信息论与编码

第5章 有噪信道编码定理

张建国

主要内容与基本要求

- 主要内容
  - 信道速率的定义；
  - 解码准则（最小误差与最大似然），误码率的计算；
  - 二元编码误差；
  - 多元编码误差，信道编码定理。
- 基本要求
  - 理解信道编码的目的；理解信道速率的概念；
  - 理解最小误差和最大似然两个解码准则，会根据最大似然解码准则划分输出子集；
  - 了解二元编码的误码率上界，会计算 $g_n(s)$；
  - 了解多元编码的误码率上界，理解编码指数的含义及使用方法，掌握编码指数的曲线变化，掌握有噪信道编码定理。

本章目录

5.0 引言

信道编码的基本概念

信息是一种抽象的内涵，它以信号的形式在信道上传输，即信号是信息的有形载体。

（广义的）信道编码是以信息在信道上的正确传输为目标的编码，它可以分为两个层次。

- **如何正确接收载有信息的信号**

目的是消除直流分量、改造信号频谱、压缩占用带宽、抑制码间串扰等，从而使信号适应信道的特性。这属于通信原理的研究范畴，如数字信号的基带传输（AMI、HDBn）与载波传输（ASK、PSK、FSK）。这些编码实际上是线路编码（line code），有时也称为信道编码。

- **如何避免少量差错信号对信息内容的影响**

信息论中的信道编码，即差错控制编码，包括各种纠错、检错码。

5.0 引言

上一章证明了信源编码定理（理想信道编码定理），结论是只要信源序列的信息量不超过信道容量，就总可以找到一种编码，使编、解码过程所造成的误码任意小。本章将证明在有噪信道中，上述结论仍然成立。这就是信道编码定理，即香农第二定理。

由于信道有噪声，因此传输过程会造成误码。信道编码的目的就是**通过编码来减小因传输错误而造成的误码**。注意这里的误码和信源编码中的误码含义是不同的，信源编码的误码是由于编解码过程造成的，而信道编码的误码是由于传输错误造成的。

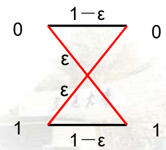
例如，对二元对称信道BSC。

如果不进行信道编码，则误码率为 ε。

假设该误码率无法达到要求，就需要编码。

比如使用三重码编码：0-000、1-111。

可以采用2/3规则判决。



5.0 引言

这种编码方式可以纠一位错。

此时的误码率为：$3\varepsilon^2(1-\varepsilon)+\varepsilon^3=3\varepsilon^2-2\varepsilon^3$。

该编码方法是把比较短的码字（1位）编成比较长的码字（3位），3位长的二进制码字共有8种可能，这里只选用了其中的2种（000和111），通过增加冗余增大了码字间的距离。

从此例可以看到，信道编码与信源编码相比，是反着来的，信源编码追求的是高效，而信道编码追求的是可靠。

可见，为了保证传输正确，需要

- **增加冗余**，利用冗余保证正确恢复。
- **合理地选择码字**，增加码字间的距离。
- **一套解码方法**。判断与哪个码字距离最近。

$$\begin{array}{c} 000 \\ 111 \end{array}\begin{bmatrix} (1-\varepsilon)^3 & (1-\varepsilon)^2\varepsilon & (1-\varepsilon)^2\varepsilon & (1-\varepsilon)^2\varepsilon & (1-\varepsilon)\varepsilon^2 & (1-\varepsilon)\varepsilon^2 & (1-\varepsilon)\varepsilon^2 & \varepsilon^3 \\ \varepsilon^3 & (1-\varepsilon)\varepsilon^2 & (1-\varepsilon)\varepsilon^2 & (1-\varepsilon)\varepsilon^2 & (1-\varepsilon)^2\varepsilon & (1-\varepsilon)^2\varepsilon & (1-\varepsilon)^2\varepsilon & (1-\varepsilon)^3 \end{bmatrix}$$

5.0 引言

信道编码中的一对矛盾——**误码率与传输效率**。

- 为减小误码率，必须增加码字距离，这就需要增加可选码字的集合，而可选码字数 $T=D^N$，它随着 N 的增加指数增加。
- 当 N 增加时，信道传递一个信源字符的时间也随之增加，这意味着传输效率下降。
- 因而在实现低误码率编码的同时，必须考虑其所要求的传输速率方面的代价。所以对信道编码的要求是要权衡这两方面因素，既要降低误码率，又不能使传输速率(或说传输效率)太低。

信道传输速率与解码准则是信道编解码的基础，接下来首先给出信道速率的定义以及两种解码准则，然后讨论二元编码误差，在此基础上证明多元信道编码定理。

5.1 信道速率

定义：信道每用一次所需要传递的信息量。

设信道编码器的输入是长为 L 的 K 进制序列，输出是 N 位长 D 进制序列。可能的输入有 $M=K^L$ 种，输出有 $T=D^N$ 种，$M<T$。

设输入序列是离散无记忆的，其每符号的熵为 $H(U)$。

根据定义，信道速率为：
$$R\triangleq\frac{LH(U)}{N}$$

- 从定义可以看出，信道速率并没有任何时间的概念，而是次的概念。
- 信道速率定义与**互信息定义**的比较。
- **编码过程不改变序列的信息量。**

信源编码理想（编出的符号等概）时：

$$R=\frac{LH(U)}{N}=\frac{L\log K}{N}=\frac{\log K^L}{N}=\frac{\log M}{N}$$

5.1 信道速率

信道编码的目的是在满足误码率要求的前提下，使信道速率尽量接近信道容量，也就是使一次需要传递的信息量尽可能地接近信道容量。从而达到在满足可靠性要求的前提下充分利用信道的目的。

而编码定理要证明的就是：**只要信道速率小于信道容量，总存在一种编码使误码率任意小。**

对理想无噪信道，编码定理需要证明 $R=\frac{LH(U)}{N}<\log D$ 时，误码任意小；（平均码长与输入熵的关系）

对有噪信道情况，编码定理需要证明 $R=\frac{LH(U)}{N}<C$ 时，误码任意小。（信道速率与信道容量的关系）

信道速率与工程中常用的数据传输速率的关系为：$R=r\tau_c$ bit。

其中 r 是每秒传递的比特信息量，τ_c 是信道传递一个数符所需要的时间。

$$R=r\tau_c\ \text{bit}=r\tau_c\ln 2\ \text{nat}$$

5.2 解码准则

前面已提到，系统的错误概率与信道统计特性（前向转移概率矩阵）有关。除此之外，错误概率还与编码方式及译码规则有关。

例如对BSC信道。$\varepsilon\le 1/2$ 与 $\varepsilon>1/2$ 。

译码规则：设离散信道的输入符号集为$X=\{x_i\},i=1,2,\cdots,K$ ，输出符号集为$Y=\{y_j\},j=1,2,\cdots,D$ 。译码规则就是一个函数 $g(y_j)$，它对每一个输出符号，确定**唯一**一个输入符号与其对应（即单值函数）。

$$g(y_j)=x_m \quad x_m\in X;\ j=1,2,\cdots,D$$

译码误差：设$g(y_j)=x_m$，当信道输出端收到y_j时，如果发送端发送的是x_m，即(x_m,y_j)联合发生时，译码是正确的，否则就会译码错误。

$$\mathbf{P}(x,y)=\begin{pmatrix} p(x_1,y_1) & p(x_1,y_2) & \cdots & p(x_1,y_N)\\ p(x_2,y_1) & p(x_2,y_2) & & \\ \vdots & \vdots & \ddots & \vdots \\ p(x_M,y_1) & \cdots & & p(x_M,y_N)\end{pmatrix}$$

5.2 解码准则

1. 解码误差的计算方法

设信源和信宿分别有K 和D个符号消息，译码函数为$g(y_j)=x_m$ ，则由概率间的相互关系可得译码后的误差概率为：

$$P_{\mathrm{E}}=\sum_{j=1}^{D}\sum_{\substack{i=1\\ i\ne m}}^{K}P(x_i,y_j)=\sum_{j=1}^{D}\sum_{\substack{i=1\\ i\ne m}}^{K}P(x_i)P(y_j\,|\,x_i)=\sum_{j=1}^{D}\sum_{\substack{i=1\\ i\ne m}}^{K}P(y_j)P(x_i|y_j)$$

从输出的角度看

$$P_{\mathrm{E}}=\sum_{j=1}^{D}\sum_{\substack{i=1\\ i\ne m}}^{K}P(y_j)P(x_i|y_j)=\sum_{j=1}^{D}P(y_j)\left[1-P(x_m|y_j)\right]$$

从输入的角度看

$$P_{\mathrm{E}}=\sum_{j=1}^{D}\sum_{\substack{i=1\\ i\ne m}}^{K}P(x_i)P(y_j\,|\,x_i)=\sum_{i=1}^{K}P(x_i)\left[1-\sum_{g(y_j)=x_i}P(y_j|x_i)\right]$$

5.2 解码准则

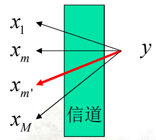
我们当然希望总译码误差越小越好。

2. 最小误差解码准则（Minimum-error）

出发点：使译码后的错误概率 P_E 为最小。

方法：收到 y 后，发那个符号的概率最大，就解码为该符号。即如果 $p(x_{m'}\,|\,y)>p(x_m\,|\,y)\quad \forall m\ne m'$，则解码为$x_{m'}$，即 $g(y)=x_{m'}$ 。

$$\mathbf{P}(x\,|\,y)=\begin{pmatrix} p(x_1|y_1) & p(x_1|y_2) & \cdots & p(x_1|y_N)\\ p(x_2|y_1) & p(x_2|y_2) & & \\ \vdots & \vdots & \ddots & \vdots \\ p(x_M|y_1) & \cdots & & p(x_M|y_N)\end{pmatrix}$$



由于这种方法是通过寻找最大后验概率来进行译码的，故又常称之为**最大后验概率（Maximum A posteriori Probability）准则**。

5.2 解码准则

最大后验概率译码方法是理论上最优（解码误差最小）的译码方法。但在实际译码时，需要计算后验概率，而后验概率的定量计算有时比较困难，因此需要寻找更为实际可行的译码准则。

3. 最大似然解码准则(Maximum-likelihood)

最小误差解码准则：$p(x_{m'}|y)>p(x_m|y)\quad \forall m\neq m'$

$$p(x_{m'}|y)p(y)>p(x_m|y)p(y)\quad \forall m\neq m'$$

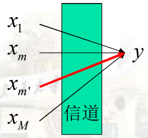
$$p(x_{m'},y)>p(x_m,y)\quad \forall m\neq m'$$

$$p(y|x_{m'})p(x_{m'})>p(y|x_m)p(x_m)\quad \forall m\neq m'$$

当所有输入等概时：$p(y|x_{m'})>p(y|x_m)\quad \forall m\neq m'$

以上式为准则的解码方法称为最大似然解码。

当输入等概时，两种解码准则一致。



5.2 解码准则

例5.2.1：一离散信源的概率空间和信道的前向转移概率矩阵如下：

$$\begin{pmatrix} x \\ p(x)\end{pmatrix}=\begin{Bmatrix} x_1 & x_2 & x_3 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4}\end{Bmatrix}\qquad \mathbf{P}(y|x)=\begin{bmatrix}1/2 & 1/3 & 1/6\\ 1/6 & 1/2 & 1/3\\ 1/3 & 1/6 & 1/2\end{bmatrix}$$

分别求出最小误差解码准则和最大似然解码准则的译码结果，并求出两种译码方法的译码误差（错误传输概率）。

解：对最小误差解码准则，需要求后验概率矩阵。

$$\mathbf{P}(x|y)=\mathbf{P}(x,y)\left[\mathrm{diag}(\mathbf{p}(y))\right]^{-1}$$

$$\mathbf{P}(x,y)=\mathrm{diag}(\mathbf{p}(x))\mathbf{P}(y|x)$$

$$=\begin{bmatrix}1/2 & & \\ & 1/4 & \\ & & 1/4\end{bmatrix}\begin{bmatrix}1/2 & 1/3 & 1/6\\ 1/6 & 1/2 & 1/3\\ 1/3 & 1/6 & 1/2\end{bmatrix}=\begin{bmatrix}1/4 & 1/6 & 1/12\\ 1/24 & 1/8 & 1/12\\ 1/12 & 1/24 & 1/8\end{bmatrix}$$

5.2 解码准则

所以：$\mathbf{p}(y)=[3/8\quad 1/3\quad 7/24]'$

$$\mathbf{P}(x|y)=\mathbf{P}(x,y)\left[\mathrm{diag}(\mathbf{p}(y))\right]^{-1}$$

$$=\begin{bmatrix}1/4 & 1/6 & 1/12\\ 1/24 & 1/8 & 1/12\\ 1/12 & 1/24 & 1/8\end{bmatrix}\begin{bmatrix}8/3 & & \\ & 3 & \\ & & 24/7\end{bmatrix}=\begin{bmatrix}2/3 & 1/2 & 2/7\\ 1/9 & 3/8 & 2/7\\ 2/9 & 1/8 & 3/7\end{bmatrix}$$

根据最小误差解码准则，应译码如下：

$$y_1\to x_1,\ y_2\to x_1,\ y_3\to x_3$$

译码正确的概率为：

$$\begin{aligned}P_C&=p(x_1,y_1)+p(x_1,y_2)+p(x_3,y_3)\\ &=p(y_1)p(x_1|y_1)+p(y_2)p(x_1|y_2)+p(y_3)p(x_3|y_3)\\ &=p(x_1)\left[p(y_1|x_1)+p(y_2|x_1)\right]+p(x_3)p(y_3|x_3)\\ &=1/4+1/6+1/8=13/24\end{aligned}$$

5.2 解码准则

所以译码误差 $P_E=1-P_C=11/24$。

可见，这个信道的传输特性是非常差的。

对最大似然解码准则，直接由信道矩阵(前向转移概率矩阵)

$$\mathbf{P}(y|x)=\begin{bmatrix}1/2 & 1/3 & 1/6\\ 1/6 & 1/2 & 1/3\\ 1/3 & 1/6 & 1/2\end{bmatrix}$$

可得：　$y_1\to x_1,\ y_2\to x_2,\ y_3\to x_3$

此时，

$$P_C'=p(x_1,y_1)+p(x_2,y_2)+p(x_3,y_3)=1/4+1/8+1/8=1/2$$

$$P_E'=1-P_C'=1/2$$

可见，当输入不等概时，最大似然译码的误差比最小误差译码的误差要大一些。

5.2 解码准则

- 可以把上述概念扩大到 L 重扩展的符号序列，这时只要把单符号消息变成符号消息序列，以上各种结论都是成立的。
- 采用最大后验概率准则，或在输入符号先验等概时采用最大似然准则作为译码规则，可使平均错误译码概率 P_E 达到最小值 $P_{E\min}$。但无论采用什么样的译码准则，只要存在噪声，总不能仅通过改进译码方法而使 $P_{E\min}$ 减小到0。
- 实际的通信信道总会存在干扰和噪声，为寻求通信的可靠性，有必要研究专门针对通信可靠性的编码技术。

4. 解码过程

解码过程的本质是对信道的输出划分子集。设信道输入为M个N长的符号序列，输出为 L 个 N 长的符号序列。根据解码准则对输出码字集合划分出 $M+E, E\geq 0$ 个子集，每个子集与输入码字或检出有错一一对应。如输出码字集合划分为：$Y^N=\{y_1^N, y_2^N, \cdots, y_M^N, y_{M+1}^N, \cdots, y_{M+E}^N\}$

5.2 解码准则

其中$\{y_1^N, y_2^N, \cdots, y_M^N\}$分别对应$M$个输入码字。而$\{y_{M+1}^N, \cdots, y_{M+E}^N\}$代表出错集合。

当信道输入 x_m^N 时，如果信道输出 $y^N\in y_m^N$，则传输正确。

令 y_m^C 为 y_m^N 的补集，即 $y_m^C=Y^N\setminus y_m^N$。则输入 x_m^N 时的误码率为：

$$P_{e,m}=\sum_{y^N\in y_m^c} p(y^N\mid x_m^N)$$

平均误码率为：

$$P_e=\sum_{m=1}^{M} p(x_m^N)P_{e,m}$$

例5.2.2：信道编码器输入“0”、“1”两种符号，采用三重码编码方法，其输出通过 $\varepsilon=0.1$ 的BSC信道。用最大似然准则译码，对信道输出划分子集并求出译码误差 $P_{e,1}$、$P_{e,2}$以及 P_e。

5.2 解码准则

解：根据三重码编码：0-000、1-111。信道可能的输出有8种。

首先根据最大似然解码准则划分子集，计算序列的前向转移概率。

$p(y^3\mid x_1^3)=$

$p(000\mid 000)=0.9^3=0.729$

$p(001\mid 000)=0.9^2\times 0.1=0.081$

$p(010\mid 000)=0.9^2\times 0.1=0.081$

$p(100\mid 000)=0.9^2\times 0.1=0.081$

$p(011\mid 000)=0.9\times 0.1^2=0.009$

$p(101\mid 000)=0.9\times 0.1^2=0.009$

$p(110\mid 000)=0.9\times 0.1^2=0.009$

$p(111\mid 000)=0.1^3=0.001$

$p(y^3\mid x_2^3)=$

$p(000\mid 111)=0.1^3=0.001$

$p(001\mid 111)=0.9\times 0.1^2=0.009$

$p(010\mid 111)=0.9\times 0.1^2=0.009$

$p(100\mid 111)=0.9\times 0.1^2=0.009$

$p(011\mid 111)=0.9^2\times 0.1=0.081$

$p(101\mid 111)=0.9^2\times 0.1=0.081$

$p(110\mid 111)=0.9^2\times 0.1=0.081$

$p(111\mid 111)=0.9^3=0.729$

5.2 解码准则

通过比较，输出可以划分为下面两个子集：

$$y_1^3=\{000,001,010,100\},\quad y_2^3=\{111,110,101,011\}$$

由于 $Y^3=y_1^3\cup y_2^3$，$y_1^3\cap y_2^3=\varnothing$，所以 $y_1^c=y_2^3$，$y_2^c=y_1^3$。

下面计算误码率：

$$P_{e,1}=\sum_{y^3\in y_1^c} p(y^3\mid 000)=p(011\mid 000)+p(110\mid 000)+p(101\mid 000)+p(111\mid 000)$$
$$=0.009\times 3+0.001=0.028$$

$$P_{e,2}=\sum_{y^3\in y_2^c} p(y^3\mid 111)=0.009\times 3+0.001=0.028$$

$$P_e=\sum_{m=1}^{2} p(x_m^3)P_{e,m}=p(x_1^3)p_{e,1}+p(x_2^3)p_{e,2}$$
$$=p(x_1^3)\times 0.028+p(x_2^3)\times 0.028=\left[p(x_1^3)+p(x_2^3)\right]\times 0.028=0.028$$

5.3 二元编码误差

理论上，对任何编码都可以用5.2节的方法计算解码误差。然而，这种计算既无法告诉我们应如何选择编码的码长以及合适的码字集合，也不能用于证明编码定理。

因此，我们应该寻找误码率能达到的上界。通过上界不仅可以证明编码定理，还可以深入理解编码参数的选择问题。

信道编码定理的证明非常复杂，为此先讨论二元编码的情况。

信道编码器只有两种输入，经编码后分别对应如下码字：

$$x_1^N=\left[x_{1,1},x_{1,2},\cdots x_{1,n},\cdots x_{1,N}\right],\quad x_2^N=\left[x_{2,1},x_{2,2},\cdots x_{2,n},\cdots x_{2,N}\right]$$

信道输出为：$y^N=\left[y_1,y_2,\cdots y_n,\cdots y_N\right]$，所有输出构成的集合记为 Y^N

采用最大似然解码准则进行解码，

y_1^N 表示解码为 x_1^N 的码字集合，$y_1^C=Y^N\setminus y_1^N$。

y_2^N 表示解码为 x_2^N 的码字集合，$y_2^C=Y^N\setminus y_2^N$。

5.3 二元编码误差

根据5.2节的结论，发 x_1^N 时的误码率为：$P_{e,1}=\sum\limits_{y^N\in y_1^c}p(y^N\mid x_1^N)$。

由于采用最大似然解码，必有

$$p\left(y^N\mid x_2^N\right)\ge p\left(y^N\mid x_1^N\right)\quad \text{if } y^N\in y_1^C$$

所以：

$$P_{e,1}\le\sum_{y^N\in y_1^c}p\left(y^N\mid x_1^N\right)\left[\frac{p\left(y^N\mid x_2^N\right)}{p\left(y^N\mid x_1^N\right)}\right]^s\qquad \forall s\in(0,1)$$

$$=\sum_{y^N\in y_1^c}p\left(y^N\mid x_1^N\right)^{1-s}p\left(y^N\mid x_2^N\right)^s$$

扩大求和范围，进一步放大，得：

$$P_{e,1}\le\sum_{y^N}p\left(y^N\mid x_1^N\right)^{1-s}p\left(y^N\mid x_2^N\right)^s\quad \forall s\in(0,1)$$

同理，$P_{e,2}\le\sum\limits_{y^N}p\left(y^N\mid x_2^N\right)^{1-r}p\left(y^N\mid x_1^N\right)^r\quad \forall r\in(0,1)$

5.3 二元编码误差

取 $r=1-s$，则发不同码字的误码率可写为如下统一形式：

$$P_{e,m}\le\sum_{y^N}p\left(y^N\mid x_1^N\right)^{1-s}p\left(y^N\mid x_2^N\right)^s\quad m=1,2\quad 0<s<1$$

对离散无记忆信道：

$$P_{e,m}\le\sum_{y_1}\sum_{y_2}\cdots\sum_{y_N}\prod_{n=1}^{N}p\left(y_n\mid x_{1,n}\right)^{1-s}p\left(y_n\mid x_{2,n}\right)^s$$

$$=\prod_{n=1}^{N}\sum_{y_n}p\left(y_n\mid x_{1,n}\right)^{1-s}p\left(y_n\mid x_{2,n}\right)^s$$

定义辅助函数：$\quad g_n(s)\triangleq\sum\limits_{y_n}p\left(y_n\mid x_{1,n}\right)^{1-s}p\left(y_n\mid x_{2,n}\right)^s$

则

$$P_{e,m}\le\prod_{n=1}^{N}g_n(s)\qquad m=1,2\qquad 0<s<1$$

得到离散无记忆信道二元编码误码率的上界。

5.3 二元编码误差

下面讨论 $g_n(s)$ 的性质：

- 边值，$s\to 0^+$及 $s\to 1^-$时的极限。

$$g_n(0)\triangleq\sum_{\substack{y_n\\ p(y_n|x_{2,n})\ne 0}}p\left(y_n\middle|x_{1,n}\right)\le 1\qquad g_n(1)\triangleq\sum_{\substack{y_n\\ p(y_n|x_{1,n})\ne 0}}p\left(y_n\middle|x_{2,n}\right)\le 1$$

- 曲线的形状，$g_n(s)$是一个凹函数。

$$g_n(s)=\sum_{y_n}p(y_n\mid x_{1,n})\left[\frac{p(y_n\mid x_{2,n})}{p(y_n\mid x_{1,n})}\right]^s$$

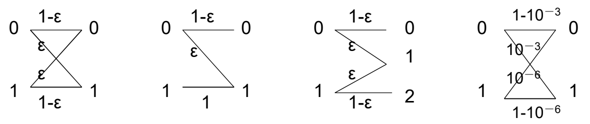
$$g_n'(s)=\sum_{y_n}p(y_n\mid x_{1,n})\left[\frac{p(y_n\mid x_{2,n})}{p(y_n\mid x_{1,n})}\right]^s[\ln p(y_n\mid x_{2,n})-\ln p(y_n\mid x_{1,n})]$$

$$g_n''(s)=\sum_{y_n}p(y_n\mid x_{1,n})\left[\frac{p(y_n\mid x_{2,n})}{p(y_n\mid x_{1,n})}\right]^s[\ln p(y_n\mid x_{2,n})-\ln p(y_n\mid x_{1,n})]^2\ge 0$$

5.3 二元编码误差

例5.3.1：取 $x_{1,n}=0,\ x_{2,n}=1$， $g_n(s)\triangleq\sum_{y_n}p(y_n|x_{1,n})^{1-s}\,p(y_n|x_{2,n})^s$

计算以下四种信道的 $g_n(s)$。



当输出只有两种可能时：

$$g_n(s)\triangleq\sum_{y_n}p(y_n|0)^{1-s}\,p(y_n|1)^s=p(0|0)^{1-s}p(0|1)^s+p(1|0)^{1-s}p(1|1)^s$$

对BSC： $g_n(s)=(1-\varepsilon)^{1-s}\varepsilon^s+\varepsilon^{1-s}(1-\varepsilon)^s$

Figure

对Z信道： $g_n(s)=(1-\varepsilon)^{1-s}0^s+\varepsilon^{1-s}1^s=\varepsilon^{1-s}$

对BEC： $g_n(s)=p(0|0)^{1-s}p(0|1)^s+p(1|0)^{1-s}p(1|1)^s+p(2|0)^{1-s}p(2|1)^s$

$$=(1-\varepsilon)^{1-s}0^s+\varepsilon^{1-s}\varepsilon^s+0^{1-s}(1-\varepsilon)^s=\varepsilon$$

5.3 二元编码误差

计算三种不同编码方式下BSC信道的 $g_n(s)$

（1）所有位均不同，取 $x_1^N=\underbrace{00\cdots0}_{N个}$， $x_2^N=\underbrace{11\cdots1}_{N个}$。（最容易想到的方式）

由例5.3.1， $g_n(s)=(1-\varepsilon)^{1-s}\varepsilon^s+\varepsilon^{1-s}(1-\varepsilon)^s$。

把 s 与 $1-s$ 互换，$g_n(s)$ 值不变，说明其关于 $s=1/2$ 对称，又由于它是凹的，所以 $g_n(s)$ 在 $s=1/2$ 处取得最小值。（几何上）

$$g_n(s)=(1-\varepsilon)\left(\frac{\varepsilon}{1-\varepsilon}\right)^s+\varepsilon\left(\frac{1-\varepsilon}{\varepsilon}\right)^s$$

$$g_n'(s)=(1-\varepsilon)\left(\frac{\varepsilon}{1-\varepsilon}\right)^s\ln\left(\frac{\varepsilon}{1-\varepsilon}\right)+\varepsilon\left(\frac{1-\varepsilon}{\varepsilon}\right)^s\ln\left(\frac{1-\varepsilon}{\varepsilon}\right)=0$$

解得 $s=1/2$。

因此：$\min g_n(s)=g_n(1/2)=2\sqrt{\varepsilon(1-\varepsilon)}$。

5.3 二元编码误差

所以：$P_{e,m}\le\left[g_n(1/2)\right]^N=\left[2\sqrt{\varepsilon(1-\varepsilon)}\right]^N \qquad m=1,2$ 。

（2）一半码元符号相同，一半不同。

$$g_n(s)=p(0|x_{1,n})^{1-s}\,p(0|x_{2,n})^s+p(1|x_{1,n})^{1-s}\,p(1|x_{2,n})^s$$

当 $x_{1,n}=0,\ x_{2,n}=1$ 时，与（1）相同，$g_n(s)=(1-\varepsilon)^{1-s}\varepsilon^s+\varepsilon^{1-s}(1-\varepsilon)^s$。

当 $x_{1,n}=1,\ x_{2,n}=0$ 时， $g_n(s)=p(0|1)^{1-s}\,p(0|0)^s+p(1|1)^{1-s}\,p(1|0)^s$

$$=\varepsilon^{1-s}(1-\varepsilon)^s+(1-\varepsilon)^{1-s}\varepsilon^s$$

当 $x_{1,n}=0,\ x_{2,n}=0$ 时，$g_n(s)=p(0|0)^{1-s}\,p(0|0)^s+p(1|0)^{1-s}\,p(1|0)^s$

$$=p(0|0)+p(1|0)=1$$

当 $x_{1,n}=1,\ x_{2,n}=1$ 时，$g_n(s)=p(0|1)^{1-s}\,p(0|1)^s+p(1|1)^{1-s}\,p(1|1)^s$

$$=p(0|1)+p(1|1)=1$$

5.3 二元编码误差

所以：

$$P_{e,m}\le\left[\min g_n(s)\right]^N=1^{\frac{N}{2}}\cdot\left[2\sqrt{\varepsilon(1-\varepsilon)}\right]^{\frac{N}{2}}=\left[\sqrt{2\sqrt{\varepsilon(1-\varepsilon)}}\right]^N \qquad m=1,2$$

比较（1）与（2）的结果：

由于 $0\le2\sqrt{\varepsilon(1-\varepsilon)}\le1$，所以 $\left[2\sqrt{\varepsilon(1-\varepsilon)}\right]^N\le\left[2\sqrt{\varepsilon(1-\varepsilon)}\right]^{N/2}$ 。

- 对半数位不同的编码，其误码率上界要大于所有位不同的编码。
- 另一方面，方法（2）的可选码字个数要比方法（1）的多。

这说明误码率上界的增加换来了可选码字集的扩大。

（3）随机取码

设随机选中 x_1^N、x_2^N 的概率分别为 $Q(x_1^N)$、$Q(x_2^N)$，选取独立进行，则选中某两个特定码字的概率为 $Q(x_1^N)Q(x_2^N)$。相应的误码率上界为：

$$P_{e,m}\le\sum_{y^N}p(y^N|x_1^N)^{1-s}\,p(y^N|x_2^N)^s \qquad m=1,2 \quad 0<s<1$$

5.3 二元编码误差

由于码字是随机选取的，因此要求其平均误码率。

$$\bar{P}_{e,m} \le \sum_{x_1^N}\sum_{x_2^N} Q\left(x_1^N\right)Q\left(x_2^N\right)\sum_{y^N} p\left(y^N|x_1^N\right)^{1-s} p\left(y^N|x_2^N\right)^s$$

$$= \sum_{y^N}\left[\sum_{x_1^N} Q\left(x_1^N\right)p\left(y^N|x_1^N\right)^{1-s}\right]\left[\sum_{x_2^N}Q\left(x_2^N\right)p\left(y^N|x_2^N\right)^s\right]$$

$\forall s\in(0,1)$ 上式成立，取 $s=1/2$ 得：（事实上，上式关于 $s=1/2$ 对称）

$$\bar{P}_{e,m} \le \sum_{y^N}\left[\sum_{x^N} Q(x^N)\sqrt{p(y^N|x^N)}\right]^2$$

对无记忆信道：

$$\bar{P}_{e,m} \le \sum_{y^N}\left[\sum_{x^N}\prod_{n=1}^{N} Q(x_n)\sqrt{p(y_n|x_n)}\right]^2$$

$$= \sum_{y^N}\left[\prod_{n=1}^{N}\sum_{x_n} Q(x_n)\sqrt{p(y_n|x_n)}\right]^2$$

5.3 二元编码误差

再次交换求和与连乘符号：$\bar{P}_{e,m} \le \prod_{n=1}^{N}\sum_{y_n}\left[\sum_{x_n} Q(x_n)\sqrt{p(y_n|x_n)}\right]^2$

对时不变信道，有：

$$\bar{P}_{e,m} \le \left\{\sum_{y_n}\left[\sum_{x_n} Q(x_n)\sqrt{p(y_n|x_n)}\right]^2\right\}^N$$

对BSC信道，由于 x_1^N、x_2^N 独立随机选取，所以 $Q(0)=Q(1)=1/2$：

$$\bar{P}_{e,m} \le \left\{\sum_{j=0}^{1}\left[\sum_{k=0}^{1} Q(k)\sqrt{p(j|k)}\right]^2\right\}^N$$

$$= \left\{\sum_{j=0}^{1}\left[Q(0)\sqrt{p(j|0)}+Q(1)\sqrt{p(j|1)}\right]^2\right\}^N$$

$$= \left\{\left[Q(0)\sqrt{p(0|0)}+Q(1)\sqrt{p(0|1)}\right]^2+\left[Q(0)\sqrt{p(1|0)}+Q(1)\sqrt{p(1|1)}\right]^2\right\}^N$$

5.3 二元编码误差

代入相应的概率值，得：

$$\bar{P}_{e,m} \le \left\{\left[\frac{1}{2}\sqrt{1-\varepsilon}+\frac{1}{2}\sqrt{\varepsilon}\right]^2+\left[\frac{1}{2}\sqrt{\varepsilon}+\frac{1}{2}\sqrt{1-\varepsilon}\right]^2\right\}^N$$

$$= \left[\frac{1}{2}\left(\sqrt{\varepsilon}+\sqrt{1-\varepsilon}\right)^2\right]^N$$

下面比较三种编码方式的误码率上界：

编码方式	所有位不同	半数位不同	随机选码	
$\bar{P}_{e,m}$	$\left[2\sqrt{\varepsilon(1-\varepsilon)}\right]^N$	$\left[\sqrt{2\sqrt{\varepsilon(1-\varepsilon)}}\right]^N$	$\left[\frac{1}{2}\left(\sqrt{\varepsilon}+\sqrt{1-\varepsilon}\right)^2\right]^N$	1
$\varepsilon=0$	0	0	$\bar{P}_{e,m}=2^{-N}$	
$\varepsilon=1/2$	1	1	1	

当 $N\to\infty$ 时，除强噪信道外，三种上界均趋于零。

5.4 多元编码定理

一、多元编码的误码率上界

定理5.4.1：设信道的前向转移概率矩阵为 $\mathbf{P}(\mathbf{y}|\mathbf{x})$，$M\ge 2$ 为编码符号数。采用最大似然译码准则，则对任意 $1\le m\le M$ 和 $0\le\rho\le 1$，有：

$$\bar{P}_{e,m} \le (M-1)^\rho\sum_{\mathbf{y}}\left[\sum_{\mathbf{x}} q(\mathbf{x})p(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}}\right]^{\rho+1}$$

其中，$\mathbf{x}$ 表示编出的某个码字矢量，设长为 N，$N\ge 1$，即 x^N。

$\mathbf{y}$ 是信道输出，即 y^N。

$q(\mathbf{x})$ 是码字 $\mathbf{x}$ 出现的概率。

总误码率上界

$$\bar{P}_e = \sum_{m=1}^{M} p(\mathbf{x}_m)\bar{P}_{e,m} \le (M-1)^\rho\sum_{\mathbf{y}}\left[\sum_{\mathbf{x}} q(\mathbf{x})p(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}}\right]^{\rho+1}$$

5.4 多元编码定理

证明：

编码方法：随机编码

码字长为N，设采用D进制编码，则共有D^N个可选码字组成可选码字集合。设共有M种消息待传（$M < D^N$）。对可选码字集合进行M次独立随机抽取（抽取后放回），将抽出的码字组成一个码字集合C（码字有可能重复）。$C=\{\mathbf{x}_1,\ldots,\mathbf{x}_m,\ldots,\mathbf{x}_M\}$

总共可以构成D^{NM}个这样的码集。

由随机编码的构造过程可知，抽取任何一个码字的概率相同，且相互独立。

Gallager上界

按照最大似然译码准则译码，若收到$\mathbf{y}$译码为$\mathbf{x}_k$，则必有：

$$p(\mathbf{y}|\mathbf{x}_k) > p(\mathbf{y}|\mathbf{x}_m) \qquad \forall m \neq k$$

5.4 多元编码定理

即有：

$$\frac{p(\mathbf{y}|\mathbf{x}_k)}{p(\mathbf{y}|\mathbf{x}_m)} > 1 \quad \Rightarrow \quad \left[\frac{p(\mathbf{y}|\mathbf{x}_k)}{p(\mathbf{y}|\mathbf{x}_m)}\right]^{\lambda} > 1,\ \forall \lambda \in (0,1)$$

定义如下指示函数：

$$I_k(\mathbf{y}) = \begin{cases} 0 & \text{if } \forall m \neq k\ p(\mathbf{y}|\mathbf{x}_k) > p(\mathbf{y}|\mathbf{x}_m) \\ 1 & \text{if } \exists m \neq k\ p(\mathbf{y}|\mathbf{x}_k) \leq p(\mathbf{y}|\mathbf{x}_m) \end{cases} \qquad m \in \{1,2,\cdots,M\}$$

则发送码字$\mathbf{x}_k$时，判错的概率为：

$$P_e(\mathbf{x}_k) = \sum_{\mathbf{y}} I_k(\mathbf{y}) p(\mathbf{y}|\mathbf{x}_k)$$

$\forall \lambda, \rho \in [0,1]$

$$\left\{\sum_{m \neq k}\left[\frac{p(\mathbf{y}|\mathbf{x}_m)}{p(\mathbf{y}|\mathbf{x}_k)}\right]^{\lambda}\right\}^{\rho} \geq \begin{cases} 0 & p(\mathbf{y}|\mathbf{x}_k) > p(\mathbf{y}|\mathbf{x}_m) \\ 1 & p(\mathbf{y}|\mathbf{x}_k) \leq p(\mathbf{y}|\mathbf{x}_m) \end{cases}$$

5.4 多元编码定理

上式右边为指示函数，即

$$I_k(\mathbf{y}) \leq \left\{\sum_{m \neq k}\left[\frac{p(\mathbf{y}|\mathbf{x}_m)}{p(\mathbf{y}|\mathbf{x}_k)}\right]^{\lambda}\right\}^{\rho}$$

再代入误码率公式，得

$$P_e(\mathbf{x}_k) \leq \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_k)\left\{\sum_{m \neq k}\left[\frac{p(\mathbf{y}|\mathbf{x}_m)}{p(\mathbf{y}|\mathbf{x}_k)}\right]^{\lambda}\right\}^{\rho}$$

取$\lambda = 1/(1+\rho)$，

$$P_e(\mathbf{x}_k) \leq \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_k)^{\frac{1}{1+\rho}}\left\{\sum_{m \neq k}\left[p(\mathbf{y}|\mathbf{x}_m)\right]^{\frac{1}{1+\rho}}\right\}^{\rho}$$

上式即为发送码字矢量$\mathbf{x}_k$时的误码率上界。

5.4 多元编码定理

对确定的码字集合$C=\{\mathbf{x}_1,\ldots,\mathbf{x}_m,\ldots,\mathbf{x}_M\}$，系统的误码率上界为：

$$P_e(C) \leq \sum_{k=1}^{M} p(\mathbf{x}_k) P_e(\mathbf{x}_k)$$

随机编码错误概率上界

随机选码方法，进行M次独立抽取，每次从可选码字集合（共D^N个可选码字）中随机抽取一个作为码字。因此，选中$\{\mathbf{x}_1,\cdots,\mathbf{x}_m,\cdots,\mathbf{x}_M\}$作为码字集合的概率为：$q(\mathbf{x}_1,\cdots,\mathbf{x}_m,\cdots,\mathbf{x}_M) = \prod_{m=1}^{M} q(\mathbf{x}_m)$

对所有可能码字集合的错误概率求统计平均，可得系统的平均错误概率上界：

$$\overline{P}_e = \sum_{\mathbf{x}_1,\cdots,\mathbf{x}_M} q(\mathbf{x}_1,\cdots,\mathbf{x}_m,\cdots,\mathbf{x}_M) P_e(C)$$

$$\leq \sum_{\mathbf{x}_1,\cdots,\mathbf{x}_M} q(\mathbf{x}_1,\cdots,\mathbf{x}_m,\cdots,\mathbf{x}_M) \sum_{k=1}^{M} p(\mathbf{x}_k) P_e(\mathbf{x}_k)$$

5.4 多元编码定理

XI'AN JIAOTONG UNIVERSITY

$$=\sum_{k=1}^{M}p(\mathbf{x}_k)\sum_{\mathbf{x}_1,\cdots,\mathbf{x}_M}q(\mathbf{x}_1,\cdots,\mathbf{x}_m,\cdots,\mathbf{x}_M)P_e(\mathbf{x}_k)$$

定义：

$$\bar{P}_{e,k}=\sum_{\mathbf{x}_1,\cdots,\mathbf{x}_M}q(\mathbf{x}_1,\cdots,\mathbf{x}_m,\cdots,\mathbf{x}_M)P_e(\mathbf{x}_k)$$

对该部分进行放大，使其与k无关。

$$\bar{P}_{e,k}=\sum_{\mathbf{x}_1,\cdots,\mathbf{x}_M}q(\mathbf{x}_1,\cdots,\mathbf{x}_m,\cdots,\mathbf{x}_M)P_e(\mathbf{x}_k)$$

$$\le\sum_{\mathbf{x}_1,\cdots,\mathbf{x}_M}q(\mathbf{x}_1)q(\mathbf{x}_2)\cdots q(\mathbf{x}_M)\sum_{\mathbf{y}}p(\mathbf{y}|\mathbf{x}_k)^{\frac{1}{1+\rho}}\left\{\sum_{m\ne k}\left[p(\mathbf{y}|\mathbf{x}_m)\right]^{\frac{1}{1+\rho}}\right\}^{\rho}$$

$$=\sum_{\mathbf{y}}\sum_{\mathbf{x}_k}q(\mathbf{x}_k)p(\mathbf{y}|\mathbf{x}_k)^{\frac{1}{1+\rho}}\sum_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_M\\ \text{除}\mathbf{x}_k}}q(\mathbf{x}_1)\cdots q(\mathbf{x}_{k-1})q(\mathbf{x}_{k+1})\cdots q(\mathbf{x}_M)\left[\sum_{m\ne k}p(\mathbf{y}|\mathbf{x}_m)^{\frac{1}{1+\rho}}\right]^{\rho}$$

5.4 多元编码定理

XI'AN JIAOTONG UNIVERSITY

对任意 $\rho\in[0,1]$，x^{ρ}关于 x 是上凸的。

$$\sum_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_M\\ \text{除}\mathbf{x}_k}}q(\mathbf{x}_1)\cdots q(\mathbf{x}_{k-1})q(\mathbf{x}_{k+1})\cdots q(\mathbf{x}_M)\left[\sum_{m\ne k}p(\mathbf{y}|\mathbf{x}_m)^{\frac{1}{1+\rho}}\right]^{\rho}$$

$$\le\left[\sum_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_M\\ \text{除}\mathbf{x}_k}}q(\mathbf{x}_1)\cdots q(\mathbf{x}_{k-1})q(\mathbf{x}_{k+1})\cdots q(\mathbf{x}_M)\sum_{m\ne k}p(\mathbf{y}|\mathbf{x}_m)^{\frac{1}{1+\rho}}\right]^{\rho}$$

$$=\left[(M-1)\sum_{\mathbf{x}}q(\mathbf{x})p(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}}\right]^{\rho}$$

代入平均错误概率表达式，得：

$$\bar{P}_{e,k}\le\sum_{\mathbf{y}}\sum_{\mathbf{x}_k}q(\mathbf{x}_k)p(\mathbf{y}|\mathbf{x}_k)^{\frac{1}{1+\rho}}\left[(M-1)\sum_{\mathbf{x}}q(\mathbf{x})p(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}}\right]^{\rho}$$

$$=(M-1)^{\rho}\sum_{\mathbf{y}}\left[\sum_{\mathbf{x}}q(\mathbf{x})p(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}}\right]^{\rho+1}$$

5.4 多元编码定理

XI'AN JIAOTONG UNIVERSITY

离散无记忆信道（DMC）的错误概率上界

长为N的序列通过DMC信道的前向转移概率 $p(\mathbf{y}|\mathbf{x})=\prod_{i=1}^{N}p(y_i|x_i)$ 。

对随机编码 $q(\mathbf{x})=\prod_{i=1}^{N}q(x_i)$

将上面的关系式代入，可得

$$\bar{P}_e\le(M-1)^{\rho}\sum_{y_1}\cdots\sum_{y_N}\left[\sum_{x_1}\cdots\sum_{x_N}q(x_1)p(y_1|x_1)^{\frac{1}{1+\rho}}\cdots q(x_N)p(y_N|x_N)^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

$$=(M-1)^{\rho}\prod_{i=1}^{N}\sum_{y_i}\left[\sum_{x_i}q(x_i)p(y_i|x_i)^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

码字序列中每个符号取自同一个符号集合，具有相同的概率分布。

$$\sum_{y_i}\left[\sum_{x_i}q(x_i)p(y_i|x_i)^{\frac{1}{1+\rho}}\right]^{1+\rho}=\sum_{y}\left[\sum_{x}q(x)p(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

5.4 多元编码定理

XI'AN JIAOTONG UNIVERSITY

所以，DMC的错误概率上界为：

$$\bar{P}_e\le(M-1)^{\rho}\left\{\sum_{y}\left[\sum_{x}q(x)p(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right\}^{N}$$

$$<M^{\rho}\left\{\sum_{y}\left[\sum_{x}q(x)p(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right\}^{N}$$

可靠性函数

由5.2节，信源编码理想时，信道速率 $R=\frac{\ln M}{N}\Rightarrow M=e^{RN}$，则上式可写为：

$$\bar{P}_e<e^{RN\rho}e^{\ln\left\{\sum_{y}\left[\sum_{x}q(x)p(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right\}^{N}}$$

令：$-\ln\sum_{y}\left[\sum_{x}q(x)p(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}=E_0(\rho,\bar{Q})$，$0\le\rho\le1$

5.4 多元编码定理

则：$\overline{P}_e < \exp\left\{-N\left[E_0\left(\rho,\overline{Q}\right)-R\rho\right]\right\} = \exp\left[-NE_r\left(\rho,\overline{Q},R\right)\right]$

可见，错误概率上界与信道速率、码符概率以及参数 ρ 有关。

二、多元编码定理

接下来讨论在什么条件下这个上界可以取得最小值，即达到最紧上界的条件。讨论步骤如下：

- 讨论 $E_0\left(\rho,\overline{Q}\right) = -\ln\sum_j\left[\sum_k Q(k)P(j|k)^{1/(1+\rho)}\right]^{1+\rho}$ 的性质。
- 通过作图讨论 $E_r\left(R,\overline{Q}\right) \triangleq \max_{0\le\rho\le1}\left[E_0\left(\rho,\overline{Q}\right)-\rho R\right]$
- 讨论编码指数 $E_r(R) \triangleq \max_{\overline{Q}} E_r\left(R,\overline{Q}\right)$

$\overline{P}_e \le e^{-NE_r(R)}$

最后，由该式得到误码率与信道速率及码长之间的关系。

5.4 多元编码定理

数学准备：Holder不等式

令 a_i, b_i, p_i 为非负数，$i=1,2,\cdots,N$，且 $\sum_i p_i = 1$，$0<\lambda<1$。

Holder不等式：
$$\sum_i a_i b_i \le \left[\sum_i a_i^{\frac{1}{\lambda}}\right]^{\lambda}\left[\sum_i b_i^{\frac{1}{1-\lambda}}\right]^{1-\lambda}$$

当且仅当对所有 i，$a_i^{1-\lambda} = b_i^{\lambda}\cdot C$ 时上式取等号，其中 C 为常数。

Holder不等式变形1：
$$\sum_i p_i a_i b_i \le \left[\sum_i p_i a_i^{\frac{1}{\lambda}}\right]^{\lambda}\left[\sum_i p_i b_i^{\frac{1}{1-\lambda}}\right]^{1-\lambda}$$

当且仅当对所有 i，$p_i a_i^{\frac{1}{\lambda}} = p_i b_i^{\frac{1}{1-\lambda}}\cdot C$ 时上式取等号，其中 C 为常数。

5.4 多元编码定理

Holder不等式变形2：
$$\left(\sum_i p_i a_i^{r}\right)^{\frac{1}{r}} \le \left(\sum_i p_i a_i^{s}\right)^{\frac{1}{s}} \qquad 0<r<s$$

当且仅当对所有 $p_i>0$，a_i 为常数时上式取等号。

引理5.4.2：令 $\overline{Q} = [Q(0),Q(1),\cdots,Q(K-1)]$ 为一概率向量，令 $a_i \ge 0, i=0,1,\cdots,K-1$。则
$$f(s) = \ln\left[\sum_{k=0}^{K-1} Q(k)a_k^{1/s}\right]^s$$

在 $s>0$ 区间关于 s 非增，是凹函数。且除对于所有 $Q(k)>0$，a_k 为常数外，$f(s)$ 严格减，严格凹。

5.4 多元编码定理

证明：首先证明非增。

由Holder不等式变形2知 $\left[\sum_{k=0}^{K-1} Q(k)a_k^{1/s}\right]^s$

在 $s>0$ 区间关于 s 非增，且除对所有 $Q(k)>0$，a_k 为常数外严格减。对其取自然对数，因为自然对数是严格增函数，所以 $f(s)$ 是非增的，且除 $Q(k)>0$，a_k 为常数外严格减。

再证明是凹函数（弦在曲线上）。用定义证明：

任取 $s_1,s_2>0, s_1\ne s_2$，令 $s=\theta s_1+(1-\theta)s_2$，其中 $0<\theta<1$。

则需要证明：$f(s) \le \theta f(s_1)+(1-\theta)f(s_2)$

令 $\lambda = \dfrac{\theta s_1}{s}$，则 $1-\lambda = \dfrac{(1-\theta)s_2}{s}$，显然 $0<\lambda,1-\lambda<1$

$\because s_1,s_2,s,\theta,1-\theta>0,\ \dfrac{1}{s} = \dfrac{1-\theta}{s}+\dfrac{\theta}{s} = \dfrac{1-\lambda}{s_2}+\dfrac{\lambda}{s_1}$

5.4 多元编码定理

所以：

$$\sum_k Q(k)a_k^{1/s} = \sum_k Q(k)a_k^{\frac{\lambda}{s_1}+\frac{1-\lambda}{s_2}} = \sum_k Q(k)a_k^{\frac{\lambda}{s_1}}a_k^{\frac{1-\lambda}{s_2}} \le \left(\sum_k Q(k)a_k^{\frac{1}{s_1}}\right)^{\lambda}\left(\sum_k Q(k)a_k^{\frac{1}{s_2}}\right)^{1-\lambda}$$

于是：

$$\left(\sum_k Q(k)a_k^{1/s}\right)^{s} \le \left(\sum_k Q(k)a_k^{\frac{1}{s_1}}\right)^{\lambda s}\left(\sum_k Q(k)a_k^{\frac{1}{s_2}}\right)^{(1-\lambda)s}$$

$$= \left(\sum_k Q(k)a_k^{\frac{1}{s_1}}\right)^{\theta s_1}\left(\sum_k Q(k)a_k^{\frac{1}{s_2}}\right)^{(1-\theta)s_2}$$

两边取对数，得：

$$\ln\left(\sum_k Q(k)a_k^{1/s}\right)^{s} \le \theta\ln\left(\sum_k Q(k)a_k^{\frac{1}{s_1}}\right)^{s_1} + (1-\theta)\ln\left(\sum_k Q(k)a_k^{\frac{1}{s_2}}\right)^{s_2}$$

即 $f(s)\le \theta f(s_1)+(1-\theta)f(s_2)$，所以 $f(s)$ 是凹函数。

5.4 多元编码定理

等号成立的条件是：$Q(k)a_k^{\frac{1}{s_1}} = Q(k)a_k^{\frac{1}{s_2}}\cdot C$。

对 $Q(k)>0$，$a_k^{1/s_1-1/s_2}=C$，即 a_k 为常数。

步骤1：$E_0\left(\rho,\bar{Q}\right)$的性质

- 性质1　$\dfrac{\partial^2 E_0\left(\rho,\bar{Q}\right)}{\partial\rho^2}\le 0 \qquad 0\le\rho\le 1$
- 性质2　$I\left(\bar{P};\bar{Q}\right)=\left.\dfrac{\partial E_0\left(\rho,\bar{Q}\right)}{\partial\rho}\right|_{\rho=0}\ge \dfrac{\partial E_0\left(\rho,\bar{Q}\right)}{\partial\rho}\ge \left.\dfrac{\partial E_0\left(\rho,\bar{Q}\right)}{\partial\rho}\right|_{\rho=1}>0 \qquad 0\le\rho\le 1$
- 性质3　$E_0\left(\rho,\bar{Q}\right)\ge 0 \qquad 0\le\rho\le 1$

根据性质绘制 $E_0\left(\rho,\bar{Q}\right)\sim\rho$ 曲线。

5.4 多元编码定理

证明：性质1，任取 $0<\rho_1<\rho_2$，$0<\theta<1$，令 $\rho_3=\theta\rho_1+(1-\theta)\rho_2$。

显然，$1+\rho_3=1+\theta\rho_1+(1-\theta)\rho_2=\theta(1+\rho_1)+(1-\theta)(1+\rho_2)$

所以：$1=\dfrac{\theta(1+\rho_1)}{1+\rho_3}+\dfrac{(1-\theta)(1+\rho_2)}{1+\rho_3}$

令 $\lambda=\dfrac{\theta(1+\rho_1)}{1+\rho_3}$，则 $1-\lambda=\dfrac{(1-\theta)(1+\rho_2)}{1+\rho_3}$

因此，$\dfrac{1}{1+\rho_3}=\dfrac{\theta}{1+\rho_3}+\dfrac{1-\theta}{1+\rho_3}=\dfrac{\lambda}{1+\rho_1}+\dfrac{1-\lambda}{1+\rho_2}$

根据引理5.4.2，令 $s=1+\rho_3, s_1=1+\rho_1, s_2=1+\rho_2$，可得：

$$\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_3}}\right)^{1+\rho_3} \le \left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_1}}\right)^{\theta(1+\rho_1)}\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_2}}\right)^{(1-\theta)(1+\rho_2)}$$

5.4 多元编码定理

两边对 j 求和，并再次利用Holder不等式，

$$\sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_3}}\right)^{1+\rho_3}$$

$$\le \sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_1}}\right)^{\theta(1+\rho_1)}\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_2}}\right)^{(1-\theta)(1+\rho_2)}$$

$$\le \left\{\sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_1}}\right)^{(1+\rho_1)}\right\}^{\theta}\left\{\sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_2}}\right)^{(1+\rho_2)}\right\}^{(1-\theta)}$$

两边取对数，有

$$\ln\sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_3}}\right)^{1+\rho_3}$$

$$\le \theta\ln\left\{\sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_1}}\right)^{(1+\rho_1)}\right\}+(1-\theta)\ln\left\{\sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_2}}\right)^{(1+\rho_2)}\right\}$$

5.4 多元编码定理

所以，$E_0(\rho,\bar{Q})$是关于 ρ的凸函数。性质一得证。 $\frac{\partial^2 E_0(\rho,\bar{Q})}{\partial\rho^2}\le 0$

取等号的条件，（共两次放大）

第一次：对所有 $p(j|k)\neq 0, p(j|k)=\text{const}$

第二次：对所有 j，

$$\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_1}}\right)^{(1+\rho_1)\theta(1-\theta)} = c\cdot\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_2}}\right)^{(1+\rho_2)\theta(1-\theta)}$$

即 $$\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_1}}\right)^{(1+\rho_1)} = c\cdot\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho_2}}\right)^{(1+\rho_2)}$$

性质2：由性质1立即可知，$\partial E_0(\rho,\bar{Q})/\partial\rho$是个减函数。

$$\left.\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}\right|_{\rho=0}\ge \frac{\partial E_0(\rho,\bar{Q})}{\partial \rho}\ge \left.\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}\right|_{\rho=1}$$

5.4 多元编码定理

由Holder不等式变形2，除对所有 $p(j|k)\neq 0, p(j|k)=\text{const}$ 外，

$\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho}}\right)^{(1+\rho)}$ 是 ρ的严格减函数。

因此，

$\sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho}}\right)^{(1+\rho)}$ 也是 ρ的严格减函数。

所以，

$E_0(\rho,\bar{Q})=-\ln\sum_j\left(\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho}}\right)^{(1+\rho)}$ 是 ρ的严格增函数。

即 $\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}>0$

接下来，求偏导，令 $f(\rho)=\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho}}$，$g(\rho)=1+\rho$

$$\left[f(\rho)^{g(\rho)}\right]'=\left[f(\rho)^{g(\rho)}\right]\left[g'(\rho)\ln f(\rho)+g(\rho)\frac{f'(\rho)}{f(\rho)}\right]$$

5.4 多元编码定理

因此， $E_0(\rho,\bar{Q})=-\ln\sum_j f(\rho)^{g(\rho)}$。

$$\left.\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}\right|_{\rho=0} = -\frac{\sum_j\left[f(\rho)^{g(\rho)}\right]'}{\sum_j f(\rho)^{g(\rho)}} = -\left.\frac{\sum_j \left[f(\rho)^{g(\rho)}\right]'}{\sum_j\left[\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho}}\right]^{1+\rho}}\right|_{\rho=0}$$

$$=-\sum_j f(\rho)^{g(\rho)}\cdot\left.\left[g'(\rho)\ln f(\rho)+(1+\rho)\frac{\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho}}\ln p(j|k)\cdot\frac{-1}{(1+\rho)^2}}{\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho}}}\right]\right|_{\rho=0}$$

$$=-\sum_j\left\{\sum_k Q(k)p(j|k)\left[\ln\sum_k Q(k)p(j|k)-\frac{\sum_k Q(k)p(j|k)\ln p(j|k)}{\sum_k Q(k)p(j|k)}\right]\right\}$$

5.4 多元编码定理

$$=-\sum_j\sum_k Q(k)p(j|k)\ln p(j)+\sum_j\sum_k Q(k)p(j|k)\ln p(j|k)$$

$$=I(X;Y)=I(\bar{P};\bar{Q})$$

性质3： $$E_0(0,\bar{Q})=-\ln\sum_j\left.\left[\sum_k Q(k)p(j|k)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right|_{\rho=0}$$

$$=-\ln\sum_j\sum_k Q(k)p(j|k)=-\ln 1=0$$

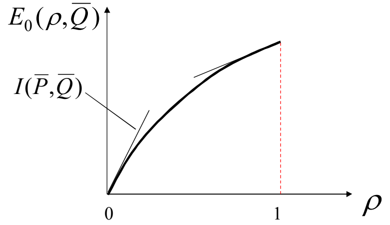
由性质2，其一阶导数大于零，所以 $E_0(\rho,\bar{Q})\ge E_0(0,\bar{Q})=0$。

总结性质

- 性质1，它是凸函数，其一阶导函数是个减函数。
- 性质2，它是增函数。
- 性质3，函数恒为正，即在第一象限。

5.4 多元编码定理

根据以上性质，绘出$E_0(\rho,\bar{Q}) \sim \rho$的曲线如下：



步骤2：求 $E_r(R,\bar{Q}) \triangleq \max\limits_{0\le\rho\le1}\left[E_0(\rho,\bar{Q})-\rho R\right]$

$$\frac{\partial\left[E_0(\rho,\bar{Q})-\rho R\right]}{\partial\rho}=\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}-R$$

5.4 多元编码定理

因此驻点应满足：$\partial E_0(\rho,\bar{Q})/\partial\rho = R$。

由$E_0(\rho,\bar{Q})$性质2，当

$$\left.\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}\right|_{\rho=0} \ge R \ge \left.\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}\right|_{\rho=1}$$

时，在 $0\le\rho\le1$ 区间有驻点，否则最大值只能在边界取得。

此时，$E_r(R,\bar{Q})=E_0(\rho,\bar{Q})-\rho\,\partial E_0(\rho,\bar{Q})/\partial\rho$。

随着 ρ从0变到1，R 从 $I(\bar{P};\bar{Q})$ 单调减小到 $\partial E_0/\partial\rho|_{\rho=1}$；

而 $E_r(R,\bar{Q})$ 从0单调增加到 $E_0(1,\bar{Q})-\partial E_0(\rho,\bar{Q})/\partial\rho|_{\rho=1}$。

在此区间内，$E_r(R,\bar{Q})\sim R$ 曲线的斜率为：

$$\frac{\partial E_r(R,\bar{Q})}{\partial R}=-\rho$$

5.4 多元编码定理

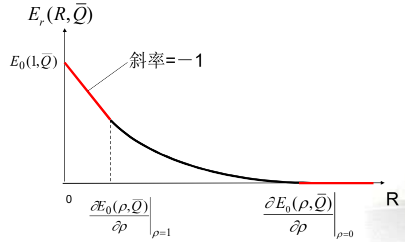
当$\rho=0$时，$E_0(\rho,\bar{Q})-\rho R=0$

当$\rho=1$时，$E_0(\rho,\bar{Q})-\rho R=E_0(1,\bar{Q})-R$

因此，

当 $R<\partial E_0(\rho,\bar{Q})/\partial\rho|_{\rho=1}$ 时，$E_r(R,\bar{Q})=E_0(1,\bar{Q})-R$。

而当$R>I(\bar{P};\bar{Q})$时，$E_r(R,\bar{Q})=0$。



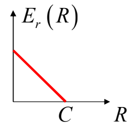
5.4 多元编码定理

步骤3：求 $E_r(R)=\max\limits_{Q}E_r(R,\bar{Q})$，绘$E_r(R)\sim R$曲线

对$\bar{Q}$优化，可以使互信息熵$I(\bar{P};\bar{Q})$最大达到信道容量C，曲线朝右走。

特殊情况：当 $\frac{\partial^2 E_0(\rho,\bar{Q})}{\partial\rho^2}=0$时，$\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}=\text{const}$。

$$\left.\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}\right|_{\rho=0}=\left.\frac{\partial E_0(\rho,\bar{Q})}{\partial\rho}\right|_{\rho=1}$$



定理5.4.3（有噪信道编码定理）：对于任何离散无记忆信道（DMC），随机编码指数 $E_r(R)$在 $0\le R<C$范围内，是凹的、单调减和正函数。

$$\bar{P}_e\le e^{-NE_r(R)}\quad R<C,\quad E_r(R)>0,\quad N\uparrow,\bar{P}_e\downarrow,\quad N\to\infty,\bar{P}_e\to0$$

5.5 本章小结

- 信道速率的定义、意义及计算
- 最大似然解码准则、最小误差解码准则，划分子集，解码误差的计算。
- 二元编码误码率的计算

$$M=2,\quad P_e\le\prod_{n=1}^{N}g_n(s)\qquad P_e\le\prod_{n=1}^{N}\min g_n(s)$$

- 多元编码误码率

$$M>2,\quad \overline{P}_e\le e^{-N\left[E_0(\rho,\overline{Q})-\rho R\right]}\qquad \overline{P}_e\le e^{-NE_r(R)}$$

- $E_0\left(\rho,\overline{Q}\right)$的性质、曲线
- 编码指数的性质、曲线（有噪信道编码定理）

$$\mathbf{P}(y|x)=\begin{pmatrix} p(y_1|x_1) & p(y_2|x_1) & \cdots & p(y_N|x_1)\\ p(y_1|x_2) & p(y_2|x_2) & & \\ \vdots & \vdots & \ddots & \vdots\\ p(y_1|x_M) & \cdots & & p(y_N|x_M)\end{pmatrix}$$

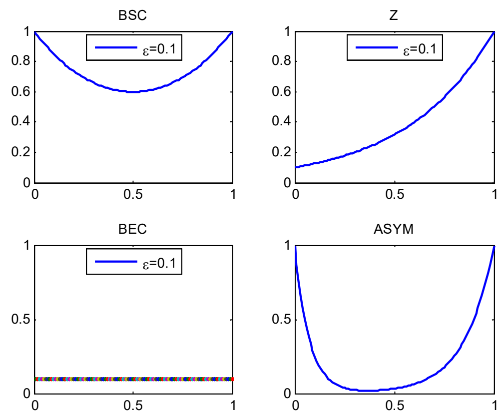
$$\mathbf{P}(x,y)=\begin{pmatrix} p(x_1,y_1) & p(x_1,y_2) & \cdots & p(x_1,y_N)\\ p(x_2,y_1) & p(x_2,y_2) & & \\ \vdots & \vdots & \ddots & \vdots\\ p(x_M,y_1) & \cdots & & p(x_M,y_N)\end{pmatrix}\qquad \begin{pmatrix} p(x_1)\\ p(x_2)\\ \vdots\\ p(x_M)\end{pmatrix}=\mathbf{p}(x)$$

$$\mathbf{p}(y)=\left(p(y_1)\quad p(y_2)\quad\cdots\quad p(y_N)\right)'$$

$$\mathbf{P}(x,y)=\mathrm{diag}\left(\mathbf{p}(x)\right)\mathbf{P}(y|x)$$

$$\mathbf{P}(x|y)=\begin{pmatrix} p(x_1|y_1) & p(x_1|y_2) & \cdots & p(x_1|y_N)\\ p(x_2|y_1) & p(x_2|y_2) & & \\ \vdots & \vdots & \ddots & \vdots\\ p(x_M|y_1) & \cdots & & p(x_M|y_N)\end{pmatrix}$$

$$\mathbf{P}(x,y)=\mathbf{P}(x|y)\,\mathrm{diag}\left(\mathbf{p}(y)\right)$$



几种信道的$g_n(s)$

5.3 二元编码误差

比较后两者：

第3种可化为：$\overline{P}_{e,m}=\left[\frac{1}{2}\left(\sqrt{\varepsilon}+\sqrt{1-\varepsilon}\right)^2\right]^N=\left[\frac{1}{2}+\sqrt{\varepsilon(1-\varepsilon)}\right]^N$

由

$$\left[\frac{1}{2}-\sqrt{\varepsilon(1-\varepsilon)}\right]^2=\frac{1}{4}-\sqrt{\varepsilon(1-\varepsilon)}+\varepsilon(1-\varepsilon)\ge 0$$

可得：$\frac{1}{4}+\varepsilon(1-\varepsilon)\ge\sqrt{\varepsilon(1-\varepsilon)}$ *

又 $\left[\frac{1}{2}+\sqrt{\varepsilon(1-\varepsilon)}\right]^2=\frac{1}{4}+\sqrt{\varepsilon(1-\varepsilon)}+\varepsilon(1-\varepsilon)$

代入*式，可得：

$$\left[\frac{1}{2}+\sqrt{\varepsilon(1-\varepsilon)}\right]^2\ge 2\sqrt{\varepsilon(1-\varepsilon)}\Rightarrow\frac{1}{2}+\sqrt{\varepsilon(1-\varepsilon)}\ge\sqrt{2\sqrt{\varepsilon(1-\varepsilon)}}$$