Chapter 1

Convex sets and convex functions

In the sequel E denotes a finite dimensional vector space and $|\cdot|$ its euclidean norm.

1.1 Convex sets

Definition 1.1.1 *A subset C of E is called a convex set if*

$$\lambda x + (1-\lambda) y \quad \in \quad C$$

for any $x, y \in C$ and any $\lambda \in [0,1]$.

Definition 1.1.2 (Convex combination) *For $x_1, \ldots, x_n \in E$ and $\lambda_1, \cdots, \lambda_n \in [0,1]$ such that $\sum_{i=1}^n \lambda_i = 1$, the point $\sum_{i=1}^n \lambda_i x_i$ is called convex combination of the points $x_1, \ldots, x_n$.*

Proposition 1.1.1 *A subset C of E is convex if and only if the convex combination*

$$\sum_{i=1}^n \lambda_i x_i \quad \in \quad C$$

for any $x_1, \ldots, x_n \in C$ and $\lambda_1, \ldots, \lambda_n \in [0,1]$ such that $\sum_{i=1}^n \lambda_i = 1$.

Proof. The proof follows from an induction argument. $\square$

Proposition 1.1.2 *Let $(C_i)_{i\in I}$ be a une family of convex subsets of E. Then the intersection $\bigcap_{i\in I} C_i$ is a convex subset of E.*

Proof. The proof follows from the definition of the convexity. $\square$

Definition 1.1.3 (Convex hull) *Let A be a subset of E. The convex hull of A, denoted by $conv(A)$, is defined as the intersection of all convex subsets C of E such that $A \subset E$. By the previous Proposition $conv(A)$ is a convex subset of E.*

Proposition 1.1.3 *$conv(A)$ is the set of all convex combinations of points of A*

$$conv(A) \quad = \quad \Big\{ \sum_{i=1}^{n} \lambda_i x_i \ : \ n \geq 1 \ , \ x_i \in A \ , \ \lambda_i \geq 0 \ , \ \sum_{i=1}^{n} \lambda_i = 1 \Big\} \ .$$

Proof. Define the set

$$\tilde{A} \quad = \quad \Big\{ \sum_{i=1}^{n} \lambda_i x_i \ : \ n \geq 1 \ , \ x_i \in A \ , \ \lambda_i \geq 0 \ , \ \sum_{i=1}^{n} \lambda_i = 1 \Big\} \ .$$

From Proposition 1.1.1, we have $\tilde{A} \subset C$ for any convex set C such that $A \subset C$. Therefore $\tilde{A} \subset conv(A)$. We next notice that $\tilde{A}$ is a convex set. Therefore $conv(A) \subset \tilde{A}$. $\square$

Definition 1.1.4 (Extremal point) *Let C be a convex subset of E. A point $x \in C$ is called an extremal point of C if*

$$x \quad \neq \quad \lambda y + (1-\lambda) z$$

for any $y, z \in C$ such that $y \neq z$ and any $\lambda \in (0,1)$. The set of extremal points of C will be denoted by $ext(C)$.

Proposition 1.1.4 *The following assertions are equivalent.*

(i) $x \in ext(C)$.

(ii) $x \neq \frac{y+z}{2}$ *for all* $y, z \in C$ *such that* $y \neq z$

(iii) $x = \frac{y+z}{2}$ *with* $y, z \in C \Rightarrow x = y = z$.

Proof. We obviously have $(i) \Rightarrow (ii) \Rightarrow (iii)$. Consider $x \in C$ satisfying (iii), and let $y, z \in C$ and $\lambda \in (0,1)$ such that $x = \lambda y + (1-\lambda)z$. If $\lambda = 1/2$ then from (iii) we have $x = y = z$. If $\lambda \in (0, 1/2)$, then $z' := 2x - z \in C$ as C is convex and $z' = 2\lambda y + (1-2\lambda)z$. Since $x = (z+z')/2$ we get from (iii) that $x = z$ and therefore $x = y$. If $\lambda \in (0, 1/2)$, we apply the same argument with y in place of z. □

Proposition 1.1.5 *(i) $x \in ext(C)$ if and only if $C \setminus \{x\}$ is convex. (ii) We have $ext(conv(A)) \subset A$ for $A \subset E$.*

Proof. (i) Let $x \in ext(C)$, $y, z \in C \setminus \{x\}$ and $\lambda \in [0,1]$. We then have $\lambda y + (1-\lambda)z \in C$. If $\lambda y + (1-\lambda)z = x$ then $y = z = x$ which is not possible. Therefore $\lambda y + (1-\lambda)z \in C \setminus \{x\}$. Conversely, suppose that $C \setminus \{x\}$ is convex. Let $y, z \in C$ with $y \neq z$ and $\lambda \in (0,1)$ such that $x = \lambda y + (1-\lambda)z$. If $y = x$ or $z = x$ we have $x = y = z$ from $x = \lambda y + (1-\lambda)z$. Therefore, $y, z \in C \setminus \{x\}$. From the convexity of $C \setminus \{x\}$ we get $x \in C \setminus \{x\}$ which is impossible. Hence $x \neq \lambda y + (1-\lambda)z$.
(ii) Let $x \in conv(A) \setminus A$. Then, there exists $n \geq 1$, $x_1, \ldots, x_n \in A$ and $\lambda_1, \ldots, \lambda_n \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$ and $\sum_{i=1}^n \lambda_i x_i = x$. Without loss of generality, we suppose that $\lambda_1 \geq \cdots \geq \lambda_n > 0$ and $x_n \notin conv(x_1, \ldots, x_{n-1})$. We then notice that $n \geq 2$. Therefore, $\lambda := \sum_{i=1}^{n-1} \lambda_i \in (0,1)$ and

$$x = \lambda y + (1-\lambda)z$$

with $y = \sum_{i=1}^{n-1} \frac{\lambda_i}{\lambda} x_i \in conv(A)$ and $z = x_n \neq y$. Therefore, $x \notin ext(conv(A))$.
□

Theorem 1.1.1 (Minkowski (or Krein-Milman in infinite dimension)) *Let A be a convex compact subset of E. Then $conv(ext(A)) = A$.*

1.2 Convex functions

Definition 1.2.5 (Convex function) *A function $f : E \to \mathbb{R} \cup \{+\infty\}$ is said to be convex if*

$$f\big(\lambda x + (1-\lambda)y\big) \quad \leq \quad \lambda f(x) + (1-\lambda) f(y)$$

for all $x, y \in E$ *with* $f(x) < +\infty$, $f(y) < +\infty$ *and all* $\lambda \in [0,1]$.

We notice that for f convex function, the sets $A_f(a) = \{x \in E \ : \ f(x) \leq a\}$ and $B_f(a) = \{x \in E \ : \ f(x) < a\}$ are convex sets. Unfortunately, the reverse implication is not true. Indeed, for $f : \mathbb{R} \to \mathbb{R}$ monotone function, the sets $A_f(a)$ and $B_f(a)$ are convex sets as they are intervals of $\mathbb{R}$.

Definition 1.2.6 (Domain of a function) *The domain of a function* $f : E \to \mathbb{R} \cup \{+\infty\}$ *is the set*

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

Remark 1.2.1 *The function* $f : E \to \mathbb{R} \cup \{+\infty\}$ *is convex if and only if 1) its domain* $dom(f)$ *is convex and 2) its restriction* $f|_{dom(f)}$ *to its domain is convex.*

Definition 1.2.7 (Strictly convex function) *A function* $f : E \to \mathbb{R} \cup \{+\infty\}$ *is said to be strictly convex if it is convex and*

$$f\big(\lambda x + (1-\lambda)y\big) \quad < \quad \lambda f(x) + (1-\lambda) f(y)$$

for all $x, y \in dom(f)$ *with* $x \neq y$ *and all* $\lambda \in (0,1)$.

Proposition 1.2.6 *let* C *be a convex subset of* E *and* $f : E \to \mathbb{R} \cup \{+\infty\}$ *be a strictly convex function such that* $C \subset dom(f)$. *Let* $x_0 \in C$ *such that*

$$f(x_0) \quad = \quad \max_{x \in C} f(x) \ .$$

Then $x_0 \in ext(C)$.

Proof. Let $M = \sup_{x \in C} f(x)$. If $x_0 \in C$ such that $f(x_0) = M$ and $x \notin ext(C)$, we can find $y, z \in C$ with $x \neq z$ such that $x_0 = (y+z)/2$. We then have $M = f(x_0) < (f(y) + f(z))/2 \leq \max(f(y), f(z))$. Thus f takes a value strictly greater than M at y or z, which contradicts the definition of M. $\square$

Proposition 1.2.7 (Convex extension) *If* $f : C \to \mathbb{R}$ *is a convex function defined on the convex set* C, *it can be extended to convex function* $\tilde{f} : E \to \mathbb{R} \cup \{+\infty\}$ *by setting*

$$\tilde{f}(x) \quad = \quad \begin{cases} f(x) & \textit{if } x \in C, \\ +\infty & \textit{else.} \end{cases}$$

This result allows to manipulate simultaneously convex functions defined on different convex sets. In particular, if $A \subset E$ we define its characteristic function χ_A by

$$\chi_A(x) \quad = \quad \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{else.} \end{cases}$$

Then, A is convex if and only if χ_A is convex.

Definition 1.2.8 (Epigraph of a function) *Let $f : E \to \mathbb{R} \cup \{+\infty\}$. Its epigraph is the subset of $E \times \mathbb{R}$ defined by*

$$epi(f) \quad = \quad \big\{(x,t) \in E \times \mathbb{R} \ : \ f(x) \leq t\big\} \ .$$

The strict epigraph of f is defined by

$$stepi(f) \quad = \quad \big\{(x,t) \in E \times \mathbb{R} \ : \ f(x) < t\big\} \ .$$

Proposition 1.2.8 *i) The function $f : E \to \mathbb{R} \cup \{+\infty\}$ is convex if and only if $epi(f)$ is convex.*

ii) If the function $f : E \to \mathbb{R} \cup \{+\infty\}$ is strictly convex then if its $stepi(f)$ is convex.

Proof. The proof follows from the definitions of convexity for functions and sets. □

Proposition 1.2.9 *(i) Let $(f_i)_{i \in I}$ be a family of convex functions. Its upper envelope f defined by $f = \sup_{i \in I} f_i$ is a convex function.*

(ii) If g is the pointwise limit of a sequence $(g_n)_{n \geq 0}$ of convex functions, then g is convex.

Proof. Fix x, y and $\lambda \in [0,1]$.

(i) Since each f_i is convex we have

$$f_i(\lambda x + (1-\lambda)y) \quad \leq \quad \lambda f_i(x) + (1-\lambda) f_i(y)$$

for any $i \in I$. Taking the supremum we get

$$\begin{aligned} \sup_{i\in I} f_i(\lambda x + (1-\lambda)y) &\leq \sup_{i\in I} \big\{\lambda f_i(x) + (1-\lambda) f_i(y)\big\} \\ &\leq \lambda \sup_{i\in I} f_i(x) + (1-\lambda) \sup_{i\in I} f_i(y) \end{aligned}$$

(ii) From the convexity of g_n we have

$$g_n(\lambda x + (1-\lambda)y) \leq \lambda g_n(x) + (1-\lambda) g_n(y) .$$

Taking the limit as n goes to ∞, we get

$$g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda) g(y) .$$

□

We may also need to consider convex functions valued in $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. We then take the convexity property of the epigraph as definition.

Definition 1.2.9 (Convex function valued in $\bar{\mathbb{R}}$) . *The function $f : E \to \bar{\mathbb{R}}$ is said to be convex if its epigraph*

$$epi(f) = \big\{(x,t) \in E \times \mathbb{R} \ : \ f(x) \leq t\big\} .$$

is convex.

In the case where f is valued in $\mathbb{R} \cup \{+\infty\}$, the definition of convexity coincides with the initial one. For the case where f is valued in $\bar{\mathbb{R}}$, we have the following result.

Proposition 1.2.10 *A function $f : E \to \bar{\mathbb{R}}$ is convex if and only if*

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda) f(y)$$

for all $x, y \in E$ with $f(x) < +\infty$, $f(y) < +\infty$ and all $\lambda \in [0,1]$, with the convention $-\infty + t = -\infty$, for $t < +\infty$, $a \cdot (-\infty) = -\infty$, for $a > 0$, and $0 \cdot (-\infty) = 0$.

Proof. The proof follows from the definition of convexity and the conventions mentioned above. □

Remark 1.2.2 *We notice that Proposition 1.2.9 holds functions valued in* $\bar{\mathbb{R}}$.

We end this chapter by the following definition.

Definition 1.2.10 (Proper convex function) *A convex function is said to be proper if it does not take the value* $-\infty$*, and is not identically equal to* $+\infty$*. If not, the function is said to be unproper.*