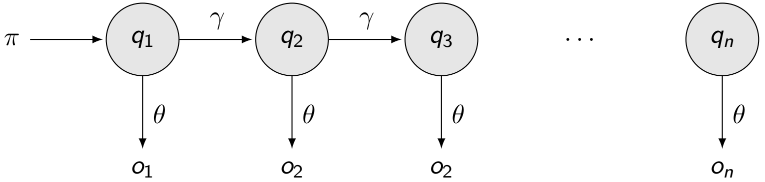
Hidden Markove Model

Dihui Lai

dlai@wustl.edu

December 30, 2018

Hidden Markov Model



Hidden Markov Model

$$
\begin{aligned}
& p(o_1, o_2, \ldots, o_T, q_1, q_2, \ldots, q_T) \\
& = p(q_1, q_2, \ldots, q_T) \prod_{i=1}^{T} p(o_i \mid q_i) \\
& = p(q_T \mid q_1, q_2, \ldots, q_{T-1}) p(q_1, q_2, \ldots, q_{T-1}) \prod_{i=1}^{T} p(o_i \mid q_i) \\
& = p(q_T \mid q_{T-1}) p(q_1, q_2, \ldots, q_{T-1}) \prod_{i=1}^{T} p(o_i \mid q_i) \\
& = \prod_{i=1}^{T} p(q_i | q_{i-1}) \prod_{i=1}^{T} p(o_i \mid q_i)
\end{aligned}
$$

At each time step, the probability of being in state j after seeing the first t observations, given the automaton λ can be calculated as

$$\alpha_t(j) = \sum_{i=1}^{N} \alpha_{t-1}(i)\gamma_{ij}\theta_j(o_t),$$

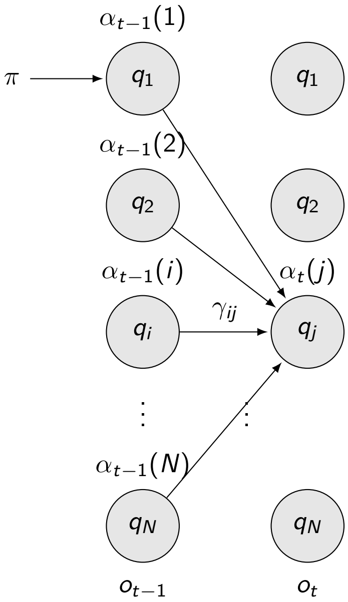
where

$$1 \leq t \leq T, 1 \leq j \leq N$$

where $\theta_j(o_t)$ is the likelihood of observing o_t given the current state is j, λ_{ij} is the probability of transition from hidden state i to hidden state j. The probability of state q_j at time t is a summation over all paths that could lead to the state. Therefore, the probability of seeing the observed sequence can be calculated as

$$P(o \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

Hidden Markov Model



HMM: Viterbi Algorithm

Given that we know the HMM $\lambda = (\Gamma, \Theta)$ and an observed squence of $O = o_1, o_2, ...o_T$, find the most probable sequence of states $Q = q_1, q_2, q_2, ..., q_T$.
The viterbi path probablity: the probability of the most probable path that leads to the jth state at time step t. Denoted as

$$\nu_t(j) = \max_{q1,q2,q3,...q_{t-1}} p(q_1, q_2, ...q_{t-1}, q_t = j, o_1, o_2, ..., o_t \mid \lambda)$$

We can calculate the viterbi path probabilty of being at state q_j at time t as

$$\nu_t(j) = \max_{i=1}^{N} \nu_{t-1}(i)\gamma_{ij}\theta_j(o_t)$$

The Viterbie Algorithm needs to back trace the most likely state sequence that leads to the probability. $bt_t(j)$, which can be found by

$$bt_t(j) = \operatorname*{argmax}_{i=1}^{N} \nu_{t-1}(i)\gamma_{ij}\theta_j(o_t)$$

HMM: Backward Algorithm

Denote the probability of seeing the observation after $t+1$ as $o_{t+1}, o_{t+2}, ...o_T$, given that we are at state j at time t.

$$\beta_t(j) = p(o_{t+1}, o_{t+2}, ..., o_T \mid q_t = j, \lambda)$$

The $\beta_t(j)$ can be expressed as the following formula

$$\begin{aligned}\beta_t(j) &= \sum_{k=1}^{N} \gamma_{jk}\theta_k(o_{t+1})p(o_{t+2}, o_{t+3}, ..., o_T \mid q_{t+1} = k, \lambda) \\ &= \sum_{k=1}^{N} \gamma_{jk}\theta_k(o_{t+1})\beta_{t+1}(k)\end{aligned}$$

The transition probability of going from *ith* state at time $t-1$ to *jth* state at time t can be calculated as

$$p(O|\lambda) = \sum_{t=1}^{N}\sum_{j=1}^{N}\alpha_t(j)\beta_t(j)$$

$$p(q_t = i, q_{t+1} = j, O \mid \lambda) = \sum_{t=1}^{N}\sum_{j=1}^{N}\alpha_t(i)\gamma_{ij}\theta_j(o_{t+1})\beta_{t+1}(j)$$

$$\begin{aligned}
\hat{\gamma}_{ij} &= p(q_{t-1} = i, q_t = j, \mid O, \lambda) \\
&= p(q_{t-1} = i, q_t = j, O \mid \lambda)/p(O \mid \lambda) \\
&= \frac{\sum_{t=1}^{N}\alpha_t(i)\gamma_{ij}\theta_j(o_{t+1})\beta_{t+1}(j)}{\sum_{t=1}^{N}\sum_{j=1}^{N}\alpha_t(j)\beta_t(j)}
\end{aligned}$$

Similarly, to calculate the emission probability $\theta_j(o_t)$

$$\begin{aligned}\eta_t(j) &= p(q_t = j \mid O, \lambda)) \\ &= \frac{p(q_t = j, O \mid \lambda)}{p(O \mid \lambda)} \\ &= \frac{\alpha_t(j)\beta_t(j)}{\sum_{j=1}^{N} \alpha_t(j)\beta_t(j)}\end{aligned}$$

The emission probability can be estimated as

$$\hat{b}_j(k) = \frac{\sum_{t=1, s.t. o_t = k}^{T} \eta_t(j)}{\sum_{t=1}^{T} \eta_t(j)}$$