Conditions d'Optimalité I

I) Définitions, motivations

On considère le problème

$$(P_X) \begin{cases} \min f(x) \\ x \in X \end{cases}$$

où $X \subset$ espace vectoriel $\mathbb{E}$ (produit scalaire noté $\langle \cdot, \cdot \rangle$) et $f : X \to \mathbb{R}$

Que sont les CO (conditions d'optimalité) ?

- rappel de l'optimisation sans contrainte (T)
  - $\Rightarrow$ on essaie de généraliser cela à (P_X)
  - $\Rightarrow$ c'est un ensemble d'$=$ ou d'$\leq$ ou de propriétés décrivant / caractérisant une solution
- x_X solution de (P_X) $\Rightarrow$ CN d'optimalité

  CS d'optimalité en x_* $\Rightarrow$ x_* "solution" de (P_X)
- CO du 1er ordre $\Rightarrow$ ne fait intervenir que les dérivées premières, pas plus

  CO du 2e ordre $\Rightarrow$ fait intervenir les dérivées secondes, pas plus

CO sans contrainte (rappel, § 4.2)

Le problème à résoudre :

$$\begin{cases} \min f(x) \\ x \in \mathbb{E}. \end{cases}$$

On note $\nabla f(x)$ et $\nabla^2 f(x)$ les gradient et hessien de f en x pour $\langle \cdot, \cdot \rangle$.

- **CN1** :

$$x_* \text{ min local} \quad \Longrightarrow \quad \nabla f(x_*) = 0.$$

  (Si f est convexe, c'est une **CS1** globale.)

- **CN2** :

$$x_* \text{ min local} \quad \Longrightarrow \quad \begin{cases} \nabla f(x_*) = 0 \\ \nabla^2 f(x_*) \succcurlyeq 0. \end{cases}$$

- **CS2** pour un minimum local strict :

$$\begin{cases} \nabla f(x_*) = 0 \\ \nabla^2 f(x_*) \succ 0 \end{cases} \quad \Longrightarrow \quad x_* \text{ min local strict}.$$

A quoi ça sert ?

- vérifier l'optimalité de x_* (CS)
- donner des informations sur la solution x_* (CN)
- calculer une solution analytiquement (sur papier) ou numériquement (CN)
- définir des tests d'arrêt d'algorithme (CN, CS)

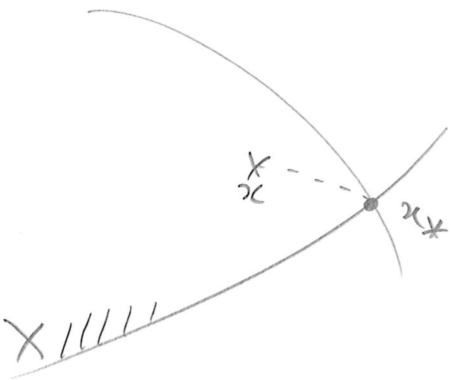
2) Une CN1 générale

↳ c'est elle que l'on va appliquer ensuite à des cas particuliers

$\boxed{A}$ Cône tangent

$$(P_X)\ \begin{cases}\min f(x)\\ x\in X\end{cases}$$

- On cherche à exprimer que f croît si l'on va de x_* vers l'"intérieur" de X



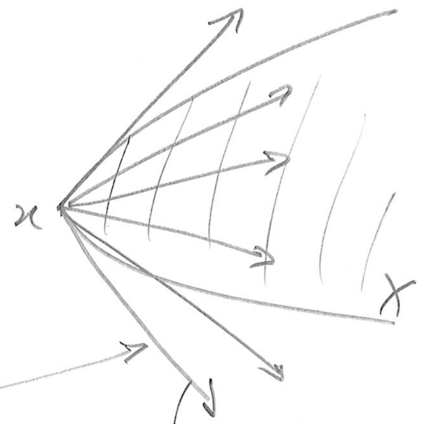
$f(x)\geq f(x_*)$
$\forall x\in X$ voisin de x_*
(c'est 1 condition à l'ordre 0, trop verbale pour être utile)

- Si l'on veut exprimer l'optimalité à l'ordre 1, il faut pouvoir "linéariser" X en x_* (comme on linéarise f en x_* dans le cas sans contrainte pour écrire $Df(x_*)=0$)
  - → conduit à la notion de cône tangent
  - → 2e raison pour laquelle on a besoin d'analyse convexe (un cône $\notin$ algèbre linéaire) (1re raison : unicité de solution)

- $K\subset \mathbb{E}$ est un cône si $\mathbb{R}_{++}K\subset K$
  c-à-d.
  $\forall t>0,\ \forall x\in K :\ tx\in K$

  ↑ pas "≥ 0" pour que l'on puisse parler du cône des matrices définies positives

– vecteur tangent à X en x

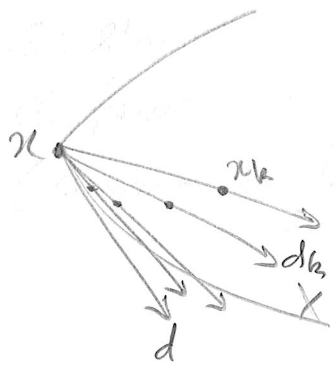


• il ne suffit pas de demander que $x+td \in X$ pour tout $t>0$ petit ; cela éliminerait les vecteurs extrêmes (ou toutes directions tangentes à une sphère)

tous des vecteurs tangents (translatés en $x \in X$)

• un passage à la limite est donc nécessaire (normal)

Dfn : d est tangent à X en x si



$\exists (d_k) \to d$
$\exists (t_k) \downarrow 0$ tels que
$x_k := x + t_k d_k \in X$

ou (plus opérationnel)

$$\boxed{\exists (x_k) \subset X,\ \exists (t_k) \downarrow 0 \text{ t.q. } \frac{x_k - x}{t_k} \to d}$$

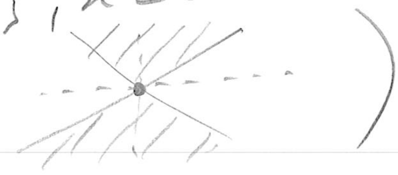
(forcément $x_k \to x$)

• cône tangent, noté $T_x X$ ou $T_X(x)$
c'est l'ensemble des directions tangentes

– c'est un cône fermé non vide ($0 \in T_x X$)

– $\boxed{X \text{ convexe}} \Rightarrow \boxed{T_x X \text{ convexe}}$

(faux si X n'est pas convexe ; exemple
$X = \{x \in \mathbb{R}^2 : x_1^2 \le x_2^2\}$, $x=0$,
$T_x X \equiv X$



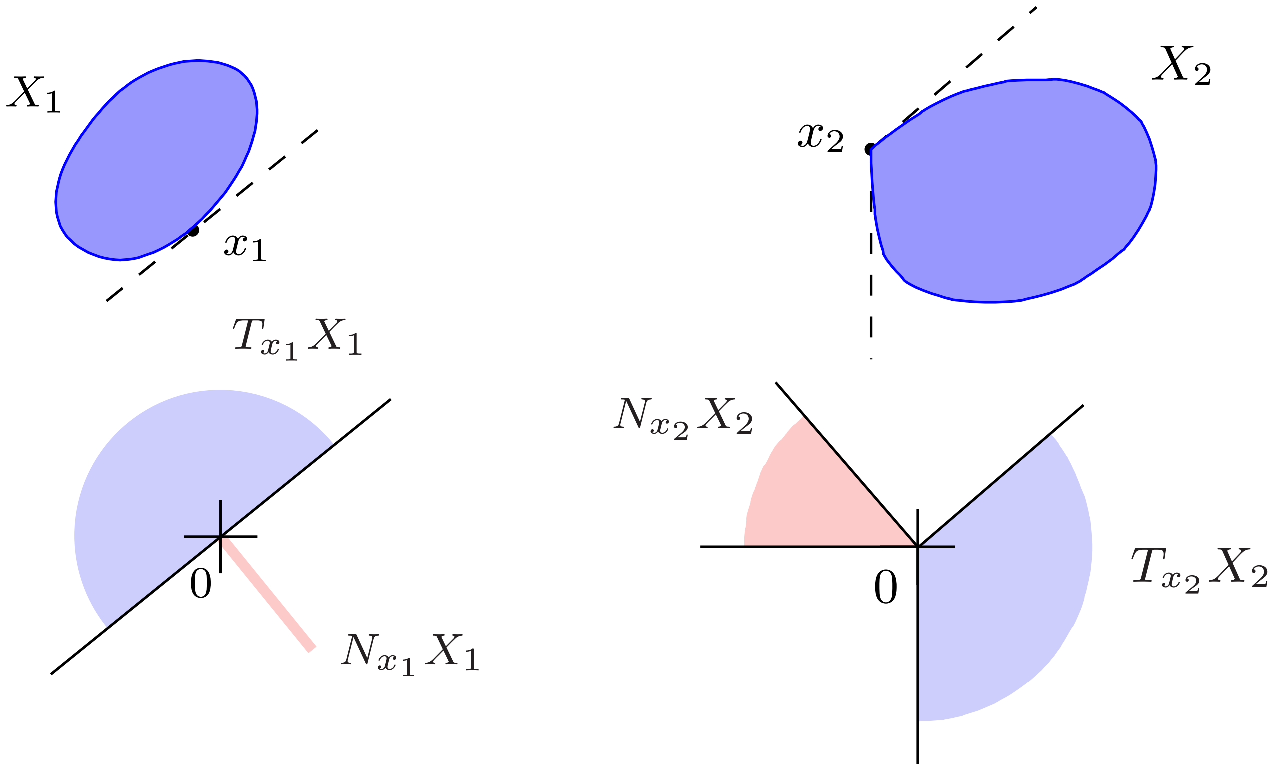
CN1 de Peano-Kantorovitch (§ 4.1)

Le problème à résoudre :

$$(P_X) \quad \begin{cases} \min f(x) \\ x \in X. \end{cases}$$

Dfn. Cône tangent : $d \in \mathrm{T}_x X \iff$

$$\exists \{x_k\} \subset X, \quad \exists \{t_k\} \downarrow 0 : \quad \frac{x_k - x}{t_k} \to d.$$



B CN1 de PEANO - KANTOROVITCH

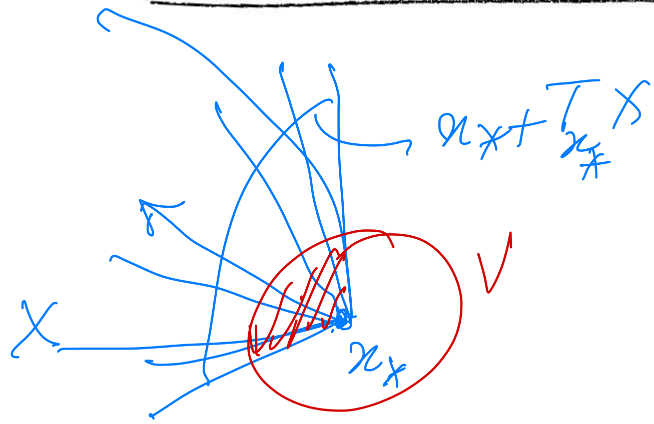
$$(P_X)\ \begin{cases} \min f(x) \\ x \in X \end{cases}$$

Théor

si • x_* minimum local de (P_X)

• f dérivable en x_*

alors $\forall d \in T_{x_*} X, \quad f'(x_*) . d \geq 0$



Dém

$d \in T_{x_*} X \setminus \{0\}$

$$\Rightarrow \exists (x_k) \subset X, \ \exists (t_k) \downarrow 0, \ \frac{x_k - x_*}{t_k} \to d$$

$$\Rightarrow x_k \to x_*$$

$$\cancel{f(x_*)} \underset{k \text{ grand}}{\leq} f(x_k) = \cancel{f(x_*)} + f'(x_*)(x_k - x_*) + o(\|x_k - x_*\|)$$

$$0 \leq f'(x_*) \underbrace{\frac{x_k - x_*}{t_k}}_{\to\, d} + \underbrace{\frac{o(\|x_k - x_*\|)}{\|x_k - x_*\|}}_{\to\, 0} \underbrace{\frac{\|x_k - x_*\|}{t}}_{\to\, \|d\|}$$

$\underbrace{0 \quad \|d\|}_{\to 0}$

$$0 \leq f'(x_*) . d$$

$\square$

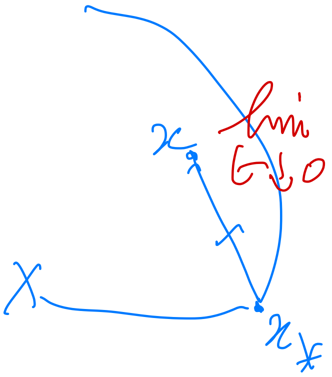
C Problèmes avec convexité

(alors on n'a pas besoin de cône tangent)

$$\boxed{\begin{array}{l} x_* \text{ min local de } (P_X) \\ X \text{ convexe} \\ f \text{ a des DD en } x_* \end{array}} \Rightarrow \boxed{\begin{array}{l} \forall x \in X \\ f'(x_*; x - x_*) \geq 0 \end{array}}$$

$$f'(x; d) \overset{\text{def}}{=} \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t}$$

<u>Dem</u> $x \in X$, $t > 0$ petit



$$\lim_{t \downarrow 0} \frac{f(x_* + t(x - x_*)) - f(x_*)}{t} \geq 0$$

$$f'(x_*; x - x_*) \geq 0$$

$\square$

<u>réciproque</u> (CS1 en optimisation <u>convexe</u>)

$$\boxed{\begin{array}{l} X \text{ convexe}, \; x_* \in X \\ f \text{ convexe sur } X \\ f'(x_*; x - x_*) \geq 0, \; \forall x \in X \end{array}} \Rightarrow \boxed{\begin{array}{l} x_* \text{ est un min} \\ \text{global de } (P_X) \end{array}}$$

<u>Dem</u> $x \in X$

$$f(x) \geq f(x_*) + \underbrace{f'(x_*; x - x_*)}_{\geq 0}$$

$$\Rightarrow f(x) \geq f(x_*), \quad \forall x \in X$$

$\square$

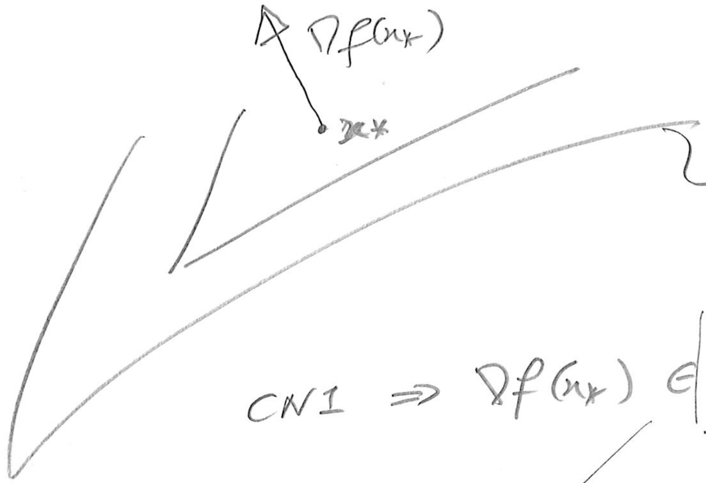
D Que fait-on maintenant ?

Pour aller plus loin, on a besoin d'analyse convexe

→ pas s'il n'y a que des égalités

$$(P_E)\ \begin{cases}\min f(x)\\ c(x)=0\end{cases}$$

$$\nabla f(x_*)^T d = 0 \quad \forall d \in N(c'(x_*))$$



$X_E = \{x : c(x) = 0\}$

$$\text{CN1} \Rightarrow \nabla f(x_*) \in \boxed{N(c'(x_*))^{\perp} = R(c'(x_*)^T)}$$

relation d'algèbre linéaire essentielle en optimisation

$$\nabla f(x_*) = c'(x_*)^T(-\lambda_*) \qquad \exists \lambda_* : \boxed{\nabla f(x_*) + c'(x_*)^T \lambda_* = 0}$$

→ s'il y a des inégalités on a besoin de remplacer la relation

$$N(A)^{\perp} := \{x : Ax = 0\}^{\perp} = R(A^T)$$

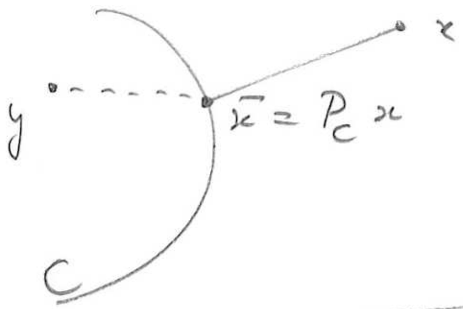
par quelque chose du type

$$\{x : Ax \underset{\geq}{\leq} 0\}^{+} = \ldots.$$

c'est le <u>lemme de Farkas</u>

3) Analyse convexe

A Projection sur un convexe fermé



$\mathbb{E}$ espace euclidien
de produit scalaire $\langle \cdot, \cdot \rangle$
norme associée $\|\cdot\| = \langle \cdot, \cdot \rangle^{1/2}$

si • C convexe, fermé, non vide de $\mathbb{E}$
• $x \in \mathbb{E}$
alors le problème $\min \{ \|y - x\| : y \in C \}$
a une solution unique

Dem $\inf_{y \in C} \left(\frac{1}{2} \|y - x\|^2 = \varphi(y) \right)$

• existence coercivité de φ

$$\frac{1}{2}\|y-x\|^2 = \frac{1}{2}\|y\|^2 - \underbrace{\langle x, y \rangle}_{\leq \|x\| \, \|y\|} + \frac{1}{2}\|x\|^2$$

$$\geq \underbrace{\underbrace{\|y\|}_{\to \infty} \underbrace{\left(\frac{1}{2}\|y\| - \|x\| \right)}_{\geq 1 \text{ si } \|y\| \text{ est grand}} + \frac{1}{2}\|x\|^2}_{\to \infty}$$

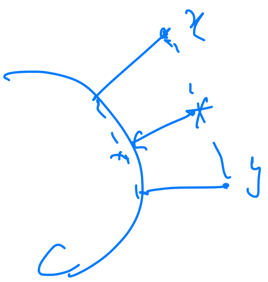
• unicité

$D^2 \varphi(x) = I$

$\to \varphi$ strict convexe
et C convexe

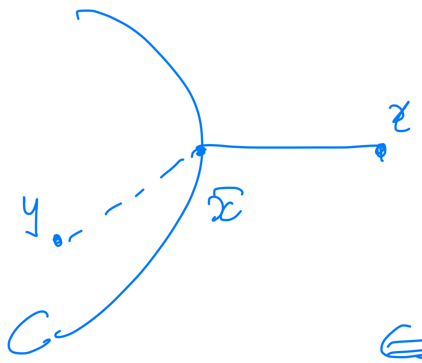
□

Dfn • la solution du problème est le projeté de x sur C
noté $P_C x$
• l'application non linéaire
$P_C : x \in \mathbb{E} \longmapsto P_C x$
est appelée le projecteur sur C
• la projection est l'opération consistant à projeter sur C



Caractérisation de $P_C x$

$$\boxed{\bar{x} = P_C x} \iff \boxed{\forall y \in C, \ \langle y - \bar{x}, \bar{x} - x \rangle \geq 0}$$



Dém:

$$\min_{y \in C} \left(\frac{1}{2} \|y - x\|^2 = \varphi(y) \right)$$

$\bar{x}$ est solution de ce pb

$$\iff \varphi'(\bar{x}; y - \bar{x}) \geq 0, \quad \forall y \in C$$

$$\langle \bar{x} - x, y - \bar{x} \rangle \geq 0, \quad \forall y \in C$$

$\square$

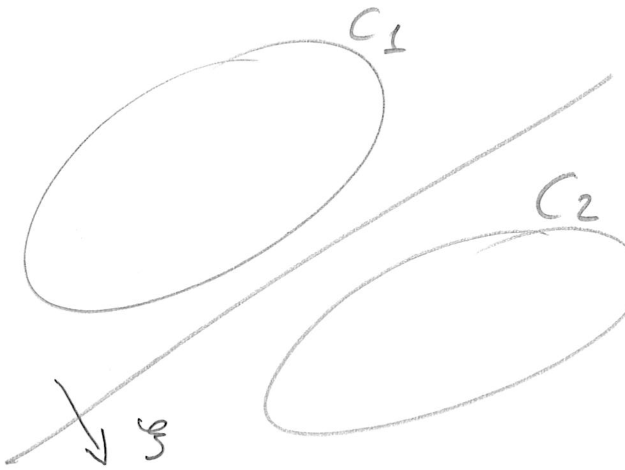
On a aussi

$$\boxed{\bar{x} = P_C x} \iff \boxed{\forall y \in C : \ \langle y - \bar{x}, y - x \rangle \geq 0}$$

$$\boxed{\forall y \in C : \ \langle \bar{x} - x, y - x \rangle \geq \|\bar{x} - x\|^2}$$

besoin de cette "sécurité" si $y - \bar{x}$ n'est pas dans le membre de gauche

B Séparation de deux convexes



$\mathbb{E}$ espace euclidien (produit scalaire $\langle \cdot , \cdot \rangle$)

C_1 et C_2 deux convexes disjoints de $\mathbb{E}$ ($C_1 \cap C_2 = \emptyset$) et non vides

- On cherche à séparer C_1 et C_2 par un hyperplan c-à-d à trouver $\xi \in \mathbb{E}$, non nul, tel que

$$\forall x_i \in C_i : \quad \langle \xi, x_1 \rangle \leq \langle \xi, x_2 \rangle$$

ou

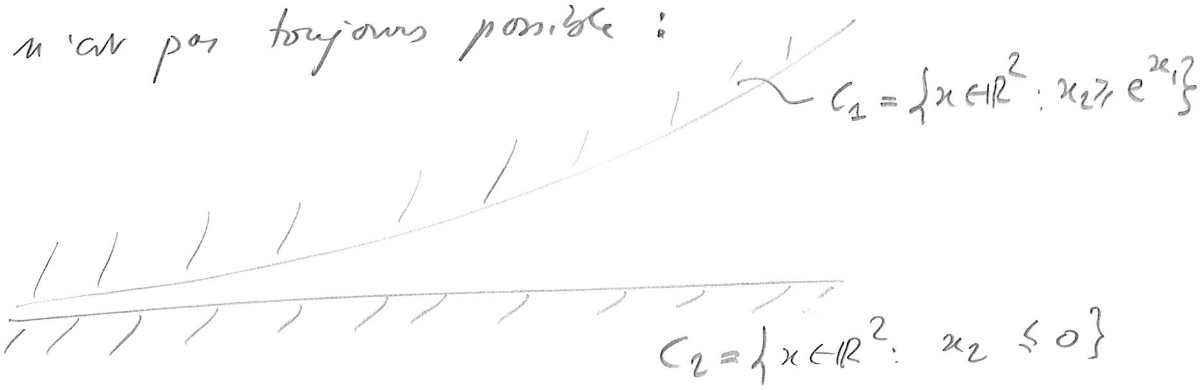
$$\boxed{\sup_{x_1 \in C_1} \langle \xi, x_1 \rangle \leq \inf_{x_2 \in C_2} \langle \xi, x_2 \rangle}$$

C'est toujours possible en dimension finie si C_1 et C_2 sont deux convexes disjoints (mais pas en dimension infinie)

- On dira que la séparation est stricte si

$$\boxed{\sup_{x_1 \in C_1} \langle \xi, x_1 \rangle < \inf_{x_2 \in C_2} \langle \xi, x_2 \rangle}$$

Ce n'est pas toujours possible :



Théo (HAHN - BANACH)

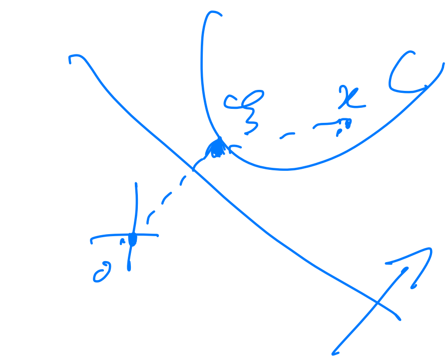
Si • C_1 et C_2 convexes, non vides, disjoints
• l'un est fermé, l'autre est compact

alors on peut séparer C_1 et C_2 strictement

Dém

1) Cas particulier $0 \notin C$ convexe fermé nonvide

$\xi := P_C(0)$



$\xi \neq 0$

$\langle x - \xi, \xi - 0 \rangle \geq 0, \quad \forall x \in C$

$\forall x \in C \quad \langle \xi, x \rangle \geq \|\xi\|^2 > 0 = \langle \xi, 0 \rangle$

2) Cas général $\quad C = C_2 - C_1 = \{ x_2 - x_1 : x_i \in C_i \}$

. C convexe : C_1, C_2 convexe

. C fermé : $C \ni y^k = x_2^k - x_1^k \longrightarrow y \overset{?}{\in} C$

⊗ $(x_1^k) \subset C_1$ compact $\Rightarrow \exists$ s.s. $(x_1^k) \to x_1 \in C_1$

⊗ $x_2^k = y^k + x_1^k \longrightarrow \underbrace{y + x_1}_{= x_2} \in C_2$ fermé

$y = x_2 - x_1 \in C$

. $C \neq \phi$ $\quad (C_1$ et $C_2 \neq \phi)$

. $0 \notin C$ $\quad (C_1 \cap C_2 = \phi)$

Cas 1 $\Rightarrow \exists\, \xi \neq 0$:

$$\langle \xi, x_2 - x_1 \rangle \geq \|\xi\|^2 > 0$$

$\forall x_i \in C_i$:

$$\langle \xi, x_1 \rangle + \|\xi\|^2 \leq \langle \xi, x_2 \rangle$$

$$\sup_{x_1 \in C_1} \langle \xi, x_1 \rangle \cancel{+ \underbrace{\|\xi\|^2}_{>0}} \leq \inf_{x_2 \in C_2} \langle \xi, x_2 \rangle$$

$\square$

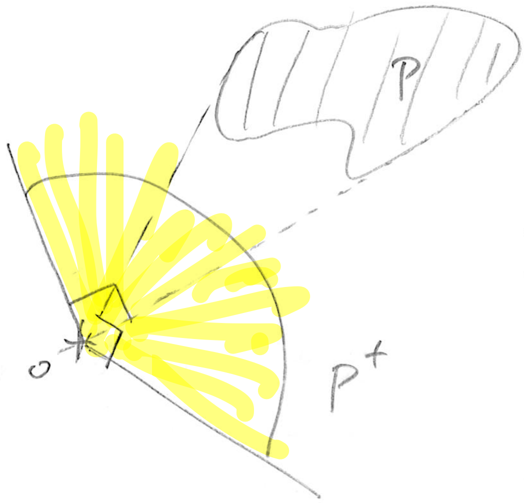
C Cône dual et lemme de Farkas

$\mathbb{E}$ e. euclidien (produit scalaire $\langle\cdot,\cdot\rangle$)

$P \subset \mathbb{E}$

Le cône dual de P est

$$\boxed{P^+ := \{x \in \mathbb{E} : \langle x, y\rangle \geqslant 0, \ \forall y \in P\}}$$



$$P^+ = \underbrace{\bigcap_{y \in P} \underbrace{\{x \in \mathbb{E} : \langle x, y\rangle \geqslant 0\}}_{\text{convexe fermé}}}_{\text{convexe fermé}}$$

$$\boxed{P^+ \text{ est un } \underline{\text{cône}} \ \underline{\text{convexe fermé}} \ \underline{\text{non vide}}}$$

motivation : On sait que $x_* \in \operatorname{argmin}\{f(x) : x \in X\}$ vérifie

$$\langle \nabla f(x_*), d\rangle = f'(x_*).d \geqslant 0, \quad \forall d \in T_{x_*}X$$

$$\Longleftrightarrow \quad \boxed{\nabla f(x_*) \in \underbrace{(T_{x_*}X)^+}}$$

pour des X particuliers, il va falloir calculer $(T_x X)^+$

Lemme de FARKAS (généralisé) (1902)

- $\mathbb{E}$ et $\mathbb{F}$ deux espaces vectoriels euclidiens
- $A : \mathbb{E} \to \mathbb{F}$ linéaire
- $A^* : \mathbb{F} \to \mathbb{E}$ adjointe de A défini par

$$\forall x \in \mathbb{E}, \forall y \in \mathbb{F} : \quad \langle Ax, y \rangle = \langle x, A^* y \rangle$$

- K un cône convexe non vide de $\mathbb{E}$ (pas nécessairement fermé)

$\mathbb{E} \xrightarrow{A} \mathbb{F}$

$K \qquad A(K)$?

$$\boxed{\overline{A(K)} = \{ y \in \mathbb{F} : A^* y \in K^+ \}^+}$$

toujours fermé

$A(K) = \{ Ax : x \in K \}$ n'est pas nécessairement fermé

Exemple 1 $\quad K = \mathbb{E} \quad P^+ := \{x : \langle y, x \rangle \geq 0 \ , \forall y \in P\}$

$$\Rightarrow \quad R(A) = \{ y \in \mathbb{F} : A^* y = 0 \}^+$$

1 s.e.v., donc $(\,)^+ = (\,)^\perp$

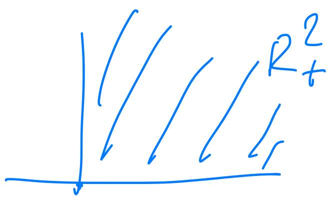
$$\Rightarrow \quad \boxed{R(A) = N(A^*)^\perp}$$

relation d'algèbre linéaire bien connue.

Exemple 2 $\quad \mathbb{E} = \mathbb{R}^n, \ K = \mathbb{R}^n_+ \quad (\Rightarrow K^+ = \mathbb{R}^n_+)$

alors $A(K) = \{ Ax : x \geq 0 \}$ est fermé (voir cours suivant)

$$\boxed{\{ Ax : x \geq 0 \} = \{ y \in \mathbb{F} : A^* y \geq 0 \}^+}$$



$$\mathbb{R}^n_+ = \{ x \in \mathbb{R}^n : x \geq 0 \}$$

orthant positif

$$(\mathbb{R}^n_+)^+ = \mathbb{R}^n_+$$

Dém

$$\overline{A(K)} = \underbrace{\{y \in F,\ A^* y \in K^+\}}_{D}{}^{+} \quad \text{fermée}$$

$\boxed{\subset}$ il suffit de montrer que $A(K) \subset D^+$

$x \in K \overset{?}{\Rightarrow} A(x) \overset{?}{\in} D^+$

Soit $y \in D \overset{?}{\Rightarrow} \langle Ax, d\rangle \geq 0$?

$\langle \underbrace{x}_{\in K}, \underbrace{A^* d}_{\in K^+}\rangle \geq 0$? ✓

$\boxed{\supset}$ (par l'absurde)

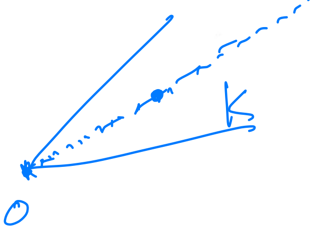
Si $D^+ \not\subset \overline{A(K)}$

$\Rightarrow \exists p \in D^+,\ p \notin \overline{A(K)}$

$\Rightarrow \underbrace{\{p\}}_{\text{convexe compact}} \cap \underbrace{\overline{A(K)}}_{\text{convexe fermé}} = \emptyset$

(HB) $\Rightarrow \exists y_0 \in F,\ \exists \alpha \in \mathbb{R},\ \underline{\forall x \in K}$

$\langle y_0, p\rangle < \alpha \leq \langle y_0, Ax\rangle$



- $x \to tx,\ t \downarrow 0$: $\alpha \leq t\langle y_0, Ax\rangle,\ t \downarrow 0$, $\alpha \leq 0$

$$\Rightarrow \langle y_0, p\rangle < 0$$

- $x \to tx,\ t \nearrow \infty$

$\frac{\alpha}{t} \leq \cancel{t} \langle y_0, Ax\rangle$, $\forall x \in K,\ \forall t \to \infty$

$$0 \leq \langle y_0, Ax \rangle, \quad \forall x \in K$$

$$0 \leq \langle A^* y_0, x \rangle, \quad \text{———}$$

$$A^* y_0 \in K^+$$

$$y_0 \in D$$

$$\langle y_0, p \rangle \geq 0 \quad (\text{as } p \in D^+)$$

contradiction

$\square$