Homework 11

Self grades are due at 11 PM on April 19, 2024.

1. **Project Logistics**

   Fill out this form to let us know whether:

   - you plan on doing the project in a group;
   - you plan on doing the project by yourself; or
   - you are not planning to do the project;

   as well as some auxiliary information. Even if you do not plan to do the project, you **must** fill out the form.

   To get credit for the problem, please write down the secret word in the confirmation message displayed after completing the survey.

2. Linear Programming

Express the following problem as an LP.

$$\min_{\vec{x}\in\mathbb{R}^k} \quad \sum_{i=1}^{k} |x_i| \tag{1}$$

$$\text{s.t.} \quad A\vec{x} = \vec{b} \tag{2}$$

Solution: We solve this problem using slack variables. First, consider the simple, related scalar minimization

$$\min_{x\in\mathbb{R}} |x|. \tag{3}$$

We can reformulate this problem using slack variable t as follows:

$$\min_{x,t\in\mathbb{R}} \quad t \tag{4}$$

$$\text{s.t.} \quad |x| \leq t \tag{5}$$

and further linearize the constraint to get LP

$$\min_{x,t\in\mathbb{R}} \quad t \tag{6}$$

$$\text{s.t.} \quad x \leq t,\ -x \leq t. \tag{7}$$

These constraints are equivalent because if $x \geq 0$ then $|x| \leq t$ if and only if $x \leq t$. And if $x < 0$ then $|x| \leq t$ if and only if $-x \leq t$. Further, when $x \geq 0$, if $x \leq t$ then $-x \leq t$ always; and when $x < 0$, if $-x \leq t$ then $x \leq t$ always. Returning to our original minimization, we can replace each absolute value in the summation with a slack variable, resulting in final LP

$$\min_{\vec{x},\vec{t}\in\mathbb{R}^k} \quad \sum_{i=1}^{k} t_i \tag{8}$$

$$\text{s.t.} \quad A\vec{x} = \vec{b} \tag{9}$$

$$x_i \leq t_i,\ -x \leq t_i,\ i = 1, \ldots, k \tag{10}$$

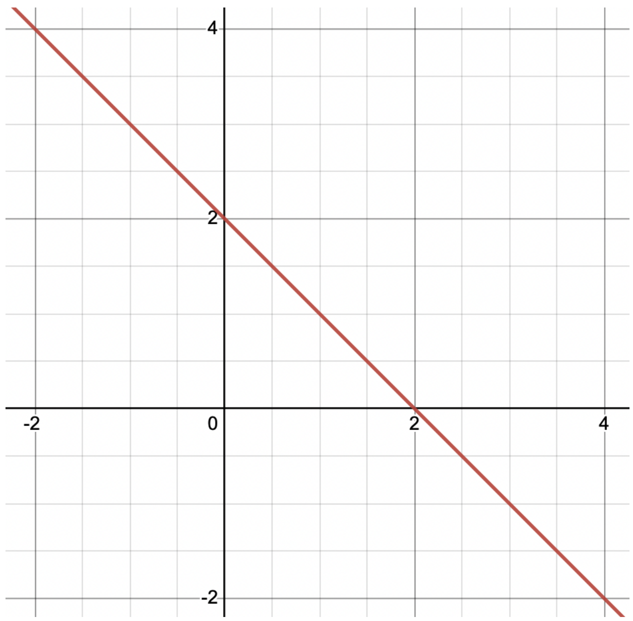
as desired.

3. 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 $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^2 \mid \begin{bmatrix} 1 & 1 \end{bmatrix} \vec{x} = 2\}$.

Solution: See the following figure:



(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 $\dim(\mathcal{H})$?

Solution: We have $\mathcal{H} = \mathcal{N}(\vec{c}^\top)$, where $\vec{c}^\top$ is interpreted as a $1 \times n$ matrix. Thus it is a linear subspace. By rank-nullity, we have $\dim(\mathcal{R}(\vec{c}^\top)) + \dim(\mathcal{N}(\vec{c}^\top)) = n$, and $\dim(\mathcal{R}(\vec{c}^\top)) = \dim(\mathcal{R}(\vec{c})) = 1$, so $\dim(\mathcal{H}) = \dim(\mathcal{N}(\vec{c}^\top)) = n - 1$.

Alternatively, you could show that $\mathcal{H}$ is closed under linear combination.

(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.

Solution: We propose the vector $-\vec{x}_\star$. Indeed, we have

$$\vec{c}^\top(-\vec{x}_\star) = -\vec{c}^\top \vec{x}_\star < 0. \tag{11}$$

Thus $-\vec{x}_\star$ is on the other side of the plane.

(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.

Solution: Suppose that without loss of generality we have $\vec{c}^\top(\vec{x}_\star - \vec{x}_0) > 0$. Then we have

$$\vec{c}^\top(\vec{x}_\star - \vec{x}_0) > 0 \implies \vec{c}^\top(\vec{x}_0 - \vec{x}_\star) < 0. \tag{12}$$

We want to find $\vec{z}$ such that $\vec{z} - \vec{x}_0 = \vec{x}_0 - \vec{x}_\star$. By algebra, $\vec{z} = 2\vec{x}_0 - \vec{x}_\star$. This gives $\vec{c}^\top(\vec{z} - \vec{x}_0) = \vec{c}^\top(\vec{x}_0 - \vec{x}_\star) < 0$, so $\vec{z}$ is the vector we want.

(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 $\vec{x}_0 = \vec{0}$.

Solution: We first claim that, for a fixed $\vec{c} \in \mathbb{R}^n$, that $\vec{0} \in \mathcal{H}(\vec{c})$ if and only if $\vec{c}$ is orthogonal to $\vec{x}_0$. Indeed,

$$\begin{aligned}
\vec{c}^\top \vec{x}_0 = 0 &\iff -\vec{c}^\top \vec{x}_0 = 0 && (13)\\
&\iff \vec{c}^\top \vec{0} - \vec{c}^\top \vec{x}_0 = 0 && (14)\\
&\iff \vec{c}^\top (\vec{0} - \vec{x}_0) = 0 && (15)\\
&\iff \vec{0} \in \mathcal{H}(\vec{c}). && (16)
\end{aligned}$$

Thus $\vec{0} \in \mathcal{H}(\vec{c})$ for every $\vec{c} \in \mathbb{R}^n$ if and only if $\vec{x}_0$ is orthogonal to every $\vec{c} \in \mathbb{R}^n$. But this is equivalent to $\vec{x}_0 = \vec{0}$, and the claim is proved.

4. Formulating problems as LPs or QPs

Formulate the problem

$$p_j^* \doteq \min_{\vec{x}} \; f_j(\vec{x}), \tag{17}$$

for different functions f_j, $j = 1, \ldots, 4$, as convex **QPs** or **LPs**, or, if you cannot, explain why. In our formulations, we always use $\vec{x} \in \mathbb{R}^n$ as the variable, and assume that $A \in \mathbb{R}^{m \times n}$, $\vec{y} \in \mathbb{R}^m$. If you obtain a convex **LP** or **QP** formulation, state precisely what the variables, objective, and constraints are.

(a) $f_1(\vec{x}) = \|A\vec{x} - \vec{y}\|_\infty + \|\vec{x}\|_1$.

Solution: For p_1^*, we replace the ℓ_∞ norm with constraints on the maximum absolute value of each element of $Ax - y$, and rewrite absolute values as linear constraints. In addition, we introduce slack variables $\vec{z} \in \mathbb{R}^n$. This gives us the convex **LP** formulation:

$$\begin{aligned} p_1^* = \min_{\vec{x},t,\vec{z}} \quad & t + \vec{1}^\top \vec{z} && (18)\\ \text{s.t.} \quad & z_i \geq x_i \geq -z_i, \quad i = 1, \ldots, n && (19)\\ & t \geq (A\vec{x} - \vec{y})_i \geq -t, \quad i = 1, \ldots, m. && (20) \end{aligned}$$

(b) $f_2(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 + \|\vec{x}\|_1$.

Solution: For p_2^*, we follow a similar approach to the previous part, except we do not need to worry about replacing the ℓ_2 norm. Hence we obtain the convex **QP**

$$\begin{aligned} p_2^* = \min_{\vec{x},\vec{z}} \quad & \vec{x}^\top (A^\top A)\vec{x} - 2\vec{y}^\top A\vec{x} + \vec{y}^\top \vec{y} + \vec{1}^\top \vec{z} && (21)\\ \text{s.t.} \quad & z_i \geq x_i \geq -z_i, \quad i = 1, \ldots, n. && (22) \end{aligned}$$

(c) $f_3(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 - \|\vec{x}\|_1$.

Solution: For p_3^*, the problem is not convex. Consider for instance the special case with $n = 1$, $A = 1$, $y = 0$: plot $f_3(x) = x^2 - |x|$ to verify it is not convex. In general, the objective of the *difference* of two convex functions is not necessarily convex.

(d) $f_4(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 + \|\vec{x}\|_1^2$.

Solution: For p_4^*, we have the convex **QP**

$$\begin{aligned} p_4^* = \min_{\vec{x},\vec{z}} \quad & \vec{x}^\top (A^\top A)\vec{x} - 2\vec{y}^\top A\vec{x} + \vec{y}^\top \vec{y} + \left(\sum_{i=1}^n z_i\right)^2 && (23)\\ \text{s.t.} \quad & z_i \geq x_i \geq -z_i, \quad i = 1, \ldots, n. && (24) \end{aligned}$$

Notice that $\left(\sum_{i=1}^n z_i\right)^2 = \vec{z}^\top Q\vec{z}$, where we define Q as an $n \times n$ matrix of all ones. Thus, our problem can be written as:

$$\begin{aligned} p_4^* = \min_{\vec{x},\vec{z}} \quad & \vec{x}^\top (A^\top A)\vec{x} - 2\vec{y}^\top A\vec{x} + \vec{y}^\top \vec{y} + \vec{z}^\top Q\vec{z} && (25)\\ \text{s.t.} \quad & z_i \geq x_i \geq -z_i, \quad i = 1, \ldots, n. && (26) \end{aligned}$$

5. 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.

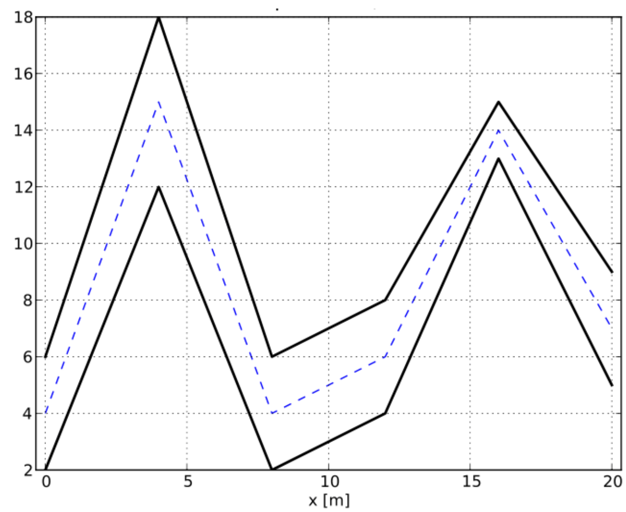


Figure 1: Slalom problem with $n = 6$ gates. The initial and final positions are fixed and not included in the figure. The skier slides from left to right. The middle path is dashed and connects the center points of gates.

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.

Solution: Assume that (x_i, z_i) is the crossing point of gate i, the path length minimization problem is thus

$$\min_{\vec{z}} \quad \sum_{i=1}^{n+1} \left\| \begin{bmatrix} x_i \\ z_i \end{bmatrix} - \begin{bmatrix} x_{i-1} \\ z_{i-1} \end{bmatrix} \right\|_2 \tag{27}$$

$$\text{s.t.} \quad y_i - c_i/2 \leq z_i \leq y_i + c_i/2, \text{ for } i = 1, \ldots, n \tag{28}$$

$$z_0 = y_0, z_{n+1} = y_{n+1}, \tag{29}$$

which is equivalent to

$$\min_{\vec{z}, \vec{t}} \quad \sum_{i=1}^{n+1} t_i \tag{30}$$

$$\text{s.t.} \quad y_i - c_i/2 \leq z_i \leq y_i + c_i/2, \text{ for } i = 0, \ldots, n+1 \tag{31}$$

$$\left\| \begin{bmatrix} x_i \\ z_i \end{bmatrix} - \begin{bmatrix} x_{i-1} \\ z_{i-1} \end{bmatrix} \right\|_2 \leq t_i, \text{ for } i = 1, \ldots, n+1. \tag{32}$$

with the convention $c_0 = c_{n+1} = 0$. Hence, the problem is 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`.*

Solution: The code can be found in the corresponding Jupyter notebook.

6. Dual Norms and SOCP

Consider the problem

$$p^\star = \min_{\vec{x}\in\mathbb{R}^n} \|A\vec{x}-\vec{y}\|_1 + \mu \|\vec{x}\|_2\,, \tag{33}$$

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. *HINT: you can make the objective function linear by introducing slack variables.*

Solution: Introducing slack variables $\vec{z} \in \mathbb{R}^m$, $t \in \mathbb{R}$, we can write

$$\min_{\vec{x},\vec{z},t} \vec{z}^\top \vec{1} + \mu t \tag{34}$$
$$\text{s.t. } |(A\vec{x}-\vec{y})_i| \le z_i, \quad i = 1,\dots,m \tag{35}$$
$$\|\vec{x}\|_2 \le t, \tag{36}$$

which can be written as

$$\min_{\vec{x},\vec{z},t} \vec{z}^\top \vec{1} + \mu t \tag{37}$$
$$\text{s.t. } \left|\vec{a}_i^\top \vec{x} - y_i\right| \le z_i, \quad i = 1,\dots,m \tag{38}$$
$$\|\vec{x}\|_2 \le t, \tag{39}$$

where $\vec{a}_i$ are the rows of A and y_i are the entries of $\vec{y}$.

This expression now satisfies our definition of an SOCP: the objective is linear, and all constraints are SOC constraints, of the form

$$\left\| \begin{bmatrix} A_i & B_i & C_i \end{bmatrix} \begin{bmatrix} \vec{x} \\ \vec{z} \\ t \end{bmatrix} - \vec{d}_i \right\|_2 \le \begin{bmatrix} \vec{f}_i \\ \vec{g}_i \\ h_i \end{bmatrix}^\top \begin{bmatrix} \vec{x} \\ \vec{z} \\ t \end{bmatrix} + k_i. \tag{40}$$

For the first m constraints, we just take $A_i = \vec{a}_i^\top$, $B_i = 0$, $C_i = 0$, $\vec{d}_i = y_i$, $\vec{f}_i = \vec{0}$, $\vec{g}_i = \vec{e}_i$ (i.e., the i^{th} standard basis vector), $h_i = 0$, and $k_i = 0$. For the last constraint, we take $A_i = I$, $B_i = 0$, $C_i = 0$, $\vec{d}_i = \vec{0}$, $\vec{f}_i = \vec{0}$, $\vec{g}_i = \vec{0}$, $h_i = 1$, and $k_i = 0$.

(b) Find a dual to the problem and express it in standard SOCP form. *HINT: You may find Problem 1 in Discussion 11 to be helpful. Recall that for every vector $\vec{z}$, the following dual norm equalities hold:*

$$\|\vec{z}\|_2 = \max_{\vec{u}\,:\,\|\vec{u}\|_2\le 1} \vec{u}^\top \vec{z}, \qquad \|\vec{z}\|_1 = \max_{\vec{u}\,:\,\|\vec{u}\|_\infty\le 1} \vec{u}^\top \vec{z}. \tag{41}$$

Additionally, you may use the following fact: If the primal problem is expressed as $p^\star = \min_{\vec{x}} \max_{\vec{y}} L(\vec{x},\vec{y})$ (which is $p^\star = \min_{\vec{x}} f_0(x)$ for $f_0(x) = \max_{\vec{y}} L(\vec{x},\vec{y})$), then the dual problem can be obtained by swapping the max and min: $d^\star = \max_{\vec{y}} \min_{\vec{x}} L(\vec{x},\vec{y})$.

Solution: Using the hint, we can rewrite the objective function of the original problem as

$$\|A\vec{x}-\vec{y}\|_1 + \mu\|\vec{x}\|_2 = \max_{\vec{u}:\|\vec{u}\|_\infty\le 1} \vec{u}^\top (A\vec{x}-\vec{y}) + \mu \max_{\vec{v}:\|\vec{v}\|_2\le 1} \vec{v}^\top \vec{x}. \tag{42}$$

We can then express the original (primal) problem as

$$p^\star = \min_{\vec{x}} \max_{\vec{u},\vec{v}:\|\vec{u}\|_\infty\le 1, \|\vec{v}\|_2\le 1} \vec{u}^\top (A\vec{x}-\vec{y}) + \mu \vec{v}^\top \vec{x}. \tag{43}$$

To form the dual, we reverse the order of min and max:

$$d^\star = \max_{\vec{u},\vec{v}:\|\vec{u}\|_\infty \leq 1, \|\vec{v}\|_2 \leq 1} \min_{\vec{x}} \vec{u}^\top (A\vec{x} - \vec{y}) + \mu \vec{v}^\top \vec{x} \tag{44}$$

$$\doteq \max_{\vec{u},\vec{v}:\|\vec{u}\|_\infty \leq 1, \|\vec{v}\|_2 \leq 1} g(\vec{u}, \vec{v}), \tag{45}$$

where g is defined as

$$g(\vec{u}, \vec{v}) \doteq \min_{\vec{x}} \vec{u}^\top (A\vec{x} - \vec{y}) + \mu \vec{v}^\top \vec{x} \tag{46}$$

$$= \min_{\vec{x}} (\vec{u}^\top A + \mu \vec{v}^\top)\vec{x} - \vec{u}^\top \vec{y} \tag{47}$$

$$= \begin{cases} -\vec{u}^\top \vec{y} & \text{if } A^\top \vec{u} + \mu \vec{v} = \vec{0} \\ -\infty & \text{otherwise.} \end{cases} \tag{48}$$

We can thus rewrite the dual problem as

$$d^\star = \max_{\vec{u},\vec{v}} \quad -\vec{u}^\top \vec{y} \tag{49}$$

$$\text{s.t.} \quad A^\top \vec{u} + \mu \vec{v} = \vec{0} \tag{50}$$

$$\|\vec{u}\|_\infty \leq 1, \quad \|\vec{v}\|_2 \leq 1. \tag{51}$$

Noting that the first constraint fully restricts the value of $\vec{v}$ — rewriting it, $\vec{v} = -\frac{A^\top \vec{u}}{\mu}$ — we can plug this value into the third constraint and eliminate $\vec{v}$ from our optimization altogether:

$$d^\star = \max_{\vec{u}} \quad -\vec{u}^\top \vec{y} \tag{52}$$

$$\text{s.t.} \quad \left\|A^\top \vec{u}\right\|_2 \leq \mu \tag{53}$$

$$\|\vec{u}\|_\infty \leq 1, \tag{54}$$

generating our final SOCP dual. If desired, we can further rewrite the final constraint as $\|\vec{u}\|_\infty = \max_i |u_i| \leq 1 \Leftrightarrow |u_i| \leq 1, \ i = 1, \ldots, m \Leftrightarrow u_i \leq 1$ and $u_i \geq -1, \ i = 1, \ldots, m$ to make the linearity of that constraint more explicit.

(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.

Solution: To determine the rough computational complexity of each problem, we examine the number of variables and the number of constraints in each problem. The primal SOCP has $\sim 10^6$ variables and 201 constraints, while the dual has 100 variables and 201 constraints. The dual problem is thus much more efficient to solve.

[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.

