Fenchel Duality

3.1 Subgradients and Convex Functions

Ryota Iwamoto

February 7, 2023

We use the book; Convex Analysis and Nonlinear Optimization (author: J.M.BORWEIN and A.S.LEWIS), pp.33-36.

> We have already seen, in the First order sufficient condition (2.1.2), one benefit of convexity in optimization: critical points of convex functions are global minimizers. In this section we extend the types of functions we consider in two important ways:
>
> (i) We do not require f to be differentiable.
>
> (ii) We allow f to take the value $+\infty$.

This book Chapter 2 explains a optimization of convex functions with good conditions, which is differentiable and not including infinity points. In this section, we consider extended functions like being not differentiable and allowed to take the value $+\infty$.

> Our derivation of first order conditions in Section 2.3 illustrates the utility of considering nonsmooth functions even in the context of smooth problems. Allowing the value $+\infty$ lets us rephrase a problem like
>
> $$\inf \{g(x) \mid x \in C\}$$
>
> as $\inf (g + \delta_C)$, where the indicator function $\delta_C(x)$ is 0 for x in C and $+\infty$ otherwise.

Here We consider the definition of indicator function.

Definition (Indicator Funciton)

The indicator function of a set C of $\mathbb{E}$, denoted by δ_C, is defined by

$$\delta_C(x) = \begin{cases} 0 & if\ x \in C, \\ \infty & if\ otherwise. \end{cases}$$

The domain of a function $f : \mathbb{E} \to (-\infty, +\infty]$ is the set

$$dom f = \{x \in \mathbb{E} \mid f(x) < +\infty\}.$$

We say f is convex if it is convex on its domain, and proper if its domain is nonempty. We call a function $g : \mathbb{E} \to [-\infty, +\infty)$ concave if $-g$ is convex, although for reasons of simplicity we will consider primarily convex functions. If a convex function f satisfies the stronger condition

$$f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y),\ for\ all\ x, y \in \mathbb{E},\ \lambda, \mu \in \mathbb{R}_+$$

we say f is *sublinear*. If $f(\lambda x) = \lambda f(x)$ for all x in $\mathbb{E}$ and λ in $\mathbb{R}_+$ then f is *positively homogeneous*: in particular this implies $f(0) = 0$. (Recall the convention $0 \cdot (+\infty) = 0$.) If $f(x + y) \leq f(x) + f(y)$ for all x and y in $\mathbb{E}$ then we say f is *subadditive*. It is immediate that if the function f is sublinear then $-f(x) \leq f(-x)$ for all x in $\mathbb{E}$. The *lineality space* of a sublinear function f is the set

$$lin f = \{x \in \mathbb{E} \mid -f(x) = f(-x)\}.$$

We describe some definitions and the figure below.

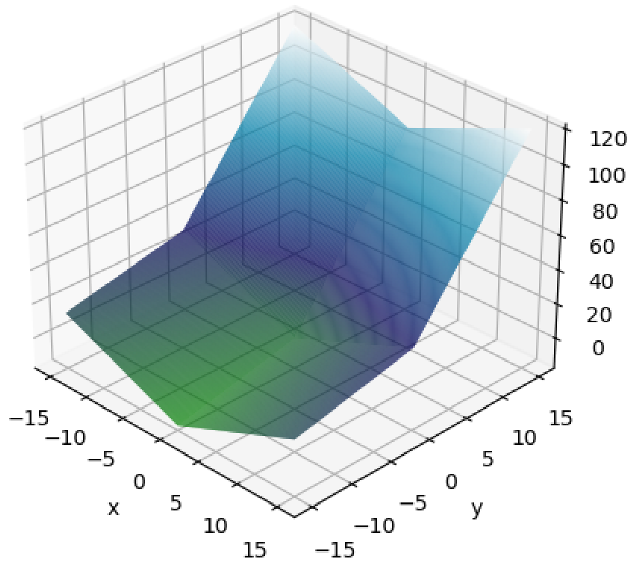
Definition (Sublinear)

A convex function f is **sublinear** if this f satisfies the condition

$$f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y),\ for\ all\ x, y \in \mathbb{E},\ \lambda, \mu \in \mathbb{R}_+.$$

Figure:

$$f(x, y) = \begin{cases} 2|x| + y & if\ y \leq 0, \\ 2|x| + 6y & if\ y > 0. \end{cases}$$



Definition (Positively Homogeneous and Subadditive)

A convex function f is **positively homogeneous** if this f satisfies the condition

$$f(\lambda x) = \lambda f(x), \ for\ all\ x \in \mathbb{E},\ \lambda \in \mathbb{R}_+.$$

And a convex function f is **subadditive** if this f satisfies the condition

$$f(x+y) \leq f(x) + f(y), \ for\ all\ x, y \in \mathbb{E}.$$

Proposition

If the function f is sublinear then $-f(x) \leq f(-x)$ for all x in $\mathbb{E}$.

Proof. We show that it holds $-f(x) \leq f(-x)$ for all x in $\mathbb{E}$.

By the assumption of sublinear of f, we can use the inequality

$$f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y), \ for\ all\ x, y \in \mathbb{E},\ \lambda, \mu \in \mathbb{R}_+.$$

For any x and y, we take $y = -x$ since $\mathbb{E}$ is a vector space, and also take $\lambda = \mu = 1$. These values and the above assumption, we get the inequality

$$f(x + (-x)) \leq f(x) + f(-x).$$

By $f(0) = 0$,

$$0 \leq f(x) + f(-x).$$

Therefore we got $-f(x) \leq f(-x)$ for all x in $\mathbb{E}$. $\square$

Proposition (the former of Proposition 3.1.1)

If the function f is sublinear if and only if f is positively homogeneous and subadditive.

Proof. We show that

$$f \text{ is sublinear} \Leftrightarrow \text{f is positively homogeneous and subadditive.}$$

($\Rightarrow$) By the definition of sublinear of f, for any y we take $y = 0$ and $\mu = 0$. Then f is positively homogeneous.

($\Leftarrow$) By the definition of subadditive and positively homogeneous of f, $\forall x, y \in \mathbb{E}$, $\lambda, \mu \in \mathbb{R}_+$,

$$f(\lambda x + \mu y) \leq f(\lambda x) + f(\mu y) = \lambda f(x) + \mu f(y).$$

Therefore we completed to prove it. $\square$

Definition (Lineality Space)

The *lineality space* of a sublinear function f, denoted by $linf$, is the set

$$linf = \{x \in \mathbb{E} \mid -f(x) = f(-x)\}.$$

The following result (whose proof is left as an exercise) shows this set is a subspace.

Proposition 3.1.1 (Sublinearity)

A function $f : \mathbb{E} \to (-\infty, +\infty]$ is sublinear if and only if it is positively homogeneous and subadditive. For a
sublinear function f, the lineality space $linf$ is the largest subspace of $\mathbb{E}$ on which f is linear.

Proof. The former proposition has already been proved. Accordingly we show the latter one, that the way of proof is two steps;

(I) $linf$ is a subspace of $\mathbb{E}$.
(II) $linf$ is the "largest" subspace of $\mathbb{E}$.

(I) We show that

(i) $\forall x, y \in linf$, $x + y \in linf$.
(ii) $\forall x \in linf$, $\alpha \in \mathbb{R}$, $\alpha x \in linf$.

(i) We prove that $\forall x, y \in linf$,

$$-f(x+y) = f(-x-y).$$

For any x and y, by the definition of linearity space, it holds that

$$-f(x) = f(-x) \text{ and } -f(y) = f(-y).$$

And, Since f is sublinear, we have $f(x+y) \leq f(x) + f(y)$. Using the result of the above proposition about a sublinear property, we can get the inequality:

$$-f(x+y) \geq -f(x) - f(y) = f(-x) + f(-y) \geq f(-x-y) = f(-(x+y)) \geq -f(x+y).$$

This inequality means that $-f(x+y) = f(-x-y)$.

(ii) We show that $\forall x \in linf$, $\alpha \in \mathbb{R}$,

$$-f(\alpha x) = f(-\alpha x).$$

By the definition of positively homogeneous of f and the definition of lineality space,

$$-f(\alpha x) = -\alpha f(x) = \alpha f(-x) = f(-\alpha x).$$

For the result of (i) and (ii), $linf$ is a subspace of $\mathbb{E}$.

(II) We show that for a sublinear function f it holds that

$$\forall V \subset \mathbb{E}, \text{ where } V \text{ is subspace of } \mathbb{E},\ V \subset linf$$
$$\Leftrightarrow \exists V_0 \subset \mathbb{E}, \text{ where } V \text{ is subspace of } \mathbb{E},\ V_0 \not\subset linf.$$

Suppose to the contrary that there exists a subspace V_0 which is not included in $linf$,

$$\exists x_0 \in V_0\ s.t.\ x_0 \notin linf. \tag{1}$$

By (1) and the definition of linearity space,

$$-f(x_0) \neq f(-x_0).$$

However V_0 is a subspace of $\mathbb{E}$ so it holds that $-f(x_0) = f(-x_0)$.

We can find a contradiction to the assumption that V_0 is a subspace of $\mathbb{E}$, therefore $linf$ is the largest subspace of $\mathbb{E}$. □

As in the First order sufficient condition (2.1.2), it is easy to check that if the point $\bar{x}$ lies in the domain of the convex function f then the directional derivative $f'(\bar{x};\cdot)$ is well-defined and positively homogeneous, taking values in $[-\infty, +\infty]$. The core of a set C (written $core(C)$) is the set of points x in C such that for any direction d in $\mathbb{E}$, $x + td$ lies in C for all small real t. This set clearly contains the interior of C, although it may be larger (Exercise 2).

Definition (the core of a set)

Let $C \subset \mathbb{E}$. The **core** of the set C is defined as

$$core(C) := \{x \in \mathbb{E} \mid \forall d \in \mathbb{E}, \exists t_0 > 0 \; s.t. \; \forall t(0 < t < t_0), x + td \in C\} \; .$$

Figure:

Exercise 2 (Core versus interior)

Considering the set in $\mathbb{R}^2$

$$D = \{(x, y) \mid y = 0 \; or \; |y| \geq x^2\} \; .$$

Prove $0 \in core(D) \backslash int(D)$

Proof. □

Proposition 3.1.2 (Sublinearity of the directional derivative)

If the function $f : \mathbb{E} \to (-\infty, +\infty]$ is convex then for any points $\bar{x}$ in $core(dom f)$, the directional derivative $f'(\bar{x}; \cdot)$ is everywhere finite and sublinear.

Proof. □