# CS5340 - Lab3 Report

Liu Yichao A0304386A

October 25, 2024

## 1. E-step

**Emission probability** we first compute emission probability with *scipy.stats.norm.pdf* function.

$$emission[i, n, k] = p(x_n^{(i)} | z_{nk} = 1, \mu_k, \sigma_k)$$

Here $i = 1, 2, \ldots, T$ is the index of observed sequence, $n = 1, 2, \ldots, N$ is the index of node inside each sequence, and $k = 1, 2, \ldots, K$ refers to the hidden state.

**$\hat{\alpha}^{(i)}(z_1)$ and $\hat{\beta}^{(i)}(z_N)$** initialize $\hat{\alpha}^{(i)}(z_1) = \frac{\alpha^{(i)}(z_1)}{c_1}$, with $\alpha^{(i)}(z_1) = p(x_1^{(i)}, z_1) = p(x_1^{(i)} | z_1) p(z_1)$ and normalizer $c_1 = \sum_{z_1} \alpha^{(i)}(z_1)$, and initialize $\hat{\beta}^{(i)}(z_N) = \mathbf{1}$

**forward process** for n = 2 to N, compute $\hat{\alpha}^{(i)}(z_n), c_n$ with

$$c_n \hat{\alpha}^{(i)}(z_n) = p(x_n^{(i)} | z_n) \sum_{z_{n-1}} \hat{\alpha}^{(i)}(z_{n-1}) p(z_n | z_{n-1}) := \tilde{\alpha}^{(i)}(z_n)$$

$$c_n = \sum_{z_n} \tilde{\alpha}^{(i)}(z_n)$$

**backward process** for n = N-1 to 1, compute $\hat{\beta}^{(i)}(z_n)$ with

$$c_{n+1} \hat{\beta}^{(i)}(z_n) = \sum_{z_{n+1}} p(x_{n+1}^{(i)} | z_{n+1}) p(z_{n+1} | z_n) \hat{\beta}^{(i)}(z_{n+1}) := \tilde{\beta}^{(i)}(z_n)$$

**compute $\gamma$ and $\xi$** with

$$\gamma^{(i)}(z_n) = \hat{\alpha}^{(i)}(z_n) \hat{\beta}^{(i)}(z_n)$$

$$\xi^{(i)}(z_{n-1}, z_n) = \frac{\hat{\alpha}^{(i)}(z_{n-1}) p(x_n^{(i)} | z_n) p(z_n | z_{n-1}) \hat{\beta}^{(i)}(z_n)}{c_n}$$

## 2. M-step

**update $\pi$** with

$$\pi_k = \frac{\sum_{i=1}^{T} \gamma^{(i)}(z_{1k})}{T}$$

**update $A$** with

$$A_{jk} = \frac{\sum_{i=1}^{T} \sum_{n=2}^{N} \xi^{(i)}(z_{n-1,j}, z_{nk})}{\sum_{i=1}^{T} \sum_{n=2}^{N} \sum_{l=1}^{K} \xi^{(i)}(z_{n-1,j}, z_{nl})}$$

**update $\mu_k$ and $\sigma_k$** here $x_n^{(i)} \in \mathbb{R}$ is sampled from Gaussian distribution $\mathcal{N}(\mu_k, \sigma_k)$

$$\mu_k = \frac{\sum_{i=1}^{T} \sum_{n=1}^{N} \gamma^{(i)}(z_{nk}) x_n^{(i)}}{\sum_{i=1}^{T} \sum_{n=1}^{N} \gamma^{(i)}(z_{nk})}$$

$$\sigma_k = \frac{\sum_{i=1}^{T} \sum_{n=1}^{N} \gamma^{(i)}(z_{nk}) (x_n^{(i)} - \mu_k)^2}{\sum_{i=1}^{T} \sum_{n=1}^{N} \gamma^{(i)}(z_{nk})}$$

### 3. EM algorithm

**initialize** $\pi, A, \mu, \sigma$  with K-means clustering
**update params**  run E-step and M-step iteratively until convergence.

### 4. Results

```
yichao@macbook:~/Documents/CS5340/lab3$ python test_lab3.py
e_step: PASSED
m_step: PASSED
Running on seq_short
---------------------
Loaded 200 sequences, with average length = 8.0
Groundtruth pi:
    [0.9 0.1 0. ]
Groundtruth A:
    [[0.3 0.7 0. ]
    [0.  0.6 0.4]
    [0.  0.  1. ]]
Groundtruth phi:
    {'mu': array([-1.5,  0.5, -0.2]), 'sigma': array([0.5, 0.2, 0.3])}
Your pi:
    [0.   0.91 0.09]
Your A:
    [[1.   0.   0.  ]
    [0.   0.28 0.72]
    [0.4  0.   0.6 ]]
Your phi:
    {'mu': array([-0.21, -1.5 ,  0.52]), 'sigma': array([0.29, 0.51, 0.21])}
Time running fit_hmm() on dataset seq_short : 1.21s

Running on seq_long
---------------------
Loaded 5 sequences, with average length = 1000.0
Groundtruth pi:
    [0.6 0.  0.  0.4]
Groundtruth A:
    [[0.3 0.  0.  0.7]
    [0.7 0.3 0.  0. ]
    [0.  0.6 0.4 0. ]
    [0.  0.  0.4 0.6]]
Groundtruth phi:
    {'mu': array([0. , 1.1, 2. , 1. ]), 'sigma': array([0.5, 0.3, 0.4, 0.2])}
Your pi:
    [0.64 0.   0.36 0.  ]
Your A:
    [[0.3  0.   0.7  0.  ]
    [0.7  0.3  0.   0.  ]
    [0.   0.   0.57 0.43]
    [0.   0.58 0.   0.42]]
Your phi:
    {'mu': array([0.  , 1.07, 1.  , 2.  ]), 'sigma': array([0.52, 0.31, 0.19, 0.41])}
Time running fit_hmm() on dataset seq_long : 3.56s

Running on seq_short2
---------------------
```

```
Loaded 100 sequences, with average length = 9.0
Groundtruth pi:
[0.4 0.3 0.  0.3]
Groundtruth A:
[[0.4 0.6 0.  0. ]
[0.2 0.2 0.6 0. ]
[0.  0.  0.1 0.9]
[0.  0.  0.  1. ]]
Groundtruth phi:
{'mu': array([ 0. , -1. ,  0.5, -0.5]), 'sigma': array([0.2, 0.2, 0.1, 0.1])}
Your pi:
[0.33 0.25 0.42 0.  ]
Your A:
[[0.18 0.   0.2  0.62]
[0.   1.   0.   0.  ]
[0.58 0.   0.42 0.  ]
[0.   0.96 0.   0.04]]
Your phi:
{'mu': array([-0.99, -0.5 ,  0.01,  0.5 ]), 'sigma': array([0.22, 0.1 , 0.17, 0.09])}
Time running fit_hmm() on dataset seq_short2 : 0.45s

Running on seq_long2
---------------------
Loaded 5 sequences, with average length = 800.0
Groundtruth pi:
    [1. 0. 0. 0. 0. 0.]
Groundtruth A:
    [[0.3 0.7 0.  0.  0.  0. ]
    [0.  0.6 0.4 0.  0.  0. ]
    [0.2 0.  0.3 0.5 0.  0. ]
    [0.  0.  0.  0.4 0.5 0.1]
    [0.  0.  0.  0.  0.8 0.2]
    [0.5 0.  0.  0.1 0.  0.4]]
Groundtruth phi:
    {'mu': array([ 0.5,  0. ,  2. , -1. , -0.2,  1.1]), 'sigma': array([0.5 , 0.25,
    0.2 , 0.2 , 0.4 , 0.1 ])}
Your pi:
    [0. 0. 0. 0. 1. 0.]
Your A:
    [[0.6  0.4  0.   0.   0.   0.  ]
    [0.   0.26 0.53 0.   0.2  0.  ]
    [0.   0.   0.43 0.09 0.   0.47]
    [0.   0.   0.08 0.42 0.49 0.  ]
    [0.69 0.   0.   0.   0.31 0.  ]
    [0.   0.   0.   0.21 0.   0.79]]
Your phi:
    {'mu': array([-0.01,  1.99, -0.98,  1.1 ,  0.49, -0.21]), 'sigma': array([0.25, 0.19,
    0.2 , 0.1 , 0.49, 0.39])}
Time running fit_hmm() on dataset seq_long2 : 3.33s
```