Notes on the Derivation

Junchen Feng

December 16, 2015

1 Derivation on the EM Algorithm

All notation follows Streeter(2015)'s paper.
The log likelihood function for the observed data is

$$\mathcal{L} = \sum_s \{log[\sum_j p_j \prod_t \mathcal{B}(q_t^j, v_t)]\}$$

Also borrow the notation

$$L_{s,j} = p_j \prod_t \mathcal{B}(q_t^j, v_t)$$
$$z_{s,j} = \frac{L_{s,j}}{\sum_{j'} L_{s,j'}}$$

1.1 q_t^j

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial q_t^j} &= \sum_s \frac{\frac{\partial L_{s,j}}{\partial q_t^j}}{\sum_j L_{s,j}} \\
&= \sum_s (\frac{\partial L_{s,j}}{\partial q_t^j} \frac{1}{L_{s,j}}) \frac{L_{s,j}}{\sum_j L_{s,j}} \\
&= \sum_s \frac{\partial log(L_{s,j})}{\partial q_t^j} z_{s,j} \\
&= \frac{1}{q_t^j(1-q_t^j)} \sum_s (v_t^s(1-q_t^j) - (1-v_t^s)q_t^j) z_{s,j} \\
&= \frac{1}{q_t^j(1-q_t^j)} \sum_s (v_t^s - q_t^j) z_{s,j} \\
&= 0
\end{aligned}$$

From here it is easy to derive that

$$\hat{q}_{t+1}^j = \frac{\sum_s v_t^s z_{s,j}}{\sum_s z_{s,j}}$$

Streeter then use beta prior to shrink the estimate by

$$q^j_{t+1} = \frac{\alpha - 1 + \sum_s v^s_t z_{s,j}}{\alpha + \beta - 2 + \sum_s z_{s,j}}$$

Note that the original v^j_s is likely a typo.

1.2 p_j

To add the constraint that $\sum_j p_j = 1$, use the lagrange multipler so that

$$\frac{\partial \mathcal{G}}{\partial p^j} = -\lambda + \sum_s \frac{1}{p^j} \frac{L_{s,j}}{\sum_j L_{s,j}} = -\lambda + \frac{1}{p^j} \sum_s z_{s,j} = 0$$
$$\frac{\partial \mathcal{G}}{\partial \lambda} = 1 - \sum_j p^j = 0$$

Some simple algebra yields

$$p^j_{t+1} = \frac{\sum_s z_{s,j}}{\sum_{j'} \sum_s z_{s,j'}}$$

2 Derivation on the Equivalence to BKT

The weight can be derived from the fact that the probability of mastering the knowledge component has to be the same at time t has to be the same.

$$P(M_t = 1) = \sum_j p_j I(q^j_t = 0.8)$$

Let j = 1,2,3 denotes the solid line, dash line and dot line respectively.

In the first period

$$P(M_1 = 1) = p_1 = 0.5$$

In the second period

$$\begin{aligned}
P(M_2 = 1) &= p_1 + p_2 \\
P(M_2 = 1) &= P(M_2 = 1, M_1 = 1) + P(M_2 = 1, M_1 = 0) \\
&= P(M_2 = 1|M_1 = 1)P(M_1 = 1) + P(M_2 = 1|M_1 = 0)P(M_1 = 0) \\
&= 1 * 0.5 + 0.5 * 0.5 \\
p_2 &= 0.25
\end{aligned}$$

It can be proved by generalization that

$$p_j = 0.5^j$$