Homework 5

Name: Edward Ajayi

Andrew ID: eaajayi

Program: MS EAI

Institution: Carnegie Mellon University Africa

Semester: Spring 2025

Unit: 18-661 - Introduction to Machine Learning for Engineers

Faculty: Prof. Carlee Joe-Wong and Prof. Gauri Joshi

Date: April 29, 2025

Question 1: Gaussian Mixture Models

We are given a 1-D mixture model where each component is an exponential distribution:

$$p(x) = \sum_{k=1}^{K} \omega_k \cdot \text{Exp}(x \mid \mu_k)$$

where:

- ω_k: mixture weight for component k, with $\omega_k \geq 0$ and $\sum_{k=1}^{K} \omega_k = 1$
- μ_k: rate of the exponential distribution for component k
- $\text{Exp}(x \mid \mu) = \mu e^{-x\mu}$ for $x \geq 0$

Let $\{x_n\}_{n=1}^{N}$ be the observed data, and $z_n \in \{1, \ldots, K\}$ be the latent indicator of the component generating x_n. Let $r_{nk} = \mathbb{I}[z_n = k]$ be the binary indicator variable for $z_n = k$.

(a) Complete Log-Likelihood

We want to write down the complete data log-likelihood:

$$\log p(\{x_n, z_n\}_{n=1}^{N}) = \sum_{n=1}^{N} \log p(x_n, z_n)$$

Assuming $p(z_n = k) = \omega_k$ and $p(x_n \mid z_n = k) = \mu_k e^{-x_n \mu_k}$, we get:

$$p(x_n, z_n = k) = \omega_k \mu_k e^{-x_n \mu_k}$$

Therefore, the complete log-likelihood becomes:

$$\log p(\{x_n, z_n\}) = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \log(\omega_k \mu_k e^{-x_n \mu_k}) = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \left(\log \omega_k + \log \mu_k - x_n \mu_k\right)$$

$$\boxed{\log p(\{x_n, z_n\}) = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \left(\log \omega_k + \log \mu_k - x_n \mu_k\right)}$$

(b) M-Step: Maximizing with respect to μ_k

We isolate the relevant terms of the log-likelihood for component k:

$$L(\mu_k) = \sum_{n=1}^{N} r_{nk} \left(\log \mu_k - x_n \mu_k \right)$$

Differentiate with respect to μ_k:

$$\frac{dL}{d\mu_k} = \sum_{n=1}^{N} r_{nk} \left(\frac{1}{\mu_k} - x_n \right)$$

Set the derivative to zero:

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

$$\boxed{\mu_k = \frac{\sum_{n=1}^{N} r_{nk}}{\sum_{n=1}^{N} r_{nk} x_n}}$$

(c) E-Step: Computing Soft Labels r_{nk}

We want to compute the posterior:

$$r_{nk} = P(z_n = k \mid x_n) = \frac{P(x_n \mid z_n = k) P(z_n = k)}{\sum_{j=1}^{K} P(x_n \mid z_n = j) P(z_n = j)}$$

Substitute in the exponential pdf and priors:

$$r_{nk} = \frac{\omega_k \mu_k e^{-x_n \mu_k}}{\sum_{j=1}^{K} \omega_j \mu_j e^{-x_n \mu_j}}$$

$$\boxed{r_{nk} = \frac{\omega_k \mu_k e^{-x_n \mu_k}}{\sum_{j=1}^{K} \omega_j \mu_j e^{-x_n \mu_j}}}$$