| Roll. No: | Name: |
|---|---|

Total Marks: 14, Total Time: 35mins

# Instructions

1. The quiz is divided into two sections: short answer questions, and problems that require a detailed solution. For the first section, provide the final answer ONLY. For the second section, provide detailed answers showing all the necessary steps.
2. Use rough sheets for any calculations *if necessary,* and do not submit the rough sheets. Do not use a pencil for writing the answers.
3. Assume standard data whenever you feel that the given data is insufficient. However, do quote your assumptions explicitly.

# I Short answer questions (Answer any four)

*Note:*
*1) If more than four questions are answered, then the first four answers will be considered for evaluation.*
*2) 2 marks for the correct answer.*

1. The convex hull of a set $C$, denoted Conv$(C)$ is defined as

$$\text{Conv}(C) = \alpha_1 x_1 + \ldots + \alpha_k x_k \mid |x_i \in C, \alpha_i \geq 0, \forall i, \alpha_1 + \ldots + \alpha_k = 1.$$

For each of the following sets in $\mathbb{R}^2$, provide a visual depiction of their convex hulls by sketching:

$\boxed{1}$ (a) $C = \{(0,1), (0,4), (-2,-1), (0,0), (3,-2), (-1,2)\}$.

$\boxed{1}$ (b) Union of two unit circles, centered at $(1,1)$ and $(-1,-1)$, respectively.

$\boxed{2}$ 2. Let $\{X_n\}$ be a martingale sequence.
Consider the following two statements:
**I**: For all $n \geq 1$, $\mathbb{E}[X_n] = E[X_1]$.
**II**: For all $n \geq 1$, Variance$(X_n) =$ Variance$(X_1)$.
**III**: For all $n \geq 1$, Variance$(X_n) \geq$ Variance$(X_1)$.
**IV**: For all $n \geq 1$, Variance$(X_n) \leq$ Variance$(X_1)$.
Which of the statements is/are true?

$\boxed{2}$ 3. Exhibit a real-valued function $f$ that is $L$-smooth but not $L$-Lipschitz.

$\boxed{2}$ 4. Let $y^+ = f(x + \delta\Delta) + \xi^+$, $y^- = f(x - \delta\Delta) + \xi^-$ and $y = f(x)$, where $\Delta = (\Delta_1, \ldots, \Delta_d)^\intercal$ is a $d$-vector of independent Rademacher random variables. Consider the following gradient estimators: For $i = 1, \ldots, d$, the $i$ co-ordinate of the gradient estimators $g_1^{;} g_2^2, g_3^i$ are defined by

$$g_1^i = \left[\frac{y^+ - y^-}{2\delta\Delta_i}\right], \; g_2^i = \left[\frac{y^+ - y}{\delta\Delta_i}\right], \; g_3^i = \left[\frac{y^+}{\delta\Delta_i}\right]. \tag{1}$$

Assume $f$ is three times continuously differentiable and $\xi^{\pm}, \xi$ are zero mean. Order these estimators using "bias" as the metric.

$\boxed{2}$ 5. Show via an example that the second-order sufficient condition for optimality, i.e., the Hessian $\nabla^2 f(x^*) \succ 0$, is not necessary. In other words, exhibit a function with a minimum, where the aforementioned condition is not satisfied.

## II. Problems that require a detailed solution (Answer any one)

*Note: If more than one question is answered, then the first answer will be considered for evaluation.*

1. Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ is a convex function with a $L$-Lipschitz gradient and a minimizer $x^*$ with function value $f^* = f(x^*)$.

$\boxed{3}$ (a) Show that for any $x \in \mathbb{R}^d$, we have

$$f(x) - f^* \geq \frac{1}{2L} \left\| \nabla f(x) \right\|^2 .$$

$\boxed{3}$ (b) Prove the following property: For any $x, y \in \mathbb{R}^d$, we have

$$\left( \nabla f(x) - \nabla f(y) \right)^{\mathsf{T}} (x - y) \geq \frac{1}{L} \left\| \nabla f(x) - \nabla f(y) \right\|^2 .$$

$\boxed{4}$ 2. (a) Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x_1, x_2) = \frac{x_1^2}{2} + x_1 x_2 + 2x_2^2 - 4x_1 - 4x_2 - x_2^3.$$

Find points where the first-order necessary condition for optima is satisfied. For each of these points, characterize whether they are minimizers or maximizers or neither.

$\boxed{2}$ (b) Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = x^{\mathsf{T}} \begin{bmatrix} 2 & 5 \\ -1 & 1 \end{bmatrix} x + \begin{bmatrix} 3 \\ 4 \end{bmatrix}^{\mathsf{T}} x + 7.$$

Does this function have a minimizer? Justify your answer.