

第3章

信道与容量

广义地讲,信道是信息传输或存储的通道,是通信系统最重要的组成部分。我们研究信道的目的在于研究信道中能传输的最大信息量,即信道容量。

在一个典型的通信系统中,信源发出携带着一定信息量的消息,并转换成适于在信道中传输的信号,然后通过信道传送到收端。信道在传送信号的同时,会引入各种干扰和随机噪声,使得信号产生失真,从而导致接收错误。由于存在干扰和噪声,因此信道的输入与输出信号之间是一种统计依赖关系,而不再是确定的函数关系。只要知道了信道的输入信号、输出信号的特性,以及它们之间的统计依赖关系,就可以确定信道的全部特性。

实际的通信系统有很多种,如卫星通信系统、公用电话网、微波通信系统、光纤通信系统等,相应的信道形态也是多种多样的,它是一种物理传输媒介,在空间或时间上把信息从一端传输到另一端。电缆、光纤、无线电波、声波、磁带、光盘都可以构成信道,信息论的研究将不细致考虑这些具体介质的物理现象、特性和规律,而是从信息传输的角度对其进行抽象,建立与各种通信系统相适应的数学模型,研究信息经过这些模型的普遍规律,指导通信系统的优化设计。一般性而言,信道可以看成是一个概率转换器,我们不关心这个转换器内部的过程,仅考虑其输入、输出及其相互关系。因此,根据输入和输出信号的形式、信道的统计特性及信道用户数等方面来对信道进行分类。

根据输入、输出信号的时间特性和取值特性分类:

(1) 离散信道:输入、输出信号在取值和时间上均为离散的信道,该信道上的信号可以离散型随机变量或随机矢量描述。

(2) 连续信道:输入、输出信号取值均为连续的信道,该信道上的信号可以连续型随机变量或随机矢量描述。

(3) 半离散信道(或半连续信道):输入与输出中一个为离散型随机变量而另一个为连续型随机变量的信道。

(4) 波形信道:输入和输出是时间上连续的随机信号 $\{x(t)\}$ 与 $\{y(t)\}$,即信道输入和输出随机变量取值均为连续,且随时间连续变化,因此可用随机过程来描述。

根据信道的统计特性分类:

(1) 恒参信道:信道的统计特性不随时间而变化,如卫星信道一般可视作恒参信道。

(2) 随参信道:信道的统计特性随时间而变化,如短波信道即是一种典型的随参信道。

根据信道用户的数量分类:

(1) 单用户信道:只有一个输入端和一个输出端的单向通信的信道。

(2) 多用户信道:在输入端或输出端中至少有一端有两个以上的用户,并且可以双向通信的信道。目前实际的通信信道绝大多数是多端信道。多端信道又可分为多元接入信道与广播信道。

根据信道的记忆特性分类:

(1) 无记忆信道:输出仅与当前的输入有关,而与过去的输入和输出无关。

(2) 有记忆信道:输出不仅与当前输入有关,而且与过去的输入和输出有关。



本章首先讨论信道的分类及信道的数学模型,然后按照离散信道、连续信道、波形信道的次序定量研究信道传输的平均互信息及其重要性质,导出信道容量概念及计算方法,进一步推导并分析香农公式。

3.1 信道模型

信道的数学模型要描述清楚三个要素:信道的输入/输出统计特性以及输入与输出之间的转换关系。考虑到信息传输的随机性,可用概率函数描述各要素的特性。离散信道的数学模型如图 3.1.1 所示。

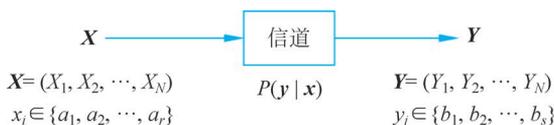


图 3.1.1 离散信道的数学模型

考虑到一般性,图中输入与输出信号均用随机矢量表示,输入 $\mathbf{X} = (X_1, X_2, \dots, X_N)$, 输出 $\mathbf{Y} = (Y_1, Y_2, \dots, Y_N)$, \mathbf{X} 或 \mathbf{Y} 中每个随机变量 X_i 或 Y_j 的取值 x_i 与 y_j 分别取自于输入、输出符号集 $A = (a_1, a_2, \dots, a_r)$, $B = (b_1, b_2, \dots, b_s)$ 。由于噪声或干扰的存在,信道输入与输出之间没有确定的函数关系,而是一种统计依赖关系,用条件概率 $P(\mathbf{y} | \mathbf{x})$ 反映,信道噪声与干扰的影响也包含在其中。同时,离散信道的数学模型可表示为

$$\{\mathbf{X}, P(\mathbf{y} | \mathbf{x}), \mathbf{Y}\}$$

根据信道的统计特性(即条件概率)的不同,离散信道又可分成如下三种情况。

(1) 无干扰(无噪)信道:信道中没有随机性的干扰,输出信号 \mathbf{Y} 与输入信号 \mathbf{X} 之间有确定的对应关系,即 $\mathbf{y} = f(\mathbf{x})$,故条件概率 $P(\mathbf{y} | \mathbf{x})$ 满足

$$P(\mathbf{y} | \mathbf{x}) = \begin{cases} 1, & \mathbf{y} = f(\mathbf{x}) \\ 0, & \mathbf{y} \neq f(\mathbf{x}) \end{cases} \quad (3.1.1)$$

(2) 有干扰无记忆信道:这种信道存在干扰,为实际中常见的信道类型,其输出符号与输入符号之间不存在确定的对应关系,且其条件概率不再具有式(3.1.1)的形式,但信道任一时刻的输出符号仅依赖同一时刻的输入符号,是**无记忆信道**。利用概率关系转换的方法可以证明无记忆信道的条件概率满足

$$P(\mathbf{y} | \mathbf{x}) = P(y_1 y_2 \dots y_N | x_1 x_2 \dots x_N) = \prod_{i=1}^N P(y_i | x_i) \quad (3.1.2)$$

(3) 有干扰有记忆信道:这是更一般的情况,这种信道某一时刻的输出不仅与当时的输入有关,还与其他时刻的输入及输出有关,其条件概率不再满足式(3.1.2),对它的分析也更复杂。

如果信道的输入与输出均是单个符号,那么可以用随机变量 X 和 Y 分别表示而不是矢量形式,这称为单符号信道。其中输入随机变量 X 的字符集为 $\{a_1, a_2, \dots, a_r\}$, 输出随机变量 Y 的字符集为 $\{b_1, b_2, \dots, b_s\}$, 信道输入与输出之间的转换关系可用条件概

率 $P(y=b_j | x=a_i)$ 描述, 该条件概率称为信道转移概率或传递概率, 则该信道的统计特性可由下面的矩阵描述:

$$\mathbf{P} = \begin{bmatrix} P(b_1 | a_1) & P(b_2 | a_1) & \cdots & P(b_s | a_1) \\ P(b_1 | a_2) & P(b_2 | a_2) & \cdots & P(b_s | a_2) \\ \vdots & \vdots & \ddots & \vdots \\ P(b_1 | a_r) & P(b_2 | a_r) & \cdots & P(b_s | a_r) \end{bmatrix}_{r \times s}$$

该矩阵称为**信道矩阵(或信道转移矩阵)**, 该矩阵每一行元素之和等于 1, 即

$$\sum_{j=1}^s P(b_j | a_i) = 1$$

其物理意义是发出输入符号 $X=a_i$ 后收到的一定是输出字符集中的某一个。

下面通过例子介绍两种重要的单符号离散信道。

例 3.1.1 二元对称信道(Binary Symmetric Channel, BSC)。

这是很重要的一种信道, 其输入与输出符号均取值于 $\{0, 1\}$ 。记 $a_1 = b_1 = 0, a_2 = b_2 = 1$, 则转移概率为

$$P(b_1 | a_1) = 1 - p$$

$$P(b_2 | a_2) = 1 - p$$

$$P(b_1 | a_2) = p$$

$$P(b_2 | a_1) = p$$

可得 BSC 的信道转移矩阵为

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

二元对称信道也可用图 3.1.2 表示。

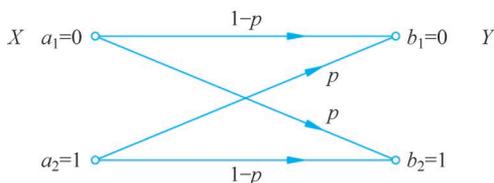


图 3.1.2 二元对称信道

不难得出, 该信道的误码率为

$$P_e = P(a_2)P(b_1 | a_2) + P(a_1)P(b_2 | a_1) = p$$

若采用纠错编码技术, 则将可有效地降低信道误码率。

例 3.1.2 二元删除信道(Binary Eraser Channel, BEC)。

它的输入 X 取值于 $\{0, 1\}$, 输出 Y 取值于 $\{0, 1, 2\}$, 且信道转移矩阵为

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

二元删除信道也可如图 3.1.3 表示。

这种信道在下述情况下存在：当信号波形传输中失真较大时，在接收端不是对接收信号硬性判为0或1，而是根据最佳接收机额外给出的信道失真信息增加一个中间状态2(称为删除符号)，采用特定的纠删编码，可有效地恢复出这个中间状态的正确取值。

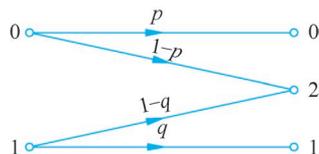


图 3.1.3 二元删除信道

3.2 平均互信息

3.2.1 信道疑义度

以图 3.2.1 所示的单符号离散无记忆信道(Discrete Memoryless Channel, DMC)为基本信道。

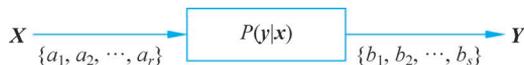


图 3.2.1 基本信道

信道输入即信源 X 的熵为

$$H(X) = - \sum_{i=1}^q P(a_i) \log P(a_i) = - \sum_X P(x) \log P(x)$$

它表示了信源 X 中各符号的平均不确定性，即输入的先验不确定性，故可称为先验熵。

信源 X 发出一个符号 a_i 后，通过信道到达收方，设 Y 接收到的符号为 b_j 。若信道中没有干扰，则信道的输入与输出符号一一对应，信源发出的信息量全部被接收端收到，也即关于信道输入的先验不确定性全部消除。

当信道中存在干扰时，信道输入、输出符号间的关系是统计依赖关系，收方收到 $y = b_j$ 后关于信道输入 $x = a_i$ 的后验概率为 $P(a_i | b_j)$ ，此时关于输入 a_i 的不确定性应为 $-\log P(a_i | b_j)$ ，将 $-\log P(a_i | b_j)$ 对所有输入符号求统计平均，可得

$$H(X | b_j) = - \sum_{i=1}^r P(a_i | b_j) \log P(a_i | b_j) \triangleq - \sum_X P(x | b_j) \log P(x | b_j)$$

它表示接收到输出符号为 b_j 后关于输入 X 的后验熵，以及收到 b_j 后关于各输入符号的平均不确定性。

再将 $H(X | b_j)$ 对所有输出符号求统计平均，可得

$$\begin{aligned} H(X | Y) &= \sum_{j=1}^s P(b_j) H(X | b_j) \\ &= - \sum_{i=1}^r \sum_{j=1}^s P(b_j) P(a_i | b_j) \log P(a_i | b_j) \\ &= - \sum_{i=1}^r \sum_{j=1}^s P(a_i b_j) \log P(a_i | b_j) \\ &\triangleq - \sum_X \sum_Y P(xy) \log P(x | y) \end{aligned}$$



中文教学
录像



视频



双语教学
录像

定义 3.2.1 设离散信道的输入、输出分别为 X 和 Y , 信道后验概率为 $P(a_i | b_j)$ ($i = 1, 2, \dots, r; j = 1, 2, \dots, s$), 定义条件熵 $H(X|Y)$ 为该信道的**信道疑义度**。

信道疑义度 $H(X|Y)$ 反映了收到输出 Y 的全部符号后关于发出 X 各符号的平均不确定性, 这个对 X 尚存的不确定性是信道干扰引起的。

若信道输入、输出是一一对应的, 则接收到输出 Y 后对 X 的不确定性将完全消除, 即有 $H(X|Y) = 0$ 。干扰情况下一般 $H(X|Y) > 0$, 说明信源符号经过有干扰信道传输后总要残留一部分不确定性, 这是由于传输过程中部分信息损失在信道中, 因此信道疑义度又称为**损失熵**。由熵的性质可知, 条件熵是不大于无条件熵的, 即 $H(X|Y) \leq H(X)$ 总成立, 这说明接收端收到信道传送来的符号 Y 后得到了一些关于 X 的信息量, 使得传输过程结束后总能消除一些关于输入 X 的不确定性, 从而获得了一些信息(只有在 Y 与 X 独立时, 才有 $H(X|Y) = H(X)$), 为此引出了收发双方的平均互信息概念。

3.2.2 平均互信息及其性质

根据上述讨论可知, $H(X)$ 代表接收到输出符号以前关于信源 X 的先验不确定性, 而 $H(X|Y)$ 代表接收到输出符号后残存的关于 X 的不确定性, 两者之差即为传输过程获得的信息量。

定义 3.2.2 令

$$I(X; Y) = H(X) - H(X|Y) \quad (3.2.1)$$

定义 $I(X; Y)$ 为信道输入 X 与输出 Y 之间的**平均互信息**。

平均互信息 $I(X; Y)$ 代表了接收到每个输出符号 Y 后获得的关于 X 的平均信息量, 单位为比特/符号。

根据 $I(X; Y)$ 的定义式可得

$$\begin{aligned} I(X; Y) &= \sum_X P(x) \log \frac{1}{P(x)} - \sum_{XY} P(xy) \log \frac{1}{P(x|y)} \\ &= \sum_{XY} P(xy) \log \frac{1}{P(x)} - \sum_{XY} P(xy) \log \frac{1}{P(x|y)} \\ &= \sum_{XY} P(xy) \log \frac{p(x|y)}{P(x)} \end{aligned} \quad (3.2.2)$$

$$I(X; Y) = \sum_{XY} P(xy) \log \frac{P(xy)}{P(x)p(y)} \quad (3.2.3)$$

$$I(X; Y) = \sum_{XY} P(xy) \log \frac{P(y|x)}{P(y)} \quad (3.2.4)$$

接收端收到某消息 y 后, 关于输入为 x 的信息量记为 x 与 y 之间的互信息 $I(x; y)$, 且

$$I(x; y) = \log \frac{1}{P(x)} - \log \frac{1}{P(x|y)} = \log \frac{P(x|y)}{P(x)} \quad (3.2.5)$$

将式(3.2.5)与式(3.2.2)对比可知, 对互信息 $I(x; y)$ 求统计平均后正是平均互信息 $I(X; Y)$, 两者分别代表了互信息的局部和整体含义, 在本质上是统一的。



视频



视频



双语教学
录像

互信息 $I(x; y)$ 的取值可能为正,也可能为负。但是,后面将看到平均互信息 $I(X; Y)$ 的值不可能为负。

从平均互信息 $I(X; Y)$ 的定义中可以进一步理解熵只是对不确定性的描述,而不确定性的消除才是接收端所获得的信息量。因此,平均互信息又称为信道的**信息传输率**,记为 R 。

平均互信息具有如下重要性质。

1. 非负性

即 $I(X; Y) \geq 0$ 。

当 X 与 Y 统计独立时等号成立。

由

$$H(X) \geq H(X | Y)$$

可得

$$I(X; Y) = H(X) - H(X | Y) \geq 0$$

而当 X 与 Y 统计独立时,有 $P(xy) = P(x)P(y)$,由式(3.2.2)可得

$$\begin{aligned} I(X; Y) &= \sum_{XY} P(xy) \log \frac{P(xy)}{P(x)P(y)} \\ &= \sum_{XY} P(x)P(y) \log 1 = 0 \end{aligned}$$

可见,平均互信息不会取负值,且一般情况下总大于0,仅当 X 与 Y 统计独立时才等于0。这个性质印证了通信系统建立的意义:通过一个信道传输总体上讲是不会被误读的(平均获得信息量不为负值),一般总能获得一些信息量,只有在 X 与 Y 统计独立的极端情况下才接收不到任何信息。

2. 极值性

即 $I(X; Y) \leq H(X)$ 。

根据熵的非负性,易得出信道疑义度也是非负的,即

$$H(X | Y) \geq 0$$

所以

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

这一性质的直观含义为接收者通过信道获得的信息量不可能超过信源本身提供的信息量。只有当信道为无损信道,即信道疑义度 $H(X | Y) = 0$ 时,才能获得信源的全部信息量。

综合性质(1)和性质(2)可得

$$0 \leq I(X; Y) \leq H(X)$$

当信道输入 X 与输出 Y 统计独立时,上式左边的等号成立;而当信道为无损信道时,上式右边的等号成立。

3. 对称性

即 $I(X; Y) = I(Y; X)$ 。

由于

$$P(xy) = P(yx)$$

故

$$\begin{aligned} I(X; Y) &= \sum_{XY} P(xy) \log \frac{P(xy)}{P(x)P(y)} \\ &= \sum_{XY} P(xy) \log \frac{P(yx)}{P(y)P(x)} \\ &= I(Y; X) \end{aligned}$$

$I(X; Y)$ 表示接收到 Y 后获得的关于 X 的信息量, $I(Y; X)$ 为发出 X 后得到的关于 Y 的信息量, 这二者是相等的, 当 X 与 Y 统计独立时, 有

$$I(X; Y) = I(Y; X) = 0$$

该式表明, 此时不可能由一个随机变量获得关于另一个随机变量的信息(全损)。而当输入 X 与输出 Y 一一对应时, 则有

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

即从一个随机变量可获得另一个随机变量的全部信息(无损)。

4. $I(X; Y)$ 与各类熵的关系

由平均互信息的计算表达式

$$\begin{aligned} I(X; Y) &= \sum_{XY} P(xy) \log \frac{P(x|y)}{P(x)} \\ &= \sum_{XY} P(xy) \log \frac{P(y|x)}{P(y)} \\ &= \sum_{XY} P(xy) \log \frac{P(xy)}{P(x)P(y)} \end{aligned}$$

可得

$$I(X; Y) = H(X) - H(X|Y) \quad (3.2.6)$$

$$I(X; Y) = H(Y) - H(Y|X) \quad (3.2.7)$$

$$I(X; Y) = H(X) + H(Y) - H(XY) \quad (3.2.8)$$

式中: $H(X|Y)$ 为信道疑义度, 表示符号通过有噪信道传输后平均损失的信息量; $H(Y|X)$ 为噪声熵(或散布度), 表示在已知 X 的条件下, 对于随机变量 Y 存在的平均不确定性。噪声熵是信道中的噪声引起, 反映了信道中噪声源的不确定性, 会影响信宿对信源信息的获取和识别; $H(XY)$ 为输入 X 与输出 Y 的联合熵。

由式(3.2.8)可得

$$\begin{aligned} H(XY) &= H(X) + H(Y) - I(X; Y) \\ &= H(X) + H(Y|X) \end{aligned} \quad (3.2.9)$$

$$H(XY) = H(Y) + H(X|Y) \quad (3.2.10)$$

由式(3.2.6)和式(3.2.7)可得

$$H(X|Y) = H(X) - I(X; Y) \quad (3.2.11)$$

$$H(Y | X) = H(Y) - I(X; Y) \quad (3.2.12)$$

至此得到了平均互信息 $I(X; Y)$ 与信源熵 $H(X)$ 、信宿熵 $H(Y)$ 、联合熵 $H(XY)$ 、信道疑义度 $H(X|Y)$ 及信道噪声熵 $H(Y|X)$ 等的一系列关系式, 它们之间的相互关系可用图 3.2.2 表示。

图 3.2.2 中圆 $H(X)$ 减去其左边阴影部分 $H(X|Y)$, 即得到中间部分 $I(X; Y)$, 这正是式(3.2.6)所表达的结果。以此类推, 式(3.2.6)~式(3.2.12)所描述的所有关系都可以通过图 3.2.2 得到形象的解释。

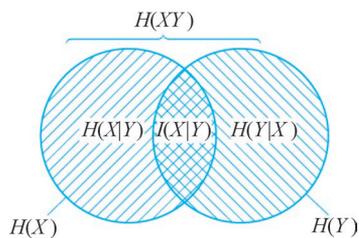


图 3.2.2 $I(X; Y)$ 与各类熵的关系图

5. $I(X; Y)$ 的凸函数特性

由 $I(X; Y)$ 的计算表达式

$$I(X; Y) = \sum_{XY} P(xy) \log \frac{P(y | x)}{P(y)}$$

又

$$P(xy) = P(x)P(y | x)$$

及

$$P(y) = \sum_X P(xy) = \sum_X P(x)P(y | x)$$

所以

$$I(X; Y) = \sum_{XY} P(x)P(y | x) \log \frac{P(y | x)}{\sum_X P(x)P(y | x)} \\ \triangleq f[P(x), P(y | x)] \quad (3.2.13)$$

即 $I(X; Y)$ 完全是信源 X 的概率分布 $P(x)$ 及信道转移概率 $P(y | x)$ 的函数。进一步分析可知, $I(X; Y)$ 与两个自变量 $P(x)$ 与 $P(y | x)$ 分别是 \cap 型凸函数与 \cup 型凸函数的关系。为避免烦琐的证明过程, 可以用一个例子加以说明。

例 3.2.1 设信源 X 的概率空间为

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \omega & 1 - \omega \triangleq \bar{\omega} \end{bmatrix}$$

BSC 信道转移矩阵为

$$\mathbf{P} = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$

而

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

首先分析上式中的 $H(Y | X)$ 项:

$$H(Y | X) = \sum_X P(x) \sum_Y P(y | x) \log \frac{1}{p(y | x)}$$

$$= \sum_x P(x) \left[p \log \frac{1}{p} + \bar{p} \log \frac{1}{\bar{p}} \right]$$

由于上式中 $\left[p \log \frac{1}{p} + \bar{p} \log \frac{1}{\bar{p}} \right]$ 与 $P(x)$ 无关, 因此可将 $\sum_x P(x)$ 单独列出, 又有

$$\sum_x P(x) = 1$$

则有

$$H(Y | X) = p \log \frac{1}{p} + \bar{p} \log \frac{1}{\bar{p}} \triangleq H(p)$$

其次分析 $H(Y)$ 项:

$$p(y=0) = \omega \bar{p} + \bar{\omega} p$$

$$p(y=1) = \omega p + \bar{\omega} \bar{p} = 1 - (\omega \bar{p} + \bar{\omega} p)$$

则有

$$\begin{aligned} H(Y) &= (\omega \bar{p} + \bar{\omega} p) \log \frac{1}{\omega \bar{p} + \bar{\omega} p} + [1 - (\omega \bar{p} + \bar{\omega} p)] \log \frac{1}{1 - (\omega \bar{p} + \bar{\omega} p)} \\ &\triangleq H(\omega \bar{p} + \bar{\omega} p) \end{aligned}$$

由上可得

$$\begin{aligned} I(X; Y) &= H(Y) - H(Y | X) \\ &= H(\omega \bar{p} + \bar{\omega} p) - H(p) \end{aligned} \quad (3.2.14)$$

$I(X; Y)$ 的凸函数特性如下。

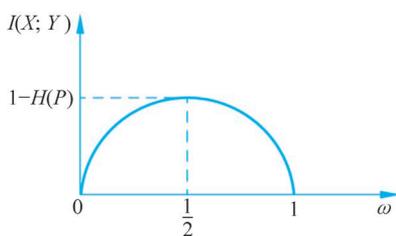


图 3.2.3 $I(X; Y)$ 与 ω 的关系曲线

(1) $I(X; Y)$ 是信源概率分布 $P(x)$ 的 \cap 型凸函数。

对于例 3.2.1, $I(X; Y)$ 与信源分布 ω 的关系曲线如图 3.2.3 所示。

可见, 此时 $I(X; Y)$ 与信源分布 $P(x)$ (例中为 ω) 确实呈 \cap 型凸函数关系。当信源等概分布, 即 $\omega = \bar{\omega} = \frac{1}{2}$ 时, 有 $I(X; Y) = H(1/2) - H(p) = 1 -$

$H(p)$, 达到极大值。此时在信道接收端平均每个符号获得最大的信息量。而当 $\omega = 0$ 或 1 时, 有 $I(X; Y) = 0$, 因为此时信源本身为确定的, 不提供任何信息量。

这一性质表明: 当信道固定时, 信源分布不同, 该信道接收端获得的平均信息量也不同。在例 3.2.1 中, 当两个信源符号等概分布时, $I(X; Y)$ 取极大值。后面将看到, $I(X; Y)$ 的这一凸函数特性是导出信道容量概念的依据。

(2) $I(X; Y)$ 是信道转移概率 $P(y|x)$ 的 \cup 型凸函数。

对于例 3.2.1, $I(X; Y)$ 与信道转移概率 $P(y|x)$ 的关系曲线如图 3.2.4 所示(例中 $P(y|x)$ 以 p 表示)。由式(3.2.14)可知, 当 $p = 0$ 时, 有

$$I(X; Y) = H(\omega \bar{p} + \bar{\omega} p) - H(p)$$

$$I(X; Y) = H(\omega) - H(0) = H(\omega)$$

当 $p=1$ 时,有

$$I(X; Y) = H(\bar{\omega}) - H(1) = H(\bar{\omega}) = H(\omega)$$

此时, $I(X; Y)$ 取得极大值 $H(\omega)$ 。实际上这就是无损信道的情况。注意: 当 $p=1$ 时, 信道的输出正好完全相反。0 变为 1, 1 变为 0。此时只要将接收译码规则颠倒, 仍能无损失地接收到全部信息量, 故 $I(X; Y)$ 仍然能取极大值 $H(\omega)$ 。

当 $p=1/2$ 时, 有

$$I(X; Y) = H\left(\frac{1}{2}\right) - H\left(\frac{1}{2}\right) = 0$$

此时 $I(X; Y)$ 取极小值 0, 这说明对任何一种信源, 总存在最差的信道, 使 $I(X; Y) = 0$ 。

实际上, $I(X; Y)$ 这一凸函数特性也是导出率失真函数概念的重要依据, 在此不再赘述。

以上在讨论平均互信息概念和性质时是基于单符号信道传输情况的。在实际中传输或存储的符号往往是一串离散符号为一组表示的, 即以序列形式传输, 下面讨论此种情况的平均互信息及相关性质。

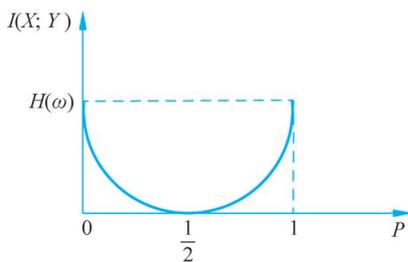


图 3.2.4 $I(X; Y)$ 与信道转移概率 $P(y|x)$ 的关系曲线

3.2.3 序列形式的平均互信息

前几节讨论了最简单的离散信道, 即信道的输入和输出都只是单个随机变量的信道。然而一般离散信道的输入和输出往往是一系列时间(或空间)离散的随机变量, 即为随机序列。该信道输入是 N 个符号 X_1, X_2, \dots, X_N 构成的随机序列, 输出端对应的也是以 N 维随机序列 $\mathbf{Y} = (Y_1, Y_2, \dots, Y_N)$ 表示, 信道的转移概率将以 N 维联合转移概率 $P(\mathbf{Y} | \mathbf{X}) = P(Y_1 Y_2 \dots Y_N | X_1 X_2 \dots X_N)$ 描述。多符号信道的平均互信息与 3.2.2 节所定义的平均互信息一致, 只不过信道中传递的符号以 N 维随机序列表示, 具体为

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

若输入或输出随机序列中每个随机变量都取值于同一输入或输出的符号集, 则这种离散信道称为单符号离散信道(图 3.1.1)的 N 次扩展信道。其数学模型基本与单符号离散信道的数学模型相同, 可用 $[X^N, P(\mathbf{y} | \mathbf{x}), Y^N]$ 概率空间来描述, 不同的是各概率函数均是联合概率形式, 如图 3.2.5 所示。

$$\left. \begin{array}{ccc} X^N & & Y^N \\ \alpha_1 = (a_1 \cdots a_1) & & \beta_1 = (b_1 \cdots b_1) \\ \alpha_2 = (a_1 \cdots a_2) & \rightarrow P(\beta_k | \alpha_k) \rightarrow & \beta_2 = (b_1 \cdots b_2) \\ \vdots & & \vdots \\ \alpha_{r^N} = (a_r \cdots a_r) & & \beta_{s^N} = (b_s \cdots b_s) \end{array} \right\}$$

图 3.2.5 N 次扩展信道的数学模型

其信道矩阵为

$$\mathbf{\Pi} = [P(\beta_h | \alpha_k)] = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1s^N} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{r^N 1} & \pi_{r^N 2} & \cdots & \pi_{r^N s^N} \end{bmatrix}$$

式中

$$\pi_{kh} = P(\beta_h | \alpha_k)$$

$$\alpha_k = (a_{k_1}, a_{k_2}, \dots, a_{k_N}), \quad a_{k_i} \in \{a_1, a_2, \dots, a_r\}, \quad i = 1, 2, \dots, N$$

$$\beta_h = (b_{h_1}, b_{h_2}, \dots, b_{h_N}), \quad b_{h_i} \in \{b_1, b_2, \dots, b_s\}, \quad i = 1, 2, \dots, N$$

且满足

$$\sum_{h=1}^{s^N} \pi_{kh} = 1, \quad k = 1, 2, \dots, r^N$$

如果某个输出 Y_i ($i = 1, 2, \dots, N$) 的取值仅取决于该时刻的输入 X_i ($i = 1, 2, \dots, N$), 与其他时刻的输入无关, 则为无记忆扩展信道, 即 DMC 的 N 次扩展信道。此时, 输入随机序列与输出随机序列之间的转移概率等于对应时刻的随机变量的转移概率的乘积, 即

$$\pi_{kh} = P(\beta_h | \alpha_k) = \prod_{i=1}^N P(b_{h_i} | a_{k_i})$$

无记忆信道的 N 次扩展信道的平均互信息为

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}) &= I(\mathbf{X}^N; \mathbf{Y}^N) = H(\mathbf{X}^N) - H(\mathbf{X}^N | \mathbf{Y}^N) = H(\mathbf{Y}^N) - H(\mathbf{Y}^N | \mathbf{X}^N) \\ &= \sum_{\mathbf{x}^N} \sum_{\mathbf{y}^N} P(\alpha_k \beta_h) \log \frac{P(\alpha_k | \beta_h)}{P(\alpha_k)} = \sum_{\mathbf{x}^N} \sum_{\mathbf{y}^N} P(\alpha_k \beta_h) \log \frac{P(\alpha_k \beta_h)}{P(\alpha_k) P(\beta_h)} \\ &= \sum_{\mathbf{x}^N} \sum_{\mathbf{y}^N} P(\alpha_k \beta_h) \log \frac{P(\beta_h | \alpha_k)}{P(\beta_h)} \end{aligned}$$

3.2.2 节中讨论的平均互信息的性质同样适用于序列形式的平均互信息。

另外可以证明, 在一般离散信道中信道输入和输出两个离散随机序列之间的平均互信息 $I(\mathbf{X}; \mathbf{Y})$, 与两序列中对应的离散随机变量之间的平均互信息 $I(X_i; Y_i)$ 存在以下重要的结论。

定理 3.2.1 设离散信道的输入序列 $\mathbf{X} = (X_1 X_2 \cdots X_N)$ 通过信道传输, 接收到的随机序列为 $\mathbf{Y} = (Y_1 Y_2 \cdots Y_N)$, 而信道的转移概率 $P(\mathbf{y} | \mathbf{x}) = P(y_1 y_2 \cdots y_N | x_1 x_2 \cdots x_N)$, 其中 $X_i \in A = \{a_1, a_2, \dots, a_r\}$, $Y_i \in B = \{b_1, b_2, \dots, b_s\}$, 且有 $x \in X, y \in Y, x_i \in X_i, y_i \in Y_i$ ($i = 1, 2, \dots, N$)。

若信道是无记忆的, 则存在

$$I(\mathbf{X}; \mathbf{Y}) \leq \sum_{i=1}^N I(X_i; Y_i)$$

若信源是无记忆的, 则存在

$$I(\mathbf{X}; \mathbf{Y}) \geq \sum_{i=1}^N I(X_i; Y_i)$$

若信源和信道都是无记忆的,则存在

$$I(\mathbf{X}; \mathbf{Y}) = \sum_{i=1}^N I(X_i; Y_i)$$

(备注:证明留作习题。)

该定理其实也是序列形式的平均互信息凸函数特性的表现。当给定信道是无记忆时,平均互信息 $I(\mathbf{X}; \mathbf{Y})$ 也是信源概率分布的 \cap 型凸函数,其最大点位于信源各符号独立时;而当给定信源是无记忆时,平均互信息 $I(\mathbf{X}; \mathbf{Y})$ 将随着信道转移概率呈 \cup 型凸函数变化,其最小值点位于信道无记忆状态。

3.3 离散信道及其容量

3.3.1 信道容量的定义

研究信道的目的是获得尽可能高的信息传输率,即希望信道中平均每个符号所能传递的信息量尽可能大。

由前面的讨论可知,平均互信息 $I(X; Y)$ 代表了平均每个接收符号获得的关于输入 X 的信息量,因此它又称为信道的信息传输率,即

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

信息传输率的概念可扩展到更宽的范围。如信源的熵也可直接称为信源的信息传输率,当对信源进行某种变换(如信源编码或信道编码)后,得到新的信源的熵即为变换后的信息传输率(平均每个变换后的符号 Y 承载关于变换前符号 X 的信息量)。而信源信息经过信道传输后的信息传输率就是平均互信息 $I(X; Y)$ 。

$I(X; Y)$ 是信源分布 $P(x)$ 及信道转移概率 $P(y|x)$ 的函数,即

$$I(X; Y) \triangleq f[P(x), P(y|x)]$$

而且 $I(X; Y) \sim P(x)$ 为 \cap 型凸函数关系,当信源分布 $P(x)$ 变化到某一点时, $I(X; Y)$ 可达到最大值,即此时信道的信息传输率最大,这正是我们希望的。故对某一特定信道,当信源取某种最佳概率分布时, $I(X; Y)$ 能达到最大值。

定义 3.3.1 设某信道的平均互信息为 $I(X; Y)$, 其输入(即信源)的分布为 $P(x)$, 则定义

$$\max_{P(x)} \{I(X; Y)\} \triangleq C \quad (3.3.1)$$

为该信道的**信道容量**。

信道容量 C 即为信道的最大信息传输率(单位为比特/符号),此最大信息传输率是在信源取最佳分布时获得的。对于例 3.1.1 中的 BSC 信道,最佳信源分布为等概分布,而其信道容量为 $C=1-H(p)$,可见 C 只与信道转移概率 p 有关。

需要注意的是:一个给定信道的信道容量 C 是确定的,不随信源分布而变,信道容



中文教学
录像



视频



双语教学
录像

量 C 是信道本身固有的属性,信道矩阵给定了,信道容量就是一个确定值,其取值的大小直接反映了信道传信能力的高低。

另外,式(3.3.1)中 $I(X; Y)$ 的最大值通过调整信源 X 的概率分布 $P(x)$ 得到。由 $I(X; Y)$ 的凸函数特性知,并不是所有的信源概率分布都能使 $I(X; Y)$ 达到最大值,对于给定信道,只有在与之匹配的最佳输入分布 $P^*(x)$ 时,信道的信息传输率才能达到此最大值,即达到了信道容量。即使如此,信息传输过程中还会存在差错,这是由信道的转移概率决定的,是信息传输固有的现象。当传输错误超过可靠性容限时,采用信道编码的方法将特定的冗余码元加入信源符号中,以便在接收端纠正错误,从而提高传输可靠性。这样做势必降低信道的信息传输率,有效性与可靠性是信息传输的一对基本矛盾。而香农信息理论告诉我们,这个矛盾理论上是可以解决的。在后续章节介绍的有噪信道编码定理将指出,总存在最佳的信道编码,保证在信道的信息传输率不超过信道容量时能获得任意高的传输可靠性。

有时我们关心的是信道在单位时间内平均每个符号传输的信息量。若传输一个符号需要 t 秒,则信道平均每秒传输的信息量为

$$R_t = \frac{1}{t} I(X; Y)$$

R_t 称为信息传输速率,单位为 b/s。而该信道单位时间内传输的最大信息量为

$$C_t = \frac{1}{t} \max_{P(x)} \{I(X; Y)\}$$

一般仍称 C_t 为信道容量,但增加一个下标 t ,区别于 C 。

下面从单符号信道到序列形式传递的信道,从简单到复杂,讨论各种离散信道的容量计算。

3.3.2 简单离散信道的信道容量

计算信道容量是对信道研究的重要课题,对一般信道而言计算其信道容量并非易事,下面两节介绍几种特殊信道,根据其信道特点,可简化对平均互信息 $I(X; Y)$ 求极大值的过程。当离散信道的输入与输出之间为确定关系或简单的统计依赖关系时,可称其为简单离散信道,包括无噪无损信道、有噪无损信道、无噪有损信道。图 3.3.1 为这三类信道的具体例子。

1. 无噪无损信道

图 3.3.1(a) 所示信道的输入与输出符号之间存在确定的一一对应关系,其信道转移概率为

$$P(b_j | a_i) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \quad (i, j = 1, 2, 3)$$

它的信道转移矩阵为单位矩阵,即

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

这种信道疑义度 $H(X|Y)$ 必为 0, 所以平均互信息为

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

它表示接收端收到符号 Y 后平均获得的信息量等于信源发出每个符号所包含的平均信息量, 信道传输中没有信息损失。而且由于噪声熵也等于 0, 因此图 3.3.1(a) 所示的信道称为无噪无损信道, 其信道容量为

$$C = \max_{P(x)} \{I(X; Y)\} = \max_{P(x)} \{H(X)\} = \log r \text{ (比特 / 符号)}$$

2. 有噪无损信道

图 3.3.1(b) 所示的有噪无损信道的转移矩阵为

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{5} & \frac{3}{10} & \frac{1}{10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

每个输入符号通过信道后可能变成几种输出符号, 因此其噪声熵 $H(Y|X) \geq 0$ 。但各个输入符号所对应的输出符号不相重合, 这些输出符号可分成不相交的三个集合, 且这些集合与各输入符号一一对应。这就意味着, 接收到输出符号 Y 后, 对发送 X 的符号可以完全确定, 信道疑义度 $H(X|Y) = 0$, 故其平均互信息为

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

由于噪声熵 $H(Y|X) > 0$, 而 $I(X; Y) = H(Y) - H(Y|X)$, 故 $I(X; Y) < H(Y)$ 。因此, 有噪无损信道的信道容量为

$$C = \max_{P(x)} \{I(X; Y)\} = \max_{P(x)} \{H(X)\} = \log r \text{ (比特 / 符号)}$$

由上述两种信道的特点可以看出: 若信道转移矩阵中每一列有且仅有一个非零元素(每个输出符号对应着唯一的一个输入符号), 则该信道一定是无损信道。其信息传输率即等于信源熵, 信道容量等于 $\log r$ 。

3. 无噪有损信道

图 3.3.1(c) 所示的无噪有损信道的转移矩阵为

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

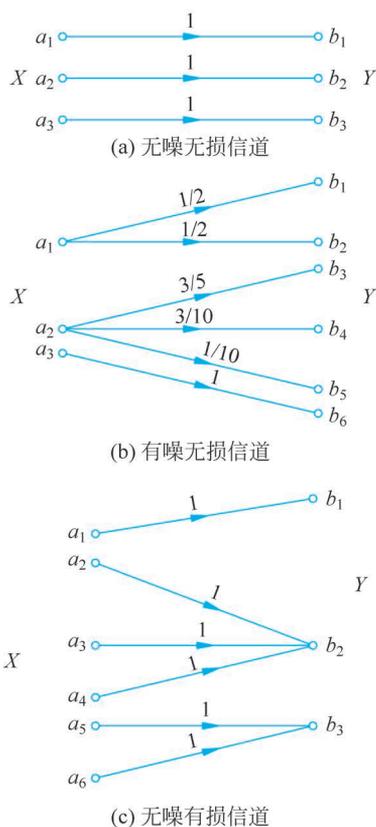


图 3.3.1 三类简单离散信道

每个输入符号都确定地转变成某一输入信号,因此其噪声熵 $H(Y|X)=0$,而接收到输出符号却不能确切地判断发出的是什么符号,因此信道疑义度 $H(Y|X)>0$,从而

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

设 Y 有 s 个符号,则 Y 等概分布时其熵 $H(Y)$ 最大。由图 3.3.1(c) 所示的信道转移关系容易证明存在最佳的输入分布使输出 Y 达到等概分布,故这种无噪有损信道的信道容量为

$$C = \max_{P(x)} \{I(X; Y)\} = \max_{P(x)} \{H(Y)\} = \log s \text{ (比特 / 符号)}$$

由上述分析可知,若信道转移矩阵中每一行有且仅有一个非零元素,则该信道一定是无噪信道。

至此分析了无损或无噪的简单离散信道及其信道容量的计算方法。更一般的离散信道既是有噪的又是有损的(后面统一称为有噪信道),其信道转移矩阵中至少有一行存在一个以上的非零元素,同时至少有一列存在一个以上的非零元素。这种情况下信道容量计算将十分复杂。下面讨论一类特殊的有噪有损信道——对称离散信道。

3.3.3 对称离散信道的信道容量

离散信道中有一类特殊的信道,其转移矩阵具有很强的对称性,即信道矩阵 \mathbf{P} 中每一行都是由同一集合 $\{p'_1, p'_2, \dots, p'_s\}$ 的诸元素排列而成,并且每一列也都是由同一集合 $\{q'_1, q'_2, \dots, q'_r\}$ 诸元素排列而成,具有这种对称性的信道称为**对称离散信道**。例如,信道矩阵

$$\mathbf{P}_1 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

与

$$\mathbf{P}_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$$

对应的信道都是对称离散信道。信道矩阵

$$\mathbf{P}_3 = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.1 & 0.7 \end{bmatrix}$$

对应的信道不是对称离散信道,因为它的第 1 列、第 2 列及第 3 列的构成元素都不完全相同。而且由于 $r=2, s=3$, 因此无论怎样调整信道矩阵 \mathbf{P}_3 的元素都不会满足对称性。

同样,信道矩阵



中文教学
录像



视频

$$\mathbf{P}_4 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} \end{bmatrix}$$

对应的信道也不是对称离散信道,因为它的第1列与第2列的构成元素都不完全相同。若将信道矩阵 \mathbf{P}_4 进行列置换,则可得到如下等效信道矩阵:

$$\mathbf{P}_4 = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \vdots & \frac{1}{3} & \vdots & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \vdots & \frac{1}{6} & \vdots & \frac{1}{3} \end{bmatrix}$$

该矩阵可按列划分成3个子矩阵,每个子矩阵均是对称矩阵,这类信道称为**准对称信道**。即如果把信道输出符号集 Y 划分成若干子集 $Y_h (h=1, 2, \dots, s)$, 每个子集 Y_h 所对应的信道矩阵中列所组成的子阵满足“每行元素是第一行的置换,每一列元素是第一列的置换”的对称矩阵条件。不难看出,对称信道既对输入对称也对输出对称,准对称信道仅对输入对称而对输出是局部对称,因此对称信道是准对称信道的特例。

下面分析对称离散信道的信道容量。

由

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

而

$$H(Y | X) = - \sum_X P(x) \sum_Y P(y | x) \log P(y | x)$$

由于信道的对称性,信道矩阵的每一列中的元素均取自同样的集合,因此上式中第二个和式 $\sum_Y P(y | x) \log P(y | x)$ 与 x 无关,仅与各信道转移概率值 p'_i 有关,并且其形式与信源熵计算公式相同,故

$$\begin{aligned} H(Y | X) &= \left[- \sum_Y P(y | x) \log P(y | x) \right] \sum_X P(x) \\ &= - \sum_Y P(y | x) \log P(y | x) \\ &\triangleq H(p'_1, p'_2, \dots, p'_s) \end{aligned}$$

因此可得

$$I(X; Y) = H(Y) - H(p'_1, p'_2, \dots, p'_s)$$

信道容量为

$$C = \max_{P(x)} \{ H(Y) \} - H(p'_1, p'_2, \dots, p'_s)$$

这就变成求一种输入分布 $P(x)$ 使 $H(Y)$ 取最大值的问题。

共有 s 个输出符号,故 $H(Y)$ 的极大值为 $\log s$,且只有当输入符号等概分布时 $H(Y)$ 才达到此最大值。此种情况下不一定存在一种输入分布 $P(x)$ 能使输出符号达到等概率分布。但对于对称离散信道,当输入 X 等概分布时,输出 Y 恰好也取等概分布。这是由于

$$P(y) = \sum_X P(xy) = \sum_X P(x)P(y|x)$$

当 X 等概分布时有 $P(x)=1/r$, 故

$$P(y) = \sum_X \frac{1}{r} P(y|x) = \frac{1}{r} \sum_X P(y|x)$$

由于对称离散信道的矩阵 \mathbf{P} 中每一列之和为常数, 即 $\sum_X P(y|x) = \sum_{i=1}^r q'_i$, 故 $P(y) =$

$\frac{1}{r} \sum_{i=1}^r q'_i$, 与 y 无关, Y 中所有符号必然等概分布, 由此可得对称离散信道的信道容量为

$$C = \log s - H(p'_1, p'_2, \dots, p'_s) \quad (3.3.2)$$

上式是对称离散信道能够传输的最大平均信息量, 它只与对称信道矩阵中的行矢量 $\{p'_1, p'_2, \dots, p'_s\}$ 有关, 该信道的最佳输入分布为等概分布。

例 3.3.1 某对称离散信道的信道矩阵为

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

由式(3.3.2)可得其信道容量为

$$\begin{aligned} C &= \log 4 - H\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right) \\ &= 2 + \frac{1}{3} \log \frac{1}{3} + \frac{1}{3} \log \frac{1}{3} + \frac{1}{6} \log \frac{1}{6} + \frac{1}{6} \log \frac{1}{6} \\ &= 0.0817 (\text{比特/符号}) \end{aligned}$$

若信道的输入符号与输出符号数都等于 r , 且信道矩阵为

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

式中: $\bar{p} = 1 - p$ 。

此类信道称为**强对称信道**或**均匀信道**, 它的总错误概率为 p , 对称地平均分配给 $r-1$ 个输出符号。它是对称离散信道的一个特例, 其信道容量为

$$\begin{aligned} C &= \log r - H\left(\bar{p}, \frac{p}{r-1}, \frac{p}{r-1}, \dots, \frac{p}{r-1}\right) \\ &= \log r + \bar{p} \log \bar{p} + (r-1) \frac{p}{r-1} \log \frac{p}{r-1} \\ &= \log r - p \log(r-1) + (\bar{p} \log \bar{p} + p \log p) \end{aligned}$$

$$= \log r - p \log(r-1) - H(p) \quad (3.3.3)$$

对于 $r=2$ 的二元对称信道,由式(3.3.3)可计算得其信道容量为

$$C = 1 - H(p) \text{ (比特 / 符号)}$$

可见结果与前述结论一致。

3.3.4 一般离散信道的信道容量

信道容量是在固定信道的条件下平均互信息的最大值,由于 $I(X; Y)$ 是输入概率分布 $P(x)$ 的凹型凸函数,所以最大值一定存在。而 $I(X; Y)$ 是 r 个变量 $\{P(a_1), P(a_2), \dots, P(a_r)\}$ 的多元函数,其中 $P(x)$ 要满足非负和完备性的约束,因此求信道容量的问题就是多元函数求约束极值的问题。对离散信道, $P(x)$ 是离散的数值集合,最大值可以用拉格朗日乘子法求解; 对于连续信道,可以用变分法求解。对于一般 DMC 信道来说,有时还得不到一个清晰的解析解。本节介绍求信道容量的一般性规律及通用方法。

3.3.4.1 DMC 的信道容量定理

为了推导信道容量,首先介绍微分学中关于凸函数极值的一个定理。

定理 3.3.1 设 $f(x)$ 是定义在所有分量均非负的半无限向量空间上的可微上凸函数, $M = \max f(x)$ 是 $f(x)$ 在此空间上的最大值。则 $x = x^*$ 时能达到此最大值 M 的充要条件是

$$\begin{cases} \left. \frac{\partial f(x)}{\partial x} \right|_{x=x^*} \leq 0, & x = 0 \\ \left. \frac{\partial f(x)}{\partial x} \right|_{x=x^*} = 0, & x > 0 \end{cases} \quad (3.3.4)$$

考虑图 3.2.1 所示的单符号离散无记忆信道,我们要完成的工作就是在以下约束条件下求平均互信息的最大值,即

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

$$\begin{cases} \sum_{i=1}^r P(a_i) = 1 \\ P(a_i) \geq 0 \end{cases}$$

引入一个新函数

$$\phi = I(X; Y) - \lambda \left[\sum_{i=1}^r P(a_i) - 1 \right]$$

式中: λ 为拉格朗日乘子(待定常数)。

解方程组

$$\left. \frac{\partial \phi}{\partial P(a_l)} \right|_{P=P^*} = 0 \quad (l = 1, 2, \dots, r)$$

其中的关键步骤的结果:

$$\frac{\partial I(X; Y)}{\partial P(a_l)} = \sum_{j=1}^s P(b_j | a_l) \log P(b_j | a_l) - \sum_j \frac{\partial \left[\sum_i P(a_i) P(b_j | a_i) \right]}{\partial P(a_l)} \log P(b_j) -$$

$$\sum_{i,j} P(a_i b_j) \frac{\partial \log P(b_j)}{\partial P(a_l)}$$

由于

$$\frac{\partial \log P(b_j)}{\partial P(a_l)} = \frac{1}{P(b_j)} \frac{\partial P(b_j)}{\partial P(a_l)} \log e = \frac{P(b_j | a_l)}{P(b_j)} \log e \quad (l = 1, 2, \dots, r)$$

则有

$$\begin{aligned} \frac{\partial I(X; Y)}{\partial P(a_l)} &= \sum_{j=1}^s P(b_j | a_l) \log \frac{P(b_j | a_l)}{P(b_j)} - \sum_{i,j} P(a_i b_j) \frac{P(b_j | a_l)}{P(b_j)} \log e \\ &= \sum_{j=1}^s P(b_j | a_l) \log \frac{P(b_j | a_l)}{P(b_j)} - \sum_{j=1}^s P(b_j | a_l) \log e \\ &= \sum_{j=1}^s P(b_j | a_l) \log \frac{P(b_j | a_l)}{P(b_j)} - \log e \end{aligned}$$

因此

$$\frac{\partial \phi}{\partial P(a_l)} = \sum_{j=1}^s P(b_j | a_l) \log \frac{P(b_j | a_l)}{P(b_j)} - \log e - \lambda = 0 \quad (l = 1, 2, \dots, r)$$

综合可得

$$\begin{cases} \sum_{j=1}^s P(b_j | a_l) \log \frac{P(b_j | a_l)}{P(b_j)} = \log e + \lambda & (l = 1, 2, \dots, r) \\ \sum_{l=1}^r P(a_l) = 1 \end{cases} \quad (3.3.5)$$

在满足式(3.3.5)时平均互信息 $I(X; Y)$ 为最大, 即

$$C = \max_{P(X)} I(X; Y) = \sum_{l=1}^r P(a_l^*) \sum_{j=1}^s P(b_j | a_l^*) \log \frac{P(b_j | a_l^*)}{P(b_j)} = \log e + \lambda$$

令

$$I(x = a_i; Y) = \sum_{i=1}^r P(a_i) \sum_{j=1}^s P(b_j | a_i) \log \frac{P(b_j | a_i)}{P(b_j)}$$

它可表示输出端接收到 Y 后获得关于 $x = a_i$ 的信息量, 即是信源符号 $x = a_i$ 对输出端 Y 平均提供的互信息。由式(3.3.5)不难看出, 当平均互信息 $I(X; Y)$ 达到最大时, 每个输入符号 x 向输出 Y 提供的平均互信息量一样, 这个值就是该信道的信道容量。

由此, 得到离散无记忆信道有下述信道容量定理。

定理 3.3.2 一般离散信道的平均互信息 $I(X; Y)$ 达到极大值(等于信道容量)的充要条件是输入概率分布 $\{p_i\}$ 满足

$$I(x_i; Y) = C, \quad \text{对所有 } x_i \text{ 其 } p_i \neq 0$$

$$I(x_i; Y) \leq C, \quad \text{对所有 } x_i \text{ 其 } p_i = 0$$

定理 3.3.2 说明对于达到信道容量的最佳输入分布 $P^* = \{P^*(x_1), \dots, P^*(x_r)\}$, 所有概率非 0 的信源符号为输出端提供了相同的信息量。且根据此定理还可得出对称信道的容量定理。

定理 3.3.3 达到准对称 DMC 信道容量的输入概率分布相等。

证明: 若信道为准对称,则该信道输出符号集 Y 划分成若干个子集 Y_h ($h=1,2,\dots,s$),每个子集 Y_h 所对应的信道矩阵中列所组成的子阵是对称的。且设信道输入等概,即

$$P(x=a_k) = \frac{1}{r} \quad (k=1,2,\dots,r)$$

则

$$\begin{aligned} I(x=a_k; Y) &= \sum_{j=1}^s P(b_j | a_k) \log \frac{P(b_j | a_k)}{P(b_j)} \\ &= \sum_{j=1}^s P(b_j | a_k) \log \frac{P(b_j | a_k)}{\frac{1}{r} \sum_{i=1}^r P(b_j | a_i)} \\ &= \sum_h \sum_{j \in Y_h} P(b_j | a_k) \log \frac{P(b_j | a_k)}{\frac{1}{r} \sum_{i=1}^r P(b_j | a_i)} \end{aligned} \quad (3.3.6)$$

由于 Y_h 对应的子阵是对称矩阵,因此对每一个 $b_j \in Y_h$, $\frac{1}{r} \sum_{i=1}^r P(b_j | a_i)$ 是一样的,即属于同一个输出子集 Y_h 中的 b_j ,式(3.3.6)中分母部分相等(同一个输出子集中的符号概率 $P(b_j)$ 相等)。而同一子阵中各行元素 $\{P(b_j | a_k)\}$ 都可以由第一行置换得到,因此对于任何 a_k ,

$$\sum_{j \in Y_h} P(b_j | a_k) \log \frac{P(b_j | a_k)}{\frac{1}{r} \sum_{i=1}^r P(b_j | a_i)}$$

相等,式(3.3.6)中第一个求和是对 h 进行的,即对所有输出子集求和,与 k 无关,因此对任何 a_k , $I(x=a_k; Y)$ 相等,满足定理 3.3.2,即当信道输入等概时,平均互信息达到最大(信道容量)。

[证毕]

例 3.3.2 删除信道如图 3.3.2 所示,其信道矩阵为

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

可知该信道是准对称信道,其中 $Y_1 = \{0,1\}$, $Y_2 = \{2\}$,由定理 3.3.3 可求其信道容量。当输入等概时, $P(x=0) = P(x=1) = 0.5$,并且

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

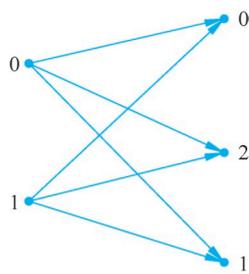


图 3.3.2 删除信道

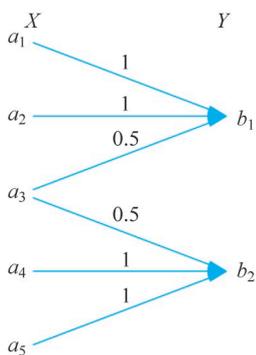


图 3.3.3 离散信道

例 3.3.3 设离散信道如图 3.3.3 所示,其信道矩阵为

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0.5 & 0.5 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

该信道仅对输出对称,可利用定理 3.3.2 求其信道容量。观察此信道,输入符号 a_3 传递到 b_1 和 b_2 是等概率的,可设想若 a_3 的概率分布等于零,则该信道变成无噪信道,而且 a_1, a_2 与 a_4, a_5 都分别传递到 b_1 与 b_2 ,因此可只取 a_1 和 a_5 。所以假设输入概率分布 $P(a_1)=P(a_5)=0.5, P(a_2)=P(a_3)=P(a_4)=0$,可计算得输出符号概率 $P(b_1)=P(b_2)=0.5$ 。不难计算

$$I(x=a_1; Y) = I(x=a_5; Y) = \log 2$$

$$I(x=a_2; Y) = I(x=a_3; Y) = I(x=a_4; Y) = 0$$

可见此假设分布满足定理 3.3.2,因此信道容量为

$$C = \log 2 = 1(\text{比特 / 符号})$$

最佳分布为

$$\begin{bmatrix} X \\ P^*(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ 0.5 & 0 & 0 & 0 & 0.5 \end{bmatrix}$$

实际上,还可以假设输入分布为 $P(a_1)=P(a_2)=P(a_4)=P(a_5)=0.25, P(a_3)=0$,同样可得输出符号概率 $P(b_1)=P(b_2)=0.5$,以及

$$\begin{cases} I(x_i; Y) = \log 2 & (x_i = a_1, a_2, a_4, a_5) \\ I(x_i; Y) = 0 < \log 2 & (x_i = a_3) \end{cases}$$

此假设分布也满足定理 3.3.2,因此信道容量仍为

$$C = \log 2 = 1(\text{比特 / 符号})$$

但最佳分布为

$$\begin{bmatrix} X \\ P^*(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

当然,还可找到此信道的其他最佳输入分布。这也反映出信道容量只与信道特性相关,与信道输入分布无关。因此,信道的最佳输入分布不一定是唯一的。从式(3.3.5)可知,互信息 $I(x_i; Y)$ 仅直接与信道转移概率及输出概率分布有关,因而达到信道容量的输入概率分布不是唯一的,但输出概率分布是唯一的。

3.3.4.2 信道矩阵为非奇异方阵的信道容量

定理 3.3.2 是一个构造性定理,它只给出了最佳输入概率分布应该满足的条件。为了得到具体的最佳输入概率分布 $P^* = \{P^*(x_1), P^*(x_2), \dots, P^*(x_r)\}$ 和信道容量 C ,

需要求解式(3.3.5)所示的方程组。这是一个包括 $r+1$ 个方程、 $r+1$ 个未知量 ($\{P^*(x_1), P^*(x_2), \dots, P^*(x_r)\}$ 和 C) 的方程组。当 $r < s$ 时, 求解方程组很困难, 需要反复试验或借助于计算机迭代计算法。当信道矩阵为非奇异方阵时, 方程组有唯一解。下面讨论这种情况。

可将式(3.3.5)中前 r 个方程

$$\sum_{j=1}^s P(b_j | a_i) \log \frac{P(b_j | a_i)}{P(b_j)} = C$$

等效改写为

$$\begin{aligned} \sum_{j=1}^s P(b_j | a_i) \log P(b_j | a_i) &= C + \sum_{j=1}^s P(b_j | a_i) \log P(b_j) \\ &= \sum_{j=1}^s P(b_j | a_i) [C + \log P(b_j)] \end{aligned}$$

进行变量代换, 令 $\beta_j = C + \log P(b_j)$, 上式可写为

$$\sum_{j=1}^s P(b_j | a_i) \log P(b_j | a_i) = \sum_{j=1}^s P(b_j | a_i) \beta_j \quad (i = 1, 2, \dots, r) \quad (3.3.7)$$

写成矩阵形式, 即

$$-\begin{bmatrix} H(Y | a_1) \\ H(Y | a_2) \\ \vdots \\ H(Y | a_r) \end{bmatrix} = \begin{bmatrix} P(b_1 | a_1) & \cdots & P(b_s | a_1) \\ \vdots & \ddots & \vdots \\ P(b_1 | a_r) & \cdots & P(b_s | a_r) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{bmatrix} \quad (3.3.8)$$

式(3.3.7)是一个包含 r 个方程、 s 个未知量的方程组, 其系数矩阵即为信道转移概率矩阵, 当 $r=s$ 且信道矩阵非奇异方阵, 即 $r=s$ 且信道矩阵是可逆矩阵时, 方程有唯一的解, 可解出 $\beta_1, \beta_2, \dots, \beta_s$ 。进一步地, 根据 $\beta_j = C + \log P(b_j)$ 和输出集的完备性可求得信道容量为

$$C = \log \left(\sum_{j=1}^r 2^{\beta_j} \right)$$

对应的输出概率分布为

$$P(b_j) = 2^{\beta_j - C} \quad (j = 1, 2, \dots, s)$$

再根据

$$P(b_j) = \sum_{i=1}^r P(a_i) P(b_j | a_i) \quad (j = 1, 2, \dots, s)$$

即可解出达到信道容量的最佳输入分布 $P^*(x = a_i)$ 。

例 3.3.4 计算如图 3.3.4 所示的 Z 信道的信道容量和最佳输入分布 (其中 $0 < p < 1$)。

解: 该信道矩阵为

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

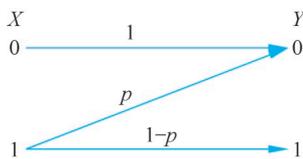


图 3.3.4 Z 信道

可知该信道不是对称类信道,但其信道矩阵为非奇异方阵,所以根据式(3.3.7)得

$$\begin{cases} \beta_1 = 1 \log 1 = 0 \\ p\beta_1 + (1-p)\beta_2 = p \log p + (1-p) \log(1-p) \end{cases}$$

求解方程组即得

$$\beta_1 = 0, \quad \beta_2 = \frac{H(p)}{p-1}$$

式中: $H(p)$ 为以 p 为独立自变量的二元信源熵。

所以信道容量为

$$C = \log[1 + 2^{\frac{H(p)}{p-1}}]$$

最佳输出分布为

$$P(b_1) = (1 + 2^{\frac{H(p)}{p-1}})^{-1}$$

$$P(b_2) = 1 - P(b_1) = \frac{2^{\frac{H(p)}{p-1}}}{1 + 2^{\frac{H(p)}{p-1}}}$$

最佳输入分布为

$$P(a_1) = 1 - \frac{p^{\frac{p}{1-p}}}{1 + (1-p)p^{\frac{p}{1-p}}}$$

$$P(a_2) = 1 - P(a_1) = \frac{p^{\frac{p}{1-p}}}{1 + (1-p)p^{\frac{p}{1-p}}}$$

3.3.4.3 信道容量的迭代算法

比较经典的离散无记忆信道容量的迭代算法是 Blahut-Arimoto 算法,由 S. Arimoto 和 R. E. Blahut 于 1972 年给出,它是一种有效的数值算法,能以任意给定的精度及有限步数计算出任意离散无记忆信道的信道容量,因此可以作为一般 DMC 的信道容量通用算法。该算法的理论基础就是定理 3.3.2,本节仅直接给出算法流程,其详细推导及收敛性证明参见文献[4]。

下面的定理给出了 Blahut-Arimoto 算法的基本思想。

定理 3.3.4 设信道的转移概率矩阵 $\mathbf{Q} = [P(b_j | a_i)]_{r \times s}$, P^0 是任给的输入符号的一个初始概率分布,其所有分量 $P^0(a_i)$ 均不为 0。按照下式不断对概率分布进行迭代、更新:

$$P^{r+1}(a_i) = P^r(a_i) \frac{\beta_i(P^r)}{\sum_{l=1}^K P^r(a_l) \beta_l(P^r)}$$

式中

$$\beta_i(P^r) = \exp[I(x = a_i; Y)] \Big|_{P=P^r} = \exp \left\{ \sum_{j=1}^s P(b_j | a_i) \log \frac{P(b_j | a_i)}{\sum_{l=1}^K P^r(a_l) P(b_j | a_l)} \right\}$$



则由此所得的 $I(P^r, Q)$ 序列收敛于信道容量 C 。

算法首先初始化输入符号的概率分布,使所有概率分量不为 0(一般情况下,往往将初始分布设为等概率分布),而后计算每个输入符号对输出集的平均互信息 $I(x=a_i; Y)$ 。在迭代过程中,不断提高具有较大平均互信息 $I(x=a_i; Y)$ 的输入符号 a_i 的概率,降低较小互信息的输入符号概率。当平均互信息与最大互信息足够接近时,迭代结束,认为平均互信息达到信道容量。其算法流程如图 3.3.5 所示。

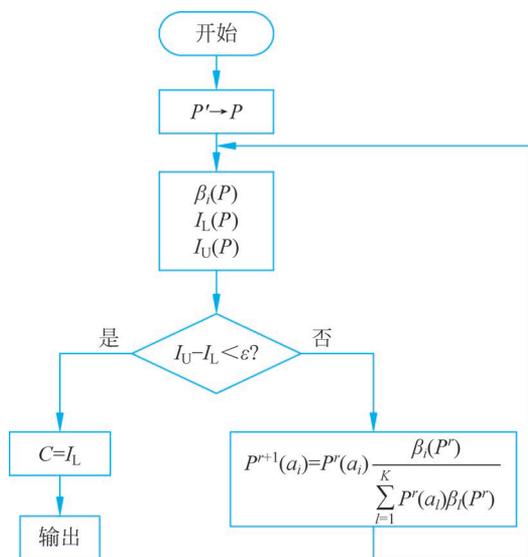


图 3.3.5 一般 DMC 信道容量迭代算法流程

图 3.3.5 中:

$$I_L(P) = \log \left\{ \sum_{i=1}^K P(a_i) \beta_i(P) \right\}$$

$$I_U(P) = \log \left\{ \max_i \beta_i(P) \right\}$$

3.3.5 离散无记忆 N 次扩展信道的信道容量

对于一般的离散无记忆信道,其信道容量的计算需附加许多条件,并通过复杂的迭代运算才能求得。不过一旦得到了离散无记忆信道的信道容量,就较易求得它的 N 次扩展信道的信道容量。

设基本信道为离散无记忆信道,其输入符号取自集合 $A = \{a_1, a_2, \dots, a_r\}$, 输出符号集合为 $B = \{b_1, b_2, \dots, b_s\}$, 信道矩阵为

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1s} \\ P_{21} & P_{22} & \cdots & P_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ P_{r1} & P_{r2} & \cdots & P_{rs} \end{bmatrix}$$

且满足

$$\sum_{j=1}^s P_{ij} = 1 \quad (i=1,2,\dots,r)$$

则此无记忆信道的 N 次扩展信道的数学模型可用图 3.2.5 表示。

该 N 次扩展信道的输入矢量 \mathbf{X} 的可能取值有 r^N 个,而输出矢量 \mathbf{Y} 的可能取值有 s^N 个。其信道转移矩阵为

$$\pi = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1s^N} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2s^N} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{r^N 1} & \pi_{r^N 2} & \cdots & \pi_{r^N s^N} \end{bmatrix}$$

式中:

$$\begin{aligned} \pi_{kh} &= P(\beta_h | \alpha_k) \\ &= P(b_{h_1} b_{h_2} \cdots b_{h_N} | a_{k_1} a_{k_2} \cdots a_{k_N}) \\ &= \prod_{i=1}^N P(b_{h_i} | a_{k_i}), k_i \in \{1,2,\dots,r^N\}, h_i \in \{1,2,\dots,s^N\} \end{aligned}$$

且满足

$$\sum_{h=1}^{s^N} \pi_{kh} = 1 \quad (k=1,2,\dots,r^N)$$

由定理 3.2.1,离散无记忆 N 次扩展信道的平均互信息满足

$$I(\mathbf{X}; \mathbf{Y}) \leq \sum_{i=1}^N I(X_i; Y_i)$$

当信源也为离散无记忆时,上式等号成立。故当信源 $\mathbf{X}=(X_1, X_2, \dots, X_N)$ 取无记忆信源,且各分信源 X_i 均取得最佳分布时,可得其信道容量为

$$\begin{aligned} C^N &= \max_{P(\mathbf{x})} I(\mathbf{X}; \mathbf{Y}) = \max_{P(\mathbf{x})} \sum_{i=1}^N I(X_i; Y_i) \\ &= \sum_{i=1}^N \max_{P(x_i)} I(X_i; Y_i) = \sum_{i=1}^N C_i \end{aligned} \quad (3.3.9)$$

式中令 $C_i = \max_{P(x_i)} I(X_i; Y_i)$,这是输入随机变量 $\mathbf{X}=(X_1, X_2, \dots, X_N)$ 中第 i 个随机变量 X_i 通过离散无记忆信道传输的最大信息量。若已求得离散无记忆信道的信道容量 C ,则由于输入随机序列 \mathbf{X} 中各变量在同一信道中传输,故有 $C_i=C(i=1,2,\dots,N)$,即任何时刻通过离散无记忆信道传输的最大信息量都相同。

由式(3.3.9)得

$$C^N = NC$$

此式说明离散无记忆 N 次扩展信道的信道容量等于原单符号离散信道的信道容量的 N 倍,且只有当输入信源是无记忆的,并且每一输入变量 X_i 的分布 $P(x)$ 各自达到最佳分布时,才能达到这个信道容量值 NC 。



视频

3.3.6 并联信道与串联信道的信道容量

在研究较复杂的系统时,往往可以将其分解成若干已知信道容量的简单信道的组合,从物理连接上看,组合多为信道间的串联或并联。并联信道中的各个子信道是并行的,根据输入与输出占用信道的方式又可分为输入并接信道、并用信道以及和信道。

设有 N 个子信道,其输入分别是 X_1, X_2, \dots, X_N , 输出分别是 Y_1, Y_2, \dots, Y_N , 各子信道的转移概率分别是 $P(y_1|x_1), P(y_2|x_2), \dots, P(y_N|x_N)$, 各子信道的信道容量分别是 C_1, C_2, \dots, C_N 。

3.3.6.1 输入并接信道

N 个信道的并接信道如图 3.3.6 所示,各子信道的输入集取自相同的符号集合 X , 输出集合取自不同的符号集 Y_1, Y_2, \dots, Y_N , 则该信道的平均互信息可表示为 $I(X; Y_1, \dots, Y_N)$ 。

不难得出以下结论:

$$I(X; Y_1, \dots, Y_N) = I(X; Y_1) + I(X; Y_2 | Y_1) + \dots + I(X; Y_N | Y_1 \dots Y_{N-1})$$

$$I(X; Y_1, \dots, Y_N) \geq I(X; Y_i)$$

$$C \geq \max\{C_1, C_2, \dots, C_N\}$$

$$C \leq \max_{P(x)} H(X)$$

在实际中,多次测量、多路径传输、协同通信等可以考虑用输入并接信道进行建模分析。

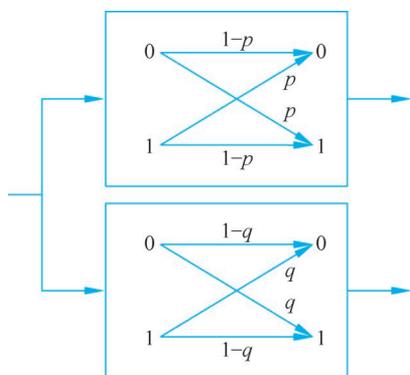


图 3.3.7 并接信道

例 3.3.5 将 $BSC(p)$ 与 $BSC(q)$ 组成输入并接信道如图 3.3.7 所示,求该组合信道的容量。

解: 组合信道的输入集为 $\{0, 1\}$, 输出集为 $\{00, 01, 10, 11\}$, 等效信道转移矩阵为

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

这是一个准对称信道,可以分成两个子矩阵

$$\mathbf{P}_1 = \begin{bmatrix} \bar{p}\bar{q} & pq \\ pq & \bar{p}\bar{q} \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} \bar{p}q & p\bar{q} \\ p\bar{q} & \bar{p}q \end{bmatrix}$$

当输入等概时,平均互信息最大,其信道容量为

$C = \log 2 - H(\bar{p}\bar{q}, \bar{p}q, p\bar{q}, pq) - (\bar{p}q + p\bar{q}) \log(\bar{p}q + p\bar{q}) - (\bar{p}\bar{q} + pq) \log(\bar{p}\bar{q} + pq)$
其容量曲线如图 3.3.8 所示。可见,该信道的容量依然是两个信道转移概率的下凸型函数。

3.3.6.2 并用信道

N 个信道的并用信道如图 3.3.9 所示,各子信道的输入和输出均取自不同的输入符

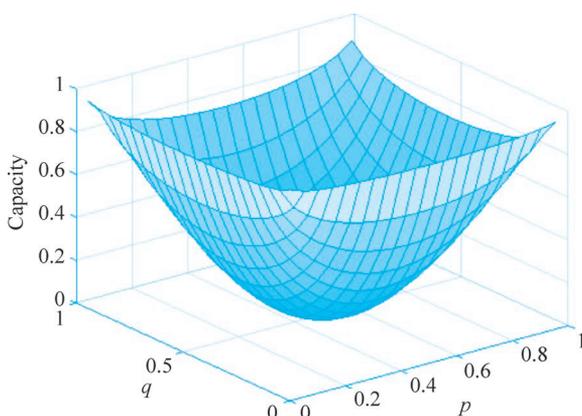


图 3.3.8 输入并接信道容量

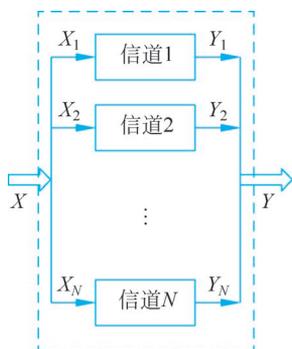


图 3.3.9 N 个信道的并用信道

号集合 X_1, X_2, \dots, X_N 和输出符号集 Y_1, Y_2, \dots, Y_N , 则该信道的平均互信息可表示为 $I(X_1 X_2 \dots X_N; Y_1 Y_2 \dots Y_N)$ 。

并用信道中 N 个子信道独立并行使用, 其等效信道转移概率为

$$P(y_1 y_2 \dots y_N | x_1 x_2 \dots x_N) = \prod_{i=1}^N P(y_i | x_i)$$

该信道相当于离散无记忆信道的 N 次扩展信道。因此有

$$I(\mathbf{X}; \mathbf{Y}) = I(X_1 X_2 \dots X_N; Y_1 Y_2 \dots Y_N) \leq \sum_{i=1}^N I(X_i; Y_i)$$

所以并用信道的信道容量为

$$C = \max \sum_{i=1}^N I(X_i; Y_i) = \sum_{i=1}^N \max I(X_i; Y_i) = \sum_{i=1}^N C_i$$

例 3.3.6 将 $BSC(p)$ 与 $BSC(q)$ 组合成并用信道, 求该组合信道的容量。

解: 并用信道的输入集和输出集均可表示为 $\{00, 01, 10, 11\}$, 等效转移矩阵为

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

该信道为对称信道, 其容量为

$$C = \log 4 - H(\bar{p}\bar{q}, \bar{p}q, p\bar{q}, pq) = 2 - H(p) - H(q) = C_1 + C_2$$

由此可见, 并用信道容量等于两个并联的子信道容量之和。

在实际的通信中, 诸如各类复用系统及近代提出的多输入多输出(MIMO)系统都可抽象为并用信道, 这也就不难解释为何可以利用 MIMO 信道成倍地提高无线信道容量, 成为现代的无线宽带移动通信系统的主要支撑技术。

3.3.6.3 和信道

和信道如图 3.3.10 所示,由 N 个独立的子信道并联组成,但信息传输每次只通过其中的一个信道。 r_1, r_2, \dots, r_N 表示各子信道的利用率,显然 $0 \leq r_i \leq 1 (i=1, 2, \dots, N)$ 且 $r_1 + r_2 + \dots + r_N = 1$ 。

下面以两个信道组成和信道为例,讨论该类信道的平均互信息与信道容量。

设信道 1 和信道 2 的转移概率矩阵分别为 \mathbf{P}_1 和 \mathbf{P}_2 , 两个信道被选用的概率(称为信道利用率)分别为 r_1 和 $r_2, r_1 + r_2 = 1$ 。输入符号集(信源)分别为 X_1 和 X_2 , 其数学模型分别如下:

$$\begin{aligned} \begin{bmatrix} X_1 \\ P(x_1) \end{bmatrix} &= \begin{bmatrix} a_1 & a_2 & \cdots & a_{q_1} \\ p_1 & p_2 & \cdots & p_{q_1} \end{bmatrix}, \quad \sum_{i=1}^{q_1} p_i = 1 \\ \begin{bmatrix} X_2 \\ P(x_2) \end{bmatrix} &= \begin{bmatrix} a'_1 & a'_2 & \cdots & a'_{q_2} \\ p'_1 & p'_2 & \cdots & p'_{q_2} \end{bmatrix}, \quad \sum_{i=1}^{q_2} p'_i = 1 \end{aligned}$$

那么组合成的和信道的输入数学模型为

$$\begin{bmatrix} \mathbf{X} \\ P(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & a_{q_2} & a'_1 & \cdots & a'_{q_2} \\ r_1 p_1 & \cdots & r_1 p_{q_1} & r_2 p'_1 & \cdots & r_2 p'_{q_2} \end{bmatrix}$$

显然, $\sum_{\mathbf{x}} P(\mathbf{x}) = 1$ 。

则和信道的输入熵为

$$\begin{aligned} H(\mathbf{X}) &= \sum_{\mathbf{x}} P(\mathbf{x}) \log \frac{1}{P(\mathbf{x})} = \sum_{i=1}^{q_1} r_1 p_i \log \frac{1}{r_1 p_i} + \sum_{i=1}^{q_2} r_2 p'_i \log \frac{1}{r_2 p'_i} \\ &= r_1 H(X_1) + r_2 H(X_2) + H(r_1) \end{aligned}$$

该信道的转移概率为

$$P(\mathbf{y} | \mathbf{x}) = \begin{cases} P_1(b_j | a_i) & (a_i \in X_1, b_j \in Y_1) \\ P_2(b_j | a_i) & (a_i \in X_2, b_j \in Y_2) \\ 0 & (\text{其他}) \end{cases}$$

可见,和信道的转移概率矩阵实际是一个由 \mathbf{P}_1 和 \mathbf{P}_2 构成的分块对角矩阵,即

$$\mathbf{Q} = \begin{bmatrix} \mathbf{P}_1 & 0 \\ 0 & \mathbf{P}_2 \end{bmatrix}$$

和信道的损失熵为

$$\begin{aligned} H(\mathbf{X} | \mathbf{Y}) &= H(X_1 X_2 | Y_1 Y_2) = \sum P(Y_i) H(X_i | Y_i) \\ &= r_1 H(X_1 | Y_1) + r_2 H(X_2 | Y_2) \end{aligned}$$

则平均互信息为

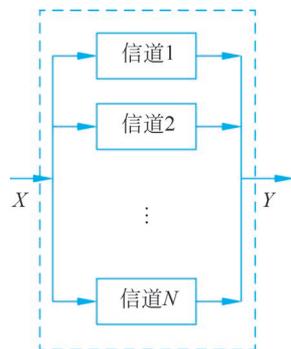


图 3.3.10 和信道

$$\begin{aligned}
 I(\mathbf{X}; \mathbf{Y}) &= H(\mathbf{X}) - H(\mathbf{X} | \mathbf{Y}) \\
 &= r_1 [H(X_1) - H(X_1 | Y_1)] + r_2 [H(X_2) - H(X_2 | Y_2)] + H(r_1, r_2) \\
 &= r_1 I(X_1; Y_1) + r_2 I(X_2; Y_2) + H(r_1, r_2)
 \end{aligned}$$

下面计算该信道的信道容量:

$$\begin{aligned}
 C &= \max_{r_1, P(x_1), P(x_2)} I(\mathbf{X}; \mathbf{Y}) = \max_{r_1, P(x_1), P(x_2)} \{r_1 I(X_1; Y_1) + r_2 I(X_2; Y_2) + H(r_1, r_2)\} \\
 &= \max_{r_1} \{r_1 \max_{P(x_1)} I(X_1; Y_1) + r_2 \max_{P(x_2)} I(X_2; Y_2) + H(r_1, r_2)\} \\
 &= \max_{r_1} \{r_1 C_1 + r_2 C_2 + H(r_1, r_2)\}
 \end{aligned}$$

令

$$\frac{\partial I(\mathbf{X}; \mathbf{Y})}{\partial r_1} = 0$$

有

$$C_1 - C_2 - \log r_1 + \log r_2 = 0$$

$$C_1 - \log r_1 = C_2 - \log r_2 \stackrel{\text{def}}{=} \lambda$$

因此

$$r_1 = 2^{C_1 - \lambda}$$

$$r_2 = 2^{C_2 - \lambda}$$

又

$$r_1 + r_2 = 1$$

可得

$$\lambda = \log(2^{C_1} + 2^{C_2})$$

进而可得

$$\lambda = C = \log(2^{C_1} + 2^{C_2})$$

同理,对 N 个信道的和信道,有

$$\mathbf{Q} = \begin{bmatrix} \mathbf{P}_1 & & & \\ & \mathbf{P}_2 & & \\ & & \ddots & \\ & & & \mathbf{P}_N \end{bmatrix}$$

设各个信道的利用率为 $P_i(C) = r_i$, 则其平均互信息为

$$I(\mathbf{X}; \mathbf{Y}) = \sum_{i=1}^N r_i I(X_i; Y_i) + H(r_1 r_2 \cdots r_N)$$

其信道容量为

$$C = \log_2 \sum_{i=1}^N 2^{C_i}$$

各子信道要满足各自的最佳输入分布,且最佳信道利用率 $P_i(C) = 2^{(C_i - C)}$ 。

例 3.3.7 将 $\text{BSC}(p)$ 与 $\text{BSC}(q)$ 组合成和信道,求该组合信道的容量。

解：和信道的等效信道转移矩阵为

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

信道容量为

$$C = \log[2^{1-H(p)} + 2^{1-H(q)}]$$

不同参数下的和信道容量如图 3.3.11 所示。

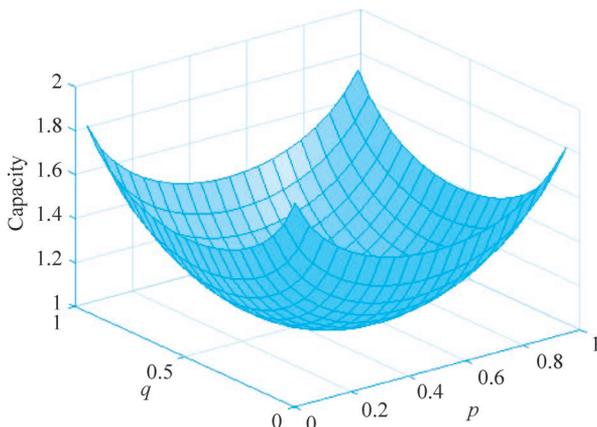


图 3.3.11 和信道容量

将上述三个例子的结果进行对比,如图 3.3.12 所示。由图可以看出和信道带来的优势。当 $p=0.5, q=0.1$ 时,信道 $\text{BSC}(p)$ 的容量 $C_1=0$, $\text{BSC}(q)$ 的容量 $C_2=0.531$,最佳信道利用率 $r_1=0.409, r_2=0.591$,可得容量 $C=1.2898$ 。比其中任意子信道的容量都大,其原因是和信道在信道选择(切换)过程中选择的不确定性也携带了一部分信息。利用这一原理就可解释空间调制(Spatial Modulation, SM)技术提高分集增益的本质。

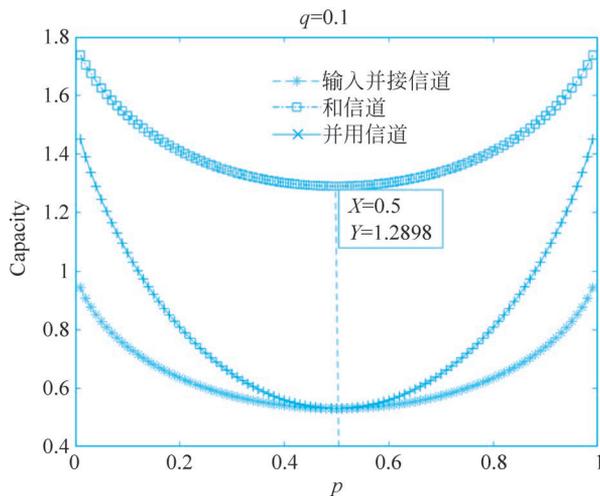


图 3.3.12 三种并联类组合信道的容量对比



空间调制的基本原理(图 3.3.13)是将一组信息比特分为空间维和符号域两部分:一部分经过传统的调制进行发送,如 PSK、QAM 等;另一部分是将部分调制信息映射到空间调制系统中相应的天线上进行发送,而这种处理就相当于组建了一个和信道,所以空间调制系统的天线序号也承载了一部分发送信息,这样携带的信息为 $k = \log N_t + \log M$ (N_t 为发送天线数目, M 为调制方式的阶数)。由此可见,系统的频谱利用率大大提高了。

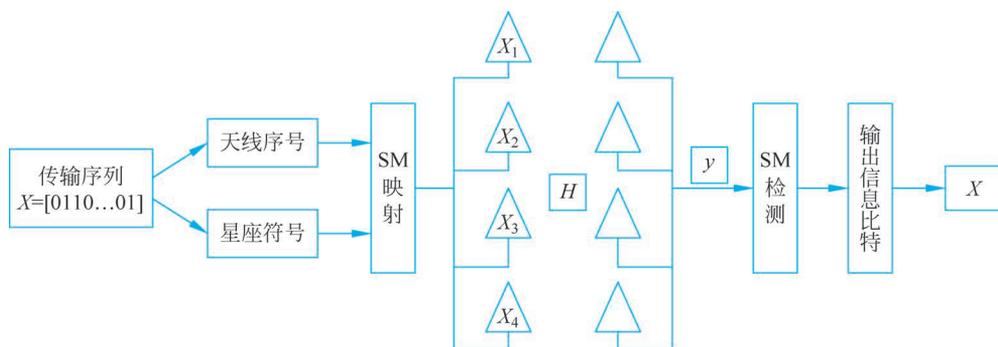


图 3.3.13 SM 的基本原理

例如,传输比特序列为 0110,其中前面两个比特 01 对应于选择天线 2(按照表 3.3.1 中的映射规则),后面两比特 10 则用来选择调制符号 3,所以 0110 比特信息在空间调制系统中是激活第二根天线发送第三个符号。

表 3.3.1 映射规则

输入信息比特	天线索引	调制符号
00	1	1
01	2	2
10	3	3
11	4	4

空间调制是一种新型的 MIMO 无线通信技术,与 MIMO 系统相比较,因其只需要激活一根发射天线来进行信息的传输,所以能够有效地避免多天线系统天线间同步(IAS)和天线间干扰(ICI)等问题。空间调制技术已被广泛运用于 5G 通信系统。

3.3.6.4 串联信道

串联信道是一种基本的信道组合方式,在卫星中继通信、微波中继接力等实际应用中广泛存在。当然,还包括在许多数据处理系统中,往往是根据执行功能或任务需求将整个系统划分成若干个独立处理的单元构成的串联信道。信道串联的一个关键前提是前一个子信道的输出集与下一级的子信道输入集相同。如两个信道组成的串联信道如图 3.3.14 所示。

图中串联信道的等效信道转移概率为 $P(z|x)$,若 X, Y, Z 满足马尔可夫链 $X \rightarrow Y \rightarrow Z$,即 X 和 Y 是统计相关的, Y 和 Z 是统计相关的,所以 X 和 Z 也是统计相关的;但若 Y 已知,则 X 和 Z 就是统计无关,即 $I(X; Z|Y) = 0$ 。总信道的传递概率矩阵为子信道的

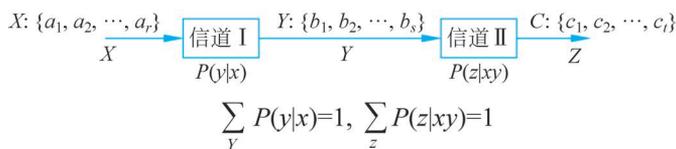


图 3.3.14 串联信道

信道矩阵乘积,即

$$\mathbf{P} = \mathbf{P}_1 \mathbf{P}_2$$

串联信道的信息传输测度函数存在如下规律。

定理 3.3.5 (数据处理定理) 对于马尔可夫链 $X \rightarrow Y \rightarrow Z$, 有

$$I(X; Z) \leq I(X; Y)$$

$$I(X; Z) \leq I(Y; Z)$$

证明: 根据平均互信息的链式法则可得

$$I(X; YZ) = I(X; Y) + I(X; Z | Y) = I(X; Z) + I(X; Y | Z)$$

对于马尔可夫链 $X \rightarrow Y \rightarrow Z$, 有 $I(X; Z | Y) = 0$, 且由平均互信息的非负性可知 $I(X; Y | Z) \geq 0$, 故有 $I(X; Y) \geq I(X; Z)$ 。

同理, 有

$$I(XY; Z) = I(Z; XY) = I(Z; Y) + I(Z; X | Y) = I(Z; X) + I(Z; Y | X)$$

$$I(Z; X | Y) = 0, I(Z; Y | X) \geq 0$$

因此有

$$I(Z; Y) \geq I(Z; X)$$

根据平均互信息的对称性可得

$$I(Y; Z) \geq I(X; Z) \quad [\text{证毕}]$$

数据处理定理也称为信息不增性或数据处理不等式, 它说明串联信道的传输只会丢失信息而不会额外增加信息, 即在数据处理中每增加一次处理环节, 一般会损失一部分信息, 最多保持原来获得的信息, 不可能比原来获得的信息有所增加。只有当串联信道的等效信道矩阵等于第一级信道矩阵时(如图 3.3.14 串联信道中第二个信道是无噪一对应信道), 通过串联信道传输后不会增加信息的损失。

例 3.3.8 将以下两个信道进行串联:

$$\mathbf{P}_1 = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

不难看出, 串联信道的信道矩阵为

$$\mathbf{P} = \mathbf{P}_1 \mathbf{P}_2 = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

故有 $I(X; Z) = I(X; Y)$ 。该串联信道不会使信道中信息损失增加。

下面考虑串联信道的信道容量问题。

由于串联信道的信道矩阵等于各子信道的信道矩阵的乘积, 即

$$\mathbf{P} = \mathbf{P}_1 \mathbf{P}_2 \cdots \mathbf{P}_N$$

基于 \mathbf{P} 可以计算等效串联信道的信道容量。而根据定理 3.3.5 可知, 串联信道的容量不大于组成它的各子信道的容量 $C_i (i=1, 2, \dots, N)$, 因此有

$$C \leq \min\{C_1, C_2, \dots, C_N\}$$

例 3.3.9 求 N 个相同的二元对称信道 BSC(p) 串联的信道容量。

解: 设 BSC(p) 的信道矩阵为

$$\mathbf{P}_0 = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

BSC(p) 的信道容量为

$$C = 1 - H(p)$$

将 BSC(p) 进行两级串联, 其信道矩阵为

$$\mathbf{P} = \mathbf{P}_0^2 = \begin{bmatrix} (1-p)^2 + p^2 & 2p(1-p) \\ 2p(1-p) & (1-p)^2 + p^2 \end{bmatrix}$$

该信道仍为二元对称信道, 只是错误传输概率由 p 变为 $2p(1-p)$, 因此其信道容量为

$$C_2 = 1 - H[2p(1-p)]$$

可以验证, $C_2 \leq C_1$, 当且仅当 $p=0.5$ 时等式成立。

再考虑 BSC(p) 的 N 级串联情况, 其等效信道矩阵 $\mathbf{P} = \mathbf{P}_0^N$, 通过正交变换可以把 \mathbf{P}_0 分解为

$$\mathbf{P}_0 = \mathbf{T}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1-2p \end{bmatrix} \mathbf{T}$$

式中

$$\mathbf{T} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

因此有

$$\begin{aligned} \mathbf{P} = \mathbf{P}_0^N &= \mathbf{T}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1-2p \end{bmatrix}^N \mathbf{T} \\ &= \mathbf{T}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & (1-2p)^N \end{bmatrix} \mathbf{T} \\ &= \frac{1}{2} \begin{bmatrix} 1+(1-2p)^N & 1-(1-2p)^N \\ 1-(1-2p)^N & 1+(1-2p)^N \end{bmatrix} \end{aligned}$$

可见, N 级 BSC 的串联信道也是二元对称信道, 其信道容量为

$$C_N = 1 - H[1-(1-2p)^N]$$

若 $p \neq 0$, 当 N 趋于无穷大时, 有

$$\mathbf{P}_0^\infty = \lim_{N \rightarrow \infty} \mathbf{P}_0^N = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

此时的信道容量为 0。因此, 当 $N \rightarrow \infty$ 时, $C_N \rightarrow 0$, 即信息完全损失。图 3.3.15 表示了

BSC(p)串联信道的容量变化情况,随着级联次数的增加,信道容量逐渐减小,值得注意的是,当BSC的错误传输概率 p 很小时这种减小并不大,实际数字通信系统中二元对称信道的 p 一般在 10^{-6} 以下。所以,若干次串接后信道容量的减少并不很明显,仍可保持一定的数值。

以上的结论都是在单符号离散无记忆信道中加以讨论和证明的。对于输入和输出都是随机序列的一般离散信道,数据处理定理仍然成立。在一般通信系统中,信源序列 S 经过编码模块形成 X ,然后经过信道输出 Y ,再经过译码模块得到 Z 。其中,噪声信道的输出矢量 Y 中所包含信源序列 S 的平均信息量大于数据处理后所估算的矢量 Z 中所包含信源序列 S 的平均信息量。这一点在理论上是正确的,但并不说明不需要进行数据处理。在实际应用中往往为了获得更有用和更有效的信息,还是需要进行适当的数据处理的。

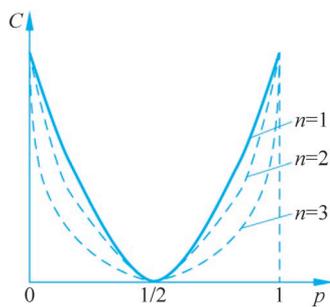


图 3.3.15 N 级 BSC(p)串联信道的信道容量变化

3.4 连续信道及其容量

连续信道强调在其中传输的信号幅度连续取值,即考察某时刻信道的输入与输出时,以连续型随机变量 X 和 Y 表示,若考察若干离散时刻信道的输入与输出,则以连续型随机序列 $\mathbf{X}=(X_1 X_2 \cdots X_N)$ 和 $\mathbf{Y}=(Y_1 Y_2 \cdots Y_N)$ 表示。

3.4.1 基本连续信道

基本连续信道就是单符号连续信道,输入和输出以单个连续型随机变量表示,如图 3.4.1 所示。其输入 X 取值于连续幅度区间 $[a, b]$ 或实数域 \mathbf{R} ,输出 Y 取值于连续幅度区间 $[c, d]$ 或实数域 \mathbf{R} ,输入与输出之间的变换依靠信道的转移概率密度函数 $p(y|x)$ 联系起来,并满足

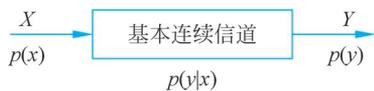


图 3.4.1 基本连续信道

$$\int_{\mathbf{R}} p(y|x) dy = 1 \quad (3.4.1)$$

基本连续信道数学模型可以用 $[X, p(y|x), Y]$ 表示。其输入信源 X 为

$$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} (a, b) \\ p(x) \end{bmatrix} \quad \text{或} \quad \begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} \mathbf{R} \\ p(x) \end{bmatrix}$$

$$\int_a^b p(x) dx = 1 \quad \text{或} \quad \int_{\mathbf{R}} p(x) dx = 1$$

其输出信宿为

$$\begin{bmatrix} Y \\ p(y) \end{bmatrix} = \begin{bmatrix} (c, d) \\ p(y) \end{bmatrix} \quad \text{或} \quad \begin{bmatrix} Y \\ p(y) \end{bmatrix} = \begin{bmatrix} \mathbf{R} \\ p(y) \end{bmatrix}$$

$$\int_a^b p(y) dy = 1 \quad \text{或} \quad \int_{\mathbf{R}} p(y) dy = 1$$



中文教学
录像

而信道的传递概率密度函数为 $p(y|x)$, 满足式(3.4.1)。

第2章已经介绍了连续分布随机变量的各类微分熵, 对连续信道的平均互信息来说, 不但它的这些关系式和离散信道下平均互信息的关系式完全类似, 而且它保留了离散信道的平均互信息的所有含义和性质, 只是表达式中用连续信源的微分熵替代了离散信源的熵。可以将3.2.2节中的平均互信息引入描述两个单符号连续分布随机变量之间的平均互信息:

$$I(X; Y) = \iint_{\mathbf{R}} p(xy) \log \frac{p(x|y)}{p(x)} dx dy = h(X) - h(X|Y) \quad (3.4.2)$$

$$I(X; Y) = \iint_{\mathbf{R}} p(xy) \log \frac{p(y|x)}{p(y)} dx dy = h(Y) - h(Y|X) \quad (3.4.3)$$

$$I(X; Y) = \iint_{\mathbf{R}} p(xy) \log \frac{p(xy)}{p(x)p(y)} dx dy = h(X) + h(Y) - h(XY) \quad (3.4.4)$$

可见, 连续信道的平均互信息与各种微分熵之间的关系和离散信道的相应关系是一致的, 同样单符号连续信道的信息传输率 $R = I(X; Y)$ (比特/自由度)。其信道容量为

$$C = \max_{p(x)} I(X; Y) = \max_{p(x)} [h(Y) - h(Y|X)] \text{ (比特/自由度)} \quad (3.4.5)$$

当扩展到多维连续信道时也能得到类似结论。

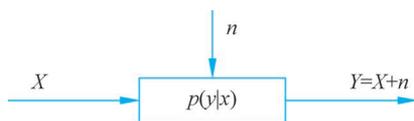


图 3.4.2 连续加性信道

下面考虑连续加性信道, 其输入、输出和噪声分别是连续型随机变量 X 、 Y 和 n , 且有 $Y = X + n$, 如图 3.4.2 所示。

定理 3.4.1 对于加性噪声信道, 有

$$p(y|x) = p(n), \quad h(Y|X) = h(n|X) = h(n)$$

式中: $p(n)$ 、 $h(n)$ 分别为噪声源的概率密度函数和微分熵。

证明: 由加性信道特点做坐标变换, $X = X, n = Y - X$ 。

再根据随机向量变换的联合概率密度函数计算方法, 有

$$\begin{aligned} p(xy) &= p(xn) \left| \frac{\partial(X, n)}{\partial(X, Y)} \right| = p(xn) \begin{vmatrix} \frac{\partial X}{\partial X} & \frac{\partial X}{\partial Y} \\ \frac{\partial n}{\partial X} & \frac{\partial n}{\partial Y} \end{vmatrix} \\ &= p(xn) \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = p(xn) \end{aligned}$$

由于信道输入 X 与噪声 n 是统计独立的, 可进一步得到

$$p(xy) = p(x)p(n)$$

因此

$$p(n) = \frac{p(xy)}{p(x)} = p(y|x)$$

进一步代入微分熵计算公式, 并利用坐标变换 $dx dy = dx dn$, 可得

$$h(Y|X) = - \iint_{\mathbf{R}} p(xy) \log p(y|x) dx dy$$

$$\begin{aligned}
&= -\iint_{\mathbf{R}} p(xn) \log p(n) dx dn \\
&= -\int_{-\infty}^{\infty} p(x) dx \int_{-\infty}^{\infty} p(n) \log p(n) dn \\
&= h(n) \quad \quad \quad \text{[证毕]}
\end{aligned}$$

定理 3.4.1 揭示了加性信道的一个重要性质,即信道的转移概率密度函数就是噪声的概率密度函数,同时噪声熵即为噪声源的微分熵。

基于此,结合式(3.4.3)和式(3.4.4)可得单符号加性信道 $[X, p(y|x), Y]$ 的信息传输率和信道容量分别为

$$\begin{aligned}
R &= I(X; Y) = h(Y) - h(n) \text{ (比特/自由度)} \\
C &= \max_{p(x)} I(X; Y) = \max_{p(x)} [h(Y) - h(n)] \\
&= \max_{p(x)} h(Y) - h(n) \text{ (比特/自由度)}
\end{aligned}$$

由于噪声 n 与输入 X 统计独立,因此加性信道中 $h(n)$ 与 $p(x)$ 无关,即改变输入 X 的概率分布只能影响 $h(Y)$,因此求加性信道的信道容量就是求某个最佳输入概率密度 $p^*(x)$ 使 $h(Y)$ 达到最大。

对于单符号加性连续信道,若加入信道的噪声概率密度函数服从高斯分布(正态分布),该信道称为加性高斯信道。这是在通信工程中最常用的具有实输入/输出的噪声信道模型。

设信道叠加的噪声 n 是均值为零、方差为 σ^2 的一维高斯噪声,则其概率密度函数为

$$p(n) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

噪声熵为

$$h(n) = \frac{1}{2} \log(2\pi e \sigma^2)$$

因此,单符号加性高斯信道的信道容量为

$$C = \max_{p(x)} h(Y) - h(n) = \max_{p(x)} h(Y) - \frac{1}{2} \log(2\pi e \sigma^2)$$

可见,只有 $h(Y)$ 与输入信号的概率密度函数 $p(x)$ 有关。当信道输出信号 Y 的平均功率限制在 P_0 时,由第2章的微分熵性质知,若 Y 是均值为零的高斯变量,其熵 $h(Y)$ 为最大。

对于加性信道 $Y=X+n$ 而言,输出信号的均方值为

$$E(Y^2) = E[(X+n)^2] = E(X^2) + 2E(X)E(n) + E(n^2)$$

平稳随机过程的均方值就是其平均功率。而且在一般通信系统的研究中,常假设输入信号和噪声都是均值为0的随机过程,此时信号的方差就等于其平均功率,即信道的平均噪声功率 $P_n = \sigma^2$,而输出信号功率 P_0 等于输入信号功率 P_s 与噪声功率 P_n 之和,即

$$P_0 = P_s + P_n$$

则输入信号平均功率受限意味着输出平均功率也受限,如限制在 P_o 。在此条件下,只有当输出信号 Y 服从高斯分布时, $h(Y)$ 才能取得最大值,即

$$\max h(Y) = \frac{1}{2} \log(2\pi e P_o)$$

因此,单符号加性高斯信道的信道容量为

$$\begin{aligned} C &= \frac{1}{2} \log(2\pi e P_o) - \frac{1}{2} \log(2\pi e \sigma^2) \\ &= \frac{1}{2} \log\left(\frac{P_o}{\sigma^2}\right) \\ &= \frac{1}{2} \log\left(1 + \frac{P_s}{\sigma^2}\right) = \frac{1}{2} \log\left(1 + \frac{P_s}{P_n}\right) \quad (\text{比特 / 自由度}) \end{aligned} \quad (3.4.6)$$

最佳输入分布 $p^*(x)$ 是均值为 0、方差(平均功率)为 P_s 的高斯分布。

3.4.2 多维连续信道

多维连续信道输入是 N 维连续型随机序列 $\mathbf{X} = (X_1 X_2 \cdots X_N)$, 输出是 N 维连续型随机序列 $\mathbf{Y} = (Y_1 Y_2 \cdots Y_N)$, 信道传递概率密度函数 $p(\mathbf{y} | \mathbf{x}) = p(y_1 y_2 \cdots y_N | x_1 x_2 \cdots x_N)$, 满足

$$\iint_{\mathbf{R}} \cdots \int_{\mathbf{R}} p(y_1 y_2 \cdots y_N | x_1 x_2 \cdots x_N) dy_1 dy_2 \cdots dy_N = 1$$

式中: \mathbf{R} 为实数域。

用 $[\mathbf{X}, p(\mathbf{y} | \mathbf{x}), \mathbf{Y}]$ 来描述多维连续信道。与离散无记忆信道的定义一样,若连续信道在任一时刻输出的变量只与对应时刻的输入变量有关,与以前时刻的输入、输出变量无关,也与以后的输入变量无关,则此信道为**无记忆连续信道**。满足连续无记忆信道的充要条件:

$$p(\mathbf{y} | \mathbf{x}) = p(y_1 y_2 \cdots y_N | x_1 x_2 \cdots x_N) = \prod_{i=1}^N p(y_i | x_i) \quad (3.4.7)$$

一般情况,式(3.4.7)不满足,也就是多维连续信道任何时刻的输出变量与其他任何时刻的输入与输出变量一般都是相关的,则此信道称为**连续有记忆信道**。

由式(3.4.2)~式(3.4.5)推广可得多维连续信道的平均互信息及信道容量:

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{X}) - h(\mathbf{X} | \mathbf{Y}) = \iint_{\mathbf{R}} p(\mathbf{xy}) \log \frac{p(\mathbf{x} | \mathbf{y})}{p(\mathbf{x})} dx dy \quad (3.4.8)$$

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{Y} | \mathbf{X}) = \iint_{\mathbf{R}} p(\mathbf{xy}) \log \frac{p(\mathbf{y} | \mathbf{x})}{p(\mathbf{y})} dx dy \quad (3.4.9)$$

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{X}) + h(\mathbf{Y}) - h(\mathbf{XY}) = \iint_{\mathbf{R}} p(\mathbf{xy}) \log \frac{p(\mathbf{xy})}{p(\mathbf{x})p(\mathbf{y})} dx dy \quad (3.4.10)$$

$$C = \max_{p(\mathbf{x})} I(\mathbf{X}; \mathbf{Y}) = \max_{p(\mathbf{x})} [h(\mathbf{Y}) - h(\mathbf{Y} | \mathbf{X})] \quad (3.4.11)$$

同理,多维加性信道的输入、输出及噪声为随机序列形式:

$$\mathbf{X} = (X_1 X_2 \cdots X_N), \mathbf{Y} = (Y_1 Y_2 \cdots Y_N), \mathbf{n} = (n_1 n_2 \cdots n_N)$$

则其信息传输率和信道容量分别为

$$\begin{aligned} R &= I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{n}) \text{ (比特 / 自由度)} \\ C &= \max_{p(x)} I(\mathbf{X}; \mathbf{Y}) = \max_{p(x)} [h(\mathbf{Y}) - h(\mathbf{n})] \text{ (比特 / 自由度)} \end{aligned} \quad (3.4.12)$$

3.5 波形信道与香农公式

波形信道的输入是随机过程 $\{x(t)\}$, 输出也是随机过程 $\{y(t)\}$ 。实际波形信道的带宽总是受限的, 所以在有限观测时间内满足限频 F 、限时 T 的条件。根据时间采样定理可将波形信道的输入 $\{x(t)\}$ 和输出 $\{y(t)\}$ 的随机过程信号离散化, 转换成 N 维随机序列 $\mathbf{X} = (X_1 X_2 \cdots X_N)$ 和 $\mathbf{Y} = (Y_1 Y_2 \cdots Y_N)$, 这样波形信道就可转化成多维连续信道描述, 如图 3.5.1 所示。

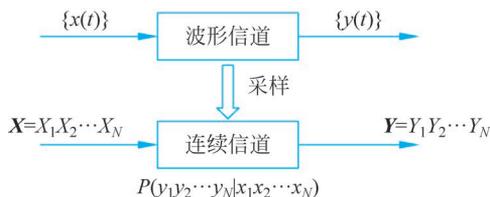


图 3.5.1 波形信道及其与多维连续信道关系

故而可得波形信道的平均互信息为

$$\begin{aligned} I(x(t); y(t)) &= \lim_{N \rightarrow \infty} I(\mathbf{X}; \mathbf{Y}) \\ &= \lim_{N \rightarrow \infty} [h(\mathbf{X}) - h(\mathbf{X} | \mathbf{Y})] \\ &= \lim_{N \rightarrow \infty} [h(\mathbf{Y}) - h(\mathbf{Y} | \mathbf{X})] \\ &= \lim_{N \rightarrow \infty} [h(\mathbf{X}) + h(\mathbf{Y}) - h(\mathbf{XY})] \end{aligned} \quad (3.5.1)$$

一般情况下, 波形信道都是研究其单位时间内的信息传输率, 即信息传输速率 R_t 。设 $\{x(t)\}$ 和 $\{y(t)\}$ 的持续时间为 T , 波形信道的信息传输速率和单位时间信道容量为

$$\begin{aligned} R_t &= \frac{1}{T} I(\mathbf{X}; \mathbf{Y}) = \frac{1}{T} \lim_{N \rightarrow \infty} [h(\mathbf{Y}) - h(\mathbf{Y} | \mathbf{X})] \text{ (b/s)} \\ C_t &= \max_{p(x)} R_t = \max_{p(x)} \left\{ \frac{1}{T} \lim_{N \rightarrow \infty} [h(\mathbf{Y}) - h(\mathbf{Y} | \mathbf{X})] \right\} \text{ (b/s)} \end{aligned} \quad (3.5.2)$$

式中: $p(x)$ 为输入序列 \mathbf{X} 的概率密度函数。

若研究加性信道, 由定理 3.4.1 对于加性噪声信道, 有加性信道的噪声熵 $h(\mathbf{Y} | \mathbf{X})$ 就是噪声源的熵 $h(\mathbf{n})$ 。结合式 (3.5.2) 可得一般加性波形信道单位时间的信道容量为

$$C_t = \max_{p(x)} \left\{ \frac{1}{T} \lim_{N \rightarrow \infty} [h(\mathbf{Y}) - h(\mathbf{n})] \right\} \text{ (b/s)} \quad (3.5.3)$$

由于在不同限制条件下连续随机变量有不同的最大微分熵取值, 所以由式 (3.5.2) 和式 (3.5.3) 知, 加性信道的信道容量 C 取决于噪声的统计特性和输入随机矢量 \mathbf{X} 所受的限制条件。对于一般通信系统, 无论是输入信号还是噪声, 它们的平均功率或能量总



中文教学
录像



视频

是有限的。所以本节只讨论在平均功率受限的条件下波形信道的信道容量。

3.5.1 带限高斯白噪声加性波形信道的信道容量

在模拟通信系统中最常用的信道模型是高斯白噪声加性 (Additive White Gaussian Noise, AWGN) 信道。其输入和输出分别是随机过程 $\{x(t)\}$ 和 $\{y(t)\}$, 信道噪声是加性高斯白噪声 $\{n(t)\}$, 其均值为零, 单边带功率谱密度为 N_0 , 所以输出信号满足 $\{y(t)\} = \{x(t)\} + \{n(t)\}$ 。另外, 实际信道的频带宽度总是有限的, 设其带宽为 W 。这样, 信道的输入、输出信号和噪声都是频带受限的随机过程。进一步假设输入与输出信号的持续时间为 T 。(需要说明的是, 根据时频域物理定律, 持续时间有限的信号, 其频带无限宽; 反之, 频带宽度有限的信号, 其持续时间必然无限长。因此, 既限频带又限时长的信号是物理不可能的, 但如果信号的主要能量集中在带宽 W 内, 这一假设不会造成太大的误差。)

根据奈奎斯特采样定理, 一个频带受限 W 内的信号可以通过 $2W$ 采样率得到的样本唯一表示, 由于信号持续时间为 T , 因此这个波形信道可以通过采样变成一个时间离散化 $2WT$ 维的连续信道, 即以自由度为 $N = 2WT$ 的随机序列形式处理, 如图 3.5.2 所示。由于是加性信道, 所以随机序列信道也满足 $\mathbf{Y} = \mathbf{X} + \mathbf{n}$ 。

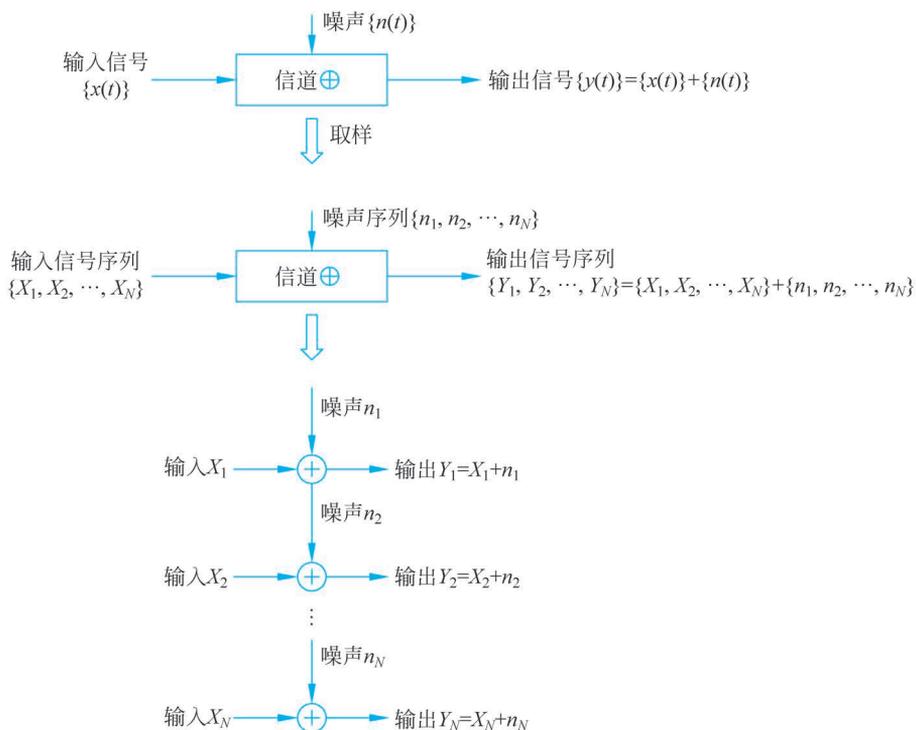


图 3.5.2 带限高斯白噪声加性信道变换成 N 个独立并联高斯加性信道

因为信道的频带是受限的, 所以加入信道的噪声成为带限高斯白噪声。而低频带限高斯白噪声的各样本值彼此统计独立, 所以限频的高斯白噪声过程可分解为 N 维统计独立的随机序列, 其中每个分量 n_i 都是均值为零, 方差为

$$\sigma_{n_i}^2 = P_{n_i} = \sigma^2 = N_0 WT / (2WT) = N_0 / 2 \quad (3.5.4)$$

可得 N 维的联合概率密度为

$$p(\mathbf{n}) = p(n_1 n_2 \cdots n_N) = \prod_{i=1}^N p(n_i) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{n_i^2}{2\sigma^2}}$$

对加性信道来说,若上式成立,意味着信道是无记忆的。那么,随机序列信道就可等效成 N 个独立高斯加性信道的并联。结合式(3.4.6)及并联信道容量计算方法可得等效信道容量为

$$C^N = \sum_{i=1}^N \frac{1}{2} \log\left(1 + \frac{P_{s_i}}{P_{n_i}}\right)$$

由式(3.5.4)可得高斯白噪声的每个样本值的方差 $\sigma_{n_i}^2 = P_{n_i} = N_0 / 2$ 。而信号的平均功率受限为 P_s , T 时间内总平均功率为 $P_s T$, 每个信号样本值的平均功率为 $(P_s T) / (2WT) = P_s / 2W$ 。所以可得 $[0, T]$ 时刻内信道的信道容量为

$$C^N = \frac{N}{2} \log\left(1 + \frac{P_s / 2W}{N_0 / 2}\right) = \frac{N}{2} \log\left(1 + \frac{P_s}{N_0 W}\right) \quad (\text{比特} / N \text{ 个自由度})$$

又因为在单位时间 $(1/T)$ 内采集到 $2W$ 个信道符号,归一化后得到带限 AWGN 信道单位时间信道容量为

$$C_t = \frac{1}{T} C^N = W \log\left(1 + \frac{P_s}{N_0 W}\right) = W \log\left(1 + \frac{P_s}{P_n}\right) \quad (\text{b/s}) \quad (3.5.5)$$

这就是著名的香农信道容量公式(简称**香农公式**)。式中的对数取以 2 为底。它给出了频带受限、平均功率受限的 AWGN 信道的信道容量计算方法。当信道输入信号是平均功率受限的高斯分布信号时,信息传输率才达到此信道容量。

例 3.5.1 给定某平均功率受限的 AWGN 信道,带宽为 4MHz。

- (1) 若输入信号与噪声的平均功率比 $\left(\frac{P_s}{P_n} \triangleq \text{SNR}\right)$ 为 7, 求其信道容量。
- (2) 若 SNR 降为 3, 带宽需调整为多少, 才能达到与(1)相同的信道容量?
- (3) 若带宽为 3MHz, 求达到与(1)同样信道容量的 SNR。

解:

$$C_t = W \log(1 + \text{SNR}) = 4 \times 10^6 \log(1 + 7) = 12 \text{ (Mb/s)}$$

$$W = \frac{C_t}{\log(1 + \text{SNR})} = \frac{12 \times 10^6}{\log(1 + 3)} = 6 \text{ (MHz)}$$

$$\text{SNR} = 2^{C_t/W} - 1 = 2^{12/3} - 1 = 15$$

一些实际信道是非高斯波形信道,对于平均功率受限非高斯加性信道,其信道容量的精确计算一般较困难。但根据平均功率受限条件下的最大熵定理,加性信道在噪声服从高斯分布时噪声熵最大,即高斯加性信道是平均功率受限条件下的最差信道。所以,香农公式可适用于其他波形信道,由其得到的值是加性非高斯波形信道的信道容量的下限值。

3.5.2 香农公式的意义与启发

香农公式是平均功率受限下的带限 AWGN 信道的信道容量计算公式,从式中可以看出,香农公式把信道的统计信息参量(信道容量)和信道的实际物理量(频带宽度 W 、信噪功率比 P_s/P_n 等)联系起来。它表明带限 AWGN 信道的信道容量 C_t 与带宽 W 、信号平均功率 P_s 以及噪声功率谱密度 N_0 有关。香农公式也建立了带宽 W 、信噪比 P_s/P_n 、信道容量 C_t 三者之间的制约关系。下面分析在不同情况下由香农公式得到的一些重要结论,反映了通信系统优化设计中的一般规律。

1. 增大信道容量的一般方法

(1) 信噪比不变,扩展频带宽度 W 也可增大信道容量 C_t

由式(3.5.5)可见,信道容量与所传输信号的有效带宽成正比,信号的有效带宽越宽,信道容量越大。如通信系统构建中不断开发新的传输媒体,有线通信中从明线(150kHz)、对称电缆(600kHz)、同轴电缆(1GHz)到光纤(25THz);无线由中波、短波、超短波到毫米波、微米波。又如,采用信道均衡技术开辟可传输的频带等。

但需要注意的是,噪声功率 $P_n = N_0 W$,所以随着带宽的增大,噪声功率也会变大,信噪比随之减小,又使信道容量变小。因此,增加信道带宽 W (也就是信号的带宽),并不能无限制地使信道容量增大。也就是说,当发送信号功率 P_s 不变时,增加带宽可以在一定程度上增大信道容量。但随着带宽增大趋于无穷时,信道容量将趋于某一极限,如图 3.5.3 所示。具体推导如下:

$$\begin{aligned}\lim_{W \rightarrow \infty} C_t &= \lim_{W \rightarrow \infty} W \log \left(1 + \frac{P_s}{N_0 W} \right) \\ &= \frac{P_s}{N_0} \lim_{W \rightarrow \infty} W \frac{N_0}{P_s} \log \left(1 + \frac{P_s}{N_0 W} \right)\end{aligned}$$

令

$$x = \frac{P_s}{N_0 W}$$

则有

$$\lim_{W \rightarrow \infty} C_t = \frac{P_s}{N_0} \lim_{x \rightarrow 0} \frac{1}{x} \log(1+x)$$

又

$$\lim_{x \rightarrow 0} \frac{1}{x} \ln(x+1) = 1$$

所以有

$$\lim_{W \rightarrow \infty} C_t = \frac{P_s}{N_0} \log_e \approx 1.44 \frac{P_s}{N_0} \quad (3.5.6)$$

此式说明当频带很宽,或信噪比很低时,信道容量正比于信号功率与噪声功率密度比。此值是加性高斯白噪声信道在无限带宽时的信道容量,它是信息传输率的极限值。

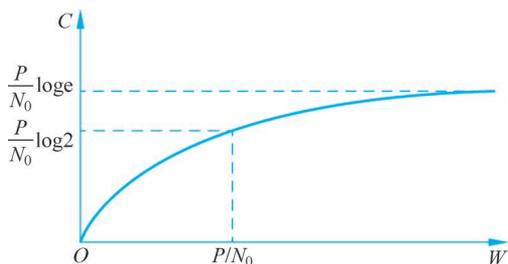
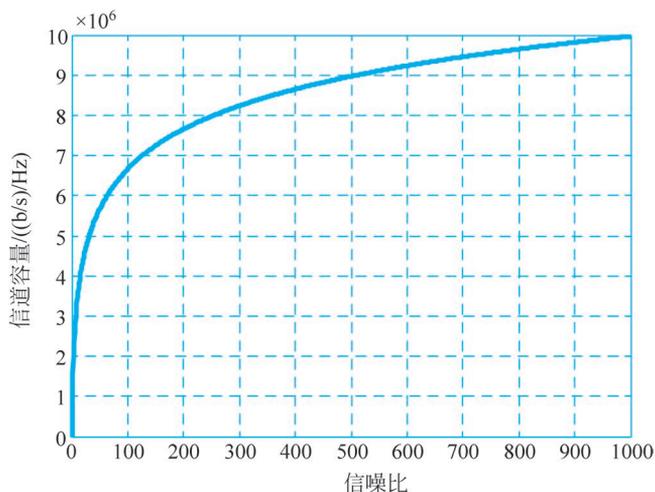


图 3.5.3 加性高斯白噪声信道容量与带宽关系

(2) 带宽不变,增加信号功率或者提高信噪比,可使 C_t 增大

信道容量与信道上的信号噪声比有关,信噪比越大,信道容量也越大,但其制约规律呈对数关系。通信工程中很多技术手段目的都是提升系统信噪比,获得大的信道容量。例如,通过提高发送功率、提高天线增益、将无方向的漫射改为方向性强的波束或点波束、采用分集接收等加大信号功率,采用低噪声器件、滤波、屏蔽、接地、低温运行等技术降低噪声功率。

通过增大信号发送功率来增大信道容量在理论上没有上限,如果能使信道输入功率无穷大,那么 AWGN 信道的容量将为无穷大。但随着信号功率的增加, $\lim_{P_s \rightarrow \infty} \frac{dC_t}{dP_s} \rightarrow 0$, 信道容量增长将越来越慢,如图 3.5.4 所示。

图 3.5.4 加性高斯白噪声信道容量与信噪比关系(带宽 $W=1\text{MHz}$)

2. C_t 一定时,带宽与信噪比可相互补偿

香农公式把信道的统计信息参量(信道容量)和信道的实际物理量(频带宽度、信噪功率比等)联系起来。它从理论上阐明一个理想通信系统的最大可靠信息传输率可完全由带宽 W 、传输时间 T 和信噪功率比 P_s/P_n 三个物理量确定。在信道容量要求一定时, W 、 T 、SNR 三者之间可以互相转换,互相补偿。虽然并未提出具体解决的实际方法,但

是指出了努力的方向。

(1) 若传输时间 T 固定, 则扩展信道带宽 W 就可以降低对信噪比的要求; 反之, 带宽变窄, 就要增加信噪比。也就是说, 可以通过带宽和信噪比的互换而保持信息传输能力不变。

将例 3.5.1 中在信道容量 $C_t = 12\text{Mb/s}$ 时, 所需带宽 W 、信噪比 SNR 以及输入信号功率 P_s 的情况列于表 3.5.1。由表可以看出: 从情况一到情况二, 带宽减少了 25%, 信噪比需提升 1.1 倍, 输入信号功率必须增加约 61%; 从情况一到情况三, 带宽增大 50%, 信噪比可降低 57%, 输入信号功率随之降低约 36%。带宽很小地改变, 信噪比或信号功率就有较大的改变。若增加较少的带宽, 就能节省较大的信号功率, 这在深空通信中是很重要的。

表 3.5.1 例 3.5.1 中在 $C_t = 12\text{Mb/s}$ 时的 W 、SNR 及 P_s

项目	带宽 W/MHz	信噪比 SNR	输入信号功率 P_s
情况一	4	7	$28 \times 10^6 N_0$
情况二	3	15	$45 \times 10^6 N_0$
情况三	6	3	$18 \times 10^6 N_0$

注: $P_s = \text{SNR} \cdot W \cdot N_0$ 。

(2) 如果信噪功率比固定不变, 增加信道的带宽 W 就可以缩短传送时间 T , 换取传输时间的节省, 或者花费较长的传输时间来换取频带的节省, 也就是实现频带和通信时间的互换。

例如, 为了能在窄带电缆信道中传送电视信号, 往往用增加传送时间的办法来压缩电视信号的带宽, 首先把电视信号高速记录在录像带上, 然后慢放这个磁带, 慢到使输出频率降低到足以在窄带电缆信道中传送的程度。在接收端, 将接收到的慢录像信号进行快放, 于是恢复了原来的电视信号。

(3) 如果保持频带不变, 那么可以采用增加时间 T 来改善信噪比。

这一原理已被应用于弱信号接收技术中, 即所谓积累法。这种方法是将重复多次收到的信号叠加起来。由于有用信号直接相加, 而干扰则是按功率相加(噪声是随机的, 叠加后会相互抵消; 而信号是相关的, 叠加会增强), 因而经积累相加后, 信噪比得到改善, 但所需接收时间相应增加。如何换取, 一般根据实际情况而决定。例如, 宇宙飞船与地面通信由于信号与噪声的功率比很小, 故着重考虑增加带宽和增加传输时间来换取对信噪比的要求。倘若信道频带十分紧张, 则考虑提高信噪比或传输时间来降低对带宽的要求。

3. 香农公式对极低信噪比通信的启发

由香农公式可以发现, 当发送信噪比小于 1 时, 信道的信道容量并不等于 0, 这说明此时信道仍具有传输信息的能力。也就是说, 发送信号功率很弱, 即便是淹没在噪声中仍有可能进行可靠的通信, 这对于卫星通信、深空通信等具有特别重要的意义。同时, 该结论也为极低信噪比下的信号检测、隐蔽通信等技术方向奠定理论基础。

4. E_b/N_0 与香农限

香农公式给出的信道容量是连续信道中可靠传输的最大信息传输率。是否可以用无限制加大信号有效带宽的方法来减小发射功率,或在任意低的信噪比情况下仍能实现可靠通信?从香农公式中还可以找出了达到无错误(无失真)通信的传输速率的理论极限值,称为香农极限。

若以最大信息速率,即信道容量($C_t = \max R_t$)来传输信息,又令每传送 1bit 信息所需的能量为 E_b ,得总的信号功率 $P_s = R_t E_b$ (当信息传输速率达最大时 $P_s = C_t E_b$)。代入式(3.5.6)可得

$$\lim_{w \rightarrow \infty} C_t = \frac{P_s}{N_0} \log e = \frac{R_t E_b}{N_0} \log e$$

式中: E_b/N_0 表示单位频带内传输 1bit 信息的信噪比,称为归一化信噪比,且有

$$\frac{E_b}{N_0} = \frac{\lim_{w \rightarrow \infty} C_t}{R_t \log e}$$

由带宽与信噪比的互换关系可知, E_b/N_0 的最小值发生在带宽趋于无穷大时,且令此时的信息传输速率达最大,即 $R_t = C_t$,则有

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{1}{\log e} = \ln 2 = -1.6(\text{dB})$$

这个值称为**香农限**,它表明可靠传输 1bit 信息所需要的最小能量为 $0.693N_0$ 。香农限(-1.6dB)是在带宽趋于无穷大时达到的,是在理论上能实现可靠通信的 E_b/N_0 的最小值。

图 3.5.5 示出了 E_b/N_0 与归一化信息传输速率 R_t/W 之间的关系。所以在实际通信系统的评估与分析中常用此香农限来衡量实际系统的潜力,以及各种纠错编码性能的好坏。如何达到和接近这个理论极限香农并没有给出具体方案,而这正是通信研究人员所面临的任务。

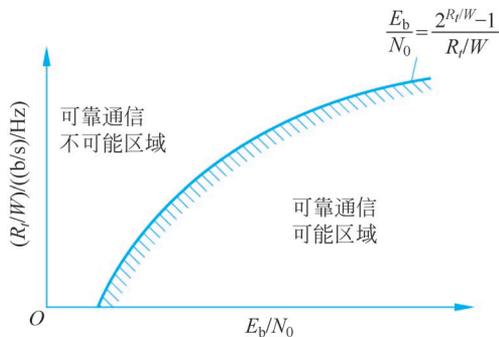


图 3.5.5 E_b/N_0 与 R_t/W 的关系

小结

平均互信息:

$$I(X; Y) = H(X) - H(X | Y) \text{ (比特/符号)}$$

代表了接收到每个输出符号 Y 后获得的关于 X 的平均信息量。

$$I(X; Y) = \sum_{XY} P(xy) \log \frac{p(x|y)}{P(x)}$$

$$I(X; Y) = \sum_{XY} P(xy) \log \frac{P(xy)}{P(x)P(y)}$$

$$I(X; Y) = \sum_{XY} P(xy) \log \frac{P(y|x)}{P(y)}$$

互信息:

$$I(x; y) = \log \frac{1}{P(x)} - \log \frac{1}{P(x|y)} = \log \frac{P(x|y)}{P(x)}$$

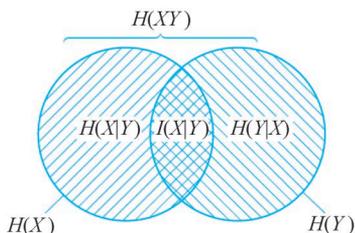
平均互信息的性质:

- (1) 非负性: $I(X; Y) \geq 0$, 当 X 与 Y 统计独立时等号成立。
- (2) 极值性: $I(X; Y) \leq H(X)$ 。
- (3) 对称性: $I(X; Y) = I(Y; X)$ 。
- (4) 与各熵的关系:

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

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

$$I(X; Y) = H(X) + H(Y) - H(XY)$$



- (5) 凸函数特性: $I(X; Y)$ 是 $P(x)$ 的 \cap 型凸函数, $I(X; Y)$ 是 $P(y|x)$ 的 \cup 型凸函数。

序列形式的平均互信息:

$$I(\mathbf{X}; \mathbf{Y}) = I(X_1 X_2 \cdots X_N; Y_1 Y_2 \cdots Y_N) = H(\mathbf{X}) - H(\mathbf{X} | \mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y} | \mathbf{X})$$

- (1) 若信道是无记忆的, 则存在

$$I(\mathbf{X}; \mathbf{Y}) \leq \sum_{i=1}^N I(X_i; Y_i)$$

- (2) 若信源是无记忆的, 则存在

$$I(\mathbf{X}; \mathbf{Y}) \geq \sum_{i=1}^N I(X_i; Y_i)$$

- (3) 若信源和信道都是无记忆的, 则存在

$$I(\mathbf{X}; \mathbf{Y}) = \sum_{i=1}^N I(X_i; Y_i)$$

信道容量:

$$\max_{P(x)} \{I(\mathbf{X}; \mathbf{Y})\} \triangleq C$$

(1) 无噪无损信道:

$$C = \max_{P(x)} \{I(\mathbf{X}; \mathbf{Y})\} = \max_{P(x)} \{H(X)\} = \log r$$

(2) 有噪无损信道:

$$C = \max_{P(x)} \{I(\mathbf{X}; \mathbf{Y})\} = \max_{P(x)} \{H(X)\} = \log r$$

(3) 无噪有损信道:

$$C = \max_{P(x)} \{I(\mathbf{X}; \mathbf{Y})\} = \max_{P(x)} \{H(Y)\} = \log s$$

(4) 对称离散信道:

$$C = \log s - H(p'_1, p'_2, \dots, p'_s)$$

式中: $\{p'_1, p'_2, \dots, p'_s\}$ 为信道矩阵中的行矢量。

(5) 离散无记忆的 N 次扩展信道:

$$C^{(N)} = \max_{P(x)} I(\mathbf{X}; \mathbf{Y}) = \max_{P(x)} \sum_{i=1}^N I(X_i; Y_i) = \sum_{i=1}^N C_i = NC$$

(6) 输入并接信道:

$$C \leq \max_{P(x)} H(X)$$

(7) 并用信道:

$$C = \max \sum_{i=1}^N I(X_i; Y_i) = \sum_{i=1}^N \max I(X_i; Y_i) = \sum_{i=1}^N C_i$$

(8) 和信道:

$$C = \log_2 \sum_{i=1}^N 2^{C_i}$$

最佳信道利用率为

$$P_i(C) = 2^{(C_i - C)}$$

数据处理定理: 对于马尔可夫链 $X \rightarrow Y \rightarrow Z$, 有

$$I(X; Z) \leq I(X; Y), \quad I(X; Z) \leq I(Y; Z)$$

串联信道的容量:

$$C \leq \min\{C_1, C_2, \dots, C_N\}$$

式中: $C_i (i=1, 2, \dots, N)$ 为各子信道的容量。

连续分布随机变量之间的平均互信息:

$$I(X; Y) = \iint_{\mathbf{R}} p(xy) \log \frac{p(x|y)}{p(x)} dx dy = h(X) - h(X|Y)$$

$$I(X; Y) = \iint_{\mathbf{R}} p(xy) \log \frac{p(y|x)}{p(y)} dx dy = h(Y) - h(Y|X)$$

$$I(X; Y) = \iint_{\mathbf{R}} p(xy) \log \frac{p(xy)}{p(x)p(y)} dx dy = h(X) + h(Y) - h(XY)$$

单符号连续信道的信道容量:

$$C = \max_{p(x)} I(X; Y) = \max_{p(x)} [h(Y) - h(Y|X)] \text{ (比特 / 自由度)}$$

单符号加性高斯信道的信道容量:

$$\begin{aligned} C &= \frac{1}{2} \log(2\pi e P_o) - \frac{1}{2} \log(2\pi e \sigma^2) = \frac{1}{2} \log\left(\frac{P_o}{\sigma^2}\right) \\ &= \frac{1}{2} \log\left(1 + \frac{P_s}{\sigma^2}\right) \\ &= \frac{1}{2} \log\left(1 + \frac{P_s}{P_n}\right) \text{ (比特 / 自由度)} \end{aligned}$$

式中: P_o 为输出信号功率; σ^2 为噪声方差; P_s 为发送信号功率; P_n 为噪声功率。

多维连续信道的平均互信息及信道容量:

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{X}) - h(\mathbf{X} | \mathbf{Y}) = \iiint_{\mathbf{R}} p(\mathbf{xy}) \log \frac{p(\mathbf{x} | \mathbf{y})}{p(\mathbf{x})} dx dy$$

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{Y} | \mathbf{X}) = \iiint_{\mathbf{R}} p(\mathbf{xy}) \log \frac{p(\mathbf{y} | \mathbf{x})}{p(\mathbf{y})} dx dy$$

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{X}) + h(\mathbf{Y}) - h(\mathbf{XY}) = \iiint_{\mathbf{R}} p(\mathbf{xy}) \log \frac{p(\mathbf{xy})}{p(\mathbf{x})p(\mathbf{y})} dx dy$$

$$C = \max_{p(x)} I(\mathbf{X}; \mathbf{Y}) = \max_{p(x)} [h(\mathbf{Y}) - h(\mathbf{Y} | \mathbf{X})]$$

多维加性信道的信道容量:

$$C = \max_{p(x)} I(\mathbf{X}; \mathbf{Y}) = \max_{p(x)} [h(\mathbf{Y}) - h(\mathbf{n})] \text{ (比特 / 自由度)}$$

一般加性波形信道单位时间的信道容量:

$$C_t = \max_{p(x)} \left\{ \frac{1}{T} \lim_{N \rightarrow \infty} [h(\mathbf{Y}) - h(\mathbf{n})] \right\} \text{ (b/s)}$$

带限高斯白噪声加性信道单位时间信道容量(香农公式):

$$C_t = \frac{1}{T} C^N = W \log\left(1 + \frac{P_s}{N_0 W}\right) = W \log\left(1 + \frac{P_s}{P_n}\right) \text{ (b/s)}$$

式中: P_s 为发送信号功率; P_n 为噪声功率; $P_n = N_0 W$; N_0 为单边带功率谱密度。

加性高斯白噪声信道在无限带宽时的信道容量:

$$\lim_{W \rightarrow \infty} C_t = \frac{P_s}{N_0} \log e \approx 1.44 \frac{P_s}{N_0}$$

香农限:

理论上能实现可靠通信的 E_b/N_0 的最小值,即

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{1}{\log e} = \ln 2 = -1.6 \text{ (dB)}$$

习题

3.1 设有一离散无记忆信源,其概率空间为

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0.6 & 0.4 \end{bmatrix}$$

它们通过一干扰信道,信道输出端的接收符号集 $Y = [0 \ 1]$,信道矩阵为

$$P = \begin{bmatrix} p(0/0) & p(1/0) \\ p(0/1) & p(1/1) \end{bmatrix} = \begin{bmatrix} 5/6 & 1/6 \\ 3/4 & 1/4 \end{bmatrix}$$

- 试求: (1) 信源 X 中事件 X_1 和 X_2 分别含有的自信息。
 (2) 收到消息 y_j ($j=1,2$) 后,获得的关于 x_i ($i=1,2$) 的信息量。
 (3) 输出符号集 Y 的平均信息量 $H(Y)$ 。
 (4) 信道疑义度 $H(X|Y)$ 及噪声熵 $H(Y|X)$ 。
 (5) 接收到消息 Y 后获得的平均互信息。

3.2 设有扰离散信道的输入端是以等概率出现的 A, B, C, D 四个字母。该信道的正确传输概率为 $1/2$, 错误传输概率均匀分布在其他三个字母上。验证在该信道上每个字母传输的平均信息量为 0.21bit 。

3.3 设有下述消息将通过一个有噪二元对称信道传送,消息为 $M_1=00, M_2=01, M_3=10, M_4=11$,这4种消息在发送端是等概的。试求:

- (1) 输入为 M_1 , 输出第一个数字为 0 的互信息量是多少?
 (2) 如果第二个数字也是 0, 这时又带来多少附加信息?

3.4 通过某二元无损信道传输一个由字母 A, B, C, D 组成的符号集,把每个字母编码成两个二元码脉冲序列,以 00 代表 $A, 01$ 代表 $B, 10$ 代表 $C, 11$ 代表 D ,每个二元码元脉冲宽度为 5ms 。

(1) 不同字母等概出现时,试计算传输的平均信息速率。

(2) 若每个字母出现的概率分别为 $P_A = \frac{1}{5}, P_B = \frac{1}{4}, P_C = \frac{1}{4}, P_D = \frac{3}{10}$, 试计算传输的平均信息速率。

3.5 设有一批电阻,按阻值分 70% 是 $2\text{k}\Omega$, 30% 是 $5\text{k}\Omega$; 按功耗分 64% 是 $1/8\text{W}$, 其余是 $1/4\text{W}$ 。已知 $2\text{k}\Omega$ 阻值的电阻中 80% 是 $1/8\text{W}$, 试问通过测量阻值可以平均得到的关于功耗的信息量是多少?

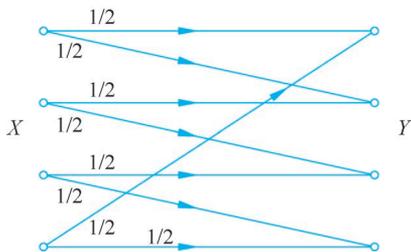
3.6 设 BSC 信道矩阵为

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

- (1) 若 $P(0)=3/4, P(1)=1/4$, 求 $H(X), H(X|Y), H(Y|X)$ 及 $I(X; Y)$ 。
 (2) 求该信道容量及其最佳输入分布。

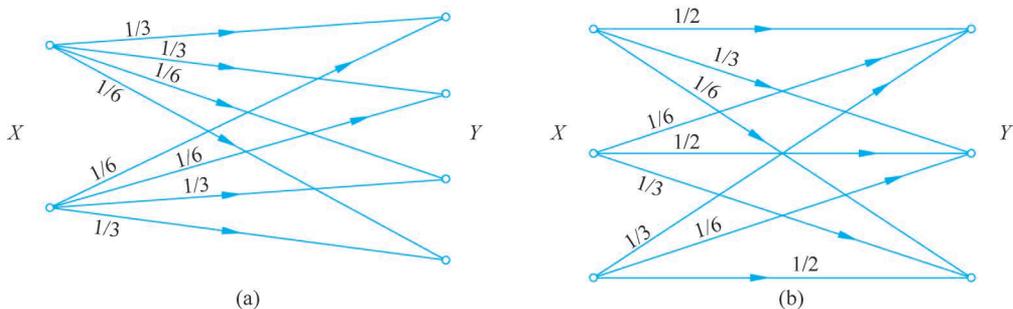
3.7 在有扰离散信道上传输符号 0 和 1, 在传输过程中每 100 个符号发生一个错误, 已知 $P(0)=P(1)=1/2$, 信源每秒内发出 1000 个符号, 求此信道的信道容量。

3.8 设有扰离散信道如题 3.8 图所示, 试求此信道的信道容量及最佳输入分布。



题 3.8 图

3.9 求题 3.9 图所示信道的信道容量及其最佳输入概率分布。



题 3.9 图

3.10 BSC 信道矩阵为

$$\mathbf{P} = \begin{bmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{bmatrix}$$

设该信道以每秒 1500 个二元符号的速度传输输入符号, 现有一消息序列共有 14000 个二元符号, 并设在这消息中 $P(0)=P(1)=1/2$, 试问从信息传输的角度来考虑, 10s 内能否将这消息序列无失真地传送完。

3.11 设一离散无记忆信道的信道矩阵为

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

(1) 计算信道容量 C 。

(2) 找出一个长度为 2 的码, 其信息传输率为 $\frac{1}{2} \log 5$ (5 个码字) 如果按最大似然译码规则设计译码器, 求译码器输出的平均错误概率 P_E (输入码字等概条件下)。

3.12 如果一个统计人员面对转移概率为 $p(y|x)$ 且信道容量 $C = \max_{p(x)} I(X; Y)$ 的通信信道, 会对输出做出很有帮助的预处理 $\hat{Y} = g(Y)$, 并且断定这样做能够严格地改进容量。

(1) 证明他错了。

(2) 在什么条件下他不会严格地减小容量?

3.13 考虑时变离散无记忆信道。设 Y_1, Y_2, \dots, Y_N 在已知 X_1, X_2, \dots, X_N 的条件下是条件独立的, 并且条件概率分布

$$p(\mathbf{y} | \mathbf{x}) = \prod_{i=1}^N p(y_i | x_i).$$

设 $\mathbf{X} = (X_1, X_2, \dots, X_N), \mathbf{Y} = (Y_1, Y_2, \dots, Y_N)$ 。求 $C = \max_{p(x)} I(X; Y)$ 。

3.14 求下列两个信道的信道容量, 并加以比较。

$$(1) \begin{bmatrix} \bar{p} - \epsilon & p - \epsilon & 2\epsilon \\ p - \epsilon & \bar{p} - \epsilon & 2\epsilon \end{bmatrix}$$

$$(2) \begin{bmatrix} \bar{p} - \epsilon & p - \epsilon & 2\epsilon & 0 \\ p - \epsilon & \bar{p} - \epsilon & 0 & 2\epsilon \end{bmatrix}$$

其中 $\bar{p} + p = 1$ 。

3.15 试计算下述信道的信道容量 (以 p, q 作为变量):

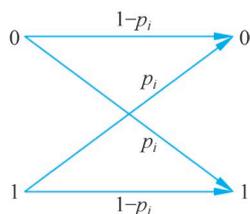
$$\mathbf{P} = \begin{bmatrix} \bar{p} & p & 0 & 0 \\ p & \bar{p} & 0 & 0 \\ 0 & 0 & \bar{q} & q \\ 0 & 0 & q & \bar{q} \end{bmatrix}$$

3.16 两个串联信道如图 3.3.14 所示, 其输入、输出分别为 X, Y, Z , 并且满足 XYZ 是马尔可夫链。试证明若第二信道是无损信道, 即 $H(Y|Z) = 0$, 则得到串联信道 $I(X; Z) = I(X; Y)$ 。

3.17 把 n 个 BSC(p) 串联起来, 证明该串联信道可以等效为一个 BSC, 其错误转移概率为 $\frac{1}{2} [1 - (1 - 2p)^n]$, 且 $\lim_{n \rightarrow \infty} I(X_0; X_n) = 0 (p \neq 0, 1)$ 。

3.18 若有两个串联的离散信道, 它们的信道矩阵都为

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



题 13 图

并设第一个信道的输入符号 $X \in \{a_1, a_2, a_3, a_4\}$ 是等概分布, 求 $I(X; Z)$ 和 $I(X; Y)$ 并加以比较。

3.19 设两连续随机变量 X 和 Y , 它们的联合概率密度为均值为零、协方差矩阵为 $\mathbf{R} = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix}$ 的正态分布, 在 ρ 分别为 0、1、-1 的情况下, 计算 $I(X; Y)$ 。

3.20 一对独立并列信道, $Y_1 = X_1 + n_1, Y_2 = X_2 + n_2$, 其中 n_1, n_2 是均值为零、协方差矩阵为 $\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$ 的正态分布随机变量, 并且信号平均功率均受限于 $2P$, 又假设 $\sigma_1^2 > \sigma_2^2$ 。试求:

- (1) 信号平均功率受限值 P 取什么值时相当于只有一个噪声方差为 σ_2^2 的信道。
- (2) 要使这一对信道都工作, P 该取什么值。

3.21 设某一信号的信息传输率为 5.6kb/s, 噪声功率谱 $N = 5 \times 10^{-6}$ mW/Hz, 在带限 $B = 4$ kHz 的高斯信道中传输, 无差错传输需要的最小输入功率 P 是多少?

3.22 一个平均功率受限制的连续信道, 其通频带为 1MHz, 信道上存在白色高斯噪声。

- (1) 已知信道上的信号与噪声的平均功率之比为 10, 求该信道的信道容量。
- (2) 若信道上的信号与噪声的平均功率之比降至 5, 要达到相同的信道容量, 信道通频带应该多大?
- (3) 若信道通频带减少为 0.5MHz, 要保持相同的信道容量, 信道上的信号与噪声的平均功率比值应该多大?

3.23 阅读香农著作 *A Mathematical Theory of Communication* 的 PART II: The Discrete Channel With Noise 中的第 12、13 部分, 对照本章讲述的基本原理, 以自己的理解写一篇文献学习报告, 并结合通信技术对理论知识进行一定拓展。

3.24 查阅有关 MIMO 信道容量的文献, 完成一篇综合论述报告。

3.25 查阅相关资料, 完成一篇关于香农公式指导通信技术发展的论文。