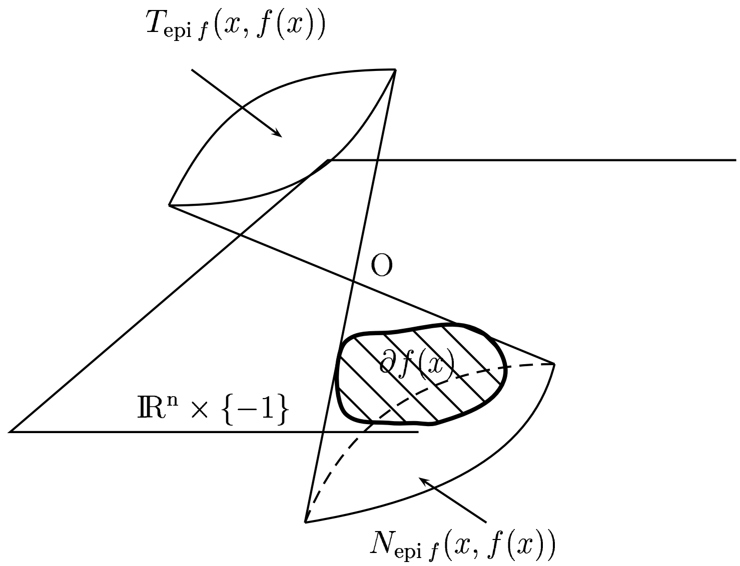
UNIVERSITY OF KAISERSLAUTERN

Department of Mathematics



Convex Analysis

Dr. Stefan Nickel

Preface

Convex analysis is one of the mathematical tools which is used both explicitly and indirectly in many mathematical disciplines. However, there are not so many courses which have convex analysis as the main topic. More often, parts of convex analysis are taught in courses like linear or nonlinear optimization, probability theory, geometry, location theory, etc.. This manuscript gives a systematic introduction to the concepts of convex analysis. A focus is set to the geometrical interpretation of convex analysis. This focus was one of the reasons why I have decided to restrict myself to the finite dimensional case. Another reason for this restriction is that in the infinite dimensional case many proofs become more difficult and more technical. Therefore, it would not have been possible (for me) to cover all the topics I wanted to discuss in this introductory text in the infinite dimensional case, too. Anyway, I am convinced that even for someone who is interested in the infinite dimensional case this manuscript will be a good starting point.

When I offered a course in convex analysis in the Wintersemester 1997/1998 (upon which this manuscript is based) a lot of students asked me how this course fits in their own studies. Because this manuscript will (hopefully) be used by some students in the future, I will give here some of the possible statements to answer this very question.

- Convex analysis can be seen as an extension of classical analysis, in which still we get many of the results, like a mean-value theorem, with less assumptions on the smoothness of the function.

- Convex analysis can be seen as a foundation of linear and nonlinear optimization which provides many tools to handle concepts in optimization much easier (for example the Lemma of Farkas).

- Finally, convex analysis can be seen as a link between abstract geometry and very algorithmic oriented computational geometry.

As already explained before, this manuscript is based on a one semester course and therefore cannot cover all topics and discuss all aspects of convex analysis in detail. To guide the interested reader I have included a list of nice books about this subject at the end of the manuscript. It should be noted that the philosophy of this course follows [3], [4] and THE BOOK of modern convex analysis [6]. The geometrical emphasis however, is also related to intentions of [1].

At the end of this preface I would like to thank Dr. Matthias Ehrgott for preparing the exercises for this manuscript and his general support. I also would like to thank Chokri Hamdaoui, who typed this manuscript with a great enthusiasm. Last but not least I thank all the members of the research group of Prof. Hamacher for their help in proof-reading and for their useful remarks.

Stefan Nickel, October 1998

Contents

List of Tables

List of Figures

Part I

CONVEX SETS

Chapter 1

BASIC CONCEPTS

1.1 Definition and Important Examples

Definition 1 *A set* $C \subset \mathbb{R}^n$ *is called* **convex** *if* $\alpha x + (1-\alpha)x' \in C$ *whenever* $x, x' \in C$, *and* $\alpha \in [0,1]$ *(or equivalently,* $\alpha \in (0,1)$*).*

To give a more geometrical definition of a convex set, we need

Definition 2 *For* $x,\ x' \in \mathbb{R}^n$, *the* **line segment** $[\boldsymbol{x}, \boldsymbol{x'}]$ *connecting* x *and* x' *is defined as* $[x, x'] := \{\alpha x + (1-\alpha)x' : 0 \leq \alpha \leq 1\}$.

Definition 3 (Geometrical Definition of a Convex Set) *A set* C *is called* **convex** *if the line segment* $[x, x'] \subset C$ *whenever* $x,\ x' \in C$.

From a geometrical point of view, there is another type of sets closely related to convex sets.

Definition 4 *A set* $C \subset \mathbb{R}^n$ *is called* **star-shaped with respect to** $\boldsymbol{x} \in \mathbb{R}^n$ *whenever* $[x, x'] \subset C$ *for all* $x' \in C$. *We simply say* C *is* **star-shaped** *if* C *is star-shaped with respect to* $0 \in \mathbb{R}^n$.

Therefore, we can say:

- A set $C \subset \mathbb{R}^n$ is convex if it is star-shaped with respect to x whenever $x \in C$, or
- A set $C \subset \mathbb{R}^n$ is convex if $C - \{x\} := \{y : y = C - x \text{ with } c \in C\}$ is star-shaped whenever $x \in C$.

Remark As a consequence, we get that a convex set is path-connected.

Example 1.1 *Some examples of sets are given in Figure 1.1.*

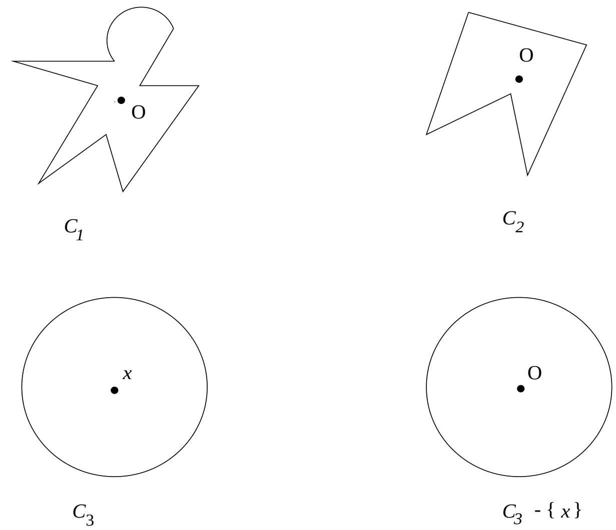


Figure 1.1: C_1 is not convex, not star-shaped; C_2 is star-shaped, but not convex; C_3 is convex.

Important Examples of Convex Sets

- Intervals in $\mathbb{R}$.

Definition 5 **An (affine) hyperplane** *is is a set associated with* $(r, s) \in \mathbb{R}^n \times \mathbb{R}$, $s \neq 0$, *and defined by*

$$H_{s,r} := \{x \in \mathbb{R}^n : \langle s, x \rangle \geq r\}.$$

- $H_{s,r}$ is a convex set.
  (Convexity:

$$\begin{array}{rcl} x, x' \in H_{s,r}, \alpha \in [0,1], \langle s, \alpha x + (1-\alpha)x' \rangle & = & \alpha \langle s, x \rangle + (1-\alpha)\langle s, x' \rangle \\ & = & \alpha r + (1-\alpha) r = r). \end{array}$$

Remark $H_{s,0} = \{s\}^{\perp} = \{x \in \mathbb{R}^n : \langle s, x \rangle = 0\}$. s is **normal** to $H_{s,0}$.

- Any affine subspace (or affine manifold) is convex by virtue of the vector space axioms.

Definition 6 *The* **half-spaces** *of* $\mathbb{R}^n$ *with respect to (w.r.t.)* $(s, r) \in \mathbb{R}^n \times \mathbb{R}$, $s \neq 0$, *are defined by*

$$H^{\leq}_{s,r} := \{x \in \mathbb{R}^n : \langle s, x \rangle \leq r\}$$

*(**closed** half-space) and*

$$H^{<}_{s,r} := \{x \in \mathbb{R}^n : \langle s, x\rangle < r\}$$

*(**open** half-space)*

- $H^{\leq}_{s,r}$ and $H^{<}_{s,r}$ are convex sets.

Definition 7 *Call* $\alpha = (\alpha_1, \ldots, \alpha_k)$ *the* **generic point** *of* $\mathbb{R}^{\mathrm{k}}$*. The* **unit simplex** *in* $\mathbb{R}^k$ *is defined as*

$$\begin{aligned}\Delta_k : &= \{\alpha \in \mathbb{R}^k : \sum_{i=1}^{k} \alpha_i = 1, \alpha_i \geq 0, i = 1, \ldots, k\} \\ &= \{\alpha \in \mathbb{R}^k : e^T\alpha = 1, e_i^T\alpha \geq 0, i = 1, \ldots, k\},\end{aligned}$$

where $\{e_1, \ldots, e_k\}$ *is the canonical basis of* $\mathbb{R}^k$*,* e_l *is the* l^{th} *unit vector, and* $e = (1, 1, \ldots, 1)^T$*. (default: row vectors* $e = (1, \ldots, 1) \hat{=} (1, \ldots, 1)^T$*, and* e^T *is a column vector).*

Remark Unit simplices are convex, compact and have empty interior (w.r.t. $\mathbb{R}^k$).

Example 1.2 *Refer to Figure 1.2.*

$$\Delta_2 = \{\alpha \in \mathbb{R}^2 : \alpha_1 + \alpha_2 = 1, \alpha_i \geq 0, i = 1, 2\}.$$

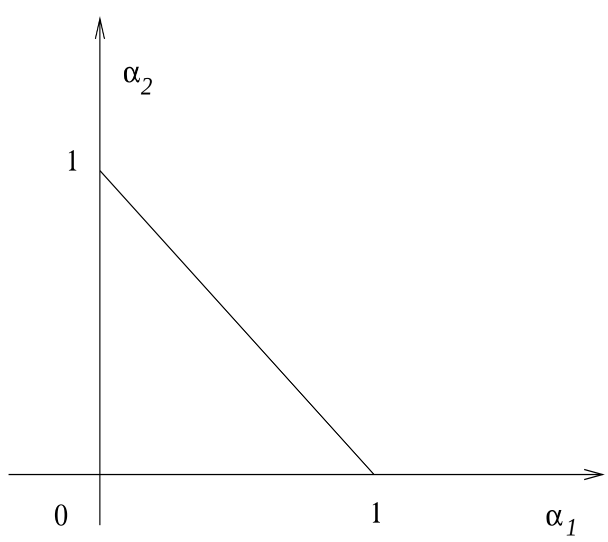


Figure 1.2: Example of simplex in $\mathbb{R}^2$.

We will often refer to an $\alpha \in \Delta_k$ as a set of (k) **convex multipliers**. We can embed Δ_k in $\mathbb{R}^m$, $m > k$, by appending $m-k$ zeros to α and obtaining a vector of Δ_m. In this sense, the unit simplex of $\mathbb{R}^k$ is a simplex of $\mathbb{R}^n$ in the hyperplane given by $e^T\alpha = 1$.

Example 1.3 *Refer to Figure 1.3.*

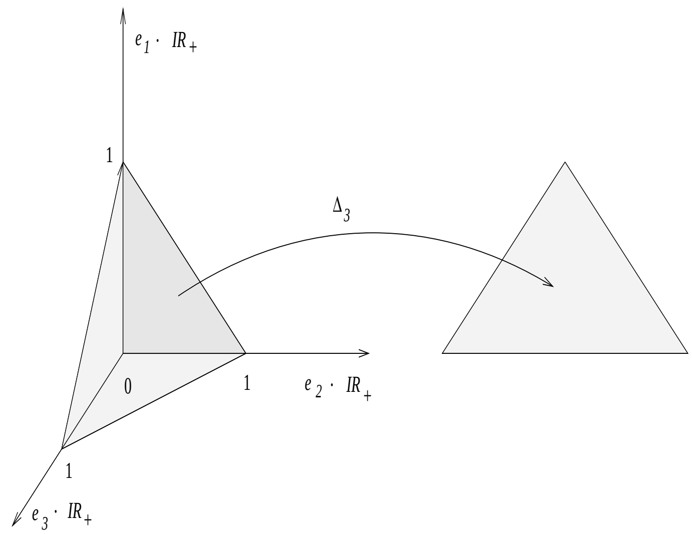


Figure 1.3: Example of simplex in $\mathbb{R}^3$.

To get sets with nonempty interior, we define

$$\Delta_k' := \{\alpha \in \mathbb{R}^k : e^T\alpha \leq 1, \alpha_i \geq 0, i = 1, \ldots, k\}.$$

We have

$$\alpha \in \Delta_k' \Longleftrightarrow \exists\, \alpha_{k+1} \geq 0 : (\alpha, \alpha_{k+1}) \in \Delta_{k+1}.$$

Therefore, $\Delta_k' \subset \mathbb{R}^k$ can be identified with Δ_{k+1} by a projection operator.

Definition 8 *A set $K \subset \mathbb{R}^n$ is called a* **cone** *if $\alpha x \in K$ whenever $x \in K, \alpha > 0$. A cone K is called a* **convex cone** *if K is convex.*

Remark Checking for convexity of a cone K can be reduced to checking if $x + y \in K$ whenever $x, y \in K$ (or $K + K \subset K$).

Example 1.4 *The set*

$$K = \{x \in \mathbb{R}^n : \langle s_j, x\rangle = 0,\ j = 1, \ldots, m, \langle s_{m+j}, x\rangle \le 0,\ j = 1, \ldots, p\}$$

with $s_j \in \mathbb{R}^n$ *is a convex cone.*

1. *K is a cone:* $x \in K, \alpha > 0 \Longrightarrow$
   - $\langle s_j, \alpha x\rangle = \alpha\langle s_j, x\rangle \ge 0$
   - $\langle s_{m+j}, \alpha x\rangle = \alpha\langle s_{m+j}, x\rangle \le 0$
2. *K is a convex cone:* $x, y \in K \Longrightarrow$
   - $\langle s_j, x + y\rangle = \langle s_j, x\rangle + \langle s_j, y\rangle = 0$
   - $\langle s_{m+j}, x + y\rangle = \langle s_{m+j}, x\rangle + \langle s_{m+j}, y\rangle \le 0$

Therefore,

$$\begin{aligned} \Omega_+ : &= \{x \in \mathbb{R}^n : \langle e_i, x\rangle \ge 0, i = 1, \ldots, n\} \\ &= \{x = (\xi^1, \ldots, \xi^n) : \xi^i \ge 0, i = 1, \ldots, n\}. \end{aligned}$$

The **nonnegative orthant** of $\mathbb{R}^n$ is a convex cone.

1.2 Operations on Convex Sets Preserving Convexity

Proposition 1.1

a) *Let* $\{C_j\}_{j\in\mathcal{J}}, C_j \subset \mathbb{R}^n$ *be an arbitrary family of convex sets. Then* $C := \bigcap_{j\in\mathcal{J}} C_j$ *is convex.*

b) *Let* $C_1, \ldots, C_k$, *with* $C_i \subset \mathbb{R}^{n_i}$. *Then* $C_1, \ldots, C_k$ *are convex sets* $\Longleftrightarrow C_1 \times \ldots \times C_k$ *is a convex set in* $\mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_k}$.

c) *Let* $A : \mathbb{R}^n \to \mathbb{R}^m$ *be an affine mapping, i.e.*

$$A(\alpha x + (1 - \alpha)x') = \alpha A(x) + (1 - \alpha)A(x'), x, x' \in \mathbb{R}^n, \alpha \in \mathbb{R},$$

and $C \subset \mathbb{R}^n$ *convex. Then* $A(C) \subset \mathbb{R}^m$ *is convex. If* $D \in \mathbb{R}^m$ *is a convex set then* $A^{-1}(D) := \{x \in \mathbb{R}^n : A(x) \in D\}$ *is convex in* $\mathbb{R}^n$.

Proof:

a) Let $x,\ x' \in C \Rightarrow x,\ x' \in C_j,\ j \in \mathcal{J}$
$\Rightarrow x'' := \alpha x + (1-\alpha)x' \in C_j, j \in \mathcal{J}, \alpha \in [0,1] \Rightarrow x'' \in C.$

b) "$\Longrightarrow$": $C_1, \dots, C_k$ are convex. Let $x, x' \in C := C_1 \times \dots \times C_k$

$$\begin{aligned} &\Rightarrow \quad x = (x_1, \dots, x_k), x = (x_1^{'}, \dots, x_k^{'}) \text{ with } x_i, x_i^{'} \in C_i, i = 1, \dots, k \\ &\Rightarrow \quad \alpha x + (1-\alpha)x' = (\alpha x_1, \dots, \alpha x_k) + ((1-\alpha)x_1^{'}, \dots, (1-\alpha)x_k^{'}) \in C. \end{aligned}$$

"$\Longleftarrow$": Follows from c) since projection on the components is an affine mapping.

c) Let $x, x' \in \mathrm{IR^n}$. Then $A([x,x']) = [A(x), A(x')]$. Hence $[x,x'] \subset C$ implies $A([x,x']) \subset A(C)$. Therefore C is convex $\Rightarrow A(C)$ is convex.
Let $D \subset \mathrm{IR^m}$ be convex and $y, y' \in D$ with $y = A(x), y' = A(x')$
$\Rightarrow [y, y'] \subset D$ implies $[x, x'] \subset A^{-1}$.

□

Remark Let $C_1, \dots, C_k$ be convex. Then it is in general not true that $C := \bigcup_{i=1}^{k} C_i$ is convex.

Consequences of Proposition 1.1

- The polyhedron $P := \{x \in \mathrm{IR^n} : \langle s_j, x\rangle \le r_j,\ \ j = 1, \dots, m\}$ with $(s_1, r_1), \dots, (s_m, r_m) \in \mathrm{IR^n} \times \mathrm{IR}$ is convex as the intersection of m half-spaces. An example is given in Figure 1.4.
- If $C \subset \mathrm{IR^n}$ is convex then $-C := \{x \in \mathrm{IR^n} : -x \in C\}$ is convex.
- **The Minkowski sum** of two convex sets $C_1, C_2 \subset \mathrm{IR^n}$, defined as

$$C_1 + C_2 := \{x = x_1 + x_2 : x_1 \in C_1, x_2 \in C_2\},$$

is convex. (From Proposition 1.1 b) C_1, C_2 convex $\Rightarrow C_1 \times C_2$ convex, and by Proposition 1.1 c) with the affine mapping
$C_1 \times C_2 \to \mathrm{IR^n}, (x_1, x_2) \mapsto \alpha_1 x_1 + \alpha_2 x_2, \quad \alpha_1, \alpha_2 \in \mathrm{IR}$).

Remark Closed sets need not be mapped to closed sets by the Minkowski sum:

$$C_1 := \{(\xi, \eta) : \xi \ge 0, \xi\eta \ge 1\},\ C_2 := \mathrm{IR} \times \{0\}.$$

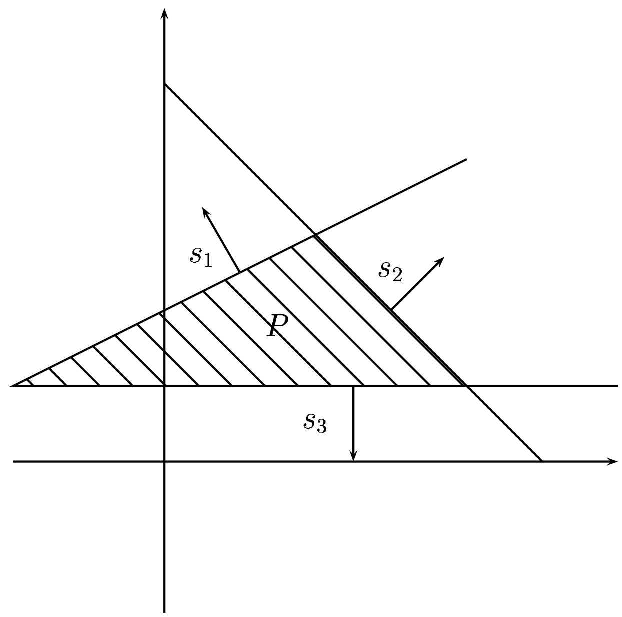


Figure 1.4: Example of a polyhedron.

$C_1 + C_2 = \{(x_1, x_2) : x_1 \in \mathbb{R},\ x_2 > 0\}$ is not closed, as shown in Figure 1.2.

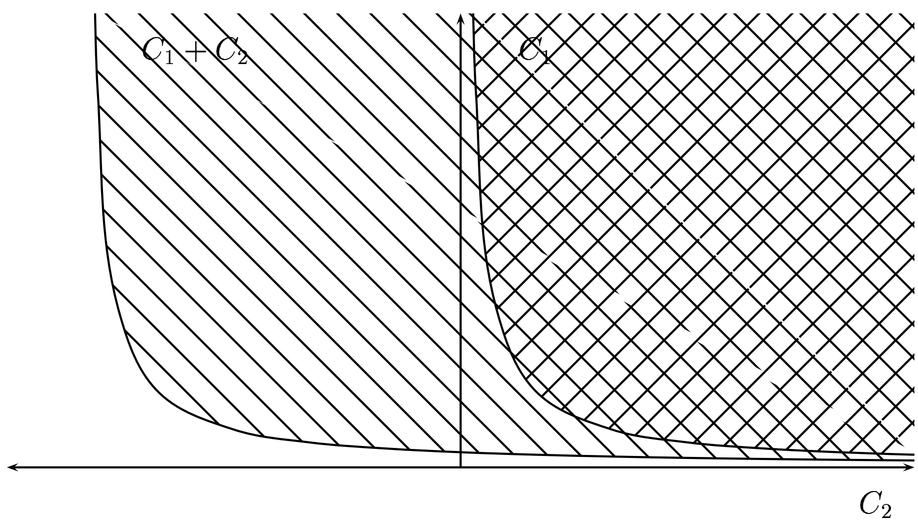


Figure 1.5: $C_1 + C_2 = \{(x_1, x_2) : x_1 \in \mathbb{R}, x_2 > 0\}$, is not closed.

Example 1.5 *Let $C \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ be convex. Use a projection $A : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{n_1}$ to see that the 'slice' of C along y, $C(y) := \{x \in \mathbb{R}^{n_1} : (x, y) \in C\}$ and the shadow of C over $\mathbb{R}^{n_1}$ $C_1 := \{x \in \mathbb{R}^{n_1} : (x, y) \in C \text{ for some } y \in \mathbb{R}^{n_2}\}$ are convex. Refer to Figure 1.6.*

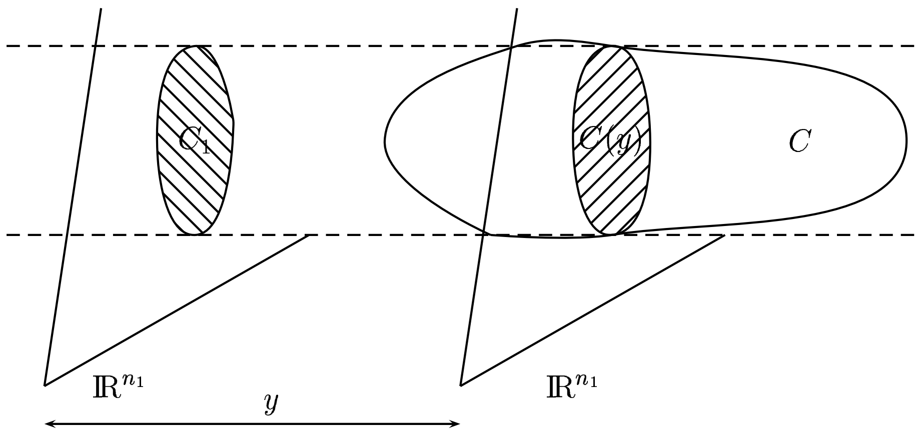


Figure 1.6: Illustration for Example 1.5.

Proposition 1.2 *If $C \subset \mathbb{R}^n$ is convex then its interior* int C *and its closure* cl C *are convex.*

Proof: For $x, x' \in C, x \neq x'$ and $\alpha \in (0, 1)$ we set $x'' := \alpha x + (1 - \alpha)x' \in [x, x']$ and $x'' \neq x', x'' \neq x$.

a) Let $x, x' \in \text{int } C$. Then we can choose $\delta > 0$, such that

$$B(x', \delta) := \{y \in \mathbb{R}^n : \|y - x'\| \leq \delta\} \subset C.$$

From $\frac{\|x'' - x\|}{\|x' - x\|} = (1 - \alpha)$ we get

$$B(x'', (1 - \alpha)\delta) = \alpha x + (1 - \alpha)B(x', \delta) \subset C,$$

since $x \in \text{int } C$ and $B(x', \delta) \subset \text{int } C \Rightarrow x'' \in \text{int } C$ (Refer to Figure 1.7).

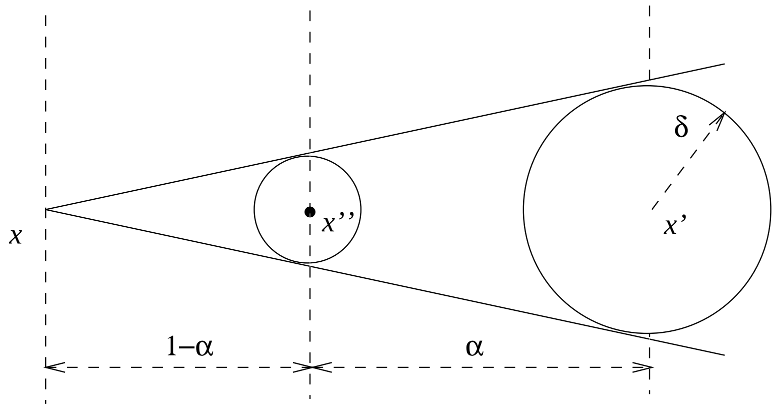


Figure 1.7: Illustration for proof of Proposition 1.2.

b) Now let $x, x' \in \text{cl } C$. Select two sequences $\{x_k\}$ and $\{x_k'\}$ in C converging to x and x' respectively. Then $\alpha x_k + (1-\alpha)x_k' \in C$ for all k and converges to $x'' \Rightarrow x'' \in \text{cl } C$.

□

1.3 Convex Combinations and Convex Hulls

First we recall some definitions from linear algebra.

Definition 9 *Let* $x_1, \ldots, x_k \in \mathbb{R}^n, \alpha_1, \ldots, \alpha_k \in \mathbb{R}$.

a) $\sum_{i=1}^{k} \alpha_i x_i$ *is called a* **linear combination of $x_1, \ldots, x_k$**. *A* **linear subspace of $\mathbb{R}^n$** *is a set containing all its linear combinations. For* $S \subset \mathbb{R}^n$ *the set*

$$\{x \in \mathbb{R}^n : x \textit{ is a linear combination of elements from } S\}$$

is called **linear hull of S** *(* $\text{lin } S$*).*

b) $\sum_{i=1}^{k} \alpha_i x_i$ *with* $\sum_{i=1}^{k} \alpha_i = 1$ *is called an* **affine combination of $x_1, \ldots, x_k$**. *An* **affine manifold of $\mathbb{R}^n$** *is a set containing all its affine combinations. For* $S \subset \mathbb{R}^n$*, the set*

$$\{x \in \mathbb{R}^n : x \textit{ is an affine combination of elements from } S\}$$

is called **affine hull of S** *(*$\text{aff } S$*).*

Definition 10 *A* **convex combination** *of* $x_1, \dots, x_k \in \mathbb{R}^n$ *is an element of the form* $\sum_{i=1}^{k} \alpha_i x_i$ *with* $\sum_{i=1}^{k} \alpha_i = 1$ *and* $\alpha_i \in \mathbb{R}_{+0}, i = 1, \dots, k$ *(*$\mathbb{R}_{+0} := \{x \in \mathbb{R} : x \geq 0\}$*).*

Remark Every convex combination is an affine combination and every affine combination is a linear combination.

Proposition 1.3 *Let* $C \subset \mathbb{R}^n$*.* C *is convex* $\iff$ C *contains every convex combination of its elements.*

Proof: "$\Longleftarrow$": Clear, since the convex combinations of two elements are just line segments.
"$\Longrightarrow$": Let $C \subset \mathbb{R}^n$ be convex and take $x_1, \dots, x_k \in C$, $\alpha = (\alpha_1, \dots, \alpha_k) \in \Delta_k (\alpha_i \geq 0, \sum \alpha_i = 1)$. At least one $\alpha_i > 0$ $(\sum \alpha_i = 1)$. Without loss of generality (*wlog*) assume that $\alpha_1 > 0$. Now define

$$y_2 := \frac{\alpha_1}{\alpha_1 + \alpha_2} x_1 + \frac{\alpha_2}{\alpha_1 + \alpha_2} x_2 \in C,$$

since $\frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2} = 1$.
Therefore

$$y_3 := \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + \alpha_3} y_2 + \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} x_3 = \frac{1}{\alpha_1 + \alpha_2 + \alpha_3} \sum_{i=1}^{3} \alpha_i x_i \in C.$$

We continue with the same argument until

$$\begin{aligned} y_k : &= \frac{\alpha_1 + \dots + \alpha_{k-1}}{\sum_{i=1}^{k} \alpha_i = 1} y_{k-1} + \frac{\alpha_k}{\sum_{i=1}^{k} \alpha_i = 1} x_k \\ &= \sum_{i=1}^{k} \alpha_i x_i \in C. \end{aligned}$$

□

Example 1.6 *See Figure 1.8.*

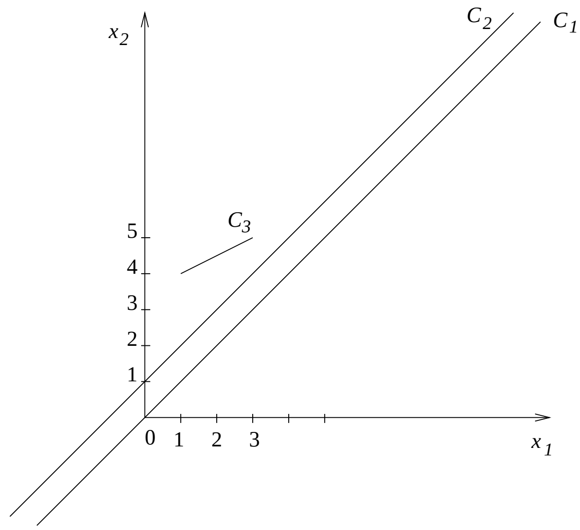


Figure 1.8: Example of sets in $\mathbb{R}^2$.

$C_1 = \{(x,y) \in \mathbb{R}^2 : x = y\}$ *is a linear subspace of* $\mathbb{R}^2$.
$C_2 = \{(x,y) \in \mathbb{R}^2 : x = y - 1\}$ *is an affine subspace of* $\mathbb{R}^2$.
$C_3 = \{x' \in \mathbb{R}^2 : x' = \alpha(1,4)^T + (1-\alpha)(3,5)^T, \alpha \in [0,1]\}$ *is a convex set in* $\mathbb{R}^2$.

Definition 11 *Let* $S \subset \mathbb{R}^n, S \neq \emptyset$. *The* **convex hull of** S (co S) **is defined as**

$$\text{co } S := \bigcap \{C : S \subset C, C \text{ convex}\}.$$

Proposition 1.4 co $S = \{x \in \mathbb{R}^n$*: for some* $k \in \mathbb{N}$*, there exist* $x_1, \ldots, x_k \in S$ *and* $\alpha = (\alpha_1, \ldots, \alpha_k) \in \Delta_k$ *such that* $\sum_{i=1}^{k} \alpha_i x_i = x\} =: T$.

Proof:

- We have (by def. of T) $T \supset S$.
- Let $C \supset S$, C convex. Proposition 1.3 $\Longrightarrow$ C contains all convex combinations of $S \Rightarrow C \supset T$.
- Now we show that T is convex which implies $T \supset$ co S, and with co $S = C$, co $S \supset T$.

Take $x = \sum_{i=1}^{k} \alpha_i x_i, y = \sum_{j=1}^{l} \beta_j y_j \in T$, $\alpha_1, \ldots, \alpha_k \in \mathbb{R}_{+0}$, $\beta_1, \ldots, \beta_l \in \mathbb{R}_{+0}, \alpha = (\alpha_1, \ldots, \alpha_k) \in \Delta_k, \beta = (\beta_1, \ldots, \beta_l) \in \Delta_l$. Furthermore, let $\lambda \in (0,1) \Rightarrow \lambda x + (1-\lambda)y$ be a certain combination of $k+l$ elements of T.

$$\lambda\alpha_i \geq 0, (1-\lambda)\beta_j \geq 0, i = 1, \ldots, k, j = 1, \ldots, l,$$

$$\text{and } \lambda \sum_{i=1}^{k} \alpha_i + (1-\lambda) \sum_{j=1}^{l} \beta_j = \lambda + 1 - \lambda = 1$$

$\Rightarrow \lambda x + (1-\lambda)y \in T$, since we found

$$(\lambda\alpha_1, \ldots, \lambda\alpha_k, (1-\lambda)\beta_1, \ldots, (1-\lambda)\beta_l) \in \Delta_{k+l}.$$

□

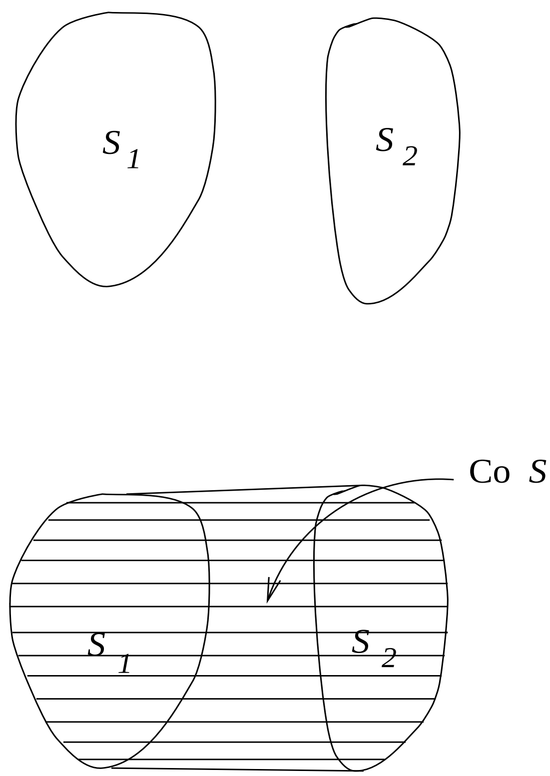


Figure 1.9: Convex hull, co S, of the union of S_1 and S_2.

Example 1.7 *Let S_1, S_2 be convex sets in $\mathbb{R}^n$. $S = S_1 \cup S_2$.* co S *consists only of convex combination with $x_1 \in S_1, x_2 \in S_2$. Refer to Figure 1.9.*

Theorem 1.5 (C. Caratheodory (1873-1950)) *Let $S \subset \mathbb{R}^n$. Any $x \in \operatorname{co} S$ can be represented as a convex combination of $n+1$ elements of S.*

Proof: Let $x = \sum_{i=1}^{k} \alpha_i x_i \in \operatorname{co} S$ be a convex combination of $x_1, \ldots, x_k \in S, k > n+1$. Now we show that one α_i can be set to 0 without changing x. If an $\alpha_i = 0$ we are done. Therefore, assume $\alpha_i > 0$, $i = 1, \ldots, k$. Since $k > n+1$, $x_1, \ldots, x_k$ are affinely dependent $\Rightarrow \exists\ \delta_1, \ldots, \delta_k$ with $(\delta_1, \ldots, \delta_k \neq 0)$, such that $\sum_{i=1}^{k} \delta_i x_i = 0$, $\sum_{i=1}^{k} \delta_i = 0$.
There is at least one $\delta_j > 0$ ($\sum \delta_i = 0$) and we can set

$$\alpha_i^{'} := \alpha_i - t^* \delta_i, i = 1, \ldots, k,$$

where

$$\begin{aligned} t^* &:= \max\{t \geq 0 : \alpha_i - t\delta_i \geq 0, i = 1, \ldots, k\} \\ &= \min_{\delta_j \geq 0}\{\frac{\alpha_j}{\delta_j}\} \end{aligned}$$

We have

- $\alpha_i^{'} \geq 0$, $i = 1, \ldots, k$ (either $\delta_i \leq 0$ or by construction of t^*)
- $\sum_{i=1}^{k} \alpha_i^{'} = \sum_{i=1}^{k} \alpha_i - t^* \sum_{i=1}^{k} \delta_i = 1$
- $\sum_{i=1}^{k} \alpha_i^{'} x_i = x - t^* \sum_{i=1}^{k} \delta_i x_i = x$
- $\exists\ i_0$ such that $i_0^{'} = 0$ by construction of t^*.

Therefore, we have expressed x as a convex combination of $k-1$ elements of S. If $k-1 = n+1$ we are finished. Otherwise, we can repeat the above construction until we reach $n+1$ elements (which may be affinely independent).

□

Remark In the proof we geometrically start from $\alpha = (\alpha_1, \ldots, \alpha_k) \in \Delta_k$. We compute a direction $\delta = (\delta_1, \ldots, \delta_k)$, which is in the linear subspace parallel to aff Δ_k so that $\alpha - t\delta \in \operatorname{aff} \Delta_k$ and x is kept invariant. t^* is the maximal step size such that $\alpha - t\delta \in \Delta_k \Rightarrow \alpha - t^*\delta$ is on the boundary of Δ_k, i.e. it is in Δ_{k-1}. See Figure 1.10 for an illustration.

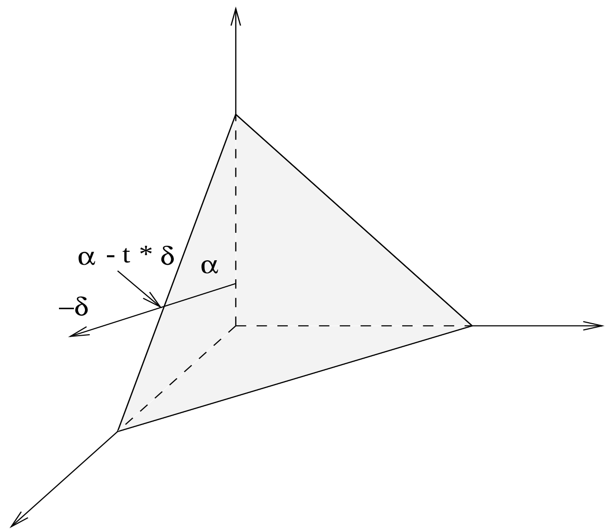


Figure 1.10: Illustration for proof of Theorem 1.5.

Theorem 1.6 (Fenchel, Bunt) *If $S \subset \mathbb{R}^n$ has no more than n connected components then any $x \in \operatorname{co} S$ can be expressed as a convex combination of n elements of S.*

Proof: Let $x \in \operatorname{co} S$. Then we know by Theorem 1.5 that x can be written as

$$x = \sum_{i=1}^{n+1} \alpha_i x_i, \quad x_i \in S, i = 1, \ldots, n+1, \alpha = (\alpha_1, \ldots, \alpha_{n+1}) \in \Delta_{n+1}. \quad (1.1)$$

We assume that all $\alpha_i > 0$, $i = 1, \ldots, n+1$, since otherwise we have a sum with at most n summands and the result is shown.*wlog* $x = 0$, otherwise do a translation. Let $x_i^{'}$ be the reflexion of x_i in x, i.e.

$$x_i^{'} := (2x) - x_i, \ i = 1, \ldots, n+1. \quad (1.2)$$

Refer to Figure 1.11 for an illustration. Moreover, let C_j be the cone defined by the n points $x_1^{'}, x_2^{'}, \ldots, x_{j-1}^{'}, x_{j+1}^{'}, \ldots, x_{n+1}^{'}$.

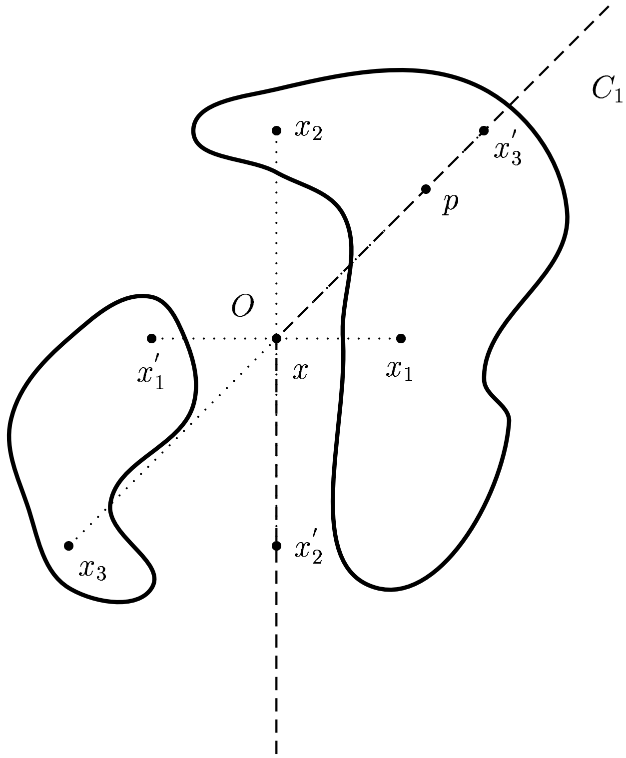


Figure 1.11: Illustration for proof of Theorem 1.6.

Consider the cone C_1. A point $z \in C_1$ can be written as

$$z = \lambda(\mu_2 x_2' + \ldots + \mu_{n+1} x_{n+1}') \tag{1.3}$$

with $\mu_i \geq 0, \mu_2 + \ldots + \mu_{n+1} = 1, \lambda \geq 0$. The same construction works for general $C_j,\ j = 1, \ldots, n+1$. By (1.1) and (1.2) $(x = 0)$ we can write

$$\begin{aligned} x_j' &= -x_j \\ &= \frac{1}{\alpha_j}\left(\overbrace{\alpha_1 x_1}^{>0} + \ldots + \overbrace{\alpha_{j-1}}^{>0} x_{j-1} + \overbrace{\alpha_{j+1}}^{>0} x_{j+1} + \ldots + \overbrace{\alpha_{n+1}}^{>0} x_{n+1}\right) \\ &= \frac{1}{\alpha_j}\left(\overbrace{(-\alpha_1)}^{<0} {x_1}' + \ldots + \overbrace{(-\alpha_{j-1})}^{<0} x_{j-1}' + \overbrace{(-\alpha_{j+1})}^{<0} x_{j+1}' + \ldots + \overbrace{(-\alpha_{n+1})}^{<0} x_{n+1}'\right) \end{aligned}$$

$\Rightarrow x_j \in \text{int } C_j$. Since there are by construction $C_1, \ldots, C_{n+1}$ and S has at most n components, there is a $p \in S,\ p \neq x_j',\ j = 1, \ldots, n+1$, with $p \in \text{bd } C_j$ (not each int C_j can contain a connected component). *wlog* assume $j = 1$

$$\Rightarrow p \quad = \quad \lambda(\mu_2 x_2' + \ldots + \mu_{n+1} x_{n+1}')$$

with at least one $\mu_i = 0$ ($p \in \text{bd } C_j$). Assume $\mu_2 = 0$

$$\begin{aligned} \Rightarrow p &= \mu_3^{'} x_3^{'} + \ldots + \mu_{n+1}^{'} x_{n+1}^{'} \text{ with } \mu_i^{'} = \lambda \mu_i \\ \Rightarrow 0 &= x = \frac{p + \mu_3^{'} x_3 + \ldots + \mu_{n+1}^{'} x_{n+1}}{1 + \mu_3^{'} + \ldots + \mu_{n+1}^{'}} \end{aligned}$$

$\Rightarrow x$ can be written as a convex combination of n elements from S.

□

Theorem 1.7 (Helly's Theorem, 1923) *Let* $C_1, \ldots, C_k \subset \mathbb{R}^n$, $k \geq n+1$, C_i *convex,* $i = 1, \ldots, k$. *Then*

$$\bigcap_{i=1}^{k} C_i \neq \emptyset \Longleftrightarrow \forall \{i_1, \ldots, i_{n+1}\} \subset \{1, \ldots, k\} \quad C_{i_1} \cap C_{i_2} \cap \ldots \cap C_{i_{n+1}} \neq \emptyset.$$

Proof:
"$\Longrightarrow$": Obvious.
"$\Longleftarrow$": By induction.

- $k = n + 1$ Obvious.
- $k > n+1$, $k \to k+1$: By the induction hypothesis, for each i such that $1 \leq i \leq k+1$ there exists $x_i \in \mathbb{R}^n$ such that

$$x_i \in \bigcap_{j=1, j \neq i}^{k+1} C_j.$$

Since $k \geq n+1$ the points $x_1, \ldots, x_{k+1}$ are affinely dependent $\Rightarrow$ $\exists\, \delta_1, \ldots, \delta_{k+1} (\delta_1, \ldots, \delta_{k+1}) \neq 0$, such that

$$\sum_{i=1}^{k+1} \delta_i x_i = 0 \textit{ and } \sum_{i=1}^{k+1} \delta_i = 0.$$

wlog we assume that $\delta_1 \geq 0, \ldots, \delta_s \geq 0$ and $\delta_{s+1} < 0, \ldots, \delta_{k+1} < 0$. Let

$$y := \frac{\sum_{i=1}^{s} \delta_i x_i}{\sum_{i=1}^{s} \delta_i}.$$

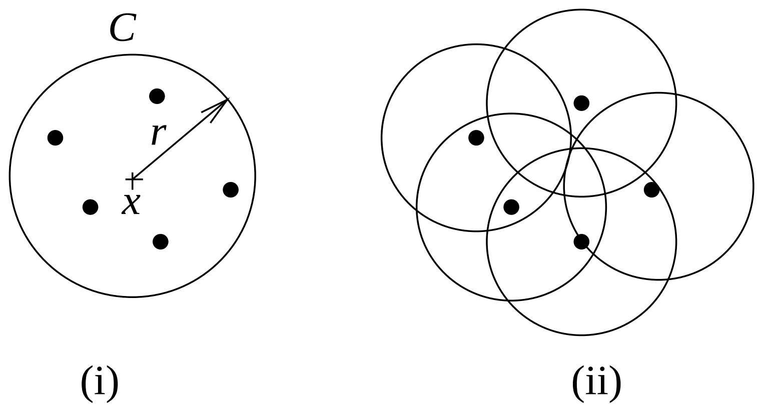


Figure 1.12: The problem for Example 1.8: (i) the covering problem, and (ii) the dual problem.

Then we have also

$$y := \frac{\sum_{i=s+1}^{k+1} (-\delta_i) x_i}{\sum_{i=s+1}^{k+1} (-\delta_i)}$$

$$\Rightarrow y \in \bigcap_{i=1}^{s} C_i \text{ and } y \in \bigcap_{i=s+1}^{k+1} C_i \Rightarrow y \in \bigcap_{i=1}^{k+1} C_i.$$

□

Remark The theorem has an analogue for an arbitrary (not necessarily finite) collection of compact convex sets in $\mathbb{R}^n$.

Example 1.8 *Given $x_1, \dots, x_k \in \mathbb{R}^2$, $k > 2$. Find the smallest covering circle for $x_1, \dots, x_k$ $C(x,r)$ with center x and radius r. (For all $x_i : \|x_i - x\| \leq r$)*

Dual Interpretation: Find the smallest radius r such that $\bigcap_{i=1}^{k} C(x_i, r) \neq \emptyset$.

Use Helly's Theorem (Theorem 1.7) for arbitrary r.

$$\bigcap_{i=1}^{k} C(x_i, r) \neq \emptyset \Longleftrightarrow C(x_{i_1}, r) \cap C(x_{i_2}, r) \cap C(x_{i_3}, r) \neq \emptyset, \{i_1, i_2, i_3\} \subset \{1, 2, 3\}.$$

But this is easy:

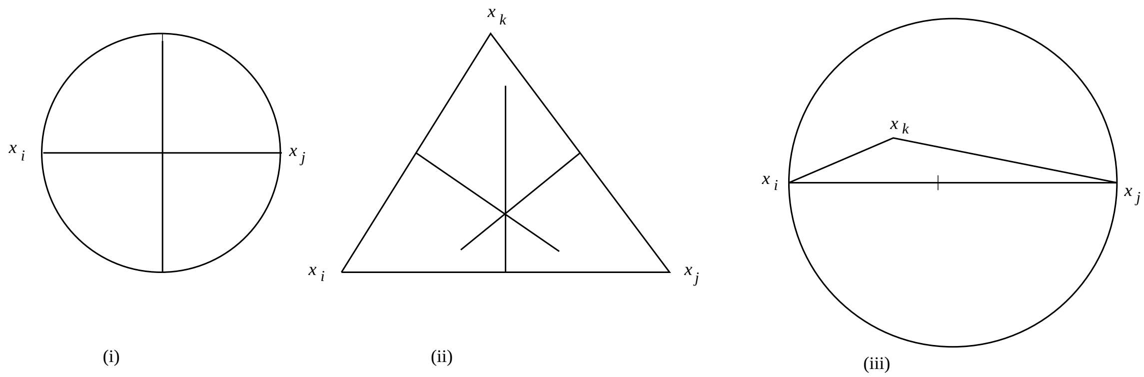


Figure 1.13: Illustration for Example in $\mathbb{R}^2$: (i) two points; (ii) three points with all angles less than 120 degrees; (iii) three points with an angle greater than or equal to 120 degrees.

- *For 2 points: The minimum covering circle for x_i, x_j is the circle with x = midpoint of $[x_i, x_j]$ and $r = \frac{1}{2}\|x_i - x_j\|$.*
- *For 3 points: Either the circumscribing circle of the triangle x_i, x_j, x_k, or if one angle is obtuse, the minimum covering circle of the longest side of the triangle x_i, x_j, x_k.*

  We generate $O(k^3)$ many candidates and select the one with smallest r.

1.4 Closed Convex Sets and Hulls

Definition 12 *The* **closed convex hull** *of $S \subset \mathbb{R}^n$, $S \neq \emptyset$ (denoted $\overline{\text{co}}\ S$) is defined as the intersection of all closed convex sets containing S.*

Proposition 1.8 *Let $S \subset \mathbb{R}^n$, $S \neq \emptyset$. $\overline{\text{co}}\ S = \text{cl}\,(\text{co}\ S)$. [cl: closure]*

Proof:

- "$\text{cl}\,(\text{co}\ S) \supset \overline{\text{co}}\ S$": Clear, since $\text{cl}\,(\text{co}\ S)$ is convex, closed and contains S.
- "$\text{cl}\,(\text{co}\ S) \subset \overline{\text{co}}\ S$": Take $C \supset S$, with C closed and convex $\Rightarrow C \supset \text{cl}\,(\text{co}\ S)$. Since this is true for any closed convex set $C \supset S \Rightarrow \overline{\text{co}}\ S = \bigcap \{C : C \supset S,\ C \text{ closed, convex}\} \supset \text{cl}\,(\text{co}\ S)$.

$\square$

Remark In general cl (co S) $\neq$ co (cl S). For example:
$S = \{(0,0)\} \cup \{(\xi, 1) : \xi \geq 0\}$ is a closed set but co S is not closed. The halfline $\{(x,y) : x > 0, y = 0\}$ is missing. See Figure 1.14.

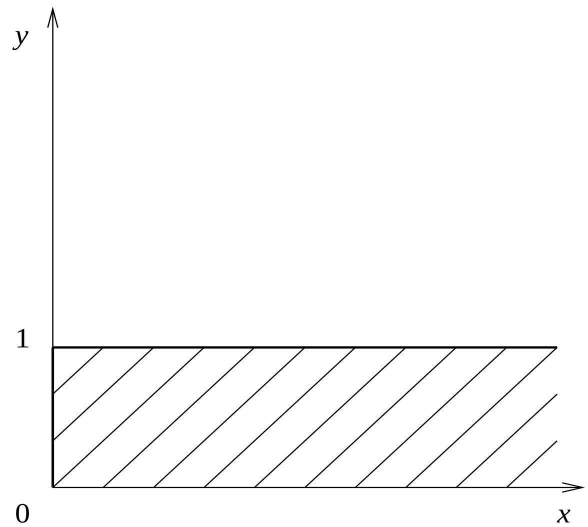


Figure 1.14: Illustration for cl (co S) $\neq$ co (cl S): notice the missing $\{(x,y) : x > 0, y = 0\}$.

Theorem 1.9 *Let $S \subset \mathbb{R}^n$. If S is bounded (respectively compact), then* co S *is bounded (respectively compact).*

Proof: Let $x = \sum_{i=1}^{n+1} \alpha_i x_i \in$ co S. If S is bounded, say by M, we can write

$$\|x\| \leq \sum_{i=1}^{n+1} \alpha_i \|x_i\| \leq M \sum_{i=1}^{n+1} \alpha_i = M.$$

Now take a sequence $\{x^k\} \subset$ co S. For each k we can choose $x_1^k \ldots, x_{n+1}^k \in S$ and $\alpha^k = (\alpha_i^k, \ldots, \alpha_i^k) \in \Delta_{n+1}$ such that $x^k = \sum_{i=1}^{n+1} \alpha_i^k x_i^k$. Note that Δ_{n+1} is compact. If S is compact, we can extract a converging subsequence $\{\alpha^k\}$, $\{x_i^k\}$ as many times as necessary (not more than $n+2$ times). We end up with an index set $\mathcal{K}$ such that when $k \to \infty$, $\{x_i^k\}_{k \in \mathcal{K}} \to x_i \in S$ and $\{\alpha^k\}_{k \in \mathcal{K}} \to \alpha \in \Delta_{n+1}$. Therefore $\{x^k\}_{k \in \mathcal{K}}$ converges to a point x which can be expressed as a convex combination of points of $S \Rightarrow x \in$ co $S \Rightarrow$ co S is compact.

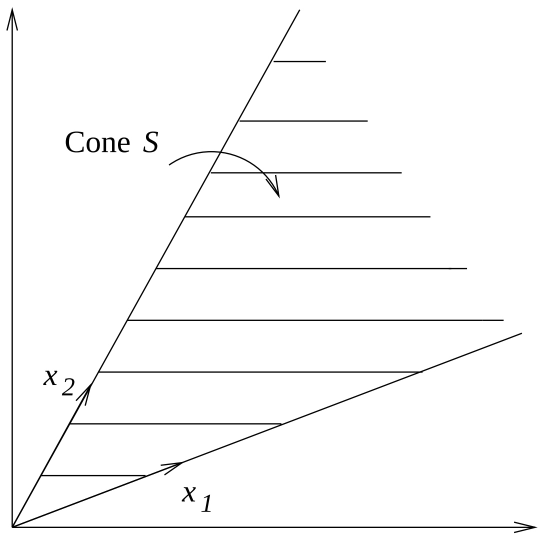


Figure 1.15: Example of Conical Hull of $S = \{x_1, x_2\}$.

□

Remark $S \subset \mathbb{R}^n$ bounded $\Longrightarrow \overline{\text{co}}\ S = \text{cl co } S = \text{co cl } S$.

Definition 13 *Let* $x_1, \ldots, x_k \in \mathbb{R}^n$*,* $\alpha_i, \ldots, \alpha_k \in \mathbb{R}_{+0}$*. An element of the form* $\sum_{i=1}^k \alpha_i x_i$ *is called a* **conical combination** *of* $x_1, \ldots, x_k$*. The set of all conical combinations from a given set* $S \subset \mathbb{R}^n$*,* $S \neq \emptyset$*, is called the* **conical hull** *of* S *(denoted:* cone S*).*

Example 1.9 *(Refer to Figure 1.15).* $S = \{x_1, x_2\}$.

Remark cone $S = \mathbb{R}_{+0}$ (co S) = co ($\mathbb{R}_{+0}\ S$).

Definition 14 *The* **closed conical hull** *of* $S \subset \mathbb{R}^n$*,* $S \neq \emptyset$*, is*

$$\overline{\text{cone}}\ S := \text{cl (cone } S) = \text{cl } \left\{ \sum_{i=1}^{k} \alpha_i x_i : \alpha_i \geq 0, x_i \in S, i = 1, \ldots, k \right\}.$$

Proposition 1.10 $S \subset \mathbb{R}^n$*,* $S \neq \emptyset$*,* S *compact,* $0 \notin \text{co } S$*. Then*

$$\overline{\text{cone}}\ S = \mathbb{R}_{+0}\ (\text{co } S) = \text{cone } S.$$

Proof: $C := \text{co } S$ is compact (Proposition 1.8) and $0 \notin C$. We show that $\mathbb{R}_{+0}\ C$ is closed. Let $\{t_k x_k\} \subset \mathbb{R}_{+0}\ C$ converge to y. *wlog* $x_k \to x \in C$ (otherwise take a subsequence) and $x \neq 0$.
We have $t_k \frac{x_k}{||x_k||} \to \frac{y}{||x||} \Rightarrow t_k \to \frac{||y||}{||x||} =: t > 0$. Then $t_k x_k \to tx = y \in \mathbb{R}_{+0}\ C$.

□

1.5 Exercises

1. Let A be a symmetric positive definite $n \times n$ matrix ($A \in \mathrm{Mat(n, IR)}$) such that $x^T Ax > 0 \quad \forall\, x \in \mathrm{IR^n} \quad x \neq 0$). Show that

   a) $\langle x, y\rangle_A = x^T Ay$ is a scalar product on $\mathrm{IR^n}$.

   b) $||x||_A = \sqrt{\langle x, x\rangle_A}$ is a norm on $\mathrm{IR^n}$.

2. a) Show that the following sets S_i are convex

$$\begin{aligned} S_1 &= \{x \in \mathrm{IR}^4 : 2x_1 + 3x_4 \leq 5,\ -2x_1 + 5x_2 - x_3 + 4x_4 \leq 3 \\ &\qquad\qquad x_1 + x_2 + x_3 + x_4 = 2,\ x_i \geq 0\} \\ S_2 &= \left\{x \in \mathrm{IR^n} : \sum a_i x_i^2 \leq c;\ a_i > 0,\ c > 0\right\} \end{aligned}$$

   b) Is S convex, star-shaped (with respect to a point), none of the two, or both?

$$S = \{(x, y) \in \mathrm{IR}^2 : y \geq x^4 - 3x^2,\ y \leq 5\}$$

3. Show that a convex cone K is a subspace if and only if $K = -K$.

4. Let $K \subseteq \mathrm{IR^n}$ be a cone. K is called pointed if $x \in K \Rightarrow -x \notin K$. K is called acute if there exists an open halfspace $H^<_{s,0}$ such that cl $(K) \subset H^<_{s,0} \cup \{0\}$, where cl (K) is the closure of K.

   a) Is a convex cone acute or pointed?

   b) Give a graphical example (in IR^2) of a pointed cone which is not acute.

5. Cones are useful to define orderings on IR^n. Given a cone K we say that K is the set of nonnegative elements and define

$$x \leq_K y \Leftrightarrow y - x \in K$$

   a) Show that if $0 \in K$ and K is convex and pointed then $\leq_K$ is a partial order (i.e. a reflexive, transitive, antisymmetric relation). What about the converse?

b) What are the cones related to the componentwise order

$$x \leq y \Leftrightarrow x_i \leq y_i \quad i = 1, \ldots, n$$

and the lexicographic order
$x \leq_{lex} y \Leftrightarrow x = y$ or $\exists\, 1 \leq k < n \quad x_i = y_i, \quad i = 1, \ldots, k;$
$x_{k+1} < y_{k+1}$.

6. Let $S = C_1 \cup \ldots \cup C_m$ where $C_i \subset \mathbb{R}^n$ are convex, $i = 1, \ldots, m$. Show that

$$\text{co } S = \left\{ \sum_{i=1}^{m} \alpha_i x_i : \alpha_i \geq 0, \quad \sum \alpha_i = 1, x_i \in C_i \right\}.$$

7. Let S_1 and S_2 be open subsets of $\mathbb{R}^n$. Is $S_1 + S_2$ open? Let

$$\begin{aligned} S_1 &= \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 2\} \\ S_2 &= \Big\{(x, y) \in \mathbb{R}^2 : \max\{|x - 2|, |y - 2|\} \leq 1\Big\} \setminus \{(2, 2)\}. \end{aligned}$$

Graphically determine $S_1 + S_2$.

8. Let $S \subset \mathbb{R}^n$ contain at least $n + 2$ points.

Prove: There exist $S_1, S_2 \subset \mathbb{R}^n$ s.t.

$$S_1 \cap S_2 = \emptyset \quad S_1 \cup S_2 = S$$

and

$$\text{co } S_1 \cap \text{co } S_2 \neq \emptyset.$$

(Hint: Helly's Theorem and its proof)

9. a) Prove: If $S \subset \mathbb{R}^n$ is open then co S is open.

   b) Prove: $\dim C = n \Leftrightarrow \text{int } C \neq \emptyset$, where C is a convex set.

10. a) Let $S = \{(x, y) : (x - 2)^2 + (y - 2)^2 \leq 1\} \cup \{(0, 0)\}$. Draw the convex, conical and affine hull of S.

b) Give an example of a compact set S such that $\overline{\text{cone}}\ S \neq \text{cone}\ S$.

c) Give a condition for S such that cone $S = \mathbb{R}^n$ holds and prove it.

11. $S \subset \mathbb{R}^n$ is called radial in $x^r \in S$ if $\forall\, y \in \mathbb{R}^n\ \exists\, t > 0$ such that $\forall\, \alpha \in (-t, t) \quad \alpha y \in (S - x)$. Then x^r is called a radial point.

    Prove: Let $C \in \mathbb{R}^n$ be a convex set such that int $C \neq \emptyset$.

    Then every radial point is an interior point.

Chapter 2

CONVEX SETS ATTACHED TO A CONVEX SET

2.1 The Relative Interior

Definition 15 *The* **relative interior ri** $\boldsymbol{C}$ *of a convex set* $C \subset \mathbb{R}^n$ *is the interior of* C *for the topology relative to the affine hull of* C*. In other words,*

$$x \in \text{ri } C :\Longleftrightarrow x \in \text{aff } C \text{ and } \exists\, \delta > 0 : (\text{aff } C) \cap B(x, \delta) \subset C.$$

The **dimension** of a convex set C ($\dim C$) is the dimension of its affine hull (the dimension of the linear subspace parallel to aff C).
Remark The cluster points of a set C are in aff C (aff C is closed) $\Rightarrow$ relative cl $C =$ cl C. But we need the **relative boundary**

$$\text{rbd } C := \text{cl } C \setminus \text{ri } C.$$

Theorem 2.1 *Let* $C \subset \mathbb{R}^n$*. If* $C \neq \emptyset$ *then* ri $C \neq \emptyset$ *and* $\dim(\text{ri } C) = \dim C$*.*

C	aff C	$\dim C$	ri C
$\{x\}$	$\{x\}$	0	$\{x\}$
$[x, x'],\ x \neq x'$	affine line generated by x, x'	1	(x, x')
Δ_n	affine manifold of equation $e^T \alpha = 1$	$n - 1$	$\{\alpha \in \Delta_n : \alpha_i > 0\}$
$B(x_0, \delta)$	$\mathbb{R}^n$	n	int $B(x_0, \delta)$

Table 2.1: Affine hull, dimension and relative interior of some sets.

Proof: Let $k := 1 + \dim C \Rightarrow \text{aff } C$ has dimension $k-1$. Therefore, there exist $x_1, \dots, x_k \in C$ which are affinely independent.
Further, let $\Delta := \text{co}\,\{x_1, \dots, x_k\}$ the simplex generated by $x_1, \dots, x_k$.
We have aff $\Delta =$ aff C because $\Delta \subset C$ and $\dim \Delta = k-1$. We are finished if we can show that ri $\Delta \neq \emptyset$. Take $\bar{x} = \frac{1}{k}\sum_{i=1}^{k} x_i$ and describe the affine hull of the simplex, aff Δ, by points of the form

$$\bar{x} + y = \bar{x} + \sum_{i=1}^{k} \alpha_i(y) x_i = \sum_{i=1}^{k} [\frac{1}{k} + \alpha_i(y)] x_i,$$

where $\alpha(y) = (\alpha_1(y), \dots, \alpha_k(y)) \in \mathbb{R}^k$ solves

$$\sum_{i=1}^{k} \alpha_i x_i = y, \ \sum_{i=1}^{k} \alpha_i = 0.$$

This system of equations has a unique solution ($x_1, \dots, x_k$ are affinely independent) and $y \mapsto \alpha(y)$ is linear and continuous. Therefore, we can find $\delta > 0$ such that

$$\|y\| \leq \delta \Rightarrow |\alpha_i(y)| \leq \frac{1}{k} \ \forall\, i = 1, \dots, k.$$

Hence

$$\bar{x} + y \in \Delta \Rightarrow \bar{x} \in \text{ri } \Delta \subset \text{ri } C.$$

In particular, $\dim \text{ri } C = \dim \Delta = \dim C$.

□

Remark ri $\Delta = \{\sum_{i=1}^{k} \alpha_i x_i : \sum_{i=1}^{k} \alpha_i = 1, \alpha_i > 0\}$.

Lemma 2.2 *Let $x \in \text{cl } C$, $x' \in \text{ri } C$. Then (the half-open segment)*

$$(x, x'] = \{\alpha x + (1-\alpha)x' : 0 \leq \alpha < 1\} \subset \text{ri } C.$$

Proof: Take $x'' = \alpha x + (1-\alpha)x', 1 > \alpha \geq 0$. We assume *wlog* that aff $C = \mathbb{R}^n$, to avoid writing "$\bigcap$aff C" and use int instead of ri. Since $x \in \text{cl } C$, $\forall\, \epsilon > 0$, we have $x \in C + B(0, \epsilon)$ and we can write

$$\begin{aligned} B(x'', \epsilon) &= \alpha x + (1-\alpha)x' + B(0, \epsilon) \\ &\subset \alpha C + (1-\alpha)x' + (1+\alpha)B(0, \epsilon) \\ &= \alpha C + (1-\alpha)(x' + B(0, \frac{1+\alpha}{1-\alpha}\epsilon)). \end{aligned}$$

Since $x' \in \text{int } C$, we can choose ϵ so small that

$$x' + B(0, \frac{1+\alpha}{1-\alpha}) \subset C \Rightarrow B(x'', \epsilon) \subset \alpha C + (1-\alpha)C = C.$$

□

Remark Consequence is that a half-line starting from $x' \in \text{ri } C$ can not cut the relative boundary of C in more than one point, as depicted in Figure 2.1. A line meeting ri C can not cut cl C at more than 2 points. The relative boundary is a fairly regular object.

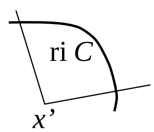


Figure 2.1: A half-line from $x' \in \text{ri } C$ can not cut rbd C in more than one point.

Proposition 2.3 *The three convex sets* ri $C, C,$ cl C *have the same affine hull (and hence the same dimension), the same relative interior, and the same closure, (and the same relative boundary).*

Proof: Assume $C \neq \emptyset$ (otherwise, the result is obvious).

- Closure. We show that cl (ri C) = cl C.
  "$\subseteq$": cl (ri C) $\subseteq$ cl C.
  "$\supseteq$" Let $x \in$ cl C, $x' \in$ ri C (by Theorem 2.1 ri C is not empty). Lemma 2.2 $\Rightarrow (x, x'] \subset$ ri $C \Rightarrow x$ is the limit of points in ri $C \Rightarrow x \in$ cl (ri C).

- Relative interior. We show that ri (clC) = ri C.
  "$\supseteq$": ri (cl C) $\supseteq$ ri C.
  "$\subseteq$": Let $x \in$ ri (cl C) $\Rightarrow \exists\, \delta$ such that aff (cl C) $\cap B(x, \delta) \subset$ cl $C \Rightarrow$ aff $C \cap B(x, \frac{\delta}{2}) \subset C \Rightarrow x \in$ ri $C \Rightarrow$ ri (cl C) $\subset$ ri C.

- Theorem 2.1 $\Rightarrow \dim C = \dim \text{ri } C = \dim \text{cl } C$. Since aff (ri C) $\subset$ aff $C \subset$ aff (cl C) $\Rightarrow$ The affine hulls are the same.

□

Proposition 2.4 *Let $C_1, C_2 \subset \mathbb{R}^n$, be convex. Let* ri $C_1 \cap$ ri $C_2 \neq \emptyset$. *Then the following hold*

a) ri $(C_1 \cap C_2) = (\text{ri } C_1) \cap (\text{ri } C_2)$.

b) cl $(C_1 \cap C_2) = (\text{cl } C_1) \cap (\text{cl } C_2)$.

Proof:

b) "cl $(C_1 \cap C_2) \subset (\text{cl } C_1) \cap (\text{cl } C_2)$": Follows from monotonicity of closure.
"$(\text{cl } C_1) \cap (\text{cl } C_2) \subset$ cl $(C_1 \cap C_2)$": Given $x \in (\text{cl } C_1) \cap (\text{cl } C_2)$. Take $x' \in (\text{ri } C_1) \cap (\text{ri } C_2)$ Lemma 2.2 $\Rightarrow (x, x'] \subset (\text{ri } C_1) \cap (\text{ri } C_2)$ $\Rightarrow x \in \text{cl } ((\text{ri } C_1) \cap (\text{ri } C_2)) \subset \text{cl } (C_1 \cap C_2)$.

a) "ri $(C_1 \cap C_2) \subset (\text{ri } C_1) \cap (\text{ri } C_2)$": Proposition 2.3 $\Rightarrow$

$$\text{ri } (C_1 \cap C_2) = \text{ri } ((\text{ri } C_1) \cap (\text{ri } C_2)) \subset (\text{ri } C_1) \cap (\text{ri } C_2).$$

"$(\text{ri } C_1) \cap (\text{ri } C_2) \subset$ ri $(C_1 \cap C_2)$": Let $y \in (\text{ri } C_1) \cap (\text{ri } C_2)$. Take $x' \in C_1$ (resp. C_2) $\Rightarrow [x', y] \subset$ aff C_1 (resp. aff C_2). By definition of relative interior, $[x', y]$ can be stretched beyond y and stays in C_1 (resp. C_2). See Figure 2.2.

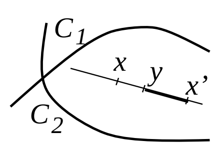


Figure 2.2: Illustration for proof of Proposition 2.4.

Now take $x' \in \text{ri } (C_1 \cap C_2) \subset C_1$, $x' \neq x$. (If no such point x' exists, we are done $(\text{ri } C_1) \cap (\text{ri } C_2) = y = \text{ri } (C_1 \cap C_2)$). The above described stretching procedure yields an $x \in C_1 \cap C_2$ such that $y \in (x, x')$: $y = \alpha x + (1 - \alpha)x'$, $\alpha \in (0, 1)$. Apply Lemma 2.2 $\Rightarrow y \in \text{ri } (C_1 \cap C_2)$.

□

Proposition 2.5 *Let $C_i \subset \mathbb{R}^n$ be convex, $i = 1, \dots, k$. Then*

$$\text{ri}\,(C_1 \times \dots \times C_k) = (\text{ri}\ C_1) \times \dots \times (\text{ri}\ C_k).$$

Proof: Apply the definition of relative interior, and note that

$$\text{aff}\,(C_1 \times \dots \times C_k) = (\text{aff}\ C_1) \times \dots \times (\text{aff}\ C_k).$$

□

Proposition 2.6 *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be an affine mapping and $C \subset \mathbb{R}^n$ be convex. Then*

- $\text{ri}\,(A(C)) = A(\text{ri}\,(C))$.
- *If $D \subset R^m$ is convex satisfying $A^{-1}(\text{ri}\ D) \neq \emptyset$, then*

$$\text{ri}\,(A^{-1}(D)) = A^{-1}(\text{ri}\ D).$$

Proof: Use monotonocity of closure, Lemma 2.2 and the stretching mechanism of Proposition 2.4 (the rest is left as an exercise).

□

Remark From Propositions 2.6, 2.5 $\Rightarrow$ $\text{ri}\ (\alpha_1 C_1 + \alpha_2 C_2) = \alpha_1 \text{ri}\ C_1 + \alpha_2 \text{ri}\ C_2$, C_1, C_2 convex. For $\alpha_1 = -\alpha_2 = 1$, we get

$$0 \in \text{ri}\,(C_1 - C_2) \Longleftrightarrow (\text{ri}\ C_1) \cap (\text{ri}\ C_2) \neq \emptyset.$$

2.2 The Asymptotic Cone

Let $C \subset \mathbb{R}^n$ be closed and convex. Define:

$$C_\infty(x) := \{d \in \mathbb{R}^n : x + td \in C \text{ for all } t > 0\}$$

or equivalently

$$C_\infty(x) = \bigcap_{t>0} \frac{C - x}{t}.$$

C_∞ is a closed convex cone containing 0.

Proposition 2.7 *The closed convex cone C_∞ does not depend on $x \in C$.*

Proof: Take $x_1, x_2 \in C$, $x_1 \neq x_2$. We prove $C_\infty(x_1) \subset C_\infty(x_2)$ (The other inclusion follows by exchanging x_1 and x_2). Let $d \in C_\infty(x_1), t > 0$. Take $\epsilon \in (0,1)$ and consider

$$\begin{aligned} \bar{x}_\epsilon : &= x_1 + td + (1-\epsilon)(x_2 - x_1) \\ &= \epsilon(x_1 + \frac{t}{\epsilon} d) + (1-\epsilon) x_2 \end{aligned}$$

$\Rightarrow \bar{x}_\epsilon \in C$. On the other hand $x_2 + \lim_{\epsilon \to 0} \bar{x}_\epsilon \in \text{cl } C = C$.

□

Definition 16 *The* **asymptotic cone** *(or* **recession cone***) of a closed convex set $C \subset \mathbb{R}^n$ is the closed convex cone C_∞ $(:= C_\infty(x)$ for some $x \in C)$.*

Example 2.1 *Refer to Figure 2.3 for examples of sets and their asymptotic cone.*

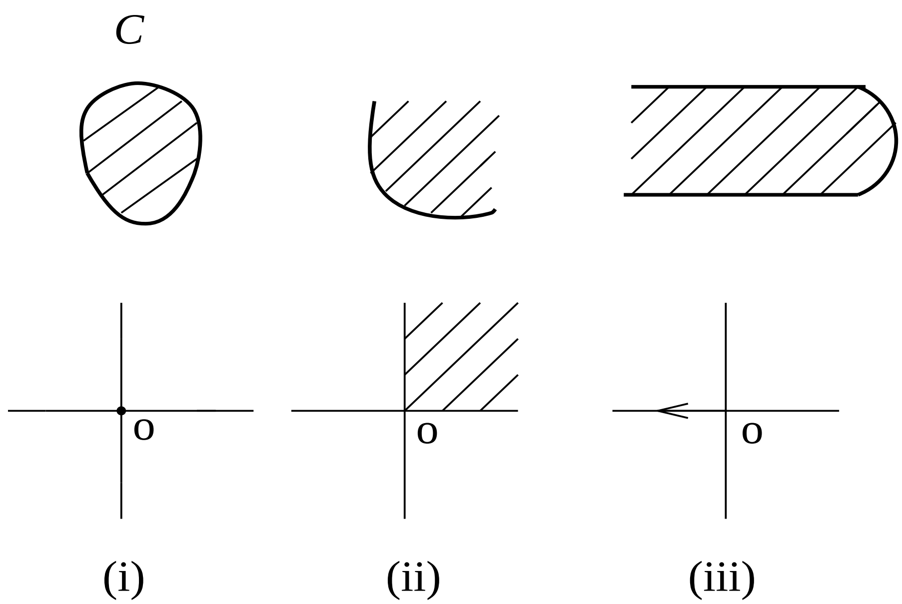


Figure 2.3: Examples of sets and their asymptotic cone.

Proposition 2.8 *A closed convex set C is compact $\Longleftrightarrow C_\infty = \{0\}$.*

Proof: "$\Longrightarrow$": C is bounded $\Rightarrow$ there is no $d \neq 0$ with $x + td \in C$ for all $t > 0$.

"$\Longleftarrow$": Assume C is not bounded. Let $\{x_k\} \subset C$ such that $\|x_k\| \to +\infty$ (assume $x_k \neq 0$). The sequence $\{d_k := \frac{x_k}{\|x_k\|}\}$ is bounded and we can extract a convergent subsequence $\{d_k\}$ with $\lim\limits_{k \in K} d_k = d$ ($K \subset \mathbb{N}$, $\|d\| = 1$). Given $x \in C$, take k so large that $\|x_k\| \geq t$. Then

$$x + td = \lim_{k \in K} \left((1 - \frac{t}{\|x_k\|})x + \frac{t}{\|x_k\|} x_k \right) \in C$$

since C is closed $\Rightarrow d \in C_\infty$, $d \neq 0$.

□

Other Useful Properties of C_∞

- If $\{C_j\}_{j \in \mathcal{J}}$ is a family of closed convex sets with $\bigcap\limits_{j \in \mathcal{J}} C_j \neq \emptyset$ then

$$(\bigcap_{j \in \mathcal{J}} C_j)_\infty = \bigcap_{j \in \mathcal{J}} (C_j)_\infty.$$

- For $C_j \subset \mathbb{R}^{n_i}$, $j = 1, \ldots, m$, closed convex sets, we have

$$(C_1 \times \ldots \times C_m)_\infty = (C_1)_\infty \times \ldots \times (C_m)_\infty.$$

- $A : \mathbb{R}^n \to \mathbb{R}^m$ linear. If $C \subset \mathbb{R}^n$ is closed, convex and $A(C)$ is closed then

$$A(C_\infty) \subset [A(C)]_\infty.$$

- If $D \subset \mathbb{R}^m$ is closed and convex with $A^{-1}(D) \neq \emptyset$ then

$$[A^{-1}(D)]_\infty = A^{-1}(D_\infty).$$

2.3 Extreme Points and Exposed Faces

In this section, $C \subset \mathbb{R}^n$, C convex and $C \neq \emptyset$.

Definition 17 *Let $x \in C \subset \mathbb{R}^n$, $C \neq \emptyset$, C convex. x is an* **extreme point** *of C if there are no $x_1, x_2 \in C$, $x_1 \neq x_2$ such that $x = \frac{1}{2}(x_1 + x_2)$. Equivalently:*

- *if* $(x = \alpha x_1 + (1 - \alpha) x_2, x_i \in C, \alpha \in (0, 1) \Rightarrow x = x_1 = x_2)$.

- *if* ($x = \sum_{i=1}^{k} \alpha_i x_i$ *convex combination,* $\alpha_i \in (0,1)$
  $\Rightarrow x = x_1 = x_2 = \ldots = x_k$).
- *if* $C - \{x\}$ *is still convex.*

The set of extreme points of C is denoted by **ext** $\boldsymbol{C}$.

Example 2.2

- $C = B(0,1) = \{x : \|x\| \leq 1\}$. *Find* x *for which*
  $x = \frac{1}{2}(x_1 + x_2)$ *implies* $x = x_1 = x_2$.
  We have $\|x_1\| \leq 1$, $\|x_2\| \leq 1$, $\|x\| = \frac{1}{2}\|x_1 + x_2\|$.

  Consider the following relation:

  $$\frac{1}{2}\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2 - \frac{1}{2}\|x_2 - x_1\|^2.$$

  Take x *with* $\|x\| = 1 \Rightarrow \frac{1}{2}\|x_1 + x_2\|^2 = \frac{1}{2}2^2 = 2$.
  But $\|x_1\|^2 + \|x_2\|^2 - \frac{1}{2}\|x_2 - x_1\|^2 \leq 2$

  $$\begin{aligned} &\Rightarrow \|x_2 - x_1\| = 0 \Rightarrow x_1 = x_2 \text{ and } \|x_2\| = \|x_1\| = 1 \\ &\Rightarrow \frac{1}{2}(x_1 + x_2) = \frac{1}{2}(2x_1) \Rightarrow x = x_1 = x_2 \\ &\Rightarrow \text{ext } B(0,1) \supset \{x : \|x\| = 1\}. \end{aligned}$$

  If $\|x\| < 1$, x *is no extreme point (equation can be fulfilled with* $x_1 \neq x_2$*).*

- $C = \{x = (\xi, \eta) \in \mathbb{R}^2 : |\xi| + |\eta| \leq 1\}$.
  ext $C = \{(1,0), (0,-1), (-1,0), (0,1)\}$.
  $(1,0) \in$ ext C : $(1,0) = \frac{1}{2}(x_1 + x_2) \Rightarrow x_1 = (\xi^1, 0), x_2 = (\xi^2, 0)$ *with* $\xi^1 + \xi^2 = 2 \Rightarrow \xi^1 = \xi^2 = 1 \Rightarrow x = x_1 = x_2$.
- *If* C *is a convex cone then no* $x \neq 0$ *can be an extreme point.*
- *An affine manifold or a half-space has no extreme points.*

Proposition 2.9 *If* $C \neq \emptyset$ *is compact, then* ext $C \neq \emptyset$.

Proof: Since C is compact

$$\max_{x\in C} \|x\|^2 = \|\bar{x}\|^2$$

exists, with $\bar{x} \in C$. We show that $\bar{x}$ is an extreme point of C. Suppose there are $x_1, x_2 \in C, x_1 \neq x_2$ with $x = \frac{1}{2}(x_1 + x_2)$. We use again

$$\frac{1}{2}\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2 - \frac{1}{2}\|x_2 - x_1\|^2.$$

$$\begin{aligned} \|\bar{x}\|^2 = \|\frac{1}{2}(x_1 + x_2)\|^2 &< \frac{1}{2}(\|x_1\|^2 + \|x_2\|^2) \\ &\leq \frac{1}{2}(\|\bar{x}\|^2 + \|\bar{x}\|^2) = \|\bar{x}\|^2, \end{aligned}$$

and that is a contradiction.

□

The next result will be proved later in Section 4.2.

Theorem 2.10 (H. Minkowski) *Let $C \subset \mathbb{R}^n$, C compact, convex. Then $C = \text{co}\,(\text{ext}\ C)$.*

Remark

- Theorem 2.10 and Theorem 1.5 (Caratheodory) $\Longrightarrow$ if $\dim C = k$ then $x = \sum_{i=1}^{k} \alpha_i x_i,\ x_i \in \text{ext}\ C$.
- If $C = \text{co}\,\{x_1, \ldots, x_m\}$ then ext $C \subset \{x_1, \ldots, x_m\}$.

Definition 18 *Let $C \subset \mathbb{R}^n$, $C \neq \emptyset$, C convex and $F \subset C$, $F \neq \emptyset$, F convex. F is called a* **face** *if*

$$(x_1, x_2) \in C \times C \text{ and } \exists\, \alpha \in (0,1) : \alpha x_1 + (1-\alpha)x_2 \in F \Longrightarrow [x_1, x_2] \subset F.$$

Example 2.3 $x \in \text{ext}\ C \Longleftrightarrow \{x\}$ *is a face of* C.

Remark In general $x \in C' \subset C$ is an extreme point of $C \Longrightarrow x \in \text{ext}\ C'$.

For faces we have also

$\dim F$	F
0	extreme points of C
1	edges of C
$\vdots$	$\vdots$
$\dim C - 1$	facet of C
$\dim C$	C

Table 2.2: Dimension and extreme points of some sets.

Proposition 2.11 *Let $F \subset C$ be a face of C, C convex. Then*

$$(\text{ext } F) \subset (\text{ext } C).$$

Proof: Let $x \in F \subset C$ and assume that $x \notin \text{ext } C$

$$\Rightarrow \exists\ x_1, x_2 \in C,\ \alpha \in (0,1) : x = \alpha x_1 + (1-\alpha)x_2 \in F.$$

F face $\Rightarrow [x_1, x_2] \subset F \Rightarrow x \notin \text{ext } F$.

□

Definition 19 *An affine hyperplane H_s^r is said to* **support** *the set C when $\langle s, y\rangle \leq r$ for all $y \in C$. (H_s^r is called then a* **supporting hyperplane***). A supporting hyperplane H_s^r is said to* **support C at $x \in C$** *when $x \in H_s^r$ and $\langle s, x\rangle = r$.*

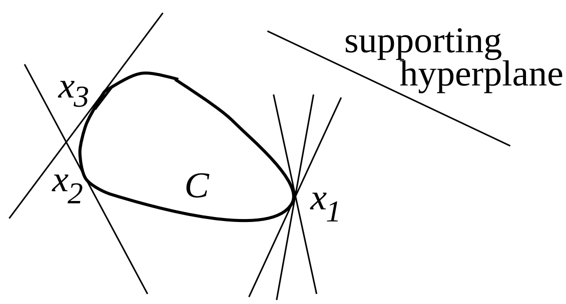


Figure 2.4: Supporting hyperplanes.

Definition 20 *$F \subset C$ is an* **exposed face** *of C, (C convex) if there is a supporting hyperplane H_s^r of C such that $F = C \cap H_s^r$. An* **exposed point (or vertex)** *is a 0-dim exposed face, i.e., $\{x\} = C \cap H_s^r$.*

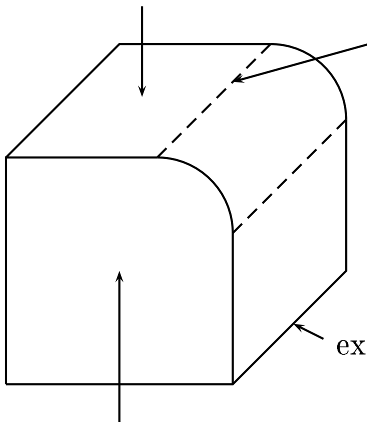


Figure 2.5: Example of extreme and/or exposed points.

Proposition 2.12 *An exposed face is a face.*

Proof: Let $F \subset C$ be an exposed face of C defined by H_s^r, i.e. $F = C \cap H_s^r$. For $x_1, x_2 \in C$ we have

$$\langle s, x_i \rangle \leq r, \ i = 1, 2. \tag{2.1}$$

Take $\alpha \in (0,1)$ such that $\alpha x_1 + (1-\alpha)x_2 \in F \subset H_s^r \Rightarrow \langle s, x_i \rangle = r, \ i = 1, 2$. Suppose we would have strict inequality in Inequality (2.1):

$$\langle s, x_i \rangle < r, \ i = 1, 2. \ \alpha \in (0,1)$$

$$\Rightarrow \langle s, \alpha x_1 + (1-\alpha)x_2 \rangle = \alpha \langle s, x_1 \rangle + (1-\alpha)\langle s, x_2 \rangle < r,$$

a contradiction.

□

Remark

- Not every extreme point is exposed. See Figure 2.5 for an illustration.

- Proposition 2.11 also holds for exposed faces.
- Straszewicz (1935) showed: exp $C \subset$ ext $C \subset$ cl (exp C).
- Facets are always exposed.

Let $F = C \cap H_s^r$ be an exposed face $\Rightarrow \langle s, y\rangle \leq \langle s, x\rangle$ for all $y \in C$, $x \in F$. Therefore we can say: F is an exposed face of when there is a $s \in \mathbb{R}^n$, $s \neq 0$, with $F = \{x \in C : \langle s, x\rangle = \sup_{y \in C}\langle s, y\rangle\} =: F_C(s)$ (the face of C exposed by s). We can write $C= F_C(0)$ (if we allow $s = 0$ we get the whole set C).

Proposition 2.13 *Let $C \subset \mathbb{R}^n$ be convex and compact. For $s \in \mathbb{R}^n$ it holds:*

- $\max_{x \in C}\langle s, x\rangle = \max_{x \in \text{ext } C}\langle s, x\rangle$ *and*
- $\operatorname{argmax}_{x \in C}\langle s, x\rangle = \text{co}\,\{\operatorname{argmax}_{x \in \text{ext } C}\langle s, x\rangle\}$.

Proof: Since C is compact, $\langle s, \cdot\rangle$ attains its maximum on $F_C(s)$. $F_C(s)$ is convex, compact and by Theorem 2.10 (Minkowski) $F_C(s) = \text{co}\,\{\text{ext } F_C(s)\}$. By Proposition 2.11 and the last remarks ext $F_C(s) \subset$ ext C.

□

2.4 Exercises

1. Properties of the relative interior. Let C_1, C_2 be convex such that ri $(C_1) \cap$ ri $(C_2) \neq \emptyset$.

   a) ri $(C_1 + C_2) =$ ri $C_1 +$ ri C_2

   b) ri $(\lambda C_1) = \lambda$ri C_1, $\lambda \in \mathbb{R}$

2. Properties of the asymptotic cone

   a) Let C_j, $j \in J$ be convex sets, $\bigcap_{j \in J} C_j \neq \emptyset$, then $(\bigcap_{j \in J} C_j)_\infty = \bigcap_{j \in J}(C_j)_\infty$.

   b) Let $C_j \subseteq \mathbb{R}^{n_j}$ be convex $j = 1, \ldots, m$ then $(C_1 \times \ldots \times C_m)_\infty = (C_1)_\infty \times \ldots \times (C_m)_\infty$.

c) Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be affine. If $C \subset \mathbb{R}^n$ is closed convex and $A(C)$ is closed then

$$A(C_\infty) \subset [A(C)]_\infty.$$

d) If $D \subset \mathbb{R}^n$ is closed, convex, $A^{-1}(D) \neq \emptyset$ then $[A^{-1}(D)]_\infty = A^{-1}(D_\infty)$

3. $$C := \text{co}\left(\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\-1\end{pmatrix}\right\} \cup \{(x,y,z) : (x-1)^2 + y^2 = 1,\ z = 0\}\right).$$

   a) Determine ext C.

   b) C is closed, what about ext C?

   c) Can you find such an example in $\mathbb{R}^2$?

4. Let $C \subset \mathbb{R}^n$ be convex. A convex set $F \subset C$ is an extremal subset of C if there are no two points $x_1, x_2 \in C \backslash F$ such that $\frac{1}{2}(x_1 + x_2) \in F$. Prove

   a) If F_1 is an extremal subset of F_2 and F_2 is an extremal subset of F_3 than F_1 is an extremal subset of F_3.

   b) Let F_j, $j \in J$ be extremal subsets of C. Then $\bigcap_{j \in J} F_j$ is an extremal subset of C.

5. Prove that $\exp C \subset \text{ext } C$ for a convex set C. Does the converse inclusion hold?

6. A set $K \subset \mathbb{R}^n$, int $K \neq \emptyset$, is called strictly convex if for all $x, y \in K$, $x \neq y$ and all $t \in (0,1)$ $tx + (1-t)y \in$ int K. Prove that the three following statements are equivalent, where K is a closed convex set.

   1) K is strictly convex.

   2) ext $K =$ bd K.

   3) exp $K =$ bd K.

Chapter 3

PROJECTION ONTO CLOSED CONVEX SETS

A projection p_V onto a subspace $V \subset \mathrm{I\!R^n}$ is linear, symmetric, positive semidefinite and idempotent ($p_V \circ p_V = p_V$). In the following, let $C \subset \mathrm{I\!R^n}$ be nonempty, closed and convex. We are solving for fixed $x \in \mathrm{I\!R^n}$ the problem:

$$\inf\{\frac{1}{2}\|y-x\|^2 : y \in C\}, \tag{3.1}$$

in other words: look for $y \in C$ which are closest (in Euclidean distance) to x. To do this, define

$$f_x(y) := \frac{1}{2}\|y-x\|^2$$

and for $c \in C$ the sublevel set

$$S := \{y \in \mathrm{I\!R}^n : f_x(y) \leq f_c(x)\}.$$

Now we can write (3.1) as

$$\inf\{f_x(y) : y \in C \cap S\}.$$

Lemma 3.1 *The infimum in (3.1) is attained and this minimum is unique.*

Proof: f_x is continous and S (and therefore $C \cap S$) is compact $\Rightarrow \inf\{f_x(y) : y \in C \cap S\} = \min\{f_x(y) : y \in C \cap S\}$. Let y_1, y_2 be two solutions of (3.1). Use

$$\frac{1}{2}\|x_1+x_2\|^2 = \|x_1\|^2 + \|x_2\|^2 - \frac{1}{2}\|x_2-x_1\|^2$$

with $x_i = y_i - x$ to obtain for $y_0 = \frac{1}{2}(y_1 + y_2) \in C$

$$f_x(y_0) = \frac{1}{2}(f_x(y_1) + f_x(y_2)) - \frac{1}{8}\|y_2 - y_1\|^2$$

$$\frac{1}{2}\|\frac{1}{2}(y_1 + y_2) - x\| = \frac{1}{8}\|x_1 + x + x_2 + x - 2x\|.$$

Since $f_x(y_0) \geq f_x(y_1),\ f_x(y_0) \geq f_x(y_2) \Rightarrow \|y_2 - y_1\| = 0 \Rightarrow y_2 = y_1 = y_0$.

□

We have defined a projection operator p_C with $x \mapsto p_C$ assigns to x the unique solution of (3.1).

Theorem 3.2 *A point $y_x \in C$ is the projection $p_C(x)$ $(p_C(x) = y_x)$ $\Longleftrightarrow \langle x - y_x, y - y_x\rangle \leq 0$ for all $y \in C$.*

Proof: "$\Longrightarrow$": Call y_x the solution of (3.1) w.r.t. x. Take $y \in C$ so that $y_x + \alpha(y - y_x) \in C$ for any $\alpha \in (0, 1)$

$$\begin{aligned}
\Rightarrow f_x(y_x) &\leq f_x(y_x + \alpha(y - y_x)) \\
&= \frac{1}{2}\|y_x - x + \alpha(y - y_x)\|^2 \\
&= \frac{1}{2}\|y_x - x\|^2 + \alpha\langle y_x - x, y - y_x\rangle + \frac{1}{2}\alpha^2\|y - y_x\|^2 \\
&\Leftrightarrow\ 0 \leq \alpha\langle y_x - x, y - y_x\rangle + \frac{1}{2}\alpha^2\|y - y_x\|^2 \\
&\Rightarrow\ 0 \leq \langle y_x - x, y - y_x\rangle + \frac{1}{2}\alpha\|y - y_x\|^2.
\end{aligned}$$

With $\alpha \to 0 \Rightarrow 0 \leq \langle y_x - x, y - y_x\rangle$.

"$\Longleftarrow$": Take $y_x \in C$ with $\langle x - y_x, y - y_x\rangle \leq 0$ for all $y \in C$. If $y_x = x \Rightarrow y_x$ solves (3.1). If $y_x \neq x$ write for $y \in C$

$$\begin{aligned}
0 &\geq \langle x - y_x, y - y_x\rangle \\
&= \langle x - y_x, y - x + x - y_x\rangle \\
&= \|x - y_x\|^2 + \langle x - y_x, y - x\rangle \geq \|x - y_x\|^2 - \|x - y\|\|x - y_x\|.
\end{aligned}$$

Divide by $\|x - y_x\| \neq 0 \Rightarrow \|x - y\| \geq \|x - y_x\|$ for all $y \in C \Rightarrow y_x$ solves (3.1).

□

Geometrical Interpretations

(Refer to Figure 3.1). $\cos\Theta := \frac{\langle v,w\rangle}{\|v\|\|w\|} \in [-1,1]$, $v, w \in \mathbb{R}^n$.

$$\langle x - y_x, y - y_x\rangle \leq 0 \;\forall\; y \in C \tag{3.2}$$

means that the angle between $(y - y_x)$ and $(x - y_x)$ is obtuse for any $y \in C$. Rewriting Inequality (3.2) yields

$$\langle x - p_C(x), y\rangle \leq \langle x - p_C(x), p_C(x)\rangle \text{ for all } y \in C$$

$\Rightarrow p_C(x)$ lies in the face of C exposed by $x - p_C(x)$ $(F_C(x - p_c(x)))$.

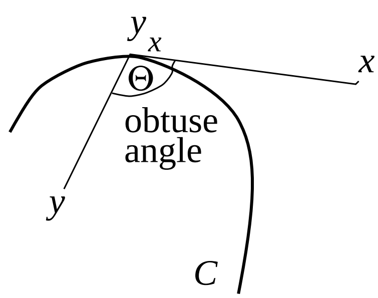


Figure 3.1: Geormetrical interpretation of projection.

Properties of p_C

- $\{x \in \mathbb{R}^n : p_C(x) = x\} = C$.
- $p_C \circ p_C = p_C$.
- p_C is linear $\Longleftrightarrow C$ is a subspace.

Proposition 3.3 *For all $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$, it holds that*

$$\|p_C(x_1) - p_C(x_2)\|^2 \leq \langle p_C(x_1) - p_C(x_2), x_1 - x_2\rangle.$$

Proof: Write $\langle y - y_x, x - y_x\rangle \leq 0$ with $x = x_1, y = p_C(x_2) \in C$:

$$\langle p_C(x_2) - p_C(x_1), x_1 - p_C(x_1)\rangle \leq 0, \tag{3.3}$$

and analogous with $x = x_2$, $y = p_C(x_1)$:

$$\langle p_C(x_1) - p_C(x_2), x_2 - p_C(x_2)\rangle \leq 0. \tag{3.4}$$

Inequality (3.3) + Inequality (3.4)
$\Rightarrow \langle p_C(x_1) - p_C(x_2), x_2 - x_1 + p_C(x_1) - p_C(x_2)\rangle \leq 0$
$\Rightarrow \|p_C(x_1) - p_C(x_2)\|^2 - \langle p_C(x_1) - p_C(x_2), x_1 - x_2\rangle \leq 0.$

□

Consequences

- From C.S.I (Cauchy-Schwarz Inequality) we get

  $\|p_C(x_1) - p_C(x_2)\| \leq \|x_1 - x_2\|, 0 \in C \Rightarrow \|p_C(x)\| \leq \|x\|$ non-expensive.

- $0 \leq \langle p_C(x_1) - p_C(x_2), x_1 - x_2\rangle$ for all $x_1, x_2 \in \mathbb{R}^n$. p_C is "monotone increasing".

Definition 21 *Let K be a convex cone. The* **polar cone** *of K is defined as $K^\circ := \{s \in \mathbb{R}^n : \langle s, x\rangle \leq 0$ for all $x \in K\}$.*

- K° is a closed convex cone (closedness from continuity of the scalar product).

- $K' \subset K \Longrightarrow (K')^\circ \supset K^\circ$.

Example 3.1 *For $x_1, \ldots, x_m \in \mathbb{R}^m$ let*

$$K = \left\{ \sum_{j=1}^{m} \alpha_j x_j : \alpha_j \geq 0, j = 1, \ldots, m \right\}.$$

We compute K°:

$$\begin{aligned} K^\circ &= \{s \in \mathbb{R}^n : \langle s, x\rangle \leq 0 \; \forall \; x \in K\} \\ &= \{s \in \mathbb{R}^n : \langle s, \sum_{j=1}^{m} \alpha_j x_j\rangle \leq 0 \; \forall \; \alpha_j\} \\ &= \{s \in \mathbb{R}^n : \sum_{j=1}^{m} \alpha_j \langle s, x_j\rangle \leq 0 \; \forall \; \alpha_j \geq 0\} \\ &= \{s \in \mathbb{R}^n : \langle s, x_j\rangle \leq 0, j = 1, \ldots, m\}. \end{aligned}$$

$K = \text{cone}\,\{x_1, \ldots, x_m\} \Rightarrow K^\circ = \{s \in \mathbb{R}^n : \langle s, x_j\rangle \leq 0, j = 1, \ldots, m\}$. *In* $\mathbb{R}^2$ *let* $K = \text{cone}\,\{(1, 0), (0, 1)\}$, $s \in \mathbb{R}^2$, *for* K°, $\langle s, (1, 0)\rangle \leq 0$, $\langle s, (0, 1)\rangle \leq$

$0 \Rightarrow K^\circ=\{s = (\eta,\xi) : \eta \leq 0, \xi \leq 0\}$. *The sets K and K° are shown in Figure 3.2.*

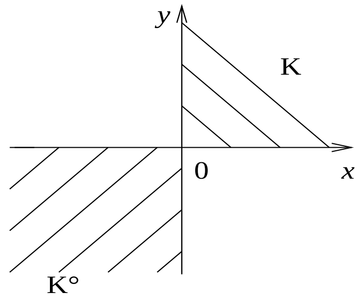


Figure 3.2: Example of a convex cone, K, and its polar cone, K°.

Proposition 3.4 *Let K be a closed convex cone. Then*

$$y_x = p_K(x) \Longleftrightarrow y_x \in K,\ x - y_x \in K^\circ, \langle x - y_x, y_x\rangle \geq 0.$$

Proof: "$\Longrightarrow$": From Theorem 3.2 we know that for $y_x = p_K(x)$ we have

$$\langle x - y_x, y - y_x\rangle \leq 0 \;\forall\; y \in K. \tag{3.5}$$

Taking $y = \alpha y_x, \alpha \geq 0$, we get $(\alpha - 1)\langle x - y_x, y_x\rangle \leq 0$ for $\alpha \geq 0$. Since $(\alpha - 1) \gtrless 0$ we have $\langle x - y_x, y_x\rangle = 0$ and Inequality (3.5) becomes $\langle y, x - y_x\rangle \leq 0$ for all $y \in K \Rightarrow x - y_x \in K$.
"$\Longleftarrow$": Let $y_x \in K$ and $x - y_x \in K^\circ$, $\langle x - y_x, y_x\rangle = 0$.
For $y \in K$ we write

$$\begin{aligned} f_x(y) &= \frac{1}{2}\|x - y_x + y_x - y\|^2 \\ &= \frac{1}{2}\|x - y_x\|^2 + \frac{1}{2}\|y_x - y\|^2 + \|x - y_x\|\|y_x - y\| \\ &\geq f_x(y_x) + \langle x - y_x, y_x - y\rangle \end{aligned}$$

with $\langle x - y_x, y_x\rangle = 0$ we have

$$\begin{aligned} \langle x - y_x, y_x - y\rangle &= \langle x - y_x, -y\rangle \\ &\geq 0 \end{aligned}$$

since $(x - y_x) \in K^\circ \Rightarrow f_x(y) \geq f_x(y_x)$.

□

Other Properties in The Conical Case

- $p_K(x) = 0 \Longleftrightarrow x \in K^\circ$.
- $p_K(\alpha x) = \alpha p_K(x)$ for all $\alpha \geq 0$.
- $p_K(-x) = -p_{-K}(x)$.
- $p_K(x) + p_{K^\circ}(x) = x$.

Theorem 3.5 *Let K be a closed and convex cone. For x_1, x_2 the following properties are equivalent*

a) $x = x_1 + x_2$ with $x_1 \in K, x_2 \in K^\circ$ and $\langle x_1, x_2 \rangle = 0$.

b) $x_1 = p_K(x)$ and $x_2 = p_{K^\circ}(x)$.

Proof: $p_K(x)$ is characterized by $p_K(x) \in K, x - p_K(x) \in K^\circ$ and $\langle x - p_K(x), p_K(x) \rangle = 0$.
We show that $x - p_K(x) = p_{K^\circ}(x)$, i.e. $p_{K^\circ}(x) \in K^\circ \Leftrightarrow x - p_K(x) \in K^\circ$. $x - p_{K^\circ}(x) \in K \Leftrightarrow p_K(x) \in K^{\circ\circ}$. $\langle x - (x - p_K(x)), x - p_K(x) \rangle = \langle p_K(x), x - p_K(x) \rangle = 0$. Therefore $x = p_K(x) + p_{K^\circ}(x)$. We show that $K = K^{\circ\circ}$, $K \subset K^{\circ\circ}$. "$K^{\circ\circ} \subset K$": K is closed, convex. Let $x_0 \notin K \Rightarrow \exists$ separating hyperplane $H_{s,1}$ such that $\langle s, x \rangle \leq 1 \; \forall \, x \in K$ and $\langle s, x_0 \rangle > 1 \Rightarrow x_0 \notin K^{\circ\circ} \Rightarrow K^{\circ\circ} \subset K$. "$K \subset K^{\circ\circ}$": $K^\circ = \{s : \langle s, x \rangle \leq 0 \; \forall \, x \in K\}$, $K^{\circ\circ} = \{x : \langle x, s \rangle \leq 0 \; \forall \, s \in K^\circ\}$. Let $x \in K \Rightarrow \langle s, x \rangle \leq 0 \; \forall \, s \in K^\circ \Rightarrow x \in K^{\circ\circ} \Rightarrow K \subset K^{\circ\circ}$.

□

Remark $x = x_1 + x_2, x_1 \in K, x_2 \in K^\circ$

$$\Longrightarrow \|x_1\| \geq \|p_K(x)\|, \|x_2\| \geq \|p_{K^\circ}(x)\|.$$

3.1 Exercises

1. a) Let $K_1, K_2 \subset \mathbb{R}^n$ be cones. Prove that

   1. $(K_1 + K_2)^0 = K_1^0 \cap K_2^0 = (K_1 \cup K_2)^0$.
   2. $(K_1 \cap K_2)^0 \supset K_1^0 + K_2^0 = \text{co}\,(K_1^0 \cup K_2^0)$.

   b) The 1-polar of a set $K \subset \mathbb{R}^n$ is $K^1 := \{s : \langle s, x \rangle \leq 1\}$. Show that

1. $(\lambda K)^1 = \frac{1}{\lambda} K^1$
2. $K \subset K^{11}$ and $K^1 = K^{111}$
3. If K is a subspace then $K^1 = K^\perp$
4. If K is a cone then $K^1 = K^0$

2. Let C be a closed convex set. show that

a) p_C is continuous.

b) Let

$$\begin{aligned} f_x^\infty(y) : &= \max_{i=1}^{n} |y_i - x_i| \\ S &= \{y : f_x^\infty(x) \leq f_x^\infty(c)\} \end{aligned}$$

for some $c \in C$. The infimum

$$\inf_{y \in C \cap S} f_x^\infty(y)$$

is attained, but not necessarily unique.

Chapter 4

SEPARATION RESULTS

4.1 Separation of Convex Sets

Theorem 4.1 *Let $C \subset \mathbb{R}^n$, $C \neq \emptyset$, C convex, closed and $x \notin C$. Then there exists $s \in \mathbb{R}^n$ such that $\langle s, x\rangle > \sup\{\langle s, y\rangle : y \in C\}$.*

Proof: Set $s := x - p_C(x)$ ($\neq 0$ since $x \notin C$).
Use Theorem 3.2 ($\langle x - p_C(x), y - p_C(x)\rangle \leq 0\ \forall y \in C$) and get $0 \geq \langle s, y - x + s\rangle = \langle s, y\rangle - \langle s, x\rangle + \|s\|^2$
$\Rightarrow \langle s, x\rangle - \|s\|^2 \geq \langle s, y\rangle$ for all $y \in C$.

□

Geometrical Interpretation

(Refer to Figure 4.1). If $s \neq 0$ then $\langle s, x\rangle = r$ defines a hyperplane $H_{s,r}$. If we vary r, we get a set of parallel hyperplanes. Choose $s \neq 0$ as in Proof of Theorem 4.1 and take

$$r = r_s := \frac{1}{2}(\langle s, x\rangle + \sup_{y \in C}\langle y, s\rangle)$$

$\Rightarrow \langle s, x\rangle - r_s \geq 0$ and $\langle s, y\rangle - r_s < 0$ for all $y \in C$. Hence $H_{s,r}$ separates C and $\{x\}$.

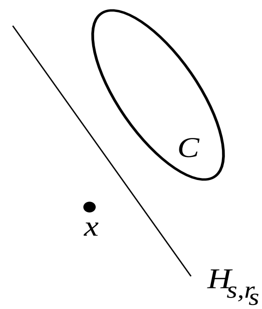


Figure 4.1: Geometrical interpretation of separation.

Corollary 4.2 (Strict Separation of Convex Sets) *Let $C_1,\ C_2 \subset \mathbb{R}^n$, $C_1 \neq \emptyset$, $C_2 \neq \emptyset$, C_1, C_2 closed and convex. $C_1 \cap C_2 = \emptyset$. If C_2 is bounded then there exists $s \in \mathbb{R}^n$ such that*

$$\sup_{y \in C_1} \langle s, y \rangle < \min_{y \in C_2} \langle s, y \rangle.$$

Proof: We know that $C_1 - C_2$ is convex (Proposition 1.1) and closed (C_2 is compact). From $C_1 \cap C_2 = \emptyset$, it follows that $0 \notin C_1 - C_2$. Theorem 4.1 $\Rightarrow \exists\ s \in \mathbb{R}^n : s$ separates $C_1 - C_2$ from $\{0\}$.

$$\sup\{\langle s, y \rangle : y \in C_1 - C_2\} < \langle s, 0 \rangle = 0$$

$$\begin{aligned} \Rightarrow 0 &> \sup_{y_1 \in C_1} \langle s, y_1 \rangle + \sup_{y_2 \in C_2} \langle s, -y_2 \rangle \\ &= \sup_{y_1 \in C_1} \langle s, y_1 \rangle - \inf_{y_2 \in C_2} \langle s, y_2 \rangle. \end{aligned}$$

Since C_2 is bounded

$$\sup_{y_1 \in C_1} \langle s, y_1 \rangle - \inf_{y_2 \in C_2} \langle s, y_2 \rangle = \sup_{y_1 \in C_1} \langle s, y_1 \rangle - \min_{y_2 \in C_2} \langle s, y_2 \rangle.$$

$\Rightarrow \min_{y_2 \in C_2} \langle s, y_2 \rangle > \sup_{y_1 \in C_1} \langle s, y_1 \rangle.$

□

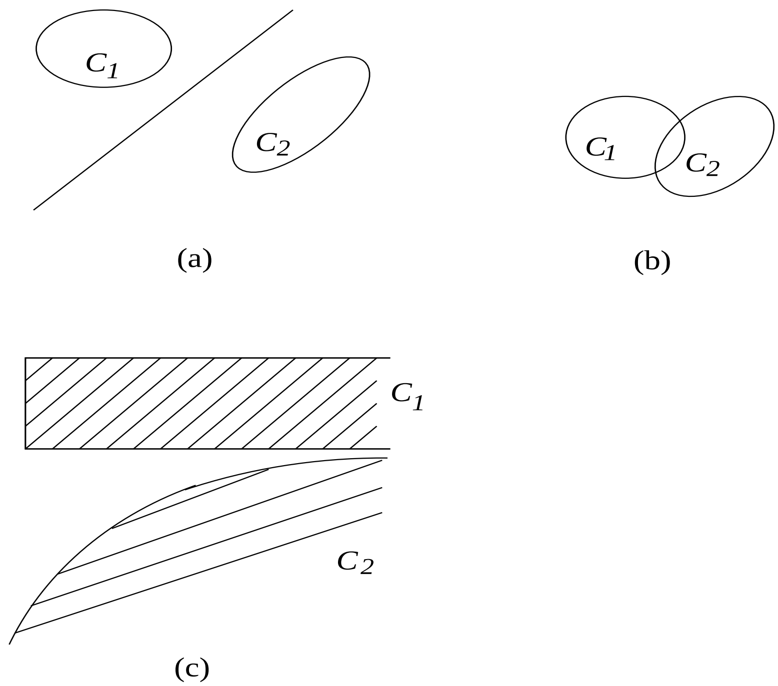


Figure 4.2: Examples of separation: (a) strict separation is possible; (b) no separation is possible; (c) no strict separation is possible (compactness is needed).

Definition 22 $C_1, C_2 \subset \mathbb{R}^n$, *non-empty convex sets are called* **properly separated by $s \in \mathbb{R}^n$** *if*

$$\sup_{y_1 \in C_1} \langle s, y_1 \rangle \leq \inf_{y_2 \in C_2} \langle s, y_2 \rangle$$

and

$$\inf_{y_1 \in C_1} \langle s, y_1 \rangle < \sup_{y_2 \in C_2} \langle s, y_2 \rangle.$$

Theorem 4.3 (Proper Separation of Convex Sets) *If C_1, $C_2 \in \mathbb{R}^n$, C_1,*
$C_2 \neq \emptyset$, convex and satisfying $(\text{ri } C_1) \cap (\text{ri } C_2) = \emptyset$, then C_1 and C_2 can be properly separated.

4.2 Consequences of Separation

Existence of supporting hyperplanes

Let $C \subsetneq \mathbb{R}^n$, C convex. We have cl $C \neq \mathbb{R}^n$. ($C \supset$ ri $C =$ ri cl $C =$ ri $\mathbb{R}^n = \mathbb{R}^n$, see Proposition 2.3). Therefore we know by Theorem 4.1 that there exists a hyperplane $H_{s,r}$ separating cl C from $x \notin$ cl C, and $H_{s,r}$ is supporting.

Lemma 4.4 *$C \subsetneq \mathbb{R}^n$, $C \neq \emptyset$, C convex, $x \in$ bd C. Then there exists a hyperplane supporting C at x.*

Proof: Since C, cl C and their complements have the same boundary, (Section 2.1), we can find a sequence $\{x_k\}$ with $x_k \notin$ cl C, $k = 1, 2, \ldots$ and $\lim_{k \to +\infty} x_k = x$. For each k, we have by Theorem 4.1 s_k with $\|s_k\| = 1$ such that $\langle s_k, x_k - y\rangle > 0$ for all $y \in C \subset$ cl C. Find a subsequence such that $s_k \to s \neq 0$ and take the limit to obtain

$$\langle s, x - y\rangle \geq 0 \text{ for all } y \in C \Rightarrow \langle s, x\rangle = r \geq \langle s, y\rangle \text{ for all } y \in C.$$

□

Problem

(Assume $\dim C \leq n - 1$). In the proof of Lemma 4.4 we found $s \in \mathbb{R}^4$: $\langle s, x - y\rangle \geq 0$ for all $y \in C \Rightarrow \langle s, x\rangle \leq \langle s, y\rangle = r$. Make a construction relative to aff C, as shown below in Figure 4.3.

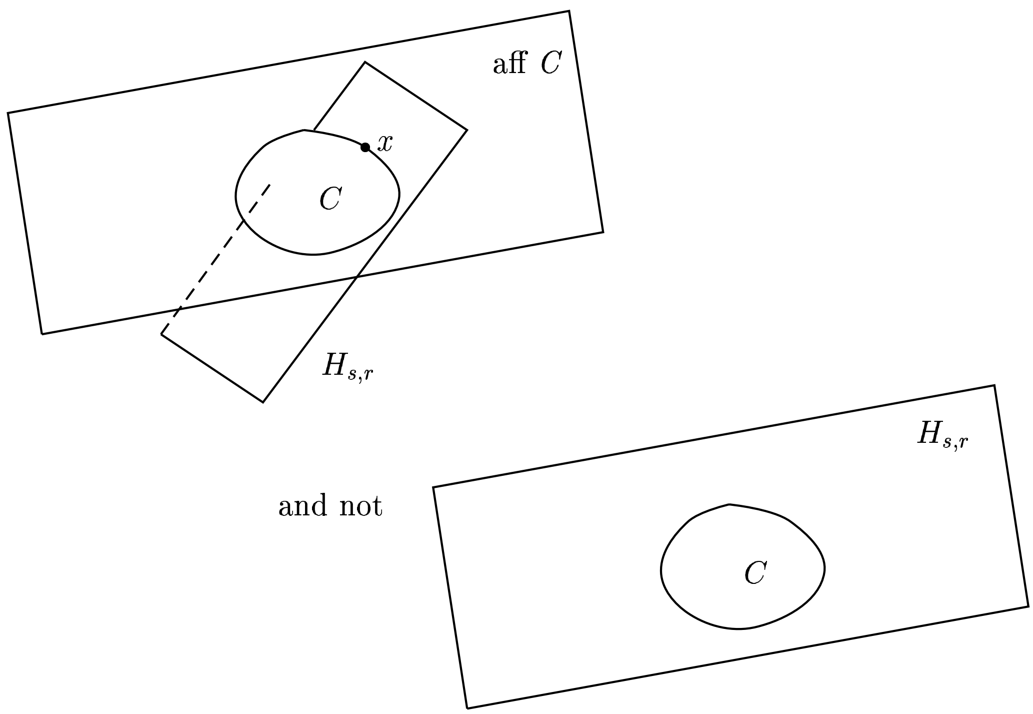


Figure 4.3: Illustration for Remark to Lemma 4.4

Let V be the subspace parallel to aff C, $U = V^{\perp}$. We have $\langle s, y-x\rangle = 0$ for all $s \in U, y \in C$. Assume $x \in$ rbd C and translate C to $C_0 := C - \{x\} \Rightarrow$ C_0 is a convex set in V and $0 \in$ rbd C_0. Build a sequence $\{x_k\} \subset V \setminus$ cl C_0 tending to 0 (and this is analogous to proof of Lemma 4.4) and $s_k \in V$, $\|s_k\| = 1$ separating x_k from $C_0 \Rightarrow s_k \to s (s \neq 0) \in V$ separates (not strictly) $\{0\}$ and $C_0 \Rightarrow s$ separates $\{x\}$ and C.

Definition 23 *$H_{s,r}$ is a* **non-trivial support (at x)** *if* $\boldsymbol{s \notin U}$, *i.e., for* $\boldsymbol{s = s_v + s_u}$, $\boldsymbol{\mathbb{R}^n = V \oplus U}$ *we have* $\boldsymbol{s_v \neq 0}$.

Then $C \not\subseteq H_{s,r}$ since otherwise $r = \langle s, y\rangle = \langle s_v, y\rangle + \langle s_u, x\rangle$ for all $y \in C \Rightarrow$ $\langle s_v, \cdot\rangle$ is constant on $C \Rightarrow s_v \in U \Rightarrow s_v = 0$ and that is a contradiction.
We may moreover assume that $s_u = 0$. If $s = s_v + s_u$ is a non-trivial support $\Rightarrow s_v = s_v + 0$ is also a non-trivial support. $H_{s_v,r}$ is a hyperplane orthogonal to C.

Outer Description of Closed Convex Sets

We have already seen that for a convex set $C \subsetneq \mathbb{R}^n$ we have a supporting hyperplane $H_{s,r}$ with $C \subset H^{\leq}_{s,r}$. Then

$$\begin{aligned} \Sigma_C &:= \{(s,r) \in \mathbb{R}^n \times \mathbb{R} : C \subset H^{\leq}_{s,r}\} \\ &= \{(s,r) \in \mathbb{R}^n \times \mathbb{R} : \langle s, y\rangle \leq r \text{ for all } y \in C\} \neq \emptyset. \end{aligned}$$

Then we can intersect all halfspaces $H^{\leq}_{s,r}$ with $(s,r) \in \Sigma_C$, as shown in Figure 4.4.

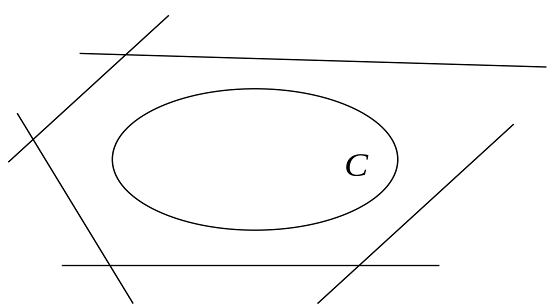


Figure 4.4: Outer description of a closed convex set C.

$$\begin{aligned} C \subset C^* &:= \bigcap_{(s,r)\in\Sigma_C} H^{\leq}_{s,r} \\ &= \{z \in \mathbb{R}^n : \langle s, z\rangle \leq r \text{ whenever } \langle s, y\rangle \leq r \ \forall y \in C\}. \end{aligned}$$

Theorem 4.5 *Let $\emptyset \neq C \subsetneq \mathbb{R}^n$, C convex. Then $C^* = \text{cl } C$.*

Proof:

- "$C^* \supset \text{cl } C$": By construction, intersection of halfspaces.
- "$C^* \subset \text{cl } C$": Take $x \notin \text{cl } C \Rightarrow$ we can separate x and cl C, i.e.,

$$\exists\, s_0 \neq 0 : \langle s_0, x\rangle > \sup_{y\in C}\langle s_0, y\rangle := r_0.$$

  Then $(s_0, r_0) \in \Sigma_C$ but $x \notin H^{\leq}_{s_0,r_0} \Rightarrow x \notin C^* \Rightarrow C^* \subset \text{cl } C$.

□

Corollary 4.6 *The data $(s_j, r_j) \in \mathbb{R}^n \times \mathbb{R}$ for $j \in \mathcal{J}$, $\mathcal{J}$ arbitrary set, is equivalent to the data of a closed convex set C via the relation*

$$C = \bigcap_{j\in\mathcal{J}} \{x \in \mathbb{R}^n : \langle s_j, x\rangle \leq r_j\}.$$

Proof: For given C, define $\mathcal{J} := \Sigma_C$. If $\mathcal{J}$ is given C is a closed convex set.

□

Definition 24 *A* **closed convex polyhedron** *is an intersection of finitely many halfspaces.*
Take $(s_1, r_1), \ldots, (s_m, r_m) \in \mathbb{R}^n \times \mathbb{R}$, with $s_i \neq 0$, $i = 1, \ldots, m$.
Then define

$$P := \{x \in \mathbb{R}^n : \langle s_j, x\rangle \leq r_j, j = 1, \ldots, m\} = \{x \in \mathbb{R}^n : Ax \leq b\},$$

where A is the matrix with row s_j and $b = (r_1, \ldots, r_m) \in \mathbb{R}^m$. A **closed convex polyhedral cone** *is the special case where $b = 0$.*

Inner Description of Convex Sets

Proof of Minkowski's Theorem (Theorem 2.10):
Proof: By induction over $\dim C =: k$.

- $k = \dim C = 0$. $C = \{x\}$.
- $k - 1 \to k$: Take $x \in C$:

- Case 1: $x \in \text{rbd } C$: By the Remark to Lemma 4.4 there exists a nontrivial hyperplane H supporting C at x. We have $\dim C \cap H \leq k-1$, $C \cap H$ is nonempty, convex, compact. By the induction hypothesis $\Rightarrow C \cap H$ is a convex combination of extreme points in that set. Moreover, $C \cap H$ is an exposed face of C. By Proposition (2.11) $\Rightarrow \text{ext } (C \cap H) \subset \text{ext } C$.
- Case 2: $x \in \text{ri } C$: Take $x' \in C$, $x' \neq x$ ($\dim C > 0$). The affine line defined by $[x, x']$ cuts rbd C in at most two points (Remark to Lemma 2.2). Since C is compact, they intersect at exactly two points y, z. By case 1 $y, z \in \text{rbd } C$ can be written as convex combinations from elements of ext C
$\Rightarrow x = \lambda y + (1-\lambda) z$, $\lambda \in (0,1)$ can also be written as a convex combination of ext C.

□

Bipolar of a Convex Cone

Definition 25 *Let $C \subset \mathbb{R}^n$ be nonempty, closed, convex. The function*

$$\sigma_C(s) := \sup\{\langle s, y\rangle : y \in C\}$$

is called the **support function** *of C. For $x \in C$ we have by definition: $\langle s, x\rangle \leq \sigma_C(s)$ for all $s \in \mathbb{R}^n$.*

Proposition 4.7 *Let $K \subset \mathbb{R}^n$ be a convex cone with polar K°. Then*

$$K^{\circ\circ} := (K^\circ)^\circ = \text{cl } K.$$

Proof:

$$\sigma_{\text{cl } K}(s) = \begin{cases} \langle s, 0\rangle = 0 & \text{if } \langle s, x\rangle \leq 0 \ \ \forall x \in \text{cl } K \\ +\infty & \text{otherwise} \end{cases}$$

$\Rightarrow \sigma_{\text{cl } K}(s) = 0$ if $x \in K^\circ$ and $+\infty$ otherwise. Using Theorem 4.1: $x \in \text{cl } K$ $\Longleftrightarrow \langle s, \cdot\rangle \leq \sigma_{\text{cl } K}(\cdot)$, which becomes here

$$x \in \text{cl } K \Longleftrightarrow \left\{ \begin{array}{ll} \langle s, x\rangle \leq 0 & \forall\, s \in K^\circ \\ \langle s, x\rangle \text{ arbitrary} & \forall\, s \notin K^\circ \end{array} \right\} = (K^\circ)^\circ = K^{\circ\circ}.$$

□

Remark If cl $K = K$ then $K = K^{\circ\circ}$.

4.3 The Lemma of (Minkowski-) Farkas

In linear algebra: Let A be a $n \times m$ matrix and $b \in \mathbb{R}^n$.

$$\begin{gathered} A\alpha = b \text{ has a solution in } \mathbb{R}^m \\ \Longleftrightarrow \\ b \in \text{lin } A = [\ker A^T]^\perp \\ \Longleftrightarrow \\ b^\perp \supset \ker A^T \\ \Longleftrightarrow \\ \{x \in \mathbb{R}^n : b^T x = 0\} \supset \{x \in \mathbb{R}^n : A^T x = 0\}. \end{gathered}$$

Denote by $s_1, \ldots, s_m$ the columns of A. This is equivalent to $b \in \text{lin } \{s_1, \ldots, s_m\}$ if and only if $\langle b, x\rangle = 0$ whenever $\langle s_j, x\rangle = 0$, $j = 1, \ldots, m$.

Lemma 4.8 (Farkas I) *Let $b, s_1, \ldots, s_m$ be given in $\mathbb{R}^n$.*

$$\{x \in \mathbb{R}^n : \langle s_j, x\rangle \leq 0, j = 0, \ldots, m\} \subset \{x \in \mathbb{R}^n : \langle b, x\rangle \leq 0\} \tag{1}$$

$$\Longleftrightarrow b \in \text{cone } \{s_1, \ldots, s_m\}. \tag{2}$$

Lemma 4.9 (Farkas II) *Let $b, s_1, \ldots, s_m \in \mathbb{R}^n$. Then exactly one of the following statements is true*

$$P{:=} \textit{ There exist } \alpha_1, \ldots, \alpha_m, \alpha_i \geq 0, i = 1, \ldots, m \textit{ such that } b = \sum_{j=1}^m \alpha_j s_j. \tag{3}$$

$$Q{:=} \textit{ The system of inequalities } \langle b, x\rangle > 0, \langle s_j, x\rangle \leq 0, j = 1, \ldots, m \textit{ has a solution } x \in \mathbb{R}^n. \tag{4}$$

Note that (3) = (2) and (4) = not (1).
Now a complete geometric version follows. Let $K := \text{cone } \{s_1, \ldots, s_m\}$. Then K° is the left hand side of (1), i.e., $\{x \in \mathbb{R}^n : \langle s_j, x\rangle \leq 0,\ j = 1, \ldots, m\}$. Therefore Farkas' Lemma asserts: $b \in K$ (2) $\Longleftrightarrow$ $b \in K^{\circ\circ}$ (1) or shorter $K = K^{\circ\circ} \Longleftrightarrow K$ is closed (by Proposition 4.7 and Remark).

Lemma 4.10 (Farkas III) *Let $s_1, \ldots, s_m \in \mathbb{R}^n$. Then the convex cone*

$$K := \text{cone } \{s_1, \ldots, s_m\} = \{\sum_{j=1}^m \alpha_j s_j : \alpha_j \geq 0, j = 1, \ldots, m\} \textit{ is closed.}$$

Proof:
Case 1: Vectors $s_1, \dots, s_m$ are linearly independent. Then the convergence of $x^k = \sum_{j=1}^m \alpha_j^k s_j$ is equivalent to the convergence of each $\{\alpha_j^k\}$ to some α_j with $\alpha_j \geq 0$ since each $\alpha_j^k \geq 0 \Rightarrow x = \sum \alpha_j s_j \in K$.
Case 2: Vectors $s_1, \dots, s_m$ are linearly dependent. Then $\sum_{j=1}^m \beta_j x_j = 0$ has a solution $\beta \in \mathbb{R}^m$, $\beta \neq 0$. Assume $\beta_j \leq 0$ for some $j \in \{1, \dots, m\}$ (change to $-\beta$ if necessary). Write as in the proof of Theorem 1.5 (Caratheodory) each $x \in K$ as

$$x = \sum_{j=1}^m \alpha_j s_j = \sum_{j=1}^m (\alpha_j + t^*(x)\beta_j) s_j = \sum_{j=1, j \neq i(x)}^m \alpha_j' s_j, \tag{4.1}$$

where $i(x) \in \underset{\beta_j<0}{\operatorname{argmin}} \frac{-\alpha_j}{\beta_j}$, $t^* := \frac{-\alpha_{i(x)}}{\beta_{i(x)}}$
$\Rightarrow \alpha_j' := \alpha_j + t^*(x)\beta_j \geq 0$. If we vary $x \in K$ we construct a decomposition $K = \bigcup \{K_i : i = 1, \dots, m\}$ where $K_i = \text{cone } s_j,\ j \in \{1, \dots, m\},\ j \neq i$. If the generators of some K_i are still linearly dependent then repeat the above construction. After finitely many iterations we get a decomposition of K as a finite union of convex cones K_i with linearly independent generators. Using Case 1 $\Rightarrow$ each K_i is closed $\Rightarrow K = \bigcup K_i$ is closed.

□

Theorem 4.11 (Generalized Farkas) *Let (b, r), $(s_j, \rho_j) \in \mathbb{R}^n \times \mathbb{R}$, $j \in \mathcal{J}$. Furthermore, let*

$$\langle s_j, x \rangle \leq \rho_j \ \forall\, j \in \mathcal{J} \tag{4.2}$$

have a solution $x \in \mathbb{R}^n$ (the system is consistent). Then the following two properties are equivalent:

i) $\langle b, x \rangle \leq r$ for all x satisfying Inequality (4.2).

ii) $(b, r) \in \text{cl cone } s = \overline{\text{cone}}\ S,\ S := \{(0, 1)\} \cup \{s_j, \rho_j\}_{j \in \mathcal{J}}$.

Proof: Idea of the proof: Show ii) $\Rightarrow$ i) directly and not ii) $\Rightarrow$ not i) by separating (b, r) from cl (cone S).

□

Relation to Farkas' Lemma:

1) r, $\rho_j = 0$, $j \in \mathcal{J} \Rightarrow x = 0$ satisfies Inequality (4.2) ($\langle s_j, 0 \rangle \leq 0$).
So we do not need Inequality (4.2) to state i). We can write directly
i') $\langle s_j, x \rangle \leq 0$, $j \in \mathcal{J} \Rightarrow \langle b, x \rangle \leq 0$ is equivalent to
ii') $b \in \overline{\text{cone}}\ \{s_j : j \in \mathcal{J}\}$

2) $\mathcal{J} = \{1, \dots, m\} \Rightarrow$ we can use the dot product for $\langle \cdot, \cdot \rangle$.
Let A be a matrix with columns $s_j's$, and $\rho = (\rho_1, \dots, \rho_m)$.
$i'')$ $\{x \in \mathbb{R}^n : A^T x \leq \rho\} \subset \{x \in \mathbb{R}^n : b^T x \leq r\}$
is equivalent to
$ii'')$ $\exists\, \alpha \in \mathbb{R}^m_{+0} : A\alpha = b, \rho\alpha \leq r$.
1) and 2) give
$i''')$ $\{x \in \mathbb{R}^n : A^T x \leq 0\} \subset \{x \in \mathbb{R}^n : b^T x \leq 0\}$ is equivalent to
$ii''')$ $b \in \text{cone}\,\{s_1, \dots, s_m\} = \overline{\text{cone}}\,\{s_1, \dots, s_m\}$, by Farkas' Lemma.

4.4 Exercises

1. For $S = [0, 1]$ and $S = [1, 2] \times [1, 2]$ compute the support function $\sigma_S(x)$.

2. Let $V \subset \mathbb{R}^n$ be a linear subspace such that $V \cap \mathbb{R}^n{}_{+0} = \{0\}$. Prove that $V^\perp \cap \text{int}\, \mathbb{R}^n_{+0} \neq \emptyset$ using the separation theorem.

3. Let $A : \mathbb{R}^n \to \mathbb{R}^k$ be a linear mapping. Then either

   a) $A^T x > 0$ $(i.e. (A^T x)_i > 0,\ i = 1, \dots, n)$ has a solution $x \in \mathbb{R}^k$, or

   b) $Ay = 0,\ y \geq 0,\ y \neq 0$ has a solution $y \in \mathbb{R}^n$

   but never both.

Chapter 5

CONICAL APPROXIMATIONS OF CONVEX SETS

For $S \subset \mathbb{R}^n$ "smooth" we have a unique hyperplane tangent to a $x \in S$, or if $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable gr $f \approxeq \{(y,r) : r - f(x) = \langle \nabla f(x), y - x \rangle\}$. Convex sets need not be smooth. See Figure 5.1.

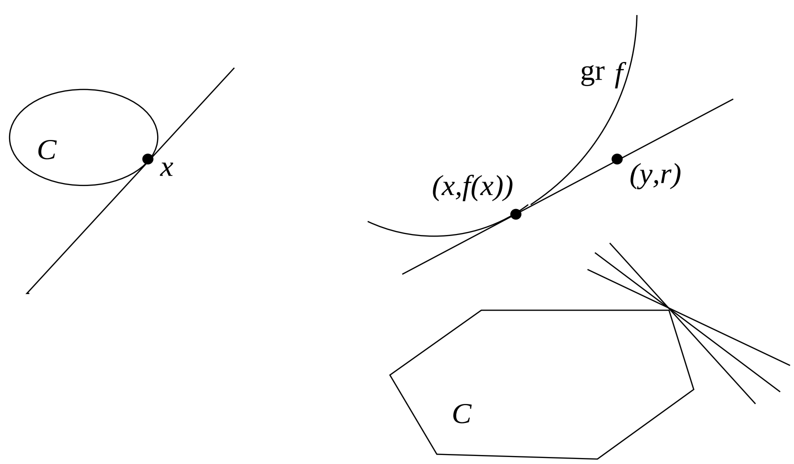


Figure 5.1: Examples of sets and tangent hyperplane.

5.1 Tangent Cones to General Closed Sets

Let $S \subset \mathbb{R}^n$ be closed, not necessarily convex.

Definition 26 *Let $S \subset \mathbb{R}^n$, $S \neq \emptyset$, closed. Then $d \in \mathbb{R}^n$ is a direction*

tangent *to* S *at* $x \in S$ *if there exists a sequence* $\{x_k\} \subset S$ *and a sequence* $\{t_k\}$*, such that for* $k \to \infty$ *we have*

$$x_k \to x,\ t_k \searrow 0,\ \frac{x_k - x}{t_k} \to d. \tag{5.1}$$

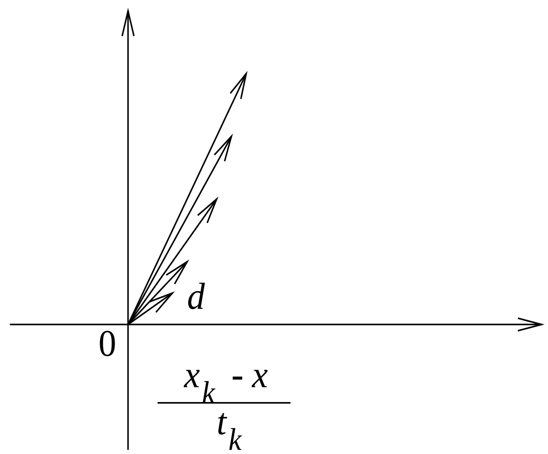


Figure 5.2: Tangent direction.

The set of all tangent directions to S *at* $x \in S$ *is called the* **tangent cone** *(or contingent cone) to* S *at* $x \in S$*, and is denoted by* $T_S(x)$*.*

Remark

- $0 \in T_S(x)$ ($x_k = x$ for all k).
- $d \in T_S(x) \Longrightarrow \alpha d \in T_S(x)$ for all $\alpha > 0$ (change the sequence $\{t_k\}$ to $\{\frac{t_k}{\alpha}\}$).
- $x \in \text{int } S \Longrightarrow T_S(x) = \mathbb{R}^n$. Hence only $x \in \text{bd } S$ are interesting.

Proposition 5.1 *A direction* $d \in \mathbb{R}^n$ *is tangent to* S *at* $x \in S$

$$\Longleftrightarrow \exists\ \{d_k\} \to d, \exists\ \{t_k\} \searrow 0 \text{ such that } x + t_k d_k \in S \text{ for all } k.$$

Proof: Set in (5.1) $d_k := \frac{x_k - x}{t_k} (\to d) \Rightarrow x_k = x + t_k\, d_k (\in S)$.

□

Proposition 5.2 *The tangent cone is closed.*

Proof: Take $\{d_l\} \subset T_S(x)$, $x \in S$ with $\{d_l\} \to d$. For each l take sequences $\{x_{l,k}\}_k$ and $\{t_{l,k}\}_k$ associated with d_l as defined in (5.1) . For a fixed $l > 0$ we can find k_l such that $\|\frac{x_{l,k_l}-x}{t_{l,k_l}} - d_l\| \leq \frac{1}{l}$ $(\frac{x_{k,l}-x}{t_{l,k}} \to d_l$ by definition). As $l \to \infty$ we get sequences $\{x_{l,k_l}\}_l$ and $\{t_{l,k_l}\}_l$ defining d as an element of $T_S(x)$. An illustration is given in Figure 5.3.

□

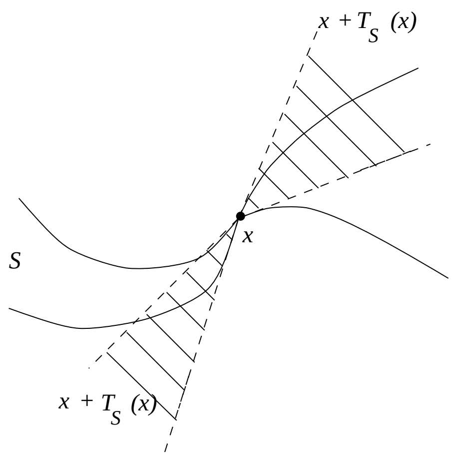


Figure 5.3: Illustration for proof of Proposition 5.2.

Example 5.1 *Let* $S := \{x \in \mathbb{R}^n : c(x) \leq 0\}$, $c : \mathbb{R}^n \to \mathbb{R}$ *continuous differentiable. Take* $x \in S$ *with* $c(x) = 0$ *and* $\nabla c(x) \neq 0$.
$T_S(x) = \{d \in \mathbb{R}^n : \langle \nabla c(x), d\rangle \leq 0\}$ *(half-space).*

5.2 Tangent and Normal Cones to a Convex Set

In this subsection, $C \subset \mathbb{R}^n$ is closed and convex.

Proposition 5.3

$$\begin{aligned} T_C(x) &= \overline{\text{cone}}\,(C - x) \\ &= \text{cl}\,(\mathbb{R}_{+0}(C-x)) \\ &= \text{cl}\,\{d \in \mathbb{R}^n : d = \alpha(y-x), y \in C, \alpha \geq 0\}. \end{aligned}$$

Proof: "$T_C(x) \supset \overline{\text{cone}}\ (C - x)$": Since C is convex, $C - \{x\} \subset T_C(x)$. Since by Proposition 5.2 $T_C(x)$ is closed $\Rightarrow$ cl $(\mathbb{R}_{+0}(C - x)) \subset T_C(x)$.
"$\overline{\text{cone}}\ (C-x) \supset T_C(x)$": For $\in T_C(x)$ take the corresponding $\{x_k\}, \{t_k\}$ from the definition $\frac{x_k - x}{t_k} \in \mathbb{R}_{+0}(C-x) \Rightarrow d \in$ cl $(\mathbb{R}_{+0}(C-x)) \Rightarrow d \in \overline{\text{cone}}\ (C-x)$.

□

Since $T_C(x)$ is a closed convex set, it can be described as an intersection of closed half-spaces.

Definition 27 *A direction $s \in \mathbb{R}^n$ is* **normal** *to C at $x \in C$ if $\langle s, y-x\rangle \le 0$ for all $y \in C$. The set of all normal directions to C at $x \in C$ is called the* **normal cone** *to C at x, and is denoted by $N_C(x)$.*

Remark

- The angle between s and $y - x$ is obtuse.
- From Section 4.2 we know that there exists a nonzero normal direction at each $x \in$ bd C (Lemma 4.4).
- Theorem 3.2 implies $v - p_C(v) \in N_C(p_C(v))$ for all $v \in \mathbb{R}^n$.
- For $x \in$ int C, $N_C(x) = \{0\}$ (all directions are in $B(x, \delta)$).

Example 5.2 $C := H^{\le}_{s,r} = \{y \in \mathbb{R}^n : \langle s, y\rangle \le r\}$.
Let $x \in H^{\le}_{s,r} \Rightarrow N_C(x) = \{\lambda s : \lambda \ge 0\}$.

$\langle \lambda s, y - x\rangle = \lambda(\overbrace{\langle s, y\rangle}^{\le r} - \overbrace{\langle s, x\rangle}^{=r}) \le 0$. $y \in C$. *Refer to Figure 5.4.*

Proposition 5.4 *The normal cone is the polar of the tangent cone*

$$N_C(x) = (T_C(x))^\circ,\ x \in C.$$

Proof: "$N_C(x) \subset (T_C(x))^\circ$": Let $s \in N_C(x)$, i.e. $\langle s, d\rangle \le 0$ for all $d \in C - x$
$\Rightarrow \langle s, d\rangle \le 0$ for all $d \in \mathbb{R}_{+0}(C - x)$
$\Rightarrow \langle s, d\rangle \le 0$ for all $d \in$ cl $(\mathbb{R}_{+0}(C - x)) = T_C$ by Proposition 5.3
$\Rightarrow s \in (T_c(x))^\circ$.
"$(T_C(x))^\circ \subset N_C(x)$": Take $s \in (T_C(x))^\circ$
$\Rightarrow \langle s, d\rangle \le 0$ for all $d \in T_C(x) \supset C - x$
$\Rightarrow \langle s, d\rangle \le 0$ for all $d \in C - x \Rightarrow s \in N_C(x)$.

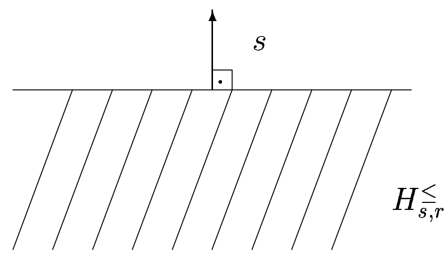


Figure 5.4: Illustration for Example 5.2.

□

Corollary 5.5 *The tangent cone is the polar of the normal cone.*

$$\begin{aligned} T_C(x) &= (N_C(x))^\circ. \\ T_C(x) &= \{d \in \mathbb{R}^n : \langle s, d\rangle \leq 0 \; \forall \; s \in N_C(x)\}. \end{aligned}$$

Proof: By Proposition 5.2 $T_C(x)$ is closed. By Proposition 4.7 $\Rightarrow (T_C(x))^{\circ\circ} = T_C(x) \Rightarrow N_C(x) = (T_C(x))^\circ \Leftrightarrow (N_C(x))^\circ = T_C(x)$.

□

Remark In Corollary 5.5 $T_C(x)$ is written as intersection of homogeneous half-spaces.

Example 5.3

a) Let $C = K$, K closed convex cone.

- $T_K(0) = K \Rightarrow K^\circ = N_K(0)$.
- *If $x \neq 0$, $x \in K$ then $\mathbb{R}\{x\} \subset T_K(x)$. We even have*

$$N_K(x) = \{s \in K^\circ : \langle s, x\rangle = 0\}, s \neq 0.$$

b) $C := \{x \in \mathbb{R}^n : \langle s_j, x\rangle \leq r_j,\ j = 1, \dots, m\}$,
$\mathcal{J}(x) := \{j = 1, \dots, m : \langle s_j, x\rangle = r_j\}$ (active constraints). Then

- $T_C(x) = \{d \in \mathbb{R}^n : \langle s_j, d\rangle \leq 0$ *for* $j \in \mathcal{J}(x)\}$ *and*
- $(T_C(x))^\circ = N_C(x) = \text{cone}\ \{s_j : j \in \mathcal{J}(x)\} = \{\sum_{j\in\mathcal{J}(x)} \alpha_j s_j : \alpha_j \geq 0\}$.

5.3 Some Properties of Tangent and Normal Cones

Proposition 5.6 *All C_i are nonempty closed convex sets.*

i) For $x \in C_1 \cap C_2$:

- $T_{C_1 \cap C_2} \subset T_{C_1}(x) \cap T_{C_2}(x)$,
- $N_{C_1 \cap C_2}(x) \supset N_{C_1}(x) + N_{C_2}(x)$.

ii) $C_i \subset \mathbb{R}^{n_i}$, $i = 1, 2$, $(x_1, x_2) \in C_1 \times C_2$,

- $T_{C_1 \times C_2}(x_1, x_2) = T_{C_1}(x_1) \times T_{C_2}(x_2)$,
- $N_{C_1 \times C_2}(x_1, x_2) = N_{C_1}(x_1) \times N_{C_2}(x_2)$.

iii) Let $A(x) = y_0 + A_0 x$ be an affine mapping with A_0 linear, $x \in C$. Then

- $T_{A(C)}(A(x)) = \mathrm{cl}\ (A_0\, T_C(x))$,
- $N_{A(C)}(A(x)) = A_0^{*-1}[N_C(x)]$,

where A_0^* is the adjoint of A.

iv) In particular:

- $T_{C_1 + C_2}(x_1 + x_2) = \mathrm{cl}\ (T_{C_1}(x_1) + T_{C_2}(x_2))$.
- $N_{C_1 + C_2}(x_1 + x_2) = N_{C_1}(x_1) \cap N_{C_2}(x_2)$.

Proof:

i) $d \in T_{C_1 \cap C_2}(x)$
$\Rightarrow \exists\, d_k \to d,\ \exists\, t_k \searrow 0$ such that $x + t_k d_k \in C_1 \cap C_2$
$\Rightarrow \exists\, d_k \to d,\ \exists\, t_k \searrow 0$ such that $x + t_k d_k \in C_1,\ x + t_k d_k \in C_2$
$\Rightarrow d \in T_{C_1}(x) \cap T_{C_2}(x)$, i.e. $T_{C_1 \cap C_2}(x) \subset T_{C_1}(x) \cap T_{C_2}(x)$.
$(T_{C_1}(x) \cap T_{C_2}(x))^\circ \subset (T_{C_1 \cap C_2}(x))^\circ = N_{C_1 \cap C_2}(x)$, by Proposition 5.4.
$(T_{C_1}(x))^\circ + (T_{C_2}(x))^\circ = N_{C_1}(x) + N_{C_2}(x)$.

ii) "$\subset$": $d \in T_{C_1 \times C_2}(x) \Rightarrow \exists\, d_k = (d_k^1, d_k^2) \to d,\ \exists\, t_k \searrow 0$ such that $(x_1, x_2) + t_k(d_k^1, d_k^2) \in C_1 \times C_2$
$\Rightarrow x_1 + t_k d_k^1 \in C_1 \Rightarrow d_1 \in T_{C_1}(x)$
$\Rightarrow x_2 + t_k d_k^2 \in C_2 \Rightarrow d_2 \in T_{C_2}(x)$.
$\Rightarrow d \in T_{C_1}(x_1) \times T_{C_2}(x_2)$.
"$\supset$": $d \in T_{C_1}(x_1) \times T_{C_2}(x_2) \Rightarrow d = (d^1, d^2),\ \exists\, (d_k^1, d_k^2) \to (d^1, d^2)$,

$t_k^1 \searrow 0$ $t_k^2 \searrow 0$ such that $x_1 + t_k^1 d_k^1 \in C_1,\ x_2 + t_k^2 d_k^2 \in C_2$
$\Rightarrow (x_1, x_2) + (t_k^1 d_k^1, t_k^2 d_k^2) \in C_1 \times C_2$
$\Rightarrow (x_1, x_2) + \min\{t_k^1, t_k^2\}(d_k^1, d_k^2) \in C_1 \times C_2$
$\Rightarrow d \in T_{C_1 \times C_2}(x_1, x_2)$.
$(T_{C_1 \times C_2}(x_1, x_2))^\circ = (T_{C_1}(x_1) \times T_{C_2}(x_2))^\circ$
$N_{C_1 \times C_2}(x_1, x_2) = N_{C_1}(x_1) \times N_{C_2}(x_2)$. We use: $(K_1 \times K_2)^\circ = {K_1}^\circ \times {K_2}^\circ$.
Proof: "$\supset$": Tivial.
"$\subset$": Let $s \in (K_1 \times K_2)^\circ \Rightarrow \langle s, (x_1, x_2)\rangle \leq 0\ \forall\ (x_1, x_2) \in K_1 \times K_2$
$\Rightarrow \langle s_1, x_1\rangle + \langle s_2, s_2\rangle \leq 0$. Suppose $s_1 \notin {K_1}^\circ \Rightarrow \langle s_1, x_1\rangle > 0$. Taking $x_2 = \lambda x, \lambda \to 0$ gives a contradiction $\Rightarrow s_1 \in {K_1}^\circ, s_2 \in {K_2}^\circ$ analogously $\Rightarrow s \in {K_1}^\circ \times {K_2}^\circ$.

iii) $d \in T_{A(C)}(A(x))$

$$\begin{aligned}
&\Leftrightarrow\ \exists\ A(x_k) \to A(x)\ t_k \searrow 0 \text{ such that } \frac{A(x_k) - A(x)}{t_k} \to d \\
&\Leftrightarrow\ \frac{y_0 + A_0(x_k) - A_0(x) - y_0}{t_k} \to d \\
&\Leftrightarrow\ \frac{A_0(x_k - x)}{t_k} \to d \\
&\Leftrightarrow\ A_0\left(\frac{1}{t_k}(x_k - x)\right) \to d \\
&\Leftrightarrow\ \frac{x_k - x}{t_k} \to d' \in T_C(x) \\
&\Leftrightarrow\ d \in \text{cl}\ (A_0(T_C(x))).
\end{aligned}$$

$$\begin{aligned}
s \in N_{A(C)}(A(x)) \ &\Leftrightarrow\ \langle s, A(y) - A(x)\rangle \leq 0\ \forall\, y \in C \\
&\Leftrightarrow\ \langle s, A_0(y - x)\rangle \leq 0\ \forall\, y \in C \\
&\Leftrightarrow\ \langle A_0^* s, y - x\rangle \leq 0\ \forall\, y \in C \\
&\Leftrightarrow\ A_0^* s \in N_C(x) \\
&\Leftrightarrow\ s \in A_0^{*-1}[N_C(x)].
\end{aligned}$$

$$\begin{aligned}
T_{A(C)}(A(x)) &= \overline{\text{cone}}\ (A(C) - A(x)) \\
&= \overline{\text{cone}}\ (A_0(C - x)) \\
&= \text{cl}\ (\mathbb{R}_{+0} A_0(C - x)) \\
&= \text{cl}\ (A_0(T_C(x))).
\end{aligned}$$

iv) $A(x_1, x_2) = x_1 + x_2$. $A : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, $A = A_0$.

$$\begin{aligned} T_{A(C)}(A(x)) &= T_{C_1+C_2}(x_1 + x_2) \\ &\overset{iii)}{=} \text{cl}\,(A(T_{C_1 \times C_2}(x))) \\ &\overset{ii)}{=} \text{cl}\,(A(T_{C_1}(x_1) \times T_{C_2}(x_2))) \\ &= \text{cl}\,(T_{C_1}(x_1) + T_{C_2}(x_2)). \end{aligned}$$

Find A^*:

$$\begin{aligned} \langle A^* y, x\rangle &= \langle y, Ax\rangle \\ &= \langle y, x_1 + x_2\rangle \\ &= \langle y, x_1\rangle + \langle y, x_2\rangle \end{aligned}$$

$\Rightarrow A^* y = (y, y)$. $A^* : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$.

$$\begin{aligned} N_{C_1+C_2}(x_1 + x_2) &= A^{*-1}[N_{C_1 \times C_2}(x)] \\ &\overset{ii)}{=} A^{*-1}[N_{C_1}(x_1) \times N_{C_2}(x_2)] \\ &= N_{C_1}(x_1) \cap N_{C_2}(x_2). \end{aligned}$$

□

Proposition 5.7 *Let $x \in C$, $s \in \mathbb{R}^n$. Then the following properties are equivalent:*

i) $s \in N_C(x)$.

ii) $x \in F_C(s)$, *i.e.* $\langle s, x\rangle = \max_{y \in C} \langle s, y\rangle$.

iii) $x = p_C(x + s)$.

Proof: "i) $\Leftrightarrow$ ii)":

$$\begin{aligned} s \in N_C(x) &\Leftrightarrow \langle s, y - x\rangle \leq 0 \text{ for all } y \in C \\ &\Leftrightarrow \langle s, y\rangle \leq \langle s, x\rangle \text{ for all } y \in C. \end{aligned}$$

"iii) $\Leftrightarrow$ i)":

$$\begin{aligned} x = p_C(x + s) \text{ by Theorem 3.1} &\Leftrightarrow \langle x + s - x, y - x\rangle \leq 0 \text{ for all } y \in C \\ &\Leftrightarrow \langle s, y - x\rangle \leq 0 \text{ for all } y \in C \\ &\Leftrightarrow s \in N_C(x). \end{aligned}$$

□

Geometric Interpretation

(Refer to Figure 5.5). $p_C^{-1} = \{x\} + N_C(x)$ for all $x \in C$ and
$x \neq x' \Rightarrow (\{x\} + N_C(x)) \cap (\{x'\} + N_C(x')) \neq \emptyset$ (otherwise projection would not be single valued).

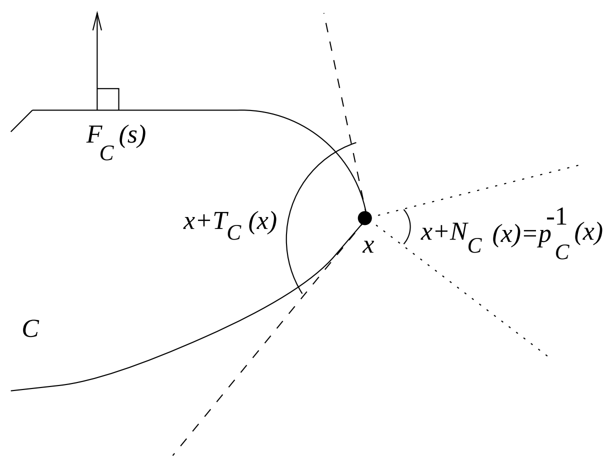


Figure 5.5: Geometric interpretation of tangent and normal cones.

5.4 Exercises

1.
   - Is the tangent cone $T_S(x)$ always convex? What if S is convex?
   - Let $C = K$, be a closed, convex cone. Determine $T_K(x)$, $N_K(x)$ for $x \in K$
   - Find the tangent cone $T_C(\alpha)$ and the normal cone $N_C(\alpha)$ of $C = \Delta_n = \{\alpha \in \mathbb{R}^n : \Sigma\alpha_i = 1,\ 0 \leq \alpha_i \leq 1\}$. Distinguish between $\alpha \in \text{ri } C$ and $\alpha \in \text{rbd } C$.

Part II

CONVEX FUNCTIONS

Chapter 6

DEFINITIONS AND FIRST EXAMPLES

6.1 Classical and Modern Definition of Convex Functions

Definition 28 (Classical Version) *Let $C \subset \mathbb{R}^n$, $C \neq \emptyset$, C convex. A function $f : C \to \mathbb{R}$ is called* **convex** *on C if for all $x, x' \in C$ and all $\alpha \in (0,1)$*

$$f(\alpha x + (1-\alpha)x') \leq \alpha f(x) + (1-\alpha) f(x')$$

holds.
A function $f : C \to \mathbb{R}$ is called **strictly convex** *on C if for all $x, x' \in C$ and all $\alpha \in (0,1)$*

$$f(\alpha x + (1-\alpha)x') < \alpha f(x) + (1-\alpha) f(x')$$

holds.
A function $f : C \to \mathbb{R}$ is called **strongly convex** *on C (with modulus c) if there exists $c > 0$ such that for all $x, x' \in C$ and all $\alpha \in (0,1)$*

$$f(\alpha x + (1-\alpha)x') \leq \alpha f(x) + (1-\alpha) f(x') - \frac{1}{2} c\alpha(1-\alpha)\|x - x'\|^2$$

holds.

Proposition 6.1 *f is strongly convex on C with modulus $c \iff f - \frac{1}{2}c\|\cdot\|$ is convex on C.*

Proof: Follows from

$$\begin{aligned} f(\alpha x+(1-\alpha)x') &- \frac{1}{2}c\|\alpha x+(1-\alpha)x'\|^2 \\ &\leq \alpha f(x)+(1-\alpha)f(x')-\frac{1}{2}c(\alpha\|x\|^2+(1-\alpha)\|x'\|^2) \\ \Leftrightarrow f(\alpha x+(1-\alpha)x') &\leq \alpha f(x)+(1-\alpha)f(x')-\frac{1}{2}c\alpha(1-\alpha)\|x-x'\|^2. \end{aligned}$$

□

Definition 29 (Modern Version) *A function $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$, not identically $+\infty$, is* **convex** *if for all $x,x'\in\mathbb{R}^n$ and all $\alpha\in(0,1)$*

$$f(\alpha x+(1-\alpha)x')\leq \alpha f(x)+(1-\alpha)f(x')$$

holds (in $\mathbb{R}\cup\{+\infty\}$). The class of convex functions is denoted by Conv $\mathbb{R}^n$.

Remark

- To convert a convex function $f:C\to\mathbb{R}$ from the classical definition to a convex function $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ we just define $f(x):=+\infty$ for all $x\notin C$.
- We can define modern versions of strictly convex functions and strongly convex functions analogously.

Definition 30 *The* **domain** *of $f\in$ Conv $\mathbb{R}^n$ is defined as*

$$\operatorname{dom} f:=\{x\in\mathbb{R}^n : f(x)<+\infty\}.$$

If $f\in$ Conv $\mathbb{R}^n \Longrightarrow \operatorname{dom} f\subset\mathbb{R}^n$ is convex.
Therefore, we get the classical version back by defining $f:\operatorname{dom} f\to\mathbb{R}$. For a function f, the **graph** of f (gr f) is the set of couples $(x,f(x))$.

Definition 31 *Given $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$, not identically $+\infty$, the* **epigraph of f** *is the nonempty set* $\operatorname{epi} f:=\{(x,r)\in\mathbb{R}^n\times\mathbb{R} : r\geq f(x)\}$. *The* **strict epigraph** *is defined as* $\operatorname{epi}_s f:=\{(x,r)\in\mathbb{R}^n\times\mathbb{R} : r>f(x)\}$.

In terms of sublevel sets we can write

$$(x,r)\in\operatorname{epi} f\Longleftrightarrow x\in S_r(f)\ (=\{x\in\mathbb{R}^n : f(x)\leq r\}).$$

Proposition 6.2 *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be not identically $+\infty$. Then the following three properties are equivalent:*

i) $f \in \mathrm{Conv}\,\mathbb{R}^n$.

ii) $\mathrm{epi}\, f \subset \mathbb{R}^n \times \mathbb{R}$ *is a convex set.*

iii) $\mathrm{epi_s}\, f \subset \mathbb{R}^n \times \mathbb{R}$ *is a convex set.*

Proof: "$i) \Longrightarrow ii)$": Let $(x_1, r_1), (x_2, r_2) \in \mathrm{epi}\, f$ and let $\alpha \in (0,1)$.

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) &\leq \alpha f(x_1) + (1-\alpha) f(x_2) \\ &\leq \alpha r_1 + (1-\alpha) r_2 \end{aligned}$$

$\Rightarrow (\alpha x_1 + (1-\alpha)x_2,\ \alpha r_1 + (1-\alpha)r_2) \in \mathrm{epi}\, f$.
"$ii) \Longrightarrow i)$": Let $(x_1, r_r),\ (x_2, r_2) \in \mathrm{epi}\, f,\ \alpha \in (0, 1)$.

$$\begin{aligned} &\Rightarrow (\alpha x_1 + (1-\alpha)x_2,\ \alpha r_1 + (1-\alpha)r_2) \in \mathrm{epi}\, f \\ &\Rightarrow f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha r_1 + (1-\alpha) r_2. \end{aligned}$$

With $r_1 = f(x_1),\ r_2 = f(x_2)$ the result follows.
"$i) \Longrightarrow iii)$": Analogous to "$i) \Longrightarrow ii)$" replacing $\leq$ by $<$.
"$iii) \Longrightarrow i)$": Analogous as "$ii) \Longrightarrow i)$". Choosing $r_i = f(x_i) + \epsilon,\ \epsilon > 0,\ i = 1, 2$, we obtain

$$f(\alpha x_1 + (1-\alpha)x_2) < \alpha f(x_1) + (1-\alpha) f(x_2) + \epsilon$$

$\Rightarrow f$ is convex, since ϵ is arbitrary.

□

Definition 32 *A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ not identically $+\infty$ is* **concave** *if $-f \in \mathrm{Conv}\,\mathbb{R}^n$.*

Remark Geometrically, this means that the **hypograph** of f ($\{(x, r) \in \mathbb{R}^n \times \mathbb{R} : r \leq f(x)\}$ is a convex set. Figure 6.1 shows the graph, epigraph and sublevel set of a function.

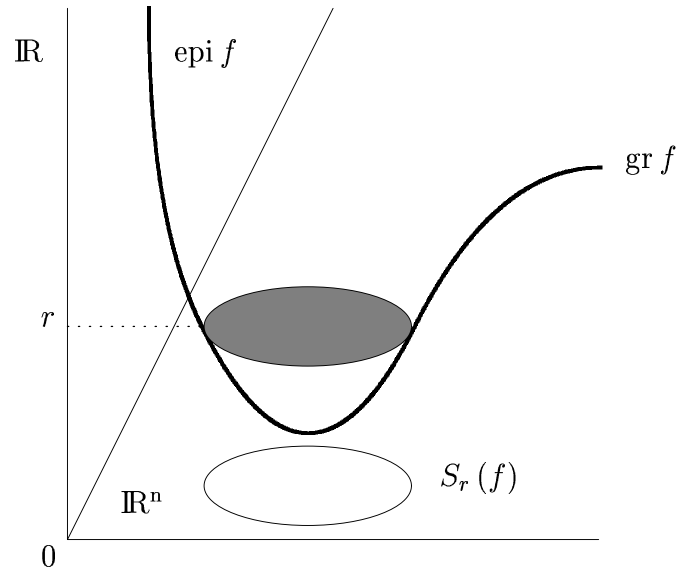


Figure 6.1: Illustration of graph, epigraph and sublevel set of a function.

Remark

- If $f \in \text{Conv } \mathbb{R}^n \Longrightarrow S_r(f)$ is convex because it can be viewed as intersection with epi f and projection onto $\mathbb{R}^n$.
  $S_r(f) = p_{\mathbb{R}^n}(\text{epi } f \cup (\mathbb{R}^n \times \{r\}))$.
- $S_r(f)$ convex $\not\Rightarrow f \in \text{Conv } \mathbb{R}^n$.
- $S_r(f)$ convex $\Rightarrow f$ is quasi-convex.

Theorem 6.3 *Let* $f \in \text{Conv } \mathbb{R}^n$. *Then for all* $x_1, \dots, x_k \in \text{dom } f$ *and all* $\alpha = (\alpha_1, \dots, \alpha_k) \in \Delta_k$ *we have* $f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leq \sum_{i=1}^{k} \alpha_i f(x_i)$.

Proof: By induction over k.

- $k = 2$: The inequality follows from the definition of convexity.
- $k-1 \to k$: If $\alpha_k = 0$ or $\alpha_k = 1$ there is nothing to show. Assume all $\alpha_k \in (0,1)$. Set $\bar{\alpha} = \sum_{i=1}^{k-1} \alpha_i$, $\bar{\alpha} \in (0,1)$ (so $\alpha_k = 1 - \bar{\alpha} \in (0,1)$) and $\bar{\alpha}_i := \frac{\alpha_i}{\bar{\alpha}}$, $i = 1, \dots, k-1$.

$$
\begin{aligned}
&\Rightarrow \bar{\alpha}_i \geq 0, \sum_{i=1}^{k-1} \bar{\alpha}_i = 1 \\
&\Rightarrow \sum_{i=1}^{k} \alpha_i x_i = \bar{\alpha} \sum_{i=1}^{k-1} \bar{\alpha}_i x_i + (1-\bar{\alpha}) x_k \\
&\Rightarrow f\left(\sum_{i=1}^{k} \alpha_i x_i\right) = f\left(\bar{\alpha} \sum_{i=1}^{k-1} \bar{\alpha}_i x_i + (1-\bar{\alpha}) x_k\right) \\
&\qquad \leq \bar{\alpha} f\left(\sum_{i=1}^{k-1} \bar{\alpha}_i (x_i)\right) + (1-\bar{\alpha}) f(x_k) \\
&\qquad \leq \sum_{i=1}^{k-1} \bar{\alpha} \bar{\alpha}_i f(x_i) + \alpha_k f(x_k) = \sum_{i=1}^{k} \alpha_i f(x_i).
\end{aligned}
$$

□

Proposition 6.4 *Let* $f \in \text{Conv } \mathbb{R}^n$. *Then* ri epi $f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : x \in \text{ri dom } f, r > f(x)\}$.

Proof: Since dom f is the projection of epi f onto $\mathbb{R}^n$, Proposition 1.1 and Proposition 2.6

$$\Rightarrow \text{ri dom } f \text{ is the projection of ri epi } f \text{ onto } \mathbb{R}^n. \tag{6.1}$$

Take $x \in \text{ri dom } f$. The preimage of x w.r.t the projection is

$$\begin{aligned} (\{x\} \times \mathbb{R}) \cap \text{ri epi } f &= \text{ri } (\{x\} \times \mathbb{R} \cap \text{epi } f) \text{ by Proposition 2.4} \\ &= \{(x,r) : r \in (f(x), +\infty)\} \end{aligned}$$

$\Rightarrow$ For $x \in \text{ri dom } f$, $(x,r) \in \text{ri epi } f \Leftrightarrow r > f(x)$. Together with (6.1), the result follows.

□

Remark In general $\text{epi}_s f \neq \text{ri epi } f$.

6.2 Special Convex Functions

6.2.1 Linear and Affine Functions

A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is called **linear** if

$$\text{epi } f = \{(x,r) \in \mathbb{R}^n \times \mathbb{R} : \langle s, x\rangle - r \leq 0\}.$$

epi f is characterized by $s \in \mathbb{R}^n$.
A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is called **affine** if

$$\begin{aligned} \text{epi } f &= \{(x,r) : r \geq f(x_0) + \langle s, x - x_0\rangle\} \\ &= \{(x,r) : \langle s, x\rangle - r \leq \langle s, x_0\rangle - f(x_0)\}. \end{aligned}$$

epi f is characterized by a constant term and a vector $(s, -1) \in \mathbb{R}^n \times \mathbb{R}$.

Proposition 6.5 *Any $f \in \text{Conv } \mathbb{R}^n$ is minorized by some affine function. More precisely, for any $x_0 \in \text{ri dom } f$, there is a $s \in \mathbb{R}^n$ in the subspace parallel to* aff dom f *such that $f(x) \geq f(x_0) + \langle s, x - x_0\rangle$ for all $x \in \mathbb{R}^n$.*

Proof: dom f is the projection of epi f onto $\mathbb{R}^n$ $\Rightarrow$ aff (epi f) = aff (dom f)$\times\mathbb{R}$. Let V be the subspace parallel to aff (dom f) $\Rightarrow$ aff (dom f) $= \{x_0\} + V$ with $x_0 \in \text{dom } f$ $\Rightarrow$ aff (epi f) $= \{x_0 + V\} \times \mathbb{R}$. Now equip $V \times \mathbb{R}$ and $\mathbb{R}^n \times \mathbb{R}$ with the scalar product of the product spaces. Chose $x_0 \in \text{ri dom } f$. Proposition 6.4 $\Rightarrow$ $(x_0, f(x_0)) \in \text{rbd epi } f$. Section 4.2 $\Rightarrow$

we can find a nontrivial hyperplane supporting epi f at $(x_0, f(x_0))$ and there are $s = s_v \in V$ and $\alpha \in \mathbb{R}$, $(s, \alpha) \neq 0$ such that

$$\langle s, x\rangle + \alpha r \leq \langle s, x_0\rangle + \alpha f(x_0) \text{ for all } (x, r) \text{ with } f(x) \leq r. \tag{6.2}$$

Since $s \in V$ and $x_0 \in$ ri dom f we can find $\delta > 0$ such that $x_0+\delta s \in$ dom f for which Inequality (6.2) gives $\delta\|s\|^2 \leq \alpha\,(f(x_0) - f(x_0 + \delta s)) < \infty \Rightarrow \alpha \neq 0$ (otherwise $s, \alpha = 0$). *wlog* assume $\alpha = -1 \Rightarrow$ Inequality (6.2) gives the affine function we are looking for.

□

6.2.2 Closed Convex Functions

Definition 33 *A function $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ is* **lower semi-continuous** *(l.s.c.) (on $\mathbb{R}^n$) if for each $x \in \mathbb{R}^n$*

$$\liminf_{y\to x} f(y) \geq f(x). \tag{6.3}$$

Proposition 6.6 *For $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, the following three properties are equivalent:*

i) f is l.s.c on $\mathbb{R}^n$.

ii) epi f *is a closed set on $\mathbb{R}^n \times \mathbb{R}$.*

iii) $S_r(f)$ is closed (possibly empty) for all $r \in \mathbb{R}$.

Proof: "$i) \Longrightarrow ii)$": Let $\{(y_k, r_k)\}$ be a sequence in epi f converging to (x, r) for $k \to \infty$. Since $f(y_k) \leq r_k$ for all k, Inequality (6.3)
$\Rightarrow r := \lim r_k \geq \liminf f(y_k) \geq \liminf_{y\to x} f(y) \geq f(x) \Rightarrow (x, r) \in$ epi f.
"$ii) \Longrightarrow iii)$": Construct $S_r(f)$ as in Remark to Proposition 6.2.
$\Rightarrow S_r(f) = p_{\mathbb{R}^n}$ (epi $f \cap (\mathbb{R}^n \times \{r\})$) is closed.
"$iii) \Longrightarrow i)$": Suppose f is not l.s.c. for some $x \Rightarrow \exists$ (sub)sequence $\{y_k\} \to x$ such that $f(y_k) \to \rho < f(x) \leq +\infty$. Take $r \in (\rho, f(x))$. For k large enough $f(y_k) \leq r < f(x) \Rightarrow S_r(f)$ contains $\{y_k\}$ but not the limit $x \Rightarrow S_r(f)$ is not closed.

□

Definition 34 $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is* **closed** *if f is l.s.c. everywhere.*

Definition 35 *The* **closure** *(or l.s.c. hull) of* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is* $\text{cl}\, f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$, $\text{cl}\, f(x) := \liminf_{y \to x} f(y)$ *for all* $x \in \mathbb{R}^n$*, or equivalently* $\text{epi}\,(\text{cl}\, f) := \text{cl}\,(\text{epi}\, f)$.

With convexity, f is minorized by an affine function (Proposition 6.5) $\Rightarrow -\infty$ is not needed.

Proposition 6.7 (Radial or Inner Construction) *Let* $f \in \text{Conv}\, \mathbb{R}^n$, $x' \in \text{ri dom}\, f$. *Then* $(\text{cl}\, f)(x) = \lim_{t \searrow 0} f\,(x + t(x' - x))$ *for all* $x \in \mathbb{R}^n$.

Proof: For $x_t := x + t(x' - x) \to x$ for $t \searrow 0 \Rightarrow (\text{cl}\, f)(x) \leq \liminf_{t \searrow 0} f(x + t(x' - x))$. We will now show that $\limsup_{t \searrow 0} f(x + t(x' - x)) \leq r$ for all $r \geq (\text{cl}\, f)(x)$ (If no such r exists $\Rightarrow \text{cl}\, f(x) = +\infty$). Let $(x, r) \in \text{epi}\,(\text{cl}\, f) = \text{cl}\,(\text{epi}\, f)$. Pick $r' > f(x')$. Proposition 6.4 $\Rightarrow (x', r') \in \text{ri epi}\, f$. Apply Lemma 2.2 to epi f to get

$$t(x', r') + (1 - t)(x, r) \in \text{ri epi}\, f \subset \text{epi}\, f \quad \text{for all } t \in [0, 1]$$

$\Rightarrow f(x + t(x' - x)) \leq tr' + (1 - t)r$ for all $t \in (0, 1]$. For $t \searrow 0$ we get $\limsup_{t \searrow 0} f(x + t(x' - x)) \leq r$.

$\square$

Proposition 6.8 *For* $f \in \text{Conv}\, \mathbb{R}^n$ *we have*

i) $\text{cl}\, f \in \text{Conv}\, \mathbb{R}^n$.

ii) $\text{cl}\, f = f$ *on* $\text{ri dom}\, f$.

Proof: "*i)*": From Proposition 1.2 $\Rightarrow$ epi cl f = cl epi f is a convex set. Also $\text{cl}\, f \leq f \not\equiv +\infty$. Proposition 6.5 guarantees that $\text{cl}\, f(x) > -\infty$ for all $x \Rightarrow \text{cl}\, f \in \text{Conv}\, \mathbb{R}^n$.
"*ii)*": Take $x \in \text{ri dom}\, f \Rightarrow \phi(t) = f(x + td)$ is continuous at $t = 0 \Rightarrow \text{cl}\, f = f$ on ri dom f. Moreover, $\text{cl}\, f(x) \equiv +\infty$ for $x \notin \text{cl dom}\, f$.

$\square$

The set of all closed convex functions is denoted by $\overline{\text{Conv}}\, \mathbb{R}^n$.

Proposition 6.9 (Outer Construction of $\operatorname{cl} f$**)** *Let* $f \in \operatorname{Conv} \mathbb{R}^n$*. Then*

$$\operatorname{cl} f(x) = \sup_{(s,b) \in \mathbb{R}^n \times \mathbb{R}} \{\langle s, x\rangle - b : \langle s, y\rangle - b \le f(y) \textit{ for all } y \in \mathbb{R}^n\}.$$

(Refer to Figure 6.2).

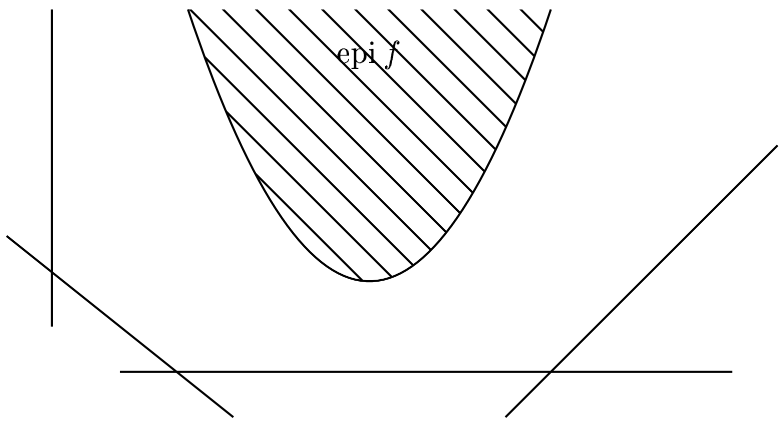


Figure 6.2: Illustration for Proposition 6.9.

Proof: A closed halfspace containing $\operatorname{epi} f$ is characterized by $0 \neq (s, \alpha) \in \mathbb{R}^n \times \mathbb{R}$, $b \in \mathbb{R}$:

$$\langle s, x\rangle + \alpha r \le b \text{ for all } (x, r) \in \operatorname{epi} f. \tag{6.4}$$

Denote by $\Sigma \subset \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ all triples $\sigma = (s, \alpha, b)$ fulfilling Inequality (6.4) and by $H_\sigma^{\le} := \{(x, r) : \langle s, x\rangle + \alpha r \le b\}$ the corresponding halfspaces. From Proposition 6.6 $\Rightarrow$ $\operatorname{epi}(\operatorname{cl} f) = \operatorname{cl}(\operatorname{epi} f) = \bigcap_{\sigma \in \Sigma} H_\sigma^{\le}$, by Theorem 4.5. Note that by Inequality (6.4) $\alpha \le 0$ and by positive homogeneity $\Rightarrow$ $\alpha = 0$ or $\alpha = -1$ is enough to consider. Partition

$$\Sigma_1 := \{(s, -1, b) : \text{ Inequality (6.4) holds with } \alpha = -1\} \text{ and}$$
$$\Sigma_0 := \{(s, 0, b) : \text{ Inequality (6.4) holds with } \alpha = 0\}.$$

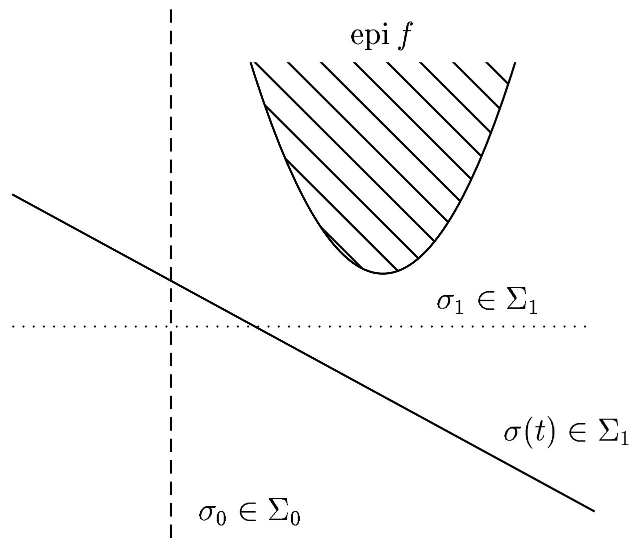


Figure 6.3: Illustration for proof of Proposition 6.9.

Σ_1 corresponds to affine functions minorizing f (Proposition 6.5 $\Rightarrow \Sigma_1 \neq \emptyset$) and Σ_0 corresponds to closed halfspaces in $\mathbb{R}^n$ containing dom f ($\Sigma_0 = \emptyset$ if dom $f = \mathbb{R}^n$). Now we have to show that $\bigcap_{\sigma\in\Sigma} H_\sigma^{\leq} = \bigcap_{\sigma\in\Sigma_1} H_\sigma^{\leq} = \text{cl}\,(\text{epi}\, f)$ even if $\Sigma_0 \neq \emptyset$. Take $\sigma_0 = (s_0, 0, b_0) \in \Sigma_0$, and $\sigma_1 = (s_1, -1, b_1) \in \Sigma_1$. Set $\sigma(t) := (s_1 + ts_0, -1, b_1 + tb_0) \in \Sigma_1$ for all $t \geq 0$. Now we prove that $H_{\sigma_0}^{\leq} \cap H_{\sigma_1}^{\leq} = \bigcap_{t\geq 0} H_{\sigma_t}^{\leq} =: H^{\leq}$ ($\Rightarrow \bigcap_{\sigma\in\Sigma} H_\sigma^{\leq} = \bigcap_{\sigma\in\Sigma_1} H_\sigma^{\leq}$). See Figure 6.3 for an illustration.

- "$H_{\sigma_0}^{\leq} \cap H_{\sigma_1}^{\leq} \subset H^{\leq}$": From the definition of $H_\sigma^{\leq}$ we have for $(x, r) \in H_{\sigma_0}^{\leq} \cap H_{\sigma_1}^{\leq}$

$$\langle s_1 + ts_0, x\rangle - (b_1 + tb_0) \leq r \text{ for all } t \geq 0 \tag{6.5}$$

$\Rightarrow (x, r) \in H_{\leq}$.

- "$H^{\leq} \subset H^{\leq}_{\sigma_0} \cap H^{\leq}_{\sigma_1}$": Take $(x,r) \in H^{\leq}$. Set $t = 0$ in Inequality (6.5) $\Rightarrow$ $(x,r) \in H^{\leq}_{\sigma_1}$. Divide by $t > 0$ and take $t \to +\infty \Rightarrow (x,r) \in H^{\leq}_{\sigma_0}$.

□

6.3 Examples

6.3.1 Indicator and Support Functions

Definition 36 *Let $S \subset \mathbb{R}^n$, $S \neq \emptyset$. The* **indicator function of *S*** I_S : $\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is defined by*

$$I_S(x) := \begin{cases} 0 & \text{if } x \in S \\ +\infty & \text{otherwise} \end{cases}$$

- I_S is (closed and) convex $\iff S$ is (closed and) convex. Clear, since epi $I_S = S \times \mathbb{R}_{+_0}$.

- $\sigma_S(x) := \sup\{\langle s, x\rangle : s \in S\}$ is closed and convex. σ_S is called the **support function** of S. For $\alpha > 0$, $\sup_{s \in S}\langle s, \alpha x\rangle = \alpha \sup_{s \in S}\langle s, x\rangle \Rightarrow \sigma(\alpha x)$ $= \alpha\sigma(x) \Rightarrow$ epi σ_S is a cone in $\mathbb{R}^n \times \mathbb{R}$. Also dom $\sigma_S = \{a \in \mathbb{R}^n\text{: } \exists\, r$ such that $\langle s, a\rangle \leq r$ for all $s \in S\}$ is a convex cone in $\mathbb{R}^n$.

6.3.2 Piecewise Affine and Polyhedral Functions

Consider for $x \in \mathbb{R}^n$, $x \mapsto \check{f}(x) := \max\{\langle s_j, x\rangle - b_j : j = 1, \ldots, m\}$ where $(s_1, b_1), \ldots, (s_m, b_m) \in \mathbb{R}^n \times \mathbb{R}$. The function $\check{f}$ is called **piecewise affine**. $\mathbb{R}^n$ is divided into ($\leq m$) regions in which $\check{f}$ is affine. The region $k \in \{1, \ldots, m\}$, possibly empty, is the closed convex polyhedron $\{x \in \mathbb{R}^n : \langle s_k, x\rangle - b_k \geq \langle s_j, x\rangle - b_j, j = 1, \ldots, m\}$. It is easy to see that epi f is a closed convex polyhedron. With $\check{f}$, we cannot describe all polyhedral epigraphs. Define a polyhedral function as a function whose epigraph is a closed convex polyhedron. epi $f = \{(x,r) \in \mathbb{R}^n \times \mathbb{R} : \langle s_j, x\rangle + \alpha_j r \leq b_j$ for $j \in \mathcal{J}\}$ where $\mathcal{J}$ is a finite set, $(s, \alpha, b)_j \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$. For an epigraph $\Rightarrow \alpha_j \leq 0$ ($\alpha_j = 0$ or $\alpha_j = -1$). Let us for $\mathcal{J} = \{1, \ldots, m + p\}$ denote by $\{1, \ldots, m\} \subset \mathcal{J}$ the indices with $\alpha_j = -1$ and $\mathcal{J} \setminus \{1, \ldots, m\}$ the indices with $\alpha_j = 0 \Rightarrow$ $f(x) = \check{f}(x)$ if $\langle s_j, x\rangle \leq b_j$ for $j = m + 1, \ldots, m + p$, otherwise $f(x) = +\infty$. $\Rightarrow$ A polyhedral function $f = \check{f} + I_P$, where P is a closed convex polyhedron.

6.3.3 Norms and Distances

We consider $\emptyset \neq C \subset \mathbb{R}^n$, C convex and $|||\cdot|||$ an arbitrary norm. The **distance function** to C is defined by

$$d_C(x) := \inf\{|||\, y - x \,|||: y \in C\}.$$

- $d_C \in \text{Conv}\,\mathbb{R}^n$: Take $\{y_k\}, \{y_k'\} \subset C$ such that for $k \to +\infty$ $|||\, y_k - x \,||| \to d_C(x)$, $|||\, y_k' \,||| \to d_C(x')$. Now form a subsequence $z_k := \alpha y_k + (1-\alpha) y_k' \in C$ with $\alpha \in (0,1)$.

$$\begin{aligned} d_C(\alpha x + (1-\alpha)x') &\leq |||\, z_k - \alpha x - (1-\alpha)x' \,||| \\ &= |||\, \alpha(y_k - x) + (1-\alpha)(y_k' - x') \,||| \\ &\leq \alpha \,|||\, (y_k - x) \,||| + (1-\alpha)\, |||\, y_k' - x' \,||| \\ &\overset{k\to\infty}{=} \alpha d_C(x) + (1-\alpha) d_C(x'). \end{aligned}$$

$\Longrightarrow$ dom $d_C = \mathbb{R}^n$. Proposition 6.8 $\Rightarrow$ d_C is closed and d_c is l.s.c. It is clear that $d_C = d_{\text{cl}\,C}$. Proposition 2.3 $\Rightarrow$ C, cl C, ri C have the same distance function. For all $x \in \text{cl}\,C$ we get $d_C(x) = 0$.

- Modified Definition:

$$D_C(x) := \begin{cases} d_C(x) & \text{if } x \in C^c, C^c := \mathbb{R}^n \setminus C \\ -d_{C^c}(x) & \text{if } x \in C \end{cases}$$

Again (if C, $C^c \neq \emptyset$) D_C is convex, dom $D_C = \mathbb{R}^n$ and

  - int $C = \{x \in \mathbb{R}^n : D_C(x) < 0\}$.
  - bd $C = \{x \in \mathbb{R}^n : D_C(x) = 0\}$.
  - $(\text{cl}\,C)^c = \{x \in \mathbb{R}^n : D_C(x) > 0\}$.

6.3.4 Epigraphical Hull

Let $C \subset \mathbb{R}^n \times \mathbb{R}$, $C \neq \emptyset$, C convex.

Definition 37 *The* **epigraphical hull** *of C is the smallest epigraph containing C.*

We need for a given C to build the epigraphical hull:

- For each $(x, r) \in C$ add to C all (x, r') with $r' > r$.

- Add to C the point (x,r) if $(x,r') \in C$ and $r' \to r$.

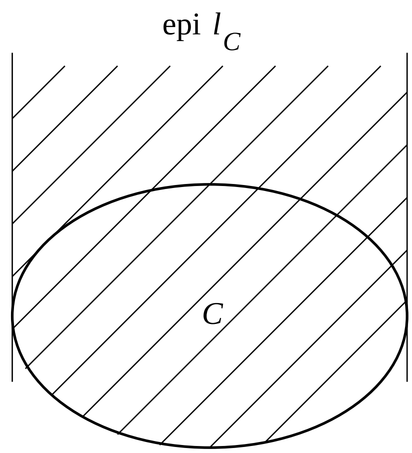


Figure 6.4: Illustration for the epigraphical hull.

Definition 38 *The* **lower-bound function** *of C is defined as* $x \in \mathbb{R}^n \mapsto l_C(x) := \inf\{r \in \mathbb{R} : (x,r) \in C\}$.

Then epi l_c is the epigraphical hull of C. $l_C(x) > -\infty$ for all x if $\{r \in \mathbb{R} : (x,r) \in C\}$ is minorized for all $x \in \mathbb{R}^n$.
We have: $\text{epi}_s\, l_c \subset C + \{0\} \times \mathbb{R}_{+0} \subset \text{epi}\, l_C \subset \text{cl}\,(C + \{0\} \times \mathbb{R}_{+0})$.

Theorem 6.10 *Let $C \subset \mathbb{R}^n \times \mathbb{R}$, $C \neq \emptyset$ and $\{r \in \mathbb{R} : (x,r) \in C\}$ be minorized for all $x \in \mathbb{R}^n$. Then*

i) If C is convex then $l_C \in \text{Conv}\,\mathbb{R}^n$.

ii) If C is closed, convex then $l_C \in \overline{\text{Conv}}\,\mathbb{R}^n$.

Proof: "i)": Take $\epsilon > 0, \alpha \in (0,1)$ and $(x_i, r_i) \in C$, $i = 1,2$ such that $r_i \leq l_C(x_i) + \epsilon$ for $i = 1,2$. Since C is convex $(\alpha x_1 + (1-\alpha)x_2, \alpha r_1 + (1-\alpha)r_2) \in C$ $\Rightarrow l_C(\alpha x_1 + (1-\alpha)x_2) \leq \alpha r_1 + (1-\alpha)r_2 \leq \alpha l_C(x_1) + (1-\alpha)l_C(x_2) + \epsilon$. Since $\epsilon > 0$ is arbitrary i) holds.
"ii)": Take a sequence $\{(x_k, \rho_k)\} \subset \text{epi}\, l_C$ converging to (x, ρ). We show that epi l_C is closed ($l_C(x) \leq \rho$). Proposition 6.6 $\Rightarrow l_C \in \overline{\text{Conv}}\,\mathbb{R}^n$. By definition of $l_C(x_k)$, we can select for each $k \in \mathbb{N}$, $r_k \in \mathbb{R}$ such that $(x_k, r_k) \in C$ and

$$l_C(x_k) \leq r_k \leq l_C(x_k) + \frac{1}{k} \leq \rho_k + \frac{1}{k} \tag{6.6}$$

$\Rightarrow \{r_k\}$ is bounded from above. With Proposition 6.8 we have $\{r_k\}$ is bounded from below. $\Rightarrow$ By eventually extracting a subsequence we may assume $r_k \to r$. C closed $\Rightarrow (x,r) \in C \Rightarrow l_C(x) \leq r$. With $k \to +\infty$ in Inequality (6.6) $\Rightarrow r \leq \rho$.

□

6.4 Exercises

1. Let $0 < x_i < \infty$ and $0 < t_i < 1 \quad i = 1, \dots, n \; \Sigma t_i = 1$.

   Then

$$\prod_{i=1}^{n} x_i^{t_i} \leq \sum_{i=1}^{n} t_i x_i$$

$$\left(\sum_{i=1}^{n} t_i x_i \right)^{\sum\limits_{i=1}^{n} t_i x_i} \leq \prod_{i=1}^{n} x_i^{t_i x_i}$$

2. A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is called quasi-convex if

$$f(\alpha x + (1-\alpha)x') \leq \max\{f(x),\; f(x')\}$$

   $\forall\, \alpha \in (0,1)$.

   Prove that f is quasi convex if and only if $S_r(f)$ is convex for all $r \in \mathbb{R}$.

3. Let $f = I \to \mathbb{R} \cup \{+\infty\}$, $I \subset \mathbb{R}$ be an interval. Prove that f is convex if and only if

$$s(x) = \frac{f(x) - f(x_0)}{x - x_0} \quad \text{is}$$

   increasing on $I \backslash \{x_0\}$.

4. A convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ not identically $+\infty$ is sometimes called properly convex.

   A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty, -\infty\}$ is improperly convex, if it is convex, but not properly convex.

   Prove that:

a) If f is improperly convex then $f(x) = -\infty \;\; \forall\, x \in \operatorname{ri} \operatorname{dom} f$.

b) If f is closed then $\operatorname{dom} f$ is closed and $f(x) = -\infty \; \forall\, x \in \operatorname{dom} f$.

5. Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be convex.
   Let $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be defind by

$$g(x) = \inf \; \{f(x, y) : y \in \mathbb{R}^m\}.$$

   Prove that g is convex.

6. $f : \mathbb{R}^n \to \mathbb{R}$ is logarithmically convex if $\log f$ is convex. Prove that a log-convex function is convex.
   Give an example of a log-convex function and of a convex function which is not log-convex.

Chapter 7

FUNCTIONAL OPERATIONS PRESERVING CONVEXITY

7.1 Basic Operations

Proposition 7.1 *Let* $f_1, \dots, f_m \in \text{Conv } \mathbb{R}^n$ *(resp.* $\overline{\text{Conv}}\ \mathbb{R}^n$*),* $t_1, \dots, t_m$ *be positive numbers and let exist* $x_0 \in \mathbb{R}^n$ *with* $f_i(x_0) < +\infty$*,* $i = 1, \dots, m$*. Then* $f := \sum_{j=1}^{m} t_j f_j \in \text{Conv } \mathbb{R}^n$ *(resp.* $\overline{\text{Conv}}\ \mathbb{R}^n$*).*

Proof: Convexity follows directly from the definition. For closedness

$$\liminf_{y \to x} t_j f_j(y) = t_j \liminf_{y \to x} f_j(y) \geq t_j f_j(x),$$

as f_j are closed, $t_j > 0$. The result follows from

$$\sum_{j=1}^{m} \liminf_{y \to x} t_j f_j(y) \leq \liminf_{y \to x} \sum_{j=1}^{m} t_j f_j(y).$$

□

Proposition 7.2 *Let* $\{f_j\}_{j \in \mathcal{J}}$*,* $f_j \in \text{Conv } \mathbb{R}^n$ *(resp.* $\overline{\text{Conv}}\ \mathbb{R}^n$*). If there exists* x_0 *with* $\sup_{j \in \mathcal{J}} f_j(x_0) < +\infty$ *then*

$$f := \sup\{f_j : j \in \mathcal{J}\} \in \text{Conv } \mathbb{R}^n \ (\overline{\text{Conv}}\ \mathbb{R}^n).$$

Proof: Follows from $\text{epi } f = \bigcap_{j \in \mathcal{J}} \text{epi } f_j$ and by assumption $\bigcap_{j \in \mathcal{J}} \text{epi } f_j \neq \emptyset$ $\Rightarrow$ epi f is closed, convex.

□

Example 7.1 ***Definition 39*** *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a function not identically* $+\infty$ *and minorized by an affine function (For some* $(s_0, b) \in \mathbb{R}^n \times \mathbb{R}$*:* $f(\cdot) \geq \langle s_0, \cdot\rangle - b$ *on* $\mathbb{R}^n$*). Then*

$$f^* : \mathbb{R}^n \ni s \mapsto \sup\{\langle s, x\rangle - f(x) : x \in \mathrm{dom}\, f\}$$

is called the **conjugate function** *of* f.

- *Since* $f^*(s_0) \leq b$ *and* $f^*(s) > -\infty$ *for all* s *because* $\mathrm{dom}\, f \neq \emptyset$*, we have* $f^* \in \overline{\mathrm{Conv}}\ \mathbb{R}^n$ *(as a supremum of closed convex functions). True for any* f.
- *Let* $S \subset \mathbb{R}^n$*,* $S \neq \emptyset$*,* $\mathbb{R}^n \ni s \mapsto \phi_S(x) := \frac{1}{2}(\|x\|^2 - d_S^2(x))$*. Then* ϕ_S *is always convex:*

$$\begin{aligned} d_S^2(x) &= \inf_{c\in S} \|x - c\|^2 \\ &= \|x\|^2 - \sup_{c\in S}(2\langle c, x\rangle - \|c\|^2) \end{aligned}$$

$\Rightarrow \phi_S(x) = \sup\{\langle s, x\rangle - \frac{1}{2}\|x\|^2 : c \in S\}$.
Also note that $\phi_S = (\frac{1}{2}\|\cdot\|^2 + I_S)^*$.

Proposition 7.3 *Let* $f \in \mathrm{Conv}\ \mathbb{R}^n$ *(resp.* $\overline{\mathrm{Conv}}\ \mathbb{R}^n$*),* A *be an affine mapping from* $\mathbb{R}^m \to \mathbb{R}^n$ *such that* $(\mathrm{Im}\ A) \cap (\mathrm{dom}\, f) \neq \emptyset$*. Then the function*

$$f \circ A : \mathbb{R}^m \ni x \mapsto (f \circ A)(x) = f(A(x)) \in \mathrm{Conv}\ \mathbb{R}^m \text{ (resp. } \overline{\mathrm{Conv}}\ \mathbb{R}^m\text{).}$$

Proof: It is clear that $(f \circ A)(x) > -\infty$ for all x, and from $(\mathrm{Im}\ A) \cap (\mathrm{dom}\, f) \neq \emptyset$ there exists $y = A(x) \in \mathbb{R}^m : f(y) < +\infty$.
For convexity:

$$\begin{aligned} A(\alpha x + (1-\alpha)x') &= \alpha A(x) + (1-\alpha)A(x') \\ \Rightarrow (f \circ A)(\alpha x + (1-\alpha)x') &= f(\alpha A(x) + (1-\alpha)A(x')) \\ &\overset{f\in \mathrm{Conv}\ \mathbb{R}^n}{\leq} \alpha f(A(x)) + (1-\alpha)f(A(x)'), \alpha \in [0,1]. \end{aligned}$$

Since A is continuous $\Rightarrow f \circ A \in \overline{\mathrm{Conv}}\ \mathbb{R}^m$ if $f \in \overline{\mathrm{Conv}}\ \mathbb{R}^n$.

□

Proposition 7.4 *Let* $f \in \mathrm{Conv}\ \mathbb{R}^n$ *(resp.* $\overline{\mathrm{Conv}}\ \mathbb{R}^n$*) and let* $g \in \mathrm{Conv}\ \mathbb{R}^n$ *(resp.* $\overline{\mathrm{Conv}}\ \mathbb{R}^n$*) be increasing. Assume that there exists* $x_0 \in \mathbb{R}^n$ *such that* $f(x_0) \in \mathrm{dom}\ g$ *and set* $g(+\infty) := +\infty$*. Then* $g \circ f : x \mapsto g(f(x)) \in \mathrm{Conv}\ \mathbb{R}^n$ *(resp.* $\overline{\mathrm{Conv}}\ \mathbb{R}^n$*).*

Proof: Convexity follows from the definition:

$$\begin{aligned} g \circ f(\alpha x + (1-\alpha)x') &= g(f(\alpha x + (1-\alpha)x')) \\ &\leq g(\alpha f(x) + (1-\alpha) f(x')) \\ &\leq \alpha g(f(x)) + (1-\alpha) g(f(x')), \end{aligned}$$

as g is increasing and convex, respectively.

□

7.2 Infimal Convolution

Let f_1, f_2 be two functions from $\mathbb{R}^n$ to $\mathbb{R} \cup \{+\infty\}$. We form epi f_1 + epi $f_2 \subset \mathbb{R}^n \times \mathbb{R}$:

$$C := \{(x_1 + x_2, r_1 + r_2) : r_j \geq f_j(x_j) \text{ for } j = 1, 2\}.$$

What is $l_C(x)$ and when does it exist?

Definition 40 *Let f_1, f_2 be two functions from $\mathbb{R}^n$ to $\mathbb{R} \cup \{+\infty\}$. Their* **infimal convolution** *is the function from $\mathbb{R}^n$ to $\mathbb{R} \cup \{\pm\infty\}$, defined by*

$$\begin{aligned} (f_1 \overset{+}{\vee} f_2)(x) &:= \inf\{f_1(x_1) + f_2(x_2) : x_1 + x_2 = x\} \\ &= \inf_{y \in \mathbb{R}^n} (f_1(y) + f_2(x - y)). \end{aligned} \tag{7.1}$$

The infimal convolution is called **exact** at $x = \bar{x}_1 + \bar{x}_2$ if the infimum is attained at $(\bar{x}_1, \bar{x}_2)$ (not necessarily unique).

Proposition 7.5 *Let $f_1, f_2 \in$ Conv $\mathbb{R}^n$. Suppose f_1, f_2 have a common affine minorant: $\exists\ (s, b) \in \mathbb{R}^n \times \mathbb{R}$: $f_j(x) \geq \langle s, x\rangle - b, \quad j = 1, 2, x \in \mathbb{R}^n$. Then $f_1 \overset{+}{\vee} f_2 \in$ Conv $\mathbb{R}^n$.*

Proof: Let $x \in \mathbb{R}^n$ and x_1, x_2 such that $x = x_1 + x_2$. By assumption $\Rightarrow f_1(x_1) + f_2(x_2) \geq \langle s, x\rangle - 2b > -\infty$. This inequality is also true for $(f_1 \overset{+}{\vee} f_2)(x)$. Take $x_j \in$ dom f_j, $j = 1, 2 \Rightarrow x_1 + x_2 \in$ dom $(f_1 \overset{+}{\vee} f_2) \Rightarrow$ dom $(f_1 \overset{+}{\vee} f_2) \neq \emptyset$.
For convexity:

$$\begin{aligned} (f_1 \overset{+}{\vee} f_2)(x) &= \inf\{f_1(x_1) + f_2(x_2) : x_1 + x_2 = x\} \\ &= \inf\{r_1 + r_2 : r_1 \geq f_1(x_1),\ r_2 \geq f_2(x_2),\ x_1 + x_2 = x)\} \\ &= l_C(x), \end{aligned}$$

where $C =$ epi f_1 + epi $f_2 \Rightarrow C$ convex. Theorem 6.10 $\Rightarrow l_C \in$ Conv $\mathbb{R}^n \Rightarrow f_1 \overset{+}{\vee} f_2 \in$ Conv $\mathbb{R}^n$.

□

Some Properties of $\overset{+}{\vee}$

- $f_1 \overset{+}{\vee} f_2 = f_2 \overset{+}{\vee} f_1$ (commutativity).
- $(f_1 \overset{+}{\vee} f_2) \overset{+}{\vee} f_3 = f_1 \overset{+}{\vee} (f_2 \overset{+}{\vee} f_3)$ (associativity).
- $f \overset{+}{\vee} I_{\{0\}} = f$ (neutral element).
- $f_1 \leq f_2 \;\Rightarrow\; f_1 \overset{+}{\vee} g \leq f_2 \overset{+}{\vee} g$ ($\overset{+}{\vee}$ preserves the order).

Example 7.2

- *Let* C_1, C_2 *be convex,* $C_i \neq \emptyset$, $i = 1,2$, $C_i \subset \mathbb{R}^n$, $i = 1,2$. *We have* $I_{C_1} \overset{+}{\vee} I_{C_2} = I_{C_1+C_2}$, *clear since*

$$I_C = \begin{cases} 0 & \textit{if } x \in C \\ +\infty & \textit{otherwise} \end{cases}$$

  $\Rightarrow$ *Since the sum of closed sets need not be closed,* $\overset{+}{\vee}$ *need not be closed.*

- $C \subset \mathbb{R}^n$, C *convex,* $C \neq \emptyset$, $|||\cdot|||$ *arbitrary norm. Then* $I_C \overset{+}{\vee} |||\cdot||| = d_C$.

$$\begin{aligned} (I_C \overset{+}{\vee} |||\cdot|||)(x) &= \inf\{I_C(x_1) + |||\, x_2 \,|||,\ x = x_1 + x_2\} \\ &= \inf_{y \in \mathbb{R}^n} [I_C(y) + |||\, x - y \,|||] \\ &= \inf_{y \in C} |||\, x - y \,||| \\ &= d_C(x). \end{aligned}$$

7.3 Image of a Function Under a Linear Mapping

Definition 41 *Let* $A : \mathbb{R}^m \to \mathbb{R}^n$ *be linear,* $g : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$. *The* **image** *of* g *under* A *is the function* $Ag : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ *defined by*

$$(Ag)(x) := \inf\{g(y) : Ay = x\}.$$

Theorem 7.6 *Let $g \in \text{Conv}\, \mathbb{R}^n$. Assume that for all $x \in \mathbb{R}^n$, g is bounded from below on $A^{-1}(x) = \{y \in \mathbb{R}^m : Ay = x\}$. Then $Ag \in \text{Conv}\, \mathbb{R}^n$.*

Proof:

- By assumption $Ag > -\infty$.
- Also $(Ag)(x) < +\infty$, whenever $x = Ay, y \in \text{dom}\, g$.
- Consider $A' : \mathbb{R}^m \times \mathbb{R} \ni (y, r) \mapsto A'(y, r) := (Ay, r) \in \mathbb{R}^m \times \mathbb{R}$. Since A' is linear $\Rightarrow A'(\text{epi}\, g) =: C$ is convex in $\mathbb{R}^n \times \mathbb{R}$. Compute

$$\begin{aligned} l_C(x) &= \inf_{r \in \mathbb{R}} \{r : (x, r) \in C\} \\ &= \inf_{y,r} \{r : Ay = x,\ g(y) \le r\} \\ &= \inf_{y} \{g(y) : Ay = x\} = (Ag)(x) \end{aligned}$$

$\Rightarrow Ag = l_C \Rightarrow Ag$ is convex.

□

Example 7.3 *(Refer to Figure 7.1).*

- *Let $f_1, f_2 \in \text{Conv}\, \mathbb{R}^n$. Define $g \in \text{Conv}\, (\mathbb{R}^n \times \mathbb{R}^n)$ by $g(x_1, x_2) := f_1(x_1) + f_2(x_2)$ and $A : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ by $A(x_1, x_2) := x_1 + x_2$. $\Rightarrow Ag = f_1 \overset{+}{\vee} f_2$. Ag need not preserve closedness since $\overset{+}{\vee}$ does not.*
- *The* **marginal** *function of $g \in \text{Conv}\, (\mathbb{R}^n \times \mathbb{R}^m)$ is $\mathbb{R}^n \ni x \mapsto \gamma(x) := \inf\{g(x, y) : y \in \mathbb{R}^m\}$. Here $\gamma = Ag$, where A is the projection of $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ onto $x \in \mathbb{R}^n$.*

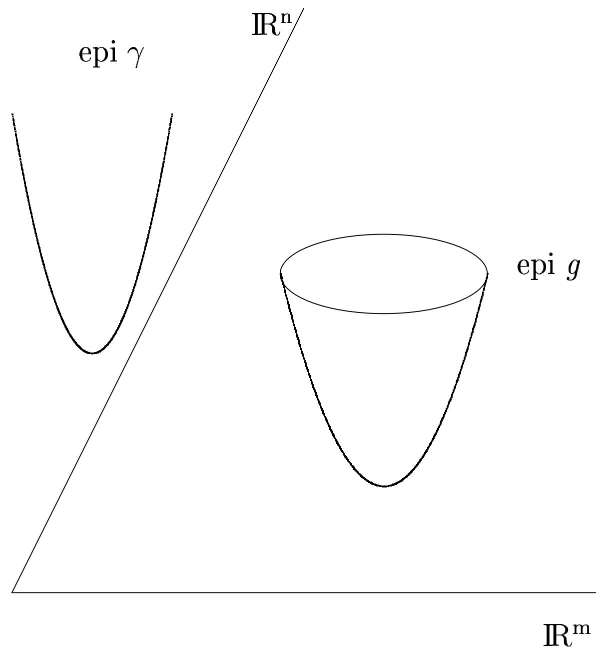


Figure 7.1: Illustration for Example 7.3.

7.4 Exercises

1. a) Let $f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle$, where A is a symmetric positive definite $n \times n$ matrix. Determine $f^*(s)$.

   b) Prove the following rules for conjugate functions

      i) $g(x) = \alpha f(x) \Rightarrow g^*(s) = \alpha f^*\left(\frac{s}{\alpha}\right)$.

      ii) Let A be a $n \times n$ matrix, $det\ A \neq 0$ then $(f \circ A)^* = f^* \circ (A^{-1})^T$.

      iii) $f_1 \leq f_2 \Rightarrow f_1^* \geq f_2^*$.

2. a) Let $f_j(x) = \frac{1}{2}\langle x, A_j x\rangle$, $j = 1, 2$ where A_j are symmetric positive definite $n \times n$ matrices. Find $f_1 \overset{+}{\vee} f_2$.

   b) Prove the following rules for the infimal convolution

      i) $f_1 \overset{+}{\vee} f_2 = f_2 \overset{+}{\vee} f_1$.

      ii) $f_1 \leq f_2 \Rightarrow f_1 \overset{+}{\vee} g \leq f_2 \overset{+}{\vee} g$.

      iii) $(f_1 \overset{+}{\vee} f_2)^* = f_1^* + f_2^*$.

3. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a convex function and $n > 0$. Then $f_u : \mathbb{R}^n \to \mathbb{R}$ defined by $f_u(x) = uf(\frac{x}{u})$ is called the dilation of f. The dilation is convex because epi $f_u = u$ epi f.

   Prove that $\tilde{f}(x,u) = \begin{cases} uf(\frac{x}{u}) & u > 0 \\ +\infty & u \leq 0 \end{cases}$ is in Conv $\mathbb{R}^{n+1}$.

4. Consider $\tilde{f}(x,u)$ of Exercise 3 above. Let $f \in \overline{\text{Conv}}\ \mathbb{R}^n$. Then

$$\text{cl}\,\tilde{f}(x,u) = \begin{cases} uf(\frac{x}{u}) & u > 0 \\ \lim_{\alpha \searrow 0} \alpha f(x' - x + \frac{x}{\alpha}) & u = 0 \\ +\infty & u < 0 \end{cases}$$

5. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ such that $\exists\, x \in \mathbb{R}^n$ with $f(x) < +\infty$ and $\exists\, a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$ s.t. $f(x) \geq \langle a, x \rangle - \beta > -\infty\ \forall\, x \in \mathbb{R}^n$.

   Show that $f^*(s) = \sigma_{\text{epi}\, f}(s, -1)$
   and $\sigma_{\text{epi}\, f}(s, -u) = \begin{cases} uf^*(\frac{s}{u}) & u > 0 \\ \sigma_{\text{dom}\, f}(x) & u = 0 \\ +\infty & u < 0 \end{cases}$

6. Let f be as in 5.

   a) $f^*(s) \geq \langle s, x \rangle - f(x)$

   b) Let $f^{**} = (f^*)^*$. Show that

$$f^{**}(x) = \sup_{(s,r)} \{\langle s, y \rangle - r : f(y) \geq \langle s, y \rangle - r\ \forall\, y \in \mathbb{R}^n\}.$$

   c) f^{**} is a closed function.

   d) $f^{***} = f^*$

7. Let

$$f(x) = \begin{cases} -x & x \leq -2 \\ -x^2 + 6 & -2 \leq x \leq 2 \\ 1 & x \geq 2 \end{cases}$$

   Draw f^{**} using 6 b).

8. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a function. Then the function $\text{co}\, f(x) := \inf\{r : (x, r) \in \text{co epi}\, f\}$ is called the convex hull of f.

   Prove: $f^{**} = \text{cl}\,(\text{co}\, f)$.

Chapter 8

CONTINUITY PROPERTIES OF CONVEX FUNCTIONS

Lemma 8.1 *Let $f \in \text{Conv } \mathbb{R}^n$. If there are x_0, δ, m and M such that*

$$m \leq f(x) \leq M \text{ for all } x \in B(x_0, 2\delta)$$

then f is **Lipschitzian** *on $B(x_0, \delta)$, i.e. for all $y, y' \in B(x_0, \delta)$:*

$$|f(y) - f(y')| \leq \frac{M-m}{\delta}\|y - y'\|. \tag{8.1}$$

Proof: If $y = y'$ then Inequality (8.1) holds. Take $y, y' \in B(x_0, \delta)$, $y \neq y'$, and define $y'' := y' + \delta\frac{y'-y}{\|y'-y\|}$. By construction,

$$y' \in [y, y''], y' = \frac{\|y'-y\|}{\delta + \|y'-y\|}y'' + \frac{\delta}{\delta + \|y'-y\|}y.$$

Refer to Figure 8.1. Now we get

$$\begin{aligned}
f(y') - f(y) &\leq \frac{\|y'-y\|}{\delta + \|y'-y\|}f(y'') + \frac{\delta}{\delta + \|y'-y\|}f(y) - f(y) \\
&= \frac{\|y'-y\|}{\delta + \|y'-y\|}f(y'') - \frac{\|y'-y\|}{\delta + \|y'-y\|}f(y) \\
&= \frac{\|y'-y\|}{\delta + \|y'-y\|}(f(y'') - f(y)) \\
&\leq \frac{\|y'-y\|}{\delta}(M - m).
\end{aligned}$$

By exchanging y and y' we get $|f(y) - f(y')| \leq \frac{\|y-y'\|}{\delta}(M - m)$.

□

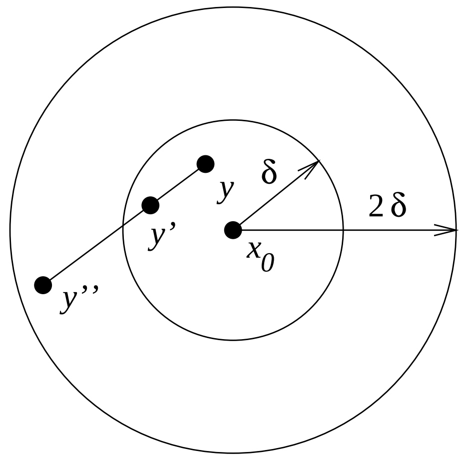


Figure 8.1: Illustration for proof of Lemma 8.1.

Theorem 8.2 *Let* $f \in \text{Conv } \mathbb{R}^n$. $S \subset \text{ri dom} f$, S *convex, compact. Then there exists* $L = L(S) \geq 0$ *such that*

$$|f(x) - f(x')| \leq L\|x - x'\| \textit{ for all } x, x' \in S. \tag{8.2}$$

Proof:

1. Since we are interested only in $x \in \text{dom} f$, we can state Inequality (8.2) in $\mathbb{R}^d$, where $d = \dim \text{dom} f$.
   Alternatively, we may assume ri dom f = int dom f, to simplify the writing.

2. Let $x_0 \in S$. We will prove that there exist $\delta = \delta(x_0) > 0$ and $L = L(x_0, \delta)$ such that $B(x_0, \delta) \subset \text{int dom} f$ and

$$|f(y) - f(y')| \leq L\|y - y'\| \text{ for all } y, y' \in B(x_0, \delta). \tag{8.3}$$

3. If 2. holds for all $x_0 \in S$, the balls $B(x_0, \delta)$ constitute a covering of S. S compact $\Rightarrow$ we can extract a finite covering

$$(x_1, \delta(x_1), L(x_1, \delta(x_1)), \ldots, (x_k, \delta(x_k), L(x_k, \delta(x_k)))).$$

4. Set $L := \max\{L(x_1, \delta(x_1)), \dots, L(x_k, \delta(x_k))\}$. Then for $x, x' \in S$ we can subdivide $[x, x']$ by the finite covering of 3. in subsegments $[x, y_1], [y_1, y_2], \dots, [y_k, x']$ where f is Lipschitzian with the common constant L. (Order these subsegments and their endpoints) $\Rightarrow$ with $x - x' = x - y_1 + y_1 - y_2 + y_2 - \dots - x'$. We have

$$\begin{aligned} |f(x) - f(x')| &= |f(x) - f(y_1) + f(y_1) - f(y_2) + f(y_2) - \dots - f(x')| \\ &\leq |f(x) - f(y_1)| + |f(y_1) - f(y_2)| + \dots + |f(y_k) - f(x')| \\ &\leq L\|x - y_1\| + \dots + L\|y_k - x'\| \\ &\leq L\|x - y_1 + y_1 - y_2 - \dots - x'\| = L\|x - x'\|. \end{aligned}$$

Now we still have to show 2. We show that the assumptions of Lemma 8.1 can be fulfilled. We construct, as in the proof of Theorem 2.1, a simplex $\Delta = \text{co}\,\{v_0, \dots, v_n\} \subset \text{dom}\, f$, with $x_0 \in \text{int}\,\Delta \Rightarrow \exists\, \delta > 0 : B(x_0, 2\delta) \subset \Delta$. Now any $y \in B(x_0, 2\delta)$ can be written as $y = \sum_{i=1}^n \alpha_i v_i$, $\alpha \in \Delta_{n+1}$. f convex $\Rightarrow f(y) \leq \sum_{i=1}^n \alpha_i f(v_i)$ by Proposition 6.5 f is bounded from below, say by m on $B(x_0, 2\delta)$: With $M := \max\{f(x_0), \dots, f(v_n)\}$ we get $m \leq f(y) \leq M$ for all $y \in B(x_0, 2\delta)$. Lemma 8.1 $\Rightarrow$ 2. holds.

□

Remark Theorem 8.2 means in particular that f is continuous on ri dom f. ($x, x_0 \in$ ri dom $f, x \to x_0 \Rightarrow f(x) \to f(x_0)$). We could state Theorem 8.2 also as: f is **locally** Lipschitzian on ri dom f: for all $x_0 \in$ ri dom f, there are $L(x_0), \delta(x_0)$ such that

$$|f(x) - f(x')| \leq L(x_0)\|x - x'\| \text{ for all } x, x' \in S(x_0),$$

where $S(x_0) := B(x_0, \delta(x_0)) \cap \text{aff dom}\, f \subset \text{ri dom}\, f$.

Proposition 8.3 (Lipschitzian Extension) *Let $C \subset \mathbb{R}^n$, $C \neq \emptyset$, C convex. Further, let $f \in \text{Conv}\,\mathbb{R}^n$ be Lipschitzian, with constant L on C. Then there exists a convex function f_1 such that*

1) $f_1(x) = f(x)$ for all $x \in C$, and

2) f_1 is Lipschitzian with constant L on $\mathbb{R}^n$.

Moreover, there is a largest function satisfying 1) and 2) namely the infimal convolution

$$\begin{aligned} \mathbb{R}^n \ni x \mapsto (f + I_C)_{[L]}(x) &:= ((f + I_C) \overset{+}{\vee} (L\|\cdot\|))(x) \\ &= \inf\{f(y) + L\|x - y\| : y \in C\}. \qquad (8.4) \end{aligned}$$

Proof: Let $\check{f} := (f + I_C)_{[L]}$.

- $\check{f} \in \text{Conv } \mathbb{R}^n$: $\check{f} > -\infty$ for all $x \in \mathbb{R}^n$. Let $x_0 \in \text{ri } C$ and apply Proposition 6.5 to $(f + I_C)$ with domain $C \Rightarrow$ There is a $s \in V$ (with V the subspace parallel to aff C) such that $f(x) \geq f(x_0) + \langle s, x - x_0 \rangle$ for all $x \in \mathbb{R}^n$. Now take $\delta > 0$ so small that $x = x_0 + \delta s \in C$ and use the Lipschitz property of f on C: $L\delta\|s\| \geq f(x) - f(x_0) \geq \delta\|s\|^2$ $\Rightarrow \|s\| \leq L$. Therefore we have $\langle s, x \rangle \leq \|s\|\|x\| \leq L\|x\|$ for all $x \in \mathbb{R}^n$ $\Rightarrow (f + I_C)$ and $L\|\cdot\|$ are minorized by a common affine function with slope s. Proposition 7.5 $\Rightarrow \check{f} \in \text{Conv } \mathbb{R}^n$.

- $\check{f}$ is Lipschitzian on $\mathbb{R}^n$: For $x, x' \in \mathbb{R}^n$ and $\epsilon > 0$, let $y' \in C$ be such that $f(y') + L\|x' - y'\| \leq \check{f}(x') + \epsilon$. By definition, we have

$$\begin{aligned} \check{f}(x) &\leq f(y') + L\|x - y'\| \leq f(y') + L\|x - x'\| + L\|x' - y'\| \\ \Rightarrow \check{f}(x) &\leq \check{f}(x') + L\|x - x'\| + \epsilon. \end{aligned}$$

  Since this holds for arbitrary $x, x', \epsilon \Rightarrow f$ is Lipschitzian.

- $\check{f} = f$ on C: Let $x \in C$. By definition, $\check{f}(x) \leq f(x)$. (We add $L\|x - y\|$). By the Lipschitz property of f on C

$$\begin{aligned} &\Rightarrow f(x) \leq f(y) + L\|y - x\| \text{ for all } y \in C \\ &\Rightarrow f(x) \leq \check{f}(x) \\ &\Rightarrow \check{f}(x) = f(x) \text{ for all } x \in C. \end{aligned}$$

- $\check{f}$ is maximal: Let f_1 satisfy 1) and 2)

$$\begin{aligned} &\Rightarrow f_1(x) - f(y) \leq L\|x - y\| \text{ for all } x \in \mathbb{R}^n, y \in C \\ &\Rightarrow f_1(x) \leq \check{f}(x) \text{ for all } x \in \mathbb{R}^n. \end{aligned}$$

□

Remark On rbd dom f, no continuity can be guaranteed.
$f(x) = \sup_{\alpha,\beta}\{\xi\alpha + \eta\beta : \frac{1}{2}\alpha^2 \leq \beta\}$. $f \in \overline{\text{Conv}}\ \mathbb{R}^n \Rightarrow f$ is l.s.c. But f is not u.s.c (upper semi-continuous). $f(0) = 0$,

$$\begin{aligned} f(\xi, \eta) &= \sup_{\alpha}(\frac{1}{2}\eta\alpha^2 + \xi\alpha) \\ &= \begin{cases} 0 & \text{if } \xi = \eta = 0 \\ \frac{-\xi^2}{2\eta} & \text{if } \eta < 0 \\ +\infty & \text{otherwise} \end{cases} \end{aligned}$$

If $x \to 0$ with $\eta = -\frac{1}{2}\xi^2 \Rightarrow f(x) \equiv 1 > 0 = f(0) \Rightarrow f$ is not u.s.c at 0.

Theorem 8.4 *Let f_k: $\mathbb{R}^n \to \mathbb{R}$ be convex and let the f_k converge pointwise for $k \to +\infty$ to f: $\mathbb{R}^n \to \mathbb{R}$. Then f is convex and for each compact set S, the convergence of f_k to f is uniform on S.*

Proof:

1. By definition of convexity, f is convex.

2. We want to use Lemma 8.1 to show the uniform convergence. Therefore, we need to bound f_k on S independently of k. Let $r > 0$ such that $S \subset B(0, r)$.

   a) $g := \sup_k f_k$ is convex (Proposition 7.2) and $g(x) < +\infty$ for all $x \in \mathbb{R}^n$. Since $\{f_k(x)\}$ is bounded, Theorem 8.2 $\Rightarrow$ g is continuous and therefore bounded by M on the compact set $B(0, 2r)$. $f_k(x) \leq g(x) \leq M$ for all k and all $x \in B(0, 2r)$. Moreover, the sequence $\{f_k(0)\}$ is bounded from below (f_k convex): $\mu \leq f_k(0)$ for all k. $\Rightarrow$ For $x \in B(0, 2r)$ and all k, we use the convexity relation on $[x, x'] \subset B(0, 2r)$:

   $$2\mu \leq 2f_k(0) \leq f_k(x) + f_k(-x) \leq f_k(x) + M$$

   $\Rightarrow$ the f_k 's are bounded from below (independently of k) $\Rightarrow$ We can use Lemma 8.1: There is some L (independent of k) such that

   $$|f_k(y) - f_k(y')| \leq L\|y - y'\| \text{ for all } k \text{ and all } y, y' \in B(0, r). \tag{8.5}$$

   Of course, the same Lipschitz property holds for f.

   b) Fix $\epsilon > 0$ and cover S by balls $B(x, \epsilon)$ for all x describing S. Now extract a finite covering $B(x_1, \epsilon), \ldots, B(x_m, \epsilon) \supset S$. With $x \in S$, arbitrary, take x_i such that $x \in B(x_i, \epsilon)$. There is $k_{i,\epsilon}$ such that for all $k > k_{i,\epsilon}$:

   $$\begin{aligned} |f_k(x) - f(x)| &\leq |f_k(x) - f_k(x_i)| + |f_k(x_i) - f(x_i)| \\ &\quad + |f(x_i) - f(x)| \\ &\leq (2L + 1)\epsilon, \; x, x_i \in S \subset B(0, r). \end{aligned}$$

   This inequality is valid uniformly in x, providing that

   $$k \geq \max\{k_{1,\epsilon}, \ldots, k_{m,\epsilon}\} =: k_\epsilon.$$

$\square$

8.1 Exercise

Let $f : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ be a convex function. First show that

$$f\left(x_2 + \frac{x_1 - x_2}{\alpha}\right) \geq \frac{1}{\alpha}\Big(f(x_1) - f(x_2)\Big) + f(x_2)$$

for all $0 < \alpha \leq 1$.

Use this inequality to show that f is continuous on $\operatorname{int} \operatorname{dom} f$ $(= \operatorname{ri} \operatorname{dom} f)$.

Chapter 9

BEHAVIOR AT INFINITY

We assume that $f \in \overline{\text{Conv}}\, \mathbb{R}^n$ to investigate $(\text{epi}\, f)_\infty$. $(\text{epi}\, f)_\infty$ is a closed convex cone in $\mathbb{R}^n \times \mathbb{R}$ and contains $\{0\} \times \mathbb{R}_{+0}$.

$$(\text{epi}\, f)_\infty = \{(d, \rho) \in \mathbb{R}^n \times \mathbb{R} : (x_0, r_0) + t(d, \rho) \subset \text{epi}\, f \ \forall\ t > 0\}, \qquad (9.1)$$

where $(x_0, r_0) \in \text{epi}\, f$ (arbitrary). This can be rewritten

$$(\text{epi}\, f)_\infty = \{(d, \rho) : \text{epi}\, f + t(d, \rho) \subset \text{epi}\, f \ \forall\ t > 0\}$$

and since $(\text{epi}\, f)_\infty$ is a convex cone,

$$(\text{epi}\, f)_\infty = \{(d, \rho) : \text{epi}\, f + (d, \rho) \subset \text{epi}\, f\}.$$

Proposition 9.1 *Let* $f \in \overline{\text{Conv}}\, \mathbb{R}^n$. *Then* $(\text{epi}\, f)_\infty = \text{epi}\, f'_\infty$, *with* $f'_\infty \in \overline{\text{Conv}}\, \mathbb{R}^n$ *defined by*

$$\begin{aligned} d \mapsto f'_\infty(d) \quad &:= \quad \sup_{t>0} \frac{f(x_0 + td) - f(x_0)}{t} \\ &= \quad \lim_{t \to +\infty} \frac{f(x_0 + td) - f(x_0)}{t}, \end{aligned}$$

where $x_0 \in \text{dom}\, f$ *(arbitrary).*

Proof: Since $(x_0, f(x_0)) \in \text{epi}\, f$ (9.1) yields that $(d, \rho) \in (\text{epi}\, f)_\infty$ $\Leftrightarrow f(x_0 + td) \le f(x_0) + t\rho$ for all $t > 0$. This means

$$\sup_{t>0} \frac{f(x_0 + td) - f(x_0)}{t} \le \rho \qquad (9.2)$$

$\Rightarrow (\text{epi}\, f)_\infty$ is the epigraph of the function whose value at d is the left-hand side of Inequality (9.2) (independent of x). Moreover, the difference quotient in Inequality (9.2) is closed convex in d and increasing in t ($t \mapsto f(x_0 + td)$ is convex). Proposition 7.4 $\Rightarrow f'_\infty \in \overline{\text{Conv}}\, \mathbb{R}^n$.

□

Remark By construction, f'_∞ is positively homogeneous:
$f'_\infty(\alpha d) = \alpha f'_\infty(d)$ for all $\alpha > 0$.

Definition 42 *The function f'_∞, defined in Proposition 9.1, is called the* **asymptotic function** *(or recession function, or auto-deconvolution) of f.*

Example 9.1 $f = I_C$, *C is a closed convex set.* $I_C(x_0 + td) = 0$ *for all* $t > 0 \Leftrightarrow d \in C_\infty \Rightarrow (I_C)'_\infty = I_{C_\infty}$.

Proposition 9.2 *Let $f \in \overline{\text{Conv}}\ \mathbb{R}^n$. Then all nonempty sublevel-sets of f have the same asymptotic cone, which is the sublevel-set of f'_∞ at level 0: $\forall\, r \in \mathbb{R}$ with $S_r(f) \neq \emptyset$ we have*

$$(S_r(f))_\infty = \{d \in \mathbb{R}^n : f'_\infty(d) \leq 0\} = S_0(f'_\infty).$$

In particular, the following statements are equivalent:

i) $\exists\, r \in \mathbb{R} : S_r(f) \neq \emptyset$, *compact.*

ii) All sublevel-sets of f are compact.

iii) $f'_\infty(d) > 0$ *for all* $d \in \mathbb{R}^n, d \neq 0$.

Proof: By definition, $d \in (S_r(f))_\infty$ if and only if $x \in S_r(f) \Rightarrow (x + td \in S_r(f)$ for all $t > 0)$, or $(x, r) \in \text{epi}\, f \Rightarrow (x + td, r + 0 \cdot t) \in \text{epi}\, f$ for all $t > 0$ $(\Leftrightarrow d \in S_0(f'_\infty))$, which means $(d, 0) \in (\text{epi}\, f)_\infty = \text{epi}\, f'_\infty$.
iii) means the particular case where $S_0(f'_\infty) = \{0\}$. This is equivalent to $(S_r(f))_\infty = \{0\}$ for all $r \in \mathbb{R}$ with $S_r(f) \neq \emptyset$. Proposition 2.8 $\Leftrightarrow S_r(f)$ is compact.

□

Definition 43 (Coercivity) *$f \in \overline{\text{Conv}}\ \mathbb{R}^n$ satisfying i) ii) or iii) in Proposition 9.2 are called* **0-coercive**. *Equivalently, the 0-coercive functions are those that "increase at infinity". $f(x) \to +\infty$ whenever $\|x\| \to +\infty$, and closed convex 0-coercive functions achieve their minimum over $\mathbb{R}^n$.*
If $f'_\infty(d) = +\infty$ for all $d \neq 0$ ($f'_\infty = I_{\{0\}}$) we have $\frac{f(x)}{\|x\|} \to +\infty$ whenever $\|x\| \to +\infty$ (take a cluster point of $\left\{\frac{x_k}{\|x_k\|}\right\}$ and use the fact that f'_∞ is l.s.c.). In this case f is called **1-coercive** *or* **coercive**.

Proposition 9.3 *Let $f \in \overline{\text{Conv}}\,\mathbb{R}^n$. f is Lipschitzian on $\mathbb{R}^n \iff f'_\infty < \infty$ on $\mathbb{R}^n$. The best Lipschitz constant for f is then* $\sup\{f'_\infty(d) : \|d\| = 1\}$.

Proof: " $\Longleftarrow$ ": $f'_\infty < \infty$ on $\mathbb{R}^n$. Chapter 8 $\Rightarrow f'_\infty$ is continuous and therefore bounded on the comapct unit sphere: $\sup\{f'_\infty(d) : \|d\| = 1\} =: L < +\infty$. Since f'_∞ is positively homogeneous, we have $f'_\infty(d) \leq L\|d\|$ for all $d \in \mathbb{R}^n$. We can express $f'_\infty(d)$ as $\sup\limits_{t>0}\left\{\frac{f(x_0+td)-f(x_0)}{t} : x_0 \in \text{dom}\, f\right\}$ and we get $f(x_0+td) - f(x_0) \leq L\|td\|$ for all $x_0 \in \text{dom}\, f$, $d \in \mathbb{R}^n$. But $\text{dom}\, f = \mathbb{R}^n$ since $f(x_0+td) < +\infty$ for all $d \Rightarrow L$ is a global Lipschitz constant for f.
" $\Longrightarrow$ ": Let f have a global Lipschitz constant L. Take $x_0 \in \text{dom}\, f$ and use Proposition 9.1: $f(x_0+td) - f(x_0) \leq Lt\|d\|$ for all $t > 0$ and $d \in \mathbb{R}^n$. With the definition of f'_∞ we get $f'_\infty(d) \leq L\|d\|$ for all $d \in \mathbb{R}^n$. We have $f'_\infty < +\infty$ and $L \leq \sup\{f'_\infty(d) : \|d\| = 1\}$.

□

9.1 Exercises

1. Which of the following functions f_i are 1-coercive?

   a) $f_1(x) = \sqrt{|x|}, \quad f_1 : \mathbb{R} \to \mathbb{R}.$

   b) $f_2(x) = \prod\limits_{i=1}^{n} |x_i|, \quad f_2 : \mathbb{R}^n \to \mathbb{R},\ n \geq 1.$

   Find co epi *f_1*.

2. a) $f_1, \ldots, f_m \in \overline{\text{Conv}}\,\mathbb{R}^n$, $t_1, \ldots, t_m > 0$ and $\exists\ x_0 : f_j(x_0) < \infty \quad \forall j = 1, \ldots, m$. Then for

   $$f = \sum_{j=1}^{m} t_j f_j \quad \Rightarrow \quad f'_\infty = \sum_{j=1}^{m} t_j (f_j)'_\infty$$

   b) Let $f_j,\ j \in J$ be a family of functions in $\overline{\text{Conv}}\,\mathbb{R}^n$ s.t. $\exists\ x_0\ \sup\limits_{j\in J}\ f_j(x_0) < \infty$. Then for

   $$f = \sup_{j\in J} f_j \quad \Rightarrow \quad f'_\infty = \sup_{j\in J} (f_j)'_\infty$$

   c) $A : \mathbb{R}^n \to \mathbb{R}^n$ affine, $A(x) = A_0 x + y_0$ and $f \in \overline{\text{Conv}}\,\mathbb{R}^n$. $A(\mathbb{R}^n) \cap \text{dom}\ f \neq \emptyset$. Then $(f \circ A)'_\infty = f'_\infty \circ A_0$.

Chapter 10

DIFFERENTIABLE CONVEX FUNCTIONS

Theorem 10.1 *Let f be a function differentiable on $\Omega \subset \mathbb{R}^n$, Ω open. Moreover, let $C \subset \Omega$, C convex. Then*

i) f is convex on C

$$\Leftrightarrow f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle \text{ for all } (x, x_0) \in C \times C. \quad (10.1)$$

ii) f is strictly convex on $C \Leftrightarrow$ (10.1) holds with strict inequality for $x \neq x_0$.

iii) f is strongly convex with modulus c on C

$$\Leftrightarrow f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2} c \|x - x_0\|^2 \;\; \forall \, (x, x_0) \in C \times C.$$

Proof: (Refer to Figure 10.1).

i) " $\Longrightarrow$ ": Let f be convex on $C \Rightarrow$ For $(x_0, x) \in C \times C, \alpha \in (0,1)$ we have

$$\begin{aligned} f(\alpha x + (1-\alpha)x_0) - f(x_0) &\leq \alpha(f(x) - f(x_0)) \\ \Rightarrow \frac{1}{\alpha}(f(\alpha x + (1-\alpha)x_0) - f(x_0)) &\leq f(x) - f(x_0). \end{aligned}$$

With $\alpha \downarrow 0$ we get
$\langle \nabla f(x_0), x - x_0 \rangle \leq f(x) - f(x_0) \Rightarrow$ (10.1) holds.
" $\Longleftarrow$ ": Let $x_1, x_2 \in C, \alpha \in (0,1)$ and define $x_0 := \alpha x_1 + (1-\alpha)x_2 \subset C$.

By assumption, we have $f(x_i) \geq f(x_0) + \langle \nabla f(x_0), x_i - x_0 \rangle$ for $i = 1, 2$. Convex combination of these inequalities yields:

$$\alpha f(x_1) + (1-\alpha) f(x_2) \geq f(x_0) + \langle \nabla f(x_0), \alpha x_1 + (1-\alpha) x_2 - x_0 \rangle.$$

Since $x_0 := \alpha x_1 + (1-\alpha) x_2$ we get

$$\alpha f(x_1) + (1-\alpha) f(x_2) \geq f(\alpha x_1 + (1-\alpha) x_2) + \overbrace{\langle \nabla f(x_0), 0 \rangle}^{=0}$$

$\Rightarrow f$ is convex.

ii) " $\Longrightarrow$ ": f is strictly convex $\Rightarrow$ For $x, x_0 \in C, x \neq x_0, \alpha \in (0,1)$ we have $f(x_0 + \alpha(x - x_0)) - f(x_0) < \alpha(f(x) - f(x_0))$. f is convex $\Rightarrow f(x_0 + \alpha(x - x_0)) - f(x_0) \geq \langle \nabla f(x_0), \alpha(x - x_0) \rangle$.
" $\Longleftarrow$ " Analogous to i).

iii) Use Proposition 6.1 and the fact that $f - \frac{1}{2} c \|\cdot\|^2$ is also differentiable, and apply i).

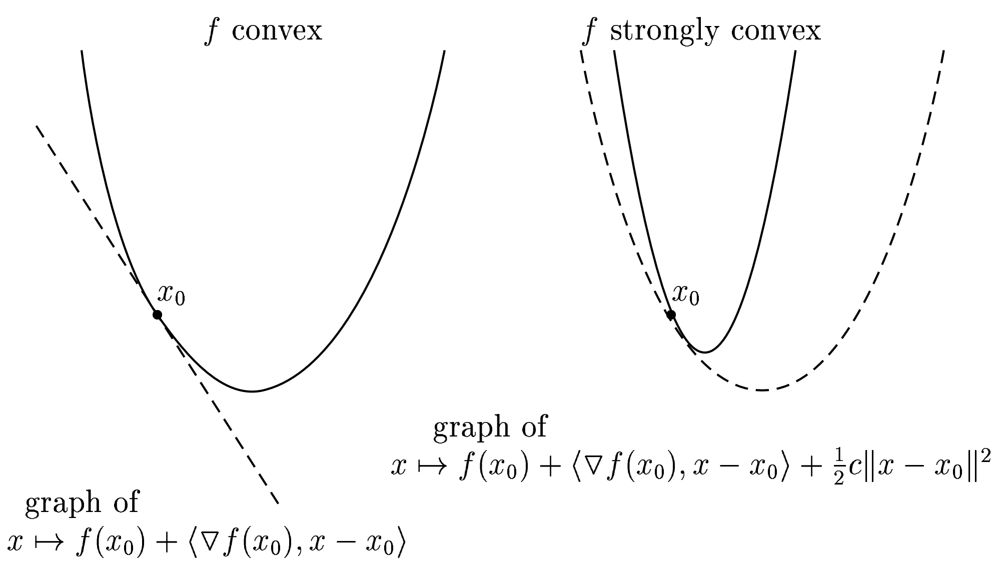


Figure 10.1: Illustration for proof of Theorem 10.1.

□

Remark If we write (10.1) as equality: $f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + r(x_0, x)$. f is convex $\Rightarrow r(\cdot, \cdot) \geq 0$ for all x, x_0 and $r(x_0, \cdot)$ is convex.

Definition 44 *Let $C \subset \mathbb{R}^n$ be convex. $F : C \to \mathbb{R}^n$ is* **monotone** *(resp. strictly monotone, strongly monotone with modulus $c > 0$) on C if for all $x, x' \in C$:*

$$\langle F(x) - F(x'), x - x' \rangle \geq 0$$

$$(\textit{resp. } \langle F(x) - F(x'), x - x' \rangle > 0 \;\; \textit{whenever}\, x \neq x',$$

$$\textit{resp. } \langle F(x) - F(x'), x - x' \rangle \geq c \|x - x'\|^2).$$

Theorem 10.2 *Let f be differentiable on $\Omega \subset \mathbb{R}^n, \Omega$ open and let $C \subset \Omega, C$ convex. Then f is convex (resp. strictly convex, resp. strongly convex with modulus $c > 0$) on $C \iff \nabla\, f$ is monotone (resp. strictly monotone, resp. strongly monotone with modulus $c > 0$) on C.*

Proof: We show the result for strongly convex and strongly monotone with $c \geq 0$ (includes the convex, monotone case).
" $\Longrightarrow$ ": Let f be strongly convex on C. Theorem 10.1 $\Rightarrow$ For $x, x_0 \in C$ we have

$$\begin{aligned} f(x) &\geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2} c \, \|x - x_0\|^2 \\ f(x_0) &\geq f(x) + \langle \nabla f(x), x_0 - x \rangle + \frac{1}{2} c \, \|x_0 - x\|^2. \end{aligned}$$

Adding these two inequalities up yields: $\langle \nabla f(x) - \nabla f(x_0), x - x_0 \rangle \geq c \, \|x - x_0\|^2 \Rightarrow \nabla f$ is strongly monotone.
" $\Longleftarrow$ ": Take $x_0, x_1 \in C$ and consider the function $(\mathbb{R} \to \mathbb{R})$ $t \mapsto \phi(t) := f(x_t)$, with $x_t := x_0 + t(x_1 - x_0)$ and $t \in (a, b) \supset [0, 1], x_t \in \Omega$. ϕ is well-defined and differentiable $\Rightarrow \phi'(t) = \langle \nabla f(x_t), x_1 - x_0 \rangle$. Therefore we have for all $0 \leq t' < t \leq 1$

$$\begin{aligned} \phi'(t) - \phi'(t') &= \langle \nabla f(x_t) - \nabla f(x_{t'}), x_1 - x_0 \rangle \\ &= \frac{1}{t - t'} \langle \nabla f(x_t) - \nabla f(x_{t'}), x_t - x_{t'} \rangle. \end{aligned} \tag{10.2}$$

Since ∇f is monotone $\Rightarrow \phi'$ is increasing. Now we use the following result, known as the Criterion of Increasing Slopes (see [3], pp 3-4): Let I be a nonempty interval of $\mathbb{R}$. A function $f : \to \mathbb{R}$ is convex on I if and only if, for all $x_0 \in I$, the slope-function

$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$$

is increasing on $I \setminus \{x_0\}$.
$\Rightarrow \phi$ is convex. For strong convexity, set $t' = 0$ in (10.2) and use strong monotonicity for ∇f to get from (10.2) $\phi'(t) - \phi'(0) \geq \frac{1}{t} c\|x_t - x_0\|^2 = t\, c\, \|x_1 - x_0\|^2$. Since ϕ is the integral of ϕ' we can write

$$\phi(1) - \phi(0) - \phi'(0) = \int_0^1 (\phi'(t) - \phi'(0))\, d\, t \geq \frac{1}{2} c\|x_1 - x_0\|^2$$

$$\Rightarrow \quad f(x_1) - f(x_0) - \langle \nabla f(x_0), x_1 - x_0 \rangle \geq \frac{1}{2}\, c\, \|x_1 - x_0\|^2.$$

Theorem 10.1 $\Rightarrow f$ is strongly convex. The (strictly convex $\Longleftrightarrow$ strictly monotone) case follows analogously.

□

Example 10.1 *$C \subset \mathbb{R}^n$, $C \neq \emptyset$, C convex. Consider $\phi_C(x) := \frac{1}{2}(\|x\|^2 - d_C^2(x))$ which is convex and finite on $\mathbb{R}^n$ (Example 39). What is $\nabla \phi_C$?: Denote $\Delta := d_C^2(x + h) - d_C^2(x)$. Since $d_C^2(x) \leq \|x - p_C(x + h)\|^2$, we have $\Delta \geq \|x + h - p_C(x + h)\|^2 - \|x - p_C(x + h)\|^2 = \|h\|^2 + 2\langle h, x - p_C(x + h)\rangle$. By exchanging the role of x and $x + h$ we obtain:*

$$\Delta \leq \|x + h - p_C(x)\|^2 - \|x - p_C(x)\|^2 = \|h\|^2 + 2\langle h, x - p_C(x)\rangle.$$

*From Chapter 3, we know that p_C is nonexpensive ($\|p_C(x_1) - p_C(x_2)\| \leq \|x_1 - x_2\|$) $\Rightarrow \Delta = 2\langle x - p_C(x), h\rangle + o(\|h\|) \Rightarrow \nabla d_C^2(x) = 2(x - p_C(x))$ $\Rightarrow \Delta\phi_C(x) = \frac{1}{2}(2x - [2(x - p_C(x))]) = p_C(x)$. For $x \notin C$ we have $d_C(x) > 0$ and we get: $\nabla d_C(x) = \nabla\sqrt{d_C^2(x)} = \frac{x - p_C(x)}{\|x - p_C(x)\|}$.
Using Proposition 3.3 we get for all $x, x' \in C$*

$$\begin{aligned} \langle \nabla\phi_C(x) - \nabla\phi_C(x', x - x'\rangle &= \langle p_C(x) - p_C(x'), x - x'\rangle \\ &\geq \|p_C(x) - p_C(x')\|^2 \\ &= \|x - x'\|^2 \end{aligned} \qquad (10.3)$$

*$\Rightarrow \nabla\phi_C$ is strongly monotone with modulus 1 on C. Theorem 10.2 $\Rightarrow \phi_C$ is strongly convex with modulus 1 on C.
What about ϕ_C outside C?
Take $p \in \operatorname{bd} C$, $x, x' \in N_C(p) \Rightarrow p_C(x) = p_C(x') = p \Rightarrow$ left-hand side of (10.3) is $0 \Rightarrow \phi_C$ is affine on $\{p\} + N_C(p)$. Refer to Figure 10.2.*

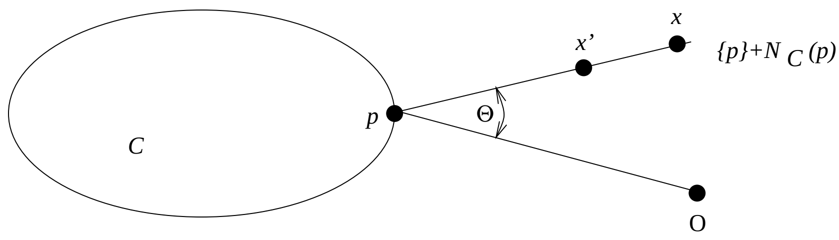


Figure 10.2: Illustration for Example 10.1.

We apply the triangular relation

$$\|p\|^2 + \|x - p\|^2 = \|x\|^2 - 2\|p\|\|x - p\| \cos \Theta$$

to see that

$$\phi_C(x) = \frac{1}{2} (\|x\|^2 - \|x - p\|^2) = \frac{1}{2} \|p\|^2 + \|p\|\|x - p\| \cos \Theta$$

is affine with respect to $\|x - p\|$ *when* p *and* Θ *are fixed.* $\nabla \phi_C$ *cannot be strongly convex outside* $C \Rightarrow \nabla \phi_C$ *cannot be strongly monotone outside* C.

Theorem 10.3 *Let* $f \in \text{Conv}\, \mathbb{R}^n$. *The subset of* $\text{int dom} f$ *where* f *is not differentiable is of zero Lebesgue measure.*

Theorem 10.4 *Let* f *be twice differentiable on* $\Omega \subset \mathbb{R}^n$, Ω *open, convex. Then*

i) f *is convex on* $\Omega \Longleftrightarrow \nabla^2 f(x_0)$ *is positive semi-definite for all* $x_0 \in \Omega$.

ii) $\nabla^2 f(x_0)$ *is positive definite for all* $x_0 \in \Omega \Longrightarrow f$ *is strictly convex on* Ω.

iii) f *is strongly convex with modulus* c *on* $\Omega \Longleftrightarrow$ *the smallest eigenvalue of* $\nabla^2 f(x)$ *is minorized by* c *on* Ω : *for all* $x_0 \in \Omega$, $d \in \mathbb{R}^n$:

$$\langle \nabla^2 f(x_0) d, d \rangle \geq c \|d\|^2.$$

Proof: For given $x_0 \in \Omega, d \in \mathbb{R}^n$ and $t \in \mathbb{R}$ such that $x_0 + td \in \Omega$, we set $x_t := x_0 + td, \phi(t) := f(x_t) = f(x_0 + td), \phi''(t) = \langle \nabla^2 f(x_t) d, d \rangle$.

i) " $\Longrightarrow$ ": f is convex on $\Omega, x_0 \in \Omega, d \in \mathbb{R}^n, d \neq 0 \Rightarrow \phi$ is convex on $I := \{t \in \mathbb{R} : x_0 + td \in \Omega\}$. We get from the Criterion of Increasing Slopes [3]

$$0 \leq \phi''(t) = \langle \nabla^2 f(x_t) d, d \rangle \ \text{ for all } t \in I$$

$\Rightarrow \nabla^2 f(x_0)$ is positive semi-definite.
" $\Longleftarrow$ ": Take $[x_0, x_1] \subset \Omega$ and set $d := x_1 - x_0$. Since by assumption $\nabla^2 f(x_t)$ is positive semi-definite we have $\phi''(t) \geq 0$. The Criterion of Increasing Slopes [3] $\Rightarrow \phi$ is convex on $[0,1] \Rightarrow f$ is convex on Ω since x_0, x_1 were chosen arbitrarily.

ii) Show that ∇f is strictly monotone on Ω and the result follows then by Theorem 10.2. Take $[x_0, x_1] \subset \Omega, x_1 \neq x_0, d := x_1 - x_0$ and apply the Mean-Value Theorem to ϕ' (which is differentiable on $[0,1]$): For some $\tau \in (0,1)$ we have $\phi'(1) - \phi'(0) = \phi''(\tau) = \langle \nabla^2 f(x_\tau) d, d \rangle > 0$. The result follows since $\phi'(1) - \phi'(0) = \langle \nabla f(x_1) - \nabla f(x_0), x_1 - x_0 \rangle$.

iii) Use Proposition 6.1 and apply i) to $f - \frac{1}{2} c \|\cdot\|^2$, whose Hessian operator is $\nabla^2 f - cI_n$ and has the eigenvalues $\lambda - c$, where λ describes the eigenvalues of $\nabla^2 f$.

□

10.1 Exercises

1. Let $\Omega = \{x : x_i > 0,\ i = 1 \ldots n\}$ and $f : \Omega \to \mathbb{R}$ defined by $f(x) = -\left(\prod_{i=1}^{n} x_i\right)^{1/n}$. Show that f is convex using the criterion of Theorem 10.4.

2. $F : D \to \mathbb{R}^n,\ D \subseteq \mathbb{R}^n$ is monotone (strictly monotone, strongly monotone with modulus $c > 0$) if

$$\begin{array}{lrcll} \forall\, x, x' \in D & \langle F(x) - F(x'), x - x' \rangle & \geq & 0 & \\ & (\langle F(x) - F(x'), x - x' \rangle & > & 0 \quad \text{whenever} & x \neq x' \\ & \langle F(x) - F(x'), x - x' \rangle & \geq & c\|x - x'\|^2). & \end{array}$$

   a) Show that $f : \mathbb{R} \to \mathbb{R}$ which is monotone increasing ($x \leq y \Rightarrow f(x) \leq f(y)$) is monotone.
   b) Suppose F is monotone then $F_\gamma(x) = F(x) + \gamma x$ is strongly monotone with modulus $\gamma > 0$.
   c) If F is a Q-contraction, i.e. $\|F(x) - F(x')\| \leq Q\|x - x'\|,\ Q < 1$ then $F'(x) = x - F(x)$ is strongly monotone with modulus $1 - Q$.

Chapter 11

SUBLINEAR FUNCTIONS AND GAUGES

Definition 45 *A function* $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is* **sublinear** *if it is convex* ($\sigma \in \mathrm{Conv}\, \mathbb{R}^n$) *and positive homogeneous* ($\sigma(tx) = t\sigma(x)$ *for all* $x \in \mathbb{R}^n$, $t > 0$).

Remark

- σ is positive homogeneous $\Longleftrightarrow$

$$\sigma(tx) \leq t\sigma(x) \text{ for all } x \in \mathbb{R}^n, t > 0, \tag{11.1}$$

  since it follows $\sigma(x) = \sigma(t^{-1} \cdot tx) \leq t^{-1}\sigma(tx)$

$$\Rightarrow t\sigma(x) \leq \sigma(tx), \tag{11.2}$$

  and from Inequality (11.1) and Inequality (11.2) we get equality.

- $\sigma(0) = \sigma(t \cdot 0) = t\sigma(0)$ for all $t > 0 \Rightarrow \sigma(0) = 0$ or $\sigma(0) = +\infty$.

Proposition 11.1 *A function* $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is sublinear* $\Longleftrightarrow$ epi σ *is a nonempty cone in* $\mathbb{R}^n \times \mathbb{R}$.

Proof: From Proposition 6.2 we know σ is convex $\Longleftrightarrow$ epi σ is a nonempty convex set. It is left to show that σ is positive homogeneous $\Longleftrightarrow$ epi σ is a cone.
" $\Longrightarrow$ ": Let σ be positive homogeneous. For $(x, r) \in$ epi σ from $\sigma(x) \leq r$ we get

$$\sigma(tx) = t\sigma(x) \leq t\, r \quad \text{for all } t > 0$$

$\Rightarrow t \cdot (x, r) \in \text{epi } \sigma \Rightarrow \text{epi } \sigma$ is a cone.
" $\Longleftarrow$ ": Let epi σ be a cone. Therefore $(x, \sigma(x)) \in \text{epi } \sigma$ implies $(tx, t\sigma(x)) \in \text{epi } \sigma \Rightarrow \sigma(tx) \leq t\sigma(x)$ for all $t > 0$. Using the above Remark $\Rightarrow \sigma$ is positive homogeneous.

□

Definition 46 *A function* $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is* **subadditive** *if it satisfies* $\sigma(x_1 + x_2) \leq \sigma(x_1) + \sigma(x_2)$ *for all* $x_1, x_2 \in \mathbb{R}^n$.

Proposition 11.2 *A function* $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, *not identically equal to* $+\infty$, *is sublinear if and only if one of the following two properties holds:*

(P1) $\sigma(t_1 x_1 + t_2 x_2) \leq t_1 \sigma(x_1) + t_2 \sigma(x_2)$ *for all* $x_1, x_2 \in \mathbb{R}^n, t_1, t_2 > 0$.

(P2) σ *is positive homogeneous and subadditive.*

Proof:

- "*sublinearilty* $\Longrightarrow$ $(P1)$": For $x_1, x_2 \in \mathbb{R}^n, t_1, t_2 > 0$ set $t := t_1 + t_2 > 0$ and compute

$$\begin{aligned} \sigma(t_1 x_1 + t_2 x_2) &= \sigma\left(t\left(\frac{t_1}{t}x_1 + \frac{t_2}{t}x_2\right)\right) \\ &= t\sigma\left(\frac{t_1}{t}x_1 + \frac{t_2}{t}x_2\right) \\ &\leq t\left(\frac{t_1}{t}\sigma(x_1) + \frac{t_2}{t}\sigma(x_2)\right) \end{aligned}$$

  $\Rightarrow$ (P1) holds.

- "$(P1) \Longrightarrow (P2)$": From (P1) we get σ is subadditive by setting $t_1 = t_2 = 1$. Positive homogeneity follows with $x_1 = x_2 = x, t_1 = t_2 = \frac{1}{2}t$,

$$\sigma(tx) \leq t\sigma(x) \text{ for all } x \in \mathbb{R}^n, t > 0$$

  $\Rightarrow$ (P2) holds.

- "$(P2) \Longrightarrow$ *sublinearity*": Take $t_1, t_2 > 0$ with $t_1 + t_2 = 1$, and apply subadditivity and then positive homogeneity:

$$\sigma(t_1 x_1 + t_2 x_2) \leq \sigma(t_1 x_1) + \sigma(t_2 x_2) = t_1 \sigma(x_1) + t_2 \sigma(x_2)$$

  $\Rightarrow \sigma$ is convex $\Rightarrow \sigma$ is sublinear.

□

Corollary 11.3 *If σ is sublinear, then $\sigma(x) + \sigma(-x) \geq 0$ for all $x \in \mathbb{R}^n$.*

Proof: Take $x_2 = -x_1$. Subadditivity $\Rightarrow \sigma(x) + \sigma(-x) \geq \sigma(0) \geq 0$.

□

Remark Convexity and subadditivity do not necessarily imply sublinearity. $f(x) \equiv a \in \mathbb{R}_+$.

$$f(\lambda x + (1-\lambda)x') \leq \lambda f(x) + (1-\lambda) f(x')$$

$$f(x + x') \leq f(x) + f(x'),$$

but $f(tx) \nleq t f(x) = t\,a$ for all $t > 0$.

The classes of functions can be summarized in Figure 11.1.

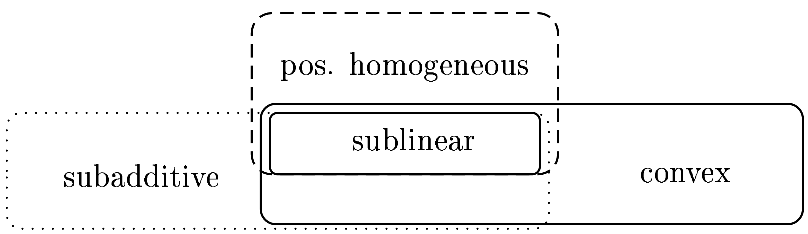


Figure 11.1: Classes of functions.

Definition 47 *Let $C \subset \mathbb{R}^n$, $0 \in C$, C closed and convex. The function γ_C defined by $\gamma_C(x) := \inf\{\lambda > 0 : x \in \lambda C\}$ is called the* **gauge** *of C. We set $\gamma_C(x) := +\infty$ if $x \notin \lambda C$ for all $\lambda > 0$.*

Remark If we take the unit ball in $\mathbb{R}^n$ as C, $a > 0$, then $\gamma_C(x) = \|x\|$ (Euclidean norm).
Figure 11.2 illustrates the epigraph of the gauge function.

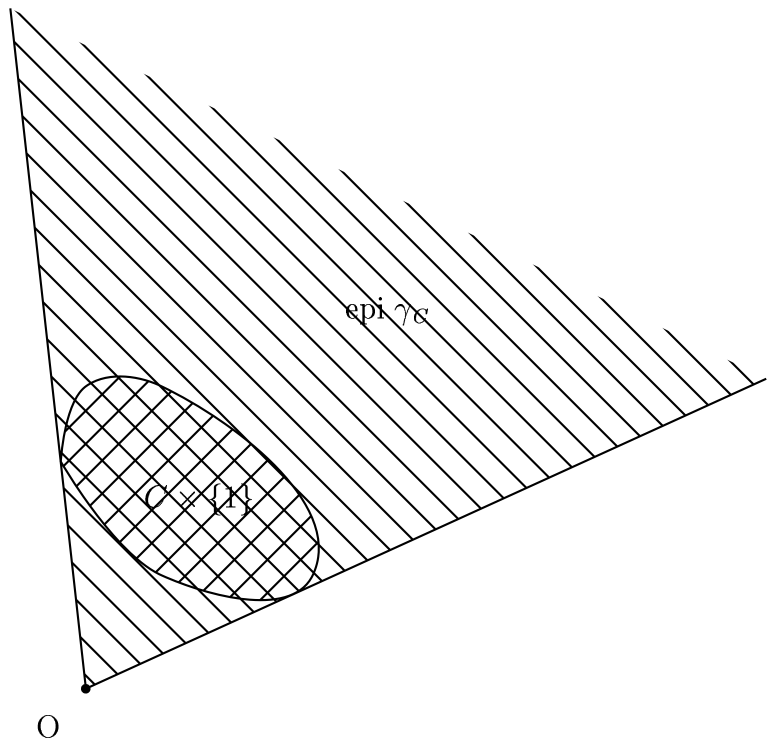


Figure 11.2: Illustration for the epigraph of the gauge function.

Theorem 11.4 *Let C be a closed convex set containing the origin. Then*

i) γ_C *is a nonnegative closed sublinear function.*

ii) γ_C *is finite everywhere* $\Longleftrightarrow 0 \in \operatorname{int} C$.

iii) $r\,C = \{x \in \mathbb{R}^n : \gamma_C(x) \le r\}$ *for all* $r > 0$ *and*
$C_\infty = \{x \in \mathbb{R}^n : \gamma_C(x) = 0\}$.

Proof: "$i), ii)$":

- Nonnegativity follows from the definition of γ_C.
- Positive homogeneity follows from the definition of γ_C
  $(\forall\, t > 0, \gamma_C(tx) = \inf\{\lambda > 0 : tx \in \lambda C\} = t\lambda(x))$.
- $\gamma_C(0) = 0$ since $0 \in C$.

- Convexity: Let $K_C :=$ cone $(C \times \{1\}) = \{(\lambda_C, \lambda) \in \mathbb{R}^n \times \mathbb{R} : c \in C, \lambda \geq 0\}$. K_C is convex (but need not be closed) and γ_C is given by $\gamma_C(x) = \inf\{\lambda : (x, \lambda) \in K_C\}$. $\Rightarrow \gamma_C$ is the lower-bound funtion of K_C. Theorem 6.10 $\Rightarrow \gamma_C$ is convex. Positive homogeneity $\Rightarrow \gamma_C$ is sublinear.

- $C = \{x \in \mathbb{R}^n : \gamma_C(x) \leq 1\} = S_1(\gamma_C)$. From the definition of γ_C we get if $x \in C \Rightarrow \gamma_C(x) \leq 1$. If $\gamma_C(x) \leq 1$ we have to show that $x \in C$. Let $x_k := (1 - \frac{1}{k})x$ for $k = 1, 2, \ldots$. $\gamma_C(x_k) = (1 - \frac{1}{k})\gamma_C(x) < 1 \Rightarrow \exists\, \lambda_k \in (0, 1)$ such that $x_k \in \lambda_k C \Leftrightarrow \frac{x_k}{\lambda_k} \in C$. Since C is convex, $0 \in C$ we have $\lambda_k(\frac{x_k}{\lambda_k}) + (1 - \lambda_k) \cdot 0 = x_k \in C$. Since $x_k \in C$ for $k = 1, 2, \ldots$ and C is closed $\lim\limits_{k\to\infty} x_k = x \in C$. From $C = \{x \in \mathbb{R}^n : \gamma_C(x) \leq 1\}$ we get by positive homogeneity $r\,C = \{x \in \mathbb{R}^n : \gamma_C(x) \leq r\}$. Section 2.2 $\Rightarrow C_\infty = \bigcap\{r\,C : r > 0\} = \bigcap\limits_{r>0}\{x \in \mathbb{R}^n : \gamma_C(x) \leq r\} = \{x \in \mathbb{R}^n : \gamma_C(x) = 0\} \Rightarrow$ iii) holds, and also from Proposition 6.6 γ_C is closed (sublevel sets are closed) $\Rightarrow$ i) holds.

- ii) " $\Longleftarrow$ ": $C = \{x \in \mathbb{R}^n : \gamma_C(x) \leq 1\}$. Let $0 \in \operatorname{int} C \Rightarrow$ there exists $\epsilon > 0$ such that for all $x \neq 0 : x_\epsilon := \epsilon\frac{x}{\|x\|} \in C \Rightarrow \gamma_C(x_\epsilon) \leq 1$. Positive homogeneity $\Rightarrow \gamma_C(x) = \frac{\|x\|}{\epsilon}\gamma_C(x_\epsilon) \leq \frac{\|x\|}{\epsilon}$. This holds for all $x \in \mathbb{R}^n(\gamma_C(0) = 0) \Rightarrow \gamma_C$ is finite everywhere.
  " $\Longrightarrow$ ": Suppose that γ_C is finite everywhere. By Theorem 8.2 $\Rightarrow \gamma_C$ has an upper bound $L > 0$ on the unit ball: $\|x\| \leq 1 \Rightarrow \gamma_C(x) \leq L$. iii)$\Rightarrow x \in L\,C \Rightarrow B(0, \frac{1}{L}) \subset C$.

□

Remark γ_C is the lower-bound function of the cone $K_C (= K_C + \{0\} \times \mathbb{R}_{+0})$. Section 6.3.4 $\Rightarrow K_C \subset$ epi $\gamma_C \subset$ cl K_C and since γ_C has a closed epigraph $\Rightarrow$ epi $\gamma_C =$ cl $K_C = \overline{\text{cone}}\ (C \times \{1\})$.

Corollary 11.5 *C is compact $\Longleftrightarrow \gamma_C(x) > 0$ for all $x \neq 0$.*

Proof: C is compact $\Leftrightarrow C_\infty = \{0\}$, Proposition 2.8
$C_\infty = \{x \in \mathbb{R}^n : \gamma_C(x) = 0\}$, by iii).

□

11.1 Exercises

1. Let σ be a sublinear function and $x_1, \dots, x_m \in \operatorname{dom} \sigma$ such that

$$\sigma(x_j) + \sigma(-x_j) = 0 \quad j = 1, \dots, m.$$

   Then σ is linear on

$$U = \left\{ \sum \alpha_j e_j : \alpha_j \in \mathbb{R} \right\} = \operatorname{lin} \{x_1, \dots, x_n\}.$$

2. a) Let $b_1, \dots, b_H \in \mathbb{R}^n$ and $B := \operatorname{Conv} \{b_1, \dots, b_H\}$. Show that

$$\gamma_B(x) = \min \left\{ \sum_{q=1}^{H} \lambda_q : x = \sum_{q=1}^{H} \lambda_q b_q \right\}.$$

   Furthermore, if $H = 2G$ and $b_i = -b_{i-G} \quad i = G+1, \dots, H$ (i.e. B is symmetric w.r.t. 0) then

$$\gamma_B(x) = \min \left\{ \sum_{q=1}^{G} |\lambda_q| : x = \sum_{q=1}^{G} \lambda_q b_q \right\}$$

   and γ_B is a norm.

   b) Let $f \in \overline{\operatorname{Conv}}\, \mathbb{R}^n$ be s.t. $0 \le f(tx) = t^2 f(x) \quad \forall\, x \in \mathbb{R}^n, t > 0$ then $\sqrt{f}$ is a gauge. Especially let $f = \langle x, Ax \rangle$ where A is positive semi-definite and symmetric, then the above result is true.

3. Let σ be a sublinear function. Show that

   a) $\operatorname{cl} \sigma$ is sublinear.

   b) If σ is closed then $\sigma(0) = 0$ and $\sigma'_\infty = \sigma$.

4. Let $\{\sigma_j\}_{j \in J}$ be a family of sublinear functions, all minorized by some linear function.

   a) $\sigma = \operatorname{co} (\inf \sigma_j)$ is sublinear.

   b) If $J = \{1, \dots m\}$ then $\operatorname{co}\ \min\{\sigma_1, \dots \sigma_m\} = \sigma_1 \overset{+}{\vee} \dots \overset{+}{\vee} \sigma_m$.

Chapter 12

SUPPORT FUNCTIONS

Definition 48 *$S \subset \mathbb{R}^n$, $S \neq \emptyset$. The function $\sigma_S : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by $\mathbb{R}^n \ni x \mapsto \sigma_S(x) := \sup\{\langle s, x\rangle : s \in S\}$ is called the* **support function** *of S.*

Proposition 12.1 *A support function is closed and sublinear.*

Proof: From Proposition 7.2 we get closedness and convexity. Positive homogeneity follows from positive homogeneity of $\langle \cdot, x\rangle$. Also $\sigma_S \not\equiv +\infty$, since $\sigma_S(0) = 0$.

$\square$

Proposition 12.2 *σ_S ($S \neq \emptyset$) is finite everywhere $\Longleftrightarrow$ S is bounded.*

Proof: " $\Longleftarrow$ ": Let S be bounded, $S \subset B(0, L)$ for some $L > 0$. Then $\langle s, x\rangle \leq \|s\|\|x\| \leq L\|x\|$ for all $s \in S \Rightarrow \sigma_S(x) \leq L\|x\|$ for all $x \in \mathbb{R}^n$.
" $\Longrightarrow$ ": σ_S is finite on $\mathbb{R}^n$. Theorem 8.2 $\Rightarrow \exists\, L$: $\langle s, x\rangle \leq \sigma_S(x) \leq L$ for all $(s, x) \in S \times B(0, 1)$. If $s \neq 0 \Rightarrow$ we can take $x = \frac{s}{\|s\|}$, which implies $\|s\| \leq L$. Refer to Figure 12.1 for an illustration.

$\square$

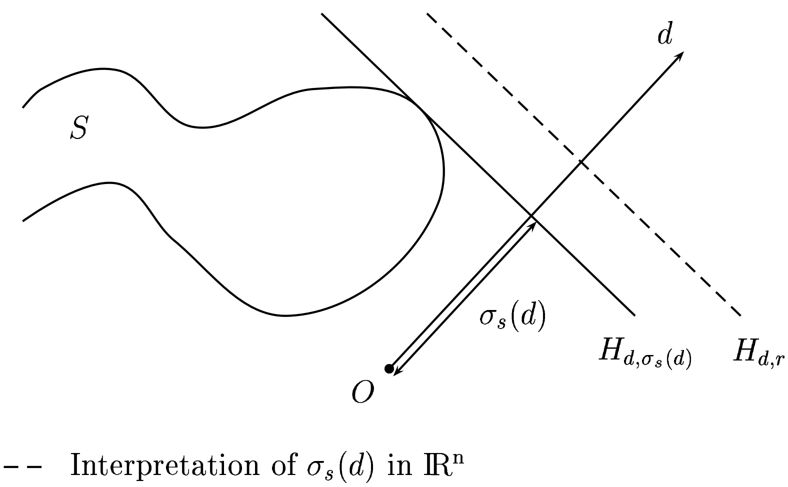


Figure 12.1: Illustration for proof of Proposition 12.2.

Proposition 12.3 *For $S \subset \mathbb{R}^n$, $S \neq \emptyset$ it holds $\sigma_S = \sigma_{\mathrm{cl}\, S} = \sigma_{\mathrm{co}\, S} = \sigma_{\overline{\mathrm{co}}\, S}$.*

Proof:

- Convexity of $\langle s, \cdot \rangle \Rightarrow \sigma_S = \sigma_{\mathrm{cl}\, S}$.
- Convexity of $\langle s, \cdot \rangle \Rightarrow \sigma_S = \sigma_{\mathrm{co}\, S}$.
- From Proposition 1.8 $\Rightarrow \sigma_S = \sigma_{\overline{\mathrm{co}}\, S}$.

□

Theorem 12.4 *Let $S \subset \mathbb{R}^n$, $S \neq \emptyset$. Then $s \in \overline{\mathrm{co}}\, S \iff (\langle s, d \rangle \leq \sigma_S(d)$ for all $d \in X)$ where X can be taken as $\mathbb{R}^n, B(0,1), \mathrm{bd}\, B(0,1)$ or $\mathrm{dom}\, \sigma_S$.*

Proof: Since σ_S is positive homogeneous the choices for X are equivalent ($0 \in \mathrm{dom}\, \sigma_S$).
" $\Longrightarrow$ ": $s \in \overline{\mathrm{co}}\, S$. Proposition 12.3 $\Rightarrow \langle s, d \rangle \leq \sigma_S(d)$ for all $d \in X$.
" $\Longleftarrow$ ": We take $X = \mathbb{R}^n$. Suppose $s \notin \overline{\mathrm{co}}\, S \Rightarrow \{s\}$ and $\overline{\mathrm{co}}\, S$ can be strictly separated (Theorem 4.1): $\exists\, d_0 \in \mathbb{R}^n : \langle s, d_0 \rangle > \sup\{\langle s', d_0 \rangle : s' \in \overline{\mathrm{co}}\, S\} = \sigma_S(d_0)$.

□

Figure 12.2 shows the relationship between closed convex sets and support functions.

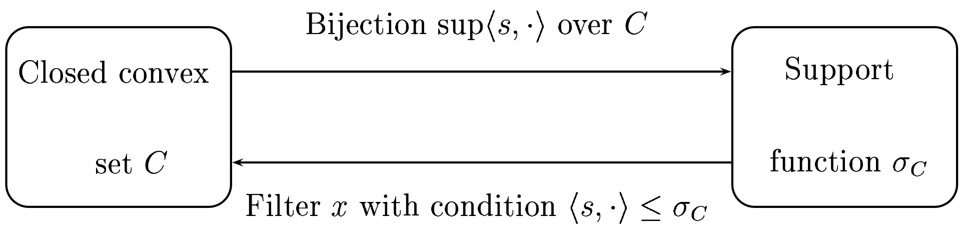


Figure 12.2: Relation between closed convex sets and support functions.

Proposition 12.5 *Let $S \subset \mathbb{R}^n$, $S \neq \emptyset$, S closed and convex. Then* $S \in \text{int } S \iff \langle s, d\rangle < \sigma_S(d)$ *for all* $d \neq 0$.

Proof: We do the proof for $\tilde{B} := \text{bd } B(0, 1)$, which is enough by posititive homogeneity.
" $\Longrightarrow$ ": For $s \in \text{int } S, \exists\, \epsilon > 0$: $s + \epsilon d \in S$ for all $d \in \tilde{B}$. Then we get from the definition of σ_S: $\sigma_S(d) \geq \langle s + \epsilon d, d\rangle = \langle s, d\rangle + \epsilon$ for all $d \in \tilde{B}$.
" $\Longleftarrow$ ": Let $s \in \mathbb{R}^n$ such that $\sigma_S(d) - \langle s, d\rangle > 0$ for all $d \in \tilde{B}$. σ_S closed, $\tilde{B}$ compact $\Rightarrow 0 < \epsilon := \inf\{\sigma_S(d) - \langle s, d\rangle : d \in \tilde{B}\} \leq +\infty \Rightarrow \langle s, d\rangle + \epsilon \leq \sigma_S(d)$ for all $d \in \tilde{B}$. Now take u with $\|u\| < \epsilon$. C.S.I. $\Rightarrow$ For all $d \in \tilde{B}\langle s + u, d\rangle = \langle s, d\rangle + \langle u, d\rangle \leq \langle s, d\rangle + \epsilon \leq \sigma_S(d)$. Theorem 12.4 $\Rightarrow s + u \in S \Rightarrow s \in \text{int } S$.

□

Remark

- If $S = \{s\}$ we get $\sigma_S = \langle s, \cdot\rangle$ a linear function. Therefore support functions (which are sublinear) can be seen as a generalization of linear functions.
- For $S = B(0, 1)$ we can compute $\sigma_S(d) \geq \langle \frac{d}{\|d\|}, d\rangle = \|d\|$ (if $d \neq 0$). For $s \in B(0, 1)$ we get $\langle s, d\rangle \leq \|s\| \cdot \|d\| \leq \|d\|$ by C.S.I. $\Rightarrow \sigma_{B(0,1)}(d) = \|d\|$.

Theorem 12.6 *Let σ be a closed sublinear function. Then σ is the support function of the nonempty closed convex set* $S_\sigma := \{s \in \mathbb{R}^n : \langle s, d\rangle \leq \sigma(d)$ *for all* $d \in \mathbb{R}^n\}$.

Proof: σ convex. Proposition 6.5 $\Rightarrow$ σ is minorized by some affine function: $\exists\ (s,r) \in \mathbb{R}^n \times \mathbb{R}$:

$$\langle s, d\rangle - r \leq \sigma(d) \text{ for all } d \in \mathbb{R}^n. \tag{12.1}$$

Since $\sigma(0) = 0 \Rightarrow r \geq 0$. By positive homogeneity, we get $\langle s, d\rangle - \frac{1}{t}\cdot r \leq \sigma(d)$ for all $d \in \mathbb{R}^n, t > 0$. By letting $t \to +\infty$ we see that

$$\langle s, d\rangle \leq \sigma(d) \text{ for all } d \in \mathbb{R}^n. \tag{12.2}$$

Inequality (12.2) is sharper than Inequality (12.1) since we can express σ as the supremum of linear functions (not general affine functions) as in Proposition 6.9 $\Rightarrow \sigma(d) = \sup\{\langle s, d\rangle : \langle s, \cdot\rangle \text{ minorizes } \sigma\}$.

□

Remark The bijection of Theorem 12.4 between support functions and closed convex sets is a bijection between closed sublinear functions and closed convex sets.

Corollary 12.7 *For a closed convex set S, $S \neq \emptyset$ and a closed sublinear function, the following are equivalent:*

i) *σ is the support function of S*

ii) $S = \{s : \langle s, d\rangle \leq \sigma(d) \text{ for all } d \in X\}$,
$X \in \{\mathbb{R}^n, B(0,1), \text{bd } B(0,1), \text{dom } \sigma\}$.

Proof: $X = \mathbb{R}^n$ is Theorem 12.6. The rest follows by positive homogeneity.

□

Definition 49 *$S \subset \mathbb{R}^n$, S closed, convex, $S \neq \emptyset$. σ the support function of S. For given $d \neq 0$, the set $F_S(d) := \{s \in S : \langle s, d\rangle = \sigma(d)\}$ is called the* **exposed face of *S*** *associated with d, or the* **face exposed by *d***. *(We may allow $d = 0$ to get $S = F_S(0)$). (See Section 2.3).*

Proposition 12.8 *For $s \in S$, $S \neq \emptyset$, S closed convex, it holds that*

$$s \in F_S(d) \Longleftrightarrow d \in N_S(s).$$

Proof: Restatement of Proposition 5.7.

□

Proposition 12.9 *Let $S \neq \emptyset$, S closed, convex. Then*

$$\text{bd } S = \bigcup\{F_S(d) : d \in X\}.\ X \in \{\mathbb{R}^n \setminus \{0\}, \text{bd } B(0,1), \text{dom } \sigma_S \setminus \{0\}\}.$$

Proof: First note that $F_S(d)$ does not depend on $\|d\| \Rightarrow X = \mathbb{R}^n \setminus \{0\}$ is the same as $X = \text{bd } B(0,1)$. Since $F_S(d) = \emptyset$ for $d \notin \text{dom } \sigma_S$, the choice $X = \text{dom } \sigma_S \setminus \{0\}$ is also equivalent. If $s \in \text{int } S$, $d \neq 0$ then $s + \epsilon d \in S$ and s cannot be a maximizer of $\langle \cdot, d\rangle (\langle s + \epsilon d, d\rangle = \langle s, d\rangle + \epsilon\|d\| \geq \langle s, d\rangle) \Rightarrow s \notin F_S(d)$. For $s \in \text{bd } S \Rightarrow N_S(s)$ contains a vector $d \neq 0$. Proposition 12.8 $\Rightarrow s \in F_S(d)$.

□

Proposition 12.10 *Let C be a closed, convex set, $0 \in C$. Then γ_C is the support function of a closed convex set containing the origin. Namely*

$$C^\circ := \{s \in \mathbb{R}^n : \langle s, d\rangle \leq 1 \textit{ for all } d \in C\},$$

where C° is called the **polar (set) of C**.

Proof: By Theorem 11.4 we know that γ_C is closed, sublinear and nonnegative. Theorem 12.6 $\Rightarrow \gamma_C$ is the support function of some closed set D, with $0 \in D$, $D = \{s \in \mathbb{R}^n : \langle s, d\rangle \leq r$ for all $(d, r) \in \text{epi } \gamma_C\}$. From the remarks to Theorem 11.4, we know that $\text{epi } \gamma_C = \overline{\text{cone}}\ C \times \{1\} \Rightarrow$ we can use positive homogenity to write $D = \{s \in \mathbb{R}^n : \langle s, d\rangle \leq 1$ for all d such that $\gamma_C(d) \leq 1\}$ $D = \{s \in \mathbb{R}^n : \langle s, d\rangle \leq 1$ for all $d \in C\}$, by Theorem 11.4 iii) $\Rightarrow D = C^\circ$.

□

Corollary 12.11 *Let C be a closed convex set containing the origin. Then $\sigma_C = \gamma_{C^\circ}$.*

Proof: From Proposition 12.10 $\Rightarrow \gamma_{C^\circ} = \sigma_{C^{\circ\circ}}$. To show $(C^\circ)^\circ = C$:

1) $C^\circ = \{s \in \mathbb{R}^n : \langle s, d\rangle \leq r$ for all $(d, r) \in \text{epi } \gamma_C\}$.

2) $(\text{epi } \gamma_C)^\circ = \{(d', r') : \langle d, d'\rangle + rr' \leq 0$ for all $(d, r) \in \text{epi } \gamma_C\}$.

   1) and 2)

$$\begin{aligned} \Rightarrow C^\circ &= p_{\mathbb{R}^n}((\text{epi } \gamma_C)^\circ \cap (\mathbb{R}^n \times \{-1\})) \\ C^\circ &= p_{\mathbb{R}^n}((\text{epi } \gamma_{C^\circ}) \cap (\mathbb{R}^n \times \{1\})), \end{aligned}$$

   by definition.

Analogously, we get

$$\begin{aligned}(C^\circ)^\circ &= p_{\mathbb{R}^n}((\text{epi}\,\gamma_{C^\circ})^\circ \cap (\mathbb{R}^n \times \{-1\})) \\ &= p_{\mathbb{R}^n}((\text{epi}\,\gamma_C)^{\circ\circ} \cap (\mathbb{R}^n \times \{1\}))\ \text{(by (3))} \\ &= p_{\mathbb{R}^n}((\text{epi}\,\gamma_C) \cap (\mathbb{R}^n \times \{1\}))\ \text{(as epi } \gamma_C \text{ is closed)} \\ &= C,\ \text{by definition.}\end{aligned}$$

□

12.1 Exercises

1. Let S be a nonempty closed convex set.

$$\text{cl dom}\,\sigma_S = (S_\infty)^0.$$

2. Let $|||\ \ |||: \mathbb{R}^n \to \mathbb{R}$ be a norm and define $B := \{x \in \mathbb{R}^n : |||x||| \leq 1\}$ and $B^* := \{s \in \mathbb{R}^n : \langle s, x\rangle \leq |||x||| \quad \forall\, x \in \mathbb{R}^n\}$.

   Then $\sigma_B = \gamma_{B^*} = |||\ \ |||^*$ defined by

$$|||\ \ |||^* = \max\{\langle s, x\rangle : |||x||| \leq 1\}.$$

   Furthermore, $|||\ \ |||^*$ is a norm, the so-called dual norm.

Chapter 13

THE SUBDIFFERENTIAL

In this section $f : \mathbb{R}^n \to \mathbb{R}$ is convex, i.e. f is continuous and locally Lipschitz continuous.

13.1 Definitions

Definition 50 *Let $x, d \in \mathbb{R}^n$. Then the* **difference quotient** *of f at x in direction d is defined as $q(t) := \frac{f(x+td)-f(x)}{t}$ for $t > 0$. $t \mapsto q(t)$ is increasing (the Criterion of Increasing Slopes [3]), $q(t)$ is bounded near 0. The* **directional derivative** *of f at x in direction d is $f'(x,d) = \lim_{t \downarrow 0} q(t) = \inf_{t>0} q(t)$.*

Remark If we look at $\phi : \mathbb{R} \to \mathbb{R}$, $t \mapsto \phi(t) = f(x + td)$ then

$$\begin{aligned} f'(x,d) &= \inf_{t>0} \frac{f(x+td) - f(x)}{t} \\ &= \inf_{t>0} \frac{\phi(t) - \phi(0)}{t - 0} \\ &= D_+\phi(0) \end{aligned}$$

is the right derivative of ϕ at 0. Moreover,

$$\begin{aligned} f'(x,d) &= \inf_{t>0} \frac{f(x-td) - f(x)}{t} \\ &= \sup_{t<0} \frac{f(x+td) - f(x)}{-t} \\ &= \sup_{t<0} \frac{\phi(t) - \phi(0)}{-t - 0} \\ &= -D_-\phi(0) \end{aligned}$$

is the negative of the left derivative of ϕ at 0.

Proposition 13.1 *Let $x \in \mathbb{R}^n$. Then $f'(x,\cdot)$ is sublinear finite.*

Proof:

- Convexity: Let $d_1, d_2 \in \mathbb{R}$, $\alpha_1, \alpha_2 \geq 0$ with $\alpha_1 + \alpha_2 = 1$. From convexity of f we get

$$\begin{aligned} & f(x + t(\alpha_1 d_1 + \alpha_2 d_2)) - f(x) \\ & \quad = f(\alpha_1(x + td_1) + \alpha_2(x + td_2)) \\ & \qquad\quad -\alpha_1 f(x) - \alpha_2 f(x) \\ & \quad \leq \alpha_1(f(x + td_1) - f(x)) + \alpha_2(f(x + td_2) - f(x)). \end{aligned}$$

  Dividing by t and taking $t \to 0$ we get

$$f'(x, \alpha_1 d_1 + \alpha_2 d_2) \leq \alpha_1 f'(x, d_1) + \alpha_2 f'(x, d_2).$$

- Positive homogeneity: Let $\lambda > 0$:

$$\begin{aligned} f'(x, \lambda d) &= \lim_{t \searrow 0} \lambda \left(\frac{f(x + \lambda t d) - f(x)}{\lambda t} \right) \\ &= \lambda \lim_{\tau \searrow 0} \frac{f(x + \tau d) - f(x)}{\tau} \\ &= \lambda f'(x, d). \end{aligned}$$

- Finitenesss: Let $\|d\| = 1$. f is Lipschitz continuous around x:

$$\begin{aligned} & \exists\, \epsilon > 0, L > 0 : \ |f(x + td) - f(x)| \leq Lt \ \forall\, 0 \leq t \leq \epsilon \\ & \quad \Rightarrow |f'(x, d)| \leq L. \text{ By positive homogeneity} \\ & \quad \Rightarrow |f'(x, d)| \leq L\|d\| \text{ for all } d \in \mathbb{R}^n. \end{aligned}$$

□

From Proposition 13.1 and Theorem 12.6, $f'(x,\cdot)$ is a support function for some nonempty, compact, convex set.

Definition 51 (Subdifferential I) *The* **subdifferential $\partial f(x)$** *of f at x is the nonempty compact convex set of $\mathbb{R}^n$ whose support function is $f'(x,\cdot)$:*

$$\partial f(x) := \{s \in \mathbb{R}^n : \langle s, d \rangle \leq f'(x, d) \ \ \forall\, d \in \mathbb{R}^n\}. \tag{SI}$$

A vector $s \in \partial f(x)$ is called a **subgradient** *of f at x.*

Remark All results of the correspondence between compact convex sets and finite sublinear functions can be applied.

Proposition 13.2 *For $d \mapsto \sigma(d) := f'(x,d)$ the following hold:*

i) $\sigma'(0,\delta) = f'(x,\delta)$ *for all* $\delta \in \mathbb{R}^n$.

ii) $\sigma(\delta) = \sigma(0) + \sigma'(0,\delta) = \sigma'(0,\delta)$ *for all* $\delta \in \mathbb{R}^n$.

iii) $\partial\sigma(0) = \partial f(x)$.

Proof: "i) ii)": From the positive homogeneity of σ and since $\sigma(0) = 0$ (see Chapter 11)

$$\begin{aligned} \Rightarrow \sigma'(0,\delta) &= \frac{\sigma(t\delta) - \sigma(0)}{t} \\ &= \sigma(\delta) \\ &= f'(x,\delta) \text{ for all } t > 0 \Rightarrow \text{ i) and ii).} \end{aligned}$$

"iii)":

$$\begin{aligned} \partial f(x) &= \{s : \langle s,d\rangle \le f'(x,d) \text{ for all } d \in \mathbb{R}^n\} \\ \partial\sigma(0) &= \{s : \langle s,d\rangle \le \sigma'(0,d) = f'(x,d) \text{ for all } d \in \mathbb{R}^n\}, \end{aligned}$$

and the supported set is unique $\Rightarrow$ iii).

□

Geometric Interpretation

gr $f'(x,\cdot)$ is made up of all half-lines tangent to gr f at $(x, f(x))$. These are the same as the half-lines tangent to gr $f'(x,\cdot)$

Definition 52 (Subdifferential II)

$$\partial f(x) := \{s \in \mathbb{R}^n : f(y) \ge f(x) + \langle s, y - x\rangle \textit{ for all } y \in \mathbb{R}^n\}. \qquad \text{(SII)}$$

Remark

- In SII, $f(x)$ is compared to $f(y)$ for all $y \in \mathbb{R}^n$. (This is not explicitly done in SI).
- Difference between SII and differentiation:

  - SII is global, whereas differentiation is local.
  - In differentiation, a remainder term $o(\|x-y\|)$ is added.

- In SII all $s \in \partial f(x)$ are described at the same time. For $d \neq 0$ (by Proposition 12.9) $f'(x,d) = \langle s_d, d\rangle$ plots only bd $\partial f(x)$ and only one exposed face at a time.

Theorem 13.3 *Definitions SI and SII are equivalent.*

Proof: Let s satisfy SI $\Leftrightarrow \langle s,d\rangle \leq f'(x,d)$ for all $d \in \mathbb{R}^n \Leftrightarrow \langle s,d\rangle \leq \frac{f(x+td)-f(x)}{t}$ for all $d \in \mathbb{R}^n, t > 0$. Now use $y = x+td$, $d = \frac{y-x}{t}$, y will attain all values in $\mathbb{R}^n \Leftrightarrow \langle s, \frac{y-x}{t}\rangle \leq \frac{f(y)-f(x)}{t} \Leftrightarrow \langle s, y-x\rangle \leq f(y)-f(x)$ $\Leftrightarrow f(y) \leq f(x) + \langle s, y-x\rangle$ for all $y \in \mathbb{R}^n \Leftrightarrow$ SII.

□

Remark

- We can get S2 → directional derivative. In SII we take $y = x + \delta\frac{s}{\|s\|}$, $\delta > 0 \Rightarrow f(x) + L(\delta) \geq f(y) \geq f(x) + \delta\|s\| \Rightarrow f(x)$ has a finite support function $\Rightarrow$ this support function is $f'(x,\cdot)$.

- A finite sublinear function is convex $\Longrightarrow \sigma$ has a subdifferential $\partial\sigma(0) = \{s \in \mathbb{R}^n : \langle s,d\rangle \leq \sigma(d)$ for all $d \in \mathbb{R}^n\}$, which is Theorem 12.4 $\Rightarrow$ To construct the convex set associated with a support function: A finite sublinear function is the support of its subdifferential at 0 ($\sigma = \sigma_{\partial\sigma(0)}$), as summarized in Figure 13.1.

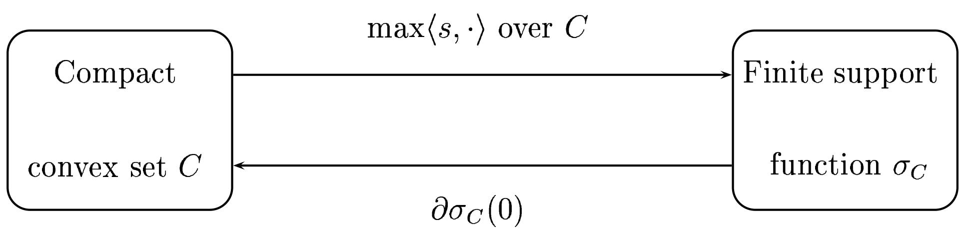


Figure 13.1: Relation between compact convex sets and finite support functions.

13.2 Geometric Constructions

From SI we know that the elements of $\partial f(x)$ are the slopes of the hyperplanes supporting f at $(x, f(x)) \in \mathbb{R}^n \times \mathbb{R}$.

Proposition 13.4

i) $s \in \mathbb{R}^n \in \partial f(x) \iff (s, -1) \in \mathbb{R}^n \times \mathbb{R}$ *is in* $N_{\text{epi } f}(x, f(x))$*, or* $N_{\text{epi } f}(x, f(x)) = \{(\lambda s, -\lambda) : s \in \partial f(x), \lambda \geq 0\}$.

ii) $T_{\text{epi } f}(x, f(x))$ *is the epigraph of the directional derivative of* $d \mapsto f'(x, d)$*, or* $T_{\text{epi } f}(x, f(x)) = \{(d, r) : r \geq f'(x, d)\}$.

Proof:

i) From the definition of the normal cone $(s, -1) \in N_{\text{epi } f}(x, f(x))$

$$\begin{aligned}
&\Leftrightarrow \quad \langle s, y - x\rangle + (-1)(r - f(x)) \leq 0 \ \forall\, y \in \mathbb{R}^n,\ r \geq f(y) \\
&\Leftrightarrow \quad r \geq f(x) + \langle s, y - x\rangle \ \forall\, y \in \mathbb{R}^n,\ r \geq f(y) \\
&\Leftrightarrow \quad f(y) \geq f(x) + \langle s, y - x\rangle \ \forall\, y \in \mathbb{R}^n, r \geq f(y)
\end{aligned}$$

$\Leftrightarrow s \in \partial f(x)$ since $0 \in N_{\text{epi } f}(x, f(x))$ and this is a cone. i) follows.

ii) We use Corollary 5.5 to get $T_{\text{epi } f}(x, f(x)) = (N_{\text{epi } f}(x, f(x)))^\circ$

$$\begin{aligned}
&\Leftrightarrow T_{\text{epi } f}(x, f(x)) \\
&\quad = \quad \{(d, r) \in \mathbb{R}^n \times \mathbb{R} : \langle \lambda s, d\rangle + r\,(-\lambda) \leq 0 \ \forall\, s \in \partial f(x), \lambda \geq 0\} \\
&\quad \overset{\lambda \neq 0}{=} \quad \{(d, r) \in \mathbb{R}^n \times \mathbb{R} : \langle s, d\rangle \leq r \ \forall\, s \in \partial f(x)\} \\
&\quad = \quad \{(d, r) \in \mathbb{R}^n \times \mathbb{R} : r \geq \max\{\langle s, d\rangle : s \in \partial f(x)\} \\
&\quad = \quad f'(x, d)\}.
\end{aligned}$$

The result follows also in the case $\lambda = 0$. Refer to Figure 13.2.

□

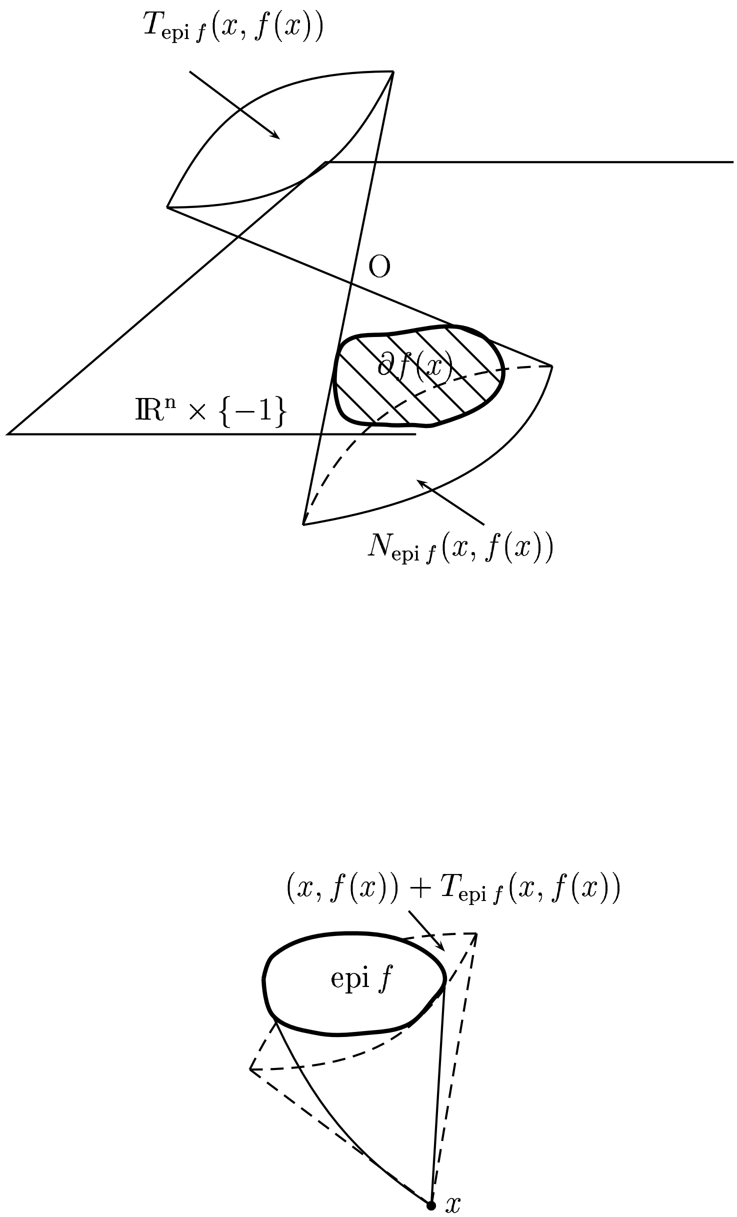


Figure 13.2: Illustration for proof of Proposition 13.4.

Let $Sf(x) := S_{f(x)}(f) = \{y \in \mathbb{R}^n : f(y) \leq f(x)\}$.

Lemma 13.5 *Let $f \in \mathbb{R}^n \to \mathbb{R}$ be convex. Then $T_{Sf(x)}(x) \subset \{d : f'(x,d) \leq 0\}$.*

Proof: Take arbitrary $y \in Sf(x)$, $t > 0$ and set $d := t(y - x)$. From the definition of $f'(x,d)$

$$0 \geq t(f(y) - f(x)) = t(f(x + \frac{d}{t}) - f(x)) \geq f'(x,d)$$

$$\Rightarrow \mathbb{R}_{+0}(Sf(x) - x) \subset \{d : f'(x,d) \leq 0\}. \tag{13.1}$$

(For $d = 0$, note that $0 \in Sf(x) - x$). We know that $f'(x,\cdot)$ is a closed function $\Rightarrow \{d : f'(x,d) \leq 0\}$ is closed. From Proposition 5.3 $\Rightarrow T_{Sf(x)}(x) = \mathrm{cl}\,(\mathbb{R}_{+0}(Sf(x) - x))$. The result follows by taking the closure in (13.1).

□

Remark The converse inclusion in Lemma 13.5 is not true in general. Let $f(x) = \frac{1}{2}\|x\|^2$. $Sf(0) = 0, f'(0,d) = 0$ for all $d \Rightarrow T_{Sf(x)}(x) = \{0\} \subsetneq \{d : f'(x,d) \leq 0\} = \mathbb{R}^n$.

Proposition 13.6 *Let $g : \mathbb{R}^n \to \mathbb{R}$ be convex and suppose $g(x_0) < 0$ for some $x_0 \in \mathbb{R}^n$. Then*

i) $\mathrm{cl}\,\{\xi : g(\xi) < 0\} = \{\xi : g(\xi) \leq 0\}$

ii) $\{\xi : g(\xi) < 0\} = \mathrm{int}\,\{\xi : g(\xi) \leq 0\}$. *It follows*

iii) $\mathrm{bd}\,\{\xi : g(\xi) \leq 0\} = \{\xi : g(\xi) = 0\}$.

Proof:

- g is finite, convex $\Rightarrow g$ is l.s.c $\Rightarrow \mathrm{cl}\,\{\xi : g(\xi) < 0\} \subset \{\xi : g(\xi) \leq 0\}$.
- Take $\bar{\xi}$ arbitrary with $g(\bar{\xi}) \leq 0$ and set for $k > 0$ $z_k := \frac{1}{k}x_0 + (1 - \frac{1}{k})\bar{\xi}$. Since g is convex: $g(z_k) \leq \frac{1}{k}g(x_0) + (1 - \frac{1}{k})g(\bar{\xi}) < 0 \Rightarrow$ with $k \to +\infty$ $\bar{\xi} \in \mathrm{cl}\,\{\xi : g(\xi) < 0\} \Rightarrow$ i) holds.
- Take the interior of both sides of i). From Proposition 2.3, "int cl " = "int " and the "int " operator has no effect because g is continuous $\Rightarrow$ ii) holds.

□

Theorem 13.7 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be convex,* $0 \notin \partial f(x)$. *Then*

i) $T_{Sf(x)}(x) = \{d \in \mathbb{R}^n : f'(x,d) \leq 0\}$.

ii) $\text{int } T_{Sf(x)}(x) = \{d \in \mathbb{R}^n : f'(x,d) < 0\} \neq 0$.

Proof: From SI: $0 \notin \partial f(x) \Rightarrow \exists\, d : f'(x,d) < 0 \Rightarrow f(x+td) < f(x)$ for $t > 0$ small enough. Rewrite d as $\frac{x+td-x}{t}$ with $x + td \in Sf(x)$ and we get

$$\{d : f'(x,d) < 0\} \subset \mathbb{R}_{+0}(Sf(x) - x) \subset T_{Sf(x)}. \tag{13.2}$$

From Proposition 13.6 for $g = f'(x,\cdot) \Rightarrow \text{cl } \{d : f'(x,d) < 0\} = \{d : f'(x,d) \leq 0\} \Rightarrow$ part i) follows by taking the closure of both sides of (13.2), and using Lemma 13.5. Now take the interior in i) and apply ii) of Proposition 13.6 with $g = f'(x,\cdot)$ to show part ii) of the theorem.

□

This result can also be formulated in terms of normal cones.

Theorem 13.8 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be convex,* $0 \notin \partial f(x)$. *Then*
d is normal to $Sf(x)$ at x $\Longleftrightarrow$
$\exists\, t \geq 0$ *and* $s \in \partial f(x)$ *such that* $d = ts : N_{Sf(x)}(x) = \mathbb{R}_{+0}\partial f(x)$.

Proof: (Refer to Figure 13.3). Write part i) of Theorem 13.7 as

$$\begin{aligned} T_{Sf(x)}(x) &= \{d \in \mathbb{R}^n : \langle s, d\rangle \leq 0\ \forall\, s \in \partial f(x)\} \\ &= \{d \in \mathbb{R}^n : \langle \lambda s, d\rangle \leq 0\ \forall\, \lambda \geq 0, s \in \partial f(x)\} \\ &= (\mathbb{R}_{+0}\partial f(x))^\circ. \end{aligned}$$

Now take the polar on both sides

$$\begin{aligned} \Rightarrow (T_{Sf(x)}(x))^\circ &= (\mathbb{R}_{+0}\partial f(x))^{\circ\circ} \\ \Leftrightarrow N_{Sf(x)}(x) &= \mathbb{R}_{+0}\partial f(x). \end{aligned}$$

□

Remark The assumption $0 \notin \partial f(x)$ can be formulated in different equivalent ways:

- In view of SI it means $f'(x, d_0) < 0$ for some d_0.
- In view of SII it means $\exists\, x_0$ such that $f(x_0) < f(x)$.

- This implies that the assumption also holds for every y with $f(y) = f(x)$.

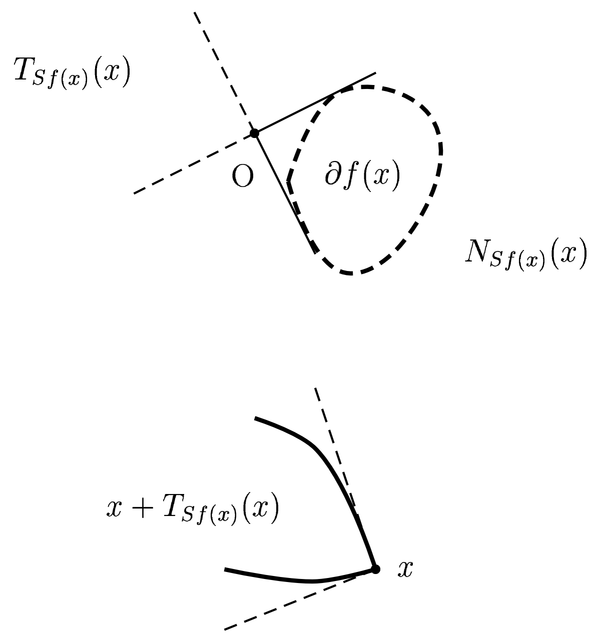


Figure 13.3: Illustration for proof of Theorem 13.8.

13.3 First-Order Developments

Lemma 13.9 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be convex,* $x \in \mathbb{R}^n$*. For any* $\epsilon > 0$*, there exists* $\delta > 0$ *such that*

$$\|h\| \leq \delta \Longrightarrow |f(x+h) - f(x) - f'(x,h)| \leq \epsilon\|h\|. \tag{13.3}$$

Proof: (By contradition) Suppose there exists $\epsilon > 0$ and a sequence $\{h_k\}$ with $\|h_k\| =: t_k \leq \frac{1}{k}$ such that $|f(x+h_k) - f(x) - f'(x,h_k)| \leq \epsilon t_k$ for $k = 1, 2, \ldots$. We can assume that $\frac{h_k}{t_k} \to d$ for some d, $\|d\| = 1$ (if necessary extract a subsequence). Now take a local Lipschitz constant L of f (as in the proof of Proposition 13.1) and compute:

$$\begin{aligned}\epsilon t_k &< |f(x+h_k) - f(x) - f'(x,h_k)| \\ &\leq |f(x+h_k) - f(x+t_k d)| + |f(x+t_k d) - f(x) \\ &\quad -f'(x,t_k d)| + |f'(x,t_k d) - f'(x,h_k)| \\ &\leq 2L\|h_k - t_k d\| + |f(x+t_k d) - f(x) - t_k f'(x,d)|.\end{aligned}$$

Divide both sides by $t_k > 0$ and pass to the limit $(k \to +\infty)$ $\Rightarrow \epsilon \leq 2L\|d-d\| + |\frac{f(x+t_k d)-f(x)}{t_k} - f'(x,d)| = 0$. The contradiction establishes the result.

□

The Inequality (13.3) of Lemma 13.9 can also be written as

$$f(x+h) = f(x) + f'(x,h) + o(\|h\|), \tag{13.4}$$

or $\lim\limits_{\alpha \searrow 0,\, d' \to d} \frac{f(x+td')-f(x)}{t} = f'(x,d)$.

Definition 53 *We call a function* $D : \mathbb{R}^n \to \mathbb{R}$ *a* **derivative** *of* $f : \mathbb{R}^n \to \mathbb{R}$ *if*

$$\frac{f(x+h) - f(x) - D(h)}{\|h\|} \to 0 \text{ for } h \to 0. \tag{13.5}$$

Depending on the type of convergence for h we distinguish between:

i) f is differentiable in the view of Gâteaux at x: (13.5) holds for $h = td$, d fixed in $\mathbb{R}^n$, $t \to 0$.

ii) f is differentiable in the view of Fréchet at x: (13.5) holds for $h \to 0$ arbitrary

iii) f has a directional derivative at x: (13.5) holds as in i) but with $t > 0$.

iv) f is differentiable in the view of Dini at x: (13.5) holds with $h = td'$, $t \searrow 0$, $d' \to d \in \mathbb{R}^n$.

If f is Lipschitzian around x we get from Lemma 13.9 (even without convexity) that these four concepts coincide.

Corollary 13.10 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be convex. At any* $x \in \mathbb{R}^n$ *we have*

$$f(x+h) = f(x) + \langle s, h \rangle + o(\|h\|)$$

whenever one of the following properties holds:

i) $s \in F_{\partial f(x)}(h)$.

ii) $h \in N_{\partial f(x)}(s)$.

iii) $s = p_{\partial f(x)}(s+h)$. *Moreover, i)* $\iff$ *ii)* $\iff$ *iii).*

Proof: By definition, $\partial f(x) = \{s \in \mathbb{R}^n : \langle s, d\rangle \leq f'(x,d)$ for all $d \in \mathbb{R}^n\}$. Since $\partial f(x)$ is compact we can write (see Chapters 11, 12) $f'(x,d) = \max\{\langle s,d\rangle : s \in \partial f(x)\}$. From Chapter 12 $\Rightarrow \exists\ s_d \in F_{\partial f(x)}(d)$ such that $f'(x,d) = \langle s_d, d\rangle$. This is equivalent to having $d \in N_{\partial f(x)}(d)$ (by Proposition 12.8). This is also equivalent, by Proposition 5.7, to $s_d = p_{\partial f(x)}$. Therefore the result follows directly from Lemma 13.9 by setting $d = h$. See Figure 13.4 for an illustration.

□

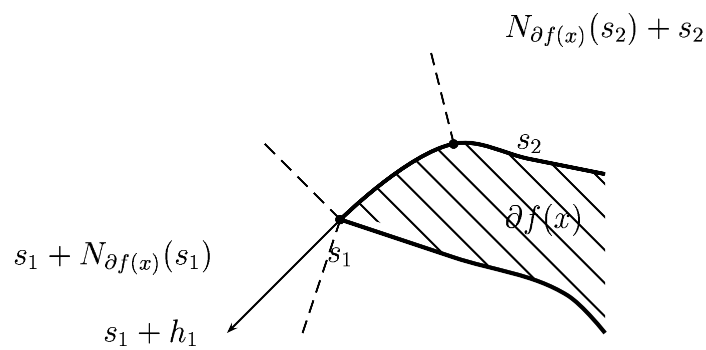


Figure 13.4: Illustration for proof of Corollary 13.10.

Corollary 13.11 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be convex. If* f *is (Gâteaux)-differentiable at* $x \implies \partial f(x) = \{s\} = \nabla f(x)$. *If* $\partial f(x) = \{s\} \implies f$ *is (Fréchet)-differentiable at* x *with* $\nabla f(x) = s$. *Moreover, by the local Lipschitz property* *of f around x, all the 4 differentiability concepts coincide and we can simply say:*

$$\partial f(x) = \{s\} \Longleftrightarrow f \textit{ is differentiable at } x \textit{ with } \nabla f(x) = s.$$

Proof: Let $\lim\limits_{t\to 0} \frac{f(x+td)-f(x)}{t} \langle s, x\rangle$ for all $d \in \mathrm{I\!R}$, by Corollary 13.10. Local Lipschitzian property $\Rightarrow F_{\partial f(x)}(d) = \{s\} \Leftrightarrow f(x+h) = f(x) + \langle s, h\rangle + o(\|h\|)$ which means Fréchet-differentiabililty.

□

Proposition 13.12 *Let $f : \mathrm{I\!R}^n \to \mathrm{I\!R}$ be convex. For all $x, d \in \mathrm{I\!R}^n$ we have $F_{\partial f(x)}(d) = \partial[f'(x,\cdot)](d)$.*

Proof:

- "$F_{\partial f(x)}(d) \subset \partial[f'(x,\cdot)](d)$": If $s \in \partial f(x)$ then $f'(x, d') \geq \langle s, d'\rangle$ for all $d' \in \mathrm{I\!R}^n$ since $f'(x,\cdot)$ is the support function of $\partial f(x)$. If $\langle s, d\rangle = f'(x, d)$ ($s \in F_{\partial f(x)}(d)$)we have $f'(x, d') \geq f'(x, d) + \langle s, d' - d\rangle$ for all $d' \in \mathrm{I\!R}^n \Rightarrow s \in \partial[f'(x,\cdot)](d)$, by definition.
- "$\partial[f'(x,\cdot)](d) \subset F_{\partial f(x)}(d)$": Let s satisfy

  $$f'(x, d') \geq f'(x, d) + \langle s, d' - d\rangle \text{ for all } d' \in \mathrm{I\!R}^n (\text{i.e. } s \in \partial[f'(x,\cdot)](d)). \tag{13.6}$$

  Set $d'' := d' - d$. Subadditivity

  $$\begin{aligned} &\Rightarrow\ f'(x, d) + f'(x, d'') \geq f'(x, d') \geq f'(x, d) + \langle s, d''\rangle \ \text{ for all } d'' \in \mathrm{I\!R}^n \\ &\Rightarrow\ f'(x,\cdot) \geq \langle s, \cdot\rangle \\ &\Rightarrow\ s \in \partial f(x). \end{aligned}$$

  Moreover, with $d' = 0$ in Inequality (13.6), we get $\langle s, d\rangle \geq f'(x, d) \Rightarrow s \in F_{\partial f(x)}(d)$. See Figure 13.5 for an illustration.

□

Example 13.1 *For $t > 0$, $\partial[f'(x,\cdot)](td)$ does not depend on t. Proposition 13.4 $\Rightarrow \partial[f'(x,\cdot)](0) = F_{\partial f(x)}(0) - \partial f(0)$.*

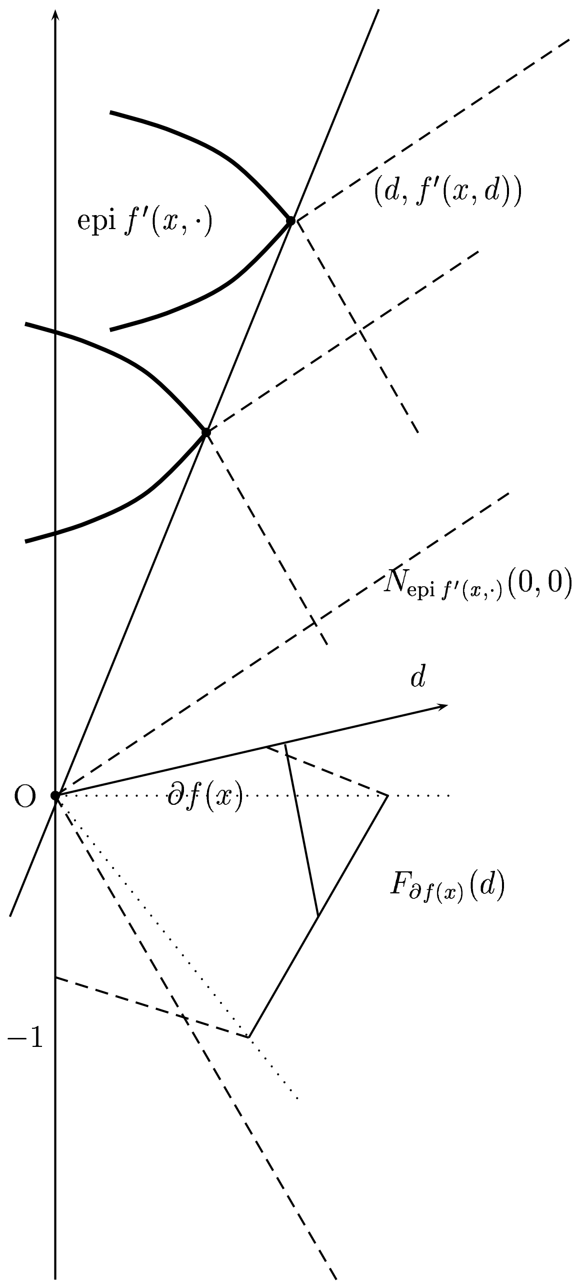


Figure 13.5: Illustration for proof of Proposition 13.12.

13.4 Minimality Conditions

Theorem 13.13 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. Then the following properties are equivalent:*

i) *f is minimized at x over $\mathbb{R}^n$: $f(y) \geq f(x)$ for all $y \in \mathbb{R}^n$.*

ii) $0 \in \partial f(x)$.

iii) *$f'(x, d) \geq 0$ for all $d \in \mathbb{R}^n$.*

Proof: "*i) $\iff$ ii)*": (directly from SII) $f(y) \geq f(x) + \langle 0, y - x\rangle$ for all $y \in \mathbb{R}^n$.
"*ii) $\iff$ iii)*": (directly from SI) $\langle 0, d\rangle \leq f'(x, d)$ for all $d \in \mathbb{R}^n$.

□

Remark

- We get that a local minimum (iii)) is always a global minimum (ii)).
- In the differentiable case we get $\nabla f(x) = \{0\}$.

Proposition 13.14 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and $x \in \mathbb{R}^n$ be a minimizer of f. For all $\epsilon > 0$ there exists $\delta > 0$ such that*

i) $h \in N_{\partial f(x)}(0) \cap B(0, \delta) \implies f(x+h) \leq f(x) + \epsilon\|h\|$. *On the other hand*

ii) $h \notin N_{\partial f(x)}(0) \implies f(x + h) > f(x)$.

Proof: "i)": $h \in N_{\partial f(x)}(0) \Leftrightarrow \langle s, h\rangle \leq 0$ for all $s \in \partial f(x)$. Theorem 13.13 iii) $\Rightarrow$ $f'(x, h) = 0$. Lemma 13.9 $\Rightarrow$ $f(x + h) = f(x) + o(\|h\|)$ $\Rightarrow$ For all $\epsilon > 0 \; \exists \; \delta > 0$: $f(x + h) \leq f(x) + \epsilon\|h\|$.
"ii)": $h \notin N_{\partial f(x)}(0) \Leftarrow f'(x, h) > 0$. Lemma 13.9 $\Rightarrow$ $f(x + h) = f(x) + f'(x, d) + \epsilon\|h\| \Rightarrow f(x + h) > f(x)$.

□

Proposition 13.15 *There exists ϵ such that $f(x + h) \geq f(x) + \epsilon\|h\|$ for all $h \in \mathbb{R}^n \iff 0 \in \text{int}\, \partial f(x)$.*

Proof: $0 \in \text{int}\, \partial f(x) \Leftrightarrow \exists\; \epsilon > 0$: $B(0, \epsilon) \subset \partial f(x) \Leftrightarrow f'(x, \cdot) \geq \epsilon\| \cdot \|$ $\Leftrightarrow f(x + h) \geq f(x) + \epsilon\|h\|$ for all $h \in \mathbb{R}^n$, by the definition of f'.

□

Remark With Proposition 13.15, we can only have a direction $d \neq 0$, $f'(x, d) = 0$ if $0 \in \text{bd}\, \partial f(x)$. Thoerem 13.7 characterizes the set of descent directions.

13.5 Mean-Value Theorems

We use $\phi(t) := f(ty+(1-t)x)$ for all $t \in [0,1]$, where $f : \mathbb{R}^n \to \mathbb{R}$ is convex (hence $\phi : \mathbb{R} \to \mathbb{R}$ is convex), and $x_t := ty + (1-t)x$ for fixed $x, y \in \mathbb{R}^n$.

Lemma 13.16 $\partial\phi(t) = \{\langle s, y-x\rangle : s \in \partial f(x_t)\} = \langle \partial f(x_t), y-x\rangle$.

Proof: Let

$$\begin{aligned} D_+\phi(t) &:= \lim_{\tau\downarrow 0} \frac{f(x_t + \tau(y-x)) - f(x_t)}{\tau} \\ &= f'(x_t, y-x), \\ D_-\phi(t) &:= \lim_{\tau\uparrow 0} \frac{f(x_t + \tau(y-x)) - f(x_t)}{\tau} \\ &= -f'(x_t, -(y-x)). \end{aligned}$$

With

$$\begin{aligned} f'(x_t, y-x) &= \max_{s\in\partial f(x_t)} \langle s, y-x\rangle, \\ -f'(x_t, -(y-x)) &= \min_{s\in\partial f(x_t)} \langle s, y-x\rangle, \end{aligned}$$

we get $\partial\phi(t) = [D_-\phi(t), D_+\phi(t)] = \{\langle s, y-x\rangle : s \in \partial f(x)\}$.

□

Theorem 13.17 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. For $x, y \in \mathbb{R}^n, x \neq y$, there exists $t \in (0,1)$ and $s \in \partial f(x_t)$ such that*

$$f(y) - f(x) = \langle s, y-x\rangle$$

or

$$f(y) - f(x) \in \bigcup_{t\in(0,1)} \{\langle \partial f(x_t), y-x\rangle\}.$$

Proof: We reformulate in terms of $\phi : \mathbb{R} \to \mathbb{R}$ using Lemma 13.16. To show: $\phi(1) - \phi(0) \in \partial\phi(t)$ for some $t \in [0,1]$. Define $\psi(t) := \phi(t) - \phi(0) - (\phi(1) - \phi(0)) \cdot t \Rightarrow \psi$ is convex and continuous on $[0,1]$. Moreover, $\psi(0) = 0$ and $\psi(1) = 0$ by construction $\Rightarrow \exists\, t \in (0,1)$ such that ψ is minimized $\Rightarrow$ For this t we have $0 \in \partial\psi(t)$. Now $\partial\psi(t) = \partial\phi(t) - (\phi(1) - \phi(0))$ for $t \in \partial\psi(t)$ $\Rightarrow \partial\phi(t) = \phi(1) - \phi(0)$.

□

The Mean-Value Theorem can also be stated in integral form.

Theorem 13.18 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. For $x, y \in \mathbb{R}^n$ we have*

$$f(y) - f(x) = \int_0^1 \langle \partial f(x_t), y - x \rangle dt.$$

Remark The meaning of the integral is: if $\{s_t : t \in [0,1]\}$ with $s_t \in \partial f(x_t)$ for all $t \in [0,1]$, then $\int_0^1 \langle s_t, y - x \rangle dt$ is independent of the selection, and its value is $f(y) - f(x)$.

Example 13.2

- **Support Functions:** *Let C be compact, convex, $C \neq \emptyset$, with support function σ_C. From definition SII we ge*

$$\begin{aligned}
\partial \sigma_C(0) &= \{s \in \mathbb{R}^n : \sigma_C(y) \geq \overset{=0}{\sigma_C(0)} + \langle s, y - 0 \rangle \ \forall\, y \in \mathbb{R}^n\} \\
&= C \quad \textit{and} \\
(\sigma_C)'(0, d) &= \inf_{t>0} \frac{\sigma_C(0 + td) + \sigma_C(0)}{t} \\
&= \frac{t\sigma_C(d)}{t} = \sigma_C(d) \\
\Rightarrow (\sigma_C)'(0, \cdot) &= \sigma_C.
\end{aligned}$$

  Note the relation to Proposition 13.2 with $f = \sigma_C, x = 0$. Now use Proposition 13.12 in the sense that $f'(x, \cdot)$ is the support function of $\partial f(x)$. Therefore Proposition 13.12 can also be used for σ_C and C. It follows that $\partial \sigma_C(x) = F_C(x)$ and $(\sigma_C)'(x, \cdot) = \sigma_{F_C}(x)$. This can be seen as the optimal value of the problem (for variable $s \in C$, and fixed d, x) $\max\limits_{\langle s,x \rangle = \sigma_C(x)} \langle d, s \rangle$.

- **Norms and Gauges:** *Let $||| \cdot |||$ be a norm with unit ball $B = \{x :||| x |||\leq 1\}$. $||| \cdot |||$ is (see Chapter 12) the support function of*

$$B^* = \{s \in \mathbb{R}^n : \langle s, x \rangle \leq ||| x ||| \ \ \forall\, x \in \mathbb{R}^n\}$$

with associated dual norm $||| \cdot |||^*$.

$$\begin{aligned}\Rightarrow \partial||| \cdot |||(0) &= \partial\sigma_{B^*}(0)\\ &= \{s \in \mathbb{R}^n : \max_{|||d||| \le 1, d \in B} \langle s, d\rangle \le 1\}\\ &= B^* \text{ and}\\ \partial||| \cdot |||(x) &= \partial\sigma_{B^*}(x) = F_{B^*}(x)\\ &= \{s \in B^* : \langle s, x\rangle = \max_{u \in B^*}\langle u, x\rangle = ||| x |||\}.\end{aligned}$$

For a general closed convex set $C, 0 \in C$, *we use the polar set* $C^\circ := \{x : \langle s, x\rangle \le 1$ *for all* $s \in C\} \Rightarrow \partial\gamma_C(0) = \partial_{\gamma_{C^\circ}}(0) = C^\circ$ *and* $(\gamma_C)'(0, \cdot) = \gamma_C$. *Moreover,* $\partial\gamma_C(x) = F_{C^\circ}(x)$ *and* $(\gamma_C)'(x, \cdot) = \sigma_{F_{C^\circ}}(x)$.

- **Distance Functions:** $d_C(x) := \min\{\|y - x\| : y \in C\}$, C *closed and convex.*

$$\partial_{d_C}(x) = \begin{cases} N_C(x) \cap B(0,1) & \text{if } x \in C\\ \frac{x - p_C(x)}{\|x - p_C(x)\|} & \text{if } x \notin C.\end{cases}$$

The case $x \notin C$ *is shown in Example 10.1. Here even* ∇d_C *exists. For* $x \in C$, *let* $s \in \partial_{d_C}(x)$

$$\begin{aligned}&\Rightarrow\ d_C(x') \ge \langle s, x' - x\rangle\ \forall\ x' \in \mathbb{R}^n\\ &\Rightarrow\ \langle s, x' - x\rangle \le 0\ \forall\ x' \in C\\ &\Rightarrow\ s \in N_C(x).\end{aligned}$$

With $x' = x + s$ *we get*

$$\|s\|^2 \le d_C(x + s) \le \|x + s - x\| = \|s\|$$

$\Rightarrow \|s\| \le 1 \Rightarrow s \in N_C(x) \cap B(0,1)$. *Let* $s \in N_C(x) \cap B(0,1)$ *and, for all* $x' \in \mathbb{R}^n$, *write*

$$\langle s, x' - x\rangle = \langle s, x' - p_C(x')\rangle + \langle s, p_C(x') - x\rangle.$$

Now use $\|s\| \le 1$ *and C.S.I.*

$$\begin{aligned}&\langle s, x' - p_C(x')\rangle \le \|x' - p_C(x')\|\|s\| = d_C(x')\\ \Rightarrow\ & d_C(x') \ge \langle s, x' - x\rangle\ \forall\ x' \in \mathbb{R}^n\\ \Rightarrow\ & s \in \partial d_C.\end{aligned}$$

13.6 Computational Rules for Subdifferentials

Theorem 13.19 *Let f_1, f_2 be two convex functions from $\mathbb{R}^n$ to $\mathbb{R}$ and t_1, t_2 be positive. Then*

$$\partial(t_1 f_1 + t_2 f_2)(x) = t_1 \partial f_1(x) + t_2 \partial f_2(x) \quad \textit{for all} \quad x \in \mathbb{R}^n.$$

Theorem 13.20 *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be an affine mapping ($Ax = A_0 x + b$, with A_0 linear and $b \in \mathbb{R}^m$) and let g be a finite convex function on $\mathbb{R}^m$. Then*

$$\partial(g \circ A)(x) = A_0^* \partial g(Ax) \quad \textit{for all} \quad x \in \mathbb{R}^n.$$

Let $f_1, \dots, f_m$ be m convex functions from $\mathbb{R}^n$ to $\mathbb{R}$. Let F be defined by

$$\mathbb{R}^n \ni x \mapsto F(x) := (f_1(x), \dots, f_m(x)) \in \mathbb{R}^m.$$

Equip $\mathbb{R}^m$ with the dot-product and let $g : \mathbb{R}^m \to \mathbb{R}$ be convex and increasing componentwise, i.e.

$$y^i \geq z^i \quad \text{for} \quad i = 1, \dots, m \Longrightarrow g(y) \geq g(z).$$

Theorem 13.21 *Let f, F and g be as defined above. For all $x \in \mathbb{R}^n$,*

$$\partial(g \circ F)(x) = \left\{ \sum_{i=1}^{m} \rho^i s_i : (\rho^1, \dots, \rho^m) \in \partial g(F(x)), s_i \in \partial f_i(x) \quad \textit{for} \quad i = 1, \dots, m \right\}.$$

REFERENCES

[1] T. Bonnesen and W. Fenchel. *Theorie der Konvexen Körper.* Berlin, 1934.

[2] H. G. Eggleston. *Convexity.* The Cyndics of the Cambridge University Press, 1969.

[3] J-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms I.* Springer-Verlag, Berlin Heidelberg, 1993.

[4] J-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms II.* Springer-Verlag, Berlin Heidelberg, 1993.

[5] R. R. Phelps. *Convex Functions, Monotone Operators and Differentiability.* Springer-Verlag, Berlin Heidelberg, 1989. Lecture Notes in Mathematics.

[6] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.

[7] R. T. Rockafellar. Conjugate duality and optimization. In *Regional Conference Series in Applied Mathematics*, volume 16, Philadelphia, Pennsylvania 19103, 1974. Conference Board of the Mathematical Sciences.

[8] R. T. Rockafellar. *The Theory of Subgradients and its Applications to Problems of Optimization. Convex and Nonconvex Functions.* Heldermann Verlag, Berlin, 1981.

[9] R. T. Rockafellar and R. J-B. Wets. *Variational Analysis.* Springer-Verlag, 1998.

[10] F. A. Valentine. *Convex Sets.* McGraw-Hill Inc., New York, 1964.

[11] F. A. Valentine. *Konvexe Mengen.* Hochschultaschenbücher-Verlag, Bibliographisches Institut, Mannheim, 1968.

Index