Convex Analysis for Optimization

Olga Kuryatnikova (Erasmus University Rotterdam)
kuryatnikova@ese.eur.nl

September-October 2024
Lecture 1

Organization

- Format: weekly lectures for 9 weeks
- Obligatory attendance of at least 7 lectures (Sept 9 to Nov 4)
- Grade: take-home assignment, groups of up to two students
- Weekly exercises, not graded, published on the course website
- Office hours or mistakes in the course material: contact us during the lecture or via email

Prerequisites

- Real analysis and linear algebra at bachelor's level

Literature

- D. Bertsekas, **Convex Optimization Theory**, Athena Scientific, 2009 (main book), online version
- S. Boyd and L. Vandenberghe, **Convex Optimization**, Cambridge University Press, 2004 (for more applications and details), online version
- R. T. Rockafellar. **Convex analysis**. Princeton University Press, 1970 or later editions (for somewhat more theory), online version

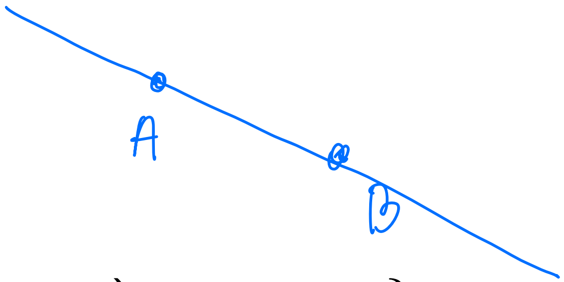
Course plan

- Week 1: Introduction to convexity
- Week 2: More on convex sets
- Week 3: More on convex functions
- Week 4: Dual description of convex functions
- Week 5: Duality and optimization
- Week 6: Introduction to algorithms, descend methods
- Week 7: Proximal methods, projected gradients
- Weeks 8 - 9: Fix point approach, averaged operators

On which sets we work

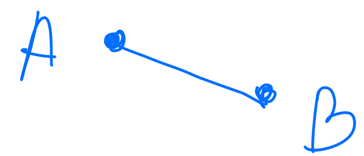
- Usually we just use $\mathbb{R}^n$
- Sometimes extended reals: $\overline{\mathbb{R}}^n \cup \{\infty\} \cup \{-\infty\}$
- All we do is generalizable to topological vector spaces

Convex set



Line L between points $x, y \in \mathbb{R}^n$ is

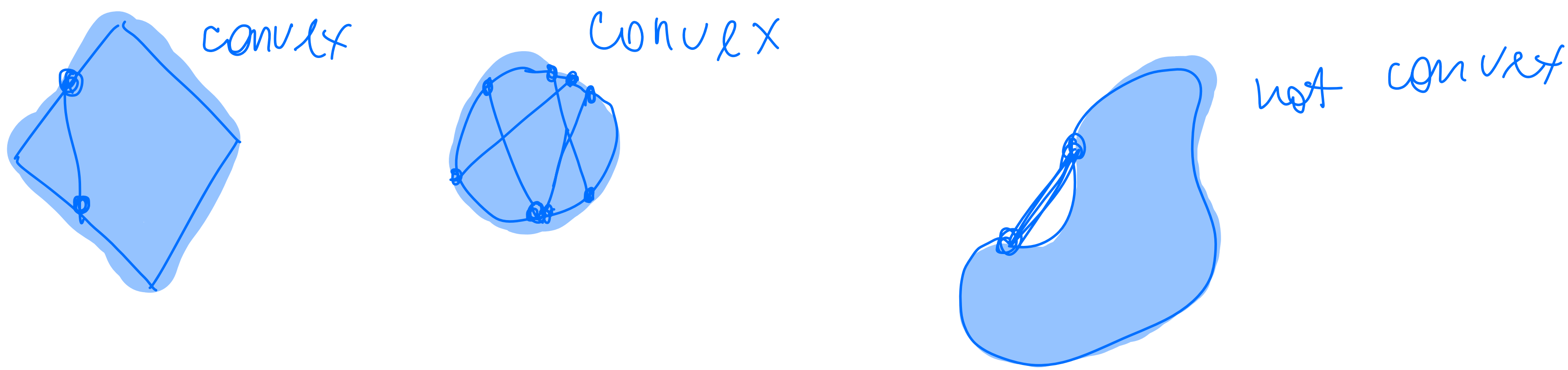
$$L := \{z \in \mathbb{R}^n : z = \alpha x + (1-\alpha)y,\ \alpha \in \mathbb{R}\}$$



Line segment LS between points $x, y \in \mathbb{R}^n$ is

$$LS := \{z \in \mathbb{R}^n : z = \alpha x + (1-\alpha)y,\ 1 \geq \alpha \geq 0\}$$

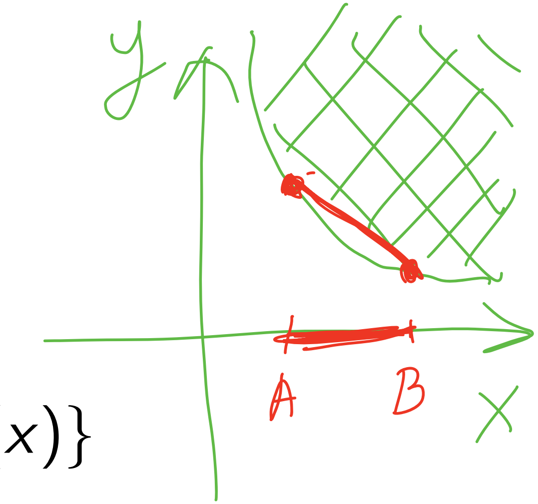
Def: convex set contains the line segment between its any two points



$\forall A, B \in S$: $\alpha A + (1-\alpha) B \in S$ $\rightarrow$ convex set

for any $0 \leq \alpha \leq 1$

Convex function



Epigraph of a function $f : S \to \overline{\mathbb{R}}$ is

$$\mathrm{epi}(f) := \{(x, t) \in S \times \mathbb{R} : x \in S, t \geq f(x)\}$$

- Def: a function is convex if it lies below the line segment between any two points in its domain

$$f(\alpha A + (1-\alpha) B) \leq \alpha f(A) + (1-\alpha) f(B)$$

- Another def: a function is convex if its epigraph is a convex set

(fits better for functions on extended line if they can be equal to ∞, $-\infty$)

Functions onto extended line

Indicator f-n of set S: could also be other constant

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

proper, defined on $\overline{\mathbb{R}}^n$

Domain of a function is the set where it is defined

Effective domain of a function $f : S \to \overline{\mathbb{R}}$ is

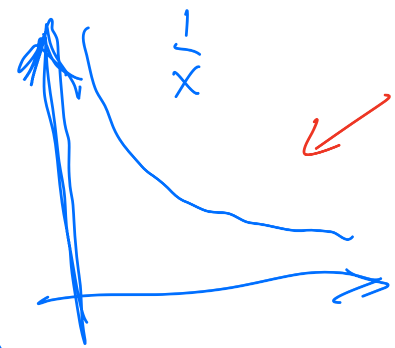
$$\mathrm{dom}(f) := \{x \in \mathbb{R}^n : f(x) < \infty\}$$

Def: f is proper if $f(x) < \infty$ for some $x \in S$ and $f(x) > -\infty$ for all $x \in S$ (i.e., its epigraph is non-empty and contains no vertical lines)

extended

$$\begin{cases} f(x) = \frac{1}{x} \text{ on } \underbrace{[0, \infty)}_{\text{domain}} \\ f(0) = \infty \end{cases}$$

Effective domain is $(0, \infty)$



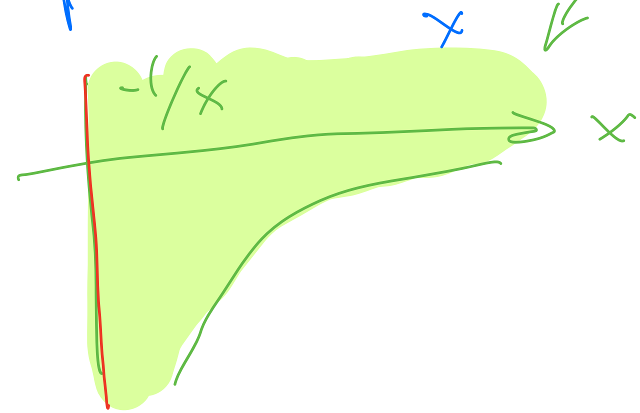
proper

extended f-n

$$\begin{cases} f(x) = -\frac{1}{x} \text{ on } [0, \infty) \\ f(0) = -\infty \end{cases}$$

not proper, epi contains vert. line $x = 0$

epi$(-\frac{1}{x})$



Convex optimization problem

epigraph reformulation:

A problem

$$\min_x \; f_0(x)$$
$$\text{s.t. } f_i(x) \leq 0, \qquad i = 1, \ldots, m,$$

$=$ $\min_{x,t} \; t$ s.t. $f_0(x) \leq t$, $f_i(x) \leq 0$, $\forall i = 1..m$

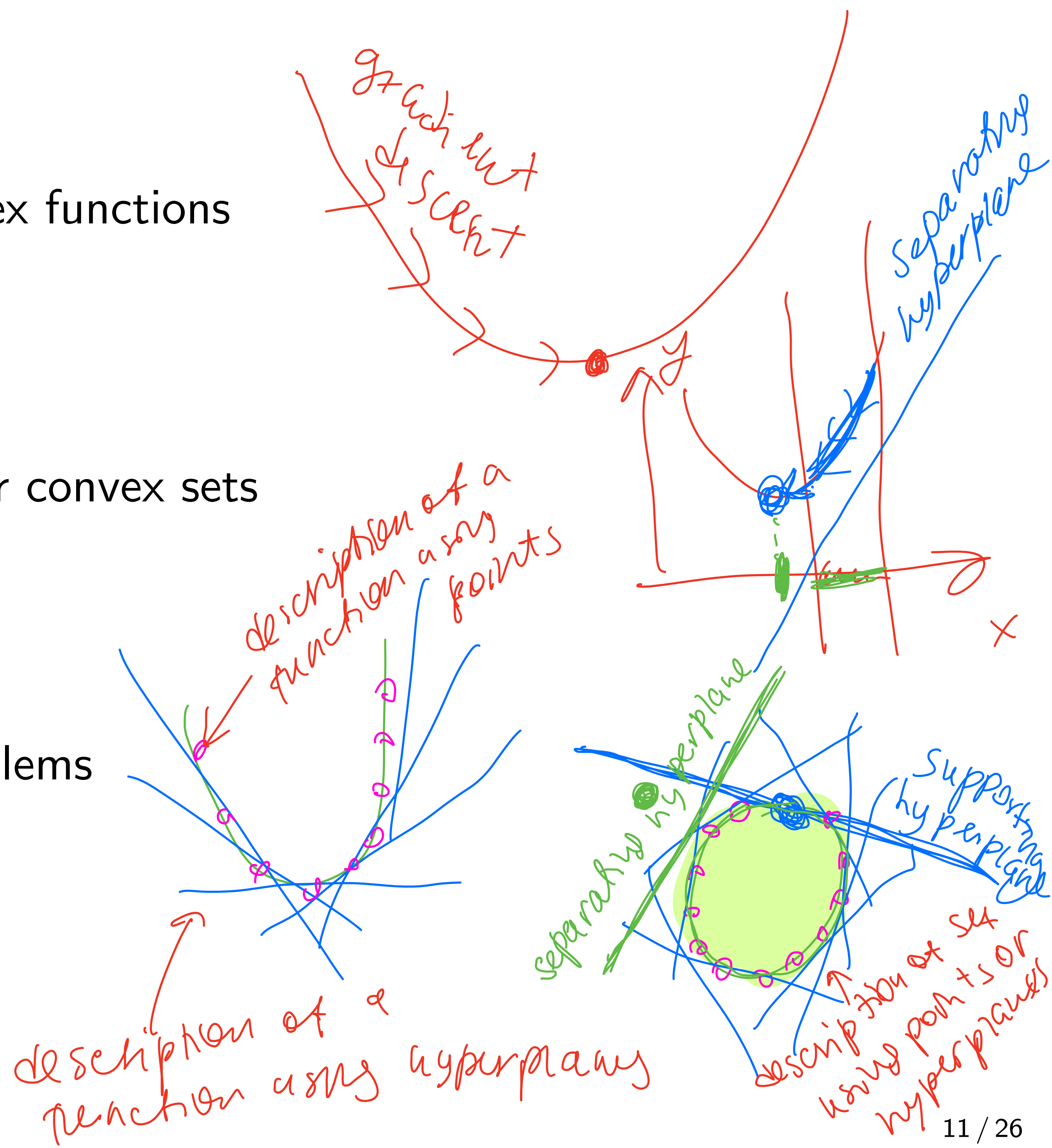
where all functions are convex.

Why convexity?

Global minima of convex functions

Separation theorems for convex sets

Duality for convex problems



Usage of convexity

- Convexity is a basis for more complex problems

stochastic component.

integer constraints

convex opt.

- Many data science problems (e.g., most regressions, SVM, PCA)
- Problems in physics (e.g., power, water, gas, signal processing)

Usage of convexity

- Convexity is a basis for more complex problems

stochastic component.

integer constraints

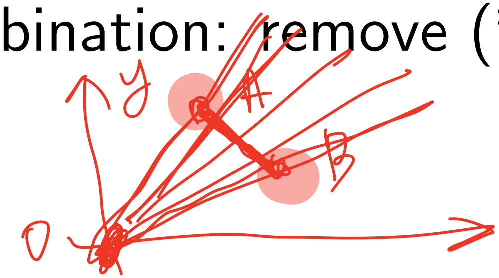
convex opt.

- Many data science problems (e.g., most regressions, SVM, PCA)
- Problems in physics (e.g., power, water, gas, signal processing)
- Other problems, e.g., neural networks, are not convex, but algorithms from this course help to find local optima
- Can also use convex approximations (e.g., McCormick envelopes, difference-of-convex algorithms, high-dimensional liftings)

Combinations

Def: Convex combination of $x_1, \dots, x_n$ is $\sum_{i=1}^{n} \alpha_i x_i$ for some $\alpha_1, \dots, \alpha_n$ where $\alpha_1, \dots, \alpha_n \geq 0$ (*) and $\sum_{i=1}^{n} \alpha_i = 1$ (**) [DCC]

Conic combination: remove (**) from [DCC]

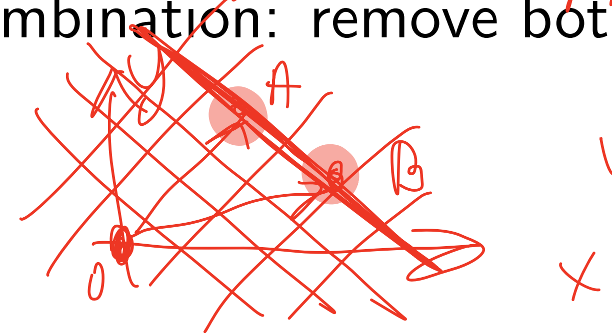


Affine combination: remove (*) from [DCC]

line R^1

$\rightarrow$ whole R^2

Linear combination: remove both (*) and (**) from [DCC]

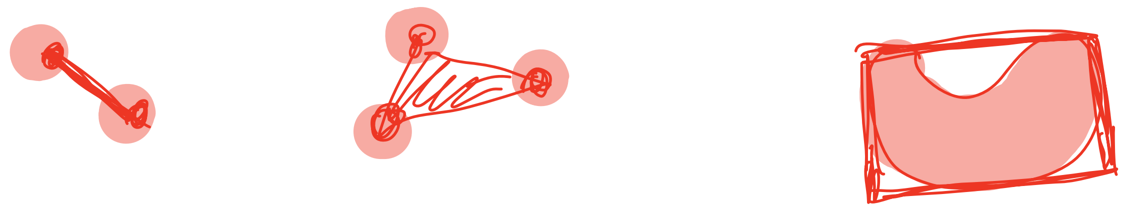


whole R^2

Convexifying sets

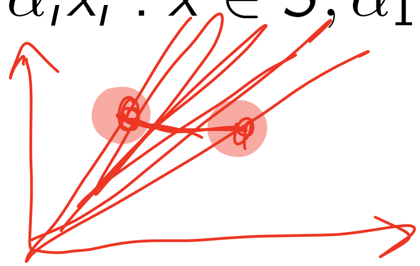
Convex hull of set S: (S is denoted by ●)

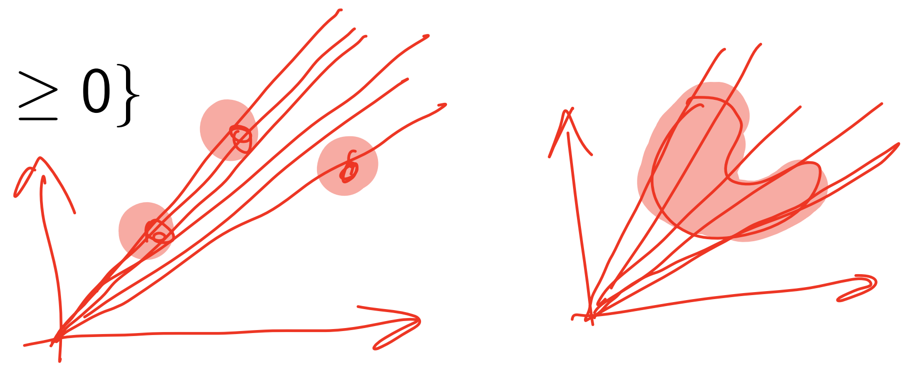
$$\mathrm{conv}(S) = \{\sum_{i=1}^{n} \alpha_i x_i : x \in S, \alpha_1, \ldots, \alpha_n \geq 0, \textstyle\sum_{i=1}^{n} \alpha_i = 1\}$$



Conic hull of set S:

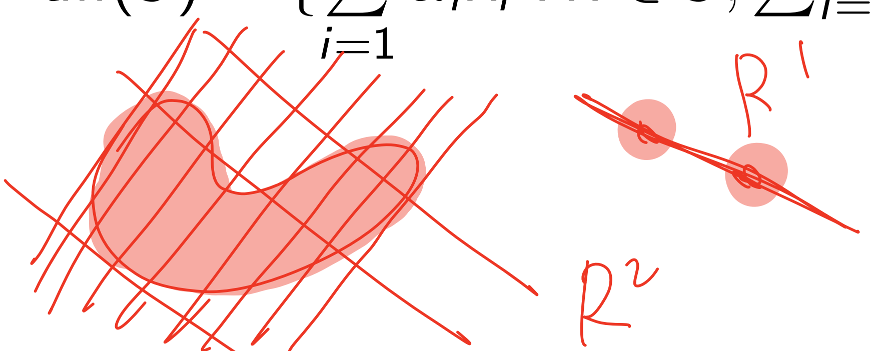
$$\mathrm{cone}(S) = \{\sum_{i=1}^{n} \alpha_i x_i : x \in S, \alpha_1, \ldots, \alpha_n \geq 0\}$$

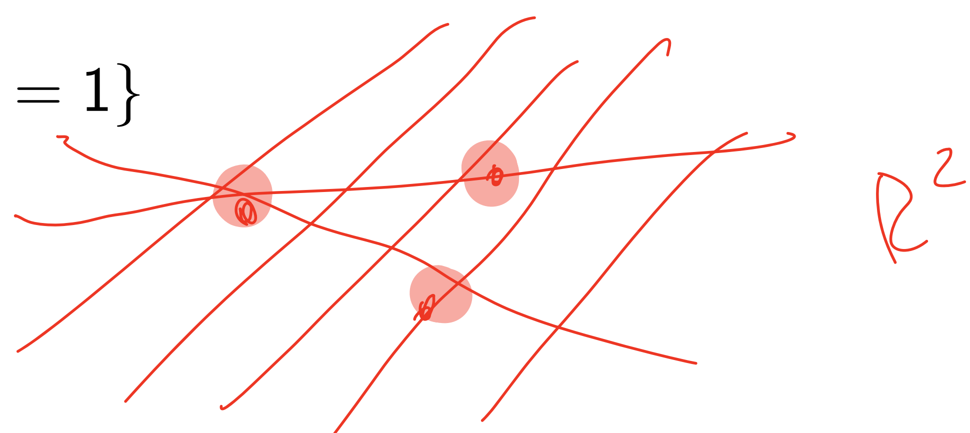




Affine hull of set S:

$$\mathrm{aff}(S) = \{\sum_{i=1}^{n} \alpha_i x_i : x \in S, \textstyle\sum_{i=1}^{n} \alpha_i = 1\}$$

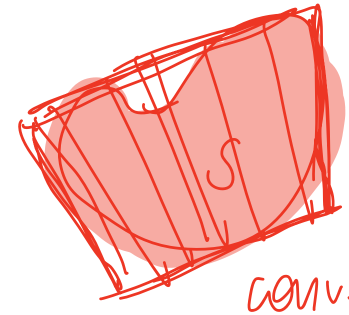




Dimension of a convex set

Dimension of a convex set is equal to the dimension of its affine hull

Caratheodory's Theorem



convex hull S

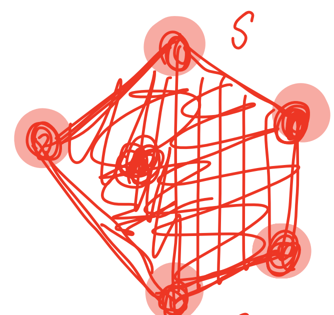
Let S be a nonempty subset of $\mathbb{R}^n$. Then

(a) Every $y \in \text{cone}(S), y \neq 0$ can be written as $\sum_{i=1}^{n} \alpha_i x_i$, where $x_1, \ldots, x_n \in S$ are linearly independent and $\alpha_1, \ldots, \alpha_n$ are positive.

(b) Every $y \in \text{conv}(S)$ is a convex combination of no more than $n+1$ elements from S.

↓

As a consequence, any x in S



convex hull of S, S are points

is a convex comb. of at most n+1 vertices of S, if S is convex and compact.

Proof of Caratheodory's Theorem

(See page 21 of the main book for the full proof of (a),(b))

(B) $x \in \mathrm{conv}(S) \Rightarrow x = \sum_{i=1}^{n+1} \alpha_i y_i$, $y_i \in S$: To prove

$S = [A, B]$, define a lifting

$\bar{S} = \{(x, 1)$ for all $x \in S\}$

We know $x = \sum \gamma_i y_i$, $\gamma_i \geq 0$, $\sum \gamma_i = 1$, $y_i \in S$

We need lifting $\bar{S}$ to make sure the conic combination has $\sum \alpha_i = 1$

$\bar{x} = (x, 1)$, then

$$\bar{x} = \sum \gamma_i \begin{bmatrix} 1 \\ y_i \end{bmatrix} = \begin{bmatrix} 1 \\ \sum \gamma_i y_i \end{bmatrix} \in \mathrm{cone}(\bar{S}) \Rightarrow$$

$$\bar{x} = \sum_{i=1}^{n+1} \alpha_i \bar{z}_i, \quad \bar{z}_i \in \bar{S} \Rightarrow \bar{x} = \sum_{i=1}^{n+1} \alpha_i \begin{bmatrix} 1 \\ z_i \end{bmatrix}, \quad z_i \in S$$

$$= \begin{bmatrix} \sum \alpha_i \\ \sum \alpha_i z_i \end{bmatrix} = \begin{bmatrix} 1 \\ \sum \alpha_i z_i \end{bmatrix}$$

$\sum \alpha_i = 1$, we have $n+1$ elements $z_i \in S$ ☒

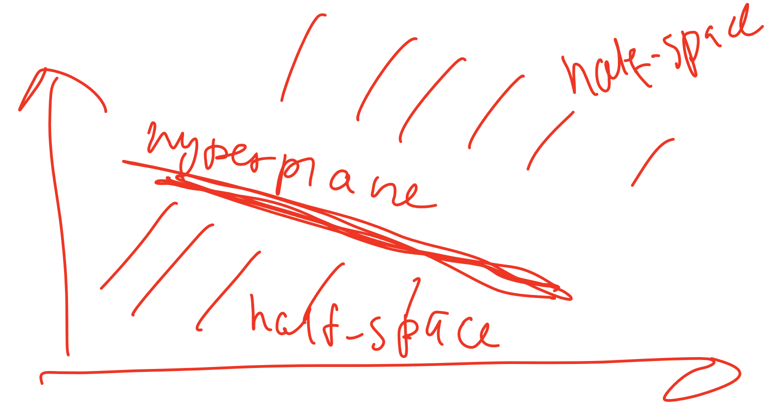
Affine transformation

An affine transformation L from vector space X to vector space Y:

$$L(x \in X) = Ax + b \in Y, \text{ for some linear operator } A \text{ and } b \in Y.$$

When $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, A is a matrix in $\mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Frequently used convex sets



- Hyperplane for some given $a \in \mathbb{R}^n$, $b \in \mathbb{R}$:

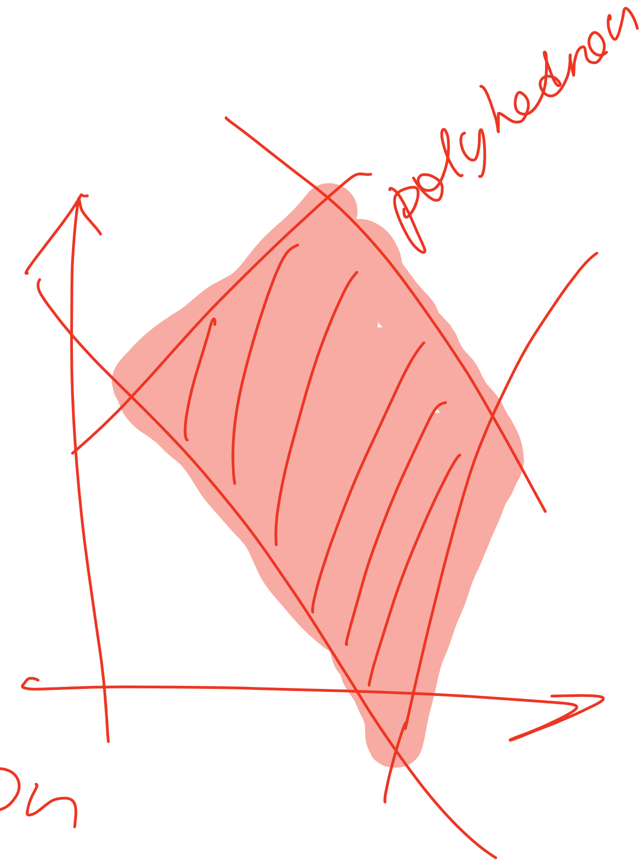
$$HP := \{x \in \mathbb{R}^n : a^\top x = b\}$$

- Half-space for some given $a \in \mathbb{R}^n$, $b \in \mathbb{R}$:

$$HS := \{x \in \mathbb{R}^n : a^\top x \leq b\}$$

- Polyhedron for some given $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^m$:

$$P := \{x \in \mathbb{R}^n : A^\top x \leq b\}$$

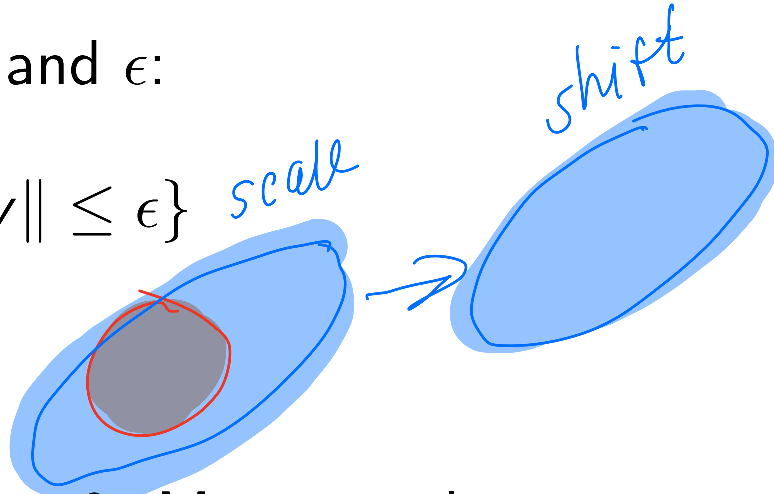


polytope is bounded polyhedron

More of frequently used sets

- Ball B for some given norm $\|\cdot\|$, center y, and ϵ:

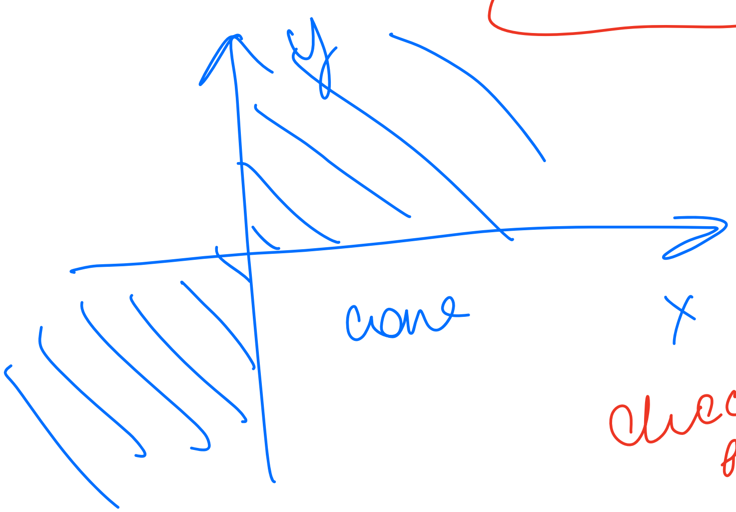
$$B(y,\epsilon) = \{x \in \mathbb{R}^n : \|x - y\| \leq \epsilon\}$$

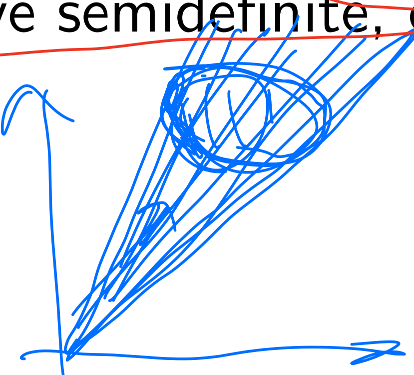


- Ellipsoid: affine transformation of a ball

$$\mathcal{E}(A, y, b, \epsilon) = \{x \in \mathbb{R}^n : \|A(x-y)+b\| \leq \epsilon\}$$

- Cone C: for all $x \in C$ we have $\alpha x \in S$ if $\alpha > 0$. Most popular convex cones: second-order, positive semidefinite, exponential.





convex cone

check Boyd and Vandenberghe book for more info about these cones

Closure of a set

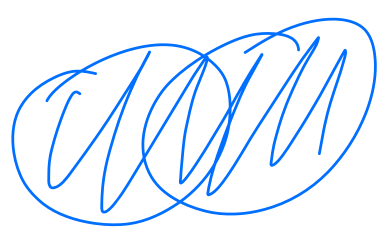
Closure of a set S is the set together with all its limit points (aka points that are limits of sequences belonging to S), denoted by $\text{cl}(S)$.

Convexity preserving operations on sets

- Intersection of any number of convex sets
- Cartesian product of convex sets
- Closure of a convex set
- Affine transformation (including projection onto some coordinates)
- Sum of elements of convex sets:
  $S = \{\sum_i x_i,\ x_i \in A_i,\ A_i \text{ are convex for all } i\}$
- Perspective mapping $S = \{x/t : [x, t] \in A,\ A \text{ is convex}\}$
- Linear-fractional mapping $S = \{\frac{Ax+b}{c^\top x+d} : x \in A,\ A \text{ is convex}\}$
- These are the main ones but not the only

prove

Counterexample: union of two convex sets can be non-convex



How to show a set is convex

- Apply definition
- Show the set is defined by convex functions
- Show the set is obtained from other convex sets via convexity preserving operations

Proof that linear-fractional map preserves convexity

$x \in S$, $y \in S$: $\dfrac{Ax+b}{c^Tx+d} = LFM(x)$

$\dfrac{Ay+b}{c^Ty+d} = LFM(y)$, need to show that

$$\alpha\left(\frac{Ax+b}{c^Tx+d}\right) + (1-\alpha)\left(\frac{Ay+b}{c^Tx+d}\right) \overset{*}{=} \frac{Az+b}{c^Tz+d}$$

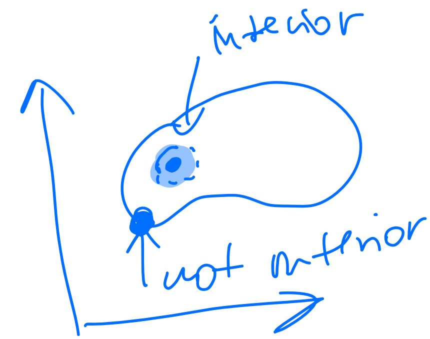
for some $z \in S$

Solution: as $z \in S$, $z = \lambda x + (1-\lambda)y$ for some $0 \le \lambda \le 1$, and to get the equality $*$ we set $\lambda = \dfrac{\alpha(c^Ty+d)}{\alpha(c^Ty+d)+(1-\alpha)(c^Tx+d)}$:

$$\frac{Az+b}{c^Tz+d} = \frac{A(\lambda x+(1-\lambda)y)+b}{c^T(\lambda x+(1-\lambda)y)+d} = \frac{\lambda(Ax+b)+(1-\lambda)(Ay+b)}{\lambda(c^Tx+d)+(1-\lambda)(c^Ty+d)} = \frac{\alpha(c^Ty+d)(Ax+b)+(1-\alpha)(c^Tx+d)(Ay+b)}{\underbrace{\alpha(c^Ty+d)(c^Tx+d)+(1-\alpha)(c^Tx+d)(c^Ty+d)}_{=(c^Ty+d)(c^Tx+d)}}$$

$$= \alpha\left(\frac{Ax+b}{c^Tx+d}\right)+(1-\alpha)\left(\frac{Ay+b}{c^Ty+d}\right) \quad \boxtimes$$

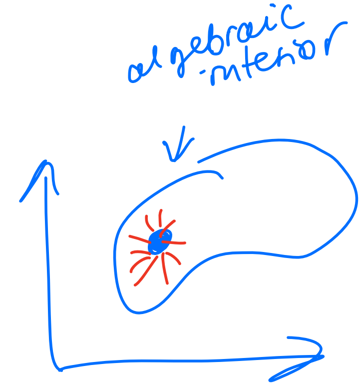
Concepts of interior



Let $S \subseteq \mathbb{R}^n$

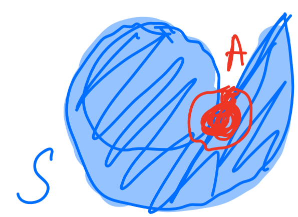
can be just "set"

- Interior:
  $\text{int}(S) := \{x \in S : \exists \text{ open ball } A \text{ such that } x \in A \subseteq S\}$



line segment

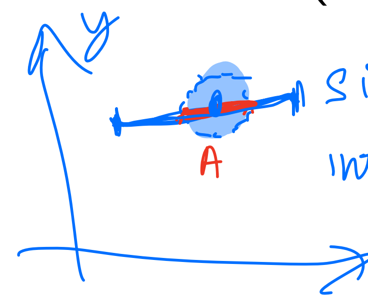
- Algebraic interior:
  $\text{core}(S) := \{x \in S : \forall z \in \mathbb{R}^n \; \exists \delta > 0 \text{ such that } [x, x + \delta z] \subseteq S\}$



here point $A \in$ core S, but $A \notin \text{int}(S)$, so $\text{core}(S) \neq \text{int}(S)$

- Relative interior:
  $\text{ri}(S) := \{x \in S : \exists \text{ open ball } A \text{ such that } x \in A \cap \text{aff}(S) \subseteq S\}$

facial reduction is used to reduce optimization search to $\text{aff}(S)$

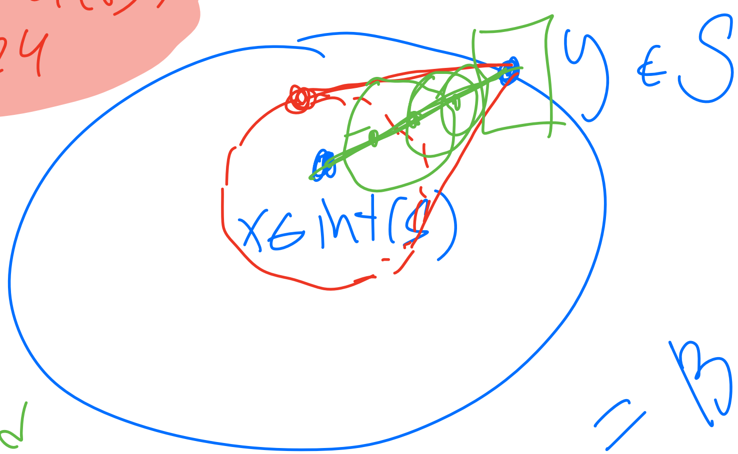


S is line segment in $\mathbb{R}^2$
$\text{int}(S) = \emptyset$, but $A \in \text{ri}(S)$,
so $\text{ri}(S) \neq \emptyset$

Line segment principle

Let $S \subseteq \mathbb{R}^n$ be a convex set. If $x \in \text{int}(S)$ (resp. $\text{ri}(S)$) and $y \in \text{cl}(S)$, then $[x, y) \subset \text{int}(S)$ (resp. $\text{ri}(S)$). In particular, $\text{int}(S)$ (resp. $\text{ri}(S)$) is a convex set. This is called "Line segment principle".

for the full proof including cl(S), see the main book, page 24



$= B(\varepsilon x + (1-\varepsilon)y, \varepsilon r)$

$B(x, r) \xrightarrow{\text{construct smaller shifted balls}} \varepsilon B(x, r) + (1-\varepsilon)y \to$ all smaller balls for $\varepsilon \in (0, 1]$

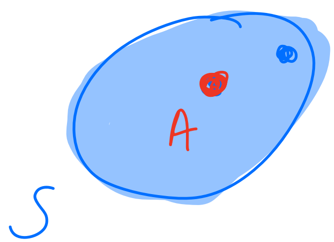
$\varepsilon \hat{x} + (1-\varepsilon)y \in S \ \forall \hat{x} \in B(x, r) \Rightarrow B(\varepsilon x + (1-\varepsilon)y, \varepsilon r) \subseteq S \Rightarrow$

$\Rightarrow \varepsilon x + (1-\varepsilon)y \in \text{int}(S)$ for $\forall \varepsilon \in (0, 1]$ by def. of $\text{int}(S)$ $\boxtimes$

Algebraic interior of convex sets

For convex sets, the definition of algebraic interior reduces to:
$\text{core}(S) := \{x \in S : \forall z \in \mathbb{R}^n\ \exists \delta > 0 \text{ such that } x + \delta z \in S\}$

To prove $A \in \text{core}(S)$ in this case, it is sufficient to prove that for each direction, there is a small enough step so that schift of A in this direction belongs to S.



$\text{core}(S) = \text{int}(S)$ for convex $S \subseteq \mathbb{R}^n$: can use them interchangeably in proofs. Can show using the Line Segment Principle for $\text{int}(S)$.

For non-convex sets, we can have $\text{core}(S) \neq \text{int}(S)$, see previous slide. For infinite-dimensional convex sets, we can also have $\text{core}(S) \neq \text{int}(S)$.