Optimization Methods **Fall 2023**

Homework 1

Instructor: Lijun Zhang *Name:* Ji Qiankun, *StudentId:* 221300066

Problem 1: Norms

A function $f : \mathbb{R}^n \to \mathbb{R}$ with $\operatorname{dom} f = \mathbb{R}^n$ 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

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

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

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

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

Solution. a) To prove that $\|\cdot\|_*$ is a valid norm, need to show that it satisfies the four properties of a norm:

1. Non-negativity: By definition, $\|z\|_* = \sup_{\|x\|\leq 1} z^{\mathrm{T}} x$. Since the supremum of a set of real numbers is always greater than or equal to any element of the set, we have $z^{\mathrm{T}} x \leq \|z\|_*$ for all x with $\|x\| \leq 1$. In particular, setting $x = 0$, we have $z^{\mathrm{T}} 0 = 0 \leq \|z\|_*$, so $\|\cdot\|_*$ is non-negative.

2. Definiteness: If $z = 0$, then clearly $\|\cdot\|_*(z) = 0$. Conversely, suppose that $\|\cdot\|_*(z) = 0$. Then by definition, $z^{\mathrm{T}} x \leq 0$ for all x with $\|x\| \leq 1$. In particular, setting $x = z/\|z\|$ (which has norm $\|x\| = 1$), we have $z^{\mathrm{T}}(z/\|z\|) = \|z\|^2/\|z\| = \|z\|$ (since $\|\cdot\|$ is a norm), so $\|z\| = 0$, which implies that $z = 0$.

3. Homogeneity: For any scalar α, we have

$$\|\alpha z\|_* = \sup_{\|x\|\leq 1} (\alpha z)^{\mathrm{T}} x = \sup_{\|y\|\leq|\alpha|} z^{\mathrm{T}}(y/\alpha) = |\alpha| \sup_{\|y/\alpha\|\leq 1} z^{\mathrm{T}}(y/\alpha) = |\alpha| \, \|\cdot\|_*(z).$$

4. Triangle inequality: For any vectors z, w, we have

$$\|z+w\|_* = \sup_{\|x\|\leq 1} (z+w)^{\mathrm{T}} x = \sup_{\substack{\|y_1+y_2\|\leq 1\\ y_1,y_2}} z^{\mathrm{T}} y_1 + w^{\mathrm{T}} y_2 \leq \sup_{y_1}\left(z^{\mathrm{T}} y_1 + \sup_{y_2}\left(w^{\mathrm{T}} y_2 : \|y_2\| \leq 1 - \|y_1\|\right)\right) = \|z\|_* + \|w\|_*.$$

Therefore, it is shown that $\|\cdot\|_*$ is therefore a valid norm.

b) To prove that the dual of the Euclidean norm is the Euclidean norm itself, need to show that

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

By Cauchy-Schwarz inequality, we have

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

where the last inequality follows since $\|x\| = 1$. Therefore, we have

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

□

Problem 2: Inequalities

Let $x \in \mathbb{R}^n, y \in \mathbb{R}^n$, where n is a positive integer. Let $\|\cdot\|$ denote the Euclidean norm.

a) Prove the triangle inequality $\|x+y\| \leq \|x\| + \|y\|$.

b) Prove $\|x+y\|^2 \leq (1+\epsilon)\|x\|^2 + (1+\frac{1}{\epsilon})\|y\|^2$ for any $\epsilon > 0$.

(*Hint*: You may need the Young's inequality for products, i.e. if a and b are nonnegative real numbers and p and q are real numbers greater than 1 such that $1/p + 1/q = 1$, then $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.)

Solution. a) First square both sides and then expand:

$$\begin{aligned}\|x+y\|^2 &= (x+y)^T(x+y)\\ &= x^Tx + 2x^Ty + y^Ty\\ &= \|x\|^2 + 2x^Ty + \|y\|^2.\end{aligned}$$

Since $x^Ty \leq \|x\|\|y\|$ by the Cauchy-Schwarz inequality, we have

$$\|x+y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

Taking the square root of both sides, then get the triangle inequality $\|x+y\| \leq \|x\| + \|y\|$
b) Using Young's inequality. Let $a = \sqrt{\epsilon}\|x\|$ and $b = \frac{1}{\sqrt{\epsilon}}\|y\|$ in Young's inequality, then we have

$$ab \leq \frac{a^2}{2} + \frac{b^2}{2},$$

which is equivalent to

$$\sqrt{\epsilon}\|x\|\frac{1}{\sqrt{\epsilon}}\|y\| \leq \frac{\epsilon}{2}\|x\|^2 + \frac{1}{2}\|y\|^2.$$

Rearranging terms gives us

$$(1+\epsilon)\|x\|^2 + (1+\frac{1}{\epsilon})\|y\|^2 - 2\|x\|\|y\| \geq 0,$$

Since $\|x+y\|^2 = \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$, it is equivalent to the desired inequality $\|x+y\|^2 \leq (1+\epsilon)\|x\|^2 + (1+\frac{1}{\epsilon})\|y\|^2$ for any $\epsilon > 0$.

$\square$

Problem 3: Definition of convexity

Which of the following sets are convex? Please provide explanations for your choices.

a) a set of the form $\{x \in \mathbf{R}^n \mid \alpha \leq a^Tx \leq \beta\}$.

b) a set of the form $\{x \in \mathbf{R}^n \mid \alpha_i \leq a^Tx_i \leq \beta_i, i = 1, 2, ..., n\}$.

c) a set of the form $\{x \in \mathbf{R}^n \mid a_1^Tx \leq b_1, a_2^Tx \leq b_2\}$.

d) The set of points closer to a given point than a given set, *i.e.*,

$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$$

where $S \subseteq \mathbf{R}^n$.

e) The set of points closer to one set than another, *i.e.*,

$$\{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\},$$

where $S, T \subseteq \mathbf{R}^n$, and

$$\mathbf{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}.$$

f) The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, *i.e.*, the set $\{x \mid \|x - a\|_2 \leq \theta\|x - b\|_2\}$. You can assume $a \neq b$ and $0 \leq \theta \leq 1$.

Solution. a) The set is convex. This is because it represents a hyperplane in $\mathbf{R}^n$, and all points between any two points in the set also belong to the set.

b) The set is not necessarily convex. This is because each constraint applies to a different dimension of x, and there may exist two points in the set such that the line segment connecting them does not entirely belong to the set.

c) The set is convex. This is because it represents the intersection of two half-spaces, which are both convex sets.

d) The set is convex. This is known as a Voronoi cell, which is always convex.

e) The set is convex. This is because it represents the region of space where every point is closer to the set S than to the set T, which forms a convex region.

f) The set is not convex when $\theta = 1$ and $a \neq b$. This is because it forms an ellipse with foci at a and b, but excludes the segment between a and b. However, for other values of θ, it forms a convex region.

□

Problem 4: Examples

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

$$C = \{x \in \mathbb{R}^n | x^T Ax + b^T x + c \leqslant 0\}$$

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

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

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

Solution. a) The function $f(x) = x^T Ax + b^T x + c$ is a convex function when A is positive semi-definite (i.e., $A \succeq 0$). The sublevel set of a convex function, which is the set of points for which the function value is less than or equal to a certain value, is always a convex set. Therefore, the set C is convex if $A \succeq 0$.

b) It is true. This is because the intersection of two convex sets is always a convex set. The set C is convex as shown in a), and the hyperplane defined by $g^T x + h = 0$ is also a convex set. Therefore, their intersection is convex. Furthermore, if $A + \lambda gg^T \succeq 0$, then the function defining the set C remains convex, so the intersection of C with any hyperplane (including the one defined by $g^T x + h = 0$) remains convex.

□

Problem 5: Dual cones

Describe the dual cone for each of the following cones.

a) $K = \{0\}$.

b) $K = \mathbf{R}^2$.

c) $K = \{(x_1, x_2) \, || x_1 | \leq x_2\}$.

d) $K = \{(x_1, x_2) \mid x_1 + x_2 = 0\}$.

Solution. The dual cone of a cone K is defined as the set of all vectors y such that $y^T x \leq 0$ for all x in K.
a) the dual cone is R^n, where n is the dimension of the space.
b) the dual cone is also R^2.
c) the dual cone is $\{(y1, y2) | y1 >= 0 \text{ and } y2 >= 0\}$.
d) the dual cone is $\{(y1, y2) | y1 - y2 >= 0 \text{ and } y1 + y2 >= 0\}$

□