Continuous Optimization
Date: 5 October 2023
Submission deadline:
12 October, 12:00

Exercise 1 (1pt)**.** Give an example of a function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\|\nabla f(x)\|_2 \leq 1$ for all $x \in \mathbb{R}^n$ but the Lipschitz constant of its gradient is unbounded.

Exercise 2 (2pts)**.** Consider the function $f(x) := \frac{x^2}{|x|+1}$. As once can see, the function f is a "smooth" variant of $x \mapsto |x|$ in the sense that it is differentiable everywhere and $|f(x) - |x|| \to 0$ when x tends to either $+\infty$ or $-\infty$. Similarly, once can consider a multivariate extension $F : \mathbb{R}^n \to \mathbb{R}$ given by

$$F(x) := \sum_{i=1}^{n} \frac{x_i^2}{|x_i| + 1}.$$

This can be seen as a smoothening of the ℓ_1-norm $\|x\|_1$. Prove that $\|\nabla F(x)\|_\infty \leq 1$ for all $x \in \mathbb{R}^n$.

Exercise 3 (2pts)**.** Since the function $x \mapsto \max\{x_1, \ldots, x_n\}$ is not differentiable, one often considers the so-called soft-max function

$$s_\alpha(x) := \frac{1}{\alpha} \log\left(\sum_{i=1}^{n} e^{\alpha x_i}\right)$$

for some $\alpha > 0$, as a replacement for the maximum. Prove that

$$\max\{x_1, \ldots, x_n\} \leq s_\alpha(x) \leq \frac{\log n}{\alpha} + \max\{x_1, \ldots, x_n\},$$

thus the larger α is, the better approximation we obtain. Further, prove that for every $x \in \mathbb{R}^n$,

$$\|\nabla s_\alpha(x)\|_\infty \leq 1.$$

Exercise 4 (1pt)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a strictly convex function. Prove that the function $x \mapsto D_F(x, y)$ for a fixed $y \in \mathbb{R}^n$ is strictly convex.

Exercise 5 (1pt)**.** Prove that for all $p \in \Delta_n$, $D_{KL}(p, p^1) \leq \log n$. Here p^1 is the uniform probability distribution with $p_i^1 = \frac{1}{n}$ for $i = 1, \ldots, n$.

Exercise 6 (4pts)**.** For a nonempty subset $K \subseteq \mathbb{R}^n$, we can define the distance function $\operatorname{dist}(\cdot, K)$ and the projection operator $\operatorname{proj}_K : \mathbb{R}^n \to K$ as $\operatorname{dist}(x, K) := \inf_{y \in K} \|x - y\|$ and $\operatorname{proj}_K(x) := \operatorname{arginf}_{y \in K} \|x - y\|$.

(a) Prove that proj_K is well-defined when K is a closed and convex set, i.e., show that the minimum is attained at a unique point.

(b) Prove that for all $x, y \in \mathbb{R}^n$, we have $\|\operatorname{proj}_K(x) - \operatorname{proj}_K(y)\| \leq \|x - y\|$.

(c) Prove that for any set $K \subseteq \mathbb{R}^n$, the function $x \mapsto \operatorname{dist}^2(x, K)$ is convex.

(d) Let $K = B_{m,\infty} = \{x \in \mathbb{R}^m : \ \|x\|_\infty \leq 1\}$. Prove that the function $f(x) := \operatorname{dist}^2(x, K)$ has a Lipschitz continuous gradient with Lipschitz constant equal to 2.