Projection de plus courte distance

On a rappelé dans la leçon sur les espaces complets le fait suivant :

A. *Dans un espace de Hilbert, tout convexe fermé non vide* C *contient un élément de norme minimale*, élément d'ailleurs unique.

Cet énoncé revient à trouver le point de C qui est le plus proche de 0 ; il est donc équivalent, par simple translation, au théorème de projection de plus courte distance sur les convexes fermés non vides d'un espace de Hilbert,

B. *Pour tout convexe fermé non vide* C *dans un espace de Hilbert* H *et pour tout point* x *de* H*, il existe un unique point* y *de* C *qui réalise la distance de* x *à* C,

$$\|x - y\| = d(x, \mathrm{C}) = \inf\{\|x - c\| : c \in \mathrm{C}\}.$$

L'énoncé **A** reste valable dans certains espaces de Banach non hilbertiens ; c'est assez facile à voir dans l'espace $\mathrm{L}^4(0,1)$ (voir les détails dans la section « *Compléments* ») ; l'énoncé **A** est vrai aussi dans tous les espaces L^p tels que $1 < p < +\infty$. Mais il n'est pas vrai dans tous les espaces de Banach : il y a des espaces où la projection de plus courte distance existe toujours, mais n'est pas nécessairement unique, et des espaces où cette projection peut ne pas exister. Voici un exemple où elle n'existe pas, suivi de quelques commentaires généraux.

Exemple. On considère dans $\mathrm{L}^1(0,1)$ (ici, nous prendrons des fonctions à valeurs réelles) l'ensemble C_1 formé de toutes les fonctions f réelles intégrables telles que

$$\int_0^1 t f(t)\,\mathrm{d}t = 1.$$

Il s'agit d'un hyperplan affine, fermé dans L^1. Mais C_1 ne contient pas d'élément de norme minimale, comme on va le prouver maintenant.

On voit que pour toute $f \in \mathrm{C}_1$,

$$1 = \int_0^1 t f(t)\,\mathrm{d}t \le \int_0^1 |f(t)|\,\mathrm{d}t = \|f\|_1,$$

donc tous les éléments de C_1 sont de norme ≥ 1. Posons pour $0 < a < 1$

$$\forall t \in [0,1], \qquad f_a(t) = \frac{1}{1-a}\,\frac{\mathbf{1}_{[a,1]}(t)}{t}\,;$$

on voit que $f_a \in \mathrm{C}_1$ et

$$\|f_a\|_1 = \frac{1}{1-a}\int_a^1 \frac{\mathrm{d}t}{t} = \frac{\ln(1/a)}{1-a}$$

tend vers 1 quand $a \to 1$. L'inf des normes des éléments de C_1 est donc égal à 1. Mais il n'est pas atteint : si $f \in \mathrm{C}_1$, la différence

$$\|f\|_1 - 1 = \int_0^1 |f(t)|\,\mathrm{d}t - \int_0^1 t f(t)\,\mathrm{d}t = \int_0^1 |f(t)|\big(1 - t\,\mathrm{sign}(f(t))\big)\,\mathrm{d}t$$

est égale à l'intégrale d'une fonction ≥ 0 : cette intégrale ne peut être nulle que si la fonction sous l'intégrale est presque partout nulle. Comme $1 - t\,\mathrm{sign}(f(t)) > 0$ pour tout $t < 1$, on déduit que f doit être nulle presque partout, ce qui est impossible car $f = \mathbf{0}_{\mathrm{L}^1}$ n'est pas dans C_1.

Moralement, la seule façon d'y arriver serait de prendre pour f la mesure de Dirac au point 1 : mais ça n'est pas une fonction ! (Voir la fin de la remarque **5** plus loin, où on développe ce point.)

Remarques.

1. Si on considère un élément quelconque $f_0 \in \mathrm{C}_1$ et l'hyperplan linéaire fermé C_0 défini par

$$\mathrm{C}_0 = \{f \in \mathrm{L}^1(0,1) : \int_0^1 t f(t)\,\mathrm{d}t = 0\},$$

on voit par translation que f_0 n'admet pas de projection de plus courte distance (pour la distance L^1) sur l'hyperplan linéaire fermé C_0.

2. On voit facilement (petit exercice) que le fait que l'hyperplan C_1 n'admet pas d'élément de norme minimale équivaut à dire que la forme linéaire ℓ continue sur $\mathrm{L}^1(0,1)$ définie par

$$\forall f \in \mathrm{L}^1(0,1), \qquad \ell(f) = \int_0^1 t f(t)\,\mathrm{d}t$$

n'atteint pas son supremum sur la boule unité fermée de $\mathrm{L}^1(0,1)$; rappelons que le supremum sur la boule unité fermée d'une forme linéaire continue ℓ sur un espace normé réel E est, par définition, la norme de ℓ dans le dual topologique E',

$$\|\ell\|_{\mathrm{E}'} = \sup\{|\ell(x)| : x \in \mathrm{E},\ \|x\| \le 1\} = \sup\{\ell(x) : x \in \mathrm{E},\ \|x\| \le 1\}.$$

On parlera donc dans la suite de forme linéaire *qui n'atteint pas sa norme sur la boule unité*. Un théorème de Robert C. James, dont la première version a été publiée en 1957, explique la généralité de ce phénomène :

un espace de Banach réel E *est réflexif si et seulement si toute forme linéaire continue sur* E *atteint sa norme sur la boule unité fermée de* E.

3. On peut remplacer l'espace $\mathrm{L}^1(0,1)$ par $\mathrm{C}([0,1])$ et la forme linéaire précédente par la forme linéaire ℓ_0, continue sur $\mathrm{C}([0,1])$, qui est définie par

$$\forall \varphi \in \mathrm{C}([0,1]), \qquad \ell_0(\varphi) = \int_0^{1/2} \varphi(t)\,\mathrm{d}t - \int_{1/2}^1 \varphi(t)\,\mathrm{d}t.$$

On verra encore que l'hyperplan affine $\{\varphi : \ell_0(\varphi) = 1\}$ n'admet pas d'élément de norme minimale (ou bien que la forme linéaire ℓ_0 n'atteint pas sa norme sur la boule unité fermée de $\mathrm{C}([0,1])$).

4. Pour la culture : un résultat (délicat, comme est le théorème de James) de Bishop-Phelps indique que *pour tout espace de Banach réel* E*, l'ensemble des formes linéaires continues sur* E *qui atteignent leur norme sur la boule unité fermée de* E*, est dense dans le dual (topologique)* E'.

À titre d'illustration, on pourra voir que pour tout ε tel que $0 < \varepsilon < 1/2$, la forme linéaire ℓ_ε définie sur $\mathrm{E} = \mathrm{C}([0,1])$ par

$$\forall \varphi \in \mathrm{C}([0,1]), \qquad \ell_\varepsilon(\varphi) = \int_0^{1/2-\varepsilon} \varphi(t)\,\mathrm{d}t - \int_{1/2+\varepsilon}^1 \varphi(t)\,\mathrm{d}t$$

atteint sa norme sur la boule unité de $\mathrm{C}([0,1])$, et qu'elle approche ℓ_0 (en norme dans l'espace E') quand $\varepsilon \to 0$.

5. Si E est un espace normé et $x \in \mathrm{E}$, on peut définir une forme linéaire continue ℓ_x sur le dual E' en posant

$$\forall x' \in \mathrm{E}', \qquad \ell_x(x') = x'(x).$$

Cette forme linéaire ℓ_x atteint toujours sa norme sur la boule unité fermée de E' (c'est une conséquence du théorème de Hahn-Banach : il existe une forme linéaire x' de norme ≤ 1 telle que $x'(x) = \|x\|$; cela implique que le plongement canonique $x \to \ell_x$ de E dans le bidual $\mathrm{E}'' = (\mathrm{E}')'$ est isométrique,

$$\|\ell_x\|_{(\mathrm{E}')'} = \|x\| \,;$$

pour mémoire, l'espace E est réflexif quand ce plongement est surjectif de E sur E'').

Changeons un peu l'éclairage du premier exemple qui a été donné dans cette feuille. La fonction continue $x : t \in [0,1] \to x(t) = t$ est un vecteur de norme 1 de l'espace normé (réel) $\mathrm{E} = \mathrm{C}([0,1])$; le dual de E est formée des mesures réelles μ sur la tribu borélienne de $[0,1]$. La forme linéaire ℓ_x est donc donnée par

$$\mu \in \mathrm{E}' \to \int_{[0,1]} t \,\mathrm{d}\mu(t) \,;$$

comme prévu, elle atteint son maximum sur la boule unité de E', au point $\mu = \delta_1$, la mesure (de probabilité) de Dirac au point 1 (et de plus, δ_1 est *l'unique* point de la boule unité où la forme linéaire atteint son maximum). Les éléments f de $\mathrm{L}^1(0,1)$ permettent de définir un sous-espace fermé de E', formé de toutes les mesures de la forme $\mathrm{d}\mu(t) = f(t)\,\mathrm{d}t$, et la norme de L^1 coïncide avec la norme induite par E' ; ainsi, l'espace $\mathrm{L}^1(0,1)$ s'identifie à un sous-espace fermé de E' ; restreinte à ce sous-espace, la forme ℓ_x s'écrit

$$f \in \mathrm{L}^1(0,1) \to \int_0^1 t f(t)\,\mathrm{d}t.$$

C'est la forme linéaire ℓ du premier exemple : elle n'atteint pas sa norme sur la boule unité de $\mathrm{L}^1(0,1)$.

6. Même dans un espace de Hilbert, les convexes fermés bornés non vides n'admettent pas toujours d'élément de norme *maximale*.

Considérons par exemple dans $\mathrm{H} = \ell^2(\mathbb{N})$ le convexe fermé borné C formé de toutes les suites $\mathbf{x} = (x_n)_{n \geqslant 0}$ de carré sommable telles que

$$\sum_{n \geqslant 0} (1 + 2^{-n})|x_n|^2 \leq 1.$$

On verra que le sup des normes des éléments de C est 1 (chercher des vecteurs $\mathbf{x} \in \mathrm{C}$ de la forme $\mathbf{x} = \lambda_n \mathbf{e}_n$, multiple du vecteur e_n de la base hilbertienne naturelle de H, pour $n \to +\infty$), mais qu'il n'est pas atteint.

Compléments.

Si x et y sont deux nombres réels, le développement de la puissance quatrième par la formule du binôme montre que

$$(x+y)^4 + (x-y)^4 = 2x^4 + 6x^2y^2 + 2y^4 \geq 2x^4 + 2y^4 \,;$$

il en résulte, par intégration, que pour deux fonctions réelles $t \to x(t)$ et $t \to y(t)$ de $\mathrm{L}^4(0,1)$, on a

$$\int_0^1 \bigl(x(t) + y(t)\bigr)^4 \,\mathrm{d}t + \int_0^1 \bigl(x(t) - y(t)\bigr)^4 \,\mathrm{d}t \geq 2\int_0^1 (x(t))^4 \,\mathrm{d}t + 2\int_0^1 (y(t))^4 \,\mathrm{d}t,$$

ou encore, en utilisant la norme de l'espace L^4,

$$\frac{\|x+y\|_4^4+\|x-y\|_4^4}{2} \ge \|x\|_4^4+\|y\|_4^4,$$

ou bien en posant $u=x+y$ et $v=x-y$

$$(*) \qquad \frac{\|u\|_4^4+\|v\|_4^4}{2} \ge \Bigl\|\frac{u+v}{2}\Bigr\|_4^4+\Bigl\|\frac{u-v}{2}\Bigr\|_4^4.$$

Cette relation, analogue à (une moitié de) la relation du parallélogramme dans un espace de Hilbert, est tout ce qu'il nous faut pour montrer l'existence et l'unicité d'un élément de norme minimale dans un convexe C non vide fermé dans $\mathrm{L}^4(0,1)$, en recopiant la preuve qu'on donne dans un Hilbert :

on peut sélectionner une suite $(u_n) \subset \mathrm{C}$ telle que $\|u_n\|$ tende vers $\mathrm{D} := d(0,\mathrm{C})$, l'infimum des normes des éléments de C ; alors cette suite (u_n) est de Cauchy dans l'espace complet L^4, puisque d'après $(*)$

$$\Bigl\|\frac{u_n-u_m}{2}\Bigr\|_4^4 \le \frac{\|u_n\|_4^4+\|u_m\|_4^4}{2}-\Bigl\|\frac{u_n+u_m}{2}\Bigr\|_4^4 \le \frac{\|u_n\|_4^4+\|u_m\|_4^4}{2}-\mathrm{D}^4,$$

où on a utilisé la convexité de C pour déduire que $(u_n+u_m)/2$, étant élément de C, a une norme $\ge \mathrm{D}$, et donc

$$\limsup_{n,m}\Bigl\|\frac{u_n-u_m}{2}\Bigr\|_4^4 \le \frac{\mathrm{D}^4+\mathrm{D}^4}{2}-\mathrm{D}^4=0\,;$$

la suite de Cauchy (u_n) converge vers un élément u de norme minimale dans le convexe fermé C ; l'unicité résulte de la même inégalité $(*)$.

L'énoncé **A** est valable dans tous les espaces L^p tels que $1<p<+\infty$; les cas $2 \le p < +\infty$ s'obtiennent en généralisant l'inégalité vue pour les nombres réels et pour $p=4$: pour tous réels x,y et tout réel $p \ge 2$, on a

$$(**) \qquad |x+y|^p+|x-y|^p \ge 2|x|^p+2|y|^p.$$

On en déduit dans l'espace L^p, pour $2 \le p < +\infty$,

$$\|x+y\|_p^p+\|x-y\|_p^p \ge 2\|x\|_p^p+2\|y\|_p^p.$$

et par la même preuve que pour L^4 on prouve l'existence et l'unicité de la projection de plus courte distance dans L^p, $2 \le p < +\infty$.

Pour $1<p\le 2$ on peut démontrer une inégalité de la forme

$$\frac{\|x+y\|_p^2+\|x-y\|_p^2}{2} \ge \|x\|_p^2+c_p^2\|y\|_p^2$$

pour une certaine constante $c_p>0$. Cette propriété est une forme de la notion *d'uniforme convexité* pour un espace normé ; consulter le livre de Brézis, *Analyse fonctionnelle*.

Pour démontrer l'inégalité réelle $(**)$, on peut se lancer dans une étude de la fonction

$$y \in \mathbb{R} \to |1+y|^p+|1-y|^p-2-2|y|^p,$$

exercice un peu fastidieux, ou bien s'appuyer sur deux bons vieux principes : les normes L^p sont croissantes en p pour un espace de probabilité, et décroissantes pour $\ell^p(\mathbb{N})$; ainsi, pour tout $p \ge 2$,

$$\Bigl(\frac{|x+y|^p+|x-y|^p}{2}\Bigr)^{1/p} \ge \Bigl(\frac{|x+y|^2+|x-y|^2}{2}\Bigr)^{1/2} = \bigl(|x|^2+|y|^2\bigr)^{1/2} \ge \bigl(|x|^p+|y|^p\bigr)^{1/p}.$$