24/03/2021

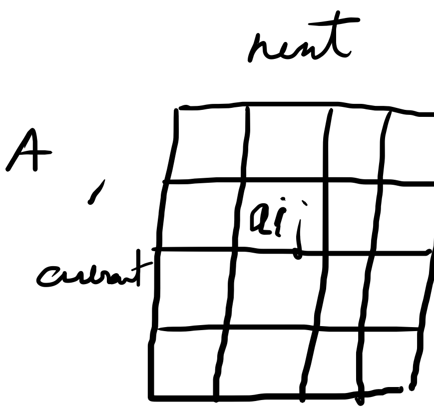
# Hidden Markov Models

An HMM is specified by the following components:

$Q = q_1 q_2 \ldots q_N$ — a set of N states

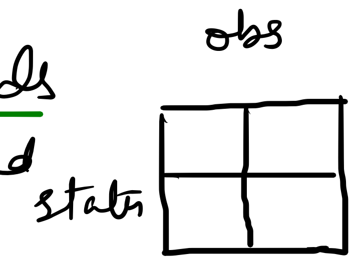
✓ $A = a_{11} a_{12} \ldots a_{n1} \ldots a_{nn}$ — a transition probability matrix A, each $a_{ij}$ represents of moving from state i to state j

s.t. $\sum_{j=1}^{n} a_{ij} = 1 \quad \forall i$




✓ $O = o_1 o_2 \ldots o_T$ — a sequence of T observations

$B = b_i(o_t)$ — sequence of observation likelihoods an observation $o_t$ beign generated from state i




$q_0, q_F$ — start state and end state

$\pi = \pi_1, \pi_2 \ldots \pi_N$ — an initial probability distribution over states

$QA = \{q_x, q_y \ldots\}$ — a set $QA \subset Q$ of legal accepting states.

$P(q_i | q_1 \ldots q_{i-1}) = P(q_i | q_{i-1})$ — Markov Assumption

$P(o_i | q_1 \ldots q_i, o_1 \ldots o_i \ldots o_T) = P(o_i / q_i)$ — Output Assumption

Problem 1 (Likelihood): Given an HMM $\lambda = (A, B)$ and an observation sequence $O$, determine the likelihood $P(O/\lambda)$

Problem 2 (Decoding): Given an observation sequence $O$ and HMM $\lambda = (A, B)$, discover the best hidden state sequence $Q$.

Problem 3 (Learning): Given an observation sequence $O$ and the set of states, learn the HMM parameters $A$ and $B$.

Problem 1 (Likelihood) : Given an HMM $\lambda = (A, B)$ and an observation sequence $O$, determine the likelihood $P(O/\lambda)$

Transition

$$\begin{bmatrix} A \end{bmatrix}$$

Observation

$$\begin{bmatrix} B \end{bmatrix}$$

Observation sequence $O$

$3 \quad , \quad 1 \quad , \quad 3$

$$P(3,1,3/\lambda) = ?$$

Problem 2 (Decoding) : Given an observation sequence $O$ and HMM $\lambda = (A, B)$, discover the best hidden state sequence $Q$.

$O \rightarrow 3, 1, 3$

best $P(\text{states} \mid 3, 1, 3)$

$P(\text{hot}, \text{hot}, \text{hot} \mid 3, 1, 3)$

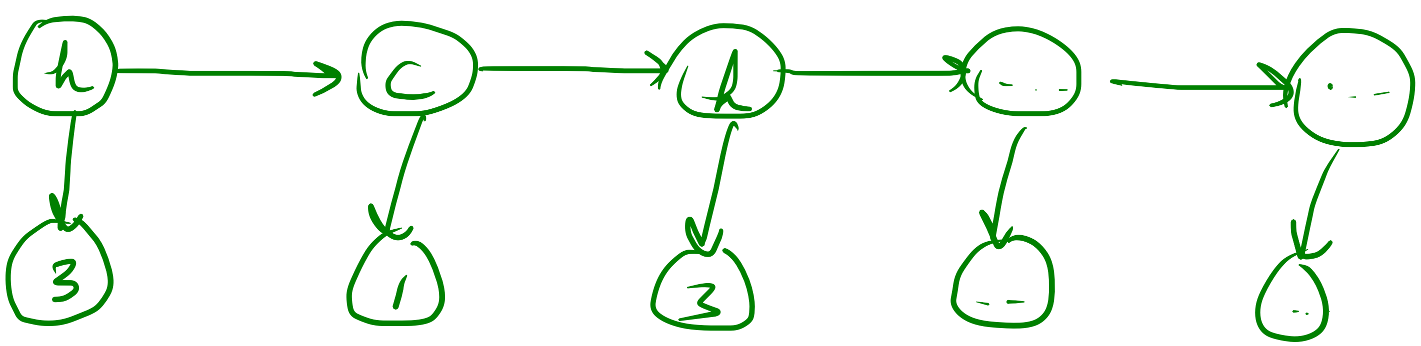
$\vdots$

$P(\text{hot cold hot} \mid 3, 1, 3)$

Problem 3 (Learning) : Given an observation sequence O and the set of states, learn the HMM parameters A and B .

S hot cold hot cold, hot, cold, hot, hot

O $\longrightarrow$ 3, 1, 3 - - - - - - - - 1, 6, 2, 8, 9




$\begin{bmatrix} ? \\ A \end{bmatrix}$ $\begin{bmatrix} ? \\ B \end{bmatrix}$

Problem 1    Given $\lambda = (A, B)$

$$P(O \mid A) = ?$$

| | | next | |
|---|---|---|---|
| | | h | c |
| Cur | h | | |
| | c | | |

| | h | c |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |

Let take $O = 3, 1, 3$

Assumption:- suppose corresponding sequence of state is hot hot cold

$$\begin{tikzcd}
h \arrow[r] \arrow[d] & h \arrow[r] \arrow[d] & c \arrow[d] \\
3 & 1 & 3
\end{tikzcd}$$

$$P(3,1,3 \mid h, h, c) = P(3 \mid h) \cdot P(1 \mid h) \cdot P(3 \mid c)$$

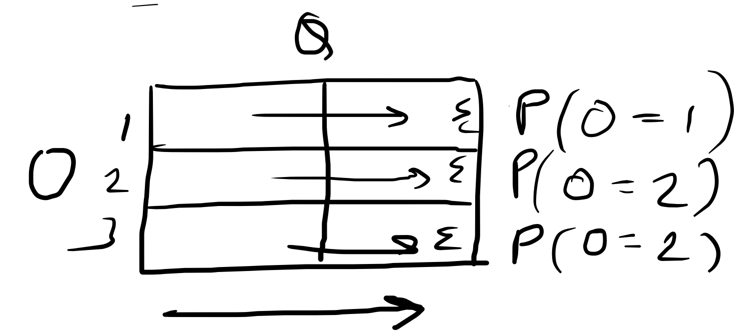
Since we dont know the hidden state sequence, we need to compute the probability of 3, 1, 3 by summing over all possible state sequences weighted by their prob.

$$P(O,Q) = P(O|Q)\cdot P(Q) = \prod_{i=1}^{N} P(o_i|q_i) \cdot \prod_{i=1}^{N} (q_i|q_{i-1})$$

$$P(3\,1\,3 \mid \text{hot hot cold}) = P(h|\text{start})\cdot P(h|h)\cdot P(c|h)\cdot P(3|h)\cdot P(1|h)\cdot P(3|c)$$

$$P(O) = \sum_{Q} P(O,Q)$$

$$P(O) = \sum_{Q} P(O|Q)\cdot P(Q)$$




$$P(3,1,3) = P(3,1,3 \mid \text{hot}\cdot\text{hot}\cdot\text{hot})\cdot P(\text{hot}\cdot\text{hot}\cdot\text{hot}) +$$

$$P(3,1,3 \mid \underline{x}\ \underline{x}\ \underline{x})\cdot P(\underline{x}\ \underline{x}\ \underline{x}) +$$

$$P(3,1,3 \mid \underline{y}\ \underline{y}\ \underline{y})\cdot P(y\ y\ y) +$$

This direct method not computationally feasible.

Solution:- Dynamic Programming base FORWARD ALGORITHM