

Convex Analysis

Daniel Mao

Contents

1	Affine Sets	1
1.1	Definitions	1
2	Generalized Interiors	3
2.1	Definitions	3
2.2	Basic Properties	4
2.3	Arithmetic Properties	4
3	Convex Sets	7
3.1	Definitions (bug)	8
3.2	Arithmetic Properties of Convex Sets	10
3.3	The Convex Hull Operator	11
3.4	The Closed Convex Hull Operator	14
3.5	Stability of Convexity	15
3.6	Topological Properties of Convex Sets	20
3.7	Examples of Convex Sets	23
3.8	The Carathéodory Theorem	24
4	Geometric Objects	25
4.1	Definitions	25
4.2	Properties	25
5	Cones	27
5.1	Definitions	27
5.2	Stability of the Cone Structure	28
5.3	Other Properties	28
5.4	Closed Conical Hull	30
5.5	The cone and $\overline{\text{cone}}$ Operators	30
5.6	Dual Cone	32
5.7	Polar Cone	34

5.8	Extreme Rays	34
6	Tangent Cones and Normal Cones	37
6.1	Definitions	37
6.2	Basic Properties	38
6.3	Arithmetic Properties	42
6.4	Other Properties	43
7	Extreme Points and Faces	45
7.1	Extreme Points	46
7.2	Faces	48
7.3	Results for Compact Convex Sets	51
8	Projection Operators	53
8.1	Definitions	53
8.2	Properties (bug)	54
8.3	Examples	56
8.4	Characterizations	57
9	Separation	59
9.1	Definitions	59
9.2	Main Results	59
10	Convex Functions	63
10.1	Preliminaries	64
10.2	The Indicator Function	65
10.3	Definitions	68
10.4	Basic Properties	69
10.5	Differentiable Convex Functions	70
10.6	Convexity and Lipschitz-ness	74
10.7	Stability of Convexity	76
10.8	Examples	80
11	More Convex Functions	83
11.1	Strictly Convex	84
11.2	Strongly Convex	85
11.3	Uniformly Convex	88
11.4	Quasiconvex	89

12 Support	91
12.1 Definitions	91
12.2 Properties	91
12.3 Supporting Hyperplane	94
13 Conjugacy	95
13.1 Definition and Examples	96
13.2 Basic Properties	97
13.3 Double Conjugate	98
13.4 Conjugates and Sub-Differentials	100
14 Proximal Operators	103
14.1 Definitions	103
14.2 Examples	103
14.3 Basic Properties	103
14.4 Prox Calculus Rules	105
14.5 The Second Prox Theorem	105
14.6 Moreau Decomposition	107
15 Ellipsoids	109
15.1 Properties	109

Chapter 1

Affine Sets

1.1 Definitions

DEFINITION 1.1 (Affine Combination). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . 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 $n \in \mathbb{N}$, $v_i \in S$, $\forall i \in [n]$, $\lambda_i \in \mathbb{R}$, $\forall i \in [n]$, and $\sum_{i=1}^n \lambda_i = 1$.

DEFINITION 1.2 (Affine Span). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . We define the **affine span** of S , denoted by $\text{affspan}(S)$, to be the set of all affine combinations of S .

DEFINITION 1.3 (Affine Set). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . We say that S is an **affine set** if and only if $S = \text{affspan}(S)$.

DEFINITION 1.4 (Affine Hull). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset

of \mathcal{V} . We define the **affine hull** of S , denoted by $\text{affhull}(S)$, to be the smallest affine set containing S .

THEOREM 1.5. Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . Then

$$\text{affspan}(S) = \text{affhull}(S).$$

Chapter 2

Generalized Interiors

2.1 Definitions

DEFINITION 2.1 (Relative Interior - 1). Let \mathcal{V} be a normed linear space over \mathbb{R} . Let S be a subset of \mathcal{V} . 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

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

DEFINITION 2.2 (Relative Interior - 2). Let \mathcal{V} be a normed linear space over \mathbb{R} . Let S be a subset of \mathcal{V} . We define the **relative interior** of S , denoted by $\text{ri}(S)$, to be a subset of S given by

$$\text{ri}(S) := \left\{ x \in S : \text{cone}(S - x) = \text{span}(S - x) \right\}.$$

DEFINITION 2.3 (Strong Relative Interior). Let \mathcal{V} be a normed linear space over \mathbb{R} . Let S be a subset of \mathcal{V} . We define the **strong relative interior** of S , denoted by $\text{sri}(S)$, to be a subset of S given by

$$\text{sri}(S) := \left\{ x \in S : \text{cone}(S - x) = \overline{\text{span}}(S - x) \right\}.$$

DEFINITION 2.4 (Quasi-Relative Interior). Let \mathcal{V} be a normed linear space over \mathbb{R} . Let S be a subset of \mathcal{V} . We define the **quasi-relative interior** of S , denoted by $\text{qri}(S)$, to be a subset of S given by

$$\text{qri}(S) := \left\{ x \in S : \overline{\text{cone}}(S - x) = \overline{\text{span}}(S - x) \right\}.$$

2.2 Basic Properties

PROPOSITION 2.5. For a singleton set S , $\text{ri}(S) = S = \text{cl}(S)$.

PROPOSITION 2.6. For any set S , we have $\text{ri}(S) \subseteq S$.

REMARK 2.7. The relative interior operator is not monotonic. 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 2.8. Let \mathcal{V} be a normed linear space over \mathbb{R} . Let S be a subset of \mathcal{V} . Then if $\text{int}(S) \neq \emptyset$ we have $\text{ri}(S) = \text{int}(S)$.

Proof. It suffices to show that $\text{aff}(S) = \mathcal{V}$. Since $\text{int}(S) \neq \emptyset$, we can take $x \in \text{int}(S)$. Then $\exists r > 0$, $\text{ball}(x, r) \subseteq S$. Then

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

So $\text{aff}(S) = \mathcal{V}$. □

2.3 Arithmetic Properties

PROPOSITION 2.9 (Linearity). Let \mathcal{V} be a normed linear space over \mathbb{R} . Let C_1

and C_2 be convex subsets of \mathbb{R} . Let $\lambda_1, \lambda_2 \in \mathbb{R}$. Then

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

PROPOSITION 2.10. 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

Contents

3.1	Definitions (bug)	8
3.2	Arithmetic Properties of Convex Sets	10
3.3	The Convex Hull Operator	11
3.4	The Closed Convex Hull Operator	14
3.5	Stability of Convexity	15
3.6	Topological Properties of Convex Sets	20
3.7	Examples of Convex Sets	23
3.8	The Carathéodory Theorem	24

3.1 Definitions (bug)

DEFINITION 3.1 (Convex Combination). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . We define a **convex combination** of S to be a point x in \mathcal{V} of the form

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

where (1) $n \in \mathbb{N}$, (2) $v_1, \dots, v_n \in S$, (3) $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$, and (4) $\sum_{i=1}^n \lambda_i = 1$.

DEFINITION 3.2 (Convex Span). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . We define the **convex span** of S , denoted by $\text{convspan}(S)$, to be the set of all possible convex combinations of S .

DEFINITION 3.3 (Convex). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathbb{E} . We say that S is **convex** if and only if $S = \text{convspan}(S)$, or equivalently, if

$$\forall x, y \in S, \forall \alpha, \beta \in [0, 1] : \alpha + \beta = 1, \quad \alpha x + \beta y \in S.$$

DEFINITION 3.4 (Convex Hull). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . We define the **convex hull** of S , denoted by $\text{convhull}(S)$, to be the smallest convex set containing S .

PROPOSITION 3.5. Let \mathcal{V} be a vector space over \mathbb{R} . For any subset S of \mathcal{V} , we have $\text{convspan}(S) = \text{convhull}(S)$. They will both be denoted by $\text{conv}(S)$ from now on.

Proof. **Forward Inclusion:** Let x be an arbitrary element of $\text{convspan}(S)$. I will show that $x \in \text{convhull}(S)$. Let C be an arbitrary convex set containing S . Since x is a convex combination of elements in S , x is also a convex combination of elements in C . So $x \in C$. This holds for any convex set in \mathcal{V} containing S . So $x \in \text{convhull}(S)$. So $\text{convspan}(S) \subseteq \text{convhull}(S)$.

Backward Inclusion: I will show that $\text{convhull}(S) \subseteq \text{convspan}(S)$. □

not finished

DEFINITION 3.6 (Pointed). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . We say that S is **pointed** if and only if S contains no whole lines.

3.2 Arithmetic Properties of Convex Sets

PROPOSITION 3.7. Let \mathcal{V} be a vector space over \mathbb{R} . Let C be a convex subset of \mathcal{V} . Let $\lambda_1, \lambda_2 \in \mathbb{R}_+$. Then

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

Proof. The case where $\lambda_1 = 0$ or $\lambda_2 = 0$ is trivial. Now suppose that $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$.

Forward Inclusion: Let x be an arbitrary element of $(\lambda_1 + \lambda_2)C$. I will show that $x \in \lambda_1 C + \lambda_2 C$. Since $x \in (\lambda_1 + \lambda_2)C$, $x = (\lambda_1 + \lambda_2)c$ for some $c \in C$. Since \mathcal{V} is a vector space over \mathbb{R} , $c \in C \subseteq \mathcal{V}$, and $\lambda_1, \lambda_2 \in \mathbb{R}_+ \subseteq \mathbb{R}$, we get $x = (\lambda_1 + \lambda_2)c = \lambda_1 c + \lambda_2 c$. Notice $\lambda_1 c \in \lambda_1 C$, and $\lambda_2 c \in \lambda_2 C$. So $x \in \lambda_1 C + \lambda_2 C$. So $(\lambda_1 + \lambda_2)C \subseteq \lambda_1 C + \lambda_2 C$.

Backward Inclusion: Let x be an arbitrary element of $\lambda_1 C + \lambda_2 C$. I will show that $x \in (\lambda_1 + \lambda_2)C$. Since $x \in \lambda_1 C + \lambda_2 C$, $x = \lambda_1 c_1 + \lambda_2 c_2$ for some $c_1, c_2 \in C$. Notice

$$x = \lambda_1 c_1 + \lambda_2 c_2 = (\lambda_1 + \lambda_2) \left(\underbrace{\frac{\lambda_1}{\lambda_1 + \lambda_2}}_{\in [0,1]} \underbrace{c_1}_{\in C} + \underbrace{\frac{\lambda_2}{\lambda_1 + \lambda_2}}_{\in [0,1]} \underbrace{c_2}_{\in C} \right).$$

Notice the second term is a convex combination of two points in C and hence is in C . So $x \in (\lambda_1 + \lambda_2)C$. So $\lambda_1 C + \lambda_2 C \subseteq (\lambda_1 + \lambda_2)C$. \square

3.3 The Convex Hull Operator

PROPOSITION 3.8 (The Convex Hull Operator). Let \mathcal{V} be a vector space over \mathbb{R} .

1. Expansive

$$\forall S \subseteq \mathcal{V}, \quad S \subseteq \text{conv}(S).$$

2. Monotonic Increasing

$$\forall S_1, S_2 \subseteq \mathcal{V} : S_1 \subseteq S_2, \quad \text{conv}(S_1) \subseteq \text{conv}(S_2).$$

3. Idempotent

$$\forall S \subseteq \mathcal{V}, \quad \text{conv}(\text{conv}(S)) = \text{conv}(S).$$

PROPOSITION 3.9 (Bounded). The convex hull of a bounded set is bounded.

Proof. Let \mathcal{V} be a normed linear space over \mathbb{R} . Let C be a bounded subset of \mathcal{V} . Then $\exists R > 0$ such that $\forall c \in C, \|c\| < R$. Let x be an arbitrary element of $\text{conv}(C)$. Then $\exists n \in \mathbb{Z}_{++}, \exists \lambda_1, \dots, \lambda_n \in [0, 1], \exists c_1, \dots, c_n \in C$ such that $\sum_{i \in [n]} \lambda_i = 1$ and $x = \sum_{i \in [n]} \lambda_i c_i$. Then by the triangle inequality of $\|\cdot\|$, we get

$$\|x\| = \left\| \sum_{i \in [n]} \lambda_i c_i \right\| \leq \sum_{i \in [n]} \lambda_i \|c_i\| < \sum_{i \in [n]} \lambda_i R = 1 \cdot R = R.$$

So $\forall x \in \text{conv}(C), \|x\| < R$. So $\text{conv}(C)$ is bounded. \square

PROPOSITION 3.10 (Open). The convex hull of an open set is open.

Proof Approach (1). Let \mathcal{V} be a topological vector space. Let G be an open subset of \mathcal{V} . I will show that $\text{conv}(G)$ is open. Let $x \in \text{conv}(G)$ be arbitrary. Then $x = \sum_{i \in [n]} \lambda_i g_i$ for some $n \in \mathbb{Z}_{++}, \lambda_i \in [0, 1], \forall i \in [n]$, and $g_i \in G, \forall i \in [n]$. Since G is an open set and $g_i \in G, \forall i \in [n]$, there exist neighborhoods $\mathcal{N}_1, \dots, \mathcal{N}_n$ of g_1, \dots, g_n , respectively, such that $g_i \in \mathcal{N}_i \subseteq G, \forall i \in [n]$. Define $\mathcal{N} := \sum_{i \in [n]} \lambda_i \mathcal{N}_i$. Then \mathcal{N} is a neighborhood of x , and $\mathcal{N} \subseteq \text{conv}(G)$. So $x \in \text{int}(\text{conv}(G))$. So $\text{conv}(G) \subseteq \text{int}(\text{conv}(G))$. So $\text{conv}(G) = \text{int}(\text{conv}(G))$. So $\text{conv}(G)$ is open. \square

Proof Approach (2). Let \mathcal{V} be a topological vector space. Let G be an open subset of \mathcal{V} . I will show that $\text{conv}(G)$ is open. Let $x \in \text{conv}(G)$ be arbitrary. Then $x = \sum_{i \in [n]} \lambda_i g_i$ for some $n \in \mathbb{Z}_{++}$, $\lambda_i \in [0, 1]$, $\forall i \in [n]$, and $g_i \in G$, $\forall i \in [n]$. Let $i_0 \in [n]$ be such that $\lambda_{i_0} \neq 0$. Then

$$x = \sum_{i=1}^n \lambda_i g_i = \left(\sum_{i \neq i_0} \lambda_i g_i \right) + \lambda_{i_0} g_{i_0} \in \left(\sum_{i \neq i_0} \lambda_i g_i \right) + i_0 G \subseteq \text{conv}(G).$$

So

$$\text{conv}(G) = \bigcup_{x \in \text{conv}(G)} \left\{ \sum_{i \neq i_0} \lambda_i g_i + i_0 G \right\}.$$

Note that for each $x \in \text{conv}(G)$, the function $f_x : \mathcal{V} \rightarrow \mathcal{V}$ given by $f_x(v) := \sum_{i \neq i_0} \lambda_i g_i + i_0 v$ is a homeomorphism. So

$$\text{conv}(G) = \bigcup_{x \in \text{conv}(G)} f_x(G)$$

is the union of a collection of open sets and hence is open. \square

REMARK 3.11 (Closed). The convex hull of a closed set need not be closed.

- Example in \mathbb{R}^2 : The set $S := \{(x, y) \in \mathbb{R}^2 : y \geq \frac{1}{1+x^2}\}$ is closed. However, $\text{conv}(S) = \{(x, y) \in \mathbb{R}^2 : y > 0\}$ is open.
- Example in ℓ^∞ : Define for each $n \in \mathbb{Z}_{++}$ a sequence x_n by $x_n^{(n)} := \frac{1}{n}$ and $x_n^{(i)} := 0$, $\forall i \neq n$. Consider the set $S := \{x_n\}_{n \in \mathbb{N}} \cup \{0\}$. Then S is a compact subset of ℓ^∞ . However, $\text{conv}(S)$ contains the elements $\sum_{n=1}^k 2^{-n} x_n$ for $k \in \mathbb{Z}_{++}$. Notice they converge to the sequence $\sum_{n=1}^{\infty} 2^{-n} x_n$, which is not in $\text{conv}(K)$ (it has infinitely many non-zero entries).

PROPOSITION 3.12 (Compact in \mathbb{R}^n). Let K be a compact subset of \mathbb{R}^n . Then $\text{conv}(K)$ is also compact.

PROPOSITION 3.13. Let \mathcal{V} be a normed linear space. Let K be a compact subset of \mathcal{V} . Then $\overline{\text{conv}}(K)$ is pre-compact (totally bounded). Moreover, if \mathcal{V} is complete, then

$\overline{\text{conv}}(K)$ is also compact.

3.4 The Closed Convex Hull Operator

DEFINITION 3.14 (Closed Convex Hull). Let S be a set in some Euclidean space. We define the **closed convex hull** of S , denoted by $\overline{\text{conv}}(S)$, to be the smallest closed convex containing S .

PROPOSITION 3.15. The closed convex hull is the closure of the convex hull.

PROPOSITION 3.16. 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 3.17. 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 3.18. The closure operation and the convex hull operation do not commute in general.

3.5 Stability of Convexity

PROPOSITION 3.19 (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. \square

PROPOSITION 3.20 (Affine Map). Convexity is stable under affine mapping. i.e., the affine image of a convex set is convex.

PROPOSITION 3.21 (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

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

$$= \lambda_1(\lambda x_1 + (1 - \lambda)y_1) + \lambda_2(\lambda x_2 + (1 - \lambda)y_2).$$

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. \square

COROLLARY 3.22. The Minkowski sum of two convex sets is convex.

LEMMA 3.23. Let \mathcal{V} be a normed linear space. Let C be a convex subset of \mathcal{V} . Let $x \in \text{int}(C)$. Let $y \in \text{cl}(C)$. Then

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

Proof Approach (1). Let $t \in (0, 1]$ be arbitrary. Define a point $z \in \mathcal{V}$ by $z := tx + (1 - t)y$. If $t = 1$, then $z = x \in \text{int}(C) \subseteq C$ and we are done. Otherwise, $t \in (0, 1)$. Since $x \in \text{int}(C)$, $\exists r_x > 0$ such that $\text{ball}(x, r_x) \subseteq C$. Define $r_y := \frac{t}{1-t}r_x$. Since $y \in \text{cl}(C)$, $\exists y' \in \text{ball}(y, r_y) \cap C$. Define a point $z' \in \mathcal{V}$ by $z' := tx + (1 - t)y'$. Since $x, y' \in C$, $t \in (0, 1)$, and C is convex, we get $z' \in C$. Define a point $x' \in \mathcal{V}$ by $x' := \frac{1}{t}(z - (1 - t)y')$ so that $z = tx' + (1 - t)y'$. Notice $x = \frac{1}{t}(z' - (1 - t)y')$. So

$$\begin{aligned} \|x - x'\| &= \left\| x - \frac{1}{t}(z' - (1 - t)y') \right\| = \left\| \frac{1}{t}(z' - (1 - t)y') - \frac{1}{t}(z - (1 - t)y') \right\| \\ &= \frac{1}{t} \|z - z'\| = \frac{1}{t} \|tx + (1 - t)y - tx - (1 - t)y'\| \\ &= \frac{1-t}{t} \|y - y'\| \leq \frac{1-t}{t} r_y = \frac{1-t}{t} \frac{t}{1-t} r_x = r_x. \end{aligned}$$

That is, $\|x - x'\| \leq r_x$. So $x' \in \text{ball}(x, r_x) \subseteq C$. Since $x', y' \in C$, $t \in (0, 1)$, and C is convex, we get $z \in C$. \square

Proof Approach (2). Let $t \in (0, 1]$ be arbitrary. If $t = 1$, then $tx + (1 - t)y = x \in \text{int}(C) \subseteq C$ and we are done. Otherwise, $t \in (0, 1)$. Define $B := \text{ball}(0, 1)$. Then for some small enough $\varepsilon > 0$, we have

$$\begin{aligned} tx + (1 - t)y + \varepsilon B &\subseteq tx + (1 - t)(C + \varepsilon B) + \varepsilon B, \text{ since } y \in \text{cl}(C) \\ &= tx + (1 - t)C + (1 - t)\varepsilon B + \varepsilon B \end{aligned}$$

$$\begin{aligned}
&= tx + (2-t)\varepsilon B + (1-t)C \\
&= t\left(x + \frac{2-t}{t}\varepsilon B\right) + (1-t)C \\
&\subseteq tC + (1-t)C, \text{ since } x \in \text{int}(C) \\
&= C.
\end{aligned}$$

□

LEMMA 3.24. 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 3.25 (Interior). Convexity is stable under interior. i.e., the interior of a convex set is convex.

Proof. Let \mathcal{V} be a normed linear space. Let S be a convex subset of \mathcal{V} . 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 3.26 (Relative Interior). Convexity is stable under relative interior. i.e., the relative interior of a convex set is convex.

Proof. Let \mathcal{V} be a normed linear space. Let S be a convex subset of \mathcal{V} . If $\text{ri}(S) = \emptyset$, then we are done. Otherwise, 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. \square

PROPOSITION 3.27 (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. \square

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. \square

PROPOSITION 3.28 (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.

□

3.6 Topological Properties of Convex Sets

THEOREM 3.29. Let \mathcal{V} be a topological vector space. Let C be a convex subset of \mathcal{V} with $\text{int}(C) \neq \emptyset$. Then

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

Proof of 1. Since $C \subseteq \text{cl}(C)$, we have $\text{int}(C) \subseteq \text{int}(\text{cl}(C))$. So there remains only to show that $\text{int}(\text{cl}(C)) \subseteq \text{int}(C)$. Let x be an arbitrary element of $\text{int}(\text{cl}(C))$. Then $\exists \varepsilon > 0$ such that $\text{ball}(x, \varepsilon) \subseteq \text{cl}(C)$. \square

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 3.30. Let C be a convex set. Then

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

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

PROPOSITION 3.31. Let C be a convex set. Then

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

Proof. Forward Direction: Assume that $C \neq \emptyset$. I will show 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 $\text{ri}(C) \neq \emptyset$.

Backward Direction: Assume that $\text{ri}(C) \neq \emptyset$. I will show that $C \neq \emptyset$. Since $\text{ri}(C) \neq \emptyset$ and $\text{ri}(C) \subseteq C$, we get $C \neq \emptyset$. \square

3.7 Examples of Convex Sets

EXAMPLE 3.32. 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 $C = \bigcap_{i \in I} C_i$.

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

□

3.8 The Carathéodory Theorem

THEOREM 3.33 (Carathéodory). Let S be a subset of \mathbb{R}^n . Let x be some point in $\text{conv}(S)$. Then x can be represented as a convex combination of at most $n + 1$ points in S . i.e., x lies in some r -simplex with vertices in S , where $r \leq n$.

Chapter 4

Geometric Objects

4.1 Definitions

DEFINITION 4.1 (Hyperplane). Let \mathbb{E} be a Euclidean space over \mathbb{R} . Let H be a subset of \mathbb{E} . We say that H is a **hyperplane** if and only if H can be expressed as

$$H = \{x \in \mathbb{E} : a^\top x = b\}$$

for some $a \in \mathbb{E} \setminus \{0\}$ and $b \in \mathbb{R}$.

DEFINITION 4.2 (Closed Half-Space). Let \mathbb{E} be a Euclidean space over \mathbb{R} . Let P be a subset of \mathbb{E} . We say that P is a **closed half-space** if and only if P can be expressed as

$$P = \{x \in \mathbb{E} : a^\top x \leq b\}$$

for some $a \in \mathbb{E} \setminus \{0\}$ and $b \in \mathbb{R}$.

DEFINITION 4.3 (Polyhedron). Let \mathbb{E} be a Euclidean space over \mathbb{R} . Let P be a subset of \mathbb{E} . We say that P is a **polyhedron** if and only if P can be expressed as the intersection of finitely many closed half-spaces in \mathbb{E} .

4.2 Properties

PROPOSITION 4.4. Polyhedrons are convex.

Chapter 5

Cones

5.1 Definitions

DEFINITION 5.1 (Conical Combination). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . 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 5.2 (Cone). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . We say that S is a **cone** if and only if $S = \mathbb{R}_{++}S$.

DEFINITION 5.3 (Conical Hull). Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . We define the **conical hull** of S , denoted by $\text{cone}(S)$, to be the intersection of all cones containing S .

PROPOSITION 5.4. Let \mathcal{V} be a vector space over \mathbb{R} . Let S be a subset of \mathcal{V} . Then $\text{cone}(S) = \mathbb{R}_{++}S$.

Proof. Forward Direction: I will show that $\text{cone}(S) \subseteq \mathbb{R}_{++}S$. 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

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

Backward Direction: I will show that $\mathbb{R}_{++}S \subseteq \text{cone}(S)$. 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

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

□

5.2 Stability of the Cone Structure

PROPOSITION 5.5. The closure of a cone is a cone.

Proof. Let \mathcal{V} be a normed linear space. Let C be a cone in \mathcal{V} . I will show that $\text{cl}(C)$ is also a cone. Let x be an arbitrary element of $\text{cl}(C)$. Then $\exists (x_n)_{n \in \mathbb{N}} \subseteq C$ such that $\lim_{n \in \mathbb{N}} x_n = x$. Let $\lambda \in \mathbb{R}_{++}$ be arbitrary. Since C is a cone and $x_n \in C$, $\forall n \in \mathbb{N}$, we have $\lambda x_n \in C$. So $(\lambda x_n)_{n \in \mathbb{N}} \subseteq C$. Moreover, notice $\lim_{n \in \mathbb{N}} \lambda x_n = \lambda \lim_{n \in \mathbb{N}} x_n = \lambda x$. So $\lambda x \in \text{cl}(C)$. This holds for any $\lambda \in \mathbb{R}_{++}$. So $\mathbb{R}_{++}\text{cl}(C) \subseteq \text{cl}(C)$. It is clear that $\text{cl}(C) \subseteq \mathbb{R}_{++}\text{cl}(C)$. So we have $\text{cl}(C) = \mathbb{R}_{++}\text{cl}(C)$. So $\text{cl}(C)$ is a cone. □

5.3 Other Properties

PROPOSITION 5.6. 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 5.7. 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 $0 \in \text{int}(C)$.

□

5.4 Closed Conical Hull

DEFINITION 5.8 (Closed Conical Hull). Let \mathcal{V} be a normed linear space. Let S be a subset of \mathcal{V} . We define the **closed conical hull** of S , denoted by $\overline{\text{cone}}(S)$, to be the intersection of all closed cones containing C .

PROPOSITION 5.9. Let \mathcal{V} be a normed linear space. Let S be a subset of \mathcal{V} . Then

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

Proof. Forward Direction: I will show that $\overline{\text{cone}}(S) \subseteq \text{cl}(\text{cone}(S))$. By Proposition 5.5, $\text{cl}(\text{cone}(S))$ is a closed cone containing S . Since $\overline{\text{cone}}(S)$ is the smallest closed cone containing S , we have $\overline{\text{cone}}(S) \subseteq \text{cl}(\text{cone}(S))$.

Backward Direction: I will show that $\text{cl}(\text{cone}(S)) \subseteq \overline{\text{cone}}(S)$.

$$\begin{aligned} S &\subseteq \overline{\text{cone}}(S) \\ \implies \text{cone}(S) &\subseteq \text{cone}(\overline{\text{cone}}(S)), \text{ since cone is monotonic increasing} \\ \implies \text{cl}(\text{cone}(S)) &\subseteq \text{cl}(\text{cone}(\overline{\text{cone}}(S))), \text{ since cl is monotonic increasing} \\ \iff \text{cl}(\text{cone}(S)) &\subseteq \overline{\text{cone}}(S), \text{ since } \overline{\text{cone}}(S) \text{ is a closed cone.} \end{aligned}$$

This completes the proof.

□

5.5 The cone and $\overline{\text{cone}}$ Operators

PROPOSITION 5.10 (The cone Operator). The cone operator has the following properties.

1. Expansive: $\forall S \subseteq \mathbb{E}$,

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

2. Monotone: $\forall S_1, S_2 \subseteq \mathbb{E}$,

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

3. Idempotence: $\forall S \subseteq \mathbb{E}$,

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

PROPOSITION 5.11 (Bauschke-Combettes, 2017 Book). Let \mathcal{V} be a normed linear space. Let S be a subset of \mathcal{V} . Then

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

This is the smallest convex cone containing 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 $\text{conv}(\text{cone}(S)) = \text{cone}(\text{conv}(S))$.

□

PROPOSITION 5.12 (Bauschke-Combettes, 2017 Book). Let \mathcal{V} be a normed linear space. Let S be a subset of \mathcal{V} . Then

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

This is the smallest closed convex cone containing S .

5.6 Dual Cone

DEFINITION 5.13 (Dual Cone). Let \mathfrak{X} be an inner product space over \mathbb{R} with inner product $\langle \cdot, \cdot \rangle$. Let C be a subset of \mathfrak{X} . We define the **dual cone** of C , denoted by C^* , to be a subset of \mathfrak{X} given by

$$C^* := \{x \in \mathfrak{X} : \forall y \in C, \langle x, y \rangle \geq 0\}.$$

EXAMPLE 5.14. The set of positive semidefinite matrices \mathbb{S}_+^n is self-dual. i.e., $(\mathbb{S}_+^n)^* = \mathbb{S}_+^n$.

PROPOSITION 5.15. The dual cone of any set is a cone.

PROPOSITION 5.16. The dual cone of a convex cone is always a closed convex cone.

PROPOSITION 5.17. Let \mathbb{E} be a Euclidean space. Let K be a convex cone in \mathbb{E} . Then $K^{**} = \text{cl}(K)$.

PROPOSITION 5.18. Let \mathbb{E} be a Euclidean space. Let K be a pointed, closed convex cone with nonempty interior. Then so is K^* .

PROPOSITION 5.19. Let \mathbb{E} be a Euclidean space. Let K_1 and K_2 be nonempty convex cones. Then

1. $(K_1 + K_2)^* = K_1^* \cap K_2^*$.
2. $(\text{cl}(K_1) \cap \text{cl}(K_2))^* = \text{cl}(K_1^* + K_2^*)$.
3. If K_1 and K_2 are closed and $\text{ri}(K_1) \cap \text{ri}(K_2) \neq \emptyset$, then $(K_1 \cap K_2)^* = K_1^* + K_2^*$.

Proof of (1). Forward Direction: Let x be an arbitrary element of $(K_1 + K_2)^*$. I will show that $x \in K_1^* \cap K_2^*$. Since $x \in (K_1 + K_2)^*$, $\forall k \in K_1 + K_2$, we have $\langle x, k \rangle \geq 0$. Let k_1 be an arbitrary element of K_1 . Let k_2 be an arbitrary element of K_2 . Then

$$\begin{aligned} \langle x, k_1 \rangle &= \left\langle x, \lim_{n \rightarrow \infty} \left(k_1 + \frac{1}{n} k_2 \right) \right\rangle \\ &= \lim_{n \rightarrow \infty} \left\langle x, k_1 + \frac{1}{n} k_2 \right\rangle, \text{ since } \langle x, \cdot \rangle \text{ is continuous} \\ &\geq \lim_{n \rightarrow \infty} 0, \text{ since } k_1 + \frac{1}{n} k_2 \in K_1 + K_2 \\ &= 0. \end{aligned}$$

That is, $\langle x, k_1 \rangle \geq 0$. A similar argument can show that $\langle x, k_2 \rangle \geq 0$. So $x \in K_1^*$ and $x \in K_2^*$. So $x \in K_1^* \cap K_2^*$.

Backward Direction: Let x be an arbitrary element of $K_1^* \cap K_2^*$. I will show that $x \in (K_1 + K_2)^*$. Let k be an arbitrary element of $K_1 + K_2$. Then k can be written as $k = k_1 + k_2$ where $k_1 \in K_1$ and $k_2 \in K_2$. Since $x \in K_1^* \cap K_2^*$, $x \in K_1^*$. Since $x \in K_1^*$ and

$k_1 \in K_1$, we get $\langle x, k_1 \rangle \geq 0$. A similar argument can show that $\langle x, k_2 \rangle \geq 0$. So

$$\langle x, k \rangle = \langle x, k_1 + k_2 \rangle = \langle x, k_1 \rangle + \langle x, k_2 \rangle \geq 0 + 0 = 0.$$

That is, $\langle x, k \rangle \geq 0$. So $x \in (K_1 + K_2)^*$. \square

5.7 Polar Cone

DEFINITION 5.20 (Polar Cone). Let \mathcal{H} be a Hilbert space over \mathbb{R} . Let S be a subset of \mathcal{H} . We define the **polar cone** of S , denoted by S° , to be a subset of \mathcal{H} given by

$$S^\circ := \{x \in \mathcal{H} : \forall y \in S, \langle x, y \rangle \leq 0\}.$$

PROPOSITION 5.21 (Bauschke-Combettes, 2017 Book). The polar cone of any set is a closed convex cone.

PROPOSITION 5.22 (Bauschke-Combettes, 2017 Book). Let \mathcal{H} be a Hilbert space over \mathbb{R} . Let $A, B \subseteq \mathcal{H}$. Then $A \subseteq B \implies B^\circ \subseteq A^\circ$.

PROPOSITION 5.23. Let \mathcal{H} be a Hilbert space over \mathbb{R} . Let C be a subset of \mathcal{H} . Then $C^\circ = -C^*$.

PROPOSITION 5.24. If S is a linear subspace of some Euclidean space \mathbb{E} , then $S^\circ = S^\perp$.

5.8 Extreme Rays

DEFINITION 5.25 (Rays). Let \mathcal{V} be a vector space. Let R be a subset of \mathcal{V} . We say that R is a **ray** if and only if R can be expressed as

$$R = \{\alpha v : \alpha \in \mathbb{R}_+\}$$

for some $v \in \mathbb{E} \setminus \{0\}$.

DEFINITION 5.26 (Extreme Rays). Let \mathcal{V} be a vector space. Let K be a convex cone in \mathcal{V} . Let R be a ray in K . We say that R is an **extreme ray** in K if and only if for any pair of rays R_1 and R_2 in K such that $R_1 + R_2 \supseteq R$, we have either $R_1 = R$ or $R_2 = R$ (or both).

Chapter 6

Tangent Cones and Normal Cones

6.1 Definitions

DEFINITION 6.1 (Tangent Cone - 1). Let $C \subseteq \mathbb{R}^n$ be nonempty and convex. Let $x \in \mathbb{R}^n$. We define the **tangent cone** to C at point x , denoted by $T_C(x)$, to be a subset of \mathbb{R}^n given by

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

DEFINITION 6.2 (Tangent Cone - 2). Let $C \subseteq \mathbb{R}^n$ be nonempty and convex. Let $x, d \in \mathbb{R}^n$. We say that d is **tangent** to C at point x if and only if there is a sequence $\{z_i\}_{i \in \mathbb{Z}_{++}} \subseteq C$ and a sequence $\{t_i\}_{i \in \mathbb{Z}_{++}} \subseteq \mathbb{R}$ such that $\lim_{i \in \mathbb{Z}_{++}} z_i = x$, $\lim_{i \in \mathbb{Z}_{++}} t_i = 0$, and $\lim_{i \in \mathbb{Z}_{++}} \frac{z_i - x}{t_i} = d$. We define the **tangent cone** to C at point x , denoted by $T_C(x)$, to be the set of all all tangent vectors to C at point x .

PROPOSITION 6.3. Are the two definitions equivalent?

DEFINITION 6.4 (Normal Cones). Let $C \subseteq \mathbb{R}^n$ be nonempty and convex. Let $x \in \mathbb{R}^n$. We define the **normal cone** to C at point x , denoted by $N_C(x)$, to be a subset of \mathbb{R}^n given by

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

6.2 Basic Properties

PROPOSITION 6.5. Let C be a closed convex set in \mathbb{E} . Let $x \in \mathbb{R}^n$. 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, $T_C(x) = \overline{\text{cone}}(C - x)$. So $T_C(x)$ is a closed.

Tangent Cone is Convex:

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

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

Tangent Cone is a Cone

By definition, $T_C(x) = \overline{\text{cone}}(C - x)$. 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 $v \in N_C(x)$, $\langle v, y \rangle \leq 0$.

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

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 6.6. 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

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

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 $C - x \subseteq \overline{\text{cone}}(C - x) = T_C(x)$ 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 6.7. Let C be a closed convex set in \mathbb{E} such that $\text{int}(C) \neq \emptyset$. Let $x \in \mathbb{R}^n$. 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 $\overline{\text{cone}}(C - x) = \mathbb{E}$.

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

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

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

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

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,

$$2p \in T_C(x). \tag{2}$$

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

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

□

6.3 Arithmetic Properties

PROPOSITION 6.8. Let C and D be convex subsets of \mathbb{E} . Let $x \in \mathbb{R}^n$. 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 6.9. 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

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

6.4 Other Properties

PROPOSITION 6.10. 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.

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

$$\Longleftrightarrow (u, -1) \in N_{\text{epi}(f)}(x, f(x)).$$

□

Chapter 7

Extreme Points and Faces

Contents

4.1	Definitions	25
4.2	Properties	25

7.1 Extreme Points

DEFINITION 7.1 (Extreme Points - 1). Let \mathcal{V} be a vector space. Let C be a nonempty convex subset of \mathcal{V} . Let z be some point in C . We say that z is an **extreme point** of C if and only if it does not lie between any two distinct points in C . i.e.,

$$\forall x, y \in C, \forall t \in (0, 1), \quad tx + (1 - t)y = z \implies x = y = z.$$

DEFINITION 7.2 (Extreme Points - 2). Let \mathcal{V} be a vector space. Let C be a nonempty convex subset of \mathcal{V} . Let x be some point in C . We say that x is an **extreme point** of C if and only if $C \setminus \{x\}$ is still convex.

PROPOSITION 7.3. The two definitions of extreme point are equivalent.

Proof. Forward Direction: Assume that x does not lie between any two distinct points in C . I will show 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.

Backward Direction: Assume that $C \setminus \{x\}$ is convex. I will show 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 any two distinct points in C . \square

PROPOSITION 7.4. If C is nonempty, convex, and compact, then $\text{Ext}(C) \neq \emptyset$.

PROPOSITION 7.5. Let \mathcal{V} be a locally convex space. Let K be a nonempty, compact, and convex set in \mathcal{V} . Then $\text{Ext}(K) \neq \emptyset$.

7.2 Faces

DEFINITION 7.6 (Faces - 1). Let \mathcal{V} be a vector space. Let C be a nonempty convex subset of \mathcal{V} . Let $F \subseteq \mathcal{V}$. We say that F is a **face** of C , denoted by $F \trianglelefteq C$, if and only if F is a nonempty convex subset of C such that

$$\forall x, y \in C, \forall t \in (0, 1), \quad tx + (1 - t)y \in F \implies x, y \in F.$$

DEFINITION 7.7 (Faces - 2). Let \mathcal{V} be a vector space. Let C be a nonempty convex subset of \mathcal{V} . Let $F \subseteq \mathcal{V}$. We say that F is a **face** of C , denoted by $F \trianglelefteq C$, if and only if F is a nonempty convex subset of C such that

$$\forall n \in \mathbb{N}, \forall x \in C^n, \forall t \in (0, 1)^n : \sum_{i=1}^n t_i = 1, \quad \sum_{i=1}^n t_i x_i \in F \implies x \in F^n.$$

PROPOSITION 7.8. The two definitions of faces above are equivalent.

Proof. Forward Direction: Suppose that $\forall x, y \in C, \forall t \in (0, 1)$ such that $tx + (1 - t)y \in F$, we have $x, y \in F$. Let $n \in \mathbb{N}$, $x \in C^n$, $t \in (0, 1)^n$ be arbitrary such that $\sum_{i=1}^n t_i = 1$. Define a point z by $z := \sum_{i=1}^n t_i x_i$. Suppose that $z \in F$. I will show that $x \in F^n$. Note that $\forall i \in \{1, \dots, n\}$, we have

$$z = \sum_{j=1}^n t_j x_j = t_i x_i + (1 - t_i) \sum_{j \neq i} \frac{t_j}{1 - t_i} x_j.$$

Consider the point $z_i := \sum_{j \neq i} \frac{t_j}{1 - t_i} x_j$. Note that $\forall j \neq i, \frac{t_j}{1 - t_i} \in (0, 1)$ and that $\sum_{j \neq i} \frac{t_j}{1 - t_i} = 1$. So since $\forall j \neq i, x_j \in C$ and C is convex, we get $z_i \in C$. By assumption, we get $x_i, z_i \in F$. In particular, $x_i \in F$. So $x \in F^n$.

Backward Direction: Suppose that $\forall n \in \mathbb{N}, \forall x \in C^n, \forall t \in (0, 1)^n$ such that $\sum_{i=1}^n t_i = 1$ and that $\sum_{i=1}^n t_i x_i \in F$, we have $x \in F^n$. Take $n := 2$, then $\forall x, y \in C, \forall t \in (0, 1)$ such that $tx + (1 - t)y \in F$, we have $x, y \in F$. \square

Faces are generalizations of extreme points.

PROPOSITION 7.9 (Transitivity). Let \mathcal{V} be a vector space. Let A , B , and C be nonempty convex subsets of \mathcal{V} . Suppose that $A \trianglelefteq B$ and $B \trianglelefteq C$. Then $A \trianglelefteq C$.

Proof. Let x and y be two arbitrary elements of C . Let t be an arbitrary element of $(0, 1)$. Define a point z by $z := tx + (1 - t)y$. Suppose that $z \in A$. I will show that $x, y \in A$. Note that since $A \trianglelefteq B$, we have $A \subseteq B$. So $z \in A \subseteq B$. Since $x, y \in C$, $t \in (0, 1)$, $z \in B$, and $B \trianglelefteq C$, we get $x, y \in B$. Since $x, y \in B$, $t \in (0, 1)$, $z \in A$, and $A \trianglelefteq B$, we get $x, y \in A$. So $A \trianglelefteq C$. \square

PROPOSITION 7.10 (Intersection). Let \mathcal{V} be a vector space. Let C be a nonempty convex subset of \mathcal{V} . Let $A, B \trianglelefteq C$. Then $(A \cap B) \trianglelefteq C$.

Proof. Let x and y be two arbitrary elements of C . Let t be an arbitrary element of $(0, 1)$. Define a point z by $z := tx + (1 - t)y$. Suppose that $z \in A \cap B$. I will show that $x, y \in A \cap B$. Since $A \trianglelefteq C$, $x, y \in C$, $t \in (0, 1)$, and $z \in A \cap B \subseteq A$, we get $x, y \in A$. Similarly, we get $x, y \in B$. So $x, y \in A \cap B$. So $(A \cap B) \trianglelefteq C$. \square

7.2.1 Exposed Faces

DEFINITION 7.11 (Exposed Face of a Convex Cone). Let \mathcal{V} be an inner product space with inner product $\langle \cdot, \cdot \rangle$. Let C be a nonempty convex conic subset of \mathcal{V} . Let $F \trianglelefteq C$. We say that F is **exposed** if and only if $\exists a \in \mathcal{V} \setminus \{0\}$ such that

$$F = \{x \in C : \langle a, x \rangle = 0\} \text{ and } C \subseteq \{x \in \mathcal{V} : \langle a, x \rangle \leq 0\}.$$

PROPOSITION 7.12. Let \mathcal{V} be a vector space. Let C be nonempty convex subset of \mathcal{V} . Then every face of C is contained in some exposed face of C .

7.2.2 Relation Between Extreme Points and Faces

DEFINITION 7.13 (Extreme Points - 3). Let \mathcal{V} be a vector space. Let C be a nonempty convex subset of \mathcal{V} . Let x be some point in C . We say that x is an **extreme**

point of C if and only if $\{x\}$ is a face of C .

PROPOSITION 7.14. This definition of extreme points is equivalent to the previous two.

PROPOSITION 7.15. If F is a face of C , then $\text{Ext}(F) \subseteq \text{Ext}(C)$.

7.3 Results for Compact Convex Sets

THEOREM 7.16. Let $K \subseteq \mathbb{R}^n$ be a nonempty compact convex set. Then K has an extreme point.

Proof. Consider the following optimization problem

$$(P) \quad \begin{array}{ll} \max_x & \|x\|_2^2 \\ \text{subject to:} & x \in K. \end{array}$$

Since K is compact and the function $x \mapsto \|x\|_2^2$ is continuous, the maximum is attained, say, at \bar{x} . Assume for the sake of contradiction that \bar{x} is not an extreme point of K . Then $\exists \varepsilon \in \mathbb{R}^n \setminus \{0_n\}$ such that $\bar{x} \pm \varepsilon \in K$. By the optimality of \bar{x} , we must have $\|\bar{x} + \varepsilon\|_2^2 \leq \|\bar{x}\|_2^2$ and $\|\bar{x} - \varepsilon\|_2^2 \leq \|\bar{x}\|_2^2$. So

$$2\|\bar{x}\|_2^2 \geq \|\bar{x} + \varepsilon\|_2^2 + \|\bar{x} - \varepsilon\|_2^2 \quad (7.1a)$$

$$= \|\bar{x}\|_2^2 + \|\varepsilon\|_2^2 + 2\langle \bar{x}, \varepsilon \rangle + \|\bar{x}\|_2^2 + \|\varepsilon\|_2^2 - 2\langle \bar{x}, \varepsilon \rangle \quad (7.1b)$$

$$= 2\|\bar{x}\|_2^2 + 2\|\varepsilon\|_2^2 + 0. \quad (7.1c)$$

So $\|\varepsilon\|_2^2 \leq 0$. So $\varepsilon = 0_n$. This contradicts to the choice of ε that $\varepsilon \neq 0_n$. \square

LEMMA 7.17. Let \mathcal{V} be a locally convex space. Let K be a nonempty compact convex subset of \mathcal{V} . Let $\rho \in \mathcal{V}^*$. Define $r := \sup\{\Re\rho(x) : x \in K\}$. Define $F := \{x \in K : \Re\rho(x) = r\}$. Then F is a nonempty compact face of K .

Proof. Nonempty: Since $\Re\rho$ is continuous and K is compact, $\{\Re\rho(x) : x \in K\}$ is a compact set in \mathbb{R} . So $r = \sup\{\Re\rho(x) : x \in K\}$ is attained. So $F \neq \emptyset$.

Compact: Notice $F = (\Re\rho)^{-1}(\{r\})$. Since $\Re\rho$ is continuous and $\{r\} \subseteq \mathbb{R}$ is closed, F is closed. Since F is a closed subset of K and K is compact, F is compact.

Convex: Let x and y be arbitrary elements of F . Let $t \in (0, 1)$. Since $x, y \in F$, we have $\Re\rho(x) = \Re\rho(y) = r$. So

$$\Re\rho(tx + (1-t)y) = t\Re\rho(x) + (1-t)\Re\rho(y) = tr + (1-t)r = r.$$

So $tx + (1-t)y \in F$. So F is convex.

Face: Let x and y be arbitrary elements of K . Let $t \in (0, 1)$. Suppose that $tx + (1-t)y \in F$. Since $x, y \in K$, we have $\Re\rho(x) \leq r$ and $\Re\rho(y) \leq r$. Since $tx + (1-t)y \in F$, we have

$$t\Re\rho(x) + (1-t)\Re\rho(y) = \Re\rho(tx + (1-t)y) = r.$$

So we must have $\Re\rho(x) = \Re\rho(y) = r$. So $x, y \in F$. So F is a face of K .

□

THEOREM 7.18 (Krein-Milman Theorem). A compact convex set in a (Hausdorff) locally convex space is the closed convex hull of its extreme points.

Proof. Let \mathcal{V} be a locally convex space. Let K be a nonempty, compact, and convex set in \mathcal{V} .

Forward Direction: Show that $K \subseteq \overline{\text{conv}}(\text{Ext}(K))$. Let m be an arbitrary element of K . Assume for the sake of contradiction that $m \notin \overline{\text{conv}}(\text{Ext}(K))$. By the Hahn-Banach Theorem, there is some $\tau \in \mathcal{V}^*$ and $\alpha, \beta \in \mathbb{R}$ such that $\alpha > \beta$ and

$$\forall b \in \overline{\text{conv}}(\text{Ext}(K)), \quad \Re\tau(m) \geq \alpha > \beta \geq \Re\tau(b).$$

Define $s := \sup\{\Re\tau(w) : w \in K\}$. Define $L := \{z \in K : \Re\tau(z) = s\}$. Then L is a nonempty compact face of K . So $\text{Ext}(L) \neq \emptyset$. Let e be an element of $\text{Ext}(L)$. Then $e \in \text{Ext}(L) \subseteq L$. So $\Re\tau(e) = s$. So

$$\forall b \in \overline{\text{conv}}(\text{Ext}(K)), \quad \Re\tau(e) = s \geq \Re\tau(m) \geq \alpha > \beta \geq \Re\tau(b).$$

That is, $\Re\tau(e) > \Re\tau(b)$. Since L is a face of K , $\text{Ext}(L) \subseteq \text{Ext}(K)$. Notice $e \in \text{Ext}(L) \subseteq \text{Ext}(K) \subseteq \overline{\text{conv}}(\text{Ext}(K))$. So in particular, $\Re\tau(e) > \Re\tau(e)$. This is a contradiction. So $m \in \overline{\text{conv}}(\text{Ext}(K))$. So $K \subseteq \overline{\text{conv}}(\text{Ext}(K))$.

Backward Direction: Show that $\overline{\text{conv}}(\text{Ext}(K)) \subseteq K$. Note that $\text{Ext}(K) \subseteq K$. Since K is closed and convex and $\text{Ext}(K) \subseteq K$, we get $\overline{\text{conv}}(\text{Ext}(K)) \subseteq K$.

□

PROPOSITION 7.19. Let \mathcal{V} be a vector space. Let K be a nonempty compact convex subset of \mathcal{V} . Let $\mathcal{F}(K)$ denote the set of faces of K , partially ordered by inclusion. Then the minimal proper faces in $\mathcal{F}(K)$ are the extreme points of K .

PROPOSITION 7.20. Let \mathcal{V} be a vector space. Let K be a nonempty compact convex subset of \mathcal{V} . Let $\mathcal{F}(K)$ denote the set of faces of K , partially ordered by inclusion. Then the maximal proper faces in $\mathcal{F}(K)$ are exposed.

Chapter 8

Projection Operators

8.1 Definitions

DEFINITION 8.1 (Projection). Let \mathcal{H} be a Hilbert space over \mathbb{R} with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Let $S \subseteq \mathcal{H}$ be nonempty and let $x \in \mathcal{H}$. We define the **projection** of x onto S , denoted by $\text{proj}_S(x)$, to be a point in \mathcal{H} 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 8.2 (Existence). If S is nonempty 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

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

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 $p = \text{proj}_S(x)$. \square

PROPOSITION 8.3 (Uniqueness). If S is nonempty, 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. \square

8.2 Properties (bug)

PROPOSITION 8.4 (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 $\text{proj}_C = \text{proj}_C \text{proj}_C$. \square

PROPOSITION 8.5. Let $C \subseteq \mathbb{R}^n$ be nonempty closed and convex. Then $\text{Fix}(\text{proj}_C) = C$.

Proof. (\Rightarrow) Let $x \in \text{Fix}(\text{proj}_C)$ be arbitrary. Then $x = \text{proj}_C(x)$. By definition of projection, we have $\text{proj}_C(x) \in C$. So $x = \text{proj}_C(x) \in C$.

(\Leftarrow) Let $x \in C$ be arbitrary. Then $x = \text{proj}_C(x)$. So x is a fixed point of proj_C . \square

PROPOSITION 8.6 (Linearity). Let C be a nonempty closed convex set in \mathcal{H} . Then $\text{proj}_C \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ if and only if C is a linear subspace of \mathcal{H} .

Proof. (\Rightarrow) Suppose that $\text{proj}_C \in \mathcal{L}(\mathcal{H}, \mathcal{H})$. I will show that C is a linear subspace. Let $x, y \in C$ be arbitrary. Then we have

$$\text{proj}_C(x + y) = \text{proj}_C(x) + \text{proj}_C(y) = x + y, \text{ since } x, y \in C.$$

Since $\text{proj}_C(x + y) \in C$, we get $x + y \in C$. So C is closed under addition. Let $x \in C$ and $k \in \mathbb{R}$ be arbitrary. Then we have

$$\text{proj}_C(kx) = k \text{proj}_C(x) = kx, \text{ since } x \in C.$$

Since $\text{proj}_C(kx) \in C$, we get $kx \in C$. So C is closed under scalar multiplication. So C is a linear subspace of \mathcal{H} .

(\Leftarrow) Suppose that C is a linear subspace of \mathcal{H} . I will show that $\text{proj}_C \in \mathcal{L}(\mathcal{H}, \mathcal{H})$. Let $x, y \in \mathcal{H}$ be arbitrary. Since C is linear, $\text{proj}_C(x) \in C$, $\text{proj}_C(y) \in C$, we get $\text{proj}_C(x) + \text{proj}_C(y) \in C$. Let $z \in C$ be arbitrary. Consider the distances $\|(x + y) - z\|$ and $\|(x + y) - (\text{proj}_C(x) + \text{proj}_C(y))\|$. **not finished** \square

PROPOSITION 8.7 (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 8.8. Let $C \subseteq \mathbb{R}^n$ be nonempty closed and convex. Then $\text{proj}_C \in \mathcal{Lip}(1)$. i.e.,

$$\forall x, y \in \mathbb{R}^n, \quad \|\text{proj}_C(x) - \text{proj}_C(y)\| \leq \|x - y\|.$$

Proof. Let $x, y \in \mathbb{R}^n$ be arbitrary. For brevity, let p_x denote $\text{proj}_C(x)$ and p_y denote $\text{proj}_C(y)$. If $\|p_x - p_y\| = 0$, then trivially $\|p_x - p_y\| \leq \|x - y\|$ and we are done. So we may assume that $\|p_x - p_y\| \neq 0$. Then by the Projection Theorem, we have

$$\langle p_y - p_x, x - p_x \rangle \leq 0 \text{ and} \tag{8.1a}$$

$$\langle p_x - p_y, y - p_y \rangle \leq 0. \tag{8.1b}$$

Now

$$\|p_x - p_y\|^2 = \langle p_x - p_y, p_x - p_y \rangle \quad (8.2a)$$

$$= \underbrace{\langle p_y - p_x, x - p_x \rangle}_{\leq 0} + \underbrace{\langle p_x - p_y, y - p_y \rangle}_{\leq 0} + \langle p_x - p_y, x - y \rangle \quad (8.2b)$$

$$\leq \langle p_x - p_y, x - y \rangle \quad (8.2c)$$

$$\leq \|p_x - p_y\| \|x - y\|, \text{ by the Cauchy-Schwarz inequality.} \quad (8.2d)$$

That is, $\|p_x - p_y\|^2 \leq \|p_x - p_y\| \|x - y\|$. Since we have assumed that $\|p_x - p_y\| \neq 0$, we get $\|p_x - p_y\| \leq \|x - y\|$. So $\text{proj}_C \in \mathcal{Lip}(1)$, as desired. \square

PROPOSITION 8.9 (Firmly Non-expansive). Let $C \subseteq \mathbb{R}^n$ be nonempty closed and convex. Then proj_C is firmly non-expansive.

Proof. This is to prove.

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

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

\square

8.3 Examples

EXAMPLE 8.10 (Projection onto Half Spaces, Bauschke-Combettes, 2017 Book). Let \mathcal{H} be a Hilbert space over \mathbb{R} with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$. Let $C \subseteq \mathcal{H}$ be a half-space given by

$$C := \{x \in \mathcal{H} : \langle x, u \rangle \leq \eta\}$$

where $u \in \mathcal{H}$ and $\eta \in \mathbb{R}$ are constants. Then if $u = 0$ and $\eta \geq 0$, we have $C = \mathcal{H}$ and

hence $\text{proj}_C = \text{Id}$; if $u = 0$ and $\eta < 0$, we have $C = \emptyset$; if $u \neq 0$, we have

$$\text{proj}_C(x) = \begin{cases} x, & \text{if } \langle x, u \rangle \leq \eta \\ x + \frac{\eta - \langle x, u \rangle}{\|u\|^2} u, & \text{if } \langle x, u \rangle > \eta. \end{cases}$$

8.4 Characterizations

THEOREM 8.11 (Projection Theorem). Let $C \subseteq \mathbb{R}^n$ be nonempty closed and convex. Let $x, p \in \mathbb{R}^n$. Then $p = \text{proj}_C(x)$ if and only if $p \in C$ and

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

Proof. Let y be an arbitrary point in C . Define for each $\alpha \in [0, 1]$ a point $y_\alpha \in \mathbb{R}^n$ by $y_\alpha := \alpha y + (1 - \alpha)p$. Since $y, p \in C$ and C is convex, we always have $y_\alpha \in C$. Now

$$p = \text{proj}_C(x) \tag{8.3a}$$

$$\iff p \in C \text{ and } \forall y \in C, \|x - p\|^2 \leq \|x - y\|^2 \tag{8.3b}$$

$$\iff p \in C \text{ and } \forall y \in C, \forall \alpha \in [0, 1], \|x - p\|^2 \leq \|x - y_\alpha\|^2 \tag{8.3c}$$

$$\iff p \in C \text{ and } \forall y \in C, \forall \alpha \in [0, 1], \|x - p\|^2 \leq \|x - p - \alpha(y - p)\|^2 \tag{8.3d}$$

$$\iff p \in C \text{ and } \forall y \in C, \forall \alpha \in [0, 1], \alpha \langle x - p, y - p \rangle \leq \frac{1}{2} \alpha^2 \|y - p\|^2 \tag{8.3e}$$

$$\iff p \in C \text{ and } \forall y \in C, \langle x - p, y - p \rangle \leq 0. \tag{8.3f}$$

□

THEOREM 8.12 (Bauschke-Combettes, 2017 Book). Let \mathcal{H} be a Hilbert space over \mathbb{R} with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$. Let $K \subseteq \mathcal{H}$ be a nonempty closed convex cone. Let $x, p \in \mathcal{H}$. Then $p = \text{proj}_K(x)$ if and only if $p \in K$, $\langle x - p, p \rangle = 0$, and $x - p \in K^\circ$.

Chapter 9

Separation

9.1 Definitions

DEFINITION 9.1 (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

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

DEFINITION 9.2 (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 9.3 (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

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

9.2 Main Results

PROPOSITION 9.4. 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 9.5. 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. \square

THEOREM 9.6. 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 $\left\{ C_1 \text{ is non-empty closed and convex } D_n \text{ is non-empty compact and convex} \right\}$ 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. \square

PROPOSITION 9.7. 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

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

Chapter 10

Convex Functions

Contents

3.1	Definitions (bug)	8
3.2	Arithmetic Properties of Convex Sets	10
3.3	The Convex Hull Operator	11
3.4	The Closed Convex Hull Operator	14
3.5	Stability of Convexity	15
3.6	Topological Properties of Convex Sets	20
3.7	Examples of Convex Sets	23
3.8	The Carathéodory Theorem	24

10.1 Preliminaries

DEFINITION 10.1 (Epigraph). Let f be a function from \mathbb{R}^n to \mathbb{R}^* . We define the **epigraph** of f , denoted by $\text{epi}(f)$, to be a subset of $\mathbb{R}^n \oplus \mathbb{R}$ given by

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

DEFINITION 10.2 (Domain). Let f be a function from \mathbb{R}^n to \mathbb{R}^* . We define the **domain** of f , denoted by $\text{dom}(f)$, to be a subset of \mathbb{R}^n given by

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

DEFINITION 10.3 (Proper). Let f be a function from \mathbb{R}^n to \mathbb{R}^* . We say that f is **proper** if and only if all of the following conditions hold

- $\exists x \in \mathbb{R}^n, f(x) \neq +\infty$; and
- $\forall x \in \mathbb{R}^n, f(x) \neq -\infty$

10.2 The Indicator Function

DEFINITION 10.4 (The Indicator Function). Let S be a subset of \mathbb{R}^n . We define the **indicator function** of S , denoted by δ_S , to be a function from \mathbb{R}^n to \mathbb{R}^* given by

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

PROPOSITION 10.5. Let S be a subset of \mathbb{R}^n . 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 $\delta_S(p) = 0$, $\exists x_0 \in \mathbb{R}^n$ such that $\delta_S(x_0) \neq +\infty$.

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

Since $\exists x_0 \in \mathbb{R}^n$ such that $\delta_S(x_0) \neq +\infty$ and $\forall x \in \mathbb{R}^n$, $\delta_S(x) \neq -\infty$, 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{R}^n .

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

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

Since $\forall x \in \mathbb{R}^n$, $\delta_S(x) = +\infty$, 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 $\alpha \rightarrow \alpha_\infty$, $\alpha_\infty \geq 0$.

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 10.6 (Convex Function). Let f be a function from \mathbb{R}^n to \mathbb{R}^* . We say that f is **convex** if and only 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 10.7 (Convex Function). Let f be a function from \mathbb{R}^n to \mathbb{R}^* . We say that f is **convex** if and only if $\text{epi}(f) \subseteq \mathbb{R}^{n+1}$ is convex.

PROPOSITION 10.8. 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 $\text{dom}(f), \text{epi}(f) \neq \emptyset$.

(\Rightarrow) Suppose that $\forall x, y \in \text{dom}(f), \forall \lambda \in [0, 1]$, we have $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$. I will show that $\text{epi}(f) \subseteq \mathbb{R}^{n+1}$ 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 $f(z) \leq \gamma$, $(z, \gamma) \in \text{epi}(f)$.

(\Leftarrow) Suppose that $\text{epi}(f) \subseteq \mathbb{R}^{n+1}$ is convex. I will show 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, $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$. \square

10.4 Basic Properties

PROPOSITION 10.9. Let \mathcal{V} be a vector space over \mathbb{R} . Let $f : \mathcal{V} \rightarrow \mathbb{R}$ be a convex function. Then $\forall x, y \in \mathbb{R}^n, \forall t \notin (0, 1)$, we have $f(tx + (1 - t)y) \geq tf(x) + (1 - t)f(y)$.

Proof. Assume for the sake of contradiction that $\exists x, y \in \mathbb{R}^n, \exists t \notin (0, 1)$ such that $f(tx + (1 - t)y) < tf(x) + (1 - t)f(y)$. Assume without loss of generality that $t > 1$. The case where $t < 0$ can be proved similarly. Define $z := tx + (1 - t)y$. Then $x = \frac{1}{t}z + \frac{t-1}{t}y$. Notice that x is now a convex combination of z and y . By convexity of f , we have

$$\begin{aligned} f(x) &= f\left(\frac{1}{t}z + \frac{t-1}{t}y\right) \leq \frac{1}{t}f(z) + \frac{t-1}{t}f(y) < \frac{1}{t}\left[tf(x) + (1-t)f(y)\right] + \frac{t-1}{t}f(y) \\ &= f(x) + \frac{1-t}{t}f(y) + \frac{t-1}{t}f(y) = f(x). \end{aligned}$$

That is, $f(x) < f(x)$, which is itself a contradiction. So $\forall x, y \in \mathbb{R}^n, \forall t > 1$, we have $f(tx + (1 - t)y) \geq tf(x) + (1 - t)f(y)$. \square

PROPOSITION 10.10 (Necessary Condition). The domain of a convex function is convex.

Proof. Follows from the fact that convexity of sets is stable under affine transformations. Define an affine mapping $A : \mathbb{R}^n \oplus \mathbb{R} \rightarrow \mathbb{R}^n$ by $A((x, \alpha)) := x$. Then $\text{dom}(f) = A(\text{epi}(f))$ and hence is convex. \square

PROPOSITION 10.11. The level sets of a convex function are convex.

PROPOSITION 10.12 (Restriction to a Line). A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if and only if $\forall x \in \text{dom}(f), \forall v \in \mathbb{R}^n$, 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 10.13. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be differentiable and convex. Then $f \in \mathcal{C}^1$. i.e., f is continuously differentiable.

PROPOSITION 10.14. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be differentiable. Then f is convex if and only if f' is non-decreasing.

PROPOSITION 10.15. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper and convex. Let $x \in \text{dom}(f)$. Then the following conditions are equivalent:

- f is differentiable at x ;
- the subgradient $\partial f(x)$ at x is singleton.

Moreover, when these equivalent conditions are met, we have $\partial f(x) = \{\nabla f(x)\}$.

PROPOSITION 10.16 (First-Order Condition). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper. Suppose that $\text{dom}(f)$ is convex and open and that f is differentiable on $\text{dom}(f)$. Then the following conditions are equivalent:

- f is convex;
- $\forall x, y \in \text{dom}(f)$, $f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle$. i.e., the first-order approximation of f is a global under-estimator.

Proof. (\Rightarrow) Suppose that f is convex. I will show that $\forall x, y \in \text{dom}(f)$, we have $f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle$. Let $x, y \in \text{dom}(f)$ be arbitrary. Since f is convex and differentiable at point x , $\partial f(x) = \{\nabla f(x)\}$. So $\nabla f(x)$ satisfies the subgradient inequality. That is, $f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle$.

(\Leftarrow) Suppose that $\forall x, y \in \text{dom}(f)$, we have $f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle$. I will show that f is convex. Let $x, y \in \text{dom}(f)$ and $t \in (0, 1)$ be arbitrary. Define a point $z \in \mathbb{R}^n$ by $z := tx + (1 - t)y$. Since $\text{dom}(f)$ is convex, we have $z \in \text{dom}(f)$. Applying the assumption, we get $f(x) - f(z) \geq \langle \nabla f(z), x - z \rangle$ and $f(y) - f(z) \geq \langle \nabla f(z), y - z \rangle$. Now

$$f(z) = tf(z) + (1 - t)f(z) \tag{10.1a}$$

$$\leq t \left[f(x) - \langle \nabla f(z), x - z \rangle \right] + (1-t) \left[f(y) - \langle \nabla f(z), y - z \rangle \right] \quad (10.1b)$$

$$= tf(x) + (1-t)f(y) - \langle \nabla f(z), tx + (1-t)y - z \rangle \quad (10.1c)$$

$$= tf(x) + (1-t)f(y) - \langle \nabla f(z), \mathbb{0}_n \rangle, \text{ by definition of } z \quad (10.1d)$$

$$= tf(x) + (1-t)f(y). \quad (10.1e)$$

That is, $f(z) \leq tf(x) + (1-t)f(y)$. This holds for any $x, y \in \text{dom}(f)$ and $t \in (0, 1)$. So f is convex. The proof is now complete. \square

PROPOSITION 10.17. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper. Suppose 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,

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle. \quad (1)$$

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 10.18 (Second-Order Condition). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper. Suppose that $\text{dom}(f)$ is convex and that $f \in \mathcal{C}^2$ on $\text{dom}(f)$. Then the following conditions are equivalent:

- f is convex;

- $\forall x \in \text{dom}(f), \nabla^2 f(x) \in \mathbb{S}_+^n.$

10.6 Convexity and Lipschitz-ness

THEOREM 10.19. Let f be a differentiable convex function from \mathbb{R}^n to \mathbb{R} . Then the following statements are equivalent.

1. ∇f is Lipschitz with constant L .
2. $\forall x, y \in \mathbb{R}^n$, we have

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

3. $\forall x, y \in \mathbb{R}^n$, 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. $\forall x, y \in \mathbb{R}^n$, 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{R}^n .

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

That is,

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

□

THEOREM 10.20. Let f be a twice continuously differentiable function from \mathbb{R}^n to \mathbb{R} . Let L be some non-negative number. Then the following statements are equivalent.

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

10.7 Stability of Convexity

PROPOSITION 10.21 (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

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

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,

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

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

PROPOSITION 10.22 (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

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

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 10.23 (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{R}^n . 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,

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

So g is convex. □

PROPOSITION 10.24. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be convex. Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be convex and non-decreasing. Then $g \circ f : \mathbb{R} \rightarrow \mathbb{R}$ is convex.

PROPOSITION 10.25 (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 $\bigcap_{i \in I} \text{epi}(f_i)$ is convex. That is, $\text{epi}(F)$ is convex. Since $\text{epi}(F)$ is convex, F is convex. \square

PROPOSITION 10.26 (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 10.27. Affine functions are convex.

EXAMPLE 10.28. 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 10.29. 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{R}^n . Let x and y be two points in \mathbb{R}^n . 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,

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

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

□

EXAMPLE 10.30. The distance function to a convex set is convex.

EXAMPLE 10.31. The perspective of a convex function is convex. i.e., if $f : \mathbb{R}^n \rightarrow \mathbb{R}$

Chapter 11

More Convex Functions

Contents

1.1	Definitions	1
-----	-----------------------	---

11.1 Strictly Convex

DEFINITION 11.1 (Strictly Convex). Let \mathcal{V} be a vector space over \mathbb{R} . Let f be a proper function from \mathcal{V} to \mathbb{R}^* . We say that f is **strictly convex** if and only if $\forall x, y \in \text{dom}(f)$, $\forall \lambda \in (0, 1)$, we have

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y). \quad (11.1)$$

PROPOSITION 11.2. Strictly convex functions are convex.

11.2 Strongly Convex

11.2.1 Definition and Equivalent Conditions

DEFINITION 11.3 (Strongly Convex). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper. Let $\beta \in \mathbb{R}_{++}$. We say that f is **β -strongly convex** if and only if $\forall x, y \in \text{dom}(f)$, $\forall \lambda \in [0, 1]$, we have

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{\beta}{2} \lambda(1 - \lambda) \|x - y\|_2^2. \quad (11.2)$$

We say that f is **strongly convex** if and only if $\exists \beta \in \mathbb{R}_{++}$ such that f is β -strongly convex.

PROPOSITION 11.4. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper and $\beta \in \mathbb{R}_{++}$. Then the following conditions are equivalent:

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

Proof. Let $g := f - \frac{\beta}{2} \|\cdot\|^2$. Then $g(x) = +\infty$ if and only if $f(x) = +\infty$ and hence $\text{dom}(g) = \text{dom}(f)$. Let $x, y \in \text{dom}(f) = \text{dom}(g)$ be arbitrary. Let $\lambda, \mu \in [0, 1]$ be arbitrary such that $\lambda + \mu = 1$. Then

$$\lambda\mu \|x - y\|_2^2 \quad (11.3a)$$

$$= \lambda\mu (\|x\|_2^2 + \|y\|_2^2 - 2\langle x, y \rangle) \quad (11.3b)$$

$$= \lambda\|x\|_2^2 - \lambda^2\|x\|_2^2 + \mu\|y\|_2^2 - \mu^2\|y\|_2^2 - 2\lambda\mu\langle x, y \rangle \quad (11.3c)$$

$$= \lambda\|x\|_2^2 + \mu\|y\|_2^2 - \|\lambda x\|_2^2 - \|\mu y\|_2^2 - 2\langle \lambda x, \mu y \rangle \quad (11.3d)$$

$$= \lambda\|x\|_2^2 + \mu\|y\|_2^2 - \|\lambda x + \mu y\|_2^2 \quad (11.3e)$$

So

$$f \text{ is } \beta\text{-strongly convex} \quad (11.4a)$$

$$\iff f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y) - \frac{\beta}{2} \lambda\mu \|x - y\|_2^2 \quad (11.4b)$$

$$\iff f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y) - \frac{\beta}{2} \left[\lambda\|x\|_2^2 + \mu\|y\|_2^2 - \|\lambda x + \mu y\|_2^2 \right] \quad (11.4c)$$

$$\iff f(\lambda x + \mu y) - \frac{\beta}{2} \|\lambda x + \mu y\|_2^2 \leq \lambda \left[f(x) - \frac{\beta}{2} \|x\|_2^2 \right] + \mu \left[f(y) - \frac{\beta}{2} \|y\|_2^2 \right] \quad (11.4d)$$

$$\iff g(\lambda x + \mu y) \leq \lambda g(x) + \mu g(y) \quad (11.4e)$$

$$\iff g \text{ is convex.} \quad (11.4f)$$

That is, f is β -strongly convex if and only if g is convex, as desired. \square

11.2.2 Properties

PROPOSITION 11.5. Strongly convex functions are strictly convex.

PROPOSITION 11.6. Let $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^*$. Suppose f is β -strongly convex for some positive constant β and g is convex. Then $f + g$ is also β -strongly convex.

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

Proof.

$$\begin{aligned} & f \text{ is } \beta\text{-strongly convex} \\ \implies & f - \frac{\beta}{2} \|\cdot\|^2 \text{ is convex} \\ \implies & f + g - \frac{\beta}{2} \|\cdot\|^2 \text{ is convex} \\ \implies & f + g \text{ is } \beta\text{-strongly convex.} \end{aligned}$$

\square

PROPOSITION 11.7 (Direct Sum). Let $\beta \in \mathbb{R}_{++}$ and $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}^*$ be proper and β -strongly convex. Define $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ by $f := \bigoplus_{i=1}^n f_i$. Then f is also β -strongly convex.

Proof. Let $x, y \in \text{dom}(f)$ and $t \in [0, 1]$ be arbitrary. Since $x, y \in \text{dom}(f)$, we have $\forall i \in \{1, \dots, n\}$, $x_i, y_i \in \text{dom}(f_i)$. So

$$f(tx + (1-t)y) = \sum_{i=1}^n f_i(tx_i + (1-t)y_i) \quad (11.5a)$$

$$\leq \sum_{i=1}^n \left[tf_i(x_i) + (1-t)f_i(y_i) - \frac{\beta}{2}t(1-t)(x_i - y_i)^2 \right] \quad (11.5b)$$

$$= t \sum_{i=1}^n f_i(x_i) + (1-t) \sum_{i=1}^n f_i(y_i) - \frac{\beta}{2}t(1-t) \sum_{i=1}^n (x_i - y_i)^2 \quad (11.5c)$$

$$= tf(x) + (1-t)f(y) - \frac{\beta}{2}t(1-t)\|x-y\|_2^2. \quad (11.5d)$$

So f is also β -strongly convex. \square

11.2.3 Differentiable Strongly Convex Functions

PROPOSITION 11.8 (First-Order Conditions). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper and $\beta \in \mathbb{R}_{++}$. Suppose that $f \in \mathcal{C}^1$. Then the following conditions are equivalent:

- f is β -strongly convex;
- $\forall x, y \in \text{dom}(f)$, we have $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \beta \|y - x\|_2^2$;
- $\forall x, y \in \text{dom}(f)$, we have $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\beta}{2} \|y - x\|_2^2$;

PROPOSITION 11.9 (Second-Order Conditions). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper and $\beta \in \mathbb{R}_{++}$. Suppose that $f \in \mathcal{C}^2$. Then the following conditions are equivalent:

- f is β -strongly convex;
- $\forall x \in \text{dom}(f)$, $\nabla^2 f(x) \succeq \beta I$.

11.3 Uniformly Convex

DEFINITION 11.10 (Uniformly Convex). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper. Let $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ be such that $\varphi(x) \geq 0$ and $\varphi(x) = 0$ if and only if $x = 0$. We say that f is **uniformly convex** with modulus φ if and only if $\forall x, y \in \text{dom}(f)$, $\forall t \in [0, 1]$, we have

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - t(1-t)\varphi(\|x - y\|_2). \quad (11.6)$$

11.4 Quasiconvex

DEFINITION 11.11 (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 11.12 (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 11.13. 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 12.1 (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 12.2 (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 12.3. 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 12.4. 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 12.5. Let S be a subset of \mathbb{E} . Then $\sigma_S = \sigma_{\text{conv}(S)} = \sigma_{\overline{\text{conv}}(S)}$.

Proof.

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

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

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,

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

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

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

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

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

□

12.3 Supporting Hyperplane

THEOREM 12.6 (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

Contents

2.1	Definitions	3
2.2	Basic Properties	4
2.3	Arithmetic Properties	4

13.1 Definition and Examples

DEFINITION 13.1 (Convex Conjugate/Legendre-Fenchel Conjugate). Let f be a function from \mathbb{R}^n to \mathbb{R}^* . We define the **convex conjugate**, or **Legendre-Fenchel conjugate** of f , denoted by f^* , to be a function from \mathbb{R}^n to \mathbb{R}^* given by

$$f^*(x) := \sup_{y \in \mathbb{R}^n} \left\{ \langle y, x \rangle - f(y) \right\} \quad (13.1)$$

EXAMPLE 13.2. Let S be a subset of \mathbb{R}^n . 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 $x \in \mathbb{R}^n$, we have

$$\delta_S^*(x) = \sup_{y \in S} \left\{ \langle x, y \rangle - \underbrace{\delta_S(y)}_{=0} \right\} = \sup_{y \in S} \langle x, y \rangle = \sigma_S(x).$$

So $\delta_S^* = \sigma_S$. □

13.2 Basic Properties

PROPOSITION 13.3. 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. \square

PROPOSITION 13.4. The convex conjugate function is lower semi-continuous.

PROPOSITION 13.5. The convex conjugate transform is a convex operator. i.e.,

$$\forall f, g : \mathbb{R}^n \rightarrow \mathbb{R}^*, \forall t \in [0, 1], \quad (tf + (1-t)g)^* \leq tf^* + (1-t)g^*. \quad (13.2)$$

PROPOSITION 13.6. If $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ is strictly convex, then f^* is differentiable.
(not sure)

13.3 Double Conjugate

PROPOSITION 13.7. Let f be any function from \mathbb{R}^n to \mathbb{R}^* . Then $f^{**} \leq f$.

Proof. Let $x \in \mathbb{R}^n$ be arbitrary.

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

That is, $f^{**}(x) \leq f(x)$. This holds for all $x \in \mathbb{R}^n$. So $f^{**} \leq f$. \square

PROPOSITION 13.8. Let f be a proper function. Then the following conditions are equivalent:

- f is convex and lower semi-continuous;
- $f^{**} = f$.

PROPOSITION 13.9 (Order-Reversing). Let f and g be functions from \mathbb{R}^n to \mathbb{R}^* . Then $f \leq g$ implies $f^* \geq g^*$ and $f^{**} \leq g^{**}$.

Proof. Let $x \in \mathbb{R}^n$ be arbitrary.

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

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

Let $x \in \mathbb{R}^n$ be arbitrary.

$$f^{**}(x)$$

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

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

PROPOSITION 13.10.

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

13.4 Conjugates and Sub-Differentials

THEOREM 13.11 (Fenchel-Young Inequality). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^*$ be proper. Then $\forall x, y \in \mathbb{R}^n$, we have

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

PROPOSITION 13.12. Let f be a proper closed convex function from \mathbb{R}^n to \mathbb{R}^* . Then $\forall x, y \in \mathbb{R}^n$,

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

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 $z \in \mathbb{R}^n$. Since $y \in \partial f(x)$, we get

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

Rearranging yields

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

It follows that

$$\sup_{u \in \mathbb{R}^n} (\langle y, u \rangle - f(u)) \leq \langle y, x \rangle - f(x). \quad (1)$$

By definition of supremum, we have

$$\sup_{u \in \mathbb{R}^n} (\langle y, u \rangle - f(u)) \geq \langle y, x \rangle - f(x). \quad (2)$$

From (1) and (2), we get

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

That is,

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

Then

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

$$= \langle z - y, x \rangle.$$

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

Chapter 14

Proximal Operators

14.1 Definitions

DEFINITION 14.1 (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 Examples

EXAMPLE 14.2 (Soft Threshold). Let $\lambda \geq 0$. Let f be a function from \mathbb{R} to \mathbb{R} given by $f(x) := \lambda|x|$. Then

$$\text{prox}_f(x) = \begin{cases} x + \lambda, & \text{if } x < -\lambda \\ 0, & \text{if } -\lambda \leq x \leq \lambda \\ x - \lambda, & \text{if } x > \lambda. \end{cases}$$

14.3 Basic Properties

(bug)

PROPOSITION 14.3. 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.

Proof. Let x be an arbitrary element of \mathbb{E} . Define a function $h : \mathbb{E} \rightarrow \mathbb{R}^*$ by $h(y) := \frac{1}{2}\|y-x\|^2$. Define a function $g : \mathbb{E} \rightarrow \mathbb{R}^*$ by $g(y) := f(y) + h(y)$. Then $\text{prox}_f(x) = \underset{y \in \mathbb{E}}{\operatorname{argmin}} g(y)$. Note that h is proper, lower semi-continuous, and β -strongly convex for any $\beta \in (0, 1)$. Since f and h are proper, g is proper (why?). Since f and h are lower semi-continuous, g is lower semi-continuous. Since f is convex and h is β -strongly convex, g is β -strongly convex. Since g is proper, lower semi-continuous, and strongly convex, g has a unique minimizer (why?). So $\text{prox}_f(x)$ is a singleton set. \square

not fully understood

PROPOSITION 14.4. Let C be a nonempty closed convex subset of \mathbb{E} . Then prox_{δ_C} and proj_C are both singleton and $\text{prox}_{\delta_C} = \text{proj}_C$.

Proof. Since C is nonempty, convex, and closed, δ_C is proper, convex, and lower semi-continuous and hence prox_{δ_C} is singleton. Since C is nonempty, convex, and closed, proj_C is singleton. Let x and p be arbitrary elements of \mathbb{E} . Then

$$\begin{aligned}
 & p \in \text{prox}_{\delta_C}(x) \\
 \iff & p \in \underset{y \in \mathbb{E}}{\operatorname{argmin}} \{ \delta_C(y) + \frac{1}{2}\|y-x\|^2 \} \\
 \iff & \forall y \in \mathbb{E}, \delta_C(y) + \frac{1}{2}\|y-x\|^2 \geq \delta_C(p) + \frac{1}{2}\|p-x\|^2 \\
 \iff & p \in C \text{ and } \forall y \in C, \frac{1}{2}\|y-x\|^2 \geq \frac{1}{2}\|p-x\|^2 \\
 \iff & p \in C \text{ and } \forall y \in C, \|y-x\| \geq \|p-x\| \\
 \iff & p \in \underset{y \in C}{\operatorname{argmin}} \|y-x\| \\
 \iff & p \in \text{proj}_C(x).
 \end{aligned}$$

\square

PROPOSITION 14.5 (Firmly Non-Expansive). Let f be a proper closed convex function. Then prox_f is firmly non-expansive.

14.4 Prox Calculus Rules

PROPOSITION 14.6 (Scaling and Translation).

THEOREM 14.7 (Norm Composition).

PROPOSITION 14.8. Let f_1, \dots, f_m be proper, convex, and lower semi-continuous functions from \mathbb{R} to \mathbb{R}^* . Define a function $f : \mathbb{R}^m \rightarrow \mathbb{R}^*$ by $f((x_i)_{i=1}^m) := \sum_{i=1}^m f_i(x_i)$. Then

$$\text{prox}_f((x_i)_{i=1}^m) = (\text{prox}_{f_i}(x_i))_{i=1}^m.$$

Proof. Since each f_i is proper, convex, and lower semi-continuous, f is proper, convex, and lower semi-continuous. Let $(x_i)_{i=1}^m$ and $(p_i)_{i=1}^m$ be arbitrary elements of \mathbb{R}^m . Then

$$(p_i)_{i=1}^m = \text{prox}_f((x_i)_{i=1}^m)$$

$$\iff$$

□

14.5 The Second Prox Theorem

PROPOSITION 14.9. 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 14.10. Let f be a proper, convex, and lower semi-continuous function from \mathbb{E} to \mathbb{R}^* . Let x and p be elements of \mathbb{E} . Then $p = \text{prox}_f(x)$ if and only if

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

Proof. Forward Direction:

Assume that $p = \text{prox}_f(x)$. I will show that $\forall y \in \mathbb{E}$, $\langle y - p, x - p \rangle \leq f(y) - f(p)$. Let y be an arbitrary element of \mathbb{E} . Define for each $\lambda \in (0, 1)$ a point p_λ by $p_\lambda := \lambda y + (1 - \lambda)p$. Then

$$\begin{aligned}
& p = \text{prox}_f(x) \\
& \implies f(p) + \frac{1}{2}\|x - p\|^2 \leq f(p_\lambda) + \frac{1}{2}\|x - p_\lambda\|^2 \\
& \iff f(p) \leq f(p_\lambda) + \frac{1}{2}\|x - p_\lambda\|^2 - \frac{1}{2}\|x - p\|^2 \\
& \iff f(p) \leq f(p_\lambda) + \frac{1}{2} \left\langle \left[(x - p_\lambda) + (x - p) \right], \left[(x - p_\lambda) - (x - p) \right] \right\rangle \\
& \iff f(p) \leq f(p_\lambda) + \frac{1}{2} \left\langle \left[2x - \lambda y - (1 - \lambda)p - p \right], \left[p - \lambda y - (1 - \lambda)p \right] \right\rangle \\
& \iff f(p) \leq f(p_\lambda) + \frac{1}{2} \left\langle \left[2(x - p) + \lambda(p - y) \right], \left[\lambda(p - y) \right] \right\rangle \\
& \iff f(p) \leq f(p_\lambda) + \lambda \langle x - p, p - y \rangle + \frac{1}{2}\lambda^2\|p - y\|^2 \\
& \iff f(p) \leq f(\lambda y + (1 - \lambda)p) + \lambda \langle x - p, p - y \rangle + \frac{1}{2}\lambda^2\|p - y\|^2 \\
& \implies f(p) \leq \lambda f(y) + (1 - \lambda)f(p) + \lambda \langle x - p, p - y \rangle + \frac{1}{2}\lambda^2\|p - y\|^2 \\
& \iff \lambda \langle y - p, x - p \rangle \leq \lambda f(y) - \lambda f(p) + \frac{1}{2}\lambda^2\|p - y\|^2 \\
& \iff \langle y - p, x - p \rangle \leq f(y) - f(p) + \frac{1}{2}\lambda\|p - y\|^2 \\
& \iff \langle y - p, x - p \rangle \leq f(y) - f(p).
\end{aligned}$$

Backward Direction:

Assume that $\forall y \in \mathbb{E}$, $\langle y - p, x - p \rangle \leq f(y) - f(p)$. I will show that $p = \text{prox}_f(x)$. Let y be an arbitrary element of \mathbb{E} . Then

$$\begin{aligned}
& \langle y - p, x - p \rangle \leq f(y) - f(p) \\
& \iff f(p) + \frac{1}{2}\|x - p\|^2 \leq f(y) + \langle x - p, p - y \rangle + \frac{1}{2}\|x - p\|^2 \\
& \implies f(p) + \frac{1}{2}\|x - p\|^2 \leq f(y) + \langle x - p, p - y \rangle + \frac{1}{2}\|x - p\|^2 + \frac{1}{2}\|p - y\|^2 \\
& \iff f(p) + \frac{1}{2}\|x - p\|^2 \leq f(y) + \frac{1}{2}\|(x - p) + (p - y)\|^2 \\
& \iff f(p) + \frac{1}{2}\|x - p\|^2 \leq f(y) + \frac{1}{2}\|x - y\|^2 \\
& \implies p = \text{prox}_f(x).
\end{aligned}$$

This completes the proof. □

14.6 Moreau Decomposition

THEOREM 14.11 (Moreau Decomposition - Version 1). 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

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

□

THEOREM 14.12 (Moreau Decomposition - Version 2). Let $S \subseteq \mathbb{R}^n$ be a nonempty closed convex cone. Let $z \in \mathbb{R}^n$. Then $\bar{z} = \text{proj}_S(z)$ if and only if $\bar{z} \in S$ and

$$\exists \bar{y} \in S^* \text{ such that } z = \bar{z} - \bar{y} \text{ and } \bar{z}^\top \bar{y} = 0.$$

In this case, we have $\bar{y} = \text{proj}_{S^*}(-z)$. Therefore, $\forall z \in \mathbb{R}^n$, we have

$$z = \text{proj}_S(z) - \text{proj}_{S^*}(-z) = [z]_+ - [-z]_+.$$

Chapter 15

Ellipsoids

DEFINITION 15.1 (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 15.2. The eigenvectors of A define the principal axes of the ellipsoid.