HMM formula collection

Standard HMMs

Hidden Markov Process: $X_1^T = X_1, \dots, X_T$
Observed process: $Y_1^T = Y_1, \dots, Y_T$

State space: $S = \{s_1, \dots, s_N\}$
Initial probabilities: $\pi_i = P(X_1 = i),\ i = 1, \dots, N$
Transition probabilities: $a_{ij} = P(X_{t+1} = j | X_t = i), i, j \in S$
Emission probabilities: $b_j(Y_t | Y_1^{t-1}) = P(Y_t | X_t = j, Y_1^{t-1})$

$P(X_1^T, Y_1^T) = \pi_{X_1} b_{X_1}(Y_1) \prod_{t=2}^{T} a_{X_{t-1},X_t} b_{X_t}(Y_t | Y_1^{t-1})$

Forward Algorithm

$\alpha_i(t) = P(Y_1^t, X_t = i)$

Initialization: $\alpha_i(0) = \pi_i, i = 1, \dots, N$

Tabular computation: $\alpha_i(t) = \sum_{j \in S} \alpha_j(t-1) a_{ji} b_i(Y_t | Y_1^{t-1}), i = 1, \dots, N, t = 1, \dots, T$

Termination: $\alpha_i(T+1) = \sum_{j \in S} \alpha_j(T) a_{ji}, i = 1, \dots, N$

Backward Algorithm

$\beta_i(t) = P(Y_{t+1}^T | Y_1^t, X_t = i)$

Initialization: $\beta_i(T+1) = 1, i = 1, \dots, N$

Tabular computation: $\beta_i(t) = \sum_{j \in S} \beta_j(t+1) a_{ij} b_j(Y_{t+1} | Y_1^t), i = 1, \dots, N, t = T, T-1, \dots, 0$

Termination: none

The Viterbi Algorithm

$\delta_i(t) = \max_{X_1^{t-1}} P(Y_1^t, X_1^{t-1}, X_t = i)$

Initialization: $\delta_i(0) = \pi_i, i = 1, \dots, N$

Tabular computation: $\delta_i(t) = \max_{1 \le j \le N} \delta_j(t-1) a_{ji} b_i(Y_t | Y_1^{t-1}), i = 1, \dots, N, t = 1, \dots, T$
$\psi_i(t) = \operatorname{argmax}_{1 \le j \le N} \delta_j(t-1) a_{ji} b_i(Y_t | Y_1^{t-1}), i = 1, \dots, N, t = 1, \dots, T$

Termination: $\delta_i(T+1) = \max_{1 \le j \le N} \delta_j(T) a_{ji}, i = 1, \dots, N$
$\psi_i(T+1) = \underset{1 \le j \le N}{\operatorname{argmax}} \delta_j(T) a_{ji}, i = 1, \dots, N$

Traceback: $X_{T+1}^* = \psi_i(T+1)$
$X_t^* = \psi_{X_{t+1}^*}(t+1), t = T, T-1, \dots, 0$

The Baum-Welch Algorithm

$\gamma_i(t) = P(X_t = i | Y_1^T) = \frac{\alpha_i(t)\beta_i(t)}{\sum_{j \in S} \alpha_j(t)\beta_j(t)}, i = 1, \dots, N, t = 1, \dots, T$

$\xi_{ij}(t) = P(X_t = i, X_{t+1} = j | Y_1^T), i = 1, \dots, N, t = 1, \dots, T$

Initialization: pick arbitrary model parameters $\{\pi_i, a_{ij}, b_j : i, j = 1, \dots, N\}$

Recursion:

1. E-phase: calculate the forward and the backward algorithms for the current parameter settings.
2. M-phase: produce new parameter estimates using the re-estimation formulas below.

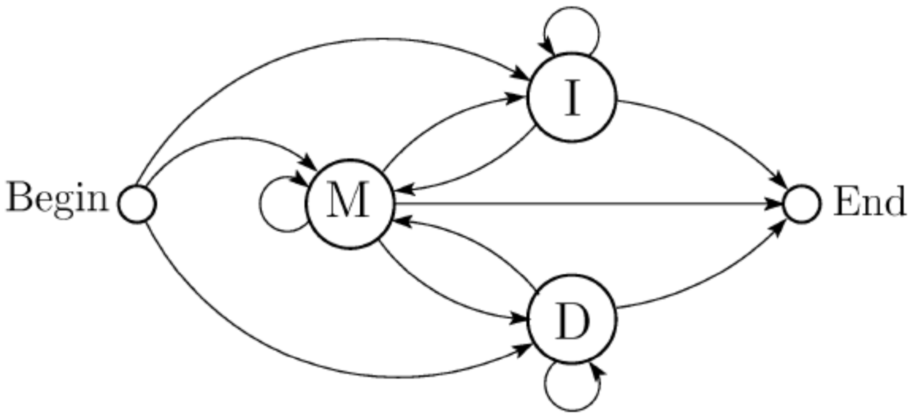
Re-estimation formulas:

$\hat{\pi}_i = \gamma_i(1)$

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_{ij}(t)}{\sum_{t=1}^{T-1} \gamma_i(t)}$$

$$\hat{b}_j(c) = \frac{\sum_{t=1, Y_t = c}^{T-1} \gamma_j(t)}{\sum_{t=1}^{T-1} \gamma_j(t)}$$

Pair HMMs



State durations:

$$(d_l, e_l) = \begin{cases} (1,1) \text{ if } X_l = M \\ (1,0) \text{ if } X_l = I \\ (0,1) \text{ if } X_l = D \end{cases}$$

$$p_l = \sum_{k=1}^{L} d_k, \qquad q_l = \sum_{k=1}^{L} e_k$$

$p_0 = q_0 = 0, p_L = T, q_L = U$

The Viterbi algorithm for PHMMs

Initialization: $\delta_i(0,0) = \pi_i$

$\delta_i(t, 0) = \delta_i(0, u) = 0, t > 0, u > 0$

Tabular computation:

$$\delta_M(t,u) = b_M(Y_t, Z_u \middle| Y_1^{t-1}, Z_1^{u-1}) \cdot \max \begin{cases} \delta_M(t-1, u-1)a_{MM} \\ \delta_I(t-1, u)a_{IM} \\ \delta_D(t, u-1)a_{DM} \end{cases}$$

$$\delta_I(t,u) = b_I(Y_t, - \middle| Y_1^{t-1}, Z_1^{u}) \cdot \max \begin{cases} \delta_M(t-1, u-1)a_{MI} \\ \delta_I(t-1, u)a_{II} \end{cases}$$

$$\delta_D(t,u) = b_D(-, Z_u \middle| Y_1^{t}, Z_1^{u-1}) \cdot \max \begin{cases} \delta_M(t-1, u-1)a_{MD} \\ \delta_D(t, u-1)a_{DD} \end{cases}$$

Termination:

$$\delta_{\mathrm{M}}(T+1, U+1) = \max \begin{cases} \delta_M(T, U)a_{MM} \\ \delta_I(T, U+1)a_{IM} \\ \delta_D(T+1, U)a_{DM} \end{cases}$$

$$\delta_I(T+1, U+1) = \max \begin{cases} \delta_M(T, U)a_{MI} \\ \delta_I(T, U+1)a_{II} \end{cases}$$

$$\delta_I(T+1, U+1) = \max \begin{cases} \delta_M(T, U)a_{MD} \\ \delta_D(T+1, U)a_{DD} \end{cases}$$