CSC246 Sequence Project Math writeup

Tianyi Ma

April 24, 2021

Calculating $\alpha_t(j)$:
1. Initialization:

$$\alpha_1(j) = \pi_j b_j(o_1) 1 \leq j \leq N$$

2. Recursion:

$$\alpha_t(j) = \sum_{i=1}^{N} \alpha_{t-1}(i) a_{ij} b_j(o_t); 1 \leq j \leq N, 1 < t \leq T$$

3. Termination:

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

π_j - the initial probability of hidden state j
a_{ij} - the transition probability from state i to j
$b_j(o_t)$ - the emission probability of state j to the t-th observation
N - the number of possible hidden states
T - the total number of tokens in the sequence
$P(O|\lambda)$ - the probability of the sample sequence based on current parameters

Calculating $\beta_t(i)$:
1. Initialization:

$$\beta_T(i) = 1;\ 1 \leq i \leq N$$

2. Recursion:

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(o_{t+1}) \beta_{t+1}(j);\ 1 \leq i \leq N, 1 \leq t < T$$

3. Termination:

$$P(O|\lambda) = \sum_{j=1}^{N} \pi_j b_j(o_1) \beta_1(j)$$

EM-algorithm:
E-step:

$$\gamma_t(j) = \frac{\alpha_t(j)\beta_t(j)}{\alpha_T(q_F)} \ \forall t \text{ and } j$$

$$\xi_t(i,j) = \frac{\alpha_t(j)a_{ij}b_j\beta_{t+1}(j)}{\alpha_T(q_F)} \ \forall t, i, \text{and } j$$

M-step:

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1}\sum_{k=1}^{N} \xi_t(i,k)}$$

$$\hat{b}_j(v_k) = \frac{\sum_{t=1 s.t. O_t=v_k}^{T} \gamma_t(i,j)}{\sum_{t=1}^{T} \gamma_t(j)}$$

Two questions:
1. How to compute the log likelihood?
2. What is q_F in the E-step?