Lecture 1

January 5, 2022

Introduction

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. Consider the problem:

$$(P) \qquad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in C \subseteq \mathbb{R}^n . \end{array}$$

In the special case, when $C = \mathbb{R}^n$ and f is "nice", the minimizers of f (if any) will occur at the critical points of f, namely, $x \in \mathbb{R}^n$ such that

$$\nabla f(x) = 0 .$$

This is known as "Fermat's rule", which we will learn about more later.

In this course

X is an Euclidean space with inner product $\langle \cdot , \cdot \rangle$ and induced norm $\|x\| = \sqrt{\langle x, x \rangle}$

Examples:

$\boxed{1}$ $X = \mathbb{R}^d$, $(\forall x \in X)(\forall y \in X)$

$\langle x, y \rangle = x^T y$

$\boxed{2}$ $X = \mathbb{R}^{n \times n}$, $(\forall x \in X)(\forall y \in X)$

$\langle x, y \rangle = \operatorname{tr}(x^T y)$.

In this Course, we will discuss and learn about Convexity of sets and functions and how we can approach problem (P) in the more general settings of :

1) Absence of differentiability of the Function f, f is Convex (this is called the objective function)

and / or

2) $\phi \neq C \subsetneqq X$, C convex (C is called the Constraint set).

Affine sets and affine subspaces in $\mathbb{R}^n$:

Definition 1.1.1

Let $S \subseteq X$.

S is an affine subspace if

$S \neq \phi$
and

$(\forall x \in S)\ (\forall y \in S)\ (\forall \lambda \in \mathbb{R})$
$\lambda x + (1-\lambda) y \in S$.

Let $S \subseteq X$. The affine hull of S, denoted by $\operatorname{aff}(S)$ is the intersection of all affine subspaces Containing S, i.e., the smallest affine subspace of X Containing S.

Examples of affine sets of $\mathbb{R}^n$:

1. L, where $L \subseteq \mathbb{R}^n$ is a linear subspace.

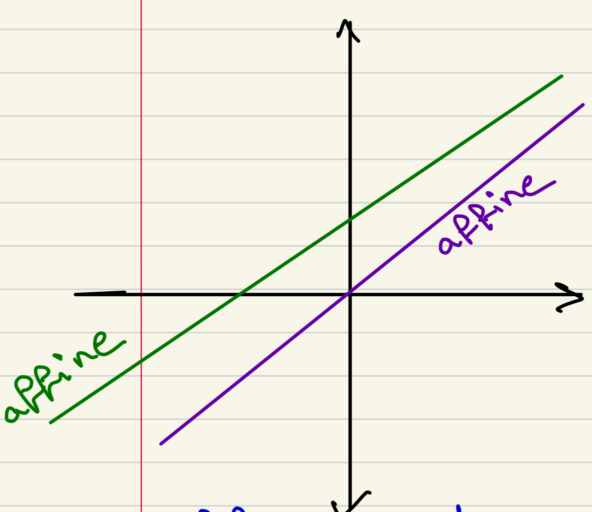
2. $a + L$, where $a \in \mathbb{R}^n$, $L \subseteq \mathbb{R}^n$ is a linear subspace.

3. $\mathbb{R}^n$

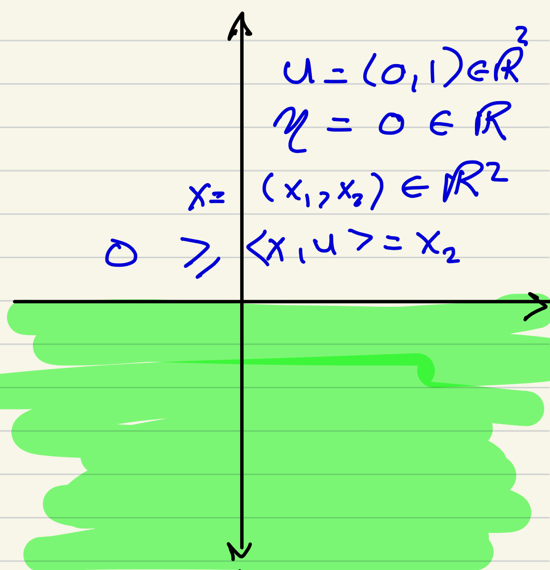
Geometrically speaking :

A nonempty subset $S \subseteq X$ is an affine subspace if the line connecting any two points in the set lies entirely in the set.

Examples in $\mathbb{R}^2$:



Affine Sets ✓



$\{ (x_1, x_2) \mid x_2 \leq 0 \}$

NOT affine

Convex sets in $\mathbb{R}^n$:

Definition L1-2

A subset C of X is **Convex** if $(\forall \lambda \in]0,1[)$ $(\forall x \in C)$ $(\forall y \in C)$

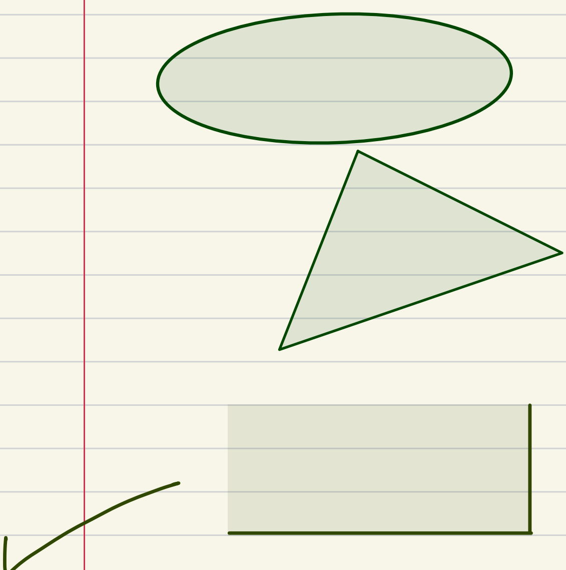
$$\lambda x + (1-\lambda) y \in C.$$

The following are examples of Convex subsets of $\mathbb{R}^d$:

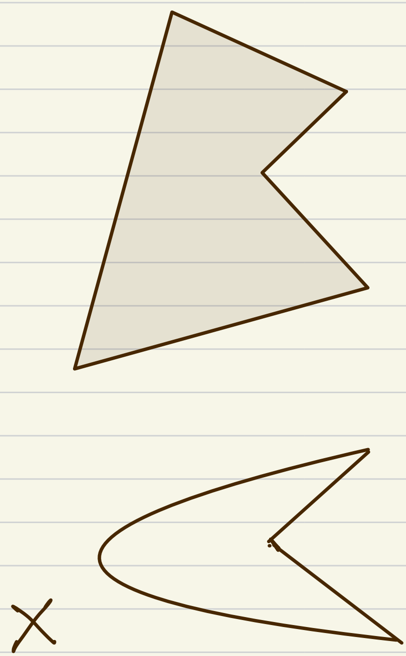
1. ϕ , $\mathbb{R}^d$.
2. C , where C is a ball.
3. C , where C is an affine subspace
4. C , where C is a half-space i.e., $C = \{ x \in \mathbb{R}^d \mid \langle x, u \rangle \leq \eta \}$ where $u \in \mathbb{R}^d$, $\eta \in \mathbb{R}$ are fixed.

Geometrically speaking :

A subset $C \subseteq X$ is convex if given any two points $x \in C$, $y \in C$, the line segment joining x and y, denoted by $[x, y]$, lies entirely in C.



Convex Sets



Not Convex

Theorem L1-3

The intersection of an arbitrary collection of convex sets in X is convex.

Proof:

Let I be an indexed set (not necessarily finite).

Let $(C_i)_{i\in I}$ be a collection of convex subsets of X.

Set $C := \bigcap_{i\in I} C_i$.

Goal: C is convex ?

If $C = \phi \Rightarrow C$ is convex.

Suppose that $C \neq \phi$.

Let $\lambda \in]0,1[$ and let $(x,y) \in C \times C$.

Since C_i is convex $(\forall i \in I)$,
we learn that
$(\forall i \in I) \quad \lambda x + (1-\lambda) y \in C_i$.
Hence,
$$\lambda x + (1-\lambda) y \in \bigcap_{i \in I} C_i = C.$$
Hence, C is convex.

The proof is complete. $\square$

Corollary L1-4

Let $b_i \in X$, $\beta_i \in \mathbb{R}$ for $i \in I$, where I is an arbitrary index set. Then the set:

$$C = \{ x \in X \mid \langle x, b_i \rangle \leq \beta_i, \forall i \in I \}$$

is Convex.

Proof:

See HW 1

Convex combinations of vectors:

Definition

A vector sum

$\lambda_1 x_1 + \cdots + \lambda_m x_m$

is called a convex combination of vectors $x_1, \ldots, x_m$ if $\forall i \in \{1, \ldots, m\}$ $\lambda_i \geq 0$, and $\sum_{i=1}^{m} \lambda_i = 1$.

Theorem L1-5

A subset C of X is convex if and only if it contains all the convex combinations of its elements.

Proof :

($\Leftarrow$) Suppose C contains all the convex combinations of its elements.

Let $\lambda \in]0,1[$ and let $x \in C$, $y \in C$.

By assumption, the convex combination

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

lies in C.

Therefore, C is convex

($\Rightarrow$) Suppose C is convex. ✓

We proceed by induction on m, where m is the number of elements in the convex combination.

Base case: when $m=2$, the conclusion is clear by the convexity of C.

Now, suppose that for some $m \geq 2$ it holds that any convex combination of m vectors lies in C.

Let $\{x_1, \dots, x_m, x_{m+1}\} \subseteq C$,

let $\lambda_1, \dots, \lambda_m, \lambda_{m+1} \geq 0$,

s.th. $\boxed{\sum_{i=1}^{m+1} \lambda_i = 1 .}$

Our goal is to show that

$$z := \sum_{i=1}^{m+1} \lambda_i x_i \in C . \quad ?$$

Observe that, there must exist at least one $\lambda_i \in [0,1[$ or else

if all $\lambda_i = 1 \Rightarrow 1 = \sum_{i=1}^{m+1} \lambda_i = m+1 > 3$

which is absurd $\#$

Without loss of generality, we can and do assume that $\lambda_{m+1} \in [0,1[$.

Now :

$$z = \sum_{i=1}^{m+1} \lambda_i x_i$$

$$= \sum_{i=1}^{m} \lambda_i x_i + \lambda_{m+1} x_{m+1}$$

$$= (1-\lambda_{m+1}) \sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} x_i + \lambda_{m+1} x_{m+1}$$

$$= \boxed{(1-\lambda_{m+1})} \sum_{i=1}^{m} \hat{\lambda}_i x_i + \boxed{\lambda_{m+1}} \underbrace{x_{m+1}}_{\in C}.$$

Observe that, $\hat{\lambda}_i := \frac{\lambda_i}{1-\lambda_{m+1}} \geq 0$

and that $\sum_{i=1}^{m} \hat{\lambda}_i = \frac{\lambda_1 + \cdots + \lambda_m}{1-\lambda_{m+1}}$

$$= \frac{1-\lambda_{m+1}}{1-\lambda_{m+1}} \quad \left[\text{by using } \lambda_1+\lambda_2+\cdots+\lambda_m+\lambda_{m+1}=1\right]$$

$$= 1$$

Using the inductive hypothesis,

we learn that

$$\sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} x_i \in C .$$

Hence,

$$z = (1-\lambda_{m+1}) \underbrace{\sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} x_i}_{\in C} + \lambda_{m+1} \underbrace{x_{m+1}}_{\in C}$$

$$\in C \qquad (C \text{ is Convex}).$$

The proof is complete.

$\square$

Definition L1-6 "Convex hull"

Let $S \subseteq X$. The intersection of all convex sets containing S is called the convex hull of S and is denoted by $\operatorname{Conv} S$.

By Theorem 2.1, $\operatorname{Conv} S$ is convex. In fact, it is the smallest convex set containing S.

Theorem L1-7

Let $S \subseteq X$. Then $\operatorname{conv} S$ consists of all the convex combinations of the elements of S, i.e.,

$$\operatorname{Conv} S = \left\{ \sum_{i \in I} \lambda_i x_i \;\middle|\; I \text{ is a finite index set}, (\forall i \in I)\ x_i \in S,\ \lambda_i \geqslant 0, \sum_{i \in I} \lambda_i = 1 \right\}.$$

Proof:

Set

$$D := \left\{ \sum_{i\in I} \lambda_i x_i \;\middle|\; I \text{ is a finite index set}, (\forall i \in I)\ x_i \in S, \lambda_i \geq 0, \sum_{i\in I} \lambda_i = 1 \right\}.$$

$\mathrm{Conv}\, S \subseteq D$??

Clearly, $S \subseteq D$. Moreover, D is convex. Indeed, let $d_1, d_2 \in D$, and let $\lambda \in]0,1[$.

Then, there exist

$\lambda_1, \ldots, \lambda_K \geq 0$, $\sum_{i=1}^{K} \lambda_i = 1$,

$\mu_1, \ldots, \mu_r \geq 0$, $\sum_{j=1}^{r} \mu_j = 1$

$d_1 = \sum_{i=1}^{K} \lambda_i x_i$, $\{x_1, \ldots, x_K\} \subseteq S$

$d_2 = \sum_{j=1}^{r} \mu_j y_j$, $\{y_1, \ldots, y_r\} \subseteq S$.

.

Therefore,

$$\lambda d_1 + (1-\lambda) d_2$$

$$= \lambda \lambda_1 x_1 + \cdots + \lambda \lambda_k x_k$$

$$+ (1-\lambda)\mu_1 y_1 + \cdots + (1-\lambda)\mu_r y_r .$$

Observe ~~that~~ $\lambda \lambda_i, (1-\lambda)\mu_j \geq 0$

$$i \in \{1, \ldots, k\}, \quad j \in \{1, \ldots, r\},$$

and that

$$\lambda \lambda_1 + \cdots + \lambda \lambda_k + (1-\lambda)\mu_1 + \cdots + (1-\lambda)\mu_r$$

$$= \lambda \sum_{i=1}^{k} \lambda_i + (1-\lambda) \sum_{j=1}^{r} \mu_j$$

$$= \lambda (1) + (1-\lambda)(1) = \lambda + 1 - \lambda = 1 .$$

Altogether, we conclude that

D is convex set $\supseteq S$.

Hence, $\operatorname{Conv} S \subseteq D$.

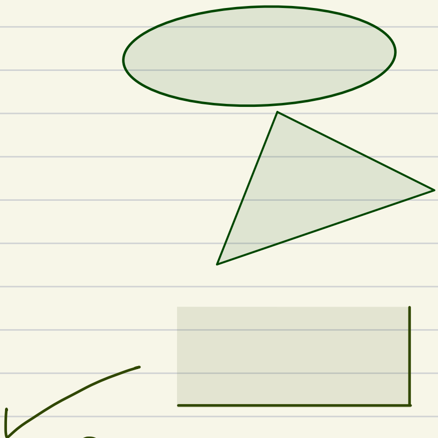
$D \subseteq \operatorname{Conv} S$?

Observe that $S \subseteq \operatorname{Conv} S$.

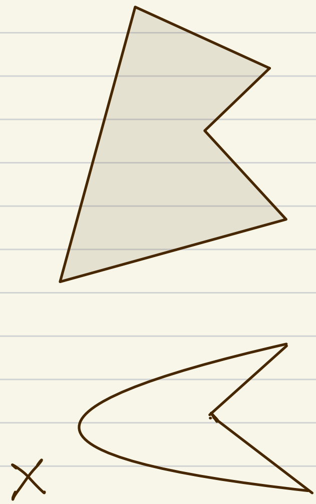
Now, Combine with Theorem 2.2 to learn that the Convex Combinations of elements of S lie in $\operatorname{Conv} S$.

$\square$

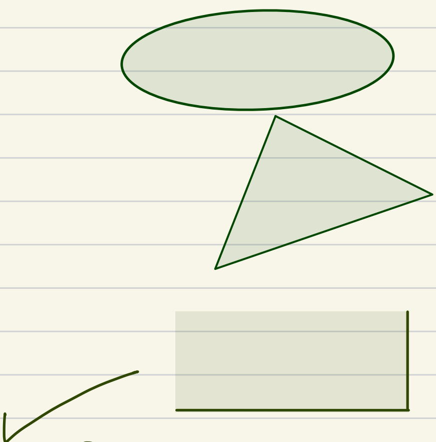
Convex hull: Examples.



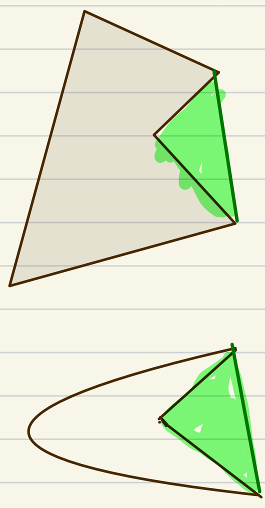
Convex sets



Not convex



Convex sets



Conv S