Lycée Henri IV
MPSI

Cours
Durée : 2 h

CONVEXITÉ

Les énoncés et parties suivis du symbole [✠] ne seront pas traités en cours.

MPSI – Lycée Henri IV – Sébastien Kerner

Prérequis

Revoir les chapitres sur :

- les fonctions ;
- le langage des limites ;
- la continuité ;
- la dérivabilité.

Dans tout ce chapitre, I désigne un intervalle non vide et non réduit à un point.

Les trois lettres AQT signifient « Âne Qui Trotte » et sont utilisées pour désigner une démonstration facile laissée au lecteur.

A. Convexité

A.1. Définitions

Définition 1

Une fonction $f : I \longrightarrow \mathbb{R}$ est dite convexe sur I lorsque

$$\forall x, y \in I, \quad \forall t \in [0;1]\,, \quad f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y).$$

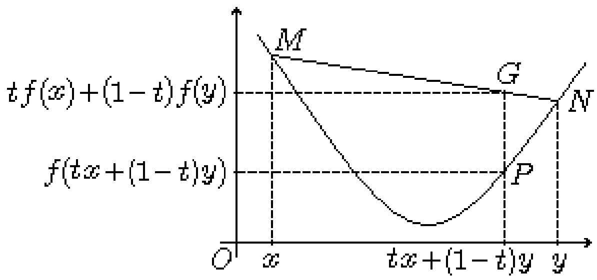
Une fonction $f : I \longrightarrow \mathbb{R}$ est dite concave sur I lorsque l'inégalité ci-dessus est vérifiée avec un $\geqslant$ à la place du $\leqslant$.

Notons que la convexité et la concavité sont intimement liées puisque f est convexe sur un intervalle si, et seulement si, $-f$ est concave sur le même intervalle.

Interprétation graphique de la convexité

Notons M et N les points de la courbe représentative de f d'abscisses respectives x et y. Le point G, barycentre du système $\{(M,t);(N,1-t)\}$, est alors de coordonnées $\big(tx + (1-t)y\,;\,tf(x) + (1-t)f(y)\big)$. Soit P le point de la courbe représentative de f d'abscisse $tx + (1-t)y$. La convexité de f revient donc à dire que l'ordonnée du point P est inférieure à celle du point G, c'est-à-dire que P est en dessous de G.

Une fonction est convexe si, et seulement si tout arc de sa courbe représentative est en dessous de la corde correspondante.



Pour une fonction concave, la propriété est inversée : tout arc est au dessus de sa corde.

On peut également définir la stricte convexité et la stricte concavité en remplaçant l'inégalité large par une inégalité stricte (et en imposant les conditions $x \neq y$ et $t \in \,]0;1[$).

Exemples :

- Toute fonction affine est à la fois convexe et concave sur $\mathbb{R}$. On peut même remarquer que ce sont les seules applications ayant cette double propriété.
- La fonction valeur absolue est convexe sur $\mathbb{R}$ puisque, en vertu de l'inégalité triangulaire, on a $\forall x, y \in \mathbb{R},\ \forall t \in [0;1]\,,\ |tx + (1-t)y| \leqslant t|x| + (1-t)|y|$.
- Une somme de fonctions convexes (respectivement concaves) sur I est une fonction convexe (respectivement concave) sur I.

Définition 2

Soit $x_0 \in I$. On dit que x_0 est un point d'inflexion de la fonction $f : I \longrightarrow \mathbb{R}$ lorsque f est de convexité différente à droite et à gauche de x_0 (c'est-à-dire convexe au voisinage de ${x_0}^+$ et concave au voisinage de ${x_0}^-$ ou l'inverse).

A.2. Inégalité de Jensen

La proposition suivante établit l'inégalité de Jensen, qui est une généralisation de l'inégalité de convexité.

Proposition 1

Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe sur I. Pour tout $n \geqslant 1$, tous $x_1, \ldots, x_n \in I$ et tous $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_+$ tels que $\lambda_1 + \cdots + \lambda_n = 1$, on a

$$f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leqslant \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n).$$

Si f est concave sur I, l'inégalité ci-dessus est vérifiée avec un signe $\geqslant$ à la place de $\leqslant$.

■ Nous démontrons le résultat par récurrence dans le cas d'une fonction convexe.

Initialisation : Le cas $n = 1$ est trivial puisqu'alors $\lambda_1 = 1$. Le cas $n = 2$ correspond à la définition de la convexité avec $\lambda_1 = t$ et $\lambda_2 = 1 - t$.

Hérédité : Soit $n \geqslant 2$. Supposons le résultat vrai au rang n et démontrons le au rang $n+1$.
Soient $y_1, \ldots, y_n, y_{n+1} \in I$ et $\mu_1, \ldots, \mu_n, \mu_{n+1} \in \mathbb{R}_+$ tels que $\mu_1 + \cdots + \mu_n + \mu_{n+1} = 1$.
Si $\mu_{n+1} = 1$, alors $\mu_1 = \cdots = \mu_n = 0$ et la propriété au rang $n+1$ est clairement vraie.
On suppose donc dorénavant que $\mu_{n+1} \neq 1$. On écrit alors

$$f\Big(\sum_{i=1}^{n+1} \mu_i y_i\Big) = f\Big((1 - \mu_{n+1}) \sum_{i=1}^{n} \frac{\mu_i}{1 - \mu_{n+1}} y_i + \mu_{n+1} y_{n+1}\Big)$$

et on utilise la propriété de convexité de f, ce qui donne

$$f\Big(\sum_{i=1}^{n+1} \mu_i y_i\Big) \leqslant (1 - \mu_{n+1}) f\Big(\sum_{i=1}^{n} \frac{\mu_i}{1 - \mu_{n+1}} y_i\Big) + \mu_{n+1} f(y_{n+1}).$$

On constate alors que $\displaystyle\sum_{i=1}^{n} \frac{\mu_i}{1 - \mu_{n+1}} = 1$, ce qui permet d'appliquer l'hypothèse de récurrence :

$$f\Big(\sum_{i=1}^{n+1} \mu_i y_i\Big) \leqslant (1 - \mu_{n+1}) \sum_{i=1}^{n} \frac{\mu_i}{1 - \mu_{n+1}} f(y_i) + \mu_{n+1} f(y_{n+1})$$

d'où

$$f\Big(\sum_{i=1}^{n+1} \mu_i y_i\Big) = \sum_{i=1}^{n+1} \mu_i f(y_i).$$

On a ainsi démontré la propriété au rang $n+1$.

Conclusion : Le principe de récurrence permet de conclure. ■

On peut aussi énoncer l'inégalité de Jensen sans faire l'hypothèse que la somme des poids est égale à 1. On obtient alors le résultat suivant.

Corollaire 1

Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe sur I. Pour tout $n \geqslant 1$, tous $x_1, \ldots, x_n \in I$ et tous $\mu_1, \ldots, \mu_n \in \mathbb{R}_+$ tels que $\mu_1 + \cdots + \mu_n \neq 0$, on a

$$f\Big(\frac{\mu_1 x_1 + \cdots + \mu_n x_n}{\mu_1 + \cdots + \mu_n}\Big) \leqslant \frac{\mu_1 f(x_1) + \cdots + \mu_n f(x_n)}{\mu_1 + \cdots + \mu_n}.$$

Si f est concave sur I, l'inégalité ci-dessus est vérifiée avec un signe $\geqslant$ à la place de $\leqslant$.

■ Immédiat en appliquant la proposition précédente avec $\forall k \in [\![1;n]\!]$, $\lambda_k = \mu_k/(\mu_1 + \cdots + \mu_n)$. ■

A.3. Les lemmes des pentes

Voici le lemme de la pente croissante.

Lemme 1

Une fonction $f : I \longrightarrow \mathbb{R}$ est convexe (respectivement concave) sur I si et seulement si, pour tout $a \in I$, la fonction

$$\tau_a : x \longmapsto \frac{f(x) - f(a)}{x - a}$$

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

■ On démontre la propriété dans le cas d'une fonction convexe.

$\Rightarrow$ Supposons f convexe sur I. Soient $a \in I$ et $x, y \in I \backslash \{a\}$ tels que $x < y$. On distingue trois cas : ou bien $a < x < y$, ou bien $x < a < y$ ou encore $x < y < a$. On traite le cas où $a < x < y$. Les deux autres cas sont similaires. On écrit alors que $x = ta + (1-t)y$ pour $t \in [0;1]$. La convexité de f permet alors d'en déduire que

$$\begin{aligned}
\tau_a(x) &= \frac{f(x) - f(a)}{x - a} \\
&= \frac{f\big(ta + (1-t)y\big) - f(a)}{ta + (1-t)y - a} \\
&\leqslant \frac{tf(a) + (1-t)f(y) - f(a)}{(1-t)(y-a)} \\
&= \frac{(1-t)\big(f(y) - f(a)\big)}{(1-t)(y-a)} \\
&= \frac{f(y) - f(a)}{y - a} \\
&= \tau_a(y),
\end{aligned}$$

ce qui montre que la fonction τ_a est croissante sur $I \cap]a; +\infty[$.

$\Leftarrow$ Supposons réciproquement que, pour tout $a \in I$, la fonction τ_a est croissante sur $I \setminus \{a\}$. Soient $x, y \in I$ tel que $x < y$ et $t \in [0;1]$. On veut démontrer que $f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y)$. Les cas $t = 0$ et $t = 1$ sont triviaux. On suppose donc que $t \in]0;1[$. Dès lors, $tx + (1-t)y \in]x;y[$ et l'on peut écrire, en vertu de la croissance de la fonction τ_x, que $\tau_x\big(tx + (1-t)y\big) \leqslant \tau_x(y)$, c'est-à-dire

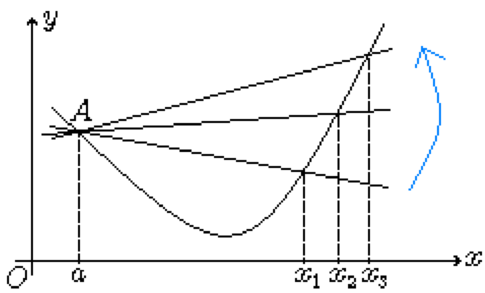
$$\frac{f\big(tx + (1-t)y\big) - f(x)}{\big(tx + (1-t)y\big) - x} \leqslant \frac{f(y) - f(x)}{y - x}.$$

Après transformation et simplification par $1-t$, on obtient $f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y)$. La fonction est donc bien convexe sur I. ■

Attention, il s'agit bien de la croissance de τ_a sur $I \setminus \{a\}$ et pas seulement de sa croissance sur chacun des deux intervalles $I \cap]-\infty; a[$ et $I \cap]a; +\infty[$ (Y réfléchir !).

Interprétation graphique du lemme de la pente croissante

Pour une fonction convexe, la pente d'une sécante dont on fixe une extrémité est une fonction croissante de l'autre extrémité.



On peut généraliser le lemme de le pente croissante. On obtient le lemme de la pente glissante.

Lemme 2

Une fonction $f : I \longrightarrow \mathbb{R}$ est convexe sur I si, et seulement si, pour tous $a, b, x, y \in I$ tels que $a \neq x$ et $b \neq y$, on a

$$(a \leqslant b \text{ et } x \leqslant y) \implies \big(\tau_a(x) \leqslant \tau_b(y)\big).$$

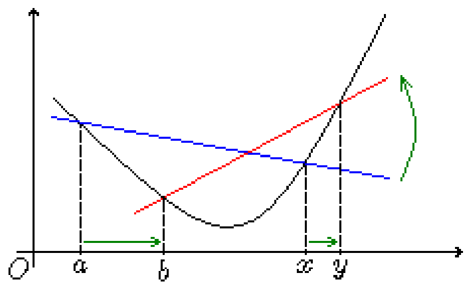
On obtient un résultat similaire pour les fonctions concaves en remplaçant $\leqslant$ par $\geqslant$.

■ On démontre la propriété dans le cas d'une fonction convexe.

- $\Rightarrow$ Supposons que f est convexe sur I. Soient $a, b, x, y \in I$ tels que $a \leqslant x$ et $b \leqslant y$.
  Si $a \neq y$, on écrit $\tau_a(x) \leqslant \tau_a(y) = \tau_y(a) \leqslant \tau_y(b) = \tau_b(y)$ où les inégalités découlent du lemme de la pente croissante,
  Si $a = y$, on écrit $\tau_a(x) = \tau_y(x) \leqslant \tau_y(b) = \tau_b(y)$ où l'inégalité découle du lemme de la pente croissante et du fait que $b \geqslant a = y \geqslant x$.
- $\Leftarrow$ Supposons réciproquement que, pour tous $a, b, x, y \in I$ tels que $a \neq x$ et $b \neq y$, on a l'implication $(a \leqslant b \text{ et } x \leqslant y) \implies \big(\tau_a(x) \leqslant \tau_b(y)\big)$. En choisissant $a = b$, on retrouve la croissance sur $I \setminus \{a\}$ de la fonction τ_a pour tout $a \in I$, ce qui donne la convexité de f en vertu du lemme de la pente croissante. ■

Interprétation graphique du lemme de la pente glissante

Lorsqu'elle glisse dans le sens des abscisses croissantes, la pente d'une fonction convexe croit.



A.4. Continuité d'une fonction convexe

Le lemme de la pente glissante permet d'établir la continuité en tout point de l'intérieur de l'intervalle.

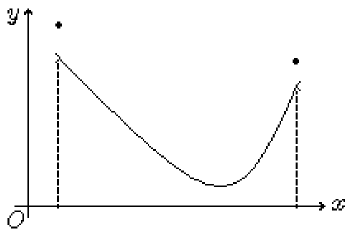
Théorème 1

Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe (ou concave) sur I. Pour tous $a, b \in \text{Int}(I)$, f est lipschitzienne sur $[a; b]$. En conséquence, f est continue sur $\text{Int}(I)$.

■ On démontre la propriété dans le cas d'une fonction convexe.

- ▷ Soient a, b deux points de $\text{Int}(I)$. Il existe alors $\varepsilon > 0$ tel que $[a-\varepsilon; b+\varepsilon] \subset I$. Soient $x, y \in [a; b]$ tels que $x \neq y$. Le lemme de la pente glissante nous dit que $\tau_{a-\varepsilon}(a) \leqslant \tau_x(y) \leqslant \tau_b(b+\varepsilon)$. La valeur absolue du taux d'accroissement de f est donc majorée sur $[a; b]$ par $k = \max\{|\tau_{a-\varepsilon}(a)| \,;\, |\tau_b(b+\varepsilon)|\}$ et f est ainsi k-lipschitzienne sur $[a; b]$.
- ▷ Si $x_0 \in I$ est différent des extrémités de I, il existe $a, b \in I$ différents des extrémités de I tels que $a < x_0 < b$. La fonction f est donc lipschitzienne sur $[a; b]$ et par conséquent continue en x_0. ■

Les seules fonctions convexes qui ne sont pas continues le sont donc seulement au bord de leur intervalle de définition :



B. Convexité et dérivabilité

B.1. Dérivabilité lattérale d'une fonction convexe

Le lemme de la pente croissante permet également d'établir la dérivabilité à gauche et à droite en tout point de l'intérieur de l'intervalle.

Théorème 2

Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe (ou concave) sur I. Alors f admet une dérivée à gauche et à droite en chaque point a de $\mathrm{Int}(I)$.

■ On démontre la propriété dans le cas où f est convexe. Soit $a \in \mathrm{Int}(I)$ de sorte qu'il existe $r > 0$ tel que $[a - r; a + r] \subset I$. Dès lors, sur $[a - r; a + r] \setminus \{a\}$, la fonction τ_a est alors croissante, minorée par $\tau_a(a - \varepsilon)$ et majorée par $\tau_a(a + \varepsilon)$. Le théorème de la limite monotone nous dit en conséquence que τ_a admet une limite à gauche et une limite à droite en a. La fonction f admet donc une dérivée à gauche et une dérivée à droite en a. ■

De plus, on peut noter que, si $f : I \longrightarrow \mathbb{R}$ est convexe et si a désigne un point de $\mathrm{Int}(I)$, alors $f'(a^-) \leqslant f'(a^+)$. Si f est concave, l'inégalité est évidemment dans l'autre sens.

Au bord de l'intervalle, le raisonnement tenu dans cette démonstration permet d'affirmer que le taux d'accroissement admet une limite finie ou infinie. Par conséquent, en un tel point, la fonction admet une tangente, mais celle-ci peut être verticale !

L'existence de dérivées lattérales en tout point de $\mathrm{Int}(I)$ assure la continuité à gauche et à droite de la fonction convexe sur $\mathrm{Int}(I)$. On retrouve ainsi la continuité sur $\mathrm{Int}(I)$.

B.2. Caractérisations différentielles de la convexité

Voici un critère différentiel très pratique pour savoir si une fonction est convexe.

Proposition 2

Soit $f : I \longrightarrow \mathbb{R}$ une fonction dérivable sur I. La fonction f est convexe (respectivement concave) si, et seulement si, la fonction f' est croissante (respectivement décroissante) sur I.

■ On démontre la propriété dans le cas d'une fonction convexe.

- $\Rightarrow$ Supposons f convexe sur I. Soient $x, y \in I$ tel que $x < y$. Pour tout $z \in]x; y[$, on a $\tau_x(z) \leqslant \tau_x(y)$ et $\tau_y(x) \leqslant \tau_y(z)$. En faisant tendre z vers x^+ dans la première inégalité, on a $f'(x) \leqslant \tau_x(y)$. En faisant tendre z vers y^- dans la seconde inégalité, on a $\tau_x(y) \leqslant f'(y)$. Comme $\tau_x(y) = \tau_y(x)$, on obtient $f'(x) \leqslant f'(y)$. Cela démontre que f' est croissante sur I.
- $\Leftarrow$ Supposons réciproquement que f' est croissante sur I. Soient $x, y \in I$ tel que $x < y$ et $t \in [0; 1]$. On veut démontrer que $f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y)$. Comme les cas $t = 0$ et $t = 1$ sont triviaux, on peut supposer que $t \in]0; 1[$. Dès lors, $tx + (1-t)y \in]x; y[$ et l'on peut écrire, en vertu du théorème des accroissements finis, qu'il existe $c_1 \in]x; tx + (1-t)y[$ et $c_2 \in]tx + (1-t)y; y[$ tels que $\tau_x\big(tx + (1-t)y\big) = f'(c_1)$ et $\tau_{tx+(1-t)y}(y) = f'(c_2)$. Comme $c_1 < c_2$ et comme f' est croissante, on obtient $\tau_x\big(tx + (1-t)y\big) \leqslant \tau_{tx+(1-t)y}(y)$. Après quelques calculs (laissés aux soins du lecteur), on en déduit que $f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y)$. La fonction f est donc convexe sur I. ■

On peut démontrer que la stricte convexité est équivalente à la stricte croissance de la dérivée.

La proposition précédente fournit, en corollaire, une caractérisation très pratique de la convexité par la dérivée seconde.

Théorème 3

Une fonction $f : I \longrightarrow \mathbb{R}$ deux fois dérivable sur I est convexe (respectivement concave) si, et seulement si, $\forall x \in I,\ f''(x) \geqslant 0$ (respectivement $\forall x \in I,\ f''(x) \leqslant 0$).

■ C'est une conséquence immédiate de la proposition précédente. ■

Attention, s'il est vrai qu'une fonction dont la dérivée seconde est strictement positive est strictement convexe, la réciproque est fausse : il existe des fonctions strictement convexes dont la dérivée seconde peut s'annuler (par exemple $x \longmapsto x^4$).

Exemples :

- $x \longmapsto x^2$ est convexe sur $\mathbb{R}$ et $x \longmapsto \sqrt{x}$ est concave sur $\mathbb{R}_+$.
- exp est convexe sur $\mathbb{R}$ et ln est concave sur $\mathbb{R}_+^*$.
- sin est concave sur $[0; \pi/2]$. Cela implique que $\forall x \in [0; \pi/2]\,,\ \sin x \geqslant 2x/\pi$.

On peut caractériser les points d'inflexion d'une fonction deux fois dérivables.

Corollaire 2

Soit $f : I \longrightarrow \mathbb{R}$ une fonction deux fois dérivable sur I. Un point $x_0 \in I$ en lequel f'' s'annule et change de signe est un point d'inflexion.

■ AQT ■

Exemples :

- arctan admet un point d'inflexion en 0 car $x \longmapsto \arctan''(x) = -2x/(1+x^2)^2$ est positive sur $\mathbb{R}_-$ et négative sur $\mathbb{R}_+$.

B.3. Position relative d'une fonction convexe et de ses tangentes

Proposition 3

Soit $f : I \longrightarrow \mathbb{R}$ une fonction dérivable et convexe sur I. La courbe représentative de f est alors au dessus de chacune de ses tangentes, c'est-à-dire que, pour tout $a \in I$, on a

$$\forall x \in I, \qquad f(x) \geqslant f(a) + (x-a)f'(a).$$

Si f est concave sur I, l'inégalité ci-dessus est vérifiée avec un signe $\geqslant$ à la place de $\leqslant$. Autrement dit, la courbe représentative d'une fonction concave est en dessous de chacune de ses tangentes.

■ Soient $a, x \in I$.
Si $a = x$, le résultat à démontrer est évident (puisqu'il est équivalent à $f(a) \geqslant f(a)$).
Supposons dorénavant que $a \neq x$.
D'après le théorème des accroissements finis appliqué à f entre a et x, il existe $c \in]a\,;x[$ tel que $f(x) = f(a) + (x-a)f'(c)$.
Si $a \leqslant c \leqslant x$ alors, comme f' est croissante sur I, on a $f'(a) \leqslant f'(c)$, d'où $f'(a)(x-a) \leqslant f'(c)(x-a)$ car $x - a > 0$. Donc $f(x) \geqslant f(a) + (x-a)f'(a)$.
Si $x \leqslant c \leqslant a$ alors, comme f' est croissante sur I, on a $f'(c) \leqslant f'(a)$, d'où $f'(a)(x-a) \leqslant f'(c)(x-a)$ car $x - a < 0$. Donc $f(x) \geqslant f(a) + (x-a)f'(a)$. ■

Il découle de ce résultat qu'en un point d'inflexion, la courbe traverse sa tangente

Exemples :

- $\forall x \in \mathbb{R}, \quad e^x \geqslant x + 1$
- $\forall x > 0, \quad \ln x \leqslant x - 1,$
- $\forall x \geqslant 0, \quad \sin x \leqslant x.$

2 h 00