

Chapitre 18

Convexité

Dans tout ce chapitre, I désigne un intervalle non trivial de $\mathbb{R}$ et f une fonction de I dans $\mathbb{R}$.

I Généralités

1. Définition

Paramétrisation d'un segment Si $x, y \in \mathbb{R}$ vérifient $x < y$, il sera utile de remarquer que le segment $[x, y]$ peut être décrit de manière paramétrique :

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

En effet, on voit aisément que z est un élément du segment $[x, y]$ si et seulement si on peut écrire $z = x + \lambda(y - x)$ avec $\lambda \in [0, 1]$. Il suffit alors de remarquer que $x + \lambda(y - x) = (1 - \lambda)x + \lambda y$.

Remarques.

- Le cas $\lambda = 0$ correspond au point x , le cas $\lambda = 1$ au point y , et le cas $\lambda = \frac{1}{2}$ au milieu du segment $\frac{x+y}{2}$.



- Si $a = (1 - \lambda)x + \lambda y$ avec $\lambda \in [0, 1]$, alors on a :

$$\lambda = \frac{a - x}{y - x}, \quad \text{et} \quad 1 - \lambda = \frac{y - a}{y - x}.$$



Cordes Si $x, y \in I$, on dit que le segment reliant M_x et M_y de coordonnées respectives $\begin{pmatrix} x \\ f(x) \end{pmatrix}$ et $\begin{pmatrix} y \\ f(y) \end{pmatrix}$ est une *corde* de f . De la même manière que ci-dessus, on peut paramétrer le segment $[M_x, M_y]$: il s'agit de l'ensemble des points de coordonnées :

$$(1 - \lambda) \begin{pmatrix} x \\ f(x) \end{pmatrix} + \lambda \begin{pmatrix} y \\ f(y) \end{pmatrix} = \begin{pmatrix} (1 - \lambda)x + \lambda y \\ (1 - \lambda)f(x) + \lambda f(y) \end{pmatrix}.$$

On appelle par ailleurs *sécante* de f la droite $(M_x M_y)$.

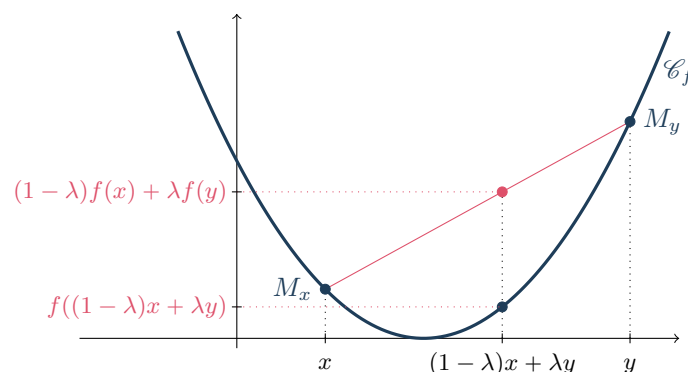
Définition - Fonction convexe, concave

On dit que f est convexe sur I si sa courbe est en-dessous de ses cordes sur I , c'est-à-dire :

$$\forall x, y \in I, \forall \lambda \in [0, 1], \quad f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y).$$

On dit que f est concave sur I si sa courbe est au-dessus de ses cordes sur I , c'est-à-dire :

$$\forall x, y \in I, \forall \lambda \in [0, 1], \quad f((1 - \lambda)x + \lambda y) \geq (1 - \lambda)f(x) + \lambda f(y).$$



Une fonction convexe et l'une de ses cordes

Remarque. La fonction f est concave sur I si et seulement si $-f$ est convexe. Pour cette raison, nous nous contenterons d'énoncer les résultats dans le cadre de fonctions convexes, le cas des fonctions concaves s'en déduisant directement.

Exemple. La fonction valeur absolue $x \mapsto |x|$ est convexe sur $\mathbb{R}$. En effet, si $x, y \in \mathbb{R}$ et $\lambda \in [0, 1]$, alors

$$|(1-\lambda)x + \lambda y| \leq |(1-\lambda)x| + |\lambda y| = (1-\lambda)|x| + \lambda|y|.$$

2. Croissance des pentes, lemme des pentes

Théorème - Croissance des pentes, inégalité des pentes

- (i) La fonction f est convexe sur I si et seulement si pour tout $a \in I$, la fonction $\tau_a : x \mapsto \frac{f(x)-f(a)}{x-a}$ est croissante sur $I \setminus \{a\}$.
- (ii) Si f est convexe sur I et $x, y, z \in I$ sont tels que $x < y < z$, alors

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

Remarque. Il est utile, notamment pour la preuve ci-dessous, de remarquer que si $x, y \in I$ et $x \neq y$, alors $\tau_x(y) = \tau_y(x)$.

Démonstration. Montrons (i) par double implication.

- Supposons que f est convexe sur I . Soient $a \in I$ et $x, y \in I$ tels que $x < y$, nous allons montrer que $\tau_a(x) \leq \tau_a(y)$. On suppose d'abord que $a < x < y$. On a alors $x = (1-\lambda)a + \lambda y$, avec $\lambda = \frac{x-a}{y-a}$. Alors par convexité,

$$f(x) \leq (1-\lambda)f(a) + \lambda f(y), \text{ c'est-à-dire } f(x) - f(a) \leq \lambda(f(y) - f(a)).$$

En divisant par $x - a > 0$, on obtient $\tau_a(x) \leq \tau_a(y)$. Le cas où $x < y < a$ se traite de la même manière. Il reste à considérer le cas où $x < a < y$. Il découle de ce qui précède que $\tau_x(a) \leq \tau_x(y)$, et $\tau_y(x) \leq \tau_y(a)$, alors

$$\tau_x(a) \leq \tau_x(y) = \tau_y(x) \leq \tau_y(a), \text{ donc } \tau_x(a) \leq \tau_y(a) \text{ puis } \tau_a(x) \leq \tau_a(y).$$

- Réciproquement, supposons que pour tout $a \in I$, la fonction τ_a est croissante sur $I \setminus \{a\}$. Soient $x, y \in I$ et $\lambda \in [0, 1]$. On peut supposer $x < y$, et on note $a = (1-\lambda)x + \lambda y$. La croissance de τ_a donne alors

$$\frac{f(x)-f(a)}{x-a} \leq \frac{f(y)-f(a)}{y-a}, \text{ donc } (f(x)-f(a))(y-a) \geq (f(y)-f(a))(x-a)$$

en multipliant l'inégalité par $(x-a)(y-a) < 0$. L'inégalité se réécrit $f(a)(y-x) \leq f(x)(y-a) - f(y)(x-a)$. Or $x-a = -\lambda(y-x)$ et $y-a = (1-\lambda)(y-x)$, ce qui entraîne $f(a) \leq (1-\lambda)f(x) + \lambda f(y)$ en divisant par $y-x > 0$, qui est l'inégalité souhaitée.

La propriété (ii) est une simple conséquence de ce qui précède : on a $\tau_x(y) \leq \tau_x(z)$, qui donne la première inégalité, puis $\tau_z(x) \leq \tau_z(y)$, donc $\tau_x(z) \leq \tau_y(z)$, ce qui donne la seconde. $\square$

Exercice 1. Montrer que si f est convexe sur un intervalle I et admet un minimum local en $a \in I$, alors le minimum est global.

Théorème - Régularité des fonctions convexes

Si f est convexe sur l'intervalle I , alors f est continue sur $\tilde{I}$.

Démonstration. Soit $a \in \tilde{I}$. Il existe $\alpha, \alpha', \beta, \beta' \in I$ tels que $\alpha' < \alpha < a < \beta < \beta'$, du fait que a n'est pas une extrémité de I . Nous allons montrer que f est lipschitzienne sur $[\alpha, \beta]$, ce qui conclura. L'inégalité des pentes donne que pour tout $x \in [\alpha, \beta]$ tel que $x \neq a$,

$$m \leq \frac{f(x)-f(a)}{x-a} \leq M, \text{ où } m = \frac{f(\alpha)-f(\alpha')}{\alpha-\alpha'} \text{ et } M = \frac{f(\beta')-f(\beta)}{\beta'-\beta}.$$

Ainsi, on a $m(x-a) \leq f(x) - f(a) \leq M(x-a)$, et en posant $k = \max(|m|, |M|)$, on a $|f(x) - f(a)| \leq k|x-a|$. On en déduit que $|f(x) - f(a)| \xrightarrow{x \rightarrow a} 0$, et donc $f(x) \xrightarrow{x \rightarrow a} f(a)$, ce qui donne la continuité en a . $\square$

Remarque. Nous avons en fait prouvé un résultat plus fort : une fonction convexe sur un intervalle I est lipschitzienne sur tout segment de I . On dit qu'elle est localement lipschitzienne.

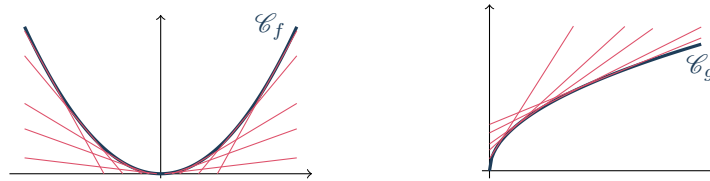
⚠ Une fonction convexe sur un intervalle I n'est pas nécessairement continue sur cet intervalle : elle peut éventuellement présenter des discontinuités aux extrémités de I .

II Convexité et dérivabilité

Théorème - Caractérisation des fonctions convexes dérivables

Si f est dérivable sur l'intervalle I , les propriétés suivantes sont équivalentes.

- (i) f est convexe sur I ,
- (ii) f' est croissante sur I ,
- (iii) la courbe de f est au-dessus de ses tangentes sur I , i.e. $\forall x, y \in I, f(y) \geq f(x) + f'(x)(y - x)$.



Le graphe d'une fonction convexe f , et le graphe d'une fonction concave g

Démonstration.

- Montrons (i) $\Rightarrow$ (ii) : si f est convexe sur I et $x, y \in I$ avec $x < y$, alors pour tout $t \in]x, y[$, on a

$$\frac{f(t) - f(x)}{t - x} \leq \frac{f(y) - f(x)}{y - x}, \quad \text{donc} \quad f'(x) = \lim_{t \rightarrow x} \frac{f(t) - f(x)}{t - x} \leq \frac{f(y) - f(x)}{y - x}.$$

De même, comme $\frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(t)}{y - t}$, on a $\frac{f(y) - f(x)}{y - x} \leq \lim_{t \rightarrow y} \frac{f(t) - f(y)}{t - y} = f'(y)$. Ceci entraîne $f'(x) \leq f'(y)$.

- Montrons (ii) $\Rightarrow$ (iii) : on suppose f' est croissante sur I , et on fixe $x \in I$. Nous allons montrer que la fonction

$$g : y \mapsto f(y) - f(x) - f'(x)(y - x)$$

est positive sur I . La fonction g est dérivable sur I , et pour tout $y \in I$, on a $g'(y) = f'(y) - f'(x)$. Par croissance de f' , on en déduit que g' est positive sur $I \cap]-\infty, x]$ et négative sur $I \cap [x, +\infty[$. Ainsi, g est minimale en x , et pour tout $y \in I$, on a $g(y) \geq g(x) = 0$. Ceci conclut.

- Montrons (iii) $\Rightarrow$ (i) : on suppose que la courbe de f est au-dessus de ses tangentes, et on considère $x, y \in I$ et $\lambda \in [0, 1]$. On pose $a = (1 - \lambda)x + \lambda y$ et on cherche à montrer $f(a) \leq (1 - \lambda)f(x) + \lambda f(y)$. On remarque que $x - a = -\lambda(y - x)$ et $y - a = (1 - \lambda)(y - x)$. En appliquant deux fois l'inégalité, on obtient

$$\begin{cases} f(x) \geq f(a) + f'(a)(x - a) \\ f(y) \geq f(a) + f'(a)(y - a) \end{cases} \quad \text{donc} \quad \begin{cases} f(x) - f(a) \geq -\lambda f'(a)(y - x) & (1) \\ f(y) - f(a) \geq (1 - \lambda) f'(a)(y - x) & (2) \end{cases}$$

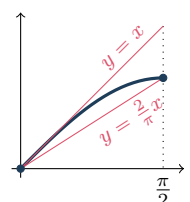
L'opération $(1 - \lambda) \times (1) + \lambda \times (2)$ donne $(1 - \lambda)(f(x) - f(a)) + \lambda(f(y) - f(a)) \geq 0$, i.e. $(1 - \lambda)f(x) + \lambda f(y) \geq f(a)$. $\square$

Exemples. – La fonction $\exp$ est de dérivée croissante, donc elle est convexe sur $\mathbb{R}$.
 – La fonction $\ln$ est de dérivée décroissante sur $\mathbb{R}_+^*$, donc elle est concave sur $\mathbb{R}_+^*$.

Exercice 2. Pour tout $x \in [0, \frac{\pi}{2}]$, montrer que $\frac{2}{\pi}x \leq \sin x \leq x$.

La fonction $\sin$ a pour dérivée $\cos$ qui décroît sur $[0, \frac{\pi}{2}]$, donc elle est concave sur $[0, \frac{\pi}{2}]$.

- Ainsi, la courbe de $\sin$ est en-dessous de sa tangente au point d'abscisse 0, qui a pour équation $y = x$. Ceci entraîne que pour tout $x \in [0, \frac{\pi}{2}]$, on a $\sin x \leq x$.
- Par ailleurs, la courbe de $\sin$ est au-dessus de sa corde entre les points d'abscisse 0 et $\frac{\pi}{2}$, d'équation $y = \frac{2}{\pi}x$. Ceci entraîne que $\frac{2}{\pi}x \leq \sin x$ pour tout $x \in [0, \frac{\pi}{2}]$.



Corollaire - Cas deux fois dérivable

Si f est deux fois dérivable sur l'intervalle I , alors f est convexe sur I si et seulement si f'' est positive sur I .

Démonstration. Comme f' est dérivable sur I , nous avons vu que f' est croissante si et seulement si $f'' \geq 0$ sur I , donc le théorème ci-dessus permet de conclure. $\square$

III Inégalité de Jensen**Théorème - Inégalité de Jensen**

Soit f une fonction convexe sur l'intervalle I . Si $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$ avec $\sum_{k=1}^n \lambda_k = 1$ et $x_1, \dots, x_n \in I$, alors

$$f\left(\sum_{k=1}^n \lambda_k x_k\right) \leq \sum_{k=1}^n \lambda_k f(x_k).$$

Démonstration. Montrons par récurrence que $\mathcal{P}(n)$ est vraie pour tout $n \geq 2$, où

$\mathcal{P}(n)$: pour tous $x_1, \dots, x_n \in I$ et $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$ tels que $\sum_{i=1}^n \lambda_i = 1$, on a $f\left(\sum_{k=1}^n \lambda_k x_k\right) \leq \sum_{k=1}^n \lambda_k f(x_k)$.

– La propriété $\mathcal{P}(2)$ est vraie : c'est la définition de la convexité.

– Soit $n \geq 2$. Supposons que $\mathcal{P}(n)$ est vraie, et montrons que $\mathcal{P}(n+1)$ l'est également. On a

$$f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) = f\left(\sum_{k=1}^n \lambda_k x_k + \lambda_{n+1} x_{n+1}\right) \leq f\left((1 - \lambda_{n+1}) \left(\sum_{k=1}^n \frac{\lambda_k}{1 - \lambda_{n+1}} x_k\right) + \lambda_{n+1} x_{n+1}\right).$$

Donc, en utilisant $\mathcal{P}(2)$, on obtient $f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) \leq (1 - \lambda_{n+1}) f\left(\sum_{k=1}^n \frac{\lambda_k}{1 - \lambda_{n+1}} x_k\right) + \lambda_{n+1} f(x_{n+1})$.

Or $\sum_{k=1}^n \frac{\lambda_k}{1 - \lambda_{n+1}} = \frac{\sum_{k=1}^n \lambda_k}{1 - \lambda_{n+1}} = \frac{1 - \lambda_{n+1}}{1 - \lambda_{n+1}} = 1$, donc $\mathcal{P}(n)$ donne $f\left(\sum_{k=1}^n \frac{\lambda_k}{1 - \lambda_{n+1}} x_k\right) \leq \sum_{k=1}^n \frac{\lambda_k}{1 - \lambda_{n+1}} f(x_k)$. Ainsi,

$$f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) \leq (1 - \lambda_{n+1}) \sum_{k=1}^n \frac{\lambda_k}{1 - \lambda_{n+1}} f(x_k) + \lambda_{n+1} f(x_{n+1}) = \sum_{k=1}^{n+1} \lambda_k f(x_k),$$

ce qui montre $\mathcal{P}(n+1)$, et conclut. $\square$

Remarque. Sous les mêmes conditions, mais avec f concave, l'inégalité est : $f\left(\sum_{k=1}^n \lambda_k x_k\right) \geq \sum_{k=1}^n \lambda_k f(x_k)$.

Exemple. Une application classique est l'inégalité dite arithmético-géométrique : si $x_1, \dots, x_n > 0$, alors

$$\left(\prod_{k=1}^n x_k\right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{k=1}^n x_k.$$

Comme $\ln$ est concave, on a par inégalité de Jensen (avec $\lambda_1 = \dots = \lambda_n = \frac{1}{n}$, ce qui donne $\sum_{k=1}^n \lambda_k = 1$) :

$$\ln\left(\sum_{k=1}^n \frac{1}{n} x_k\right) \geq \sum_{k=1}^n \frac{1}{n} \ln(x_k).$$

Ainsi, on a : $\ln\left(\prod_{k=1}^n x_k\right)^{\frac{1}{n}} = \frac{1}{n} \sum_{k=1}^n \ln(x_k) = \sum_{k=1}^n \frac{1}{n} \ln(x_k) \leq \ln\left(\sum_{k=1}^n \frac{1}{n} x_k\right) = \ln\left(\frac{1}{n} \sum_{k=1}^n x_k\right)$.

On obtient l'inégalité en composant par la fonction $\exp$ qui est croissante.