Homework 1 Solution

For the following problem, please use LaTex to type your solution. Please submit your solution in PDF. Handwritten solution will not be accepted. You may use the template to write down your solution: https://www.cs.ucsb.edu/~leili/course/dl23w/hw_template.tex.

Problem 1: Probability Basics (20')

Problem 1a: Counting and combinatorics (10')

If 2n kids are randomly divided into two equal subgroups, find the probability that the two tallest kids will be: (i) in the same subgroup; (ii) in different subgroups.

(Hint: Try putting the tallest and the second tallest into Group 1 and Group 2, then count the total number of different group assignments in each of the four possibilities. Hint 2: Check if your solution is correct for n = 2.)

Solution:

$$P(\text{same subgroups}) = \frac{n-1}{2n-1}$$

$$P(\text{different subgroups}) = \frac{n}{2n-1}$$

Problem 1b: Bayes Theorem (10')

In answering a question on a multiple choice test, a candidate either knows the answer with probability p ($0 \leq$ p $<$ 1) or does not know the answer with probability 1- p. If she knows the answer, she puts down the correct answer with a probability of 0.99, whereas if she guesses, the probability of her putting down the correct result is 1/k (k choices to the answer). Find the conditional probability that the candidate knew the answer to a question, given that she has made the correct answer.

Solution:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

=

$$\frac{p \times 0.99}{p \times 0.99 + \frac{1-p}{k}}$$

Where

- A: the event that the candidate knew the answer to the question

- B: the event that the candidate put down the correct answer
- k: the number of choices

Problem 2: Calculus Basics (10')

Suppose

$$f(x) = \frac{e^x logx}{x^2}$$

What is differential $f'(x)$?

Solution:

$$f(x) = \frac{e^x logx}{x^2}$$

$$\begin{aligned} f'(x) = \frac{df}{dx} = \frac{x^2 \cdot \left(\mathrm{e}^x \log(x) + \frac{\mathrm{e}^x}{x}\right) - 2x\mathrm{e}^x \log(x)}{x^4} \\ = \frac{\mathrm{e}^x \log(x)}{x^2} - \frac{2\mathrm{e}^x \log(x)}{x^3} + \frac{\mathrm{e}^x}{x^3} \\ = \frac{\mathrm{e}^x \cdot ((x-2)\log(x) + 1)}{x^3} \end{aligned}$$

Problem 3: Vector Basics (10')

Suppose

$$x = \begin{bmatrix} 0 \\ -1 \\ 2 \\ -2 \\ 4 \end{bmatrix}$$

What is the 1- norm and 2-norm of $|x|$?

Solution: 1 norm:

$$|x||_1 = |x_1| + |x_2| + |x_3| + ... + |x_n|$$

2 norm:

$$||x||_2 = \sqrt{x_1^2 + x_2^2 + x_3^2 + ... + x_n^2}$$

With the vector x given

1 norm:

$$||x||_1 = |x_1| + |x_2| + |x_3| + |x_4| + |x_5| = |0| + |-1| + |2| + |-2| + |4| = 9$$

2 norm:

$$||x||_2 = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2} = \sqrt{0^2 + (-1)^2 + 2^2 + (-2)^2 + 4^2} = \sqrt{29}$$

Problem 4: Vector Calculus (20')

Suppose x is a 3-d vector.

$$f(x) = |e^{A\cdot x+b} - c|_2^2$$

where

$$A = \begin{bmatrix} 2 & -2 & 3 \\ 3 & 2.5 & -1 \end{bmatrix}, b = \begin{bmatrix} -2 \\ -3 \end{bmatrix}, c = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}$$

$|\cdot|_2$ is 2-norm: $|x|_2 = \sqrt{x_1^2 + x_2^2 + \ldots}$.
What is the differential $\frac{\partial f}{\partial x}$?

Solution: Let $g(x) = e^{A\cdot x+b}$.

$$f(x) = f(g(x))$$

$$\begin{aligned}
\nabla f &= \frac{\partial g}{\partial x}\frac{\partial f}{\partial g} \\
&= A^\top \cdot \mathrm{diag}(e^{Ax+b}) \cdot 2(e^{Ax+b} - c) \\
&= 2A^\top \cdot \left(e^{Ax+b} \odot (e^{Ax+b} - c)\right) \\
&= 2A^\top \cdot (e^{2Ax+2b} - c \odot e^{Ax+b}) \\
&= \begin{bmatrix} 4 & 6 \\ -4 & 5 \\ 6 & -2 \end{bmatrix} \cdot \left(\exp\left(\begin{bmatrix} 4 & -4 & 6 \\ 6 & 5 & -2 \end{bmatrix} \cdot x + \begin{bmatrix} -4 \\ -6 \end{bmatrix}\right) - \exp\left(\begin{bmatrix} 2 & -2 & 3 \\ 3 & 2.5 & -1 \end{bmatrix} \cdot x + \begin{bmatrix} -2 \\ -3 \end{bmatrix}\right)\right)
\end{aligned}$$

Problem 5: MLE (20')

Consider the following distribution with parameter α

$$p(x|\alpha) = \begin{cases} \frac{\alpha e^\alpha}{x^{\alpha+1}} & x \in [e, +\infty] \\ 0 & \text{otherwise} \end{cases}$$

Suppose we are given a dataset $D = \{x_1, x_2, \dots, x_n\}$.

1. (20') Derive the log-likelihood function $\ell(\alpha; D)$

   Solution:

$$\begin{aligned}
\ell(\alpha; D) &= \frac{1}{n}\sum_{i=1}^{n} \log p(x_i|\alpha) \\
&= \frac{1}{n}\sum_{i=1}^{n} \log \frac{\alpha e^{\alpha}}{x^{\alpha+1}} \\
&= \frac{1}{n}\sum_{i=1}^{n} (\log \alpha + \alpha - (\alpha+1)\log x_i) \\
&= \alpha + \log \alpha - \frac{1}{n}(\alpha+1)\sum_{i=1}^{n} \log x_i \\
&= (1 - \frac{1}{n}\sum_{i=1}^{n} \log x_i)\alpha + \log \alpha - \frac{1}{n}\sum_{i=1}^{n} \log x_i
\end{aligned}$$

2. (20') Give the maximum likelihood estimate (MLE) of α.

   Solution:

$$\frac{d\ell}{d\alpha} = (1 - \frac{1}{n}\sum_{i=1}^{n} \log x_i) + \frac{1}{\alpha}$$

   Let $\frac{d\ell}{d\alpha} = 0$, we obtain

$$\hat{\alpha} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \log x_i - 1}$$