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

Exercise 1 (2pts)**.** Let $\mathcal{M}$ be a nonempty family of subsets of $\{1,\ldots,n\}$. For a set $M \in \mathcal{M}$, let $1_M \in \mathbb{R}^n$ be the indicator vector of M, i.e., $1_M(i) = 1$ if $i \in M$ and $1_M(i) = 0$ otherwise. Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x) := \log\left(\sum_{M\in\mathcal{M}} e^{\langle x, 1_M\rangle}\right).$$

Prove that the gradient of f is L-Lipschitz continuous for some $L > 0$ that depends polynomially on n with respect to the Eucledian norm.

Exercise 2 (3pts)**. Online convex optimization.** Given a sequence of convex and differentiable functions $f^1, f^2, \cdots : K \to \mathbb{R}$ and a sequence of points $x^1, x^2, \cdots \in K$, define the **regret** up to time T to be

$$\mathrm{Regret}_T := \sum_{t=1}^{T} f^t(x^t) - \min_{x\in K}\sum_{t=1}^{T} f^t(x).$$

Consider the following strategy inspired by the mirror descent algorithm, called **follow the regularized leader**:

$$x^{t+1} := \operatorname*{argmin}_{x\in K} \sum_{i=1}^{t} f^i(x) + R(x)$$

for a convex regularizer $R : K \to \mathbb{R}$ and $x^1 := \operatorname{argmin}_{x\in K} R(x)$. Assume that the gradient of each f_i is bounded everywhere by G and that the diameter of K is bounded by D. Prove the following:

(a) $\mathrm{Regret}_T \le \sum_{t=1}^{T}(f^t(x^t) - f^t(x^{t+1})) - R(x^1) + R(x^*)$ for every $T = 1,\ldots$, where $x^* := \operatorname{argmin}_{x\in K}\sum_{t=1}^{T} f^t(x)$.

(b) Given $\varepsilon > 0$, use this method for $R(x) := \frac{1}{\eta}\|x\|_2^2$ for an appropriate choice of η and T to get $\frac{1}{T}\mathrm{Regret}_T \le \varepsilon$.

Exercise 3 (3pts)**.** Let $G = (V, E)$ be an undirected graph with n vertices and m edges, and let $s, t \in V$ distinct vertices. Let $B \in \mathbb{R}^{n\times m}$ denote the vertex-edge incidence matrix of the graph. Consider the $s - t$ maximum flow problem:

$$\begin{aligned} \max_{x\in\mathbb{R}^m, F\ge 0} \quad & F \\ s.t. \quad & Bx = Fb, \\ & \|x\|_\infty \le 1, \end{aligned}$$

where $b = e_s - e_t$.

(a) Prove that the dual of this formulation is equivalent to the following:

$$\begin{aligned} \min_{y\in\mathbb{R}^n} \quad & \sum_{uv\in E} |y_u - y_v| \\ s.t. \quad & y_s - y_t = 1. \end{aligned}$$

(b) Prove that the optimal value of the above problem is equal to $\mathrm{MinCut}_{s,t}(G)$, the minimum number of edges one needs to remove from G to disconnect s from t. This latter problem is known as the $s - t$ minimum cut problem.

(c) Reformulate the dual as follows:

$$\begin{aligned} \min_{x\in\mathbb{R}^m} \quad & \|x\|_1 \\ s.t. \quad & x \in \mathrm{Im}(B^T), \\ & \langle x, z\rangle = 1 \end{aligned}$$

for some $z \in \mathbb{R}^m$ that depends on G and s, t. Write an explicit formula for z.