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

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

- 平稳的随机过程：

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

马尔可夫过程

- 马尔可夫过程（马尔可夫链）：

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

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

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

- 时间不变的马尔可夫过程：

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

- 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：

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

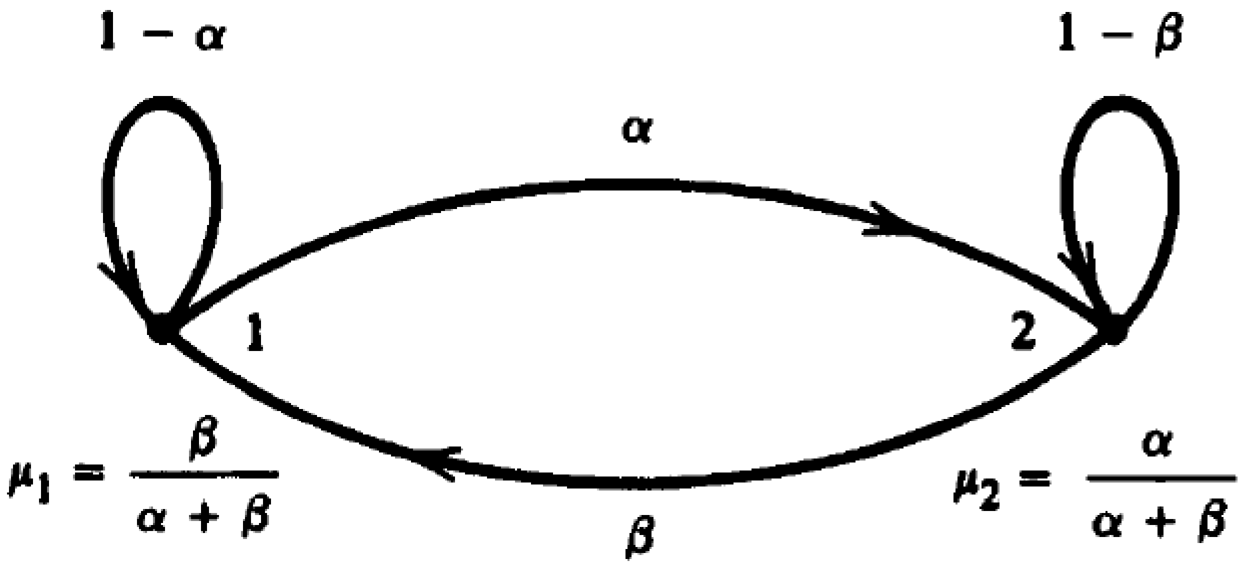
平稳分布

- 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：

$$H(\mathscr{X}) = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n)$$

  - 打字机：输出m个等可能的字母 $H(\mathscr{X}) = \log m$
  - i.i.d.随机变量序列 $X_1, X_2, \ldots$

$$H(\mathscr{X}) = \lim \frac{H(X_1, X_2, \ldots, X_n)}{n} = \lim \frac{nH(X_1)}{n} = H(X_1)$$

  - 独立但非同分布的随机变量序列

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

熵率的一个相关量

- 当如下极限存在时，定义：

$$H'(\mathscr{X}) = \lim_{n\to\infty} H(X_n|X_{n-1}, X_{n-2}, \ldots, X_1)$$

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

马尔可夫链的熵率

- 对于平稳的马尔可夫链：

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

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

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

例

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

$$H(\mathscr{X}) = H(X_2|X_1) = \frac{\beta}{\alpha+\beta} H(\alpha) + \frac{\alpha}{\alpha+\beta} H(\beta)$$

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：$H(TX)\geq H(X)$