Homework - October 1, 2024

Student: Ayrton Chilibeck, achilibe@ualberta.ca
Lecturer: Lili Mou, UoA.F24.466566@gmail.com

Problem 1: Baye's Theorem

Prove Baye's Theorem:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{\sum_j P(X_j)P(Y|X_j)}$$

To prove Baye's Theorem, we recall the definition of conditional probablity:

$$\begin{aligned} P(X|Y) =& \frac{P(X \cap Y)}{P(Y)} \\ P(Y|X) =& \frac{P(X \cap Y)}{P(X)} \end{aligned}$$

Solving for $P(X \cap Y)$ in both equations, we can then substitute, and set the two systems equal to each other

$$P(X|Y)P(Y) = P(Y|X)P(Y) = P(X \cap Y)$$

Rearranging gives the default form of the equation

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

We can then use the multiplication rule for conditional probablilities to write

$$P(Y|X) = \frac{P(X|Y)P(Y)}{\sum_j P(X_j)P(Y|X_j)} \tag{1}$$

as required. □

Problem 2: Monty Hall Problem

Consider the Monty Hall game in a TV show. There are three closed doors, behind which are a car and two goats placed randomly.

1. You are asked to open a door by the host. Say, you would like to open Door 1. What is the probability of getting a car?

2. The host knows where the car is but he/she does not tell you. Instead, the host will open another door with a goat.

   (a) If the car is behind Door 2, the host can only open Door 3.

   (b) If the car is behind Door 1, the host can open either Door 2 or Door 3. He/she will do it with equal probability.

   Say, the host has opened Door 3. What is the probability of having the car behind

Door 1 now? What is the probability of having the car behind Door 2 now?

3. If my goal is to get the car, should I change my first choice (i.e., open Door 1 or Door 2)?

We solve each section in turn.

1 Part 1

The probability of the car being behind the door is $\frac{1}{3}$

2 Part 2

The assignment provides the following hints:

1. We can analyze this using Baye's Theorem:

$$P(A=1|B=1) = \frac{P(A=1)P(B=1|A=1)}{P(A=1)P(B=1|A=1)+P(A=2)P(B=1|A=2)+P(A=3)P(B=1|A=3)}$$

2. Numerator: $\frac{1}{6}$
3. Denominator: $\frac{3}{6}$

We can plug the numbers in and we get that $P(A=1|B=1) = \frac{2}{3}$

3 Part 3

We should switch doors. We can compute $P(A=2|B=1) = \frac{1}{3}$, and so it is more likely that the car is behind the other door.

Problem 3: Linearity of Expectation

Prove that the Expectation function is linear:

$$\mathbb{E}[\alpha f(X) + \beta g(X)] = \alpha\mathbb{E}[f(X)] + \beta\mathbb{E}[g(X)]$$

Recall the independence of observations for the expectation function, and the rule of constants:

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$
$$\mathbb{E}[\alpha] = \alpha$$

From these two rules, we can first seperate the two functions from each other:

$$\mathbb{E}[\alpha f(X) + \beta g(X)] = \mathbb{E}[\alpha f(X)] + \mathbb{E}[\beta g(X)]$$

Since we know that the expectation of a constant is that constant, we can extract the constants

$$\mathbb{E}[\alpha f(X) + \beta g(X)] = \alpha \mathbb{E}[f(X)] + \beta \mathbb{E}[g(X)]$$

as required. $\square$

Problem 4: Parameter Estimation

et X $U[a, b]$ be a continuous random variable uniformly distributed in the interval $[a, b]$, where a and b are unknown parameters.
We have a dataset $\{x^{(m)}\}_{m=1}^{M}$, where each data sample is iid drawn from the above distribution, and we would like to estimate the parameters a and b.
Give the likelihood of parameters.
Give the maximum likelihood estimation of parameters.

If X is a uniform distribution, then the probability density function of X is

$$f(x \in X|a, b) = \frac{1}{b-a} \qquad\qquad a \leq x \leq b$$

We can compute the likelihood of the data given the parameters a and b with the likelihood function $\mathcal{L}(a, b)$ which is the joint probability of observing all data in our set $x^{(m)}$. We can thus write

$$\begin{aligned}\mathcal{L}(a, b) &= \Pi_{m=1}^{M} f(x|a, b) \\ &\Pi_{m=1}^{M} \frac{1}{b-a} \\ &= \left(\frac{1}{b-a}\right)^n\end{aligned}$$

and we need to add a term to ensure that all data is within a and b since observing data outside of the bounds has probability 0

$$\mathcal{L}(a, b) = \left(\frac{1}{b-a}\right)^n \cdot \nVdash_{a \leq \min \vec{x}} \cdot \nVdash_{b \geq \max \vec{x}}$$

and thus, minimizing $b - a$ maximizes the likelihood function, so our most likely estimate for the parameters a and b is

$$\begin{aligned}a &= \min \vec{x} \\ b &= \max \vec{x}\end{aligned}$$

$\square$

Problem 5: MSE

Suppose M samples, each of which $x^{(m)}$ $\mathcal{N}(\mu, 1)$ is iid generated. Show that the estimate

$$\hat{\mu} = \frac{1}{M} \sum_{m=1}^{M} x^{(m)}$$

is an unbiased estimate of μ.

This is a duplicate of the proof we covered in class for the MSE being an unbiased estimate of the value μ.

First, the sample mean for the set is

$$\bar{\mu} = \frac{1}{M} \sum_{m=1}^{M} x^{(m)}$$

Now, we need to prove that the mean is a fair estimator, so let us compute the expectation of the gaussian normal distribution as given:

$$\begin{aligned}
\mathbb{E}[\mathcal{N}(\mu, \sigma^2)] &= \text{argmax}_\theta \Pi P(x|\mu, \sigma) \\
&= \text{argmax}_\theta \log\left[\text{argmax}_\theta \Pi P(x|\mu, \sigma)\right] \\
&= \text{argmax}_\mu \sum_{m=1}^{M} \log\left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{t^{(m)} - \mu^T x^{(m)}}{\sigma}\right)\right\}\right] \\
\mathbb{E}[\mathcal{N}(\mu, \sigma^2)] &= \frac{1}{M} \sum_{m=1}^{M} x^{(m)}
\end{aligned}$$

Which we can find by a slight change to the derivation done in class. $\square$