Anaël Marit ENS Rennes

THÉORÈME DE HAHN-BANACH GÉOMATRIQUE

- 213, 229, 253 –

—

On va, dans ce développement, établir le théorème de Hahn-Banach dans sa forme géométrique dans le cas particulier des espaces de Hilbert. Ce théorème permet une vision très géométrique de la notion de partie convexe qui trouve des applications très concrètes. On détaillera d'ailleurs l'une d'entre-elles, utile notamment pour prouver l'un de mes théorèmes préférés en analyse (rien que ça).

On considère $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert réel, dont on note H' le dual topologique. Pour toute partie $C \subset H$ convexe, fermée et non-vide, on notera p_C la projection orthogonale sur C fournie par le théorème de projection. On utilisera les définitions suivantes :

Définition 1 (Hyperplan affine). *Une partie F de H est appelée* hyperplan affine *de H lorsqu'il existe une forme linéaire continue $\varphi : H \to \mathbb{R}$ et un réel α tels que :*

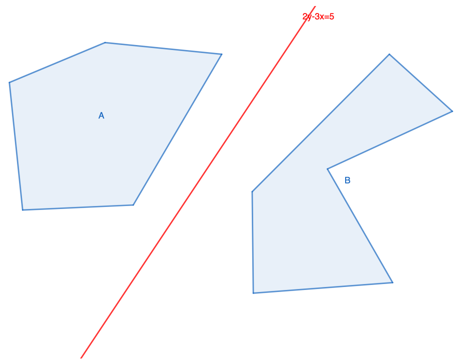
$$F = \varphi^{-1}(\{\alpha\}) \tag{1}$$

Définition 2 (Hyperplan séparateur). *On dit que deux parties disjointes A et B de H sont séparées au sens strict (resp. au sens large) par un hyperplan affine lorsque :*

$$\exists \varphi \in H',\ \exists \alpha \in \mathbb{R}\ :\ \forall (x, y) \in A \times B,\ \varphi(x) < \alpha < \varphi(y) \tag{2}$$

(resp. avec une inégalité large)

Remarque : Géométriquement, cela se traduit bien par le fait que l'hyperplan affine d'équation $\varphi(x) = \alpha$ sépare strictement A et B, comme sur le dessin :



♦

Nous allons démontrer ici le théorème suivant :

Théorème 3 (Hahn-Banach (forme géométrique) ([1], application 3.16)). *Soient A et B deux parties disjointes de H. On suppose :*

1. *A convexe et fermée.*
2. *B convexe, non vide et compacte.*

Alors A et B sont séparées au sens strict par un hyperplan affine de H.

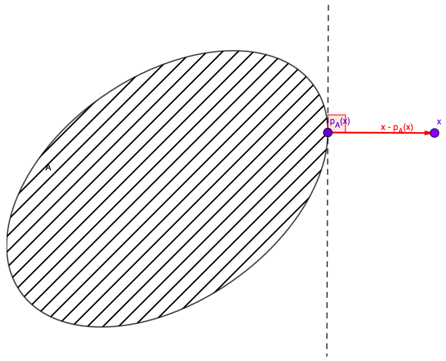
Théorème de Hahn-Banach

Remarque : Le cas où A est vide est vrai par vacuité, car toute forme linéaire continue de H est majorée sur B par compacité. On supposera donc dans la suite que A est également non vide (ce qui, admettons-le, est nettement plus intéressant).

Attention toutefois, le théorème est faux si B est vide : on pourrait alors prendre $A = H$... ♦

Démonstration.

Supposons d'abord que *B* est un singleton. On note alors $B = \{x\}$, avec $x \notin A$. La clef de la démonstration réside dans le théorème de projection sur un convexe fermé : on considère p_A la projection sur A. Le dessin suivant, en dimension 2, contient alors à lui seul toutes les informations nécessaires au bon déroulement de la preuve :



On part de l'intuition géométrique suivante : l'hyperplan séparateur recherché peut être pris orthogonal à $x - p_A(x)$. Ainsi, on cherche un hyperplan de la forme :

$$F := \mathrm{Ker}(\langle x - p_A(x), - \rangle) + \alpha \tag{3}$$

Il ne nous reste qu'à choisir correctement α. Notons $(\varphi : y \mapsto \langle x - p_A(x), y \rangle) \in H'$. On remarque alors, par caractérisation de l'angle obtus, et parce que $x \notin A$:

$$\forall y \in A,\ \varphi(y - p_A(x)) \leqslant 0 < \|x - p_A(x)\|^2 = \varphi(x - p_A(x)) \tag{4}$$

Ainsi, par linéarité :

$$\forall y \in A,\ \varphi(y) \leqslant \varphi(p_A(x)) < \varphi(x) \tag{5}$$

On prend alors $\alpha := \dfrac{\varphi(x) + \varphi(p_A(x))}{2}$ et le tour est joué ! L'inégalité précédente donne :

$$\forall y \in A,\ \varphi(y) < \alpha < \varphi(x) \tag{6}$$

Revenons au cas général. On considère :

$$C := B - A := \{x - y,\ (x, y) \in B \times A\} \tag{7}$$

Il s'agit d'une partie non vide, convexe et fermée de H. En effet, le produit $A \times B$ est convexe dans H^2 par convexité de A et de B. Comme C est son image par l'application continue $(y, x) \mapsto x - y$, c'est un convexe. Si de plus $(x_n - y_n)$ est une suite d'éléments de C convergeant dans H vers l, avec $(x_n) \in B^{\mathbb{N}}$ et $(y_n) \in A^{\mathbb{N}}$, par compacité de B, on peut extraire une sous-suite le long de laquelle (x_n) converge vers $x \in B$. En passant à la limite le long de cette sous suite dans l'expression $y_n = x_n - (x_n - y_n)$, on trouve que y_n admet $x - l$ comme valeur d'adhérence, qui se trouve dans A puisque cette partie est fermée. Il vient alors $l = x - (x - l) \in C$. Donc C est bien fermée.

Puisque A et B sont disjointes, C ne contient pas 0. Par le travail précédemment effectué, il existe un hyperplan affine qui sépare C et $\{0\}$. En d'autres termes, on dispose d'une forme linéaire continue $\varphi \in H'$ et de $\alpha \in \mathbb{R}$ tels que :

$$\forall (y, x) \in A \times B,\ \varphi(y - x) < \alpha < \varphi(0) = 0 \tag{8}$$

et donc :

$$\forall (y, x) \in A \times B,\ \varphi(y) < \alpha + \varphi(x) < \varphi(x) \tag{9}$$

enfin, comme c'est vrai pour tout x et y, et comme α est strictement négatif :

$$\sup_{y\in A} \varphi(y) \leqslant \inf_{x\in B} \varphi(x) + \alpha < \inf_{x\in B} \varphi(x) \tag{10}$$

Quitte à modifier la valeur de α, on peut supposer qu'on a l'inégalité stricte, ce qui est une exacte reformulation du fait que A et B sont séparés au sens strict par un hyperplan affine.

□

Remarque : Ce théorème reste vrai dans le cas plus général des espaces de Banach, mais la preuve est très différente car on n'a plus accès aux commodes outils hilbertiens. Celle-ci recquiert en particulier l'axiome du choix et la notion de jauge d'un convexe. Vous pouvez la trouver dans [2]. A titre personnel, je pense que ce cadre élargi n'a pas d'intérêt au niveau de l'agreg, voire même risque de desservir à se placer dans un cadre plus abstrait, qui repose sur plus d'outils hors-programme, à moins d'avoir de vraies applications à en proposer. En plus, ça se recase moins-bien... ♦

Application à l'étude des fonctions convexes

On va maintenant fournir à l'aide du théorème de Hahn-Banach un critère pratique de convexité d'une fonction définie sur un espace de Hilbert réel.

Théorème 4 ([4], chap. VI, exercice 3). *Soit C une partie convexe fermée et non-vide de H et $f : C \to \mathbb{R}$ une application continue. On a alors l'équivalence suivante :*

1. *f est convexe.*
2. *Il existe $(g_i)_{i\in I}$ une famille d'applications affines de H telle que :*

$$\forall x \in C,\ \ f(x) = \sup_{i\in I} g_i(x) \tag{11}$$

Démonstration.

Dans le sens indirect , il suffit de vérifier que le suppremum d'une famille d'applications convexes est convexe, ce qui est élémentaire. Comme les applications affines sont convexes, f est bien convexe.

Dans le sens direct , on va exploiter le théorème de Hahn-Banach géométrique. On considère *l'épigraphe* de f, c'est-à-dire la partie de $H \times \mathbb{R}$:

$$\mathscr{E}(f) := \{(x, s) \in C \times \mathbb{R} \mid s \geqslant f(x)\} \tag{12}$$

Par convexité de f, $\mathscr{E}(f)$ est un convexe de $H \times \mathbb{R}$. Il est par ailleurs fermé par continuité de f et fermeture de C. Fixons alors $x_0 \in C$ et un réel $r < f(x_0)$, de façon à ce que (x_0, r) n'appartienne pas à $\mathscr{E}(f)$. Puisqu'on a une structure d'espace de Hilbert sur $H \times \mathbb{R}$ donnée par le produit scalaire

$$\langle (x_1, s_1), (x_2, s_2)\rangle := \langle x_1, x_2\rangle_H + s_1 s_2 \tag{13}$$

on peut séparer strictement $\mathscr{E}(f)$ et (x_0, r) par un hyperplan affine d'équation $\Phi(x,s) = \alpha$, avec $\Phi \in (H \times \mathbb{R})'$ et $\alpha \in \mathbb{R}$. Posons $(\psi : x \mapsto \Phi(x,0)) \in H'$ et $\lambda \in \mathbb{R}$ tel que :

$$\forall s \in \mathbb{R},\ \Phi(0,s) = \lambda s \tag{14}$$

de sorte à pouvoir écrire :

$$\forall (x,s) \in H \times \mathbb{R},\ \Phi(x,s) = \psi(x) + \lambda s \tag{15}$$

On a alors, par définition :

$$\forall (x,s) \in \mathscr{E}(f),\ \psi(x_0) + \lambda r < \alpha < \psi(x) + \lambda s \tag{16}$$

En particulier, $(x,s) = (x_0, f(x_0)) \in \mathscr{E}(f)$, on trouve en basculant tous les termes à gauche de l'inégalité :

$$\lambda(r - f(x_0)) < 0 \tag{17}$$

et comme r est choisi inférieur strict à $f(x_0)$, il vient que λ est strictement positif. Autrement dit, l'hyperplan d'équation $\Phi(x,s) = \alpha$ est le graphe de l'application affine :

$$\begin{aligned} g_{r,x_0} : H &\to \mathbb{R} \\ x &\mapsto \frac{1}{\lambda}(\alpha - \psi(x)) \end{aligned} \tag{18}$$

et par construction, on a :

$$\forall x \in C,\ \psi(x_0) + \lambda r < \underbrace{\psi(x) + \lambda g_{x_0,r}(x)}_{=\alpha} < \psi(x) + \lambda f(x) \tag{19}$$

D'où :

$$\forall x \in C, g_{x_0,r}(x) < f(x) \tag{20}$$

et donc, puisque ceci vaut quelque soit x_0 et r, on peut passer au sup :

$$\sup_{\substack{x_0 \in C \\ r < f(x_0)}} g_{x_0,r}(x) \leqslant f(x) \tag{21}$$

Mais pour x_0 fixé, on trouve :

$$\forall r < f(x_0),\ r < g_{x_0,r}(x_0) < f(x_0) \tag{22}$$

et donc par passage au sup :

$$\sup_{r < x_0} g_{x_0,r}(x_0) = f(x_0) \tag{23}$$

d'où finalement l'égalité :

$$\sup_{\substack{x_0 \in C \\ r < f(x_0)}} g_{x_0,r}(x) = f(x) \tag{24}$$

$\square$

Inégalité de Jenssen et applications

Je ne traite pas cette partie lors du développement par manque de temps, mais il est bon de l'avoir en tête comme corollaire du résultat précédent. J'ai l'impression que l'inégalité de Jenssen passe généralement un peu inaperçue lors des cours de théorie de la mesure, et pourtant, elle peut permettre de simplifier un bon nombre de preuves usuelles qui sont souvent menées de façon plus calculatoire. Je me suis permise de mettre l'inégalité de Hölder en exemple.

Théorème 5 (Inégalité de Jenssen). *Soient $d \in \mathbb{N}^*$ et $(X, \mathfrak{F}, \mu)$ un espace probabilisé. On considère $f : X \to \mathbb{R}^d$ une fonction intégrable et $\varphi : \mathbb{R}^d \to \mathbb{R}$ une fonction convexe telle que $\varphi \circ f$ est intégrable. Alors :*

$$\varphi\left(\int_X f(x)\mathrm{d}\mu(x)\right) \leqslant \int_X \varphi \circ f(x)\mathrm{d}\mu(x) \tag{25}$$

Démonstration. On se donne $(g_i)_{i\in I}$ une famille de fonctions affines de $\mathbb{R}^d$ dans $\mathbb{R}$ telle que :

$$\varphi = \sup_{i\in I} g_i \tag{26}$$

Chaque fonction affine g_i peut se décomposer sous la forme $g_i = l_i + c_i$, avec l_i une forme linéaire et c_i une constante réelle. Or par définition de l'intégrale, l'opérateur d'intégration commute avec les formes linéaires continues. De plus, comme μ est une mesure de probabilités, les constantes sont invariantes par intégrale contre μ. Autrement dit :

$$\forall i \in I,\ g_i\left(\int_X f(x)\mathrm{d}\mu(x)\right) = \int_X g_i \circ f(x)\mathrm{d}\mu(x) \tag{27}$$

Par croissance de l'intégrale, on peut passer au sup sous l'intégrale de droite :

$$\forall i \in I,\ g_i\left(\int_X f(x)\mathrm{d}\mu(x)\right) \leqslant \int_X \underbrace{\sup_{j\in I} g_j}_{=\varphi} \circ f(x)\mathrm{d}\mu(x) \tag{28}$$

et comme c'est vrai pour tout $i \in I$, on obtient par passage au sup :

$$\varphi\left(\int_X f(x)\mathrm{d}\mu(x)\right) \leqslant \int_X \varphi \circ f(x)\mathrm{d}\mu(x) \tag{29}$$

□

Remarque : Dans le cas particulier $d = 1$, il existe une preuve beaucoup plus élémentaire de l'inégalité de Jenssen (voir par exemple [3]). Ce cas particulier suffit notamment à montrer bon nombre de résultats usuels, comme l'inégalité de Hölder, ou le théorème d'approximation par convolution. ♦

Corollaire 6 (Inégalité de Hölder). *Soit $(X, \mathfrak{F}, \mu)$ un espace mesuré. Soit $p \in [1, +\infty]$. Soit q son exposant conjugué, c'est-à-dire l'unique réel tel que $\frac{1}{p} + \frac{1}{q} = 1$. Pour toute fonction $f \in L^p(\mu)$ et toute fonction $g \in L^q(\mu)$, le produit fg est intégrable et on a :*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q \tag{30}$$

Démonstration. Le cas $p \in \{1, +\infty\}$ est immédiat et résulte de la croissance de l'inégrale. On supposera donc que ni p ni q n'est égal à $+\infty$. On écrit alors :

$$\int_X |f(x)||g(x)|\mathrm{d}\mu(x) = \|f\|_p^p \left(\left(\int_X |f(x)|^{1-p}|g(x)|\frac{|f(x)|^p}{\|f\|_p^p}\mathrm{d}\mu(x)\right)^q\right)^{1/q} \tag{31}$$

Mais la fonction $x \mapsto \frac{|f(x)|^p}{\|f\|_p^p}$ est intégrable, positive et d'intégrale 1. Ainsi, la mesure $\frac{|f|^p}{\|f\|_p^p}\mathrm{d}\mu$ est une mesure de probabilités sur X. Comme la fonction $x \mapsto |x|^q$ est convexe sur $\mathbb{R}$, on peut appliquer l'inégalité de Jenssen et pousser la puissance q-ième sous l'intégrale :

$$\int_X |f(x)||g(x)|\mathrm{d}\mu(x) \leqslant \|f\|_p^p \left(\int_X |f(x)|^{q(1-p)}|g(x)|^q\frac{|f(x)|^p}{\|f\|_p^p}\mathrm{d}\mu(x)\right)^{1/q} \tag{32}$$

Il ne reste qu'à remarquer que $q = \frac{p}{p-1}$ pour conclure, car on a alors :

$$\forall x \in X,\ |f(x)|^{q(1-p)}|f(x)|^p = 1 \tag{33}$$

et donc :

$$\int_X |f(x)||g(x)|\mathrm{d}\mu(x) \leqslant \|f\|_p^p \times \|f\|_p^{-p/q}\left(\int_X |g(x)|^q\mathrm{d}\mu(x)\right)^{1/q} \tag{34}$$

$$= \|f\|_p\|g\|_q \tag{35}$$

□

Références

[1] Vincent Beck, Jerôme Malick et Gabriel Peyré. *Objectif agrégation.* H&K, 2005. isbn : 2914010923.

[2] Haïm Brezis. *Analyse fonctionnelle, théorie et applications.* Masson, 1987. isbn : 2-225-77198-7.

[3] Marc Briane et Gilles Pagès. *Analyse, Théorie de l'intégration.* Deboeck.

[4] Daniel Li. *Cours d'analyse fonctionnelle.* Ellipses.