Concavité, convexité des fonctions réelles à variable réelle

1) **Motivation**

Rappelons d'abord qu'une partie d'un plan est dite **convexe** si tout segment reliant deux points quelconques de cette partie se situe intégralement dans cette partie et qu'elle est concave si son complémentaire est convexe.

Exemples : Un disque (avec ou sans le cercle en déterminant la frontière) est convexe, tout comme l'intérieur d'un triangle ou d'un rectangle.

On cherche alors à caractériser les fonctions définies sur un intervalle I et telles que leur courbe représentative $\mathcal{C}$ dans un plan $\mathcal{P}$ muni d'un repère orthogonal, présente la propriété suivante :

L'ensemble des points $M(x, y)$ du plan $\mathcal{P}$ tels que : $y \geq f(x)$ c'est-à-dire des points situés au-dessus de la courbe $\mathcal{C}$ est soit convexe (on dira que la fonction est convexe sur I) soit concave (on dira que la fonction est concave sur I)

A noter que pour une fonction dérivable sur I, être convexe signifie que sa courbe représentative $\mathcal{C}$ se situe au-dessus de chacune de ses tangentes et pour une fonction concave, en-dessous de chacune de ses tangentes.

Exemples : La fonction $x \to x^2$ est une fonction convexe sur $\mathbb{R}$. La fonction $x \to x^3$ est une fonction convexe sur $[0, +\infty[$ et concave sur $]-\infty, 0]$.

Remarques :

Pour une fonction convexe sur I, si on se donne deux points $M_1(x_1, y_1), M_2(x_2, y_2)$ situés au-dessus de la courbe $\mathcal{C}$ un point $M(x, y)$ du segment $[M_1, M_2]$ peut être paramétré par un réel $t \in [0,1]$ tel que :

$$\overrightarrow{M_1M} = t\,\overrightarrow{M_1M_2}$$

Ce qui conduit à :

$$\begin{cases} x - x_1 = t\,(x_2 - x_1) \\ y - y_1 = t\,(y_2 - y_1) \end{cases}$$

Soit encore :

$$\begin{cases} x = (1-t)\,x_1 + t\,x_2 \\ y = (1-t)\,y_1 + t\,y_2 \end{cases}$$

Que l'on peut encore écrire de façon plus compacte, en utilisant le produit externe sur l'espace vectoriel $\mathbb{R}^2$, sous forme :

$$(x, y) = (1-t).(x_1, y_1) + t.(x_2, y_2)$$

2) Définition

Soit $\boldsymbol{f}$ une fonction d'un intervalle $\boldsymbol{I}$ de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$ et le sous ensemble de $\mathbb{R}^2$ défini par :

$$\boldsymbol{C^+ = \{(x,y) \in \mathbb{R}^2 : y \geq f(x)\}}$$

On dit que $\boldsymbol{f}$ est convexe si $\boldsymbol{C^+}$ est convexe c'est-à-dire :

$$\forall \left((x_1,y_1),(x_2,y_2)\right) \in \left(\mathbb{R}^2\right)^2 :$$

$$\left((x_1,y_1),(x_2,y_2)\right) \in (C^+)^2 \Rightarrow \forall\, t \in [0,1] : (1-t).(x_1,y_1) + t.(x_2,y_2) \in C^+$$

A noter que :

$$(1-t).(x_1,y_1) + t.(x_2,y_2) \in \mathcal{C}^+ \Leftrightarrow (1-t)\, y_1 + t\, y_2 \geq f\big((1-t)\, x_1 + t\, x_2\big)$$

Ceci amène à une caractérisation équivalente :

$\boldsymbol{f}$ est convexe sur un intervalle $\boldsymbol{I}$

$$\Leftrightarrow \forall\, t \in [0,1]\ \forall\, (x_1,x_2) \in I^2 :\ f\big((1-t)\, x_1 + t\, x_2\big) \leq (1-t)\, f(x_1) + t\, f(x_2)$$

qui s'interprète géométriquement sur la représentation graphique de f par le fait qu'un point situé sur la corde reliant deux points de la courbe $M_1\big(x_1, f(x_1)\big)$ et $M_2\big(x_2, f(x_2)\big)$ se situe au-dessus du point de la courbe de même abscisse.

De façon analogue, on définit une fonction concave sur un intervalle I comme une fonction telle que $\mathcal{C}^+$ soit concave et on aboutit à la caractérisation équivalente :

$\boldsymbol{f}$ est concave sur un intervalle $\boldsymbol{I}$

$$\Leftrightarrow \forall\, t \in [0,1]\ \forall\, (x_1,x_2) \in I^2 :\ f\big((1-t)\, x_1 + t\, x_2\big) \geq (1-t)\, f(x_1) + t\, f(x_2)$$

On peut noter également :

$\boldsymbol{f}$ est concave sur l'intervalle $\boldsymbol{I} \Leftrightarrow -\boldsymbol{f}$ est convexe sur l'intervalle $\boldsymbol{I}$

Remarque :

On définit également les concepts de fonction **strictement convexe** ou **strictement concave sur $\boldsymbol{I}$** en remplaçant l'inégalité large par une inégalité stricte mais pour $t \in]0,1[$

3) Caractérisation de la concavité et de la convexité pour une fonction dérivable

Soit f une fonction définie et dérivable sur un intervalle I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$. Alors :

$\boldsymbol{f}$ **est (resp. strictement) convexe sur** $\boldsymbol{I}$ $\Leftrightarrow$ $\boldsymbol{f'}$ **est (resp. strictement) croissante sur** $\boldsymbol{I}$

Dont on déduit

$\boldsymbol{f}$ **est (resp. strictement) concave sur** $\boldsymbol{I}$ $\Leftrightarrow$ $\boldsymbol{f'}$ **est (resp. strictement) décroissante sur** $\boldsymbol{I}$

Si de plus, f est deux fois dérivable sur I :

$\boldsymbol{f}$ **est convexe sur** $\boldsymbol{I}$ $\Leftrightarrow \forall\, \boldsymbol{x} \in \boldsymbol{I} :\ \boldsymbol{f''(x)} \geq \mathbf{0}$

Dont on déduit :

$\boldsymbol{f}$ **est concave sur** $\boldsymbol{I}$ $\Leftrightarrow \forall\, \boldsymbol{x} \in \boldsymbol{I} :\ \boldsymbol{f''(x)} \leq \mathbf{0}$

Pour le caractère strictement convexe ou strictement concave, il suffit d'ajouter que $\boldsymbol{f''}$ **ne s'annule qu'en des points isolés.**

Preuve :

Seule la première caractérisation est à prouver, les autres étant évidentes. Nous allons le faire dans le cas de la convexité.

Supposons que f' soit croissante sur I. Pour un couple $(x_1, x_2) \in I^2$ considérons la fonction g définie sur $[0,1]$ par :

$$g(t) = f\big((1-t)\,x_1 + t\,x_2\big) - \big((1-t)\,f(x_1) + t\,f(x_2)\big)$$

Cette fonction est dérivable sur I et :

$$g'(t) = (x_2 - x_1)\,f'\big((1-t)\,x_1 + t\,x_2\big) - \big(f(x_2) - f(x_1)\big)$$

$$= (x_2 - x_1)\left(f'\big(x_1 + t\,(x_2 - x_1)\big) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}\right)$$

Or, le théorème des accroissements finis permet d'affirmer l'existence d'un réel c du segment $]x_1, x_2[$ donc pouvant être mis sous forme $c = (1-t_0)\,x_1 + t_0\,x_2$ avec $t_0 \in\,]0,1[$ et tel que :

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'\big(x_1 + t_0\,(x_2 - x_1)\big)$$

Donc :

$$g'(t) = (x_2 - x_1)\,\Big(f'\big(x_1 + t\,(x_2 - x_1)\big) - f'\big(x_1 + t_0\,(x_2 - x_1)\big)\Big)$$

Ainsi :

$$g'(t) \leq 0 \text{ sur } [0, t_0] \text{ et } g'(t) \geq 0 \text{ sur } [t_0, 1]$$

De plus :

$$g(0) = g(1) = 0$$

On en déduit :

$$\forall\, t \in [0,1] :\; g(t) \leq 0$$

Ce qui prouve que f est convexe sur I

Réciproquement : si f est convexe sur I alors soit $(x_1, x_2) \in I^2$ et $t \in [0,1]$. On a :

$$f\big((1-t)\, x_1 + t\, x_2\big) \leq (1-t)\, f(x_1) + t\, f(x_2)$$

Donc :

$$f\big(x_1 + t\,(x_2 - x_1)\big) \leq f(x_1) + t\,\big(f(x_2) - f(x_1)\big)$$

D'où :

$$\frac{f\big(x_1 + t\,(x_2 - x_1)\big) - f(x_1)}{t\,(x_2 - x_1)} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

En faisant tendre t vers 0 on en déduit par passage à la limite :

$$f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Considérons alors l'ensemble :

$$J = \left\{x \in [x_1, x_2] :\; \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}\right\}$$

J n'est pas vide car il contient x_1 et il est majoré parx_2 donc il admet une borne supérieure s. Celle-ci est limite d'une suite (s_n) de points de J qui vérifient donc :

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(s_n)}{x_2 - s_n} \quad (1)$$

Par passage à la limite et continuité de f en s, on en déduit :

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(s)}{x_2 - s}$$

Supposons par l'absurde que $s < x_2$. En appliquant le théorème des accroissements finis, nous aurions l'existence d'un point $d \in]s, x_2[$ tel que :

$$\frac{f(x_2) - f(s)}{x_2 - s} = f'(d)$$

Or d'après ce qui précède :

$$f'(d) \leq \frac{f(x_2) - f(d)}{x_2 - d}$$

Donc $d \in J$ et $d > s$ ce qui est absurde donc $s = x_2$. Ainsi x_2 est limite d'une suite (s_n) de points de J.

Et en faisant tendre n vers l'infini dans (1) on a la suite (s_n) qui tend vers x_2 et le membre de droite qui tend vers $f'(x_2)$. On en déduit, par passage à la limite :

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'(x_2)$$

D'où :

$$\boxed{f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'(x_2)}$$

f' est donc croissante sur I.

Remarque : l'encadrement précédent s'interprète par le fait que le coefficient directeur d'une corde est compris entre le coefficient directeur de la tangente au point de l'extrémité gauche de la corde et le coefficient directeur de la tangente au point de l'extrémité droite, ce qui parait évident d'un point de vue géométrique, mais devait être démontré d'un point de vue analytique.

4) **Propriétés** :

Si f est une fonction convexe sur un intervalle I alors :

$$\forall\, (x_1, x_2, .., x_n) \in I^n \; \forall\, (t_1, t_2, .., t_n) \in \mathbb{R}^n :$$

$$\begin{cases} x_1 < x_2 <..< x_n, \\ \forall\, i \in [\![1, n]\!] : t_i > 0 \\ t_1 + t_2 + .. + t_n = 1 \end{cases} \Rightarrow f\left(\sum_{i=1}^{n} t_i\, x_i\right) \leq \sum_{i=1}^{n} t_i\, f(x_i)$$

Preuve : par récurrence sur n :

L'initialisation a déjà été vue, ne reste que l'hérédité : On suppose donc la propriété établie au rang n.

Soit $(x_1, x_2, .., x_n, x_{n+1}) \in I^{n+1}$ $(t_1, t_2, .., t_n, t_{n+1}) \in \mathbb{R}^{n+1}$ tels que :

$x_1 < x_2 <..< x_n < x_{n+1},\; t_1 + t_2 + .. + t_n + t_{n+1} = 1$ et $\forall\, i \in [\![1, n+1]\!] :\; t_i > 0$ alors :

$$f\left(\sum_{i=1}^{n+1} t_i\, x_i\right) = f\left(\sum_{i=1}^{n} t_i\, x_i + t_{n+1}\, x_{n+1}\right)$$

$$= f\left((1 - t_{n+1}) \frac{\sum_{i=1}^{n} t_i\, x_i}{1 - t_{n+1}} + t_{n+1}\, x_{n+1}\right) \leq (1 - t_{n+1})\, f\left(\frac{\sum_{i=1}^{n} t_i\, x_i}{1 - t_{n+1}}\right) + t_{n+1}\, f(x_{n+1})$$

Posons pour $i \in [\![1, n]\!]$:

$$h_i = \frac{t_i}{1 - t_{n+1}} = \frac{t_i}{\sum_{i=1}^{n} t_i}$$

Alors :

$$\forall\, i \in [\![1, n]\!] :\ h_i > 0\,, \qquad \sum_{i=1}^{n} h_i \ = 1$$

Donc :

$$f\left(\frac{\sum_{i=1}^{n} t_i\, x_i}{1 - t_{n+1}}\right) = f\left(\sum_{i=1}^{n} h_i\, x_i\right) \leq \sum_{i=1}^{n} h_i\, f(x_i) = \frac{1}{1 - t_{n+1}} \sum_{i=1}^{n} t_i\, f(x_i)$$

D'où :

$$f\left(\sum_{i=1}^{n+1} t_i\, x_i\right) \leq \sum_{i=1}^{n} t_i\, f(x_i) + t_{n+1}\, f(x_{n+1}\,)$$

Ce qui montre l'hérédité

5) **<u>Applications</u>** :

La convexité ou la concavité présente de nombreux intérêts dans l'analyse mathématique dont voici deux exemples :

- Un encadrement de la fonction sinus sur $\left[0, \frac{\pi}{2}\right]$: on encadre la fonction qui est concave sur cet intervalle par sa corde et sa tangente en 0, ce qui donne :

$$\frac{2}{\pi}\, x \leq sin(x) \leq x$$

- Une limite (oral polytechnique 1982) :

$$\lim_{n \longrightarrow +\infty} \left(\frac{1^{\frac{1}{n}} + 2^{\frac{1}{n}} + \cdots + n^{\frac{1}{n}}}{n}\right)^{n}$$

Pour cette dernière on commence par établir une propriété qui permet de comparer la moyenne arithmétique et la moyenne géométrique de n réels strictement positifs $x_1, x_2, \dots, x_n$ définies comme étant :

- Pour la moyenne arithmétique :

$$\frac{x_1 + x_2 + \cdots + x_n}{n}$$

- Pour la moyenne géométrique :

$$(x_1 x_2 \dots x_n)^{\frac{1}{n}}$$

On considère pour cela la fonction :

$$f(x) = Ln(x)$$

Cette fonction est concave sur $\mathbb{R}$ donc :

$$Ln\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) \geq \frac{Ln(x_1) + Ln(x_2) + \cdots + Ln(x_n)}{n} = Ln(x_1 x_2 \dots x_n)^{\frac{1}{n}}$$

On en déduit :

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq (x_1 x_2 \dots x_n)^{\frac{1}{n}}$$

La moyenne arithmétique est donc supérieure ou égale à la moyenne géométrique.

Appliqué à notre problème de limite, nous en déduisons :

$$\frac{1^{\frac{1}{n}} + 2^{\frac{1}{n}} + \cdots + n^{\frac{1}{n}}}{n} \geq \left(1^{\frac{1}{n}}\, 2^{\frac{1}{n}} \dots n^{\frac{1}{n}}\right)^{\frac{1}{n}}$$

Soit :

$$\left(\frac{1^{\frac{1}{n}} + 2^{\frac{1}{n}} + \cdots + n^{\frac{1}{n}}}{n}\right)^n \geq 1^{\frac{1}{n}}\, 2^{\frac{1}{n}} \dots n^{\frac{1}{n}} = n!^{\frac{1}{n}}$$

Or :

$$n!^{\frac{1}{n}} = e^{\frac{1}{n}Ln(n!)}$$

Et :

$$\frac{1}{n}\,Ln(n!) = \frac{Ln(1) + Ln(2) + \cdots + Ln(n)}{n}$$

Le théorème de Césaro indique que cette suite a la même limite que la suite $\big(Ln(n)\big)$ donc tend vers $+\infty$.

Par comparaison, la suite initiale tend vers $+\infty$.