Self grades are due at 11 PM on November 18, 2022.

1. 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 \( \mathcal{H} = \{ \vec{x} \in \mathbb{R}^2 \mid [1 \ 1] \vec{x} = 2 \} \).

Solution: See the following figure:

![Hyperplane Sketch](image)

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

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

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

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

Solution: We propose the vector \( -\vec{x}_* \). Indeed, we have

\[
\vec{c}^T (-\vec{x}_*) = -\vec{c}^T \vec{x}_* < 0.
\]

Thus \( -\vec{x}_* \) 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 \( \mathcal{H} = \{ \vec{x} \in \mathbb{R}^n \mid \vec{c}^T (\vec{x} - \vec{x}_0) = 0 \} \). Suppose \( \vec{x}_* \in \mathbb{R}^n \) 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:** We propose the vector \( 2\vec{x}_0 - \vec{x}_* \). Indeed, we have
\[
\vec{c}^T ((2\vec{x}_0 - \vec{x}_*) - \vec{x}_0) = \vec{c}^T (\vec{x}_0 - \vec{x}_*)
\]
\[
= -\vec{c}^T (\vec{x}_* - \vec{x}_0)
\]
\[
< 0.
\]

(e) Let \( \vec{x}_0 \in \mathbb{R}^n \) be arbitrary. For a vector \( \vec{c} \in \mathbb{R}^n \), let \( \mathcal{H}(\vec{c}) = \{ \vec{x} \in \mathbb{R}^n \mid \vec{c}^T (\vec{x} - \vec{x}_0) = 0 \} \). 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,
\[
\vec{c}^T \vec{x}_0 = 0 \iff -\vec{c}^T \vec{x}_0 = 0
\]
\[
\iff \vec{c}^T \vec{0} - \vec{c}^T \vec{x}_0 = 0
\]
\[
\iff \vec{c}^T (\vec{0} - \vec{x}_0) = 0
\]
\[
\iff \vec{0} \in \mathcal{H}(\vec{c}).
\]

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.
2. LP at Boundary

Consider the LP:

$$\min_{\vec{x} \in \mathbb{R}^n} \vec{c}^T \vec{x} \quad (9)$$

$$\text{s.t.} \quad A\vec{x} \leq \vec{b}. \quad (10)$$

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

$$\mathcal{P} = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}_i^T \vec{x} - b_i \leq 0, \ 1 \leq i \leq n\} \quad (11)$$

$$= \bigcap_{i=1}^m \{\vec{x} \in \mathbb{R}^n \mid \vec{a}_i^T \vec{x} - b_i \leq 0\}. \quad (12)$$

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 $\vec{c}^T \vec{v} < 0$, then $p^* = -\infty$.

Solution: If $\vec{c}^T \vec{v} < 0$, then for any point $\vec{y}_\alpha = \vec{x}_0 + \alpha \vec{v} \in \mathcal{L}$, we have that

$$\vec{c}^T \vec{y}_\alpha = \vec{c}^T \vec{x}_0 + \alpha \vec{c}^T \vec{v}. \quad (13)$$

If we take $\alpha \to \infty$,

$$\lim_{\alpha \to \infty} \vec{c}^T \vec{y}_\alpha = \vec{c}^T \vec{x}_0 + \lim_{\alpha \to \infty} \alpha \vec{c}^T \vec{v} \quad (14)$$

$$= -\infty. \quad (15)$$

(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 $\mathcal{B} := \{\vec{y} \in \mathbb{R}^n \mid \|\vec{y} - \vec{x}_0\|_2 \leq \epsilon\} \subset \mathcal{P}$. 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 $1 \leq i \leq n$, $\vec{a}_i^T \vec{x} - b_i = 0$). For this, consider a proof by contradiction. Show that for any $\vec{x}_0$ on the interior, there exists another point

$$\vec{x}_1 := \vec{x}_0 - \epsilon \frac{\vec{c}}{\|\vec{c}\|_2}, \quad (16)$$

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

Solution: We begin by showing that $\vec{x}_1 \in \mathcal{P}$. This is done by showing that $\vec{x}_1 \in \mathcal{B} \subset \mathcal{P}$:

$$\|\vec{x}_1 - \vec{x}_0\|_2 = \|\epsilon \frac{\vec{c}}{\|\vec{c}\|_2}\|_2 \quad (17)$$

$$= \epsilon. \quad (18)$$
Since $\bar{x}_1$ is within an $\epsilon$-ball of $\bar{x}_0$, we must have that $\bar{x}_1 \in P$.

Next, we show that $\bar{x}_1$ is more optimal:

$$c^T \bar{x}_1 = c^T \bar{x}_0 - \epsilon \frac{c^T c}{\|c\|_2}$$

$$= c^T \bar{x}_0 - \epsilon \|c\|_2$$

$$< c^T \bar{x}_0.$$  \hspace{1cm} (19) (20) (21)

Note: it is also entirely possible to show that $B \subseteq \text{Int}(P)$ in which case we can prove a stronger condition: namely, that given any point in the interior of the polytope, we can find another point, also in the interior, that is more optimal.
3. Duality
Consider the function
\[ f(\vec{x}) = \vec{x}^\top A\vec{x} - 2\vec{b}^\top \vec{x}. \]

First, we consider the unconstrained optimization problem
\[ p^* = \min_{\vec{x} \in \mathbb{R}^n} f(\vec{x}) = \min_{\vec{x} \in \mathbb{R}^n} \vec{x}^\top A\vec{x} - 2\vec{b}^\top \vec{x} \]
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 \( p^* = -\infty \). Let \( \vec{x}^* \) denote the minimizing argument of the optimization problem.

(a) Suppose \( A \succeq 0 \) (positive semidefinite) and \( \vec{b} \in \mathcal{R}(A) \). Let \( \text{rk}(A) = n \). Find \( p^* \).

\text{HINT: What does } A \succeq 0 \text{ tell you about the function } f? \text{ How can you leverage the rank of } A \text{ to compute } p^*?\]

\textbf{Solution:} If \( \text{rk}(A) = n \), then \( A > 0 \), and therefore the objective is strictly convex. Setting the gradient to 0 we obtain,
\[
\nabla_{\vec{x}} f(\vec{x}) = 2A\vec{x} - 2\vec{b} = 0 \quad \Rightarrow \quad A\vec{x} = \vec{b} \quad \Rightarrow \quad \vec{x}^* = A^{-1}\vec{b}.
\]

Where in the last step, we used that fact that a full rank square matrix is invertible. Plugging this back into our objective function we get,
\[
f(\vec{x}^*) = (\vec{b}^\top (A^{-1})^\top)A(A^{-1}\vec{b}) - 2\vec{b}^\top A^{-1}\vec{b} \quad \text{(27)}
\]
\[
= \vec{b}^\top A^{-1} A^{-1}\vec{b} - 2\vec{b}^\top A^{-1}\vec{b} \quad \text{(28)}
\]
\[
= -\vec{b}^\top A^{-1}\vec{b} \quad \text{(29)}
\]
\[
p^* = -\vec{b}^\top A^{-1}\vec{b} \quad \text{(30)}
\]

(b) Suppose \( A \succeq 0 \) (positive semidefinite) and \( \vec{b} \in \mathcal{R}(A) \) as before. Let \( A \) be rank-deficient, i.e., \( \text{rk}(A) = r < n \). Let \( A \) have the compact/thin and full SVD as follows, with diagonal positive definite \( \Lambda_r \in \mathbb{R}^{r \times r} \):
\[
A = U_r \Lambda_r U_r^\top = \begin{bmatrix} A_r & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} U_r^\top \\ 0 \end{bmatrix} = \begin{bmatrix} U_r^\top \Lambda_r \end{bmatrix} \quad \text{(31)}
\]

Show that the minimizer \( \vec{x}^* \) of the optimization problem (23) is not unique by finding a general form for the family of solutions for \( \vec{x}^* \) in terms of \( U_r, U_1, \Lambda_r, \vec{b} \).

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

\textbf{Solution:} Since \( A \succeq 0 \), \( f(\vec{x}) \) is convex and we can attempt to find the minimizer by setting the gradient to zero. Doing this we obtain,
\[
\nabla_{\vec{x}} f(\vec{x}) = 2A\vec{x} - 2\vec{b} = 0 \quad \Rightarrow \quad A\vec{x} = \vec{b} \quad \text{(32)}
\]
as in the part (a) of this problem.
However, now this equation has infinite solutions since $\vec{b}$ lies in the range of $A$ and $A$ is rank-deficient. Indeed we can add any vector from the (non-trivial) nullspace of $A$ to any particular solution $\vec{x}_0$ of Equation (32) and get another solution.

By the Fundamental Theorem of Linear Algebra we have,

\[
\vec{x} = U_r \vec{\alpha} + U_1 \vec{\beta} \tag{33}
\]

\[
\vec{b} = U_r \vec{\gamma}, \tag{34}
\]

where we used the fact that $\vec{b} \in \mathcal{R}(A)$. Using this we obtain,

\[
U_r \Lambda_r U_r^T (U_r \vec{\alpha} + U_1 \vec{\beta}) = U_r \vec{\gamma} \tag{35}
\]

Since the columns of $U_1$ and $U_r$ are orthogonal to each other and because $U_r^T U_r = I, \Lambda_r$ is invertible we have,

\[
U_r \Lambda_r U_r^T U_r \vec{\alpha} = U_r \vec{\gamma} \tag{36}
\]

\[
\Rightarrow \quad \vec{\alpha} = \Lambda_r^{-1} \vec{\gamma} \tag{37}
\]

\[
= \Lambda_r^{-1} U_r^T \vec{b}. \tag{38}
\]

Thus any solution to Equation (32) and hence a minimizer to the optimization problem (23) can be written as,

\[
\vec{x}^* = U_r \Lambda_r^{-1} U_r^T \vec{b} + U_1 \vec{\beta}. \tag{39}
\]

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

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

**Solution:** Since $A \not\succeq 0$ there exists an eigenvalue, eigenvector pair $(\mu, \vec{v})$ such that

\[
\vec{v}^T A \vec{v} = \mu < 0. \tag{40}
\]

Assuming without loss of generality that $-2\vec{b}^T \vec{v} \leq 0$ (If it is positive then multiply $\vec{v}$ by $-1$) we can take $\vec{x} = \alpha \vec{v}$ to obtain,

\[
f(\vec{x}) = f(\alpha \vec{v}) = \alpha^2 \vec{v}^T A \vec{v} + \alpha (-2 \vec{b}^T \vec{v}), \tag{41}
\]

which goes to $-\infty$ as $\alpha$ goes to $\infty$ since $\vec{v}^T A \vec{v} < 0$ and $-2\vec{b}^T \vec{v} \leq 0$.

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

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

**Solution:** First, note that since $A$ is symmetric, we have $\mathcal{R}(A) = \mathcal{R}(A^T)$. We have $\vec{b} = \vec{v}_1 + \vec{v}_2$ with $\vec{v}_1 \in \mathcal{R}(A) = \mathcal{R}(A^T)$ and $\vec{v}_2 \in \mathcal{N}(A)$ as $\mathbb{R}^n = \mathcal{R}(A) \oplus \mathcal{N}(A)$ from the Fundamental Theorem of Linear Algebra. We cannot have $\vec{v}_2 = 0$ as otherwise we’d get $\vec{b} = \vec{v}_1 \in \mathcal{R}(A)$ which is a contradiction. Now, let $\vec{v} = \vec{v}_2$. We get from this:

\[
f(\alpha \vec{v}) = \alpha^2 \vec{v}^T A \vec{v} - 2\alpha (\vec{v}_1 + \vec{v}_2)^T \vec{v}_2 = 0 - 2\alpha \|\vec{v}_2\|^2 \tag{42}
\]
where we used the fact that $\vec{v}_2 \in \mathcal{N}(A)$ and $\vec{v}_1 \in \mathcal{R}(A)$. As $\alpha \to \infty$, we get that $f(\alpha \vec{v}) \to -\infty$ from which we conclude that $p^* = -\infty$.

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 $\text{rk}(A) = r$, where $0 \leq r \leq n$. Now we consider the constrained optimization problem

$$
p^* = \min_{\vec{x} \in \mathbb{R}^n} \vec{x}^T A \vec{x} - 2\vec{b}^T \vec{x}
\text{s.t. } \vec{x}^T \vec{x} \geq 1.
$$

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

**Solution:**

$$
\mathcal{L}(\vec{x}, \lambda) = \vec{x}^T A \vec{x} - 2\vec{b}^T \vec{x} + \lambda (1 - \vec{x}^T \vec{x})
= \vec{x}^T A \vec{x} - \vec{x}^T \lambda \vec{x} - 2\vec{b}^T \vec{x} + \lambda
= \vec{x}^T (A - \lambda I) \vec{x} - 2\vec{b}^T \vec{x} + \lambda
$$

(f) For any matrix $C \in \mathbb{R}^{n \times n}$ with $\text{rk}(C) = r \leq n$ and compact SVD

$$
C = U_r \Lambda_r V_r^T,
$$
we define the pseudoinverse as

$$
C^\dagger = V_r \Lambda_r^{-1} U_r^T.
$$

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 (43) can be written as,

$$
d^* = \max_{\lambda \geq 0} -\vec{b}^T (A - \lambda I)^\dagger \vec{b} + \lambda.
$$

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, $\min_{\vec{x}} \mathcal{L}(\vec{x}, \lambda) = -\vec{b}^T (A - \lambda I)^\dagger \vec{b} + \lambda$.

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

**Solution:**

$$
g(\lambda) = \min_{\vec{x}} \mathcal{L}(\vec{x}, \lambda) = \min_{\vec{x}} \vec{x}^T (A - \lambda I) \vec{x} - 2\vec{b}^T \vec{x} + \lambda
$$

Drawing from parts (c) and (d), we can see that if $A - \lambda I \not\succeq 0$ or if $A - \lambda I \succeq 0, \vec{b} \not\in \mathcal{R}(A - \lambda I)$, then we can choose $\vec{x}$ to drive the Lagrangian to $-\infty$. If the constraints are satisfied, however, then we can proceed like in part (b) by taking the gradient:

$$
\nabla_{\vec{x}} \mathcal{L} = 2(A - \lambda I) \vec{x} - 2\vec{b} = 0
\Rightarrow (A - \lambda I) \vec{x} = \vec{b}
\Rightarrow \vec{x}^* = (A - \lambda I)^\dagger \vec{b}
$$
where in the last step, we used the fact that the PSD constraint on $A - \lambda I$ is satisfied and $\tilde{b}$ lies in the range of $A - \lambda I$, so we can use the pseudoinverse and the gradient-zero point is indeed the minimum. Plugging this back into the Lagrangian, we get:

$$
\mathcal{L}(\tilde{x}^*, \lambda) = \tilde{b}^\top ((A - \lambda I)^\dagger)(A - \lambda I)(A - \lambda I)^\dagger\tilde{b} - 2\tilde{b}^\top (A - \lambda I)^\dagger\tilde{b} + \lambda
$$

(54)

$$
= \tilde{b}^\top (A - \lambda I)^\dagger(A - \lambda I)(A - \lambda I)^\dagger\tilde{b} - 2\tilde{b}^\top (A - \lambda I)^\dagger\tilde{b} + \lambda
$$

(55)

$$
= \tilde{b}^\top (A - \lambda I)^\dagger\tilde{b} - 2\tilde{b}^\top (A - \lambda I)^\dagger\tilde{b} + \lambda
$$

(56)

$$
= -\tilde{b}^\top (A - \lambda I)^\dagger\tilde{b} + \lambda
$$

(57)

where we used the fact that $(A - \lambda I)^\dagger$ is symmetric and by properties of pseudo inverse,

$$(A - \lambda I)^\dagger(A - \lambda I)(A - \lambda I)^\dagger = (A - \lambda I)^\dagger.$$

(58)

Now, we have a full expression for our dual function:

$$
g(\lambda) = \begin{cases} 
-\tilde{b}^\top (A - \lambda I)^\dagger\tilde{b} + \lambda & \text{if } A - \lambda I \succeq 0, \tilde{b} \in \mathcal{R}(A - \lambda I) \\
-\infty & \text{else}
\end{cases}
$$

(59)

The dual problem follows, as it is just a maximization of the dual function:

$$
d^* = \max_{\lambda \geq 0} g(\lambda)
$$

(60)
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](image)

Table 1: Problem data for Problem 2. Here \( n = 5 \).

<table>
<thead>
<tr>
<th>( i )</th>
<th>( x_i )</th>
<th>( y_i )</th>
<th>( c_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>4</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>4</td>
<td>N/A</td>
</tr>
</tbody>
</table>

(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

\[
\min_{\vec{z}} \sum_{i=1}^{n+1} \left\| \begin{bmatrix} x_i \\ z_i \end{bmatrix} - \begin{bmatrix} x_{i-1} \\ z_{i-1} \end{bmatrix} \right\|_2
\]
\[
\text{s.t. } y_i - c_i/2 \leq z_i \leq y_i + c_i/2, \text{ for } i = 1, \ldots, n \\
\]
\[
z_0 = y_0, z_{n+1} = y_{n+1},
\]
which is equivalent to
\[
\min_{\vec{z}} \sum_{i=1}^{n+1} t_i 
\]
\[
\text{s.t. } y_i - c_i/2 \leq z_i \leq y_i + c_i/2, \text{ for } i = 0, \ldots, n + 1 \\
\left\| \begin{bmatrix} x_i \\ z_i \end{bmatrix} - \begin{bmatrix} x_{i-1} \\ z_{i-1} \end{bmatrix} \right\|_2 \leq t_i, \text{ for } i = 1, \ldots, n + 1.
\]

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. \textit{HINT: You should be able to use packages such as cvxpy and numpy.}

\textbf{Solution:} The code can be found in the corresponding Jupyter notebook.
5. Sphere Enclosure

Let $B_i$, $i = 1, \ldots, m$, be $m$ Euclidean balls in $\mathbb{R}^n$, with centers $\vec{x}_i$, and radii $\rho_i \geq 0$. We wish to find a ball $B$ of minimum radius that contains all the $B_i$, $i = 1, \ldots, m$. Cast this problem as an SOCP.

**Solution:** Let $\vec{c} \in \mathbb{R}^n$ and $r \geq 0$ denote the center and radius of the enclosing ball $B$, respectively. We express the given balls $B_i$ as

$$B_i = \{ \vec{x} : \vec{x} = \vec{x}_i + \vec{\delta}_i, \| \vec{\delta}_i \|_2 \leq \rho_i \}, \quad i = 1, \ldots, m. \quad (67)$$

We have that $B_i \subseteq B$ if and only if

$$\max_{\vec{x} \in B_i} \| \vec{x} - \vec{c} \|_2 \leq r. \quad (68)$$

Note that

$$\max_{\vec{x} \in B_i} \| \vec{x} - \vec{c} \|_2 = \max_{\| \vec{\delta}_i \|_2 \leq \rho_i} \| \vec{x}_i - \vec{c} + \vec{\delta}_i \|_2 = \| \vec{x}_i - \vec{c} \|_2 + \rho_i. \quad (69)$$

The last step follows by choosing $\vec{\delta}_i$ in the direction of $\vec{x}_i - \vec{c}$. The problem is then cast as the following SOCP

$$\min_{\vec{c}, r} \quad r \quad (70)$$

$$\text{s.t.} \quad \| \vec{x}_i - \vec{c} \|_2 + \rho_i \leq r, \quad i = 1, \ldots, m. \quad (71)$$
6. Dual Norms and SOCP

Consider the problem

\[ p^* = \min_{\vec{x} \in \mathbb{R}^n} \|A\vec{x} - \vec{y}\|_1 + \mu \|\vec{x}\|_2, \quad (72) \]

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.

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

\[
\begin{align*}
\min & \quad \vec{z}^T \vec{1} + \mu t \\
\text{s.t.} & \quad |(A\vec{x} - \vec{y})_i| \leq z_i, \quad i = 1, \ldots, m \quad (74) \\
& \quad \|\vec{x}\|_2 \leq t. \quad (75)
\end{align*}
\]

This expression now satisfies our definition of an SOCP: the objective is linear, and all constraints are SOC constraints. (To see this, recall that the first set of absolute value constraints can be written equivalently as \( (A\vec{x} - \vec{y})_i \leq z_i \) and \( (A\vec{x} - \vec{y})_i \geq -z_i \), which are linear constraints; we keep the absolute value version for conciseness. The second set of constraints are SOC constraints in canonical 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:

\[
\|\vec{z}\|_2 = \max_{\vec{u} : \|\vec{u}\|_2 \leq 1} \vec{u}^T \vec{z}, \quad \|\vec{z}\|_1 = \max_{\vec{u} : \|\vec{u}\|_\infty \leq 1} \vec{u}^T \vec{z}. \quad (76)
\]

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

\[
\|A\vec{x} - \vec{y}\|_1 + \mu \|\vec{x}\|_2 = \max_{\vec{u} : \|\vec{u}\|_\infty \leq 1} \vec{u}^T (A\vec{x} - \vec{y}) + \mu \max_{\vec{v} : \|\vec{v}\|_2 \leq 1} \vec{v}^T \vec{x}. \quad (77)
\]

We can then express the original (primal) problem as

\[
p^* = \min_{\vec{x}} \max_{\vec{u}, \vec{v} : \|\vec{u}\|_\infty \leq 1, \|\vec{v}\|_2 \leq 1} \vec{u}^T (A\vec{x} - \vec{y}) + \mu \vec{v}^T \vec{x}. \quad (78)
\]

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

\[
d^* = \max_{\vec{u}, \vec{v} : \|\vec{u}\|_\infty \leq 1, \|\vec{v}\|_2 \leq 1} \min_{\vec{x}} \vec{u}^T (A\vec{x} - \vec{y}) + \mu \vec{v}^T \vec{x} \quad (79)
\]

\[
= \max_{\vec{u}, \vec{v} : \|\vec{u}\|_\infty \leq 1, \|\vec{v}\|_2 \leq 1} g(\vec{u}, \vec{v}), \quad (80)
\]

where \( g \) is defined as

\[
g(\vec{u}, \vec{v}) = \min_{\vec{x}} \vec{u}^T (A\vec{x} - \vec{y}) + \mu \vec{v}^T \vec{x} \quad (81)
\]

\[
= \min_{\vec{x}} (\vec{u}^T A + \mu \vec{v}^T) \vec{x} - \vec{u}^T \vec{y} \quad (82)
\]

\[
= \begin{cases} 
-\vec{u}^T \vec{y} & \text{if } A^T \vec{u} + \mu \vec{v} = \vec{0} \\
-\infty & \text{otherwise.} 
\end{cases} \quad (83)
\]
We can thus rewrite the dual problem as

\[ d^* = \max_{\bar{u}, \bar{v}} - \bar{u}^T \bar{y} \quad (84) \]

subject to

\[ A^T \bar{u} + \mu \bar{v} = 0 \quad (85) \]

\[ \|\bar{u}\|_\infty \leq 1, \|\bar{v}\|_2 \leq 1. \quad (86) \]

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

\[ d^* = \max_{\bar{u}} - \bar{u}^T \bar{y} \quad (87) \]

subject to

\[ \|A^T \bar{u}\|_2 \leq \mu \quad (88) \]

\[ \|\bar{u}\|_\infty \leq 1. \quad (89) \]

generating our final SOCP dual. If desired, we can further rewrite the final constraint as \( \|\bar{u}\|_\infty = \max_i |u_i| \leq 1 \Leftrightarrow |u_i| \leq 1, \ i = 1, \ldots, m \Leftrightarrow u_i \leq 1 \) 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 \( 100 \times 10^6 \). 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.