CIE6020/MAT3350 HW Assignment 2

Due: 23:55, 18 Feb 2019

1. Let X, Y, Z be three random variables with a joint probability mass function $p(x, y, z)$. The relative entropy between the joint distribution and the product of the marginals is

$$D(p(x,y,z)||p(x)p(y)p(z)) = \mathbb{E}\left[\log \frac{p(x,y,z)}{p(x)p(y)p(z)}\right].$$

Expand this in terms of entropies. When is this quantity zero?

Solution:

$$\begin{aligned} D(p(x,y,z)||p(x)p(y)p(z)) &= \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y,z)}{p(x)p(y)p(z)} \\ &= \sum_{x,y,z} p(x,y,z) \log p(x,y,z) \\ &\quad - \sum_{x} p(x) \log p(x) - \sum_{y} p(y) \log p(y) \\ &\quad - \sum_{z} p(z) \log p(z) \\ &= H(X) + H(Y) + H(Z) - H(X,Y,Z). \end{aligned}$$

Clearly, if we have $p(x,y,z) = p(x)p(y)p(z)$, the quantity is zero.

2. Let the random variable X have three possible outcomes $\{a, b, c\}$. Consider two distributions on this random variable:

symbol	$p(x)$	$q(x)$
a	$\frac{1}{2}$	$\frac{1}{3}$
b	$\frac{1}{4}$	$\frac{1}{3}$
c	$\frac{1}{4}$	$\frac{1}{3}$

Calculate $H(p)$, $H(q)$, $D(p||q)$ and $D(q||p)$. Verify that in this case, $D(p||q) \neq D(q||p)$.

Solution:

$$\begin{aligned}
H(p) &= -\frac{1}{2}\log\frac{1}{2} - 2\times\frac{1}{4}\log\frac{1}{4} = \frac{3}{2}.\\
H(q) &= -3\times\frac{1}{3}\log\frac{1}{3} = \log 3.\\
D(p||q) &= \frac{1}{2}\log\frac{1/2}{1/3} + 2\frac{1}{4}\log\frac{1/4}{1/3} = \log 3 - \frac{3}{2}.\\
D(q||p) &= \frac{1}{3}\log\frac{1/3}{1/2} + 2\frac{1}{3}\log\frac{1/3}{1/4} = \frac{5}{3} - \log 3.
\end{aligned}$$

We see that $D(p||q) = \log 3 - \frac{3}{2} \neq \frac{5}{3} - \log 3 = D(q||p)$.

3. Show that $\ln x \geq 1 - \frac{1}{x}$ for $x > 0$, where the equality holds when $x = 1$.

Solution: Let $f(x) = lnx - 1 + \frac{1}{x}$. The derivative $f'(x) = \frac{1}{x} - \frac{1}{x^2}$. We have

- $f(1) = 0$;
- when $x \in (0, 1)$, $f'(x) < 0$;
- when $x > 1$, $f'(x) > 0$.

Therefore, for $x > 0$, $f(x) \geq 0$, i.e., $\ln x \geq 1 - \frac{1}{x}$, where the equality holds when $x = 1$.

4. *Conditioning reduces entropy.* Show that $H(Y|X) \leq H(Y)$ with equality iff X and Y are independent.

Solution: $H(Y) - H(Y|X) = I(X;Y) \geq 0$ where the equality holds when X and Y are independent.

5. Show that $I(X;Y|Z) \geq 0$ with equality iff $X \to Z \to Y$.

Solution: $I(X;Y|Z) = \sum_z p(z)I(X;Y|Z = z) \geq 0$ as mutual information is nonnegative, where the equality holds iff for all z, $I(X;Y|Z = z) = 0$, i.e., for all x, y and z, $p(x, y|z) = p(x|z)p(y|z)$.

6. *Data processing.* Let $X_1 \to X_2 \to X_3 \to \cdots \to X_n$ form a Markov chain, i.e.,

$$p(x_1, x_2, ..., x_n) = p(x_1)p(x_2|x_1)\cdots p(x_n|x_{n-1}).$$

Reduce $I(X_1; X_2, ..., X_n)$ to its simplest form.

Solution: First, the Markov chain implies $p(x_2, \ldots, x_n) = p(x_2)p(x_3|x_2)\cdots p(x_n|x_{n-1})$. Then,

$$
\begin{aligned}
I(X_1; X_2, ..., X_n) &= \sum_{x_1,...,x_n} p(x_1, ..., x_n) \log \frac{p(x_1, ..., x_n)}{p(x_1)p(x_2, ..., x_n)} \\
&= \sum_{x_1,...,x_n} p(x_1, ..., x_n) \log \frac{p(x_1)p(x_2|x_1)\cdots p(x_n|x_{n-1})}{p(x_1)p(x_2)p(x_3|x_2)\cdots p(x_n|x_{n-1})} \\
&= \sum_{x_1,x_2} p(x_1, x_2) \log \frac{p(x_2|x_1)}{p(x_1)} \\
&= I(X_1, X_2).
\end{aligned}
$$

7. Let X and Y be two random variables and let Z be independent of (X, Y). Show that $I(X; Y) \geq I(X; g(Y, Z))$ for any function g.

8. *Bottleneck.* Suppose that a (nonstationary) Markov chain starts in one of n states, necks down to $k < n$ states, and then fans back to $m > k$ states. Thus, $X_1 \to X_2 \to X_3$, that is, $p(x_1, x_2, x_3) = p(x_1)p(x_2|x_1)p(x_3|x_2)$, for all $x_1 \in \{1, 2, ..., n\}, x_2 \in \{1, 2, ..., k\}, x_3 \in \{1, 2, ..., m\}$.

   (a) Show that the dependence of X_1 and X_3 is limited by the bottleneck by proving that $I(X_1; X_3) \leq \log k$.

   (b) Evaluate $I(X_1; X_3)$ for $k = 1$, and conclude that no dependence can survive such a bottleneck.

Solution:

(a) Since $X_1 \to X_2 \to X_3$, by the data processing inequality, $I(X_1; X_3) \leq I(X_1; X_2) \leq H(X_2) \leq \log k$.

(b) By (a), we have $I(X_1, X_3) \leq \log 1 = 0$. Since mutual information is negative, we have $I(X_1, X_3) = 0$, i.e., X_1 and X_3 are independent.