MGMT 690 - Pset 4

Spring 2024

Instructions:

- This pset is due on **Thursday, May 2** at 11:59pm.
- Completed psets should be submitted to Gradescope.
- **Exercises** are for your own review only. They do not need to be submitted and will not be graded.
- **Complete all problems 1–3.**

Exercises

- Recall that a function $f : \mathbb{R}^n \to [-\infty, \infty]$ is *closed* if the epigraph

$$\{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \; f(x) \le t\}$$

is a closed set.

  - Prove that any convex function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous.
  - Prove that any continuous function from $\mathbb{R}^n \to \mathbb{R}$ is closed.
  - Give an example of a convex function $f : \mathbb{R}^n \to [-\infty, \infty]$ that is not closed.

- Suppose $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ for each $i \in [k]$. Define $g : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k}$ by

$$g(x_1, \dots, x_k) = \sum_{i=1}^{k} f_i(x_i).$$

Show that $g^*(y_1, \dots, y_k) = \sum_{i=1}^{k} f_i^*(y_i)$.

Problems

1. [15pts] Let $f, g : \mathbb{R}^n \to \mathbb{R}$. The infimal convolution of f and g, denoted $f \Box g$, is a function $(f \Box g) : \mathbb{R}^n \to [-\infty, \infty]$ defined as

$$(f \Box g)(x) := \inf_{z \in \mathbb{R}^n} f(z) + g(x - z).$$

Prove that if $f, g : \mathbb{R}^n \to \mathbb{R}$, then

$$(f \Box g)^*(y) = f^*(y) + g^*(y) \qquad \forall y \in \mathbb{R}^n.$$

2. [40pts] This question derives a mirror descent setup on the simplex with the ℓ_1 norm.

   Let

$$\Delta_n := \left\{ x \in \mathbb{R}^n : \begin{array}{l} x \geq 0 \\ 1^\intercal x \leq 1 \end{array} \right\}.$$

   Define $\omega : \Delta_n \to \mathbb{R}$ by

$$\omega(x) = \sum_{i=1}^{n} x_i \log(x_i)$$

   with the convention that $0 \log(0) = 0$.

   (a) [10 pts] Verify that ω is closed and convex and differentiable on $\operatorname{dom}(\partial\omega)$

   (b) [15pts] Prove that $\omega(x)$ is 1-strongly convex on Δ_n with respect to the ℓ_1 norm.

   Hint: Recall, this is equivalent to saying that $x \mapsto \omega(x) - \frac{1}{2}\|x\|_1^2$ is convex on the set Δ_n. You may use without proving that this function is convex if and only if the Hessian is positive semidefinite on the interior of Δ_n.

   (c) [15pts] The basic step in mirror descent with this setup is: given $x_t \in \operatorname{dom}(\partial\omega)$, $g_t \in \partial f(x_t)$ and $\eta_t > 0$, set

$$x_{t+1} = \operatorname*{arg\,min}_{x \in \Delta_n} \left\{ \langle \eta_t g_t - \nabla\omega(x_t), x \rangle + \omega(x) \right\}.$$

   Give a closed form solution to this problem.

3. [45pts] This problem improves the Frank–Wolfe convergence rate by assuming that the domain is strongly convex and the objective is strongly convex.

   Fix an arbitrary norm on $\mathbb{R}^n$. We say that a set $\Omega \subseteq \mathbb{R}^n$ is μ-strongly convex if for all $x, y \in \Omega$, $\gamma \in [0, 1]$

$$\mathbb{B}((1-\gamma)x + \gamma y, \gamma(1-\gamma)\frac{\mu}{2}\|x - y\|^2) \subseteq \Omega.$$

   Here, $\mathbb{B}(x_0, r) = \{x \in \mathbb{R}^n : \|x_0 - x\| \leq r\}$.

   Let $f : \mathbb{R}^n \to \mathbb{R}$ be an L_f-smooth μ_f-strongly convex function w.r.t. $\|\cdot\|$. Let $\Omega \subseteq \mathbb{R}^n$ be a compact convex set with diameter D. Assume that Ω is μ_Ω-strongly convex.

Now, consider the following algorithm

Algorithm 1 Frank–Wolfe for strongly convex sets and objectives

Given: $x_0 \in \Omega$

- For $t = 0, \ldots,$
  - Set $y_t \in \arg\min_{y \in \Omega} \langle \nabla f(x_t), y_t \rangle$
  - Set $x_{t+1} = (1 - \eta_t)x_t + \eta_t y_t$ where

$$\eta_t = \begin{cases} 1 & \text{if } t = 0 \\ 1 & \text{if } L_f \leq \frac{\mu_\Omega}{2} \|\nabla f(x_t)\|_* \\ \frac{\mu_\Omega \|\nabla f(x_t)\|_*}{4L} & \text{else} \end{cases}.$$

Let $\delta_t \coloneqq f(x_t) - f^\star$.

(a) [0pts] Recall that $\delta_1 \leq \frac{LD^2}{2}$.

(b) [15pts] Prove that for all $t \geq 1$,

$$\langle \nabla f(x_t), x_t - y_t \rangle \geq \frac{1}{2}\delta_t + \frac{\mu}{4} \|x_t - y_t\|^2 \|\nabla f(x_t)\|_*.$$

Hint: First, show that $\langle \nabla f(x_t), x_t - y_t \rangle \geq \frac{\mu_\Omega}{2} \|x_t - y_t\|^2 \|\nabla f(x_t)\|_*$ using the strong convexity of Ω.

(c) [15pts] Prove that for all $t \geq 1$,

$$\delta_{t+1} \leq \max\left(\frac{1}{2}, \left(1 - \frac{\mu_\Omega \|\nabla f(x_t)\|_*}{8L_f}\right)\right) \cdot \delta_t.$$

(d) [15pts] Now, use the fact that f is μ_f-strongly convex to show that if $0 < \epsilon \ll 1$, then $\delta_T \leq \epsilon$ for

$$T = O\left(\frac{L}{\mu_\Omega \sqrt{\mu_f}\sqrt{\epsilon}}\right).$$

Hint 1: Strong convexity allows you to upper bound $\|\nabla f(x_t)\|_*$ in terms of δ_t.

Hint 2: Consider the sequence $\delta_1, \delta_2, \ldots$. Fix $k \geq 1$. For how many indices $t \geq 1$, can it be the case that $\delta_t \in (\frac{LD^2}{2^{k+1}}, \frac{LD^2}{2^k}]$?