01 离散信源与信息熵

- 自信息、平均信息量与信息熵
- 杂项式计算
- 马尔可夫信源
- 冗余度

常考题库

傅叶云：2.05 2.17 2.18 2.19 2.20
岳殿武：2.11

INFORMATION
THEORY AND CODING

01 自信息、平均信息量与信息熵

$$\boxed{\begin{bmatrix} X \\ P(x) \end{bmatrix}} = \begin{bmatrix} x_1 & x_2 & x_n \\ P(x_1) & P(x_2) & P(x_3) \end{bmatrix}$$

自信息

概率

信源

解题类 1-1 对某一信源求自信息数目

Step 1：使用公式求出每个自信息的信息量

$$I_{(x_i)} = \log \frac{1}{P(x_i)}$$

Step 2：数出消息中每个符号存在的个数为n_i

Step 3：求出自信息为每个自信息的信息量与个数相乘再求和，即为自信息I

$$I = \sum n_i I_{(x_i)}$$

题型解题引导 信息论与编码(B)

@GhostKING学长

01 自信息、平均信息量与信息熵

傅叶云 | 2.05 信息论与编码(B)

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} a_1 = 0 & a_2 = 1 & a_3 = 2 & a_4 = 3 \\ 3/8 & 1/4 & 1/4 & 1/8 \end{bmatrix}$$

设离散无记忆信源发出的消息为(2021 2013 0213 0012 0321 0110 3210 1002 1032 0112 2321 0),求：

(1)此消息的自信息是多少?

(2)在此消息中平均每个符号携带的信息量是多少?

解：(1) 每个符号的信息量为

$$I(a_1 = 0) = \log\frac{8}{3} = 1.415\ bit$$

$$I(a_2 = 1) = \log 4 = 2\ bit$$

$$I(a_3 = 2) = \log 4 = 2\ bit$$

$$I(a_4 = 3) = \log 8 = 3\ bit$$

在发出的消息中，共有 14 个“0”符号，13 个“1”符号，12 个“2”符号，6 个“3”符号，则得到消息的自信息为：

$$I = 14 \times 1.415 + 13 \times 2 + 12 \times 2 + 6 \times 3 = 87.81\ bit$$

(2) 此消息平均每个符号携带的信息量为

$$\bar{I} = \frac{I}{N} = \frac{87.81}{45} = 1.95\ bit/symbol$$

01 自信息、平均信息量与信息熵

岳殿武 | 2.11 信息论与编码(B)

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} a_1 = 0 & a_2 = 1 & a_3 = 2 & a_4 = 3 \\ 3/8 & 1/4 & 1/4 & 1/8 \end{bmatrix}$$

求该信源产生60个符号构成的消息的信息熵。

解： 信源的信息熵为

$$H(X) = \sum_{i=1}^{n} P(x_i)[log_2 \frac{1}{P(x)}]$$

$$= \frac{3}{8}\log\frac{8}{3} + 2 \times \frac{1}{4}\log 4 + \frac{1}{8}\log 8$$

$$= 1.9056\ bit/symbol$$

60个符号构成的消息的信息熵为

$$60H(X) = 60 \times 1.9056 = 114.34\ bit$$

02 杂项式计算

解题类 1-4 计算$H^k(X)$，$H_k(X)$，$H(X_L|X_{L-1}X_{L-2}\dots)$，$\lim\limits_{N\to\infty} H_k(X)$

Step 1：求解$H(X)$

Step 2：公式求解

$$H^k(X) = k \cdot H(x)$$

$$H_k(X) = \frac{1}{k} H^k(X)$$

$$H(X_L|X_{L-1}X_{L-2}\dots) = H(X)$$

$$\lim_{N\to\infty} H_k(X) = H(X)$$

@GhostKING学长

02 杂项式计算

设问类 2 列出X^k信源所有可能信号

即每个信号由 k 个位置可以使用自信息进行填充排序，列出所有组合形式即可。

02 杂项式计算

傅叶云 | 2.17 信息论与编码(B)

设有一个信源，它产生 0、1 序列的消息。它在任意时间而且不论以前输出过什么符号，均按$P(0)=0.4$，$P(1)=0.6$的概率输出符号。

1) 试问这个信源是否是平稳的？
2) 试计算$H(X^2)$，$H(X_3|X_1X_2)$及 $\lim[N\to\infty]\ H_N(X)$；
3) 试计算$H(X^4)$并写出X^4信源中可能有的所有符号。

解：(1) 信源任意时间发出自信息的概率与时间无关，即为平稳信源。

(2) $$H(X)=\sum_{i=1}^{n}P(x_i)[log_2\frac{1}{P(x)}]$$

$$=0.4\log\frac{1}{0.4}+0.6\log\frac{1}{0.6}$$

$$=0.971\ bit/symbol$$

$$H(X^2)=2H(X)=1.924\ bit/symbol$$

$$H(X_3|X_2X_1)=H(X)=0.971\ bit/symbol$$

$$\lim_{N\to\infty}H_N(X)=H(X)=0.971\ bit/symbol$$

(3) $$H(X^4)=4H(X)=3.884\ bit/symbol$$

0000	0100	1000	1100
0001	0101	1001	1101
0010	0110	1010	1110
0011	0111	1011	1111

03 马尔可夫信源

解题类 2-1 绘制马尔可夫信源状态转移图

Step 1：以题目中"当x_i为 A 时，为 B 的概率为 p"做箭头指向绘图 A→B，并标记为 b：p

Step 2：对于题目给出 P(B|A)=p，则做箭头指向绘图 A→B，并标记为 b：p

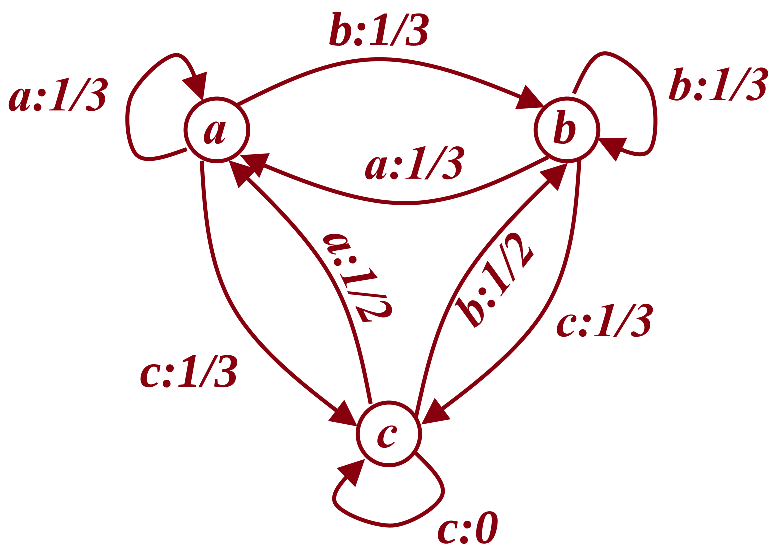
题型解题引导 信息论与编码(B)

@GhostKING学长

03 马尔可夫信源

傅叶云 | 2.18 信息论与编码(B)

设有一个信源，它在开始时以$P(a)=0.6$, $P(b)=0.3$, $P(c)=0.1$的概率输出X_1。如果X_1为 a 时，则X_2为 a、b、c 的概率为 1/3；如果X_1为 b，则X_2为 a、b、c 的概率为 1/3；如果X_1为 c，则X_2为 a、b 的概率为 1/2，为 c 的概率为 0。而且后面输出X_i的概率只与X_i-1有关，又$P(X_i|X_i-1)=P(X_2|X_1)\ i\geq 3$。试利用马尔可夫信源的图示法画出状态转移图，并计算信息熵H_∞。



03 马尔可夫信源

解题类 2-2 根据马尔可夫信源状态转移图求信息熵H_∞

Step 1：观察对于每个自信息x_i，有哪些箭头从另一个自信息$\overleftarrow{x_\iota}$（包括它自己本身，对应着标记的概率 p）指向它，列出方程组，每个方程组为

$$Q(x_i)=\sum p_i Q(\overleftarrow{x_\iota})=p_1Q(\overleftarrow{x_1})+p_2Q(\overleftarrow{x_2})+\cdots+p_iQ(\overleftarrow{x_\iota})$$

注意：x_i和$\overleftarrow{x_\iota}$在解题时要填具体的字母噢！

Step 2：补一条方程

$$Q(x_1)+Q(x_2)+\cdots+Q(x_i)=1$$

Step 3：求解出方程组的各个$Q(x_i)$

Step 4：将方程组中，同一字母前的概率（包括相同的也要写，有几个写几个）写入$H(X_i|\ldots)$的…中，逗号隔开，并计算出结果

$$H(X_i|P_1,P_2,P_n)=\sum P_n\log\frac{1}{P_n}$$

Step 5：求解信息熵H_∞为各个$Q(x_i)$与$H(X_i|\ldots)$的乘积求和

$$H_\infty=\sum Q(x_i)H(X_i|\ldots)$$

题型解题引导

信息论与编码(B)

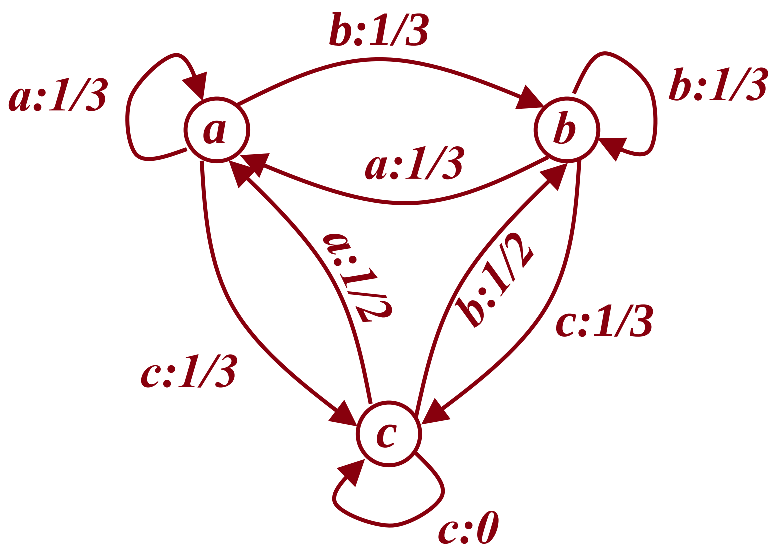
@GhostKING学长

03 马尔可夫信源

傅叶云 | 2.18

信息论与编码(B)

设有一个信源，它在开始时以$P(a)=0.6$, $P(b)=0.3$, $P(c)=0.1$的概率输出X_1。如果X_1为 a 时，则X_2为 a、b、c 的概率为 1/3；如果X_1为 b，则X_2为 a、b、c 的概率为 1/3；如果X_1为 c，则X_2为 a、b 的概率为 1/2，为 c 的概率为 0。而且后面输出X_i的概率只与X_i-1有关，又$P(X_i|X_i-1)=P(X_2|X_1)$ $i\geq 3$。试利用马尔可夫信源的图示法画出状态转移图，并计算信息熵H_∞。



$$Q(a)=\frac{1}{3}Q(a)+\frac{1}{3}Q(b)+\frac{1}{2}Q(c)$$

$$Q(b)=\frac{1}{3}Q(a)+\frac{1}{3}Q(b)+\frac{1}{2}Q(c)$$

$$Q(c)=\frac{1}{3}Q(a)+\frac{1}{3}Q(b)$$

$$Q(a)+Q(b)+Q(c)=1$$

$$Q(a)=Q(b)=\frac{3}{8}\qquad Q(c)=\frac{1}{4}$$

03 马尔可夫信源

傅叶云 | 2.18

信息论与编码(B)

$$Q(a)=\frac{1}{3}Q(a)+\frac{1}{3}Q(b)+\frac{1}{2}Q(c)$$

$$Q(b)=\frac{1}{3}Q(a)+\frac{1}{3}Q(b)+\frac{1}{2}Q(c)$$

$$Q(c)=\frac{1}{3}Q(a)+\frac{1}{3}Q(b)$$

$$Q(a)+Q(b)+Q(c)=1$$

$$Q(a)=Q(b)=\frac{3}{8}\qquad Q(c)=\frac{1}{4}$$

$$H\left(a\middle|\frac{1}{3},\frac{1}{3},\frac{1}{2}\right)=2\times\frac{1}{3}\log 3+\frac{1}{2}\log 2=1.5566$$

$$H\left(b\middle|\frac{1}{3},\frac{1}{3},\frac{1}{2}\right)=2\times\frac{1}{3}\log 3+\frac{1}{2}\log 2=1.5566$$

$$H\left(c\middle|\frac{1}{3},\frac{1}{3}\right)=2\times\frac{1}{3}\log 3=1.0566$$

$$H_\infty=\sum Q(E_i)H(X|\,E_i)=1.439\ bit/symbol$$

03 马尔可夫信源

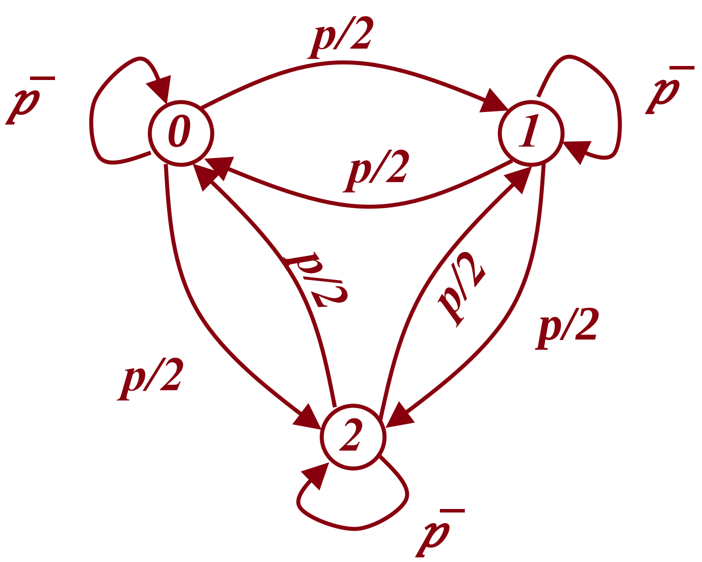
傅叶云 | 2.19 信息论与编码(B)

一阶马尔可夫信源的状态图如图所示，信源X的符号集为$\{0,1,2\}$并定义$\bar{p}=1-p$。

1) 求信源平稳后的概率分布$P(0)$、$P(1)$和$P(2)$；

2) 求此信源的熵；

3) 近似认为此信源为无记忆时，符号的概率分布等于平稳分布。求近似信源的熵$H(X)$并与$H\infty$进行比较。



$$Q(0)=\bar{p}Q(0)+\frac{p}{2}Q(1)+\frac{p}{2}Q(2)$$

$$Q(1)=\frac{p}{2}Q(0)+\bar{p}Q(1)+\frac{p}{2}Q(2)$$

$$Q(2)=\frac{p}{2}Q(1)+\frac{p}{2}Q(1)+\bar{p}Q(2)$$

$$Q(0)+Q(1)+Q(2)=1$$

$$Q(0)=Q(1)=Q(2)=\frac{1}{3} \qquad P(0)=P(1)=P(2)=\frac{1}{3}$$

题型解题引导 信息论与编码(B)

@GhostKING学长

03 马尔可夫信源

傅叶云 | 2.19 信息论与编码(B)

一阶马尔可夫信源的状态图如图所示，信源X的符号集为$\{0,1,2\}$并定义$\bar{p}=1-p$。

1) 求信源平稳后的概率分布$P(0)$、$P(1)$和$P(2)$；

2) 求此信源的熵；

3) 近似认为此信源为无记忆时，符号的概率分布等于平稳分布。求近似信源的熵$H(X)$并与$H\infty$进行比较。

$$Q(0)=\bar{p}Q(0)+\frac{p}{2}Q(1)+\frac{p}{2}Q(2)$$

$$Q(1)=\frac{p}{2}Q(0)+\bar{p}Q(1)+\frac{p}{2}Q(2)$$

$$Q(2)=\frac{p}{2}Q(1)+\frac{p}{2}Q(1)+\bar{p}Q(2)$$

$$Q(0)+Q(1)+Q(2)=1$$

$$Q(0)=Q(1)=Q(2)=\frac{1}{3}$$

$$H\left(0\middle|\bar{p},\frac{p}{2},\frac{p}{2}\right)=\bar{p}\times\frac{p}{2}+2\log\frac{2}{p}$$

$$H\left(1\middle|\bar{p},\frac{p}{2},\frac{p}{2}\right)=\bar{p}\times\frac{p}{2}+2\log\frac{2}{p}$$

$$H\left(2\middle|\bar{p},\frac{p}{2},\frac{p}{2}\right)=\bar{p}\times\frac{p}{2}+2\log\frac{2}{p}$$

$$H_\infty=\sum Q(E_i)H(X|\,E_i)=\bar{p}\log\frac{1}{\bar{p}}+p\log\frac{1}{p}+p\ bit/symbol$$

题型解题引导 信息论与编码(B)

@GhostKING学长

03 马尔可夫信源

傅叶云 | 2.19 信息论与编码(B)

一阶马尔可夫信源的状态图如图所示，信源X的符号集为$\{0,1,2\}$并定义$\bar{p}=1-p$。

1) 求信源平稳后的概率分布$P(0)$、$P(1)$和$P(2)$；

2) 求此信源的熵；

3) 近似认为此信源为无记忆时，符号的概率分布等于平稳分布。求近似信源的熵$H(X)$并与H_∞进行比较。

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

$$H(X) = 1.585\ bit/symbol$$

$$H_\infty = \bar{p}\log\frac{1}{\bar{p}} + p\log\frac{1}{p} + p = \frac{2}{3}\log\frac{3}{2} + \frac{1}{3}\log 3 + \frac{1}{3} = 1.251\ bit/symbol$$

$$H(X) > H_\infty$$

04 冗余度

傅叶云 | 2.20 信息论与编码(B)

黑白气象传真图的消息只有黑色和白色两种，即信源 X={黑,白}，设黑色出现的概率为 P(黑)=0.3，白色出现的概率 P(白)=0.7。

1) 假设图上黑白消息出现前后没有关联，求熵$H(X)$；

2) 假设消息出现前后有关联，其依赖关系为 P(白|白)=0.9，P(黑|白)=0.1，P(白|,黑)=0.2，P(黑|黑)=0.8，求此一阶马尔可夫信源的熵；

3) 分别求上述两种信源的剩余度。

如果出现黑白消息前后没有关联，信息熵为：

$$H(X) = -\sum p_i \log p_i = 0.881 \text{ 比特/符号}$$

题型解题引导 信息论与编码(B)

@GhostKING学长

04 冗余度

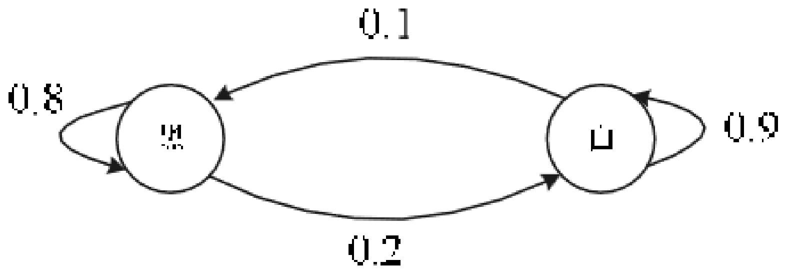
傅叶云 | 2.20 信息论与编码(B)

黑白气象传真图的消息只有黑色和白色两种，即信源X={黑,白}，设黑色出现的概率为P(黑)=0.3，白色出现的概率P(白)=0.7。

1) 假设图上黑白消息出现前后没有关联，求熵$H(X)$；

2) 假设消息出现前后有关联，其依赖关系为P(白|白)=0.9，P(黑|白)=0.1，P(白|,黑)=0.2，P(黑|黑)=0.8，求此一阶马尔可夫信源的熵；

3) 分别求上述两种信源的剩余度。



设黑白两个状态的极限概率为Q(黑)和Q(白)，根据切普曼—柯尔莫哥洛夫方程可得：

$$\begin{cases} Q(黑) = 0.8Q(黑) + 0.1Q(白) \\ Q(白) = 0.2Q(黑) + 0.9Q(白) \\ Q(黑) + Q(白) = 1 \end{cases}$$

解得：

$$Q(黑) = \frac{1}{3}, \quad Q(白) = \frac{2}{3}$$

此信源的信息熵为：

$$H_\infty = \sum Q(E_i)H(X \mid E_i) = 0.553 \text{比特/符号}$$

傅叶云 | 2.20

信息论
与编码(B)

解题类 2-6 求冗余度

对于 q 元信息，自信息的冗余度：

$$\gamma = 1 - \frac{H(X)}{\log q}$$

对于 q 元信息，马尔可夫信源的冗余度：

$$\gamma = 1 - \frac{H_\infty}{\log q}$$

04 冗余度

傅叶云 | 2.20 信息论与编码(B)

黑白气象传真图的消息只有黑色和白色两种，即信源 X={黑,白}，设黑色出现的概率为 P(黑)=0.3，白色出现的概率 P(白)=0.7。

1) 假设图上黑白消息出现前后没有关联，求熵$H(X)$；

2) 假设消息出现前后有关联，其依赖关系为 P(白|白)=0.9，P(黑|白)=0.1，P(白|,黑)=0.2，P(黑|黑)=0.8，求此一阶马尔可夫信源的熵；

3) 分别求上述两种信源的剩余度。

两信源的冗余度分别为：

$$\gamma_1 = 1 - \frac{H(X)}{\log 2} = 0.119$$

$$\gamma_1 = 1 - \frac{H_\infty}{\log 2} = 0.447$$