1 基本概念

Shannon 信道模型

信道的构成 $\{\mathcal{X}, q(y|x), \mathcal{Y}\}$。

- 输入：符号集合 $\{1, 2, \ldots, M\}$,
- 编码函数 $X^n : \{1, 2, \ldots, M\} \to \mathcal{X}^n, X^n(i) \in \mathcal{X}^n$。称 $\{X^n(1), \ldots, X^n(M)\}$ 为码本 (codebook)。
- 译码函数 $g : \mathcal{Y}^n \to \{1, 2, \ldots, M\}$

信道编码 (M, n) 码。

- 条件错误概率

$$\begin{aligned} \lambda_i &= \Pr\{g(Y^n) \neq i | X^n = x^n(i)\} \\ &= \sum_{y^n} q(y^n | x^n(i)) I(g(y^n) \neq i) \end{aligned}$$

  其中 i 为输入符号，$I(\cdot)$ 为指示函数。

- 最大错误概率

$$\lambda^{(n)} = \max_{i \in \{1, \ldots, M\}} \lambda_i$$

- 算术平均错误概率

$$P_e^{(n)} = \frac{1}{M} \sum_{i=1}^{M} \lambda_i$$

令 $k = \log_{|\mathcal{X}|} M$，称为信息码元，码字长为 n，定义码率为 n 长的码字中实际有效携带信息的位数：$R = \frac{k}{n} \in [0, 1]$。

称码率为可达的，如果存在二元序列 $(\lceil 2^{nR} \rceil, n)$ 满足：$\lambda^{(n)} \to 0$, 则称码率是可达的。

2 典型设计

重复码，若每个输入符号重复 $2n+1$ 次，码率为 $R = \frac{1}{(2n+1)}$ 。考虑使用重复码编码 $(M = 2, n = 2n+1, \mathcal{X}, \mathcal{Y} = \{0, 1\})$，信道为二进制对称信道，误码率 $p_e = p < \frac{1}{2}$ 使用“多数判决原则”进行解码，发生错误为 $(2n+1)$ 次

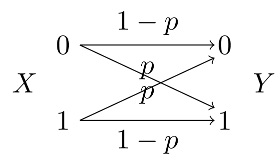

图 1: 二进制对称信道

传输有至少 $n+1$ 次传错

$$p_e = \sum_{i=n+1}^{2n+1} \binom{2n+1}{i} p^i (1-p)^{2n+1-i}$$

令 $q = \frac{p}{1-p} \in (0,1)$，则

$$\begin{aligned} p_e &= \frac{q^{n+1} \sum_{j=0}^{n} q^j \binom{2n+1}{n-j}}{(1+q)^{2n+1}} \\ &\le \frac{q^{n+1} \sum_{j=0}^{n} \binom{2n+1}{n-j}}{(1+q)^{2n+1}} \\ &= \frac{q}{1+q} \left(\frac{2\sqrt{q}}{1+q} \right)^{2n} \end{aligned}$$

因为 $q \neq 1$, $\frac{2\sqrt{q}}{1+q} < 1$ 所以当 $n \to \infty$ 时，误码率 $p_e \to 0$。

考虑二进制独立信源和对称信道，码率 R 和误码率 P_e 的关系为：

$$R \le \frac{1 - H(p)}{1 - H(P_e)}$$

3 信道编码定理

联合典型序列 (Jointly Typical Sequence)：设 $(X,Y) \sim p(x,y)$, 随机序列对 $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ 是联合典型序列，当 $x^n, y^n \in A_\epsilon^{(n)}$ 且 $|-\frac{1}{n} \log p(x^n, y^n) - H(X,Y)| < \epsilon$ 对给定的 ϵ 成立。

设 $\widetilde{X}^n$ 与 $\widetilde{Y}^n$ 独立，$\widetilde{X}^n \sim p(x^n), \widetilde{Y}^n \sim p(y^n)$ 可以证明

$$\Pr\{(\widetilde{X}^n, \widetilde{Y}^n) \in A_\epsilon^{(n)}\} \le 2^{-n(I(X;Y) - 3\epsilon)}$$

因为

$$\begin{aligned}\Pr\{(\widetilde{X}^n,\widetilde{Y}^n)\in A_\epsilon^{(n)}\} &= \sum_{(x^n,y^n)\in A_\epsilon^{(n)}} p(x^n,y^n)\\ &= \sum_{(x^n,y^n)\in A_\epsilon^{(n)}} p(x^n)p(y^n)\end{aligned}$$

由定义 $p(x^n)\le 2^{-n[H(x)-\epsilon]}, p(y^n)\le 2^{-n[H(Y)-\epsilon]}$ 所以

$$\begin{aligned}\Pr\{(\widetilde{X}^n,\widetilde{Y}^n)\in A_\epsilon^{(n)}\} &\le \sum_{(x^n,y^n)\in A_\epsilon^{(n)}} 2^{-n[H(x)+H(Y)-2\epsilon]}\\ &= 2^{-n[H(x)+H(Y)-2\epsilon]}|A_\epsilon^{(n)}|\end{aligned}$$

由典型集的性质 $|A_\epsilon^{(n)}|\le 2^{n[H(X,Y)+\epsilon]}\Rightarrow$

$$\Pr\{(\widetilde{X}^n,\widetilde{Y}^n)\in A_\epsilon^{(n)}\}\le 2^{-n[I(X;Y)-3\epsilon]}$$

4 Hamming 码

基本概念

所有码字 c_i 的集合记为一个码本 $C^{(n)}=\{c_i\}_{i=1}^M$

码重：二进制序列中 1 的个数，$w(c_i)$

最小码重：$w_{\min}=\min_i\{w(c_i)\}$ （全零码除外）。

Hamming 距离：两个二进制序列中对应位不相等的个数，$d(c_i,c_j)$。

最小 Hamming 距离：$d_{\min}=\min_{i\neq j}\{d(c_i,c_j)\}$。

4.1 Hamming(7,4)

4 位数据位，3 位校验位。$d_{\min}=w_{\min}=3$。码率 $R=\frac{4}{7}$。顺序为 $[p_1,p_2,d_1,p_3,d_2,d_3,d_4]$。其中

$$\begin{aligned}p_1&=d_1+d_2+d_4(\mathrm{mod}2)\\ p_2&=d_1+d_3+d_4(\mathrm{mod}2)\\ p_3&=d_2+d_3+d_4(\mathrm{mod}2)\end{aligned}$$

生成矩阵 G、解码矩阵 R 和校验矩阵 H 具有下面的形式:

$$G = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} R = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} H = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$