Chapter 2

Xuan

2020 年 9 月 14 日

1 条件互信息

在有三个变量的情况下，符号 x_i 与符号对 (y_j, z_k) 之间的互信息量定义为

$$I(x_i; y_j, z_k) = log\frac{p(x_i|y_j, z_k)}{p(x_i)}$$

条件互信息量是在给定 z_k 条件下，x_i 与 y_j 之间的互信息量，定义为

$$I(x_i; y_j|z_k) = log\frac{p(x_i|y_j, z_k)}{p(x_i|z_k}$$

结合上述公式可写为

$$I(x_i; y_j, z_k) = I(x_i; z_k) + I(x_i; y_j|z_k)$$

1.1 相关公式推导

$$\begin{aligned} I(X;Y|Z) &= \sum p(x,y,z)log\frac{p(x,y|z)}{p(x|z)p(y|z)} \\ &= \sum_z D(p(x,y|z)||p(x|z)p(y|z)) * p(z) \end{aligned} \tag{1}$$

$$\begin{aligned} \frac{p(x,y|z)}{p(x|z)p(y|z)} &= \frac{p(x|y,z)p(y|z)p(x)}{p(x)p(x|z)p(y|z)} \\ &= \frac{p(x,y,z)}{p(x)p(y,z)} * \frac{p(x)}{p(x|z)} \end{aligned} \tag{2}$$

将 (2) 带入 (1)

$$(1) = \sum p(x,y,z) log \frac{p(x,y,z)}{p(x)p(y,z)} - \sum p(x,y,z) log \frac{p(x|z)p(z)}{p(x)p(z)} \tag{3}$$

$$\begin{aligned} \sum p(x,y,z) log \frac{p(x|z)p(z)}{p(x)p(z)} &= \sum_{x,y,z} (p(x,z) log \frac{p(x,z)}{p(x)p(z)}) p(y|x,z) \\ &= \sum_{x,z} \sum_{y} \\ &= \sum_{x,z} p(x,z) log \frac{p(x,z)}{p(x)p(z)} \sum_{y} p(y|x,z) \end{aligned} \tag{4}$$

将 (4) 带入 (3)

$$\begin{aligned} (3) &= \sum p(x,y,z) log \frac{p(x,y,z)}{p(x)p(y,z)} - \sum p(x,z) log \frac{p(x,z)}{p(x)p(z)} \\ &= I(X;Y,Z) - I(X;Z) \\ &= H(X) - H(X|Y,Z) - (H(X) - H(X|Z)) \\ &= H(X|Z) - H(X|Y,Z) \end{aligned} \tag{5}$$

2 信息不增性

对于马尔可夫链 $X- > Y- > Z$，有

$$I(X;Z) \leq I(Y;Z)$$

2.1 公式证明

$$\left.\begin{aligned} I(X;Y,Z) &= I(X;Y) + I(X;Z|Y) \\ I(X;Z|Y) &= \sum_{y} D(p(x,z|y)||p(x|y)p(z|y))p(y) = 0 \end{aligned}\right\} I(X;Y,Z) = I(X;Y) \tag{6}$$

$$\begin{aligned} I(X;Y,Z) &= I(X;Z) + I(X;Y|Z) \\ p(x,y,z) = p(y)p(x|y)p(z|y) &\iff p(x,z|y) = p(x|y)p(z|y) \end{aligned} \tag{7}$$

综上所述，$I(X;Y) = I(X;Y,Z) \geq I(X;Z)$

2.2 文氏图证明

见 chapter2.jpg

3 连续信源的熵和互信息

3.1 连续信源熵

$$H_c(x) = -\int_{-\infty}^{\infty} p_x(x) log p_x(x) dx \geq 0$$

3.2 能量确定条件下，$\int_{-infty}^{\infty} x^2 p(x) dx$ 在高斯分布情况下最大化熵率

$$H_c(x) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-m)^2}{2\sigma^2}} log[\frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-m)^2}{2\sigma^2}}] dx \tag{8}$$

3.3 范围确定条件下，均匀分布最大化熵率

变量 X 的幅度取值限定在 [a,b]，则有 $\int_a^b p_x(x)dx = 1$，当任意 $p_x(x)$ 符合平均分条件，

$$p_x(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & other \end{cases}$$

时，信源达到最大熵

$$\begin{aligned} H_c(x) &= -\int_a^b \frac{1}{b-a} log \frac{1}{b-a} dx \\ &= log(b-a) \end{aligned} \tag{9}$$