Optimisation

Azgal ABICHOU

ESSAI, 2018

Plan du cours

Introduction et motivation

- Plusieurs données, Plusieurs paramètres, ...
- Tout est à optimiser !
- Un ingénieur doit savoir optimiser !
- Optimiser, c'est gagner : du temps, coût, de l'argent, ...

Ensembles convexes

On se donne un espace de Banach (evn complet) intuitivement un espace sans trous ! ou aucun point manquant !
Remarque : Tout evn de dimension finie est un banach.
Souvent dans ce cours, on va travailler dans R^n.

- Segments : Soient u et v dans $V \equiv R^n$ on a
  - $[u, v] = \{w = (1-t)u + tv, t \in [0, 1]\}$ = segment fermé d'extrémité u et v.
  - $]u, v[= \{w = (1-t)u + tv, t \in]0, 1[\}$ = segment ouvert d'extrémité u et v.
  - $[u, v[$, ...
- Ensembles convexes : Soit K une partie de V ; On dit que K est convexe si pour tout u et v dans K, $[u, v] \subset K$.
- Quelques propriétés :
  - Toute intersection de convexe est convexe
  - le vide est par convention convexe !
  - La réunion n'est toujours convexe.

Ensembles convexes

- Combinaison convexe : Pour toute famille finie d'éléments de V, $u_1, ..., u_n$ et tout système $\lambda_1, ..., \lambda_n \geq 0$ tel que $\Sigma\lambda_i = 1$, l'élément $\Sigma\lambda_i u_i$ s'appelle combinaison convexe des points $u_1, ..., u_n$.
- Proposition : Soit K convexe de V alors pour toute famille $u_1, ..., u_n$ de points de K et tout $\lambda_1, ..., \lambda_n$ de réels > 0 avec $\Sigma\lambda_i = 1$, le point $\Sigma\lambda_i u_i \in K$.
  Démontrer !

Ensembles convexes

- Enveloppe convexe : Soit K un se de V $K \subset V$. On appelle enveloppe convexe le + petit convexe contenant K et on la note $Co(K) \equiv$ l'$\bigcap$ de tous les convexes contenus dans K.
- Proposition : $Co(K) =$ l'ensemble de toutes les combinaisons convexes d'éléments de $K = \{\Sigma\lambda_i u_i \ \lambda_i \geq 0 \ \Sigma\lambda_i = 1 \ u_i \in K\}$. Démontrer !
- Remarque : A convexe $\Leftrightarrow Co(A) \equiv A$
- Exemple : $V = R^2$ et soit $K = \{(x, y) \in R^2 \quad -1 \leq x \leq 1 \ \ y \leq x^3\}$. Trouver $Co(K)$.

Fonctions convexes

Soit K une partie convexe de V et $f : K \to \bar{R}$

- Définitions :
  - f est convexe sur K si $\forall x, y \in K \quad t \in [0,1]$, $f((1-t)x+ty) \leq (1-t)f(x)+tf(y)$.
  - f est strictement convexe si $\forall x \neq y \in K \quad t \in]0,1[$, $f((1-t)x+ty) < (1-t)f(x)+tf(y)$.
  - f est fortement convexe ou ($\alpha-$convexe) sur K si il existe un $\alpha > 0$ et $t \in [0,1]$ tel que $\frac{\alpha}{2}t(1-t)\|x-y\|^2 + f((1-t)x+ty) \leq (1-t)f(x)+tf(y)$
- Convention : f est à valeurs dans $\bar{R}$ donc $(1-t)f(x)+tf(y)$ peut être indéterminée si $f(x)=+\infty$ et $f(y)=-\infty$ ou le contraire ! on impose donc par convention pour $\forall a \in \bar{R}$, $a+(+\infty)=+\infty$ et $\forall a \in \bar{R}$, $a \neq +\infty$, $a+(-\infty)=-\infty$.
- Remarque : forte convexité $\Rightarrow$ stricte convexité $\Rightarrow$ convexité.

Fonctions convexes

- Proposition : Soit K une partie convexe de V et $f : K \to \bar{R}$ convexe sur K, $\forall \alpha \in \bar{R}$ les ensembles $\{u \in K, f(u) \leq \alpha\}$ et $\{u \in K, f(u) < \alpha\}$ sont convexes.
  Démontrer !
- Attention ! : La réciproque est en général fausse !
  Un contre-exemple !
- Domaine effectif : Soit $f : K \to \bar{R}$, on appelle domaine effectif de f et on note $dom(f) = \{u \in K, f(u) < +\infty\}$.
- Proposition : si f est convexe $\Rightarrow dom(f)$ est convexe.
- Attention ! : La réciproque est en général fausse !
  Un contre-exemple !

Fonctions convexes

- Proposition : Soit K une partie convexe de V et $f : K \to \bar{R}$ convexe sur K, $\forall \alpha \in \bar{R}$ les ensembles $\{u \in K, f(u) \leq \alpha\}$ et $\{u \in K, f(u) < \alpha\}$ sont convexes.
  Démontrer !
- Attention ! : La réciproque est en général fausse !
  Un contre-exemple !
- Domaine effectif : Soit $f : K \to \bar{R}$, on appelle domaine effectif de f et on note $dom(f) = \{u \in K, f(u) < +\infty\}$.
- Proposition : si f est convexe $\Rightarrow$ $dom(f)$ est convexe.
- Attention ! : La réciproque est en général fausse !
  Un contre-exemple !

Fonctions convexes

- Fonction indicatrice : Soit K une partie de V. On appelle fonction indicatrice χ_K de K, la fonction définie de V dans $\bar{R}$ par $\chi_K(v) = 0$ si v dans K et $\chi_K(v) = +\infty$ sinon.
- Remarque : on a $dom(\chi_K) = K$.
- Proposition : K convexe $\Leftrightarrow \chi_K$ convexe. Démontrer !
- Fonction propre : Soit $f : V \to \bar{R}$. On dit que f est propre si $f(v) > -\infty$ pour tout v dans V et $\exists v_0 \in V$ avec $f(v_0) < +\infty$.
- Remarque : Si $K \subset V$ non vide alors χ_K est propre.
- Epigraphe : On appelle épigraphe de f noté $epi(f)$ le se de $V \times R$ défini par $epi(f) = \{(x, \alpha) \in V \times R, f(x) \leq \alpha\}$.
- Proposition : f convexe $\Leftrightarrow epi(f)$ est convexe. Démontrer !.

Fonctions convexes

- Remarque : Une fonction de deux variables (x, y) qui est convexe par rapport à x pour tout y et convexe par rapport à y pour tout x n'est pas nécessairement convexe par rapport au couple (x, y). Un exemple !
- Une autre remarque : La forte convexité est le cadre agréable pour de nombreux problèmes d'optimisation car elle donne facilement l'existence, l'unicité et des algorithmes très performants mais tout de même c'est une hypothèse très forte ! la bonne chose c'est elle inclut le cas des fonctions quadratiques, important en pratique dans plusieurs domaines ! !

Convexité et continuité

Soit $f : R^n \to \bar{R}$.

- Proposition : Soit $f : R^n \to \bar{R}$. On suppose qu'il existe $x_0 \in R^n$ et $M \in R$ tels que pour tout x dans un voisinage de R^n on a $f(x) \leq M$ et $f(x_0) > -\infty$ alors f est continue au voisinage de x_0. C'est à dire une fonction convexe bornée est localement continue !
  Un petit exercice d'analyse à faire pour les passionnés !
- Remarque : On peut noter au passage que, même en dimension finie, les fonctions convexes ne sont pas nécessairement continues au bord de leur domaine de définition ! ! Un exemple !

Convexité et continuité

Soit $f : R^n \to \bar{R}$.

- Semi-continuité inférieure (sci) : On dit que f est semi-continue inférieurement si pour tout $\alpha \in R$, l'ensemble $\{x | f(x) \leq \alpha\}$ est fermé.
- Semi-continuité supérieure (scs) : On dit que f est semi-continue supérieurement si pour tout $\alpha \in R$, l'ensemble $\{x | f(x) \geq \alpha\}$ est fermé.
- Remarque :
  - f est sci $\Leftrightarrow$ $-f$ est scs !
  - f est sci et scs $\Leftrightarrow$ f est continue !
- Propriété : f est sci $\Leftrightarrow$ pour tout x s'il existe $x_n \longrightarrow x$ alors $f(x) \leq \liminf f(x_n)$ $\Leftrightarrow$ $epi(f)$ est fermé !

Dérivabilité

Soit $f : R^n \to \bar{R}$.

- Dérivée directionnelle : On dit que f admet en u une dérivée dans la direction v si $\lim_{t\to 0^+} \dfrac{f(u+tv)-f(u)}{t}$ existe, et on note donc cette limite $f'(u,v)$.
- Gâteaux-dérivabilité : Si $f'(u,v)$ est linéaire par rapport à v, on dira alors que f est Gateaux-dérivable en u et on note cette dérivée $f'(u)$. C'est à dire si f est Gâteaux-dérivable en u, on a $\lim_{t\to 0^+} \dfrac{(f(u+tv)-f(u))}{t} = f'(u)v$ pour tout v !
- Remarques :
  - Si $f'(u)$ existe, elle est unique !
  - Si on est dans R, Gâteaux-dérivabilité coïncide avec la dérivabilité ordinaire (au sens classique !)

Dérivabilité

- Fréchet-dérivabilité : On dit que f est dérivable au sens de Fréchet en u et on note $f'(u)$ sa dérivée si $\|f(u+h) - f(u) - f'(u)h\| = o(\|h\|)$ quand $\|h\| \to 0$. Ceci est équivaut à $\lim_{\|h\|\to 0} \dfrac{\|f(u+h) - f(u) - f'(u)h\|}{\|h\|} = 0$
- Remarque : Fréchet dérivable $\Rightarrow$ Gâteaux-dérivable $\Rightarrow$ dérivable dans toute direction ! La réciproque est évidemment fausse !
- Exemple : Dans R^n, l'application $x \mapsto \|x\|$ est dérivable dans toute direction en 0 mais pas Gâteaux-dérivable en 0 ! !

Convexité et dérivabilité

Soit $f : K \in R^n \to \bar{R}$, K convexe non vide, f Gâteaux-dérivable dans K.

- Proposition 1 : On a équivalence entre :
  - f est α-convexe sur K (i.e. $\frac{\alpha}{2}t(1-t)\|x-y\|^2 + f((1-t)x+ty) \leq (1-t)f(x)+tf(y)$)
  - $f(y) - f(x) \geq (f'(x), y-x) + \frac{\alpha}{2} \parallel y-x \parallel^2$ pour tous x et y dans K.

  Démontrer !
- Proposition 2 : On a équivalence entre :
  - f est α-convexe sur K
  - $(f'(y) - f'(x), y-x) \geq \alpha \parallel y-x \parallel^2$ pour tous x et y dans K.

  Démontrer !
- Remarque : Récap ! On a équivalence entre f est α-convexe sur $K \Leftrightarrow f(y) - f(x) \geq (f'(x), y-x) + \frac{\alpha}{2} \parallel y-x \parallel^2 \Leftrightarrow$ $(f'(y) - f'(x), y-x) \geq \alpha \parallel y-x \parallel^2$ pour tous x et y dans K.

Dérivabilité d'ordre 2

Soit $f : K \in R^n \to \bar{R}$.

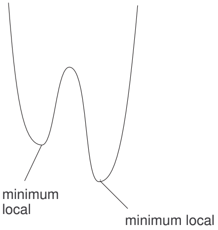
- Définition : On dit que f est deux fois Gâteaux-dérivable en x dans K, si f est dérivable en x et s'il existe un opérateur linéaire continue noté $f"(x)$ tel que :
  $(f"(x)u, v) = \lim_{\|t\| \to 0^+} \dfrac{(f'(x+tu), v) - (f'(x), v)}{t}$.
  $f"(x)$ est appelé le Hessien associé à f en x.
- Proposition : On a équivalence entre :
  - f est α-convexe sur K.
  - $(f"(x)(u-v), (u-v)) \geq \alpha \| u - v \|^2$ pour tous x et y dans K.

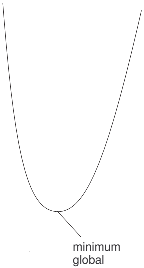
  Démontrer !
- Exercice : Montrer que $f(x) = \dfrac{1}{2}(Ax, x) - (b, x) + c$ où $x \subset R^n$ est dérivable et calculer $f'(x)$ et $f"(x)$.

Optimisation convexe

Soit $f : K \in R^n \to \bar{R}$.

- Minimum local : On dit que x^* est un minimum local de f s'il existe un voisinage $O(x^*)$ de x^* tel que on a $f(x) \geq f(x^*)$ pour tout $x \in O(x^*)$.
- Minimum global : On dit que x^* est un minimum global de f sur K si pour tout x dans K, on $f(x) \geq f(x^*)$.
- Remarque : Si f est convexe, un minimum local est forcément global !.





Conditions d'existence

- Résultat classique : Si f est continue sur K compact alors le problème $(\mathcal{P})\begin{cases} \min f(x) \\ x \in K \end{cases}$ admet au moins une solution.
- Un autre résultat d'existence : Si f est continue (ou s.c.i suffit !), propre sur K fermé non vide de R^n et si $\lim f(x) = +\infty$ quand $\|x\|$ tend vers ∞ alors le problème $(\mathcal{P})\begin{cases} \min f(x) \\ x \in K \end{cases}$ admet au moins une solution.
- Remarque : Si f est strictement convexe alors on aura l'unicité !

Conditions d'optimalité

Soit $f : K \in R^n \to \bar{R}$. K convexe non vide, f Gâteaux-dérivable dans K.

- CNS du premier ordre : x^* solution du problème $(\mathcal{P}) \begin{cases} \min f(x) \\ x \in K \end{cases} \Leftrightarrow (f'(x^*), y - x^*) \geq 0$ pour tout y dans K.
- Corollaire : Si $K = R^n$ ou K un s.e.v de R^n x^* solution du problème $(\mathcal{P}) \begin{cases} \min f(x) \\ x \in K \end{cases} \Leftrightarrow f'(x^*) = 0$.
- Remarque : Si $K = R^n$ ou K un s.e.v de R^n et si f n'est pas convexe, la condition $f^{'}(x^*) = 0$ n'est pas suffisante !
  Un contre-exemple !
  Penser à $f(x) = x^3$ sur R !

Optimisation sans et avec contraintes

Optimiser sans contraintes !

$$(\mathcal{P}) \begin{cases} \min f(x), \\ x \in R^n \end{cases}$$

Optimiser avec contraintes !

$$(\mathcal{P}) \begin{cases} \min f(x), \\ x \in \mathrm{K} \end{cases}$$

- Si f et K sont convexes, alors le problème est dit problème convexe.
- CN d'optimalité sans contraintes ! : f est dérivable. Si f admet un extremum local en x^* alors $f^{'}(x^*) = 0$.
- CN d'optimalité avec contraintes ! :f est dérivable. Si f admet un extremum local en x^* alors $(f^{'}(x^*), x - x^*) \geq 0$ pour tout $x \in K$.
- La CN devient suffisante dans le cas convexe !

- CN du 2ème ordre ! : f est deux fois dérivables au voisinage de (x^*). Si f admet un minimum local x^* alors le Hessien $f''(x^*)$ est SDP (semi-défini positif) !
- CS du 2ème ordre ! : $\begin{cases} f'(x) = 0 \\ f''(x^*) DP \end{cases} \Rightarrow x^*$ est un minimum local !
- Exemples ! :
  - $\min f(x, y) = x^3 + 2y^3 - 4xy$ pour $x \in R^2$.
  - $\min f(x, y) = \sqrt{x} + \sqrt{y} - 4x - 4y$ pour $x \in R^2$

Optimisation avec contraintes

Le problème est de la forme $(\mathcal{P}) \begin{cases} \min f(x) \\ g_i(x) = 0, \ \ 1 \leq i \leq p \end{cases}$

-
- Exemple ! : Soit $(\mathcal{P}) \begin{cases} \max xy, \\ 3x + 5y = 8 \end{cases}$ On peut le résoudre par substitution (y en fonction de x) !
- Mais cette méthode ne s'applique pas toujours ! faut une méthode plus générale $\Rightarrow$ la méthode des multiplicateurs de Lagrange !

Optimisation avec contraintes d'égalité

Soit le problème avec p contraintes d'égalité !

$$(\mathcal{P}) \begin{cases} \min f(x) \\ g_i(x) = 0, \;\; 1 \leq i \leq p \end{cases}$$

- CN du premier ordre ! :
  - f et g_i sont de classe $\mathcal{C}^1$ sur un voisinage de x^*.
  - $g_i^{'}(x^*)$ sont linéairement indépendants. (les contraintes sont dites qualifiées ou régulières !)

  Alors si x^* solution locale du problème $(\mathcal{P})$ alors il existe λ_1, λ_2,.., λ_p tels que $f'(x^*) + \sum_{i=1}^{p} \lambda_i g_i'(x^*) = 0$
- CS du 2ème ordre ! :
  - f et g_i sont de classe $\mathcal{C}^2$ sur un voisinage de x^*.
  - $g_i^{'}(x^*)$ sont linéairement indépendants.

  Si $\exists\ \lambda_1, \lambda_2, \ldots \lambda_p$ tels que $f'(x^*) + \sum_{i=1}^{p} \lambda_i g_i'(x^*) = 0$ et $\forall h$ tel que $g_i^{'}(x^*).h = 0$, $h^t[f"(x^*) + \sum_{i=1}^{p} \lambda_i g"_i(x^*)]h > 0$ Alors x^* est solution locale du problème $(\mathcal{P})$.

Cas convexe ! CN → CNS

On suppose que :

- f convexe de classe $\mathcal{C}^1$ sur un voisinage de x^*.
- g_i sont affine pour tout $1 \leq i \leq p$.
- $g_i'(x^*)$ sont linéairement indépendants.
- $g_i(x^*) = 0$ pour tout $1 \leq i \leq p$.

on a

x^* solution du problème $(\mathcal{P}) \iff$

$\exists\ \lambda_1, \lambda_2, \ldots \lambda_p$ tels que $f'(x^*) + \sum_{i=1}^{p} \lambda_i g_i'(x^*) = 0$

Exemple ! :

$$(\mathcal{P}) \begin{cases} \min x^2 + 2y^2, \\ x + y = 1 \end{cases}$$

Optimisation avec contraintes d'inégalité

Soit le problème sur R^n $(\mathcal{P}) \begin{cases} \min f(x) \\ g_i(x) = 0, \;\; 1 \leq i \leq p \\ h_k(x) = 0, \;\; 1 \leq k \leq q \end{cases}$

- p contraintes d'égalité et q contraintes d'inégalité.

Conditions d'optimalité ! :

- CN du premier ordre : On suppose que :
  - f convexe de classe $\mathcal{C}^1$ sur un voisinage de x^*.
  - g_i sont affine et h_k sont convexes.

  avec
  - $g_i^{'}(x^*)$, $\{h_j'(x^*) \quad h_j(x^*) = 0\}$ sont linéairement indépendants.
  - $\exists\ \overline{x}$ tel que $g_i(\overline{x}) = 0$ et $h_k(\overline{x}) < 0 \rightarrow$ Condition de Slater !.
  - $g_i(x^*) = 0\ \forall\ 1 \leq i \leq p$ et $h_k(x^*) \leq 0\ \forall\ 1 \leq k \leq q$.

  On a x^* solution du problème $(\mathcal{P}) \Longrightarrow \exists\ \lambda_1, \lambda_2, \ldots \lambda_p$ et $\mu_1, \mu_2, \ldots \mu_q$ tels que
  - $f^{'}(x^*) + \sum_{i=1}^{p} \lambda_i g_i^{'}(x^*) + \sum_{k=1}^{q} \mu_k h_k^{'}(x^*) = 0$
  - $\mu_k h_k(x^*) = 0$ pour tout $1 \leq k \leq q$.
  - $\mu_k \geq 0$ pour tout $1 \leq k \leq q$.

Quelques remarques et un exemple !

- λ_i et μ_j sont les multiplicateurs de Lagrange.
- Les trois dernières conditions sont dites Conditions de Karush-Kuhn-Tucker (KKT) !
- Cas convexe : CN $\rightarrow$ CNS
  - f convexe, g_i affines et h_k convexes !
  - $g_i^{'}(x^*)$ sont linéairement indépendants, $g_i(x^*) = 0$ et $h_k(x^*) \leq 0$, Condition de Slater satisfaite !

  on a x^* solution du problème $(\mathcal{P}) \Longleftrightarrow \exists\ \lambda_1, \lambda_2, \ldots \lambda_p$ et $\mu_1, \mu_2, \ldots \mu_q$ tels que
  - Conditions de KKT satisfaites !.
- Pour un problème de maximisation, on a
  - f concave et $\mu_j \leq 0$

Des exemples !

- $$\begin{cases} \min 2x^2 + y^2 + 2xy - 10x - 10y, \\ 3x + y \leq 6 \\ x^2 + y^2 \leq 5 \end{cases}$$
- $$\begin{cases} \min 3x^2 + y^2, \\ x + y = \frac{1}{2} \\ x^2 + y^2 \leq \frac{1}{2} \end{cases}$$
- $$\begin{cases} \min x^2, \\ 2x + 5 \leq 0 \end{cases}$$

Au travail, les solutions !