University of Bergen

Information theory
INF242

Second mandatory exercise

Student :
Arthur Gautier

November 8, 2023

Contents

1 Introduction

The exercise is to make a program in Python that implements the Blahut-Arimoto algorithm for computing the capacity

$$C = \max_{f_X} I(X, Y)$$

of a discrete memoryless channel (DMC) with input X and output Y described by the channel transition probability distribution $f_{Y|X}(y|x)$.

2 Implementation

We first need to initialized a distribution, let's do it uniformly :

```
fx = np.ones((1, m)) / m
```

In an infinite loop we start by calculating $r^{<r>}(x|y) = \frac{f_X^{<r>}(x) f_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} f_X^{<r>}(x') f_{Y|X}(y|x')}$

```
for iteration in range(10000):
    fxt = fx.reshape(-1, 1) #transpose
    r_x_y = fxt * f_y_x / np.sum(fxt * f_y_x, axis=0)
```

Next, afterward, we calculate $f_X^{<r+1>}(x) = \frac{\prod_{y \in \mathcal{Y}} r^{<r>}(x|y)^{f_{Y|X}(y|x)}}{\sum_{x' \in \mathcal{X}} \prod_{y \in \mathcal{Y}} r^{<r>}(x'|y)^{f_{Y|X}(y|x')}}$ in a same loop.

```
    f_x_plus_1 = np.prod(np.power(r_x_y, f_y_x), axis=1)
    f_x_plus_1 = f_x_plus_1 / np.sum(f_x_plus_1)
```

At the end of each iterations we compute $\|f_x^{<r+1>} - f_x^{<r>}\|$ to check if the norm of this difference is small enough to stop the loop.

```
    difference = np.linalg.norm(f_x_plus_1 - fx)
    fx = f_x_plus_1
    if difference < 1e-12:
        # stopping condition
        break
```

With the output of the loop we can now compute the capacity

$$\begin{aligned} C = \max_{f_X} I(X, Y) &= \max_{f_X(x)} \max_{r(x|y)} \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} f_X(x) f_{Y|X}(y|x) \log \left(\frac{r(x|y)}{f_X(x)} \right) \\ &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} f_X(x) f_{Y|X}(y|x) \log \left(\frac{r^*(x|y)}{f_X(x)} \right) \end{aligned}$$

```
# Calculate the capacity
c = 0
for i in range(m):
    if fx[i] > 0:
        c += np.sum(fx[i] * f_y_x[i, :] *
                    np.log2(r_x_y[i, :] / fx[i] + 1e-16))
return c
```

3 Applications

3.1 Channel A

$$f_{Y|X}(y|x) = \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix}$$

$C = 0.029049405545331426$

3.2 Channel B

$$f_{Y|X}(y|x) = \begin{bmatrix} 1.0 & 0.0 \\ 0.3 & 0.7 \end{bmatrix}$$

$C = 0.5036919334848172$

3.3 Channel C

$$f_{Y|X}(y|x) = \begin{bmatrix} 0.7 & 0.0 & 0.3 \\ 0.21 & 0.68 & 0.11 \\ 0.0 & 0.2 & 0.8 \end{bmatrix}$$

$C = 0.6418067659013819$

3.4 Channel D

$C = 0.6322355905586438$

3.5 Channel E

$C = 0.5087646986882709$