COMS21202: Symbols, Patterns and Signals

Lab 4: Maximum Likelihood Estimation

NOTE: You will need to refer to lecture 4 and Matlab help pages to complete this exercise.

1. You believe that your data follows a normal distribution. Assuming the standard deviation (σ) is 0.5, you wish to estimate the mean μ of the normal distribution representing your data. You thus have a model with a single parameter μ you wish to tune/train. Discuss with your lab partner how the likelihood $p(D|\mu)$ for some observations $D = \{d_1, \cdots, d_N\}$ can be represented by (assuming independent observations):

$$p(D|\theta) = \prod_i \mathcal{N}(d_i|\theta, 0.25) \tag{1}$$

$$= \prod_i \frac{2}{\sqrt{2\pi}} e^{-2(d_i-\mu)^2} \tag{2}$$

2. Use the Maximum Likelihood Estimate (MLE) recipe to find μ_{ML}. [*Note:* This is done on paper, not using Matlab]

3. In this lab, we aim to help you understand MLE by experimenting with different values of μ to find $\mu_{ML} = arg\,max_\mu\, p(D|\mu)$.

   (a) Load the data from file 'data1.dat' and plot a histogram of the data.

   (b) Write a function *computeLikelihood(mu)* that takes a value of μ (e.g. μ = 0), and computes $p(D|\mu)$ using Equation (2) for the data in 'data1.dat'. Note: Do NOT use a for loop. You may use the matlab function *prod* in the calculation.

   (c) Write a function *loopLikelihood()* that loops through possible values of $\mu \in \{0.00, 0.01, 0.02, ..., 1.00\}$, calls *computeLikelihood(mu)* for each value and stores an array of all likelihood values. You can use a loop for this function.

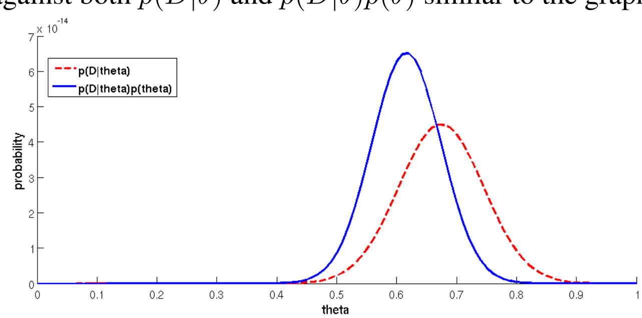
   (d) Based on your calculation, what would $\max p(D|\mu)$ be? What would $arg\,max_\mu\, p(D|\mu)$ be?

   Discuss: Make sure you understand the difference between the two.

   (e) Plot μ against $p(D|\mu)$ for different μ values. Can you visually spot μ_{ML}.

   (f) Assume you have prior knowledge of what μ_{ML} should be, $p(\mu) = \mathcal{N}(0.5, 0.01)$. Write functions *computePosterior(mu)* and *loopPosterior()* to find $\mu_{MAP} = \arg\max_\mu p(D|\mu)p(\mu)$.

   (g) plot μ against both $p(D|\theta)$ and $p(D|\theta)p(\theta)$ similar to the graph below.



   (h) Repeat the above calculations for 'data2.dat' and 'data3.dat' and explain your observations.

EXTRA: Until now, you assumed σ = 0.5. Remove this assumption and estimate $\theta_{MAP} = [\mu_{MAP}, \sigma_{MAP}]$ experimentally by looping through possible values of μ and σ. Assume the prior probability for $p(\sigma)$ is $\mathcal{N}(0.5, 0.16)$.

EXTRA: Plot (μ, σ) against $p(D|\theta)p(\theta)$ similar to the mesh graph below (use the function $meshc$ in Matlab)