Chapitre 8

Convexité (fin du précédent)

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 8.1 (Barycentre de deux points)

Pour tous $x, y \in I$, tels que $x \leq y$,

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

Par exemple, lorsque $\alpha = \frac{1}{2}$, le point $\alpha x + (1-\alpha)y = \frac{x+y}{2}$ est le milieu du segment $[x, y]$.
Lorsque α tend vers 1, $\alpha x + (1-\alpha)y$ tend vers x.
Lorsque α tend vers 0, $\alpha x + (1-\alpha)y$ tend vers y.

Définition 8.2 (Fonction convexe)

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

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

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 $y \leq x$, $P(y,x,1-\alpha)$ est vraie. Or,

$$P(y,x,1-\alpha) \iff P(x,y,\alpha)$$

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 $\alpha f(x) + (1-\alpha) f(y)$ parcourt le segment qui relie $f(x)$ à $f(y)$.

Proposition 8.3 (Position par rapport aux cordes)

Soit $f : I \to \mathbb{R}$ une fonction. f est convexe si et seulement si sa courbe C_f est *en-dessous* toute corde qui relie deux de ses points.
Autrement dit, pour tous $x, y \in I$ tels que $x < y$, si on note $d_{x,y}$ la fonction affine qui passe par les points $(x, f(x))$ et $(y, f(y))$, alors

$$\forall t \in [x,y] \qquad f(t) \leq d_{x,y}(t)$$

Idée de la preuve. Pour tout $\alpha \in [0,1]$, si on pose $u_\alpha := \alpha x + (1-\alpha)y$ alors on peut montrer que

$$d_{x,y}(u_\alpha) = \alpha f(x) + (1-\alpha) f(y)$$

et donc f est convexe si et seulement si, pour tous $x, y \in I$ et $\alpha \in [0,1]$, on a $f(u_\alpha) \leq d_{x,y}(u_\alpha)$. Or, quand α parcourt $[0,1]$, u_α parcourt $[x,y]$, si bien que cela revient à avoir $f(t) \leq d_{x,y}(t)$ pour tout $t \in [x,y]$. □

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 8.4 (Fonction concave)

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

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

En particulier, f est concave si et seulement si C_f est *au-dessus* de toutes ses cordes.

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 est convexe et concave. 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 8.5 (Barycentre de n points)

Soit $n \in \mathbb{N}^*$. Pour tous $\alpha_1, \ldots, \alpha_n \in \mathbb{R}_+$ tels que $\sum_{i=1}^{n} \alpha_i = 1$, et pour tous points $x_1, \ldots, x_n \in I$, on a

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

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

Démonstration. On pose $m = \min(x_1, \ldots, x_n)$ et $M = \max(x_1, \ldots, x_n)$. 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 $(x_i)_{1\leq i\leq n}$, pondérés par les poids $(\lambda_i)_{1\leq i\leq n}$.

Proposition 8.6 (Inégalité de Jensen)

Soient $n \in \mathbb{N}^*$ et $f : I \to \mathbb{R}$ une fonction. Si f est convexe, alors pour tous $\alpha_1, \dots, \alpha_n \in \mathbb{R}_+$ tels que $\sum_{i=1}^{n} \alpha_i = 1$, et pour tous points $x_1, \dots, x_n \in I$, alors

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

Cette inégalité est très souvent utilisée en prenant $\alpha_i = \frac{1}{n}$.

Exemple 3. Avec les hypothèses de la Proposition ci-dessus, montrer que $(x_1 + \cdots + x_n)^2 \leq n(x_1^2 + \cdots + x_n^2)$.

Remarque (Inégalité de Jensen concave)**.** Lorsque f est concave, on a

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

Autrement dit, il suffit de changer le sens de l'inégalité. De même, toutes les propriétés qu'on verra pour les fonctions convexes ont leur équivalent pour les fonctions concaves en changeant le sens des inégalités.

1.3 Caractérisation par l'inégalité des pentes

Proposition 8.7 (Inégalité des pentes)

Soit $f : I \to \mathbb{R}$ une fonction convexe. Alors pour tous $x, y, z \in I$

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

Démonstration.

□

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 8.9.

Proposition 8.8 (Croissance du taux d'accroissement)

Soit $f : I \to \mathbb{R}$ une fonction convexe. Alors pour tout $a \in I$, le taux d'accroissement de f en a est une fonction croissante, càd l'application

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

est croissante (sur $I \setminus \{a\}$).

Démonstration. Soient $a \in I$ 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,

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

□

2 Convexité et dérivabilité

2.1 Caractérisation par la dérivée

Proposition 8.9 (Convexité et dérivée première)

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

Démonstration. On procède par double implication.
Sens direct : supposons que f soit convexe. Soient $x, y, z \in I$ 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. 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, 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.

□

Corollaire 8.10

Soit $f : I \to \mathbb{R}$ une fonction deux fois dérivable. f est convexe si et seulement si $f'' \geq 0$.

Ceci permet de justifier facilement les Exemples 1 et 2.

2.2 Position par rapport à la tangente

Proposition 8.11

Soit $f : I \to \mathbb{R}$ une fonction convexe et dérivable. Alors pour tout $a \in I$, la courbe C_f est au-dessus de sa tangente en a :

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

Démonstration. Soit $a \in I$. Comme f est convexe, le taux d'accroissement τ_a est croissant sur $I \setminus \{a\}$. On distingue deux cas :

- Si $x < a$, alors par le théorème de la limite monotone,
$$\frac{f(x)-f(a)}{x-a} = \tau_a(x) \leq \lim_{x\to a^-} \tau_a(x) = f'(a)$$
et donc $\dfrac{f(x)-f(a)}{x-a} \leq f'(a)$ En multipliant par $x-a<0$, on obtient $f(x) \geq f(a) + f'(a)(x-a)$.
- Si $x > a$, alors par le théorème de la limite monotone,
$$f'(a) = \lim_{x\to a^+} \tau_a(x) \leq \tau_a(x) = \frac{f(x)-f(a)}{x-a}$$
et donc $\dfrac{f(x)-f(a)}{x-a} \geq f'(a)$ En multipliant par $x-a>0$, on obtient $f(x) \geq f(a) + f'(a)(x-a)$.

□

Remarque. En particulier, si a est un point critique de f, alors $f(x) \geq f(a)$. Autrement dit, tout point critique d'une fonction convexe est un minimum global.

Exemple 4. La fonction exp est convexe. Sa tangente en 0 a pour équation

$$y = e^0(x-0) + e^0 = x+1$$

Ainsi, pour tout $x \in \mathbb{R}$, on a $e^x \geq 1+x$.

3 Adaptation pour les fonctions concaves

Soit $f : I \to \mathbb{R}$ une fonction. On a vu la définition d'une fonction concave et l'inégalité de Jensen pour les fonctions concaves. À cela, il faut rajouter les résultats suivants :

- Si f est concave, alors pour tout $a \in I$ son taux d'accroissement τ_a est décroissant sur $I \setminus \{a\}$.
- Si f est concave, alors C_f est au-dessus de ses cordes.
- Si f est concave, alors C_f est en-dessous de ses tangentes (en les points où f est dérivable).
- Si f est dérivable, alors f est concave si et seulement si f' est décroissante.
- Si f est deux fois dérivable, alors f est concave si et seulement si $f'' \leq 0$.

4 Complément : sécante

Soient $f : I \to \mathbb{R}$ une fonction convexe, $a, b \in I$ avec $a < b$. On note A, B les points de coordonnées $(a, f(a))$ 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 $]-\infty, a] \cup [b, +\infty[$.