信息论基础

李　莹
liying2009@ecust.edu.cn

第四章：信道及信道容量

一、信道分类

二、离散单符号信道及其信道容量

三、离散多符号信道及其信道容量

四、组合信道及其信道容量

4. 离散对称信道的信道容量

定义1： 若信道矩阵**P**中每行都是第一行的排列，则称此信道是**行对称信道**。

$$H(Y|X)=-\sum_i p(x_i)H(Y|x_i)$$
$$=-\sum_i p(x_i)\sum_j p(y_j|x_i)\log p(y_j|x_i)$$
$$=-\sum_j p(y_j|x_i)\log p(y_j|x_i)=H(Y|x_i)$$

$$\mathbf{P}=\begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6}\\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3}\end{bmatrix}$$

$$H(Y|X)=p(x_1)H(\frac{1}{3},\frac{1}{3},\frac{1}{6},\frac{1}{6})+p(x_2)H(\frac{1}{6},\frac{1}{3},\frac{1}{6},\frac{1}{3})=H(\frac{1}{6},\frac{1}{3},\frac{1}{6},\frac{1}{3})$$

定义2：若信道矩阵中每行都是第一行的排列，并且每列都是第一列的排列，则称之为**对称信道**。

$$\mathbf{P} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} \end{bmatrix}$$

定义3：虽然不是对称信道，但是信道矩阵可以按列分为一些对称的子阵，则称之为**准对称信道**。

$$\mathbf{P}=\begin{bmatrix}0.8 & 0.1 & 0.1\\ 0.1 & 0.1 & 0.8\end{bmatrix}$$

$$\Rightarrow\quad \mathbf{P}_1=\begin{bmatrix}0.8 & 0.1\\ 0.1 & 0.8\end{bmatrix}\qquad \mathbf{P}_2=\begin{bmatrix}0.1\\ 0.1\end{bmatrix}$$

$$\mathbf{P}=\begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6}\\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3}\end{bmatrix}$$

定义4： 若$r=s$，且对于每一个输入符号，正确传输概率都相等，且错误传输概率 p 均匀地分配到 r-1 个符号，则称此信道为**强对称信道或均匀信道**。

$$\mathbf{P}=\begin{bmatrix} \bar{p} & \frac{p}{r-1} & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \frac{p}{r-1} & \bar{p} & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{p}{r-1} & \frac{p}{r-1} & \frac{p}{r-1} & \cdots & \bar{p} \end{bmatrix}$$

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} \end{array}$$

强对称信道具备四个特征：

1. 矩阵中的每一行都是第一行的排列；(行对称)

   矩阵中的每一列都是第一列的排列。(列对称)

2. 信道输入与输出消息（符号）数相等，即 $r=s$。

3. 错误分布是均匀的：信道矩阵中正确传输概率都相等，且错误传输概率均匀地分配到r-1个符号上。

4. 不仅每一行元素之和为1，每一列元素之和也为1。

显然，对称性的基本条件是1，而2、3、4是加强条件。

放松对信道的约束，仅满足条件1，就构成一般的对称信道。

例1：

$$\mathbf{P}=\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

例2：

$$\mathbf{P}=\begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$$

再进一步放松条件，信道矩阵按列分成若干子阵，如果子阵是对称的，则称为准对称信道。

$$\mathbf{P}=\begin{pmatrix}1-\varepsilon_1-\varepsilon_2 & \varepsilon_2 & \varepsilon_1\\ \varepsilon_2 & 1\text{-}\varepsilon_1-\varepsilon_2 & \varepsilon_1\end{pmatrix}=\left(\begin{array}{cc|c}1-\varepsilon_1-\varepsilon_2 & \varepsilon_2 & \varepsilon_1\\ \varepsilon_2 & 1\text{-}\varepsilon_1-\varepsilon_2 & \varepsilon_1\end{array}\right)=\left(\mathbf{P}_1 \ \vdots\ \mathbf{P}_2\right)$$

$$\mathbf{P}=\begin{bmatrix}1-p-q & q & p\\ p & q & 1-p-q\end{bmatrix}$$

定理4.1 对于对称信道，当信道输入概率分布为等概分布时，输出概率分布**必**为等概分布。

证明：当输入为等概分布时 $p(x_i)=\frac{1}{r}, i=1,2,\cdots,r$

则输出 $p(y_j)=\sum_i p(x_i)p(y_j|x_i)=\frac{1}{r}\sum_i p(y_j|x_i)=\frac{1}{r}H_j,$

其中 $H_j=\sum_{i=1}^{r}p(y_j|x_i)$ H_j为信道矩阵第j列元素之和。

而对称信道每一列是第一列的不同排列。因此

$$H_1=H_2=\cdots=H_s=H$$

又因为 $sH=r\cdot 1$ $p(y_j)=\frac{1}{r}H=\frac{1}{s}$

即当信道输入为等概分布时，输出亦为等概分布。

定理4.2 **对称信道** 当信道**输出**概率分布为等概的情况下达到信道容量：

$$C=\log s-H(p_1',p_2',\cdots,p_s')$$

其中 $\boldsymbol{p_1',p_2',\cdots,p_s'}$ 是信道矩阵中的任意一行中的元素。

证明：

$$H(Y|X)=H(Y|x_i)=H(p_1',p_2',\cdots,p_s')$$

$$I(X;Y)=H(Y)-H(Y|X)=H(Y)-H(p_1',p_2',\cdots,p_s')$$

$$C=\max_{p(x)}I(X;Y)$$

$$=\max_{p(x)}\left\{H(Y)-H(p_1',p_2',\cdots,p_s')\right\}$$

$$\boxed{=\max_{p(x)}H(Y)-H(p_1',p_2',\cdots,p_s')}=\log s-H(p_1',p_2',\cdots,p_s')$$

推论：对于强对称信道有：$C=\log r - p\log(r-1) - H(p)$

$$
\begin{aligned}
C &= \log s - H(p_1', p_2', \cdots, p_s') \\
&= \log r - H(\bar{p}, \frac{p}{r-1}, \cdots, \frac{p}{r-1}) \\
&= \log r + \bar{p}\log\bar{p} + \frac{p}{r-1}\log\frac{p}{r-1} \cdots + \frac{p}{r-1}\log\frac{p}{r-1} \\
&= \log r + \bar{p}\log\bar{p} + p\log\frac{p}{r-1} \\
&= \log r - p\log(r-1) + \bar{p}\log\bar{p} + p\log p \\
&= \log r - p\log(r-1) - H(p)
\end{aligned}
$$

$s = r$

例4.2 求对称信道的信道容量，

$$\mathbf{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{3} & \frac{1}{6}\\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2}\end{bmatrix}$$

解：

$$C=\log s-H(p_1',p_2',\cdots,p_s')$$

$$=\log 3-H(\frac{1}{2},\frac{1}{3},\frac{1}{6})$$

$$=0.126 \text{ 比特/符号}$$

准对称信道

$$C = \max_{p(x)} H(Y) - H(p_1', p_2', \cdots, p_s')$$

$$C \le \log s - H(p_1', p_2', \cdots, p_s')$$

当信道输入概率分布为等概的情况下达到信道容量：

$$C = \log r - H(p_1', p_2', \cdots, p_s') - \sum_{k=1}^{n} N_k \log M_k$$

设信道矩阵可划分为n个子矩阵，其中N_k是第k个子矩阵中行元素之和，M_k是第k个子矩阵中列元素之和。

例4.3：求准对称信道的信道容量。二元对称删除信道：

$$\mathbf{P} = \begin{bmatrix} 1-p-q & q & p \\ p & q & 1-p-q \end{bmatrix}$$

解： N_1=1-q, M_1=1-q, N_2=q, M_2=2q

$$C = \log r - H(p_1', p_2', \cdots, p_s') - \sum_{k=1}^{n} N_k \log M_k$$

$$= \log 2 - H(1-p-q, q, p) - (1-q)\log(1-q) - q\log(2q)$$

5. 一般离散信道的信道容量

信道容量

约束条件：

$$\begin{cases} \sum_{i} p(x_i) = 1 \\ p(x_i) \geq 0, \qquad i = 1,\ 2,\ \cdots,\ r \end{cases}$$

求信道容量转化为求 $I(X;Y)$ **对信源概率分布** $P(X)$**的条件极值。**

解：　引入辅助函数

$$F=I(X;Y)-\lambda\left[\sum_{i=1}^{r}p(x_i)-1\right]$$　（λ为待定系数）

$$\frac{\partial F}{\partial p(x_i)}=\frac{\partial}{\partial p(x_i)}\left\{I(X;Y)-\lambda\left[\sum_{i=1}^{r}p(x_i)-1\right]\right\}\qquad (i=1,..,r)$$

$$=\frac{\partial I(X;Y)}{\partial p(x_i)}-\lambda$$

$$I(X;Y)=\sum_{i=1}^{r}\sum_{j=1}^{s}p(x_i)p(y_j\mid x_i)\log\frac{p(y_j\mid x_i)}{p(y_j)}$$

$$=\sum_{i=1}^{r}p(x_i)\sum_{j=1}^{s}p(y_j\mid x_i)\log p(y_j\mid x_i)-\sum_{j=1}^{s}p(y_j)\log p(y_j)$$

$$p(y_j)=\sum_{i=1}^{r}p(x_i)p(y_j\mid x_i)\quad\Rightarrow\quad\frac{\partial}{\partial p(x_i)}p(y_j)=p(y_j\mid x_i)$$

$$\Rightarrow\quad\frac{\partial\log p(y_j)}{\partial p(x_i)}=\frac{p(y_j\mid x_i)}{p(y_j)}\log e$$

$$\frac{\partial}{\partial p(x_i)}I(X;Y)$$

$$=\sum_{i=1}^{s}p(y_j\mid x_i)\log p(y_j\mid x_i)-\sum_{i=1}^{s}p(y_j\mid x_i)\log p(y_j)-\sum_{j=1}^{s}p(y_j\mid x_i)\log e$$

$$=\sum_{j=1}^{s}p(y_j\mid x_i)\log\frac{p(y_j\mid x_i)}{p(y_j)}-\log e$$

令 $\dfrac{\partial F}{\partial p(x_i)} = 0$

则 $\dfrac{\partial F}{\partial p(x_i)} = \sum_{j=1}^{s} p(y_j \mid x_i) \log \dfrac{p(y_j \mid x_i)}{p(y_j)} - \log e - \lambda = 0$

$$\sum_{j=1}^{s} p(y_j \mid x_i) \log \frac{p(y_j \mid x_i)}{p(y_j)} = \lambda + \log e \qquad i = 1, 2, \cdots, r$$

$$\sum_{i=1}^{r} p(x_i) \sum_{j=1}^{s} p(y_j \mid x_i) \log \frac{p(y_j \mid x_i)}{p(y_j)} = \sum_{i=1}^{r} p(x_i)(\log e + \lambda)$$

$$\Longrightarrow \quad C = \log e + \lambda$$

在某些条件下利用这个方法可以计算C:

$$\sum_j p(y_j|x_i)\log\frac{p(y_j|x_i)}{p(y_j)}=\log e+\lambda=C$$

$$\sum_j p(y_j|x_i)\log p(y_j|x_i)=\sum_j p(y_j|x_i)\log p(y_j)+C$$

$$=\sum_j p(y_j|x_i)[\log p(y_j)+C]$$

令 $\beta_j=\log p(y_j)+C$

$$\sum_j p(y_j|x_i)\beta_j=\sum_j p(y_j|x_i)\log p(y_j|x_i)$$

$$\mathbf{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4}\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2}\end{bmatrix}$$

这是一个含有s个未知数、由r个方程组成的方程组。
当$r=s$，且信道矩阵是可逆矩阵时，该方程组有唯一解。

$$\Rightarrow \quad \beta_j \qquad j=1,2,\cdots,s$$

$$\beta_j = \log p(y_j) + C \Rightarrow \begin{matrix} p(y_j) = 2^{\beta_j - C} \\ \sum_j p(y_j) = 1 \end{matrix} \Rightarrow \sum_j 2^{\beta_j - C} = 1$$

$$\Rightarrow \quad C = \log \sum_j 2^{\beta_j} \quad \Rightarrow \quad \boxed{p(y_j) = 2^{\beta_j - C} \qquad j=1,2,\cdots,s}$$

$$p(y_j) = \sum_i p(x_i) p(y_j \mid x_i) \Rightarrow \boxed{p(x_i) \qquad (i=1,2,\cdots,r)}$$

例4.5：求以下信道的信道容量。

信道矩阵

$$\mathbf{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4}\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2}\end{bmatrix}$$

解：
$$\begin{cases} \frac{1}{2}\beta_1 + \frac{1}{4}\beta_2 + \frac{1}{4}\beta_4 = \frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{4}\log\frac{1}{4} \\ \beta_2 = 0 \\ \beta_3 = 0 \\ \frac{1}{4}\beta_1 + \frac{1}{4}\beta_3 + \frac{1}{2}\beta_4 = \frac{1}{4}\log\frac{1}{4} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{2}\log\frac{1}{2} \end{cases}$$

$$\beta_2 = \beta_3 = 0, \quad \beta_1 = \beta_4 = -2$$

$$C = \log\sum_j 2^{\beta_j} = \log(2^{-2} + 2^0 + 2^0 + 2^{-2}) = \log 5 - 1 \text{ 比特/符号}$$

$$p(y_j) = 2^{\beta_j - C} \qquad j = 1, 2, \cdots, s$$

$$p(y_1) = p(y_4) = 2^{-2-\log 5+1} = \frac{1}{10} \qquad p(y_2) = p(y_3) = 2^{0-\log 5+1} = \frac{4}{10}$$

$$p(y_j) = \sum_i p(x_i) p(y_j | x_i) \qquad \begin{cases} p(y_1) = \frac{1}{2} p(x_1) + \frac{1}{4} p(x_4) \\ p(y_2) = \frac{1}{4} p(x_1) + p(x_2) \\ p(y_3) = p(x_3) + \frac{1}{4} p(x_4) \\ p(y_4) = \frac{1}{4} p(x_1) + \frac{1}{2} p(x_4) \end{cases}$$

$$p(x_1) = p(x_4) = \frac{4}{30},$$

$$p(x_2) = p(x_3) = \frac{11}{30}$$

例：　有一信道矩阵 $\begin{bmatrix} 1 & 0 \\ \varepsilon & 1-\varepsilon \end{bmatrix}$，求$C$.

补充：

1）采用上述方法求出信道容量以后，还必须解出 $p(x_i)$，因为在采用拉格朗日数乘法时并没有加上 $p(x_i)\geq 0$ 的约束条件，因此算出的$p(x_i)$可能是负值。

当计算结果为负值时，此解无效。它表明最大值在边界上，即某些输入符号的概率为0。设某些输入符号的概率为0，然后重新进行计算。

2）如果r=2，则可以直接对I(X;Y)求导，得到信道容量和最佳输入分布。

例4.4：已知信道的转移矩阵为 $\mathbf{P}=\begin{bmatrix}0.5 & 0.3 & 0.2\\ 0.3 & 0.5 & 0.2\end{bmatrix}$

求信道容量。

解：设输入概率分布 $p(x_1)=\alpha$，$p(x_2)=1-\alpha$

$$\mathbf{P}_Y=\mathbf{P}_X\mathbf{P}_{Y|X}=\begin{bmatrix}\alpha & 1-\alpha\end{bmatrix}\begin{bmatrix}0.5 & 0.3 & 0.2\\ 0.3 & 0.5 & 0.2\end{bmatrix}$$

$$=\begin{bmatrix}0.3+0.2\alpha & 0.5-0.2\alpha & 0.2\end{bmatrix}$$

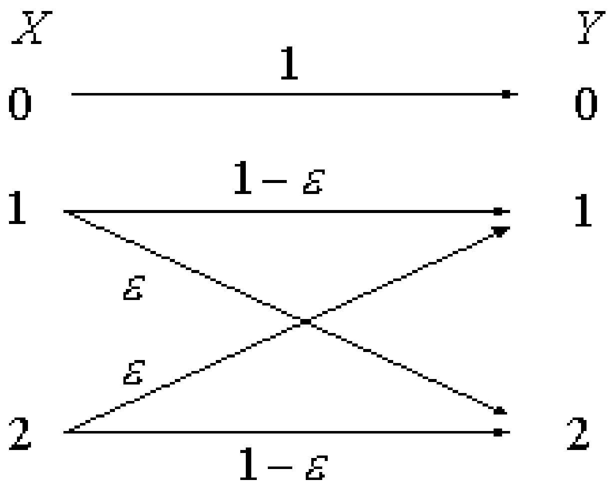
$$I(X;Y)=H(Y)-H(Y\mid X)$$
$$=-\sum_j p(y_j)\log p(y_j)-H(Y\mid x_i)$$
$$=-(0.3+0.2\alpha)\log(0.3+0.2\alpha)-(0.5-0.2\alpha)\log(0.5-0.2\alpha)$$
$$-0.2\log 0.2+0.5\log 0.5+0.3\log 0.3+0.2\log 0.2$$

$$\frac{\partial I(X;Y)}{\partial \alpha}=0$$

$$0.2\log(0.3+0.2\alpha)-0.2+0.2\log(0.5-0.2\alpha)+0.2=0$$

$$\alpha=1/2 \qquad C=\max I(X;Y)=0.036$$

练习：信道及它的输入、输出如图所示：



(1)　求最佳输入分布。

(2)　求 $\varepsilon = 0,\ \dfrac{1}{2}$ 时的信道容量。

6. 信道容量定理

定理4.3　$I(X;Y)$ 达到信道容量的充要条件是输入分布 $p(x_i)$ 满足以下充要条件：

$p(x_i)\neq 0$ 时　　$I(x_i;Y) = C$

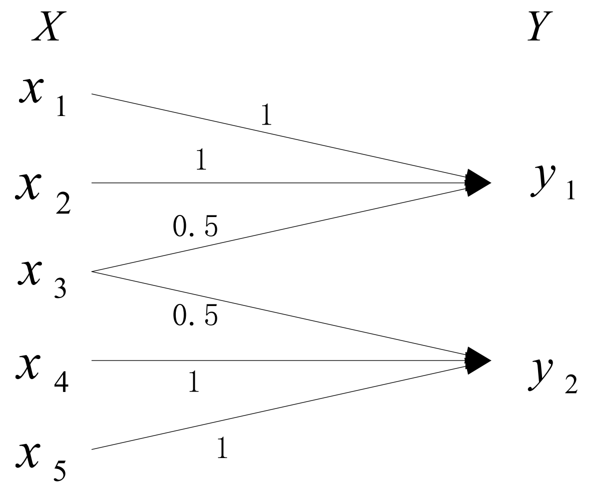
$p(x_i) = 0$ 时　　$I(x_i;Y)\leq C$

$$I(x_i;Y)=\sum_j p(y_j\mid x_i)\log\frac{p(y_j\mid x_i)}{p(y_j)}$$

某些特殊矩阵可以利用这个方法可以推导得到C。

例4.7

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



$p(x_3)=0,\quad p(x_2)=p(x_4)=0,$
$p(x_1)=p(x_5)=1/2$

$$I(x_1;Y)=I(x_5;Y)=\log 2$$
$$I(x_2;Y)=I(x_4;Y)=0$$
$$I(x_3;Y)=0$$

或者

$p(x_3)=0,\quad p(x_2)=p(x_4)$

$=p(x_1)=p(x_5)=1/4$

$$I(x_1;Y)=I(x_5;Y)=\log 2$$
$$I(x_2;Y)=I(x_4;Y)=\log 2$$
$$I(x_3;Y)=0$$

例 4.6　当输入等概时准对称信道达到信道容量。

$$p(x_i)=\frac{1}{r} \qquad I(x_i;Y)=\sum_{j=1}^{s} p(y_j\,|\,x_i)\log\frac{p(y_j\,|\,x_i)}{p(y_j)}$$

$$=\sum_{j=1}^{s} p(y_j\,|\,x_i)\log\frac{p(y_j\,|\,x_i)}{\sum_{k=1}^{r} p(x_k)p(y_j\,|\,x_k)}$$

$$=\sum_{j=1}^{s} p(y_j\,|\,x_i)\log\frac{p(y_j\,|\,x_i)}{\frac{1}{r}\sum_{k=1}^{r} p(y_j\,|\,x_k)}$$

在同一子阵$\mathbf{P}_l$中

$$p(y_j)=\frac{1}{r}\sum_{k=1}^{r}p(y_j\,|\,x_k) \quad \text{相等}$$

对于不同的 x_i ，

$$\sum_{y_j\in Y_l} p(y_j\,|\,x_i)\log\frac{p(y_j\,|\,x_i)}{\frac{1}{r}\sum_{k=1}^{r}p(y_j\,|\,x_k)} \quad \text{相等}$$

所以，对于任意 x_i

$$I\left(x_i;Y\right)=\sum_{l}\sum_{y_j\in Y_l}p(y_j\,|\,x_i)\log\frac{p(y_j\,|\,x_i)}{\frac{1}{r}\sum_{k=1}^{r}p(y_j\,|\,x_k)}$$ 相等

即当输入等概时，准对称信道达到信道容量。