Advanced Graph Algorithms and Optimization Spring 2024

Course Introduction & Convex Optimization

R. Kyng & M. Probst *Problem Set 1 — Monday, February 19th*

These exercises will not count toward your grade, but you are encouraged to solve them all. This exercise sheet contains exercises relating to lectures in Week 1.

To get feedback, you must hand in your solutions by 23.59 pm on February 29th. Both hand-written and LaTeX solutions are acceptable, but we will only attempt to read legible text.

Exercise 1

Let us define the *α-sub-level set* of a function $f : \mathbb{R}^n \to \mathbb{R}$ to be the set $S_\alpha \stackrel{\text{def}}{=} \{\boldsymbol{x} : f(\boldsymbol{x}) \leq \alpha\}$.

(i) Prove that if a function f is convex, then all its sub-level sets are convex sets.

(ii) Is it true that a function whose sub-levels sets are all convex is necessarily convex?

Exercise 2

Recall the definition of Laplacian $\boldsymbol{L} = \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top$.

(i) We can also define Laplacian as $\boldsymbol{L} \stackrel{\text{def}}{=} \boldsymbol{D} - \boldsymbol{A}$, where $\boldsymbol{A}$ is the weighted adjacency matrix, i.e. $\boldsymbol{A}(u,v) = 1/\boldsymbol{r}(u,v)$, and $\boldsymbol{D} \stackrel{\text{def}}{=} \text{diag}_{v\in V}\, \boldsymbol{w}(v)$ for $\boldsymbol{w}(v) := \sum_{(u,v)\in E} 1/\boldsymbol{r}(u,v)$. Prove that these two definitions are equivalent.

(ii) Given a function on the vertices, $\boldsymbol{x} \in \mathbb{R}^V$, the Laplacian quadratic form is

$$\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = \sum_{(u,v)\in E} \frac{(\boldsymbol{x}(u) - \boldsymbol{x}(v))^2}{\boldsymbol{r}(u,v)}.$$

Prove the above equality and building on that, show that $\boldsymbol{L}$ is positive semi-definite.

(iii) What is the kernel of $\boldsymbol{L}$, which is denoted by $\text{Ker}(\boldsymbol{L})$?

Exercise 3

(i) Prove that for a matrix $\boldsymbol{A}$ we have $\text{im}(\boldsymbol{A}) = \ker(\boldsymbol{A}^\top)^\perp$, where $\text{im}(\boldsymbol{A})$ denotes the image of $\boldsymbol{A}$ and $\ker(\boldsymbol{A}^\top)^\perp$ is the orthogonal complement to $\ker(\boldsymbol{A}^\top)$.

(ii) Building on part (i), prove that in a connected graph with resistances $\boldsymbol{r} \in \mathbb{R}^E_{>0}$, an electrical flow $\boldsymbol{f}$ routing demand $\boldsymbol{d}$ exists if and only if $\boldsymbol{1}^\top \boldsymbol{d} = 0$.

Exercise 4

We define[1] the gradient of a multivariate function $f : \mathbb{R}^n \to \mathbb{R}$ as the function $\boldsymbol{\nabla} f : \mathbb{R}^n \to \mathbb{R}^n$ with $\boldsymbol{\nabla} f(\boldsymbol{x})(i) = \frac{d}{d\boldsymbol{x}(i)} f(\boldsymbol{x})$, i.e. we consider f at a point $\boldsymbol{x}$, treat the ith coordinate as a variable $\boldsymbol{x}(i)$, take a derivative w.r.t. it and then evaluate it at the point $\boldsymbol{x}$.

Now, prove that the system of linear equations $\boldsymbol{L}\boldsymbol{x} = \boldsymbol{d}$ is the same as the system obtained by setting the gradient of the function $c(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{d}$ equal to zero.

Exercise 5

The goal of this exercise is to prove that

$$\max_{\boldsymbol{x} \in \mathbb{R}^V} \boldsymbol{x}^\top \boldsymbol{d} - \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = \min_{\boldsymbol{f} \in \mathbb{R}^E} \frac{1}{2} \sum_e \boldsymbol{r}(e)\boldsymbol{f}(e)^2 \quad \text{s.t. } \boldsymbol{B}\boldsymbol{f} = \boldsymbol{d}.$$

We'll break that down into a few steps.

Let $\boldsymbol{f} \in \mathbb{R}^E$ be an arbitrary flow that satisfies $\boldsymbol{B}\boldsymbol{f} = \boldsymbol{d}$, i.e. it routes the demand $\boldsymbol{d}$. Let $\boldsymbol{x} \in \mathbb{R}^V$ be arbitrary voltages. *Arbitrary* means you cannot assume these are the electrical flow and voltages.

(i) Prove that

$$\frac{1}{2}\sum_e \boldsymbol{r}(e)\boldsymbol{f}(e)^2 = \boldsymbol{x}^\top \boldsymbol{d} - \left(\sum_{(u,v) \in E} (\boldsymbol{x}(u) - \boldsymbol{x}(v))(\boldsymbol{f}(u,v)) - \frac{1}{2}\boldsymbol{r}(u,v)\boldsymbol{f}(u,v)^2 \right)$$

Hint: use that $\boldsymbol{x}^\top(\boldsymbol{B}\boldsymbol{f} - \boldsymbol{d}) = 0$.

(ii) Prove that

$$(\boldsymbol{x}(u) - \boldsymbol{x}(v))(\boldsymbol{f}(u,v)) - \frac{1}{2}\boldsymbol{r}(u,v)\boldsymbol{f}(u,v)^2 \leq \frac{1}{2}\frac{(\boldsymbol{x}(u) - \boldsymbol{x}(v))^2}{\boldsymbol{r}(u,v)}.$$

(iii) Conclude that $\frac{1}{2}\boldsymbol{f}^\top \boldsymbol{R}\boldsymbol{f} \geq \boldsymbol{x}^\top \boldsymbol{d} - \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x}$.

(iv) Assume we are given $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{f}}$ such that

$$\boldsymbol{L}\tilde{\boldsymbol{x}} = \boldsymbol{d} \text{ and } \tilde{\boldsymbol{f}} = \boldsymbol{R}^{-1}\boldsymbol{B}^\top \tilde{\boldsymbol{x}}$$

Prove that $\boldsymbol{B}\tilde{\boldsymbol{f}} = \boldsymbol{d}$ and

$$\tilde{\boldsymbol{x}}^\top \boldsymbol{d} - \frac{1}{2}\tilde{\boldsymbol{x}}^\top \boldsymbol{L}\tilde{\boldsymbol{x}} = \frac{1}{2}\tilde{\boldsymbol{f}}^\top \boldsymbol{R}\tilde{\boldsymbol{f}}.$$

(v) Show

$$\tilde{\boldsymbol{x}} \in \arg\max_{\boldsymbol{x} \in \mathbb{R}^V} \boldsymbol{x}^\top \boldsymbol{d} - \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x}$$

and

$$\tilde{\boldsymbol{f}} \in \arg\min_{\boldsymbol{f} \in \mathbb{R}^E} \frac{1}{2}\sum_e \boldsymbol{r}(e)\boldsymbol{f}(e)^2 \quad \text{s.t. } \boldsymbol{B}\boldsymbol{f} = \boldsymbol{d}.$$

[1] We will give a more formal definition of Frechet derivatives later, which is formally what we mean by 'gradient'.

Exercise 6

Recall that the following theorem gives us a sufficient (though not necessary) condition for optimality.

Theorem (Extreme Value Theorem)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a continuous function and* $\mathcal{F} \subseteq \mathbb{R}^n$ *be nonempty, bounded, and closed. Then, the optimization problem* $\min f(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{F}$ *has an optimal solution.*

Prove the above theorem. You might use the following two theorems.

Theorem (Bolzano-Weierstrass)**.** *Every bounded sequence in* $\mathbb{R}^n$ *has a convergent subsequence.*

Theorem (Boundedness Theorem)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a continuous function and* $\mathcal{F} \subseteq \mathbb{R}^n$ *be nonempty, bounded, and closed. Then* f *is bounded on* $\mathcal{F}$*.*

Exercise 7

Prove or sketch a proof of Taylor's Theorem.

Theorem (Taylor's Theorem, multivariate first-order remainder form)**.** *If* $f : S \to \mathbb{R}$ *is continuously differentiable over* $[\boldsymbol{x}, \boldsymbol{y}]$*, then for some* $\boldsymbol{z} \in [\boldsymbol{x}, \boldsymbol{y}]$*, we have* $f(\boldsymbol{y}) = f(\boldsymbol{x}) + \boldsymbol{\nabla} f(\boldsymbol{z})^\top (\boldsymbol{y} - \boldsymbol{x})$.

Exercise 8

Let $f_1(\boldsymbol{x}), \cdots, f_k(\boldsymbol{x})$ be a collection of convex functions all with the same domain and define $f(\boldsymbol{x}) \stackrel{\text{def}}{=} \max_{1 \le i \le k} f_i(\boldsymbol{x})$. Prove that $f(\boldsymbol{x})$ is convex.

Exercise 9

Assume that $f(x, y)$ is a convex function and S is a convex non-empty set. Prove that

$$g(x) = \inf_{y \in S} f(x, y)$$

is convex, provided $g(x) > -\infty$ for all x.

Exercise 10

For each function below, determine whether it is convex or not.

1. $f(x) = |x|^6$ on $x \in \mathbb{R}$
2. $f(x) = \exp(x)$ on $x \in (0, \infty)$
3. $f(x, y) = \sqrt{x + y}$ on $(x, y) \in (0, 1) \times (0, 1)$
4. $f(x, y) = xy$ on $(x, y) \in (-1, 1) \times (-1, 1)$

Bonus Exercise 11: Jensen's Inequality

This exercise will teach you about Jensen's inequality, one of the most important inequalities that we use when studying convex functions.

1. Assume that $S \subseteq \mathbb{R}^n$ is a convex set and that the function $f : S \to \mathbb{R}$ is convex. Suppose that $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n \in S$ and $\theta_1, \cdots, \theta_n \geq 0$ with $\theta_1 + \cdots + \theta_n = 1$. Prove that

$$f(\theta_1 \boldsymbol{x}_1 + \cdots + \theta_n \boldsymbol{x}_n) \leq \theta_1 f(\boldsymbol{x}_1) + \cdots + \theta_n f(\boldsymbol{x}_n).$$

   Remark. *This is typically known as Jensen's inequality and can be extended to infinite sums. If $\mathcal{D}$ is a probability distribution on S, and $\boldsymbol{X} \sim \mathcal{D}$, then $f(\mathbb{E}[\boldsymbol{X}]) \leq \mathbb{E}[f(\boldsymbol{X})]$ whenever both integrals are finite.*

2. Prove that $\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^n x_i$.

3. Prove that $\frac{1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}} \leq \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}$.