

基础信息论

连续信源的信息率失真函数

华中科技大学电信学院

连续信源的信息率失真函数

- 一般情况下，信息在传输过程中必然会存在一定的噪声和干扰，使得信源的消息在传输过程中存在一定的误差和失真。
- 对于连续信源而言，在传输过程中总会有波形失真，连续信源的信息率失真理论就是在一定意义上定量分析信号的失真程度。
- 本节主要讨论连续信源的信息率失真函数，连续信源的率失真理论与离散信源情况基本相同。

平均失真度定义

- 定义： 设连续信源（随机变量） X , 其概率密度为 $p_X(x)$, 设另一变量 Y , 且 X 和 Y 之间失真函数是某一非负的二元函数 $d(x, y)$, 则平均失真度定义为

$$\begin{aligned}\bar{D} &= E\{d(X, Y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(xy) d(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{p_X(x)} \mathbf{p(y|x)} d(x, y) dx dy\end{aligned}$$

式中 $p(y|x)$ 为信道特征, 满足 $\int_{-\infty}^{\infty} p(y|x) dy = 1$

连续信源的信息率失真函数相关定义

定义： 设所有试验信道的集合为 P_D ,在满足一定失真度 $\bar{D} \leq D$ 时, 连续信源的信息率失真函数为

$$R(D) = \inf_{p(y|x) \in P_D} I(X, Y)$$

式中 \inf 表示下界, 试验集合为 $P_D : \{p(y|x), \bar{D} \leq D\}$

分析：连续信源的信息率失真函数具有离散信源的信息率失真函数的性质。

由于 $R(D)$ 的求解是一个求极值问题, 并且 $p(y|x)$ 是一个二元函数, 计算 $R(D)$ 非常复杂。

不失一般的讨论均方误差失真准则下高斯信源的 $R(D)$ 。

高斯信源的信息率失真函数

高斯信源的信息率失真函数推导

设高斯信源 X 的均值为 m_X 、方差为 σ_X^2 定义失真函数为平均误差失真。

求上述条件下的 $R(D)$ 实际上是求在条件

$$D = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(xy)(x - y)^2 dx dy$$

保真度准则

$$= \int_{-\infty}^{\infty} p(x) \int_{-\infty}^{\infty} (x - y)^2 p(y | x) dy dx$$

和 $\int_{-\infty}^{\infty} p(y | x) dy = 1$ 下,

信道转移概率的归一化

对于 $p(y | x)$ 为某一分布时交互信息量 $I(X; Y)$ 的极值,

即 $R(D) = \min_{p(y|x)} I(X; Y)$

连续信源的信息率失真函数求法

■ 有两种方法可求（略）：

1. 应用拉格朗日算子，与离散的算法类似。
2. 用反向信道。

■ 其结果是：
$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma_X^2}{D} & D \leq \sigma_X^2 \\ 0 & D > \sigma_X^2 \end{cases}$$

当信源均值不为0时，仍有这个结果，因为高斯信源的熵只与随机变量的方差有关，与均值无关。

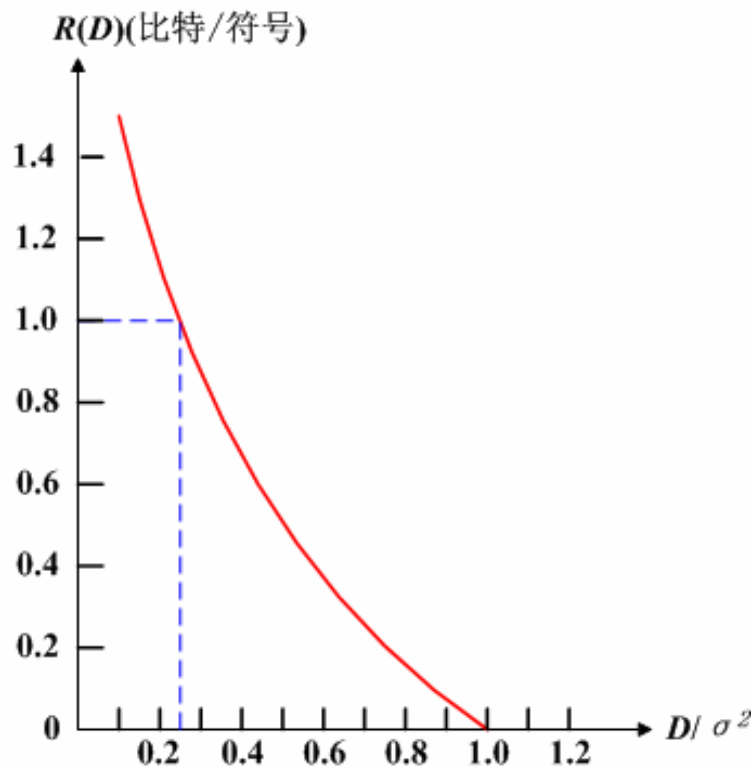


图4.3.2 高斯信源在均方误差准则下的 $R(D)$ 函数

连续信源的信息率失真函数曲线

✓当 $D=\sigma^2$ 时, $R(D)=0$: 这说明若允许失真 (均方误差) 等于信源方差, 只需用均值 m 来表示信源的输出, 不需要传送信源的任何实际输出;

✓当 $D=0$ 时, $R(D)\rightarrow\infty$: 这点说明连续信源要毫无失真地传送信源的输出是不可能的。即毫无失真地传送信源的输出必须要求信道具有无限大的容量;

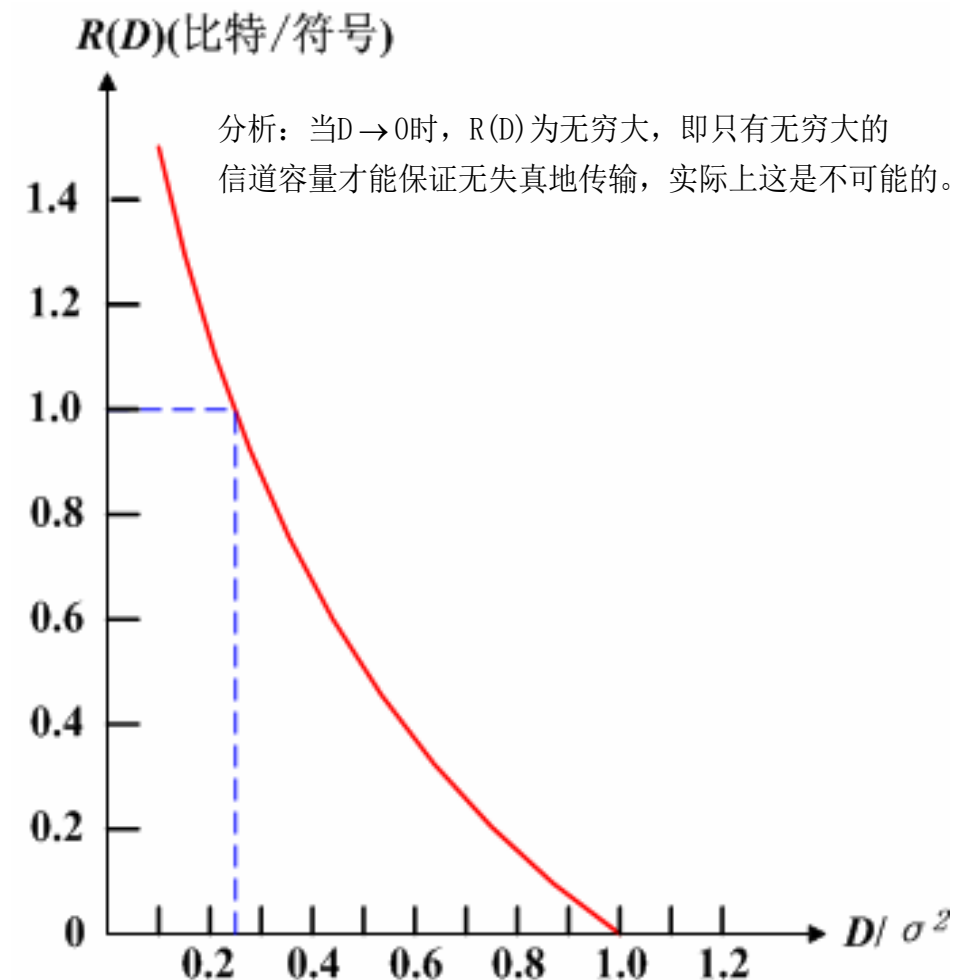


图4.3.2 高斯信源在均方误差准则下的 $R(D)$ 函数

连续信源的信息率失真函数曲线

✓当 $0 < D < \sigma^2$ 时：允许一定失真，则传送信息率可降低，意味着信源可压缩，连续信源的率失真理论正是连续信源量化、压缩的理论基础。

✓当 $D = 0.25\sigma^2$ 时， $R(D) = 1$ 比特/符号：这说明在允许均方误差小于或等于 $0.25\sigma^2$ 时，连续信号的每个样本值最少需用一个二进制符号来传输。由香农第三定理证明了这种压缩编码是存在的，然而实际上要找到这种可实现的最佳编码方法很困难。

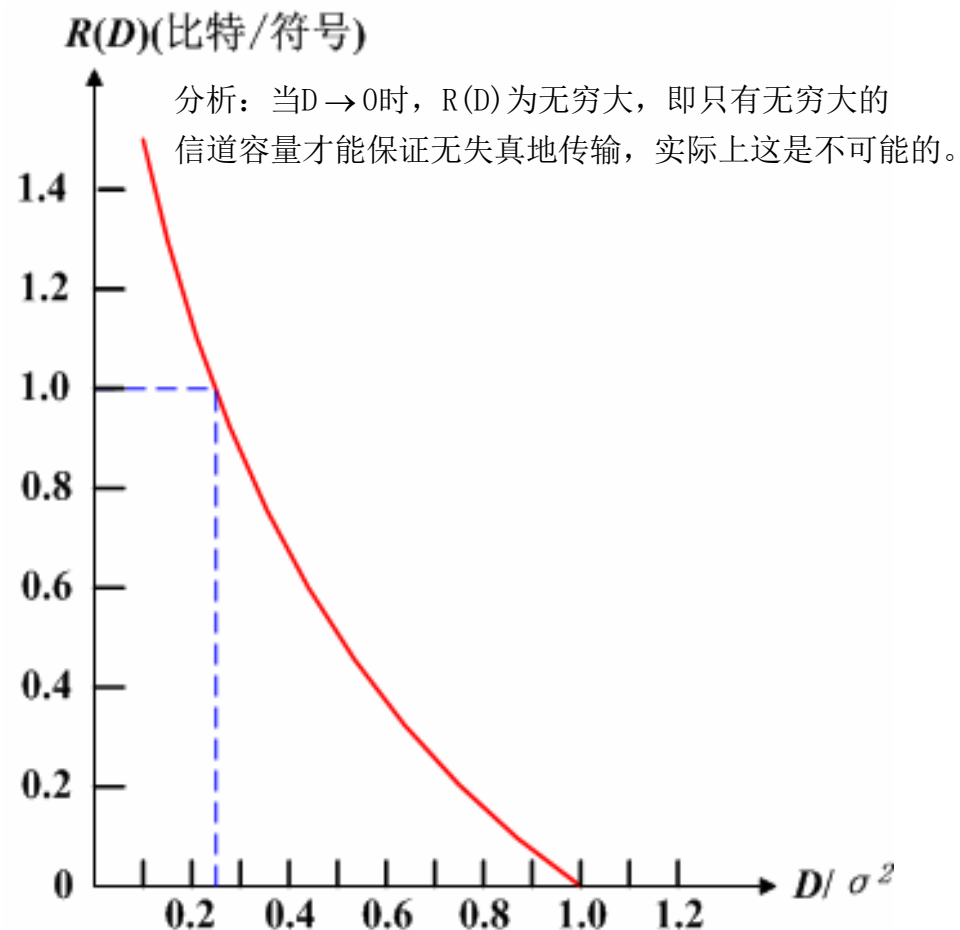


图4.3.2 高斯信源在均方误差准则下的 $R(D)$ 函数

谢谢!

黑晓军

华中科技大学

电子信息与通信学院

Email: heixj@hust.edu.cn

网址: <http://eic.hust.edu.cn/aprofessor/heixiaojun>

参考资料

- 陈运, 信息论与编码, 第3版, 第7章7.3节, 电子工业出版社, 2015