Optimization Methods **Fall 2019**

Homework 1

Instructor: Lijun Zhang *Name:* Renzhe Zhou, *StudentId:* 181220076

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 $f : \mathbb{R}^n \leftarrow \mathbb{R}$ with $\mathbf{dom} f = \mathbb{R}^n$ 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 $x, y \in v$

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. Question a):
Make g(z)=$\|z\|_*$, $\sup\{z^T x\} = sup\{\sum_{i=1}^n z_i x_i\}$, s.t. $0 \leq \|x\| \leq 1$. That is, to maximize it if $z_i < 0$ then $x_i \leq 0$ too, if $z_i \geq 0$ then $x_i \geq 0$ too. Therefore, $\sum_{i=1}^n z_i x_i \geq 0$, which means $\|\cdot\|_*$ is nonnegative.
$sup\{\sum_{i=1}^n z_i x_i\} = 0$ and $z_i x_i \geq 0$, so $z_i x_i = 0$ for any x, so $z_i = 0$, so $z = 0$. And when z=0, obviously $g(z) = 0$. Therefore, $g(z) = 0$ if and only is $z = 0$.
$g(tz) = \sup\{tz^T x | \|x\| \leq 1\} = |t| sup\{z^T x | \|x\| \leq 1\} = |t| g(z)$,so g is homogeneous.
To prove $g(z_1 + z_2) \leq g(z_1) + g(z_2)$, is to prove $\sup((z_1 + z_2)^T x) \leq \sup(z_1^T x) + \sup(z_2^T x), s.t. 0 \leq \|x\| \leq 1$,assume they are $(z_1 + z_2)^T x_0, z_1^T x_1, z_2^T x_2$ then we need to prove $z_1^T x_0 + z_2^T x_0 \leq z_1^T x_1 + z_2^T x_2$, because x_1 and x_2 respectively maximize $z_1^T x$ and $z_2^T x$, so $z_1^T x_0$ and $z_2^T x_0$ will be no more than $z_1^T x_1$ and $z_2^T x_2$, so $g(z_1 + z_2) \leq g(z_1) + g(z_2)$.
Above all, $\|\cdot\|_*$ is a norm.
Question b):
Because $|z^T x| \leq \|z\|_2 \|x\|_2$, so $\sup\{z^T x | \|x\|_2 \leq 1\} \leq \|z\|_2 \|x\|_2$, I am going to prove $\sup\{z^T x | \|x\|_2 \leq 1\} = \|z\|_2 \|x\|_2$.
Let's see if $\sup\{z^T x | \|x\|_2 \leq 1\}$ can up to $\|z\|_2 \|x\|_2$. Conditions for Cauchy–Schwarz inequality to hold is $x = kz$ in which $k \in R$.
When z=0, Obviously $\sup\{z^T x | \|x\|_2 \leq 1\} = \|z\|_2 \|x\|_2 = 0$.
When $z \neq 0$, $\|x\|_2 \leq 1 \rightarrow \|kz\|_2 \leq 1 \rightarrow |k| \leq \frac{1}{\|z\|_2}$, suck a k is available. So $\sup\{z^T x | \|x\|_2 \leq 1\} = \|z\|_2 \|x\|_2$ when x is kz, and $\|x\|_2 \leq 1$, to maximize it $\|x\|_2$ should be 1. Therefore $\sup\{z^T x | \|x\|_2 \leq 1\} = \|z\|_2 \|x\|_2 = \|z\|_2$. □

Problem 2: Affine and Convex Sets

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

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

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

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. Question a):
Make $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k)^T$,$T = (1, t, t^2, \cdots, t^{(k-1)})^T$,so $p(t) = \alpha^T T$,assume m,n in the set, we have $m_1 = 1, n_1 = 1$, $|m^T T| \leq 1, |n^T T| \leq 1$, make $k = \theta m + (1-\theta)n$, then $k_1 = \theta m_1 + (1-\theta)n_1 = \theta + 1 - \theta = 1$, and $|k^T T| = |\{\theta m + (1-\theta)n\}^T T| = |\{\theta m^T + (1-\theta)n^T\}T| = |\theta m^T T + (1-\theta)n^T T|$, let's take $m^T T = 1, n^T T = -1$ and $\theta = 100$, then it becomes $|\theta m^T T| + |(1-\theta)n^T T| = \theta m^T T + (\theta - 1)|n^T T| = 199$, so it is not in the set, which means the set is not affine.
Question b):
1):Assume $(x_1, y_1), (x_2, y_2)$ in the set, we have $x_1 \leq y_1, x_2 \leq y_2$, so for $(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2)$, we have $\frac{\theta x_1 + (1-\theta)x_2}{\theta y_1 + (1-\theta)y_2} \leq \frac{\theta y_1 + (1-\theta)y_2}{\theta y_1 + (1-\theta)y_2} = 1$, so it is convex.
2):The solution is similar to 1), and it is convex too.
3):Assume $(x_1, y_1), (x_2, y_2)$ in the set, we have $x_1 y_1 \leq 1, x_2 y_2 \leq 1$, then $((\theta x_1 + (1-\theta)x_2)(\theta y_1 + (1-\theta)y_2)) = \theta^2 x_1 y_1 + \theta(1-\theta)(x_1 y_2 + x_2 y_1) + (1-\theta)^2 x_2 y_2$, when $\theta = \frac{1}{2}, x_1 = 10, y_1 = \frac{1}{10}, x_2 = \frac{1}{10}, y_2 = 10$, it is equal to $25 + \frac{201}{400}$, so it is not convex.
4):the solution is similar to 3), and it is not convex too.
5):if the set is convex, then $y = tanh(x)$ should be linear additive, that is to say $\tanh(ax_1 + bx_2) = a\tanh(x_1) + b\tanh(x_2)$, obviously this is not a linear addictive function, so the set is not convex.

□

Problem 3: Examples

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

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

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 affine.
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. Question a):
1):A set is convex implies its intersection with any line is convex.
We take a line $\{x' + tv|t \in \mathbb{R}\}$, and make $(x'+tv)^\top A(x'+tv) + b^\top(x'+tv) + c = \alpha t^2 + \beta t + \gamma$ where $\alpha = v^T Av, \beta = b^T v + 2x'^T Av, \gamma = x'^T Ax' + b^T x' + c$.
Its intersection with C becomes $\{x' + tv|\alpha t^2 + \beta t + \gamma \le 0\}$.
When $\alpha \ge 0$, it is convex because the solution would be a line segment or a dot or empty. When $\alpha < 0$, it is not convex because the solution would be two rays.
And because $A \succeq 0$, so $\alpha = v^T Av \ge 0$ for all v, so C is convex.
2):The statement is true.
Let $D = \{x \in \mathbb{R}^n | g^T x + h = 0\}, E = C \cap D$, and in line $\{x' + tv|t \in \mathbb{R}\}$ we take $x' \in D$ so $g^T x' + h = 0$. Then line $\{x' + tv|t \in \mathbb{R}\}$'s intersection with E is $\{x' + tv|\alpha t^2 + \beta t + \gamma \le 0, g^T(x'+tv) + h = 0\}$, because $g^T(x'+tv) + h = g^T x' + h + g^T v = g^T vt = 0$, the intersection becomes

$$\{x' + tv|\alpha t^2 + \beta t + \gamma \le 0, g^T vt = 0\} \tag{4}$$

If $g^T v \ne 0$, then t=0, which means the intersection solution is a dot or empty, in this case E is convex.
If $g^T v = 0$, then we are back to case 1): $\{x' + tv|\alpha t^2 + \beta t + \gamma \le 0\}$, and $v^T Av = v^T Av + 0 = v^T Av + v^T g = v^T Av + v^T gg^T v = v^T Av + \lambda v^T gg^T v = v^T(A + \lambda gg^T)v$, because $(A + \lambda gg^T) \succeq 0$, so $v^T(A + \lambda gg^T)v \ge 0$, so in this case E is convex too.
Above all, E is convex.
Question b):
1):C° is the intersection of a number of halfspaces, so it is convex.
2):The polar of a polyhedra is unit simplex.
3):The unit ball of $\|x\|$ is $\{x \in \mathbb{R}^n | \|x\| \le 1\}$, the dual norm is $\|z\|_* = \sup\{z^T x| \|x\| \le 1\}$, and the unit ball of $\|z\|_*$ is $\{z \in \mathbb{R}^n | \|z\|_* \le 1\}$, which equivalents to $\{z \in \mathbb{R}^n | z^T x \le 1, \|x\| \le 1\}$, and that is the polar of $\{x \in \mathbb{R}^n | \|x\| \le 1\}$, so the polar of norm $||\cdot||$ is the dual norm's unit ball.
4):Make $K = C^\circ = \{y \in \mathbb{R}^n | y^\top x \le 1 \text{ for all } x \in C\}$,then $K^\circ = \{z \in \mathbb{R}^n | z^\top y \le 1 \text{ for all } y \in K\}$, which can be written as $\{z \in \mathbb{R}^n | z^\top y \le 1, x^\top y \le 1 \text{ for all } x \in C\}$, we can guess z=x, and this assumption is true because C is closed and convex and $0 \in C$, so such a transformation keeps all x in C remaining as z. Therefore, $K^\circ = C$, so $(C^\circ)^\circ = C$, and we are done.

□

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

$$\phi(x) = \frac{Ax+b}{c^\top x + d}, \psi(y) = \frac{Ey+f}{g^\top y + h}, \tag{5}$$

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

$$\begin{bmatrix} A & b \\ c^\top & d \end{bmatrix}, \begin{bmatrix} E & f \\ g^\top & h \end{bmatrix}, \tag{6}$$

respectively.

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

$$\mathbf{dom}\Gamma = \{x \in \mathbf{dom}\ \phi | \phi(x) \in \mathbf{dom}\ \psi\}. \tag{7}$$

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

$$\begin{bmatrix} E & f \\ g^\top & h \end{bmatrix} \begin{bmatrix} A & b \\ c^\top & d \end{bmatrix}. \tag{8}$$

Solution. Firstly we have $A \in \mathbb{R}^{m\times n}, E \in \mathbb{R}^{n\times m}, b, g \in \mathbb{R}^m, f, c \in \mathbb{R}^n, d, h \in \mathbb{R}$.

$$\psi(\phi(x)) = \frac{E\phi(x) + f}{g^\top \phi(x) + h} = \frac{E\frac{Ax+b}{c^\top x+d} + f}{g^\top \frac{Ax+b}{c^\top x+d} + h} = \frac{E(Ax+b) + f(c^\top x + d)}{g^\top(Ax+b) + h(c^\top x + d)} = \frac{(EA + fc^\top)x + Eb + fd}{(g^\top A + hc^\top)x + g^\top b + hd} \tag{9}$$

and $EA + fc^\top \in \mathbb{R}^{n\times n}, Eb + fd \in \mathbb{R}^n, g^\top A + hc^\top \in \mathbb{R}^n, g^\top b + hd \in \mathbb{R}$. So $\psi(\phi(x))$ is linear-fractional. Because

$$\begin{bmatrix} E & f \\ g^\top & h \end{bmatrix} \begin{bmatrix} A & b \\ c^\top & d \end{bmatrix} = \begin{bmatrix} EA + fc^\top & Eb + fd \\ g^\top A + hc^\top & g^\top b + hd \end{bmatrix} \tag{10}$$

so the the matrix associate with it is the product above.

□

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 $K_1^* \subseteq K_2^*$.

Solution. Question 1):
$K^* = \{y | x^T y \geq 0, \forall x \in K\}$, which means K^* is the intersection of a set of homogeneous halfspaces, so it is a convex cone.
Question 2):
Assume $y \in K_2^*$, because $x^T y \geq 0, \forall x \in K_2$ and $K_1 \subseteq K_2$, so $x^T y \geq 0, \forall x \in K_1$, which means $y \in K_1^*$, so $K_1^* \subseteq K_2^*$.

$\square$