Assignment 3: HMM

Luca Brena

November 2019

1 Exercise 1

- $z_k = 1$ we are selecting the k-th curve

$$p(x) = \sum_{k=1}^{K} p(x \wedge z_k = 1) = \sum_{k=1}^{K} p(z_k = 1)p(x \mid z_k = 1) = \sum_{k=1}^{K} \pi_k P(x; \lambda_k)$$

- $$\gamma_{n,k} = p(z_{n,k} = 1 \mid x_n) = \frac{\pi_k P(x_n; \lambda_k)}{\sum_{j=1}^{K} \pi_j P(x_j; \lambda_j)}$$

- For the N instances

$$m_c = \sum_{i=1}^{N} \gamma_{i,c}$$

$$m = \sum_{i=1}^{c} m_i$$

To update the c-th π

$$\pi_c = \frac{m_c}{m}$$

To update the c-th λ

$$\lambda_c = \frac{1}{m_c} \sum_{i=1}^{N} \gamma_{i,c} x_i$$

2 Exercise 2

The HMM associated with the problem is rapresented in the figure 1

- (to semplify the notation I use 1 instead of Urn1 and 2 instead of Urn2)

$$S = \{Urn1, Urn2\} = \{1, 2\}$$

$$O = \{b, r, y\}$$

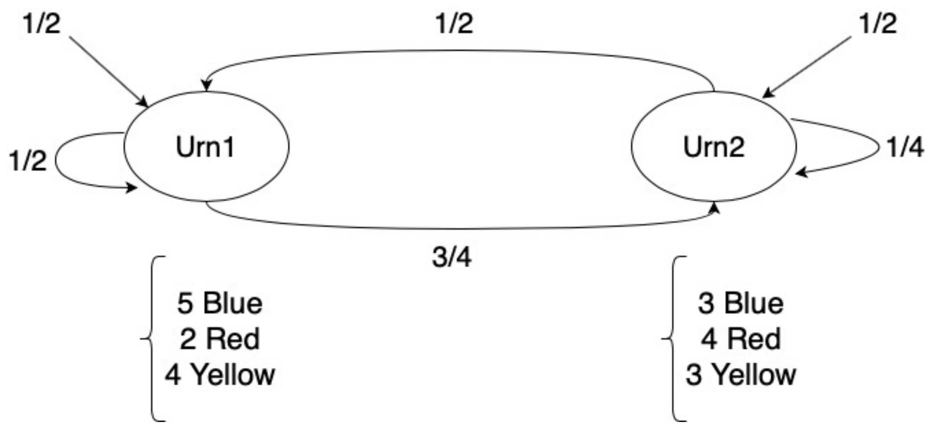

Figure 1: HMM

$$\pi = \{0.5, 0.5\}$$
$$A = \begin{bmatrix} 0.5 & 0.5 \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$$
$$B = \begin{bmatrix} \frac{5}{11} & \frac{2}{11} & \frac{4}{11} \\ \frac{3}{10} & \frac{4}{10} & \frac{3}{10} \end{bmatrix}$$

- $$p(1, 2, 1 \mid y, r, b) = \frac{p(y, r, b \mid 1, 2, 1)p(1, 2, 1)}{p(y, r, b)} = \frac{p(1, y, 2, r, 1, b)}{p(y, r, b)}$$

  To compute the numerator:

  $$\alpha_0(1) = \pi_1 \cdot B_1(y) = 0.5 \cdot \frac{4}{11} = 0.18$$
  $$\alpha_1^*(2) = \alpha_0(1) \cdot A_{1,2} \cdot B_2(r) = 0.18 \cdot 0.5 \cdot \frac{4}{10} = 0.036$$
  $$\alpha_2^*(1) = \alpha_1^*(2) \cdot A_{2,1} \cdot B_1(b) = 0.036 \cdot \frac{3}{4} \cdot \frac{5}{11} = 0.0123$$

  To compute the denominator
  For every possible state i:

  $$\alpha_0(i) = \pi_i B_i(O_0)$$
  $$\alpha_0(1) = \pi_1 B_1(y) = 0.5 \cdot \frac{4}{11} = \frac{2}{11}$$
  $$\alpha_0(2) = \pi_2 B_2(y) = 0.5 \cdot \frac{3}{10} = \frac{3}{20}$$

For every possible state i and j:

$$\alpha_1(i) = \sum_{j=0}^{S} \alpha_0(j) \cdot A_{j,i} \cdot B_i(O_1)$$

$$\alpha_1(1) = \alpha_0(1) \cdot A_{1,1} \cdot B_1(r) + \alpha_0(2) \cdot A_{2,1} \cdot B_1(r) = \frac{2}{11} \cdot 0.5 \cdot \frac{2}{11} + \frac{3}{20} \cdot \frac{3}{4} \cdot \frac{2}{11} = 0.0370$$

$$\alpha_1(2) = \alpha_0(1) \cdot A_{1,1} \cdot B_2(r) + \alpha_0(2) \cdot A_{2,1} \cdot B_2(r) = \frac{2}{11} \cdot 0.5 \cdot \frac{4}{10} + \frac{3}{20} \cdot \frac{1}{4} \cdot \frac{4}{10} = 0.0514$$

For every possible state i and j:

$$\alpha_2(i) = \sum_{j=0}^{S} \alpha_1(j) \cdot A_{j,i} \cdot B_i(O_2)$$

$$\alpha_2(1) = \alpha_1(1) \cdot A_{1,1} \cdot B_1(b) + \alpha_1(2) \cdot A_{2,1} \cdot B_1(b) = 0.0370 \cdot 0.5 \cdot \frac{5}{11} + 0.0514 \cdot \frac{3}{4} \cdot \frac{5}{11} = 0.0259$$

$$\alpha_2(2) = \alpha_1(1) \cdot A_{1,2} \cdot B_2(b) + \alpha_1(2) \cdot A_{2,2} \cdot B_2(b) = 0.0370 \cdot 0.5 \cdot \frac{3}{10} + 0.0514 \cdot \frac{1}{4} \cdot \frac{3}{10} = 0.009405$$

$$p(y, r, b) = \sum_{i=1}^{S} \alpha_2(i) = \alpha_2(1) + \alpha_2(2) = 0.0259 + 0.009405 = 0.035305$$

So, finally

$$\frac{p(1, y, 2, r, 1, b)}{p(y, r, b)} = \frac{0.0123}{0.035305} = 0.3484$$

- We have to compute the following probability:

$$\operatorname*{argmax}_{x,y,z} p(x, y, z \mid r, y, b) = \operatorname*{argmax}_{x,y,z} \frac{p(x, y, z, r, y, b)}{p(r, y, b)}$$

  Considering the denominator constant we can use the proportional form:

$$\operatorname*{argmax}_{x,y,z} p(x, y, z, r, y, b)$$

  Let's use Viterbi's algorithm
  For $t = 0$:

$$\delta_0(x) = \pi_x \cdot B_x(O_0)$$

$$\delta_0(1) = \pi_1 \cdot B_1(r) = \frac{2}{11} \cdot 0.5 = 0.09$$

$$\delta_0(2) = \pi_2 \cdot B_2(r) = \frac{4}{10} \cdot 0.5 = 0.2$$

For $t = 1$:

$$\delta_1(y) = \max_x[\delta_{t0}(x) \cdot A_{x,y}] \cdot B_y(O_1)$$

$$\delta_1(1) = \max_x[\delta_{t0}(1)\cdot A_{1,1}; \delta_{t0}(2)\cdot A_{2,1}]\cdot B_1(y) = \max_x[0.09\cdot 0.5; 0.2\cdot\frac{3}{4}]\cdot\frac{4}{11} = 0.0545$$

$$\psi_1(1) = \delta_{t0}(2) \cdot A_{2,1}$$

$$\delta_1(2) = \max_x[\delta_{t0}(1)\cdot A_{1,2}; \delta_{t0}(2)\cdot A_{2,2}]\cdot B_2(y) = \max_x[0.09\cdot 0.5; 0.2\cdot\frac{1}{4}]\cdot\frac{3}{10} = 0.015$$

$$\psi_1(2) = \delta_{t0}(2) \cdot A_{2,2}$$

For $t = 2$:

$$\delta_2(y) = \max_x[\delta_{t1}(x) \cdot A_{x,y}] \cdot B_y(O_2)$$

$$\delta_2(1) = \max_x[\delta_{t1}(1)\cdot A_{1,1}; \delta_{t1}(2)\cdot A_{2,1}]\cdot B_1(b) = \max_x[0.0545\cdot 0.5; 0.015\cdot\frac{3}{4}]\cdot\frac{5}{11} = 0.0124$$

$$\psi_2(1) = \delta_{t1}(1) \cdot A_{1,1}$$

$$\delta_2(2) = \max_x[\delta_{t1}(1)\cdot A_{1,2}; \delta_{t1}(2)\cdot A_{2,2}]\cdot B_2(b) = \max_x[0.0545\cdot 0.5; 0.015\cdot\frac{1}{4}]\cdot\frac{3}{10} = 0.00817$$

$$\psi_2(2) = \delta_{t1}(1) \cdot A_{1,2}$$

Now we can backpropagating the results, we pick the max in $t = 2$ then we follow the path backwards using ψ. The set of most likely status sequence is $\{2, 1, 1\}$

3 Exercise 3

Refer to ex3.py

4 Exercise 4

- Considering $f(x) = \frac{y_2 - y_1}{2}$:
  1. Train set: {(-1,0), (0,1)}, Validation set:{(1,0)}

     $$f(x) = \frac{1}{2} \qquad e_1 = (0 - \frac{1}{2})^2 = \frac{1}{4}$$

  2. Train set: {(-1,0), (1,0)}, Validation set:{(0,1)}

     $$f(x) = 0 \qquad e_2 = (1 - 0)^2 = 1$$

3. Train set: {(0,1), (1,0)}, Validation set:{(-1,0)}

$$f(x) = -\frac{1}{2} \qquad e_3 = (0 + \frac{1}{2})^2 = \frac{1}{4}$$

So we can say that the l.o.u cross validation of the model is:

$$E = \frac{1}{3}(\frac{1}{4} + 1 + \frac{1}{4}) = \frac{1}{2}$$

- Considering $f(x) = ax + b$, (x_v, y_v) denotes the validation points

1. Train set: {(-1,0), (0,1)}, Validation set:{(1,0)}

$$f(x) = x + 1 \qquad e_1 = (y_v - f(x_v))^2 = (0 - 2)^2 = 4$$

2. Train set: {(-1,0), (1,0)}, Validation set:{(0,1)}

$$f(x) = 0 \qquad e_2 = (y_v - f(x_v))^2 = (1 - 0)^2 = 1$$

3. Train set: {(0,1), (1,0)}, Validation set:{(-1,0)}

$$f(x) = -x + 1 \qquad e_3 = (y_v - f(x_v))^2 = (0 - 2)^2 = 4$$

So we can say that the l.o.u cross validation of the model is:

$$E = \frac{1}{3}(4 + 4 + 1) = 3$$