**University of Illinois** Spring 2018

# CS 446: Machine Learning
## Homework 9

Due on Tuesday, April 3, 2018, 11:59 a.m. Central Time

1. **[16 points]** Gaussian Mixture Models & EM
Consider a Gaussian mixture model with $K$ components ($k \in \{1, \ldots, K\}$), each having mean $\mu_k$, variance $\sigma_k^2$, and mixture weight $\pi_k$. All these are parameters to be learned, and we subsume them in the set $\theta$. Further, we are given a dataset $X = \{x_i\}$, where $x_i \in \mathbb{R}$. We also use $Z = \{z_i\}$ to denote the latent variables, such that $z_i = k$ implies that $x_i$ is generated from the $k^{th}$ Gaussian.

(a) What is the log-likelihood of the data $\log p(X; \theta)$ according to the Gaussian Mixture Model? (use $\mu_k$, $\sigma_k$, $\pi_k$, $K$, $x_i$, and $X$). Don't use any abbreviations.

| Your answer: |
|---|

(b) For learning $\theta$ using the EM algorithm, we need the conditional distribution of the latent variables $Z$ given the current estimate of the parameters $\theta^{(t)}$ (we will use the superscript $(t)$ for parameter estimates at step $t$). What is the posterior probability $p(z_i = k|x_i; \theta^{(t)})$? To simplify, wherever possible, use $\mathcal{N}(x_i|\mu_k, \sigma_k)$ to denote a Gaussian distribution over $x_i \in \mathbb{R}$ having mean $\mu_k$ and variance $\sigma_k^2$.

| Your answer: |
|---|

(c) Find $\mathbb{E}_{z_i|x_i;\theta^{(t)}}[\log p(x_i, z_i; \theta)]$. Denote $p(z_i = k|x_i; \theta^{(t)})$ as $z_{ik}$, and use all previous notation simplifications.

| Your answer: |
|---|

(d) $\theta^{(t+1)}$ is obtained as the maximizer of $\sum_{i=1}^{N} \mathbb{E}_{z_i|x_i;\theta^{(t)}}[\log p(x_i, z_i; \theta)]$. Find $\mu_k^{(t+1)}$, $\pi_k^{(t+1)}$, and $\sigma_k^{(t+1)}$, by using your answer to the previous question.

Your answer:

(e) How are kMeans and Gaussian Mixture Model related? (There are three conditions)

Your answer: