Teoria informacji kolokwium 4 grudnia 2024

Piotr Styczyński[1]

[1]MIM UW (ps386038)

2024-12-08 (updated: CURRENT_DATE)

1 Problem 3

I we use the "triangle" inequality i.e $H(C \mid B) + H(B \mid A) \geq H(C \mid A)$ for:

$$\begin{cases} A = Y \\ B = g(Y) \\ C = f(X) \end{cases} \text{and} \begin{cases} A = X \\ B = f(X) \\ C = g(Y) \end{cases} \tag{1}$$

We get:

$$\begin{cases} H(f(X) \mid g(Y)) + H(g(Y) \mid Y) \geq H(f(X) \mid Y) \\ H(g(Y) \mid f(X)) + H(f(X) \mid X) \geq H(g(Y) \mid X) \end{cases} \implies \begin{cases} H(f(X) \mid g(Y)) + 0 \geq H(f(X) \mid Y) \\ H(g(Y) \mid f(X)) + 0 \geq H(g(Y) \mid X) \end{cases} \tag{2}$$

If we add those equalities side-by-side and move to the right, we get:

$$0 \geq -H(f(X) \mid g(Y)) - H(g(Y) \mid f(X)) + H(f(X) \mid Y) + H(g(Y) \mid X)$$

Let's consider $X = Y$ it can be alternativelly expressed as:

$$\begin{gathered} H(f(X) \mid g(X)) + H(g(X) \mid f(X)) \geq 0 \\ H(f(X), g(X)) - I(f(X); g(X)) \geq 0 \\ H(f(X), g(X)) \geq I(f(X); g(X)) \end{gathered}$$

Now with that if we look at this equation we can observe that for:

1. $f = g = id$: the equality happen
2. $f, g := const$: the equality happen and every term is zeroed
3. $X = Y, f = g$: the equality happen
4. $X \perp Y$: the equality happen

Looking at those observations (and first of the equations) we can conclude that condition implying strict inequality is as follows:

$$\underset{\substack{x,x' \in X \\ y,y' \in Y \\ Pr(X=x \wedge Y=y)>0}}{\exists} Pr(X = x \wedge Y = y) \neq Pr(X = x) * Pr(Y = y) \wedge g(x) = g(x') \wedge f(y) \neq f(y') \quad (3)$$

Now to prove the inequality in the exercice, we add and substract joint entropy of $f(X)$ and $g(Y)$:

$$0 \geq H(f(X), g(Y)) - H(f(X) \mid g(Y)) - H(g(Y) \mid f(X)) + (H(f(X) \mid Y) + H(g(Y) \mid X) - H(f(X), g(Y))) \quad (4)$$

We now use the fact that we can "extend" entropy to joint entropy for functions: $H(f(X) \mid Y) = H(f(X), g(Y) \mid Y)$ (analogical for $H(g(Y) \mid X)$):

$$0 \geq (H(f(X), g(Y)) - H(f(X) \mid g(Y)) - H(g(Y) \mid f(X)))+$$
$$(H(f(X), g(Y) \mid Y) + H(f(X), g(Y) \mid X) - H(f(X), g(Y)))$$

Adding empty term $H(f(X), g(Y) \mid X, Y)$:

$$0 \geq (H(f(X), g(Y)) - H(f(X) \mid g(Y)) - H(g(Y) \mid f(X)))+$$
$$(-H(f(X), g(Y) \mid X, Y) + H(f(X), g(Y) \mid Y) + H(f(X), g(Y) \mid X) - H(f(X), g(Y)))$$

Now we use the fact that $I(A; B) = H(A, B) - H(A \mid B) - H(B \mid A)$

$$0 \geq I(f(X); g(Y)) + (-H(f(X), g(Y) \mid X, Y) + H(f(X), g(Y) \mid Y) + H(f(X), g(Y) \mid X) - H(f(X), g(Y))) \quad (5)$$

We add $I(X; Y)$ on both sides:

$$I(X; Y) \geq I(f(X); g(Y))+$$
$$(I(X; Y) - H(f(X), g(Y) \mid X, Y) + H(f(X), g(Y) \mid X) + H(f(X), g(Y) \mid Y) - H(f(X), g(Y)))$$

We assume $Z := f(X), g(Y)$ and look at the terms in bracket:

$$I(X; Y) - H(Z \mid X, Y) + H(Z \mid X) + H(Z \mid Y) - H(Z) \quad (6)$$

We expand by definition of mutual information and use the "condition inversion rule"
Rule: $H(Y \mid X) = H(X \mid Y) + H(Y) - H(X)$

$$\begin{aligned} &(H(X) + H(Y) - H(X, Y)) \\ -&(H(X, Y \mid Z) + H(Z) - H(X, Y)) \\ +&(H(X \mid Z) + H(Z) - H(X)) \\ +&(H(Y \mid Z) + H(Z) - H(Y)) - H(Z) \end{aligned}$$

After adding everything up:

$$-H(X, Y \mid Z) + H(X \mid Z) + H(Y \mid Z) \quad (7)$$

Which is just a mutual information:

$$-H(X, Y \mid Z) + H(X \mid Z) + H(Y \mid Z) = I(X; Y \mid Z) \quad (8)$$

If we summarize that we get desired relation:

$$I(X;Y) \geq I(f(X);g(Y)) + I(X;Y \mid f(X),g(Y)) \tag{9}$$

As a side note (as I often used it and I'm not entirely sure if that is considered "basic" enough, so provide quick side-note for completeness sake) we can deliver "triangle" inequality from:

$$H(Y \mid X) + H(Z \mid Y) \geq H(Y \mid X) + H(Z \mid Y,X) = H(Y,Z \mid X) \geq H(Z \mid X) \tag{10}$$