# LEÇON N° 74 :

# Fonctions convexes d'une variable réelle. Applications.

**Pré-requis** :
- *Convexité d'une partie de* $\mathbb{R}^2$ *;*
- *Notions de continuité, dérivabilité (« normale », à gauche et à droite).*

Soit $I = [a, b]$ un intervalle de $\mathbb{R}$. On note $I^\circ$ l'intervalle $I$ privé de ses bornes. On considère dans cette leçon une fonction $f : I \longrightarrow \mathbb{R}$.

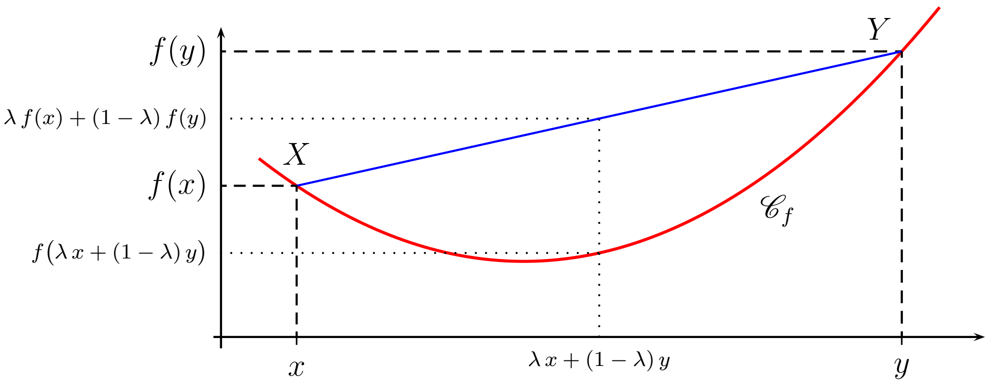
## 74.1 Fonctions convexes

**Définition 1 : La fonction $f$ est dite *convexe* si**

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

**Elle est dite *concave* si $-f$ est convexe.**

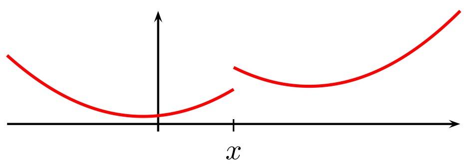
**Interprétation graphique** :




*Exemples* : Les fonctions $x \longmapsto x^2$, $x \longmapsto |x|$ sont convexes.

`Remarques 1:`

1. Soient $f$ et $g$ convexes sur $I$, et $\lambda > 0$. Alors $\lambda f$ et $f + g$ sont convexes sur $I$.
2. Si $f : \mathbb{R} \longrightarrow \mathbb{R}$, alors pour tout réel $x$, $f|_{]-\infty,x[}$ et $f|_{[x,+\infty[}$ convexes $\not\Rightarrow$ $f$ convexe sur $\mathbb{R}$.



3. $f$ est une fonction affine si et seulement si $f$ est convexe et concave.

**Théorème 1 : $f$ est convexe si et seulement si $\mathscr{A} = \{(x,y) \in \mathbb{R}^2 \mid f(x) \leqslant y\}$ est une partie convexe de $\mathbb{R}^2$.**

***démonstration*** *:*

**"$\Longrightarrow$" :** *Soient $M(x_1,y_1), N(x_2,y_2) \in \mathscr{A}$, de sorte que $f(x_1) \leqslant y_1$ et $f(x_2) \leqslant y_2$. On note ces deux inégalités $(\bigstar)$. Soit $T(x,y) \in [MN]$, on veut montrer que $T \in \mathscr{A}$, c'est-à-dire $f(x) \leqslant y$. Puisque $T \in [MN]$, il existe $\lambda \in [0,1]$ tel que $x = \lambda\, x_1 + (1-\lambda)\, x_2$ et $y = \lambda\, y_1 + (1-\lambda)\, y_2$. Alors*

$$f(x) = f\big(\lambda\, x_1 + (1-\lambda)\, x_2\big) \overset{f \text{ convexe}}{\leqslant} \lambda\, f(x_1) + (1-\lambda)\, f(x_2) \overset{(\bigstar)}{\leqslant} \lambda\, y_1 + (1-\lambda)\, y_2 = f(y).$$

**"$\Longleftarrow$" :** *Soient $(a_1,a_2) \in I^2, M\big(a_1, f(a_1)\big), N\big(a_2, f(a_2)\big)$. $M, N \in \mathscr{A}$, donc $[MN] \subset \mathscr{A}$ (car $\mathscr{A}$ est convexe), et par suite, pour tout $\lambda \in [0,1]$, on a*

$$T\big(\lambda\, a_1 + (1-\lambda)\, a_2, \lambda\, f(a_1) + (1-\lambda)\, f(a_2)\big) \in \mathscr{A}.$$

*Cette appartenance s'écrit aussi*

$$f\big(\lambda\, a_1 + (1-\lambda)\, a_2\big) \leqslant \lambda\, f(a_1) + (1-\lambda)\, f(a_2),$$

*donc $f$ est convexe.*

*Notons que $\mathscr{A}$ est appelé* épigraphe *de $f$.* ■

**Théorème 2 : Les trois propositions suivantes sont équivalentes :**

**(i) $f$ est convexe ;**

**(ii) Pour tout $(x,y,z) \in I^3$ tel que $x < y < z$,**

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

**(iii) Pour tout $x_0 \in I$, la fonction suivante est croissante :**

$$\begin{array}{rccc} \varphi_{x_0} : & I\backslash\{x_0\} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \dfrac{f(x)-f(x_0)}{x-x_0}. \end{array}$$

***démonstration*** *:*

**(i) $\Rightarrow$ (ii) :** *$x < y < z$ et $f$ est convexe, donc il existe $\lambda \in [0,1]$ tel que*

$$(\sharp) \qquad \left\{ \begin{array}{l} y = \lambda\, x + (1-\lambda)\, z \\ f(y) \leqslant \lambda\, f(x) + (1-\lambda)\, f(z). \end{array} \right.$$

*Par suite, et puisque $1-\lambda \neq 0$, on a*

$$\frac{f(y)-f(x)}{y-x} \overset{(\sharp)}{\leqslant} \frac{\lambda\, f(x) + (1-\lambda)\, f(z) - f(x)}{\lambda\, x + (1-\lambda)\, z - x} = \frac{(1-\lambda)\big(f(z)-f(x)\big)}{(1-\lambda)(z-x)} = \frac{f(z)-f(x)}{z-x}.$$

$(\sharp)$ *donne aussi, puisque* $\lambda \neq 0$,

$$\begin{cases} x = \frac{1}{\lambda}\big(y - (1-\lambda)\,z\big) \\ f(x) \geqslant \frac{1}{\lambda}\big(f(y) - (1-\lambda)\,f(z)\big). \end{cases}$$

*En procédent comme tout à l'heure, on obtient alors*

$$\frac{f(z)-f(x)}{z-x} \leqslant \frac{f(z) - \frac{1}{\lambda}\big(f(y)-(1-\lambda)\,f(z)\big)}{z - \frac{1}{\lambda}\big(y-(1-\lambda)\,z\big)} = \frac{f(z)-f(y)}{z-y}.$$

**(ii) ⇒ (iii) :** *Soit* $(x, y, x_0) \in I^3$ *tel que* $x_0 < x < y$*. Alors par hypothèse,*

$$\varphi_{x_0}(x) = \frac{f(x)-f(x_0)}{x-x_0} \leqslant \frac{f(y)-f(x_0)}{y-x_0} = \varphi_{x_0}(y).$$

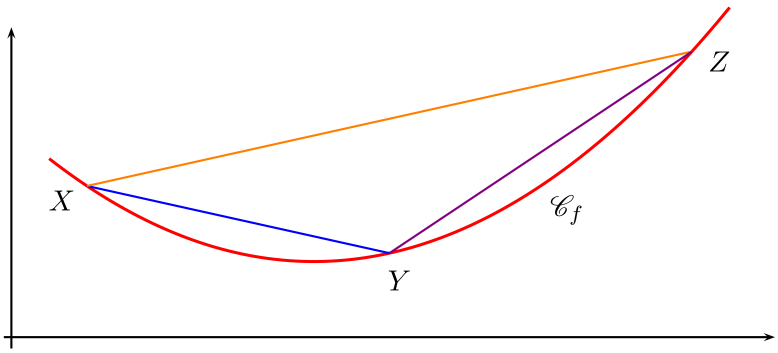
*De même si* $x < x_0 < y$ *et* $x < y < x_0$*. On en déduit que* $\varphi_{x_0}$ *est croissante sur* $I \backslash \{x_0\}$.

**(iii) ⇒ (i) :** *Soient* $(x, y) \in I^2$ *tel que* $x < y$*,* $\lambda \in \,]0, 1[$ *et* $x_0 = \lambda\,x + (1-\lambda)\,y \in I$*. D'après (iii), et en notant que* $x - x_0 < 0$*,* $y - x_0 > 0$ *et* $y - x > 0$*, on a :*

$$\begin{aligned}
\varphi_{x_0}(x) \leqslant \varphi_{x_0}(y) \quad &\Leftrightarrow \quad \frac{f(x)-f(x_0)}{x-x_0} \leqslant \frac{f(y)-f(x_0)}{y-x_0} \\
&\Rightarrow \quad (y-x_0)\big(f(x)-f(x_0)\big) \geqslant (x-x_0)\big(f(y)-f(x_0)\big) \\
&\Leftrightarrow \quad \lambda(y-x)\big(f(x)-f(x_0)\big) \geqslant (\lambda-1)(y-x)\big(f(y)-f(x_0)\big) \\
&\Leftrightarrow \quad \lambda\,f(x) + (1-\lambda)\,f(y) \geqslant (\lambda-1-\lambda)\,\big(-f(x_0)\big) \\
&\Leftrightarrow \quad \lambda\,f(x) + (1-\lambda)\,f(y) \geqslant f(x_0) = f(\lambda\,x + (1-\lambda)\,y\big).
\end{aligned}$$

■

`Remarque 2` : Si l'on note $X\big(x, f(x)\big)$, $Y\big(y, f(y)\big)$ et $Z\big(z, f(z)\big)$, alors (i) traduit le fait que la pente de la droite $(XY)$ est inférieure à celle de $(XZ)$, elle-même inférieure à celle de $(YZ)$, comme le montre clairement l'illustration ci-dessous :




## 74.2 Résultats utiles

**Théorème 3 : Si $f : I \longrightarrow \mathbb{R}$ est convexe, alors pour tout $x_0 \in I^\circ$, $f$ admet une dérivée à gauche $f'_g$ et une dérivée à droite $f'_d$, $f$ est continue sur $I^\circ$ et $f'_g$, $f'_d$ sont croissantes sur $I^\circ$.**

***démonstration*** *:*

• *Soit* $x_0 \in I^\circ$. *Pour tous* $x \in ]a, x_0[$ *et* $y \in ]x_0, b[$, *le théorème 2 nous assure que* $\varphi_{x_0}$ *est croissante et majorée sur* $]a, x_0[$ *par* $\varphi_{x_0}(y)$. *Donc la limite à gauche en* $x_0$ *de* $\varphi_{x_0}$ *existe et*

$$\lim_{x \to x_0^-} \varphi_{x_0}(x) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} = f'_g(x_0).$$

*On procède de la même manière pour la dérivée à droite, et ces deux résultats amènent directement la continuité de* $f$ *sur* $I^\circ$.

• *Soit* $x_0 \in I^\circ$. *On a alors que pour tous* $x \in ]a, x_0[$ *et* $y \in ]x_0, b[$,

$$\frac{f(x) - f(x_0)}{x - x_0} \leqslant \frac{f(y) - f(x_0)}{y - x_0}.$$

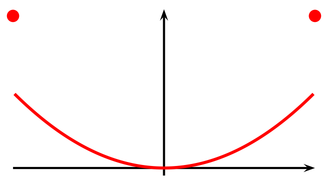
*En passant à la limite (*$x \to {x_0}^-$ *et* $y \to {x_0}^+$*), on obtient l'inégalité* $f'_g(x_0) \leqslant f'_d(x_0)$, *notée* $(1)$. *Soit alors* $x_1 \in I^\circ$ *tel que* $x_0 < x_1$. *D'après le point (ii) du théorème 2, on a pour tout* $xin$ $]x_0, x_1[$,

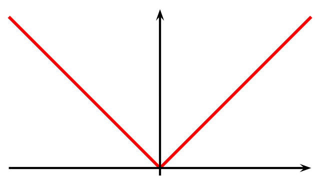
$$\begin{aligned}
& \frac{f(x) - f(x_0)}{x - x_0} \leqslant \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \frac{f(x) - f(x_1)}{x - x_1} \\
\Rightarrow \quad & \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leqslant \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \lim_{x \to x_1^-} \frac{f(x) - f(x_1)}{x - x_1} \\
\Leftrightarrow \quad & f'_d(x_0) \leqslant \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant f'_g(x_1). \qquad (2)
\end{aligned}$$

*Ainsi, les inégalités* $(1)$ *et* $(2)$ *nous donnent* $f'_g(x_0) \leqslant f'_g(x_1)$, *ce qui justifie la croissance de* $f'_g$ *sur* $I^\circ$. *On procède de la même manière pour* $f'_d$. ■

`Remarques 3 :`

1. $\boldsymbol{f}$ **convexe sur** $\boldsymbol{I} \not\Rightarrow \boldsymbol{f}$ **continue sur** $\boldsymbol{I}$**.** Considérer par exemple la fonction $f$ définie sur $[-1, 1]$ par $f(-1) = f(1) = 2$ et pour tout $x \in ]-1, 1[$, $f(x) = x^2$.
2. $\boldsymbol{f}$ **convexe sur** $\boldsymbol{I} \not\Rightarrow \boldsymbol{f}$ **dérivable sur** $\boldsymbol{I}$**.** Considérer par exemple la valeur absolue sur un intervalle ouvert contenant 0.





**Théorème 4 : Soit** $\boldsymbol{f : I \longrightarrow \mathbb{R}}$ **une fonction continue sur** $\boldsymbol{I}$ **et dérivable sur** $\boldsymbol{I^\circ}$**. Alors** $\boldsymbol{f}$ **est convexe sur** $\boldsymbol{I}$ **si et seulement si** $\boldsymbol{f'}$ **est croissante sur** $\boldsymbol{I^\circ}$**.**

***démonstration*** *: Le sens direct est assuré par le théorème 3. En effet, puisque* $f$ *est dérivable sur* $I^\circ$, *on a* $f'_g = f'_d = f'$ *sur* $I^\circ$, *ce qui donne directement la croissance de* $f'$ *sur* $I^\circ$.

*Démontrons alors le sens indirect. Soient* $x_1, x_2 \in I$ *tels que* $x_1 < x_2$ *et* $x_0 \in ]x_1, x_2[$. *Il existe donc* $\lambda \in ]0, 1[$ *tel que* $x_0 = \lambda x_1 + (1 - \lambda) x_2$. *On applique le théorème des accroissements finis sur* $]x_1, x_0[$ *et* $]x_0, x_2[$, *de sorte que*

$$\begin{aligned}
\exists\, c_1 \in ]x_1, x_0[ \quad | \quad f(x_0) - f(x_1) = f'(c_1)(x_0 - x_1) \Leftrightarrow f'(c_1) = \frac{f(x_0) - f(x_1)}{x_0 - x_1}, \\
\exists\, c_2 \in ]x_0, x_2[ \quad | \quad f(x_2) - f(x_0) = f'(c_2)(x_2 - x_0) \Leftrightarrow f'(c_2) = \frac{f(x_2) - f(x_0)}{x_2 - x_0}.
\end{aligned}$$

*Puisque $f'$ est croissante sur $I^\circ$ par hypothèse, et $c_1 < c_2$, alors*

$$\begin{aligned} f'(c_1) \leqslant f'(c_2) \quad &\Leftrightarrow \quad \frac{f(x_0)-f(x_1)}{x_0-x_1} \leqslant \frac{f(x_2)-f(x_0)}{x_2-x_0} \\ &\Rightarrow \quad \lambda(x_2-x_1)\big(f(x_0)-f(x_1)\big) \leqslant (1-\lambda)(x_2-x_1)\big(f(x_2)-f(x_0)\big) \\ &\Leftrightarrow \quad f(x_0) = f\big(\lambda\, x_1 + (1-\lambda)\, x_2\big) \leqslant \lambda\, f(x_1) + (1-\lambda)\, f(x_2), \end{aligned}$$

*donc $f$ est convexe sur $I$.* ■

**Corollaire 1 : Soit $f : I \longrightarrow \mathbb{R}$ une fonction continue sur $I$ et dérivable sur $I^\circ$. Alors $f$ est convexe sur $I$ si et seulement si la courbe représentative de $f$ est au-dessus de toutes ses tangentes.**

***démonstration :*** *Soit $x_0 \in I^\circ$. L'équation de la tangente en $x_0$ à $f$ est donnée pour tout $x \in I$ par $y = f'(x_0)(x-x_0) + f(x_0)$. Soit alors $g(x) = f(x) - f'(x_0)(x-x_0) - f(x_0)$ définie sur $I$. $g$ est dérivable sur $I^\circ$ et $g'(x) = f'(x) - f'(x_0)$. On remarque que $g(x_0) = 0$, et on note $\mathscr{C}_f$ la courbe représentative de $f$. On a alors les équivalences suivantes :*

$$\begin{aligned} & f \text{ est au-dessus de toutes ses tangentes} \\ \Leftrightarrow \quad & \begin{cases} g' < 0 \text{ sur } ]a, x_0[ \\ g' > 0 \text{ sur } ]x_0, a[ \end{cases} \Leftrightarrow \begin{cases} \forall\, x \leqslant x_0,\, f'(x) \leqslant f'(x_0) \\ \forall\, x_0 \leqslant x,\, f'(x_0) \leqslant f'(x) \end{cases} \\ \Leftrightarrow \quad & f' \text{ croissante sur } I^\circ \quad \overset{\text{thm 4}}{\Leftrightarrow} \quad f \text{ convexe sur } I. \end{aligned}$$

■

**Corollaire 2 : Soit $f : I \longrightarrow \mathbb{R}$ est continue sur $I$ et de classe $\mathscr{C}^2$ sur $I^\circ$. Alors $f$ est convexe sur $I$ si et seulement si $f''$ est positive sur $I^\circ$.**

***démonstration :*** *Le résultat est évident. En effet, le théorème 4 nous assure déjà que $f$ convexe sur $I$ équivaut à $f'$ croissante sur $I^\circ$. Puisque $f$ est $\mathscr{C}^2$ sur $I^\circ$, ceci équivaut encore à $f'' \geqslant 0$ sur $I^\circ$.* ■

# 74.3 Applications

## 74.3.1 Extremums

**Proposition 1 : Soient $f : I \longrightarrow \mathbb{R}$ une fonction dérivable et convexe sur $I$, et $a \in I$ tel que $f'(a) = 0$. Alors $f$ admet un minimum en $a$.**

***démonstration :*** *D'après le corollaire 1, $f$ est au-dessus de toutes ses tangentes. En particulier, au point $a$, on a donc*

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

*donc $f$ admet un minimum en $a$.* ■

### 74.3.2 Inégalité de convexité

#### Exemples

*Exemple 1* : $f(x) = e^x$ définie sur $\mathbb{R}$. $f$ est $\mathscr{C}^2$, et $f'' \geqslant 0$ sur $\mathbb{R}$, donc $f$ est convexe. Par le corollaire 1, appliqué en particulier au point d'abscisse nulle, on trouve alors que

$$\forall\, x \in \mathbb{R}, \qquad f(x) \geqslant f'(0)(x-0) + f(0) \Leftrightarrow e^x \geqslant x+1.$$

*Exemple 2* : $g(x) = \sin(x)$ définie sur $[0, \pi/2]$. $g$ y est $\mathscr{C}^2$ et $g'' = -\sin \leqslant 0$ sur $[0, \pi/2]$, donc $g$ est concave. On obtient ainsi la double-inégalité (obtenue de la généralité $g$ concave $\Rightarrow$ cordes $\leqslant$ g $\leqslant$ tangentes) :

$$\forall\, x \in \left[0, \frac{\pi}{2}\right], \qquad \frac{2}{\pi} x \leqslant \sin(x) \leqslant x.$$

#### Convexité avec plusieurs points

**Proposition 2 : Soient $f : I \longrightarrow \mathbb{R}$ convexe, $n \geqslant 2$, $(\lambda_1, x_1), \dots, (\lambda_n, x_n) \in \mathbb{R} \times I$ tels que $\sum_{i=1}^n \lambda_i = 1$. Alors on a l'inégalité**

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

***démonstration*** : *On procède par récurrence sur l'entier $n \geqslant 2$.*

**Initialisation :** *Pour $n = 2$, c'est exactement la définition donnée.*

**Hérédité :** *On suppose le résultat vrai jusqu'au rang $n-1$, et l'on se donne $(\lambda_1, x_1), \dots, (\lambda_n, x_n) \in \mathbb{R} \times I$ tels que $\sum_{i=1}^n \lambda_i = 1$. Alors*

$$\begin{aligned} f\left(\sum_{i=1}^n \lambda_i\, x_i\right) &= f\left(\left(1 - \sum_{i=3}^n \lambda_i\right) \underbrace{\frac{\lambda_1\, x_1 + \lambda_2\, x_2}{1 - \sum_{i=3}^n \lambda_i}}_{=:\, y} + \sum_{i=3}^n \lambda_i\, x_i\right) \\ &\overset{\text{H.R.}}{\leqslant} \left(1 - \sum_{i=3}^n \lambda_i\right) f(y) + \sum_{i=3}^n \lambda_i\, f(x_i). \end{aligned}$$

*Or on a aussi (d'après la formule vraie au rang 2) que*

$$f(y) \leqslant \frac{\lambda_1}{1 - \sum_{i=3}^n \lambda_i} f(x_1) + \frac{\lambda_2}{1 - \sum_{i=3}^n \lambda_i} f(x_2)$$

*(on aura pris soin de vérifier que la somme des coefficients vaut bien 1), et en remplaçant ce résultat dans l'inégalité obtenue par l'hypothèse de récurrence, on aboutit directement au résultat recherché.*

■

#### Comparaison de la moyenne

**Proposition 3 : Pour tous $x_1, \dots, x_n \in \mathbb{R}_+^*$, on a la double-inégalité**

$$\sqrt[n]{x_1 \cdots x_n} \leqslant \frac{x_1 + \cdots + x_n}{n} \leqslant \sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}}.$$

***démonstration :*** *La fonction* $\ln$ *est concave sur* $\mathbb{R}_+^*$*. On a donc*

$$\begin{aligned}\sum_{i=1}^{n} \frac{1}{n} \ln(x_i) \leqslant \ln\left(\sum_{i=1}^{n} \frac{1}{n} x_i\right) \quad &\Leftrightarrow \quad \ln\left(\sqrt[n]{x_1 \cdots x_n}\right) \leqslant \ln\left(\frac{x_1 + \cdots + x_n}{n}\right) \\ &\Leftrightarrow \quad \sqrt[n]{x_1 \cdots x_n} \leqslant \frac{x_1 + \cdots + x_n}{n},\end{aligned}$$

*car* $\exp$ *est une bijection strictement croissante. Il ne reste plus qu'à montrer la seconde inégalité, qui est une conséquence de la convexité de la fonction* $x \longmapsto x^2$*. En effet,*

$$\left(\sum_{i=1}^{n} \frac{1}{n} x_i\right)^2 \leqslant \sum_{i=1}^{n} \frac{1}{n}\, x_i^2 = \frac{x_1^2 + \cdots + x_n^2}{n} \quad \Leftrightarrow \quad \frac{x_1 + \cdots + x_n}{n} \leqslant \sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}},$$

*car la fonction racine carrée est aussi une bijection strictement croissante.* ■