Développement: Optimisation dans un espace de Hilbert

Arthur Maritch-Roy

Dans ce développement, on introduit la notion de convergence faible dans un espace de Hilbert, dans l'optique d'étudier un problème de minimisation qui ne se résout pas "simplement" à l'aide du théorème de Lax-Milgram.

Définition 1. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y \rangle \to \langle x, y \rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème 1. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Remarque 1. En général, on n'a aucune chance qu'une suite bornée admette une sous-suite convergente dans H. En effet, le théorème de Riesz assure qu'une boule n'est jamais compacte en dimension infinie.

Démonstration. On note $(h_n)_n$ une suite dense de H et on note $M = \sup \|x_n\|$. Pour chaque k, la suite $(\langle x_n, h_k \rangle)_n$ est bornée. On peut donc en extraire une sous suite qui converge vers un $\varphi(h_k)$. Par extraction diagonale, on peut exhiber une extractrice qui va faire converger simultanément tous les $(\langle x_n, h_k \rangle)_n, k \in \mathbb{N}$. Notons ψ cette extractrice et montrons alors que pour un $y \in H$, la suite $(\langle x_{\psi(n)}, y \rangle)$ converge, *i.e.* qu'elle est de Cauchy. Pour p et q deux entiers, et $k \in \mathbb{N}$, on a :

$$|\langle x_{\psi(p)}, y \rangle - \langle x_{\psi(q)}, y \rangle| \leq |\langle x_{\psi(p)}, y - h_k \rangle| + |\langle x_{\psi(p)}, h_k \rangle - \langle x_{\psi(q)}, h_k \rangle| + |\langle x_{\psi(q)}, h_k - y \rangle|.$$

Pour $\varepsilon > 0$, en prenant h_k assez proche de y, on peut rendre les termes extrémaux plus petits que ε par Cauchy-Schwarz. Pour le terme du milieu, on peut le rendre plus petit que ε pour $p, q \to +\infty$ puisque la suite $(\langle x_{\psi(p)}, h_k \rangle)_n$ converge donc est de Cauchy. Somme toute, on a bien :

$$\forall y \in H, \langle x_{\psi(p)}, y \rangle \to \varphi(y).$$

Par linéarité de la limite et bilinéarité du produit scalaire, φ est une forme linéaire. Elle est de plus continue (avec M comme module de continuité par Cauchy-Schwarz), donc on peut la représenter par un vecteur $x^* \in H$, ce qui achève la démonstration. $\square$

Théorème 2. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Démonstration. Soit (x_n) une suite minimisante, c'est-à-dire telle que $J(x_n) \to \inf_H J$. Cette suite est bornée, sinon cela contredirait l'hypothèse de coercitivité. Comme c'est une suite bornée, on peut appliquer le résultat du théorème précédent, et exhiber un $x^* \in H$ tel que (quitte à extraire) l'on ait $x_n \rightharpoonup x^*$.

Posons $\alpha > \inf_H J$ et $C_\alpha = J^{-1}(]-\infty, \alpha])$, qui est non vide par définition de la borne inférieure et fermé par continuité de J. On peut aussi voir simplement que la convexité de J entraîne celle de C_α. On peut donc définir p_α le projection sur C_α. Pour k assez grand, x_k appartient à C_α et on a alors, d'après le théorème de projection sur le convexe C_α :

$$\langle x^* - p_\alpha(x^*), x_k - p_\alpha(x^*)\rangle \leq 0.$$

Par continuité du produit scalaire, on obtient, pour k qui tend vers l'infini, $\|x^* - p_\alpha(x^*)\|^2 \leq 0$ soit $x^* \in C_\alpha$. Ceci valant pour tout α, on peut le faire tendre vers $\inf_H J$, ce qui donne $J(x^*) \leq \inf_H J$ et qui permet donc de conclure. □

Application 1. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}_0^1(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens des distributions.

Démonstration. Qui dit au sens faible dit formulation faible. On considère la fonctionnelle J définie sur $\mathrm{H}_0^1(0,1)$ définie par :

$$J : u \longmapsto \int_0^1 \left(\frac{(u')^2}{2} + \frac{|u|^{p+1}}{p+1} - fu\right)$$

Pour alléger les notations, on note $\|\cdot\|$ la norme sur H_0^1 et $\|\cdot\|_2$ la norme L^2.

Soient $u, v \in \mathrm{H}_0^1(0,1)$:

$$\begin{aligned} J(u+v) &= \int_0^1 \left(\frac{(u'+v')^2}{2} + \frac{|u+v|^{p+1}}{p+1} - f(u+v)\right) \\ &J(u) + \int_0^1 \left(u'v' + \frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - fv\right) + \frac{1}{2}\|v'\|_2^2 \\ &= J(u) + \int_0^1 \left(u'v' + |u|^{p-1}uv - fv\right) + \int_0^1 \left(\frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - |u|^{p-1}uv\right) + o(\|v\|). \end{aligned}$$

Montrons donc que $\varepsilon_u(v) := \int_0^1 \left(\frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - |u|^{p-1}uv\right) = o(\|v\|)$.

Pour x, y deux *réels*, on pose $\varphi : t \mapsto \varepsilon_x(ty)$. C'est une fonction de t dérivable avec :

$$\varphi'(t) = y(p+1)\sigma(x+ty)|x+ty|^p - x|x|^{p-1}y,$$

où σ est la fonction signe, qui vaut par convention 0 en 0. Ainsi d'après l'inégalité des accroissements finis sur le segment $[0,1]$:

$$\left|\frac{|x+y|^{p+1}}{p+1} - \frac{|x|^p}{p+1} - |x|^{p-1}xy\right| \leq |y| \sup_{|t|\leq|y|}\{|\sigma(x+t)|x+t|^p| - \sigma(x)|x|^p|\}.$$

En rentrant cette inégalité dans l'intégrale, on a alors :

$$\begin{aligned} \left|\int_0^1 \left(\frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - |u|^{p-1}uv\right)\right| &\leq \int_0^1 |v| \sup_{|t|\leq\|v\|_\infty}\{|\sigma(u+t)|u+t|^p| - \sigma(u)|u|^p|\} \\ &\leq \|v\|_2 \sqrt{\int_0^1 (\sup_{|t|\leq\|v\|_\infty}\{|\sigma(u+t)|u+t|^p| - \sigma(u)|u|^p|\})^2} \\ &\leq \|v\| \sqrt{\int_0^1 (\sup_{|t|\leq\|v\|_\infty}\{|\sigma(u+t)|u+t|^p| - \sigma(u)|u|^p|\})^2}. \end{aligned}$$

Il nous reste à montrer que cette dernière intégrale tend vers 0 lorsque $\|v\|$ tend vers 0. Par continuité de la fonction $x \mapsto \sigma(x)|x|^p$, la borne supérieure dans l'intégrale converge simplement vers 0 quand $\|v\|$ tend vers 0. On peut alors supposer $\|v\| \leq 1$ dans perte de généralité et alors dominer l'intégrande par $((|u|+1)^p + |u|^p)^2 \in \mathrm{L}^1(0,1)$ car u admet un représentant continu. Par convergence dominée, on a la convergence vers 0. Ainsi, la différentielle de J en u vaut :

$$\mathrm{d}J_u(v) = \int_0^1 \big(u'v' + |u|^{p-1}uv - fv\big).$$

En effet, on vient de voir que l'on avait le développement à l'ordre un souhaité, il faut maintenant voir que c'est bien une forme linéaire continue. Le caractère linéaire est évident. Pour la continuité :

$$\begin{aligned}
|\mathrm{d}J_u(v)| &\leq \int_0^1 (|u'v'| + |u|^{p-1}uv + |fv|) \\
&\leq \sqrt{\int_0^1 |u'|^2}\|v'\|_2 + \sqrt{\int_0^1 |u|^{2p}}\|v\|_2 + \|f\|_2\|v\|_2 \\
&\leq C\|v\|,
\end{aligned}$$

puisque toutes les intégrales convergent (la première car $u' \in \mathrm{L}^2$, la deuxième car u admet un représentant continu).

Montrons maintenant que J est strictement convexe. L'intégrande est convexe en u ainsi il est facile de voir que J est convexe. Si on a égalité pour un $\lambda \in]0,1[$ de $J(\lambda u + (1-\lambda)v)$ et $\lambda J(u) + (1-\lambda)J(v)$, alors on a :

$$\frac{1}{p+1}\int_0^1 (|\lambda u + (1-\lambda v)|^{p+1} - \lambda|u|^{p+1} - (1-\lambda)|v|^{p+1}) = \frac{1}{2}\int_0^1 (\lambda u'^2 - (\lambda u' + (1-\lambda)v')^2).$$

Par convexité du carré et de $|\cdot|^{p+1}$ sur $\mathbb{R}$, la première intégrande est négative presque partout et la deuxième positive presque partout. Nécessairement, les deux intégrandes sont nulles presque partout. En particulier, on a presque partout, et même partout par continuité :

$$|\lambda u + (1-\lambda)v|^{p+1} = \lambda|u|^{p+1} + (1-\lambda)|v|^{p+1}.$$

On a alors $u = v$ puisque $t \mapsto |t|^{p+1}$ est strictement convexe.

Enfin, J est coercitive puisque :

$$J(u) \geq \frac{1}{2}\int_0^1 u'^2 - \|u\|_2\|f\|_2 \geq C\|u\|^2 - \|u\|_2\|f\|_2 \xrightarrow[\|u\|\to\infty]{} +\infty,$$

où l'on a utilisé l'inégalité de Poincaré. D'après le théorème précédent, J admet donc un minimiseur u^*. En écrivant la condition nécessaire de minimalité, on retombe sur notre problème de départ au sens faible.

Pour l'unicité, il faut voir que si on a un autre u qui minimise, alors la différentielle de J en ce u s'annule sur l'espace des fonctions test. Par continuité de la différentielle et densité des fonctions test dans H_0^1, u est un point critique de J. Par stricte convexité de J, c'est forcément u^*. $\square$

Références : la preuve de Banach-Alaoglu doit être dans Brezis, *Analyse Fonctionnelle*. L'application provient de la page web de Thomas Cavallazzi.