FONCTIONS CONVEXES D'UNE VARIABLE RÉELLE

I. Définition

I.1. Définition

Soit f une fonction définie sur un intervalle I. On dit que f est convexe sur I lorsque :

$$\forall x, y \in I, \forall \lambda \in [0 ; 1], f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y)$$

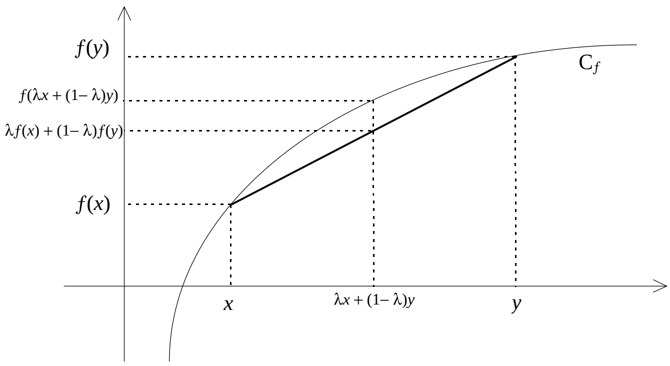
Lorsque l'on a l'inégalité dans l'autre sens, on dit que f est concave sur I.

Si l'inégalité est stricte, on parle alors de stricte convexité ou concavité.

Notons que f est concave sur I si et seulement si $-f$ est convexe sur I.

Remarquons que si $x = y$ ou si $\lambda = 0$ ou $\lambda = 1$ alors l'inégalité $f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y)$ est triviale.

Illustration dans le cas d'une fonction concave : la courbe est au dessus de toute corde.



Exemples :

- La fonction $f : x \mapsto |x|$ est convexe puisque d'après l'inégalité triangulaire : $|\lambda x + (1-\lambda)y| \leqslant \lambda|x| + (1-\lambda)|y|$.
- La fonction $f : x \mapsto x^2$ est convexe. En effet :

$$\lambda f(x) + (1-\lambda)f(y) - f(\lambda x + (1-\lambda)y) = \lambda x^2 + (1-\lambda) y^2 - \lambda^2 x^2 - 2\lambda(1-\lambda)xy - (1-\lambda)^2 y^2$$

$$= \lambda(1-\lambda)[x^2 - 2xy + y^2] = \lambda(1-\lambda)(x-y)^2 \geqslant 0$$

- Les fonctions affines sont à la fois convexes et concaves.

I.2. Théorème

Soient f une fonction convexe définie sur un intervalle I, $x_1, \dots, x_n$ des points de I et $\lambda_1, \dots, \lambda_n$ des réels positifs vérifiant $\sum_{i=1}^{n} \lambda_i = 1$. Alors :

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

En particulier :
$$f\left(\frac{x_1 + \dots + x_n}{n}\right) \leqslant \frac{f(x_1) + \dots + f(x_n)}{n}$$

Preuve : par récurrence sur n. La formule est vraie pour $n = 2$ (c'est la définition avec $\lambda_1 = \lambda$ et $\lambda_2 = 1 - \lambda_1$)

Considérons la propriété suivante :

$$H_n = \text{"}\forall x_1, \dots, x_n \in I, \forall \lambda_1, \dots, \lambda_n \in [0 ; 1] \text{ tels que } \sum_{i=1}^{n} \lambda_i = 1 \text{ on a : } f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i) \text{"}$$

Comme f est convexe sur I, on a : $\forall x, y \in I$, $\forall \lambda \in [0 ; 1]$, $f(\lambda x + (1 - \lambda)y) \leqslant \lambda f(x) + (1 - \lambda)f(y)$

En posant $x_1 = x$, $x_2 = y$, $\lambda_1 = \lambda$ et $\lambda_2 = 1 - \lambda$ on obtient H_2.

Supposons H_n vraie pour un certain entier $n \geqslant 2$ et montrons que cela entraîne H_{n+1}.

Soient $\lambda_1, \dots, \lambda_{n+1} \in [0 ; 1]$ tels que $\sum_{i=1}^{n+1} \lambda_i = 1$. Posons $S = \sum_{i=1}^{n} \lambda_i$.

Si $S = 0$ alors $\lambda_i = 0$ pour tout $i \in [\![1 ; n]\!]$ et donc $\lambda_{n+1} = 1$. L'inégalité $f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n+1} \lambda_i f(x_i)$ est alors vérifiée puisque trivialement réduite à $f(x_{n+1}) \leqslant f(x_{n+1})$.

Si $S \neq 0$ alors posons pour tout $i \in [\![1 ; n]\!]$, $\mu_i = \dfrac{\lambda_i}{S}$. Ainsi, on a $\sum_{i=1}^{n} \mu_i = \sum_{i=1}^{n} \dfrac{\lambda_i}{S} = \dfrac{1}{S} \times S = 1$ ce qui va permettre d'utiliser l'hypothèse de récurrence.

$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) = f\left(\sum_{i=1}^{n} \lambda_i x_i + \lambda_{n+1} x_{n+1}\right) = f\left(S\sum_{i=1}^{n} \mu_i x_i + (1-S)x_{n+1}\right) \text{ car } \lambda_{n+1} = 1 - \sum_{i=1}^{n} \lambda_i = 1 - S.$$

Et comme f est convexe (ou d'après H_2), on a :

$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \leqslant S\, f\left(\sum_{i=1}^{n} \mu_i x_i\right) + (1-S) f(x_{n+1})$$

Et comme $\sum_{i=1}^{n} \mu_i = 1$ et H_n est vraie par hypothèse de récurrence, on a : $f\left(\sum_{i=1}^{n} \mu_i x_i\right) \leqslant \sum_{i=1}^{n} \mu_i f(x_i)$

Or,
$$S\sum_{i=1}^{n} \mu_i f(x_i) = \sum_{i=1}^{n} \lambda_i f(x_i)$$

D'où finalement :
$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i) + \lambda_{n+1} f(x_{n+1}) \text{ d'où } H_{n+1}.$$

Le principe de récurrence achève la démonstration.

Le cas particulier en résulte en choisissant $\lambda_1 = \lambda_2 = \dots = \lambda_n = \dfrac{1}{n}$.

<u>Exemple</u> : $\forall x_1, \dots, x_n \in \mathbb{R}$: $\left(\sum_{i=1}^{n} x_i\right)^2 \leqslant n \sum_{i=1}^{n} x_i^2$ (convexité de $x \longmapsto x^2$)

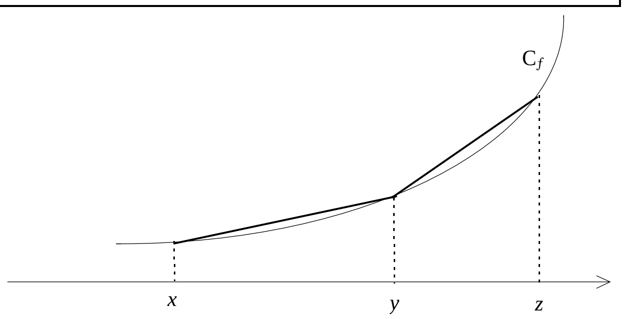
II. Différentes caractérisations

<u>II. 1. Caractérisation 1</u>

Soit f une fonction définie sur un intervalle I.

$$f \text{ est convexe sur } I \iff \forall x, y, z \in I,\ x < y < z \Rightarrow \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(z) - f(y)}{z - y}$$

<u>Preuve</u> :



$\Rightarrow$ Soient $x, y, z \in I$ tels que $x < y < z$.

Puisque $y \in]x ; z[$, on peut écrire :

$y = \lambda x + \mu z$ où $\lambda = \dfrac{z - y}{z - x}$ et $\mu = 1 - \lambda = \dfrac{y - x}{z - x} \in]0 ; 1[$

Comme f est convexe sur I, on a :

$$f(y) \leqslant \lambda f(x) + \mu f(z)$$

Ajoutons $\lambda f(y)$ à chaque membre, ainsi : $\lambda f(y) - \lambda f(x) \leqslant \mu f(z) - \mu f(y)$

Soit en divisant par $\lambda\mu > 0$: $\dfrac{f(y)-f(x)}{\mu} \leqslant \dfrac{f(z)-f(y)}{\lambda}$

Or, $\lambda = \dfrac{z-y}{z-x}$, $\mu = \dfrac{y-x}{z-x}$ et $z - x > 0$ d'où : $\dfrac{f(y)-f(x)}{y-x} \leqslant \dfrac{f(z)-f(y)}{z-y}$

$\Leftarrow$ Soient $x, z \in I$ avec (pour fixer les idées $x < z$, le cas $x = z$ étant trivial) et soit $\lambda \in \,]0 ; 1[$.

Posons $y = \lambda x + \mu z$ où $\mu = 1 - \lambda$. Ainsi $x < y < z$.

On a donc $\dfrac{f(y)-f(x)}{y-x} \leqslant \dfrac{f(z)-f(y)}{z-y}$ d'où $\lambda[f(y) - f(x)] \leqslant \mu[f(z) - f(y)]$ d'où$f(y) \leqslant \lambda f(x) + \mu f(z)$.

L'inégalité étant trivialement valable lorsque $\lambda = 0$ ou $\lambda = 1$, on en déduit que f est convexe sur I.

II.2. Caractérisation 2

Soit f une fonction définie sur un intervalle I. Pour tout $\alpha \in I$ on pose $\varphi_\alpha(x) = \dfrac{f(x)-f(\alpha)}{x-\alpha}$ $\forall x \in I \setminus \{\alpha\}$.

$$f \text{ est convexe sur } I \iff \forall \alpha \in I,\ \varphi_\alpha \text{ est croissante sur } I \setminus \{\alpha\}$$

Preuve :

$\Rightarrow$ Soient $x, y \in I \setminus \{\alpha\}$ tels que $x < y$. Montrons que $\varphi_\alpha(x) \leqslant \varphi_\alpha(y)$

Supposons que l'on ait $x < y < \alpha$. Comme $y \in \,]x ; \alpha[$, on peut écrire :

$$y = \lambda x + \mu\alpha \text{ où } \lambda = \frac{\alpha - y}{\alpha - x} \in \,]0 ; 1[\text{ et } \mu = 1 - \lambda = \frac{y-x}{\alpha - x} \in \,]0 ; 1[$$

Comme f est convexe sur I, on en déduit : $f(y) \leqslant \lambda f(x) + \mu f(\alpha)$

Retranchons $f(\alpha)$ à chaque membre : $f(y) - f(\alpha) \leqslant \lambda[f(x) - f(\alpha)]$

D'où ($x - \alpha < 0$ et $y - \alpha < 0$) : $\dfrac{f(y)-f(\alpha)}{y-\alpha} \geqslant \dfrac{f(x)-f(\alpha)}{x-\alpha}$

$$\varphi_\alpha(y) \geqslant \varphi_\alpha(x)$$

Le cas $\alpha < x < y$ est analogue et le cas $x < \alpha < y$ immédiat d'après la caractérisation 1.

Bilan : la fonction φ_α est bien croissante sur $I \setminus \{\alpha\}$.

$\Leftarrow$ Soient $x, y, z \in I$ tels que $x < y < z$.

Comme φ_y est croissante, on a $\varphi_y(x) \leqslant \varphi_y(z)$ donc (caractérisation 1) f est convexe.

Et évidemment, on a : f est concave sur $I \iff \forall \alpha \in I$, φ_α est décroissante sur $I \setminus \{\alpha\}$

II.3. Conséquences

Soit f une fonction convexe sur I . Alors :

1) f admet en tout point α intérieur à I une dérivée à droite et une dérivée à gauche

2) f est continue en tout point α intérieur à I.

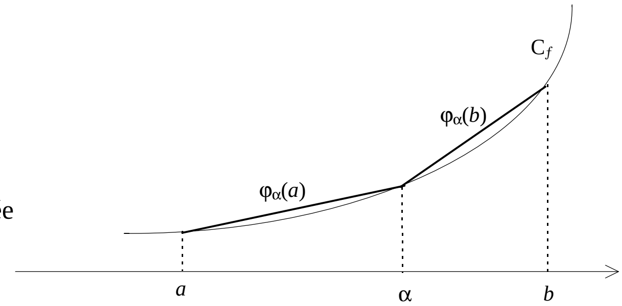
Preuve :

1) Soit $\alpha \in \overset{\circ}{I}$. La fonction φ_α est croissante sur $I \setminus \{\alpha\}$ et bornée au voisinage de α. Elle admet donc une limite à droite et à gauche de α (voir leçon sur les fonctions monotones). Donc f admet en α une dérivée à droite et une dérivée à gauche et comme $\varphi_\alpha(\alpha^-) \leqslant \varphi_\alpha(\alpha^+)$, on a $f'_g(\alpha) \leqslant f'_d(\alpha)$

2) Soit $\alpha \in \overset{\circ}{I}$. Donc $\exists a, b \in I$ tels que $a < \alpha < b$.

La fonction φ_α est croissante sur $[a ; \alpha[$ minorée par $\varphi_\alpha(a)$ et majorée par $\varphi_\alpha(b)$:

$$\forall x \in [a ; \alpha[: \varphi_\alpha(a) \leqslant \varphi_\alpha(x) \leqslant \varphi_\alpha(b)$$



De même sur $]\alpha ; b]$, φ_α est croissante et bornée (toujours par $\varphi_\alpha(a)$ et $\varphi_\alpha(b)$) :

$$\forall x \in]\alpha ; b] : \varphi_\alpha(a) \leqslant \varphi_\alpha(x) \leqslant \varphi_\alpha(b)$$

Bilan : $\exists M \in \mathbb{R}$ tel que $\forall x \in I \setminus \{\alpha\}$ on a :

$$|f(x) - f(\alpha)| \leqslant M|x - \alpha| \quad (M = \sup\{|\varphi_\alpha(a)| ; |\varphi_\alpha(b)|\})$$

Et comme l'inégalité ci-dessus est encore valable pour $x = \alpha$, on en déduit que f est continue en α.

(En effet, $\forall \varepsilon > 0$, pour $\eta < \dfrac{\varepsilon}{M}$, $|x - \alpha| < \eta$ entraîne $|f(x) - f(\alpha)| < M|x - \alpha| < M\eta < \varepsilon$)

Remarque : il existe des fonctions convexes non continues. (Mais les points de discontinuité sont situés sur Fr(I))

Par exemple la fonction f définie sur $[0 ; 1]$ par $f(x) = \begin{cases} 1 \text{ si } x = 0 \\ 0 \text{ si } x \in \ ]0\ ;\ 1] \end{cases}$

En effet, on a toujours $f(\lambda x + (1 - \lambda)y) = 0$ puisque $\lambda x + (1 - \lambda)y \neq 0$ et $\lambda f(x) + (1 - \lambda)f(y) \geqslant 0$

La conséquence II.3. ci-dessus s'applique aussi aux fonctions concaves.

II.4. Caractérisation 3

Soit f une fonction définie et dérivable sur un intervalle I.

$$f \text{ est convexe sur } I \iff f' \text{ est croissante sur } I$$

Preuve :

$\Rightarrow$ Soient $a, b \in I$ avec $a \leqslant b$. Nous savons que f est dérivable sur I donc :

$f'(a) = \lim\limits_{x \to a} \varphi_a(x)$ avec $f'(a) \leqslant \varphi_a(b)$ puisque φ_a est croissante

$f'(b) = \lim\limits_{x \to b} \varphi_b(x)$ avec $f'(b) \geqslant \varphi_b(a)$ puisque φ_b est croissante

Et comme $\varphi_a(b) = \varphi_b(a)$, on a $f'(a) \leqslant f'(b)$ donc f' est croissante sur I.

$\Leftarrow$ Soient $a, b, c \in I$ avec $a < b < c$.

D'après le théorème des accroissements finis appliqué à f sur $]a ; b[$ puis sur $]b ; c[$:

$\exists \alpha \in \ ]a ; b[$ tel que $f(b) - f(a) = f'(\alpha)(b - a)$ c'est-à-dire $f'(\alpha) = \varphi_b(a)$

$\exists \beta \in \ ]b ; c[$ tel que $f(c) - f(b) = f'(\beta)(c - b)$ c'est-à-dire $f'(\beta) = \varphi_b(c)$

Or, f' est croissante et $\alpha < \beta$ donc $f'(\alpha) \leqslant f'(\beta)$ c'est-à-dire $\varphi_b(a) \leqslant \varphi_b(c)$.

Ainsi, d'après la caractérisation 1, f est convexe.

Et pour les fonctions concaves, on a : f est concave sur $I \Leftrightarrow f'$ est décroissante sur I

On a vu, que la convexité se traduit par le fait que les cordes sont au dessus du graphe. La caractérisation suivante valable pour une fonction dérivable traduit la convexité par le fait que les tangentes sont en dessous du graphe.

II.5. Caractérisation 4

Soit f une fonction définie et dérivable sur un intervalle I.

$$f \text{ est convexe sur } I \Leftrightarrow \forall \alpha \in I, \forall x \in I, f(x) \geqslant f(\alpha) + (x - \alpha)f'(\alpha)$$

Preuve :

$\Rightarrow$ Soit $\alpha \in I$ et soit $x \in I$. Si $x = \alpha$, l'inégalité est triviale.

Si $x > \alpha$, alors $f'(\alpha) = \varphi_\alpha(\alpha^+) \leqslant \varphi_\alpha(x)$ car φ_α est croissante. D'où $f'(\alpha) \leqslant \dfrac{f(x) - f(\alpha)}{x - \alpha}$ c'est-à-dire :

$$f(x) \geqslant f(\alpha) + (x - \alpha)f'(\alpha)$$

Si $x < \alpha$, alors $\varphi_\alpha(x) \leqslant \varphi_\alpha(\alpha^-) = f'(\alpha)$n d'où $\dfrac{f(x) - f(\alpha)}{x - \alpha} \leqslant f'(\alpha)$ et comme $x - \alpha < 0$:

$$f(x) \geqslant f(\alpha) + (x - \alpha)f'(\alpha)$$

$\Leftarrow$ Soient $x, y \in I$ avec $x \leqslant y$. On a donc, par hypothèse :

$$f(y) \geqslant f(x) + (y - x)f'(x) \text{ et } f(x) \geqslant f(y) + (x - y)f'(y)$$

$$f'(x) \leqslant \frac{f(y) - f(x)}{y - x} \leqslant f'(y)$$

Donc f' est croissante et donc f est convexe.

Notons que l'inégalité est tournée dans l'autre sens en cas de concavité.

II.6. Caractérisation 5

Soit f une fonction définie et deux fois dérivable sur un intervalle I.

$$f \text{ est convexe sur } I \Leftrightarrow f''(x) \geqslant 0 \text{ sur } I$$

Preuve : f convexe sur $I \Leftrightarrow f'$ croissante sur $I \Leftrightarrow f'' \geqslant 0$ sur I

Et évidemment on a : f concave sur I si et seulement si $f'' \leqslant 0$ sur I.

Exemples :

- La fonction exponentielle est convexe sur $\mathbb{R}$.
- La fonction ln est concave sur $]0 ; +\infty[$.
- La fonction sinus est concave sur $[0 ; \pi[$; la fonction cosinus concave sur $[0 ; \dfrac{\pi}{2}]$ et convexe sur $[\dfrac{\pi}{2} ; \pi]$.

- Cas des fonctions puissances : $f : x \mapsto x^h$ pour $h \in\ ]0 ; +\infty[$. f est de classe C^∞ et on a :

  $f''(x) = h(h-1)\ x^{h-2} \quad \forall x > 0$. D'où :

$$f \text{ convexe } \Leftrightarrow h \leqslant 0 \text{ ou } h \geqslant 1$$

$$f \text{ concave } \Leftrightarrow h \in [0 ; 1]$$

Etc ...

III. Quelques propriétés des fonctions convexes

<u>III.1. Propriété</u>

Soit f une fonction définie sur $\mathbb{R}$, convexe, positive et admettant deux zéros distincts $t_1 < t_2$. Alors f est nulle sur $[t_1\,;\,t_2]$.

<u>Preuve</u> :

Par l'absurde, si f n'est pas nulle sur $[t_1\,;\,t_2]$ alors $\exists t \in [t_1\,;\,t_2]$ tel que $f(t) > 0$. Ce qui entraînerait :

$\varphi_t(t_1) = \dfrac{f(t_1) - f(t)}{t_1 - t} = -\dfrac{f(t)}{t_1 - t} > 0$ et $\varphi_t(t_2) = \dfrac{f(t_2) - f(t)}{t_2 - t} = -\dfrac{f(t)}{t_2 - t} < 0$ ce qui contredit II.2. (φ_t croissante)

<u>III.2. Propriété</u>

Soit f une fonction définie sur $\mathbb{R}$, convexe et majorée. Alors f est constante.

<u>Preuve</u> :

Soit M un réel tel que $|f(x)| \leqslant M \ \forall x \in \mathbb{R}$.Considérons $\varphi_0(x) = \dfrac{f(x) - f(0)}{x}$ pour $x \in \mathbb{R}^*$.

Pour $x > 0$, on a : $\varphi_0(x) \leqslant \dfrac{M - f(0)}{x}$. φ_0 est croissante sur $\mathbb{R}^*$ donc admet une limite en $+\infty$ et $\lim\limits_{x\to+\infty} \varphi_0(x) \leqslant 0$.

Pour $x < 0$, on a : $\varphi_0(x) \geqslant \dfrac{M - f(0)}{x}$. φ_0 est croissante sur $\mathbb{R}^*$ donc admet une limite en $-\infty$ et $\lim\limits_{x\to-\infty} \varphi_0(x) \geqslant 0$.

Donc φ_0 est nulle sur $\mathbb{R}^*$ et, par suite, $f(x) = f(0) \ \forall x \in \mathbb{R}$: f est constante sur $\mathbb{R}$.

Notons que ce résultat ne subsiste pas si f est définie sur $[0\,;+\infty[$. ($f : x \mapsto \dfrac{1}{1+x}$ pour $x > 0$ est majorée par 1 et décroissante)

<u>III.3. Propriété</u>

Soit f une fonction définie sur $\mathbb{R}$, convexe et croissante et non constante. Alors $\lim\limits_{+\infty} f = +\infty$.

<u>Preuve</u> :

Puisque f est croissante, $\lim\limits_{+\infty} f$ existe. Si $\lim\limits_{+\infty} f = \ell \in \mathbb{R}$, alors $\lim\limits_{+\infty} \varphi_0 = 0$. Or $\varphi_0(1) = f(1) - f(0) \geqslant 0$ puisque f est croissante. Et comme φ_0 est croissante, on φ_0 nulle sur $\mathbb{R}^*$ donc f constante sur $\mathbb{R}$, ce qui contredit l'hypothèse. Donc $\lim\limits_{+\infty} f = +\infty$.

IV. Applications

Démontrer que la fonction $x \mapsto \dfrac{\mathbf{e}^x - 1}{x}$ est croissante sur $\mathbb{R}^*$.

La fonction exponentielle étant convexe, on a en utilisant la caractérisation 2 pour $\alpha = 0$: $\varphi_0(x) = \dfrac{\mathbf{e}^x - 1}{x}$ d'où le résultat.

Démontrer que la fonction $x \mapsto \dfrac{\sin x}{x}$ est décroissante sur $]0 ; \pi]$.

La fonction sinus étant concave sur $]0 ; \pi]$, on a d'après la caractérisation 2 pour $\alpha = 0$: $\varphi_0(x) = \dfrac{\sin x}{x}$ d'où le résultat.

Démontrer que $\forall x \in [0 ; \dfrac{\pi}{2}]$, $\dfrac{2}{\pi} x \leqslant \sin x \leqslant x$

Comme la fonction sinus est concave sur $[0 ; \pi]$, on a : $\sin(\lambda x + (1-\lambda)y) \geqslant \lambda \sin x + (1-\lambda) \sin y$

Avec $x = \dfrac{\pi}{2}$, $y = 0$ et $\lambda = \dfrac{2}{\pi} x \in [0 ; 1]$, il vient : $\sin x \geqslant \dfrac{2}{\pi} x$.

De plus, d'après II.5. le graphe de la fonction sinus est en dessous de ses tangentes (concavité) :

$$\sin x \leqslant \sin \alpha + (x - \alpha)\cos \alpha$$

Avec $\alpha = 0$, il vient $\sin x \leqslant x$.

MOYENNES ARITHMÉTIQUES ET GÉOMÉTRIQUES

La moyenne arithmétique de n nombres réels strictement positifs est toujours au moins égale à la moyenne géométrique de ces nombres :

$$\forall x_1, ..., x_n \in]0 ; +\infty[: \frac{x_1 + ... + x_n}{n} \geqslant \sqrt[n]{x_1 ... x_n}$$

<u>Preuve</u> :

La fonction ln est concave sur $]0 ; +\infty[$ donc :

$$\forall x_1, ..., x_n \in]0 ; +\infty[\text{ et } \forall \lambda_1, ..., \lambda_n \in [0 ; 1] \text{ tels que } \sum_{i=1}^{n} \lambda_i = 1 : \ln(\lambda_1 x_1 + ... + \lambda_n x_n) \geqslant \lambda_1 \ln x_1 + ... + \lambda_n \ln x_n$$

En composant par la fonction exponentielle qui est croissante : $\boxed{\lambda_1 x_1 + ... + \lambda_n x_n \geqslant x_1^{\lambda_1} ... x_n^{\lambda_n}}$

En choisissant $\lambda_1 = ... = \lambda_n = \dfrac{1}{n}$, on obtient : $\dfrac{x_1 + ... + x_n}{n} \geqslant \sqrt[n]{x_1 ... x_n}$

INÉGALITÉ DE HÖLDER

Soient p et q deux nombres réels strictement supérieurs à 1 tels que $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

Soient également $a_1, \dots, a_n, b_1, \dots, b_n$ $2n$ nombres réels strictement positifs. Alors on a :

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} b_i^q\right)^{\frac{1}{q}}$$

<u>Preuve</u> :

Comme $p > 1$ la fonction puissance $f : x \mapsto x^p$ est convexe sur $]0\,;+\infty[$. On a donc :

$$\forall\, x_1, \dots, x_n \in\ ]0\,;+\infty[\text{ et } \forall \lambda_1, \dots, \lambda_n \in [0\,;1] \text{ tels que} \sum_{i=1}^{n} \lambda_i = 1 : f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i)$$

$$\left(\sum_{i=1}^{n} \lambda_i x_i\right)^p \leqslant \sum_{i=1}^{n} \lambda_i x_i^p$$

En particulier pour $x_i = \dfrac{a_i}{b_i^{q-1}}$ et $\lambda_i = \dfrac{b_i^q}{\sum_{i=1}^{n} b_i^q}$ il vient : $\left(\dfrac{\sum_{i=1}^{n} a_i b_i}{\sum_{i=1}^{n} b_i^q}\right)^p \leqslant \dfrac{\sum_{i=1}^{n} a_i^p b_i^{q-p(q-1)}}{\sum_{i=1}^{n} b_i^q}$

Or, $q - pq + p = 0$ puisque $\dfrac{1}{p} + \dfrac{1}{q} = 1$ d'où : $\left(\sum_{i=1}^{n} a_i b_i\right)^p \leqslant \sum_{i=1}^{n} a_i^p \left(\sum_{i=1}^{n} b_1^q\right)^{p-1}$

Puis en extrayant les racine $p^{\text{èmes}}$ des deux membres (la fonction $u \mapsto u^{\frac{1}{p}}$ étant croissante car $p > 1$) et en remarquant que $\dfrac{p-1}{p} = \dfrac{1}{q}$, il vient : $\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} b_i^q\right)^{\frac{1}{q}}$

Cas particulier : pour $p = q = 2$, on récupère l'inégalité de Cauchy-Schwarz : $< a \mid b > \leqslant \|a\|\,\|b\|$ (pour la structure euclidienne usuelle de $\mathbb{R}^n$)

INÉGALITÉ DE MINKOWSKI

Soit $p > 1$ et soient $a_1, \dots, a_n, b_1, \dots, b_n$ $2n$ nombres réels strictement positifs. Alors on a :

$$\left(\sum_{i=1}^{n} (a_i + b_i)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}}$$

<u>Preuve</u> :

On considère la fonction f définie sur $]0\,;+\infty[$ par $f(x) = \left(1 + x^{\frac{1}{p}}\right)^p$. f est de classe C^∞ et on a :

$$f'(x) = x^{\frac{1}{p}-1}\left(1 + x^{\frac{1}{p}}\right)^{p-1} \quad \text{et} \quad f''(x) = \frac{1-p}{p} x^{\frac{1}{p}-2}\left(1 + x^{\frac{1}{p}}\right)^{p-1} + x^{\frac{1}{p}-1} \frac{p-1}{p} x^{\frac{1}{p}-1}\left(1 + x^{\frac{1}{p}}\right)^{p-2}$$

$$f''(x) = \frac{1-p}{p} x^{\frac{1}{p}-2}\left(1+x^{\frac{1}{p}}\right)^{p-2}\left[1+x^{\frac{1}{p}}-x^{\frac{1}{p}}\right] = \frac{1-p}{p} x^{\frac{1}{p}-2}\left(1+x^{\frac{1}{p}}\right)^{p-2} \leqslant 0 \text{ car } 1-p \leqslant 0$$

Donc f est concave sur $]0\,;+\infty[$. D'où

$$\forall\, x_1, \dots, x_n \in\ ]0\,;+\infty[\text{ et } \forall \lambda_1, \dots, \lambda_n \in [0\,;1] \text{ tels que } \sum_{i=1}^{n} \lambda_i = 1 : \left(1+\left(\sum_{i=1}^{n} \lambda_i x_i\right)^{\frac{1}{p}}\right)^p \geqslant \sum_{i=1}^{n} \lambda_i \left(1+x_i^{\frac{1}{p}}\right)^p$$

c'est-à-dire :
$$\left(1+\left(\sum_{i=1}^{n} \lambda_i x_i\right)^{\frac{1}{p}}\right)^p \geqslant \sum_{i=1}^{n} \left(\lambda_i^{\frac{1}{p}} + (\lambda_i x_i)^{\frac{1}{p}}\right)^p$$

Soient alors $a_1, \dots, a_n, b_1, \dots, b_n$ $2n$ nombres réels strictement positifs. Posons :

$$S = \sum_{i=1}^{n} a_i^p \text{ et } \lambda_i = \frac{a_i^p}{S} \text{ (ainsi } \sum_{i=1}^{n} \lambda_i = 1\text{) et } x_i = \frac{b_i^p}{a_i^p}$$

D'où :
$$\left(1+\left(\sum_{i=1}^{n} \frac{b_i^p}{S}\right)^{\frac{1}{p}}\right)^p \geqslant \sum_{i=1}^{n} \left(\frac{a_i}{S^{\frac{1}{p}}} + \frac{b_i}{S^{\frac{1}{p}}}\right)^p$$

En multipliant tout par S, il vient :
$$\left(S^{\frac{1}{p}}+\left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}}\right)^p \geqslant \sum_{i=1}^{n} (a_i + b_i)^p$$

Puis en extrayant les racine $p^{\text{èmes}}$ des deux membres, (la fonction $u \mapsto u^{\frac{1}{p}}$ étant croissante car $p > 1$) :

$$\left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}} \geqslant \left(\sum_{i=1}^{n} (a_i + b_i)^p\right)^{\frac{1}{p}}$$

Cas particulier : si $p = 2$ on récupère l'inégalité triangulaire : $\|a + b\| \leqslant \|a\| + \|b\|$

AUTRE DÉMONSTRATION DE L'INÉGALITÉ DE MINKOWSKI À L'AIDE D'UN THÉOREME SUR LA CONVEXITÉ DES FONCTIONS POSITIVEMENT HOMOGÈNES

<u>Définition</u>

Soit E un espace vectoriel sur $\mathbb{R}$ et $f : E \longrightarrow \mathbb{R}$ une application.

On dit que f est positivement homogène si :

$(\forall x \in E), (\forall \lambda \in \mathbb{R}^+)$: $\qquad f(\lambda x) = \lambda f(x)$

<u>Exemple</u> : $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x = (x_1, ..., x_n) \mapsto \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$ est positivement homogène, $\forall p \in\ ]0 ; +\infty[$.

<u>Théorème</u>

Soit E un espace vectoriel sur $\mathbb{R}$ et $f : E \longrightarrow \mathbb{R}$ une application positivement homogène.

1) f convexe sur $E \iff \forall x, y \in E \ \ f(x+y) \leqslant f(x) + f(y)$

De plus, si $f \geqslant 0$ sur E :

2) f convexe sur $E \iff B = \{x \in E \mid f(x) \leqslant 1\}$ convexe

<u>Démonstration</u> :

1) $\Rightarrow$: Soient $x, y \in E$. Comme f est convexe, on a : $f(\frac{1}{2}x + \frac{1}{2}y) \leqslant \frac{1}{2}f(x) + \frac{1}{2}f(y)$

Et comme f est positivement homogène, il vient $f(x+y) \leqslant f(x) + f(y)$

$\Leftarrow$: Soient $x, y \in E$, et $\lambda, \mu \in [0 ; 1]$ tels que $\lambda + \mu = 1$.

Par hypothèse, on a : $f(\lambda x + \mu y) \leqslant f(\lambda x) + f(\mu y)$

Et comme f est positivement homogène : $f(\lambda x + \mu y) \leqslant \lambda f(x) + \mu f(y)$, donc f est convexe sur E

2) $\Rightarrow$: Soient $x, y \in B$, et $\lambda, \mu \in [0 ; 1]$ tels que $\lambda + \mu = 1$.

Par hypothèse, on a : $f(\lambda x + \mu y) \leqslant \lambda f(x) + \mu f(y)$

Et comme $x, y \in B$: $f(\lambda x + \mu y) \leqslant \lambda + \mu \leqslant 1$

Donc $\lambda x + \mu y \in B$, donc B est convexe.

Notons que la démonstration ci-dessus n'utilise pas l'homogénéité de f.

$\Leftarrow$: Soient $x, y \in E$. Il suffit, d'après 1) de montrer que $f(x+y) \leqslant f(x) + f(y)$

Soient $\alpha \in\ ]f(x) ; +\infty[$ et $\beta =]f(y) ; +\infty[$. Ainsi, on a : $\frac{f(x)}{\alpha} < 1$ et $\frac{f(y)}{\beta} < 1$.

C'est-à-dire, f étant homogène : $f\left(\frac{x}{\alpha}\right) < 1$ et $f\left(\frac{y}{\beta}\right) < 1$

Donc, $\frac{x}{\alpha}$ et $\frac{y}{\beta}$ sont éléments de B. Ce dernier étant convexe, on a : $\forall\lambda,\ \in [0 ; 1]$ tels que $\lambda + \mu = 1$:

$$\lambda \frac{x}{\alpha} + \mu \frac{y}{\beta} \in B,$$

En particulier, en choisissant $\lambda = \frac{\alpha}{\alpha+\beta}$ et $\mu = \frac{\beta}{\alpha+\beta}$, il vient : $\frac{x}{\alpha+\beta} + \frac{y}{\alpha+\beta} \in B$.

Autrement dit, f étant positivement homogène : $f(x+y) \leqslant \alpha + \beta$

En passant à la limite (lorsque α tend vers $f(x)$ et lorsque β tend vers $f(y)$), il vient :

$$f(x+y) \leqslant \alpha + \beta$$

C'est-à-dire $f(x+y) \leqslant f(x) + f(y)$ et d'après 1), f est convexe sur E.

Démonstration de l'inégalité de Minkowski :

On considère l'application : $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x = (x_1, ..., x_n) \longmapsto \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$ ou $p \in [1 ; +\infty[$.

Déjà, Comme $p \geqslant 1$, f^p est convexe :

Considérons, pour tout $i \in [\![1 ; n]\!]$ les fonctions suivantes :

$$g_i : \mathbb{R}^n \longrightarrow \mathbb{R},\ x = (x_1, ..., x_n) \longmapsto |x_i|^p \quad \text{et} \quad h_i : \mathbb{R} \longrightarrow \mathbb{R},\ x_i \longmapsto |x_i|^p.$$

Les fonctions h_i sont convexes sur $\mathbb{R}^+$ (puisque dérivables sur $\mathbb{R}^+$ lorsque $p > 1$ et $h_i'(x_i) = px^{p-1}$ avec h_i' croissante sur $[0 ; +\infty[$ avec le prolongement $h_i'(0) = 0$). Par parité, on en déduit que les fonctions h_i sont convexes sur $\mathbb{R}$.

On en déduit que les fonctions g_i sont convexes sur $\mathbb{R}^n$: $\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, $\forall (\lambda, \mu) \in [0 ; 1]^2$ tels que $\lambda + \mu = 1$, on a : $g_i(\lambda x + \mu y) = |\lambda x_i + \mu y_i|^p = h_i(\lambda x_i + \mu y_i) \leqslant \lambda h_i(x_i) + \mu h_i(y_i) = \lambda g(x) + \mu g(y)$.

Donc, d'après le théorème (assertion 2 $\Rightarrow$) appliqué à f^p, l'ensemble $B_p = \{x \in E \mid f^p(x) \leqslant 1\}$ est convexe.

Or, $f^p(x) \leqslant 1 \Leftrightarrow f(x) \leqslant 1$, donc $B_p = \{x \in E \mid f(x) \leqslant 1\}$. Ce dernier étant convexe, on en déduit d'après le théorème (assertion 2 $\Leftarrow$) que f est convexe d'où (théorème assertion 1 $\Rightarrow$) :

$$\forall x, y \in \mathbb{R}^n : f(x+y) \leqslant f(x) + f(y)$$

Ce qui s'écrit encore (en notant $x = (x_1, ..., x_n)$ et $y = (y_1, ..., y_n)$) :

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}$$