OUTILS D'OPTIMISATION POUR LES SCIENCES DES DONNÉES ET DE LA DÉCISION

Séance 1 (20/09/2024)

- Présentation du cours
- Introduction à l'optimisation (CM+TD)

Logistique

Page du cours

https://www.lamsade.dauphine.fr/~croyer/coursODD.html

Mon mail

clement.royer@lamsade.dauphine.fr

Organisation :

8 séances de 3h

Vendredi matin
8h30-11h45

1h30 CM (C. Royer)
1h30 TD/TP (S. Kerleau)

⚠ Créneau du 04/10 va peut-être être déplacé

Notation

60% examen (13/12)

NB : Autorisé : une feuille A4 de notes

40% projet (notebook Python) → individuel

] Date de rendu : fin décembre janvier

INTRODUCTION À L'OPTIMISATION

Définition informelle : Optimiser = prendre la meilleure décision possible parmi un ensemble de possibilités

Optimisation : Domaine scientifique

Exemples de domaines d'application de l'optimisation :
Finance (optimisation de portefeuille), Réseaux d'énergie, sciences de données, travaux à Dauphine

Mathématiquement : On écrit un problème d'optimisation sous la forme :

on cherche la plus petite valeur possible

"sous les contraintes"

$$\underset{x \in \mathbb{R}^d}{\text{minimiser}} \; f(x) \quad \text{s.c.} \quad x \in \mathcal{X}$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}$$

$x \in \mathbb{R}^d$

ensemble des vecteurs à d coordonnées réelles

f : fonction objectif (quantifie la qualité d'une décision)

$$f : \mathbb{R}^d \to \mathbb{R}$$

x : variables de décision (paramètres sur lesquels on peut jouer pour changer la décision)

$\mathcal{X} \subseteq \mathbb{R}^d$: ensemble admissible (conditions sur x pour la décision soit acceptable)

NB : Dans ce cours, on considèrera surtout des problèmes de minimisation, car les problèmes de maximisation peuvent toujours être écrits comme problèmes de minimisation

$$\underset{x \in \mathbb{R}^d}{\text{maximiser}} \; f(x) \;\; \text{s.c.} \;\; x \in \mathcal{X} \quad \text{équivaut à} \quad \underset{x \in \mathbb{R}^d}{\text{minimiser}} \; (-f)(x) \;\; \text{s.c.} \;\; x \in \mathcal{X}$$

Solutions d'un problème d'optimisation

Dans la suite, on se concentre sur des problèmes sans contraintes ($\mathcal{X} = \mathbb{R}^d$)

$$\boxed{(P) \qquad \underset{x \in \mathbb{R}^d}{\text{minimiser}} \; f(x)}$$

" minimiser $f(x)$ par rapport à $x \in \mathbb{R}^d$ "

Définitions : • On dit que $x^* \in \mathbb{R}^d$ est une solution de (P) (ou encore un minimum global de (P)) si

$$\forall x \in \mathbb{R}^d, \quad f(x^*) \leq f(x)$$

L'ensemble des minima globaux de (P) est noté $\underset{x \in \mathbb{R}^d}{\operatorname{argmin}} f(x) \subseteq \mathbb{R}^d$

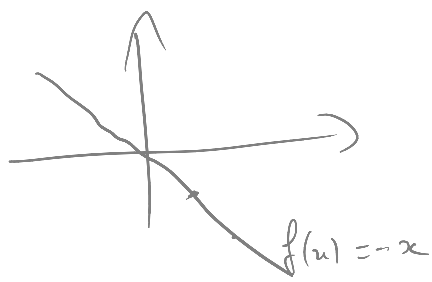
• La valeur minimale (aussi appelée la valeur optimale ou l'optimum) de (P) est notée $\min_{x \in \mathbb{R}^d} f(x) \in \mathbb{R} \cup \{-\infty, \infty\}$

↑ cas pathologiques : pas de solution

⇒ Lorsque le problème a une solution, on a

$$\min_{x \in \mathbb{R}^d} f(x) = f(x^*) \quad \forall x^* \in \underset{x \in \mathbb{R}^d}{\operatorname{argmin}} f(x)$$

Ex) $d = 1$, $f(x) = -x$



$\underset{x \in \mathbb{R}^d}{\text{minimiser}} \; f(x)$ n'a pas de solution

$$\Rightarrow \underset{x \in \mathbb{R}^d}{\operatorname{argmin}} f(x) = \emptyset \quad \text{et} \quad \min_{x \in \mathbb{R}^d} f(x) = -\infty$$

↳ En pratique, vérifier qu'un point/vecteur $x \in \mathbb{R}^d$ est une solution à partir de la définition est en général impossible. (infinité d'évaluations de f à réaliser)

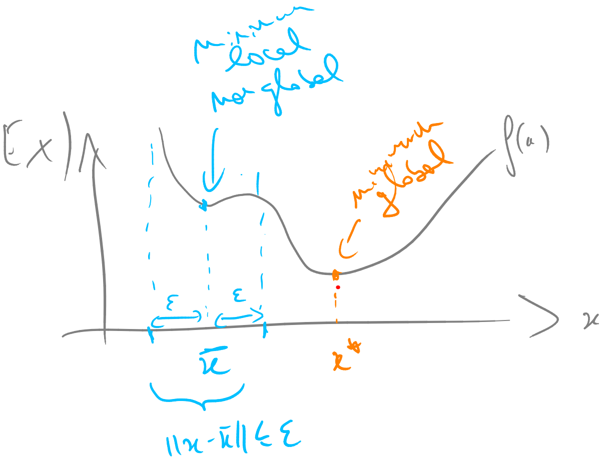
$\Rightarrow$ Deux possibilités :
- × chercher des solutions approchées
- × considérer des classes de fonctions pour lesquelles les solutions sont faciles à déterminer

<u>Définition</u> : Un point $\bar{x} \in \mathbb{R}^d$ est un <u>minimum local</u> de (P)

si $\exists\, \varepsilon > 0$, $\forall\, x \in \mathbb{R}^d$, $\|x - \bar{x}\| \leq \varepsilon \Rightarrow f(\bar{x}) \leq f(x)$

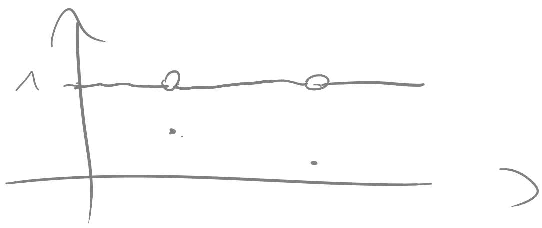
Norme (euclidienne) dans $\mathbb{R}^d$

$\forall\, v \in \mathbb{R}^d$, $v = \begin{bmatrix} v_1 \\ \vdots \\ v_d \end{bmatrix}$, $\|v\| = \sqrt{\sum_{i=1}^{d} v_i^2}$



<u>Remarque</u> : Tout minimum global est local, mais l'inverse n'est pas vraie en général

↳ Les minima locaux sont en général plus faciles à déterminer que des minima globaux (au sens pratique), mais vérifier qu'un point est un minimum local via la définition est aussi difficile que vérifier que c'est un minimum global



<u>But</u> : Trouver des classes de problèmes pour lesquels on peut caractériser les minima avec des objets finis

Fonctions C^1 et C^2

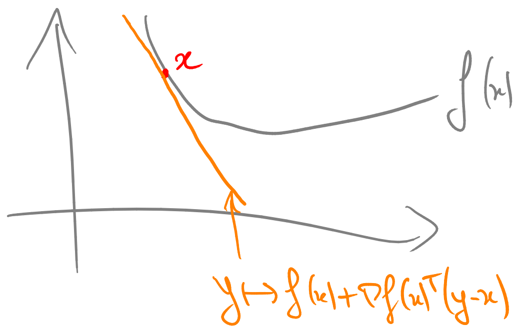
Déf : • $f : \mathbb{R}^d \to \mathbb{R}$ est C^1 si elle est continûment différentiable

$$\Longrightarrow \begin{cases} \forall x \in \mathbb{R}^d,\ \exists\ \underset{\text{gradient de } f \text{ en } x}{\underset{\uparrow}{\nabla f(x)}} \in \mathbb{R}^d \text{ et } \nabla f : \begin{array}{l}\mathbb{R}^d \to \mathbb{R}^d \\ x \mapsto \nabla f(x)\end{array} \text{ est continue.} \\[2em] \forall (x,y) \in (\mathbb{R}^d)^2,\ f(y) \simeq f(x) + \nabla f(x)^T (y-x) \\ \qquad\qquad \text{lorsque } \|y-x\| \text{ est petite} \end{cases}$$

$\forall (u,v) \in (\mathbb{R}^d)^2$

$u^T v = \sum_{i=1}^{d} u_i v_i$ "produit scalaire entre u et v"

f $C^1 \Rightarrow f$ est approchée localement par une fonction linéaire qui dépend du gradient



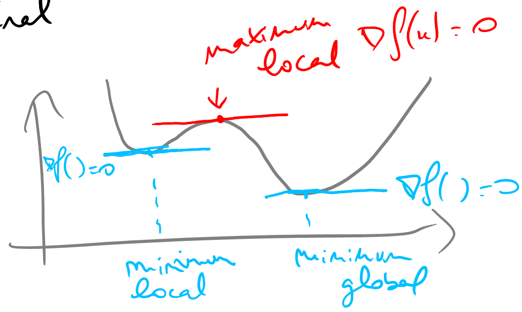
Théorème (Condition d'optimalité nécessaire à l'ordre 1)

Soit $f : \mathbb{R}^d \to \mathbb{R}$ (de classe) C^1. et $\bar{x} \in \mathbb{R}^d$

Si $\bar{x}$ est un minimum local de f **ALORS** $\nabla f(\bar{x}) = 0_{\mathbb{R}^d} \left(= \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \right) \updownarrow d$

↑ vecteur de taille finie

⚠ Ce n'est qu'une condition nécessaire (Si ... ALORS ...), il n'y a pas équivalence en général



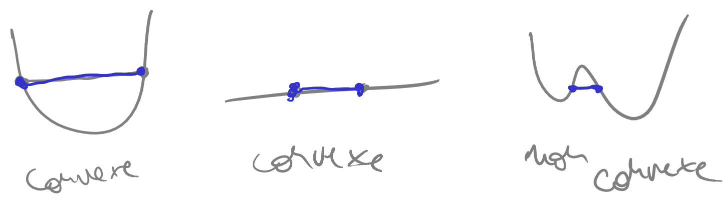
Un point stationnaire d'ordre 1, càd $\bar{x} \in \mathbb{R}^d$ tel que $\nabla f(\bar{x}) = 0_{\mathbb{R}^d}$ peut être

- un minimum local, voire global ;
- un maximum local, voire global ;
- ni l'un ni l'autre (point selle / point col)

Def : (Fonctions convexes)

• $f : \mathbb{R}^d \to \mathbb{R}$ est convexe si :

$$\forall (x,y) \in (\mathbb{R}^d)^2, \quad \forall \alpha \in [0,1], \quad f(\alpha x + (1-\alpha) y) \leq \underbrace{\alpha f(x) + (1-\alpha) f(y)}$$



Propriétés : Soit f convexe. Alors,

Tout minimum local de $\underset{x \in \mathbb{R}^d}{\text{minimiser}}\ f(x)$

1) Tout minimum local de f est un minimum global

2) $\forall \bar{x} \in \mathbb{R}^d, \quad \left[\nabla f(\bar{x}) = 0_{\mathbb{R}^d}\right] \Longleftrightarrow \left[\bar{x} \in \operatorname{argmin}_{x \in \mathbb{R}^d} f(x)\right]$

Équivalence !

Pour des fonctions C^1 convexes, le gradient permet de vérifier en temps de calcul fini qu'un point est un minimum global.

Def : $f : \mathbb{R}^d \to \mathbb{R}$ est de classe C^2 si elle est C^1 et que ∇f est aussi C^1

$\Longrightarrow \forall x \in \mathbb{R}^d, \quad \exists\, \nabla^2 f(x) \in \mathbb{R}^{d \times d}$ → matrices de taille $d \times d$: d lignes, d colonnes

telle que $f(y) \simeq f(x) + \nabla f(x)^T (y-x) + \frac{1}{2} (y-x)^T \nabla^2 f(x) (y-x)$

Lorsque $\|y-x\|$ est petite

$\nabla^2 f(x)$: matrice hessienne de f en x

$\forall x \in \mathbb{R}^d$, $\nabla^2 f(x) = \nabla^2 f(x)^T$ (matrice symétrique)

$\forall A \in \mathbb{R}^{d\times m}$, $A^T \in \mathbb{R}^{m\times d}$, $[A^T]_{ji} = [A]_{ij}$ $\forall i = 1..d$, $\forall j = 1..m$

$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 3}$ $\quad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \in \mathbb{R}^{3\times 2}$

Théorème : (Conditions d'optimalité à l'ordre 2)

(P) $\underset{x\in\mathbb{R}^d}{\text{minimiser}}\ f(x)$

Soit $f : \mathbb{R}^d \to \mathbb{R}$ de classe C^2 et $\bar{x} \in \mathbb{R}^d$.

Condition nécessaire : $[\bar{x}$ minimum local de $(P)]$

$$\Downarrow$$

pt stationnaire à l'ordre 2 $\rightarrow$ $[\nabla f(\bar{x}) = 0_{\mathbb{R}^d}$ et $\nabla^2 f(\bar{x}) \succeq 0]$

$\forall v \in \mathbb{R}^d$, $v^T \nabla^2 f(\bar{x}) v \geq 0$

"semi-définie positive"

Condition suffisante

$[\nabla f(\bar{x}) = 0_{\mathbb{R}^d}$ et $\nabla^2 f(\bar{x}) \succ 0]$

$v^T \nabla^2 f(\bar{x}) v > 0$ si $v \neq 0_{\mathbb{R}^d}$

$$\Downarrow$$

$[\bar{x}$ minimum local de $(P)]$

Pour les fonctions convexes

Si f convexe $+ C^2$, alors $\nabla^2 f(x) \succeq 0$ est toujours vrai

Vérifier $\nabla^2 f(x) \succeq 0$ $\forall x \in \mathbb{R}^d$ garantit que la fonction f est convexe.