Chapter 3: 信道与信道容量

Xuan

2020 年 10 月 11 日

1 信道的基本概念

1.1 二进制离散信道 (BSC)

该信道模型的输入和输出信号的符号数都是 2，即 $X \in A = 0\,1$ 和 $Y \in B = 0,1$，转移概率为

$$\begin{aligned} p(Y=0|X=1) &= p(Y=1|X=0) = p \\ p(Y=1|X=1) &= p(Y=0|X=0) = 1-p \end{aligned} \tag{1}$$

1.2 加性高斯白噪声信道 (AWGN)

$$Y = X + G$$

G 是一个零均值、方差为 σ^2 的高斯随机变量，当 $X = a_i$ 给定后，Y 是一个均值为 a_i、方差为 σ^2 的高斯随机变量

$$p_Y(y|a_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-a_i)^2/2\sigma^2}$$

2 信道

信道可以看成是转移概率

对于信息 $M \in \mathcal{M}$，传输速率 $R = \frac{log(\mathcal{M})}{n}$，信道传输的过程即可表示为：

$$M \to x^n(M) \xrightarrow{p(y|x)} y^n \to \hat{M}$$

1. 设计一个方案，该方案可以达到某传输概率
2. 证明超出该传输速率无法传输 $\iff$ 能够传输的都不超过这个速率

2.1 1. 设计方案

2.1.1 典型序列

Define: x^n is (n, ε) typical, when $|N(x|x^n) - p(x)| < \varepsilon n$, for all $x \in \mathcal{X}$, where $N(x|x^n)$ is the empirical distribution.

$Pr(x^n istypical) \to 1$, when n $\to \infty$, which can be proved by Law of Large Numbers

2.2 典型集

Set $T(n, \varepsilon)$ 为典型序列的集合

1. $Pr(x^n \in T(n, \varepsilon)) \to 1$
2. $|T(n, \varepsilon)| \approx 2^{nH(x)}$

   2.1 $x^n \in T(\varepsilon, n), p(X^n = x^n) \approx 2^{-nH(x)}$

2.2.1 Proof

$$p(x^n) = \prod_{i=1}^{n} p(x_i) = \prod_{x \in \mathcal{X}} p(x)^{N(x|x^n)n} (*)$$

$\Rightarrow$ 所有典型序列的概率都差不多大

$$log(*) = \sum_{x \in \mathcal{X}} nN(x|x^n) logp(x) \approx n \sum_{x \in \mathcal{X}} p(x) logp(x) = -nH(x)$$

Pr(x^n: x^n is typical) = * $\approx 2^{-nH(x)}$

2.3 典型集和散度的关系

$Pr(x^n \in T(n,\varepsilon)) =?$
$x^n \in T(n,\varepsilon), N(x|x^n)\ p(x)$

$$\begin{aligned} Pr(x^n) &= \prod q(x)^{nN(x|x^n)} \\ &\approx \prod q(x)^{np(x)} \\ &= 2^{log \prod q(x)^{np(x)}} \\ &= 2^{np(x) \sum logq(x)} \\ &= 2^{-np(x)log\frac{1}{q(x)}} \end{aligned} \tag{2}$$

$$\begin{aligned} Pr(T(n,\varepsilon)) &= \sum_{x^n \in T(\varepsilon,n)} Pr(x^n) \\ &\approx 2^{-np(x)logp(x)} * 2^{-np(x)log\frac{1}{q(x)}} \\ &= 2^{-nD(p||q)} \end{aligned} \tag{3}$$

2.4 条件典型集

典型条件集的大小：$2^{nH(Y|X)}$

3 随机码簿 (Random Codebook)

$X - Y, p(y|x)$，n 长编码，$|M| \approx 2^{nI(X;Y)}$
具体步骤如下：

1. 生成码簿：给定一个概率 $p(x)$，按 $p(x)$ $i.i.d.$ 生成 $x^n(i)$ 序列，重复 $2^{nI(X;Y)}$ 次
2. 发送 $x^n(M)$
3. 解码：y^n 查表寻找 $(x^n(i), y^n)$ 联合典型
4. 错误率分析：
   (a) 错误一：发送 x^n，但收到的 y^n 和 x^n 不联合典型

(b) 错误二：发送 x^n，收到 y^n，但存在 $(x^n)'$ 与 y^n 联合典型

$$p(mistake1 \cup mistake2) \leq p(mistake1) + p(mistake2)$$

1. $p(mistake1) \to 0$, when $n \to \infty$(由大数定律可知)

2. $p(mistake2) \to 0$, when $|M| \leq 2^{nI(X;Y)}$

$$\begin{aligned} p(mistake2) &\leq \sum_{x^n \neq (x^n)'} p((x^n)', y^n) \\ &= 2^{n[I(X;Y)-\varepsilon]} - 2^{-nI(X;Y)} \end{aligned} \tag{4}$$

4 Converse theory of channel capacity

假设 $|M|$ 等概率发生

$$\begin{aligned} R &= \frac{log|M|}{n} \\ nR &= log|M| = H(M) \\ &= H(M|Y^n) + I(M;Y^n) \end{aligned} \tag{5}$$

其中

$$\begin{aligned} I(M;Y^n) &= I(X^n;Y^n) \\ &= H(Y^n) - H(Y^n|X^n) \\ &\leq \sum_{i=0}^{n} H(y_i) - H(Y^n|X^n) \\ &= \sum H(y_i) - \sum H(y_i|X^n, y_1, ..., y_n)(ChainRule) \\ &= \sum H(y_i) - \sum H(y_i|x_i) \\ &= \sum I(x_i;y_i) \\ &\leq nmax_{p(x_i)} I(X;Y) \end{aligned} \tag{6}$$

4.1 Fano's Inequality

For Markov Chain, $X \to Y \to \hat{X}$

$$H(X|\hat{X}) \leq H(Pe) + log(|X| - 1) \leq 1 + log(|H| - 1)$$

where Pe is the probability of making errors, $H(Pe)$ obey 0-1 distribution, such that $H(Pe) \leq 1$

4.1.1 Proof of Fano's Inequality

Import Indicator Variable E, where has the property:

$$E = \begin{cases} 0, X = \hat{X} \\ 1, X \neq \hat{X} \end{cases}$$

such that $H(E|X, \hat{X}) = 0$

$$\begin{aligned} H(X|\hat{X}) &= H(X|\hat{X}) + H(E|X, \hat{X}) \\ &= H(E, X|\hat{X}) \\ &= H(X|E, \hat{X}) + H(E|\hat{X}) \\ &= (Pe)H(X|E = 1, \hat{X}) + (1 - Pe)H(X|E = 0, \hat{X}) + H(E|\hat{X}) \\ &\leq (Pe)log(|X| - 1) + H(Pe) \end{aligned} \tag{7}$$

$$H(X|E = 0, \hat{X}) = 0$$

4.2 $R \leq maxI(X;Y)$

From equation (5), together with equation (6) and (7), we can get

$$\begin{gathered} H(M|Y^n) + I(M;Y^n) \leq (Pe)log(|M|) + nmaxI(X;Y) \\ (1 - Pe)R \leq maxI(X;Y) \end{gathered} \tag{8}$$

Because $Pe \to 0$, so that we can conclude that $R \leq max_{p(x_i)}I(X;Y)$

5 Extension: 一阶马尔可夫链的概率

$X_1 - X_2 - ... - X_n$,

$$P_{x_n}|P_{x_1} = P_{x_n}|P_{x_{n-1}}...P_{x_2}|P_{x_1} = (P_{x_{i+1}}|P_{x_i})^{n-1}$$

稳态：转移矩阵的特征向量

6 微分熵的性质

1. $h(x) = h(x + C)$，C 是常数
2. $h(aX) = h(X) + log|a|$
   (a) $h(A\vec{x}) = h(\vec{x}) + log|A|$

6.1 性质 2 的证明

令 $Y = aX$，则有 $dy = adX$，故 $P(Y \le y) = P(X \le \frac{y}{a})$

$$P_Y(y) = \frac{dP(Y \le y)}{dy} = \frac{dP(X \le \frac{y}{a})}{d\frac{y}{a}} * \frac{1}{a} = f_x(\frac{y}{a}) * \frac{1}{a}$$

$$\begin{aligned} -\int p(y)logp(y)dy &= -\int \frac{1}{a}f_x(\frac{y}{a})log\frac{1}{a}f_x(\frac{y}{a})dy \\ &= -\int f_x(\frac{y}{a})log\frac{1}{a}f_x(\frac{y}{a})d\frac{y}{a} \\ &= -\int f_x(x)log\frac{1}{a}f_x(x)dx \\ &= -\int f_x(x)logf_x(x)dx - \int f_x(x)log\frac{1}{a}dx \\ &= -\int f_x(x)logf_x(x)dx - log|a| \end{aligned} \tag{9}$$

7 连续信道

Example 对于 $X \sim N(0, \sigma^2)$，有 $h(x) = \frac{1}{2}log(2\pi e\sigma^2)$.
则对于 $aX \sim N(0, a^2\sigma^2)$，有 $h(ax) = \frac{1}{2}log(2\pi ea^2\sigma^2) = \frac{1}{2}log(2\pi e\sigma^2) + \frac{1}{2}loga^2 = \frac{1}{2}log(2\pi e\sigma^2) + loga$

7.1 加性噪声信道 (Additative Noise Channel)

x 与 z 相互独立
$p(y) = p(x + z)$
$P(Y \le y) = P(X + Z \le y) = \int_{X+Z\le y} p(x, z)dxdz = \int_{X+Z\le y} p(x)p(z)dxdz$

1. 对任何 X,Z 是高斯分布时，最小化信道容量 $I(X;Y)$，因为该情况下混乱程度最大（Proof: Entropy Power Inequality）

2. 在方差一定的情况下，如果 Z 是高斯分布，则 X 也是高斯分布时，最大化信道容量

7.1.1 证明第二条

$$\begin{aligned} I(Y;X) &= I(X+Z;X) = H(x+z) - H(x+z|x) \\ &= H(x+z) - H(z) \end{aligned} \tag{10}$$

$D(x), Var(x)$ 为给定值的情况下，$Var(x+z) = Var(x) + Var(z)$ 为定值
因为当 X+Z 为高斯分布时，H(X+Z) 取到最大值，由于 Z 为高斯分布，故 X 也服从高斯分布

7.2 信道容量不等式

$X \sim N(0,P), Z \sim N(0,N)$

$$\begin{aligned} C &= I(X;Y) = h(x+z) - h(z) \\ &= \frac{1}{2}log(2\pi e(P+N)) - \frac{1}{2}log(2\pi eN) \\ &= \frac{1}{2}(\frac{2\pi e(P+N)}{2\pi eN}) \\ &= \frac{1}{2}log(\frac{P+N}{N}) \\ &= \frac{1}{2}log(1+\frac{P}{N}) \end{aligned} \tag{11}$$

信噪比 SNR：$\frac{P}{N}$

7.3 注水法 (Water-Filling)

$x_1, ..., x_N$ 相互独立,$Var(z_i) = N_i$,$\sum Var(X_i) = P_s$,问 $max \sum_{i=1}^{n} I(X_i;Y_i)$
$\frac{1}{2}log(1+\frac{P_i}{N_i}) \Leftrightarrow max \sum_{i=1}^{n} \frac{1}{2}log(1+\frac{P_i}{N_i}), s.t. P_1 + ... + P_N = P_s$

$L(\vec{P}) = \sum_{i=1}^{n} \frac{1}{2} log(1 + \frac{P_i}{N_i}) - \lambda(\sum_{i=1}^{n} P_i - P_s)$

$$\begin{aligned} \frac{\partial L}{\partial P_i} &= \frac{\partial \sum_{i=1}^{n} \frac{1}{2} log(1 + \frac{P_i}{N_i}) - \lambda(\sum_{i=1}^{n} P_i - P_s)}{\partial P_i} \\ &= \frac{1}{2} \frac{1}{P_i + N_i} - \lambda & = 0 \end{aligned} \tag{12}$$

As a result, $P_i + N_i = \frac{1}{2\lambda}$, so that $\sum P_I + N_i = \frac{n}{2\lambda}$