信息论

信号传输与处理的理论基础

例题：第八章部分习题选解

典型随机变量的信息熵

8.1 微分熵。计算下列各密度函数的微分熵 $h(X)=-\int f\ln f$：

(a) 指数密度函数 $f(x)=\lambda e^{-\lambda x}, x\geqslant 0$。

(b) 拉普拉斯密度函数 $f(x)=\frac{1}{2}\lambda e^{-\lambda|x|}$。

(c) X_1 与 X_2 的和的密度函数，其中 X_1 与 X_2 是独立的正态分布，均值为 μ_i，方差为 $\sigma_i^2, i=1,2$。

* 答案【试给出完整的计算】：
* (a) $\log\frac{e}{\lambda}$ (b) $\log\frac{2e}{\lambda}$
* (c) $\frac{1}{2}\log 2\pi e(\sigma_1^2+\sigma_2^2)$ 提示：X_1+X_2也具有正态分布，先确定其方差。

熵与行列式的凹性不等式

8.2 行列式的凹性。设 K_1 和 K_2 为两个 $n\times n$ 对称非负定矩阵。证明下列由樊畿(Ky Fan)[199]给出的结果：

$$|\lambda K_1+\bar{\lambda} K_2|\geqslant|K_1|^{\lambda}|K_2|^{\bar{\lambda}},\ \text{对于 } 0\leqslant\lambda\leqslant 1,\ \bar{\lambda}=1-\lambda$$

其中 $|K|$ 表示 K 的行列式。[提示：先假设 $\mathbf{X}_1\sim N(0,K_1)$，$\mathbf{X}_2\sim N(0,K_2)$，以及 $\theta=$ Bernoulli(λ)，令 $\mathbf{Z}=\mathbf{X}_\theta$，然后利用结论 $h(\mathbf{Z}|\theta)\leqslant h(\mathbf{Z})$。]

* 解答:

（1）设置一个二元离散随机变量θ：$P[\theta=1]=\lambda$, $P[\theta=2]=1-\lambda$, $0<\lambda<1$.

* 再设置一个随机变量$Z=X_\theta$，即$\theta=1$时$Z=X_1$，$\theta=2$时$Z=X_2$，因此

$p(z|\theta=1) = X_1$的概率密度$=N(0,K_1)$

* $p(z|\theta=2) = X_2$的概率密度$=N(0,K_2)$
* 计算Z的均值$m = \int_{-\infty}^{\infty} dz\, z p(z|\theta=1)p[\theta=1] + \int_{-\infty}^{\infty} dz\, z p(z|\theta=2)p[\theta=2] = 0$
* 计算Z^2的均值$m = \int_{-\infty}^{\infty} dz\, z^2 p(z|\theta=1)p[\theta=1] + \int_{-\infty}^{\infty} dz\, z^2 p(z|\theta=2)p[\theta=2]$
* $= \lambda K_1+(1-\lambda)K_2$ 【试完成以上计算】
* 因此 Z的方差 $\sigma_Z^2 = \lambda K_1+(1-\lambda)K_2$

(2) 具有给定方差$\lambda K_1+(1-\lambda)K_2$的随机变量中，Gauss变量具有最大熵

$$\frac{1}{2}\ln(2\pi e)^n|\lambda K_1+(1-\lambda)K_2|$$

熵与行列式的凹性不等式（续）

8.2 行列式的凹性。设 K_1 和 K_2 为两个 $n\times n$ 对称非负定矩阵。证明下列由樊畿(Ky Fan)[199]给出的结果：

$$|\lambda K_1+\bar{\lambda}K_2|\geqslant|K_1|^{\lambda}|K_2|^{\bar{\lambda}}，对于 0\leqslant\lambda\leqslant1,\bar{\lambda}=1-\lambda$$

其中 $|K|$ 表示 K 的行列式。[提示：先假设 $\mathbf{X}_1\sim N(0,K_1)$，$\mathbf{X}_2\sim N(0,K_2)$，以及 $\theta=$ Bernoulli(λ)，令 $\mathbf{Z}=\mathbf{X}_\theta$，然后利用结论 $h(\mathbf{Z}|\theta)\leqslant h(\mathbf{Z})$。]

* 解答：

(1) 设置一个二元离散随机变量θ：$P[\theta=1]=\lambda$, $P[\theta=2]=1-\lambda$, $0<\lambda<1$.

* 再设置一个随机变量 $Z=X_\theta$，即θ=1时 $Z=X_1$，θ=2时 $Z=X_2$，因此

Z的方差 = $\lambda K_1+(1-\lambda)K_2$

(2) 具有给定方差 $\lambda K_1+(1-\lambda)K_2$ 的随机变量中，Gauss变量具有最大熵

$$\frac{1}{2}\ln(2\pi e)^n|\lambda K_1+(1-\lambda)K_2|$$

(3) Z并非Gauss变量，因此

$$\frac{1}{2}\ln(2\pi e)^n|\lambda K_1+(1-\lambda)K_2| \geq H(Z)\geq H(Z|\theta)$$

(4) 【子问题】计算条件熵

$$H(Z|\theta) = \lambda\frac{1}{2}\ln(2\pi e)^n|K_1|+(1-\lambda)\frac{1}{2}\ln(2\pi e)^n|K_2|$$

综合以上结果，结论得证。

有限幅度信号与噪声情形互信息量

8.3 均匀分布噪声。设一个信道的输入随机变量 X 服从区间 $-1/2\leqslant x\leqslant 1/2$ 上的均匀分布，而信道的输出信号为 $Y=X+Z$，其中 Z 是噪声随机变量，服从区间 $-a/2\leqslant z\leqslant a/2$ 上的均匀分布。

(a) 求 $I(X;Y)$ 作为 a 的函数。

* 解(*a*) I(X;Y) = H(Y) – H(Y|X)
* 计算H(Y)，为此需要计算Y的概率密度$p_Y(y)=p_X(x)*p_Z(z)$【即卷积积分】，
* 【试完成计算】结果分两种情形：$a\leqq 1$时

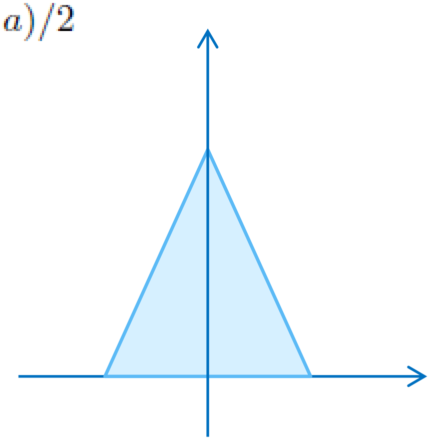
$$p_Y(y)=\begin{cases}(1/2a)(y+(1+a)/2) & -(1+a)/2\le y\le -(1-a)/2\\ 1 & -(1-a)/2\le y\le +(1-a)/2\\ (1/2a)(-y-(1+a)/2) & +(1-a)/2\le y\le +(1+a)/2\end{cases}$$

* $a>1$时

$$p_Y(y)=\begin{cases}y+(a+1)/2 & -(a+1)/2\le y\le -(a-1)/2\\ 1/a & -(a-1)/2\le y\le +(a-1)/2\\ -y-(a+1)/2 & +(a-1)/2\le y\le +(a+1)/2\end{cases}$$

* 因此【试完成计算】$a\leqq 1$时 $H[Y]=a/2$, $a>1$时 $H[Y]=\ln a+1/2a$.
* 此外，$H[Y|X]=H[Z]=\ln a$【试完成计算】，因此

$$I(X;Y)=\begin{cases}a/2-\ln a & \text{if } a\le 1\\ 1/2a & \text{if } a\ge 0\,.\end{cases}$$



$a=1$时的$p_Y(y)$

有限幅度信号与噪声情形互信息量(续)

(b) 对于 $a=1$，当输入信号 X 是峰值约束的时候，即 X 的取值范围限制于 $-1/2 \leqslant x \leqslant 1/2$ 时，求信道容量。为使得互信息 $I(X;Y)$ 达到最大值，X 应该服从什么概率分布？

* 解 (*b*) I(X;Y) = H(Y) – H(Y|X)

  【注意以下的分析是针对X具有任何可能分布的一般情况，X的具体分布正是本小题要解决的内容】
* 由于Y=X+Z在区间[-1,+1]上取值，由于Y在该区间上具有均匀分布时
* H[Y]最大且最大值为1【验证之】，又由于*a*=1时H[Z]=0【参考上页的计算】
* 因此，最大互信息量
* C = *max* I(X;Y) = *max* H(Y) – H(Z) = 1
*
*

连续随机变量的信息熵的尺度变换性质

8.5 尺度性质。设 $h(\boldsymbol{X}) = -\int f(\boldsymbol{x})\log f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$，证明 $h(A\boldsymbol{X}) = \log|\det(A)| + h(\boldsymbol{X})$。

* 求解概要:
* 运用多重积分变量变换的Jaccobi公式:
* $$\iiint d\boldsymbol{x}F(\boldsymbol{x}) = \iiint d\boldsymbol{y}|\partial\boldsymbol{\Phi}_i/\partial\boldsymbol{y}_j|F(\boldsymbol{\Phi}(\boldsymbol{y}))$$
*
* 本题将上述公式应用于线性变换 $\boldsymbol{\Phi}(\boldsymbol{y})=\mathbf{A}\boldsymbol{y}$。【试完成相应的计算】

典型集合

8.10 典型集的形态。令 X_i 为服从 $f(x)$ 的独立同分布序列，其中

$$f(x)=ce^{-x^4}$$

令 $h=-\int f\ln f$。描述典型集 $A_\varepsilon^{(n)}=\{x^n\in\mathcal{R}^n: f(x^n)\in 2^{-n(h\pm\epsilon)}\}$ 的形态。

- 求解概要:
- h=X的信息熵，X^n=$(X_1,\dots,X_n)$是i.i.d.随机序列，相应的典型
- 集按定义是
- $A(n,\varepsilon)$ = $\{X^n$: $2^{-n(h-\epsilon)}\le f(x^n)\le 2^{-n(h+\epsilon)}$ $\}$
- 其中

$$\begin{aligned} f(x^n) &= \prod_{i=1}^{n} f(x) \\ &= \prod_{i=1}^{n} ce^{-x_i^4} \\ &= e^{nln(c)-\sum_{i=1}^{n} x_i^4} \end{aligned}$$

- 因此$(X_1,\dots,X_n)\in A(n,\varepsilon)$当且仅当

$$n(ln(c)+(h-\epsilon)ln(2))\ge\sum_{i=1}^{n}x_i^4\ge n(ln(c)+(h+\epsilon)ln(2))$$

下一单元预告

* 信息论知识在MIMO通信技术中的应用
* 【注】下一单元中的习题均作为思考题，不要求提交。