Chaîne de Markov cachée

Données

N le nombre d'états cachés et $Q=\{q_1,\dots,q_N\}$ l'ensemble des états cachés.

$P=(p_{i,j})\in\mathcal{M}_N(\mathbb{R})$ la matrice stochastique des transitions de la chaîne de Markov cachée.

$\Pi=(\pi_i)\in\mathbb{R}^N$ répartition initiale des probabilités sur la chaîne de Markov cachée.

M le nombres de symboles observables $V=\{v_1,\dots,v_M\}$ l'ensemble d'iceux

$B=(b_{i,k})\in\mathcal{M}_{N,M}(\mathbb{R})$ la matrice dont les coefficients donnent les probabilités d'observation selon l'état de la chaîne de Markov :

$$b_{i,k}=b_i(v_k)=\text{probabilité d'observer } v_k \text{ quand l'état est } q_i$$

Observations

On réalise T observations générant une séquence :

$$O_T=(o_1,o_2,\dots,o_T)\in V^T$$

Ces observations sont issues d'une séquence d'états cachés

$$I_t=(i_1,i_2,\dots i_T)\in[\![1,N]\!]^T$$

avec i_t l'indice de l'état caché de la chaîne de Markov à l'instant t.

Le problème

Etant donnée une séquence d'observation O_T, quelle est la succession d'état I_T la plus probable ?

On cherche donc I_T tel que la probabilité $p(I_T/O_T)$ soit maximale.
Or

$$p(I_T/O_T)=\frac{p(I_T \text{ et } O_T)}{p(O_T)}$$

le problème est donc équivalent à déterminer I_T maximisant $p(I_T \text{ et } O_T)$.

L'algorithme de Viterbi

On pose

$$\delta_t(j)=\max_{1\le i_1,\dots i_{t-1}\le N} p(i_1,\dots,i_{t-1},o_1,\dots,o_{t-1},i_t=j,o_t)$$

où $p(i_1,\dots,i_{t-1},o_1,\dots,o_{t-1},i_t=j,o_t)$ se comprend comment étant la probabilité d'avoir les t premières observations $o_1,\dots,o_{t-1},o_t$ en passant par la succession d'états $i_1,\dots,i_{t-1}$ et j

Le nombre $\delta_t(j)$ se comprend comment étant la probabilité maximale d'être dans l'état j et d'avoir observé $o_1,\dots,o_{t-1}$ et o_t.

On a

$$\delta_1(i)=b_i(o_1)\pi_i \text{ et pour } t\ge 2,\ \delta_t(j)=\max_{1\le i\le N}\left\{\delta_{t-1}(i)p_{i,j}b_j(o_t)\right\}$$

Posons aussi

$$\psi_t(j)=\text{l'indice } i \text{ réalisant le max définissant } \delta_t(j)$$

En déterminant i_T maximisant $\delta_T(i)$ pour $i\in\{1,\dots,N\}$, on détermine l'état final le plus probable et on peut reconstruire la séquence de succession d'états la plus probable par

$$i_t=\psi_{t+1}(i_{t+1}) \text{ pour } t=(T-1),\dots,1$$