计算题整理

2024年1月6日 11:41

熵

熵的可加性

2.1 A 村有一半人说真话，3/10 人总说假话，2/10 人拒绝回答；B 村有3/10人诚实，一半人说谎，2/10 人拒绝回答。现随机地从 A 村和 B 村抽取人，p 为抽到 A 村人的概率，$1-p$ 为抽到 B 村人的概率，问通过测试某人说话的状态平均能获得多少关于该人属于哪个村的信息？通过改变 p，求出该信息的最大值。

记X为村庄，Y为说话的状态。

> 问说话状态平均能获得多少关于村庄的信息量，是对说话状态的平均而非信息量的平均

采用熵的可加性，$H(Y|X) = P(A)H(Y|X = A) + P(B)H(Y|X = B) = (p + 1 - p)H(0.5,0.3,0.2) = 1.49\ bits$

$I(X;Y) = H(Y) - (Y|X) = H(0.3 + 0.2p, 0.5 - 0.2p, 0.2) - 1.49$

当且仅当$0.3 + 0.2p = 0.5 - 0.2p,\ p = 0.5$时，$I(X;Y) = H(0.4,0.4,0.2) - 1.49 = 0.032\ bit$最大。

马尔科夫链

$I(X;Z|Y) = 0$

2.4 随机掷 3 颗骰子，以 X 表示第一颗骰子抛掷的结果，以 Y 表示第一颗和第二颗骰子抛掷之和，以 Z 表示 3 颗骰子的点数之和，试求 $H(X|Y)$，$H(Y|X)$，$H(Z|X,Y)$，$H(X,Z|Y)$ 和 $H(Z|X)$。

马尔科夫，令$Y = X_1 + X_2$，$Z = X_1 + X_2 + X_3$

$H(Y|X) = H(X_1 + X_2|X_1) = H(X_2) = H(X) = log6$

同理，$H(Z|X,Y) = H(X) = log6$

$H(Y) = \cdots = 3.274\ bit$

$H(X|Y) = H(X) - H(Y) + H(Y|X) = 1.90\ bit$

$H(Z|X) = H(Y) = 3.274\ bit$

$H(X,Z|Y) = H(X|Y) + H(Z|Y) = 4.48\ bit$

熵证明题

2.9 若3个随机变量 X,Y,Z，有 $X+Y=Z$ 成立，其中 X 和 Y 独立，试证

(a) $H(X)\leqslant H(Z)$

(b) $H(Y)\leqslant H(Z)$

(c) $H(X,Y)\geqslant H(Z)$

(d) $I(X;Z)=H(Z)-H(Y)$

(e) $I(X,Y;Z)=H(Z)$

(f) $I(X;Y,Z)=H(X)$

(g) $I(Y;Z|X)=H(Y)$

(h) $I(X;Y|Z)=H(X|Z)=H(Y|Z)$

发现每题Venn图特殊的点。

比如此题$H(Z|X)=H(Y|X),\quad H(Z|Y)=H(X|Y),\quad I(X;Y)=0$

2.12 证明下列关系

(a) $H(Y,Z|X)\leqslant H(Y|X)+H(Z|X)$ 等号成立的充要条件为对一切 i,j,k 有 $p(y_j,z_k|x_i)=p(y_j|x_i)p(z_k|x_i)$ 成立。

(b) $H(Y,Z|X)=H(Y|X)+H(Z|X,Y)$。

(c) $H(Z|X,Y)\leqslant H(Z|Y)$ 等号成立的充要条件为对一切 i,j,k 有 $p(y_j,z_k,|x_i)=p(y_j|x_i)p(z_k|x_i)$ 成立。

(a) $H(Y,Z|X)=H(Y|X)+H(Z|X)-I(Y;Z|X)\Rightarrow I(Y;Z|X)=0$

马尔科夫信源计算

2.21 1个二状态平稳马尔可夫信源，输出为 $\cdots,X_0,X_1,X_2,\cdots$，信源状态转移矩阵为 $\boldsymbol{P}=\begin{bmatrix}\frac{1}{10} & \frac{9}{10}\\ \frac{3}{5} & \frac{2}{5}\end{bmatrix}$，对任何 n 求 $\frac{1}{n}H(X_1X_2\cdots X_n)$ 和 $H(X_n|X_1X_2\cdots X_{N-1})$。

马尔科夫信源：

(1) 求解稳态概率$\mu_0,\ \mu_1$

(2) 得到$H(X)=H(\mu_0,\mu_1),\quad H(X_1|X_0)=\mu_0H(X|S_0)+\mu_1H(X|S_1)$

此题还要注意对n分类讨论。

信源编码

切比雪夫多项式

3.2 设一个DMS，$U=\begin{bmatrix} a_1 & a_2 \\ p(a_1)=0.25 & p(a_2)=0.75 \end{bmatrix}$，其熵为 $H(U)$。考察长度为 L 的输出序列，当 $L \geqslant L_0$ 时满足

$$P_r\left\{\left|\frac{I(u^L)}{L}-H(U)\right|>\delta\right\}\leqslant\varepsilon$$

(a) 在 $\delta=0.05$，$\varepsilon=0.1$ 时求 L_0 值。

(b) 在 $\delta=10^{-3}$，$\varepsilon=10^{-6}$ 时求 L_0 值。

(c) $$A=\left\{u^L:\ \left|\frac{I(u^L)}{L}-H(U)\right|<\delta\right\}$$

求在(a)、(b)给定的 $L=L_0$ 情况下 A 中元素数目的上、下限。

$P\{|\frac{I(U^L)}{L}-H(U)|>\delta\}\leq\frac{\sigma^2}{L\delta^2}=\epsilon$

(a)、(b) $L=\frac{\sigma^2}{\epsilon\delta^2}=\frac{\sum p(a_i)[I(a_i)-H(U)]^2}{\epsilon\delta^2}$

(c) $|A|\geq(1-\epsilon)2^{L[H(U)-\epsilon]}$, $|A|\leq 2^{L[H(U)+\epsilon]}$

后缀分解集

3.4 确定下面码是否唯一可译，若不是则构造一个模糊序列。

(a)

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
010	0001	0110	1100	00011	00110	11110	101011

(b)

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
abc	abcd	e	dba	bace	ceac	ceab	eabd

$S_0\leftrightarrow S_0$

$S_0\leftrightarrow S_1$

$S_0\leftrightarrow S_2$

以此类推

Huffman编码

3.6 令DMS为

$$U=\begin{Bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} \\ 0.16 & 0.14 & 0.13 & 0.12 & 0.1 & 0.09 & 0.08 & 0.07 & 0.06 & 0.05 \end{Bmatrix}$$

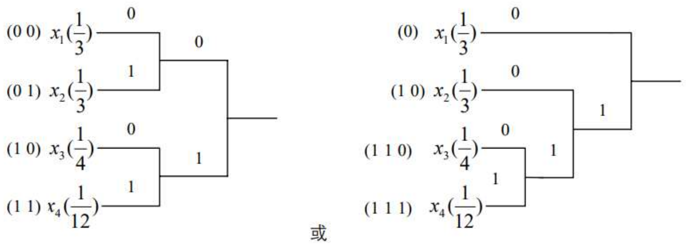
(a) 求二元Huffman码，计算 $\bar{n}$ 和 η。

(b) 求三元Huffman码，计算 $\bar{n}$ 和 η。

码字条件：$码字数=(D-1)*i+1$

效率：$\eta=\frac{H(U)}{\bar{n}\log D}$

需要注意有的编码有多种Huffman构筑方式，而平均码长相等：



Shannon编码

3.12 令 DMS 为

$$U=\begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_K \\ p(a_1)\geq & p(a_2)\geq & p(a_3)\geq & \dots & \geq p(a_K) \end{bmatrix}$$

定义 $\begin{cases} Q_i=\sum_{k=1}^{i-1} p(a_k) & i>1 \\ Q_1=0 \end{cases}$

按如下的方法进行编码，消息 a_k 用 Q_i 的二进制小数展开的截短来表示，截短的长度为保留小数点后 l_k 位，$l_k=\left\lceil \log\frac{1}{p(a_k)} \right\rceil$，$\lceil x \rceil$ 表示大于，等于 x 的最小整数。

(a) 对 DMS

$$U=\begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} \end{bmatrix}$$

进行上述编码。

(b) 证明上述编码满足异字头条件，且平均码长 $\bar{n}$ 满足 $H(U)\leq \bar{n} \leq H(U)+1$

（b）证明编码满足异字头条件：

反证法：

假设上述编码不符合异字头条件，则总存在a_i是a_j的码字前缀。

则$i<j,\ n_i<n_j$，且记$a_j=a_1 d_{n_i+1} \dots d_{n_j}$

$p_j=0.a_1 d_{n_i+1} \dots d_{n_j}((binary))>p_i$

矛盾。

综合编码

3.14 假定 DMS 为

$$X=\begin{pmatrix} a_1 & a_2 & \cdots & a_m \\ p_1 & p_2 & \cdots & p_m \end{pmatrix}$$

令 l_i 表示对应于消息 a_i 的二元码字的长度，C_i 表示消息 a_i 重要性的加权，于是这个码的平均代价为

$$C=\sum_{i=1}^{m} p_i l_i \cdot C_i$$

(a) 在 $\sum_{i=1}^{m} 2^{-l_i}\leq 1$ 约束下最小化 C，求出最小化 C 的值 C^* 和相应的 l_i^*，$i=1,2,\cdots,m$（这里忽略对于 l_i 是整数的限制）。

(b) 如何利用 Huffman 编码方法对所有唯一可译码进行最小化 C，这个最小化 C 记为 C_{Huffman}。

(c) 证明

$$C^* \leq C_{\text{Huffman}} \leq C^* + \sum_{i=1}^{m} p_i c_i$$

(a)

$$J=\sum p_i\, c_i l_i+\lambda \sum 2^{-l_i}$$

$$\frac{\partial J}{\partial l_i}=p_i c_i-\frac{\lambda l_i}{ln2}2^{-l_i}=0 \Rightarrow 2^{-l_i}=\frac{p_i c_i ln2}{\lambda l_i}$$

$$\sum 2^{-l_i}=1=\frac{ln2}{\lambda l_i}\sum p_i\, c_i \Rightarrow \lambda=\frac{ln2}{l_i}\sum p_i\, c_c$$

Thus, $l_i^*=log\frac{\sum p-Ic_i}{p_i c_i}$, $\quad C^*=\sum p_i\, c_i l_i=\sum p_i\, c_i[log\frac{\sum p_i c_i}{p_i c_i}]$

(b)

加权后各消息概率$q_i = \frac{p_i c_i}{\sum p_i c_i}$

此时码长为l_i，则$\bar{n} = \sum q_i\, l_i$，　$C_{Huffman} = \bar{n} \sum c_i\, p_i$

(c)

对Huffman码，满足$H(q_i) \le \bar{n} < H(q_i) + 1$

$$H(q_i) = \sum -q_i log(q_i) = -\sum \frac{p_i c_i}{\sum p_i c_i} log(\frac{p_i c_i}{\sum p_i c_i}) = C^*/(\sum p_i\, c_i)$$

得证

算数码

3.15 已知二元信源{0,1}，$p_0 = \frac{1}{8}$，$p_1 = \frac{7}{8}$，试用算术码对以下序列进行编码

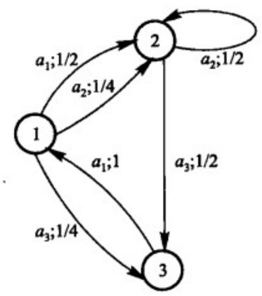
11111110111110

算数码的编码过程参考下表：

表 3.3.5　例 3.3.7 的编码过程

输入符号	$p(x^i)$	$l(x^i)$	$F(x^i)$	$C(x^i)$
	$p(\lambda)=1$	0	$F(\lambda)=0$	
1	0.11	1	0.01	0.1
1	0.1001	1	0.0111	0.1
1	0.011011	2	0.100101	0.11
0	0.00011011	4	0.100101	0.1010
1	0.0001010001	4	0.1001101011	0.1010
0	0.000001010001	6	0.1001101011	0.100111
1	0.00000011110011	7	0.10011100000001	0.1001111
1	0.0000001011011001	7	0.1001110011110111	0.1001111

马尔科夫信源编码



习题 3.16 图

(a) 求稳态下各状态概率 $q(i)$，以及各字母 a_i 的出现概率，$i=1,2,3$。

(b) 求 $H(U|s_i)$，$i=1,2,3$。　(c) 求 $H_\infty(U)$。

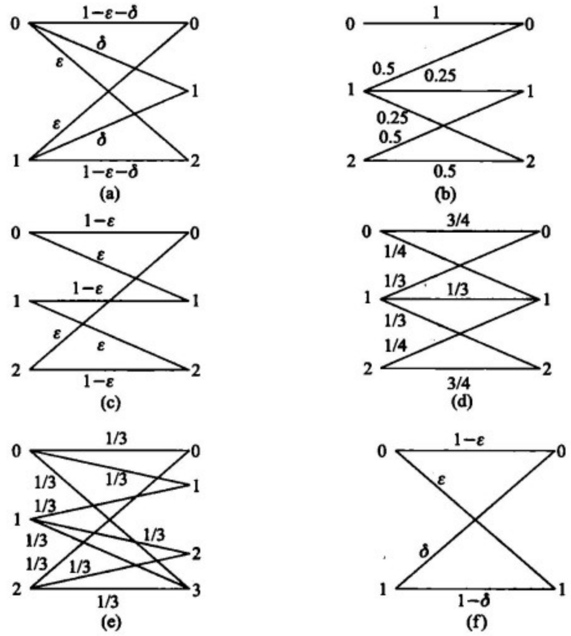
(d) 对各状态 S_j 求最佳二元码。(e) 计算平均码长。

参考总复习。

信道编码

DMC信道容量的计算

4.1 计算如习题4.1图所示离散无记忆信道的容量。



(a) 准对称信道

(b) 定理4.2.2，对输入概率矢量(Q_k)达到转移概率矩阵P的DMC信道容量满足对$\forall k, Q_k > 0$，$I(X=k;Y)=C$；$\forall k, Q_k = 0$，$I(X+k;Y)\le C$

$$P=\begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

当 $P(X=0)=P(X=2)=0.5$，$P(X)=0$时，

$$P(Y=0)=0.5,\quad P(Y=1)=0.25,\quad P(Y=2)=0.25$$

$$I(X=0;Y)=\sum_{j=0}^{2}p(j\mid 0)\log\frac{p(j\mid 0)}{p(j)}=1\ \text{bit}$$

$$I(X=2;Y)=\sum_{j=0}^{2}p(j\mid 2)\log\frac{p(j\mid 2)}{p(j)}$$

$$=0.5\log\frac{0.5}{0.25}+0.5\log\frac{0.5}{0.25}=1\ \text{bit}$$

$$I(X=1;Y)=0\le 1;$$

(c) 对称信道

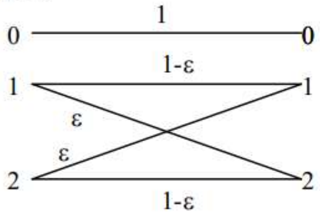
(d) 普通信道，经常取(0.5,0.5,0)概率分布。

(e) 准对称信道

(f) 常规计算

组合信道

4.2 **计算如下信道的容量，**



习题 4.2 图

和信道的信道容量 $= log(\sum 2^C)$

记得说明取得信道容量时的信道利用概率分布

带限、加性高斯白噪声信道

4.14 给定系统带宽为 W，噪声功率谱密度为 N_0，试证明传送 1 bit 信息所需最小能量为 $0.693N_0$(W)。如果要求 $\frac{P}{WN_0}>4\,000$，其中 P 为信号功率，试证明所需的信号能量至少为此最小值的 482 倍。

频带效率η下，每传1bit所需能量$E_b(\eta) \geq \frac{2^\eta-1}{\eta}N_0$

当$\eta \to 0$时，$E_{min} = ln2N_0 = 0.693N_0$

根据Shannon公式，每秒传递信息量为 $C = Wlog(1+\frac{S}{N_0W})$

设需要T_b时间传输1bit信息量，则$WT_b log(1+\frac{S}{N_0W}) = 1\ bit$

此时信号能量$E_b = S/R = ST_b = \frac{S\ bit}{Wlog(1+\frac{S}{N_0W})}$

$\frac{E_b}{E_{min}} = \frac{S}{WN_0 ln2 log(1+\frac{S}{N_0W})} \geq \frac{4000}{log(4001)ln2}$

计算理论

图灵机

[例] 假设：b表示空格

q_1表示机器的初始状态

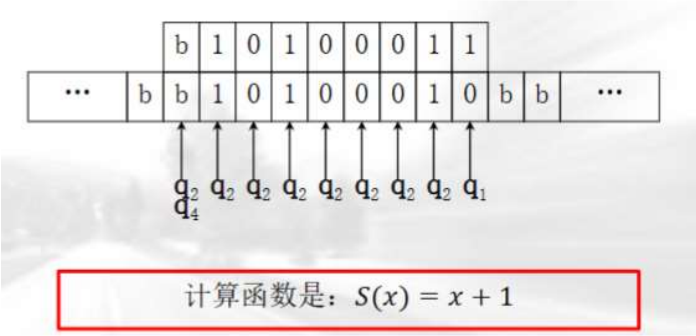
q_4表示机器的结束状态

如果带子上的输入信息为10100010，读写头位对准最右边第一个为0的方格，且状态为q_1。

按照以下五元组指令集执行后，输出正确的计算结果是什么？

指令集： $q_1$01Lq_2; $q_1$10Lq_3; $q_1$bbNq_4; $q_2$00Lq_2; $q_2$11Lq_2; $q_2$bbNq_4; $q_3$01Lq_2; $q_3$10Lq_3; $q_3$bbNq_4

指令集的前两位表示对状态和纸带数值的判断，后三位表示对数值的修改和状态转移。



K氏复杂度

1. 令x, y为任意长度的0, 1序列，证明：$K(x,y) \le K(x) + K(y) + c$

   Suppose we are given two sequences, x and y, with Kolmogorov complexity $K(x)$ and $K(y)$ respectively. Then there must exist programs p_x and p_y, of length $K(x)$ and $K(y)$ respectively, which print out x and y. Form the following program:

   Run the following two programs, not halting after the first;

   Run the program p_x , interpreting the halt as a jump to the next step;

   Run the program p_y.

2. 整数和的复杂度 证明(a) $K(n) \le logn + 2loglogn + c$；(b) $K(n_1 + n_2) \le K(n_1) + K(n_2) + c$

   (a) to describe an integer n, we will tell the computer the length of n, and then tell it the bits in n.

   (b) Given two programs to print out n_1 and n_2, we can modify them so that they write on the work tape, rather than the output tape. Then we can add an instruction to add the two numbers together and print them out.

贝叶斯判别

给定两个正太分布$p(x|C_1) \sim N(\mu_1, \sigma_1^2)$, $p(x|C_2) \sim N(\mu_2, \sigma_2^2)$，$P(C_1)$和$P(C_2)$，计算贝叶斯判别点。

$$P(C_1|x) = \frac{f_x(x|C_1)P(C_1)}{f_x(x)},\ P(C_2|x) = \frac{f_x(x|C_2)P(C_2)}{f_x(x)}$$

$$M = \frac{P(C_1|x)}{P(C_2|x)} = \frac{\sigma_2 P(C_1)}{\sigma_1 P(C_2)} e^{\frac{1}{2}(\frac{x-\mu_1}{\sigma_1})^2 - \frac{1}{2}(\frac{x-\mu_2}{\sigma_2})^2}$$

$M = 1$时的解为贝叶斯判别点。

感知器设计

1. 设计一个两输入感知器来实现布尔函数A AND (!B)

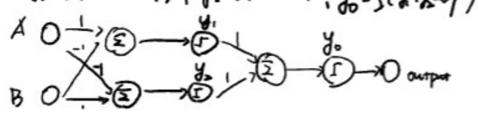
   列出真值表，设计一个线性函数做区分即可。

2. 设计一个两层感知器网络实现异或布尔函数

   对于一层实现不了的，做分解。

A	B	$A \wedge \neg B$	$\neg A \wedge B$	A XOR B
0	0	0	0	0
0	1	0	1	1
1	0	1	0	1
1	1	0	0	0

$y_1 = S(A - B - 0.7)$, $y_2 = S(B - A - 0.7)$, $y_0 = S(y_1 + y_2 - 0.7)$



机器学习

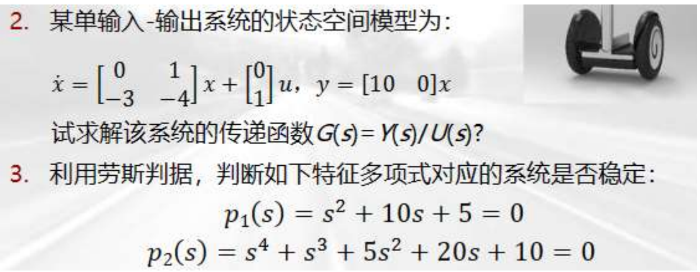
推导输出为o的单个单元的梯度下降训练法则，其中$o = w_0 + w_1x_1 + w_1x_1^2 + \cdots + w_nx_n + w_nx_n^2$

$$E(\vec{w}) = \frac{1}{2}\sum(t_d - o)^2$$

$$\frac{\partial E}{\partial w_i} = \begin{cases} -\sum(t_d - o) & i = 0 \\ -\sum(t_d - o)(x_i + x_i^2) & otherwise \end{cases}$$

$$\delta w_i = \begin{cases} \eta\sum(t_d - o) & i = 0 \\ \eta\sum(t_d - o)(x_i + x_i^2) & otherwise \end{cases}$$

控制论



非常简单，不解答。