MÉTHODE DES FONCTIONS ORTHOGONALES (A4, C5)
(17 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

La **méthode des fonctions orthogonales** est une méthode d'**estimation** fondée sur des résultats classiques de l'analyse fonctionnelle (**espaces de HILBERT**). Elle est notamment utilisée pour l'**estimation de la densité** d'une **loi de probabilité**.

(i) Soit $(\Omega, \mathscr{T}, \mathscr{P})$ un **modèle statistique** non paramétré, $\xi : \Omega \mapsto \mathbf{R}$ une **vars** et $X = (X_1, ..., X_N)$ un **échantillon iid** selon la **loi** P^ξ de la **variable parente** ξ. : ie $P^\xi = \xi(P)$, $\forall P \in \mathscr{P}$. On suppose que chaque P^ξ admet une **densité** $f = dP^\xi / d\lambda_1$ pr à la **mesure de LEBESGUE** (cf **modèle dominé**).

Si $f \in L_{\mathbf{R}}^2(\mathbf{R}, \mathscr{B}_{\mathbf{R}}, \lambda_1)$ (**espace de HILBERT** des fonctions numériques de carré intégrable), on peut la décomposer dans une **base** orthonormale complète $(f_\alpha)_{\alpha \in \mathbf{N}}$ constituée d'éléments $f_\alpha \in L_{\mathbf{R}}^2(\mathbf{R}, \mathscr{B}_{\mathbf{R}}, \lambda_1)$ selon l'égalité suivante, entendue au sens de la convergence dans L^2 (cf **convergence dans L^p**) :

(1) $\quad f = \Sigma_{\alpha \in \mathbf{N}}\ b_\alpha f_\alpha$,

les **coefficients** $b_\alpha \in \mathbf{R}$ étant donnés par les relations :

(2) $\quad b_\alpha = \int_{\mathbf{R}} f_\alpha(x)\, dP^\xi(x) = \int f_\alpha(x)\, f(x)\, d\lambda_1(x), \qquad \forall \alpha \in \mathbf{N}$.

(ii) La **méthode des fonctions orthogonales** consiste à estimer f (ie la famille $b = (b_\alpha)_{\alpha \in \mathbf{N}}$) à l'aide d'une famille de vars tq :

(3) $$b_{N\alpha}^\# \text{ ou } b_{N\alpha}^\#(X) = \begin{cases} 0 & \text{si } \alpha > q(N), \\ N^{-1} \Sigma_{n=1}^N f_\alpha(X_n) & \text{si } \alpha \le q(N), \end{cases}$$

où $N \mapsto q(N) \in \mathbf{N}$ est une fonction à valeurs entières permettant une **approximation** de f qui soit la « meilleure » possible (cf **approximation d'une densité**).

L'**estimateur ponctuel** $f_N^\#$ de la densité est alors :

(4) $\quad f_N^\#(x) = \Sigma_{\alpha=0}^{q(N)}\ b_{N\alpha}^\#\ f_\alpha(x), \qquad \forall x \in \mathbf{R}$.

Lorsque $\alpha \le q(N)$ dans (3), $b_{N\alpha}^\#$ se calcule en remplaçant P^ξ par la **loi empirique** P_N dans (2).

(iii) L'**estimateur par les fonctions orthogonales** $f_N^\#$ ainsi défini possède les propriétés suivantes :

(a) ce n'est pas un **estimateur strict** de f. En effet, il peut prendre des valeurs négatives ou même ne pas sommer à 1. Il n'appartient donc pas (en général) à la **famille** des **densités de probabilité** ;

(b) c'est un estimateur biaisé de f (cf **biais**). En effet :

(5) $\quad E\, f_N^{\#} = \Sigma_{\alpha=0}^{q(N)}\ b_\alpha\, f_\alpha = f_{q(N)}$

(troncature du développement de f dans (1)). Cependant, sous certaines conditions, l'**écart quadratique moyen intégré** de $f_N^{\#}$ tend vers 0 lorsque $\lim_N q\,(N) = +\infty$ et que $\lim_N q\,(N) / N = 0$;

(c) sous diverses **conditions de régularité** (portant notamment sur $(f_\alpha)_{\alpha\,\in\,\mathbf{N}}$ et sur q), $f_N^{\#}$ converge presque sûrement vers f (pour la topologie de la **convergence uniforme** définie sur l'ensemble des densités uniformément continues sur **R**), ie (cf **convergence presque sûre**, **fonction uniformément continue**) :

(6) $\quad \sup_{x\,\in\,\mathbf{R}} |f_N^{\#}\,(x) - f\,(x)| \ \to_{N\,\to\,+\infty}\ 0, \qquad P\text{-p.s.}, \forall\, P \in \mathscr{P}$;

(d) sous des conditions du même type que les précédentes, la **loi asymptotique** de $f_N^{\#}$ est une **loi gaussienne**, ie :

(7) $\quad \mathscr{L}\ ((f_N^{\#} - f_{q(N)}) / \sigma\,(f_N^{\#})\,) \to_{N\,\to\,+\infty}\ \mathscr{N}\,(0,\,1)$ (**loi normale réduite**),

où $\sigma\,(f_N^{\#})$ désigne l'**écart-type** de $f_N^{\#}$.

(iv) On peut estimer $f^{\,(j)}$ (**dérivée** d'ordre $j \in \mathbf{N}^*$ de f) par la méthode des fonctions orthogonales :

(a) soit partir d'un système orthonormé complet $(f_\alpha)_{\alpha\,\in\,\mathbf{N}}$ tq f_α soit de classe C^j ($\forall\ \alpha \in \mathbf{N}$ et $\forall\ j \in \mathbf{N}^*$), puis estimer $f^{\,(j)}$ par la dérivée d'ordre j (pr à $x \in \mathbf{R}$) de $f_N^{\#}$, ie :

(8) $\quad f_N^{\#\,(j)} = (d^j / dx_j)\, f_N^{\#}$.

Dans ce cas, on montre que, sous certaines conditions :

(9) $\quad ||f_N^{\#} - f||_j \ \to_{N\,\to\,+\infty}\ 0, \ P\text{-p.s.}, \qquad \forall\ P \in \mathscr{P}$,

au sens de la norme $||g||_i = \sup_{i=1}^{j}\ \sup_{x\,\in\,\mathbf{R}} |g^{(j)}\,(x)|$, et $f_N^{\#\ (j)}$ admet aussi une loi asymptotique gaussienne ;

(b) soit estimer directement $f^{\,(j)}$ (supposée calculable analytiquement) à partir de sa décomposition sur une base orthonormée complète $(g_\alpha)_{\alpha\,\in\,\mathbf{N}}$ de $L_{\mathbf{R}}^2\,(\mathbf{R}, \mathscr{B}_{\mathbf{R}}\,, \lambda_1)$.

(v) il existe diverses variantes ou extensions de la méthode :

(a) changement de mesure dominante. Cette variante consiste à remplacer λ_1 par une **mesure positive** μ bornée sur $\mathscr{B}_{\mathbf{R}}$ (cf **mesure bornée**) et de masse totale unité (eg une **mesure de probabilité** ou une « **fonction de poids** »), puis à se donner une **suite** de fonctions μ-orthogonales $(\varphi_\alpha)_{\alpha \in \mathbf{N}}$ et normée (suite orthonormée), ie tq :

$$(10)\quad \begin{cases} \int_{\mathbf{R}} \varphi_\alpha \, \varphi_\beta \, d\mu = \delta_{\alpha\beta} & (\textbf{symbole de KRONECKER}), \\ \int_{\mathbf{R}} \varphi_0^2 \, d\mu = 1. \end{cases}$$

L'**estimateur par les fonctions orthogonales** est alors défini comme précédemment ;

(b) estimation multidimentionnelle. Celle-ci considère la densité d'un **vecteur aléatoire** $\xi : \Omega \mapsto \mathbf{R}^K$ (ie une densité multidimensionnelle) et requiert l'usage de fonctions orthonormales à K variables.

(vi) En pratique, on choisit a priori la suite $(f_\alpha)_{\alpha \in \mathbf{N}}$ en fonction du type de densité f considérée.

Ainsi, lorsque f est à **support** compact, cette suite est souvent celle des fonctions trigonométriques. A titre d'exemple, si Supp μ = [a, b] (avec a < b) et si $\varphi_\alpha (x) = \cos \{\alpha \, \pi \, (b - a)^{-1} (x - a)\}$ (avec $\varphi_0 (x) = (b - a)^{-1}$), alors :

$$(11)\quad f_N^{\#} (x) = (b - a)^{-1} + \Sigma_{\alpha=1}^{q(N)} \, b_{N\alpha}^{\#} \, \cos \{\alpha \, \pi \, (X^{(N)} - X^{(1)})^{-1} (x - X^{(1)})\},$$

avec :

$$(12)\quad b_{N\alpha}^{\#} = 2 \, . \, N^{-1} (X^{(N)} - X^{(1)})^{-1} \, . \, \Sigma_{n=1}^{N} \cos \{\alpha \, \pi \, (X^{(N)} - X^{(1)})^{-1} (x - X^{(1)})\},$$

où $X^{(.)}$ désigne l'échantillon ordonné associé à X (cf **statistique d'ordre**).

(vii) Diverses fonctions orthogonales sont utilisés dans ce type de problèmes, eg : fonctions de BERNSTEIN (cf **polynômes de BERNSTEIN**), fonctions de HERMITE (cf **polynômes de HERMITE**) ou fonctions LEGENDRE (cf **polynômes de LEGENDRE**), resp associées aux **polynômes orthogonaux** de même nom.