IOE 611: Homework 1

Yulun Zhuang

September 24, 2024

Problem 1

For a vector $x \in \mathbf{R}^n$, its ℓ_∞-norm is defined as $\|x\|_\infty = \max_{1\leq i\leq n} |\alpha_i|$. Consider the following two optimization problems:

$$\begin{aligned} (P1) \min_x \ & \|Ax - b\|_\infty \\ \text{s.t. } & x \succeq 0 \end{aligned}$$

$$\begin{aligned} (P2) \min_{y,t} \quad & t \\ \text{s.t.} \quad & a_i^\top y - t \leq b_i, \\ & -a_i^\top y - t \leq -b_i, \forall i = 1, \ldots, k, \\ & y \succeq 0. \end{aligned}$$

where $x \in \mathbf{R}^n$, and $a_i^\top$ is the i th row of A. Note that $(P2)$ is a linear program. Show that $(P1)$ and $(P2)$ are equivalent by showing that:

(a) For every feasible solution x of $(P1)$, there exists a feasible solution (y, t) of $(P2)$ such that $t \leq \|Ax - b\|_\infty$

Proof. Notice that

$$\begin{aligned} \|Ax - b\|_\infty &= \max_{i=1,\ldots,k} |(Ax - b)_i| \\ &= \max_{i=1,\ldots,k} \left|a_i^\top x - b_i\right| \end{aligned}$$

For any feasible solution $x \succeq 0$ of $(P1)$, let

$$\begin{cases} \hat{y} = x \succeq 0 \\ \hat{t} = \|Ax - b\|_\infty = \max_i \left|a_i^\top \hat{y} - b_i\right| \end{cases}$$

For any $i = 1, \ldots, k$,

$$\begin{cases} \hat{t} \geq \left|a_i^\top \hat{y} - b_i\right| \\ \hat{t} \geq a_i^\top x - b_i \Leftrightarrow b_i \geq a_i^\top x - \hat{t}, \\ \hat{t} \geq b_i - a_i^\top x \Leftrightarrow -b_i \geq -a_i^\top x - \hat{t}, \end{cases}$$

which implies $(\hat{y}, \hat{t})$ is a feasible solution of (P2). $\square$

(b) For every feasible solution (y, t) of $(P2)$, there exists a feasible solution x of $P1$ such that $\|Ax - b\|_\infty \leq t$

Proof. For any feasible solution (y, t) of $(P2)$, let $\hat{x} = y$. By definition

$$
\begin{aligned}
& y \succeq 0 \\
& a_i^\top y - t \leq b_i, \forall i = 1, \ldots, k \\
& -a_i^\top y - t \leq -b_i, \forall i = 1, \ldots, k
\end{aligned}
$$

For any $i = 1, \ldots, k$

$$
\begin{aligned}
& -t \leq a_i^\top y - b_i \leq t \\
& t \geq \left| a_i^\top y - b_i \right| \\
& t \geq \max_{i=1,\ldots,k} \left| a_i^\top y - b_i \right| \\
& t \geq \|Ay - b\|_\infty \\
& t \geq \|A\hat{x} - b\|_\infty
\end{aligned}
$$

which implies $\hat{x}$ is a feasible solution of $(P1)$. $\square$

Problem 2

Derive a linear program equivalent to

$$\min_x \|Ax - b\|_\infty + \rho\|x\|_1$$

where $\rho \geq 0$.

Solution.

$$\begin{aligned}
\min_{t,y,z} \quad & t + \rho \sum_{j=1}^{n} z_j \\
s.t. \quad & a_i^\top - t \leq b_i \\
& - a_i^\top - t \leq -b_i \\
& y_j \leq z_j \\
& - y_j \leq z_j \\
& i = 1, \dots, k \\
& j = 1, \dots, n
\end{aligned}$$

$\blacksquare$

Problem 3

Let $C \subseteq \mathbf{R}^n$ be a convex set, with $x_1, \ldots, x_k \in C$, and let $\theta_1, \ldots, \theta_k \in \mathbf{R}$ satisfy $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$. Show that $\theta_1 x_1 + \cdots + \theta_k x_k \in C$.

Proof. By induction, for $k = 2$, $\theta_1 x_1 + \theta_2 x_2 \in C$ holds by the definition of convexity.

Assume this holds for $k = m, m \geq 2$.

$$\theta_1 x_1 + \cdots + \theta_m x_m \in C$$

For $k = m+1$, $x_1, \ldots, x_m, x_{m+1} \in C$, and $\theta_1, \ldots, \theta_m, \theta_{m+1} \in \mathbb{R}$ satisfy $\theta_i \geq 0$, $\sum_{i=1}^{m+1} \theta_i = 1$.

$$\begin{aligned}
&\sum_{i=1}^{m+1} \theta_i \alpha_i \\
&= \sum_{i=1}^{m} \theta_i \alpha_i + \theta_{m+1} x_{m+1} \\
&= (1 - \theta_{m+1}) \sum_{i=1}^{m} \frac{\theta_i}{1 - \theta_{m+1}} \alpha_i + \theta_{m+1} x_{m+1}
\end{aligned}$$

Since $\sum_{i=1}^{m} \frac{\theta_i}{1-\theta_{m+1}} \alpha_i \in C$ by the induction hypothesis, and

$$\begin{aligned}
&\sum_{i=1}^{m+1} \theta_i = 1 \\
&\sum_{i=1}^{m} \theta_i = 1 - \theta_{m+1} \\
&\sum_{i=1}^{m} \frac{\theta_i}{1 - \theta_{m+1}} = 1
\end{aligned}$$

Original proposition holds for arbitrary k. $\square$

Problem 4

(a) Show that if a matrix is symmetric and has non-negative eigenvalues, then it must be PSD.

Proof. Suppose $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix and has non-negative eigenvalues, we can decompose A as

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^\top, i = 1, \ldots, n$$

where $\lambda_i \geq 0$ are eigenvalues of A and v_i is the corresponding eigenvectors. Note that $\{v_i\}_{i\in[n]}$ is an orthonormal basis of $\mathbb{R}^n$, i.e.

$$v_i^T v_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

For any $x \in \mathbb{R}^n$, $\exists \alpha_i \in \mathbb{R}, i = 1, \ldots, n$, we can represent it as

$$x = \sum_{i=1}^{n} \alpha_i v_i$$

and then

$$\begin{aligned}
Ax &= \left(\sum_{i=1}^{n} \lambda_i v_i v_i^\top\right)\left(\sum_{j=1}^{n} \alpha_j v_j\right) \\
&= \sum_{i,j=1}^{n} \lambda_i \alpha_j \left(v_i^\top v_j\right) v_j \\
&= \sum_{i=1}^{n} \lambda_i \alpha_i v_i, \\
x^\top Ax &= \left(\sum_{i=1}^{n} \alpha_i v_i^\top\right)\left(\sum_{j=1}^{n} \lambda_j \alpha_j v_j\right) \\
&= \sum_{i,j=1}^{n} \lambda_j \alpha_i \alpha_j \left(v_i^\top v_j\right) \\
&= \sum_{i=1}^{n} \lambda_i \alpha_i^2 \geqslant 0.
\end{aligned}$$

Therefore, A is PSD.

□

(b) Show that a symmetric matrix X is PSD if and only if it can be written as $X = VV^\top$, for some matrix V .

Proof. $\Rightarrow$: If a symmetric matrix X is PSD, it can be decomposed as

$$
\begin{aligned}
X &= \sum_{i=1}^{n} \lambda_i v_i v_i^\top \\
&= Q \Lambda Q^\top
\end{aligned}
$$

where $\lambda_i \geq 0$ are eigenvalues of A; v_i is the corresponding eigenvectors; $Q = [v_1, \ldots, v_n]$ is an orthogonal matrix and $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix.

Let $V = Q\sqrt{\Lambda}$, where $\sqrt{\Lambda} = diag(\sqrt{\lambda}_1, \ldots, \sqrt{\lambda}_n)$, we have

$$
\begin{aligned}
X &= Q\sqrt{\Lambda}\sqrt{\Lambda}^\top Q^\top \\
&= (Q\sqrt{\Lambda})(Q\sqrt{\Lambda})^\top \\
&= VV^\top
\end{aligned}
$$

$\Leftarrow$: If a symmetric matrix $X = VV^\top$, for any $z \in \mathbb{R}^n$, we have

$$
\begin{aligned}
z^\top X z &= z^\top V V^\top z \\
&= (V^\top z)^\top (V^\top z) \\
&= \|V^\top z\| \succeq 0
\end{aligned}
$$

Therefore, X is PSD. $\square$

Problem 5

Show that the PSD cone is self-dual, that is, $(\mathbb{S}^n_+)^* = \mathbb{S}^n_+$.

Proof. By the definition of dual cone, we have

$$(\mathbb{S}^n_+)^* = \{B \in \mathbb{S}^n_+ \,|\, \langle A, B\rangle \geq 0, \forall A \in \mathbb{S}^n_+\}$$

where $\langle A, B\rangle = tr(AB)$.

We now show that $\forall A \in \mathbb{S}^n_+$, $\langle A, B\rangle \geq 0$ if and only if $B \in \mathbb{S}^n_+$.

$\Rightarrow$: Suppose $B \notin \mathbb{S}^n_+$, then there exists $z \in \mathbb{R}^n$, such that $z^\top Bz = tr(zz^\top B) < 0$.

Let $A = zz^\top$ is PSD, then

$$\langle A, B\rangle = tr(AB) = tr(zz^\top B) < 0$$

which shows $B \notin (\mathbb{S}^n_+)^*$

$\Leftarrow$: Since $A \in \mathbb{S}^n_+$, A can be decomposed as

$$A = \sum_{i=1}^n \lambda_i z_i z_i^\top$$

where $\lambda_i \geq 0$ are eigenvalues. Then,

$$\begin{aligned}
\langle A, B\rangle &= tr(AB) \\
&= tr(B \sum_{i=1}^n \lambda_i z_i z_i^\top) \\
&= \sum_{i=1}^n \lambda_i z_i^\top B z_i \geq 0
\end{aligned}$$

which shows $B \in (\mathbb{S}^n_+)^*$

□

Problem 6

The second-order cone is defined as $C = \{(x, t) : \|x\|_2 \leq t, t \geq 0\}$.

(a) Prove that C is a cone ant it is convex.

Proof. To show C is a cone, for any $\lambda > 0, (x, t) \in C$ we have

$$\|\lambda x\|_2 = \lambda \|x\|_2 \leq \lambda t$$

which implies $\lambda(x, t) = (\lambda x, \lambda t) \Leftrightarrow C$ is a cone.

To show C is convex, for any $(x_1, t_1), (x_2, t_2) \in C$, $\theta \in [0, 1]$, we have

$$\begin{aligned} \|\theta x_1 + (1-\theta) x_2\|_2 &\leq \theta \|x_1\|_2 + (1-\theta) \|x_2\|_2 \\ &\leq \theta t_1 + (1-\theta) t_2 \end{aligned}$$

which implies $\theta (x_1, t_1) + (1-\theta)(x_2, t_2) = (\theta x_1 + (1-\theta) x_2, \theta t_1 + (1-\theta) t_2) \Leftrightarrow C$ is convex. $\square$

(b) Let P be a PSD matrix. Prove that the set $C = \{x | x^\top P x \leq (a^\top x)^2, a^\top x \geq 0\}$ is a cone and it is convex. This cone is called hyperbolic cone.

Proof. To show C is a cone, for any $\lambda > 0, x \in C$,

$$\begin{aligned} (\lambda x)^\top P (\lambda x) &= \lambda^2 \left(x^\top P x \right) \\ &\leqslant \lambda^2 \left(a^\top x \right)^2 \\ &= \left(a^\top (\lambda x) \right)^2 \end{aligned}$$

Since $a^\top \lambda x = \lambda a^\top x \geq 0$, $\lambda x \in C$. C is a cone. $\square$