QIT Homework Lecture 4

Jesús Mosteiro García

October 2023

1. **Let** $(X^n, Y^n) \sim p(x^n, y^n)$. Prove:

$$\sum_{i=1}^{n} I(X_{i+1}^n; Y_i | Y_1^{i-1}) = \sum_{i=1}^{n} I(X_i; Y_1^{i-1} | X_{i+1}^n)$$

where $X_i^j = (X_i, ..., X_j), Y_i^j = (Y_i, ..., Y_j)$ and $X_{n+1}^n = Y_1^0 = \varnothing$.
Hint: Note that by application of the chain rule of mutual information:

$$I(X_{i+1}^n; Y_1^i) = I(X_{i+1}^n; Y_1^{i-1}) + I(X_{i+1}^n; Y_i | Y_1^{i-1})$$

and also:

$$I(X_i^n; Y_1^{i-1}) = I(X_{i+1}^n; Y_1^{i-1}) + I(X_i; Y_1^{i-1} | X_{i+1}^n)$$

Solution:
We can rearrange the hints to obtain:

$$\begin{gathered} I(X_{i+1}^n; Y_1^{i-1}) = I(X_{i+1}^n; Y_1^i) - I(X_{i+1}^n; Y_i | Y_1^{i-1}) \\ I(X_{i+1}^n; Y_1^{i-1}) = I(X_i^n; Y_1^{i-1}) - I(X_i; Y_1^{i-1} | X_{i+1}^n) \\ I(X_{i+1}^n; Y_i | Y_1^{i-1}) - I(X_{i+1}^n; Y_1^i) = I(X_i; Y_1^{i-1} | X_{i+1}^n) - I(X_i^n; Y_1^{i-1}) \end{gathered} \tag{1}$$

Inserting our summation, we know that:

$$\sum_{i=1}^{n} [I(X_{i+1}^n; Y_i | Y_1^{i-1}) - I(X_{i+1}^n; Y_1^i)] = \sum_{i=1}^{n} [I(X_i; Y_1^{i-1} | X_{i+1}^n) - I(X_i^n; Y_1^{i-1})] \tag{2}$$

As such, our original equality holds if:

$$\sum_{i=1}^{n} I(X_{i+1}^n; Y_1^i) = \sum_{i=1}^{n} I(X_i^n; Y_1^{i-1}) \tag{3}$$

On the left side, when $i = n$, we have:

$$I(X_{n+1}^n; Y_1^n) = I(\varnothing; Y_1^n) = 0 \tag{4}$$

Similarly, the right side equals 0 when $i = 1$:

$$I(X_1^n; Y_1^0) = I(X_1^n; \varnothing) = 0 \tag{5}$$

Thus,

$$\sum_{i=1}^{n-1} I(X_{i+1}^n; Y_1^i) = \sum_{i=2}^{n} I(X_i^n; Y_1^{i-1}) \tag{6}$$

Substituting $j = i + 1$ on the left side, we show that the equality holds.

$$\sum_{j=2}^{n} I(X_j^n; Y_1^{j-1}) = \sum_{i=2}^{n} I(X_i^n; Y_1^{i-1}) \tag{7}$$

2. **Inequalities.** Label each of the following statements with $=, \leq, \geq$. Justify your answers.

a. $H(X|Z)$ versus $H(X|Y) + H(Y|Z)$.
Hint: Consider $H(X, Y|Z)$.

Solution:

We know that

$$H(X|Z) \leq H(X, Y|Z) \tag{8}$$

because inserting Y may only increase the entropy, and that

$$H(X|Y, Z) \leq H(X|Y), \tag{9}$$

since the extra information from Z cannot increase the entropy.
We now make use of

$$H(X|Y, Z) = H(X, Y|Z) - H(Y|Z) \tag{10}$$

and get

$$H(X, Y|Z) \leq H(X|Y) + H(Y|Z). \tag{11}$$

From there follows that

$$H(X|Z) \leq H(X|Y) + H(Y|Z). \tag{12}$$

b. $H(X + Y)$ versus $H(X)$ when X and Y are independent.
Hint: Consider $H(X + Y|Y)$.

Solution:

When Y is given, the only entropy contributing to the total is that of X. Since the variables are independent, this doesn't give any more information on X, and so:

$$H(X) = H(X + Y|Y) \tag{13}$$

Of course, we know that

$$H(X+Y|Y) \leq H(X+Y), \tag{14}$$

and thus:

$$H(X) \leq H(X+Y) \tag{15}$$

c. $I(X_1, X_2; Y_1, Y_2)$ versus $I(X_1; Y_1) + I(X_2; Y_2)$,
if $p(y_1, y_2|x_1, x_2) = p(y_1|x_1)p(y_2|x_2)$.
Hint: Apply the chain rule sequentially and note that
$Y_1 \leftrightarrow X_1 \leftrightarrow X_2 \leftrightarrow Y_2$ form a Markov chain.

Solution:

We can express the mutual information in terms of entropy:

$$I(X_1, X_2; Y_1, Y_2) = H(Y_1, Y_2) - H(Y_1, Y_2|X_1, X_2) \tag{16}$$

And, since we are working with a Markov chain:

$$\begin{gathered} I(X_1, X_2; Y_1, Y_2) = H(Y_1, Y_2) - H(Y_1|X_1, X_2) - H(Y_2|X_1, X_2) \\ I(X_1, X_2; Y_1, Y_2) = H(Y_1, Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \end{gathered} \tag{17}$$

Finally, we get to the following inequality:

$$\begin{gathered} I(X_1, X_2; Y_1, Y_2) \leq H(Y_1) - H(Y_1|X_1) + H(Y_2) - H(Y_2|X_2) \\ I(X_1, X_2; Y_1, Y_2) \leq I(X_1; Y_1) + I(X_2; Y_2) \end{gathered} \tag{18}$$

d. $I(X_1, X_2; Y_1, Y_2)$ versus $I(X_1; Y_1) + I(X_2; Y_2)$,
if $p(x_1, x_2) = p(x_1)p(x_2)$.
Hint: Apply the chain rule sequentially again.

Solution:

This case is similar to c, with the difference that we now know that X_1 and X_2 are independent. This separates our Markov chain, implying that Y_1 and Y_2 are also indepenent. Thus,

$$H(Y_1, Y_2) = H(Y_1) + H(Y_2), \tag{19}$$

and so Inequality 18 becomes an equality

$$\begin{gathered} I(X_1, X_2; Y_1, Y_2) = H(Y_1) - H(Y_1|X_1) + H(Y_2) - H(Y_2|X_2) \\ I(X_1, X_2; Y_1, Y_2) = I(X_1; Y_1) + I(X_2; Y_2) \end{gathered} \tag{20}$$

3. **Z channel.** The Z channel has binary input and output alphabets, and conditional pmf $p(0|0) = 1, p(1|1) = p(0|1) = 0.5$. Find the capacity C.

Solution:
Our target function is:

$$C = \max_p I(i;o) = \max_p \frac{p}{2}(\log_2 \frac{1}{2-p} + \log_2 \frac{1}{p}) + (1-p)\log_2 \frac{2}{2-p} \tag{21}$$

where p is the probability of the input being 1.
Taking the derivative of the mutual information with respect to p:

$$\frac{dI}{dp} = \frac{1}{2}(\log_2 \frac{1}{4p} - \log_2 \frac{1}{2-p}) = \frac{1}{2}\log_2 \frac{2-p}{p} - 1 \tag{22}$$

The root of this function is the probability at which the mutual information is maximal:

$$\log_2 \frac{2-p}{p} = 2 \quad \Rightarrow \quad p = \frac{2}{5} \tag{23}$$

Thus, we find the capacity:

$$C = 0.2(\log_2 0.625 + \log_2 2.5) + 0.6\log_2 1.25 \approx 0.3219 \tag{24}$$

4. **The Noisy typewriter channel.** Compute the capacity of a noisy typewriter channel where $p(o_n|i_n) = p(o_{n+1}|i_n) = 0.5$, with $0 \leq n < 26$. Provide two solutions for the maximizing pmf $p(x)$, that yield capacity.

Solution:
It is clear by observation that maximizing the entropy of the input independently yields $\log_2 26$ (uniform distribution across the 26 inputs). Since the distribution of every output once an input is known boils down to a Bernoulli with $p = 0.5$, the conditional entropy is 1:

$$C = \max_{p(x)} I(i;o) = \max_{p(x)} H(i) - H(i|o) = \log_2 26 - 1 = \log_2 13 \tag{25}$$

Another not so obvious solution is to distribute the input evenly making use of every other character in the alphabet. This causes the input entropy to drop to $\log_2 13$, but also means that an output is fully controlled by one and only one input, resulting in $H(i|o) = 0$:

$$C = \max_{p(x)} I(i;o) = \max_{p(x)} H(i) - H(i|o) = \log_2 13 - 0 = \log_2 13 \tag{26}$$