Machine Learning

Quiz 6

Student Name: _____

Consider points $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbb{R}^D$ and integer $K > 1$. K-means aims at minimizing the within cluster variance:

$$J(\boldsymbol{\mu}, l|\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_{n=1}^{N} \sum_{k=1}^{K} l_{nk} |\mathbf{x}_n - \boldsymbol{\mu}_k|^2$$

where each $\boldsymbol{\mu}_k \in \mathbb{R}^D$ is the center of cluster k, and $l_{nk} \in \{0, 1\}$, s.t. $l_{nk} = 1$ iff sample $\mathbf{x}_n$ is assigned to cluster k.

1. (1 point) What are the two alternating steps of the k-means algorithm?

   1. 2.

2. (1 point) Does the algorithm converge? If yes, to what?

3. (1 point) Label the two steps of k-means above as either **E** or **M** step in a EM-framework.

4. (1 point) Briefly: What is the main difference between k-means and EM for Gaussian mixture?