Homework 1

Submitted by,

Name: Sayan Chakraborty

Roll No.: EE18MTECH11030

Subject: Probabilistic Graphical Models

Ans 9: We have,

$$Q(\theta, \theta^{old}) = \sum_{Z} p(Z|X, \theta^{old}) \ln(p(X, Z|\theta)) \quad (1)$$

We have,

$$p(X, Z|\theta) = p(z_1|\pi)\left(\prod_{n=2}^{N} p(z_n|z_{n-1}, A)\right) \prod_{n=1}^{N} p(x_n|z_n, \phi) \quad (2)$$

Substituting (2) in (1),

$$Q(\theta, \theta^{old}) = \sum_{Z} p(Z|X, \theta^{old}) \ln(p(X, Z|\theta))$$

$$\Rightarrow Q(\theta, \theta^{old}) = \sum_{Z} p(Z|X, \theta^{old}) \cdot \left[\ln(p(z_1|\pi)) + \sum_{n=2}^{N} \ln(p(z_n|z_{n-1}, A)) + \sum_{n=1}^{N} \ln(p(x_n|z_n, \phi))\right]$$

Now, $p(z_1|\pi) = \prod_{k=1}^{K} (\pi_k)^{z_{1k}}$, $\sum_{k=1}^{K} \pi_k = 1$

$$p(x_n|z_n, \phi) = \prod_{k=1}^{K} p(x_n|\phi_k)^{z_{nk}}$$

$$A_{jk} = p(z_{nk} = 1 | z_{n-1,j} = 1),$$

$$0 \le A_{jk} \le 1, \quad \sum_{k} A_{jk} = 1$$

$$p(z_n|z_{n-1}, A) = \prod_{k=1}^{K} \prod_{j=1}^{K} A_{jk}^{z_{n-1,j}\, z_{n,k}}$$

Now,

$$Q(\theta, \theta^{old}) = \sum_{Z} p(Z|X, \theta^{old}) \cdot \left[\sum_{k=1}^{K} z_{1k} \ln(\pi_k) + \sum_{n=2}^{N} \sum_{k=1}^{K} \sum_{j=1}^{K} (z_{n-1,j}\, z_{nk}) \ln(A_{jk}) + \sum_{n=1}^{N} \sum_{k=1}^{K} z_{nk} \ln(p(x_n|\phi_k))\right]$$

From,

$$\gamma(z_{nk}) = \mathbb{E}[z_{nk}] = \sum_{z} \gamma(z)\, z_{nk},$$

$$\xi(z_{n-1,j}, z_{nk}) = \mathbb{E}[z_{n-1,j}, z_{nk}] = \sum_{z} \gamma(z)\, z_{n-1,j}\, z_{nk},$$

it is clear that

$$Q(\theta, \theta^{old}) = \sum_{z} p(z \mid x, \theta^{old}) \cdot \Big[\sum_{k=1}^{K} z_{1k} \ln(\pi_k) + \sum_{n=2}^{N} \sum_{k=1}^{K} \sum_{j=1}^{K} (z_{n-1,j}, z_{nk}) \ln(A_{jk}) + \sum_{n=1}^{N} \sum_{k=1}^{K} z_{nk} \ln(p(x_n \mid \phi_k)) \Big]$$

$$\Rightarrow Q(\theta, \theta^{old}) = \sum_{k=1}^{K} \gamma(z_{1k}) \ln(\pi_k) + \sum_{n=2}^{N} \sum_{k=1}^{K} \sum_{j=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln(A_{jk}) + \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \ln(p(x_n \mid \phi_k))$$

Ans 2:

Now, we have,

$$Q(\theta, \theta^{old}) = \sum_{k=1}^{K} \gamma(z_{1k}) \ln(\pi_k) +$$

$$\sum_{n=2}^{N} \sum_{k=1}^{K} \sum_{j=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln(A_{jk})$$

$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \ln(p(x_n \mid \phi_k))$$

For π_K

$$Q(\theta, \theta^{old}) = \sum_{k=1}^{K} \gamma(z_{1k}) \ln(\pi_k) +$$

$$\sum_{n=2}^{N} \sum_{k=1}^{K} \sum_{j=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln(A_{jk})$$

$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \ln(p(x_n \mid \phi_k)) + \lambda\left(\sum_{k=1}^{K} \pi_k - 1\right)$$

Now,

$$\frac{\partial Q(\theta, \theta^{old})}{\partial \pi_K} = \gamma(z_{1k}) \, 1/\pi_k + \lambda = 0$$

$$\Rightarrow \lambda = \frac{-\gamma(z_{1k})}{\pi_k}$$

$$\Rightarrow \pi_k \lambda = -\gamma(z_{1k})$$

$$\Rightarrow \sum_{k=1}^{K} \pi_k \lambda = -\sum_{k=1}^{K} \gamma(z_{1k})$$

$$\Rightarrow \lambda = -\sum_{k=1}^{K} \gamma(z_{1k})$$

Now,

$$\gamma(z_{1k}) + \pi_k \lambda = 0$$

$$\Rightarrow \gamma(z_{1k}) + \pi_k\left(-\sum_{k=1}^{K} \gamma(z_{1k})\right) = 0$$

$$\Rightarrow \quad \gamma(z_{1k}) + \pi_k \left(-\sum_{k=1}^{K} \gamma(z_{1k})\right) = 0$$

$$\Rightarrow \boxed{\pi_k = \frac{\gamma(z_{1k})}{\sum_{k=1}^{K} \gamma(z_{1k})}}$$

For A_{jk}

$$Q(\theta, \theta^{old}) = \sum_{k=1}^{K} \gamma(z_{1k}) \ln(\pi_k) + \sum_{n=2}^{N} \sum_{k=1}^{K} \sum_{j=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln(A_{jk})$$
$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \ln(f(x_n | \phi_k)) + \lambda\left(\sum_{k=1}^{K} A_{jk} - 1\right)$$

Now,

$$\frac{\partial Q(\theta, \theta^{old})}{\partial A_{jk}} = 0$$

$$\Rightarrow \sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk}) \frac{1}{A_{jk}} + \lambda = 0$$

$$\Rightarrow \sum_{k=1}^{K} \sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk}) + \sum_{k=1}^{K} A_{jk} \lambda = 0$$

$$\Rightarrow \lambda = -\sum_{k=1}^{K} \sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk})$$

Now,

$$\sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk}) + A_{jk}\lambda = 0$$

$$\Rightarrow \sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk}) + A_{jk}\left(-\sum_{k=1}^{K} \sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk})\right) = 0$$

$$\Rightarrow \sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk}) + A_{jk}\left(-\sum_{k=1}^{K}\sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk})\right) = 0$$

$$\Rightarrow \boxed{A_{jk} = \frac{\sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk})}{\sum_{k=1}^{K}\sum_{n=2}^{N} \xi(z_{n-1,j}, z_{nk})}}$$

For $p(x_n \mid \phi_k)$

$$Q(\theta, \theta^{old}) = \sum_{k=1}^{K} \gamma(z_{1k}) \ln(\pi_k) + \sum_{n=2}^{N}\sum_{k=1}^{K}\sum_{j=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln(A_{jk}) + \sum_{n=1}^{N}\sum_{k=1}^{K} \gamma(z_{nk}) \ln(p(x_n \mid \phi_k))$$

Assuming Gaussian emission densities, we have,

$$p(x_n \mid \phi_k) = \mathcal{N}(x_n \mid \mu_k, \Sigma_k)$$

then,

$$Q(\theta, \theta^{old}) = \sum_{k=1}^{K} \gamma(z_{1k}) \ln(\pi_k) + \sum_{n=2}^{N}\sum_{k=1}^{K}\sum_{j=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln(A_{jk}) + \sum_{n=1}^{N}\sum_{k=1}^{K} \gamma(z_{nk}) \ln(\mathcal{N}(x_n \mid \mu_k, \Sigma_k))$$

Now,

$$Q(\theta, \theta^{old}) = \sum_{k=1}^{K} \gamma(z_{1k}) \ln(\pi_k) + \sum_{n=2}^{N} \sum_{k=1}^{K} \sum_{j=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln(A_{jk})$$

$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \ln(\mathcal{N}(x_n \mid \mu_k, \Sigma_k))$$

$\therefore \phi = \{\mu_k, \Sigma_k\}$

$$\therefore \frac{\partial Q(\theta, \theta^{old})}{\partial \mu_k} = 0$$

$$\Rightarrow \sum_{n=1}^{N} \frac{\gamma(z_{nk})}{\mathcal{N}(x_n \mid \mu_k, \Sigma_k)} \frac{\partial}{\partial \mu_k} \ln(\mathcal{N}(x_n \mid \mu_k, \Sigma_k)) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \frac{\gamma(z_{nk})}{\mathcal{N}(x_n \mid \mu_k, \Sigma_k)} \frac{\partial}{\partial \mu_k} \left(\frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} \exp\left(-\frac{1}{2} (x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k)\right)\right) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \frac{\gamma(z_{nk})}{\mathcal{N}(x_n \mid \mu_k, \Sigma_k)} \mathcal{N}(x_n \mid \mu_k, \Sigma_k) \frac{\partial}{\partial \mu_k} \left(-\frac{1}{2} (x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k)\right) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \gamma(z_{nk}) \left\{ \Sigma^{-1} (x_n - \mu_k) \right\} = 0$$

$$\Rightarrow \sum_{n=1}^{N} \gamma(z_{nk}) x_n - \sum_{n=1}^{N} \gamma(z_{nk}) \mu_k = 0$$

$$\Rightarrow \boxed{\mu_K = \frac{\sum_{n=1}^{N} \gamma(z_{nk})\, x_n}{\sum_{n=1}^{N} \gamma(z_{nk})}}$$

Now,

$$\frac{\partial Q(\theta, \theta^{old})}{\partial \Sigma_k} = 0 \quad \left(\sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \ln\left(\mathcal{N}(x_n \mid \mu_k, \Sigma_k)\right) \right)$$

$$\Rightarrow \sum_{n=1}^{N} \frac{\gamma(z_{nk})}{\mathcal{N}(x_n \mid \mu_k, \Sigma_k)} \frac{\partial}{\partial \Sigma_k} \left(\frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} \exp\left(-\tfrac{1}{2} (x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k)\right) \right) = 0$$

—— ③

Now,

$$\frac{\partial}{\partial \Sigma_k} \left(\frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} \exp\left(-\tfrac{1}{2} (x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k)\right) \right)$$

$$= -\tfrac{1}{2} \frac{\cancel{(2\pi)^d}}{\cancel{(2\pi)^d}} \frac{\cancel{|\Sigma_k|}\, \Sigma_k^{-1}}{\cancel{|\Sigma_k|} \sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\tfrac{1}{2} (x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k)\right)$$

$$+ \frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} \exp\left(-\tfrac{1}{2} (x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k)\right) \left(\Sigma_k^{-1} (x_n - \mu_k)(x_n - \mu_k)^T \Sigma_k^{-1} \right)$$

$$= -\tfrac{1}{2} \frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} \exp\left(-\tfrac{1}{2} (x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k)\right).$$

$$\left[\Sigma_k^{-1} - \Sigma_k^{-1} (x_n - \mu_k)(x_n - \mu_k)^T \Sigma_k^{-1} \right]$$

$$= -\tfrac{1}{2} \mathcal{N}(x_n \mid \mu_k, \Sigma_k) \left[\Sigma_k^{-1} - \Sigma_k^{-1} (x_n - \mu_k)(x_n - \mu_k)^T \Sigma_k^{-1} \right]$$

—— ④

Using ④ in ③,

③ $\Rightarrow$

$$\sum_{n=1}^{N} \frac{\gamma(z_{nk})}{\mathcal{N}(x_n|\mu_k,\Sigma_k)} \frac{\partial}{\partial \Sigma_k}\left(\frac{1}{\sqrt{(2\pi)^d|\Sigma_k|}} \exp\left(-\tfrac{1}{2}(x_n-\mu_k)^T\Sigma_k^{-1}(x_n-\mu_k)\right)\right) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \frac{\gamma(z_{nk})}{\mathcal{N}(x_n|\mu_k,\Sigma_k)} \cdot \left\{-\tfrac{1}{2}\,\mathcal{N}(x_n|\mu_k,\Sigma_k)\left[\Sigma_k^{-1} - \Sigma_k^{-1}(x_n-\mu_k)(x_n-\mu_k)^T\Sigma_k^{-1}\right]\right\} = 0$$

$$\Rightarrow \sum_{n=1}^{N} \gamma(z_{nk})\left[\Sigma_k^{-1} - \Sigma_k^{-1}(x_n-\mu_k)(x_n-\mu_k)^T\Sigma_k^{-1}\right] = 0$$

by premultiplying by Σ_k, we get,

$$\Rightarrow \sum_{n=1}^{N} \gamma(z_{nk})\left[I - (x_n-\mu_k)(x_n-\mu_k)^T\Sigma_k^{-1}\right] = 0$$

by post multiplying by Σ_k, we get,

$$\Rightarrow \sum_{n=1}^{N} \gamma(z_{nk})\left[\Sigma_k - (x_n-\mu_k)(x_n-\mu_k)^T\right] = 0$$

$$\boxed{\Rightarrow \Sigma_k = \frac{\sum_{n=1}^{N} \gamma(z_{nk})(x_n-\mu_k)(x_n-\mu_k)^T}{\sum_{n=1}^{N} \gamma(z_{nk})}}$$

Ans: 3)

(A) Two phones are picked for the Hidden Markov model implementation:

(i) bha
(ii) bhe

(B) A training set and a testing set was prepared for both these phones with no overlap between the training and testing data. The training and testing sets are:

(i) training_bha.wav
(ii) testing_bha.wav
(iii) training_bhe.wav
(iv) testing_bhe.wav

(b) MFCC was found from the raw speech samples of both the training and testing data of both the phones. Basic 13 element version of the MFCC was used.

(c) Using the training data the HMM parameters $\theta = \{\pi, A, \phi\}$ we found, where $\phi = \{\mu, \Sigma\}$

(d) Finally, a basic two-class classifier was implemented. Here, we first find the HMM parameters for a given training data.

Then, the obtained parameters are compared with the same parameters obtained from the training sets of phones bha and bhe.

The, training set belongs to the phone which yield the least error after comparison.

(e) The classifier was tested for different values of $k = \{3, 5, 7\}$. It was seen that the classifier could correctly classify a given testing set.