Homework 5 - Introduction to Machine Learning For Engineers

kipngeno koech - bkoech

April 18, 2025

1 Gaussian Mixture Models

Consider an exponential mixture model for a 1-D dataset $\{x_n\}$ with the density function

$$p(x) = \sum_{k=1}^{K} \omega_k \mathrm{Exp}(x|\mu_k),$$

where K is the number of mixture components, μ_k is the rate parameter, and ω_k is the mixture weight corresponding to the k-th component. The exponential distribution is given by

$$\mathrm{Exp}(x|\mu) = \mu \exp(-x\mu) \quad \text{for all } x \geq 0. \tag{1}$$

We would like to derive the model parameters (ω_k, μ_k) for all k using the EM algorithm. Consider the hidden labels $z_n \in \{1, \dots, K\}$ and indicator variables r_{nk} that are 1 if $z_n = k$ and 0 otherwise. The complete log-likelihood (assuming base e for the log) is then written as

$$\sum_n \log p(x_n, z_n) = \sum_n \sum_{z_n=k} \left[\log p(z_n = k) + \log p(x_n|z_n = k)\right].$$

1. Write down and simplify the expression for the complete log-likelihood for the exponential mixture model described above. Plugging the definition of the exponential distribution here immediately gives

$$\sum_n \log p(x_n, z_n) = \sum_k \sum_n r_{nk} \left[\log \omega_k + \log \mathrm{Exp}(x_n|\mu_k)\right] = \sum_k \sum_n r_{nk} \left[\log \omega_k + \log \mu_k - x_n \mu_k\right].$$

2. Solve the M step of the EM algorithm and find μ_k for $k = 1, \dots, K$ that maximizes the complete log-likelihood. Taking the derivative of the log-likelihood with respect to μ_k and setting it to zero, we have:

$$\frac{1}{\mu_k} \sum_n r_{nk} - \sum_n r_{nk} x_n = 0. \tag{2}$$

$$\mu_k = \frac{\sum_n r_{nk}}{\sum_n r_{nk} x_n}. \tag{3}$$

3. Perform the E step of the EM algorithm and write the equation to update the soft labels $r_{nk} = P(z_n = k|x_n)$. Using Bayes' rule, we have:

$$r_{nk} = \frac{P(x_n, z_n = k)}{P(x_n)} = \frac{\omega_k \mu_k \exp(-x_n \mu_k)}{\sum_{k'} \omega_{k'} \mu_{k'} \exp(-x_n \mu_{k'})}.$$