A0286188L Vishwanath Dattatreya Doddamani E1237250@u.nus.edu

CS5340 Assignment 3 – Hidden Markov Models Report

1. E Step:
   In e_step() function implementation, $\gamma(Z_n)$ *and* $\xi(Z_{n-1,n})$ are computed. Alpha-Beta variant of forward backward algorithm is used for computation. First the alpha is calculated using equations:

$$\alpha(\mathbf{z}_1) = p(\mathbf{x}_1, \mathbf{z}_1) = p(\mathbf{z}_1)p(\mathbf{x}_1|\mathbf{z}_1) = \prod_{k=1}^{K} \{\pi_k p(\mathbf{x}_1|\phi_k)\}^{z_{1k}}$$

$$\alpha(\mathbf{z}_n) = p(\mathbf{x}_n|\mathbf{z}_n) \sum_{\mathbf{z}_{n-1}} \alpha(\mathbf{z}_{n-1}) p(\mathbf{z}_n|\mathbf{z}_{n-1})$$

   Here $p(X_n|Z_n)$, is calculated using normal distribution probability density function. It is then normalized with scaling factor. Then beta is initialized to ones and then calculated using equation:

$$\beta(\mathbf{z}_n) = \sum_{\mathbf{z}_{n+1}} \beta(\mathbf{z}_{n+1}) p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1}) p(\mathbf{z}_{n+1}|\mathbf{z}_n)$$

   Once alpha and beta are calculated, $\gamma(Z_n)$ is calculated using equation:

$$\gamma(\mathrm{Z_n}) = \frac{\alpha(\mathrm{Z_n})\beta(\mathrm{Z_n})}{\mathrm{p(X)}}$$

   Here $\mathrm{p(X)}$ is the sum of alpha multiplied by beta. Then, $\xi(Z_{n-1,n})$ is calculated using:

$$\frac{\alpha(z_{n-1})p(x_n|z_n)p(z_n|z_{n-1})\beta(z_n)}{p(X)}$$

   The resulting gamma_list and xi_list is returned.

2. M Step:
   In m_step() function implementation, pi (Probability of each state), A (transition matrix), phi (Mean and Standard Deviation of each state) is calculated. For each sequence of observation, pi, A, and phi are computed using equations:

$$\pi_k = \frac{\gamma(z_{1k})}{\sum_{j=1}^{K}\gamma(z_{1j})}, \quad A_{jk} = \frac{\sum_{n=2}^{N}\xi(z_{n-1,j}, z_{nk})}{\sum_{l=1}^{K}\sum_{n=2}^{N}\xi(z_{n-1,j}, z_{nl})} \quad \boldsymbol{\mu}_k = \frac{\sum_{n=1}^{N}\gamma(z_{nk})\mathbf{x}_n}{\sum_{n=1}^{N}\gamma(z_{nk})}, \quad \boldsymbol{\Sigma}_k = \frac{\sum_{n=1}^{N}\gamma(z_{nk})(\mathbf{x}_n - \boldsymbol{\mu}_k)(\mathbf{x}_n - \boldsymbol{\mu}_k)^{\mathrm{T}}}{\sum_{n=1}^{N}\gamma(z_{nk})}$$

   Initially pi is calculated, then the transition matrix. Later Mu (mean) is calculated for the entire sequence of observations. Then using this Mu, Sigma (std dev) is computed.

3. Finally, fit_hmm() is implemented using e_step() and m_step(). At first, the initialization of states and distributions is done. Then e_step() and m_step is computed. After that, the convergence is checked (each parameter is less that -10^4). If convergence is not achieved, then e_step() and m_step are again computed. Once convergence is achieved, the parameters of each state distribution is returned.

Challenges faced:
Figuring the equation for N sequences takes time especially in M-step. Normalization takes time.

Conclusion:
This assignment helps in deeply understand Hidden Markov Models and helps declutter complex equations. Once equations are understood and extrapolated for n sequences, the implementation becomes easier (Note that all equation used in this report are in reference to the lecture slides).