第二章离散信源信息度量

第四节离散马氏信源

陈兴同

中国矿业大学 数学学院

2021 年 8 月

内容提要

内容提要

内容提要

1. 马氏信源
2. 马氏信源的联合分布：
3. 马氏信源转移概率

内容提要

内容提要

内容提要

字符的依赖性

信源发出的符号之间可能有依赖关系，比如语言符号间就是有依赖关系的，例如：当发出四个字符“Engl”之后，第五个字符发出“i”概率 P{i|Engl}就比较大；当发出三个字符“一马当”之后，第四个符发出“先”的概率几乎为 1！这就是信源符号间的依赖性，称为**记忆性**，马氏信源就是表示这种有记忆信源的数学模型。通常有记忆的信源要比无记忆的信源复杂。

如果离散随机序列 $X_1, X_2, \cdots, X_n, \cdots$，存在正整数 k，使得对任意 n 个字符 $x_1, x_2, \cdots, x_n$，都成立条件概率：

$$
\begin{aligned}
& P\{X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_2 = x_2, X_1 = x_1\} \\
= \quad & P\{X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_{n-k} = x_{n-k}\},
\end{aligned}
$$

则称这个随机序列为 k **阶 Markov 序列**或 k **阶马氏信源**，特别一阶 Markov 序列又称为 **Markov 链或马氏信源**，其中 k 称为记忆长度。

马氏信源定义

如果离散随机序列 $X_1, X_2, \cdots, X_n, \cdots$，存在正整数 k, 使得对任意 n 个字符 $x_1, x_2, \cdots, x_n$，都成立条件概率：

$$
\begin{aligned}
& P\{X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_2 = x_2, X_1 = x_1\} \\
= \ & P\{X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_{n-k} = x_{n-k}\},
\end{aligned}
$$

则称这个随机序列为 k **阶 Markov 序列**或 k **阶马氏信源**，特别一阶 Markov 序列又称为 **Markov 链或马氏信源**，其中 k 称为记忆长度。

离散马氏信源特点：

k **阶离散马氏信源的模型**是 k 阶马氏链，在时刻 n 发出的字符只依赖于它前面 k 个时刻已经发出的字符，与它更前面的时刻发出的字符无关。

（1） 如果 $k=1$ 就称为马氏信源；

（2） 马氏信源是一种有记忆的信源，参数 k 可以理解为记忆长度。

（3） 若信源字符空间为有限集则称为**有限状态马氏信源**，本节只讨论有限状态马氏信源。

联合分布：

计算联合概率 $p(x_1, x_2, \cdots, x_n) = P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\}$ 时要利用记忆性质。比如：对一阶马氏信源有联合概率公式

$$p(x_1, x_2, \cdots, x_n) = p(x_1)p(x_2|x_1)p(x_3|x_2)\cdots p(x_n|x_{n-1}), \tag{2.\cancel{1}3}$$

对 k 阶马氏信源有联合概率公式

$$\begin{aligned} p(x_1, x_2, \cdots, x_n) &= p(x_1, x_2, \cdots, x_k)p(x_{k+1}|x_1, x_2, \cdots, x_k) \\ & p(x_{k+2}|x_2, \cdots, x_{k+1})\cdots p(x_n|x_{n-k}, x_{n-k+1}, \cdots, x_{n-1}). \end{aligned} \tag{2.\cancel{2}4}$$

有了联合分布就完全确定了信源的统计规律。

确定马氏信源统计规律

对于 k 阶马氏信源，下标连续的 k 维随机变量 $(X_{n-k},\cdots,X_{n-1})$ 所有可能的取值 $(x_1,x_2,\cdots,x_k)$ 称为在时刻 $n-1$ 时的**状态矢量**，它们全在笛卡尔积集合 $\mathcal{X}^k$ 中，共有 N^k 种可能的取值，可以对所有这些状态矢量进行编号，记成 $s_i, i=1,2,\cdots,M$，每个 s_i 具有向量形式 $(x_1,x_2,\cdots,x_k)$ 或字符串形式 $x_1x_2\cdots x_k$。

根据公式（2.~~2~~4），如果能够确定 M 种状态的概率分布

$$p(s_i)=P\left\{X_{n-k}=x_1,X_{n-k+1}=x_2,\cdots,X_{n-1}=x_k\right\}=p(x_1,x_2,\cdots,x$$

及在每种状态下发出一个字符的条件概率 $p(x_{k+1}|s_i)$, 则就可以完全确定这个 k 阶马氏信源的统计规律性。

转移概率矩阵

设在 $n-1$ 时刻信源处于第 i 个状态 $s_i=(x_1,x_2,\cdots,x_k)$，当第 n 个时刻发出一个字符 x_{k+1} 后，信源处于第 j 个状态 $s_j=(x_2,x_3,\cdots,x_{k+1})$，称条件概率

$$p_{j|i}(n)=p(s_j|s_i)=p(x_{k+1}|x_1x_2\cdots x_k),i,j=1,2,\cdots,M$$

为第 n 时刻从状态 s_i 到状态 s_j 的**一步转移概率**。可以将这些转移概率构造成在时刻 n 处的**转移概率矩阵**

$$P(n)=(p_{j|i}(n))=\begin{pmatrix} p_{1|1}(n) & p_{2|1}(n) & \cdots & p_{M|1}(n) \\ p_{1|2}(n) & p_{2|2}(n) & \cdots & p_{M|2}(n) \\ \vdots & \vdots & \ddots & \vdots \\ p_{1|M}(n) & p_{2|M}(n) & \cdots & p_{M|M}(n) \end{pmatrix},$$

齐次马氏信源

如果这个转移概率矩阵关于时间不变，则称该马氏信源具有**齐次性**，这时转移概率矩阵写成

$$P = (p_{j|i}) = \begin{pmatrix} p_{1|1} & p_{2|1} & \cdots & p_{M|1} \\ p_{1|2} & p_{2|2} & \cdots & p_{M|2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1|M} & p_{2|M} & \cdots & p_{M|M} \end{pmatrix}.$$

多步转移概率矩阵

如果在时刻 n 时信源处于第 i 个状态 $s_i = (x_1, x_2, \cdots, x_k)$，经过 l 个时刻后第 $n+l$ 时刻信源处于第 j 个状态 $s_j = (y_1, y_2, \cdots, y_k)$，记这种状态转移概率为

$$p_{j|i}^{(l)} = P\left\{\text{时刻}n+l\text{时状态为}s_j|\text{时刻}n\text{时状态为}s_i)\right\}, i, j = 1, 2, \cdots, M,$$

称为 l **步转移概率**，由它所构成的矩阵称为 l 步转移概率矩阵

$$P^{(l)} = (p_{j|i}^{(l)}) = \begin{pmatrix} p_{1|1}^{(l)} & p_{2|1}^{(l)} & \cdots & p_{M|1}^{(l)} \\ p_{1|2}^{(l)} & p_{2|2}^{(l)} & \cdots & p_{M|2}^{(l)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1|M}^{(l)} & p_{2|M}^{(l)} & \cdots & p_{M|M}^{(l)} \end{pmatrix}.$$

练习：

证明：对齐次马氏信源，一步转移概率矩阵与 l 步转移概率矩阵之间有关系 $P^{(l)} = P^{l}$..

可以先证 $l=2$

定义：

可遍历

将 k 阶马氏信源的状态转移用有向图来表示，称为**状态转移图或仙农图**，每个状态对应一个结点，如果一步转移概率 $p_{j|i} > 0$，则称从状态 s_i 到状态 s_j 是**一步可达的**，并用一个带权 $p_{j|i}$ 的有向边来表示；如果 $\exists l \geq 1, p_{j|i}^{(l)} > 0$，则称从状态 s_i 到状态 s_j 是 l 步**可达的**；如果齐次马氏链的任意两个状态都是可达的，则这个齐次马氏信源是**可遍历**的。

例题 2.4.1：

设一阶马氏信源的字符集为 $\mathcal{X} = \{0, 1, 2\}$，状态转移矩阵为

$$\begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0 & 1/3 & 2/3 \end{pmatrix},$$

试确定这个马氏信源的状态，状态转移图

解：

由于是一阶马氏信源，故它的状态为 $s_1=0, s_2=1, s_3=2$，并且

$$
\begin{aligned}
&p(s_1|s_1)=0.5, p(s_2|s_1)=0.5, p(s_3|s_1)=0,\\
&p(s_1|s_2)=0.5, p(s_2|s_2)=0.25, p(s_3|s_2)=0.25,\ ,\\
&p(s_1|s_3)=0, p(s_2|s_3)=1/3, p(s_3|s_3)=2/3,
\end{aligned}
$$

其转移图为图 2-3。

图 2-3:

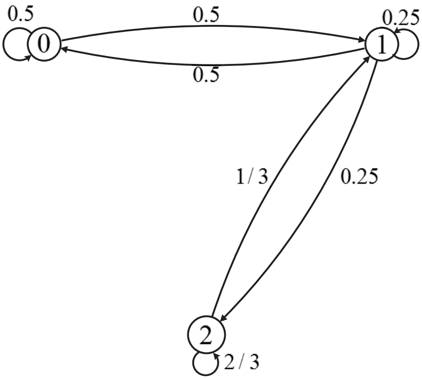


Figure: 例题 2.4.1 图 2-3

例题 2.4.2：

设一阶马氏信源的字符集为 $\mathcal{X} = 0, 1, 2, 3$，状态转移矩阵为

$$\begin{pmatrix} 0 & 0 & 0.5 & 0.5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

试确定这个马氏信源的状态，状态转移图。

解：

由于是一阶马氏信源，故它的状态为
$s_1=0, s_2=1, s_3=2, s_4=3$，并且

$$
\begin{aligned}
&p(s_1|s_1)=0, p(s_2|s_1)=0, p(s_3|s_1)=0.5, p(s_4|s_1)=0.5,\\
&p(s_1|s_2)=1, p(s_2|s_2)=0, p(s_3|s_2)=0, p(s_4|s_2)=0,\\
&p(s_1|s_3)=0, p(s_2|s_3)=1, p(s_3|s_3)=0, p(s_4|s_3)=0,\\
&p(s_1|s_4)=0, p(s_2|s_4)=1, p(s_3|s_4)=0, p(s_4|s_4)=0,
\end{aligned}
$$

其转移图为图 2-4。

图 2-4：

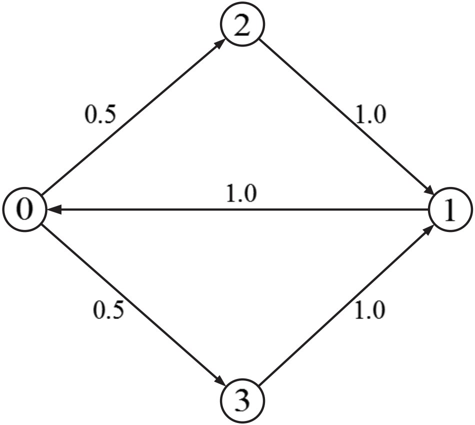


Figure: 例题 2.4.1 图 2-4

例题 2.4.3:

一个二阶马氏信源的字符集为 $\mathcal{X} = 0, 1$, 状态转移概率为 $p(0|00) = p(1|11) = 0.8, p(1|00) = p(0|11) = 0.2, p(0|01) = p(0|10) = p(1|01) = p(1|10) = 0.5$，试确定这个马氏信源的状态，状态转移矩阵，状态转移图。

解：

由于是二阶马氏信源，故它的状态为
$s_1=00, s_2=01, s_3=10, s_4=11$，并且

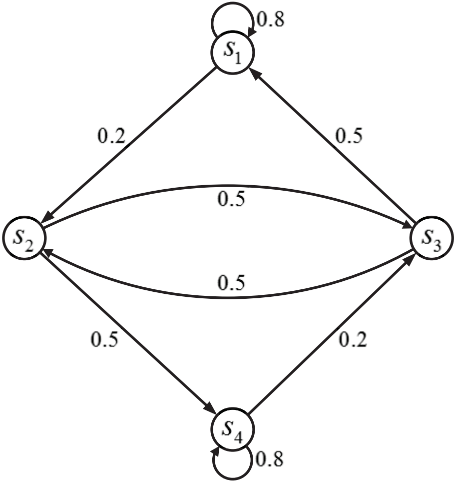
$$\begin{aligned}
&p(s_1|s_1)=0.8, p(s_2|s_1)=0.2, p(s_3|s_1)=0, p(s_4|s_1)=0,\\
&p(s_1|s_2)=0, p(s_2|s_2)=0, p(s_3|s_2)=0.5, p(s_4|s_2)=0.5,\\
&p(s_1|s_3)=0.5, p(s_2|s_3)=0.5, p(s_3|s_3)=0, p(s_4|s_3)=0,\\
&p(s_1|s_4)=0, p(s_2|s_4)=0, p(s_3|s_4)=0.2, p(s_4|s_4)=0.8,
\end{aligned}$$

故状态矩阵为

$$\begin{pmatrix} 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix},$$

其转移图为图 2-5。

图 2-5:



例题 2.4.4：

一个一阶马氏信源的字符集为 $\mathcal{X} = 1, 2, 3, 4$，状态转移图为（??）。试确定这个马氏信源的状态，状态转移矩

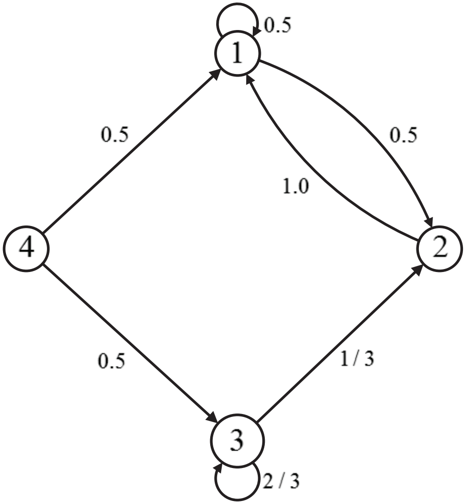
2-6

解：

由于是一阶马氏信源，故每个字符就是它的一种状态，其转移矩阵为

$$\begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix}.$$

图 2-6：



所有状态的联合分布：

设 $\pi(n) = \begin{pmatrix} \pi_1(n) & \pi_2(n) & \cdots & \pi_M(n) \end{pmatrix}$ 表示 k 阶马氏信源在第 n 时刻时所有状态的概率分布，其中 $\pi_i(n)$ 表示第 n 时刻时信源处于第 i 个状态 s_i 时的概率，则在第 $n+1$ 时刻信源的各个状态的概率分布为

2-5

$$\pi(n+1) = \pi(n)P \quad \text{或} \quad \pi_j(n+1) = \sum_{i=1}^{M} \pi_i(n) p_{j|i}, \qquad \cancel{(5.1)}$$

其中 P 为一步转移概率矩阵。事实上：因为第 $n+1$ 时刻的状态 s_j 可以由第 n 个时刻的任何一种状态转移得到，故由全概率公式得

$$\pi_j(n+1) = p(s_j) = \sum_{i=1}^{M} p(s_i)p(s_j|s_i) = \sum_{i=1}^{M} \pi_i(n) p_{j|i}.$$

定义：

平稳分布

若成立极限

$$\lim_{n\to\infty}\pi_j(n)=\pi_j\geq 0, j=1,2,\cdots,M \quad 并且 \quad \sum_{j=1}^{M}\pi_j=1,$$

称分布 $\pi=(\pi_1,\pi_2,\cdots,\pi_M)$ 为 k 阶马氏信源的平稳分布，对 (2-5) 式两边取极限得 $\pi P=\pi$，于是给出如下定义。

定义2.4.1

*设 $X_1,X_2,\cdots,X_n,\cdots$ 是一个 k 阶齐次马氏信源，它的一步转移概率矩阵为 M 阶矩阵 $P=(p_{j|i})$。如果存在信源状态空间中的概率分布 $\pi=(\pi_1,\pi_2,\cdots,\pi_M)$ 使 $\pi=\pi P$，则称分布 π 为这个齐次马氏信源的**平稳分布**。*

平稳分布存在性：

平稳分布表明：当马氏信源在某个时刻有状态分布 π 时，则在之后任何时刻状态的分布都为 π。事实上：如果在第 n 时刻状态分布为 π，则在第 $n+1$ 时刻状态分布为

$$p(s_j) = \sum_{i=1}^{M} p(s_j|s_i)p(s_i) = \sum_{i=1}^{M} p_{j|i}\pi_i,$$

这正是 πP 的第 j 列，由 $\pi = \pi P$ 知，也是 π 的第 j 列。

命题2.4.1

如果一个 k 阶齐次马氏信源的任意 l 步转移概率矩阵 $P^{(l)} = P^l$ 是正矩阵，则该信源是可遍历的，并且存在平稳分布。

证明参随机过程的教材。

例题 2.4.5：怎么求平稳分布

已知三个状态的马氏信源的一步转移概率矩阵如下，判断它的遍历性，并求平稳分布。

$$\begin{pmatrix} 3/4 & 1/4 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 3/4 & 1/4 \end{pmatrix}.$$

图 2-7：

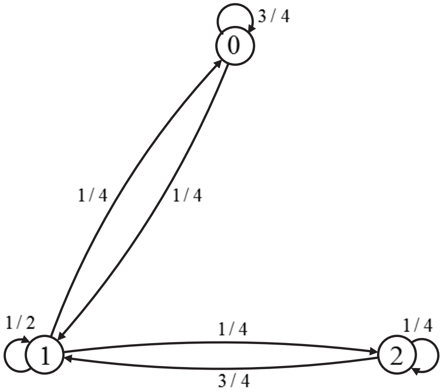


Figure: 例题 2.4.5 图 2-7

解：

（1）它的两步转移概率矩阵

$$\begin{pmatrix} 5/8 & 5/16 & 1/16 \\ 5/16 & 1/2 & 3/16 \\ 3/16 & 9/16 & 1/4 \end{pmatrix} > 0,$$

故这是一个遍历信源，如图 2-7 。

续解：

（2）设 $\pi=(\pi_1,\pi_2,\pi_3)$ 为平稳分布，由 $\pi P=\pi$ 即方程组

$$\begin{cases} \frac{3}{4}\pi_1+\frac{1}{4}\pi_2=\pi_1 \\ \frac{1}{4}\pi_1+\frac{1}{2}\pi_2+\frac{3}{4}\pi_3=\pi_2 \\ \frac{1}{4}\pi_2+\frac{1}{4}\pi_3=\pi_3 \end{cases},\begin{cases} \pi_1=\pi_2 \\ \pi_2=3\pi_3 \end{cases},$$

再利用 $\pi_1+\pi_2+\pi_3=1$ 可得平稳分布为

$$\pi=(\frac{3}{7},\frac{3}{7},\frac{1}{7})$$

$$=(0.42857142857143, 0.42857142857143, 0.14285714285714).$$

续解：

（3）数值试验：假如本题中初始状态分布为

$$X_1 \sim p(x) \sim \begin{pmatrix} s_1 & s_2 & s_3 \\ 1/4 & 1/2 & 1/4 \end{pmatrix},$$

则在不同时刻时信源状态的分布见表 2-1。

例题 2.4.6：怎么求平稳分布

已知四个状态的马氏信源的一步转移概率矩阵如下，求判断它的遍历性，并求平稳分布。

$$P=\begin{pmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix}.$$

图 2-8:

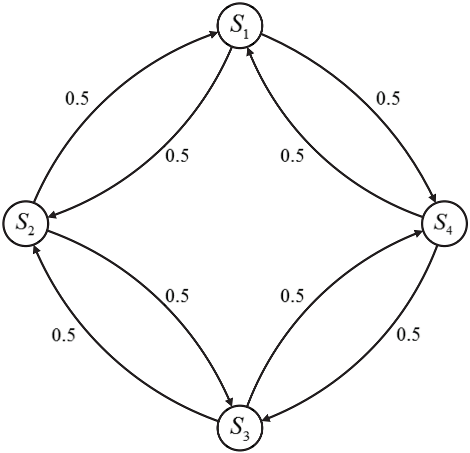


Figure: 例题 2.4.6 图 2-8

解：

可以计算（1）当 l 为奇数时，l 步转移概率矩阵为 $P^l = P$。
（2）当 l 为偶数时，l 步转移概率矩阵当为

$$P^l = \begin{pmatrix} 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \end{pmatrix},$$

虽然 P^l 不大于 0，但根据它的状态转移图 2-8，可知这是一个遍历信源。

续解：

（3）设 $\pi=(\pi_1,\pi_2,\pi_3,\pi_4)$ 为平稳分布，由 $\pi P=\pi$ 即方程组

$$\begin{cases} 0.5\pi_2+0.5\pi_4=\pi_1 \\ 0.5\pi_1+0.5\pi_3=\pi_2 \\ 0.5\pi_2+0.2\pi_4=\pi_3 \\ 0.5\pi_1+0.5\pi_3=\pi_4 \end{cases},\pi_1=\pi_2=\pi_3=\pi_4,$$

再利用 $\pi_1+\pi_2+\pi_3+\pi_4=1$ 可得平稳分布为 $\pi=(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4})$。

定理 2.4.1：马氏信源熵率公式

一旦 k 阶马氏信源经过很长时间达到平稳状态后，联合概率（2-4）式不随时间变化而改变，它就变成了平稳信源，这时就可以用平稳信源的熵率公式求熵率。

(定理 2.4.1)

对 k 阶齐次马氏信源 $X_1, X_2, \cdots, X_n, \cdots$ *的熵率是*

$$H_\infty(X) = \sum_{i=1}^{M} \pi_i H(X|s_i), \qquad (\cancel{6.1})\ 2\text{-}6$$

其中 $H(X|s_i), i = 1, 2, \cdots, M$ *是状态转移概率矩阵中每种状态下的条件熵，* π *为平稳分布。*

证明：

$$
\begin{aligned}
H_\infty(X) &= \lim_{n\to\infty} H(X_n|X_1X_2\cdots X_{n-1}) \\
&= \lim_{n\to\infty} H(X_n|X_{n-k}\cdots X_{n-1}) \\
&= \lim_{n\to\infty} \sum_{i=1}^{M} \pi_i(n) H(X|s_i) \\
&= \sum_{i=1}^{M} \left[\lim_{n\to\infty} \pi_i(n)\right] H(X|s_i) \\
&= \sum_{i=1}^{M} \pi_i H(X|s_i).
\end{aligned}
$$

根据定理，需要由状态转移概率矩阵求出信源的平稳分布后才能求出熵率。

例题 2.4.7

一个二阶齐次马氏信源字符空间为 $\mathcal{X}=\{0,1\}$，转移概率矩阵如下，试求它的四个状态、状态转移图、平稳分布和熵率。

$$\begin{pmatrix} 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix}$$

解：

（1）四个状态为：$s_1 = 00, s_2 = 01, s_3 = 10, s_4 = 11$。
（2）状态转移图为图 2-5。
（3）设 $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$ 为平稳分布，由 $\pi P = \pi$ 即方程组

$$\begin{cases} 0.8\pi_1 + 0.5\pi_3 = \pi_1 \\ 0.2\pi_1 + 0.5\pi_3 = \pi_2 \\ 0.5\pi_2 + 0.2\pi_4 = \pi_3 \\ 0.5\pi_2 + 0.8\pi_4 = \pi_4 \end{cases}, \begin{cases} \pi_1 = \pi_4 \\ \pi_2 = \pi_3 = \frac{2}{5}\pi_4 \end{cases},$$

再利用 $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$ 可得平稳分布为
$\pi = (\frac{5}{14}, \frac{2}{14}, \frac{2}{14}, \frac{5}{14})$。

续解：

（4）为了求熵率，先求出每个状态下的条件熵：

$$\begin{aligned}
H(X|s_1) &= -(0.8\ln 0.8 + 0.2\ln 0.2) = 0.5004 \text{ nats},\\
H(X|s_2) &= -(0.5\ln 0.5 + 0.5\ln 0.5) = \ln 2 = 0.6931 \text{ nats},\\
H(X|s_3) &= -(0.5\ln 0.5 + 0.5\ln 0.5) = \ln 2 = 0.6931 \text{ nats},\\
H(X|s_4) &= -(0.2\ln 0.2 + 0.8\ln 0.8) = 0.5004 \text{ nats},
\end{aligned}$$

从而熵率为：

$$\begin{aligned}
H_\infty(X) &= \pi_1 H(X|s_1) + \pi_2 H(X|s_2) + \pi_3 H(X|s_3) + \pi_4 H(X|s_4)\\
&= 0.5556 \text{ nats}.
\end{aligned}$$