§ 4. Constrained optimization

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t. } x \in C \qquad C \subseteq \mathbb{R}^n$$

constraints domain

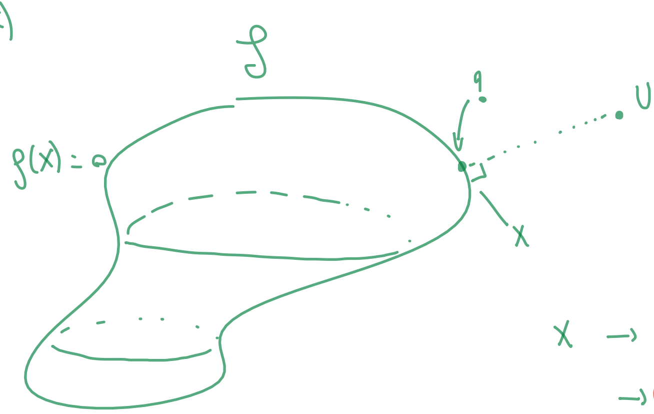
examples:

- surface $\mathcal{S}$ — implicit model $f(x) = 0$

  ex sphere $\underbrace{\|x - c\|^2 - r^2}_{f(x)} = 0$

  $U \in \mathbb{R}^3$

  projection of U on the surface



$x \to$ closest to U

$\to \in \mathcal{S}$

$$\left.\begin{array}{rcl} J : \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \|x - U\|^2 \end{array}\right) \quad \min_{x \in \cancel{\mathbb{R}^3}\ \mathcal{S}} ?$$

$\underbrace{\phantom{x\in\mathcal{S}}}$ modeled by $\Updownarrow$ $f(x) = 0$

$$\min_{x \in \mathcal{S}} J(x)$$

domain modeled by equation $\varphi(x) = 0$ or $\varphi(x) \leq 0$

with $\mathcal{S}$ described by

$$\mathcal{S} = \{ x \in \mathbb{R}^3 ;\ f(x) = 0 \}$$

- coca-cola problem § 1

Constrained optimization problem

$f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$\min_{x \in C} f(x)$ with $C \subseteq \mathbb{R}^n$ modeled by :

$C = \{ x \;;\; \varphi_1(x) = 0 \;\ldots\; \varphi_k(x) = 0$

and $\varphi_{k+1}(x) \leqslant 0 \;\ldots\; \varphi_m(x) \leqslant 0 \}$

$\varphi_j : \mathbb{R}^n \longrightarrow \mathbb{R}$

m constrains → k equality constraints
→ $m-k$ inequality constraints

ex :

→ projection
- → $m = 3$
- → $m = 1$ (1 equality constraint)

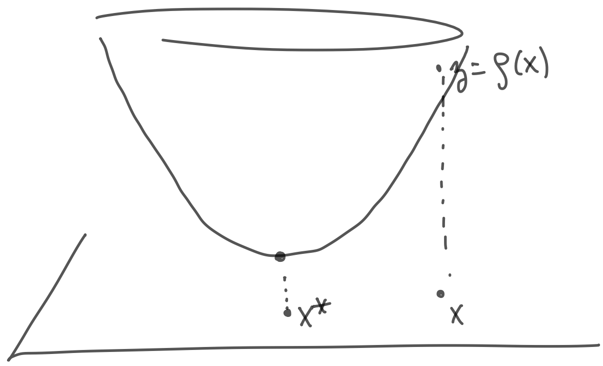
→ coca-cola
- → $m = 2$ ($x = (r, h)$)
- → $m = 1$ (1 equality constraint — volume $= V_0$)

How to compute this min ?

ex : $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

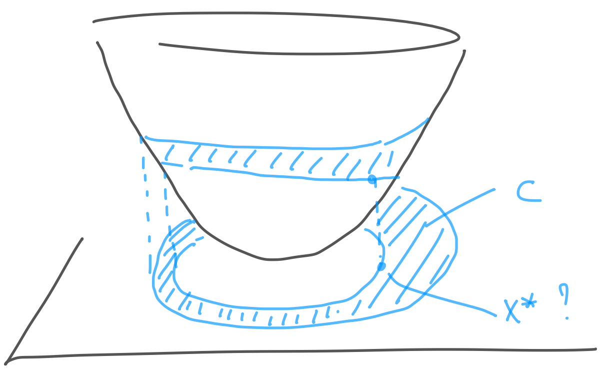
→ no constraints f C^1

x^* min $\underset{N.C.}{\Longrightarrow}$ $\nabla f(x^*) = \vec{0}$



→ with constraints

$\cancel{\nabla f(x^*) = \vec{0}}$



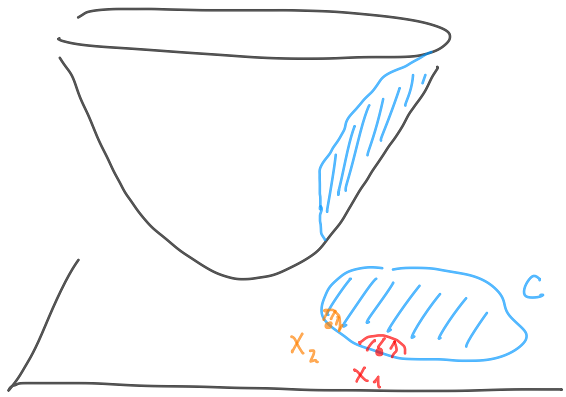
I. Domain and constraints equations

① How to decide if x^* is a local solution ?

x^* is a local solution of the constr. prob. if :

$\exists$ neigh. of x^* in C

$f(x^*) \leqslant f(x$



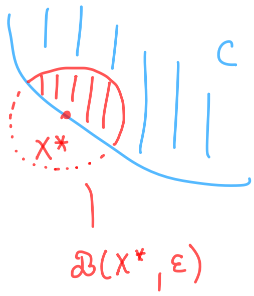
1) to be defined ...

2) simplify the idea

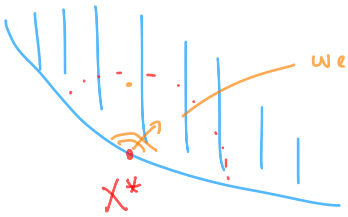
small ball centred at x^*

1) $\mathcal{B}(x^*) \cap C$

neighborhood of x^*



2) Admissible directions



we are interested in pts close from x^*

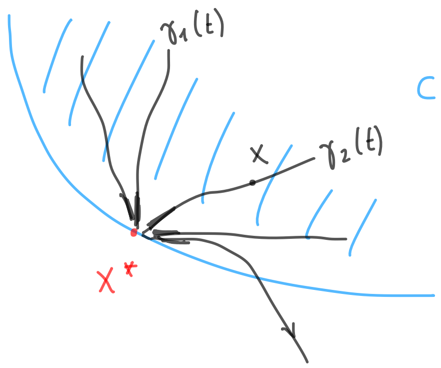
s.t. we are leaving x^* inside C

$\boxed{\text{Admissible curves at } x^*}$

curves starting from x^* and remaining inside C "for some time"

$\gamma(0) = x^*$

$\gamma(t) \in C$ with $t \leqslant \alpha$ small



$C \subseteq \mathbb{R}^n$

curves modeled by a parametric model :

$$\gamma : \mathbb{R}^+ \longrightarrow \mathbb{R}^n$$
$$t \longmapsto \gamma(t)$$

x^* constr. min if $\forall \gamma$ admissible curve at x^*

$f(x^*) \leq f(\gamma(t))$ for t small enough

v.o

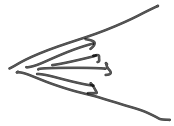
curves close to $t = 0$

$\Updownarrow$

comes to consider their tangents

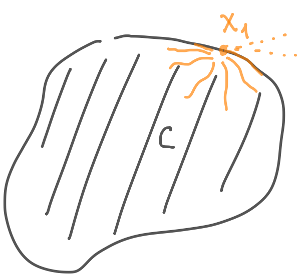
Cone of Admissible directions (at x^*) are tangents of admissible curves

It is a cone



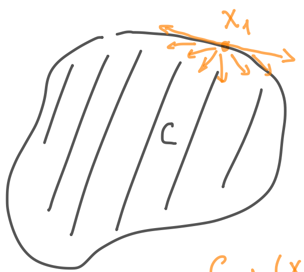
$\boxed{C_{ad}(x^*)}$

ex :

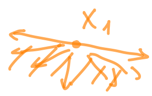


admissible curves

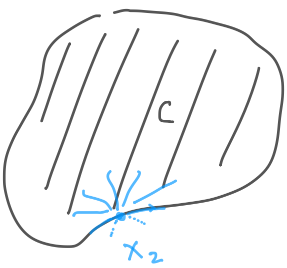
tangents



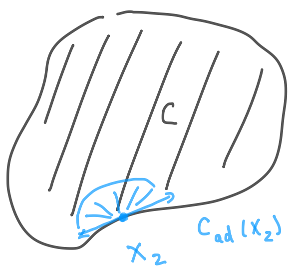
$C_{ad}(x_1)$:

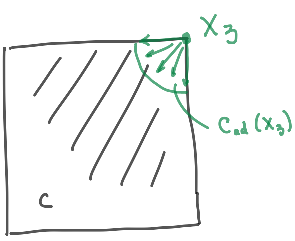


half-space



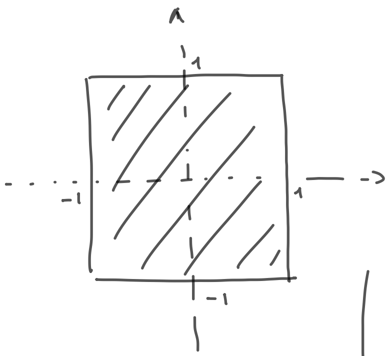
tangents





model

constraints equations



$|x| \leq 1$

$|y| \leq 1$

$$\begin{cases} \varphi_1(x,y) = |x| - 1 \ (\leq 0) \\ \varphi_2(\underbrace{x,y}_{X}) = |y| - 1 \ (\leq 0) \end{cases}$$

Summary

pb: $f : X^n \to \mathbb{R}$

$$\min_{x \in C} f(x) \quad \text{with} \quad C = \left\{ x \in \mathbb{R}^n ; \begin{array}{ll} \varphi_i(x) = 0 & i = 1 \dots k \\ \varphi_j(x) \leq 0 & j = k+1 \dots m \end{array} \right\}$$

1) x^* constrained min ? / solution of this pb ? $\longrightarrow$ decide it from the equations we have

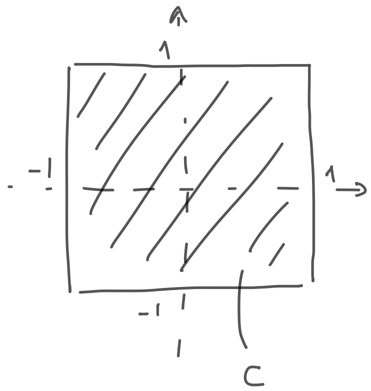
?

$\swarrow$ f $\qquad$ $\searrow$ φ_i / φ_j

We are going to see that:

$C_{ad}(x^*) \sim$ given by $\{ \nabla \varphi_i(x^*) \,/\, \nabla \varphi_j(x^*) \}$

We come back to



model 1 $\rightarrow$ 2 inequality constraints equations

$$\begin{cases} \varphi_1(x,y) = |x| - 1 \\ \varphi_2(x,y) = |y| - 1 \end{cases}$$

⊖ non differentiable

⊕ 2 constraints

model 2 $\rightarrow$ 4 inequalities $\quad (\leq 0)$

$x \leq 1$	$\varphi_1(x,y) = x - 1$
$x \geq -1$	$\varphi_2(x,y) = -x - 1$
$y \leq 1$	$\varphi_3(x,y) = y - 1$
$y \geq -1$	$\varphi_4(x,y) = -y - 1$

⊖ 4 constraints

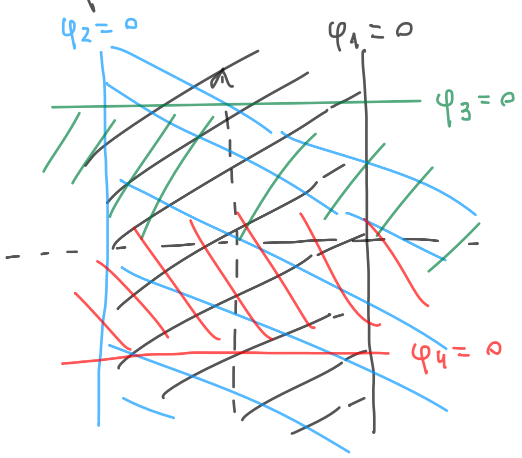
⊕ φ_i $\mathcal{C}^\infty$ linear

we consider this model $\leftarrow$

$\varphi_1(x) \leq 0$

$\downarrow$

half-plane

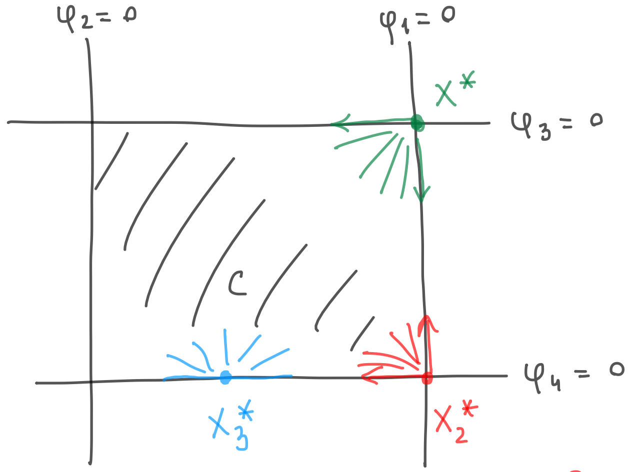


φ_2

φ_3

φ_4

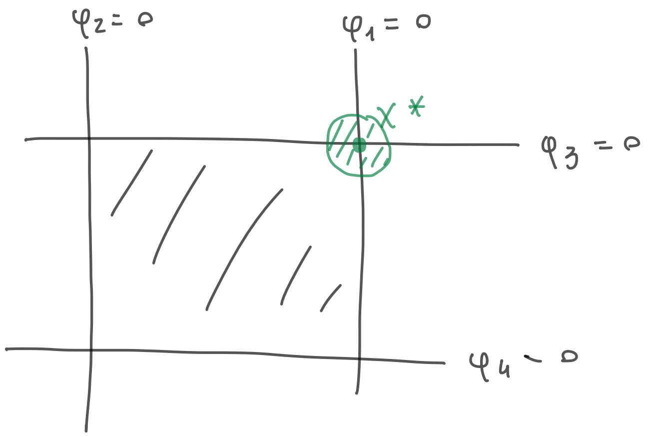
$C \sim$ intersection

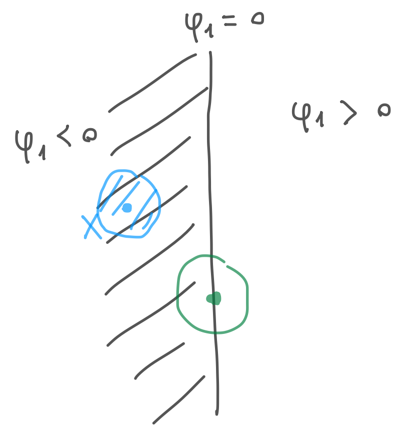


$C_{ad}(x^*)$? — tangents of φ_1 / φ_3

$C_{ad}(x^*)$ — tangents of φ_1 / φ_4

$C_{ad}(x^*)$ limited by tangent of φ_4





$\varphi_1(x^*) = 0$ ⟵

$\cancel{\varphi_2(x^*) < 0}$

$\varphi_3(x^*) = 0$ ⟵

$\cancel{\varphi_4(x^*) < 0}$

remain < 0 in a neighborhood

↓

discard them

$\varphi_1(x) < 0$

↓

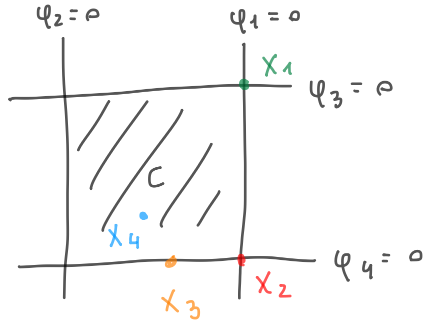
still true in a neighborhood

Constraints that are "critical" and may become satisfied / not satisfied around x^* are called

$\boxed{\text{saturated constraints}}$ $\longleftrightarrow$ $\underline{\underline{\varphi_i(x^*) = 0}}$

$\underline{I_0(x^*)}$: set of indices of saturated constraints.

ex!



$I_0(x_1) = \{1, 3\}$

$I_0(x_2) = \{1, 4\}$

$I_0(x_3) = \{4\}$

$I_0(x_4) = \emptyset$

Intuitively

$C_{ad}(x^*)$ $=$ limited by $\{ \nabla \varphi_i (x^*)$ with $i \in I_0(x^*) \}$

to get equality / prove

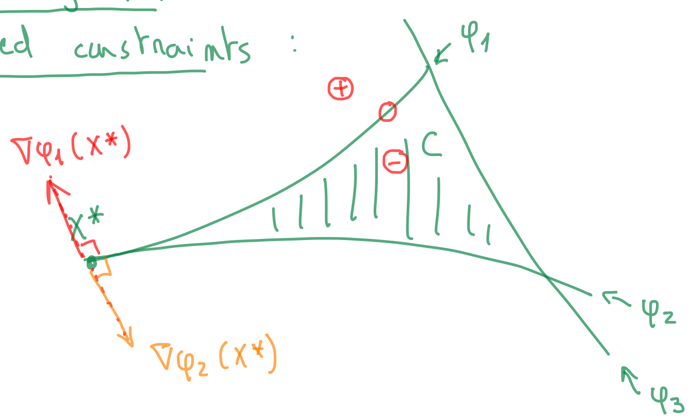
we need one more property on the modeling / φ_i functions

contraintes qualifiées

* def: we say that constraints are qualified in x^* if

$\{ \nabla \varphi_i (x^*) \; ; \; i \in I_0(x^*) \}$ is a free family

examples of non qualified constraints :



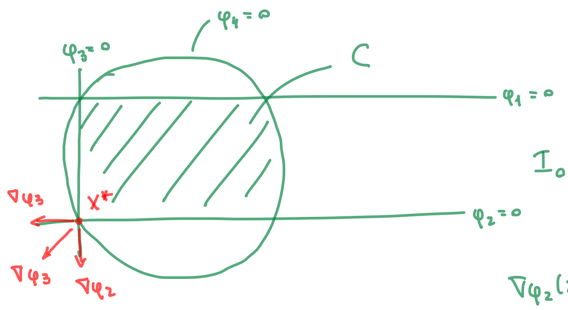
$I_0(x^*) = \{1, 2\}$

$\downarrow$

$\nabla \varphi_1 (x^*)$ / $\nabla \varphi_2 (x^*)$

// $\rightarrow$ not free !

⚠ constraints not qualified in x^*



$\mathbb{R}^2$

$$I_0(X^*) = \{2,3,4\}$$

$$\downarrow$$

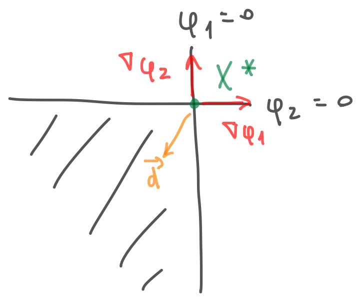
$$\underbrace{\nabla\varphi_2(X^*),\ \nabla\varphi_3(X^*),\ \nabla\varphi_4(X^*)}_{\text{3 vectors in dim 2}}$$

$\Downarrow$

not free !

Prop If constraints are qualified in X^*

$$C_{ad}(X^*) = \{ \vec{d} \ ;\ \forall i \in I_0(X^*)\ \ \nabla\varphi_i(X^*)\cdot\vec{d} \leq 0 \}$$

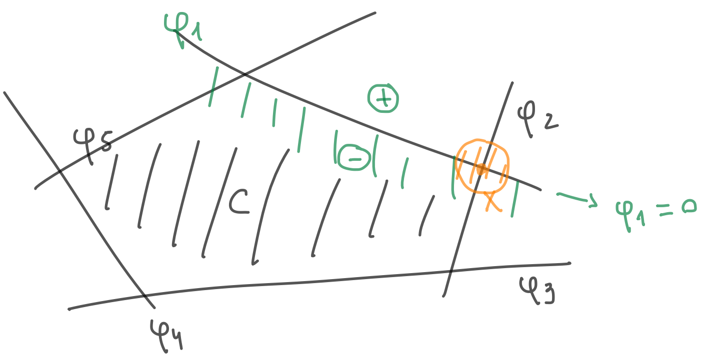


---

min f sur C

$\uparrow$

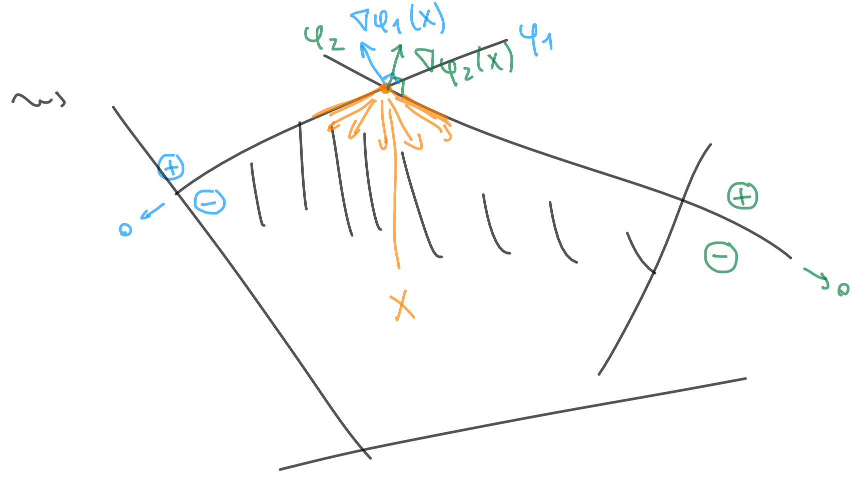
décrit par éq. φ_i $\nearrow \leq 0$ $\searrow = 0$



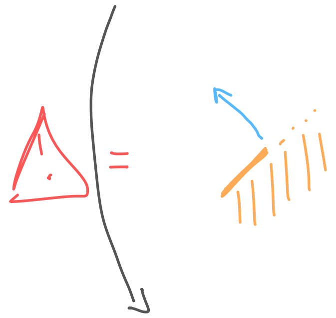
$\leadsto$ X $\left| \begin{array}{l} \left. \begin{array}{l} \varphi_1 = 0 \\ \varphi_2 = 0 \end{array} \right) \text{⚠ au vois. le signe change} \\ \left. \begin{array}{l} \varphi_3 \\ \vdots \\ \varphi_5 \end{array} \right) < 0 \quad \text{au vois. } < 0 \end{array} \right.$

Contraintes "importantes" en X : contraintes saturées

$$I_0(X) = \{1;2\}$$



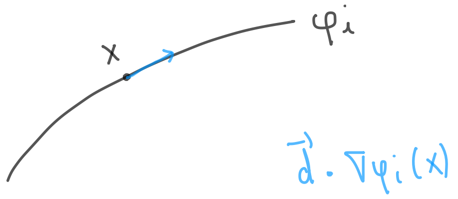
$C_{ad}(x)$



$$\left\{ \vec{d} \text{ tq } \forall i \in I_0(x) \quad \vec{d} \cdot \nabla\varphi_i(x) \leq 0 \right\}$$

Pour avoir égalité : les contraintes doivent être qualifiées en X $\longleftrightarrow$

$\{\nabla\varphi_i(x) \; ; \; i \in I_0(x)\}$ famille libre



$\vec{d} \cdot \nabla\varphi_i(x) = 0$

$\Updownarrow$

$\vec{d} \;//\;$ tangente

II. Théorème de Lagrange

⚠ Dédié aux pbs n'ayant que des contraintes égalité

$\cancel{\varphi_i \leq 0}$ $\quad\to\quad$ $\varphi_i = 0$

① Lagrangien

Pb : (*)

$\boxed{\begin{array}{l} f : \mathbb{R}^n \longrightarrow \mathbb{R} \\ \downarrow \\ \text{min sur } C \subseteq \mathbb{R}^n \\ \qquad\qquad \downarrow \\ \qquad\qquad \text{modélisé par } m \text{ contraintes égalité } (\varphi_i \;\; i : 1 \to m) \end{array}}$

dim n | m contraintes

$\xrightarrow{\text{Lagrange}}$

$\underbrace{\text{dim } n+m}_{\text{Lagrangien}}$ / " sans contraintes "

Lagrangien du pb (*) (optim sous contraintes) :

$$L : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}$$

$$(\underbrace{x_1 \dots x_n}_{X}, \underbrace{\lambda_1 \dots \lambda_m}_{}) \longmapsto f(X) + \sum_{i=1}^{m} \lambda_i \cdot \varphi_i(X)$$

X : inconnue de f...

m nouvelles variables : multiplicateurs de Lagrange

Théorème de Lagrange → CN de min sous contraintes

Si f et les φ_i sont $\mathcal{C}^2$

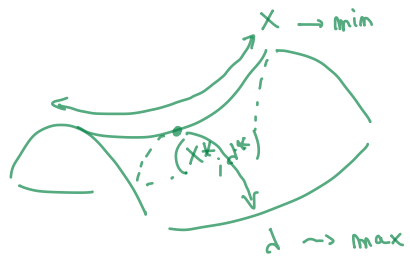
Si X^* est une solution du pb. contraint (*)

si les contraintes sont qualifiées en X^*

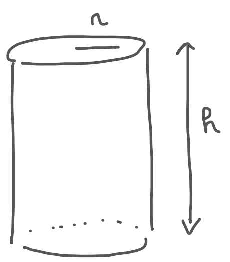
alors $\exists\, \lambda_1^* \dots \lambda_m^*$ tq $\nabla L(X^*, \lambda_1^* \dots \lambda_m^*) = \vec{0}$

intuitivement proche de l'optim sans contraintes de L

⚠ $(X^*, \lambda_1^* \dots \lambda_n^*)$ est un pt selle de L



III. Retour au pb coca-cola

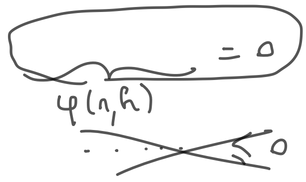


1) $n = 2$ $X = (r, h)$ — $\cancel{\mathbb{R}^2}$

2) Fonction à minimiser : surface

$$f(r,h) = 2\pi r^2 + 2\pi r h$$

3) Contrainte : volume $= V_0$ $\longleftrightarrow$

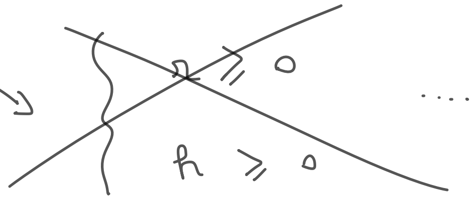


Restriction de $\mathbb{R}^2$ à certains (r,h) correspondant aux cylindres recherchés

$$\varphi(r,h) = \pi r^2 h - V_0 \quad (=0)$$

$\longrightarrow$ contrainte égalité

$m = 1$



⇝ si $r \leq 0$ et $h \leq 0$

$$f(r,h) = f(\underbrace{-r,-h}_{\geq 0})$$

↓

fonction symétrique ...

... ⇝ si r ou h sont négatifs

↓

on gère "à la main"

4) <u>Contraintes qualifiées</u> ?

X ⇝ contraintes saturées $(=0)$

↓

toutes (ici 1) car on n'a que des contr. égalité

⇝ $\{ \nabla\varphi(r,h) \}$ libre ? (1 vect)

⇕

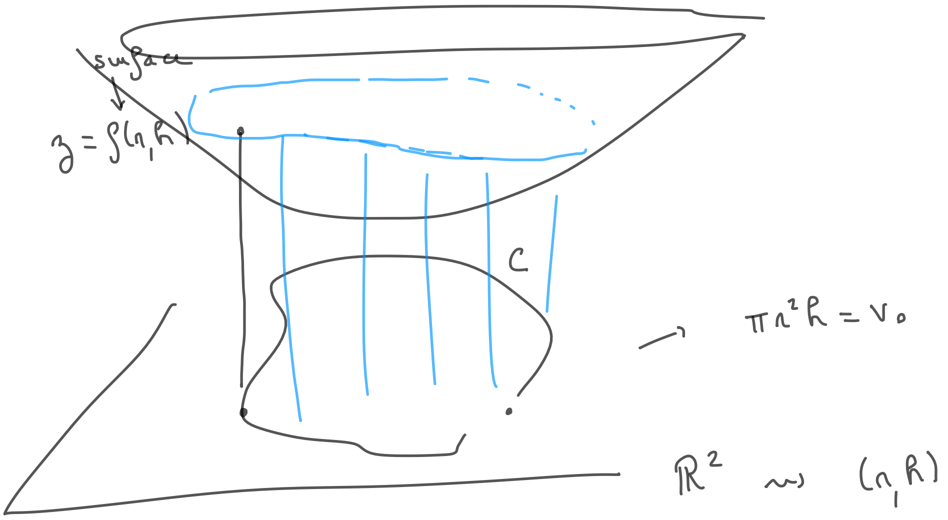
Libre si $\neq \vec{0}$

$$\nabla\varphi(r,h) = \begin{pmatrix} 2\pi r h \\ \pi r^2 \end{pmatrix} = \vec{0} \iff \begin{cases} h \text{ quelconque} \\ r = 0 \end{cases}$$

$\varphi(r,h) = \pi r^2 h - V_0$

$\nabla\varphi(r,h) = \vec{0}$ pour $\underbrace{(r,h) \text{ tq } r = 0}_{\notin C\ ?}$

cylindres de vol V_0



5) <u>Lagrangien</u> : $\quad f : \mathbb{R}^2 \longrightarrow \mathbb{R} \ / \ m = 1$ contrainte

$\partial / \partial \lambda \longrightarrow \varphi$

$$L : \mathbb{R}^{2+1} = \mathbb{R}^3 \longrightarrow \mathbb{R}$$

$$\underset{(r,h)}{\underset{\|}{(X}}, \lambda) \longmapsto f(X) + \boxed{\lambda \cdot \varphi(X)} = 2\pi r h + 2\pi r^2 + \lambda(\pi r^2 h - V_0)$$

6) $$\underset{(r,h,\lambda)}{\nabla L(X,\lambda)} = \begin{pmatrix} 2\pi h + 4\pi r + \lambda(2\pi r h) \\ 2\pi r + \lambda \cdot \pi r^2 \\ \pi r^2 h - V_0 \end{pmatrix} = \vec{0}$$

$$\Leftrightarrow \begin{cases} \cancel{2\pi} h + \overset{2}{\cancel{4\pi}} r + \lambda(\cancel{2\pi} r h) = 0 \\ 2\cancel{\pi r} + \lambda \cancel{\pi r^2} = 0 \quad (r \neq 0 \Leftrightarrow \mathrm{Vol} = \underset{\neq 0}{V_0}) \\ \pi r^2 h = V_0 \end{cases}$$

$$\Leftrightarrow \begin{cases} h + 2r + \lambda r h = 0 \\ 2 + \lambda r = 0 \longrightarrow \lambda = -\frac{2}{r} \\ \pi r^2 h = V_0 \end{cases}$$

$\boxed{\text{Système non linéaire } n+m \text{ eq / } n+m \text{ inconnues}}$

$\downarrow$

Newton–Raphson sur $\nabla L : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^{n+m}$

$$\iff \begin{cases} \alpha = -\frac{2}{r} \\ \cancel{h} + 2r - \frac{2}{\cancel{r}} \cancel{r} h = 0 \rightsquigarrow h = 2r \\ \pi r^2 h = V_0 \end{cases}$$

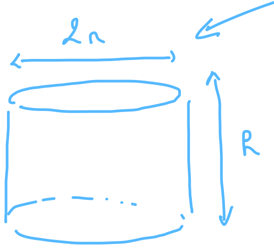
$\cancel{2h}$

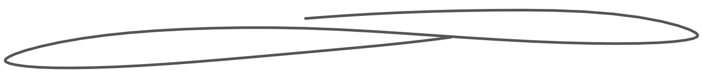
$-h$

$$2\pi r^3 = V_0$$

$$\begin{cases} r = \sqrt[3]{\frac{V_0}{2\pi}} \\ h = 2r \end{cases}$$

$$\alpha = -\frac{2}{r}$$





$\nabla L = \vec{0} \longrightarrow$ Newton $\longrightarrow$ zéros de fct

$g : \mathbb{R}^n \longrightarrow \mathbb{R}^n$

Newton$(g, J(g))$

matrice $\sim g'$

Newton$(\nabla L, H(L))$

g $\quad J(g)$ $\quad$ "g'"

Ici