Chapitre I

Parties convexes d'un espace vectoriel et rappels sur les fonctions convexes

Table des matières

Partie A : Parties convexes d'un espace vectoriel **2**
1. Barycentres 2
2. Parties convexes 4

Partie B : Fonctions convexes d'une variable réelle **8**
1. Fonctions convexes 8
2. Convexité et épigraphe 9
3. Inégalité des pentes 12

Partie C : Convexité et dérivabilité **15**
1. Fonctions convexes dérivables 15
2. Fonctions convexes deux fois dérivables 16
3. Exemples d'inégalités classiques 17

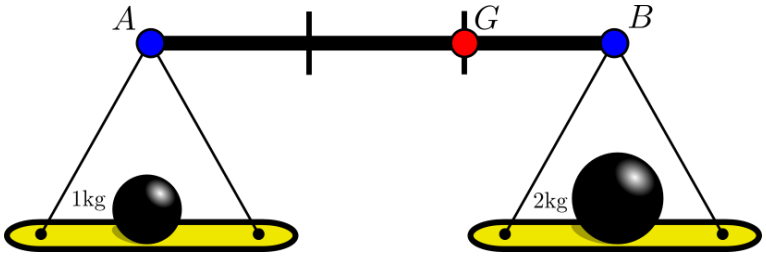
Partie A

Parties convexes d'un espace vectoriel

Dans cette partie, E désigne un espace vectoriel réel. On le muni de sa structure d'espace affine canonique et on note O son origine. Pour un vecteur u de E, on identifie (abusivement) le point $O+u$ et le vecteur u. De plus, compte tenu de cette identification, si u et v sont des points de E, on note $\overrightarrow{uv} = v - u$.

1. Barycentres

La notion de barycentre est le pendant mathématique de la notion physique de centre de gravité (ou centre de masses) : considérons une barre sur laquelle on fixe deux plateaux. Si on place sur le premier plateau une masse de 1kg et sur le deuxième, une masse de 2kg, le centre de gravité du système (son point d'équilibre) se trouvera aux deux tiers de la barre. On modélise ce système dans le plan affine :



On remarque que le centre de gravité G vérifie l'égalité vectorielle $\overrightarrow{GA} + 2\overrightarrow{GB} = \overrightarrow{0}$, ou de manière équivalente, avec O l'origine du plan affine :

$$\overrightarrow{OG} = \frac{1}{2+1}\left(\overrightarrow{OA} + 2\overrightarrow{OB}\right).$$

On généralise alors cette situation :

Définition 1. *Barycentres*

Soit $n \in \mathbb{N}^*$ et $x_1, ..., x_n$ des points de E.
On dit que l'élément g de E est un **barycentre** de $x_1, ..., x_n$ s'il existe des réels $\lambda_1, ..., \lambda_n$ vérifiant $\sum_{i=1}^{n} \lambda_i \neq 0$ et tels que :

$$g = \frac{1}{\sum_{i=1}^{n} \lambda_i} \sum_{i=1}^{n} \lambda_i x_i.$$

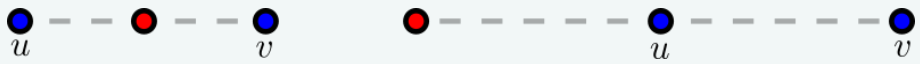
Dans ce cas, on dit que g est le **barycentre de la famille de points pondérés** $((x_i, \lambda_i))_{1 \leq i \leq n}$.

Remarque 1.

Si tous les réels de la famille $(\lambda_i)_{1\leq i\leq n}$ sont égaux à un certain réel non nul λ, le barycentre est appelé **isobarycentre** et est égal à $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$. On remarque que l'isobarycentre ne dépend pas de la valeur commune λ des coefficients.

Exemple 1.

1. L'isobarycentre de deux points u et v est le point $\frac{1}{2}(u+v)$. Le barycentre de $((u,2),(v,-1))$ est le point $2u-v$.



2. Soit X une variable aléatoire prenant un nombre fini de valeurs réelles $x_1,...,x_n$ de probabilités respectives $p_1,...,p_n$. Le barycentre de la famille $((x_i,p_i))_{1\leq i\leq n}$ n'est autre que l'espérance de X.

> En effet, comme $\sum_{i=1}^{n} p_i = 1$, on a, si g est le barycentre de la famille $((x_i,p_i))_{1\leq i\leq n}$:
>
> $$g = \frac{1}{\sum_{i=1}^{n} p_i}\sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} x_i P(X=x_i) = E(X).$$

Exercice 1.

1. Dans $\mathbb{R}^2$, déterminer l'isobarycentre des points $(-1,0),(1,0)$ et $(1,1)$. A quel point particulier du triangle formé par ces trois points ce barycentre correspond-il ?
2. Soit u et v deux vecteurs de E. Montrer que l'ensemble des barycentres de u et v est égal à $u+\mathrm{Vect}(\overrightarrow{uv})$.
   En déduire la nature géométrique de l'ensemble des barycentres de deux points.

Correction.

1. Le calcul donne $g=(\frac{1}{3},\frac{1}{3})$. Il s'agit du centre de gravité du triangle (i.e. le point d'intersection des médianes du triangle).
2. On note $\mathcal{G}$ l'ensemble des barycentres de u,v. Soit $\lambda,\mu\in\mathbb{R}$ avec $\alpha=\lambda+\mu\neq 0$. On a

$$\frac{1}{\alpha}(\lambda u+\mu v) = \frac{1}{\alpha}((\alpha-\mu)u+\mu v) = u+\frac{\mu}{\alpha}(v-u). \quad (*)$$

   Par suite, le barycentre $g=\frac{1}{\alpha}(\lambda u+\mu v)$ appartient à $u+\mathrm{Vect}(v-u)$. D'où $\mathcal{G}\subset u+\mathrm{Vect}(\overrightarrow{uv})$.

   Soit $x\in u+\mathrm{Vect}(\overrightarrow{uv})$. Alors il existe $\beta\in\mathbb{R}$ tel que $x=u+\beta(v-u)$. En posant $\mu=\beta$ $\lambda=1-\beta$ et $\alpha=1$, on a bien $\mu+\lambda=\alpha$ et, d'après l'égalité $(*)$, $x=\frac{1}{\alpha}(\lambda u+\mu v)\in\mathcal{G}$. D'où $u+\mathrm{Vect}(\overrightarrow{uv})\subset\mathcal{G}$.
   Il en résulte que $\mathcal{G}=u+\mathrm{Vect}(\overrightarrow{uv})$.

Proposition 1.

Soit $x_1, ..., x_n$ des points de E, $\lambda_1, ..., \lambda_n$ des réels tels que $\sum_{i=1}^{n} \lambda_i \neq 0$ et μ un réel non nul. Le barycentre de la famille de points pondérés $((x_i, \mu\lambda_i))_{1 \leq i \leq n}$ est bien défini et est égal au barycentre de la famille de points pondérés $((x_i, \lambda_i))_{1 \leq i \leq n}$.

Démonstration.

Il suffit de remarquer que l'on peut simplifier par μ dans la définition du barycentre de la famille $((x_i, \mu\lambda_i))_{1 \leq i \leq n}$. □

Remarque 2.

En vertu de la proposition précédente, quitte à diviser par la somme des pondérations, tout barycentre peut être considéré comme celui d'une famille dont la somme des pondérations vaut 1.

Définition 2. *Barycentre à coefficients positifs*

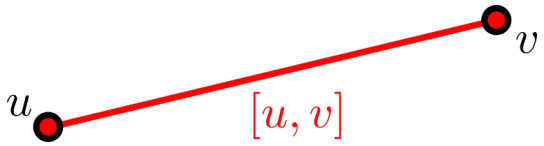
Soit $x_1, ..., x_n$ des points de E et $\lambda_1, ..., \lambda_n$ des réels tels que $\sum_{i=1}^{n} \lambda_i \neq 0$. On dit que le barycentre de la famille de points pondérés $((x_i, \lambda_i))_{1 \leq i \leq n}$ est **à coefficients positifs** si, pour tout $i \in \{1, ..., n\}$, $\lambda_i \geq 0$.

2. Parties convexes

Définition 3. *Segment*

Soit u, v des points de E. Le **segment** entre u et v, noté $[u, v]$ est l'ensemble défini par :

$$[u, v] := \{tu + (1-t)v \mid t \in [0, 1]\}$$



Proposition 2.

Soit $x, y, z \in \mathbb{R}$ avec $x \leq z$. On a $x \leq y \leq z$ si et seulement si, il existe $t \in [0, 1]$ tel que $y = tx + (1-t)z$.
Dans ce cas, et si $z \neq x$, on a $t = \dfrac{z-y}{z-x}$.

Démonstration.

($\Rightarrow$) Tout d'abord, si $x = z$, alors $x = y = z$ et tout $t \in [0,1]$ convient.

On suppose $x \neq z$. On peut remarquer que y est le barycentre de la famille $((x, z-y), (z, y-x))$. En effet, on a :

$$(z-y)x + (y-x)z = (z-x)y.$$

Ainsi, en posant $t = \dfrac{z-y}{z-x}$, on obtient $1-t = \dfrac{y-x}{z-x}$, d'où, finalement :

$$y = tx + (1-t)z.$$

($\Leftarrow$) On suppose qu'il existe $t \in [0,1]$ tel que $y = tx + (1-t)z$. Comme $x \leq z$, $t \geq 0$ et $1-t \geq 0$, on a :

$$x = tx + (1-t)x \geq y = tx + (1-t)z \leq tz + (1-t)z = z$$

et de plus, si $z \neq x$, comme on vient de montrer que $x \leq y \leq z$ d'après le point précédent, $y = t'x + (1-t')z$ avec $t' = \frac{z-y}{z-x}$.
Ainsi, on a :

$$\begin{array}{rcl} 0 & = & y - y \\ & = & tx + (1-t)z - (t'x + (1-t')z) \\ 0 & = & (t-t')(x-z) \end{array}$$

Or comme $x \neq z$, et que $\mathbb{R}$ est un anneau intègre, on obtient $t = t' = \frac{z-y}{z-x}$.

□

Exercice 2.

Soit $x, z \in \mathbb{R}$ avec $x \leq z$. Montrer que la notation $[x, z]$ en tant qu'intervalle de $\mathbb{R}$ coïncide bien avec la définition précédente de segment.

Correction.

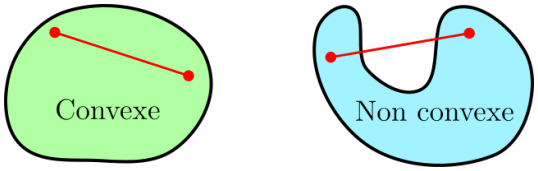
On note $[x,z]_1 = \{y \in \mathbb{R} \mid x \leq y \leq z\}$ (notation en tant qu'intervalle de $\mathbb{R}$) ;
et $[x,z]_2 = \{tx + (1-t)z \mid t \in [0,1]\}$ (notation définie précédemment).

D'après la proposition précédente, on a $y \in [x,z]_1$ i.e. $x \leq y \leq z$ si, et seulement si, il existe $t \in [0,1]$ tel que $y = tx + (1-t)z$ i.e. $y \in [x,z]_1$.
D'où $[x,z]_1 = [x,z]_2$. Les notations coïncident donc bien !

Définition 4. *Partie convexe*

Soit A une partie de E.
On dit que A est **convexe** si, pour tous $u, v \in A$, le segment $[u, v]$ est inclus dans A ; i.e.

$$\forall u, v \in A,\ \forall t \in [0,1], \quad tu + (1-t)v \in A.$$



Exemple 2.

Les segments, les sous-espaces vectoriels et les sous-espaces affines sont des parties convexes.

— **Segments**. Soit $x, y \in E$. Soit $u, v \in [x, y]$. Alors il existe $\alpha, \beta \in [0, 1]$ tels que $u = \alpha x + (1 - \alpha)y$ et $v = \beta x + (1 - \beta)y$.
Par suite, on a, pour tout $t \in [0, 1]$:

$$tu + (1 - t)v = t\,(\alpha x + (1 - \alpha)y) + (1 - t)\,(\beta x + (1 - \beta)y) = sx + (1 - s)y,$$

où $s = t\alpha + (1 - t)\beta$.
Or s est compris entre 0 et 1 car $s \in [\alpha, \beta] \subset [0, 1]$.
Il en résulte que $tu + (1 - t)v \in [x, y]$; d'où $[x, y]$ est une partie convexe de E.

— **Sous-espace vectoriel**. Il suffit de remarquer qu'un sous-espace vectoriel est stable par combinaison linéaire.

Théorème 1. *Caractérisation des convexes par les barycentres*

Soit A une partie de E.
A est convexe si, et seulement si, le barycentre à coefficients positifs de toute famille de points pondérés de A appartient à A.

Démonstration.

Soit $A \subset E$. Comme dans la remarque 2, on supposera sans perte de généralité que les sommes de pondérations associées aux barycentres valent 1.

- ($\Leftarrow$). Immédiat : pour tous $u, v \in A$ et tout $t \in \mathbb{R}$, on remarque que $tu + (1 - t)v$ est le barycentre à coefficients positifs de $((u, t), (v, (1 - t))$.
- ($\Rightarrow$). On suppose A convexe. On montre, par récurrence sur $\mathbb{N}^*$, que, pour tout $n \in \mathbb{N}^*$, $\forall x_1, ...x_n \in A$, $\forall \lambda_1, ..., \lambda_n \in \mathbb{R}_+$ tels que $\sum_{i=1}^n \lambda_i = 1$, $\sum_{i=1}^n \lambda_i x_i \in A$.
  — La propriété est triviale pour $n = 1$.
  — Soit $n \in \mathbb{N}^*$. On suppose la propriété vraie pour l'entier n. Montrons qu'elle est vraie pour $n + 1$.
  Soit $x_1, ..., x_{n+1} \in A$ et $\lambda_1, ..., \lambda_{n+1} \in \mathbb{R}$ tels que $\sum_{i=1}^{n+1} \lambda_i = 1$.
  1er cas : $\lambda_{n+1} = 1$. Alors pour $i = 1, ..., n$, $\lambda_i = 0$ d'où $\sum_{i=1}^{n+1} \lambda_i x_i = x_{n+1} \in A$.
  2eme cas : $\lambda_{n+1} \neq 1$. Alors $\sum_{i=1}^n \lambda_i = 1 - \lambda_{n+1} \neq 0$. On a donc :

$$\sum_{i=1}^{n+1} \lambda_i x_i = \left(\frac{1 - \lambda_{n+1}}{\sum_{i=1}^n \lambda_i} \sum_{i=1}^n \lambda_i x_i \right) + \lambda_{n+1} x_{n+1}.$$

Or, par hypothèse de récurrence, $g := \sum_{i=1}^{n} \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i} x_i$ appartient à A.

Par suite, $\sum_{i=1}^{n+1} \lambda_i x_i = (1 - \lambda_{n+1})g + \lambda_{n+1} x_{n+1}$ appartient à A car $\lambda_{n+1} \in [0, 1]$ et A est convexe.

Ce qui achève le raisonnement par récurrence.

$\square$

Partie B

Fonctions convexes d'une variable réelle

Dans cette partie, I désigne un intervalle de $\mathbb{R}$ d'intérieur non vide.

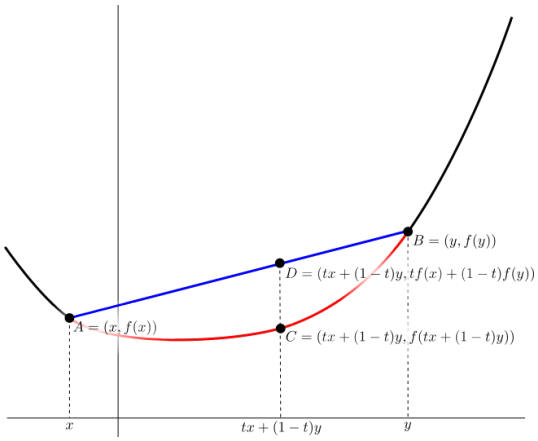
1. Fonctions convexes

Définition 5. *Fonction convexe / concave*

Soit $f : I \to \mathbb{R}$ une fonction. On dit que f est **convexe** si, pour tous $x, y \in I$ et pour tout $t \in [0,1]$:

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

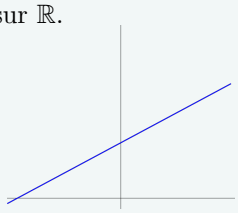
On dit que f est **concave** si $-f$ est convexe.



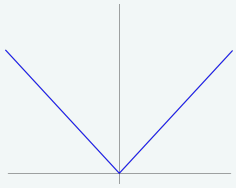
Interprétation graphique : On considère le graphe de d'une fonction f de I dans $\mathbb{R}$, c'est-à-dire l'ensemble $\mathcal{G}_f = \{(x, f(x)) \mid x \in I\}$, dans le plan.
La fonction f est convexe sur I si, pour tous x, y dans I, le graphe de la fonction $f_{|[x,y]}$ (i.e. f restreinte au segment $[x, y]$) est en dessous de la corde qui relie $A = (x, f(x))$ et $B = (y, f(y))$.
En effet, tout $z \in [x, y]$ peut s'écrire $z = tx + (1-t)y$ avec $t \in [0,1]$ et les points d'abscisse z sur le graphe de f et sur la corde admettent pour ordonnées $f(tx + (1-t)y)$ et $tf(x) + (1-t)f(y)$, respectivement.

Exemple 3.

— Toute fonction affine est convexe sur $\mathbb{R}$.



— La fonction valeur absolue est convexe sur $\mathbb{R}$.



Exercice 3.

1. Démontrer ces deux exemples.
2. Démontrer que $x \mapsto x^2$ est convexe sur $\mathbb{R}$ et que $x \mapsto x^3$ n'est ni convexe ni concave sur $\mathbb{R}$.

Correction.

1. Il est clair qu'une fonction affine vérifie le cas d'égalité dans la définition de fonction convexe. Pour la fonction valeur absolue, il suffit d'appliquer l'inégalité triangulaire.
2. On note $f : x \mapsto x^2$. Soit $x, y \in \mathbb{R}$ et $t \in [0, 1]$. On a
$$\begin{aligned} f(tx + (1-t)y) - (tf(x) - (1-t)f(y)) &= t(1-t)(2xy - x^2 - y^2) \\ &= -t(1-t)(x-y)^2 \leq 0. \end{aligned}$$
Donc $x \mapsto x^2$ est convexe sur $\mathbb{R}$.

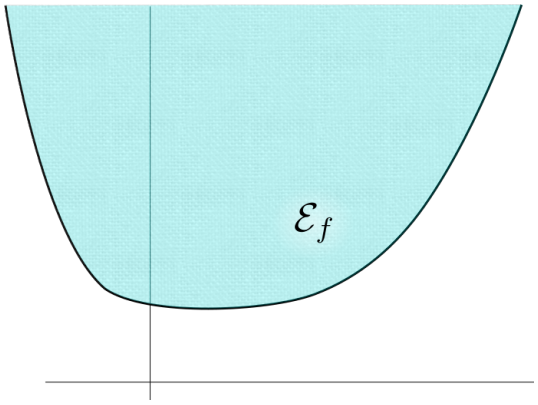
   Pour la fonction $x \mapsto x^3$, on remarque que la corde entre entre les points d'abscisse -1 et 1 est au dessous puis au dessus de la courbe. Les abscisses $-1/2$ et $1/2$ (par exemple) fournissent un contre-exemple convenable.

2. Convexité et épigraphe

La convexité (ou la concavité) d'une fonction peut se lire géométriquement à partir de son graphe. Pour cela, on définit l'épigraphe, l'ensemble des points au dessus de la courbe de la fonction :

Définition 6. *Épigraphe*

On appelle **épigraphe** de f le sous-ensemble de $I \times \mathbb{R}$, noté $\mathcal{E}_f$, définit par :
$$\mathcal{E}_f = \{(x, y) \in I \times \mathbb{R} \mid f(x) \leq y\}.$$



Théorème 2. *Caractérisation géométrique de la convexité*

Soit $f : I \to \mathbb{R}$. La fonction f est convexe sur I si, et seulement si, son épigraphe $\mathcal{E}_f$ est une partie convexe de $\mathbb{R}^2$.

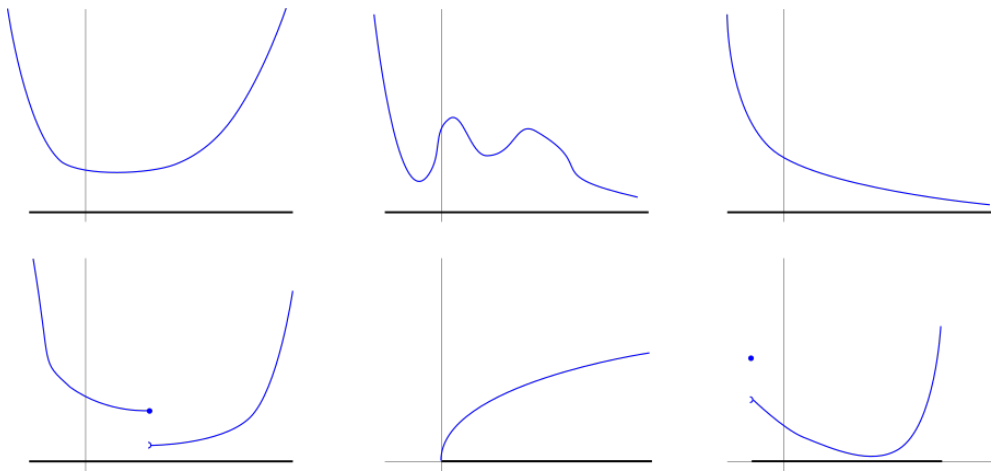
Démonstration.

- ($\Rightarrow$). On suppose f convexe sur I. Montrons que $\mathcal{E}_f$ est convexe.
  Soit $M_1 = (x_1, y_1)$, $M_2 = (x_2, y_2)$ des points de $\mathcal{E}_f$ et $M = (x, y) \in [M_1, M_2]$. Alors il existe $t \in [0, 1]$ tel que $M = tM_1 + (1-t)M_2$, d'où :
  $$x = tx_1 + (1-t)x_2 \text{ et } y = ty_1 + (1-t)y_2.$$
  La fonction f étant convexe, on a $f(x) \leq tf(x_1) + (1-t)f(x_2)$. Or, M_1, M_2 appartiennent à l'épigraphe donc $f(x_1) \leq y_1$ et $f(x_2) \leq y_2$. Par suite, comme t et $(1-t)$ sont positifs,
  $$f(x) \leq ty_1 + (1-t)y_2 = y.$$
  Donc M appartient à $\mathcal{E}_f$.
  Il en résulte que tout segment de $\mathcal{E}_f$ est inclus dans $\mathcal{E}_f$ donc $\mathcal{E}_f$ est convexe.
- ($\Leftarrow$). On suppose $\mathcal{E}_f$ convexe. Montrons que f est convexe sur I.
  Soit $x_1, x_2 \in I$. Alors $M_1 = (x_1, f(x_1))$ et $M_2 = (x_2, f(x_2))$ appartiennent à $\mathcal{E}_f$ (car $f(x_i) \leq f(x_i)$, $i = 1, 2$). Par convexité de $\mathcal{E}_f$, le segment $[M_1, M_2]$ est inclus dans $\mathcal{E}_f$. Autrement dit, pour tout $t \in [0, 1]$, $(x, y) := t(x_1, f(x_1)) + (1-t)(x_2, f(x_2))$ vérifie $f(x) \leq y$. Donc, pour tout $t \in [0, 1]$:
  $$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$
  Ceci étant vrai pour tous x_1, x_2 dans I, il en résulte que f est convexe sur I.

$\square$

Exercice 4.

1. A partir des graphes ci-dessous, déterminer si les fonctions qui leur sont associées sont convexes ou non.



2. En examinant le graphe de $f : x \mapsto \frac{1}{x^2}$, discuter de la convexité de f sur son domaine de définition et sur certains de ses sous-domaines.

Théorème 3. *Inégalité de Jensen*

Soit $f : I \to \mathbb{R}$ une fonction. Si f est convexe sur I, alors, pour tout $n \in \mathbb{N}^*$, tous points $x_1, ..., x_n$ de I et tous réels positifs $\lambda_1, ..., \lambda_n$ tels que $\sum_{i=1}^n \lambda_i = 1$, on a :

$$f(\sum_{i=1}^n \lambda_i x_i) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

Démonstration.

On suppose f convexe sur I.
Soit $n \in \mathbb{N}^*$ et $((x_i, \lambda_i))_{1 \leq i \leq n}$ une famille de points pondérés avec, pour tout $1 \leq i \leq n$, $x_i \in I$ et $\lambda_i \geq 0$; et $\sum_{i=1}^n \lambda_i = 1$.
Pour chaque $1 \leq i \leq n$, le point $M_i = (x_i, f(x_i))$ du graphe de f appartient à son épigraphe $\mathcal{E}_f$ qui, d'après le théorème 2, est une partie convexe de $\mathbb{R}^2$. Par suite, d'après le théorème 1, le barycentre G à coefficients positifs de la famille $((M_i, \lambda_i))_{1 \leq i \leq n}$ appartient à $\mathcal{E}_f$.
Or on a :

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

et donc $f(\sum_{i=1}^n \lambda_i x_i) \leq \sum_{i=1}^n \lambda_i f(x_i)$ car G appartient à $\mathcal{E}_f$. □

Exercice 5.

1. Démontrer que pour tout $n \in \mathbb{N}^*$ et pour tous $x_1, ..., x_n \in \mathbb{R}$,

$$\left(\sum_{i=1}^{n} x_i\right)^2 \leq n \sum_{i=1}^{n} x_i^2.$$

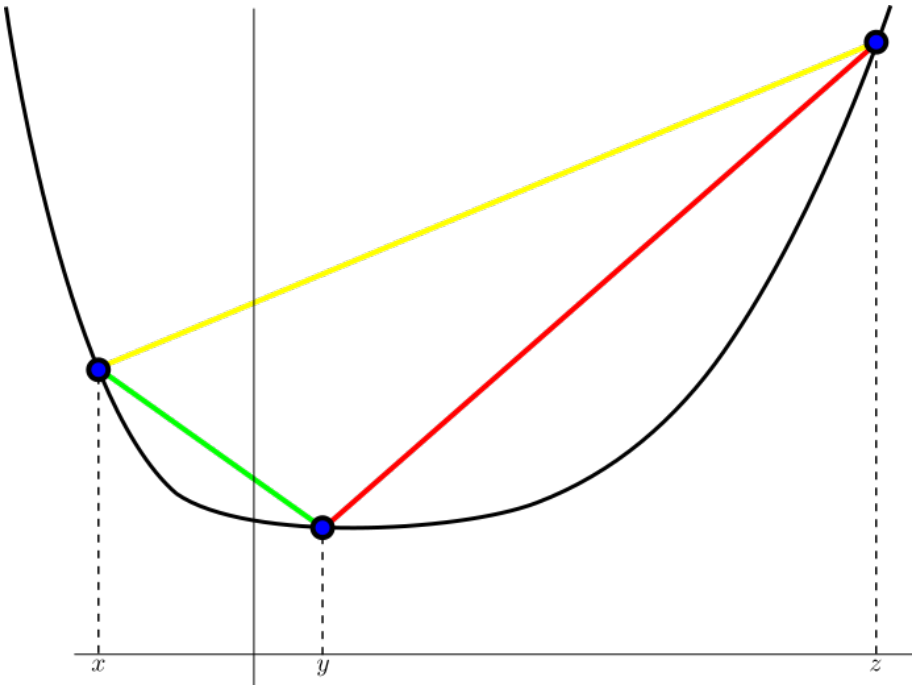
2. Le théorème précédent est-il une équivalence ?
3. Donner une autre démonstration de l'inégalité de Jensen en utilisant un raisonnement par récurrence.

Correction.

1. On applique l'inégalité de Jensen à la fonction convexe $x \mapsto x^2$ (voire Exercice 3) et à la famille de points pondérés $((x_i, \frac{1}{n}))_{1 \leq i \leq n}$. On multiplie ensuite l'inégalité obtenue par n^2 pour obtenir le résultat.
2. Oui, il suffit de remarquer que l'inégalité de Jensen pour $n = 2$ correspond à la définition de la convexité.

3. Inégalité des pentes

On établit dans ce paragraphe l'inégalité des pentes des cordes d'une fonction convexe : en examinant la figure ci-dessous, on peut remarquer que l'ordre des valeurs des coefficients directeurs (= des pentes) des trois cordes représentées ne dépend que de l'ordre des abscisses de leurs extrémités. Cette configuration usuelle peut servir de moyen mnémotechnique pour retenir l'inégalité des pentes.



Théorème 4. *Inégalité des pentes*

Soit $f : I \to \mathbb{R}$ une fonction. f est convexe sur I si, et seulement si, pour tous $x, y, z \in I$ avec $x < y < z$, on a :

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

Démonstration.

- ($\Rightarrow$). On suppose f convexe sur I.
  Soit $x, y, z \in I$ avec $x < y < z$. Alors $t = \frac{z-y}{z-x} \in [0, 1]$ vérifie $y = tx + (1 - t)z$ (voire Exercice 2) et la fonction f est convexe donc :

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

  On obtient alors $f(y) - f(z) \leq \frac{z-y}{z-x}(f(x) - f(z))$ ce qui donne la deuxième inégalité de l'énoncé.

  Pour la première inégalité, on procède de même. On a :

  $$f(y) - f(x) \leq (t - 1)f(x) + (1 - t)f(z) = (1 - t)(f(z) - f(x)) = \frac{y - x}{z - x}(f(z) - f(x)).$$

- ($\Leftarrow$). On suppose l'inégalité vraie pour tous $x, y, z \in I$ avec $x < y < z$.
  Soit $x, z \in I$ avec $x \leq z$, $t \in [0, 1]$ et $y = tx + (1 - t)z$. Montrons que $f(y) \leq tf(x) + (1 - t)f(z)$.
  Si $x = z$ ou si $t = 0$ ou 1, l'inégalité de convexité est vérifiée. On suppose donc $x < y < z$ (on remarque alors que $t = \frac{z-y}{z-x}$). La deuxième inégalité nous donne :

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

  d'où $f(y) \leq \frac{y-x}{z-x}(f(z) - f(x)) + f(x) = (1 - t)(f(z) - f(x)) + f(x) = tf(x) - (1 - t)f(z)$.
  Par suite, f est convexe sur I.

□

On établit alors la proposition suivante :

Proposition 3.

Soit $f : I \to \mathbb{R}$ une fonction. f est convexe sur I si, et seulement si, pour tout $x_0 \in I$, l'application :

$$\tau_{x_0} : \left| \begin{array}{lcl} I \setminus \{x_0\} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{f(x) - f(x_0)}{x - x_0}. \end{array} \right.$$

est croissante.

Démonstration.

- ($\Rightarrow$). On suppose f convexe sur I.
  Soit $x_0 \in I$. Soit $x, y \in I \setminus \{x_0\}$ avec $x < y$. On obtient immédiatement le résultat en appliquant l'inégalité des pentes aux trois cas possibles suivants : $a < x$, $x < a < y$ et enfin $a < y$.
- ($\Leftarrow$). On suppose que pour tout x_0, τ_{x_0} est croissante.
  Soit $x, y \in I$ et $t \in [0, 1]$. On pose $x_0 = tx + (1-t)y$. Si $x = y$ ou $t = 0$ ou 1, l'inégalité de convexité pour f est clairement vérifiée. On suppose $x < y$ et $t \in]0, 1[$ (et donc $x < x_0 < y$). La fonction τ_{x_0} est croissante donc on a :

  $$\frac{f(x_0) - f(x)}{x_0 - x} \leq \frac{f(y) - f(x_0)}{y - x_0}.$$

  Or $x_0 - x = (1-t)(y-x)$ et $y - x_0 = t(y-x)$, donc

  $$t(f(x_0) - f(x)) \leq (1-t)(f(y) - f(x_0)) \text{ ce qui donne } f(x_0) \leq tf(x) + (1-t)f(y).$$

  Il en résulte que f est convexe sur I.

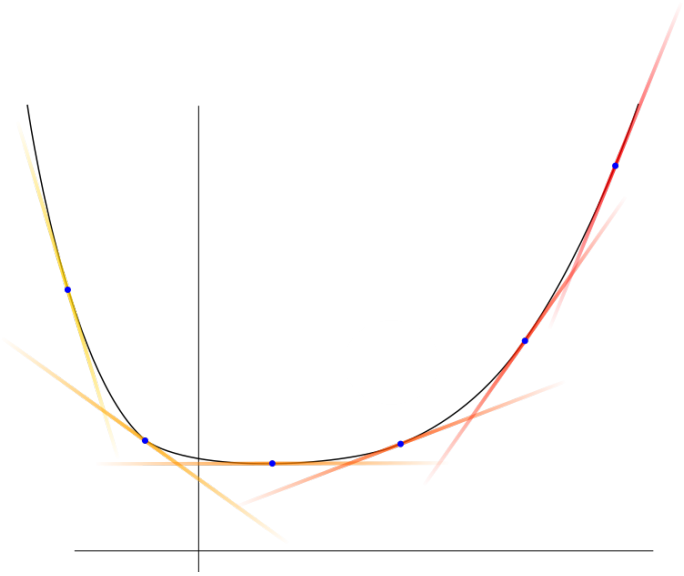
$\square$

Partie C

Convexité et dérivabilité

Grâce aux résultats obtenus dans la partie précédente, on va pouvoir obtenir des caractérisations de la convexité plus "maniables" en ajoutant des hypothèses de régularité à nos fonctions.

1. Fonctions convexes dérivables



Théorème 5.

Soit $f : I \to \mathbb{R}$ une fonction *dérivable* sur I. f est convexe sur I si, et seulement si, f' est croissante sur I.

Démonstration.

- ($\Rightarrow$). On suppose f convexe sur I.
  Soit $x, z \in I$ avec $x < z$. D'après l'inégalité des pentes (théorème 4), pour tout $y \in I$ avec $x < y < z$, on a :

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

  En passant à la limite dans l'inégalité $(*)$ avec $y \to x$, on obtient :

$$f'(x) \leq \frac{f(z) - f(x)}{z - x},$$

puis en prenant $y \to z$ dans $(*)$, on obtient :

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

Par suite, $f'(x) \leq f'(z)$. Donc f' est croissante sur I.

- ($\Leftarrow$). On suppose que f' est croissante.
  Soit $x, z \in I$ et $t \in [0,1]$. On pose $y = tx + (1-t)z$ (et donc $t = \frac{z-y}{z-x}$). Comme d'habitude, l'inégalité de convexité pour f est triviale pour $x = z$ ou $t = 0$ ou 1. On suppose donc $x < y < z$. Alors, f étant dérivable sur $[x, y]$ et sur $[y, z]$, d'après le théorème des accroissements finis, il existe $c_1 \in]x, y[$ et $c_2 \in]y, z[$ tels que :

  $$\frac{f(y)-f(x)}{y-x} = f'(c_1) \text{ et } \frac{f(z)-f(y)}{z-y} = f'(c_2).$$

  Or f' est croissante, donc $\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(y)}{z-y}$. Par suite,

  $$t(f(y)-f(x)) = \frac{1}{z-x}(z-y)(f(y)-f(x)) \leq \frac{1}{z-x}(y-x)(f(z)-f(y)) = (1-t)(f(z)-f(y)).$$

  D'où $f(y) \leq tf(x) + (1-t)f(z)$.
  Il en résulte que f est convexe sur I.

□

On en déduit la position relative du graphe d'une fonction convexe par rapport à ses tangentes :

Proposition 4.

Soit $f : I \to \mathbb{R}$ une fonction dérivable sur I. Si f est convexe sur I, alors son graphe est au dessus de chacune de ses tangentes i.e. pour tout $x_0 \in I$, on a :

$$\forall x \in I, \quad f(x) \geq f'(x_0)(x - x_0) + f(x_0).$$

Démonstration.

On suppose f convexe. Soit $x_0 \in I$. La tangente T en x_0 à pour équation :

$$T : y = f'(x_0)(x - x_0) + f(x_0).$$

On considère la fonction $g : I \to \mathbb{R}$ définie par $g(x) = f(x) - (f'(x_0)(x - x_0) + f(x_0))$, pour $x \in I$. La fonction g est dérivable sur I comme différence de deux fonctions dérivables sur I et pour tout $x \in I$, $g'(x) = f'(x) - f'(x_0)$.
D'après le théorème 5, f' est croissante sur I donc g' l'est aussi. De plus, $g'(x_0) = 0$ donc, par croissance de g', on a :

$$\forall x \leq x_0, \quad g'(x) \leq 0 \text{ et } \forall x \geq x_0, \quad g'(x) \geq 0.$$

Par suite, g admet un minimum global en x_0, d'où, pour tout $x \in I$, $g(x) \geq g(x_0) = 0$. Il en résulte que pour tout $x_0 \in I$, quelque soit $x \in I$, $f(x) \geq f'(x_0)(x - x_0) + f(x_0)$. □

2. Fonctions convexes deux fois dérivables

On énonce ici le théorème qui, dans la pratique, est le plus couramment utilisé pour démontrer la convexité d'une fonction quand on a des hypothèses suffisantes de régularité (et c'est souvent le cas !).

Théorème 6. *Caractérisation de la convexité par la dérivée seconde*

Soit $f : I \to \mathbb{R}$ une fonction deux fois dérivable sur I. f est convexe si, et seulement si, f'' est positive i.e.

$$\forall x \in I, \quad f''(x) \geq 0.$$

Démonstration.

On applique le théorème 5 en remarquant que pour f une fonction deux fois dérivable sur I, il y a équivalence entre f' croissante et f'' positive. □

Exercice 6.

Traduire toutes les caractérisations précédemment établies dans le cas *concave*.

3. Exemples d'inégalités classiques

Grâce aux outils développés autour de la notion de convexité dans ce chapitre, il devient aisé d'établir des inégalités qui auraient pu demander des efforts plus fournis pour être prouvées.

Exemple 4.

1. $\forall x \in \mathbb{R}$, $e^x \geq 1 + x$.
2. $\forall x \in \mathbb{R}_+^*$, $\ln(x) \leq x - 1$
3. $\forall x \in [0, \pi/2]$, $\frac{2}{\pi}x \leq \sin(x) \leq x$.

1. On note $f : x \mapsto e^x$, fonction C^∞ sur $\mathbb{R}$. Alors, pour tout $x \in \mathbb{R}$, on a $f''(x) = e^x > 0$. Ainsi, f est une fonction convexe sur $\mathbb{R}$. Par suite, f étant dérivable sur $\mathbb{R}$, le graphe de f est au dessus de sa tangente en 0, d'où :
$$e^x = f(x) \geq f'(0)(x-0) + f(0) = x + 1.$$

2. On note $f : x \mapsto \ln(x)$, fonction C^∞ sur $\mathbb{R}_+^*$. Alors, pour tout $x \in \mathbb{R}_+^*$, on a $f''(x) = -\dfrac{1}{x^2} < 0$. Ainsi, f est une fonction concave sur $\mathbb{R}_+^*$. Par suite, f étant dérivable sur $\mathbb{R}_+^*$, le graphe de f est en dessous de sa tangente en 1, d'où :
$$\ln(x) = f(x) \leq f'(1)(x-1) + f(1) = x - 1.$$

3. On note $f : x \mapsto \sin(x)$, fonction C^∞ sur $\mathbb{R}$ et en particulier sur $\left[0, \frac{\pi}{2}\right]$. Alors, pour tout $x \in \left[0, \frac{\pi}{2}\right]$, on a $f''(x) = -\sin(x) < 0$. Ainsi, f est une fonction concave sur $\left[0, \frac{\pi}{2}\right]$. Par suite, f étant dérivable sur $\left[0, \frac{\pi}{2}\right]$, le graphe de f est en dessous de sa tangente en 0 et au dessus de sa corde entre 0 et $\frac{\pi}{2}$, d'où :
$$\sin(x) = f(x) \leq f'(0)(x-0) + f(0) = x;$$
et, la corde étant d'équation $y = \dfrac{sin(\pi/2) - sin(0)}{\pi/2 - 0}x + 0 = \dfrac{2}{\pi}x$,
$$\sin(x) \geq \frac{2}{\pi}x.$$

Exemple 5. *Moyenne géométrique VS moyenne arithmétique*

Soit $n \in \mathbb{N}^*$ et $x_1, ..., x_n \in \mathbb{R}_+$. On a :
$$\sqrt[n]{x_1 ... x_n} \leq \frac{x_1 + ... + x_n}{n};$$
ou autrement écrit :
$$\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^{n} x_i.$$
En déduire que pour tout $n \in \mathbb{N}^*$,
$$\sqrt[n]{n!} \leq \frac{n+1}{2}.$$

Si au moins un des x_i est nul, alors l'inégalité est vérifiée.
On suppose que, pour tout $i \in [\![1, n]\!]$, $x_i \neq 0$. On remarque que $\sqrt[n]{x_1 ... x_n} > 0$ et on a :

$$\ln\left(\sqrt[n]{x_1 ... x_n}\right) = \ln\left(\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}\right) = \sum_{i=1}^{n} \frac{1}{n} \ln(x_i).$$

Le fonction ln étant concave sur $\mathbb{R}_+^*$, on applique la formule de Jensen à ln et la famille de points pondérée $(x_i, \frac{1}{n})_{1 \leq i \leq n}$ d'où :

$$\ln\left(\sum_{i=1}^{n} \frac{1}{n} x_i\right) \geq \sum_{i=1}^{n} \frac{1}{n} \ln(x_i) = \ln\left(\sqrt[n]{x_1 ... x_n}\right).$$

On obtient alors le résultat voulu en appliquant la fonction croissante $x \mapsto e^x$ à l'inégalité précédente.

Pour obtenir la dernière inégalité, il suffit d'utiliser le résultat précédent avec $x_i = i$ et de remarquer que $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.