
Convex Optimization

Adrien Taylor

Class notes by Antoine Groudiev



Last modified 15th September 2024

Contents

1	Introduction	2
2	Convex sets	2
2.1	Definitions	2
2.2	Examples	2
2.2.1	Hyperplanes and halfspaces	2
2.2.2	Euclidian balls and ellipsoids	3
2.2.3	Cones	4
2.3	Convexity-preserving operations	5
2.3.1	Intersection and union	6
2.3.2	Affine functions	7
3	Convex functions	7
4	Convex problems	7

Abstract

This document is Antoine Groudiev's class notes while following the class *Deep Learning* at the Computer Science Department of ENS Ulm. It is freely inspired by the lectures of Adrien Taylor.

1 Introduction

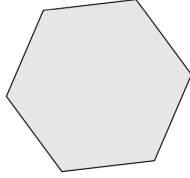
2 Convex sets

2.1 Definitions

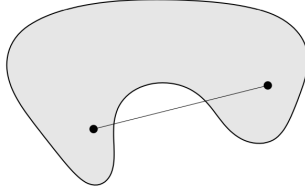
Definition (Convex set). A set C is a *convex set* if every segment that connects two points in C is in C . Formally:

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

Example. Here are some examples of convex and non-convex sets:



Convex



Non-convex



Non-convex

In many cases, we will use proper (i.e. non-empty) convex sets, and closed convex sets.

Definition (Convex hull). The *convex hull* of S , denoted $\text{Conv}(S)$, is the smallest convex set that contains S .

Definition (Convex combinations). The *convex combinations* of x_1, \dots, x_k are all the point x of the form:

$$x = \theta_1 x_1 + \dots + \theta_k x_k$$

with $\theta_1, \dots, \theta_k \geq 0$ and $\sum_{i=1}^k \theta_i = 1$.

Property 2.1. The convex hull of a set S is the set of all convex combinations of points in S :

$$\text{Conv}(S) = \left\{ \sum_{i=1}^k \theta_i x_i \mid (x_i) \in S^k, (\theta_i) \in \mathbb{R}_+^k, \sum_{i=1}^k \theta_i = 1 \right\}$$

2.2 Examples

2.2.1 Hyperplanes and halfspaces

Definition (Hyperplane). A *hyperplane* is the set of the form:

$$H = \{ x \mid a^\top x = b \}$$

for some $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$. a is called the *normal vector* of H . Hyperspaces are affine and convex.

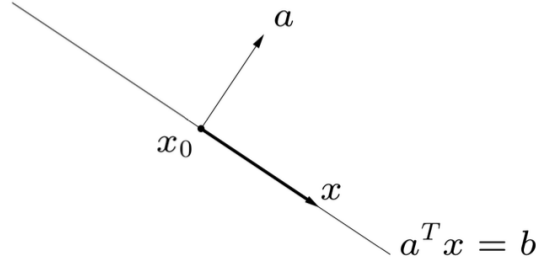


Figure 2.1: Hyperplane

Definition (Halfspace). A *halfspace* is the set of the form:

$$H = \{ x \mid a^\top x \leq b \}$$

for some $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$. a is called the *normal vector* of H . Halfspaces are convex.

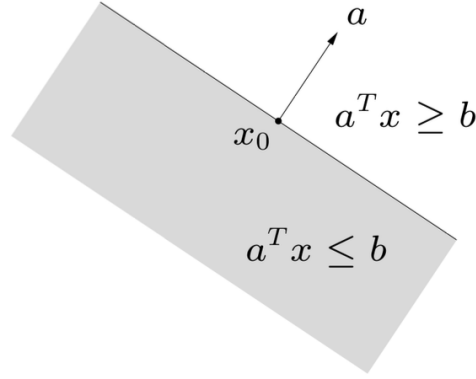


Figure 2.2: Halfspace

2.2.2 Euclidian balls and ellipsoids

Definition (Euclidian ball). The *Euclidian ball* of center x_c and radius r is the set:

$$B(x_c, r) = \{ x \mid \|x - x_c\|_2 \leq r \} = \{ x_c + ru \mid \|u\|_2 \leq 1 \}$$

Euclidian balls are convex.

Definition (Ellipsoid). An *ellipsoid* is the set of the form:

$$E = \{ x \mid (x - x_c)^\top P^{-1} (x - x_c) \leq 1 \}$$

with $P \in \mathbb{S}_{++}^n$ ¹ and $x_c \in \mathbb{R}^n$. Ellipsoids are convex.

¹ \mathbb{S}_{++}^n denotes the set of symmetric positive definite matrices of size n

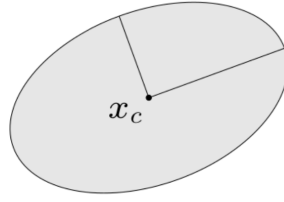


Figure 2.3: Ellipsoid

An alternative representation of an ellipsoid is:

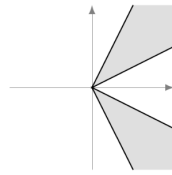
$$E = \{ x_c + Au \mid \|u\|_2 \leq 1 \}$$

for some nonsingular matrix $A \in \text{GL}_n(\mathbb{R})$. We can choose A symmetric and positive definite without loss of generality, for instance by choosing $A = P^{1/2}$.

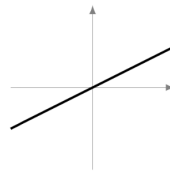
2.2.3 Cones

Definition (Cones). A set K is a *cone*, or a *nonnegative homogeneous set*, if:

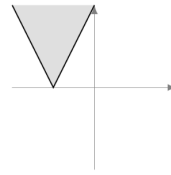
$$\forall x \in K, \forall \theta \in \mathbb{R}_+^*, \quad \theta x \in K$$



Cone



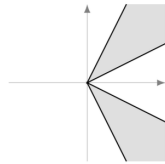
Cone



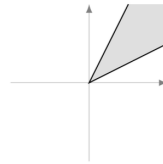
Not cone

Definition (Convex cone). A set K is a *convex cone* if:

$$\forall x_1, x_2 \in K, \forall \theta_1, \theta_2 \in \mathbb{R}_+^*, \quad \theta_1 x_1 + \theta_2 x_2 \in K$$



Non-convex



Convex

In the followings, we will denote by:

- $\mathbb{S}^n \subset \mathbb{R}^{n \times n}$ the set of symmetric matrices of size n
- \mathbb{S}_+^n the set of positive semidefinite matrices of size n , that is matrices verifying:

$$\forall z \in \mathbb{R}^n, z^\top X z \geq 0$$

also denoted $X \succcurlyeq 0$.

- \mathbb{S}_{++}^n the set of positive definite matrices of size n , that is matrices verifying:

$$\forall z \in \mathbb{R}^n, z^\top X z > 0$$

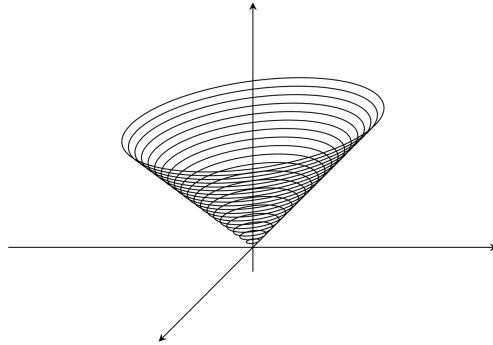
also denoted $X \succ 0$.

\mathbb{S}_+^n and \mathbb{S}_{++}^n are convex cones.

Special cases of cones include:

Positive orthant $K = \mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x_i \geq 0, \forall i\}$

Norm cones $K = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\| \leq t\}$. A particular case is the second-order cone (SOC), based on the ℓ_2 norm.



Positive polynomials $K_n = \{x \in \mathbb{R}^{n+1} \mid \forall t \in \mathbb{R}, \sum_{i=0}^n x_i t^i \geq 0\}$

Positive semidefinite cone $\mathbb{S}_+^n = \{X \in \mathbb{S}^n \mid \forall z \in \mathbb{R}^n, z^\top X z \geq 0\}$

Co-positive cone $\mathbb{S}_+^n = \{X \in \mathbb{S}^n \mid \forall z \in \mathbb{R}_+^n, z^\top X z \geq 0\}$

Exponential cone $\{(x, y, z) \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R} \mid z \geq y e^{x/y}\}$

Definition (Dual cones). The *dual cone* to a convex cone K is the set:

$$K^* = \{y \mid \forall x \in K, y^\top x \geq 0\}$$

Convex cones and their duals are particularly useful for convex duality. A convex cone that satisfies $K = K^*$ is called *self-dual*.

Definition (Polar cones). The *polar cone* to a convex cone K is the set:

$$K^\diamond = \{y \mid \forall x \in K, y^\top x \leq 0\}$$

We have the identity $K^\diamond = -K^*$.

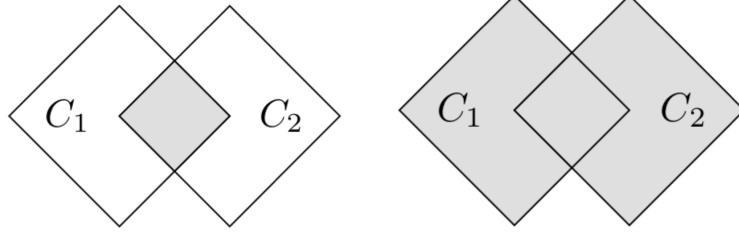
2.3 Convexity-preserving operations

To establish the convexity of a set C , the most basic approach is to apply the definition by proving that every segment that connects two points in C is in C . However, this can be tedious in practice. Instead, we can use operations that preserve convexity.

2.3.1 Intersection and union

Property 2.2 (Convexity is preserved by intersection). For any convex sets C_1 and C_2 , the intersection $C_1 \cap C_2$ is convex.

Likewise, the intersection of an arbitrary number of convex sets is convex.

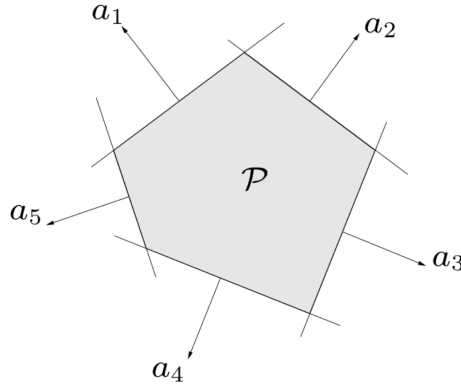


Remark. The union of convex sets is not necessarily convex. For instance in \mathbb{R} , both $[0, 1]$ and $[2, 3]$ are convex, but their union $[0, 1] \cup [2, 3]$ is not.

Definition (Polyhedron). A *polyhedron* is the solution set of finitely many linear inequalities and equalities:

$$S = \{ x \in \mathbb{R}^n \mid Ax \leq b, Cx = d \}$$

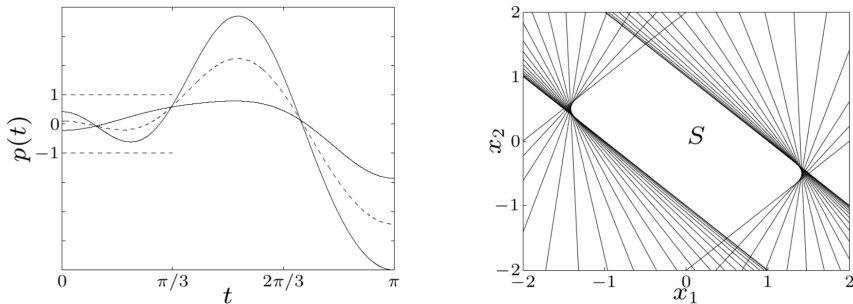
for $A \in \mathcal{M}_{m,n}(\mathbb{R})$ and $C \in \mathcal{M}_{p,n}(\mathbb{R})$. Polyhedra are convex, since they are the intersection of halfspaces and hyperplanes which are convex.



Example. Let:

$$S = \left\{ x \in \mathbb{R}^m \mid \forall t \in \mathbb{R}, \quad |t| \leq \frac{\pi}{3} \implies \left| \sum_{k=1}^m x_k \cos(kt) \right| \leq 1 \right\}$$

S is convex, since it can be written as the intersection of convex sets.



Example. \mathbb{S}_+^n is convex since it is the intersection of convex sets:

$$\mathbb{S}_+^n = \left\{ X \in \mathbb{S}^n \mid \forall z \in \mathbb{R}^n, z^\top X z \geq 0 \right\} = \mathbb{S}^n \cap \bigcap_{z \in \mathbb{R}^n} \left\{ X \in \mathcal{M}_n(\mathbb{R}) \mid z^\top X z \geq 0 \right\}$$

Each set $\left\{ X \in \mathcal{M}_n(\mathbb{R}) \mid z^\top X z \geq 0 \right\}$ being convex, their intersection is convex. In particular, it doesn't matter if the number of sets is finite, countable or uncountable.

2.3.2 Affine functions

Property 2.3 (The image of a convex set by an affine function is convex). If $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is an affine function, then if C is convex, $L(C)$ is convex.

More explicitly, let $A \in \mathcal{M}_{m,n}(\mathbb{R})$ and $b \in \mathbb{R}^m$. The affine function $L(x) = Ax + b$ maps C to $L(C) = \{ y \in \mathbb{R}^m \mid \exists x \in C, y = Ax + b \}$, which is convex if C is convex.

Property 2.4 (The pre-image of a convex set by an affine function is convex). If $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is an affine function, then $L^{-1}(C)$, the pre-image of C by L defined by:

$$L^{-1}(C) = \{ x \in \mathbb{R}^n \mid L(x) \in C \}$$

is convex if C is convex.

Example (Linear matrix inequalities). Let $A_1, \dots, A_m \in \mathbb{S}^n(\mathbb{R})$. The set:

$$\left\{ x \in \mathbb{R}^m \mid \sum_{i=1}^m x_i A_i \succcurlyeq 0 \right\}$$

is an affine pre-image of \mathbb{S}_+^n for the mapping $L : \mathbb{R}^m \rightarrow \mathbb{S}^n$ defined by:

$$L(x) = \sum_{i=1}^m x_i A_i$$

\mathbb{S}_+^n being convex, the set is convex. $\sum_{i=1}^m x_i A_i \succcurlyeq 0$ is called a *linear matrix inequality*.

3 Convex functions

4 Convex problems