Homework 3 Recitation: Forward and Viterbi Algorithms

Sezai Artun Ozyegin

May 11, 2019

Content

Content

Introduction

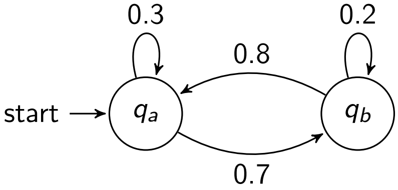
- The formulas in this presentation are from [1].
- There are 3 tasks in Hidden Markov Models(HMM)
  1. Evaluation Task
  2. Decoding Task
  3. Learning Task
- This homework focuses only on Evaluation and Decoding tasks.
- You are asked to implement **Forward** and **Viterbi** algorithms for evaluation and decoding tasks, respectively.

Content

Hidden Markov Models

- **Markov assumption:** The next state is dependent only upon the current state.
- State transition probabilities are independent of the actual time at which transitions take place.
- Current output is statistically independent of the previous outputs.

State transtition matrix



- State transition matrix A:

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}.$$

Observation Probability Matrix, Initial State Probabilities

- Example observation probability matrix B:

$$B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}.$$

- Example initial state probabilites π:

$$\pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Content

Forward Algorithm

- Let $\alpha_t(i)$ be the probability of the partial observation until t, when it terminates at the state i.

$$\alpha_t(i) = p(o_1, ..., o_t, q_t = i | A, B, \pi).$$

- Then, the following recursion holds:

$$\alpha_{t+1}(j) = b_j(o_{t+1}) \sum_{i=1}^{N} a_{ij} \alpha_t(i),$$

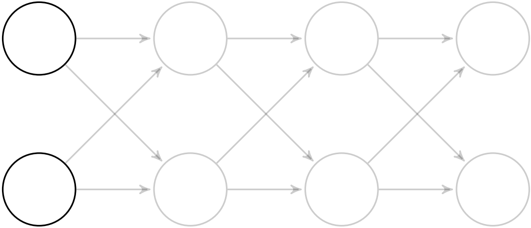
where the initialization is

$$\alpha_1(j) = \pi_j b_j(o_1).$$

- We can find $p(O|A, B, \pi)$ by adding all $\alpha_T(i)$'s for $i = 1, ..., N$.

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

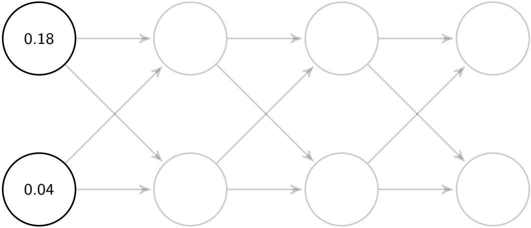


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

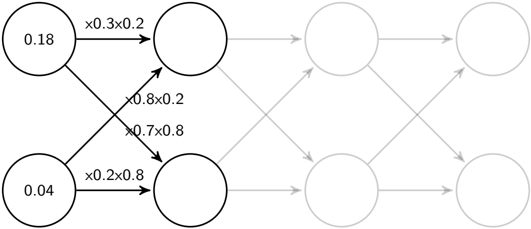


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

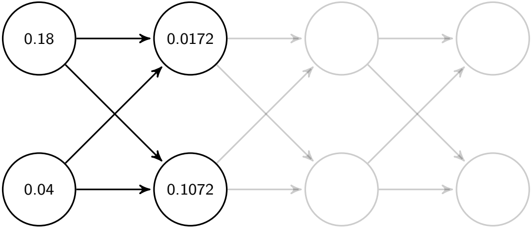


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

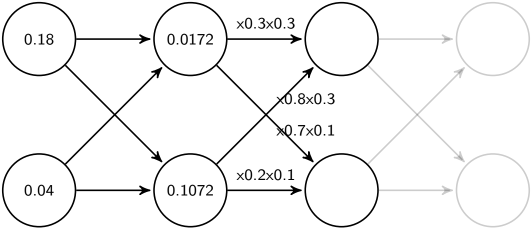


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

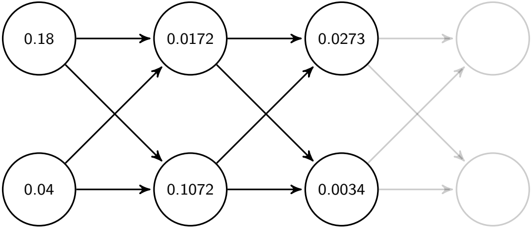


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

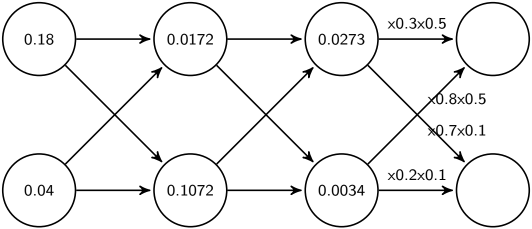


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

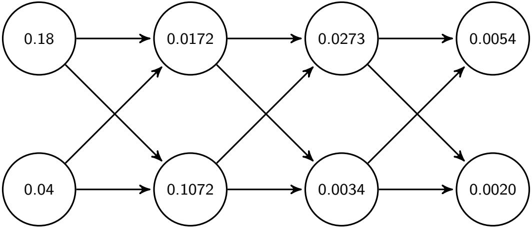


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

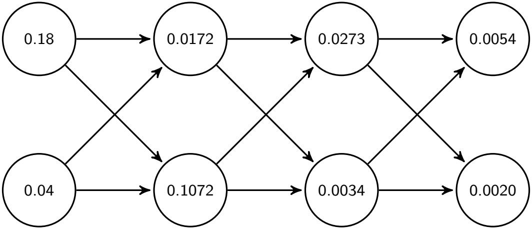


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.



$$p(O|A, B, \pi) = 0.0054 + 0.0020 = 0.74 \tag{1}$$

Content

Viterbi Algorithm

- Let $\delta_t(i)$ be the highest probability that the partial observation sequence and state sequence until t, when it terminates at the state i.

$$\delta(i) = \max_{q_1, q_2, \dots q_{t-1}} p(q_1, \dots, q_{t-1}, q_t = i, o_1, \dots, o_t | A, B, \pi).$$

- Then, the following recursion holds:

$$\delta_{t+1}(j) = b_j(o_{t+1}) \max_i (a_{ij}\delta_t(i)),$$

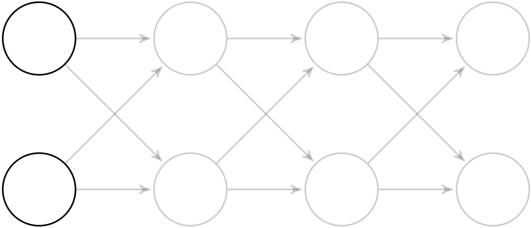
where the initialization is

$$\delta_1(j) = \pi_j b_j(o_1).$$

- By always keeping pointer back to the winning state, we can find the most likely state sequence.

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

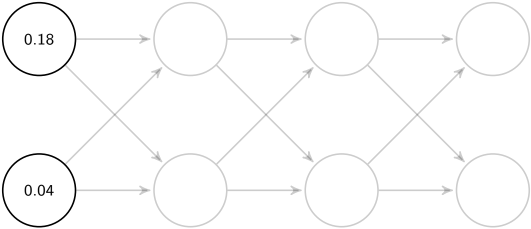


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

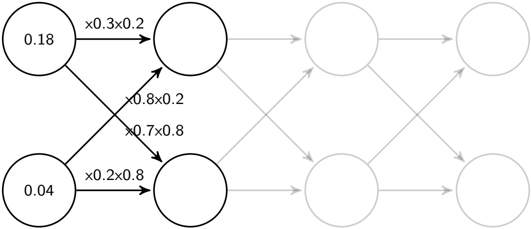


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

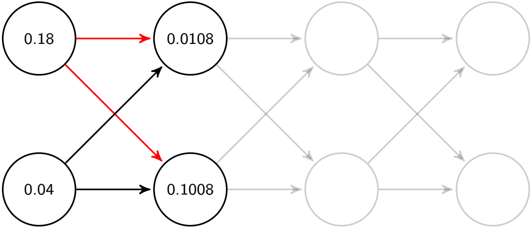


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

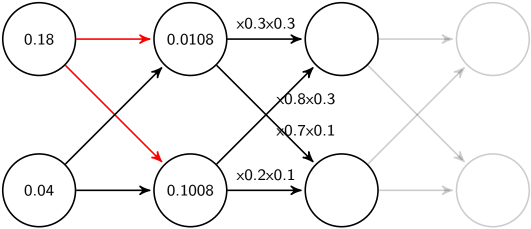


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

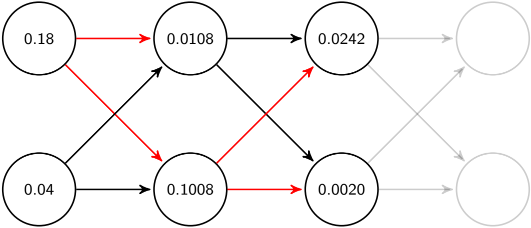


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

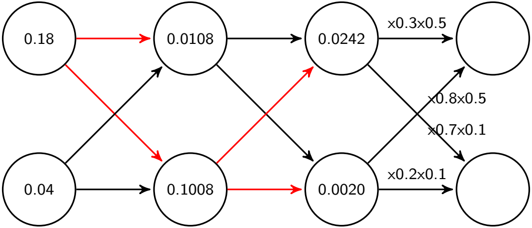


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

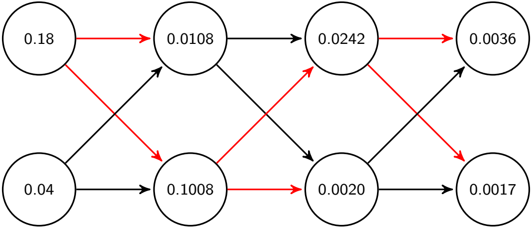


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.

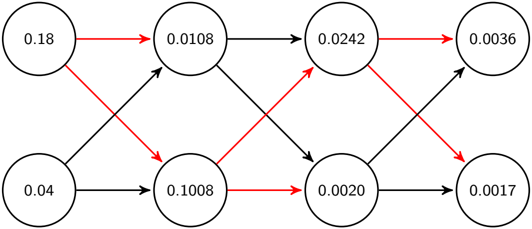


- State transition matrix A, observation probability matrix B, initial state probabilities π are shown below.

$$A = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}, B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \qquad \pi = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}.$$

Trellis Diagram

- For observation sequence $O = [1, 2, 1, 0]$, let's create a Trellis Diagram.



- The highest probability among the right-most nodes is 0.0036.
- If we trace the red path backwards we get the most probable state sequence: $[0, 1, 0, 0]$.

Content

Conclusion

- No report is required for the homework.
- Only need to fill the **forward** and **viterbi** functions provided in the homework files.
- You can test your code using the tester provided. However, even if your code passes all the test, additional tests are going to be applied. Therefore, you may not get full points.
- Further details are mentioned in the homework pdf. You can ask questions on COW or ODTUCLASS. You can send me an email anytime.

References I

Narada Warakagoda.
A hybrid ann-hmm asr system with nn based adaptive preprocessing m.sc. thesis.
Available at
http://jedlik.phy.bme.hu/~gerjanos/HMM/hoved.html.