Continuous Optimization Exam

04/06/2021
3 hours – documents allowed

The exam is made of 4 independent exercises, in roughly increasing difficulty. If necessary, you can admit the results of previous questions. When using the recalls, cite them.

Some useful recalls

i) If $X \sim \mathcal{N}(\mu, \Sigma)$ is a Gaussian vector, then, for any vector u, we have $u^\top X \sim \mathcal{N}(u^\top \mu, u^\top \Sigma u)$.

ii) The Fenchel transform of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is given by $f^\star(x) = \sup_{y \in \mathbb{R}^n} x^\top y - f(y)$.

iii) Assume that f is a convex proper lsc function. f is μ-strongly convex, iff $f^\star$ is differentiable with $\frac{1}{\mu}-$Lipschitz gradient.

iv) Assume that f is a convex proper lsc function. Then $\lambda \in \partial f(x)$ iff $x \in \partial f^\star(\lambda)$.

v) Assume that f is a convex proper lsc function. Then $\lambda \in \partial f(x)$ iff $x \in \arg\max_y \lambda^\top y - f(y)$

Exercice 1: Warm-up *0 points*

(a) (1 point) On what conditions on the set C is $\mathbb{I}_C$ a proper lower semicontinuous, convex function ?

(b) (2 points) Write the KKT conditions for the following problem.

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \ln\left(\sum_{i=1}^n e^{x_i}\right) \\ \text{s.t.} & \sum_{i=1}^n x_i = 0 \\ & \sum_{i=1}^n x_i^2 \leq 1 \end{aligned}$$

Are they necessary and/or sufficient conditions of optimality for this problem ?

(c) (2 points) We consider the following problem

$$(P) \quad \min_{x \in \mathbb{R}^n} \quad f(x) \qquad \text{s.t.} \quad Ax = b, \quad x \leq 0$$

with value v and the following penalized versions

$$(P_t^{in}) \quad \min_{x \in \mathbb{R}^n} \quad f(x) - t\sum_{i=1}^n \ln(-x_i) \quad \text{s.t.} \quad Ax = b, \quad x < 0 \qquad \text{and} \qquad (P_t^{out}) \quad \min_{x \in \mathbb{R}^n} \quad f(x) + t\sum_{i=1}^n (x_i)^+ \quad \text{s.t.} \quad Ax = b$$

with associated value v_t^{in} and v_t^{out}, and an optimal solution x_t^{in} and x_t^{out}.
Intuitively, assuming that f is "well behaved", for t going to which value does (P_t^{in}) tends to the original problem (P) ? In which sense ? What can you say about x_t^{in} ? Can you compare v_t^{in} and v ? Same questions for (P_t^{out})

Exercice 2: Support function (4 points) *0 points*

For any set $C \subset \mathbb{R}^n$, we define its support function

$$\sigma_C : x \mapsto \sup_{c \in C} c^\top x$$

(a) (2 points) Assume that C and D are closed convex sets. Using a separation theorem, show that $C = D$ if and only if their support functions are equal.

(b) (2 points) For any set C, recall that the indicator function $\mathbb{I}_C$ take value 0 on C and $+\infty$ outside. Show that, for any non empty set C, the Fenchel transform of its indicator function of set C is its support function, i.e. $\mathbb{I}_C^\star = \sigma_C$. Deduce a second proof for the previous question.

Exercice 3: A linear problem with Gaussian cost *0 points*

In the following we assume that $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^m$ are given matrices ; c is a gaussian random variable with mean $\bar{c} \in \mathbb{R}^n$ and variance $\Sigma \in \mathbb{R}^{n^2}$.

(a) (2 points) We consider the following optimization problem

$$(P_\gamma) \quad \min_{x \in \mathbb{R}^n} \quad \mathbb{E}\big[c^\top x\big] + \gamma Var(c^\top x)$$

Show that P_γ is a quadratic program. Comment on the complexity of solving P_γ. (Hint : answer should depend on the value of the parameter $\gamma \in \mathbb{R}$).

(b) (2 points) We now consider the following problem

$$\begin{aligned}(P'_\alpha) \quad \min_{x \in \mathbb{R}^n, z \in \mathbb{R}} \quad & z \\ \text{s.t.} \quad & Ax \leq b \\ & \mathbb{P}\Big[c^\top x \geq z\Big] \leq \alpha\end{aligned}$$

Show that, for $\alpha \in]0, 0.5]$, (P'_α) is equivalent to an SOCP, using $\phi(t) = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-u^2/2} du$ (which is 1 minus the cdf of a centered gaussian), or its inverse ϕ^{-1}.

What happen if $\alpha \in]0.5, 1]$?

Exercice 4: Prox operator and Moreau-regularization *0 points*

For any $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ proper convex function we define the *proximal operator*

$$\text{prox}_f : x \mapsto \arg\min_{y \in \mathbb{R}^n} f(y) + \frac{1}{2}\|y - x\|^2$$

and Moreau regularization of parameter $\mu > 0$

$$f_\mu : x \mapsto \inf_{y \in \mathbb{R}^n} f(y) + \frac{1}{2\mu}\|y - x\|^2$$

We want to study the *proximal point algorithm* given by the following sequence

$$x^{(k+1)} = \text{prox}_{\mu f}(x^{(k)}).$$

(a) (1 point) Show that prox_f and f_μ are well defined. For C closed convex non empty, and $f = \mathbb{I}_C$ recognize prox_f and f_μ.

(b) (1 point) Show that $x^\sharp$ is a minimizer of f if and only if it minimizes f_μ, if and only if $x^\sharp = \text{prox}_f(x^\sharp)$.

(c) (1 point) Show that $f_\mu(x) = \frac{1}{2\mu}\|x\|^2 - \frac{1}{\mu}(\mu f + \frac{1}{2}\|\cdot\|^2)^\star(x)$.

(d) (1 point) Show that $\mathrm{prox}_{\mu f}(x) = \arg\max_y x^\top y - \mu f(y) - \frac{1}{2}\|y\|^2$.

(e) (1 point) Show that $\nabla f_\mu(x) = \frac{1}{\mu}(x - \mathrm{prox}_{\mu f}(x))$.

(f) (1 point) Interpret the proximal point algorithm as a gradient algorithm.

(g) (2 points) Writing

$$\begin{aligned} f_\mu(x) = \min_{y,z} \quad & f(y) + \frac{1}{2\mu}\|z\|^2 \\ \text{s.t.} \quad & x - y = z \end{aligned}$$

and using duality show that $f_\mu(x) = (f^\star + \frac{\mu}{2}\|\cdot\|^2)^\star(x)$. Deduce that f_μ has $\frac{1}{\mu}$-Lipschitz gradient.

(h) (1 point) Show that, if f is a proper convex lowersemicontinuous function, admitting a minimizer, the proximal point alogrithm converges toward a minimizer of f.

(i) (1 point) For the following problem

$$\min_{x \in \mathbb{R}^n} \quad g(x) + h(x)$$

with $g \in \mathcal{C}^1$, we introduce the proximal gradient algorithm given as

$$x^{(k+1)} = \mathrm{prox}_{\mu h}\Big(x^{(k)} - \mu \nabla g(x^{(k)})\Big)$$

Recognize the proximal gradient algorithm for $h = 0$ first and then for $h = \mathbb{I}_C$, with g proper convex lowersemicontinuous, and $C \subset \mathrm{dom}(g)$ closed convex non-empty.