信息论笔记

1-绪论

信息论的产生：Shannon 通信的数学理论，1984

信息论的研究对象是通信系统，研究目的是使信号有效可靠的传输。

通信系统组成：信源、编码器、信道（噪声）、译码器、信宿

信息：本质是不确定性，包括信号、消息等形式。通信系统中实际传输的是信号，本质内容是信息。

信息论的主要内容：一个概念，三个定理

1. 信息的度量：
2. 无失真信源编码:$R \geq H \Leftrightarrow \exists$
3. 信道容量和可靠传输:$R < C \Leftrightarrow P_E \to 0$
4. 信息率失真理论: $R > R(D) \Leftrightarrow \overline{D} \leq D$

2-信息的度量

自信息: 随机变量某一取值所含的信息 $I(x) = -\log p(x)$

联合自信息: $I(x,y) = -\log p(x,y)$

条件自信息: $I(x|y) = -\log p(x|y)$ 满足 $I(x,y) = I(x) + I(x|y)$

信息熵: 随机事件所含信息 $H(X) = E_{p_{X(x)}}[I(x)] = \sum_x -p(x)\log p(x)$

联合熵: $H(XY) = \sum_x \sum_y -p(xy)\log p(xy)$

条件熵: $H(X|Y) = \sum_x \sum_y -p(xy)\log p(x|y)$

相对熵:(又称 K-L 散度) $D(P\|Q) = \sum_x -P(x)\log \frac{P(x)}{Q(x)}$

平均互信息: $I(X|y) = \sum_x p(x|y)\log \frac{p(xy)}{p(x)p(y)}$ $I(X;Y) = \sum_x \sum_y p(xy)\log \frac{p(xy)}{p(x)p(y)}$

平均条件互信息：$I(X;Y|Z) = \sum_{x,y,z} p(xyz)\log \frac{p(x|yz)}{p(x|z)}$

2.1 性质

熵的性质: 对称性、非负性、扩展性、可加性、极值性、一一对应变换下的不变性

互信息的性质:

THE $D(P\|Q) \geq 0)$

proof $D(P\|Q) \geq \sum_x P\left(1-\frac{Q}{P}\right) = \sum_x P - Q = 0$

THE $H(X|Y) \leq H(X)$

THE (凸函数的性质)$f(\sum a_i b_i) \geq \sum a_i f(b_i) \quad (\sum a_i = 0, a_i \to= 0)$

THE $H(XY) = H(X) + H(Y|X)$

EXT 熵的链规则

THE $H(X) \leq \log n$ 仅当等概分布

THE $I(X|y) \geq 0$ 仅当$X, \{y, \overline{y}\}$独立

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

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

THE $I(X;Y|Z) \geq 0$ 仅当$p(x|z) = p(x|yz)$

EXT 平均互信息的链原则

3-离散信源

3.1 分类

按符号集的有限性; 按信源和符号的依赖关系(记忆性); 按统计特性(平稳性)

3.2 信源的描述

?对单符号信源, 可用p_i描述,

3.3 信源的扩展

1. 等长消息扩展:
2. 变长消息扩展:

3.4 信源的熵

DEF N 次扩展源平均符号熵: $H_N(X) = N^{-1}H(X_1...X_N)$

EXT 极限符号熵 $H_\infty(X)$

THE 若$H(X) < \infty$有

$$H(X_N|X_1..) \text{ 不增}$$

$$H_N(X) \geq H(X_N|X_1...)$$

$$H_N(X) \text{ 不增}$$

$$H_\infty(X) = \lim H(X_N|X_1...)$$

3.5 Markov 链

DEF 随机序列中任一变量只依赖前一个变量

DEF 齐次 markov 链: p_{ij}与时间无关

DEF p_{ij} 称由 i 转移到 j 的概率, 矩阵形式为 $\boldsymbol{P}$

EQ Kolmogorov 等式 $P^{m+n} = P^m P^n$

Markov 链的平稳分布满足 $\pi^T = \pi^T P$,总存在概率矢量解。 考虑$\lim P^k = \begin{pmatrix} \pi^T \\ ... \end{pmatrix}$的收敛性?

3.6 markov 信源

DEF：有限状态机模型$x_k \to s_k \to ...$,状态s_k的数目为 Markov 链阶数。

PROP markov 性：输出符号只与当前状态有关; Unifiliarity(单线性)：s_k, x_k唯一确定s_{k+1}

3.7 markov 链 N 次扩展源熵的计算

将 mth markov 链变为 mth markov 信源得到 $H(X_1...X_N) = H(S_{m+1}) + \sum_{k=m+1}^{N} H(S_{k+1}|S_k) = H(\pi) + (N-m)\pi^T \cdot h$ 其中$[h_i] = \left[\sum_y -p_{ij}\log p_{ij}\right]$ 从而得到平均符号熵$H_N(X) = N^{-1}(H(\pi) + (N-m)\pi^T h)$

3.8 markov 源符号熵的计算

(假设平稳)markov 链熵率 $H_\infty(X) = \pi^T h = H(S_{k+1}|S_k)$

THE $H(X_k|s_1 = j, X_1...) = \sum_i p(s_k = i|s_1 = j)H(X|s_k = j)$

THE markov 源熵率 $H_\infty(X) = H(X|S) = \sum \pi_i H(X|s = i)$

3.9 相关性与剩余度

平稳有记忆信源可用马氏链近似：

$$H_0(X) = \log n \quad (n\text{为符号集大小})(\text{-1 阶熵})$$

$$H_1(X) = H(X_1)(\text{零阶熵})$$

$$H_2(X) = H(X_2|X_1)$$

$$...(\text{熵率})$$

DEF

$$\begin{cases} \text{剩余度 } \gamma = 1 - \frac{H_\infty}{H_0} \\ \text{信源效率 } \eta = \frac{H_\infty}{H_0} \end{cases}$$

5-无失真信源编码

4.1 分类

分组码: 对信源符号分组编码，码符号序列称为码字; 奇异码: 码字互不相同的编码; 唯一可译码；异前置码(和即时码等价)；

4.2 定长码

N 长信源序列唯一可译条件：$n^N \le r^l$

THE 渐进均分定理

1. 对离散无记忆信源，$\forall \varepsilon, \delta > 0, \exists N_0, \forall N \ge N_0$,

$$|N^{-1} \log p(\vec{x}) + H(X)| < \delta \text{ x 称典型序列,否则 x 出现的概率小于} \varepsilon$$

2. *典型序列概率估计 $p(\vec{x}) = 2^{-N(H(X)\pm\delta)}$
3. *典型序列个数估计 $N_G = (1 - \frac{\sigma}{2} \pm \frac{\sigma}{2})2^{N(H(X)\pm\delta)}$
4. *AEP(平稳遍历信源) (推论)
5. 定长码编码定理: 若 $\frac{l}{N} \log r \ge H(X) + \delta$ 则 N 足够大时存在译码差错无限小
6. 定长码长度下界: $N_0 \ge \frac{\sigma^2}{\varepsilon^2\delta^2} = \left(\frac{\eta}{1-\eta}\right)^2 \frac{1}{H^2(X)} \frac{\sigma^2}{\varepsilon}$

4.3 变长码

THE 异前置码存在的充要条件 $\sum r^{-l_i} \le 1$

THE 长 N 信源序列变长码编码定理 $\frac{H(X)}{\log r} \le \bar{l} < \frac{H(X)}{\log r} + \frac{1}{N}$

4.4 最优编码

THE 二元 Huffman 编码是最小平均码长准则下的最优编码,Shannon 编码是(渐进意义/竞争最优准则)下的最优编码

其他编码方法: 算术编码, LZ 编码

Huffman 编码方法: (1)按概率从大到小排列, 最小 2 项编码末尾大 1 小 0,2 项概率相加得到新信源,重复该过程

算数编码方法:

$$(L_i, H_i) = L_{i-1} + \Delta_{i-1}(l_{i'}, h_{i'})$$

$$\Delta_i = H_i - L_i$$

$$L_0 = 0, H_0 = 1$$

4.5 码率,编码效率,信息传输速率的定义

$$R = \frac{l \log r}{N} \vee \bar{l} \log r$$

$$\eta = \frac{NH}{l \log r} \vee \frac{H}{\bar{l} \log r}$$

$$R_c = \frac{NH}{l} \vee \quad \frac{H}{\bar{l}}$$

4.6 无失真编码定理(1st)

$$R \geq H \Leftrightarrow \text{存在无失真信源编码}$$

6-离散信道及容量

5.1 信道的分类

无噪信道:分为无损信道(1-n,y 无损恢复 x),确定信道(n-1,x 确定 y),无扰信道(1-1)

平稳无记忆信道模型: $\{X, p_{ij}, Y\}$

5.2 单符号离散信道容量

DEF $C = \max_{p(x)} I(X;Y) = \max_{p_i} \sum p_i p_{ij} \log \frac{p_{ij}}{q_j}$, 其中$q_j = \sum_k p_k p_{kj}$

信道容量计算:

(1)对对称信道,有:

THE 对称信道输入等概率时达到容量 $C = H(Y)|_{p(x)=\frac{1}{r}} - H(p_{11}, p_{12}, ...)$

(2) 迭代法求解

(3) 利用充要条件$I(a_i;Y) \leq C \quad$ (当$p(a_i) > 0$时取等)求解

5.3 信道容量代数

THE 对马氏链 $X \to Y \to Z$,$I(X;Z) \leq I(X;Y) \wedge I(X;Z) \leq I(Y;Z)$

EXT 数据处理定理 $I(U;V) \leq I(X;Y)$

级联信道$C \leq C$,并联信道$C' = \sum C_i$;和信道$C' = \log 2^{C_i}$;无记忆 N 次扩展信道 $C^N = NC$

5.4 多维矢量信道容量

THE

$$\begin{cases} H(Y^N|X^N) \leq \sum H(Y_i|X_i) \quad \text{仅当信道无记忆} \\ H(X^N|Y^N) \leq \sum H(X_i|Y_i) \end{cases}$$

THE 对离散无记忆信道 $I(X^N;Y^N) \leq \sum (X_i;Y_i)$

5.5 有约束信道容量

?

7-有噪信道编码

6.1 判决和译码规则

DEF 判决函数 $g(y = b_i) = a^*$

DEF 平均错误率 $P_E = 1 - \sum_y P(X = g(b_i), Y = b_i) \triangleq p(x^* y)$

6.2 最佳判决准则

最大后验概率(maximum a posteriori,map):$g(y) = \mathrm{argmax}_x p(x|y)$

最大似然准则(maximum likelihood,ml): $g(y) = \mathrm{argmax}_x p(y|x)$

6.3 信道编码和最佳译码

DEF 汉明距离: $d_{H(x,y)} = \sum |x_i - y_i|$

DEF (n,k)二元线性分组码:?

THE $d_H \geq 2t + 1 \Leftrightarrow$ 可以纠正 r 个错误

THE 对二元无记忆信道，最大似然准则等价于最小汉明距离准则；对无记忆加性高斯信道，最大似然准则等价于最小欧式距离准则

常见的分组码：重复码、奇偶校验码等,非分组码包括算数编码等

6.4 Fano 不等式

THE Fano 不等式(疑义度上界)

$$H(X|Y) \leq H(p_E; 1 - p_E) + p_E \log(r - 1) \quad \text{仅当 } P(X \neq x^*, Y = y) = \frac{P_E}{r - 1} \text{时取等}$$

6.5 有噪信道编码定理

DEF* $\delta -$ 联合典型序列: $n_{ij} = np_i p_{ij}(1 \pm \delta)$

THE* 联合典型序列的概率估计:$p(\vec{x}, \vec{y}) = 2^{-nH(XY)(1 \pm \delta)}$

THE* 联合典型序列个数估计：若 X,Y 独立,$\vec{x}, \vec{y}$为 $\delta -$ 典型序列，$\vec{x}\vec{y}$构成 $\delta -$ 联合典型序列,则给定$\vec{y}$,$\vec{x}$的个数 $|F_{\vec{y}}| \leq 2^{nH(X|Y) + \delta(H(XY) + H(Y))}$

THE 有噪信道编码定理: 对离散无记忆平稳信道,信道容量 C,信息传输速率 R,若$R < C$ 存在(M,n)码,n 足够大时可使P_E无限小;反之对任意编码$P_E > 0$

THE 信源信道编码定理: 对离散无记忆平稳信道,每秒容量 C,信源每秒熵 H,若$H < C$ 则存在编码系统…

6.6 纠错码技术

*

4-连续信息与连续信源

7.1 差熵的导出:量化逼近法

DEF 差熵(微分熵)将区间[a,b]分为 N 份,

$$H(X)=-\lim_{N\to\infty}\sum_{i=1}^{N}p(x_i)\Delta\log p(x_i)-\lim_{N\to\infty}\log\Delta$$

其中定义差熵$h(X)=-\lim_{N\to\infty}\sum_{i=1}^{N}p(x_i)\Delta\log p(x_i)=\int_a^b -p(x)\log p(x)\,\mathrm{d}x$

例题:高斯信源的熵$h(X)=\frac{1}{2}\log 2\pi e\sigma^2$

类似可定义$h(XY),h(X|Y),I(X;Y)$

7.2 差熵的性质

条件不增性;可加性(链法则)

不具备的性质:非负性(p(x)可以未归一化);一一对应变换下的不变性

THE $y=f(x),f:\mathbb{R}^N\to\mathbb{R}^N$则 $h(Y^N)=h(X^N)-\int\mathrm{d}V(p(\vec{x})\log|\det(J)|)$,式 中$J_{ij'}=\frac{\partial x_i}{\partial y^{j'}}$

EXT $h\left(A\cdot X^N+\alpha\right)=h\left(X^N\right)+\log|\det(A)|$

THE $D(p\|q)\geq 0$

例题:多维相关高斯随机矢量的熵 若

$$p(\vec{x})=\frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\det\left(\sigma_{ij}^2\right)}}\cdot\exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T\left[\sigma_{ij}^2\right]^{-1}(\vec{x}-\vec{\mu})\right)$$

则

$$h(X^N)=\frac{N}{2}\log\left(2\pi e\left(\det\left(\sigma_{ij}^2\right)\right)^{\frac{1}{N}}\right)$$

7.3 平均互信息的性质

对称性;非负性;与差熵的关系;线性变换下的不变性;

7.4 连续最大熵定理

THE 振幅受限的随机变量,均匀分布时取得最大熵

THE 平均功率受限的随机变量,高斯分布时取得最大熵,$h(X)=\frac{1}{2}\log(2\pi eP)$

8-连续信道

8.1 定义

时间离散连续信道:对应随机矢量 ;波形信道:对应随机过程

单符号连续信道容量: $C(\beta)=\max\{I(X;Y)|E[f(x)]\leq\beta\}$

8.2 加性噪声信道容量

DEF $Y=X+Z$,X 与 Z 无关

PROP:…

THE 若输入限制 $f(x)=\overline{x^2}\le\beta=E$,加性高斯噪声方差$\sigma_z$,则$C(E)=\frac{1}{2}\log\left(1+\frac{E}{\sigma_z^2}\right)$

EXT 并联加性高斯噪声信道容量

8.3 加性高斯白噪声(AWGN)信道

DEF 是一种波形信道,满足 $y(t)=x(t)*g(t)+n(t)\quad$ (白噪声 $g(t)=\delta(t)$)

波形信道的离散化:...

THE $C=W\log\left(1+\frac{P}{N_0W}\right)$, 其中$N_0$为单边谱密度(双边为$\frac{N_0}{2}$),$W$为带宽,$P$为输入信号功率限制

THE AWGN 信道编码定理 $R\le C$

记$E_b=\frac{P}{R}$为单位传输速率需要的功率,记功率利用率$\frac{E_b}{N_0}$,频谱利用率$\frac{R}{W}$, 则

$$\frac{R}{W}\le\log\left(1+\frac{E_b}{N_0}\frac{R}{W}\right)$$

$$\frac{E_b}{N_0}\ge\frac{2^{\frac{R}{W}}-1}{\frac{R}{W}}$$

9-信息率失真函数

9.1 R(D)定义和性质

率失真函数的定义: $R(D)=\min_{p(y|x)\in P_D}I(X;Y)$, 其中$D$为失真测度(保真度准则),$P_D=\{p(y|x)|E[d(x,y)]\le D\}$, 其中$E[d(x,y)]=\overline{D}=\sum_{x,y}p_ip_{ij}d_{ij}$, 保真度准则规定$\overline{D}\le D$

率失真函数的性质: 单调性,下凸性

D 的上下界:

$$D_{\min}=\sum_i p_i\min_j d_{ij}$$

$$D_{\max}=\min_j\sum_i p_id_{ij}$$

9.2 常见 R(D)函数

在汉明失真测度$d_{ij}=1-\delta_{ij}$下:

对二元对称?信源$R(D)=H(p)-H(D)\quad(0\le D\le p)$

对 r 元等概信源$R(D)=\log r-D\log(r-1)-H(D)\quad\left(0\le D\le 1-\frac{1}{r}\right)$

9.3 限失真信源编码定理

$$R>R(D)\Leftrightarrow\exists\text{ 信源编码 }E[d]\le D$$

$$C_t > R_{t(D)} \Leftrightarrow \exists \text{ 信源信道编码 } E[d] \leq D$$

十、主要论述

经典信息论的研究对象是信源编码和信道编码

通信的结果是获得信息

改变表示不改变信息量

唯一可译是无失真/满足 Kraft 不等式的充分条件

$l,\bar{l}$的区别

MAP 可由联合概率阵/后验概率阵判决

影响译码错误率的因素: 信道统计特性+信道编译码方法

ML/MAP 等价的条件: 输入等概/统计特性未知

两步编码如何达到一步编码效果: 容量大于信源熵+信源编码效率达到最大+错误率无限小?

率失真函数建立了信息传输率和平均失真度的关系.

十一、主要证明

1. 对 r 元信源X^N做 huffman 编码,等价为一 r 元信源,证明 N 足够大时,为等概信源
2. 证明噪声熵为$H(n)$
3. 证明对离散无记忆强对称信道, 最小汉明距离准则等价于最大似然译码准则.

十二、约定

$$p(x) \leftrightarrow p_X(X=a_i)\text{(i 可以为哑指标也可以为自由指标)}$$