EECS 127/227AT   Optimization Models in Engineering   UC Berkeley   Spring 2023

Homework 10

This homework is due at 11 PM on April 6, 2023.

Submission Format: Your homework submission should consist of a single PDF file that contains all of your answers (any handwritten answers should be scanned).

1. **Does strong duality hold?**

   Consider

$$\min_{(x,y)\in\mathcal{D}} e^{-x} \tag{1}$$

$$\text{s.t. } x^2/y \leq 0 \tag{2}$$

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:

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

Explain why $g(\lambda)$ is lower bounded by 0 for $\lambda \geq 0$. **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 $\lambda \geq 0$. *HINT: To show that the infimum in Equation (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. LP at Boundary

Consider the LP:

$$\min_{x\in\mathbb{R}^n} \quad \vec{c}^\top \vec{x} \tag{4}$$
$$\text{s.t.} \quad A\vec{x} \le \vec{b}. \tag{5}$$

Here we have non-zero $\vec{c} \in \mathbb{R}^n$, $\vec{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, and $\vec{b} \in \mathbb{R}^m$. The feasible set forms a polyhedron

$$\mathcal{P} = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}_i^\top \vec{x} - b_i \le 0,\ 1 \le i \le n\} \tag{6}$$
$$= \bigcap_{i=1}^{m} \{\vec{x} \in \mathbb{R}^n \mid \vec{a}_i^\top \vec{x} - b_i \le 0\}. \tag{7}$$

In the second line, $\mathcal{P}$ is defined as the intersection of half-spaces. When this set is bounded, it is often referred to as a polytope instead.

In this problem we will prove facts about linear programs, including the crucial fact that the solution to an LP exists at the boundary of a polytope.

(a) First, consider the case where we have an unbounded polyhedron. For any such $\mathcal{P}$, assume there exists some $\vec{x}_0, \vec{v} \in \mathbb{R}^n$ with $\vec{x}_0 \neq \vec{v}$ such that $\mathcal{L} := \{\vec{x}_0 + \alpha\vec{v} \mid \alpha \in [0,\infty)\} \subset \mathcal{P}$. This is saying that there exists a line segment, unbounded on one side, that is contained within $\mathcal{P}$. Show that if $\vec{c}^\top \vec{v} < 0$, then $p^* = -\infty$.

(b) Now, suppose our feasible set is defined by a bounded polytope (i.e. a polytope that contains its boundary). We say a point belongs to the interior of $\mathcal{P}$ (i.e. $\vec{x}_0 \in \text{Int}(\mathcal{P})$) if there exists some ball of radius $\epsilon > 0$ such that $\mathcal{B} := \{\vec{y} \in \mathbb{R}^n \mid \|\vec{y} - \vec{x}_0\|_2 \le \epsilon\} \subset \mathcal{P}$. Show that the optimal point for the LP cannot be obtained in the interior of $\mathcal{P}$ and must be obtained on the boundary (that is, when for some $1 \le i \le n$, $\vec{a}_i^\top \vec{x} - b_i = 0$). For this, consider a proof by contradiction. Show that for any x_0 on the interior, there exists another point

$$\vec{x}_1 := \vec{x}_0 - \epsilon\frac{\vec{c}}{\|\vec{c}\|_2}, \tag{8}$$

such that $\vec{x}_1 \in \mathcal{P}$ and $\vec{c}^\top \vec{x}_1 < \vec{c}^\top \vec{x}_0$.

3. Sensitivity and Dual Variables

In this problem, we look into the interpretation of dual variables as sensitivity parameters of the primal problem. Recall the canonical, convex primal problem

$$\min_{\vec{x}} \quad f_0(\vec{x}) \tag{9}$$

$$\text{subject to} \quad f_i(\vec{x}) \le 0, \quad i = 1, \ldots, m \tag{10}$$

$$h_j(\vec{x}) = 0, \quad j = 1, \ldots, p \tag{11}$$

where f_0, f_i are convex and h_j are affine (assume the problem has strong duality). Consider the *perturbed* problem

$$\min_{\vec{x}} \quad f_0(\vec{x}) \tag{12}$$

$$\text{subject to} \quad f_i(\vec{x}) \le u_i, \quad i = 1, \ldots, m \tag{13}$$

$$h_j(\vec{x}) = v_j, \quad j = 1, \ldots, p \tag{14}$$

and define

$$p^\star(\vec{u}, \vec{v}) = \min\{f_0(\vec{x}) \mid \vec{x} \text{ such that } f_i(\vec{x}) \le u_i \ \ \forall i, h_j(\vec{x}) = v_j \ \ \forall j\} \tag{15}$$

In words, $p^\star(\vec{u}, \vec{v})$ is the optimal value for the perturbed problem if it is feasible, and defined to be $+\infty$ (infeasible) otherwise. Note $p^\star(\vec{0}, \vec{0})$ is the original problem.

(a) Prove $p^\star(\vec{u}, \vec{v})$ is jointly convex in $(\vec{u}, \vec{v})$.

HINT: Consider $\mathcal{D} = \{(\vec{x}, \vec{u}, \vec{v}) \mid f_i(\vec{x}) \le u_i \ \ \forall i, h_j(\vec{x}) = v_j \ \ \forall j\}$, which is the set of triples $(\vec{x}, \vec{u}, \vec{v})$ such that $\vec{x}$ is a feasible point for the perturbed problem with the perturbations $(\vec{u}, \vec{v})$. Is $\mathcal{D}$ convex? Also, define $F(\vec{x}, \vec{u}, \vec{v})$ to be a function that is equal to $f_0(\vec{x})$ on $\mathcal{D}$ and $+\infty$ otherwise.

HINT: Prove and use the following lemma:

Let S_1, S_2 be convex sets with a function $f : S_1 \times S_2 \to \mathbb{R}$ which is jointly convex in both arguments. Define $g(x) = \min_{y \in S_2} f(x, y)$. Then $g(x)$ is convex in $x \in S_1$.

(b) Assume that strong duality holds, and that the dual optimum is attained. Let $(\vec{\lambda}^\star, \vec{\nu}^\star)$ be the optimal dual variables for the dual of the unperturbed primal problem (9). Show that for any point $\hat{\vec{x}}$ that is feasible for the perturbed problem, we have

$$f_0(\hat{\vec{x}}) \ge p^\star(0, 0) - \vec{u}^\top \vec{\lambda}^\star - \vec{v}^\top \vec{\nu}^\star \tag{16}$$

HINT: Consider the Lagrangian for the original problem $L(\vec{x}, \vec{\lambda}, \vec{\nu})$ and $g(\vec{\lambda}, \vec{\nu}) = \min_{\vec{x}} L(\vec{x}, \vec{\lambda}, \vec{\nu})$. Now relate this to the perturbed problem.

(c) Using the previous subpart, show that for all $\vec{u}, \vec{v}$, we have

$$p^\star(\vec{u}, \vec{v}) \ge p^\star(\vec{0}, \vec{0}) - \vec{u}^\top \vec{\lambda}^\star - \vec{v}^\top \vec{\nu}^\star. \tag{17}$$

(d) Suppose we only have 1 equality and 1 inequality constraint (that is u, v are scalars). For each of the following situations, argue whether the value of $p^\star(u, v)$ increases or decreases as compared to $p^\star(0, 0)$ or whether we are unsure as to whether it increases or decreases.

i. If $\lambda^\star$ is large and we pick $u < 0$.

ii. If $\lambda^\star$ is large and we pick $u > 0$.

iii. If $\nu^\star$ is large and positive (resp. negative) and we pick $v < 0$ (resp. $v > 0$).

4. KKT with circles

Consider the problem

$$\min_{\vec{x}\in\mathbb{R}^2} \quad x_1^2 + x_2^2 \tag{18}$$

$$\text{s.t.} \quad (x_1 - 1)^2 + (x_2 - 1)^2 \leq 4 \tag{19}$$

$$(x_1 - 1)^2 + (x_2 + 1)^2 \leq 4 \tag{20}$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in \mathbb{R}^2$.

(a) Sketch the feasible region and the level sets of the objective function. Find the optimal point $\vec{x}^*$ and the optimal value p^*.

(b) Does strong duality hold?

(c) Write the KKT conditions for this optimization problem. Do there exist Lagrange multipliers λ_1^* and λ_2^* that prove the optimality of $\vec{x}^*$?

5. Water Filling

Consider the following problem:

$$\begin{aligned} \text{minimize} \quad & -\sum_{i=1}^{n} \log(\alpha_i + x_i) && (21) \\ \text{subject to} \quad & \vec{x} \geq 0,\ \vec{1}^\top \vec{x} = c, && (22) \end{aligned}$$

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 $\lambda \in \mathbb{R}^n$ 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 (λ^*, ν^*) that verify the KKT conditions, then $\vec{x}^*$ will be a primal optimal point, (λ^*, ν^*) 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. we show how λ^* 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 preceding part of this question.

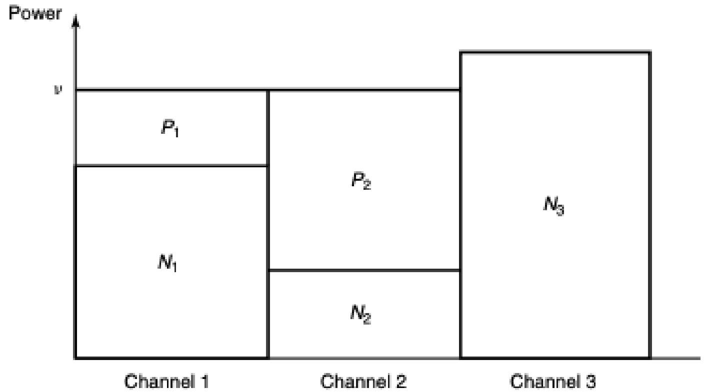

Figure 1: This graphic depicts a solution to the water-filling problem. On the x-axis we have n communication channels and on the y-axis we have the power in each channel. There is a base amount of noise N_i, which for us corresponds to α_i. Water-filling tells us that we should fill each channel until $\frac{1}{\nu^*}$, 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.

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.

