Information-Theoretic System Analysis and Design

Homework Assignment

January 10, 2018

The homework has to be delivered in paper form (handwritten, LaTeX, etc.) by **February 15th, 2018**. By the same time, your simulation code should have been sent to `geiger@ieee.org`. You may work alone or in pairs. Do not forget to add your names and matriculation numbers on a cover sheet.

Problem 1 (/10): Assume that $X = (X_1, X_2)$ is jointly Gaussian with zero mean and with covariance matrix $\underline{\mathbf{C}}_X$. Connect the differential entropy of X_1 with its second moment, the conditional differential entropy of X_1 given X_2 with the minimum mean-squared error obtained by estimating X_1 from X_2, and the mutual information between X_1 and X_2 with the covariance coefficient between X_1 and X_2.

Octave Problem 3 (/15): Let X be a Gaussian RV with zero mean and unit variance. Implement the Lloyd-algorithm to design a quantizer with a resolution of 2 bits. Calculate the entropy of the quantizer output Y. What do you observe? Repeat the procedure for a resolution constraint of 3, 4, and 5 bits.

Problem 6 (/10): Assume that X is a one-dimensional RV with even PDF, i.e., $f_X(x) = f_X(-x)$. Assume further that g has even symmetry, and is bijective on $[0, \infty)$ and $(-\infty, 0]$. Show that the information loss satisfies $L(X \to Y) = 1$.

(Octave) Problem 8 & 9 (/15): Assume X is uniformly distributed on $[-a+m, a+m]$ and assume that $g(x) = |x|$.

- Compute the information loss $L(X \to Y)$ as a function of the expected value m of X.
- Implement a numerical estimator of $L(X \to Y)$ as described in Section 2.3.3 of the course notes. Plot the computed information loss as a function of m, for $a = 1$ and $m \in [-1.5, 1.5]$ and compare it to the analytical results.

Problem 10 (/5): Show that the order of two systems has an influence on the information loss of the cascade. **Hint:** Take a uniformly distributed input $X \sim \mathcal{U}([-a, a])$, and let the two systems be a rectifier $g(x) = |x|$ and an offset device $h(x) = x + m$ for some constant $m \neq 0$. **Second Hint:** You may use the results from Problems 6 and 8.

Problem 20 (/15): Assume that the distribution of X has an absolutely continuous component and a discrete component. The absolutely continuous component is uniform on $[-a, a]$, with $P_X^{ac}(\mathcal{X}) = 0.6$. The discrete component has mass points at 0 and a, with $P_X^d(0) = P_X^d(a) = 0.2$. Let $Y = |X - m|$. Compute the information loss $L(X \to Y)$ as a function of m.

Problem 28 (/15): Let X be uniformly distributed on $[-b, b]$. Let $Y = g(X)$, where g is the center clipper with clipping region $[-c, c]$, $c < b$:

$$g(x) = \begin{cases} 0, & x \in [-c, c] \\ x, & \text{otherwise.} \end{cases}$$

Compute the information dimension of Y and the $d(Y)$-dimensional entropy of Y. Compute the conditional information dimension $d(X|Y)$ and show that $d(X) = d(X|Y) + d(Y)$ in this case.

Problem 32 (/10): The half-wave rectifier is defined as

$$g(x) = \begin{cases} x, & x \geq 0 \\ 0, & \text{else} \end{cases}. \tag{1}$$

Let P_e be the minimum reconstruction error, i.e.,

$$P_e := \min_{r\colon \mathcal{Y}\to\mathcal{X}} \mathbb{P}(X \neq r(Y)).$$

In other words, you may choose a deterministic reconstructor $r\colon \mathcal{Y} \to \mathcal{X}$, that minimizes the reconstruction error probability. This reconstructor may depend on the distribution of X, but it may of course not depend on the realizations of X. Show that the relative information loss satisfies $l(X \to Y) = P_e$ for a continuously distributed input X.

Problem 33 (/5): Assume a one-dimensional continuous input, i.e., let X have a PDF f_x. Which system causes higher information loss? The full-wave rectifier or the half-wave rectifier? Justify your answer!

Problem 34 (/5): Find an input RV X such that the full-wave rectifier causes higher information loss than the half-wave rectifier.

Problem 42 (/10): Let S and N be independent Gaussian RVs with zero mean and unit variance. Show that, for $X = S + N$, S and N are not conditionally independent given X.

Problem 43 (/10): Let $S = (\tilde{S}, 0)$ and $N = (0, \tilde{N})$, where $\tilde{S}$ and $\tilde{N}$ are independent Gaussian RVs with zero mean and unit variance. Show that S and N are independent. Show that, for $X = S + N$, S and N are conditionally independent given X.

Octave Problem (/25): Download the two files `S.mat` and `X.mat` from the course website. These two files contain 100.000 realizations of the RVs S and X, respectively, where S is a one-dimensional RV and where X is three-dimensional satisfying

$$X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} S + \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}.$$

The coefficients a_i are unknown, and the noise components N_i are Gaussian and independent (and independent of S), but not necessarily having the same variance. Perform the following tasks:

- For every i, compute the mutual informations $I(X_i, S)$. Which coordinate contains most information about S?
- Compute the sample covariance matrix $\underline{\hat{\mathbf{C}}}_X = \frac{1}{L}\underline{\mathbf{X}}^T\underline{\mathbf{X}}$, where L is 100.000 and where $\underline{\mathbf{X}}$ is the $L \times 3$ matrix containing the realizations of X.
- Compute the eigenvalues and eigenvectors of $\underline{\hat{\mathbf{C}}}_X$.
- We now want to compress X to a one-dimensional RV Y, and we try by performing PCA. In other words, we let $Y = \mathbf{w}^T X$, where $\mathbf{w}$ is the eigenvector of $\underline{\hat{\mathbf{C}}}_X$ corresponding to its largest eigenvalue. Compute $\mathbf{w}$.
- With $\mathbf{w}$ computed above, compute Y and $I(Y, S)$. What can you observe? Did PCA work in this case? Give at least one possible explanation.