Convexité et Optimisation

Présentation construite à partir du polycopié

Convexité et Optimisation de Guy Cohen

J.-Ph. Chancelier

Cermics, École des Ponts ParisTech.

Février 2019

Sous-espaces affines

Definition

Soit $\mathbb{X}$ un espace vectoriel réel. Un sous-ensemble A de $\mathbb{X}$ est un sous-espace affine si

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

Autrement dit, un sous-espace affine contient toujours la "droite" passant par deux de ses points x et y.

Enveloppe affine

- L'intersection de sous-espaces affines est un sous-espace affine
- L'ensemble des sous-espaces affines contenant A n'est pas vide (puisque $\mathbb{X}$ en est un)

Definition

L'enveloppe affine d'un sous-ensemble A, notée $\mathrm{aff}\,(A)$, est le plus petit sous-espace affine contenant A.

On l'obtient en considérant les combinaisons affines des points de A :

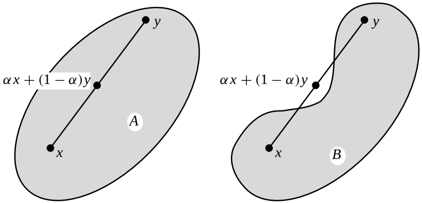
$$\mathrm{aff}\,(A) = \left\{ x \;\middle|\; n \in \mathbb{N} \setminus \{0\}, x = \sum_{i=1}^{n} \alpha_i x_i, x_i \in A, \sum_{i=1}^{n} \alpha_i = 1 \right\} .$$

Premières définitions sur les sous-ensembles convexes

Definition

Soit $\mathbb{X}$ un espace vectoriel réel. Un sous-ensemble A de $\mathbb{X}$ est convexe si

$$\forall x \in A, y \in A, \forall \alpha \in [0,1], \quad \alpha x + (1-\alpha) y \in A \,.$$



- A est convexe,
- B n'est pas convexe

- Un convexe contient toujours le segment $[x, y]$ joignant deux de ses points x et y.
- Un sous-espace affine est évidemment convexe.

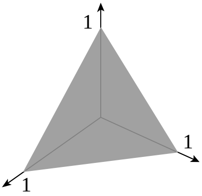
Simplexe de $\mathbb{R}^n$

Definition

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

$$\Sigma_n \stackrel{\text{def}}{=} \left\{ \alpha \in \mathbb{R}^n \;\middle|\; \alpha_i \geq 0, \;\; i = 1, \ldots, n, \;\; \sum_{i=1}^{n} \alpha_i = 1 \right\} . \qquad (2)$$

Notons que nécessairement $\alpha_i \leq 1$ pour $i = 1, \ldots, n$.



Le simplexe de $\mathbb{R}^3$

Combinaison convexe

Definition

On appelle combinaison convexe de n points $\{x_i\}_{i=1,\ldots,n}$ tout point y obtenu par la formule

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

- Une combinaison convexe de n points peut être calculée récursivement par $n-1$ combinaisons convexes de 2 points.
- Dans la Définition de la convexité on peut remplacer la combinaison convexe de 2 points par celle de n points en acceptant des valeurs de n quelconques.

Lemma (Caratheodory)

Dans un espace vectoriel de dimension n, toute combinaison convexe de m points, $m > n+1$, se ramène à une combinaison convexe de $n+1$ points au plus.

Démonstration.

Soit

$$y = \sum_{i=1}^{m} \alpha_i x_i, \;\; \alpha \in \Sigma_m \; .$$

On montre que si $m > n+1$, on peut faire décroître m de 1. Si l'un des α_i est nul, c'est immédiat. $\square$

Caratheodory : metre en forme

Démonstration.

Les vecteurs $z_i \stackrel{\text{def}}{=} x_i - x_1$ pour $i = 2, \ldots, m$, sont liés.
Il existe des β_i non tous nuls tels que

$$\sum_{i=2}^{m} \beta_i z_i = 0 \quad \Rightarrow \quad \sum_{i=1}^{m} \beta_i x_i = 0 \text{ avec } \sum_{i=1}^{m} \beta_i = 0 \ ,$$

en ayant posé $\beta_1 \stackrel{\text{def}}{=} -\sum_{i=2}^{m} \beta_i$. On considère alors

$$\gamma \stackrel{\text{def}}{=} \max_{i=1,\ldots,m} \left(\frac{\beta_i}{\alpha_i} \right) \text{ et } \delta_i = \alpha_i - \frac{\beta_i}{\gamma}, i \ = 1, \ldots, m \ .$$

Noter que $\gamma \neq 0$ car les β_i ne sont pas tous nuls. On vérifie maintenant que $y = \sum_{i=1}^{m} \alpha_i x_i = \sum_{i=1}^{m} \delta_i x_i$, et que au moins un δ_i est égal à 0. □

Examples

- L'ensemble des combinaisons convexes de n points isolés est un sous-ensemble convexe (appelé polytope ou polyèdre).
- L'intersection de sous-ensembles convexes est convexe.
- Pour $a \in \mathbb{X}$, le translaté $a + C \stackrel{\mathrm{def}}{=} \{a + x \mid x \in C\}$ est convexe.
- Pour $\alpha \in \mathbb{R}$, l'homothétique $\alpha C \stackrel{\mathrm{def}}{=} \{\alpha x \mid x \in C\}$ est convexe.
- L'image de C par une application affine de $\mathbb{X}$ dans un autre espace vectoriel $\mathbb{Y}$ est convexe.
- L'image réciproque de C par une application affine f de $\mathbb{Y}$ dans $\mathbb{X}$, $f^{-1}(C) \stackrel{\mathrm{def}}{=} \{y \in \mathbb{Y} \mid f(y) \in C\}$, est convexe.

Examples

- La somme vectorielle de C et d'un autre sous-ensemble convexe $C' \subset \mathbb{X}$, c'est-à-dire $C + C' \stackrel{\mathrm{def}}{=} \{x + x' \mid x \in C, x' \in C'\}$, est convexe.
- Le produit cartésien de $C \subset \mathbb{X}$ et $C' \subset \mathbb{Y}$, c'est-à-dire $C \times C' \stackrel{\mathrm{def}}{=} \{(x, y) \mid x \in C, y \in C'\}$, est un sous-ensemble convexe de $\mathbb{X} \times \mathbb{Y}$.
- Inversement, la projection d'un sous-ensemble convexe d'un espace produit sur l'un de ses sous-espaces composants est convexe.
- L'union de sous-ensembles convexes n'est pas convexe en général, mais l'union croissante de convexes (famille emboîtée) est convexe.

Une conséquence de ce qui précède est que les solutions d'un ensemble d'égalités et d'inégalités affines constitue un sous-ensemble convexe.

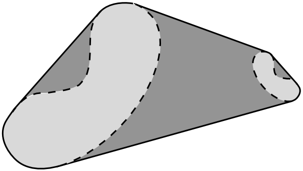
Enveloppe convexe $\mathrm{co}(A)$

Soit $A \subset \mathbb{X}$.

- l'espace $\mathbb{X}$ est un convexe contenant A.
- l'intersection de convexes contenant A étant convexes et contenant A

Definition

L'enveloppe convexe d'un sous-ensemble $A \subset \mathbb{X}$ quelconque est le plus petit convexe (au sens de l'inclusion) qui contient A. Elle est notée $\mathrm{co}(A)$



Enveloppe convexe (en sombre) d'un sous-ensemble (en clair)

Construction "interne" de $\mathrm{co}\,(A)$

Theorem

L'enveloppe convexe de A est égale à l'ensemble des combinaisons convexes d'éléments de A.

Démonstration.

- Soit B l'ensemble des combinaisons convexes d'éléments de A.
- On a $B \subset \mathrm{co}\,(A)$ puisque $\mathrm{co}\,(A)$, en tant que sous-ensemble convexe, doit contenir les combinaisons convexes de ses points, et en particulier ceux de A.
- Si on montre que B est convexe, alors comme évidemment $B \supset A$, on aura que $B \supset \mathrm{co}\,(A)$ d'après la définition de ce dernier sous-ensemble.
- Si on prend deux combinaisons convexes d'éléments de A, on doit montrer qu'une combinaison convexe de ces deux combinaisons convexes est encore une combinaison convexe d'éléments de A. Tout tient dans la remarque que, pour n nombres α_i, m nombres β_j et un nombre γ, tous dans $[0, 1]$, on a

$$\sum_{i=1}^{n} \alpha_i = 1 \text{ et } \sum_{j=1}^{m} \beta_j = 1 \quad \Rightarrow \quad \sum_{i=1}^{n} \gamma\alpha_i + \sum_{j=1}^{m} (1-\gamma)\beta_j = 1 \ .$$

$\square$

Théorème de Caratheodory

Theorem (Caratheodory)

Dans un espace vectoriel de dimension n, l'enveloppe convexe d'un sous-ensemble A est égale à l'ensemble des combinaisons convexes de $n+1$ points de A.

Cônes

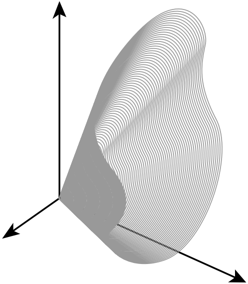
Rôle essentiel dans la formulation des contraintes inégalités.

Definition

Un sous-ensemble C est un cône si

$$\forall x \in C, \ \forall \alpha \geq 0, \ \alpha x \in C \ . \tag{3}$$

Un cône est donc une union de demi-droites fermées issues de l'origine.



Un cône non convexe de $\mathbb{R}^3$.

Cônes

- Les espaces vectoriels sont évidemment des cônes
- Un cône est dit saillant si $C \cap (-C) = \{0\}$
- Un espace vectoriel n'est pas un cône saillant.
- Un espace vectoriel est un cône convexe.

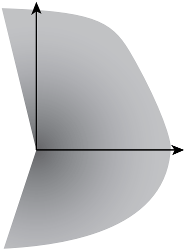


FIGURE – Un cône non saillant dans $\mathbb{R}^2$

Theorem

Un cône C est convexe si et seulement si il est stable par addition :

$$\forall x \in C, \forall y \in C, \ \ x + y \in C \ . \tag{4}$$

- préordre : relation réflexive et transitive
- relation d'ordre : relation de préordre qui est de plus antisymétrique
- La relation $x \succeq y \Leftrightarrow x - y \in C$ est une relation de préordre, et que c'est une relation d'ordre si de plus C est saillant.

Ordre de $\mathbb{R}^n$

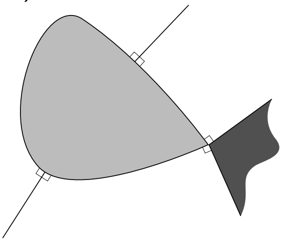
- L'utilisation la plus classique dans $\mathbb{R}^n$:
- le cône constitué par le premier "orthant". L'ensemble des $x \in \mathbb{R}^n$ tels que $x_i \geq 0, i = 1, \ldots, n$.
- La "non négativité" de x revient alors à celle de toutes ses composantes.

Definition

Dans un espace de Hilbert $\mathbb{X}$, le cône normal ou orthogonal à un sous-ensemble convexe A au point $a \in A$ est défini par

$$A_a^{\perp} = \{z \in \mathbb{X} \mid \langle z\,, x - a \rangle \leq 0, \forall x \in A\} \ .$$

C'est un cône convexe fermé comme intersection d'une collection de demi-espaces fermés. On notera $A^{\perp}$ pour $A_0^{\perp}$, (lorsque bien sûr $0 \in A$).



Cônes normaux en divers points de A.

Cône normal (ou orthogonal) à un sous-ensemble convexe

- Si $a \in \mathring{A}$ (intérieur de A, supposé donc non vide) alors $A_a^\perp = \{0\}$.
- Si A est sous-espace vectoriel ou affine, $A_a^\perp$ ne dépend pas de a. C'est le sous-espace vectoriel des z tels que $\langle z\,, x - a\rangle = 0, \forall x \in A, \forall a \in A$.
- Si C un cône convexe $C_a^\perp = C^\perp \cap \{a\}^\perp$, où $\{a\}$ désigne le sous-espace vectoriel de dimension 1 engendré par le vecteur a. D'où $\{a\}^\perp \cap C^\perp = \{0\}$ si $a \in \mathring{C}$.

Cône polaire

Dans la littérature, on trouve la notion de cône polaire d'un sous-ensemble A, généralement notée A°, et définie comme

Definition

$$A^\circ \stackrel{\mathrm{def}}{=} \{z \in \mathbb{X} \mid \langle z\,, x\rangle \leq 0, \forall x \in A\} \ .$$

Et qui vérifie :

$$A^\circ = \big(\angle(A)\big)^\circ = \big(\angle(A)\big)^\perp \ .$$

De même, on trouve les définitions suivantes

Definition

$$A^+ \stackrel{\mathrm{def}}{=} \{z \in \mathbb{X} \mid \langle z\,, x\rangle \geq 0, \forall x \in A\} \ .$$

$$A^- \stackrel{\mathrm{def}}{=} \{z \in \mathbb{X} \mid \langle z\,, x\rangle \leq 0, \forall x \in A\} \ .$$

Ordre dual dans les espaces de Hilbert

- Soit C un cône positif (convexe, fermé, saillant) dans $\mathbb{X}$, on pose

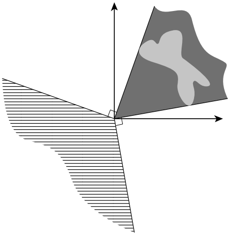
$$C^\star \stackrel{\text{def}}{=} \{\widehat{x} \in \mathbb{X}^* \mid \langle \widehat{x} \, , x \rangle \geq 0, \forall x \in C \subset \mathbb{X}\} \ .$$

- Modulo l'identification de $\mathbb{X}$ à $\mathbb{X}^*$ on a $C^\star = -C^\perp$
- $C^\star$ est un cône convexe fermé de $\mathbb{X}^*$.
- Même si on identifie $\mathbb{X}$ à son dual, en général $C \neq C^\star$: "être positif" ne veut pas dire la même chose dans le primal et dans le dual.
- Préfèrer la notation explicite $x \in C$ ou $x \in C^\star$ à la notation ambiguë $x \succeq 0$.

- $\mathbb{X}$ est un cône convexe contenant A ;
- L'intersection de cônes convexes contenant A contient aussi A et est un cône convexe.

Definition

L'enveloppe cônique d'un sous-ensemble $A \subset \mathbb{X}$ quelconque est le plus petit cône convexe (au sens de l'inclusion) qui contient A. Elle est notée $\angle(A)$.



Enveloppe cônique (en sombre) d'un sous-ensemble (en clair) et cône polaire associé (rayé)

Enveloppe cônique

Theorem

- *L'enveloppe cônique de A est égale à l'enveloppe cônique de* $\mathrm{co}\,(A)$ *;*
- *elle est également obtenue en effectuant toutes les combinaisons linéaires à coefficients non négatifs d'éléments de A.*

Theorem

Dans un espace vectoriel de dimension n, l'enveloppe cônique d'un sous-ensemble A est égale à l'ensemble des combinaisons linéaires à coefficients non négatifs de $n+1$ points de A.

Cône tangent à un sous-ensemble

Definition

Soit A un sous-ensemble d'un espace vectoriel normé $\mathbb{X}$. On dit que x est tangent à A au point $a \in A$ si il existe une suite $\{\alpha^k\}_{k\in\mathbb{N}}$ de réels positifs et une suite $\{a^k\}_{k\in\mathbb{N}}$ d'éléments de A telles que

$$\lim_{k\to+\infty} a^k = a \quad \text{et} \quad \lim_{k\to+\infty} \alpha^k(a^k - a) = x \,.$$

Definition

L'ensemble des x tangents à A en a est appelé cône tangent à A en a (il est noté $A_a^\top$).

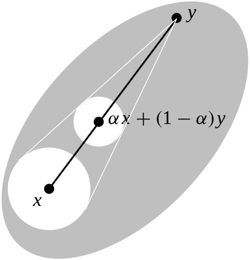
- On note $A^\top$ pour $A_0^\top$
- C' est une notion locale et si $a \in \overset{\circ}{A} \subset \mathbb{X}$, alors $A_a^\top = \mathbb{X}$.
- Si A est convexe, il existe une définition globale.

Topologie : Ouverture et fermeture des convexes

Theorem

Si C est convexe, alors l'intérieur $\overset{\circ}{C}$ de C et son adhérence $\overline{C}$ sont aussi convexes.

- $\overline{C}$: Si x et y appartiennent à $\overline{C}$, il existe des suites $\{x_n\}$ et $\{y_n\}$ d'éléments de C convergeant respectivement vers x et y, et pour tout $\alpha \in [0,1]$, $\alpha x_n + (1-\alpha)y_n$ appartient à C et converge vers $\alpha x + (1-\alpha)y$ qui appartient donc à $\overline{C}$, donc $\overline{C}$ est convexe.
- $\overset{\circ}{C}$:



Une boule centrée sur x et son homothétique de centre y

Intérieur relatif

Definition

Pour un sous-ensemble convexe C non vide d'un espace vectoriel normé de dimension finie, on appelle intérieur relatif de C, noté $\mathrm{ri}\,(C)$, l'intérieur de C pour la topologie induite sur $\mathrm{aff}\,(C)$. On appelle dimension de C la dimension du sous-espace vectoriel parallèle à $\mathrm{aff}\,(C)$ (c'est-à-dire $\mathrm{aff}\,(C) - a$, pour tout $a \in \mathrm{aff}\,(C)$).

Theorem

Pour tout sous-ensemble convexe C non vide d'un espace vectoriel normé de dimension finie, $\mathrm{ri}\,(C)$ est non vide.

Intérieur relatif

Soit n la dimension de $\mathrm{aff}\,(C)$ et soit $x_0 \in C$. Il existe n vecteurs x_i de C tels que $\{x_i - x_0\}_{i=1,\dots,n}$ constituent une base de $\mathrm{aff}\,(C) - x_0$:

$$y = \sum_{i=1}^{n} \alpha_i (x_i - x_0) + x_0 = \sum_{i=0}^{n} \alpha_i x_i \;, \forall y \in \mathrm{aff}\,(C)$$

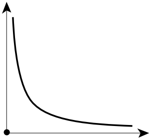
en ayant posé $\alpha_0 \overset{\text{def}}{=} 1 - \sum_{i=1}^{n} \alpha_i$. L'application f qui à $\alpha \in \mathbb{R}^n$ associe le $y \in \mathrm{aff}\,(C)$ défini par la formule ci-dessus est un isomorphisme continu entre $\mathbb{R}^n$ et $\mathrm{aff}\,(C)$.
Considérons alors l'ouvert suivant de $\mathbb{R}^n$:

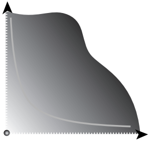
$$\Omega \overset{\text{def}}{=} \left\{ \alpha \in \mathbb{R}^n \;\middle|\; \alpha_i > 0, i = 1, \dots, n, \sum_{i=1}^{n} \alpha_i < 1 \right\} .$$

Alors $f(\Omega)$ est un ouvert de $\mathrm{aff}\,(C)$ et $f(\Omega) \subset C$. L'intérieur de C dans $\mathrm{aff}\,(C)$ est non vide.

Enveloppe convexe fermée

- L'enveloppe convexe d'un sous-ensemble ouvert est ouverte.
- L'enveloppe convexe d'un sous-ensemble fermé n'est pas nécessairement fermée.
- En dimension finie, si l'ensemble est fermé et borné l'enveloppe convexe est aussi compacte, donc fermée.





Un sous-ensemble fermé et son enveloppe convexe non fermée

enveloppe convexe fermée

Definition

L'enveloppe convexe fermée d'un sous-ensemble $A \subset \mathbb{X}$ quelconque est le plus petit convexe fermé (au sens de l'inclusion) qui contient A. Elle est notée $\overline{\mathrm{co}}\,(A)$.

Justifié par le fait que l'intersection de convexes fermés est un convexe fermé. Le théorème ci-dessous indique une autre façon de définir $\overline{\mathrm{co}}\,(A)$.

Theorem

L'enveloppe convexe fermée d'un sous-ensemble A est égale à la fermeture (ou l'adhérence) de son enveloppe convexe, c'est-à-dire $\overline{\mathrm{co}\,(A)}$.

Le cas de la dimension finie

Theorem

Dans un espace vectoriel normé $\mathbb{X}$ de dimension finie, si $A \subset \mathbb{X}$ est borné (resp. compact), alors $\mathrm{co}\,(A)$ est bornée (resp. compacte).

- Si n est la dimension de l'espace et si $x \in \mathrm{co}\,(A)$ alors

$$\|x\| \leq \sum_{i=1}^{n+1} \alpha_i \|x_i\| \quad \text{avec} \quad \{x_i\}_{i=1,\cdots,n+1} \in A$$

- Soit $\{x^k\}_{k\in\mathbb{N}}$ une suite d'éléments de $\mathrm{co}\,(A)$, On a :

$$x^k = \sum_{i=1}^{n+1} \alpha_i^k x_i^k \quad \text{avec} \quad \alpha^k \in \Sigma_{n+1}$$

Le simplexe Σ_{n+1} est compact. Par extraction successives on trouve un point d'accumulation de $\{x^k\}$ dans $\mathrm{co}\,(A)$.

Enveloppe cônique fermée

Definition

L'enveloppe cônique fermée d'un sous-ensemble $A \subset \mathbb{X}$ quelconque est le plus petit cône convexe fermé (au sens de l'inclusion) qui contient A. Elle est notée $\overline{\angle}(A)$.

$$\overline{\angle}(A) = \overline{\angle(A)} = \overline{\angle}\left(\overline{A}\right) = \overline{\angle\left(\overline{A}\right)} \supset \angle\left(\overline{A}\right) \ . \tag{5}$$

- Il est faux que l'enveloppe cônique d'un compact A est fermée.
- Considérer dans $\mathbb{R}^2$ la boule fermée de centre $(1, 0)$ et de rayon 1)
- Cela devient vrai avec l'hypothèse que $0 \notin \mathrm{co}\,(A)$.

Theorem

Soit A un sous-ensemble convexe d'un espace vectoriel normé $\mathbb{X}$, et soit $a \in A$. Alors $A_a^{\top}$ est un cône convexe fermé et c'est en fait l'enveloppe cônique fermée de $A - a$

- $A_a^{\top}$ est convexe car stable par addition :

$$
\begin{aligned}
x_1 + x_2 &= \lim_{k \to +\infty} \left(\alpha_1^k (a_1^k - a) + \alpha_2^k (a_2^k - a)\right) \\
&= \lim_{k \to +\infty} (\alpha_1^k + \alpha_2^k) \left(\frac{\alpha_1^k}{\alpha_1^k + \alpha_2^k} (a_1^k - a) + \frac{\alpha_2^k}{\alpha_1^k + \alpha_2^k} (a_2^k - a) \right) \\
&= \lim_{k \to +\infty} (\alpha_1^k + \alpha_2^k) \left(\frac{\alpha_1^k}{\alpha_1^k + \alpha_2^k} a_1^k + \frac{\alpha_2^k}{\alpha_1^k + \alpha_2^k} a_2^k - a \right) .
\end{aligned}
$$

- $A_a^\top$ est fermé.
  Pour cela, on considère une suite $\{x_n\}_{n\in\mathbb{N}} \subset A_a^\top$ convergeant vers x et on montre que $x \in A_a^\top$.
  À chaque élément x_n de $A_a^\top$ sont associées deux suites $\{\alpha_n^k\}_{k\in\mathbb{N}}$ et $\{a_n^k\}_{k\in\mathbb{N}}$: il existe alors un indice k_n tel que

$$\left\|\alpha_n^{k_n}\left(a_n^{k_n} - a\right) - x_n\right\| \leq 1/n \ .$$

  Les suites $\{\alpha_n^{k_n}\}_{n\in\mathbb{N}}$ et $\{a_n^{k_n}\}_{n\in\mathbb{N}}$ sont telles que $\alpha_n^{k_n}\left(a_n^{k_n} - a\right)$ converge vers x, ce qui montre le résultat recherché.

Orthogonalité du cône tangent et du cône normal

- Inclusion $\overline{\angle}(A-a) \subset A_a^\top$.
  Soit $x \in A$, Les suites $\{\alpha^k\}_{k\in\mathbb{N}}$, avec $\alpha^k = k$, et $\{a^k\}_{k\in\mathbb{N}}$, avec $a^k = (1-1/k)a + x/k$, sont telles que $\lim \alpha^k(a^k - a) = x - a$, et de plus $a^k \in A$ parce que A est convexe. On a donc montré que $x - a \in A_a^\top$. Donc $A - a \subset A_a^\top$. Mais $A_a^\top$ étant un cône convexe fermé qui contient $A - a$, il contient aussi $\overline{\angle}(A-a)$ qui est le plus petit cône convexe fermé ayant cette propriété.
- Inclusion inverse. Tout $x \in A_a^\top$ s'écrit comme la limite d'une suite $\alpha^k(a^k - a)$ avec $a^k \in A$. Chacun de ces termes appartient à $\overline{\angle}(A-a)$ par définition de cet ensemble. Celui-ci étant fermé, la limite appartient aussi à cet ensemble.

Le cas convexe

Corollary

Si A est convexe, et $a \in A$, le cône tangent à A en a est l'orthogonal du cône normal à A en a, c'est-à-dire que

$$\left(A_a^{\top}\right)^{\perp} = A_a^{\perp} . \tag{6}$$

Pour A convexe et $a \in A$, grâce au théorème précédent, nous bénéficions d'une description simplifiée de $A_a^{\top}$:

$$A_a^{\top} = \overline{\angle}(A - a) = \overline{\{y \mid y = \alpha(x - a), \alpha \in \mathbb{R}_+, x \in A\}} .$$

On en déduit immédiatement que

$$\left(A_a^{\top}\right)^{\perp} = \{z \mid \langle z \,, \alpha(x - a)\rangle \leq 0, \alpha \in \mathbb{R}_+, x \in A\}$$

(qui est automatiquement fermé). Il est évident que la présence du $\alpha \geq 0$ n'apporte rien de plus à la définition de l'ensemble ci-dessus, et on reconnaît alors précisément la définition de $A_a^{\perp}$.

Theorem

Soit x un élément d'un espace de Hilbert $\mathbb{X}$ et A un sous-ensemble convexe fermé de $\mathbb{X}$. Il existe un unique point $y \in A$ tel que

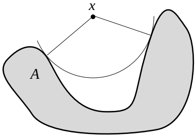
$$\|y - x\| = \min_{z \in A} \|z - x\| \ . \tag{7}$$

Cet élément y est caractérisé par l'inéquation variationnelle

$$\forall z \in A, \ \langle x - y \, , z - y \rangle \leq 0 \ . \tag{8}$$

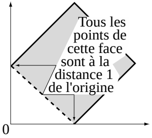
L'élément y est appelé projection de x sur A et sera noté $\Pi_A(x)$.

- A fermé est nécessaire pour l'existence de la projection de x.
- Il suffit de chercher le minimum sur $A \cap B_f(x, \beta')$ avec $\beta' > \beta = \inf_{z \in A} \|z - x\|$. En dimension finie, le minimum d'une fonction continue sur un compact est atteint.
- L'importance de la convexité de A pour l'unicité :



Le point x a deux projections sur A.

- La distance est liée à une norme issue d'un produit scalaire joue un rôle également important pour l'unicité de la projection. Considérons les boules pour la norme ℓ^1.



Non unicité de la projection avec la norme ℓ^1

Propriétés de la projection

Theorem

La projection Π_A est un opérateur Lipschitzien de constante 1 et monotone, c'est-à-dire que

$$\left\| \Pi_A(x') - \Pi_A(x) \right\| \leq \left\| x' - x \right\| \;, \tag{9}$$

$$\left\langle \Pi_A(x') - \Pi_A(x) \,, x' - x \right\rangle \geq 0 \;. \tag{10}$$

Utilisons la caractérisation (8) de la projection pour x, $y = \Pi_A(x)$ et $z = \Pi_A(x')$:

$$\left\langle \Pi_A(x') - \Pi_A(x) \,, x - \Pi_A(x) \right\rangle \leq 0 \;,$$

puis additionnons cette inéquation avec celle analogue qui est obtenue en inversant le rôle de x et x'.

$$\left\langle \Pi_A(x') - \Pi_A(x) \,, x' - x \right\rangle \geq \left\| \Pi_A(x') - \Pi_A(x) \right\|^2 \;.$$

Projection sur un cône

Theorem

Soit C un cône convexe fermé d'un espace de Hilbert $\mathbb{X}$ et $x \in \mathbb{X}$. la projection $\Pi_C(x) \in C$ est caractérisée par

$$\langle \Pi_C(x) \,, x - \Pi_C(x) \rangle = 0 \,,$$
$$\forall z \in C, \;\; \langle z \,, x - \Pi_C(x) \rangle \leq 0 \,.$$

La projection Π_C sur un cône est un opérateur positivement homogène de degré 1, c'est-à-dire que

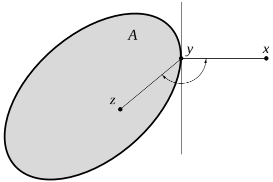
$$\forall \alpha \geq 0, \;\; \Pi_C(\alpha x) = \alpha \Pi_C(x) \,.$$

De plus, $x - \Pi_C(x)$ est égal à la projection $\Pi_{-C^\star}(x)$ et x se décompose de façon unique en

$$x = \Pi_C(x) + \Pi_{-C^\star}(x) \;\; \textit{avec} \;\; \langle \Pi_C(x) \,, \Pi_{-C^\star}(x) \rangle = 0 \,.$$

Séparation des convexes

Interprétation géométriques des conditions d'optimalité de la projection :



Séparation du point x et du convexe A

Theorem

Soit A un sous-ensemble convexe fermé d'un espace de Hilbert $\mathbb{X}$ et $x \notin A$. Alors il existe $r \in \mathbb{X}$ tel que

$$\sup_{z \in A} \langle r \,, z \rangle < \langle r \,, x \rangle \ . \tag{12}$$

Il suffit de prendre $r = x - \Pi_A(x)$ qui est non nul si $x \notin A$.

Séparation des convexes

Definition

Deux sous-ensembles convexes A_1 et A_2 d'un espace de Hilbert sont séparés strictement par la forme linéaire $x \mapsto \langle r \, , x \rangle$ si

$$\sup_{x_1 \in A_1} \langle r \, , x_1 \rangle < \inf_{x_2 \in A_2} \langle r \, , x_2 \rangle \ , \tag{13a}$$

ou bien

$$\sup_{x_2 \in A_2} \langle r \, , x_2 \rangle < \inf_{x_1 \in A_1} \langle r \, , x_1 \rangle \ . \tag{13b}$$

Géométriquement, les deux convexes appartiennent chacun à l'un des demi-espaces délimité par l'hyperplan $\langle r \, , x \rangle = \alpha$ pour tout α compris strictement entre les deux nombres définis dans les deux membres de l'inégalité (13a) ou (13b) (selon celle qui est valide).

Corollary

Dans un espace de Hilbert, on considère deux convexes fermés A_1 et A_2 et on suppose de plus que A_2 est compact. Ces deux convexes peuvent être séparés strictement si et seulement si ils sont disjoints.

Si A_1 et A_2 sont disjoints, alors considérons $A_1 - A_2 = \{x_1 - x_2 \mid x_1 \in A_1, x_2 \in A_2\}$ qui est convexe ; l'origine n'appartient pas à $A_1 - A_2$.

Autre lecture : Le vecteur $x - \Pi_A(x) \in A_a^\perp$ et ce vecteur n'est pas nul si $x \notin A$.
Réciproquement

Lemma

Soit A un convexe fermé d'un espace de Hilbert $\mathbb{X}$ de dimension finie et a appartenant à la frontière de A ($a \in A \setminus \overset{\circ}{A}$). Alors, $A_a^\perp$ n'est pas réduit à 0.

Si $aA \setminus \overset{\circ}{A}$, il existe une suite $\{x^k \notin A\}_{k \in \mathbb{N}}$ et tendant vers a. De plus par séparation :

$$\exists r^k \in \mathbb{X} : \sup_{x \in A} \left\langle r^k , x \right\rangle < \left\langle r^k , x^k \right\rangle . \tag{14}$$

On peut choisir les r^k de norme 1 et par compacité de la boule unité on trouve r un point d'accumulation de la suite $\{r^k\}_{k \in \mathbb{N}}$:

$$\forall x \in A, \ \langle r , x \rangle \leq \langle r , a \rangle ,$$

Hyperplans d'appui et description externe des convexes

Parmi les hyperplans qui séparent un point $x \notin A$ du convexe fermé A, on considère l'hyperplan

$$\{z \mid \langle x - \Pi_A(x) \,, z - \Pi_A(x) \rangle = 0\}$$

On appelle un tel hyperplan (ayant au moins un point commun avec A et incluant A dans l'un de ses demi-espaces fermés) un hyperplan d'appui de A (au point qu'il a en commun avec A).

Hyperplans d'appui et description externe des convexes

Theorem

Tout convexe fermé A est égal à l'intersection de tous les demi-espaces fermés qui le contiennent et qui sont délimités par tous les hyperplans d'appui en tous les points de A.

Corollary

Dans un espace de Hilbert, un convexe fermé pour la topologie forte est aussi fermé pour la topologie faible.

Si un sous-ensemble A quelconque est fermé pour la topologie faible, il est fermé pour la topologie forte. La réciproque est en général fausse. Le corollaire ci-dessus affirme que c'est vrai si A est de plus convexe.

On obtient donc une "description externe" des convexes fermés.

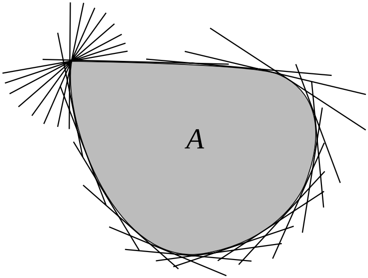


Figure – Description externe d'un convexe fermé

Lemme de Farkas

Une autre conséquence du théorème de séparation est le résultat suivant connu sous le nom de "Lemme de Farkas".

Lemma (Farkas)

Considérons une famille $\{r_j\}_{j \in J}$ *d'éléments d'un espace de Hilbert* $\mathbb{X}$ *et une famille* $\{\alpha_j\}_{j \in J}$ *de nombres réels associés. On suppose que le système d'inégalités*

$$\langle r_j \,, x \rangle \leq \alpha_j, \;\; \forall j \in J \,,$$

admet au moins une solution. Alors, étant donné $s \in \mathbb{X}$ *et* $\beta \in \mathbb{R}$, *les deux énoncés suivants sont équivalents :*

$$\forall x \in \mathbb{X}, \;\; [\langle r_j \,, x \rangle \leq \alpha_j, \forall j \in J] \Rightarrow [\langle s \,, x \rangle \leq \beta] \,, \tag{15a}$$

$$(s, \beta) \in \overline{\angle}(\{(r_j, \alpha_j)\}_{j \in J} \cup \{(0,1)\}) \subset \mathbb{X} \times \mathbb{R} \,, \tag{15b}$$

où $(0,1) \in \mathbb{X} \times \mathbb{R}$.

C est un cône convexe fermé

Lemma

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

$$C^{\star\star} = \{ y \mid \langle y \, , \widehat{x} \rangle \geq 0, \forall \widehat{x} : [\forall x \in C, \langle \widehat{x} \, , x \rangle \geq 0] \} \ .$$

Autrement dit,

$$y \in C^{\star\star} \iff \big[[\forall x \in C : \langle x \, , \widehat{x} \rangle \geq 0] \Rightarrow [\langle y \, , \widehat{x} \rangle \geq 0] \big] \ ,$$

ce qui est équivalent, d'après le lemme de Farkas à ce que $y \in \overline{\angle}(\{x\}_{x \in C}) = \overline{C} = C$ puisque C est fermé. Donc $C^{\star\star} = C$.

Épigraphe et convexité

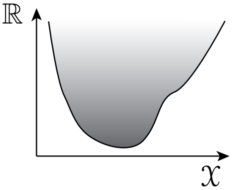
Definition

Soit $\mathbb{X}$ un espace de Hilbert et $f : \mathbb{X} \to \overline{\mathbb{R}}$. On appelle épigraphe de f, noté $\mathrm{epi}\,(f)$, le sous-ensemble

$$\mathrm{epi}\,(f) \stackrel{\mathrm{def}}{=} \{(x,\xi) \mid x \in \mathbb{X}, \xi \in \mathbb{R} : \xi \geq f(x)\} \ .$$

On aura parfois besoin de l'épigraphe strict, noté $\mathrm{epi}_{\mathrm{s}}\,(f)$:

$$\mathrm{epi}_{\mathrm{s}}\,(f) \stackrel{\mathrm{def}}{=} \{(x,\xi) \mid x \in \mathbb{X}, \xi \in \mathbb{R} : \xi > f(x)\} \ .$$



Épigraphe d'une fonction $f : \mathbb{X} \to \overline{\mathbb{R}}$.

Épigraphe et convexité

- Un sous-ensemble $A \subset \mathbb{X} \times \mathbb{R}$ est un épigraphe si et seulement
  - il est "infini vers le haut" : c'est-à-dire que si $(x, \xi) \in A$, alors $(x, \xi') \in A$ dès que $\xi' \geq \xi$
  - $\forall x \in \mathbb{X}, \inf_{(x,\xi)\in A}$ est atteint (quitte à valoir $-\infty$).
- On a alors, pour A verifiant ces deux propriétés :

$$\mathrm{epi}\,(f) = A \iff \forall x \in \mathbb{X}, f(x) = \inf_{(x,\xi)\in A} \xi \ ,$$

- Cela permet d'échanger un coût par une contrainte
- Les fonctions sont à valeurs dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ mais l'épigraphe est un sous-ensemble de $\mathbb{X} \times \mathbb{R}$ et non $\mathbb{X} \times \overline{\mathbb{R}}$.

Lemma

Pour une fonction $f : \mathbb{X} \to \overline{\mathbb{R}}$, $\mathrm{epi}\,(f)$ *est convexe si et seulement si* $\mathrm{epi}_{\mathrm{s}}\,(f)$ *est convexe.*

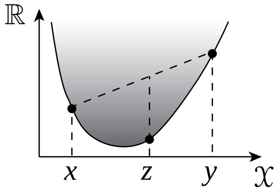
Definition

Soit $\mathbb{X}$ un espace de Hilbert et $f : \mathbb{X} \to \overline{\mathbb{R}}$. La fonction f est convexe si son épigraphe est convexe.

On trouve alors que cette définition équivaut à :

$$\forall x \in \mathbb{X}, y \in \mathbb{Y}, \forall \alpha \in [0, 1], f\big(\alpha x + (1-\alpha)y\big) \leq \alpha f(x) + (1-\alpha) f(y) \ .$$

La valeur de la fonction f en tout point z du segment $[x, y]$ est située sous la valeur de la “corde” (interpolation linéaire de f entre x et y) au même point z.



Interprétation de la convexité

- Une fonction f est concave si $-f$ est convexe.
- Une fonction f est concave ssi son “hypographe” (la partie de l'espace située sous le graphe) est convexe.
- Une fonction est affine si et seulement si elle est à la fois convexe et concave.

Domaine de f

Definition

On appelle domaine de $f : \mathbb{X} \to \overline{\mathbb{R}}$, noté $\mathrm{dom}\,(f)$, le sous-ensemble de $\mathbb{X}$ où f prend des valeurs finies :

$$\mathrm{dom}\,(f) \stackrel{\mathrm{def}}{=} \{x \in \mathbb{X} \mid f(x) < +\infty\} \ .$$

On observe que $\mathrm{dom}\,(f)$ est précisément la projection de $\mathrm{epi}\,(f)$ sur $\mathbb{X}$. Le domaine d'une fonction convexe est donc convexe.

Definition

Une fonction $f : \mathbb{X} \to \overline{\mathbb{R}}$ est strictement convexe si elle est convexe et si l'inégalité stricte a lieu dans l'inégalité de convexité dès que $x \neq y$ et $\alpha \in]0,1[$ (Le graphe de la fonction n'a pas de partie affine).

Definition

Une fonction $f : \mathbb{X} \to \overline{\mathbb{R}}$ est fortement convexe (de module b) si

$$\exists b > 0 : \forall x \in \mathbb{X}, y \in \mathbb{Y}, \forall \alpha \in [0,1],$$

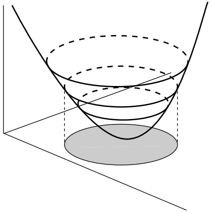
$$f\big(\alpha x + (1-\alpha)y\big) \leq \alpha f(x) + (1-\alpha) f(y) - b\alpha(1-\alpha) \left\| x - y \right\|^2 /2 \,.$$

Convexité forte implique convexité stricte. Réciproque fausse : $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^4$.

Definition

Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ et $\beta \in \overline{\mathbb{R}}$. On appelle ensemble de niveau de f au niveau β, noté $\triangledown_\beta(f)$, le sous-ensemble (éventuellement vide)

$$\triangledown_\beta(f) \stackrel{\text{def}}{=} \{x \in \mathbb{X} \mid f(x) \leq \beta\} \ .$$



Ensembles de niveau.

- Si f est convexe alors $\triangledown_\beta(f)$ est convexe pour tout β.
- Ne caractérise pas les fonctions convexes.
- Une fonction est dite quasi-convexe si $\forall \beta \, \triangledown_\beta(f)$ est convexe.

Fonction indicatrice

Definition

Soit A un sous-ensemble d'un espace de Hilbert $\mathbb{X}$. On appelle fonction indicatrice de A, noté I_A, la fonction définie sur $\mathbb{X}$ à valeurs dans $\overline{\mathbb{R}}$ suivante :

$$\mathrm{I}_A(x) = \begin{cases} 0 & \text{si } x \in A\,, \\ +\infty & \text{sinon.} \end{cases}$$

- I_A est convexe si et seulement si A est convexe.
- $A = \overline{\nabla}_\beta\left(\mathrm{I}_A\right)$ pour tout β non négatif et fini.
- A est aussi le domaine de I_A
- La fonction I_A caractérise sans ambiguïté le sous-ensemble A.

Fonction support

Definition

Soit A un sous-ensemble non vide d'un espace de Hilbert $\mathbb{X}$. On appelle fonction support de A, noté σ_A, la fonction définie sur $\mathbb{X}$ à valeurs dans $\overline{\mathbb{R}}$ suivante :

$$\sigma_A(p) = \sup_{x \in A} \langle p \, , x \rangle \ .$$

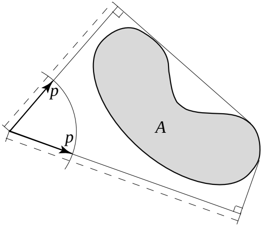
Avec la convention de prendre la valeur $-\infty$ (respectivement, $+\infty$) lorsque le sup (respectivement, l'inf) est pris sur un ensemble vide.

- Convexe même lorsque A ne l'est pas :
- C'est une enveloppe supérieure de fonctions linéaires.
- Elle est aussi positivement homogène de degré 1 :

$$\forall \alpha > 0, \forall p \in \mathbb{X}, \ \ \sigma_A(\alpha p) = \alpha \sigma_A(p) \ .$$

Fonction support

- La fonction support ne caractérise pas A.
- En effet, A, $\mathrm{co}\,(A)$ et $\overline{\mathrm{co}}\,(A))$ ont même fonction support.



Valeurs de $\sigma_A(p)$ (longueurs en pointillés) pour deux valeurs particulières de p avec $\|p\| = 1$

- Par définition de la fonction support,

$$x \in A \Rightarrow x \in B = \{\forall p \in \mathbb{X}, \langle p\,, x\rangle \leq \sigma_A(p)\}\,.$$

- B est un convexe fermé. On a donc $A \subset \overline{\mathrm{co}}\,(A) \subset B$

Theorem

$$B = \{x \in \mathbb{X} \mid \forall p \in \mathbb{X}, \langle p\,, x\rangle \leq \sigma_A(p)\} = \overline{\mathrm{co}}\,(A) \quad \textit{et } \sigma_{\overline{\mathrm{co}}(A)} = \sigma_A \; .$$

On a déjà $B \supset \overline{\mathrm{co}}\,(A)$. S'il existe $a \in B$ et $a \notin \overline{\mathrm{co}}\,(A)$, alors on trouve $r \in \mathbb{X}$ t.q.

$$\sup_{x \in \overline{\mathrm{co}}(A)} \langle r\,, x\rangle < \langle r\,, a\rangle \; .$$

Mais, par ailleurs,

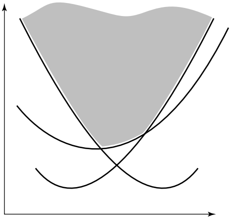
$$\sigma_A(r) \stackrel{\mathrm{def}}{=} \sup_{x \in A} \langle r\,, x\rangle \leq \sup_{x \in \overline{\mathrm{co}}(A)} \langle r\,, x\rangle \; .$$

Il existe donc r tel que $\sigma_A(r) < \langle r\,, a\rangle$, qui contredit $a \in B$. La définition de B implique que $\sigma_B \leq \sigma_A$. mais évidemment $\sigma_A \leq \sigma_{\overline{\mathrm{co}}(A)} = \sigma_B$ d'où l'égalité.

Opérations préservant la convexité des fonctions

L'intersection d'épigraphes correspond à prendre l'enveloppe supérieure des fonctions.

$$f : x \mapsto \sup_{j \in J} f_j(x) \text{ est convexe si } f_j \text{ est convexe, } \forall j \in J \ .$$



Enveloppe supérieure de fonctions convexes.

La fonction support est convexe dès que $A \neq \emptyset$: c'est l'enveloppe supérieure d'une famille de fonctions affines.

Opérations préservant la convexité des fonctions

- La somme pondérée à coefficients positifs de fonctions convexes est convexe.
- Transformation affine : $f : x \mapsto f(Ax + b)$ est convexe si f est convexe, et $A \in \mathcal{L}(\mathbb{X}, \mathbb{X})$, $b \in \mathbb{X}$.
- L'homothétique d'un sous-ensemble convexe reste convexe.

$$\forall \alpha > 0, \ \ \mathrm{epi}\,(g) = \alpha\,\mathrm{epi}\,(f) \iff g(x) = \alpha f(x/\alpha) \ .$$

  La fonction g est convexe lorsque f est convexe.
- composition d'une fonction $f : \mathbb{X} \to \mathbb{R}$ avec une fonction $g : \mathbb{R} \to \mathbb{R}$ est convexe si f est convexe et si g est convexe et monotone non décroissante.

Opérations préservant la convexité des fonctions

- La restriction d'une application $f : \mathbb{X} \times \mathbb{Y} \to \overline{\mathbb{R}}$ à l'un des espaces, c'est-à-dire par exemple l'application

$$g_y : \mathbb{X} \to \overline{\mathbb{R}}, x \mapsto f(x, y) \text{ } (y \text{ fixé})$$

  est convexe si f est convexe sur l'espace $\mathbb{X} \times \mathbb{Y}$ (on dira qu'elle est conjointement convexe en (x, y)).
- Il ne suffit pas que toutes les restrictions g_y pour tout y et les restrictions analogues à x fixé soient convexes pour pouvoir affirmer que f est (conjointement) convexe.

Inf-convolution

Definition

On définit une nouvelle fonction h, appelée inf-convolution de f et g, opération notée $f \mathbin{\square} g$, de la façon suivante :

$$h(x) = (f \mathbin{\square} g)(x) \stackrel{\text{def}}{=} \inf_{y \in \mathbb{X}} \big(f(y) + g(y - x) \big) \ . \tag{16}$$

- Si les deux fonctions f et g sont minorées par une même fonction affine alors h ne prend pas la valeur $-\infty$.
- L'épigraphe strict de h vérifie $\mathrm{epi}_s\,(h) = \mathrm{epi}_s\,(f) + \mathrm{epi}_s\,(g)$.
- Si f et g sont convexes, alors leur inf-convolution est convexe.

Marginales

Soit f une fonction sur le produit $\mathbb{X} \times \mathbb{Y}$ de deux espaces de Hilbert conjointement convexe en (x, y) et bornée inférieurement sur $\mathbb{X} \times \mathbb{Y}$.

Definition

On appelle fonction "marginale" g définie sur $\mathbb{X}$ la fonction :

$$g(x) \stackrel{\text{def}}{=} \inf_{y \in \mathbb{Y}} f(x, y) \; .$$

- g est bornée inférieurement sur $\mathbb{X}$
- $\mathrm{epi_s}\,(g)$ est la projection de $\mathrm{epi_s}\,(f)$ sur la composante $\mathbb{X}$
- La fonction marginale g est convexe

Introduction de la transformée de Fenchel

Pour une fonction $f : \mathbb{X} \to \mathbb{R}$, sur $\mathbb{X} \times \mathbb{R}$ muni du produit scalaire $\langle \cdot \, , \cdot \rangle_{\mathbb{X}} + \cdot \times \cdot$, on considère $\sigma_{\mathrm{epi}(f)}$.

- Si $\varpi > 0$ alors $\sigma_{\mathrm{epi}(f)}(p, \varpi) = +\infty$.
- Si $\varpi = 0$, calculer $\sigma_{\mathrm{epi}(f)}(p, 0)$, revient à calculer au point p, la fonction support de la projection de $\mathrm{epi}\,(f)$ sur $\mathbb{X}$ qui n'est autre que $\mathrm{dom}\,(f)$.
- Si $\varpi < 00$, on utilise $\varpi = -1$ (positive homogènéité de σ_A)

$$
\begin{aligned}
\sigma_{\mathrm{epi}(f)}(p, -1) &= \sup_{(x,\xi) \in \mathrm{epi}(f)} \big(\langle p \, , x \rangle - \xi \big) \\
&= \sup_{\xi \geq f(x)} \big(\langle p \, , x \rangle - \xi \big) \\
&= \sup_{x \in \mathrm{dom}\,(f)} \big(\langle p \, , x \rangle - f(x) \big) \, .
\end{aligned}
$$

Cette dernière formule définit ce qui est connu sous le nom de transformée de Fenchel de f.

Definition

Pour une fonction $f : \mathbb{X} \to \mathbb{R}$, on appelle transformée de Fenchel, notée $\mathcal{F}(f)$, la fonction de $\mathbb{X} \to \overline{\mathbb{R}}$ suivante :

$$\mathcal{F}(f) : \mathbb{X} \to \overline{\mathbb{R}}, \;\; \mathcal{F}(f)(p) = \sup_{x \in \mathrm{dom}\,\big(f\big)} \Big(\langle p \,, x \rangle - f(x) \Big) \;.$$

- $\mathcal{F}(f)$ est convexe : enveloppe supérieure d'une famille de fonctions affines en p
- Par ailleurs,

$$\forall p, \forall x, \;\; \mathcal{F}(f)(p) + f(x) \geq \langle p \,, x \rangle \;\; .$$

Quelques exemples

- Fonction indicatrice

$$\mathcal{F}(\mathrm{I}_A) = \sigma_A(p) \ .$$

- Si $h(x) \stackrel{\text{def}}{=} \min\big(f(x), g(x)\big)$ alors

$$\mathcal{F}(h) = \sup(\mathcal{F}(f), \mathcal{F}(g)) \ .$$

- L'épigraphe (strict) de $h = f \,\square\, g$ est égal à la somme des épigraphes (stricts) de f et g :

$$\mathcal{F}(f \,\square\, g) = \mathcal{F}(f) + \mathcal{F}(g)$$

- L'épigraphe de la fonction marginale g est la projection de l'épigraphe de la fonction initiale f.

$$\mathcal{F}(g)(p) = \mathcal{F}(f)(p, 0)$$

Enveloppe convexe

Enveloppe convexe de l'épigraphe d'une fonction quelconque :

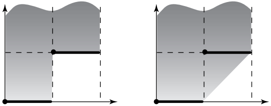
- Éviter que l'enveloppe convexe ait des valeurs $-\infty$

$$f : x \in \mathbb{R} \mapsto -x^2 \in \mathbb{R} \tag{17}$$

Demander que la fonction admet une minorante affine.

- L'enveloppe convexe d'un épigraphe n'est pas un épigraphe :

$$x \mapsto \begin{cases} +\infty & \text{si } x \notin [0, 2[\\ \lfloor x \rfloor \text{ (partie entière de } x) & \text{sinon.} \end{cases}$$



Une fonction, son épigraphe et l'enveloppe convexe de son épigraphe

Récupérer ces bords inférieurs par une opération d'infimum.

Enveloppe convexe

Theorem (Enveloppe convexe d'une fonction)

Pour une fonction f donnée, si elle est minorée par une fonction affine, il existe une plus grande fonction convexe inférieure ou égale à f. Elle est appelée enveloppe convexe de f notée $\mathrm{co}\,(f)$. On a que

$$\mathrm{co}\,(f)\,(x) = \inf\,\{\xi \mid (x,\xi) \in \mathrm{co}\,(\mathrm{epi}\,(f))\}\ .$$

Par ailleurs, $\mathrm{co}\,(f)$ est aussi donnée par la formule :

$$\mathrm{co}\,(f)\,(x) = \inf\left\{\sum_{i=1}^{n} \alpha_i f(x_i) \mid n \in \mathbb{N}, \alpha \in \Sigma_n, \sum_{i=1}^{n} \alpha_i x_i = x\right\}\ .$$

Lemma

Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ une fonction convexe sur un espace de Hilbert $\mathbb{X}$ et $a \in \mathbb{X}$ un point au voisinage duquel la fonction est bornée. Alors, f est continue en a.

On se ramène à l'origine. On suppose qu'il existe un voisinage V (Symétrique quitte à utiliser $V \cap (-V)$) de 0 où f est bornée par une constante M. Soit $\alpha \in]0, 1[$ et $x \in \alpha V$, alors $(x/\alpha) \in V$ et $(-x/\alpha) \in V$, et donc $f(x/\alpha) \leq M$ et $f(-x/\alpha) \leq M$:

$$\begin{aligned}
x = (1-\alpha)0 + \alpha(x/\alpha) &\Rightarrow f(x) \leq (1-\alpha)f(0) + \alpha f(x/\alpha) \\
&\Rightarrow f(x) \leq \alpha M \, , \\
0 = \frac{1}{1+\alpha}x + \frac{\alpha}{1+\alpha}(-x/\alpha) &\Rightarrow f(0) \leq \frac{1}{1+\alpha}f(x) + \frac{\alpha}{1+\alpha}f(-x/\alpha) \\
&\Rightarrow f(x) \geq (1+\alpha)f(0) - \alpha f(-x/\alpha) \\
&\Rightarrow f(x) \geq -\alpha M \, .
\end{aligned}$$

Theorem

Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ une fonction convexe sur un espace de Hilbert $\mathbb{X}$, f est continue sur l'intérieur de son domaine qui est non vide si et seulement si il existe un ouvert sur lequel f est bornée.

Supposons qu'il existe un ouvert V sur lequel f est bornée par M.

Alors, $V \subset \mathrm{dom}\,(f)$, par définition de $\mathrm{dom}\,(f)$ et $\overset{\circ}{\mathrm{dom}}\,(f) \neq \varnothing$.

Soit $z \in \overset{\circ}{\mathrm{dom}}\,(f)$, $x \in V$ et α assez petit, alors :

$y = (z - \alpha x)/(1 - \alpha) \in \overset{\circ}{\mathrm{dom}}\,(f)$. Le point z est maintenant une combinaison convexe de x et de y. Une homothétie de centre y et de rapport α amène x sur z et V se transforme en un voisinage W centré autour de z.

$$f(a) \leq \alpha M + (1 - \alpha) f(y) \leq M' \,, \forall a \in W$$

f est bien bornée au voisinage de z.

Le cas de la dimension finie

Corollary

Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ une fonction convexe sur un espace de Hilbert $\mathbb{X}$ de dimension finie. Alors f restreinte à l'intérieur relatif de son domaine est continue.

- On écarte le cas $f(x) = +\infty$ pour tout $x \in \mathbb{X}$.
- $\mathrm{dom}\left(f\right)$ est un convexe non vide et donc $\mathrm{ri}\left(\mathrm{dom}\left(f\right)\right) \neq \varnothing$
- Soit n la dimension de $\mathrm{aff}\left(\mathrm{dom}\left(f\right)\right)$. Il faut vérifier que $\mathrm{ri}\left(\mathrm{dom}\left(f\right)\right)$ contient un ouvert sur lequel f est bornée.
- Or $\mathrm{ri}\left(\mathrm{dom}\left(f\right)\right)$ contient un polyèdre ouvert P de $(n+1)$ sommets $\{x_i\}_{i=1,\dots,n}$ et par convexité :

$$\forall x \in P \quad |f(x)| \leq max_{i=1}^{n+1}|f(x_i)| \leq K \tag{18}$$

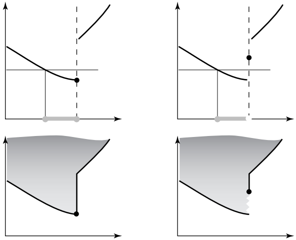
Fermeture de l'épigraphe et semi-continuité inférieure

Les fonctions dont l'épigraphe est fermé ?

Definition

Une fonction est semi-continue inférieurement en un point x si

$$f(x) \leq \liminf_{y \to x} f(y) \; .$$



Une fonction s.c.i., une modification non s.c.i., les ensembles de niveau et les épigraphes

Theorem

Pour une fonction $f : \mathbb{X} \to \mathbb{R}$ sont équivalents :

1. *f est s.c.i. ;*
2. *son épigraphe est fermé ;*
3. *ses ensembles de niveau sont fermés.*

$1 \Rightarrow 2$: Soit $\{(x^k, \xi^k)\} \subset \mathrm{epi}\,(f)$ convergeant vers (x, ξ). $\xi = \lim \xi^k \geq \liminf f(x^k) \geq f(x)$ d'où $\mathrm{epi}\,(f)$ est fermé.

$2 \Rightarrow 3$: $\nabla_\beta\,(f)$ est la section de $\mathrm{epi}\,(f)$ par l'hyperplan "horizontal" au niveau β ramenée sur $\mathbb{X}$ par une translation de $(0, -\beta)$.

$3 \Rightarrow 1$: Si f n'est pas s.c.i. en x, il existe une suite $x^k \to x$ et $f(x^k) \to \beta < f(x)$. Soit $\beta' \in]\beta, f(x)[$, à partir d'un certain rang $\{x^k\} \in \nabla_{\beta'}\,(f)$. Comme $x \notin \nabla_{\beta'}\,(f)$, $\nabla_{\beta'}\,(f)$ ne peut être fermé.

Topologie faible

Corollary

Une fonction convexe $f : \mathbb{X} \to \overline{\mathbb{R}}$ qui est s.c.i. pour $\mathbb{X}$ muni de la topologie forte est aussi s.c.i. pour $\mathbb{X}$ muni de la topologie faible.

Corollary

L'enveloppe supérieure d'une famille de fonctions s.c.i. est s.c.i.. En particulier, l'enveloppe supérieure d'une famille de fonctions affines continues ou convexes s.c.i. est convexe et s.c.i.. Donc la classe des fonctions convexes s.c.i. est stable par enveloppe supérieure.

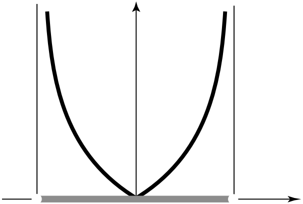
L'enveloppe supérieure d'une famille de fonctions a pour épigraphe l'intersection des épigraphes des fonctions de cette famille et cette intersection est fermée si tous ces épigraphes le sont.

Domaine

Le domaine $\operatorname{dom}(f)$ d'une fonction convexe s.c.i. n'est pas nécessairement fermé.

$$f : \mathbb{R} \to \overline{\mathbb{R}}, \quad f(x) = \begin{cases} |\operatorname{tang} x| & \text{pour } -\pi/2 < x < \pi/2 \\ +\infty & \text{sinon.} \end{cases} \tag{19}$$

Cela montre au passage que la projection d'un convexe fermé n'est pas nécessairement fermée.



Graphe de la fonction (19) et son domaine

Fonction indicatrice

La fonction indicatrice d'un sous-ensemble A est convexe s.c.i. si et seulement si le sous-ensemble est convexe fermé.

Enveloppe convexe s.c.i.

Theorem (Enveloppe convexe s.c.i. d'une fonction)

Pour une fonction f donnée, si elle est minorée par une fonction affine, il existe une plus grande fonction convexe s.c.i. inférieure ou égale à f. Elle est appelée enveloppe convexe s.c.i. de f notée $\overline{\mathrm{co}}\,(f)$. On a que

$$\mathrm{epi}\,(\overline{\mathrm{co}}\,(f)) = \overline{\mathrm{co}}\,(\mathrm{epi}\,(f)) = \overline{\mathrm{co}\,(\mathrm{epi}\,(f))}\ .$$

Par ailleurs,

$$\overline{\mathrm{co}}\,(f)\,(x) = \liminf_{y \to x} \mathrm{co}\,(f)\,(y)\ .$$

Lemma

Considérons une fonction $f : \mathbb{X} \to \overline{\mathbb{R}}$ *convexe s.c.i.. Son épigraphe est égal à l'intersection des épigraphes de tous les fonctions affines continues qui minorent* f *ainsi que de tous les demi-espaces "verticaux" dans* $\mathbb{X} \times \mathbb{R}$ *dont la projection (ou trace) sur* $\mathbb{X}$ *est un hyperplan d'appui de* $\overline{\mathrm{dom}\,(f)}$.

Corollary

Toute fonction convexe s.c.i. est égale à l'enveloppe supérieure de toutes les fonctions affines continues qui minorent la fonction.

Enveloppe convexe s.c.i.

Corollary

L'enveloppe convexe s.c.i. d'une fonction $f : \mathbb{X} \to \overline{\mathbb{R}}$ *est donnée par la formule*

$$\overline{\mathrm{co}}\,(f)\,(x) = \sup_{(r,\alpha)} \left\{ \langle r\,, x\rangle + \alpha \mid \forall y \in \mathbb{X}, \langle r\,, y\rangle + \alpha \leq f(y) \right\} \ .$$

$H(f)(x) \overset{\mathrm{def}}{=} \sup_{(r,\alpha)} \left\{ \langle r\,, x\rangle + \alpha \mid \forall y \in \mathbb{X}, \langle r\,, y\rangle + \alpha \leq f(y) \right\}$

- $H(f)(x) \leq f(x)$ pour tout x ;
- $H(f)$ est convexe s.c.i. (sup de fonctions affines) ;
- $H(f)(x)$ est une fonction croissante de la fonction f.
- $H\big(\overline{\mathrm{co}}\,(f)\big) = \overline{\mathrm{co}}\,(f)$

Comme $\overline{\mathrm{co}}\,(f) \leq f$, alors $\overline{\mathrm{co}}\,(f) = H\big(\overline{\mathrm{co}}\,(f)\big) \leq H(f) \leq f$. Comme $H(f)$ est convexe s.c.i. et minorante on a aussi $H(f) \leq \overline{\mathrm{co}}\,(f)$.

On peut être plus précis en ce qui concerne la famille de fonctions affines continues à considérer. En effet pour r fixé, le plus grand α possible tel que

$$\alpha \leq f(y) - \langle r\,, y \rangle\,, \;\; \forall y \in \mathbb{X}\,,$$

est égal à

$$\begin{aligned} \alpha &= \inf_{y \in \mathbb{X}} \left(f(y) - \langle r\,, y \rangle \right) = - \sup_{y \in \mathbb{X}} \left(\langle r\,, y \rangle - f(y) \right) \\ &= - \sup_{y \in \mathrm{dom}\left(f\right)} \left(\langle r\,, y \rangle - f(y) \right) = -\mathcal{F}(f)(r)\,, \end{aligned}$$

Par conséquent :

$$\overline{\mathrm{co}}\,(f)\,(x) = \sup_{r \in \mathbb{X}} \left(\langle r\,, x \rangle - \mathcal{F}(f)(r) \right)$$

qui n'est rien d'autre que $\mathcal{F}(\left(\mathcal{F}(f)\right))(x)$ noté $\mathcal{F}^{(2)}(f)(x)$.
On peut donc énoncer le corollaire suivant.

Enveloppe convexe s.c.i.

Corollary

L'enveloppe convexe s.c.i. d'une fonction f est égale à $\mathcal{F}^{(2)}(f)$. La transformée de Fenchel est involutive pour les fonctions convexes s.c.i..

Definition (Dérivée directionnelle)

On appelle dérivée directionnelle de $f : \mathbb{X} \to \mathbb{Y}$ ($\mathbb{X}$ et $\mathbb{Y}$ sont deux espaces de Hilbert) au point $x \in \mathbb{X}$ et dans la direction $d \in \mathbb{X}$, notée $Df(x; d)$, la limite suivante, si elle existe dans $\mathbb{Y}$:

$$Df(x; d) = \lim_{\varepsilon \to 0_+} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon} .$$

Par exemple, pour la fonction $|x|$ de $\mathbb{R}$ dans $\mathbb{R}$

$$Df(0; 1) = 1 \quad \text{et} \quad Df(0; -1) = 1 .$$

Sur cet exemple, il apparaît que la dérivée directionnelle en un point x peut ne pas être une fonction linéaire de la direction d.

Dérivée de Gâteaux

Definition (Dérivée de Gâteaux)

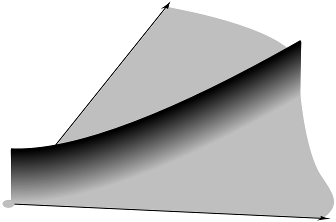
Si f admet en x des dérivées directionnelles pour toutes les directions d et si $Df(x; d)$ est une fonction linéaire continue de d, c'est-à-dire qu'il existe un opérateur $A \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$ telle que $Df(x; d) = A(d)$ pour tout d, alors la fonction f est dite différentiable au sens de Gâteaux au point x ; l'opérateur A, qui sera noté $f'(x)$, est appelée dérivée de Gâteaux de f au point x. Si une telle dérivée existe en tout point où f est définie, on dira que f est différentiable au sens de Gâteaux.

Diverses notions de différentiabilité

L'existence d'une dérivée de Gâteaux en un point n'implique pas la continuité de la fonction en ce point.

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = \begin{cases} 1 & \text{si } y = x^2 \text{ et } x \neq 0 , \\ 0 & \text{sinon,} \end{cases}$$

La dérivée de Gâteaux existe en $(0,0)$ mais la fonction n'est pas continue en ce point.



Gâteaux-différentiable
mais
non continue en $(0,0)$.

Dérivée de Fréchet

Il existe une notion plus forte que la dérivée de Gâteaux qui interdit ce genre de paradoxe.

Definition (Dérivée de Fréchet)

La fonction $f : \mathbb{X} \to \mathbb{Y}$ admet une dérivée de Fréchet au point $x \in \mathbb{X}$ si il existe un opérateur linéaire continu, encore noté $f'(x) \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$, tel que :

$$\lim_{\|d\| \to 0} \frac{f(x+d) - f(x) - f'(x)(d)}{\|d\|} = 0$$

et ce pour tout $d \in \mathbb{X}$ tendant vers 0.

En Optimisation où l'on est surtout préoccupé de comparer des valeurs de fonctions coût, la notion de dérivée de Gâteaux suffit. C'est ce que désignera désormais la notation $f'(x)$.

Lemma

Si $f : \mathbb{X} \to \overline{\mathbb{R}}$ *est convexe, l'application* $\varepsilon \mapsto \big(f(x + \varepsilon d) - f(x)\big)/\varepsilon$ *est une fonction non décroissante de* $\varepsilon \in \mathbb{R}$ *pour tout* $x \in \mathbb{X}$ *et* $d \in \mathbb{X}$ *(partout où elle est définie).*

Corollary

Une fonction convexe $f : \mathbb{X} \to \overline{\mathbb{R}}$ *admet toujours des dérivées directionnelles éventuellement égales à* $\pm\infty$*. De plus,*

$$Df(x; d) = \inf_{\varepsilon > 0} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon} .$$

Enfin, $-Df(x; -d) \leq Df(x; d)$*.*

Dérivée directionnelle

- La fonction $d \mapsto Df(x;d)$ est positivement homogène de degré 1
- Elle est convexe si f est convexe.
- Elle est sous-additive, c'est-à-dire que

$$Df(x;d+d') \leq Df(x;d) + Df(x;d') \ .$$

En fait, Les fonctions positivement homogènes de degré 1 sont convexes si et seulement si elles sont sous-additives. Observer que l'épigraphe d'une fonction positivement homogène est un cône.

Sous-gradients et interprétation géométrique

Une“ minorante affine continue” d'une fonction $f : \mathbb{X} \to \overline{\mathbb{R}}$ est définie par un “coefficient directeur” $r \in \mathbb{X}$ et une “ordonnée à l'origine” α tels que

$$f(y) \geq \langle r , y \rangle + \alpha \,, \ \ \forall y \in \mathbb{X} \,,$$

et cette minorante est “exacte” en x si l'égalité a lieu en $y = x$ dans l'égalité précédente. Alors, $\alpha = f(x) - \langle r , x \rangle$ et l'inégalité précédente se réécrit

$$f(y) - f(x) \geq \langle r , y - x \rangle \ , \ \ \forall y \in \mathbb{X} \,.$$

Sous-gradients et interprétation géométrique

Definition (Sous-gradient, sous-différentiel)

Pour une fonction $f : \mathbb{X} \to \overline{\mathbb{R}}$, on appelle sous-gradient de f en x (avec $f(x) < +\infty$) tout vecteur $r \in \mathbb{X}$ qui vérifie :

$$f(y) \geq \langle r \,, y - x \rangle + f(x) \,, \quad \forall y \in \mathbb{X}$$

L'ensemble (vide ou non vide) de ces vecteurs est appelé sous-différentiel de f en x et est noté $\partial f(x)$.

- Forme équivalente : Si $r \in \partial f(x)$ l'hyperplan d'équation $\langle r \,, y \rangle - \xi = \langle r \,, x \rangle - f(x)$ est un hyperplan d'appui de $\mathrm{epi}\,(f)$ au point $\big(x, f(x)\big)$.
- En particulier, $(r, -1)$ est le “coefficient directeur” de cet hyperplan.

Sous-gradients et interprétation géométrique

Lemma

Un vecteur r appartient à $\partial f(x)$ pour une fonction convexe $f : \mathbb{X} \to \overline{\mathbb{R}}$ si et seulement si $(r, -1)$ appartient au cône normal à $\mathrm{epi}\,(f)$ au point $\big(x, f(x)\big)$ et si r appartient au cône normal à l'ensemble de niveau $\triangledown_{f(x)}(f)$ au point x.

Pour un convexe fermé A, $\partial \mathrm{I}_A(x)$ est égal à $A_x^{\perp}$.

Theorem

Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ une fonction convexe qui est finie et continue en au moins un point x. Alors, le sous-différentiel de f est non vide en tout point de l'intérieur de son domaine (qui est lui-même non vide) et en particulier en x.

Sous-différentiel et lien avec la dérivée directionnelle

Theorem

Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ une fonction convexe. En un point $x \in \mathbb{X}$, on a :

$$\partial f(x) = \{ r \in \mathbb{X} \mid \forall d \in \mathbb{X}, \langle r \,, d \rangle \leq Df(x; d) \} \ . \tag{20}$$

Corollary

Le sous-différentiel $\partial f(x)$ d'une fonction f convexe en un point x est un sous-ensemble convexe fermé (éventuellement vide).

C'est clair ici puisque $\partial f(x)$ apparaît dans (20) comme l'intersection d'une famille de demi-espaces fermés.

Sous-différentiel et lien avec la dérivée directionnelle

Corollary

Si $f : \mathbb{X} \to \overline{\mathbb{R}}$ est une fonction convexe et si elle est finie et continue en x, alors la fonction $d \mapsto Df(x; d)$ est la fonction support de $\partial f(x)$. Autrement dit,

$$Df(x, d) = \sup_{r \in \partial f(x)} \langle r \, , d \rangle \ . \tag{21}$$

Si f est finie et continue en x, elle est bornée au voisinage de x et $\partial f(x)$ est non vide. Soit $r \in \partial f(x)$. Alors,

$$\langle r \, , d \rangle \leq Df(x; d) \leq f(x + d) - f(x) \ .$$

$Df(x; d)$ est bornée pour toute direction d (si l'on borne $\|d\|$). Comme elle est convexe en d on en déduit qu'elle est continue.

Sous-différentiel et lien avec la dérivée directionnelle

Theorem

Soit $h : \mathbb{X} \to \overline{\mathbb{R}}$ une fonction convexe s.c.i. et positivement homogène de degré 1. Alors $h = \sigma_{S_h}$ défini par

$$S_h \stackrel{\text{def}}{=} \{r \in \mathbb{X} \mid \langle r , d \rangle \leq h(d), \ \forall d \in \mathbb{X}\} \ .$$

On considère les minorantes affines de h (qui est convexe s.c.i.) : $\langle r , d \rangle + \alpha \leq h(d), \ \forall d \in \mathbb{X}$. h est homogène, alors $h(0) = 0$ et donc $\alpha \leq 0$. De même pour tout $\varepsilon > 0$, $\varepsilon \langle r , d \rangle + \alpha \leq \varepsilon h(d)$, La fonction linéaire $d \mapsto \langle r , d \rangle$ est une meilleure minorante.

$$\forall d, \ \ h(d) = \sup \left\{ \langle r , d \rangle \mid \forall r, \forall y, \langle r , y \rangle \leq h(y) \right\} \ ,$$

mais, d'après la définition de S_h, ceci se réécrit :

$$\forall d, \ \ h(d) = \sup \left\{ \langle r , d \rangle \mid \forall r \in S_h \right\} = \sup_{r \in S_h} \langle r , d \rangle \ .$$

Cas des fonctions convexes différentiables

Theorem

Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ une fonction convexe, finie et continue en $x \in \mathbb{X}$. Elle admet une dérivée $f'(x)$ si et seulement si $\partial f(x)$ est réduit à un seul point, à savoir $f'(x)$, et donc

$$f(y) - f(x) \geq \left\langle f'(x) \, , y - x \right\rangle \ , \ \ \forall y \in \mathbb{X} \, . \tag{22}$$

Si $\partial f(x)$ est réduit à un singleton, $Df(x; d)$ est une fonction linéaire de d, f est alors (Gâteaux-) différentiable au point x.
Si f est différentiable, alors on a $Df(x; y - x) = \langle f'(x) \, , y - x \rangle$.
Mais, $Df(x; y - x) \leq f(y) - f(x)$, d'où (22).
Si il existe un autre $r \in \partial f(x)$. $Df(x; d) = \langle f'(x) \, , d \rangle$ pour tout d, mais par ailleurs, $Df(x; d) \geq \langle r \, , d \rangle$ et donc $\langle f'(x) \, , d \rangle \geq \langle r \, , d \rangle$ pour tout d. pour $d = r - f'(x)$, on obtient $-\left\| r - f'(x) \right\|^2 \geq 0$.

Theorem

Une fonction $f : \mathbb{X} \to \mathbb{R}$ différentiable est convexe si et seulement si elle vérifie (22) *pour tous x et y.*

Il reste à montrer le "si". On suppose (22) pour tous x et y et on utilise cette inégalité pour $y = u$ et $x = u + \alpha(v - u)$, $\alpha \in [0, 1]$, $u \in \mathbb{X}$, $v \in \mathbb{X}$, ce qui donne

$$f(u) - f\big(u + \alpha(v - u)\big) \geq \alpha \left\langle f'\big(u + \alpha(v - u)\big), u - v \right\rangle .$$

Avec $y = v$ et le même x, on obtient

$$f(v) - f\big(u + \alpha(v - u)\big) \geq (1 - \alpha) \left\langle f'\big(u + \alpha(v - u)\big), v - u \right\rangle .$$

La combinaison convexe de ces deux inégalités avec les "poids" $(1 - \alpha)$, respectivement α, conduit finalement à l'inégalité de la convexité.

Le sous-différentiel comme opérateur multivoque

La correspondence $x \mapsto \partial f(x)$ est une application multivoque ou multi-application, c'est-à-dire qu'elle associe à un point $x \in \mathbb{X}$ le sous-ensemble $\partial f(x)$ de $\mathbb{X}$. De plus, elle est ici à "valeurs" convexes fermées.

Definition

Une multi-application $A : \mathbb{X} \to 2^{\mathbb{X}}$ est monotone si

$$\forall x \in \mathbb{X}, \forall y \in \mathbb{X}, \forall r \in A(x), \forall s \in A(y), \ \ \langle r - s \,, x - y \rangle \geq 0 \,.$$

Elle est strictement monotone si l'inégalité stricte a lieu ci-dessus dès que $y \neq x$. Elle est fortement monotone (de module b) si

$$\exists b > 0 : \ \ \forall x \in \mathbb{X}, \forall y \in \mathbb{X}, \forall r \in A(x), \forall s \in A(y),$$
$$\langle r - s \,, x - y \rangle \geq b \left\| x - y \right\|^2 \ .$$

Lemma

Le sous-différentiel d'une fonction convexe est une multi-application monotone.
Réciproquement, pour une fonction sous-différentiable, si le sous-différentiel est monotone, la fonction est convexe.

Le lemme précédent peut s'énoncer avec strictement convexe et strictement monotone et avec convexe fortement monotone de module b et fortement monotone de module b.

Calcul sous-différentiel

- Soit f et g deux fonctions convexes sous-différentiables de $\mathbb{X}$ dans $\overline{\mathbb{R}}$.

$$\forall x \in \mathbb{X}, \;\; \partial(f+g)(x) \supset \partial f(x) + \partial g(x) \; . \tag{23}$$

  L'égalité peut être démontrée si il existe un point dans $\mathrm{dom}\left(f\right) \cap \mathrm{dom}\left(g\right)$ où au moins l'une des fonctions est continue.

- Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ une fonction convexe et $A : \mathbb{Y} \to \mathbb{X}$ un opérateur linéaire continu. Montrer que

$$\forall x \in \mathbb{X}, \;\; \partial(f \circ A)(x) \supset A^{*} \circ \partial f(Ax) \; . \tag{24}$$

  L'égalité peut être démontrée si il existe un point dans $\mathrm{dom}\left(f\right) \cap \mathrm{im}\, A$ où f est continue.

Soit $f : \mathbb{X} \to \overline{\mathbb{R}}$ définie comme l'enveloppe supérieure d'une famille de fonctions convexes $\{f_j\}_{i \in J}$:

$$f(x) \stackrel{\text{def}}{=} \sup_{j \in J} f_j(x) \ .$$

A priori, rien n'est dit sur la nature de l'ensemble J. On pose

$$J(x) = \arg\max_{j \in J} f_j(x) \stackrel{\text{def}}{=} \{j \in J \mid f(x) = f_j(x)\} \ .$$

On a

$$\partial f(x) \supset \overline{\mathrm{co}} \left(\bigcup_{j \in J(x)} \partial f_j(x) \right) \ . \tag{25}$$

La formule (25) montre en tout cas que lorsque $J(x)$ n'est pas réduit à un point il y a peu de chance que f soit différentiable même si toutes les fonctions f_j le sont.

Soit $f : \mathbb{X} \times \mathbb{Y} \to \overline{\mathbb{R}}$ une fonction convexe et sous-différentiable.

$$\Pi_{\mathbb{X}}\big(\partial f(x,y)\big) = \partial_x f(x,y) \ , \tag{26}$$

où $\partial_x f(x,y)$ sous-différentiel de la fonction $x \mapsto f(x,y)$ à y donné.

$$\partial f(x,y) \subset \partial_x f(x,y) \times \partial_y f(x,y) \ . \tag{27}$$

Soit $f : \mathbb{X} \times \mathbb{Y} \to \overline{\mathbb{R}}$ une fonction convexe et soit g la marginale. On pose

$$Y(x) = \arg\min_{y \in \mathbb{Y}} f(x,y) \stackrel{\text{def}}{=} \{y \in \mathbb{Y} \mid g(x) = f(x,y)\} \ .$$

Montrer que

$$\partial g(x) \subset \bigcap_{y \in Y(x)} \partial_x f(x,y) \ . \tag{28}$$

Theorem

Soit f une fonction convexe. L'inégalité

$$\forall p, \forall x, \ \ \mathcal{F}(f)(p) + f(x) \geq \langle p \, , x \rangle \ .$$

est une égalité si et seulement si $p \in \partial f(x)$. De plus, si f est s.c.i., alors l'égalité est aussi équivalente au fait que $x \in \partial \mathcal{F}(f)(p)$.

De $\mathcal{F}(f)(p) + f(x) = \langle p \, , x \rangle$ et $\mathcal{F}(f)(p) + f(y) \geq \langle p \, , y \rangle$ on obtient que $p \in \partial f(x)$.
Réciproquement, si $p \in \partial f(x)$, on a

$$\forall y, \ \ \langle p \, , y \rangle - f(y) \leq \langle p \, , x \rangle - f(x) \ ,$$

C'est x qui réalise le sup dans la définition de $\mathcal{F}(f)$ au point p.
Utiliser le couple de fonctions $(\mathcal{F}(f), \mathcal{F}^{(2)}(f) = f)$.

Inverse ensembliste

Introduisons la notion d'inverse ensembliste d'une application, voire même d'une multi-application $A : \mathbb{X} \to 2^{\mathbb{Y}}$: c'est la multi-application de $\mathbb{Y}$ dans $2^{\mathbb{X}}$ définie par

$$A^{-1}(y) = \{x \in \mathbb{X} \mid y \in A(x)\} \ .$$

Autrement dit, $y \in A(x)$ est par définition équivalent à $x \in A^{-1}(y)$. Alors, du Théorème 82, et pour une fonction convexe s.c.i. et différentiable, on déduit que

$$x \in \left(\partial f\right)^{-1}(p) \Leftrightarrow p \in \partial f(x) \Leftrightarrow x \in \partial \mathcal{F}(f)(p) \ ,$$

et par conséquent

$$\left(\partial f\right)^{-1} = \partial \mathcal{F}(f) \ . \tag{29}$$

Ceci montre que, considérée plutôt au niveau des sous-différentiels qu'à celui des fonctions, la transformée de Fenchel agit comme une inversion !

Formulation du problème et existence d'une solution

On va traiter du problème générique suivant qualifié de "problème d'optimisation convexe"

$$\mathcal{P} \qquad \min_{u \in U^{\mathrm{ad}}} J(u) \ ,$$

- U^{ad} est un sous-ensemble convexe fermé d'un espace de Hilbert $\mathbb{U}$ (dit "ensemble admissible")
- J est une fonction convexe s.c.i. de $\mathbb{U}$ dans $\overline{\mathbb{R}}$.
- On appellera "solution" tout u où $J(u)$ atteint effectivement sa valeur infimale sur U^{ad}.
- L'ensemble des solutions est noté $\arg\min_{u \in U^{\mathrm{ad}}} J(u)$.

Formulation équivalente

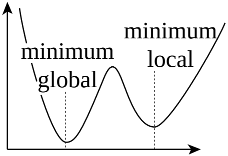
- On s'est autorisé des fonctions J prenant éventuellement la valeur $+\infty$.
- On pourrait se limiter à des problèmes de minimisation où U^{ad} est toujours égal à l'espace $\mathbb{U}$ tout entier
- Le problème $\mathcal{P}$ est équivalent au problème

$$\min_{u \in \mathbb{U}} \left(J(u) + \mathrm{I}_{U^{\mathrm{ad}}}(u) \right) , \tag{30}$$

Une première conséquence importante de la convexité est que toute solution locale (on dit aussi “minimum local”) est une solution globale (“minimum global”). Par “solution locale” ou “minimum local”, on entend un point $u^\circ \in U^{\mathrm{ad}}$ tel que

$$J(u^\circ) \leq J(u), \ \ \forall u \in U^{\mathrm{ad}} \cap \overline{B}(u^\circ, \rho) \ ,$$

pour un certain $\rho > 0$. Autrement dit, l'ensemble admissible est réduit à un voisinage du prétendu minimum local.



Minimum local et minimum global

Minimum global

Lemma

Pour un problème convexe, toute solution locale est globale.

Démonstration.

Soit $u \in U^{\mathrm{ad}} \setminus \overline{B}(u^\circ, \rho)$, $\alpha = \rho / \|u - u^\circ\| < 1$ et $u_\alpha = (1-\alpha)u^\circ + \alpha u$. On a $u_\alpha - u^\circ = \alpha(u - u^\circ)$ et ce vecteur est de norme égale à ρ, donc $x_\alpha \in \overline{B}(u^\circ, \rho)$, et par ailleurs, évidemment $u_\alpha \in U^{\mathrm{ad}}$. Alors,

$$J(u^\circ) \leq J(u_\alpha) \leq (1-\alpha)J(u^\circ) + \alpha J(u) \Rightarrow J(u^\circ) \leq J(u) \ .$$

Ceci étant vrai pour $u \in U^{\mathrm{ad}} \setminus \overline{B}(u^\circ, \rho)$, mais également pour $u \in U^{\mathrm{ad}} \cap \overline{B}(u^\circ, \rho)$ (par définition de u°) est donc vrai pour $u \in U^{\mathrm{ad}}$. $\square$

Terminologie

Definition

Une fonction J est coercive sur U^{ad} si

$$\lim J(u) = +\infty \quad \text{lorsque} \quad \|u\| \to +\infty \quad \text{et} \quad u \in U^{\mathrm{ad}} \; .$$

- Dans toutes les directions de U^{ad} où u peut tendre vers l'infini, $J(u)$ tend vers $+\infty$.
- Si J est fortement convexe, alors elle est coercive sur tout sous-ensemble U^{ad}.
- Il est équivalent de dire que J est coercive sur U^{ad} ou que $J + \mathrm{I}_{U^{\mathrm{ad}}}$ est coercive sur $\mathbb{U}$.

Existence

Theorem

Si J est une fonction convexe s.c.i. et coercive sur U^{ad}, si U^{ad} est convexe et fermé, si $\mathrm{dom}\,(J) \cap U^{\mathrm{ad}} \neq \varnothing$, alors il existe au moins une solution au problème $\mathcal{P}$. L'ensemble des solutions est un sous-ensemble convexe fermé. Il est réduit à un point (la solution est donc unique) si J est strictement convexe.

- Sous $\mathrm{dom}\,(J) \cap U^{\mathrm{ad}} \neq \varnothing$ il existe au moins un $u \in U^{\mathrm{ad}}$ où $J(u)$ est finie.
- Une suite minimisante $\{u^k\}_{k\in\mathbb{N}} \subset U^{\mathrm{ad}}$, c'est-à-dire telle que $\lim_{k\to+\infty} J(u^k) = \inf_{u\in U^{\mathrm{ad}}} J(u)$ est bornée car sinon, avec la coercivité de J sur U^{ad}, on obtiendrait une contradiction.
- Elle est donc compacte dans la topologie faible, et on peut extraire une sous-suite $\{u^{k_i}\}$ faiblement convergente vers un certain $u^\sharp$.

- du fait que U^{ad} est un convexe fermé, donc faiblement fermé $u^\sharp \in U^{\mathrm{ad}}$.
- Une fonction convexe s.c.i. est aussi faiblement s.c.i. on a

$$\inf_{u \in U^{\mathrm{ad}}} J(u) = \lim_{k_i \to +\infty} J(u^{k_i}) \geq J(u^\sharp) \;,$$

- L'inf est fini et atteint.
- Toutes les autres solutions peuvent être obtenues comme le sous-ensemble

$$U^\sharp \stackrel{\mathrm{def}}{=} \nabla_{J(u^\sharp)}(J) \cap U^{\mathrm{ad}} \;.$$

- Si J est strictement convexe et si on suppose que l'on a deux solutions distinctes $u_1^\sharp$ et $u_2^\sharp$ (évidemment dans U^{ad}), alors le point $(u_1^\sharp + u_2^\sharp)/2$ (également dans U^{ad} par convexité de U^{ad}) donne nécessairement à J une valeur strictement inférieure à $\Big(J\big(u_1^\sharp\big) + J\big(u_2^\sharp\big)\Big)/2 = J\big(u_1^\sharp\big) = J\big(u_2^\sharp\big)$ par convexité stricte ce qui est une contradiction.

Theorem

En rajoutant l'hypothèse que J est bornée dans un ouvert et que $\overset{\circ}{\mathrm{dom}}\,(J) \cap U^{\mathrm{ad}} \neq \varnothing$ *chacune des conditions ci-dessous constitue une condition nécessaire et suffisante pour que $u^\sharp$ soit une solution du problème $\mathcal{P}$, en plus de la condition évidemment nécessaire que $u^\sharp \in U^{\mathrm{ad}}$:*

$$\forall u \in U^{\mathrm{ad}}, \;\; DJ(u^\sharp; u - u^\sharp) \geq 0 \, ,$$

$$0 \in \partial\big(J + \mathrm{I}_{U^{\mathrm{ad}}}\big)(u^\sharp) \, ,$$

$$\exists r^\sharp \in \partial J(u^\sharp) \cap \Big(-(U^{\mathrm{ad}})^{\perp}_{u^\sharp}\Big) \, ,$$

$$\exists r^\sharp \in \partial J(u^\sharp) : \;\; \forall u \in U^{\mathrm{ad}}, \;\; \Big\langle r^\sharp \, , u - u^\sharp \Big\rangle \geq 0 \, .$$

Autre caractérisation des solutions optimales

Lemma

Sous les hypothèses du Théorème précédent et si J est égale à la somme de deux fonctions convexes J_1 et J_2, on obtient la condition équivalente à $DJ_1(u^\sharp; u - u^\sharp)$:

$$\forall u \in U^{\mathrm{ad}}, \;\; DJ_1(u^\sharp; u - u^\sharp) + J_2(u) - J_2(u^\sharp) \geq 0 \; . \tag{32}$$

Autre caractérisation des solutions optimales

$$
\begin{aligned}
D(J_1 + J_2)(u^\sharp; u - u^\sharp) \\
&= \inf_{\varepsilon > 0} \Big((J_1 + J_2)\big(u^\sharp + \varepsilon(u - u^\sharp)\big) - (J_1 + J_2)(u^\sharp) \Big) / \varepsilon \\
&\leq \inf_{\varepsilon > 0} \Big(J_1\big(u^\sharp + \varepsilon(u - u^\sharp)\big) - J_1(u^\sharp) \\
&\qquad\qquad + (1 - \varepsilon) J_2(u^\sharp) + \varepsilon J_2(u) - J_2(u^\sharp) \Big) / \varepsilon
\end{aligned}
$$

par convexité de J_2

$$
\begin{aligned}
&= \inf_{\varepsilon > 0} \Big(J_1\big(u^\sharp + \varepsilon(u - u^\sharp)\big) - J_1(u^\sharp) + \varepsilon\big(J_2(u) - J_2(u^\sharp)\big) \Big) / \varepsilon \\
&= DJ_1(u^\sharp; u - u^\sharp) + J_2(u) - J_2(u^\sharp) \,,
\end{aligned}
$$

et donc la non négativité du premier membre entraîne celle du second membre.

Autre caractérisation des solutions optimales

Réciproquement, puisque $DJ_1(u^\sharp; u - u^\sharp) \leq J_1(u) - J_1(u^\sharp)$, alors $DJ_1(u^\sharp; u - u^\sharp) + J_2(u) - J_2(u^\sharp) \leq (J_1 + J_2)(u) - (J_1 + J_2)(u^\sharp)$, et la non négativité du premier membre entraîne celle du second.

Satisfaction d'un ensemble d'égalités et d'inégalités

L'ensemble admissible U^{ad} est décrit par :

$$\Theta_i(u) \leq 0, \ \ i = 1, \ldots, m,$$
$$\Omega_j(u) = 0, \ \ j = 1, \ldots, p.$$

Pour que cet ensemble de contraintes définisse un sous-ensemble U^{ad} convexe :

- les fonctions $\Theta_i : \mathbb{U} \to \mathbb{R}$ sont convexes, car la contrainte i définit l'ensemble de niveau $\overline{\nabla}_0 \left(\Theta_i\right)$ qui est alors convexe,
- les fonctions $\Omega_j : \mathbb{U} \to \mathbb{R}$ sont affines ; en effet, la contrainte j égalité peut être vue comme la réunion de deux contraintes

$$\Omega_j(u) \leq 0 \ \text{ et } \ -\Omega_j(u) \leq 0 \ ;$$

  si chacune doit correspondre à une fonction convexe, alors Ω_j doit être affine.

Description de contraintes explicites

- Sous les hypothèses que la fonction coût J est convexe, que les fonctions Θ_i sont convexes, et que les fonctions Ω_j sont affines, on parle de "programmation convexe".
- Le caractère fermé de U^{ad} est assuré si les fonctions Ω_j sont continues et si les fonctions Θ_i sont au moins s.c.i..
- Il n'y a pas unicité de la représentation d'un sous-ensemble admissible U^{ad} par un jeu de contraintes égalité et inégalité.

$$\Xi(u) \stackrel{\mathrm{def}}{=} \sum_{i=1}^{m} \max\big(0, \Theta_i(u)\big) + \sum_{j=1}^{p} |\Omega_j(u)| \ ,$$

$$\Psi(u) \stackrel{\mathrm{def}}{=} \max\left(\max_{i=1,\dots,m} \Big(\max\big(0, \Theta_i(u)\big)\Big), \max_{j=1,\dots,p} \Big(|\Omega_j(u)| \Big) \right) ,$$

U^{ad} peut être aussi décrit par $\Xi(u) \leq 0$ ou $\Psi(u) \leq 0$, ou même $\Xi(u) = 0$ ou $\Psi(u) = 0$

Lemma

On suppose les fonctions Θ_i sous-différentiables et on écrit les fonctions Ω_j affines et continues globalement sous la forme $A(u) = b$ où $b \in \mathbb{R}^p$ et $A \in \mathcal{L}(\mathbb{U}, \mathbb{R}^p)$. Pour $u \in U^{\mathrm{ad}}$ on définit l'ensemble $\Lambda(u) \subset \mathbb{R}^m_+$ de la façon suivante

$$\Lambda(u) \stackrel{\mathrm{def}}{=} \{\lambda \in \mathbb{R}^m \mid \lambda_i \geq 0, \;\; \lambda_i \Theta_i(u) = 0, \;\; i = 1, \dots, m\} \; . \qquad (35)$$

On définit le cône

$$C(u) \stackrel{\mathrm{def}}{=} \sum_{i=1}^{m} \lambda_i \partial\Theta_i(u) + A^* \mu \; , \;\; \lambda \in \Lambda(u) \; , \;\; \mu \in \mathbb{R}^p \; . \qquad (36)$$

Alors, $C(u) \subset \left(U^{\mathrm{ad}}\right)^{\perp}_{u}$.

Preuve

Soit $r \in C(u)$: $r = \sum_{i=1}^{m} \lambda_i \theta_i + A^* \mu$ avec $\theta_i \in \partial\Theta_i(u)$, $\mu \in \mathbb{R}^p$ et $\lambda \in \Lambda(u)$ Alors,

$$\langle r \,, v - u \rangle = \sum_{i=1}^{m} \lambda_i \langle \theta_i \,, v - u \rangle + \langle \mu \,, A(v-u) \rangle$$

$$\leq \sum_{i=1}^{m} \lambda_i \big(\Theta_i(v) - \Theta_i(u) \big) + \langle \mu \,, A(v) - A(u) \rangle$$

puisque $\lambda_i \geq 0$ et $\theta_i \in \partial\Theta_i(u)$,

$$= \sum_{i=1}^{m} \lambda_i \Theta_i(v)$$

du fait que $\lambda \in \Lambda(u)$, $A(u) = A(v) = b$ (u et v sont dans U^{ad})

$$\leq 0$$

puisqu'à nouveau $v \in U^{\mathrm{ad}}$, donc $\Theta_i(v) \leq 0$ et $\lambda_i \geq 0$ pour tout i.

Les écarts complémentaires

- Soit $u \in U^{\mathrm{ad}}$, la condition sur les λ_i dans la définition de $\Lambda(u)$ peut-être réécrite sous les formes équivalentes suivantes :

$$\lambda_i \geq 0 \ , \ \ \Theta_i(u) \leq 0 \ , \ \ \sum_{i=1}^{m} \lambda_i \Theta_i(u) = 0 \ ,$$

$$\Theta_i(u) \leq 0 \ , \ \ I(u) \stackrel{\mathrm{def}}{=} \{ i \in [1, \dots, m] \mid \Theta_i(u) = 0 \} \\ , \ \ \lambda_i \geq 0 \ , \ \ \lambda_i = 0 \text{ pour } i \notin I(u) \ ,$$

- Elle est connue sous le nom de condition d'écarts complémentaires.

Les écarts complémentaires

- L'ensemble $I(u)$ est l'ensemble des indices des contraintes dites saturées ou actives au point u.
- La condition d'écarts complémentaires indique que si $\Theta_i(u) < 0$, alors nécessairement $\lambda_i = 0$, et aussi, si $\lambda_i > 0$, alors nécessairement $\Theta_i(u) = 0$.
- Seules les contraintes saturées en u sont à prendre en compte pour construire $C(u)$
- Ceci est naturel : le bord de U^{ad} est localement défini en u par les contraintes saturées (et les contraintes égalité qui sont toujours saturées par définition).

L'égalité des deux cônes n'a pas toujours lieu

Considérons dans $\mathbb{U} = \mathbb{R}^2$ les deux contraintes suivantes :

$$x^2 + y^2 \leq 1 \text{ et } -y \leq -1 \,.$$

- U^{ad} se réduit au point $u = (0, 1)$ et donc $\left(U^{\mathrm{ad}}\right)_u^{\perp} = \mathbb{R}^2$
- Le cône $C(u)$ est constitué par la droite $x = 0$.
- $C(u) \neq \left(U^{\mathrm{ad}}\right)_u^{\perp}$
- Du fait que U^{ad} est réduit à un singleton, n'importe quelle fonction coût J conduira à la solution $u^\sharp = (0, 1)$
- $J'(u^\sharp)$ pourra prendre une valeur quelconque, puisque le cône normal à U^{ad} est tout $\mathbb{R}^2$,
- On ne pourra pas toujours affirmer que $J'(u^\sharp) \in -C(u^\sharp)$.

Qualification des contraintes et conditions de Lagrange-Karush-Kuhn-Tucker

Faire l'hypothèse

$$C(u^\sharp) = \left(U^{\mathrm{ad}}\right)^{\perp}_{u^\sharp} , \tag{39}$$

est appelé "hypothèse de qualification des contraintes dans sa forme minimale".

Lemma

Si les fonctions qui définissent U^{ad} sont affines continues et si $U^{\mathrm{ad}} \neq \emptyset$, alors (39) est satisfaite en tout point $u \in U^{\mathrm{ad}}$.

Qualification des contraintes et conditions de Lagrange-Karush-Kuhn-Tucker

Corollary

Si les fonctions qui définissent U^{ad} *sont affines continues et si* $U^{\mathrm{ad}} \neq \emptyset$, *alors une solution* $u^\sharp$ *de* $\mathcal{P}$ *est caractérisée par l'existence de "multiplicateurs"* $\lambda_i^\sharp, i = 1, \ldots, m$, *vérifiant la conditions des écarts complémentaires et de "multiplicateurs"* $\mu_j^\sharp, j = 1, \ldots, p$, *tels que*

$$0 \in \partial J(u^\sharp) + \sum_{i=1}^{m} \lambda_i^\sharp \theta_i + \sum_{j=1}^{p} \mu_j^\sharp a_j \ .$$

Une condition suffisante de qualification

- Condition de Slater :

$$\exists u^\circ : \Theta_i(u^\circ) < 0\ ,\ \ i = 1, \dots, m,\ \ A(u^\circ) = b\ .$$

- Les contraintes égalité s'écrivent globalement $A(u) = b$.
- Cette hypothèse dit qu'il existe un point u° satisfaisant strictement les contraintes (le mot “strictement” ne s'applique bien sûr qu'aux contraintes inégalité).

Theorem

- J : *convexe s.c.i. sous-différentiable ;*
- Θ_i : *convexes, continues et sous-différentiables ;*
- Ω_j : *affines et continues* ($A(u) = b$) *;*
- *les conditions de qualification sont satisfaites.*

Une condition nécessaire et suffisante pour que $u^\sharp$ soit une solution du problème $\mathcal{P}$ est qu'il existe des multiplicateurs $\lambda_i^\sharp, i = 1, \ldots, m$, satisfaisant la condition des écarts complémentaires (pour $u = u^\sharp$) et des multiplicateurs $\mu_j^\sharp, j = 1, \ldots, p$, ($\mu^\sharp$ désigne le vecteur de $\mathbb{R}^p$ de coordonnées $\mu_j^\sharp$) tels que

$$0 \in \partial J(u^\sharp) + \sum_{i=1}^{m} \lambda_i^\sharp \partial \Theta_i(u^\sharp) + A^* \mu^\sharp \ .$$

Le cas différentiables

Corollary

Sous les mêmes hypothèses, en supposant de plus que les fonctions J et Θ_i, $i = 1, \ldots, m$, sont différentiables. Alors, une condition nécessaire et suffisante pour que $u^\sharp$ soit une solution du problème $\mathcal{P}$ est qu'il existe des multiplicateurs $\lambda_i^\sharp, i = 1, \ldots, m$, stisfaisant la condition des écarts complémentaires (pour $u = u^\sharp$) et des multiplicateurs $\mu_j^\sharp$, $j = 1, \ldots, p$, tels que

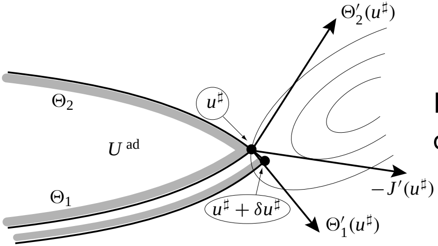
$$J'(u^\sharp) + \sum_{i=1}^{m} \lambda_i^\sharp \Theta_i'(u^\sharp) + \sum_{j=1}^{p} \mu_j^\sharp \Omega_j'(u^\sharp) = 0 \,. \tag{40}$$

On trouve parfois l'hypothèse supplémentaire concernant les contraintes égalité :

"A est surjective"

- Si A n'est pas surjective :
  - ou bien $b \notin \operatorname{im} A$ et alors $U^{\mathrm{ad}} = \varnothing$
  - ou bien $b \in \operatorname{im} A$ et les contraintes égalité sont redondantes.
- Elle évite quand $b \in \operatorname{im} A$, la non unicité et le caractère non borné des multiplicateurs.
- En effet, comme le vecteur $\mu^\sharp$ des multiplicateurs intervient sous la forme $A^*\mu^\sharp$, si A^* n'est pas injective, on peut ajouter à un $\mu^\sharp$ qui convient, n'importe quel élément du noyau de A^*.

Interprétation marginaliste des multiplicateurs



Perturbation d'une contrainte dans $\mathbb{U} = \mathbb{R}^2$

- Contraintes inégalité Θ_1 et Θ_2, toutes les deux actives à la solution $u^\sharp$,
- Les ensembles de niveau de la fonction coût J
- Perturbation de $\Theta_1 \leq 0$ qui devient $\Theta_1(u) \leq \varepsilon$
- Élargissement de U^{ad} et diminution de $J(u^\sharp)$
- De combien s'améliore le coût optimal ?

Interprétation marginaliste des multiplicateurs

- Θ_2 était active avant la perturbation. La solution $u^\sharp + \delta u^\sharp$ vérifie $\left\langle \Theta_2'(u^\sharp) \, , \delta u^\sharp \right\rangle = 0$ ($\Theta_2'(u^\sharp)$ est orthogonal à la frontière $\Theta_2 = 0$.
- Par ailleurs, $\left\langle \Theta_1'(u^\sharp) \, , \delta u^\sharp \right\rangle = \varepsilon$
- La condition d'optimalité est ici

  $$J'(u^\sharp) + \lambda_1^\sharp \Theta_1'(u^\sharp) + \lambda_2^\sharp \Theta_2'(u^\sharp) = 0 \; .$$

  En faisant le produit scalaire avec $\delta u^\sharp$ on obtient que :

  $$\begin{aligned} 0 &= \left\langle J'(u^\sharp) + \lambda_1^\sharp \Theta_1'(u^\sharp) + \lambda_2^\sharp \Theta_2'(u^\sharp) \, , \delta u^\sharp \right\rangle \\ &= \left\langle J'(u^\sharp) \, , \delta u^\sharp \right\rangle + \lambda_1^\sharp \varepsilon \; . \end{aligned}$$

- L'interprétation marginaliste des multiplicateurs : c'est au signe près la sensibilité du coût optimal par rapport à une variation "marginale" du second membre de la contrainte correspondante.

Formulation compacte des contraintes

$$(\mathcal{P}) \qquad \min_{u \in U^{\mathrm{ad}} \cap U^{\mathrm{ad}}_{\Theta}} J(u)$$

- $\mathbb{C}$: "espace des contraintes" qui est un espace de Hilbert,
- $\Theta : \mathbb{U} \to \mathbb{C}$: une "fonction contrainte",
- C : cône convexe et saillant C
- Les contraintes sont maintenant

$$\Theta(u) \in -C \text{ ce qui signifie } \Theta(u) \preceq 0 \tag{41}$$

- $U^{\mathrm{ad}}_{\Theta} = \{u \in \mathbb{U} \mid \Theta(u) \in -C\}$.
- Contraintes égalité : $C = \{0\}$
- $\mathbb{C} = \mathbb{C}_1 \times \mathbb{C}_2$ et $C = C_1 \times C_2$ où $C_2 = \{0\}$
- $u \in U^{\mathrm{ad}}$: contraintes implicites où U^{ad} est convexe fermé.

Lagrangien

On considère $\lambda \in \mathbb{C}^*$ et on pose

$$L(u, \lambda) = J(u) + \langle \lambda \,, \Theta(u) \rangle \ . \tag{42}$$

- Cette expression est appelée Lagrangien associé au problème d'optimisation $\mathcal{P}$
- la variable λ est appelée multiplicateur ou variable duale.
- Par opposition, u est la variable primale.
- u est restreinte au sous-ensemble admissible U^{ad},
- λ va être restreinte au sous-ensemble admissible $C^\star$

Lemma

Soit

$$P(u) \overset{\mathrm{def}}{=} \sup_{\lambda \in C^\star} L(u, \lambda) \; .$$

Alors $P(u) = J(u) + \mathrm{I}_{U^{\mathrm{ad}}_{\Theta}}(u)$ *et le problème*

$$\min_{u \in U^{\mathrm{ad}}} P(u)$$

est équivalent au problème $\mathcal{P}$ *(même valeurs optimales et mêmes solutions).*

Preuve

- Si $u \in U^{\mathrm{ad}}_{\Theta}$, alors $\langle \lambda \,, \Theta(u) \rangle \leq 0$ pour tout $\lambda \in C^{\star}$
- la valeur 0 est atteinte pour $\lambda = 0$

$$u \in U^{\mathrm{ad}}_{\Theta} \Rightarrow \sup_{\lambda \in C^{\star}} \langle \lambda \,, \Theta(u) \rangle = 0 \; .$$

- Inversement si $\langle \lambda \,, \Theta(u) \rangle \leq 0$ pour tout $\lambda \in C^{\star}$, alors $\Theta(u) \in -C^{\star\star} = -C$
- Donc si $u \notin U^{\mathrm{ad}}_{\Theta}$ il existe $\lambda \in C^{\star}$ t.q. $\langle \lambda \,, \Theta(u) \rangle > 0$.
- Puisque $C^{\star}$ est un cône :

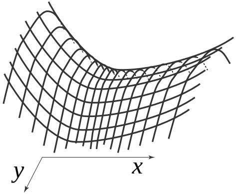
$$u \notin U^{\mathrm{ad}}_{\Theta} \Rightarrow \sup_{\lambda \in C^{\star}} \langle \lambda \,, \Theta(u) \rangle = +\infty \; .$$

- De tout ceci, on déduit que $P = J + \mathrm{I}_{U^{\mathrm{ad}}_{\Theta}}$.

Definition

Pour une fonction $f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ et des sous-ensembles convexes fermés $X \subset \mathbb{X}$ et $Y \subset \mathbb{Y}$, on dit que $(x^\sharp, y^\sharp) \in X \times Y$ est un point selle de f sur $X \times Y$ si

$$\forall x \in X, \;\; \forall y \in Y, \;\; f(x^\sharp, y) \leq f(x^\sharp, y^\sharp) \leq f(x, y^\sharp) \; .$$



Point selle

Point selle

- On cherche à maximiser f en y et à minimiser en x.
- Interprétation "théorie des jeux" ici jeu à "somme nulle". Les joueurs cherchent à minimiser leur propre fonction coût la fonction coût du joueur x est $f(x,y)$ alors que celle du joueur y est $-f(x,y)$.
- Un autre formulation équivalente :

$$f(x^\sharp, y) \leq f(x, y^\sharp) \;, \;\; \forall x \in X, \;\; \forall y \in Y \;. \tag{43}$$

- Un point selle peut-être vu comme le point fixe d'une multi-application.

$$X^\sharp(y) = \arg\min_{x \in X} f(x,y) \;; \quad Y^\sharp(x) = \arg\max_{y \in Y} f(x,y) \;,$$

et la multi-application $F^\sharp : \mathbb{X} \times \mathbb{Y} \to 2^{\mathbb{X} \times \mathbb{Y}}$ définie par $F^\sharp(x,y) = X^\sharp(y) \times Y^\sharp(x)$.

Point selle

Lemma

Le point $(x^\sharp, y^\sharp)$ est un point selle de f sur $X \times Y$ si et seulement si

$$
\begin{aligned}
f(x^\sharp, y^\sharp) = \sup_{y \in Y} f(x^\sharp, y) &= \min_{x \in X} \sup_{y \in Y} f(x, y) \\
&= \max_{y \in Y} \inf_{x \in X} f(x, y) = \inf_{x \in X} f(x, y^\sharp) ,
\end{aligned}
$$

autrement dit $\inf_{x \in X} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} \inf_{x \in X} f(x, y)$ *et les* $\inf$ *et* $\sup$ *les plus extérieurs sont atteints en* $x^\sharp$, *respectivement* $y^\sharp$.

On note tout d'abord qu'on a tojours :

$$
\sup_y \inf_x f(x, y) \leq \inf_x \sup_y f(x, y) \tag{44}
$$

Preuve

Soit $(x^\sharp, y^\sharp)$ un point selle, on a avec l'inégalité de droite :

$f(x^\sharp, y^\sharp) = \min_{x \in X} f(x, y^\sharp)$ et comme $\inf_{x \in X} f(x, y^\sharp) \leq \sup_{y \in Y} \inf_{x \in X} f(x, y)$,

on a

$$f(x^\sharp, y^\sharp) \leq \sup_{y \in Y} \inf_{x \in X} f(x, y) \,.$$

On conlut en raisonnant de façon symétrique avec l'inégalité de gauche et en utilisant (44).
Réciproquement, supposons vraie l'équation du Th.. Alors

$$\forall y \in Y, \;\; f(x^\sharp, y) \leq \sup_{y \in Y} f(x^\sharp, y) = f(x^\sharp, y^\sharp) \,.$$

et de façon symétrique

$$\forall x \in X, \;\; f(x, y^\sharp) \geq \inf_{x \in X} f(x, y^\sharp) = f(x^\sharp, y^\sharp) \,,$$

Cas convexe-concave

$f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ est convexe-concave si $f(\cdot, y)$ est convexe pour tout y et $f(x, \cdot)$ est concave pour tout x.
Les inégalités de point selle sont équivalentes à

$$\exists r^\sharp \in \partial_x f(x^\sharp, y^\sharp) : \ \forall x \in X \ , \ \left\langle r^\sharp \ , x - x^\sharp \right\rangle \geq 0 \ ,$$

$$\exists s^\sharp \in \partial_y f(x^\sharp, y^\sharp) : \ \forall y \in Y \ , \ \left\langle s^\sharp \ , y - y^\sharp \right\rangle \leq 0 \ .$$

En changeant $s^\sharp$ en $t^\sharp = -s^\sharp$:

$$\exists t^\sharp \in -\partial_y f(x^\sharp, y^\sharp) : \ \forall y \in Y \ , \ \left\langle t^\sharp \ , y - y^\sharp \right\rangle \geq 0 \ .$$

Posons $z = (x, y) \in \mathbb{W} \overset{\mathrm{def}}{=} \mathbb{X} \times \mathbb{Y}$, $Z = X \times Y \subset \mathbb{W}$, et définissons l'opérateur (éventuellement multivoque)

$$A : \ \ z = (x, y) \in \mathbb{W} \mapsto \partial_x f(x, y) \times \big(- \partial_y f(x, y) \big) \subset \mathbb{X} \times \mathbb{Y} \ .$$

$$\exists a^\sharp \in A(z^\sharp) : \ \forall z \in Z \ , \ \left\langle a^\sharp \ , z - z^\sharp \right\rangle \geq 0 \ .$$

Cas convexe-concave

- Pour une fonction convexe-concave et sous-différentiable, l'opérateur A est monotone.
- Le problème de point selle se raméne à celui d'une inéquation variationnelle avec un opérateur A monotone.
- A "ne dérive pas d'un potentiel". Il n'est pas en général le sous-différentiel d'une fonction (qui serait nécessairement convexe puisque A est monotone) sauf dans le cas particulier où f serait de la forme $f(x, y) = g(x) + h(y)$ (alors A est le sous-différentiel de $(x, y) \mapsto g(x) - h(y)$).

Condition suffisante d'optimalité

Theorem

Si $(u^\sharp, \lambda^\sharp)$ *est un point selle du Lagrangien sur* $U^{\mathrm{ad}} \times C^\star$*, alors* $u^\sharp$ *est solution du problème* $\mathcal{P}$.

Hypothèses de convexité

- La fonction coût J est convexe, s.c.i.et sous-différentiable.
- En ce qui concerne $\Theta : \mathbb{U} \to \mathbb{C}$,

Definition

On dit que l'application $\Theta : \mathbb{U} \to \mathbb{C}$ est *C-convexe* si, pour tout u et v dans $\mathbb{U}$ et tout α dans $[0, 1]$,

$$\Theta\big(\alpha u + (1-\alpha)v\big) - \alpha\Theta(u) - (1-\alpha)\Theta(v) \in -C \ .$$

- U_{Θ}^{ad} est convexe et fermé si Θ est continue.

Definition

On appellera "C-sous-gradient" de Θ au point u tout élément $\theta \in \mathcal{L}(\mathbb{U}, \mathbb{C})$ t.q.

$$\forall v, \ \ \Theta(v) - \Theta(u) - \theta\big(v - u\big) \in C \ ,$$

Hypothèses de convexité

- On notera $\partial\Theta(u)$ l'ensemble de ces applications linéaires que l'on appellera "sous-différentiel" de Θ en u.
- Lorsque $C = \{0\}$ la C-convexité signifie que Θ est affine
- Lorsque $\lambda \in C^\star$, la fonction $u \mapsto \langle \lambda \,, \Theta(u) \rangle$ est convexe et

$$\partial_u \langle \lambda \,, \Theta(u) \rangle \supset \{ r \in \mathbb{C} \mid r = \theta^*(\lambda), \;\; \theta \in \partial\Theta(u) \} \;.$$

- Lorsque l'égalité a lieu, on dit que Θ est "régulièrement sous-différentiable"

Point selle du Lagrangien et multiplicateurs de Kuhn et Tucker

Lemma

L'inégalité de gauche du point selle est équivalente au fait que

$$\lambda^\sharp \in C^\star \; , \; \; \Theta(u^\sharp) \in -C \; , \; \; \left\langle \lambda^\sharp \, , \Theta(u^\sharp) \right\rangle = 0 \; .$$

Sous l'hypothèse que Θ *est continue et régulièrement sous-différentiable, l'inégalité de droite du point selle est équivalente à*

$$u^\sharp \in U^{\mathrm{ad}} \; , \; \; \exists r^\sharp \in \partial J(u^\sharp) \; , \; \; \exists \theta^\sharp \in \partial \Theta(u^\sharp) : \; \; \forall u \in U^{\mathrm{ad}} \; , \quad \left\langle r^\sharp + \left(\theta^\sharp\right)^* (\lambda^\sharp) \, , u - u^\sharp \right\rangle \geq 0 \; .$$

Corollary

On suppose que le problème convexe $\mathcal{P}$ admet une solution $u^\sharp$ et un multiplicateur $\lambda^\sharp$ associé à la contrainte $\Theta(u^\sharp) \in -C$ c'est-à-dire vérifiant avec $u^\sharp$ les conditions de Kuhn-Tucker. Si la solution primale $u^\sharp$ n'est pas unique, alors l'ensemble des multiplicateurs $\lambda^\sharp$ (eux-mêmes possiblement non uniques) ne dépend pas de la solution $u^\sharp$ particulière choisie.

Application d'un théorème général d'existence de point selle

- Sans aucune hypothèse de convexité, l'existence d'un point selle du Lagrangien constitue une condition suffisante d'optimalité du problème d'optimisation sous contraintes en ce qui concerne la partie primale de ce point selle.
- L'existence d'une solution primale et de multiplicateurs de Lagrange-Kuhn-Tucker pour un problème d'optimisation sous contraintes a été mise en correspondance directe, dans le cas convexe uniquement, avec l'existence d'un point selle du Lagrangien.
- Dans le cas convexe, lorsqu'une solution primale existe, et lorsque l'existence de multiplicateurs peut être assurée, l'existence d'un point selle du Lagrangien devient aussi une condition nécessaire d'optimalité.

Cependant, l'existence de multiplicateurs associés à une solution primale n'a été établie que pour le cas d'un nombre fini de contraintes égalité (affines) et inégalité (convexes).

Theorem

Soit $f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ *une fonction convexe-concave s.c.i.-s.c.s.. Soit* X *et* Y *deux sous-ensembles convexes fermés de* $\mathbb{X}$ *et* $\mathbb{Y}$*, respectivement. On suppose aussi qu'il existe* $x^\circ \in X$ *tel que*

$$\lim f(x^\circ, y) = -\infty \;\; \textit{lorsque} \;\; \|y\| \to +\infty \;\; \textit{et} \;\; y \in Y \; , \tag{46a}$$

et qu'il existe $y^\circ \in Y$ *tel que*

$$\lim f(x, y^\circ) = +\infty \;\; \textit{lorsque} \;\; \|x\| \to +\infty \;\; \textit{et} \;\; x \in X \; , \tag{46b}$$

ou bien Y*, respectivement* X*, est borné. Alors, il existe un point selle de* f *sur* $X \times Y$*.*

- On considère ici le Lagrangien associé au problème $\mathcal{P}$ (avec $X = U^{\mathrm{ad}}$ et $Y = C^{\star}$).
- C'est une fonction affine, donc concave de λ.
- Avec les hypothèses de convexité faites sur J et Θ, C'est une fonction convexe de u pour tout $\lambda \in C^{\star}$.
- Pour l'application du Théorème, on complètera donc la définition de L de la façon suivante :

$$L(u, \lambda) \stackrel{\mathrm{def}}{=} \begin{cases} J(u) + \langle \lambda \,, \Theta(u) \rangle & \text{si } \lambda \in C^{\star} \,, \\ -\infty & \text{sinon.} \end{cases} \tag{47}$$

- Pour la continuité, les hypothèses que J est s.c.i. et que Θ est continue sont suffisantes pour assurer le caractère s.c.i. par rapport à u.
- Par rapport à λ, on a affaire à une fonction affine continue sur le convexe fermé $C^{\star}$ et le passage à $-\infty$ lorsque λ sort de $C^{\star}$ assure le caractère s.c.s..

Théorème d'existence et Lagrangien

- Si U^{ad} est bornée, l'hypothèse de coercivité est assurée.
- Sinon, on fait l'hypothèse que J est coercive sur U^{ad}. Cette hypothèse est plus forte que coercive sur $U^{\mathrm{ad}} \cap U^{\mathrm{ad}}_{\Theta}$. C'est le prix que l'on paie pour la technique Lagrangienne.
- Comme $C^{\star}$ n'est pas borné, on choisit $u^{\circ} \in \mathrm{dom}\left(J\right)$ et on suppose de plus que u° vérifie l'hypothèse de Slater :

$$\exists u^{\circ} \in U^{\mathrm{ad}} \cap \mathrm{dom}\left(J\right) : \; \Theta(u^{\circ}) \in -\overset{\circ}{C}$$

- Alors $L(u^{\circ}, \lambda)$ tend vers $-\infty$ lorsque λ part à l'infini en restant dans $C^{\star}$.
- Cependant, cette hypothèse n'a de sens que lorsque $\overset{\circ}{C}$ est non vide, ce qui exclut les contraintes égalité

La fonction "perturbation"

On introduit le sous-ensemble suivant :

$$S \stackrel{\text{def}}{=} \{(\theta, y) \in \mathbb{C} \times \mathbb{R} | \; \exists u \in U^{\text{ad}}, \exists c \in C, \exists z \in \mathbb{R}_+ : \\ \theta = \Theta(u) + c, y = J(u) + z\} \; .$$

Ce sous-ensemble est l'épigraphe d'une fonction appelée "fonction perturbation".

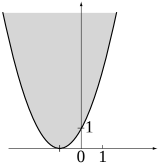
$$\varphi(\theta) \stackrel{\text{def}}{=} \min \left\{ J(u) \;\middle|\; u \in U^{\text{ad}}, \Theta(u) \in \theta - C \right\} \; .$$

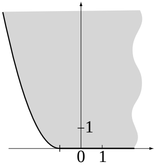
La résolution de $\mathcal{P}$revient à calculer $\varphi(0)$.

Considérons le problème "perturbé"

$$\min_{u\in\mathbb{R}} (u+1)^2 \ \text{ sous } \ u=\theta \ \text{ ou } \ u \leq \theta \ .$$

On observe que la fonction φ est convexe et qu'elle est non décroissante dans le cas inégalité.





La fonction perturbation φ et son épigraphe S pour des contraintes égalité ou inégalité

La fonction "perturbation"

Lemma

Sous les hypothèses de convexité et de semi-continuité inférieure, respectivement de continuité, formulées pour J et Θ (ainsi évidemment que le caractère convexe fermé de U^{ad} et C), le sous-ensemble S est convexe fermé et donc la fonction φ est convexe s.c.i.. Elle est de plus C-non croissante, ce qui signifie que

$$\theta' - \theta \in C \Rightarrow \varphi(\theta') \leq \varphi(\theta) \ , \tag{48}$$

et en conséquence, tout $r \in \partial\varphi(\theta)$ appartient à $-C^\star$.

Expression alternative

On peut donner l'expression alternative suivante de cette fonction :

$$\varphi(\theta) = \inf_{u \in U^{\mathrm{ad}}} \sup_{\lambda \in C^\star} \big(J(u) + \langle \lambda \,, \Theta(u) - \theta \rangle \big) \,. \tag{49}$$

Multiplicateurs et sous-différentiel de la fonction perturbation

Lemma

- *Soit* $(u^\sharp, \lambda^\sharp)$ *un point selle du Lagrangien associé à* P *sur* $U^{\mathrm{ad}} \times C^\star$.
- *Alors,* $-\lambda^\sharp$ *est un sous-gradient de la fonction* φ *au point* 0.
- *Réciproquement, si* $\partial\varphi(0)$ *est non vide (et cela suppose bien sûr que* $0 \in \mathrm{dom}\left(\varphi\right)$, *donc qu'il existe une solution* $u^\sharp$ *au problème* P *que* $\varphi(0) = J(u^\sharp)$ *et que* $\Theta(u^\sharp) \in -C$*), tout élément dans* $-\partial\varphi(0)$ *constitue avec* $u^\sharp$ *un point selle du Lagrangien associé.*

Existence d'un sous-gradient de φ en 0

Étudier le cône normal à $S = \operatorname{epi}(\varphi)$ au point $\big(0, J(u^\sharp)\big)$.

Theorem

On suppose qu'il existe une solution $u^\sharp$ au problème (**??**) [a] *et on fait l'hypothèse de qualification des contraintes suivante :*

$$0 \in \overbrace{\Theta(U^{\mathrm{ad}}) + C}^{\circ} .$$

Alors il existe un multiplicateur optimal $\lambda^\sharp$ et donc un point selle au Lagrangien.

a. Cela nécessite donc que $\operatorname{dom}(J) \cap U^{\mathrm{ad}} \cap U^{\mathrm{ad}}_{\Theta} \neq \emptyset$.

Dans le cas de contraintes égalité (c'est-à-dire lorsque $C = \{0\}$), la condition prend la forme

$$0 \in \overbrace{\Theta(U^{\mathrm{ad}})}^{\circ} \ . \tag{50}$$

D'une manière générale, cette condition de qualification exprime que la contrainte peut être satisfaite avec un $u \in U^{\mathrm{ad}}$ non seulement sous sa forme originale ($\theta = 0$) mais pour toute perturbation $\theta \in \mathbb{C}$ suffisamment petite et dans toutes les directions de l'espace $\mathbb{C}$.

Fonction perturbation, T. de Fenchel et Lagrangien

- On considère la fonction duale :

$$\psi(\lambda) \stackrel{\text{def}}{=} \begin{cases} \inf_{u \in U^{\mathrm{ad}}} L(u, \lambda) & \text{si } \lambda \in C^\star, \\ -\infty & \text{sinon.} \end{cases}$$

- $\psi(\lambda)$ est une fonction concave et s.c.s.
- Lorsque le Lagrangien admet un point selle $(u^\sharp, \lambda^\sharp)$ sur $U^{\mathrm{ad}} \times C^\star$, le coût optimal du problème $\mathcal{P}$ est aussi égal au $\max_{\lambda \in C^\star} \inf_{u \in U^{\mathrm{ad}}} L(u, \lambda)$ qui n'est autre que

$$\max_{\lambda \in \mathbb{C}} \psi(\lambda) .$$

- Ce maximum est atteint évidemment pour tout $\lambda^\sharp$

Theorem

On a la formule : $\forall \lambda \in \mathbb{C}, \;\; \mathcal{F}(\varphi)(\lambda) = -\psi(-\lambda)$.

$$
\begin{aligned}
\mathcal{F}(\varphi)(\lambda) &= \sup_{\theta \in \mathbb{C}} \big(\langle \lambda \,, \theta \rangle - \varphi(\theta) \big) \\
&= \sup_{\theta \in \mathbb{C}} \Big(\langle \lambda \,, \theta \rangle - \min \Big\{ J(u) \;\Big|\; u \in U^{\mathrm{ad}}, \Theta(u) \in \theta - C \Big\} \Big) \\
&= \max_{u \in U^{\mathrm{ad}}} \sup_{\theta \in C + \Theta(u)} \big(\langle \lambda \,, \theta \rangle - J(u) \big) \\
&= \begin{cases} +\infty & \text{si } \lambda \notin -C^\star, \\ \sup_{u \in U^{\mathrm{ad}}} \big(\langle \lambda \,, \Theta(u) \rangle - J(u) \big) & \text{si } \lambda \in -C^\star, \end{cases}
\end{aligned}
$$

Si $\lambda \notin -C^\star$, $\theta \to \infty$ dans la “bonne” direction fait croître arbitrairement l'expression. Si $\lambda \in -C^\star$ choisir $\lambda = \Theta(u)$.

- Considérons dans $\mathbb{C} \times \mathbb{R}$, $y + \langle \lambda \,, \theta \rangle = a$, où les variables sont $\theta \in \mathbb{C}$ et $y \in \mathbb{R}$, où $\lambda \in \mathbb{C}$ (ou $\mathbb{C}^*$) est un paramètre, et où $a \in \mathbb{R}$ est un autre paramètre.
- Décrit un hyperplan, $-\lambda$ est sa “pente” tandis que a est l'“ordonnée à l'origine”.
- L'espace $\mathbb{C} \times \mathbb{R}$ est celui dans lequel a été défini S (voir (**??**)) et on peut s'intéresser à l'intersection de l'hyperplan avec ce sous-ensemble.
- Cela revient à ne considérer dans l'équation de l'hyperplan que les points de la forme $\theta = \Theta(u) + c$ et $y = J(u) + z$ pour $u \in U^{\mathrm{ad}}$, $c \in C$ et $z \geq 0$.
- On va s'intéresser en fait à “abaisser” le plus possible l'hyperplan en le déplaçant parallèlement à lui-même (donc en gardant λ fixe mais en cherchant l'ordonnée à l'origine a la plus petite possible) sous la contrainte de conserver une intersection non vide avec S.

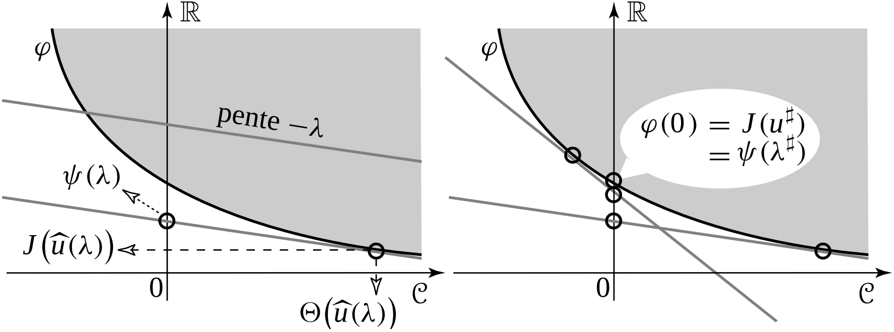


FIGURE – La dualité par l'image

- Lorsque le bord inférieur de S (le graphe de φ) a l'allure décroissante représentée sur la figure (lorsque $C = \mathbb{R}_+$ dans le cas où $\mathbb{C} = \mathbb{R}$ dessiné ici),
- On voit que la plus petite ordonnée à l'origine possible reste finie (n'est pas égale à $-\infty$) si la "pente" n'est pas positive, donc si $\lambda \in \mathbb{R}_+$ (plus généralement si $\lambda \in C^\star$).
- Si de plus S a une forme convexe, la plus petite ordonnée à l'origine sera obtenue dans la position limite où la droite (plus généralement l'hyperplan) "tangente" l'ensemble S sur son bord inférieur.
- Le point de contact étant de la forme $(\Theta(u), J(u))$ (c'est-à-dire qu'on a intérêt à prendre $z = 0$ et $c = 0$ pour minimiser a en restant dans S à partir du moment où λ a le bon signe ou la bonne orientation).
- Autrement dit, le plus petit a associé à un λ donné est obtenu en minimisant $y + \langle \lambda \,, \theta \rangle = J(u) + \langle \lambda \,, \Theta(u) \rangle$ pour $u \in U^{\mathrm{ad}}$.

Explication géométrique des résultats

- Finalement, calculer la plus petite ordonnée à l'origine correspondant à un λ fixé, c'est minimiser $L(u, \lambda)$ pour $u \in U^{\mathrm{ad}}$ et c'est donc calculer $\psi(\lambda)$. Cette constatation est d'ailleurs conforme avec le fait démontré précédemment que ψ a un rapport avec $\mathcal{F}(\varphi)$.
- On va chercher maintenant à obtenir l'ordonnée à l'origine $\psi(\lambda)$ de l'hyperplan d'appui de $S = \mathrm{epi}\,(\varphi)$ la plus élevée possible en agissant sur la "pente" $-\lambda$.
- On peut voir sur la partie droite de la Figure **??** que la plus haute ordonnée à l'origine que l'on peut obtenir par ce procédé est la valeur $\varphi(0)$ (qui est égale, comme on le sait à $J(u^\sharp)$).
- Cela se démontre facilement : puisque $y = \langle -\lambda\,, \theta\rangle + \psi(\lambda)$ définit un hyperplan d'appui de $\mathrm{epi}\,(S)$, le graphe de φ se trouve au dessus de tout hyperplan d'appui et donc

$$\forall \theta, \forall \lambda, \;\; \varphi(\theta) \geq \langle -\lambda\,, \theta\rangle + \psi(\lambda) \Rightarrow \forall \lambda, \;\; \varphi(0) \geq \psi(\lambda)\,.$$

Explication géométrique des résultats

- Pour obtenir la valeur maximale de l'ordonnée à l'origine $\psi(\lambda)$, il faut donc choisir $\lambda = \lambda^\sharp$ telle que $\psi(\lambda^\sharp) = \varphi(0) = J(u^\sharp)$. Le point de contact de cet hyperplan avec S a comme coordonnées $(0, J(u^\sharp))$, c'est-à-dire que $\left\langle \lambda^\sharp , \Theta(u^\sharp) \right\rangle = 0$.
- Enfin, la récapitulation de toutes ces opérations montre que l'on a réalisé $\max_{\lambda \in C^\star} \inf_{u \in U^{\mathrm{ad}}} L(u, \lambda)$ et que l'on ne parvient à atteindre $\varphi(0) = J(u^\sharp) = \min_{u \in U^{\mathrm{ad}}} \sup_{\lambda \in C^\star} L(u, \lambda)$ que dans le cas où ces deux valeurs sont égales, ce qui se réalise dans le cas convexe.