1. Investigating the EM Algorithm for Gaussian Micture Models

a) $r_{ik} = \frac{\lambda_k Norm_{x_i}[\mu_k, \sum k]}{\sum_{j=1}^{K} \lambda_j Norm_{x_i}[\mu_j, \sum j]} = \frac{\frac{\lambda_k}{2\pi^{1/2} |\sum k|^{1/2}} \exp\left(-\frac{1}{2}(x_i - \mu_k)^T \sum_k^{-1}(x_i - \mu_k)\right)}{\sum_{j=1}^{K} \frac{\lambda_j}{2\pi^{1/2} |\sum j|^{1/2}} \exp\left(-\frac{1}{2}(x_i - \mu_j)^T \sum_j^{-1}(x_i - \mu_j)\right)}$

b) We can remove $2\pi^{1/2}$

$$C_i = \frac{1}{\sum_{j=1}^{K} \frac{\lambda_j}{|\sum j|^{1/2}} \exp\left(-\frac{1}{2}(x_i - \mu_j)^T \sum_j^{-1}(x_i - \mu_j)\right)}$$

and because $\lambda_j = \frac{1}{K}$ $\sum = I$, so

$$C_i = \frac{K}{\sum_{j=1}^{K} \exp\left(-\frac{1}{2}(x_i - \mu_j)^T (x_i - \mu_j)\right)}$$

$$\log C_i = \log K - \log \sum_{j=1}^{K} \exp\left(-\frac{1}{2}(x_i - \mu_j)^T (x_i - \mu_j)\right)$$

$$\leq -\sum_{j=1}^{K} \log \exp\left(-\frac{1}{2}(x_i - \mu_j)^T (x_i - \mu_j)\right)$$

Because all mixtures have unit variance and $\sum \lambda_j = 0$, so

$$= -\sum_{j=1}^{K} \left(-\frac{1}{2}(x_i - \mu_j)^T (x_i - \mu_j)\right)$$

$$= \frac{1}{2} \sum_{j=1}^{K} (x_i - \mu_j)^T (x_i - \mu_j)$$

I think $\log C_i$ is the value of x_i in the curve of Gaussian Mixture Model.

c) The formula is $\mu_k = \frac{\sum_{i=1}^{I} r_{ik} x_i}{\sum_{i=1}^{I} r_{ik}}$

If $r_{ik} = 1$ for the k with highest probability, the above formula will become $\mu_k = \frac{\sum_{i=1}^{I} x_i}{I} = E(x)$ it will not be updated when training and the EM algorithm for Gaussian Mixture Models will break down to many totally same single Gaussian models. According to λ_k is uniform and $\sum k = I$, the final model will sum up to one uniform Gaussian model.