Des $\mathcal{L}^p$ aux L^p

Résumé

On a vu que, sur un espace mesuré $(X, \mathscr{T}, \mu)$, l'ensemble

$$\mathcal{L}^p(X) = \left\{ f : X \to \mathbb{R} \text{ mesurable }, \int_X |f|^p d\mu < \infty \right\},$$

muni de l'application $N_p : \mathcal{L}^p(X) \to \mathbb{R},\ f \mapsto \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}}$, était *presque* un espace vectoriel normé. On va voir ici comment utiliser les relations d'équivalences pour construire un espace vectoriel sur lequel N_p soit bel et bien une norme.

Plus précisément, pour $p \in [1, \infty]$, on a obtenu que $\mathcal{L}^p$ est un espace vectoriel, tel que

- $N_p(f) \geq 0$ pour tout $f \in \mathcal{L}^p$;
- $N_p(\lambda f) = |\lambda| N_p(f)$ pour tout $f \in \mathcal{L}^p$, pour tout $\lambda \in \mathbb{R}$;
- $N_p(f+g) \leq N_p(f) + N_p(g)$, pour tous $f, g \in \mathcal{L}^p$;
- $N_p(f) = 0$ si, et seulement si, $f = 0$ μ-presque partout. Autrement dit, si, et seulement si $\mu(\{x \in X,\ f(x) \neq 0\}) = 0$.

On dit que N_p est une *semi-norme* sur $\mathcal{L}^p(X)$. Pour que N_p définisse une norme sur $\mathcal{L}^p(X)$, il faudrait avoir $N_p(f) = 0$ si, et seulement si, f est la fonction nulle, autrement dit $f = 0$ *partout*. Comment contourner ce problème ?

1 Relations d'équivalence, classes d'équivalence

Définition 1. *Une relation $\sim$ sur un ensemble E est une* relation d'équivalence *si elle est*

1. réflexive : *pour tout $x \in E$, $x \sim x$;*
2. symétrique : *pour tous $x, y \in E$, $x \sim y$ ssi $y \sim x$.*
3. transitive : *pour tous $x, y, z \in E$, si $x \sim y$ et $y \sim z$, alors $x \sim z$.*

Exemple 2.

1. L'égalité est une relation d'équivalence sur $\mathbb{R}$, mais la relation $\geq$ ne l'est pas : $x \geq y$ n'implique pas $y \geq x$.
2. Sur $\mathbb{N}$, si on fixe $p \in \mathbb{N}^*$, la relation $n \equiv m$ ssi $n - m$ est divisible par p est une relation d'équivalence[1].
3. La relation "$D \sim D'$ ssi D et D' sont parallèles ou confondues" est une relation d'équivalence sur les droites du plan.
4. Sur $M_n(\mathbb{R})$, la relation "$A \sim B$ ssi il existe une matrice inversible P telle que $B = P^{-1}AP$" est une relation d'équivalence.
5. La relation d'équivalence des normes sur un e.v.n. E est... une relation d'équivalence sur l'ensemble des normes sur E.

Et, ce qui nous intéresse d'avantage :

1. Cette relation, appelée "congruence", est aux fondements de la théorie des nombres et de la cryptographie

Proposition 3. *Soit $(X, \mathscr{T}, \mu)$ un espace mesuré. La relation $\sim$ définie sur $\mathcal{L}^p(X)$ par "$f \sim g$ ssi $f = g$ μ-presque partout" est une relation d'équivalence sur $\mathcal{L}^p$.*

Preuve. On vérifie les trois points de la définition 1.

- *Réflexivité* : Pour tout $f \in \mathcal{L}^p(X)$, l'ensemble $\{x \in X, f(x) \neq f(x)\} = \emptyset$ est de mesure nulle, donc $f \sim f$.
- *Symétrie* : Pour tous $f, g \in \mathcal{L}^p(X)$, si $f \sim g$, alors $\mu(\{x \in X f(x) \neq g(x)\}) = 0$, donc $\mu(\{x \in X g(x) \neq f(x)\}) = 0$, donc $g \sim f$.
- *Transitivité :* Supposons que $f \sim g$ et $g \sim h$. Alors $\{x \in X, f(x) \neq h(x)\} \subset \{x \in X, f(x) \neq g(x)\} \cup \{x \in X, g(x) \neq h(x)\}$, donc

$$\mu(\{x \in X, f(x) \neq h(x)\}) \leq \mu(\{x \in X, f(x) \neq g(x)\}) + \mu(\{x \in X, g(x) \neq h(x)\}) = 0,$$

  autrement dit, $f \sim h$. □

Une fois qu'on a une relation d'équivalence $\sim$ sur un ensemble E, on peut regrouper les éléments équivalents entre eux dans des paquets appelés classes d'équivalence :

Définition 4. *Soit $\sim$ une relation d'équivalence sur E et $x \in E$. On appelle* classe d'équivalence de x *(pour la relation $\sim$) l'ensemble $[x] = \{y \in E,\ x \sim y\}$.*

On a alors :

Proposition 5. *Les classes d'équivalence forment une partition de E. Autrement dit, chaque élément de E est dans un unique paquet.*

Démonstration. Soit $x \in E$. Alors $x \in [x]$, donc x est dans au moins un paquet.

Supposons que x soit aussi dans le paquet $[y]$, et montrons que c'est le même paquet, autrement dit que $[x] = [y]$. Puisque $x \in [y]$, on a $y \sim x$ et, par symétrie, $x \sim y$. Donc :

- si $z \in [x]$, alors $x \sim z$. Mais $y \sim x$ donc, par transitivité, $y \sim z$, donc $z \in [y]$
- si $z \in [y]$ alors $y \sim z$. Mais on a $x \sim y$ donc, par transitivité, $x \sim z$, donc $z \in [x]$.

On a donc bien $[x] = [y]$. □

Dans le cas de la relation d'égalité μ-presque partout, on a, pour $f \in \mathcal{L}^p(X)$,

$$[f] = \{g \in \mathcal{L}^p(X),\ f - g = 0\ \mu - p.p.\} = \{f + z,\ z = 0\ \mu - p.p.\}$$

Au lieu de regarder l'ensemble des éléments de E, on peut alors regarder l'ensemble des paquets :

Définition 6. *Soit $\sim$ une relation d'équivalence sur un ensemble E. Alors l'ensemble*

$$E/\sim\ = \{[x], x \in E\}$$

est appelé ensemble quotient *de E par $\sim$.*

Les ensembles de Lebesgue sont alors les quotients des $\mathcal{L}^p(X)$ pour la relation d'égalité μ-p.p. :

Définition 7. *Pour $p \in [1, \infty]$, on note $L^p = \mathcal{L}^p(X)/\sim$, où $\sim$ est la relation d'égalité μ-presque partout décrite à la proposition 3. Autrement dit,*

$$L^p(X) = \{[f],\ f \in \mathcal{L}^p(X)\}.$$

Les éléments de $L^p(X)$ ne sont donc pas des fonctions à proprement parler, mais des “paquets” de fonctions toutes égales μ-presque partout. Souvent, on écrit “Soit $f \in L^p(X)$” ; c’est un abus de langage, on veut dire par là qu’on choisit un paquet $A \in L^p$ et on le représente par une des fonctions du “paquet” $A = [f]$. Deux représentants différents sont égaux presque partout, mais peuvent différer sur un ensemble de mesure nulle.

Par exemple, dans le cas de la mesure de Lebesgue λ sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, les singletons $\{x\}$ sont de mesure nulle, donc deux représentants d’une même classe $C \in L^p(\mathbb{R})$ peuvent prendre des valeurs différentes en un x donné. On ne peut donc pas parler de la valeur en un point d’une “fonction” de $L^p(\mathbb{R})$.

En revanche, sur $\mathbb{N}$ muni de la mesure de comptage, cette distinction ne change rien :

Exercice : Montrer que sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ où μ est la mesure de comptage, si $f = g$ $\mu-$p.p alors $f(n) = g(n)$ pour tout $\mathbb{N}$. En déduire que chaque classe de $L^p(\mathbb{N})$ contient un seul élément.

Montrons qu’on peut munir $L^p(X)$ d’une structure d’espace vectoriel. On aimerait poser

$$\begin{cases} [f] + [g] = [f+g] \\ \lambda \cdot [f] = \lambda[f] \end{cases}$$

Il peut être utile de prendre le temps de bien comprendre ce qu’on entend par là.

Pour l’addition, on prend deux classes C_1 et C_2 dans $L^p(X)$, et on veut définir la somme de C_1 et C_2. Pour cela, on prend un représentant $f \in C_1$ (autrement dit, $C_1 = [f]$) et $g \in C_2$ (i.e. $C_2 = [g]$). Alors $f, g \in \mathcal{L}^p(X)$ sont des fonctions, on peut donc définir leur somme $f + g \in \mathcal{L}^p(X)$. On pose alors $C_1 + C_2 = [f + g] \in L^p(X)$. Il faut donc vérifier que si on avait choisi d’autres représentants $\tilde{f} \in C_1$, $\tilde{g} \in C_2$, on aurait obtenu le même résultat pour la somme $C_1 + C_2$; autrement dit, on doit vérifier que si $[f] = C_1 = [\tilde{f}]$ et $[g] = C_2 = [\tilde{g}]$ alors $[f + g] = [\tilde{f} + \tilde{g}]$.

De même, pour la multiplication scalaire, on souhaite définir $\lambda \cdot C$ pour $\lambda \in \mathbb{R}$ et $C \in L^p(X)$. On prend un représentant $f \in C$: c’est une fonction de $\mathcal{L}^p(X)$, donc on peut la multiplier par λ pour obtenir la fonction $\lambda f \in \mathcal{L}^p(X)$. On pose alors $\lambda \cdot C = [\lambda f] \in L^p(X)$. Et, comme précédemment, il faut alors vérifier que si $\tilde{f}$ est un autre représentant de C, alors $[\lambda \tilde{f}] = [\lambda f]$.

Il faut donc montrer :

Lemme 8. *Soient $f, g, \tilde{f}, \tilde{g} \in \mathcal{L}^p(X)$, $\lambda \in \mathbb{R}$. Alors :*

1. *Si $[f] = [\tilde{f}]$ et $[g] = [\tilde{g}]$ alors $[f + g] = [\tilde{f} + \tilde{g}]$.*
2. *Si $[f] = [\tilde{f}]$ alors $[\lambda f] = [\lambda \tilde{f}]$.*

Preuve.

1. Supposons que $[f] = [\tilde{f}]$ et $[g] = [\tilde{g}]$. Alors $A = \{x \in X, f(x) \neq \tilde{f}(x)\}$ et $B = \{x \in X, g(x) \neq \tilde{g}(x)\}$ sont tous deux de mesures nulle. Or,
$$\{x \in X, (f + g)(x) \neq (\tilde{f} + \tilde{g})(x)\} \subset A \cup B,$$
fonc $\mu(\{x \in X, (f + g)(x) \neq (\tilde{f} + \tilde{g})(x)\}) \leq \mu(A) + \mu(B) = 0$. Donc $f + g$ et $\tilde{f} + \tilde{g}$ sont égales $\mu-$presque partout, autrement dit, $[f + g] = [\tilde{f} + \tilde{g}]$.
2. Si $\lambda = 0$, c’est immédiat. Sinon, $f(x) = \tilde{f}(x)$ équivaut à $\lambda f(x) = \lambda \tilde{f}(x)$ donc $\{x \in X, (\lambda f)(x) \neq (\lambda \tilde{f})(x)\} = \{x \in X, f(x) \neq \tilde{f}(x)\}$ et est donc de mesure nulle. Donc, quel que soit λ, λf et $\lambda \tilde{f}$ sont égales μ-presque partout, donc $[\lambda f] = [\lambda \tilde{f}]$. □

Corollaire 9. *Muni de ces deux lois, $L^p(\mu)$ est un espace vectoriel.*

Remarque 10. Dans l’espace vectoriel $L^p(\mu)$, le vecteur nul est alors la classe $[0]$, qui contient la fonction nulle mais aussi toutes les fonctions nulles presque partout.

On souhaite aussi que N_p définisse une norme sur $L^p(\mu)$. Comme pour les lois qui font de $L^p(\mu)$ un espace vectoriel, il nous faut définir $N_p(C)$, où C est un élément de $L^p(\mu)$, c’est-à-dire une classe d’équivalence. On aimerait procéder comme suit :

- on choisit un représentant $f \in C$: autrement dit $f \in \mathcal{L}^p(\mu)$ et $[f] = C$

- on pose $N_p(C) = N_p([f]) = N_p(f) = \int_X |f|^p d\mu$.

Comme précédemment, il s'agit de vérifier que le résultat ne dépend pas de quel représentant de C on choisit. Autrement dit, on doit vérifier que si $[f] = C = [\tilde{f}]$, alors $N_p(f) = N_p(\tilde{f})$. Or si $[f] = [\tilde{f}]$ alors $f = \tilde{f}$ μ-presque partout, donc $|f|^p - |\tilde{f}|^p = 0$ μ-presque partout. Donc

$$N_p(f)^p - N_p(\tilde{f})^p = \int_X |f|^p - |\tilde{f}|^p d\mu = 0$$

donc $N_p(f) = N_p(\tilde{f})$.

Alors N_p définit une norme sur l'espace vectoriel $L^p(\mu)$. On a déjà vu que

1. $N_p([0]) = 0$
2. Pour tout $\lambda \in \mathbb{R}$, pour tout $[f] \in L^p(\mu)$, $N_p(\lambda[f]) = \lambda N_p(f)$;
3. Pour tous $[f], [g] \in L^p(\mu)$, $N_p([f] + [g]) \leq N_p([f]) + N_p([g])$.

Il reste à vérifier que si $N_p([f] = 0$, alors $[f] = [0]$. Or si $N_p([f]) = 0$ alors $\int_X |f|^p d\mu = 0$, donc $|f|^p = 0$ μ-presque partout, donc $f = 0$ μ-presque partout. Autrement dit, $[f] = [0]$.

On a en fait mieux que ça :

Théorème (Riesz-Fisher). *Pour tout $p \geq 1$, $(L^p(\mu), N_p)$ est un espace vectoriel normé complet.*

Preuve. Soit $([f_n])_n$ une suite de Cauchy dans $L^p(\mu)$, que l'on identifie à ses représentants $f_n \in \mathcal{L}^p(\mu)$. On cherche $f \in \mathcal{L}^p(\mu)$ telle que $N_p(f_n - f) \xrightarrow{n \to \infty} 0$.

Puisque $(f_n)_n$ est de Cauchy, pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que pour tous $p, q \geq N$, $N_p(f_p - f_q) < \varepsilon$. En appliquant ceci avec $\varepsilon = \frac{1}{2^k}$, on construit par récurrence une sous-suite $(f_{n_k})_k$ de (f_n) telle que, pour tout $k \in \mathbb{N}$:

$$N_p(f_{n_{k+1}} - f_{n_k}) < \frac{1}{2^k}.$$

Il nous suffit de montrer que cette sous-suite a une limite dans $L^p(\mu)$: en effet, ce sera alors une valeur d'adhérence de la suite de Cauchy $(f_n)_n$, et une suite de Cauchy qui a une valeur d'adhérence converge vers cette valeur d'adhérence.

Posons, pour $N \in \mathbb{N}$,

$$g_N : x \in X \mapsto \sum_{k=0}^{N} |f_{n_{k+1}}(x) - f_{n_k}(x)|$$

Alors pour tout N, $g_N \in \mathcal{L}^p(\mu)$ et, par l'inégalité de Minkowski,

$$N_p(g_N) \leq \sum_{k=0}^{N} N_p(f_{n_{k+1}} - f_{n_k}) < 2.$$

Ainsi, $(g_N^p)_N$ est une suite croissante de fonctions mesurables, donc, par convergence monotone,

$$\int_X \lim_{N \to \infty} g_N^p \, d\mu = \lim_{N \to \infty} \int_X g_N^p \, d\mu \leq 2^p \tag{1}$$

En d'autres termes, $\lim g_N^p$ est intégrable, donc finie μ-presque partout : l'ensemble

$$E = \{x \in X, \lim_{N \to \infty} g_N^p = \infty\}$$

est de mesure nulle. Posons alors

$$g(x) = \begin{cases} \sum_{k=0}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| \text{ si } x \in X \setminus E \\ 0 \text{ sinon.} \end{cases}$$

On a donc $g \in \mathcal{L}^p(\mu)$ et $N_p(g) \leq 2$.

Remarquons que pour tout $k \in \mathbb{N}$, $f_{n_k}(x) = f_{n_0}(x) + \sum_{i=0}^{k-1} f_{n_{i+1}}(x) - f_{n_i}(x)$, et sur $X \setminus E$, la somme converge absolument, donc converge, quand $k \to \infty$.

Posons donc

$$f(x) = \begin{cases} \lim_{k\to\infty} f_{n_k}(x) \text{ si } x \in X \setminus E \\ 0 \text{ sinon.} \end{cases}$$

Alors $f_{n_n}(x) \xrightarrow{k\to\infty} f(x)$ μ-presque partout et, pour tout k, $|f_{n_k}|$ est majoré par la fonction intégrable $|f_{n_0}| + g$. Donc, par le théorème de convergence dominée dans $\mathcal{L}^p(\mu)$, $f \in \mathcal{L}^p(\mu)$ et $N_p(f_{n_k} - f) \xrightarrow{k\to\infty} 0$.

Donc f est une valeur d'adhérence de $(f_n)_n$ dans l'espace vectoriel normé $L^p(\mu)$. Ceci implique $N_p(f_n - f) \xrightarrow{n\to\infty} 0$, comme on le souhaitait. □

En bonus, au cours de la preuve, on a montré le résultat suivant :

Proposition 12. *Soit $p \geq 1$, $(f_n)_n$ une suite qui converge vers f dans $L^p(\mu)$. Alors $(f_n)_n$ admet une sous-suite qui converge simplement vers f μ-presque partout.*