Hidden Markov Model

Wei Wang @ CSE, UNSW

April 3, 2019

- $P(S_{1:k} \mid O_{1:k}) \stackrel{\text{def}}{=} P(S_1, \ldots, S_k \mid O_1, \ldots, O_k)$
- What do we need?
  - $\arg\max P(S_{1:k} \mid O_{1:k})$
  - Can we derive even $P(S_1 \mid O_1)$?

k=1的时候

s是state O是observation

- $P(S_{1:k} \mid O_{1:k}) \stackrel{\text{def}}{=} P(S_1, \ldots, S_k \mid O_1, \ldots, O_k)$
- What do we need? 在给定观测样本序列前提下,得到最大概率的状态序列排序
  - $\arg\max P(S_{1:k} \mid O_{1:k})$
  - Can we derive even $P(S_1 \mid O_1)$?
  - Need to use the Bayes rule:

$$P(S_{1:k} \mid O_{1:k}) = \frac{P(O_{1:k} \mid S_{1:k}) \cdot P(S_{1:k})}{P(O_{1:k})}$$

  - Because of the "arg max", we have

$$\begin{aligned} \arg\max_{S_{1:k}} P(S_{1:k} \mid O_{1:k}) &= \arg\max_{S_{1:k}} P(O_{1:k} \mid S_{1:k}) \cdot P(S_{1:k}) \\ &= \arg\max_{S_{1:k}} P(O_{1:k}, S_{1:k}) \end{aligned}$$

- Since we can compute the probabilities, the naive method is to consider all $O(N^k)$ state sequences, where N is the number of states.

- A common trick (for Markov Chains) is to check if we can use recursion (thanks for the abundant conditional independences given by the Markov assumption).

Note:

$$P(O_{1:k+1} \mid S_{1:k+1}) = P(O_{k+1} \mid O_{1:k}, S_{1:k+1}) \cdot P(O_{1:k} \mid S_{1:k+1})$$
$$= P(O_{k+1} \mid S_{k+1}) \cdot P(O_{1:k} \mid S_{1:k})$$

$$P(S_{1:k+1}) = P(S_{1:k}) \cdot P(S_{k+1} \mid S_k)$$

Then

$$\begin{aligned}
& P(O_{1:k+1}, S_{1:k+1}) \\
=& P(O_{k+1} \mid S_{k+1}) \cdot P(O_{1:k} \mid S_{1:k}) \cdot P(S_{1:k}) \cdot P(S_{k+1} \mid S_k) \\
=& P(O_{k+1} \mid S_{k+1}) \cdot P(S_{k+1} \mid S_k) \cdot P(O_{1:k}, S_{1:k})
\end{aligned}$$

$$
\begin{aligned}
P(O_{1:k+1}, S_{1:k+1}) =& P(O_{1:k}, S_{1:k}) \cdot \Big(P(O_{k+1} \mid S_{k+1}) \cdot P(S_{k+1} \mid S_k) \Big) \\
=& P(O_{1:k}, S_{1:k}) \cdot f(S_k, S_{k+1})
\end{aligned}
$$

- Note that the value of S_k is important when considering applying "max".
- Define $\delta(k, v) \stackrel{\text{def}}{=} \max_{S_{1:k-1}} P(O_{1:k}, S_{1:k-1}, v)$, we have the recurrence:

  $$\delta(k+1, v) = \max_u (\delta(k, u) \cdot f(u, v))$$

  with the boundary condition:

  $$\delta(1, v) = P(O_1, v)$$

- Then

  $$\max_{S_{1:k+1}} P(O_{1:k+1}, S_{1:k+1}) = \max_v \max_{S_{1:k}} P(O_{1:k+1}, S_{1:k}, v) = \max_v \delta(k+1, v)$$

- The arg max solution can be obtained via backtracking or some additional data structure.