

OPTIMISATION DANS UN HILBERT

- 203, 205, 213, 219, 229, 253 -

L'objectif de ce développement est de généraliser un résultat d'optimisation de fonctionnelles convexes, valables dans les espaces de dimension finie par des arguments très simples de compacité, au cadres des espaces de Hilbert. La preuve exploitera un de mes théorèmes d'analyse fonctionnelle préférés : le théorème de Banach-Alaoglu, qu'on démontrera en première partie. C'est un théorème qui est vrai dans un cadre beaucoup plus général en travaillant avec la notion de convergence faible-, mais qui s'obtient dans le cadre hilbertien en utilisant seulement des notions du programme d'agreg.*

L'intégralité du développement se retrouve dans [2]. J'ai légèrement modifié l'organisation de la preuve pour mettre plus en valeur le théorème de Banach-Alaoglu, mais tous les ingrédients se retrouvent dans leur livre.

J'ai ajouté une annexe à la fin du document qui fournit un contre-exemple lorsque on sort du cadre hilbertien, et qui donne une application concrète du théorème (qui que très hors-programme pour l'agrégation).

Soit H un espace de Hilbert. On considère $J : H \rightarrow H$ une fonction convexe, continue et coercive⁽ⁱ⁾. Le but est de démontrer :

Théorème 1. *Il existe $x^* \in H$ tel que $J(x^*) = \inf_{x \in H} J(x)$.*

Lorsque H est de dimension finie, ce théorème s'obtient très simplement de la façon suivante : on choisit $y \in \mathbb{R}$ une valeur prise par J , et $R > 0$ tel que $J(x) \geq y$ dès que y est de norme plus grande que R , donné par coercivité de J . Comme J est continue, elle atteint un minimum sur le compact $B(0, R)$ et ce minimum est global de par le choix de R . Mais en dimension infinie, c'est une autre paire de manches !

Théorème de Banach-Alaoglu

La clef de la démonstration sera le théorème suivant :

Théorème 2 (Banach-Alaoglu). *Soit $(x_n) \in H^{\mathbb{N}}$ une suite bornée de vecteurs de H . Il existe $\psi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante et $x^* \in H$ tel que :*

$$\forall y \in H, \langle x_{\psi(n)}, y \rangle \xrightarrow{n \rightarrow +\infty} \langle x^*, y \rangle \quad (1)$$

On dit alors que (x_n) converge faiblement vers x^ , ce qu'on note $x_n \rightharpoonup x^*$.*

(i). C'est-à-dire telle que $|J(x)| \xrightarrow{\|x\| \rightarrow +\infty} +\infty$.

Démonstration. On considère le sous-espace vectoriel fermé de H :

$$F := \overline{\text{Vect}(x_n \mid n \in \mathbb{N})} \quad (2)$$

F est un espace séparable, c'est-à-dire qu'il contient une partie dénombrable et dense (par exemple le \mathbb{Q} -espace vectoriel engendré par les (x_n)). Soit (y_n) une telle suite dense. On va alors construire itérativement une suite d'extractrice de la façon suivante :

- Comme (x_n) est bornée, la suite $(\langle x_n, y_0 \rangle)$ est bornée dans \mathbb{R} par l'inégalité de Cauchy-Schwarz. Ainsi, il existe φ_0 une extractrice telle que $(\langle x_{\varphi_0(n)}, y_0 \rangle)$ converge.
- Si on suppose contruite une famille $\varphi_0, \dots, \varphi_k$ d'extractrices telle que pour tout $i \in \llbracket 0, k \rrbracket$, la suite $(\langle x_{\varphi_0 \circ \dots \circ \varphi_i(n)}, y_i \rangle)$ converge dans \mathbb{R} , il suffit de prendre φ_{k+1} une extractrice telle que $(\langle x_{\varphi_0 \circ \dots \circ \varphi_{k+1}(n)}, y_{k+1} \rangle)$ converge pour obtenir le terme suivant.

Sans surprise, on procède à une extraction diagonale : posons

$$\begin{aligned} \psi : \mathbb{N} &\rightarrow \mathbb{N} \\ k &\mapsto \varphi_0 \circ \dots \circ \varphi_k(k) \end{aligned} \quad (3)$$

Pour tout $k \in \mathbb{N}$, la suite $(\langle x_{\psi(n)}, y_k \rangle)$ est alors convergente dans \mathbb{R} .

Montrons maintenant que $(\langle x_{\psi(n)}, y \rangle)$ converge encore pour tout vecteur $y \in F$ en montrant que cette suite est de Cauchy. Soit $y \in F$. Par densité :

$$\exists k \in \mathbb{N} : \|y - y_k\| \leq \frac{\varepsilon}{2 \sup_{n \in \mathbb{N}} \|x_n\|} \quad (4)$$

De plus, grâce au travail précédent :

$$\exists N \in \mathbb{N}, : \forall n, p \geq N, |\langle x_{\psi(n)} - x_{\psi(p)}, y_k \rangle| \leq \frac{\varepsilon}{2} \quad (5)$$

Ainsi, pour tout $n, p \geq N$, on a :

$$|\langle x_{\psi(n)} - x_{\psi(p)}, y \rangle| \leq |\langle x_{\psi(n)} - x_{\psi(p)}, y - y_k \rangle| + |\langle x_{\psi(n)} - x_{\psi(p)}, y_k \rangle| \quad (6)$$

$$\leq (\|x_{\psi(n)}\| + \|x_{\psi(p)}\|) \|y - y_k\| + |\langle x_{\psi(n)} - x_{\psi(p)}, y_k \rangle| \quad (7)$$

$$\leq \varepsilon \quad (8)$$

Cette suite est donc de Cauchy dans \mathbb{R} : elle converge. On peut donc légitimement poser :

$$\begin{aligned} L : F &\rightarrow \mathbb{R} \\ y &\mapsto \lim_{n \rightarrow +\infty} \langle x_{\psi(n)}, y \rangle \end{aligned} \quad (9)$$

Par linéarité et unicité de la limite, ainsi que par bilinéarité du produit scalaire, L est linéaire. Montrons qu'elle est continue. Pour $y \in F$ et $n \in \mathbb{N}$, on a :

$$|\langle x_{\psi(n)}, y \rangle| \leq \sup_{n \in \mathbb{N}} \|x_n\| \|y\| \quad (10)$$

d'où en passant à la limite :

$$|L(y)| \leq \sup_{n \in \mathbb{N}} \|x_n\| \|y\| \quad (11)$$

et L est bien continue. On peut donc appliquer le théorème de représentation de Riesz :

$$\exists! x^* \in F : L = \langle x^*, - \rangle \quad (12)$$

ce qui est exactement dire que $(x_{\psi(n)})$ converge faiblement vers x^* **dans F** . Montrons que x^* est encore valeur d'adhérence au sens faible dans H . Pour cela, on exploite la décomposition orthogonale :

$$H = F \oplus F^\perp \quad (13)$$

Pour $y \in H$, il existe une unique écriture $y = y_F + y_\perp$ avec $y_F \in F$ et $y_\perp \in F^\perp$. On a alors :

$$\langle x_{\psi(n)}, y \rangle = \langle x_{\psi(n)}, y_F \rangle \xrightarrow{n \rightarrow +\infty} \langle x^*, y_F \rangle = \langle x^*, y \rangle \quad (14)$$

car $x^* \in F$. Donc la convergence faible vers x^* a en fait lieu dans H . \square

Minimisation de J

On va maintenant exploiter le théorème de Banach-Alaoglu pour prouver que J admet un minimum sur H . Notons $\alpha := \inf_{x \in H} J(x) \in \mathbb{R} \cup \{-\infty\}$. Par définition de l'infimum, il existe (x_n) une suite d'éléments de H telle que :

$$J(x_n) \xrightarrow{n \rightarrow +\infty} \alpha \quad (15)$$

Par coercivité de J , (x_n) est bornée, puisque si (x_n) admettait une sous-suite dont la norme diverge, $J(x_n)$ tendrait vers $+\infty$ le long de cette sous-suite. Ainsi, le théorème de Banach-Alaoglu fournit un vecteur $x^* \in H$ et une extractrice φ telle que $x_{\varphi(n)} \rightharpoonup x^*$. Montrons que $J(x^*) = \alpha$ (et donc en particulier que α est fini).

On considère pour cela $\beta > \alpha$ et on définit :

$$C_\beta := J^{-1}(] - \infty, \beta]) \quad (16)$$

On va montrer que $x^* \in C_\beta$, c'est-à-dire $J(x^*) \leq \beta$. Si on a cette appartenance pour tout β , on aura nécessairement que $J(x^*) = \alpha$.

Comme J est continue, C_β est fermé dans H . De plus, par convexité de J , C_β est convexe. Enfin, la définition de β implique que C_β est non-vide. On peut donc appliquer le théorème de projection sur un convexe fermé : notons p_β la projection sur C_β .

Puisque $J(x_n) \rightarrow \alpha$, il existe un rang n_0 à partir duquel les termes de la suite (x_n) sont dans C_β . Par caractérisation de l'angle obtus, on a alors :

$$\Re(\langle x^* - p_\beta(x^*), x_{\varphi(n)} - p_\beta(x^*) \rangle) \leq 0 \quad (17)$$

Or en passant à la limite $[n \rightarrow +\infty]$, par convergence faible de $(x_{\varphi(n)})$:

$$\Re(\langle x^* - p_\beta(x^*), x^* - p_\beta(x^*) \rangle) = \|x^* - p_\beta(x^*)\|^2 \leq 0 \quad (18)$$

et donc $x^* = p_\beta(x^*)$: on a bien prouvé $J(x^*) \leq \beta$! Ainsi, $J(x^*)$ ne peut-être que α , et α est fini ! \square

Annexe : à quoi ça sert ?

Bonne question en effet, le genre qu'il faut s'être posée avant de passer à l'oral. Le fait est que ça n'est pas évident de trouver une « vraie » application de ce résultat.

Remarquons tout d'abord que le résultat qu'on vient de montrer est plus fort que le théorème de Lax-Milgram : on s'attend à pouvoir montrer le même genre de résultat avec. En l'occurrence, il peut servir à montrer que certains types d'équations différentielles non linéaires admettent des solutions. Cela dit, le cadre habituel de travail pour faire des équations diff est l'ensemble des fonctions de classe C^k selon la régularité dont on a besoin, et... ça n'est pas un espace de Hilbert ! On va devoir introduire certains espaces qui vont permettre de chercher des solutions au sens faible :

Définition 3 (Espace de Sobolev ([1], chap X.1)). Pour $k \in \mathbb{N}$, on note $H^k(I)$ l'ensemble des fonctions de classe L^2 sur I dont les k premières dérivées au sens des distributions sont représentées par une fonction L^2 . En d'autres termes :

$$H^k(I) = \left\{ f \in L^2 \mid \forall j \in \llbracket 0, k \rrbracket, \exists g_j \in L^2(I) : \forall \varphi \in C_c^\infty(I), (-1)^k \int_I f \varphi^{(j)} = \int_I g_j \varphi \right\} \quad (19)$$

On muni $H^k(I)$ du produit scalaire :

$$\langle f, g \rangle_{H^k} := \sum_{j=0}^k \int_I f^{(j)} g^{(j)} \quad (20)$$

C'est alors un espace de Hilbert.

On note H_0^k l'adhérence des fonctions C^∞ à support compact dans H^k .

Ces espaces offrent un cadre de travail pour étudier des équations différentielles tout en profitant des outils hilbertiens. Notre théorème permet notamment de montrer :

Théorème 4. Soient f une fonction L^2 et p un entier non nul. Il existe une unique fonction $u \in H^2(]0, 1[) \cap H_0^1(]0, 1[)$ telle que :

$$\begin{cases} (u'^2)' + u|u|^{p-1} = f \\ u(0) = u(1) = 0 \end{cases} \quad (21)$$

La démonstration est TRÈS longue et technique (elle est faite intégralement dans la version de Thomas Courant). Je pense que c'est complètement déraisonnable de vouloir placer ce genre de choses dans son plan (sauf si vous avez une machine de guerre sur ce genre de sujet), mais à mon avis c'est toujours bon d'avoir de vagues rudiments pour pouvoir répondre à cette bête question : « à quoi ça sert ? ».

Bonus : un contre-exemple dans le cas non hilbertien

Pour pouvoir affirmer que ce théorème a de la valeur, il m'a semblé important de trouver un contre-exemple lorsque H n'est plus supposé être un espace de Hilbert (mettons un espace de Banach par exemple). Je vais donner ici un exemple de fonctionnelle convexe, continue et coercive qui n'atteint pas de minimum, mais il est très possible que mon exemple soit beaucoup trop compliqué pour ce que c'est. Si jamais vous trouvez quelque chose de beaucoup plus simple, n'hésitez pas à me contacter !

On se place dans l'espace E des suites réelles qui convergent vers 0 qu'on munit de la norme infinie. Soit (a_n) une suite l^1 de réels strictement positifs dont la somme vaut 1. On pose alors :

$$\begin{aligned} \varphi : E &\rightarrow \mathbb{R} \\ (x_n) &\mapsto \sum_{n=0}^{+\infty} a_n x_n \end{aligned} \tag{22}$$

φ est une forme linéaire continue de norme d'opérateur 1. Remarquons que φ n'atteint pas sa norme d'opérateur, dans le sens où pour toute suite non nulle (x_n) de E , $|\varphi((x_n))| < \|(x_n)\|_\infty$, puisque l'égalité est atteinte si, et seulement si, (x_n) est constante (et n'est donc pas un élément de E). On pose alors :

$$\begin{aligned} J : E &\rightarrow \mathbb{R} \\ (x_n) &\mapsto \max\{1, \|(x_n)\|_\infty\} + |1 - \varphi((x_n))| \end{aligned} \tag{23}$$

J est clairement continue et coercive. Montrons qu'elle est de plus convexe. Soit $(x, y) \in E^2$ et soit $\lambda \in [0, 1]$. On a :

$$|1 - \varphi(\lambda x + (1 - \lambda)y)| \leq \lambda |1 - \varphi(x)| + (1 - \lambda) |1 - \varphi(y)| \tag{24}$$

par inégalité triangulaire. De plus :

$$\max\{1, \|\lambda x + (1 - \lambda)y\|_\infty\} \leq \max\{1, \lambda \|x\|_\infty + (1 - \lambda)\|y\|_\infty\} \tag{25}$$

$$\leq \max\{1, \lambda \|x\|_\infty\} + \max\{1, (1 - \lambda)\|y\|_\infty\} \tag{26}$$

$$\leq \lambda \max\{1, \|x\|_\infty\} + (1 - \lambda) \max\{1, \|y\|_\infty\} \tag{27}$$

Il est clair que $J \geq 1$ et que si $((x_n^{(p)}))_{p \in \mathbb{N}}$ est une suite de suites de E de norme 1 telle que $\varphi((x_n^{(p)})) \xrightarrow{p \rightarrow +\infty} 1$, alors $J((x_n^{(p)})) \rightarrow 1$. Donc 1 est l'infimum de J . Mais pourtant :

$$J(x_n) = 1 \iff \max\{1, \|(x_n)\|_\infty\} = 1 \wedge |1 - \varphi((x_n))| = 0 \tag{28}$$

$$\iff \|(x_n)\|_\infty \leq 1 \wedge \varphi((x_n)) = 1 \tag{29}$$

ce qui est impossible puisque φ n'atteint pas sa norme. Donc J n'a pas de minimum sur E .

Références

- [1] Francis HIRSCH et Gilles LACOMBE. *Elements d'analyse fonctionnelle*. Dunod, 2009. ISBN : 978-2-10-053594-1.
- [2] Lucas ISENMANN et Timothée PECATTE. *L'oral à l'agrégation de mathématiques, une sélection de développements*. Ellipses, 2024.