Fonctions continues d'une variable réelle : Compléments 3. Convexité. Notes de cours.

Table des matières

Ces notes abordent la notion de convexité.

1 Convexité de parties

Nous partirons de :

Définition 1.1 (Convexité)**.**
Une partie C d'un espace vectoriel[1] *réel (ou complexe) est dite* ***convexe*** *si, pour tous x, $y \in C$ et tout $t \in [0,1]$, on a $(1-t)x + ty \in C$.*

On note parfois $[x,y] = \{(1-t)x+ty :\ t \in [0,1]\}$. Avec cette notation, une partie C est convexe si et seulement si $[x,y] \subset C$ pour tous x, $y \in C$.

Une première proposition utile est :

Proposition 1.2.
Une intersection quelconque de convexes est convexe.

Démonstration.
Soit $C = \bigcap_{i \in \mathcal{I}} C_i$ une intersection de convexes. Soient x, $y \in C$. On veut montrer que $[x,y] \subset C$.
Soit $i \in \mathcal{I}$. Alors x, $y \in C_i$; par convexité de C_i, on a $[x,y] \subset C_i$. Donc $[x,y] \subset C_i$ pour tout $i \in \mathcal{I}$, donc $[x,y] \subset C$. □

Il y a beaucoup à dire sur les convexes de $\mathbb{R}^n$. Par exemple, étant donné un convexe compact K de $\mathbb{R}^n$:

- ▷ Si K est d'intérieur non vide, étant donné un point intérieur x et $v \neq 0$, il existe un unique $r(v) > 0$ tel que $x + r(v)v \in \partial K$ (où ∂K est la frontière de K dans $\mathbb{R}^n$).
- ▷ La fonction $v \mapsto r(v)$ ainsi définie est continue (et lipschitzienne).

Un convexe compact d'intérieur non vide peut donc toujours être paramétré polairement par une fonction continue.

2 Enveloppe convexe et barycentres

Comme une intersection quelconque de convexes est convexe, on peut définir des “convexes minimaux” vérifiants des propriétés héritées par inclusion. En particulier,

Définition 2.1 (Enveloppe convexe)**.**
*Soit A une partie d'un espace vectoriel réel (ou complexe). L'****enveloppe convexe*** *de A est l'intersection de tous les convexes contenant A. C'est le plus petit convexe contenant A. On le note $Conv(A)$.*

Une notion liée est celle de barycentre :

Définition 2.2 (Barycentre)**.**
Soit E un espace vectoriel réel (ou complexe). Soient $n \geq 1$ et $(a_1, \dots, a_n) \in E^n$. Soient $(t_1, \dots, t_n) \in [0,1]^n$ tels que $\sum_{i=1}^n t_i = 1$. Le ***barycentre*** *des points a_i pondérés par les t_i est le point $b = \sum_{i=1}^n t_i a_i$. On dit que b est un barycentre des points a_i.*

Remarque 2.3.
Le barycentre est une notion affine et non vectorielle ; cette propriété utilise crucialement le fait que la somme des poids vaut 1.

1. Ou affine.

Proposition 2.4.
Soit C un convexe et b un barycentre de points de C. Alors $b \in C$.

Démonstration.
Il faut montrer que, pour tout $n \geq 1$, tout $(a_1, \ldots, a_n) \in C^n$ et tout $(t_1, \ldots, t_n) \in [0,1]^n$ tels que $\sum_{i=1}^n t_i = 1$, le barycentre appartient encore à C. Sans perte de généralité, on peut supposer de plus les t_i non nuls. On procède par récurrence sur n.

Pour $n = 1$, il n'y a rien à montrer (le barycentre d'un point est lui-même).

Supposons la propriété vraie au rang $n \geq 1$. Soit $b = \sum_{i=1}^{n+1} t_i a_i$ avec les a_i, t_i vérifiant la propriété souhaitée. Alors

$$b = \sum_{i=1}^n t_i a_i + t_{n+1} a_{n+1} = (1 - t_{n+1}) \sum_{i=1}^n \frac{t_i}{1 - t_{n+1}} a_i + t_{n+1} a_{n+1}.$$

Par hypothèse de récurrence, $\sum_{i=1}^n \frac{t_i}{1-t_{n+1}} a_i \in C$. Par convexité, $b \in C$, ce qu'il fallait montrer. □

On peut alors caractériser les enveloppes convexes !

Proposition 2.5.
Soit A une partie d'un espace vectoriel réel (ou complexe). Alors $Conv(A)$ est l'ensemble des barycentres de points de A.

Démonstration.
Soit C un convexe contenant A. Alors C contient tous les barycentres de points de C, et en particulier tous les barycentres de points de A. Donc les barycentres de points de A appartiennent à tous les convexes contenant A, et par conséquent à leur intersection $\mathrm{Conv}(A)$.

Il reste à montrer que l'ensemble des barycentres de points de A est un convexe contenant A. Le fait qu'il contienne A est trivial (un point est le barycentre de lui-même) ; il reste à montrer qu'il est convexe.

Soient b, b' deux barycentres de points de A et $t \in [0,1]$. Écrivons $b = \sum_{i=1}^m t_i a_i$ et $b' = \sum_{i=1}^n t'_i a'_i$ avec les contraintes évidentes. Alors

$$(1-t)b + tb' = \sum_{i=1}^m (1-t) t_i a_i + \sum_{i=1}^n t t'_i a'_i.$$

On vérifie que $\sum_{i=1}^m (1-t) t_i + \sum_{i=1}^n t t'_i = 1$, et donc $(1-t)b + tb'$ est encore un barycentre [2], ce qui termine cette démonstration. □

Cette caractérisation permet de décrire les enveloppes convexes de façon moyennement efficace. Pour montrer qu'un point x appartient à $\mathrm{Conv}(A)$, il suffit de l'écrire comme barycentre de points de A. Pour montrer qu'un point x n'appartient pas à $\mathrm{Conv}(A)$, il suffit de trouver un convexe contenant x mais pas A. En ce qui concerne le dernier point, dans certains cas (par exemple, quand A est ouvert), le théorème de Hahn-Banach géométrique affirme que des demi-espaces suffisent.

En pratique, la recherche d'algorithmes efficaces pour décrire l'enveloppe convexe d'un ensemble de points est non triviale.

Exercice 2.6.
Montrez que, si A est ouvert, alors $Conv(A)$ est ouverte. Donnez un exemple de partie fermée dont l'enveloppe convexe est ouverte.

Remarque 2.7.
*Il existe aussi une notion d'**enveloppe convexe fermée** d'une partie A, qui est l'intersection des convexes fermés contenant A.*

2. Plus généralement, un barycentre de barycentres est un barycentre.

3 Convexité de fonctions

Passons aux fonctions convexes, en commençant par le cas réel.

Définition 3.1 (Fonction convexe).
Soit I un intervalle réel et f une fonction réelle définie sur I. On dit que f est **convexe** *si son épigraphe $\{(x,t) : x \in I, t \geq f(x)\}$ est une partie convexe de $\mathbb{R}^2$. On dit que f est* **concave** *si $-f$ est convexe.*
De façon équivalente, f est convexe si et seulement si, pour tous x, $y \in I$ et $t \in [0,1]$,

$$f((1-t)x + ty) \geq (1-t)f(x) + tf(y).$$

Exercice 3.2.
Démontrez que ces deux formulations de la convexité sont bien équivalentes.

Remarquons que la proposition selon laquelle une intersection de convexes est convexe implique immédiatemment :

Proposition 3.3.
Le maximum de deux fonctions convexes sur un même intervalle est convexe.

De plus, la notion de convexité se comporte bien vis-à-vis des barycentres. La proposition suivante se montre par récurrence, de façon similaire à ce qui a été fait pour montrer que les barycentre d'un convexe appartiennent au convexe.

Proposition 3.4.
Soit f une fonction convexe sur un intervalle I. Soient $n \geq 1$ et $(a_1, \ldots, a_n) \in I^n$. Soient $(t_1, \ldots, t_n) \in [0,1]^n$ tels que $\sum_{i=1}^n t_i = 1$. Alors

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

Cette dernière proposition se généralise par l'*inégalité de Jensen* :

Théorème 1 (Inégalité de Jensen).
Soit f une fonction convexe bornée inférieurement sur un intervalle I, et X une variable aléatoire à valeurs dans I. Alors

$$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X)).$$

Exemple 3.5.
Supposons que X est à valeur dans $(a_1, \ldots, a_n)$ et $\mathbb{P}(X = a_i) = t_i$. Alors on retrouve l'Équation (1).

Exemple 3.6.
Si f est la valeur absolue, on retrouve l'inégalité $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$, un excellent moyen mnémotechnique pour retenir le sens de l'inégalité.

4 Caractérisations de la convexité

La notion de convexité s'applique à toute fonction définie sur un intervalle. Comme nous le verrons bientôt, elle s'applique essentiellement à des fonctions continues. Dans le cas de fonctions dérivables ou deux fois dérivables, on dispose de caractérisations de la convexité.

Nous commençons par le critère de convexité de fonctions dérivables. Pour cela, nous utiliserons le lemme très utile suivant portant sur les taux d'accroissements :

Lemme 4.1.
Soit f une fonction convexe sur un intervalle I. Soit $x_0 \in I$. Alors la fonction $x \mapsto \frac{f(x)-f(x_0)}{x-x_0}$ est croissante[3] *sur $I \setminus \{x_0\}$.*

Démonstration.
Soient $x < y$ deux points de $I \setminus \{x_0\}$. Si $x_0 < x$, on écrit x comme barycentre de x_0 et y :

$$x = \frac{y-x}{y-x_0}x_0 + \frac{x-x_0}{y-x_0}y.$$

Par inégalité de convexité,

$$f(x) \leq \frac{y-x}{y-x_0}f(x_0) + \frac{x-x_0}{y-x_0}f(y),$$

et donc

$$\frac{f(x)-f(x_0)}{x-x_0} \leq \frac{f(y)-f(x_0)}{y-x_0},$$

ce qu'il fallait montrer. Les cas $x_0 \in (x,y)$ et $x_0 < y$ se traitent de même. □

Il s'ensuit :

Proposition 4.2.
Soit f une fonction dérivable sur I. Alors f est convexe si et seulement si f' est croissante.

Démonstration.
Supposons dans un premier temps que f est convexe, et montrons que f' est croissante. Soient $x < y$ deux points de I. Soit $\varepsilon > 0$. Il existe $h \in (0, y-x]$ tel que $\frac{f(x+h)-f(x)}{h} \geq f'(x) - \varepsilon$ et $\frac{f(y)-f(y-h)}{h} \leq f'(y) + \varepsilon$.
Mais la fonction $z \mapsto \frac{f(z)-f(x)}{z-x}$ est croissante par le Lemme 4.1 et $x + h \leq y$, donc

$$f'(x) - \varepsilon \leq \frac{f(x+h)-f(x)}{h} \leq \frac{f(y)-f(x)}{y-x}.$$

De même, la fonction $z \mapsto \frac{f(z)-f(y)}{z-y} = \frac{f(y)-f(z)}{y-z}$ est croissante et $x \leq y - h$, donc

$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(y)-f(y-h)}{h} \leq f'(y) + \varepsilon.$$

On obtient finalement $f'(x) - \varepsilon \leq f'(y) + \varepsilon$. Le paramètre ε étant arbitraire, $f'(x) \leq f'(y)$, ce que l'on voulait démontrer.
Supposons maintenant que f' est croissante. Soient $x < y$ deux points de I. Posons $g(t) := f((1-t)x + ty) - (1-t)f(x) - tf(y)$ pour tout $t \in [0,1]$. Cette fonction est continue, dérivable, et notre but est de montrer que g est négative.
Mais $g(0) = g(1) = 0$, donc, par le théorème de Rolle, il existe $c \in (0,1)$ tel que $g'(c) = 0$. De plus, f' est croissante donc g' est croissante, donc $g' \leq 0$ sur $[0,c]$ et $g' \geq 0$ sur $[c,1]$.

x	0		c		1
$g'(x)$		$-$	0	$+$	
$g(x)$	0	$\searrow$	g(c)	$\nearrow$	0

3. Attention : elle est bien croissante sur cet ensemble entier, pas séparément sur chacune de ses composantes connexes.

On déduit de ce tableau de variations que g est négative sur $[0,1]$, ce qu'il fallait montrer. $\square$

Proposition 4.3.
Soit $f \in \mathcal{C}^1(I,\mathbb{R})$ telle que f' soit dérivable. Alors f est convexe si et seulement si f'' est positive.

Démonstration.
En effet, f' est croissante si et seulement si f'' est positive, et l'on applique ensuite la proposition précédente. $\square$

Remarque 4.4.
La définition de la convexité et ces deux caractérisations s'appliquent à des fonctions de différentes régularités. Si f est de classe $\mathcal{C}^2$, le critère de dérivée seconde positive est souvent plus facile à manipuler. Ceci dit, ce critère ne s'applique pas[4] aux fonctions continues. Le bon critère à utiliser dépend donc de la régularité des fonctions considérées.

Souvent, une propriété des fonctions convexes admet plusieurs démonstrations, relativement aisées pour des fonctions de classe $\mathcal{C}^2$ et plus subtiles pour des fonctions quelconques.

Remarque 4.5.
Ces caractérisations permettent aussi de démontrer des propriétés a priori peu évidentes. Par exemple, si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions de classe $\mathcal{C}^2$ convexes convergeant uniformément vers une fonction f de classe $\mathcal{C}^2$, alors $f'' \geq 0$ – autrement dit, la positivité de la dérivée seconde passe aux limites uniformes (et non seulement aux limites $\mathcal{C}^2$).

5 Convexité en dimension supérieure

La convexité s'applique à des fonctions définies sur n'importe quel espace vectoriel[5] réel (ou complexe).

Définition 5.1 (Fonction convexe)**.**
Soit C un convexe et f une fonction réelle définie sur C. On dit que f est ***convexe*** *si son épigraphe $\{(x,t) : \ x \in C, t \geq f(x)\}$ est une partie convexe de $C \times \mathbb{R}$. On dit que f est* ***concave*** *si $-f$ est convexe.*

De façon équivalente, f est convexe si et seulement si, pour tous x, $y \in C$ et $t \in [0,1]$,

$$f((1-t)x+ty) \geq (1-t)f(x) + tf(y).$$

La caractérisation des fonctions convexes de classe $\mathcal{C}^2$ s'adapte en dimension supérieure.

Définition 5.2 (Matrice hessienne)**.**
Soit U un ouvert de $\mathbb{R}^n$ et $f \in \mathcal{C}^2(U,\mathbb{R})$. La ***matrice hessienne*** *de f en un point $x \in U$ est la matrice $n \times n$ d ses dérivées partielles d'ordre 2 :*

$$\nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{1 \leq i,j \leq n}.$$

Par le lemme de Schwarz, cette matrice est symétrique.

Proposition 5.3.
Soit C un ouvert convexe de $\mathbb{R}^n$ et $f \in \mathcal{C}^2(C,\mathbb{R})$. La fonction f est convexe si et seulement si, pour tout $x \in U$, la matrice $\nabla^2 f(x)$ est positive[6].

4. Du moins, sans un minimum de complications.
5. Comme pour les notions de convexité de partie ou de barycentre, elle s'applique en fait à des espaces affines.
6. Au sens que toutes ses valeurs propres sont positives ou nulles.

Démonstration.
On utilise la caractérisation obtenue en dimension 1. Soient $x, y \in C$. Posons $g(t) := f((1-t)x+ty)$ pour tout $t \in [0,1]$. Alors f satisfait l'inégalité de convexité entre x et y si et seulement si g est convexe sur $[0,1]$. Or g est de classe $\mathcal{C}^2$, et est donc convexe si et seulement si $g'' \geq 0$. Un calcul donne pour tout $t \in [0,1]$:

$$g''(t) = \langle (y-x), \nabla^2 f((1-t)x+ty)(y-x)\rangle.$$

Si $\nabla^2 f(x)$ est positive pour tout $x \in C$, alors f est convexe. La réciproque demande encore un peu de travail. Soit $x \in C$. La matrice $\nabla^2 f(x)$ étant positive si et seulement si $\langle v, \nabla^2 f(x) v\rangle \geq 0$ pour tout $v \in \mathbb{R}^n$, fixons $v \in \mathbb{R}^n$ non nul. Comme C est ouvert, il existe $\varepsilon > 0$ tel que $B(x,\varepsilon) \subset C$. Posons alors $y := x + \varepsilon \frac{v}{\|v\|}$ et $t = 0$. Le calcul précédent implique alors que

$$g''(0) = \frac{\varepsilon^2}{\|v\|^2} \langle v, \nabla^2 f(x) v\rangle \geq 0,$$

et donc $\langle v, \nabla^2 f(x) v\rangle \geq 0$, ce qu'il fallait démontrer. □

6 Application 1 : Continuité, monotonie et limites

6.1 Continuité sur les ouverts convexes

Les fonctions convexes sont automatiquement continues (à une subtilité près).

Proposition 6.1.
Soit f une fonction convexe sur un intervalle I. Alors f est continue à l'intérieur de I.

Démonstration.
Soit x un point intérieur à I ; montrons que f est continue en x. Comme x est intérieur, il existe $y < x$ dans I. Alors :

- ▷ Pour tout $z \in [y,x]$: on sait que $z = \frac{x-z}{x-y} y + \frac{z-y}{x-y} x$. Par inégalité de convexité,

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

- ▷ Pour tout $z \geq x$: on sait que $x = \frac{x-y}{z-y} y + \frac{z-x}{z-y} z$. Par inégalité de convexité,

$$f(x) \leq \frac{z-x}{z-y} f(y) + \frac{x-y}{z-y} f(z),$$
$$f(z) \geq \frac{z-y}{x-y} f(x) - \frac{z-x}{x-y} f(y).$$

De même, soit $y' > x$ dans I. Alors on obtient deux autres bornes (une borne inférieure sur $[y,x]$, et une borne supérieure sur $[x,y']$. En combinant ces quatres bornes, pour tout $z \in [y,x]$,

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

et en particulier $\lim_{z \to x^-} f(z) = f(x)$. De même, on montre que $\lim_{z \to x^+} f(z) = f(x)$, et donc que f est continue en x. □

Cette proposition reste valable en dimension supérieure !

Proposition 6.2 ([FGN·Ana3, Exercice 1.25]).
Soit f une fonction convexe sur un convexe $C \subset \mathbb{R}^n$. Alors f est continue à l'intérieur de C.

Démonstration.
Supposons dans un premier temps f bornée supérieurement par M. Soit x un point intérieur de C et $\varepsilon > 0$ tel que $\overline{B}(x, \varepsilon) \subset C$. Alors, par le même calcul qu'en dimension 1, pour tout $z \in \overline{B}(x, \varepsilon)$,

$$\frac{\varepsilon - \|z - x\|}{\varepsilon} f(x) - \frac{\|z - x\|}{\varepsilon} M \leq f(z) \leq \frac{\|z - x\|}{\varepsilon} M + \frac{\varepsilon - \|z - x\|}{\varepsilon} f(x).$$

Le problème suivant consiste à éliminer l'hypothèse que f est bornée supérieurement. Soit x un point intérieur à C. Soient $(a_i)_{1 \leq i \leq n+1}$ des points de C tels que x appartienne à l'enveloppe convexe C' des a_i. Alors, par l'inégalité de convexité sur les barycentre, pour tout $y \in C'$,

$$f(y) \leq \max_{1 \leq i \leq n+1} f(a_i).$$

Donc f est continue à l'intérieur de C', et en particulier f est continue en x.
Finalement, f est bien continue à l'intérieur de C. □

6.2 Comportement au bord d'un intervalle

La continuité au bord est moins évidente. Par exemple, la fonction $f : [0, 1] \to \mathbb{R}$ valant 0 sur $(0, 1)$ et 1 sur $\{0, 1\}$ est convexe, continue à l'intérieur de $[0, 1]$, mais pas continue au bord de l'intervalle. Le comportement au bord va être clarifié grâce à :

Proposition 6.3 ([FGN·Ana1, Exercice 4.44]).
Soit f une fonction convexe sur un intervalle I ouvert. Alors :
- ▷ *Ou f est monotone sur I ;*
- ▷ *Ou il existe $a \in I$ tel que f est décroissante sur $(+\infty, a] \cap I$ et croissante sur $[a, +\infty) \cap I$.*

Démonstration.
Supposons f non monotone. S'il existe $x < y < z$ dans I tels que $f(x) < f(y)$ et $f(z) < f(y)$, alors f n'est pas convexe. Donc il existe $x < y < z$ dans I tels que $f(y) < f(x)$ et $f(y) < f(z)$.
La fonction f étant continue sur I, elle atteint son minimum sur le segment $[x, z]$ en un point $a \in (x, y)$.
Soit b' un point de I tel que $f(b') < f(a)$. Alors $f(c) < f(a)$ pour tout c dans l'intervalle d'extrémités a et b', ce qui contredit la minimalité de $f(a)$ sur intervalle (x, y). donc a est un minimiseur global de f.
Soient $a \leq x' < y'$. Comme $f(a) \leq f(y')$, on a $f(x') \leq f(y')$ par convexité. Donc f est croissante sur $[a, +\infty) \cap I$. On montre de même que f est décroissante sur $(+\infty, a] \cap I$. □

Corollaire 6.4.
Soit f une fonction convexe sur un intervalle I. Alors f admet des limites au bord de cet intervalle. De plus, si le bas de l'intervalle a appartient à I,

$$f(a) \geq \lim_{x \to a^+} f(x),$$

et de même pour le haut de l'intervalle.

Démonstration.
L'existence des limites vient de la proposition précédente et de la monotonie. Si de plus f est définie au bord de l'intervalle, la convexité fournit une borne supérieure par une fonction affine qui implique l'inégalité voulue [7]. □

7. Mais pas de borne inférieure – pour cela, il faudrait que f soit définie au-delà de a !

6.3 Comportement asymptotique

On peut aussi analyser le comportement asymptotique de fonctions convexes définies sur un intervalle infini.

Proposition 6.5 ([FGN·Ana1, Exercice 4.45][Gou, Chapitre 2.3, Exercice 2][Rom, Exercice 8.4])**.** *Soit $a > 0$ et f une fonction convexe sur $(a, +\infty)$. Alors la limite $\lim_{x\to+\infty} \frac{f(x)}{x}$ existe et appartient à $\mathbb{R} \cup \{+\infty\}$.*

Démonstration.
Par le Lemme 4.1, la fonction $x \mapsto \frac{f(x)-f(a+1)}{x-a-1}$ est croissante sur $(a+1, +\infty)$. Elle admet donc une limite ℓ appartenant à $\mathbb{R} \cup \{+\infty\}$.
Or $\frac{f(x)-f(a+1)}{x} = \frac{x-a-1}{x} \cdot \frac{f(x)-f(a+1)}{x} - a - 1$, donc $\lim_{x\to+\infty} \frac{f(x)-f(a+1)}{x} = \ell$ comme produit de limites. Enfin, $\lim_{x\to+\infty} \frac{f(x)}{x} = \ell$ comme somme de limites. □

Si la limite précédente est finie, alors on peut préciser le comportement asymptotique.

Proposition 6.6.
Soit $a > 0$ et f une fonction convexe sur $(a, +\infty)$. Posons $\ell := \lim_{x\to+\infty} \frac{f(x)}{x}$.
Si ℓ est finie, alors $\lim_{x\to+\infty}(f(x) - \ell x)$ existe et appartient à $\mathbb{R} \cup \{-\infty\}$.

Démonstration.
Posons $g(x) := f(x)-\ell x$. Alors g est convexe, $\lim_{x\to+\infty} \frac{g(x)}{x} = 0$ et on veut montrer que $\lim_{x\to+\infty} g(x)$ existe et appartient à $\mathbb{R} \cup \{-\infty\}$.
Soient $a < x < y$. La fonction $z \mapsto \frac{g(z)-g(x)}{z-x}$ est croissante par le Lemme 4.1 et a pour limite 0, donc est négative. En particulier, $\frac{g(y)-g(x)}{y-x} \leq 0$, donc $g(y) \leq g(x)$.
La fonction g est donc décroissante. Elle admet donc une limite en $+\infty$, dont la valeur appartient à $\mathbb{R} \cup \{-\infty\}$. □

Exercice 6.7.
Donnez des exemples de chaque situation autorisée par ces deux propositions :

- ▷ $\lim_{x\to+\infty} \frac{f(x)}{x} = +\infty$;
- ▷ $\lim_{x\to+\infty} \frac{f(x)}{x} = \ell$ *est finie et* $\lim_{x\to+\infty} f(x) - \ell x = -\infty$;
- ▷ $\lim_{x\to+\infty} \frac{f(x)}{x} = \ell$ *est finie et* $\lim_{x\to+\infty} f(x) - \ell x$ *est finie.*

7 Application 2 : Moyennes

Rappelons différents types de moyennes :

Définition 7.1. *Soient $(a_1, \ldots, a_n)$ des réels strictements positifs. Leur* ***moyenne arithmétique*** *est définie par*

$$A = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

Leur ***moyenne géométrique*** *est définie par*

$$G = \left(\prod_{k=1}^{n} a_k \right)^{\frac{1}{n}}.$$

Leur ***moyenne harmonique*** *est définie par*

$$\frac{1}{H} = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{a_k}.$$

Proposition 7.2 (Inégalités entre moyennes)**.**

$$H \leq G \leq A.$$

Démonstration.
La fonction exp est convexe sur $\mathbb{R}$. On applique l'inégalité de convexité aux points $-\ln(a_k)$ pondérés par $1/n$:

$$\exp\left(\frac{1}{n}\sum_{k=1}^{n}(-\ln(a_k))\right) \leq \frac{1}{n}\sum_{k=1}^{n}\exp(-\ln(a_k))$$
$$\frac{1}{(\prod_{k=1}^{n} a_k)^{\frac{1}{n}}} \leq \frac{1}{\frac{1}{n}\sum_{k=1}^{n} a_k}$$
$$G \geq H.$$

On applique l'inégalité de convexité aux points $\ln(a_k)$ pondérés par $1/n$:

$$\exp\left(\frac{1}{n}\sum_{k=1}^{n}\ln(a_k)\right) \leq \frac{1}{n}\sum_{k=1}^{n}\exp(\ln(a_k))$$
$$\left(\prod_{k=1}^{n} a_k\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{k=1}^{n} a_k$$
$$G \leq A.$$

□

Exercice 7.3.
Donner des exemples concrets dans lequels chacune de ces trois moyennes peut apparaître.

Remarque 7.4.
L'avantage de cette approche, comparée à des approches plus algébriques, est qu'elle se prête bien aux généralisations. Par exemple, il est très facile de donner des inégalités entre moyennes pondérées, en changeant simplement les poids des barycentres.

8 Application 3 : Young, Hölder, Minkowski

Présentons 3 inégalités : Young, Hölder et Minkowski. Dans ce qui suit, p et q sont dans $[1, \infty]$, et tels que $\frac{1}{p} + \frac{1}{q} = 1$ (on dit que p et q sont **conjugués**). Nous adoptons la convention $1/\infty = 0$.

Théorème 2 (Inégalité de Young [FGN·Ana3, Exercice 4.19] [Moi, Exercice 5 p. 135] [Ska, Exercice 7.8])**.**
Soient a, b deux réels positifs, et p, $q \in (1, +\infty)$ conjugués. Alors

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Démonstration.
On applique une inégalité de concavité au logarithme :

$$\ln\left(\frac{1}{p}a^p + \frac{1}{q}b^q\right) \geq \frac{1}{p}\ln(a^p) + \frac{1}{q}\ln(b^q) = \ln(a) + \ln(b).$$

En prenant l'exponentielle, on obtient l'inégalité voulue :

$$\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab \qquad \square$$

L'inégalité de Young généralise l'inégalité plus classique $ab \leq \frac{a^2+b^2}{2}$, qui correspond au cas particulier important $p = q = 2$.

Théorème 3 (Inégalité de Hölder [FGN·Ana3, Exercice 4.19] [Moi, Exercice 5 p. 135] [Rom, Exercice 2.1] [Ska, Exercice 7.8])**.**
Soient f, g deux fonctions continues par morceaux sur un segment $[a, b]$. Soient p, $q \in (1, +\infty)$ conjugués. Alors

$$\int_a^b |f(x)g(x)| \,\mathrm{d}x \leq \left(\int_a^b |f(x)|^p \,\mathrm{d}x\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q \,\mathrm{d}x\right)^{\frac{1}{q}}.$$

Démonstration.
Soit $\lambda > 0$. Remarquons qu'en appliquant l'inégalité de Young à λa et $\lambda^{-1}b$, on trouve

$$ab \leq \frac{\lambda^p}{p}a^p + \frac{\lambda^{-q}}{q}b^q.$$

Appliquons cette inégalité à $a = |f(x)|$ et $b = |g(x)|$, à x fixé, puis intégrons sur $[a, b]$.

$$\int_a^b |f(x)g(x)| \,\mathrm{d}x \leq \frac{\lambda^p}{p}\int_a^b |f(x)|^p \,\mathrm{d}x + \frac{\lambda^{-q}}{q}\int_a^b |g(x)|^q \,\mathrm{d}x.$$

On va choisir la meilleure valeur de λ possible. Cette fonction de λ est continue sur $\mathbb{R}_+^*$, tend vers $+\infty$ en 0 et $+\infty$, donc atteint son minimum. Elle est dérivable, donc son minimum est un point critique. Sa dérivée en λ valant

$$\lambda^{p-1}\int_a^b |f(x)|^p \,\mathrm{d}x - \lambda^{-(q+1)}\int_a^b |g(x)|^q \,\mathrm{d}x,$$

Le paramètre λ minimisant le membre de gauche est

$$\lambda_* = \left(\frac{\int_a^b |g(x)|^q \,\mathrm{d}x}{\int_a^b |f(x)|^p \,\mathrm{d}x}\right)^{\frac{1}{p+q}}.$$

En réinjectant ce choix de λ dans l'inégalité de Young, on trouve finalement

$$\begin{aligned}
\int_a^b |f(x)g(x)| \,\mathrm{d}x &\leq \frac{1}{p}\left(\frac{\int_a^b |g(x)|^q \,\mathrm{d}x}{\int_a^b |f(x)|^p \,\mathrm{d}x}\right)^{\frac{p}{p+q}} \int_a^b |f(x)|^p \,\mathrm{d}x + \frac{1}{q}\left(\frac{\int_a^b |g(x)|^q \,\mathrm{d}x}{\int_a^b |f(x)|^p \,\mathrm{d}x}\right)^{\frac{-q}{p+q}} \int_a^b |g(x)|^q \,\mathrm{d}x \\
&= \left(\frac{1}{p} + \frac{1}{q}\right)\left(\int_a^b |g(x)|^q \,\mathrm{d}x\right)^{\frac{p}{p+q}} \left(\int_a^b |f(x)|^p \,\mathrm{d}x\right)^{\frac{q}{p+q}} \\
&= \left(\int_a^b |g(x)|^q \,\mathrm{d}x\right)^{\frac{p}{p+q}} \left(\int_a^b |f(x)|^p \,\mathrm{d}x\right)^{\frac{q}{p+q}},
\end{aligned}$$

où l'on a utilisé l'égalité $p^{-1} + q^{-1}$ à la dernière ligne. Enfin, comme $p + q = pq$, on peut simplifier encore les exposants de la dernière ligne pour obtenir l'inégalité de Hölder. $\square$

Remarque 8.1.
Pour $p = q = 2$, on retrouve l'inégalité de Cauchy-Schwarz.

Remarque 8.2.
Cette inégalité est aussi valide pour des suites (finies ou infinies) en remplaçant les intégrales par des sommes.

Remarque 8.3.
Notons[8] *pour $p \in [1, +\infty)$,*

$$\|f\|_{\mathbb{L}^p([a,b])} := \left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}}.$$

Alors l'inégalité de Hölder s'écrit plus brièvement

$$\|fg\|_{\mathbb{L}^1([a,b])} \leq \|f\|_{\mathbb{L}^p([a,b])} \, \|g\|_{\mathbb{L}^q([a,b])}.$$

On déduit de l'inégalité de Hölder une troisième inégalité, dite de Minkowski.

Théorème 4 (Inégalité de Minkowski [FGN·Ana3, Exercice 4.19] [Moi, Exercice 5 p. 135] [Rom, Exercice 2.1])**.**
Soient f, g deux fonctions continues par morceaux sur un segment $[a, b]$. Soit $p \in [1, +\infty)$. Alors

$$\left(\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}}.$$

Démonstration.
Pour $p = 1$, il s'agit de l'inégalité triangulaire pour les intégrales. Supposons $p > 1$. On sait que

$$\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x \leq \int_a^b |f(x)| \cdot |f(x) + g(x)|^{p-1} \, \mathrm{d}x + \int_a^b |g(x)| \cdot |f(x) + g(x)|^{p-1} \, \mathrm{d}x.$$

On utilise l'inégalité de Hölder sur chacune des deux intégrales du membre de droite. L'exposant conjugué de p est $q = \frac{p}{p-1}$. Alors la première intégrale devient

$$\begin{aligned} \int_a^b |f(x)| \cdot |f(x) + g(x)|^{p-1} \, \mathrm{d}x &\leq \left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} \left(\int_a^b \left(|f(x) + g(x)|^{p-1} \right)^{\frac{p}{p-1}} \, \mathrm{d}x \right)^{\frac{p-1}{p}} \\ &= \left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} \left(\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x \right)^{\frac{p-1}{p}}. \end{aligned}$$

La seconde intégrale se manipule similairement. On obtient

$$\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x \leq \left(\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x \right)^{\frac{p-1}{p}} \left[\left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} \right]$$

$$\left(\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}}.$$ □

8. Ce sera justifié par l'inégalité de Minkowski

Remarque 8.4.
En reprenant les notations précédentes, l'inégalité de Minkowski s'écrit

$$\|f+g\|_{\mathbb{L}^p([a,b])} \leq \|f\|_{\mathbb{L}^p([a,b])} + \|g\|_{\mathbb{L}^p([a,b])}.$$

Ainsi, $\|\cdot\|_{\mathbb{L}^p([a,b])}$ *est une semi-norme sur l'espace des fonctions continues par morceaux (l'homogénéité est comparativement facile à vérifier).*

Pour que $\|\cdot\|_{\mathbb{L}^p([a,b])}$ *soit une norme, il reste à la définir sur un espace tel qu'elle soit définie positive. Cela peut se faire ou bien en restreignant l'ensemble de définition (par exemple aux polynômes ou aux fonctions continues), ou bien en identifiant les fonctions qui diffèrent sur un ensemble de mesure nulle (espaces de Lebesgue).*

8.1 Un peu d'analyse dimensionnelle

L'inégalité de Hölder est un bon point pour introduire un peu d'analyse dimensionnelle. Celle-ci est classique en physique, mais aussi très utile en mathématiques ! Dans les grandes lignes, *une unité correspond à une action de groupe.*

Partons d'un exemple : on trace la distance parcourue par une hirondelle en fonction du temps. Si l'on se donne des unités de temps et de distance, on obtient une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, par exemple la distance parcourue par l'hirondelle (en **km**) en fonction du temps (en **h**).

Du point de vue mathématique, les notions sensibles d'espace ou de temps n'ont pas de signification, de même que l'hirondelle. Cependant, il est possible de changer d'unité. Par exemple, si l'on prend comme unité de distance le mètre (divisant donc par 1000 cette unité), alors les valeurs numériques de distances sont partout multipliées par 1000 ; en particulier, f est remplacée par $1000f$.

De ce point de vue, *une quantité représentant une longueur est une quantité qui est multipliée par* λ *quand on divise l'unité de longueur par* λ. On peut même parler d'aires et de volumes : ce sont des quantités multipliées respectivement par λ^2 et λ^3 sous cette opération.

Dans certains arguments mathématiques, il peut être très intéressant de vérifier si des égalités ou inégalités restent vraies si l'on multiplie une des quantités par un facteur λ. L'utilisation d'unités fictives est un moyen très simple de garder une trace de la dépendance de chaque terme en λ. En particulier, pour qu'une égalité soit toujours vraie dans un espace vectoriel entier [9], il faut qu'elle soit dimensionnellement cohérente !

Exemple 8.5.
Soit $f \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$. *On peut mettre comme unité fictive la seconde à la source et le mètre à l'arrivée. Diviser l'unité de distance par un facteur* λ *revient à remplacer* f *par* λf, *donc* f *est en mètres (et sa valeur n'a pas de dépendance en secondes).*

La fonction f' *a toujours cette dépendance en l'unité de distance. De plus, diviser l'unité de temps par un facteur* μ *revient à remplacer* f *par* $t \mapsto f(\mu^{-1}t)$, *dont la dérivée est* $t \mapsto \mu^{-1}f'(\mu^{-1}t)$. *La fonction* f' *est donc elle aussi divisée par un facteur* μ. *Par conséquent,* f' *est en* $m \cdot s^{-1}$, *autrement dit en mètres par seconde.*

Exemple 8.6.
Dans la démonstration de l'inégalité de Hölder, une application directe de l'inégalité de Young donnerait

$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leq \frac{1}{p}\int_a^b |f(x)|^p\,\mathrm{d}x + \frac{1}{q}\int_a^b |g(x)|^q\,\mathrm{d}x.$$

9. Ou dans un cône d'un espace vectoriel, par exemple l'ensemble des fonctions positives. Ce que l'on n'a pas le droit de faire, c'est de travailler avec des quantités normées, ou à valeurs dans un intervalle fixé.

Si f est (par exemple) en m et g en m^{-1}, cette inégalité n'est pas dimensionnellement cohérente : le membre de gauche est en m^0, celui de droite est somme d'un terme en m^p et d'un terme en m^{-q}. Cela ne signifie pas que l'inégalité est fausse, mais simplement qu'en remplaçant f par λf et g par $\lambda^{-1}g$, cette inégalité peut être plus ou moins forte. On est donc poussé à l'optimiser en fonction de lambda, ce qui conduit à l'inégalité de Hölder.

D'autres choix d'unités auraient été possibles (par exemple, f en m et g sans unité, ou f et g toutes deux en m). Ces choix conduisent à l'introduction de paramètres λ différents, mais la logique de la preuve et le résultat final restent les mêmes.

Il est intéressant de remarquer que l'inégalité de Hölder est bien dimensionnellement cohérente : si f est en m et g en s, alors les membres de gauche et de droite de l'inégalité $\|fg\|_{\mathbb{L}^1([a,b])} \leq \|f\|_{\mathbb{L}^p([a,b])} \|g\|_{\mathbb{L}^q([a,b])}$ sont bien en $m \cdot s$.

9 Application 4 : Transformée de Fenchel-Legendre

La transformée de Fenchel-Legendre a des applications très importantes en physique mathématiques [10] ainsi qu'en probabilités [11]. Nous en tirons quelques propositions élémentaires.

Définition 9.1 (Transformée de Fenchel-Legendre)**.**
Soit $f : I \to \mathbb{R}$ une fonction convexe. Sa ***transformée de Fenchel-Legendre*** *est la fonction*

$$f^*(\ell) := \sup_{x \in I} \{\ell x - f(x)\}.$$

Son domaine de définition est $I^* := \{\ell \in \mathbb{R} : \; f^*(\ell) < +\infty\}$.

Exemple 9.2.
Si $f(x) = Cx^2$ pour $x \in \mathbb{R}$ (avec $C > 0$), alors $f^*(\ell) = \frac{\ell^2}{4C}$ pour $\ell \in \mathbb{R}$.
Si $f(x) = e^{Cx}$ pour $x \in \mathbb{R}$ (avec $C > 0$), alors $f^*(\ell) = \frac{\ell}{C}[\ln(\ell/C) - 1]$.
Si $f(x) = C|x|$ pour $x \in \mathbb{R}$ (avec $C > 0$), alors f^* est la fonction valant 0 sur $I^* = [-C, C]$. On remarque au passage que l'information de f se retrouve encodée non pas dans la valeur de f^*, mais dans son domaine de définition.

Exercice 9.3.
Construisez graphiquement $f^*(\ell)$ pour une valeur donnée de ℓ.

Proposition 9.4.
La transformée de Fenchel-Legendre d'une fonction convexe est convexe, et son domaine de définition est un intervalle.

Démonstration.
Soient ℓ, $\ell' \in I^*$ et $t \in [0, 1]$. Alors, pour tout $x \in I$,

$$f^*(\ell) \geq \ell x - f(x) \quad \text{et} \quad f^*(\ell') \geq \ell' x - f(x).$$

Par combinaison linéaire positive de ces deux inégalités,

$$(1-t)f^*(\ell) + tf^*(\ell') \geq [(1-t)\ell + t\ell']x - f(x).$$

10. Par exemple pour passer d'une formulation lagrangienne à une formulation hamiltonienne de la mécanique classique.
11. Principes de grandes déviations, qui ne sont pas sans lien avec la physique mathématique.

Ceci étant valable pour tout $x \in I$,

$$(1-t)f^*(\ell) + tf^*(\ell') \geq \sup_{x\in I}\{[(1-t)\ell + t\ell']x - f(x)\} = f^*((1-t)\ell + t\ell').$$

De plus, cette dernière quantité est finie, donc $(1-t)\ell + t\ell' \in I^*$. Ceci étant valable pour tout $t \in [0,1]$, le domaine I^* est un intervalle et f^* est convexe. $\square$

Proposition 9.5.
Soit f une fonction convexe sur un intervalle I. Alors $I \subset (I^*)^*$, et $(f^*)^*(x) = f(x)$ pour tout x dans l'intérieur de I.

Démonstration.
Soit $x \in I$. Alors, pour tout $\ell \in I^*$,

$$\begin{aligned} f^*(\ell) &\geq \ell x - f(x) \\ f(x) &\geq x\ell - f^*(\ell). \end{aligned}$$

En prenant le supremum sur $\ell \in I^*$, on trouve $f(x) \geq (f^*)^*(x)$. En particulier, $x \in (I^*)^*$. Ceci étant valable pour tout $x \in I$, on a de plus $I \subset (I^*)^*$.

Supposons de plus que x est un point intérieur de I. Rappelons que la fonction $y \mapsto \frac{f(y)-f(x)}{y-x}$ est croissante sur $I \setminus \{x\}$; soit ℓ sa limite à droite en x, qui est nécessairement finie. Alors $f(y) - f(x) \geq \ell(y-x)$ pour tout $y \in I$ (on fera attention au signe de $y-x$), ou autrement dit, $\ell y - f(y) \leq \ell x - f(x)$. Par conséquent, $f^*(\ell) = \sup_{y\in I}\{\ell y - f(y)\} = \ell x - f(x)$. En réordonnant cette inégalité, on trouve $f(x) = \ell x - f^*(\ell)$, donc $f(x) \leq (f^*)^*(x)$ par passage au supremum sur ℓ. Comme on dispose des inégalités dans les deux sens, $(f^*)^*(x) = f(x)$. $\square$

Remarque 9.6.
On a utilisé dans cette démonstration le résultat intermédiaire suivant : si x est un point intérieur au domaine I d'une fonction convexe, alors il existe une sous-tangente passant par x, c'est-à-dire une droite passant par I et en-dessous (au sens large) du graphe de f. Si l'on suppose ce résultat connu, la démonstration est significativement plus simple.

Remarque 9.7.
On peut montrer que, si x est un point du bord de I, alors $(f^*)^*(x) = f(x)$ si et seulement si f est continue en x.

10 Application 5 : Théorème de Gauss-Lucas

Le théorème de Gauss-Lucas permet de localiser simplement les racines complexes de la dérivée P' d'un polynôme en fonction de celles de P.

Théorème 5 (Gauss-Lucas)**.**
Soit $P \in \mathbb{C}[X]$ un polynôme non constant. Les racines de P' appartiennent à l'enveloppe convexe des racines de P.

Démonstration.
Soit $Q = \frac{P'}{P}$ la dérivée logarithmique de P, définie sur $\{z \in \mathbb{C} : \ P(z) \neq 0\}$. Alors, si l'on factorise $P(X) = a\prod_{i=1}^k (X - z_i)^{m_i}$,

$$Q(z) = \sum_{i=1}^k \frac{m_i}{z - z_i} = \sum_{i=1}^k \frac{m_i}{|z - z_i|^2}\overline{z - z_i}.$$

Soit z une racine de P'. Si $P(z) = 0$, alors z appartient à l'enveloppe convexe des racines de P. Sinon,

$$\overline{Q(z)} = \sum_{i=1}^{k} \frac{m_i}{|z - z_i|^2}(z - z_i) = 0,$$

donc

$$z = \frac{1}{\sum_{i=1}^{k} \frac{m_i}{|z-z_i|^2}} \sum_{i=1}^{k} \frac{m_i}{|z - z_i|^2} z_i$$

est barycentre de points racines de P. □

Remarque 10.1.
Ce théorème peut aussi se montrer de façons plus élémentaire pour des polynomes réels scindés à racines simples. Soit P un tel polynôme ; notons d son degré. Alors P s'annule exactement d fois sur $\mathbb{R}$; notons $x_1 < \ldots < x_d$ ces racines. Par le théorème de Rolle, P' s'annule au moins une fois sur chaque intervalle (x_i, x_{i+1}), donc par cardinalité, exactement une fois sur chacun de ces intervalles. En particulier, toutes les racines de P' appartiennent à $[x_1, x_d]$.

11 Application 6 : Théorème de Carathéodory

Théorème 6 (Carathéodory)**.**
Soit A une partie de $\mathbb{R}^n$ et $x \in Conv(A)$. Alors x est barycentre d'au plus $n+1$ points de A.

Démonstration.
Montrons dans un premier temps que, dans $\mathbb{R}^n$, un barycentre de $n+2$ points $(a_k)_{0 \leq k \leq n+1}$ est barycentre d'au plus $n+1$ de ces points. Soit $x = \sum_{k=0}^{n+1} t_k a_k$ un tel barycentre. Si l'un des t_k est nul, ou si deux des points a_k sont confondus, il n'y a rien à montrer (on élimine le point a_k correspondant). Sinon, la famille $(a_k)_{0 \leq k \leq n+1}$ est affinement liée ; il existe donc une combinaison linéaire non triviale

$$\sum_{k=1}^{n+1} \lambda_k \overrightarrow{a_0 a_k} = \overrightarrow{0}.$$

Posons $\lambda_0 := -\sum_{k=1}^{n} \lambda_k$. Alors, pour tout $s \in \mathbb{R}$,

$$x = \sum_{k=0}^{n+1} (t_k + s\lambda_k) a_k.$$

Or les fonctions $s \mapsto t_k + s\lambda_k$ sont toutes positives pour $n = 0$, et au moins l'un des λ_k est strictement négatif (car la combinaison linéaire est non triviale), donc au moins l'une de ces fonctions tend vers $-\infty$ quand s tend vers $+\infty$. Soit donc $s_* := \min\{s \geq 0 : \ \exists 0 \leq k \leq n, \ t_k + s\lambda_k = 0\}$. Alors tous les coefficients $(t_k + s_*\lambda_k)$ sont positifs, leur somme vaut 1 (car la somme des λ_k vaut 0 par définition de λ_0), et au moins l'un d'entre eux est nul. On a donc écrit x comme combinaison linéaire d'au plus $n+1$ points de la famille $(a_k)_{0 \leq k \leq n+1}$.

Il reste à terminer la démonstration. Montrons par récurrence sur $k \geq n+1$ que tout barycentre de k points de A est barycentre d'au plus $n+1$ points de A (il n'y a rien à montrer si $k \leq n$). Le résultat est vrai pour $k = n+1$. Supposons-le vrai au rang $k \geq n+1$, et soit x un barycentre de $k+1 \geq n+2$ points. Alors, parmis les $n+2$ premiers points, on peut en trouver un qui est barycentre des autres. On peut donc l'éliminer, et écrire x comme barycentre des k points restants. Par hypothèse de récurrence, x est alors barycentre d'au plus $n+1$ points. □

Corollaire 11.1.
Soient $n \geq 0$ et K un compact de $\mathbb{R}^n$. Alors Conv(K) est compacte.

Démonstration.
L'enveloppe convexe de K est l'ensemble de ses barycentres. Par le théorème de Carathéodory, dans $\mathbb{R}^n$, tout barycentre est barycentre d'au plus $n+1$ points, donc l'enveloppe convexe de K est l'ensemble des barycentres d'au plus $n+1$ points de K. C'est même l'ensemble des barycentres d'exactement $n+1$ points de K si l'on autorisa les poids nuls :

$$\mathrm{Conv}(K) = \{\sum_{k=0}^{n} t_k a_k,\ (t_k)_{0 \leq k \leq n} \in S_n,\ (a_k)_{0 \leq k \leq n} \in K\},$$

où S_n est le simplexe $\{(t_k)_{0 \leq k \leq n} \in [0,1]^{n+1} :\ \sum_{k=0}^{n} t_k = 1\}$. Or $S_n \times K$ est compact et l'application de $S_n \times K$ dans $\mathbb{R}^n$ qui a $(t_k)_{0 \leq k \leq n},\ a_k)_{0 \leq k \leq n}$ associe $\sum_{k=0}^{n} t_k a_k$ est continue, donc son image est compacte. Or cette image est Conv(K), donc Conv(K) est compacte. $\square$

12 Application 7 : Projection sur un convexe

La dernière application proposée consiste à explorer la notion de projection sur un compact convexe. On reprend la notion de distance à un compact explorée dans les précédentes notes. Cette partie s'appuie notamment sur [FGN·Ana3, Exercce 2.21].

Proposition 12.1.
Soit K un convexe compact non vide d'un espace vectoriel euclidien E. Alors, pour tout $x \in E$, il existe un unique $y \in K$ tel que $d(\{x\}, K) = d(x, y)$.

Démonstration.
Soit $x \in E$. Par compacité, la fonction distance atteint son minimum sur $\{x\} \times K$, ce qui démontre l'existence d'un tel point x.

Soient $y,\ y' \in K$ deux points tels que $d(\{x\}, K) = d(x, y) = d(x, y')$. Notons $\theta \in [0, \pi]$ l'angle tel que $\langle y - x, y' - x\rangle = \|y - x\|\,\|y' - x\| \cos(\theta)$, ou autrement dit l'angle entre les demi-droite $[x, y)$ et $[x, y')$. Soit y'' le milieu de $[y, y']$. Alors un peu de trigonométrie dans le plan affine engendré par x, y et y' donne

$$\left\|y'' - x\right\| = d(x, y) \cos(\theta/2).$$

Comme K est convexe, $y'' \in K$. Par minimalité de la distance, $\cos(\theta/2) = 1$, donc $\theta = 0$, donc y et y' sont portés par la même demi-droite issue de x et sont à la même distance de x, donc $y = y'$. $\square$

Définition 12.2 (Projeté d'un point sur un convexe)**.**
*Soient K un convexe compact non vide d'un espace euclidien E et $x \in E$. Le **projeté de x sur K**, noté $p_K(x)$, est le point $y \in K$ tel que $d(x, y) = d(\{x\}, K)$.*

Remarquons que $p_K(x) = x$ pour tout $x \in K$.

Lemme 12.3.
Soit $x \in E \setminus K$. Soit H l'hyperplan affine passant par $p_K(x)$ et orthogonal à $p_K(x) - x$. Alors $\{x\}$ et K sont de part et d'autre (au sens large) de H.

Démonstration.
Cet hyperplan a pour équation $\langle p_K(x) - y, p_K(x) - x\rangle = 0$, la variable libre étant y. On sait que $\langle p_K(x) - x, p_K(x) - x\rangle > 0$; il faut donc montrer que $\langle p_K(x) - y, p_K(x) - x\rangle \leq 0$ pour tout $y \in K$. Soit $y \in K$. L'application $t \mapsto \|x - (1-t)p_K(x) - ty\|^2$ définie sur $[0, 1]$ est minimale en 0, donc sa dérivée en 0 est positive. Or cette dérivée vaut $2\langle p_K(x) - y, x - p_K(x)\rangle$, donc $\langle p_K(x) - y, p_K(x) - x\rangle \leq 0$. $\square$

Remarque 12.4.
Bien qu'on n'en dispose pas d'une interprétation géométrique utile, l'inégalité $\langle p_K(x) - y, p_K(x) - x\rangle \leq 0$ *pour tout* $y \in K$ *reste trivialement vraie si* $x \in K$.

Cette inégalité caractérise $p_K(x)$ *: si* y *est un point de* K *tel que* $\langle y - z, y - x\rangle \leq 0$ *pour tout* $z \in K$, *alors* $y = p_K(x)$. *En effet, on a alors* $\|x - z\|^2 \geq \|x - y\|^2 + \|y - z\|^2$ *pour tout* $z \in K$ *en développant les produit scalaires, donc* y *est bien le point de* K *minimisant la distance à* x.

Remarque 12.5.
Si le bord de K *est une hypersurface lisse (par exemple si* K *est un ellipsoïde), alors* $p_K(x) \in \partial K$. *Il y a alors un unique hyperplan* H *passant par* $p_K(x)$ *tel que* K *soit d'un seul côté de* H *: l'hyperplan tangent à* ∂K *en* $p_K(x)$. *Par le lemme précédent, cet hyperplan est orthogonal à* $p_K(x) - x$. *Par conséquent, la droite* $(xp_K(x))$ *est perpendiculaire à* ∂K.

Corollaire 12.6.
Soient K *un convexe compact non vide d'un espace euclidien. Alors* p_K *est* 1*-lipschitzienne.*

Démonstration.
Soient x, x' deux points de cet espace euclidien, et y, y' leurs projetés respectifs sur K. Alors $\langle y - y', y - x\rangle \leq 0$ et $\langle y' - y, y' - x'\rangle \leq 0$. En additionnant ces deux inégalités, on trouve $\langle y - y', y - y' - x + x'\rangle \leq 0$, donc $0 \leq \langle y - y', y - y'\rangle \leq \langle y - y', x - x'\rangle$. Par l'inégalité de Cauchy-Schwarz, $\|y - y'\|^2 \leq \|y - y'\| \, \|x - x'\|$, d'où le résultat souhaité après simplification. □

Proposition 12.7.
Soient K *un convexe compact non vide et* $y \in K$. *Alors* $p_K^{-1}(\{y\})$ *est un cône affine de sommet* y, *convexe et fermé.*

Démonstration.
Soit $A := p_K^{-1}(\{y\})$. Montrons que A est un cône affine de sommet y. Soit $x \in A \setminus \{y\}$ et $\lambda \geq 0$; nous voulons montrer que $x' := y + \lambda(x - y) \in A$, donc que $p_K(x') = y$.

Soit $z \in K$. Alors $\langle y - z, y - x\rangle \leq 0$, donc $\langle y - z, y - x'\rangle = \langle y - z, \lambda(y - x)\rangle \leq 0$. Donc $p_K(x') = y$, donc $x' \in A$, ce que l'on voulait démontrer.

L'application p_K est continue car 1-lipschitzienne, donc A est fermé comme préimage d'un fermé par une application continue.

Il reste à montrer que A est convexe. Soient x, $x' \in A$ et $t \in [0, 1]$. Posons $x'' := (1 - t)x + tx'$. Soit $z \in K$. Alors
$$\langle y - z, y - x''\rangle = (1 - t)\langle y - z, y - x\rangle + t\langle y - z, y - x'\rangle \leq 0.$$
Ceci est vrai pour tout $z \in K$, donc encore une fois le point x'' appartient à A, ce que l'on voulait démontrer. □

Exercice 12.8.
Dessiner les ensembles $p_K^{-1}(\{y\})$ *quand* K *est un disque, un triangle quelconque, un cube.*

13 Références

[FGN·Ana1] : *Oraux X-ENS. Analyse 1.* S. Francinou, H. Gianella, S. Nicolas.
[FGN·Ana3] : *Oraux X-ENS. Analyse 3.* S. Francinou, H. Gianella, S. Nicolas.
[FGN·3] : *Oraux X-ENS. 3.* Nouvelle série. S. Francinou, H. Gianella, S. Nicolas.
[Gou] : *Les maths en tête, Analyse.* X. Gourdon.
[KM] : *Agrégation interne de mathématiques. Analyse pour le second oral.* S. Kobeissi et D. Meneu.

[Moi] : *Mathématiques supérieures, Analyse*. J. Moisan, F. Chanet, F. Delmas, N. Tosel.
[Rom] : *Éléments d'analyse réelle*. J.-E. Rombaldi.
[Ska] : *Analyse*. G. Skandalis.
[RW·L2] : *Mathématiques tout-en-un pour la Licence 2*. J.P. Ramis et A. Warusfel.