

Homework 1

Instructor: Lijun Zhang

Name: Chenxiao Gao, StudentId: 181220014

Notice

- The submission email is: **njuoptfall2019@163.com**.
- Please use the provided L^AT_EX file as a template. If you are not familiar with L^AT_EX, you can also use Word to generate a **PDF** file.

Problem 1: Norms

A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with $\text{dom} f = \mathbb{R}^n$ is called a *norm* if

- f is nonnegative: $f(x) \geq 0$ for all $x \in \mathbb{R}^n$
- f is definite: $f(x) = 0$ only if $x = 0$
- f is homogeneous: $f(tx) = |t|f(x)$, for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$
- f satisfies the triangle inequality: $f(x+y) \leq f(x) + f(y)$, for all $x, y \in \mathbb{R}^n$

We use the notation $f(x) = \|x\|$. Let $\|\cdot\|$ be a norm on \mathbb{R}^n . The associated dual norm, denoted $\|\cdot\|_*$, is defined as

$$\|z\|_* = \sup\{z^T x \mid \|x\| \leq 1\}$$

- a) Prove that $\|\cdot\|_*$ is a valid norm.
 b) Prove that the dual of the Euclidean norm (ℓ_2 -norm) is the Euclidean norm, *i.e.*, prove that

$$\|z\|_{2*} = \sup\{z^T x \mid \|x\|_2 \leq 1\} = \|z\|_2$$

(Hint: Use Cauchy-Schwarz inequality.)

Solution.

- a) We can prove that $f(x)$ satisfies the abovementioned properties:

- 1) randomly choose an vector x that satisfies $\|x\| \leq 1$, and it is trivial that $\| -x \| \leq 1$. If $z^T x \geq 0$, then we have $\|z\|_* = \sup\{z^T x \mid \|x\| \leq 1\} \geq 0$, or else $\|z\|_* \geq z^T(-x) \geq 0$. Thus, $\|\cdot\|_*$ is nonnegative.
- 2) based on 1), if $\|z\|_* = 0$ then we have $z^T x, z^T(-x) = -z^T x \leq 0$, which denotes that for all $x \in \mathbb{R}^n$, $z^T x = -z^T x = 0$, and $z^T = 0$. if $z^T = 0$, it is trivial that $\|z\|_* = 0$. Thus, $\|\cdot\|_*$ is definite.
- 3) for all $x \in \mathbb{R}^n$ and for all $t \in \mathbb{R}$,
 if $t \geq 0$, then $\|tz\|_* = \sup\{tz^T x \mid \|x\| \leq 1\} = t \sup\{z^T x \mid \|x\| \leq 1\} = |t|\|z\|_*$ (t is a constant), which means $\|\cdot\|_*$ is homogeneous.
 else if $t \leq 0$, then

$$\begin{aligned} \|tz\|_* &= \| -t(-z) \|_* \\ &= \sup\{(-t)(-z)^T(-x) \mid \| -x \| \leq 1\} \\ &= (-t) \sup\{(-z)^T(-x) \mid \| -x \| \leq 1\} \\ &= |t| \sup\{z^T x \mid \| -x \| \leq 1\} \\ &= |t| \sup\{z^T x \mid \|x\| \leq 1\} \\ &= |t| \|z\|_* \end{aligned}$$

Therefore, $\|\cdot\|_*$ is homogeneous.

4) For all $x, y \in v$, $\|x + y\|_* = \sup\{x^T z + y^T z \mid \|z\| \leq 1\} \leq \sup\{x^T z \mid \|z\| \leq 1\} + \sup\{y^T z \mid \|z\| \leq 1\} = \|x\|_* + \|y\|_*$, and the equality holds if and only if $x^T z$ and $y^T z$ achieves their maximum at the same point. Thus $\|\cdot\|_*$ satisfies the triangle inequality.

b) From Cauchy-Schwarz inequality, we can show that $|z^T x|$ achieves maximum when $x = \frac{z}{\|z\|_2}$. Noticing that $\frac{z}{\|z\|_2} \leq 1$, we can conclude that $\|z\|_* = \sup\{z^T x \mid \|x\| \leq 1\} = \frac{z^T z}{\|z\|_2} = \|z\|_2$, and the equality holds when $x = \frac{z}{\|z\|_2}$ or $\frac{-z}{\|z\|_2}$, depending on the sign of $z^T x$. Q.E.D.

□

Problem 2: Affine and Convex Sets

Affine sets C_a and convex C_c sets are the sets satisfying the constraints below:

$$\begin{aligned} \theta x_1 + (1 - \theta)x_2 &\in C_a \\ \text{s.t. } x_1, x_2 &\in C_a \end{aligned} \quad (1)$$

$$\begin{aligned} \theta x_1 + (1 - \theta)x_2 &\in C_c \\ \text{s.t. } x_1, x_2 &\in C_c, 0 \leq \theta \leq 1 \end{aligned} \quad (2)$$

a) Is the set $\{\alpha \in \mathbb{R}^k \mid p(0) = 1, |p(t)| \leq 1 \text{ for } \alpha \leq t \leq \beta\}$, where $p(t) = \alpha_1 + \alpha_2 t + \dots + \alpha_k t^{k-1}$, affine?

b) Determine if each set below is convex.

- 1) $\{(x, y) \in \mathbf{R}_{++}^2 \mid x/y \leq 1\}$.
- 2) $\{(x, y) \in \mathbf{R}_{++}^2 \mid x/y \geq 1\}$.
- 3) $\{(x, y) \in \mathbf{R}_+^2 \mid xy \leq 1\}$.
- 4) $\{(x, y) \in \mathbf{R}_+^2 \mid xy \geq 1\}$.
- 5) $\{(x, y) \in \mathbf{R}^2 \mid y = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}\}$.

Solution.

a) Not necessarily.

proof:

Define $S = \{\alpha \in \mathbb{R}^k \mid p(0) = 1, |p(t)| \leq 1 \text{ for } \alpha \leq t \leq \beta\}$ for all α_1 and $\alpha_2 \in \mathbb{R}^n$

Consider $\alpha_1 = (1, 0, -1), \alpha_2 = (1, 0, -2), t \in [-1, 1]$, we can see that both α_1 and α_2 belongs to S .

However, for $\alpha_* = \theta\alpha_1 + (1 - \theta)\alpha_2$, the absolute value of the polynomial $p(t) = 1 + (\theta - 2)t^2$ is not necessarily below 1.

Thus, the set S is not necessarily affine.

b) Define S as the following sets.

1) Convex.

proof:

For all (x_1, y_1) and $(x_2, y_2) \in \mathbf{R}_{++}^2$, we have $x_1 \leq y_1, x_2 \leq y_2$, so $\theta x_1 + (1 - \theta)x_2 \leq \theta y_1 + (1 - \theta)y_2$.

$\therefore \theta, (1 - \theta), y_1, y_2 \geq 0$

$\therefore \frac{\theta x_1 + (1 - \theta)x_2}{\theta y_1 + (1 - \theta)y_2} \leq 1$

Therefore the set S is convex.

2) Convex.

proof:

For all (x_1, y_1) and $(x_2, y_2) \in \mathbf{R}_{++}^2$, we have $x_1 \geq y_1, x_2 \geq y_2$, so $\theta x_1 + (1 - \theta)x_2 \geq \theta y_1 + (1 - \theta)y_2$.

$\therefore \theta, (1 - \theta), y_1, y_2 \geq 0$

$\therefore \frac{\theta x_1 + (1 - \theta)x_2}{\theta y_1 + (1 - \theta)y_2} \geq 1$

Therefore the set S is convex.

3) Not convex.

proof:

Assign $(x_1, y_1) = (2, 1/2), (x_2, y_2) = (2, 1/2), \theta = 1/2$

$$\therefore (\theta x_1 + (1 - \theta)x_2)(\theta y_1 + (1 - \theta)y_2) = (5/4)^2 = 25/16 > 1$$

Thus, the set S is not convex.

4) Convex.

proof:

$$\therefore x_1 y_1 \geq 1, x_2 y_2 \geq 1$$

$$\therefore x_1 y_2 + x_2 y_1 \geq \frac{x_1}{x_2} + \frac{x_2}{x_1} \geq 2$$

$$\therefore (\theta x_1 + (1 - \theta)x_2)(\theta y_1 + (1 - \theta)y_2) = (x_2 y_2) + (\theta - \theta^2)(2 - x_1 y_2 - x_2 y_1)$$

$$\therefore (\theta - \theta^2) \leq 0, 2 - x_1 y_2 - x_2 y_1 \leq 0$$

$$\therefore (\theta x_1 + (1 - \theta)x_2)(\theta y_1 + (1 - \theta)y_2) \geq 1$$

Therefore the set S is convex.

5) Not Convex.

proof:

Assign $(x_1, y_1) = (0, 0), (x_2, y_2) = (1, \frac{e-e^{-1}}{e+e^{-1}}), \theta = 1/2$

$$\text{However, } \frac{1}{2}y_1 + \frac{1}{2}y_2 \neq \tanh(\frac{1}{2}x_1 + \frac{1}{2}x_2)$$

Therefore, the set S is not convex.

□

Problem 3: Examples

a) Let $C \subseteq \mathbb{R}^n$ be the solution set of a quadratic inequality,

$$C = \{x \in \mathbb{R}^n | x^\top A x + b^\top x + c \leq 0\}, \quad (3)$$

with $A \in \mathbb{S}^n, b \in \mathbb{R}^n$, and $c \in \mathbb{R}$.

1) Show that C is convex if $A \succeq 0$.

2) Is the following statement true? The intersection of C and the hyperplane defined by $g^\top x + h = 0$ is convex if $A + \lambda g g^\top \succeq 0$ for some $\lambda \in \mathbb{R}$.

b) The polar of $C \subseteq \mathbb{R}^n$ is defined as the set

$$C^\circ = \{y \in \mathbb{R}^n | y^\top x \leq 1 \text{ for all } x \in C\}$$

1) Show that C° is convex.

2) What is a polar of a polyhedra?

3) What is the polar of the unit ball for a norm $\|\cdot\|$?

4) Show that if C is closed and convex, with $0 \in C$, then $(C^\circ)^\circ = C$

Solution.

a) 1) proof:

C is convex

$$\Leftrightarrow \text{for any } x_1 \in C, x_2 \in C, \text{ and } \theta \in [0, 1], \theta x_1 + (1 - \theta)x_2 \in C$$

$$\Leftrightarrow (\theta x_1 + (1 - \theta)x_2)^\top A (\theta x_1 + (1 - \theta)x_2) + b^\top (\theta x_1 + (1 - \theta)x_2) + c \leq 0$$

$$\Leftrightarrow \theta^2 x_1^\top A x_1 + (1 - \theta)^2 x_2^\top A x_2 + \theta(1 - \theta)x_1^\top A x_2 + \theta(1 - \theta)x_2^\top A x_1 + b^\top (\theta x_1 + (1 - \theta)x_2) + c \leq 0$$

(considering the shapes of $x_1^\top A x_2$ and $x_2^\top A x_1$ are both 1×1 , $x_1^\top A x_2$ actually equals to $x_2^\top A x_1$)

$$\Leftrightarrow \theta^2 x_1^\top A x_1 + (1 - \theta)^2 x_2^\top A x_2 + 2\theta(1 - \theta)x_1^\top A x_2 + b^\top (\theta x_1 + (1 - \theta)x_2) + c \leq 0$$

$$\begin{aligned}
& \because b^T x_1 \leq -c - x_1^T A x_1, b^T x_2 \leq -c - x_2^T A x_2 \\
& \therefore P = \theta^2 x_1^T A x_1 + (1-\theta)^2 x_2^T A x_2 + 2\theta(1-\theta)x_1^T A x_2 + b^T(\theta x_1 + (1-\theta)x_2) + c \\
& \leq \theta^2 x_1^T A x_1 + (1-\theta)^2 x_2^T A x_2 + 2\theta(1-\theta)x_1^T A x_2 + \theta(-c - x_1^T A x_1) + (1-\theta)(-c - x_2^T A x_2) + c \\
& = \theta(\theta-1)x_1^T A x_1 + \theta(\theta-1)x_2^T A x_2 + 2\theta(1-\theta)x_1^T A x_2 \\
& = \theta(1-\theta)(2x_1^T A x_2 - x_1^T A x_1 - x_2^T A x_2) \\
& = \theta(1-\theta)(x_1^T A(x_2 - x_1) + (x_1 - x_2)^T A x_2) \\
& = \theta(1-\theta)(-(x_1 - x_2)^T A(x_1 - x_2)) \\
& \because A \in \mathbb{S}_+^n \\
& \therefore \theta(1-\theta)(-(x_1 - x_2)^T A(x_1 - x_2)) \leq 0 \\
& \therefore P \leq 0
\end{aligned}$$

This means the original proposition 'C is convex' is true.

Q.E.D.

2) the statement is true.

proof:

Let D stands for the hyperplane defined by $g^T x + h = 0$

First, it is trivial that any convex combination of $x_1, x_2 \in C \cup D$ also belongs to D , for hyperplanes are always convex.

So we only need to prove that the convex combination belongs to C .

For any $x_1 \in C \cup D$ and $x_2 \in C \cup D$, we have $g^T x_1 + h = g^T x_2 + h = 0$

$$\therefore g^T x_1 + h = g^T x_2 + h = 0$$

$$\therefore x_1^T g = x_2^T g = 0$$

$$\therefore x_1^T g g^T x_1 = x_2^T g g^T x_2 = 0$$

$$\therefore (x_1 - x_2)^T g g^T (x_1 - x_2) = 0$$

$$\therefore A + \lambda g g^T \in \mathbb{S}_+^n$$

$$\therefore (x_1 - x_2)^T (A + g g^T) (x_1 - x_2) \geq 0$$

$$\therefore (x_1 - x_2)^T g g^T (x_1 - x_2) = 0$$

$$\therefore (x_1 - x_2)^T A (x_1 - x_2) = (x_1 - x_2)^T (A + g g^T) (x_1 - x_2) - (x_1 - x_2)^T g g^T (x_1 - x_2) \geq 0$$

This means the polynomial P defined in 1) satisfies $P \leq 0$

Therefore, the set C is convex, and $\theta x_1 + (1-\theta)x_2 \in C \cup D$

Q.E.D.

b) 1) proof:

For any $y_1, y_2 \in C^\circ$, we have $y_1^T x \leq 1$ (for all $x \in C$) and $y_2^T x \leq 1$ (for all $x \in C$).

$$\therefore \text{For all } x \in C \text{ and } \theta \in [0, 1], (\theta y_1 + (1-\theta)y_2)^T x = \theta y_1^T x + (1-\theta)y_2^T x \leq \theta + (1-\theta) = 1$$

$$\therefore \theta y_1 + (1-\theta)y_2 \in C^\circ$$

Therefore C° is convex.

2) solution:

Suppose P is a bounded polyhedron. (I am not sure whether polyhedra are all bounded, but here I suppose they are)

Therefore, P can be represented by $\text{conv}\{v_1, v_2, \dots, v_k\}$, for k finite, and any element from P can be represented by $\theta_1 v_1 + \dots + \theta_k v_k$, where $\theta_i \geq 0$ and $1^T \theta = 1$

For $x \in P^\circ$, the inner product of x and $\theta_1 v_1 + \dots + \theta_k v_k$ is below 1 no matter how θ varies.

$$\text{So we let } x \text{ satisfies } x^T v_i \leq 1, i = 1, 2, \dots, k, \text{ so that } x^T(\theta_1 v_1 + \dots + \theta_k v_k) = \theta_1 x^T v_1 + \dots + \theta_k x^T v_k \leq \theta_1 + \theta_2 + \dots + \theta_k = 1$$

In the meantime, the set $\{x \in \mathbb{R}^n | x^T v_i \leq 1, i = 1, 2, \dots, k\}$ is actually a polyhedron.

So, the polar of a polyhedron is a polyhedron.

3) solution:

According to the definition of dual norm $\|\cdot\|_*: \|z\|_* = \sup\{z^T x | \|x\| \leq 1\}$,

the polar of the unit ball for $\|\cdot\|$ is $\{y | \|y\|_* \leq 1\}$

4) proof:

We prove $(C^\circ)^\circ = C$ by 2 steps.

(\supseteq) :

Suppose $x \in C$.

\therefore for all $y \in C^\circ$, $y^T x \leq 1$

$\therefore y^T x = (y^T x)^T = x^T y$

\therefore for all $y \in C^\circ$, $x^T y \leq 1$

$\therefore x \in (C^\circ)^\circ$

$\therefore (C^\circ)^\circ \supseteq C$.

(\subseteq) :

Suppose there exists at least one element $x \in (C^\circ)^\circ$ but $x \notin C$.

According to the Separating Hyperplane Theorem, there exists a separating hyperplane $a^T x = b$ between the non-intersecting sets C and $(C^\circ)^\circ - C$.

By adjusting the value of a , we can scale down the value of b to 1, so we simply assume the hyperplane here is $a^T x = 1$

$\therefore 0 \in C$ and C is a closed set

\therefore for all $z \in C$, $a^T z \leq 1$; for all $w \in (C^\circ)^\circ$, $a^T w > 1$

$\therefore a \in C^\circ$

However, the supposition we made before indicates that there exists one element $x \in (C^\circ)^\circ - C$, and for x we have $a^T x > 1$, and this contradicts with $x \in (C^\circ)^\circ$

Therefore $(C^\circ)^\circ - C = \emptyset$.

So $(C^\circ)^\circ \subseteq C$

In conclusion, $(C^\circ)^\circ = C$.

□

Problem 4: Operations That Preserve Convexity

Suppose $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\psi : \mathbb{R}^m \rightarrow \mathbb{R}^p$ are the linear-fractional functions

$$\phi(x) = \frac{Ax + b}{c^\top x + d}, \psi(y) = \frac{Ey + f}{g^\top y + h}, \quad (4)$$

with domains $\mathbf{dom} \phi = \{x | c^\top x + d > 0\}$, $\mathbf{dom} \psi = \{y | g^\top y + h > 0\}$. We associate with ϕ and ψ the matrices

$$\begin{bmatrix} A & b \\ c^\top & d \end{bmatrix}, \begin{bmatrix} E & f \\ g^\top & h \end{bmatrix}, \quad (5)$$

respectively.

Now, consider the composition Γ of ϕ and ψ , i.e., $\Gamma(x) = \psi(\phi(x))$, with domain

$$\mathbf{dom} \Gamma = \{x \in \mathbf{dom} \phi | \phi(x) \in \mathbf{dom} \psi\}. \quad (6)$$

Show that Γ is linear-fractional, and that the matrix associate with it is the product

$$\begin{bmatrix} E & f \\ g^\top & h \end{bmatrix} \begin{bmatrix} A & b \\ c^\top & d \end{bmatrix}. \quad (7)$$

Solution.

$$\begin{aligned} \therefore \Gamma(x) &= \psi(\phi(x)) \\ &= \frac{E \frac{Ax+b}{c^\top x + d} + f}{g^\top \frac{Ax+b}{c^\top x + d} + h} \\ &= \frac{(EA + fc^\top)x + Eb + fd}{(g^\top A + hc^\top)x + g^\top b + hd} \end{aligned} \quad (8)$$

So apparently Γ is linear-fractional and it can be associated with the matrix

$$\begin{bmatrix} EA + fc^T & Eb + fd \\ g^T A & hc^T \end{bmatrix}, \quad (9)$$

which is exactly the product

$$\begin{bmatrix} E & f \\ g^T & h \end{bmatrix} \begin{bmatrix} A & b \\ c^T & d \end{bmatrix}. \quad (10)$$

Q.E.D.

□

Problem 5: Generalized Inequalities

Let K^* be the dual cone of a convex cone K . Prove the following

- 1) K^* is indeed a convex cone.
- 2) $K_1 \subseteq K_2$ implies $K_2^* \subseteq K_1^*$.

Solution.

- 1) proof:

\therefore for any $y_1, y_2 \in K^*$, we have $x^T y_1 \geq 0, x^T y_2 \geq 0$ (for any $x \in K$).

\therefore for any $x \in K$ and $\theta_1, \theta_2 \in [0, +\infty]$, $x^T(\theta_1 y_1 + \theta_2 y_2) = \theta_1 x^T y_1 + \theta_2 x^T y_2 \geq 0$.

$\therefore \theta_1 y_1 + \theta_2 y_2 \in K^*$

Therefore, K^* is indeed a cone.

Q.E.D.

- 2) proof:

For any $y \in K_2^*$, we always have $x^T y \geq 0$ where x is arbitrarily chosen from K_2 .

Considering $K_1 \subseteq K_2$, $x^T y \geq 0$ still holds when x is from K_1 .

So for any $y \in K_2^*$, we always have $x^T y \geq 0$ where x is arbitrarily chosen from K_1 , which means $y \in K_1^*$

Therefore, $K_2^* \subseteq K_1^*$

Q.E.D.

□