第三章离散信道及其容量

第三节离散无记忆信道容量

陈兴同

中国矿业大学 数学学院

2021 年 8 月

内容提要

1. 信道容量定义

内容提要

内容提要

内容提要

信道容量释义:

信道容量是由信道的统计特征所决定的。对离散无记忆信道，输入信号 X 与输出信号 Y 之间的互信息 $I(X;Y)$ 就是信道传输的信息量。

显然对不同的输入分布互信息是不同的值，信道传输的信息量也不同。人们总是希望对于给定的离散无记忆信道能用它的最大容量进行传输。信道所能传输的最大信息量称为信道容量。

输入输出公共信息量:

设有一个离散无记忆信道 $\{\mathcal{X}, Q(y|x), \mathcal{Y}\}$。
（1）对应于**输入字符集** $\mathcal{X} = \{a_1, a_2, \cdots, a_N\}$ 上的分布 $X \sim p_X(x), x \in \mathcal{X}$ 称为这个信道的**输入分布**。
（2）对于所给的输入分布可求出**联合分布**

$$p(x, y) = P\{X = x, Y = y\} = p_X(x)Q(y|x), (x, y) \in \mathcal{X} \times \mathcal{Y}.$$

（3）对应**输出字符集** $\mathcal{Y} = \{b_1, b_2, \cdots, b_N\}$ 上的分布 $X \sim p_Y(y), y \in \mathcal{Y}$ 称为这个信道的**输出分布**。对于所给的输入分布可求出**输出分布**

$$p_Y(y) = \sum_x p(x, y) = \sum_x p_X(x)Q(y|x), y \in \mathcal{Y}.$$

续：信道输入输出公共信息量

（4）对于给定的输入分布 p_X，输入随机变量 X 与输出随机变量 Y 之间互信息为

$$I(X;Y) = \sum_x p_X(x) \sum_y Q(y|x) \log \frac{Q(y|x)}{\sum_{x'} p_X(x')Q(y|x')} = f(p, Q),$$

因此互信息是信道的输入分布 $p = (p_1, p_2, \cdots, p_N)$ 及信道矩阵 $Q(y|x)$ 的多元函数。但是信道矩阵已知，因此互信息仅为输入分布的多元函数。

续：信道输入输出公共信息量

（5）记全体输入分布的集合

$$\mathcal{P}_{in} = \left\{ p = (p_1, p_2, \cdots, p_N) | p_i \geq 0, \sum_{i=1}^{N} p_i = 1 \right\},$$

则它构成 N 维欧氏空间 $\mathbb{R}^n$ 中的有界凸集合，可以证明它也是闭集。

（6）由于当信道矩阵 $Q(y|x)$ 固定时，互信息是信道的输入分布 p 的凹函数，从而互信息 $f(p, Q)$ 在闭凸集合 P_{in} 上有最大值。

信道容量定义 3.3.1

*离散无记忆**信道容量**是*

$$C = \max_{p \in \mathcal{P}_{in}} I(X;Y) = \max_{p \in \mathcal{P}_{in}} f(p, Q).$$

*称达到信道容量的输入分布称为**最大输入分布**，对应的输出分布称为**最大输出分布**。*

若信道发送一个比特平均需要 t 秒，则信道单位时间内最大信息传输量为 $C_t = C/t$，称为信道的**信息传输速率**，单位为 bps，表示每秒发送的比特数。信息传输速率也可以作为信道容量的度量。通常数字信道也有一定的带宽，它指单位时间内信道能够传输的最大比特数，称为**数据传输速率**，单位是 bps 或 Kbps,Mbps 等。

例题 3.3.1

如图 3-3 称为**二进删除信道**，现在设 $p = 0.2$。如果输入分布为 $p(a_1) = q, p(a_2) = 1 - q$ ，求这个信道的联合分布、输出分布以及**反向传输矩阵**。

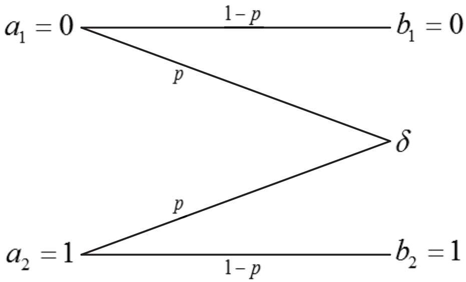


Figure: 二进删除信道

3-3

解：

信道矩阵为

$$Q = \begin{pmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.2 & 0.8 \end{pmatrix},$$

从而由乘法公式 $p(x,y) = p(y|x)p_X(x)$ 得联合分布律

$(X,Y) \sim p(x,y) \sim$

	b_1	δ	b_2
a_1	$0.8q$	$0.2q$	0
a_2	0	$0.2(1-q)$	$0.8(1-q)$

比如：

$$p(a_1,b_1) = p(b_1|a_1)p(a_1) = 0.8q,$$

$$p(a_2,\delta) = p(\delta|a_2)p(a_2) = 0.2(1-q)$$

并且输出分布为

$$p_Y(b_1) = 0.8q, p_Y(\delta) = 0.2, p_Y(b_2) = 0.8(1-q),$$

续解：

于是互信息量 $I(X;Y)$

$$
\begin{aligned}
&= \sum_x p_X(x) \sum_y p(y|x) \log \frac{p(y|x)}{p_Y(y)} \\
&= p(a_1)\left[p(b_1|a_1)\log\frac{p(b_1|a_1)}{p_Y(b_1)} + p(b_2|a_1)\log\frac{p(b_2|a_1)}{p_Y(b_2)}\right. \\
&\quad \left. + p(\delta|a_1)\log\frac{p(\delta|a_1)}{p_Y(\delta)}\right] \\
&\quad + p(a_2)\left[p(b_1|a_2)\log\frac{p(b_1|a_2)}{p_Y(b_1)} + p(b_2|a_2)\log\frac{p(b_2|a_2)}{p_Y(b_2)}\right. \\
&\quad \left. + p(\delta|a_2)\log\frac{p(\delta|a_2)}{p_Y(\delta)}\right] \\
&= 0.8\left(q\log\frac{1}{q} + (1-q)\log\frac{1}{1-q}\right) = 0.8h(q),
\end{aligned}
$$

续解：

显然，当 $h(q)$ 达到最大值 1bit 时 $I(X;Y)$ 达到最大值 0.8bits，故这个信道的容量为 0.8bits。另外可由 Bayes 公式求出反向条件概率

$$p(a_1|b_1)=\frac{p(a_1,b_1)}{p_Y(b_1)}=\frac{0.8q}{0.8q}=1, p(a_2|b_1)=0,$$
$$p(a_1|\delta)=\frac{p(a_1,\delta)}{p_Y(\delta)}=\frac{0.2q}{0.2}=q, p(a_2|\delta)=1-q,$$
$$p(a_1|b_2)=\frac{p(a_1,b_2)}{p_Y(b_2)}=0, p(a_2|b_2)=\frac{p(a_2,b_2)}{p_Y(b_2)}=\frac{0.8(1-q)}{0.8(1-q)}=1,$$

故反向条件概率矩阵为

$$\tilde{Q}(x|y)=\begin{pmatrix}1 & 0\\ q & 1-q\\ 0 & 1\end{pmatrix}.$$

例题 3.3.2：理想信道

如图 3-4，这种信道的输入在传输过程没有受到任何噪声干扰，并且输入输出之间有一一对应关系，因而信道矩阵是单位矩阵，条件熵 $H(X|Y)=0$，

$$I(X;Y)=H(X)-H(X|Y)=H(X),$$

从而信道容量为

$$C=\max_{p_X(x)} I(X;Y)=\max_{p_X(x)} H(X)=\log N,$$

最大输入、输出分布都是等可能分布：$p_X(x)=1/N, x\in\mathcal{X}, p_Y(y)=1/N, y\in\mathcal{Y}$。

传输图 3-4

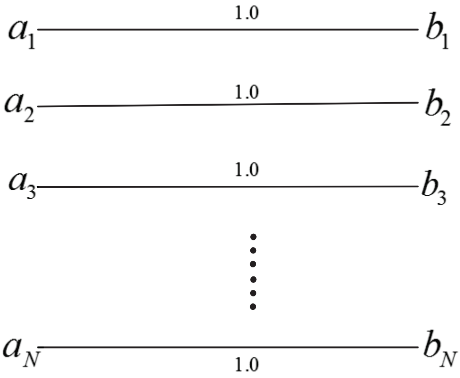


Figure: 图 3-4

例题 3.3.3

如图 3-5，这种信道也没有受到任何噪声干扰，它的输出对应多个互不相交的输入。若已知输入则输出就确定了，故条件概率 $p(y|x)=0$或1，条件熵为 $H(Y|X)=0$，从而

$$I(X;Y)=H(Y)-H(Y|X)=H(Y)\leq\log M,$$

$$C=\max_{p_X(x)}I(X;Y)=\max_{p_X(x)}H(Y)\leq\log M,$$

但是可以构造一个输入分布使输出 Y 服从等可能分布，从而信道容量为 $C=\log M$。

传输图 3-5

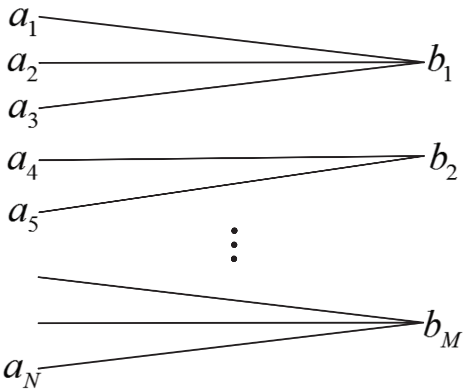


Figure: 图 3-4

例题 3.3.4

如图 3-6，它的输入对应多个互不相交的的输出。从输出可以准确断定输入，因而条件概率 $p(x|y)=1$或0，条件熵为 $H(X|Y)=0$，从而

$$I(X;Y)=H(X)-H(X|Y)=H(X),$$

信道容量为

$$C=\max_{p_X(x)} I(X;Y)=\max_{p_X(x)} H(X)=\log N,$$

这说明输入信息在传输过程中，没有损失。最大输入分布是等可能分布 $p_X(x)=1/N, x\in\mathcal{X}$，但最大输出分布未必是等可能分布。

传输图 3-6 P56

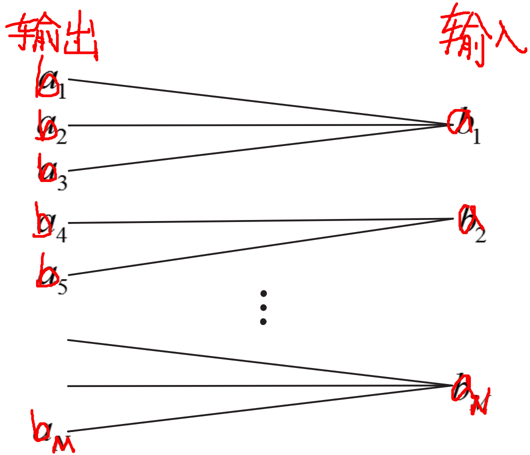


Figure: 图 3-~~4~~6

例题 3.3.5

3-2a

（**二进对称信道**）如图 ??，简记为 BSC，它的每个输入都以相同概率错成另一个输出，求信道容量。

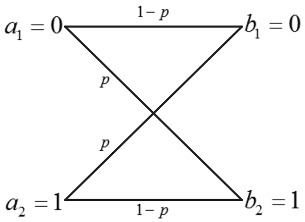


Figure: 图 3-2a

解：

其信道矩阵恰为对称矩阵

$$Q = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix},$$

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

当输入分布 $p_X(x) = 1/2$ 时，输出分布 $p_Y(y) = 1/2$，恰好也是等可能分布，从而 $C = 1 - h(p)$bits。

例题 3.3.6

（无用信道）如图 3-7，它的每个输入字符传错的概率都与另一个字符正确传输的概率相同，求信道容量。

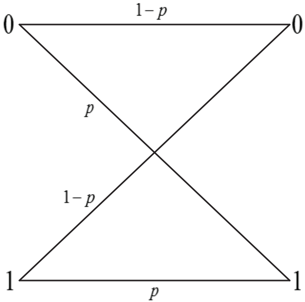


Figure: 图 3-7

解：

其信道矩阵为

$$Q = \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix},$$

对任何输入分布 $p_X(0) = q, p_X(1) = 1-q$，其输出分布为 $p_Y(0) = 1-p, p_Y(1) = p$，从而

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) = H(Y) - \sum_x p_X(x) H(Y|X=x) \\ &= H(Y) - \sum_x p_X(x) h(p) = H(Y) - h(p) \\ &= h(p) - h(p) = 0, \end{aligned}$$

故信道容量 $C = 0$。对此可以这样理解：当收到一个字符时推断发出字符不可能，因为发出 a_1, a_2 的概率相同。

例题 3.3.7

（二元删除信道）如图 3-3，求信道容量。

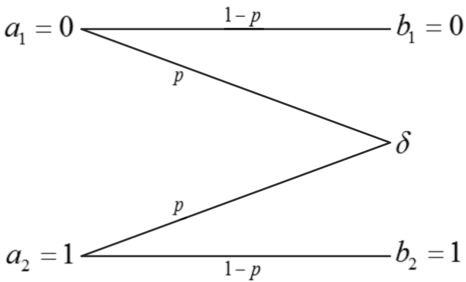


Figure: 二进删除信道

3-3

解：

信道矩阵为

$$Q = \begin{pmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{pmatrix}$$

对任何输入分布 $p_X(0) = q, p_X(1) = 1-q$，可求出输出分布为

$$p_Y(0) = q(1-p), p_Y(\delta) = p, p_Y(1) = (1-q)(1-p).$$

续解：

从而可求得

$$
\begin{aligned}
H(Y) &= -q(1-p)\log(q(1-p)) - p\log p - (1-q)(1-p)\log(1-q) \\
&= -q(1-p)\log(q(1-p)) - p\log p - (1-p)\log(1-q)(1-p) \\
&\quad +q(1-p)\log(1-q)(1-p) \\
&= q(1-p)\log\frac{1-q}{q} - p\log p \\
&\quad -(1-p)\log(1-p) - (1-p)\log(1-q) \\
&= (1-p)\left[q\log\frac{1-q}{q} - \log(1-q)\right] + h(p) \\
&= (1-p)\left[-q\log q - (1-q)\log(1-q)\right] + h(p) \\
&= (1-p)h(q) + h(p),
\end{aligned}
$$

续解：

$$
\begin{aligned}
I(X;Y) &= H(Y)-H(Y|X) = H(Y)-\sum_x p_X(x)H(Y|X=x) \\
&= H(Y)-\sum_x p_X(x)h(p) \\
&= H(Y)-h(p) \\
&= (1-p)h(q),
\end{aligned}
$$

当输入分布为等可能分布时，$h(q)$ 达到最大值 1bits，所以信道容量为：

$$C = \max_{p_X\in\mathcal{P}} I(X;Y) = \max_{0\leq q\leq 1}(1-p)h(q) = 1-p \text{ bits},$$

此时输出分布为：

$$Y \sim p_Y(y) = \begin{pmatrix} 0 & \delta & 1 \\ \frac{1-p}{2} & p & \frac{1-p}{2} \end{pmatrix}.$$

几个常用的离散无记忆信道

现在来建立几个常用的离散无记忆信道容量计算公式，为此根据信道矩阵的特点对信道进行分类：

（1）**输入对称信道**指信道矩阵的每一行元素都是另一行元素的重新排列。

（2）**输出对称信道**指信道矩阵的每一列元素都是另一列元素的重新排列。

（3）**对称信道**指既是输入对称又是输出对称的信道。

（4）**弱对称信道**指输入对称的信道，同时矩阵的每一列元素之和相等。

几个常用的离散无记忆信道

（5）**准对称的信道**指将信道矩阵的列进行适当重新排序后，再进行分块，使得在每个块矩阵中每一行都是其它的行的重新排列，每一列也是其它的列的重新排列；即将输出集 $\mathcal{Y}$ 划分成若于个互不相交的子集 $\mathcal{Y}_1, \mathcal{Y}_2, \cdots, \mathcal{Y}_m$，使得每个子集所的对应的信道矩阵的列构成的块矩阵 $Q_1, Q_2, \cdots, Q_m$ 中的每个都具有性质：每行都是其它行的重新排列，每列也是其它的列的重新排列。

例题 3.3.8

下面几个信道矩阵决定的信道属于哪种类型？

$$Q_1 = \begin{pmatrix} 1/3 & 1/3 & 1/6 & 1/6 \\ 1/6 & 1/3 & 1/6 & 1/3 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 1/3 & 1/6 & 1/2 \\ 1/3 & 1/2 & 1/6 \end{pmatrix},$$

$$Q_3 = \begin{pmatrix} 1/3 & 1/3 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/3 & 1/3 \end{pmatrix}, \quad Q_4 = \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/6 & 1/2 & 1/3 \\ 1/3 & 1/6 & 1/2 \end{pmatrix}.$$

$Q_1 - --$ 准对称信道，但不是弱对称信道。$Q_2 - --$ 准对称信道，又是弱对称信道。$Q_3 - --$ 对称信道。$Q_4 - --$ 对称信道。对称信道的矩阵未必是对称信道。

引理 3.3.1

离散无记忆对称信道、弱对称信道、输出对称信道，当输入分布为等可能分布时，输出分布也是等可能分布。

证明：设输入分布为 $X \sim p_X(x) = 1/N, x \in \mathcal{X}$，则输出分布为

$$Y \sim p_Y(y) = \sum_x p_X(x)p(y|x) = \frac{1}{N}\sum_x p(y|x).$$

但是对于对称信道、弱对称信道、输出对称信道其信道矩阵中任一列的和是相同的，即

$$\sum_x p(y|x) = S$$

与列号 y 无关，从而对每个 y 概率 $p_Y(y)$ 是常数，故只能是等可能分布。

命题 3.3.1：弱对称信道容量公式

对于对称信道、弱对称信道，记 $Q^{(i)} = (p_{1|i}, p_{2|i}, \cdots, p_{M|i})$ *为信道矩阵的第* i *行，则有信道容量公式*

$$C = \log M - h(Q^{(i)}), \tag{3.1}$$

其中 $h(Q^{(i)})$ *是概率分布* $Q^{(i)}$ *的熵；最大输入分布与最大输出分布都为等可能分布* $p_X(x) = 1/N, p_Y(y) = 1/M$*。*

证明：

证明：对于对称与弱对称信道，它们的信道矩阵中任一行都是相同元素的排列，故每行条件熵相同即 $H(Y|X=x)=h(Q^{(i)})$，并且

$$H(Y|X)=\sum_x p(x)H(Y|X=x)=h(Q^{(i)})\sum_x p(x)=h(Q^{(i)}),$$

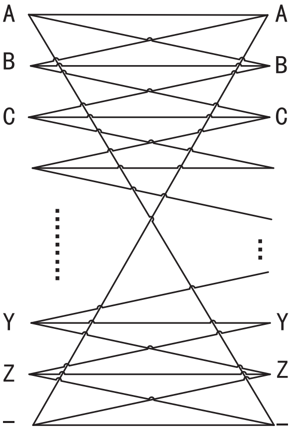
由引理 3.3.1，最大输出分布为等可能分布, 故可达到最大熵 $\log M$，从而信道容量为

$$I(X;Y)=H(Y)-H(Y|X)=H(Y)-h(Q^{(i)})\leq \log M-h(Q^{(i)})=C.$$

利用这个定理可以计算一些对称与弱对称信道的容量及其最大输入、输出分布。

例题 3.3.9

设信道输出字符集分别为 $\mathcal{X}, \mathcal{Y} = \{A, B, \cdots, Z, -\}$。当输入一个字符时，以等概率输出相邻的三个字符，如图 3-8 所示，求这个信道的容量。



解：

它的信道矩阵为

$$Q = \begin{pmatrix} 1/3 & 1/3 & 0 & \cdots & \cdots & 1/3 \\ 1/3 & 1/3 & 1/3 & \cdots & \cdots & 0 \\ 0 & 1/3 & 1/3 & 1/3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1/3 & 1/3 \\ 1/3 & 0 & 0 & \cdots & 1/3 & 1/3 \end{pmatrix},$$

显然这是对称信道，信道容量为

$$C = \log 27 - \log 3 = \log 9,$$

最大输入、输出分布都为等可能分布

$$p_X(x) = p_Y(y) = 1/27.$$

例题 3.3.10：

如图 3-9 所示，求这个信道的容量。

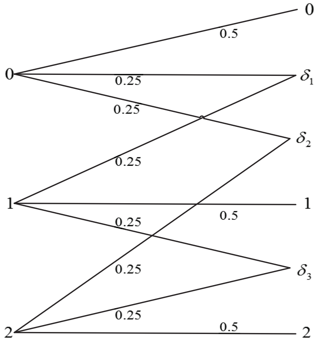


Figure: 图 3-9

解：

信道矩阵为

$$\begin{pmatrix} 0.5 & 0.25 & 0.25 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0.5 & 0.25 & 0 \\ 0 & 0 & 0.25 & 0 & 0.25 & 0.5 \end{pmatrix},$$

显然它是一个弱对称信道，并且

$$h(Q^{(i)}) = -0.5 \ln 0.5 - 0.25 \ln 0.25 - 0.25 \ln 0.25 = \ln \sqrt{8},$$

因此信道容量为

$$C = \ln 6 - h(Q^{(i)}) = \ln \frac{6}{\sqrt{8}} = 0.752 \text{ nats}.$$

例题 3.3.11：

（模 K 的加法信道）如图 3-10，随机变量 X, Y, Z 字符集均为 $\{0, 1, \cdots, K-1\}$，并且随机变量 Z 具有指定的分布 $Z \sim p_Z(z), Y = (X + Z)mod(K)$，求信道容量。

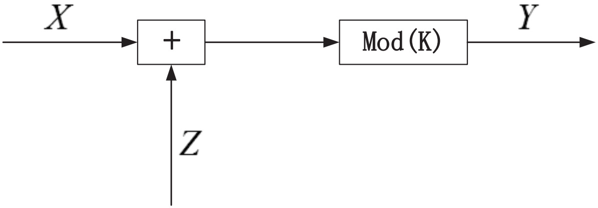


Figure: 图 3-10

解：

这需要先求出信道矩阵。以 $K=3$ 为例信道矩阵为

$$\begin{pmatrix} p_Z(0) & p_Z(1) & p_Z(2) \\ p_Z(2) & p_Z(0) & p_Z(1) \\ p_Z(1) & p_Z(2) & p_Z(0) \end{pmatrix},$$

事实上：

$$\begin{aligned} p(0|2) &= P\{Y=0|X=2\} = P\{Z=1\} = p_Z(1), \\ p(1|2) &= P\{Y=1|X=2\} = P\{Z=2\} = p_Z(2), \\ p(2|2) &= P\{Y=2|X=2\} = P\{Z=0\} = p_Z(0). \end{aligned}$$

这是一个对称信道，从而

$$C = \log K - H(Z).$$

例题 3.3.12:

有一个二元对称信道如图 3-2a 所示，其中 $p = 0.02$，该信道数据传输速率为 1500 bits/秒。现有一消息序列共有 14000 bits，并设在这消息中 $p(0) = p(1) = 1/2$。问从信息传输的角度来考虑，10 秒钟内能否将这消息序列无失真地通过信道传送完？

解：

本题要弄清信道**数据传输速率与信息传输速率**的区别。因为是二元对称信道，故 $C = 1 - h(p)$。由 $p = 0.02$ 得 $C = 1 - h(0.02) = 1 + 0.02 \times \log 0.02 = 0.887$ bits，它表示每个信道字符即每个比特实际上只能传送 0.887bits 的信息量，信息传输速率 $C_t = 1500 \times 0.887 = 1330.5\text{bits}/$ 秒。但是消息序列的字符服从均匀分布，每个字符包含 1bit 信息量，总共包含 14000bits 信息量，需无失真传输 14000bits，所需时间 $t = 14000/1330.5 = 10.5s > 10$ 秒，故传不完。

信道容量的最优化问题：

将信道容量改成一个最优化问题

$$C = \max_{p \in \mathcal{P}_{in}} f(p, Q), \tag{4.1}$$

$$\sum_{i=1}^{N} p_i = 1, \tag{4.2}$$

$$p_i \geq 0, i = 1, 2, \cdots, N, \tag{4.3}$$

它的可行域为 $\mathcal{P}_{in}$，注意到它不包含分量全为 0 的原点 $p = (0, 0, \cdots, 0)^T$，因此可行点 p 的分量不会全为 0，根据 Lagrange 乘数法可得关于信道容量充要条件。

引理 3.3.2:

如果多元函数 $y=f(p), p=(p_1,p_2,\cdots,p_N)$ *是可行域* $\mathcal{P}_{in}$ *上的连续可微的凹函数，则函数* $f(p)$ *在* $\mathcal{P}_{in}$ *上最大值点为* $p* = (p_1^*, p_2^*, \cdots, p_N^*)$ *的充要条件是存在一个常数* C *使得*

$$\begin{cases} \left.\dfrac{\partial f}{\partial p_i}\right|_{p^*} = C & \text{若} i \in \{i|p_i^* > 0\} \\ \left.\dfrac{\partial f}{\partial p_i}\right|_{p^*} \leq C & \text{若} i \in \{i|p_i^* = 0\} \end{cases} . \tag{4.4}$$

3.5

定理 3.3.2:

离散无记忆信道输入分布 $p^* = (p_1^*, p_2^*, \cdots, p_N^*)$ 为最大输入分布的充要条件是存在常数 C 使

$$\begin{cases} D(Q^{(i)}||q^*) = C & 若 i \in \{i|p_i^* > 0\} \\ D(Q^{(i)}||q^*) \le C & 若 i \in \{i|p_i^* = 0\} \end{cases} \tag{4.5}$$

3.8

并且 C 就是信道容量。其中

（1） $Q^{(i)} = (p_{1|i}, p_{2|i}, \cdots, p_{M|i})$ 为信道矩阵第 i 行的条件分布。

（2） q^* 是最大输出分布。

证明：

根据信道容量所对应的最优化问题（4.1），由定理 **??** 函数 $f(p,Q)$ 是输入分布的凹函数，约束条件（4.2）、（4.3）所构成的可行域 $\mathcal{P}_{in}$ 是闭凸集，故 $f(p,Q)$ 存在最大值。于是由引理 4.1，最大值点为 $p^*=(p_1^*,p_2^*,\cdots,p_N^*)$ 的充要条件为存在常数 C_1 使得

$$\begin{cases} \left.\dfrac{\partial f}{\partial p_i}\right|_{p^*} = C_1 & \text{若} \quad i\in\{i|p_i^*>0\} \\ \left.\dfrac{\partial f}{\partial p_i}\right|_{p^*} \le C_1 & \text{若} \quad i\in\{i|p_i^*=0\} \end{cases} . \tag{4.6}$$

39

续证明：

为了方便求导，将 $f(p,Q)$ 改写成

$$
\begin{aligned}
f(p,Q) &= \sum_{i=1}^{N} p_i \sum_{j=1}^{M} p_(j|i) \left(\log p_(j|i) - \log q_j\right), q_j = \sum_{i=1}^{N} p_i p_(j|i), \\
&= \sum_{i=1}^{N} p_i \sum_{j=1}^{M} p_(j|i) \log p_(j|i) - \sum_{i=1}^{N} p_i \sum_{j=1}^{M} p_(j|i) \log q_j .
\end{aligned}
$$

续证明：

现在求函数 $f(p,Q)$ 的偏导数

$$
\begin{aligned}
\frac{\partial f}{\partial p_k} &= \sum_{j=1}^{M} p_{j|k}\log p_{j|k} - \sum_{j=1}^{M} p_{j|k}\log p_j - \left[\sum_{i=1}^{N} p_i \sum_{j=1}^{M} p_(j|i)\frac{p_{j|k}}{q_j}\right]\log e \\
&= \sum_{j=1}^{M} p_{j|k}\log\frac{p_{j|k}}{q_j} - \left[\sum_{j=1}^{M}\frac{p_{j|k}}{q_j}\sum_{i=1}^{N} p_i p_(j|i)\right]\log e \\
&= D(Q^{(k)}||q) - \left[\sum_{j=1}^{M}\frac{p_{j|k}}{q_j}q_j\right]\log e \\
&= D(Q^{(k)}||q) - \left[\sum_{j=1}^{M} p_{j|k}\right]\log e \\
&= D(Q^{(k)}||q) - \log e,
\end{aligned}
$$

续证明：

从而充要条件（4.6）等价于条件（4.5），其中 $C=C_1+\log e$。现在证明 C 就是信道容量，事实上：

$$
\begin{aligned}
I(X;Y) &= \sum_x p^*(x)\sum_y p(y|x)\log\frac{p(y|x)}{q^*(y)} \\
&= \sum_{i=1}^{N} p_i^* D(Q^{(i)}||q^*) \\
&= \sum_{p_i^*>0} p_i^* D(Q^{(i)}||q^*) = \sum_{p_i^*>0} p_i^* C = C.
\end{aligned}
$$

命题 3.3.2：离散准对称信道容量公式

离散准对称信道最大输入分布为等可能分布，并且信道容量公式为

$$C = H(Y) - h(Q^{(i)}),$$

其中

（1） $H(Y)$ *为对最大输出分布的熵。*

（2） $h(Q^{(i)})$ *为信道矩阵的任意一行所构成的条件分布的熵。*

证明：

当输入分布为等概率分布 $p_X(x)=1/N$ 时，相应的输出分布为

$$p_Y(y)=\sum_x p(x)p(y|x)=\sum_{i=1}^N \frac{p(y|a_i)}{N}=\frac{1}{N}\sum_{i=1}^N p(y|a_i),$$

根据准对称信道定义得

（1）在每个小块矩阵 Q_k 中不同列仅仅元素排列不同，故当 $y\in\mathcal{Y}_k$ 时 $\sum_{i=1}^N p(y|a_i)$ 是只与 k 有关的常数，从而当 $y\in\mathcal{Y}_k$ 时 $p_Y(y)$ 也是仅与 k 有关的常数，记为 r_k。

（2）在每个小块矩阵 Q_k 中不同行仅仅元素排列不同，故和 $\sum_{y\in\mathcal{Y}_k} p(y|a_i)$ 是与行数无关只与 k 有关的常数, 记为 s_k。

（3）与第二条一样，在每个小块矩阵 Q_k 中和 $\sum_{y\in\mathcal{Y}_k} p(y|a_i)\log p(y|a_i)$ 也是与行数无关只与 k 有关的常数, 记为 t_k。

续证明：

从而 $D(Q^{(i)}||p_Y)$

$$
\begin{aligned}
&= \sum_{y\in\mathcal{Y}} p(y|a_i)\log\frac{p(y|a_i)}{p_Y(y)} = \sum_{k=1}^{m}\sum_{y\in\mathcal{Y}_k} p(y|a_i)\log\frac{p(y|a_i)}{p_Y(y)} \\
&= \sum_{k=1}^{m}\sum_{y\in\mathcal{Y}_k} p(y|a_i)\log\frac{p(y|a_i)}{r_k} \\
&= \sum_{k=1}^{m}\left[\sum_{y\in\mathcal{Y}_k} p(y|a_i)\log p(y|a_i) + \sum_{y\in\mathcal{Y}_k} p(y|a_i)\log\frac{1}{r_k}\right] \\
&= \sum_{k=1}^{m}\left[t_k + \log\frac{1}{r_k}\sum_{y\in\mathcal{Y}_k} p(y|a_i)\right] \\
&= \sum_{k=1}^{m}\left[t_k + s_k\log\frac{1}{r_k}\right] = C\text{常数},
\end{aligned}
$$

续证明:

意即对任何 $i=1,2,\cdots,N$ 都有 $D(Q^{(i)}||q)=C$。由信道容量的充要条件（定理 3.3.2）得准对称信道最大输入分布为等概率分布，这时

$$\begin{aligned} H(Y|X) &= \sum_x p(x)\sum_y p(y|x)\log\frac{1}{p(y|x)}, \\ &= \sum_x \frac{1}{N}h(Q^{(i)}) = h(Q^{(i)}), \\ C &= I(X;Y) = H(Y)-H(Y|X) = H(Y)-h(Q^{(i)}). \end{aligned}$$

例题 3.3.13：

求下面信道的容量。

$$Q = \begin{pmatrix} 1/3 & 1/3 & 1/6 & 1/6 \\ 1/6 & 1/3 & 1/6 & 1/3 \end{pmatrix}.$$

解:

这是一个准对称信道，最大输入分布为等可能分布或等概率分布即 $X \sim p_X(x) = 0.5, x = 0, 1$，故最大输出分布为

$$p_Y(0) = 1/4 \quad , \quad p_Y(1) = 1/3 \quad , \quad p_Y(2) = 1/6 \quad , \quad p_Y(3) = 1/4,$$

从而

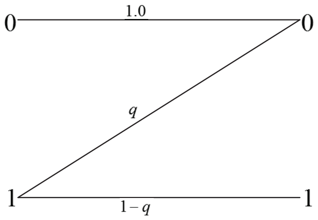
$$H(Y) = \frac{1}{4}\ln 4 + \frac{1}{3}\ln 3 + \frac{1}{6}\ln 6 + \frac{1}{4}\ln 4 = 1.358 \text{ nats},$$

$$h(Q^{(i)}) = \frac{1}{3}\ln 3 + \frac{1}{3}\ln 3 + \frac{1}{6}\ln 6 + \frac{1}{6}\ln 6 = 1.3297 \text{ nats},$$

信道容量为 $C = H(Y) - h(Q^{(i)}) = 0.0283$ nats。

例题 3.3.14：

求 **Z** 信道的容量。



解：它的信道矩阵为

$$Q = \begin{pmatrix} 1 & 0 \\ q & 1-q \end{pmatrix},$$

没有任何对称性，可以按定义求容量，也可以按信道容量的充要条件（4.5）求容量。

解：

下面给出按定义求容量的方法。设输入分布为
$X \sim p_X(x) = (1-p, p)$，则输出分布为

$$Y \sim p_Y(y) = p_X Q = (1-p(1-q), p(1-q)) \,,$$

从而

$$\begin{aligned}
H(Y) &= h\left(p(1-q)\right), \\
H(Y|X) &= \sum_x p_X(x) H(Y|X=x) = pH(Y|X=1) = ph(q), \\
I(X;Y) &= H(Y) - H(Y|X) = h\left(p(1-q)\right) - ph(q) = f(p).
\end{aligned}$$

续解:

为求信道容量，只须求函数 $f(p)$ 的最大值。先求驻点，再求最值。

$$f'(p) = h'(p(1-q))(1-q) - h(q) = (1-q)\ln\frac{1-p(1-q)}{p(1-q)} - h(q) = 0,$$

由此解出最大值点

$$p^* = \frac{\frac{1}{1-q}}{1+e^{\frac{h(q)}{1-q}}},$$

信道容量为

$$C = h(p^*(1-q) - p^*h(q).$$