UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Tuesday, 3 June, 2014 1:30 pm to 4:30 pm

PAPER 65

CONVEX OPTIMISATION WITH APPLICATIONS IN IMAGE PROCESSING

*Attempt no more than **FOUR** questions.*

*There are **FIVE** questions in total.*

The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury Tag
Script 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) Prove that every closed convex set $C \subseteq \mathbb{R}^n$ is the intersection of all closed half-spaces that contain C.

b) State the central theorem about the Legendre–Fenchel transform and explain the most important steps in the proof leading to the final relation between f^{**} and $\mathrm{cl\ con} f$. How is the theorem in a) used in the proof?

c) Consider the infimal convolution $f : \mathbb{R}^n \to \bar{\mathbb{R}}$,

$$f(x) := \inf_{(y,z),y+z=x} \{g(y) + h(z)\},$$

where $g(y) = \frac{1}{2}\|y\|_2^2$ and $h(z) = \|z\|_1$. This is also known as a Huber-type functional. Characterise the subdifferential ∂f using g^* and h^*, and derive an explicit representation of ∂f for this specific choice of g and h. Sketch f for the case $n = 1$. Now assume that $u \in \mathbb{R}^n$ is a dicretisation of a one-dimensional signal. Given what you know about L^2 and L^1-based regularisation of the gradient, what properties would you expect the regulariser $R(u) := f(\nabla u)$ to have? Where could it be useful?

2

a) Assume $f : \mathbb{R}^n \times \mathbb{R}^m \to \bar{\mathbb{R}}$ is a given perturbation function. Define the primal and dual objectives φ, ψ, primal and dual problems, and primal and dual marginal functions p, q. State a sufficient condition for strong duality. Explain how a sensitivity analysis can be performed in the perturbation framework and what it is useful for, and state a sufficient condition under which it is possible.

b) Consider the noise-*constrained* TV problem

$$\min_{u \in \mathbb{R}^n} \mathrm{TV}(u) \text{ s.t. } \|u - g\|_2^2 \leqslant \sigma, \tag{1}$$

where $\sigma \geqslant 0$, TV is a standard discretisation of the total variation – in particular real-valued, proper, lower semi-continuous, convex, and non-negative – and $g \in \mathbb{R}^n$ is the discretisation of a given image. Compared to the more common *penalised* version

$$\min_{u \in \mathbb{R}^n} \{\mathrm{TV}(u) + \lambda \|u - g\|_2^2\}, \tag{2}$$

the constrained formulation (1) has the advantage that finding a good estimate for the maximum noise level σ is often easier than determining the optimal parameter λ.

Define a perturbation function for problem (1) where the scalar σ is the perturbed variable. Derive the associated saddle-point formulation and show that the set of saddle points is non-empty if $\sigma > 0$.

c) Using the saddle-point formulation from b), show that for every $\sigma > 0$ and every solution u' of the constrained problem (1) there exists a $\lambda \geqslant 0$ so that u' is the solution of the penalised problem (2).

Conversely, show that for every $\lambda > 0$, we can find a $\sigma \geqslant 0$ so that the solution of the penalised problem (2) is a solution of the constrained problem (1).

3

a) State the definition of forward- and backward-steps $F_{\tau f}$ and $B_{\tau f}$ in terms of set-valued mappings and motivate their definition. Show that the backward step $B_{\tau f}$ is at most single-valued for every proper, lower semi-continuous, convex function f and $\tau > 0$. Carefully explain each step of the proof. In particular, point out where the assumptions on f are used.

b) Show that the backward step is always non-empty, i.e., it is always exactly single-valued.

c) Consider the two-dimensional TGV-regularised problem

$$\inf_{u \in \mathbb{R}^n} \sup_{v \in \mathbb{R}^{n \times 4}} \left\{ \frac{1}{2}\|u - g\|_2^2 + \langle u, \operatorname{Div}_2 v\rangle - \delta_C(v) - \delta_D(\operatorname{Div}_1 v) \right\},$$

where $C := \{v \in \mathbb{R}^{n\times 4} | \|v_{i,\cdot}\|_2 \leqslant \alpha_1, i = 1, \ldots, n\}$, $D := \{w \in \mathbb{R}^{n \times 2} | \|w_{i,\cdot}\|_2 \leqslant \alpha_2, i = 1, \ldots, n\}$, and $\operatorname{Div}_1 : \mathbb{R}^{n\times 4} \to \mathbb{R}^{n \times 2}, \operatorname{Div}_2 : \mathbb{R}^{n \times 4} \to \mathbb{R}^n$ are linear operators discretising the divergence and second-order divergence of v.

Choose a suitable first-order method from the lectures that allows to recover primal and dual iterates, and apply it to the problem in such a way that all sub-steps can be solved in closed form. You may have to introduce additional variables. State the complete iteration including the solutions for the sub-steps.

4

a) Consider the conic problem in standard form,

$$\inf_{x \in \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad Ax - b \geqslant_K 0, \tag{1}$$

where K is a proper, closed, convex, self-dual cone with associated canonical barrier function F and $A \in \mathbb{R}^{m \times n}$ has full column rank. State the dual problem. Define the primal-dual central path and state a joint characterisation of the points on the primal-dual central path. Derive the set of equations that define the Newton step for tracing the central path.

Why is this method called an interior-point method, and where in the iteration is this property of the iterates required? Explain a strategy to ensure that the iterates are always interior points. How can we find an interior starting point in practise?

b) Consider the one-dimensional denoising problem

$$\min_{u \in \mathbb{R}^n} \|u - g\|_2 + \lambda h(Gu),$$

where $\lambda > 0$, $G : \mathbb{R}^n \to \mathbb{R}^n$ discretises the gradient (derivative) operator, and the regulariser $h : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$h(v) = \sum_{i=1}^{n} \sigma(v_i), \quad \sigma(s) = \begin{cases} s^2, & s \geqslant 0, \\ 0, & s < 0, \end{cases}$$

where $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$.

Reformulate the problem in standard conic form (1), introducing additional variables as necessary. You do not need to define A, b and c explicitly, it is sufficient to list all constraints in a way that makes the linearity obvious.

To which class of conic problems does the reformulation belong? Find a suitable canonical barrier function for this problem.

5

Write an essay about Kernel Support Vector Machines. Make sure to state all necessary definitions and important equations, and in particular the primal and dual problems. The essay should also explain the role of duality and the kernel trick, and where and how the latter is used. Discuss the connection to Mercer kernels and show how they relate to nonlinear embeddings into Hilbert spaces.

END OF PAPER