UNIVERSITY OF CAMBRIDGE

MAMA/339, NST3AS/339, MAAS/339

MAT3

MATHEMATICAL TRIPOS Part III

Friday, 9 June, 2023 9:00 am to 11:00 am

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 Let $a_1, \ldots, a_m \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and consider the function $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) = \max_{i=1,\ldots,m} (\langle a_i, x \rangle + b_i).$$

(a) Show that f is convex. State the definition of *subgradient*. For $x \in \mathbb{R}^n$ give a subgradient of f at x and justify your answer.

(b) Is the function f Lipschitz (with respect to the Euclidean norm)? Justify your answer, and if yes, give an upper bound on the Lipschitz constant. Write down the subgradient method for minimizing f. State, without proof, an upper bound of the form $O(\epsilon^{-p})$ (where $p > 0$) on the number of iterations needed to find a point x such that $f(x) - \min f \leqslant \epsilon$.

For $\beta > 0$, let

$$f_\beta(x) = \beta^{-1} \log \sum_{i=1}^{m} \exp(\beta(\langle a_i, x \rangle + b_i)).$$

(c) Show that f_β is convex, and that for all x,

$$f(x) \leqslant f_\beta(x) \leqslant f(x) + \beta^{-1} \log m.$$

(d) Give an expression for the gradient of f_β. Show that f_β is L-smooth with respect to the Euclidean norm for some constant $L > 0$ that you should specify.

(e) Explain how, using Nesterov's accelerated gradient, one can compute x such that $f(x) - \min f \leqslant \epsilon$ in at most $O(\epsilon^{-1})$ iterations.

2 A nonlinear map $F : \mathbb{R}^N \to \mathbb{R}^N$ is *monotone* with respect to an inner product $\langle \cdot, \cdot \rangle$ if $\langle F(v) - F(w), v - w \rangle \geqslant 0$ for all $v, w \in \mathbb{R}^N$.

(a) Show that if $f : \mathbb{R}^N \to \mathbb{R}$ is convex and differentiable, then $F(w) = \nabla f(w)$ is a monotone map.

(b) Let $M : \mathbb{R}^N \to \mathbb{R}^N$ be a linear map which is self-adjoint with respect to $\langle \cdot, \cdot \rangle$ and positive definite. Define the (nonlinear) map $T : \mathbb{R}^n \to \mathbb{R}^n$ for each $w \in \mathbb{R}^N$ via the implicit equation

$$M(T(w)) + F(T(w)) = M(w) \tag{1}$$

which we assume admits a unique solution $T(w)$ for all $w \in \mathbb{R}^N$, where $F : \mathbb{R}^N \to \mathbb{R}^N$ is a generic monotone map with respect to $\langle \cdot, \cdot \rangle$. In short, we write $T(w) = (M+F)^{-1}(M(w))$. Verify that the fixed points of T are zeros of F, i.e., $T(w) = w \implies F(w) = 0$. After recalling the definition of a *firmly nonexpansive map*, show that the map T is firmly nonexpansive with respect to the inner product $\langle \cdot, \cdot \rangle_M$ defined by $\langle v, w \rangle_M = \langle v, M(w) \rangle = \langle M(v), w \rangle$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex differentiable function, and consider the problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad Ax = b \tag{2}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

(c) Write down the Lagrangian and the dual optimization problem. Show that (x, z) is a pair of primal-dual optimal points if, and only if, $F\begin{pmatrix} x \\ z \end{pmatrix} = 0$, where $F : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ is the nonlinear map defined by

$$F\begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} \nabla f(x) - A^T z \\ Ax - b \end{pmatrix}.$$

Show that the map F is monotone with respect to the Euclidean inner product.

(d) Consider a symmetric matrix M of size $n + m$ of the form

$$M = \begin{bmatrix} \alpha I & A^T \\ A & \beta I \end{bmatrix}. \tag{3}$$

where α, β are scalars, and I represents an identity matrix of suitable size. Show that if $\alpha\beta > \|A\|_2^2$, then M is positive definite. [Here $\|A\|_2$ is the operator norm of A.]

(e) Let $T : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ be the nonlinear map defined as in (1), i.e., $T(w) = (M + F)^{-1}(M(w))$, with the linear map $M : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ given in (3). Consider the fixed point iterations for T, namely

$$w_{k+1} = T(w_k).$$

By writing $w_k = \begin{pmatrix} x_k \\ z_k \end{pmatrix} \in \mathbb{R}^{n+m}$, simplify these iterations, and show that they can be expressed solely in terms of the *proximal operator* of $\alpha^{-1} f$ and the linear map A.

END OF PAPER