Information Theory for Data Science

Assignment 3

Prof. Giorgio Taricco

Politecnico di Torino – DET

Introduction

- Include in the report all necessary elements for the evaluation of your work
- Include the code, preferably written in MATLAB, properly commented, explained (like you would for a user manual), and referred to in the report

1 – Distortion (12 points)

- Consider the random variable X with pdf:

$$f_X(x) = \frac{1}{2} e^{-|x|}.$$

- Given N, h (N is a positive odd integer, h is a positive real number) calculate analytically the distortion based on the following thresholds and quantized values:

$$(2k - N - 1)h \Big|_{k=1}^{N}, \qquad (2k - N - 2)h \Big|_{k=1}^{N+1}$$

- Plot the corresponding rate distortion function for $N = 3,5,7, \dots, 31$, $h = \frac{5}{N}, \frac{10}{N}, \frac{20}{N}$ after suitable transformation of N into R
- Comment the result

2 – Distortion in two dimensions (8 points)

- Consider the random vector $(X, Y)^T$ with pdf:

$$f_{XY}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}.$$

- Given N, h (N is a positive odd integer, h is a positive real number), quantize **both** the horizontal and vertical axes by using the following thresholds and quantized values:

$$(2k - N - 1)h\Big|_{k=1}^{N}, \qquad (2k - N - 2)h\Big|_{k=1}^{N+1}$$

Hint: use the fact that X, Y are independent

- Write a (preferably) MATLAB program to calculate the distortion as a function of N, h
- Plot the corresponding rate distortion function for $N = 3,5,7, \dots, 31$, $h = \frac{1}{N}, \frac{3}{N}, \frac{10}{N}$ after suitable transformation of N into R
- Comment the result

3 – Lloyd algorithm (12 points)

- Consider the random variable X with pdf:

$$f_X(x) = \frac{1}{2} e^{-|x|}.$$

- Apply the Lloyd algorithm and write a program to find the N thresholds, the $(N+1)$ quantization intervals and quantization levels, and the average quantization error
- The equations required must be derived and written in closed form with all simplifications in the report
- Initialize with the symmetric intervals:

$$\left(-\infty, -\frac{N-1}{2}\right), \dots, \left(\frac{N-1}{2}, \infty\right)$$

- For example,
  - if $N = 4$, they are: $\left(-\infty, -\frac{3}{2}\right), \left(-\frac{3}{2}, -\frac{1}{2}\right), \left(-\frac{1}{2}, \frac{1}{2}\right) \left(\frac{1}{2}, \frac{3}{2}\right), \left(\frac{3}{2}, \infty\right)$
  - if $N = 5$: $(-\infty, -2), (-2, -1), (-1,0), (0,1), (1,2), (2, \infty)$
- In the report include the numerical results for the following values of N: $10,20,100$