

Chapter 2

Extended Real-Valued Functions

Underlying Space: Recall that in this book, the underlying spaces (denoted usually by \mathbb{E} or \mathbb{V}) are finite-dimensional inner product vector spaces with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$.

2.1 Extended Real-Valued Functions and Closedness

An *extended real-valued function* is a function defined over the entire underlying space that can take any real value, as well as the infinite values $-\infty$ and ∞ . Since infinite values are allowed, we also define the appropriate arithmetic operations with $-\infty$ and ∞ as follows:

$$\begin{aligned}
 a + \infty &= \infty + a = \infty & (-\infty < a < \infty), \\
 a - \infty &= -\infty + a = -\infty & (-\infty < a < \infty), \\
 a \cdot \infty &= \infty \cdot a = \infty & (0 < a < \infty), \\
 a \cdot (-\infty) &= (-\infty) \cdot a = -\infty & (0 < a < \infty), \\
 a \cdot \infty &= \infty \cdot a = -\infty & (-\infty < a < 0), \\
 a \cdot (-\infty) &= (-\infty) \cdot a = \infty & (-\infty < a < 0), \\
 0 \cdot \infty &= \infty \cdot 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0 = 0.
 \end{aligned}$$

In a sense, the only “unnatural” rule is the last one, since the expression “ $0 \cdot \infty$ ” is considered to be undefined in some branches of mathematics, but in the context of extended real-valued functions, defining it as zero is the “correct” choice in the sense of consistency. We will also use the following natural order between finite and infinite numbers:

$$\begin{aligned}
 \infty &> a & (-\infty \leq a < \infty), \\
 -\infty &< a & (-\infty < a \leq \infty).
 \end{aligned}$$

For an extended real-valued function $f : \mathbb{E} \rightarrow [-\infty, \infty]$, the *effective domain* or just the *domain* is the set

$$\text{dom}(f) = \{\mathbf{x} \in \mathbb{E} : f(\mathbf{x}) < \infty\}.$$

The notation “ $f : \mathbb{E} \rightarrow [-\infty, \infty]$ ” means that f is (potentially) extended real-valued (even if not explicitly stated). The notation “ $f : \mathbb{E} \rightarrow (-\infty, \infty]$ ” means that f is extended real-valued and does not attain the value $-\infty$.

The simplest examples of extended real-valued functions are indicators.

Example 2.1 (indicator functions). For any subset $C \subseteq \mathbb{E}$, the *indicator function* of C is defined to be the extended real-valued function given by

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

We obviously have

$$\text{dom}(\delta_C) = C.$$

The *epigraph* of an extended real-valued function $f : \mathbb{E} \rightarrow [-\infty, \infty]$ is defined by

$$\text{epi}(f) = \{(\mathbf{x}, y) : f(\mathbf{x}) \leq y, \mathbf{x} \in \mathbb{E}, y \in \mathbb{R}\}.$$

The epigraph is a subset of $\mathbb{E} \times \mathbb{R}$. Note that if $(\mathbf{x}, y) \in \text{epi}(f)$, then obviously $\mathbf{x} \in \text{dom}(f)$. A function $f : \mathbb{E} \rightarrow [-\infty, \infty]$ is called *proper* if it does not attain the value $-\infty$ and there exists at least one $\mathbf{x} \in \mathbb{E}$ such that $f(\mathbf{x}) < \infty$, meaning that $\text{dom}(f)$ is nonempty. The notion of closedness will play an important role in much of the analysis in this book.

Definition 2.2 (closed functions). A function $f : \mathbb{E} \rightarrow [-\infty, \infty]$ is **closed** if its epigraph is closed.

The indicator function δ_C is closed if and only if its underlying set C is closed.

Proposition 2.3 (closedness of indicators of closed sets). The indicator function δ_C is closed if and only if C is a closed set.

Proof. The epigraph of δ_C is given by

$$\text{epi}(\delta_C) = \{(\mathbf{x}, y) \in \mathbb{E} \times \mathbb{R} : \delta_C(\mathbf{x}) \leq y\} = C \times \mathbb{R}_+,$$

which is evidently closed if and only if C is closed. \square

We thus obtained in particular that the domain of a closed indicator function is necessarily a closed set. However, in general, we note that the domain of a closed function might not be closed. A classical example for this observation is given below.

Example 2.4. Consider the function $f : \mathbb{R} \rightarrow [-\infty, \infty]$ given by

$$f(x) = \begin{cases} \frac{1}{x}, & x > 0, \\ \infty, & \text{else.} \end{cases}$$

The domain of the function, which is the open interval $(0, \infty)$, is obviously not closed, but the function is closed since its epigraph

$$\text{epi}(f) = \{(x, y) : xy \geq 1, x > 0\}$$

is a closed set; see Figure 2.1. ■

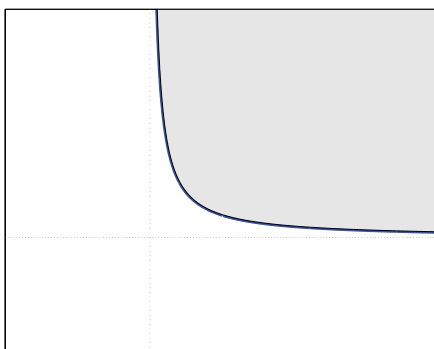


Figure 2.1. The epigraph of the function $f(x) = \frac{1}{x}$ for $x > 0$ and ∞ otherwise.

A property that will be later shown to be equivalent to closedness is *lower semicontinuity*.

Definition 2.5 (lower semicontinuity). A function $f : \mathbb{E} \rightarrow [-\infty, \infty]$ is called **lower semicontinuous at $\mathbf{x} \in \mathbb{E}$** if

$$f(\mathbf{x}) \leq \liminf_{n \rightarrow \infty} f(\mathbf{x}_n)$$

for any sequence $\{\mathbf{x}_n\}_{n \geq 1} \subseteq \mathbb{E}$ for which $\mathbf{x}_n \rightarrow \mathbf{x}$ as $n \rightarrow \infty$. A function $f : \mathbb{E} \rightarrow [-\infty, \infty]$ is called **lower semicontinuous** if it is lower semicontinuous at each point in \mathbb{E} .

For any $\alpha \in \mathbb{R}$, the α -level set of a function $f : \mathbb{E} \rightarrow [-\infty, \infty]$ is the set

$$\text{Lev}(f, \alpha) = \{\mathbf{x} \in \mathbb{E} : f(\mathbf{x}) \leq \alpha\}.$$

The following theorem shows that closedness and lower semicontinuity are equivalent properties, and they are both equivalent to the property that all the level sets of the function are closed.

Theorem 2.6 (equivalence of closedness, lower semicontinuity, and closedness of level sets). Let $f : \mathbb{E} \rightarrow [-\infty, \infty]$. Then the following three claims are equivalent:

- (i) f is lower semicontinuous.
- (ii) f is closed.
- (iii) For any $\alpha \in \mathbb{R}$, the level set

$$\text{Lev}(f, \alpha) = \{\mathbf{x} \in \mathbb{E} : f(\mathbf{x}) \leq \alpha\}$$

is closed.

Proof. (i \Rightarrow ii) Suppose that f is lower semicontinuous. We will show that $\text{epi}(f)$ is closed. For that, take $\{(\mathbf{x}_n, y_n)\}_{n \geq 1} \subseteq \text{epi}(f)$ such that $(\mathbf{x}_n, y_n) \rightarrow (\mathbf{x}^*, y^*)$ as $n \rightarrow \infty$. Then for any $n \geq 1$,

$$f(\mathbf{x}_n) \leq y_n.$$

Therefore, by the lower semicontinuity of f at \mathbf{x}^* , we have

$$f(\mathbf{x}^*) \leq \liminf_{n \rightarrow \infty} f(\mathbf{x}_n) \leq \liminf_{n \rightarrow \infty} y_n = y^*,$$

showing that $(\mathbf{x}^*, y^*) \in \text{epi}(f)$ and hence that f is closed.

(ii \Rightarrow iii) Suppose that f is closed, namely, that $\text{epi}(f)$ is closed. Let $\alpha \in \mathbb{R}$. We will show that $\text{Lev}(f, \alpha)$ is closed. If $\text{Lev}(f, \alpha) = \emptyset$, we are done. Otherwise, take a sequence $\{\mathbf{x}_n\}_{n \geq 1} \subseteq \text{Lev}(f, \alpha)$ that converges to $\bar{\mathbf{x}}$. Obviously $(\mathbf{x}_n, \alpha) \in \text{epi}(f)$ for any n and $(\mathbf{x}_n, \alpha) \rightarrow (\bar{\mathbf{x}}, \alpha)$ as $n \rightarrow \infty$. By the closedness of $\text{epi}(f)$, it follows that $(\bar{\mathbf{x}}, \alpha) \in \text{epi}(f)$, establishing the fact that $\bar{\mathbf{x}} \in \text{Lev}(f, \alpha)$.

(iii \Rightarrow i) Suppose that all the level sets of f are closed. We will show that it is lower semicontinuous. Assume by contradiction that f is not lower semicontinuous, meaning that there exists $\mathbf{x}^* \in \mathbb{E}$ and $\{\mathbf{x}_n\}_{n \geq 1} \subseteq \mathbb{E}$ such that $\mathbf{x}_n \rightarrow \mathbf{x}^*$ and $\liminf_{n \rightarrow \infty} f(\mathbf{x}_n) < f(\mathbf{x}^*)$. Take α that satisfies

$$\liminf_{n \rightarrow \infty} f(\mathbf{x}_n) < \alpha < f(\mathbf{x}^*). \quad (2.1)$$

Then there exists a subsequence $\{\mathbf{x}_{n_k}\}_{k \geq 1}$ such that $f(\mathbf{x}_{n_k}) \leq \alpha$ for all $k \geq 1$. By the closedness of the level set $\text{Lev}(f, \alpha)$ and the fact that $\mathbf{x}_{n_k} \rightarrow \mathbf{x}^*$ as $k \rightarrow \infty$, it follows that $f(\mathbf{x}^*) \leq \alpha$, which is a contradiction to (2.1), showing that (iii) implies (i). \square

The next result shows that closedness of functions is preserved under affine change of variables, summation, multiplication by a nonnegative number, and maximization. Before stating the theorem, we note that in this book we will not use the inf/sup notation but rather use only the min/max notation, where the usage of this notation does not imply that the maximum or minimum is actually attained.

Theorem 2.7 (operations preserving closedness).

- (a) Let $\mathcal{A} : \mathbb{E} \rightarrow \mathbb{V}$ be a linear transformation from \mathbb{E} to \mathbb{V} and $\mathbf{b} \in \mathbb{V}$ and let $f : \mathbb{V} \rightarrow [-\infty, \infty]$ be an extended real-valued closed function. Then the function $g : \mathbb{E} \rightarrow [-\infty, \infty]$ given by

$$g(\mathbf{x}) = f(\mathcal{A}(\mathbf{x}) + \mathbf{b})$$

is closed.

- (b) Let $f_1, f_2, \dots, f_m : \mathbb{E} \rightarrow (-\infty, \infty]$ be extended real-valued closed functions and let $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{R}_+$. Then the function $f = \sum_{i=1}^m \alpha_i f_i$ is closed.
- (c) Let $f_i : \mathbb{E} \rightarrow (-\infty, \infty], i \in I$ be extended real-valued closed functions, where I is a given index set. Then the function

$$f(\mathbf{x}) = \max_{i \in I} f_i(\mathbf{x})$$

is closed.

Proof. (a) To show that g is closed, take a sequence $\{(\mathbf{x}_n, y_n)\}_{n \geq 1} \subseteq \text{epi}(g)$ such that $(\mathbf{x}_n, y_n) \rightarrow (\mathbf{x}^*, y^*)$ as $n \rightarrow \infty$, where $\mathbf{x}^* \in \mathbb{E}$ and $y^* \in \mathbb{R}$. The relation $\{(\mathbf{x}_n, y_n)\}_{n \geq 1} \subseteq \text{epi}(g)$ can be written equivalently as

$$f(\mathcal{A}(\mathbf{x}_n) + \mathbf{b}) \leq y_n \text{ for all } n \geq 1.$$

Therefore, $(\mathcal{A}(\mathbf{x}_n) + \mathbf{b}, y_n) \in \text{epi}(f)$. Hence, since f is closed and $\mathcal{A}(\mathbf{x}_n) + \mathbf{b} \rightarrow \mathcal{A}(\mathbf{x}^*) + \mathbf{b}, y_n \rightarrow y^*$ as $n \rightarrow \infty$ (by the continuity of linear transformations), it follows that $(\mathcal{A}(\mathbf{x}^*) + \mathbf{b}, y^*) \in \text{epi}(f)$, meaning that

$$f(\mathcal{A}(\mathbf{x}^*) + \mathbf{b}) \leq y^*,$$

which is the same as the relation $(\mathbf{x}^*, y^*) \in \text{epi}(g)$. We have shown that $\text{epi}(g)$ is closed or, equivalently, that g is closed.

(b) We will prove that f is lower semicontinuous, which by Theorem 2.6 is equivalent to the closedness of f . Let $\{\mathbf{x}_n\}_{n \geq 1}$ be a sequence converging to \mathbf{x}^* . Then by the lower semicontinuity of f_i , for any $i = 1, 2, \dots, m$,

$$f_i(\mathbf{x}^*) \leq \liminf_{n \rightarrow \infty} f_i(\mathbf{x}_n).$$

Multiplying the above inequality by α_i and summing for $i = 1, 2, \dots, m$ gives

$$\left(\sum_{i=1}^m \alpha_i f_i \right) (\mathbf{x}^*) \leq \sum_{i=1}^m \liminf_{n \rightarrow \infty} \alpha_i f_i(\mathbf{x}_n) \leq \liminf_{n \rightarrow \infty} \left(\sum_{i=1}^m \alpha_i f_i \right) (\mathbf{x}_n),$$

where in the last inequality we used the fact that for any two sequences of real numbers $\{a_n\}_{n \geq 1}$ and $\{b_n\}_{n \geq 1}$, it holds that

$$\liminf_{n \rightarrow \infty} a_n + \liminf_{n \rightarrow \infty} b_n \leq \liminf_{n \rightarrow \infty} (a_n + b_n).$$

A simple induction argument shows that this property holds for an arbitrary number of sequences. We have thus established the lower semicontinuity and hence closedness of $\sum_{i=1}^m \alpha_i f_i$.

(c) Since f_i is closed for any $i \in I$, it follows that $\text{epi}(f_i)$ is closed for any i , and hence $\text{epi}(f) = \bigcap_{i \in I} \text{epi}(f_i)$ is closed as an intersection of closed sets, implying that f is closed. \square

2.2 Closedness versus Continuity

A relation between continuity and closedness is described in the following theorem stating that if an extended real-valued function is continuous over its domain,² which is assumed to be closed, then it is closed.

Theorem 2.8. *Let $f : \mathbb{E} \rightarrow (-\infty, \infty]$ be an extended real-valued function that is continuous over its domain and suppose that $\text{dom}(f)$ is closed. Then f is closed.*

Proof. To show that $\text{epi}(f)$ is closed (which is the same as saying that f is closed), take a sequence $\{(\mathbf{x}_n, y_n)\}_{n \geq 1} \subseteq \text{epi}(f)$ for which $(\mathbf{x}_n, y_n) \rightarrow (\mathbf{x}^*, y^*)$ as $n \rightarrow \infty$ for some $\mathbf{x}^* \in \mathbb{E}$ and $y \in \mathbb{R}$. Since $\{\mathbf{x}_n\}_{n \geq 1} \subseteq \text{dom}(f)$, $\mathbf{x}_n \rightarrow \mathbf{x}^*$ and $\text{dom}(f)$ is closed, it follows that $\mathbf{x}^* \in \text{dom}(f)$. By the definition of the epigraph, we have for all $n \geq 1$,

$$f(\mathbf{x}_n) \leq y_n. \quad (2.2)$$

Since f is continuous over $\text{dom}(f)$, and in particular at \mathbf{x}^* , it follows by taking n to ∞ in (2.2) that

$$f(\mathbf{x}^*) \leq y^*,$$

showing that $(\mathbf{x}^*, y^*) \in \text{epi}(f)$, thus establishing the closedness of $\text{epi}(f)$. \square

In particular, any real-valued continuous function over \mathbb{E} is closed.

Corollary 2.9. *Let $f : \mathbb{E} \rightarrow \mathbb{R}$ be continuous. Then f is closed.*

The above results demonstrate that there is a connection between continuity and closedness. However, these two notions are different, as the following example illustrates.

Example 2.10. Consider the function $f_\alpha : \mathbb{R} \rightarrow (-\infty, \infty]$ given by

$$f_\alpha(x) = \begin{cases} \alpha, & x = 0, \\ x, & 0 < x \leq 1, \\ \infty, & \text{else.} \end{cases}$$

²A function $g : \mathbb{E} \rightarrow (-\infty, \infty]$ is *continuous over its domain* if for any sequence $\{\mathbf{x}_n\}_{n \geq 1} \subseteq \text{dom}(g)$ satisfying $\mathbf{x}_n \rightarrow \mathbf{x}^*$ as $n \rightarrow \infty$ for some $\mathbf{x}^* \in \text{dom}(g)$, it holds that $g(\mathbf{x}_n) \rightarrow g(\mathbf{x}^*)$ as $n \rightarrow \infty$.

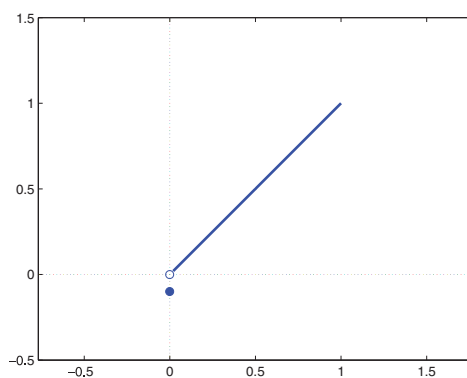


Figure 2.2. An example of a closed function, which is not continuous over its domain.

This function is closed if and only if $\alpha \leq 0$, and it is continuous over its domain if and only if $\alpha = 0$. Thus, the function $f_{-0.1}$, plotted in Figure 2.2, is closed but not continuous over its domain. ■

Example 2.11 (l_0 -norm). Consider the l_0 -norm function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$f(\mathbf{x}) = \|\mathbf{x}\|_0 \equiv \#\{i: x_i \neq 0\}.$$

That is, $\|\mathbf{x}\|_0$ is the number of nonzero elements in \mathbf{x} . Note the l_0 -norm is actually not a norm. It does not satisfy the homogeneity property. Nevertheless, this terminology is widely used in the literature, and we will therefore adopt it. Although f is obviously not continuous, it is closed. To show this, note that

$$f(\mathbf{x}) = \sum_{i=1}^n I(x_i),$$

where $I: \mathbb{R} \rightarrow \{0, 1\}$ is given by

$$I(y) = \begin{cases} 0, & y = 0, \\ 1, & y \neq 0. \end{cases}$$

The function I is closed since its level sets, which are given by

$$\text{Lev}(I, \alpha) = \begin{cases} \emptyset, & \alpha < 0, \\ \{0\}, & \alpha \in [0, 1), \\ \mathbb{R}, & \alpha \geq 1, \end{cases}$$

are closed sets. Therefore, f , as a sum of closed functions, is closed (Theorem 2.7(b)). ■

It is well known that a continuous function over a nonempty compact³ set attains a minimum. This is the well-known Weierstrass theorem. We will now show that this property also holds for closed functions.

Theorem 2.12 (Weierstrass theorem for closed functions). *Let $f : \mathbb{E} \rightarrow (-\infty, \infty]$ be a proper closed function and assume that C is a compact set satisfying $C \cap \text{dom}(f) \neq \emptyset$. Then*

- (a) f is bounded below over C .
- (b) f attains its minimal value over C .

Proof. (a) Suppose by contradiction that f is not bounded below over C . Then there exists a sequence $\{\mathbf{x}_n\}_{n \geq 1} \subseteq C$ such that

$$\lim_{n \rightarrow \infty} f(\mathbf{x}_n) = -\infty. \quad (2.3)$$

By the Bolzano–Weierstrass theorem, since C is compact, there exists a subsequence $\{\mathbf{x}_{n_k}\}_{k \geq 1}$ that converges to a point $\bar{\mathbf{x}} \in C$. By Theorem 2.6, f is lower semicontinuous, and hence

$$f(\bar{\mathbf{x}}) \leq \liminf_{k \rightarrow \infty} f(\mathbf{x}_{n_k}),$$

which is a contradiction to (2.3).

(b) Denote by f_{opt} the minimal value of f over C . Then there exists a sequence $\{\mathbf{x}_n\}_{n \geq 1}$ for which $f(\mathbf{x}_n) \rightarrow f_{\text{opt}}$ as $n \rightarrow \infty$. As before, take a subsequence $\{\mathbf{x}_{n_k}\}_{k \geq 1}$ that converges to some point $\bar{\mathbf{x}} \in C$. By the lower semicontinuity of f , it follows that

$$f(\bar{\mathbf{x}}) \leq \lim_{k \rightarrow \infty} f(\mathbf{x}_{n_k}) = f_{\text{opt}},$$

showing that $\bar{\mathbf{x}}$ is a minimizer of f over C . \square

When the set C in the premise of Theorem 2.12 is not compact, the Weierstrass theorem does not guarantee the attainment of a minimizer, but attainment of a minimizer can be shown when the compactness of C is replaced by closedness if the function has a property called *coerciveness*.

Definition 2.13 (coerciveness). *A proper function $f : \mathbb{E} \rightarrow (-\infty, \infty]$ is called coercive if*

$$\lim_{\|\mathbf{x}\| \rightarrow \infty} f(\mathbf{x}) = \infty.$$

An important property of closed coercive functions is that they possess a minimizer on any closed set that has a nonempty intersection with the domain of the function.

Theorem 2.14 (attainment under coerciveness). *Let $f : \mathbb{E} \rightarrow (-\infty, \infty]$ be a proper closed and coercive function and let $S \subseteq \mathbb{E}$ be a nonempty closed set satisfying $S \cap \text{dom}(f) \neq \emptyset$. Then f attains its minimal value over S .*

³A set is called *compact* if it is closed and bounded.

Proof. Let \mathbf{x}_0 be an arbitrary point in $S \cap \text{dom}(f)$. By the coerciveness of f , there exists an $M > 0$ such that

$$f(\mathbf{x}) > f(\mathbf{x}_0) \text{ for any } \mathbf{x} \text{ satisfying } \|\mathbf{x}\| > M. \quad (2.4)$$

Since any minimizer \mathbf{x}^* of f over S satisfies $f(\mathbf{x}^*) \leq f(\mathbf{x}_0)$, it follows from (2.4) that the set of minimizers of f over S is the same as the set of minimizers of f over $S \cap B_{\|\cdot\|}[\mathbf{0}, M]$, which is compact (both sets are closed, and $B_{\|\cdot\|}[\mathbf{0}, M]$ is bounded) and nonempty (as it contains \mathbf{x}_0). Therefore, by the Weierstrass theorem for closed functions (Theorem 2.12), there exists a minimizer of f over $S \cap B[\mathbf{0}, M]$ and hence also over S . \square

2.3 Convex Functions

2.3.1 Definition and Basic Properties

Like closedness, the definition of convexity for extended real-valued functions can be written in terms of the epigraph.

Definition 2.15 (convex functions). *An extended real-valued function $f : \mathbb{E} \rightarrow [-\infty, \infty]$ is called **convex** if $\text{epi}(f)$ is a convex set.*

It is not difficult to show that a proper extended real-valued function $f : \mathbb{E} \rightarrow (-\infty, \infty]$ is convex if and only if $\text{dom}(f)$ is convex and the restriction of f to $\text{dom}(f)$ is convex over $\text{dom}(f)$ in the sense of convexity of real-valued functions over convex domains. Using this observation, we conclude that a proper extended real-valued function f is convex if and only if

$$f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y}) \text{ for all } \mathbf{x}, \mathbf{y} \in \mathbb{E}, \lambda \in [0, 1], \quad (2.5)$$

or, equivalently, if and only if $\text{dom}(f)$ is convex and (2.5) is satisfied for any $\mathbf{x}, \mathbf{y} \in \text{dom}(f)$ and $\lambda \in [0, 1]$. Inequality (2.5) is a special case of *Jensen's inequality*, stating that for any $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in \mathbb{E}$ and $\boldsymbol{\lambda} \in \Delta_k$, the following inequality holds:

$$f\left(\sum_{i=1}^k \lambda_i \mathbf{x}_i\right) \leq \sum_{i=1}^k \lambda_i f(\mathbf{x}_i).$$

There are several operations that preserve convexity of extended real-valued convex functions. Some of them are summarized in Theorem 2.16 below. The proof can be easily deduced by combining two facts: (i) the same properties are known to hold for real-valued convex functions defined on a given convex domain, and (ii) the observation that a proper extended real-valued function is convex if and only if its domain is convex and its restriction to its domain is a real-valued convex function.

Theorem 2.16 (operations preserving convexity).

- (a) *Let $\mathcal{A} : \mathbb{E} \rightarrow \mathbb{V}$ be a linear transformation from \mathbb{E} to \mathbb{V} (two underlying vector spaces) and $\mathbf{b} \in \mathbb{V}$, and let $f : \mathbb{V} \rightarrow (-\infty, \infty]$ be an extended real-valued*

convex function. Then the extended real-valued function $g : \mathbb{E} \rightarrow (-\infty, \infty]$ given by

$$g(\mathbf{x}) = f(\mathcal{A}(\mathbf{x}) + \mathbf{b})$$

is convex.

(b) Let $f_1, f_2, \dots, f_m : \mathbb{E} \rightarrow (-\infty, \infty]$ be extended real-valued convex functions, and let $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{R}_+$. Then the function $\sum_{i=1}^m \alpha_i f_i$ is convex.

(c) Let $f_i : \mathbb{E} \rightarrow (-\infty, \infty], i \in I$, be extended real-valued convex functions, where I is a given index set. Then the function

$$f(\mathbf{x}) = \max_{i \in I} f_i(\mathbf{x})$$

is convex.

Given a nonempty set $C \subseteq \mathbb{E}$, the *distance function to C* is defined by

$$d_C(\mathbf{x}) = \min_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|.$$

The next example shows that for Euclidean spaces, the function $\frac{1}{2} (\|\mathbf{x}\|^2 - d_C^2(\mathbf{x}))$ is always convex, regardless of whether C is convex or not.

Example 2.17.⁴ Suppose that the underlying space \mathbb{E} is Euclidean (meaning that $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$). Let $C \subseteq \mathbb{E}$ be a nonempty set, and consider the function

$$\varphi_C(\mathbf{x}) = \frac{1}{2} (\|\mathbf{x}\|^2 - d_C^2(\mathbf{x})).$$

To show that φ_C is convex, note that

$$d_C^2(\mathbf{x}) = \min_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 - \max_{\mathbf{y} \in C} [2\langle \mathbf{y}, \mathbf{x} \rangle - \|\mathbf{y}\|^2].$$

Hence,

$$\varphi_C(\mathbf{x}) = \max_{\mathbf{y} \in C} \left[\langle \mathbf{y}, \mathbf{x} \rangle - \frac{1}{2} \|\mathbf{y}\|^2 \right]. \quad (2.6)$$

Therefore, since φ_C is a maximization of affine—and hence convex—functions, by Theorem 2.16(c), it is necessarily convex. ■

Another operation that preserves convexity is partial minimization of jointly convex functions.

Theorem 2.18 (convexity under partial minimization). Let $f : \mathbb{E} \times \mathbb{V} \rightarrow (-\infty, \infty]$ be a convex function satisfying the following property:

$$\text{for any } \mathbf{x} \in \mathbb{E} \text{ there exists } \mathbf{y} \in \mathbb{V} \text{ for which } f(\mathbf{x}, \mathbf{y}) < \infty. \quad (2.7)$$

⁴Example 2.17 is from Hiriart-Urruty and Lemaréchal [67, Example 2.1.4].

Let⁵ $g : \mathbb{E} \rightarrow [-\infty, \infty)$ be defined by

$$g(\mathbf{x}) \equiv \min_{\mathbf{y} \in \mathbb{E}} f(\mathbf{x}, \mathbf{y}).$$

Then g is convex.

Proof. Let $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{E}$ and $\lambda \in [0, 1]$. To show the convexity of g , we will prove that

$$g(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda g(\mathbf{x}_1) + (1 - \lambda) g(\mathbf{x}_2). \quad (2.8)$$

The inequality is obvious if $\lambda = 0$ or 1 . We will therefore assume that $\lambda \in (0, 1)$. The proof is split into two cases.

Case I: Here we assume that $g(\mathbf{x}_1), g(\mathbf{x}_2) > -\infty$. Take $\varepsilon > 0$. Then there exist $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{V}$ such that

$$f(\mathbf{x}_1, \mathbf{y}_1) \leq g(\mathbf{x}_1) + \varepsilon, \quad (2.9)$$

$$f(\mathbf{x}_2, \mathbf{y}_2) \leq g(\mathbf{x}_2) + \varepsilon. \quad (2.10)$$

By the convexity of f , we have

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, \lambda \mathbf{y}_1 + (1 - \lambda) \mathbf{y}_2) &\leq \lambda f(\mathbf{x}_1, \mathbf{y}_1) + (1 - \lambda) f(\mathbf{x}_2, \mathbf{y}_2) \\ &\stackrel{(2.9), (2.10)}{\leq} \lambda(g(\mathbf{x}_1) + \varepsilon) + (1 - \lambda)(g(\mathbf{x}_2) + \varepsilon) \\ &= \lambda g(\mathbf{x}_1) + (1 - \lambda) g(\mathbf{x}_2) + \varepsilon. \end{aligned}$$

Therefore, by the definition of g , we can conclude that

$$g(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda g(\mathbf{x}_1) + (1 - \lambda) g(\mathbf{x}_2) + \varepsilon.$$

Since the above inequality holds for any $\varepsilon > 0$, it follows that (2.8) holds.

Case II: Assume that at least one of the values $g(\mathbf{x}_1), g(\mathbf{x}_2)$ is equal $-\infty$. We will assume without loss of generality that $g(\mathbf{x}_1) = -\infty$. In this case, (2.8) is equivalent to saying that $g(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) = -\infty$. Take any $M \in \mathbb{R}$. Then since $g(\mathbf{x}_1) = -\infty$, it follows that there exists $\mathbf{y}_1 \in \mathbb{V}$ for which

$$f(\mathbf{x}_1, \mathbf{y}_1) \leq M.$$

By property (2.7), there exists $\mathbf{y}_2 \in \mathbb{V}$ for which $f(\mathbf{x}_2, \mathbf{y}_2) < \infty$. Using the convexity of f , we obtain that

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, \lambda \mathbf{y}_1 + (1 - \lambda) \mathbf{y}_2) &\leq \lambda f(\mathbf{x}_1, \mathbf{y}_1) + (1 - \lambda) f(\mathbf{x}_2, \mathbf{y}_2) \\ &\leq \lambda M + (1 - \lambda) f(\mathbf{x}_2, \mathbf{y}_2), \end{aligned}$$

which by the definition of g implies the inequality

$$g(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda M + (1 - \lambda) f(\mathbf{x}_2, \mathbf{y}_2).$$

Since the latter inequality holds for any $M \in \mathbb{R}$ and since $f(\mathbf{x}_2, \mathbf{y}_2) < \infty$, it follows that $g(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) = -\infty$, proving the result for the second case. \square

⁵The fact that g does not attain the value ∞ is a direct consequence of property (2.7).

2.3.2 The Infimal Convolution

Let $h_1, h_2 : \mathbb{E} \rightarrow (-\infty, \infty]$ be two proper functions. The *infimal convolution* of h_1, h_2 is defined by the following formula:

$$(h_1 \square h_2)(\mathbf{x}) \equiv \min_{\mathbf{u} \in \mathbb{E}} \{h_1(\mathbf{u}) + h_2(\mathbf{x} - \mathbf{u})\}.$$

A direct consequence of Theorem 2.18 is the following result stating that the infimal convolution of a proper convex function and a real-valued convex function is always convex.

Theorem 2.19 (convexity of the infimal convolution). *Let $h_1 : \mathbb{E} \rightarrow (-\infty, \infty]$ be a proper convex function and let $h_2 : \mathbb{E} \rightarrow \mathbb{R}$ be a real-valued convex function. Then $h_1 \square h_2$ is convex.*

Proof. Define $f(\mathbf{x}, \mathbf{y}) \equiv h_1(\mathbf{y}) + h_2(\mathbf{x} - \mathbf{y})$. The convexity of h_1 and h_2 implies that f is convex. In addition, property (2.7) holds since for any $\mathbf{x} \in \mathbb{E}$, we can pick any $\mathbf{y} \in \text{dom}(h_1)$ and obtain that $f(\mathbf{x}, \mathbf{y}) = h_1(\mathbf{y}) + h_2(\mathbf{x} - \mathbf{y}) < \infty$. Thus, by Theorem 2.18, the function $h_1 \square h_2$, as a partial minimization function of $f(\cdot, \cdot)$ w.r.t. the second argument is a convex function. \square

Example 2.20 (convexity of the distance function). Let $C \subseteq \mathbb{E}$ be a nonempty convex set. The distance function can be written as the following infimal convolution:

$$d_C(\mathbf{x}) = \min_{\mathbf{y}} \{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in C\} = \min_{\mathbf{y} \in \mathbb{E}} \{\delta_C(\mathbf{y}) + \|\mathbf{x} - \mathbf{y}\|\} = (\delta_C \square h_1)(\mathbf{x}),$$

where $h_1(\cdot) = \|\cdot\|$. Since δ_C is proper and convex and h_1 is real-valued convex, it follows by Theorem 2.19 that d_C is convex. \blacksquare

2.3.3 Continuity of Convex Functions

It is well known that convex functions are continuous at points in the interior of their domain. This is explicitly recalled in the next result, which actually states a stronger property of convex functions—local Lipschitz continuity.

Theorem 2.21 (local Lipschitz continuity of convex functions [10, Theorem 7.36]). *Let $f : \mathbb{E} \rightarrow (-\infty, \infty]$ be convex. Let $\mathbf{x}_0 \in \text{int}(\text{dom}(f))$. Then there exist $\varepsilon > 0$ and $L > 0$ such that $B[\mathbf{x}_0, \varepsilon] \subseteq C$ and*

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| \leq L\|\mathbf{x} - \mathbf{x}_0\| \quad (2.11)$$

for all $\mathbf{x} \in B[\mathbf{x}_0, \varepsilon]$.

Convex functions are not necessarily continuous at boundary points. Continuity is not guaranteed even when the function at hand is closed and convex (cf. Example 2.32). However, for univariate functions we will now show that closed and convex functions are continuous.

Theorem 2.22 (continuity of closed convex univariate functions). *Let $f : \mathbb{R} \rightarrow (-\infty, \infty]$ be a proper closed and convex function. Then f is continuous over $\text{dom}(f)$.*

Proof. Since f is convex, its domain is some interval $I = \text{dom}(f)$. If $\text{int}(I) = \emptyset$, then I is a singleton, and consequently the continuity of f over I is obvious. Assume then that $\text{int}(I) \neq \emptyset$. The fact that f is continuous over $\text{int}(I)$ follows from Theorem 2.21. We only need to show the continuity of f at the endpoints of I (if it exists). For that, we can assume without loss of generality that the interval I has a left endpoint a , and we will prove the right continuity of f at a . We begin by showing that $\lim_{t \rightarrow a^+} f(t)$ exists. Let $c > a$ be an arbitrary scalar in I and define the function

$$g(t) \equiv \frac{f(c-t) - f(c)}{t}.$$

Obviously, g is defined on $(0, c-a]$. We will show that g is nondecreasing and upper bounded over $(0, c-a]$. For that, take $0 < t \leq s \leq c-a$. Then

$$c-t = \left(1 - \frac{t}{s}\right)c + \frac{t}{s}(c-s),$$

and hence, by the convexity of f ,

$$f(c-t) \leq \left(1 - \frac{t}{s}\right)f(c) + \frac{t}{s}f(c-s),$$

which after some rearrangement of terms can be seen to be equivalent to

$$\frac{f(c-t) - f(c)}{t} \leq \frac{f(c-s) - f(c)}{s}.$$

Thus,

$$g(t) \leq g(s) \text{ for any } 0 < t \leq s \leq c-a. \quad (2.12)$$

Namely, g is nondecreasing over $(0, c-a]$. To show the upper boundedness, just plug $s = c-a$ into (2.12) and obtain that

$$g(t) \leq g(c-a) \text{ for any } t \in (0, c-a]. \quad (2.13)$$

We can thus conclude that $\lim_{t \rightarrow (c-a)^-} g(t)$ exists and is equal to some real number ℓ . Hence,

$$f(c-t) = f(c) + tg(t) \rightarrow f(c) + (c-a)\ell,$$

as $t \rightarrow (c-a)^-$, and consequently $\lim_{t \rightarrow a^+} f(t)$ exists and is equal to $f(c) + (c-a)\ell$. Using (2.13), we obtain that for any $t \in (0, c-a]$,

$$f(c-t) = f(c) + tg(t) \leq f(c) + (c-a)g(c-a) = f(c) + (c-a)\frac{f(a) - f(c)}{c-a} = f(a),$$

implying the inequality $\lim_{t \rightarrow a^+} f(t) \leq f(a)$. On the other hand, since f is closed, it is also lower semicontinuous (Theorem 2.6), and thus $\lim_{t \rightarrow a^+} f(t) \geq f(a)$. Consequently, $\lim_{t \rightarrow a^+} f(t) = f(a)$, proving the right continuity of f at a . \square

2.4 Support Functions

Let $C \subseteq \mathbb{E}$ be a nonempty set. Then the *support function* of C is the function $\sigma_C : \mathbb{E}^* \rightarrow (-\infty, \infty]$ given by

$$\sigma_C(\mathbf{y}) = \max_{\mathbf{x} \in C} \langle \mathbf{y}, \mathbf{x} \rangle.$$

For a fixed \mathbf{x} , the linear function $\mathbf{y} \mapsto \langle \mathbf{y}, \mathbf{x} \rangle$ is obviously closed and convex. Therefore, by Theorems 2.7(c) and 2.16(c), the support function, as a maximum of closed and convex functions, is always closed and convex, regardless of whether C is closed and/or convex. We summarize this property in the next lemma.

Lemma 2.23 (closedness and convexity of support functions). *Let $C \subseteq \mathbb{E}$ be a nonempty set. Then σ_C is a closed and convex function.*

In most of our discussions on support functions in this chapter, the fact that σ_C operates on the dual space \mathbb{E}^* instead of \mathbb{E} will have no importance—recall that we use the convention that the elements of \mathbb{E}^* and \mathbb{E} are the same. However, when norms will be involved, naturally, the dual norm will have to be used (cf. Example 2.31).

Additional properties of support functions that follow directly by definition are given in Lemma 2.24 below. Note that for two sets A, B that reside in the same space, the sum $A + B$ stands for the *Minkowski sum* given by

$$A + B = \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}.$$

Also, for a scalar $\alpha \in \mathbb{R}$ and a set $A \subseteq \mathbb{E}$, the set αA is

$$\alpha A = \{\alpha \mathbf{a} : \mathbf{a} \in A\}.$$

Lemma 2.24.

- (a) **(positive homogeneity)** For any nonempty set $C \subseteq \mathbb{E}$, $\mathbf{y} \in \mathbb{E}^*$ and $\alpha \geq 0$,

$$\sigma_C(\alpha \mathbf{y}) = \alpha \sigma_C(\mathbf{y}).$$

- (b) **(subadditivity)** For any nonempty set $C \subseteq \mathbb{E}$ and $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{E}^*$,

$$\sigma_C(\mathbf{y}_1 + \mathbf{y}_2) \leq \sigma_C(\mathbf{y}_1) + \sigma_C(\mathbf{y}_2).$$

- (c) For any nonempty set $C \subseteq \mathbb{E}$, $\mathbf{y} \in \mathbb{E}^*$ and $\alpha \geq 0$,

$$\sigma_{\alpha C}(\mathbf{y}) = \alpha \sigma_C(\mathbf{y}).$$

- (d) For any two nonempty sets $A, B \subseteq \mathbb{E}$ and $\mathbf{y} \in \mathbb{E}^*$,

$$\sigma_{A+B}(\mathbf{y}) = \sigma_A(\mathbf{y}) + \sigma_B(\mathbf{y}).$$

Proof. (a) $\sigma_C(\alpha \mathbf{y}) = \max_{\mathbf{x} \in C} \langle \alpha \mathbf{y}, \mathbf{x} \rangle = \alpha \max_{\mathbf{x} \in C} \langle \mathbf{y}, \mathbf{x} \rangle = \alpha \sigma_C(\mathbf{y})$.

(b)

$$\begin{aligned} \sigma_C(\mathbf{y}_1 + \mathbf{y}_2) &= \max_{\mathbf{x} \in C} \langle \mathbf{y}_1 + \mathbf{y}_2, \mathbf{x} \rangle = \max_{\mathbf{x} \in C} [\langle \mathbf{y}_1, \mathbf{x} \rangle + \langle \mathbf{y}_2, \mathbf{x} \rangle] \\ &\leq \max_{\mathbf{x} \in C} \langle \mathbf{y}_1, \mathbf{x} \rangle + \max_{\mathbf{x} \in C} \langle \mathbf{y}_2, \mathbf{x} \rangle = \sigma_C(\mathbf{y}_1) + \sigma_C(\mathbf{y}_2). \end{aligned}$$

(c)

$$\sigma_{\alpha C}(\mathbf{y}) = \max_{\mathbf{x} \in \alpha C} \langle \mathbf{y}, \mathbf{x} \rangle = \max_{\mathbf{x}_1 \in C} \langle \mathbf{y}, \alpha \mathbf{x}_1 \rangle = \alpha \max_{\mathbf{x}_1 \in C} \langle \mathbf{y}, \mathbf{x}_1 \rangle = \alpha \sigma_C(\mathbf{y}).$$

(d)

$$\begin{aligned} \sigma_{A+B}(\mathbf{y}) &= \max_{\mathbf{x} \in A+B} \langle \mathbf{y}, \mathbf{x} \rangle = \max_{\mathbf{x}_1 \in A, \mathbf{x}_2 \in B} \langle \mathbf{y}, \mathbf{x}_1 + \mathbf{x}_2 \rangle \\ &= \max_{\mathbf{x}_1 \in A, \mathbf{x}_2 \in B} [\langle \mathbf{y}, \mathbf{x}_1 \rangle + \langle \mathbf{y}, \mathbf{x}_2 \rangle] = \max_{\mathbf{x}_1 \in A} \langle \mathbf{y}, \mathbf{x}_1 \rangle + \max_{\mathbf{x}_2 \in B} \langle \mathbf{y}, \mathbf{x}_2 \rangle \\ &= \sigma_A(\mathbf{y}) + \sigma_B(\mathbf{y}). \quad \square \end{aligned}$$

Following are some basic examples of support functions.

Example 2.25 (support functions of finite sets). Suppose that

$$C = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\},$$

where $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m \in \mathbb{E}$. Then

$$\sigma_C(\mathbf{y}) = \max\{\langle \mathbf{b}_1, \mathbf{y} \rangle, \langle \mathbf{b}_2, \mathbf{y} \rangle, \dots, \langle \mathbf{b}_m, \mathbf{y} \rangle\}.$$

■

Recall that $S \subseteq \mathbb{E}$ is called a *cone* if it satisfies the following property: for any $\mathbf{x} \in S$ and $\lambda \geq 0$, the inclusion $\lambda \mathbf{x} \in S$ holds.

Example 2.26 (support functions of cones). Let $K \subseteq \mathbb{E}$ be a cone. Define the polar cone of K as

$$K^\circ = \{\mathbf{y} \in \mathbb{E}^* : \langle \mathbf{y}, \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{x} \in K\}.$$

We will show that

$$\sigma_K(\mathbf{y}) = \delta_{K^\circ}(\mathbf{y}). \quad (2.14)$$

Indeed, if $\mathbf{y} \in K^\circ$, then $\langle \mathbf{y}, \mathbf{x} \rangle \leq 0$ for all $\mathbf{x} \in K$ and for $\mathbf{x} = \mathbf{0}$, $\langle \mathbf{y}, \mathbf{x} \rangle = 0$. Therefore,

$$\sigma_K(\mathbf{y}) = \max_{\mathbf{x} \in K} \langle \mathbf{y}, \mathbf{x} \rangle = 0.$$

If $\mathbf{y} \notin K^\circ$, then there exists $\tilde{\mathbf{x}} \in K$ such that $\langle \mathbf{y}, \tilde{\mathbf{x}} \rangle > 0$. Since $\lambda \tilde{\mathbf{x}} \in K$ for all $\lambda \geq 0$, it follows that

$$\sigma_K(\mathbf{y}) \geq \langle \mathbf{y}, \lambda \tilde{\mathbf{x}} \rangle = \lambda \langle \mathbf{y}, \tilde{\mathbf{x}} \rangle \text{ for all } \lambda \geq 0.$$

Taking $\lambda \rightarrow \infty$, we obtain that $\sigma_K(\mathbf{y}) = \infty$ for $\mathbf{y} \notin K^\circ$, and hence formula (2.14) is proven. ■

Example 2.27 (support function of the nonnegative orthant). Consider the space $\mathbb{E} = \mathbb{R}^n$. As a special case of Example 2.26, since $(\mathbb{R}_+^n)^\circ = \mathbb{R}_-^n$, it follows that

$$\sigma_{\mathbb{R}_+^n}(\mathbf{y}) = \delta_{\mathbb{R}_-^n}(\mathbf{y}).$$

■

The next example uses Farkas's lemma,⁶ which we now recall.

Lemma 2.28 (Farkas's lemma—second formulation). *Let $\mathbf{c} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then the following two claims are equivalent:*

- A. *The implication $\mathbf{Ax} \leq \mathbf{0} \Rightarrow \mathbf{c}^T \mathbf{x} \leq 0$ holds true.*
- B. *There exists $\mathbf{y} \in \mathbb{R}_+^m$ such that $\mathbf{A}^T \mathbf{y} = \mathbf{c}$.*

Example 2.29 (support functions of convex polyhedral cones). Let the underlying space be $\mathbb{E} = \mathbb{R}^n$ and let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Define the set

$$S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} \leq \mathbf{0}\}.$$

Since S is a cone, we can use Example 2.26 to conclude that

$$\sigma_S(\mathbf{y}) = \delta_{S^\circ}(\mathbf{y}).$$

Note that $\mathbf{y} \in S^\circ$ if and only if

$$\langle \mathbf{y}, \mathbf{x} \rangle \leq 0 \text{ for any } \mathbf{x} \text{ satisfying } \mathbf{Ax} \leq \mathbf{0}. \quad (2.15)$$

By Farkas's lemma (Lemma 2.28), (2.15) is equivalent to the statement

$$\text{there exists } \boldsymbol{\lambda} \in \mathbb{R}_+^m \text{ such that } \mathbf{A}^T \boldsymbol{\lambda} = \mathbf{y}.$$

Hence,

$$S^\circ = \{\mathbf{A}^T \boldsymbol{\lambda} : \boldsymbol{\lambda} \in \mathbb{R}_+^m\}.$$

To conclude,

$$\sigma_S(\mathbf{y}) = \delta_{\{\mathbf{A}^T \boldsymbol{\lambda} : \boldsymbol{\lambda} \in \mathbb{R}_+^m\}}(\mathbf{y}).$$

■

Example 2.30 (support functions of affine sets). Let the underlying space be $\mathbb{E} = \mathbb{R}^n$ and let $\mathbf{B} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$. Define the affine set

$$C = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Bx} = \mathbf{b}\}.$$

⁶The lemma and its proof can be found, for example, in [10, Lemma 10.3].

We assume that C is nonempty, namely, that there exists $\mathbf{x}_0 \in \mathbb{R}^n$ for which $\mathbf{B}\mathbf{x}_0 = \mathbf{b}$. The support function is obviously given by

$$\sigma_C(\mathbf{y}) = \max_{\mathbf{x}} \{ \langle \mathbf{y}, \mathbf{x} \rangle : \mathbf{B}\mathbf{x} = \mathbf{b} \}.$$

Making the change of variables $\mathbf{x} = \mathbf{z} + \mathbf{x}_0$, we obtain that the support function can be rewritten as

$$\begin{aligned} \sigma_C(\mathbf{y}) &= \max_{\mathbf{z}} \{ \langle \mathbf{y}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{x}_0 \rangle : \mathbf{B}\mathbf{z} = \mathbf{0} \} \\ &= \langle \mathbf{y}, \mathbf{x}_0 \rangle + \max_{\mathbf{z}} \{ \langle \mathbf{y}, \mathbf{z} \rangle : \mathbf{B}\mathbf{z} = \mathbf{0} \} \\ &= \langle \mathbf{y}, \mathbf{x}_0 \rangle + \sigma_{\tilde{C}}(\mathbf{y}), \end{aligned} \quad (2.16)$$

where $\tilde{C} = \{ \mathbf{x} \in \mathbb{R}^n : \mathbf{B}\mathbf{x} = \mathbf{0} \}$. The set \tilde{C} is a convex polyhedral cone that can be written as

$$\tilde{C} = \{ \mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{0} \},$$

where $\mathbf{A} = \begin{pmatrix} \mathbf{B} \\ -\mathbf{B} \end{pmatrix}$. By Example 2.29, it follows that

$$\sigma_{\tilde{C}} = \delta_{\tilde{C}^\circ}, \quad (2.17)$$

where \tilde{C}° is the polar cone of \tilde{C} , which is given by

$$\tilde{C}^\circ = \{ \mathbf{B}^T \boldsymbol{\lambda}_1 - \mathbf{B}^T \boldsymbol{\lambda}_2 : \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^m \}.$$

We will show that

$$\tilde{C}^\circ = \text{Range}(\mathbf{B}^T). \quad (2.18)$$

Indeed, if $\mathbf{v} \in \tilde{C}^\circ$, then there exists $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^m$ for which $\mathbf{v} = \mathbf{B}^T \boldsymbol{\lambda}_1 - \mathbf{B}^T \boldsymbol{\lambda}_2 = \mathbf{B}^T(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2) \in \text{Range}(\mathbf{B}^T)$. In the other direction, if $\mathbf{v} \in \text{Range}(\mathbf{B}^T)$, then there exists $\boldsymbol{\lambda} \in \mathbb{R}^m$ for which $\mathbf{v} = \mathbf{B}^T \boldsymbol{\lambda}$. Defining $\boldsymbol{\lambda}_1 = [\boldsymbol{\lambda}]_+$, $\boldsymbol{\lambda}_2 = [-\boldsymbol{\lambda}]_+$, we obtain that $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2$ with $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^m$, and hence

$$\mathbf{v} = \mathbf{B}^T \boldsymbol{\lambda} = \mathbf{B}^T(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2) = \mathbf{B}^T \boldsymbol{\lambda}_1 - \mathbf{B}^T \boldsymbol{\lambda}_2 \in \tilde{C}^\circ.$$

Combining (2.16), (2.17), and (2.18), we finally conclude that

$$\sigma_C(\mathbf{y}) = \langle \mathbf{y}, \mathbf{x}_0 \rangle + \delta_{\text{Range}(\mathbf{B}^T)}(\mathbf{y}).$$

■

Example 2.31 (support functions of unit balls). Suppose that \mathbb{E} is the underlying space endowed with a norm $\| \cdot \|$. Consider the unit ball given by

$$B_{\| \cdot \|}[\mathbf{0}, 1] = \{ \mathbf{x} \in \mathbb{E} : \| \mathbf{x} \| \leq 1 \}.$$

By the definition of the dual norm, we have for any $\mathbf{y} \in \mathbb{E}^*$

$$\sigma_{B_{\|\cdot\|}[0,1]}(\mathbf{y}) = \max_{\|\mathbf{x}\| \leq 1} \langle \mathbf{y}, \mathbf{x} \rangle = \|\mathbf{y}\|_*.$$

Thus, for example, for the space \mathbb{R}^n we have

$$\begin{aligned} \sigma_{B_{\|\cdot\|_p}[0,1]}(\mathbf{y}) &= \|\mathbf{y}\|_q \quad \left(1 \leq p \leq \infty, \frac{1}{p} + \frac{1}{q} = 1\right), \\ \sigma_{B_{\|\cdot\|_{\mathbf{Q}}}[0,1]}(\mathbf{y}) &= \|\mathbf{y}\|_{\mathbf{Q}^{-1}} \quad (\mathbf{Q} \in \mathbb{S}_{++}^n). \end{aligned}$$

In the first formula we use the convention that $p = 1/\infty$ corresponds to $q = \infty/1$.

The next example is also an example of a closed and convex function that is not continuous (recall that such an example does not exist for one-dimensional functions; see Theorem 2.22).

Example 2.32.⁷ Consider the following set in \mathbb{R}^2 :

$$C = \left\{ (x_1, x_2)^T : x_1 + \frac{x_2^2}{2} \leq 0 \right\}.$$

Then the support function of C is given by

$$\sigma_C(\mathbf{y}) = \max_{x_1, x_2} \left\{ y_1 x_1 + y_2 x_2 : x_1 + \frac{x_2^2}{2} \leq 0 \right\}. \quad (2.19)$$

Obviously, $\sigma_C(\mathbf{0}) = 0$. We will compute the support function at $\mathbf{y} \neq \mathbf{0}$. In this case, it is easy to see that the maximum of problem (2.19) is attained at the boundary of C .⁸ Therefore,

$$\sigma_C(\mathbf{y}) = \max_{x_1, x_2} \left\{ y_1 x_1 + y_2 x_2 : x_1 + \frac{x_2^2}{2} = 0 \right\} = \max_{x_2} \left\{ -\frac{y_1}{2} x_2^2 + y_2 x_2 \right\}.$$

If $y_1 < 0$, then the maximal value is ∞ . If $y_1 = 0$ and $y_2 \neq 0$, then the maximal value is also ∞ . If $y_1 > 0$, the maximum is attained at $x_2 = \frac{y_2}{y_1}$, and the corresponding maximal value is $\frac{y_2^2}{2y_1}$. Thus, the support function is given by

$$\sigma_C(\mathbf{y}) = \begin{cases} \frac{y_2^2}{2y_1}, & y_1 > 0, \\ 0, & y_1 = y_2 = 0, \\ \infty & \text{else.} \end{cases}$$

⁷Example 2.32 is from Rockafellar [108, p. 83].

⁸This fact can be shown by contradiction. If the maximum was attained at an interior point of C , then the gradient of the objective function, meaning \mathbf{y} , would be the zeros vector, which is a contradiction to the assumption that $\mathbf{y} \neq \mathbf{0}$.

By Lemma 2.23, σ_C is closed and convex. However, it is not continuous at $(y_1, y_2) = (0, 0)$. Indeed, taking for any $\alpha > 0$ the path $y_1(t) = \frac{t^2}{2\alpha}$, $y_2(t) = t$ ($t > 0$), we obtain that

$$\sigma_C(y_1(t), y_2(t)) = \alpha,$$

and hence the limit of $\sigma_C(y_1(t), y_2(t))$ as $t \rightarrow 0^+$ is α , which combined with the fact that $\sigma_C(0, 0) = 0$ implies the discontinuity of f at $(0, 0)$. The contour lines of σ_C are plotted in Figure 2.3. ■

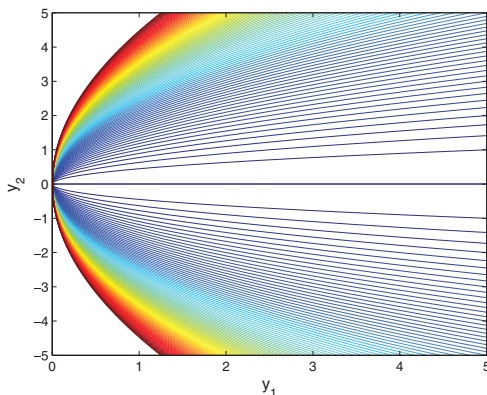


Figure 2.3. Contour lines of the closed, convex, and noncontinuous function from Example 2.32.

An important property of support functions is that they are completely determined by their underlying sets as long as these sets are closed and convex. The proof of this result requires the strict separation theorem,⁹ which is now recalled.

Theorem 2.33 (strict separation theorem). *Let $C \subseteq \mathbb{E}$ be a nonempty closed and convex set, and let $\mathbf{y} \notin C$. Then there exist $\mathbf{p} \in \mathbb{E}^* \setminus \{\mathbf{0}\}$ and $\alpha \in \mathbb{R}$ such that*

$$\langle \mathbf{p}, \mathbf{y} \rangle > \alpha$$

and

$$\langle \mathbf{p}, \mathbf{x} \rangle \leq \alpha \text{ for all } \mathbf{x} \in C.$$

Lemma 2.34. *Let $A, B \subseteq \mathbb{E}$ be nonempty closed and convex sets. Then $A = B$ if and only if $\sigma_A = \sigma_B$.*

Proof. If $A = B$, then obviously $\sigma_A = \sigma_B$. Suppose now that $\sigma_A = \sigma_B$. We will prove that $A = B$. Assume by contradiction that this is not the case, and without loss of generality suppose that there exists $\mathbf{y} \in A$ such that $\mathbf{y} \notin B$. Since $\mathbf{y} \notin B$ and B is nonempty closed and convex, by the strict separation theorem, there exists a

⁹The theorem and its proof can be found, for example, in [10, Theorem 10.1].

hyperplane separating \mathbf{y} from B , meaning that there exists $\mathbf{p} \in \mathbb{E}^* \setminus \{\mathbf{0}\}$ and $\alpha > 0$ such that

$$\langle \mathbf{p}, \mathbf{x} \rangle \leq \alpha < \langle \mathbf{p}, \mathbf{y} \rangle \text{ for any } \mathbf{x} \in B.$$

Taking the maximum over $\mathbf{x} \in B$, we conclude that $\sigma_B(\mathbf{p}) \leq \alpha < \langle \mathbf{p}, \mathbf{y} \rangle \leq \sigma_A(\mathbf{y})$, a contradiction to the assertion that the support functions are the same. \square

A related result states that the support function stays the same under the operations of closure and convex hull of the underlying set.

Lemma 2.35. *Let $A \subseteq \mathbb{E}$ be nonempty. Then*

- (a) $\sigma_A = \sigma_{\text{cl}(A)}$;
- (b) $\sigma_A = \sigma_{\text{conv}(A)}$.

Proof. (a) Since $A \subseteq \text{cl}(A)$,

$$\sigma_A(\mathbf{y}) \leq \sigma_{\text{cl}(A)}(\mathbf{y}) \text{ for any } \mathbf{y} \in \mathbb{E}^*. \quad (2.20)$$

We will show the reverse inequality. Let $\mathbf{y} \in \mathbb{E}^*$. Then by the definition of the support function, there exists a sequence $\{\mathbf{x}^k\}_{k \geq 1} \subseteq \text{cl}(A)$ such that

$$\langle \mathbf{y}, \mathbf{x}^k \rangle \rightarrow \sigma_{\text{cl}(A)}(\mathbf{y}) \text{ as } k \rightarrow \infty. \quad (2.21)$$

By the definition of the closure, it follows that there exists a sequence $\{\mathbf{z}^k\}_{k \geq 1} \subseteq A$ such that $\|\mathbf{z}^k - \mathbf{x}^k\| \leq \frac{1}{k}$ for all k , and hence

$$\mathbf{z}^k - \mathbf{x}^k \rightarrow \mathbf{0} \text{ as } k \rightarrow \infty. \quad (2.22)$$

Now, since $\mathbf{z}^k \in A$,

$$\sigma_A(\mathbf{y}) \geq \langle \mathbf{y}, \mathbf{z}^k \rangle = \langle \mathbf{y}, \mathbf{x}^k \rangle + \langle \mathbf{y}, \mathbf{z}^k - \mathbf{x}^k \rangle.$$

Taking $k \rightarrow \infty$ and using (2.21), (2.22), we obtain that

$$\sigma_A(\mathbf{y}) \geq \sigma_{\text{cl}(A)}(\mathbf{y}) + 0 = \sigma_{\text{cl}(A)}(\mathbf{y}),$$

which combined with (2.20) yields the desired result $\sigma_A = \sigma_{\text{cl}(A)}$.

(b) Since $A \subseteq \text{conv}(A)$, we have that $\sigma_A(\mathbf{y}) \leq \sigma_{\text{conv}(A)}(\mathbf{y})$ for any $\mathbf{y} \in \mathbb{E}^*$. We will show the reverse inequality. Let $\mathbf{y} \in \mathbb{E}^*$. Then by the definition of the support function, there exists a sequence $\{\mathbf{x}^k\}_{k \geq 1} \subseteq \text{conv}(A)$ such that

$$\langle \mathbf{y}, \mathbf{x}^k \rangle \rightarrow \sigma_{\text{conv}(A)}(\mathbf{y}) \text{ as } k \rightarrow \infty. \quad (2.23)$$

By the definition of the convex hull, it follows that for any k , there exist vectors $\mathbf{z}_1^k, \mathbf{z}_2^k, \dots, \mathbf{z}_{n_k}^k \in A$ (n_k is a positive integer) and $\lambda^k \in \Delta_{n_k}$ such that

$$\mathbf{x}^k = \sum_{i=1}^{n_k} \lambda_i^k \mathbf{z}_i^k.$$

Now,

$$\langle \mathbf{y}, \mathbf{x}^k \rangle = \left\langle \mathbf{y}, \sum_{i=1}^{n_k} \lambda_i^k \mathbf{z}_i^k \right\rangle = \sum_{i=1}^{n_k} \lambda_i^k \langle \mathbf{y}, \mathbf{z}_i^k \rangle \leq \sum_{i=1}^{n_k} \lambda_i^k \sigma_A(\mathbf{y}) = \sigma_A(\mathbf{y}),$$

where the inequality follows by the fact that $\mathbf{z}_i^k \in A$. Taking the limit as $k \rightarrow \infty$ and using (2.23), we obtain that $\sigma_{\text{conv}(A)}(\mathbf{y}) \leq \sigma_A(\mathbf{y})$. \square

Example 2.36 (support of the unit simplex). Suppose that the underlying space is \mathbb{R}^n and consider the unit simplex set $\Delta_n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{e}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\}$. Since the unit simplex can be written as the convex hull of the standard basis of \mathbb{R}^n ,

$$\Delta_n = \text{conv}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\},$$

it follows by Lemma 2.35(b) that

$$\sigma_{\Delta_n}(\mathbf{y}) = \sigma_{\{\mathbf{e}_1, \dots, \mathbf{e}_n\}}(\mathbf{y}) = \max\{\langle \mathbf{e}_1, \mathbf{y} \rangle, \langle \mathbf{e}_2, \mathbf{y} \rangle, \dots, \langle \mathbf{e}_n, \mathbf{y} \rangle\}.$$

Since we always assume (unless otherwise stated) that \mathbb{R}^n is endowed with the dot product, the support function is

$$\sigma_{\Delta_n}(\mathbf{y}) = \max\{y_1, y_2, \dots, y_n\}.$$

\blacksquare

The table below summarizes the main support function computations that were considered in this section.

| C | $\sigma_C(\mathbf{y})$ | Assumptions | Reference |
|--|--|---|--------------|
| $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ | $\max_{i=1,2,\dots,n} \langle \mathbf{b}_i, \mathbf{y} \rangle$ | $\mathbf{b}_i \in \mathbb{E}$ | Example 2.25 |
| K | $\delta_{K^\circ}(\mathbf{y})$ | K —cone | Example 2.26 |
| \mathbb{R}_+^n | $\delta_{\mathbb{R}_-^n}(\mathbf{y})$ | $\mathbb{E} = \mathbb{R}^n$ | Example 2.27 |
| Δ_n | $\max\{y_1, y_2, \dots, y_n\}$ | $\mathbb{E} = \mathbb{R}^n$ | Example 2.36 |
| $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{0}\}$ | $\delta_{\{\mathbf{A}^T \boldsymbol{\lambda} : \boldsymbol{\lambda} \in \mathbb{R}_+^m\}}(\mathbf{y})$ | $\mathbb{E} = \mathbb{R}^n, \mathbf{A} \in \mathbb{R}^{m \times n}$ | Example 2.29 |
| $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{B}\mathbf{x} = \mathbf{b}\}$ | $\langle \mathbf{y}, \mathbf{x}_0 \rangle + \delta_{\text{Range}(\mathbf{B}^T)}(\mathbf{y})$ | $\mathbb{E} = \mathbb{R}^n, \mathbf{B} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{B}\mathbf{x}_0 = \mathbf{b}$ | Example 2.30 |
| $B_{\ \cdot\ }[\mathbf{0}, 1]$ | $\ \mathbf{y}\ _*$ | - | Example 2.31 |