清华大学深圳研究生院

应用信息论

2018 年春季学期

作业 1

赵丰 2018 年 6 月 25 日

1.1. 设 X 和 Y 是各有均值 m_x, m_y, 方差为 σ_x^2, σ_y^2，且相互独立的高斯随机变量，已知 $U = X + Y, V = X - Y$。试求 $I(U;V)$。

解. U, V 的联合分布是均值为 $[\mu_x + \mu_y, \mu_x - \mu_y]$，协方差矩阵为

$$\Lambda_{U,V} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_x^2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^T = \begin{bmatrix} \sigma_x^2 + \sigma_y^2 & \sigma_x^2 - \sigma_y^2 \\ \sigma_x^2 - \sigma_y^2 & \sigma_x^2 + \sigma_y^2 \end{bmatrix}$$

由多元高斯分布微分熵的公式

$$h(U) = \frac{1}{2}\log(2\pi e(\sigma_x^2 + \sigma_y^2))$$

$U|V = v$ 也是高斯分布, 方差为 $\frac{4\sigma_x^2\sigma_y^2}{\sigma_x^2+\sigma_y^2}$，与 v 无关，因此

$$h(U|V) = \mathbb{E}_V[h(U|V = v)] = \frac{1}{2}\log(2\pi e\frac{4\sigma_x^2\sigma_y^2}{\sigma_x^2 + \sigma_y^2}) \Rightarrow$$

$$\begin{aligned} I(U;V) =& h(U) - h(U|V) \\ =& \frac{1}{2}\log(2\pi e(\sigma_x^2 + \sigma_y^2)) - \frac{1}{2}\log(2\pi e\frac{4\sigma_x^2\sigma_y^2}{\sigma_x^2 + \sigma_y^2}) \\ =& \log\left(\frac{\sigma_x^2 + \sigma_y^2}{2\sigma_x\sigma_y}\right) \end{aligned}$$

1.2. 设有随机变量 X, Y, Z 均取值于 $\{0, 1\}$, 已知
$I(X;Y) = 0, I(X;Y|Z) = 1$。求证 $H(Z) = 1, H(X, Y, Z) = 2$

证明.
$I(X;Y|Z) = H(X|Z) - H(X|Y,Z) \le H(X|Z) \le H(X) \le \log(2) = 1$
所以等号全都成立 $\Rightarrow X \sim B(\frac{1}{2})$。同理可知 $Y \sim B(\frac{1}{2})$。另外
$H(Y|Z) = H(Y) \Rightarrow I(Y;Z) = 0 \Rightarrow H(Z|Y) = H(Z)$

$$\begin{aligned} & H(X|Y,Z) = 0 \\ \iff & H(X,Y,Z) = H(Y,Z) \\ \iff & H(X,Y) + H(Z|X,Y) = H(Y) + H(Z|Y) \\ \iff & 2 + H(Z|X,Y) = 1 + H(Z) \\ \iff & H(Z) = 1 + H(Z|X,Y) \end{aligned}$$

由上式推出 $H(Z) \geq 1$, 又
$H(Z) \leq 1 \Rightarrow H(Z) = 1 \Rightarrow H(X,Y,Z) = 2$ □

1.3. (数据处理不等式) 设有信号 X 经过处理器 A 后获输出 Y,Y 再经处理器 B 后获输出 Z。已知处理器 A 和 B 分别独立处理 X 和 Y。试证：$I(X;Z) \leq I(X;Y)$

证明. $I(X;Z) = H(Z) - H(Z|X) = H(Z); I(X;Y) = H(Y)$ 因为 Z 是 Y 的函数 $\Rightarrow H(Z) \leq H(Y) \Rightarrow I(X;Z) \leq I(X;Y)$ □

1.4. 已知随机变量 X 和 Y 的边际概率密度 $p(a_k, b_j)$ 满足

$$p(a_1) = \frac{1}{2}, p(a_2) = p(a_3) = \frac{1}{4}, p(b_1) = \frac{2}{3}, p(b_2) = p(b_3) = \frac{1}{6}$$

试求能使 $H(X,Y)$ 取得最大值的联合概率密度分布。

解. $H(X,Y) = H(X) + H(Y) - I(X;Y) \leq H(X) + H(Y) = \frac{7}{6} + \log 3$
等号成立当且仅当 X,Y 相互独立 $\Rightarrow p(x,y) = p(x)p(y)$

1.5. 设随机变量 X,Y,Z 满足 $p(x,y,z) = p(x)p(y|x)p(z|y)$。求证
$I(X;Y) \geq I(X;Y|Z)$

证明. 因为 $p(x,y,z) = p(x)p(y|x)p(z|y,x) \Rightarrow p(z|y,x) = p(z|y) \Rightarrow x$ 与 z 关于 y 条件独立 $\Rightarrow I(X;Y|Z) = H(X|Z) - H(X|Y,Z) = H(X|Z) - H(X|Y) \leq H(X) - H(X|Y) = I(X;Y)$ □

1.6. 求证 $I(X;Y;Z) =$
$H(X,Y,Z) - H(X) - H(Y) - H(Z) + I(X;Y) + I(Y;Z) + I(Z;X)$,
其中 $I(X;Y;Z) \triangleq I(X;Y) - I(X;Y|Z)$

证明.

$$\begin{aligned}
I(X;Y;Z) =& I(X;Y) - I(X;Y|Z) \\
=& H(X) + H(Y) - H(X,Y) - (H(X|Z) - H(X|Y,Z)) \\
=& H(X) + H(Y) - H(X,Y) - (H(X,Z) - H(Z)) + H(X,Y,Z) - H(Y,Z) \\
=& H(X,Y,Z) - H(X) - H(Y) - H(Z) + (H(X) + H(Y) - H(X,Y)) \\
&+ (H(Y) + H(Z) - H(Y,Z)) + (H(Z) + H(X) - H(X,Z)) \\
=& H(X,Y,Z) - H(X) - H(Y) - H(Z) + I(X;Y) + I(Y;Z) + I(Z;X)
\end{aligned}$$

□

1.7. 令 $p=(p_1,p_2,\dots,p_a)$ 是一个概率分布，满足 $p_1\ge p_2\ge \dots p_a$, 假设 $\epsilon>0$ 使得 $p_1-\epsilon\ge p_2+\epsilon$ 成立，证明：
$H(p_1,p_2,\dots,p_a)\le H(p_1-\epsilon,p_2+\epsilon,p_3,\dots,p_a)$

证明. 设 $f(\epsilon)=(p_1-\epsilon)\log(p_1-\epsilon)+(p_2+\epsilon)\log(p_2+\epsilon)$。由已知
$0\le\epsilon\le\frac{p_2-p_1}{2}\Rightarrow f'(\epsilon)=\log\frac{p_2+\epsilon}{p_1-\epsilon}\le 0$
$\Rightarrow f(\epsilon)\le f(0)\Rightarrow H(p_1,p_2,\dots,p_a)\le H(p_1-\epsilon,p_2+\epsilon,p_3,\dots,p_a)$ □

1.8. 设 $p_i(x)\sim N(\mu_i,\sigma_i^2)$, 试求相对熵 $D(p_2||p_1)$

解.

$$\begin{aligned}
D(p_2||p_1)&=\int_{\mathbb{R}}p_2(x)\log\frac{p_2(x)}{p_1(x)}dx\\
&=\int_{\mathbb{R}}p_2(x)\left(\log\frac{\sigma_1}{\sigma_2}+\frac{1}{2}(\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{(x-\mu_2)^2}{\sigma_2^2})\log e\right)dx\\
&=\log\frac{\sigma_1}{\sigma_2}+\frac{1}{2}\left(\frac{(\mu_2-\mu_1)^2}{\sigma_1^2}+\frac{\sigma_2^2}{\sigma_1^2}-1\right)\log e
\end{aligned}$$

1.9. 若 $f(x)$ 分别是区间 $(0,0.01),(0,0.5),(0,1),(0,2),(0,5)$ 上均匀分布的分布函数，计算 $f(x)$ 的微分熵。

解. 设 U_t 是 $(0,t)$ 上的均匀分布，则 $h(U_t)=\log t$

- $h(U_{0.01})=\log 0.01$
- $h(U_{0.5})=-1$
- $h(U_1)=0$
- $h(U_2)=1$
- $h(U_5)=\log 5$

1.10. 设

$$\begin{aligned}
p_1(x,y)=&\frac{1}{2\pi\sigma_x\sigma_y}\exp[-\frac{1}{2}(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2})]\\
p_2(x,y)=&\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp[-\frac{1}{2(1-\rho^2)}(\frac{x^2}{\sigma_x^2}-2\rho\frac{xy}{\sigma_x\sigma_y}+\frac{y^2}{\sigma_y^2})]
\end{aligned}$$

试求 $D(p_2||p_1)$ 和 $I(X;Y)$，其中 $X,Y\sim p_2$

解.

$$
\begin{aligned}
D(p_2||p_1) = &\iint_{\mathbb{R}^2} p_2(x,y)\log\frac{p_2(x,y)}{p_1(x,y)}dxdy \\
&= -\frac{1}{2}\log(1-\rho^2) \\
&-\frac{1}{2}(\log e)\iint_{\mathbb{R}^2} p_2(x,y)\left[\frac{\rho^2x^2}{\sigma_x^2(1-\rho^2)}+\frac{\rho^2y^2}{\sigma_y^2(1-\rho^2)}-\frac{2\rho xy}{(1-\rho^2)\sigma_x\sigma_y}\right]dxdy \\
&= -\frac{1}{2}\log(1-\rho^2)
\end{aligned}
$$

$X|Y=y$ 服从高斯分布，方差为 $(1-\rho^2)\sigma_x^2$

$$
\begin{aligned}
I(X;Y) =&h(X)-h(X|Y) \\
=&\frac{1}{2}\log(2\pi e\sigma_x^2)-\frac{1}{2}\log(2\pi e\sigma_x^2(1-\rho^2)) \\
=&\frac{1}{2}\log(\frac{1}{1-\rho^2})
\end{aligned}
$$