



Chapitre 1 : Fonctions convexes

I Préliminaires

A) Notations

P désigne le plan muni d'un repère $(O, \vec{i}, \vec{j}, \vec{k})$

I est un intervalle (infini) de \mathbb{R} .

f est une fonction de I dans \mathbb{R} , et \mathcal{C} est sa courbe représentative dans le repère $(O, \vec{i}, \vec{j}, \vec{k})$.

B) Préliminaire

1. Soient a, b deux réels

Le segment d'extrémités a et b est exactement l'ensemble des réels qui s'écrivent $\alpha a + \beta b$ avec $(\alpha, \beta) \in (\mathbb{R}_+)^2$ et $\alpha + \beta = 1$, ou encore, ce qui revient bien sûr au même, l'ensemble des réels qui s'écrivent $\lambda a + (1 - \lambda)b$ avec $\lambda \in [0, 1]$.

Démonstration :

- Si $a = b$, c'est évident.
- Sinon, vu les rôles symétriques, on peut supposer $a < b$. Tout réel x s'écrit $a + t(b - a)$ (avec $t = \frac{x-a}{b-a}$). Or, pour tout $t \in \mathbb{R}$, on a les équivalences :

$$\begin{aligned} a \leq a + t(b - a) \leq b &\iff 0 \leq t(b - a) \leq b - a \\ &\iff 0 \leq t \leq 1 \end{aligned} \tag{1.1}$$

Il en résulte que l'on a bien, pour tout $x \in \mathbb{R}$:

$$\begin{aligned} x \in [a, b] &\iff \exists t \in [0, 1], x = a + t(b - a) \\ &\iff \exists t \in [0, 1], x = (1 - t)a + tb \end{aligned} \tag{1.2}$$

Géométriquement, cela signifie que le segment d'extrémités a et b est l'ensemble des barycentres de a et b affectés de coefficients positifs.

2. Soit $g: I \rightarrow \mathbb{R}$ une fonction affine sur I (c'est-à-dire du type $x \mapsto px + q$ avec $(p, q) \in \mathbb{R}^2$). Alors pour tous a et b de I , pour tous α, β de \mathbb{R} tels que $\alpha + \beta = 1$:

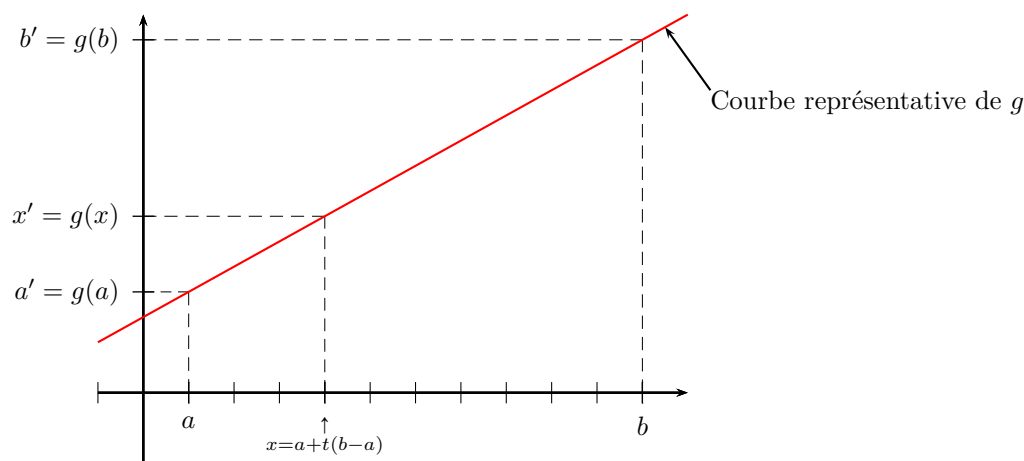
$$g(\alpha a + \beta b) = \alpha g(a) + \beta g(b) \tag{1.3}$$

Démonstration :

On a :

$$\begin{aligned} g(\alpha a + \beta b) &= p(\alpha a + \beta b) + q \\ &= p(\alpha a + \beta b) + q(\alpha + \beta) \quad (\text{car } \alpha + \beta = 1) \\ &= \alpha(pa + q) + \beta(pb + q) \\ &= \alpha g(a) + \beta g(b) \end{aligned} \tag{1.4}$$

Géométriquement, cela signifie que l'image des barycentres de deux points par une fonction affine est le barycentre des images de ces deux points, affectés des mêmes coefficients.



Sur cet exemple :

On a $x = a + t(b - a)$, avec $t = \frac{3}{10}$ (c'est-à-dire $x = \frac{7}{10}a + \frac{3}{10}b$). Alors $x' = a' + t(b' - a')$ (c'est-à-dire $x' = \frac{7}{10}a' + \frac{3}{10}b'$).

(On voit ici qu'il ne s'agit de rien d'autre que du théorème de Thalès : $\frac{x-a}{b-a} = \frac{x'-a'}{b'-a'}$).

II Fonctions convexes

A) Définition

Définition :

Soit $f: I \rightarrow \mathbb{R}$. On dit que f est convexe (sur I) lorsque :

(1) Pour tous a, b de I , pour tous α, β de \mathbb{R}_+ tels que $\alpha + \beta = 1$:

$$f(\alpha a + \beta b) \leq \alpha f(a) + \beta f(b) \quad (1.5)$$

Notons que vu les remarques suivantes :

- a et b jouent des rôles symétriques dans (1),
- L'inégalité $f(\alpha a + \beta b) \leq \alpha f(a) + \beta f(b)$ est évidemment toujours vraie si $a = b$ (et $\alpha + \beta = 1$),
- Les couples $(\alpha, \beta) \in (\mathbb{R}_+)^2$ tels que $\alpha + \beta = 1$ sont exactement les $(\lambda, 1 - \lambda)$ avec $\lambda \in [0, 1]$,

On peut déduire toute une série de variantes de la condition (1), évidemment équivalentes à (1), par exemple :

(1a) Pour tous a, b de I avec $a < b$, pour tout $(\alpha, \beta) \in (\mathbb{R}_+)^2$ tel que $\alpha + \beta = 1$:

$$f(\alpha a + \beta b) \leq \alpha f(a) + \beta f(b) \quad (1.6)$$

(1b) Pour tous a, b de I , pour tout $\lambda \in [0, 1]$:

$$f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b) \quad (1.7)$$

(1c) Pour tous a, b de I avec $a < b$, pour tout $\lambda \in [0, 1]$:

$$f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b) \quad (1.8)$$

Exemple :

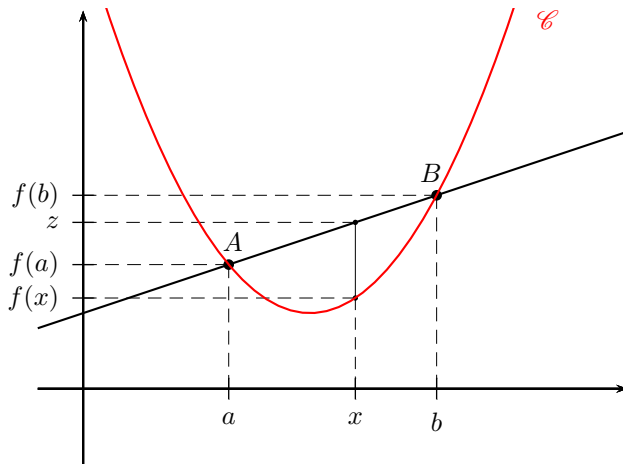
- $x \mapsto x^2$ est convexe (sur \mathbb{R}).

En effet, pour tous a, b de I et pour tout $\lambda \in [0, 1]$, on a les équivalences :

$$\begin{aligned}
 [\lambda a + (1 - \lambda)b]^2 \leq \lambda a^2 + (1 - \lambda)b^2 &\iff \lambda^2 a^2 + 2\lambda(1 - \lambda)ab + (1 - \lambda)^2 b^2 \leq \lambda a^2 + (1 - \lambda)b^2 \\
 &\iff \lambda(1 - \lambda) [-a^2 + 2ab - b^2] \leq 0 \\
 &\iff -\lambda(1 - \lambda)(a - b)^2 \leq 0 \text{ ce qui est vrai}
 \end{aligned}
 \tag{1.9}$$

- les fonctions affines sont convexes (et même mieux selon le préliminaire)

B) Interprétation graphique



On voit ici l'arc de la courbe \mathcal{C} situé entre deux points A, B de \mathcal{C} , et la corde $[A, B]$ correspondante. Ce graphique illustre la proposition :

Proposition :

Soit $f: I \rightarrow \mathbb{R}$. Alors f est convexe si et seulement si :

- (2) Tout arc de sa courbe \mathcal{C} est sous la corde correspondante.

Démonstration :

La traduction rigoureuse de la condition (2) va nous montrer que (2) \iff (1).

En effet :

Pour chaque $(a, b) \in I^2$ (avec $a < b$), notons $g_{a,b}$ la fonction affine qui coïncide avec f en a et en b , c'est-à-dire telle que : $g_{a,b}(a) = f(a)$ et $g_{a,b}(b) = f(b)$.

Alors la condition (2) équivaut à :

Pour tous a, b de I avec $a < b$, pour tout $x \in [a, b]$, $f(x) \leq g_{a,b}(x)$
 ($g_{a,b}(x)$ est le z du dessin).

Mais, d'après le préliminaire, $[a, b]$ est l'ensemble des $\lambda a + (1 - \lambda)b$ où $\lambda \in [0, 1]$. Donc (2) s'écrit aussi :

Pour tous a, b de I avec $a < b$, pour tout $\lambda \in [0, 1]$, $f(\lambda a + (1 - \lambda)b) \leq g_{a,b}(\lambda a + (1 - \lambda)b)$.

Mais, toujours d'après le préliminaire, comme $g_{a,b}$ est affine :

$$g_{a,b}(\lambda a + (1 - \lambda)b) = \lambda \underbrace{g_{a,b}(a)}_{f(a)} + (1 - \lambda) \underbrace{g_{a,b}(b)}_{f(b)}
 \tag{1.10}$$

Donc (2) est devenue :

Pour tous a, b de I avec $a < b$, pour tout $\lambda \in [0, 1]$, $f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b)$.

On reconnaît (1c).

C) Propriété de « croissance des pentes »

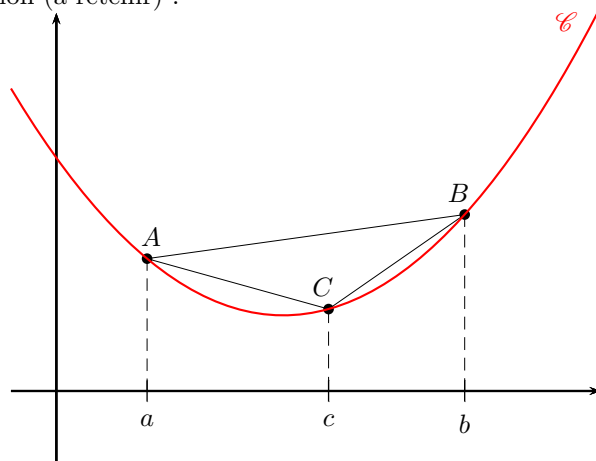
Théorème :

Soit $f: I \rightarrow \mathbb{R}$ convexe.

Alors, pour tous a, b, c de I avec $a < c < b$:

$$\frac{f(c) - f(a)}{c - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(c)}{b - c} \quad (1.11)$$

Illustration (à retenir) :



$$\frac{f(c) - f(a)}{c - a} = \text{pente de } (AC)$$

$$\frac{f(b) - f(a)}{b - a} = \text{pente de } (AB)$$

$$\frac{f(b) - f(c)}{b - c} = \text{pente de } (CB)$$

Démonstration :

En reprenant les notations de la démonstration du B) : la droite (AB) a pour équation $y = g_{a,b}(x)$.

Alors : $\frac{f(b) - f(a)}{b - a} = \frac{g_{a,b}(c) - f(a)}{c - a}$ (pente de (AB)).

Or, $f(c) \leq g_{a,b}(c)$, d'où la première inégalité.

De même, $\frac{f(b) - f(a)}{b - a} = \frac{f(b) - g_{a,b}(c)}{b - c}$ (pente de (AB)).

Or, $f(c) \leq g_{a,b}(c)$, d'où la deuxième inégalité.

Conséquence :

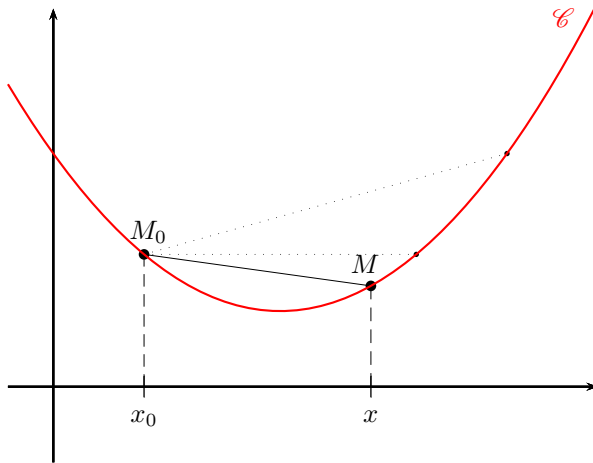
On a le théorème :

Théorème :

Soit $f: I \rightarrow \mathbb{R}$, convexe.

Alors pour tout $x_0 \in I$, l'application $x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$ est croissante sur $I \setminus \{x_0\}$.

(Autrement dit, il y a croissance des pentes des cordes dont on fixe une extrémité)



$\frac{f(x)-f(x_0)}{x-x_0}$ est la pente de la corde M_0M .

Démonstration :

Si x_1, x_2 sont deux éléments de $I \setminus \{x_0\}$ tels que $x_1 < x_2$, on a toujours :

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_2) - f(x_0)}{x_2 - x_0} \quad (1.12)$$

car :

- Si $x_0 < x_1 < x_2$, c'est la première inégalité du théorème précédent avec $(a, b, c) = (x_0, x_1, x_2)$.
- Si $x_1 < x_0 < x_2$, c'est l'inégalité obtenue par transitivité de \leq dans le théorème précédent avec $(a, b, c) = (x_1, x_0, x_2)$
- Si $x_1 < x_2 < x_0$, c'est la deuxième inégalité du théorème précédent avec $(a, b, c) = (x_1, x_2, x_0)$

III Fonctions convexes dérivables

Théorème :

Soit $f: I \rightarrow \mathbb{R}$, dérivable.

Alors f est convexe si et seulement si f' est croissante.

Corollaire :

Soit $f: I \rightarrow \mathbb{R}$, deux fois dérivable.

Alors f est convexe si et seulement si f'' est positive.

Démonstration (du théorème) :

Soit $f: I \rightarrow \mathbb{R}$, dérivable.

- Supposons f' croissante.

Soient $a, b \in I$ avec $a < b$ et soit encore $g_{a,b}$ la fonction affine qui coïncide avec f en a et en b . (ainsi, pour tout $x \in I$: $g_{a,b}(x) = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$)

Pour $x \in [a, b]$, on pose $\varphi(x) = f(x) - g_{a,b}(x)$. Alors φ est dérivable sur $[a, b]$, et, pour tout $x \in [a, b]$: $\varphi'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$.

Or, selon le théorème des accroissements finis, il existe $c \in]a, b[$ tel que $f'(c) = \frac{f(b)-f(a)}{b-a}$. Ainsi, $\forall x \in [a, b], \varphi'(x) = f'(x) - f'(c)$.

Et comme f' est supposée croissante, on en tire le tableau de variations :

x	a	c	b
$\varphi'(x)$	-	0	+
$\varphi(x)$	0		0

D'où il résulte que $\forall x \in [a, b], \varphi(x) \leq 0$.

On a donc prouvé que :

Pour tous $a, b \in I$ avec $a < b$, pour tout $x \in [a, b], f(x) \leq g_{a,b}(x)$.

Donc f est convexe (tout arc est sous la corde correspondante).

- Réciproquement, supposons f convexe. Soient alors $x_1, x_2 \in I$, avec $x_1 < x_2$.

◊ Alors, selon le théorème de croissance des pentes :

$$\forall x \in]x_1, x_2[, \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{1.13}$$

Et le passage à la limite lorsque x tend vers x_1 est possible, car f est dérivable, et donne :

$$f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{1.14}$$

◊ Le théorème de croissance des pentes donne aussi :

$$\forall x \in]x_1, x_2[, \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x} \tag{1.15}$$

Et le passage à la limite lorsque x tend vers x_2 donne :

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'(x_2) \tag{1.16}$$

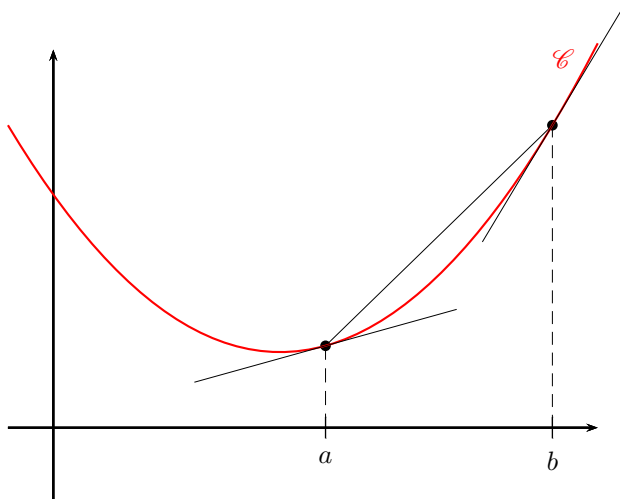
On a donc $f'(x_1) \leq f'(x_2)$, ce qui prouve la croissance de f' .

Remarque :

on a montré au passage un résultat intéressant : Si $f : I \rightarrow \mathbb{R}$ est convexe et dérivable, alors, pour tous $a, b \in I$ tels que $a < b$:

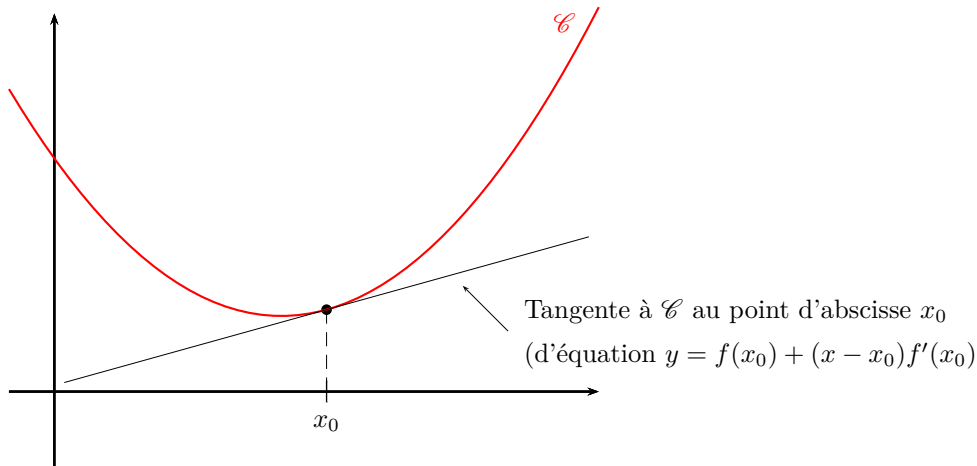
$$f'(a) \leq \frac{f(b) - f(a)}{b - a} \leq f'(b) \tag{1.17}$$

Illustration :



Théorème :

Soit $f: I \rightarrow \mathbb{R}$ convexe et dérivable. Alors la courbe \mathcal{C} de f est au dessus de ses tangentes, c'est-à-dire : $\forall x_0 \in I, \forall x \in I, f(x) \geq f(x_0) + (x - x_0)f'(x_0)$.



Démonstration :

Soit $f: I \rightarrow \mathbb{R}$ convexe et dérivable, et soit $x_0 \in I$.

Pour $x \in I$, posons $\varphi(x) = f(x) - f(x_0) - (x - x_0)f'(x_0)$.

Alors φ est dérivable sur I et :

$$\forall x \in I, \varphi'(x) = f'(x) - f'(x_0) \tag{1.18}$$

Or, f' est croissante sur I , on obtient donc le tableau de variations :

x	x_0	
$\varphi'(x)$	-	+
$\varphi(x)$	↘	↗

Ce qui prouve que $\forall x \in I, \varphi(x) \geq 0$

D'où $\forall x \in I, f(x) \geq f(x_0) + (x - x_0)f'(x_0)$.

IV Inégalité de convexité

Théorème :

Soit $f: I \rightarrow \mathbb{R}$ convexe. Alors, pour tout $n \in \mathbb{N}^*$, pour tous $a_1, a_2, \dots, a_n \in I$ et pour tous réels $\lambda_1, \lambda_2, \dots, \lambda_n$ positifs et de somme 1 :

$$f\left(\sum_{i=1}^n \lambda_i a_i\right) \leq \sum_{i=1}^n \lambda_i f(a_i) \tag{1.19}$$

Démonstration :

Par récurrence sur n ($\mathcal{P}(n)$ étant ce qui suit, dans l'énoncé du théorème, le « pour tout $n \in \mathbb{N}^*$ »)

- $\mathcal{P}(1)$ est vrai : trivial, $\forall a_1 \in I, \forall \lambda_1 \in \mathbb{R}$ tel que $\lambda_1 = 1, f(\lambda_1 a_1) \leq \lambda_1 f(a_1)$.

- $\mathcal{P}(2)$ est vrai : c'est la définition de la convexité.
- Soit $n \in \mathbb{N}^*$. Supposons $\mathcal{P}(n)$. Soient alors $a_1, a_2, \dots, a_{n+1} \in I$ et $\lambda_1, \lambda_2, \dots, \lambda_{n+1} \in \mathbb{R}_+$ tels que $\sum_{i=1}^{n+1} \lambda_i = 1$.
 - ◇ Si $\sum_{i=1}^n \lambda_i = 0$, c'est que $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$ et $\lambda_{n+1} = 1$, et on a bien l'inégalité voulue.
 - ◇ Sinon, on pose $S = \sum_{i=1}^n \lambda_i$. On peut écrire :

$$\sum_{i=1}^{n+1} \lambda_i a_i = S \frac{\sum_{i=1}^n \lambda_i a_i}{S} + \lambda_{n+1} a_{n+1} \quad (1.20)$$

et S et λ_{n+1} sont positifs, de somme 1. Donc, par définition de la convexité :

$$f\left(\sum_{i=1}^{n+1} \lambda_i a_i\right) \leq S \times f\left(\frac{\sum_{i=1}^n \lambda_i a_i}{S}\right) + \lambda_{n+1} f(a_{n+1}) \quad (1.21)$$

Or, si on pose pour chaque $i \in \llbracket 1, n \rrbracket$, $\lambda'_i = \frac{\lambda_i}{S}$, les λ'_i sont positifs et de somme 1, donc, par hypothèse de récurrence :

$$f\left(\sum_{i=1}^n \frac{\lambda_i}{S} a_i\right) \leq \sum_{i=1}^n \frac{\lambda_i}{S} f(a_i) \quad (1.22)$$

D'où finalement :

$$f\left(\sum_{i=1}^{n+1} \lambda_i a_i\right) \leq \underbrace{S \sum_{i=1}^n \frac{\lambda_i}{S} f(a_i) + \lambda_{n+1} f(a_{n+1})}_{= \sum_{i=1}^{n+1} \lambda_i f(a_i)} \quad (1.23)$$

Ce qui achève la récurrence.

Cas particulier (important, obtenu en prenant tout les λ_i égaux à $\frac{1}{n}$) :

Soit $f: I \rightarrow \mathbb{R}$ convexe. Alors pour tout $n \in \mathbb{N}^*$, pour tout $a_1, a_2, \dots, a_n \in I$, on a :

$$f\left(\frac{1}{n} \sum_{i=1}^n a_i\right) \leq \frac{1}{n} \sum_{i=1}^n f(a_i) \quad (1.24)$$

V Exemples

A) Définition

Définition :

Soit $f: I \rightarrow \mathbb{R}$. On dit que f est concave lorsque $-f$ est convexe.

Ainsi, toutes les propriétés des fonctions convexes s'appliquent immédiatement aux fonctions concaves, en retournant les inégalités.

B) Les fonctions classiques

Tout les résultats suivant se justifient en considérant la dérivée seconde, et seront dorénavant supposés connus (d'ailleurs, on les voit sur les graphes de ces fonctions, supposés connus aussi)

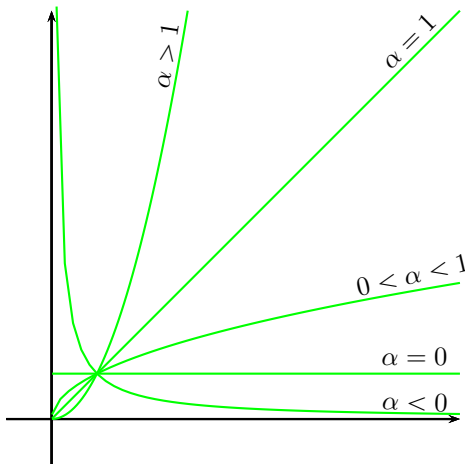
- exp est convexe sur \mathbb{R} .
- ln est concave sur \mathbb{R}_+^*

- sin est concave sur $[0, \pi]$ (et convexe sur $[-\pi, 0]$)
- cos est concave sur $[-\frac{\pi}{2}, \frac{\pi}{2}]$ (et convexe sur $[\frac{\pi}{2}, \frac{3\pi}{2}]$)
- tan est convexe sur $[0, \frac{\pi}{2}[$ (et concave sur $]-\frac{\pi}{2}, 0]$)
- fonctions puissance :

◊ Si $\alpha \in \mathbb{R}$, $x \mapsto x^\alpha$ est \mathcal{C}^∞ sur \mathbb{R}_+^* , de dérivée seconde $x \mapsto \alpha(\alpha - 1)x^{\alpha-2}$.

$$\text{Donc } x \mapsto x^\alpha \text{ est } \begin{cases} \text{convexe sur } \mathbb{R}_+^* & \text{si } \alpha < 0 \text{ ou } \alpha > 1 \\ \text{concave sur } \mathbb{R}_+^* & \text{si } 0 < \alpha < 1 \\ \text{convexe et concave (affine) sur } \mathbb{R}_+^* & \text{si } \alpha = 0 \text{ ou } 1 \end{cases}$$

Rappel des graphes :



◊ Cas particulier des exposants entiers :

— Si $n \in \mathbb{Z}$, $x \mapsto x^n$ est \mathcal{C}^∞ sur \mathbb{R}^* , de dérivée seconde $x \mapsto n(n - 1)x^{n-2}$. Donc $x \mapsto x^n$ est

$$\begin{cases} \text{convexe sur } \mathbb{R}_+^* \text{ et } \mathbb{R}_-^* \text{ si } n \text{ est pair} \\ \text{convexe sur } \mathbb{R}_+^* \text{ et concave sur } \mathbb{R}_-^* \text{ si } n \text{ est impair} \end{cases}$$

— Si $n \in \mathbb{N}$, $x \mapsto x^n$ est \mathcal{C}^∞ sur \mathbb{R} , de dérivée seconde $x \mapsto \begin{cases} n(n - 1)x^{n-2} \text{ si } n \geq 2 \\ 0 \text{ si } n = 0 \text{ ou } 1 \end{cases}$.

Donc $x \mapsto x^n$ est $\begin{cases} \text{convexe sur } \mathbb{R} \text{ si } n \text{ est pair} \\ \text{convexe sur } \mathbb{R}_+ \text{ et concave sur } \mathbb{R}_- \text{ si } n \text{ est impair} \end{cases}$

C) Exemple d'application : Comparaison des moyennes arithmétique, géométrique, harmonique, quadratique

Proposition :

Soient $a_1, a_2, \dots, a_n \in \mathbb{R}_+^*$. On pose :

$$A = \frac{a_1 + a_2 + \dots + a_n}{n} \qquad G = \sqrt[n]{a_1 a_2 \dots a_n} \qquad (1.25)$$

$$H = \left(\frac{\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_n}}{n} \right)^{-1} \qquad Q = \sqrt{\frac{a_1^2 + a_2^2 + \dots + a_n^2}{n}} \qquad (1.26)$$

Alors $H \leq G \leq A \leq Q$.

En effet : $A^2 \leq Q^2$ car $x \mapsto x^2$ est convexe sur \mathbb{R}_+^* , d'où $A \leq Q$.

$\ln G \leq \ln A$ car \ln est concave sur \mathbb{R}_+^* , d'où $G \leq A$.

$\frac{1}{G} \leq \frac{1}{H}$ car c'est l'inégalité $G \leq A$ appliquée aux $\frac{1}{a_i}$, d'où $H \leq G$.

Remarque :

Les inégalités ne sont pas faciles à démontrer de façon élémentaire...