FONCTIONS CONVEXES (programme de PCSI)

I. <u>Définition et interprétation géométrique</u>.

<u>Définition.</u> Soit I un intervalle de $\mathbb{R}$, soit $f : I \to \mathbb{R}$ une application. On dit que f est **convexe** sur I si on a

$$\forall (x,y) \in I^2 \quad \forall \lambda \in [0,1] \qquad f\big((1-\lambda)x + \lambda y\big) \leq (1-\lambda)\, f(x) + \lambda\, f(y) \qquad \textbf{(*)}\ .$$

Il est clair qu'il suffit que l'inégalité **(*)** soit satisfaite lorsque $x < y$.

<u>Commentaire.</u> Si x et y sont dans I avec $x < y$, alors $(1-\lambda)x + \lambda y$ décrit le segment $[x, y]$ (qui est inclus dans I) lorsque λ décrit $[0, 1]$, autrement dit

$$\big\{(1-\lambda)x + \lambda y\ ;\ 0 \leq \lambda \leq 1\big\} = [x,y] = \big\{z \in \mathbb{R}\ \big|\ x \leq z \leq y\big\}\ .$$

En effet, si $0 \leq \lambda \leq 1$, *alors* $x = (1-\lambda)x + \lambda x \leq (1-\lambda)x + \lambda y \leq (1-\lambda)y + \lambda y = y$, *ce qui montre l'inclusion* $\big\{(1-\lambda)x + \lambda y\ ;\ 0 \leq \lambda \leq 1\big\} \subset [x,y]$.

Réciproquement, si $z \in [x,y]$, *i.e. si* $x \leq z \leq y$, *alors en posant* $\lambda = \dfrac{z-x}{y-x}$, *on a bien* $0 \leq \lambda \leq 1$ *et* $z = (1-\lambda)x + \lambda y$, *ce qui montre l'inclusion inverse.*

<u>Interprétation géométrique.</u> Soit $f : I \to \mathbb{R}$, soient a et b des points de I avec $a < b$. Notons $\varphi_{a,b}$ l'unique fonction affine coïncidant avec f en les points a et b. On a alors

$$\varphi_{a,b}(x) = \frac{f(b)-f(a)}{b-a}\,(x-a) + f(a)\ .$$

On reconnaît en effet, en écrivant $y = \varphi_{a,b}(x)$, *l'équation de l'unique droite affine passant par les points* $A\big(a\ ,\ f(a)\big)$ *et* $B\big(b\ ,\ f(b)\big)$. Le segment de cette droite joignant les points A et B du plan, que l'on peut noter $[AB]$, est ce qu'on appelle une **sécante** (ou une **corde**) du graphe de la fonction f. Si $\lambda \in [0, 1]$, on vérifie facilement que

$$\varphi_{a,b}\big((1-\lambda)a + \lambda b\big) = \frac{f(b)-f(a)}{b-a}\,\lambda(b-a) + f(a) = (1-\lambda)\, f(a) + \lambda\, f(b)\ .$$

La convexité de la fonction f sur I s'exprime alors en disant que

$$\forall (a,b) \in I^2 \text{ avec } a < b \qquad \forall x \in [a,b] \qquad f(x) \leq \varphi_{a,b}(x)\ .$$

Autrement dit, **une fonction f est convexe sur I si et seulement si tout arc du graphe de f est situé en-dessous de sa sécante.**

<u>Définition.</u> Soit I un intervalle de $\mathbb{R}$, soit $f : I \to \mathbb{R}$ une application. On dit que f est **concave** sur I si on a

$$\forall (x,y) \in I^2 \quad \forall \lambda \in [0,1] \qquad f\big((1-\lambda)x + \lambda y\big) \geq (1-\lambda)\, f(x) + \lambda\, f(y)\ .$$

<u>Commentaires.</u> On a simplement changé le sens de l'inégalité par rapport aux fonctions convexes. Une fonction f est donc concave sur I si et seulement si $-f$ est convexe.

<u>Interprétation géométrique</u>. Une fonction f est concave sur I si et seulement si tout arc du graphe de f est situé au-dessus de sa sécante.

<u>Exemple.</u> La fonction $f : x \mapsto x^2$ est convexe sur $\mathbb{R}$.

En effet, si x et y sont des réels et si $\lambda \in [0,1]$, alors

$$\begin{aligned}
(1-\lambda)\, f(x) + \lambda\, f(y) - f\big((1-\lambda)x + \lambda y\big) &= (1-\lambda)x^2 + \lambda y^2 - \big((1-\lambda)x + \lambda y\big)^2 \\
&= \big((1-\lambda) - (1-\lambda)^2\big)\, x^2 - 2\lambda(1-\lambda)\, xy + (\lambda - \lambda^2)\, y^2 \\
&= \lambda\,(1-\lambda)\,(x-y)^2\ \geq\ 0\ .
\end{aligned}$$

II. Fonctions convexes et dérivation.

Commençons par énoncer deux lemmes dont la connaissance n'est pas exigée par le programme, mais qui sont intéressants en soi.

Lemme 1. Si $f : I \to \mathbb{R}$ est convexe, alors pour tout $a \in I$, l'application

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

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

Commentaire. Pour une fonction f convexe, le taux de variation de f entre les points a et x est donc une fonction croissante de x, le point $a \in I$ étant fixé. On parle de "**croissance de la pente des sécantes dont une extrémité est fixée**".

Preuve du Lemme 1.

• *Commençons pour cela par prouver le* **lemme des trois pentes***: si u,v,w sont trois points de I tels que $u < v < w$, alors*

$$\frac{f(v) - f(u)}{v - u} \leq \frac{f(w) - f(u)}{w - u} \leq \frac{f(w) - f(v)}{w - v} .$$

(cf. figure 3).

En effet, comme $v \in]u, w[$, il existe $\lambda \in]0, 1[$ tel que $v = (1 - \lambda)u + \lambda w$, on peut préciser que $\lambda = \dfrac{v - u}{w - u}$. Par convexité de f, on a alors **(*)***: $f(v) \leq (1 - \lambda)\, f(u) + \lambda\, f(w)$, soit $f(v) - f(u) \leq \lambda \big(f(w) - f(u)\big) = \dfrac{v - u}{w - u} \big(f(w) - f(u)\big)$, ce qui donne la première inégalité*

$$\frac{f(v) - f(u)}{v - u} \leq \frac{f(w) - f(u)}{w - u} .$$

On peut aussi réordonner **(*)** *en $f(w) - f(v) \geq (1-\lambda)\big(f(w) - f(u)\big) = \dfrac{w - v}{w - u}\big(f(w) - f(u)\big)$, ce qui donne la deuxième inégalité*

$$\frac{f(w) - f(u)}{w - u} \leq \frac{f(w) - f(v)}{w - v} .$$

• *Fixons maintenant $a \in I$. Soient x et y dans I, différents de a, tels que $x < y$, nous devons montrer que $\tau_a(x) \leq \tau_a(y)$, soit $\dfrac{f(x) - f(a)}{x - a} \leq \dfrac{f(y) - f(a)}{y - a}$. Il y a trois cas de* *figure, à savoir* $\mathbf{(1)}$*:* $x < y < a$*, ou bien* $\mathbf{(2)}$*:* $x < a < y$*, ou bien* $\mathbf{(3)}$*:* $a < x < y$*. Le lecteur s'assurera, en faisant des schémas, que cette inégalité résulte à chaque fois du lemme des trois pentes ci-dessus.*

<u>Lemme 2.</u> Soit $f : I \to \mathbb{R}$ une fonction dérivable. Alors f est convexe sur I si et seulement si sa dérivée f' est croissante sur I.

Preuve.

$\Longrightarrow$ *(sens direct). Supposons* f *convexe sur* I*. Soient* a *et* b *dans* I *avec* $a < b$*, montrons que* $f'(a) \leq f'(b)$*.*

De la définition de la dérivée comme limite d'un taux d'accroissement, on déduit que

$$f'(a) = \lim_{n \to +\infty} \tau_a\Big(a + \frac{1}{n}\Big) \qquad \text{et} \qquad f'(b) = \lim_{n \to +\infty} \tau_b\Big(b - \frac{1}{n}\Big)$$

(notations introduites dans l'énoncé du lemme 1). Or, pour n *assez grand,* $a + \dfrac{1}{n} < b - \dfrac{1}{n}$*, et a fortiori* $a + \dfrac{1}{n} < b$ *et* $a < b - \dfrac{1}{n}$*.*

Du lemme 1, on déduit alors que, pour n *grand,*

$$\tau_a\Big(a + \frac{1}{n}\Big) \leq \tau_a(b) = \tau_b(a) \leq \tau_b\Big(b - \frac{1}{n}\Big) .$$

En passant à la limite dans l'inégalité entre les membres extrêmes, on obtient $f'(a) \leq f'(b)$*.*

$\Longleftarrow$ *(sens indirect). Supposons* f' *croissante sur* I*. Soient* a *et* b *dans* I *avec* $a < b$*, soit* $\lambda \in]0, 1[$*, soit* $c = (1 - \lambda)a + \lambda b$*, on a alors* $a < c < b$*. Par l'égalité des accroissements finis, on obtient l'existence de* $u \in]a, c[$ *tel que* $\dfrac{f(c) - f(a)}{c - a} = f'(u)$*, et l'existence de* $v \in]c, b[$ *tel que* $\dfrac{f(b) - f(c)}{b - c} = f'(v)$*. Comme* $u < v$*, la croissance de* f' *entraîne* $f'(u) \leq f'(v)$*, soit*

$$\frac{f(c) - f(a)}{c - a} \leq \frac{f(b) - f(c)}{b - c} .$$

Comme $c - a = \lambda(b - a)$ *et* $b - c = (1 - \lambda)(b - a)$*, cette dernière inégalité devient*

$$\frac{f(c)-f(a)}{\lambda} \leq \frac{f(b)-f(c)}{1-\lambda} \; .$$

Enfin, en faisant les "produits en croix", on retrouve sans difficulté l'inégalité

$$f(c) = f\big((1-\lambda)a + \lambda b\big) \leq (1-\lambda)\, f(a) + \lambda\, f(b) \; ,$$

qui traduit la convexité de f.

Une conséquence immédiate de ce lemme 2 est le théorème suivant, exigible à votre programme.

Théorème. Soit $f : I \to \mathbb{R}$ une fonction deux fois dérivable. Alors f est convexe sur I si et seulement si sa dérivée seconde f'' est positive sur I.

Voici enfin un dernier résultat mentionné par votre programme:

Proposition. Soit $f : I \to \mathbb{R}$ une fonction convexe dérivable. Alors le graphe de f est situé au-dessus de chacune de ses tangentes.

Preuve. Soit $a \in I$, la tangente à la courbe $\mathcal{C}_f$ au point d'abscisse a admet pour équation

$$y = f'(a)\,(x-a) + f(a) \; .$$

Il s'agit donc de montrer que la fonction $\varphi : x \mapsto f(x) - \big(f'(a)\,(x-a) + f(a)\big)$ est à valeurs positives sur I. Or, on constate que $\varphi(a) = 0$ et que $\varphi'(x) = f'(x) - f'(a)$ est du signe de $x-a$ puisque le lemme 2 nous apprend que f' est croissante sur I. On en déduit, en dressant éventuellement un tableau de variations, que $\varphi(x) \geq \varphi(a) = 0$ partout sur I.

En rempaçant f par $-f$, on déduit bien sûr que:

- si $f : I \to \mathbb{R}$ est deux fois dérivable, alors f est concave sur I si et seulement si $f'' \leq 0$ sur I.

- si $f : I \to \mathbb{R}$ est dérivable et concave, alors le graphe de f est en-dessous de chacune de ses tangentes.

III. Inégalités de convexité.

On appelle ainsi toutes les inégalités (dont certaines seront à retenir) que l'on peut déduire de la convexité ou de la concavité de certaines fonctions, notamment en considérant la position du graphe par rapport à une tangente. Les inégalités mentionnées ci-dessous sont à connaître:

- $\boxed{\forall x \in \mathbb{R} \quad e^x \geq 1 + x}$.

En effet, la fonction exponentielle est convexe sur $\mathbb{R}$ puisque sa dérivée seconde, qui est elle-même, est positive. Le graphe de la fonction exponentielle est donc, partout sur $\mathbb{R}$, situé au-dessus de sa tangente à l'origine, qui est la droite d'équation $y = 1 + x$, puisque $\exp(0) = (\exp)'(0) = 1$.

- $\boxed{\forall x \in \,]-1, +\infty[\quad \ln(1+x) \leq x}$.

En effet, la fonction $f : x \mapsto \ln(1+x)$ est concave sur son intervalle de définition $]-1, +\infty[$ puisque $f''(x) = -\dfrac{1}{(1+x)^2} \leq 0$. Le graphe de la fonction f est donc, partout sur $]-1, +\infty[$, situé en-dessous de sa tangente à l'origine, qui est la première bissectrice d'équation $y = x$.

- $\boxed{\forall x \in \left[0, \frac{\pi}{2}\right] \quad \frac{2}{\pi}x \leq \sin(x) \leq x}$. (celle-ci n'est en fait pas au programme)

En effet, la fonction sinus est concave sur $\left[0, \frac{\pi}{2}\right]$ puisque sa dérivée seconde $(\sin)'' = -\sin$ *est négative. Donc, sur ce segment, la courbe représentative est située au-dessus de sa sécante qui est la droite d'équation $y = \dfrac{2}{\pi}\, x$, et en-dessous de sa tangente à l'origine qui est la première bissectrice d'équation $y = x$.*

Concernant la fonction sinus, l'inégalité qui est mentionnée comme exigible par le programme de PCSI est la suivante:

- $\boxed{\forall x \in \mathbb{R} \quad \big|\sin(x)\big| \leq |x|}$.

Le plus pratique pour la démontrer est de faire l'étude des variations des fonctions $g : x \mapsto x - \sin(x)$ et $h : x \mapsto x + \sin(x)$, dont il est facile de déduire que ces deux fonctions sont positives sur $\mathbb{R}_+$. On a donc

$$\forall x \in \mathbb{R}_+ \qquad -x \leq \sin(x) \leq x \ , \quad \text{soit} \quad \big|\sin(x)\big| \leq x = |x| \ .$$

Les fonctions $x \mapsto |x|$ et $x \mapsto \big|\sin(x)\big|$ étant paires, l'inégalité est aussi vérifiée sur $\mathbb{R}_-$.

UN THÉORÈME D'ALGÈBRE (programme de PCSI)

<u>Forme géométrique du théorème du rang.</u>

<u>Théorème</u>. Soient E et F deux $\mathbb{K}$-espaces vectoriels, soit u une application linéaire de E vers F. Si S est un supplémentaire de $\operatorname{Ker} u$ dans E, alors u induit un isomorphisme de S sur $\operatorname{Im} u$.

<u>Commentaire.</u> Cela signifie que l'application $v : S \to \operatorname{Im} u$ telle que $\forall x \in S \quad v(x) = u(x)$ est un isomorphisme. Cette application linéaire est obtenue par restriction de u au sous-espace S de son espace de départ E, et "corestriction" (i.e. restriction de l'espace d'arrivée) au sous-espace $\operatorname{Im} u$ de F. On note parfois $v = u\Big|_S^{\operatorname{Im} u}$, mais cette notation et ce vocabulaire ne sont pas à connaître.

Preuve. Il s'agit de montrer que tout vecteur y de $\operatorname{Im} u$ admet un unique antécédent par u dans S. Or, si on se donne $y \in \operatorname{Im} u$, ce vecteur y admet au moins un antécédent x par u dans E, soit $\exists x \in E \quad u(x) = y$. Comme $E = S \oplus \operatorname{Ker} u$, on peut décomposer x en $x = s + k$ avec $s \in S$ et $k \in \operatorname{Ker} u$, on a alors $y = u(x) = u(s) + u(k) = u(s) = v(s)$ en utilisant la notation v introduite dans le commentaire ci-dessus, d'où l'existence d'un antécédent de y dans S. Si s' est un autre vecteur de S tel que $u(s') = y$, alors $u(s') = u(s)$, donc $u(s - s') = 0_F$ et $s - s' \in S \cap \operatorname{Ker} u = \{0_E\}$, donc $s' = s$, ce qui montre l'unicité.

De cela, on déduit le **théorème du rang** habituel:

<u>Théorème</u>. Soient E et F deux $\mathbb{K}$-espaces vectoriels avec E de dimension finie, soit u une application linéaire de E vers F. On a alors $\dim E = \operatorname{rg}(u) + \dim(\operatorname{Ker} u)$.

Preuve. Comme E est de dimension finie, le sous-espace $\operatorname{Ker} u$ admet un supplémentaire S dans E, tel que

$$\dim(S) = \dim(E) - \dim(\operatorname{Ker} u) \; .$$

Le théorème ci-dessus montre alors que S est isomorphe à $\operatorname{Im} u$, ce qui entraîne l'égalité des dimensions, soit $\operatorname{rg}(u) = \dim(E) - \dim(\operatorname{Ker} u)$.