Assignment 4

Problem 1

Suppose we observe a set of data points $x_{1..4} = \{-1, 0.5, 1, 100\}$ and we would like to use Gaussian Mixture Model (GMM) for clustering D. Consider the following mixture of two Gaussians: $\mathcal{N}(0, 10)$ and $\mathcal{N}(1, 2)$ and we believe the data points are drawn from $\mathcal{N}(0, 10)$ 40% of the time and $\mathcal{N}(1, 2)$ 60% of the time.

1. Calculate the responsibility of each component for each data point in $x_{1..4}$.
2. For $x_4 = 100$ which component has higher responsibility? Briefly explain why that is the case.

Problem 2

Suppose we want to model traffic on a given road. At each hour i, we measure the number of cars y_i. We think the number of cars is well modeled by a Gaussian distribution. However, we know that some times are rush hour, so we decide to use a mixture of Gaussian distributions with two components (normal and rush hour traffic respectively.) For simplicity, imagine that traffic at each hour is independent of one another.

Draw a graphical model that represents this mixture model. You should use plate notation and include variables $\mathbf{z} = (z_1, ..., z_n)$ and $\mathbf{y} = (y_1, ..., y_n)$ as well as the parameters of the distributions π, μ_1, μ_2, σ_1 and σ_2 in your graph.

Problem 3

Consider samples $x_1, \ldots, x_n$ from a Gaussian random variable with known variance σ^2 and unknown mean μ. We further assume a prior distribution (also Gaussian) over the mean, $\mu \sim \mathcal{N}(m, s^2)$, with fixed mean m and fixed variance s^2. Thus the only unknown is μ.

1. Calculate the MAP estimate $\hat{\mu}_{MAP}$. You can state the result without proof. Alternatively, with a bit more work, you can compute derivatives of the log posterior, set to zero and solve.
2. Show that as the number of samples n increase, the MAP estimate converges to the maximum likelihood estimate.
3. Suppose n is small and fixed. What does the MAP estimator converge to if we increase the prior variance s^2?
4. Suppose n is small and fixed. What does the MAP estimator converge to if we decrease the prior variance s^2?

Assignment 4

Problem 4

Please write one thing from this course you found confusing, a topic you would like to hear more about, or something you found particularly interesting.