Machine Learning Worksheet 11

Shang-Hsin Yu – 03681048 – shanghsin.yu@tum.de

Problem 1

From a Gaussian Mixture Model we have the E step:

$$r_{nk}(\theta) = p(z_n = k \mid x_n, \theta) = \frac{\pi_k \mathcal{N}(x_n \mid \mu_k, \Sigma_k)}{\sum_{i=1}^{K} \pi_i \mathcal{N}(x_n \mid \mu_i, \Sigma_i)}$$

and the M step:

$$\pi_k = \sum_{n=1}^{N} \frac{r_{nk}(\theta)}{\sum_{i=1}^{K} r_{ni}(\theta)}$$

$$\mu_k = \sum_{n=1}^{N} \frac{x_n r_{nk}(\theta)}{r_{nk}(\theta)}$$

But as we fix all Σ to $\sigma^2 I$ and have $\sigma \to 0$, The Gaussian Distributions started to become impulses and the Gaussian that has mean closest to the point have overwhelming influence over the outcome in the E step, turning the responsibilities into 1-hot encoding. And as responsibilities turn into 1-hot encoding, π_k also becomes 0 or 1 depending if the point is closest to k or not, μ_k becomes the mean of all points "classified" as k. Which makes the EM algorithm here identical to the K-Means algorithm.