Fonctions convexes

On rappelle que, si $x, y \in \mathbb{R}$ sont tels que $x \leqslant y$, alors le segment $[x, y]$ est :

$$[x, y] = \{t \in \mathbb{R} \mid x \leqslant t \leqslant y\}$$

Le résultat permet de paramétrer un segment.

Proposition (paramétrisation d'un segment) Soient $x, y \in \mathbb{R}$ tel que $x \leqslant y$. Alors :

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

Démonstration Soient $x, y \in \mathbb{R}$ tels que $x \leqslant y$. La fonction (affine) :

$$f : \left\{ \begin{array}{ccc} [0, 1] & \longrightarrow & \mathbb{R} \\ \lambda & \longmapsto & \lambda x + (1 - \lambda y) \end{array} \right.$$

étant continue, l'image $f([0, 1])$ du segment $[0, 1]$ par f est un segment (d'après le théorème des bornes atteintes). Comme f est décroissante sur $[0, 1]$ (le coefficient directeur valant $x - y \leqslant 0$), ce segment est :

$$f([0, 1]) = [f(1), f(0)] = [x, y]$$

Par ailleurs, par définition de l'image directe d'un ensemble par une application, on sait que :

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

Le résultat s'ensuit. ■

Dans tout ce chapitre, I désigne un intervalle de $\mathbb{R}$ d'intérieur non vide et f est une fonction définie sur I à valeurs réelles, *i.e.* $f \in \mathbb{R}^I$.

I – Généralités

1) Définitions

Définition (fonction convexe) La fonction f est dite *convexe* sur I si :

$$\forall x_1, x_2 \in I,\ \forall \lambda \in [0, 1], \qquad f(\lambda x_1 + (1 - \lambda)x_2) \leqslant \lambda f(x_1) + (1 - \lambda)f(x_2)$$

La fonction f est dite *concave* si $-f$ est convexe, *i.e.* si :

$$\forall x_1, x_2 \in I,\ \forall \lambda \in [0, 1], \qquad f(\lambda x_1 + (1 - \lambda)x_2) \geqslant \lambda f(x_1) + (1 - \lambda)f(x_2)$$

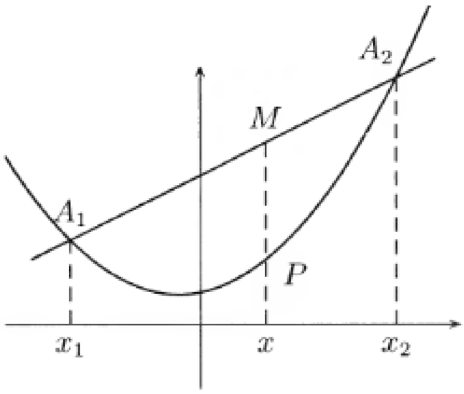
Remarque : si $x_1 = x_2$ ou si $\lambda \in \{0, 1\}$, l'inégalité est toujours vérifiée (c'est en fait une égalité).

Interprétation graphique. Notons Γ le graphe de la fonction $f : I \longrightarrow \mathbb{R}$; rappelons qu'il s'agit du sous-ensemble de $I \times \mathbb{R}$ suivant :

$$\Gamma = \{(x, f(x)) \mid x \in I\}$$

Alors :

La fonction f est convexe sur I si, et seulement si, pour tous points A_1 et A_2 de Γ d'abscisses respectives x_1 et x_2, le graphe de $f|_{[x_1,x_2]}$ est en dessous de la corde $[A_1, A_2]$.



Justification. Soient $x_1, x_2 \in I$ tels que $x_1 < x_2$ (il n'y a rien à démontrer si $x_1 = x_2$) et A_1, A_2 les points de coordonnées $(x_1, f(x_1)), (x_2, f(x_2))$ respectivement. Soit encore $x \in [x_1, x_2]$. Il existe $\lambda \in [0,1]$ tel que $x = \lambda x_1 + (1 - \lambda x_2)$. On note :

- ⋆ P le point de coordonnées $(x, f(x))$;
- ⋆ M le point de la corde $[A_1, A_2]$ d'abscisse x.

Les ordonnées de P et M sont respectivement $f(\lambda x_1 + (1-\lambda)x_2)$ et $\lambda f(x_1) + (1-\lambda)f(x_2)$. En effet, la corde a pour équation :

$$y = \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) + f(x_1)$$

donc l'ordonnée de M vaut :

$$\begin{aligned} y_M &= \frac{f(x_2) - f(x_1)}{x_2 - x_1}(\lambda x_1 + (1-\lambda)x_2 - x_1) + f(x_1) \\ &= (1-\lambda)(f(x_2) - f(x_1)) + f(x_1) \\ &= \lambda f(x_1) + (1-\lambda)f(x_2) \end{aligned}$$

Ainsi, dire que f est convexe signifie exactement l'observation géométrique évoquée.

Exemple (de fonctions convexes)

1. Toute fonction affine est à la fois convexe et concave sur $\mathbb{R}$.
   Justification. Soit $f \in \mathbb{R}^{\mathbb{R}}$ une fonction affine. Il existe alors $a, b \in \mathbb{R}$ tel que :
   $$\forall x \in \mathbb{R}, \qquad f(x) = ax + b$$
   Soient $x_1, x_2 \in \mathbb{R}$ et $\lambda \in \mathbb{R}$. Alors :
   $$\begin{aligned} f(\lambda x_1 + (1-\lambda)x_2) &= a(\lambda x_1 + (1-\lambda)x_2) + b \\ &= \lambda(ax_1 + b) + (1-\lambda)(ax_2 + b) \\ &= \lambda f(x_1) + (1-\lambda)f(x_2) \end{aligned}$$
   donc f est concave et convexe.
2. La fonction valeur absolue est convexe sur $\mathbb{R}$.
   Justification. Soient $x, y \in \mathbb{R}$ et $\lambda \in \mathbb{R}$. D'après l'inégalité triangulaire dans $\mathbb{R}$, on a :
   $$|\lambda x + (1-\lambda)y| \leqslant \lambda|x| + (1-\lambda)|y|$$
   puisque $\lambda \in [0,1]$.
3. Une somme de fonctions convexes est une fonction convexe.
   Justification. Soient f et g sont deux fonctions convexes sur I. Soit $(\lambda, x, y) \in [0,1] \times I^2$. Alors :
   $$\begin{aligned} (f+g)(\lambda x + (1-\lambda)y) &= f(\lambda x + (1-\lambda)y) + g(\lambda x + (1-\lambda)y) \\ &\leqslant \lambda f(x) + (1-\lambda)f(y) + \lambda g(x) + (1-\lambda)g(y) \quad \text{(car } f \text{ et } g \text{ sont convexes)} \\ &= \lambda(f+g)(x) + (1-\lambda)(f+g)(y) \end{aligned}$$
   On conclut donc que la fonction $f + g$ est convexe.

2) Inégalité de convexité

L'inégalité suivante généralise la définition.

Proposition (inégalité de Jensen) Soit f une fonction convexe sur I et soient p un entier naturel non nul et $\lambda_1, \ldots, \lambda_p$ des nombres réels positifs telle que $\sum_{i=1}^{p} \lambda_i = 1$. Alors :

$$\forall x_1, \ldots, x_p \in I, \qquad f\left(\sum_{i=1}^{p} \lambda_i x_i\right) \leqslant \sum_{i=1}^{p} \lambda_i f(x_i)$$

Démonstration On utilise un raisonnement par récurrence. Pour tout entier naturel p non nul, considérons la proposition $\mathscr{P}_p$: « pour tous $\lambda_1, \ldots, \lambda_p \in \mathbb{R}_+$ tels que $\sum_{i=1}^{p} \lambda_i = 1$, et pour tous $x_1, \ldots, x_p \in I$, on a :

$$\sum_{i=1}^{p} \lambda_i x_i \in I \qquad \text{et} \qquad f\left(\sum_{i=1}^{p} \lambda_i x_i\right) \leqslant \sum_{i=1}^{p} \lambda_i f(x_i) \text{ »}$$

- ⋆ La proposition $\mathscr{P}_1$ est vraie car, si $p = 1$, alors on a nécessairement $\lambda_1 = 1$.
- ⋆ Soit $p \in \mathbb{N}^*$ tel que la proposition $\mathscr{P}_p$ soit vraie. Soient $\lambda_1, \ldots, \lambda_{p+1} \in \mathbb{R}_+$ tels que $\sum_{i=1}^{p+1} \lambda_i = 1$, et soient $x_1, \ldots, x_{p+1} \in I$. On distingue deux cas.
  - **Premier cas :** $\lambda_{p+1} = 1$

    Pour tout $i \in [\![1, p]\!]$, on a nécessairement $\lambda_i = 0$. Alors $\sum_{i=1}^{p+1} \lambda_i x_i = x_{p+1} \in I$ et l'inégalité à établir est alors clairement vraie (elle est en fait une égalité).
  - **Deuxième cas :** $\lambda_{p+1} \neq 1$ *i.e.* $\lambda_{p+1} < 1$

    Alors :

    $$\sum_{i=1}^{p+1} \lambda_i x_i = \sum_{i=1}^{p} \lambda_i x_i + \lambda_{p+1} x_{p+1} = (1 - \lambda_{p+1}) \sum_{i=1}^{p} \frac{\lambda_i}{1 - \lambda_{p+1}} x_i + \lambda_{p+1} x_{p+1}$$

    La famille $\left(\frac{\lambda_i}{1 - \lambda_{p+1}}\right)_{i \in [\![1,p]\!]}$ est composée de p nombres réels positifs ou nuls et :

    $$\begin{aligned}\sum_{i=1}^{p} \frac{\lambda_i}{1 - \lambda_{p+1}} = \frac{1}{1 - \lambda_{p+1}} \sum_{i=1}^{p} \lambda_i &= \frac{1}{1 - \lambda_{p+1}} \left(\sum_{i=1}^{p+1} \lambda_i - \lambda_{p+1}\right) \\ &= \frac{1}{1 - \lambda_{p+1}} (1 - \lambda_{p+1}) \\ &= 1\end{aligned}$$

    Par hypothèse de récurrence, le nombre réel $y = \sum_{i=1}^{p} \frac{\lambda_i}{1 - \lambda_{p+1}} x_i$ appartient à I et :

    $$f(y) \leqslant \sum_{i=1}^{p} \frac{\lambda_i}{1 - \lambda_{p+1}} f(x_i)$$

Par convexité de I, on a $\sum_{i=1}^{p+1} \lambda_i x_i \in I$ puis, par convexité de f sur I,

$$\begin{aligned} f\left(\sum_{i=1}^{p+1} \lambda_i x_i\right) &= f((1-\lambda_{p+1})y + \lambda_{p+1}x_{p+1}) \\ &\leqslant (1-\lambda_{p+1})f(y) + \lambda_{p+1}f(x_{p+1}) \\ &\leqslant (1-\lambda_{p+1})\sum_{i=1}^{p} \frac{\lambda_i}{1-\lambda_{p+1}} f(x_i) + \lambda_{p+1}f(x_{p+1}) \\ &= \sum_{i=1}^{p+1} \lambda_i f(x_i), \end{aligned}$$

d'où l'inégalité annoncée. La proposition $\mathscr{P}_{p+1}$ est donc vraie.

Par principe de récurrence simple, la proposition est démontrée. ■

Remarques : soient f une fonction convexe sur I, $p \in \mathbb{N}^*$ et $x_1, \ldots, x_p \in I$.

- ⋆ si $\lambda_1, \ldots, \lambda_p$ sont des nombres réels positifs non tous nuls (*i.e.* : $\exists\, i \in [\![1,p]\!]$, $\lambda_i \neq 0$), alors :

$$f\left(\frac{\lambda_1 x_1 + \cdots + \lambda_p x_p}{\lambda_1 + \cdots + \lambda_p}\right) \leqslant \frac{\lambda_1 f(x_1) + \cdots + \lambda_p f(x_p)}{\lambda_1 + \cdots + \lambda_p}$$

- ⋆ comme $\sum_{i=1}^{p} \frac{1}{p} = 1$, on a :

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

Exemple Nous allons démontrer ultérieurement que la fonction exponentielle est convexe sur $\mathbb{R}$. On pourra donc écrire (pour $p = 3$ par exemple) :

$$\forall x, y, z \in \mathbb{R}, \qquad \mathrm{e}^{\frac{x+y+z}{3}} \leqslant \frac{\mathrm{e}^x + \mathrm{e}^y + \mathrm{e}^z}{3}$$

3) Caractérisation de la convexité par la croissance des pentes

Proposition (inégalités des pentes) Soit $f \in \mathbb{R}^I$. Les assertions suivantes sont équivalentes :

(i) f est convexe sur I ;

(ii) pour tous $x, y, z \in I$,

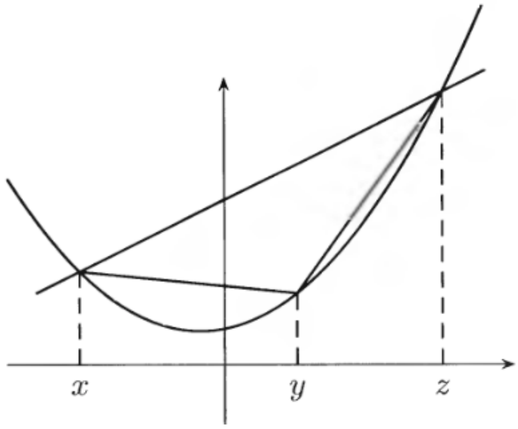
$$x < y < z \Longrightarrow \frac{f(y)-f(x)}{y-x} \leqslant \frac{f(z)-f(x)}{z-x} \qquad (*)$$

(iii) pour tous $x, y, z \in I$,

$$x < y < z \Longrightarrow \frac{f(y)-f(x)}{y-x} \leqslant \frac{f(z)-f(y)}{z-y} \qquad (**)$$

(iv) pour tous $x, y, z \in I$,

$$x < y < z \Longrightarrow \frac{f(z)-f(x)}{z-x} \leqslant \frac{f(z)-f(y)}{z-y}$$



Remarque : on retrouve facilement les inégalités à l'aide d'un dessin.

Démonstration Commençons par démontrer l'équivalence entre les propriétés (i) et (ii) en raisonnant par double implication.

$\boxed{\Longrightarrow}$ Supposons que (i) soit vérifiée. Soient $x, y, z \in I$ tels que $x < y < z$. Alors :

$$y = \lambda x + (1-\lambda)z \qquad \text{où} \qquad \lambda = \frac{z-y}{z-x} \in [0,1]$$

Comme f est convexe, on a :

$$\begin{aligned} f(y) = f(\lambda x + (1-\lambda)z) &\leqslant \lambda f(x) + (1-\lambda)f(z) \\ &= \frac{z-y}{z-x}f(x) + \frac{y-x}{z-x}f(z) \end{aligned}$$

et donc :

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

d'où l'on tire l'inégalité annoncée en divisant par $y - x > 0$.

$\boxed{\Longleftarrow}$ Réciproquement, supposons que la propriété (ii) soit vraie et montrons que f est convexe. Soient $\lambda \in [0,1]$ et $x_1, x_2 \in I$. Posons $y = \lambda x_1 + (1-\lambda)x_2$.

- ⋆ Si $x_1 = x_2$ ou si $\lambda \in \{0,1\}$, alors :

$$f(\lambda x_1 + (1-\lambda)x_2) = \lambda f(x_1) + (1-\lambda)f(x_2)$$

- ⋆ Supposons maintenant que $x_1 \neq x_2$ et que $\lambda \in]0,1[$. Sans perte de généralité, on peut supposer que $x_1 < x_2$. D'après (ii), on a :

$$\frac{f(y)-f(x_1)}{y-x_1} \leqslant \frac{f(x_2)-f(x_1)}{x_2-x_1}$$

c'est-à-dire :

$$\frac{f(\lambda x_1 + (1-\lambda)x_2) - f(x_1)}{(1-\lambda)(x_2-x_1)} \leqslant \frac{f(x_2)-f(x_1)}{x_2-x_1}$$

et comme $x_2 - x_1 > 0$, on en déduit que :

$$\begin{aligned} f(\lambda x_1 + (1-\lambda)x_2) &\leqslant (1-\lambda)(f(x_2) - f(x_1)) + f(x_1) \\ &= \lambda f(x_1) + (1-\lambda)f(x_2) \end{aligned}$$

La fonction f est donc convexe.

On remarque enfin que les inégalités (ii), (iii) et (iv) sont équivalentes à :

$$(x-y)f(z) + (y-z)f(x) + (z-x)f(y) \leqslant 0$$

Elles sont donc équivalentes entre elles. ■

Remarque : la fonction $f \in \mathbb{R}^I$ est donc convexe si, et seulement si, pour tout $a \in I$, la fonction :

$$\tau_a : \left\{ \begin{array}{ccc} I \setminus \{a\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{f(x) - f(a)}{x - a} \end{array} \right.$$

est croissante sur l'ensemble $I \setminus \{a\}$.

✎✎✎ **Exercice** Soit $f \in \mathbb{R}^{\mathbb{R}_+}$ une fonction croissante, convexe et non constante. Montrer que $f(x) \xrightarrow[x \to +\infty]{} +\infty$.

Une solution. Par hypothèse, il existe $x, y \in \mathbb{R}_+$ tels que $x < y$ et $f(x) < f(y)$. Comme f est convexe sur $\mathbb{R}_+$, on a d'après l'inégalité des pentes $(*)$:

$$\forall t \in]y, +\infty[, \qquad \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(t) - f(x)}{t - x}$$

et donc :

$$\forall t \in]y, +\infty[, \qquad f(t) \geqslant (t - x)\frac{f(y) - f(x)}{y - x} + f(x)$$

Or :

$$\frac{f(y) - f(x)}{y - x} > 0 \qquad \text{donc} \qquad (t - x)\frac{f(y) - f(x)}{y - x} + f(x) \xrightarrow[x \to +\infty]{} +\infty$$

Le théorème de comparaison permet de conclure quant à la limite de f en $+\infty$.

✎✎✎ **Exercice** Soient p et q deux nombres réels et f une fonction définie sur un intervalle I. On considère la fonction :

$$g : \left\{ \begin{array}{ccc} I & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) - px - q \end{array} \right.$$

Montrer que la fonction g est convexe si et seulement si la fonction f est convexe.

Une solution.

- ⋆ Pour tout $a \in I$, les fonctions $\tau_a(f)$ et $\tau_a(g)$ diffèrent de la constante p. La croissance de $\tau_a(f)$ équivaut donc à celle de $\tau_a(g)$.
- ⋆ Si f est convexe, alors comme toute fonction affine est convexe, g est aussi convexe. On traite la réciproque de manière analogue.

II – Convexité et dérivabilité

On peut facilement étudier la convexité/concavité d'une fonction lorsque celle-ci est dérivable voire deux fois dérivable. On rappelle que :

- ⋆ $\mathscr{D}(I, \mathbb{R})$ désigne l'ensemble des fonctions dérivables sur I à valeurs dans $\mathbb{R}$;
- ⋆ $\mathscr{D}^2(I, \mathbb{R})$ est l'ensemble des fonctions deux fois dérivables sur I à valeurs dans $\mathbb{R}$.

1) Caractérisation des fonctions dérivables convexes

Proposition (convexité et dérivabilité) Soit $f \in \mathbb{R}^I$.

- ⋆ Si $f \in \mathscr{D}(I, \mathbb{R})$, alors f est convexe sur I si et seulement si f' est croissante sur I.
- ⋆ Si $f \in \mathscr{D}^2(I, \mathbb{R})$, alors f est convexe sur I si et seulement si $f'' \geqslant 0$ sur I (*i.e.* : $\forall x \in I,\ f''(x) \geqslant 0$).

Démonstration Le deuxième point est une conséquence du premier, et du lien entre croissance d'une fonction dérivable et signe de la dérivée. On démontre maintenant le premier point en raisonnant par double implication.

⋆ Supposons que la fonction f soit convexe sur I et soient $x, y \in I$ tels que $x < y$. Pour tout $t \in]x, y[$, on a d'après les inégalités des pentes :

$$\forall t \in]x, y[, \qquad \frac{f(t)-f(x)}{t-x} \leqslant \frac{f(y)-f(x)}{y-x} \leqslant \frac{f(y)-f(t)}{y-t} = \frac{f(t)-f(y)}{t-y}$$

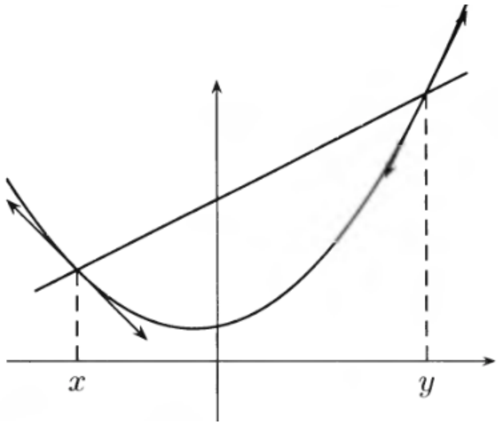
La fonction f étant dérivable sur I, elle l'est en particulier à droite au point x. On obtient donc, en faisant tendre t vers x^+ dans l'inégalité de gauche ci-dessus, on obtient :

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

De même, f est dérivable à gauche en y donc :

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

Par transitivité de la relation $\leqslant$, on obtient $f'(x) \leqslant f'(y)$. La fonction f' est donc croissante sur I.



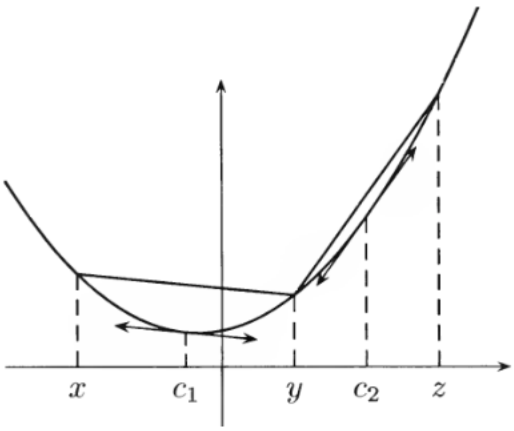
⋆ Supposons maintenant que f' soit croissante sur I et soient $x, y, z \in I$ tels que $x < y < z$. D'après le théorème des accroissements finis, il existe $c_1 \in]x, y[$ et $c_2 \in]y, z[$ tels que :

$$\frac{f(y)-f(x)}{y-x} = f'(c_1) \qquad \text{et} \qquad \frac{f(z)-f(y)}{z-y} = f'(c_2)$$

Comme $x < c_1 < y < c_2 < z$, la croissance de f' nous donne $f'(c_1) \leqslant f'(c_2)$, c'est-à-dire :

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

D'après $(**)$, la fonction f est convexe.



■

Exemple ⋆ La fonction exponentielle est convexe sur $\mathbb{R}$, la fonction logarithme est concave sur $\mathbb{R}_+^*$.

⋆ La fonction $f : x \longmapsto \ln(1 + e^x)$ est convexe sur $\mathbb{R}$. En effet, $f \in \mathscr{D}^2(\mathbb{R}, \mathbb{R})$ et pour tout $x \in \mathbb{R}$, on a

$$f'(x) = \frac{e^x}{1+e^x} \qquad \text{puis} \qquad f''(x) = \frac{e^x}{(1+e^x)^2} \geqslant 0$$

2) Position par rapport à la tangente

Proposition (convexité et tangente) Soit f une fonction convexe et dérivable sur l'intervalle I. Alors :

$$\forall a, x \in I, \qquad f(x) \geqslant f(a) + (x-a)f'(a)$$

Démonstration Soit $a, x \in I$. L'inégalité est immédiate si $a = x$. Supposons maintenant que $a \neq x$. On distingue deux cas.

- ⋆ **Premier cas :** $a < x$
  Comme $f \in \mathscr{D}(I, \mathbb{R})$, on a en particulier $f \in \mathscr{C}([a,x], \mathbb{R}) \cap \mathscr{D}(]a,x[, \mathbb{R})$. D'après le théorème des accroissements finis, il existe $c \in]a,x[$ tel que :
  $$\frac{f(x) - f(a)}{x-a} = f'(c)$$
  Or $c \geqslant a$ et f' est croissante sur I (car f est convexe et dérivable sur I) donc $f'(c) \geqslant f'(a)$. Par conséquent :
  $$\frac{f(x) - f(a)}{x-a} \geqslant f'(a)$$
  En multipliant par $x - a > 0$ puis en ajoutant $f(a)$, on obtient bien $f(x) \geqslant f(a) + (x-a)f'(a)$.
- ⋆ **Deuxième cas :** $x < a$
  De la même manière, il existe $c \in]x,a[$ tel que $\dfrac{f(a) - f(x)}{a-x} = f'(c)$ et on a ici $f'(c) \leqslant f'(a)$ (car $c \leqslant a$).
  On obtient l'inégalité souhaitée en multipliant par $a - x < 0$ puis en isolant $f(x)$.

L'inégalité est donc vérifiée dans tous les cas, ce qui achève la démonstration. ■

Remarque : le résultat de la proposition précédente peut s'exprimer en disant que le graphe d'une fonction convexe dérivable est situé au-dessus de chacune de ses tangentes.

3) Inégalités de convexité

Proposition
- ⋆ Pour tout $x \in \mathbb{R}$, on a $\mathrm{e}^x \geqslant 1 + x$.
- ⋆ Pour tout $x \in]-1, +\infty[$, on a $\ln(1+x) \leqslant x$.

Démonstration La fonction exponentielle est dérivable et convexe sur $\mathbb{R}$. D'après la proposition précédente, la courbe $\mathscr{C}_{\exp}$ se situe au-dessus de la tangente à la courbe au point d'abscisse 0. Autrement dit :

$$\forall x \in \mathbb{R}, \qquad \mathrm{e}^x \geqslant \exp'(0)(x-0) + \mathrm{e}^0,$$

d'où l'inégalité souhaitée. Pour obtenir la deuxième inégalité, la fonction $f : x \longmapsto \ln(1+x)$ étant concave sur son domaine de définition (car $f'' \leqslant 0$), son graphe se situe en-dessous de sa tangente au point d'abscisse 0. Il suffit alors d'écrire l'inégalité correspondante. ■