Practical Session 11

Clustering

1 K-medians

Problem 1: Consider a modified version of the K-means objective, where we use L_1 distance instead.

$$J(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\mu}) = \sum_{i=1}^{N} \sum_{k=1}^{K} z_{ik} \|\boldsymbol{x}_i - \boldsymbol{\mu}_k\|_1$$

This variation of the algorithm is called K*-medians*. Derive the Lloyd's algorithm for this model.

2 Gaussian mixture model

Problem 2: Derive the E step update for Gaussian mixture model.

Problem 3: Derive the M step update for Gaussian mixture model.

3 Expectation Maximization algorithm

Problem 4: Consider a mixture model where the components are given by independent Bernoulli variables. This is useful when modelling, e.g., binary images, where each of the D dimensions of the image $\boldsymbol{x}$ corresponds to a different pixel that is either black or white. More formally, we have

$$p(\boldsymbol{x} | \boldsymbol{z} = k) = \prod_{d=1}^{D} \theta_{kd}^{x_d} (1 - \theta_{kd})^{1 - x_d}.$$

That is, for a given mixture index $\boldsymbol{z} = k$, we have a product of independent Bernoullis, where θ_{kd} denotes the Bernoulli parameter for component k at pixel d.

Derive the EM algorithm for the parameters $\theta = \{\theta_{kd} \mid k = 1, \dots, K, d = 1, \dots, D\}$ of a mixture of Bernoullis.

Assume here for simplicity, that the distribution of components $p(\boldsymbol{z})$ is uniform: $p(\boldsymbol{z}) = \prod_{k=1}^{K} \pi_k^{z_k} = \prod_{k=1}^{K} (\frac{1}{K})^{z_k}$.