Chapitre 15

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 15.1 (Barycentre de deux points)

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

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

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

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

Définition 15.2 (Fonction convexe)

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

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

ou encore, ce qui est équivalent,

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

Il est clair que la première définition ci-dessus entraine la seconde. Montrons la réciproque. Posons $P(x, y, \alpha)$ l'assertion suivante :

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

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 $0 \leq \alpha \leq 1$.
- Supposons maintenant $x \geq y$. Alors, comme $y \leq x$, on en déduit que $P(y, x, 1-\alpha)$ est vraie. Or, comme $1 - (1-\alpha) = \alpha$,

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

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

Remarque (Interprétation graphique)**.** Les points x, y étant fixés. Quand α parcourt $[0,1]$,

- le point $u_\alpha := \alpha x + (1-\alpha)y$ parcourt le segment qui relie x à y.
- tandis que le point $v_\alpha := \alpha f(x) + (1-\alpha) f(y)$ parcourt le segment qui relie $f(x)$ à $f(y)$.

Propriété 15.3 (Position par rapport aux cordes)

Une fonction f est convexe si et seulement si sa courbe C_f est *en-dessous* de toute corde qui relie deux de ses points.

Autrement dit, pour tous $x, y \in I$ tels que $x < y$, si on note u_α et v_α les points ci-dessus, alors f est convexe si $\forall \alpha \in [0,1] \qquad f(u_\alpha) \leq v_\alpha$

Démonstration. Soit $a, b \in I$ tels que $a < b$. On pose g la fonction affine qui passe par les points $(x, f(x))$ et $(y, f(y))$. On a alors

$$\forall t \in [x,y] \qquad g(t) = \frac{f(y)-f(x)}{y-x}(t-x)+f(x)$$

Il suffit donc de montrer que pour tout $t \in [x,y]$, on a $f(t) \leq g(t)$. Cependant, vu le Lemme 15.1, il est suffisant de montrer que pour tout $\alpha \in [0,1]$, on a $f(u_\alpha) \leq g(u_\alpha)$, avec $u_\alpha := \alpha x + (1-\alpha)y$. Or,

$$\begin{aligned}
g(u_\alpha) &= g(\alpha x + (1-\alpha)y) \\
&= \frac{f(y)-f(x)}{y-x}\left[\alpha x + (1-\alpha)y - x\right] + f(x) \\
&= \frac{f(y)-f(x)}{y-x}\left[(1-\alpha)(y-x)\right] + f(x) \\
&= (f(y)-f(x))(1-\alpha) + f(x) \\
&= \alpha f(x) + (1-\alpha) f(y)
\end{aligned}$$

Ainsi, comme f est convexe, on a

$$\begin{aligned}
f(u_\alpha) &= f(\alpha x + (1-\alpha)y) \\
&\leq \alpha f(x) + (1-\alpha) f(y) = g(u_\alpha)
\end{aligned}$$

□

Exemple 1. On prouve ces exemples par interprétation graphique pour le moment :

- Les fonctions $x \mapsto |x|$ et $x \mapsto x^2$, ainsi que $x \mapsto e^{-x}$ sont convexes.
- La fonction $x \mapsto -x^2$ n'est pas convexe.

Définition 15.4 (Fonction concave)

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

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

Propriété 15.5

Une fonction f est concave si et seulement si sa courbe C_f est *au-dessus* de toute corde qui relie deux de ses points.

Exemple 2. On prouve ces exemples par interprétation graphique pour le moment :

- Les fonctions $x \mapsto -x^2$ et $x \mapsto \ln x$ 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.

1.2 Inégalité de Jensen

Lemme 15.6 (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

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

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 :

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

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 $(\alpha_i)_{1<i<n}$.

Propriété 15.7 (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 $\alpha_1 = \alpha_2 = \cdots = \alpha_n = \frac{1}{n}$.

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 3. Soit $n \in \mathbb{N}^*$ et $x_1, \cdots, x_n \in \mathbb{R}$. Montrer que $(x_1 + \cdots + x_n)^2 \leq n(x_1^2 + \cdots + x_n^2)$.

1.3 Caractérisation par l'inégalité des pentes

Propriété 15.8 (Inégalité des pentes)

Soit $f : I \to \mathbb{R}$ une fonction et $x, y, z \in I$.

- Si f est convexe, alors

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

- Si f est concave, alors

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

Démonstration. On ne fait la preuve que pour f convexe.

□

Ce résultat admet une réciproque : si $f : I \to \mathbb{R}$ vérifie l'inégalité des pentes alors f est convexe. Il est même suffisant de vérifier une seule des deux inégalités, par exemple

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

Pour le prouver, on peut procéder comme pour la preuve du sens réciproque de la Proposition 15.10, cf plus loin. Ce résultat fait partie du "programme gris".

Propriété 15.9 (Croissance du taux d'accroissement)

Soit $f : I \to \mathbb{R}$ une fonction. Pour tout $a \in I$, on pose τ_a le taux d'accroissement de f en a, càd l'application

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

- Si f est convexe, alors τ_a est croissante (sur $I \setminus \{a\}$).
- Si f est concave, alors τ_a est décroissante (sur $I \setminus \{a\}$).

Démonstration. On ne prouve que le premier point. Supposons f convexe. Soit $a \in I$ et $x, y \in I \setminus \{a\}$ tels que $x < y$. Montrons que $\tau_a(x) \leq \tau_a(y)$. On distingue 3 cas :

- Si $x < y < a$, alors par l'inégalité des pentes,
$$\tau_a(x) = \frac{f(a)-f(x)}{a-x} \leq \frac{f(a)-f(y)}{a-y} = \tau_a(y)$$
- Si $x < a < y$, alors par l'inégalité des pentes,
$$\tau_a(x) = \frac{f(a)-f(x)}{a-x} \leq \frac{f(y)-f(a)}{y-a} = \tau_a(y)$$
- Si $a < x < y$, alors par l'inégalité des pentes,
$$\tau_a(x) = \frac{f(x)-f(a)}{x-a} \leq \frac{f(y)-f(a)}{y-a} = \tau_a(y)$$

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

2 Convexité et dérivabilité

2.1 Caractérisation par la dérivée

Propriété 15.10 (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 $x, y, z \in I$ tels que $x < y < z$. On va montrer que $f'(x) \leq f'(z)$. Comme f est convexe, par 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 z fixé). On obtient $f'(x) \leq \dfrac{f(z)-f(x)}{z-x}$. 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 fixé), on obtient $\dfrac{f(z)-f(x)}{z-x} \leq f'(z)$. Ainsi, $f'(x) \leq f'(z)$. Par arbitraire sur x, z, on en déduit que f' est croissante.

$\square$

Corollaire 15.11

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 2.

2.2 Position par rapport à la tangente

Propriété 15.12

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

- Si f est convexe, alors pour tout $a \in I$, la courbe 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 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 $a \in I$. Montrons que pour tout $x \in I$, on a

$$f(x) \geq f'(a)(x-a) + f(a)$$

L'inégalité est évidemment vraie si $x = a$. Supposons donc $x \neq a$. Comme f est convexe, son taux d'accroissement en a, noté τ_a, est croissant sur $I \setminus \{a\}$. On distingue deux cas :

- Si $x < a$, alors pour tout $y \in]x, a[$, on a $\tau_a(x) \leq \tau_a(y)$. En passant à la limite quand y tend vers a^-, on a :

$$\tau_a(x) \leq \lim_{y \to a^-} \tau_a(y) = f'(a)$$

  D'où, en multipliant par $x - a < 0$, on en déduit que $f(x) - f(a) \geq f'(a)(x-a)$.
- Si $x > a$, alors pour tout $y \in]a, x[$, on a $\tau_a(y) \leq \tau_a(x)$. En passant à la limite quand y tend vers a^+, on a :

$$f'(a) = \lim_{y \to a^+} \tau_a(y) \leq \tau_a(x)$$

  D'où, en multipliant par $x - a > 0$, on en déduit que $f(x) - f(a) \geq f'(a)(x-a)$.

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

Exemple 4. Montrer que pour tout $x \in \mathbb{R}$, on a $e^x \geq 1 + x$.

Remarque. La propriété 15.12 montre que, si f et convexe et si a est un point critique de f, alors pour tout $x \in I$, on a $f(x) \geq f(a)$. Autrement dit, tout point critique d'une fonction convexe est un minimum global. De même, tout point critique d'une fonction concave est un maximum global.

3 Complément : sécante

Soit $f : I \to \mathbb{R}$ une fonction convexe, $a, b \in I$ avec $a < b$. On note A et B les points de coordonnées respectives $(a, f(a))$ et $(b, f(b))$. On a vu qu'alors le segment $[AB]$ est appelé une corde de la courbe C_f. La droite (AB) est appelée une sécante de C_f.
On a alors le résultat suivant :

- Sur l'intervalle $[a, b]$, la courbe C_f est en-dessous de la sécante (AB). Il s'agit d'un résultat déjà vu car, sur $[a, b]$, la sécante (AB) coïncide avec la corde $[AB]$.
- Sur $]-\infty, a] \cup [b, +\infty[$, la courbe C_f est au-dessus de la sécante (AB).

Bien entendu, ce résultat s'adapte aux fonctions concave : si f est concave, la courbe C_f est au-dessus de la sécante (AB) sur $[a, b]$, mais en-dessous de la sécante (AB) sur $]-\infty, a] \cup [b, +\infty[$.