Advanced Graph Algorithms and Optimization	Spring 2023
(Mostly) Convex Duality	
R. Kyng & M. Probst Gutenberg	*Problem Set 8 — Wednesday, May 3rd*

The exercises for this week will not count toward your grade, but you are highly encouraged to solve them all. This exercise sheet has exercises related to week 8. We encourage you to start early so you have time to go through everything.

To get feedback, you must hand in your solutions by 23:59 on May 11. Both hand-written and LaTeX solutions are acceptable, but we will only attempt to read legible text.

Exercise 1: Different Duals

Let $G = (V, E)$ be a directed graph with capacities $\boldsymbol{c} \in \mathbb{R}^E \geq \mathbf{0}$, and edge-vertex incidence matrix $\boldsymbol{B}$, and consider a demand vector $\boldsymbol{d} \in \mathbb{R}^V$ with $\boldsymbol{d} \perp \mathbf{1}$.

Let us try to minimize $\left\|\boldsymbol{C}^{-1}\boldsymbol{f}\right\|_\infty$, where $\boldsymbol{C} = \text{diag}_{e\in E}\, \boldsymbol{c}(e)$. This leads to an optimization problem

$$\min_{\boldsymbol{f}\in\mathbb{R}^E_{\geq 0}} \left\|\boldsymbol{C}^{-1}\boldsymbol{f}\right\|_\infty \quad \text{s.t. } \boldsymbol{B}\boldsymbol{f} = \boldsymbol{d}$$

Solution

- Compute the dual of this problem. Note that we encoded the "constraint" $\boldsymbol{f} \geq \mathbf{0}$ in the domain of the variable $\boldsymbol{f}$. This means that there will *not* be a dual variable associated with this constraint.

  The dual is:

$$\begin{aligned} &\max_{\boldsymbol{x}} \min_{\boldsymbol{f}\in\mathbb{R}^E_{\geq 0}} \left\|\boldsymbol{C}^{-1}\boldsymbol{f}\right\|_\infty + \boldsymbol{x}^\top(\boldsymbol{d} - \boldsymbol{B}\boldsymbol{f}) \\ &= \max_{\boldsymbol{x}} \left(\boldsymbol{x}^\top \boldsymbol{d} + \min_{\boldsymbol{f}\in\mathbb{R}^E_{\geq 0}} -\boldsymbol{f}^\top \boldsymbol{B}^\top \boldsymbol{x} + \left\|\boldsymbol{C}^{-1}\boldsymbol{f}\right\|_\infty \right) \\ &= \max_{\boldsymbol{x}} \boldsymbol{x}^\top \boldsymbol{d} \\ &\text{s.t. } \sum_e \max(\boldsymbol{c}(e)\boldsymbol{b}_e^\top \boldsymbol{x}, 0) \leq 1 \end{aligned}$$

- Does Strong Duality hold for this pair of primal and dual problems?

  Slater's condition holds when there exists $\boldsymbol{f} > \mathbf{0}$ s.t. $\boldsymbol{B}\boldsymbol{f} = \boldsymbol{d}$.

  Note that there exists a flow $\boldsymbol{f} \geq \mathbf{0}$ s.t. $\boldsymbol{B}\boldsymbol{f} = \boldsymbol{d}$ if and only if there exists a directed path from s to t.

  We can construct a flow $\boldsymbol{f} > \mathbf{0}$ s.t. $\boldsymbol{B}\boldsymbol{f} = \boldsymbol{d}$ if and only if for every directed edge $(u, v) \in E$, there exists a directed path from s to u and a directed path from v to t. We can test decide this for all edges using a BFS starting at s forward along directed edges, and a BFS backward along directed edges from t. Thus we can check if Slater's condition holds, and in this case Strong Duality holds. If some edge is not reachable from s or cannot reach t in this way, then it cannot carry non-zero flow and we can remove it and obtain an equivalent problem. However, we won't address whether Strong Duality holds when such an edge exists and hasn't been deleted (in fact it does hold, which can be argued via Linear Programming Duality).

Consider also a variant of this problem, where we instead treat $\boldsymbol{f} \geq \boldsymbol{0}$ as an explicit constraint:

$$\begin{aligned} \min_{\boldsymbol{f} \in \mathbb{R}^E} & \left\| \boldsymbol{C}^{-1} \boldsymbol{f} \right\|_\infty \\ \text{s.t. } & \boldsymbol{B}\boldsymbol{f} = \boldsymbol{d} \\ & \boldsymbol{f} \geq \boldsymbol{0} \end{aligned}$$

This problem has a different dual problem because we made $\boldsymbol{f} \geq \boldsymbol{0}$ an explicit constraint.

- Compute the dual of this variant of the problem.

  The dual is:

$$\begin{aligned} & \max_{\substack{\boldsymbol{x} \in \mathbb{R}^V, \boldsymbol{s} \in \mathbb{R}^E \\ \boldsymbol{s} \geq \boldsymbol{0}}} \min_{\boldsymbol{f} \in \mathbb{R}^E} \boldsymbol{s}^\top(-\boldsymbol{f}) + \left\| \boldsymbol{C}^{-1}\boldsymbol{f} \right\|_\infty + \boldsymbol{x}^\top(\boldsymbol{d} - \boldsymbol{B}\boldsymbol{f}) \\ &= \max_{\boldsymbol{x}} \left(\boldsymbol{x}^\top \boldsymbol{d} + \min_{\boldsymbol{f} \in \mathbb{R}^E} -\boldsymbol{f}^\top(\boldsymbol{s} + \boldsymbol{B}^\top \boldsymbol{x}) + \left\| \boldsymbol{C}^{-1}\boldsymbol{f} \right\|_\infty \right) \\ &= \max_{\substack{\boldsymbol{x} \in \mathbb{R}^V, \boldsymbol{s} \in \mathbb{R}^E \\ \boldsymbol{s} \geq \boldsymbol{0}}} \boldsymbol{x}^\top \boldsymbol{d} \\ & \text{s.t. } \sum_e \boldsymbol{c}(e) \left| \boldsymbol{s}(e) + \boldsymbol{b}_e^\top \boldsymbol{x} \right| \leq 1 \end{aligned}$$

- Explain how, given an optimal solution to the first dual problem, we can compute an optimal solution to the second dual problem.

  Observe that, given *any* $\boldsymbol{x}$, the optimal choice of $\boldsymbol{s}(e)$ is $\boldsymbol{s}(e) = \max(-\boldsymbol{b}_e^\top \boldsymbol{x}, 0)$, which gives

$$\boldsymbol{c}(e) \left| \boldsymbol{s}(e) + \boldsymbol{b}_e^\top \boldsymbol{x} \right| = \max(\boldsymbol{c}(e)\boldsymbol{b}_e^\top \boldsymbol{x}, 0).$$

  We can see that after choosing this value of $\boldsymbol{s}$, the two duals are equal and must have the same set of optimal $\boldsymbol{x}^*$, and given any such $\boldsymbol{x}^*$, we can get optimal $\boldsymbol{s}^*$ by the above relation.

Exercise 2: A "Broken" Dual

Consider the following optimization problem

$$\begin{aligned} \inf_{x \in \mathbb{R}, y \in \mathbb{R}_{>0}} & e^{-x} \\ \text{s.t. } & x^2/y \leq 0. \end{aligned}$$

- Compute the dual program.
- What is the optimal value of the primal program? And of the dual program? Does strong duality hold? Does Slater's condition hold?

Solution The *Lagrangian* of this problem is:

$$L(x, y, s) = e^{-x} + sx^2/y$$

Hence, we have the following dual:

$$\sup_{s\geq 0} \inf_{x\in\mathbb{R}, y\in\mathbb{R}_{>0}} e^{-x} + sx^2/y$$

Observe that $y \to x^3$ and $x \to \infty$ gives a value of 0 in the limit. Thus, we are left with the following dual problem:

$$\sup_{s\geq 0} 0$$

The optimal value of this problem is $\beta^* = 0$. Observe that the only primal-feasible x is $x = 0$. Hence, we have for the optimal values of the primal and dual: $\alpha^* = 1 > 0 = \beta^*$. In other words, *strong duality* does not hold. Slater's condition is not fulfilled since the inequality $x^2/y \leq 0$ does not hold strictly for any $x \in \mathbb{R}$ and $y \in \mathbb{R}_{>0}$.

Exercise 3: Norms and a Lagragian

Suppose $1 < q < p < \infty$. In this exercise, we want to prove that for $\boldsymbol{x} \in \mathbb{R}^n$, we have

$$\|\boldsymbol{x}\|_q \leq n^{1/q-1/p} \|\boldsymbol{x}\|_p . \tag{1}$$

Consider the following optimization problem:

$$\max_{\boldsymbol{x}\in\mathbb{R}^n : \|\boldsymbol{x}\|_p^p \leq 1} \|\boldsymbol{x}\|_q^q$$

- Is this a convex optimization problem?
- Informally, explain why, at any maximizing $\boldsymbol{x}$ for the above problem, there should exist a $\lambda > 0$ such that
$$\boldsymbol{\nabla}_{\boldsymbol{x}} \left(\|\boldsymbol{x}\|_q^q - \lambda \|\boldsymbol{x}\|_p^p \right) = \mathbf{0}.$$
- Use the existence of such a λ to prove Equation (1).

Solution It is not a convex optimization since we are maximizing a norm instead of minimizing it. Assume that such a λ would not exist. Then we could find a $\boldsymbol{\delta}$ such that $\boldsymbol{\delta} \perp \nabla \|\boldsymbol{x}\|_p^p$ and $\boldsymbol{\delta} \not\perp \nabla \|\boldsymbol{x}\|_q^q$. Thus we can move from $\boldsymbol{x}$ to $\boldsymbol{x} + \epsilon\boldsymbol{\delta}$ for a infinitesimally small ϵ increasing $\|\boldsymbol{x}\|_q^q$ while $\|\boldsymbol{x}\|_p^p$ stays constant. This would contradict the maximality of $\boldsymbol{x}$.
The existence of λ implies that for all $i \in [n]$:

$$q \cdot |\boldsymbol{x}_i|^{q-1} - \lambda \cdot p \cdot |\boldsymbol{x}_i|^{p-1} = 0 \quad \Leftrightarrow \quad |\boldsymbol{x}_i| = \begin{cases} \left(\frac{q}{p\cdot\lambda}\right)^{1/(p-q)} \\ 0 \end{cases}$$

In other words, for every maximizer, there exists a constant a such that $|\boldsymbol{x}_i| = \begin{cases} a \\ 0 \end{cases}$ for all $i \in [n]$.

Suppose k coordinates are non-zero. Then we see that

$$1 = \|\boldsymbol{x}\|_p^p = \sum_i a^p = k \cdot a^p \quad \Leftrightarrow \quad a = k^{-1/p}$$

Computing the q-norm of $\boldsymbol{x}$: $\|\boldsymbol{x}\|_q^q = \sum_i a^q = k^{1-q/p} \leq n^{1-q/p}$. Note that we can conclude that the maximizer of the q given our p-norm constraint is non-zero on all coordinates, and equal on all of them up to a sign.

Consider any $\boldsymbol{x} \in \mathbb{R}^n$ and let $\boldsymbol{x}^*$ be the maximizer of the problem above. Then we have:

$$n^{1-q/p} = \|\boldsymbol{x}^*\|_q^q \geq \left\| \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_p} \right\|_q^q = \frac{\|\boldsymbol{x}\|_q^q}{\|\boldsymbol{x}\|_p^q} \quad \Leftrightarrow \quad n^{1/q-1/p} \|\boldsymbol{x}\|_p \geq \|\boldsymbol{x}\|_q .$$

Exercise 4: Flows and Voltages and Other Powers

Consider a connected undirected graph $G = (V, E)$ with resistances $\boldsymbol{r} \in \mathbb{R}^E$ and edge-vertex incidence matrix $\boldsymbol{B}$, and a demand vector $\boldsymbol{d} \in \mathbb{R}^V$ with $\boldsymbol{d} \perp \boldsymbol{1}$.

Given some $p > 1$, we'll look at the flow optimization problem

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

- Is the above optimization problem convex?
- Does Slater's condition hold for the problem?
- What is the dual problem for the problem? Define $q > 0$ to be the number such that $1 = \frac{1}{q} + \frac{1}{p}$. Try to find a clean expression of the dual. Writing the expression in terms of q instead of p will simplify it.

Suppose our instance of the optimization problem has an optimal flow solution $\boldsymbol{f}^*$. Let $\alpha = \sum_e \boldsymbol{r}(e) \frac{1}{p} |\boldsymbol{f}^*(e)|^p$ be the optimal problem value. Suppose that for some particular edge $\hat{e}$ we have

$$\gamma = \frac{\boldsymbol{r}(\hat{e}) \frac{1}{p} |\boldsymbol{f}^*(\hat{e})|^p}{\alpha}.$$

Now, consider a modified instance with resistances $\tilde{\boldsymbol{r}} \in \mathbb{R}^E$ given by

$$\tilde{\boldsymbol{r}}(e) = \begin{cases} \boldsymbol{r}(e) & \text{for } e \neq \hat{e} \\ 2^{p-1} \boldsymbol{r}(e) & \text{for } e = \hat{e} \end{cases}$$

That is, we increase the resistance on edge $\hat{e}$ by a factor 2^{p-1}. Let $\tilde{\alpha}$ denote optimal value of program with the new resistances $\tilde{\boldsymbol{r}}$. Prove that

$$\tilde{\alpha} \geq \left(1 + \frac{p-1}{2} \gamma\right) \alpha.$$

Hint: use the dual problem!

Solution The optimization problem is convex. There are only linear constraints and the objective function is convex because $|\cdot|$ is convex and x^p is convex and increasing on $[0, \infty)$ for $p > 1$. Slater's condition holds since we only have linear constraints.

The *Lagrangian* corresponding to the optimization problem is:

$$L(\boldsymbol{f}, \boldsymbol{x}) = \sum_e \boldsymbol{r}(e) \frac{1}{p} |\boldsymbol{f}(e)|^p + \boldsymbol{x}^\top (\boldsymbol{d} - \boldsymbol{B}\boldsymbol{f}) = \boldsymbol{x}^\top \boldsymbol{d} + \sum_e \boldsymbol{r}(e) \frac{1}{p} |\boldsymbol{f}(e)|^p - \boldsymbol{f}(e) \boldsymbol{x}^\top \boldsymbol{b}_e$$

Thus, we get the following dual:

$$\max_{\boldsymbol{x} \in \mathbb{R}^V} \min_{\boldsymbol{f} \in \mathbb{R}^E} \boldsymbol{x}^\top \boldsymbol{d} + \sum_e \boldsymbol{r}(e) \frac{1}{p} |\boldsymbol{f}(e)|^p - \boldsymbol{f}(e) \boldsymbol{x}^\top \boldsymbol{b}_e$$

In order to compute the minimizing $\boldsymbol{f}^*$ we will set the gradient of $\nabla_{\boldsymbol{f}} L(\boldsymbol{f}, \boldsymbol{x})$ to zero:

$$0 = \frac{\partial L(\boldsymbol{f}, \boldsymbol{x})}{\partial \boldsymbol{f}(e)} = \operatorname{sign}(\boldsymbol{f}(e)) \cdot \boldsymbol{r}(e) \cdot |\boldsymbol{f}(e)|^{p-1} - \boldsymbol{x}^\top \boldsymbol{b}_e$$

Note that $p - 1 = p/q$. Hence, the above is equivalent to:

$$\boldsymbol{f}(e) = \boldsymbol{r}(e)^{-q/p} \cdot \operatorname{sign}(\boldsymbol{x}^\top \boldsymbol{b}_e) \cdot \left| \boldsymbol{x}^\top \boldsymbol{b}_e \right|^{q/p} \tag{2}$$

Using that $q - 1 = q/p$, we can rewrite the dual problem:

$$\begin{aligned} \max_{\boldsymbol{x} \in \mathbb{R}^V} L(\boldsymbol{f}^*, \boldsymbol{x}) &= \max_{\boldsymbol{x} \in \mathbb{R}^V} \boldsymbol{x}^\top \boldsymbol{d} + \sum_e \boldsymbol{r}(e)^{1-q} \frac{1}{p} \left| \boldsymbol{x}^\top \boldsymbol{b}_e \right|^q - \boldsymbol{r}(e)^{-q/p} \left| \boldsymbol{x}^\top \boldsymbol{b}_e \right|^{q/p+1} \\ &= \max_{\boldsymbol{x} \in \mathbb{R}^V} \boldsymbol{x}^\top \boldsymbol{d} + \sum_e \underbrace{\left(\frac{1}{p} - 1 \right)}_{-1/q} \boldsymbol{r}(e)^{1-q} \left| \boldsymbol{x}^\top \boldsymbol{b}_e \right|^q \end{aligned}$$

Now, we are considering the modified instance of the problem. Note that:

$$\tilde{\boldsymbol{r}}(\hat{e})^{1-q} = 2^{(p-1)\cdot(1-q)} \boldsymbol{r}(\hat{e})^{1-q} = 2^{-1} \boldsymbol{r}(\hat{e})^{1-q}$$

Let $\boldsymbol{f}^*$ and $\boldsymbol{x}^*$ be the old optimal flow and voltages respectively. Because of 2 we can write:

$$\gamma = \frac{\boldsymbol{r}(\hat{e}) \frac{1}{p} |\boldsymbol{f}^*(\hat{e})|^p}{\alpha} = \frac{\boldsymbol{r}(\hat{e})^{1-q} \frac{1}{p} \left| \boldsymbol{x}^\top \boldsymbol{b}_{\hat{e}} \right|^q}{\alpha} \tag{3}$$

By strong duality we have:

$$\begin{aligned} \tilde{\alpha} &= \max_{\boldsymbol{x} \in \mathbb{R}^V} \boldsymbol{x}^\top \boldsymbol{d} - \sum_e \frac{1}{q} \tilde{\boldsymbol{r}}(e)^{1-q} \left| \boldsymbol{x}^\top \boldsymbol{b}_e \right|^q \\ &\geq \boldsymbol{x}^{*\top} \boldsymbol{d} - \sum_e \frac{1}{q} \tilde{\boldsymbol{r}}(e)^{1-q} \left| \boldsymbol{x}^{*\top} \boldsymbol{b}_e \right|^q \\ &= \boldsymbol{x}^{*\top} \boldsymbol{d} - \sum_{e \neq \hat{e}} \frac{1}{q} \boldsymbol{r}(e)^{1-q} \left| \boldsymbol{x}^{*\top} \boldsymbol{b}_e \right|^q - \frac{1}{2q} \boldsymbol{r}(\hat{e})^{1-q} \left| \boldsymbol{x}^{*\top} \boldsymbol{b}_{\hat{e}} \right|^q \\ &= \boldsymbol{x}^{*\top} \boldsymbol{d} - \sum_e \frac{1}{q} \boldsymbol{r}(e)^{1-q} \left| \boldsymbol{x}^{*\top} \boldsymbol{b}_e \right|^q + \frac{1}{2q} \boldsymbol{r}(\hat{e})^{1-q} \left| \boldsymbol{x}^{*\top} \boldsymbol{b}_{\hat{e}} \right|^q \\ &= \alpha + \frac{p}{2q} \gamma \alpha = \left(1 + \frac{p-1}{2} \gamma \right) \alpha \end{aligned}$$

In the second to last equality we used 3 and the optimality of $\boldsymbol{x}^*$. For the last equality we used that $p/q = p - 1$.