Optimization Methods **Fall 2019**

Homework 1

Instructor: Lijun Zhang *Name:* Weikang Li, *StudentId:* 181220031

Notice

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

Problem 1: Norms

A function $f : \mathbb{R}^n \to \mathbb{R}$ 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

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

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

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

.

(*Hint:* Use Cauchy–Schwarz inequality.)

Solution. a). To prove a valid norm, we need to validate the character according to the definition.

- f is nonnegative: Without loss of generality, we let $x = 0$, it is obvious that $\|z\|_* = \sup\{z^T x | \|x\| \leq 1\} \geq 0$ for all $z \in \mathbb{R}^n$
- f is definite: $f(z) = 0$ means $\sup\{z^T x | \|x\| \leq 1\} = 0$, so $z^T x \leq 0$, it is easy to see that this will hold only if $x = 0$
- f is homogeneous: From above we have f is nonnegative, so:

$$f(tz) = \|tz\|_* = \sup\{(tz)^T x | \|x\| \leq 1\} = |t| \sup\{z^T x | \|x\| \leq 1\} = |t|\|z\|_* = |t|f(z)$$

- f satisfies the triangle inequality:

$$\begin{aligned} f(y + z) &= \sup\{(y + z)^T x | \|x\| \leq 1\} \\ &= \sup\{y^T x + z^T x | \|x\| \leq 1\} \\ &\leq \sup\{y^T x | \|x\| \leq 1\} + \sup\{z^T x | \|x\| \leq 1\} \\ &= f(y) + f(z) \end{aligned}$$

, for all $y, z \in \mathbb{R}^n$

b). To prove $\|z\|_{2*} = \sup\{z^T x | \|x\|_2 \leq 1\} = \sup\{|z^T x| | \|x\|_2 \leq 1\} = \|z\|_2$:
According to the Cauchy-Schwarz inequality, we have:

$$|z^T x| \leq \|z\|_2 \|x\|_2 \leq \|z\|_2.$$

So we just need to prove that the maximum value of $|z^T x|$ is $\|z\|_2$. Without loss of generality, we let $x = kz$:
if $k = 0$, trivial.
if $k \neq 0$, it is obvious that $\|z\|_2 \|x\|_2 = \|z\|_2$ only if $\|x\|_2 = \|kz\|_2 = k\|z\|_2 = 1$, so $k = \frac{1}{\|z\|_2}$.

□

Problem 2: Affine and Convex Sets

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

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

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

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

b) Determine if each set below is convex.

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

Solution. a). Not affine: We suppose that α_1, α_2 in the set, without loss of generality we set $\alpha_{1,1}, \alpha_{2,1} = 1$,then we have :

$$|\alpha_{1,1} + \alpha_{1,2} t + \cdots + \alpha_{1,k} t^{k-1}| \leq 1$$
$$|\alpha_{2,1} + \alpha_{2,2} t + \cdots + \alpha_{2,k} t^{k-1}| \leq 1$$

for $\theta \in \mathbb{R}$:

$$\theta \alpha_1 + (1-\theta)\alpha_2 = \theta(\alpha_{1,1}, \alpha_{1,2}, \cdots, \alpha_{1,k}) + (1-\theta)(\alpha_{2,1}, \alpha_{2,2}, \cdots, \alpha_{2,k})$$

in this case,

$$|p(t)| = |\theta|\alpha_{1,1} + \alpha_{1,2} t + \cdots + \alpha_{1,k} t^{k-1}| + (1-\theta)|\alpha_{2,1} + \alpha_{2,2} t + \cdots + \alpha_{2,k} t^{k-1}||$$

let $\theta = 2$, $|\alpha_{1,1} + \alpha_{1,2} t + \cdots + \alpha_{1,k} t^{k-1}| = 1$, $|\alpha_{2,1} + \alpha_{2,2} t + \cdots + \alpha_{2,k} t^{k-1}| = -1$, it is obvious that $|p(t)| > 1$, so $\theta\alpha_1 + (1-\theta)\alpha_2$ is not in the set, which means the set is not convex.

b). 1) Convex;
2) Convex;
3) Not convex: Consider a combination z of two points $x(\frac{1}{2}, 2)$ and $y(2, \frac{1}{2})$ in the set, and let $\theta = \frac{1}{2}$. Therefore, from $z = \theta x + (1-\theta) y$ we have $z(\frac{5}{4}, \frac{5}{4})$. It is obvious that $z \notin \{(x, y) \in \mathbf{R}^2_{+} | xy \leq 1\}$
4) Convex;
5) Not convex: Consider a combination z of two points $x(0, 0)$ and $y(-\infty, -1)$ in the set. Therefore, from $z = \theta x + (1-\theta) y$. It is obvious that for any $0 \leq \theta \leq 1$, $z \notin \{(x, y) \in \mathbf{R}^2 | y = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}\}$.

□

Problem 3: Examples

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

$$C = \{x \in \mathbb{R}^n | x^\top Ax + b^\top x + c \leq 0\}, \tag{3}$$

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 $A + \lambda gg^\top \succeq 0$ 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) To prove the set C is convex, an approach is to prove the intersection of C and any lines is convex. Let the line is defined as the set $\{x_0 + tv \,|\, t \in \mathbb{R}\}$, $v, x_0 \in \mathbb{R}^n$.
Therefore, we have:

$$(x_0 + tv)^T A(x_0 + tv) + b^T(x_0 + tv) + c \leq 0$$

let:

$$\alpha = v^T Av, \quad \beta = b^T v + 2x_0^T Av, \quad \gamma = c + b^T x_0 + x_0^T Ax_0$$

the intersection is:

$$\{x_0 + tv \,|\, \alpha t^2 + \beta t + \gamma \leq 0\}$$

If the line intersects the set, the above inequality has a solution. Let's consider only the case where t is solvable.
if $A \succeq 0$, so $\alpha = v^T at \geq 0$:
when $\alpha = 0$:

- $\beta = 0,\ \gamma \leq 0 : t \in \mathbb{R}$
- $\beta \geq 0 : t \leq \frac{-\gamma}{\beta}$
- $\beta \leq 0 : t \geq \frac{-\gamma}{\beta}$

when $\alpha > 0$: $t_1 \leq t \leq t_2$ (t_1, t_2 is two roots of the parabola)
Therefore, the intersection of C and any lines defined above is convex, so C is convex.
2) Let the set of hyperplanes $H = \{x \,|\, g^T x + h = 0\}$, we define $\delta = g^T v$, $\epsilon = g^T x_0 + h$. Assuming that $x_0 \in H$, which means $\epsilon = g^T x_0 + h = 0$.
So, the intersection of $C \cap H$ and lines defined above is:

$$\{x_0 + tv \,|\, \alpha t^2 + \beta t + \gamma \leq 0,\ \delta t = 0\}$$

If $\delta \neq 0$, so $t = 0$, the intersection is $\{x\}$;
If $\delta = 0$, the intersection is $\{x_0 + tv \,|\, \alpha t^2 + \beta t + \gamma \leq 0\}$, from above we have: this is convex if $\alpha \geq 0 \Rightarrow v^T Av \geq 0$.
This will hold if $A + \lambda gg^\top \succeq 0$ for some $\lambda \in \mathbb{R}$, which means:

$$v^T Av = v^T(A + \lambda gg^\top)v \geq 0$$

b) 1) According to the definition, it is obvious that the polar is the intersection of halfspaces $\{y \,|\, y^T x \leq 1\}$, so it is convex.
2) According to the definition of the polar and polyhedra, it is easy to see that the polar of a polyhedra is still

a polyhedra.
3) According to the definition of the unit ball and the polar of a set:

$$\mathcal{B} = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\},$$

$$\therefore \mathcal{B}^\circ = \{y \in \mathbb{R}^n | y^\top x \leq 1 \text{ for all } x \in \mathcal{B}\} = \{y \in \mathbb{R}^n | \sup\{y^T x | \|x\| \leq 1\} \leq 1\} = \{y \in \mathbb{R}^n | \|y\|_* \leq 1\}$$

4) Assume that $x \in C$ and $y \in C^\circ$, so $y^T x \leq 1$, also we have $x^T y \leq 1$ for all $y \in C^\circ$, which means $x \in (C^\circ)^\circ$, so $C \subseteq (C^\circ)^\circ$.
Assume that $x \in (C^\circ)^\circ$ and $x \notin C$, According to the Separation Theorem of Hyperplane, there must be a seperating hyperplane for C and $\{x\}$, which means for $z \in C$, $a^T z \leq b$; $a^T x > b$; because $0 \in C$, we have $b \geq 0$. Without loss of generality, we let $z \in C$, $a^T z \leq 1$; $a^T x > 1$. Therefore, $a \in C^\circ$. From the assumption above, we have $x \in (C^\circ)^\circ$, which means $x^T a \leq 1$ *i.e.* $a^T x \leq 1$, which is contradicted with hypothesis.
Therefore, $(C^\circ)^\circ = C$

□

Problem 4: Operations That Preserve Convexity

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

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

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

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

respectively.

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

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

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

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

Solution. According to the definition: $\Gamma(x) = \psi(\phi(x))$, therefore we have, for $x \in$ **dom** Γ,

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

As can be seen from the form of the upper form, Γ is linear-fractional function, and associated with the product matrix:

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

□

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. a) According to the definition of K^*: $\{y \mid x^T y \geq 0, \forall x \in K\}$ we found that K^* is the intersection of a set of homogeneous halfspaces (all halfspaces are convex) . Therefore, K^* is indeed a convex cone.

b) According to the definition : $y \in K_2^*$ infers $x^T y \geq 0$ for all $x \in K_2$. Meanwhile, $K_2 \supseteq K_1$, which infers $x^T y \geq 0$ for all $x \in K_1$. Therefore, $K_2^* \subseteq K_1^*$.

$\square$