Machine Learning - HW5

11146 3646 - Faizaan Charania

Q1. HMM with tied mixtures

The given observation model is of the form:

$$P(O_t \mid X_t = j, \theta) = \sum_{k=1}^{K} \omega_{jk} \mathcal{N}(O_t \mid \mu_k, \Sigma_k)$$

$$\forall_j \in \{1, \ldots M\}$$

We assume that we have M types of hidden states

1) List all parameters of this model.

We deal with 3 different types of parameters

a. $P(X_1) \rightarrow P(X_1)$ is a $M \times 1$ vector

$\therefore$ we have $(M-1)$ parameters since we can calculate the last parameter as

$$(1 - \text{sum(remaining)})$$

b. $P(X_{t+1} | X_t)$: the transitional probability matrix, the dimensions will be

$[M \times (M-1)]$ matrix

c. $P(O_t | X_t)$: the likelihood estimate

It has K Normal distributions

$\therefore$ we have

μ : K parameters

Σ : K parameters

w_{jk} for each state

$\therefore$ [$M \times K$ matrix]

2) E - step

We need to estimate

- $P(X_1 | O_{1...T})$

$P(X_t | O_{1...T}) \; \forall t$

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

We calculate these using the forward backward algorithm

Forward pass

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

$P(X_1)$ can be calculated using counting

$P(O_1|X_1)$ can be calculated using the given distribution ~~of~~ for observations.

For $i = 2 \dots i = T$:

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

like before

$P(O_i \mid x_i)$ is given as a gaussian mixture
$P(X_i \mid X_{i-1})$ estimated by counting
$\alpha_{i-1}(x_{i-1})$ from previous iterate

Backward pass

<u>Initialisation:</u> $\beta_T(X_T) = 1$

For $i = T-1$ to 1

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

where

$P(O_{i+1} \mid x_{i+1}) \rightarrow$ by gaussian mixture
$P(X_{i+1} \mid X_i) \rightarrow$ by counting
$\beta_{i+1}(x_{i+1}) \rightarrow$ from previous iteration

<u>Marginal distribution:</u>

$$P(X_i \mid O_{1 \ldots T}) = \frac{\alpha_i(x_i)\, \beta_i(x_i)}{\sum_{x_i} \alpha_i(x_i)\, \beta_i(x_i)} \quad \forall i$$

and

$$P(x_i, x_{i+1} \mid O_{1 \ldots T}) = \frac{\alpha_i(x_i)\, P(x_{i+1} \mid x_i) \cdot P(O_{i+1} \mid x_{i+1}) \times \beta_{i+1}(x_{i+1})}{\sum_{x_i, x_{i+1}} \alpha_i(x_i)\, P(x_{i+1} \mid x_i)\, P(O_{i+1} \mid x_{i+1})\, \beta_{i+1}(x_{i+1})}$$

3) <u>M-step</u> :

Updation of model parameters

$$P(x_1 = j) = \cancel{\frac{1}{f} \sum_{i=1} P(x_1 = j \mid O_{1 \ldots T})}$$

~~from E-step~~

$$= \frac{1}{f} \sum_{i=1}^{f} P(x_1 = j \mid O_{1 \ldots T})$$

$\hookrightarrow$ from E-step

i.e. we take weighted sum for initial probability

n = Num of sequences

ii) Intermediate term :

$$T(j, l) = \sum_{i=1}^{f} \sum_{t=1}^{T-1} P(X_t = j, X_{t+1} = l \mid O_{1 \ldots T})$$

where $j, l \in \{1 \ldots M\}$

$\therefore$ The transitional probability,

iii) $$P(j \rightarrow l) = \frac{T(j, l)}{\sum_y T(j, y)}$$

iv) For updating the parameter of the the distribution μ_k, Σ_k, w_{jk} we can use MLE as we did in previous assignment.