IST 597: Homework 1

Justin Silverman

Problem 1 (10 points)

This problem is designed to test your understanding of notation we will be using frequently in this course. For $i \in \{1, \ldots, N\}$ and for $y \in \mathcal{R}$

$$y_i \sim N(\mu, \sigma^2)$$

Write the functional form of $p(y_1, \ldots, y_N)$. I am looking for you to simplify this expression.

Hint: You need to use the product symbol $\prod_{i=1}^{N}$.

Problem 2 (40 points)

You are working with the following Bayesian linear model:

$$y \sim N(\beta x, \sigma^2) \tag{1}$$

$$\beta \sim N(\mu, \tau^2) \tag{2}$$

This joint distribution can be denoted as $p(y, \beta | \sigma^2, \mu, \tau^2)$. Derive the marginal distribution $p(y|\sigma^2, \mu, \tau^2)$. Note, here this model is for a single scalar value of y (a single observation).

Hint: If you end up using integration you are doing this the hard way.

Problem 3 (20 points)

Joe and Xiang live together. Joe is a very boring man, he only does four things per day: eat, sleep, and work. In contrast, Xiang is much more lively and does a number of activities daily including playing loud music (which bothers Joe because he is a boring husk of a man). Being the dullard that he is, Joe decides to study how his productivity (measured as the number of hours he works per day) is related to the volume of Xiang's music (measured in decibels). Over the course of 1 month Joe collects daily measurements (y_i, x_i) for where y_i is his productivity on day i and x_i is the volume on day i. Joe decides to use a linear model of the form: $y_i = \alpha + \beta x_i + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma^2)$. He fits his model to data using OLS (e.g., same result as if he used maximum likelihood) and gets an estimate $\hat{\beta} = -3$. Joe then want to predict his productivity when Xiang plays his music very loudly. Remarkably, Joe ends up predicting that his productivity would be negative! What mistake has Joe made (beyond living his life as a monotonous bore)?

Problem 4 (10 points + 10 possible extra credit points)

Solve for X and tell me how you did it.

$$\begin{bmatrix} 26 & -15 & 7 \\ 25 & 3 & -9 \\ 45 & -14 & -8 \end{bmatrix} = X \begin{bmatrix} -1 & -5 & 3 \\ 5 & -4 & 0 \\ 1 & 4 & -3 \end{bmatrix} \tag{3}$$

Extra Credit: On this problem, extra credit will be given for the easiest solutions (those where the student had to work the least hard) and those that are the fastest (in terms of computational steps). However, saying "I copied the answer from [insert other students name]" will not get credit even though that would be easy. Also, for my own safety, I would consider hacking my computer to find the answer an exceedingly slow and difficult solution and such a solution would receive negative points. To receive all extra credit points you probably need to mention two solutions: one that is fast and one that is easy.

Problem 5 (20 points)

Sudar is trying to solve the following optimization problem using gradient descent with a step size of 0.2.

$$\underset{b}{argmin} \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{b^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

Where $\nu > 0$ and $\Gamma(x)$ refers to the Gamma function. He is using the built in `optim` function in the R programming language with the an initial value of 3. He is working on a laptop running Windows 8 and using R version 3.6. His laptop memory is near capacity as he has too many Google Chrome tabs open. He keeps getting weird results and a few times his laptop even crashed. What is Sudar doing wrong?