

Convex Analysis

Daniel Mao

Contents

1	Affine Sets	1
1.1	Definitions	1
2	Relative Topology	3
2.1	Definitions	3
2.2	Basic Properties	3
2.3	Arithmetic Properties	4
3	Convex Sets	5
3.1	Definitions	5
3.2	Basic Properties	6
3.3	Arithmetic Properties	6
3.4	Stability of Convexity	7
3.5	Topological Properties	11
3.6	Examples of Convex Sets	14
4	Cones (in Analysis)	15
4.1	Definitions	15
4.2	The cone Operator	16
4.3	Other Properties	17
5	Tangent Cones and Normal Cones	19
5.1	Definitions	19
5.2	Basic Properties	19
5.3	Arithmetic Properties	24
5.4	Other Properties	24
6	Dual Cones and Polar Cones	27
6.1	Definitions	27
6.2	Properties	27

7	Extreme Points	29
7.1	Definitions	29
7.2	Properties	30
8	Projection	31
8.1	Definitions	31
8.2	Properties of the Projection Operator	32
8.3	Characterization	34
9	Separation	35
9.1	Definitions	35
9.2	Main Results	35
10	Convex Functions	39
10.1	Preliminaries	39
10.2	The Indicator Function	39
10.3	Definitions	42
10.4	Basic Properties	43
10.5	Differentiable Convex Functions	43
10.6	Convexity and Lipschitz-ness	46
10.7	Stability of Convexity	47
10.8	Examples	50
11	More Convex Functions	53
11.1	Strictly Convex	53
11.2	Strongly Convex	53
11.3	Quasiconvex	54
12	Support	57
12.1	Definitions	57
12.2	Properties	57
12.3	Supporting Hyperplane	59
13	Conjugacy	61
13.1	Definition and Examples	61
13.2	Basic Properties	62
13.3	Double Conjugate	62
13.4	Conjugates and Sub-Differentials	63

14 Proximal	65
14.1 Definition	65
14.2 Basic Properties	65
14.3 Prox Calculus Rules	65
14.4 The Second Prox Theorem	66
14.5 Moreau Decomposition	66
15 Ellipsoids	67
15.1 Properties	67

Chapter 1

Affine Sets

1.1 Definitions

Definition (Affine Combination). *Let S be a set in \mathbb{E} . We define an **affine combination** of S to be a point x in the space of the form*

$$x = \sum_{i=1}^n \lambda_i v_i$$

where (1) $n \in \mathbb{N}$, (2) $v_i \in S$ for all i , (3) $\lambda_i \in \mathbb{R}$ for all i , and (4) $\sum_{i=1}^n \lambda_i = 1$.

Definition (Affine Span). *Let S be a set in \mathbb{E} . We define the **affine span** of S , denoted by $\text{affspan}(S)$, to be the set of all affine combinations of S .*

Definition (Affine Set). *Let S be a set in \mathbb{E} . We say that S is an **affine set** if $S = \text{aff}(S)$.*

Definition (Affine Hull). *Let S be a set in \mathbb{E} . We define the **affine hull** of S , denoted by $\text{affhull}(S)$, to be the smallest affine set containing S .*

Chapter 2

Relative Topology

2.1 Definitions

Definition (Relative Interior). *Let \mathbb{E} be some Euclidean space. Let S be a set in the space. We define the **relative interior** of S , denoted by $\text{ri}(S)$, to be the interior of S for the topology relative to the affine hull $\text{aff}(S)$. i.e., the set given by*

$$\text{ri}(S) := \{x \in \text{aff}(S) : \exists r > 0, \text{ball}(x, r) \cap \text{aff}(S) \subseteq S\}.$$

A quick result. For a singleton set S , $\text{ri}(S) = S = \text{cl}(S)$.

2.2 Basic Properties

Proposition 2.2.1. *For any set S , we have $\text{ri}(S) \subseteq S$.*

Remark. *The relative interior operator is not monotonic.*

Example 2.2.1. *Consider \mathbb{R} with the usual topology and sets $\{0\}$ and $[0, 1]$. Then $\text{ri}(\{0\}) = \{0\}$ and $\text{ri}([0, 1]) = (0, 1)$.*

Proposition 2.2.2. *Let S be a set in some Euclidean space \mathbb{E} . Then if $\text{int}(S) \neq \emptyset$, $\text{ri}(S) = \text{int}(S)$.*

Proof.

It suffices to show that $\text{aff}(S) = \mathbb{R}^n$.

Since $\text{int}(S) \neq \emptyset$, $\exists x \in \text{int}(S)$.

Since $x \in \text{int}(S)$, $\exists r > 0$, $\text{ball}(x, r) \subseteq S$.

$\mathbb{E} = \text{aff}(\text{ball}(x, r)) \subseteq \text{aff}(S) \subseteq \mathbb{E}$.

This shows $\text{aff}(S) = \mathbb{E}$.

■

2.3 Arithmetic Properties

Proposition 2.3.1. *Let C_1 and C_2 be convex subsets of \mathbb{E} . Let λ_1 and λ_2 be scalars in \mathbb{R} . Then*

$$\text{ri}(\lambda_1 C_1 + \lambda_2 C_2) = \lambda_1 \text{ri}(C_1) + \lambda_2 \text{ri}(C_2).$$

Proposition 2.3.2. *Let C_1 be a convex set in \mathbb{E}_1 . Let C_2 be a convex set in \mathbb{E}_2 . Then*

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

Chapter 3

Convex Sets

3.1 Definitions

Definition (Convex Combination). *Let S be a set in \mathbb{E} . We define a **convex combination** of S to be a point x in the space of the form*

$$x = \sum_{i=1}^n \lambda_i v_i$$

where (1) $n \in \mathbb{N}$, (2) $v_i \in S$ for all i , (3) $\lambda_i \in \mathbb{R}_+$ for all i , and (4) $\sum_{i=1}^n \lambda_i = 1$.

Definition (Convex Span). *Let S be a set in \mathbb{E} . We define a **convex span** of S , denoted by $\text{convspan}(S)$, to be the set of all convex combinations of S .*

Definition (Convex Sets). *Let S be a set in \mathbb{E} . We say that S is **convex** if $S = \text{convspan}(S)$, or equivalently, $\alpha x + \beta y \in S$ for any $x, y \in S$ and any $\alpha, \beta \in [0, 1]$ such that $\alpha + \beta = 1$.*

Definition (Convex Hull). *Let S be a set in \mathbb{E} . We define the **convex hull** of S , denoted by $\text{convhull}(S)$, to be the smallest convex set containing S .*

Definition (Closed Convex Hull). *Let S be a set in some Euclidean space. We define the **closed convex hull** of S to be the intersection of all closed convex supersets of S .*

Proposition 3.1.1. *For any set S , $\text{convspan}(S) = \text{convhull}(S)$. They will both be denoted by $\text{conv}(S)$ from now on.*

Proof.

For one direction, let x be an arbitrary point in $\text{convspan}(S)$.

We are to prove that $x \in \text{convhull}(S)$.

Let C be an arbitrary convex set containing S .

Since x is a convex combination of S , x is also a convex combination of C .

Since x is a convex combination of C and C is convex, $x \in C$.

Since x is in any convex set containing S , $x \in \text{convhull}(S)$.

Since $x \in \text{convhull}(S)$ for any $x \in \text{convspan}(S)$, $\text{convspan}(S) \subseteq \text{convhull}(S)$.

For the reverse direction,

proof incomplete.

■

Proposition 3.1.2. *The closed convex hull is the closure of the convex hull.*

3.2 Basic Properties

Proposition 3.2.1 (The conv Operator).

(1) *For any set S in a Euclidean space, we have*

$$S \subseteq \text{conv}(S).$$

(2) *(Monotonic) For any sets S_1 and S_2 in \mathbb{E} , if $S_1 \subseteq S_2$, then*

$$\text{conv}(S_1) \subseteq \text{conv}(S_2).$$

(3) *(Idempotent) For any set S in \mathbb{E} , we have*

$$\text{conv}(\text{conv}(S)) = \text{conv}(S).$$

Theorem 1 (Carathéodory). *Let \mathbb{E} be some Euclidean space. Let S be some set in the space. Let x be some point in $\text{conv}(S)$. Then x can be represented as a convex combination of at most $d + 1$ points in S . i.e., x lies in some r -simplex with vertices in S , where $r \leq d$.*

3.3 Arithmetic Properties

Proposition 3.3.1. *Let C be a convex set. Let λ_1 and λ_2 be in \mathbb{R}_+ . Then*

$$(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C.$$

Proof.

The case where any of λ_1 and λ_2 is 0 is trivial. I will assume that $\lambda_1, \lambda_2 > 0$.

For one direction, let x be an arbitrary point in $(\lambda_1 + \lambda_2)C$.

Since $x \in (\lambda_1 + \lambda_2)C$, $\exists c \in C, x = (\lambda_1 + \lambda_2)c$.

Since $\begin{cases} (\lambda_1 + \lambda_2)c = \lambda_1 c + \lambda_2 c \\ x = (\lambda_1 + \lambda_2)c \end{cases}$, we get $x = \lambda_1 c + \lambda_2 c$.

Since $\begin{cases} x = \lambda_1 c + \lambda_2 c \\ \lambda_1 c \in \lambda_1 C \\ \lambda_2 c \in \lambda_2 C \end{cases}$, we get $x \in \lambda_1 C + \lambda_2 C$.

Since $x \in \lambda_1 C + \lambda_2 C$ for any $x \in (\lambda_1 + \lambda_2)C$, $\lambda_1 C + \lambda_2 C \subseteq (\lambda_1 + \lambda_2)C$.

For the reverse direction, let x be an arbitrary point in $\lambda_1 C + \lambda_2 C$.

Since $x \in \lambda_1 C + \lambda_2 C$, $\exists c_1, c_2 \in C, x = \lambda_1 c_1 + \lambda_2 c_2$.

Define scalars $\mu_1 := \frac{\lambda_1}{\lambda_1 + \lambda_2}$ and $\mu_2 := \frac{\lambda_2}{\lambda_1 + \lambda_2}$.

Then $x = (\lambda_1 + \lambda_2)c$.

Since $\lambda_1, \lambda_2 > 0$, $\mu_1, \mu_2 \in [0, 1]$.

Define point $c := \mu_1 c_1 + \mu_2 c_2$.

Since $\begin{cases} c = \mu_1 c_1 + \mu_2 c_2 \\ c_1, c_2 \in C \\ \mu_1, \mu_2 \in [0, 1] \\ \mu_1 + \mu_2 = 1 \\ C \text{ is convex} \end{cases}$, we get $c \in C$.

Since $x = (\lambda_1 + \lambda_2)c$ and $c \in C$, $x \in (\lambda_1 + \lambda_2)C$.

Since $x \in (\lambda_1 + \lambda_2)C$ for any $x \in \lambda_1 C + \lambda_2 C$, $(\lambda_1 + \lambda_2)C \subseteq \lambda_1 C + \lambda_2 C$.

Since $\lambda_1 C + \lambda_2 C \subseteq (\lambda_1 + \lambda_2)C$ and $(\lambda_1 + \lambda_2)C \subseteq \lambda_1 C + \lambda_2 C$, $(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C$ ■

3.4 Stability of Convexity

Proposition 3.4.1 (Intersection). *Convexity is stable under intersection. i.e., the intersection of any collection of convex sets is convex.*

Proof. Let $\{C_i\}_{i \in I}$ be an arbitrary collection of convex sets where I is an index set and C_i is convex for any $i \in I$. Let C denote their intersection. If $C = \emptyset$, then we are done. Else, let x and y be two arbitrary points in C . Let λ be an arbitrary number in $(0, 1)$. Define a point $z := \lambda x + (1 - \lambda)y$. Since $x \in C$ and $C = \bigcap_{i \in I} C_i$, we get $x \in C_i$ for any $i \in I$. Since $y \in C$ and $C = \bigcap_{i \in I} C_i$, we get $y \in C_i$ for any $i \in I$. Let i be an arbitrary index in I . Since $x \in C_i$ and $y \in C_i$ and $\lambda \in (0, 1)$ and C_i is convex and $z = \lambda x + (1 - \lambda)y$, we get $z \in C_i$. Since $z \in C_i$ for any $i \in I$ and $C = \bigcap_{i \in I} C_i$, we get $z \in C$. Since

$$\forall x, y \in C, \forall \lambda \in (0, 1), \quad \lambda x + (1 - \lambda)y \in C,$$

by definition of convex sets, we get C is convex. ■

Proposition 3.4.2 (Affine Map). *Convexity is stable under affine mapping. i.e., the affine image of a convex set is convex.*

Proposition 3.4.3 (Linear Combinations). *Convexity is stable under linear combinations. i.e., if C_1 and C_2 are convex sets and λ_1 and λ_2 are real numbers, then the set C defined as*

$$C := \lambda_1 C_1 + \lambda_2 C_2$$

is convex.

Proof. If $C_1 = \emptyset$ or $C_2 = \emptyset$, then $\lambda_1 C_1 + \lambda_2 C_2 = \emptyset$ and we are done. Now assume that $C_1, C_2 \neq \emptyset$. Then $C = \lambda_1 C_1 + \lambda_2 C_2 \neq \emptyset$. Let x and y be arbitrary points in C .

Since $x \in C$, $\exists x_1 \in C_1, x_2 \in C_2$ such that $x = \lambda_1 x_1 + \lambda_2 x_2$.

Since $y \in C$, $\exists y_1 \in C_1, y_2 \in C_2$ such that $y = \lambda_1 y_1 + \lambda_2 y_2$.

Let $\lambda \in [0, 1]$ be arbitrary. Define a point z as $z := \lambda x + (1 - \lambda)y$. Then

$$\begin{aligned} z &= \lambda x + (1 - \lambda)y \\ &= \lambda(\lambda_1 x_1 + \lambda_2 x_2) + (1 - \lambda)(\lambda_1 y_1 + \lambda_2 y_2) \\ &= \lambda_1(\lambda x_1 + (1 - \lambda)y_1) + \lambda_2(\lambda x_2 + (1 - \lambda)y_2). \end{aligned}$$

Since $x_1, y_1 \in C_1$, $\lambda \in [0, 1]$ and C_1 is convex, we get $\lambda x_1 + (1 - \lambda)y_1 \in C_1$.

Since $x_2, y_2 \in C_2$, $\lambda \in [0, 1]$ and C_2 is convex, we get $\lambda x_2 + (1 - \lambda)y_2 \in C_2$.

So $z = \lambda_1(\lambda x_1 + (1 - \lambda)y_1) + \lambda_2(\lambda x_2 + (1 - \lambda)y_2) \in \lambda_1 C_1 + \lambda_2 C_2$.

That is, $\forall x \in C, \forall y \in C, \forall \lambda \in [0, 1]$, we have $\lambda x + (1 - \lambda)y \in C$.

So by definition, C is convex. ■

Corollary. *The Minkowski sum of two convex sets is convex.*

Lemma 1. *Let C be a convex set in \mathbb{E} . Let $x \in \text{int}(C)$. Let $y \in \text{cl}(C)$. Then*

$$\forall \lambda \in (0, 1], \quad \lambda x + (1 - \lambda)y \in C.$$

Proof.

Since $x \in \text{int}(S)$, there exists some radius r_x such that $\text{ball}(x, r_x) \subseteq S$.

Define $r_z := \lambda r_x$.

Let z' be an arbitrary point in $\text{ball}(z, r_z)$.

Define $x' := \frac{1}{\lambda}(z' - (1 - \lambda)y)$.

Notice

$$\begin{aligned} &\|x - x'\| \\ &= \frac{1}{|\lambda|} \|\lambda x - \lambda x'\| \\ &= \frac{1}{|\lambda|} \|(z - (1 - \lambda)y) - (z' - (1 - \lambda)y)\| \\ &= \frac{1}{|\lambda|} \|z - z'\| \end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{|\lambda|} r_z, \text{ since } z' \in \text{ball}(z, r_z) \\
&= \frac{1}{|\lambda|} \lambda r_x \\
&= r_x.
\end{aligned}$$

That is,

$$\|x - x'\| \leq r_x.$$

So $x' \in \text{ball}(x, r_x)$.

Since $x' \in \text{ball}(x, r_x)$ and $\text{ball}(x, r_x) \subseteq S$, we get $x' \in S$.

$$\text{Since } \begin{cases} z' = \lambda x' + (1 - \lambda)y \\ x', y \in S \\ \lambda \in (0, 1] \\ S \text{ is convex} \end{cases}, \text{ we get } z' \in S.$$

Since $z' \in S$ for any $z' \in \text{ball}(z, r_z)$, $\text{ball}(z, r_z) \subseteq S$.

Since there exists some radius r_z such that $\text{ball}(z, r_z) \subseteq S$, $z \in \text{int}(S)$.

Alternative Expressing:

Define $B := \text{ball}(0, 1)$.

$$\begin{aligned}
&(1 - \lambda)x + \lambda y + \varepsilon B \\
&\subseteq (1 - \lambda)x + \lambda(C + \varepsilon B) + \varepsilon B \\
&= (1 - \lambda)x + \lambda C + \lambda \varepsilon B + \varepsilon B \\
&= (1 - \lambda)x + (1 + \lambda)\varepsilon B + \lambda C \\
&= (1 - \lambda)\left(x + \frac{1 + \lambda}{1 - \lambda}\varepsilon B\right) + \lambda C \\
&\subseteq (1 - \lambda)C + \lambda C \\
&= C.
\end{aligned}$$

■

Lemma 2. Let C be a convex set in \mathbb{E} . Let $x \in \text{ri}(C)$. Let $y \in \text{cl}(C)$. Then

$$\forall \lambda \in (0, 1], \quad \lambda x + (1 - \lambda)y \in C.$$

Proof.

Case 1. $\text{int}(C) \neq \emptyset$.

Then $\text{int}(C) = \text{ri}(C)$.

Since $x \in \text{int}(C)$ and $y \in \text{cl}(C)$, $\forall t \in (0, 1]$, $z := tx + (1 - t)y \in C$.

Case 2. $\text{int}(C) = \emptyset$.

Now $\dim(C) < d$.

Say $\dim(C) = l$.

Apply case 1 in \mathbb{R}^l .

■

Proposition 3.4.4 (Interior). *Convexity is stable under interior. i.e., the interior of a convex set is convex.*

Proof. Let S be a convex set in \mathbb{E} . If $\text{int}(S) = \emptyset$, then we are done. Else: let x and y be two arbitrary points in $\text{int}(S)$. Let λ be an arbitrary number in $(0, 1)$. Define a point z by $z := \lambda x + (1 - \lambda)y$. Since $x, y \in \text{int}(S)$ and $\lambda \in (0, 1)$, by the lemma, we get $z \in \text{int}(S)$. Since

$$\forall x, y \in \text{int}(S), \forall \lambda \in (0, 1), \quad \lambda x + (1 - \lambda)y \in \text{int}(S),$$

we get $\text{int}(S)$ is convex.

■

Proposition 3.4.5 (Relative Interior). *Convexity is stable under relative interior. i.e., the relative interior of a convex set is convex.*

Proof. Let S be a convex set in \mathbb{E} . If $\text{ri}(S) = \emptyset$, then we are done. Else: let x and y be two arbitrary points in $\text{ri}(S)$. Let λ be an arbitrary number in $(0, 1)$. Define a point z by $z := \lambda x + (1 - \lambda)y$. Since $x, y \in \text{ri}(S)$ and $\lambda \in (0, 1)$, by the lemma, we get $z \in \text{ri}(S)$. Since

$$\forall x, y \in \text{ri}(S), \forall \lambda \in (0, 1), \quad \lambda x + (1 - \lambda)y \in \text{ri}(S),$$

we get $\text{ri}(S)$ is convex.

■

Proposition 3.4.6 (Closure). *Convexity is stable under closure. i.e., the closure of a convex set is convex.*

Proof Approach 1.

Let $x, y \in \text{cl}(C)$.

Let $t \in [0, 1]$.

Since $x \in \text{cl}(C)$, $\exists \{x_i\}_{i \in \mathbb{N}} \subseteq C$, $\lim_{i \rightarrow \infty} x_i = x$.

Since $y \in \text{cl}(C)$, $\exists \{y_i\}_{i \in \mathbb{N}} \subseteq C$, $\lim_{i \rightarrow \infty} y_i = y$.

Since $\lim_{i \rightarrow \infty} x_i = x$ and $\lim_{i \rightarrow \infty} y_i = y$, $\lim_{i \rightarrow \infty} (tx_i + (1 - t)y_i) = tx + (1 - t)y$.

Since $x_i, y_i \in C$ and C is convex, $tx_i + (1 - t)y_i \in C$.

Since $tx_i + (1 - t)y_i \in C$, $\lim_{i \rightarrow \infty} (tx_i + (1 - t)y_i) = tx + (1 - t)y$, $tx + (1 - t)y \in \text{cl}(C)$.

Since $\forall x, y \in \text{cl}(C)$, $\forall t \in [0, 1]$, $tx + (1 - t)y \in \text{cl}(C)$, we get $\text{cl}(C)$ is convex.

■

Proof Approach 2.

$\text{cl}(C) = \bigcap_{\varepsilon > 0} [C + \varepsilon \text{ball}(0, 1)]$. This is an intersection of linear combinations of convex sets and hence convex. ■

Proposition 3.4.7 (Conical Hull). *Convexity is stable under conical hull. i.e., if C is convex, then $\text{cone}(C)$ is convex.*

Proof.

Let x and y be arbitrary points in $\text{cone}(C)$.

Let λ be an arbitrary number in $(0, 1)$.

Define point z as $z := \lambda x + (1 - \lambda)y$.

Since $x \in \text{cone}(C)$, $\exists x' \in C$ and $\exists \alpha > 0$ such that $x = \alpha x'$.

Since $y \in \text{cone}(C)$, $\exists y' \in C$ and $\exists \beta > 0$ such that $y = \beta y'$.

Define point z' as $z' := \frac{\lambda\alpha}{\lambda\alpha + (1-\lambda)\beta}x' + \frac{(1-\lambda)\beta}{\lambda\alpha + (1-\lambda)\beta}y'$.

Since $x', y' \in C$ and $\frac{\lambda\alpha}{\lambda\alpha + (1-\lambda)\beta} \in (0, 1)$ and $\frac{\lambda\alpha}{\lambda\alpha + (1-\lambda)\beta} + \frac{(1-\lambda)\beta}{\lambda\alpha + (1-\lambda)\beta} = 1$ and C is convex and $z' := \frac{\lambda\alpha}{\lambda\alpha + (1-\lambda)\beta}x' + \frac{(1-\lambda)\beta}{\lambda\alpha + (1-\lambda)\beta}y'$, we get $z' \in C$.

Since $z' \in C$ and $z = (\lambda\alpha + (1 - \lambda)\beta)z'$, $z \in \text{cone}(C)$.

That is, $\lambda x + (1 - \lambda)y \in \text{cone}(C)$.

Since $\forall x, y \in \text{cone}(C)$, $\forall \lambda \in (0, 1)$, $\lambda x + (1 - \lambda)y \in \text{cone}(C)$, we get $\text{cone}(C)$ is convex. ■

3.5 Topological Properties

Proposition 3.5.1. *A closed convex hull does not distinguish a set from its closure. i.e., for any set S , we have $\overline{\text{conv}}(S) = \overline{\text{conv}}(\text{cl}(S))$.*

Proposition 3.5.2. *The convex hull of a bounded set is bounded.*

Proposition 3.5.3. *The convex hull of a compact set is compact.*

Proposition 3.5.4. *If S is bounded, then the closure operation and the convex hull operation commute. i.e., $\text{conv}(\text{cl}(S)) = \text{cl}(\text{conv}(S))$.*

Remark. *The closure operation and the convex hull operation do not commute in general.*

Theorem 2. *Let C be a convex set such that $\text{int}(C) \neq \emptyset$. Then*

(1) $\text{int}(C) = \text{int}(\text{cl}(C))$, and

(2) $\text{cl}(C) = \text{cl}(\text{int}(C))$.

Proof of (1). $\text{int}(C) \subseteq \text{int}(\text{cl}(C))$ is clear. For $\text{int}(\text{cl}(C)) \subseteq \text{int}(C)$, let x be an arbitrary point in $\text{int}(\text{cl}(C))$.

Since $x \in \text{int}(\text{cl}(C))$,

$$\exists r > 0 \text{ such that } \text{ball}(x, r) \subseteq \text{cl}(C).$$

Since $\text{int}(C) \neq \emptyset$, pick $y \in \text{int}(C)$.

Define a scalar λ by

$$\lambda := \frac{r}{2\|x - y\|}.$$

Define a point z by

$$z := x + \lambda(x - y).$$

Since $\lambda = \frac{r}{2\|x - y\|}$ and $z = x + \lambda(x - y)$,

$$\begin{aligned} & \|z - x\| \\ &= \|x + \lambda(x - y) - x\| \\ &= \|\lambda(x - y)\| \\ &= \lambda\|x - y\| \\ &= \frac{r}{2\|x - y\|}\|x - y\| \\ &= \frac{r}{2} \\ &< r. \end{aligned}$$

That is,

$$\|z - x\| < r.$$

So $z \in \text{ball}(x, r)$. It follows that $z \in \text{cl}(C)$.

Since $z = x + \lambda(x - y)$, rearranging this yields

$$x = \frac{1}{1 + \lambda}z + \frac{\lambda}{1 + \lambda}y.$$

$$\text{Since } \begin{cases} x = \frac{1}{1 + \lambda}z + \frac{\lambda}{1 + \lambda}y \\ z \in \text{cl}(C) \\ y \in \text{int}(C) \\ \frac{1}{1 + \lambda}, \frac{\lambda}{1 + \lambda} \in (0, 1) \\ \frac{1}{1 + \lambda} + \frac{\lambda}{1 + \lambda} = 1 \end{cases}, \text{ by the lemma, we get}$$

$$x \in \text{int}(C).$$

Since $\forall x \in \text{int}(\text{cl}(C)), x \in \text{int}(C)$, we get $\text{int}(\text{cl}(C)) \subseteq \text{int}(C)$. ■

Proof of (2). $\text{cl}(\text{int}(C)) \subseteq \text{cl}(C)$ is clear. For $\text{cl}(C) \subseteq \text{cl}(\text{int}(C))$, let x be an arbitrary point in $\text{cl}(C)$.

Since $\text{int}(C) \neq \emptyset$, pick $y \in \text{int}(C)$.

Let $\lambda \in [0, 1)$.

Define a point z by

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

$$\text{Since } \begin{cases} z(\lambda) := \lambda x + (1 - \lambda)y \\ x \in \text{cl}(C) \\ y \in \text{int}(C) \\ \lambda \in [0, 1) \end{cases}, \text{ by the lemma, we get}$$

$$z(\lambda) \in \text{int}(C).$$

$$\text{Since } \begin{cases} z(\lambda) \in \text{int}(C) \\ \lim_{\lambda \rightarrow 1} z(\lambda) = x \end{cases}, \text{ we get}$$

$$x \in \text{cl}(\text{int}(C)).$$

Since $\forall x \in \text{cl}(C), x \in \text{cl}(\text{int}(C))$, we get $\text{cl}(C) \subseteq \text{cl}(\text{int}(C))$. ■

Proposition 3.5.5. *If C is convex, then*

$$(1) \text{ aff}(\text{ri}(C)) = \text{aff}(C) = \text{aff}(\text{cl}(C)),$$

$$(2) \text{ ri}(\text{ri}(C)) = \text{ri}(C) = \text{ri}(\text{cl}(C)), \text{ and}$$

$$(3) \text{ cl}(\text{ri}(C)) = \text{cl}(C) = \text{cl}(\text{cl}(C)).$$

Proposition 3.5.6. *Let C be a convex set. Then*

$$C \neq \emptyset \iff \text{ri}(C) \neq \emptyset.$$

Proof.

For one direction, assume that $C \neq \emptyset$. We are to prove that $\text{ri}(C) \neq \emptyset$. Since $C \neq \emptyset$, $\text{aff}(C) \neq \emptyset$. Since C is convex, $\text{aff}(C) = \text{aff}(\text{ri}(C))$. Since $\begin{cases} \text{aff}(C) \neq \emptyset \\ \text{aff}(C) = \text{aff}(\text{ri}(C)) \end{cases}$, we get

$$\text{aff}(\text{ri}(C)) \neq \emptyset.$$

Since $\text{aff}(\text{ri}(C)) \neq \emptyset$, we get $\text{ri}(C) \neq \emptyset$.

For the reverse direction, assume that $\text{ri}(C) \neq \emptyset$. We are to prove that $C \neq \emptyset$. Since $\text{ri}(C) \neq \emptyset$ and $\text{ri}(C) \subseteq C$, we get $C \neq \emptyset$. ■

3.6 Examples of Convex Sets

Example 3.6.1. Let I be an index set. Let b_i for $i \in I$ be vectors in \mathbb{E} . Let β_i for $i \in I$ be reals. Then the set C given by

$$C := \{x \in \mathbb{E} : \forall i \in I, \langle x, b_i \rangle \leq \beta_i\}$$

is convex.

Proof.

Each of $C_i := \{x \in \mathbb{E} : \langle x, b_i \rangle \leq \beta_i\}$ is convex and $C = \bigcap_{i \in I} C_i$.

$$\begin{aligned} \langle z, b_i \rangle &= \langle \lambda x + (1 - \lambda)y, b_i \rangle \\ &= \lambda \langle x, b_i \rangle + (1 - \lambda) \langle y, b_i \rangle \\ &\leq \lambda \beta_i + (1 - \lambda) \beta_i \\ &= \beta_i. \end{aligned}$$

■

Chapter 4

Cones (in Analysis)

4.1 Definitions

Definition (Conical Combination). *Let S be a set in \mathbb{E} . We define a **conical combination** of S to be a point x in the space of the form*

$$x = \sum_{i=1}^n \lambda_i v_i$$

where (1) $n \in \mathbb{N}$, (2) $v_i \in S$ for all i , and (3) $\lambda_i \in \mathbb{R}_{++}$ for all i .

Definition (Cone). *Let S be a set in \mathbb{E} . We say that S is a **cone** if $S = \mathbb{R}_{++}S$.*

Definition (Conical Hull). *Let S be a set in \mathbb{E} . We define the **conical hull** of S , denoted by $\text{cone}(S)$, to be the intersection of all cones containing S .*

Proposition 4.1.1. *Let S be a set in \mathbb{E} . Then $\text{cone}(S) = \mathbb{R}_{++}S$.*

Proof. For one direction, we are to prove $\text{cone}(S) \subseteq \mathbb{R}_{++}S$. Since $\mathbb{R}_{++}\mathbb{R}_{++}S = \mathbb{R}_{++}S$, $\mathbb{R}_{++}S$ is a cone. Since $1 \in \mathbb{R}_{++}$, $S \subseteq \mathbb{R}_{++}S$. Since $\mathbb{R}_{++}S$ is a cone containing S and $\text{cone}(S)$ is the smallest cone containing S , we get

$$\text{cone}(S) \subseteq \mathbb{R}_{++}S.$$

For the reverse direction, we are to prove $\mathbb{R}_{++}S \subseteq \text{cone}(S)$. Let C be an arbitrary cone containing S . Since $S \subseteq C$, $\mathbb{R}_{++}S \subseteq \mathbb{R}_{++}C$. Since C is a cone, $\mathbb{R}_{++}C = C$. So $\mathbb{R}_{++}S \subseteq C$. Since $\mathbb{R}_{++}S \subseteq C$ for any cone C containing S , we get

$$\mathbb{R}_{++}S \subseteq \text{cone}(S).$$

■

Definition (Closed Conical Hull). *Let S be a set in \mathbb{E} . We define the **closed conical hull** of S , denoted by $\text{clcone}(S)$, to be the intersection of all closed cones containing S .*

Proposition 4.1.2. *For any set S in \mathbb{E} , we have*

$$\text{clcone}(S) = \text{cl}(\text{cone}(S)).$$

Proof.

For $\text{clcone}(S) \subseteq \text{cl}(\text{cone}(S))$.

Since $\text{cl}(\text{cone}(S))$ is a closed cone containing S and $\text{clcone}(S)$ is the smallest closed cone containing S , $\text{clcone}(S) \subseteq \text{cl}(\text{cone}(S))$.

For $\text{cl}(\text{cone}(S)) \subseteq \text{clcone}(S)$.

Since $S \subseteq \text{clcone}(S)$, by the monotonicity of the cone operator, $\text{cone}(S) \subseteq \text{cone}(\text{clcone}(S))$.

Since $\text{cone}(S) \subseteq \text{cone}(\text{clcone}(S))$, by the monotonicity of the closure operator, $\text{cl}(\text{cone}(S)) \subseteq \text{cl}(\text{cone}(\text{clcone}(S)))$.

Since $\text{clcone}(S)$ is a cone, $\text{cone}(\text{clcone}(S)) = \text{clcone}(S)$.

Since $\text{clcone}(S)$ is closed, $\text{cl}(\text{clcone}(S)) = \text{clcone}(S)$.

Since $\text{cone}(\text{clcone}(S)) = \text{clcone}(S)$ and $\text{cl}(\text{clcone}(S)) = \text{clcone}(S)$, we get $\text{cl}(\text{cone}(\text{clcone}(S))) = \text{clcone}(S)$.

Since $\text{cl}(\text{cone}(S)) \subseteq \text{cl}(\text{cone}(\text{clcone}(S)))$ and $\text{cl}(\text{cone}(\text{clcone}(S))) = \text{clcone}(S)$, we get $\text{cl}(\text{cone}(S)) \subseteq \text{clcone}(S)$.

Since $\text{clcone}(S) \subseteq \text{cl}(\text{cone}(S))$ and $\text{cl}(\text{cone}(S)) \subseteq \text{clcone}(S)$, we get $\text{clcone}(S) = \text{cl}(\text{cone}(S))$. ■

4.2 The cone Operator

Proposition 4.2.1 (The cone Operator). *The cone operator has the following properties.*

$$(1) \quad \forall S \subseteq \mathbb{E},$$

$$S \subseteq \text{cone}(S).$$

$$(2) \quad \forall S_1, S_2 \subseteq \mathbb{E},$$

$$S_1 \subseteq S_2 \implies \text{cone}(S_1) \subseteq \text{cone}(S_2).$$

$$(3) \quad \forall S \subseteq \mathbb{E},$$

$$\text{cone}(\text{cone}(S)) = \text{cone}(S).$$

Proposition 4.2.2. *The conv operator and the cone operator commute. Let S be a set in \mathbb{E} . Then*

$$\text{conv}(\text{cone}(S)) = \text{cone}(\text{conv}(S)).$$

Proof.

For $\text{cone}(\text{conv}(S)) \subseteq \text{conv}(\text{cone}(S))$, let x be an arbitrary point in $\text{cone}(\text{conv}(S))$.

Since $x \in \text{cone}(\text{conv}(S))$, we get $\exists \lambda \in \mathbb{R}_+$, $\exists n \in \mathbb{N}$, $\exists v_1, \dots, v_n \in S$, $\exists \mu_1, \dots, \mu_n \in [0, 1]$, $\sum_{i=1}^n \mu_i = 1$ such that $x = \lambda \sum_{i=1}^n \mu_i v_i$.

Since $x = \lambda \sum_{i=1}^n \mu_i v_i$, $x = \sum_{i=1}^n \mu_i (\lambda v_i)$.

Since $\lambda \in \mathbb{R}_+$ and $v_i \in S$, $\lambda v_i \in \text{cone}(S)$.

Since $\lambda v_i \in \text{cone}(S)$ and $\mu_i \in [0, 1]$, $\sum_{i=1}^n \mu_i = 1$, $\sum_{i=1}^n \mu_i (\lambda v_i) \in \text{conv}(\text{cone}(S))$.

Since $\forall x \in \text{cone}(\text{conv}(S))$, $x \in \text{conv}(\text{cone}(S))$, $\text{cone}(\text{conv}(S)) \subseteq \text{conv}(\text{cone}(S))$.

For $\text{conv}(\text{cone}(S)) \subseteq \text{cone}(\text{conv}(S))$, let x be an arbitrary point in $\text{conv}(\text{cone}(S))$.

Since $x \in \text{conv}(\text{cone}(S))$, $\exists n \in \mathbb{N}$, $\exists \lambda_i \in [0, 1]$, $\sum_{i=1}^n \lambda_i = 1$, $\exists \mu_i \in \mathbb{R}_+$, $\exists v_i \in S$ such that $x = \sum_{i=1}^n \lambda_i \mu_i v_i$.

Define $\alpha := \sum_{i=1}^n \lambda_i \mu_i$.

Define $\beta_i := \lambda_i \mu_i / \alpha$.

Then $\alpha \in \mathbb{R}_+$ and $\beta_i \in [0, 1]$ and $\sum_{i=1}^n \beta_i = 1$ and $x = \alpha \sum_{i=1}^n \beta_i v_i$.

Since $\beta_i \in [0, 1]$ and $\sum_{i=1}^n \beta_i = 1$ and $v_i \in S$, we get $\sum_{i=1}^n \beta_i v_i \in \text{conv}(S)$.

Since $\alpha \in \mathbb{R}_+$ and $\sum_{i=1}^n \beta_i v_i \in \text{conv}(S)$ and $x = \alpha \sum_{i=1}^n \beta_i v_i$, we get $x \in \text{cone}(\text{conv}(S))$.

Since $\forall x \in \text{conv}(\text{cone}(S))$, $x \in \text{cone}(\text{conv}(S))$, we get $\text{conv}(\text{cone}(S)) \subseteq \text{cone}(\text{conv}(S))$.

Since $\text{cone}(\text{conv}(S)) \subseteq \text{conv}(\text{cone}(S))$ and $\text{conv}(\text{cone}(S)) \subseteq \text{cone}(\text{conv}(S))$, we get $\text{conv}(\text{cone}(S)) = \text{cone}(\text{conv}(S))$. ■

4.3 Other Properties

Proposition 4.3.1. *Let C be a convex set in \mathbb{E} . Assume $\text{int}(C) \neq \emptyset$ and $0 \in C$. Then $\text{int}(\text{cone}(C)) = \text{cone}(\text{int}(C))$.*

Proof.

For one direction, let x be an arbitrary point in $\text{int}(\text{cone}(C))$. We are to prove that $x \in \text{cone}(\text{int}(C))$.

Since $x \in \text{int}(\text{cone}(C))$, $\exists r$ such that $\text{ball}(x, r) \subseteq \text{cone}(C)$.

Since $x \in \text{int}(\text{cone}(C))$, $x \in \text{cone}(C)$.

Since $x \in \text{cone}(C)$, $\exists n \in \mathbb{N}$, $\exists \lambda_1, \dots, \lambda_n > 0$, $\exists v_1, \dots, v_n \in C$ such that $x = \sum_{i=1}^n \lambda_i v_i$.

Assume for the sake of contradiction that $\exists k \in \{1, \dots, n\}$ such that $\forall r_k > 0$, $\text{ball}(v_k, r_k) \cap \mathbb{E} \setminus C \neq \emptyset$.

not finished

For the reverse direction, let x be an arbitrary point in $\text{cone}(\text{int}(C))$. We are to prove that $x \in \text{int}(\text{cone}(C))$.

Since $x \in \text{cone}(\text{int}(C))$, $\exists n \in \mathbb{N}$, $\exists \lambda_1, \dots, \lambda_n > 0$, $\exists v_1, \dots, v_n \in \text{int}(C)$ such that $x = \sum_{i=1}^n \lambda_i v_i$.

Since $v_i \in \text{int}(C)$ for each $i \in \{1, \dots, n\}$, $\exists r_i$ such that $\text{ball}(v_i, r_i) \subseteq C$.

Define $R := \min\{\lambda_i r_i\}_{i=1}^n$.

Say $R = \lambda_k r_k$ for some $k \in \{1, \dots, n\}$.

Let y be an arbitrary point in $\text{ball}(x, R)$.

Since $y \in \text{ball}(x, R)$, $\exists w$ such that $\|w\| < R$ and $y = x + w$.

$$\begin{aligned} y &= \sum_{i=1}^n \lambda_i v_i + w \\ &= \sum_{i \neq k} \lambda_i v_i + \lambda_k v_k + w \\ &= \sum_{i \neq k} \lambda_i v_i + \lambda_k (v_k + w/\lambda_k). \end{aligned}$$

Since $\|w\| < R$, $\|w/\lambda_k\| < R/\lambda_k = r_k$.

Since $\|w/\lambda_k\| < r_k$, $v_k + w/\lambda_k \in \text{ball}(v_k, r_k)$.

So $v_k + w/\lambda_k \in C$.

So $y \in \text{cone}(C)$.

Since $\forall y \in \text{ball}(x, R)$, $y \in \text{cone}(C)$, $\text{ball}(x, R) \subseteq \text{cone}(C)$.

Since $\exists r$ such that $\text{ball}(x, r) \subseteq \text{cone}(C)$, $x \in \text{int}(\text{cone}(C))$.

This proves $\text{cone}(\text{int}(C)) \subseteq \text{int}(\text{cone}(C))$. ■

Proposition 4.3.2. *Let C be a convex set in \mathbb{E} . Assume $\text{int}(C) \neq \emptyset$ and $0 \in C$. Then*

$$0 \in \text{int}(C) \iff \text{cone}(C) = \mathbb{E}.$$

Proof. For one direction, assume that $0 \in \text{int}(C)$. We are to prove that $\text{cone}(C) = \mathbb{E}$. Clearly

$$\text{cone}(C) \subseteq \mathbb{E}.$$

Since $0 \in \text{int}(C)$, $\exists r > 0$ such that $\text{ball}(0, r) \subseteq C$. Since $\text{ball}(0, r) \subseteq C$, $\text{cone}(\text{ball}(0, r)) \subseteq \text{cone}(C)$. Since $\text{cone}(\text{ball}(0, r)) = \mathbb{E}$ and $\text{cone}(\text{ball}(0, r)) \subseteq \text{cone}(C)$, we get

$$\mathbb{E} \subseteq \text{cone}(C).$$

For the reverse direction, assume that $\text{cone}(C) = \mathbb{E}$. We are to prove that $0 \in \text{int}(C)$.

$$\mathbb{E} = \text{int}(\mathbb{E}) = \text{int}(\text{cone}(C)) = \text{cone}(\text{int}(C)).$$

If $0 \notin \text{int}(C)$, then $0 \notin \text{cone}(\text{int}(C))$. So $0 \in \text{int}(C)$. ■

Chapter 5

Tangent Cones and Normal Cones

5.1 Definitions

Definition (Tangent Cones). *Let C be a non-empty convex set in \mathbb{E} . Let x be a point in \mathbb{E} . We define the **tangent cone** to C at point x , denoted by $T_C(x)$, to be a set given by*

$$T_C(x) := \begin{cases} \text{clcone}(C - x), & \text{if } x \in C, \text{ or} \\ \emptyset, & \text{if } x \notin C. \end{cases}$$

Definition (Normal Cones). *Let C be a non-empty convex set in \mathbb{E} . Let x be a point in \mathbb{E} . We define the **normal cone** to C at point x , denoted by $N_C(x)$, to be a set given by*

$$N_C(x) := \begin{cases} \{v \in \mathbb{E} : \forall y \in C - x, \langle y, v \rangle \leq 0\}, & \text{if } x \in C, \text{ or} \\ \emptyset, & \text{if } x \notin C. \end{cases}$$

5.2 Basic Properties

Proposition 5.2.1. *Let C be a closed convex set in \mathbb{E} . Let x be a point in \mathbb{E} . Then $T_C(x)$ and $N_C(x)$ are closed convex cones.*

Proof.

If $C = \emptyset$, then $T_C(x) = N_C(x) = \emptyset$.

If $C \neq \emptyset$ and $x \notin C$, then $T_C(x) = N_C(x) = \emptyset$.

So now I assume that $C \neq \emptyset$ and $x \in C$.

Tangent Cone is Closed:

By definition, $T_C(x) = \text{clcone}(C - x)$. So $T_C(x)$ is a closed.

Tangent Cone is Convex:

$$\begin{array}{ll}
 C \text{ is convex} & \\
 \Downarrow & \text{since affine mapping preserves convexity} \\
 C - x \text{ is convex} & \\
 \Downarrow & \text{since the cone operator preserves convexity} \\
 \text{cone}(C - x) \text{ is convex} & \\
 \Downarrow & \text{since the cl operator preserves convexity} \\
 \text{cl}(\text{cone}(C - x)) \text{ is convex} & \\
 \Downarrow & \text{since } \text{cl} \circ \text{cone} = \text{clcone} \\
 \text{clcone}(C - x) \text{ is convex} &
 \end{array}$$

That is, $T_C(x)$ is convex.

Tangent Cone is a Cone

By definition, $T_C(x) = \text{clcone}(C - x)$. So $T_C(x)$ is a cone.

Normal Cone is Closed:

Let $\{x_i\}_{i \in \mathbb{N}}$ be an arbitrary sequence in $N_C(x)$ that converges to some point in \mathbb{E} .

Say $x_i \rightarrow x_\infty$.

Let y be an arbitrary point in $C - x$.

Since $x_i \in N_C(x)$ and $y \in C - x$, by definition of $N_C(x)$, we get $\langle x_i, y \rangle \leq 0$.

Since $\langle x_i, y \rangle \leq 0$ for any $i \in \mathbb{N}$ and $x_i \rightarrow x_\infty$, we get $\langle x_\infty, y \rangle \leq 0$.

Since $\forall y \in C - x$, $\langle x_\infty, y \rangle \leq 0$, by definition of $N_C(x)$, we get $x_\infty \in N_C(x)$.

Since any convergent sequence whose terms are in $N_C(x)$ has its limit also in $N_C(x)$, $N_C(x)$ is closed.

Normal Cone is Convex:

Let u and v be arbitrary points in $N_C(x)$.

Let λ be an arbitrary number in $(0, 1)$.

Define point z as $z := \lambda u + (1 - \lambda)v$.

Let y be an arbitrary point in $C - x$.

Since $u \in N_C(x)$, $\langle u, y \rangle \leq 0$.

Since $v \in N_C(x)$, $\langle v, y \rangle \leq 0$.

$$\begin{aligned}
 & \langle z, y \rangle \\
 &= \langle \lambda u + (1 - \lambda)v, y \rangle \\
 &= \lambda \langle u, y \rangle + (1 - \lambda) \langle v, y \rangle \\
 &\leq \lambda 0 + (1 - \lambda) 0
 \end{aligned}$$

$$= 0.$$

That is, $\langle z, y \rangle \leq 0$.

Since $\forall y \in C - x$, $\langle z, y \rangle \leq 0$, we get $z \in N_C(x)$.

That is, $\lambda u + (1 - \lambda)v \in N_C(x)$.

Since $\forall u, v \in N_C(x)$, $\forall \lambda \in (0, 1)$, $\lambda u + (1 - \lambda)v \in N_C(x)$, we get $N_C(x)$ is convex.

Normal Cone is a Cone:

Let v be an arbitrary point in $N_C(x)$.

Let λ be an arbitrary number such that $\lambda > 0$.

Let y be an arbitrary point in $C - x$.

Since $v \in N_C(x)$, $\langle v, y \rangle \leq 0$.

Since $\langle v, y \rangle \leq 0$ and $\lambda > 0$, $\langle \lambda v, y \rangle \leq 0$.

Since $\forall y \in C - x$, $\langle \lambda v, y \rangle \leq 0$, we get $\lambda v \in N_C(x)$.

Since $\forall v \in N_C(x)$, $\forall \lambda > 0$, $\lambda v \in N_C(x)$, we get $N_C(x)$ is a cone. ■

Proposition 5.2.2. *Let C be a non-empty closed convex set in \mathbb{E} . Let x be a point in C . Let n be a point in \mathbb{E} . Then*

$$n \in N_C(x) \iff \forall t \in T_C(x), \langle n, t \rangle \leq 0.$$

Proof.

For one direction, assume that $n \in N_C(x)$.

We are to prove that

$$\forall t \in T_C(x), \quad \langle n, t \rangle \leq 0.$$

Let t be an arbitrary point in $T_C(x)$.

Since $t \in T_C(x) = \text{cl}(\text{cone}(C - x))$,

$$\exists \{t_i\}_{i \in \mathbb{N}} \subseteq \text{cone}(C - x), \text{ such that } t_i \rightarrow t. \quad (1)$$

Since $t_i \in \text{cone}(C - x)$,

$$\forall i \in \mathbb{N}, \exists \lambda_i \in \mathbb{R}_{++}, \exists c_i \in C \text{ such that } t_i = \lambda_i(c_i - x). \quad (2)$$

Since $n \in N_C(x)$ and $c_i \in C$,

$$\langle n, c_i - x \rangle \leq 0. \quad (3)$$

Now using (2) and (3), we have

$$\begin{aligned} & \langle n, t_i \rangle \\ &= \langle n, \lambda_i(c_i - x) \rangle, & \text{since } t_i = \lambda_i(c_i - x) \\ &= \lambda_i \langle n, c_i - x \rangle \end{aligned}$$

$$\begin{aligned}
&\leq \lambda_i \cdot 0, & \text{since } \langle n, c_i - x \rangle \leq 0 \\
&= 0.
\end{aligned}$$

That is,

$$\forall i \in \mathbb{N}, \quad \langle n, t_i \rangle \leq 0.$$

Since $\langle n, t_i \rangle \leq 0$ for each $i \in \mathbb{N}$ and $t_i \rightarrow t$, we get

$$\langle n, t \rangle \leq 0.$$

For the reverse direction, assume that n is a vector such that

$$\forall t \in T_C(x), \quad \langle n, t \rangle \leq 0.$$

We are to prove that $n \in N_C(x)$.

Let y be an arbitrary point in $C - x$.

Since $C - x \subseteq \text{clcone}(C - x) = T_C(x)$ and $y \in C - x$, we get $y \in T_C(x)$.

Since $y \in T_C(x)$ and $\forall t \in T_C(x), \langle n, t \rangle \leq 0$, we get $\langle n, y \rangle \leq 0$.

Since $\forall y \in C - x, \langle n, y \rangle \leq 0$, we get $n \in N_C(x)$. ■

Theorem 3. *Let C be a closed convex set in \mathbb{E} such that $\text{int}(C) \neq \emptyset$. Let x be a point in \mathbb{E} . Then*

$$x \in \text{int}(C) \iff T_C(x) = \mathbb{E} \iff N_C(x) = \{0\}.$$

Proof.

Part 1.

$x \in \text{int}(C)$ if and only if $0 \in \text{int}(C - x)$, if and only if $\text{clcone}(C - x) = \mathbb{E}$.

Part 2.

For one direction, assume that $T_C(x) = \mathbb{E}$.

We are to prove that $N_C(x) = \{0\}$.

Consider $n = 0$.

Since

$$\forall t \in T_C(x), \quad \langle 0, t \rangle = 0 \leq 0,$$

we get $0 \in N_C(x)$.

Let n be an arbitrary vector in $N_C(x)$.

By another proposition, we have

$$\begin{aligned}
&n \in N_C(x) \\
&\iff \forall t \in T_C(x) = \mathbb{E}, \langle n, t \rangle \leq 0 \\
&\implies \text{for } t = n, \langle n, t \rangle = \langle n, n \rangle \leq 0
\end{aligned}$$

$$\implies n = 0.$$

That is, $n \in N_C(x) \implies n = 0$.

So $N_C(x) = \{0\}$.

For the reverse direction, assume that $N_C(x) = \{0\}$.

We are to prove that $T_C(x) = \mathbb{E}$.

Clearly $T_C(x) \subseteq \mathbb{E}$.

For $\mathbb{E} \subseteq T_C(x)$, let x be an arbitrary point in \mathbb{E} .

Define $p := \text{proj}_{T_C(x)}(x)$.

Since $p = \text{proj}_{T_C(x)}(x)$,

$$\forall y \in T_C(x), \quad \langle x - p, y - p \rangle \leq 0. \quad (1)$$

Since $p = \text{proj}_{T_C(x)}(x)$, $p \in T_C(x)$.

Since $p \in T_C(x)$ and $T_C(x)$ is a cone,

$$2p \in T_C(x). \quad (2)$$

Apply (1) to $y = 2p$, we get

$$\langle x - p, 2p - p \rangle = \langle x - p, p \rangle \leq 0. \quad (3)$$

Since $T_C(x)$ is a closed cone,

$$0 \in T_C(x). \quad (4)$$

Apply (1) to $y = 0$, we get

$$\langle x - p, 0 - p \rangle = \langle x - p, -p \rangle \leq 0. \quad (5)$$

From (3) and (5), we get

$$\langle x - p, p \rangle = 0.$$

So (1) becomes

$$\forall y \in T_C(x), \quad \langle x - p, y \rangle \leq 0.$$

So $x - p \in N_C(x)$.

So $x - p = 0$.

So $x = p$.

So $x \in T_C(x)$.

Since $\forall x \in \mathbb{E}, x \in T_C(x)$, we get

$$\mathbb{E} \subseteq T_C(x).$$

■

5.3 Arithmetic Properties

Proposition 5.3.1. *Let C and D be convex subsets of \mathbb{E} . Let x be a point in \mathbb{E} . Then*

$$N_C(x) + N_D(x) \subseteq N_{C \cap D}(x).$$

Proof.

If C or D is empty, then $N_C(x) + N_D(x) = N_{C \cap D}(x) = \emptyset$.

So now I assume that $C, D \neq \emptyset$.

If $x \notin C \cap D$, then $N_C(x) + N_D(x) = N_{C \cap D}(x) = \emptyset$.

So now I assume that $x \in C \cap D$.

Let v be an arbitrary point in $N_C(x) + N_D(x)$.

Since $v \in N_C(x) + N_D(x)$, $\exists u \in N_C(x)$, $\exists w \in N_D(x)$ such that $v = u + w$.

Since $u \in N_C(x)$, $\forall y \in C - x$, $\langle u, y \rangle \leq 0$.

Since $w \in N_D(x)$, $\forall y \in D - x$, $\langle w, y \rangle \leq 0$.

Let y be an arbitrary point in $C \cap D - x$.

Since $y \in C \cap D - x$, we get $y \in C - x$ and $y \in D - x$.

$$\begin{aligned} & \langle v, y \rangle \\ &= \langle u + w, y \rangle \\ &= \langle u, y \rangle + \langle w, y \rangle \\ &\leq 0 + 0 = 0. \end{aligned}$$

This is true for any $y \in C \cap D - x$.

So $v \in N_{C \cap D}(x)$.

This is true for any $v \in N_C(x) + N_D(x)$.

So $N_C(x) + N_D(x) \subseteq N_{C \cap D}(x)$.

■

Theorem 4. *Let C and D be convex sets in \mathbb{E} . Assume that $\text{ri}(C) \cap \text{ri}(D) \neq \emptyset$. Let x be a point in $C \cap D$. Then*

$$N_{C \cap D}(x) = N_C(x) + N_D(x).$$

5.4 Other Properties

Proposition 5.4.1. *Let f be a proper, convex, and lower semi-continuous function from \mathbb{E} to \mathbb{R}^* . Let x be a point in $\text{dom}(f)$. Let u be a point in \mathbb{E} . Then $u \in \partial f(x)$ if and only if $(u, -1) \in N_{\text{epi}(f)}(x, f(x))$.*

Proof.

$$\begin{aligned}
& u \in \partial f(x) \\
\iff & \forall y \in \mathbb{E}, f(y) \geq f(x) + \langle u, y - x \rangle \\
\iff & \forall y \in \text{dom}(f), f(y) \geq f(x) + \langle u, y - x \rangle \\
\iff & \forall (y, \beta) \in \text{epi}(f), f(x) + \langle u, y - x \rangle \leq \beta \\
\iff & \forall (y, \beta) \in \text{epi}(f), \langle (u, -1), (y - x, \beta - f(x)) \rangle \leq 0 \\
\iff & \forall (y, \beta) \in \text{epi}(f), \langle (u, -1), (y, \beta) - (x, f(x)) \rangle \leq 0 \\
\iff & (u, -1) \in N_{\text{epi}(f)}(x, f(x)).
\end{aligned}$$

■

Chapter 6

Dual Cones and Polar Cones

6.1 Definitions

Definition (Dual Cone). *Let K be a cone in \mathbb{E} . We define the **dual cone** of K , denoted by K^* , to be the set given by*

$$K^* := \{x \in \mathbb{E} : \forall k \in K, x \cdot k \geq 0\}.$$

Definition (Polar Cone). *Let \mathbb{E} be some Euclidean space. Let S be some set in the space. We define the **polar cone** of S , denoted by C° , to be the set given by*

$$C^\circ := \{y \in \mathbb{E} : \forall x \in C, \langle y, x \rangle \leq 0\}.$$

6.2 Properties

Proposition 6.2.1. *If S is a linear subspace of some Euclidean space \mathbb{E} , then $S^\circ = S^\perp$.*

Chapter 7

Extreme Points

7.1 Definitions

Definition (Extreme Points). *Let \mathbb{E} be some Euclidean space. Let C be a nonempty convex set in the space. Let x be some point in C . We say that x is an **extreme point** of C if it does not lie between any two distinct points in C .*

Definition (Extreme Points). *Let \mathbb{E} be some Euclidean space. Let C be a nonempty convex set in the space. Let x be some point in C . We say that x is an **extreme point** of C if $C \setminus \{x\}$ is convex.*

Definition (Extreme Points). *Let \mathbb{E} be some Euclidean space. Let C be a nonempty convex set in the space. Let x be some point in C . We say that x is an **extreme point** of C if $\{x\}$ is a face of C .*

Proposition 7.1.1. *The three definitions of extreme point are equivalent.*

Def 1 \iff Def 2.

For one direction, assume that x does not lie between any two distinct points in C . We are to prove that $C \setminus \{x\}$ is convex. Let x_1 and x_2 be two arbitrary distinct points in $C \setminus \{x\}$. Let λ be an arbitrary number in $(0, 1)$. Define a point y as $y := \lambda x_1 + (1 - \lambda)x_2$. Since C is convex, $x_1, x_2 \in C$, and $\lambda \in (0, 1)$, we get $y \in C$. Since x does not lie between any two distinct points in C , $y \neq x$. So $y \in C \setminus \{x\}$. That is, I have proved that

$$\forall x_1, x_2 \in C \setminus \{x\}, \forall \lambda \in (0, 1), \quad y = \lambda x_1 + (1 - \lambda)x_2 \in C \setminus \{x\}.$$

By definition, $C \setminus \{x\}$ is convex.

For the reverse direction, assume that $C \setminus \{x\}$ is convex. We are to prove that x does not lie between any two distinct points in C . Assume for the sake of contradiction that

x does lie between two distinct points in C . Say $x = \lambda x_1 + (1 - \lambda)x_2$ where $x_1, x_2 \in C$, $x_1 \neq x_2$, and $\lambda \in (0, 1)$. Clearly $x \neq x_1$ and $x \neq x_2$. So $x_1, x_2 \in C \setminus \{x\}$. Since $C \setminus \{x\}$ is convex, $x_1, x_2 \in C \setminus \{x\}$, and $\lambda \in (0, 1)$, we get $x = \lambda x_1 + (1 - \lambda)x_2 \in C \setminus \{x\}$. This leads to a contradiction. So the assumption that x lies between two distinct points in C does not hold. i.e. x does not lie between two distinct points in C . ■

7.2 Properties

Proposition 7.2.1. *If C is nonempty, convex, and compact, then $\text{extr}(C) \neq \emptyset$.*

Theorem 5 (Minkowski). *A compact convex set is the convex hull of its extreme points.*

Chapter 8

Projection

8.1 Definitions

Definition (Projection). *Let \mathcal{H} be a Hilbert space. Let S be a non-empty set in the space. Let x be a point in the space. We define the **projection** of x onto S , denoted by $\text{proj}_S(x)$, to be a point given by*

$$\text{proj}_S(x) := \operatorname{argmin}_{p \in S} \|p - x\|.$$

i.e., $\text{proj}_S(x)$ is the closest point in S to x .

Proposition 8.1.1 (Existence). *If S is non-empty and closed, then the projection $\text{proj}_S(x)$ exists.*

Proof. Define for an $n \in \mathbb{N}$ a point c_n to be a point in S that satisfies

$$\lim_{i \in \mathbb{N}} \|c_i - x\| = d_S(x) \text{ where } d_S(x) = \inf_{p \in S} \|p - x\|.$$

Since \mathcal{H} is a Hilbert space, the norm $\|\cdot\|$ on \mathcal{H} satisfies the Parallelogram Law. So

$$\begin{aligned} \|c_m - c_n\|^2 &= 2\|c_m - x\|^2 + 2\|c_n - x\|^2 - \|c_m + c_n - 2x\|^2 \\ &= 2\|c_m - x\|^2 + 2\|c_n - x\|^2 - 4\left\|\frac{c_m + c_n}{2} - x\right\|^2 \\ &\leq 2\|c_m - x\|^2 + 2\|c_n - x\|^2 - 4d_S(x) \\ &\rightarrow 2d_S(x) + 2d_S(x) - 4d_S(x) = 0. \end{aligned}$$

So the sequence $(c_i)_{i \in \mathbb{N}}$ is Cauchy. Since \mathcal{H} is a Hilbert space, it is complete. So $(c_i)_{i \in \mathbb{N}}$ converges. Since S is closed, and $(c_i)_{i \in \mathbb{N}}$ is a Cauchy sequence in S , $p := \lim_{i \in \mathbb{N}} c_i \in S$. So $\|p - x\| = \|\lim_{i \in \mathbb{N}} c_i - x\| = \lim_{i \in \mathbb{N}} \|c_i - x\| = d_S(x)$. So p is the minimizer of the distance to the point x over S . So $p = \text{proj}_S(x)$. ■

Proposition 8.1.2 (Uniqueness). *If S is non-empty, closed, and convex, then the projection $\text{proj}_S(x)$ is unique.*

Proof. Let p denote $\text{proj}_S(x)$. Then $\|p - x\| = d_S(x)$. Let q be a point in S such that $\|q - x\| = d_S(x)$. Then by the Parallelogram Law,

$$\begin{aligned} 0 \leq \|p - q\|^2 &= 2\|x - p\|^2 + 2\|q - x\|^2 - 4\left\|x - \frac{1}{2}(p + q)\right\|^2 \\ &\leq 2d_S^2(x) + 2d_S^2(x) - 4d_S^2(x) \\ &= 0. \end{aligned}$$

This shows $\|p - q\| = 0$ and hence $p = q$. Thus the projection is unique. ■

8.2 Properties of the Projection Operator

Proposition 8.2.1 (Idempotent). *The projection operator is idempotent. i.e., if C is a nonempty closed convex set in \mathbb{E} , then $\text{proj}_C = \text{proj}_C \text{proj}_C$.*

Proof. Let x be an arbitrary point in \mathbb{E} . By definition, $\text{proj}_C(x) \in C$. Since $\text{proj}_C(x) \in C$, the closest point in C to $\text{proj}_C(x)$ is $\text{proj}_C(x)$. So $\text{proj}_C \text{proj}_C(x) = \text{proj}_C(x)$. This is true for any $x \in \mathbb{E}$. So $\text{proj}_C = \text{proj}_C \text{proj}_C$. ■

Proposition 8.2.2. *Let C be a nonempty closed convex set in \mathbb{E} . Then the set of fixed points of the operator proj_C is C .*

Proof. For one direction, let x be an arbitrary fixed point of proj_C . We are to prove that $x \in C$. Since x is a fixed point of proj_C , $x = \text{proj}_C(x)$. By definition of projection, $\text{proj}_C(x) \in C$. So $x = \text{proj}_C(x) \in C$.

For the reverse direction, let x be an arbitrary point in C . We are to prove that x is a fixed point of C . Since $x \in C$, the closest point in C to x is x . So $x = \text{proj}_C(x)$. So x is a fixed point of proj_C . ■

Proposition 8.2.3 (Linearity). *Let C be a nonempty closed convex set in \mathbb{E} . Then the operator proj_C is linear if and only if C is a linear subspace.*

Proposition 8.2.4 (Non-expansive). *The projection operator is non-expansive. i.e., if C is a nonempty closed convex set in \mathbb{E} , then $\|\text{proj}_C(x)\| \leq \|x\|$ for any $x \in \mathbb{E}$.*

this is not true. I guess it will be true when C is a linear subspace.

Proposition 8.2.5. *Let C be a nonempty closed convex set in \mathbb{E} . Then proj_C is Lipschitz with constant 1.*

Proof. Let x and y be two arbitrary points in \mathbb{E} . If $\|\text{proj}_C(x) - \text{proj}_C(y)\| = 0$, then $\|\text{proj}_C(x) - \text{proj}_C(y)\| \leq \|x - y\|$. Otherwise,

$$\begin{aligned}
& \|\text{proj}_C(x) - \text{proj}_C(y)\|^2 \\
&= \langle \text{proj}_C(x) - \text{proj}_C(y), \text{proj}_C(x) - \text{proj}_C(y) \rangle \\
&= \langle \text{proj}_C(x) - \text{proj}_C(y), \text{proj}_C(x) - x \rangle \\
&+ \langle \text{proj}_C(x) - \text{proj}_C(y), x - y \rangle \\
&+ \langle \text{proj}_C(x) - \text{proj}_C(y), y - \text{proj}_C(y) \rangle \\
&= \langle x - \text{proj}_C(x), \text{proj}_C(y) - \text{proj}_C(x) \rangle \\
&+ \langle y - \text{proj}_C(y), \text{proj}_C(x) - \text{proj}_C(y) \rangle \\
&+ \langle \text{proj}_C(x) - \text{proj}_C(y), x - y \rangle \\
&\leq 0 + 0 + \langle \text{proj}_C(x) - \text{proj}_C(y), x - y \rangle \\
&= \langle \text{proj}_C(x) - \text{proj}_C(y), x - y \rangle \\
&\leq \|\text{proj}_C(x) - \text{proj}_C(y)\| \|x - y\|.
\end{aligned}$$

That is,

$$\|\text{proj}_C(x) - \text{proj}_C(y)\|^2 \leq \|\text{proj}_C(x) - \text{proj}_C(y)\| \|x - y\|.$$

Dividing both sides by $\|\text{proj}_C(x) - \text{proj}_C(y)\|$ gives

$$\|\text{proj}_C(x) - \text{proj}_C(y)\| \leq \|x - y\|.$$

So proj_C is Lipschitz with constant 1. ■

Proposition 8.2.6 (Firmly Non-expansive). *Let C be a nonempty closed convex set in \mathbb{E} . Then proj_C is firmly non-expansive.*

Proof. This is to prove.

$$\forall x, y \in \mathbb{E}, \quad \|\text{proj}_C(y) - \text{proj}_C(x)\|^2 \leq \langle \text{proj}_C(y) - \text{proj}_C(x), y - x \rangle.$$

$$\begin{aligned}
& \|\text{proj}_C(y) - \text{proj}_C(x)\|^2 \\
&= \langle \text{proj}_C(y) - \text{proj}_C(x), \text{proj}_C(y) - \text{proj}_C(x) \rangle \\
&= \langle \text{proj}_C(y) - \text{proj}_C(x), \text{proj}_C(y) - y \rangle \\
&+ \langle \text{proj}_C(y) - \text{proj}_C(x), y - x \rangle \\
&+ \langle \text{proj}_C(y) - \text{proj}_C(x), x - \text{proj}_C(x) \rangle \\
&\leq 0 + \langle \text{proj}_C(y) - \text{proj}_C(x), y - x \rangle + 0 \\
&= \langle \text{proj}_C(y) - \text{proj}_C(x), y - x \rangle.
\end{aligned}$$
■

8.3 Characterization

Theorem 6 (Projection Theorem). *Let C be a nonempty closed convex set in \mathbb{E} . Let x and p be points in \mathbb{E} . Then $p = \text{proj}_C(x)$ if and only if*

$$\forall y \in C, \quad \langle y - p, x - p \rangle \leq 0.$$

Proof. Let y be an arbitrary point in C . Let α be an arbitrary number in $[0, 1]$. Define $y_\alpha := \alpha y + (1 - \alpha)p$. Now

$$\begin{aligned} p &= \text{proj}_C(x) \\ \iff \forall y \in C, \forall \alpha \in [0, 1], \|x - p\|^2 &\leq \|x - y_\alpha\|^2 \\ \iff \forall y \in C, \forall \alpha \in [0, 1], \|x - p\|^2 &\leq \|x - p - \alpha(y - p)\|^2 \\ \iff \forall y \in C, \langle x - p, y - p \rangle &\leq 0. \end{aligned}$$

■

Chapter 9

Separation

9.1 Definitions

Definition (Separated). Let S_1 and S_2 be two sets in \mathbb{E} . We say that S_1 and S_2 are **separated** if $\exists b \in \mathbb{E} \setminus \{\vec{0}\}$ such that

$$\sup_{s_1 \in S_1} \langle s_1, b \rangle \leq \inf_{s_2 \in S_2} \langle s_2, b \rangle.$$

We say that they are **strongly separated** if the inequality holds strictly.

Definition (Strongly Separated). Let S_1 and S_2 be two sets in \mathbb{E} . We say that they are **strongly separated** if the inequality holds strictly.

Definition (Properly Separated). Let S_1 and S_2 be two sets in \mathbb{E} . We say that S_1 and S_2 are **properly separated** if $\exists b \in \mathbb{E}$ such that

$$\begin{aligned} \sup_{x \in S_1} \langle x, b \rangle &\leq \inf_{y \in S_2} \langle y, b \rangle, \text{ and} \\ \inf_{x \in S_1} \langle x, b \rangle &> \sup_{y \in S_2} \langle y, b \rangle. \end{aligned}$$

Definition (Hyperplane). A **hyperplane** is a set in some Euclidean space \mathbb{E} of the form $\{x \in \mathbb{E} : a^T x = b\}$ for some $a, b \in \mathbb{E}$ and $a \neq \vec{0}$.

Definition (Polyhedrons). A **polyhedron** is a set that is the solution set of finitely many linear inequalities, or equivalently, the intersection of finitely many half spaces.

9.2 Main Results

Proposition 9.2.1. Let C be a nonempty closed convex set in \mathbb{E} . Let x be a point in \mathbb{E} such that $x \notin C$. Then x and C are strongly separated.

Proof.

Define a point p by

$$p := \text{proj}_C(x).$$

Define a point a by

$$a := x - p.$$

To prove that x is strongly separated from C , it suffices to prove that

$$\forall y \in C, \quad \langle y, a \rangle < \langle x, a \rangle.$$

Since $x \notin C$ and C is closed,

$$a \neq 0. \tag{1}$$

Let y be an arbitrary point in C . Since $p = \text{proj}_C(x)$ and $y \in C$,

$$\langle y - p, x - p \rangle \leq 0. \tag{2}$$

$$\begin{aligned} & \langle y, a \rangle \\ & < \langle y, a \rangle + \langle a, a \rangle, \text{ since } a \neq 0 \\ & = \langle y + a, a \rangle \\ & = \langle y + x - p, x - p \rangle, \text{ substitute } a = x - p \\ & = \langle y - p, x - p \rangle + \langle x, x - p \rangle \\ & \leq 0 + \langle x, x - p \rangle, \text{ since } \langle y - p, x - p \rangle \leq 0 \\ & = \langle x, x - p \rangle \\ & = \langle x, a \rangle. \end{aligned}$$

That is,

$$\forall y \in C, \quad \langle y, a \rangle < \langle x, a \rangle.$$

So x is strongly separated from C . ■

Proposition 9.2.2. *Let C_1 be a non-empty closed convex set in \mathbb{E} . Let C_2 be a non-empty compact convex set in \mathbb{E} . Assume that C_1 and C_2 are disjoint. Then C_1 and C_2 are strongly separated.*

Proof. Since C_1 is non-empty closed and convex and C_2 is non-empty compact and convex, we get $C_1 - C_2$ is non-empty closed and convex. Since $C_1 \cap C_2 = \emptyset$, $0 \notin C_1 - C_2$. Since $C_1 - C_2$ is non-empty closed and convex and $0 \in C_1 - C_2$, 0 and $C_1 - C_2$ are strongly separated. Since 0 is strongly separated from $C_1 - C_2$,

$$\exists a \neq 0 \text{ such that } \forall c_1 \in C_1, c_2 \in C_2, \quad \langle c_1 - c_2, a \rangle < \langle 0, a \rangle.$$

That is,

$$\langle c_1, a \rangle < \langle c_2, a \rangle.$$

So C_1 and C_2 are strongly separated. ■

Theorem 7. *Let C_1 and C_2 be non-empty closed convex sets in \mathbb{E} . Assume that C_1 and C_2 are disjoint. Then C_1 and C_2 are separated.*

Proof. For $n \in \mathbb{N}$, define

$$D_n := C_2 \cap \text{ball}(0, n).$$

Then D_n is compact for any $n \in \mathbb{N}$. Since $\{C_1$ is non-empty closed and convex D_n is non-empty compact and convex we get C_1 and D_n are strongly separated for any $n \in \mathbb{N}$. So

$$\forall n \in \mathbb{N}, \exists a_n \in \mathbb{E}, \|a_n\| = 1 \text{ such that } \forall c_1 \in C_1, \forall d_2 \in D_n, \quad \langle c_1, a_n \rangle < \langle d_2, a_n \rangle.$$

Since $\|a_n\| = 1$ for any $n \in \mathbb{N}$, there exists a subsequence $\{a_n\}_{n \in I}$ where I is some infinite subset of \mathbb{N} such that $\{a_n\}_{n \in I}$ converges to some point $a \in \mathbb{E}$. Let x be an arbitrary point in C_1 . Let y be an arbitrary point in C_2 . For large enough n , $y \in D_n$. Since

$$\begin{cases} \langle x, a_n \rangle < \langle y, a_n \rangle \text{ for large enough } n \\ \lim_{n \in I, n \rightarrow \infty} \langle x, a_n \rangle = \langle x, a \rangle \\ \lim_{n \in I, n \rightarrow \infty} \langle y, a_n \rangle = \langle y, a \rangle \end{cases}, \text{ we get}$$

$$\langle x, a \rangle \leq \langle y, a \rangle.$$

Since

$$\exists a \neq 0 \text{ such that } \forall x \in C_1, \forall y \in C_2, \quad \langle x, a \rangle \leq \langle y, a \rangle,$$

by definition of separated, C_1 and C_2 are separated. ■

Proposition 9.2.3. *Let C_1 and C_2 be non-empty convex subsets of \mathbb{E} . Then C_1 and C_2 are properly separated if and only if*

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

Proposition 9.2.4. *Polyhedrons are convex.*

Chapter 10

Convex Functions

10.1 Preliminaries

Definition (Epigraph). Let f be a function from \mathbb{E} to \mathbb{R}^* . We define the **epigraph** of f , denoted by $\text{epi}(f)$, to be the set given by

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

Definition (Domain). Let f be a function from \mathbb{E} to \mathbb{R}^* . We define the **domain** of f , denoted by $\text{dom}(f)$, to be a set given by

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

Definition (Proper). Let f be a function from \mathbb{E} to \mathbb{R}^* . We say that f is **proper** if

$$\begin{aligned} \exists x \in \mathbb{E}, \quad f(x) &\neq +\infty, \text{ and} \\ \forall x \in \mathbb{E}, \quad f(x) &\neq -\infty \end{aligned}$$

10.2 The Indicator Function

Definition (The Indicator Function). Let S be a subset of \mathbb{E} . We define the **indicator function** of S , denoted by δ_S , to be a function from \mathbb{E} to \mathbb{R}^* given by

$$\delta_S(x) = \begin{cases} 0, & x \in S \\ +\infty, & x \notin S. \end{cases}$$

Proposition 10.2.1. Let S be a subset of \mathbb{E} . Then

(1) S is non-empty if and only if δ_S is proper.

(2) S is convex if and only if δ_S is convex.

(3) S is closed if and only if δ_S is lower semi-continuous.

Proof of (1).

For one direction, assume that S is not empty.

We are to prove that δ_S is proper.

Since $S \neq \emptyset$, pick $p \in S$.

Since $p \in S$, $\delta_S(p) = 0$.

Since $\delta_S(p) = 0$, $\exists x_0 \in \mathbb{E}$ such that $\delta_S(x_0) \neq +\infty$.

By definition of the indicator function, it never takes $-\infty$.

Since $\exists x_0 \in \mathbb{E}$ such that $\delta_S(x_0) \neq +\infty$ and $\forall x \in \mathbb{E}$, $\delta_S(x) \neq -\infty$, we get δ_S is proper.

For the reverse direction, assume that δ_S is proper.

We are to prove that S is non-empty.

Assume for the sake of contradiction that S is empty.

Let x be an arbitrary point in \mathbb{E} .

Since $S = \emptyset$, $x \notin S$.

Since $x \notin S$, $\delta_S(x) = +\infty$.

Since $\forall x \in \mathbb{E}$, $\delta_S(x) = +\infty$, by definition of proper function, δ_S is not proper.

This contradicts to the assumption that δ_S is proper.

So the assumption that $S = \emptyset$ is false.

i.e., S is non-empty. ■

Proof of (2).

For one direction, assume that S is convex.

We are to prove that δ_S is convex.

Let x and y be arbitrary points in $\text{dom}(\delta_S)$.

By definition of indicator functions, $\text{dom}(\delta_S) = S$.

So $x, y \in S$.

Let λ be an arbitrary number in $(0, 1)$.

Define point z as $z := \lambda x + (1 - \lambda)y$.

Since $x, y \in S$ and $\lambda \in (0, 1)$ and S is convex and $z = \lambda x + (1 - \lambda)y$, we get $z \in S$.

Since $z \in S$, $\delta_S(z) = 0$.

Since $\lambda \in (0, 1)$ and $\text{range}(\delta_S) = \{0, +\infty\}$, we get $\lambda\delta_S(x) + (1 - \lambda)\delta_S(y) \geq 0$.

Since $\delta_S(z) = 0$ and $\lambda\delta_S(x) + (1 - \lambda)\delta_S(y) \geq 0$, we get $\delta_S(z) \leq \lambda\delta_S(x) + (1 - \lambda)\delta_S(y)$.

That is, $\delta_S(\lambda x + (1 - \lambda)y) \leq \lambda\delta_S(x) + (1 - \lambda)\delta_S(y)$.

Since $\forall x, y \in \text{dom}(\delta_S)$, $\forall \lambda \in (0, 1)$, $\delta_S(\lambda x + (1 - \lambda)y) \leq \lambda\delta_S(x) + (1 - \lambda)\delta_S(y)$, we get δ_S is convex.

For the reverse direction, assume that δ_S is convex.

We are to prove that S is convex.

The case where S is empty is trivial.

So now I assume $S \neq \emptyset$.

Let x and y be arbitrary points in S .

Let λ be an arbitrary number in $(0, 1)$.

Define point z as $z := \lambda x + (1 - \lambda)y$.

Since $x \in S$, $\delta_S(x) = 0$.

Since $y \in S$, $\delta_S(y) = 0$.

Since $\delta_S(x) = \delta_S(y) = 0$, we get $\lambda\delta_S(x) + (1 - \lambda)\delta_S(y) = 0$.

Since $\lambda \in (0, 1)$ and δ_S is convex, $\delta_S(z) \leq \lambda\delta_S(x) + (1 - \lambda)\delta_S(y)$.

Since $\delta_S(z) \leq \lambda\delta_S(x) + (1 - \lambda)\delta_S(y)$ and $\lambda\delta_S(x) + (1 - \lambda)\delta_S(y) = 0$, we get $\delta_S(z) \leq 0$.

By definition of the indicator function, $\delta_S(z) \geq 0$.

Since $\delta_S(z) \leq 0$ and $\delta_S(z) \geq 0$, we get $\delta_S(z) = 0$.

Since $\delta_S(z) = 0$, $z \in S$.

That is, $\lambda x + (1 - \lambda)y \in S$.

Since $\forall x, y \in S, \forall \lambda \in (0, 1), \lambda x + (1 - \lambda)y \in S$, we get S is convex. ■

Proof of (3).

For one direction, assume that S is closed.

We are to prove that δ_S is lower semi-continuous.

Let $\{(x_i, \alpha_i)\}_{i \in \mathbb{N}}$ be an arbitrary sequence in $\text{epi}(\delta_S)$ that converges.

Say its limit is $(x_\infty, \alpha_\infty)$.

Since $(x_i, \alpha_i) \rightarrow (x_\infty, \alpha_\infty)$, $x_i \rightarrow x_\infty$.

Since $(x_i, \alpha_i) \in \text{epi}(\delta_S)$, $\delta_S(x_i) \leq \alpha_i$.

Since $\delta_S(x_i) \leq \alpha_i$ and $\alpha_i \in \mathbb{R}$, we get $\delta_S(x_i) \neq +\infty$.

Since $\delta_S(x_i) \neq +\infty$, $x_i \in S$.

Since $x_i \in S$ and $x_i \rightarrow x_\infty$ and S is closed, $x_\infty \in S$.

Since $x_\infty \in S$, $\delta_S(x_\infty) = 0$.

Since $x_i \in S$, $\delta_S(x_i) = 0$.

Since $\delta_S(x_i) = 0$ and $\delta_S(x_i) \leq \alpha_i$, $\alpha_i \geq 0$.

Since $(x_i, \alpha_i) \rightarrow (x_\infty, \alpha_\infty)$, $\alpha_i \rightarrow \alpha_\infty$.

Since $\alpha_i \geq 0$ and $\alpha_i \rightarrow \alpha_\infty$, $\alpha_\infty \geq 0$.

Since $\delta_S(x_\infty) = 0$ and $\alpha_\infty \geq 0$, $\delta_S(x_\infty) \leq \alpha_\infty$.

Since $\delta_S(x_\infty) \leq \alpha_\infty$, $(x_\infty, \alpha_\infty) \in \text{epi}(\delta_S)$.

Since for any convergent sequence in $\text{epi}(\delta_S)$, its limit is also in $\text{epi}(\delta_S)$, we get $\text{epi}(\delta_S)$ is closed.

For the reverse direction, assume that δ_S is lower semi-continuous.

We are to prove that S is closed.

Let $\{x_i\}_{i \in \mathbb{N}}$ be an arbitrary sequence in S that converges.

Say its limit is x_∞ .

Since $x_i \in S$, $\delta_S(x_i) = 0$.

Since $\delta_S(x_i) = 0$, $(x_i, 0) \in \text{epi}(\delta_S)$.

Since $x_i \rightarrow x_\infty$, $(x_i, 0) \rightarrow (x_\infty, 0)$.

Since $(x_i, 0) \in \text{epi}(\delta_S)$ and $(x_i, 0) \rightarrow (x_\infty, 0)$, $(x_\infty, 0) \in \text{epi}(\delta_S)$.

Since $(x_\infty, 0) \in \text{epi}(\delta_S)$, $\delta_S(x_\infty) \leq 0$.

By definition of the indicator function, $\delta_S(x_\infty) \geq 0$.

Since $\delta_S(x_\infty) \leq 0$ and $\delta_S(x_\infty) \geq 0$, we get $\delta_S(x_\infty) = 0$.

Since $\delta_S(x_\infty) = 0$, $x_\infty \in S$.

Since for any convergent sequence in S , its limit is also in S , we get S is closed. ■

10.3 Definitions

Definition (Convex Function). *Let f be a function from \mathbb{E} to \mathbb{R}^* . We say that f is **convex** if*

$$\forall x, y \in \text{dom}(f), \forall \lambda \in [0, 1], \quad f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Definition (Convex Function). *Let f be a function from \mathbb{E} to \mathbb{R}^* . We say that f is **convex** if the epigraph of f is convex.*

Proposition 10.3.1. *The two definitions of convexity of functions are equivalent.*

Proof.

The case where $\text{dom}(f), \text{epi}(f) = \emptyset$ is trivial.

So now I assume that $\text{dom}(f), \text{epi}(f) \neq \emptyset$.

For one direction, assume that $\forall x, y \in \text{dom}(f), \forall \lambda \in [0, 1]$, we have $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$.

We are to prove that the epigraph of f is convex.

Let (x, α) and (y, β) be two arbitrary points in $\text{epi}(f)$.

Since $(x, \alpha), (y, \beta) \in \text{epi}(f)$, $x, y \in \text{dom}(f)$.

Let λ be an arbitrary number in $[0, 1]$.

Define a point $(z, \gamma) := \lambda(x, \alpha) + (1 - \lambda)(y, \beta)$.

Then $z = \lambda x + (1 - \lambda)y$ and $\gamma = \lambda \alpha + (1 - \lambda)\beta$.

Since $x, y \in \text{dom}(f)$, $\lambda \in [0, 1]$, we get $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$.

Since $(x, \alpha) \in \text{epi}(f)$, $f(x) \leq \alpha$.

Since $(y, \beta) \in \text{epi}(f)$, $f(y) \leq \beta$.

Since $f(x) \leq \alpha$ and $f(y) \leq \beta$ and $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$, we get $f(\lambda x + (1 - \lambda)y) \leq \lambda \alpha + (1 - \lambda)\beta$.

Since $z = \lambda x + (1 - \lambda)y$ and $\gamma = \lambda\alpha + (1 - \lambda)\beta$ and $f(\lambda x + (1 - \lambda)y) \leq \lambda\alpha + (1 - \lambda)\beta$, we get $f(z) \leq \gamma$.

Since $f(z) \leq \gamma$, $(z, \gamma) \in \text{epi}(f)$.

For the reverse direction, assume that $\text{epi}(f)$ is convex.

We are to prove that $\forall x, y \in \text{dom}(f)$, $\forall \lambda \in [0, 1]$, we have $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$.

Let x and y be two arbitrary points in $\text{dom}(f)$.

Let λ be an arbitrary number in $[0, 1]$.

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

Define $\gamma := \lambda f(x) + (1 - \lambda)f(y)$.

Since $(x, f(x)) \in \text{epi}(f)$ and $(y, f(y)) \in \text{epi}(f)$ and $\lambda \in [0, 1]$ and $\text{epi}(f)$ is convex, we get $\lambda(x, f(x)) + (1 - \lambda)(y, f(y)) \in \text{epi}(f)$.

Since $z = \lambda x + (1 - \lambda)y$ and $\gamma = \lambda f(x) + (1 - \lambda)f(y)$ and $\lambda(x, f(x)) + (1 - \lambda)(y, f(y)) \in \text{epi}(f)$, we get $(z, \gamma) \in \text{epi}(f)$.

Since $(z, \gamma) \in \text{epi}(f)$, $f(z) \leq \gamma$.

That is, $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$. ■

10.4 Basic Properties

Proposition 10.4.1 (Necessary Condition). *The domain of a convex function is convex.*

Proof. Follows from the fact that convexity is stable under affine transformations. Define $A((x, \alpha)) := x$. Then $\text{dom}(f) = A(\text{epi}(f))$. ■

Proposition 10.4.2. *The level sets of a convex function are convex.*

Proposition 10.4.3 (Restriction to a Line). *A function $f : \mathbb{E} \rightarrow \mathbb{R}$ is convex if and only if $\forall x \in \text{dom}(f), \forall v \in \mathbb{E}$, the function $g_{x,v} : \mathbb{R} \rightarrow \mathbb{R}$ given by*

$$g_{x,v}(t) = f(x + tv)$$

is convex.

10.5 Differentiable Convex Functions

Proposition 10.5.1. *Let f be a proper convex function from \mathbb{E} to \mathbb{R}^* . Let $x \in \text{dom}(f)$. If f is differentiable at point x , then $\nabla(f)(x)$ is the unique subgradient of f at point x . i.e., $\partial(f)(x) = \{\nabla(f)(x)\}$. Conversely, if the subgradient $\partial(f)(x)$ of f at point x is a singleton set $\{v\}$, then f is differentiable at point x and $\nabla(f)(x) = v$.*

Proof. ■

Proposition 10.5.2 (First-Order Condition). *Let f be a proper function from \mathbb{E} to \mathbb{R}^* . Assume that $\text{dom}(f)$ is convex and open and that f is differentiable on $\text{dom}(f)$. Then f is convex if and only if*

$$\forall x, y \in \text{dom}(f), \quad f(y) - f(x) \geq \langle \nabla(f)(x), y - x \rangle.$$

i.e., the first-order approximation of f is a global under-estimator.

Proof.

Part 1.

For one direction, assume that f is convex. We are to prove that

$$\forall x, y \in \text{dom}(f), \quad f(y) - f(x) \geq \langle \nabla(f)(x), y - x \rangle.$$

Let x and y be arbitrary points in $\text{dom}(f)$. Since f is convex and differentiable at point x , $\nabla(f)(x) = \partial(f)(x)$. So $\nabla(f)(x)$ satisfies the subgradient inequality. That is,

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

Part 2.

For the reverse direction, assume that

$$\forall x, y \in \text{dom}(f), \quad f(y) - f(x) \geq \langle \nabla(f)(x), y - x \rangle.$$

We are to prove that f is convex.

Not Finished. ■

Proposition 10.5.3. *Let f be a proper function from \mathbb{E} to \mathbb{R}^* . Assume that $\text{dom}(f)$ is convex and open and that f is differentiable on $\text{dom}(f)$. Then f is convex if and only if*

$$\forall x, y \in \text{dom}(f), \quad \langle \nabla(f)(x) - \nabla(f)(y), x - y \rangle \geq 0.$$

Proof.

Part 1.

For one direction, assume that f is convex. We are to prove that

$$\forall x, y \in \text{dom}(f), \quad \langle \nabla(f)(x) - \nabla(f)(y), x - y \rangle \geq 0.$$

Let x and y be arbitrary points in $\text{dom}(f)$. Since f is convex and differentiable at point x , $\nabla(f)(x) = \partial(f)(x)$. So $\nabla(f)(x)$ satisfies the subgradient inequality. That is,

$$f(y) - f(x) \geq \langle \nabla(f)(x), y - x \rangle. \tag{1}$$

Since f is convex and differentiable at point y , $\nabla(f)(y) = \partial(f)(y)$. So $\nabla(f)(y)$ satisfies the subgradient inequality. That is,

$$f(x) - f(y) \geq \langle \nabla(f)(y), x - y \rangle. \quad (2)$$

Take the sum of inequalities (1) and (2), we get

$$\begin{aligned} & (f(y) - f(x)) + (f(x) - f(y)) \geq \langle \nabla(f)(x), y - x \rangle + \langle \nabla(f)(y), x - y \rangle \\ \implies & 0 \geq -\langle \nabla(f)(x), x - y \rangle + \langle \nabla(f)(y), x - y \rangle \\ \implies & \langle \nabla(f)(x), x - y \rangle - \langle \nabla(f)(y), x - y \rangle \geq 0 \\ \implies & \langle \nabla(f)(x) - \nabla(f)(y), x - y \rangle \geq 0. \end{aligned}$$

Part 2.

For the reverse direction, assume that

$$\forall x, y \in \text{dom}(f), \quad \langle \nabla(f)(x) - \nabla(f)(y), x - y \rangle \geq 0.$$

We are to prove that f is convex. Let x and y be arbitrary points in $\text{dom}(f)$. Define a function φ on $(0, 1)$ by

$$\varphi(\lambda) := f(\lambda x + (1 - \lambda)y).$$

Notice φ is differentiable and

$$\varphi'(\lambda) = \langle \nabla(f)(\lambda x + (1 - \lambda)y), x - y \rangle.$$

Let α and β be arbitrary numbers in $(0, 1)$. Assume that $\alpha < \beta$. Define two points z_α and z_β by $z_\alpha := \alpha x + (1 - \alpha)y$ and $z_\beta := \beta x + (1 - \beta)y$. Then

$$\begin{aligned} & \varphi'(\beta) - \varphi'(\alpha) \\ &= \langle \nabla(f)(\beta x + (1 - \beta)y), x - y \rangle - \langle \nabla(f)(\alpha x + (1 - \alpha)y), x - y \rangle \\ &= \langle \nabla(f)(z_\beta), x - y \rangle - \langle \nabla(f)(z_\alpha), x - y \rangle \\ &= \langle \nabla(f)(z_\beta) - \nabla(f)(z_\alpha), x - y \rangle \\ &= \left\langle \nabla(f)(z_\beta) - \nabla(f)(z_\alpha), \frac{z_\beta - z_\alpha}{\beta - \alpha} \right\rangle \\ &= \frac{1}{\beta - \alpha} \langle \nabla(f)(z_\beta) - \nabla(f)(z_\alpha), z_\beta - z_\alpha \rangle \\ &\geq \frac{1}{\beta - \alpha} \cdot 0, \text{ by assumption} \\ &= 0. \end{aligned}$$

That is,

$$\forall \alpha, \beta \in (0, 1), \quad \beta > \alpha \implies \varphi'(\beta) - \varphi'(\alpha) \geq 0.$$

So φ' is increasing. So φ is convex. So

$$\varphi(\lambda) \leq \lambda\varphi(1) + (1 - \lambda)\varphi(0).$$

That is,

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

By definition, f is convex. ■

Proposition 10.5.4 (Second-Order Condition). *A twice continuously differentiable real-valued function f defined on a convex set is convex if and only if*

$$\forall x \in \text{dom}(f), \quad \nabla^2 f(x) \geq 0$$

where $\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{bmatrix}$ denotes the Hessian matrix of f at x_0 .

Proposition 10.5.5. *Let f be a twice continuously differentiable function from \mathbb{E} to \mathbb{R} . Then f is convex if and only if $\forall x \in \mathbb{E}$, $\nabla^2 f(x)$ is positive semi-definite.*

10.6 Convexity and Lipschitz-ness

Theorem 8. *Let f be a differentiable convex function from \mathbb{E} to \mathbb{R} . Then the following statements are equivalent.*

(1) ∇f is Lipschitz with constant L .

(2) $\forall x, y \in \mathbb{E}$, we have

$$f(y) - f(x) \leq \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2.$$

(3) $\forall x, y \in \mathbb{E}$, we have

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle + \frac{1}{2L} \|\nabla f(y) - \nabla f(x)\|^2.$$

(4) $\forall x, y \in \mathbb{E}$, we have

$$L \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \|\nabla f(y) - \nabla f(x)\|^2.$$

(1) \implies (2).

Assume that ∇f is Lipschitz with constant L .

Let x and y be two arbitrary points in \mathbb{E} .

$$\begin{aligned}
& f(y) - f(x) \\
&= \int_0^1 \langle \nabla f(x + t(y - x)), y - x \rangle dt \\
&= \langle \nabla f(x), y - x \rangle + \int_0^1 \langle \nabla f(x + t(y - x)) - \nabla f(x), y - x \rangle dt \\
&\leq \langle \nabla f(x), y - x \rangle + \int_0^1 \|\nabla f(x + t(y - x)) - \nabla f(x)\| \|y - x\| dt \\
&\leq \langle \nabla f(x), y - x \rangle + \int_0^1 L \|x + t(y - x) - x\| \|y - x\| dt \\
&= \langle \nabla f(x), y - x \rangle + L \|y - x\|^2 \int_0^1 t dt \\
&= \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2
\end{aligned}$$

That is,

$$f(y) - f(x) \leq \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2.$$

■

Theorem 9. *Let f be a twice continuously differentiable function from \mathbb{E} to \mathbb{R} . Let L be some non-negative number. Then the following statements are equivalent.*

- (1) ∇f is L -Lipschitz.
- (2) $\forall x \in \mathbb{E}, \|\nabla^2 f(x)\| \leq L$.

10.7 Stability of Convexity

Proposition 10.7.1 (Non-Negative Linear Combination). *A non-negative linear combination of proper convex functions is again convex.*

Proof. It suffices to prove that non-negative scalar multiples of convex functions are convex and sums of two convex functions are convex.

Part 1.

Let f be a proper convex function. Let $\alpha \geq 0$ be an arbitrary scalar. We are to prove that αf is convex. Notice $\text{dom}(f) = \text{dom}(\alpha f)$. Since f is proper, $\text{dom}(f) \neq \emptyset$. So $\text{dom}(\alpha f) \neq \emptyset$. Let x and y be two arbitrary points in $\text{dom}(\alpha f)$. Let λ be an arbitrary number in $(0, 1)$. Define a point z as $z := \lambda x + (1 - \lambda)y$. Then

$$(\alpha f)(\lambda x + (1 - \lambda)y) = \alpha f(\lambda x + (1 - \lambda)y)$$

$$\begin{aligned}
&\leq \alpha(\lambda f(x) + (1 - \lambda)f(y)) \\
&= \lambda \alpha f(x) + (1 - \lambda) \alpha f(y) \\
&= \lambda(\alpha f)(x) + (1 - \lambda)(\alpha f)(y).
\end{aligned}$$

That is,

$$\forall x, y \in \text{dom}(\alpha f), \forall \lambda \in (0, 1), \quad (\alpha f)(\lambda x + (1 - \lambda)y) \leq \lambda(\alpha f)(x) + (1 - \lambda)(\alpha f)(y).$$

So by definition, αf is convex.

Part 2.

Let f and g be proper convex functions. We are to prove that $f + g$ is convex. Notice $\text{dom}(f + g) = \text{dom}(f) \cap \text{dom}(g)$. Since f is proper, $\text{dom}(f) \neq \emptyset$. Since g is proper, $\text{dom}(g) \neq \emptyset$. So $\text{dom}(f + g) \neq \emptyset$. Let x and y be two arbitrary points in $\text{dom}(f + g)$. Let λ be an arbitrary number in $(0, 1)$. Define a point z as $z := \lambda x + (1 - \lambda)y$. Then

$$\begin{aligned}
(f + g)(\lambda x + (1 - \lambda)y) &= f(\lambda x + (1 - \lambda)y) + g(\lambda x + (1 - \lambda)y) \\
&\leq \lambda f(x) + (1 - \lambda)f(y) + \lambda g(x) + (1 - \lambda)g(y) \\
&= \lambda(f(x) + g(x)) + (1 - \lambda)(f(y) + g(y)) \\
&= \lambda(f + g)(x) + (1 - \lambda)(f + g)(y).
\end{aligned}$$

That is,

$$\forall x, y \in \text{dom}(f + g), \forall \lambda \in (0, 1), \quad (f + g)(\lambda x + (1 - \lambda)y) \leq \lambda(f + g)(x) + (1 - \lambda)(f + g)(y).$$

So by definition, $f + g$ is convex. ■

Proposition 10.7.2 (Direct Sum). *Direct sums of convex functions are convex.*

Proof. Let z and w be two arbitrary points in $\text{dom}(f \oplus g)$. Let $\lambda \in (0, 1)$ be arbitrary. Say $z = x \oplus y$ and $w = u \oplus v$ where $x, u \in \mathbb{R}^m$ and $y, v \in \mathbb{R}^p$. Since $z \in \text{dom}(f \oplus g)$, $(f \oplus g)(z) \neq +\infty$. That is, $f(x) + g(y) \neq +\infty$. So neither $f(x)$ nor $g(y)$ is $+\infty$. So both $x \in \text{dom}(f)$ and $y \in \text{dom}(g)$. Similarly, we have $u \in \text{dom}(f)$ and $v \in \text{dom}(g)$. Consider the point

$$\begin{aligned}
&\lambda z + (1 - \lambda)w \\
&= \lambda x \oplus y + (1 - \lambda)u \oplus v \\
&= (\lambda x + (1 - \lambda)u) \oplus (\lambda y + (1 - \lambda)v).
\end{aligned}$$

Apply $f \oplus g$ to both sides, we get

$$(f \oplus g)(\lambda z + (1 - \lambda)w)$$

$$\begin{aligned}
&= (f \oplus g) [(\lambda x + (1 - \lambda)u) \oplus (\lambda y + (1 - \lambda)v)] \\
&= f(\lambda x + (1 - \lambda)u) + g(\lambda y + (1 - \lambda)v).
\end{aligned}$$

Since f and g are convex, we get

$$\begin{aligned}
f(\lambda x + (1 - \lambda)u) &\leq \lambda f(x) + (1 - \lambda)f(u), \text{ and} \\
g(\lambda y + (1 - \lambda)v) &\leq \lambda g(y) + (1 - \lambda)g(v).
\end{aligned}$$

So

$$\begin{aligned}
&(f \oplus g)(\lambda z + (1 - \lambda)w) \\
&\leq \lambda f(x) + (1 - \lambda)f(u) + \lambda g(y) + (1 - \lambda)g(v) \\
&= \lambda(f(x) + g(y)) + (1 - \lambda)(f(u) + g(v)) \\
&= \lambda(f \oplus g)(x \oplus y) + (1 - \lambda)(f \oplus g)(u \oplus v) \\
&= \lambda(f \oplus g)(z) + (1 - \lambda)(f \oplus g)(w).
\end{aligned}$$

That is,

$$(f \oplus g)(\lambda z + (1 - \lambda)w) \leq \lambda(f \oplus g)(z) + (1 - \lambda)(f \oplus g)(w).$$

This holds for any $z, w \in \text{dom}(f \oplus g)$ and any $\lambda \in (0, 1)$. So $(f \oplus g)$ is convex. ■

Proposition 10.7.3 (Composition). *The composition of a convex function with an affine function is convex. i.e., if f is convex, then $f(Ax + b)$ is convex.*

Proof. Let x and y be arbitrary points in \mathbb{E} . Let λ be an arbitrary number in $(0, 1)$. Define a point z by $z := \lambda x + (1 - \lambda)y$.

$$\begin{aligned}
&g(\lambda x + (1 - \lambda)y) \\
&= f(A(\lambda x + (1 - \lambda)y) + b) \\
&= f(\lambda Ax + (1 - \lambda)Ay + b), && \text{by linearity of } A \\
&= f(\lambda Ax + (1 - \lambda)Ay + \lambda b + (1 - \lambda)b), && \text{decompose } b \\
&= f(\lambda(Ax + b) + (1 - \lambda)(Ay + b)) \\
&\leq \lambda f(Ax + b) + (1 - \lambda)f(Ay + b), && \text{by convexity of } f \\
&= \lambda g(x) + (1 - \lambda)g(y).
\end{aligned}$$

That is,

$$\forall x, y \in \mathbb{E}, \forall \lambda \in (0, 1), \quad g(\lambda x + (1 - \lambda)y) \leq \lambda g(x) + (1 - \lambda)g(y).$$

So g is convex. ■

Proposition 10.7.4 (Supremum). *The supremum of a collection of convex functions is again convex. i.e., Let $\{f_i\}_{i \in I}$ be a collection of convex functions where I is some index set. Then the function F given by $F := \sup_{i \in I} f_i$ is convex.*

Proof.

$$\begin{aligned}
 (x, \alpha) &\in \text{epi}(F) \\
 \iff \sup_{i \in I} f_i(x) &\leq \alpha \\
 \iff \forall i \in I, f_i(x) &\leq \alpha \\
 \iff \forall i \in I, (x, \alpha) &\in \text{epi}(f_i) \\
 \iff (x, \alpha) \in \bigcap_{i \in I} &\text{epi}(f_i).
 \end{aligned}$$

So $\text{epi}(F) = \bigcap_{i \in I} \text{epi}(f_i)$. Since f_i are convex, $\text{epi}(f_i)$ are convex. Since $\text{epi}(f_i)$ are convex, $\bigcap_{i \in I} \text{epi}(f_i)$ is convex. That is, $\text{epi}(F)$ is convex. Since $\text{epi}(F)$ is convex, F is convex. ■

Proposition 10.7.5 (Pointwise Supremum). *If $f(x, y)$ is convex in x for each y in some set \mathcal{A} , then the function g given by*

$$g(x) = \sup_{y \in \mathcal{A}} f(x, y)$$

is convex.

10.8 Examples

Example 10.8.1. *Affine functions are convex.*

Example 10.8.2. *Norms are convex.*

Proof.

$$\begin{aligned}
 &\|\alpha x + \beta y\| \\
 &\leq \|\alpha x\| + \|\beta y\| \\
 &= |\alpha| \|x\| + |\beta| \|y\| \\
 &= \alpha \|x\| + \beta \|y\|.
 \end{aligned}$$
■

Example 10.8.3. *Square norms are convex.*

Proof Approach 1. Notice $\|\cdot\|^2$ is the direct sum of m squares and squares are convex. So by CO 463 Assignment 2 Problem 3, $\|\cdot\|^2$ is convex. ■

Proof Approach 2. The domain is \mathbb{E} . Let x and y be two points in \mathbb{E} . Let λ be an arbitrary number in $(0, 1)$. Define a point z as $z := \lambda x + (1 - \lambda)y$.

$$\begin{aligned}
& \|\lambda x + (1 - \lambda)y\|^2 \\
&= \|\lambda x\|^2 + \|(1 - \lambda)y\|^2 + 2\langle \lambda x, (1 - \lambda)y \rangle \\
&= \lambda^2 \|x\|^2 + (1 - \lambda)^2 \|y\|^2 + 2\lambda(1 - \lambda)\langle x, y \rangle \\
&\leq \lambda^2 \|x\|^2 + (1 - \lambda)^2 \|y\|^2 + 2\lambda(1 - \lambda)\|x\|\|y\| \\
&\leq \lambda(\lambda - 1)\|x\|^2 + \lambda(\lambda - 1)\|y\|^2 + 2\lambda(1 - \lambda)\|x\|\|y\| \\
&\quad + \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 \\
&= \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 \\
&\quad + \lambda(\lambda - 1)[\|x\|^2 + \|y\|^2 - 2\|x\|\|y\|] \\
&\leq \lambda\|x\|^2 + (1 - \lambda)\|y\|^2
\end{aligned}$$

That is,

$$\forall x, y \in \mathbb{E}, \forall \lambda \in (0, 1), \quad \|\lambda x + (1 - \lambda)y\|^2 \leq \lambda\|x\|^2 + (1 - \lambda)\|y\|^2.$$

So by definition, $\|\cdot\|^2$ is convex. ■

Example 10.8.4. *The distance function to a convex set is convex.*

Example 10.8.5. *The perspective of a convex function is convex. i.e., if $f : \mathbb{E} \rightarrow \mathbb{R}$*

Chapter 11

More Convex Functions

11.1 Strictly Convex

Definition (Strictly Convex). Let f be a proper function from \mathbb{E} to \mathbb{R}^* . We say that f is **strictly convex** if $\forall x, y \in \text{dom}(f)$, $\forall \lambda \in [0, 1]$, we have $f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$, except when $\lambda x + (1 - \lambda)y = x$ or y .

Proposition 11.1.1. Strictly convex functions are convex.

11.2 Strongly Convex

Definition (Strongly Convex). Let f be a proper function from \mathbb{E} to \mathbb{R}^* . We say that f is **strongly convex** with constant β if $\forall x, y \in \text{dom}(f)$, $\forall \lambda \in [0, 1]$, we have

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{\beta}{2}\lambda(1 - \lambda)\|x - y\|^2,$$

for some positive constant β .

Proposition 11.2.1. Strongly convex functions are strictly convex.

Proposition 11.2.2. Let f be a function from \mathbb{E} to \mathbb{R}^* . Then f is β -strongly convex if and only if $f - \frac{\beta}{2}\|\cdot\|^2$ is convex.

Proof. Let β be some positive constnat. Let g denote $f - \frac{\beta}{2}\|\cdot\|^2$. Let $x, y \in \mathbb{E}$ be arbitrary. Let $\lambda \in (0, 1)$ be arbitrary.

$$\begin{aligned} & f \text{ is } \beta\text{-strongly convex} \\ \iff & f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \\ & \quad - \frac{\beta}{2}\lambda(1 - \lambda)\|x - y\|^2 \end{aligned}$$

$$\begin{aligned}
&\iff f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \\
&\quad - \frac{\beta}{2}\lambda(1 - \lambda)(\|x\|^2 + \|y\|^2 - 2\langle x, y \rangle) \\
&\iff f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \\
&\quad - \lambda \frac{\beta}{2}\|x\|^2 + \frac{\beta}{2}\lambda^2\|x\|^2 \\
&\quad - (1 - \lambda) \frac{\beta}{2}\|y\|^2 + \frac{\beta}{2}(1 - \lambda)^2\|y\|^2 \\
&\quad + \beta\lambda(1 - \lambda)\langle x, y \rangle \\
&\iff f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \\
&\quad - \lambda \frac{\beta}{2}\|x\|^2 - (1 - \lambda) \frac{\beta}{2}\|y\|^2 \\
&\quad + \frac{\beta}{2}\|\lambda x\|^2 + \frac{\beta}{2}\|(1 - \lambda)y\|^2 + \beta\langle \lambda x, (1 - \lambda)y \rangle \\
&\iff f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \\
&\quad - \lambda \frac{\beta}{2}\|x\|^2 - (1 - \lambda) \frac{\beta}{2}\|y\|^2 \\
&\quad + \frac{\beta}{2}\|\lambda x + (1 - \lambda)y\|^2 \\
&\iff g(\lambda x + (1 - \lambda)y) \leq \lambda g(x) + (1 - \lambda)g(y) \\
&\iff f - \frac{\beta}{2}\|\cdot\|^2 \text{ is } \beta \text{ convex.}
\end{aligned}$$

■

Question: Can we allow f to take on $-\infty$? Do we need f to be proper?

Proposition 11.2.3. *Let f and g be functions from \mathbb{E} to \mathbb{R}^* . Assume that f is β -strongly convex for some positive constant β and that g is convex. Then $f + g$ is β -strongly convex.*

Question: Can we allow f or g to take on $-\infty$? Do we need f and g to be proper?

Proof.

Since f is β -strongly convex, $f - \frac{\beta}{2}\|\cdot\|^2$ is convex.

Since $f - \frac{\beta}{2}\|\cdot\|^2$ and g are convex, $f + g - \frac{\beta}{2}\|\cdot\|^2$ is convex.

Since $f + g - \frac{\beta}{2}\|\cdot\|^2$ is convex, $f + g$ is β -strongly convex.

■

11.3 Quasiconvex

Definition (Quasiconvex). *Let $f : \mathbb{E} \rightarrow \mathbb{R}$ be a function with convex domain. We say that f is **quasiconvex** if any level set of f is convex.*

Proposition 11.3.1 (Jensen's Inequality for Quasiconvex Functions). *Let f be a quasiconvex function. Then $\forall x, y \in \text{dom}(f)$, $\forall \alpha, \beta \in [0, 1]$ such that $\alpha + \beta = 1$,*

$$f(\alpha x + \beta y) \leq \max\{f(x), f(y)\}.$$

Proposition 11.3.2. *A differentiable real-valued function f with convex domain is convex if and only if $\forall x, y \in \text{dom}(f)$,*

$$f(y) \leq f(x) \implies \nabla f(x) \cdot (y - x) \leq 0. \quad ???$$

Not sure where did this come from but I don't think this is correct.

Chapter 12

Support

12.1 Definitions

Definition (Support Function). *Let S be a subset of \mathbb{E} . We define the **support function** of S , denoted by σ_S , to be a function from \mathbb{E} to \mathbb{R}^* given by*

$$\sigma_S(x) = \sup_{s \in S} \langle x, s \rangle.$$

Definition (Supporting Hyperplane). *Let S be a set in \mathbb{E} with nonempty boundary. Let x_0 be a point in the boundary of S . We define a **supporting hyperplane** H to set S at point x_0 to be a set of the form*

$$H = \{x \in \mathbb{E} : a^T x = a^T x_0\},$$

such that $a \in \mathbb{E}$ and $a \neq \vec{0}$ and $\forall x \in S, a^T x \leq a^T x_0$.

12.2 Properties

Proposition 12.2.1. *The support function of a non-empty set S is proper, convex, and lower semi-continuous.*

Proof.

Part 1. Proper.

Define f_s to be a function from \mathbb{E} to \mathbb{R} by $f_s(x) = \langle s, x \rangle$.

These functions are linear and hence proper, convex, and lower semi-continuous.

Notice $\sigma_S = \sup_{s \in S} f_s$.

So σ_S is convex and lower semi-continuous.

Since $\sigma_S(0) = \sup_{s \in S} \langle 0, s \rangle = 0$, $\exists x_0 \in \mathbb{E}, \sigma_S(x) \neq +\infty$.

Since $\sigma_S(x) = \sup_{s \in S} \langle x, s \rangle \geq \langle x, s \rangle \neq -\infty, \forall x \in \mathbb{E}, \sigma_S(x) \neq -\infty$.

Since $\exists x_0 \in \mathbb{E}, \sigma_S(x) \neq +\infty$ and $\forall x \in \mathbb{E}, \sigma_S(x) \neq -\infty$, by definition, σ_S is proper. ■

Proposition 12.2.2. *The support function of a non-empty and bounded set is continuous.*

Proof.

Let x_0 be an arbitrary point in \mathbb{E} . Let ε be an arbitrary positive number. Define $M := \sup_{y \in C} \|y\| + 1$. Since C is bounded, M is finite. Define $\delta := \varepsilon/M$. Let x be an arbitrary point such that $\|x - x_0\| < \delta$. Let y be an arbitrary point in \mathbb{E} . Then by the Cauchy Schwarz inequality, we have

$$\langle x - x_0, y \rangle \leq \|x - x_0\| \|y\|.$$

That is,

$$\langle x, y \rangle \leq \|x - x_0\| \|y\| + \langle x_0, y \rangle.$$

It follows that

$$\begin{aligned} \sup_{y \in C} \langle x, y \rangle &\leq \sup_{y \in C} (\|x - x_0\| \|y\| + \langle x_0, y \rangle) \\ &\leq \|x - x_0\| \sup_{y \in C} \|y\| + \sup_{y \in C} \langle x_0, y \rangle. \end{aligned}$$

That is,

$$\sigma_C(x) \leq \sigma_C(x_0) + \|x - x_0\| \sup_{y \in C} \|y\|.$$

By definition of δ and M ,

$$\sigma_C(x) < \sigma_C(x_0) + \varepsilon. \tag{1}$$

Similarly, reversing the role of x and x_0 , we can prove that

$$\sigma_C(x_0) < \sigma_C(x) + \varepsilon. \tag{2}$$

From (1) and (2) we get

$$|\sigma_C(x) - \sigma_C(x_0)| < \varepsilon.$$

Since $\forall \varepsilon > 0, \exists \delta > 0$ such that $|\sigma_C(x) - \sigma_C(x_0)| < \varepsilon$ whenever $\|x - x_0\| < \delta$, by definition, σ_C is continuous. ■

Proposition 12.2.3. *Let S be a subset of \mathbb{E} . Then $\sigma_S = \sigma_{\text{conv}(S)} = \sigma_{\text{clconv}(S)}$.*

Proof.

Let x be an arbitrary point in \mathbb{E} .

$$\sigma_S(x) = \sup \{ \langle x, s \rangle : s \in S \}$$

$$\begin{aligned}\sigma_{\text{conv}(S)}(x) &= \sup \{ \langle x, s \rangle : s \in \text{conv}(S) \} \\ \sigma_{\text{clconv}(S)}(x) &= \sup \{ \langle x, s \rangle : s \in \text{clconv}(S) \}.\end{aligned}$$

It is easy to see that by the linearity of inner products,

$$\text{conv} \{ \langle x, s \rangle : s \in S \} = \{ \langle x, s \rangle : s \in \text{conv}(S) \}.$$

It is easy to see that by the linearity and the continuity of inner products,

$$\text{clconv} \{ \langle x, s \rangle : s \in S \} = \{ \langle x, s \rangle : s \in \text{clconv}(S) \}.$$

It is also easy to see that for any subset A of the reals,

$$\sup(A) = \sup(\text{conv}(A)),$$

and

$$\sup(A) = \sup(\text{cl}(A)).$$

So

$$\begin{aligned}\sigma_S(x) &= \sup \{ \langle x, s \rangle : s \in S \} \\ &= \sup \text{conv} \{ \langle x, s \rangle : s \in S \} \\ &= \sup \{ \langle x, s \rangle : s \in \text{conv}(S) \} \\ &= \sigma_{\text{conv}(S)}(x).\end{aligned}$$

That is, $\sigma_S(x) = \sigma_{\text{conv}(S)}(x)$.

$$\begin{aligned}\sigma_S(x) &= \sup \{ \langle x, s \rangle : s \in S \} \\ &= \sup \text{conv} \{ \langle x, s \rangle : s \in S \} \\ &= \sup \{ \langle x, s \rangle : s \in \text{conv}(S) \} \\ &= \sup \text{cl} \{ \langle x, s \rangle : s \in \text{conv}(S) \} \\ &= \sup \{ \langle x, s \rangle : s \in \text{cl}(\text{conv}(S)) \} \\ &= \sup \{ \langle x, s \rangle : s \in \text{clconv}(S) \} \\ &= \sigma_{\text{clconv}(S)}(x).\end{aligned}$$

That is, $\sigma_S(x) = \sigma_{\text{clconv}(S)}(x)$. ■

12.3 Supporting Hyperplane

Theorem 10 (Supporting Hyperplane Theorem). *For any boundary point x_0 of a convex set C , there exists a supporting hyperplane to C at x_0 .*

Chapter 13

Conjugacy

13.1 Definition and Examples

Definition (Convex Conjugate (Legendre–Fenchel Transformation)). *Let f be a function from \mathbb{E} to \mathbb{R}^* . We define the **convex conjugate** of f , denoted by f^* , to be a function also from \mathbb{E} to \mathbb{R}^* given by*

$$f^*(x) := \sup_{y \in \mathbb{E}} \{ \langle y, x \rangle - f(y) \}.$$

Example 13.1.1. *Let S be a subset of \mathbb{E} . Then $\delta_S^* = \sigma_S$.*

Proof. Recall that

$$\delta_S(x) = \begin{cases} 0, & x \in S, \\ +\infty, & x \notin S, \end{cases}$$
$$\sigma_S(x) = \sup_{y \in S} \langle x, y \rangle.$$

Now for any $x \in \mathbb{E}$,

$$\begin{aligned} \delta_S^*(x) &= \sup_{y \in S} (\langle x, y \rangle - \delta_S(y)) \\ &= \sup_{y \in S} (\langle x, y \rangle - 0) \\ &= \sup_{y \in S} \langle x, y \rangle \\ &= \sigma_S(x). \end{aligned}$$

So $\delta_S^* = \sigma_S$. ■

13.2 Basic Properties

Proposition 13.2.1. *The convex conjugate function is convex.*

Proof. If $\text{dom}(f) = \emptyset$, then one can see that $f^* \equiv -\infty$. It is a pointwise supremum of affine functions. ■

Proposition 13.2.2. *The convex conjugate function is lower semi-continuous.*

13.3 Double Conjugate

Proposition 13.3.1. *Let f be any function from \mathbb{E} to \mathbb{R}^* . Then $f^{**} \leq f$.*

Proof. Let x be an arbitrary point in \mathbb{E} .

$$\begin{aligned}
 & f^{**}(x) \\
 &= \sup_{y \in \mathbb{E}} \{ \langle y, x \rangle - f^*(y) \} \\
 &= \sup_{y \in \mathbb{E}} \left\{ \langle y, x \rangle - \sup_{z \in \mathbb{E}} \{ \langle z, y \rangle - f(z) \} \right\} \\
 &\leq \sup_{y \in \mathbb{E}} \left\{ \langle y, x \rangle - \{ \langle x, y \rangle - f(x) \} \right\} \\
 &= \sup_{y \in \mathbb{E}} \{ f(x) \} \\
 &= f(x).
 \end{aligned}$$

That is, $f^{**}(x) \leq f(x)$. Since $\forall x \in \mathbb{E}$, $f^{**}(x) \leq f(x)$, we get $f^{**} \leq f$. ■

Proposition 13.3.2. *Let f be a proper function. Then f is convex and lower semi-continuous if and only if*

$$f^{**} = f.$$

Proposition 13.3.3. *Let f and g be functions from \mathbb{E} to \mathbb{R}^* . Then $f \leq g$ implies $f^* \geq g^*$ and $f^{**} \leq g^{**}$.*

Proof. Let x be an arbitrary point in \mathbb{E} .

$$\begin{aligned}
 & f^*(x) \\
 &= \sup_{y \in \mathbb{E}} \{ \langle y, x \rangle - f(y) \} \\
 &\geq \sup_{y \in \mathbb{E}} \{ \langle y, x \rangle - g(y) \} \\
 &= g^*(x).
 \end{aligned}$$

That is, $f^*(x) \geq g^*(x)$. Since $\forall x \in \mathbb{E}$, $f^*(x) \geq g^*(x)$, we get $f^* \geq g^*$.

Let x be an arbitrary point in \mathbb{E} .

$$\begin{aligned}
 f^{**}(x) &= \sup_{y \in \mathbb{E}} \{ \langle y, x \rangle - f^*(y) \} \\
 &= \sup_{y \in \mathbb{E}} \left\{ \langle y, x \rangle - \sup_{z \in \mathbb{E}} \{ \langle z, y \rangle - f(z) \} \right\} \\
 &\leq \sup_{y \in \mathbb{E}} \left\{ \langle y, x \rangle - \sup_{z \in \mathbb{E}} \{ \langle z, y \rangle - g(z) \} \right\} \\
 &= \sup_{y \in \mathbb{E}} \{ \langle y, x \rangle - g^*(y) \} \\
 &= g^{**}(x).
 \end{aligned}$$

That is, $f^{**}(x) \leq g^{**}(x)$. Since $\forall x \in \mathbb{E}$, $f^{**}(x) \leq g^{**}(x)$, we get $f^{**} \leq g^{**}$. ■

Proposition 13.3.4.

$$\text{epi}(f^{**}) = \text{conv}(\text{epi}(f)).$$

13.4 Conjugates and Sub-Differentials

Theorem 11 (Fenchel-Young). *Let f be a proper function from \mathbb{E} to \mathbb{R}^* . Then $\forall x, y \in \mathbb{E}$, we have*

$$f(x) + f^*(y) \geq \langle x, y \rangle.$$

Proposition 13.4.1. *Let f be a proper closed convex function from \mathbb{E} to \mathbb{R}^* . Then $\forall x, y \in \mathbb{E}$,*

$$y \in \partial f(x) \iff x \in \partial f^*(y) \iff f(x) + f^*(y) = \langle x, y \rangle.$$

Proof of $y \in \partial f(x) \iff x \in \partial f^(y)$.* For one direction, assume that $y \in \partial f(x)$. We are to prove that $x \in \partial f^*(y)$. Consider an arbitrary point $z \in \mathbb{E}$. Since $y \in \partial f(x)$, we get

$$\forall u \in \mathbb{E}, \quad \langle y, u - x \rangle \leq f(u) - f(x).$$

Rearranging yields

$$\forall u \in \mathbb{E}, \quad \langle y, u \rangle - f(u) \leq \langle y, x \rangle - f(x).$$

It follows that

$$\sup_{u \in \mathbb{E}} (\langle y, u \rangle - f(u)) \leq \langle y, x \rangle - f(x). \tag{1}$$

By definition of supremum, we have

$$\sup_{u \in \mathbb{E}} (\langle y, u \rangle - f(u)) \geq \langle y, x \rangle - f(x). \tag{2}$$

From (1) and (2), we get

$$\sup_{u \in \mathbb{E}} (\langle y, u \rangle - f(u)) = \langle y, x \rangle - f(x).$$

That is,

$$f^*(y) = \langle y, x \rangle - f(x).$$

Then

$$\begin{aligned} & f^*(z) - f^*(y) \\ &= \sup_{u \in \mathbb{E}} (\langle z, u \rangle - f(u)) - \sup_{u \in \mathbb{E}} (\langle y, u \rangle - f(u)) \\ &= \sup_{u \in \mathbb{E}} (\langle z, u \rangle - f(u)) - \langle y, x \rangle + f(x) \\ &\geq \langle z, x \rangle - f(x) - \langle y, x \rangle + f(x) \\ &= \langle z - y, x \rangle. \end{aligned}$$

That is,

$$\langle x, z - y \rangle \leq f^*(z) - f^*(y).$$

So $x \in \partial f^*(y)$. This proves

$$y \in \partial f(x) \implies x \in \partial f^*(y).$$

Since $f^{**} = f$, similarly, we can prove that

$$x \in \partial f^*(y) \implies y \in \partial f(x).$$

■

Proposition 13.4.2. *Let f be a proper convex function from \mathbb{E} to \mathbb{R}^* . Let x be a point in \mathbb{E} . Assume that $\partial f(x) \neq \emptyset$. Then $f^{**}(x) = f(x)$.*

Chapter 14

Proximal

14.1 Definition

Definition (Proximal Operator). *Let f be a function from \mathbb{E} to \mathbb{R}^* . We define the **proximal operator** of f , denoted by prox_f , to be a function from \mathbb{E} to $\mathcal{P}(\mathbb{E})$ given by*

$$\text{prox}_f(x) := \underset{y \in \mathbb{E}}{\operatorname{argmin}} \left\{ f(y) + \frac{1}{2} \|y - x\|^2 \right\}.$$

14.2 Basic Properties

Proposition 14.2.1. *Let f be a proper, convex, and lower semi-continuous function from \mathbb{E} to \mathbb{R}^* . Then $\forall x \in \mathbb{E}$, $\text{prox}_f(x)$ is a singleton set.*

Proposition 14.2.2. *Let C be a nonempty closed convex set in \mathbb{E} . Then $\text{prox}_{\delta_C} = \text{proj}_C$.*

Proposition 14.2.3 (Firmly Non-Expansive). *Let f be a proper closed convex function. Then prox_f is firmly non-expansive.*

14.3 Prox Calculus Rules

Proposition 14.3.1 (Scaling and Translation).

Theorem 12 (Norm Composition).

14.4 The Second Prox Theorem

Proposition 14.4.1. *Let f be a proper, convex, and lower semi-continuous function from \mathbb{E} to \mathbb{R}^* . Let x and p be points in \mathbb{E} . Then $p = \text{prox}_f(x)$ if and only if*

$$x - p \in \partial f(p).$$

Proposition 14.4.2. *Let f be a proper, convex, and lower semi-continuous function from \mathbb{E} to \mathbb{R}^* . Let x and p be points in \mathbb{E} . Then $p = \text{prox}_f(x)$ if and only if*

$$\forall y \in \mathbb{E}, \quad \langle y - p, x - p \rangle \leq f(y) - f(p).$$

14.5 Moreau Decomposition

Theorem 13 (Moreau Decomposition). *Let f be a proper closed convex function from \mathbb{E} to \mathbb{R}^* . Then*

$$\text{prox}_f + \text{prox}_{f^*} = \text{id}.$$

Proof. Let x be an arbitrary point in \mathbb{E} . We are to prove that

$$\text{prox}_f(x) + \text{prox}_{f^*}(x) = x.$$

Let p denote $\text{prox}_f(x)$. Since f is proper convex and lower semi-continuous and $p = \text{prox}_f(x)$, we get

$$x - p \in \partial f(p).$$

Since $x - p \in \partial f(p)$, we get $p \in \partial f^*(x - p)$. It follows that $x - p = \text{prox}_{f^*}(x)$. Substitute $p = \text{prox}_f(x)$ and rearrange the equation, we get

$$\text{prox}_f(x) + \text{prox}_{f^*}(x) = x.$$

■

Chapter 15

Ellipsoids

Definition (Ellipsoid). *Let v be a point in some Euclidean space \mathbb{E} . We define an **ellipsoid**, centered at point v , to be a set of the form*

$$\{x \in \mathbb{E} : (x - v)^T A (x - v) = 1\}$$

where A is some d by d positive definite matrix.

15.1 Properties

Proposition 15.1.1. *The eigenvectors of A define the principal axes of the ellipsoid.*