CSE 512 Machine Learning HW5

Mihir Chakradeo 111462188

Q1.

[I] HMM with tied mixtures

$$P(O_t | X_t = j, \theta) = \sum_{k=1}^{K} w_{jk} \mathcal{N}(O_t | \mu_k, \Sigma_k) \ \forall j \in \{1, \ldots M\}.$$

Q. ① List all the parameters of this HMM model.

We have 3 distributions:

(i) $P(X_1)$: $(M-1)$ parameters

→ Transition:

(ii) $P(X_{t+1} | X_t)$: $M \times (M-1)$ parameters.

→ Likelihood:

(iii) $P(O_t | X_t)$:

- μ_k : K parameters (Mean)
- Σ_k : K parameters (covariance)
- w_{jk} : MK parameters. (weights of GMM)

Q. ② Derive the E step. What do we need to estimate in the E-step?

For the E step, we need to estimate the following distributions:

(i) $P(X_1 | O_{1:T})$

(ii) $P(X_t | O_{1:T}) \ \forall t$

(iii) $P(X_t, X_{t+1} | O_{1:T}) \ \forall t$.

We calculate these using the forward backward pass as follows :-

→ FORWARD PASS

- $\alpha_1(X_1) = P(O_1|X_1)\, P(X_1)$

  Here, $P(O_1|X_1)$ is given to us

  $P(X_1)$ can be found out by counting.

- Now, we calc. $\alpha_2(X_2) \dots \alpha_T(X_T)$:

$$\alpha_i(X_i) = \sum_{x_{i-1}} P(O_i|X_i)\, P(X_i|X_{i-1}=x_{i-1})\, \alpha_{i-1}(x_{i-1})$$

here, $P(O_i|x_i)$ is given to us.

$P(X_i|X_{i-1}=x_{i-1})$ can be found by counting.

$\alpha_{i-1}(x_{i-1})$ from previous iteration.

→ BACKWARD PASS

- Initialize : $\beta_T(X_T)=1$.
- Generate backwards factors by eliminating X_{i+1} for $i = T-1$ to 1

$$\beta_i(X_i) = \sum_{x_{i+1}} P(O_{i+1}|x_{i+1})\, P(X_{i+1}=x_{i+1}|X_i)\, \beta_{i+1}(x_{i+1})$$

here, $P(O_{i+1}|x_{i+1})$ is given to us.

$P(X_{i+1}=x_{i+1}|X_i)$ can be found by counting.

$\beta_{i+1}(x_{i+1})$ from previous iteration.

∴ we can calculate:

$$P(X_i|O_{1:T}) \propto \alpha_i(X_i)\, \beta_i(X_i) \quad \forall i$$

$$P(X_i|O_{1:T}) = \frac{\alpha_i(X_i)\, \beta_i(X_i)}{\sum_{x_i} \alpha_i(X_i)\, \beta_i(X_i)}$$

→ For calculating $P(X_t, X_{t+1} | O_{1:T})$ $\forall t$ -
can be obtained as :

$$\cancel{P(X_i \ldots)} \; P(X_i, X_{i+1} | O_{1:T}) \propto \alpha_i(X_i) P(X_{i+1}|X_i) P(O_{i+1}|X_{i+1}) \beta_{i+1}(X_{i+1})$$

$$P(X_i, X_{i+1} | O_{1:T}) = \frac{\alpha_i(X_i) P(X_{i+1}|X_i) P(O_{i+1}|X_{i+1}) \beta_{i+1}(X_{i+1})}{\sum_{X_i, X_{i+1}} \alpha_i(X_i) \beta_i(X_i)}$$

Q.(3) Derive the M step. How do we update the parameters of the model.

(i) $P(X_1 = j) = \frac{1}{M} \sum_{i=1}^{M} P(X_1 = j | O^i_{1:T^i})$ $\quad j \in \{1 \ldots M\}$

(ii) $T(j, c) = \sum_{i=1}^{M} \sum_{t=1}^{T-1} P(X_t = j, X_{t+1} = c | O^i_{1:T})$.

(iii) $P(j \to c) = \frac{T(j,c)}{\sum_{c'} T(j,c')}$.

- We update the parameters (μ, Σ, w) of the model with MLE as we did in the previous assignment.