HW #4

Uziel Rivera-Lopez

04/01/2024

Problem 1

1. $\binom{n}{k} \cdot \theta^k \cdot (1-\theta)^{n-k}$
   $= \binom{10}{7} \cdot \theta^7 \cdot (1-\theta)^3$
   $= \frac{10!}{7!(10-7)!} \cdot \theta^7 \cdot (1-\theta)^3$
   $= 120 \cdot \theta^7 \cdot (1-\theta)^3$

2. Posterior $\propto$ Likelihood $\times$ Prior
   $120 \cdot \theta^7 \cdot (1-\theta)^3 \cdot \theta^{2-1} \cdot (1-\theta)^{2-1}$
   $= 120 \cdot \theta^8 \cdot (1-\theta)^4$
   Posterior is $\beta(8,4)$

Problem 2

1. $\mu = \begin{bmatrix} \text{Mean of GDP} \\ \text{Mean of Inflation Rate} \\ \text{Mean of Unemployment Rate} \end{bmatrix} = \begin{bmatrix} 2.94 \\ 2.48 \\ 5.1 \end{bmatrix}$ $\Sigma =$

$\begin{bmatrix} \sigma^2_{GDP} & \sigma^2_{GDP,InflationRate} & \sigma^2_{GDP,UnemploymentRate} \\ \sigma^2_{InflationRate,GDP} & \sigma^2_{InflationRate} & \sigma^2_{IR,UR} \\ \sigma^2_{UnemploymentRate,GDP} & \sigma^2_{UR,IR} & \sigma^2_{UnemploymentRate} \end{bmatrix} =$

$\begin{bmatrix} 0.093 & -0.1015 & 0.0825 \\ -0.1015 & 0.247 & -0.1375 \\ 0.0825 & -0.1375 & 0.1 \end{bmatrix}$

2. The variance between GDP and Unemployment rate is 0.0825, which means that GDP growth is associated with higher unemployment rates. With the variance between Inflation Rate and Unemployment Rate is -0.1375, telling us that inflation and unemployment go together in pairs. Lastly, the variance between GDP and Inflation Rate is -0.1015, which tells us that GDP growth and inflation rate are negatively correlated, with higher GDP causing lower inflation rates.

Problem 3

1. In parallel processing we have two different implementations, SIMD and MIMD. SIMD stands for Single Instruction Multiple Data, and MIMD stands for Multiple Instruction Multiple Data. In SIMD, we just have a single instruction that is executed on the variety of data we have. MIMD, we have to write multiple instructions to be executed on the variety of data. This follows how our neurons work in our brain because we can have multiple neurons working on the same task but different information, while, different neurons can be working on different tasks with different information. Like my eyes right now are taking on the task of reading this text, while my hands are typing this text, MIMD.

2. The perceptron is just a node with a function that takes in inputs and gives an output from the function. So like the terminals of a neuron, the perceptron takes in inputs, with the weights of the inputs, which indicate the level of importance for each input.

3. Online learning we don't have to have all the data at once, we can just have the data be delivered in batches. Yet, we need to be online at all times when the model is learning. With offline learning, we have to download all the data at once, and then we can start the learning process. This takes up too much time and resources, but we don't have to be online at all times.

4. Stochastic gradient descent is similar with gradient descent, where we want to find the minimum of the function. Yet, with stochastic gradient descent, we don't calculate the average of the entire dataset, we just calculate the average of a random sample of the dataset. Along with that, we update the weights of the model after each sample. This allows to have a faster learning process when we are doing online learning. Since we working in batches, we can update the weights of the model after each batch.

Thank you for a great semester!