Roll No : 123456789 **Dept.**:
e.g., 190040001 e.g., CSE

IIT Bombay
CS 217/337: AIML
Quiz 1, 2023-24-II
Date: February 7, 2024

CS 217/337: Artificial Intelligence and Machine Learning

Total: **10 × 3 = 30 marks,** *Duration:* **1 hour, ATTEMPT ALL QUESTIONS**

Instructions:

1. This question-and-answersheet booklet contains a total of 4 sheets of paper (8 pages, page 2 is blank). Please verify.
2. Write your roll number and department on **every side of every sheet** (except the blank sheet) of this booklet. Use only **black/blue ball-point pen**. The first 5 minutes of additional time is given exclusively for this activity.
3. Write final answers neatly with a pen **only in the given boxes**.
4. Use the rough sheets for scratch works / attempts to solution. **Write only the final solution (which may be a sequence of logical arguments) in a precise and succinct manner in the boxes provided**. Do not provide unnecessarily elaborate steps. The space within the boxes are sufficient for the correct and precise answers.
5. Submit your answerscripts to the teaching staff when you leave the exam hall or the time runs out (whichever is earlier). **Your exam will not be graded if you fail to return the paper**.
6. **This is a closed book, notes, internet exam. No communication device, e.g., cellphones, iPad, etc., is allowed**. Keep these switched off in your bag and keep the bag away from you. If anyone is found in possession of such devices during the exam, that answerscript may be disqualified for evaluation and DADAC may be invoked.
7. One A4 assistance sheet (text **only on one side**) is allowed for the exam.

IIT Bombay
CS 217/337: AIML
Quiz 1, 2023-24-II
Date: February 7, 2024

Roll No : ______ e.g., 190040001 **Dept.**: ______ e.g., CSE

Problem 1 (10 points). A goods delivery company has two methods to deliver their goods:

1. **slow track**, that require *two* units of resources to be deployed per unit of the track, and
2. **fast track**, that require *three* units of resources to be deployed per unit of the track.

On a given day, the company has only *ten* resource credits to spend. Also, there are *four* units of slow track and *three* units of fast track available to the company on that day.

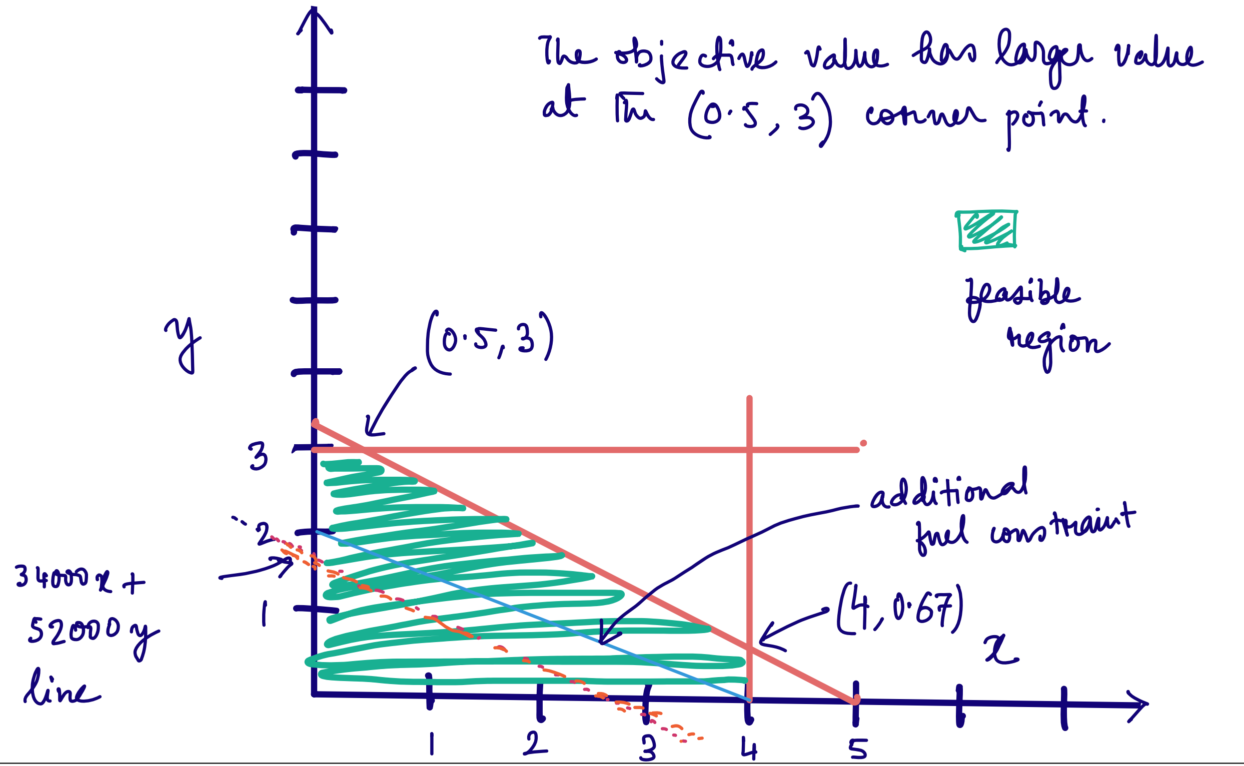
(a) Write down the constraints assuming x and y to be the number of units of slow and fast tracks respectively. **1 point.**

$2x + 3y \leq 10$

$x \leq 4$, $y \leq 3$

$x, y \geq 0$

(b) Show the feasible region on an x-y plane (make your plot roughly to scale). **2 points.**



IIT Bombay
CS 217/337: AIML
Quiz 1, 2023-24-II
Date: February 7, 2024

Roll No : e.g., 190040001 **Dept.**: e.g., CSE

(c) The company makes a profit of ₹34,000 per day from every unit of slow track and a profit of ₹52,000 per day from every unit of fast track chosen. Find the optimal number of units of each track needed to maximize the profit by writing down the linear program below. **2 points.**

maximize $34000x + 52000y$

s.t. $2x + 3y \leq 10$

$x \leq 4, \quad y \leq 3$

$x, y \geq 0$

(d) Solve the linear program (you can use graphical method by showing it on the plot of part (b)) to find:

The optimal number of units of (i) Slow tracks: 0.5 (ii) Fast tracks: 3

Maximum profit: ₹173,000 **1 + 1 + 1 points.**

(e) Suppose the company now also runs into a maximum fuel constraint of 500 liters of fuel on a given day. Each unit of slow track uses 125 liters and fast track uses 250 liters of fuel a day. How many units of each track should now be used to deliver the goods on this day to maximize profit under this fuel constraint?

Units of (i) Slow tracks: 4 (ii) Fast tracks: 0 **1 + 1 points.**

This additionally puts the constraint

$125x + 250y \leq 500$

$\Rightarrow \quad x + 2y \leq 4$

IIT Bombay
CS 217/337: AIML
Quiz 1, 2023-24-II
Date: February 7, 2024

Roll No : e.g., 190040001 | **Dept.**: e.g., CSE

Problem 2 (10 points). Consider a linear model of the form

$$g(\mathbf{x}_i, \mathbf{w}) = w_0 + \sum_{j=1}^{d} w_j x_i^{(j)}, \quad \mathbf{x}_i, \mathbf{w} \in \mathbb{R}^d, \forall i = 1, \ldots, n, \tag{1}$$

together with a sum-of-squares error function of the form

$$\texttt{Err}(\mathbf{w}, D) = \frac{1}{2} \sum_{i=1}^{n} \left(g(\mathbf{x}_i, \mathbf{w}) - y_i\right)^2, \tag{2}$$

correction as done in the exam

where the dataset is given by $D = \{(\mathbf{x}_i, y_i)_{i=1}^n\}$. Now suppose that a Gaussian noise ϵ_i with zero mean and variance σ^2 is added independently to each of the input variables $\mathbf{x}_i$. By making use of $\mathbb{E}[\epsilon_i] = 0$ and $\mathbb{E}[\epsilon_i \epsilon_j] = \sigma^2 \mathbb{I}\{i = j\}$, find the expression for the expected error averaged over the noise distribution. Note that when the noise is added to the input variables, the linear function g and consequently the error function both become random variables. This question asks about the expectation of that random error. Show every step of your derivation as guided below. **Briefly explain each step of your derivations**.

(a) Derive how the linear model after the addition of noise, say $\overline{g}$, is related to the linear model g in Equation (1). **3 points.**

$$\bar{g}(x_i, w) = w_0 + \sum_{j=1}^{d} w_j \left(x_i^{(j)} + \epsilon_i^{(j)}\right)$$

$$= w_0 + \sum_{j=1}^{d} w_j x_i^{(j)} + \sum_{j=1}^{d} w_j \epsilon_i^{(j)}$$

$$= g(x_i, w) + \sum_{j=1}^{d} w_j \epsilon_i^{(j)}$$

IIT Bombay
CS 217/337: AIML
Quiz 1, 2023-24-II
Date: February 7, 2024

Roll No : e.g., 190040001 | **Dept.**: e.g., CSE

(b) Derive the expression of the error after the addition of noise, say $\overline{\texttt{Err}}(\mathbf{w}, D)$. **4 points.**

$$\overline{Err}(w,D) = \frac{1}{2}\sum_{i=1}^{n}\left(\bar{g}(x_i,w) - y_i\right)^2$$

$$= \frac{1}{2}\sum_{i=1}^{n}\left[\bar{g}^2(x_i,w) - 2\bar{g}(x_i,w)\,y_i + y_i^2\right]$$

$$= \frac{1}{2}\sum_{i=1}^{n}\left[g_i^2 + 2g_i\sum_{j=1}^{d} w_j\epsilon_i^{(j)} + \left(\sum_{j=1}^{d} w_j\epsilon_i^{(j)}\right)^2 - 2g_iy_i - 2y_i\sum_{j=1}^{d} w_j\epsilon_i^{(j)} + y_i^2\right]$$

$$= \underbrace{\frac{1}{2}\sum_{i=1}^{n}(g_i - y_i)^2}_{Err(w,D)} + \sum_i g_i\sum_{j=1}^{d} w_j\epsilon_i^{(j)} - \sum_i y_i\sum_{j=1}^{d} w_j\epsilon_i^{(j)} + \frac{n}{2}\left(\sum_{j=1}^{d} w_j\epsilon_i^{(j)}\right)^2$$

(c) Show how the expected error $\mathbb{E}[\overline{\texttt{Err}}(\mathbf{w}, D)]$ is related to the original error $\texttt{Err}(\mathbf{w}, D)$ in Equation (2). **3 points.**

$$\mathbb{E}\left[\overline{Err}(w,D)\right] = Err(w,D) + \sum_i\sum_j g_iw_j\underbrace{\mathbb{E}\,\epsilon_i^{(j)}}_{=0} - \sum_i\sum_j y_iw_j\underbrace{\mathbb{E}\,\epsilon_i^{(j)}}_{=0} + \frac{n}{2}\mathbb{E}\left(\sum_{j=1}^{d} w_j\epsilon_i^{(j)}\right)^2$$

$$= Err(w,D) + \frac{n}{2}\sum_{j=1}^{d} w_j^2\sigma^2$$

[all cross terms $=0$ since $\mathbb{E}[\epsilon_i^{(j)}\epsilon_i^{(j')}] = 0$]

IIT Bombay
CS 217/337: AIML
Quiz 1, 2023-24-II
Date: February 7, 2024

Roll No : e.g., 190040001 **Dept.**: e.g., CSE

Problem 3 (10 points). In a study of the relationship between the daily temperature of a place (T) and monthly cost on electric bill (C) of a company, the following data was gathered.

T (in Fahrenheit)	20	30	50	60	80	90
C (in Thousand ₹)	125	110	95	90	110	130

Assume a probabilistic linear regression model for the relationship of cost and temperature as discussed in class:

$$C = \alpha + \beta T + \epsilon,$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is observational noise (σ is known). Find the *maximum likelihood estimates* (MLE) of the parameters α and β as guided by the steps below.

(a) First, derive the expression of the objective function that you need to *minimize* for the MLE. Show every step of the derivation. **4 points.**

$$c_i \sim \mathcal{N}(\alpha + \beta T_i, \sigma^2)$$

Likelihood : $P(c_i \mid T_i, \alpha, \beta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(c_i - \alpha - \beta T_i)^2}{2\sigma^2}\right\}$

Negative log likelihood

$$NLL(w) = \frac{1}{2\sigma^2} \sum_{i=1}^{n} (C_i - \alpha - \beta T_i)^2 - \text{const.}$$

$$\text{objective} = \underset{\alpha, \beta}{\operatorname{argmin}} \sum_{i=1}^{n} (c_i - \alpha - \beta T_i)^2 = E \text{ (say)}$$

Roll No : e.g., 190040001 **Dept.**: e.g., CSE

(b) Solve this optimization problem to find the system of equations you need to solve to find the MLE of α and β. **4 points.**

$$\frac{\partial E}{\partial \alpha} = 0 \Rightarrow n\alpha + \left(\sum_{i=1}^{n} T_i\right)\beta = \sum C_i$$

$$\Rightarrow 6\alpha + 330\beta = 660 \quad \text{---} \; 1$$

$$\frac{\partial E}{\partial \beta} = 0 \Rightarrow \left(\sum T_i\right)\alpha + \left(\sum T_i^2\right)\beta = \sum C_i T_i$$

$$\Rightarrow 33\alpha + 2190\beta = 3645 \quad \text{---} \; 2$$

Solving the above two equations, we get

(c) The MLE values are (rounded to two decimal places) **1 + 1 points.**

$\alpha =$ 107·8 and $\beta =$ 0·04

END OF QUESTION PAPER. GOOD LUCK!