Tutorial 5 - Markov Chains and Hidden Markov Models

COMP9418 – Advanced Topics in Statistical Machine Learning

Lecturer: Gustavo Batista

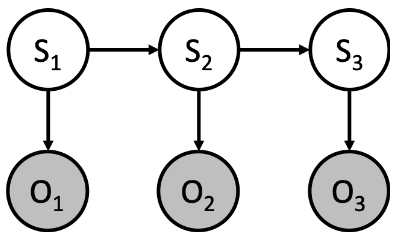
Lecture: Markov Chains and Hidden Markov Models

Topic: Questions from lecture topics

Last revision: Monday 12$^{\text{th}}$ October, 2020 at 05:58

Question 1

Answer the questions considering the Hidden Markov Model (HMM) shown in the following figure:



a. If each state has k different values, each observation has m different values, and the chain is length d, how many parameters are necessary to define this HMM?

b. Show the factorization for the joint distribution $P(S_1, S_2, S_3, O_1, O_2, O_3)$.

c. What conditional independences hold in this HMM?

Answer

a. We will need three probability distribution tables to specify this model. The first one is the initial state distribution, the second the transition probabilities and the final one the emission probabilities. The initial state is a table with k entries. Therefore, it requires $k - 1$ parameters. The transition table lists the probabilities among all possible combinations of state values. Therefore this table has k^2 entries. However, since it is a conditional probability table, the probabilities conditioned to each state must sum to one, and we will need $k^2 - k$ parameters. Finally, the emission table has km entries, where m is the number of emission values. Since it is also a conditional probability table conditioned to the states, it requires $km-k$ parameters. In total, we have $k-1+k^2-k+km-k = k^2+km-k-1$ parameters. Notice that we share those parameters along the chain. Therefore, the number of parameters is independent of the length (d) of the model.

b. $P(S_1, S_2, S_3, O_1, O_2, O_3) = P(S_1)P(O_1|S_1)P(S_2|S_1)P(O_2|S_2)P(S_3|S_2)P(O_3|S_3)$. From the factorization, we can observe that the present state (S_t) is only conditioned to the previous one (S_{t-1}). In a similar way, the current observation (O_t) is only conditioned to the current state (S_t).

c. We can use d-separation to show that the observation of a state S_t separates the chain in two independent parts: past and future, i.e., $S_1, \dots, S_{t-1}, O_1, \dots, O_{t-1} \perp S_{t+1}, \dots, S_n, O_{t+1}, \dots, O_n | S_t$

Question 2

Lisa is given a fair coin C_1 and asked to flip it eight times in a row. Lisa also has a biased coin C_2 with probability 0.8 of landing heads. All we know is that Lisa flipped the fair coin initially, but we believe that she intends to switch to the biased coin and that she tends to be 10% successful in performing the switch. Suppose that we observe the outcome of the eight coin flips and want to find out whether Lisa managed to perform a coin switch and when.

a. Describe a Bayesian network and a corresponding query that solves this problem.

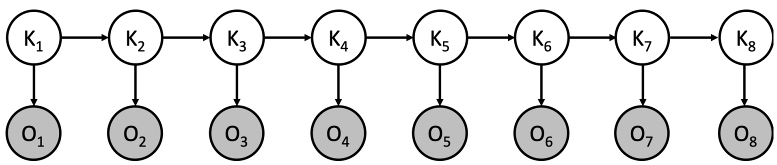
b. Write the probability tables necessary to model this problem.

c. Use the source code developed in the practical part of this tutorial to find the solution to this problem, assuming that the flips came out as follows:

   i. tails, tails, tails, heads, heads, heads, heads, heads.
   ii. tails, tails, heads, heads, heads, heads, heads, heads.

Answer

a. The Bayesian network is a Hidden Markov model (HMM) of the following form:



where K_t are the hidden states that can assume one of two values: **biased** or **fair**. O_t are the observed variables that can assume the values **heads** or **tails**.

To answer this query, we need to indicate the most likely states for the hidden variables. Therefore, this is an MLE query that can be answered with the Viterbi algorithm.

b. We need a transition probability table and an emission probability table to define this HMM. It is also useful to describe the initial state.

The transition probability table is the following:

K_{t-1}	K_t -	$P(K_t\|K_{t-1})$
fair	fair	0.9
fair	biased	0.1
biased	fair	0.0
biased	biased	1.0

Here, we are assuming that once Lisa switched from a fair to a biased coin, we will never return to a fair coin.

The emission probability table is the following:

K_t	O_t	$P(O_t\|K_t)$
fair	heads	0.5
fair	tails	0.5
biased	heads	0.8
biased	tails	0.2

Also, the initial state is the one Lisa starts with a fair coin:

K_1	$P(K_1)$
fair	1.0
biased	0.0

Question 3

Consider a cow that may be infected with a disease that can possibly be detected by performing a milk test. The test is performed on five consecutive days, leading to five outcomes. We want to determine the state of the cow's infection over these days, given the test outcomes. The prior probability of infection on day one is 1/10,000; the test false-positive rate is 5/1,000, and; its false-negative rate is 1/1,000. Moreover, the state of infection at a given day depends only on its state on the previous day. In particular, the probability of a new infection on a given day is 2/10,000, while the probability that an infection would persist to the next day is 7/10.

a. Describe a Bayesian network and a corresponding query that solves this problem.

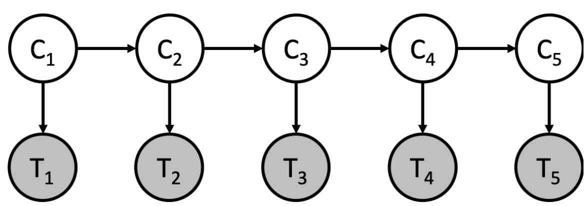
b. Write the probability tables necessary to model this problem.

c. Use the source code developed in the practical part of this tutorial to find the most likely state of the cow's infection over the five days given the following test outcomes:

   1. positive, positive, negative, positive, positive
   2. positive, negative, negative, positive, positive
   3. positive, negative, negative, negative, positive

Answer

a. The Bayesian network is a Hidden Markov model (HMM) of the following form:



where C_t are the hidden states that represent the cow's state and can assume one of two values: **healthy** or **infected**. O_t are the observed variables with test results and can assume the values **positive** or **negative**.

To answer this query, we need to indicate the most likely states for the hidden variables. Therefore, this is an MLE query that can be answered using the Viterbi algorithm.

b. We need a transition probability table and an emission probability table to define this HMM. It is also useful to describe the initial state.

The transition probability table is the following:

C_{t-1}	C_t	$P(C_t \mid C_{t-1})$
healthy	healthy	9998/10000
healthy	infected	2/10000
infected	healthy	3/10
infected	infected	7/10

The emission probability table is the following:

C_t	T_t	$P(T_t \vert C_t)$
healthy	positive	5/1000
healthy	negative	995/1000
infected	positive	999/1000
infected	negative	1/1000

Also, the initial state is related to the disease prevalence:

C_1	$P(C_1)$
healthy	9999/10000
infected	1/10000

Question 4

For each transition matrix below, is the corresponding Markov chain irreducible? Is it aperiodic? What is its stationary distribution? The matrix $T(i, j)$ represents the transition probabilities from a state i in time $t-1$ to a state j in time t.

a. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

b. $\begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}$

Answer

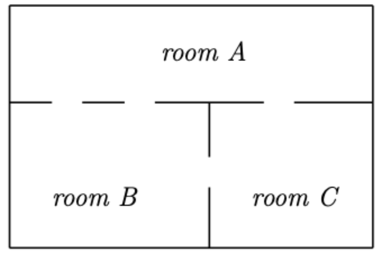
a. Irreducible since we can reach any state from any state. Periodic since starting from state 1 in time 0, we reach states 1 and 3 at even times and 2 and 4 at odd times. Stationary distribution $(1/6, 2/6, 2/6, 1/6)$.

b. Reducible since state 4 is absorbing (once entered, cannot be left). Aperiodic. Stationary distribution $(0, 0, 0, 1)$.

c. Reducible since we cannot reach states {3,4} from states {1,2} and vice-versa. Aperiodic. Two Stationary distributions $(1/2, 1/2, 0, 0)$ and $(0, 0, 1/2, 1/2)$

Question 5

A trained mouse lives in the house shown. A bell rings at regular intervals, and the mouse is trained to change rooms each time it rings. When it changes rooms, it is equally likely to pass through any of the doors in the room it is in. Approximately what fraction of its life will it spend in each room?



Answer

This problem has the following transition matrix, considering $A = 1$, $B = 2$ and $C = 3$.

$$\begin{bmatrix} 0 & 2/3 & 1/3 \\ 2/3 & 0 & 1/3 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

Therefore, we have the following equations

$$\begin{aligned} \pi(A) &= 2/3\pi(B) + 1/2\pi(C) \\ \pi(B) &= 2/3\pi(A) + 1/2\pi(C) \\ \pi(C) &= 1/3\pi(B) + 1/3\pi(A) \end{aligned}$$

Simplifying, we get

$\pi(A) = \pi(B) = 3/2\pi(C)$

And imposing the constrain $\pi(A) + \pi(B) + \pi(C) = 1$

We obtain $\pi(A) = 3/8, \pi(B) = 3/8$ and $\pi(C) = 1/4$.