可约的. 如果转移图不是强连通的, 我们称其为可约的.

我们未来会说明, 不可约性是具有唯一平稳分布的充分条件.

$p = q = 1$ 的情况

我们接着来考察 $p = q = 1$ 时会发生什么. 这个时候马尔可夫链如图 2.7 所示.

这个马尔可夫链的状态转移图是二分图. 易见 $(\frac{1}{2}, \frac{1}{2})$ 是其唯一的平稳分布. 然而, 对于初始分布 $\mu_0 = (1, 0)^T$, 可以看到 μ_t 在 $(1, 0)^T$ 和 $(0, 1)^T$ 之间振荡. 因此, 它并不总是收敛到平稳分布. 我们来推广这种“振荡”的概念.

对于一个马尔可夫链的状态 i , 我们用 C_i 来表示包含 i 的所有有向环的集合.

定义 2.9 周期性与无周期性.

如果对于马尔可夫链的一个状态 i , 满足

$$\gcd(\{|c| : c \in C_i\}) = 1,$$

那么我们就称 i 是无周期性 (aperiodic) 的, 否则, 就称它为周期性 (periodic) 的. 如果一个马尔可夫链中的每一个状态都是无周期性的, 则称该马尔可夫链是无周期性的.

我们应该这样来理解周期性的定义. 假设一个包含点 i 的所有有向环长度的最大公约数比如是 3, 那么说明, 我们从 i 出发, 一定只有在 3 的倍数步之后, 才可能回到 i . 这个性质, 便推广了我们在两个点的例子里面提到的循环振荡的现象.

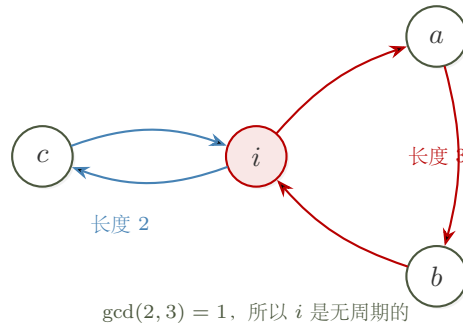


图 2.8: 状态 i 包含在一个长度为 2 的环和一个长度为 3 的环中. 由于 $\gcd(2, 3) = 1$, i 是无周期的.

实际上, 对于有限状态马尔可夫链来说, 不可约性和无周期性就保证了平稳分布是唯一存在, 并且从任意分布出发都能收敛到它. 这便是我们接下来要介绍的马尔可夫链基本定理 (*fundamental theorem of Markov chains*, F 它是我们第一个关于马尔可夫链的重要定理.

定理 2.10 马尔可夫链基本定理.

如果有限马尔可夫链 $P \in \mathbb{R}^{N \times N}$ 是不可约且无周期性的, 那么它有唯一的平稳分布 $\pi \in \mathbb{R}^N$, 并且对于任意分布 $\mu \in \mathbb{R}^N$,

$$\lim_{t \rightarrow \infty} \mu^\top P^t = \pi^\top.$$

我们之前提到过, 这里的收敛我们理解为逐点收敛, 即对于每一个 $i \in [N]$, $\lim_{t \rightarrow \infty} (\mu^\top P^t)(i) = \pi(i)$.

2.3 时间反演, 可逆链与 Metropolis-Hastings 算法

我们前面提到了 MCMC 算法, 也就是说给定一个目标分布 π , 希望能够设计一个马尔可夫链, 使得通过模拟它, 最终能够从 π 中进行采样. Metropolis-Hastings 算法就是一个实现这个目标的算法. 在介绍它之前, 我们首先介绍时间反演链 (time reversal chain) 和可逆 (reversible) 马尔可夫链的概念.

时间反演链 (time reversal chain)

假设一条马尔可夫链一直在平稳分布 π 下运行. 这时一个自然的问题是: 如果我们把这条链的演化录像倒放, 倒放后的过程还是马尔可夫链吗? 答案是肯定的, 我们可以直接写出它的转移概率:

定义 2.11 时间反演链.

设 P 是状态空间 $[N]$ 上的马尔可夫链的转移矩阵, π 是它的一个平稳分布, 并且满足 $\forall i \in [N], \pi(i) > 0$. 定义矩阵 \hat{P} 为

$$\forall i, j \in [N], \hat{P}(i, j) := \frac{\pi(j)P(j \rightarrow i)}{\pi(i)}. \quad (2.3)$$

我们称 \hat{P} 为 P 关于平稳分布 π 的时间反演链 (*time reversal chain*).

这个定义的来历很简单, 是贝叶斯 (Bayes) 公式的直接推论: 假设链已经处在平稳状态, 那么对任意两个状态 $i, j \in [N]$, 有:

$$\mathbb{P}[X_{t-1} = j \mid X_t = i] = \frac{\mathbb{P}[X_t = i \mid X_{t-1} = j] \mathbb{P}[X_{t-1} = j]}{\mathbb{P}[X_t = i]} = \frac{P(j \rightarrow i)\pi(j)}{\pi(i)} = \hat{P}(i, j).$$

因此, $\hat{P}(i, j)$ 恰好刻画了“如果把电影倒放, 当前看到状态 i 时, 上一帧是 j 的条件概率”. 这正是“时间反演”这个名称的来历.

我们还需要验证 \hat{P} 的确是一个合法的转移矩阵. 对任意 $i \in [N]$,

$$\sum_{j \in [N]} \hat{P}(i, j) = \sum_{j \in [N]} \frac{\pi(j)P(j \rightarrow i)}{\pi(i)} = \frac{1}{\pi(i)} \sum_{j \in [N]} \pi(j)P(j \rightarrow i) = \frac{\pi(i)}{\pi(i)} = 1.$$

此外, π 同时也是 \hat{P} 的平稳分布, 因为

$$[\pi^\top \hat{P}](j) = \sum_{i \in [N]} \pi(i) \hat{P}(i, j) = \sum_{i \in [N]} \pi(i) P(i \rightarrow j) = \pi(j).$$

对于可逆链, 一个平凡的例子是确定性的游走. 比如一个确定性的三环 $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ 的反向链就是 $1 \rightarrow 3 \rightarrow 2 \rightarrow 1$. 下面是一个更有意思的例子.

示例 2.12 环上的有偏随机游走.

考虑状态空间 $\Omega = \mathbb{Z}/m\mathbb{Z}$ 上的随机游走. 对每一个状态 $i \in \Omega$, 链以概率 p 从 i 走到 $i+1$, 以概率 $1-p$ 从 i 走到 $i-1$, 其中所有加减法都按模 m 计算.

这个链的平稳分布是均匀分布 $\pi(i) = \frac{1}{m}$. 这是因为对任意 $i \in \Omega$, 流入状态 i 的两种可能分别来自 $i-1$ 和 $i+1$, 因此

$$[\pi^\top P](i) = \frac{1}{m} P(i-1 \rightarrow i) + \frac{1}{m} P(i+1 \rightarrow i) = \frac{1}{m} (p + (1-p)) = \frac{1}{m} = \pi(i).$$

于是时间反演链满足

$$\hat{P}(i \rightarrow i+1) = P(i+1 \rightarrow i) = 1-p, \quad \hat{P}(i \rightarrow i-1) = P(i-1 \rightarrow i) = p.$$

也就是说, 如果原链更倾向于沿着环顺时针前进 ($p > \frac{1}{2}$), 那么反向链就更倾向于逆

给熟悉线性代数的读者: 若定义加权内积 $\langle f, g \rangle_\pi := \sum_{i \in [N]} \pi(i) f(i) g(i)$, 则 \hat{P} 满足 $\langle P f, g \rangle_\pi = \langle f, \hat{P} g \rangle_\pi$, 也就是说它是 P 在这个内积下的伴随算子. 这个观点在后文会多次出现.

这里 $\mathbb{Z}/m\mathbb{Z}$ 可以理解成 $\{0, 1, \dots, m-1\}$, 但 $m-1$ 的下一个点是 0, 也就是说, 这 m 个状态首尾相接地排成一个环.

时针前进; 正向的偏置在反向过程中被完全翻转. 只有在 $p = \frac{1}{2}$ 时, 原链与反向链才相同, 也就是 $\hat{P} = P$. 这正是下面要定义的可逆性.

下面这个命题说明了时间反演链的一个重要性质: 对于任意一条轨道, 在端点按照平稳分布开始的场合, 按照 P “正放” 和按照 \hat{P} “倒放” 的统计规律是完全相同的.

命题 2.13 轨道恒等式.

设 P 是状态空间 $[N]$ 上的马尔可夫链的转移矩阵, π 是它的一个平稳分布, 并且满足 $\forall i \in [N], \pi(i) > 0$. 记 \hat{P} 为 P 关于 π 的时间反演链. 则对于任意整数 $t \geq 1$ 和任意路径 $z_0, z_1, \dots, z_t \in [N]$, 都有

$$\pi(z_0) \prod_{k=1}^t P(z_{k-1}, z_k) = \pi(z_t) \prod_{k=1}^t \hat{P}(z_k, z_{k-1}).$$

证明. 由时间反演链的定义,

$$\hat{P}(z_k, z_{k-1}) = \frac{\pi(z_{k-1})P(z_{k-1}, z_k)}{\pi(z_k)}.$$

因此

$$\pi(z_t) \prod_{k=1}^t \hat{P}(z_k, z_{k-1}) = \pi(z_t) \prod_{k=1}^t \frac{\pi(z_{k-1})P(z_{k-1}, z_k)}{\pi(z_k)} = \pi(z_0) \prod_{k=1}^t P(z_{k-1}, z_k),$$

其中中间的 $\pi(z_1), \dots, \pi(z_t)$ 全部逐项约去. \square

可逆 (reversible) 马尔可夫链

我们刚刚构造了关于平稳分布 π 的反向链 \hat{P} . 最特殊的情形是, 这条反向链恰好和原链完全相同, 也就是

$$\forall i, j \in [N], \hat{P}(i, j) = P(i, j).$$

把时间反演链的定义代入上式, 便得到

$$\forall i, j \in [N], \pi(i)P(i, j) = \pi(j)P(j, i),$$

一般来说, 对于任意一个给定的马尔可夫链的转移矩阵 P , 如果存在某个分布 π 使得上式成立, 我们就称 P 关于 π 是(时间)可逆的. 这一组等式称为细致平衡条件 (*detailed balance conditions*). 也就是说, 可逆性正是“原链与它关于 π 的时间反演链完全相同”这一特殊情形. 我们还可以直接验证, 这时 π 必须是 P 的平稳分布:

$$[\pi^\top P](j) = \sum_{i \in [N]} \pi(i)P(i, j) = \sum_{i \in [N]} \pi(j)P(j, i) = \pi(j).$$

前面的轨道恒等式告诉我们：任意一条路径的正向权重，恰好等于它的倒放路径在反向链中的权重。而在可逆情形下， $\hat{P} = P$ ，所以整段轨道倒放后的统计规律也保持不变。也就是说，如果从平稳分布出发，任何遵循该链的变量序列 X_0, X_1, \dots, X_t 的分布，即 $(X_0, X_1, \dots, X_{t-1}, X_t)$ 的分布与 $(X_t, X_{t-1}, \dots, X_1, X_0)$ 的分布是相同的。对于所有 $x_0, x_1, \dots, x_t \in [N]$,

$$\begin{aligned} P[X_0 = x_0, X_1 = x_1, \dots, X_t = x_t] &= \pi(x_0)P(x_0, x_1) \cdots P(x_{t-1}, x_t) \\ &= \pi(x_t)P(x_t, x_{t-1}) \cdots P(x_1, x_0) \\ &= P[X_0 = x_t, X_1 = x_{t-1}, \dots, X_t = x_0]. \end{aligned}$$

这也是“可逆”这一名称的来源。显然，可逆的链是一类特殊的链，但它的表达能力也足够强大，能够满足我们大部分的要求。并且，它的特殊性在于其转移矩阵在某种意义上是对称的，因此许多线性代数中的强有力工具都可以用来研究它。我们也能够很快地验证一个分布是不是给定链的平稳分布——只需验证细致平衡条件即可。我们下面来看两个例子。

示例 2.14.

考虑在一个 N 个顶点的连通的 d -正则的无向图上进行随机游走。在这个图上，每一个顶点均正好有 d 个邻居，游走的每一步均是等概率的选择一个邻居走过去。那么均匀分布 π 是这个随机游走的一个平稳分布。我们直接来考察细致平衡条件。对于两个不同的点 i, j ，如果它们不相邻，那么细致平衡条件的左右两边都是零，自然成立。否则

$$\pi(i) \cdot P(i \rightarrow j) = \frac{1}{N} \cdot \frac{1}{d} = \pi(j) \cdot P(j \rightarrow i).$$

我们会经常说对于某一个 $i \in \Omega$ ， Ω 上的分布 μ 满足 $\mu(i) \propto w(i)$ ，即 $\mu(i)$ 正比于 $w(i)$ ，也就是说 $\mu(i) = \frac{w(i)}{\sum_{j \in \Omega} w(j)}$ 。这里 $\forall i, w(i) \geq 0$ 是一个权重函数。

示例 2.15.

在上面的例子里，如果这个图的 N 个顶点的度数分别是 d_1, \dots, d_N ，那么对于同样的随机游走（每次均匀选一个邻居走过去），它的平稳分布是对于每一个 i ， $\pi(i) \propto d_i$ 。为了验证这个，我们检查细致平衡条件。对于相邻的 i, j ，我们有

$$\pi(i) \cdot P(i \rightarrow j) = \frac{d_i}{Z} \cdot \frac{1}{d_i} = \frac{1}{Z} = \frac{d_j}{Z} \cdot \frac{1}{d_j} = \pi(j) \cdot P(j \rightarrow i),$$

这儿 $Z := \sum_i d_i$ 是归一化常数。

回忆有限状态马尔可夫链基本定理需要的两个条件：不可约性和无周期性。对于可逆链，其转移图是无向图（但每条边两个方向的权重不一样），因此不可约性等价于转移图的

实际上，对于可逆链，边 (i, j) 的权重的正确看法是定义为 $\pi(i) \cdot P(i \rightarrow j)$ 。根据细致平衡条件， (i, j) 的权重和 (j, i) 的权重是一样的。这样，就真的是无向图了。

连通性. 而无周期性则等价于图不是二分图.

Metropolis-Hastings 算法

给定状态空间 Ω 上的分布 π , 我们如何设计一个马尔可夫链 P , 使得 π 是 P 的平稳分布? Metropolis-Hastings 算法提供了一种方法, 只要转移图 G 是连通的且无向的.

Metropolis-Hastings 算法构造的马尔可夫链一定是可逆的. 给定任意一个连通的无向图 (V, E) 作为转移图, 设 Δ 为转移图中除自环外的最大度数 (即 $\Delta := \max_{i \in [N]} \sum_{j \neq i \in [N]} \mathbb{1}\{i, j\} \in E$). 设当前状态为 i , 且 i 有 d 个邻居. 一步转移按如下方式进行: 首先, 均匀随机地从 i 的 d 个邻居中选择一个 j , 作为提议状态. 然后, 以概率 $\min\left\{\frac{\pi(j)}{\pi(i)}, 1\right\}$ 接受该提议 (移动到 j), 否则拒绝 (留在 i). 此外, 为了保证所有顶点的度数一致, 我们在上述过程之前附加一个懒惰化步骤: 以概率 $\frac{d}{\Delta+1}$ 执行上面的提议-接受流程, 以概率 $\frac{\Delta+1-d}{\Delta+1}$ 直接留在 i . 这等价于给每个状态 i 补充 $\Delta+1-d$ 条自环, 使得包含自环在内每个点的度数均为 $\Delta+1$. 以下给出该算法的伪代码.

算法 2.16 Metropolis-Hastings 一步转移.

Require: 当前状态 i , 目标分布 π , 转移图 (V, E) , 最大度数 Δ

- 1: 均匀随机选取 $k \in \{1, 2, \dots, \Delta+1\}$
- 2: **if** $k \leq d$ **then** ▷ d 为 i 的邻居数
- 3: 令 $j \leftarrow i$ 的第 k 个邻居
- 4: 以概率 $\min\left\{\frac{\pi(j)}{\pi(i)}, 1\right\}$ 令 $i \leftarrow j$ ▷ 接受提议
- 5: **return** i

于是转移矩阵 P 为

$$\forall i, j \in [N]: \quad P(i, j) = \begin{cases} \frac{1}{\Delta+1} \min\left\{\frac{\pi(j)}{\pi(i)}, 1\right\}, & \text{若 } i \neq j \text{ 且 } i, j \text{ 相邻;} \\ 1 - \sum_{k \neq i} P(i, k), & \text{若 } i = j. \end{cases}$$

我们可以验证, P 对于 π 是可逆的:

$$\forall i, j \in \Omega: \quad \pi(i)P(i, j) = \pi(i) \cdot \frac{1}{\Delta+1} \min\left\{\frac{\pi(j)}{\pi(i)}, 1\right\} = \frac{\min\{\pi(i), \pi(j)\}}{\Delta+1} = \pi(j)P(j, i).$$

注意到, 由于每一个点上均有自环, 所以我们构造出来的 P 是不可约且无周期性的. 因此, 根据定理 2.10, 从任意初始状态出发, 通过不停的模拟 P , 分布均能够收敛到 π .

Metropolis-Hastings 算法的一个很重要的好处在于我们并不需要知道 π 就能实现该算法, 我们只需要知道 $\frac{\pi(j)}{\pi(i)}$ 这一量的值. 而这一个量在许多应用中更容易计算. 比如说, 我们知道每一个状态 i 都有一个权重 $w(i) \geq 0$, 并且想要的分布是 $\pi(i) \propto w(i)$, 那么 $\frac{\pi(j)}{\pi(i)} = \frac{w(j)}{w(i)}$, 我们并不需

示例 2.17.

我们通过一个简单的例子来展示该算法的工作原理. 考虑一个有 3 个顶点 $\{a, b, c\}$ 的图. 顶点之间有无向边 (a, b) 、 (b, c) 和 (a, c) , 每个顶点都有自环. 在这种情况下, $\Delta = 2$. 如果我们希望设计一个具有平稳分布 $(\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$ 的转移矩阵 P , 根据 Metropolis 算法, 我们有:

$$\begin{aligned} P(a, b) &= \frac{1}{2+1} \cdot \frac{2}{3} = \frac{2}{9}, \\ P(a, c) &= \frac{1}{2+1} \cdot \frac{1}{3} = \frac{1}{9}, \\ P(a, a) &= 1 - \frac{1}{9} - \frac{2}{9} = \frac{2}{3}. \end{aligned}$$

大家可以类似的得出从 b 出发以及从 c 出发的权重.

我们将在未来多次使用 Metropolis-Hastings 算法构造理想平稳分布的马尔可夫链.

2.4 耦合 (coupling)

我们在这一节引入称为耦合的概念, 是概率论中一个非常重要的工具, 也是我们在下一节证明马尔可夫链基本定理的核心技巧. 我们首先定义两个分布的全变差距离 (total variation distance).

定义 2.18 全变差距离.

在一个可数状态空间 Ω 上, 两个分布 μ 和 ν 的全变差距离定义为 $d_{TV}(\mu, \nu) := \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|$.

我们可以观察下图 (概率质量函数的图像) 中两个分布在同一样本空间上的关系. 全变差距离等于两条曲线所围面积的一半.

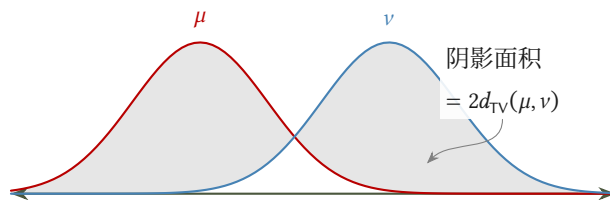


图 2.9: 两个分布的全变差距离

从图中我们可以看出如下命题成立 (见第一章练习).

命题 2.19.

定义 $\mu(A) = \sum_{x \in A} \mu(x)$, $\nu(A) = \sum_{x \in A} \nu(x)$, 则有

$$d_{TV}(\mu, \nu) = \max_{A \subseteq \Omega} |\mu(A) - \nu(A)|.$$

换句话说, 两个分布 μ 和 ν 的全变差距离, 可以看成那些所有 μ 比 ν 大的样本点上的概率差的和.

现在我们给出耦合的正式定义.

定义 2.20 耦合.

设 μ 和 ν 是分别定义在空间 Ω_1 和 Ω_2 上的两个分布. 令 ω 为定义在空间 $\Omega_1 \times \Omega_2$ 上的分布. 如果随机变量 $(X, Y) \sim \omega$ 满足 $X \sim \mu$ 且 $Y \sim \nu$, 则称 ω 为 μ 和 ν 的一个耦合.

我们有时候也会说一个新的 (定义在另外的概率空间上的) 二元随机变量 (\bar{X}, \bar{Y}) 随机变量是 X 和 Y 的耦合, 它的意思是, (\bar{X}, \bar{Y}) 的分布是 X 的分布和 Y 的分布的耦合. 在这门课里, 在不会引起误解的情况下, 我们语言上不太区分随机变量的耦合和分布的耦合.

耦合总是存在的, 比如我们让 $\omega = \mu \otimes \nu$, 即 $(X, Y) \sim \omega$ 对应于 X 和 Y 各自独立的来自 μ 和 ν , 那显然 ω 是 μ 和 ν 的耦合. 然而, 在许多应用中, 我们希望 X 和 Y 是相关的, 同时保持它们各自的边缘概率正确.

示例 2.21 硬币的耦合.

设 $\mu = \text{Ber}(\frac{1}{2})$, $\nu = \text{Ber}(\frac{2}{3})$, $X \sim \mu$, $Y \sim \nu$. 下面两个表格分别给出了 (X, Y) 的两种联合分布 (即两种耦合), 每个格子的值是 $P[X = x, Y = y]$.

	Y = 0	Y = 1
X = 0	$\frac{1}{6}$	$\frac{1}{3}$
X = 1	$\frac{1}{6}$	$\frac{1}{3}$

	Y = 0	Y = 1
X = 0	$\frac{1}{3}$	$\frac{1}{6}$
X = 1	0	$\frac{1}{2}$

左表是独立耦合 ($\omega(x, y) = \mu(x) \cdot \nu(y)$), 此时 $P[X \neq Y] = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}$. 右表则“尽量”的把概率放到了对角线上 (即 $X = Y$ 的格子), 使得 $P[X \neq Y] = \frac{1}{6}$.

事实上, 对于一般 (不一定离散) 的分布 μ 和 ν , 我们用这个式子定义它们的全变差距离, 即

$$d_{TV}(\mu, \nu) := \sup_A |\mu(A) - \nu(A)|.$$

其中 A 来自于所有的可测集.

严格的说, 我们有两个概率空间 $(\Omega_1, \mathcal{F}_1, \mu)$ 和 $(\Omega_2, \mathcal{F}_2, \nu)$, 那么我们说概率空间 $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \omega)$ 上的概率测度 ω 是 μ 和 ν 的耦合, 当且仅当如果随机变量 (X, Y) 的分布是 ω , 那么它们的边缘分布 X 和 Y 分别是 μ 和 ν .

请验证两个表格的行和与列和分别是 μ 和 ν .

在上面的例子里，第一张图对应的是两个独立的硬币 $X \sim \text{Ber}(\frac{1}{2})$, $Y \sim \text{Ber}(\frac{2}{3})$ 。而在第二张图里，可以验证，这是使得 $\mathbb{P}_{(X,Y) \sim \omega} [X = Y]$ 最大的耦合 ω 。

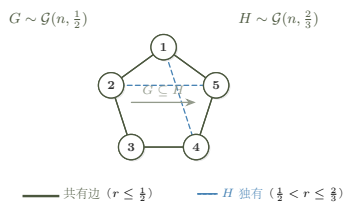


图 2.10: 通过耦合使得 $G \subseteq H$

示例 2.22 随机图的比较.

Erdős-Rényi 随机图模型 $\mathcal{G}(n, p)$ 是指这样一个包含 n 个顶点的随机图的分布，我们通过如下操作来得到一个随机图 $G \sim \mathcal{G}(n, p)$ ：对于每一对点 $\{i, j\}$ ，我们独立的扔一个 $\text{Ber}(p)$ 的硬币，如果是正面，就加入边 $\{i, j\}$ ，如果是反面，则不加入。我们现在想证明， $\mathbb{P}_{G \sim \mathcal{G}(n, 1/2)} [G \text{ 连通}] \leq \mathbb{P}_{H \sim \mathcal{G}(n, 2/3)} [H \text{ 连通}]$ 。这看似是一个很显然的结论，但似乎不太好严格的证明。我们现在用耦合的想法，来巧妙的说明。

我们首先看一下，如果要从 $\mathcal{G}(n, p)$ 中真的采样一个图，应该如何操作。一个很自然的做法就是依次检查每一个 $\{i, j\}$ ，对每一个点对，在 $[0, 1]$ 中均匀随机的选取一个实数 r ，然后看如果 $r \leq p$ ，则加上边 $\{i, j\}$ 否则不加。我们现在想象有两个人同时从 $\mathcal{G}(n, 1/2)$ 和 $\mathcal{G}(n, 2/3)$ 中采样图 G 和 H 。她们依次检查同一个 $\{i, j\}$ ，然后对每一个点对，在 $[0, 1]$ 中均匀随机的选取一个实数 r ，然后两个人使用同一个 r ，来决定自己的图里面是否有这一条边。显然，这样获得的 G 和 H 的分布是 $\mathcal{G}(n, 1/2)$ 和 $\mathcal{G}(n, 2/3)$ 的耦合。我们把这个耦合称作 ω 。此外，我们的构造保证了，如果 $\{i, j\}$ 是 G 的边，那么它一定也是 H 的边（因为 $r \leq 1/2 \implies r \leq 2/3$ ）。所以我们有下面的结论：

$$\begin{aligned} \mathbb{P}_{G \sim \mathcal{G}(n, 1/2)} [G \text{ 连通}] &= \mathbb{P}_{(G,H) \sim \omega} [G \text{ 连通}] \\ &\leq \mathbb{P}_{(G,H) \sim \omega} [H \text{ 连通}] \\ &= \mathbb{P}_{H \sim \mathcal{G}(n, 2/3)} [H \text{ 连通}]. \end{aligned}$$

实际上，我们这里在每条边上构造的耦合就是示例 2.21 里第二个表格里的耦合。

实际上，我们会特别关心这种使 $X = Y$ 概率最大的耦合，因为它与两个分布的全变差距离有密切联系。下面这个结论，称为耦合引理 (*coupling lemma*)，会在我们未来的学习中扮演关键的角色。同样，我们不妨假设我们考虑的是有限的概率空间。这个引理对于一般的概率空间也是正确的，但其证明留作练习。

引理 2.23 耦合引理.

设 μ 和 ν 是定义在同一个（离散）样本空间 Ω 上的两个分布。则对于 μ 和 ν 的任何耦合 ω ，有

$$\mathbb{P}_{(X,Y) \sim \omega} [X \neq Y] \geq d_{\text{TV}}(\mu, \nu).$$

这也可能是耦合之所以被称为耦合的原因。在文献中，我们往往把 $X = Y$ 这个事件称为“ X 和 Y 耦合住了 (coupled)”。

并且, 存在一个“最优”耦合 ω^* , 使得

$$\mathbb{P}_{(X,Y) \sim \omega^*} [X \neq Y] = d_{TV}(\mu, \nu).$$

证明. 对于有限的 Ω , 设计一个耦合等价于填充一个 $\Omega \times \Omega$ 的矩阵, 使得其边缘分布正确. 显然有

$$\mathbb{P}[X = Y] = \sum_{t \in \Omega} \mathbb{P}[X = Y = t] \leq \sum_{t \in \Omega} \mu(t) \wedge \nu(t).$$

回忆我们定义过记号 $a \wedge b := \min\{a, b\}$,
 $a \vee b := \max\{a, b\}$.

因此,

$$\mathbb{P}[X \neq Y] \geq 1 - \sum_{t \in \Omega} \mu(t) \wedge \nu(t) = \sum_{t \in \Omega} (\mu(t) - \mu(t) \wedge \nu(t)).$$

由此,

$$\mathbb{P}[X \neq Y] = \max_{A \subseteq \Omega} \{\mu(A) - \nu(A)\} = d_{TV}(\mu, \nu).$$

为了构造“最优”的耦合 ω^* 使得上述不等式取到等号, 对于每个 $t \in \Omega$, 我们可以“贪心”的令 $\mathbb{P}_{(X,Y) \sim \omega^*} [X = Y = t] = \min\{\mu(t), \nu(t)\}$. 当然, 我们还要说明剩下的格子能够合法的填下去. 至少存在这样一种填法: 把剩下的格子里的概率定义成对于 μ 和 ν 去掉对角线概率然后进行合理归一化后各自的概率乘积. \square

请务必想明白剩下的各自概率如何构造.

耦合引理为两个分布之间的距离提供了一种上界: 对于任意两个分布 μ 和 ν 及其任意耦合 ω , $\mathbb{P}_{(X,Y) \sim \omega} [X \neq Y]$ 的上界也是 $d_{TV}(\mu, \nu)$ 的上界. 这是一种非常有用的方法来约束全变差距离.

2.5 马尔可夫链基本定理的证明

我们现在回忆一下在 2.2 节中介绍的马尔可夫链的基本定理.

定理 2.24 马尔可夫链的基本定理.

若有限马尔可夫链 $P \in \mathbb{R}^{n \times n}$ 是不可约且无周期的, 则它有唯一的平稳分布 $\pi \in \mathbb{R}^n$. 并且, 对于任意分布 $\mu \in \mathbb{R}^n$, 有

$$\lim_{t \rightarrow \infty} \mu^\top P^t = \pi^\top.$$

我们在这一节给出一个基于耦合法的证明.

不可约且无周期链的代数性质

我们首先研究不可约且无周期链的转移矩阵 P 的性质. 我们将使用以下引理.

引理 2.25.

设 c_1, c_2, \dots, c_s 是一组正整数, 满足 $\gcd(c_1, \dots, c_s) = 1$. 则对于任何足够大的整数 b , 存在 $y_1, y_2, \dots, y_s \in \mathbb{N}$ 使得

$$c_1 y_1 + c_2 y_2 + \dots + c_s y_s = b.$$

证明. 根据 Bézout 恒等式, 存在 $x_1, x_2, \dots, x_s \in \mathbb{Z}$, 使得

$$c_1 x_1 + c_2 x_2 + \dots + c_s x_s = 1.$$

Bézout 恒等式是初等数论中的基本结论:
对于整数 a_1, \dots, a_s , 存在整数 x_1, \dots, x_s 使得
 $a_1 x_1 + \dots + a_s x_s = \gcd(a_1, \dots, a_s)$.

我们对 s 进行归纳. $s = 1$ 的情形显然成立. 假设 $s \geq 2$ 且引理对更小的 s 成立. 令 $g = \gcd(c_1, \dots, c_{s-1})$. 根据归纳假设, 我们知道

$$\frac{c_1}{g} \cdot x_1 + \frac{c_2}{g} \cdot x_2 + \dots + \frac{c_{s-1}}{g} \cdot x_{s-1} = b' \iff c_1 \cdot x_1 + c_2 \cdot x_2 + \dots + c_{s-1} x_{s-1} = g \cdot b'$$

对足够大的 b' 有非负解. 因此, 我们只需证明以下方程

$$g \cdot b' + c_s \cdot x_s = b \tag{1}$$

对于任何足够大的 b , 存在非负解 (b', x_s) , 且 $b' > 0$. 换句话说, 我们需要证明, 对于任意 $b_0 > 0$, 当 b 足够大时, 方程 (1) 有 $b' > b_0$ 的非负解.

注意到 $\gcd(g, c_s) = 1$, 我们可以找到整数 (y, x) 使得

$$g \cdot y + c_s \cdot x = 1 \iff g \cdot (by) + c_s \cdot (bx) = b.$$

注意, 对于任意 $k \in \mathbb{Z}_{\geq 0}$, 有 $g \cdot (by + kc_s) + c_s \cdot (bx - kg) = b$. 我们需要 $by + kc_s > b_0$ 且 $bx - kg \geq 0$, 这等价于

$$\frac{bx}{g} \geq k > \frac{b_0 - by}{c_s}.$$

当 $b \geq g(b_0 + c_s)$ 时, 总能找到这样的整数 k . □

接下来我们利用这一引理证明不可约且无周期链的以下性质.

命题 2.26.

设 $P \in \mathbb{R}^{n \times n}$ 是一个不可约且无周期的马尔可夫链. 则有

$$\exists t^*, \forall i, j \in [n]: P^{t^*}(i, j) > 0.$$

证明. 不可约的性质意味着 $\forall i, j, \exists t: P^t(i, j) > 0$. 假设从状态 i 出发并返回 i 的环路有 s 条, 其长度为 c_1, c_2, \dots, c_s . 根据无周期性, 我们有 $\gcd(c_1, c_2, \dots, c_s) = 1$.

对于任意足够大的 m 和任意状态对 (i, j) , 根据上述引理和不可约性, 存在从 i 到 j 的恰好 m 步的路径. 因此, 存在足够大 $t_{ij} > 0$, 使得从 i 出发可以先走 $t_{ij} - m$ 步回到 i , 再走 m 步到达 j , 从而 $P^{t_{ij}}(i, j) > 0$. 我们令 $t^* = \max_{i,j} t_{ij}$, 则对于任意状态对 (i, j) , 都有 $P^{t^*}(i, j) > 0$. \square

从证明中也可以看出来, 对于每一个 i, j , 对于任何 $t > t^*$, $P^t(i, j) > 0$ 也成立. 实际上, 上面命题里的论断反过来也成立:

命题 2.27.

如果一个马尔可夫链满足 $\exists t^*, \forall i, j \in [n]: P^{t^*}(i, j) > 0$, 则它一定是不可约且无周期的.

证明. 不可约是显然的. 由于我们知道对于任意 $i \in [n]$, $P^{t^*}(i), P^{t^*+1}(i) > 0$, 所以它有两个长度为 t^* 和 $t^* + 1$ 的自环. 但 $\gcd(t^*, t^* + 1) = 1$, 因此它是无周期的. \square

事实上, 对于不可约的马尔可夫链, 无周期性是一个类性质, 也就是说, 如果 i 和 j 是强连通的, 那么 i 无周期当且仅当 j 无周期. 所以, 对于不可约的马尔可夫链, 要么所有状态都是无周期的, 要么所有状态都是有周期的. 因此, 我们会称整个马尔可夫链是“无周期”或者“有周期”的. 我们把这个性质的证明留作一个练习.

Chihao: 这里加了一个练习.

定理 2.10 的证明

我们已知 P 有一个平稳分布 π . 我们需要证明, 对于任意初始分布 μ_0 , 有

$$\lim_{t \rightarrow \infty} d_{TV}(\mu_t, \pi) = 0$$

成立. 这里 $\mu_t^\top = \mu_0^\top P^t$. 我们使用耦合引理, 通过构造 μ_t 和 π 的一个耦合, 来给出它们的全变差距离的一个上界. 为了方便阐述, 我们引入马尔可夫链的耦合的概念.

设 $(X_t)_{t \geq 0}, \{Y_t\}_{t \geq 0}$ 是两个马尔可夫链, 它们的转移矩阵分别是 P 和 Q . 我们说一个 (二元取值) 的随机过程 $(\tilde{X}_t, \tilde{Y}_t)_{t \geq 0}$ 是 $(X_t)_{t \geq 0}$ 和 $(Y_t)_{t \geq 0}$ 的耦合, 如果对于任意 $t \geq 1, \forall i, j$:

$$\begin{aligned} \mathbb{P}_{(\tilde{X}_t, \tilde{Y}_t)}[\tilde{X}_t = j \mid \tilde{X}_{t-1} = i] &= P(i \rightarrow j) \\ \mathbb{P}_{(\tilde{X}_t, \tilde{Y}_t)}[\tilde{Y}_t = j \mid \tilde{Y}_{t-1} = i] &= Q(i \rightarrow j). \end{aligned}$$

换句话说, 如果我们单独盯着 \tilde{X}_t 或者 \tilde{Y}_t 看, 它的演化行为分别和 X_t 和 Y_t 是一样的. 如果我们用 μ_t 表示 X_t 的分布, 用 ν_t 表示 Y_t 的分布, 那么显然 $(\tilde{X}_t, \tilde{Y}_t)$ 的分布是 μ_t 和 ν_t 的一个耦合.

请验证这件事!

我们用上面命题中的 t^* 来定义 $Q := P^{t^*}$. 由于 $Q(i, j) \geq \delta > 0$ 对于所有状态对 (i, j) 成立, 接下来我们对 Q -链 (即每 t^* 步观察一次的 P -链) 构造耦合. 设 $\{X_k\}$ 和 $\{Y_k\}$ 是两个 Q -链, 其中 $Y_0 \sim \pi$ 而 X_0 由任意分布 μ_0 生成. 我们构造 $(\bar{X}_k, \bar{Y}_k)_{k \geq 0}$ 如下:

注意 π 也是 Q 的平稳分布: $\pi^\top Q = \pi^\top P^{t^*} = \pi^\top$.

- 初始情况, 分别独立的从 μ_0 和 π 中采样: $\bar{X}_0 \sim \mu_0, \bar{Y}_0 \sim \pi$;
- 对于 $k \geq 1$, 如果 $\bar{X}_{k-1} = \bar{Y}_{k-1}$, 则令 $\bar{X}_k = \bar{Y}_k$ (一起按 Q 演化); 否则分别独立的按照 Q 从 \bar{X}_{k-1} 和 \bar{Y}_{k-1} 演化.

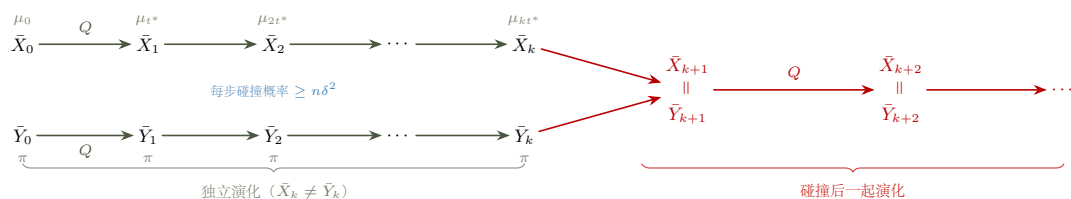


图 2.11: Q -链耦合示意图: 两条链独立演化直至碰撞, 此后一起移动

根据耦合引理, $d_{\text{TV}}(\mu_{kt^*}, \pi) \leq \mathbb{P}[\bar{X}_k \neq \bar{Y}_k]$. 现在我们估计碰撞概率. 当 $\bar{X}_{k-1} \neq \bar{Y}_{k-1}$ 时, 两链独立演化, 因此对于任意状态 j :

$$\mathbb{P}[\bar{X}_k = j \text{ 且 } \bar{Y}_k = j] = Q(\bar{X}_{k-1}, j) \cdot Q(\bar{Y}_{k-1}, j) \geq \delta^2.$$

对 j 求和, 得到

$$\mathbb{P}[\bar{X}_k = \bar{Y}_k \mid \bar{X}_{k-1} \neq \bar{Y}_{k-1}] = \sum_j Q(\bar{X}_{k-1}, j) \cdot Q(\bar{Y}_{k-1}, j) \geq n\delta^2.$$

因此 $\mathbb{P}[\bar{X}_k \neq \bar{Y}_k] \leq (1 - n\delta^2) \cdot \mathbb{P}[\bar{X}_{k-1} \neq \bar{Y}_{k-1}]$, 归纳得到

$$\mathbb{P}[\bar{X}_k \neq \bar{Y}_k] \leq (1 - n\delta^2)^k.$$

最后, 我们还有以下引理.

引理 2.28.

对于任意 $t \geq 0$, 有 $d_{\text{TV}}(\mu_{t+1}, \pi) \leq d_{\text{TV}}(\mu_t, \pi)$ 成立.

证明. 我们使用耦合引理. 设 (X_t, Y_t) 是 μ_t 和 π 的最优耦合, 即 $\mathbb{P}[X_t \neq Y_t] = d_{\text{TV}}(\mu_t, \pi)$. 我们构造 (X_{t+1}, Y_{t+1}) 如下:

- 如果 $X_t = Y_t$, 则令 X_{t+1} 从 X_t 按 P 演化, 并令 $Y_{t+1} = X_{t+1}$;
- 如果 $X_t \neq Y_t$, 则令 X_{t+1} 和 Y_{t+1} 分别独立地从 X_t 和 Y_t 按 P 演化.

我们验证 (X_{t+1}, Y_{t+1}) 是 μ_{t+1} 和 π 的一个耦合. 首先, X_{t+1} 的边缘分布为 μ_{t+1} , 因为 $X_t \sim \mu_t$ 且 X_{t+1} 按 P 从 X_t 演化. 其次, Y_{t+1} 的边缘分布为 π , 因为在两种情况下 Y_{t+1} 都按 P 从某个 π -分布的随机变量演化, 由 π 的平稳性知 $Y_{t+1} \sim \pi$.

现在, $X_{t+1} \neq Y_{t+1}$ 只可能在 $X_t \neq Y_t$ 时发生 (因为当 $X_t = Y_t$ 时我们令 $Y_{t+1} = X_{t+1}$), 因此

$$\mathbb{P}[X_{t+1} \neq Y_{t+1}] \leq \mathbb{P}[X_t \neq Y_t] = d_{TV}(\mu_t, \pi).$$

由耦合引理, $d_{TV}(\mu_{t+1}, \pi) \leq \mathbb{P}[X_{t+1} \neq Y_{t+1}] \leq d_{TV}(\mu_t, \pi)$. \square

结合上述耦合的指数衰减和 d_{TV} 的单调性, 对于任意 t , 取 $k = \lfloor t/t^* \rfloor$, 得

$$d_{TV}(\mu_t, \pi) \leq d_{TV}(\mu_{kt^*}, \pi) \leq (1 - n\delta^2)^k \rightarrow 0 \quad (t \rightarrow \infty).$$

这完成了定理 2.10 的证明.

第一种情况中, $Y_t = X_t$, 但 $Y_{t+1} = X_{t+1}$ 是从 X_t 按 P 演化的结果; 由于 $Y_t \sim \pi$ 且 $Y_t = X_t$, 这等价于从 π -分布出发按 P 演化.

2.6 马尔可夫链的收敛速率与混合时间 (mixing time)

我们现在开始研究马尔可夫链的收敛速率. 我们从混合时间的概念开始. 对于任意 $\varepsilon > 0$, 马尔可夫链 P 的 (误差不超过 ε) 的混合时间定义为最小的步数 t , 使得从任意初始分布运行马尔可夫链后, 其与平稳分布的全变差距离至多为 ε . 形式化地定义为:

$$\tau_{\text{mix}}(\varepsilon) := \min_t \sup_{\mu_0} d_{TV}(\mu_t, \pi) \leq \varepsilon.$$

在有的文献里, 人们会定义 $\tau_{\text{mix}} := \tau_{\text{mix}}(1/4)$, 并直接把 τ_{mix} 称为混合时间.

回忆刚才我们用耦合法证明马尔可夫链基本定理的时候, 我们用到了不等式

$$d_{TV}(\mu_t, \pi) \leq \mathbb{P}_{(\bar{X}_t, \bar{Y}_t)}[\bar{X}_t \neq \bar{Y}_t].$$

其中 (\bar{X}_t, \bar{Y}_t) 是一个耦合. 显然, 如果我们能够定量的控制上式右边的概率, 比如说证明对于任意初始的 (\bar{X}_0, \bar{Y}_0) , $\mathbb{P}_{(\bar{X}_t, \bar{Y}_t)}[\bar{X}_t \neq \bar{Y}_t] \leq \varepsilon$, 那么根据定义, 马上就能得到 $\tau_{\text{mix}}(\varepsilon) \leq t$. 换句话说

为什么这个不等式可以蕴含混合时间上界? 请自己严格证明.

一个好的耦合可以给出马尔可夫链的混合时间的好的上界

因此, 耦合法是我们将要介绍的证明马尔可夫链收敛速度上界的一个重要方法. 下面是几个例子.

超立方体上的随机游走

这是个可逆马尔可夫链, 因此使用细致平衡条件即可简单验证其平稳分布为均匀分布.

回忆我们在本章开头介绍过的超立方体上的随机游走（参见图 2.4），并且已经验证了其平稳分布是均匀分布。现在我们考虑它的一个 lazy 版本。状态空间为 $\Omega = \{\pm 1\}^n$ ，当且仅当 $\|x - y\|_1 = 1$ 时，状态 x 和 y 之间存在一条边。我们从一个点 $X_0 \in \Omega$ 出发，在每一步中，如果当前的位置是 $X_t \in \{\pm 1\}^n$ ：

这里以 $\frac{1}{2}$ 概率不动的主要目的是为了使得马尔可夫链无周期化。在每个点上都加上自环的马尔可夫链也被称为 lazy chain。

- 以概率 $\frac{1}{2}$ 保持不动；
- 否则，随机选择一个 $i \in [n]$ 并翻转 $X_t(i)$ （即令 $X_t(i) \leftarrow -X_t(i)$ ）。

现在我们用耦合分析该过程的混合时间。我们可以把它每一步的移动写成下面这个等价形式。

- 随机选择一个 $i \in [n]$ 和一个 $b \in \{\pm 1\}$ ；
- 将 $X_t(i)$ 改为 b 。

我们构造一组马尔可夫链的耦合 (X_t, Y_t) ，其中 X_0 和 Y_0 是任意状态。构造方法是

两个随机游走 X_t 和 Y_t 在每一步选择相同的 i 和 b 。

于是，一旦某个位置 $i \in [n]$ 被选择， $X_t(i)$ 和 $Y_t(i)$ 将永远相同。因此，这个问题等价于我们研究过的奖券收集问题。我们可以计算出对于 $t \geq n \log n + cn$ ，未选择第 i 维的概率为：

$$\left(1 - \frac{1}{n}\right)^{n \log n + cn} \leq \frac{e^{-c}}{n}.$$

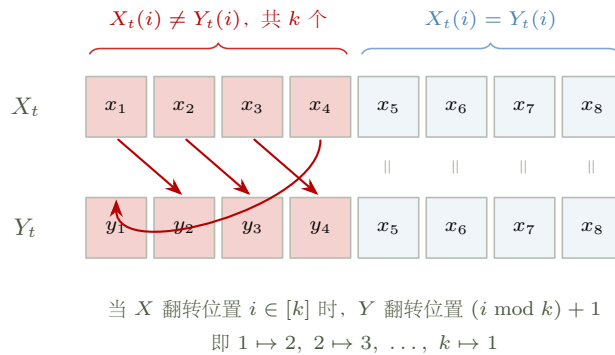
于是，至少存在一个维度未被选择的概率不超过 e^{-c} 。我们希望这个概率小于 ε ，因此选择 $c > \log \frac{1}{\varepsilon}$ 。于是，

$$\tau_{\text{mix}}(\varepsilon) \leq n \log \frac{n}{\varepsilon}.$$

接下来我们稍作修改，将随机游走里每一步不动的概率由 $\frac{1}{2}$ 改为 $\frac{1}{n+1}$ ，即以概率 $\frac{1}{n+1}$ 保持不动，使得这个 lazy walk 略为活跃一点。在这种情况下，我们描述另一种 X_t 和 Y_t 的耦合方法。我们总可以不失一般性的排列 X_t 和 Y_t 的坐标，使得所有取值不同的坐标排在前面。也就是说，存在一个索引 k ，使得当 $1 \leq i \leq k$ 时， $X_t(i) \neq Y_t(i)$ ，而当 $i > k$ 时， $X_t(i) = Y_t(i)$ 。我们的耦合规则如下：

- 如果 $k = 0$ ， Y_t 的行为与 X_t 相同。
- 如果 $k = 1$ ，当 X_t 翻转第一位时， Y_t 保持不动；当 X_t 保持不动时， Y_t 翻转第一位；其余时间两者行为一样。
- 如果 $k \geq 2$ ，我们区分 X 是否翻转了 $[k]$ 中的位置：
 - 如果 X 保持不动或翻转了一个 $i > k$ 的位置： Y 的行为相同；

- 如果 X 翻转了 $1 \leq i \leq k$: Y 翻转 $(i \bmod k) + 1$, 即将 $1 \mapsto 2, 2 \mapsto 3, \dots, k-1 \mapsto k, k \mapsto 1$.

图 2.12: 超立方体耦合的坐标循环置换示意图 ($k = 4$ 的情形)

显然, 上述方法确实是一个耦合. 事实上, 这种耦合的行为类似于“倍速奖券收集问题”, 因为在 $k \geq 2$ 的情况下, 如果 $i \in [k]$, 我们每次可以“收集”两个奖券. 因此, 可以相信,

$$\tau_{\text{mix}} \leq \frac{1}{2}n \log n + O(n).$$

随机至顶 (random-to-top) 洗牌问题

给定一副 n 张牌的牌堆, 考虑以下洗牌规则:

- 随机抽取一张牌;
- 将这张牌放到牌堆的顶端.

这一洗牌规则可以视为在 $n!$ 种排列上的随机游走, 且容易验证, 均匀分布是其平稳分布. 我们设计一个此马尔可夫链的耦合方法. 即, 设 X_t 和 Y_t 分别表示两副牌堆, 我们通过以下方式构造 X_{t+1} 和 Y_{t+1} :

抽取相同的牌.

注意, 这里是抽取“相同的牌”, 而不是位于相同位置的牌. 换句话说, 一旦我们在 X_t 中抽取了 $\heartsuit K$, 我们也在 Y_t 中抽取 $\heartsuit K$.

显然, 这是一个耦合, 而且一旦某张牌 (比如 $\heartsuit K$) 被抽取过, 那么 $\heartsuit K$ 在两副牌堆中的位置将始终相同. 因此, 如果我们问需要多少轮 T 使得 $X_T = Y_T$, 这个问题就再次等价于奖券收集问题. 我们有

$$\tau_{\text{mix}}(\epsilon) \leq n \log \frac{n}{\epsilon}.$$

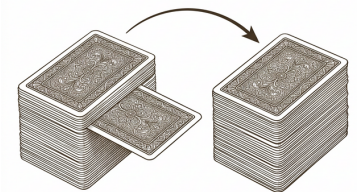


图 2.13: 随机至顶洗牌: 从牌堆中随机抽出一张牌, 放到顶端

第三章 可数无限状态马尔可夫链

我们在第二章深入讨论了有限状态的马尔可夫链，并建立了马尔可夫链基本定理。这个定理让我们对于马尔可夫链的结构及其平稳分布有了清晰的认识。

当马尔可夫链的状态变成无穷多个的时候，我们在有限状态时候的经验就不太能用了，比如下面这个在 \mathbb{N} 上的随机游走。

这一随机过程的状态空间为非负整数集 \mathbb{N} 。在每个正数状态 $i \geq 1$ ，一个人以概率 p 转移到 $i+1$ ，以概率 $1-p$ 向左转移到 $i-1$ ；而在边界状态 $i=0$ ，由于向左无路可走，我们规定它以概率 $1-p$ 停留在原地，以概率 p 依然向右。

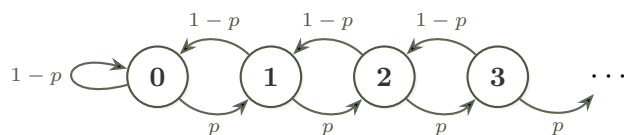


图 3.1: 概率逃逸模型：一维随机游走

假使在这个无穷多状态的马尔可夫链上，我们存在一个全局的平稳分布 π 。那么根据平稳分布的定义，我们一定有以下等式成立

$$\begin{aligned} \pi(0) &= \pi(0)(1-p) + \pi(1)(1-p) && \implies \pi(1) = \frac{p}{1-p}\pi(0), \\ \pi(1) &= \pi(0)p + \pi(2)(1-p) && \implies \pi(2) = \frac{p}{1-p}\pi(1), \\ &\dots && \\ \pi(i) &= \pi(i-1)p + \pi(i+1)(1-p) && \implies \pi(i+1) = \frac{p}{1-p}\pi(i). \\ &\dots && \end{aligned}$$

这个递推式告诉我们，对于任何 $i \geq 0$ ，都有显式解： $\pi(i) = \left(\frac{p}{1-p}\right)^i \pi(0)$ 。这是一个严格的等比数列。但是注意 π 必须是一个合法的概率分布，因此它必须满足所有状态的概率加

起来为 1:

$$\sum_{i=0}^{\infty} \pi(i) = \pi(0) \sum_{i=0}^{\infty} \left(\frac{p}{1-p}\right)^i = 1.$$

通过考察无穷级数 $\sum \left(\frac{p}{1-p}\right)^i$ 的敛散性, 我们会发现一些有趣的新现象:

- 如果 $p < \frac{1}{2}$ (向原点聚拢的力量更强), 无穷级数收敛于一个有限值. 这个时候可以得到一个合法的平稳分布 $\pi(i) = \left(\frac{p}{1-p}\right)^i \frac{1-2p}{1-p}$.
- 如果 $p > \frac{1}{2}$ (向无穷远散逸的力量更强), 无穷级数发散. 要使得该方程和能够收敛等于 1, 唯一的可能就是 $\pi(0) = 0$. 但这会导致对于所有的状态都有 $\pi(i) = 0$. 显然这说明该系统合法的平稳分布不存在.
- 在 $p = \frac{1}{2}$ (聚拢和散逸的力量取得平衡), 我们有 $\forall i \in \mathbb{N}, \pi(i) = \pi(0)$. 这同样说明平稳分布不存在.

我们之前知道, 对于有限状态的马尔可夫链, 平稳分布是一定存在的. 但从这个例子可以看出, 当状态空间变成无限时, 事情会变得复杂. 首先, 我们可能有一种使得平稳分布不存在的模式, 即“概率往无穷远处逃走了” ($p > \frac{1}{2}$); 另外, 直观上 $p = \frac{1}{2}$ 时候平稳分布的不存在来自于另外一种模式. 我们会在第 3.4 节中仔细的讨论它们.

我们这一章将要推广在上一章所发展的马尔可夫链理论, 实现状态空间从有限到无限的跨越.

3.1 信息流 (filtration)、停时 (stopping time) 与强马尔可夫性 (strong Markov property)

在接下来的讨论中, 我们事实上需要对于马尔可夫链进行更加精细的观测和操作, 这使得我们必须引入更加方便的数学语言, 即信息流 (filtration) 与停时 (stopping time) 的概念. 有了这些概念, 我们便可以证明我们关心的马尔可夫链满足一个更强的性质, 称为强马尔可夫性 (strong Markov property). 这个性质在我们未来分析马尔可夫链的时候起到至关重要的作用.

信息流: 刻画逐渐获得信息的过程

在概率论课中, 我们知道一个 σ -代数可以被看作“信息”的集合. 其中一个常见的特殊情况是一个随机变量 X 所生成的 σ -代数 $\sigma(X)$, 可以理解成关于 X 的“信息”. 那么, 对于一个随机过程 X_0, X_1, \dots , 对于 $t \geq 0$, 我们可以用 $\mathcal{F}_t := \sigma(X_0, X_1, \dots, X_t)$ 表示到时刻 t 为止

在时间离散场合, 当我们说到信息流 $\{\mathcal{F}_t\}_{t \geq 0}$ 的时候, 我们指的就是一族逐渐递增的 σ -代数 $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$. 但是在本书绝大多数场合, 我们关心的信息流是随机过程 $\{X_t\}_{t \geq 0}$ 的自然信息流, 即 $\mathcal{F}_t := \sigma(X_0, X_1, \dots, X_t)$.

该过程所揭示的所有信息. 我们把所有的这些 σ -代数放在一起, 就构成了随机过程 $\{X_t\}_{t \geq 0}$ 的一个信息流 $\{\mathcal{F}_t\}_{t \geq 0}$, 它被称之为该马尔可夫链的自然信息流. 一个显然的性质是

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$$

有了这个语言, 我们可以把马尔可夫性 (定义 2.1) 写成对于任何只与 X_{t+1}, X_{t+2}, \dots 有关的事件 $A \in \sigma(X_{t+1}, X_{t+2}, \dots)$,

$$P[A | \mathcal{F}_t] = P[A | X_t]. \quad (3.1)$$

关于条件是 σ -代数或者随机变量的条件概率的定义, 见附录「条件概率与条件期望」一节.

停时: 基于已有知识的决断

我们想象一个赌徒在赌场中决定何时离场. 他可以规定“赢满 100 块就走”, 也可以规定“连输三把就走”. 这些都是合法的策略, 因为在做出决定的每个瞬间, 他只需查阅当时已有的历史信息 \mathcal{F}_t , 便可知道是否应该离场. 然而, “在赢得最多的那一局之后走”就不是合法的策略——他需要看到未来所有结果之后才能确认“赢最多”在哪里.

我们把这些合法的策略叫做停时.

请大家用上面赌徒离场策略的例子检验一下这个定义.

定义 3.1 停时 (stopping time) .

设 τ 为一个取值为非负整数或 ∞ 的随机变量. 如果对于每一个时刻 $t \geq 0$, 事件 $[\tau \leq t]$ 都是 \mathcal{F}_t -可测的, 我们就称其为关于信息流 $\{\mathcal{F}_t\}_{t \geq 0}$ 的一个停时.

顾名思义, 停时 τ 定义了一个随机的停下来的时刻. 我们要求它对于每一个具体的时刻 t , 在观察了过程直到时刻 t 的历史后, 就能确切地知道“是否已经该停下来了”. 考虑一个随机游走, \mathcal{F}_t 为其自然信息流. “第一次访问状态 j 的时刻”是一个停时, 因为到了给定的时刻 t , 我们只要检查前 t 步是否出现过 j 就能判断; 而“最后一次访问状态 j 的时刻”通常不是停时, 因为我们需要未来的信息才能作出判定.

在马尔可夫链的研究中, 我们经常会遇到两个非常重要的停时:

定义 3.2 首次时间 (hitting time) 与返回时间 (return time) .

给定状态集合 $A \subseteq \Omega$, 定义首次时间 (hitting time) 为

$$T_A = \inf\{t \geq 0 : X_t \in A\}.$$

这两个定义的区别仅在于起始时刻: T_A 允许 $t = 0$ (如果初始状态已经在 A 里就立刻停下), 而 T_A^+ 要求至少走一步之后再停.

同时定义

$$T_A^+ = \inf\{t \geq 1 : X_t \in A\}.$$

当 $A = \{i\}$ 是单个状态时, 简记为 T_i 和 T_i^+ ; 其中 T_i^+ 正是从状态 i 出发的返回时间 (return time) .

容易验证, T_A 和 T_i^+ 都是合法的停时——要判断 $T_A \leq t$ 是否成立, 只需检查轨迹 X_0, X_1, \dots, X_t 中是否有任何一步落入了 A .

强马尔可夫性

马尔可夫性告诉我们, 在给定的固定时刻 t , 给定当前状态 X_t , 未来的演化与过去的历史 \mathcal{F}_t 独立. 一个很自然的问题是: 如果 t 不是常数, 而是一个停时 τ , 类似的独立性条件是否仍然成立?

对于离散时间的马尔可夫链, 答案是肯定的. 这就是强马尔可夫性.

定义 3.3 齐时马尔可夫链的强马尔可夫性 (strong Markov property for time homogeneous Markov chains) .

设 $(X_t)_{t \geq 0}$ 为一个齐时马尔可夫链, τ 是关于其自然信息流的一个停时. 我们称该过程具有强马尔可夫性, 若在条件 $\tau < \infty$ 且 $X_\tau = x_0$ 下, 停时之后的过程 $(X_\tau, X_{\tau+1}, X_{\tau+2}, \dots)$ 与 τ 之前的历史独立, 且其条件分布等同于一个以 x_0 为起点的全新马尔可夫链. 即对于任意相对时刻 $t_1, \dots, t_k > 0$ 和状态 x_1, \dots, x_k ,

$$\begin{aligned} \mathbb{P}[X_{\tau+t_1} = x_1, \dots, X_{\tau+t_k} = x_k \mid (X_0, X_1, \dots, X_{\tau-1}, X_\tau) = (z_0, z_1, \dots, z_{\tau-1}, x_0), \tau < \infty] \\ = \mathbb{P}[X_{t_1} = x_1, \dots, X_{t_k} = x_k \mid X_0 = x_0]. \end{aligned}$$

换句话说, 在停时 τ 的那一瞬间, 系统执行了一次无损重启——过去的历史被完全抛弃, 它以当前的状态 X_τ 为起点, 开始了一段全新的、在分布上与旧系统完全相同的演化.

为了简化记号, 对于状态空间中的任何一个状态 $i \in \Omega$, 我们在之后会使用 $\mathbb{P}_i[\cdot]$ 和 $\mathbb{E}_i[\cdot]$ 来分别表示以 i 为初始状态的概率测度以及期望算子, 也就是 $\mathbb{P}[\cdot \mid X_0 = i]$ 和 $\mathbb{E}[\cdot \mid X_0 = i]$.

命题 3.4 离散马尔可夫链满足强马尔可夫性.

设 $(X_t)_{t \geq 0}$ 为可数状态空间上的马尔可夫链, τ 为关于自然信息流 $(\mathcal{F}_t)_{t \geq 0}$ 的停时. 则对于任意 $k \geq 1$ 、状态 $x_0, x_1, \dots, x_k \in \Omega$,

$$\mathbb{P}[X_{\tau+1} = x_1, \dots, X_{\tau+k} = x_k \mid X_\tau = x_0, \tau < \infty] = \mathbb{P}_{x_0}[X_1 = x_1, \dots, X_k = x_k].$$

一般马尔可夫链的强马尔可夫性也可以使用形如式 (3.1) 的条件来表述. 这需要用到停时对应的 σ -代数 $\mathcal{F}_\tau := \{A \in \mathcal{F} : \forall t \geq 0, A \cap [\tau \leq t] \in \mathcal{F}_t\}$, 它精确地刻画了“在停时那一刻所拥有的信息”. 有了这个记号, 强马尔可夫性定义为对于任何事件 $A \in \sigma(X_{\tau+1}, X_{\tau+2}, \dots)$,

$$\mathbb{P}[A \mid \mathcal{F}_\tau] = \mathbb{P}[A \mid X_\tau, \tau] \quad \text{a.s. on } [\tau < \infty]$$

注意到这个定义里面我们没有要求马尔可夫链是齐时的.

当考虑连续时间的马尔可夫过程 $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ 时, 有可能 (X_t) 满足马尔可夫性但不满足强马尔可夫性.

证明. 根据定义,

$$\mathbb{P}[X_{\tau+1} = x_1, \dots, X_{\tau+k} = x_k \mid X_\tau = x_0, \tau < \infty] = \frac{\mathbb{P}[X_{\tau+1} = x_1, \dots, X_{\tau+k} = x_k, X_\tau = x_0, \tau < \infty]}{\mathbb{P}[X_\tau = x_0, \tau < \infty]}.$$

由于 τ 只取非负整数值, 按全概率公式将上式分子按 τ 的取值展开, 可以得到

$$\frac{\mathbb{P}[X_{\tau+1} = x_1, \dots, X_{\tau+k} = x_k, X_\tau = x_0, \tau < \infty]}{\mathbb{P}[X_\tau = x_0, \tau < \infty]} = \frac{\sum_{n=0}^{\infty} \mathbb{P}[X_{n+1} = x_1, \dots, X_{n+k} = x_k, X_n = x_0, \tau = n]}{\mathbb{P}[X_\tau = x_0, \tau < \infty]}.$$

令 $q := \mathbb{P}_{x_0}[X_1 = x_1, \dots, X_k = x_k]$, 我们只需证明, 对于每一个时间切片 $[\tau = n]$ 上, 都有

$$\mathbb{P}[X_{n+1} = x_1, \dots, X_{n+k} = x_k, X_n = x_0, \tau = n] = \mathbb{P}[X_\tau = x_0, \tau = n] \cdot q,$$

然后对所有的 n 求和就可以了.

对于每一个固定的 n , 我们有

$$\begin{aligned} \mathbb{P}[X_{\tau+1} = x_1, \dots, X_{\tau+k} = x_k, X_\tau = x_0, \tau = n] &= \mathbb{P}[X_{n+1} = x_1, \dots, X_{n+k} = x_k, X_n = x_0, \tau = n] \\ &\stackrel{\triangleright C_n := [\tau = n] \cap [X_n = x_0]}{=} \mathbb{P}[X_{n+1} = x_1, \dots, X_{n+k} = x_k, C_n] \end{aligned}$$

我们可以使用全期望公式得到

$$\begin{aligned} \mathbb{P}[X_{n+1} = x_1, \dots, X_{n+k} = x_k, C_n] &= \mathbf{E}[\mathbf{1}[X_{n+1} = x_1, \dots, X_{n+k} = x_k] \cdot \mathbf{1}[C_n]] \\ &\stackrel{\triangleright \text{全期望公式}}{=} \mathbf{E}[\mathbf{E}[\mathbf{1}[X_{n+1} = x_1, \dots, X_{n+k} = x_k] \cdot \mathbf{1}[C_n] \mid \mathcal{F}_n]] \\ &\stackrel{\triangleright C_n \in \mathcal{F}_n}{=} \mathbf{E}[\mathbf{1}[C_n] \cdot \mathbf{E}[\mathbf{1}[X_{n+1} = x_1, \dots, X_{n+k} = x_k] \mid \mathcal{F}_n]] \\ &\stackrel{\triangleright \text{马尔可夫性 (式 (3.1))}}{=} \mathbf{E}[\mathbf{1}[C_n] \cdot \mathbb{P}[X_{n+1} = x_1, \dots, X_{n+k} = x_k \mid X_n]]. \end{aligned}$$

我们知道, 如果 C_n 发生了, 那么 X_n 的值一定是 x_0 , 因此上面条件概率的取值为

$$\mathbb{P}[C_n] \cdot \mathbb{P}[X_1 = x_1, \dots, X_k = x_k \mid X_0 = x_0] = \mathbb{P}[C_n] \cdot q.$$

这便完成了证明. □

有了停时与强马尔可夫性这两个工具, 我们便可以开始将上一章有限状态马尔可夫链的基本概念逐一推广到无穷状态空间了.

3.2 从有限到无限

在上一章中, 我们使用线性代数和图论语言描述了有限状态马尔可夫链. 当状态空间 Ω 变成可数无穷时, 我们无法再将其转移概率和分布简单地表示为有限维的矩阵运算和有限图上的操作. 我们接下来将讨论如何将这些概念推广到无穷状态的情形.

转移核与查普曼-科尔莫戈罗夫方程

对于有限状态的马尔可夫链,我们使用转移矩阵 $P \in [0, 1]^{\Omega \times \Omega}$ 来编码一步转移概率. 当状态个数 $|\Omega|$ 变成无限时, 我们把 P 理解成一个 $\Omega \times \Omega \rightarrow [0, 1]$ 的函数, 即

$$\forall i, j \in \Omega, P(i, j) := \mathbb{P}[X_{t+1} = j \mid X_t = i].$$

我们把 P 称为转移核 (*transition kernel*). 我们之前在 Ω 有限场合熟悉的性质对于转移核 P 依然成立: 对于每一个状态 $i \in \Omega$, 转移核 P 在 i 这一行的所有元素加起来必须等于 1, 即 $\sum_{j \in \Omega} P(i, j) = 1$.

转移核是一个函数 $P: \Omega \times \Omega \rightarrow [0, 1]$, 它满足对于任意状态 $i \in \Omega$, 都有 $\sum_{j \in \Omega} P(i, j) = 1$. 为了与上一章的记号保持连贯, 我们在转移核中依旧使用 $P^t(i, j)$ 来表示从 i 出发, 经过 t 步之后到达 j 的概率, 即 $P^t(i, j) := \mathbb{P}[X_t = j \mid X_0 = i]$.

查普曼-科尔莫戈罗夫方程 (*Chapman-Kolmogorov equation*) 在无穷状态下依然成立: 对于任意 $s, t \geq 0$ 和状态 $i, j \in \Omega$,

$$P^{s+t}(i, j) = \sum_{k \in \Omega} P^s(i, k) P^t(k, j).$$

注意到这里查普曼-科尔莫戈罗夫方程不再像上一章那样就等价于矩阵乘法的结合律了. 但我们使用归纳法和全概率公式依然可以很容易地证明它. 我们把这个证明也留做练习.

分布的演化

在上一章中, 我们将状态空间 $[N]$ 上的分布等价地看成一个列向量 $\mu \in [0, 1]^N$, 并且用矩阵乘法 $\mu_{t+1}^\top = \mu_t^\top P$ 来描述分布随时间的演化. 当状态空间 Ω 变成可数无穷时, 分布 μ 变成了一个定义在 Ω 上的函数 $\mu: \Omega \rightarrow [0, 1]$, 满足 $\sum_{i \in \Omega} \mu(i) = 1$.

尽管我们无法再将分布的演化写成有限维的矩阵运算, 但由全概率公式推导出的演化方程在无穷状态下依然成立: 对于任意 $j \in \Omega$,

$$\mu_{t+1}(j) = \sum_{i \in \Omega} \mu_t(i) P(i, j).$$

更一般地, 从分布 μ_s 出发经过 t 步后的分布为

$$\mu_{s+t}(j) = \sum_{i \in \Omega} \mu_s(i) P^t(i, j).$$

因此, 每个时间点的分布依然完全由转移核 P 和初始分布 μ_0 决定.

平稳分布

平稳分布的定义可以直接推广到无穷状态空间.

定义 3.5 平稳分布 (无穷状态) .

设 P 为可数状态空间 Ω 上的转移核. 如果一个分布 π 满足对于任意 $j \in \Omega$,

$$\sum_{i \in \Omega} \pi(i) P(i, j) = \pi(j),$$

则称 π 为 P 的平稳分布.

在上一章中,我们利用线性代数证明了每个有限马尔可夫链都至少有一个平稳分布. 那个证明的核心是: 由 $P\mathbf{1} = \mathbf{1}$ 出发, 推出 1 是 P 的特征值, 从而也是 P^T 的特征值, 继而存在一个对应的非负特征向量. 然而, 这个论证完全依赖于有限维矩阵的性质, 在无穷维的场合不再适用.

有限和无限在这里分岔!

不过, 我们在有限状态场合所定义的可逆链概念和细致平衡条件, 在可数无穷的场合可以自然推广. 我们可以同样的验证, 如果有一个分布 π 满足对于任意 $i, j \in \Omega$,

$$\pi(i) P(i, j) = \pi(j) P(j, i),$$

那么 π 就是 P 的平稳分布.

事实上, 时间反演链的概念也完全一样地推广到可数无穷状态空间. 若 π 是转移核 P 的平稳分布且满足 $\forall i \in \Omega, \pi(i) > 0$, 则定义

$$\hat{P}(i, j) := \frac{\pi(j) P(j, i)}{\pi(i)}, \quad \forall i, j \in \Omega$$

仍然给出一个合法的转移核, 而 π 也是 \hat{P} 的平稳分布. 第二章中关于时间反演链的一切性质——包括轨道恒等式 (命题 2.13) ——在此完全成立.

可达、互通、不可约与无周期

在上一章中, 我们使用有限有向图的语言来刻画马尔可夫链的结构: 不可约等价于转移图强连通, 无周期通过有向图中包含某个顶点的环的长度来定义. 当状态空间变成可数无穷时, 转移图变成了一张无穷图——顶点集和边集都可能是无限的. 尽管“强连通”“环”这些图论概念在无穷图上依然有意义, 但在随机过程的文献中, 人们更习惯于用以下纯概率的语言来描述状态之间的关系.

定义 3.6 可达与互通.

设 $i, j \in \Omega$ 是两个状态.

- 如果存在 $t \geq 0$ 使得 $P^t(i, j) > 0$, 则称从 i 可以到达 (lead to) j , 记作 $i \rightarrow j$.
- 如果 $i \rightarrow j$ 且 $j \rightarrow i$ 同时成立, 则称 i 和 j 互通 (communicate), 记作 $i \leftrightarrow j$.

容易验证, 互通关系 \leftrightarrow 是一个等价关系 (自反性、对称性、传递性均成立), 因此它将状态空间 Ω 分割成若干个互通类 (*communicating class*). 每一个互通类内的状态可以互相到达, 而不同互通类之间则至少有一个方向无法到达.

在有限状态的场合, 不可约等价于转移图是强连通的. 在无穷状态的场合, “强连通”一般直接用互通的语言来表述.

定义 3.7 不可约 (irreducible) .

如果整个状态空间 Ω 只构成一个互通类, 即对于任意 $i, j \in \Omega$ 均有 $i \leftrightarrow j$, 则称该马尔可夫链是不可约的.

我们在定义 2.8 中定义的周期性概念在现在的场合依旧适用. 为了方便讨论, 对于一个状态 $i \in \Omega$, 我们定义

$$d(i) := \gcd\{t \geq 1 : P^t(i, i) > 0\}$$

为 i 的所有自环的长度的最大公约数. 我们知道 i 是无周期的等价于 $d(i) = 1$.

在上一章有限状态的场合, 我们已经证明了不可约链中所有状态共享同一个周期. 这个结论在无穷状态下依然成立.

命题 3.8 互通状态的周期相同.

若 $i \leftrightarrow j$, 则 $d(i) = d(j)$.

因此, 对于不可约链, 我们可以直接谈论整条链的周期 d . 如果 $d = 1$, 就称该链是无周期的.

术语对照: 有限状态 vs. 可数无穷状态

我们在本节中将上一章有限状态马尔可夫链的核心概念逐一推广到了可数无穷状态空间. 下表对这些概念在两种场合下的表述进行了汇总比较.

概念	有限状态 ($ \Omega < \infty$)	可数无穷状态 ($ \Omega = \aleph_0$)
转移概率	转移矩阵 $P \in [0, 1]^{N \times N}$	转移核 $P: \Omega \times \Omega \rightarrow [0, 1]$
分布	列向量 $\mu \in [0, 1]^N$	函数 $\mu: \Omega \rightarrow [0, 1]$, $\sum_i \mu(i) = 1$
分布演化	$\mu_{t+1}^\top = \mu_t^\top P$ (矩阵乘法)	$\mu_{t+1}(j) = \sum_{i \in \Omega} \mu_t(i) P(i, j)$ (无穷级数)
C-K 方程	$P^{s+t} = P^s P^t$ (矩阵乘法)	$P^{s+t}(i, j) = \sum_{k \in \Omega} P^s(i, k) P^t(k, j)$
平稳分布	$\pi^\top P = \pi^\top$; 总是存在	$\sum_i \pi(i) P(i, j) = \pi(j)$; 可能不存在
可逆性	$\pi(i)P(i, j) = \pi(j)P(j, i)$ (细致平衡)	定义相同, 验证方式相同
不可约	转移图强连通	$\forall i, j: i \leftrightarrow j$ (互通)
无周期	包含状态 i 的环长之最大公约数为 1	$d(i) = \gcd\{t \geq 1 : P^t(i, i) > 0\} = 1$

有了上述概念的推广, 一个核心问题自然浮现: 从某个状态出发, 能否到达给定的目标? 如果能到达, 平均需要多长时间? 这便是下一节的主题.

3.3 首达时间的相关计算

对于马尔可夫链的演化, 我们通常最关心的问题之一是: 从状态 i 出发, 到达某个目标状态集合 A 的概率有多大? 如果能到达, 平均需要多长时间?

对于目标集合 $A \subset \Omega$, 回顾我们将首达时间记作停时 $T_A = \inf\{t \geq 0 : X_t \in A\}$. 为了方便记号, 我们定义两个量:

- **首达概率 (Hitting Probability)**: 记为 $h_A(i) := \mathbb{P}_i[T_A < \infty]$, 表示从状态 i 出发, 在有限步内到达 A 的概率.
- **期望首达时间 (Expected Hitting Time)**: 记为 $e_A(i) := \mathbb{E}_i[T_A]$, 表示从状态 i 出发, 到达 A 所需的期望步数.

计算这两个量最直接的方法被称为第一步分析 (*first-step analysis*). 核心思想是利用马尔可夫性, 对第一步的状态进行全概率展开. 对于首达概率, 如果起点 $i \in A$, 毫无疑问 $h_A(i) = 1$. 如果 $i \notin A$, 走完第一步之后到了某个状态 j , 此时就像是重新出发一样. 于

当 $h_A(i) < 1$ 时, 由于有正的概率永远无法到达 A (即 $T_A = \infty$), 显然此时 $e_A(i) = \infty$.

是

写成概率会更清楚:

$$\begin{aligned} \mathbb{P}_i[T_A < \infty] &= \sum_{j \in \Omega} \mathbb{P}_i[X_1 = j] \mathbb{P}_i[T_A < \infty | X_1 = j] \\ &= \sum_{j \in \Omega} P(i, j) h_A(j). \end{aligned}$$

上面第二个等号值得细说. 我们处于 $X_0 = i \notin A$ 这一状态十分关键. 在 \mathbb{P}_i 对应的概率空间中, 事件 $[T_A < \infty]$ 由 (i, X_1, X_2, \dots) 的值完全决定. 因此, 在 $\mathbb{P}_i[\cdot | X_1 = j]$ 中, 它由 (i, j, X_2, X_3, \dots) 完全决定. 由于我们知道 $i \notin A$, 所以 $[T_A < \infty]$ 由 (j, X_2, X_3, \dots) 完全决定. 根据马尔可夫性, 在 $\mathbb{P}_i[\cdot | X_1 = j]$ 的概率空间 (i, j, X_2, X_3, \dots) 中的 (j, X_2, X_3, \dots) 的分布与 \mathbb{P}_j 下 (j, X_1, X_2, \dots) 的分布完全相同.

对于期望的等式, 我们唯一需要改变的是把全概率公式换成全期望公式.

$$h_A(i) = \sum_{j \in \Omega} P(i, j) h_A(j), \quad \forall i \notin A.$$

同理, 对于期望首达时间, 如果起点 $i \in A$, 时间为 $e_A(i) = 0$. 如果 $i \notin A$, 走了一步消耗掉 1 个时间单位后到达 j , 剩下的期望时间就是 $e_A(j)$. 因此:

$$e_A(i) = 1 + \sum_{j \in \Omega} P(i, j) e_A(j), \quad \forall i \notin A.$$

这两个方程形式上极其简单、直观, 并且在状态图有限时, 往往可以通过合适的边界条件使得线性方程组有唯一的解. 然而, 在状态无限的时候, 这个方程可能会有多解. 因此, 寻找正确的解成为一个核心问题.

方程的伪解与最小非负解定理

我们回忆本章开头的那个有偏随机游走: 在 \mathbb{N} 上以概率 p 向右走、 $1-p$ 向左走 (在 0 处向左的概率转化为原地停留). 我们想计算从状态 $i > 0$ 出发, 到达原点 $A = \{0\}$ 的概率, 即 $h_0(i) = \mathbb{P}_i[T_0 < \infty]$.

利用第一步分析, 方程为:

$$\begin{aligned} h_0(0) &= 1, \\ h_0(i) &= ph_0(i+1) + (1-p)h_0(i-1), \quad \forall i \geq 1. \end{aligned}$$

有趣的事情发生在 $p > 1/2$ 的时候. 直觉告诉我们, 无论从哪个非原点出发, 这个随机游走会有概率会跑掉而永远不回原点, 也就是真实的 $h_0(i)$ 应该严格小于 1 并随着 i 的增大而减小. 但是我们解一下这个方程会发现它有无穷组解: 对于任意 $C_2 \in \mathbb{R}$, 函数 $h_0(i) = 1 - C_2 + C_2 \left(\frac{1-p}{p}\right)^i$ 都是方程的解. 真实的解只有一个, 我们把那些并不符合真实物理意义的解称为伪解 (*spurious solution*).

为了排除伪解, 找到真实的 h_A 和 e_A , 我们需要从有限的场合入手. 想象我们把无穷的时间截断, 定义 $h_A^{(N)}(i) := \mathbb{P}_i[T_A \leq N]$ 为在 N 步内到达的概率. 直观上我们就应该有 $h_A = \lim_{N \rightarrow \infty} h_A^{(N)}$.

显然在限制步数 $N = 0$ 时, 若起点不在 A 则概率为 0. 而随着 N 增大, $h_A^{(N)}(i)$ 对于每个 i 都是单调递增的. 它的极限一定收敛到所有非负解中最小的那个, 恰好就是真实的首达概率——那些高于它的伪解自然被排除了. 这便是如下经典的最小非负解定理 (*minimal non-negative solution*).

定理 3.9 最小非负解定理.

给定马尔可夫链以及目标集合 $A \subseteq \Omega$.

$$p = 0.7 > \frac{1}{2}, \text{ 公比 } \frac{1-p}{p} = \frac{3}{7} < 1$$

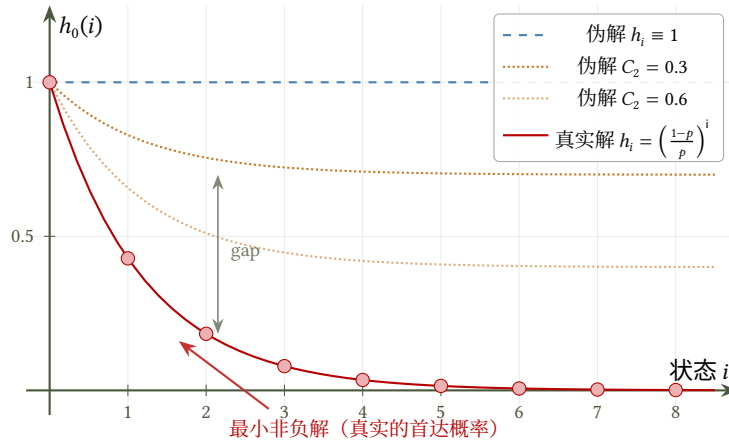


图 3.2: 真实解与伪解

1. 向量 $\vec{h}_A = (h_A(i))_{i \in \Omega}$ 是满足以下方程系统的最小非负解:

$$\begin{cases} x_i = 1, & i \in A, \\ x_i = \sum_{j \in \Omega} P(i, j) \cdot x_j, & i \notin A. \end{cases} \quad (3.2)$$

2. 向量 $\vec{e}_A = (e_A(i))_{i \in \Omega}$ 是满足以下方程系统的最小非负解:

$$\begin{cases} y_i = 0, & i \in A, \\ y_i = 1 + \sum_{j \in \Omega} P(i, j) \cdot y_j, & i \notin A. \end{cases} \quad (3.3)$$

这里的“最小”指的是坐标维度的偏序: 如果 \vec{z} 是方程的另一个非负解向量, 则一定对所有 $i \in \Omega$ 都有 $h_A(i) \leq z_i$ (或对 e_A 有 $e_A(i) \leq z_i$).

我们把这个定理的证明放在本小节最后.

利用最小非负解计算

有了最小非负解定理的保证, 我们便可以通过“列方程、求通解、取最小非负解”这一套路, 来系统地处理无穷状态马尔可夫链的首达问题了.

再次回到 \mathbb{N} 上的有偏随机游走. 为了算出从 $i \geq 1$ 走到 0 的首达概率 $h_i := h_0(i)$, 我们重写递推方程:

$$h_i = ph_{i+1} + (1-p)h_{i-1} \implies p(h_{i+1} - h_i) = (1-p)(h_i - h_{i-1}).$$

这是一个标准的常系数线性差分方程. 当 $p \neq 1/2$ 时, 特征方程有两个不同的根 1 和 $\frac{1-p}{p}$, 通解为:

$$h_i = C_1 + C_2 \left(\frac{1-p}{p} \right)^i.$$

这里的 C_1, C_2 是待定常数. 代入边界条件 $h_0 = 1$, 得到 $C_1 + C_2 = 1$, 即 $h_i = 1 - C_2 + C_2 \left(\frac{1-p}{p} \right)^i$.

- 当 $p > 1/2$ 时, 连乘因子 $\frac{1-p}{p} < 1$. 通解可以写作 $h_i = 1 - C_2 \left(1 - \left(\frac{1-p}{p} \right)^i \right)$. 为了保证对任意 i 都有 $h_i \geq 0$, 必须有 $C_2 \leq 1$. 满足约束的所有 $C_2 \in (-\infty, 1]$ 中, 哪个使 h_i 整体最小? 由于 $1 - \left(\frac{1-p}{p} \right)^i > 0$, C_2 越大则 h_i 越小, 所以取 $C_2 = 1$. 于是, 常数解 $h_i \equiv 1$ (对应 $C_2 = 0$) 作为伪解被排除, 真实的首达概率为 $h_i = \left(\frac{1-p}{p} \right)^i < 1$.
- 当 $p < 1/2$ 时, 因 $\frac{1-p}{p} > 1$, 如果 $C_2 < 0$, 随着 $i \rightarrow \infty$, 数列 h_i 将趋于负无穷, 违反非负性. 因此 $C_2 \geq 0$. 最小非负解对应 $C_2 = 0$, 即 $h_i \equiv 1$. 结论: 当向心的力量更强时, 无论初始位置距原点多远, 随机游走都以概率 1 返回 0.
- 当 $p = 1/2$ 时 (无偏游走), 特征方程有重根 1 , 通解退化为 $h_i = C_1 + C_2 i$. 代入边界条件 $h_0 = 1$ 得 $C_1 = 1$, 即 $h_i = 1 + C_2 i$. 若 $C_2 < 0$, 则当 i 足够大时 $h_i < 0$, 违反非负性. 因此 $C_2 \geq 0$, 最小非负解对应 $C_2 = 0$, 即 $h_i \equiv 1$. 结论与 $p < 1/2$ 相同: 无偏随机游走也以概率 1 返回原点.

既然 $p \leq 1/2$ 时肯定能返回, 那平均花费的时间 $e_i := e_0(i)$ 会是多少呢? 对应定理中第二组式子的差分方程为:

$$e_i = 1 + p e_{i+1} + (1-p) e_{i-1} \implies p(e_{i+1} - e_i) = (1-p)(e_i - e_{i-1}) - 1, \quad \forall i \geq 1.$$

这是一个非齐次方程. 令相邻两项的差为 $\Delta_i := e_i - e_{i-1}$, 则方程可化为关于 Δ_i 的一阶常系数差分方程:

$$p \Delta_{i+1} = (1-p) \Delta_i - 1 \implies \Delta_{i+1} = \frac{1-p}{p} \Delta_i - \frac{1}{p}.$$

我们在等式两边配凑一个常数, 寻找它的不动点, 即令 $x = \frac{1-p}{p} x - \frac{1}{p}$, 解得 $x = \frac{1}{1-2p}$. 于是方程可以改写为:

$$\Delta_{i+1} - \frac{1}{1-2p} = \frac{1-p}{p} \left(\Delta_i - \frac{1}{1-2p} \right).$$

这说明数列 $\Delta_i - \frac{1}{1-2p}$ 是一个等比数列. 我们由此可以写出 Δ_i 的通项: 对于某个待定常数 C , 有

$$\Delta_i = \frac{1}{1-2p} + C \left(\frac{1-p}{p} \right)^{i-1}.$$

由于 $e_0 = 0$, 我们有 $e_i = \sum_{j=1}^i \Delta_j$:

$$e_i = \sum_{j=1}^i \left(\frac{1}{1-2p} + C \left(\frac{1-p}{p} \right)^{j-1} \right) = \frac{i}{1-2p} + D \left(\left(\frac{1-p}{p} \right)^i - 1 \right).$$

其中 D 是由等比数列求和后重新整理得到的待定常数.

- 当 $p < 1/2$ 时, 连乘公比 $\frac{1-p}{p} > 1$. 如果选取 $D < 0$, 整个序列 e_i 很快就会因指数项的负向发散而变成负数. 所以我们必须让 $D \geq 0$. 要求其在所有满足条件的非负序列中最小, 自然是令 $D = 0$. 得到最终确切解: $e_i = \frac{i}{1-2p}$. 这与物理直觉高度吻合: 系统向着原点有一股稳定的漂移力量, 而期望游走时间随初始位置的距离线性增长.
- 当 $p = 1/2$ 时 (无偏游走), 差分方程退化为 $\Delta_{i+1} = \Delta_i - 2$, 于是 $\Delta_i = \Delta_1 - 2(i-1)$ 为等差数列. 对其求和得到

$$e_i = \sum_{j=1}^i \Delta_j = i\Delta_1 - i(i-1),$$

这是一个关于 i 的开口向下的二次函数. 无论 Δ_1 取何值, 当 i 足够大时 e_i 必将跌落至负数, 违反非负性. 这说明不存在有限的非负解, 唯一符合物理意义的解是 $e_i \equiv \infty$. 换言之, 无偏随机游走虽然以概率 1 返回原点, 但期望返回时间是无穷大的.

借助最小非负解定理, 我们将无穷状态马尔可夫链的首达问题系统地归结为求解差分方程并筛选最小非负解的代数问题. 不过, 上面的例子中显露出了截然不同的物理现象: 当 $p < 1/2$ 时不仅保证返回, 且期望时间有限; 而 $p = 1/2$ 时虽然必然返回 (概率为 1), 其期望游走时间却发散到了 ∞ . 这意味着对于无穷状态系统, 仅仅保证可达性是不够的. 我们在下一节将引入严格的状态分类体系, 来正式区分这些不同的回归行为.

最小非负解定理的证明

我们只证明 \vec{h}_A 是定理 3.2 的最小非负解, 对于 \vec{e}_A 的证明完全类似, 留作练习. 首先, \vec{h}_A 是方程的一个非负解这件事是显然的: $h_A(i) \geq 0$ 显然成立; 而对于 $i \notin A$, 对第一步落脚点 X_1 使用全概率公式和马尔可夫性, 直接得到

$$h_A(i) = P_i [T_A < \infty] = \sum_{j \in \Omega} P(i, j) P_j [T_A < \infty] = \sum_{j \in \Omega} P(i, j) h_A(j).$$

关键在于证明它是所有非负解中最小的. 为此我们引入截断逼近: 定义 N 步内的首达概率 $h_A^{(N)}(i) := P_i [T_A \leq N]$. 由于事件序列 $\{T_A \leq N\}$ 单调递增并收敛到 $\{T_A < \infty\}$, 有 $\lim_{N \rightarrow \infty} h_A^{(N)}(i) = h_A(i)$.

假设 $\vec{x} = (x_i)_{i \in \Omega}$ 是定理 3.2 的任意一个非负解, 我们对 N 进行数学归纳法, 证明对于每一个 i , 都有 $x_i \geq h_A^{(N)}(i)$ 成立. 当 $N = 0$ 时, 对于 $i \in A$ 显然有 $x_i = 1 = h_A^{(0)}(i)$; 对于 $i \notin A$, 因为前提是非负解, 所以自然有 $x_i \geq 0 = h_A^{(0)}(i)$. 这是归纳起始. 假设对某个 N , 所有 j 均有 $x_j \geq h_A^{(N)}(j)$. 那么对于 $i \notin A$, 利用归纳假设逐项放缩:

$$x_i = \sum_{j \in \Omega} P(i, j)x_j \geq \sum_{j \in \Omega} P(i, j)h_A^{(N)}(j) = h_A^{(N+1)}(i).$$

结合对于 $i \in A$ 恒有 $x_i = 1 = h_A^{(N+1)}(i)$, 归纳步骤完成. 这就说明对于一切 N 和 i , 都有 $x_i \geq h_A^{(N)}(i)$ 成立. 取 $N \rightarrow \infty$ 的极限即得 $x_i \geq h_A(i)$, 至此 \vec{h}_A 的最小性得证.

时间反演视角：逃离与命中的对称

我们在这一节介绍一个关于首达时间的一个有趣性质. 首先引入一个完全不同的视角: 利用时间反演链, 我们可以把一段轨迹倒过来看. 在正放时候“从集合 S 逃离并在 t 步后出现在 y ”与倒放时候“从 y 出发并在第 t 步首次命中 S ”描述的是同一批路径, 因此在使用平稳分布 π 对端点加权后, 二者的概率相等.

回忆对于任意集合 $S \subseteq \Omega$, 我们已经定义了

$$T_S^+ = \inf\{t \geq 1 : X_t \in S\}.$$

如果 $(\hat{X}_t)_{t \geq 0}$ 是时间反演链对应的过程, 我们相应地记

$$\hat{T}_S^+ := \inf\{t \geq 1 : \hat{X}_t \in S\}.$$

命题 3.10 逃离与命中的对偶.

设 $(X_t)_{t \geq 0}$ 是可数状态空间 Ω 上的马尔可夫链, 它有一个平稳分布 π , 并且满足 $\forall i \in \Omega, \pi(i) > 0$. 记 \hat{P} 为第 2.11 章定义的时间反演链, $(\hat{X}_t)_{t \geq 0}$ 为以 \hat{P} 为转移核的马尔可夫链. 则对于任意非空集合 $S \subseteq \Omega$ 、任意 $x \in S$ 、 $y \in \Omega$ 以及任意整数 $t \geq 1$, 有

$$\pi(x)P_x[T_S^+ \geq t, X_t = y] = \pi(y)P_y[\hat{T}_S^+ = t, \hat{X}_t = x],$$

其中左边的概率在原链下计算, 右边的概率在时间反演链下计算.

证明. 固定 $x \in S$ 、 $y \in \Omega$ 和 $t \geq 1$. 左边事件对应的正是所有满足下列条件的路径

$$z_0, z_1, \dots, z_t : z_0 = x, z_1, \dots, z_{t-1} \notin S, z_t = y.$$

也就是说, 这条路径从 S 中的点 x 出发, 在前 $t-1$ 步始终没有回到 S , 并在第 t 步到达 y .

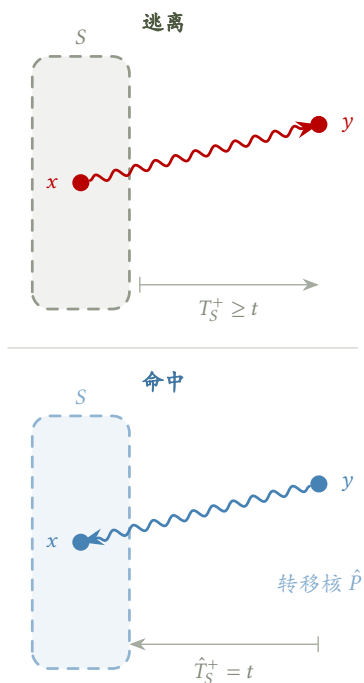


图 3.3: 逃离与命中

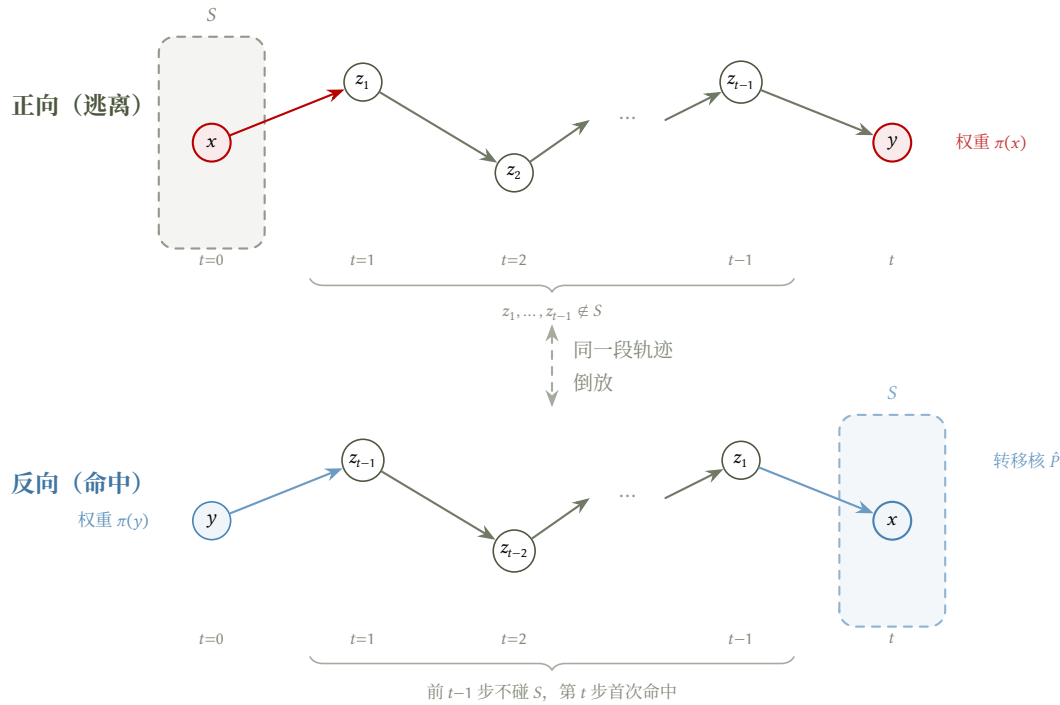


图 3.4: 逃离与命中的对偶: 同一段轨迹的正放与倒放

对于任意一条这样的路径, 把第二章中时间反演链的定义 2.3 沿着整条路径连乘, 便得到

$$\pi(z_0) \prod_{k=1}^t P(z_{k-1}, z_k) = \pi(z_t) \prod_{k=1}^t \hat{P}(z_k, z_{k-1}).$$

将 $z_0 = x$ 、 $z_t = y$ 代入, 可写成

$$\pi(x) \prod_{k=1}^t P(z_{k-1}, z_k) = \pi(y) \prod_{k=1}^t \hat{P}(z_k, z_{k-1}).$$

现在对所有满足 $z_1, \dots, z_{t-1} \notin S$ 的中间点求和. 左边恰好得到

$$\pi(x) P_x [T_S^+ \geq t, X_t = y].$$

而右边则是在反向链中, 对所有从 y 出发、前 $t-1$ 步都没有碰到 S 、并在第 t 步首次命中 S 且命中点为 x 的路径求和, 因此正是

$$\pi(y) P_y [\hat{T}_S^+ = t, \hat{X}_t = x].$$

两边相等, 命题得证.

□

这个引理的直观意义很有趣. 正向看时, x 是轨迹逃离 S 时最后一次在集合 S 的地方; 反向看时, 同一段轨迹却变成了一条在第 t 步首次命中集合 S 的路径, 而且命中点正好也是 x . 所谓“逃离”与“命中”, 只是同一段录像在正放与倒放时呈现出的两种面貌.

3.4 状态的分类

我们通过上一节的计算发现, 无穷状态的马尔可夫链并不像有限状态那样必定能在有限的时间步内返回. 有偏随机游走的例子表明: 系统离开某状态后可能永远不再返回 ($p > 1/2$); 可能必然返回, 但平均需要耗费无穷长的时间 ($p = 1/2$); 只有当向左的力量足够强时, 不但必然返回, 且期望等待时间是有限的 ($p < 1/2$). 这三种截然不同的物理现象对应了系统三种不同的状态.

常返与瞬时

我们首先关心的基本问题是: 系统离开某个状态 i 后, 在未来的演化中是否还能以概率 1 回到该状态?

定义 3.11 常返 (recurrent) 与瞬时 (transient) .

对于 $i \in \Omega$, 回忆定义 3.2 中定义的返回时间 $T_i^+ = \inf\{t \geq 1 : X_t = i\}$. 若 $\mathbb{P}_i[T_i^+ < \infty] = 1$, 则称状态 i 是常返 (recurrent) 的; 否则 (即 $\mathbb{P}_i[T_i^+ < \infty] < 1$), 称其为瞬时 (transient) 的.

常返性刻画了马尔可夫链“从 i 出发定会回到 i ”的性质. 结合强马尔可夫性, 如果 i 是常返的, 这意味着不仅第一次必定能回来, 而且在每一次回归的瞬间, 系统都能抛却历史进行一次无损的重启. 由于每次重启后的演化规律一致, 系统将以概率 1 发生无数次回归.

我们可以用一个更为宏观的视角来刻画这种现象——考察系统对状态的总访问次数. 定义随机变量 $N_i := \sum_{t=0}^{\infty} \mathbf{1}[X_t = i]$ 为系统演化过程中访问状态 i 的总次数 (包含初始的 $t = 0$). 接下来的等价定理将有关发生次数的概率问题, 转化为一道关于转移概率级数敛散性的分析题.

对于一般的非负随机变量 Y , $\mathbb{P}[Y = \infty] = 1$ 显然比 $\mathbb{E}[Y] = \infty$ 强得多. 然而, 对于特定状态的访问次数 N_i , 这两者在 \mathbb{P}_i 的测度下是等价的.

第 3 点中的等价是单调收敛定理的简单推论:

$$\begin{aligned} \mathbb{E}_i[N_i] &= \mathbb{E}_i\left[\sum_{t=0}^{\infty} \mathbf{1}[X_t = i]\right] \\ &= \sum_{t=0}^{\infty} \mathbb{E}_i[\mathbf{1}[X_t = i]] \\ &= \sum_{t=0}^{\infty} \mathbb{P}_i[X_t = i] = \sum_{t=0}^{\infty} P^t(i, i). \end{aligned}$$

定理 3.12 常返性的等价刻画.

设 i 为可数状态马尔可夫链的一个状态, 则以下三个条件等价:

1. i 是常返的.

2. 从 i 出发, 系统以概率 1 访问 i 无穷多次, 即 $\mathbb{P}_i[N_i = \infty] = 1$.
3. 期望访问总次数发散, 即 $\mathbf{E}_i[N_i] = \infty$, 并且这等价于 $\sum_{t=0}^{\infty} P^t(i, i) = \infty$.

证明. $1 \implies 2$: 假设 $\mathbb{P}_i[T_i^+ < \infty] = 1$. 对于任何正整数 k , 利用强马尔可夫性可知系统连续发生第 k 次回归的概率为 $(\mathbb{P}_i[T_i^+ < \infty])^k = 1$. 由于这对于任意的 k 都成立, 说明状态会被以概率 1 访问无穷多次.

$2 \implies 3$: 由于事件 $[N_i = \infty]$ 几乎必然发生, 其期望 $\mathbf{E}_i[N_i] = \infty$ 显然成立.

$3 \implies 1$: 我们采用反证法. 若认定 i 是瞬时的, 即 $\mathbb{P}_i[T_i^+ < \infty] = \alpha < 1$. 由强马尔可夫性, 系统每次返回 i 后都独立地以概率 α 再次返回. 因此, 除去初始访问 ($t = 0$ 时刻), 后续的回归次数 $N_i - 1$ 服从参数为 $1 - \alpha$ 的几何分布, 即 $\mathbf{E}_i[N_i - 1] = \frac{\alpha}{1 - \alpha}$. 于是 $\mathbf{E}_i[N_i] = 1 + \frac{\alpha}{1 - \alpha} = \frac{1}{1 - \alpha} < \infty$. 这与条件 $\mathbf{E}_i[N_i] = \infty$ 相悖, 于是 α 必须等于 1. \square

互通类的类性质

我们之前已经证明过, 对于不可约的马尔可夫链, 无周期性是一种类性质. 实际上, 常返与瞬时这两个性质对于不可约链也有同样的性质——在同一个互通等价类内的所有状态, 必然共享同一种回归属性.

命题 3.13 常返是类性质.

若 i 是常返的, 并且与 j 互通 ($i \leftrightarrow j$), 则 j 也是常返的.

本定理提供两个视角的证明: 纯代数视角的级数放缩十分简练; 而基于第一步耦合的分析则展现了更强的物理演化直觉.

基于查普曼-科尔莫戈罗夫方程的代数证明. 由于 $i \leftrightarrow j$, 存在有限步数 $r, s \geq 0$ 使得 $P^r(i, j) > 0$ 且 $P^s(j, i) > 0$. 利用查普曼-科尔莫戈罗夫方程, 从 j 出发先用 s 步走到 i , 在 i 处游走 t 步回到 i , 再用 r 步回到 j , 可得:

$$P^{s+t+r}(j, j) \geq P^s(j, i) \cdot P^t(i, i) \cdot P^r(i, j).$$

对所有 $t \geq 0$ 求和:

$$\sum_{t=0}^{\infty} P^{s+t+r}(j, j) \geq P^s(j, i) P^r(i, j) \sum_{t=0}^{\infty} P^t(i, i).$$

由于 i 为常返状态, 右侧的级数 $\sum P^t(i, i) = \infty$, 而系数 $P^s(j, i) P^r(i, j) > 0$, 因此右侧 $= \infty$. 左侧 $\sum_{t=0}^{\infty} P^{s+t+r}(j, j) = \sum_{n=s+r}^{\infty} P^n(j, j) \leq \sum_{n=0}^{\infty} P^n(j, j)$, 故 $\sum_{n=0}^{\infty} P^n(j, j) = \infty$, 即 j 也是常返的. \square

基于路径试验的直觉证明. 首先说明从 i 必定可以到达 j , 即 $\mathbb{P}_i[T_j < \infty] = 1$. 由于 $i \leftrightarrow j$, 假设某条有限连通路径的转移成功概率为 $q > 0$. 由于 i 是常返的, 这意味着系统会进行无数

次以 i 结尾的试错轮回. 在每一次新的轮回中, 系统都有 q 的先验概率沿着特定路径到达 j . 如果系统永远避开 j , 这要求无数次的独立几何试验均失败, 其概率界限为 $\lim_{k \rightarrow \infty} (1 - q)^k = 0$. 因此, 只要 $q > 0$, 系统必然能到达 j .

接着说明反向保证 $P_j[T_i < \infty] = 1$. 若 $P_j[T_i < \infty] = \alpha < 1$, 考虑这一组合事件: 系统从 i 优先以概率 q 跨越到 j , 随后在 j 的下层演化中遭遇到永远不再返回 i 的游走. 该事件发生的联合概率不低于 $q(1 - \alpha) > 0$. 这就意味着从 i 出发的系统存在严格大于 0 的概率不再返回 i . 这违背了初始给定的回归保证 $P_i[T_i^+ < \infty] = 1$. 因此 α 必须为 1.

于是, 我们知道如果系统从 j 出发, 有限步到达了 i , 而从 i 出发又有限步到达了 j , 那么系统一定从 j 出发有限步返回了 j . 由强马尔可夫性, 这两段旅程在停时处条件独立, 因此

$$P_j[T_j^+ < \infty] \geq P_j[T_i < \infty] \cdot P_i[T_j < \infty] = 1.$$

从而完成了 j 为常返状态的验证. \square

因此, 我们知道, 如果马尔可夫链是不可约的, 那么它的所有状态必定要么都是常返的, 要么都是瞬时的. 我们因此会说该马尔可夫链本身是常返的, 或者是瞬时的.

醉汉与醉鸟

借助级数 $\sum_{t=0}^{\infty} P^t(i, i)$ 的敛散性判据, 我们可以通过纯粹的解析计算来判定特定马尔可夫链的常返性. 下面是一个广为流传的经典例子: 在 \mathbb{Z}^d 上的随机游走是否是常返的?

“A drunk man will find his way home,
but a drunk bird may get lost forever.”

— 角谷静夫 (Shizuo Kakutani)

这里的模型与 Pólya (1921) 的经典模型稍有不同. Pólya 研究的是标准最近邻随机游走 (simple random walk), 即每步从 $2d$ 个坐标轴方向中均匀随机选取一个, 移动一个单位. 我们这里采用的是坐标独立的乘积游走 (product walk), 每步同时所有坐标上独立移动 ± 1 . 两种模型的常返/瞬时的维数阈值 $d \leq 2$ 相同, 但乘积游走的计算更为简洁, 因此我们以它为例.

示例 3.14 随机游走的维数相变.

考虑 d 维整数格点 \mathbb{Z}^d 上的乘积随机游走 (product random walk): 在每一步中, 对每个坐标 $\ell \in [d]$, 独立均匀随机地选取 $\Delta_\ell \in \{-1, +1\}$, 然后令 $X_{t+1} = X_t + (\Delta_1, \dots, \Delta_d)$. 由于各坐标独立, d 维乘积游走等价于 d 个独立的一维随机游走. 由于奇偶性, 游走只在偶数步才可能回到原点. 利用斯特林公式, 一维游走在第 $2m$ 步回到原点的概率满足

$$P_0[X_{2m} = 0] = \frac{\binom{2m}{m}}{2^{2m}} \approx \frac{1}{\sqrt{\pi m}}.$$

由于各坐标独立, d 维游走在第 $2m$ 步回到原点的概率为

$$P_0[X_{2m} = \mathbf{0}] = (P_0[X_{2m} = 0])^d \approx (\pi m)^{-d/2}.$$

于是常返性的判定归结为级数 $\sum_{m=1}^{\infty} m^{-d/2}$ 的敛散性. 由微积分中的结论, 该级数当且仅当指数 $d/2 \leq 1$ (即 $d \leq 2$) 时发散.

斯特林公式: $n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n (1 + o(1))$.

于是，我们有结论：在一维和二维上漫游的醉汉，终将以概率 1 回到出发点（常返）；而从三维开始（ $d \geq 3$ ），空间增长得太快，醉鸟有正概率永远迷失（瞬时）。

正常返与零常返

仅用常返和瞬时两个大类来界定系统的性质是不完备的。常返性保证了系统一定会返回（即 $P_i[T_i^+ < \infty] = 1$ ），但并未限定完成一次回归所需的时间。为了刻画这一差别，我们考察回归时间的期望值 $m_i := E_i[T_i^+]$ 。

定义 3.15 正常返（positive recurrent）与零常返（null recurrent）。

对于常返状态 i ：如果回归的期望时间有限，即 $m_i = E_i[T_i^+] < \infty$ ，则称其为正常返。

如果回归的期望时间发散到无穷，即 $m_i = E_i[T_i^+] = \infty$ ，则称其为零常返。

在这里我们可以回顾上一节中一维有偏随机游走的计算结果。对于原点 0，利用全期望公式对第一步展开：

$$m_0 = E_0[T_0^+] = (1-p) \cdot 1 + p \cdot (1 + E_1[T_0]) = 1 + p e_1,$$

其中 e_1 是上节求出的从状态 1 到原点的期望首达时间。代入已有结论：

- $p < 1/2$: $e_1 = \frac{1}{1-2p}$ 是有限的，因此 $m_0 < \infty$ ，系统为正常返。
- $p = 1/2$: 首达概率仍为 1（常返），但 $e_1 = \infty$ ，于是 $m_0 = \infty$ ，系统为零常返。
- $p > 1/2$: 返回概率严格小于 1，系统为瞬时。

与常返性类似，对于互通的点来说，正常返也是一种类性质——当满足 $i \leftrightarrow j$ 互通时， i 为正常返当且仅当 j 也为正常返。我们把这一性质的证明留作习题。

区分正常返与零常返是很重要的。在有限状态空间中，边界的束缚使得不可约的系统都是正常返（用下一节的结论很容易证明）。但当状态空间变为无穷时，判断是否正常返会变得复杂。在接下来的一节中，我们将看到：正常返与平稳分布的存在性密切相关。

3.5 马尔可夫链基本定理（可数无穷状态）

经过前几节对首达时间、状态分类与回归性质的讨论，我们已经为无穷状态马尔可夫链的极限定理做好了准备。这一基本定理推广了有限状态空间下的极限行为，并且将平稳分布的 $\pi(i)$ 与期望返回时间 $m_i := E_i[T_i^+]$ 建立了精确的联系。

注意到，基本定理和有限状态空间的唯一区别就是把有限状态换成了正常返。我们马上可以看到，对于不可约链，正常返是平稳分布存在的充要条件。

定理 3.16 马尔可夫链基本定理（可数无穷状态）.

设状态空间可数的马尔可夫链是不可约、非周期且正常返的，则：

1. 系统存在唯一的平稳分布 π ，且对于所有状态 $i \in \Omega$ ， $\pi(i) > 0$.
2. 平稳分布与各个状态的期望返回时间具有严格倒数关系：

$$\pi(i) = \frac{1}{m_i} = \frac{1}{\mathbf{E}_i [T_i^+]} \quad \forall i \in \Omega.$$

3. 对于任意初始状态 j ，转移概率按时间极限收敛至此平稳分布：

$$\lim_{t \rightarrow \infty} P^t(j, i) = \pi(i).$$

证明分为三部分：首先利用更新周期理论构造出平稳分布，证明其存在性；接着使用我们介绍过的逃离与命中的对称关系，证明平稳分布的唯一性；最后，我们使用与有限场合相同的方法——耦合法——证明极限收敛。

存在：返回周期的更新分解**引理 3.17 存在性.**

设状态空间可数的马尔可夫链是不可约且正常返的，则系统存在一个平稳分布 π ，且满足 $\pi(i) > 0$ 对所有 $i \in \Omega$ 成立。

证明. 固定任意一个状态 $k \in \Omega$ 。由于链是不可约且正常返的，期望回归时间 $m_k < \infty$ 。在一个完整的返回周期 $[0, T_k^+)$ 内，我们考察系统在任意状态 j 上的期望停留次数：

$$\tilde{\pi}(j) := \mathbf{E}_k \left[\sum_{t=0}^{T_k^+-1} \mathbf{1}[X_t = j] \right].$$

容易验证，由于计数从 $t = 0$ （此时 $X_0 = k$ ）开始，到 $t = T_k^+ - 1$ 截止（不含再次回到 k 的那一步），可得 $\tilde{\pi}(k) = 1$ 。对所有状态求和，总的期望停留次数等于回归周期的期望长度：

$$\sum_{j \in \Omega} \tilde{\pi}(j) = \mathbf{E}_k \left[\sum_{t=0}^{T_k^+-1} \sum_{j \in \Omega} \mathbf{1}[X_t = j] \right] = \mathbf{E}_k [T_k^+] = m_k < \infty.$$

我们要证明经过归一化 $\pi(j) = \tilde{\pi}(j)/m_k$ 后得到的就是平稳分布。为此只需验证 $\tilde{\pi}P = \tilde{\pi}$ 。对任意状态 j ，计算：

$$\begin{aligned} \sum_{i \in \Omega} \tilde{\pi}(i) P(i, j) &= \sum_{i \in \Omega} \mathbf{E}_k \left[\sum_{t=0}^{T_k^+ - 1} \mathbb{1}[X_t = i] \right] \cdot P(i, j) \\ &\stackrel{\triangleright \text{Fubini-Tonelli 定理}}{=} \mathbf{E}_k \left[\sum_{t=0}^{T_k^+ - 1} \sum_{i \in \Omega} \mathbb{1}[X_t = i] \cdot P(i, j) \right]. \end{aligned}$$

接下来注意到, 对每个固定的时刻 t , 内层求和中指示函数 $\mathbb{1}[X_t = i]$ 在 $i = X_t$ 时取值 1, 其余为 0, 因此

$$\sum_{i \in \Omega} \mathbb{1}[X_t = i] \cdot P(i, j) = P(X_t, j).$$

由马尔可夫性, $P(X_t, j) = \mathbb{P}[X_{t+1} = j | X_t]$, 故

$$\sum_{i \in \Omega} \tilde{\pi}(i) P(i, j) = \mathbf{E}_k \left[\sum_{t=0}^{T_k^+ - 1} \mathbb{P}[X_{t+1} = j | X_t] \right] = \mathbf{E}_k \left[\sum_{t=0}^{T_k^+ - 1} \mathbb{1}[X_{t+1} = j] \right].$$

平移一下求和的指标, 我们得到:

$$\begin{aligned} \mathbf{E}_k \left[\sum_{t=0}^{T_k^+ - 1} \mathbb{1}[X_{t+1} = j] \right] &= \mathbf{E}_k \left[\sum_{t=1}^{T_k^+} \mathbb{1}[X_t = j] \right] \\ &= \mathbf{E}_k \left[\sum_{t=0}^{T_k^+ - 1} \mathbb{1}[X_t = j] \right] + \mathbf{E}_k [\mathbb{1}[X_{T_k^+} = j] - \mathbb{1}[X_0 = j]] \\ &= \tilde{\pi}(j). \end{aligned}$$

至此我们证明了 $\tilde{\pi}P = \tilde{\pi}$. 由于链不可约, 对任意 $j \in \Omega$, 从 k 出发存在一条在中途不经过 k 便到达 j 的路径, 因此 $\tilde{\pi}(j) > 0$. 加之 $\tilde{\pi}$ 的总和为 $m_k < \infty$, 除以 m_k 即得一个各分量严格为正的合法概率分布. 特别地, $\pi(k) = 1/m_k$ 揭示了平稳分布与期望回归时间之间的精确联系. \square

在证明唯一性之前, 我们首先确认一个重要事实: 对于不可约链, 平稳分布的存在已经蕴含了常返性.

命题 3.18 平稳分布蕴含常返性.

设状态空间可数的马尔可夫链是不可约的, 并且存在平稳分布 π 满足 $\forall i \in \Omega, \pi(i) > 0$. 则该链是常返的.

证明. 假设链是瞬时的. 由于链不可约, 瞬时性是类性质 (命题 3.13), 故所有状态均为瞬时的. 首先, 对于瞬时状态 j , 由常返等价刻画定理知 $\sum_{t=0}^{\infty} P^t(j, j) < \infty$, 收敛级数的通项必

注意 $\tilde{\pi}$ 的构造依赖于所选取的状态 k : 选取不同的 k 会得到不同的 $\tilde{\pi}$, 归一化后各自都给出一个平稳分布. 但我们尚不知道平稳分布是否唯一——如果不唯一, 不同的 k 可能构造出不同的平稳分布. 我们目前能确认的仅仅是: 对于固定的基准点 k , 构造出来的平稳分布满足 $\pi(k) = 1/m_k$. 要推广到对所有 i 成立的 $\pi(i) = 1/m_i$, 需要接下来唯一性部分的论证.

这个命题成立的直观是: 瞬时链终将遗忘每一个状态——从任何起点出发, 在时间足够长后出现在任何给定状态 j 的概率趋于零. 但平稳分布意味着“永不遗忘”——系统出现在 j 的概率始终是 $\pi(j) > 0$. 二者不兼容.

趋于零, 故 $P^t(j, j) \rightarrow 0$. 接下来, 对于任意 $i \neq j$, 由不可约性存在 $s \geq 0$ 使得 $P^s(j, i) > 0$. 利用查普曼-科尔莫戈罗夫方程:

$$P^{s+t}(j, j) \geq P^s(j, i) \cdot P^t(i, j),$$

因此 $P^t(i, j) \leq P^{s+t}(j, j)/P^s(j, i) \rightarrow 0$. 至此我们证明了: 对所有 $i, j \in \Omega$, $P^t(i, j) \rightarrow 0$.

但 π 是平稳分布, 因此对任意 t 和 j 均有 $\pi(j) = \sum_{i \in \Omega} \pi(i) P^t(i, j)$. 取 $t \rightarrow \infty$, 由控制收敛定理 ($\pi(i) P^t(i, j) \leq \pi(i)$ 且 $\sum_i \pi(i) = 1$), 得

$$\pi(j) = \sum_{i \in \Omega} \pi(i) \lim_{t \rightarrow \infty} P^t(i, j) = 0,$$

与 $\pi(j) > 0$ 矛盾. 因此链必为常返. □

唯一性: 基于逃离与命中的对称关系

我们接着证明, 如果系统存在一个平稳分布 π , 那么一定有 $\pi(i) = \frac{1}{\mathbf{E}_i[T_i^+]}$.

引理 3.19 唯一性.

设状态空间可数的马尔可夫链是不可约且常返的, 并且其存在平稳分布 π , 满足对于任意 $i \in \Omega$, $\pi(i) > 0$. 则该平稳分布是唯一的, 且对于所有状态 $i \in \Omega$, $\pi(i) = \frac{1}{\mathbf{E}_i[T_i^+]}$. 同时, 这个马尔可夫链也是正常返的.

证明. 固定 $i \in \Omega$. 我们有

$$\mathbf{E}_i[T_i^+] = \sum_{t=1}^{\infty} \mathbb{P}_i[T_i^+ \geq t] = \sum_{t=1}^{\infty} \sum_{j \in \Omega} \mathbb{P}_i[T_i^+ \geq t, X_t = j].$$

由于各项都非负, 可用 Fubini-Tonelli 定理交换求和次序, 于是

$$\pi(i) \cdot \mathbf{E}_i[T_i^+] = \sum_{j \in \Omega} \sum_{t=1}^{\infty} \pi(i) \cdot \mathbb{P}_i[T_i^+ \geq t, X_t = j].$$

现在在 [命题 3.10](#) 中取 $S = \{i\}$, $x = i$. 于是对任意 $j \in \Omega$ 和任意 $t \geq 1$, 都有

$$\pi(i) \cdot \mathbb{P}_i[T_i^+ \geq t, X_t = j] = \pi(j) \cdot \mathbb{P}_j[\hat{T}_i^+ = t, \hat{X}_t = i].$$

代回上式, 得到

$$\begin{aligned} \pi(i) \mathbf{E}_i[T_i^+] &= \sum_{j \in \Omega} \sum_{t=1}^{\infty} \pi(j) \cdot \mathbb{P}_j[\hat{T}_i^+ = t, \hat{X}_t = i] \\ &= \sum_{j \in \Omega} \pi(j) \cdot \sum_{t=1}^{\infty} \mathbb{P}_j[\hat{T}_i^+ = t], \end{aligned}$$

其中最后一步用了事件 $[\hat{T}_i^+ = t]$ 已经蕴含 $\hat{X}_t = i$. 因此

$$\pi(i) \cdot \mathbf{E}_i [T_i^+] = \sum_{j \in \Omega} \pi(j) \cdot \mathbb{P}_j [\hat{T}_i^+ < \infty].$$

一个简单的事实是: 关于 π 的时间反演链 \hat{P} 保持不可约性与常返性 (我们留作练习). 于是 \hat{P} 也是不可约常返的, 对任意 $j \in \Omega$,

$$\mathbb{P}_j [\hat{T}_i^+ < \infty] = 1.$$

代回上式便得到

$$\pi(i) \mathbf{E}_i [T_i^+] = \sum_{j \in \Omega} \pi(j) = 1.$$

由此直接得到 $m_i = \mathbf{E}_i [T_i^+] = 1/\pi(i)$. 由于 $\pi(i) > 0$, 故 $m_i < \infty$, 即该马尔可夫链是正常返的. \square

收敛性: 耦合法证明

接下来证明, 随着时间的推移, 从任意初始状态出发的转移概率必然收敛于上面构造出的平稳分布 π , 即 $\lim_{t \rightarrow \infty} P^t(j, i) = \pi(i)$.

我们使用和第二章完全相同的耦合框架. 回忆那里的证明策略: 构造两条链 \bar{X}_t (从任意初始状态出发) 和 \bar{Y}_t (从平稳分布 π 出发), 让它们独立演化直到碰撞, 碰撞后黏合在一起. 由耦合引理, $|P^t(j, i) - \pi(i)| \leq \mathbb{P}[\tau > t]$, 所以只要证明碰撞时间 τ 几乎必然有限, 收敛性就成立了.

在第二章中, 碰撞是这样保证的: 我们利用了有限不可约无周期链的一个关键性质——存在全局的 t^* , 使得 $P^{t^*}(i, j) > 0$ 对所有状态对 i, j 同时成立 (回忆这个 t^* 的存在依赖于 $t^* = \max_{i, j} t_{ij}$, 有限状态空间保证最大值有限). 有了全局的 t^* , 我们可以定义 $Q = P^{t^*}$, 此时 $Q(i, j) \geq \delta > 0$ 对所有 i, j 一致成立, 从而每一步两条独立链碰撞的概率至少为 $n\delta^2 > 0$, 碰撞时间不仅几乎必然有限, 而且以指数速率衰减.

当状态空间变成可数无穷时, 这个论证失效了. 虽然对每一对固定的 i, j , 仍然存在 t_{ij} 使得 $P^t(i, j) > 0$ 对所有 $t \geq t_{ij}$ 成立 (这一点只依赖于不可约和无周期, 与状态空间是否有限无关), 但 $\sup_{i, j} t_{ij}$ 可能是 $+\infty$ ——全局一致的 t^* 不再存在, 从而也没有一致正下界 δ . 因此, 我们要用另外的方法来说明, 两个链依然以概率 1 会撞在一起.

我们仍然用同样的方式构造耦合过程 (\bar{X}_t, \bar{Y}_t) :

1. $\bar{X}_0 = j$ (固定初始状态), $\bar{Y}_0 \sim \pi$ (从平稳分布采样).
2. 碰撞前: 两条链各自独立地按 P 演化.

3. 碰撞后: 定义耦合时间 $\tau = \inf\{t \geq 0 : \bar{X}_t = \bar{Y}_t\}$. 一旦 $t \geq \tau$, 两链黏合在一起, 即 $\bar{X}_t \equiv \bar{Y}_t$, 共同按 P 演化.

由耦合引理, 我们有

$$|P^t(j, i) - \pi(i)| \leq \mathbb{P}[\bar{X}_t \neq \bar{Y}_t] \leq \mathbb{P}[\tau > t].$$

因此, 只需证明 $\mathbb{P}[\tau < \infty] = 1$.

为此, 我们观察碰撞发生之前乘积过程的结构. 此时两条链独立演化, (\bar{X}_t, \bar{Y}_t) 构成 $\Omega \times \Omega$ 上的马尔可夫链, 一步转移概率为

$$P_{\otimes}((x_1, y_1), (x_2, y_2)) = P(x_1, x_2) \cdot P(y_1, y_2).$$

两条链碰撞等价于乘积链到达对角线 $D = \{(x, x) : x \in \Omega\}$. 因此 $\mathbb{P}[\tau < \infty] = 1$ 等价于: 乘积链从任意初始状态出发, 几乎必然到达 D .

我们验证乘积链是不可约、无周期且正常返的——由此可推出它几乎必然访问每一个状态, 特别是 D 中的状态. 虽然全局的 t^* 不存在了, 但逐对有限的 t_{ij} 足够完成以下验证:

- **不可约:** 对任意 (x_1, y_1) 和 (x_2, y_2) , 取 $t \geq \max(t_{x_1 x_2}, t_{y_1 y_2})$, 则 $P_{\otimes}^t((x_1, y_1), (x_2, y_2)) > 0$. 注意这里只需要逐对的 t_{ij} 有限, 不需要对所有 (x_1, y_1, x_2, y_2) 统一取 t^* .
- **无周期:** 对任意状态 (x, y) , 取 $t \geq \max(t_{xx}, t_{yy})$, 则 $P_{\otimes}^t((x, y), (x, y)) = P^t(x, x) \cdot P^t(y, y) > 0$. 这对所有充分大的 t 都成立, 故周期为 1.
- **正常返:** 乘积测度 $\pi \otimes \pi$ 是 P_{\otimes} 的平稳分布 (因为两个坐标独立, 各自满足 $\pi P = \pi$), 且各分量 $(\pi \otimes \pi)(x, y) = \pi(x)\pi(y) > 0$. 由命题 3.18, 不可约链存在正的平稳分布则必为常返的; 再由引理 3.19, 该链是正常返的.

既然乘积链是不可约且正常返的, 它几乎必然访问每一个状态, 特别是对角线 D 中的状态. 因此

$$\mathbb{P}[\tau < \infty] = 1.$$

至此, 结合唯一性部分 (引理 3.19) 的结论 $\pi(i) = 1/m_i$ 和平稳分布的唯一性, 基本定理的全部三条结论均已得证.

这里的关键转变是: 在有限状态空间中, 我们通过每步碰撞概率的一致正下界来推出碰撞时间几乎必然有限 (甚至指数衰减); 在无穷状态空间中, 我们换一种方法——将碰撞问题转化为乘积链的常返性问题.

这里用到的基本事实是: 对不可约无周期链, $\forall i, j, \exists t_{ij}$ 使得 $P^t(i, j) > 0$ 对所有 $t \geq t_{ij}$ 成立. 其证明与第二章相同: 无周期条件保证 $\exists t_i$ 使得 $\forall t \geq t_i, P^t(i, i) > 0$; 再结合 $i \rightarrow j$ ($\exists s, P^s(i, j) > 0$), 得 $P^t(i, j) \geq P^{t-s}(i, i) \cdot P^s(i, j) > 0$ 对 $t \geq t_i + s$ 成立.

参考文献

- [Tao12] Terence Tao. *Topics in Random Matrix Theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012. 8