Exercise 1: Gaussian mixture model

Gaussian Mixture Model

We have the model

$$\begin{aligned} \mu_k &\sim \mathcal{N}(0, \sigma^2), \quad k = 1, \dots, K; \\ c_i &\sim \text{Categorical}(1/K, \dots, 1/K), \quad i = 1, \dots, n; \\ Y_i \mid c_i, \boldsymbol{\mu} &\sim \mathcal{N}(\mu_{c_i}, 1), \quad i = 1, \dots, n. \end{aligned}$$

where we assume σ^2 is known.

Exercise part I

Approximate the posterior

$$p(\boldsymbol{\mu}, \boldsymbol{c} \mid \boldsymbol{y}) \propto p(\boldsymbol{\mu}, \boldsymbol{c}, \boldsymbol{y}) = p(\boldsymbol{\mu}) \prod_{i=1}^{n} p(c_i) p(y_i | c_i, \boldsymbol{\mu});$$

with the variational approximation

$$q(\boldsymbol{\mu}, \boldsymbol{c}) = \prod_{k=1}^{K} q(\mu_k) \prod_{i=1}^{n} q(c_i).$$

For this model:

1. Derive $q(c_i) \propto \exp\left\{ \mathbb{E}_{q_{\boldsymbol{c}_{-i}, \boldsymbol{\mu}}} [\log p(c_i, \boldsymbol{c}_{-i}, \boldsymbol{\mu}, \boldsymbol{y})] \right\}$ for $i = 1, \dots, n$ and $q(\mu_k) \propto \exp\left\{ \mathbb{E}_{q_{\boldsymbol{c}, \boldsymbol{\mu}_{-k}}} [\log p(\boldsymbol{c}, \boldsymbol{\mu}, \boldsymbol{y})] \right\}$ for $k = 1, \dots, K$ to obtain updates;
2. Derive the ELBO $= \mathbb{E}_q [\log p(\boldsymbol{\mu}, \boldsymbol{c}, \boldsymbol{y})] - \mathbb{E}_q [\log q(\boldsymbol{\mu}, \boldsymbol{c})]$.

Solution

We have the mean-field variational approximation factors

$$q(\mu_k) \propto \mathcal{N}(m_k, s_k^2); \;\; q(c_i) \propto \phi_{i,c_i} \propto \exp\left\{ y_i m_{c_i} - \frac{1}{2} s_{c_i}^2 - \frac{1}{2} m_{c_i}^2 \right\};$$

where

$$m_k = \frac{\sum_{i=1}^{n} \phi_{i,k} y_i}{1/\sigma^2 + \sum_{i=1}^{n} \phi_{i,k}}; \quad s_k^2 = \left(\frac{1}{\sigma^2} + \sum_{i=1}^{n} \phi_{i,k} \right)^{-1}; \quad \sum_{k=1}^{K} \phi_{i,k} = 1;$$

and the ELBO was derived to be

$$\text{ELBO} = -\frac{1}{2\sigma^2} \sum_{k=1}^{K} \left[s_k^2 + m_k^2 \right] + \sum_{i=1}^{n} \sum_{k=1}^{K} \phi_{i,k} \left[y_i m_k - \frac{1}{2} s_k^2 - \frac{1}{2} m_k^2 \right] - \sum_{i=1}^{n} \sum_{k=1}^{K} \phi_{i,k} \log \phi_{i,k} + \frac{1}{2} \sum_{k=1}^{K} \log s_k^2 + \text{const.}$$

Exercise part II

Implement the CAVI algorithm, and run it on simulated data with $K = 3$, $\boldsymbol{\mu} = (-1, 1, 3)$, $\sigma^2 = 1$, $n = 300$. Use the initialisation $\phi_{i,c_i} = 1/K$ for all $i = 1, \ldots, n$, and $m_1 = 1, m_2 = 2, m_3 = 3$ and $s_k^2 = 0.5$ for all $k = 1, \ldots, K$. Use the ELBO to assess convergence. and estimate 95% credible intervals for $\mu_k, k = 1, \ldots, K$ from their estimated distribution. What happens if you instead initialise all m_k with the same value?

The R code is given in a separate file.