信息理论

第三部分：信息的传输

余官定　教授

浙江大学
信息与电子工程学院

引言

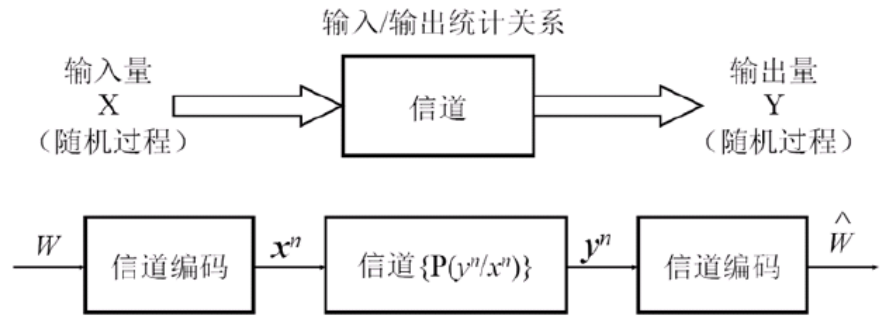
研究信道容量对于通信有及其重要的意义

- 通信的数学理论：功率和带宽等通信资源的效用极限
- 随机编码、联合译码、注水法则等重要思想
- 具体通信体制(CDMA/OFDM/MIMO)的理论基础

信息论中的信道模型是对实际物理信道的一个数学抽象，实际物理信道：噪声信道、干扰信道、无线衰落信道、存储信道等等。实际物理信道类型很多，建立信道的数学模型非常重要。

信道的抽象数学模型

$\{\mathcal{X}; p(y \mid x); \mathcal{Y}\}$



- 离散无记忆信道 $p_N\left(y^N \mid x^N\right) = \prod_{n=1}^{N} p(y_n \mid x_n), \forall N$
- 平稳信道 $p\left(y_n = j \mid x_n = k\right) = p\left(y_m = j \mid x_m = k\right), \forall m, n$

信道的分类

- **按输入/输出信道在幅度和时间上取值**
  - 幅度离散，时间离散信道
  - 幅度连续，时间离散信道
  - 幅度连续，时间连续信道
  - 幅度离散，时间连续信道
- **按输入/输出之间的记忆性**
  - 有记忆信道
  - 无记忆信道
- **按输入/输出之间关系的确定性**
  - 确定信道
  - 随机信道

信道容量

信道容量为输出序列对不同输入序列所提供的最大互信息:

$$C = \lim_{n\to\infty} \frac{1}{n} \max_{Q(x^n)} I\left(X_1X_2\cdots X_N; Y_1Y_2\cdots Y_N\right)$$

平均每次利用信道，在输入和输出符号之间所能相互提供的互信息的最大值的极限

该信道容量定义对所有形式信道均成立

离散无记忆信道容量(1)

对于离散无记忆信道(DMC)

$$I\left(X_1X_2\cdots X_N;Y_1Y_2\cdots Y_N\right)\le\sum_{n=1}^{N}I\left(X_n;Y_n\right)$$

其中等号在输入为独立随机序列时达到。
因此

$$\begin{aligned}C&=\lim_{n\to\infty}\frac{1}{n}\max_{Q(x^n)}I\left(X_1X_2\cdots X_N;Y_1Y_2\cdots Y_N\right)\\&=\max_{\{Q_k\}}I(X;Y)\end{aligned}$$

离散无记忆信道容量(2)

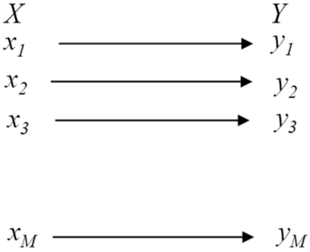
$$\begin{aligned} H(\mathbf{Y}) &= H\left(Y_1, Y_2, \cdots Y_n\right) \\ &= H\left(Y_1\right) + H\left(Y_2 \mid Y_1\right) + \cdots + H\left(Y_n \mid Y_{n-1} \cdots Y_1\right) \\ &\leq \sum_{i=1}^{n} H\left(Y_i\right) \end{aligned}$$

$$\begin{aligned} H(\mathbf{Y} \mid \mathbf{X}) &= -E\left\{\log p\left(y^n \mid x^n\right)\right\} = -E\left\{\log \prod_{i=1}^{n} p\left(y_i \mid x_i\right)\right\} \\ &= -\sum_{i=1}^{n} E\left\{\log p\left(y_i \mid x_i\right)\right\} = \sum_{i=1}^{n} H\left(Y_i \mid X_i\right) \end{aligned}$$

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y} \mid \mathbf{X}) \\ &\leq \sum_{i=1}^{n}\left\{H\left(Y_i\right) - H\left(Y_i \mid X_i\right)\right\} = \sum_{i=1}^{n} I\left(X_i, Y_i\right) \end{aligned}$$

“=”在输出Y为独立分布时成立。当X独立分布时，Y也独立分布。

DMC容量例子——无噪信道

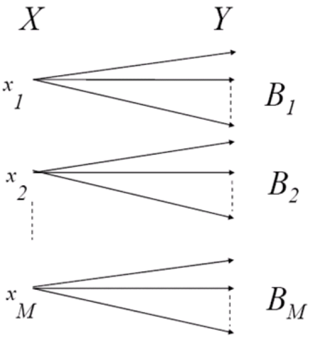


$$H(X \mid Y) = 0$$

$$I(X;Y) = H(X)$$

$$C = \max_{\{Q_k\}} I(X;Y) = \max_{\{Q_k\}} H(X) = \log M \text{ bit}$$

DMC容量例子——无损信道



$$H(X \mid Y) = 0$$

$$I(X;Y) = H(X)$$

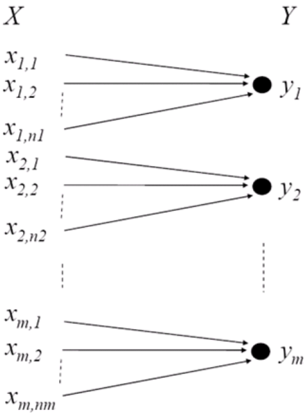
$$C = \max_{\{Q_k\}} I(X;Y)$$

$$= \max_{\{Q_k\}} H(X)$$

$$= \log M \text{ bit}$$

译码方法：？

DMC容量例子——确定信道



$$p(y_j \mid x_i) = 0 \text{ or } 1$$

$$I(X;Y) = H(Y) - H(Y \mid X) = H(Y)$$

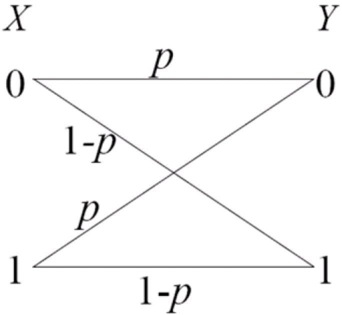
$$C = \max_{\{Q_k\}} I(X;Y)$$

$$= \max_{\{Q_k\}} H(Y)$$

$$= \log m \text{ bit}$$

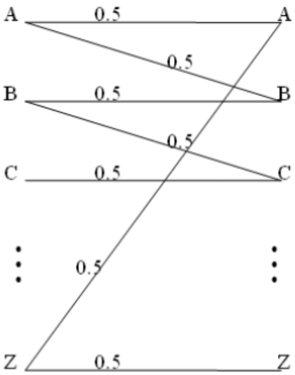
编码方法?

DMC容量例子——无用信道



$$p(y_j \mid x_i) = p(y_j)$$
$$p(x_i \mid y_j) = p(x_i)$$
$$H(X \mid Y) = H(X)$$
$$I(X;Y) = 0$$
$$C \equiv 0$$

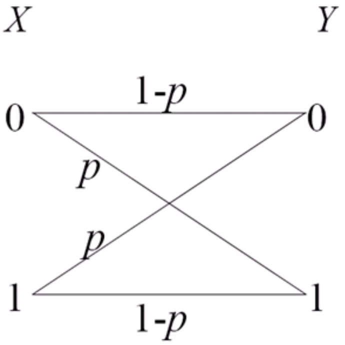
DMC容量例子——复制信道



$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y \mid X) = H(Y) - 1 \\
C &= \max_{\{Q_k\}} I(X;Y) = \max_{\{Q_k\}} H(Y) - 1 \\
&\leq \log 26 - 1 = \log 13
\end{aligned}
$$

编码方法？

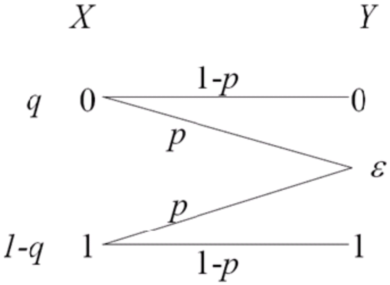
DMC容量例子——二元对称信道(BSC)



$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y \mid X) \\
&= H(Y) - \sum_x p(x) H(Y \mid X = x) \\
&= H(Y) - \sum_x p(x) H(p) \\
&= H(Y) - H(p) \\
&\leq 1 - H(p)
\end{aligned}
$$

当输入取等概分布时，输出Y也是等概分布，故等号可以成立；特殊情况：P=1, 0, 1/2

DMC容量例子——二元除删信道(BEC)



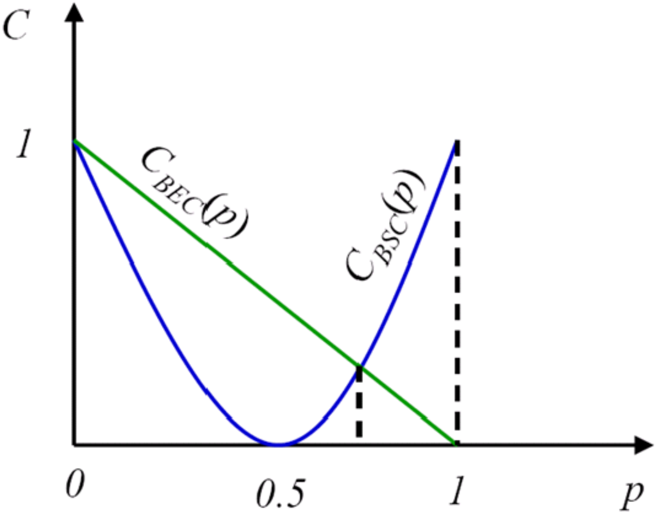
$$
\begin{aligned}
C &= \max_{\{Q_k\}} I(X;Y) \\
&= \max_{\{Q_k\}} \{H(Y) - H(Y \mid X)\} \\
&= \max_{\{Q_k\}} H(Y) - H(p)
\end{aligned}
$$

$$
\begin{aligned}
H(Y) &= H(q(1-p), p, (1-q)(1-p)) \\
&= H(p) + (1-p)H(q)
\end{aligned}
$$

$$
\begin{aligned}
C &= \max_{\{Q_k\}} H(Y) - H(p) \\
&= \max_{q} (1-p)H(q) \\
&= 1 - p
\end{aligned}
$$

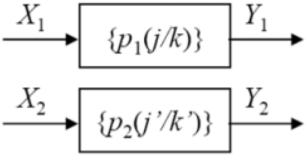
当输入取等概分布时，等号成立。

BSC与BEC的容量比较和启示

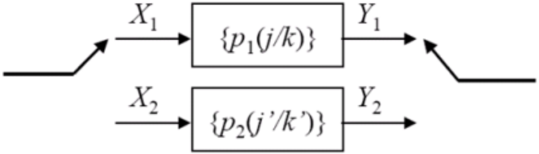


信道的组合

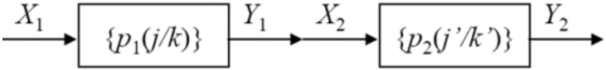
平行信道（积信道）



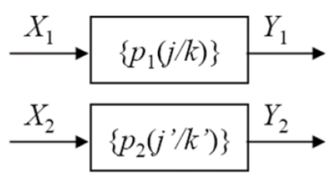
开关信道（和信道）



级联信道



平行信道（积信道）



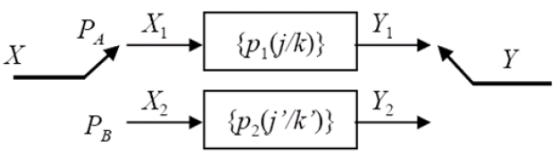
$$C = \sum_{i=1}^{n} C_i$$

$$p(jj' \mid kk') = p_1(j \mid k)p_2(j' \mid k')$$

$$\begin{aligned} I(X_1X_2;Y_1Y_2) &= H(Y_1Y_2) - H(Y_1Y_2 \mid X_1X_2) \\ &\le H(Y_1) + H(Y_2) - H(Y_1 \mid X_1) - H(Y_2 \mid X_2) \\ &= I(X_1;Y_1) + I(X_2;Y_2) \end{aligned}$$

等号成立的条件是各个子信道分别独立取最佳

开关信道（和信道）



$$p(j \mid k) = \begin{cases} p_1(j \mid k), & k \in \mathcal{X}_1,\ j \in \mathcal{Y}_1 \\ p_2(j \mid k), & k \in \mathcal{X}_2,\ j \in \mathcal{Y}_2 \\ 0, & \text{otherwise} \end{cases}, \quad Q_k = \begin{cases} P_A Q_k^1, & k \in \mathcal{X}_1 \\ P_B Q_k^2, & k \in \mathcal{X}_2 \end{cases}$$

$$I(X;Y) = P_A I(X_1;Y_1) + P_B I(X_2;Y_2) + H(P_A, P_B)$$

$$C = \max_{\{P_A, Q_k^1, Q_k^2\}} I(X;Y) = \max_{P_A} \{P_A C_1 + P_B C_2 + H(P_A, P_B)\}$$

$$C = \log\left[2^{C_1} + 2^{C_2}\right], \qquad P_A = 2^{C_1 - C},\ P_B = 2^{C_2 - C}$$

开关信道（和信道）

$$C = \max_{P_A} \{P_A C_1 + P_B C_2 - P_A \log P_A - P_B \log P_B\}$$

$$\frac{\partial C}{\partial P_A} = C_1 - \log P_A - 1 = \lambda$$

$$P_A = 2^{C_1 - \lambda}, P_B = 2^{C_2 - \lambda}$$

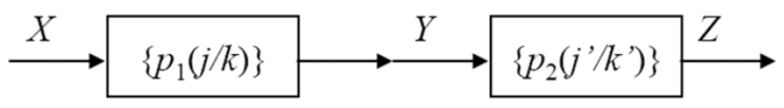
$$C = \left(C_1 2^{C_1-\lambda} + C_2 2^{C_2-\lambda} - (C_1 - \lambda) 2^{C_1-\lambda} - (C_2 - \lambda) 2^{C_2-\lambda}\right)$$

$$= \lambda \left(2^{C_1-\lambda} + 2^{C_2-\lambda}\right) = \lambda$$

$$\therefore 2^{C_1-C} + 2^{C_2-C} = 1 \Rightarrow C = \log\left(2^{C_1} + 2^{C_2}\right)$$

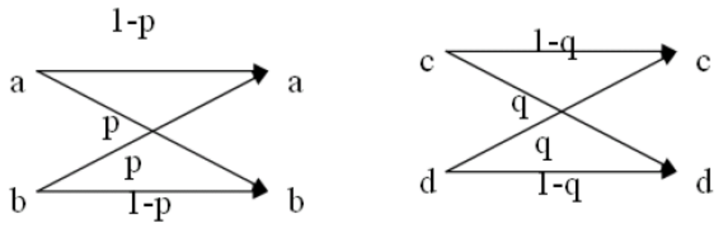
Even $C_1 = C_2 = 0,\ C = 1$, Why?

级联信道



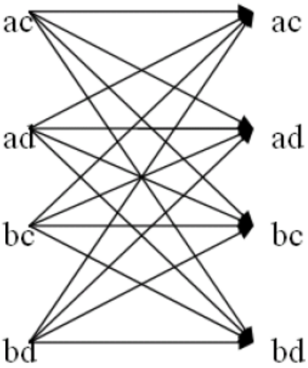
$$I(X;Y) \geq I(X;Z)$$
$$I(Y;Z) \geq I(X;Z)$$
$$C \leq \min\{C_1, C_2\}$$

两个对称信道的组合(1)



$$C_1 = 1 - H(p)$$
$$C_2 = 1 - H(q)$$

两个对称信道的组合(2)

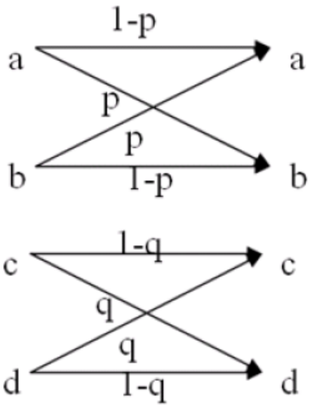


积组合（平行信道）
状态转移矩阵

$$\begin{bmatrix} \bar{p}\bar{q} & \bar{p}q & p\bar{q} & pq \\ \bar{p}q & \bar{p}\bar{q} & pq & p\bar{q} \\ p\bar{q} & pq & \bar{p}\bar{q} & \bar{p}q \\ pq & p\bar{q} & \bar{p}q & \bar{p}\bar{q} \end{bmatrix}$$

$$\begin{aligned} C &= 2 - H(pq, \bar{p}q, p\bar{q}, \bar{p}\bar{q}) \\ &= 2 - H(p) - H(q) \\ &= C_1 + C_2 \end{aligned}$$

两个对称信道的组合(3)



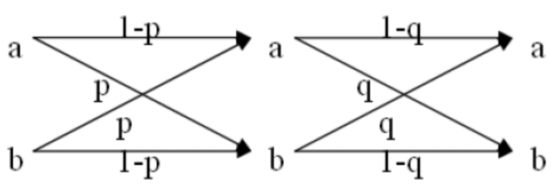
和组合（开关平行信道）
状态转移矩阵

$$\begin{bmatrix} \bar{p} & p & 0 & 0 \\ p & \bar{p} & 0 & 0 \\ 0 & 0 & \bar{q} & q \\ 0 & 0 & q & \bar{q} \end{bmatrix}$$

$$H(Y) \leq H(r) + 1, r = P(X = a) + P(X = b)$$
$$H(Y \mid X) \leq rH(p) + (1 - r)H(q)$$
$$C = \max_r \{H(r) + 1 - rH(p) - (1 - r)H(q)\}$$
$$= \max_r \{H(r) + rC_1 + (1 - r)C_2\}$$

两个对称信道的组合(4)

级联组合



状态转移矩阵，对称信道

$$\begin{bmatrix} \bar{p}\bar{q}+pq & \bar{p}q+p\bar{q} \\ \bar{p}q+p\bar{q} & \bar{p}\bar{q}+pq \end{bmatrix}$$

$$C = 1 - H\left(\bar{p}\bar{q}+pq\right) \leq \begin{cases} 1 - H(p) = C_1 \\ 1 - H(q) = C_2 \end{cases}$$

$$H\left(\lambda p_1 + (1-\lambda)p_2\right) \geq \lambda H(p_1) + (1-\lambda)H(p_2)$$

本讲小结

- 离散无记忆信道容量
- 信道的组合

第二讲：离散无记忆信道的容量定理

离散无记忆信道的容量定理

优化目标：

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

约束条件：

$$\begin{cases} Q_k \geq 0 \\ \sum_{k=0}^{K-1} Q_k = 1 \end{cases}$$

$I(X;Y)$是输入X分布的上凸函数，因此极大值存在。

凸优化（1）

凸优化：凸函数在凸集上的极值。
假定$f(x)$是定义在所有分量为非负的K维无穷凸集R上的一个凸函数，则满足下式的点称为$f(x)$的平稳点，$f(x)$在平稳点处取到极值

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_K} \end{bmatrix} = 0$$

凸优化（2）

在该点处的Hessian矩阵，为对称矩阵

$$H(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial {x_1}^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_K} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial {x_2}^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_K} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_K \partial x_1} & \frac{\partial^2 f(x)}{\partial x_K \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial {x_K}^2} \end{pmatrix}$$

如果是凸函数，则在其定义域内，Hessian矩阵为半正定。Hessian矩阵也是判断一个函数是否为凸的依据。

凸优化（3）

$\nabla f(x)=0$的取值不一定在非负的集合R上！

定理1：$f(x)$是定义在所有分量为非负的K维无穷凸集R 上的一个凸函数，如果$\nabla f(x)$在R上连续，则$f(x)$在R上取极大值的充要条件是

$$f'(x)=0, \forall x_k>0$$
$$f'(x)<0, \forall x_k=0$$

说明：如果在R上能找到，$f'(x)=0$，则极大值必然在该点取到，否则取到边界点$x_k=0$，且此时$f'(x)<0$

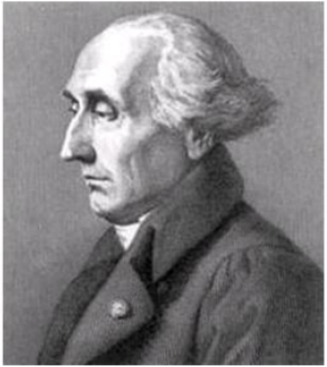
凸优化（4）

定理2：设$f(x)$是定义在K维概率空间R上的一个凸函数，如果$\nabla f(x)$在R 上连续，
则$f(x)$在R上取极大值的充要条件是

$$f'(x) = \lambda, \forall x_k > 0$$
$$f'(x) < \lambda, \forall x_k = 0$$

其中λ是拉格朗日乘数，由$\sum_k x_k = 1$决定。
问题：$\mathbf{max} \quad f(x)$约束条件：$x_k \geq 0, \ \forall k, \ \sum_k x_k = 1$
构造拉格朗日$L(x,\lambda) = f(x) - \lambda \sum_k x_k$，最大化$f(x)$，等价于最大化$L(x,\lambda)$，根据定理1，可得定理2。

凸优化（5）



Joseph-Louis , Lagrange

(1736-1813)

Albert William Tucker (1905-1995)

Harold William Kuhn (1925)

凸优化（6）

凸优化问题：

$$\text{Minimize} \quad f(x)$$

$$\text{subject to:} \qquad g_i(x) \le 0, h_j(x) = 0$$

KKT条件：

$$\begin{aligned}
&\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{l} \lambda_j \nabla h_j(x^*) = 0,\\
&g_i(x^*) \le 0, \ \text{for all } i = 1, \ldots, m\\
&h_j(x^*) = 0, \ \text{for all } j = 1, \ldots, l\\
&\mu_i \ge 0, \ \text{for all } i = 1, \ldots, m\\
&\mu_i g_i(x^*) = 0, \ \text{for all } i = 1, \ldots, m.
\end{aligned}$$

离散无记忆信道的容量定理

定理：概率分布$\{Q_0, Q_1, \cdots, Q_{K-1}\}$达到转移概率为$\{p(j|k)\}$的离散无记忆信道容量C的充要条件为：

$$I(X=k;Y) = C,\ \forall k,\ Q_k > 0$$
$$I(X=k;Y) < C,\ \forall k,\ Q_k = 0$$

其中$I(X=k;Y)$表示通过信道传送字符$X=k$时，信道的输入与输出之间可获得的互信息的期望值，即：

$$I(X=k;Y) = \sum_{j=0}^{J-1} p(j|k) \log \frac{p(j|k)}{p(j)} = \sum_{j=0}^{J-1} p(j|k) \log \frac{p(j|k)}{\sum_{i=0}^{K-1} Q_i p(j|i)}$$

离散无记忆信道的容量定理

证明：

$$I(X=l;Y)=\sum_{j=0}^{J-1}p(j|l)\log\frac{p(j|l)}{\sum_{i=0}^{K-1}Q_ip(j|i)}$$

$$\frac{\partial I(X=l;Y)}{\partial Q_k}=-\sum_{j=0}^{J-1}p(j|l)\frac{p(j|k)}{\sum_{i=0}^{K-1}Q_ip(j|i)}$$

$$\sum_{l=0}^{K-1}Q_l\frac{\partial I(X=l;Y)}{\partial Q_k}=\sum_{l=0}^{K-1}Q_l\sum_{j=0}^{J-1}p(j|l)\frac{p(j|k)}{\sum_{i=0}^{K-1}Q_ip(j|i)}$$

$$=\sum_{j=0}^{J-1}\frac{p(j|k)}{\sum_{i=0}^{K-1}Q_ip(j|i)}\sum_{l=0}^{K-1}Q_lp(j|l)=1$$

离散无记忆信道的容量定理

$$J(\{Q_k\}) = I(X;Y) - \lambda\left(\sum_{k=0}^{K-1} Q_k\right) = \sum_{k=0}^{K-1} Q_k I(X=k;Y) - \lambda\left(\sum_{k=0}^{K-1} Q_k\right)$$

由KKT条件，

$$\frac{\partial J(\{Q_k\})}{\partial Q_k}\begin{cases} = 0,\ \forall Q_k > 0 \\ < 0,\ \forall Q_k = 0 \end{cases}$$

$$\begin{aligned}\frac{\partial J(\{Q_k\})}{\partial Q_k} &= I(X=k;Y) + \sum_{l=0}^{K-1} Q_l \frac{\partial I(X=l;Y)}{\partial Q_k} - \lambda \\ &= I(X=k;Y) - (\lambda - 1)\end{aligned}$$

令$C = \lambda - 1$，得

$$\begin{aligned} I(X=k;Y) &= C,\ \forall k,\ Q_k > 0 \\ I(X=k;Y) &< C,\ \forall k,\ Q_k = 0 \end{aligned}$$

离散无记忆信道的容量定理

说明：
取到离散无记忆信道的容量的充要条件是对于好的 $I(X=k;Y)$，其值均相等，对于不好的$I(X=k;Y)$，其值均为0。

$$C = I(X;Y) = \sum_k Q_k I(X=k;Y)$$

对称离散无记忆信道（1）

离散无记忆信道的转移概率可用$K \times J$矩阵表示

$$\mathbf{P} = \{p(j|k)\} = \begin{bmatrix} p(0|0) & p(1|0) & \cdots & p(J-1|0) \\ p(0|0) & p(1|1) & \cdots & p(J-1|1) \\ \vdots & \vdots & \ddots & \vdots \\ p(0|K-1) & p(1|K-1) & \cdots & p(J-1|K-1) \end{bmatrix}$$

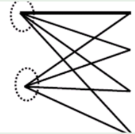
向量$\mathbf{x} = \{x_1, x_2, \cdots, x_N\}$的置换向量定义为：

$$\mathbf{x}' = \{x_{\pi(1)}, x_{\pi(2)}, \cdots, x_{\pi(N)}\}$$

$\boldsymbol{\pi} = \{\pi_1, \pi_2, \cdots, \pi_N\}$是$\{1, 2, \cdots, N\}$ 的任意排列。
例如$\mathbf{x} = \{0.2, 0.3, 0.1, 0.4\}, \mathbf{x}' = \{0.1, 0.2, 0.3, 0.4\}$

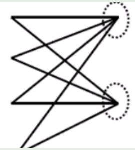
对称离散无记忆信道（2）

若P中每一行都是第一行的一个置换，则该信道关于输入对称



$$H(Y|X) = H(Y|X=k) = -\sum_{j=0}^{J-1} p(j|k)\log p(j|k)$$

若P中每一列都是第一列的一个置换，则该信道关于输出对称



$$\sum_{k=0}^{K-1} p(j|k) = \frac{K}{J} \quad j = 0, 1, \cdots, J-1$$

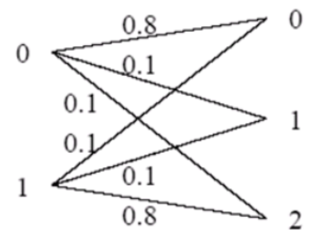
若一个信道既关于输入对称，又关于输出对称，即P中每一行都是第一行的一个置换，每一列都是第一列的一个置换，则该信道是对称的

例如：

$$\mathbf{x} = \begin{pmatrix} 0.2 & 0.3 & 0.1 & 0.4 \\ 0.3 & 0.2 & 0.4 & 0.1 \\ 0.1 & 0.4 & 0.2 & 0.3 \\ 0.4 & 0.1 & 0.3 & 0.2 \end{pmatrix}$$

对称离散无记忆信道（3）

对一个信道的转移概率矩阵P按列划分，得到若干个子信道，若划分出的所有子信道均是对称的，则称该信道是准对称的



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

$$P = \begin{pmatrix} 0.2 & 0.1 & 0.4 & 0.3 \\ 0.3 & 0.1 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0.1 & 0.3 \\ 0.3 & 0.4 & 0.1 & 0.2 \end{pmatrix}$$

准对称离散无记忆信道的容量定理

达到准对称离散无记忆信道容量的输入分布为均匀分布。
证：

$$
\begin{aligned}
\text{若} Q_k &= \frac{1}{K}, k = 0, 1, \cdots, K-1 \\
I(X=k;Y) &= \sum_{j=0}^{J-1} p(j|k) \log \frac{p(j|k)}{\frac{1}{K}\sum_{i=0}^{K-1} p(j|i)} \\
&= \sum_{s}\sum_{j\in y_s} p(j|k) \log \frac{p(j|k)}{\frac{1}{K}\sum_{i}^{K-1} p(j|i)} \\
&= \text{constant}.
\end{aligned}
$$

故满足KKT条件。

$$
P = \begin{pmatrix} 0.2 & 0.1 & 0.4 & 0.3 \\ 0.3 & 0.1 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0.1 & 0.3 \\ 0.3 & 0.4 & 0.1 & 0.2 \end{pmatrix}
$$

具有不同对称性的信道的容量

若信道关于输入对称，则：

$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&= H(Y) + \sum_{j=0}^{J-1} p(j|k)\log p(j|k)
\end{aligned}
$$

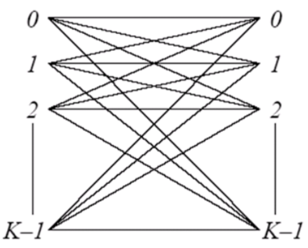
若信道同时也关于输出对称（即信道对称），则：

$$
C = \log J + \sum_{j=0}^{J-1} p(j|k)\log p(j|k)
$$

若信道只关于输入对称，则：

$$
C \le \log J + \sum_{j=0}^{J-1} p(j|k)\log p(j|k)
$$

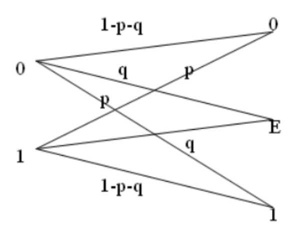
K元对称信道的容量



$$p(j|k) = \begin{cases} 1-p, k=j \\ \dfrac{p}{K-1}, k \neq j \end{cases}$$

$$\begin{aligned} C &= \log K + \sum_{j=0}^{J-1} p(j|k) \log p(j|k) \\ &= \log K + (1-p)\log(1-p) + p \log \frac{p}{K-1} \\ &= \log K - H(p) - p\log(K-1) \end{aligned}$$

除删信道（BEC）

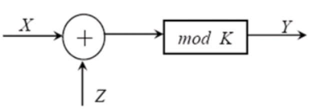


$$\mathbf{P} = \begin{pmatrix} 1-p-q & q & p \\ p & q & 1-p-q \end{pmatrix}$$

$$\begin{aligned} Q_0 &= Q_1 = 0.5 \\ C &= I(X=0;Y) = I(X=1;Y) \\ &= (1-p-q)\log\frac{1-p-q}{(1-q)/2} + q\log\frac{q}{q} + p\log\frac{p}{(1-q)/2} \\ &= (1-p-q)\log(1-p-q) + p\log p - (1-q)\log\frac{1-q}{2} \end{aligned}$$

$p=0$：二元除删信道；$q=0$：二元对称信道；

模K加法信道



该信道是一个对称信道

$Y = X + Z \bmod K$

$X, Y, Z \in \{0, 1, 2, \cdots, K-1\}$

$p(z)$为任意分布

$$
\begin{aligned}
H(Y|X) &= -\sum_x \sum_y Q(x) p(y|x) \log p(y|x) \\
&= -\sum_x \sum_z Q(x) p(z) \log p(z) \\
&= H(z)
\end{aligned}
$$

所以：$C = \log K - H(z)$

当Z为等概分布时，信道容量为0，当Z为确定性分布时，信道容量为$\log K$。

转移概率矩阵可逆信道的容量计算（1）

$$\text{若} Q_k > 0, \text{则} I(X=k;Y) = \sum_{j=0}^{J-1} p(j|k) \log \frac{p(j|k)}{\sum_{i=0}^{K-1} Q_i p(j|i)} = C$$

$$\text{令} \omega_j = \sum_{i=0}^{K-1} Q_i p(j|i), \text{则} \sum_{j=0}^{J-1} p(j|k) \log p(j|k) - \sum_{j=0}^{J-1} p(j|k) \log \omega_j = C$$

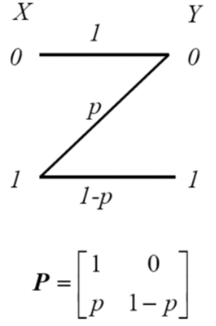
$$\text{即} \sum_{j=0}^{K-1} p(j|k) \underbrace{[C + \log \omega_j]}_{\beta_j} = \underbrace{\sum_{j=0}^{K-1} p(j|k) \log p(j|k),}_{\alpha_k} \quad k = 0, 1, \cdots, K-1$$

亦即$\mathbf{P}\beta$=α，若$\mathbf{P}$可逆则可解出唯一的β。
进一步可得$\omega_j = 2^{\beta_j - C}$，再由$\sum_j \omega_j = 1$，得$C = \log \sum_j 2^{\beta_j}$

转移概率矩阵可逆信道的容量计算（2）

进一步，再由$\omega_j = \sum_{i=0}^{K-1} Q_i p(j|k), j = 0, 1, \cdots, K-1$
求出$\{Q_k\}$。
讨论：
若存在某个$Q_k < 0$，则原假设前提不满足，必须尝试令其中某个符号不发送（不一定是符号k），再用上述过程重新计算。

例子



$$\boldsymbol{P\beta} = \boldsymbol{\alpha}$$

$$\begin{bmatrix} 1 & 0 \\ p & 1-p \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} 0 \\ -H(p) \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 0 \\ -H(p)/(1-p) \end{bmatrix}$$

$$C = \log \sum_{j=0} 2^{\beta_j} = \log \left[1 + 2^{-H(p)/(1-p)} \right]$$

$$P = \begin{bmatrix} 1 & 0 \\ p & 1-p \end{bmatrix}$$

例子

$$\omega_0 = 2^{\beta_0 - C} = \frac{1}{1 + 2^{-\frac{H(p)}{1-p}}} = q + (1-q)p$$

$$\omega_1 = 2^{\beta_1 - C} = \frac{2^{-\frac{H(p)}{1-p}}}{1 + 2^{-\frac{H(p)}{1-p}}} = (1-p)(1-q)$$

$$q = \frac{1 - \left(\frac{p}{1-p}\right) 2^{-\frac{H(p)}{1-p}}}{1 + 2^{-\frac{H(p)}{1-p}}}$$

本讲小结

- 离散无记忆信道容量定理
- 准对称信道容量定理
- 转移概率矩阵可逆信道容量计算

作业：

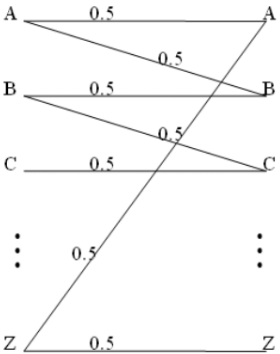
4.1

4.2

4.9

第三讲：离散无记忆信道的编码定理

复制信道

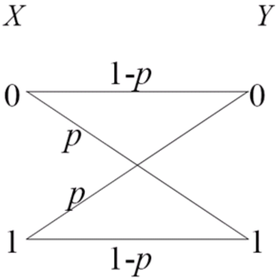


$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&= H(Y) - 1
\end{aligned}
$$

$$
\begin{aligned}
C &= \max_{\{Q_k\}} I(X;Y) \\
&= \max_{\{Q_k\}} H(Y) - 1 \\
&\le \log 26 - 1 = \log 13
\end{aligned}
$$

编码方法：?

二元对称信道



$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y \mid X) \\
&= H(Y) - \sum_x p(x) H(Y \mid X = x) \\
&= H(Y) - \sum_x p(x) H(p) \\
&= H(Y) - H(p) \\
&\leq 1 - H(p)
\end{aligned}
$$

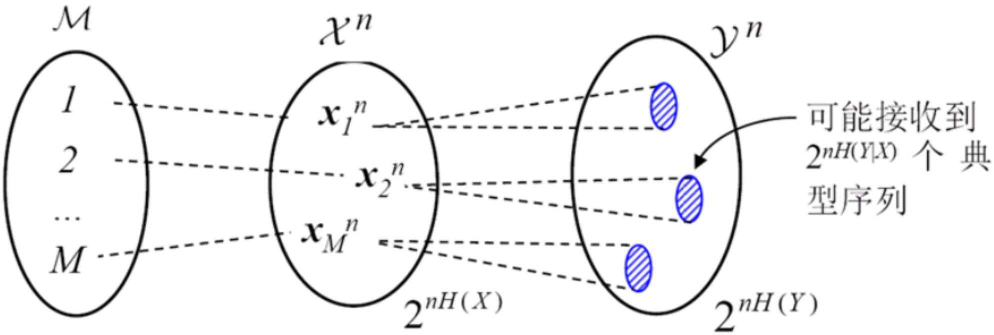
离散无记忆信道的编码

$W = \{1, 2, \cdots, M\}$

编码速率（传信率）：$R = \dfrac{\log M}{n}$。R与C的关系：

香农信道编码定理：如果信息传输速率R小于信道容量C，则总存在一种编码方法，使信息在该信道上无错误地可靠传输。

离散无记忆信道的编码直观说明



$$M \approx \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{nI(X;Y)};\ R = \frac{\log M}{n} \approx I(X;Y) \le C$$

离散无记忆信道编码的基本定义(1)

定义4.4.1 离散无记忆信道$(\mathcal{X}, p(y|x), \mathcal{Y})$上的一个$(M, n)$ 码由如下组成：

①一个与M个消息相对应的标号集合$\mathcal{W} = \{1, 2, \cdots, M\}$。

②一个编码器，把消息$w \in \mathcal{W}$映射成码字$x^n \in \mathcal{X}^n$，所得到的码字为$X^n(1), X^n(2), \cdots, X^n(M)$；一个码的全体码字构成码书。

③一个译码器$g_1 : \mathcal{Y}^n \to \mathcal{W}$：对应于确定的译码法则，帮助接收者根据接收到的序列Y^n来确定发送消息是什么。

离散无记忆信道编码的基本定义(2)

定义4.4.2 (错误概率) 发送第i个消息所发生的错误概率定义为：

$$\lambda_i = P\left\{g(Y^n) \neq i | X^n = x^n(i)\right\}$$

最大错误概率定义为：

$$\lambda^{(n)} = \max_{i \in \{1,2,\cdots,M\}} \lambda_i$$

平均错误概率：

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

编码速率：

$R = \frac{\log M}{n}$比特/传输

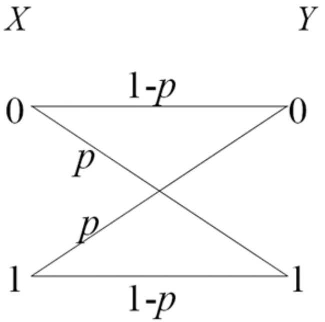
如果存在一系列$(2^{nR}, n)$码，当$n \to \infty$时，最大错误概率 $\lambda^{(n)} \to 0$，R被称为可达的。

香农编码定理的证明

证明思想：
1.所谓可靠通信是指错误概率可以任意小，但并非为零。
2.信道上不是仅传输一个符号，而是传输一串很长的符号序列。由于多次使用信道，可以利用概率论中的大数定理。
3.计算在一类随机选择的码书上的平均错误概率，平均错误概率与最大错误概率具有一样的意义。

如果$(2^{nR}, n)$码的平均错误概率$\lambda^{(n)} < \epsilon$，则至少有一半的码字其最大错误概率小于2ϵ，意味着$(2^{nR-1}, n)$码的最大错误概率 $\lambda^{(n)} < 2\epsilon \to 0$，即码率$\frac{nR-1}{n} = R - \frac{1}{n}$是可达的。

二元对称信道的编码定理(1)



$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y \mid X) \\
&= H(Y) - \sum_x p(x) H(Y \mid X = x) \\
&= H(Y) - \sum_x p(x) H(p) \\
&= H(Y) - H(p) \\
&\leq 1 - H(p)
\end{aligned}
$$

$p < 1/2$

二元对称信道的编码定理(2)

码书矩阵：

$$\begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_n(1) \\ x_1(2) & x_2(2) & \cdots & x_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(2^{nR}) & x_2(2^{nR}) & \cdots & x_n(2^{nR}) \end{bmatrix}$$

译码方法：如果接收到的y^n与某个码字$x^n(i)$的Hamming距离小于$r = n(p+\epsilon_2)$时，就宣称发送的是第i个码字，其中ϵ_2是任意小数，且满足$p+\epsilon_2 < 0.5$。

译码错误：y^n与$x^n(i)$的Hamming距离大于$r = n(p+\epsilon_2)$，或者y^n与其它$x^n(j)$的Hamming距离小于$r = n(p+\epsilon_2)$。

二元对称信道的编码定理(3)

$$P(E) = \sum_{\mathcal{L}} P(\mathcal{L}) P_e^n(\mathcal{L})$$

$$P_e^n(\mathcal{L}) = \frac{1}{2^{nR}} \sum_{i=1}^{2^{nR}} \lambda_i(\mathcal{L})$$

$$\begin{aligned} P(E) &= \frac{1}{2^{nR}} \sum_{i=1}^{2^{nR}} \sum_{\mathcal{L}} P(\mathcal{L}) \lambda_i(\mathcal{L}) \\ &= \sum_{\mathcal{L}} P(\mathcal{L}) \lambda_1(\mathcal{L}) \triangleq P(E|i=1) \end{aligned}$$

$$P(E|i=1) = P\left\{\overline{E}_1 \cup E_2 \cup \cdots \cup E_{2^{nR}}\right\} = P(\overline{E}_1) + \sum_{i=2}^{2^{nR}} E_i$$

二元对称信道的编码定理(4)

$$P(E|i=1) = P\left\{\overline{E}_1 \cup E_2 \cup \cdots \cup E_{2^{nR}}\right\} = P(\overline{E}_1) + \sum_{i=2}^{2^{nR}} E_i$$

1) 不妨令$i=1$为全0序列，则$P(\overline{E}_1)$表示输出序列中1的个数超过np的概率，根据切比雪夫不等式，该概率可以任意小，即$P(\overline{E}_1) < \frac{\epsilon}{2}$。

2) 发送其它序列时，输出序列有2^n个，但只有$\sum_{t=0}^{r} C_n^t$个落在半径为r的Hamming球内

$$P(E_1) < \frac{\epsilon}{2} + (2^{nR}-1)\frac{1}{2^n}\sum_{t=0}^{r} C_n^t$$

二元对称信道的编码定理(5)

$$
\begin{aligned}
P(E_1) &< \frac{\epsilon}{2} + (2^{nR} - 1)\frac{1}{2^n}\sum_{t=0}^{r} C_n^t \\
&= \frac{\epsilon}{2} + (2^{nR} - 1)\frac{1}{2^n}\sum_{t=0}^{n\lambda} C_n^t, \lambda < 1/2 \\
&< \frac{\epsilon}{2} + 2^{nR}2^{-n[1-H(\lambda)]} = \frac{\epsilon}{2} + 2^{n(R-C)}
\end{aligned}
$$

$$
r = n\lambda = n(p + \epsilon_2)
$$

$$
1 \geq \sum_{t=0}^{n\lambda} C_n^t \lambda^t (1-\lambda)^{n-t} \geq \sum_{t=0}^{n\lambda} C_n^t \lambda^{n\lambda}(1-\lambda)^{n-n\lambda} = \sum_{t=0}^{n\lambda} C_n^t 2^{-nH(\lambda)}
$$

联合典型列

相对于联合分布$p(x,y)$的联合典型列集合$A_\epsilon^{(n)}$
是指具有下列性质的序列对(x^n,y^n)的集合

$$A_\epsilon^{(n)} = \{(x^n,y^n) \in \mathcal{X}^n \times \mathcal{Y}^n :$$

$$\left|-\frac{1}{n}\log p(x^n) - H(X)\right| < \epsilon$$

$$\left|-\frac{1}{n}\log p(y^n) - H(Y)\right| < \epsilon$$

$$\left.\left|-\frac{1}{n}\log p(x^n,y^n) - H(XY)\right| < \epsilon\right\}$$

联合典型列的性质(1)

设$A_\epsilon^{(n)}$是与$p(x,y)$对应的联合典型列集合，若联合序列(X^n,Y^n)的每一个符号对(x_i,y_i)均是按照联合分布$p(x,y)$独立选取，即$(X^n,Y^n)\sim p(x^n,y^n)$，则

(1) $Pr\left((X^n,Y^n)\in A_\epsilon^{(n)}\right)\to 1$

(2) $(1-\epsilon)2^{n(H(XY)-\epsilon)}\leq \left|A_\epsilon^{(n)}\right|\leq 2^{n(H(XY)+\epsilon)}$

(3) 如果联合序列$(\widetilde{Y}^n,\widetilde{Y}^n)\sim p(x^n)p(y^n)$，即$\widetilde{X}^n$和$\widetilde{Y}^n$分别按照$p(x,y)$的边缘分布$p(x)$和$p(y)$来独立选取，则 $(1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}\leq Pr\left((\widetilde{X}^n,\widetilde{Y}^n)\in A_\epsilon^{(n)}\right)\leq 2^{-n(I(X;Y)-3\epsilon)}$

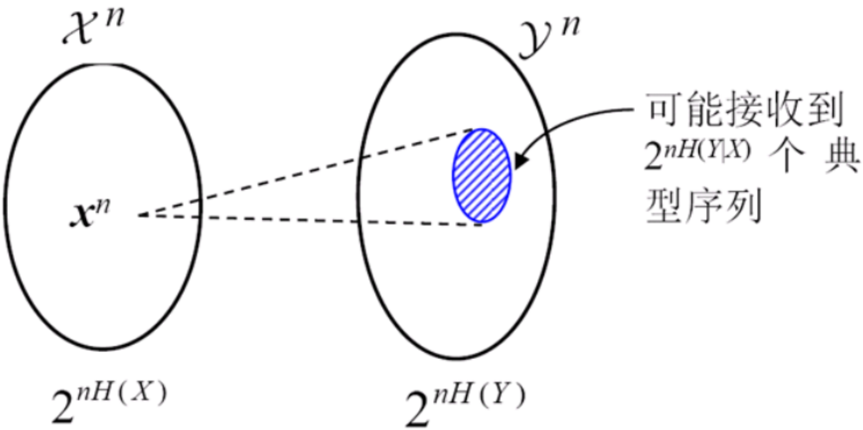
联合典型列的性质(2)

$(\widetilde{Y}^n, \widetilde{Y}^n)$与$(X^n, Y^n)$有相同的边缘分布$p(x^n)$ 和$p(y^n)$，故

$$
\begin{aligned}
Pr\left((\widetilde{X}^n, \widetilde{Y}^n) \in A_\epsilon^{(n)}\right) &= \sum_{A_\epsilon^{(n)}} p(x^n)p(y^n) \\
&\le \left|A_\epsilon^{(n)}\right| 2^{-n(H(X)-\epsilon)} 2^{-n(H(Y)-\epsilon)} \\
&\le 2^{n(H(XY)+\epsilon)} 2^{-n(H(X)-\epsilon)} 2^{-n(H(Y)-\epsilon)} \\
&= 2^{-n(I(X;Y)-3\epsilon)}
\end{aligned}
$$

$$
\begin{aligned}
Pr\left((\widetilde{X}^n, \widetilde{Y}^n) \in A_\epsilon^{(n)}\right) &= \sum_{A_\epsilon^{(n)}} p(x^n)p(y^n) \\
&\ge \left|A_\epsilon^{(n)}\right| 2^{-n(H(X)+\epsilon)} 2^{-n(H(Y)+\epsilon)} \\
&\ge (1-\epsilon) 2^{n(H(XY)-\epsilon)} 2^{-n(H(X)-\epsilon)} 2^{-n(H(Y)-\epsilon)}
\end{aligned}
$$

联合典型列的性质(3)



随机选择y^n，约有$2^{-nI(X;Y)}$的概率与x^n构成联合典型。

离散无记忆信道编码定理

香农信道编码定理：如果信息传输率R小于信道容量C，则总存在一种编码方法，使信息在该信道上无错误地可靠传输。

所有低于信道容量C的速率R均是可达的，即当$R < C$时，总存在一系列码$(2^{nR}, n)$，当$n \to \infty$时，最大误码概率$\lambda^{(n)} \to 0$。
逆定理：具有$\lambda^{(n)} \to 0$的任何$(2^{nR}, n)$码必有$R \leq C$。

码率可达区$\Longleftrightarrow$容量区内界
容量区内界如果与容量区外界相等，则为容量区。

离散无记忆信道编码定理的证明(1)

码书矩阵：

$$
\begin{bmatrix}
x_1(1) & x_2(1) & \cdots & x_n(1) \\
x_1(2) & x_2(2) & \cdots & x_n(2) \\
\vdots & \vdots & \ddots & \vdots \\
x_1(2^{nR}) & x_2(2^{nR}) & \cdots & x_n(2^{nR})
\end{bmatrix}
$$

码书随机产生，均匀分布。

译码方法：如果接收到的y^n与某个码字$x^n(\hat{w})$是联合典型的，即$(x^n(\hat{w}), y^n) \in A_\epsilon^{(n)}$，就宣称发送的是第$\hat{w}$个码字。

译码错误：(1)译码错误事件：$\hat{w} \neq w$; (2) 不可译事件：不存在任何$x^n(\hat{w})$，使得$(x^n(\hat{w}), y^n) \in A_\epsilon^{(n)}$。

离散无记忆信道编码定理的证明(2)

$$P(E) = \sum_{\mathcal{L}} P(\mathcal{L}) P_e^n(\mathcal{L})$$

$$P_e^n(\mathcal{L}) = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \lambda_w(\mathcal{L})$$

$$P(E) = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{\mathcal{L}} P(\mathcal{L}) \lambda_w(\mathcal{L})$$

$$= \sum_{\mathcal{L}} P(\mathcal{L}) \lambda_1(\mathcal{L}) \triangleq P(E|w=1)$$

$$P(E|w=1) = P\left\{\overline{E}_1 \cup E_2 \cup \cdots \cup E_{2^{nR}}\right\} = P(\overline{E}_1) + \sum_{i=2}^{2^{nR}} E_i$$

离散无记忆信道编码定理的证明(3)

$$P(E|w=1) = P\left\{\overline{E}_1 \cup E_2 \cup \cdots \cup E_{2^{nR}}\right\} = P(\overline{E}_1) + \sum_{i=2}^{2^{nR}} E_i$$

1) 根据联合典型列性质：$P(\overline{E}_1) \leq \epsilon$
2) 由于$x^n(i)$与$x^n(1)$独立无关，意味着y^n与$x^n(i)$独立，因此，

$$P(E_i) \leq 2^{-n(I(X;Y)-3\epsilon)}$$

$P(E|w=1) \leq \epsilon + 2^{-n(I(X;Y)-R-3\epsilon)}$
当$R < I(X;Y) - 3\epsilon$时，对于充分大的n，有$P(E|w=1) \leq 2\epsilon$

逆定理证明

逆定理：具有$\lambda^{(n)} \to 0$的任何$(2^{nR}, n)$码必有$R \le C$

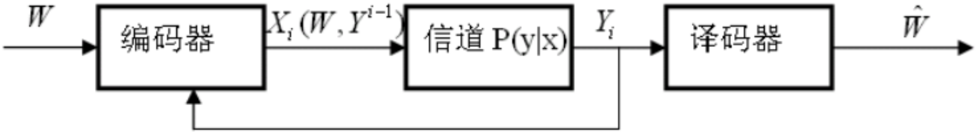
$$\begin{aligned}
nR = H(W) &= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(X^n;Y^n) \quad \to \text{数据处理定理} \\
&\le H(P_e^{(n)}) + P_e^{(n)} \cdot nR + I(X^n;Y^n) \quad \to Fano\text{不等式} \\
&\le 1 + P_e^{(n)} \cdot nR + nC \\
R &\le P_e^{(n)} R + \frac{1}{n} + C
\end{aligned}$$

$P_e^{(n)} \ge 1 - \frac{C}{R} - \frac{1}{nR}$，当$R > C$，则$P_e^{(n)}$一定大于0!

香农信道编码定理的证明思路总结

- 随机编码
- 联合典型列
- 渐进无差错
- 逆定理证明：*Fano*不等式

具有理想反馈的DMC的容量(1)



问题：反馈能否增加离散无记忆信道的容量？
答案：否

定理：$C_{FB} = C = \max_{p(x)} I(X;Y)$

具有理想反馈的DMC的容量(2)

证明：显然$C_{FB} \geq C$

$$
\begin{aligned}
nR = H(W) &= H(W|Y^n) + I(W;Y^n) \\
&\leq H(W|\hat{W}) + I(W;Y^n) \\
&\leq 1 + P_e^{(n)} \cdot nR + I(W;Y^n)
\end{aligned}
$$

$$
\begin{aligned}
I(W;Y^n) &= H(Y^n) - H(Y^n|W) \\
&= H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y^{i-1}, W) \\
&= H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y^{i-1}, W, X_i) \\
&= H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i) \leq nC \quad \Longrightarrow C_{FB} \leq C
\end{aligned}
$$

具有理想反馈的DMC的容量(3)

- 反馈不能提高离散无记忆信道的容量，但对于离散有记忆信道来说，反馈可以增加信道容量。
- 虽然反馈不能提高DMC的容量，但是利用反馈可以较简单地实现性能好的编码，比如可以用码长较短的信道编码就可以达到很低的误码率。
- 反馈如果不是理想的，情况会如何？
- 有限反馈？

信源信道分离编码和联合编码

信源、信道分离编码

$H(W) < R_s < R_c < C \ \Rightarrow P_e^{(n)} \to 0$

信源、信道联合编码

在有限集上取值的信源$\mathcal{V}$的熵速率为$H(\mathcal{V})$，若$H(\mathcal{V}) < C$，则存在一个信源信道联合编码，使得$P_e^{(n)} \to 0$；反之，若$H(\mathcal{V}) > C$，则不可能以任意小的错误概率传送信源。

信源信道联合编码定理的证明(1)

正定理:

$H(\mathcal{V}) < C \Rightarrow \exists\epsilon, H(\mathcal{V}) + \epsilon < C$

该源的ϵ典型列集合$\left|A_\epsilon^{(n)}\right| < 2^{n(H(\mathcal{V})+\epsilon)} < 2^{nC}$

所以可用不大于2^{nC}个不同的码字表达典型列集合。

此时编码速率小于信道容量，故该码率可达。

对非典型列不予编码，由此引起的差错不多于ϵ。

信源信道联合编码定理的证明(2)

逆定理：

反之，设信源信道联合编码为$V^n \to X^n \to Y^n \to \hat{V}^n$

则$H(V^n|\hat{V}^n) \leq 1 + p_e^{(n)} \log |\mathcal{V}^n| = 1 + np_e^{(n)} \log |\mathcal{V}|$

$$
\begin{aligned}
H(\mathcal{V}) &\leq \frac{H(V_1 V_2 \cdots V_n)}{n} = \frac{H(V^n)}{n} = \frac{1}{n}H(V^n|\hat{V}^n) + \frac{1}{n}I(V^n;\hat{V}^n) \\
&\leq \frac{1}{n}(1 + np_e^{(n)} \log |\mathcal{V}|) + \frac{1}{n}I(X^n;Y^n) \\
&\leq \frac{1}{n} + p_e^{(n)} \log |\mathcal{V}| + C
\end{aligned}
$$

若要求$P_e^{(n)} \to 0$，则$H(\mathcal{V}) < C$

信源信道分离编码和联合编码

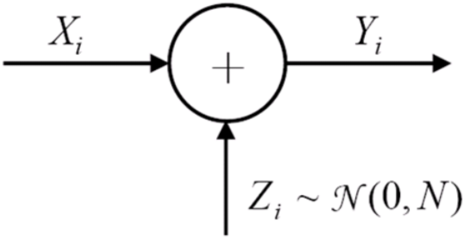
- 只要$H < C$，总可以找到可行的信源信道联合编码；也可以分别构造最优的信源编码和信道编码，使信息传输可达
- 信源信道分离编码减轻了工程师的负担，使得编码变得简单
- 信源信道联合编码不能使得可行速率极限增加，但可以简化编码
- 当前，信源信道联合编码是比较热门的课题

本讲小结

- 离散无记忆信道的编码定理
- 联合典型列
- 信道编码定理的证明
- 信源信道分离编码和联合编码

第四讲：加性高斯噪声（AWGN）信道

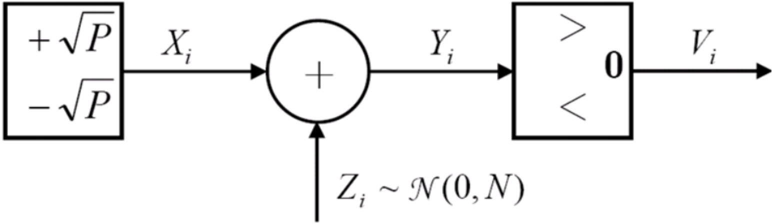
时间离散的加性高斯信道



$$Y_i = X_i + Z_i$$

$$\frac{1}{n}\sum_{i=1}^{n} {x_i}^2 \le P$$

时间离散的加性高斯信道的幅度离散化



$$Y_i = X_i + Z_i$$
$$X_i \in \{+\sqrt{P}, -\sqrt{P}\}$$

时间离散的加性高斯信道的幅度离散化

二元对称信道

$$
\begin{aligned}
P_e &= P\{Y < 0 | X = \sqrt{P}\} \\
&= P\{Y > 0 | X = -\sqrt{P}\} \\
&= P\{Z < -\sqrt{P} | X = \sqrt{P}\} \\
&= P\{Z > \sqrt{P}\} = 1 - \Phi\left(\sqrt{\frac{P}{N}}\right)
\end{aligned}
$$

二元除删信道

加性高斯信道的容量

$$C = \max_{p(x):EX^2 \leq P} I(X;Y)$$

$$\begin{aligned} I(X;Y) &= h(Y) - h(Y|X) \\ &= h(Y) - h(X+Z|X) \\ &= h(Y) - h(Z|X) \\ &= h(Y) - h(Z) \end{aligned}$$

$$\begin{aligned} h(Z) &= \frac{1}{2}\log 2\pi e N \\ E(Y^2) &= P + N \\ h(Y) &\leq \frac{1}{2}\log 2\pi e(P+N) \end{aligned}$$

$$I(X;Y) \leq \frac{1}{2}\log 2\pi e(P+N) - \frac{1}{2}\log 2\pi e N = \frac{1}{2}\log(1+\frac{P}{N})$$

当输入X为高斯分布时等号成立

加性高斯信道的容量的意义

1.发送信号采用高斯分布时，互信息最大。

实际通信系统？

2.干扰信道（噪声信号）为高斯分布时，互信息最小。

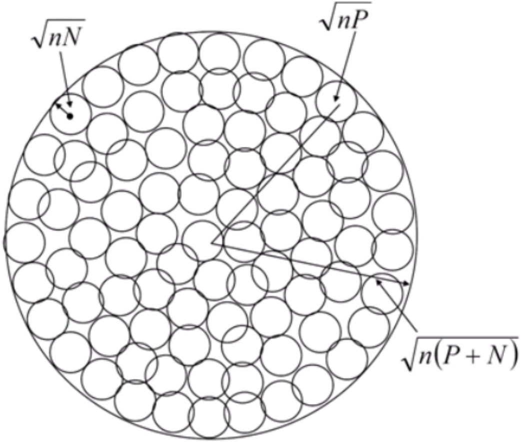
高斯干扰最有效。

$$I(X;Y) = h(Y) - h(Z)$$

高斯信道编码定理

高斯信道编码定理：在噪声方差为N，信号功率限制为P的加性高斯信道上，任何速率小于C 的码率R，是可达的，即总存在一种编码方法，使信息在该信道上无错误地可靠传播。
逆定理：任何$R > C$是不可达的。

$$\frac{A_n(n(P+N))^{\frac{n}{2}}}{A_n(nN)^{\frac{n}{2}}} = 2^{\frac{n}{2}\log(1+\frac{P}{N})}$$



逆定理证明

利用Fano不等式

$$H(W|Y^n) \le 1 + nRP_e^{(n)} \triangleq n\varepsilon_n$$

因为$P_e^{(n)} \to 0$，所以当$n \to \infty$时$\varepsilon_n \to 0$。因而

$$\begin{aligned} nR = H(W) &= I(W;Y^n) + H(W|Y^n) \\ &\le I(X^n;Y^n) + n\varepsilon_n \\ &= h(Y^n) - h(Y^n|X^n) + n\varepsilon_n \\ &= \sum_{i=1}^{n}\{h(Y_i) - h(Z_i)\} + n\varepsilon_n \end{aligned}$$

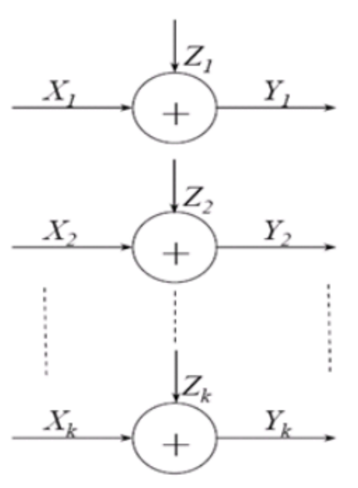
$$h(Y_i) \le \frac{1}{2}\log 2\pi e(P_i + N)$$

$$\begin{aligned} nR &\le \sum_{i=1}^{n}\{h(Y_i) - h(Z_i)\} + n\varepsilon_n \\ &\le \sum_{i=1}^{n}\frac{1}{2}\log(1 + \frac{P_i}{N}) + n\varepsilon_n \end{aligned}$$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}\log(1 + \frac{P_i}{N}) \le \frac{1}{2}\log(1 + \frac{1}{n}\sum_{i=1}^{n}\frac{P_i}{N}) \le \frac{1}{2}\log(1 + \frac{P}{N}) = C$$

高斯平行信道（1）

物理意义：OFDM；CDMA；MIMO



$$Y_i = X_i + Z_i, i = 1, 2, \cdots, k$$

$$Z_i = N(0, N_i), i = 1, 2, \cdots, k$$

$$E\left\{\sum_{i=1}^{k} X_i^2\right\} \leq P$$

$$\begin{aligned} & I(X_1, X_2, \cdots, X_k; Y_1, Y_2, \cdots, Y_k) \\ & = h(Y_1, Y_2, \cdots, Y_k) - \sum_{i=1}^{k} h(Z_i) \\ & \leq \sum_{i=1}^{k} h(Y_i) - \sum_{i=1}^{k} h(Z_i) \\ & \leq \frac{1}{2} \sum_{i=1}^{k} \log(1 + \frac{P_i}{N_i}) \end{aligned}$$

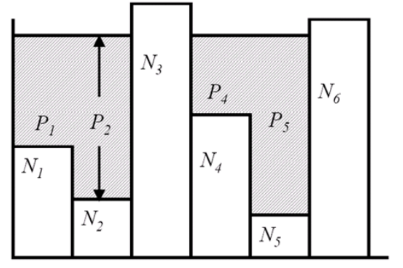
其中，$P_i = EX_i^2, \sum_{i=1}^{k} P_i \leq P$

高斯平行信道（2）

$$\max \sum_{i=1}^{k} \log\left(1+\frac{P_i}{N_i}\right)$$
$$\text{s.t} \sum_{i=1}^{k} P_i \le P$$

拉格朗日乘数法：

$$J = \sum_{i=1}^{k} \log\left(1+\frac{P_i}{N_i}\right) - \lambda \sum_{i=1}^{k} P_i$$
$$\frac{\partial J}{\partial P} = \frac{1}{P_i+N_i} - \lambda = 0$$
$$P_i = \gamma - N_i$$
$$P_i = (\gamma - N_i)^+$$



根据$\sum_{i=1}^{k} P_i = P$可以算出γ

注水（灌水，water-filling，water-pouring）法则：

带限（模拟）高斯信道的容量

$$y(t) = x(t) + z(t)$$

$x(t), z(t)$的带宽均限制在$[0, W]$Hz之内。

连续信号的离散化表示（Nyquist抽样）

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right) \mathrm{sinc} 2\pi W \left(t - \frac{n}{2W}\right)$$

其中，$\mathrm{sinc}(x) = \dfrac{\sin x}{x}$。

连续信号的离散化

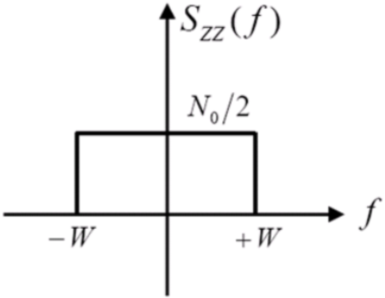
$$x(t) = \sum_{n=-\infty}^{\infty} x_n \varphi_n(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2W}\right) \operatorname{sinc} 2\pi W \left(t - \frac{n}{2W}\right)$$

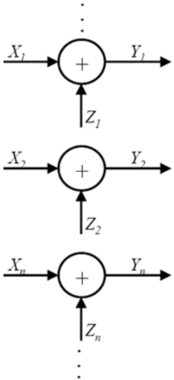
$$z(t) = \sum_{n=-\infty}^{\infty} z_n \varphi_n(t) = \sum_{n=-\infty}^{\infty} z\left(\frac{n}{2W}\right) \operatorname{sinc} 2\pi W \left(t - \frac{n}{2W}\right)$$

$$y(t) = \sum_{n=-\infty}^{\infty} y_n \varphi_n(t) = \sum_{n=-\infty}^{\infty} (x_n + z_n) \varphi_n(t)$$

$$y_n = x_n + z_n$$

等效平行信道





若噪声$z(t)$是均值为0、单边带宽为W、双边功率谱密度为$N_0/2$ 的白高斯过程，其Nyquist采样序列$\{z_n = z(\frac{n}{2W})\}$为均值为0，方差为$N_0/2$ 的独立随机序列。

模拟高斯信道的容量

考虑T秒钟的模拟传输过程，它相当于$2WT$次平行的传输。设每次传输时输入样本x_i的方差分别为P_i，则由功率限制：

$$\sum_{i=1}^{2WT} P_i \le PT$$

T秒钟的传输容量：$C_T = \frac{1}{2}\sum_{i=1}^{2WT} \log\left(1+\frac{P_i}{N_i}\right)$，

$$N_i \equiv \frac{N_0}{2}, P_i \equiv \frac{P}{2W}; C_T = WT\log\left(1+\frac{P}{N_0W}\right)$$

等效地，每秒钟的容量$C = W\log\left(1+\frac{P}{N_0W}\right)$

传输1比特需要的最小能量

传输每比特的平均能量为$\varepsilon_b = P/R$

由容量公式可得$\eta = \dfrac{R}{W} = \log\left(1 + \dfrac{\varepsilon_b R}{N_0 W}\right)$

所以在频谱效率为η时，$\varepsilon_b^*(\eta) = \dfrac{N_0}{\eta}\left(2^\eta - 1\right)$

由于是η的严格单调增函数，所以传输1比特的最小能量为

$$\varepsilon_b^*(0) = \lim_{R\to 0} \varepsilon_b^*(\eta) = N_0 \ln 2 = 0.693 N_0$$

$$f(x)=\frac{e^x-1}{x}\text{是}x\text{的增函数}$$

证明：$f'(x)=\frac{e^x}{x}-\frac{e^x-1}{x^2}=\frac{e^x(x-1)+1}{x^2}$

$g(x)=e^x(x-1)+1$是x的增函数，且$g(0)=0$

$f'(x)>0, x>0$。

推论：$\varepsilon_b^*(R)=\frac{N_0}{R}\left(2^R-1\right)$ 是R 的增函数，令$R=x\log e$。

推论：$P(W)=\left(2^{\frac{C}{W}}-1\right)N_0W$是$W$的减函数，令$W=\frac{C}{x\log e}$。

功率与频率资源的互换

$$C = W\log\left(1+\frac{P}{N_0 W}\right) \Longrightarrow P(W) = \left(2^{\frac{C}{W}}-1\right)N_0 W$$

功率与频率资源的互换：所使用的带宽越宽，需要的功率越小。

- 每比特需要的最少功率：$P_{min} = \lim_{W\to\infty}\left(2^{\frac{C}{W}}-1\right)N_0 W = N_0 C\ln 2$
- 功率效率的极限：$W\to\infty, C\to\dfrac{P}{N_0}\log e(bps)$

对通信技术的启示

- 使用的带宽越大，系统功率效率越高，或者说传输同样信息，需要的功率越小。
  CDMA,OFDM...（不利因素）
- P随C呈指数上升趋势
  16QAM的功率效率不如QPSK!
- 在AWGN信道上信息传输得越慢，则越节省能量。
  传感器网络，节点能量有限，数据量比较小，尽量使用BPSK!（矛盾!）

香农信道容量定理讨论（1）

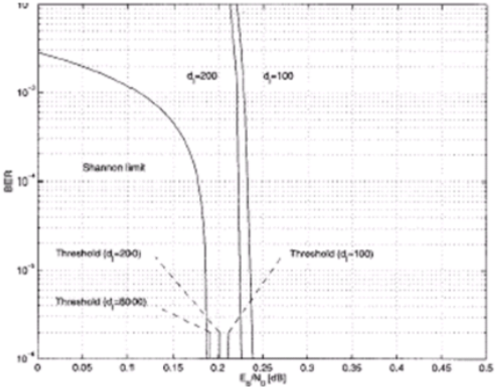
例子：假设一个高斯信道，其容量为1.5b/s/Hz，如果使用BPSK 传输，根据Shannon 信道容量定理，错误概率为0，实际系统错误概率不可能为0，为什么？

原因：码长无限长，高斯输入。

香农信道容量定理讨论（2）

On the Design of Low-Density Parity-Check Codes within 0.0045 dB of the Shannon Limit

Sae-Young Chung, *Member, IEEE*, G. David Forney, Jr., *Fellow, IEEE*, Thomas J. Richardson, and Rüdiger Urbank



本讲小结

- 高斯加性信道的容量
- 平行高斯信道的注水法则
- 模拟高斯信道的容量

作业

- 4.13
- 4.14