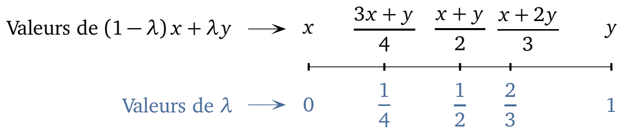
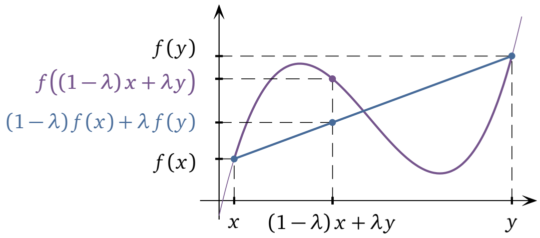
Convexité

Dans tout ce chapitre, I est un **intervalle** de $\mathbb{R}$ et non une partie quelconque. En d'autres termes, pour tous $x, y \in I$ pour lesquels $x \leqslant y$: $\quad [x, y] \subset I, \quad$ où l'on rappelle que par définition : $\quad [x, y] = \big\{ t \in \mathbb{R} \mid \quad x \leqslant t \leqslant y \big\}$.

Tâchons cependant de concevoir $[x, y]$ autrement. Quand λ décrit $[0, 1]$, $\lambda(y - x)$ décrit le segment d'extrémités 0 et $y - x$, donc $(1 - \lambda)x + \lambda y = x + \lambda(y - x)$ décrit le segment d'extrémités x et y.



Conclusion : $\quad \boxed{[x, y] = \big\{ (1 - \lambda)x + \lambda y \mid \quad \lambda \in [0, 1] \big\}.}$



Donnons-nous maintenant une fonction $f : I \longrightarrow \mathbb{R}$. Quand λ décrit $[0, 1]$, $(1 - \lambda)f(x) + \lambda f(y)$ décrit le segment d'extrémités $f(x)$ et $f(y)$, et dans le plan, le point de coordonnées $\big((1-\lambda)x+\lambda y \,,\, (1-\lambda)f(x)+\lambda f(y)\big)$ décrit le segment d'extrémités $\big(x, f(x)\big)$ et $\big(y, f(y)\big)$, appelé une *corde de f*.

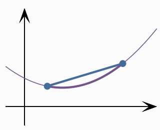
Si $x \neq y$, la droite complète qui passe par $\big(x, f(x)\big)$ et $\big(y, f(y)\big)$ est quant à elle appelée une *sécante de f*.

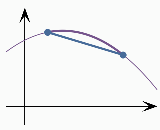
Définition (Fonction convexe/concave) Soit $f : I \longrightarrow \mathbb{R}$ une fonction.

- **Fonction convexe :** On dit que f est *convexe* si son graphe est situé en-dessous de toutes ses cordes, i.e. si :
$$\forall x, y \in I, \quad \forall \lambda \in [0, 1], \quad f\big((1 - \lambda)x + \lambda y\big) \leqslant (1 - \lambda)f(x) + \lambda f(y).$$
- **Fonction concave :** On dit que f est *concave* si son graphe est situé au-dessus de toutes ses cordes, i.e. si :
$$\forall x, y \in I, \quad \forall \lambda \in [0, 1], \quad f\big((1 - \lambda)x + \lambda y\big) \geqslant (1 - \lambda)f(x) + \lambda f(y).$$

Il est équivalent de dire que $-f$ est convexe.

Fonction convexe,
perpétuel virage à gauche





Fonction concave,
perpétuel virage à droite

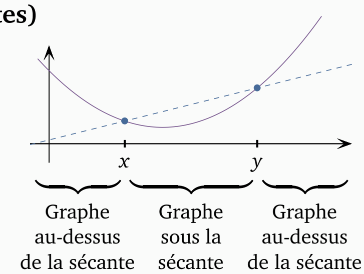
Nous n'énoncerons et ne démontrerons ci-après que des théorèmes sur les fonctions convexes, mais le cas des fonctions concaves s'en déduit par multiplication par -1. Remplacer simplement dans chaque résultat convexe le plus petit par le plus grand, le positif par le négatif, la croissance par la décroissance et le dessus par le dessous.

Exemple La fonction valeur absolue est convexe sur $\mathbb{R}$ car pour tous $x, y \in \mathbb{R}$ et $\lambda \in [0, 1]$, d'après l'inégalité triangulaire :
$$\big|(1 - \lambda)x + \lambda y\big| \leqslant |1 - \lambda| \times |x| + |\lambda| \times |y| = (1 - \lambda)|x| + \lambda|y|.$$

Théorème (Position du graphe d'une fonction convexe par rapport à ses sécantes)

Soient $f : I \longrightarrow \mathbb{R}$ une fonction convexe et $x, y \in I$ avec $x < y$.

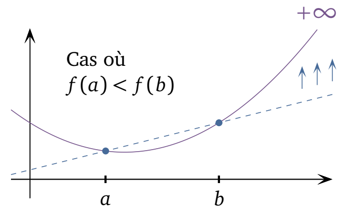
Le graphe de f est situé sous sa sécante sur $[x, y]$ et au-dessus à l'extérieur de $[x, y]$.



Démonstration Montrons que pour tout $t \in I \cap [y, +\infty[$: $\quad f(t) \geqslant \dfrac{f(y) - f(x)}{y - x}(t - y) + f(y) \quad$ — on travaillerait de même à gauche de x.

Fixons $t \in I \cap [y, +\infty[$. Il s'agit de montrer que : $\quad (t - x)f(y) \leqslant (t - y)f(x) + (y - x)f(t), \quad$ ou encore que $f(y) \leqslant \left(1 - \dfrac{y - x}{t - x}\right)f(x) + \dfrac{y - x}{t - x}f(t)$. Or $x < y \leqslant t$, donc $\dfrac{y - x}{t - x}$ appartient à $[0, 1]$, et si nous notons λ ce réel : $\quad (1 - \lambda)x + \lambda t = \dfrac{t - y}{t - x}x + \dfrac{y - x}{t - x}t = y. \quad$ On conclut par convexité de f :
$$f(y) = f\big((1 - \lambda)x + \lambda y\big) \leqslant (1 - \lambda)f(x) + \lambda f(y).$$

Exemple Soient $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction convexe et $a, b \in \mathbb{R}$ avec $a < b$. Si $f(a) < f(b)$: $\lim\limits_{x \to +\infty} f(x) = +\infty$, et si $f(a) > f(b)$: $\lim\limits_{x \to -\infty} f(x) = +\infty$.



Démonstration Le graphe de f est au-dessus de sa sécante à l'extérieur de $[a, b]$, donc pour tout $t < a$ ou $t > b$: $f(t) \geqslant \dfrac{f(b) - f(a)}{b - a}(t - a) + f(a)$. Le résultat en découle par minoration en distinguant bien le cas $f(a) < f(b)$ où l'on fait tendre t vers $+\infty$ du cas $f(a) > f(b)$ où on le fait tendre vers $-\infty$.

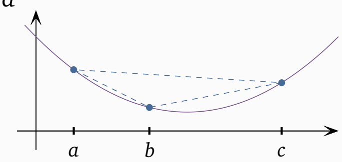
Théorème (Caractérisation de la convexité par les pentes des sécantes, inégalité des pentes) Soit $f : I \longrightarrow \mathbb{R}$ une fonction.

(i) **Caractérisation de la convexité par les pentes des sécantes :**

f est convexe sur I si et seulement si pour tout $a \in I$, la fonction $x \longmapsto \dfrac{f(x) - f(a)}{x - a}$ est croissante sur $I \setminus \{a\}$.

(ii) **Inégalité des pentes :** Si f est convexe sur I, alors pour tous $a, b, c \in I$ avec $a < b < c$:

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



Bref, les pentes des sécantes croissent en cas de convexité !

Démonstration

(i) Supposons d'abord f convexe sur I. Fixons $a \in I$ et montrons que la fonction $t \longmapsto \dfrac{f(t) - f(a)}{t - a}$ est croissante sur $I \setminus \{a\}$. Soient $x, y \in I \setminus \{a\}$ avec $x < y$. Nous allons distinguer deux cas et les traiter chacun en comparant le graphe de f à sa sécante passant par $\big(a, f(a)\big)$ et $\big(x, f(x)\big)$, i.e. en comparant les fonctions f et $t \longmapsto \dfrac{f(t) - f(a)}{t - a}(t - a) + f(a)$.

— Si $y > a$, y est situé à l'extérieur du segment $[a, x]$ ou $[x, a]$, donc $f(y) \geqslant \dfrac{f(x) - f(a)}{x - a}(y - a) + f(a)$, ou encore $\dfrac{f(x) - f(a)}{x - a} \leqslant \dfrac{f(y) - f(a)}{y - a}$ puisque $y - a > 0$.

— Si $y < a$, y appartient au segment $[a, x]$ ou $[x, a]$, donc $f(y) \leqslant \dfrac{f(x) - f(a)}{x - a}(y - a) + f(a)$, ou encore $\dfrac{f(x) - f(a)}{x - a} \leqslant \dfrac{f(y) - f(a)}{y - a}$ puisque $y - a < 0$.

C'est l'inégalité souhaitée dans les deux cas.

Réciproquement, supposons $t \longmapsto \dfrac{f(t) - f(a)}{t - a}$ croissante sur $I \setminus \{a\}$ pour tout $a \in I$ et montrons que f est convexe. Fixons $x, y \in I$ et $\lambda \in [0, 1]$. Pour montrer que $f\big((1 - \lambda)x + \lambda y\big) \leqslant (1 - \lambda)f(x) + \lambda f(y)$, on peut supposer sans perte de généralité que $\lambda \in]0, 1[$ et $x \neq y$, et même $x < y$ quitte à remplacer λ par $(1 - \lambda)$.

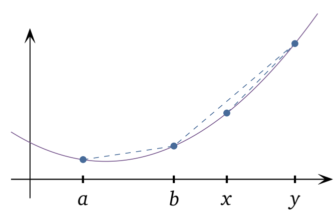
Posons maintenant $a = (1 - \lambda)x + \lambda y$, de sorte que $x < a < y$. La fonction $t \longmapsto \dfrac{f(t) - f(a)}{t - a}$ étant croissante par hypothèse : $\dfrac{f(x) - f(a)}{x - a} \leqslant \dfrac{f(y) - f(a)}{y - a}$, donc sachant que $x - a < 0$ et $y - a > 0$:

$$(y - a)\big(f(x) - f(a)\big) \geqslant (x - a)\big(f(y) - f(a)\big), \qquad \text{puis } f(a) \leqslant \frac{y - a}{y - x} f(x) + \frac{a - x}{y - x} f(y),$$

or ceci s'écrit aussi : $f(a) \leqslant (1 - \lambda)f(x) + \lambda f(y)$ par définition de a et c'est le résultat voulu.

(ii) La fonction $t \longmapsto \dfrac{f(t) - f(a)}{t - a}$ est croissante sur $I \setminus \{a\}$ d'après (i), donc $\dfrac{f(b) - f(a)}{b - a} \leqslant \dfrac{f(c) - f(a)}{c - a}$. De même, la fonction $t \longmapsto \dfrac{f(t) - f(c)}{t - c}$ est croissante sur $I \setminus \{c\}$, donc $\dfrac{f(c) - f(a)}{c - a} \leqslant \dfrac{f(c) - f(b)}{c - b}$.

Exemple Soient $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction convexe et $a, b \in \mathbb{R}$ deux réels pour lesquels $a < b$ et $f(a) \leqslant f(b)$. Alors f est croissante sur $I \cap [b, +\infty[$. Ce résultat complète bien sûr notre dernier exemple.



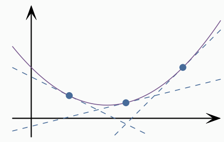
Démonstration Soient $x, y \in I \cap [b, +\infty[$ avec $x < y$. Pour montrer que $f(x) \leqslant f(y)$, une simple figure nous suggère d'appliquer deux fois l'inégalité des pentes : $\dfrac{f(y) - f(x)}{y - x} \geqslant \dfrac{f(y) - f(b)}{y - b} \geqslant \dfrac{f(b) - f(a)}{b - a} \geqslant 0$.

Théorème (Caractérisation des fonctions convexes dérivables) Soit $f \in \mathscr{D}(I,\mathbb{R})$. Les assertions suivantes sont équivalentes :

(i) f est convexe sur I.

(ii) f' est croissante sur I — ou bien $f'' \geqslant 0$ si f est deux fois dérivable sur I.

(iii) Le graphe de f est situé au-dessus de toutes ses tangentes.



Démonstration

(i) $\Longrightarrow$ (ii) Soient $x,y \in I$ avec $x<y$. Pour tout $t \in]x,y[$: $\dfrac{f(t)-f(x)}{t-x} \leqslant \dfrac{f(y)-f(x)}{y-x} \leqslant \dfrac{f(y)-f(t)}{y-t}$ par convexité de f d'après l'inégalité des pentes. Faisons indépendamment tendre t vers x à droite puis vers y à gauche : $f'(x) \leqslant \dfrac{f(y)-f(x)}{y-x} \leqslant f'(y)$. Comme voulu, f' est croissante sur I.

(ii) $\Longrightarrow$ (iii) Fixons $a \in I$ et montrons que le graphe de f est situé au-dessus de sa tangente en a. Notons pour cela φ la fonction $t \longmapsto f(t) - f'(a)(t-a) - f(a)$. Cette fonction est dérivable sur I et pour tout $x \in I$: $\varphi'(x) = f'(x) - f'(a)$. Par croissance de f', φ' est négative à gauche de a et positive à droite. Ainsi, φ est décroissante à gauche de a et croissante à droite, donc positive sur I tout entier puisque $\varphi(a)=0$.

(iii) $\Longrightarrow$ (i) Montrons que f est convexe sur I. Soient $x,y \in I$ et $\lambda \in [0,1]$. Posons $a = (1-\lambda)x + \lambda y$. Pour tout $t \in I$: $f(t) \geqslant f'(a)(t-a) + f(a)$, donc :

$$\begin{aligned}(1-\lambda)f(x)+\lambda f(y) &\geqslant (1-\lambda)\Big(f'(a)(x-a)+f(a)\Big)+\lambda\Big(f'(a)(y-a)+f(a)\Big)\\ &= f'(a)\Big((1-\lambda)x+\lambda y-a\Big)+f(a) = f\Big((1-\lambda)x+\lambda y\Big).\end{aligned}$$

Exemple La fonction logarithme est concave sur $\mathbb{R}_+^*$ car sa dérivée seconde $x \longmapsto -\dfrac{1}{x^2}$ est négative, donc pour tous $x,y>0$ et $\lambda = \dfrac{1}{2}$ dans la définition de la concavité : $\ln\dfrac{x+y}{2} \geqslant \dfrac{1}{2}\ln x + \dfrac{1}{2}\ln y$.

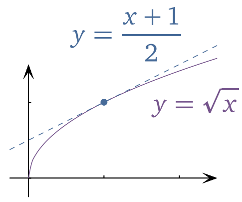
De nombreuses inégalités de convexité découlent de la définition ou des résultats qui précèdent. Rappelons-en trois qu'il faut bien connaître et comprendre graphiquement. À l'exception de la minoration du sinus, ces inégalités illustrent toutes l'idée que le graphe d'une fonction convexe (resp. concave) est au-dessus (resp. en dessous) de ses tangentes. La minoration du sinus exprime quant à elle l'idée que la corde du sinus joignant les points d'abscisses 0 et $\dfrac{\pi}{2}$ a pour équation $y = \dfrac{2x}{\pi}$.

$$\forall x \in \mathbb{R}, \quad e^x \geqslant 1+x.$$

$$\forall x > -1, \quad \ln(1+x) \leqslant x. \qquad \forall x > 0, \quad \ln x \leqslant x-1.$$

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

Exemple Pour tout $x>0$: $\sqrt{x} \leqslant \dfrac{x+1}{2}$.



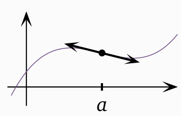
Démonstration La fonction $x \longmapsto \sqrt{x}$ est concave sur $\mathbb{R}_+^*$ car sa dérivée $x \longmapsto \dfrac{1}{2\sqrt{x}}$ y est décroissante. Son graphe est situé sous sa tangente en 1, d'équation $y = \dfrac{x+1}{2}$.

Avant de poursuivre, revenons rapidement sur la définition de la convexité. On dit qu'une fonction $f : I \longrightarrow \mathbb{R}$ est *strictement convexe sur* I si pour tous $x,y \in I$ distincts et pour tout $\lambda \in]0,1[$: $f\big((1-\lambda)x+\lambda y\big) < (1-\lambda)f(x)+\lambda f(y)$. L'intuition géométrique est la même qu'en cas de convexité simple à ceci près qu'une fonction strictement convexe ne peut pas être affine sur un intervalle $[a,b]$ avec $a<b$. Son graphe est toujours strictement au-dessus de ses cordes.

Les résultats précédents s'étendent bien au cadre de la convexité stricte. L'inégalité des pentes devient simplement stricte et une fonction $f : I \longrightarrow \mathbb{R}$ est strictement convexe si et seulement si pour tout $a \in I$, la fonction $x \longmapsto \dfrac{f(x)-f(a)}{x-a}$ est strictement croissante sur $I \setminus \{a\}$. Enfin, une fonction dérivable $f \in \mathscr{D}(I,\mathbb{R})$ est strictement convexe si et seulement si sa dérivée f' est strictement croissante, si et seulement si pour tout $a \in I$, son graphe est situé au-dessus de sa tangente en a avec $\big(a, f(a)\big)$ pour seul point de contact. Par exemple, la fonction exponentielle est strictement convexe sur $\mathbb{R}$ car sa dérivée est strictement croissante, donc dans l'inégalité de convexité $e^x \geqslant 1+x$, l'égalité n'est atteinte que pour $x=0$.

Définition-théorème (Point d'inflexion) Soient $f : I \longrightarrow \mathbb{R}$ une fonction et a un point intérieur de I. On dit que f *possède un point d'inflexion en* a si f est convexe au voisinage de a à gauche et concave au voisinage de a à droite — ou l'inverse.

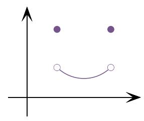
Si f est deux fois dérivable sur I, f possède un point d'inflexion en a si et seulement si f'' s'annule en a et est positive au voisinage de a à gauche et négative au voisinage de a à droite — ou l'inverse.



Exemple Soient $a \in \mathbb{R}^*$ et $b, c \in \mathbb{R}$. La fonction $x \longmapsto ax^3 + bx^2 + cx + d$ possède un et un seul point d'inflexion, à savoir en $-\dfrac{b}{3a}$, car sa dérivée seconde $x \longmapsto 6ax + 2b$ s'annule en changeant de signe en $-\dfrac{b}{3a}$ et ne le fait nulle part ailleurs.

Théorème (Convexité implique continuité) Pour une fois, l'intervalle I est supposé OUVERT. Soit $f : I \longrightarrow \mathbb{R}$ une fonction. Si f est convexe sur I, f est continue sur I.

Attention ! Il est essentiel que l'intervalle I soit OUVERT. La fonction « smiley » ci-contre est convexe mais pas continue aux bornes.



Le résultat peut être amélioré sur le terrain de la dérivabilité, nous verrons ça en TD. Attention tout de même, une fonction convexe n'est pas forcément dérivable sur son ensemble de définition, pensez à la fonction valeur absolue !

Démonstration Fixons $a \in I$. Par convexité de f, la fonction $x \longmapsto \dfrac{f(x) - f(a)}{x - a}$ est croissante sur $I \setminus \{a\}$, donc possède des limites FINIES à gauche et à droite en a d'après le théorème de la limite monotone. Il en découle que : $\quad f(x) = \dfrac{f(x) - f(a)}{x - a} \times (x - a) + f(a) \quad \underset{x \to a^\pm}{\longrightarrow} f(a), \quad$ i.e. que f est continue en a.

Théorème (Inégalité de Jensen) Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe. Pour tous $x_1, \ldots, x_n \in I$ et $\lambda_1, \ldots, \lambda_n \in [0,1]$ pour lesquels $\lambda_1 + \ldots + \lambda_n = 1$: $\quad f\left(\sum_{k=1}^n \lambda_k x_k\right) \leqslant \sum_{k=1}^n \lambda_k f(x_k).$

Démonstration Par récurrence sur n. Pour $n = 2$, le résultat est ni plus ni moins que la définition de la convexité.

- **Initialisation :** Rien à démontrer pour $n = 1$.
- **Hérédité :** Soit $n \in \mathbb{N}^*$. Faisons l'hypothèse que pour tous $x_1, \ldots, x_n \in I$ et $\lambda_1, \ldots, \lambda_n \in [0,1]$ pour lesquels $\lambda_1 + \ldots + \lambda_n = 1$: $\quad f\left(\sum_{k=1}^n \lambda_k x_k\right) \leqslant \sum_{k=1}^n \lambda_k f(x_k)$.

  Fixons $x_1, \ldots, x_{n+1} \in I$ et $\lambda_1, \ldots, \lambda_{n+1} \in [0,1]$ pour lesquels $\lambda_1 + \ldots + \lambda_{n+1} = 1$. Si $\lambda_{n+1} = 0$, l'inégalité à prouver au rang $n+1$ découle trivialement de l'hypothèse de récurrence. Supposons donc $\lambda_{n+1} > 0$ et remplaçons la moyenne coefficientée des $n+1$ réels $x_1, \ldots, x_{n+1}$ par une moyenne coefficientée de n réels auxquels nous pourrons appliquer l'hypothèse de récurrence.

  Posons pour cela $\lambda'_n = \lambda_n + \lambda_{n+1} \in [0,1]$ et $x'_n = \dfrac{\lambda_n}{\lambda'_n} x_n + \dfrac{\lambda_{n+1}}{\lambda'_n} x_{n+1} \in I$. Sachant que $\lambda_1 + \ldots + \lambda_{n-1} + \lambda'_n = 1$, appliquons l'hypothèse de récurrence aux n réels $x_1, \ldots, x_{n-1}, x'_n$ et aux n coefficients $\lambda_1, \ldots, \lambda_{n-1}, \lambda'_n$:
  $$f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) = f\big(\lambda_1 x_1 + \ldots + \lambda_{n-1} x_{n-1} + \lambda'_n x'_n\big) \overset{\text{HDR}}{\leqslant} \lambda_1 f(x_1) + \ldots + \lambda_{n-1} f(x_{n-1}) + \lambda'_n f(x'_n).$$
  Pour finir, par convexité de f : $\quad f(x'_n) = f\left(\dfrac{\lambda_n}{\lambda'_n} x_n + \dfrac{\lambda_{n+1}}{\lambda'_n} x_{n+1}\right) \leqslant \dfrac{\lambda_n}{\lambda'_n} f(x_n) + \dfrac{\lambda_{n+1}}{\lambda'_n} f(x_{n+1}), \quad$ donc comme voulu : $\quad f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) \leqslant \sum_{k=1}^{n+1} \lambda_k f(x_k)$.

Exemple Pour tous $x_1, \ldots, x_n > 0$:
$$\underbrace{\frac{n}{\displaystyle\sum_{k=1}^n \frac{1}{x_k}}}_{\text{Moyenne harmonique}} \leqslant \underbrace{\sqrt[n]{\prod_{k=1}^n x_k}}_{\text{Moyenne géométrique}} \leqslant \underbrace{\frac{1}{n}\sum_{k=1}^n x_k}_{\text{Moyenne arithmétique}}.$$
L'inégalité de droite s'appelle l'*inégalité arithmético-géométrique*.

Démonstration Comme $\displaystyle\sum_{k=1}^n \frac{1}{n} = 1$, l'inégalité de Jensen appliquée à la fonction concave $x \longmapsto \ln x$ s'écrit :
$$\ln\left(\sum_{k=1}^n \frac{1}{n} x_k\right) \geqslant \sum_{k=1}^n \frac{1}{n} \ln x_k = \ln \sqrt[n]{\prod_{k=1}^n x_k},$$
puis par croissance de l'exponentielle : $\quad \dfrac{1}{n}\displaystyle\sum_{k=1}^n x_k \geqslant \sqrt[n]{\prod_{k=1}^n x_k}. \quad$ On obtient l'inégalité de gauche en remplaçant simplement x_k par $\dfrac{1}{x_k}$ pour tout $k \in [\![1, n]\!]$ dans l'inégalité de droite.

En vue de notre dernier exemple, rappelons que par convention : $0^p = 0$ pour tout $p > 0$, ce qui rend la fonction $x \longmapsto x^p$ continue sur $\mathbb{R}_+$ tout entier. Les inégalités de Hölder et Minkowski énoncées ci-dessous dans $\mathbb{R}^n$ n'ont pas beaucoup d'importance en elles-mêmes, mais leur généralisation aux espaces vectoriels de fonctions est au contraire très fructueuse.

Théorème (Inégalités de Hölder et Minkowski)

(i) **Norme p d'un vecteur de $\mathbb{R}^n$:** Soit $p > 1$.
Pour tout $X \in \mathbb{R}^n$, on appelle *norme p de X* le réel positif $\|X\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}$.
Pour tous $X \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$: $\|\lambda X\|_p = |\lambda| . \|X\|_p$ et $\|X\|_p = 0 \iff X = 0_{\mathbb{R}^n}$.

(ii) **Inégalité de Hölder :** Soient $p, q > 1$ deux réels pour lesquels $\frac{1}{p} + \frac{1}{q} = 1$.
Pour tous $X, Y \in \mathbb{R}^n$: $\left|\sum_{k=1}^n x_k y_k\right| \leqslant \|X\|_p \, \|Y\|_p$.

(iii) **Inégalité de Minkowski :** Soit $p > 1$. Pour tous $X, Y \in \mathbb{R}^n$: $\|X + Y\|_p \leqslant \|X\|_p + \|Y\|_p$.

La norme 2 n'est jamais qu'une généralisation à n réels du concept de norme que vous calculez en géométrie du plan et de l'espace à partir des coordonnées dans une base orthonormale.

L'inégalité de Minkowski n'est finalement qu'une autre manière d'appeler l'inégalité triangulaire dans le contexte des normes p. L'inégalité de Hölder est plus mystérieuse au premier abord, mais vous l'avez déjà rencontrée dans le cas particulier où $p = q = 2$: $\left|\sum_{k=1}^n x_k y_k\right| \leqslant \sqrt{\sum_{k=1}^n x_k^2}\,\sqrt{\sum_{k=1}^n y_k^2}$ (*inégalité de Cauchy-Schwarz*).

Démonstration

(i) Pour tous $X \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$: $\|\lambda X\|_p = \left(\sum_{k=1}^n |\lambda x_k|^p\right)^{\frac{1}{p}} = \left(|\lambda|^p \sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} = |\lambda| . \|X\|_p$.

(ii) Montrons d'abord que pour tous $x, y \geqslant 0$: $xy \leqslant \frac{x^p}{p} + \frac{y^q}{q}$. Il nous suffit de montrer le résultat pour $x, y > 0$. Or par concavité de la fonction logarithme appliquée aux réels x^p et y^q, sachant que $\frac{1}{p} + \frac{1}{q} = 1$: $\ln\left(\frac{x^p}{p} + \frac{y^q}{q}\right) \geqslant \frac{1}{p}\ln(x^p) + \frac{1}{q}\ln(y^q) = \ln(xy)$. On conclut par croissance de l'exponentielle.

Montrons maintenant l'inégalité de Hölder. Fixons $X, Y \in \mathbb{R}^n$.

— Supposons dans un premier temps X et Y *unitaires* au sens suivant : $\|X\|_p = \|Y\|_q = 1$. Pour tout $k \in [\![1, n]\!]$: $|x_k y_k| \leqslant \frac{|x_k|^p}{p} + \frac{|y_k|^q}{q}$, donc : $\left|\sum_{k=1}^n x_k y_k\right| \leqslant \sum_{k=1}^n |x_k y_k| \leqslant \frac{\|X\|_p^p}{p} + \frac{\|Y\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1$.

— Si jamais $\|X\|$ est nul : $\sum_{k=1}^n |x_k|^p = 0$, or on somme ici des réels positifs, donc $x_k = 0$ pour tout $k \in [\![1, n]\!]$ et l'inégalité de Hölder est triviale dans ce cas. Même chose si $\|Y\|$ est nul.

— Enfin, si $\|X\| > 0$ et $\|Y\| > 0$, ramenons-nous au cas unitaire grâce à (i) : $\left\|\frac{X}{\|X\|_p}\right\|_p = \frac{\|X\|_p}{\|X\|_p} = 1$ et $\left\|\frac{Y}{\|Y\|_q}\right\|_q = 1$, donc : $\left|\sum_{k=1}^n \frac{x_k}{\|X\|_p} \times \frac{y_k}{\|Y\|_q}\right| \leqslant \left\|\frac{X}{\|X\|_p}\right\|_p \left\|\frac{Y}{\|Y\|_q}\right\|_q = 1$ et c'est fini.

(iii) Pour commencer : $\frac{1}{p} + \frac{1}{q} = 1$, donc $\frac{1}{q} = \frac{p-1}{p}$, donc $(p-1)q = p$. À présent, pour tous $X, Y \in \mathbb{R}^n$:

$$\|X+Y\|_p^p = \sum_{k=1}^n |x_k + y_k|^p = \sum_{k=1}^n |x_k + y_k| \times |x_k + y_k|^{p-1} \leqslant \sum_{k=1}^n |x_k| \times |x_k + y_k|^{p-1} + \sum_{k=1}^n |y_k| \times |x_k + y_k|^{p-1}$$

$$\overset{\text{Hölder}}{\leqslant} \|X\|_p \left(\sum_{k=1}^n |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}} + \|Y\|_p \left(\sum_{k=1}^n |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}}$$

$$= \|X\|_p \left(\sum_{k=1}^n |x_k + y_k|^p\right)^{\frac{1}{q}} + \|Y\|_p \left(\sum_{k=1}^n |x_k + y_k|^p\right)^{\frac{1}{q}} = \|X\|_p \, \|X+Y\|_p^{\frac{p}{q}} + \|Y\|_p \, \|X+Y\|_p^{\frac{p}{q}}.$$

Pour finir, l'inégalité de Minkowski est triviale si $\|X+Y\| = 0$. Dans le cas contraire, divisons par $\|X+Y\|_p^{\frac{p}{q}}$: $\|X\|_p + \|Y\|_p \geqslant \|X+Y\|_p^{p-\frac{p}{q}} = \|X+Y\|_p^{p\left(1-\frac{1}{q}\right)} = \|X+Y\|_p$.