

第二章 习题参考答案

2.1 $\ln \frac{\sigma_x^2 + \sigma_y^2}{2\sigma_x\sigma_y}$ nat 或 $\log_2 \frac{\sigma_x^2 + \sigma_y^2}{2\sigma_x\sigma_y}$ bit

2.2 证明:

$$\begin{aligned} \therefore I(X;Y|Z) &= H(X|Z) - H(X|YZ) = H(XZ) - H(Z) - H(XYZ) + H(YZ) \\ &= H(X) + H(Z|X) - H(Z) - H(XY) - H(Z|XY) + H(Y) + H(Z|Y) \\ &= [H(X) + H(Y) - H(XY)] + H(Z|X) - H(Z) - H(Z|XY) + H(Z|Y) \\ &= I(X;Y) + H(Z|X) - H(Z) - H(Z|XY) + H(Z|Y) \\ &\leq 0 + H(Z) - H(Z) - H(Z|XY) + H(Z) = H(Z) - H(Z|XY) \end{aligned}$$

$$\therefore 1 \leq H(Z) - H(Z|XY), \text{即 } H(Z) \geq 1 + H(Z|XY)$$

$$\text{又 } \because H(Z) \leq 1, H(Z|XY) \geq 0, \text{故 } H(Z) = 1, H(Z|XY) = 0$$

同理, 可推出 $H(X) = 1; H(Y) = 1;$

$$H(XYZ) = H(XY) + H(Z|XY) = H(X) + H(Y) + H(Z|XY) = 1 + 1 + 0 = 2$$

2.3 1) $H(X) = 0.918$ bit, $H(Y) = 0.918$ bit

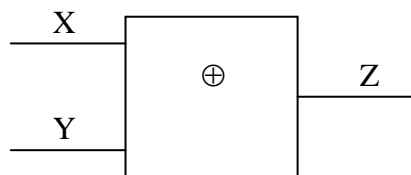
2) $H(X|Y) = \frac{2}{3}$ bit, $H(Y|X) = \frac{2}{3}$ bit, $H(X|Z) = \frac{2}{3}$ bit

3) $I(X;Y) = 0.251$ bit, $H(XYZ) = 1.585$ bit

2.4 证明: (1)根据熵的可加性, 可直接得到

$$(2) \because Y \text{ 的值取自 } (a_1, a_2, \dots, a_{k-1}), \therefore H(Y) \leq \log(k-1), \text{故原式得证}$$

2.5 考虑如下系统:



假设输入 X 、 Y 是相互独立的, 则满足 $I(X;Y) = 0$

$$\text{又 } I(X;Y|Z) = H(X|Z) - H(X|YZ) = H(X|Z) = 1 \text{ bit}$$

$$\text{不妨设 } P(Z=0) = P(Z=1) = \frac{1}{2}$$

$$\text{设 } P(X=0, Y=0|Z=0) = p \quad P(X=1, Y=1|Z=0) = 1-p$$

$$P(X=0, Y=1|Z=1) = q \quad P(X=1, Y=0|Z=1) = 1-q$$

$$\text{则 } H(X|Z) = -\frac{1}{2} [p \log p + (1-p) \log (1-p)] - \frac{1}{2} [q \log q + (1-q) \log (1-q)] = 1$$

满足上式的 p 、 q 可取: $p = \frac{1}{2}$; $q = \frac{1}{2}$

\therefore 满足条件的一个联合分布:

$$\begin{aligned} P(X=0, Y=0, Z=0) &= \frac{1}{4} & P(X=1, Y=1, Z=0) &= \frac{1}{4} \\ P(X=1, Y=1, Z=0) &= \frac{1}{4} & P(X=1, Y=0, Z=1) &= \frac{1}{4} \end{aligned}$$

2.6 解:

给出均匀分布 $p(x) = \frac{1}{b-a}$ $a \leq x \leq b$ 其中 $b-a < 1$, 则 $h(X) < 0$

2.7 证明: $I(X;Y;Z) = I(X;Y) - I(X;Y|Z)$
 $= I(X;Z) - I(X;Z|Y)$

$\because A, B$ 处理器独立, $\therefore I(X;Z|Y) = 0$

$\therefore I(X;Z) = I(X;Y) - I(X;Y|Z) \leq I(X;Y)$

等号于 $p(x/yz) = p(x)$ 下成立

2.8 $N=2$ 时, $P(0\ 0) = \frac{1}{2}$, $P(1\ 1) = \frac{1}{2}$, 其它为 0

$I(X_1; X_2) = 1$ bit

$N \neq 2$ 时,

$$\begin{aligned} & I(X_{k-1}; X_k | X_1 \cdots X_{k-2}) \quad (3 \leq k) \\ &= P(X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) I(X_{k-1}; X_k | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) \\ &+ P(X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) I(X_{k-1}; X_k | X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) \end{aligned}$$

$$P(X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{2}$$

$$P(X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{2}$$

$$P(X_{k-1}=1 | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{2}$$

$$P(X_{k-1}=0 | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{2}$$

$$P(X_k=1 | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{2}$$

$$P(X_k=0 | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{2} \quad (\text{注意, 这里 } k \leq N-1)$$

$$P(X_{k-1}=1|X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{2}$$

$$P(X_{k-1}=0|X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{2}$$

$$P(X_k=1|X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{2} \quad (\text{注意, 这里 } k \leq N-1)$$

$$P(X_k=0|X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{2}$$

$$P(X_{k-1}=0, X_k=0|X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{4}$$

$$P(X_{k-1}=0, X_k=1|X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{4}$$

$$P(X_{k-1}=1, X_k=0|X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{4}$$

$$P(X_{k-1}=1, X_k=1|X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) = \frac{1}{4}$$

$$P(X_{k-1}=0, X_k=0|X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{4}$$

$$P(X_{k-1}=0, X_k=1|X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{4}$$

$$P(X_{k-1}=1, X_k=0|X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{4}$$

$$P(X_{k-1}=1, X_k=1|X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = \frac{1}{4}$$

综上: $I(X_{k-1}; X_k | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) \quad (3 \leq k \leq N-1)$

$$\begin{aligned} &= H(X_{k-1} | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) + H(X_k | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) \\ &\quad - H(X_{k-1}; X_k | X_1 \cdots X_{k-2} \text{ 中有奇数个 } 1) \\ &= 0 \end{aligned}$$

$$I(X_{k-1}; X_k | X_1 \cdots X_{k-2} \text{ 中有偶数个 } 1) = 0$$

\therefore 当 $3 \leq k \leq N-1$ 时, $I(X_{k-1}; X_k | X_1 \cdots X_{k-2}) = 0$

当 $k = N$ 时 即

$$\begin{aligned} &I(X_{N-1}; X_N | X_1 \cdots X_{N-2}) \\ &= H(X_{N-1} | X_1 \cdots X_{N-2}) - H(X_{N-1} | X_1 \cdots X_{N-2}, X_N) \\ &= 1 \text{ bit} \end{aligned}$$

2.9 1) 实例如 2.5 题

2) 考虑随机变量 $X=Y=Z$ 的情况

$$\text{取 } P(X=0, Y=0, Z=0) = \frac{1}{2} \quad P(X=1, Y=1, Z=1) = \frac{1}{2}$$

$$\text{则 } I(X;Y|Z) = 0$$

$$I(X;Y) = 1 \quad \text{满足 } I(X;Y|Z) < I(X;Y)$$

2.10 $H(XY) \leq H(X) + H(Y)$ 等号在 X, Y 独立时取得

$$\therefore P(a_1b_1) = \frac{1}{3} \quad P(a_1b_2) = \frac{1}{12} \quad P(a_1b_3) = \frac{1}{12}$$

$$P(a_2b_1) = \frac{1}{6} \quad P(a_2b_2) = \frac{1}{24} \quad P(a_2b_3) = \frac{1}{24}$$

$$P(a_3b_1) = \frac{1}{6} \quad P(a_3b_2) = \frac{1}{24} \quad P(a_3b_3) = \frac{1}{24}$$

满足 $H(XY)$ 取最大值

2.11 证明:

$$\because p(xyz) = p(x)p(y|x)p(z|y)$$

$$\therefore I(X;Z|Y) = 0,$$

$$I(X;Y) - I(X;Y|Z) = I(X;Y;Z) = I(X;Z) - I(X;Z|Y) = I(X;Z) \geq 0$$

故 $I(X;Y) \geq I(X;Y|Z)$ 成立

2.12 证明:

$$H(XYZ) = H(XZ) + H(Y|XZ)$$

$$I(Y;Z|X) = H(Y|X) - H(Y|XZ)$$

$$\therefore H(XYZ) = H(XZ) + H(Y|X) - I(Y;Z|X)$$

2.13 证明:

$$\because I(X;Y;Z) = I(X;Y) - I(X;Y|Z)$$

$$= H(X) - H(X|Y) - H(Y|Z) + H(Y|XZ)$$

$$= H(X) - H(X|Y) - H(Y|Z) + H(XYZ) - H(XZ)$$

$$= H(XYZ) - H(X|Y) - H(Y|Z) - H(Z|X)$$

$$\text{而等式右边} = H(XYZ) - H(X) - H(Y) - H(Z)$$

$$+ H(X) - H(X|Y) + H(Y) - H(Y|Z) + H(Z) - H(Z|X)$$

$$= H(XYZ) - H(X|Y) - H(Y|Z) - H(Z|X)$$

故左式 = 右式, 原式成立

$$\mathbf{2.14} \quad P(X=n) = \left(\frac{1}{2}\right)^{n-1} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^n$$

$$H(X) = - \sum_{n=1}^{\infty} \frac{1}{2} \left(\frac{1}{2}\right)^{n-1} \log\left(\frac{1}{2}\right)^n = \sum n \left(\frac{1}{2}\right)^n = 2 \text{ bit}$$

$$2.15 \quad \frac{(\mu_1 - \mu_2)^2}{2\sigma_1^2} + \frac{\sigma_2^2 - \sigma_1^2}{2\sigma_1^2} + \log \frac{\sigma_1}{\sigma_2} \quad (\text{nat})$$

2.16 证明:

记长为 $2N$ 随机序列 $X_1 X_2 \cdots X_N X_{N+1} \cdots X_{2N}$ 中出现 a_k 的频率

$$P_{2N}(a_k) = \frac{1}{2N} \sum_{m=1}^{2N} I(X_m = a_k) = \frac{n}{2N} \text{ 的概率为 } p_k\left(\frac{n}{2N}\right), \text{ 其中 } X_1 X_2 \cdots X_N \text{ 中出现 } a_k$$

的频率为 $P_N(a_k) = \frac{1}{N} \sum_{n=1}^N I(X_n = a_k) = \frac{n_1}{N}$ 的概率为 $p_k\left(\frac{n_1}{N}\right)$, $X_{N+1} X_{N+2} \cdots X_{2N}$ 中出

现 a_k 的频率为 $P_N'(a_k) = \frac{1}{N} \sum_{n=N+1}^{2N} I(X_n = a_k) = \frac{n_2}{N} = P_N(a_k)$ 的概率为 $p_k\left(\frac{n_2}{N}\right)$, 则有

$$P_{2N}(a_k) = \frac{1}{2} P_N(a_k) + \frac{1}{2} P_N'(a_k)$$

所以

$$E\{I(P_{2N}(a_k), p(a_k))\} = E\left\{I\left(\frac{1}{2} P_N(a_k) + \frac{1}{2} P_N'(a_k), p(a_k)\right)\right\}$$

根据鉴别信息的凸性

$$I\left(\frac{1}{2} P_N(a_k) + \frac{1}{2} P_N'(a_k), p(a_k)\right) \leq \frac{1}{2} I(P_N(a_k), p(a_k)) + \frac{1}{2} I(P_N'(a_k), p(a_k))$$

$$\text{又 } \because E\left\{\frac{1}{2} I(P_N(a_k), p(a_k)) + \frac{1}{2} I(P_N'(a_k), p(a_k))\right\} = \frac{1}{2} E\{I(P_N(a_k), p(a_k))\} + \frac{1}{2} E\{I(P_N'(a_k), p(a_k))\}$$

而根据随机序列的平稳性, 有:

$$E\left\{\frac{1}{2} I(P_N(a_k), p(a_k)) + \frac{1}{2} I(P_N'(a_k), p(a_k))\right\} = E\{I(P_N(a_k), p(a_k))\} = E\{I(P_N'(a_k), p(a_k))\}$$

$$\begin{aligned} \therefore E\{I(P_{2N}(a_k), p(a_k))\} &= E\left\{I\left(\frac{1}{2} P_N(a_k) + \frac{1}{2} P_N'(a_k), p(a_k)\right)\right\} \\ &\leq E\left\{\frac{1}{2} I(P_N(a_k), p(a_k)) + \frac{1}{2} I(P_N'(a_k), p(a_k))\right\} \\ &= E\{I(P_N(a_k), p(a_k))\} \end{aligned}$$

2.17 解:

$$\because I(p_2, p_1; XY) = \iint p_2(xy) \log \frac{p_2(xy)}{p_1(xy)} dx dy ;$$

$$p_1(xy) = g(x)h(y) ;$$

$$\text{其中 } g(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) ;$$

$$h(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{y^2}{2\sigma_y^2}\right) ;$$

$$\begin{aligned} \therefore I(p_2, p_1; XY) &= \iint p_2(xy) \log \frac{p_2(xy)}{g(x)h(y)} dx dy \\ &= \iint p_2(xy) \log\left(\frac{1}{\sqrt{1-\rho^2}}\right) dx dy \\ &\quad + \iint \log e \cdot p_2(xy) \left[-\frac{1}{2(1-\rho^2)} \left(\frac{x^2}{\sigma_x^2} - 2\rho \frac{xy}{\sigma_x \sigma_y} + \frac{y^2}{\sigma_y^2} \right) + \frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] dx dy \\ &= \log\left(\frac{1}{\sqrt{1-\rho^2}}\right) - \frac{\log e}{2(1-\rho^2)} \left[\frac{\rho^2}{\sigma_x^2} E(X^2) + \frac{\rho^2}{\sigma_y^2} E(Y^2) - \frac{2\rho}{\sigma_x \sigma_y} E(XY) \right] \\ &= \log\left(\frac{1}{\sqrt{1-\rho^2}}\right) \end{aligned}$$

$$\begin{aligned} I(p_1, p_2; XY) &= \iint p_1(xy) \log \frac{p_1(xy)}{p_2(xy)} dx dy \\ &= -\log\left(\frac{1}{\sqrt{1-\rho^2}}\right) + \frac{\log e}{2(1-\rho^2)} \left[\frac{\rho^2}{\sigma_x^2} E(X^2) + \frac{\rho^2}{\sigma_y^2} E(Y^2) - \frac{2\rho}{\sigma_x \sigma_y} E(XY) \right] \\ &= -\log\left(\frac{1}{\sqrt{1-\rho^2}}\right) + \frac{\rho^2}{1-\rho^2} \log e \end{aligned}$$

$$\therefore J(p_2, p_1; XY) = I(p_2, p_1; XY) + I(p_1, p_2; XY) = \frac{\rho^2}{1-\rho^2} \log e$$

当 XY 满足 $p_1(xy)$ 分布时, $I(X; Y) = 0$;

当 XY 满足 $p_2(xy)$ 分布时, $I(X; Y) = I(p_2, p_1; XY) = \log\left(\frac{1}{\sqrt{1-\rho^2}}\right)$

2.18 $I(p_2, p_1; X) - I(q_2, q_1; X | Y)$

$$= \sum_k p_2(x_k) \log \frac{p_2(x_k)}{p_1(x_k)} - \sum_k \sum_j q_2(x_k | y_j) h_2(y_j) \log \frac{q_2(x_k | y_j)}{q_1(x_k | y_j)}$$

$$\text{其中 } \sum_k q_2(x_k, y_j) = h_2(y_j) \quad \sum_k q_1(x_k, y_j) = h_1(y_j)$$

$$\sum_j q_2(x_k, y_j) = p_2(x_k) \quad \sum_j q_1(x_k, y_j) = p_1(x_k)$$

于是 $I(p_2, p_1; X) - I(q_2, q_1; X | Y)$

$$= \sum_k \sum_j q_2(x_k, y_j) \log \frac{p_2(x_k) \cdot q_1(x_k | y_j)}{p_1(x_k) \cdot q_2(x_k | y_j)}$$

$$= \sum_k \sum_j q_2(x_k, y_j) \log \frac{q_1(x_k, y_j)}{p_1(x_k) \cdot h_1(y_j)} \cdot \frac{h_2(y) p_2(x)}{q_2(x_k, y_j)}$$

当 $q_2(x_k, y_j) = p_2(x) h_2(y)$, 且 $q_1(x_k, y_j) = p_1(x_k) h_1(y_j)$ 时

$$I(p_2, p_1; X) - I(q_2, q_1; X | Y) = 0$$

当 $q_2(x_k, y_j) = p_2(x) h_2(y)$, 且 $q_1(x_k, y_j) \neq p_1(x_k) h_1(y_j)$ 时

$$I(p_2, p_1; X) - I(q_2, q_1; X | Y)$$

$$= \sum_k \sum_j q_2(x_k, y_j) \log \frac{q_1(x_k, y_j)}{p_1(x_k) \cdot h_1(y_j)} \geq 0$$

\therefore 关系不定

2.19 解:

天平有 3 种状态, 即平衡, 左重, 左轻, 所以每称一次消除的不确定性为 $\log 3$,

12 个球中的不等重球(可较轻, 也可较重)的不确定性为: $-\log \frac{1}{12} \cdot \frac{1}{2} = \log 24$ 因

为 $3\log 3 > \log 24$

\therefore 3 次测量可以找出该球

具体称法略。

第三章习题答案

3.1 解:

$$\because \lim_{N \rightarrow \infty} \frac{1}{N} \log P(X_1 X_2, \dots X_N) = -H(U)$$

$$\therefore \lim_{N \rightarrow \infty} P(X_1 X_2, \dots X_N)^{\frac{1}{N}} = \exp(-H(U))$$

$$\text{其中 } H(U) = -\sum_{i=1}^K P_i \log P_i$$

3.2 解:

(1)

$$P\left\{\left|\frac{\log P(U)}{N} + H_{\infty}(U)\right| \leq \delta\right\}$$

$\delta = 0$ 时

设该序列中出现 0 的个数为 n_0 , 则出现 1 的个数为 $(N - n_0)$, 则上式变为

$$P\left\{\left|\frac{n_0 \log \frac{1}{4} + (N - n_0) \log \frac{3}{4}}{N} - \frac{1}{4} \log \frac{1}{4} - \frac{3}{4} \log \frac{3}{4}\right| \leq 0\right\}$$

$$\Rightarrow P\left\{\left|\left(\log \frac{3}{4} - \log \frac{1}{4}\right)\left(\frac{1}{4} - \frac{n_0}{N}\right)\right| \leq 0\right\}$$

$$\Rightarrow P\left\{\frac{n_0}{N} = \frac{1}{4}\right\}$$

$$= \begin{cases} C^{\frac{N}{4}} \left(\frac{1}{4}\right)^{\frac{N}{4}} \left(\frac{3}{4}\right)^{\frac{3N}{4}} & N \bmod 4 = 0 \\ 0 & N \bmod 4 \neq 0 \end{cases}$$

(2)

$\delta = 0.05$ 时

同样可推得典型序列的概率为

$$\begin{cases} \sum_k C_N^k \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{N-k} & k \text{ 满足 } \left|\frac{k}{N} - \frac{1}{4}\right| \leq \frac{1}{20 \log 3} \\ 0 & \text{没有满足上述条件的 } k \end{cases}$$

3.3 0.469 bit/sample

3.4 1) 不妨设 $M = 2^j + k (j \geq 0, 0 \leq k < 2^j)$, 可进行如下编码: 首先作一深度为 j

的二叉满树，并在 2^j 个叶子节点中取 k 个节点，以这 k 个节点为根节点，生成 k 个深度为 1 的子树，于是得到了一个有 $2^j + 2k - k = M$ 个叶子的二叉树，对此二叉树的叶子按 Halfman 方法进行编码，即得到最优的二元即时码。

$$2) I = \frac{1}{M} \times (j+1) \times 2k + \frac{1}{M} \times j \times k = j + \frac{2k}{M} = \log_2 M$$

当且仅当 $k=0$ ，即 $M = 2^j$ 时， $I = \log_2 M$

3.5 解：

不妨设 $u_i (i = \dots -2, -1, 0, 1, 2, \dots)$ 取自字母表 $\{a_1, a_2 \dots a_n\}$ ，设一阶转移概率为

$$\begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix}, \text{ 所以在当前码字 } u_j \text{ 进行编码时，由 } u_{j-1} = a_k, \text{ 对 } u_j \text{ 可能的}$$

取值，依概率分布 $(P_{k1} \dots P_{kn})$ 进行 Halfman 编码，即是最佳压缩方案。

3.6 0.801 bit/sample

3.7 1) $\frac{6}{7}$ bit/sample

$$2) P(1) = \frac{2}{7} \quad P(2) = \frac{3}{7} \quad P(3) = \frac{2}{7}$$

如按无记忆信源进行编码，则根据信源所处的 1, 2, 3 三个状态对应编码成 00, 1, 01。

$$\text{平均码长为: } \frac{2}{7} \times 2 + \frac{3}{7} \times 1 + \frac{2}{7} \times 2 = \frac{11}{7} \text{ bit/sample}$$

如果按马尔可夫信源进行编码：

状态 1 时：a→0, b→10, c→11

状态 2 时：a→0, b→1

状态 3 时：无需发任何码字

$$\therefore \text{平均码长: } \frac{2}{7} \times \left(\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 2 \right) + \frac{3}{7} \times \left(\frac{1}{2} \times 1 + \frac{1}{2} \times 1 \right) + \frac{2}{7} \times 0 = \frac{6}{7}$$

bit/sample

$$\mathbf{3.8} \quad I = j + 2 - \frac{2}{x}$$

3.9 1) $H(X) = - (p \log p + q \log q)$ bit/sample

$H(Y) = - (p \log p + q \log q)$ bit/sample

2) $p=q=\frac{1}{2}$ 时, $H(Y) = 1 \text{ bit/符号}$

3.10

二元 Halfman 码: $a_1=00, a_2=01, a_3=100, a_4=101, a_5=110, a_6=1110, a_7=1111$

三元 Halfman 码: $a_1=0, a_2=1, a_3=20, a_4=21, a_5=220, a_6=221, a_7=222$

3.11 1) 证明:

$$\begin{aligned} H(U) - \log_2 3\bar{l} &= -\sum_j P(a_j) \log P(a_j) - \log_2 3 \sum_j P(a_j) l_j \\ &= \sum_{j=1}^k P(a_j) [-\log P(a_j) + \log_2 3^{-l_j}] \\ &= \sum_{j=1}^k P(a_j) \log \frac{3^{-l_j}}{P(a_j)} \leq \sum_{j=1}^k P(a_j) \left[\log \frac{3^{-l_j}}{P(a_j)} - 1 \right] = \sum_{j=1}^k 3^{-l_j} - 1 \leq 0 \end{aligned}$$

当且仅当 $\frac{3^{-l_j}}{P(a_j)} = 1$ 即 $P(a_j) = (\frac{1}{3})^{l_j} = (\frac{1}{3})^k, k=l_j$ 为整数, 取得等号,

即 $H(U) = \log_2 3\bar{l}$ 或 $\bar{l} = \frac{H(U)}{\log_2 3}$ 时, 信源字母概率取 $(\frac{1}{3})^k$ 的形式。

2) 设经过 $j+1$ 缩减, 又由于最后一次缩减必剩下 3 个字母,

即 k 满足: $k - \underbrace{3+1-3+1\cdots-3+1}_{j\uparrow} = 3 \Rightarrow k = 3 + 2j, \therefore k$ 为奇数。

3.12 1) 考虑信源 U' 满足分布:

$$\begin{bmatrix} \frac{a_1}{C_1 P(a_1)} & \frac{a_2}{C_2 P(a_2)} & \cdots & \frac{a_k}{C_k P(a_k)} \\ \frac{C_1 P(a_1)}{M} & \frac{C_2 P(a_2)}{M} & \cdots & \frac{C_k P(a_k)}{M} \end{bmatrix} \quad \text{其中 } M = \sum_{k=1}^K P(a_k) C_k$$

对 U' 进行最优编码, 则其平均码长 $\bar{l}' = \sum_{k=1}^K \frac{C_k P(a_k)}{M} l_k$, 如果 \bar{l}' 取

最小, 则 $C = M \cdot \bar{l}'$ 取最小,

$$\text{又 } \bar{l}'_{\min} = \frac{H_{\infty}(U')}{\log J} = H_{\infty}(U') = -\sum_{k=1}^K \frac{C_k P(a_k)}{M} \log \frac{C_k P(a_k)}{M}$$

$$\therefore C_{\min} = M \cdot \bar{l}'_{\min} = -\sum_{k=1}^K P(a_k) \log \frac{C_k P(a_k)}{\sum_{i=1}^K C_i P(a_i)}$$

2) $\therefore H_{\infty}(U') \leq \bar{l}' \leq H_{\infty}(U') + 1$

$$\text{即 } H_{\infty}(U') \leq \frac{C}{M} \leq H_{\infty}(U') + 1 \Rightarrow M \cdot H_{\infty}(U') \leq C \leq M \cdot H_{\infty}(U') + M$$

$$\text{又 } C_{\min} = M \cdot H_{\infty}(U')$$

$$\therefore C_{\min} \leq C \leq C_{\min} + \sum_{k=1}^K P(a_k) C_k$$

第4章习题答案

4.1 设有离散无记忆信道，输入 $X: \begin{Bmatrix} a_1 & a_2 & \cdots & a_K \\ p(a_1) & p(a_2) & \cdots & p(a_K) \end{Bmatrix}$ ，输出 $Y:$

$\begin{Bmatrix} b_1 & b_2 & \cdots & b_J \\ p(b_1) & p(b_2) & \cdots & p(b_J) \end{Bmatrix}$ ，当输入/出 x 和 y 的互信息 $I(x;y)$ 也为一随机变量，试

证：当 $\text{Var}\{I(x;y)\} = 0$ 时，平均互信息 $I(X;Y)$ 达到信道容量 C 。

证明：

由 $\text{Var}\{I(x;y)\} = 0$ 可知， $I(x;y)$ 以概率1取常数 $C (= E[I(x;y)])$ ，所以有

$$I(x;Y) = \sum P(b_j/a_k) I(x=a_k; y=b_j) = C$$

满足定理4.3关于离散无记忆信道达到信道容量的充要条件。且此时平均互信息

$$I(X;Y) = \sum_{k=1}^K p(a_k) I(a_k;Y) = C$$

所以当 $\text{Var}\{I(x;y)\} = 0$ 时，平均互信息达到信道容量 C 。

<证毕>

4.2 设某信道的输入 X 取值 $\{+1, -1\}$ ，又信道有加性噪声 n ，其分布密度为

$$p(n) = \begin{cases} \frac{1}{4}, & |n| \leq 2 \\ 0, & |n| > 2 \end{cases}, \text{求信道容量。}$$

答： $C = 0.5\text{bit}$ ，当输入的概率分布为 $P(X=1) = P(X=-1) = \frac{1}{2}$ 时达到信道容量。

4.3 设在图 4.10 的一般高斯信道中 $N(f) = \frac{N_0}{2}$, $H(f) = \frac{1}{1+(f/f_0)^2}$ ，试求信道的容

量费用函数 $C(P_s)$

解：

$$\frac{1}{2\lambda} - \frac{N(f)}{|H(F_\lambda)|^2} = 0 \quad (1)$$

根据

$$\int_{-F_\lambda}^{F_\lambda} \left(\frac{1}{2\lambda} - \frac{N(f)}{|H(f)|^2} \right) df \quad (2)$$

可推得 F_λ 的一个方程，如下：

$$\frac{4}{3} \frac{F_\lambda^3}{f_0^2} + \frac{4}{5} \frac{F_\lambda^5}{f_0^4} = \frac{P_s}{N_0} \quad (3)$$

解出(3)式,即可求得 F_λ

而

$$\begin{aligned}
 C(P_s) &= \int_{f \in F_\lambda} \frac{1}{2} \log \frac{|H(f)|^2}{2\lambda N(f)} df \\
 &= \int_{-F_\lambda}^{F_\lambda} \frac{1}{2} \log \frac{1}{2\lambda} df - \int_{-F_\lambda}^{F_\lambda} \frac{1}{2} \log \left\{ \frac{N_0}{2} \left[1 + \left(\frac{f}{f_0} \right)^2 \right]^2 \right\} df \\
 &= 4[F_\lambda - f_0 \operatorname{tg}^{-1}(\frac{F_\lambda}{f_0})] \log e
 \end{aligned}$$

故 $C(P_s)$ 可求。

4.4 设 X 和 Y 为信道的输入和输出，两者均取值于集合 $A = \{a_1, a_2, \dots, a_K\}$ 。已知

$p(x = a_k) = p_k$, $p(y = a_j | x = a_k) = p_{kj}$, 定义 $P_e = \sum_k p_k \sum_{j \neq k} p_{kj}$, 求证:

$$H(X|Y) \leq P_e \log(K-1) + H(P_e)$$

其中 $H(P_e) = -P_e \log P_e - (1 - P_e) \log(1 - P_e)$ 。

证明:

法一: 构造随机变量 Z , 满足

$$Z = \begin{cases} 1, & \text{当 } X \neq Y \\ 0, & \text{当 } X = Y \end{cases}$$

于是 $P(Z=1) = P_e$, $P(Z=0) = 1 - P_e$, 从而

$$\begin{aligned}
 H(XZ|Y) &= H(X|Y) + H(Z|XY) \\
 &= H(Z|Y) + H(X|YZ)
 \end{aligned}$$

又由 $H(Z|XY) = 0$, 可得

$$H(X|Y) = H(Z|Y) + H(Z|YZ) \leq H(Z) + H(X|YZ) \quad (\text{a})$$

而 $H(X|YZ) = P(Z=1)H(X|Y, Z=1) + \underbrace{P(Z=0)H(X|Y, Z=0)}_{\text{等于0}}$

$$= P_e H(X|Y, Z=1) \leq P_e \log(K-1) \quad (\text{b})$$

$$H(Z) = H(P_e) \quad (\text{c})$$

将(b), (c)代入(a)可得

$H(X|Y) \leq P_e \log(K-1) + H(P_e)$ 成立。

<证毕>

法二:

$$H(X|Y) = -\sum_k \sum_{j \neq k} P(a_k, a_j) \log \frac{P(a_k, a_j)}{P(a_j)} - \sum_k P(a_k, a_k) \log \frac{P(a_k, a_k)}{P(a_k)}$$

于是有

$$\begin{aligned} & P_e \log(K-1) + H(P_e) - H(X|Y) \\ &= P_e \log(K-1) - P_e \log P_e - (1-P_e) \log(1-P_e) \\ &+ \sum_k \sum_{j \neq k} P(a_k, a_j) \log \frac{P(a_k, a_j)}{P(a_k)} + \sum_k P(a_k, a_k) \log \frac{P(a_k, a_k)}{P(a_k)} \\ &= \left(\sum_k \sum_{j \neq k} P(a_k, a_j) \right) \log \frac{(K-1)}{P_e} - \left(\sum_k P(a_k, a_k) \right) \log(1-P_e) \\ &+ \sum_k \sum_{j \neq k} P(a_k, a_j) \log \frac{P(a_k, a_j)}{P(a_k)} + \sum_k P(a_k, a_k) \log \frac{P(a_k, a_k)}{P(a_k)} \\ &= \sum_k \sum_{j \neq k} P(a_k, a_j) \log \frac{(K-1)}{P_e} \frac{P(a_k, a_j)}{P(a_k)} + \sum_k P(a_k, a_k) \log \frac{1}{(1-P_e)} \frac{P(a_k, a_k)}{P(a_k)} \\ &\geq \sum_k \sum_{j \neq k} P(a_k, a_j) \left(1 - \frac{P_e}{(K-1)} \frac{P(a_k)}{P(a_k, a_j)} \right) + \sum_k P(a_k, a_k) \left(1 - (1-P_e) \frac{P(a_k)}{P(a_k, a_k)} \right) \\ &= \sum_{k=1}^K \sum_{j=1}^K P(a_k, a_j) - \sum_k \sum_{j \neq k} \frac{P_e}{(K-1)} P(a_k) - \sum_k P(a_k)(1-P_e) \\ &= 1 - P_e - (1-P_e) = 0 \end{aligned}$$

所以有 $H(X|Y) \leq P_e \log(K-1) + H(P_e)$ 。

<证毕>

4.5 已知信道转移概率矩阵如下，求此信道的信道容量。

x \ y	y			
	0	1	2	3
0	1/3	1/3	1/6	1/6
1	1/6	1/3	1/6	1/3

答：0.04bit，当输入等概时达到信道容量。

4.6 设有信道，输入X的字母表为： $\{0,1,2,\dots,K-1\}$ ，噪声为独立加性噪声Z，Z的取值也在 $\{0,1,2,\dots,K-1\}$ 集合中，但两者相加为模K相加，即输出 $Y = X \oplus Z$ 模K，试求此信道的信道容量。

答： $\log K - H(Z)$ ，输入等概时达到信道容量

4.7 设有二元对称信道 $Y_n = X_n \oplus Z_n$ ，其中 \oplus 为模2和， $X_n, Y_n \in \{0,1\}$ ， $Z_n: \begin{Bmatrix} 0 & 1 \\ 1-p & p \end{Bmatrix}$ ，

但 Z_n 不是独立随机序列，试证：

$$\max I(X_1 X_2 \cdots X_N; Y_1 Y_2 \cdots Y_N) \geq NC$$

其中, $C = 1 - H(P)$ 。

证明:

对任意输入分布有下式成立

$$I(\mathbf{X}; \mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y} | \mathbf{X}) = H(\mathbf{Y}) - H(\mathbf{Z}) \geq H(\mathbf{Y}) - \sum_n H(Z_n)$$

取输入分布 $P'(X_1 X_2 \cdots X_N)$ 使得序列 $Y_1 Y_2 \cdots Y_N$ 统计独立, 且

$P(Y_n = 1) = P(Y_n = 0) = \frac{1}{2}$, 设这时的互信息为 $I'(X_1 X_2 \cdots X_N; Y_1 Y_2 \cdots Y_N)$, 这时有

$$H(\mathbf{Y}) = \sum_{n=1}^N H(Y_n) = N,$$

所以

$$\begin{aligned} & \max_{P(x_1 x_2 \cdots x_N)} I(X_1 X_2 \cdots X_N; Y_1 Y_2 \cdots Y_N) \\ & \geq I'(X_1 X_2 \cdots X_N; Y_1 Y_2 \cdots Y_N) \\ & \geq N - \sum_{n=1}^N H(Z_n) = N - NH_Z(P) = NC \end{aligned}$$

<证毕>

4.8 设有输入为 X , 输出为 $\mathbf{Y} = [Y_1 \ Y_2]$ 的高斯信道, 其中 $Y_1 = X + Z_1$, $Y_2 = X + Z_2$, X

的最大功率受限 P , $(Z_1 Z_2) \sim N_2(0, K)$, 其中 $K = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix}$, 试求:

$$1) \ I(X; Y_1 Y_2) = I(X; Y_1) + I(X; Y_2) - I(Y_1; Y_2) + I(Y_1; Y_2 | X)$$

2) $\rho = 1$ 时的信道容量。

1) 证明:

$$I(X; Y_1 Y_2) = I(X; Y_1) + I(X; Y_2 | Y_1) \tag{a}$$

$$\text{又由 } I(X; Y_1; Y_2) = I(X; Y_2) - I(X; Y_2 | Y_1) = I(Y_1; Y_2) - I(Y_1; Y_2 | X)$$

$$\text{所以 } I(X; Y_2 | Y_1) = I(X; Y_2) + I(Y_1; Y_2 | X) - I(Y_1; Y_2)$$

$$\text{代入(a), 可得 } I(X; Y_1 Y_2) = I(X; Y_1) + I(X; Y_2) - I(Y_1; Y_2) + I(Y_1; Y_2 | X)。$$

<证毕>

2) 答: $C = \frac{1}{2} \log(1 + \frac{P}{\sigma^2})$, 当输入是方差为 P 的高斯分布时达到信道容量。

第 5 章 习 题 答 案

5.1 设有码C，在用此码传输时能使 $H(\mathbf{U}|\mathbf{Y})=0$ （式中 \mathbf{U} 表示码字矢量， \mathbf{Y} 代表接收矢量），试证此时的码率 R 必小于该信道的信道容量 C 。

（提示： $H(\mathbf{U}) = H(\mathbf{U}|\mathbf{Y}) + I(\mathbf{U};\mathbf{Y})$ ）

证明：

$$R = \frac{H(\mathbf{U})}{N} = \frac{H(\mathbf{U}|\mathbf{Y}) + I(\mathbf{U};\mathbf{Y})}{N} = \frac{I(\mathbf{U};\mathbf{Y})}{N} \leq \max \frac{I(\mathbf{U};\mathbf{Y})}{N} = C$$

所以结论成立。

<证毕>

5.2 设用码字母表大小为 J ，码长为 N 的分组码试通信道容量为 C 的信道传输信息，试证：收端译码的差错概率 P_e 满足

$$P_e \geq \frac{1}{\log J} \left(R - C - \frac{1}{N} \right)$$

证明：

由4.4证明可知

$$H(X|Y) \leq P_e \log(K-1) + H(P_e) \leq P_e \log(K-1) + 1$$

$$\text{又 } R - C - \frac{1}{N} = \frac{H(X) - I(X;Y) - 1}{N} = \frac{H(X|Y) - 1}{N} \leq \frac{P_e \log(K-1)}{N}$$

而 $K = J^N$ ，故

$$R - C - \frac{1}{N} \leq \frac{P_e \log(J^N - 1)}{N} < P_e \log J$$

$$\text{故 } P_e \geq \frac{1}{\log J} \left(R - C - \frac{1}{N} \right)$$

5.7 设(6,3)二元线性码的生成矩阵

$$G = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix},$$

a. 试找出 G 的行缩减梯形法式表示；

答: $G' = \begin{vmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{vmatrix}$

b. 求监督矩阵 H ;

答: $H = \begin{vmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{vmatrix}$

c. 找出最小重陪集首项;

伴随式	陪集首
000	000000
001	000001
010	000010
011	000100
100	001000
101	010000
110	001010
111	100000

d. 在B.S.C信道中对接收矢量111010, 000011, 101010进行译码;

答: 111010译码为110010; 000011译码为000111, 101010译码为101011。

5.9 设有码如下表示:

信息	码字
00	00000
01	01101
10	10111
11	11010

a. 找出生成矩阵 G 与监督矩阵 H ;

答: $G = \begin{vmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{vmatrix};$

b. 在B.S.C下给出最大似然译码的译码表;

答:

译码表如下:

伴随式	陪集首			
000	00000	01101	10111	11010
001	00001	01100	10110	11011
010	00010	01111	10101	11000
011	00011	01110	10100	11001
100	00100	01001	10011	11110
101	01000	00101	11111	10010
110	10001	11100	00110	01011
111	10000	11101	00111	01010
译码结果	00000	01101	10111	11010

c. 求正确译码的概率。

答: $P_e = (1 - \varepsilon)^5 + 5\varepsilon(1 - \varepsilon)^4 + 2\varepsilon^2(1 - \varepsilon)^3$

5.10 试证:

- 二元线性码中码字重必全为偶数或奇偶各半;
- (n, k) 码的平均码字重不超过 $n/2$ 。

证明:

a. 首先证明如下事实: 偶重码字与偶重码字之和为偶重码字, 奇重码字与奇重码字之和为偶重码字, 奇重码字与偶重码字之和为奇重码字。

设码字 C_1, C_2 为偶重码字, 即 $w(C_1) = 2n$, $w(C_2) = 2m$, 设 C_1, C_2 码字有 h 个 1 的位置是相同的, 于是有 $w(C_1 + C_2) = 2n + 2m - 2h$, 所以偶重码字和偶重码字之和为偶重码字。

类似可得其他结论成立。

因为码字是由生成矩阵 G 得到的, 即

$$\mathbf{C} = \mathbf{uG} = [u_1, u_2 \cdots u_k] \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_k \end{bmatrix}, \text{ 其中 } \mathbf{g}_1, \mathbf{g}_2 \cdots \mathbf{g}_k \text{ 是生成矩阵的行向量。}$$

1) 如果生成矩阵的所有行向量重为偶, 则由偶重码字之和为偶重码字可知, 所有的码字重都为偶重码字;

2) 如果生成矩阵的 k 个行向量中存在至少一个向量其重为奇, 不妨设该向量为 \mathbf{g}_1 , 它对应的消息

位为 u_1 , 记 N_{1o} 为 u_1 取 1 时, 码重为奇的码字数, 相应的码字集合为 S_{1o} , N_{0o} 为 u_1 取 0

时, 码重为奇的码字数, 相应的码字集合为 S_{0o} , 记 N_{1e} 为 u_1 取1时, 码重为偶的码字数, 相应的码字集合为 S_{1e} , N_{0e} 为 u_1 取0时, 码重为偶的码字数, 相应的码字集合为 S_{0e} , 下证有 $N_{1o} = N_{0e}$ 成立: $\forall C_i \in S_{1o}$, 有 $C_i = \mathbf{g}_1 + u_{i2}\mathbf{g}_2 + \cdots + u_{ik}\mathbf{g}_k$, 因为 \mathbf{g}_1 , C_i 为奇重, 所以 $u_{i2}\mathbf{g}_2 + \cdots + u_{ik}\mathbf{g}_k$ 为偶重, 从而必存在一个 $C_i' \in S_{0e}$, 且满足 $C_i' = 0\mathbf{g}_1 + u_{i2}\mathbf{g}_2 + \cdots + u_{ik}\mathbf{g}_k$, 即 S_{1o} 和 S_{0e} 存在一一对应的关系, 所以有 $N_{1o} = N_{0e}$ 成立。

类似可证 $N_{0o} = N_{1e}$ 。

又奇重码字数 = $N_{1o} + N_{0o}$, 偶重码字数 = $N_{1e} + N_{0e}$, 所以有奇重码字数 = 偶重码字数。

<证毕>

c. 将 2^k 个 (n, k) 码排成一个矩阵

$$\begin{bmatrix} C_1^1 & C_2^1 & \cdots & C_n^1 \\ C_1^2 & C_2^2 & \cdots & C_n^2 \\ \cdots & & & \\ C_1^{2^k} & C_2^{2^k} & \cdots & C_n^{2^k} \end{bmatrix},$$

记上述矩阵的第 j 列为 $\mathbf{C}_j = [C_j^1 \ C_j^2 \ \cdots \ C_j^{2^k}]^T$, 第 i 行即第 i 个码字为 \mathbf{C}^i , 其中 i, j 满足 $1 \leq i \leq 2^k, 1 \leq j \leq n$ 。则平均码重为

$$\bar{w} = \frac{1}{2^k} \sum_{i=1}^{2^k} w(\mathbf{C}^i) = \frac{1}{2^k} \sum_{i=1}^{2^k} \sum_{j=1}^n C_j^i = \frac{1}{2^k} \sum_{j=1}^n w(\mathbf{C}_j)$$

对 \mathbf{C}_j 的情况进行讨论:

1) 如果 $\mathbf{C}_j = \mathbf{0}$, 即生成矩阵存在全零列, 则有 $w(\mathbf{C}_j) = 0$;

2) 如果 $\mathbf{C}_j \neq \mathbf{0}$, 即存在 $C_j^i = 1$, 则由线性码的性质, $\forall m$, 必 $\exists n$, 满足 $\mathbf{C}^m + \mathbf{C}^i = \mathbf{C}^n$, 所以有 $C_j^m + C_j^i = C_j^n$, 从而有 $C_j^m \neq C_j^n$, 所以第 j 列中1和0的个数必相等, 即有 $w(\mathbf{C}_j) = 2^{k-1}$ 。

$$\text{综上有 } \bar{w} = \frac{1}{2^k} \sum_{j=1}^n w(\mathbf{C}_j) \leq \frac{1}{2^k} n \cdot 2^{k-1} = \frac{n}{2}。$$

<证毕>

第6章习题答案

- 6.1 设已知离散无记忆信源在给定失真量度 $d(k, j)$, $k = 1, 2, \dots, K, j = 1, 2, \dots, J$ 下的信息速率失真函数为 $R(D)$, 现定义新的失真量度 $d'(k, j) = d(k, j) - g_k$ 。试证: 在新的失真量度下信息速率失真函数:

$$R'(D) \text{ 为 } R'(D) = R(D + G), \text{ 其中 } G = \sum_k p(a_k) g_k。$$

证明:

对离散无记忆信源

$$R(D) = \min\{I(U; V), E\{d(U, V)\} \leq D\}$$

$$\begin{aligned} \therefore R'(D) &= \min\{I(U; V), E\{d'(U, V)\} \leq D\} \\ &= \min\{I(U; V), E\{d(k, j) - g_k\} \leq D\} \\ &= \min\{I(U; V), E\{d(k, j) - E\{g_k\}\} \leq D\} \\ &= \min\{I(U; V), E\{d(k, j) \leq E\{g_k\} + D\}\} \end{aligned}$$

而 $E\{g_k\} = \sum_k p(a_k) g_k = G$, 故上式可表示为

$$R'(D) = \min\{I(U; V), E\{d(k, j) \leq D + G\}\} = R(D + G)$$

<证毕>

- 6.2 设有带宽为4kHz的限带白色高斯信源, 欲通过信道容量为16 kb/s的信道传输, 试求在理想情况下信道输出端可能得到的最大信噪比。

答: 理想情况下, 当 $R \leq C$ 可以实现无误传输, 即

$$R = F \log \frac{P}{D} \leq C = 16000$$

将数值代入, 可得

$$\frac{P}{D} \leq 2^4 = 16$$

所以可能达到的最大信噪比为16。

- 6.3 设 $\{\alpha^n\}$ ($n = 1, 2, \dots, N$) 是 N 个正交规范化的 N 维向量。A 是由 $(\alpha^1, \alpha^2, \dots, \alpha^N)$ 组成的正交矩阵。N 维随机向量 X 在正交变换下得到新的随机向量 Z, $Z = A^T X$, 其中 $X = (x_1, x_2, \dots, x_N)^T$, $Z = (z_1, z_2, \dots, z_N)^T$, $E\{X\} = 0$ 。令

$$\sigma_n^2 = E\{(z_n - \bar{z}_n)^2\}, \quad \rho_n = \frac{\sigma_n^2}{\sum_{n=1}^N \sigma_n^2}$$

$$H(Z) = -\sum_{n=1}^N \rho_n \log \rho_n$$

试证：所有正交变换中， $K-L$ 变换所对应的 Z 有最小的熵 $H(Z)$ 。

证明：

1) 设随机向量 \mathbf{X} 的相关矩阵为 \mathbf{R} ，对应的 $K-L$ 变换为矩阵 $\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_N]$ ，相应的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_N$ ，由题目条件可知

$z_n = (\mathbf{a}^n)^T \mathbf{X}$ ，由 $E[\mathbf{X}] = 0$ ，知 $E[z_n] = 0$ ， $1 \leq n \leq N$ ，从而

$$\sigma_n^2 = E[z_n^2] = E[(\mathbf{a}^n)^T \mathbf{X} \mathbf{X}^T \mathbf{a}^n] = (\mathbf{a}^n)^T \mathbf{R} \mathbf{a}^n, \quad 1 \leq n \leq N;$$

2) \mathbf{a}^n 可展开为 $\mathbf{a}^n = b_{n1}\mathbf{u}_1 + b_{n2}\mathbf{u}_2 + \cdots + b_{nN}\mathbf{u}_N$ ， $1 \leq n \leq N$ ，所以

$$\begin{aligned} \sigma_n^2 &= (b_{n1}\mathbf{u}_1^T + b_{n2}\mathbf{u}_2^T + \cdots + b_{nN}\mathbf{u}_N^T) \mathbf{R} (b_{n1}\mathbf{u}_1 + b_{n2}\mathbf{u}_2 + \cdots + b_{nN}\mathbf{u}_N) \\ &= b_{n1}^2 \lambda_1 + b_{n2}^2 \lambda_2 + \cdots + b_{nN}^2 \lambda_N \end{aligned}$$

由于 $\{\mathbf{a}^n\}$ 是正交规范化向量，所以

$$\|\mathbf{a}^n\|^2 = b_{n1}^2 + b_{n2}^2 + \cdots + b_{nN}^2 = 1$$

$$\forall i \neq j, (\mathbf{a}^i)^T (\mathbf{a}^j) = b_{i1}b_{j1} + b_{i2}b_{j2} + \cdots + b_{iN}b_{jN} = 0$$

$$\text{令矩阵 } \mathbf{B} = \begin{bmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & \vdots & \vdots \\ b_{N1} & \cdots & b_{NN} \end{bmatrix}, \text{ 于是有 } \mathbf{B}\mathbf{B}^T = \mathbf{I}, \text{ 且 } \mathbf{B}^T\mathbf{B} = \mathbf{I}, \text{ 即}$$

$$\sum_{i=1}^N b_{in}^2 = 1, \quad 1 \leq n \leq N; \quad (\text{a})$$

$$\text{所以 } \sum_{n=1}^N \sigma_n^2 = \sum_{n=1}^N (b_{n1}^2 \lambda_1 + b_{n2}^2 \lambda_2 + \cdots + b_{nN}^2 \lambda_N) = \lambda_1 + \lambda_2 + \cdots + \lambda_N;$$

3) 记

$$h(\rho_n) = -\rho_n \log \rho_n = -\left(\frac{b_{n1}^2 \lambda_1 + b_{n2}^2 \lambda_2 + \cdots + b_{nN}^2 \lambda_N}{\lambda_1 + \lambda_2 + \cdots + \lambda_N}\right) \log\left(\frac{b_{n1}^2 \lambda_1 + b_{n2}^2 \lambda_2 + \cdots + b_{nN}^2 \lambda_N}{\lambda_1 + \lambda_2 + \cdots + \lambda_N}\right)$$

$$H(Z) = \sum_{n=1}^N h(\rho_n) = \sum_{n=1}^N h\left(b_{n1}^2 \frac{\lambda_1}{\lambda_1 + \lambda_2 + \cdots + \lambda_N} + \cdots + b_{nN}^2 \frac{\lambda_N}{\lambda_1 + \lambda_2 + \cdots + \lambda_N}\right)$$

由熵函数的凸性，及 $b_{n1}^2 + b_{n2}^2 + \cdots + b_{nN}^2 = 1$ ，

$$H(Z) \geq \sum_{n=1}^N \left(b_{n1}^2 h\left(\frac{\lambda_1}{\lambda_1 + \lambda_2 + \cdots + \lambda_N}\right) + b_{n2}^2 h\left(\frac{\lambda_2}{\lambda_1 + \lambda_2 + \cdots + \lambda_N}\right) + \cdots + b_{nN}^2 h\left(\frac{\lambda_N}{\lambda_1 + \lambda_2 + \cdots + \lambda_N}\right) \right)$$

将(a)式代入上式, 可得

$$H(Z) \geq \sum_{n=1}^N h\left(\frac{\lambda_n}{\lambda_1 + \lambda_2 + \cdots + \lambda_N}\right),$$

不等式右边即为K-L变换所对应得熵, 所以结论成立。

6.4 设无记忆信源 $\begin{vmatrix} X \\ p(x) \end{vmatrix} = \begin{vmatrix} -1, & 0, & 1 \\ 1/3, & 1/3, & 1/3 \end{vmatrix}$, 接收符号 $A_Y = \left\{-\frac{1}{2}, \frac{1}{2}\right\}$, 失真矩阵

$$D = \begin{vmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 1 \end{vmatrix}, \text{ 试求: } D_{\max} \text{ 和 } D_{\min} \text{ 及达到 } D_{\max}, D_{\min} \text{ 时的转移概率矩阵。}$$

答: $D_{\min} = 1$, 当转移概率矩阵为 $\begin{bmatrix} 1 & 0 \\ \alpha & 1-\alpha \\ 0 & 1 \end{bmatrix}$, α 满足 $0 \leq \alpha \leq 1$;

$D_{\max} = \frac{4}{3}$, 当转移概率矩阵为 $\begin{bmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \\ \alpha & 1-\alpha \end{bmatrix}$, α, β 满足 $0 \leq \alpha \leq 1, 0 \leq \beta \leq 1$ 。

6.5 已知二元信源 $X: \begin{pmatrix} 0, & 1 \\ p, & 1-p \end{pmatrix}$ 以及失真矩阵 $(d_{kj}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, 试求:

- 1) D_{\min} ;
- 2) D_{\max} ;
- 3) $R(D)$ 。

答: $D_{\min} = 0$;

$$D_{\max} = \min\{p, 1-p\};$$

$$R(D) = H(p) - H(D), \text{ 其中 } H(D) = -D \ln D - (1-D) \ln(1-D), 0 \leq D \leq D_{\max}。$$

6.6 设有总功率为10 mW的限带(0~5 kHz)白色高斯信源通过加性白色高斯噪声信道传输, 后者的带宽为0~10 kHz, 噪声的单边功率谱密度为1 $\mu\text{W/Hz}$, 容许最大输入功率为20 mW。试求在理想情况下接收信号可以达到的最小均方误差。

$$\text{答: } R = F \ln \frac{P_u}{D} = 5000 \ln \frac{10}{D},$$

$$C = W \ln\left(1 + \frac{P_s}{N_0 W}\right) = 10000 \ln\left(1 + \frac{20}{10}\right),$$

理想状况下, $R \leq C$, 所以接收信号可以达到的最小均方误差为 $D = \frac{10}{9} \text{ mW}$ 。

6.7 设有平稳高斯信源 $X(t)$, 其功率谱为

$$G(f) = \begin{cases} A, & |f| \leq F_1 \\ 0, & |f| > F_1 \end{cases}$$

失真度量取 $d(x, y) = (x - y)^2$ ，容许的样值失真为 D ，试求：

1) 信息速率失真函数 $R(D)$ ；

2) 用一独立加性高斯信道(带宽 F_2 ，限功率 P ，噪声的双边功率谱密度 $\frac{N_0}{2}$)来传送

上述信源时，最小可能方差与 F_2 的关系。

答： $R(D) = F_1 \log\left(\frac{2AF_1}{D}\right)$;

$D \geq 2AF_1\left(1 + \frac{P}{N_0F_2}\right)^{\frac{F_2}{F_1}}$ ，即最小可能方差为 $2AF_1\left(1 + \frac{P}{N_0F_2}\right)^{\frac{F_2}{F_1}}$ 。

6.8 设有矢量信源，其各分量 $X_k \sim N(0, \sigma_k^2)$ ， $k = 1, 2, \dots, K$ ，是 K 个独立的随机变量，

失真 $d(x_1 x_2 \dots x_K; \hat{x}_1 \hat{x}_2 \dots \hat{x}_K) = \sum_{k=1}^K (x_k - \hat{x}_k)^2$ ，试证：在此条件下

$$R(D) = \sum_{k=1}^K \frac{1}{2} \log \frac{\sigma_k^2}{D_k}$$

其中， $D_k = \begin{cases} \lambda, & \text{当 } \lambda < \sigma_k^2 \\ \sigma_k^2, & \text{当 } \lambda \geq \sigma_k^2 \end{cases}$ ， λ 满足 $\sum_{k=1}^K D_k = D$ 。

证明：

$$R(D) = \min_{q(\hat{\mathbf{x}}|\mathbf{x})} \{I(\mathbf{X}; \hat{\mathbf{X}}), E\{d(\mathbf{X}, \hat{\mathbf{X}})\} \leq D\}$$

$$= \min_{q(\hat{\mathbf{x}}|\mathbf{x})} \left\{ \sum_i H(x_i) - H(\mathbf{X} | \hat{\mathbf{X}}), \sum_{i=1}^K (x_i - \hat{x}_i)^2 \leq D \right\}$$

$$\begin{aligned} H(\mathbf{X} | \hat{\mathbf{X}}) &= H(x_1 | \hat{x}_1 \hat{x}_2 \dots \hat{x}_K) + H(x_2 | \hat{x}_1 \hat{x}_2 \dots \hat{x}_K x_1) + \dots + H(x_K | \hat{x}_1 \hat{x}_2 \dots \hat{x}_K x_1 x_2 \dots x_{K-1}) \\ &\leq H(x_1 | \hat{x}_1) + H(x_2 | \hat{x}_2) + \dots + H(x_K | \hat{x}_K) \end{aligned}$$

所以有

$$I(\mathbf{X}; \hat{\mathbf{X}}) \geq \sum_i H(x_i) - H(x_i | \hat{x}_i) = \sum_i I(x_i; \hat{x}_i)$$

等号当且仅当在转移概率满足独立性条件时，即 $q(\hat{x}_1 \hat{x}_2 \dots \hat{x}_K | x_1 x_2 \dots x_K) = q(\hat{x}_1 | x_1) q(\hat{x}_2 | x_2) \dots q(\hat{x}_K | x_K)$

所以要使互信息最小，必定要设计编码器的转移概率独立，由于输入统计独立，所以只要设计各组成分量的互信息最小，而各组成分量的互信息的最小值为

$$R_k(D_k) = \frac{1}{2} \log \frac{\sigma_k^2}{D_k}, 0 \leq D_k \leq \sigma_k^2,$$

所以求 $R(D) = \min_{q(\hat{\mathbf{X}}|\mathbf{X})} \{I(\mathbf{X};\hat{\mathbf{X}}), E\{d(\mathbf{X},\hat{\mathbf{X}})\} \leq D\}$ 即在约束条件

$$0 \leq D_k \leq \sigma_k^2$$

$$\sum_{k=1}^K D_k = D$$

下求 $R(D) = \min \sum_k \frac{1}{2} \log \frac{\sigma_k^2}{D_k}$, 由拉格朗日求极值方法可知当

$$D_k = \begin{cases} \lambda, & \text{当 } \lambda < \sigma_k^2, \\ \sigma_k^2, & \text{当 } \lambda \geq \sigma_k^2, \end{cases} \text{ 且 } \lambda \text{ 满足 } \sum_{k=1}^K D_k = D \text{ 时 } R(D) \text{ 取得最小值。}$$

6.9 设有离散无记忆信源 X 经编码后输出 Y , 失真矩阵的所有列是集合 $\{d_1, d_2, \dots, d_m\}$ 的某一置换。定义函数

$$\Phi(D) = \max_{\mathbf{P}: \sum_{i=1}^m P_i d_i \leq D} H(\mathbf{P})$$

- 1) 证明 $\Phi(D)$ 是 D 的上凸函数;
- 2) 证明: $I(X;Y) \geq H(X) - \Phi(D)$ 。

证明

1)

设概率分布 \mathbf{P}' 满足 $\sum_i P'_i d_i \leq D_1$, 令满足 $H(\mathbf{P}') = \Phi(D_1)$ 的熵函数为 H_1 , 设 \mathbf{P}'' 满足

$\sum_i P''_i d_i \leq D_2$, 令满足 $H(\mathbf{P}'') = \Phi(D_2)$ 的熵函数为 H_2 , 定义 $\mathbf{P} = \lambda_1 \mathbf{P}' + \lambda_2 \mathbf{P}''$, 且有

$\lambda_1 + \lambda_2 = 1, 0 \leq \lambda_1, \lambda_2 \leq 1$, 则有

$$\sum_i P_i d_i = \sum_i \lambda_1 P'_i d_i + \lambda_2 P''_i d_i \leq \lambda_1 D_1 + \lambda_2 D_2 = D \quad (\text{a})$$

又由熵函数的凸性可知,

$$H(\mathbf{P}) \geq \lambda_1 H_1 + \lambda_2 H_2 = \lambda_1 \Phi(D_1) + \lambda_2 \Phi(D_2)$$

又由 $\Phi(D)$ 的定义可知 $\Phi(D) \geq H(\mathbf{P})$, 所以有

$$\Phi(D) \geq \lambda_1 \Phi(D_1) + \lambda_2 \Phi(D_2) \quad (\text{b})$$

由(a), (b)可知函数 $\Phi(D)$ 是 D 的上凸函数。

<证毕>

2)

$$D = E[d(x, y)] = \sum_{i=1}^m \sum_{j=1}^m P(x_i y_j) d(x_i, y_j) = \sum_{j=1}^m \sum_{i=1}^m P(y_j) P(x_i | y_j) d(x_i, y_j)$$

令 $D_j = \sum_{i=1}^m P(x_i | y_j) d(x_i, y_j) = \sum_{i=1}^m q_i^j d(x_i, y_j)$, 于是有

$$D = E[d(x, y)] = \sum_{j=1}^m P(y_j) D_j$$

由于失真矩阵所有的列是集合 $\{d_1, d_2, \dots, d_m\}$ 的某一置换, D_j 也可看成是是概率分布

$$[P(x_1 | y_j) \quad P(x_2 | y_j) \quad \cdots \quad P(x_m | y_j)] = [q_1^j \quad q_2^j \quad \cdots \quad q_m^j] = \mathbf{q}^j$$

的某一置换和 $[d_1 \quad d_2 \quad \cdots \quad d_m]$ 的内积, 所以可以定义函数

$$\Phi(D_j) = \max_{\mathbf{q}^j: \sum q_i^j d_i \leq D_j} H(\mathbf{q}^j), \text{ 于是由 } \Phi(D) \text{ 的凸性可得}$$

$$\Phi(D) = \Phi\left(\sum_{j=1}^m P(y_j) D_j\right) \geq \sum_{j=1}^m P(y_j) \Phi(D_j) \geq \sum_{j=1}^m P(y_j) H(\mathbf{q}^j)$$

$$= -\sum_{j=1}^m \sum_{i=1}^m P(y_j) P(x_i | y_j) \log P(x_i | y_j) = H(X | Y)$$

所以有 $I(X; Y) \geq H(X) - \Phi(D)$ 成立

<证毕>

Qbit.5d6d.com

Qbit.5d6d.com

第一题修改如下，在 $H(U)$ 前加一个“ $-$ ”号。

$$\therefore \lim_{N \rightarrow \infty} \frac{1}{N} \log P(X_1 X_2, \dots X_N) = -H(U)$$

$$\therefore \lim_{N \rightarrow \infty} P(X_1 X_2, \dots X_N)^{\frac{1}{N}} = \exp(-H(U))$$

$$\text{其中 } H(U) = -\sum_{i=1}^K P_i \log P_i$$

互信息函数 $I(P, Q)$ 的性质 2 的证明。

对于确定的条件概率矩阵 Q 互信息函数 $I(P, Q)$ 是概率矢量空间 S 上的上凸函数。

(其中 $S = \{ P : P = (p_1, p_2, \dots, p_K), 0 \leq p_k \leq 1, k = 1, 2, \dots, K, \text{ 而 } \sum_{k=1}^K p_k = 1 \}$)

证明：首先由定义知： $I(X, Y) = H(Y) - H(Y|X)$

$$\text{其中 } H(Y) = - \sum_{j=1}^J p(b_j) \log p(b_j)$$

$$= - \sum_{k=1}^K \sum_{j=1}^J p(a_k, b_j) \log \sum_{k=1}^K p(a_k) p(b_j / a_k)$$

$$= - \sum_{k=1}^K \sum_{j=1}^J p(a_k) p(b_j / a_k) \log \sum_{k=1}^K p(a_k) p(b_j / a_k)$$

$$H(Y|X) = - \sum_{k=1}^K \sum_{j=1}^J p(a_k, b_j) \log p(b_j / a_k)$$

$$= - \sum_{k=1}^K \sum_{j=1}^J p(a_k) p(b_j / a_k) \log p(b_j / a_k)$$

可知对于确定的 Q ， $H(Y)$ 和 $H(Y|X)$ 都是 S 上的函数，且 $H(Y|X)$ 关于 P 是线性的。

下面将证明 $H(Y)$ 是 S 上的上凸函数。即对 $\forall P_1 = (p_{11}, p_{12}, \dots, p_{1K})$,

$P_2 = (p_{21}, p_{22}, \dots, p_{2K}) \in S$, 及 $\lambda, \bar{\lambda}, 0 \leq \lambda \leq 1, \bar{\lambda} = 1 - \lambda$. 成立

$$\begin{aligned}
 & - \sum_{k=1}^K \sum_{j=1}^J [\lambda p_{1k}(a_k) p(b_j/a_k) + \bar{\lambda} p_{2k}(a_k) p(b_j/a_k)] \log \sum_{k=1}^K [\lambda p_{1k}(a_k) + \bar{\lambda} p_{2k}(a_k)] p(b_j/a_k) \\
 & \geq \\
 & \quad - \lambda \sum_{k=1}^K \sum_{j=1}^J p_{1k}(a_k) p(b_j/a_k) \log \sum_{k=1}^K p_{1k}(a_k) p(b_j/a_k) \\
 & \quad - \bar{\lambda} \sum_{k=1}^K \sum_{j=1}^J p_{2k}(a_k) p(b_j/a_k) \log \sum_{k=1}^K p_{2k}(a_k) p(b_j/a_k) \tag{1}
 \end{aligned}$$

事实上, 首先看不等式左边:

$$\begin{aligned}
 & - \sum_{k=1}^K \sum_{j=1}^J [\lambda p_{1k}(a_k) p(b_j/a_k) + \bar{\lambda} p_{2k}(a_k) p(b_j/a_k)] \log \sum_{k=1}^K [\lambda p_{1k}(a_k) + \bar{\lambda} p_{2k}(a_k)] p(b_j/a_k) \\
 & = \\
 & \quad - \sum_{k=1}^K \sum_{j=1}^J [\lambda p_{1k}(a_k, b_j) + \bar{\lambda} p_{2k}(a_k, b_j)] \log \sum_{k=1}^K [\lambda p_{1k}(a_k) p(b_j/a_k) + \bar{\lambda} p_{2k}(a_k) p(b_j/a_k)] = \\
 & \quad = - \sum_{j=1}^J [\lambda p_{1j}(b_j) + \bar{\lambda} p_{2j}(b_j)] \log [\lambda p_{1j}(b_j) + \bar{\lambda} p_{2j}(b_j)] \tag{2}
 \end{aligned}$$

而不等式右边:

$$- \lambda \sum_{k=1}^K \sum_{j=1}^J p_{1k}(a_k) p(b_j/a_k) \log \sum_{k=1}^K p_{1k}(a_k) p(b_j/a_k)$$

-

$$\begin{aligned}
& -\bar{\lambda} \sum_{k=1}^K \sum_{j=1}^J p_{2k}(a_k) p(b_j/a_k) \log \sum_{k=1}^K p_{2k}(a_k) p(b_j/a_k) = \\
& = -\lambda \sum_{j=1}^J p_{1j}(b_j) \log p_{1j}(b_j) - \bar{\lambda} \sum_{j=1}^J p_{2j}(b_j) \log p_{2j}(b_j) \quad (3)
\end{aligned}$$

因为 $H(Y)$ 关于 Y 的分布是上凸函数，则成立下面不等式：

$$\begin{aligned}
& -\sum_{j=1}^J [\lambda p_{1j}(b_j) + \bar{\lambda} p_{2j}(b_j)] \log [\lambda p_{1j}(b_j) + \bar{\lambda} p_{2j}(b_j)] \\
& \geq -\lambda \sum_{j=1}^J p_{1j}(b_j) \log p_{1j}(b_j) - \bar{\lambda} \sum_{j=1}^J p_{2j}(b_j) \log p_{2j}(b_j)
\end{aligned}$$

所以，综合 (2)，(3) 式，(1) 式成立。即 $H(Y)$ 是 S 上的上凸函数，又知 $H(Y|X)$

关于 P 是线性的，所以 $I(P, Q) = I(X, Y) = H(Y) - H(Y|X)$ 是概率矢量空间 S 上的上凸函数。