Roll. No:	Name:

Total Marks: 13, Total Time: 45mins

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 for the first three questions and a short justification for the last question. 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

$\boxed{1}$ 1. True or False: Consider the function $f(x) = x + \frac{1}{x}$, for any $x \geq 1$. Then, there exists a $0 < \alpha < 1$ such that

$$|f(x) - f(y)| \leq \alpha |x - y|, \forall x, y \geq 1.$$

$\boxed{2}$ 2. Provide a bound using Oh-notation for the number of iterations n of a stochastic gradient algorithm so that $f(x_n) - f(x^*) \leq \epsilon$ (with obvious notation) for an objective function that is (i) convex and smooth; and (ii) strongly-convex and smooth. Consider the case of unbiased and biased gradient information, with the latter having the usual $c_1\delta^2$ bias and c_2/δ^2 variance bounds. Ignore the constants and specify the dependence on ϵ alone.

Answer:

Function type	Unbiased case	Biased case
Convex and smooth		
Strongly-convex and smooth		

$\boxed{1}$ 3. In a finite horizon MDP setting, suppose we have time-invariant state and action spaces. Consider a modified problem, where the terminal cost $g_N(x)$ is replaced by $g'_N(x) = g_N(x) + 10$. Let J_k and J'_k denote the kth stage functions in the DP algorithm for the original and modified problems. Then, $J_k(x) \leq J'_k(x)$, for all x and k.

4. Consider a function $H : \mathbb{R}^n \to \mathbb{R}^n$ that satisfies

$$\|Hr - r^*\| \leq \alpha \leq \|r - r^*\| \,,$$

for some $r* \in \mathbb{R}^n$ and $\alpha \in (0, 1)$. Here $\|\cdot\|$ is the ℓ_2 norm.

State whether the following statements are true or false:

$\boxed{1}$ (a) r^* is the unique fixed point of H.

$\boxed{1}$ (b) Value iteration $r_{k+1} = Hr_k$ would converge to r^*.

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. An exam for the highly popular EE736 course is open book with two questions of M marks each. For each question, a student taking this exam can either (i) think and solve the problem; or (ii) search the internet for the solution. If he/she chooses to think, then he/she has a probability p_1 of finding the right solution approach, and therefore obtaining $M/3$ step marks. In the case when the right approach is found, the student has a probability p_2 to finish writing an answer that secures full M marks for that question. The corresponding probabilities for option (ii) are q_1 and q_2, and the step marks being $M/2$. The passing marks for the exam is 40% of the total $2M$ marks. Assume $0 < q_1 < p_1 < 1$ and $p_1 p_2 = q_1 q_2$, and answer the following:

$\boxed{1}$ (a) List all the open-loop policies for this problem.

$\boxed{3}$ (b) Calculate the probability of passing for each of these policies.

$\boxed{1.5}$ (c) Which is the best open loop policy?

$\boxed{1.5}$ (d) Is there a closed-loop policy (from a finite horizon MDP formulation) that does better than any of the open-loop policies? Provide an intuitive justification.

2. Let A be a $m \times d$ matrix with rows $a_1^\intercal, \ldots, a_m^\intercal$. For each $i = 1, \ldots, m$, a_i is a $d \times 1$-vector, representing the feature vector associated with example i in a ML training problem. Let $\mathbf{y}_1, \ldots, \mathbf{y}_p$ be $m \times 1$ vectors.

   Assume A has full column rank.

   Consider the following problem:

$$\min_x \alpha_1 \|Ax - \mathbf{y}_1\|^2 + \alpha_2 \|Ax - \mathbf{y}_2\|^2 + \ldots + \alpha_p \|Ax - \mathbf{y}_p\|^2, \tag{1}$$

   for some positive scalars $\alpha_1, \ldots, \alpha_p$.

   Answer the following:

$\boxed{2}$ (a) Let x_i^* denote the solution to $\min_x \|Ax - \mathbf{y}_1\|^2$. Express the solution x^* of (1) in terms of x_i^* and α_i, $i = 1, \ldots, p$.

$\boxed{2}$ (b) Write a SGD type algorithm for solving (1).

$\boxed{3}$ (c) Provide a finite time bound for the SGD iterate from the part above. This bound should quantify the rate of convergence of the SGD algorithm to x^*.

Note: A detailed proof is not necessary. You can use a non-asymptotic bound from the course notes. State the result precisely, with step-size choice and verify the assumptions for a theorem from the course textbook are met. Alternately, you can provide a proof sketch.