STAT 534
Lecture

©2019 Marina Meilă
mmp@stat.washington.edu
Scribes: Yae-Lin Sheu

1 Inference Problems

Definitions

- A: transition matrix
- B: transmission matrix
- a_{ij}: transition probability from state i to state j

$$a_{ij} = Pr[\ q_t = j \mid q_t = i\ ] \tag{1}$$

- b_{ik}: emission probability

$$b_{ik} = Pr[\ O_t = k \mid q_t = i\ ] \tag{2}$$

- π: initial distribution of q_1

Assume:

- $O_{1:T}$ observed
- $q_{1:T}$ unobserved

1. Forward-Backward

$$Pr[\ O_{1:T} \mid A, B\ ] \tag{3}$$

2. Viterbi algorithm

$$\arg\max_{q_{1:T}} Pr[\ q_{1:T} \mid O_{1:T}\ ] \tag{4}$$

3. Baum-Welch (ML Estimation): Estimate $\hat{A}$, $\hat{B}$ from $O_{1:T}$.

$$Pr[\ q_1 = i\ ] = \pi_i \tag{5}$$

$$Pr[\ q_2 = j\ ] = \sum_{i=1}^{N} \pi_i a_{ij} \tag{6}$$

Note that here the i indicates the values of q_1.

$$Pr[\ q_t = j\ ] = \sum_{i=1}^{N} Pr[\ q_{t-1} = i\ ]a_{ij} \tag{7}$$

Note that here the i indicates the values of q_{t-1}.

$$Pr[\ O_1 = k\ ] = \sum_{i=1}^{N} \pi_i b_{ik} \tag{8}$$

Note that here the i indicates the values of q_1.

$$Pr[\ O_t = k\ ] = \sum_{i=1}^{N} Pr[\ q_t = i\ ]b_{ik} \tag{9}$$

Note that here the i indicates the values of q_t.

- $\alpha_t(i) = Pr[\ O_1(t), q_t = i\ ]$
- $\beta_t(i) = Pr[\ O_{t+1:T}, q_t = i\ ]$
- $\alpha_1(i) = Pr[\ q_1 = i\ ]Pr[\ O_1 \mid q_1 = i\ ] = \pi_i b_{iO_t}$
-
$$\alpha_t(i) = \sum_{j=1}^{N} \alpha_{t-1}(j) a_{ji} b_{iO_1} \tag{10}$$

$$= \sum_{j=1}^{N} Pr[\ O_{1:t-1} \mid q_{t-1} = j\ ]Pr[\ q_t = i \mid q_{t-1} = j\ ]Pr[\ O_t \mid q_t = j\ ] \tag{11}$$

$$= \sum_{j=1}^{N} Pr[\ O_{1:t}\ ,\ q_{t-1} = j\ ,\ q_t = i] \tag{12}$$

  Note that here the j indicates the values of q_t. There are N^2 operations.

-
$$Pr[\ O_{1:T}\ ] = \sum_{i=1}^{N} \alpha_i(T) = \sum_{q_1,\ldots,q_T} Pr[\ q_1 ... q_T]P[O_{1:T} \mid q_{1:T}] \tag{13}$$

  In the first part of the equation, there are N^2T operations. In the latter part of the equation, there are N^T sequences.

- $\beta_T(i) = 1$

- $$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_{jO_{t+1}}(j) \quad (14)$$

  $$= \sum_{j=1}^{N} Pr[\ q_{t+1} = j \mid q_t = i\ ] P[O_{t+2:T} \mid q_{t+1} = j] \quad (15)$$

  Note that there are N^2 operations. The run time for the backward procedure is approximately TN^2.

- exercise: $Pr[\ O_{1:T}\ ]$ from $\beta_i(t)$, $t = 1 : T$, $i = 1 : N$.

$$Pr[\ O_{1:T}\ ] = \sum_{i=1}^{N} \alpha_i(t)\beta_i(t) = \sum_{i=1}^{N} P[O_{1:t}\ ,\ q_i] P[O_{t+1:T} \mid q_i] \quad (16)$$

$$\gamma_i(t) = P[q_t = i \mid O_{1:T}] = \frac{P[q_t = i\ ,\ O_{1:T}]}{Pr[\ O_{1:T}\ ]} \quad (17)$$