ANALYSE CONVEXE

Prof Adama COULIBALY
UFR de Mathématiques et Informatique,
Université Félix HOUPHOUET-BOIGNY,
22 BP 582 Abidjan 22, Côte d'Ivoire.

19 octobre 2016

Table des matières

Chapitre 1

Ensembles convexes

Le cadre général de ce cours est un espace vectoriel réel de dimension n. On peut donc sans perdre de généralités considérer l'espace vectoriel réel $\mathbb{R}^n$.

1.1 Sous espaces affines

Définition 1.1.1 *On appelle combinaison linéaire affine de $x^i : i = 1, \cdots, k$ des points de $\mathbb{R}^n$, toute combinaison linéaire $x = \sum_{i=1}^k \lambda_i x^i$, avec $\lambda_i \in \mathbb{R}\ \forall\, i = 1, \cdots, k$ et $\sum_{i=1}^k \lambda_i = 1$.*

Définition 1.1.2 *Etant donné $M \subset \mathbb{R}^n$, $M \neq \emptyset$, on dit que M est un sous espace affine (ou une variété) de $\mathbb{R}^n$, si M est stable par combinaison linéaire affine. C'est-à-dire :*

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

Proposition 1.1.1 *Pour un sous ensemble non vide de $\mathbb{R}^n$, les conditions suivantes sont équivalentes :*
1) M est un sous espace affine de $\mathbb{R}^n$
2) Il existe un sous espace vectoriel V unique de $\mathbb{R}^n$ tel que :

$$\forall a \in M,\ M = a + V = \{a + v : v \in V\}.$$

(Tout translaté de V par un élément de M est un élément de M.

Preuve : Supposons que M est un sous espace affine de $\mathbb{R}^n$ et soit $x \in M$. Posons $V = M - a = \{x - a : x \in M\}$.
i) $0 \in V$
ii) $\forall\, \alpha \in \mathbb{R},\ \forall\, x - a \in V,\ \alpha(x-a) = [(1-\alpha)a + \alpha x] - a \in V$
iii) $\forall\, x - a,\ y - a \in V,\ (x-a) + (y-a) = 2[(\frac{1}{2}x + \frac{1}{2}y) - a] \in V$
Donc V est un sous espace vectoriel et $M = a + V$.

Il reste à montrer que V ne depend pas de a.

Supposons que $b \in M$, W un sous espace vectoriel tel que $M = b + W$.

$0 \in W = M - b = a - b + V$. On a alors $-(a-b) \in V$ et donc $(a-b) \in V$. Par suite $W = a - b + V \subset V$.

De même on montre que $V \subset W$. D'où $V = W$.

Soit $b \in M$
$b + V = a + (b-a) + V \subset a + V + V \subset a + V$
$a + V = b + (a-b) + V = b + [-(b-a)] + V \subset b + V + V \subset b + V$
Donc $b + V = a + V = M$. □

Définition 1.1.3 *Soit M un sous espace affine de $\mathbb{R}^n$. Le sous espace vectoriel V vérifiant : $\forall\, a \in M,\ M = a + V$ est appelé la direction de M. La dimension de V est la dimension de M.*

Exemple 1.1.1 *a) Soit $a \in \mathbb{R}^n$ et d un élément non nul de $\mathbb{R}^n$. $D = a + \mathbb{R}d$ est la droite affine de $\mathbb{R}^n$ passant par a et de vecteur directeur d.*

b) Soit f un endomorphisme de $\mathbb{R}^n$ et $(a, b) \in \mathbb{R}^{n2}$ tel que $f(a) = b$. $M = \{x \in \mathbb{R}^n : f(x) = b\}$ est le sous espace affine de $\mathbb{R}^n$ passant a et de direction $V = \ker f$.

On montre facilement que

Proposition 1.1.2 *M est un sous espace affine si et seulement si M contient toute combinaison linéaire affine de toute famille finie de ses points.*

Proposition 1.1.3 *Si $\{M_i\}_{i\in I}$ est une famille de sous espaces affines d'intersection non vide, alors $\cap_{i\in I} M_i$ est un sous espace affine.*

Preuve : Prendre $a \in \cap_{i\in I} M_i$. Les éléments $E_i = M_i - a$ sont alors des sous espaces vectoriels et donc $M = \cap_{i\in I} M_i$ aussi qui n'est rien d'autre que $\cap_{i\in I} M_i - a$. □

Etant donné $S \subset \mathbb{R}^n$, on considère l'ensemble des sous espaces affines contenant S. Cet ensemble est non vide car il contient $\mathbb{R}^n$.

Définition 1.1.4 *Etant donné $S \subset \mathbb{R}^n$, on appelle sous espace affine engendré par S, l'intersection de tous les sous espaces affines contenant S. C'est le plus petit sous espace affine contenant S. On le note $\mathrm{aff}(S)$.*

En dimension finie, comme c'est le cas ici, un sous espace affine est toujours fermé ce qui n'est pas toujours le cas en dimension infinie.

Définition 1.1.5 *Etant donné $S \subset \mathbb{R}^n$, on appelle dimension de S la dimension de $\mathrm{aff}(S)$.*

Proposition 1.1.4 *Soit $\emptyset \neq S \subset \mathbb{R}^n$, $a \in S$. On a*

$$\mathrm{aff}(S) = \left\{ x \in \mathbb{R}^n : \exists p \in \mathbb{N}^*,\ \exists(\alpha_i,\, x^i) \in (\mathbb{R} \times S)^p : \begin{array}{l} \sum_{i=1}^p \alpha_i = 1, \\ \sum_{i=1}^p \alpha_i x^i = x \end{array} \right\}$$

ou encore

$$\mathrm{aff}(S) = \left\{ x = a + \sum_{i=1}^p \lambda_i (x_i - a) : \begin{array}{l} p \in \mathbb{N},\ x_i \in S,\ \lambda_i \in \mathbb{R}, \forall i = 1, \cdots, p, \\ \sum_{i=1}^p \lambda_i = 1 \end{array} \right\}.$$

Proposition 1.1.5 $S_1 \subset S_2 \implies \mathrm{aff}(S_1) \subset \mathrm{aff}(S_2)$

Définition 1.1.6 *Un hyperplan est un sous espace affine de codimension 1.*

$$H \text{ est un hyperplan } \Longleftrightarrow H = a + V \text{ et } \mathrm{codim} V = 1.$$

En dimension finie n comme c'est le cas ici,

$$\mathrm{codim} V = 1 \Longleftrightarrow \dim V = n - 1$$

Proposition 1.1.6 $\emptyset \neq H \subset \mathbb{R}^n$
Les conditions suivantes sont équivalentes
i) H est un hyperplan
ii) Il existe $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ linéaire non nulle et $\alpha \in \mathbb{R}$ tels que

$$H = \{x \in \mathbb{R}^n : \varphi(x) = \alpha\}.$$

Preuve : Supposons i) et soit V la direction de H et $a \in H$.
a) Supposons $a \in V$:
Alors $H = a + V = V$. Comme V est de codimension 1, $\exists\, x_0 \in \mathbb{R}^n \setminus V$ tel que $\mathbb{R}^n = V \oplus \mathbb{R}x_0$.
Soit φ définie par :

$$\varphi : \begin{array}{c} \mathbb{R}^n = V \oplus \mathbb{R}x_0 \longrightarrow \mathbb{R} \\ x = v + \lambda x_0 \longmapsto \lambda \end{array}$$

φ est linéaire non nulle et on a

$$\{x \in \mathbb{R}^n : \varphi(x) = 0\} = H.$$

b) Supposons $a \notin V$:
Donc $\mathbb{R}^n = V \oplus \mathbb{R}a$. Considérons l'application

$$\varphi : \begin{array}{c} \mathbb{R}^n = V \oplus \mathbb{R}a \longrightarrow \mathbb{R} \\ x = v + \lambda a \longmapsto \lambda \end{array}$$

L'application φ est linéaire non nulle et

$$\{x \in \mathbb{R}^n : \varphi(x) = 1\} = V + a = H.$$

Réciproquement supposons que

$$H = \{x \in \mathbb{R}^n : \varphi(x) = \alpha\}$$

où φ est une application linéaire non nulle de $\mathbb{R}^n$ dans $\mathbb{R}$.
Soit $a \in H$: alors on a $\forall x \in \mathbb{R}^n$,

$$\begin{array}{rcl} x \in H & \Longleftrightarrow & \varphi(x) = \varphi(a) \\ & \Longleftrightarrow & x - a \in \varphi^{-1}(0) \end{array}$$

Posons $V = \varphi^{-1}(0)$ c'est donc un sous espace vectoriel de codimension 1 et on a $H = a + V$. Alors H est hyperplan. $\square$

Dans $\mathbb{R}^n$ tout hyperplan divise l'espace en deux demi espaces fermés de frontière cet hyperplan.
Par exemple si H est un hyperplan défini par

$$H = \{x \in \mathbb{R}^n : \varphi(x) = \alpha\}$$

où φ est une application linéaire non nulle de $\mathbb{R}^n$ dans $\mathbb{R}$, les demi-espaces fermés sont

$$\{x \in \mathbb{R}^n : \varphi(x) \leq \alpha\} \text{ et } \{x \in \mathbb{R}^n : \varphi(x) \geq \alpha\}.$$

Les ensembles $\{x \in \mathbb{R}^n : \varphi(x) < \alpha\}$ et $\{x \in \mathbb{R}^n : \varphi(x)\ \alpha\}$ sont des demi espaces ouverts de frontière H.

1.2 Ensembles convexes

1.2.1 Définitions

La notion de combinaison linéaire convexe est donnée dans la définition suivante.

Définition 1.2.1 *On appelle combinaison linéaire convexe de $x^i : i = 1, \cdots, k$, k points de $\mathbb{R}^n$, tout élément de $\mathbb{R}^n$ $x = \sum_{i=1}^{k} \lambda_i x^i$ où les coefficients λ_i sont positifs et de somme 1.*

En particulier, une combinaison linéaire convexe de deux points x et y, est tout point $z = (1-\lambda)x + \lambda y$ avec $\lambda \in [0, 1]$

On a les définitions suivantes

Définition 1.2.2 *Soit $x, y \in \mathbb{R}^n$; on appelle segment "fermé" d'extrémités x et y, l'ensemble noté $[x, y]$ et défini par :*

$$[x, y] = \{z \in \mathbb{R}^n : z = (1-\lambda)\,x + \lambda y \ : \lambda \in [0, 1]\}\,.$$

C'est l'ensemble de toutes les combinaisons linéaires convexes des points x et y.

De façon analogue, on définit :

Définition 1.2.3 *On appelle segment "ouvert" d'extrémités x et y, et on le note $]x, y[$, l'ensemble*

$$]x, y[= \{z \in \mathbb{R}^n : z = (1-\lambda)\,x + \lambda y : \ \lambda \in]0, 1[\}\,.$$

On définit aussi $]x, y]$ et $[x, y[$ qui sont appelés segment semi ouvert en x respectivement en y.

$$]x, y] = \{z \in \mathbb{R}^n : z = (1-\lambda)\,x + \lambda y \ \ \lambda \in]0, 1]\}\,.$$

$$[x, y[= \{z \in \mathbb{R}^n : z = (1-\lambda)\,x + \lambda y \ \ \lambda \in [0, 1[\}\,.$$

Définition 1.2.4 *Soit C une partie de $\mathbb{R}^n$. C est convexe si seulement si pour tout $x, y \in C$, $(1-\lambda)\,x + \lambda y \in C$ pour tout $\lambda \in [0, 1]$. Autrement dit, C est convexe si seulement si C contient tout segment fermé d'extrémités deux quelconques de ses points.*

Exemple 1.2.1 *- Dans $\mathbb{R}^n$, les ensembles suivants sont convexes. $\mathbb{R}^n$, l'ensemble vide par convention, les singletons, les boules, les segments, les sous espaces affines.*

- Dans $\mathbb{R}$, les parties convexes sont les intervalles.

On a la proposition :

Proposition 1.2.1 *Une partie C de $\mathbb{R}^n$ est convexe si seulement si elle contient toute combinaison linéaire convexe de toute famille finie d'éléments qui lui appartiennent.*

Preuve : Si C contient toute combinaison linéaire convexe de familles finies d'éléments qui lui appartiennent, en particulier, prenant une famille de deux éléments x et y de C, on a $[x, y] \subset C$ et donc C est convexe.

Réciproquement, soit C un ensemble convexe de $\mathbb{R}^n$. Alors C contient toute combinaison linéaire convexe de deux quelconques de ses éléments. Donc la propriété est vraie pour une famille comportant deux éléments. Supposons qu'elle est vraie pour une famille de $k-1$ éléments.

Soit

$$\mathcal{F} = \left\{x^1, x^2, \cdots, x^k\right\}$$

une famille de k élément de C.
Soit

$$x = \sum_{i=1}^{k} \lambda_i x^i \text{ avec } \lambda_i \geq 0, \ \sum_{i=1}^{k} \lambda_i = 1.$$

On a

$$x = \sum_{i=1}^{k} \lambda_i x^i = \sum_{i=1}^{k-1} \lambda_i x^i + \lambda_k x^k.$$

Soit

$$\lambda = \sum_{i=1}^{k-1} \lambda_i.$$

On a $\lambda \in [0, 1]$.

Si $\lambda = 0$ alors $\lambda_i = 0$ pour tout $i = 1, \cdots, k-1$ et donc $\lambda_k = 1$. Il vient alors que $x = \lambda_k x^k = x^k \in C$.

Si $\lambda \neq 0$, on peut écrire

$$x = \lambda \sum_{i=1}^{k-1} \left(\frac{\lambda_i}{\lambda}\right) x^i + \lambda_k x^k.$$

L'élément

$$y = \sum_{i=1}^{k-1} \left(\frac{\lambda_i}{\lambda}\right) x^i,$$

est une combinaison linéaire convexe de $k-1$ éléments de C. C'est donc un élément de C, par hypothèse de recurrence. Donc $x = \lambda y + \lambda_k x^k$. Or $\lambda_k = 1 - \lambda$ avec $\lambda \in [0, 1]$. Donc x est combinaison linéaire convexe de deux éléments de C. Comme par hypothèse, C est convexe, on a alors $x \in C$.
□

On a les propriétés suivantes

1.2.2 Propriétés algébriques

On rappelle les notions suivantes.

Définition 1.2.5 *Une application f de $\mathbb{R}^n$ dans $\mathbb{R}^m$ est dite affine si l'une des conditions suivantes est vérifiée.*
i) Pour tout x, y dans $\mathbb{R}^n$ et $\lambda \in \mathbb{R}$, on a

$$f((1-\lambda)x + \lambda y) = (1-\lambda)f(x) + \lambda f(y).$$

ii) Il existe une application linéaire $\mathcal{L}$ de $\mathbb{R}^n$ dans $\mathbb{R}^m$ et un vecteur a dans $\mathbb{R}^m$ tels que :

$$\forall\, x \in \mathbb{R}^n, \ f(x) = \mathcal{L}(x) + a.$$

Les résultats suivants sont immédiats.

Proposition 1.2.2 *1) Si C_1 et C_2 sont convexes alors pour tous α_1 et α_2 dans $\mathbb{R}$, $\alpha_1 C_1 + \alpha_2 C_2$ est convexe.*
2) Toute intersection de parties convexes est convexe.
3) L'union de sous ensembles convexes n'est pas convexe en général, mais l'union croissante de convexes (famille mboîtée) est convexe.
4) Le produit cartésien de deux convexes $C \subset \mathbb{R}^n$ et $C' \subset \mathbb{R}^m$ est convexe dans $\mathbb{R}^n \times \mathbb{R}^m$.
5) Inversement, la projection d'un sous ensemble convexe d'un espace produit sur l'un de ses sous espaces composants est convexe.
6) L'image d'un convexe par une application affine est convexe.
7) L'image réciproque d'un convexe par une application affine est convexe.

On a la proposition suivante

Proposition 1.2.3 *Si C est convexe alors pour tout α et β positifs ou nuls, on a*

$$\alpha C + \beta C = (\alpha + \beta) C.$$

Preuve : Comme les scalaires α et β sont positifs ou nuls, le cas où $\alpha + \beta = 0$ est trivial
Considérons α et β tels que $\alpha + \beta > 0$.
L'inclusion ci-dessous est immédiate :

$$(\alpha + \beta) C \subset \alpha C + \beta C.$$

Montrons à présent que

$$\alpha C + \beta C \subset (\alpha + \beta) C.$$

Soit $z \in \alpha C + \beta C$. Alors il existe x, y dans C tels que $z = \alpha x + \beta y$.
On peut écrire :

$$z = (\alpha + \beta) \left[\frac{\alpha}{\alpha + \beta} x + \frac{\beta}{\alpha + \beta} y \right].$$

On a

$$\frac{\alpha}{\alpha + \beta} > 0, \frac{\beta}{\alpha + \beta} > 0, \frac{\alpha}{\alpha + \beta} + \frac{\beta}{\alpha + \beta} = 1.$$

Comme C est convexe, alors

$$\frac{\alpha}{\alpha + \beta} x + \frac{\beta}{\alpha + \beta} y \in C.$$

D'où le résultat □

1.2.3 Propriétés topologiques

Notons que pour $x \in \mathbb{R}^n$, et $\varepsilon > 0$, $B(x, \varepsilon)$ désigne la boule euclidienne fermée de centre x et de rayon ε.

On remarque qu'on a toujours $B(x, \varepsilon) = x + \varepsilon B(0, 1)$, $B(0, 1)$ étant la boule unité euclidienne fermée.

On rappelle que

Définition 1.2.6 *Étant donné un sous ensemble C de $\mathbb{R}^n$, son adhérence son intérieur et sa frontière sont respectivement les ensembles :*

$\overline{C} = \cap\{C + B(0,\varepsilon) : \varepsilon > 0\}$,
$\text{int}(C) = \{x \in \mathbb{R}^n : \exists\varepsilon > 0,\ B(x,\varepsilon) \subset C\}$,
$\text{Fr}(C) = \overline{C} \setminus \text{int}(C)$.

Proposition 1.2.4 *Soit C un convexe de $\mathbb{R}^n$. Alors :*
1) $\overline{C}$ est convexe.
2) $\text{int}(C)$ est convexe.
3) Si $x \in \text{int}(C)$ et $y \in \overline{C}$ alors $[x, y[\ \subset \text{int}(C)$.
4) Si $\text{int}(C) \neq \emptyset$ alors

$$\text{int}(C) = \text{int}(\overline{C}),\ \overline{C} = \overline{\text{int}(C)},\ \text{Fr}(C) = \text{Fr}(\text{int}(C)).$$

Preuve : 1) Immédiat.

2) Soient x et y deux éléments de $int(C)$, $\lambda \in [0,1]$ et $z = (1-\lambda)x + \lambda y$. D'après la définition de $\text{int}(C)$, il existe ε_1 et ε_2 tels que $B(x,\varepsilon_1) \subset C$ et $B(y,\varepsilon_2) \subset C$. Donc pour $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$, on a

$$B(x,\varepsilon) \subset C,\ B(y,\varepsilon) \subset C.$$

Comme

$$\begin{aligned} B(z,\varepsilon) &= z + B(0,\varepsilon) \\ &= (1-\lambda)x + \lambda y + B(0,\varepsilon) \\ &= (1-\lambda)x + \lambda y + ((1-\lambda)+\lambda)B(0,\varepsilon) \\ &= (1-\lambda)\left[x + B(0,\varepsilon)\right] + \lambda\left[y + B(0,\varepsilon)\right] \\ &= (1-\lambda)B(x,\varepsilon) + \lambda B(y,\varepsilon) \end{aligned}$$

et

$$(1-\lambda)B(x,\varepsilon) + \lambda B(y,\varepsilon) \subset (1-\lambda)C + \lambda C = C,$$

car C est convexe, on conclut que $B(z,\varepsilon) \subset C$. Par suite $z \in int(C)$. Donc $int(C)$ est convexe.

3) Soit $x \in \text{int}(C)$ et $y \in \overline{C}$. Soit $z = (1-\lambda)x + \lambda y$, $\lambda \in [0, 1[$. Montrons que $z \in \text{int}(C)$.

On a pour tout $\varepsilon > 0$,

$$\begin{aligned} B(z,\varepsilon) &= z + B(0,\varepsilon) \\ &= (1-\lambda)x + \lambda y + B(0,\varepsilon) \\ &\subset (1-\lambda)x + \lambda(C + B(0,\varepsilon)) + B(0,\varepsilon) \\ &= (1-\lambda)x + \lambda C + (1+\lambda)B(0,\varepsilon) \\ &= \lambda C + (1-\lambda)\left[x + B(0, \tfrac{(1+\lambda)}{(1-\lambda)}\varepsilon)\right] \end{aligned}$$

Comme $x \in \text{int}(C)$, il existe $\varepsilon' > 0$ tel que $x + B(0,\varepsilon') \subset C$.

Soit ε_0 tel que

$$\frac{(1+\lambda)}{(1-\lambda)}\varepsilon_0 < \varepsilon'.$$

On a alors

$$B(z,\varepsilon_0) \subset \lambda C + (1-\lambda)\left[x + B(0,\varepsilon_0)\right] \subset \lambda C + (1-\lambda)C = C.$$

Donc $z \in \text{int}(C)$.

4) Montrons que $\overline{\text{int}(C)} = \overline{C}$.

On sait que $\text{int}(C) \subset C$. Donc $\overline{\text{int}(C)} \subset \overline{C}$.

Soit $y \in \overline{C}$; comme $int(C) \neq \emptyset$, soit $x \in \text{int}(C)$.
D'après 3), $[x, y[\subset \text{int}(C)$: par suite $y \in \overline{\text{int}(C)}$.

Montrons maintenant que $\text{int}(\overline{C}) = \text{int}(C)$.

On a l'inclusion $\text{int}(C) \subset \text{int}(\overline{C})$.

Soit $x \in \text{int}(\overline{C})$ et $y \in \text{int}(C)$ car non vide.

Considérons les points du type $z = x + \lambda(x - y)$ avec $\lambda > 0$. Pour $\lambda_0 > 0$ suffisamment petit,

$$z_0 = x + \lambda_0(x - y) \in \text{int}(\overline{C}).$$

Donc $z_0 \in \overline{C}$ et $[y, z_0[\subset \text{int}(C)$.
Or $x = \frac{1}{1+\lambda_0} z_0 + \frac{\lambda_0}{1+\lambda_0} y$, donc $x \in [y, z_0[$ et par suite $x \in \text{int}(C)$.

On a par définition

$$\text{Fr}(C) = \overline{C} \setminus \text{int}(C), \ \text{Fr}(\text{int}(C)) = \overline{\text{int}(C)} \setminus \text{int}(C).$$

Comme $\overline{C} = \overline{\text{int}(C)}$ on a le résultat. □

1.3 Enveloppes convexes

Définition 1.3.1 *Soit S une partie de $\mathbb{R}^n$. On appelle enveloppe convexe de S, l'intersection de tous les convexes contenant S. C'est le plus petit convexe contenant S. On le note $conv(S)$.*

On a le résultat suivant.

Proposition 1.3.1 *Soit $S \subset \mathbb{R}^n$. L'enveloppe convexe de S, est l'ensemble de toutes les combinaisons linéaires convexes finies d'éléments de S. Autrement dit on a :*

$$conv(S) = \left\{ x = \sum_{i=1}^{k} \lambda_i x^i : k \in \mathbb{N}^*, \begin{array}{l} x^i \in S, \forall i = 1, \cdots, k, \\ \lambda_i \geq 0, \forall i = 1, \cdots, k, \sum_{i=1}^{k} \lambda_i = 1 \end{array} \right\}$$

Preuve :

Posons

$$\mathcal{C} = \left\{ x = \sum_{i=1}^{k} \lambda_i x^i : k \in \mathbb{N}^*, \begin{array}{l} x^i \in S, \forall i = 1, \cdots, k, \\ \lambda_i \geq 0, \forall i = 1, \cdots, k, \sum_{i=1}^{k} \lambda_i = 1 \end{array} \right\}$$

Soit C un convexe de $\mathbb{R}^n$ contenant S. Donc C contient toute combinaison linéaire convexe de toute famille finie d'éléments de C. Comme C contient S alors C contient toute combinaison linéaire convexe de toute famille finie d'éléments de S et donc C contient $\mathcal{C}$. Par suite $\mathcal{C}$ est contenu dans tous les convexes contenant S. On a alors $\mathcal{C} \subset conv(S)$.

D'autre part, on vérifie facilement que $\mathcal{C}$ est convexe et contient S. Et comme $conv(S)$ est le plus petit convexe contenant S, on a alors $conv(S) \subset \mathcal{C}$. D'où l'égalité $conv(S) = \mathcal{C}$. □

Proposition 1.3.2 *S est convexe si et seulement si $conv(S) = S$*

Preuve : Si S est convexe alors il est le plus petit convexe contenant S. Donc $conv(S) = S$. Réciproquement si on a $conv(S) = S$ alors S est convexe. □

Proposition 1.3.3 *1) $conv(conv(S)) = conv(S)$*
2) Si on a $A \subset B$ alors $conv(A) \subset conv(B)$

Preuve : 1) Comme $conv(S)$ est convexe, on a $conv(conv(S)) = conv(S)$.

2) Soit $A \subset B$: le plus petit convexe contenant B contient aussi A donc $conv(B)$ contient $conv(A)$. □

Proposition 1.3.4 *Soient A et B deux convexes de $\mathbb{R}^n$. On a*

$$conv(A \cup B) = \cup_{\lambda \in [0,1]} (1-\lambda)A + \lambda B.$$

Preuve : Posons $\mathcal{C} = \cup_{\lambda \in [0,1]} (1-\lambda)A + \lambda B$.

Montrons que $\mathcal{C}$ est convexe. Soient x, y dans $\mathcal{C}$ et $\alpha \in [0,1]$. Il existe alors λ_1 et λ_2 dans $[0,1]$ tels que $x \in (1-\lambda_1)A + \lambda_1 B$ et $y \in (1-\lambda_2)A + \lambda_2 B$ Comme A et B sont convexes on a alors

$$\begin{aligned}(1-\alpha)x + \alpha y &\in (1-\alpha)[(1-\lambda_1)A + \lambda_1 B] + \alpha[(1-\lambda_2)A + \lambda_2 B] \\ &= [(1-\alpha)(1-\lambda_1) + \alpha(1-\lambda_2)]A + [(1-\alpha)\lambda_1 + \alpha\lambda_2]B.\end{aligned}$$

Les coefficients $(1-\alpha)(1-\lambda_1) + \alpha(1-\lambda_2)$ et $(1-\alpha)\lambda_1 + \alpha\lambda_2$ sont positifs et de somme 1, donc $(1-\alpha)x + \alpha y \in \mathcal{C}$. Par suite $\mathcal{C}$ est convexe.

Il est immédiat que $\mathcal{C}$ contient $A \cup B$. Donc $conv(A \cup B) \subset \mathcal{C}$. En outre tout élément de $\mathcal{C}$ est combinaison linéaire convexe de deux éléments de $A \cup B$ et donc il appartient à $conv(A \cup B)$. D'où l'égalité des ensembles $conv(A \cup B)$ et de $\mathcal{C}$ □

Plus généralement on montre que

Proposition 1.3.5 *Si $k \in \mathbb{N}^*$, et $C_1, \cdots, C_k$ sont des convexes de $\mathbb{R}^n$, alors*

$$conv(\cup_{i=1}^k C_i) = \left\{ x = \sum_{i=1}^k \alpha_i x^i : \begin{array}{l} \alpha_i \geq 0, \forall i = 1, \cdots, k, \ \sum_{i=1}^k \alpha_i = 1 \\ x^i \in C_i \end{array} \right\}.$$

A présent rappelons la définition d'une famille de points affinement indépendants.

Définition 1.3.2 *Soit x^0, x^1, $\cdots$, x^k, $k+1$ points de $\mathbb{R}^n$. On dit qu'ils sont affinement indépendants si les vecteurs*

$$x^1 - x^0, \ x^2 - x^0, \ \cdots, \ x^k - x^0$$

sont linéairement indépendants

Définition 1.3.3 *L'enveloppe convexe de $k+1$ points affinement indépendants est appelée k-simplexe.*

Lemme 1.3.1 (Caratheodory)
Dans un $\mathbb{R}^n$ toute combinaison convexe de m points , $m > n+1$, peut se ramener à une combinaison convexe de $n+1$ points au plus.

Preuve : Soit

$$x = \sum_{i=1}^m \alpha_i x_i \quad \text{avec } \alpha_i \geq 0 \quad \forall\, i = 1, \cdots, m, \quad \sum_{i=1}^m \alpha_i = 1.$$

Il suffit de montrer que si $m > n+1$, on peut faire décroître m de 1. Ce fait, utilisé de façon répétée, prouvera le résultat.

Posons $z_i = x_i - x_1$ pour $i = 2, \cdots, m$. Ces vecteurs en nombre au moins égal à $n+1$ ne peuvent être linéairement independants. Il existe donc une combinaison linéaire nulle, à coefficients β_i non tous nuls, donc

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

en ayant posé

$$\beta_1 = \sum_{i=2}^{m} \beta_i.$$

On considère alors

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

On a $\gamma \neq 0$ car les β_i ne sont pas tous nuls. On vérifie que

$$x = \sum_{i=1}^{m} \alpha_i x_i = \sum_{i=1}^{m} \delta_i x_i,$$

et que les δ_i sont positifs de somme 1 avec au moins un qui est égal à 0. On a donc exprimé x comme une combinaison convexe de moins de m points. $\square$

On en déduit que

Théorème 1.3.1 (Caratheodory)
Dans $\mathbb{R}^n$ l'enveloppe convexe d'un sous-ensemble S est égal à l'ensemble des combinaisons convexes d'au plus $n+1$ points de S.

On a la propriété topologique suivante :

Proposition 1.3.6 *L'enveloppe convexe d'un ouvert est un ouvert.*

Preuve : Soit S un ouvert : montrons que $\mathrm{int}(conv(S)) = conv(S)$.

Supposons qu'il existe un élément x dans $S \cap \mathrm{Fr}(conv(S))$. Alors comme S est ouvert, il existerait un voisinage V de x contenu dans S et donc dans $conv(S)$. Or

$$x \in \mathrm{Fr}(conv(S)) = \overline{conv(S)} \setminus \mathrm{int}(conv(S)).$$

Ce qui est contradictoire.

Donc $S \cap \mathrm{Fr}(conv(S)) = \emptyset$. On a alors $S \subset \mathrm{int}(conv(S))$. Comme $conv(S)$ est le plus petit convexe contenant S, on a $conv(S) \subset \mathrm{int}(conv(S))$. On obtient donc $conv(S) = \mathrm{int}(conv(S))$. $\square$

Comme le montre l'exemple ci-dessous, l'enveloppe convexe d'un fermé n'est pas en général un fermé.

Soit

$$S = \{(x, y) \in \mathbb{R}^2 : x \geq 0, xy \geq 1\} \cup \{(0, 0)\}$$

On a

$$conv(S) = \{(x, y) : x > 0, y > 0\} \cup \{(0, 0)\}$$

qui n'est pas un fermé.

Proposition 1.3.7 *Soit $S \subset \mathbb{R}^n$.*

1) Si S est borné alors $conv(S)$ est borné.

2) Si S est compact alors $conv(S)$ est compact.

Preuve :

1) Immédiat

2) On a

$$conv(S) = \{x = \sum_{i=1}^{n+1} \lambda_i x_i : \lambda_i \geq 0, x_i \in S, \ \forall i = 1, \cdots, n+1, \sum_{i=1}^{n+1} \lambda_i = 1\}.$$

Posons

$$K = \{\lambda \in \mathbb{R}^{n+1} : \lambda_i \geq 0, \forall i = 1, \cdots, n+1, \sum_{i=1}^{n+1} \lambda_i = 1\}.$$

K est compact.

Soit

$$f : \begin{array}{l} S^{n+1} \times K \longrightarrow \mathbb{R}^n \\ (x_1, \cdots x_{n+1}) \longmapsto \sum_{i=1}^{n+1} \lambda_i x_i \end{array}$$

f est continue et $S^{n+1} \times K$ est compact ce qui implique que $f(S^{n+1} \times K)$ est compact. Or $f(S^{n+1} \times K) = conv(S)$. D'où la proposition. □

Définition 1.3.4 *Soit S une partie de $\mathbb{R}^n$, l'enveloppe convexe fermée de S est l'intersection de tous les convexes fermés contenant S. C'est le plus petit convexe fermé contenant S. On le note $\overline{conv}(S)$.*

Proposition 1.3.8 *1) Si A et B sont deux sous ensembles de $\mathbb{R}^n$ avec $A \subset B$, alors $\overline{conv}(A) \subset \overline{conv}(B)$.*

2) Si S est une partie de $\mathbb{R}^n$, on a :

$$\overline{conv}(S) = \overline{conv}(\overline{S}) = \overline{conv(S)}.$$

Preuve : La preuve de 1) est immédiate.

2) L'ensemble des convexes fermés contenant S est égal à l'ensemble des convexes fermés contenant $\overline{S}$. Donc $\overline{conv}(S) = \overline{conv}(\overline{S})$.

Montrons que

$$\overline{conv}(S) = \overline{conv(S)}.$$

On a

$$S \subset conv(S) \subset \overline{conv(S)}.$$

Or $\overline{conv(S)}$ est un convexe fermé ; donc,

$$\overline{conv}(S) \subset \overline{conv(S)}.$$

D'autre part, on a

$$conv(\overline{S}) \subset \overline{conv}(S)$$

et $S \subset \overline{S}$: donc

$$conv(S) \subset conv(\overline{S}) \subset \overline{conv}(S).$$

On en déduit alors que

$$\overline{conv(S)} \subset \overline{\overline{conv}(S)} = \overline{conv}(S).$$

Ce qui donne la deuxième inclusion et donc l'égalité recherchée. □

1.4 Points extrêmes

Définition 1.4.1 *Soit C un sous ensemble convexe non vide de $\mathbb{R}^n$.*
Un point x de C est un point extrême si

$$\forall\,(x_1, x_2, \alpha) \in C^2 \times]0,1[,\ \ x = (1-\alpha)x_1 + \alpha x_2 \Longrightarrow x_1 = x_2 = x.$$

i.e. Il n'existe pas deux points y et z de C distincts tels que $x \in]y, z[$.

L'ensemble $\mathrm{ext}(C)$ *des points extrêmes de C est appelé profil de C.*

On a la caractérisation suivante

Proposition 1.4.1 *Soit C un sous ensemble convexe de $\mathbb{R}^n$. Les propositions suivantes sont équivalentes.*
i) $x \in \mathrm{ext}(C)$
ii) $C \setminus \{x\}$ *est convexe.*

Preuve : Soit $x \in \mathrm{ext}(C)$.
Si

$$(x_1, x_2, \alpha) \in (C \setminus \{x\}) \times (C \setminus \{x\}) \times]0,1[$$

alors

$$(x_1, x_2, \alpha) \in C \times C \times]0,1[\text{ et } x_1 \neq x \neq x_2.$$

Comme C est convexe, on a

$$(1-\alpha)x_1 + \alpha x_2 \in C \text{ et } x_1 \neq x \neq x_2.$$

Alors

$$(1-\alpha)x_1 + \alpha x_2 \in C \setminus \{x\}$$

et donc $C \setminus \{x\}$ est convexe.

Réciproquement, supposons que $C \setminus \{x\}$ est convexe et soit

$$(x_1, x_2, \alpha) \in C \times C \times]0,1[\text{ avec } (1-\alpha)x_1 + \alpha x_2 = x.$$

Donc

$$(1-\alpha)x_1 + \alpha x_2 \notin C \setminus \{x\}.$$

Ce qui implique que :

$$x_1 \notin C \setminus \{x\} \text{ ou } x_2 \notin C \setminus \{x\}$$

C'est-à-dire que $x_1 = x$ ou $x_2 = x$. Comme $(1-\alpha)x_1 + \alpha x_2 = x$ avec $\alpha \in]0,1[$, on a nécessairement $x_1 = x = x_2$. D'où $x \in \mathrm{ext}(C)$. □

On admettra que

Théorème 1.4.1 *Tout convexe compact est l'enveloppe convexe de ses points extrêmes.*

Chapitre 2

Projection sur un convexe

Dans ce chapitre $\mathbb{R}^n$ est muni de sa structure euclidienne.

2.1 Théorème de projection sur un convexe

Le théorême ci-dessous est fondamental en optimisation.

Théorème 2.1.1 *Soit S un convexe fermé de $\mathbb{R}^n$ et $x \in \mathbb{R}^n$. Il existe un unique point $p(x) \in S$ dont la distance à x est minimale. C'est-à-dire*

$$\|p(x) - x\| \leq \|y - x\| \quad \forall y \in S.$$

$p(x)$ est appelé projection de x sur S.

Preuve : Soit $d = \inf_y [\|y - x\| : y \in S]$, $B(x, d+1)$ la boule fermée de centre x et de rayon $d+1$. Considérons $K = S \cap B(x, d+1)$. L'ensemble K est fermé et borné, c'est donc un compact. Alors l'application $y \mapsto \|y - x\|$ étant continue elle atteint ses bornes sur K. Donc il existe $p(x) \in K$ et donc $p(x) \in S$ tel que $\|p(x) - x\| = d$. D'où le résultat d'existence. Montrons maintenant l'unicité.

Soient a et b deux points de S tels que

$$\|a - x\| = \|b - x\| = d = \inf_y [\|y - x\| : y \in S].$$

Comme S est convexe alors $\frac{1}{2}a + \frac{1}{2}b \in S$. Donc

$$d \leq \|x - \frac{a+b}{2}\| = \frac{1}{2}\|2x - a - b\| = \frac{1}{2}\|(x - a) + (x - b)\|.$$

Ce qui implique que

$$d \leq \|x - \frac{a+b}{2}\| \leq \frac{1}{2}(\|x - a\| + \|x - b\|) = d.$$

On en déduit que $\|x - \frac{a+b}{2}\| = d$. Il s'ensuit que $\|x - a + x - b\| = \|x - a\| + \|x - b\|$ c'est-à-dire qu'il existe $\lambda \in \mathbb{R}$ tel que $x - a = \lambda(x - b)$. Donc on a $|\lambda| = 1$.

Si $\lambda = 1$, alors $x - a = x - b$ et donc $a = b$.

Si $\lambda = -1$, alors $2x = a + b$ donc $x = \frac{a+b}{2}$ et appartient donc à S. Dans ce cas on a $d = 0$ et donc $x = a = b$.

Dans tous les cas de figure, on a $a = b$. D'où l'unicité. □

Une caractérisation de la projection est donnée ici.

Théorème 2.1.2 *La projection $p(x)$ de x sur S un convexe fermé de $\mathbb{R}^n$ est l'unique point z de S satisfaisant la condition*

$$\langle y - z, x - z\rangle \leq 0 \quad \forall y \in S.$$

Autrement dit, $p(x)$ est la projection de x sur S si et seulement si

$$\langle x - p(x), y - p(x)\rangle \leq 0 \quad \forall y \in S.$$

Preuve : Soit $p(x)$ la projection sur S un convexe fermé de $\mathbb{R}^n$. Supposons qu'il existe un point $y_0 \in S$ tel que $\langle y_0 - p(x), x - p(x)\rangle > 0$.
Posons

$$\varphi(t) = \|x - [(1-t)p(x) + ty_0]\|^2, \quad t \in \mathbb{R}^+.$$

On a

$$\begin{aligned}\varphi(t) &= \|x - p(x)\|^2 + 2t\langle x - p(x), p(x) - y_0\rangle + t^2\|p(x) - y_0\|^2 \\ &= \|x - p(x)\|^2 + t\left(2\langle x - p(x), p(x) - y_0\rangle + t\|p(x) - y_0\|^2\right).\end{aligned}$$

Comme $\langle x - p(x), p(x) - y_0\rangle < 0$, pour t suffisamment petit, on aura $\varphi'(t) < 0$ et donc φ est strictement décroissante au voisinage de 0. On aura donc pour t suffisamment petit $\varphi(t) < \|x - p(x)\|^2$.

Or pour $t \in [0; 1]$, $z_t = (1-t)p(x) + ty_0 \in S$ car S est convexe. Soit donc $t \in [0; 1]$ suffisamment petit tel que $\varphi(t) < \|x - p(x)\|^2$. Pour un tel t on aura $z_t \in S$ et $\|x - z_t\|^2 < \|x - p(x)\|^2$. Ce qui contredit le fait que $p(x)$ est la projection de x sur S.

Réciproquement supposons que

$$\langle y - z; x - z\rangle \leq 0 \quad \forall y \in S.$$

Montrons que z est la projection de x sur S.

Soit $y \in S$. On a

$$\|x - y\|^2 = \|x - z + z - y\|^2 = \|x - z\|^2 + 2\langle x - z; z - y\rangle + \|y - z\|^2.$$

Par suite on a $\|x - y\|^2 \geq \|x - z\|^2$, et donc $p(x) = z$. □

2.2 Propriétés de la projection sur un convexe

L'application projection est lipschitzienne sur $\mathbb{R}^n$.

Proposition 2.2.1 *Soit S un convexe fermé non vide de $\mathbb{R}^n$ et pour tout $x \in \mathbb{R}^n$, $p(x)$ la projection de x sur S. On a*

$$\|p(x) - p(y)\| \leq \|x - y\| \quad \forall x,\ y \in \mathbb{R}^n.$$

Preuve : Soient x et y fixés dans $\mathbb{R}^n$. Considérons $p(x)$ et $p(y)$ respectivement leurs projections sur S. On sait d'après la caractérisation de la projection que $p(x)$ est tel que

$$\langle z - p(x), x - p(x)\rangle \leq 0 \quad \forall z \in S.$$

En prenant $z = p(y)$ on obtient

$$\langle p(y) - p(x), x - p(x)\rangle \leq 0. \tag{2.1}$$

De même en considérant $p(y)$ on a

$$\langle z - p(y), y - p(y)\rangle \leq 0 \quad \forall z \in S.$$

Pour $z = p(x)$ on a

$$\langle p(x) - p(y), y - p(y)\rangle \leq 0. \tag{2.2}$$

En additionnant (2.1)et (2.2) on obtient

$$\langle p(x) - p(y), y - p(y) + p(x) - x\rangle \leq 0,$$

qui est équivalent à

$$\langle p(x) - p(y), p(x) - p(y) - (x - y)\rangle \leq 0,$$

c'est-à-dire

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

On en déduit

$$\|p(x) - p(y)\|^2 \leq \|p(x) - p(y)\|\|x - y\|.$$

D'où le résultat. □

Corollaire 2.2.1 *l'application définie sur $\mathbb{R}^n$ qui à tout $x \in \mathbb{R}^n$ associe sa projection sur S un convexe fermé est continue.*

Chapitre 3

Fonctions convexes

3.1 Fonctions convexes d'une variable réelle

Définition 3.1.1 *Soient I un intervalle de $\mathbb{R}$. Une application $f : I \mapsto \mathbb{R}$ est convexe si*

$$\forall (x, y) \in I^2, \quad f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \quad \forall \lambda \in [0, 1].$$

On dit que f est concave si $-f$ est convexe.

L'application f est convexe si et seulement si, la courbe est en dessous de chacune de ses cordes (une corde étant ici un segment).

Proposition 3.1.1 *Soient I un intervalle de $\mathbb{R}$, $f : I \mapsto \mathbb{R}$. Pour tout $a \in I$ on note $\tau_a : I - \{a\} \to \mathbb{R}$ définie par $\tau_a(x) = \frac{f(x)-f(a)}{x-a}$ appelée taux d'accroissement en a. La fonction f est convexe si et seulement si pour tout $a \in I$, τ_a est une application croissante sur $I - \{a\}$.*

Preuve :

Soit $x < y$ dans $I - \{a\}$.

• $x < a < y$

Dans ce cas on peut écrire $a = x + t(y-a)$ avec $t = \frac{a-x}{y-x}$ et donc de part la convexité de f on a

$$f(a) \leq (1-t)f(x) + tf(y) \Leftrightarrow (1-t+t)f(a) \leq (1-t)f(x) + tf(y).$$

Soit $(1-t)(f(a) - f(x)) \leq t(f(y) - f(a))$. Donc en remplaçant t par sa valeur, on obtient :

$$(\frac{y-a}{y-x})(f(a) - f(x)) \leq (\frac{a-x}{y-x})(f(y) - f(a)).$$

C'est-à-dire

$$\frac{f(x) - f(a)}{x-a} \leq \frac{f(y) - f(a)}{y-a}.$$

• $a < x < y$.

Ici on peut écrire $x = a + t(y-a)$ avec $t = \frac{x-a}{y-a}$. Alors on a

$$\begin{aligned} f(x) &\leq (1-t)f(a) + tf(y) \\ \Leftrightarrow \quad f(x) - f(a) &\leq t(f(y) - f(a)) \\ \Leftrightarrow \quad f(x) - f(a) &\leq (\tfrac{x-a}{y-a})(f(y) - f(a)) \\ \Leftrightarrow \quad \tfrac{f(x)-f(a)}{x-a} &\leq \tfrac{f(y)-f(a)}{y-a} \end{aligned}$$

• $x < y < a$.
On a $y = x + t(a-x)$ avec $t = \frac{y-x}{a-x}$. Alors la convexité de f donne

$$\begin{array}{rll} & f(y) & \leq (1-t)f(x) + tf(a) \\ \Leftrightarrow & f(y) - f(a) & \leq (1-t)f(x) - (1-t)f(a)) \\ \Leftrightarrow & f(y) - f(a) & \leq (1-t)(f(x) - f(a)) \\ \Leftrightarrow & f(y) - f(a) & \leq (\frac{a-y}{a-x})(f(x) - f(a)) \\ \Leftrightarrow & \frac{f(y)-f(a)}{a-y} & \leq \frac{f(x)-f(a)}{a-x} \end{array}$$

Ce qui s'écrit encore

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(y) - f(a)}{y - a}.$$

Réciproquement supposons que pour tout $a \in I$, τ_a est une application croissante sur $I - \{a\}$.

Soit x et y, $x < y$ deux éléments de I et $\lambda \in]0,1[$. Posons $z = x + \lambda(y-x)$. On a $z < y$ et $\lambda = \frac{z-x}{y-x}$. Comme l'application τ_x est croissante sur $I - \{x\}$, on doit avoir $\tau_x(z) \leq \tau_x(y)$, c'est-à-dire :

$$\frac{f(z) - f(x)}{z - x} \leq \frac{f(y) - f(x)}{y - x}.$$

Soit

$$f(z) - f(x) \leq \frac{z - x}{y - x}(f(y) - f(x)).$$

Ce qui s'écrit encore $f(z) - f(x) \leq \lambda(f(y) - f(x))$.

D'où la convexité de f. □

Proposition 3.1.2 *Soient I un intervalle de $\mathbb{R}$ et $f : I \mapsto \mathbb{R}$ convexe. Alors f est derivable à droite et à gauche en tout point de $\mathrm{int}(I)$ et, pour tout $(a,b,c) \in I^3$ tel que $a < b < c$, on a*

$$\frac{f(b) - f(a)}{b - a} \leq f'_g(b) \leq f'_d(b) \leq \frac{f(c) - f(a)}{c - a}.$$

Preuve : Soit $a, b, c \in \mathrm{int}(\mathrm{dom}(f))$ tels que $a < b < c$ soit t el que $b = (1-t)a + tc$ on a alors $f(b) \leq (1-t)f(a) + tf(c)$. On en déduit

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}.$$

Fixons b et c et faisons tendre a vers b. On a

$$f\prime_{-}(b) \leq \frac{f(c) - f(b)}{c - b} \ \forall c > b.$$

Fixons a et b et faisons tendre c vers b. On a

$$\frac{f(b) - f(a)}{b - a} \leq f'_{+}(b), \ \forall a < b.$$

On en déduit :

$$f'_{+}(a) \leq \frac{f(b) - f(a)}{b - a} \leq f'_{-}(b) \leq f'_{+}(b) \leq \frac{f(c) - f(b)}{c - b} \leq f'_{-}(c).$$

□

Corollaire 3.1.1 *Si f est convexe sur I, alors f est continue sur* $\mathrm{int}(I)$.

Proposition 3.1.3 *Soient I un intervalle de $\mathbb{R}$ et $f : I \mapsto \mathbb{R}$ derivable sur I. f est convexe si et seulement si f' est croissante sur I.*

Proposition 3.1.4 *Soient I un intervalle borné de $\mathbb{R}$ et $f : I \mapsto \mathbb{R}$ convexe. Alors f est minorée sur I.*

Proposition 3.1.5 *Soient I un intervalle de $\mathbb{R}$, $f : I \mapsto \mathbb{R}$ convexe. On suppose que f est majorée et atteint sa borne supérieure en un point de* $\mathrm{int}(I)$. *Montrer que f est constante.*

Proposition 3.1.6 *Soient I un intervalle de $\mathbb{R}$ $(a, b) \in (\mathrm{int}(I))^2$, $f : I \mapsto \mathbb{R}$: montrer que si f est convexe sur I, alors f est lipschitzienne sur $[a, b]$.*

Proposition 3.1.7 *Soient I un intervalle ouvert de $\mathbb{R}$ et $f : I \mapsto \mathbb{R}$ convexe et dérivable. Alors f est de classe $\mathcal{C}^1$ sur I.*

3.2 Fonctions convexes à plusieurs variables

3.2.1 Définitions

Définition 3.2.1 *Soit $f : \mathbb{R}^n \to]-\infty, +\infty]$.*
On appelle :

- domaine effectif de f, l'ensemble

$$\mathrm{dom}(f) = \{x \in \mathbb{R}^n : f(x) < +\infty\}.$$

- épigraphe de f, l'ensemble

$$\mathrm{epi}(f) = \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq \lambda\}.$$

- épigraphe strict de f, l'ensemble

$$\widetilde{\mathrm{epi}}(f) = \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : f(x) < \lambda\}.$$

- section de niveau λ de f, l'ensemble

$$S_\lambda(f) = \{x \in \mathbb{R}^n : f(x) \leq \lambda\}.$$

- section stricte de niveau λ de f, l'ensemble

$$\widetilde{S}_\lambda(f) = \{x \in \mathbb{R}^n : f(x) < \lambda\}.$$

Il est clair qu'un sous ensemble F de $\mathbb{R}^n \times \mathbb{R}$ est un épigraphe si et seulement si il est "infini" vers le haut, c'est-à-dire que si $(x, \lambda) \in F$, alors $(x, \lambda') \in F$ dès que $\lambda' \geq \lambda$. Ainsi

$$\mathrm{epi}(f) = F \Longleftrightarrow \forall\, x \in \mathbb{R}^n, \quad f(x) = \inf_{(x,\lambda)\in F} \lambda.$$

On définit

Définition 3.2.2 *Une fonction* $f : \mathbb{R}^n \to]-\infty, +\infty]$ *est dite propre si* $\mathrm{dom}(f)$ *est non vide.*

Définition 3.2.3 *Une fonction* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *est convexe si*

$$\forall\, x,\ y \in \mathrm{dom} f, \quad f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \quad \forall\, \lambda \in [0,1].$$

Définition 3.2.4 *Une fonction* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *est strictement convexe, si*

$$\forall\, x,\ y \in \mathrm{dom} f, \quad x \neq y, \quad f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y) \quad \forall\, \lambda \in]0,1[.$$

Définition 3.2.5 *Une fonction* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *est fortement convexe de module* $r > 0$*, si*

$$\forall\, x,\ y \in \mathrm{dom} f,\ f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) - \tfrac{r}{2}\lambda(1-\lambda)\|y - x\|^2 \quad \forall\, \lambda \in [0,1].$$

Définition 3.2.6 $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ *est concave si*

$$\forall\, x,\ y \in \mathrm{dom} f, f((1-\lambda)x + \lambda y) \geq (1-\lambda)f(x) + \lambda f(y) \quad \forall\, \lambda \in [0,1].$$

Définition 3.2.7 *Une fonction* $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ *est strictement concave, si*

$$\forall\, x,\ y \in \mathrm{dom} f, x \neq y, f((1-\lambda)x + \lambda y) > (1-\lambda)f(x) + \lambda f(y) \quad \forall\, \lambda \in]0,1[.$$

Définition 3.2.8 *Une fonction* $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ *est fortement concave de module* $r > 0$*, si*

$$\forall\, x,\ y \in \mathrm{dom} f,\ f((1-\lambda)x + \lambda y) \geq (1-\lambda)f(x) + \lambda f(y) - \tfrac{r}{2}\lambda(1-\lambda)\|y - x\|^2 . \forall\, \lambda \in [0,1].$$

Il est immédiat que

Proposition 3.2.1 *Une fonction* $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ *est respectivement concave, strictement concave, fortement concave de module* $r > 0$*, si* $-f$ *définie sur* $\mathbb{R}^n$ *et à valeurs dans* $\mathbb{R} \cup \{+\infty\}$ *est respectivement convexe, strictement convexe, fortement convexe de module* $r > 0$*.*

On montre facilement que :

Proposition 3.2.2 *Soit* C *un convexe non vide de* $\mathbb{R}^n$ *et* $f : C \to \mathbb{R}$*.* f *est convexe sur* C *si et seulement si, la fonction étendue on dit aussi prolongement canonique de* f*, définie sur* $\mathbb{R}^n$ *par*

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in C \\ +\infty & \text{sinon} \end{cases}$$

est convexe sur $\mathbb{R}^n$*.*

Pour cela, pour étudier la convexité des fonctions on peut supposer sans perte de généralités, que les fonctions sont définies sur $\mathbb{R}^n$ tout entier.

On a le lemme suivant :

Lemme 3.2.1 *Soit* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\mathrm{epi}(f)$ *est convexe si et seulement si* $\widetilde{\mathrm{epi}}(f)$ *est convexe.*

On a une caractérisation géométrique de la convexité d'une fonction.

Proposition 3.2.3 *Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Les propositions suivantes sont équivalentes :*
i) f est convexe.
ii) L'épigraphe de f, $(\mathrm{epi}(f))$, est convexe.
iii) L'épigraphe strict de f est convexe.

La démonstration est immédiate.

On a pour les fonctions concaves des résultats analogues.

Définition 3.2.9 *On appelle hypographe de f, l'ensemble :*

$$hyp(f) = \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : f(x) \geq \lambda\}.$$

On appelle hypographe strict de f, l'ensemble :

$$\widetilde{hyp}(f) = \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : f(x) > \lambda\}.$$

Proposition 3.2.4 *Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$. Les propositions suivantes sont équivalentes :*
i) f est concave.
ii) L'hypographe de f, $(\mathrm{epi}(f))$, est convexe.
iii) L'hypographe strict de f est convexe.

On a les caractérisations suivantes :

Proposition 3.2.5 *Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ propre. On a les équivalences suivantes.*
i) f est convexe
ii) Pour toute combinaison linéaire convexe $x = \sum_{i=1}^{k} \lambda_i x^i$, on a

$$f(\sum_{i=1}^{k} \lambda_i x^i) \leq \sum_{i=1}^{k} \lambda_i f(x^i).$$

Proposition 3.2.6 *$f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ est convexe (respectivement strictement convexe) si et seulement si pour toute droite $D \subset \mathbb{R}^n$, la restriction de f à D est convexe (respectivement strictement convexe). C'est-à-dire, pour tout a et d dans $\mathbb{R}^n$, la fonction $\varphi_{a,d}$ définie sur $\mathbb{R}$ par $\varphi_{a,d}(t) = f(a + td)$ est convexe (respectivement strictement convexe).*

Preuve : Supposons f convexe. Soit $a,\ d \in \mathbb{R}^n$, montrons que $\varphi_{a,d}$ est convexe.
Soient t_1 et t_2 deux réels et $\lambda \in [0, 1]$. On a :

$$\begin{aligned}
\varphi_{a,d}(\lambda t_1 + (1-\lambda)t_2) &= f(a + (\lambda t_1 + (1-\lambda)t_2)d) \\
&= f(\lambda a + (1-\lambda)a + \lambda t_1 d + (1-\lambda)t_2 d) \\
&= f(\lambda(a + t_1 d) + (1-\lambda)(a + t_2 d)) \\
&\leq \lambda f(a + t_1 d) + (1-\lambda) f(a + t_2 d) \\
&= \lambda \varphi_{a,d}(t_1) + (1-\lambda)\varphi_{a,d}(t_2)
\end{aligned}$$

d'où la convexité de $\varphi_{a,d}$.

Réciproquement supposons que pour tous $a,\ d \in \mathbb{R}^n$, la fonction φ définie sur $\mathbb{R}$ par $\varphi(t) = f(a + td)$ est convexe. Montrons que f est convexe.

Soient $x,\ y \in \mathbb{R}^n$ et $\lambda \in [0,\,1]$. On a

$$\begin{aligned} f(\lambda x + (1-\lambda)y) = f(y + \lambda(x-y)) &= \varphi_{y,\,x-y}(\lambda) \\ &= \varphi_{y,\,x-y}(\lambda \times 1 + (1-\lambda)\times 0). \end{aligned}$$

Comme $\varphi_{y,\,x-y}$ est convexe, on a

$$\begin{aligned} \varphi_{y,\,x-y}(\lambda \times 1 + (1-\lambda)\times 0) &\leq \lambda\varphi_{y,\,x-y}(1) + (1-\lambda)\varphi_{y,\,x-y}(0) \\ &= \lambda f(x) + (1-\lambda) f(y) \end{aligned}$$

d'où la convexité de f.

Pour la stricte convexité, on procède de la même façon. □

Proposition 3.2.7 *Si $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ est convexe propre alors pour $a \in \mathrm{dom} f$ et $d \in \mathbb{R}^n$, l'application définie sur $]0, +\infty[$ par*

$$\varphi(t) = \frac{f(a+td) - f(a)}{t}$$

est croissante.

Preuve : Soient $0 < t_1 \leq t_2$. On a alors $0 < \frac{t_1}{t_2} \leq 1$. Donc

$$f(a + t_1 d) = f((1 - \frac{t_1}{t_2})a + \frac{t_1}{t_2}(a + t_2 d)) \leq (1 - \frac{t_1}{t_2}) f(a) + \frac{t_1}{t_2} f(a + t_2 d).$$

Il s'ensuit alors que

$$\frac{f(a+t_1 d) - f(a)}{t_1} \leq \frac{f(a+t_2 d) - f(a)}{t_2}.$$

Ce qui prouve la proposition. □

On a aussi la proposition suivante.

Proposition 3.2.8 *Si $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ est convexe alors les sections de niveau $S_\lambda(f)$, $\lambda \in \mathbb{R}$ sont convexes.*

Preuve : Soit $\lambda \in \mathbb{R}$ et x, y deux éléments de $S_\lambda(f)$ et $\alpha \in [0, 1]$.
La convexité de f et la définition de $S_\lambda(f)$ nous donne :

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

Donc $(1-\alpha)x + \alpha y \in S_\lambda(f)$ qui est donc convexe. □

Remarque 3.2.1 *La réciproque de cette proposition n'est pas vraie.*

On a dans la proposition ci-dessous une caractérisation de la forte convexité.

Proposition 3.2.9 *Une fonction $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ est fortement convexe de module r si et seulement si la fonction g définie sur $\mathbb{R}^n$ par $g(x) = f(x) - \frac{1}{2} r\|x-a\|^2$ ($a \in \mathbb{R}^n$) est convexe.*

Preuve : La fonction g est convexe si et seulement si

$$\forall\, x,\ y \in \mathbb{R}^n, \forall\, \lambda \in]0,1[,$$
$$g((1-\lambda)x + \lambda y) \leq (1-\lambda)g(x) + \lambda g(y). \tag{3.1}$$

Posons

$$\mu = \|(1-\lambda)x + \lambda y - a\|^2 - (1-\lambda)\|x-a\|^2 - \lambda\|y-a\|^2.$$

La condition (3.1) est alors équivalente à :

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) + \frac{1}{2} r\mu.$$

Or

$$\begin{aligned}
\mu &= \|(1-\lambda)x + \lambda y - a\|^2 - (1-\lambda)\|x-a\|^2 - \lambda\|y-a\|^2 \\
&= \|(1-\lambda)(x-a) + \lambda(y-a)\|^2 - (1-\lambda)\|x-a\|^2 - \lambda\|y-a\|^2 \\
&= (1-\lambda)^2\|x-a\|^2 + \lambda^2\|y-a\|^2 + 2\lambda(1-\lambda)\langle x-a, y-a\rangle \\
&\qquad -(1-\lambda)\|x-a\|^2 - \lambda\|y-a\|^2 \\
&= -\lambda(1-\lambda)\left(\|x-a\|^2 + \|y-a\|^2 - 2\langle x-a, y-a\rangle\right) \\
&= -\lambda(1-\lambda)\|y-x\|^2.
\end{aligned}$$

La condition (3.1) est donc encore équivalente à :

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) - \frac{1}{2} r\lambda(1-\lambda)\|y-x\|^2.$$

Ce qui signifie que la fonction f est fortement convexe. □

3.2.2 Opération sur les fonctions convexes

On montre facilement que

Proposition 3.2.10 *Si $f : \mathbb{R}^n \to \mathbb{R}$ est convexe et $\varphi : \mathbb{R} \to \mathbb{R}$ est convexe et croissante alors la fonction $h = \varphi \circ f$ est convexe.*

On en déduit alors que

Proposition 3.2.11 *Si $f : \mathbb{R}^n \to \mathbb{R}$ est concave et $\varphi : \mathbb{R} \to \mathbb{R}$ est concave et croissante alors la fonction $h = \varphi \circ f$ est concave.*

On peut aussi montrer que

Proposition 3.2.12 *Si $f : \mathbb{R}^n \to \mathbb{R}$ est strictement convexe et $\varphi : \mathbb{R} \to \mathbb{R}$ est convexe et strictement croissante alors la fonction $h = \varphi \circ f$ est strictement convexe.*

On tire les conséquences suivantes :

Corollaire 3.2.1 *Si $f : \mathbb{R}^n \to \mathbb{R}$ est convexe alors la fonction e^f est convexe.*

Corollaire 3.2.2 *Si $f : \mathbb{R}^n \to \mathbb{R}_+^*$ est concave alors la fonction $\frac{1}{f}$ est convexe.*

On obtient aussi des fonctions convexes en considérant les opérations suivantes :

Proposition 3.2.13 *Toute combinaison linéaire finie et positive de fonctions convexes est convexe. C'est-à-dire : Si pour tout $i = 1, \cdots, p$, $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ est convexe, alors pour tout $\alpha_i \geq 0$, $i = 1, \cdots, p$, la fonction $f = \sum_{i=1}^{p} \alpha_i f_i$ est convexe.*

Proposition 3.2.14 *L'enveloppe supérieure d'une famille de fonction convexes est convexe. Autrement dit, si $\{f_i\}_{i\in I}$ est une famille quelconque de fonctions convexes définies sur $\mathbb{R}^n$ et à valeurs dans $\mathbb{R} \cup \{+\infty\}$, alors la fonction f définie par $f(x) = \sup_{i\in I} f_i(x)$ est une fonction convexe.*

Preuve : Soient x et y deux éléments de $\mathbb{R}^n$ et $\lambda \in [0, 1]$. Comme pour tout $i \in I$, f_i est convexe, on a

$$\begin{aligned} \forall\, i \in I,\ f_i((1-\lambda)x + \lambda y) &\leq (1-\lambda) f_i(x) + \lambda f_i(y) \\ &\leq (1-\lambda) \sup_{k\in I} f_k(x) + \lambda \sup_{k\in I} f_k(y). \end{aligned}$$

Donc

$$\sup_{i\in I} f_i((1-\lambda)x + \lambda y) \leq (1-\lambda) \sup_{k\in I} f_k(x) + \lambda \sup_{k\in I} f_k(y).$$

Ce qui signifie que f est convexe. □

Remarque 3.2.2 *On peut démontrer ce résultat en vérifiant que*

$$\mathrm{epi}(f) = \cap_{i\in I} \mathrm{epi}(f_i).$$

Et comme les f_i sont convexes, leurs épigraphes sont convexes et donc l'épigraphe de f aussi. Par suite f est convexe.

On sait que l'image réciproque d'un convexe par une transformation affine est convexe. On utilise cette propriété pour montrer le résultat ci-dessous.

Proposition 3.2.15 *Soit A une transformation affine de $\mathbb{R}^n$ dans $\mathbb{R}^m$ et $f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe. Alors la fonction notée fA définie sur $\mathbb{R}^n$ par $fA(x) = f(Ax)$ est convexe.*

Preuve : Considérons la transformation affine

$$\begin{aligned} T :\ & \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m \times \mathbb{R} \\ & (x, \lambda) \longmapsto (Ax, \lambda) \end{aligned}$$

La fonction f étant convexe, epi(f) est un convexe de $\mathbb{R}^n \times \mathbb{R}$. Comme

$$\begin{aligned} T^{-1}(\mathrm{epi}(f)) &= \{(x,\lambda) : T(x,\lambda) \in \mathrm{epi}(f)\} \\ &= \{(x,\lambda) : (Ax,\lambda) \in \mathrm{epi}(f)\} \\ &= \{(x,\lambda) : f(Ax) \leq \lambda\} \\ &= \mathrm{epi}(fA) \end{aligned}$$

on en déduit que fA est convexe. □

On pouvait démontrer ce résultat directement en utilisant la définition d'une fonction convexe.

Proposition 3.2.16 *Si $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ est conjointement convexe par rapport aux deux variables, alors la fonction f définie sur $\mathbb{R}^n$ par*

$$f(x) = \inf_{y \in \mathbb{R}^m} h(x, y)$$

est convexe.

Preuve : Il faut et il suffit de montrer que

$$\forall \varepsilon > 0,\ \forall x^1,\ x^2 \in \mathbb{R}^n,\ \forall \lambda \in [0, 1],$$

$$f((1-\lambda)x^1 + \lambda x^2) \leq (1-\lambda)f(x^1) + \lambda f(x^2) + \varepsilon.$$

Soit donc $\varepsilon > 0$ et considérons x^1, x^2 dans $\mathbb{R}^n$ et $\lambda \in [0, 1]$.

Il existe $y^1 \in \mathbb{R}^m$ tel que

$$f(x^1) \leq h(x^1, y^1) \leq f(x^1) + \varepsilon.$$

Il existe $y^2 \in \mathbb{R}^m$ tel que

$$f(x^2) \leq h(x^2, y^2) \leq f(x^2) + \varepsilon.$$

Comme $(1-\lambda)y^1 + \lambda y^2 \in \mathbb{R}^m$, on a

$$f((1-\lambda)x^1 + \lambda x^2) \leq h((1-\lambda)x^1 + \lambda x^2, (1-\lambda)y^1 + \lambda y^2).$$

La fonction h est convexe. Donc

$$\begin{array}{rcl} h((1-\lambda)x^1 + \lambda x^2, (1-\lambda)y^1 + \lambda y^2) & = & h((1-\lambda)(x^1, y^1) + \lambda(x^2, y^2)) \\ & \leq & (1-\lambda)h(x^1, y^1) + \lambda h(x^2, y^2) \\ & \leq & (1-\lambda)(f(x^1) + \varepsilon) + \lambda(f(x^2) + \varepsilon) \\ & = & (1-\lambda)f(x^1) + \lambda f(x^2) + \varepsilon. \end{array}$$

On tire alors l'inégalité recherchée

$$f((1-\lambda)x^1 + \lambda x^2) \leq (1-\lambda)f(x^1) + \lambda f(x^2) + \varepsilon.$$

D'où la convexité de f. □

Proposition 3.2.17 *Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe et $\lambda > 0$. Alors la fonction notée $f\lambda$ et définie par $(f\lambda)(x) = \lambda f(\frac{x}{\lambda})$ est convexe.*

Preuve : On a en effet $\mathrm{epi}(f\lambda) = \lambda \mathrm{epi}(f)$. D'où le résultat. □

Définition 3.2.10 *Soient $f,\ g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$: on appelle inf-convolution de f et de g la fonction notée $f \square g$ et définie par*

$$(f \square g)(x) = \inf_{x_1 + x_2 = x} \left[f(x_1) + g(x_2) \right].$$

Proposition 3.2.18 *On a $\widetilde{\mathrm{epi}}(f \square g) = \widetilde{\mathrm{epi}}(f) + \widetilde{\mathrm{epi}}(g)$*

Preuve : Dans cette démonstration on utilise l'équivalence suivante.
Si a, b et c sont des réels alors

$$a + b < c \Leftrightarrow \exists\, \alpha,\ \beta,\ \alpha > a,\ \beta > b : \alpha + \beta = c. \tag{3.2}$$

On a aussi :

$$a + b \le c \Leftrightarrow \exists\, \alpha,\ \beta,\ \alpha \ge a,\ \beta \ge b : \alpha + \beta = c.$$

Soit $(x, \lambda) \in \widetilde{\text{epi}}(f \Box g)$. Cela signifie :

$$\exists\, x_1,\ x_2 \in \mathbb{R}^n,\ x_1 + x_2 = x : f(x_1) + g(x_2) < \lambda.$$

D'après (3.2), cela équivaut à :

$$\exists\, x_1\ x_2 \in \mathbb{R}^n, \exists\, \lambda_1,\ \lambda_2 \in \mathbb{R} : x_1 + x_2 = x,\ \lambda_1 + \lambda_2 = \lambda,\ f(x_1) < \lambda_1,\ g(x_2) < \lambda_2.$$

C'est-à-dire

$$\exists\, (x_1, \lambda_1) \in \widetilde{\text{epi}}(f),\ (x_2, \lambda_2) \in \widetilde{\text{epi}}(g) : (x_1, \lambda_1) + (x_2, \lambda_2) = (x, \lambda).$$

D'où le résultat. □

Comme une fonction f est convexe si et seulement si son épigraphe strict est convexe, on a :

Proposition 3.2.19 *Si les fonctions f et g sont convexes propres, alors $f \Box g$ est convexe.*

Remarque 3.2.3 *On peut montrer la proposition ci-dessus en remarquant que la fonction*

$$\begin{array}{rcl} h & : & \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \\ & & (x,\, y) \longmapsto h(x, y) = f(x - y) + g(y) \end{array}$$

est convexe conjointement par rapport aux deux variables (x, y). Donc

$$f \Box g(x) = \inf_y \left[f(x - y) + g(y)\right] = \inf_y h(x, y)$$

est convexe.

On montre que

Proposition 3.2.20 *La loi de composition inf-convolution est commutative et associative.*

3.2.3 Quelques fonctions convexes particulières

Définition 3.2.11 *Soit S un sous ensemble non vide de $\mathbb{R}^n$. On appelle fonction indicatrice de S, la fonction notée δ_S ou $\delta(., S)$ définie sur $\mathbb{R}^n$ par*

$$\delta_S(x) = \delta(x, S) = \begin{cases} 0 & \text{si } x \in S \\ +\infty & \text{sinon} \end{cases}$$

On peut se demander pourquoi aller chercher la valeur infinie (et ne pas se contenter de 1 par exemple). Une des raisons est qu'une telle fonction ne serait pas convexe même lorsque S est convexe. Il est immédiat de montrer que

Proposition 3.2.21 *S est convexe si et seulement si δ_S est convexe.*

Définition 3.2.12 *Soit C un sous ensemble non vide de $\mathbb{R}^n$: on appelle fonction support de C la fonction notée σ_C ou $\sigma(.,C)$ définie sur $\mathbb{R}^n$ par :*

$$\sigma_C(x) = \sigma(x, C) = \sup_y \left[\langle x, y\rangle : y \in C\right].$$

Proposition 3.2.22 *La fonction support est une fonction convexe.*

Définition 3.2.13 *Soit C un convexe non vide de $\mathbb{R}^n$. On appelle fonction jauge ou fonction de Minkowski de C, la fonction notée γ_C ou $\gamma(.,C)$ définie sur $\mathbb{R}^n$ par :*

$$\gamma_C(x) = \gamma(x, C) = \inf_\lambda \left[\lambda \geq 0 : x \in \lambda C\right].$$

Proposition 3.2.23 *La fonction jauge est une fonction convexe.*

Preuve : Soient x_1, x_2 dans $\mathbb{R}^n$ et $\alpha \in [0,1]$.
Soit $\varepsilon > 0$.

$$\exists \lambda_1 \geq 0 : x_1 \in \lambda_1 C : \gamma_C(x_1) \leq \lambda_1 \leq \gamma_C(x_1) + \varepsilon.$$

$$\exists \lambda_2 \geq 0 : x_2 \in \lambda_2 C : \gamma_C(x_2) \leq \lambda_2 \leq \gamma_C(x_2) + \varepsilon.$$

Comme C est convexe, on a

$$(1-\alpha)x_1 + \alpha x_2 \in (1-\alpha)\lambda_1 C + \alpha\lambda_2 C = ((1-\alpha)\lambda_1 + \alpha\lambda_2)C.$$

Comme $(1-\alpha)\lambda_1 + \alpha\lambda_2 \geq 0$, alors

$$\begin{aligned} \gamma_C((1-\alpha)x_1 + \alpha x_2) &\leq (1-\alpha)\lambda_1 + \alpha\lambda_2 \\ &\leq (1-\alpha)(\gamma_C(x_1) + \varepsilon) + \alpha(\gamma_C(x_2) + \varepsilon) \\ &= (1-\alpha)\gamma_C(x_1) + \alpha\gamma_C(x_2) + \varepsilon \end{aligned}$$

Ce qui termine la démonstration □

Définition 3.2.14 *Soit C un convexe non vide de $\mathbb{R}^n$. On appelle fonction distance Euclidienne de C, la fonction notée d_C ou $\mathrm{d}(.,C)$ définie sur $\mathbb{R}^n$ par :*

$$\mathrm{d}_C(x) = \mathrm{d}(x, C) = \inf_y \left[\|x - y\| : y \in C\right].$$

Proposition 3.2.24 *La fonction distance est une fonction convexe.*

Preuve : On peut montrer ce résultat de plusieurs façons.
1) On remarque que

$$\mathrm{d}_C(x) = \inf_y \left[\|x - y\| + \delta_C(y)\right],$$

c'est-à-dire l'inf-convolution de la norme et de la fonction indicatrice qui sont toutes les deux convexes. Par suite la fonction distance est convexe.

2) On considère la fonction h définie par : $h(x,y) = \|x - y\|$.
Elle est convexe en (x,y) et comme $\mathrm{d}_C(x) = \inf_y h(x,y)$, alors la fonction distance est convexe.
□

3.2.4 Caractérisation des fonctions convexes différentiables

Dans les résultats qui suivent, nous donnons une caractérisation de la convexité dans le cas différentiable.

Théorème 3.2.1 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable. On a les équivalences suivantes :*
1) f est convexe sur $\mathbb{R}^n$
2) $f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle, \ \forall\, x, y \in \mathbb{R}^n$.
3) $\langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0, \ \forall\, x, y \in \mathbb{R}^n$.

Preuve : 1)$\Rightarrow$ 2) Soient x, y dans $\mathbb{R}^n$ et $\lambda \in]0,1[$. Comme f est convexe, alors on a

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

Ce qui donne

$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda} \leq f(y) - f(x).$$

En passant à la limite, on obtient :

$$\langle \nabla f(x), y - x\rangle \leq f(y) - f(x).$$

D'où la proposition 2).

2)$\Rightarrow$ 1) On sait par hypothèse que

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle, \ \forall\, (x, y) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Soient x, et y dans $\mathbb{R}^n$ et $\lambda \in [0, 1]$. En considérant respectivement les couples $(x + \lambda(y - x), x)$ et $(x + \lambda(y - x), y)$, on a :

$$f(x) \geq f(x + \lambda(y - x) - \lambda\langle \nabla f(x + \lambda(y - x)), y - x\rangle \tag{3.3}$$

et

$$f(y) \geq f(x + \lambda(y - x) + (1 - \lambda)\langle \nabla f(x + \lambda(y - x), y - x\rangle \tag{3.4}$$

On multiplie (3.3) par $(1 - \lambda)$ et (3.4) par λ et on fait la somme des deux résultats. On obtient alors

$$(1 - \lambda)f(x) + \lambda f(y) \geq f(x + \lambda(y - x).$$

Ce qui prouve que f est convexe.

2)$\Rightarrow$ 3) Soient x et y dans $\mathbb{R}^n$. On a

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle \tag{3.5}$$

et

$$f(x) \geq f(y) + \langle \nabla f(y), x - y\rangle \tag{3.6}$$

En considérant la somme de (3.5) et de (3.6), on obtient

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

Montrons à présent que la proposition 3) implique 2).

3)$\Rightarrow$ 2) Soient x et y dans $\mathbb{R}^n$. Comme f est différentiable alors :

$$\exists\, z \in]x, y[\,:\, f(y) - f(x) = \langle \nabla f(z), y - x \rangle \tag{3.7}$$

Comme $z \in]x, y[$, il existe $\lambda \in]0, 1[$ tel que $z = x + \lambda(y - x)$.
D'après la proposition 3), on a :

$$\langle \nabla f(z) - \nabla f(x), z - x \rangle \geq 0.$$

Or $z - x = \lambda(y - x)$. Il vient donc

$$\lambda \langle \nabla f(z) - \nabla f(x), y - x \rangle \geq 0.$$

Soit

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

car $\lambda \in]0, 1[$. C'est-à-dire

$$\langle \nabla f(z), y - x \rangle \geq \langle \nabla f(x), y - x \rangle.$$

En utilisant (3.7), on obtient

$$f(y) - f(x) = \langle \nabla f(z), y - x \rangle \geq \langle \nabla f(x), y - x \rangle.$$

D'où la proposition 2). Ce qui termine la démonstration □

On a des résultats similaires pour la stricte convexité.

Théorème 3.2.2 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable.*
On a les équivalences suivantes :
1) f est strictement convexe sur $\mathbb{R}^n$.
2) $f(y) > f(x) + \langle \nabla f(x), y - x \rangle,\ \forall\, x, y \in \mathbb{R}^n,\ x \neq y$.
3) $\langle \nabla f(y) - \nabla f(x), y - x \rangle > 0,\ \forall\, x, y \in \mathbb{R}^n,\ x \neq y$.

Donnons à présent les résultats concernant la forte convexité.

Théorème 3.2.3 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable. On a les équivalences suivantes :*
1) f est fortement convexe de module $r > 0$ sur $\mathbb{R}^n$.
2) $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{r}{2}\|y - x\|^2,\ \forall\, x, y \in \mathbb{R}^n$.
3) $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq r\|y - x\|^2,\ \forall\, x, y \in \mathbb{R}^n$.

Preuve : 1)$\Rightarrow$ 2) Soient x, y dans $\mathbb{R}^n$ et $\lambda \in]0, 1[$. Comme f est fortement convexe de module r, alors on a

$$f(x + \lambda(y - x)) - f(x) \leq \lambda(f(y) - f(x)) - \frac{r}{2}\lambda(1 - \lambda)\|y - x\|^2.$$

Ce qui donne

$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda} \leq f(y) - f(x) - \frac{r}{2}(1 - \lambda)\|y - x\|^2.$$

En passant à la limite, on obtient :

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x) - \frac{r}{2}\|y - x\|^2.$$

D'où la proposition 2).

2)⇒ 1) On sait par hypothèse que

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{r}{2}\|y - x\|^2, \ \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Soient x, et y dans $\mathbb{R}^n$ et $\lambda \in [0, 1]$. En considérant respectivement les couples $(x + \lambda(y - x), x)$ et $(x + \lambda(y - x), y)$, on a :

$$f(x) \geq f(x + \lambda(y - x)) - \lambda\langle \nabla f(x + \lambda(y - x)), y - x\rangle + \frac{r}{2}\lambda^2\|y - x\|^2 \tag{3.8}$$

et

$$f(y) \geq f(x + \lambda(y - x) + (1 - \lambda)\langle \nabla f(x + \lambda(y - x)), y - x\rangle + \frac{r}{2}(1 - \lambda)^2\|y - x\|^2 \tag{3.9}$$

On multiplie (3.8) par $(1-\lambda)$ et (3.9) par λ et on fait la somme des deux résultats. On obtient alors

$$(1 - \lambda)f(x) + \lambda f(y) \geq f(x + \lambda(y - x) + \frac{r}{2}\lambda(1 - \lambda)\|y - x\|^2.$$

Ce qui prouve que f est fortement convexe de module r.

2)⇒ 3) Soit x et y dans $\mathbb{R}^n$. On a

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{r}{2}\|y - x\|^2 \tag{3.10}$$

et

$$f(x) \geq f(y) + \langle \nabla f(y), x - y\rangle + \frac{r}{2}\|y - x\|^2 \tag{3.11}$$

En considérant la somme de (3.10) et de (3.11), on obtient

$$\langle \nabla f(y) - \nabla f(x), y - x\rangle \geq r\|y - x\|^2.$$

3)⇒ 2) Soient x et y dans $\mathbb{R}^n$. Considérons la fonction φ définie sur $[0, 1]$ et à valeurs dans $\mathbb{R}$ par :

$$\varphi(t) = f(x + t(y - x)) - f(x) - t\langle \nabla f(x), y - x\rangle - \frac{t^2}{2}r\|y - x\|^2.$$

Par hypothèse, on a :

$$\langle \nabla f(x + t(y - x)) - \nabla f(x), x + t(y - x) - x\rangle \geq r\|x + t(y - x) - x\|^2.$$

Soit

$$t\langle \nabla f(x + t(y - x)) - \nabla f(x), y - x\rangle \geq rt^2\|y - x\|^2.$$

Donc pour $t > 0$, on a :

$$\langle \nabla f(x + t(y - x)) - \nabla f(x), y - x\rangle \geq rt\|y - x\|^2. \tag{3.12}$$

Comme f est différentiable, la fonction φ est derivable et pour $t \in]0, 1[$, on a :

$$\varphi'(t) = \langle \nabla f(x + t(y - x)) - \nabla f(x), y - x\rangle - rt\|y - x\|^2.$$

En considérant (3.12), il vient que $\varphi'(t)$ est positif pour tout $t \in]0, 1]$. Donc la fonction φ est croissante sur cet intervalle. On a alors $\varphi(1) \geq \varphi(0) = 0$. Soit

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{r}{2}\|y - x\|^2.$$

D'où le théorème. □

Dans le cas où la fonction est deux fois différentiable, on a aussi les caractérisations suivantes.

Théorème 3.2.4 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable. On a les équivalences suivantes :*
1) f est convexe sur $\mathbb{R}^n$.
2) Pour tout $x \in \mathbb{R}^n$, $\langle \nabla^2 f(x)h, h\rangle \geq 0 \; \forall\, h \in \mathbb{R}^n$.

Preuve : Soit $x \in \mathbb{R}^n$. On sait que pour tout $h \in \mathbb{R}^n$ et $t \in \mathbb{R}$, suffisamment petit, on a

$$f(x+th) = f(x) + t\langle \nabla f(x), h\rangle + \frac{t^2}{2}\langle \nabla^2 f(x)h, h\rangle + t^2\|h\|^2\varepsilon(t).$$

Par hypothèse, la fonction f est convexe, donc on a pour tout $h \in \mathbb{R}^n$ et $t \in \mathbb{R}$,

$$f(x+th) \geq f(x) + t\langle \nabla f(x), h\rangle.$$

Donc

$$f(x) + t\langle \nabla f(x), h\rangle + \frac{t^2}{2}\langle \nabla^2 f(x)h, h\rangle + t^2\|h\|^2\varepsilon(t) \geq f(x) + t\langle \nabla f(x), h\rangle.$$

Ce qui donne

$$\langle \nabla^2 f(x)h, h\rangle + \|h\|^2\varepsilon(t) \geq 0.$$

Comme la fonction ε tend vers 0 quand t tend vers 0, on obtient :

$$\langle \nabla^2 f(x)h, h\rangle \geq 0.$$

On conclut que la matrice hessienne $\nabla^2 f(x)$ est semi définie positive.

Réciproquement supposons que pour tout $x \in \mathbb{R}^n$, $\nabla^2 f(x)$ est semi définie positive.

Soit $x \in \mathbb{R}^n$. On sait que pour tout $y \in \mathbb{R}^n$, on a

$$f(y) = f(x) + \langle \nabla f(x), y-x\rangle + \frac{1}{2}\langle \nabla^2 f(z)(y-x), y-x\rangle,$$

avec $z \in]x, y[$. Comme par hypothèse la matrice $\nabla^2 f(z)$ est semi définie positive, alors on a

$$\langle \nabla^2 f(z)(y-x), y-x\rangle \geq 0.$$

Ce qui implique que

$$f(y) \geq f(x) + \langle \nabla f(x), y-x\rangle.$$

Par suite la fonction f est convexe. □

Le résultat qui suit concerne la stricte convexité.

Théorème 3.2.5 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable. Si*

$$\forall\, x \in \mathbb{R}^n, \quad \langle \nabla^2 f(x)h, h\rangle > 0 \quad \forall\, h \in \mathbb{R}^n, h \neq 0,$$

alors f est strictement convexe sur $\mathbb{R}^n$.

Preuve : Soit x, $y \in \mathbb{R}^n$ avec $x \neq y$. On a

$$f(y) = f(x) + \langle \nabla f(x), y-x\rangle + \frac{1}{2}\langle \nabla^2 f(z)(y-x), y-x\rangle,$$

avec $z \in]x, y[$. Comme par hypothèse la matrice $\nabla^2 f(z)$ est définie positive, alors on a

$$\langle \nabla^2 f(z)(y-x), y-x\rangle > 0$$

car $x \neq y$. On obtient alors

$$f(y) > f(x) + \langle \nabla f(x), y-x\rangle.$$

Ce qui signifie que f est strictement convexe. □

Remarque 3.2.4 *Il faut signaler que la réciproque de ce résultat n'est pas vraie. On peut considérer la fonction φ définie sur $\mathbb{R}$ suivante : $\varphi(t) = t^4$. Cette fonction est strictement convexe mais sa dérivée seconde en* 0 *est nulle.*

On a ici une caractérisation de la forte convexité.

Théorème 3.2.6 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable. On a les équivalences suivantes :*
1) f est fortement convexe de module $r > 0$ sur $\mathbb{R}^n$.
2) Pour tout $x \in \mathbb{R}^n$, $\langle \nabla^2 f(x)h, h\rangle \geq r\|h\|^2 \ \forall\, h \in \mathbb{R}^n$.

Preuve : Supposons f fortement convexe de module r. Alors comme f est différentiable, on a pour tout x, $y \in \mathbb{R}^n$, on a

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{r}{2}\|y - x\|^2, \ \forall\, x, y \in \mathbb{R}^n.$$

Soit $x \in \mathbb{R}^n$; pour tout $h \in \mathbb{R}^n$ et t suffisamment petit, on a

$$f(x + th) = f(x) + t\langle \nabla f(x), h\rangle + \frac{t^2}{2}\langle \nabla^2 f(x)h, h\rangle + t^2\|h\|^2\varepsilon(t).$$

avec ε continue et tendant vers 0 quand t tend vers 0. Donc

$$f(x) + t\langle \nabla f(x), h\rangle + \frac{t^2}{2}\langle \nabla^2 f(x)h, h\rangle + t^2\|h\|^2\varepsilon(t) \geq f(x) + t\langle \nabla f(x), h\rangle + \frac{r}{2}t^2\|h\|^2.$$

Soit

$$\frac{1}{2}\langle \nabla^2 f(x)h, h\rangle + \|h\|^2\varepsilon(t) \geq \frac{r}{2}\|h\|^2.$$

En passant à la limite, on obtient $\langle \nabla^2 f(x)h, h\rangle \geq r\|h\|^2$.

Réciproquement, supposons que pour tout $x \in \mathbb{R}^n$,

$$\langle \nabla^2 f(x)h, h\rangle \geq r\|h\|^2 \ \forall\, h \in \mathbb{R}^n.$$

On a pour tout x, $y \in \mathbb{R}^n$, la relation

$$f(y) = f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2}\langle \nabla^2 f(z)(y - x), y - x\rangle,$$

avec $z \in]x, y[$. Donc on a

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2}r\|y - x\|^2.$$

Ce qui signifie que f est fortement convexe de module r sur $\mathbb{R}^n$. $\square$

Bibliographie

[1] Bazaraa, Mokhtar S. and Shetty, C. M., 1979. Nonlinear Programming Theory and Algorithms, *John Wiley and Sons.*

[2] Bazaraa, Mokhtar S. and Shetty, C. M., 1976. Foundations of Optimization, *Lecture Notes in Economic and Mathematical Systems, No 122, Springer-Verlag New-York.*

[3] Bergounioux Maïtine, 2001. Optimisation et Conrôle des systèmes linéaires, *Dunod.*

[4] Bertsekas, Dimiri P. 1995. Non Linear Programming, *Athena Scientific.*

[5] Bonnans, J. Fréderic and Shapiro, Alexander 2000. Perturbations Analysis of Opimization Problems, *Springer.*

[6] Culioli, Jean-Christophe, 1994. Introduction à l'optimisation, *Ellipses.*

[7] Hiriart-Urruty, Jean-Baptiste, 1998. Optimisation et Analyse Convexe, *Presse Universitaire de France.*

[8] Hiriart-Urruty, Jean-Baptiste and Lemaréchal, Claude, 1993. Convex Analysis and Minimization algorithms, *Vol I et II Grundlehren der mathematichen Wissenschaften 305 and 306, Springer-Verlag.*

[9] Hiriart-Urruty, Jean-Baptiste, 1996. L'Optimisation, in collection "Que sais-je ?", *Presse Universitaire de France.*

[10] Minoux, Michel, 1983. Programmation mathématique : Théorie et Algorithmes, Vol I, *Dunod.*

[11] Rockafellar, R. Tyrrel, 1970. Convex Analysis, *Princeton University Press, Princeton N. J..*

[12] Roberts, A. Wayne and Varberg, Dale E., 1973. Convex Functions, *Academic Press.*