Convexité : rappels de MP2I et compléments

Table des matières

I Parties convexes

E désigne un $\mathbb{R}$-espace vectoriel.

La notion de barycentre est hors-programme, mais permet de comprendre la notion de convexité plus en profondeur.

Définition 1 (Barycentre, isobarycentre)
Soit $(a_i)_{i\in I}$ *une famille finie de vecteurs de* E *et* $(\lambda_i)_{i\in I}$ *une famille finie de réels tels que* $\sum_{i\in I} \lambda_i \neq 0$.
On appelle **barycentre de la famille (a_i) affecté des coefficients (λ_i)** *le point :*

$$m = \frac{1}{\sum_{i\in I} \lambda_i} \sum_{i\in I} \lambda_i a_i \in E.$$

On parle **d'isobarycentre** *lorsque les coefficients* $(\lambda_i)_{i\in I}$ *sont tous égaux entre eux. Dans ce cas :*

$$m = \frac{1}{\# I} \sum_{i\in I} a_i.$$

Vocabulaire
On dit aussi que m est le **barycentre du système pondéré $(a_i, \lambda_i)_{i\in I}$**.

Remarque
Le barycentre est inchangé si :

- *on retire de la famille les vecteurs* (a_i) *affectés d'un coefficient nul ;*
- *on permute les couples* (a_i, λ_i) *;*
- *on multiplie tous les coefficients* λ_i *par un même coefficient non nul.*

On peut donc toujours se ramener au cas où $\sum_{i\in I} \lambda_i = 1$.

Propriété 2 (Associativité du barycentre)
Soit I *un ensemble fini et* I_1, I_2 *des parties de* I *telles que* $I_1 \cup I_2 = I$, $I_1 \cap I_2 = \emptyset$.
Soit une famille $(\lambda_i)_{i\in I}$ *de réels telle que* $\sum_{i\in I_1} \lambda_i \neq 0$ *et* $\sum_{i\in I_2} \lambda_i \neq 0$.
Si m_1 *est le barycentre de* $(a_i, \lambda_i)_{i\in I_1}$ *et* m_2 *est le barycentre de* $(a_i, \lambda_i)_{i\in I_2}$*, alors le barycentre* m *de* $(a_i, \lambda_i)_{i\in I}$ *est aussi le barycentre de* $\left((m_1, \sum_{i\in I_1} \lambda_i), (m_2, \sum_{i\in I_2} \lambda_i)\right)$.

Preuve
Posons $\mu_k = \sum_{i\in I_k} \lambda_i \neq 0$ pour $k \in \{1,2\}$. Par hypothèse :

$$m_1 = \frac{1}{\mu_1} \sum_{i\in I_1} \lambda_i a_i, \qquad m_2 = \frac{1}{\mu_2} \sum_{i\in I_2} \lambda_i a_i,$$

donc

$$m = \frac{1}{\sum_{i\in I} \lambda_i} \sum_{i\in I} \lambda_i a_i = \frac{1}{\sum_{i\in I_1} \lambda_i + \sum_{i\in I_2} \lambda_i} \left(\sum_{i\in I_1} \lambda_i a_i + \sum_{i\in I_2} \lambda_i a_i \right) = \frac{1}{\mu_1 + \mu_2} \left(\mu_1 m_1 + \mu_2 m_2 \right).$$

Définition 3 (Segment)
Etant donnés deux points a et b de E, on appelle **segment** $[\boldsymbol{a}, \boldsymbol{b}]$ *l'ensemble des barycentres de (a, b) affectés de coefficients positifs :*

$$[a,b] = \left\{ \frac{1}{\lambda_1 + \lambda_2} (\lambda_1 a + \lambda_2 b)\,,\ \lambda_1 \geq 0,\ \lambda_2 \geq 0,\ \lambda_1 + \lambda_2 \neq 0 \right\}.$$

En se ramenant à une somme de coefficients égale à 1, cela se réécrit :

$$[a,b] = \{(1-\lambda)a + \lambda b,\ 0 \leq \lambda \leq 1\}.$$

Remarque
- *Si E est un espace vectoriel normé, alors le segment $[a,b]$ définit un chemin dans E de a vers b, puisque l'application $\gamma : [0,1] \to E$ définie par $\gamma(t) = (1-t)a + tb$ est continue (car lipschizienne, vu que $\|\gamma(t) - \gamma(u)\| = \|b-a\| \times |t-u|$) et vérifie $\gamma(0) = a$, $\gamma(1) = b$.*
- *Le point $\gamma(1/2) = \frac{a+b}{2}$ est l'isobarycentre de (a,b), appelé* **milieu** *du segment $[a,b]$.*

Dessin (Illustration de la paramétrisation du segment)

Définition 4 (Partie étoilée)
Une partie $A \subset E$ est dite **étoilée** *s'il existe un point $a \in A$ tel que pour tout $b \in A$, $[a,b] \subset A$.*

Dessin (Partie étoilée)

Définition 5 (Partie convexe)
Une partie $A \subset E$ est dite **convexe** *si pour tout $(a,b) \in A^2$, $[a,b] \subset A$.*

Remarque
A est convexe $\iff \forall (a,b) \in A^2,\ \forall \lambda \in [0,1],\ (1-\lambda)a + \lambda b \in A.$

Dessin (Partie convexe)

Exemple
- $\emptyset$ et E sont convexes.
- Tout segment $[a,b] \subset E$ est convexe : en effet, si x, y sont dans $[a,b]$, alors il existe $\lambda_1, \lambda_2 \in [0,1]$ tels que $x = (1-\lambda_1)a + \lambda_1 b$ et $y = (1-\lambda_2)a + \lambda_2 b$, donc pour tout $\lambda \in [0,1]$:

$$(1-\lambda)x + \lambda y = (1-\lambda)\left((1-\lambda_1)a + \lambda_1 b\right) + \lambda\left((1-\lambda_2)a + \lambda_2 b\right)$$

$$= (1 - \lambda_1 - \lambda(\lambda_2 - \lambda_1))a + (\lambda_1 + \lambda(\lambda_2 - \lambda_1))b = (1-t)a + tb,$$

  avec $t = \lambda_1 + \lambda(\lambda_2 - \lambda_1) \in [0,1]$ car :
  * si $\lambda_1 \leq \lambda_2$, $0 \leq \lambda_1 \leq t \leq \lambda_1 + (\lambda_2 - \lambda_1) = \lambda_2 \leq 1$;
  * si $\lambda_1 > \lambda_2$, $0 \leq \lambda_2 \leq t \leq \lambda_1 \leq 1$.

  On a donc montré que $\forall (x,y) \in [a,b]^2,\ (1-\lambda)x + \lambda y \in [a,b]$, donc $[a,b]$ est convexe.
- Tout sev F de E est convexe, car une combinaison convexe de deux vecteurs ($(1-\lambda)a + \lambda b$ avec $\lambda \in [0,1]$) est un cas particulier de combinaison linéaire, et tout sev est stable par combinaison linéaire.
- Tout sous-espace affine X de E est convexe : si $X = \emptyset$, c'est évident.
  Si X est non vide, alors si on fixe $x_0 \in X$, on a $X = x_0 + F$, où F est un sev de E. Dans ce cas, pour tout $(x,y) \in X^2$, on a $x = x_0 + a$ et $y = x_0 + b$ avec $(a,b) \in F^2$, donc pour tout $\lambda \in [0,1]$:

$$(1-\lambda)x + \lambda y = x_0 + \underbrace{(1-\lambda)a + \lambda b}_{\in F} \in X.$$

Propriété 6 (Convexité des boules dans un evn)
Soit E un espace vectoriel normé. Toute boule (ouverte ou fermée) est convexe.

Preuve
Soit $a \in E$ et $r > 0$. Montrons que la boule ouverte $B(a,r)$ est convexe.
Etant donnés x, y dans $B(a,r)$, on a $\|x-a\| < r$ et $\|y-a\| < r$, donc pour tout $\lambda \in [0,1]$:

$$\|(1-\lambda)x + \lambda y - a\| = \|(1-\lambda)(x-a) + \lambda(y-a)\| \leq (1-\lambda)\|x-a\| + \lambda\|y-a\| < (1-\lambda)r + \lambda r = r,$$

ce qui montre bien que $(1-\lambda)x + \lambda y \in B(a,r)$. On procède de même pour la boule fermée $\overline{B(a,r)}$.

Propriété 7 (Lien entre convexité et connexité par arcs)
Soit E un espace vectoriel normé et $A \subset E$ une partie non vide. Alors :
A convexe $\Longrightarrow$ A étoilée $\Longrightarrow$ A connexe par arcs.

Preuve
Si A est convexe, fixons un point $a \in A$ (possible car $A \neq \emptyset$). Alors pour tout $b \in A$, on a $[a,b] \subset A$ par convexité, donc A est étoilée.
Si A est étoilée, alors pour tous points $(x,y) \in A^2$, on a $[a,x] \subset A$ et $[a,y] \subset A$, donc il existe un chemin dans A de a vers x, et un chemin dans A de a vers y. Ainsi, par transitivité, il existe un chemin dans A de x vers y, ce qui montre que A est connexe par arcs.

Propriété 8 (Caractérisation des parties convexes)
Soit $A \subset E$. On a équivalence entre :

(i) A est convexe ;

(ii) tout barycentre de points de A à coefficients positifs appartient à A.

Preuve $\boxed{(ii) \Longrightarrow (i)}$ Si a, b sont dans A, alors tout point du segment $[a,b]$ est un barycentre de (a,b) à coefficients positifs, donc appartient à A d'après (ii). Donc $[a,b] \subset A$, ce qui montre que A est convexe.

$\boxed{(i) \Longrightarrow (ii)}$ Supposons A convexe et montrons la propriété (ii) par récurrence sur le nombre de points $(a_1, \cdots, a_n) \in A^n$ considérés.
Si $n = 1$, il n'y a rien à montrer.
Si $n = 2$, c'est direct car tout barycentre de (a_1, a_2) à coefficients positifs est un point du segment $[a_1, a_2]$, donc appartient à A par convexité.
Supposons la propriété (ii) vraie pour un entier $n \in \mathbb{N}^*$ et montrons-la pour $n+1$ points de A, notés $(a_1, \cdots, a_n, a_{n+1})$. Soient des coefficients positifs $\lambda_1, \cdots, \lambda_{n+1}$ tels que $\sum_{i=1}^{n+1} \lambda_i \neq 0$, et soit m le barycentre de $(a_i, \lambda_i)_{1 \leq i \leq n+1}$. On peut supposer que $\forall i \in \{1, \cdots, n+1\}$, $\lambda_i > 0$ (sinon, un des λ_i est nul et m se réécrit comme barycentre de n points, d'où $m \in A$ par hypothèse de récurrence).
Par associativité du barycentre (cf. prop. 2), m est aussi le barycentre de $((b,\mu), (a_{n+1}, \lambda_{n+1}))$, où b est le barycentre de $(a_i, \lambda_i)_{1 \leq i \leq n}$ et $\mu = \sum_{i=1}^{n} \lambda_i > 0$. Par hypothèse de récurrence, $b \in A$, et vu qu'on a aussi $a_{n+1} \in A$, on déduit par le cas $n = 2$ que m est dans A, en tant que barycentre à coefficients positifs de deux points de A.

II Fonctions convexes

I désigne un intervalle de $\mathbb{R}$.

1) Définition et propriétés

Définition 9 (Fonction convexe)
Une fonction $f : I \to \mathbb{R}$ est dite **convexe** *lorsque*

$$\forall (a,b) \in I^2,\ \forall \lambda \in [0,1], \qquad f((1-\lambda)a + \lambda b) \leq (1-\lambda)f(a) + \lambda f(b).$$

Remarque
Il suffit que l'inégalité précédente soit vérifiée pour tout $a \neq b$ dans I^2 et pour tout $\lambda \in]0,1[$ pour que f soit convexe. En effet, l'inégalité est toujours vraie si $a = b$ ou $\lambda = 0$ ou $\lambda = 1$ (c'est une égalité dans ces cas).

Dans la suite, on supposera ***I* d'intérieur non vide** (dans les autres cas, I est vide ou singleton, donc $f : I \to \mathbb{R}$ est automatiquement convexe dans ces cas triviaux).

Rappel (Graphe d'une fonction)
Le graphe de $f : I \to \mathbb{R}$ est l'ensemble

$$\Gamma_f = \{(x,y) \in \mathbb{R}^2,\ x \in I \text{ et } y = f(x)\}.$$

Propriété 10 (Caractérisation de la convexité par la position relative du graphe et des cordes)
Une fonction $f : I \to \mathbb{R}$ est convexe si et seulement si pour tout $(a,b) \in I^2$ avec $a \neq b$, le graphe de $f|_{[a,b]}$ est situé en dessous de la corde joignant les points d'abscisse a et b.

Preuve
Fixons $(a,b) \in I^2$ distincts et notons $A = (a, f(a))$ et $B = (b, f(b))$ les points du graphe de f d'abscisses respectives a et b. La corde joignant A et B est le segment $[A,B]$:

$$[A,B] = \{(1-\lambda)A + \lambda B,\ \lambda \in [0,1]\} = \{((1-\lambda)a + \lambda b\ ,\ (1-\lambda)f(a) + \lambda f(b))\ ,\ \lambda \in [0,1]\}.$$

Etant donné un point quelconque $M = (x_M, y_M)$ de la corde $[A,B]$, le fait que f soit convexe se traduit donc par

$$f(x_M) \leq y_M,$$

c'est-à-dire que M est situé au-dessus du point $N = (x_M, f(x_M))$ qui est le point du graphe de f de même abscisse.

Exemple
La fonction "valeur absolue" $|.| : \mathbb{R} \to \mathbb{R}$ est convexe.
Toute fonction affine est convexe.

Définition 11 (Fonction concave)
Une fonction $f : I \to \mathbb{R}$ est dite **concave** *lorsque*

$$\forall (a,b) \in I^2,\ \forall \lambda \in [0,1], \qquad f((1-\lambda)a + \lambda b) \geq (1-\lambda)f(a) + \lambda f(b).$$

Remarque
f est concave ssi pour tout $(a,b) \in I^2$ avec $a \neq b$, le graphe de $f|_{[a,b]}$ est situé au-dessus de la corde joignant les points d'abscisse a et b.

Propriété 12 (Lien entre convexité et concavité)
f est concave $\iff$ $-f$ est convexe.

Preuve
Multiplier une inégalité par -1 change son sens.

ATTENTION !
Il existe des fonctions ni convexes ni concaves ($f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^3$ par exemple).
Les fonctions affines sont à la fois convexes et concaves.

Définition 13 (Epigraphe d'une fonction)
Soit $f : I \to \mathbb{R}$. On appelle **épigraphe** *de f l'ensemble :*

$$Epi(f) = \{(x,y) \in \mathbb{R}^2,\ x \in I \text{ et } y \geq f(x)\}.$$

Remarque
$Epi(f)$ est la partie du plan située au-dessus du graphe de f.

Propriété 14 (Caractérisation de la convexité par l'épigraphe)
Une fonction $f : I \to \mathbb{R}$ est convexe si et seulement si son épigraphe $Epi(f)$ est convexe.

Preuve $\boxed{\Longrightarrow}$ Supposons f convexe. Soit $A = (x_A, y_A)$ et $B = (x_B, y_B)$ deux points de $Epi(f)$. On a $f(x_A) \leq y_A$ et $f(x_B) \leq y_B$. Considérons un point C du segment $[A, B]$:

$$C = (x_C, y_C) = (1-\lambda)A + \lambda B = ((1-\lambda)x_A + \lambda x_B, (1-\lambda)y_A + \lambda y_B).$$

Par convexité de f, on a

$$f(x_C) = f((1-\lambda)x_A + \lambda x_B) \leq (1-\lambda)f(x_A) + \lambda f(x_B) \leq (1-\lambda)y_A + \lambda y_B = y_C,$$

donc $C \in Epi(f)$, ce qui prouve que $Epi(f)$ est convexe.

$\boxed{\Longleftarrow}$ Supposons $Epi(f)$ convexe. Pour tous $(a,b) \in I^2$ distincts, les points $A = (a, f(a))$ et $B = (b, f(b))$ sont dans $Epi(f)$ (car sur le graphe de f), donc $[A, B] \subset Epi(f)$ par hypothèse. Cela signifie que la corde $[A, B]$ est située au-dessus du graphe de $f|_{[a,b]}$, et donc f est convexe par la prop. 10.

Propriété 15 (Inégalité de Jensen)
Soit $f : I \to \mathbb{R}$ une fonction convexe et $n \in \mathbb{N}^*$. Alors, pour toutes familles $(a_1, \cdots, a_n) \in I^n$ et $(\lambda_1, \cdots, \lambda_n) \in (\mathbb{R}^+)^n$ tels que $\sum\limits_{i=1}^{n} \lambda_i = 1$, on a

$$f\left(\sum_{i=1}^{n} \lambda_i a_i\right) \leq \sum_{i=1}^{n} \lambda_i f(a_i).$$

Preuve (Preuve "classique")
Par récurrence sur n.
Pour $n = 1$, il n'y a rien à faire (on a égalité).
Pour $n = 2$, c'est directement la définition de la convexité de f puisque $\lambda_1 = 1 - \lambda_2$, donc

$$f(\lambda_1 a_1 + \lambda_2 a_2) = f((1-\lambda_2)a_1 + \lambda_2 a_2) \leq (1-\lambda_2)f(a_1) + \lambda_2 f(a_2) = \lambda_1 f(a_1) + \lambda_2 f(a_2).$$

Soit $n \geq 2$. On suppose l'inégalité vraie pour n points. Montrons-la pour $n+1$ points : soit $(a_1, \cdots, a_{n+1}) \in I^{n+1}$ et $(\lambda_1, \cdots, \lambda_{n+1}) \in (\mathbb{R}^+)^{n+1}$ tels que $\sum\limits_{i=1}^{n+1} \lambda_i = 1$.
Si un des λ_i est nul, alors l'inégalité est vraie car on se ramène au cas de n points.
Si $\lambda_i > 0$ pour tout $i \in \{1, \cdots, n+1\}$, alors en posant $\lambda = \sum\limits_{i=1}^{n} \lambda_i > 0$, on a $\lambda_{n+1} = 1 - \lambda$, donc par convexité de f :

$$f\left(\sum_{i=1}^{n+1} \lambda_i a_i\right) = f\left(\lambda\left(\sum_{i=1}^{n} \frac{\lambda_i}{\lambda} a_i\right) + (1-\lambda)a_{n+1}\right) \leq \lambda f\left(\sum_{i=1}^{n} \frac{\lambda_i}{\lambda} a_i\right) + (1-\lambda)f(a_{n+1}).$$

Or, par hypothèse de récurrence appliquée à la famille $(a_1, \cdots, a_n)$ avec les coefficients $(\frac{\lambda_i}{\lambda})_{1 \leq i \leq n}$ (qui sont bien positifs et de somme 1), on a

$$f\left(\sum_{i=1}^{n} \frac{\lambda_i}{\lambda} a_i\right) \leq \sum_{i=1}^{n} \frac{\lambda_i}{\lambda} f(a_i) = \frac{1}{\lambda}\sum_{i=1}^{n} \lambda_i f(a_i),$$

donc finalement
$$f\left(\sum_{i=1}^{n+1}\lambda_i a_i\right) \leq \sum_{i=1}^{n}\lambda_i f(a_i) + \lambda_{n+1} f(a_{n+1}) = \sum_{i=1}^{n+1}\lambda_i f(a_i),$$
ce qui montre l'hérédité.

Preuve (Preuve géométrique)
Pour tout $i \in \{1, \cdots, n\}$, considérons le point $A_i = (a_i, f(a_i)) \in \mathbb{R}^2$. Chaque point A_i appartient au graphe de f, donc à son épigraphe. La fonction f étant convexe, l'épigraphe de f est convexe, donc contient aussi le barycentre du système $(A_i, \lambda_i)_{1\leq i\leq n}$, c'est-à-dire le point :
$$M = \left(\sum_{i=1}^{n}\lambda_i a_i, \sum_{i=1}^{n}\lambda_i f(a_i)\right) = (x_M, y_M).$$
Et le fait que $M \in Epi(f)$ se traduit par $f(x_M) \leq y_M$, c'est-à-dire l'inégalité de Jensen.

Définition 16 (Fonction pente en un point x_0)
Soit $f : I \to \mathbb{R}$ et $x_0 \in I$. On appelle **fonction pente en x_0** *(ou* **taux d'accroissement en x_0***) la fonction $p_{x_0} : I \setminus \{x_0\} \to \mathbb{R}$ définie par*
$$\forall x \in I \setminus \{x_0\}, \qquad p_{x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0}.$$

Remarque
Pour $(x, y) \in I^2$ distincts, on a $p_x(y) = p_y(x)$.

Propriété 17 (Caractérisation de la convexité par la croissance des pentes)
$f : I \to \mathbb{R}$ est convexe si et seulement si pour tout $x_0 \in I$, la fonction $p_{x_0} : I \setminus \{x_0\} \to \mathbb{R}$ est croissante.

Preuve $\boxed{\Longrightarrow}$ Supposons f convexe et fixons trois points distincts $(x_0, x, y) \in I^3$ avec $x < y$. Montrons que $p_{x_0}(x) \leq p_{x_0}(y)$.

* si $x_0 < x < y$, alors il existe $\lambda \in]0, 1[$ tel que $x = (1 - \lambda)x_0 + \lambda y$, donc par convexité de f :
$$f(x) \leq (1 - \lambda)f(x_0) + \lambda f(y).$$
Vu que $x - x_0 > 0$, on peut majorer la pente :
$$p_{x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0} \leq \frac{\lambda(f(y) - f(x_0))}{x - x_0} = \frac{\lambda(f(y) - f(x_0))}{\lambda(y - x_0)} = p_{x_0}(y).$$
* si $x < x_0 < y$, alors d'après le premier cas (en permutant les rôles des variables) :
$$p_{x_0}(x) = p_x(x_0) \leq p_x(y) = \frac{f(y) - f(x)}{y - x}.$$
Or, il existe $\lambda \in]0, 1[$ tel que $x_0 = (1 - \lambda)x + \lambda y$, donc par convexité de f :
$$f(x_0) \leq (1 - \lambda)f(x) + \lambda f(y).$$
On en déduit (vu que $y - x_0 > 0$) la minoration :
$$p_{x_0}(y) = \frac{f(y) - f(x_0)}{y - x_0} \geq \frac{f(y) - (1 - \lambda)f(x) - \lambda f(y)}{y - x_0} = \frac{(1 - \lambda)(f(y) - f(x))}{(1 - \lambda)(y - x)} = p_x(y),$$
donc finalement, $p_{x_0}(x) \leq p_x(y) \leq p_{x_0}(y)$.
* si $x < y < x_0$, alors d'après le deuxième cas (en permutant les rôles des variables) :
$$p_y(x) \leq p_x(x_0) \leq p_y(x_0),$$
donc $p_{x_0}(x) = p_x(x_0) \leq p_y(x_0) = p_{x_0}(y)$.

$\boxed{\Longleftarrow}$ Supposons que pour tout $x_0 \in I$, la fonction $p_{x_0} : I \setminus \{x_0\} \to \mathbb{R}$ est croissante et montrons que f est convexe. Pour cela fixons $a < b$ dans I et $\lambda \in]0,1[$. En posant $c = (1-\lambda)a + \lambda b$, on a $a < c < b$, donc par hypothèse :

$$p_a(c) \leq p_a(b),$$

c'est-à-dire :

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

En multipliant par $c - a > 0$, cela revient à :

$$f(c) - f(a) \leq \frac{c-a}{b-a}(f(b)-f(a)) = \lambda(f(b)-f(a)),$$

c'est-à-dire $f(c) \leq (1-\lambda)f(a) + \lambda f(b)$, et donc f est convexe.

Corollaire 18 (Inégalité des trois pentes)
Si $f : I \to \mathbb{R}$ est convexe, alors pour tous points $x < y < z$ dans I, on a :

$$p_x(y) \leq p_x(z) \leq p_y(z).$$

Preuve
D'après la proposition précédente, p_x est croissante donc $p_x(y) \leq p_x(z) = p_z(x)$.
Mais p_z est également croissante, donc $p_z(x) \leq p_z(y) = p_y(z)$, ce qui montre la double inégalité voulue.

2) Fonctions convexes dérivables

Théorème 19 (Caractérisation des fonctions convexes dérivables)
Soit $f : I \to \mathbb{R}$ dérivable. Alors, f est convexe ssi f' est croissante.

Preuve $\boxed{\Longrightarrow}$ Supposons f convexe. Fixons $x < y$ dans I et montrons que $f'(x) \leq f'(y)$.
Pour tout point z tel que $x < z < y$, on a l'inégalité des trois pentes :

$$p_x(z) \leq p_x(y) \leq p_z(y) = p_y(z),$$

c'est-à-dire

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

Puisque f est dérivable en tout point de I, on a

$$\lim_{z \to x^+} \frac{f(z)-f(x)}{z-x} = f'(x), \qquad \lim_{z \to y^-} \frac{f(z)-f(y)}{z-y} = f'(y),$$

donc par passage à la limite lorsque $z \to x^+$ puis lorsque $z \to y^-$ dans les inégalités précédentes (séparément !), on obtient $f'(x) \leq p_x(y)$, puis $p_x(y) \leq f'(y)$, donc $f'(x) \leq f'(y)$.

$\boxed{\Longleftarrow}$ Supposons f' croissante. Etant donnés $a < b$ dans I et $\lambda \in]0,1[$, on pose $c = (1-\lambda)a + \lambda b$. On a alors $a < c < b$, donc en appliquant l'égalité des accroissements finis sur les segments $[a,c]$ et $[c,b]$ (sur lesquels f est bien dérivable), il existe $x_a \in]a,c[$ et $x_b \in]c,b[$ tels que

$$\frac{f(c)-f(a)}{c-a} = f'(x_a), \qquad \frac{f(b)-f(c)}{b-c} = f'(x_b).$$

Vu que $a < x_a < c < x_b < b$, on obtient par croissance de f' l'inégalité $f'(x_a) \leq f'(x_b)$, c'est-à-dire :

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

ou encore

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

En multipliant par $\lambda(1-\lambda)(b-a) > 0$, on en déduit :

$$(1-\lambda)(f(c)-f(a)) \leq \lambda(f(b)-f(c)),$$

c'est-à-dire

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

et donc f est bien convexe.

Corollaire 20 (Caractérisation des fonctions convexes deux fois dérivables)
Soit $f : I \to \mathbb{R}$ deux fois dérivable. Alors, f est convexe ssi $f'' \geq 0$.

Preuve
Immédiat d'après la proposition précédente car sous ces hypothèses, f' croissante équivaut à $(f')' \geq 0$ (vu que I est un intervalle).

Remarque
On a aussi :

$$f \text{ concave } \iff (-f) \text{ convexe } \iff (-f)'' = -f'' \geq 0 \iff f'' \leq 0.$$

Propriété 21 (Position relative du graphe et des tangentes)
Soit $f : I \to \mathbb{R}$ dérivable et convexe. Alors, le graphe de f est situé entièrement au-dessus de toutes ses tangentes.

Preuve
Fixons $a \in I$. Une équation cartésienne de la tangente au graphe de f en a est

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

Dans la preuve du théorème 19, nous avons montré que sous ces hypothèses, on a pour tout $x \in I$:

$$x < a \implies \frac{f(a)-f(x)}{a-x} = p_x(a) \leq f'(a) \underset{(x-a<0)}{\implies} f(x)-f(a) \geq (x-a)f'(a),$$

$$a < x \implies f'(a) \leq p_a(x) = \frac{f(x)-f(a)}{x-a} \underset{(x-a>0)}{\implies} f(x)-f(a) \geq (x-a)f'(a).$$

Donc pour tout $x \in I \setminus \{a\}$, on a l'inégalité $f(x) \geq f'(a)(x-a) + f(a)$, et elle reste vraie pour $x = a$.

Remarque
Le graphe d'une fonction concave est situé entièrement en dessous de ses tangentes.

3) Inégalités de convexité classiques

- La fonction ln est concave, donc son graphe est en-dessous de sa tangente en $x = 1$, c'est-à-dire

$$\forall x > 0, \qquad \ln(x) \leq x - 1.$$

- La fonction exp est convexe, donc son graphe est au-dessus de sa tangente en $x = 0$, c'est-à-dire

$$\forall x \in \mathbb{R}, \qquad e^x \geq 1 + x.$$

- La fonction sin est concave sur $[0, \pi/2]$, donc son graphe est au-dessus de la corde joignant les points $(0,0)$ et $(\pi/2, 1)$ et en-dessous de la tangente en $x = 0$, ce qui donne l'encadrement :

$$\forall x \in [0, \pi/2], \qquad \frac{2}{\pi}x \leq \sin(x) \leq x.$$

- Inégalité arithmético-géométrique : pour tous réels $a_1, \cdots, a_n \geq 0$:

$$\sqrt[n]{a_1 \cdots a_n} \leq \frac{1}{n}\left(a_1 + \cdots + a_n\right)$$

(utiliser l'inégalité de Jensen avec la fonction convexe $-\ln$ et les coefficients $\lambda_i = 1/n$).

- Inégalité de Young (qui résulte elle aussi de la concavité de ln)

$$\forall (p,q) \in]1,+\infty[^2, \quad \frac{1}{p} + \frac{1}{q} = 1 \implies \forall (a,b) \in]0,+\infty[^2, \quad ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

(qui sert notamment à montrer l'inégalité de Hölder).