MATH 450 - 558: Smooth and Nonsmooth Optimization

Instructor: Emre Mengi

Fall Semester 2015
Final Exam
Tuesday January 5th, 2016
Duration: 180 minutes

#1	20	
#2	20	
#3	20	
#4	20	
#5	20	
Σ	100	

NAME ____________________

STUDENT ID ____________________

SIGNATURE ____________________

- Please put your name, student ID and signature in the space provided above.
- You may use the lecture notes and/or your own notes, but you should not be using any textbooks.

Question 1 (20 points) Express the dual problem of

$$\text{minimize}_{x \in \mathbb{R}^n} \; c^T x \quad \text{subject to} \;\; f(x) \geq 0$$

with $c \neq 0$, in terms of the Fenchel conjugate f^*.

Question 2 Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz continuous function with Lipschitz constant γ, that is

$$|f(x) - f(y)| \quad \leq \quad \gamma \|x - y\|_2$$

for all $x, y \in \mathbb{R}^n$.

(a) (10 points) Prove the following regarding the generalized directional derivative $f^{(0)}$ of f:

$$|f^{\circ}(x; p) - f^{\circ}(y; q)| \leq \gamma \left(\|x - y\|_2 + \|p - q\|_2\right) \quad \forall x, y \in \mathbb{R}^n, \quad \forall p, q \in \mathbb{R}^n.$$

(b) (10 points) Prove the following regarding the generalized gradient $\partial f(x)$ of f (at every x):

$$\|\Psi\|_2 \leq \gamma \quad \forall \Psi \in \partial f(x).$$

Question 3 Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex Lipschitz continuous (but not necessarily differentiable) function. The penalized bundle method to find the global minimizer of f keeps track of two sequences $\{x^{(k)}\}$ and $\{y^{(k)}\}$ in $\mathbb{R}^n$. It requires the solution of the nonsmooth optimization problem

$$\text{minimize}_{x\in\mathbb{R}^n} \;\; P_k(x) := \phi_k(x) + \frac{\mu}{2}\|x - x^{(k)}\|_2^2 \tag{1}$$

where $\mu > 0$ is a penalty parameter, and

$$\phi_k(x) := \max\left\{ f(y^{(j)}) + s_j^T(x - y^{(j)}) \;\mid\; j = 0, \ldots, k \right\}$$

for some $s_j \in \partial f(y^{(j)})$ for $j = 0, \ldots, k$, repeatedly. The point $y^{(k+1)}$ is defined as the global minimizer of $P_k(x)$. Furthermore, $x^{(k+1)} := y^{(k+1)}$ if $f(y^{(k+1)})$ satisfies a sufficient decrease condition compared to $f(x^{(k)})$, otherwise $x^{(k+1)} := x^{(k)}$.

(a) (10 points) Express (1) as a convex optimization problem with a quadratic objective function subject to linear constraints.

(b) (10 points) Write down the centrality conditions (that is the KKT conditions but with the complementarity condition replaced by a centering equation in the primal-dual space) for the convex optimization problem in part **(a)**.

(Bonus) (5 points) Write down one iteration of Newton's method for the solution of the centrality conditions in part **(b)**.

Question 4 Let A be a matrix-valued function defined by

$$A(x) := A_0 + x_1 A_1 + \cdots + x_d A_d$$

where $A_0, A_1, \ldots, A_d$ are given $n \times n$ symmetric positive semidefinite matrices, and let $\lambda_{\max} : \mathbb{R}^d \to \mathbb{R}, \;\; \lambda_{\max}(x) := \lambda_{\max}(A(x))$. Furthermore, assume $A_0 \neq 0$.

(a) (10 points) Express the unconstrained eigenvalue optimization problem

$$\text{minimize}_{x \in \mathbb{R}^d} \; \lambda_{\max}(x) \tag{2}$$

as a constrained optimization problem with a linear objective subject to a positive semidefiniteness constraint.

(b) (10 points) Derive a semidefinite program that yields an upper bound for (2). State also conditions that guarantee that this semidefinite program is equivalent to (2).
(Hint: Try to view the constrained optimization problem in part **(a)** as the dual of a semidefinite program.)

Question 5 (20 points) Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ with $m > n$ be given. Write down a necessary and sufficient condition for a point to be a global minimizer for the following problem using generalized gradients:

$$\text{minimize}_{x \in \mathbb{R}^n} \; \|b - Ax\|_\infty.$$

For simplicity, assume the following:

(i) $\text{rank}(A) = n$;

(ii) every set consisting of n rows of A is linearly independent;

(iii) letting $r(x) := b - Ax$, at every $x \in \mathbb{R}^n$ no more than n components of $r(x)$ are identical in absolute value.

Note that the function $f(x) := \|b - Ax\|_\infty$ is differentiable everywhere excluding a set Ω of (Lebesgue) measure zero; you can use this fact in your answer.