限失真信源编码简介

信息论与编码原理不挂科 第六讲

限失真信源编码简介

信息论与编码原理不挂科 第六讲

4大模块

2道题目

信息论与编码原理不挂科 第六讲

限失真信源编码简介

模块1 限失真信源编码的基本概念

模块2 失真度和平均失真度

模块3 信息率失真函数

模块4 限失真信源编码定理

限失真信源编码的基本概念

小节1 限失真信源编码概述

小节2 限失真信源编码的研究方法

限失真信源编码的基本概念

小节1 限失真信源编码概述

小节2 限失真信源编码的研究方法

通信系统中失真存在的原因

1. 信道噪声的干扰使得信息传输过程会产生差错。
2. 信息传输率超过信道容量时，必然产生差错。
3. 信源熵是信源无失真压缩的极限，若再继续压缩则会带来失真。

通信系统中失真存在的合理性

1. 信宿的灵敏度和分辨率有限，不要求绝对无失真。
2. 允许失真的存在，可以提高信息传输率，从而降低通信成本。

本讲的研究内容，为**在允许一定程度的失真的条件下，能够把信源信息压缩到什么程度**，即**最少**需要多少 bit 才能描述信源。

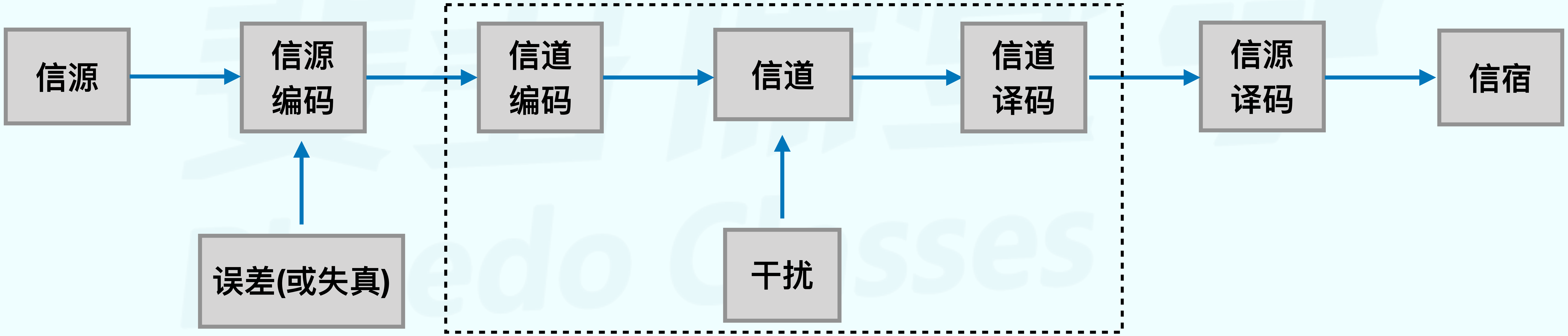
限失真信源编码的基本概念

小节1 限失真信源编码概述

小节2 限失真信源编码的研究方法

限失真信源编码的研究方法

用研究信道的方法来研究有失真信源压缩问题。



只由信源编码产生

广义无干扰信道，可省略

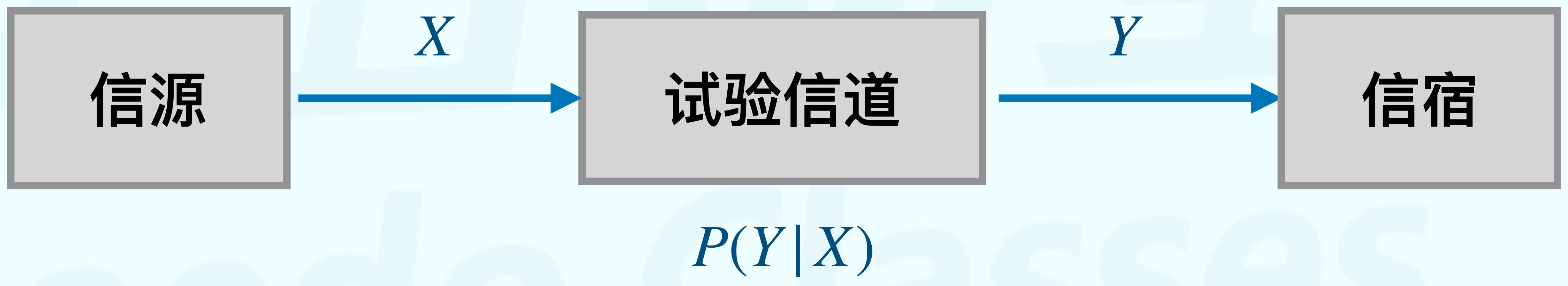
允许失真越大，信息传输率可越小；
允许失真越小，信息传输率须越大。

即R与失真有关。

限失真信源编码的研究方法

用虚拟手法用信道表示信源编码的作用，将信源编码和信源译码等价为一个信道。

由于是失真编码，信道不是一一对应的，用传递概率描述编译码前后的关系，可将通信系统简化为下图：



选择不同的试验信道，相当于不同的编码方法，其所得的平均失真度 D 不同，有的满足 $D \le D^*$，有些则 $D > D^*$。

我们要研究在**给定允许失真**的条件下，是否可以设计一种信源编码使得**信息传输率R最低**，首先需讨论失真测度。

失真度和平均失真度

小节1 失真函数

小节2 平均失真度

失真度和平均失真度

小节1 失真函数

小节2 平均失真度

失真函数 $d(x, y)$

失真函数为非负函数，即 $d(x_i, y_j) \geq 0$，函数形式可根据需要定义。

用失真函数定量测度信源发生一个符号 x_i，而在接收端再接收 y_j 所引起的误差或失真。

通常较小的 d 表示较小的失真，而 $d(x_i, y_j) = 0$ 表示无失真。

假设信源分布为 $\begin{bmatrix} X \\ p \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ p(x_1) & p(x_2) & \dots & p(x_n) \end{bmatrix}$，输出为 $Y: \{y_1, y_2, \dots y_m\}$ 。

定义失真矩阵 $D = \begin{bmatrix} d(x_1, y_1) & d(x_1, y_2) & \dots & d(x_1, y_m) \\ d(x_2, y_1) & d(x_2, y_2) & \dots & d(x_2, y_m) \\ \vdots & \vdots & \vdots & \vdots \\ d(x_n, y_1) & d(x_n, y_2) & \dots & d(x_n, y_m) \end{bmatrix}$ 。

说明

失真函数是根据人们的实际需要和失真引起的损失，风险大小而人为规定的。

常用的失真函数有：

1 汉明失真：$d(x_i, y_j) = \begin{cases} 0 & x_i = y_j \\ 1 & x_i \neq y_j \end{cases}$

2 平方误差失真函数：$d(x_i, y_j) = (x_i - y_j)^2$

例题6-1 设信道输入为$X=\{0,1\}$，输出为$Y=\{0,?,2\}$，规定失真函数：

$$d(0,0)=d(1,1)=0$$
$$d(0,1)=d(1,0)=1$$
$$d(0,?)=d(1,?)=0.5$$

求失真矩阵。

解析6-1

$$D=\begin{bmatrix} d(0,0) & d(0,?) & d(0,1) \\ d(1,0) & d(1,?) & d(1,1) \end{bmatrix}=\begin{bmatrix} 0 & \frac{1}{2} & 1 \\ 1 & \frac{1}{2} & 0 \end{bmatrix}$$

矢量传输中的失真函数$d(\alpha, \beta)$

输入N长符号序列$\alpha = X_1X_2\cdots X_N$，其中每个消息$X$的取值均来自$\{x_1, x_2, ..., x_n\}$；

输出N长符号序列$\beta = Y_1Y_2\cdots Y_N$，其中每个消息Y的取值均来自$\{y_1, y_2, ..., y_m\}$；

用失真函数定量测度信源发生一个符号x_i，而在接收端再接收y_j所引起的误差或失真。

定义矢量失真函数 $d(\alpha, \beta) = \sum_{i=1}^{N} d(X_i, Y_i)$ 。

即矢量的失真度等于序列中对应单个信源符号的失真度之和。

同样可以定义失真度的矢量失真函数矩阵，而该矩阵中共有 $n^N \times m^N$个元素。

失真度和平均失真度

小节1 失真函数

小节2 平均失真度

平均失真度

$$\overline{D} = E[d(x_i, y_j)] = \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)d(x_i, y_j)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i)p(y_j\,|\,x_i)d(x_i, y_j)$$

平均失真度表示信源 P_X 在给定信道 $P_{Y|X}$ 中传输时所引起的失真的总体量度。

若平均失真度 $\overline{D}$ 不大于所允许的失真 D^*，即 $\overline{D} \le D^*$，则称此为保真度准则。

N长信源序列的平均失真度$\overline{D}(N)$

$$\overline{D}(N) = E[d(\alpha_i, \beta_j)] = \sum_{i=1}^{n^N} \sum_{j=1}^{m^N} p(\alpha_i, \beta_j) d(\alpha_i, \beta_j)$$

$$= \sum_{i=1}^{n^N} \sum_{j=1}^{m^N} p(\alpha_i) p(\beta_j \mid \alpha_i) d(\alpha_i, \beta_j)$$

$$= \sum_{i=1}^{n^N} \sum_{j=1}^{m^N} p(\alpha_i) p(\beta_j \mid \alpha_i) \sum_{l=1}^{N} d(\alpha_{il}, \beta_{jl})$$

信源的平均失真度（单个符号的失真度）$\overline{D}_N = \dfrac{1}{N}\overline{D}(N)$

当信源与信道均无记忆时, $\overline{D}(N) = \sum_{l=1}^{N} D_l$, D_l是信源序列中第l个分量的平均失真度；

若离散信源平稳，则$\overline{D}(N) = N\overline{D}$，离散无记忆平稳信源通过无记忆试验信道时，序列平均失真度为单符号平均失真度的N倍。

信息率失真函数

小节1 保真度准则

小节2 信息率失真函数

信息率失真函数

小节1 保真度准则

小节2 信息率失真函数

保真度准则的定义与限失真试验信道

保真度准则

若预先规定的平均失真度为 D^*，则称信源压缩后的平均失真度 $\overline{D}$ 不大于D^*的准则称为保真度准则，即$\overline{D} \le D^*$。

D^*允许的试验信道

将满足保真度准则的所有信道称为失真度 D^*允许信道（也称 D^*允许的实验信道），记作：

$$B_{D^*} = \{p(y_j|x_i) : \overline{D} \le D^*\}$$

对于确定信源和失真函数，不同编码对应不同的试验信道。

信息率失真函数

小节1 保真度准则

小节2 信息率失真函数

信息率失真函数的定义

在满足失真要求的信道中，寻找一种信道，使给定的信源经过这个信道传输时，其信息传输率，即平均互信息 $I(X;Y)$ 最小。定义这个最小值为信息率失真函数。

在满足保真度准则的前提下, $R(D)$是信息率允许压缩到的最小值，为信源特有的参数。

$$R(D) = \min_{p(y_j|x_i)\in B_D} \{I(X;Y)\} = \min_{p(y_j|x_i)\in B_D} \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i)p(y_j|x_i)\frac{\log p(y_j|x_i)}{p(y_j)}$$

注：此时满足 $I(X;Y) = H(Y) - H(Y|X) = H(Y)$。

信道容量与信息率失真函数的比较

信道容量$C = \max_{p(x_i)} I(X;Y)$是在信道固定前提下，选取一种信源概率分布使信息传输率最大（求极大值），它反映了信道传输信息的能力，是信道可靠传输的最大信息传输率，信道容量与信源无关，是反映信道特性的参量。

率失真函数$R(D) = \min_{p(y_j|x_i)\overline{D}\leq D} I(X;Y)$是在信源固定，满足保真度准则的条件下的信息传输率的最小值，反映了满足一定失真度的条件下，信源可压缩的程度，即满足失真要求而再现信源消息所必须获得的最少平均信息量。

信息率失真函数的性质

1 $R(D)$的定义域和值域：

平均失真度D的取值范围：　$[D_{\min}, D_{\max}]$

率失真函数的取值范围：　$[H(X), 0]$

- 对$D_{\min}$的讨论：

由信源给定

$$D_{\min} = \min[\sum_X \sum_Y p(x_i)p(y_j|x_i)d(x_i, y_j)] = \boxed{\sum_{i=1}^{r} p(x_i)} \min\boxed{[\sum_{j=1}^{s} p(y_j|x_i)d(x_i, y_j)]}$$

选择合适的试验信道

让求和号中的每一项最小化

当信源输入（即x_i）固定时，对于不同的y_j，失真矩阵对应行的元素中必然有一个最小值，或若干个相同的最小值。我们可以通过构造试验信道使平均失真度最小化。

信息率失真函数的性质

1 $R(D)$的定义域和值域：

平均失真度D的取值范围： $[D_{\min}, D_{\max}]$

率失真函数的取值范围： $[H(X), 0]$

- 对$D_{\min}$的讨论：

由信源给定

$$D_{\min} = \min[\sum_X \sum_Y p(x_i)p(y_j|x_i)d(x_i, y_j)] = \boxed{\sum_{i=1}^{r} p(x_i)} \min\boxed{[\sum_{j=1}^{s} p(y_j|x_i)d(x_i, y_j)]}$$

选择合适的试验信道

让求和号中的每一项最小化

我们构造试验信道满足：

$\sum_j p(y_j|x_i) = 1$ $d(x_i, y_j)$ 为失真矩阵每一行中的最小值

$\sum_j p(y_j|x_i) = 0$ $d(x_i, y_j)$不为失真矩阵每一行中的最小值

因此最小平均失真度 $D_{\min} = \sum_i p(x_i)\boxed{\min d(x_i, y_j)}$ 信源输入x_i固定时，失真矩阵对应行的最小值

信息率失真函数的性质

1 $R(D)$的定义域和值域：

平均失真度D的取值范围：　$[D_{\min}, D_{\max}]$

率失真函数的取值范围：　$[H(X),0]$

- 对$D_{\min}$的讨论：

$D_{\min} = \sum_i p(x_i)\boxed{\min d(x_i, y_j)}$ **信源输入x_i固定时，失真矩阵对应行的最小值**

当**失真矩阵中每行至少有一个零元素**时，$D_{\min} = 0$，此时意味着不允许任何失真的情况，一般情况下$D_{\min} = 0$，此时信息传输率至少应该等于信源输出的信息量，即信源熵，故$R(0) = H(X)$；

当失真矩阵不符合上述条件时 $D_{\min} \neq 0$；

信息率失真函数的性质

1 $R(D)$的定义域和值域:

平均失真度D的取值范围: $[D_{\min}, D_{\max}]$

率失真函数的取值范围: $[H(X),0]$

- 对$D_{\max}$的讨论:

$D_{\max}$是满足信息传输率$R(D)=0$的最小失真，即对于$\forall D \ge D_{\max}$，$R(D)=0$恒成立;

对于$\forall D \ge D_{\max}$，$R(D)=0$，从而$I(X;Y)=0$，此时X,Y统计独立，故$p(y_j|x_i)$与x_i无关;

我们不难发现会有很多种试验信道满足条件，但我们要研究的是下界$D=D_{\max}$的情形;

$$D_{\max} = \min \sum_{i,j} p(x_i)p(y_j|x_i)d(x_i,y_j)$$

$$= \min \sum_{i,j} p(y_j|x_i) \sum_{i} p(x_i)d(x_i,y_j)$$

信息率失真函数的性质

1 $R(D)$的定义域和值域：

平均失真度D的取值范围： $[D_{\min}, D_{\max}]$

率失真函数的取值范围： $[H(X), 0]$

- 对$D_{\max}$的讨论：

$$D_{\max} = \min \sum_{i,j} p(x_i)p(y_j|x_i)d(x_i, y_j)$$

$$= \min \sum_{i,j} p(y_j|x_i)\boxed{\sum_i p(x_i)d(x_i, y_j)}$$ **最小化**

我们构造试验信道满足：

对于固定的x_i，当$\sum_i p(x_i)d(x_i, y_j)$ 最小时，令 $p(y_j|x_i) = q(y_j) = 1$

否则，令 $p(y_j|x_i) = q(y_j) = 0$，则$D_{\max} = \min \sum_i p(x_i)d(x_i, y_j)$

例题6-2 设信道输入为$X=\{a_1,a_2,a_3\}$，且为等概分布，已知失真矩阵为$D=\begin{bmatrix}1&2&3\\2&1&3\\3&2&1\end{bmatrix}$；

试求$D_{\min}$与$D_{\max}$ 以及相应试验信道的转移概率矩阵。

解析6-2

$$D_{\min}=\sum_i p(x_i)\min d(x_i,y_j)$$
$$=\frac{1}{3}\min(1,2,3)+\frac{1}{3}\min(2,1,3)+\frac{1}{3}\min(3,2,1)$$
$$=1$$

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

试验信道为

$$P_{Y|X}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

$$D_{\max}=\min\sum_i p(x_i)d(x_i,y_j)$$
$$=\min\begin{bmatrix}p(a_1)\times1+p(a_2)\times2+p(a_3)\times3\\p(a_1)\times2+p(a_2)\times1+p(a_3)\times2\\p(a_1)\times3+p(a_2)\times3+p(a_3)\times1\end{bmatrix}=\frac{5}{3}$$

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

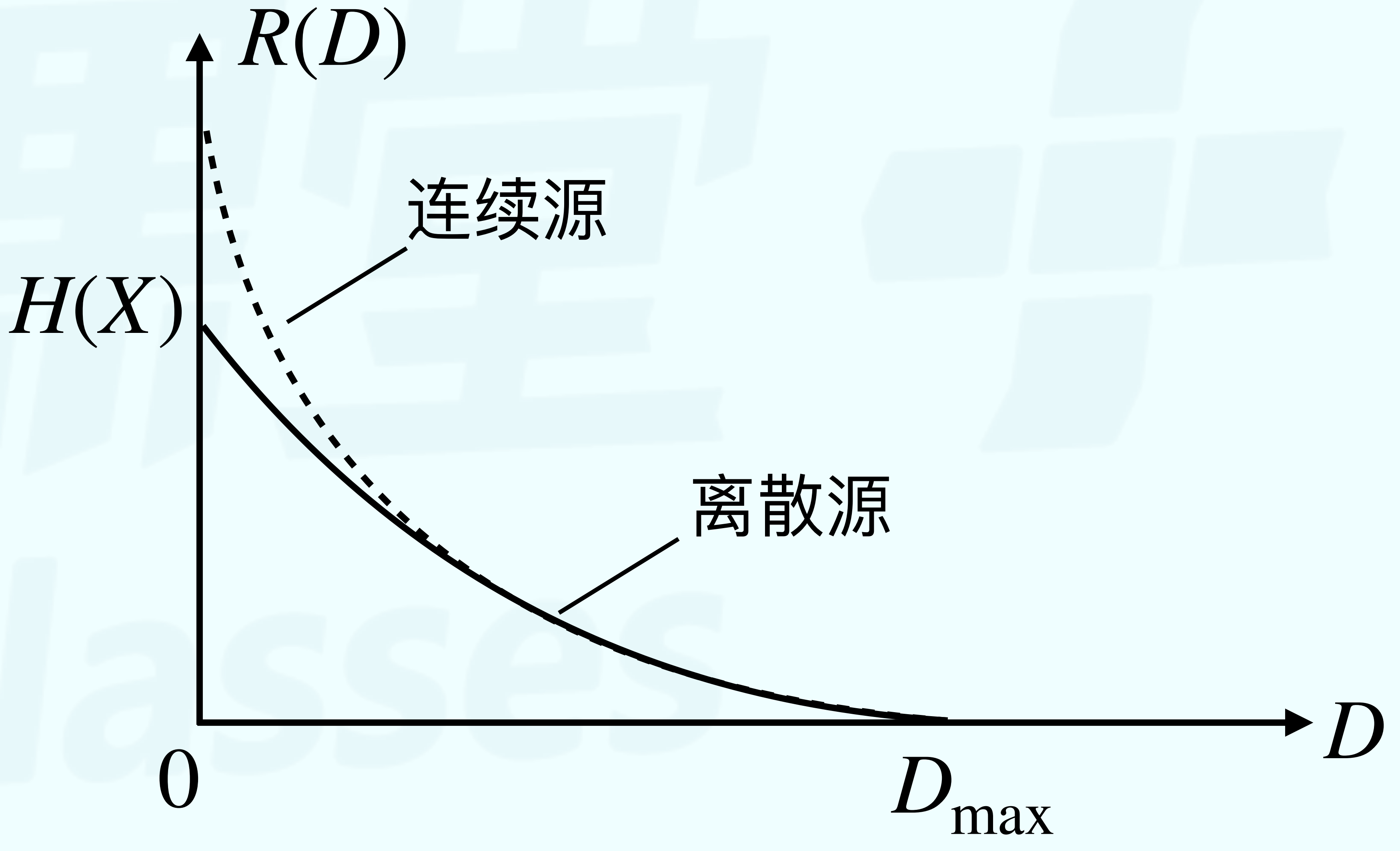
试验信道为

$$P_{Y|X}=\begin{bmatrix}0&1&0\\0&1&0\\0&1&0\end{bmatrix}$$

信息率失真函数的性质

2 $R(D)$是关于D的下凸函数。

3 $R(D)$在定义区间是严格递减函数。



限失真信源编码定理

小节1 香农第三定理

香农第三定理

设离散无记忆平稳信源的信息率失真函数为$R(D)$，只要满足$R \geq R(D)$，信源序列长度N足够大时，一定存在一种编码方法，其译码失真小于或等于$D+\varepsilon$，其中ε是任意小的正数。

反之，若$R < R(D)$，则无论采用什么样的编码方法，其译码失真必大于D。

即在允许失真D的条件下，信源最小的、可达的信息传输率是信源的$R(D)$，提供了压缩的下界。

说明

该定理为存在性定理。

正定理：$R \geq R(D)$ 时，译码失真小于或等于$D+\varepsilon$的码肯定存在，但定理本身未给出码的构造方法。

逆定理：$R < R(D)$ 时，译码失真必大于D，肯定找不到满足条件的码，因此用不着浪费时间与精力。

香农信息论的三个概念：信源熵、信道容量和信息率失真函数，都是临界值，是从理论上衡量通信能够满足要求的重要极限。