Roll. No:	Name:

Total Marks: 25, Total Time: 120mins

Instructions

1. The exam 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

Note: Max marks is nine. If you score above this limit, your marks will be truncated to nine.

1 1. **True/False:** If $f_1 : \mathbb{R}^d \to \mathbb{R}$ is L_1-smooth and $f_2 : \mathbb{R}^d \to \mathbb{R}$ is L_2-smooth, then $f_1 + f_2$ is $(L_1 + L_2)$-smooth.

1.5 2. Let $\{X_i\}$ be a i.i.d. sequence of positive random variables with mean one. Define $Y_n = \prod_{i=1}^n X_i$.

Consider the following two statements:
I: $\{Y_n\}$ is a martingale sequence.
II: $\{\sqrt{Y_n}\}$ is a super-martingale sequence.
II: $\{\sqrt{Y_n}\}$ is a sub-martingale sequence.

Which of the above statements is/are true?

3. Let $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ and let α_1, α_2 be two positive scalars. Answer with a "true" or "false" for each of these statements:

.5 (a) If f_1, f_2 are convex, then $\max(\alpha_1 f_1, \alpha_2 f_2)$ is convex.

.5 (b) If f_1, f_2 are concave, then $\max(\alpha_1 f_1, \alpha_2 f_2)$ is concave.

4. Suppose we want to minimize the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x_1, x_2) = (x_1 - x_2)^4 + x_1^2 - x_2^2 - 2x_1 + 2x_2 + 1.$$

1 (a) Find points where the first-order necessary condition for a minimum is satisfied.

1 (b) For each of these points, characterize whether the second-order necessary condition is satisfied.

1 5. Let $f : \mathbb{R}^d \to \mathbb{R}$ be differentiable. Then, if $\alpha, \beta \in \mathbb{R}$,

$$\int_0^\alpha \langle \nabla f(x + t(y - x)), \beta(y - x) \rangle \, dt = ?$$

(A) $f(y) - f(x)$ (B) $\beta(f(y) - f(x))$ (C) $\frac{\beta}{\alpha}\left(f(\alpha y + (1 - \alpha)x) - f(x)\right)$ (D) $\beta\left(f(\alpha y + (1 - \alpha)x) - f(x)\right)$

$\boxed{1}$ 6. **True/False:** If $\{X_n\}$ and $\{Y_n\}$ are two martingales, then their product $\{X_nY_n\}$ is also a martingale.

$\boxed{1.5}$ 7. Write down the Gaussian smoothing-based gradient estimator using two function measurements, and its bias/variance bounds when the objective f is L-smooth and convex. Specify the setting with necessary conditions, when the variance bound is better.

$\boxed{2}$ 8. Consider the following function: For some scalars a, b, c,

$$f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2.$$

Suppose you have a function measurement oracle that output $\hat{f}(x) = f(x) + \epsilon$, where ϵ is a standard normal random variable. Using the simultaneous perturbation method, devise a gradient estimator using two function measurements from the above oracle. Specify bounds on the bias and variance of this estimator.
Note: Only the optimal (lowest) bias/variance bounds would fetch full marks.

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

Note: If more than four questions are answered, then the first four answers will be considered for evaluation.

1. Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables (r.v.s). Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $\mathbb{E}[|f(X_n)|] < \infty, \forall n$. Let $a_n = \mathbb{E}(f(X_n))] \neq 0, \forall n$. Define

$$S_n = \frac{\prod_{m=1}^{n} f(X_m)}{\prod_{m=1}^{n} a_m}, \forall n \geq 1.$$

$\boxed{2}$ (a) Is $\mathbb{E}[|S_n|] < \infty,\ \forall n$?

$\boxed{2}$ (b) Is $\{S_n\}$ a martingale sequence?

2. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2.$$

Answer the following:

$\boxed{2}$ (a) Prove or disprove: f is strongly convex if $a > 0$ and $c > 0$.

$\boxed{1}$ (b) Derive a necessary and sufficient condition for strong convexity of f. This condition should be in terms of a, b, c.

$\boxed{1}$ (c) Under the condition from the part above, characterize the minimizer, say x^* of f.

3. Let $h(x^{(1)}, x^{(2)}) = \begin{bmatrix} 2x^{(1)} + 2x^{(2)} + 5 \\ 2x^{(1)} + 3x^{(2)} + 7 \end{bmatrix}$.

Answer the following:

$\boxed{1}$ (a) Find x^* such that $h(x^*) = 0$.

$\boxed{1}$ (b) Consider a root-finding algorithm with the following update iteration:

$$x_{n+1} = x_n - a(n)h(x_n). \tag{1}$$

Specify a value for $a(n)$ that ensures $x_n \to x^*$ as $n \to \infty$. Justify your answer.

$\boxed{2}$ (c) Suppose h is not directly observable. Instead, for any x, we have a noisy observation $\hat{h}(x)$ that satisfies

$$\mathbb{E}\left[\hat{h}(x) \mid x\right] = h(x) \text{ and } \mathbb{E}\left[\left\|\hat{h}(x)\right\|^2\right] \leq \sigma^2.$$

Specify a stochastic approximation variant of (1) and establish asymptotic convergence of the stochastic approximation iterate to x^*.

$\boxed{4}$ 4. Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function. For any three points x_1, x_2, x_3 such that $x_1 < x_2 < x_3$, show that

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

5. Let $x \in \mathbb{R}^d$, with x_i denoting the ith coordinate. Are the functions defined below convex? Justify your answer.

$\boxed{2}$ (a) $f(x) = \log\left(\exp(x_1) + \ldots + \exp(x_d)\right)$.

$\boxed{2}$ (b) $f(x) = \exp(x^\top Ax)$, where A is a positive semi-definite matrix.

$\boxed{2}$ 6. (a) Let $f : \mathbb{R} \to \mathbb{R}$. Recall that $f'(x^*) = 0$ and $f''(x^*) \geq 0$ are the first and second-order necessary conditions for a local minimizer. In a similar spirit, derive a third-order necessary condition, assuming f is three-times continuously differentiable.

$\boxed{2}$ (b) Show an example function f and a point x^* that satisfies the first, second and third-order necessary conditions, but x^* is not a local minimizer of the function f.