Spring 2016

6.441 - Information Theory

Homework 8

Due: Tue, Apr 19, 2016 (in class)

`Prof. Y. Polyanskiy`

1 Reading (optional)

1. Read [1, Chapter 7]

2 Exercises

NOTE: Each exercise is 10 points. Only 3 exercises per assignment will be graded. If you submit more than 3 solved exercises please indicate which ones you want to be graded.

1 (Information density and types.) Let $P_{Y|X} : \mathcal{A} \to \mathcal{B}$ be a DMC and let P_X be some input distribution. Take $P_{X^nY^n} = P_{XY}^n$ and define $i(a^n; b^n)$ with respect to this $P_{X^nY^n}$.

1. Show that $i(x^n; y^n)$ is a function of only the "joint-type" $\hat{P}_{XY}$ of (x^n, y^n), which is a distribution on $\mathcal{A} \times \mathcal{B}$ defined as

$$\hat{P}_{XY}(a, b) = \frac{1}{n}\#\{i : x_i = a, y_i = b\}\,,$$

where $a \in \mathcal{A}$ and $b \in \mathcal{B}$. Therefore $\{\frac{1}{n}i(x^n; y^n) \geq \gamma\}$ can be interpreted as a constraint on the joint type of (x^n, y^n).

2. Assume also that the input x^n is such that $\hat{P}_X = P_X$. Show that

$$\frac{1}{n}i(x^n; y^n) \leq I(\hat{P}_X, \hat{P}_{Y|X})\,.$$

The quantity $I(\hat{P}_X, \hat{P}_{Y|X})$, sometimes written as $I(x^n \wedge y^n)$, is an *empirical mutual information*[1]. Hint:

$$\mathbb{E}_{Q_{XY}}\left[\log \frac{P_{Y|X}(Y|X)}{P_Y(Y)}\right] = D(Q_{Y|X}||Q_Y|Q_X) + D(Q_Y||P_Y) - D(Q_{Y|X}||P_{Y|X}|Q_X) \qquad (1)$$

2 Consider the following (memoryless) channel. It has a side switch U that can be in positions `ON` and `OFF`. If U is on then the channel from X to Y is $BSC(\delta)$ and if U is off then Y is Bernoulli (1/2) regardless of X. The receiving party sees Y but not U. A design constraint is that U should be in the `ON` position no more than the fraction s of all channel uses, $0 \leq s \leq 1$. Questions:

1. One strategy is to put U into `ON` over the first sn time units and ignore the tail $(1-s)n$ readings of Y. What is the maximal rate in bits per channel use achievable with this strategy?

[1] Invented by V. Goppa for his maximal mutual information (MMI) decoder:

$$\hat{W} = \operatorname*{argmax}_{i=1,\ldots,M} I(c_i \wedge y^n)\,.$$

2. How much does the communication rate increase if the encoder is allowed to modulate the U switch together with the input X (while still satisfying the s-constraint on U).
3. Now assume nobody has access to U, which is random, independent of X, memoryless across different channel uses and

$$P[U = \texttt{ON}] = s.$$

Find capacity.

3 (Capacity-cost at $P = P_0$.) Recall that we have shown that for stationary memoryless channels and $P > P_0$ capacity equals $f(P)$:

$$C(P) = f(P)\,, \tag{2}$$

where

$$P_0 \triangleq \inf_{x \in \mathcal{A}} \mathsf{c}(x) \tag{3}$$

$$f(P) \triangleq \sup_{X:\mathbb{E}[c(X)] \le P} I(X;Y)\,. \tag{4}$$

Show:

1. If P_0 is not admissible, i.e., $c(x) > P_0$ for all $x \in \mathcal{A}$, then $C(P_0)$ is undefined (even $M = 1$ is not possible)
2. If there exists a unique x_0 such that $\mathsf{c}(x_0) = P_0$ then

$$C(P_0) = f(P_0) = 0\,.$$

3. If there are more than one x with $\mathsf{c}(x) = P_0$ then we still have

$$C(P_0) = f(P_0)\,.$$

4. Give example of a channel with discontinuity of $C(P)$ at $P = P_0$. (Hint: select a suitable cost function for the channel $Y = (-1)^Z \cdot \mathrm{sign}(X)$, where Z is Bernoulli and $\mathrm{sign} : \mathbb{R} \to \{-1, 0, 1\}$)

4 Consider a stationary memoryless additive non-Gaussian noise channel:

$$Y_i = X_i + Z_i\,, \qquad \mathbb{E}[Z_i] = 0, \quad \mathrm{Var}[Z_i] = 1$$

with the input constraint

$$\|x^n\|_2 \le \sqrt{nP} \quad \iff \quad \sum_{i=1}^{n} x_i^2 \le nP\,.$$

1. Prove that capacity $C(P)$ of this channel satisfies

$$\frac{1}{2}\log(1+P) \le C(P) \le \frac{1}{2}\log(1+P) + D(P_Z \| \mathcal{N}(0,1))\,,$$

where P_Z is the distribution of the noise. (Hints: Gaussian saddle point and the golden formula $I(X;Y) \le D(P_{Y|X} \| Q_Y | P_X)$.)

2. If $D(P_Z||\mathcal{N}(0,1)) = \infty$ (Z is very non-Gaussian), then it is possible that the capacity is infinite. Consider Z is ± 1 equiprobably. Show that the capacity is infinite by a) proving the maximal mutual information is infinite; b) giving an explicit scheme to achieve infinite capacity.

5 (Time varying channel, Problem 9.12 [1]) A train pulls out of the station at constant velocity. The received signal energy thus falls off with time as $1/i^2$. The total received signal at time i is

$$Y_i = \left(\frac{1}{i}\right) X_i + Z_i,$$

where $Z_1, Z_2, \ldots \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$. The transmitter constraint for block length n is

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2(w) \leq P, \quad w \in \{1, 2, \ldots, 2^{nR}\}.$$

Using Fano's inequality, show that the capacity C is equal to zero for this channel.

6 Consider the additive noise channel with $\mathcal{A} = \mathcal{B} = \mathbb{F}_2$ and $P_{Y^n|X^n} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ specified by

$$Y^n = X^n + Z^n\,,$$

where $Z^n = (Z_1, \ldots, Z_n)$ is a stationary Markov chain with $P_{Z_2|Z_1}(0|1) = P_{Z_2|Z_1}(1|0) = \tau$. Find the Shannon capacity

$$C = \lim_{\epsilon \to 0+} \liminf_{n \to \infty} \frac{1}{n} \log M^*(n, \epsilon)\,.$$

(Hint: your proof should work for an arbitrary stationary ergodic noise process $Z^\infty = (Z_1, \ldots))$. Can the capacity be achieved by linear codes?

References

[1] T. Cover and J. Thomas, *Elements of Information Theory,* Second Edition, Wiley, 2006

MIT OpenCourseWare
https://ocw.mit.edu

6.441 Information Theory
Spring 2016

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.