Gaussian Mixture Model Parameter Estimation with Prior Hyper Parameters

Learning Objective: Gain insight into how the Expectation-Maximization (EM) algorithm can be used to fit a two-component Gaussian mixture model.

Expectation Maximization

Implementing the Expectation Maximization (EM) algorithm for Gaussian mixtures[1]. Assume that the sample data consist of two 1D Gaussians with equal mixing coefficients, where each Gaussian has variance equal to 1, but they have different expectations. The sample data can be visualized as a histogram.

The EM algorithm is comprised of two main parts: the E step and the M step that are run once per iteration. The first step calculates the expected value of each hidden variable given the current hypothesis (expectation in our case). This can be seen on the left-hand side of Equation 1, where x_i is the i^{th} sample, k is the number of Gaussian components, μ is the mean value, σ is the standard deviation, and pdf_N is the probability density function for a univariate Gaussian distribution.

$$E[z_{ij}] = \frac{pdf_{\mathcal{N}}(x_i, \mu_j, \sigma)}{\sum_{n=1}^{k} pdf_{\mathcal{N}}(x_i, \mu_n, \sigma)} \qquad pdf_{\mathcal{N}}(x, \mu, \sigma) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \tag{1}$$

The second step calculates maximum likelihood assuming that the hidden variables take on the expected values from the E step. This can be seen in Equation 2, where m is the number of samples.

$$\mu_j = \frac{\sum_{i=1}^{m} E[z_{ij}] x_i}{\sum_{i=1}^{m} E[z_{ij}]} \tag{2}$$

----

[1]In literature, a Gaussian mixture model (GMM) may sometimes be referred to as a mixture of Gaussians (MoG). In general, a MoG is simply a weighted sum or *mixture* of *k* Gaussian distributions with different means and covariances.