第3章

离散信源

本章主要内容

3.1离散信源的分类与数学模型

3.1.1 离散信源的分类

3.1.2 离散无记忆信源数学模型

3.1.3 离散有记忆信源数学模型

3.1.4 离散平稳信源数学模型

3.2 离散无记忆信源的扩展

3.2.1 等长消息扩展

3.2.2 变长消息扩展

本章主要内容

3.3 离散平稳信源的熵

3.3.1　单符号信源的熵

3.3.2　等长无记忆扩展源的熵

3.3.3　变长无记忆扩展源的熵

3.3.4　平稳有记忆信源的熵

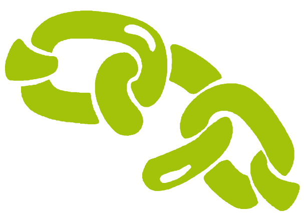
3.4 有限状态马尔可夫链

3.4.1马氏链的基本概念

3.4.2 齐次马氏链

3.4.3 马氏链状态分类

3.4.4 马氏链的平稳分布



本章主要内容

3.5 马尔可夫信源

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

3.6 信源的相关性与剩余度

- 3.6.1 信源的相关性
- 3.6.2 信源的剩余度
- 3.6.3 文本信源

3.1 离散信源的分类与数学模型

3.1.1 离散信源的分类

3.1.2 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

根据符号集的取值范围：有限/无限

根据信源统计特性：平稳/非平稳

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

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

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

- A={a_1，…，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

  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:

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

Xi的符号集 $A = \{ a_1 \ \cdots \ a_n \}$

X^N的符号集 $A^N = \{ \alpha_1 \ \cdots \ \alpha_M \}$

3.1.2 离散无记忆信源的数学模型

因为信源是无记忆的，所以：

$$p(\mathbf{x}) = p(x_1, x_2, \cdots, x_N) = \prod_{i=1}^{N} p(x_i)$$

可见，多维离散无记忆信源模型总可以通过单符号离散信源模型得到。也可以说，离散无记忆信源模型可以用多个单符号离散信源模型来描述。

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。

3.1.4 离散平稳信源数学模型

- 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。

3.1.4 离散平稳信源数学模型

平稳序列的统计特性与时间的推移无关

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

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

3.1.4 离散平稳信源数学模型

■ 例3.2

一平稳信源X的符号集A={0，1}，产生随机序列{xn}，

其中P(x1=0)=p， 求P(xn=1)（n >1）的概率。

解：

$$\text{平稳性} \Rightarrow P(x_n = 0) = p$$

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

3.1.4 离散平稳信源数学模型

■ 例3.2续

对同一信源，若P(x1=0，x2=1)=b，求P(x4=1/x3=0)。

解：

$$\text{平稳性} \Rightarrow P(x_3=0,x_4=1)=P(x_1=0,x_2=1)=b$$

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

$$\text{而} \quad P(x_2=1/x_1=0)=P(x_1=0,x_2=1)/P(x_1=0)=b/p$$

3.2 离散无记忆信源的扩展

3.2.1 等长消息扩展

3.2.2 变长消息扩展

3.2.1 等长消息扩展

▲ 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 $X^N = X_1X_2,,,X_N,(X_i\text{与}X\text{同分布})$

▲ 信源与其扩展源的关系：

$$X \longrightarrow \boxed{\text{N个连续输出的符号合并}} \xrightarrow{X^N = X_1X_2,,,X_N}$$

3.2.1 等长消息扩展

例3.3　求 例3.1 中信源的二次扩展模型。

解：

① 二元信源X的符号集为 {0, 1}

$$\Rightarrow 二次扩展源: X^2 = X_1X_2$$

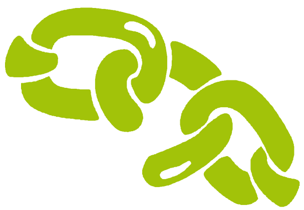
$$符号集:\{00,01,10,11\}$$

$$\Rightarrow X^2的模型:$$

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

② X^2各符号的概率为:

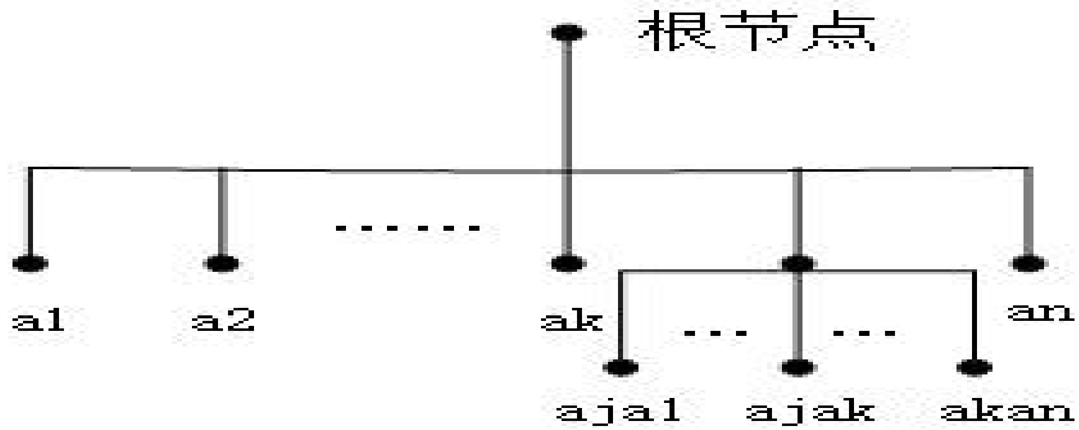
$$p(\alpha_1) = p^2, p(\alpha_2) = p(1-p) = p(\alpha_3), p(\alpha_4) = (1-p)^2$$



3.2.2 变长消息扩展

构造消息图

设离散无记忆信源X，符号集为A，n=|A|

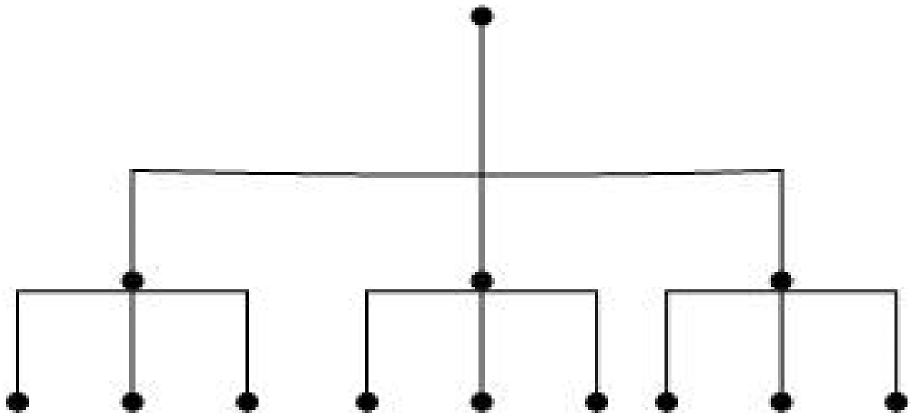


分裂终止时每片树叶表示的是从根节点到该叶路径上的信源消息，叶的概率就是消息的概率，叶的阶数就是消息长度。

如果消息构成满树，消息概率也满足归一化条件，这时消息集中的消息可视为某个信源的输出。这个信源称为信源X的变长扩展源

3.2.2 变长消息扩展

如果消息树是全树



就对应着信源的等长扩展。所以等长扩展可以视为变长扩展的特例。

3.2.2 变长消息扩展

什么消息集可以作为某信源的扩展？

概率满足归一化条件 → **完备消息集** → **适定消息集**

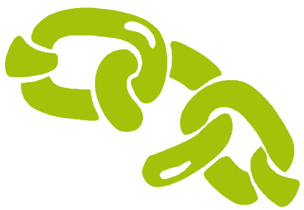
↑

异前置

例3.4 设例3.1中信源的消息集为A*={0, 10, 11}，求以此消息集进行变长扩展的信源符号的概率。

解 消息集A*中各个符号概率为

$$p(0)=p \qquad p(10)=p(1-p) \qquad p(11)=(1-p)^2$$



很明显，上面消息的概率满足归一化条件，并且是适定消息集。

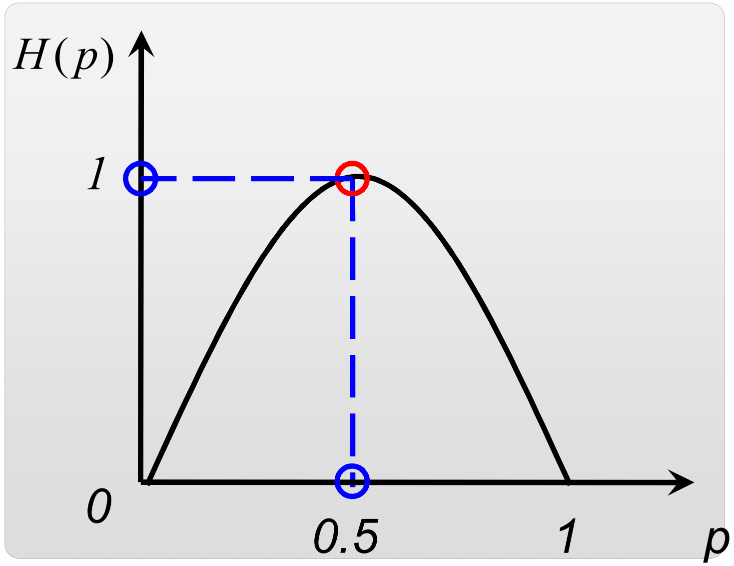
3.3 离散平稳信源的熵

无记忆信源的熵

平稳有记忆信源的熵

3.3.1 单符号信源的熵

H(p)



（1）具有熵的一切性质

（2）对p的导函数为 $H'(p)=\log\frac{1-p}{p}$

（3）p=0.5时，H(p)达到最大值1bit

（4）H(p)是p的上凸函数 $H''(p)=-(\log e)\frac{1}{p(1-p)}\prec 0$

3.3.2等长无记忆扩展源的熵

离散无记忆信源X的N次扩展源的熵等于信源X熵的N倍，即

$$H(X^N)=NH(X)$$

证明：

$$\text{熵的可加性}\xrightarrow[\text{② } X_i\text{互相独立且分布相同}]{\text{① 无记忆信源}}H(X^N)=\sum_{i=1}^{n}H(X_i)=NH(X)$$

3.3.2等长无记忆扩展源的熵

■ 例3.6

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：

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

$$= 2 \times 1.5 = 3 \quad bit/\text{扩展符号}$$

3.3.3变长无记忆扩展源的熵

离散无记忆信源的变长扩展源X*的熵为消息平均长度 与信源熵的乘积，即

$$H(X^{*}) = E(L)H(X)$$

$$\text{变长扩展} \xrightarrow{\text{特例}} \text{等长扩展}$$

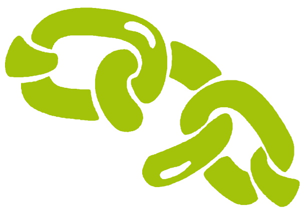
$$H(X^{*}) = E(L)H(X) \xrightarrow{\text{特例}} H(X^{N}) = N\ H(X)$$

3.3.3变长无记忆扩展源的熵

■ 例3.7

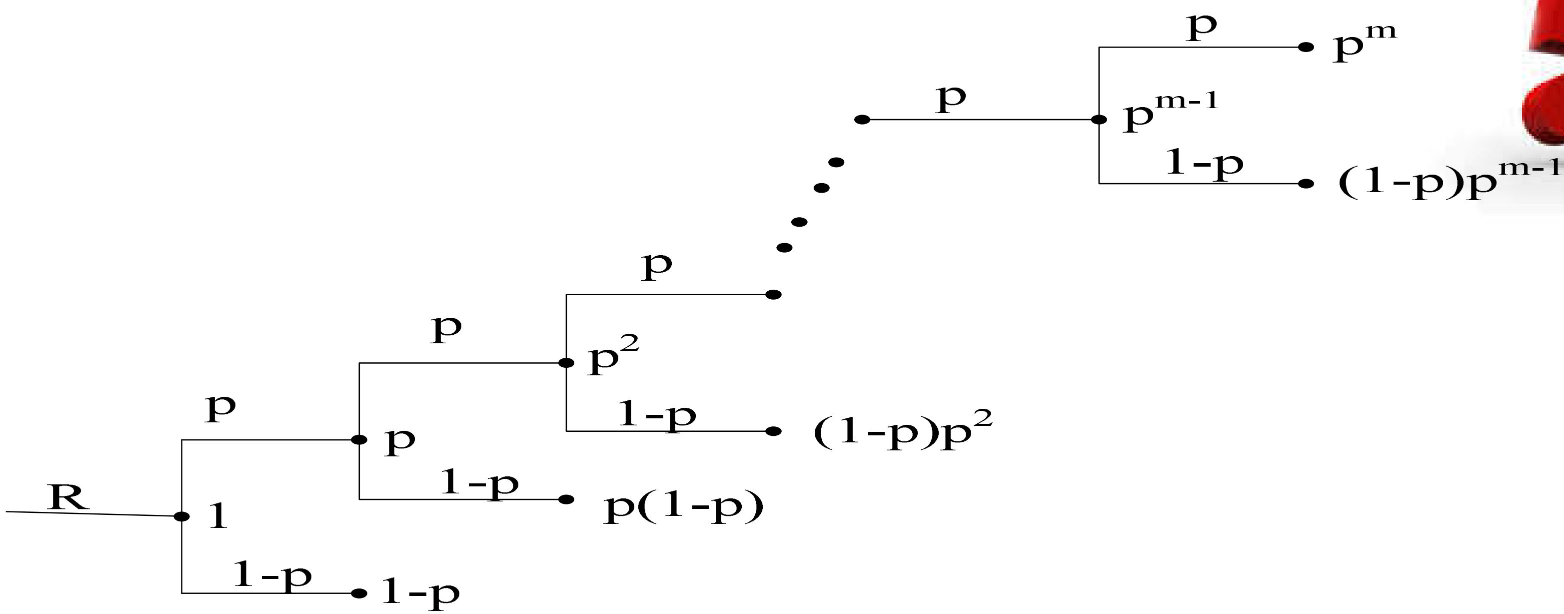
有一个二元无记忆信源X，发“ 0 ”的概率为p，对信源符号按下表进行分组得到一个新信源，符号集为Sn={s1, s2, s3, …, sm+1}，信源符号分组与信息信源符号的对应关系如下表：

信源消息	1	01	001	……	000…01 （m-1个“0”，1个“1”）	000…0 （m个“0”）
新信源符号	s_1	s_2	s_3	……	s_n	s_{m+1}



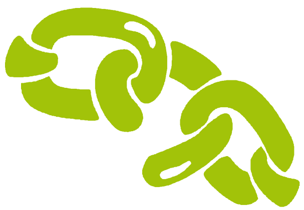
求新信源的熵Hn。

3.3.3变长无记忆扩展源的熵



消息的平均长度： $E(L)=1+p+\cdots+p^{m-1}=(1-p^m)/(1-p)$

新信源的熵Hn： $H_m=E(L)H(p)=H(p)(1-p^m)/(1-p)$



其中 $H(p)=-p\log p-(1-p)\log(1-p)$

当 $m\to\infty$ 时 $H_m\to\lim\limits_{m\to\infty}(1-p^m)/(1-p)=1/(1-p)$

3.3.4平稳有记忆信源的熵

根据平稳性和熵的不增原理

$$H(X^N) \le N\,H(X_1)$$，仅当无记忆信源时等式成立。

对于X的N次扩展源，定义平均符号熵为

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

3.3.4平稳有记忆信源的熵

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：符号熵或熵率

3.3.4平稳有记忆信源的熵

定理3.2 对任意离散平稳信源，若

$$H_1(X) < \infty$$

$H(X_N / X_1 \cdots X_{N-1})$不随N而增加

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

$H_N(X)$ 不随N而增加

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

说明：有记忆信源的符号熵也可通过计算极限条件熵得到

3.3.4平稳有记忆信源的熵

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

$$\text{证明：}\left.\begin{array}{l}\text{信源的平稳性}\\ \\ \text{熵的不增原理}\end{array}\right\}\Rightarrow$$

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

这说明对于平稳信源，条件越多，条件熵越不增加

3.3.4平稳有记忆信源的熵

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

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

$$
\begin{aligned}
NH_N(X) &= H(X_1 \cdots X_N) \\
&= H(X_1) + H(X_2 / X_1) + \cdots + H(X_N / X_1 \cdots X_{N-1}) \\
&\ge N\,H(X_N / X_1 \cdots X_{N-1})
\end{aligned}
$$

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

平均符号熵不小于条件熵

3.3.4平稳有记忆信源的熵

（3） $H_N(X)$ 不随N而增加

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

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

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

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

平均符号熵不随序列的长度而增加

3.3.4平稳有记忆信源的熵

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

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

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

3.3.4平稳有记忆信源的熵

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

利用(1)的结果与平稳性，有

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

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

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

3. 3. 4平稳有记忆信源的熵

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

先令 $j\to\infty$ ，后令 $N\to\infty$ ，得

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

另外，由（2）的结果，当 $N\to\infty$时, 有

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

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

所以 证毕。

3.3.4平稳有记忆信源的熵

定理3.3的注释

（1）信源熵率等于最小的平均符号熵；

（2）该定理提供了通过计算极限条件熵计算平稳信源熵率的方法

（3）当信源记忆长度有限时，计算极限条件熵通常要比计算极限平均符号熵容易得多。

3. 4 有限状态马尔科夫链

内容

- 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

定义

▲$1,\cdots,J$ 构成的集合Ω：状态集合

▲ $1,\cdots,J$ ：状态

▲ 随机变量 x_n ：马氏链在n时刻的状态

▲ *设随机序列* $\{x_n, n \geq 0\}$ ，
每个随机变量 x_n *对* $x_n, x_{n-1}, \cdots$ *的依赖只通过* x_{n+1}*实现，即*
$p(x_{n+1}=j|x_n=i, x_{n-1}=k, \cdots, x_0=m)=$
$p(x_{n+1}=j|x_n=i)$

3.4.1 马氏链的基本概念

对于一阶马氏链，下述说法等价：

- ●随机变量 x_{n+1} 对 $x_n, x_{n-1}, \cdots$ 的依赖只通过 x_n 实现；
- ●在 x_{n-1} 给定条件下，x_n 与 $x_{n-2}, x_{n-3}, \cdots$ 无关；
- ●在 x_{n-1} 给定条件下，x_n 与 $x_{n-2}, x_{n-3}, \ldots$ 是条件独立的；
- ●在 x_{n-k} 给定条件下，x_n与 $x_{n-k-1}, x_{n-k-2}, \cdots$ 是条件独立的。

上述等价说法可以推广到多阶马氏链。

3.4.1 马氏链的基本概念

- 一阶马氏链的当前状态只与前一个状态有关
- m阶马氏链的当前状态只与前m个状态有关
- 马氏链是时间离散，状态也离散的随机过程
- 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

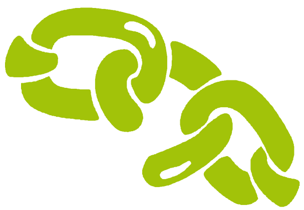
3.4.1 马氏链的基本概念

描述马氏链的最重要的参数：状态转移概率

对于离散时刻n、l，相应的状态转移概率可表示为：

$$p(s_l = j / s_n = i) = p_{ij}(n,l)$$

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



3.4.1 马氏链的基本概念

转移概率的主要性质

$$0 \le p_{ij}(n,l) \le 1, \quad i, j \in \Omega;$$

$$\sum_j p_{ij}(n,l) = 1;$$

一步转移概率 $p_{ij}(n,n+1) = p(x_{n+1} = j / x_n = i) = p_{ij}(n)$

其中n为起始时刻$i, j \in \Omega$

3.4.1 马氏链的基本概念

转移概率的主要性质

K步转移概率 $p_{ij}^{(k)}(n) = p(x_{n+k} = j / x_n = i)$

0步转移概率 $p_{ij}^{(0)} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

系统在任何时刻必处于 Ω 中某一状态

3.4.2 齐次马氏链（1）

若马氏链转移概率与起始时刻无关，即对任意n, 有

$$p_{ij}(n) = p(x_{n+1} = j / x_n = i) = p_{ij}$$

则称为齐次马氏链。对齐次马氏链，仍然有

$$p_{ij} \geq 0, \sum_j p_{ij} = 1$$

齐次马氏链K步转移概率也与起始时刻无关，写成$p_{ij}^{(k)}$

3.4.2 齐次马氏链（2）

表示方法

- 转移概率矩阵

$$\mathbf{P}=\left(p_{ij}\right)=\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1J} \\ p_{21} & p_{22} & \cdots & p_{2J} \\ \cdots & \cdots & \cdots & \cdots \\ p_{J1} & p_{J2} & \cdots & p_{JJ} \end{pmatrix}$$

- 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

3.4.2 齐次马氏链（3）

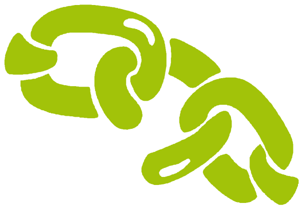
例3.8　一个矩阵，验证此矩阵对应一个齐次马氏链的转移概率矩阵并确定此马氏链的状态数

$$\mathbf{P}=\begin{array}{c}\begin{bmatrix}1/3 & 1/3 & 1/3\\ 1/4 & 1/2 & 1/4\\ 1/4 & 1/4 & 1/2\end{bmatrix}\end{array}\begin{array}{l}=1\\ =1\\ =1\end{array}$$

解：

$$\left.\begin{array}{l}\text{① 元素非负}\\ \text{每行和为1}\end{array}\right\}\ \text{齐次马氏链转移概率矩阵}$$

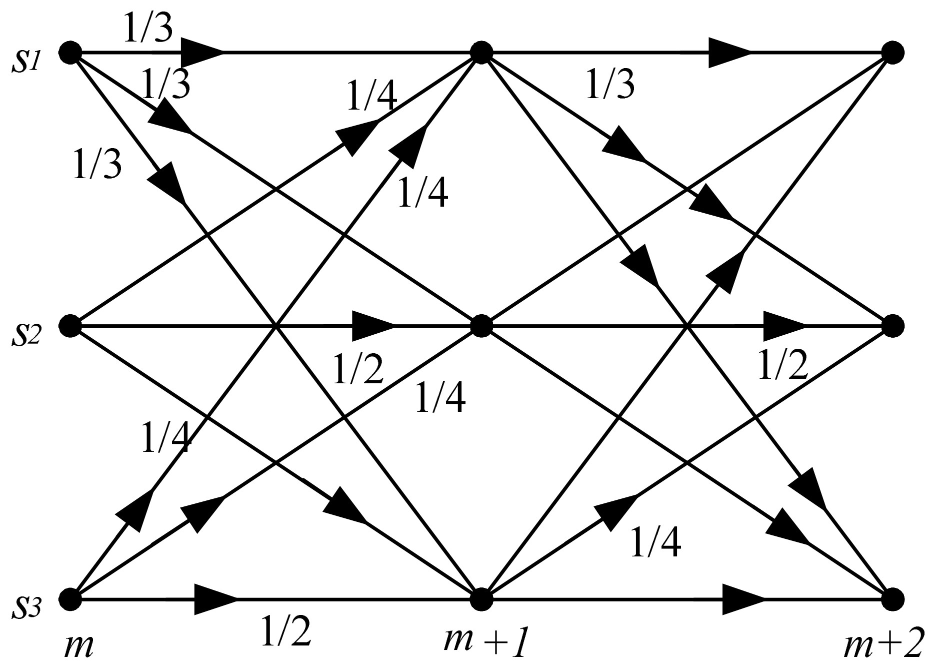
② 状态数=3



3.4.2 齐次马氏链（4）

例3.8 续　　画出此马氏链的网格图。

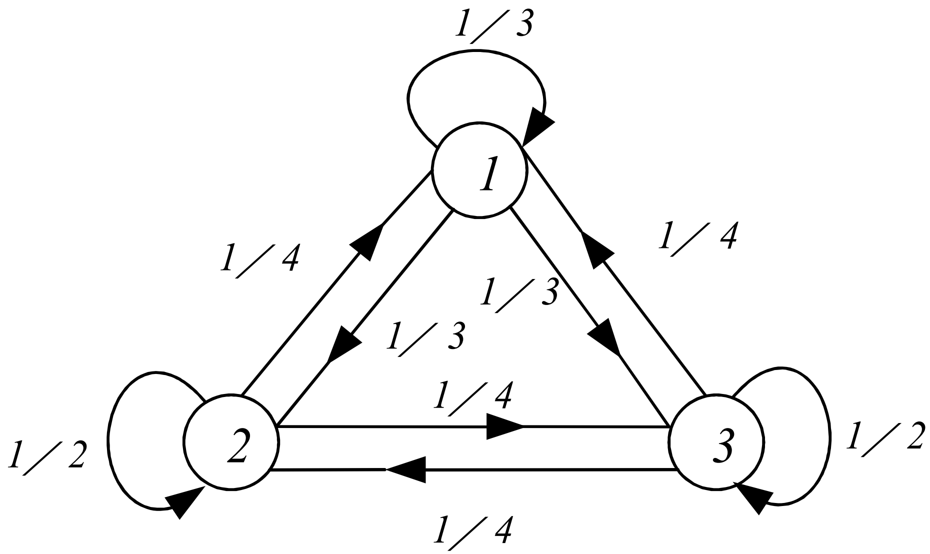
解：



3.4.2 齐次马氏链（5）

例3.8 续　画出此马氏链的状态转移图

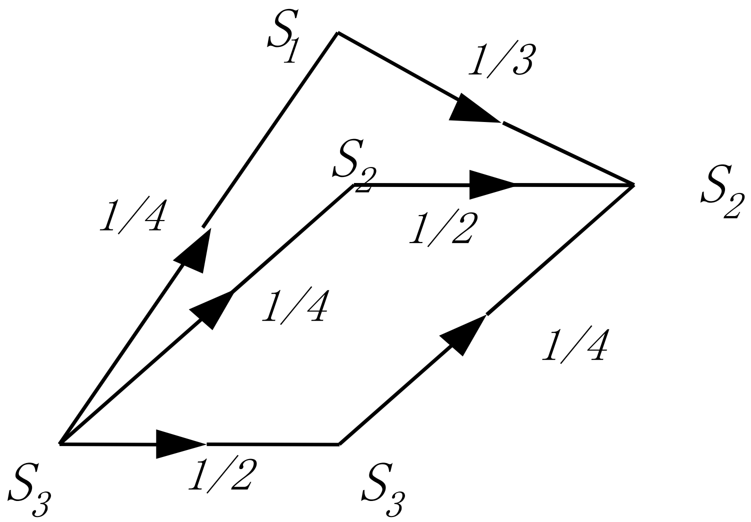
解：



3.4.2 齐次马氏链（6）

例3.8 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（6）

解：

下面分两步来计算：

1）计算从m时刻从s3经m+1时刻某状态sk到m+2时刻s2的转移概率

2）对1）中计算的经m+1时刻的所有状态sk（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 $\mathbf{P}^2$ 的第(i, j)个元素

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

4 $\mathbf{P}^{m+n} = \mathbf{P}^m \mathbf{P}^n \Rightarrow p_{ij}^{(m+n)} = \sum_{k} p_{ik}^{(m)} p_{kj}^{(n)}$

Kolmogorov-Chapman方程(2)

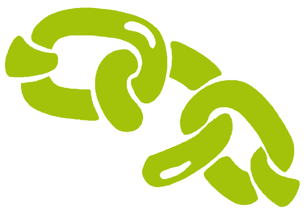
$$p_{ij}^{(m+n)} = p(x_{m+n+l} = j \mid x_l = i) = \sum_k p(x_{m+n+l} = j, x_{l+m} = k \mid x_l = i)$$

$$= \sum_k p(x_{l+m} = k \mid x_l = i) p(x_{m+n+l} = j \mid x_l = i, x_{l+m} = k)$$

$$= \sum_k p(x_{l+m} = k \mid x_l = i) p(x_{m+n+l} = j \mid x_{l+m} = k)$$

$$= \sum_k p_{ik}^{(m)} p_{kj}^{(n)}$$

提供了计算多步转移概率的方法



Kolmogorov-Chapman方程(3)

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

其中J为状态数，经k步转移后的状态概率分

布为 $\mathbf{p}^{(k)}=(p_1^{(k)},p_2^{(k)},\cdots,p_J^{(k)})^T$ ，则有：

$$(\mathbf{p}^{(k)})^T=(\mathbf{p}^{(0)})^T\mathbf{P}^k=(\mathbf{p}^{(m)})^T\mathbf{P}^{k-m}$$

- 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(4)

例3.9

设例3.8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。

Kolmogorov-Chapman方程(4)

解：

$$\mathbf{p}^{(1)} = \mathbf{p}^{(0)}\mathbf{P} = \begin{pmatrix}1/2 & 1/4 & 1/4\end{pmatrix}\begin{pmatrix}1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2\end{pmatrix}$$

$$= \begin{pmatrix}7/24 & 17/48 & 17/48\end{pmatrix}$$

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

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

3.4.3 马氏链状态分类（1）

▲ 若对某一k≥1，有 $p_{ij}^{(k)} > 0$ ，则称状态j可由状态i到达，记为 i→j

▲ 如果i→j ，且 j→i ，则称状态i与状态j可互通，记为**i↔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）

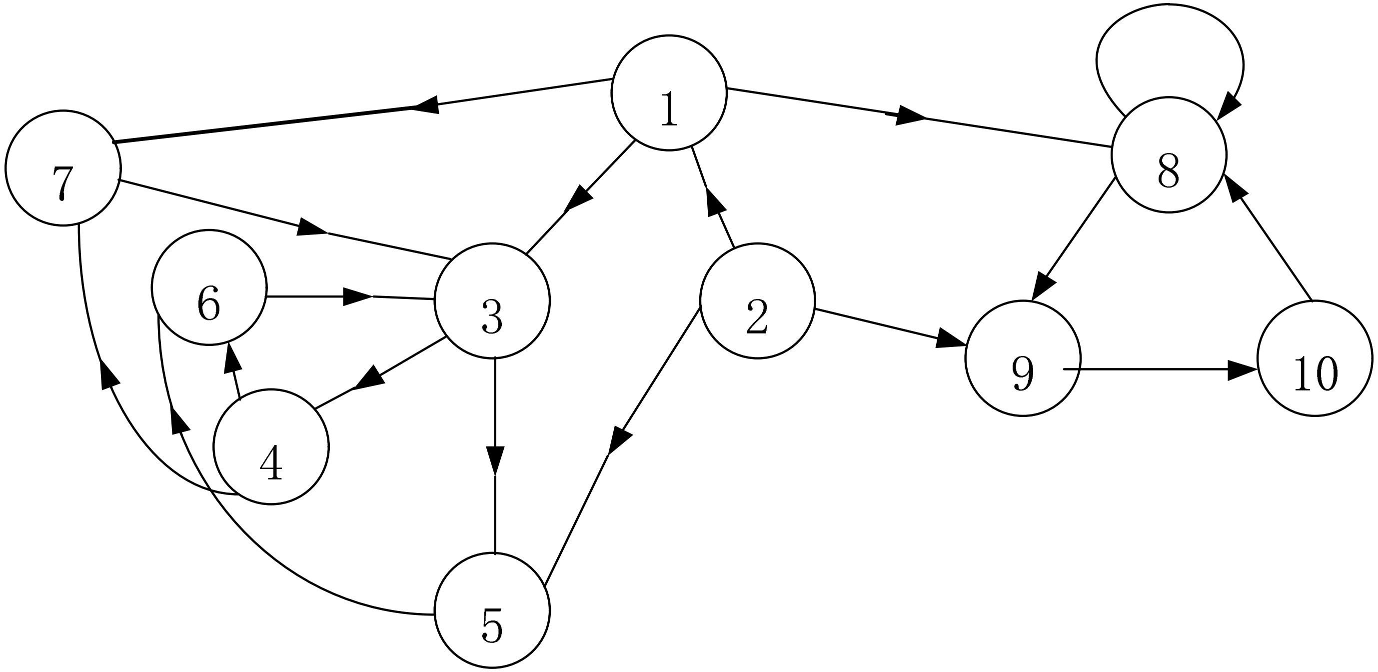
常返态　如果状态i是经过有限步后迟早要返回的状态；

过渡态　不是常返态。即存在某状态j经过若干步以后总能达到某一其它状态，但不能从其它状态返回。

在一个常返类中，$p_{ii}^{(k)}>0$ 表示从状态i出发经过k步首次返回到i状态的概率。设 d_i 为正整数集 $\{k:k\geq 1, p_{ii}^{(k)}>0\}$ 的最大公约数。若 $d_i=1$ ，则称为周期的，若 $d_i>1$ ，则称为非周期的或称为遍历的。

3.4.3 马氏链状态分类（3）

例3.10　按互通性将状态分成若干类（子集）



3.4.3 马氏链状态分类（4）

解：

$$C_1 = \{1\}$$

$$C_2 = \{2\}$$

过渡态

$$C_3 = \underbrace{\{3,4,5,6,7\}}$$

$$C_4 = \underbrace{\{8,9,10\}}$$

常返态

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

$d_i > 1$ 周期的

$d_i = 1$ 非周期

3.4.3 马氏链状态分类（4）

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。

非周期的常返态称为遍历态

3.4.4 马氏链的平稳分布（1）

若对任意整数m, n, 马氏链的状态分布满足

$$P(x_m = i) = P(x_n = i) = \pi_i \qquad （3.25）$$

则称 $\{\pi_i\}$ 为平稳分布，或稳态分布。

其中， $\boldsymbol{\pi} = \left(\pi_1, \pi_2, \cdots, \pi_J\right)^T$ 为平稳分布列矢量

3.4.4 马氏链的平稳分布（2）

$$\mathbf{p}^{(0)} = (p_1^{(0)}, p_2^{(0)}, \cdots, p_J^{(0)})^T = \boldsymbol{\pi}, \quad \mathbf{p}^{(n)} = \boldsymbol{\pi}, \qquad n > 0$$

平稳马氏链

定理3.5 如果一个遍历有限状态马氏链的转移概率矩阵为 P 那么

$$\lim_{k\to\infty} \mathbf{P}^k = \mathbf{e}\boldsymbol{\pi}^T$$

其中，$\mathbf{e} = (1,1,\cdots,1)^T$，为$J$维列矢量，$\boldsymbol{\pi}$为平稳分布列矢量

因此，$\lim\limits_{k\to\infty} \mathbf{P}^k$是一个每行都相同，都等于$\boldsymbol{\pi}^T$的矩阵。

3.4.4 马氏链的平稳分布（3）

$$\lim_{k\to\infty}(\mathbf{p}^{(k)})^T = \lim_{k\to\infty}\{(\mathbf{p}^{(0)})^T\mathbf{P}^k\} = (p^{(0)})^T\mathbf{e}\boldsymbol{\pi}^T = \boldsymbol{\pi}^T$$

- 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稳定值（稳态分布），与初始状态概率分布无关。
- k时刻： $\boldsymbol{p}^{(k)} = (p_1^{(k)}, p_2^{(k)}, \cdots; p_J^{(k)})^T$
- 初始概率：$\mathbf{p}^{(0)}$
- 平稳分布：$\boldsymbol{\pi}$

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；

●如果马氏链中仅存在一个常返类，则方程(3.31)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；

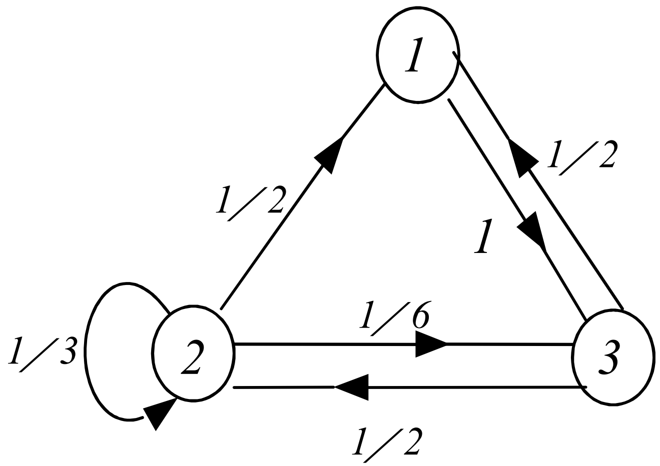
●如果马氏链中仅存在一个常返类而且是非周期的（即遍历的），则(3.32)式成立；如果有多个常返类，但都是非周期的，则Pn 也收敛，但矩阵的每行可能不同；如果马氏链具有一个或多个周期常返类，则 Pn 不收敛

3.4.4 马氏链的平稳分布（5）

例3.11

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 $\lim\limits_{k\to\infty}\mathbf{P}^k$ 的值。

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



3.4.4 马氏链的平稳分布（6）

解：

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

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

3.4.4 马氏链的平稳分布（7）

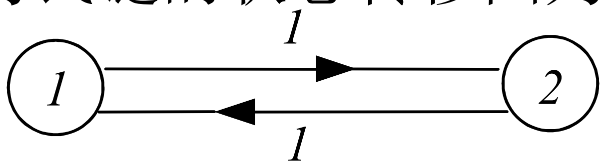
例3.12

一马氏链的状态转移矩阵为 $\mathbf{P}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$

确定它的n次幂是否收敛并求其平稳分布

3.4.4 马氏链的平稳分布（7）

解：该马氏链的状态转移图为：



所以马氏链是一个周期常返类，周期=2。

由于 $\mathbf{P}^{2k}=\mathbf{I} \qquad \mathbf{P}^{2k+1}=\mathbf{P}$

另外，马氏链是有限状态的，因此存在唯一的平稳分布。

$$\left.\begin{array}{l}[\pi_1\ \pi_2]\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}=[\pi_1\ \pi_2] \\ \pi_1+\pi_2=1\end{array}\right\}\Rightarrow \pi_1=\pi_2=\frac{1}{2}$$

3.5 马尔可夫信源

1. 马氏源的基本概念

2. 马氏源的产生模型

3. 马氏源N次扩展源熵的计算

4. 马氏源符号熵的计算

3.5 马尔可夫信源的基本概念（1）

马氏源：

▲ 当前时刻输出符号的概率仅与当前时刻的信源状态有关

▲下一时刻的信源状态由当前信源状态和当前输出符号唯一确定。

$$p\{s_{k+1}=j/x_k=a_l,s_k=i\}=\begin{cases}1 & s_{k+1}=g\left(s_k,x_k\right)\\0 & s_{k+1}\neq g\left(s_k,x_k\right)\end{cases}$$

m阶马氏链的处理方法(2)

★

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$(x_{k-m}\cdots x_{k-1})\rightarrow s_k$

x取遍 $A=\{a_1\cdots a_n\}$ ，k=m, m+1, …； 状态 s_k

取自 $\Omega=\{1,2,\cdots,n^m\}$ ， 为状态数；n^m

m阶马氏链的处理方法(2)

3）符号转移概率转换成状态转移概率

$p(s_{k+1}/s_k)=p(x_k/s_k)$ 其中 （$x_{k-m}\ \cdots\ x_{k-1}$）$\rightarrow s_k$ ，（$x_{k-m+1}\ \cdots\ x_k$）$\rightarrow s_{k+1}$

4）得到马氏源模型：

$$\mathbf{P}=\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1,n^m} \\ p_{21} & p_{22} & \cdots & p_{2,n^m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n^m,1} & p_{n^m,2} & \cdots & p_{n^m,n^m} \end{pmatrix}$$

其中，$p_{ij}=p(s_{k+1}=j|s_k=i)(i,j=1,\cdots,n^m)$

3.5.1马氏源的基本概念(3)

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，

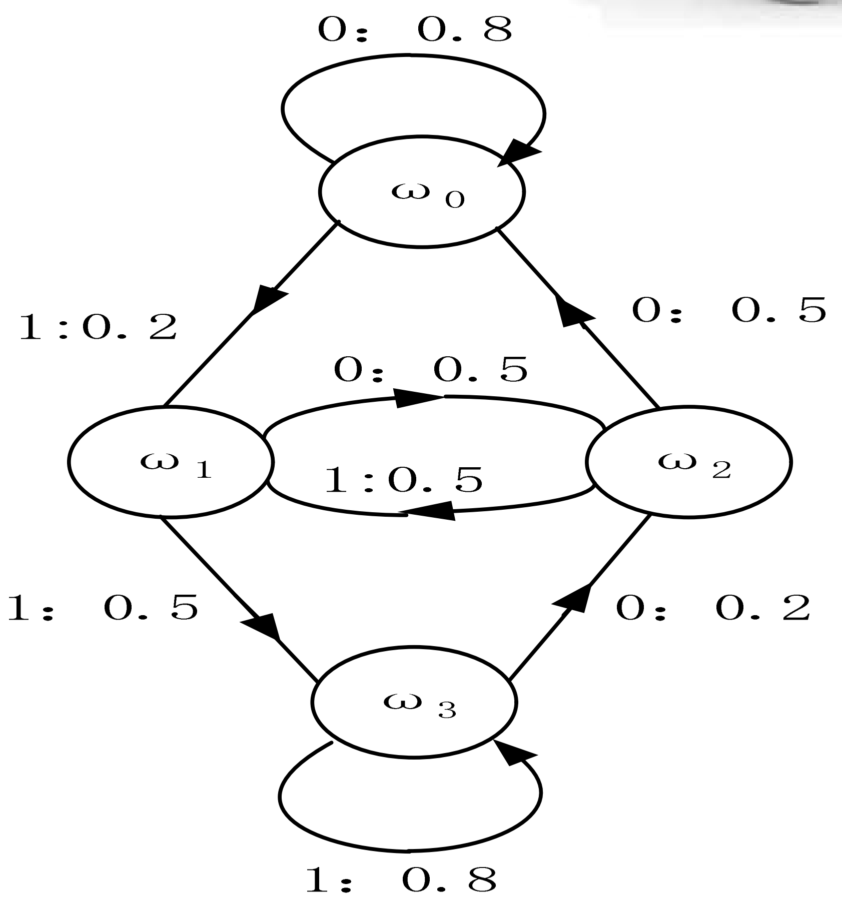
p(0/01)=p(0/10)=p(1/01)=p(1/10)=0.5

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(4)

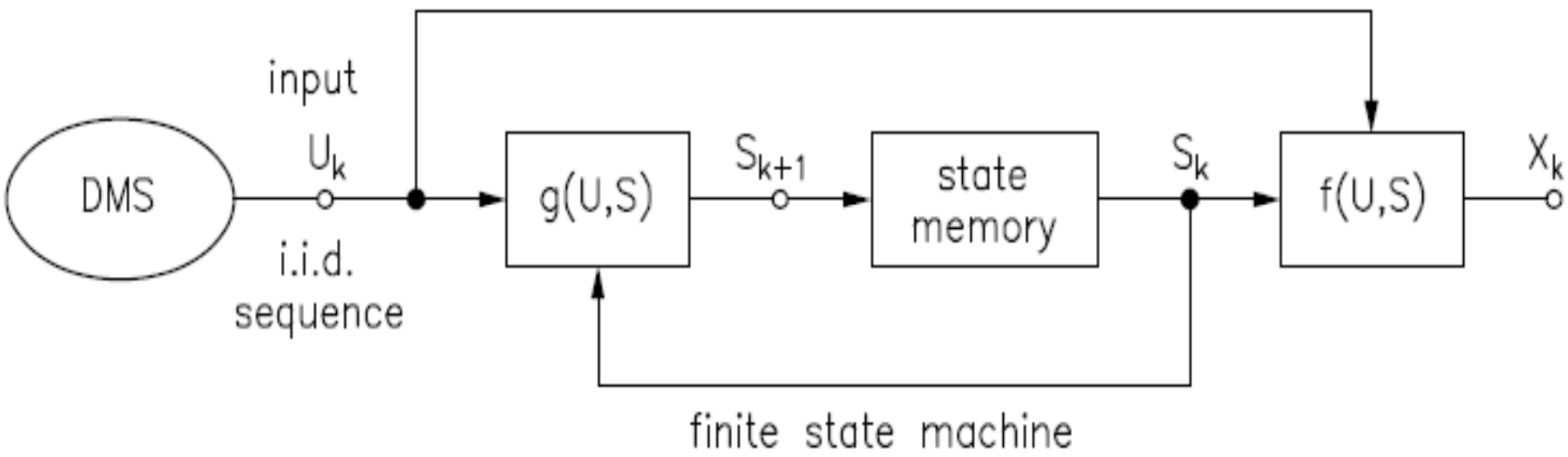
解：

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



3.5.2 马氏源的产生模型(1)

- 平稳转移概率的马氏源的产生模型



3.3.2 离散平稳有记忆信源的熵

$$p(x_k \mid s_k) = \sum_{u_k} p(x_k, u_k \mid s_k) = \sum_{u_k} p(u_k \mid s_k) p(x_k \mid s_k, u_k)$$

$$= \sum_{u_k} p(u_k) p(x_k \mid u_k, s_k)$$

$$\because\ p(x_k \mid u_k, s_k) = \begin{cases} 1 & x_k = u_k, s_k \\ 0 & x_k \neq u_k, s_k \end{cases}$$

$$\therefore\ p(x_k \mid s_k) = \sum_{u_k : x_k = f(u_k, s_k)} p(u_k)$$

当x_k, u_k为二元信源、一一对应关系时，

$$p(x_k \mid s_k) = \begin{cases} p(u_k) & x_k = f(u_k, s_k) \\ 1 - p(u_k) & x_k \neq f(u_k, s_k) \end{cases}$$

3.5.2 马氏源的产生模型(3)

例3.14

设独立随机列$\{u_n\}$，$p(u_n=0)=p$，$p(u_n=1)=q$，$p+q=1$，随机序列$\{x_n\}$与$\{u_n\}$的关系为$x_n=u_n\oplus x_{n-1}\oplus x_{n-2}$ 其中$\oplus$为模2加；问：

（1）随机序列$\{x_n\}$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(4)

解：

序列$\{x_n\}$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$\{x_n\}$构成二阶马氏链。设 $s_n = x_{n-2}x_{n-1}$ 条件概率：

$$p(s_{n+1} | s_n) = \sum_{n:u_n = x_n = x_n \oplus x_{n-1} \oplus x_{n-2}} p(u_n)$$

$$= p(u_n)|_{u_n = x_n \oplus x_{n-1} \oplus x_{n-2}}, s_{n+1} = (x_{n-1} \ \ x_n)$$

3.5.2 马氏源的产生模型(5)

解：

所求条件概率：

$$
\begin{aligned}
& p(x_n=0\mid x_{n-2}=0,x_{n-1}=0)=p(x_n=1\mid x_{n-2}=0,x_{n-1}=1) \\
& =p(x_n=1\mid x_{n-2}=1,x_{n-1}=0)=p(x_n=0\mid x_{n-2}=1,x_{n-1}=1) \\
& =p(u_n=0)=p
\end{aligned}
$$

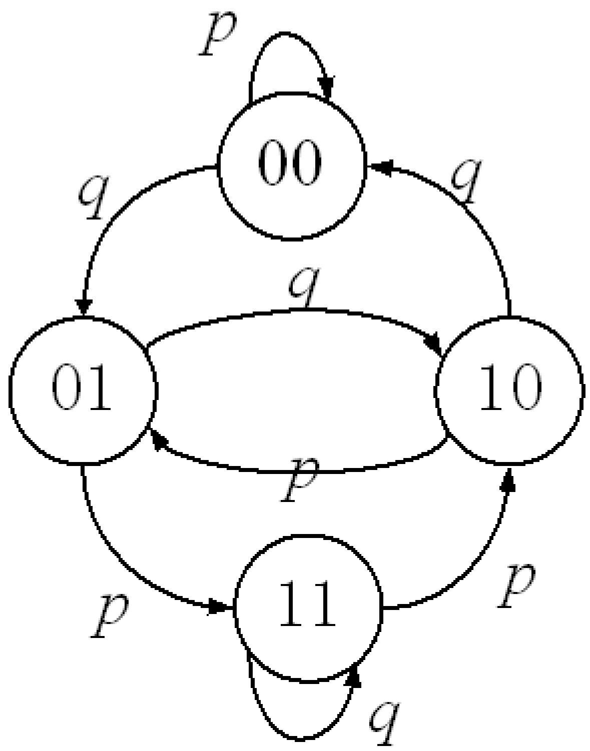
$$
\begin{aligned}
& p(x_n=1\mid x_{n-2}=0,x_{n-1}=0)=p(x_n=0\mid x_{n-2}=0,x_{n-1}=1) \\
& =p(x_n=0\mid x_{n-2}=1,x_{n-1}=0)=p(x_n=1\mid x_{n-2}=1,x_{n-1}=1) \\
& =p(u_n=1)=1-p
\end{aligned}
$$

3.5.2 马氏源的产生模型(6)

解：

因此，马氏链的

状态转移概率图：



3.5.3 马氏链N次扩展源的熵的计算(1)

■ 例3.15

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

3.5.3 马氏链N次扩展源的熵的计算(2)

解：

首先求平稳分布

$$\begin{cases} (p_0 \quad p_1) = (p_0 \quad p_1)\begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} \\ p_0 + p_1 = 1 \end{cases}$$

$$\Rightarrow \quad p_0 = 3/5, p_1 = 2/5$$

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

3.5.3 马氏链N次扩展源的熵的计算(3)

解：

类似得到

$$p(001)=0.6\times0.8\times0.2=0.096$$

$$p(010)=0.6\times0.2\times0.3=0.036$$

$$p(011)=0.6\times0.2\times0.7=0.084$$

$$p(100)=0.4\times0.3\times0.8=0.096$$

$$p(101)=0.4\times0.3\times0.2=0.024$$

$$p(110)=0.4\times0.7\times0.3=0.084$$

$$p(111)=0.4\times0.7\times0.7=0.196$$

3.5.3 马氏链N次扩展源的熵的计算(4)

做映射 $(x_{1+k} \cdots x_{m+k}) \rightarrow s_{m+1+k}(j),\ \mathrm{k}=0,\cdots,N-m$

其中k为时间标号，j为状态序号。

$H(X_1X_2\cdots X_N)=H(S_{m+1}S_{m+2}\cdots S_{N+1})$，其中，$S_k = X_{k-m}X_{k-m+1}\cdots X_{k-1}$

利用熵的可加性，将上式展开，并利用马氏性得

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

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

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

3.5.3 马氏链N次扩展源的熵的计算（5）

对于平稳马氏链，状态序列也是平稳的，状态条件熵也不随时间平移而改变。设状态平稳分布为

$$\boldsymbol{\pi}=\left(\pi_1,\pi_2,\cdots,\pi_{n^m}\right)^T$$

则上式变为：

$$H(X_1X_2\cdots X_N)=H(S_{m+1})+(N-m)H(S_{k+1}/S_k)$$

$$=H(\pi)+(N-m)\sum_{i=1}^{n^m}\pi_i h_i=H(\pi)+(N-m)\pi^T\mathrm{h}$$

其中，H(π)表示利用状态平稳分布计算的熵，

$h=(h_1h_2\cdots h_{n^m})^T$ 为列矢量。

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

N次扩展源的平均符号熵为:

$$H(X_1 X_2 \cdots X_N) = H(\boldsymbol{\pi}) + (N-m)\boldsymbol{\pi}^T \mathbf{h}$$

$$H_N(X) = \frac{1}{N} H(X_1 X_2 \cdots X_N) = \frac{1}{N}[H(\boldsymbol{\pi}) + (N-m)\boldsymbol{\pi}^T \mathbf{h}]$$

3.5.4 马氏源符号熵的计算(1)

计算方法1：

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:

$$H_{\infty}(X)=\lim_{N\to\infty}H_N(X)=\boldsymbol{\pi}^T\mathbf{h}$$

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)

- 例3.13续

  求信源的极限符号熵

解：

$$H_\infty(X) = \boldsymbol{\pi}^T \mathbf{h}$$

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

$$= 0.801 bit/\text{符号}$$

3.5.4 马氏源符号熵的计算(3)

计算方法2

●（FSM马氏源）当信源从某一状态转移到另一个新状态时，存在多个信源序列对应一个状态。这样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。

●给定当前信源状态条件下信源的输出符号熵为：

$$H(X/s=i)=-\sum_{l=1}^{n}p_i(a_l)\log p_i(a_l)$$

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

3.5.4 马氏源符号熵的计算(4)

▲ 在给定某特殊状态s1=j和以前的输出X1,X2,…Xk-1条件下当前输出符号Xk的熵满足：

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

对s1取平均

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

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

3.5.4 马氏源符号熵的计算(5)

对于平稳信源，状态概率与时间起点无关，所以

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

对于m阶平稳马氏链的符号熵为

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

利用平稳性及马氏性，有

$$H_\infty = H(X_{k+1} / X_1 \cdots X_k)$$

$$\triangleq H_{k+1}$$

3.5.4 马氏源符号熵的计算(6)

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

$$
\begin{aligned}
H(X_{k+1} / X_1 \cdots X_k) &= H(X_{k+1} / S_1, X_1 \cdots X_k) \\
&= \sum_{i=1}^{J} p(s = i) H(X / s = i) \\
&= \boldsymbol{\pi}^T \mathbf{h}
\end{aligned}
$$

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(7)

定理3.7 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(8)

（1）定理3.7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；

（4）比方法1更通用。

3.6 信源的相关性与剩余度

3.6.1 信源的相关性

3.6.2 信源剩余度（冗余度）

3.6.3 文本信源

3.6.1 信源的相关性

1、信源的相关性就是信源符号间的依赖程度。

由平稳性与熵的不增原理，有：
可见，符号相关程度越大，熵越小，反之亦然。

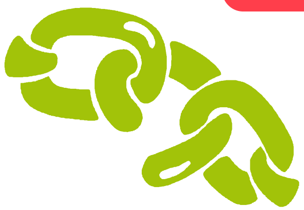
设信源有m个符号，那么对于不同情况可以分别计算信源的熵为

$H_0 = \log m$ （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）



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

3.6.2 信源剩余度（冗余度）

信源的冗余度是指信源中”多余”(即可以被无损压缩掉的)成分所占的比例。

信源效率：$\eta = \frac{H_\infty}{H_0}$

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

3.6.3 文本信源

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

3.6.3 文本信源

对于实际英文字母组成的信源：

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

$H_\infty=1.4$（比特/符号）　　$\gamma=1\text{-}0.29=0.71$

3.6.3 文本信源

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 文本信源

根据上表，可近似估算汉语信源的信息熵：

$$H(X) = -\sum_{i=1}^{10000} p_i \log p_i$$

$$= 9.773 \text{ （比特/符号）}$$

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

3.6.3 文本信源

例3.17 根据上面的统计结果，分别计算汉语语与英语信源的效率和剩余度。

对于汉语，信源效率：$\eta_2 = 4.1/13.29 = 0.31$

剩余度：$\gamma_2 = 1 - \eta_2 = 0.69$

对于英语，信源效率：$\eta_1 = 1.4/\log_2 27 = 0.29$

剩余度：$\gamma_1 = 1 - \eta_1 = 0.71$

本 章 小 结

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；

2 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、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：

$$H_\infty(X)=\sum_j \pi_j H(X/s=j)$$

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

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

4．信源剩余度

$$\gamma=1-\eta=1-\frac{H_\infty}{H_0}$$

谢谢!