UNIVERSITY OF CAMBRIDGE

MAMA/339, NST3AS/339, MAAS/339

MAT3

MATHEMATICAL TRIPOS Part III

Wednesday 5 June 2024 1:30 pm to 3:30 pm

PAPER 339

TOPICS IN CONVEX OPTIMISATION

Before you begin please read these instructions carefully

Candidates have TWO HOURS to complete the written examination.

Attempt **BOTH** questions.
There are **TWO** questions in total.
The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury tag
Script paper
Rough paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1

(a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, $C \subset \mathbb{R}^n$ a closed convex set and consider the problem

$$f^* = \min_{x \in \mathbb{R}^n} \{f(x) : x \in C\}.$$

The projected subgradient method has iterates $x_{i+1} = P_C(x_i - t_i g_i)$ where g_i is a subgradient of f at x_i, P_C is the Euclidean projection on C and $t_i > 0$ is a step size. Assuming that f is G-Lipschitz with respect to the Euclidean norm, and that x^* is an optimal solution of the problem, show that after k iterations we have

$$\min\{f(x_0), \ldots, f(x_{k-1})\} - f^* \leqslant \frac{G\|x_0 - x^*\|_2}{\sqrt{k}}$$

for a suitable choice of constant step size $t_0 = \cdots = t_{k-1} = t$ (allowed to depend on k, G, x_0 and x^*). [15]

Consider a linear program with optimal value p^*,

$$p^* = \min_{x \in \mathbb{R}^n} \{\langle c, x\rangle : Ax \leqslant b\} \tag{1}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ where $m \geqslant n$. We assume that the problem is strictly feasible, and bounded (i.e., p^* is finite).

(b) Derive the dual optimization program. Verify that weak duality holds. Explain what is meant by strong duality. Does this linear program exhibit strong duality? [10]

(c) Show that there is $M > 0$ large enough such that the solution of the linear program (1) is the same as the solution of the *unconstrained* minimization problem:

$$\min_{x \in \mathbb{R}^n} \{\langle c, x\rangle + M \max\{0, (Ax - b)_1, \ldots, (Ax - b)_m\}\}. \tag{2}$$

[*Hint: you may want to consider the optimal solution of the dual program of* (1).] [15]

(d) Let $f(x)$ be the objective function in (2). Show that f is convex. Is f Lipschitz? If yes, give a bound on the Lipschitz constant. For any $x \in \mathbb{R}^n$ give an expression for a subgradient of f at x. [10]

2

(a) Consider the following convex minimization problem

$$\min_{x\in\mathbb{R}^n}\left\{\frac{1}{2}\|x-y\|_2^2 : 0 \leqslant x \leqslant 1 \text{ and } \langle a,x\rangle = b\right\} \qquad (1)$$

where $y \in \mathbb{R}^n, a \in \mathbb{R}^n, b \in \mathbb{R}$ and the inequalities $0 \leqslant x \leqslant 1$ are interpreted componentwise. We assume the problem is strictly feasible. Use strong duality to write down necessary and sufficient conditions for a point x^* to be a solution of (1). [*Hint: your optimality conditions (after simplification, if necessary) should only involve a single dual variable associated to the linear equality constraint.*]

Deduce that the solution x^* of (1) reduces to solving a one-dimensional nonlinear equation. [10]

(b) State the definition of the proximal operator of a convex lower semi-continuous function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Give necessary and sufficient conditions, in terms of the subdifferential of f, to have $\mathbf{prox}_f(y) = x$ for some $x \in \mathbb{R}^n$. Prove the generalized Moreau identity: $\mathbf{prox}_{tf}(y) = y - t\mathbf{prox}_{t^{-1}f^*}(y/t)$ where $t > 0$ and f^* is the Fenchel conjugate of f. [10]

(c) Let $C \subset \mathbb{R}^n$ be a compact convex set and define

$$\phi(x) = \max_{v\in C}\langle x,v\rangle\,. \qquad (2)$$

Show that for $t > 0$, $\mathbf{prox}_{t\phi}(y) = y - tP_C(y/t)$ where P_C is the Euclidean projection on C. [10]

(d) Given an integer $k \in \{1,\ldots,n-1\}$, and $x \in \mathbb{R}^n$, let $h(x)$ be the sum of the k largest components of x, i.e.,

$$h(x) = x_{[1]} + \cdots + x_{[k]} \qquad (3)$$

where $x_{[1]} \geqslant \ldots \geqslant x_{[n]}$ are the components of $x \in \mathbb{R}^n$ sorted in nonincreasing order. Show that h can be put in the form (2) where C is a convex set of the form

$$C = \{v \in \mathbb{R}^n : 0 \leqslant v \leqslant 1, \langle a,v\rangle = b\}$$

for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ to be specified. Use parts (a) and (c) to explain how the proximal operator of h can be evaluated. [*You do not need to explain how to solve one-dimensional nonlinear equations.*] [10]

(e) Consider a minimization problem of the form

$$\min_{x\in\mathbb{R}^n}\frac{1}{2}\|Ax-b\|_2^2 + \lambda h(x)$$

where $\lambda \geqslant 0$ and $h(x)$ is the function defined in (3). Write down a proximal gradient algorithm for the problem above, and explain how the step size should be chosen. Comment on the convergence rate. [10]

END OF PAPER