Module A1
Éléments de topologie

Sauf mention contraire, $\mathcal{X}$ (resp. $\mathcal{Y}$) est un espace de Hilbert, muni du produit scalaire noté $\langle \cdot, \cdot \rangle$ et on note $\| \cdot \|$ la norme qui découle du produit scalaire.

L'objectif de ce module est d'introduire brièvement des résultats de topologie qui nous seront utiles dans la suite du cours pour l'analyse du comportement des méthodes d'optimisation du premier ordre.

1 Droite réelle achevée

Dans ce cours, pour des raisons de lisibilité, au lieu de considérer un couple de fonction $\tilde{f}$ et de son domaine de définition $\mathcal{D} \subset \mathcal{X}$, on va considérer des fonctions f dont le domaine de définition est défini de manière *implicite*, en "prolongeant" la définition de $\tilde{f}$ sur l'espace $\mathcal{X}$ par la valeur $+\infty$; plus précisément, on va remplacer la fonction $\tilde{f} : \mathcal{D} \to \mathbb{R}$ par la fonction $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ définie par

$$\forall\, x \in \mathcal{X}, \qquad f(x) = \begin{cases} \tilde{f}(x) & \text{si } x \in \mathcal{D} \\ +\infty & \text{sinon} \end{cases}$$

Cette opération implique de considérer la définition suivante :

Définition 1 (Droite réelle achevée)

On appelle *droite réelle achevée* l'ensemble suivant, noté $\overline{\mathbb{R}}$:

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\} = [\,-\infty\,;+\infty\,]$$

> Puisque le "prolongement" d'une fonction non définie initialement sur tout l'espace $\mathcal{X}$ se fait en attribuant aux points non définis une image infinie, on pourrait se contenter de considérer l'ensemble $\mathbb{R} \cup \{+\infty\} =]\,-\infty\,;+\infty\,]$. Cependant, dans le module **A6 : Dualité min-max**, on verra qu'exceptionnellement, il peut être utile d'attribuer à certains points l'image $-\infty$.

En adjoignant à la droite réelle les deux valeurs infinies $+\infty$ et $-\infty$, on est amené à étendre les opérations élémentaires (règles de calcul) à ces deux points supplémentaires. Certaines sont simples à définir ; on commence par exemple avec la somme d'une valeur infinie avec un scalaire :

$$\forall\, \beta \in \mathbb{R}, \qquad (+\infty) + \beta = +\infty \qquad \text{et} \qquad (-\infty) + \beta = -\infty$$

(la valeur finie est "absorbée" par la valeur infinie). Le produit d'une valeur infinie avec un scalaire non nul est également simple à définir : pour tout $\alpha \neq 0$,

$$\alpha \times (+\infty) = \begin{cases} +\infty & \text{si } \alpha > 0 \\ -\infty & \text{si } \alpha < 0 \end{cases} \qquad \text{et} \qquad \alpha \times (-\infty) = \begin{cases} -\infty & \text{si } \alpha > 0 \\ +\infty & \text{si } \alpha < 0 \end{cases}$$

Le cas du produit avec 0 est plus délicat ; on songera par exemple aux fonctions $f : x \mapsto x$ et $g : x \mapsto 1/x$, pour lesquelles on a

$$f(0^+) = 0, \qquad g(0^+) = +\infty \qquad \text{et} \qquad f(0^+) \times g(0^+) = 1$$

car $f \times g = 1$, alors que

$$f(0^+) = 0, \qquad \big(g(0^+)\big)^2 = +\infty \qquad \text{et} \qquad f(0^+) \times \big(g(0^+)\big)^2 = +\infty$$

car $(f \times g^2)(x) = 1/x$ quel que soit $x \in \mathbb{R}$. On voit donc qu'il n'est pas raisonnable d'attribuer une valeur unique à ce produit. Concernant les opérations entre valeurs infinies, on a d'une part

$$(+\infty) + (+\infty) = +\infty \qquad \text{et} \qquad (-\infty) + (-\infty) = -\infty$$

et d'autre part

$$(+\infty) \times (+\infty) = (-\infty) \times (-\infty) = +\infty \qquad \text{et} \qquad (+\infty) \times (-\infty) = (-\infty) \times (+\infty) = -\infty$$

Il n'est en revanche pas non plus raisonnable de définir la somme $(+\infty) + (-\infty)$ (ou, de manière équivalente, $+\infty - (+\infty)$) ; en effet, si l'on considère la fonction $g : x \mapsto 1/x$, on a par exemple

$$g(0^+) - g(0^+) = 0 \qquad \text{et} \qquad \big(g(0^+)\big)^2 - g(0^+) = +\infty$$

(car $g(x) - g(x) = 0$ et $(g(x))^2 - g(x) = (1 - x)/x^2$). On peut résumer ces observations en introduisant les *conventions* suivantes :

Définition 2 (Lois de composition sur $\overline{\mathbb{R}}$)

On peut étendre la définition de la somme et du produit usuels dans $\mathbb{R}$ sur la droite réelle achevée en posant pour la somme

$$\forall\, \beta \in \,]-\infty\,;+\infty\,], \qquad (+\infty) + \beta = +\infty$$
$$\forall\, \beta \in [-\infty\,;+\infty\,[, \qquad (-\infty) + \beta = -\infty$$

et pour le produit

$$\forall\, \alpha \in \,]\,0\,;+\infty\,], \qquad \alpha \times (+\infty) = +\infty \qquad \text{et} \qquad \alpha \times (-\infty) = -\infty$$
$$\forall\, \alpha \in [-\infty\,;0\,[, \qquad \alpha \times (+\infty) = -\infty \qquad \text{et} \qquad \alpha \times (-\infty) = +\infty$$

ces deux lois de composition internes restant commutatives.

> Il n'est donc pas possible de munir de manière raisonnable la droite réelle achevée d'une structure de groupe ou de corps.

On revient maintenant au cas des fonctions dont on étend la “définition” sur l'espace entier $\mathcal{X}$. Puisque les fonctions que nous étudierons seront maintenant définies sur $\mathcal{X}$, il faut distinguer le domaine de définition formel (ici $\mathcal{X}$) de l'ensemble des points sur lequel elles prennent une valeur finie (l'ensemble de définition classique).

Définition 3 (Domaine d'une fonction)

Soit $f : \mathcal{X} \to \overline{\mathbb{R}}$ une fonction. On appelle *domaine* (ou *domaine effectif*) de f l'ensemble

$$\operatorname{dom} f = \Big\{x \in \mathcal{X} \mid f(x) \notin \{+\infty, -\infty\}\Big\}$$

> **Attention !** Dorénavant, dans ce cours, on ne précisera plus l'ensemble de définition des fonctions considérées, celui-ci étant implicitement donné par le domaine de la fonction. En particulier, une fonction $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ semble par défaut "définie" partout, mais c'est un leurre car elle peut prendre en certains points une valeur infinie. En revanche, une fonction $f : \mathcal{X} \to \mathbb{R}$ est explicitement à valeurs **finies** (dans $\mathbb{R}$). Seules de telles fonctions peuvent être continues ou différentiables (sous-entendu : sur tout l'espace).

Considérons un premier exemple de fonctions pouvant prendre une valeur infinie :

Définition 4 (Indicatrice d'un ensemble)

Soit $\mathcal{A} \subset \mathcal{X}$. On appelle *indicatrice* de $\mathcal{A}$ la fonction $\chi_{\mathcal{A}} : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ définie par

$$\forall\, x \in \mathcal{X}, \qquad \chi_{\mathcal{A}}(x) = \begin{cases} 0 & \text{si } x \in \mathcal{A} \\ +\infty & \text{si } x \notin \mathcal{A} \end{cases}$$

> Attention ! Il faut bien distinguer la fonction indicatrice de la *fonction caractéristique* de l'ensemble $\mathcal{A}$
>
> $$\forall\, x \in \mathcal{X}, \qquad \mathbb{1}_{\mathcal{A}}(x) = \begin{cases} 1 & \text{si } x \in \mathcal{A} \\ 0 & \text{si } x \notin \mathcal{A} \end{cases}$$
>
> Notons que $\operatorname{dom} \chi_{\mathcal{A}} = \mathcal{A}$ mais que $\operatorname{dom} \mathbb{1}_{\mathcal{A}} = \mathcal{X}$.

D'autres opérations sur la droite réelle achevée nous seront utiles dans ce cours. Il s'agit tout d'abord de la relation d'ordre

$$\forall\, \alpha \in \mathbb{R}, \qquad -\infty < \alpha < +\infty$$

de sorte que $+\infty$ (resp. $-\infty$) est le plus grand (resp. le plus petit) élément de $\overline{\mathbb{R}}$. Tout ensemble de la droite réelle achevée admet une borne supérieure et une borne inférieure dans $\overline{\mathbb{R}}$; plus précisément, si $\mathcal{E} \subset \overline{\mathbb{R}}$ est non vide et différent du singleton $\{-\infty\}$ (resp. $\{+\infty\}$), alors

$$\sup \mathcal{E} \in \mathbb{R} \cup \{+\infty\} \qquad (\text{resp.} \qquad \inf \mathcal{E} \in \mathbb{R} \cup \{-\infty\})$$

tandis que $\quad \sup \emptyset = \sup\{-\infty\} = -\infty \quad$ et $\quad \inf \emptyset = \inf\{+\infty\} = +\infty$

Remarque : Puisque les bornes supérieure et inférieure d'une fonction $f : \mathcal{X} \to \overline{\mathbb{R}}$ sur un ensemble $\mathcal{E} \subset \mathcal{X}$ sont définies comme

$$\sup_{x \in \mathcal{E}} f(x) = \sup \Big\{ f(x) \mid x \in \mathcal{E} \Big\} \qquad \text{et} \qquad \inf_{x \in \mathcal{E}} f(x) = \inf \Big\{ f(x) \mid x \in \mathcal{E} \Big\}$$

l'ensemble $\mathcal{E}$ n'étant pas nécessairement inclus dans le domaine de f, les ensembles considérés au-dessus peuvent contenir $+\infty$ ou $-\infty$. Autrement dit, contrairement au cas réel ordinaire, il est techniquement possible d'avoir

$$\sup_{x \in \mathcal{E}} f(x) = -\infty \qquad \text{et} \qquad \inf_{x \in \mathcal{E}} f(x) = +\infty$$

même lorsque $\mathcal{E}$ est non vide.

Dans le cas des fonctions à valeurs sur la droite réelle achevée, il est nécessaire d'adapter la définition de la caractérisation séquentielle de la continuité en un point de façon à lever toute ambiguïté. En effet, si l'on considère la fonction $\chi_{[\,0\,;1\,]}$ comme "prolongement" de la restriction de la fonction nulle au segment $[0;1]$ (que l'on notera $\tilde{f}$), la définition usuelle de la continuité assure que $\tilde{f}$ est continue en 0. En revanche, si l'on considère la fonction $\chi_{[0;1]}$ dont l'ensemble de définition est formellement tout $\mathbb{R}$, la caractérisation séquentielle de la continuité assure que $\chi_{[0;1]}$ n'est pas continue en 0 ; en effet, selon que l'on considère la suite $(1/k)_{k\in\mathbb{N}^*}$ ou $(-1/k)_{k\in\mathbb{N}^*}$, toutes deux de limite nulle, on aura

$$\lim_{k\to+\infty} \chi_{[0;1]}\left(\frac{1}{k}\right) = \lim_{k\to+\infty} 0 = 0 \qquad \text{et} \qquad \lim_{k\to+\infty} \chi_{[0;1]}\left(-\frac{1}{k}\right) = \lim_{k\to+\infty} +\infty = +\infty$$

Cette incohérence n'est pas acceptable, car $\chi_{[0;1]} : \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ et $\tilde{f} : [0;1] \to \mathbb{R}$ sont censées définir le même objet. On est donc amené à considérer la définition modifiée suivante :

Définition 5 (Fonction continue)

Soit $f : \mathcal{X} \to \overline{\mathbb{R}}$ une fonction. Soit $x^0 \in \operatorname{dom} f$. On dit que f est *continue* en x^0 si pour toute suite $(x_k)_{k\in\mathbb{N}}$ de $\operatorname{dom} f$, on a

$$\lim_{k\to+\infty} x_k = x^0 \qquad \lim_{k\to+\infty} f(x_k) = f(x^0)$$

On dit que f est *continue sur son domaine* si f est continue en tout point de son domaine. On dit que f est *continue* si f est continue sur $\mathcal{X}$.

> Ainsi, dans ce cours (par convention), toute fonction continue (et *a fortiori* différentiable) ne peut prendre de valeur infinie. En particulier, si $\mathcal{A} \neq \mathcal{X}$, la fonction indicatrice $\chi_{\mathcal{A}}$ **n'est pas** continue (ni différentiable).

2 Semi-continuité inférieure

Dans cette section, on va introduire une notion plus faible que la continuité afin de conserver des hypothèses raisonnables sur des fonctions à valeurs sur $\overline{\mathbb{R}}$.

2.1 Définition et exemples

Définition 6 (Fonction semi-continue inférieurement)

Soit $f : \mathcal{X} \to \overline{\mathbb{R}}$ une fonction. Soit $x^0 \in \mathcal{X}$. On dit que f est *semi-continue inférieurement* en x^0, abrégé en *s.c.i.* en x^0, si pour toute suite $(x_k)_{k\in\mathbb{N}} \in \mathcal{X}^{\mathbb{N}}$,

$$\lim_{k\to+\infty} x_k = x^0 \qquad \Longrightarrow \qquad \liminf_{k\to+\infty} f(x_k) \geq f(x^0)$$

On dit que f est *s.c.i.* si elle est s.c.i. en tout point $x^0 \in \mathcal{X}$.

Cette définition est connue sous le nom de *caractérisation séquentielle* de la semi-continuité inférieure. Notons qu'elle ne fait pas appel au domaine de f. Elle est équivalente à :

$$\liminf_{x\to x^0} f(x) \geq f(x^0)$$

On rappelle que
$$\liminf_{k\to+\infty} u_k = \sup_{k\in\mathbb{N}} \left\{ \inf_{n\geq k} u_n \right\}$$

et que
$$\liminf_{x \to x^0} f(x) = \lim_{t \to 0^+} \inf \left\{ f(x) \mid 0 < \|x - x^0\| < t \right\}$$

EXEMPLE

Exemple d'une fonction s.c.i. Considérons la fonction réelle définie par

$$\forall\, x \in \mathbb{R}, \qquad f(x) = \begin{cases} x+1 & \text{si } x > 0 \\ x^2 & \text{si } x \leq 0 \end{cases}$$

Son graphe est représenté en figure 1. Cette fonction est continue, donc s.c.i., en tout point $\mathbb{R} \setminus \{0\}$. En 0, on a

$$f(0) = 0 \qquad \text{et} \qquad \forall\, x \neq 0, \qquad f(x) \in \{0, 1\}$$

En particulier, on a $f(x) \geq 0$. Il s'ensuit que pour toute suite $(x_k)_{k \in \mathbb{N}}$ de limite nulle,

$$\liminf_{k \to +\infty} f(x_k) \geq 0 = f(0)$$

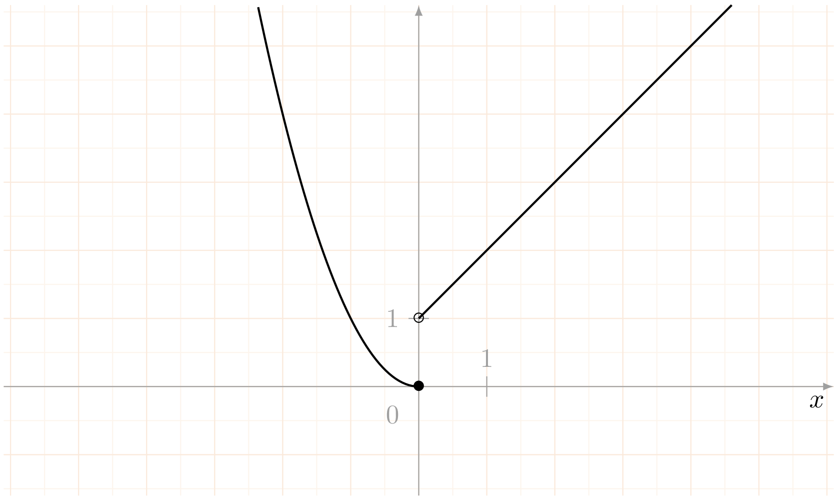

FIGURE 1 – Exemple de fonction s.c.i. non continue.

Comme on l'a vu dans l'exemple précédent, une fonction s.c.i. n'est pas nécessairement continue. On a cependant immédiatement le résultat suivant lorsque f ne prend pas la valeur $-\infty$:

Proposition 1

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction. Soit $x^0 \in \operatorname{dom} f$. On suppose que f est continue en x^0. Alors f est s.c.i. en x^0.

DÉMONSTRATION : Soit $(x_k)_{k \in \mathbb{N}}$ une suite de points de $\mathcal{X}$ convergeant vers x^0. Puisque $\mathcal{X} = \operatorname{dom} f \cup {}^c\operatorname{dom} f$, deux cas de figure sont possibles.

- **On peut extraire de la suite $(x_k)_{k \in \mathbb{N}}$ une sous-suite dans $\operatorname{dom} f$.** Dans ce cas, on a pour tout $n \in \mathbb{N}$

$$\liminf_{k \to +\infty} f(x_k) = \inf_{n \geq k} f(x_n) = \inf_{\substack{n \geq k \\ x_n \in \operatorname{dom} f}} f(x_n) = \lim_{\substack{k \to +\infty \\ x_k \in \operatorname{dom} f}} f(x_k) = f(x^0)$$

où la limite est une conséquence de la continuité de f en x^0 ; en passant à la borne supérieure sur les $n \in \mathbb{N}$, on obtient

$$\sup_{n\in\mathbb{N}} \inf_{n\geq k} f(x_n) = f(x^0)$$

- **On ne peut pas extraire de la suite** $(x_k)_{k\in\mathbb{N}}$ **une sous-suite dans** dom f**.** Dans ce cas, il existe un rang $k^0 \in \mathbb{N}$ à partir duquel les x_k n'appartiennent pas à dom f. Il s'ensuit puisque f ne prend pas la valeur $-\infty$ que

$$\liminf_{k\to+\infty} f(x_k) = \lim_{k\to+\infty} f(x_k) = +\infty > f(x^0)$$

Dans les deux cas, on a démontré que f est s.c.i. en x^0. ■

Remarque : Lorsque f est continue en x^0, on a pour toute suite $(x_k)_{k\in\mathbb{N}}$ de dom f

$$\lim_{k\to+\infty} x_k = x^0 \qquad \Longrightarrow \qquad \lim_{k\to+\infty} f(x_k) = f(x^0)$$

Supposons à présent que f ne soit que s.c.i. en x^0. Alors si $(x_k)_{k\in\mathbb{N}}$ est une suite de dom f convergeant vers x^0, on a par définition du domaine $f(x_k) \in \mathbb{R}$ pour tout $k \in \mathbb{N}$. Soit $(x_{k_j})_{j\in\mathbb{N}}$ une sous-suite telle que la suite $(f(x_{k_j}))_{j\in\mathbb{N}}$ soit convergente, de limite ℓ. On peut vérifier que

$$\liminf_{j\to+\infty} f(x_{k_j}) = \lim_{j\to+\infty} f(x_{k_j}) = \ell$$

de sorte que

$$\ell \geq f(x^0)$$

Un exemple important est celui des indicatrices d'ensembles fermés.

Proposition 2

Soit $\mathcal{A} \subset \mathcal{X}$ un ensemble **fermé**. Alors la fonction $\chi_\mathcal{A}$ est s.c.i.

Démonstration : Soit $x^0 \in \mathcal{X}$ et soit $(x_k)_{k\in\mathbb{N}}$ une suite de points de $\mathcal{X}$ convergeant vers x^0. Puisque $\mathcal{X} = \mathcal{A} \cup {}^c\mathcal{A}$, on peut considérer deux cas de figure :

- Si $x^0 \in \mathcal{A}$, alors on a

$$\chi_\mathcal{A}(x^0) = 0 \qquad \text{et} \qquad \forall\, k \in \mathbb{N}, \quad \chi_\mathcal{A}(x_k) \in \{0; +\infty\}$$

ce qui implique en particulier que

$$0 = \chi_\mathcal{A}(x^0) \leq \liminf_{n\to+\infty} \chi_\mathcal{A}(x_k) \in \{0; +\infty\}$$

La fonction $\chi_\mathcal{A}$ est donc semi-continue inférieurement en tout $x^0 \in \mathcal{A}$.

- Si $x^0 \notin \mathcal{A}$, alors on peut vérifier que l'ensemble

$$\Big\{k \in \mathbb{N} \mid x_k \in \mathcal{A}\Big\}$$

est fini. En effet, dans le cas contraire, on pourrait extraire une sous-suite convergente de $(x_k)_{k\in\mathbb{N}}$ dont tous les termes appartiennent à $\mathcal{A}$. Puisque $\mathcal{A}$ est supposé fermé, il s'ensuivrait que la limite de cette sous-suite appartient également à $\mathcal{A}$; par unicité de la limite, cela contredit l'hypothèse considérée dans ce cas de figure. Ainsi, il existe un indice $k_0 \in \mathbb{N}$ tel que

$$\forall\, k \geq k_0, \qquad x_k \notin \mathcal{A}$$

de sorte que, pour tout $k \geq k_0$, on a $\chi_\mathcal{A}(x_k) = +\infty$. Par conséquent,

$$\liminf_{n\to+\infty} \chi_\mathcal{A}(x_k) = +\infty = \chi_\mathcal{A}(x^0)$$

ce qui assure la semi-continuité inférieure de $\chi_\mathcal{A}$ en tout point $x^0 \notin \mathcal{A}$. ■

On peut généraliser ce résultat aux fonctions continues sur leur domaine fermé :

Proposition 3

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction. On suppose que $\operatorname{dom} f$ est **fermé** et que f est **continue sur son domaine**. Alors f est s.c.i.

DÉMONSTRATION : Commençons par remarquer que, pour tout $x^0 \in \operatorname{dom} f$, la fonction f est continue en x^0, donc s.c.i. en x^0 (proposition 1). On s'intéresse donc aux points $x^0 \notin \operatorname{dom} f$. Par hypothèse sur le domaine de f, son complémentaire est ouvert, donc il existe $\varepsilon > 0$ tel que

$$\forall\, x \in \mathcal{X}, \qquad \|x - x^0\| \leq \varepsilon \qquad \Longrightarrow \qquad x \notin \operatorname{dom} f$$

Autrement dit, pour tout tel x, on a $f(x) = +\infty$. Ainsi, pour toute suite $(x_k)_{k\in\mathbb{N}}$ une suite de points de $\mathcal{X}$ convergeant vers x^0, il existe un rang $k^0 \in \mathbb{N}$ tel que

$$\forall\, k \geq k^0, \qquad f(x_k) = +\infty \qquad \text{donc} \qquad \liminf_{k\to+\infty} f(x_k) = +\infty = f(x^0)$$

et f est s.c.i. en x^0. ■

2.2 Propriétés

Établissons quelques résultats sur les fonctions s.c.i. qui nous seront utiles dans la suite du cours. Notons que la notion de semi-continuité **inférieure** est particulièrement adaptée aux fonctions à valeurs dans $\mathbb{R} \cup \{+\infty\}$.

Proposition 4

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction. Alors on a équivalence entre les deux énoncés suivants :

(i) f est s.c.i. ;

(ii) pour tout $t \in \mathbb{R}$, l'ensemble de niveau inférieur
$$\operatorname{niv}_{\leq t} f = \Big\{x \in \mathcal{X} \mid f(x) \leq t\Big\}$$
est fermé.

DÉMONSTRATION : On démontre séparément les deux sens de l'équivalence.

- **Sens direct.** Soit $t \in \mathbb{R}$ et $(x_k)_{k\in\mathbb{N}}$ une suite de $\operatorname{niv}_{\leq t} f$, c'est-à-dire que
$$\forall\, k \in \mathbb{N}, \qquad f(x_k) \leq t$$
On suppose que cette suite est convergente, de limite x^*. Puisque f est s.c.i.,
$$\liminf_{k\to+\infty} f(x_k) \geq f(x^*)$$
Or, par définition de $\liminf_{k\to+\infty} f(x_k)$, on a
$$\liminf_{k\to+\infty} f(x_k) \leq t$$
de sorte que $x^* \in \operatorname{niv}_{\leq t} f$. Ainsi, on prouve que cet ensemble est fermé.
- **Réciproque.** Soit $(x_k)_{k\in\mathbb{N}}$ une suite convergente de limite x^0. Par l'absurde, supposons que
$$\underbrace{\liminf_{k\to+\infty} f(x_k)}_{\in\, \mathbb{R}\cup\{-\infty\}} < f(x^0) \in \mathbb{R} \cup \{+\infty\}$$
En particulier, il existe $\ell \in \mathbb{R}$ tel que
$$\liminf_{k\to+\infty} f(x_k) < \ell < f(x^0)$$

Il s'ensuit qu'il existe $\varepsilon > 0$ tel que, pour tout $k \in \mathbb{N}$,

$$\inf_{n \geq k} f(x_n) \leq \ell - \varepsilon$$

Pour tout $k \in \mathbb{N}$, on construit par récurrence une sous-suite $(x_{k_k})_{k\in\mathbb{N}}$ telle que

$$f(x_{n_k}) \leq \ell - \frac{\varepsilon}{2}$$

L'ensemble de niveau inférieur $\ell - \varepsilon/2$ étant fermé, on a

$$\lim_{k\to+\infty} x_{n_k} = x^0 \in \mathrm{niv}_{\leq \ell-\varepsilon/2} f \qquad \text{soit} \qquad f(x^0) \leq \ell - \frac{\varepsilon}{2} < f(x^0)$$

ce qui est absurde. ■

On termine cette section avec un dernier résultat :

Définition 7 (Enveloppe supérieure)

Soit $\mathcal{I} \subset \mathbb{R}$ et soit $(f_k)_{k\in\mathcal{I}}$ une famille de fonctions $\mathcal{X} \to \mathbb{R} \subset \{+\infty\}$. On appelle *enveloppe supérieure* de la famille de fonctions $\{f_i\}_{i\in\mathcal{I}}$ la fonction définie par

$$\left\{ \begin{array}{ccc} \mathcal{X} & \to & \mathbb{R} \cup \{+\infty\} \\ x & \mapsto & \sup\limits_{i\in\mathcal{I}} f_i(x) \end{array} \right.$$

Proposition 5 (Enveloppe supérieure d'une famille de fonctions s.c.i)

Soit $(f_i)_{i\in\mathcal{I}}$ une famille de fonctions $\mathcal{X} \to \mathbb{R} \cup \{+\infty\}$. Soit $x^0 \in \mathcal{X}$. On suppose que, pour tout $i \in \mathcal{I}$, la fonction f_i est s.c.i. en x^0. Alors l'enveloppe supérieure de la famille de fonctions $\{f_i\}_{i\in\mathcal{I}}$ est s.c.i. en x^0.

Démonstration : Il suffit de remarquer que

$$\mathrm{niv}_{\leq t} f = \mathcal{X} \setminus \Big\{ x \in \mathcal{X} \mid f(x) > t \Big\} = \mathcal{X} \setminus \Big\{ x \in \mathcal{X} \mid \sup_{i\in\mathcal{I}} f_i(x) > t \Big\}$$

Or, on a

$$\Big\{ x \in \mathcal{X} \mid \sup_{k\in\mathcal{I}} f_i(x) > t \Big\} = \bigcup_{i\in\mathcal{I}} \Big\{ x \in \mathcal{X} \mid f_i(x) > t \Big\} = \bigcup_{i\in\mathcal{I}} \Big(\mathcal{X} \setminus \mathrm{niv}_{\leq t} f_i \Big)$$

où chaque ensemble de niveau inférieur est fermé d'après la proposition 4, on en déduit, puisque l'union d'ensembles ouverts est ouvert, que l'ensemble de niveau inférieur de f est fermé. ■

3 Convexité

Dans cette section, on rappelle les principaux résultats concernant les ensembles et fonctions convexes qui nous seront utiles dans ce cours. Pour plus de détails (en particulier les démonstrations), le lecteur est invité à consulter les modules **A3 : Fonctions convexes différentiables** (3MA261) et **B6 : Projection sur un convexe** (3MA261).

3.1 Ensembles convexes

On commence par quelques rappels sur les ensembles convexes. Les ensembles convexes sont les ensembles stables par combinaison convexe :

Définition 8 (Ensemble convexe)

Soit $\mathcal{C} \subset \mathcal{X}$. On dit que $\mathcal{C}$ est *convexe* si

$$\forall\,(x_1,x_2) \in \mathcal{C}^2, \forall\,\lambda \in [\,0\,;1\,], \qquad \lambda\,x_1 + (1-\lambda)\,x_2 \in \mathcal{C}$$

EXEMPLE

Quelques exemples. Les ensembles suivants sont convexes :

- les (sous-)espaces vectoriels ;
- les singletons ;
- les demi-espaces d'un espace vectoriel ;
- les boules (fermées ou ouvertes).

On va maintenant s'intéresser aux opérations préservant la convexité qui, combinées avec la liste de l'exemple ci-dessus, permettent de démontrer rapidement qu'un ensemble est convexe :

Proposition 6

Soit $\mathcal{I} \subset \mathbb{R}$ et $\mathcal{C}_i$ un ensemble convexe de $\mathcal{X}$ pour tout $i \in \mathcal{I}$. Soit $\mathcal{C}$ un ensemble convexe de $\mathcal{Y}$ et $A : \mathcal{X} \to \mathcal{Y}$ une application linéaire. Soit $(i,j) \in \mathcal{I}^2$. Alors

- la somme $\qquad \mathcal{C}_i + \mathcal{C}_j = \big\{ x_i + x_j \in \mathcal{X} \mid x_i \in \mathcal{C}_i \text{ et } x_j \in \mathcal{C}_j \big\}$
- le produit cartésien
$$\mathcal{C}_i \times \mathcal{C} = \big\{ (x,y) \in \mathcal{X} \times \mathcal{Y} \mid x \in \mathcal{C}_i \text{ et } y \in \mathcal{C} \big\}$$
- l'intersection $\qquad \bigcap_{i\in\mathcal{I}} \mathcal{C}_i = \big\{ x \in U \mid \forall\,i \in \mathcal{I}, \quad x \in \mathcal{C}_i \big\}$
- l'image directe $\qquad A(\mathcal{C}_i) = \Big\{ A\,x \in \mathcal{Y} \mid x \in \mathcal{C}_i \Big\}$
- l'image réciproque $\qquad A^{-1}(\mathcal{C}) = \Big\{ x \in \mathcal{X} \mid A\,x \in \mathcal{C} \Big\}$

sont des ensembles convexes.

REMARQUE : En revanche, l'union de deux ensembles convexes n'est pas toujours convexe (il suffit de considérer l'union de deux singletons disjoints).

EXERCICE

Rectangles du plan. Démontrer que tout rectangle du plan est convexe.

EXEMPLE

Contraintes d'égalité et d'inégalité affines. Soit $G : \mathcal{X} \to \mathbb{R}$ une fonction affine. Alors les ensembles suivants

$$\Big\{ x \in \mathcal{X} \mid G(x) = 0 \Big\} \qquad \text{et} \qquad \Big\{ x \in \mathcal{X} \mid G(x) \leq 0 \Big\}$$

sont convexes, comme images réciproques par la forme linéaire $x \mapsto \langle a, x \rangle$ du singleton $\{-b\}$ et de l'intervalle fermé $]-\infty\,;-b\,]$.

Intéressons-nous à l'intérieur d'un convexe. L'intérieur topologique d'un convexe est parfois vide ; on songera par exemple à la plupart des sous-espaces affines. Or, dans de nombreux énoncés relatifs aux ensembles convexes, il est nécessaire de considérer une forme d'intérieur (qui n'est pas nécessairement l'intérieur topologique usuel). Aussi, on introduit la notion suivante :

Définition 9 (Intérieur relatif d'un ensemble convexe)

Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe non vide. On appelle *intérieur relatif* de $\mathcal{C}$ l'ensemble noté $\mathrm{intr}(\mathcal{C})$ défini par

$$\mathrm{intr}(\mathcal{C}) = \Big\{ x \in \mathcal{C} \mid \exists\, V \text{ voisinage de } x \text{ tel que } (V \cap \mathrm{aff}(\mathcal{C})) \subset \mathcal{C} \Big\}$$

où $\mathrm{aff}(\mathcal{C})$ est le *sous-espace affine engendré par* $\mathcal{C}$, c'est-à-dire le plus petit sous-espace affine contenant $\mathcal{C}$.

> Dans la littérature anglophone, l'intérieur relatif de $\mathcal{C}$ est noté $\mathrm{ri}(\mathcal{C})$.

Il découle immédiatement de la définition que l'intérieur topologique est contenu dans l'intérieur relatif. Considérons un exemple moins trivial :

EXEMPLE

Intérieur relatif d'un segment. On considère le segment reliant l'origine et le point $(1, 0)$ dans le plan :

$$\mathcal{C} = [\,0\,;1\,] \times \{0\}$$

Cet ensemble est visiblement fermé. Son intérieur topologique est vide. En revanche, puisque $\mathrm{aff}(\mathcal{C}) = \mathbb{R} \times \{0\}$ l'axe des abscisses, on a $\mathrm{intr}(\mathcal{C}) = \,]\,0\,;1\,[\, \times \{0\}$. Autrement dit, il s'agit du même segment privé de ses deux extrémités.

Proposition 7 (Intérieur relatif en dimension finie)

On suppose que $\mathcal{X}$ est de dimension finie. Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe non vide. Alors $\mathrm{intr}(\mathcal{C})$ est non vide.

DÉMONSTRATION : Admis.

On termine ce paragraphe en rappelant un autres des intérêts des ensembles convexes en optimisation, à savoir la projection :

Définition 10 (Projection sur un ensemble)

Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble et $x^0 \in \mathcal{X}$. On appelle *projection* de x^0 sur $\mathcal{C}$ l'ensemble noté $\mathrm{proj}_{\mathcal{C}}(x^0)$ des points x^* de $\mathcal{C}$ vérifiant

$$\forall\, x \in \mathcal{C}, \qquad \|x^* - x^0\| \le \|x - x^0\|$$

Lorsque $\mathcal{C}$ est convexe, fermé et non vide, la projection bien définie :

Théorème 1 (Projection sur un ensemble convexe fermé non vide)

Soit $\mathcal{C} \subset \mathcal{X}$ est un ensemble convexe fermé non vide. Alors la projection de tout point x_0 sur $\mathcal{C}$ contient un unique élément.

DÉMONSTRATION : Cf. Module **B6 : Projection sur un convexe** (3MA261).

> La démonstration dont la référence est mentionnée ci-dessus utilise des arguments simples de convexité. Dans le module **B1 : Méthodes d'optimisation du premier ordre**, on verra une preuve alternative utilisant des arguments d'optimisation de plus haut niveau.

3.2 Fonctions convexes

On va à présent s'intéresser aux fonctions convexes, qui occupent une place importante en optimisation, comme on le verra tout au long de ce cours.

Définition 11 (Fonction convexe)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction. On dit que f est *convexe* si

$$\forall (x_1, x_2) \in \mathcal{X}^2, \forall \lambda \in [0;1], \quad f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

On dit par ailleurs que f est *propre* si son domaine est non vide.

> Noter qu'une fonction convexe ne peut prendre la valeur $-\infty$.

EXEMPLE

Quelques fonctions convexes définies sur un espace de HILBERT. Les fonctions suivantes sont convexes et propres :

- les formes linéaires et affines ;
- les normes ;
- les normes au carré.

REMARQUE : Le domaine d'une fonction convexe est nécessairement convexe. En effet, si x_1 et x_2 sont deux points du domaine d'une fonction convexe f, alors $f(x_1)$ et $f(x_2)$ sont finies par définition du domaine, tandis que par définition de la convexité, on a nécessairement $f(\lambda x_1 + (1-\lambda) x_2) < +\infty$, ce qui implique que toute combinaison convexe de x_1 et x_2 est dans le domaine de f.

Un exemple important de fonctions convexes est donné dans la proposition suivante :

Proposition 8 (Indicatrice d'un ensemble convexe)

Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe. Alors $\chi_{\mathcal{C}}$ est une fonction convexe. De plus, si $\mathcal{C}$ est non vide, alors $\chi_{\mathcal{C}}$ est propre.

REMARQUE : Ainsi, si $\mathcal{C}$ est convexe, non vide et fermé, $\chi_{\mathcal{C}}$ est convexe, propre et s.c.i.

DÉMONSTRATION : Il suffit de vérifier que $\operatorname{dom} \chi_{\mathcal{C}} = \mathcal{C}$. ■

De la même manière que pour les ensembles convexes, certaines opérations préservent la convexité des fonctions.

Proposition 9 (Combinaison linéaire)

Soit $f_1 : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ et $f_2 : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ deux fonctions convexes. Alors pour tout $\alpha_1 \geq 0$ et $\alpha_2 \geq 0$, la fonction $\alpha_1 f_1 + \alpha_2 f_2$ est convexe.

DÉMONSTRATION : Laissé au lecteur.

La proposition précédente peut être interprétée comme le fait que certaines compositions *externes* par une fonction linéaire de fonctions convexes sont convexes. Pour la composition *interne* par une application linéaire (ou affine), aucune restriction n'est nécessaire quant à la partie linéaire de A, comme le montre la proposition suivante :

Proposition 10 (Composition par une application affine)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe et $A : \mathcal{Y} \to \mathcal{X}$ une application affine. Alors la fonction $f \circ A$ est convexe.

DÉMONSTRATION : Il suffit d'utiliser la linéarité partielle de A et de remarquer que

$$b = \lambda\, b + (1-\lambda)\, b$$

pour tout $b \in \mathbb{R}$ et tout $\lambda \in [\,0\,;1\,]$. ■

La proposition 10 suggère que certaines compositions de fonctions convexes ne sont pas nécessairement convexes. C'est effectivement le cas : il suffit de considérer par exemple la fonction $-\exp$. De manière générale, si les fonctions en jeu ne sont pas affines, seule la composition *externe* par une fonction convexe croissante d'une fonction convexe est assurée d'être convexe, comme on peut le voir dans l'exemple suivant.

EXEMPLE

Composition d'une fonction convexe par une fonction convexe croissante. Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe et $g : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe croissante. Alors $g \circ f$ est convexe. En effet, si on suppose que son domaine est non vide (le résultat étant immédiat dans le cas contraire), alors on a pour tout $(x_1, x_2) \in \mathcal{X}^2$ et $\lambda \in [\,0\,;1\,]$

$$f\big(\lambda\, x_1 + (1-\lambda)\, x_2\big) \leq \lambda\, f(x_1) + (1-\lambda)\, f(x_2)$$

par convexité de f ; par croissance de g, on a alors

$$g \circ f\big(\lambda\, x_1 + (1-\lambda)\, x_2\big) \leq g\big(\lambda\, f(x_1) + (1-\lambda)\, f(x_2)\big)$$

On conclut à l'aide de la convexité de g.

On signale enfin une propriété qui nous sera très utile dans la suite de ce cours :

Proposition 11 (Enveloppe supérieure)

Soit $\mathcal{I} \subset \mathbb{R}$ et $f_i : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe pour tout $i \in \mathcal{I}$. Alors l'*enveloppe supérieure* de la famille de fonctions $\{f_i\}_{i\in\mathcal{I}}$ est une fonction convexe.

Démonstration : Il suffit d'écrire la définition de la convexité des f_i et de passer à la borne supérieure. ■

On termine ce paragraphe avec quelques considérations sur la continuité des fonctions convexes en dimension finie. On va admettre le résultat suivant :

Proposition 12

On suppose que $\mathcal{X}$ est de dimension finie. Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe. Soit $x^0 \in \operatorname{intr}(\operatorname{dom} f)$. Alors f est continue en x^0.

En particulier, toute fonction convexe de domaine ouvert est continue sur son domaine, puisque son intérieur relatif contient son intérieur topologique qui est son domaine. On rappelle par ailleurs qu'en dimension finie, tout ensemble convexe possède un intérieur relatif non vide, ce qui implique que l'ensemble des points sur lequel une fonction convexe est continue est non vide en dimension finie.

Démonstration : Admis. Pour le cas où $\operatorname{dom} f = \mathcal{X}$, la preuve se trouve dans le module **A3 : Fonctions convexes différentiables** (3MA261).

Remarque : Puisque l'intérieur topologique est contenu dans l'intérieur relatif, on peut remplacer dans la proposition précédente $\operatorname{intr}(\operatorname{dom} f)$ par $\operatorname{int}(\operatorname{dom} f)$. L'intérêt d'une telle manœuvre est de faciliter la vérification des hypothèses, car l'intérieur topologique est généralement plus simple à obtenir que l'intérieur relatif. Cependant, l'énoncé obtenu est plus faible, d'autant plus que, parfois, l'intérieur topologique du domaine est vide. Ainsi, l'énoncé original est utile lorsqu'il s'agit de justifier l'existence d'un point où f est continue (sans pouvoir nécessairement le déterminer), tandis que l'énoncé plus faible est plus exploitable lorsque l'intérieur topologique du domaine est non vide et qu'il est nécessaire de trouver de manière explicite un point où f est continue.

Il est important de considérer des points situés à l'intérieur relatif du domaine, comme le démontre le contre-exemple suivant :

Contre-exemple

Discontinuité au bord. Considérons la fonction $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ définie par

$$\forall\, x \in \mathbb{R}, \qquad f(x) = \begin{cases} 0 & \text{si } x > 0 \\ 1 & \text{si } x = 0 \\ +\infty & \text{si } x < 0 \end{cases}$$

Le graphe de cette fonction – discontinue en 0 – est affiché en figure 2. Vérifions qu'elle est convexe. L'inégalité définissant la convexité étant immédiatement vérifiée pour deux points x_1 et x_2 strictement positifs, il reste le cas où $x_1 = 0$ et $x_2 > 0$. Puisque pour tout $\lambda \in \,]\,0\,;1\,[$, on a $\lambda\, x_1 + (1-\lambda)\, x_2 > 0$, et que

$$f\big(\lambda\, x_1 + (1-\lambda)\, x_2\big) = 0 \leq \lambda = \lambda\, f(x_1) + (1-\lambda)\, f(x_2)$$

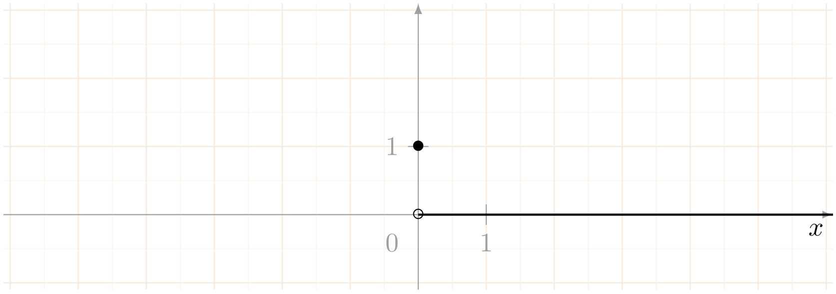

FIGURE 2 – Cas d'une fonction convexe discontinue au bord de son domaine.

3.3 Forte convexité

Parmi les fonctions convexes, deux familles de fonctions jouent un rôle particulier en optimisation :

Définition 12 (Fonction strictement convexe)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction. On dit que f est *strictement convexe* si

$$\forall\, x_1 \neq x_2 \in \operatorname{dom} f, \forall \lambda \in \,]\,0\,;1\,[\,, \quad f(\lambda\, x_1 + (1-\lambda)\, x_2) < \lambda\, f(x_1) + (1-\lambda)\, f(x_2)$$

Remarquons que, dans la définition ci-dessus, l'inégalité stricte est imposée aux points distincts dans le domaine.

Définition 13 (Fonction fortement convexe)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction et soit $\alpha > 0$. On dit que f est *fortement convexe de module* α si

$$\forall\, (x_1, x_2) \in \mathcal{X}^2, \forall \lambda \in [\,0\,;1\,]\,,$$
$$f\big(\lambda\, x_1 + (1-\lambda)\, x_2\big) \leq \lambda\, f(x_1) + (1-\lambda)\, f(x_2) - \frac{\alpha}{2}\, \lambda\, (1-\lambda)\, \|x_1 - x_2\|^2$$

On dit parfois que f est *α-convexe*. Il est aisé de vérifier qu'une fonction fortement convexe est strictement convexe.

EXERCICE

Module de forte convexité. Montrer que toute fonction fortement convexe de module α est fortement convexe de module α' pour tout $0 < \alpha' \leq \alpha$.

Considérons un premier exemple important de fonctions fortement convexes :

EXEMPLE

Distance au carré. Soit $x^0 \in \mathcal{X}$ et $\alpha > 0$. On pose

$$f : \left\{ \begin{array}{ccc} \mathcal{X} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{\alpha}{2}\, \|x - x^0\|^2 \end{array} \right.$$

Soit $(x_1, x_2) \in \mathcal{X}^2$ et $\lambda \in [\,0\,;1\,]$. En remarquant que $x^0 = \lambda\, x^0 + (1-\lambda)\, x^0$ et en développant le premier carré, on obtient que

$$\begin{aligned} &\frac{\alpha}{2}\,\big\|\lambda\, x_1 + (1-\lambda)\, x_2 - x^0\big\|^2 - \frac{\alpha}{2}\,\lambda\,\|x_1 - x^0\|^2 - \frac{\alpha}{2}\,(1-\lambda)\,\|x_2 - x^0\|^2 \\ &\qquad = -\frac{\alpha}{2}\,\lambda\,(1-\lambda)\big(\|x_1 - x^0\|^2 + \|x_2 - x^0\|^2 - 2\,\langle x_1 - x^0, x_2 - x^0\rangle\big) \end{aligned}$$

En utilisant une identité remarquable pour simplifier le terme entre parenthèses dans l'expression précédente, on prouve que f est une fonction fortement convexe de module α.

La définition de la forte convexité est en général difficile à vérifier. C'est pourquoi les résultats suivants sont utiles pour déterminer la forte convexité d'une fonction.

EXERCICE

Somme d'une fonction convexe et d'une fonction fortement convexe. Montrer que la somme d'une fonction convexe et d'une fonction fortement convexe est fortement convexe.

La réciproque est vraie, comme le montre la proposition suivante :

Proposition 13 (Caractérisation des fonctions fortement convexes)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction. Alors on a l'équivalence entre les énoncés suivants :

(i) f est fortement convexe de module α ;

(ii) pour tout $x^0 \in \mathcal{X}$, la fonction

$$x \mapsto f(x) - \frac{\alpha}{2}\,\|x - x^0\|^2$$

est une fonction convexe.

DÉMONSTRATION : Admis. La preuve se trouve dans le module **A3 : Fonctions convexes différentiables** (3MA261).

Autrement dit, en posant g la fonction apparaissant dans l'énoncé (ii) de la proposition 13, on assure l'existence d'une fonction convexe $g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ telle que

$$\forall\, x \in \mathcal{X}, \qquad f(x) = g(x) + \frac{\alpha}{2}\,\|x - x^0\|^2$$

c'est-à-dire que toute fonction fortement convexe est la somme de d'une fonction convexe et d'une fonction fortement convexe. Notons par ailleurs que f et g ont même domaine.

Pour aller plus loin

Enveloppe supérieure d'une famille de fonctions. Cette notion sera utilisée dans le contexte de la dualité min-max (module **A6 : Dualité min-max**). En particulier, on verra qu'il est possible de représenter toute fonction convexe comme l'enveloppe supérieure d'une famille de fonctions affines, ce qui conduira à la notion de conjuguée convexe (module **A8 : Dualité de Fenchel et conjuguée convexe**).

Généralisation de la notion de différentiabilité. Le gradient d'une fonction convexe possède des propriétés remarquables, que l'on préservera en définissant une généralisation de la différentielle aux fonctions convexes non différentiables, ou encore non convexes et non différentiables (module **A2 : Sous-différentiabilité**). Cette généralisation, appelée *sous-différentiel*, se révélera en particulier très utile pour caractériser les minima globaux d'une fonction convexe (module **B1 : Méthodes d'optimisation du premier ordre**). On verra par ailleurs qu'un des intérêts de l'optimisation convexe réside dans l'absence de minima locaux.

Forte convexité. La forte convexité est une propriété intéressante pour l'optimisation, car elle garantit à la fois l'existence et l'unicité du minimiseur (module **B1 : Méthodes d'optimisation du premier ordre**). De plus, elle permet d'assurer des comportements plus intéressants pour les algorithmes d'optimisation, tant en termes de stabilité qu'en termes de vitesse de convergence (module **B1 : Méthodes d'optimisation du premier ordre**). On verra en outre que la forte convexité est une forme de régularité dans le domaine dual (**A8 : Dualité de Fenchel et conjuguée convexe**).