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

Homework 11

This homework is due at 11 PM on April 13, 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. **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:

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

   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, 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 a linear program (minimizing a linear objective subject to linear constraints), provided one allows for infinitely many 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

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

$$\text{such that } \ |x| \leq 1 \tag{3}$$

   (e) Now, suppose for the formulation in (1), $f_0(\vec{x})$ is a convex function, $\mathcal{X}$ is a convex set, and all $f_i(\vec{x})$ are 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

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

   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 (1) at some point $\vec{\bar{x}}$ that satisfies $f_i(\vec{\bar{x}}) > 0$, then these exists a point $\vec{y}$ between $\vec{\bar{x}}$ and $\vec{x}^\star$ that is feasible to the original problem in (1). 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 (1).

2. Fun with Hyperplanes

In this problem we work with hyperplanes, which are key components of linear programming as well as future topics such as support vector machines.

(a) Sketch the hyperplane $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^2 \mid \begin{bmatrix} 1 & 1 \end{bmatrix} \vec{x} = 2\}$.

(b) Let $\vec{c} \in \mathbb{R}^n$ be nonzero, and let $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top \vec{x} = 0\}$. Show that $\mathcal{H}$ is a linear subspace of $\mathbb{R}^n$. What is $\dim(\mathcal{H})$?

(c) Let $\vec{c} \in \mathbb{R}^n$ be nonzero, and let $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top \vec{x} = 0\}$. Suppose $\vec{x}_\star \in \mathbb{R}^n$ is on one side of the hyperplane, i.e., $\vec{c}^\top \vec{x}_\star > 0$. Give any vector which is on the other side of the hyperplane but not on the hyperplane itself.

(d) Let $\vec{c} \in \mathbb{R}^n$ be nonzero, and let $\vec{x}_0 \in \mathbb{R}^n$ be arbitrary. Let $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top (\vec{x} - \vec{x}_0) = 0\}$. Suppose $\vec{x}_\star \in \mathbb{R}^n$ is on one side of the hyperplane. Give any vector which is on the other side of the hyperplane but not on the hyperplane itself.

(e) Let $\vec{x}_0 \in \mathbb{R}^n$ be arbitrary. For a vector $\vec{c} \in \mathbb{R}^n$, let $\mathcal{H}(\vec{c}) \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top (\vec{x} - \vec{x}_0) = 0\}$. Show that $\vec{0} \in \mathcal{H}(\vec{c})$ for every $\vec{c} \in \mathbb{R}^n$ if and only if $\vec{x}_0 = \vec{0}$.

3. Duality

Consider the function

$$f(\vec{x}) = \vec{x}^\top A\vec{x} - 2\vec{b}^\top \vec{x}. \tag{5}$$

First, we consider the unconstrained optimization problem

$$p^\star = \min_{\vec{x}\in\mathbb{R}^n} f(\vec{x}) = \min_{\vec{x}\in\mathbb{R}^n} \vec{x}^\top A\vec{x} - 2\vec{b}^\top \vec{x} \tag{6}$$

for a real $n \times n$ symmetric matrix $A \in \mathbb{S}^n$ and $\vec{b} \in \mathbb{R}^n$. If the problem is unbounded below, then we say $p^\star = -\infty$. Let $\vec{x}^\star$ denote the minimizing argument of the optimization problem.

(a) Suppose $A \succeq 0$ (positive semidefinite) and $\vec{b} \in \mathcal{R}(A)$. Let $\operatorname{rank}(A) = n$. Find $p^\star$.

HINT: What does $A \succeq 0$ tell you about the function f? How can you leverage the rank of A to compute $p^\star$?

(b) Suppose $A \succeq 0$ (positive semidefinite) and $\vec{b} \in \mathcal{R}(A)$ as before. Let A be rank-deficient, i.e., $\operatorname{rank}(A) = r < n$. Let A have the compact/thin and full SVD as follows, with diagonal positive definite $\Lambda_r \in \mathbb{R}^{r\times r}$:

$$A = U_r\Lambda_r U_r^\top = \begin{bmatrix} U_r & U_1 \end{bmatrix} \begin{bmatrix} \Lambda_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_r^\top \\ U_1^\top \end{bmatrix}. \tag{7}$$

Show that the minimizer $\vec{x}^\star$ of the optimization problem (6) is not unique by finding a general form for the family of solutions for $\vec{x}^\star$ in terms of $U_r, U_1, \Lambda_r, \vec{b}$.

HINT: As before, $A \succeq 0$ gives you some information about the objective function f. Can you use this information along with the fact that $b \in \mathcal{R}(A)$ to obtain a general form for the minimizers of f? Use the fact that any vector $\vec{x} \in \mathbb{R}^n$ can be written as $\vec{x} = U_r\vec{\alpha} + U_1\vec{\beta}$ for unique $\vec{\alpha}, \vec{\beta}$.

(c) If $A \not\succeq 0$ (A not positive semi-definite) show that $p^\star = -\infty$ by finding $\vec{v}$ such that $f(\alpha\vec{v}) \to -\infty$ as $\alpha \to \infty$.

HINT: $A \not\succeq 0$ means that there exists $\vec{v}$ such that $\vec{v}^\top A\vec{v} < 0$.

(d) Suppose $A \succeq 0$ (positive semidefinite) and $\vec{b} \notin \mathcal{R}(A)$. Find $p^\star$. Justify your answer mathematically.

HINT: From FTLA, we know that $\mathbb{R}^n = \mathcal{R}(A^\top) \oplus \mathcal{N}(A)$. Therefore, $\vec{b} = \vec{v}_1 + \vec{v}_2$ where $\vec{v}_1 \in \mathcal{R}(A) = \mathcal{R}(A^\top)$ and $\vec{v}_2 \in \mathcal{N}(A)$.

For parts (e) and (f), consider real $n \times n$ symmetric matrix $A \in \mathbb{S}^n$ and $\vec{b} \in \mathbb{R}^n$. Let $\operatorname{rank}(A) = r$, where $0 \le r \le n$. Now we consider the constrained optimization problem

$$\begin{aligned} p^\star = \min_{\vec{x}\in\mathbb{R}^n} \ & \vec{x}^\top A\vec{x} - 2\vec{b}^\top \vec{x} \\ \text{s.t. } & \vec{x}^\top \vec{x} \ge 1. \end{aligned} \tag{8}$$

(e) Write the Lagrangian $\mathcal{L}(\vec{x}, \lambda)$, where λ is the dual variable corresponding to the inequality constraint.

(f) For any matrix $C \in \mathbb{R}^{n\times n}$ with $\operatorname{rank}(C) = r \le n$ and compact SVD

$$C = U_r\Lambda_r V_r^\top, \tag{9}$$

we define the pseudoinverse as

$$C^\dagger = V_r\Lambda_r^{-1}U_r^\top. \tag{10}$$

We use the "dagger" operator to represent this. If $\vec{d}$ lies in the range of C, then a solution to the equation $C\vec{x} = \vec{d}$, can be written as $\vec{x} = C^{\dagger}\vec{d}$, even when C is not full rank. Show that the dual problem to the primal problem (8) can be written as,

$$d^{\star} = \max_{\substack{\lambda \geq 0 \\ A - \lambda I \succeq 0 \\ \vec{b} \in \mathcal{R}(A - \lambda I)}} -\vec{b}^{\top} (A - \lambda I)^{\dagger} \vec{b} + \lambda. \tag{11}$$

HINT: To show this, first argue that when the constraints are not satisfied $\min_{\vec{x}} \mathcal{L}(\vec{x}, \lambda) = -\infty$. *Then show that when the constraints are satisfied,* $\min_{\vec{x}} \mathcal{L}(\vec{x}, \lambda) = -\vec{b}^{\top} (A - \lambda I)^{\dagger} \vec{b} + \lambda$.

HINT: Compute $g(\lambda)$ *and explore its behavior under the constraints.*

4. A Slalom Problem

A skier must slide from left to right by going through n parallel gates of known position (x_i, y_i) and width c_i, $i = 1, \ldots, n$. The initial position (x_0, y_0) is given, as well as the final one, (x_{n+1}, y_{n+1}). Before reaching the final position, the skier must go through gate i by passing between the points $(x_i, y_i - c_i/2)$ and $(x_i, y_i + c_i/2)$ for each $i \in \{1, \ldots, n\}$.

Figure 1 is an example and does not have the right value of n nor show the true (x_i, y_i, c_i) values. Use values for (x_i, y_i, c_i) from Table 1.

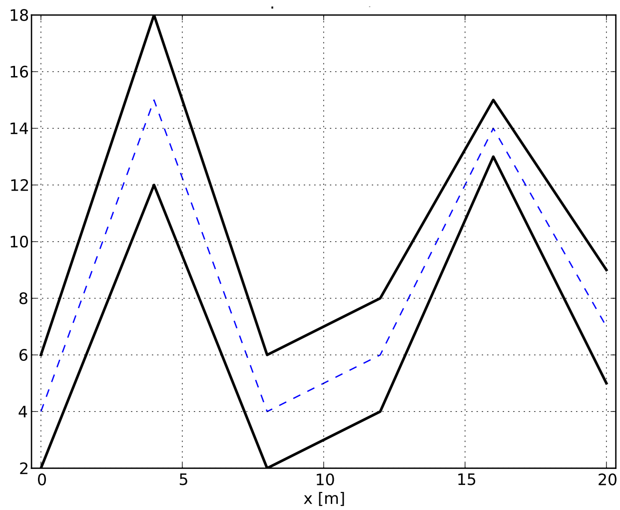

Figure 1: Slalom problem with $n = 6$ gates. The initial and final positions are fixed and not included in the figure. The skier slides from left to right. The middle path is dashed and connects the center points of gates.

Table 1: Problem data for Problem 2. Here $n = 5$.

i	x_i	y_i	c_i
0	0	4	N/A
1	4	5	3
2	8	4	2
3	12	6	2
4	16	8	1
5	20	7	2
6	24	4	N/A

(a) Given the data $\{(x_i, y_i, c_i)\}_{i=0}^{n+1}$, write an optimization problem that minimizes the total length of the path. Your answer should come in the form of an SOCP.

(b) Solve the problem numerically with the data given in Table 1. *HINT: You should be able to use packages such as* `cvxpy` *and* `numpy`*.*

5. 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

$$p^\star = \min_{\vec{h}, b, z} \quad f_0(\vec{h}, b, z) \tag{12}$$

$$s.t. \quad \vec{h}^\top \vec{x}_i + b \leq 0 \qquad \forall i : y_i = 1 \tag{13}$$

$$\vec{h}^\top \vec{x}_i + b \geq z \qquad \forall i : y_i = -1 \tag{14}$$

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.

6. Dual Norms and SOCP

Consider the problem

$$p^\star = \min_{\vec{x} \in \mathbb{R}^n} \|A\vec{x} - \vec{y}\|_1 + \mu \|\vec{x}\|_2 \,, \tag{15}$$

where $A \in \mathbb{R}^{m \times n}$, $\vec{y} \in \mathbb{R}^m$, and $\mu > 0$.

(a) Express this (primal) problem in standard SOCP form.

(b) Find a dual to the problem and express it in standard SOCP form. *HINT: Recall that for every vector $\vec{z}$, the following dual norm equalities hold:*

$$\|\vec{z}\|_2 = \max_{\vec{u} \,:\, \|\vec{u}\|_2 \le 1} \vec{u}^\top \vec{z}, \qquad \|\vec{z}\|_1 = \max_{\vec{u} \,:\, \|\vec{u}\|_\infty \le 1} \vec{u}^\top \vec{z}. \tag{16}$$

(c) Assume strong duality holds[1] and that $m = 100$ and $n = 10^6$, i.e., A is 100×10^6. Which problem would you choose to solve using a numerical solver: the primal or the dual? Justify your answer.

[1] In fact, you can show that strong duality holds using Sion's theorem, a generalization of the minimax theorem that is beyond the scope of this class.

7. 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.

