14 | Convexité

« Si je devais me réveiller après avoir dormi pendant mille ans, ma première question serait : l'hypothèse de Riemann a-t-elle été prouvée ? »

David Hilbert (1862–1943)

Plan de cours

Dans tout ce chapitre, E désignera un $\mathbb{R}$-espace vectoriel de dimension quelconque.

I | Parties convexes d'un espace vectoriel réel

A – Introduction aux barycentres

Définition 14.1 : Point pondéré et barycentre d'une famille finie de points pondérés

- On appelle point pondéré tout couple $(A, \lambda) \in E \times \mathbb{R}$.
- On appelle système pondéré toute famille $((A_1, \lambda_1), \dots, (A_n, \lambda_n))$ de points pondérés dont le poids total $\Lambda = \sum_{k=1}^{n} \lambda_k$ est non nul. Le barycentre du système pondéré est alors défini par $G = \frac{1}{\Lambda} \sum_{k=1}^{n} \lambda_k A_k$.

Le barycentre, lorsqu'il existe, est une moyenne (pondérée) de points. Lorsque les coefficients sont tous égaux, le barycentre est qualifié d'isobarycentre.

On utilise parfois la notation $\mathrm{bar}((A_i, \lambda_i))$ pour désigner le barycentre du système pondéré associé.

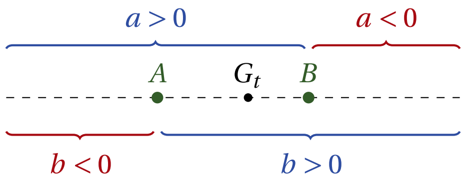
Exemple – cas du plan affine euclidien

> G est l'unique point tel que $\sum_{k=1}^{n} \lambda_k \overrightarrow{GA_k} = \overrightarrow{0}$. Pour tout point M, $\overrightarrow{MG} = \frac{1}{\Lambda} \sum_{k=1}^{n} \lambda_k \overrightarrow{MA_k}$.
> Si les λ_k représentent des masses associées aux points A_k, le barycentre G s'interprète naturellement comme le centre de masse.

Notons que le barycentre du système pondéré $((A_1, \alpha\lambda_1), \dots, (A_n, \alpha\lambda_n))$ coïncide, pour tout $\alpha \neq 0$, avec celui du système pondéré $((A_1, \lambda_1), \dots, (A_n, \lambda_n))$. C'est seulement le poids total, $\alpha\Lambda$, qui diffère. Un poids total non nul peut donc toujours être ramené à 1.

Exemple – retour au plan affine euclidien

- Cas d'un système pondéré à deux points (A, a) et (B, b) de poids total 1



Représentation du barycentre G

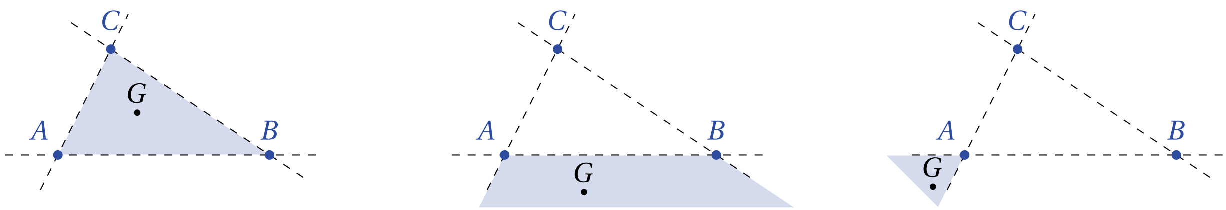
$$a\overrightarrow{GA} + b\overrightarrow{GB} = \overrightarrow{0}$$

Si le poids total est ramené à 1, on peut écrire :

$$G_t = (1-t)A + tB$$

Quand t parcourt $[0, 1]$, G_t parcourt le segment $[A, B]$

- Cas d'un système pondéré à trois points (A,a), (B,b) et (C,c) de poids total 1



$a \geqslant 0,\ b \geqslant 0,\ c \geqslant 0$ $\quad$ $a \geqslant 0,\ b \geqslant 0,\ c \leqslant 0$ $\quad$ $a \geqslant 0,\ b \leqslant 0,\ c \leqslant 0$

Le barycentre des barycentres (affectés des bons poids) reste le barycentre du système pondéré initial.

Proposition 14.2 : Associativité du barycentre

Soient deux entiers n, p tels que $1 \leqslant p \leqslant n$ et $((A_1,\lambda_1),\dots,(A_n,\lambda_n))$ un système pondéré.

On pose $G_1 = \text{bar}\big((A_1,\lambda_1),\dots,(A_p,\lambda_p)\big)$, $G_2 = \text{bar}\big((A_{p+1},\lambda_{p+1}),\dots,(A_n,\lambda_n)\big)$, $\mu_1 = \sum_{k=1}^{p} \lambda_k$ et $\mu_2 = \sum_{k=p+1}^{n} \lambda_k$.

Alors le barycentre du système pondéré $((A_1,\lambda_1),\dots,(A_n,\lambda_n))$ est aussi celui de $((G_1,\mu_1),(G_2,\mu_2))$.

Démonstration

En reprenant les notations de l'énoncé, $\mu_1 G_1 + \mu_2 G_2 = \sum_{k=1}^{p} \lambda_k A_k + \sum_{k=p+1}^{n} \lambda_k A_k = \sum_{k=1}^{n} \lambda_k A_k$. ■

B – Partie convexe d'un espace vectoriel réel

Nous noterons par la suite en minuscule les points du $\mathbb{R}$-espace vectoriel E.

Définition 14.3 : Segment

On appelle segment d'extrémités a et b l'ensemble de points :

$$[a,b] = \{tb + (1-t)a \mid t \in [0,1]\}$$

Un segment est donc l'ensemble des barycentres à coefficients positifs des points a et b.

Définition 14.4 : Partie convexe

Une partie C de E est dite convexe si pour tous $a, b \in C$, $[a,b] \subset C$.

Partie convexe $\quad$ *Partie non convexe*

Proposition 14.5

Une partie C est convexe si et seulement si C contient tout barycentre à coefficients positifs de toute famille de points de C.

Autrement dit, une partie est convexe si et seulement si elle est stable par « barycentration ».

Démonstration

Par récurrence sur le nombre de points, en utilisant l'associativité du barycentre. ■

Ainsi, toute partie convexe C contient automatiquement tout polygone dont les sommets sont dans C.

II | Fonctions convexes d'une variable réelle

A – Généralités

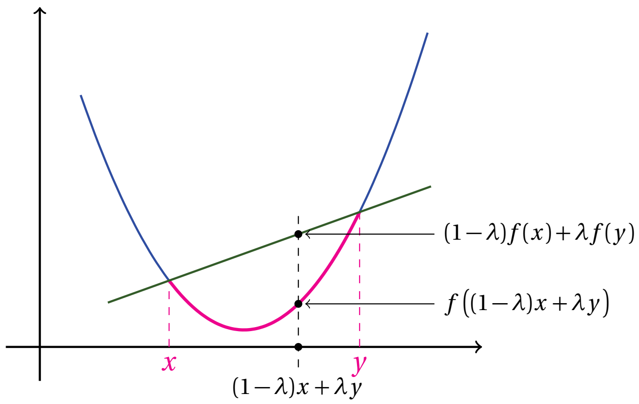
On considère des fonctions définies sur un intervalle I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$.

Définition 14.6 : Fonction convexe, fonction concave

Une fonction f est convexe sur l'intervalle I de $\mathbb{R}$ si pour tout $(x, y) \in I^2$ et tout $\lambda \in [0,1]$:

$$f\big((1-\lambda)x + \lambda y\big) \leqslant (1-\lambda)f(x) + \lambda f(y)$$

Une application f est dite concave lorsque $-f$ est convexe.



La définition reste valable pour f définie sur un $\mathbb{R}$-espace vectoriel et à valeurs dans $\mathbb{R}$.

Interprétation géométrique – Une fonction est convexe si et seulement si son graphe est au-dessous de toutes ses cordes. Nous verrons d'autres caractérisations géométriques de la convexité.

Exemple

Montrons que $t \mapsto t^2$ est convexe sur $\mathbb{R}$. Soient $x, y \in \mathbb{R}$ et $\lambda \in [0,1]$.

$$\begin{aligned}\big((1-\lambda)x + \lambda y\big)^2 &= (1-\lambda)^2x^2 + 2\lambda(1-\lambda)xy + \lambda^2y^2 \\ &\leqslant (1-\lambda)^2x^2 + \lambda(1-\lambda)(x^2+y^2) + \lambda^2y^2 = (1-\lambda)x^2 + \lambda y^2\end{aligned}$$

Une fonction est convexe si et seulement si l'image de tout barycentre est au-dessous du barycentre des images.

Proposition 14.7 : Inégalité de Jensen

Soit $f : I \to \mathbb{R}$ une fonction convexe. Pour tous $x_1, \dots, x_n \in I$ et $\lambda_1, \dots, \lambda_n$ des réels positifs de somme 1. Alors,

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i)$$

Démonstration

Procédons là encore par récurrence, grâce à l'associativité du barycentre.

- **Initialisation –** Pour $n = 2$, on retrouve la définition d'une fonction convexe.
- **Hérédité –** Supposons la propriété vraie pour n points, montrons qu'elle l'est encore pour $n+1$ points $x_1, \dots, x_{n+1} \in I$, associés à des poids positifs $\lambda_1, \dots, \lambda_{n+1}$ de somme 1. Pour appliquer l'hypothèse de récurrence en manipulant des poids de somme 1, écrivons :

$$\begin{aligned}f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) = f\left((1-\lambda_{n+1})\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i + \lambda_{n+1}x_{n+1}\right) &\leqslant (1-\lambda_{n+1}) f\left(\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i\right) + \lambda_{n+1} f(x_{n+1}) \\ &\leqslant (1-\lambda_{n+1})\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} f(x_i) + \lambda_{n+1} f(x_{n+1}) = \sum_{i=1}^{n+1} \lambda_i f(x_i)\end{aligned}$$

Ainsi s'achève notre démonstration par récurrence. ∎

B – Caractérisations géométriques

Définition 14.8

Soit $f : I \to \mathbb{R}$. On appelle épigraphe de f l'ensemble des points du plan situés au-dessus du graphe de f. Autrement dit,

$$\mathscr{E} = \{(x, y) \in \mathbb{R}^2 \mid y \geqslant f(x)\}$$

Proposition 14.9

Une fonction est convexe si et seulement si son épigraphe est une partie convexe de $\mathbb{R}^2$.

Démonstration

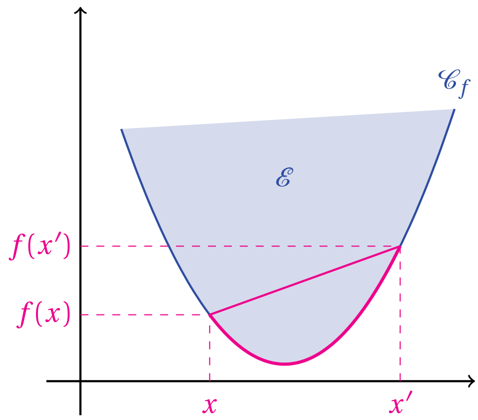
$\boxed{\Longrightarrow}$ Supposons f convexe et donnons-nous deux points (x, y) et (x', y') de l'épigraphe $\mathscr{E}$ ainsi que $\lambda \in [0,1]$. On a bien $((1-\lambda)x + \lambda x', (1-\lambda)y + \lambda y') \in \mathscr{E}$ puisque :

$$\begin{aligned} f\big((1-\lambda)x+\lambda x'\big) &\leqslant (1-\lambda)f(x)+\lambda f(x') \quad \text{(par convexité de } f\text{)} \\ &\leqslant (1-\lambda)y+\lambda y' \quad \text{(car } (x,y),(x',y')\in\mathscr{E}\text{)} \end{aligned}$$

$\boxed{\Longleftarrow}$ Réciproquement, si l'épigraphe est convexe, le segment d'extrémités $(x, f(x))$ et $(x', f(x'))$ est inclus dans l'épigraphe.

$$\forall \lambda \in \mathbb{R}, \quad (1-\lambda)f(x)+\lambda f(x') \geqslant f((1-\lambda)x+\lambda x')$$

C'est la définition de la convexité ! ■



Théorème 14.10 : Croissance de la pente

Soit $f : I \to \mathbb{R}$. f est convexe si et seulement si pour tout $x \in I$, $t \mapsto \dfrac{f(t)-f(x)}{t-x}$ est croissante sur $I \setminus \{x\}$.

Démonstration

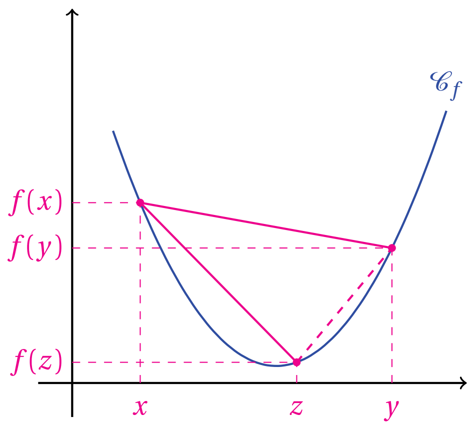
$\boxed{\Longrightarrow}$ On suppose f convexe. Soient $x, y, z \in I$ avec $x < z < y$.

- On a $z = (1-\lambda)x + \lambda y$ avec $\lambda = \dfrac{z-x}{y-x} \in [0,1]$.
- Par convexité, $f(z) \leqslant (1-\lambda)f(x) + \lambda f(y)$. Autrement dit, $f(z) - f(x) \leqslant \lambda(f(y) - f(x))$. On vient d'établir :

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

On fait de même pour $z < y < x$ et $z < x < y$.

$\boxed{\Longleftarrow}$ Soient $x, y \in I$ distincts. On applique l'inégalité précédente (croissance de la pente) à $z = (1-\lambda)x + \lambda y$. On retrouve bien $f\big((1-\lambda)x+\lambda y\big) \leqslant (1-\lambda)f(x)+\lambda f(y)$. ■



Corollaire 14.11 : Inégalité des pentes

Soient $f : I \to \mathbb{R}$ une fonction convexe et $x, y, z \in I$ vérifiant $x < z < y$. Alors,

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

Démonstration

L'inégalité de gauche traduit la croissance de $t \mapsto \dfrac{f(t)-f(x)}{t-x}$ et celle de droite de $t \mapsto \dfrac{f(t)-f(y)}{t-y}$. ■

C – Caractérisation par la dérivée

La croissance de la dérivée permet d'établir facilement la convexité d'une fonction dérivable.

Théorème 14.12

Soit f une fonction dérivable sur l'intervalle I. Alors, les trois propriétés suivantes sont équivalentes :

(i) f est convexe

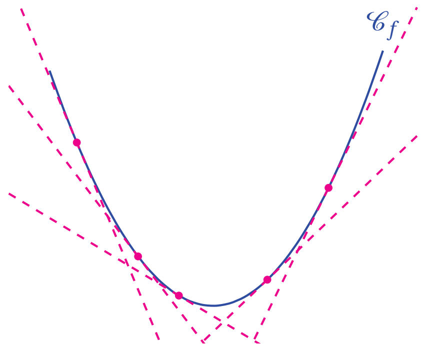
(ii) f' est croissante

(iii) le graphe de f est au-dessus de toutes ses tangentes

Démonstration

$\boxed{\text{(i)}\Longrightarrow\text{(ii)}}$ Supposons f convexe et soient $x, y \in I$ avec $x < y$.
Par inégalité des pentes,

$$\forall z \in]x, y[, \quad \frac{f(z)-f(x)}{z-x} \leqslant \frac{f(y)-f(x)}{y-x} \leqslant \frac{f(y)-f(z)}{y-z}$$



En faisant tendre z vers x à gauche dans la première inégalité et z vers y à droite dans la seconde,

$$f'(x) \leqslant \frac{f(y)-f(x)}{y-x} \leqslant f'(y)$$

On a ainsi établi la croissance de f'.

$\boxed{\text{(ii)}\Longrightarrow\text{(iii)}}$ Supposons f' croissante. Montrons que le graphe de f est au-dessus de sa tangente en $a \in I$ quelconque. La tangente a pour équation $y = f'(a)(x-a)+f(a)$. Considérons alors la fonction dérivable $\varphi : x \mapsto f(x)-f'(a)(x-a)-f(a)$.

$$\forall x \in I, \quad \varphi'(x) = f'(x)-f'(a)$$

Par croissance de f', φ est décroissante à gauche de a et croissante à droite de a. Comme $\varphi(a) = 0$, φ est positive sur I. Le graphe de f est bien au-dessus de toutes ses tangentes.

$\boxed{\text{(iii)}\Longrightarrow\text{(i)}}$ Supposons le graphe de f au-dessus de toutes ses tangentes. Soient $x, y \in I$ et $\lambda \in [0,1]$.

$$\forall a \in I, \quad f(x) \geqslant f'(a)(x-a)+f(a) \quad \text{et} \quad f(y) \geqslant f'(a)(y-a)+f(a)$$

Ainsi, $(1-\lambda)f(x)+\lambda f(y) \geqslant \big[(1-\lambda)x+\lambda y-a\big]f'(a)+f(a)$. Il suffit de prendre $a = (1-\lambda)x+\lambda y$ et il vient :

$$f\big((1-\lambda)x+\lambda y\big) \leqslant (1-\lambda)f(x)+\lambda f(y)$$

D'où la convexité de f. ■

Corollaire 14.13

Soit f une fonction deux fois dérivable sur l'intervalle I. Alors f est convexe (resp. concave) si et seulement si $f'' > 0$ (resp. $f'' < 0$).

Exemples

exp est convexe sur $\mathbb{R}$, ln est concave sur $\mathbb{R}_+^*$ et sin sur $[0, \pi/2]$. En particulier,

$$\forall x \in \mathbb{R}, \quad \mathrm{e}^x \geqslant x+1\,; \quad \forall x > -1, \quad \ln(1+x) \leqslant x\,; \quad \forall x \in [0, \pi/2], \quad \frac{2}{\pi}x \leqslant \sin(x) \leqslant x$$

Exercice 1 – Inégalité arithmético-géométrique

Prouver que pour toute famille de réels strictement positifs $(x_1, \ldots, x_n)$, $\sqrt[n]{x_1 \times \cdots \times x_n} \leqslant \dfrac{x_1+\cdots+x_n}{n}$.

D – Synthèse

Soit f une fonction définie sur un intervalle I et à valeurs dans $\mathbb{R}$.

$$\begin{aligned}
f \text{ est convexe sur } I \quad &\iff \quad f' \text{ est croissante sur } I \quad \iff \quad f'' \text{ est positive sur } I \\
&\iff \quad \forall x \in I, \quad t \mapsto \frac{f(t)-f(x)}{t-x} \quad \text{est croissante sur } I \setminus \{x\} \\
&\iff \quad f \text{ vérifie l'inégalité des pentes} \\
&\iff \quad \text{le graphe de } f \text{ est en-dessous de toutes ses cordes} \\
&\iff \quad \text{le graphe de } f \text{ est au-dessus de toutes ses tangentes} \\
&\iff \quad \text{l'épigraphe de } f \text{ est convexe}
\end{aligned}$$