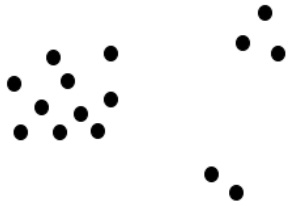
Week 4-Seminar 2

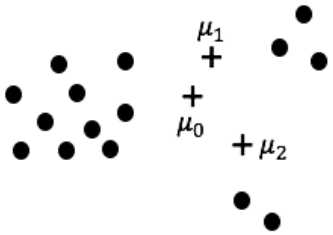
Q1: GMM Theory

Consider the set of training data in the graph below, let's assume it contains three clusters. For GMM, the means and variances of three Gaussians are μ_0 and σ_0, μ_1 and σ_1, and μ_2 and σ_2, respectively. Additionally, we have π_0, π_1, π_2 to denote the mixture proportions of the three Gaussians (i.e., $p(x) = \pi_0 N(\mu_0, \sigma_0 I) + \pi_1 N(\mu_1, \sigma_1 I) + \pi_2 N(\mu_2, \sigma_2 I)$), where I is the identity matrix and $\pi_0 + \pi_1 + \pi_2 = 1$. We will also use θ to refer to the entire collection of parameters $(\mu_0, \mu_1, \mu_2, \sigma_0, \sigma_1, \sigma_2, \pi_0, \pi_1, \pi_2)$ defining the mixture model $p(x)$.



(a) Would K-Means ($K = 3$) and our 3-cluster GMM trained using EM produce the same cluster centers (means) for this data set above? Justify your answer. (Answer without any justification will get zero point.)

(b) In the following, we apply EM to train our 3-cluster GMM on the data below. The '+' points indicate the current means μ_0, μ_1, and μ_2 of the three Gaussians after the kth iteration of EM.



(b.1) On the figure, draw the directions in which μ_0, μ_1 and μ_2 will move in the next EM iteration.

(b.2) Will the marginal likelihood of the data, $\prod_j P(x^j|\theta)$ increase or decrease on the next EM iteration? Explain your reasoning.

(b.3) Will t he estimate of π_0 increase or decrease on t he next EM iteration? Explain your reasoning.