

# Cours n°0: De l'optimisation sans contrainte à l'optimisation sous contrainte dans $\mathbb{R}^n$

Nicolas DUPIN

<http://nicolasdupin2000.wixsite.com/research>

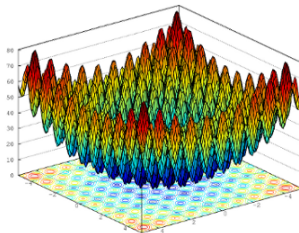
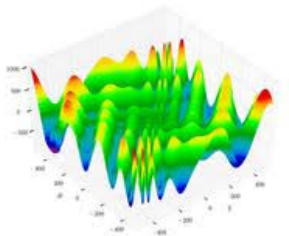
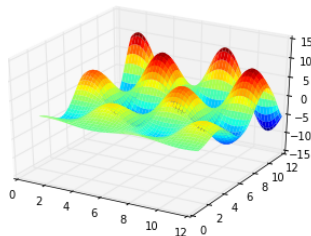
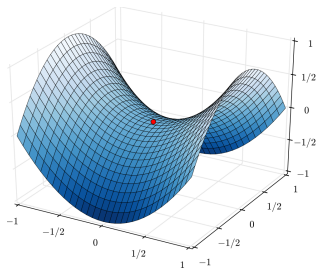
version du 20 mars 2022

Cours distribué sous licence CC-BY-NC-SA

Issu et étendu d'enseignements donnés à l'ENSTA Paris, Polytech' Paris-Saclay, et à l'université Paris-Saclay

# De l'optimisation continue en dimension 1 à la dimension $n$

- ▶ Optimisation de la variable réelle : dimension 1, une unique décision à optimiser. On a naturellement de l'optimisation d'une fonction à valeurs réelle avec plusieurs variables réelles.
- ▶ Quelles propriétés de l'optimisation en dimension 1 se généralisent en dimension supérieure ?
- ▶ Formalisation : soit  $K \subset \mathbb{R}^n$  et  $J : K \longrightarrow \mathbb{R}$ . Comment minimiser/maximiser  $J$  ?
- ▶ Aspects algorithmiques : Si on a les propriétés d'existence, comment avoir des procédés constructifs permettant une implémentation efficace ?



⇒ Même en dimension 2, on peut oublier la notion de monotonie et les tableaux de variation !!

⇒ Que peut on généraliser en dimension  $n$ ?

# Plan

Rappels de topologie et analyse en dimension  $n$

Optimisation continue sans contrainte en dimension  $n$

Optimisation discrète sans contrainte en dimension  $n$  ?

Optimisation sous contraintes, cadre général

Optimisation continue sous contrainte

Algorithmes numériques, Algorithmes de gradient

Conclusions et perspectives

# Plan

Rappels de topologie et analyse en dimension  $n$

Optimisation continue sans contrainte en dimension  $n$

Optimisation discrète sans contrainte en dimension  $n$  ?

Optimisation sous contraintes, cadre général

Optimisation continue sous contrainte

Algorithmes numériques, Algorithmes de gradient

Conclusions et perspectives

# Normes, distances dans $\mathbb{R}^n$

Soit  $\mathbb{R}^n$  a une structure de  $\mathbb{R}$ - espace vectoriel de dimension  $n$ .

C'est même un espace vectoriel normé, en définissant les normes :

$$||x||_1 = \sum_{i=1}^n |x_i|$$

$$||x||_2 = \sqrt{\sum_{i=1}^n x_i^2} \text{ (norme Euclidienne)}$$

$$||x||_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \text{ (norme de Minkowski)}$$

$$||x||_\infty = \max_{i \in [1, n]} |x_i|$$

En dimension finie, théorème d'équivalence des normes, les résultats de ce cours sont indépendants de la norme choisie. On peut choisir la norme Euclidienne par exemple.

distance associée :  $d : (x_1, \dots, x_n), (y_1, \dots, y_n) \in (\mathbb{R}^n)^2 \mapsto ||x - y||$

# Boules ouvertes, boules fermées dans $\mathbb{R}^n$

On définit à partir de la distance une structure topologique de *boule fermée* et de *boule ouverte*.

Soit  $x_0 \in \mathbb{R}^n$  et  $r > 0$ .

$$\text{BO}(x_0, r) = \{x \in \mathbb{R}^n \mid d(x, x_0) < r\}$$

Boule ouverte de centre  $x_0$  et de rayon  $r$ . Généralise  $]x_0 - r, x_0 + r[$  en dimension 1

$$\text{BF}(x_0, r) = \{x \in \mathbb{R}^n \mid d(x, x_0) \leq r\}$$

Boule fermée de centre  $x_0$  et de rayon  $r$ . Généralise  $[x_0 - r, x_0 + r]$  en dimension 1

# Intérieur et ouverts de $\mathbb{R}^n$

Soit  $A \subset \mathbb{R}^n$ .

On dit que  $a$  est un *point intérieur* de  $A$  si  $a \in A$  et s'il existe  $\varepsilon > 0$  tel que  $\text{BO}(a, \varepsilon) \subset A$ .

L'*intérieur* de  $A$ , noté  $\mathring{A}$  est l'ensemble des points intérieurs de  $A$ .  
On a toujours  $\mathring{A} \subset A$ .

$A$  est *ouvert* de  $\mathbb{R}^n$  si et seulement si  $\mathring{A} = A$ . Exemples :  $\text{BO}(a, \varepsilon)$ ,  $\{x \in \mathbb{R}^n \mid \forall i \in \llbracket 1; n \rrbracket, x_i > 0\}$ ,  $\mathbb{R}^n$ .

Une intersection d'un ensemble fini d'ouverts est un ouvert.

Une union quelconque d'ouverts est un ouvert.



# Adhérence, fermeture et fermés

Soit  $A \subset \mathbb{R}^n$ .

On dit que  $a$  est un *point d'adhérence* de  $A$  si on peut toujours trouver des points de  $A$  aussi proches que possibles de  $a$  :

$\forall \varepsilon > 0, \exists a' \in A \cap \text{BO}(a, \varepsilon)$ .

La *fermeture* de  $A$ , noté  $\overline{A}$  est l'ensemble des points d'adhérence de  $A$ . On a toujours  $A \subset \overline{A}$ .

$A$  est *fermé* de  $\mathbb{R}$  si et seulement si  $\overline{A} = A$ . Exemples :  $\text{BF}(a, \varepsilon)$ ,  $\{x \in \mathbb{R}^n \mid \forall i \in \llbracket 1; n \rrbracket, x_i > 0\}$ ,  $\mathbb{R}^n$ .

Un ensemble compact  $A \subset \mathbb{R}^n$  se définit comme un ensemble fermé et borné de  $\mathbb{R}^n$ .

Une union d'un ensemble fini de fermés est un fermé.

Une intersection quelconque de fermés est un fermé.

# Définitions de limite d'une suite de $\mathbb{R}^d$

Soit  $(u_n)_{n \in \mathbb{N}}$  une suite de  $\mathbb{R}^d$ , aussi noté  $u \in \mathcal{F}(\mathbb{N}, \mathbb{R}^d)$ . On dit que la suite  $u$  est convergente vers  $l \in \mathbb{R}^d$ , noté  $\lim_{n \rightarrow +\infty} u_n = l$  si et seulement si :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, \|u_n - l\| \leq \varepsilon$$

Dans les autres cas,  $u$  est dite divergente.

Dans  $\mathbb{R}^d$ , pas de notion intéressante de suite divergeant vers l'infini, on peut se ramener à la dimension 1 et utiliser  $\lim_{n \rightarrow +\infty} \|u_n\| = +\infty$

# Caractérisation séquentielle des fermés et des compacts

$f$  est point d'adhérence de  $F \subset \mathbb{R}^d$  si et seulement si il existe  $(u_n) \in \mathcal{F}(\mathbb{N}, F)$  une suite d'éléments de  $F$  avec  $\lim_{n \rightarrow +\infty} u_n = f$ .

$K$  est un compact de  $\mathbb{R}^d$  si et seulement si de toute suite d'éléments de  $K$ , on peut en extraire une sous-suite convergente.  
(Théorème de Bolzano-Weierstrass)

$O$  est un ouvert de  $\mathbb{R}^d$  si et seulement  $\mathbb{R} \setminus O$  est un fermé, la caractérisation séquentielle d'un fermé s'applique au complémentaire  $\mathbb{R} \setminus O$

# Définitions de limite d'une fonction en un point

Soit  $A \subset \mathbb{R}^d$ . Soit  $f : A \rightarrow \mathbb{R}$  une application, soit  $a \in \overline{A}$ .

On définit  $\lim_{x \rightarrow a} f(x) = l \in \mathbb{R}$  par :

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \text{BF}(a, \alpha), |f(x) - l| \leq \varepsilon$$

Caractérisation séquentielle des limites :  $\lim_{x \rightarrow a} f(x) = l$  si et seulement si pour toute suite  $(u_n)_{n \in \mathbb{N}}$  de  $\mathbb{R}^d$  à valeur dans  $A$  telle que  $\lim_{n \rightarrow +\infty} u_n = a$ , on a  $\lim_{n \rightarrow +\infty} f(u_n) = l$

Dans le cas contraire, pour une suite divergente, on n'a pas de définition bien générale pour une limite infinie.

Propriété intéressante qui peut être suffisante pour l'application de résultats : quand pour toute suite  $u_n$  telle que  $\lim_{n \rightarrow +\infty} u_n = +\infty$ , on a  $\lim_{n \rightarrow +\infty} |f(u_n)| = +\infty$

# Fonctions continues de $\mathbb{R}^d$ dans $\mathbb{R}$

Soit  $f : A \subset \mathbb{R}^d \rightarrow \mathbb{R}$ .

$f$  est continue en  $a \in A$  si on a  $\lim_{x \rightarrow a} f(x) = l$

$f$  est continue sur  $A$  si pour tout  $a \in A$ ,  $f$  est continue en  $a$ .

# Dérivées pour des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$

Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ .

Il y a plusieurs définitions pour la différentiabilité de fonctions, différentiabilité de Fréchet, Gateaux ...

Pour la suite, on utilisera juste des cas basiques de  $\mathbb{R}^n$  dans  $\mathbb{R}$ , très réguliers.

On voit  $f$  comme une fonction de  $n$  variables  $f(x_1, x_2, \dots, x_n)$

Pour un  $i$  fixé, si en fixant les  $n - 1$  autres variables à des valeurs atteignables dans  $A$ , on a une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$  dérivable, on dit que  $f$  admet une *dérivée partielle* selon la variable  $i$ , notée  $x \mapsto \frac{\partial f}{\partial x_i}(x_1, x_2, \dots, x_n)$ .

Si  $f$  admet toutes ses dérivées partielles, on appelle *gradient* de  $f$  la fonction vectorielle :

$$\nabla f : x \in A \mapsto \nabla f(x) = \left( \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_i}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) \in \mathbb{R}^n$$

$f$  est de classe  $\mathcal{C}^1$  si toutes les dérivées partielles sont continues.

# Dérivées d'ordre 2

$f$  est de classe  $\mathcal{C}^2$  si toutes les dérivées partielles sont dérivables par rapport à toutes les variables.

On note  $x \mapsto \frac{\partial^2 f}{\partial x_i^2}(x_1, x_2, \dots, x_n)$  la fonction obtenue en dérivant deux fois par rapport à la variable  $x_i$  en supposant les autres variables constantes.

On note  $x \mapsto \frac{\partial^2 f}{\partial x_i \partial x_j}(x_1, x_2, \dots, x_n)$  la fonction obtenue en dérivant d'abord par rapport à la variable  $x_i$  en supposant les autres variables constantes, puis selon la variable  $x_j$ .

Théorème de Schwartz :  $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$

Pour tout  $a \in A$ , on associe la matrice Hessienne  $\mathcal{H}_f(a) = \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{1 \leq i, j \leq n}$ , cette matrice est symétrique d'après le thm de Schwartz.

# Développements limités

Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  admet un gradient en un point  $a \in A$ .

Alors on peut écrire un développement limité du premier ordre pour tout  $h = (h_1, \dots, h_i, \dots, h_n)$  :

$$f(a+h) = f(a) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a) + o(\|h\|) = f(a) + \langle h, \nabla f(a) \rangle + o(\|h\|)$$

Si  $f$  est de classe  $\mathcal{C}^2$ ,

$$f(a+h) = f(a) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a) + \sum_{i=1}^n h_i^2 \frac{\partial^2 f}{\partial x_i^2}(a) + \sum_{i=1}^n \sum_{j=i+1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + o(\|h\|^2)$$



# Ensembles convexes et connexes de $\mathbb{R}^n$

Soit  $A \subset \mathbb{R}^n$ .

$A$  est un ensemble convexe de  $\mathbb{R}^n$  si pour tout couple de points de  $(a_1, a_2) \in A^2$ , le segment reliant  $a_1$  à  $a_2$  reste inclus dans  $A$  :

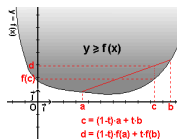
$$\forall (a_1, a_2) \in A^2, \forall \lambda \in [0, 1], \quad \lambda a_1 + (1 - \lambda)a_2 \in A$$

$A$  est connexe par arcs si pour tout couple de points de  $(a_1, a_2) \in A^2$ , on peut construire un chemin continu reliant  $a_1$  à  $a_2$  en en traversant que des points de  $A$  :

$A$  connexe par arcs  $\implies A$  convexe  $\implies A$  fermé.

N.B : La connexité formalise "d'un seul tenant", qui se généralise en informatique sur des structures discrètes, par exemple connexité d'un graphe.

# Fonctions convexes de $\mathbb{R}^n$ dans $\mathbb{R}$

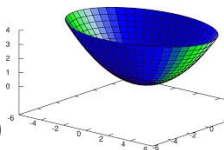


Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ .

$f$  est convexe si :

$$\forall (x, x') \in A^2, \forall \lambda \in [0; 1],$$

$$f(\lambda x + (1-\lambda)x') \leq \lambda f(x) + (1-\lambda)f(x')$$



$f$  est convexe ssi  $\{(x, y) \in A \times \mathbb{R} \mid f(x) \leq y\}$  est un convexe de  $\mathbb{R}^{n+1}$

$f$  est strictement convexe si on a de plus pour tout  $x \neq x' \in A$  :

$$\forall x, x', \forall \lambda \in ]0; 1[, f(\lambda x + (1-\lambda)x') < \lambda f(x) + (1-\lambda)f(x')$$

$f$  est dite  $\alpha$ -convexe pour  $\alpha > 0$  si pour tout  $(x, x') \in A^2$  et  $\lambda \in [0; 1]$ ,

$$f(\lambda x + (1-\lambda)x') \leq \lambda f(x) + (1-\lambda)f(x') - \frac{\alpha}{2} \lambda(1-\lambda) \|x - x'\|^2$$

$f$  est fortement convexe s'il existe  $\alpha > 0$  tq  $f$  est  $\alpha$ -convexe.

$f$  est  $\alpha$ -convexe  $\implies f$  est strictement convexe  $\implies f$  est convexe.

# Plan

Rappels de topologie et analyse en dimension  $n$

Optimisation continue sans contrainte en dimension  $n$

Optimisation discrète sans contrainte en dimension  $n$  ?

Optimisation sous contraintes, cadre général

Optimisation continue sous contrainte

Algorithmes numériques, Algorithmes de gradient

Conclusions et perspectives

# Minimum/maximum local/global

Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . On dit que

$x_0$  est un minimum local si

$$\exists \varepsilon > 0, \forall x \in BO(x_0, \varepsilon) \cap A, f(x_0) \leq f(x)$$

$x_0$  est un maximum local si

$$\exists \varepsilon > 0 \forall x \in BO(x_0, \varepsilon) \cap A, f(x_0) \geq f(x)$$

$x_0$  est un minimum global si

$$\forall x \in A, f(x_0) \leq f(x)$$

$x_0$  est un maximum global si

$$\forall x \in A, f(x_0) \geq f(x)$$

# Quelques résultats d'existence (1)

## Théorème

Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue. Si  $A$  est un compact (ie fermé et borné),  $f$  est bornée et atteint ses bornes (minimum global et maximum global).  $\min_{x \in A} f(x)$  et  $\max_{x \in A} f(x)$  sont bien définis.

Preuve : construction de suite approchant le minimum par propriété de borne inférieure, puis utilisation du théorème de Bolzano Weierstrass pour avoir une sous suite convergente, qui convergera alors vers un minimum global.

Exemple 1 :  $A = [0, 1]^n$  : hypercube de dimension  $n$

Exemple 2 :  $A = [a_1, b_1] \times [a_1, b_1] \times \cdots \times [a_n, b_n]$

## Quelques résultats d'existence (2)

### Théorème

Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue minorée telle que pour toute suite  $(u_n)_{n \in \mathbb{N}}$  de  $\mathbb{R}^n$  avec  $\lim_{n \rightarrow +\infty} \|u_n\| = +\infty$ , on a  $\lim_{n \rightarrow +\infty} f(u_n) = +\infty$ .

Alors  $f$  admet un minimum,  $\min_{x \in \mathbb{R}^n} f(x)$  est bien défini.

### Théorème

Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue périodique, ie telle qu'il existe  $T > 0$  tel que pour tout  $x \in \mathbb{R}^n$ ,  $f(x + T) = f(x)$ . Alors  $f$  est bornée et atteint ses bornes,  $\min_{x \in \mathbb{R}^n} f(x)$  et  $\max_{x \in \mathbb{R}^n} f(x)$  sont bien définis.

# Extremums locaux et gradient

Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$ .

Si  $a \in \mathring{A}$ , est un extremum local, cela impose

$$\forall i \in \llbracket 1, n \rrbracket, \frac{\partial f}{\partial x_i}(a) = 0$$

autrement dit,  $\nabla_f(a) = 0$

Sinon, des extremums locaux peuvent se trouver sur le bord (ou frontière) de  $A$

Généralise la dimension 1, avec  $f'(x) = 0$  pour un extremum local en un point intérieur  $x$ .

# Extremums locaux et matrice Hessienne

Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$ .

Pour tout  $a \in A$ , la matrice Hessienne  $\mathcal{H}_f(a) = \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)$ , matrice symétrique.

Soit  $a \in \mathring{A}$  un point tels que  $\nabla_f(a) = 0$ .

Si  $\mathcal{H}_f(a)$  est définie positive,  $f$  atteint un minimum local strict en  $a$  ;

Si  $-\mathcal{H}_f(a)$  est définie positive,  $f$  atteint un maximum local strict en  $a$  ;

Si  $\mathcal{H}_f(a)$  admet au moins une valeur propre strictement positive et une valeur propre strictement négative,  $a$  n'est pas un extremum local (c'est un point col).

Généralise la dimension 2, sur le signe de  $f''(x) = 0$  pour un extremum local en un point intérieur  $x$ . En pratique, ces caractérisations analytiques sont peu exploitables en calcul numérique.



# Matrices symétriques (définies) positives ?

Soit  $(M_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket^2} \in \mathcal{M}_n(\mathbb{R})$  une matrice carrée réelle.

La matrice  $M$  est dite symétrique si  $M_{i,j} = M_{j,i}$  pour tout  $i,j$ , i.e.  $M = M^\top$

Une matrice symétrique  $M$  est positive si toutes ses valeurs propres sont positives.

Une matrice symétrique  $M$  est définie positive si toutes ses valeurs propres sont strictement positives.

Une matrice symétrique  $M$  est (définie) négative si  $-M$  est (définie) positive.

Si une matrice symétrique  $M$  a une valeur propre strictement positive et une valeur propre strictement négative,  $M$  n'est ni symétrique positive, ni symétrique négative.

N.B : plusieurs définitions équivalentes, celles-ci permettent le plus facilement des calculs analytiques

# Attention à la réciproque!!!!

Dans le cas où il existe une valeur propre nulle, tout est possible.

Un minimum local,  $f(x_1, x_2) = x_1^4 + x_2^4$

Un minimum local,  $f(x_1, x_2) = -x_1^4 - x_2^4$

Un point selle,  $f(x_1, x_2) = x_1^4 - x_2^4$

Dans ces trois cas, on a une Hessienne nulle en  $(0, 0)$

La Hessienne donne des conditions suffisantes, mais pas nécessaires. Ce sont les conditions d'optimalité d'ordre 2, qui affinent les conditions d'optimalité d'ordre 1.

Remarque : il est équivalent :

$M$  non inversible  $\iff \det(M) = 0 \iff M$  admet 0 comme valeur propre

# Matrice Hessienne, cas particulier d'une fonction séparable

Soit  $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^2$  et convexes.

On considère  $g : (x_1, \dots, x_n) \in \mathbb{R}^n \mapsto f_1(x_1) + \dots + f_n(x_n) = \sum_{i=1}^n f_i(x_i)$

Pour tout  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ , la matrice Hessienne  $\mathcal{H}_g(x)$  est une matrice diagonale,  $\mathcal{H}_g(x) = \text{diag}(f_1''(x_1), \dots, f_n''(x_n))$

Si on a pour tous  $x$  et  $i$ ,  $f_i''(x_i) > 0$  alors la matrice Hessienne est définie positive pour tout  $x$  et  $g$  est strictement convexe.

# Calculs de valeurs propres ?

Soit  $(M_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket^2} \in \mathcal{M}_n(\mathbb{R})$  une matrice carrée réelle. Comment calculer les valeurs propres de  $M$  ?

On peut utiliser Python numpy ou toute autre bibliothèque d'algèbre linéaire.

Méthode algorithmique :

- ▶ on calcule le polynôme de degré  $n$   $P(X) = \det(M - XI_n)$
- ▶ les valeurs propres sont les racines de  $P$ , savoir résoudre  $P(x) = 0$ .

A la main, un calcul de déterminant est facile avec  $n = 2$ ,  $n = 3$ , ça peut être plus laborieux en général, idem pour la résolution de  $P(x) = 0$ .

Pour dire qu'une matrice est (définie) négative ou positive, doit on obligatoirement calculer les valeurs propres ?

Remarque : On voit que numériquement on pourrait s'arrêter avant le calcul à une bonne précision des valeurs propres pour prouver uniquement que le signe des valeurs propres est constant ou que ce n'est pas le cas.

# Critère de Sylvester

Soit  $(M_{i,j})_{(i,j) \in \llbracket 1, n \rrbracket^2} \in \mathcal{M}_n(\mathbb{R})$  une matrice carrée réelle symétrique.

$M$  est définie positive si et seulement si les  $n$  sous matrices  $(M_{i,j})_{(i,j) \in \llbracket 1, p \rrbracket^2}$  pour  $p \in \llbracket 1, n \rrbracket$  ont un déterminant strictement positif

Critère algorithmique plus facile à vérifier, à la main ou en implémentation.

# Extremums locaux et matrice Hessienne en dimension 2

Soit  $M$  une matrice réelle symétrique  $2 \times 2$ ,  $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ .

En appliquant la règle de Sylvester :

- ▶  $M$  est définie positive ssi  $a > 0$  et  $ac - b^2 > 0$
- ▶  $M$  est définie négative ssi  $a < 0$  et  $ac - b^2 > 0$

Les autres cas possibles :

- ▶  $ac - b^2 = 0 \Leftrightarrow$  il y a une valeur propre nulle, on ne peut rien dire.
- ▶  $ac - b^2 < 0$ , permet d'affirmer qu'on a deux valeurs propres de signe différentes.
- ▶  $ac - b^2 > 0$  et  $a = 0$  impliquant  $-b^2 > 0$ , ce n'est pas possible.

$\Rightarrow$  On peut déterminer très facilement si une Hessienne d'une fonction de deux variables est définie positive ou définie négative, ou si on a deux valeurs propres de signe différentes.

# Utilisation du critère de Sylvester pour une Hessienne

Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , soit  $(x, y) \in \mathbb{R}^2$ .

On note  $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = \mathcal{H}_f(x, y)$  la matrice Hessienne de  $f$  en  $(x, y)$ .

- Si  $(x, y)$  est un point critique de  $f$  et si  $M$  est définie positive alors  $(x, y)$  est un minimum local.
- Si  $(x, y)$  est un point critique de  $f$  et si  $M$  est définie négative alors  $(x, y)$  est un maximum local.

En appliquant la règle de Sylvester :

- $a > 0$  et  $ac - b^2 > 0$  ssi  $M$  est définie positive  $\Rightarrow$  min local
- $a < 0$  et  $ac - b^2 > 0$  ssi  $M$  est définie négative  $\Rightarrow$  max local

Les autres cas possibles :

- $ac - b^2 = 0 \Leftrightarrow$  il y a une valeur propre nulle, on ne peut rien dire.
- $ac - b^2 < 0$ , permet d'affirmer qu'on a un point col.

# Un critère alternatif en dimension 2

Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , soit  $(x, y) \in \mathbb{R}^2$ .

On note  $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = \mathcal{H}_f(x, y)$  la matrice Hessienne de  $f$  en  $(x, y)$ .

En regardant le polynôme caractéristique de  $M$ ,

$\chi_M(X) = \begin{vmatrix} a-X & b \\ b & c-X \end{vmatrix} = X^2 - (a+c)X + ac - b^2 = (X - \lambda_1)(X - \lambda_2)$  avec  $\lambda_1, \lambda_2 \in \mathbb{R}$  les deux valeurs propres de  $M$  (sont réelles car  $M$  symétrique, se prouve facilement)

On a  $\lambda_1 + \lambda_2 = a + c$  et  $\lambda_1 \times \lambda_2 = ac - b^2$

Si  $ac - b^2 < 0$ , on a alors  $\lambda_1, \lambda_2$  de signe différent et strict, on a un point col.

On a deux vp strictement positives ssi  $ac - b^2 > 0$  et  $a + c > 0$

On a deux vp strictement négatives ssi  $ac - b^2 > 0$  et  $a + c < 0$

Si le déterminant  $ac - b^2 = 0$ , une valeur propre est nulle, les conditions d'optimalité d'ordre 2 ne permettent pas de conclure.

Le dernier cas serait  $ac - b^2 > 0$  et  $a + c = 0$ , implique  $-a^2 - b^2 > 0$ , ce n'est pas possible



# Quelle différence entre ces deux critères ?

- ▶ Ils sont complètement équivalents en dimension 2.
- ▶ Le second critère se voit sur des cours d'optimisation spécifiques en dimension 2, il est très facile à démontrer (cf planche précédente), contrairement au critère de Sylvester qui utilise des théorèmes plus forts.
- ▶ Le critère de Sylvester se généralise en dimension générale  $n$  comme dans les planches de l'amphi 2, contrairement à l'équivalence entre signe des valeurs propres et coefficients du polynôme caractéristique.

# Extremums locaux et matrice Hessienne en dimension 3

Soit  $M$  une matrice réelle symétrique  $3 \times 3$ ,  $M = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ .

En appliquant la règle de Sylvester et la règle de Sarrus pour un déterminant de taille 3 :

- ▶  $M$  est définie positive ssi  $a > 0$  et  $ad - b^2 > 0$  et  $adf + (bec)^2 - ae^2 - b^2f - c^2d > 0$
- ▶  $M$  est définie négative ssi  $a < 0$  et  $ad - b^2 > 0$  et  $adf + (bec)^2 - ae^2 - b^2f - c^2d < 0$
- ▶  $M$  non inversible  $\iff M$  admet 0 comme valeur propre  $\iff \det(M) = 0$   
ici  $\det(M) = 0 = adf + (bec)^2 - ae^2 - b^2f - c^2d$
- ▶ Comme on a des équivalences précédemment, les autres cas correspondent à avoir au moins une vp strictement positive et une vp strictement négative, et pas de vp nulle.

$\implies$  On peut déterminer assez facilement si une Hessienne d'une fonction de trois variables est définie positive ou définie négative, ou admet deux valeurs propres non nulles de signe différent.

# Hessiennes et fonctions convexes

Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^2$ .

Si pour tout  $x \in \mathbb{R}^n$ , la matrice Hessienne  $\mathcal{H}_f(x)$  est positive (ie a des valeurs propres positives ou nulles), alors  $f$  est convexe.

Si pour tout  $x \in \mathbb{R}^n$ , la matrice Hessienne  $\mathcal{H}_f(x)$  est définie positive (ie a des valeurs propres strictement positives), alors  $f$  est strictement convexe.

N.B : généralise le lien dérivée/convexité vu en dimension 1.

# Optimisation continue et fonctions convexes

Soit  $K$  un ensemble convexe de  $\mathbb{R}^n$ ,  
soit  $f : K \rightarrow \mathbb{R}$  une fonction convexe

Alors tout un minimum local de  $f$  est un optimum global.

Si  $f$  est strictement convexe, alors il y a au plus unicité du minimum local qui est alors minimum global.

Comme en dimension 1, il peut ne pas exister de minimum local.

$$\text{ex : } x \in \mathbb{R} \mapsto e^x - x \in \mathbb{R}$$

$$\text{ex : } (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto e^{x_1} + e^{x_2} - x_1 + x_3^2 \in \mathbb{R}$$

N.B : Ces résultats généralisent la dimension 1

# Application data mining

On a  $N$  points dans  $\mathbb{R}^D$ , notés  $X_i = (x_{i,1}, x_{i,2}, \dots, x_{i,D})$  pour  $i \in \llbracket 1, N \rrbracket$

Quel est le point  $z = (z_1, z_2, \dots, z_D) \in \mathbb{R}^D$  qui minimise

$$\sum_{i=1}^N \|z - X_i\|^2 = \sum_{i=1}^N \sum_{d=1}^D (z_d - x_{i,d})^2$$

$$f : z = (z_1, z_2, \dots, z_D) \in \mathbb{R}^D \mapsto \sum_{i=1}^N \|z - X_i\|^2 = \sum_{i=1}^N \sum_{d=1}^D (z_d - x_{i,d})^2 \in \mathbb{R}$$

admet-elle un minimum ?

La Hessienne  $\mathcal{H}_f$  est constante,  $\mathcal{H}_f = \text{diag}(2, 2, \dots, 2) = 2I_N$ , 2 est la seule valeur propre.  $\mathcal{H}_f$  est toujours définie positive.

Donc  $f$  est strictement convexe et atteint au plus un minimum global si  $\nabla_f(x) = 0$  admet une solution (aussitôt unique).

# Application data mining, résolution de $\nabla_f(x) = 0$

$$f(z_1, z_2, \dots, z_D) = \sum_{i=1}^N \sum_{d=1}^D (z_d - x_{i,d})^2$$

$$\frac{\partial f}{\partial z_d}(z_1, z_2, \dots, z_D) = \sum_{i=1}^N 2(z_d - x_{i,d}) = 2Nz_d - 2 \sum_{i=1}^N x_{i,d}$$

$$\frac{\partial f}{\partial z_d}(z) = 0 \iff 0 = Nz_d - \sum_{i=1}^N x_{i,d} \iff z_d = \frac{1}{N} \sum_{i=1}^N x_{i,d}$$

$$\text{Il y a une unique solution à } \nabla_f(z) = 0, \text{ c'est } z = \frac{1}{N} \sum_{i=1}^N x_i$$

Ne connaissiez-vous pas ce résultat différemment ?

# Clustering k-means, alias k-moyennes

Soit  $E = \{X_1, \dots, X_N\}$ ,  $N$  éléments de  $\mathbb{R}^d$ .

Soit  $\Pi_K(E)$ , l'ensemble des partitions de  $E$  en  $K$  éléments :

$$\Pi_K(E) = \left\{ P \subset \mathcal{P}(E) \mid \forall p, p' \in P, p \cap p' = \emptyset \text{ and } \bigcup_{p \in P} p = E \text{ and } \text{card}(P) = K \right\}$$

Le problème des  $k$ -moyennes, (k-means) se définit par :  $\min_{\pi \in \Pi_K(E)} \sum_{P \in \pi} f(P)$

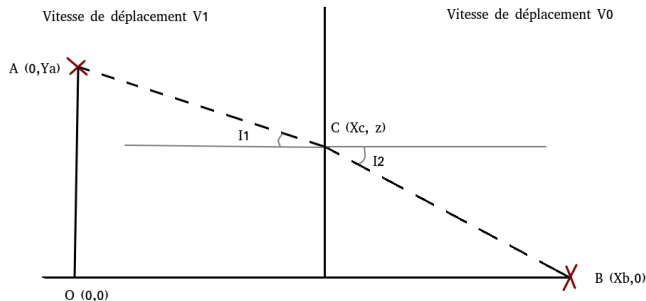
$$\text{avec } f(P) = \sum_{x \in P} \left\| x - \frac{1}{\text{card}(P)} \sum_{y \in P} y \right\|^2 = \min_{y \in \mathbb{R}^2} \sum_{x \in P} \|x - y\|^2 !!!$$

De manière générale, on peut considérer le clustering induit par :

$$f(P) = \min_{y \in \mathbb{R}^2} \sum_{x \in P} \|x - y\|^\alpha$$

mais la solution analytique n'existe pas forcément ! (ex  $\alpha = 1$ )

# Retour loi de Snell Descartes



On exprime le problème en fonction de  $i_1, i_2$ .

$$AC = \frac{x_c}{\cos i_1} \text{ donc } t_{AC} = \frac{x_c}{v_1 \cos i_1}$$

$$BC = \frac{x_b - x_c}{\cos i_0} \text{ donc } t_{BC} = \frac{x_b - x_c}{v_0 \cos i_0}$$

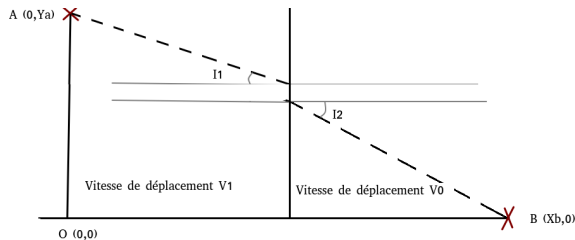
Dans la suite, on suppose pour simplifier les calculs que  $x_c = 1$  et  $x_b - x_c = 1$ ,  
on cherche donc à minimiser  $f(i_1, i_0) = \frac{1}{v_1 \cos i_1} + \frac{1}{v_0 \cos i_0}$



# Premier essai

En minimisant  $f(i_1, i_0) = \frac{1}{v_1 \cos i_1} + \frac{1}{v_0 \cos i_0}$  pour  $i_0, i_1 \in [0, \frac{\pi}{2}[$

On a une solution triviale,  $i_0 = i_1 = 0$ , mais qui ne convient pas !



Il manque une contrainte, les angles doivent permettre de taper au même point !

Contrainte portée sur l'axe vertical :  $x_c \tan i_1 + (x_b - x_c) \tan i_0 = y_A$

Dans la suite, on suppose pour simplifier les calculs que  $y_A = x_c = 1$  et  $x_b - x_c = 1$

# Approche intuitive

On se donne une valeur  $M$  très grande, et on cherche à minimiser :

$$f(i_1, i_0) = \frac{1}{v_1 \cos i_1} + \frac{1}{v_0 \cos i_0} + M (\tan i_1 + \tan i_0 - 1)^2 \text{ pour } i_0, i_1 \in \left[0, \frac{\pi}{2}\right[$$

Justification intuitive :  $M$  très grande, l'optimisation devrait éviter d'avoir une grande pénalisation en satisfaisant  $\tan i_1 + \tan i_0 = 1$ .

On serait alors ramené à de l'optimisation sans contrainte, on sait faire mntnt !

$$\begin{aligned} \frac{\partial f}{\partial i_1}(i_1, i_0) &= \frac{\sin i_1}{v_1 \cos^2 i_1} + 2M (\tan i_1 + \tan i_0 - 1) \times \frac{1}{\cos^2 i_1} \\ \frac{\partial f}{\partial i_0}(i_1, i_0) &= \frac{\sin i_0}{v_0 \cos^2 i_0} + 2M (\tan i_1 + \tan i_0 - 1) \times \frac{1}{\cos^2 i_0} \end{aligned}$$

$$\frac{\partial f}{\partial i_1}(i_1, i_0) = 0 \text{ et } \frac{\partial f}{\partial i_0}(i_1, i_0) = 0 \text{ impliquent :}$$

$$\frac{\sin i_1}{v_1} = \frac{\sin i_0}{v_0} \left( = -2M (\tan i_1 + \tan i_0 - 1) \right)$$

⇒ On a justifié l'intuition, la preuve "tirée du chapeau" est valide pour prouver le résultat avec cette intuition heuristique

⇒ On va pouvoir utiliser un tel formalisme en toute rigueur

# Optimisation sans contrainte en dimension $n$ , bilan partiel

- ▶ Optimisation sans contrainte en dimension  $n$  :  $n$  variables de décision à optimiser.
- ▶ En dimension  $n$ , on n'a pas la notion de monotonie (et les tableaux de variation) de la dimension 1, mais on a bcp de résultats généralisant la dimension 1.
- ▶ Extremums global/local définis. Pour une fonction dérivable, extremums locaux et gradients sont liés. Résoudre des équations  $\nabla_f(x) = 0$  est une base pour déterminer des minimum/maximums locaux. Savoir résoudre ces équations permet de calculer analytiquement des optimums.
- ▶ Des arguments de convexité, localement ou globalement peuvent aider à prouver que l'on a un minimum ou maximum local/global.
- ▶ Problème majeur récurrent : La résolution d'équations  $\nabla_f(x) = 0$  ne peut pas toujours se faire de manière analytique. On a en général un système d'équation non linéaire à  $n$  inconnues ...

# Plan

Rappels de topologie et analyse en dimension  $n$

Optimisation continue sans contrainte en dimension  $n$

**Optimisation discrète sans contrainte en dimension  $n$  ?**

Optimisation sous contraintes, cadre général

Optimisation continue sous contrainte

Algorithmes numériques, Algorithmes de gradient

Conclusions et perspectives

## Retour exercice 5 du TD1

Si un fermier effectue sa récolte de riz aujourd'hui, il obtiendra 1200 kg valant 0,40 euros le kg. Pour chaque semaine d'attente, la récolte augmente de 100 kg mais le prix baisse de 0,02 euros par kg. Quand devrait-il effectuer sa récolte pour maximiser ses revenus ?

Ici, la variable est un entier, nombre de semaine. On peut regarder numériquement les premières valeurs :

Soit  $n$  le nombre de semaines d'attente. On veut maximiser pour  $n \in \mathbb{N}$  :

$$P(n) = (1200 + 100n)(0,40 - 0,02n) = (12 + n)(40 - 2n)$$

Problème de type maximiser  $u_n = f(n)$  pour  $n \in \mathbb{N}$ .

# Retour exercice 5 du TD1

Ici, la variable est un entier, nombre de semaine. On peut regarder numériquement les premières valeurs :

Moment de la récolte	Quantité (kg)	Prix(Euro) / kg	Prix de la récolte totale (Euro)
Aujourd'hui	1200 kg	40 cents/kg	1200 . 40 = 48000 cents
1 semaine d'attente	1300 kg	40 - 2 = 38 cents/kg	1300 . 38 = 49400 cents
2 semaines d'attente	1400 kg	38 - 2 = 36 cents/kg	1400 . 36 = 50400 cents
3 semaines d'attente	1500 kg	36 - 2 = 34 cents/kg	1500 . 34 = 51000 cents
4 semaines d'attente	1600 kg	34 - 2 = 32 cents/kg	1600 . 32 = 51200 cents
5 semaines d'attente	1700 kg	32 - 2 = 30 cents/kg	1700 . 30 = 51000 cents
6 semaines d'attente	1800 kg	30 - 2 = 28 cents/kg	1800 . 28 = 50400 cents

GAIN

GAIN

GAIN

GAIN

PERTE

PERTE

Intuition, d'abord croissant puis décroissant.

Facile à montrer en étudiant la fonction continue  $x \in \mathbb{R}_+ \mapsto P(x) \in \mathbb{R}$ , la monotonie sur  $\mathbb{R}_+$  implique une monotonie sur  $\mathbb{N}$

$\forall x \in \mathbb{R}_+, P(x) \geq P(4)$ , 4 est l'optimum en valeurs continues et entières.

# Cas de la dimension 1, de manière plus générale

Problème de type maximiser  $u_n = f(n)$  pour  $n \in \mathbb{N}$ .

Si on peut calculer les variations de  $f$  avec des monotonies simples, on utilise la monotonie sur les sous intervalles où cela s'applique, et il resterait un nombre fini de possibilités.

ex :  $f$  croissante sur  $[0, 3.5]$  et décroissante sur  $[3.5, +\infty[$ , le maximum est  $\max(f(3), f(4))$

De manière générale on peut être plus embêté avec peu de monotonie locale.

ex :  $u_n = \sin n$

# QUBO : Quadratic Unconstrained Binary Optimization

Un QUBO est un problème qui s'écrit du type :

$$\min_{x \in \{0,1\}^n} \sum_{i=1}^n c_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j \quad (1)$$

Existence :  $x \in \{0,1\}^n$  définit un compact de  $\mathbb{R}^n$ , la fonction quadratique est continue (et même  $\mathcal{C}^\infty$  car polynomiale)

$2^n$  possibilités, on peut tout énumérer pour des petites valeurs de  $n$  pour prendre la meilleure solution



# ATTENTION!!!!

Quadratique n'implique pas convexe ou concave en dimension  $n$ , contrairement à la dimension 1!!

L'optimisation d'une fonction continue quadratique convexe permet de calculer une solution continue. On pourrait alors arrondir autour de la solution fractionnaire pour trouver des bonnes solutions, on restreindre l'espace de recherche autour de l'optimum continu.

Ce serait trop beau, on verra que les QUBO sont des problèmes d'optimisation très généraux avec bcp des applications pratiques. Les QUBO sont difficiles à résoudre de manière générale.

Fonction quadratique en dimension  $n$  : Hessienne constante, qui a les  $n^2$  coefficients de degré 2.

Quelle probabilité que les valeurs propres soient toutes négatives ou toutes positives ?

# Optimisation "boite noire"

- ▶ Une petite remarque, on pensait toujours à des fonctions numériques données, que l'on sait calculer en  $O(1)$  de manière analytique.
- ▶ En pratique, la fonction à optimiser peut être évaluée en un point uniquement avec des expériences (mesures physiques), qui prennent du temps, l'évaluation de la fonction n'est pas neutre
- ▶ Des méthodes algorithmique dédiées existent, notamment l'optimisation "boite noire" ou les algorithmes génétiques "surrogate"
- ▶ Optimisation "black-box", cours en M2.

# Plan

Rappels de topologie et analyse en dimension  $n$

Optimisation continue sans contrainte en dimension  $n$

Optimisation discrète sans contrainte en dimension  $n$  ?

**Optimisation sous contraintes, cadre général**

Optimisation continue sous contrainte

Algorithmes numériques, Algorithmes de gradient

Conclusions et perspectives

# Optimisation sous contraintes, dans le cadre de ce cours

- Optimisation sous contrainte :  $\min_{x \in X} f(x)$  où  $X = \{x \in \mathbb{Z}^n \times \mathbb{R}^p, g(x) \leq 0\}$

- Cela se note aussi 
$$\min_{\substack{x \\ \text{s.c : } g(x) \leq 0 \\ x \in \mathbb{Z}^n \times \mathbb{R}^p}} f(x) \quad (\text{s.c ou s.t : pour "sous contrainte"})$$

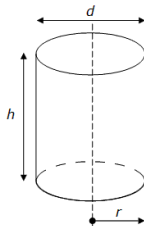
- $X$  est l'ensemble des solutions réalisables,  $X \subset \mathbb{Z}^n \times \mathbb{R}^p$ , décisions discrètes et continues. Pour ce cours,  $X$  est borné (cas majeur applicatif, qui garantit l'existence d'un minimum si  $f$  et  $g$  sont continues).
- $g : x \in \mathbb{Z}^n \times \mathbb{R}^p \mapsto \mathbb{R}^m$ , définit  $m$  **contraintes** d'inégalités (peut modéliser des égalités) : ce que la solution DOIT satisfaire,
- $f : X \mapsto \mathbb{R}$  : fonction **objectif** unique, définit une fonction de score pour classer les solutions, l'optimisation oriente vers des solutions que l'on SOUHAITE selon le choix de la fonction  $f$ .

# Optimisation continue/discrète/combinatoire

- Optimisation sous contraintes :
- $$\begin{array}{ll} \min_x & f(x) \\ \text{s.c :} & g(x) \leq 0 \\ & x \in \mathbb{Z}^n \times \mathbb{R}^p \end{array}$$
- Si  $n = 0$ , uniquement des variables continues, cadre de l'*optimisation continue*. (ce qu'on a vu dans les exemples jusque là)
- Si  $p = 0$ , uniquement des variables entières, cadre de l'*optimisation discrète*. En particulier, des décisions binaires OUI ou NON opérer une action.
- Si les variables peuvent être indicées comme des permutations/partitions d'un ensemble, cadre de l'*optimisation combinatoire*.
- Les algorithmes de résolution vont être fortement dépendant de ces structures.

# Retour application boîte de conserve (1)

Une boîte de conserve est un cylindre, de hauteur  $h$  et de rayon  $r$ . On veut optimiser les dimensions (ie en fonction de  $h$  et  $r$ ), de sorte à minimiser la surface  $S(r, h) = 2\pi r^2 + 2\pi rh$ , à un volume donné  $V_0 = \pi r^2 h$  de la boîte de conserve



Cela s'écrit  $\min_{(r,h) \in \{(r,h) \in \mathbb{R}^2 \mid g(r,h) \leq 0\}} S(r, h)$  avec  $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$  définie par :

$$g(r, h) = \begin{pmatrix} \pi r^2 h - V_0 \\ -\pi r^2 h + V_0 \\ -r \\ -h \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

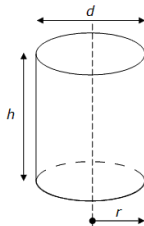
## Retour application boîte de conserve (2)

Une boîte de conserve est un cylindre, de hauteur  $h$  et de rayon  $r$ . On veut optimiser les dimensions, de sorte à maximiser le volume  $V(r, h) = \pi r^2 h$  à surface donnée  $S_0 = 2\pi r^2 + 2\pi rh$ .

Cela s'écrit :

$\max_{(r,h) \in \{(r,h) \in \mathbb{R}^2 \mid g(r,h) \leq 0\}}$   $V(r, h)$  avec  $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$  définie par :

$$g(r, h) = \begin{pmatrix} 2\pi r^2 + 2\pi rh - S_0 \\ S_0 - 2\pi r^2 - 2\pi rh \\ -r \\ -h \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$



# Minimiser la surface d'un pavé à volume fixé

On veut optimiser les dimensions d'un pavé, ie parallélépipède rectangle, en fonction de  $l_1$ ,  $l_2$  et  $l_3$ , de sorte à minimiser la surface

$S(l_1, l_2, l_3) = 2l_1 \times l_2 + 2l_2 \times l_3 + 2l_1 \times l_3$ , à un volume donné  $V_0 = l_1 \times l_2 \times l_3$  de la boîte de conserve.

Cela s'écrit 
$$\min_{(l_1, l_2, l_3) \in \{(l_1, l_2, l_3) \in \mathbb{R}^3 \mid g(l_1, l_2, l_3) \leq 0\}} l_1 \times l_2 + l_2 \times l_3 + l_1 \times l_3$$

avec  $g : \mathbb{R}^3 \longrightarrow \mathbb{R}^5$  définie par :

$$g(r, h) = \begin{pmatrix} l_1 \times l_2 \times l_3 - V_0 \\ -l_1 \times l_2 \times l_3 + V_0 \\ -l_1 \\ -l_2 \\ -l_3 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$\Rightarrow$  Cas que l'on peut résoudre en se ramenant à de l'optimisation sans contrainte en dimension 2, en exprimant  $l_3$  en fonction de  $l_1, l_2$



# Retour lois de Snell-Descartes

Pour la justification des lois de Snell-Descartes, on a vu qu'on se ramenait au problème d'optimisation sous contrainte, minimiser

$$f(i_1, i_0) = \frac{1}{v_1 \cos i_1} + \frac{1}{v_0 \cos i_0} \text{ sous contraintes } \tan i_1 + \tan i_0 = 1 \text{ pour } i_0, i_1 \in \left[0, \frac{\pi}{2} - \varepsilon\right]$$

Cela s'écrit 
$$\min_{(i_0, i_1) \in \{(i_0, i_1) \in [0, \frac{\pi}{2} - \varepsilon]^2 \mid g(i_0, i_1) \leq 0\}} \frac{1}{v_1 \cos i_1} + \frac{1}{v_0 \cos i_0}$$

avec  $g : \mathbb{R}^3 \longrightarrow \mathbb{R}^5$  définie par :

$$g(i_0, i_1) = \begin{pmatrix} \tan i_1 + \tan i_0 - 1 \\ 1 - \tan i_1 - \tan i_0 \\ -i_1 \\ -i_0 \\ i_1 - \frac{\pi}{2} - \varepsilon \\ i_0 - \frac{\pi}{2} - \varepsilon \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Justification intuitive :  $M$  très grande, l'optimisation devrait éviter d'avoir une grande pénalisation en satisfaisant  $\tan i_1 + \tan i_0 = 1$ .

# Variables et paramètres

- ▶ Optimisation sous contraintes :  $\min_{x \in X} f(x)$  où  $X = \{x \in \mathbb{N}^n \times \mathbb{R}_+^p, g(x) \leq 0\}$
- ▶  $x \in X$  désignent les variables, décisions ou leviers d'optimisation.
- ▶ A ne pas confondre avec des PARAMETRES, qui peuvent apparaître dans la fonction objectif ou dans les contraintes.
- ▶ Exemple :  $V_0$ ,  $S_0$ , pour les boîtes de conserves .
- ▶ A un jeu de paramètres donné est associé une solution optimale.
- ▶ Dans un premier temps, on supposera les paramètres connus et déterministes. La programmation stochastique traitera de données d'entrées soumises à des incertitudes, et cherchera à minimiser l'impact des aléas sur la solution optimisée.

# Objectif ou contrainte ?

- ▶ En pratique, les notions de contraintes et d'objectif ne sont pas toujours naturelles et peuvent causer des ambiguïtés.
- ▶ La fonction objectif définit une fonction de score pour classer les solutions. Exemples :
  - ▶ Temps : réaliser au plus tôt un ensemble de tâches.
  - ▶ Minimiser un coût financier/maximiser un profit.
- ▶ Pas forcément d'objectif unique et GLOBAL peuvent se dégager d'une solution, des notions locales de préférences peuvent apparaître. Cela induit des difficultés de modélisation. Cas de plusieurs objectifs potentiellement antagonistes : optimisation multi-objectif.
- ▶ Si on vous dit : "mon objectif numéro 1, c'est satisfaire mes contraintes ?"
  - ▶ S'il n'est pas possible de satisfaire toutes les contraintes, quel est le compromis à trouver ?
  - ▶ S'il est possible de satisfaire toutes les contraintes, quelle sont les solutions à favoriser ?

# Contraintes d'égalité ou d'inégalité ?

On peut écrire un pb d'optimisation contrainte sous deux formes équivalentes :

$$\begin{array}{ll} \min_x & f(x) \\ \text{s.c :} & g(x) \leq 0 \\ & x \in \mathbb{Z}^n \times \mathbb{R}^p \end{array} \quad \text{ou} \quad \begin{array}{ll} \min_x & f(x) \\ \text{s.c :} & g(x) = 0 \\ & x \in \mathbb{N}^n \times \mathbb{R}_+^p \end{array}$$

Sous la forme d'égalités, on se ramène à des inégalités en changeant une égalité en deux inégalités, et on met la positivité des variables comme une contrainte. (fait jusqu'ici)

Sous la forme d'inégalités, on rajoute une variables par contrainte d'inégalité, soit un vecteur  $\eta \geq 0$ , et on transforme chaque variable  $x_i$  et  $x_i = z_i - y_i$  avec  $z_i, y_i \geq 0$  ( $y, z$  représentent les parties positives et négatives)

$$\begin{array}{ll} \min_x & f(x) \\ \text{s.c :} & g(x) \leq 0 \\ & x \in \mathbb{Z}^n \times \mathbb{R}^p \end{array} = \begin{array}{ll} \min_{y,z,\eta} & f(z - y) \\ \text{s.c :} & g(z - y) = -\eta \\ & x, y \in \mathbb{N}^n \times \mathbb{R}_+^p, \eta \geq 0 \end{array}$$

# Résultat d'existence

## Théorème

Soit  $g : \mathbb{R}^{n+p} \rightarrow \mathbb{R}$  une fonction continue. On suppose que  $X = \{x \in \mathbb{N}^n \times \mathbb{R}_+^p, g(x) \leq 0\}$  est borné. Alors  $X = \{x \in \mathbb{N}^n \times \mathbb{R}_+^p, g(x) \leq 0\}$  est un compact et  $\min_{x \in X} f(x)$  et  $\max_{x \in X} f(x)$  sont bien définis.

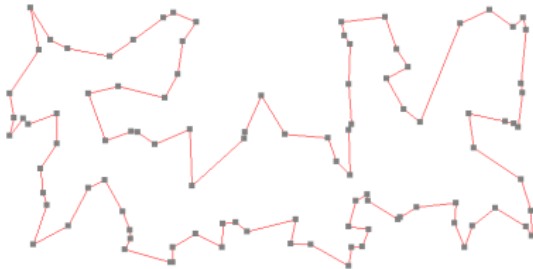
Preuve : Si  $g$  est continue,  $A = \{x \in \mathbb{R}^{n+p}, g(x) \leq 0\} = g^{<-1>}(\mathbb{R}_-)$  est fermé, image réciproque d'un fermé.  $X = A \cap \mathbb{N}^n \times \mathbb{R}_+^p$  est alors fermé par intersection de fermés.  $X$  est donc bien un compact.

# Exemple d'optimisation peu structurée

- ▶ On a  $n$  points  $(x_n)$  de  $\mathbb{R}^d$  sans aucune structure donnée.  $n$  très grand (contexte big data)
- ▶ On a un point donné  $y$  de  $\mathbb{R}^d$ . On veut déterminer quel est le point  $x_n$  le plus proche de  $y$ .
- ▶ Algorithme ; énumérer toutes les  $n$  possibilités, et garder en mémoire le point le plus proche trouvé.
- ▶ On n'aura pas d'autre alternative pour trouver l'optimum que de tout énumérer

# Exemple d'optimisation très structurée

- Un voyageur de commerce doit transiter par  $n$  villes données.
- Minimiser la longueur de son trajet.



# Exemples d'optimisation très structurée : le voyageur de commerce

- Structure du problème : trouver une liste ordonnée des villes minimisant la somme des distances entre deux points successifs.
- Cadre optimisation combinatoire : trouver une permutation  $f$  de  $\llbracket 1, n \rrbracket$  dans  $\llbracket 1, n \rrbracket$  avec  $f(1) = 1$  (choix arbitraire de première ville) minimisant la somme des distances entre deux points successifs.
- $(n - 1)!$  possibilités, impossible d'énumérer toutes les possibilités même avec des petites valeurs de  $n$ .
- On pourra formuler le problème en optimisation sous contraintes avec un nombre de variables et de contraintes polynomial en  $n$ .
- On pourra prouver l'optimalité de solutions sans tout énumérer.

Les algorithmes de résolution de ce cours rentrent dans un cadre structuré.



# Plan

Rappels de topologie et analyse en dimension  $n$

Optimisation continue sans contrainte en dimension  $n$

Optimisation discrète sans contrainte en dimension  $n$  ?

Optimisation sous contraintes, cadre général

**Optimisation continue sous contrainte**

Algorithmes numériques, Algorithmes de gradient

Conclusions et perspectives

# Motivation

On considère un pb d'optimisation continue sous contrainte sous l'une des formes suivantes où  $f$ , et  $g$  sont différentiables :

$$\begin{array}{lll} \min_{x \in \mathbb{R}^n} f(x) & \text{ou} & \min_{x \in \mathbb{R}_+^n} f(x) & \text{ou} & \min_{x \in \mathbb{R}_+^n} f(x) \\ \text{s.c : } g(x) \leq 0 & & \text{s.c : } g(x) = 0 & & \text{s.c : } h(x) = 0 \\ & & & & g(x) \leq 0 \end{array}$$

Sans contrainte, on a les résultats, peut on s'y ramener ?

Si on résout  $\min_{x \in \mathbb{R}^n} f(x)$  et que la solution optimale  $x^*$  trouvée réalise les contraintes, ie  $g(x^*) = 0$  par exemple. on a résolu le problème d'optimisation.

Preuve :  $f(x^*) = \min_{x \in \mathbb{R}^n} f(x) \leq \min_{x \in \mathbb{R}^n : g(x)=0} f(x)$

Mais si ce n'est pas le cas, peut on se ramener à un calcul de type  $\nabla_h(x) = 0$  ?

# Contraintes qualifiées en un point

Définition générale assez théorique et peu exploitable en calcul.

Ici, on survole et on explique les grandes lignes et intuitions sur ces types de problèmes, on laisse de côté ce détail technique qui ne servira pas dans la suite du cours, et on donne des conditions suffisantes. Des résultats plus généraux existent.

Les contraintes seront qualifiées en un point  $x$  si on se trouve dans un des trois cas suivants :

- ▶ Toutes les contraintes sont linéaires.
- ▶ Le gradient des contraintes actives en  $x$  sont linéairement indépendants.
- ▶ S'il n'existe qu'une seule contrainte (cas que l'on prends dans les illustrations)

# L'opérateur Lagrangien

Pour un problème d'optimisation sous contrainte  $\mathcal{P}$  avec  $n$  variables et  $m$  contraintes :

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.c : } & g(x) \leq 0 \end{aligned}$$

Le Lagrangien de  $\mathcal{P}$  se définit comme l'opérateur défini pour  $x \in \mathbb{R}^n$  et  $u \in \mathbb{R}_+^m$  :

$$\mathcal{L}(x, u) = f(x) + \langle u, g(x) \rangle = f(x) + \sum_{i=1}^m u_i g_i(x)$$

Avec des contraintes d'égalité, le Lagrangien s'écrit de même, sauf que les variables  $u_i$  correspondant aux contraintes d'égalité sont définies sur tout  $\mathbb{R}$  et plus sur  $\mathbb{R}_+$ .

# Conditions de KKT, optimisation locale (1)

KKT : Karush, Kuhn et Tucker

## Théorème

Soit  $\mathcal{P} = \min_{x \in X} f(x)$  où  $X = \{x \in \mathbb{R}^n, g(x) = 0\}$  un problème d'optimisation sous contrainte avec  $n$  variables et  $m$  contraintes d'égalité,  $f$ , et  $g_i$  étant  $C^1$ .

Soit  $x_0 \in \mathbb{R}$  tel que les vecteurs  $g'_i(x_0)$  soient linéairement indépendants. Si  $x_0$  est un minimum local de  $f$  sur  $X \{x \in \mathbb{R}^n, g(x) \leq 0\}$ , alors il existe  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  tels que :

$$\nabla f(x_0) + \sum_{i=1}^m \lambda_i \nabla g_i(x_0) = 0$$

$\implies$  Les conditions d'optimalité d'ordre 1 portent sur le Lagrangien et plus sur la fonction objectif

# Conditions de KKT, optimisation locale (2)

## Théorème

Soit  $\mathcal{P} = \min_{x \in X} f(x)$  où  $X = \{x \in \mathbb{R}^n, g(x) \leq 0\}$  un problème d'optimisation sous contrainte avec  $n$  variables et  $m$  contraintes d'inégalité,  $f$ , et  $g_i$  étant  $\mathcal{C}^1$ .

Soit  $x_0 \in \mathbb{R}$  tel que les contraintes  $g_i$  soient qualifiées en  $x_0$ . Si  $x_0$  est un minimum local de  $f$  sur  $X \{x \in \mathbb{R}^n, g(x) \leq 0\}$ , alors il existe  $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$  tels que :

$$\nabla f(x_0) + \sum_{i=1}^m \lambda_i \nabla g_i(x_0) = 0$$

$\implies$  Les conditions d'optimalité d'ordre 1 portent sur le Lagrangien et plus sur la fonction objectif

# Conditions de KKT convexes, optimisation globale

## Théorème

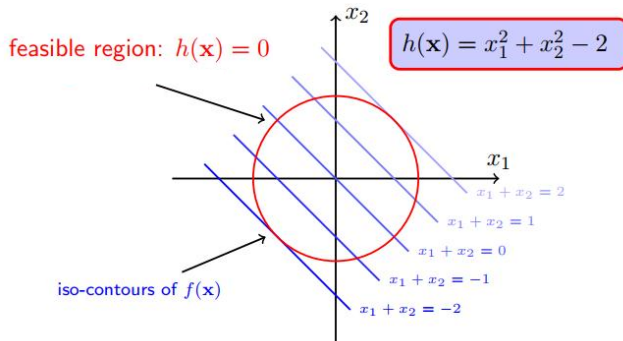
Soit  $\mathcal{P} = \min_{x \in X} f(x)$  où  $X = \{x \in \mathbb{R}^n, g(x) \leq 0\}$  un problème d'optimisation sous contrainte avec  $n$  variables et  $m$  contraintes d'inégalité,  $f$ , et  $g_i$  étant  $\mathcal{C}^1$  et convexes.

Soit  $x_0 \in \mathbb{R}^n$  tel que les contraintes  $g_i$  soient qualifiées en  $x_0$ .  $x_0$  est un minimum global de  $f$  sur  $X = \{x \in \mathbb{R}^n, g(x) \leq 0\}$  si et seulement si  $g(x_0) \leq 0$  et il existe

$$\lambda_1, \dots, \lambda_m \in \mathbb{R}_+ \text{ tels que : } \nabla f(x_0) + \sum_{i=1}^m \lambda_i \nabla g_i(x_0) = 0$$

Si  $f$  est strictement convexe, il y a au plus unicité du minimum global.

# Illustration (1)



Avec une contrainte  $h(x) \leq 0$  ou  $h(x) = 0$ , l'ensemble réalisable est défini comme un cercle ou un disque.

La fonction objectif à minimiser est  $f(x_1, x_2) = x_1 + x_2$ , définit des lignes de niveaux parallèles de points ayant la même valeur pour  $f$ .

A l'optimum, un seul point de tangence entre une ligne de niveau et la frontière  $\{x, h(x) = 0\}$



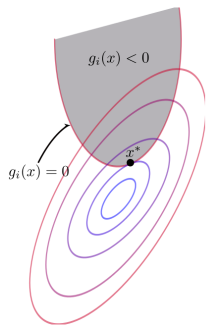
## Illustration (2)

On a ici une unique contrainte  $g(x) \leq 0$  et une fonction objectif  $f$  non linéaire à minimiser (couleur bleue pour valeur basse).

L'optimum en relâchant la contrainte ne vérifie pas  $g(x) \leq 0$ .

À l'optimum  $x^*$  du problème contraint, un seul point de tangence entre la ligne de niveau optimale et la frontière  $\{x, g(x) = 0\}$ .

Les gradients de la contrainte et de l'objectif en  $x^*$  ont la même direction, c'est ce que généralisent le Lagrangien et les conditions KKT!!



## Exemple numérique

$$\begin{array}{ll}\min_{x_1, x_2 \geq 0} & x_1^4 + x_2^4 + 12x_1^2 + 6x_2^2 - x_1x_2 - x_1 - x_2 \\ \text{s.c.} & x_1 + x_2 \geq 6 \\ & 2x_1 - x_2 \geq 3\end{array}$$

On montre aisément que  $g(x_1, x_2) = x_1^4 + x_2^4 + 12x_1^2 + 6x_2^2 - x_1x_2 - x_1 - x_2$  est strictement convexe par le calcul de la Hessienne.

Les contraintes sont convexes car linéaires. La version convexe du théorème KKT assure l'existence et l'unicité d'une solution optimale.

$$\begin{aligned}
& \min_{x_1, x_2 \geq 0} && x_1^4 + x_2^4 + 12x_1^2 + 6x_2^2 - x_1x_2 - x_1 - x_2 \\
& \text{s.c.} && x_1 + x_2 \geq 6 \\
& && 2x_1 - x_2 \geq 3
\end{aligned}$$

Le Lagrangien s'écrit :  $\mathcal{L}(x_1, x_2, u_1, u_2, u_3, u_4) = x_1^4 + x_2^4 + 12x_1^2 + 6x_2^2 - x_1x_2 - x_1 - x_2 + (6 - x_1 - x_2)u_1 + (3 - 2x_1 + x_2)u_2 - u_3x_1 - u_4x_2$

Il existe une unique solution au système KKT :

$$\begin{cases}
4x_1^3 + 24x_1 - x_2 - u_1 - 2u_2 - u_3 = 1 \\
4x_2^3 + 12x_2 - x_1 - u_1 + u_2 - u_4 = 1 \\
x_1 + x_2 \geq 6 \\
2x_1 - x_2 \geq 3 \\
(6 - x_1 - x_2)u_1 = 0 \\
(3 - 2x_1 + x_2)u_2 = 0 \\
u_3x_1 = 0 \\
u_4x_2 = 0 \\
x_1, x_2, u_1, u_2, u_3, u_4 \geq 0
\end{cases}$$

Résoudre un tel système KKT linéaire en général nécessite de distinguer des cas sur les 4 produits nuls de contrainte de complémentarité :

$$\begin{cases} 4x_1^3 + 24x_1 - x_2 - u_1 - 2u_2 - u_3 = 1 \\ 4x_2^3 + 12x_2 - x_1 - u_1 + u_2 - u_4 = 1 \\ x_1 + x_2 \geq 6 \\ 2x_1 - x_2 \geq 3 \\ (6 - x_1 - x_2)u_1 = 0 \\ (3 - 2x_1 + x_2)u_2 = 0 \\ u_3x_1 = 0 \\ u_4x_2 = 0 \\ x_1, x_2, u_1, u_2, u_3, u_4 \geq 0 \end{cases}$$

Ici, la solution est au plus unique par convexité ! Il suffit d'en trouver une.

Intuition :  $x_1 = 3$  et  $x_2 = 3$  définit une solution optimale. Plus facile à vérifier alors que de résoudre frontalement le système KKT.

# Retour lois de Snell-Descartes

Pour la justification des lois de Snell-Descartes, on a vu qu'on se ramenait à minimiser  $f(i_1, i_0) = \frac{1}{v_1 \cos i_1} + \frac{1}{v_0 \cos i_0}$  sous contraintes  $\tan i_1 + \tan i_0 = 1$  pour  $i_0, i_1 \in [0, \frac{\pi}{2}[$

On écrit le Lagrangien pour  $i_0, i_1 \in [0, \frac{\pi}{2}[$  et  $\lambda \in \mathbb{R}$

$$\begin{aligned}\mathcal{L}(i_1, i_0, \lambda) &= \frac{1}{v_1 \cos i_1} + \frac{1}{v_0 \cos i_0} + \lambda (\tan i_1 + \tan i_0 - 1) \\ \frac{\partial \mathcal{L}}{\partial i_1}(i_1, i_0, \lambda) &= \frac{\sin i_1}{v_1 \cos^2 i_1} + \lambda (\tan i_1 + \tan i_0 - 1) \times \frac{1}{\cos^2 i_1} \\ \frac{\partial \mathcal{L}}{\partial i_0}(i_1, i_0, \lambda) &= \frac{\sin i_0}{v_0 \cos^2 i_0} + \lambda (\tan i_1 + \tan i_0 - 1) \times \frac{1}{\cos^2 i_0}\end{aligned}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial i_1}(i_1, i_0) = 0 \text{ et } \frac{\partial \mathcal{L}}{\partial i_0}(i_1, i_0) = 0 \text{ impliquent :} \\ \frac{\sin i_1}{v_1} = \frac{\sin i_0}{v_0} \left( = -\lambda (\tan i_1 + \tan i_0 - 1) \right)\end{aligned}$$

$\implies$  Cette fois-ci, la preuve est rigoureuse !

# Retour application boîte de conserve (1)

$$\max_{(r,h) \in \mathbb{R}_+^2} r^2 h \text{ sous contrainte } 2\pi r^2 + 2\pi rh = S_0, \text{ ie } r^2 + rh = S'_0 \text{ avec } S'_0 = \frac{S_0}{2\pi}$$

On écrit le Lagrangien pour  $r, h \geq 0$  et  $\lambda \in \mathbb{R}$

$$\mathcal{L}(r, h, \lambda) = r^2 h + \lambda (r^2 + rh - S'_0)$$

$$\frac{\partial \mathcal{L}}{\partial r}(r, h, \lambda) = 2rh + \lambda(2r + h)$$

$$\frac{\partial \mathcal{L}}{\partial h}(r, h, \lambda) = r^2 + \lambda r$$

$$\frac{\partial \mathcal{L}}{\partial h}(r, h, \lambda) = \frac{\partial \mathcal{L}}{\partial r}(r, h, \lambda) = 0 \text{ impliquent } \lambda = -r$$

$$2rh - r(2r + h) = 0$$

$2h = 2r + h$  donc  $h = 2r$ , on le retrouve plus facilement !

Ensuite on reporte dans  $2\pi r^2 + 2\pi rh = S_0 = 4\pi r^2$  pour trouver  $r = \sqrt{\frac{S_0}{2\pi}}$

## Retour application boîte de conserve (2)

$$\min_{(r,h) \in \mathbb{R}_+^2} r^2 + rh \text{ sous contrainte } r^2 h = V_0$$

On écrit le Lagrangien pour  $r, h \geq 0$  et  $\lambda \in \mathbb{R}$

$$\mathcal{L}(r, h, \lambda) = \lambda(r^2 h - V_0) + r^2 + rh$$

$$\frac{\partial \mathcal{L}}{\partial r}(r, h, \lambda) = \lambda 2rh + (2r + h)$$

$$\frac{\partial \mathcal{L}}{\partial h}(r, h, \lambda) = \lambda r^2 + r$$

$$\frac{\partial \mathcal{L}}{\partial h}(r, h, \lambda) = \frac{\partial \mathcal{L}}{\partial r}(r, h, \lambda) = 0 \text{ impliquent } 1/\lambda = -r \text{ ie } \lambda r = -1$$

$\lambda 2rh + (2r + h) = 0$  implique  $-2h + 2r + h = 0$  donc  $h = 2r$ , on le retrouve plus facilement !

$$\text{Ensuite on reporte dans } 2\pi r^2 + 2\pi rh = S_0 = 4\pi r^2 \text{ pour trouver } r = \sqrt{\frac{S_0}{2\pi}}$$

# Retour application boîte de conserve, bilan

On voit que les deux problèmes induisent les mêmes calculs à  $1/\lambda$  près, et la même forme optimale,  $h = 2r$ , la hauteur est égale au diamètre (vous pouvez vérifier à la maison).

Idem avec le rectangle d'aire maximale à périmètre défini ou de périmètre minimal à aire définie. Ce sont les mêmes problèmes d'optimisation intuitivement !

Il y a une théorie derrière cela, théorie de la dualité.

On reverra la dualité, le Lagrangien dans le cadre de la Programmation Linéaire, si le Covid le permet !

Autre exemple, quelle est la forme optimale d'une casserole ? Comme un cylindre, mais un seul fond au lieu de 2 couvercles.



# Application forme optimale de casserole

$$\min_{(r,h) \in \mathbb{R}_+^2} \pi r^2 + 2\pi rh \text{ sous contrainte } r^2 h = V_0$$

On écrit le Lagrangien pour  $r, h \geq 0$  et  $\lambda \in \mathbb{R}$

$$\mathcal{L}(r, h, \lambda) = \lambda(r^2 h - V_0) + \pi r^2 + 2\pi rh$$

$$\frac{\partial \mathcal{L}}{\partial r}(r, h, \lambda) = \lambda 2rh + (2\pi r + 2\pi h)$$

$$\frac{\partial \mathcal{L}}{\partial h}(r, h, \lambda) = \lambda r^2 + 2\pi r$$

$$\frac{\partial \mathcal{L}}{\partial h}(r, h, \lambda) = \frac{\partial \mathcal{L}}{\partial r}(r, h, \lambda) = 0 \text{ impliquent } \lambda r = -2\pi$$

$$\lambda 2rh + (2\pi r + 2\pi h) = 0 \text{ implique } -4\pi h + 2\pi r + 2\pi h = 0 \text{ donc } h = r.$$

La forme optimale d'une casserole est définie par "hauteur = rayon" (vous pouvez vérifier à la maison).

# Plan

Rappels de topologie et analyse en dimension  $n$

Optimisation continue sans contrainte en dimension  $n$

Optimisation discrète sans contrainte en dimension  $n$  ?

Optimisation sous contraintes, cadre général

Optimisation continue sous contrainte

Algorithmes numériques, Algorithmes de gradient

Conclusions et perspectives

# Pourquoi des algorithmes numériques ?

- Problème majeur : La résolution d'équations  $\nabla f(x_0) = 0$  ne peut pas toujours se faire de manière analytique, même en dimension 1.
- En dimension 1, si  $f'$  est continue, et qu'a a  $a, b$  tels que  $f'(a)f'(b) < 0$ , une recherche dichotomique calcule efficacement avec une bonne précision numérique de telles racines.
- $\nabla f(x_1, \dots, x_n) = 0$  donne en général un système d'équations non linéaires à  $n$  variables
- En dimension  $n$  quelconque, comment faire pour avoir une résolution numérique approchée ?

# Comment interpréter le gradient ?

Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  admet un gradient en un point  $a \in A$ .

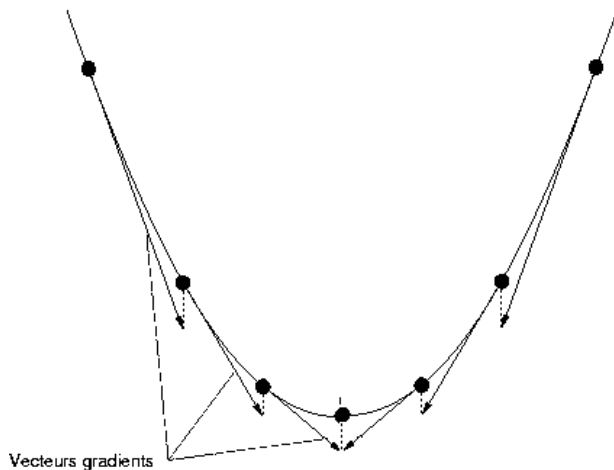
Alors on peut écrire un développement limité du premier ordre pour tout  $h = (h_1, \dots, h_i, \dots, h_n)$  :

$$f(a + h) = f(a) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a) + o(\|h\|) = f(a) + \langle h, \nabla f(a) \rangle + o(\|h\|)$$

$\nabla f(a)$  donne la direction pour augmenter le plus possible la valeur de  $f(x)$ .

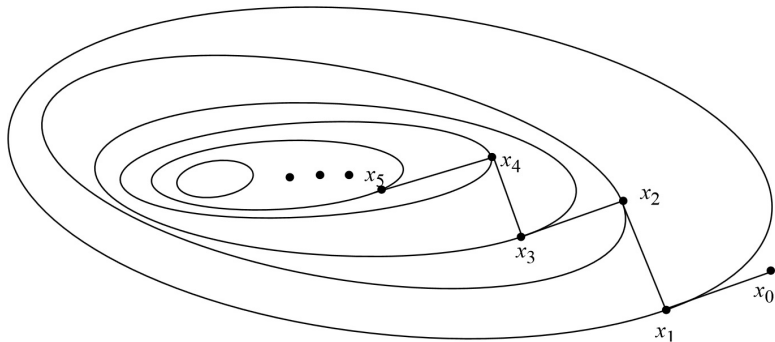
A partir d'un point  $a$  tel que  $\nabla f(a) \neq 0$ , si on veut diminuer la valeur de  $f(x)$ , on a intérêt à suivre la direction  $-\nabla f(a)$

# Illustration en dimension 1



Si  $f'(x) < 0$ , on diminue la fonction objectif en allant à "droite", vers direction  $f'(x+h)$  avec  $h \geq 0$ .

# Illustration en dimension 2

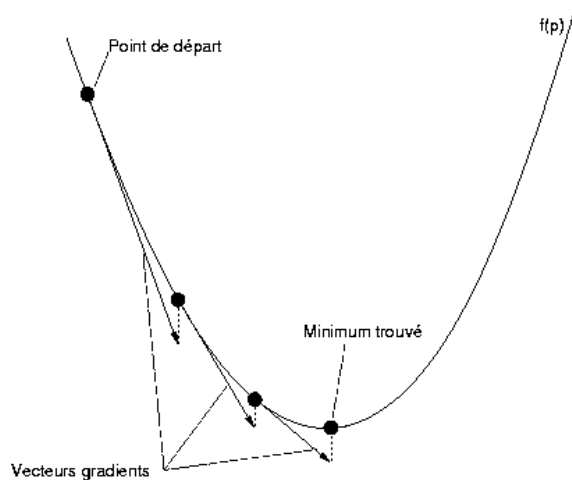


En projetant en dimension 2 avec des lignes de niveaux (et couleurs) une fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$ , le gradient est orthogonal aux lignes de niveaux

# Algorithme de gradient à pas fixe

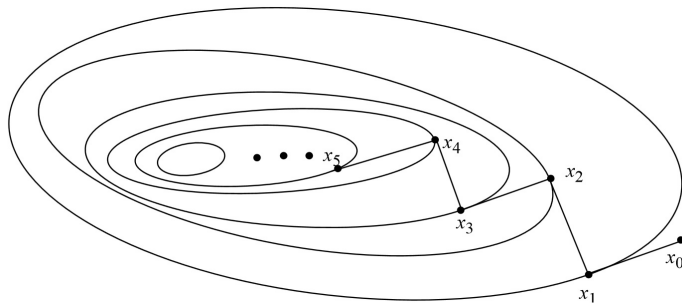
- ▶ Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est une fonction  $\mathcal{C}_1$  algorithme pour converger vers un minimum local de  $f$ , qui vérifiera  $\nabla f(a) = 0$ .
- ▶ Soit  $a \in \mathbb{R}^n$  un point initial,  $\varepsilon > 0$  une tolérance numérique,  $p > 0$  un coefficient de proportionnalité
- ▶ Algorithme : Tant que  $\|\nabla f(a)\| \neq 0$  (ou autre critère d'arrêt), on remplace  $a$  par  $a - p\nabla f(a)$
- ▶ Remarque : critère d'arrêt peut être ou comprendre en pratique :
  - ▶  $\|\nabla f(a)\| \geq \varepsilon$ , pour éviter de soucis d'arrondis numériques.
  - ▶  $|f(new) - f(old)| \leq \varepsilon$  : la convergence devient lente (utilisé en ML notamment pour éviter des effets d'over-fitting spécifique problématiques ML)
  - ▶ un nombre maximal d'itérations ou un temps maximal pour l'optimisation. (ex : robotique, calculs embarqués)

# Illustration en dimension 1

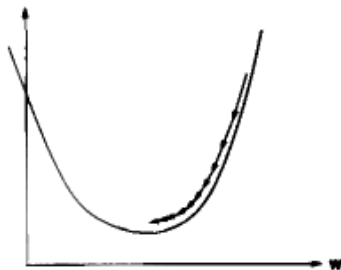




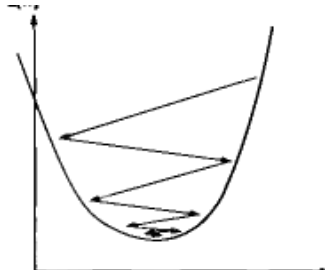
# Illustration en dimension 2



# Illustration des vitesses de convergence



coefficient faible



coefficient élevé

Connu des utilisateurs ML, le dosage du coefficient proportionnel  $p$  est délicat.

Pour des valeurs trop élevées, l'algorithme diverge.

Pour des valeurs suffisamment basses, l'algorithme converge (si la fonction est suffisamment régulière).

Pour des valeurs suffisamment basses où l'algorithme converge, plus le coefficient est faible, plus la convergence est longue. La convergence est très lente si  $p$  est trop petit

# Résultat de convergence

## Théorème

Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$ ,  $\alpha$ -convexe avec  $\alpha > 0$  qui admet un minimum global et telle que  $x \mapsto \nabla f(x)$  est  $L$ -Lipschitzienne, ie :

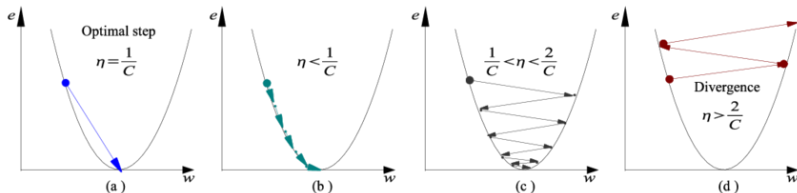
$$\forall x, y \in \mathbb{R}^n, \quad \|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

Alors, l'algorithme de gradient à pas fixe converge vers le minimum global pour tout pas  $p \in \left]0, \frac{2\alpha}{L^2}\right[$

Quelques précautions avec ce résultat :

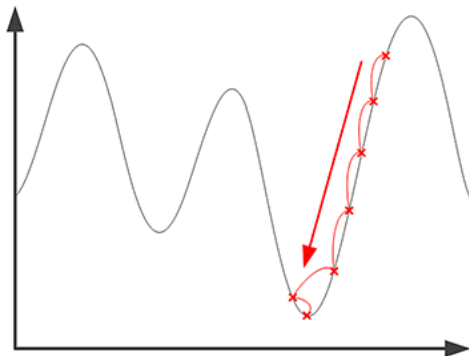
- ▶ L'hypothèse  $\alpha$ -convexe est forte.
- ▶ Les paramètres  $\alpha, L$  ne se calculent pas facilement de manière générale, et sont le plus souvent inconnus.

# Illustration des ratios de convergence



La paramétrisation reste souvent empirique, on peut avoir en tête les cas ci dessus même sans connaître  $\alpha, L$  pour calibrer la méthode.

# Illustration cas non convexe



⇒ Avec critère localement  $\alpha$ -convexe, l'algorithme de gradient à pas fixe peut converger vers un optimum local différent d'un optimum global

N.B : pour tous les raffinements des algorithmes de gradients, il en sera de même en absence de convexité globale.

# Algorithme du gradient à pas optimal

Idée : on fait varier le pas  $p$  à chaque itération. en prenant un pas "optimal", on pourra converger en moins d'itérations vite, en passant plus de temps à chaque itération pour recalculer les pas.

Soit  $x_0$  le point initial ou un point atteint par itération d'algorithme de gradient, tel que  $d_0 = \nabla f(x_0) \neq 0$ .

Pour  $h > 0$  proche de 0, on aura  $f(x_0 - hd) < f(x_0)$ .

On considère l'application  $h \mapsto f(x_0 - hd)$ . Le pas choisi est  $p = h^* = \operatorname{argmin}_{h \geq 0} f(x_0 - hd)$

On se ramène à un problème de dimension 1, la recherche dichotomique peut s'appliquer pour une résolution numérique très rapide et de très bonne qualité.

Remarque : numériquement, on n'est pas obligé de considérer exactement l'optimum, une très bonne solution obtenue rapidement suffit. On peut aussi utiliser des calculs indépendants effectués en parallèle.

# Algorithme du gradient à pas optimal, exemple

On considère le problème d'optimisation sans contrainte :

$$\min_{x_1, x_2} 2x_1^2 - x_1x_2 + x_2^2 - 3x_1 + e^{2x_1+x_2}.$$

On pose pour  $x = (x_1, x_2) \in \mathbb{R}^2$ ,  $f(x) = 2x_1^2 - x_1x_2 + x_2^2 - 3x_1 + e^{2x_1+x_2}$ . Les conditions du premier ordre correspondent à la recherche de solution à l'équation  $\nabla_f(x) = 0$ , i.e.

$$\nabla_f(x) = \begin{pmatrix} 4x_1 - x_2 - 3 + 2e^{2x_1+x_2} \\ -x_1 + 2x_2 + e^{2x_1+x_2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Un tel système ne se résout pas analytiquement.

# Algorithme du gradient à pas optimal, exemple

On calcule la Hessienne de  $f$ ,  $\mathcal{H}_f(x) = \begin{pmatrix} 4+4e^{2x_1+x_2} & 2e^{2x_1+x_2}-1 \\ 2e^{2x_1+x_2}-1 & 2+e^{2x_1+x_2} \end{pmatrix}$ .

On a  $4 + 4e^{2x_1+x_2} > 0$  pour tout  $x_1, x_2$ , on regarde si le critère de Sylvester est vérifié en calculant le déterminant  $\det(\mathcal{H}_f(x))$  :

$$\det(\mathcal{H}_f(x)) = \begin{vmatrix} 4+4e^{2x_1+x_2} & 2e^{2x_1+x_2}-1 \\ 2e^{2x_1+x_2}-1 & 2+e^{2x_1+x_2} \end{vmatrix} = \\ (4 + 4e^{2x_1+x_2})(2 + e^{2x_1+x_2}) - (2e^{2x_1+x_2} - 1)^2$$

On pose  $Z = e^{2x_1+x_2}$ , on sait que  $Z > 0$  et

$$\det(\mathcal{H}_f(x)) = 4(1+Z)(2+Z) - (2Z-1)^2 = \\ 4(Z^2 + 3Z + 2) - (4Z^2 - 2Z + 1) = 14Z + 7 > 7 > 0$$

Le critère de Sylvester assure donc que la Hessienne de  $f$  est définie positive, et donc  $f$  est strictement convexe, et définie sur un ensemble convexe  $\mathbb{R}^2$ , il y a au plus unicité d'un minimum local qui serait minimum global.



# Algorithme du gradient à pas optimal, exemple

Soit  $O = (0, 0)$   $\nabla f(O) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  donc  $O$  n'est pas un point de minimum local. Suivant la direction  $(1, 0)$ , on a une direction de descente partant de  $O$ . Suivant l'algorithme de gradient à pas optimal, on considérerait la direction  $(1, -1)$ .

Soit  $d = (1, -1)$ . On cherche alors à minimiser  
 $\lambda \in \mathbb{R}_+ \mapsto h(\lambda) = f(O + \lambda d) \in \mathbb{R}$ .

Soit  $\lambda \geq 0$ .  $h(\lambda) = f(O + \lambda d) = 2\lambda^2 + \lambda^2 + \lambda^2 - 3\lambda + e^\lambda = e^\lambda + 4\lambda^2 - 3\lambda$

$h$  est de classe  $\mathcal{C}^\infty$

$$h'(\lambda) = e^\lambda + 8\lambda - 3$$

$h''(\lambda) = e^\lambda + 8 > 0$  donc  $h$  est strictement convexe.

$h'(0) = -2 < 0$  et  $h'(1) = e + 5 > 0$ , donc  $h'(x) = 0$  admet une solution dans  $[0, 1]$  qui est unique par convexité, ce qui définit un minimum global pour  $h$ . Numériquement, on trouve à l'optimum  $\lambda \approx 0,219$  avec un algorithme dichotomique.

# Variantes d'algorithmes de gradient, optim sans contrainte

Toujours dans le cas d'optimisation sans contraintes, plusieurs variantes des algorithmes de gradient existent.

L'algorithme de gradient stochastique est couramment utilisé en ML pour bien correspondre à certaines propriétés numériques du ML.

Les variantes suivantes utilisent aussi la Hessienne :

- ▶ L'algorithme de gradient conjugué
- ▶ Méthode de Newton et pseudo-Newton
- ▶ Méthode de Broyden-Fletcher-Goldfarb-Shanno (BFGS)
- ▶ Algorithme de Levenberg-Marquardt (LM)

L'algorithme de sous-gradient s'applique à des fonctions convexes mais non différentiables (par exemple linéaire ou quadratique convexe par morceaux)

# Rappel : Développement limités d'ordre 1 et 2

Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  admet un gradient en un point  $a \in A$ .

Alors on peut écrire un développement limité du premier ordre pour tout  $h = (h_1, \dots, h_i, \dots, h_n)$  :

$$f(a+h) = f(a) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a) + o(\|h\|) = f(a) + \langle h, \nabla f(a) \rangle + o(\|h\|)$$

Si  $f$  est de classe  $\mathcal{C}^2$ ,

$$f(a+h) = f(a) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a) + \sum_{i=1}^n h_i^2 \frac{\partial^2 f}{\partial x_i^2}(a) + \sum_{i=1}^n \sum_{j=i+1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + o(\|h\|^2)$$

# Extensions d'algorithmes de gradient avec contraintes

Passer des algorithmes numériques sans contraintes à l'ajout de contraintes, est-ce appliquer les algorithmes de gradient précédent au Lagrangien et à la recherche de solutions  $\nabla \mathcal{L}(x_0) = 0$  ?

En théorie, oui, en pratique, ça se corse.

Une telle famille de méthodes : méthodes de pénalisation

Difficulté : numériquement on ne converge pas à l'optimum avec les algos de gradient sans contrainte. Implications pour les méthodes de pénalisation :

- ▶ On peut converger numériquement vers un point où les contraintes ne sont pas satisfaites alors que théoriquement, on convergerait vers un minimum local vérifiant les contraintes.
- ▶ On peut être coincé dans une zone où des contraintes ne sont pas satisfaites et avoir des minimums locaux du Lagrangien dans ces zones.

En pratique, cela devient difficile sur des problèmes fortement contraints (bcp de contraintes), plus gérable avec peu de contraintes.

Dans la suite, une autre famille de méthodes, les algorithmes de gradient projeté

# Algorithme du gradient projeté

On considère un problème  $\min_{x \in X} f(x)$  où  $X = \{x \in \mathbb{R}^n, g(x) \leq 0\}$

Idee : On suppose disposer d'un opérateur de projection  $p_X$  de  $\mathbb{R}^n$  dans l'ensemble réalisable  $X = \{x \in \mathbb{R}^n, g(x) \leq 0\}$

Au lieu de considérer une expression  $a - \eta \nabla f(a)$ , on considère  $p_X(a - \eta \nabla f(a))$  dans les algorithmes de gradients

Résultats théoriques étendant les résultat de gradient à pas fixe, avec les mêmes hypothèse de fonction  $\alpha$ -convexe et de gradient L-lipschitzien.

Problème supplémentaire : Si l'opérateur de projection  $p_X$  existe car  $X$  est fermé, son calcul n'est pas facile dans le cas général.

Remarque : une projection est une optimisation, cela revient à minimiser la distance d'un point à un sous ensemble, ici  $X$ .

# Algorithme d'Uzawa, gradient avec contraintes de bornes

Algorithme d'Uzawa : Pour des problèmes avec des contraintes de bornes, ie : des problèmes  $\min_{x \in X} f(x)$  où  $X = [a_1, b_1] \times \cdots \times [a_n, b_n]$  (un pavé de dimension  $n$ )

Alors l'opérateur de projection est connu et simple.

Si  $(a - \eta \nabla f(a))_i < a_i$ , on considère  $a_i$  au lieu de  $(a - \eta \nabla f(a))_i$

Si  $(a - \eta \nabla f(a))_i > b_i$ , on considère  $b_i$  au lieu de  $(a - \eta \nabla f(a))_i$

Et on a un algorithme qui étend les propriétés du gradient à pas fixe, sans que le coût des calculs de projection soit coûteux.

En pratique : utilisé parfois sans trop le dire en ML, de telles opérations (tout comme des renormalisations) sont bien valides grâce à ces théorèmes.

# Critère d'arrêt pour l'Algorithme d'Uzawa

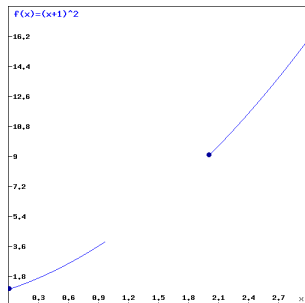
Sur cet exemple qu'on a déjà vu :

$$f : x \in [0, 1] \cup [2, 3] \mapsto (x + 1)^2 \in \mathbb{R}$$

On a ici deux minimums locaux avec  $\nabla f(x_0) \neq 0$  : en  $x_0 = 0$  et  $x_0 = 2$

Partant de ces deux points (avec un  $p$  suffisamment petit pour  $x_0 = 2$ ), on reste coincé, et on n'en bouge plus, itérant à l'infini si on avait uniquement un critère d'arrêt de type  $\|\nabla f(a)\| \neq 0$  ou  $\|\nabla f(a)\| \geq \varepsilon$ .

Stagner sur un point implique ici une boucle infinie, on peut s'y arrêter et ne pas oublier de l'ajouter comme critère d'arrêt !



# Algorithmes de gradient et Optimisation "boite noire"

- ▶ Remarque : en pensant à des fonctions numériques données, on sait évaluer en  $O(1)$  de manière analytique la fonction, son gradient, voire sa Hessienne.
- ▶ En pratique, la fonction à optimiser peut être évaluée en un point uniquement avec des expériences (mesures physiques), qui prennent du temps, l'évaluation de la fonction n'est pas neutre.
- ▶ Pire ! Dans un tel cadre, on n'a pas la dérivée.
- ▶ "Gradient-free optimisation" et Optimisation "black-box", cours en M2.



# Plan

Rappels de topologie et analyse en dimension  $n$

Optimisation continue sans contrainte en dimension  $n$

Optimisation discrète sans contrainte en dimension  $n$  ?

Optimisation sous contraintes, cadre général

Optimisation continue sous contrainte

Algorithmes numériques, Algorithmes de gradient

Conclusions et perspectives

# Pour aller plus loin

Ce cours peut faire l'objet d'une UE entière, en approfondissant les aspects théoriques, généralisant les résultats, décrivant bien plus précisément les nombreuses variantes d'algorithmes

Vous pourrez avoir besoin de ces notions sur d'autres cours, notamment en IA et ML. Vous avez les pointeurs vers les noms des méthodes, pour approfondir par vous même si vous en avez besoin.

Ressources en français conseillées pour approfondir :

Cours polycopié d'Aude Rondepierre à l'INSA-Toulouse, accessible en ligne  
<http://www.math.univ-toulouse.fr/~rondep/CoursTD/polyGMM4.pdf>

Une "Bible du domaine" : Éléments d'Optimisation Différentiable, JC Gilbert.  
Peut être demandée à l'auteur :

<https://who.rocq.inria.fr/Jean-Charles.Gilbert/ensta/optim.html>

# Bilan optimisation sans contrainte

- ▶ Optimisation sans contrainte en dimension  $n$  :  $n$  variables de décision à optimiser.
- ▶ En dimension  $n$ , on n'a pas la notion de monotonie (et les tableaux de variation), mais on a bcp de résultats généralisant la dimension 1.
- ▶ Extremums global/local définis. Pour une fonction dérivable, extremums locaux et gradients sont liés. Résoudre des équations  $\nabla f(x) = 0$  est une base pour déterminer des minimum/maximums locaux. Savoir résoudre ces équations permet de calculer analytiquement des optimums.
- ▶ Des arguments de convexité, localement ou globalement peuvent aider à prouver que l'on a un minimum unique.
- ▶ Problème majeur récurrent : La résolution d'équations  $\nabla f(x) = 0$  ne peut pas toujours se faire de manière analytique. L'algorithme dichotomique efficace de dimension 1 ne se généralise pas, on a l'algorithme de gradient qui est très utilisé en pratique.

# Bilan optimisation sous contrainte

- ▶ On a défini l'optimisation sous contrainte, de manière générale avec des décisions discrètes et continues, avec un résultat d'existence qui suffit pour la suite du cours.
- ▶ Les conditions de Karush-Kuhn-Tucker (KKT) permettent d'étendre des résultats théoriques d'optim sans contrainte, en faisant porter les résultats sur le Lagrangien.
- ▶ Les algorithmes de gradient se généralisent mal à l'optimisation sous contrainte, particulièrement pour les problèmes fortement contraints.

# Bilan des applications

- ▶ En optique géométrique, le principe de Fermat indique que la lumière suit des chemins correspondant à des minimums locaux des temps de trajet pour la lumière.
- ▶ En mécanique Newtonienne ou physique quantique, les minimas locaux (d'énergie) ont également des significations physiques, ce cours donne les outils pour résoudre de tels problèmes.
- ▶ Boîte de conserve, casserole, pour l'anecdote ! Mais ça illustre que les théorèmes KKT permettent de résoudre plus facilement ces problèmes plutôt que de "convertir " ces problèmes en optimisation sans contrainte.
- ▶ Application ML : on a une caractérisation équivalente des k-means / k-moyennes, et on peut ouvrir vers des généralisations.
- ▶ Application ML : on peut prouver l'existence et calculer les formules de régression linéaire.

# Perspectives algorithmes de gradients

Les algorithmes de gradients, des outils simples et fondamentaux, que vous reverrez :

- ▶ En ML, bien sûr ! La phase rétro-propagation des réseaux de neurones.
- ▶ En robotique ou dans le cadre étendu hilbertien sur un espace de fonction avec méthodes d'éléments finis.
- ▶ le simplexe : un algorithme de gradient projeté spécialisé
- ▶ PLNE : algo de gradient + Lagrangien servent pour des méthodes avancées d'optimisation discrète.
- ▶ Méta-heuristiques : on aura l'équivalent des algorithmes de descente pour des problèmes d'optimisation discrète.