Optimisation

I) Convex sets

$E = (\mathbb{R}^N, \langle \cdot, \cdot \rangle)$

Def (Convex Set) A subset $K \subset \mathbb{R}^N$ is convex iff $\forall t \in [0,1], \forall x, y \in K,$

$$t x + (1-t) y \in K$$

Def (Convex Hull)

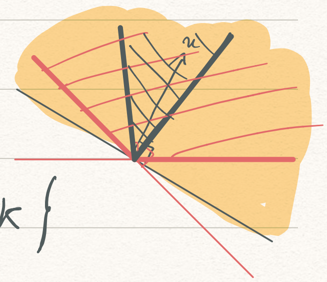
- $\mathrm{Conv}(T) = \left\{ \sum_{j=1}^{m} t_j x_j \; : \; m \geq 1, \; t_1, \dots, t_m \geq 0, \; \sum_{j=1}^{m} t_j = 1, \; x_1, \dots, x_m \in T \right\}$

- it is the smallest convex set containing T

- $\mathrm{Conv}(T) = \bigcap_{\substack{K \text{ convex} \\ T \subset K}} K$

Def (Cone) $K \subset \mathbb{R}^N$ is a cone if
$\forall t \geq 0, \forall x \in K, \quad tx \in K$

Dual cone: $K \subset \mathbb{R}^N$ a cone



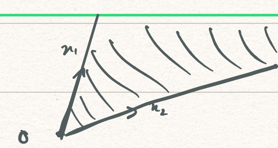
$$K^* := \{ z \in \mathbb{R}^N : \langle x, z \rangle \geq 0, \forall x \in K \}$$

- K^* is a closed convex cone referred to as "dual cone" of K.
- If K is a closed nonempty cone then $K^{**} = \mathrm{conv}(K)$

Polar Cone

$$K^\circ := \{ z \in \mathbb{R}^N : \langle x, z \rangle \leq 0, \forall z \in K \}$$
$$= -K^*$$

Def (Conic hull)



$$\mathrm{Cone}(T) = \left\{ \sum_{j=1}^{m} t_j x_j : m \geq 1, t_1, \dots, t_m \geq 0, x_1, \dots, x_m \in T \right\}$$

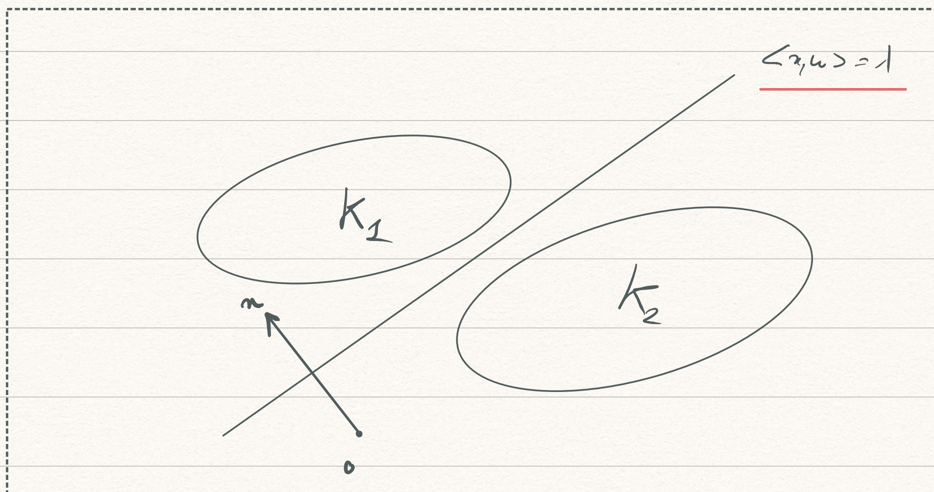
Theorem Let $K_1, K_2 \subset \mathbb{R}^N$ be convex sets

whose interiors have empty intersection

Then there exists $w \in \mathbb{R}^N$ and $\lambda \in \mathbb{R}$ s.t.

$$K_1 \subset \{ x \in \mathbb{R}^N : \langle x, w \rangle \leq \lambda \}$$

$$K_2 \subset \{ x \in \mathbb{R}^N : \langle x, w \rangle \geq \lambda \}$$

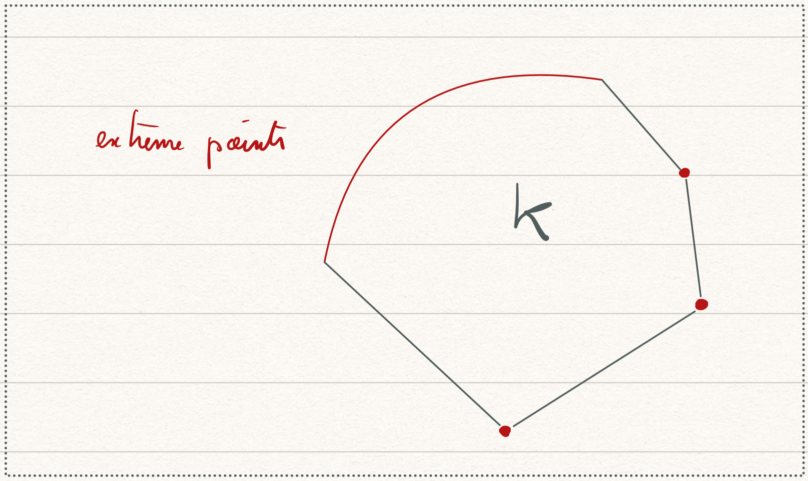


Def (Extreme point)

Let $K \subset \mathbb{R}^N$ be a convex set. A point $x \in K$ is an extreme point iff $x = ty + (1-t)z$ and $y,z \in K$, $t \in]0,1[$

$$\Rightarrow y = z = x$$

Theorem A compact convex set is the convex hull of its extreme points



II) Convex functions

$$F : \mathbb{R}^N \longrightarrow]-\infty, \infty] = \mathbb{R} \cup \{\infty\}$$

with the convention $\underset{x \in \mathbb{R}}{x + \infty = \infty}$, $\underset{\lambda > 0}{\lambda \cdot \infty = \infty}$

$$\mathrm{dom}(F) = \{ x \in \mathbb{R}^N, \ F(x) \neq \infty \}$$

- A function with $\mathrm{dom}(F) \neq \emptyset$ is called

<u>PROPER</u>

<u>Def</u> $F : \mathbb{R}^N \to]-\infty, \infty[$ is called:

- convex : $F(tx + (1-t)y) \leq tF(x) + (1-t)F(y)$
  $\forall x, y \in \mathbb{R}^N, \forall t \in [0,1]$

- strictly convex : $F(tx + (1-t)y) < tF(x) + (1-t)F(y)$
  $\forall x, y \in \mathbb{R}^N, \forall t \in]0,1[$

- strongly convex with parameter $\gamma > 0$ if

$$F(tx + (1-t)y) \leq tF(x) + (1-t)F(y) - \frac{\gamma}{2} t(1-t) \|x-y\|_2^2$$

<u>Prop</u> $F : \mathbb{R} \to \mathbb{R}$ convex <u>non decreasing</u>
and $G : \mathbb{R}^N \to \mathbb{R}$ convex
<u>then</u> $F \circ G$ is convex

ex: • A PSD (positive semi definite) symmetric matrix (i.e all eigenvalues are nonnegative)

$F(x) = x^T A x$ is convex

• A Positive Definite (i.e. all eigenvalues positive)

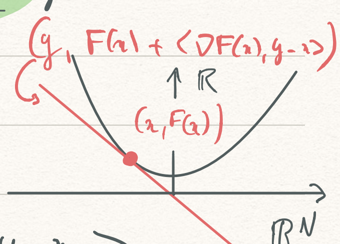
$F(x) = x^T A x$ is strongly convex

• For a convex set K, the characteristic function

$$\chi_K(x) := \begin{cases} 0 & \text{if } x \in K \\ \infty & \text{otherwise} \end{cases}$$

is convex

Prop: Let $F : \mathbb{R}^N \to \mathbb{R}$ be a differentiable function.



• F is convex iff $\forall x, y \in \mathbb{R}^N$

$$F(y) \geq F(x) + \langle \nabla F(x), y - x \rangle$$

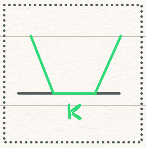
• F is strongly convex iff $\forall x, y \in \mathbb{R}^N$,

$$F(x) \geq F(y) + \langle \nabla F(y), x - y \rangle + \frac{\gamma}{2} \|x - y\|_2^2$$

- If F twice differentiable then it is convex iff $\forall x \in \mathbb{R}^N$, $\nabla^2 F(x) \succeq 0$ (all eigenvalues are nonnegative)

Prop A convex function is continuous on the interior of its domain

Prop $F : \mathbb{R}^N \to]-\infty, \infty]$ be a convex function

- A local minimizer of F is a global minimizer
- The set of minimizers of F is convex 

- If F is strictly convex then F has at most one minimizer

Theorem Let $f : \mathbb{R}^n \times \mathbb{R}^m \to]-\infty, \infty]$ be a convex function

Then $g(x) = \inf_{y \in \mathbb{R}^m} f(x, y)$, if well defined, is a convex function.

Theorem Let $K \subset \mathbb{R}^N$ be a compact convex set.
Let F be a convex function
then F attains its maximum at an extreme point of K

III] Convex conjugate

Def (Fenchel dual) Given $F : \mathbb{R}^N \to]-\infty, \infty]$,

the convex conjugate (or Fenchel dual) of F is

$$F^*(y) := \sup_{x \in \mathbb{R}^N} \left\{ \langle x, y \rangle - F(x) \right\} \in]-\infty, \infty]$$

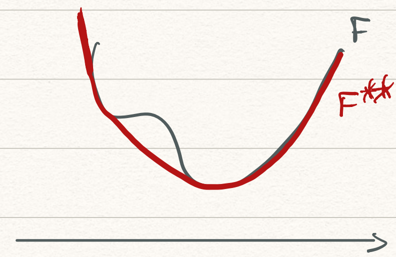
$$= - \inf_{x} \left\{ F(x) - \langle x, y \rangle \right\}$$

- F^* is convex, no matter whether F is convex or not

- Fenchel-Young $\forall x, y \in \mathbb{R}^N$,

$$\langle x, y \rangle \leq F(x) + F^*(y)$$

Prop $F : \mathbb{R}^N \to]-\infty, \infty]$



- F^* is lower semi continuous
- F^{**} is the largest lower semi continuous convex function satisfying $F^{**}(x) \leq F(x) \quad \forall x \in \mathbb{R}^N$.
  In particular, if F is convex and lower semi continuous then $\underline{F^{**} = F}$.
- For $c \neq 0$. If $F_c(x) = F(cx)$ then $F_c^*(y) = F^*(y/c)$
- For $c > 0$, $(cF)^*(y) = c\, F^*(y/c)$
- For $z \in \mathbb{R}^N$. If $F^{(z)} := F(\cdot - z)$ then $(F^{(z)})^*(y) = \langle z, y \rangle + F^*(y)$

Remark: Last week:

$$\boxed{\log \mathbb{P}[X - \mu \geq t] \leq \inf_{\lambda \in [0,b]} \left\{ \log[\varphi(\lambda)] - \lambda t \right\}}$$

$$= - F^*(t)$$

where $\left| \begin{array}{l} F(\lambda) = (\log \varphi)(\lambda) + \chi_{[a,b]}(\lambda) \\ \quad\;\; = (\log \varphi)(\lambda) \end{array} \right.$

<u>examples</u> • $F(x) = \frac{1}{2} \|x\|_2^2$, $F^*(y) = \frac{1}{2} \|y\|_2^2$

• $F(x) = e^x$, $F^*(y) = \left| \begin{array}{ll} y \ln y - y & \text{if } y > 0 \\ 0 & \text{if } y = 0 \\ \infty & \text{otherwise} \end{array} \right.$

<u>Fenchel-Young</u>: $xy \leq e^x + y \ln y - y$

$\forall x \in \mathbb{R}, \forall y > 0$

• $F(x) = \|x\|$ $\quad F^*(y) = \chi_{B_{\|.\|_*}}(y)$

$= \left| \begin{array}{ll} 0 & \text{if } \|y\|_* \leq 1 \\ \infty & \text{otherwise} \end{array} \right.$

• $F = \chi_K$ and K convex set

$F^*(y) = \sup_{x \in K} \langle x, y \rangle$

IV] Subdifferential

Def The subdifferential of a convex function $F : \mathbb{R}^N \to]-\infty, +\infty]$ at point $x \in \mathbb{R}^N$ is

$$\boxed{\partial F(x) = \left\{ v \in \mathbb{R}^N : F(y) \geq F(x) + \langle v, y - x \rangle, \forall y \in \mathbb{R}^N \right\}}$$

The elements of $\partial F(x)$ are called subgradients

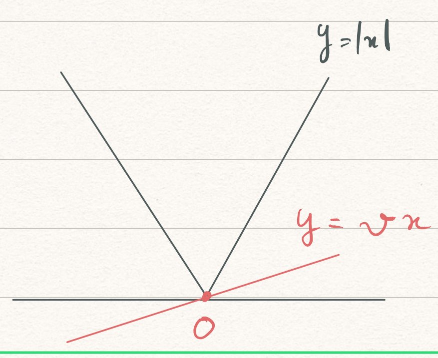
- The subdifferential of a convex function is always nonempty.
- If F is differentiable at x then

$$\partial F(x) = \left\{ \nabla f(x) \right\}$$

example $F : x \longmapsto |x|$

$$\partial F(x) = \begin{cases} \{\mathrm{sgn}(x)\} & \text{if } x \neq 0 \\ [-1, 1] & \text{if } x = 0 \end{cases}$$

$\underline{|y| \geq v y}$ $\quad (v y \leq |v| |y| \leq |y|)$



$y = v x$ when $v \in \partial F(0)$

Theorem x minimum of F convex iff $0 \in \partial F(x)$

Theorem $F: \mathbb{R}^N \to]-\infty, \infty]$ convex
$x, y \in \mathbb{R}^N$

(a) $y \in \partial F(x) \iff$ (b) $F(x) + F^*(y) = \langle x, y \rangle$

If F l.s.c then (a) $\iff$ (b) $\iff$ (c) where

(c) $x \in \partial F^*(y)$

Proximal Mapping

$F: \mathbb{R}^N \to]-\infty, \infty]$ convex

$$P_F(z) = \arg\min_{x \in \mathbb{R}^N} \left\{ F(x) + \frac{1}{2} \|x - z\|_2^2 \right\}$$

example K convex $\quad F = \chi_K \quad P_F = $ orth. proj. of K

Prop $F : \mathbb{R}^N \to]-\infty, \infty]$ convex

$$x = P_F(z) \iff z = x + \partial F(x)$$

Remark Some authors write $P_F = (Id + \partial F)^{-1}$

Theorem (Moreau's identity)

$F : \mathbb{R}^N \to]-\infty, \infty]$ convex l.s.c

then $\forall z \in \mathbb{R}^N$, $\quad P_F(z) + P_{F^*}(z) = z$

Theorem $\forall z, z'$, $\quad \| P_F(z) - P_F(z') \|_2 \leq \| z - z' \|_2$

example: $F(x) = |x|$, $\tau > 0$

$$P_{\tau F}(y) = \arg\min_{x \in \mathbb{R}} \left\{ \frac{1}{2}(x-y)^2 + \tau |x| \right\}$$

$$= \begin{cases} y - \tau & \text{if } y \geq \tau \\ 0 & \text{if } |y| \leq \tau \\ y + \tau & \text{if } y \leq -\tau \end{cases}$$

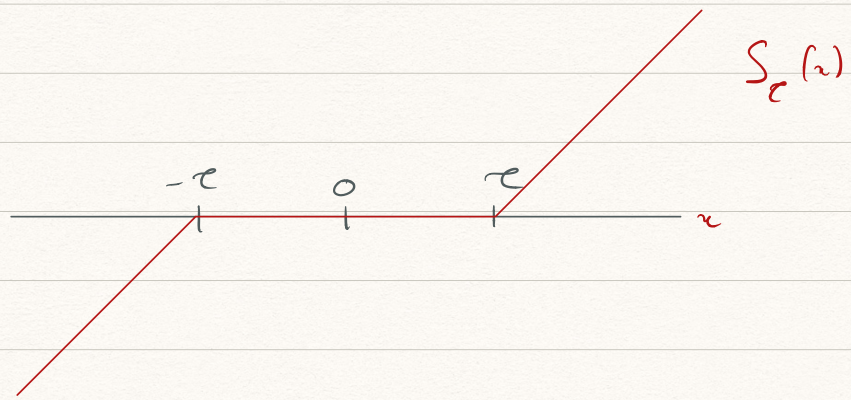
Soft thresholding S_τ

$x - y + \tau \partial|.|(x) = 0$

x tq

$x = y - \tau \partial|.|(x)$

$$= \begin{cases} y - \tau & x > 0 \\ y + \tau & x < 0 \\ y + \tau v & x = 0 \end{cases}$$



IV] Convex Optimization Problems

$$(P_0) \qquad \min_{x \in \mathbb{R}^N} F_0(x) \quad \text{subject to} \quad Ax = y \quad \text{and} \quad F_j(x) \le b_j, \quad j = 1, \dots, M$$

- F_0 : Objective
- F_j : constraint

- $K = \{ x \in \mathbb{R}^N : Ax = y \text{ and } F_j(x) \le b_j, \ \forall j \in [M] \}$

  set of constraints.

$(\mathcal{P}_0)$ is equivalent to:

$$\boxed{(\tilde{\mathcal{P}}_1) \qquad \min_{x \in K} F_0(x)}$$

or again

$$\boxed{(\mathcal{P}_1) \qquad \min_{x \in \mathbb{R}^N} \Big\{ \underbrace{F_0(x) + \chi_K(x)}_{F(x)} \Big\}}$$

Convex optimization $=$ F_0, F_ℓ convex $=$ F convex

Dual Problem of a convex program

Lagrangian: $\alpha \in \mathbb{R}^m$ $(\leftrightarrow A \in \mathbb{R}^{m \times N})$, $\beta \in \mathbb{R}^M$

$$\mathcal{L}(x, \alpha, \beta) = F_0(x) + \langle \alpha, Ax - y \rangle + \sum_{\ell=1}^{M} \beta_\ell \left(F_\ell(x) - b_\ell \right)$$

$$\boxed{\sup_{\substack{\alpha, \beta \\ \beta \geq 0}} \mathcal{L}(x, \alpha, \beta) = F_0(x) + \chi_K(x)}$$

$$\underline{\text{Primal}} = \operatorname*{Inf}_{x} \operatorname*{Sup}_{\substack{\alpha,\beta \\ \beta\geq 0}} \mathcal{L} = \operatorname{Inf} \left\{ F_0 + \chi_K \right\}$$

$$\underline{\text{Dual}} = \operatorname*{Sup}_{\substack{\alpha,\beta \\ \beta\geq 0}} \operatorname*{Inf}_{x} \mathcal{L} = \operatorname*{Sup}_{\alpha,\beta} H(\alpha,\beta)$$

$$H(\alpha, \underset{\geq 0}{\beta}) = \operatorname*{Inf}_{x\in\mathbb{R}^N} \left\{ F_0(x) + \langle \alpha, Ax-y \rangle + \sum_{\ell} \beta_\ell \left(F_\ell(x) - b_\ell\right) \right\}$$

· In particular for $x \in K$ one has:

$\forall x \in K,\ \forall \alpha \in \mathbb{R}^m,\ \forall \beta \geq 0,$

$$\boxed{\underline{H(\alpha,\beta) \leq F_0(x)}}$$

It shows that "weak duality holds":

$$\boxed{\operatorname*{Sup}_{\substack{\alpha,\beta \\ \beta\geq 0}} H(\alpha,\beta) \leq \operatorname*{Inf}_{x\in K} F_0(x)} \qquad (*)$$

Strong duality : equality in (*)

Theorem : If there exists x s.t

$$A x = b$$

and $F_\ell(x) < b_\ell \quad \ell = 1, \dots, M$

then Strong duality holds.

example : $\cdot \min \|x\|_1$ s.t $Ax = y$

$\cdot\ \mathcal{L}(x, \alpha) = \|x\|_1 + \langle \alpha, Ax - y \rangle$

$$H(\alpha) = \inf_{x \in \mathbb{R}^N} \left\{ \|x\|_1 + \langle A^T \alpha, x \rangle - \langle \alpha, y \rangle \right\}$$

$- F^*(-A^T\alpha)$ where $F(x) = \|x\|_1$

$F^* = \chi_{B_{\|.\|_\infty}}$

$$= - \langle \alpha, y \rangle + \inf_{x \in \mathbb{R}^N} \left\{ \|x\|_1 + \langle A^T \alpha, x \rangle \right\}$$

• <u>If</u> $\|A^T\alpha\|_\infty > 1$ then $\exists\, x \in \mathbb{R}^N$ t.q

$$\langle A^T\alpha, x\rangle < -\|x\|_1$$

$$\overset{\lambda>0}{\iff} \lambda\|x\|_1 + \lambda\langle A^T\alpha, x\rangle = \lambda\big(\underbrace{\|x\|_1 + \langle A^T\alpha, x\rangle}_{<0}\big)$$

<u>then</u> $\inf_x \{\|x\|_1 + \langle A^T\alpha, x\rangle\} = -\infty$

• If $\|A^T\alpha\|_\infty \leq 1$ then $|\langle A^T\alpha, x\rangle| \leq \|A^T\alpha\|_\infty \|x\|_1 \leq \|x\|_1$

then $\|x\|_1 + \langle A^T\alpha, x\rangle \geq 0$ and $\inf_x \{\|x\|_1 + \langle A^T\alpha, x\rangle\} = 0$

• $$\inf_{x\in\mathbb{R}^N} \{\|x\|_1 + \langle A^T\alpha, x\rangle\} = \begin{cases} 0 & \text{if } \|A^T\alpha\|_\infty \leq 1 \\ -\infty & \text{if } \|A^T\alpha\|_\infty > 1 \end{cases}$$

$$\boxed{H(\alpha) = \begin{cases} -\langle\alpha, y\rangle & \text{if } \|A^T\alpha\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}}$$

Dual :

$$\sup_{\alpha} H(\alpha) = \sup_{\substack{\alpha \text{ s.t.} \\ \|A^T\alpha\|_\infty \leq 1}} \{ - \langle \alpha, y \rangle \}$$

$$= - \inf_{\|A^T\alpha\|_\infty \leq 1} \{ \langle \alpha, y \rangle \}$$