EE 5345
Biomedical Instrumentation
Lecture 22: slides 417-427

Carlos E. Davila, Electrical Engineering Dept.
Southern Methodist University
slides can be viewed at:
http:// www.seas.smu.edu/~cd/ee5345.html

The Viterbi Algorithm

- Initialization ($t = 0$):

$$d_0(i) = p_i b_i(O_0), \quad 0 \le i \le N-1$$

$$\Psi_1(i) = 0$$

- Time Recursion

$$\text{For } 1 \le \text{t} \le \text{T-1}, \;\; 0 \le \text{j} \le \text{N-1}$$

$$d_{\text{t}}(j) = \max_{0 \le i \le N\text{-}1}\left[d_{\text{t-1}}(\text{i}) a_{ij}\right] b_j(O_t)$$

$$\Psi_{\text{t}}(j) = \arg\max_{0 \le i \le N\text{-}1}\left[d_{\text{t-1}}(\text{i}) a_{ij}\right]$$

The Viterbi Algorithm (cont.)

- Termination:

$$P_{\max} = \max_{0 \le i \le N-1} \left[d_{T-1}(i) \right]$$

$$i_{T-1} = \underset{0 \le i \le N-1}{\arg\max} \left[d_{T-1}(i) \right]$$

- State sequence backtracking:

For $t = T$-2, T-3, …, 0

$$i_t = \Psi_{t+1}\left(i_{t+1}\right)$$

Backward Variable

$$b_t(i) = \Pr\left(O_{t+1}, O_{t+2}, \ldots, O_{T-1} | i_t = Q_i, l\right)$$

To understand this variable, assume that the current time step is "t", the current state is "Q_i", and we know the probabilities:

$$b_{t+1}(j), \quad j = 0, \ldots N-1$$

then it should be clear that:

$$b_t(i) = \sum_{j=0}^{N-1} a_{ij} b_j(O_{t+1}) b_{t+1}(j), \quad 0 \le i \le N-1, \quad 0 \le t \le T-2$$

since each of the N events:

$$O_{t+1}, O_{t+2}, \ldots, O_{T-1} | i_t = Q_j, \quad j = 0, \ldots, N-1$$

are disjoint.

Backward Variable (cont.)

The backward variable can be computed recursively, moving backward in time.

1. initialize at $t = T - 1$,

$$b_{T-1}(i) = 1, \quad i = 0, \ldots N-1$$

2. for $t = T - 2 : -1 : 0$

$$b_t(i) = \sum_{j=0}^{N-1} a_{ij} b_j(O_{t+1}) b_{t+1}(j), \quad 0 \le i \le N-1$$

More Definitions

The probability of landing in state Q_i at time t, given the observation sequence O is:

$$g_t(i) \equiv \Pr\left(i_t = Q_i \middle| O, l\right)$$

consider the previous definitions:

$$a_t(i) = \Pr\left(O_0, O_1, \ldots, O_t, i_t = Q_i \middle| l\right)$$

$$b_t(i) = \Pr\left(O_{t+1}, O_{t+2}, \ldots, O_{T-1} \middle| i_t = Q_i, l\right)$$

hence, for a given model λ:

$$a_t(i) b_t(i) = \Pr\left(O_0, O_1, \ldots, O_{T-1} \cap i_t = Q_i\right)$$

More Definitions (cont.)

Hence: $$g_t(i) = \frac{a_t(i) b_t(i)}{\Pr(O | l)}$$

now consider the probability that we go from state Q_i at time t to state Q_j at time $t+1$ given the observation O:

$$\xi_t(\mathrm{i, j}) \equiv \Pr\left(\mathrm{i}_t = \mathrm{Q}_\mathrm{i}, \mathrm{i}_{t+1} = \mathrm{Q}_\mathrm{j} | \mathrm{O}, \lambda\right)$$

it follows that

$$x_{\mathrm{t}}(i, j) = \frac{a_t(i) a_{ij} b_j(O_{t+1}) b_{t+1}(j)}{\Pr(O | l)}$$

More Definitions (cont.)

the average number of transitions made from Q_i:

$$\sum_{t=0}^{T-2} g_t(i)$$

the average number of transitions made from Q_i to Q_j:

$$\sum_{t=0}^{T-2} x_t(i, j)$$

Solution to Problem 3: Baum-Welch Algorithm

0. Initialize A, B, and Π

1. Compute $a_t(i)$, $\beta_t(\mathrm{i})$ and $\Pr(O|\, l\,)$

2. Compute $x_t(i,j)$ and $g_t(i)$

$$x_t(i,j) = \frac{a_t(i) a_{ij} b_j(O_{t+1}) b_{t+1}(j)}{\Pr(O|\, l\,)} \qquad (\ ) \qquad \frac{{}_t(\ )\ \ {}_t(\ )}{(\qquad)}$$

3. Compute $p_i = g\ \ (i), \qquad i \le N-1$

4. Compute

$$\frac{\sum_{t}^{T-2} {}_t(\qquad)}{\sum_{t=0}^{T-2} g_t(\ )}$$

Baum-Welch Algorithm (cont.)

5. Compute

$$b_j(k) = \frac{\sum_{\substack{t=0 \\ O_t = v_k}}^{T-1} g_t(j)}{\sum_{t=0}^{T-1} g_t(j)}$$

7. go to step 2

$\Pr(O|l)$ should continue to increase until A, B, and Π converge to optimum values, at which point the algorithm is terminated.

Case Study: Coast et al.

- Used continuous density for observations:

$$b_i(v) = \frac{1}{\sqrt{2ps_i}} e^{-0.5((v-m_i)/s_i)^2}$$

This alters most of the formulas we looked at but the basic ideas remain the same.

- Observations consisted of actual ECG samples.
- Used several rhythm HMM models in parallel
- Viterbi algorithm was used to select the most likely sequence (and hence rhythm type).