

Chapitre 7

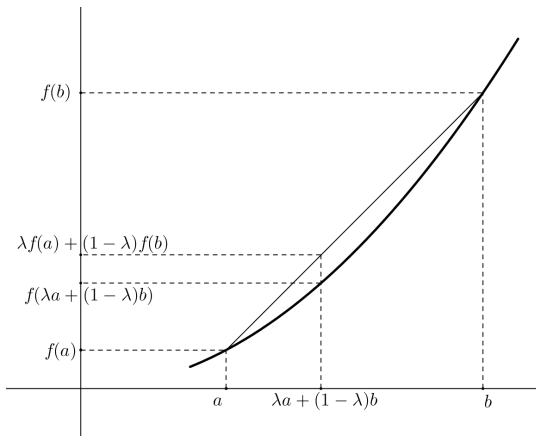
Convexité

7.1 Convexité : définition et premières propriétés

On commence par la définition générale d'une fonction convexe. Nous verrons dans la partie suivante des caractérisations plus simples dans le cas où f est dérivable ou deux fois dérivable.

Définition 7.1. Une fonction $f : I \rightarrow \mathbb{R}$ où I est intervalle de \mathbb{R} est dite *convexe* si elle vérifie l'*inégalité de convexité* suivante :

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



En observant la représentation graphique ci-dessus, on obtient la traduction géométrique suivante de la convexité d'une fonction :

Propriété 7.2. Soit $f : I \rightarrow \mathbb{R}$ et C sa courbe représentative. Alors f est convexe si et seulement si pour tout couple (A, B) de points distincts de C , l'arc \widehat{AB} sur C est au-dessous de la corde $[AB]$.

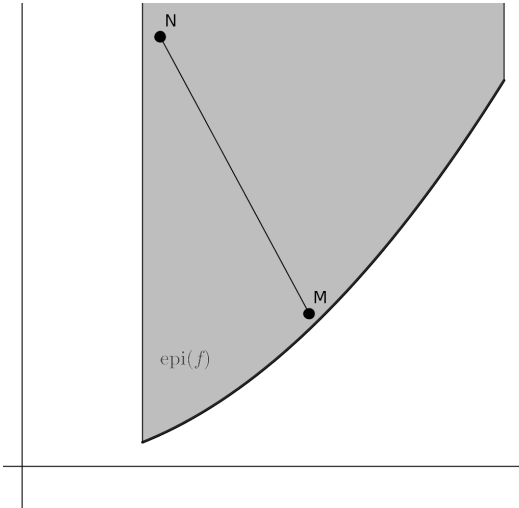
Définition 7.3. Une fonction $g : I \rightarrow \mathbb{R}$ où I est intervalle de \mathbb{R} est dite *concave* si $-g$ est convexe, ou encore si elle vérifie l'*inégalité de concavité* suivante :

$$\forall a, b \in I, \forall \lambda \in [0, 1], g(\lambda a + (1 - \lambda)b) \geq \lambda g(a) + (1 - \lambda)g(b);$$

ou de manière équivalente si pour tout couple (A, B) de points distincts de sa courbe représentative, l'arc \widehat{AB} est au-dessus de la corde $[AB]$.

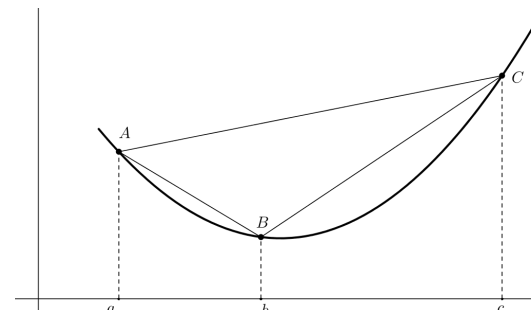
On peut également traduire la convexité d'une fonction par la convexité de son *épigraphe* en tant que partie de \mathbb{R}^2 :

Propriété 7.4. Soit $f : I \rightarrow \mathbb{R}$. On appelle *épigraphe* de f la partie de \mathbb{R}^2 notée $\text{epi}(f)$ et définie par $\text{epi}(f) = \{(x, y) \in \mathbb{R}^2 : y \geq f(x)\}$. Alors f est convexe si et seulement si $\text{epi}(f)$ est une partie convexe de \mathbb{R}^2 , c'est-à-dire pour tout couple (M, N) de points de $\text{epi}(f)$, on a $[M, N] \subseteq \text{epi}(f)$.



Propriété 7.5. Une fonction convexe $f : I \rightarrow \mathbb{R}$ vérifie l'*inégalité des pentes* :

$$\forall a, b, c \in I, a < b < c \Rightarrow \frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}.$$



Cette propriété suit géométriquement du fait que le point B est au-dessous de la corde $[AC]$, et est en fait équivalente à la convexité de f .

Exercice 7.6. Démontrer algébriquement l'inégalité des pentes, en utilisant la définition 7.1.

Pour une fonction $f : I \rightarrow \mathbb{R}$ et $a \in I$, on considère la fonction $\Delta_a f$ définie par $\Delta_a f(x) = \frac{f(x) - f(a)}{x - a}$ pour tout $x \in I \setminus \{a\}$. Les inégalités précédentes (inégalité des pentes) sur une fonction f se traduisent en terme de croissance des fonctions $\Delta_a f$:

Propriété 7.7. Une fonction $f : I \rightarrow \mathbb{R}$ est convexe si et seulement si pour tout $a \in I$, la fonction $\Delta_a f$ est croissante sur $I \setminus \{a\}$.

Exercice 7.8. Vérifier que l'inégalité des pentes est équivalente au fait que pour tout $a \in I$, la fonction $\Delta_a f$ est croissante sur $I \setminus \{a\}$.

Nous allons maintenant démontrer un lemme un peu plus difficile qui permettra de montrer que toute fonction convexe sur un intervalle ouvert est continue et admet des dérivées à gauche et à droite en tout point de cet intervalle ouvert.

Lemme 7.9. Soit I un intervalle de \mathbb{R} et $a \in \overset{\circ}{I}$. Si g est une fonction à valeurs réelles définie et croissante sur $I \setminus \{a\}$ alors g admet en a une limite à gauche et une limite à droite finies telles que

$$\lim_{x \rightarrow a^-} g(x) \leq \lim_{x \rightarrow a^+} g(x).$$

Démonstration. Comme $a \in \overset{\circ}{I}$, il existe $\eta > 0$ tel que $[a - \eta, a + \eta] \subset I$ et cela a donc un sens d'étudier l'existence d'éventuelles limites à gauche et à droite de g . On considère

$$X = \{g(x) : x \in I \text{ et } x < a\}.$$

Par croissance de g , cette partie X est majorée par $g(a + \eta)$. Notons l_1 sa borne supérieure. Encore par croissance de g , il suit que l_1 est la limite à gauche de g en a : en effet, pour $\varepsilon > 0$, il existe $x_0 \in X$ tel que $l_1 - \varepsilon < g(x_0) \leq l_1$ et alors pour tout x tel que $x_0 < x < a$, on a $l_1 - \varepsilon < g(x_0) \leq g(x) \leq l_1$ et en particulier $|l_1 - g(x)| < \varepsilon$.

De manière analogue, l'ensemble $X' = \{g(x') : x' \in I \text{ et } x' > a\}$ est minoré et sa borne inférieure l_2 correspond à la limite à droite de g en a . Enfin, comme pour tout $x \in X$ et tout $x' \in X'$, on a $x < a < x'$ et donc $g(x) \leq g(x')$, il suit que $l_1 \leq l_2$.

□

Proposition 7.10. Soit I un intervalle ouvert et f une fonction à valeurs réelles, définie et convexe sur I . Alors f est continue sur I et admet des dérivées à gauche et à droite en tout point a de I , vérifiant de plus $f'_g(a) \leq f'_d(a)$.

De plus, pour tout $a, b \in I$ tels que $a < b$, on a $f'_d(a) \leq \frac{f(b) - f(a)}{b - a} \leq f'_g(b)$.

Démonstration. Soit $a \in I$. Comme I est ouvert, on a $a \in \overset{\circ}{I}$, et comme f est convexe, la fonction $\Delta_a f$ est croissante. On peut appliquer le lemme précédent sur cette fonction, ce qui permet de conclure car les limites à gauche et droite de $\Delta_a f$ en a correspondent exactement aux dérivées à gauche et à droite de f en a . Bien sûr, la dérivabilité à droite et à gauche en a impliquent en particulier que $f(x) - f(a)$ tend vers 0 quand x tend vers a , et donc que f est continue en a .

Pour démontrer les inégalités apparaissant à la dernière ligne de l'énoncé ci-dessus, choisissons $x \in]a, b[$; la croissance de la fonction $\Delta_x f$ nous donne $\frac{f(a) - f(x)}{a - x} \leq \frac{f(b) - f(x)}{b - x}$. En faisant tendre x vers a et en utilisant la continuité de f , on obtient l'inégalité de gauche ; en faisant tendre x vers b on obtient celle de droite.

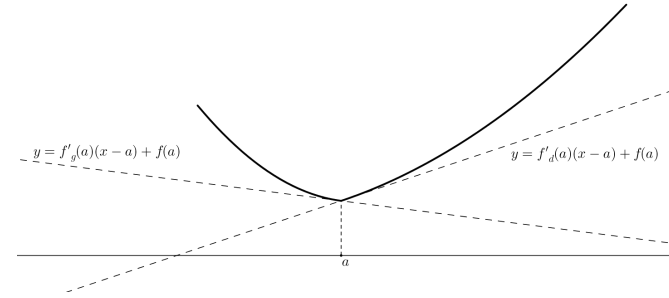
□

Notons qu'il est immédiat que $f'_g(a) \leq f'_d(a)$ pour tout $a \in \mathbb{R}$ (la limite à gauche d'une fonction croissante est toujours inférieure à sa limite à droite!) ; en particulier, si f est convexe alors f'_g et f'_d sont toutes deux des fonctions croissantes.

Théorème 7.11. Soit I un intervalle ouvert de \mathbb{R} , et $f : I \rightarrow \mathbb{R}$ une fonction convexe. Alors pour tout $a \in I$ on a pour tout $x \in I$ que $f(x) \geq f'_d(a)(x - a) + f(a)$ et $f(x) \leq f'_g(a)(x - a) + f(a)$. En particulier, il existe une fonction affine g telle que $g(a) = f(a)$ et $g(x) \leq f(x)$ pour tout $x \in I$.

Démonstration. Soit $a \in I$. D'après les inégalités à la fin de la proposition 7.10 on a pour tout $x > a$ que $\frac{f(x) - f(a)}{x - a} \geq f'_d(a)$ et donc $f(x) \geq f'_d(a)(x - a) + f(a)$. De même, pour tout $x < a$ on a $\frac{f(x) - f(a)}{x - a} \leq f'_g(a)$; en multipliant par $x - a$ (qui est négatif!) on a donc que pour tout $x < a$ $f(x) \geq f(a) + f'_g(a)(x - a)$.

De plus, on sait que $f'_g(a) \leq f'_d(a)$; par conséquent, pour $x < a$ $f'_g(a)(x - a) \geq f'_d(a)(x - a)$, et on voit finalement que l'inégalité $f(x) \geq f'_d(a)(x - a) + f(a)$ est valide pour tout $x \in \mathbb{R}$. Le même raisonnement s'applique pour montrer que l'autre inégalité est vraie pour tout $x \in \mathbb{R}$.



Proposition 7.12. Toute fonction convexe sur \mathbb{R} et majorée est constante.

Démonstration. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction convexe majorée. D'après le résultat précédent, on a $f(x) \geq f'_d(a)(x - a) + f(a)$ pour tout $a \in \mathbb{R}$; comme f est majorée ceci impose que $f'_d(a) = 0$ pour tout $a \in \mathbb{R}$ (faites un dessin et considérez les limites en $+\infty$ ou $-\infty$ selon le signe de $f'_d(a)$). De même on doit avoir $f'_g(a) = 0$ pour tout $a \in \mathbb{R}$; par conséquent $f'_d(a) = f'_g(a) = 0$ pour tout $a \in \mathbb{R}$. Autrement dit, f est dérivable et $f' = 0$: f est constante. □

Exercice 7.13. Donner un exemple de fonction convexe sur $[0, 1]$ qui n'est ni continue en 0, ni continue en 1.

7.2 Fonctions convexes dérivables

Proposition 7.14. Soit I un intervalle ouvert et $f : I \rightarrow \mathbb{R}$ dérivable sur I . Alors les propriétés suivantes sont équivalentes :

1. La fonction f est convexe sur I .
2. Pour tout $x, y \in I$ $f(x) \geq (x - y)f'(y) + f(y)$.
3. La fonction f' est croissante sur I .

Notons l'interprétation géométrique du second point ci-dessus : le graphe de f est au-dessus de ses tangentes.

Démonstration. Supposons f convexe et dérivable. Alors en tout $x \in I$ on a $f'(x) = f'_d(x) = f'_g(x)$, donc la deuxième condition est vraie d'après le théorème 7.11.

Supposons maintenant la deuxième condition satisfaite, et considérons $x < y \in I$. Alors on a à la fois $f(y) \geq f(x) + (y - x)f'(x)$ et $f(x) \geq f(y) + (x - y)f'(y)$, ce qui nous donne $f'(x)(y - x) \leq f(y) - f(x) \leq f'(y)(y - x)$. On en déduit que $f'(x) \leq f'(y)$ et donc f' est croissante.

Enfin, supposons f' croissante, fixons $x < y \in I$ et considérons la fonction $g : [0, 1] \rightarrow \mathbb{R}$ définie par

$$g(t) = tf(x) + (1 - t)f(y) - f(tx + (1 - t)y).$$

Notre but est de démontrer que g est à valeurs positives. Notons qu'on a $g(0) = g(1) = 0$, et que g est dérivable, de dérivée donnée par la formule

$$g'(t) = f(x) - f(y) + (y - x)f'(tx + (1 - t)y).$$

L'hypothèse selon laquelle f' est croissante entraîne donc que g' est décroissante ; et le théorème de Rolle nous garantit l'existence de $c \in]0, 1[$ tel que $g'(c) = 0$. Ainsi, g' est positive sur $[0, c]$ et négative sur $[c, 1]$, ce qui signifie que g est croissante sur $[0, c]$ et décroissante sur $[c, 1]$. Comme $g(0) = g(1) = 0$, on vient d'établir que g ne prend que des valeurs positives. □

Exercice 7.15. Soit I un intervalle ouvert de \mathbb{R} , f une fonction convexe sur I et dérivable. Montrer que $a \in I$ est un minimum de f sur I si et seulement si $f'(a) = 0$.

Proposition 7.16. Soit I un intervalle ouvert et $f: I \rightarrow \mathbb{R}$ une fonction deux fois dérivable. Alors f est convexe si et seulement si f'' est positive sur I .

Démonstration. Il suffit de remarquer que f'' est positive si et seulement si f' est croissante, et d'appliquer le résultat précédent. \square

Remarquons ici qu'une fonction dérivée vérifie toujours la propriété des valeurs intermédiaires (c'est le contenu du théorème de Darboux); et une fonction croissante vérifiant la propriété des valeurs intermédiaires est nécessairement continue. Ainsi, si une fonction convexe définie sur un intervalle ouvert est dérivable, sa dérivée est nécessairement continue.

7.3 Inégalités de convexité

La convexité (ou la concavité) est souvent utilisée pour établir des inégalités. Citons un exemple important.

Proposition 7.17. Soit $f:]0, +\infty[\rightarrow]0, +\infty[$ une fonction concave. Alors pour tout $x, y \geq 0$ on a $f(x+y) \leq f(x) + f(y)$.

Démonstration. Fixons $y \geq 0$ et considérons la fonction $g:]0, +\infty[\rightarrow \mathbb{R}$ définie par $g(x) = f(x) + f(y) - f(x+y)$.

Alors, pour tout $a < b \in]0, +\infty[$, on a

$$\frac{g(b) - g(a)}{b - a} = \frac{f(b) - f(a)}{b - a} - \frac{f(b+y) - f(a+y)}{b - a}.$$

Puisque $\frac{f(b+y) - f(a+y)}{b - a} = \frac{f(b+y) - f(a+y)}{(b+y) - (a+y)}$ est le taux d'accroissement de f entre $(b+y)$ et $(a+y)$,

l'inégalité des pentes nous donne donc que $\frac{g(b) - g(a)}{b - a} \geq 0$, autrement dit g est croissante.

Par conséquent, on a pour tout x que $g(x) \geq g(0) = f(0)$, et donc $f(x) + f(y) - f(x+y) \geq f(0) \geq 0$, ce qu'on voulait démontrer. \square

Voyons maintenant l'inégalité de convexité la plus importante de notre cours.

Théorème 7.18 (Inégalité de Jensen). Soit (X, \mathcal{A}, μ) un espace de probabilité, g une fonction μ -intégrable à valeurs dans un intervalle I , et $\varphi: I \rightarrow \mathbb{R}$ une fonction convexe. Alors on a

$$\varphi\left(\int_X g d\mu\right) \leq \int_X \varphi \circ g d\mu$$

(l'intégrale de droite peut être égale à $+\infty$!).

Démonstration. Posons $m = \int_X g d\mu$. Notons que $m \in I$; si jamais m est le minimum de I (s'il existe!) alors on a $\int_X (g - m) d\mu = 0$ et $g - m \geq 0$, donc $g - m$ est nulle presque partout, par conséquent on a

$$\int_X \varphi \circ g d\mu = \int_X \varphi(m) d\mu = \varphi(m) = \varphi\left(\int_X g d\mu\right).$$

On traite de même le cas où m est le maximum de I ; finalement, le cas qui nous reste est celui où m appartient à l'intérieur de I .

Alors, on sait que $\varphi'_g(m)$ existe et en posant $\alpha = \varphi'_g(m)$, le théorème 7.11 donne que

$$\forall t \in I \quad \varphi(t) - \varphi(m) \geq \alpha(t - m).$$

En particulier, pour tout $x \in X$ on a $\varphi(g(x)) \geq \varphi(m) + \alpha(g(x) - m)$. Comme g est intégrable, on en déduit que la partie négative de $\varphi \circ g$ est d'intégrale finie; et en intégrant cette inégalité, on obtient aussi que

$$\int_X \varphi \circ g d\mu \geq \int_X \varphi(m) d\mu + \alpha \int_X (g - m) d\mu = \varphi(m) + \alpha \left(\int_X g d\mu - m \right) = \varphi(m).$$

\square

Le corollaire suivant est un cas (très) particulier de l'inégalité de Jensen, qui peut se montrer élémentairement, sans théorie de la mesure.

Corollaire 7.19. Soit I un intervalle de \mathbb{R} , $\alpha_1, \dots, \alpha_n$ des réels positifs tels que $\sum_{i=1}^n \alpha_i = 1$, et φ une fonction convexe sur I . Alors, pour tout $x_1, \dots, x_n \in I$ on a

$$\varphi\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i \varphi(x_i).$$

Démonstration. On fixe $x_1, \dots, x_n \in I$ et on considère l'espace mesuré d'ensemble sous-jacent $X = \{x_1, \dots, x_n\}$, où toutes les parties sont mesurables et $\mu = \sum_{i=1}^n \alpha_i \delta_{x_i}$, où δ_{x_i} désigne la mesure de Dirac en x_i . Alors μ est une mesure de probabilité; de plus pour toute fonction $g: X \rightarrow \mathbb{R}$ on a

$$\int_X g d\mu = \sum_{i=1}^n \alpha_i g(x_i).$$

En considérant pour g la fonction identité, on a donc $\int_X \varphi \circ g d\mu = \sum_{i=1}^n \alpha_i \varphi(x_i)$, et $\int_X g d\mu = \sum_{i=1}^n \alpha_i x_i$.

L'inégalité de Jensen nous donne donc comme attendu

$$\varphi\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i \varphi(x_i).$$

\square

Remarque 7.20. Dans le corollaire ci-dessus, le cas $n = 2$ correspond exactement à la définition de la convexité. En particulier, une application φ qui satisfait l'inégalité de Jensen pour toute fonction intégrable sur un espace de probabilité, est nécessairement convexe.

Exercice 7.21. Donner une démonstration du corollaire ci-dessus qui n'utilise que la définition d'une fonction convexe (et le principe de récurrence).

Le corollaire 7.19 permet d'établir plusieurs inégalités classiques, comme par exemples celles de l'exercice suivant.

Exercice 7.22. 1. Montrer que la fonction $x \mapsto -\ln(x)$ est convexe sur $]0, +\infty[$. En déduire l'inégalité arithmético-géométrique : pour tout x_1, \dots, x_n positifs on a

$$\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^n x_i.$$

2. Montrer que pour $p \geq 1$ la fonction $x \mapsto x^p$ est convexe sur $]0, +\infty[$ et en déduire que pour tout x_1, \dots, x_n positifs on a

$$\frac{1}{n} \sum_{k=1}^n x_k \leq \left(\frac{1}{n} \sum_{k=1}^n x_k^p\right)^{\frac{1}{p}}.$$

Remarque 7.23. Dans l'exercice ci-dessus, pour montrer que $x \mapsto x^p$ est convexe sur $]0, +\infty[$, on peut montrer qu'elle est convexe sur $]0, +\infty[$ et continue sur $]0, +\infty[$, conditions qui impliquent la convexité sur $]0, +\infty[$.

Proposition 7.24. Supposons que $0 < p < q < +\infty$ et que (X, \mathcal{A}, μ) est un espace de probabilité. Alors, pour toute fonction positive mesurable f , on a

$$\left(\int_X f^p d\mu\right)^{\frac{1}{p}} \leq \left(\int_X f^q d\mu\right)^{\frac{1}{q}}$$

Démonstration. Posons $r = \frac{q}{p}$ et $g = f^p$. Comme $x \mapsto x^r$ est convexe sur $[0, +\infty[$, l'inégalité de Jensen nous dit que quand g est intégrable on a

$$\left(\int_X g d\mu \right)^r \leq \int_X g^r d\mu .$$

Autrement dit, $\left(\int_X f^p d\mu \right)^{\frac{q}{p}} \leq \int_X f^q d\mu$, et cette inégalité est équivalente à celle qu'on souhaitait établir.

Reste à traiter le cas où g n'est pas intégrable ; dans ce cas, on peut par exemple poser $f_n = \min(f, n)$, et noter que les suites (f_n^p) et (f_n^q) sont croissantes vers f^p et f^q respectivement ; on peut donc appliquer le théorème de convergence monotone pour conclure que

$$\int_X f^p d\mu = \lim_{n \rightarrow +\infty} \int_X f_n^p d\mu \quad \text{et} \quad \int_X f^q d\mu = \lim_{n \rightarrow +\infty} \int_X f_n^q d\mu .$$

Comme chaque f_n^p est intégrable, la première étape de notre raisonnement nous donne $\left(\int_X f_n^p d\mu \right)^{\frac{1}{p}} \leq \left(\int_X f_n^q d\mu \right)^{\frac{1}{q}}$, et l'inégalité souhaitée s'en déduit par passage à la limite. \square

Notation. Si (X, \mathcal{A}, μ) est un espace mesuré, f est une fonction mesurable sur X et $0 < p < +\infty$, on notera

$$\|f\|^p = \left(\int_X |f|^p d\mu \right)^{\frac{1}{p}} .$$

Ainsi, la Proposition 7.24 consiste à dire que, sur un espace de probabilité, on a pour toutes fonctions f, g mesurables positives l'inégalité $\|f\|_p \leq \|f\|_q$ dès que $0 < p < q < +\infty$.

Définition 7.25. Soit $p, q \in [1, +\infty[$. On dit que p et q sont des *exposants conjugués* si on a $\frac{1}{p} + \frac{1}{q} = 1$.

On dit qu'alors q est l'exposant conjugué de p .

Notons que, si q est l'exposant conjugué de p , alors p est l'exposant conjugué de q ; 1 est l'exposant conjugué de $+\infty$, tandis que l'exposant conjugué de 2 est 2.

Cette notion est importante en grande partie à cause de l'inégalité suivante.

Théorème 7.26 (Inégalité de Hölder). *Soit (X, \mathcal{A}, μ) un espace mesuré, $p, q \in]1, +\infty[$ deux exposants conjugués et f, g deux fonctions mesurables à valeurs réelles telles que $|f|^p$ et $|g|^q$ sont intégrables. Alors le produit fg est intégrable, et on a*

$$\int_X |fg| d\mu \leq \left(\int_X |f|^p d\mu \right)^{\frac{1}{p}} \left(\int_X |g|^q d\mu \right)^{\frac{1}{q}} .$$

De manière plus condensée : $\|fg\|_1 \leq \|f\|_p \|g\|_q$ dès que $\|f\|_p$ et $\|g\|_q$ sont finies.

Démonstration. Comme seules les valeurs absolues de f, g interviennent dans l'énoncé, on peut supposer f, g à valeurs positives. Notons ensuite que, si f ou g est nulle presque partout alors il en va de même du produit fg et l'inégalité désirée est vraie. On peut donc supposer que g n'est pas nulle presque partout, auquel cas $\|g\|_q > 0$.

Considérons l'espace mesurable (X, \mathcal{A}) muni de la mesure à densité ν définie par $\nu(A) = \int_A \frac{g^q}{\int_X g^q d\mu} d\mu$. C'est un espace de probabilité.

Considérons la fonction $h = fg^{1-q}$; alors en utilisant la proposition 7.24 et le fait que, comme p et q sont

conjugués, on a $p(1-q) + q = 0$, on obtient :

$$\begin{aligned} \int_X fg d\mu &= \left(\int_X g^q d\mu \right) \left(\int_X fg^{1-q} \frac{g^q}{\int_X g^q d\mu} d\mu \right) \\ &= \left(\int_X g^q d\mu \right) \int_X h d\nu \\ &\leq \left(\int_X g^q d\mu \right) \left(\int_X h^p d\nu \right)^{\frac{1}{p}} \\ &\leq \left(\int_X g^q d\mu \right) \left(\int_X f^p g^{p(1-q)} \frac{g^q}{\int_X g^q d\mu} d\mu \right)^{\frac{1}{p}} \\ &= \left(\int_X g^q d\mu \right) \left(\int_X \frac{f^p}{\int_X g^q d\mu} d\mu \right)^{\frac{1}{p}} \\ &= (\|g\|_q)^{q(1-\frac{1}{p})} \|f\|_p \\ &= \|g\|_q \|f\|_p . \end{aligned}$$

\square

L'inégalité de Hölder est une des inégalités fondamentales concernant les espaces L^p , qui constitue la matière du prochain chapitre ; dans le cas particulier où $p = q = 2$, on retrouve l'inégalité de Cauchy-Schwarz, valable dès que f^2 et g^2 sont intégrables :

$$\int_X |fg| d\mu \leq \left(\int_X f^2 d\mu \right)^{\frac{1}{2}} \left(\int_X g^2 d\mu \right)^{\frac{1}{2}} .$$