EECS 127/227AT Optimization Models in Engineering UC Berkeley Spring 2024

Homework 10

This homework is due at 11 PM on April 5, 2024.

Submission Format: Your homework submission should consist of a single PDF file that contains all of your answers (any handwritten answers should be scanned), as well as a printout of your completed Jupyter notebook(s).

1. **Does Strong Duality Hold?**

   Consider

$$\min_{(x,y)\in\mathcal{D}} \quad e^{-x} \tag{1}$$
$$\text{s.t.} \quad x^2/y \le 0 \tag{2}$$

   where $\mathcal{D} = \{(x,y) \mid y > 0\}$.

   (a) Prove the problem is convex. Find the optimal value. *HINT: To prove the constraint function is convex, you will have to prove it is convex with respect to the vector* $\begin{bmatrix} x & y \end{bmatrix}^\top$. *Consider computing the Hessian of the constraint function, its determinant and trace, and show that it is PSD by analyzing signs of its eigenvalues.*

   (b) Next, we will proceed to find an optimal solution and an optimal value for the dual problem. The Lagrangian dual function $g(\lambda)$, can be written as:

$$g(\lambda) = \inf_{(x,y)\in\mathcal{D}} e^{-x} + \lambda\frac{x^2}{y}. \tag{3}$$

   Explain why $g(\lambda)$ is lower bounded by 0 for $\lambda \ge 0$. *NOTE*: Here we are not dualizing the constraint $y > 0$ that is in the definition of $\mathcal{D}$ — this is only dualizing the other constraint.

   (c) Show that $g(\lambda) = 0$ for $\lambda \ge 0$. *HINT: To show that the infimum in (3) is 0, we want to show there exist* (x, y) *such that both* e^{-x} *and* $\lambda\frac{x^2}{y}$ *can get arbitrarily close to 0. HINT: Consider a sequence* $\{x_k\}$ *going to* $+\infty$ *and a sequence* $\{y_k\}$ *also going to* $+\infty$ *such that* $\lim_{k\to\infty} \frac{x_k^2}{y_k} = 0$. *Simply put, we want to drive* x *to infinity in order to drive* e^{-x} *to* 0, *while having* y *grow faster than* x^2, *so that the second term also goes to* 0.

   (d) Now, write the dual problem and find an optimal solution λ^* and an optimal value d^* for the dual problem using the results above. What is the duality gap?

   (e) Does Slater's Condition hold for this problem? Does Strong Duality hold?

2. About General Optimization

In this exercise, we test your understanding of the general framework of optimization and its language. We consider an optimization problem in standard form:

$$p^\star = \min_{\vec{x} \in \mathbb{R}^n} f_0(\vec{x}) \;:\; f_i(\vec{x}) \leq 0, \;\; i = 1, \ldots, m. \tag{4}$$

In the following we denote by $\mathcal{X}$ the feasible set. Note that the feasible set is a subset of $\mathbb{R}^n$ that satisfies the inequalities $f_i(\vec{x}) \leq 0$, i.e $\mathcal{X} = \{\vec{x} \in \mathbb{R}^n \mid f_i(\vec{x}) \leq 0, i = 1, \ldots, m\}$. We make no assumption about the convexity of $f_0(\vec{x})$ and $f_i(\vec{x}),\; i = 1, \ldots, m$. For the following statements, answer whether the statement is true or false and provide a proof or counter-example.

(a) A general optimization problem can be expressed as one with a linear objective.

(b) A general optimization problem can be expressed as an unconstrained problem with a different objective function which could possibly take a value of ∞ for some values of $\vec{x}$.

(c) A general optimization problem can be recast as minimizing a linear objective subject to (possibly infinitely many) linear constraints.

(d) If any of the constraint inequalities is strict (and therefore not active) at the optimum point, then we can remove the constraint from the original problem and obtain the same optimum value.

Note: Review the definition of active constraints from the textbooks: Boyd Section 4.1.1 and El Ghaoui Section 8.3.

Hint: Consider the problem

$$\min_x \; f(x) \; = \begin{cases} x^2 & \text{if } |x| \leq 1, \\ -1 & \text{otherwise} \end{cases} \tag{5}$$

$$\text{such that } \; |x| \leq 1 \tag{6}$$

(e) Now, suppose for the formulation in (4), $f_0(\vec{x})$ is a convex function, $\mathcal{X}$ is a convex set, and all $f_i(\vec{x})$ are convex and continuous functions. Suppose $p^\star$ is achieved at a point $\vec{x}^\star$ where for some i, $f_i(\vec{x}^\star) < 0$. Prove that we can remove this inequality constraint and still retain the same optimum. In other words, show that

$$p^\star = \min_{\vec{x} \in \mathbb{R}^n} f_0(\vec{x}) \;:\; f_j(\vec{x}) \leq 0, \;\; j = 1, \ldots i-1, i+1, \ldots, m. \tag{7}$$

HINT: Argue by contradiction that if by removing the inequality constraint $f_i(\vec{x}) \leq 0$, we achieve a different optimal for the problem in (4) at some point $\vec{\bar{x}}$ that satisfies $f_i(\vec{\bar{x}}) > 0$, then there exists a point $\vec{y}$ between $\vec{\bar{x}}$ and $\vec{x}^\star$ that is feasible to the original problem in (4). Use the continuity of f_i and the intermediate value theorem to come up with a $\vec{y}$ then show that it must be more optimal than $\vec{x}^\star$ in (4).

3. Formulating Optimization problems

(a) **Linear Separability.** Let $(\vec{x}_i, y_i)$ be given data points with $\vec{x}_i \in \mathbb{R}^n$ and binary labels $y_i \in \{-1, 1\}$. We want to know if it is possible to find a hyperplane $\mathcal{L} = \{\vec{x} \in \mathbb{R}^n : \vec{h}^\top \vec{x} + b = 0\}$ that separates all the points with labels $y_i = -1$ from all the points with labels $y_i = 1$. In other words, can we find a vector $\vec{h} \in \mathbb{R}^n$ and a scalar $b \in \mathbb{R}$ such that $\vec{h}^\top \vec{x}_i + b \leq 0$ for all i such that $y_i = 1$ and $\vec{h}^\top \vec{x}_i + b > 0$ for all i such that $y_i = -1$. We want to cast this task as the following LP

$$p^\star = \min_{\vec{h}, b, z} \quad f_0(\vec{h}, b, z) \tag{8}$$

$$s.t. \quad \vec{h}^\top \vec{x}_i + b \leq 0 \qquad \forall i : y_i = 1 \tag{9}$$

$$\vec{h}^\top \vec{x}_i + b \geq z \qquad \forall i : y_i = -1 \tag{10}$$

Complete this formulation by specifying a linear objective function f_0. What does the solution $p^\star$ say about the existence of the separating hyperplane?

(b) **Chebyshev Center.** Let $\mathcal{P} \subset \mathbb{R}^n$ be a non-empty polyhedron defined as the intersection of m hyperplanes $\mathcal{P} = \{\vec{x} : \vec{a}_i^\top \vec{x} \leq b_i \ \forall i = 1, 2, \ldots, m\}$. We define the Euclidean ball in $\mathbb{R}^n$ with radius R and center $\vec{x}_0$ as the set $\mathcal{B}(\vec{x}_0, R) = \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\|_2 \leq R\}$. We want to find a point $\vec{x}_0 \in \mathcal{P}$ that is the center of the largest Euclidean ball contained in $\mathcal{P}$. Cast this problem as an LP.

4. Water Filling

Consider the following problem:

$$\text{minimize} \quad -\sum_{i=1}^{n} \log(\alpha_i + x_i) \tag{11}$$

$$\text{subject to} \quad \vec{x} \geq 0,\ \vec{1}^\top \vec{x} = c, \tag{12}$$

where $\alpha_i > 0$ for each $i = 1, \ldots, n$.

This problem arises in information theory, in allocating power to a set of n communication channels. The variable x_i represents the transmitter power allocated to the ith channel, and $\log(\alpha_i + x_i)$ gives the capacity or communication rate of the channel, so the problem is to allocate a total power of c to the channels, in order to maximize the total communication rate.

(a) Verify that this is a convex optimization problem with differentiable objective and constraint functions. Find the domain $\mathcal{D}$ of the objective function $-\sum_{i=1}^{n} \log(\alpha_i + x_i)$ where it is well defined.

(b) Let $\vec{\lambda} \in \mathbb{R}^n$ and $\nu \in \mathbb{R}$ be the dual variables corresponding to the constraints $x_i \geq 0, i = 1, \ldots, n$ and $\vec{1}^\top \vec{x} = c$, respectively. Write a Lagrangian for the optimization problem based on these dual variables.

(c) Write the KKT conditions for the problem.

(d) Since our problem is a convex optimization problem with differential objective and constraint functions, the KKT conditions provide sufficient conditions for optimality. Hence, we know that if we can find $\vec{x}^*$ and $(\vec{\lambda}^*, \nu^*)$ that satisfy the KKT conditions, then $\vec{x}^*$ will be a primal optimal point, $(\vec{\lambda}^*, \nu^*)$ will be dual optimal. We therefore attempt to find solutions for the KKT conditions. As a first step, show how to simplify the KKT conditions so that they are expressed in terms of only $\vec{x}^*$ and ν^*, i.e. show how $\vec{\lambda}^*$ can be eliminated from these conditions.

(e) Solve for $x_i^*, 1 \leq i \leq n$, in terms of ν^* from the simplified KKT conditions derived in the previous subpart.

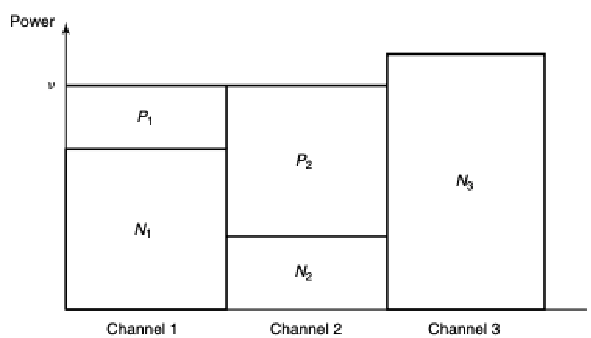

Figure 1: This graphic depicts a solution to the water-filling problem. On the x-axis we have n communication channels and on the y-axis we have the power in each channel. There is a base amount of noise N_i, which for us corresponds to α_i. Water-filling tells us that we should fill each channel until $\frac{1}{\nu^*}$, adding $\frac{1}{\nu^*} - \alpha_i$ power (in this graphic written as P_i), unless α_i already exceeds $\frac{1}{\nu^*}$. One algorithm for achieving this is to allot power to the channel with the least noise until it matches the channel with the second-least noise. Then we fill both simultaneously until they match the level of the channel with the third-least noise. Repeating this process until we run out of power to allot. This distribution of power is akin to filling connected basins with water, hence the name 'water filling'. Figure taken from *Elements of Information Theory* by Cover and Thomas.

5. Linear Programs

Consider the following linear program:

$$\max_{x_1, x_2 \in \mathbb{R}} \quad 2x_1 + 3x_2 \tag{13}$$
$$\text{s.t.} \quad x_1 + 2x_2 \leq 8 \tag{14}$$
$$x_1 - x_2 \leq 2 \tag{15}$$
$$x_2 + x_1 \geq 2 \tag{16}$$
$$x_1 \geq 0 \tag{17}$$
$$x_2 \geq 0 \tag{18}$$

(a) Sketch the feasible region of the linear program as well as the 5-, 10-, and 15-level sets of the objective function.

HINT: Recall that for $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $\alpha \in \mathbb{R}$, the α-level set of f is defined as $\{\vec{x} \in \mathbb{R}^n : f(\vec{x}) = \alpha\}$.

(b) Express the linear program in the following form:

$$\max_{\vec{x} \in \mathbb{R}^n} \quad \vec{c}^\top \vec{x} \tag{19}$$
$$\text{s.t.} \quad A\vec{x} \leq \vec{b} \tag{20}$$

Specify the values of $\vec{x}$, $\vec{c}$, A, and $\vec{b}$.

(c) Express the linear program in the following form:

$$\min_{\vec{y} \in \mathbb{R}^n} \quad \vec{c}^\top \vec{y} \tag{21}$$
$$\text{s.t.} \quad A\vec{y} = \vec{b} \tag{22}$$
$$\vec{y} \geq 0 \tag{23}$$

Specify the values of $\vec{y}$, $\vec{c}$, A, and $\vec{b}$.

HINT: Consider adding additional slack variables to the optimization problem.

(d) List the extreme points or vertices of the feasible region of the linear program given by equations (13)–(18).

(e) Find the optimal value $p^\star$ and the optimal point $\vec{x}^\star = (x_1^\star, x_2^\star)$ of the linear program given by equations (13)–(18).

HINT: Recall that for a linear program with a bounded feasible region, at least one optimal point is a vertex of the feasible region.

6. Homework Process

With whom did you work on this homework? List the names and SIDs of your group members.

NOTE: If you didn't work with anyone, you can put "none" as your answer.

