Chapter 2

Convex Analysis

This chapter presents the elements of convex analysis that will be useful for studying optimization problems and the algorithms that solve them.

2.1 Convex Sets

Definition 2.1.1 (Segment)**.** Let $x, y \in \mathbb{R}^n$. The segment in $\mathbb{R}^n$ is the set denoted and defined by

$$[x, y] = \{(1-\lambda)x + \lambda y : \lambda \in [0,1]\}.$$

Definition 2.1.2 (Convex Set)**.** Let $C \subset \mathbb{R}^n$. The set C is said to be convex if

$$\forall \lambda \in [0,1],\ \forall x, y \in C:\ (1-\lambda)x + \lambda y \in C, \tag{2.1}$$

or equivalently,

$$\forall \alpha, \beta \in \mathbb{R}^+,\ \alpha + \beta = 1,\ \forall x, y \in C:\ \alpha x + \beta y \in C,$$

or also,

$$\forall x, y \in C, \quad [x, y] \subset C.$$

Example 2.1.3. $\emptyset$ and $\mathbb{R}^n$ are two convex sets.

Example 2.1.4. A vector subspace is obviously convex, as well as an affine subspace which is nothing but the translation of a vector subspace.

Example 2.1.5. Let the set

$$S = \{x \in \mathbb{R}^3 : x_1 + 2x_2 - x_3 = 4\},$$

which can also be written as

$$S = \{x \in \mathbb{R}^3 : x^\top p = 4\},$$

with $p = (1, 2, -1)^\top$. The set S is called a hyperplane, and the vector p is called the normal vector.

Is S convex? Let $\lambda \in [0,1]$ and $x, y \in S$. We show that $(1-\lambda)x + \lambda y \in S$, i.e.

$$p^\top \big((1-\lambda)x + \lambda y\big) \stackrel{?}{=} 4.$$

We have

$$\langle p, (1-\lambda)x + \lambda y\rangle = (1-\lambda)\langle p, x\rangle + \lambda\langle p, y\rangle = (1-\lambda)4 + \lambda 4 = 4.$$

Thus, $(1-\lambda)x + \lambda y \in S$. The convexity of S follows.

In general, a hyperplane $H \subset \mathbb{R}^n$ is defined from a normal vector p and a constant α, that is

$$H = \{x \in \mathbb{R}^n : x^\top p = \alpha\}.$$

Example 2.1.6. The set

$$H^- = \{x \in \mathbb{R}^n : x^\top p \leq \alpha\}$$

is called the negative half-space. Equivalently, one defines the positive half-space

$$H^+ = \{x \in \mathbb{R}^n : x^\top p \geq \alpha\}$$

associated with the hyperplane H. The reader can check that H^+ and H^- are also convex sets.

Proposition 2.1.7.

- If C is a convex set and $\beta \in \mathbb{R}$, then

  $$\beta C = \{y : y = \beta x,\ x \in C\}$$

  is also convex.

- If C_1 and C_2 are two convex sets, then

  $$C_1 + C_2 = \{x + y : x \in C_1,\ y \in C_2\}$$

  is convex.

- Let $\{C_i\}_{i \in I}$ be a family of convex sets (where I is any index set, finite or infinite). Then

  $$\bigcap_{i \in I} C_i$$

  is convex.

2.2 Convex Functions

Definition 2.2.1. Let $C \subset \mathbb{R}^n$ be a convex set and let $f : C \to \mathbb{R}$.

- f is said to be **convex** if

$$\forall\lambda \in [0,1],\ \forall x, y \in C:\ f\big((1-\lambda)x + \lambda y\big) \leq (1-\lambda)f(x) + \lambda f(y), \tag{2.2}$$

  or equivalently,

$$\forall\lambda \in [0,1],\ \forall x, y \in C:\ f\big(x + \lambda(y-x)\big) \leq f(x) + \lambda\big(f(y) - f(x)\big),$$

  or also,

$$\forall p, q \geq 0,\ p+q=1, \quad f(px+qy) \leq pf(x) + qf(y).$$

- f is said to be **strictly convex** if

$$\forall\lambda \in (0,1),\ \forall x, y \in C,\ x \neq y:\ f\big((1-\lambda)x + \lambda y\big) < (1-\lambda)f(x) + \lambda f(y).$$

Example 2.2.2. Let $L : \mathbb{R}^n \to \mathbb{R}$ be a linear mapping and let $\alpha \in \mathbb{R}$. Define an affine function h by

$$h(x) = L(x) + \alpha.$$

For $\lambda \in [0,1]$ and $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, we have

$$h\big((1-\lambda)x$$

Remark 2.2.3. Any affine mapping is both convex and concave. The following notion of set clearly shows the link between a function and a convex set.

Definition 2.2.4 (Epigraph and Hypograph)**.** The **epigraph** of f is the subset of $C\times\mathbb{R}$ that lies above its graph. It is denoted by epi f and defined as

$$\operatorname{epi} f = \{(x,\alpha) : f(x) \leq \alpha\}.$$

The **hypograph** of f is the subset of $C\times\mathbb{R}$ that lies below its graph. It is denoted by hyp f and defined as

$$\operatorname{hyp} f = \{(x,\alpha) : f(x) \geq \alpha\}.$$

Proposition 2.2.5. Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$, with C convex. Then f is convex (resp. concave) if and only if its epigraph (resp. hypograph) is convex.

Proof. Suppose that f is convex and let us show that epi f is convex, i.e.,

$$(1-\lambda)(x_1,\alpha_1) + \lambda(x_2,\alpha_2) \in \operatorname{epi} f,$$

which is equivalent to

$$\big((1-\lambda)x_1 + \lambda x_2, (1-\lambda)\alpha_1 + \lambda\alpha_2\big) \in \operatorname{epi} f \quad \Leftrightarrow \quad f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)\alpha_1 + \lambda\alpha_2.$$

Let $(x_1,\alpha_1), (x_2,\alpha_2) \in \operatorname{epi} f$. Since f is convex, we have

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2) \leq (1-\lambda)\alpha_1 + \lambda\alpha_2,$$

where the last inequality holds because $(x_1, \alpha_1), (x_2, \alpha_2) \in \operatorname{epi} f$. Thus $\operatorname{epi} f$ is convex.

Conversely, suppose that $\operatorname{epi} f$ is convex and let us show the convexity of f. Let $x_1, x_2 \in C$ and $\lambda \in [0, 1]$. Then

$$(x_1, f(x_1)) \in \operatorname{epi} f, \quad (x_2, f(x_2)) \in \operatorname{epi} f.$$

By convexity of the epigraph, we deduce that

$$(1-\lambda)(x_1, f(x_1)) + \lambda(x_2, f(x_2)) \in \operatorname{epi} f.$$

This means

$$\big((1-\lambda)x_1 + \lambda x_2, (1-\lambda)f(x_1) + \lambda f(x_2)\big) \in \operatorname{epi} f,$$

hence

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2).$$

Therefore, f is convex. □

Example 2.2.6. The following functions are convex or concave according to the convexity of their epigraph or hypograph:

- $x \mapsto x^2$ is convex,
- $x \mapsto \sqrt{x}$ is concave,
- $x \mapsto \exp(x)$ is convex,
- $(x_1, x_2) \mapsto \frac{1}{2}\big(x_1^2 + x_2^2\big)$ is convex.

2.3 Properties of Convex Functions

Proposition 2.3.1. Let $f, g : C \subset \mathbb{R}^n \to \mathbb{R}$, with C convex, and f and g two convex functions.

1. $f + g$ is convex.
2. If $\alpha \geq 0$, then αf is convex.
3. The function $h(x) = \max\{f(x), g(x)\}$ is convex.

Convex functions can only have points of discontinuity at the boundary of their domains, as shown in the following theorem.

Theorem 2.3.2 ([**?**, Theorem 2.1])**.** *Let C be a convex set in $\mathbb{R}^n$ with nonempty interior, and let $f : C \to \mathbb{R}$ be a convex function. Then f is continuous on* $\operatorname{int}(C)$.

Theorem 2.3.3 (Characterization of convexity via the gradient)**.** *Let $C \subset \mathbb{R}^n$ be a convex open set and $f : C \to \mathbb{R}$ a differentiable function. Then f is convex on C if and only if*

$$\forall x, y \in C : \quad f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle. \tag{2.3}$$

Proof. Suppose f is convex, i.e.

$$\forall x, y \in C,\ \lambda \in [0,1]: \quad f(x+\lambda(y-x)) \leq f(x) + \lambda\big(f(y) - f(x)\big).$$

Subtracting $f(x)$ from both sides and dividing by λ, we obtain

$$\frac{f(x+\lambda(y-x)) - f(x)}{\lambda} \leq f(y) - f(x).$$

Taking the limit as $\lambda \to 0$ and applying Lemma 1.1, we get

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x),$$

which yields (**??**).

Conversely, let $x, y \in C$ and $\lambda \in [0,1]$. Setting $x_\lambda = x + \lambda(y-x)$, one finds

$$x_\lambda - x = \lambda(y-x), \quad x_\lambda - y = (1-\lambda)(x-y).$$

Applying (**??**) to f at the points (x, x_λ) and (y, x_λ), we obtain

$$f(x) \geq f(x_\lambda) - \lambda\langle \nabla f(x_\lambda), y - x \rangle,$$

$$f(y) \geq f(x_\lambda) + (1-\lambda)\langle \nabla f(x_\lambda), y - x \rangle.$$

Multiplying the first inequality by $(1-\lambda)$ and the second by λ, then summing, yields exactly the convexity condition for f. □

Theorem 2.3.4 (Characterization of convexity via the Hessian)**.** *Let $C \subset \mathbb{R}^n$ be a convex open set and $f : C \to \mathbb{R}$ a $\mathcal{C}^2$ function. Then:*

- *f is convex on C if and only if for all $x \in C$, the Hessian matrix $\nabla^2 f(x)$ is positive semidefinite,*
- *f is strictly convex on C if and only if $\nabla^2 f(x)$ is positive definite for all $x \in C$.*

That is,

$$\forall y \in \mathbb{R}^n : \quad \langle y, \nabla^2 f(x) y \rangle \geq 0. \tag{2.4}$$

Proof. Suppose f is convex and let $x \in C$. We want to show (**??**). Since C is open, for any $y \in \mathbb{R}^n$, there exists λ small enough with $\lambda \neq 0$ such that $x + \lambda y \in C$. From the previous theorem and the second-order differentiability of f, we have

$$f(x+\lambda y) \geq f(x) + \lambda\langle \nabla f(x), y \rangle, \tag{2.5}$$

and the Taylor expansion gives

$$f(x+\lambda y) = f(x) + \lambda\langle \nabla f(x), y \rangle + \tfrac{1}{2}\lambda^2\langle \nabla^2 f(x) y, y \rangle + o(\lambda^2). \tag{2.6}$$

Substituting (**??**) into (**??**), we obtain

$$\tfrac{1}{2}\lambda^2\langle \nabla^2 f(x) y, y \rangle + o(\lambda^2) \geq 0.$$

Dividing by λ^2 and letting $\lambda \to 0$ gives (**??**).

Conversely, assume the Hessian is positive semidefinite at every point in C. For $x, y \in C$, Taylor's theorem with remainder gives

$$f(y) = f(x) + \langle \nabla f(x), y - x \rangle + \tfrac{1}{2}\langle \nabla^2 f(\xi)(y - x), (y - x) \rangle,$$

for some $\xi = \lambda x + (1 - \lambda) y \in C$ with $\lambda \in (0, 1)$. Since $\nabla^2 f(\xi)$ is positive semidefinite, the last term is nonnegative, and thus

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle.$$

Therefore, f is convex. $\square$

2.4 Exercises

Exercise 2.1 A hyperplane $H \subset \mathbb{R}^n$ is the set of points satisfying the equation $\langle a, x\rangle = \alpha$, i.e.

$$H = \{x \in \mathbb{R}^n : \langle a, x\rangle = \alpha\}.$$

1. Prove that H is closed.
2. Prove that H is convex.

Exercise 2.2 Plot the following sets in an orthonormal plane and indicate which of them are convex:

1. $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$.
2. $\{(x, y) \in \mathbb{R}^2 : (x-1)^2 + (y-3)^2 \leq 3\}$.
3. $\{(x, y) \in \mathbb{R}^2 : (x-1)^2 + y^2 > 5\}$.
4. $\{(x, y) \in \mathbb{R}^2 : 0 \leq x^2 + y^2 \leq 11\}$.
5. $\{(x, y) \in \mathbb{R}^2 : 2 \leq x^2 + y^2 \leq 4\}$.
6. $\{(x, y) \in \mathbb{R}^2 : y > x^2\}$.
7. $\{(x, y) \in \mathbb{R}^2 : |y| \leq |x| \leq 1\}$.
8. $\{(x, y) \in \mathbb{R}^2 : y > \dfrac{1}{1-x^2}\}$.

Exercise 2.3 Let the sets

$$\begin{aligned}
S_1 &= \{x = (x_1, x_2) : 2 \leq x_1 \leq 5,\ x_2 = 4\},\\
S_2 &= \{x = (x_1, x_2, x_3) : x_1 - x_2 \leq 5,\ -x_1 - x_2 - x_3 \leq 7,\ x_1, x_2, x_3 \geq 0\},\\
S_3 &= \{x = (x_1, x_2, x_3) : x_1^2 - x_2^2 - x_3^2 \leq 9,\ x_1 - x_3 = 2\}.
\end{aligned}$$

1. Are these sets convex?
2. Find the interior and the closure of each set. Are those (interior/closure) sets convex?

Exercise 2.4 Show that C is convex if and only if

$$C = \Big\{ \sum_{i=1}^{k} \lambda_i x_i : k \in \mathbb{N},\ x_i \in C,\ \lambda_i \in [0, 1],\ i = 1, \ldots, k,\ \sum_{i=1}^{k} \lambda_i = 1 \Big\}.$$

Hint: Prove the result by induction on k.

Exercise 2.5 Let $(f_i)_{i\in I}$ be a family of convex functions with index set I (arbitrary). Define

$$f = \sup_{i\in I} f_i.$$

1. Show that $\operatorname{epi} f = \bigcap_{i\in I} \operatorname{epi} f_i$.
2. Deduce that f is convex.

Exercise 2.6 Let $C \subset \mathbb{R}^n$ be a convex set and $f : C \to \mathbb{R}$ a convex function. If $x_1, x_2, \ldots, x_k \in C$ and $\lambda_i \in [0,1]$ satisfy $\sum_{i=1}^k \lambda_i = 1$, prove that

$$f\Big(\sum_{i=1}^k \lambda_i x_i\Big) \le \sum_{i=1}^k \lambda_i f(x_i).$$