Introductory Lectures on Optimization

General Convex Problem (1)

Hui Qian
qianhui@zju.edu.cn

College of Computer Science, Zhejiang University

November 13, 2024

Outline

Part I
Motivation and Definitions

Motivation

We consider methods for solving the most general convex minimization problem

$$\boxed{\begin{aligned} &\min f_0(\boldsymbol{x}), \\ \text{s.t.} \quad &f_i(\boldsymbol{x}) \leq 0, \ i = 1 \ldots m, \\ &\boldsymbol{x} \in Q \subseteq \mathbb{R}^n, \end{aligned} \qquad (1)}$$

where Q is a closed convex set and $f_i(\boldsymbol{x}), i = 0, \ldots m$ are genearl convex functions. The term "general" means that these functions can be nondifferentiable.

Clearly, such a problem is more difficult than a problem with differentiable.

Motivation

Remark. Interior Point:
An element $\boldsymbol{x} \in C \subset \mathbb{R}^n$ is called an *interior* point of C if there exists an $\epsilon > 0$ for which

$$\left\{ \boldsymbol{y} \middle| \, \|\boldsymbol{y} - \boldsymbol{x}\|_2 \leq \epsilon \right\} \subset C.$$

Remark. Open and Closed Set:
A set C is open if int $C = C$, i.e., every point in C is an interior point. A set $C \subset \mathbb{R}^n$ is closed if its complement $\mathbb{R}^n \backslash C = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{x} \notin C\}$ is open.

Motivation

Note that nonsmooth minimization problems arise frequently in different applications.

- Quite often, some components of a model are composed of max-type functions:

$$f(\boldsymbol{x}) = \max_{1 \le j \le p} \phi_j(\boldsymbol{x}),$$

- Another source of nondifferentiable functions is the situation when some components of the problem (1) are given implicitly, as solutions of some auxiliary problems.

Denote by

$$\operatorname{dom} f = \{\boldsymbol{x} \in \mathbb{R}^n : |f(\boldsymbol{x})| < \infty\},$$

the domain of function f. We always assume that $\operatorname{dom} f \neq \emptyset$.

Definition

Definition 1 (Definition 3.1.1)

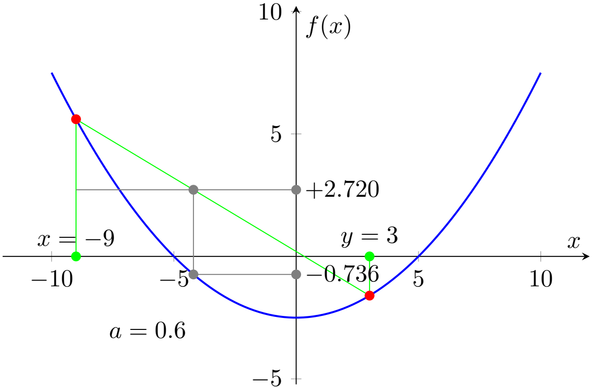
A function $f(\boldsymbol{x})$ is called convex if its domain is convex and for all $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$ and $\alpha \in [0, 1]$ the following inequality holds:

$$f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}).$$

If this inequality is strict, the function is called strictly convex. We call f concave if $-f$ is convex.

Our optimization schemes were based on gradients of smooth functions. For nonsmooth functions, such objects do not exist and we have to find something to replace them.

Definition



Jensen's inequality

Lemma 2 (Jensen's Inequality)

For any $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in \operatorname{dom} f$ (Here, f is convex) and positive coefficients $\alpha_1, \ldots, \alpha_m$ such that

$$\sum_{i=1}^{m} \alpha_i = 1, \quad \alpha_i \geq 0, i = 1 \ldots m, \tag{2}$$

we have

$$f\left(\sum_{i=1}^{m} \alpha_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^{m} \alpha_i f(\boldsymbol{x}_i).$$

Jensen's inequality

Proof. Let us prove this statement by induction over m. Definition 1 justifies inequality (2) for $m = 2$. For a set $m + 1$ points we have

$$\sum_{i=1}^{m+1} \alpha_i \boldsymbol{x}_i = \alpha_1 \boldsymbol{x}_1 + (1 - \alpha_1) \sum_{i=1}^{m} \beta_i \boldsymbol{x}_{i+1},$$

where $\beta_i = \frac{\alpha_{i+1}}{1-\alpha_1}$. Clearly

$$\sum_{i=1}^{m} \beta_i = 1, \quad \beta_i \geq 0, i = 1 \ldots m.$$

Remark. $1 - \alpha_1 = \alpha_2 + \cdots + \alpha_{i+1}.$

Jensen's inequality

Proof. (Continued.) Therefore, using Definition 1 and our inductive assumption, we have

$$f\left(\sum_{i=1}^{m+1}\alpha_i\boldsymbol{x}_i\right) = f\left(\alpha_1\boldsymbol{x}_1 + (1-\alpha_1)\sum_{i=1}^{m}\beta_i\boldsymbol{x}_i\right)$$

$$\underbrace{\leq}_{\text{convex}} \alpha_1 f(\boldsymbol{x}_1) + (1-\alpha_1) f\left(\sum_{i=1}^{m}\beta_i x_i\right) \underbrace{\leq}_{\text{assumption}} \sum_{i=1}^{m+1}\alpha_i f(\boldsymbol{x}_i).$$

□

A point $\boldsymbol{x} = \sum_{i=1}^{m}\alpha_i\boldsymbol{x}_i$ with positive coefficients α_i satisfying the normalizing condition (2) is called a convex combination of points $\{\boldsymbol{x}_i\}_{i=1}^{m}$.

Corollaries

Corollary 3 (Corollary 3.1.1)

Let $\boldsymbol{x}$ be a convex combination of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$. Then

$$f(\boldsymbol{x}) \leq \max_{1 \leq i \leq m} f(\boldsymbol{x}_i).$$

Proof. Indeed, by Jensen's inequality and condition (2), we have

$$f(\boldsymbol{x}) = f\left(\sum_{i=1}^{m} \alpha_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^{m} \alpha_i f(\boldsymbol{x}_i) \leq \sum_{i=1}^{m} \left(\alpha_i \max_{1 \leq i \leq m} f(\boldsymbol{x}_i)\right) \leq \max_{1 \leq i \leq m} f(\boldsymbol{x}_i).$$

$\square$

Corollaries

Corollary 4 (Corollary 3.1.2)

Let

$$\Delta = \mathrm{Conv}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} \equiv \left\{ \boldsymbol{x} = \sum_{i=1}^{m} \alpha_i \boldsymbol{x}_i | \alpha_i \geq 0, \sum_{i=1}^{m} \alpha_i = 1 \right\}.$$

Then $\max_{\boldsymbol{x} \in \Delta} f(\boldsymbol{x}) = \max_{1 \leq i \leq m} f(\boldsymbol{x}_i)$.

Remark. in view of Corollary 3, we have

$$f(\boldsymbol{x}) \leq \max_{1 \leq i \leq m} f(\boldsymbol{x}_i).$$

Therefore, if $f(\boldsymbol{x}_i)$ obtain the max value, $\boldsymbol{x}^* = \boldsymbol{x}_i$.

Equivalent Definition of Convex Function

Theorem 5 (Theorem 3.1.1)

A function f is convex if and only if for all $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$ and $\beta \geq 0$ such that $\boldsymbol{y} + \beta(\boldsymbol{y} - \boldsymbol{x}) \in \operatorname{dom} f$, we have

$$f(\boldsymbol{y} + \beta(\boldsymbol{y} - \boldsymbol{x})) \geq f(\boldsymbol{y}) + \beta(f(\boldsymbol{y}) - f(\boldsymbol{x})). \tag{3}$$

Proof. Necessary conditions: Let f be convex. Define $\alpha = \frac{\beta}{1+\beta}$ and $\boldsymbol{u} = \boldsymbol{y} + \beta(\boldsymbol{y} - \boldsymbol{x})$. Then

$$\boldsymbol{y} = \frac{1}{1+\beta}(\boldsymbol{u} + \beta\boldsymbol{x}) = (1-\alpha)\boldsymbol{u} + \alpha\boldsymbol{x}.$$

Equivalent Definition of Convex Function

Proof. (Continued.) Therefore,

$$f(\boldsymbol{y}) \leq (1-\alpha)f(\boldsymbol{u}) + \alpha f(\boldsymbol{x}) = \frac{1}{1+\beta}f(\boldsymbol{u}) + \frac{\beta}{1+\beta}f(\boldsymbol{x}).$$

Sufficient conditions: Assume now that (3) holds. Let us fix $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$. Define $\alpha = \frac{1}{1+\beta}$ (eg. $\beta = \frac{1-\alpha}{\alpha}$) and $\boldsymbol{u} = \alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}$. Then

$$\boldsymbol{x} = \frac{1}{\alpha}(\boldsymbol{u} - (1-\alpha)\boldsymbol{y}) = \boldsymbol{u} + \beta(\boldsymbol{u} - \boldsymbol{y}).$$

Therefore,

$$f(\boldsymbol{x}) \geq f(\boldsymbol{u}) + \beta(f(\boldsymbol{u}) - f(\boldsymbol{y})) = \frac{1}{\alpha}f(\boldsymbol{u}) - \frac{1-\alpha}{\alpha}f(\boldsymbol{y}).$$

□

Equivalent Definition of Convex Function

Theorem 6 (Theorem 3.1.2)

A function f is convex if and only if its epigraph

$$\text{epi}(f) = \{(\boldsymbol{x}, t) \in \text{dom } f \times \mathbb{R} | t \geq f(\boldsymbol{x})\}$$

is a convex set.

Proof. Necessary conditions: Let f be convex. If $(\boldsymbol{x}_1, t_1) \in \text{epi}(f)$ and $(\boldsymbol{x}_2, t_2) \in \text{epi}(f)$, then for any $\alpha \in [0, 1]$ we have

$$\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\boldsymbol{x}_1) + (1 - \alpha) f(\boldsymbol{x}_2) \geq f(\alpha \boldsymbol{x}_1 + (1 - \alpha) \boldsymbol{x}_2).$$

Thus, $(\alpha \boldsymbol{x}_1 + (1 - \alpha) \boldsymbol{x}_2, \alpha t_1 + (1 - \alpha) t_2) \in \text{epi}(f)$.

Equivalent Definition of Convex Function

Proof. (Continued.) Sufficient conditions: Let $\mathrm{epi}(f)$ be convex. Note that for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in$ dom f, the corresponding points of the graph of the function belong to the epigraph:

$$(\boldsymbol{x}_1, f(\boldsymbol{x}_1)) \in \mathrm{epi}(f), \quad (\boldsymbol{x}_2, f(\boldsymbol{x}_2)) \in \mathrm{epi}(f).$$

Therefore $(\underbrace{\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2}_{\boldsymbol{x}}, \underbrace{\alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2)}_{t}) \in \mathrm{epi}(f)$. In view of the definition of epigraph we have

$$f(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2).$$

$\square$

Other Properties

Theorem 7 (Theorem 3.1.3)

If a function f is convex, then all level sets

$$\mathcal{L}_f(\beta) = \{\boldsymbol{x} \in \text{dom } f | f(\boldsymbol{x}) \leq \beta\}$$

are either convex or empty.

Proof. Indeed, if $\boldsymbol{x}_1 \in \mathcal{L}_f(\beta)$ and $\boldsymbol{x}_2 \in \mathcal{L}_f(\beta)$, then for any $\alpha \in [0, 1]$ we have

$$f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2) \leq \alpha\beta + (1-\alpha)\beta = \beta.$$

□

Other Properties

We will see, in the examples section, that behavior of a general convex function on the boundary of its domain is sometimes out of any control. Therefore, we need to introduce one convenient notion, which will be very useful in our analysis.

Definition 8 (Definition 3.1.2)

A function f is called closed and convex if its epigraph is a closed set.

Theorem 9 (Theorem 3.1.4)

If convex function f is closed, then all its level sets are either empty or closed.

Proof. By its definition, $(\mathcal{L}_f(\beta), \beta) = \text{epi}(f) \cap \{(x, t) | t = \beta\}$. Therefore, the epigraph $\mathcal{L}_f(\beta)$ is closed and convex as an intersection of two closed convex sets. $\square$

Examples

In general, a closed and convex function is not necessarily continuous.

Example 10 (Example 3.1.1)

1. Linear function is closed and convex.
2. $f(\boldsymbol{x}) = |\boldsymbol{x}|, \boldsymbol{x} \in \mathbb{R}^1$, is closed and convex since its epigraph is
$$\{(\boldsymbol{x}, t) | t \geq \boldsymbol{x} \text{ and } t \geq -\boldsymbol{x}\},$$
the intersection of two closed convex sets.
3. All differentiable and convex on $\mathbb{R}^n$ functions belong to the class of general closed and convex functions.

Examples

Example 10

(Continued.)

4. Function $f(\boldsymbol{x}) = \frac{1}{\boldsymbol{x}}$, $\boldsymbol{x} > 0$, is convex and closed. However, its domain $\operatorname{dom} f = \operatorname{int} \mathbb{R}^1_+$ is open
5. Function $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$, where $\|\cdot\|$ is any norm, is closed and convex:

$$
\begin{aligned}
f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) &= \|\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2\| \\
&\leq \|\alpha \boldsymbol{x}_1\| + \|(1-\alpha)\boldsymbol{x}_2\| \\
&= \alpha \|\boldsymbol{x}_1\| + (1-\alpha)\|\boldsymbol{x}_2\| .
\end{aligned}
$$

for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ and $\alpha \in [0,1]$.

Examples

Example 10

5 (Continued.) The most important norms in numerical analysis are so-called ℓ_p-norms :

$$\|\boldsymbol{x}\|_p = \left[\sum_{i=1}^{n} |\boldsymbol{x}^{(i)}|^p\right]^{1/p}, \quad p \geq 1.$$

Among them there are three norms, which are commonly used:

- Euclidean norm (Euclidean norm): $\|\boldsymbol{x}\| = \left[\sum_{i=1}^{n} \left(\boldsymbol{x}^{(i)}\right)\right]^{1/2}, \ p = 2$。
- ℓ_1-norm: $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |\boldsymbol{x}^{(i)}|, \ p = 1$. (Taxicab norm or Manhattan norm)
- ℓ_∞-norm (Chebyshev distance, Infinity norm or Maximum norm):

$$\|\boldsymbol{x}\|_\infty = \max_{1\leq i\leq n} |\boldsymbol{x}^{(i)}|.$$

Examples

Example 10

5 (Continued.) Any norm defines a system of balls,

$$B_{\|\cdot\|}(\boldsymbol{x}_0, r) = \{\boldsymbol{x} \in \mathbb{R}^n | \, \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq r\}, r \geq 0,$$

where r is the radius of the ball and $\boldsymbol{x}_0 \in \mathbb{R}^n$ is its center. We call the ball $B_{\|\cdot\|}(0, 1)$ the unit ball of the norm $\|\cdot\|$. Clearly, these balls are convex set (see Theorem 7). For ℓ_p-ball of the radius r we use the notation

$$B_p(\boldsymbol{x}_0, r) = \{\boldsymbol{x} \in \mathbb{R}^n | \, \|\boldsymbol{x} - \boldsymbol{x}_0\|_p \leq r\}.$$

Examples

Example 10

5. (Continued.) Note that the following relation between Euclidean and ℓ_1-ball holds:

$$B_1(\boldsymbol{x}_0, r) \subset B_2(\boldsymbol{x}_0, r) \subset B_1(\boldsymbol{x}_0, r\sqrt{n}).$$

That is true because of the standard inequalities:

$$\sum_{i=1}^{n} \left(\boldsymbol{x}^{(i)}\right)^2 \leq \left(\sum_{i=1}^{n} |\boldsymbol{x}^{(i)}|\right)^2,$$

$$\left(\frac{1}{n}\sum_{i=1}^{n} |\boldsymbol{x}^{(i)}|\right)^2 \leq \frac{1}{n}\sum_{i=1}^{n} \left(\boldsymbol{x}^{(i)}\right)^2.$$

Examples

Example 10

6 Up to now, all our examples did not show up any pathological behavior. However, let us look at the following function of two variables:

$$f(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0, & \text{if } \boldsymbol{x}^2 + \boldsymbol{y}^2 < 1, \\ \phi(\boldsymbol{x}, \boldsymbol{y}), & \text{if } \boldsymbol{x}^2 + \boldsymbol{y}^2 = 1, \end{cases}$$

where $\phi(\boldsymbol{x}, \boldsymbol{y})$ is an arbitrary nonnegative function defined on a unit circle. The domain of this function is the unit Euclidean disk, which is closed and convex. Moreover, it is easy to see that f is convex. However, it has no reasonable properties on the boundary of its domain. Definitely, we want to exclude such functions from our considerations. That was the reason for introducing the notion of closed function. It is clear that $f(\boldsymbol{x}, \boldsymbol{y})$ is not closed unless $\phi(\boldsymbol{x}, \boldsymbol{y}) \equiv 0$.

Part II
Operations with Convex Function

Invariant Operations

Let us describe a set of invariant operations to create more complicated objects.

Theorem 11 (Theorem 3.1.5)

Let function f_1 and f_2 be closed and convex and let $\beta \geq 0$. Then all functions below are closed and convex:

1. $f(\boldsymbol{x}) = \beta f_1(\boldsymbol{x}),\ \ \text{dom } f = \text{dom } f_1.$
2. $f(\boldsymbol{x}) = f_1(\boldsymbol{x}) + f_2(\boldsymbol{x}),\ \ \text{dom } f = (\text{dom } f_1) \cap (\text{dom } f_2).$
3. $f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\},\ \ \text{dom } f = (\text{dom } f_1) \cap (\text{dom } f_2).$

Proof.

1. $f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) = \beta f_1(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \leq \beta(\alpha f_1(\boldsymbol{x}_1) + (1-\alpha) f_1(\boldsymbol{x}_2)).$

Invariant Operations

Proof. (Continued)

2. For all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in (\text{dom } f_1) \cap (\text{dom } f_2)$ 且 $\alpha \in [0,1]$, we have

$$\begin{aligned} f_1(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) + f_2(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \\ &\le \alpha f_1(\boldsymbol{x}_1) + (1-\alpha) f_1(\boldsymbol{x}_2) + \alpha f_2(\boldsymbol{x}_1) + (1-\alpha) f_2(\boldsymbol{x}_2) \\ &= \alpha(f_1(\boldsymbol{x}_1) + f_2(\boldsymbol{x}_1)) + (1-\alpha)(f_1(\boldsymbol{x}_2) + f_2(\boldsymbol{x}_2)). \end{aligned}$$

Thus $f(\boldsymbol{x})$ is convex. Let us prove that it is closed. Consider a sequence $\{(\boldsymbol{x}_k, t_k)\} \subset \text{epi}(f)$:

$$t_k \ge \underbrace{f_1(\boldsymbol{x}_k) + f_2(\boldsymbol{x}_k)}_{f(\boldsymbol{x}_k)}, \quad \lim_{k\to\infty} \boldsymbol{x}_k = \bar{\boldsymbol{x}} \in \text{dom } f, \quad \lim_{k\to\infty} t_k = \bar{t}.$$

Remark. Please refer to p178, Theorem 9.3 of Convex Analysis for more formal proof.

Invariant Operations

Proof.

2 (Continued.) Since f_1 and f_2 are closed, we have

$$\lim_{k\to\infty}\inf f_1(\boldsymbol{x}_k) \geq f_1(\bar{\boldsymbol{x}}), \quad \lim_{k\to\infty}\inf f_2(\boldsymbol{x}_k) \geq f_2(\bar{\boldsymbol{x}}),$$

where $\boxed{\lim_{k\to\infty}\inf \boldsymbol{x}_k \leftrightarrow \lim_{k\to\infty}\left(\inf_{m\geq k} \boldsymbol{x}_m\right)}$. Therefore

$$\bar{t} = \lim_{k\to\infty} t_k \geq \lim_{k\to\infty}\inf f_1(\boldsymbol{x}_k) + \lim_{k\to\infty}\inf f_2(\boldsymbol{x}_k) \geq f(\bar{\boldsymbol{x}}).$$

Thus, $(\bar{\boldsymbol{x}}, \bar{t}) \in \text{epi}\ f$. (For a set, if the limit of any sequence (in it) is also in the same set, it is a closed set.)

Invariant Operations

Lower-Semicontinuity: A function f is *lower-semicontinuous* at a given vector $\bar{\boldsymbol{x}}$ if for every sequence $\{\boldsymbol{x}_k\}$ converging to $\bar{\boldsymbol{x}}$, we have

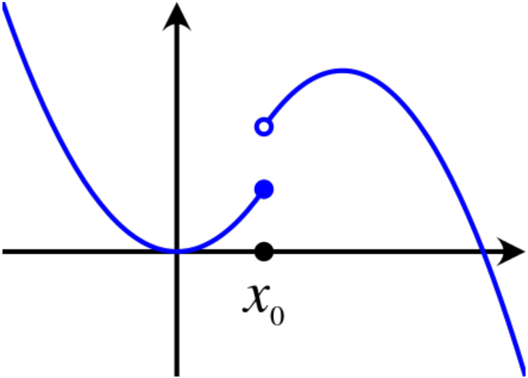
$$\liminf_{k\to\infty} f(\boldsymbol{x}_k) \geq f(\bar{\boldsymbol{x}}).$$

We say that f is *lower-semicontinuous* over a set X if f is *lower-semicontinuous* at every $\boldsymbol{x} \in X$.

Theorem 12

For a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$ the following statements are equivalent: (1)f is closed. (2)Every level set of f is closed. (3)f is lower-semicontinuous over $\mathbb{R}^n$.

Invariant Operations



Invariant Operations

Proof.

3. The epigraph of function $f(\boldsymbol{x})$ is

$$\begin{aligned} \text{epi } f &= \{(\boldsymbol{x}, t) | t \geq f_1(\boldsymbol{x}), t \geq f_2(\boldsymbol{x}), \boldsymbol{x} \in (\text{dom } f_1 \cap \text{dom } f_2)\} \\ &\equiv \text{epi } f_1 \cap \text{epi } f_2. \end{aligned}$$

Thus, epi f is closed and convex as an intersection of two closed convex sets. It remains to use Theorem 6.

$\square$

Invariant Operations

The following theorem demonstrates that convexity is an affine-invariant property.

Theorem 13 (Theorem 3.1.6)

Let function $\phi(\boldsymbol{y})$, $\boldsymbol{y} \in \mathbb{R}^m$, be convex and closed. Consider a linear operator

$$\mathcal{A}(\boldsymbol{x}) = A\boldsymbol{x} + b : \quad \mathbb{R}^n \to \mathbb{R}^m.$$

Then $f(\boldsymbol{x}) = \phi(\mathcal{A}(\boldsymbol{x}))$ is a closed and convex function with domain

$$\text{dom } f = \{\boldsymbol{x} \in \mathbb{R}^n | \mathcal{A}(\boldsymbol{x}) \in \text{dom}\phi\}.$$

Invariant Operations

Proof. For $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ from dom f, denote $\boldsymbol{y}_1 = \mathcal{A}(\boldsymbol{x}_1)$, $\boldsymbol{y}_2 = \mathcal{A}(\boldsymbol{x}_2)$. Then for $\alpha \in [0,1]$ we have

$$\begin{aligned} f(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) &= \phi(\mathcal{A}(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2)) \\ &= \phi(\alpha\boldsymbol{y}_1 + (1-\alpha)\boldsymbol{y}_2) \\ &\leq \alpha\phi(\boldsymbol{y}_1) + (1-\alpha)\phi(\boldsymbol{y}_2) \\ &= \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2). \end{aligned}$$

Thus, $f(\boldsymbol{x})$ is convex. The closedness of its epigraph follows from continuity of the linear operator $\mathcal{A}(\boldsymbol{x})$.

Invariant Operations

The next theorem is one of the main suppliers of convex functions with implicit structure.

Theorem 14 (Theorem 3.1.7)

Let Δ be seome set and

$$f(\boldsymbol{x}) = \sup_{\boldsymbol{y}}\{\phi(\boldsymbol{y}, \boldsymbol{x}) | \boldsymbol{y} \in \Delta\}.$$

Suppose that for any fixed $\boldsymbol{y} \in \Delta$, the function $\phi(\boldsymbol{y}, \boldsymbol{x})$ is closed and convex in $\boldsymbol{x}$. Then $f(\boldsymbol{x})$ is a closed and convex function with domain

$$\text{dom } f = \{\boldsymbol{x} \in \cap_{\boldsymbol{y}\in\Delta} \text{dom } \phi(\boldsymbol{y}, \cdot) | \exists\gamma : \phi(\boldsymbol{y}, \boldsymbol{x}) \leq \gamma, \forall \boldsymbol{y} \in \Delta\}. \tag{4}$$

Invariant Operations

Proof. Check domain: Indeed, if $\boldsymbol{x}$ belongs to the right-hand side of equation (4), then $f(\boldsymbol{x}) < \infty$ and we conclude that $\boldsymbol{x} \in \text{dom } f$. If $\boldsymbol{x}$ does not belong to this set, then there exists a sequence $\{\boldsymbol{y}_k\}$ such that $\phi(\boldsymbol{y}_k, \boldsymbol{x}) \to \infty$. Therefore $\boldsymbol{x}$ does not belong to dom f.

Finally, it is clear that $(\boldsymbol{x}, t) \in \text{epi } f$ if and only if for all $\boldsymbol{y} \in \Delta$ we have

$$\boldsymbol{x} \in \text{dom } \phi(\boldsymbol{y}, \cdot), \quad t \geq \phi(\boldsymbol{y}, \boldsymbol{x}).$$

This means that

$$\text{epi } f = \cap_{\boldsymbol{y} \in \Delta} \text{epi } \phi(\boldsymbol{y}, \cdot).$$

Therefore f is convex and closed since each epi $\phi(\boldsymbol{y}, \cdot)$ is convex and closed. □

注：无限个凸集的交集是凸的；无限个闭集的交集是闭的。

Examples

Example 15 (Example 3.1.2)

1. Function $f(\boldsymbol{x}) = \max_{1\le i\le n}\{\boldsymbol{x}^{(i)}\}$is closed and convex.
2. Let $\lambda = (\lambda^{(1)}, \dots, \lambda^{(m)})$ where $\lambda^{(k)} \ge 0$ for $k = 1, \dots, m$, and Δ be a set in $\mathbb{R}^m_+$. Consider the function

$$f(\boldsymbol{x}) = \sup_{\lambda\in\Delta} \sum_{i=1}^{m} \lambda^{(i)} f_i(\boldsymbol{x}),$$

where f_i are closed and convex. In view of Theorem 11, the epigraphs of functions

$$\phi_\lambda = \sum_{i=1}^{m} \lambda^{(i)} f_i(\boldsymbol{x})$$

are convex and closed. Thus, in view of Theorem 14, $f(\boldsymbol{x})$ is convex and closed. Note that we did not assume anything about the structure of the set Δ.

Examples

Example 15

3 Let Q be a convex set. Consider the function

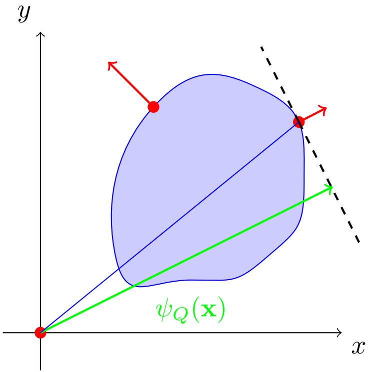
$$\psi_Q(\boldsymbol{x}) = \sup\{\langle g,\ \boldsymbol{x}\rangle | g \in Q\}.$$

Function $\psi_Q(\boldsymbol{x})$ is called support function of the set Q. Note that $\psi_Q(\boldsymbol{x})$ is closed and convex in view of Theorem 14. This function is homogeneous of degree one:

$$\psi_Q(t\boldsymbol{x}) = t\psi_Q(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathrm{dom}\ \psi_Q, \quad t \geq 0.$$

If the set Q is bounded then $\mathrm{dom}\ \psi_Q = \mathbb{R}^n$.

Examples



Remark. For $\|\boldsymbol{x}\| = 1$, $\psi_Q(\boldsymbol{x})$ is the distance from the origin to a hyperplane with the normal vector $\boldsymbol{x}$.

Examples

Example 15

4 Let Q be a set in $\mathbb{R}^n$. Consider the function $\psi(g,\gamma) = \sup_{\boldsymbol{y}\in Q} \phi(\boldsymbol{y}, g, \gamma)$ where

$$\phi(\boldsymbol{y}, g, \gamma) = \langle g,\ \boldsymbol{y}\rangle - \frac{\gamma}{2}\|\boldsymbol{y}\|^2 .$$

In view of Theorem 14, the function $\psi(g,\gamma)$ is closed and convex in (g,γ).
(1) If Q is bounded, then dom $\psi = \mathbb{R}^{n+1}$. (2) Consider the case $Q = \mathbb{R}^n$. Let us describe the domain of ψ: (2.1) If $\gamma < 0$, then for any $g \neq 0$ we can take $\boldsymbol{y}_\alpha = \alpha g$. Clearly, along this sequence $\phi(\boldsymbol{y}_\alpha, g, \gamma) \to \infty$ as $\alpha \to \infty$.
(2.2) If $\gamma = 0$, the only possible value for g is zero since otherwise the function $\phi(\boldsymbol{y}, g, 0)$is unbounded.

Examples

Example 15

4 (Continued.) Finally, (2.3) if $\gamma > 0$ then the point maximizing $\phi(\boldsymbol{y}, g, \gamma)$ with respect to $\boldsymbol{y}$ is $\boldsymbol{y}^*(g,\gamma) = \frac{1}{\gamma} g$ and we get the following expression for ψ:

$$\psi(g,\gamma) = \frac{\|g\|^2}{2\gamma}.$$

Thus,

$$\psi(g,\gamma) = \begin{cases} 0, & \text{if } g = 0, \gamma = 0, \\ \frac{\|g\|^2}{2\gamma}, & \text{if } \gamma > 0, \end{cases}$$

with the domain $\mathrm{dom}\, \psi = (\mathbb{R}^n \times \{\gamma > 0\}) \cup (0, 0)$.

Examples

Example 15

4. (Continued) Note that this is a convex set, which is neither closed nor open. Nevertheless, ψ is a closed convex function. At the same time, this function is not continuous at the origin:

$$\psi(\sqrt{\gamma}\hat{g}, \gamma) \equiv \frac{1}{2}\|\hat{g}\|^2, \quad \gamma \neq 0.$$

注：缩小 γ，则沿着 $(\sqrt{\gamma}\hat{g}, \gamma)$ 的轨迹，一直都有 $\frac{1}{2}\|\hat{g}\|^2$，但是当 $\gamma = 0$ 的时候，跳变为 0。

References I

Yurii Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media, 2013.

Yurii Nesterov. *Lectures on convex optimization*, volume 137. Springer, 2018.

Sébastien Bubeck et al. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4):231–357, 2015.

Thank You!

Email:qianhui@zju.edu.cn