Chapitre 16

Convexité

Plan du chapitre

Hypothèse

Dans tout ce chapitre, I est un intervalle de $\mathbb{R}$ non vide et non réduit à un point.

1 Généralités

1.1 Définition

Lemme 16.1 – Barycentre de deux points

Soit $x, y \in I$ tels que $x \leq y$. On a :

$$\left[x, y\right] = \left\{ \alpha x + (1-\alpha) y \mid \alpha \in \left[0,1\right] \right\}$$

Dit autrement, si on note $u_\alpha := \alpha x + (1-\alpha)y$, le point u_α parcourt le segment $\left[x,y\right]$ lorsque α parcourt $\left[0,1\right]$:

Le point u_α est appelé barycentre des points (x, y) pondéré par les poids $(\alpha, 1-\alpha)$ (notion techniquement hors-programme).

Définition 16.2 – Fonction convexe

Soit $f : I \to \mathbb{R}$ une fonction. On dit que f est convexe sur I si

$$\forall x, y \in I \qquad \forall \alpha \in \left[0, 1\right] \qquad f\left(\alpha x + (1-\alpha) y\right) \leq \alpha f(x) + (1-\alpha) f(y)$$

ou encore, ce qui est équivalent,

$$\forall x, y \in I \qquad \forall \alpha \in \left]0, 1\right[\qquad x < y \implies f\left(\alpha x + (1-\alpha) y\right) \leq \alpha f(x) + (1-\alpha) f(y)$$

Si on note $u_\alpha = \alpha x + (1-\alpha) y$, cela revient donc à dire que $f(u_\alpha) \leq \alpha f(x) + (1-\alpha) f(y)$.
Justifions cette équivalence. Il est clair que la première définition ci-dessus entraine la seconde. Montrons le sens réciproque. Soit $x, y \in I$ et $\alpha \in \left[0, 1\right]$. Posons $P(x, y, \alpha)$ l'assertion suivante :

$$P(x, y, \alpha) : \qquad f\left(\alpha x + (1-\alpha) y\right) \leq \alpha f(x) + (1-\alpha) f(y)$$

et supposons que $P(x, y, \alpha)$ est vraie lorsque $x < y$ et $0 < \alpha < 1$.

- Tout d'abord, $P(x, y, \alpha)$ est trivialement vraie lorsque $\alpha = 0$, $\alpha = 1$ ou $x = y$. Ainsi, $P(x, y, \alpha)$ est vraie lorsque $x \leq y$ et $0 \leq \alpha \leq 1$.
- Il reste à montrer $P(x, y, \alpha)$ lorsque $x > y$ (et $0 \leq \alpha \leq 1$). Or, comme $y < x$, on en déduit que $P(y, x, 1-\alpha)$ est vraie. De plus, comme $1 - (1-\alpha) = \alpha$, on en déduit :

$$\begin{aligned} P(y, x, 1-\alpha) &\iff f\left((1-\alpha) y + \alpha x\right) \leq (1-\alpha) f(y) + \alpha f(x) \\ &\iff f\left(\alpha x + (1-\alpha) y\right) \leq \alpha f(x) + (1-\alpha) f(y) \\ &\iff P(x, y, \alpha) \end{aligned}$$

  Ainsi, $P(x, y, \alpha)$ est vraie également dans le cas $x > y$. Les deux définitions données sont bien équivalentes.

1.2 Inteprétation de la convexité via les cordes

Définition 16.3

Soit $f : I \to \mathbb{R}$ une fonction. On appelle corde de $\mathscr{C}_f$ tout segment qui relie deux points distincts de la courbe $\mathscr{C}_f$.

Dit autrement, pour tous $x, y \in I$ distincts, si on note A le point de coordonnées $(x, f(x))$ et B le point de coordonéees $(y, f(y))$, alors le segment $\left[AB\right]$ est une corde de $\mathscr{C}_f$.

Remarque. On suppose $x < y$. La corde qui relie $(x, f(x))$ à $(y, f(y))$ est une droite : c'est donc la représentation graphique d'une fonction affine. On pose g la fonction :

$$\begin{aligned} g : \left[x, y\right] &\to \mathbb{R} \\ t &\mapsto \frac{f(y) - f(x)}{y - x}(t - x) + f(x) \end{aligned}$$

On vérifie que g est une fonction affine, $g(x) = f(x)$ et $g(y) = f(y)$. Ainsi la courbe $\mathscr{C}_g$ correspond à la corde qui relie $(x, f(x))$ à $(y, f(y))$. On notera que la pente de cette corde est donc $\dfrac{f(y) - f(x)}{y - x}$.

Théorème 16.4 – Position par rapport aux cordes

Une fonction f est convexe si et seulement si sa courbe $\mathscr{C}_f$ est *en-dessous* de toute corde qui relie deux de ses points.

Démonstration. Soit $x, y \in I$ tels que $x < y$. Par arbitraire sur x et y, il suffit de montrer que l'assertion

$$P(x,y): \quad \forall \alpha \in [0,1] \qquad f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

équivaut à l'assertion

$$Q(x,y): \quad \text{“}\mathscr{C}_f \text{ est en-dessous de la corde qui relie } (x, f(x)) \text{ à } (y, f(y))\text{”}$$

Or, avec g la fonction de la remarque ci-dessus, l'assertion $Q(x,y)$ se réécrit :

$$Q(x,y): \quad \forall t \in [x,y] \quad f(t) \leq g(t)$$

ou encore, par le Lemme 16.1, et en posant $u_\alpha = \alpha x + (1-\alpha)y$

$$Q(x,y): \quad \forall \alpha \in [0,1] \quad f(u_\alpha) \leq g(u_\alpha)$$

Cependant, on a $f(u_\alpha) = f(\alpha x + (1-\alpha)y)$ tandis que :

On remarque alors que

$$\begin{aligned} Q(x,y) &\iff \forall \alpha \in [0,1] \quad f(u_\alpha) \leq g(u_\alpha) \\ &\iff \forall \alpha \in [0,1] \quad f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) \\ &\iff P(x,y) \end{aligned}$$

D'où le résultat. □

Exemple 1. Pour démontrer ces exemples, on se contente de la représentation graphique pour le moment :
- Les fonctions suivantes sont convexes : $x \mapsto x^2$; $x \mapsto e^x$ et $x \mapsto e^{-x}$.
- La fonction $x \mapsto -x^2$ n'est pas convexe.

Exemple 2. Montrer que la fonction $f : x \mapsto |x|$ est convexe (sur $\mathbb{R}$).

1.3 Fonction concave

Définition 16.5 – Fonction concave

Soit $f : I \to \mathbb{R}$ une fonction. On dit que f est concave si $-f$ est convexe, c'est-à-dire :

$$\forall x, y \in I \qquad \forall \alpha \in [0,1] \qquad f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha) f(y)$$

Théorème 16.6

Une fonction f est concave si et seulement si sa courbe $\mathscr{C}_f$ est *au-dessus* de toute corde qui relie deux de ses points.

Exemple 3. On prouve ces exemples grâce à la représentation graphique pour le moment :

- Les fonctions $\quad x \mapsto -x^2 \quad$ et $\quad x \mapsto \ln x \quad$ sont concaves.
- Toute fonction affine (en particulier toute fonction constante) est convexe et concave.
- La fonction $x \mapsto x^3$ n'est ni convexe, ni concave.

1.4 Inégalité de Jensen

Lemme 16.7 – Barycentre de n points

Soit $n \in \mathbb{N}^*$ et $\alpha_1, \cdots, \alpha_n \in \mathbb{R}_+$ tels que $\sum_{i=1}^{n} \alpha_i = 1$. Soit enfin $x_1, \cdots, x_n \in I$. Alors

$$\sum_{i=1}^{n} \alpha_i x_i \in \Big[\min(x_1, \cdots, x_n) \quad , \quad \max(x_1, \cdots, x_n)\Big]$$

En particulier, $\sum_{i=1}^{n} \alpha_i x_i \in I$.

Démonstration. On pose $m = \min(x_1, \cdots, x_n)$ et $M = \max(x_1, \cdots, x_n)$. Montrons que $m \leq \sum_{i=1}^{n} \alpha_i x_i \leq M$. Il est clair que pour tout $i \in [\![1, n]\!]$,

$$m \leq x_i \leq M \qquad \text{donc} \qquad \alpha_i m \leq \alpha_i x_i \leq \alpha_i M$$

On somme la dernière inégalité pour i allant de 1 à n :

$$m \sum_{i=1}^{n} \alpha_i \leq \sum_{i=1}^{n} \alpha_i x_i \leq M \sum_{i=1}^{n} \alpha_i$$

Comme $\sum_{i=1}^{n} \alpha_i = 1$, on en déduit que $m \leq \sum_{i=1}^{n} \alpha_i x_i \leq M$. □

Le point $\sum_{i=1}^{n} \alpha_i x_i$ est appelé le barycentre des points $(x_i)_{1<i<n}$, pondérés par les poids $(\alpha_i)_{1<i<n}$ (hors-programme).

Théorème 16.8 – Inégalité de Jensen

Soit $n \in \mathbb{N}^*$ et $x_1, \cdots, x_n \in I$. Soit $\alpha_1, \cdots, \alpha_n \in \mathbb{R}_+$ tels que $\sum_{i=1}^{n} \alpha_i = 1$.

- Si $f : I \to \mathbb{R}$ est convexe, alors
$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$
- Si $f : I \to \mathbb{R}$ est concave, alors
$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \geq \sum_{i=1}^{n} \alpha_i f(x_i)$$

Attention à bien vérifier que les $\alpha_1, \cdots, \alpha_n$ soient tous positifs et que leur somme fasse 1. **Cette inégalité est très souvent utilisée en prenant $\alpha_1 = \alpha_2 = \ldots = \alpha_n = \frac{1}{n}$**, auquel cas la conclusion est :

$$f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \quad \left(\begin{matrix}\geq \\ \leq\end{matrix}\right) \quad \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

La formule pour une fonction concave f se déduit en appliquant la première formule à la fonction convexe $-f$, puis à multiplier par -1, ce qui change le sens de l'inégalité. Plus généralement, les propriétés qui suivent des fonctions convexes ont leur équivalent pour les fonctions concaves en changeant le sens des inégalités où intervient f.

Exemple 4. Soit $n \in \mathbb{N}^*$ et $x_1, \cdots, x_n > 0$. Montrer que

$$\sqrt[n]{\prod_{k=1}^{n} x_k} \leq \frac{1}{n} \sum_{k=1}^{n} x_k$$

Exemple 5. Soit $n \in \mathbb{N}^*$ et $x_1, \cdots, x_n > 0$. Montrer que

$$\frac{n}{\sum_{k=1}^{n} \frac{1}{x_k}} \leq \sqrt[n]{\prod_{k=1}^{n} x_k}$$

1.5 Caractérisation par l'inégalité des pentes

Théorème 16.9 – Inégalité des pentes

Soit $f : I \to \mathbb{R}$ une fonction

- f est convexe si et seulement si pour tous $x, y, z \in I$,

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

- Si f est concave si et seulement si pour tous $x, y, z \in I$,

$$x < y < z \implies \frac{f(y)-f(x)}{y-x} \geq \frac{f(z)-f(x)}{z-x} \geq \frac{f(z)-f(y)}{z-y}$$

Démonstration. On ne fait la preuve que pour f convexe. Seule la preuve du sens direct est au programme.

□

Remarque. En réalité, pour le sens réciproque, il est suffisant de vérifier une seule des inégalités. Par exemple :

$$\left(\forall x,y,z \in I \qquad x<y<z \implies \frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x}\right) \implies f \text{ est convexe}$$

Pour le prouver, on peut procéder comme pour la preuve du sens réciproque du Théorème 16.11, cf plus loin.

1.6 Caractérisation par le taux d'accroissement

Rappel : pour tout $a \in I$, le taux d'accroissement de f en a, est l'application qu'on notera :

$$\begin{aligned} \tau_a : I \setminus \{a\} &\to \mathbb{R} \\ x &\mapsto \frac{f(x)-f(a)}{x-a} \end{aligned}$$

Corollaire 16.10

Soit $f : I \to \mathbb{R}$ une fonction.

- f est convexe si et seulement si pour tout $a \in I$, τ_a est croissante (sur $I \setminus \{a\}$).
- f est concave si et seulement si pour tout $a \in I$, τ_a est décroissante (sur $I \setminus \{a\}$).

Démonstration. On ne prouve que le premier point. Par ce qui précède, il suffit de montrer que f vérifie l'inégalité des pentes (pour une fonction convexe) si et seulement si pour tout $a \in I$, τ_a est croissante.

- Supposons que τ_a est croissante pour tout $a \in I$. Alors pour tous $x,y,z \in I$

$$\begin{aligned} x<y<z &\implies \tau_x(y) \leq \tau_x(z) \\ &\implies \frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} \end{aligned}$$

  et donc f vérifie une des inégalités des pentes pour une fonction convexe donc est convexe (cf Remarque après le Théorème 16.9).

- Réciproquement, supposons f convexe, et montrons que pour tout $a \in I$, τ_a est croissante. Soit $x,y \in I \setminus \{a\}$ avec $x<y$. Montrons que $\tau_a(x) \leq \tau_a(y)$. On distingue 3 cas :
  - Si $a<x<y$, alors par l'inégalité des pentes,

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

– Si $x < a < y$, alors par l'inégalité des pentes,

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

– Si $x < y < a$, alors par l'inégalité des pentes,

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

Par arbitraire sur x et y, la fonction τ_a est donc croissante.

□

2 Convexité et dérivabilité

2.1 Caractérisation par la dérivée

Théorème 16.11 – Convexité et dérivée première

Soit $f : I \to \mathbb{R}$ une fonction dérivable.

- f est convexe si et seulement si f' est croissante.
- f est concave si et seulement si f' est décroissante.

Démonstration. On ne prouve que le premier point, en procédant par double implication.

Sens direct : supposons que f est convexe. Soit $x, z \in I$ tels que $x < z$. Montrons que $f'(x) \leq f'(z)$. Soit y un point quelconque de $]x, z[$. Comme f est convexe, on déduit de l'inégalité des pentes :

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

Comme f est dérivable en x, on peut passer à la limite quand y tend vers x (en laissant x, z fixés). On obtient $f'(x) \leq \frac{f(z)-f(x)}{z-x}$. Or, par l'inégalité des pentes, on a aussi

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

Et en passant à la limite quand y tend vers z (en laissant x, z fixés), on obtient $\frac{f(z)-f(x)}{z-x} \leq f'(z)$. Ainsi,

$$f'(x) \leq \frac{f(z)-f(x)}{z-x} \leq f'(z) \qquad \text{donc} \qquad f'(x) \leq f'(z)$$

Par arbitraire sur x, z, on en déduit que f' est croissante.

$\square$

Remarque. La preuve du sens direct ci-dessus montre en particulier que si f est convexe et dérivable, alors

$$\forall x,z \in I \qquad x < z \implies \left(f'(x) \leq \frac{f(z)-f(x)}{z-x} \leq f'(z) \right)$$

Corollaire 16.12

Soit $f : I \to \mathbb{R}$ une fonction deux fois dérivable.

- f est convexe si et seulement si $f'' \geq 0$.
- f est concave si et seulement si $f'' \leq 0$.

Ceci permet de justifier facilement les Exemples 1 et 3.

2.2 Position par rapport à la tangente

Théorème 16.13

Soit $f : I \to \mathbb{R}$ une fonction dérivable.

- Si f est convexe, alors pour tout $a \in I$, la courbe $\mathscr{C}_f$ est *au-dessus* de sa tangente en a :

$$\forall x \in I \qquad f(x) \geq f'(a)(x-a) + f(a)$$

- Si f est concave, alors pour tout $a \in I$, la courbe $\mathscr{C}_f$ est *en-dessous* de sa tangente en a :

$$\forall x \in I \qquad f(x) \leq f'(a)(x-a) + f(a)$$

Démonstration. On ne prouve que le premier point. Soit $a \in I$.

- Soit $x > a$. comme f est convexe et dérivable sur $[a, x]$, par la Remarque sous la preuve du Théorème 16.11, on en déduit que

$$\frac{f(x) - f(a)}{x - a} \geq f'(a) \qquad \text{ou encore} \qquad f(x) \geq f'(a)(x - a) + f(a)$$

- Soit $x < a$. De même en appliquant la Remarque sur $[x, a]$, on en déduit que

$$\frac{f(a) - f(x)}{a - x} \leq f'(a) \qquad \text{ou encore} \qquad f(x) \geq f'(a)(x - a) + f(a)$$

Dans tous les cas, on a donc l'inégalité voulue. □

Exemple 6. Montrer que pour tout $x \in \mathbb{R}$, on a $e^x \geq 1 + x$.

3 Sécante

Définition 16.14

Soit $f : I \to \mathbb{R}$ une fonction. On appelle <u>sécante</u> de $\mathscr{C}_f$ toute droite qui passe par deux points distincts de $\mathscr{C}_f$.

Dit autrement, pour tous $a, b \in I$ distincts, si on note A le point de coordonnées $(a, f(a))$ et B le point de coordonéees $(b, f(b))$, alors la droite (AB) est une sécante de $\mathscr{C}_f$.

Théorème 16.15

Soit $f : I \to \mathbb{R}$ une fonction convexe et $a, b \in I$ avec $a < b$.

- Sur $[a, b]$, la courbe $\mathscr{C}_f$ est *en-dessous* de la sécante (AB).
- Sur $]-\infty, a] \cup [b, +\infty[$, la courbe $\mathscr{C}_f$ est *au-dessus* de la sécante (AB).

Le premier point est un résultat déjà connu car, sur $[a, b]$, la sécante (AB) coïncide avec la corde $[AB]$.
Bien entendu, ce résultat s'adapte aux fonctions concaves : si f est concave, la courbe $\mathscr{C}_f$ est *au-dessus* de la sécante (AB) sur $[a, b]$, mais *en-dessous* de la sécante (AB) sur $]-\infty, a] \cup [b, +\infty[$.

4 Méthodes pour les exercices

Les méthodes sont présentées dans le cas d'une fonction convexe, mais s'adaptent au cas d'une fonction concave.

Méthode

Pour montrer qu'une fonction f est convexe, on peut :

- Si f est deux fois dérivable, montrer que f'' est positive.
- Si f est dérivable, montrer que f' est croissante.
- Sinon, utiliser la définition.

Fausse Bonne Idée : éviter de montrer que τ_a est croissante ou que f vérifie l'inégalité des pentes.

Méthode

Pour montrer des inégalités du type $f(x)$ et une expression affine $\alpha x+\beta$:

- Si l'inégalité est de la forme $f(x) \geq \alpha x+\beta$ avec f convexe, on peut regarder si $y=\alpha x+\beta$ est l'équation d'une tangente de $\mathscr{C}_f$.
- Si l'inégalité est de la forme $f(x) \leq \alpha x+\beta$ avec f convexe, on peut regarder si $y=\alpha x+\beta$ est l'équation d'une corde de $\mathscr{C}_f$, notamment si cette inégalité n'est demandée que pour tout x dans un segment $[a,b]$.

Le second point peut également s'adapter pour une inégalité du type $f(x) \geq \alpha x+\beta$ en dehors du segement $[a,b]$, au moyen d'une sécante, mais son cas d'utilisation est plus rare.

Méthode

Pour montrer une inégalité faisant intervenir n points $x_1,\cdots,x_n$, il faut le plus souvent utiliser l'inégalité de Jensen. Le plus difficile est de trouver la bonne fonction f et les points $y_1,\cdots,y_n$ en lesquels il faut l'appliquer. On peut avoir $y_k \neq x_k$, par exemple $y_k=\dfrac{1}{x_k}$ comme à l'Exemple 5 !

Conseil : partir de l'inégalité à montrer et raisonner par équivalences / réécritures jusqu'à reconnaitre un terme de la forme $f\left(\dfrac{1}{n}\sum_{k=1}^{n} y_k\right)$ et/ou $\dfrac{1}{n}\sum_{k=1}^{n} f(y_k)$, ce qui nous guide vers la bonne fonction f ou les bons réels $y_1,\cdots,y_n$.