Deep learning
Date: 5 March 2025
Submission deadline:
12 March, 15:30

Exercise 1 (1pt)**.** Let $F\colon \mathbb{R}^n \to \mathbb{R}$ 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 2 (2pts)**.** 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 $p^1_i = \frac{1}{n}$ for $i = 1, \ldots, n$.

Exercise 3 (2pts)**.** Prove that for every $x, y \in \Delta_n$ we have $D_{KL}(x, y) \geq \frac{1}{2}\|x - y\|_1^2$.

Exercise 4 (2pts)**.** Let $F\colon K \to \mathbb{R}$ be a convex, differentiable function and let $S \subseteq K$ be a closed, convex subset of K. Let $x, y \in S$ and $z \in K$ such that $y \coloneqq \operatorname{argmin}_{u \in S} D_F(u, z)$. Prove that $D_F(x, y) + D_F(y, z) \leq D_F(x, z)$. [Recall that D_F denotes the Bregman-divergence corresponding to F.]

Exercise 5 (2pts)**.** Prove that if $f\colon \mathbb{R}^n \to \mathbb{R}$ is a convex function which is G-Lipschitz continuous, then $\|\nabla f(x)\| \leq G$ for all $x \in \mathbb{R}^n$.

Exercise 6 (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 (we consider an arbitrary orientation of every edge and write $+1$ to the head and -1 to the tail of the edge). Consider the $s - t$ maximum flow problem:

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

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.