Analyse de données et modèles stochastiques

October 29, 2022

Duration: 1h30

> Lecture handouts and notes authorized. Calculators are authorized.
>
> Exercises are independent. Question within exercises can mostly be considered as independent. The number of points in front of each exercise is given as an indication and is subject to changes within reasonable bounds.
>
> When answering the questions, do not limit yourself to reporting just a number or a trivial answer. Rather provide the details of your calculations and of your reasoning.

Exercice 1 – Hidden Markov Chains (12 points)

> For numerical applications, use the table of simplified logarithms given in Appendix.

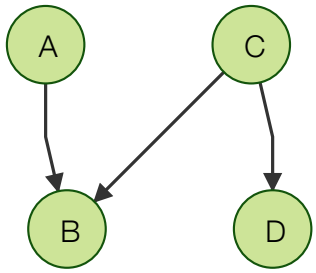
Consider a three-state hidden Markov model $(1, 2, 3)$ on an observation space with two symbols (a, b) and whose parameters are the following:

$$\pi = \begin{pmatrix} 0.6 & 0.3 & 0.1 \end{pmatrix} \qquad A = \begin{pmatrix} 0.3 & 0.7 & 0.0 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.8 & 0.2 \end{pmatrix} \qquad B = \begin{pmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \\ 0.2 & 0.8 \end{pmatrix}.$$

Questions

1. What is the (log-)probability of observing the sequence $(1, a)\ (1, b)\ (2, b)\ (3, b)\ (2, a)$?
2. Using the Viterbi algorithm, give the most likely state sequence for the sequence of observations $a\ b\ b\ b\ a$?
3. Still for the observed sequence $a\ b\ b\ a$, we wish to know the probability $P[S_3 = 1 \mid O = a\ b\ b\ a]$ that we are in state 1 at the time of the emission of the third observation. How can we calculate this probability based on some steps of the Baum-Welch algorithm?

Exercice 2 – Réseaux bayésiens (8 points)



We consider the above Bayesian network. The variables are binary and the model is defined by the following parameters:

$P[A = 0] = 0.3$ $P[C = 0] = 0.6$ $P[B = 0 \mid A, C] =$

A	C	
0	0	0.3
0	1	0.1
1	0	0.8
1	1	0.5

$P[D = 0 \mid C] =$

C	
0	0.2
1	0.6

.

1. Give the factorized form of the joint probability $P[A, B, C, D]$.
2. Compute $P[B = 1|A = 0, C = 0]$.
3. State two variables that are independent conditionally on another variable.
4. Using the factorization of the joint probability $P[A, B, C, D]$, show that the variables A and C are independent.

Appendix

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$ln(x)$	-2.3	-1.6	-1.2	-0.9	-0.7	-0.5	-0.4	-0.2	-0.1