第三章　离散信源

中山大学信息科技学院

本章内容

- ▲ 离散信源的分类与数学模型
- ▲ 离散无记忆信源的熵
- ▲ 离散平稳信源的熵
- ▲ 有限状态马尔可夫链
- ▲ 马尔可夫信源
- ▲ 信源的相关性与剩余度

3.1 离散信源的分类与数学模型

- 信源离散信源的分类
- 离散信源的数学模型

3.1.1 离散信源的分类

- ▲ 根据信源符号取值→连续/离散
- ▲ 根据输入符号间的依赖关系→无记忆/有记忆
- ▲ 有限离散信源/无限离散信源
- ▲ 平稳信源/非平稳信源

3.1.2 离散无记忆信源的数学模型

▲ 单符号离散无记忆信源的数学模型：

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & a_n \\ p(a_1) & \cdots & p(a_n) \end{bmatrix}$$

$$p(a_i) \geq 0, \quad \sum_{i=1}^{n} p(a_i) = 1$$

注释
- → $A=\{a_1, \ldots, a_n\}$ →信源的符号集
- → n →符号集的大小
- → a_i →随机变量的取值
- → $p(a_i)$ → $X= a_i$的概率。

单符号离散无记忆信源

例 3.1.1 一个二元无记忆信源，符号集A={0,1}，p为X=0的概率,q为X=1的概率，q=1-p；写出信源的模型。

解：　　信源的模型：

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ p & q \end{bmatrix}$$

3.1.2 离散无记忆信源的数学模型

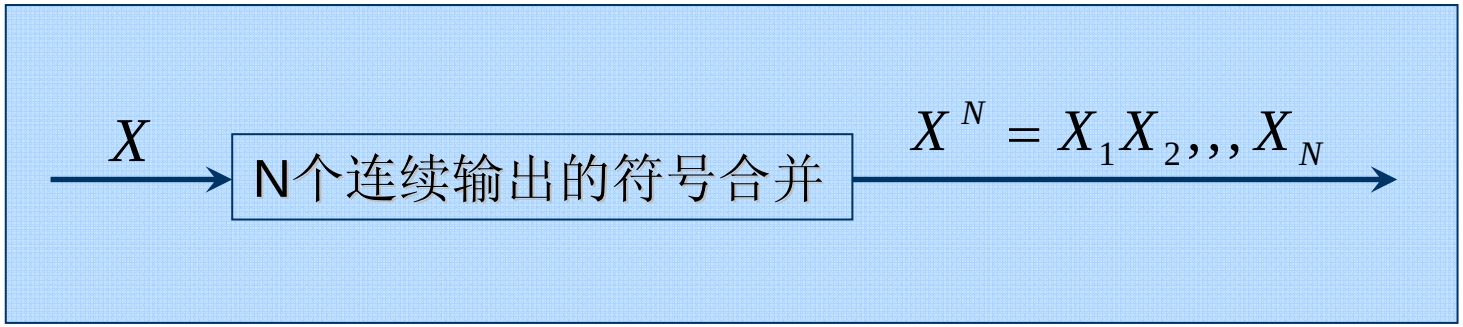
▲多维离散无记忆信源数学模型：

$$\begin{bmatrix} X^N \\ P \end{bmatrix} = \begin{bmatrix} \alpha_1 & \cdots & \alpha_M \\ p(\alpha_1) & \cdots & p(\alpha_M) \end{bmatrix}$$

离散无记忆信源的N次扩展源

▲ 信源X的N次扩展源 ： 设信源为X，由X构成的N维随机矢量集合 $X^N = X_1 X_2,,, X_N, (X_i$与$X$同分布$)$

▲ 信源与其扩展源的关系：



说明:

（1）对于离散无记忆信源 X，其 N 次扩展记为

X^N，$X^N = X_1 X_2 \cdots X_N$

（2）每个 X_i 取自同一个字母表 $A = \{a_1, a_2, \cdots, a_n\}$，且 X_i 与 X 同分布

（3）X^N 的符号集为 $A^N = \{\alpha_1, \alpha_2, \cdots, \alpha_{n^N}\}$，

$\alpha_j = (\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_N})$，$p(\alpha_j) = \prod_{k=1}^{N} p_{j_k}$

离散无记忆信源的N次扩展源

例 3.1.2 求 例3.1.1 中信源的二次扩展模型。

解： ① 二元信源X的符号集为 {0,1}

$\Rightarrow$ 二次扩展源：$X^2 = X_1X_2$，符号集：{00,01,10,11}

$\Rightarrow$ X^2的模型：

$$\begin{bmatrix} X^2 \\ p(\alpha) \end{bmatrix} = \begin{bmatrix} \alpha_1(00) & \alpha_2(01) & \alpha_3(10) & \alpha_4(11) \\ p(\alpha_1) & p(\alpha_2) & p(\alpha_3) & p(\alpha_4) \end{bmatrix}$$

② X^2各符号的概率为：

$$p(\alpha_1) = p^2, p(\alpha_2) = p(1-p) = p(\alpha_3), p(\alpha_4) = (1-p)^2$$

3.1.3 离散有记忆信源的数学模型

离散马尔可夫信源

- 马氏链是随机过程，因此可看成信源，即马尔可夫信源；这种信源是有记忆信源。
- 有限记忆的系统可以用有限状态机来描述。在有限状态机中，既包含状态之间的转移关系，也包含输出与状态之间的关系。
- 可以从有限状态机的概念出发定义马尔可夫信源。

3.2 离散无记忆信源的熵

▲单符号离散无记忆信源的熵

▲离散无记忆信源N次扩展源的熵

3.2.1 单符号离散无记忆信源的熵

例 **3.2.1** 写出例**3.1.1** 中的二元无记忆信源的熵的表达式。

解：

$$H(X) = -p\log p - q\log q$$

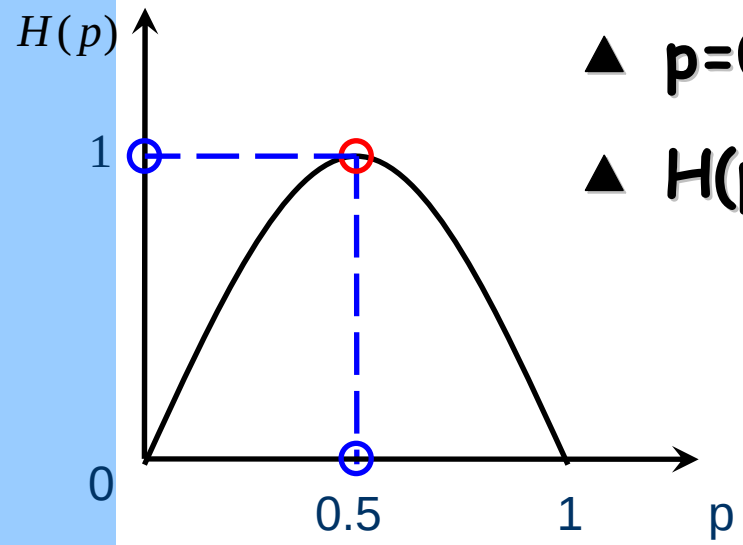
$$= -p\log p - (1-p)\log(1-p)$$

$$= H(p)$$

3.2.1 单符号离散无记忆信源的熵

H(p)的主要性质：

- ▲ 具有熵的一切性质
- ▲ 对p的导函数为 $H'(p)=\log\frac{1-p}{p}$
- ▲ p=0.5时，H(p)达到最大值1bit
- ▲ H(p)是p的上凸函数 $H''(p)=-(\log e)\frac{1}{p(1-p)}\prec 0$



3.2.2 离散无记忆信源N次扩展源的熵

定理3.2.1 离散无记忆信源X的N次扩展源 X^N 的熵等于信源X熵的N倍，即

$$H(X^N) = N\,H(X)$$

证明：

① 无记忆信源

② X_i 互相独立且分布相同

} ⇒ 熵的可加性

$$H(X^N) = \sum_{i=1}^{n} H(X_i) = NH(X)$$

3.2.2 离散无记忆信源N次扩展源的熵

例 **3.2.2** 给定离散无记忆信源模型：$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ 1/2 & 1/4 & 1/4 \end{bmatrix}$

求其二次扩展源熵。

解：

$$H(X^2) = 2H(X) = 2[-\frac{1}{2}\log\frac{1}{2} - (\frac{1}{4}\log\frac{1}{4}) \times 2]$$

$$= 2 \times 1.5 = 3 \quad bit/符号$$

注：这里的符号 - 包含2个信源符号

3.3 离散平稳信源的熵

- 离散平稳信源
- 离散平稳有记忆信源的熵

3.3.1 离散平稳信源(1)

定义:

- ▲ 信源X具有有限符号集 $A=\{a_1,a_2,\cdots,a_n\}$
- ▲ 信源产生随机序列$\{x_i\}$ $i=\cdots,1,2,\cdots$
- ▲ 对所有 $i_1,\cdots,i_N,h,j_1,\cdots,j_N$,及$x_i\in X$ 有

$$p(x_{i_1}=a_{j_1},x_{i_2}=a_{j_2},\cdots;x_{i_N}=a_{j_N})=p(x_{i_1+h}=a_{j_1},x_{i_2+h}=a_{j_2},\cdots;x_{i_N+h}=a_{j_N})$$

则称信源为离散平稳信源，所产生的序列为平稳序列

- 注意2个概念的引入:
  - 离散信源: $\begin{bmatrix} X \\ P(x) \end{bmatrix}$
  - 离散随机序列: $X_1, X_2, \cdots, X_N$ 产生自离散信源，用于研究符号间的依赖（记忆）关系

3.3.1 离散平稳信源(2)

▲ 平稳序列的统计特性与时间的推移无关

$$p(x_{i_1}, x_{i_2}, \cdots, x_{i_N}) = p(x_{i_1+h}, x_{i_2+h}, \cdots, x_{i_N+h})$$

$$p(x_i, x_{i+1} \cdots, x_{i+N}) = p(x_j, x_{j+1} \cdots, x_{j+N})$$

3.3.1 离散平稳信源(3)

例 3.3.1 一平稳信源X的符号集A={0，1}，产生随机序列，其中$P(x_1=0)=p$，求$P(x_n=1)$（n >1）的概率。

解：

平稳性 $\Rightarrow P(x_n=0)=p$

$\Rightarrow P(x_n=1)=1-p$

3.3.1 离散平稳信源(4)

例 3.3.1续 对同一信源，若$P(x_1=0, x_2=1)=b$

求$P(x_4=1/x_3=0)$。

解：平稳性 $\Rightarrow P(x_3=0, x_4=1)=P(x_1=0, x_2=1)=b$

$$\Rightarrow P(x_4=1/x_3=0)=P(x_3=0, x_4=1)/P(x_3=0)=b/p$$

$$P(x_2=1/x_1=0)=P(x_1=0, x_2=1)/P(x_1=0)=b/p$$

3.3.1 离散平稳信源(5)

▲对于平稳信源，条件概率也是平稳的。一般地，有

$$P(x_{i+N} / x_i, x_{i+1}, \cdots, x_{i+N-1}) = P(x_{j+N} / x_j, x_{j+1}, \cdots, x_{j+N-1})$$

▲ 平稳信源的熵与时间起点无关，即

$$H(X_i X_{i+1} \cdots X_{i+N}) = H(X_j X_{j+1} \cdots X_{j+N})$$

$$H(X_{i+N} / X_i, X_{i+1} \cdots, X_{i+N-1}) = H(X_{j+N} / X_j, X_{j+1} \cdots, X_{j+N-1})$$

3.3.2 离散平稳有记忆信源的熵(1)

▲ 根据平稳性和熵的不增原理

信源X输出长度为N序列的平均不确定度

$$H(X^N) \le N H(X_1)$$，仅当无记忆信源时等式成立。

▲对于X的N次扩展源，定义平均符号熵：

可用平均符号熵作为信源X的近似

$$H_N(X) = \frac{1}{N} H(X^N) = \frac{1}{N} H(X_1 \cdots X_N)$$

3.3.2 离散平稳有记忆信源的熵(2)

▲ 信源X的极限符号熵:

$$H_\infty(X\ ) = \lim_{N\to\infty}\frac{1}{N}H(X^N) = \lim_{N\to\infty}\frac{1}{N}H(X_1\cdots X_N)$$

简称: 符号熵/熵率

3.3.2 离散平稳有记忆信源的熵(3)

定理 3.3.1： 任意离散平稳信源，若 $H_1(X)<\infty$

1） $H(X_N / X_1 \cdots X_{N-1})$ 不随N而增加

2） $H_N(X) \geq H(X_N / X_1 \cdots X_{N-1})$

3） $H_N(X)$ 不随N而增加

4） $H_\infty(X)$ 存在，且 $H_\infty(X)=\lim_{N\to\infty} H(X_N / X_1 \cdots X_{N-1})$

说明：有记忆信源的符号熵也可通过计算极限条件熵得到

3.3.2 离散平稳有记忆信源的熵(4)

1）信源的平稳性
　　熵的不增原理 } ⇒ 平稳性假设

$$H(X_N / X_1 \cdots X_{N-1}) = H(X_{N+1} / X_2 \cdots X_N) \geq H(X_{N+1} / X_1 \cdots X_N)$$

这说明对于平稳信源，条件越多，条件熵越不增加

3.3.2 离散平稳有记忆信源的熵(5)

2） 只要证明N个 $H_N(X)$ 的和不小于 $N\ H(X_N / X_1 \cdots X_{N-1})$

$$NH_N(X) = H(X_1 \cdots X_N)$$

条件熵不随N增加

$$= H(X_1) + H(X_2 / X_1) + \cdots + H(X_N / X_1 \cdots X_{N-1})$$

$$\geq N\ H(X_N / X_1 \cdots X_{N-1})$$

$$\Rightarrow \quad H_N(X) \geq H(X_N / X_1 \cdots X_{N-1})$$

平均符号熵不小于条件熵

3.3.2 离散平稳有记忆信源的熵(6)

3） 由于 $NH_N(X) = H(X_1 \cdots X_{N-1}) + H(X_N / X_1 \cdots X_{N-1})$

根据平均符号熵的定义和2）的结果，有

$$N H_N(X) = (N-1)H_{N-1}(X) + H(X_N / X_1 \cdots X_{N-1})$$
$$\leq (N-1)H_{N-1}(X) + H_N(X)$$

移项到不等式左边

$$\Rightarrow \quad H_N(X) \leq H_{N-1}(X)$$

平均符号熵不随序列的长度而增加

3.3.2 离散平稳有记忆信源的熵(7)

4）通过以上证明可得， $0 \le H_N(X) \le H_{N-1}(X) \le \cdots \le H_1(X) < \infty$

$$\therefore 0 \le \lim_{N\to\infty} H_N(X) \le H_1(X) \qquad \text{即} H_\infty(X) \text{存在}$$

计算 $(N+j)H_{(N+j)}(X) = H(X_1 \cdots X_{N-1} X_N \cdots X_{N+j})$

$$= H(X_1 \cdots X_{N-1}) + H(X_N / X_1 \cdots X_{N-1}) + \cdots + H(X_{N+j} / X_1 \cdots X_{N+j-1})$$

(j+1)个

利用1)的结果与平稳性，

$$H(X_{N+j} / X_1 \cdots X_{N+j-1}) \le \cdots \le H(X_N / X_1 \cdots X_{N-1})$$

$$\Rightarrow (N+j)H_{(N+j)}(X) \le H(X_1 \cdots X_{N-1}) + (j+1)H(X_N / X_1 \cdots X_{N-1})$$

$$H_{(N+j)}(X) \le \frac{1}{N+j} H(X_1 \cdots X_{N-1}) + \frac{j+1}{N+j} H(X_N / X_1 \cdots X_{N-1})$$

极限为1

3.3.2 离散平稳有记忆信源的熵(8)

令 $j \to \infty$，不等式右边第1项为0，第2项—》$H(X_N / X_1 \cdots X_{N-1})$

再令 $N \to \infty$，有：$H_\infty(X) \le \lim_{N\to\infty} H(X_N / X_1 \cdots X_{N-1})$

另外，由（2）的结果，$H_\infty(X) \ge \lim_{N\to\infty} H(X_N / X_1 \cdots X_{N-1})$

同时大/小于一个式子,两边夹原则

所以

$$H_\infty(X) = \lim_{N\to\infty} H(X_N / X_1 \cdots X_{N-1})$$

证毕。

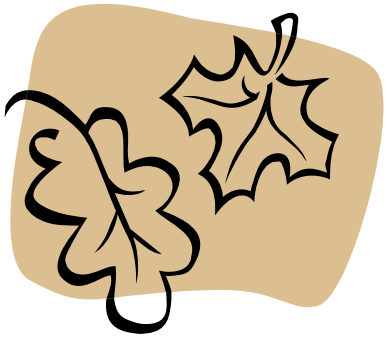
3.3.2 离散平稳有记忆信源的熵(9)

定理3.3.1的注释:

▲ 该定理提供了通过计算极限条件熵得到信源符号熵的方法；当信源为有限记忆时，极限条件熵的计算要比极限平均符号熵的计算容易得多。

—对于记忆长度为 m 的平稳信源

$$H_{\infty}(X)=\lim_{N\to\infty}H(X_N\mid X_{N-1}X_{N-2}\cdots X_1)$$
$$=H\left(X_{m+1}\mid X_mX_{m-1}\ldots X_1\right)$$



▲ 极限熵等于最小的平均符号熵。

3.4 有限状态马尔可夫链

马尔柯夫信源的输出序列

- ▲ 马氏链的基本概念
- ▲ 齐次马氏链
- ▲ 马氏链状态分类
- ▲ 马氏链的平稳分布

3.4.1 马氏链的基本概念(1)

定义:

▲ 随机序列 $\{x_n, n \geq 0\}$

假定一阶马氏序列，符号取值为状态

▲ 每个随机变量$x_n (n \geq 1)$ 仅依赖于 x_{n-1}

$$p\{x_n = j / x_{n-1} = i, x_{n-2} = k, \cdots, x_0 = m\} = p\{x_n = j / x_{n-1} = i\}$$

▲ 随机变量 x_n ： 马氏链在n时刻的状态

▲ $1, \cdots, J$ ： 状态

▲ $1, \cdots, J$ 的集合S: 状态集合

3.4.1 马氏链的基本概念(2)

1) 一阶马氏链的当前状态只与前一个状态有关

2) n阶马氏链的当前状态只与前n个状态有关

3) 马氏链是时间离散，状态也离散的随机过程

4) 状态集合为有限集→有限状态马氏链

状态集合为无限集→无穷状态马氏链

3.4.1 马氏链的基本概念(3)

▲ 描述马氏链的最重要的参数: 状态转移概率 **刻画马氏链的模型**

▲ 对于离散时刻m、n，相应的状态转移概率可表示为:

$$p(x_n = j / x_m = i) = p_{ij}(m,n)$$

表示从时刻m的i状态转移到时刻n的j状态的概率，m-n表示转移的步数。 $p_{ij}(m,n)$ 是经 $n-m(n>m)$ 步转移的概率。

讨论:有(m-n)步转移概率不表明n时的状态依赖m.因可能m取不同状态时的转移概率都相同

3.4.1 马氏链的基本概念(4)

转移概率的主要性质：

- $0 \le p_{ij}(m,n) \le 1, \quad i, j \in S;$
- $\sum_j p_{ij}(m,n) = 1;$
- 一步转移概率 $p_{ij}(m,m+1) = p(x_{m+1} = j / x_m = i) = p_{ij}(m)$

  其中，m 为起始时刻，$i, j \in S$
- K步转移概率 $p_{ij}^{(k)}(m) = p(x_{m+k} = j / x_m = i)$
- 0步转移概率 $p_{ij}^{(0)} = \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases}$

  系统在任何时刻只能处于S中一个状态

3.4.2 齐次马氏链 (1)

▲ 转移概率与起始时刻无关，具有平稳性

$$p_{ij}(m) = p(x_{m+1} = j / x_m = i) = p_{ij}$$

▲ $p_{ij} \geq 0, \sum_j p_{ij} = 1$

▲ K步转移概率也与起始时刻无关，写成 $p_{ij}^{(k)}$

如果没有上标，
约定为1步转移

3.4.2 齐次马氏链 (2)

齐次马氏链的表示方法

转移概率矩阵

$$[P]=[p_{ij}]=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1J} \\ p_{21} & p_{22} & \cdots & p_{2J} \\ \cdots & \cdots & \cdots & \cdots \\ p_{J1} & p_{J2} & \cdots & p_{JJ} \end{bmatrix}$$

=1

由状态 j 转移到状态 2 的概率；非负

网格图

每时刻的网格节点与马氏链的状态一一对应

状态转移图

状态转移图与矩阵有一一对应关系

具有时间性

3.4.2 齐次马氏链 (3)

例 **3.4.1** 一个矩阵，验证此矩阵对应一个齐次马氏链的转移概率矩阵并确定此马氏链的状态数

解：

① 元素非负

每行和为1

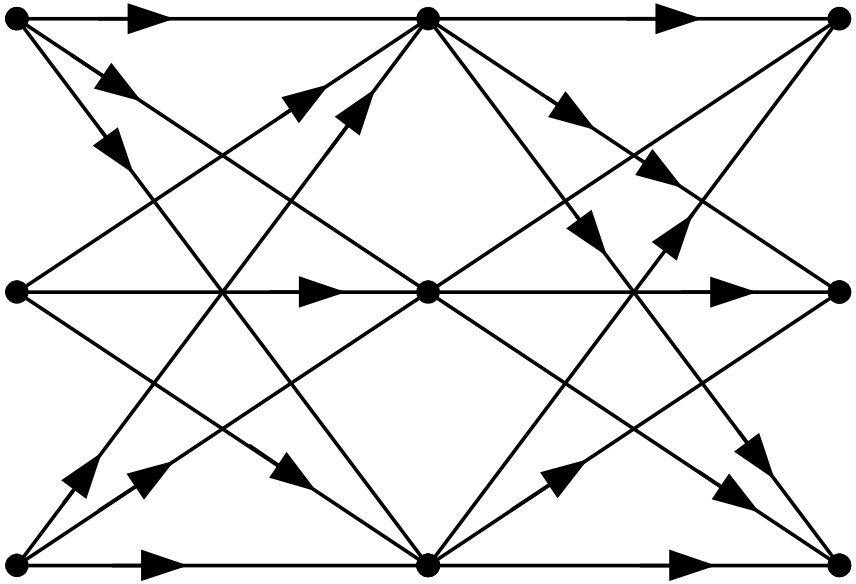
$$[p]=\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}\begin{matrix} =1 \\ =1 \\ =1 \end{matrix}$$

② 状态数 =3

3.4.2 齐次马氏链 (4)

例 3.4.1续　画出此马氏链的网格图。

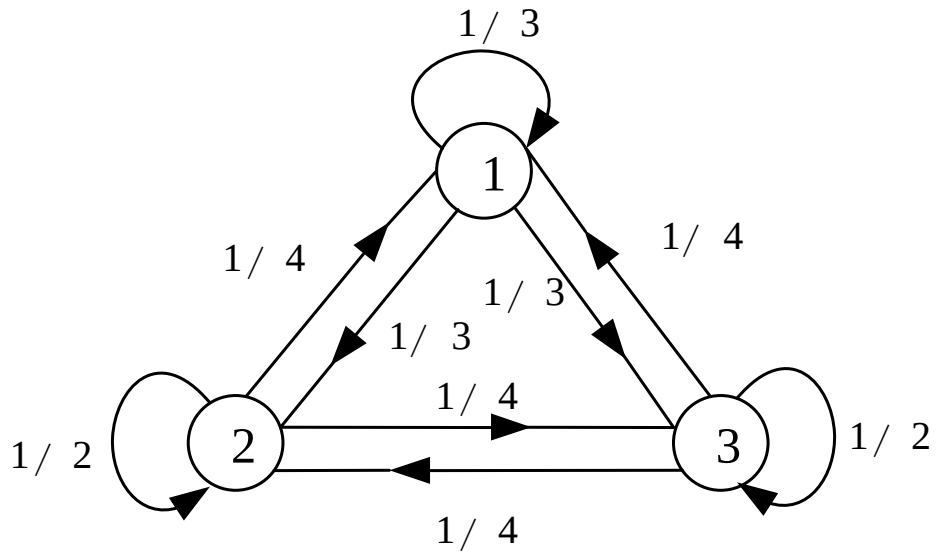
解:



3.4.2 齐次马氏链 (5)

例 3.7.1续　画出此马氏链的状态转移图

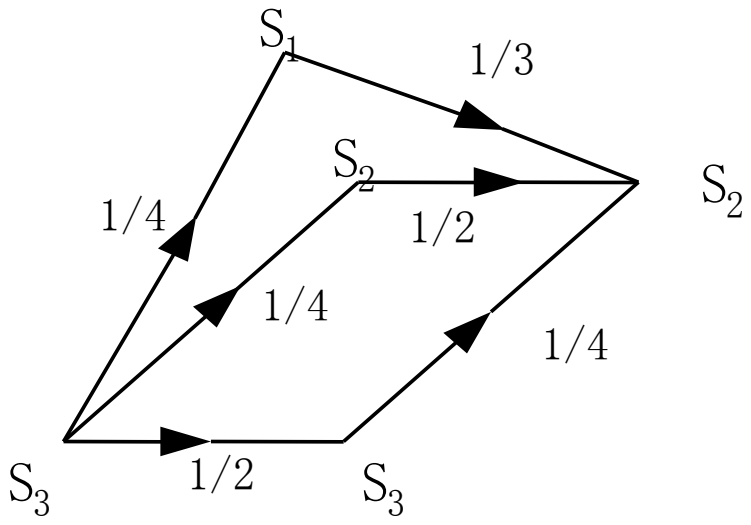
解:



3.4.2 齐次马氏链 (6)

例 3.7.1续　求从状态3到状态2的2步转移概率

解:



3.4.2 齐次马氏链 (7)

下面分两步来计算：

1) 计算从m时刻从s_3经m+1时刻某状态s_k到m+2时刻s_2的转移概率

2) 对1）中计算的经m+1时刻的所有状态s_k（k=1，2，3）概率相加，得到所求结果。计算得

$$p\{x_{m+2}=s_2 / x_m=s_3\}=\frac{1}{4}\times\frac{1}{3}+\frac{1}{4}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{4}=\frac{1}{3}$$

Kolmogorov-Chapman方程(1)

由此例可以看出：

1) 计算从状态i到状态j的2步转移概率可通过下式：

$$p_{ij}^{(2)} = \sum_k p_{ik} p_{kj}$$

2) $p_{ij}^{(2)}$ 是 $[P]^2$ 的第(i,j)个元素

3) $p_{ij}^{(m)}$ 是 $[P]$ 的m次幂 $[P]^m$ 的第(i,j)个元素

4) $[P]^{m+n} = [P]^m [P]^n \Rightarrow p_{ij}^{(m+n)} = \sum_k p_{ik}^{(m)} p_{kj}^{(n)}$

Kolmogorov-Chapman方程(2)

5) 设马氏链的初始状态概率分布为 $[p^{(0)}]=[p_1^{(0)}, p_2^{(0)}, \cdots, p_J^{(0)}]$

其中J为状态数，经k步转移后的状态概率分布为

$[p^{(k)}]=[p_1^{(k)}, p_2^{(k)}, \cdots, p_J^{(k)}]$，则有：

注意分布p，状态转移P

$$[p^{(k)}]=[p^{(0)}][P]^k=[p^{(m)}][P]^{k-m}$$

因此，一个齐次马氏链，当初始状态概率分布给定后，可计算转移后任何时刻的状态概率分布。

Kolmogorov-Chapman方程(2)

例 **3.4.2** 设例3.4.1中马氏链的初始状态的概率分布为1/2、1/4、1/4，分别求1步转移后和2步转移后的状态的概率分布。

解:

$$\boldsymbol{p}^{(1)} = \boldsymbol{p}^{(0)}\boldsymbol{P} = \begin{pmatrix}1/2 & 1/4 & 1/4\end{pmatrix}\begin{pmatrix}1/3 & 1/3 & 1/3\\ 1/4 & 1/2 & 1/4\\ 1/4 & 1/4 & 1/2\end{pmatrix} = \begin{pmatrix}7/24 & 17/48 & 17/48\end{pmatrix}$$

$$\boldsymbol{p}^{(2)} = \boldsymbol{p}^{(0)}\boldsymbol{P}^2 = \begin{pmatrix}1/2 & 1/4 & 1/4\end{pmatrix}\begin{pmatrix}5/18 & 13/36 & 13/36\\ 13/48 & 19/48 & 1/3\\ 13/48 & 1/3 & 19/48\end{pmatrix}$$

$$= \begin{pmatrix}79/288 & 209/576 & 209/576\end{pmatrix}$$

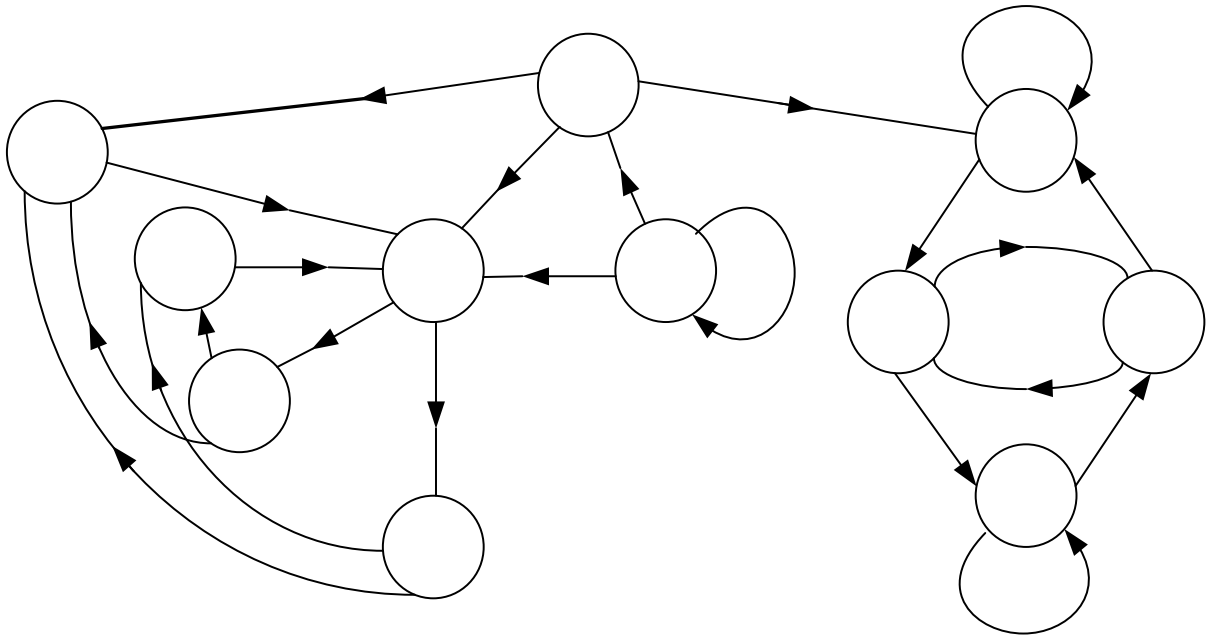
3.4.3 马氏链状态分类 (1)

- 若对某一$n \geqslant 1$，有 $p_{ij}^{(n)} > 0$ ，则称状态j可由状态i到达，记为 $i \to j$
- 如果 $i \to j$ ，且 $j \to i$ ，则称状态i与状态j可互通，记为 $i \leftrightarrow j$
- 定义每个状态都与该状态本身互通，即互通关系满足自反性
- 互通关系还满足对称性和传递性

3.4.3 马氏链状态分类 (2)

例 3.4.3　按互通性将状态分成若干类（子集）

解:



3.4.3 马氏链状态分类 (3)

$C_1 = \{s_1\}$

$C_2 = \{s_2\}$

过渡类

过渡态

$C_3 = \{s_3, s_4, s_5, s_6, s_7\}$

$C_4 = \{s_8, s_9, s_{10}, s_{11}\}$

常返类

常返态

$\{n : n \geq 1, p_{ii}^{(n)} > 0\}$ 最大公约数 d_i

所有n的最大公约数d_i

d_i >1 周期的

d_i =1 非周期

有关马氏链状态分类

常返态：有限状态马氏链中，经过有限步迟早要返回的状态（返回概率大于 **0**）

过渡态：经过该状态可以到达某一其他状态，但不能从其他状态返回

指子集中状态不能减少

马氏链按互通关系分成的不可约子集中的状态或者是常返的，称为常返类；或者是过渡的，称为过渡类。

注：周期或遍历是针对某一个类

对于常返类，$d_i > 1$为周期的，表明从该类任意状态出发，至少经过kd_i步才能返回该状态；$d_i = 1$为非周期或遍历的，表明从该类任意状态出发，在有限步后可以转移到同类中任何其他状态。

3.4.3 马氏链状态分类 (4)

定理3.4.1:

对任何马氏链（有限或无限可数状态），同一类中所有状态都有相同周期。

因取所有状态的公约数

非周期的常返态称为遍历态

3.4.4 马氏链的平稳分布 (1)

▲ 定义：若对任意整数m,n,马氏链的状态分布满足

$$P(x_m = i) = P(x_n = i) = \pi_i$$

则称 $\{\pi_i\}$为平稳分布，或稳态分布

▲ $$\pi_j = \sum_i \pi_i p_{ij} \quad (j = 1,2,\cdots,J)$$

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{P} \qquad (3.4.11)$$

其中， $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_J)^T$ 为平稳分布行矢量

3.4.4 马氏链的平稳分布 (2)

▲ 平稳马氏链

$$[p^{(0)}]=[p_1^{(0)}, p_2^{(0)},\cdots, p_J^{(0)}]=[\pi] \qquad [p^{(n)}]=[\pi]$$

从任一状态出发总有可能到其他任意状态

▲ 定理3.4.2 如果一个遍历有限状态马氏链的转移概率矩阵为[P]，那么

$$\lim_{k\to\infty}[P]^k=[e][\pi] \qquad (3.4.12)$$

其中，$[e]=[1,1,\cdots,1]^t$，为J维列矢量。$[\pi]$为平稳分布行矢量

$\Rightarrow$ $\lim_{k\to\infty}[P]^k$是一个每行都相同（都等于$[\pi]$）的矩阵

3.4.4 马氏链的平稳分布 (3)

- 对于遍历马氏链，无论初始分布如何，当转移步数足够大时，状态概率分布总是趋于平稳值，与初始状态概率分布无关。

对于转移概率P，满足 $\boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{P}$ 的分布

$[p^{(n)}]=[p_1^{(n)}, p_2^{(n)}, \cdots, p_J^{(n)}]$ $[p^{(0)}]$ $[\pi]$

$$\lim_{k\to\infty}[p^{(k)}]=\lim_{k\to\infty}\{[p^{(0)}][P]^k\}=[p^{(0)}][e][\pi]=[\pi_1\ \pi_2\ \cdots \pi_J]$$

$=1$

3.4.4 马氏链的平稳分布 (4)

- 对于有限状态马氏链，稳态分布恒存在

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T \mathbf{P}$$

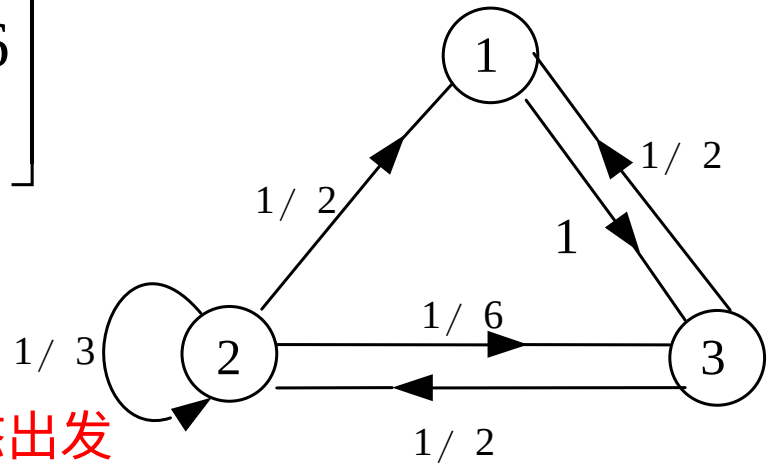
- 如果马氏链中仅存在一个常返类，则方程(3.4.11)的解是唯一的；如果存在r个常返类，则具有r个线性独立的矢量解
- 如果马氏链中仅存在一个常返类而且是非周期的（即遍历的），则(3.4.12)式成立；如果有多个常返类，但都是非周期的，则 $[P]^n$ 也收敛，但矩阵的每行可能不同；如果马氏链具有一个或多个周期常返类，则 $[P]^n$ 不收敛

虽然状态转移矩阵不收敛，但仍然存在稳态分布

3.4.4 马氏链的平稳分布 (5)

例 **3.4.4** 一马氏链的转移概率矩阵如下，问此马氏链是否具有遍历性并求平稳分布和 $\lim_{k\to\infty}[P]^k$ 的值

$$[P]=\begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/3 & 1/6 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$



解:

注：由右边状态图，从任一状态出发经有限步可以到达任意其他状态，说明是遍历的。

3.4.4 马氏链的平稳分布 (6)

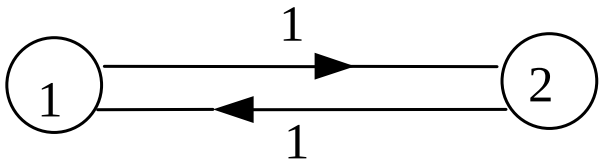
$$\begin{cases} [\pi_1\ \pi_2\ \pi_3]\begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/3 & 1/6 \\ 1/2 & 1/2 & 0 \end{bmatrix} = [\pi_1\ \pi_2\ \pi_3]) \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 2/7 \\ 8/21 \end{bmatrix} \qquad \lim_{k\to\infty}[P]^k = \begin{bmatrix} 1/3 & 2/7 & 8/21 \\ 1/3 & 2/7 & 8/21 \\ 1/3 & 2/7 & 8/21 \end{bmatrix}$$

3.4.4 马氏链的平稳分布 (7)

例 3.4.5 一马氏链的状态转移矩阵如下，确定它的n次幂是否收敛并求其平稳分布 $[P]=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$

解：



$\Rightarrow$ 马氏链包含一个周期常返类，周期 $=2$

$$\Rightarrow \begin{cases}(1)[P]^n \text{不收敛} & [P]^{2k} \quad [P]^{2k+1}\\ (2)\text{有唯一平稳分布} & \Rightarrow[\pi_1\ \pi_2]\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}=[\pi_1\ \pi_2]\\ & \pi_1+\pi_2=1\end{cases} \Rightarrow \pi_1=\pi_2=\frac{1}{2}$$

3.5 马尔可夫信源

- 马氏源的基本概念
- 马氏源的产生模型
- 马氏源N次扩展源熵的计算
- 马氏源符号熵的计算

3.5.1马氏源的基本概念(1)

马氏源的定义:

▲ 当前时刻输出符号的概率仅与当前时刻的信源状态有关

$$p\{x_l = a_k / s_l = j\} = p\{x_l = a_k / s_l = j, x_{l-1}, s_{l-1}, \cdots\}$$

▲ 当前时刻的信源状态由前一时刻信源状态和前一时刻输出符号唯一确定。

$$p\{s_l = i / x_{l-1} = a_k, s_{l-1} = j\} = \begin{cases} 1 \\ 0 \end{cases}$$

3.5.1马氏源的基本概念(2)

例 3.5.1 给定二阶马氏源符号集A={0，1}，转移概率分别为：p(0/00)=p(1/11)=0.8，

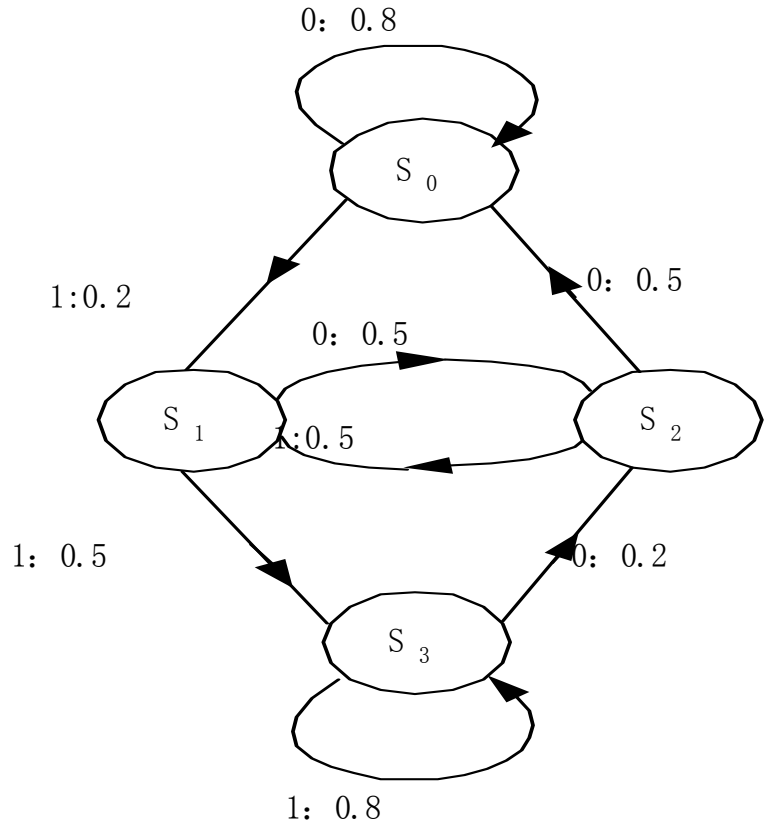
p(1/00)=P(0/11)=0.2，

p(0/01)=p(0/10)=p(1/01)=p(1/10)=0.5，确定该马氏源的状态，写出状态转移矩阵，画出信源的状态转移图。

3.5.1马氏源的基本概念(3)

解:

$$[P]=\begin{bmatrix} 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \end{bmatrix}$$



1. 状态确定

二阶马氏链符号序列—》$\cdots x_1 x_2 x_3 \cdots$—》两两合并形成状态序列$\{\omega_0 = 00, \omega_1 = 01, \omega_2 = 10, \omega_3 = 11\}$

2. 顺序： 前—后 + 输出

例如： $p(0|01)$ —》“ 01〈— 0” —》10

—》状态$\omega_1 \to \omega_2$ 条件概率描述了状态1—》状态2的转移概率

3. 当前状态决定当前符号输出 **p(0|01) 和 p(1|01)**

4. 当前状态由上一时刻状态和输出符号决定

M阶马氏链转为马氏源

m阶马氏链的处理方法

1) m阶马氏链的符号转移概率已给定 $p(x_{m+1} / x_1 \cdots x_m)$

其中x_i取自$A = \{a_1 \cdots a_n\}$

2) 做m长符号序列到信源状态的映射$(x_1 \cdots x_m) \to s_j$，x_i 取遍 $A = \{a_1 \cdots a_n\}$，i=1,…,m; s_j 状态取自 $A^m = \{\omega_1, \omega_2, \cdots \omega_{n^m}\}$，$n^m$为状态数;

马氏链

马氏源

3) 符号转移概率转换成状态转移概率 $p(x_{m+1} / x_1 \cdots x_m) = p(s_{j+1} / s_j)$

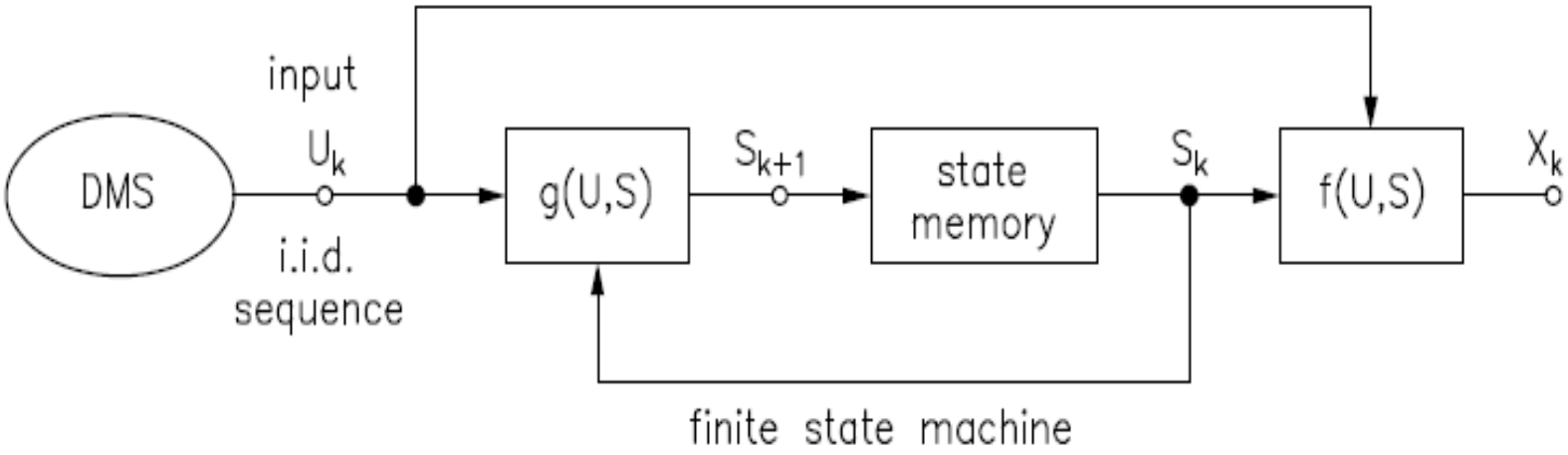
其中 $(x_2 \cdots x_{m+1}) \to s_{j+1}$ $(x_1 \cdots x_m) \to s_j$

马氏源刻画

4) 得到一阶马氏源模型:

$$\begin{bmatrix} \omega_1 & \omega_2 & \cdots & \omega_{n^m} \\ p(\omega_l / \omega_k) & k,l = 1,\ldots,n^m \end{bmatrix}$$

3.5.2 马氏源的产生模型(1)



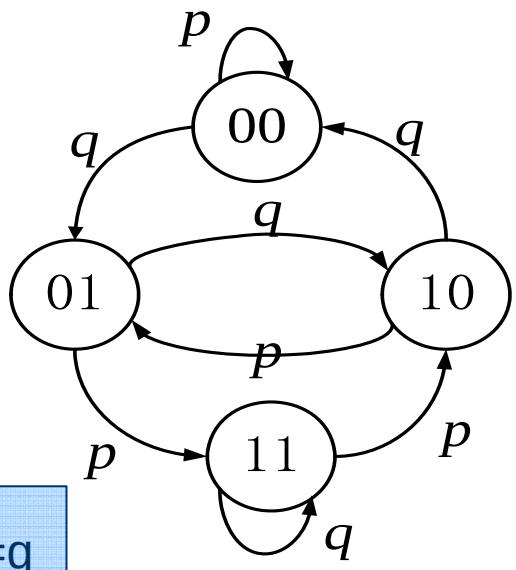
3.5.2 马氏源的产生模型(2)

例 3.5.2

设独立随机序列 $\{x_n\}$，$p(x_n=0)=p$，$p(x_n=1)=q$，$p+q=1$，随机序列 $\{y_n\}$ 与 $\{x_n\}$ 的关系为 $y_n=x_n\oplus y_{n-1}\oplus y_{n-2}$ 其中 $\oplus$ 为模2加；问：（1）随机序列 $\{y_n\}$ 是否为马氏链？（2）如果是马氏链，那么求状态转移概率并画状态转移概率图。

3.5.2 马氏源的产生模型(3)

解：序列 $\{y_n\}$ 为有记忆序列，在n时刻的取值仅与n-1时刻与n-2时刻有关，而与以前的时间无关，因此 $\{y_n\}$ 构成二阶马氏链。



说明：

1. y_{n-1}, y_{n-2}（状态）+ x_n（符号）—》 y_n, y_{n-1}（下一状态）

2. $p(y_n = 0 \mid y_{n-1} = 0, y_{n-2} = 1)$ —》 $y_n = 0, y_{n-1} = 0$ —》 $s_1 \xrightarrow{q} s_0$

P(x_n=1)=q

3.5.3 马氏链N次扩展源的熵的计算(1)

例 3.5.3 有一个二元马氏链X，符号集为{0，1}，其中符号转移概率为 $p(0/0)=0.8$ ，$p(1/1)=0.7$ ；计算该信源三次扩展源的所有符号的概率。

解： 首先求平稳分布

$$\begin{cases} (p_0 \quad p_1)=(p_0 \quad p_1)\begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} \\ p_0+p_1=1 \end{cases} \Longrightarrow \quad p_0=3/5,\ p_1=2/5$$

3.5.3 马氏链N次扩展源的熵的计算(2)

$$p(000) = p_0\,p(0/0)\,p(0/0) = 0.6\times 0.8\times 0.8 = 0.384$$

类似得到

$$p(001) = 0.6\times 0.8\times 0.2 = 0.096$$

$$p(010) = 0.6\times 0.2\times 0.3 = 0.036$$

$$p(011) = 0.6\times 0.2\times 0.7 = 0.084$$

$$p(100) = 0.4\times 0.3\times 0.8 = 0.096$$

$$p(101) = 0.4\times 0.3\times 0.2 = 0.024$$

$$p(110) = 0.4\times 0.7\times 0.3 = 0.084$$

$$p(111) = 0.4\times 0.7\times 0.7 = 0.196$$

3.5.3 马氏链N次扩展源的熵的计算(3)

m阶马氏链，i+m+1时符号的概率取决于前m个符号i+1到i+m，状态有n^m个

做映射 $(x_{1+i}\cdots x_{m+i})\to s_{m+1+i}(j)$ ，i = 0,…,N-m，其中i为时间标号，j为状态序号。

$j=n^m$

S(m+1): Xm — X1;
S(N+1):XN — X(N+1-m)

$$H(X_1X_2\ldots X_N)=H(S_{m+1}S_{m+2}\ldots S_{N+1})$$

符号序列涉及1-N中的符号；状态s(m+1) –X1;状态s(N+1) — XN；二序列涉及符号相同，故熵相同

所有状态涉及的符号在X1-XN间

其中，$S_i=X_{i-m}X_{i-m+1}\ldots X_{i-1}$

利用熵的可加性，将上式展开，并利用马氏性得

$$H(X_1X_2\ldots X_N)=H(S_{m+1})+H(S_{m+2}/S_{m+1})+\ldots+H(S_{N+1}/S_{m+1}S_{m+2}\ldots S_N)$$

$$=H(S_{m+1})+H(S_{m+2}/S_{m+1})+\ldots+H(S_{N+1}/S_N)$$

$$=H(S_{m+1})+\sum_{i=m+1}^{N}H(S_{i+1}/S_i)$$

仅依赖最近状态

3.5.3 马氏链N次扩展源的熵的计算(4)

$$\sum_{i=m+1}^{N} H(S_{i+1}/S_i) = -\sum_{i=m+1}^{N}\sum_{j=1}^{n^m} p[s_i(j)]\sum_{k=1}^{n^m} p[(s_{i+1}(k)/s_i(j)]\log p[(s_{i+1}(k)/s_i(j)]$$

$$= \sum_{i=m+1}^{N}\sum_{j=1}^{n^m} p[s_i(j)]\, h_j$$

条件熵定义；而可能的状态数为 n^m 个

$$h_j = -\sum_{k=1}^{n^m} p[(s_{i+1}(k)/s_i(j)]\log p[(s_{i+1}(k)/s_i(j)]$$

← 由状态转移阵的第j行确定

3.5.3 马氏链N次扩展源的熵的计算(5)

h_j 由状态转移概率矩阵[P]的第j行所确定。写成矩阵形式

$$\sum_{i=m+1}^{N} H(S_{i+1}/S_i) = \sum_{i=m+1}^{N} [p(s_i)][h]^t$$

每1项为时刻 i 的状态分布

其中，$[p(S_i)] = [p(s_i(1))\ p(s_i(2)) \cdots p(s_i(n^m))]$ ，为第i状态概率分布行矢量；$[h] = [h_1 h_2 \cdots h_{n^m}]$ ，为行矢量，其中每个元素由[P]的每一行所确定。

考虑从时刻（m+1）开设，转移（N-m）步

$$H(X_1 X_2 \cdots X_N) = H(S_{m+1}) + [p(s_{m+1})] \sum_{i=0}^{N-m-1} [P]^i [h]^t$$

3.5.3 马氏链N次扩展源的熵的计算(6)

▲ 如果起始状态概率为平稳分布，则

注意书中为列向量，这里为行向量

$$H(X_1X_2\cdots X_N)=H([\pi])+(N-m)[\pi][h]^t$$

▲ N次扩展源的平均符号熵为：

$$H_N(X)=\frac{1}{N}H(X_1X_2\cdots X_N)=\frac{1}{N}\{H([\pi])+(N-m)[\pi][h]^t\}$$

这里$h=[h_j]$，h_j 为由P的第 j 行所确定的条件熵

例3.5.3（续1）

问题：对于此二阶马氏链，求信源 8 次扩展的熵，相当于对该马氏源，求$H(X_1X_2\cdots X_8)$的熵

解：（1）二阶马氏源—》4 个状态—》转移概率阵

$$[P]=\begin{bmatrix} 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \end{bmatrix}$$

计算$[\mathbf{h}]=[h_1 \quad h_2 \quad h_3 \quad h_4]$

例如：$h_1=h_4=-0.8\times\log 0.8-0.2\times\log 0.2=0.722$

例3.5.3（续2）

（2）确定稳态分布

$$\begin{pmatrix}\pi_1 & \pi_2 & \pi_3 & \pi_4\end{pmatrix}P=\begin{pmatrix}\pi_1 & \pi_2 & \pi_3 & \pi_4\end{pmatrix}$$

$$\pi_1+\pi_2+\pi_3+\pi_4=1$$

故有：$[\pi_1\pi_2\pi_3\pi_4]=[5/14 \quad 1/7 \quad 1/7 \quad 5/14]$

$$H([\pi])=1.863bit$$

（3）求 8 次扩展源的熵

$$H(X_1X_2\cdots X_8)=H([\pi])+(8-2)[\pi]^{\tau}[\mathbf{h}]=6.669/\text{符号}$$

注意：这里符号长为 8

例3.5.3（续3）

讨论：和相同无记忆信源的比较

（1）原 2 元信源的平稳分布为 $p=[3/5 \quad 2/5]$

（2）计算其熵 $H(p)=0.971bit/$符号

（3）假定无记忆 8 次扩展

$$H_1(X_1X_2\cdots X_8)=8\times H_1(X_1)=7.768/\text{符号}$$

结论：利用记忆信源符号间的相关性可以获得高的压缩效率

3.5.4 马氏源符号熵的计算(1)

计算方法1：　直接由前H_N表达式，取极限

▲ 当信源从某一状态转移到另一状态时，输出符号唯一，则一个m阶马氏源的符号熵为：

$$H_{\infty}(X)=\lim_{N\to\infty}H_N(X)=[\pi][h]^t$$

▲ m阶马氏源符号熵仅由平稳分布和状态转移概率矩阵所决定。

3.5.4 马氏源符号熵的计算(2)

计算方法2:

▲ 当信源从某一状态转移到另一个新状态时，如果存在多个信源输出符号对应一个到达状态，这时由状态转移概率矩阵不能确定信源的熵，而只能以状态条件下信源的输出符号的概率求信源的熵。

从另一个视角推导

▲ 给定当前信源状态条件下信源的输出符号熵为:

$$H(X/s=j)=-\sum_{i=1}^{n}p_j(a_i)\log p_j(a_i)$$

其中，$a_i \in A=(a_1,\cdots,a_n)$，$A$ 为信源符号集

3.5.4 马氏源符号熵的计算(3)

引理3.5.1:

▲ 在给定某特殊状态$s_1=j$和以前的输出$X_1,X_2,\ldots X_{m-1}$条件下当前输出符号X_m的熵满足:

$$H(X_m / s_1 = j, X_1 \cdots X_{m-1}) = \sum_{i=1}^{J} p(s_m = i / s_1 = j) H(X / s = i)$$

对S_1取平均

$$H(X_m / S_1, X_1 \cdots X_{m-1}) = \sum_{j=1}^{J} \sum_{i=1}^{J} p(s_1 = j)\, p(s_m = i / s_1 = j) H(X / s = i)$$

$$= \sum_{i=1}^{J} p(s_m = i) H(X / s = i)$$

3.5.4 马氏源符号熵的计算(4)

对于平稳信源，状态概率与时间起点无关，所以

$$H(X_m / S_1, X_1 \cdots X_{m-1}) = \sum_{i=1}^{J} p(s = i) H(X / s = i)$$

对于m 阶平稳马氏源的符号熵为:

$$H_\infty = \lim_{N \to \infty} H(X_N / X_1 \cdots X_{N-1})$$

利用平稳性及马氏性，有

$$H_\infty = H_{m+1}$$

$$= H(X_{m+1} / X_1 \cdots X_m)$$

3.5.4 马氏源符号熵的计算(5)

当 $X_1 \cdots X_m$ 给定条件下，X_{m+1} 与以前的状态无关

$$H(X_{m+1}/X_1 \cdots X_m) = H(X_{m+1}/S_1, X_1 \cdots X_m)$$

$$= \sum_{i=1}^{J} p(s=i) H(X/s=i)$$

$$= [\pi][h]^t$$

注：在s=i时，输出J个符号对应J个状态，这个条件熵由P矩阵的第i行确定，即hi

- ▲ $[\pi]$ 为状态平稳分布行矢量
- ▲ h的各分量由状态转移概率矩阵的每一行得出的条件熵

3.5.4 马氏源符号熵的计算(6)

定理3.5.2:

平稳马氏源的符号熵为

$$H_{\infty}(X)=\sum_{i=1}^{J}\pi_i H(X/s=i)$$

3.5.4 马氏源符号熵的计算(7)

几点注释：

1) 定理3.5.2给出了马氏源符号熵的计算方法：

先求每个状态下的条件符号熵，再用状态的概率平均；

2) 计算符号熵要用状态的平稳分布；

3) 单位为比特/符号。

3.5.4 马氏源符号熵的计算(8)

例 3.5.1 续　求信源的极限符号熵

解：$H_\infty(X)=[\pi][h]^t$

$$=\frac{5}{14}[-0.8\log 0.8-0.2\log 0.2]\times 2+\frac{1}{7}[-0.5\log 0.5-0.5\log 0.5]\times 2$$

$$=0.801bit/\text{符号}$$

3.6 信源的相关性与剩余度

▲信源的相关性

▲信源剩余度（冗余度）

▲自然语言的相关性和剩余度

3.6.1 信源的相关性

信源的相关性就是信源符号间的依赖程度。设信源有m个符号，那末对于不同情况可以分别计算信源的熵为：

$H_0 = \log m$ （符号独立等概）

$H_1 = H(X_1)$ （独立信源）

$H_2 = H(X_2 / X_1)$ （一阶马氏源）

$H_n = H(X_n / X_1 \cdots X_{n-1})$ （n-1阶马氏源）

H_{n+1}

$$H_\infty = \lim_{n\to\infty} H(X_n / X_1 \cdots X_{n-1})$$

由平稳性与熵的不增原理，有： $H_0 \geq H_1 \geq H_2 \geq \cdots H_\infty$

信源符号提供的平均自信息随符号间依赖关系长度增加而减少。

3.6.2 信源剩余度（冗余度）

为描述信源的相关性，引入信源效率和剩余度的概念。

信源效率：$$\eta = \frac{H_\infty}{H_0}$$

信源剩余度：$$\gamma = 1 - \frac{H_\infty}{H_0}$$

γ 越大表示信源可以被无失真压缩的程度越高

说明：H_∞ 是信源的实际熵，H_0 是信源的最大熵

γ 表示信源剩余度的比例

选择适当的马氏源描述实际信源

- 物理世界的信源大多有记忆，以马氏源描述实际信源是重要的工程问题;
- 对于非平稳信源，其 H_∞ 不存在。工程上一般可以合理假定信源平稳
- 对于平稳信源，可以进一步假定为 n 阶马氏源，以其平均信息熵 H_{n+1} 来近似大多数平稳信源，即:

$$H_\infty = H_{n+1} = H(X_{n+1} \mid X_1 X_2 \cdots X_n)$$

n=1时有: $H_{n+1} = H_{1+1} = H_2 = H(X_2 \mid X_1)$

n=0—》无记忆信源 $H_1 = H(X)$; 无记忆等概: $H_0 = \log q$

3.6.3 自然语言的相关性和剩余度(1)

英文字母概率表

字母	概率	字母	概率	字母	概率
空格	0.1859	I	0.0575	R	0.0484
A	0.0642	J	0.0008	S	0.0514
B	0.0127	K	0.0049	T	0.0796
C	0.0218	L	0.0321	U	0.0228
D	0.0317	M	0.0198	V	0.0083
E	0.1031	N	0.0574	W	0.0175
F	0.0208	O	0.0632	X	0.0013
G	0.0152	P	0.0152	Y	0.0164
H	0.0467	Q	0.0008	Z	0.0005

有关英语信源的马氏近似

考虑到英语字母间的依赖关系，可以将英语信源以马氏源描述，例如一阶/二阶马氏源。

如果以一阶马氏源近似，需要计算字母间的一维条件概率 $p(a_j \mid a_i)$，$a_i, a_j \in$[英文字母表]。$H_2 = 3.32 bit/s$

如果以二阶马氏源近似，需要计算字母间的二维条件概率 $p(a_k \mid a_i a_j)$，$a_k, a_i, a_j \in$[英文字母表]。这时需要计算$27^3 = 19,638$项二维条件概率，计算量巨大。$H_3 = 3.1 bit/s$

3.6.3 自然语言的相关性和剩余度(2)

对于实际英文字母组成的信源：

$$H_\infty = 1.4 \text{ (比特/符号)}$$

$$\eta = \frac{H_\infty}{H_0} = 0.29$$

$$\gamma = 1 - 0.29 = 0.71$$

表明：100页英语书可以压缩71%

3.6.3 自然语言的相关性和剩余度(3)

汉字近似概率表

类别	汉字个数	所占概率P	每个汉字的概率P_i
Ⅰ	140	0.5	0.5/140
Ⅱ	625-140=485	（0.85-0.5）=0.35	0.35/485
Ⅲ	2400-625=1775	（0.997-0.85）=0.147	0.147/1775
Ⅳ	7600	0.003	0.003/7600

3.6.3 自然语言的相关性和剩余度(4)

根据上表，可近似估算汉语信源的信息熵：

$$H(X) = -\sum_{i=1}^{10000} p_i \log p_i \quad = 9.773 \text{(比特/符号)}$$

$$\gamma = 1 - \frac{H(X)}{H_0} \cong 0.264$$

本 章 小 结 (1)

1 离散信源X的N次扩展源的熵 $H(X^N) \le N H(X)$，仅当信源无记忆时等式成立；离散信源X的N次扩展源的平均符号熵 $H_N(X) = \frac{1}{N} H(X^N) \le H(X)$，仅当信源无记忆时等式成立。

本 章 小 结 (2)

2. 有记忆信源的符号熵:

$$H_\infty(X)=\lim_{N\to\infty}\frac{1}{N}H(X^N)=\lim_{N\to\infty}H(X_N/X_1\cdots X_N)$$

并且 $H(X)\geq H_2(X)\geq\cdots\geq H_N(X)\geq\cdots\geq H_\infty(X)$

3. 马氏源的符号熵: $H_\infty(X)=\sum_j \pi_j H(X/s=j)$

$$H(X/s=j)=-\sum_{i=1}^{n}p_j(a_i)\log p_j(a_i)\ ,\ [\pi]=[\pi][P]$$

4. 信源剩余度 $\gamma=1-\eta=1-\frac{H_\infty}{H_0}$

课后习题

Page.60

3.5、3.10、3.11、3.17