Chapitre 3

Fonctions convexes

3.0.1 Définition de la convexité

Définition 3.1.

Soient I un intervalle de $\mathbb{R}$ et $f : I \longrightarrow \mathbb{R}$ une fonction.

i) On dit que f est convexe sur I, si pour tout $x \in I$ et pour tout $y \in I$, avec $x < y$, on a

$$\forall t \in [0,1] \ \ f((1-t)x+ty)) \leq (1-t)f(x)+tf(y)$$

ii) On dit que f est concave sur I, si pour tout $x \in I$ et pour tout $y \in I$, avec $x < y$, on a

$$\forall t \in [0,1] \ \ f((1-t)x+ty)) \geq (1-t)f(x)+tf(y)$$

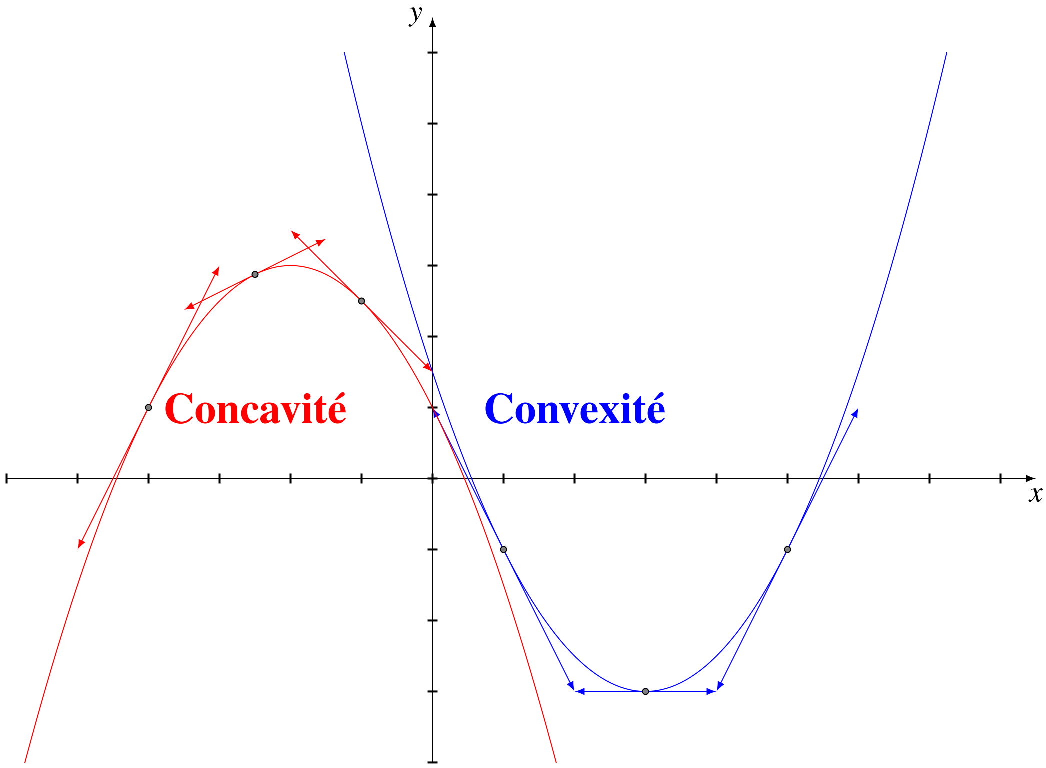
Remarques

1. f est convexe, si et seulement si, $-f$ est concave.

2. Si f est convexe sur I, alors d'après la définition, on a

$$\forall x \in I, \forall y \in I, \ f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$

3. Si f est convexe, alors pour tout $x \in I$ et pour tout $y \in I$, avec $x < y$, la courbe de f sur l'intervalle $[x,y]$ est en-dessous du segment joignant les points de coordonnées $(x, f(x))$ et $(y, f(y))$ et toutes les tangentes sont en-dessous de la courbe de f.



3.1.1 Continuité et dérivabilité des fonctions convexes

Lemme 3.2.

Soient I un intervalle de $\mathbb{R}$ et $f : I \longrightarrow \mathbb{R}$ une fonction. Alors f est convexe sur I, si et seulement si,

$$\forall x \in I, \forall y \in I, \forall z \in I,\ x < y < z \Longrightarrow \frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(y)}{z - y}$$

Preuve

$(\Longrightarrow)$ Supposons que f est convexe sur I et soient $x, y, z \in I$, tels que $x < y < z$.

Comme $y \in]x, z[$, alors $y = (1 - \lambda)x + \lambda z$, avec $\lambda = \dfrac{y - x}{z - x}$, donc $\lambda \in]0, 1[$.

Comme f est convexe, alors $f(y) \leq (1 - \lambda)f(x) + \lambda f(z)$.

Donc $(1 - \lambda)f(y) \leq (1 - \lambda)f(x) + \lambda(f(z) - f(y))$.

On a $\lambda = \dfrac{y - x}{z - x}$ et $1 - \lambda = \dfrac{z - y}{z - x}$, donc on aura

$$(z - y)f(y) \leq (z - y)f(x) + (y - x)(f(z) - f(y))$$

Par suite, on a

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

($\Longleftarrow$) Supposons que pour tout $x,y,z \in I$, avec $x<y<z$, on a

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

Puis montrons que f est convexe sur I.

Pour cela, pour $x,z \in I$, avec $x<z$, et pour $\lambda \in [0,1]$, on doit montrer que

$$f((1-\lambda)x+\lambda z) \leq (1-\lambda)f(x)+\lambda f(z)$$

Si $\lambda=0$ ou $\lambda=1$, alors il est trivial que $f((1-\lambda)x+\lambda f(z)) \leq (1-\lambda)f(x)+\lambda f(z)$.

Donc on peut supposer que $\lambda \in]0,1[$. Soit $y=(1-\lambda)x+\lambda z$, alors on a $x<y<z$, donc, par hypothèse, on a

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

avec $y-x=\lambda(z-x)$ et $z-y=(1-\lambda)(z-x)$, donc on aura

$$(1-\lambda)(f(y)-f(x)) \leq \lambda(f(z)-f(y))$$

par suite, on a $f(y) \leq (1-\lambda)f(x)+\lambda f(z)$.

Lemme 3.3.

Soit f une fonction définit sur un intervalle I de $\mathbb{R}$. Pour chaque $a \in I$, on considère la fonction φ_a définie sur $I \setminus \{a\}$ par

$$\forall x \in I \setminus \{a\},\ \varphi_a(x)=\frac{f(x)-f(a)}{x-a}$$

Alors f est convexe sur I, si et seulement si, pour tout $a \in I$, la fonction φ_a est croissante sur $I \setminus \{a\}$.

Preuve

($\Longrightarrow$) Soient $x \in I \setminus \{a\}$ et $y \in I \setminus \{a\}$, avec $x<y$.

Pour montrer que $\varphi_a(x) \leq \varphi_a(y)$, on considère trois cas :

Cas où $x<y<a$, alors on a $y \in]x,a[$, donc $y=(1-\lambda)x+\lambda a$, avec $\lambda \in]0,1[$.

Comme f est convexe, alors $f(y) \leq (1-\lambda)f(x)+\lambda f(a)$.

Donc $f(y)-f(a) \leq (1-\lambda)(f(x)-f(a))$, avec $1-\lambda=\dfrac{a-y}{a-x}$.

Ainsi, on aura $\dfrac{f(y)-f(a)}{a-y} \leq \dfrac{f(x)-f(a)}{a-x}$.

On en déduit donc que $\dfrac{f(x)-f(a)}{x-a} \leq \dfrac{f(y)-f(a)}{y-a}$, par suite $\varphi_a(x) \leq \varphi_a(y)$.

Le cas où $a<x<y$ est identique au cas précédent, il suffit de remplacer, dans la démonstration, x par a et a par x.

Cas où $x<a<y$, dans ce cas on obtient le résultat en appliquant le lemme précédent.

($\Longleftarrow$) Soient $x \in I$, $y \in I$ et $\lambda \in]0,1[$. Montrons que

$$f((1-\lambda)x+\lambda y) \leq (1-\lambda)f(x)+\lambda f(y)$$

Pour cela, on considère la fonction φ_x définie sur $I \setminus \{x\}$ par

$$\varphi_x(t) = \frac{f(t)-f(x)}{t-x}$$

alors par hypothèse, φ_x est croissante sur $I \setminus \{x\}$. Comme $(1-\lambda)x+\lambda y \in I \setminus \{x\}$ et comme $(1-\lambda)x+\lambda y \leq y$, alors on a $\varphi_x((1-\lambda)x+\lambda y)) \leq \varphi_x(y)$, ainsi on aura

$$\frac{f((1-\lambda)x+\lambda y))-f(x)}{\lambda(y-x)} \leq \frac{f(y)-f(x)}{y-x}$$

On en déduit donc que $f((1-\lambda)x+\lambda y)) \leq (1-\lambda)f(x)+\lambda f(y)$.

Théorème 3.4.

Soit f une fonction convexe sur un intervalle ouvert de $\mathbb{R}$. Alors

i) f est dérivable à droite et à gauche sur I.

ii) f est continue sur I.

Preuve

i) Soit $x_0 \in I$. Montrons que f est dérivable à droite et à gauche de x_0.

I est un intervalle ouvert, donc il existe $a, b \in I$, tel que $a < x_0 < b$. Soit g la fonction définie sur $[a,b] \setminus \{x_0\}$ par

$$g(x) = \frac{f(x)-f(x_0)}{x-x_0}$$

alors d'après le lemme précédent, g est croissante sur $[a,b] \setminus \{x_0\}$ et elle est majorée par $f(b)$ et minorée par $f(a)$, donc g possède une limite finie à droite et à gauche au point x_0, par suite f est dérivable à droite et à gauche au point x_0.

ii) f est dérivable à droite et à gauche au point x_0, donc f est continue à droite et à gauche au point x_0, par suite f est continue au point x_0.

3.4.1 Caractérisation de la convexité

Théorème 3.5.

Soit f une fonction définit et dérivable sur un intervalle I de $\mathbb{R}$. Alors f est convexe, si et seulement si, f' est croissante sur I.

Preuve

($\Longrightarrow$) Supposons que f est convexe et montrons que f' est croissante.

Pour cela, pour $a \in I$ et $b \in I$, avec $a < b$, on considère les fonctions φ_a et φ_b définies respectivement sur $I \setminus \{a\}$ et sur $I \setminus \{b\}$ par

$$\varphi_a(x) = \frac{f(x) - f(a)}{x - a} \text{ et } \varphi_b(x) = \frac{f(x) - f(b)}{x - b}$$

Alors, on a $f'(a) = \lim\limits_{x \to a} \varphi_a(x)$.

Or d'après le lemme précédent, φ_a est croissante et on a $a < b$, donc $f'(a) \leq \varphi_a(b)$.

On a aussi $f'(b) = \lim\limits_{x \to b} \varphi_b(x)$, donc $\varphi_b(a) \leq f'(b)$, car φ_b est croissante.

Comme $\varphi_a(b) = \varphi_b(a)$, alors $f'(a) \leq f'(b)$ et par suite, $'f$ est croissante.

$(\Longleftarrow)$ Supposons que f' est croissante et montrons que f est convexe.

Pour cela, d'après le lemme1, il suffit de montrer que si $a \in I$, $b \in I$ et $c \in I$, avec $a < b < c$, alors on a

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

En appliquant le théorème des accroissements finis à f sur $]a, b[$ puis sur $]b, c[$, on voit qu'il existe $\alpha \in]a, b[$ et il existe $\beta \in]b, c[$, tels que

$$f(b) - f(a) = f'(\alpha)(b - a) \text{ et } f(c) - f(b) = f'(\beta)(c - b)$$

Ainsi, on aura

$$\frac{f(b) - f(a)}{b - a} = f'(\alpha) \text{ et } \frac{f(c) - f(b)}{c - b} = f'(\beta)$$

Or f' est croissante et on a $\alpha < \beta$, donc $\dfrac{f(b) - f(a)}{b - a} \leq \dfrac{f(c) - f(b)}{c - b}$.

Corollaire 3.6.

Soit f une fonction deux fois dérivable sur un intervalle I de $\mathbb{R}$. Alors f est convexe sur I, si et seulement si, pour tout $x \in I$, on a $f''(x) \geq 0$.

Preuve

La démonstration de ce corollaire est une conséquence directe du théorème précédent, car on sait qu'une fonction dérivable sur un intervalle I est croissante, si et seulement si, sa dérivée est positive sur I.

3.6.1 Extremums d'une fonction convexe

Théorème 3.7.

Soit f une fonction convexe sur un intervalle ouvert I de $\mathbb{R}$ et soit $x_0 \in I$.

i) Si x_0 est un minimum local de f, alors x_0 est un minimum global de f sur I.

ii) Si f est dérivable en x_0 et si $f'(x_0) = 0$, alors x_0 est un minimum global de f sur I.

Preuve

i) x_0 est un minimum local de f, donc il existe $\alpha > 0$, tel que $]x_0 - \alpha, x_0 + \alpha[\subseteq I$ et tel que

$$\forall x \in I,\ x \in]x_0 - \alpha, x_0 + \alpha[\Longrightarrow f(x) \geq f(x_0)$$

Soit $x \in I$, avec $x \neq x_0$.
Nous allons choisir $\lambda \in]0,1[$, tel que $(1-\lambda)x_0 + \lambda x \in]x_0 - \alpha, x_0 + \alpha[$.
Pour cela, il suffit de choisir λ, tel que $0 < \lambda < \dfrac{\alpha}{|x - x_0|}$, alors on aura

$$\lambda \in]0,1[\text{ et } x_0 - \alpha < (1-\lambda)x_0 + \lambda x < x_0 + \alpha$$

Donc $f((1-\lambda)x_0 + \lambda x) \geq f(x_0)$ et comme f est convexe, alors on aura
$(1-\lambda)f(x_0) + \lambda f(x) \geq f(x_0)$, par suite, on a $\lambda(f(x) - f(x_0)) \geq 0$.
Donc pour tout $x \in I$, on a $f(x) \geq f(x_0)$.

ii) On considère la fonction φ_{x_0} définie sur $I \setminus \{x_0\}$ par $\varphi_{x_0}(x) = \dfrac{f(x) - f(x_0)}{x - x_0}$. Comme f est convexe sur I, alors on sait que φ_{x_0} est croissante et comme $\lim\limits_{x \to x_0} \varphi_{x_0}(x) = f'(x_0) = 0$, alors pour $x < x_0$, on a $\varphi_{x_0}(x) \leq 0$ et pour $x > x_0$, on a $\varphi_{x_0}(x) \geq 0$.
Donc, dans les deux cas, on voit que $f(x) \geq f(x_0)$, ainsi on a établi que pour tout $x \in I$, on a $f(x) \geq f(x_0)$, donc x_0 est un minimum global de f sur I.

Théorème 3.8.

Soit f une fonction convexe sur un intervalle ouvert I de $\mathbb{R}$.
Si f possède un maximum global x_0 sur I, alors f est constante sur I.

Preuve

Supposons, par absurde, que f n'est pas constante sur I.
Comme x_0 est un maximum global de f sur I, alors il existe $x \in I$, tel que $f(x) < f(x_0)$.
On a donc $x \neq x_0$, par suite, on a ou bien $x < x_0$ ou bien $x > x_0$.
On suppose, par exemple, que $x < x_0$.
Comme I est un intervalle ouvert et comme $x_0 \in I$, alors il existe $y \in I$, tel que $y \in I$, tel que $y > x_0$, donc $x_0 \in]x, y[$, donc il existe $\lambda \in]0,1[$, tel que $x_0 = (1-\lambda)x + \lambda y$.

Comme f est convexe, alors on aura

$$f(x_0) \leq (1-\lambda)f(x) + \lambda f(y) < (1-\lambda)f(x_0) + \lambda f(x_0) = f(x_0)$$

Ce qui est absurde.

Corollaire 3.9.

Soit $f : [a,b] \longrightarrow \mathbb{R}$ une fonction convexe et continue sur $[a,b]$. Alors f atteint son maximum en a ou en b.

Preuve
La démonstration est une conséquence du théorème du maximum et du théorème précédent.

3.9.1 Quelques inégalités de convexité

3.9.1.1 Inégalité de la tangente

Théorème 3.10.

Soit f une fonction convexe et dérivable sur un intervalle I. Alors on a

$$\boxed{\forall a \in I, \forall x \in I,\ f(x) \geq f(a) + (x-a)f'(a)}$$

Preuve
Si $x = a$, alors l'inégalité est trivial.

Si $x \neq a$, on considère la fonction φ_a définie sur $I \setminus \{a\}$ par

$$\varphi_a(x) = \frac{f(x)-f(a)}{x-a}$$

Comme f est convexe, alors on sait que φ_a est croissante.
Or on a $f'(a) = \lim\limits_{x \to a} \varphi_a(x)$.

Donc si $x > a$, alors $f'(a) \leq \varphi_a(x)$, par suite, on obtient le résultat.

Ei si $x < a$, alors $f'(x) \geq \varphi_a(x)$, et comme $x - a < 0$, alors on a le résultat.

Exercice
Soit f une fonction dérivable sur un intervalle I, telle que

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

Montrer que f est convexe.

Solution

Soient $x \in I$, $y \in I$ et $\lambda \in [0,1]$, alors, par hypothèse, on a

$$f(x) \geq f((1-\lambda)x+\lambda y)) - \lambda(y-x)f'((1-\lambda)x+\lambda y))$$

On a aussi

$$f(y) \geq f((1-\lambda)x+\lambda y)) + (1-\lambda)(y-x)f'((1-\lambda)x+\lambda y))$$

On en déduit donc que

$$(1-\lambda)f(x)+\lambda f(y) \geq f((1-\lambda)x+\lambda y))$$

Donc f est convexe.

3.10.0.1 Inégalité de Jensen

Théorème 3.11.

Soient f une fonction convexe sur un intervalle I et n un enteir ≥ 2. Alors pour tout $x_1, x_2, \dots, x_n$ éléments de I et pour tout $\lambda_1, \lambda_2, \dots, \lambda_n$ des nombres réels positifs vérifiant $\sum\limits_{k=1}^{n} \lambda_k = 1$, on a

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_n f(x_n)$$

Et en particulier, on a

$$\boxed{f\left(\frac{x_1+x_2+\ldots+x_n}{n}\right) \leq \frac{f(x_1)+f(x_2)+\ldots+f(x_n)}{n}}$$

Preuve

On procède par récurrence sur n, avec $n \geq 2$.

Pour $n=2$, le résultat est vrai d'après la définition de la convexité.

Supposons que la propriété est vraire pour n et montrons qu'elle est aussi vraie pour $n+1$.

Soit $x_1, x_2, \dots, x_n, x_{n+1}$ des éléments de I et soit $\lambda_1, \lambda_2, \dots, \lambda_n, \lambda_{n+1}$ des nombres réels positifs tels que $\sum\limits_{i=1}^{n+1} \lambda_i = 1$. Montrons que

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n + \lambda_{n+1} x_{n+1}) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_n f(x_n) + \lambda_{n+1} f(x_{n+1})$$

pour cela, posons $\lambda = \sum\limits_{i=1}^{n} \lambda_i$, alors deux cas sont possibles :

Si $\lambda = 0$, alors pour tout $i \in \{1,2,\dots,n\}$, on a $\lambda_i = 0$, donc dans ce cas le résultat est trivial, car $\lambda_{n+1} = 1$, donc $f(\lambda_{n+1}x_{n+1}) \leq \lambda_{n+1} f(x_{n+1})$.

Si $\lambda \neq 0$, pour chaque $i \in \{1,2,\dots,n\}$, posons $\alpha_i = \dfrac{\lambda_i}{\lambda}$, donc on aura $\sum\limits_{i=1}^{n} \alpha_i = 1$.

Comme I est un intervalle, alors $\sum_{i=1}^{n} \alpha_i x_i \in I$. Ainsi on aura

$$\begin{aligned}
f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) &= f\left(\sum_{i=1}^{n} \lambda_i x_i + \lambda_{n+1} x_{n+1}\right) \\
&= f\left(\lambda \sum_{i=1}^{n} \alpha_i x_i + \lambda_{n+1} x_{n+1}\right) \\
&= f\left(\lambda \sum_{i=1}^{n} \alpha_i x_i + (1-\lambda) x_{n+1}\right) \quad \left(\text{car } \sum_{i=1}^{n+1} \lambda_i = 1\right) \\
&\leq \lambda f\left(\sum_{i=1}^{n} \alpha_i x_i\right) + (1-\lambda) f(x_{n+1}) \quad (\text{car } f \text{ est convexe})
\end{aligned}$$

Donc on aura

$$\begin{aligned}
f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) &\leq \lambda \sum_{i=1}^{n} \alpha_i f(x_i) + (1-\lambda) f(x_{n+1}) \quad (\text{d'après l'hypothèse de récurrence}) \\
&\leq \sum_{i=1}^{n} \lambda_i f(x_i) + (1-\lambda) f(x_{n+1}) \quad (\text{car pour tout } i \in \{1,2,\ldots,n\},\ \lambda\alpha_i = \lambda_i) \\
&\leq \sum_{i=1}^{n+1} \lambda_i f(x_i)
\end{aligned}$$

Le cas particulier s'obtient en posant $\lambda_1 = \lambda_2 = \ldots = \lambda_n = \frac{1}{n}$.

Exemple

La fonction définit sur $\mathbb{R}$ par $f(x) = x^2$ est convexe sur $\mathbb{R}$, donc pour tout entier $n \geq 2$, en appliquant l'inégalité de Jensen pour $\lambda_1 = \lambda_2 = \ldots = \lambda_n = \frac{1}{n}$, on aura

$$\forall x_1, x_2, \ldots, x_n \in \mathbb{R}, \quad \left(\sum_{i=1}^{n} x_i\right)^2 \leq n \sum_{i=1}^{n} x_i^2$$

3.11.0.1 Inégalité de Hölder

Théorème 3.12.

Soient $a_1, a_2, \ldots, a_n$ et $b_1, b_2, \ldots, b_n$ des nombres réels strictement positifs, p et q deux nombres réels strictement positifs, tels que $\frac{1}{p} + \frac{1}{q} = 1$, alors on a

$$\boxed{\sum_{i=1}^{n} a_i b_i \leq \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{q}}}$$

Preuve

Soient u et v deux nombres réels strictement positifs, comme $\frac{1}{p}+\frac{1}{q}=1$, alors la concavité de la fonction ln permet d'écrire

$$\ln\left(\frac{1}{p}u^p+\frac{1}{q}v^q\right)\geq\frac{1}{p}\ln(u^p)+\frac{1}{q}\ln(v^q)$$

avec $\frac{1}{p}\ln(u^p)+\frac{1}{q}\ln(v^q)=\ln(uv)$, ainsi on obtient

$$uv\leq\frac{1}{p}u^p+\frac{1}{q}v^q$$

Posons $A=\left(\sum\limits_{i=1}^{n}a_i^p\right)^{\frac{1}{p}}$, $B=\left(\sum\limits_{i=1}^{n}b_i^q\right)^{\frac{1}{q}}$ et pour chaque $i\in\{1,2,\dots,n\}$, posons $u_i=\frac{a_i}{A}$ et $v_i=\frac{b_i}{B}$, alors on aura

$$\forall i\in\{1,2,\dots,n\},\ u_iv_i\leq\frac{1}{p}u_i^p+\frac{1}{q}v_i^q$$

avec

$$\sum_{i=1}^{n}u_iv_i=\frac{\sum\limits_{i=1}^{n}a_ib_i}{AB}$$

$$\sum_{i=1}^{n}u_i^p=\frac{\sum\limits_{i=1}^{n}a_i^p}{A^p}=1$$

$$\sum_{i=1}^{n}u_i^p=\frac{\sum\limits_{i=1}^{n}b_i^q}{B^q}=1$$

Comme $\frac{1}{p}+\frac{1}{q}=1$, alors on aura le résultat

3.12.0.1 Inégalité de Minkowski

Théorème 3.13.

Soient $a_1,a_2,\dots,a_n$, $b_1,b_2,\dots,b_n$ des nombres réels strictement positifs et p un réel, avec $p>1$. Alors on a

$$\boxed{\left(\sum_{i=1}^{n}(a_i+b_i)^p\right)^{\frac{1}{p}}\leq\left(\sum_{i=1}^{n}a_i^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{n}b_i^p\right)^{\frac{1}{p}}}$$

Preuve
Posons $q = \dfrac{p}{p-1}$, alors $q > 0$ et $\dfrac{1}{p} + \dfrac{1}{q} = 1$, donc d'après l'inégalité de Hölder, on a

$$\sum_{i=1}^{n} a_i(a_i+b_i)^{p-1} \leq \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} (a_i+b_i)^{(p-1)q}\right)^{\frac{1}{q}}$$

On a aussi

$$\sum_{i=1}^{n} b_i(a_i+b_i)^{p-1} \leq \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} (a_i+b_i)^{(p-1)q}\right)^{\frac{1}{q}}$$

En faisant la somme de ces deux inégalités, on obtient,

$$\sum_{i=1}^{n} (a_i+b_i)^p \leq \left[\left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}}\right] \left(\sum_{i=1}^{n} (a_i+b_i)^p\right)^{1-\frac{1}{p}}$$

Donc en simplifiant, on aura

$$\left(\sum_{i=1}^{n} (a_i+b_i)^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}}$$