Développement : Optimisation dans un HILBERT

ANALYSE & PROBABILITÉS

Références :
- [IP] ISENMANN L., PECATTE T., *L'oral à l'agrégation de mathématiques - Une sélection de développements*, ellipses, 2019, p335. $\longrightarrow$ La référence était interdite aux oraux pendant mon année (2024), donc vérifiez dans la nouvelle édition autorisée que ce développement y est toujours.
- [B] BRÉZIS H., *Analyse fonctionnelle - Théorie et applications*, Masson, 1987. $\longrightarrow$ Pour les rappels sur les espaces de SOBOLEV (voir application du théorème).

Pour les leçons :
- 213 : Espaces de HILBERT. Exemples d'applications.
- 219 : Extremums : existence, caractérisation, recherche. Exemples et applications.
- 223 : Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.
- 229 : Fonctions monotones. Fonctions convexes. Exemples et applications.
- 253 : Utilisation de la notion de convexité en analyse.

Définition 1.

Soit $f : H \to \mathbb{R}$. On dit que f est *coercive* lorsque

$$f(x) \underset{\|x\|\to+\infty}{\longrightarrow} +\infty.$$

Théorème 1.

Soient $(H, \|\cdot\|)$ un espace de HILBERT et $J : H \to \mathbb{R}$ une fonction convexe, continue et coercive.
Alors, il existe $a \in H$ tel que :

$$J(a) = \inf_{x\in H} J(x).$$

En particulier, l'infimum de J sur H est fini.

PREUVE : Soit $(x_k)_{k\in\mathbb{N}} \in H^{\mathbb{N}}$ telle que $J(x_k) \underset{k\to+\infty}{\longrightarrow} \inf_{x\in H} J(x)$.

⋆ ÉTAPE 1 : Montrons que $(x_k)_{k\in\mathbb{N}}$ est bornée. Montrons-le par l'absurde : supposons $(x_k)_{k\in\mathbb{N}}$ non bornée.
Il existe une extractrice $\varphi : \mathbb{N} \to \mathbb{N}$ telle que $\|x_{\varphi(k)}\| \underset{k\to+\infty}{\longrightarrow} +\infty$.
Comme J est coercive, il vient $J(x_{\varphi(k)}) \underset{k\to+\infty}{\longrightarrow} +\infty$, ce qui est absurde puisque $J(x_{\varphi(k)}) \underset{k\to+\infty}{\longrightarrow} \inf_{x\in H} J(x)$.
$(x_k)_{k\in\mathbb{N}}$ est donc bornée (disons, par une constante $C > 0$).

⋆ ÉTAPE 2 : On a appliquer un procédé d'extraction diagonale.
On pose, pour tout $k \in \mathbb{N}$, $u_k = \langle x_0 | x_k\rangle$. D'après l'inégalité de CAUCHY-SCHWARZ, on a en outre, pour tout $k \in \mathbb{N}$:

$$\|u_k\| \leqslant \|x_0\|\|x_k\| \leqslant C\|x_0\|,$$

grâce à l'ÉTAPE 1.
Montrons par récurrence la propriété $\mathcal{P}(m)$ suivante :

" Il existe des extractrices $\varphi_0, \dots, \varphi_m$ telles que, pour tout $i \leqslant m$, $(\langle x_i | x_{\varphi_0\circ\cdots\circ\varphi_i(k)}\rangle)_{k\in\mathbb{N}}$ converge ".

$\longrightarrow$ Initialisation : Si $m = 0$, on vient de montrer que $(u_k)_{k\in\mathbb{N}}$ est bornée. D'après le théorème de BOLZANO-WEIERSTRASS, il existe une extractrice φ_0 telle que $(\langle x_0 | x_{\varphi_0(k)}\rangle)_{k\in\mathbb{N}}$ converge. $\mathcal{P}(0)$ est donc vraie.
$\longrightarrow$ Hérédité : Soit $m \in \mathbb{N}$. Supposons $\mathcal{P}(m)$, et montrons $\mathcal{P}(m+1)$. D'après $\mathcal{P}(m)$, il existe des extractrices $\varphi_0, \dots, \varphi_m$ telles que, pour tout $i \leqslant m$, $u_i := (\langle x_i | x_{\varphi_0\circ\cdots\circ\varphi_i(k)}\rangle)_{k\in\mathbb{N}}$ converge et, pour $k \in \mathbb{N}$:

$$|\langle x_{m+1} | x_{\varphi_0\circ\cdots\circ\varphi_m(k)}\rangle| \leqslant \|x_{m+1}\|\|x_{\varphi_0\circ\cdots\circ\varphi_m(k)}\| \leqslant C\|x_{m+1}\|.$$

La suite $(\langle x_{m+1} | x_{\varphi_0\circ\cdots\circ\varphi_m(k)}\rangle)_{k\in\mathbb{N}}$ est donc bornée. D'après le théorème de BOLZANO-WEIERSTRASS, il existe une extractrice $\varphi_{m+1} : \mathbb{N} \to \mathbb{N}$ telle que $(\langle x_{m+1} | x_{\varphi_0\circ\cdots\circ\varphi_m\circ\varphi_{m+1}(k)}\rangle)_{k\in\mathbb{N}}$ converge.
Donc $\mathcal{P}(m+1)$ est vraie. Cela achève la récurrence.

Soit $\psi : \mathbb{N} \to \mathbb{N}$ définie par :

$$\forall k \in \mathbb{N} \quad \psi(k) = (\varphi_0 \circ \cdots \circ \varphi_k)(k).$$

Alors, pour tout $i \in \mathbb{N}$, $(\langle x_i | x_{\psi(k)}\rangle)_{k\in\mathbb{N}}$ est une suite extraite de $(\langle x_i, | x_{\varphi_0\circ\cdots\circ\varphi_i(k)}\rangle)_{k\in\mathbb{N}}$ (à partir d'un certain rang). Donc elle converge.

⋆ ÉTAPE 3 : Pour $k \in \mathbb{N}$, on pose $y_k = x_{\psi(k)}$. Montrons qu'il existe $a \in H$ tel que, pour tout $u \in H$, $\langle u, y_k\rangle \underset{k\to+\infty}{\longrightarrow} \langle u|a\rangle$.
Soit $F = \mathrm{Vect}(\{x_p \mid p \in \mathbb{N}\})$. Alors, par linéarité et par l'ÉTAPE 2, pour tout $v \in F$, $(\langle v | x_{\psi(k)}\rangle)_{k\in\mathbb{N}}$ converge.

De plus, on a :

$$H = \bar{F} \oplus F^{\perp}.$$

Soient $u \in H$ et $\varepsilon > 0$. Il existe $v \in \bar{F}$ et $w \in F^{\perp}$ tel que $u = v + w$, et il existe $\tilde{v} \in F$ tel que $\|v - \tilde{v}\| < \dfrac{\varepsilon}{4C}$.
Soient $k, \ell \in \mathbb{N}$. On a :

$$\begin{array}{rcl} |\langle u|y_k - y_\ell\rangle| & = & |\langle v|y_k - y_\ell\rangle| \\ & \leqslant & |\langle v - \tilde{v}|y_k - y_\ell\rangle| + |\langle \tilde{v}, y_k - y_\ell\rangle| \\ & \leqslant & \underbrace{\|v - \tilde{v}\|}_{< \frac{\varepsilon}{4C}} \underbrace{\|y_k - y_\ell\|}_{\leqslant 2C} + |\langle \tilde{v}, y_k - y_\ell\rangle|, \end{array}$$

d'après l'inégalité de CAUCHY-SCHWARZ. On sait de plus que $(\langle \tilde{v}|y_k\rangle)_{k\in\mathbb{N}}$ est une suite de CAUCHY, puisqu'elle converge. Il existe donc $k_0, \ell_0 \in \mathbb{N}$ tels que, pour tout $k \geqslant k_0$ et $\ell \geqslant \ell_0$, $|\langle \tilde{v}, y_k - y_\ell\rangle| < \dfrac{\varepsilon}{2}$.
On a montré que :

$$\forall k \geqslant k_0 \quad \forall \ell \geqslant \ell_0 \quad |\langle u|y_k - y_\ell\rangle| < \varepsilon.$$

La suite $(\langle u|y_k\rangle)_{k\in\mathbb{N}}$ est donc une suite de CAUCHY réelle. Comme $\mathbb{R}$ est complet, elle converge vers un réel ℓ_u.
Maintenant, soit $\varphi : H \to \mathbb{R}$ définie par :

$$\forall u \in H \quad \varphi(u) = \ell_u.$$

Alors, φ est une forme linéaire (la linéarité provient de la linéarité du produit scalaire), et elle est continue puisque, encore grâce à l'inégalité de CAUCHY-SCHWARZ[1] :

$$\forall u \in H \quad |\varphi(u)| \leqslant C\|u\|.$$

D'après le théorème de représentation de RIESZ, il existe $a \in H$ tel que, pour tout $u \in H$, $\varphi(u) = \langle u|a\rangle$, ce qui achève la preuve de l'ÉTAPE 3.
⋆ <u>ÉTAPE 4 :</u> Concluons en montrant que $J(a) = \inf\limits_{x\in H} J(x)$.
Soient $\beta > \inf\limits_{x\in H} J(x)$ et $C_\beta = \{x \in H \mid J(x) \leqslant \beta\}$.
Comme J est convexe et continue, C_β est convexe et fermé (facile à vérifier).
D'après le théorème de projection sur un convexe fermé, pour tout $u \in H$:

$$\exists ! p(u) \in C_\beta \quad \|u - p(u)\| = \mathrm{d}(u, C_\beta).$$

On sait que $J(x_k) \underset{x\to+\infty}{\longrightarrow} \inf\limits_{x\in H} J(x)$, donc $(J(y_k))_{k\in\mathbb{N}}$ converge vers la même limite. Par définition de C_β :

$$\exists N \in \mathbb{N} \quad \forall k > N \quad y_k \in C_\beta.$$

Le théorème de projection sur un convexe fermé fournit également l'inégalité :

$$\langle a - p(a)|y_k - p(a)\rangle \leqslant 0.$$

Or, d'après l'ÉTAPE 3, $\langle y_k|a - p(a)\rangle \underset{k\to+\infty}{\longrightarrow} \langle a|a - p(a)\rangle$. Donc :

$$\begin{array}{rcl} \|a - p(a)\|^2 & = & \langle a - p(a)|a - p(a)\rangle \\ & = & \lim\limits_{y\to+\infty} \langle a - p(a)|y_k - p(a)\rangle \\ & \leqslant & 0. \end{array}$$

Il vient $a = p(a)$, et donc $a \in C_\beta$, pour tout $\beta > \inf\limits_{x\in H} J(x)$.
Cela prouve finalement que $J(a) = \inf\limits_{x\in H} J(x)$. □

Remarques 2.

- Ce théorème n'a lieu d'être qu'en dimension **infinie**. En effet, en dimension finie (si $H = \mathbb{R}^n$ par exemple), comme J est coercive, on peut ramener l'étude à un fermé borné (donc à un compact), et l'infimum existe et est atteint dès lors que J est continue.
Donc attention, en dimension finie, ce théorème n'a aucun intérêt.

- Un exemple d'application du théorème est donné ci-dessous. Notez qu'il n'apparaît pas dans les références du développement.

[1] Décidément !

Proposition 3. ***Rappels sur les espaces de* Sobolev.**

On note $H^1([0,1])$ l'ensemble des fonctions admettant une dérivée faible. On munit cet espace de la norme $\|\cdot\| : u \longmapsto \|u\| = \sqrt{\|u\|_2^2 + \|u'\|_2^2}$. pose également $H := H_0^1([0,1])$ l'adhérence de $\mathcal{C}_c^\infty([0,1])$ dans $H^1([0,1])$. On rappelle/admet :
(1) que les opérations algébriques et l'intégration par parties sont valables pour les fonctions de $H^1([0,1])$;
(2) que $(H, \|\cdot\|)$ est un $\mathbb{R}$-espace de Hilbert ;
(3) l'inégalité de Poincaré : il existe $C_P > 0$ tel que pour tout $u \in H$, $\|u\|_2 \leqslant C_P \|\nabla u\|_2$. C_P est appelée *constante de* Poincaré.
On pourra trouver de tels rappels (et leurs preuves) dans [B].

Application 4.

Soit $f \in L^2([0,1])$. Alors il existe une unique fonction $u \in H$ solution de $-u'' + u = f$ au sens faible. Ce "sens faible" sera défini dans la démonstration.

Preuve : Raisonnons en plusieurs étapes.
⋆ Étape 1 : Déterminons la formulation variationnelle du problème, autrement dit : définissons le terme "solution faible" employé dans l'énoncé. Soit $\varphi \in \mathcal{C}_c^\infty([0,1])$. Pour tout $u \in \mathcal{C}^2([0,1])$, si, sur $[0,1]$, on a $-u'' + u = f$, alors $-u''\varphi + u\varphi = f\varphi$. En intégrant sur $[0,1]$, on obtient

$$-\int_0^1 u''(x)\varphi(x)\mathrm{d}x + \int_0^1 u(x)\varphi(x)\mathrm{d}x = \int_0^1 f(x)\varphi(x)\mathrm{d}x.$$

En intégrant par parties, on a

$$\int_0^1 u''(x)\varphi(x)\mathrm{d}x = \underbrace{[u'(x)\varphi(x)]_0^1}_{=0 \text{ car } \varphi \in \mathcal{C}_c^\infty([0,1])} - \int_0^1 u'(x)\varphi'(x)\mathrm{d}x,$$

donc

$$\int_0^1 u'(x)\varphi'(x)\mathrm{d}x + \int_0^1 u(x)\varphi(x)\mathrm{d}x = \int_0^1 f(x)\varphi(x)\mathrm{d}x.$$

On obtient la formulation variationnelle du problème : trouver $u \in H$ vérifiant

$$\forall v \in H \quad \int_0^1 u'(x)v'(x)\mathrm{d}x + \int_0^1 u(x)v(x)\mathrm{d}x = \int_0^1 f(x)v(x)\mathrm{d}x. \tag{1}$$

Une telle solution u est appelée *solution faible du problème* $-u'' + u = f$ *sur* $[0,1]$, $u(0) = u(1) = 0$.

Remarque 5.

Ici, on prend bien $u \in H$ (donc pas forcément muni d'une dérivée faible seconde) parce qu'il n'y a que des dérivées premières qui interviennent dans la formulation variationnelle (1).
En réalité, on pourrait presque garder les dérivées secondes u'', mais dans un souci de compréhension vis-à-vis de l'espace H, j'ai préféré les éliminer.

⋆ Étape 2 : Soit $J : H \to \mathbb{R}$ définie par :

$$\begin{aligned} \forall u \in H \quad J(u) &= \int_0^1 \left(\frac{1}{2}u'(x)^2 + \frac{1}{2}u(x)^2 - u(x)f(x)\right)\mathrm{d}x \\ &= \frac{1}{2}\|u'\|_2^2 + \frac{1}{2}\|u\|_2^2 - \int_0^1 u(x)f(x)\mathrm{d}x. \end{aligned}$$

Montrons que J est strictement convexe et continue sur H.
Pour $u \in H$, on pose $Q(u) = \|u\|^2$ et $L(u) = \int_0^1 f(t)u(t)\mathrm{d}t$, de sorte que $J(u) = \frac{1}{2}Q(u) - L(u)$.
On sait que $-L$ est convexe sur H : c'est une application linéaire.
Ensuite, si $t \in]0;1[$ et $u, v \in H$ avec $u \neq v$, on a :

$$Q(tu + (1-t)v) = \|tu + (1-t)v\|^2 = t^2\|u\|^2 + (1-t)^2\|v\|^2 + 2t(1-t)\langle u, v\rangle.$$

Donc :

$$\begin{aligned} tQ(u) + (1-t)Q(v) - Q(tu + (1-t)v) &= t\|u\|^2 + (1-t)\|v\|^2 - t^2\|u\|^2 - (1-t)^2\|v\|^2 \\ &\quad -2t(1-t)\langle u, v\rangle \\ &= t(1-t)\|u\|^2 + (1-t)(1-(1-t))\|v\|^2 \\ &\quad -2t(1-t)\langle u, v\rangle \end{aligned}$$

$$
\begin{aligned}
&= t(1-t)\|u\|^2 + t(1-t)\|v\|^2 - 2t(1-t)\langle u,v\rangle \\
&= t(1-t)(\|u\|^2 - 2\langle u,v\rangle + \|v\|^2) \\
&= t(1-t)\|u-v\|^2 \\
&> 0,
\end{aligned}
$$

et donc Q est strictement convexe sur H.
Ainsi, J est strictement convexe sur H en tant que somme d'une application convexe sur H avec une application strictement convexe sur H.
Elle est également immédiatement continue sur H car Q et $-L$ le sont.
⋆ Étape 3 : Montrons que J est coercive. En partant du fait que $(a-b)^2 \geqslant 0$ pour tout $a, b \in \mathbb{R}$, on peut aisément montrer que $ab \leqslant \dfrac{a^2+b^2}{2}$. On a donc :

$$\|u\|_2\|f\|_2 \leqslant \frac{\|u\|_2^2 + \|f\|_2^2}{2},$$

et donc grâce à l'inégalité de CAUCHY-SCHWARZ,

$$
\begin{aligned}
J(u) &\geqslant \frac{1}{2}\|u'\|_2^2 + \frac{1}{2}\|u\|_2^2 - \|u\|_2^2\|f\|_2^2 \\
&\geqslant \frac{1}{2}\|u'\|_2^2 + \frac{1}{2}\|u\|_2^2 - \frac{\|u\|_2^2 + \|f\|_2^2}{2} \\
&\geqslant \frac{1}{2}\|u'\|_2^2 - \frac{\|f\|_2^2}{2}.
\end{aligned}
$$

Enfin, en utilisant l'inégalité de POINCARÉ,

$$\|u\|^2 = \|u\|_2^2 + \|u'\|_2^2 \leqslant (C_P+1)\|u'\|_2^2.$$

Donc :

$$J(u) \geqslant \frac{1}{2(1+C_P)}\|u\|^2 - \frac{\|f\|_2^2}{2} \underset{\|u\|\to+\infty}{\longrightarrow} +\infty,$$

et J est coercive.
⋆ Étape 4 : Concluons en montrant qu'il existe une solution au sens faible de notre problème, puis qu'elle est unique.
Par un calcul direct, J est différentiable sur H (pour le voir, revenir à la définition de la différentiabilité par exemple) et :

$$
\begin{aligned}
\forall u, v \in H \quad \mathrm{d}J(u)\cdot v &= \int_0^1 (u'(x)v'(x) + u(x)v(x) - v(x)f(x))\mathrm{d}x \\
&= \int_0^1 u'(x)v'(x)\mathrm{d}x + \int_0^1 u(x)v(x)\mathrm{d}x - \int_0^1 f(x)v(x)\mathrm{d}x.
\end{aligned}
$$

On reconnaît les termes de la formulation variationnelle (1) à droite. On obtient l'équivalence suivante : pour tout $u \in H$, on a :

$$[\forall v \in H \quad \mathrm{d}J(u)\cdot v = 0] \iff -u'' + u = f \text{ au sens faible sur } [0,1].$$

Maintenant, comme J est convexe, continue et coercive sur H, d'après le théorème d'optimisation dans un HILBERT, il existe $u \in H$ tel que $J(u) = \inf\limits_{x\in H} J(x)$.
Comme u minimise J, u est un point critique de J, c'est-à-dire qu'on a $\mathrm{d}J(u) = 0$. L'équivalence ci-dessus montre alors que u est solution faible de notre problème.
Enfin, concernant l'unicité, J étant strictement convexe, u est unique, et cela achève la preuve. □

Remarque 6.

On aurait aussi pu appliquer le théorème de LAX-MILGRAM :
Pour $u, v \in H$, on pose $a(u,v) = \displaystyle\int_0^1 u'(x)v'(x)\mathrm{d}x + \int_0^1 u(x)v(x)\mathrm{d}x$ et $\ell(v) = \displaystyle\int_0^1 f(x)v(x)\mathrm{d}x$. On obtient la formulation variationnelle du problème : trouver $u \in H$ tel que

$$\forall v \in H \quad a(u,v) = \ell(v).$$

Alors :
(1) Pour tout $v \in H$, d'après l'inégalité de CAUCHY-SCHWARZ, $|\ell(v)| \leqslant \|f\|_2\|v\|_2$, donc ℓ est une forme linéaire continue.
(2) a est clairement bilinéaire.
(3) Pour tous $u, v \in H$, l'inégalité de CAUCHY-SCHWARZ fournit encore

$$
\begin{aligned}
|a(u,v)| &\leqslant \|u'\|_2\|v'\|_2 + \|u\|_2\|v\|_2 \\
&\leqslant \|u'\|\|v'\| + \|u'\|\|v'\| \\
&\leqslant 2\|u'\|\|v'\|,
\end{aligned}
$$

donc a est continue.

(4) Pour tout $u \in H$, on a directement $a(u,v) = \|u\| \geqslant \|u\|$ donc a est coercive sur H (en tant que forme bilinéaire : ce n'est a priori pas la même notion de coercivité que celle définie plus haut).
D'après le théorème de Lax-Milgram, il existe une unique solution faible u à notre problème.

On remarque une chose : les deux notions de coercivité semblent se rejoindre puisqu'on conclut à une unique solution grâce à celles-ci. D'ailleurs, a étant symétrique, le théorème de Lax-Milgram nous donne également l'existence et l'unicité du problème d'optimisation étudié.
Si je ne l'ai pas fait ici, c'était clairement pour appliquer le théorème. Sinon, en pratique, c'est bien le théorème de Lax-Milgram qu'on applique.