Quantum Information, Lecture 4 HW

Guillermo Abad

10 November, 2023

1 Let$(X^n, Y^n) \sim p(x^n, y^n)$

We can prove it, by separating the original expression into:

$$\begin{aligned}
\sum_{i=1}^{n} I\left(X_{i+1}^{n}; Y_i \mid Y_1^{i-1}\right) &= \sum_{i=1}^{n}\left[I\left(X_{i+1}^{n}; Y_1^{i}\right) - I\left(X_{i+1}^{n}; Y_1^{i-1}\right)\right] \\
&=^{*} \sum_{i=1}^{n}\left[I\left(X_{i+1}^{n}; Y_1^{i}\right) - I\left(X_{i}^{n}; Y_1^{i-1}\right) + I\left(X_i; Y_1^{i-1} \mid X_{i+1}^{n}\right)\right] \\
&=^{**} \sum_{i=1}^{n} I\left(X_i; Y_1^{i-1} \mid X_{i+1}^{n}\right)
\end{aligned}$$

where in $=^*$, we have used the second provided Hint for the application of the chain rule of mutual information, on the right term:

$$-I\left(X_{i+1}^{n}; Y_1^{i-1}\right) = -I\left(X_{i}^{n}; Y_1^{i-1}\right) + I\left(X_i; Y_1^{i-1} \mid X_{i+1}^{n}\right)$$

and in $=^{**}$, for canceling the two left terms, we have moved the sum index of the second one, obtaining:

$$\begin{aligned}
\sum_{i=1}^{n}\left[I\left(X_{i+1}^{n}; Y_1^{i}\right) - I\left(X_{i}^{n}; Y_1^{i-1}\right)\right] &= \sum_{i=1}^{n} I\left(X_{i+1}^{n}; Y_1^{i}\right) - \sum_{i=0}^{n-1} I\left(X_{i+1}^{n}; Y_1^{i}\right) \\
&= I\left(X_{n+1}^{n}; Y_1^{n}\right) - I\left(X_{1}^{n}; Y_1^{0}\right) \\
&= I\left(\emptyset; Y_1^{n}\right) - I\left(X_{1}^{n}; \emptyset\right) = 0
\end{aligned}$$

2 Inequalities

a) Answer: $\leq$

$$H(X \mid Z) \leq H(X, Y \mid Z) = H(X \mid Y, Z) + H(Y \mid Z) \leq H(X \mid Y) + H(Y \mid Z)$$

b) Answer: $\geq$

If X, Y are independent:

$$H(X+Y) \geq H(X+Y \mid Y) = H(X \mid Y) + H(Y \mid Y) = H(X)$$

c) Answer: $\leq$

From the given expression:

$$p(y_1, y_2 \mid x_1, x_2) = p(y_1 \mid x_1)\, p(y_2 \mid x_2)$$

we can easily see that X_1 to X_2 to Y_2 or that Y_1 to X_1 to X_2 form Markov chains. Therefore:

$$\begin{aligned}
I(X_1, X_2; Y_1, Y_2) &= I(X_1, X_2; Y_1) + I(X_1, X_2; Y_2 \mid Y_1) \\
&= I(X_1; Y_1) + I(X_2; Y_1 \mid X_1) + I(X_2; Y_2 \mid Y_1) + I(X_1; Y_2 \mid Y_1, X_2) \\
&\leq I(X_1; Y_1) + I(X_2; Y_1 \mid X_1) + I(X_2; Y_2) + I(X_1; Y_2 \mid X_2) \\
&= I(X_1; Y_1) + I(X_2; Y_2)
\end{aligned}$$

where at the first and last inequalities, we used that Y_1 to X_1 to X_2 to Y_2 is a Markov chain!

d) Answer: $\geq$

From the given expression $p(x_1, x_2) = p(x_1)\, p(x_2)$, we see that X_1 and X_2 are independent. Therefore:

$$\begin{aligned}
I(X_1, X_2; Y_1, Y_2) &= I(X_1; Y_1, Y_2) + I(X_2; Y_1, Y_2 \mid X_1) \\
&= I(X_1; Y_1) + I(X_1; Y_2 \mid Y_1) + I(X_2; Y_2 \mid X_1) + I(X_2; Y_1 \mid X_1, Y_2) \\
&\geq I(X_1; Y_1) + I(X_1; Y_2 \mid Y_1) + I(X_2; Y_2) + I(X_2; Y_1 \mid Y_2) \\
&\geq I(X_1; Y_1) + I(X_2; Y_2)
\end{aligned}$$

where at the first inequality we used the independence of X_1 and X_2.

3 Z channel

The Z channel has a conditional PMF $p(y \mid x)$:

$$p(0 \mid 0) = 1 \quad p(1 \mid 1) = \frac{1}{2} \quad \text{and} \quad p(0 \mid 1) = \frac{1}{2}$$

which if we define the probability of $p(x = 1)$ as p, gets us:

$$p(0,0) = 1 - p \quad p(1,0) = 0 \quad p(1,1) = \frac{p}{2} \quad \text{and} \quad p(0,1) = \frac{p}{2}$$

now, using $I(y \mid x) = -p(y, x) \log \frac{p(y,x)}{p(x)}$ we obtain:

$$I(0 \mid 0) = I(1 \mid 0) = 0 \quad \text{and} \quad I(0 \mid 1) = I(1 \mid 1) = \frac{p}{2}$$

and so $H(Y \mid X) = \sum I(y \mid x) = p$. On the other hand, $p(y) = \sum_x p(y|x)p(x)$ is:

$$p(0) = 1(1 - p) + \frac{1}{2}p = 1 - \frac{p}{2} \quad \text{and} \quad p(1) = \frac{1}{2}p = \frac{p}{2}$$

This means that the mutual information is:

$$I(X;Y) = H(Y) - H(Y \mid X) = H\left(\frac{p}{2}\right) - p = -\frac{p}{2}\log\frac{p}{2} - \left(1 - \frac{p}{2}\right)\log\left(1 - \frac{p}{2}\right) - p$$

Finally, the channel capacity will be the maximum of this quantity:

$$0 = \frac{dI(X;Y)}{dp} = \frac{\log\left(\frac{1}{2} - \frac{p}{4}\right) - \log(p)}{log(4)} = \frac{1}{2}\log\left(\frac{1}{2p} - \frac{1}{4}\right) \quad \Rightarrow \quad p = \frac{2}{5}$$

which is:

$$C = I(X;Y)\big|_{p=\frac{2}{5}} \simeq 0.3219 \text{ bits}$$

4 The Noisy typewriter channel

Now, in the Noisy Typewriter channel we have a conditional PDF given by:

$$p(y \mid x) = \frac{1}{2}\delta_{x,y} + \frac{1}{2}\delta_{x-1,y} \quad x, y = 1, 2, \ldots, N$$

And then the joint probability will be:

$$p(y, x) = p(y \mid x)p_X(x) = \frac{1}{2}\delta_{x,y}p(x) + \frac{1}{2}\delta_{x-1,y}p_X(x)$$

from where we can obtain:

$$p_Y(y) = \sum_x p(y, x) = \sum_x \left(\frac{1}{2}\delta_{x,y}p_X(x) + \frac{1}{2}\delta_{x-1,y}p_X(x)\right) = \frac{1}{2}\left(p_X(y) + p_X(y+1)\right)$$

Now, the entropy of Y will then be:

$$H(Y) = -\sum_y \frac{1}{2}\left(p_X(y) + p_X(y+1)\right)\log\left(\frac{1}{2}\left(p_X(y) + p_X(y+1)\right)\right)$$

We can also evaluate the conditional entropy now:

$$\begin{aligned} H(Y \mid X) &= -\sum_{x,y} p(y, x)\log p(y \mid x) \\ &= -\sum_y \sum_x \left(\frac{1}{2}\delta_{x,y}p_X(x) + \frac{1}{2}\delta_{x-1,y}p_X(x)\right)\log\left(\frac{1}{2}\delta_{x,y} + \frac{1}{2}\delta_{x-1,y}\right) \\ &= -\frac{1}{2}\sum_y \left(p_X(y) + p_X(y+1)\right)\log\frac{1}{2} = \sum_y p_X(y) = 1 \end{aligned}$$

and the mutual information of X and Y is:

$$I(X; Y) = H(Y) - H(Y \mid X) = H(Y) - 1$$

So, we now only need to find the maximums of $H(Y)$, which at first sight, one would guess that a homogenous distribution, with $p(x) = 1/26$ for the 26 cases would do the trick, and evaluating it, we get:

$$H(Y) = -\sum_0^{26} \frac{\frac{1}{26} + \frac{1}{26}}{2}\log\left(\frac{\frac{1}{26} + \frac{1}{26}}{2}\right) = -\sum_0^{26} \frac{1}{26}\log\left(\frac{1}{26}\right) = \log_2(26)$$

which is the maximum entropy for a random variable with 26 outcomes, so we have found our first maximum!

For the second one, we can also think that given the symmetry that the system has, a probability with only the odd or even x having $p(x) = 2/26$ and the others $p(x) = 0$, would also maximize it! And its also easy to evaluate this, by:

$$H(Y) = -\sum_0^{26} \frac{0 + \frac{2}{26}}{2}\log\left(\frac{0 + \frac{2}{26}}{2}\right) = -\sum_0^{26} \frac{1}{26}\log\left(\frac{1}{26}\right) = \log_2(26)$$

finding our other maximums!

To end, let's compute the channel capacity:

$$C = \max_{p(x)} I(X; Y) = \max_{p(x)} H(Y) - 1 = \log_2 26 - 1 = \log_2 13$$