Math 156 Term Project 5

Deadline: Wednesday June 10th 11:59pm

Problem 1

Finish Textbook problem 9.9 (p. 456). (50 points)

Problem 2

In Problem 1, we permit different covariance matrix $\mathbf{\Sigma}_k$ for each cluster k. Now suppose we impose that every cluster has identity covariance matrix $\mathbf{I}$. This changes the expected value of the complete-data log likelihood function (from textbook equation (9.40)) to

$$\mathbb{E}_{\mathbf{Z}}[\ln p(\mathbf{X}, \mathbf{Z}|\boldsymbol{\mu}, \boldsymbol{\pi})] = \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk})(\ln \pi_k + \ln \mathcal{N}(\mathbf{x}_n|\mu_k, \mathbf{I})).$$

(Note the difference in notations for the textbook and the lecture notes.) Show that maximizing the above function with regard to $\boldsymbol{\mu}_k$ $(k = 1, \cdots, K)$ is equivalent to minimizing the γ-weighted least square error function

$$E(\boldsymbol{\mu}_1, \cdots, \boldsymbol{\mu}_K) = \sum_{i=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \|\mathbf{x}_n - \boldsymbol{\mu}_k\|^2 .(50 \text{ points})$$