PCSI - mathématiques

# Fonctions convexes

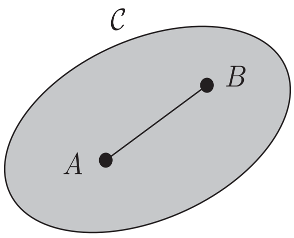
## 1 Convexes d'un $\mathbb{R}$-espace affine

Soit $\mathcal{E}$ un $\mathbb{R}$-espace vectoriel muni de sa structure affine canonique. Si $A$, $B \in \mathcal{E}$, le *segment* $[A, B]$ est

$$\begin{aligned}[A,B] &= \{(1-t)\,A + tB \mid t \in [0,1]\} \\ &= \left\{A + t\overrightarrow{AB} \mid t \in [0,1]\right\}.\end{aligned}$$

**Définition 1** *Une partie $\mathcal{C}$ de $\mathcal{E}$ est* convexe *si :*

$$\forall A \in \mathcal{C},\ \forall B \in \mathcal{C},\ [A,B] \subset \mathcal{C}.$$




**Exemple 1**

1. *$\mathcal{E}$, $\varnothing$, les singletons*
2. *les sva de $\mathcal{E}$ ;*
3. *les segments ;*
4. *les boules ouvertes (*resp. *fermées) (pour toute norme sur $\mathcal{E}$).*

## 2 Fonctions convexes

$I$ est un intervalle de $\mathbb{R}$ ; $f$ est une application de $I$ dans $\mathbb{R}$.

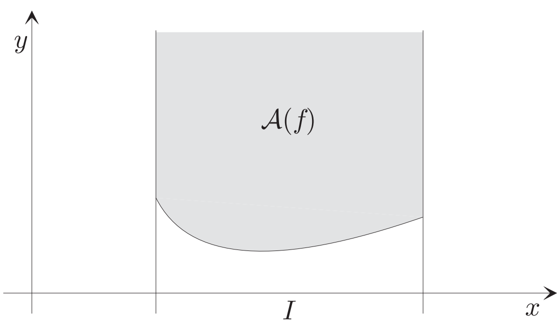
**Définition 2** *La fonction $f$ est* convexe *sur $I$ si pour tout couple $(x, y)$ de points de $I$ et pour tout $t \in [0, 1]$ on a :*

$$f\left((1-t)\,x + ty\right) \leq (1-t)\,f(x) + tf(y)\,.$$

Géométriquement, cela signifie que la partie

$$\mathcal{A}(f) \overset{\text{déf.}}{=} \{(x,y) \mid x \in I \text{ et } y \geq f(x)\}$$

est une partie convexe de $\mathbb{R}^2$ (au sens du § précédent).




**Remarque 1** *Pour prouver que $f$ est convexe sur $I$ il suffit d'établir que pour tout couple $(x, y)$ de points de $I$ tels que $x < y$ et pour tout réel $t$* de *$]0, 1[$ on a $f\left((1-t)\,x + ty\right) \leq (1-t)\,f(x) + tf(y)$.*

La définition se généralise à un nombre quelconque de points :

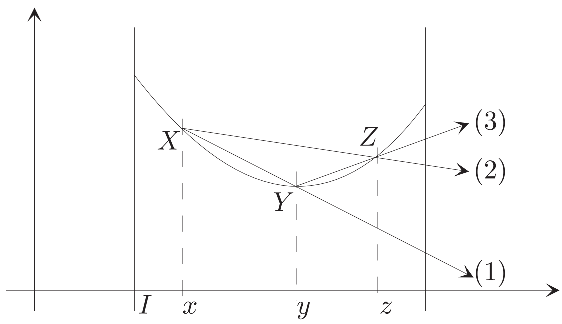
**Théorème 1** *On suppose que $f$ est convexe sur $I$. Si $n \in \mathbb{N}^*$ et si $x_1$, ..., $x_n \in I$ ; si $t_1$, ..., $t_n \in \mathbb{R}_+$ et si $\sum_{i=1}^{n} t_i = 1$ on a :*

$$f\left(\sum_{i=1}^{n} t_i\, x_i\right) \leq \sum_{i=1}^{n} t_i\, f(x_i)\,.$$

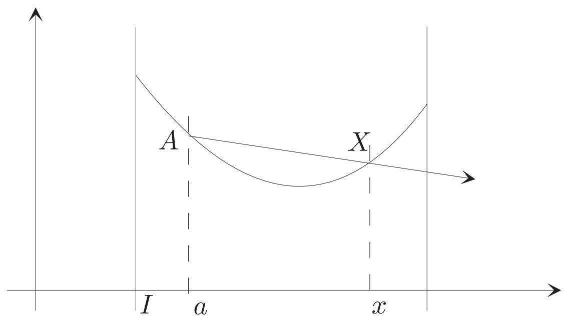
Ce dernier résultat est parfois appelé inégalité de Jensen.

**Théorème 2** *Il y a équivalence entre*

1. *$f$ est convexe sur $I$ ;*
2. *Pour tout triplet $(x, y, z)$ de points de $I$ tel que $x < y < z$, $\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} \leq \frac{f(z)-f(y)}{z-y}$ ;*
3. *Pour tout point $a$ de $I$ la fonction $q_a$ définie par $q_a(x) = \frac{f(x)-f(a)}{x-a}$ est croissante sur $I - \{a\}$.*

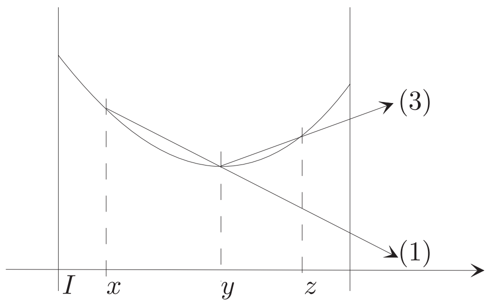



(2) pente $(XY) \leq$ pente $(XZ) \leq$ pente $(YZ)$




(3) $(x \mapsto \text{pente}\,(AX)) \nearrow$ sur $I - \{a\}$

**Remarque 2** *Pour démontrer que $f$ est convexe sur $I$ il suffit de prouver que pour tout triplet $(x,y,z)$ de points de $I$ on a $\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(y)}{z-y}$.*




## 2.1 Propriétés d'une fonction convexe sur un intervalle $I$

Une fonction convexe $f$ sur un intervalle $I$ admet des dérivées à gauche et à droite en tout point de $\mathring{I}$. Elle est donc continue en chaque tel point. Toutefois, attention !

**Remarque 3** *$f$ n'a aucune raison d'être continue en $\alpha = \inf I$ (*resp. *$\beta = \sup I$). La convexité exige seulement, dans ces cas : $f(\alpha) \geq \lim_{\alpha^+} f$ (*resp. *$f(\beta) \geq \lim_{\beta^-} f$).*

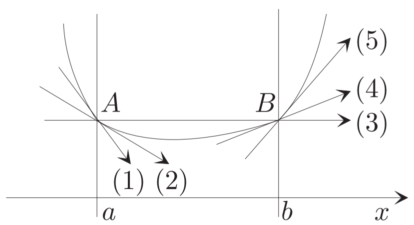
**Théorème 3** *Si $f$ est convexe sur $I$ et si $a \in \mathring{I}$, $f$ est dérivable à droite et à gauche en $a$ et :*

$$\sup_{\substack{x<a \\ x\in I}} \frac{f(x)-f(a)}{x-a} = f'_g(a) \leq f'_d(a) = \inf_{\substack{x>a \\ x\in I}} \frac{f(x)-f(a)}{x-a}.$$

Le th. suivant précise le rapport entre les dérivées à gauche et à droite en $a$ et $b$ $(a<b)$ :

**Théorème 4** *Si $f$ est convexe sur $I$, $f$ est continue sur $\mathring{I}$, dérivable à droite et à gauche sur $\mathring{I}$, les fonctions $f'_g$ et $f'_d$ sont croissantes sur $\mathring{I}$ et pour tout couple $(a,b)$ de points de $\mathring{I}$ tel que $a<b$ on a :*

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




Ainsi, les fonctions $f'_g$ et $f'_d$ sont croissantes sur $\mathring{I}$.

La réciproque est-elle vraie ? Cela nécessite une hypothèse complémentaire pour préciser la situation en les bornes de $I$ :

**Théorème 5** *Si $f : I \to \mathbb{R}$ est* continue *sur $I$ et dérivable à droite (*resp. *à gauche) sur $\mathring{I}$, $f$ est convexe sur $I$* ssi *la fonction $f'_d$ (*resp. *$f'_g$) est croissante sur $\mathring{I}$.*

L'hypothèse de continuité est plus forte que celle de la remarque 3. Ce résultat fournit plusieurs moyens pratiques de vérifier qu'une application est convexe.

**Corollaire 1**

1. *Si $f : I \to \mathbb{R}$ est continue sur $I$ et dérivable sur $\mathring{I}$, $f$ est convexe sur $I$* ssi *$f'$ est croissante sur $\mathring{I}$.*
2. *Si $f : I \to \mathbb{R}$ est continue sur $I$ et deux fois dérivable sur $\mathring{I}$, $f$ est convexe sur $I$* ssi *$f'' \geq 0$ sur $\mathring{I}$.*

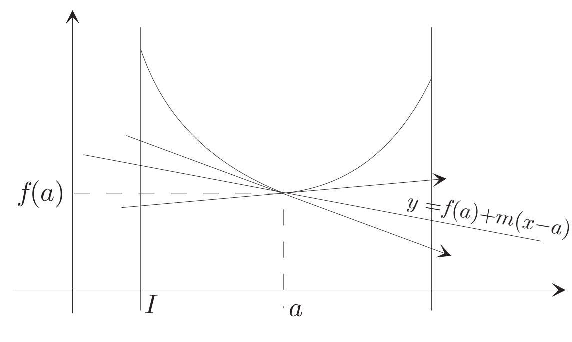
**Exemple 2** *obtenus en appliquant le corollaire 1. :*

1. *$x \mapsto x^2$ est convexe sur $\mathbb{R}$.*
2. *$x \mapsto \mathrm{e}^x$ est convexe sur $\mathbb{R}$.*
3. *Si $a > 1$, $x \mapsto a^x$ est convexe sur $\mathbb{R}$.*
4. *$x \mapsto \frac{1}{x}$ est convexe sur $\mathbb{R}_+^*$.*
5. *$x \mapsto -\ln x$ est convexe sur $\mathbb{R}_+^*$.*

*Les exemples 2.2. et 2.5. sont fondamentaux.*

On a enfin le théorème suivant, qui exprime que la courbe représentative d'une fonction convexe est au-dessus de toutes ses (demi-)tangentes :

**Théorème 6** *On suppose que $f$ est convexe sur $I$. Si $a \in \mathring{I}$ et si $m$ est un réel tel que $f'_g(a) \leq m \leq f'_d(a)$, on a pour tout $x \in I$ : $f(x) \geq f(a) + m(x-a)$.*




## 2.2 Fonctions concaves

**Définition 3** *$f : I \to \mathbb{R}$ est* concave *sur $I$ si $-f$ est convexe sur $I$.*

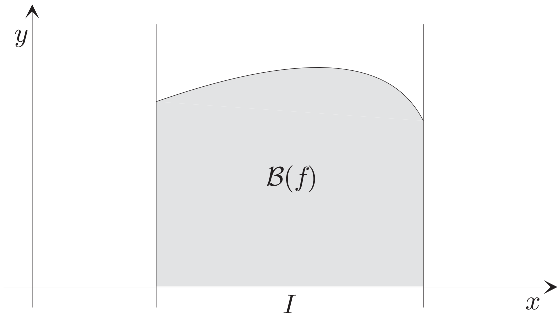
Cela signifie que pour tout couple $(x,y)$ de points de $I$ et pour tout $t \in [0,1]$ :

$$f((1-t)x + ty) \geq (1-t)f(x) + tf(y)$$

ou encore que

$$\mathcal{B}(f) \overset{\text{déf.}}{=} \left\{(x,y) \in \mathbb{R}^2 \,|\, x \in I \text{ et } y \leq f(x)\right\}$$

est une partie convexe de $\mathbb{R}^2$.




**Exemple 3** ln *est concave sur* $]0, +\infty[$.

Les fonctions affines sur $I$ sont les applications de la forme $f : I \to \mathbb{R}\,;\, x \mapsto \alpha x + \beta$ où $\alpha,\ \beta \in \mathbb{R}$.

**Théorème 7**

1. *Une fonction affine sur* $I$ *est convexe et concave sur* $I$.
2. *Si* $f : I \to \mathbb{R}$ *est une fonction convexe et concave sur* $I$, $f$ *est une fonction affine sur* $I$.

Ce deuxième point n'est pas entièrement trivial, car on ne suppose pas *a priori* que $f$ est deux fois dérivable sur $I$.

# 3 Relations de Hölder et de Minkowski

## 3.1 Cas des familles finies

La concavité de la fonction ln (ex. 3) va ici être poussée dans ses derniers retranchements.

**Lemme 1** *Soient* $\alpha,\ \beta \in \mathbb{R}$ *tels que* $\alpha,\ \beta \geq 0$, $\alpha + \beta = 1$. *Quels que soient* $x,\ y \in \mathbb{R}_+$ *on a :*

$$x^\alpha y^\beta \leq \alpha x + \beta y.$$

L'inégalité de Hölder consiste en la généralisation à un nombre quelconque de points :

**Théorème 8 (Hölder)** *Soient* $\alpha,\ \beta \in \mathbb{R}$ *tels que* $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta = 1$. *Si* $n \in \mathbb{N}^*$ *et si* $(a_i)_{1\leq i\leq n}$ *et* $(b_i)_{1\leq i\leq n}$ *sont des familles d'éléments de* $\mathbb{R}_+$ *on a :*

$$\sum\nolimits_{i=1}^n a_i^\alpha b_i^\beta \leq \left(\sum\nolimits_{i=1}^n a_i\right)^\alpha \left(\sum\nolimits_{i=1}^n b_i\right)^\beta.$$

Cet énoncé généralise l'inégalité de Schwarz de l'algèbre bilinéaire. On le voit facilement en modifiant légèrement l'écriture comme suit :

**Corollaire 2** *Si* $p,\ q \in \mathbb{R}$, $p > 0$, $q > 0$, $\frac{1}{p} + \frac{1}{q} = 1$ *; si* $n \in \mathbb{N}^*$ *et si* $(a_i)_{1\leq i\leq n}$ *et* $(b_i)_{1\leq i\leq n}$ *sont des familles d'éléments de* $\mathbb{R}_+$ *on a :*

$$\sum\nolimits_{i=1}^n a_i b_i \leq \left(\sum\nolimits_{i=1}^n a_i^p\right)^{\frac{1}{p}} \left(\sum\nolimits_{i=1}^n b_i^q\right)^{\frac{1}{q}}.$$

**Remarque 4** *Soient* $p,\ q \in \mathbb{R}$ *tels que* $p > 0$, $q > 0$, $\frac{1}{p} + \frac{1}{q} = 1$. *Si* $(x_i)_{1\leq i\leq n}$ *et* $(y_i)_{1\leq i\leq n}$ *sont des familles d'éléments de* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *on a :*

$$\left|\sum\nolimits_{i=1}^n x_i y_i\right| \leq \left(\sum\nolimits_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \left(\sum\nolimits_{i=1}^n |y_i|^q\right)^{\frac{1}{q}}.$$

Ce dernier énoncé, lorsque $\mathbb{K} = \mathbb{R}$ et $p = q = 2$, redonne l'inégalité de Schwarz.

Le théorème de Minkowski admet les mêmes généralisations :

**Théorème 9 (Minkowski)** *Soit* $p \in \mathbb{R}$, $p \geq 1$. *Si* $(a_i)_{1\leq i\leq n}$ *et* $(b_i)_{1\leq i\leq n}$ *sont des familles d'éléments de* $\mathbb{R}_+$ *on a :*

$$\left(\sum\nolimits_{i=1}^n (a_i + b_i)^p\right)^{\frac{1}{p}} \leq \left(\sum\nolimits_{i=1}^n a_i^p\right)^{\frac{1}{p}} + \left(\sum\nolimits_{i=1}^n b_i^p\right)^{\frac{1}{p}}.$$

**Corollaire 3** *Soit* $p \in \mathbb{R}$, $p \geq 1$. *Si* $(x_i)_{1\leq i\leq n}$ *et* $(y_i)_{1\leq i\leq n}$ *sont des familles d'éléments de* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *on a :*

$$\left(\sum\nolimits_{i=1}^n |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum\nolimits_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} + \left(\sum\nolimits_{i=1}^n |y_i|^p\right)^{\frac{1}{p}}.$$

Ceci redonne l'inégalité de Minkowski usuelle lorsque $\mathbb{K} = \mathbb{R}$ et $p = q = 2$.

En outre, remarquons que si l'on pose pour $x = (x_i)_{1\leq i\leq n} \in \mathbb{R}^n$, $p \in \mathbb{R}$, $p \geq 1$ :

$$\|x\|_p = \left(\sum\nolimits_{i=1}^n |x_i|^p\right)^{\frac{1}{p}},$$

alors $\|\cdot\|_p$ est une *norme* sur $\mathbb{R}^n$.

## 3.2 Cas des intégrales

$a,\ b \in \mathbb{R}$. Si $f : [a, b] \to \mathbb{R}$ est intégrable et si $f \geq 0$ (càd : $f(t) \geq 0$ pour tout $t$), $f^\alpha$ est intégrable sur $[a, b]$ donc intégrable. Le passage des inégalités du § 3.1. aux intégrales est assuré par le biais des sommes de Riemann.

**Théorème 10 (Hölder)** *Soient* $\alpha,\ \beta \in \mathbb{R}$ *tels que* $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta = 1$. *Si* $f,\ g : [a, b] \to \mathbb{R}$ *sont intégrables et positives sur* $[a, b]$ *on a :*

$$\int_{[a,b]} f^\alpha g^\beta \leq \left(\int_{[a,b]} f\right)^\alpha \left(\int_{[a,b]} g\right)^\beta.$$

**Corollaire 4** *Soient* $p,\ q \in \mathbb{R}$, $p > 0$, $q > 0$, $\frac{1}{p} + \frac{1}{q} = 1$. *Si* $f,\ g : [a, b] \to \mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *sont intégrables sur* $[a, b]$ *on a :*

$$\left|\int_{[a,b]} fg\right| \leq \left(\int_{[a,b]} |f|^p\right)^{\frac{1}{p}} \left(\int_{[a,b]} |g|^q\right)^{\frac{1}{q}}.$$

Lorsque $p = q = 2$ on retrouve l'inégalité de Schwarz pour les intégrales.

**Théorème 11 (Minkowski)** *Soit* $p \in \mathbb{R}$, $p \geq 1$. *Si* $f$, $g : [a, b] \to \mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *sont intégrables sur* $[a, b]$ *on a :*

$$\left(\int_{[a,b]} |f + g|^p\right)^{\frac{1}{p}} \leq \left(\int_{[a,b]} |f|^p\right)^{\frac{1}{p}} + \left(\int_{[a,b]} |g|^p\right)^{\frac{1}{p}}.$$

Ainsi, de même qu'au 3.1, la quantité

$$\|f\|_p = \left(\int_{[a,b]} |f|^p\right)^{\frac{1}{p}}$$

définit une norme sur $\mathcal{C}^0([a, b], \mathbb{R})$ (mais seulement une semi-norme sur $\mathrm{Int}([a, b], \mathbb{R})$).