《信息论基础》

第7章 信息率失真函数

吉小鹏

E-mail: 003163@nuist.edu.cn

南京信息工程大学 电子与信息工程学院 尚贤楼209




提纲

7.1 基本概念

7.2 离散信源信息率失真函数

实际通信系统允许一定的失真存在。

1 打电话；

2 放电影，视觉暂留性。

允许压缩信源输出的信息率。

研究内容：信息率 ⟺ 允许失真

7.1 基本概念

失真函数与平均失真度

设离散无记忆信源$\begin{pmatrix} X \\ P(X) \end{pmatrix} = \begin{Bmatrix} a_1 & a_2 & \cdots & a_n \\ p(a_1) & p(a_2) & \cdots & p(a_n) \end{Bmatrix}$，通过信道传送到接收端$\begin{pmatrix} Y \\ P(Y) \end{pmatrix} = \begin{Bmatrix} b_1 & b_2 & \cdots & b_m \\ p(b_1) & p(b_2) & \cdots & p(b_m) \end{Bmatrix}$。信道传递概率

$$P(Y/X) = \begin{bmatrix} p(b_1/a_1) & p(b_2/a_1) & \cdots & p(b_m/a_1) \\ p(b_1/a_2) & p(b_2/a_2) & \cdots & p(b_m/a_2) \\ \vdots & \vdots & \cdots & \vdots \\ p(b_1/a_n) & p(b_2/a_n) & \cdots & p(b_m/a_n) \end{bmatrix}$$

对每一对(a_i,b_j)，指定一个非负的函数，称为失真函数或失真度

$$d(a_i,b_j) \geq 0, \quad i = 1,2,\cdots,n; \quad j = 1,2,\cdots,m \tag{7.1.1}$$

7.1 基本概念

■ 失真函数与平均失真度

失真度表示成矩阵形式，称为失真矩阵

$$[D]=\begin{bmatrix} d(a_1,b_1) & d(a_1,b_2) & \dots & d(a_1,b_m) \\ d(a_2,b_1) & d(a_2,b_2) & \dots & d(a_2,b_m) \\ \dots & \dots & \dots & \dots \\ d(a_n,b_1) & d(a_n,b_2) & \dots & d(a_n,b_m) \end{bmatrix} \qquad (7.1.2)$$

例：设信源符号集为X={0,1}, 接收端收到符号集为Y= {0,1,2},规定失真函数为

d(0,0)＝d(1,1)= 0

d(0,1)＝d(1,0)= 1

d(0,2)＝d(1,2)= 0.5

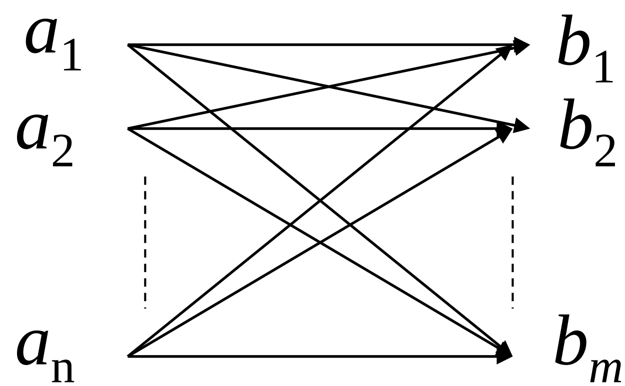
$$D=\begin{bmatrix} 0 & 1 & 0.5 \\ 1 & 0 & 0.5 \end{bmatrix}$$

7.1 基本概念

失真函数与平均失真度

常见失真函数： 1
$$d(a_i,b_j)=\begin{cases}0 & a_i=b_j\\ a & a_i\neq b_j\end{cases} \qquad (7.1.3)$$

$$\begin{bmatrix}0 & a & \dots & a\\ a & 0 & \dots & a\\ \dots & \dots & \dots & \dots\\ a & a & \dots & 0\end{bmatrix}$$



$a=1$ ⟹ 汉明失真
误差概率失真

$$\begin{bmatrix}0 & 1 & \cdots & 1\\ 1 & 0 & \cdots & 1\\ \cdots & \cdots & \cdots & \cdots\\ 1 & 1 & \cdots & 0\end{bmatrix}$$

$$Ed(X,Y)=\mathrm{P}(X\neq Y)$$

7.1 基本概念

■ **失真函数与平均失真度**

② 平方误差失真函数 $d(a_i,b_j)=(b_j-a_i)^2$ (7.1.5)

由于a_i和b_j都是随机变量，所以单个符号的失真函数$d(a_i,b_j)$也是随机变量，限失真时的失真值，只能用它的数学期望或统计平均值，因此将失真函数的数学期望称为**平均失真度**，记为

$$\bar{D}=E[d(a_i,b_j)]=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_i)p(b_j/a_i)d(a_i,b_j) \tag{7.1.7}$$

保真度准则 $\rightarrow$ $\boxed{\bar{D}\le D}$ $\leftarrow$ 允许失真上界 (7.1.8)

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_i)p(b_j/a_i)\log\frac{p(b_j/a_i)}{p(b_j)}$$

7.1 基本概念

失真函数与平均失真度

失真度量的概念是定义在字母表×字母表上的，下面把这个定义推广到序列上去。

定义：序列 x^n 和序列 y^n 间的失真定义为

$$d(x^n,y^n)=\sum_{i=1}^{n}d(x_i,y_i)$$

定义：N维信源符号序列的平均失真度

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

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

当信源和信道都是无记忆的，N维信源符号序列的平均失真度 $\overline{D}(N)=\sum_{l=1}^{N}\overline{D}_l$，信源的平均失真度 $\overline{D}_N=\frac{1}{N}\sum_{l=1}^{N}\overline{D}_l$ 。

如果离散信源是平稳的，$\overline{D}_l=\overline{D}$，$\overline{D}(N)=N\overline{D}$ 。

7.1 基本概念

■ 失真函数与平均失真度

对于N次无记忆扩展信源和信道，定义平均失真度为

$$\bar{D}(N)=\bar{D}_1+\bar{D}_2+\cdots+\bar{D}_N=\sum_{k=1}^{N}\bar{D}_k \tag{7.1.12}$$

如果N次无记忆扩展信源是同一单符号信源在不同时刻通过同一信道，则

$$\bar{D}_k=\bar{D}=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)d(a_i,b_j),\ \ k=1,2,\cdots,N \tag{7.1.14}$$

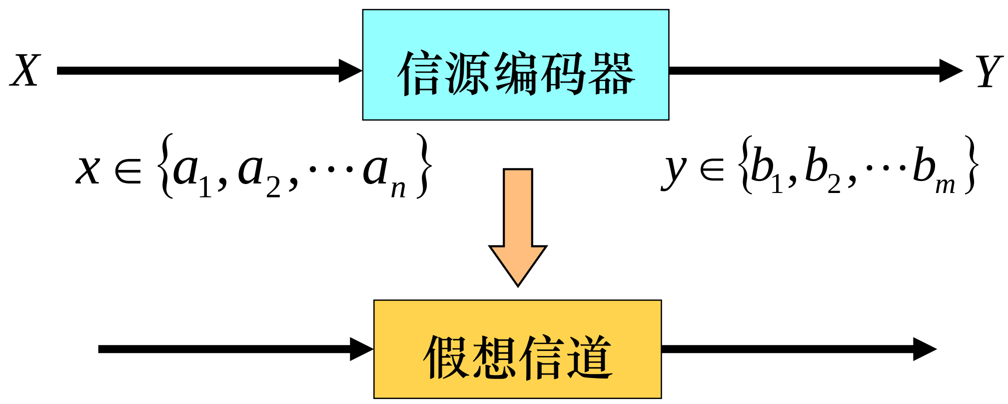
$$\bar{D}(N)=N\bar{D} \tag{7.1.15}$$

$$\bar{D}(N)\le ND \tag{7.1.16}$$

N次无记忆扩展信源通过N次无记忆扩展信道的保真度准则

7.1 基本概念

信息率失真函数的定义



信源编码器的目的是使码字信息传输率R尽量小，然而R越小，引起的平均失真就越大。给出一个失真的限制值D，在满足平均失真$\overline{D} \le D$的条件下，选择一种编码方法使信息率R尽可能小。如果将信源编码器看作信道，则限失真信源编码问题就变成由输出端Y获取输入端X最小平均互信息量问题。

7.1 基本概念

■ 信息率失真函数的定义

$$\bar{D} = E[d(a_i,b_j)] = \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j/a_i)d(a_i,b_j)$$

平均失真由信源分布$p(a_i)$、假想信道的转移概率$p(b_j/a_i)$和失真函数$d(a_i,b_j)$决定，若$p(a_i)$和$d(a_i,b_j)$已定，则调整$p(b_j/a_i)$ (选择假想信道) 使满足保真度准则

$$P_D = \{p(b_j/a_i): \bar{D} \le D,\ i=1,2,\cdots,n;\ j=1,2,\cdots,m\} \qquad (7.1.17)$$

称P_D为*D*失真许可的试验信道。

对于离散无记忆N次扩展信源和信道，相应的试验信道集合为

$$P_{D(N)} = \{p(b_j/a_i): \bar{D}(N) \le ND,\ i=1,2,\cdots,n^N;\ j=1,2,\cdots,m^N\} \quad (7.1.18)$$

信息率失真函数的定义

由于互信息量是$p(a_i)$和$p(b_j/a_i)$得二元函数，当$p(a_i)$一定时，$I(X;Y)$是关于$p(b_j/a_i)$ 的下凸函数，存在极小值。因而在P_D中寻找一种信道$p(b_j/a_i)$ ，使给定的信源$p(a_i)$经过此信道传输后，互信息量$I(X;Y)$达到最小。该最小值就称为信息率失真函数(简称率失真函数)$R(D)$，即

$$\boxed{R(D)=\min_{p(b_j/a_i)\in P_D} I(X;Y)} \tag{7.1.19}$$

对于离散无记忆N次扩展信源和信道，相应的率失真函数为

$$\boxed{R_N(D)=\min_{p(\beta_j/\alpha_i)\in P_{D(N)}} I(X^N;Y^N)} \tag{7.1.20}$$

由信源和信道的无记忆性，易证

$$R_N(D)=NR(D) \tag{7.1.21}$$

7.1 基本概念

■ 信息率失真函数的性质

① 信息率失真函数的定义 $[0, D_{\max}]$ 域

$D=0 \Rightarrow R(D)=H(X)$　　连续：$\lim\limits_{D\to 0} R(D) \to \infty$

$D\uparrow \Rightarrow R\downarrow$

$D = D_{\max} \Rightarrow R(D)=0$　　$D \geq D_{\max} \Rightarrow R(D) \equiv 0$

当给定信源[X P(X)]和失真矩阵D，信源的最小平均失真度为

$$D_{\min} = \min[\sum_X \sum_Y p(a_i)p(b_j \mid a_i)d(a_i,b_j)]$$

只有当D中每行至少有一个零元素时，信源的平均失真度才能达到零。当Dmin=0时，表示信源不允许任何失真存在。直观理解要求信息传输率至少应该等于信源输出的信息量，即信息熵。

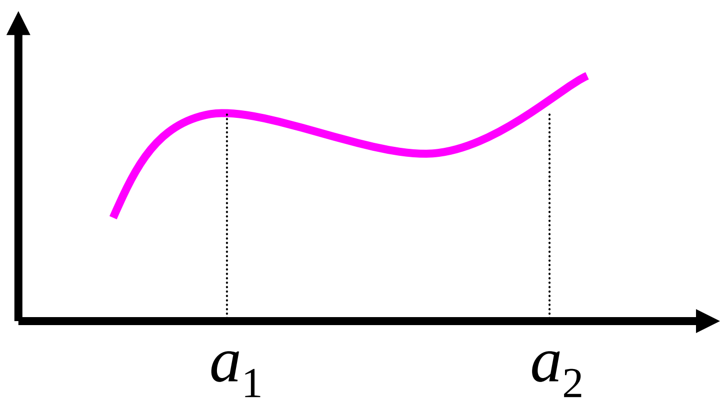
7.1 基本概念

■ **信息率失真函数的性质**

令 $p(b_j/a_i)=p(b_j)$, $\sum_{i=1}^{n}p(a_i)d(a_i,b_j)=D_j$

$$\begin{aligned}D_{\max}&=\min_{p(b_j)}\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_i)p(b_j)d(a_i,b_j)\\&=\min_{p(b_j)}\sum_{j=1}^{m}p(b_j)\sum_{i=1}^{n}p(a_i)d(a_i,b_j)=\min_{p(b_j)}\sum_{j=1}^{m}p(b_j)D_j\end{aligned}\qquad(7.1.26)$$

线性分配



$0<\alpha<1,\ \ \alpha a_1+(1-\alpha)a_2$

信息率失真函数的性质

假定所有D_j中，D_s最小，令

$$p(b_j)=\begin{cases}1 & j=s\\ 0 & j\neq s\end{cases}$$

$$\boxed{D_{\max}=\min_j D_j} \qquad (7.1.27)$$

$$\begin{array}{cc}\begin{bmatrix} d(a_1,b_1) & \dots & d(a_n,b_1)\\ d(a_1,b_2) & \dots & d(a_n,b_2)\\ \dots & \dots & \dots\\ d(a_1,b_m) & \dots & d(a_n,b_m)\end{bmatrix} & \begin{array}{c}D_1\\ D_2\\ \dots\\ D_m\end{array}\end{array}$$

7.1 基本概念

信息率失真函数的性质

2 信息率失真函数对允许平均失真度的下凸性

对于$0 \le \theta \le 1$，由线性分配原理

$$R[\theta D' + (1-\theta)D''] \le \theta R(D') + (1-\theta)R(D'') \qquad (7.1.28)$$

$$p_1(b_j/a_i),\ p_2(b_j/a_i) \Rightarrow R(D'),\ R(D'')$$

$$\bar{D}_1 = \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p_1(b_j/a_i)d(a_i,b_j) \le D' \qquad (7.1.29)$$

$$\bar{D}_2 = \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p_2(b_j/a_i)d(a_i,b_j) \le D'' \qquad (7.1.30)$$

令 $p_1(b_j) = \sum_{i=1}^{n} p(a_i)p_1(b_j/a_i),\quad p_2(b_j) = \sum_{i=1}^{n} p(a_i)p_2(b_j/a_i)$

7.1 基本概念

- 信息率失真函数的性质

$$I(X;Y_1)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_i)p_1(b_j/a_i)\log\frac{p_1(b_j/a_i)}{p_1(b_j)}=R(D') \qquad (7.1.31)$$

$$I(X;Y_2)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_i)p_2(b_j/a_i)\log\frac{p_2(b_j/a_i)}{p_2(b_j)}=R(D'') \qquad (7.1.32)$$

定义新试验信道：

$$p(b_j/a_i)=\theta p_1(b_j/a_i)+(1-\theta)p_2(b_j/a_i) \qquad (7.1.35)$$

$$\bar{D}=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_i)p(b_j/a_i)d(a_i,b_j)=\theta\bar{D}_1+(1-\theta)\bar{D}_2\le\theta D'+(1-\theta)D''=D$$

$$\bar{D}\le D$$

$$I(X;Y)\ge R(D)=R[\theta D'+(1-\theta)D'']$$

7.1 基本概念

- 信息率失真函数的性质

由$I(X;Y)$对$p(b_j/a_i)$的下凸性：

$$\begin{aligned} I(X;Y) &\le \theta I(X;Y_1)+(1-\theta)I(X;Y_2) \\ &= \theta R(D')+(1-\theta)R(D'') \end{aligned}$$

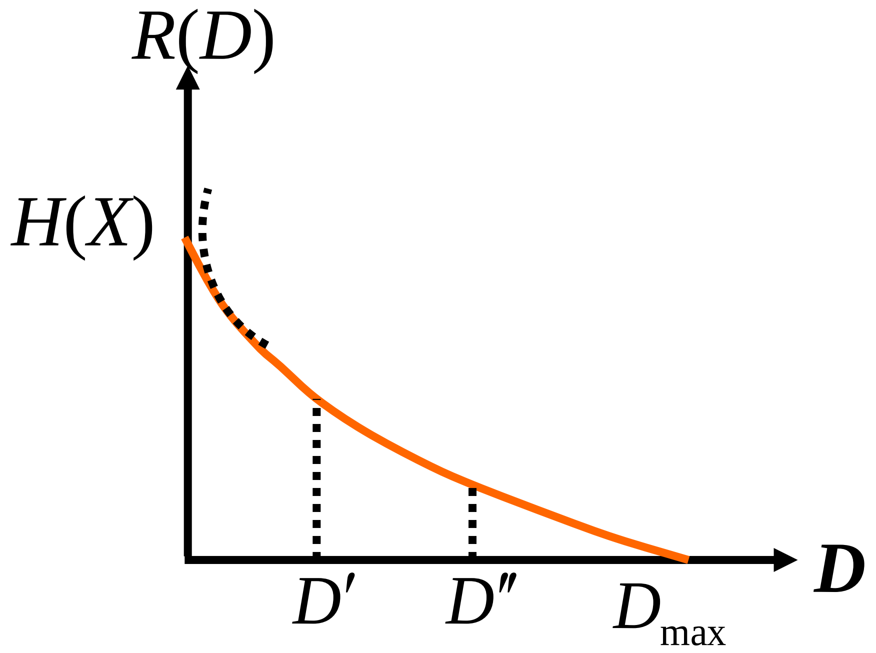
$$\Downarrow$$

$$R[\theta D'+(1-\theta)D''] \le \theta R(D')+(1-\theta)R(D'')$$

7.1 基本概念

■ **信息率失真函数的性质**

3 信息率失真函数的单调递减和连续性



提纲

7.1 基本概念

7.2 离散信源信息率失真函数

7.2 离散信源信息率失真函数

■ **离散信源信息率失真函数的参量表达式**

$p(a_i), d(a_i, b_j), p(b_j/a_i) \in P_D, \quad \bar{D} \le D$

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j/a_i)\ln\frac{p(b_j/a_i)}{p(b_j)} \tag{7.2.4}$$

$$\boxed{\begin{aligned} &\bar{D} = \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j/a_i)d(a_i, b_j) \\ &\sum_{j=1}^{m} p(b_j/a_i) = 1, \quad i = 1, 2, \cdots, n \end{aligned}} \tag{7.2.5}$$

$$p(b_j) = \sum_{i=1}^{n} p(a_i)p(b_j/a_i)$$

7.2 离散信源信息率失真函数

■ **离散信源信息率失真函数的参量表达式**

构造新函数

$$\Phi = I(X;Y) - S[\sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j/a_i)d(a_i,b_j) - \bar{D}] - \mu_i\left[\sum_{j=1}^{m} p(b_j/a_i) - 1\right] \tag{7.2.6}$$

令

$$\frac{\partial \Phi}{\partial p(b_j/a_i)} = 0 \tag{7.2.7}$$

$$p(a_i)\ln\frac{p(b_j/a_i)}{p(b_j)} + p(a_i) - p(a_i) - Sp(a_i)d(a_i,b_j) - u_i = 0 \tag{7.2.8}$$

7.2 离散信源信息率失真函数

■ 离散信源信息率失真函数的参量表达式

令 $\ln\lambda_i = \dfrac{u_i}{p(a_i)}$ ⇨ $\ln\dfrac{p(b_j/a_i)}{p(b_j)} - Sd(a_i,b_j) - \ln\lambda_i = 0$

$$p(b_j/a_i) = \lambda_i p(b_j) e^{Sd(a_i,b_j)},\quad i=1,2,\cdots,n;\ j=1,2,\cdots,m \qquad (7.2.10)$$

对j求和，有

$$1 = \lambda_i \sum_{j=1}^{m} p(b_j) e^{Sd(a_i,b_j)} \qquad (7.2.11)$$

(7.2.10)两边乘以$p(a_i)$，再对i求和

$$p(b_j) = p(b_j)\sum_{i=1}^{n} \lambda_i p(a_i) e^{Sd(a_i,b_j)}$$

$$\sum_{i=1}^{n} \lambda_i p(a_i) e^{Sd(a_i,b_j)} = 1,\ p(b_j)\neq 0,\ j=1,2,\cdots,m \qquad (7.2.12)$$

7.2 离散信源信息率失真函数

- 离散信源信息率失真函数的参量表达式

$$(7.2.12) \Rightarrow \lambda_i \Rightarrow (7.2.11) \Rightarrow p(b_j)$$

$$\lambda_i,\ p(b_j) \Rightarrow (7.2.10) \Rightarrow p(b_j/a_i)$$

$$D(S)=\sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)\lambda_i p(b_j)e^{Sd(a_i,b_j)}d(a_i,b_j) \qquad (7.2.14)$$

$$R(S)=\sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)\lambda_i p(b_j)e^{Sd(a_i,b_j)}\ln\frac{\lambda_i p(b_j)e^{Sd(a_i,b_j)}}{p(b_j)}$$

$$=SD(S)+\sum_{i=1}^{n}p(a_i)\ln\lambda_i \qquad (7.2.15)$$

7.2 离散信源信息率失真函数

- 离散信源信息率失真函数的参量表达式

$$\frac{dR}{dD} = S$$

$$\frac{dR}{dD} = \frac{\partial R}{\partial D} + \frac{\partial R}{\partial S}\frac{dS}{dD} + \sum_{i=1}^{n}\frac{\partial R}{\partial \lambda_i}\frac{d\lambda_i}{dD} = S + D\frac{dS}{dD} + \sum_{i=1}^{n}\frac{p(a_i)d\lambda_i}{\lambda_i dD}$$

$$= S + [D + \sum_{i=1}^{n}\frac{p(a_i)}{\lambda_i}\frac{d\lambda_i}{dS}]\frac{dS}{dD} = S \qquad (7.2.16)$$

负值

在公式(7.2.12)两边对S取导数

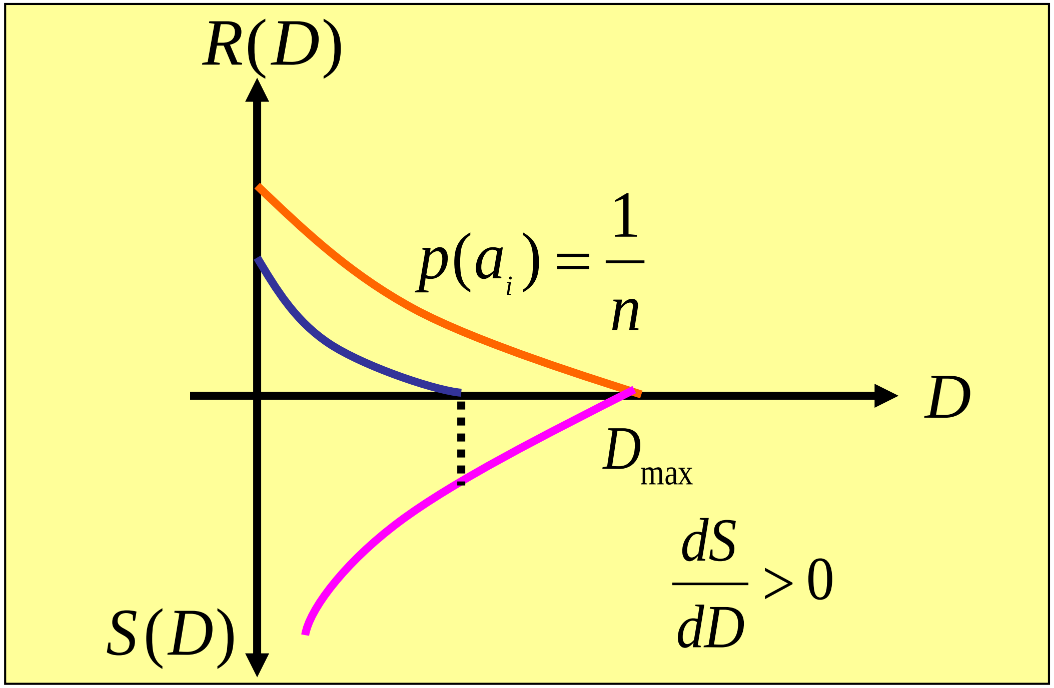
$$\sum_{i=1}^{n}\left[\frac{d\lambda_i}{dS}p(a_i)e^{Sd(a_i,b_j)} + \lambda_i p(a_i)d(a_i,b_j)e^{Sd(a_i,b_j)}\right] = 0 \qquad (7.2.17)$$

7.2 离散信源信息率失真函数

■ **离散信源信息率失真函数的参量表达式**

两边乘以$p(b_j)$ 并对j求和

$$\sum_{i=1}^{n}\frac{p(a_i)}{\lambda_i}\frac{d\lambda_i}{dS}+D=0 \Longrightarrow \frac{dR}{dD}=S \qquad (7.2.18)$$



7.2 离散信源信息率失真函数

- 二元及等概率离散信源的信息率失真函数

$$\begin{pmatrix} X \\ P(X) \end{pmatrix} = \begin{bmatrix} a_1 & a_2 \\ p & 1-p \end{bmatrix}, p \le \frac{1}{2} \quad [D] = \begin{bmatrix} 0 & \alpha \\ \alpha & 0 \end{bmatrix}, \alpha > 0$$

$$D_{\max} = \min_j D_j \quad D_j = \sum_{i=1}^{n} p(a_i) d(a_i, b_j)$$

$$D_1 = (1-p)\alpha \quad D_2 = p\alpha \quad D_{\max} = D_2 = \alpha p$$

$$1 = \sum_{i=1}^{n} \lambda_i p(a_i) e^{Sd(a_i, b_j)}$$

7.2 离散信源信息率失真函数

二元及等概率离散信源的信息率失真函数

1

$$D_{\max}=\min_j D_j=\min_j \sum_{i=1}^{n} p(a_i)d(a_i,b_j)$$

2

$$\sum_{i=1}^{n}\lambda_i p(a_i)e^{Sd(a_i,b_j)}=1$$

$\Longrightarrow \lambda_i$

3

$$\frac{1}{\lambda_i}=\sum_{j=1}^{m} p(b_j)e^{Sd(a_i,b_j)}$$

$\Longrightarrow p(b_j)$

4

$$p(b_j/a_i)=p(b_j)\lambda_i e^{Sd(a_i,b_j)}$$

$\Longrightarrow p(b_j/a_i)$

5

$$D(S)=\sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j)\lambda_i d(a_i,b_j)e^{Sd(a_i,b_j)}$$

6

$$R(S)=SD(S)+\sum_{i=1}^{n} p(a_i)\ln\lambda_i$$

7

验证 $p(b_j/a_i)$ 是否大于等于零

7.2 离散信源信息率失真函数

- 二元及等概率离散信源的信息率失真函数

$$\sum_{j=1}^{m} p(b_j)e^{Sd(a_i,b_j)} = \frac{1}{\lambda_i}$$

$$\lambda_1 p + \lambda_2(1-p)e^{s\alpha} = 1$$
$$\lambda_1 p e^{S\alpha} + \lambda_2(1-p) = 1$$

$$\lambda_1 = \frac{1}{p(1+e^{S\alpha})}$$
$$\lambda_2 = \frac{1}{(1-p)(1+e^{S\alpha})}$$

$$p(b_1) + p(b_2)e^{S\alpha} = p(1+e^{S\alpha})$$
$$p(b_1)e^{S\alpha} + p(b_2) = (1-p)(1+e^{S\alpha})$$

$$p(b_1) = \frac{p-(1-p)e^{S\alpha}}{1-e^{S\alpha}}$$
$$p(b_2) = \frac{(1-p)-pe^{S\alpha}}{1-e^{S\alpha}}$$

二元及等概率离散信源的信息率失真函数

$$p(b_j/a_i)=\lambda_i p(b_j)e^{Sd(a_i,b_j)}$$

$$p(b_1/a_1)=\frac{p-(1-p)e^{S\alpha}}{p(1-e^{2S\alpha})} \qquad p(b_2/a_1)=\frac{p-(1-p)e^{S\alpha}}{(1-p)(1-e^{2S\alpha})}$$

$$p(b_1/a_2)=\frac{(1-p)-pe^{S\alpha}}{p(1-e^{2S\alpha})} \qquad p(b_2/a_2)=\frac{(1-p)-pe^{S\alpha}}{(1-p)(1-e^{2S\alpha})} \tag{7.2.28}$$

7.2 离散信源信息率失真函数

■ 二元及等概率离散信源的信息率失真函数

$$D(S)=\sum_{i=1}^{n}\sum_{j=1}^{m}\lambda_i p(a_i)p(b_j)d(a_i,b_j)e^{Sd(a_i,b_j)}=\frac{\alpha e^{S\alpha}}{1+e^{S\alpha}} \tag{7.2.29}$$

$$R(S)=SD(S)+\sum_{i=1}^{n}p(a_i)\ln\lambda_i=\frac{S\alpha e^{S\alpha}}{1+e^{S\alpha}}-p\ln p-\frac{-\ln(1+e^{S\alpha})}{(1-p)\ln(1-p)} \tag{7.2.30}$$

$$S=\frac{1}{2}\ln\frac{D/\alpha}{1-D/\alpha} \tag{7.2.31}$$

$$\begin{aligned}R(D)&=\frac{D}{\alpha}\ln\frac{D}{\alpha}-\frac{D}{\alpha}\ln(1-\frac{D}{\alpha})+\ln(1-\frac{D}{\alpha})-p\ln p-(1-p)\ln(1-p)\\&=H(p)-H(\frac{D}{\alpha})\end{aligned} \tag{7.2.36}$$

容忍失真允许
压缩的信息率

■ 二元及等概率离散信源的信息率失真函数

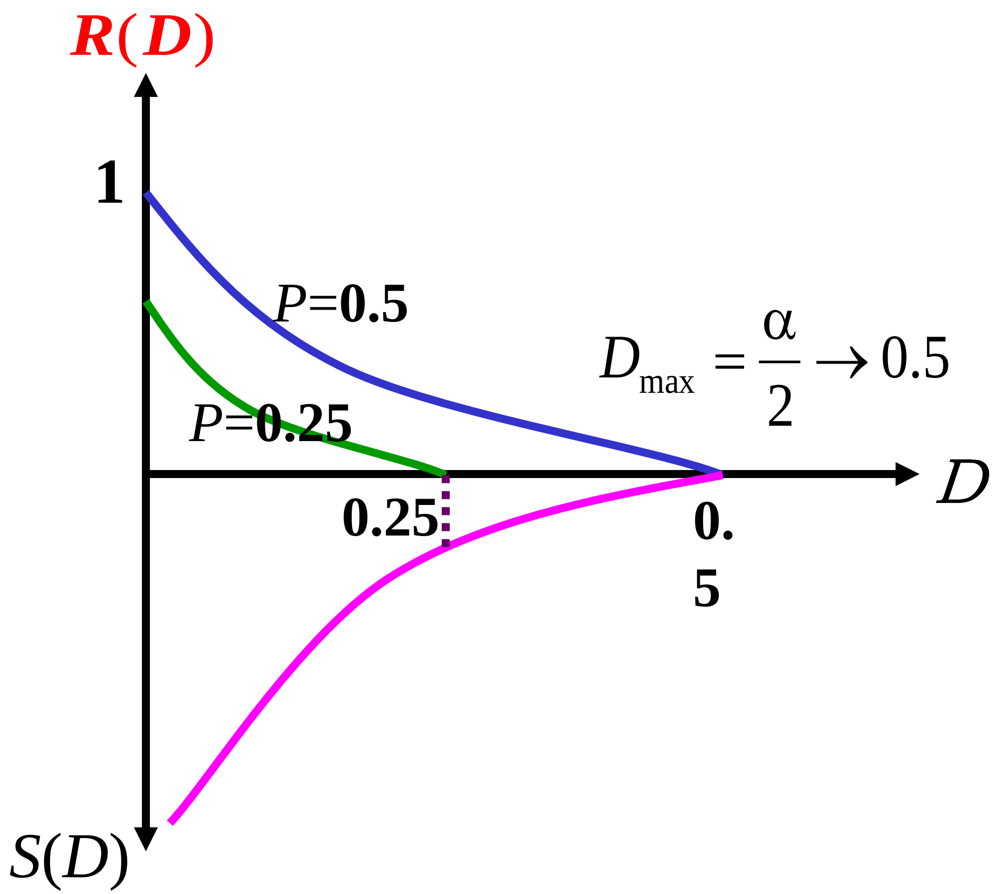


图7.2.2 二元信源和对称失真函数的$R(D)$和$S(D)$曲线