CME 307/MS&E 311/OIT 676
Qualifying Exam

June 18th, 2025

Name: ________________________________

SUNetID: ________________________________

(Your email address is yourSUNetID@stanford.edu.)

Instructions

This is a 90-minute exam that consists of two questions. Please answer them to the best of your ability.

Question 1. *Quadratic models and smooth convex optimization.*
Consider the optimization problem

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad f(x), \tag{1}$$

where $f : \mathbf{R}^n \to \mathbf{R}$ is strongly convex, L-smooth, and has Lipschitz continuous Hessian.

(a) Consider an algorithm that iteratively solves (1) by minimizing the quadratic model

$$x^{k+1} = \underset{x \in \mathbf{R}^n}{\text{argmin}} \left\{ f(x_k) + \langle \nabla f(x_k), x - x_k \rangle + \frac{1}{2\eta} \|x - x_k\|_{A_k}^2 \right\}, \tag{2}$$

given a sequence of symmetric positive-definite matrices $A_k \in \mathbf{S}^n_{++}(\mathbf{R})$ and a scalar $\eta > 0$. The term $\| \cdot \|_{A_k}$ in (2) is the A_k-norm, which is defined as $\|x\|_{A_k} = \sqrt{x^T A_k x}$.

Solve the optimization problem given in (2) to explicitly write x_{k+1} as a function of x_k.

(b) In the special case where $A_k = \nabla^2 f(x_k)$ at each iteration k, what is the name of this algorithm?

(c) How would you recommend selecting the stepsize in practice for the algorithm in (b) to ensure descent? Provide a 2-3 sentence well-reasoned argument for your approach to selecting η; note your argument does not need to be a mathematical proof.

(d) Please state the *strongest* local convergence rate of the algorithm in (b) when applied to optimize f. The possible options are sublinear, linear, or quadratic. Illustrate this convergence rate on a semi-log plot, where the vertical axis represents the suboptimality and the horizontal axis the iteration counter.

Question 2. *Bounding optimal values using duality.*

(a) Let $c \in \mathbf{R}^n$, $a_i \in \mathbf{R}^n$, $b_i \in \mathbf{R}$ (for $i \in \{1, 2, \dots, m\}$), and $\mu > 0$. Suppose we have the optimization problem

$$\text{minimize} \quad c^T x + \tfrac{1}{\mu} \textstyle\sum_{i=1}^m \log(1 + \exp(\mu(a_i^T x - b_i))), \tag{3}$$

where $x \in \mathbf{R}^n$ is the optimization variable. Show that (3) is a convex optimization problem.

(b) Derive the Lagrange dual of the equivalent problem

$$\begin{array}{ll} \text{minimize} & c^T x + \frac{1}{\mu} \sum_{i=1}^m \log(1 + \exp(\mu y_i)) \\ \text{subject to} & Ax - b \leq y, \end{array}$$

where the optimization variables are $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$, and A is the $m \times n$ matrix whose ith row is a_i^T.

(c) Suppose the pair of primal and dual linear programs

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array} \qquad \begin{array}{ll} \text{maximize} & -b^T \lambda \\ \text{subject to} & A^T \lambda + c = 0 \\ & \lambda \geq 0 \end{array}$$

has a finite optimal value $p^\star = d^\star$ and a dual optimal solution $\lambda^\star$ that satisfies $\lambda^\star \leq 1$. Let $q^\star$ be the optimal value of (3). Show that

$$p^\star \leq q^\star \leq p^\star + \frac{m \log 2}{\mu}.$$

You may use the following fact without proof: The function $f(z) = -(1 - z)\log(1 - z) - z \log z$ has range $[0, \log 2]$ over the interval $[0, 1]$.