Optimisation convexe

26 janvier 2017

Table des matières

Première partie

Ensembles convexes

1 Définitions et premières propriétés

Définition: Ensemble affine

Soit E un espace vectoriel réel. Un sous-ensemble A de E est affine si

$$\forall x \in A; y \in A; \forall \alpha \in \mathbb{R}, \alpha x + (1-\alpha)y \in A$$

Autrement dit, un ensemble affine contient toujours la droite passant par deux de ses points x et y

Définition: Ensemble affine

Soit E un espace vectoriel réel. Un sous-ensemble C de E est convexe si

$$\forall x \in C; y \in C; \forall \alpha \in [0,1], \alpha x + (1-\alpha)y \in C$$

Autrement dit, un ensemble affine contient toujours le segment $[x, y]$.

Définition: Simplexe

On appelle simplexe de $\mathbb{R}^n$ le sous-ensemble

$$\Delta_n = \left\{ \alpha \in \mathbb{R}^n; \alpha_i \geq 0, i = 1, ..., n, \sum_{i=1}^{n} \alpha_i = 1 \right\}$$

Définition: Combinaison convexe

On appelle combinaison convexe de n points $\{x_i\}_{i=1}^{n}$ tout point y s'écrivant

$$y = \sum_{i=1}^{n} \alpha_i x_i \text{ avec } \alpha \in \Delta_n$$

Théorème:

Un sous ensemble C de E est convexe si et seulement s'il contient toutes les combinaisons convexes de ses éléments.

Proposition: Opérations conservant la convexité

— Si $(C_i)_{i\in I}$ est une famille quelconque de convexes de E, alors l'intersection $\cap_{i\in I} C_i$ est encore un convexe.
— Pour $a \in E$, le translaté $a + C = \{a + x; x \in C\}$ est convexe
— Le produit cartésien de $C \in E$ et $C' \in E'$, ie $C \times C' = \{(x; y); x \in C; y \in C'\}$ est un sous ensemble convexe de $E \times E'$
— La somme $C_1 + C_2 = \{x_1 + x_2; x_1 \in C_1; x_2 \in C_2\}$ de deux ensembles convexes C_1 et C_2 est convexe.
— L'union de sous-ensembles convexes n'est en général pas convexe, mais l'union croissante de convexes est convexe.
— Soit $f : \mathbb{R}^n \to \mathbb{R}^k$ une fonction affine, et $C \subseteq \mathbb{R}^n$, $S \subseteq \mathbb{R}^k$ deux convexes. Alors $f(C)$ et $f^{-1}(S)$ sont également convexes.
— La somme partielle de deux convexes S et $C \subseteq \mathbb{R}^n \times \mathbb{R}^m$:

$$\{(x, y_1 + y_2)|(x, y_1) \in S,\ (x, y_2) \in C\}$$

est aussi convexe.
— L'image d'un convexe $C \subset \mathbb{R}^{n+1}$ par la fonction perspective $P : \mathbb{R}^n \times \mathbb{R}_*^+ \to \mathbb{R}^n$:

$$P(z, t) = z/t$$

est convexe. De même, la pré-image d'un convexe S par P

$$P^{-1}(S) = \{(x, t) \in \mathbb{R}^n \times \mathbb{R}_*^+ | x/t \in C\}$$

est convexe.

Définition:

Soit C un ensemble convexe. Une partie convexe F de C est appelée face (ou partie extrémale) de C si la propriété suivante est vérifiée

$$\left.\begin{array}{c}(x, y) \in C \times C \text{ et} \\ \exists \alpha \in]0, 1[\text{ tel que } \alpha x + (1-\alpha)y \in F\end{array}\right\} \Rightarrow [x, y] \in F$$

On appelle point extrémal une face réduite à un seul point. En d'autres termes, $\bar{x} \in C$ est un point extrémal de C s'il n'est pas possible d'avoir $\bar{x} = \alpha y + (1-\alpha)z$ avec y et z deux points distincts de C et $\alpha \in]0, 1[$. On note $Ext(C)$ l'ensemble des points extrémaux de C.

Définition: Polyèdre

Soit $A \in \mathbb{R}^{m\times n}$, $b\mathbb{R}^m$, $C \in \mathbb{R}^{p\times n}$, $d \in \mathbb{R}^d$. On définit le polyèdre $\mathcal{P}$ par :

$$\mathcal{P} = \{x \in \mathbb{R}^n | Ax \leq b,\ Cx = d\}$$

2 Enveloppe affine et enveloppe convexe

Définition: Enveloppe affine

Soit A une partie de E. L'enveloppe affine de A, notée $Aff(A)$, est l'intersection de tous les espaces affines contenant A.
On appelle *dimension affine* d'un ensemble C la dimension de $\mathrm{Aff}(C)$.

Proposition:

Soit A une partie de E. On a :

$$Aff(A) = \left\{ \sum_{i=1}^{n} \alpha_i x_i ; n \geq 1, x_i \in A, \alpha_i \in \mathbb{R}, \sum_{i=1}^{n} \alpha_i = 1 \right\}$$

Définition: Enveloppe convexe

Soit A une partie de E. L'enveloppe convexe de A, notée $conv(A)$, est l'intersection de tous les espaces convexes contenant A

Proposition:

Soit A une partie de E. On a :

$$Conv(A) = \left\{ \sum_{i=1}^{n} \alpha_i x_i ; n \geq 1, x_i \in A, \alpha_i \geq 0, \sum_{i=1}^{n} \alpha_i = 1 \right\}$$

Théorème: Carathéodory

Soit A une partie d'un espace vectoriel E de dimension n. Alors tout élément de $conv(A)$ peut s'écrire comme une combinaison convexe de $n+1$ éléments de A.

Définition: Simplexe

Soient $v_0, ..., v_k \in \mathbb{R}^n$, $k+1$ vecteurs affinement indépendants, i.e. $v_1 - v_0$, ...,$v_k - v_0$ sont linéairement indépendants. On appelle simplexe l'ensemble défini par :

$$C = \mathrm{conv}\{v_0, ..., v_k\}$$

Propriété:

Un simplexe est un polyèdre

3 Propriétés topologiques des convexes

3.1 Ouverture et fermeture des convexes

Théorème:

Soit C un ensemble convexe. Alors son intérieur $int(C)$ et son adhérence $\overline{C}$ sont aussi convexes.

3.2 Intérieur relatif

En analyse convexe, on rencontre souvent des ensembles convexes dont l'intérieur est vide : c'est le cas d'un segment dans $\mathbb{R}^2$. Il est donc utile d'introduire la notion d'intérieur relatif.

Définition: Intérieur relatif

Soit P une partie d'un espace vectoriel E. L'intérieur relatif de P, noté $ri(P)$, est son intérieur dans son enveloppe affine $Aff(P)$,munie de la topologie induite de celle de E, i.e.

$$ri(P) = \{x \in P; \exists r > 0; (B(x,r) \cap Aff(P)) \subset P\}$$

Théorème:

Soient E un espace vectoriel de dimension finie et C un convexe non vide de E. Alors $ri(C)$ est non vide.

Lemme:

Soient E un espace vectoriel de dimension finie et C un convexe non vide. Alors

$$x \in ri(C) \text{ et } y \in \overline{C} \Rightarrow [x; y[\subset ri(C)$$

Ainsi, un point $x \in E$ est dans l'intérieur relatif de C si et seulement si pour tout $y \in C$ (ou $y \in Aff(C)$), il existe $\alpha > 1$ tel que $(1 - \alpha)y + \alpha x \in C$

4 Opérations sur les ensembles convexes

4.1 Projection sur un convexe fermé

Théorème: Projection sur un convexe fermé

Soient H un espace de Hilbert et x un élément de H. Soit également C un sous-ensemble convexe fermé de H. Il existe un unique point $y \in C$ tel que

$$\|y - x\| = \min_{z \in C} \|z - x\|$$

Cet élément y est appelé la projection de x sur C et sera noté $P_C(x)$. Il est caractérisé par l'inéquation suivante

$$\forall z \in C, \langle x - P_c(x), z - P_C(x) \rangle \leq 0$$

Propriété: de la projection

L'application projection : $x \mapsto P_C(x)$ sur un convexe fermé non vide C possède les propriétés suivantes :

1. $\forall x_1, x_2 \in H$, $\langle P_C(x_2) - P_C(x_1), x_2 - x_1 \rangle \geq \|P_C(x_2) - P_C(x_1)\|^2$
2. elle est monotone : $\forall x_1, x_2 \in H$, $\langle P_C(x_2) - P_C(x_1), x_2 - x_1 \rangle \geq 0$
3. elle est Lipschitzienne de constante 1 :

$$\forall x_1, x_2 \in H, \|P_C(x_2) - P_C(x_1)\| \leq \|x_2 - x_1\|$$

4.2 Séparation des ensembles convexes

Un outil essentiel en analyse convexe est le théorème de Hahn-Banach sur la séparation des ensembles convexes. Etant donné un espace de Hilbert H,la séparation de deux convexes se fait géométriquement dans H en utilisant un hyperplan affine K de la forme

$$K = \{x \in H, \langle \xi, x \rangle = \alpha\}$$

où $\xi \in H$ est non nul et $\alpha \in \mathbb{R}$.

Définition:

On dit qu'un hyperplan $K := \{x \in H; \langle \xi, x \rangle = \alpha\}$ sépare deux convexes C_1 et C_2 si l'on a

$$\forall x_1 \in C_1, \forall x_2 \in C_2, \langle \xi, x_1 \rangle \leq \alpha \leq \langle \xi, x_2 \rangle$$

On dit qu'un hyperplan $K := \{x \in H; \langle \xi, x \rangle = \alpha\}$ sépare strictement deux convexes C_1 et C_2 s'il existe deux scalaires α_1 et α_2 tels que $\alpha_1 < \alpha < \alpha_2$ et

$$\forall x_1 \in C_1, \forall x_2 \in C_2, \langle \xi, x_1 \rangle \leq \alpha_1 < \alpha_2 \leq \langle \xi, x_2 \rangle$$

Remarque:

Une condition nécessaire et suffisante pour que C_1 et C_2 puisse être séparé par un hyperplan est qu'il existe un $\xi \in H$ non nul tel que

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

Une condition nécessaire et suffisante pour que C_1 et C_2 puisse être séparé strictement par un hyperplan est qu'il existe un $\xi \in H$ tel que

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

Théorème: Séparation d'un convexe et d'un point

Soient H un espace de Hilbert, C un sous-ensemble convexe fermé de H et $x \notin C$. Alors il existe $r \in H$ tel que

$$sup_{z \in C} \langle r, z \rangle < \langle r, x \rangle$$

Théorème: Séparation de deux convexes

Soient H un espace de Hilbert et C_1 et C_2 deux convexes non vides disjoints de H, l'un étant fermé et l'autre étant compact. Alors on peut séparer strictement C_1 et C_2.

Théorème: Séparation de deux convexes en dimension finie

Soient H un espace de Hilbert de dimension finie et C_1 et C_2 deux convexes non vides disjoints de H. Alors on peut séparer C_1 et C_2 au sens large, i.e., il existe $\xi \in H$ non nul tel que

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

4.3 Enveloppe convexe fermée

L'enveloppe convexe d'un fermé n'est pas nécessairement fermée.
Exemple : Dans $\mathbb{R}^2$, $C = \{xy \geq 1\} \cup \{0\}$: fermé.
$\mathrm{conv}(C) = \{x > 0, y > 0\} \cup \{0\}$: non fermé.

Définition:

$A \subset E$. On définit l'enveloppe convexe fermée, noté $\overline{conv}(A)$, comme l'intersection de tous les convexes fermés contenant A.

Propriété:

- $A_1 \subset A_2 \Rightarrow \overline{conv}(A_1) \subset \overline{conv}(A_2)$
- $A \subset conv(A) \subset conv(\bar{A}) \subset \overline{conv}(A)$ et $\overline{conv}(A) = \overline{conv}(\bar{A}) = \overline{conv(A)}$

Définition:

Soit H un Hilbert.
Un demi-espace fermé de H est un ensemble de la forme :

$$H^-(\xi, \alpha) = \{x \in H; (x, \xi) \leq \alpha\}$$

où $\xi \in H \neq \{0\}$ et $\alpha \in \mathbb{R}$

Proposition:

$\overline{conv}(A)$ est l'intersection de tous les demi-espaces fermés contenant A.

Corollaire:

Soit C un ensemble convexe.
Alors l'intersection de tous les demi-espaces fermés contenant C est $\overline{C}$.

Corollaire:

C convexe fermés $\Leftrightarrow$ C est l'intersection de tous les demi-espaces fermés contenant C.

Théorème:

Soient H de dimension finie et A un compact de H. Alors $conv(A)$ est compact.

5 Cônes convexes

Définition: Cône

Un ensemble C est un cône si $\lambda \in \mathbb{R}_+$, $\forall x \in C$, $\lambda x \in C$

Définition: Enveloppe conique

Soit $A \subset E$. L'enveloppe conique A, notée $cone(A)$, est l'intersection de tous les coônes convexes contenant A.

Définition: Combinaison conique

On appelle combinaison conique d'élements de A un point x tel que $x = \sum_{i=1}^{n} \lambda_i x_i$, $\lambda_i \geq 0$, $x_i \in A$

Proposition:

— C est un cône convexe si et seulement s'il contient toutes les combinaisons coniques de ses éléments.

—
$$cone(A) = \left\{ \sum_{i=1}^{n} \lambda_i x_i, n \in \mathbb{N}^*, \lambda_i \geq ^\circ, x_i \in A \right\}$$

Définition: Enveloppe conique fermée

On définit l'enveloppe conique fermée de A, notée $\overline{cone}A$, comme étant l'intersection de tous les cônes convexes fermés contenant A.

Propriété:

— $A \subset B \Rightarrow \overline{cone}(A) \subset \overline{cone}(B)$

— $A \subset cone(A) \subset cone(\bar{A}) \subset \overline{cone}(A)$ et $\overline{cone}(A) = \overline{cone}(\bar{A}) = \overline{cone(A)}$

Définition: Cône induit par une norme

Soit $\|\cdot\|$ une norme (quelconque) sur $\mathbb{R}^n$. On définit le cône induit par cette norme par :

$$C = \{(x,t) | \|x\| \leq t\} \subseteq \mathbb{R}^{n+1}\}$$

5.1 Cône propre et inégalités généralisées

Définition: Cône propre

Un cône $K \subseteq \mathbb{R}^n$ est dit *propre* si
- K est convexe
- K est fermé
- K est solide, i.e. il est d'intérieur non vide
- K est pointé, ce qui signifie qu'il ne contient aucune droite, ou autrement dit

$$x \in K,\ -x \in K \implies x = 0$$

Définition: Équations généralisées

On définit un pré-ordre sur $\mathbb{R}^n$ de la façon suivante : à un cône propre K, on associe l'ordre partiel défini par

$$x \preccurlyeq_K y \Leftrightarrow y - x \in K$$

On définit aussi un ordre partiel strict par :

$$x \prec_K y \Leftrightarrow y - x \in \text{int } K$$

Propriété:

La relation d'ordre $\preccurlyeq_K$ est transitive, refléxive, antisymétrique, stable par addition et multiplication par un scalaire positif.

5.1.1 Minimum et élément minimal

Définition: Élément minimum et minimal

$x \in S$ est un minimum de S (en vue de la relation $\preccurlyeq_K$) si pour tout $y \in S$, $x \preccurlyeq_K y$.
x est un élément minimal de S si pour $y \in S$, $y \preccurlyeq_K x$ seulement si $y = x$.

Propriété:

- S'il existe, le minimum est unique.
- $x \in S$ est un minimum de S si et seulement si $S \subset x + K$.
- $x \in S$ est un élément maximal de S si et seulement si $(x - K) \cap S = \{x\}$

5.2 Cône normal

Définition:

Soient H de Hilbert, $C \subset H$, $x \in C$.
On définit le cône normal à C en x, noté $\mathcal{N}_x C$ ou $\mathcal{N}_C(x)$ par :

$$\mathcal{N}_C(x) = \{d \in H; (d, y - x) \leq 0 \forall y \in C\}$$

Les éléments de $\mathcal{N}_x C$ sont appelés les normales à C en x.

Proposition:

Soit H de Hilbert de dimension fnie.
Si $C \subset H$ et $x \in \partial C$, alors $\mathcal{N}_x C$ contient au moins un élément non nul.

Remarque : Le résultat reste vrai en dimension infini si $\mathring{C}$ est non vide.

5.3 Cône dual

Définition: Cône dual, bidual, polaire

Soit $P \subset H$. On appelle cône dual de P, noté P^*, l'ensemble :

$$P^* = \{x \in H; (x, y) \geq 0 \ \forall y \in P\}$$

On appelle cône bidual de P : $P^{**} = (P^*)^*$
On appelle cône polaire (ou dual négatif) P^- l'ensemble

$$P^- = \{x \in H; (x, y) \leq 0 \ \forall y \in P\} = -P^*$$

Proposition:

— P^* est un cône convexe fermé non vide.
— $K_1 \subseteq K_2 \implies K_2^* \subseteq K_1^*$
— Si K est d'intérieur non vide, alors K^* est pointé
— Si la fermeture de K est pointé, alors K^* est d'intérieur non vide

Propriété: Caractérisation duale du minimum

Soit K un cône propre : on prend la relation d'ordre partielle $\preccurlyeq_K$. x est le minimum de S selon la relation d'ordre partielle $\preccurlyeq_K$ si et seulement si pour tout $\lambda \succ_{K^*} 0$, x est l'unique minimiseur de $\langle \lambda, z \rangle$ pour $z \in S$

Propriété: Caractérisation dual de l'élément minimal

— Si $\lambda \succ_{K^*} 0$ et x minimise $\langle \lambda, z\rangle$ pour $z \in S$, alors x est minimal.
— Si S est convexe, alors $x \in S$ est minimal s'il existe $\lambda \succcurlyeq_{K^*} 0$ non nul tel que x minimise $\langle \lambda, z\rangle$ pour $z \in S$.

6 Hyperplan d'appui

Définition:

Un hyperplan d'affine d'équation $(s, x) = r$ est appelé hyperplan d'appui à C en $\bar{x}$ si :

$$(s, x) \leq r \ \forall x \in C$$

$$(s, \bar{x}) = r$$

Théorème:

Soit C un ensemble convexe d'un Hilbert H. On suppose soit que H est de dimension finie soit que $\mathring{C} \neq \emptyset$. Soit $\bar{x} \in \partial C$. Alors il existe un hyperplan d'appui à C en $\bar{x}$.

7 Lemme de Farkas

Lemme:

Soient H un espace de Hilbert, $(\xi_j)_{j\in J} \subset H$ et $(\alpha_j)_{j\in J} \subset \mathbb{R}$. On suppoer que le système

$$(\xi_j, x) \leq \alpha_j \ \forall j \in J$$

admet au moins une solution.
Soit $(s, \beta) \in H \times \mathbb{R}$. On a équivalence entre les 2 propositions :

1. $$\forall x \in H, [\forall j \in J, \ (\xi_j, x) \leq \alpha_j \Rightarrow (s, x) \leq \beta]$$

2. $$(s, \beta) \in \overline{cone}\left((\xi_j, \alpha_j)_{j\in J} \cup (0, 1)\right) \subset H \times \mathbb{R}$$

Corollaire:

Sous les mêmes hypothèses avec $\alpha_j = 0 \ \forall j \in J$. On a pour $s \in H$:

1. $$\forall x \in H, [\forall j \in J, \ (\xi_j, x) \leq 0 \Rightarrow (s, x) \leq 0]$$

2. $$s \in \overline{cone}\left((\xi_j)_{j\in J}\right)$$

Lemme:

Si C est un cône convexe fermé, alors $C^{**} = C$.

Deuxième partie

Fonctions convexes

1 Définitions et propriétés

1.1 Fonctions convexes

$f : H \to \overline{\mathbb{R}} = \mathbb{R} \bigcup \{+\infty\}$

Définition:

$$Dom(f) = \{x \in H; f(x) < +\infty\}$$
$$epi(f) = \{(\alpha, x) \in \mathbb{R} \times H, \alpha \geq f(x)\}$$
$$epi_S(f) = \{(\alpha, x) \in \mathbb{R} \times H, \alpha > f(x)\}$$

On dit que f est propre si f n'est pas identiquement égal à $+\infty$.
On dit que f est convexe si $epi(f)$ est convexe.
On dit que f est concave si $-f$ est convexe.

Proposition:

Si f est convexe, alors $Dom(f)$ est convexe.
De plus, f est convexe si et seulement si : $\forall x, y \in Dom(f)$, $\forall \lambda \in [0, 1]$,

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

Définition:

On dit que f est strictement convexe si $\forall x, y \in Dom(f)$, $x \neq y$, $\forall \lambda \in]0, 1[$

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

Définition:

On dit que f est fortement convexe de module α si $\forall x, y \in Dom(f)$, $\forall \lambda \in]0,1[$

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

Propriété: opérations conservant la convexité

1. Pour $(f_i)_{i\in I}$ une famille quelconque de fonctions convexes, $\sup_{i\in I} f_i$ est convexe.
2. $\alpha \geq 0$, si f convexe, alors αf est convexe
3. Si f_1 et f_2 convexes, alors $f_1 + f_2$ convexes.

Définition:

Soit $f : H \to \overline{\mathbb{R}}$ et $\alpha \in \overline{\mathbb{R}}$.
On appelle sous ensemble de niveau de f au niveau α noté $\Gamma_\alpha(f)$ l'ensemble

$$\Gamma_\alpha(f) = \{x \in H; f(x) < \alpha\}$$

Remarque : f convexe $\Rightarrow \Gamma_\alpha(f)$ convexe $\forall \alpha \in \overline{\mathbb{R}}$
Si $\Gamma_\alpha(f)$ est convexe $\forall \alpha \in \mathbb{R}$, alors on dit que f est quasi-convexe.

Définition:

Soit $P \subset H$. On appelle fonction indicatrice de P la fonction :

$$\mathbb{1}_P(x) = \begin{cases} 0 & \text{si} \quad x \in P \\ +\infty & \text{sinon} \end{cases}$$

Remarque : Si P convexe, alors $\mathbb{1}_P$ est convexe.
Si $\alpha > 0$, $\alpha \in \mathbb{R}$, alors $\Gamma_\alpha(\mathbb{1}_P) = P$ donc $\mathbb{1}_P$ caractérise P.

1.2 Fonctions quasi-convexes

Définition: Quasi-convexe

$f : \mathbb{R}^n \to \mathbb{R}$ est dite quasi-convexe (ou unimodale) si son domaine et tous ses ensembles de niveau sont convexes.
f est quasi-concave si $-f$ est quasi-convexe. On fonction à la fois quasi-convexe et quasi-concave est dite quasi-linéaire.

Propriété:

- $f : \mathbb{R}^n \to \mathbb{R}$ est quasi-convexe si et seulement si dom f est convexe et pour tout $x, y \in$ dom f, et $0 \leq \lambda \leq 1$:
$$f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}$$
- Supposons f Fréchet-différentiable. f est quasi-convexe si et seulement si dom f est convexe et pour tout $x, y \in \text{dom} f$,
$$f(y) \leq f(x) \implies \langle \nabla f(x), y - x \rangle \leq 0$$
Si f est deux fois différentiable et quasi-convexe, alors pour tout $x \in \text{dom} f$, tout $y \in \mathbb{R}^n$,
$$\langle \nabla f(x), y \rangle = 0 \implies \langle y, \nabla^2 f(x) y \rangle \geq 0$$
On a équivalence si l'inégalité est stricte.
- $f : \mathbb{R} \to \mathbb{R}$, continue, est quasiconvexe si et seulement si une de ces conditions est vérifiée :
  - f est croissante
  - f est décroissante
  - il existe $c \in$ dom f tel que pour tout $t \leq c$ (et $t \in$ dom f), f est décroissante, et pour tout $t \geq c$ (et $t \in$ dom f), f est croissante.

1.3 Fonctions log-convexes

Définition: Logarithmiquement convexe

$f : \mathbb{R}^n \to \mathbb{R}_*^+$ est dite log-convexe si $\log f$ est convexe.
f est log-concave si $\log f$ est concave, ou de manière équivalente, $1/f$ est log-convexe.

Propriété:

- $f : \mathbb{R}^n \to \mathbb{R}_*^+$, avec dom$f$ convexe, est log-convave si et seulement si, pour tout $x, y \in \text{dom} f$, et $0 \leq \theta \leq 1$, on a
$$f(\theta x + (1-\theta)y) \geq f(x)^\theta f(y)^{1-\theta}$$
- f est log-concave si et seulement si pour tout $x \in \text{dom} f$,
$$f(x)\nabla^2 f(x) \preccurlyeq \nabla f(x) \nabla f(x)^\intercal$$

1.4 Convexité en regard des inégalités généralisées

1.4.1 Monotonicité

Soit $K \subseteq \mathbb{R}^n$ un cône propre, associé à l'inégalité généralisée $\preccurlyeq_K$.

Définition: K-croissant

Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dite K-croissante si
$$x \preccurlyeq_K y \implies f(x) \leq f(y)$$

et K-strictement croissante si

$$x \preccurlyeq_K y,\ x \neq y \implies f(x) < f(y)$$

On définit la décroissance de manière similaire.

Propriété:

— f, de domaine convexe, et K-croissante si et seulement si

$$\forall x \in \mathrm{dom} f,\ \nabla f(x) \succcurlyeq_{K^*} 0$$

Si l'inégalité est stricte, alors f est K-strictement croissante (mais la réciproque est fausse !)

1.4.2 Convexité

Soit $K \subseteq \mathbb{R}^m$ un cône propre, associé à l'inégalité généralisée $\preccurlyeq_K$.

Définition: K-croissant

Une fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ est dite K-convexe si pour tout x, y et $0 \leq \theta \leq 1$,

$$f(\theta x + (1-\theta)y) \preccurlyeq_K \theta f(x) + (1-\theta) f(y)$$

f est strictement K-convexe si l'inégalité est stricte pour tout tout $x \neq y$ et $0 < \theta < 1$.

Propriété:

— f est K-convexe si et seulement si pour tout $w \succcurlyeq_{K^*} 0$, la fonction scalaire $\langle w, f\rangle$ est convexe. f est strictement K-convexe si et seulement si pour tout $w \succcurlyeq_{K^*} 0$ non nul, la fonction scalaire $\langle w, f\rangle$ est strictement convexe

— f différentiable est K-convexe si et seulement si son domaine est convexe et pour tout $x, y \in \mathrm{dom} f$,

$$f(y) \succcurlyeq_K f(x) + Df(x)(y-x)$$

(où $Df(x)$ est la jacobienne de f en x). f est stricement K-convexe si l'inégalité est stricte pour tout $x \neq y$.

2 Fonctions d'appui

Définition:

Soit $S \subset H$.
On appelle fonction d'appui à S et on note σ_S la fonction définie par :

$$\sigma_S(d) = \sup_{s \in S} (s, d)$$

Remarque : σ_S est toujours convexe (même si S ne l'est pas).

Théorème:

Soit S un sous-ensemble non vide de H. Alors $s \in \overline{conv}(S)$ si et seulement si

$$\forall d \in H, \ (s, d) \leq \sigma_S(d)$$

De plus, $\sigma_S = \sigma_{\overline{conv}(S)}$

Remarque : Soient S_1 et S_2 2 convexes fermés. $S_1 = S_2 \Leftrightarrow \sigma_{S_1} = \sigma_{S_2}$.

Propriété:

Soient S_1 et S_2 deux sous-ensembles de H non vides.

1. $\sigma_{S_1+S_2} = \sigma_{S_1} + \sigma_{S_2}$
2. $\sigma_{S_1 \cup S_2} = \max\{\sigma_{S_1}, \sigma_{S_2}\}$

3 Transformée de Fenchel

On va chercher les fonctions affines minorantes :

$$\langle p, x \rangle + \alpha \leq f(x)$$

$$-\alpha \geq \langle p, x \rangle - f(x)$$

On va prendre $-\alpha = \sup_{x \in H}\{\langle p, x \rangle - f(x)\} = f^*(p)$.

Définition: Transformée de Fenchel

Soit H un Hilbert et $f : H \to \overline{\mathbb{R}}$. On définit la transformée de Fenchel de f, notée $f^* : H \to \overline{\mathbb{R}}$ par :

$$f^*(p) = \sup_{x \in H}\{\langle p, x \rangle - f(x)\}$$

Remarquons que, peu importe si f est convexe ou non, f^* est convexe.

Propriété: Inégalité de Young

$$\forall p, x \in H, \ f^*(p) + f(x) \geq \langle p, x \rangle$$

Proposition: Semi-continue inférieurement

Soit $f : H \to \mathbb{R}$. On dit que f est semi-continue inférieurement (sci) si l'une des deux propositions équivalentes suivantes est vérifiée :

1. $\forall x \in H,\ \forall x_n \to x$, $\liminf_{n\to+\infty} f(x_n) \geq f(x)$
2. $epi(f)$ est fermé.

Proposition:

Soit $(f_i)_{i\in I}$ une famille de fonctions sci. Alors $\sup_{i\in I} f_i$ est sci.

Corollaire:

f^* est sci et convexe.

Définition: Biconjuguée

La biconjuguée de f, notée f^{**} est définie par :

$$f^{**}(x) = \sup_{p\in H}\{\langle p, x\rangle - f^*(p)\}$$

Propriété:

— $f^{**}(x) + f^*(p) \geq \langle p, x\rangle$
— $f(x) \geq f^{**}(x)$

Proposition:

Si f est convexe, sci et propre, alors f^* est convexe, sci et propre.

Théorème: Fenchel-Moreau

Soit $f : H \to \overline{\mathbb{R}}$ une fonction propre. alors f est convexe et sci si et seulement si $f = f^{**}$

Remarque:

On peut définir la transformée de Fenchel sur un espace normé E refléxif :

$$f^* : \begin{array}{ccc} E' & \to & \mathbb{R} \\ p & \mapsto & \sup_{x \in E}\{\langle p, x\rangle_{E'E} - f(x)\} \end{array}$$

Dans ce cas, les propositions précédentes et le théorème de Fenchel-Moreau restent vraies.

Corollaire:

Soit f propre. alors f est convexe et sci si et seulement si f est l'enveloppe supérieure de ses minorantes affines.

Propriété:

— Supposons $f : \mathbb{R}^n \to \mathbb{R}$ convexe, Fréchet-différentiable, avec dom $f = \mathbb{R}^n$. Soit $z \in \mathbb{R}^n$. Posons $y = \nabla f(z)$. Alors

$$f^*(y) = \langle z, \nabla f(z)\rangle - f(z)$$

— Supposons $A \in \mathbb{R}^{n\times n}$ inversible, $b \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ convexe et propre. Alors la conjuguée de $g(x) = f(Ax+b)$ est :

$$g^*(p) = f^*\left(\left(A^{-1}\right)^{\intercal} y\right) - \langle A^{-1}b, y\rangle$$

et dom $g^* = A^{\intercal}$ dom f^*.

4 Continuité des fonctions convexes

Proposition:

Soit $f : H \to \overline{\mathbb{R}}$ convexe et propre. On suppose qu'il existe une boule ouverte sur laquelle f est bornée. Alors f est continue sur l'intérieur de son domaine qui est non vide.

Remarque:

Si f est continue en un point, alors f est bornée sur une boule, et donc f est continue sur l'intérieur de son domaine.

Corollaire:

Si $f : H \to \overline{\mathbb{R}}$ est convexe et propre avec H de dimension finie, alors f restreinte à l'intérieur relatif de son domaine est continue.

Remarque : Si $f : H \to \mathbb{R}$, alors f continue sur H.

5 Différentiabilité des fonctions convexes

5.1 Dérivées directionnelles des fonctions convexes

Théorème:

Soient $f : H \to \overline{\mathbb{R}}$, $x \in Dom(f)$, $d \in H$.

1. $\varepsilon \in \mathbb{R}_+^* \mapsto \frac{f(x+\varepsilon d)-f(x)}{\varepsilon}$ est croissante
2. $f'(x,d)$ existe toujours et vaut éventuellement $\pm\infty$. De plus, $f'(x,d) = +\infty$ si et seulement si $x + \varepsilon d \notin Dom(F)$ pour tout ε petit, et :
$$f'(x,d) = \inf_{\varepsilon \in \mathbb{R}_+^*} \frac{f(x+\varepsilon d) - f(x)}{\varepsilon}$$
$$f'(x,d) \leq f(x+d) - f(x)$$
3. $f'(x,d) \geq -f'(x,-d)$

5.2 Reconnaître une fonction convexe à l'aide de ses dérivées

Théorème:

Soit $f : H \to \overline{\mathbb{R}}$ une fonction propre. On suppose que f est différentiable sur un ouvert Ω de $Dom(f) \subset H$. On a équivalence entre les propositions suivantes :

1. f est convexe (resp. strictement convexe) sur Ω
2. $\forall x, y \in \Omega$, $f(y) \geq f(x) + f'(x, y-x)$ (resp. $f(y) > f(x) + f'(x, y-x)$)
3. $\forall x, y \in \Omega$, $(f'(y) - f'(x))(y-x) \geq 0$ (resp. $(f'(y) - f'(x))(y-x) > 0$)

Théorème:

Soit $f : H \to \overline{\mathbb{R}}$ propre et 2 fois différentiable sur un ouvert $\Omega \subset Dom(f)$.
Alors f est convexe si et seulement si $D^2 f(d,d) \geq 0$ $\forall d \in H$.
De plus, si $D^2 f(d,d) > 0$, alors f est strictement convexe (réciproque fausse : penser à $f(x) = x^4$)

6 Sous-différentiabilité des fonctions convexes

6.1 Définitions et premières propriétés

Définition: Fonction affine

a est affine si $\forall x, y \in H$, $\forall t \in \mathbb{R}$,

$$a(tx + (1-t)y) = ta(x) + (1-t)a(y)$$

Pour toute fonction affine, il existe x^* (la pente) et α (l'ordonnée) telles que $a(x) = \langle x^*, x \rangle + \alpha$.

Définition: Minorante affine

On dit que a est une minorante affine de f si a est affine et si :

$$\forall x \in H,\ f(x) \geq a(x)$$

On dit qu'une minorante affine est exacte en x_0 si $f(x_0) = a(x_0)$. Dans ce cas,

$$a(x) = \langle x^*, x - x_0 \rangle + f(x)$$

Théorème: Existence d'une minorante affine

Soit $f : H \to \overline{\mathbb{R}}$ convexe et propre. Alors f admet une minorante affine.
De plus, celle-ci peut être choisie exacte en un point de $ri(Dom(f))$, ie : si $x \in ri(Dom(f))$,

$$\exists x^* \in H;\ f(y) \geq \langle x^*, y - x \rangle + f(x)$$

Définition: Sous-différentiable

On dit que f convexe et propre est sous-différentiable en x s'il existe $x^* \in H$ tel que :

$$\forall y \in H,\ f(y) \geq \langle x^*, y - x \rangle + f(x)$$

Les éléments x^* sont appelés les sous-gradients de f en x, et on note $\partial f(x)$ l'ensemble des sous-gradients de f en x.
Par convention, si $x \notin Dom(f)$, alors $\partial f(x) = \emptyset$

Proposition: Sur l'optimalité

Soit $f : H \to \overline{\mathbb{R}}$ convexe et propre. Alors f atteint un minimum en x si et seulement si $0 \in \partial f(x)$.

Proposition:

Sous les mêmes hypothèses :

$$\partial f(x) = \{x^* \in H;\ f'(x,d) \geq \langle x^*, d\rangle,\ \forall d \in H\}$$

Théorème:

Soit $f : H \to \overline{\mathbb{R}}$ convexe et propre, et soit $x \in Dom(f)$. Les assertions suivantes sont équivalentes :

1. $\partial f(x) \neq \emptyset$
2. $\exists y \in ri(Dom(f))\,;\ f'(x, y-x) > -\infty$
3. $f'(x, \bullet) \neq -\infty$

Corollaire:

Si f est convexe et propre et si f est continue en $x \in Dom(f)$, alors $\partial f(x) \neq \emptyset$

Proposition:

Soit f convexe et propre tel que f est continue en x. Alors

$$f'(x,d) = \sigma_{\partial f(x)}(d) = \sup_{p \in \partial f(x)} \langle d, p\rangle$$

6.2 Sous-différentiabilité et transformée de Fenchel

Proposition:

Soit $f : H \to \overline{\mathbb{R}}$ convexe et propre. Alors

$$\partial f(x) = \{p \in H;\ f(x) + f^*(p) = \langle p, x\rangle\}$$

On définit de la même manière :

$$\partial f^*(p) = \{x \in H;\ f^{**}(x) + f^*(p) = \langle p, x\rangle\}$$

Proposition:

Soit $f : H \to \overline{\mathbb{R}}$ convexe, propre et sci.

$$x \in \partial f^*(p) \Leftrightarrow p \in \partial f(x)$$

6.3 Liens avec la différentiabilité

Proposition:

Soit $f : H \to \overline{\mathbb{R}}$ convexe, sci et propre. On suppose que f est continue en x.

1. Si f est Gâteaux-différentiable en x, alors

$$\partial f(x) = \{\nabla f(x)\}$$

2. Réciproquement, si $\partial f(x)$ est réduit à un seul élément, alors f est Gâteaux-différentiable en x et $\partial f(x) = \{\nabla f(x)\}$

6.4 Quelques règles de calcul

Dans toute la suite, on supposera la dimension de H finie.

Définition: Homogène et sous linéaire

$f'(x, \bullet)$ est dite homogène de degré $n \in \mathbb{R}^*$ si :

$$\forall \lambda \in \mathbb{R},\ f'(x, \lambda d) = \lambda^n f'(x, d)$$

$f'(x, \bullet)$ est dite sous-linéaire si :

$$\forall d \in H,\ \exists L > 0;\ |f'(x, d)| \leq L\|d\|$$

Proposition:

Soient $f : H \to \mathbb{R}$ une fonction convexe et propre et $x \in H$. Alors $f'(x, \bullet)$ est convexe, homogène de degré 1 et sous-linéaire.

Corollaire:

Sous les mêmes hypothèses, $\partial f(x)$ est un convexe compact non vide.

ℹ Proposition:

Soient $f_1, f_2 : H \to \mathbb{R}$ deux fonctions convexes, et $t_1, t_2 > 0$. Alors

$$\partial(t_1 f_1 + t_2 f_2)(x) = t_1 \partial f_1(x) + t_2 \partial f_2(x)$$

ℹ Proposition:

Soient $A : \mathbb{R}^n \to \mathbb{R}^m$ une fonction affine ($Ax = A_0 x + b$, $A_0 \in \mathcal{M}_{m \times n}$, $b \in \mathbb{R}^m$) et $g : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe.

$$\partial(g \circ A)(x) = A_0^* \partial g(Ax)$$

Troisième partie
Conditions d'optimalité

1 Introduction aux problèmes d'optimisation

1.1 Terminologie

On considère un problème du type

$$\begin{aligned} &\min\ f_0(x) \\ &\text{t.q. } f_i(x) \leq 0, \quad i = 1, ..., m \\ &\qquad h_i(x) = 0, \quad i = 1, ..., p \end{aligned} \tag{1}$$

On appelle

— $\mathcal{U}_{ad}$ *l'ensemble admissible* des points vérifiant les contraintes d'inégalité et d'égalité

— *point admissible* un point de $\mathcal{U}_{ad}$

— *valeure optimale* p^* de (1) :

$$p^* = \inf\{f_0(x)|f_i(x) \leq 0, i = 1, ..., m, h_i(x) = 0, i = 1, ..., p\}$$

— *ensemble optimal* l'ensemble des *points optimaux* x^* réalisant l'infimum :

$$X_{opt} = \{x^*|f_i(x^*) \leq 0, i = 1, ..., m, h_i(x^*) = 0, i = 1, ..., p,\ f(x^*) = p^*\}$$

— *ϵ-suboptimal* un point admissible x tel que $f_0(x) \leq p^* + \epsilon$

— un point x *locallement optimal* s'il existe $R > 0$ tel que

$$f_0(x) = \inf\{f_0(z)|f_i(z) \leq 0, i = 1, ..., m, h_i(z) = 0, i = 1, ..., p,\ \|x - z\| < R\}$$

1.2 Quelques formes équivalentes

Problème sous forme d'épigraphe On peut remettre le problème (1) sous la forme :

$$\begin{aligned} &\min\ t \\ &\text{t.q. } f_0(x) - t \leq 0 \\ &\qquad f_i(x) \leq 0, \qquad i = 1, ..., m \\ &\qquad h_i(x) = 0, \qquad i = 1, ..., p \end{aligned}$$

Problèmes différentiables avec contraintes linéaires : première approche des multiplicateurs Prenons le problème suivant où f_0 est Fréchet-différentiable :

$$\begin{aligned} &\min\ f_0(x) \\ &\text{t.q. } Ax = b \end{aligned}$$

Les conditions d'optimalité pour un x admissibles sont :

$$\langle \nabla f_0(x), y - x\rangle \geq 0$$

pour tout y vérifiant $Ay = b$. Comme x est admissible, tout y admissible est de la forme $y = x + v$ avec $v \in \ker(A)$. On peut donc réécrire la condition d'optimalité comme :

$$\langle \nabla f_0(x), v\rangle \geq 0,\ \forall v \in \ker(A)$$

et comme v est quelconque dans un espace vectoriel (rempalcer v par $-v$) :

$$\nabla f_0(x) \perp \ker(A)$$

Mais comme $\ker(A)^\perp =\text{Im}(A^\intercal)$, on peut dire que $\nabla f_0(x) \in\text{Im}(A^\intercal)$, ou encore qu'il existe $\nu \in \mathbb{R}^p$ tel que :

$$\nabla f_0(x) + A^\intercal \nu = 0$$

2 Problèmes convexes

$$\min_{u \in \mathcal{U}_{ad}} J(u)$$

$\mathcal{U}_{ad} \subset \mathbb{R}^n$ est l'ensemble (non vide) admissible. On suppose $\mathcal{U}_{ad}$ fermé convexe, et J convexe.

Théorème:

Si J est coercive ou si $\mathcal{U}_{ad}$ est borné, alors il existe un point de minimum.

2.1 Une condition nécessaire générale d'optimalité

Définition: Cône tangent

On dit que $d \in \mathbb{R}^n$ est une tangente à X en $\bar{x}$ si $\exists x_k \to \bar{x}$ avec $(x_k \subset X$, $t_k \to 0$, $t_k > 0$ tel que :

$$\frac{x_k - \bar{x}}{t_k} \to d$$

L'ensemble de toutes les directions tangentes est appelé le cône tangent et est noté $T_{\bar{x}}X$.

Définition: équivalente

$d \in T_{\bar{x}}X$ si $\exists t_k > 0$, $t_k \to 0$ et $\exists d_k \in X$, $d_k \to d$ tel que $\bar{x} + t_k d_k \in X$.

Proposition:

$T_{\bar{x}}X$ est un cône fermé. Il est convexe si X est convexe.

Proposition:

Soients X un ensemble convexe et $\bar{x} \in X$. Alors

$$T_{\bar{x}}X = \overline{cone}(X - \bar{x}) = \overline{\mathbb{R}_+(X - \bar{x})}$$

Définition:

Soient $X \subset \mathbb{R}^n$, $\bar{x} \in X$.
On dit que p est une direction normale à X en $\bar{x}$ si

$$\langle p, d \rangle \leq 0 \ \forall d \in T_{\bar{x}} X$$

L'ensemble des normales est appelé le cône normal, noté $\mathcal{N}_{\bar{x}} X$.

Remarque:

$\mathcal{N}_{\bar{x}} X = (T_{\bar{x}} X)^{-} = -(T_{\bar{x}} X)^{*}$
$\mathcal{N}_{\bar{x}} X$ est donc un cône convexe.

Théorème:

Soit $\mathcal{U}_{ad}$ un ensemble convexe fermé non vide, $J : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe, et $\bar{u} \in \mathcal{U}_{ad}$. Les assertions suivantes sont équivalentes :

1. $\bar{u}$ minimise J sur $\mathcal{U}_{ad}$
2. $J'(\bar{u}, u - \bar{u}) \geq 0 \ \forall u \in \mathcal{U}_{ad}$
3. $J'(\bar{u}, d) \geq 0 \ \forall d \in T_{\bar{u}} \mathcal{U}_{ad}$
4. $0 \in \partial J(\bar{u}) + \mathcal{N}_{\bar{u}} \mathcal{U}_{ad}$.

2.2 Cas où les contraintes sont explicites

$$\mathcal{U}_{ad} = \left\{ \begin{array}{lll} u; & \langle a_i, u \rangle = b_i & i = 1, ..., m \\ & c_j(u) \leq 0 & j = 1, ..., p \end{array} \right\}$$

$a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $c_j : \mathbb{R}^n \to \mathbb{R}$ convexe.
$\mathcal{U}_{ad}$ est convexe.
On note

$$\begin{array}{llll} A : & \mathbb{R}^n & \to & \mathbb{R}^m \\ & u & \mapsto & (\langle a_i, u \rangle)_{i=1}^m \end{array}$$

$$b = (b_i)_{i=1}^m$$

$$\{\langle a_i, u \rangle = b_i, \ i = 1, ..., m\} = \{u; Au = b\}$$

On note (a, b, c) les contraintes de manière générique.

Lemme:

Pour $u \in \mathcal{U}_{ad}$, on définit l'ensemble $\Lambda(u)$ par :

$$\Lambda(u) = \{\lambda \in \mathbb{R}^p; \lambda_i \geq 0, \lambda_i c_i(u) = 0 \ \forall j = 1, ..., p\}$$

On définit le cône :

$$\mathscr{N}_{(a,b,c)}(u) = \{A^* \mu + \sum_{j=1}^{p} \lambda_j s_j, \ \mu \in \mathbb{R}^n, \ \lambda \in \Lambda(u), \ s_j \in \partial c_j(u)\}$$

avec $A^* \mu = \sum_{i=1}^m \mu_i a_i$
Alors

$$\mathscr{N}_{(a,b,c)}(u) \subset \mathscr{N}_{\mathcal{U}_{ad}}(u)$$

Théorème:

Soit $J : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe, et $\bar{u} \in \mathcal{U}_{ad}$.

1. $\bar{u}$ minimise J sur $\mathcal{U}_{ad}$
2. $0 \in \partial J(\bar{u}) + \mathcal{N}_{\bar{u}}\mathcal{U}_{ad}$.
3. $\exists \mu \in \mathbb{R}^n$, $\exists \lambda \in \Lambda(\bar{u})\,;\, 0 \in \partial J(\bar{u}) + \sum_{i=1}^{m} a_i\mu_i + \sum_{j=1}^{p} \lambda_j \partial c_j(\bar{u})$

Alors (3) $\Rightarrow$ (1) $\Leftrightarrow$ (2)

2.2.1 Qualification des contraintes

Les contraintes sont qualifiées si :

$$\mathscr{N}_{(a,b,c)}(u) = \mathscr{N}_{\mathcal{U}_{ad}}(u) \text{ et alors } (3) \Leftrightarrow (1)$$

Lemme:

Si les contraintes c_j sont affines et si $\mathcal{U}_{ad}$ est non vide, alors

$$\mathscr{N}_{(a,b,c)}(u) = \mathscr{N}_{\mathcal{U}_{ad}}(u)$$

Lemme:

Si $s_1, ..., s_m \in \mathbb{R}^n$, alors $cone(s_1, ..., s_m)$ est fermé.

On oublie le cas affine :

Théorème:

On fait l'hypothèse dite de Slatter :

$$\exists u_0; \begin{cases} Au_0 = b \\ c_j(u_0) < 0 \quad \forall j = 1, ..., p \end{cases}$$

$\bar{u}$ minimise J sur $\mathcal{U}_{ad}$ si et seulement si :

$$\exists \mu \in \mathbb{R}^n, \exists \lambda \in \Lambda(\bar{u}); 0 \in \partial J(\bar{u}) + \sum_{i=1}^{m} a_i\mu_i + \sum_{j=1}^{p} \lambda_j \partial c_j(\bar{u})$$