

第四章 随机过程的熵率

- 随机过程 $\{X_i\}$: 带下标的随机变量序列

$$\Pr\{(X_1, X_2, \dots, X_n) = (x_1, x_2, \dots, x_n)\} = p(x_1, x_2, \dots, x_n),$$
$$(x_1, x_2, \dots, x_n) \in \mathcal{X}^n \text{ for } n = 1, 2, \dots$$

- 平稳的随机过程:

$$\Pr\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}$$
$$= \Pr\{X_{1+l} = x_1, X_{2+l} = x_2, \dots, X_{n+l} = x_n\}$$

马尔可夫过程

- 马尔可夫过程（马尔可夫链）：

$$\Pr(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_1 = x_1)$$

$$= \Pr(X_{n+1} = x_{n+1} | X_n = x_n)$$

$$p(x_1, x_2, \dots, x_n) = p(x_1)p(x_2|x_1)p(x_3|x_2) \cdots p(x_n|x_{n-1}).$$

- 时间不变的马尔可夫过程：

$$\Pr\{X_{n+1} = b | X_n = a\} = \Pr\{X_2 = b | X_1 = a\}, \text{ for all } a, b \in \mathcal{X}.$$

- 如无特别声明，总假定马尔可夫链是时间不变的。

马尔可夫链的表征

- 若 $\{X_i\}$ 为马尔可夫链，则称 X_n 为n时刻的状态。
- 一个时间不变的马尔可夫链完全由其初始状态和概率转移矩阵 $P=[P_{ij}]$ 所表征，

$$P_{ij} = \Pr\{X_{n+1} = j | X_n = i\}$$

- $n+1$ 时刻的随机变量概率密度函数：

$$p(x_{n+1}) = \sum_{x_n} p(x_n) P_{x_n x_{n+1}}$$

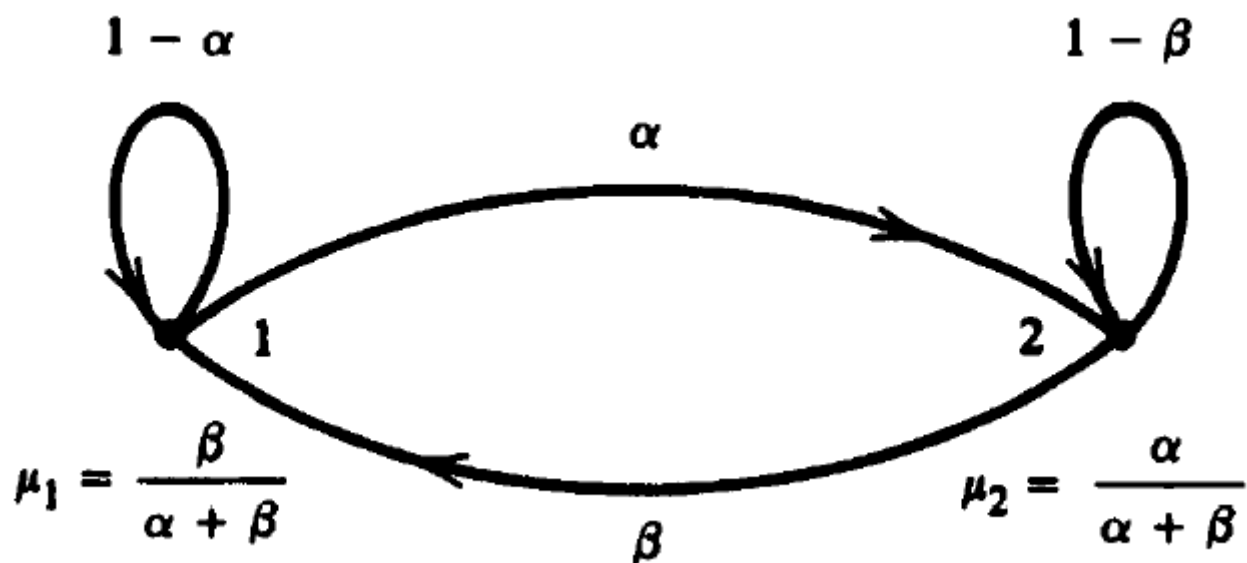
平稳分布

- 若马尔可夫链可以从任意状态经过有限步转移到另一任意状态，且转移概率为正，则称此马尔可夫链是**不可约的**。
- 如果从一个状态转移到它自身的不同路径长度的最大公因子为1，则称该马尔可夫链**是非周期的**。
- 若在 $n+1$ 时刻状态空间的分布与 n 时刻的分布相同，则称此分布为**平稳分布**。
- 若马尔可夫链的初始状态服从平稳分布，则该马尔可夫链为**平稳过程**。
- 若有限状态马尔可夫链是不可约的和非周期的，则它的平稳分布惟一，从任意的初始分布出发，当 n 趋向于无穷时， X_n 的分布必趋向于此平稳分布。

马尔可夫链的例子

例

$$P = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}$$



熵率

- 当如下极限存在时，随即过程 $\{X_i\}$ 的熵率定义为：

$$H(\mathcal{X}) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n)$$

- 打字机：输出 m 个等可能的字母 $H(\mathcal{X}) = \log m$

- i.i.d.随机变量序列 $X_1, X_2, \dots$

$$H(\mathcal{X}) = \lim_{n \rightarrow \infty} \frac{H(X_1, X_2, \dots, X_n)}{n} = \lim_{n \rightarrow \infty} \frac{nH(X_1)}{n} = H(X_1)$$

- 独立但非同分布的随机变量序列

$$H(X_1, X_2, \dots, X_n) = \sum_{i=1}^n H(X_i)$$

熵率的一个相关量

- 当如下极限存在时，定义：

$$H'(\mathcal{X}) = \lim_{n \rightarrow \infty} H(X_n | X_{n-1}, X_{n-2}, \dots, X_1) :$$

- **定理** 对于平稳随机过程， $H(\mathcal{X})$ and $H'(\mathcal{X})$ 均存在且相等： $H(\mathcal{X}) = H'(\mathcal{X})$
- **定理** 对于平稳随机过程， $H(X_n | X_{n-1}, \dots, X_1)$ 随 n 递减且存在极限。
- **定理** 若 $a_n \rightarrow a$ 且 $b_n = \frac{1}{n} \sum_{i=1}^n a_i$ ，则 $b_n \rightarrow a$ 。

马尔可夫链的熵率

➤ 对于平稳的马尔可夫链：

$$H(\mathcal{X}) = H'(\mathcal{X}) = \lim H(X_n | X_{n-1}, \dots, X_1) = \lim H(X_n | X_{n-1}) = H(X_2 | X_1)$$

➤ **定理** 设 $\{X_i\}$ 为平稳马尔科夫链，其平稳分布为 $\mathbf{u}$ ，转移矩阵为 $\mathbf{P}$ ，则熵率为：

$$H(\mathcal{X}) = - \sum \mu_i P_{ij} \log P_{ij}$$

例

$$\mathbf{P} = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}$$
$$H(\mathcal{X}) = H(X_2 | X_1) = \frac{\beta}{\alpha + \beta} H(\alpha) + \frac{\alpha}{\alpha + \beta} H(\beta)$$

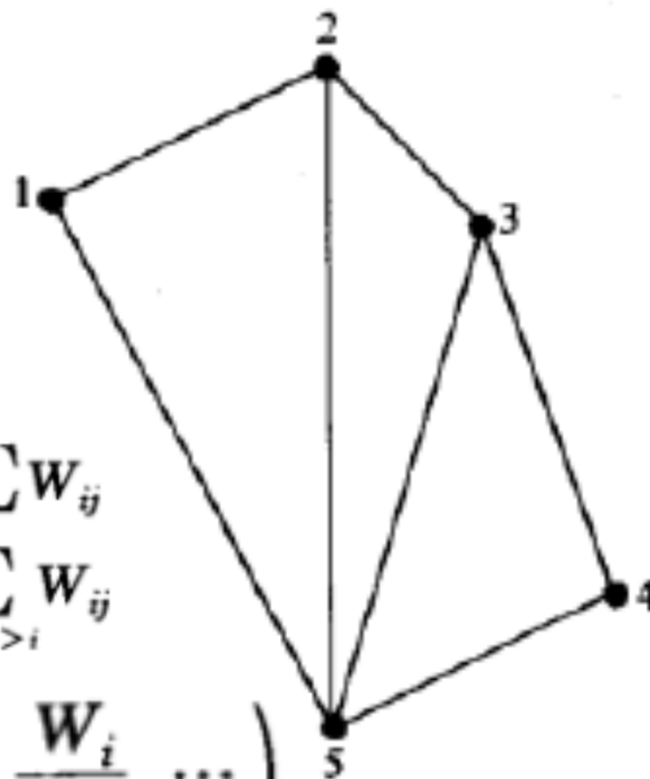
加权图上随机游动的熵率

- 节点: $\{1, 2, \dots, m\}$
- 边权重: $W_{ij} \geq 0, W_{ij} = W_{ji}$
- 若两节点之间没有边, $W_{ij} = 0$
- 随机移动 $\{X_n\}, X_n \in \{1, 2, \dots, m\}$
- 转移概率: $P_{ij} = W_{ij} / \sum_k W_{ik}$
- 平稳分布:

$$\mu_i = \frac{W_i}{2W} \quad \begin{aligned} W_i &= \sum_j W_{ij} \\ W &= \sum_{i,j: j>i} W_{ij} \end{aligned}$$

- 熵率:

$$H\left(\dots, \frac{W_{ij}}{2W}, \dots\right) - H\left(\dots, \frac{W_i}{2W}, \dots\right)$$



国际象棋棋盘上的随机游动

- 若所有的边具有相同的权重
 - 平稳分布: $E_i/2E$;
 - 熵率: $H(\mathcal{X}) = \log(2E) - H\left(\frac{E_1}{2E}, \frac{E_2}{2E}, \dots, \frac{E_m}{2E}\right)$
- 棋盘上的随机移动
 - 王: 三个可能的状态, 平稳分布分别为 $\frac{8}{420}$, $\frac{5}{420}$ 和 $\frac{3}{420}$
熵率: $0.92\log 8$
 - 车: 熵率 $\log 14$
 - 相: ...
 - 王后: ...

热力学第二定律

- 热力学第二定律：孤立系统的熵总是不减的。
- 统计热力学中，熵通常定义为物理系统的微观状态数的对数值。若所有状态都是等概率发生，则和信息论的熵的概念一致。
- 把孤立系统看作一个马尔可夫链
 - 相对熵 $D(u_n || u'_n)$ 随 n 递减；
 - 在 n 时刻状态空间上的分布 u_n 与平稳分布 u 之间的相对熵 $D(u_n || u)$ 随 n 递减；
 - 若平稳分布是均匀分布，则熵增加；
 - 对于平稳的马尔科夫过程，条件熵 $H(X_n | X_1)$ 随 n 增加；
 - 洗牌使熵增加： $H(TX) \geq H(X)$