Sur le problème des moments

1 Un théorème de représentation

Dans toute la suite, $\mathcal{X}$ désigne un espace de Hausdorff (i.e. un espace topologique séparé) localement compact, par exemple l'espace euclidien $\mathbb{R}^d$. On désigne par $C(\mathcal{X}, \mathbb{R})$ l'ensemble des fonctions continues sur $\mathcal{X}$ et $C_c(\mathcal{X}, \mathbb{R})$ l'ensemble des fonctions continues sur $\mathcal{X}$ à support compact. Si E est un sous-ensemble de $C(\mathcal{X}, \mathbb{R})$, on notera

$$E_+ := \{f \in E,\, f(x) \geq 0,\, \forall x \in \mathcal{X}\}.$$

Définition 1.1 (Sous-espace adapté)**.** *On dit qu'un sous-espace vectoriel de E de $C(\mathcal{X}, \mathbb{R})$ est adapté si les conditions suivantes sont satisfaites*

1. *$E = E_+ - E_+$;*
2. *pour tout $x \in \mathcal{X}$, il existe $f_x \in E_+$ telle que $f_x(x) > 0$;*
3. *pour tout $f \in E_+$, il existe $g \in E_+$ telle que pour tout $\varepsilon > 0$, il existe un compact K_ε de $\mathcal{X}$ tel que $|f(x)| \leq \varepsilon g(x)$ pour $x \in \mathcal{X} \backslash K_\varepsilon$.*

Exemple 1.1. *Dans le cas où $\mathcal{X}$ est un fermé de $\mathbb{R}$, l'ensemble E des fonctions polynomiales restreintes à $\mathcal{X}$ est un sous-espace adapté de $C(\mathcal{X}, \mathbb{R})$. Pour la condition 1), on écrit simplement*

$$f(x) = \frac{1}{4}\left((f+1)^2 - (f-1)^2\right).$$

Pour la condition 2), on peut prendre le polynôme constant égal à 1. La condition 3) signifie grossièrement que $|f(x)/g(x)|$ tend vers zéro à l'infini. Ainsi, f est un polynôme positif sur $\mathcal{X}$, la fonction $g(x) := x^2 f(x)$ est un polynôme positif et l'on a bien $|f(x)/g(x)| = o(1)$ à l'infini.

Lemme 1.1. *Si E est un sous-espace vectoriel adapté de $C(\mathcal{X}, \mathbb{R})$, alors pour tout $f \in C_c(\mathcal{X}, \mathbb{R})_+$, il existe une fonction $g \in E_+$ telle que $g(x) \geq f(x)$ pour tout $x \in \mathcal{X}$.*

Démonstration. D'après la définiton d'espace adapté, pour tout $x_0 \in \mathcal{X}$, il existe $g_0 \in E_+$ avec $g_0(x_0) > 0$. Quitte à multiplier la fonction g_0 par un scalaire positif assez grand, on peut même supposer que $g_0(x_0) > f(x_0)$. Par continuité, cette inégalité reste vraie dans un voisinage ouvert V_0 de x_0, i.e. pour tout $x \in V_0$, on a $g_0(x) > f(x)$. Comme f est à support compact, on peut recouvrir son support par un nombre fini de tels voisinages $(V_i)_{0 \leq i \leq n}$ et sur chacun d'entre eux, il existe une fonction $g_i \in E_+$, telle que $g_i(x) > f(x)$. On pose alors $g(x) := \sum_{i=0}^n g_i(x)$. On a alors bien $g \in E_+$ et $g(x) \geq f(x)$ pour tout $x \in \mathcal{X}$. □

On rappelle maintenant une version du théorème d'extension de Hahn–Banach.

Lemme 1.2 (Hahn-Banach)**.** *Soit E un sous-espace vectoriel d'un espace vectoriel réel F, et soit C un cône convexe de F tel que $F = E + C$. Alors toute fonctionnelle linéaire L sur E qui est positive sur $C \cap E$, s'étend en une fonctionnelle linéaire $C-$positive $\widetilde{L}$ sur l'espace F tout entier.*

On a alors le théorème suivant de représentation des fonctionnelles linéaires.

Théorème 1.1. *Soit E un sous-espace vectoriel adapté de $C(\mathcal{X},\mathbb{R})$. Pour toute fonctionnelle linéaire $L : E \to \mathbb{R}$, les assertions suivantes sont équivalentes :*

1. *La fonctionnelle L est E_+-positive, i.e. $L(f) \geq 0$ pour tout $f \in E_+$;*
2. *Il existe une mesure positive sigma-finie μ sur $\mathcal{X}$ telle que*

$$L(f) = \int_{\mathcal{X}} f(x)\mu(dx), \quad \textit{pour tout } f \in E.$$

Démonstration. L'implication 2) $\implies$ 1) est claire. Pour l'implication 1) $\implies$ 2), on introduit l'ensemble

$$\widetilde{E} := \{f \in C(\mathcal{X},\mathbb{R}),\, \exists g \in E,\, \forall x \in \mathcal{X},\, |f(x)| \leq g(x)\}.$$

Remarquons que $E_+ \subset \widetilde{E}$ (il suffit de prendre $g = f$) et même $E \subset \widetilde{E}$ puisque si $f \in E$ s'écrit $f = f_1 - f_2$ avec $f_1, f_2 \in E_+$ alors $g = f_1 + f_2 \in E_+ \subset E$ et $|f| \leq g$. On a ainsi $E + (\widetilde{E})_+ \subset \widetilde{E}$. Montrons que l'on a en fait $\widetilde{E} = E + (\widetilde{E})_+$. Soit $f \in \widetilde{E}$ et $g \in E$ telle que $|f| \leq g$. On a alors $f + g \in (\widetilde{E})_+$, $-g \in E$ car E est un espace vectoriel et on peut écrire $f = -g + (f + g)$, d'où l'inclusion réciproque.

D'après la condition 1) et le lemme 1.2 (Hahn–Banach), la fonctionnelle linéaire L positive sur E_+ s'étend en une fonctionnelle linéaire $(\widetilde{E})_+-$positive sur $\widetilde{E}$ tout entier. Par ailleurs, d'après le lemme 1.1, on a l'inclusion $C_c(\mathcal{X},\mathbb{R}) \subset \widetilde{E}$. D'après le théorème de représentation de Riesz, on déduit alors qu'il existe une mesure sigma-finie μ telle que

$$\widetilde{L}(f) = \int_{\mathcal{X}} f(x)\mu(dx), \;\; \forall f \in C_c(\mathcal{X},\mathbb{R}).$$

Rappelons que par définition d'espace adapté, on a $E = E_+ - E_+$. Pour conclure, il suffit de montrer que toute fonction $f \in E_+$ est $\mu-$intégrable et vérifie

$$L(f) = \widetilde{L}(f) = \int_{\mathcal{X}} f(x)\mu(dx).$$

Soit ainsi $f \in E_+$, on pose

$$\mathcal{U} := \{\eta \in C_c(\mathcal{X},\mathbb{R}),\, 0 \leq \eta(x) \leq 1,\, \forall x \in \mathcal{X}\}.$$

Naturellement, pour $\eta \in \mathcal{U}$, on a $f\eta \in C_c(\mathcal{X},\mathbb{R})$ de sorte que

$$\widetilde{L}(f\eta) = \int_{\mathcal{X}} f(x)\eta(x)\mu(dx).$$

Par positivité (et donc monotonie) de $\widetilde{L}$, on déduit alors que

$$\int_{\mathcal{X}} f(x)\mu(dx) = \sup_{\eta\in\mathcal{U}} \int_{\mathcal{X}} f(x)\eta(x)\mu(dx) = \sup_{\eta\in\mathcal{U}} \widetilde{L}(f\eta) \leq \widetilde{L}(f) = L(f) < +\infty.$$

Ainsi, f est $\mu-$intégrable. D'après la dernière égalité, pour conclure, il suffit de montrer que $L(f) \leq \int f d\mu$. D'après le point 3) de la définition 1.1, il existe $g \in E_+$ telle que pour tout $\varepsilon > 0$, il existe un compact K_ε tel que $|f/g| \leq \varepsilon$ sur $\mathcal{X}\backslash K_\varepsilon$. Choisissons alors $\eta_\varepsilon \in \mathcal{U}$ valant 1 sur K_ε. Alors $f \leq \varepsilon g + f\eta_\varepsilon$ et puisque $f\eta_\varepsilon \leq f$

$$L(f) = \widetilde{L}(f) \leq \varepsilon\widetilde{L}(g) + \widetilde{L}(f\eta_\varepsilon) = \varepsilon L(g) + \int_{\mathcal{X}} f\eta_\varepsilon d\mu \leq \varepsilon L(g) + \int f d\mu.$$

En faisant tendre ε vers zéro, on a bien $L(f) \leq \int f d\mu$, d'où le résultat. $\square$

2 Fonctions polynomiales positives

Comme on l'a vu dans l'exemple 1.1, les fonctions polynomiales forment un sous-espace adapté des fonctions continues. Plus généralement, si F est un fermé de $\mathbb{R}^d$, alors les restrictions à F des fonctions polynomiales (identifiées à $\mathbb{R}[x_1,\ldots,x_d]_{|F}$) forment un sous-espace adapté de $C(F,\mathbb{R})$. On note

$$\mathrm{Pos}(F) := \{p \in \mathbb{R}[x_1,\ldots,x_d],\, p(x_1,\ldots,x_d) \geq 0,\, \forall x \in F\}.$$

Le théorème 1.1 dans ce contexte porte le nom de théorème d'Haviland

Théorème 2.1 (Haviland)**.** *Soit F un fermé de $\mathbb{R}^d$ et L une fonctionnelle linéaire sur l'ensemble des fonctions polynomiales identifiées à $\mathbb{R}[x_1,\ldots,x_d]$. Les assertions suivantes sont équivalentes :*

1. *$L(f) \geq 0$ pour tout $f \in Pos(F)$;*
2. *Il existe une mesure positive sigma-finie μ sur $\mathbb{R}^d$, supportée dans F telle que*

$$L(f) = \int_F f(x)\mu(dx), \quad \forall f \in \mathbb{R}[x_1,\ldots,x_d].$$

Lorsque F est un intervalle de $\mathbb{R}$, il n'est pas difficile d'expliciter l'ensemble correspondant $\mathrm{Pos}(F)$.

Lemme 2.1. *On désigne ici par $\Sigma\mathbb{R}[x]^2$ l'ensemble des sommes finies de carrés de polynômes.*

1. *$Pos(\mathbb{R}) = \Sigma\mathbb{R}[x]^2$;*
2. *$Pos(\mathbb{R}^+) = \{f + xg,\ f,g \in \Sigma\mathbb{R}[x]^2\}$;*
3. *$Pos([a,b]) = \{f + (x-a)g,\ f,g \in \Sigma\mathbb{R}[x]^2\}$;*

Démonstration. Si $p \in \mathbb{R}[x]$, il est scindé sur $\mathbb{C}$ et on peut écrire sa factorisation

$$p(x) = \alpha \underbrace{\prod_{i=1}^{r}(x-\lambda_i)^{n_i}}_{\text{racines réelles}} \underbrace{\prod_{j=1}^{k}(x-\mu_j)^{m_j}}_{\text{racines complexes}} = \alpha \prod_{i=1}^{r}(x-\lambda_i)^{n_i} \prod_{j=1}^{k}\left((x-\Re(\mu_j))^2 + \Im(\mu_j)^2\right)^{m_j}. \tag{1}$$

Si p est positif sur $\mathbb{R}$, nécessairement les exposants n_i ci-dessus sont pairs et par ailleurs l'identité

$$(a^2+b^2)(c^2+d^2) = (ac-bd)^2 + (ad+bc)^2$$

permet d'écrire un produit de somme de carrés comme une somme de carrés, aussi p s'écrit bien comme une somme de carrés. Pour le point 2), si on désigne par $Q := \{f + xg,\ f,g \in \Sigma\mathbb{R}[x]^2\}$, on a clairement $Q \subset \mathrm{Pos}(\mathbb{R}^+)$. Ensuite on remarque que Q est stable par multiplication car on peut écrire

$$(f_1 + xf_2)(g_1 + xg_2) = (f_1f_2 + x^2g_1g_2) + x(f_2g_1 + f_1g2).$$

Aussi, il suffit de montrer que tous les facteurs dans la décomposition (1) sont dans Q. C'est clair pour les facteurs quadratiques issues des racines complexes et ainsi que pour les racines réelles lorsque les n_i sont pairs. Ensuite, si $p \in \mathrm{Pos}(\mathbb{R}^+)$ en faisant $x \to +\infty$, on a nécessairement $\alpha \geq 0$. Enfin, pour les puissances impaires des facteurs $(x-\lambda_i)$, en prenant x au voisinage de zéro, on obtient $\lambda_i \leq 0$ de sorte que $(x-\lambda_i) = (-\lambda_i) + x \times 1 \in Q$. La preuve de 3) est similaire mais un peu plus technique, cf [Sch17]. On a en fait plus précisément (en prescrivant les degrés)

$$\begin{aligned}&\mathrm{Pos}([a,b])_{2n} = \{f + (x-a)(b-x)g,\ f \in \Sigma_n\mathbb{R}[x]^2, g \in \Sigma_{n-1}\mathbb{R}[x]^2\},\\ &\mathrm{Pos}([a,b])_{2n+1} = \{f(x-a) + (b-x)g,\ f,g \in \Sigma_n\mathbb{R}[x]^2\}.\end{aligned}$$

□

3 Application au problème des moments

Sous l'étiquette "problème des moments" se cachent en fait deux questions :

- (existence) Étant donnée une suite réelle $(s_n)_{n\geq 0}$, existe-t-il une variable aléatoire réelle X de loi μ telle que
$$s_n = \mathbb{E}[X^n] = \int x^n \mu(dx), \quad \forall n \geq 0 \quad ?$$
- (unicité) La loi μ est-elle caractérisée par la suite des moments $(s_n)_{n\geq 0}$?

Concernant l'unicité, si la loi μ est à support compact, la réponse est affirmative : les moments caractérisent la loi. Dans le cas général, si μ est à support non-borné, les moments ne caractérisent pas la loi, cf contre exemple classique de la loi log-normale.

Concernant l'existence, le travail effectué plus haut permet de donner des conditions nécessaires et suffisantes au problème des moments. Étant donnée la suite $s = (s_n)_{n\geq 0}$, on lui associe de façon naturelle un opérateur linéraire L_s sur $\mathbb{R}[x]$ défini sur la base $\{1, x, x^2, x^3, \ldots\}$
$$L_s(x^j) = s_j, \quad \forall j \in \mathbb{N}.$$
Par linéarité, si $p(x) = \sum_{j=0}^{d} c_j x^j$ alors $L_s(p) = \sum_{j=0}^{d} c_j s_j$.

Théorème 3.1 (Problème des moments de Hamburger)**.** *Soit $s = (s_n)_{n\geq 0}$ une suite réelle et L_s l'opérateur linéaire associé. Les assertions suivantes sont équivalentes :*

1. *Il existe une mesure positive sigma-finie μ sur $\mathbb{R}$ telle que $s_n = \int x^n \mu(dx)$, $\forall n \geq 0$*
2. *Pour tout $n \geq 0$, la matrice de Hankel $H_n(s)$ ci-dessous est semi-définie positive*
$$H_n(s) := \begin{pmatrix} s_0 & s_1 & s_2 & \ldots & s_n \\ s_1 & s_2 & s_3 & \ldots & s_{n+1} \\ s_2 & s_3 & s_4 & \ldots & s_{n+2} \\ \ldots & \ldots & \ldots & \ldots & \\ s_n & s_{n+1} & s_{n+2} & \ldots & s_{2n} \end{pmatrix}.$$
3. *La suite $s = (s_n)_{n\geq 0}$ est semi-définie positive, i.e. pour tout suite $(z_n)_{n\geq 0} \in \mathbb{C}^{\mathbb{N}}$ nulle à partir d'un certain rang $\sum_{i,j\geq 0} s_{i+j} z_i \bar{z}_j \geq 0$.*
4. *L'opérateur L_s et positif sur $Pos(\mathbb{R})$, i.e. on a $L_s(p^2) \geq 0$ pour tout $p \in \mathbb{R}[X]$.*

Remarque 1. *Les restrictions éventuelles sur le support de F de la mesure μ se transcrivent alors naturellement sur l'ensemble $Pos(F)$:*

- *(problème des moments de Stieljes) La suite $s = (s_n)_{n\geq 0}$ est la suite des moments d'une mesure supportée dans $\mathbb{R}^+$ si et seulement si L_s est positif sur $Pos(\mathbb{R}^+)$, i.e. $L_s(p^2) \geq 0$ et $L_s(xp^2) \geq 0$ pour tout $p \in \mathbb{R}[X]$; si et seulement si les suites $(s_n)_{n\geq 0}$ et $\theta(s_n)_{n\geq 0} := (s_{n+1})_{n\geq 0}$ sont semi-définie positives ; si et seulement si les matrices $H_n(s)$ et $H_n(\theta s)$ sont semi-définies positives.*
- *(problème des moments de Hausdorff) La suite $s = (s_n)_{n\geq 0}$ est la suite des moments d'une mesure supportée dans $[a, b]$, si et seulement si L_s est positif sur $Pos([a, b])$.*

Dans les assertions ci-dessus, remplacer semi-définie positive par définie positive revient à supposer que le support de μ est infini.

Références

[Sch17] Konrad Schmüdgen. The moment problem, volume 277 of Grad. Texts Math. Cham : Springer, 2017.