Analyse de données et modèles stochastiques

29 octobre 2022

Durée : 1h30

> Supports et notes de cours autorisées. Calculatrices autorisées.
>
> Les exercices sont indépendants. Les questions au sein des exercices peuvent pour la plupart être considérées comme indépendantes. Le nombre de point par exercice est donné à titre indicatif et pourra être modifié a posteriori, dans les limites du raisonnable.
>
> Quand vous répondez aux questions, ne vous limitez pas à un simple résultat. Donnez le détail de vos calculs et précisez le raisonnement suivi.

Exercice 1 – Chaînes de Markov cachées (12 points)

> On utilisera pour cet exercice la table des logarithmes simplifiés donnée en annexe.

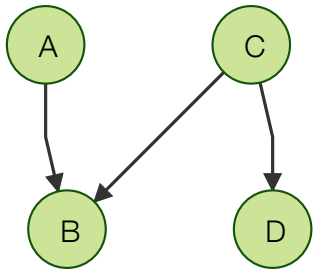
On considère un modèle de Markov caché à trois états $(1, 2, 3)$ portant sur un espace d'observation comportant deux symboles (a, b) et dont les paramètres sont les suivants :

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

Questions

1. Quelle est la (log-)probabilité d'observer la séquence $(1, a)\ (1, b)\ (2, b)\ (3, b)\ (2, a)$?
2. À l'aide de l'algorithme de Viterbi, donnez la séquence d'état la plus vraisemblable pour la séquence d'observation $a\ b\ b\ b\ a$?
3. Toujours pour la séquence observée $a\ b\ b\ b\ a$, on désire connaître la probabilité $P[S_3 = 1 \mid O = a\ b\ b\ b\ a]$ que l'on se trouve dans l'état 1 à lors de l'emission de la troisième observation. Comment peut-on calculer cette probabilité en s'appuyant sur certaines étapes de l'algorithme de Baum-Welch ?

Exercice 2 – Réseaux bayésiens (8 points)



On considère le réseau bayésien ci-dessus. Les variables sont binaires et le modèle est défini par les paramètres suivants :

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

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. Donnez la forme factorisé de la probabilité conjointe $P[A, B, C, D]$.
2. Calculez $P[B = 1|A = 0, C = 0]$.
3. Indiquez deux variables qui sont indépendantes conditionnellement à une autre variable.
4. En vous appuyant sur la factorisation de la probabilité conjointe $P[A, B, C, D]$, montrez que les variables A et C sont indépendantes.

Annexe

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