M1 finance: Statistiques non paramétriques
Fiche 9: Régression

Emeline Schmisser, `emeline.schmisser@math.univ-lille1.fr`, bureau 314 (bâtiment M3).

Dans les exercices qui suivent, on observe des couples de variables $((X_1,Y_1)\dots,(X_n,Y_n))$. Les variables X_i sont iid de densité f On suppose que les variables Y_i sont telles que

$$Y_i = g(X_i) + \varepsilon_i$$

où
- $(\varepsilon_1,\dots,\varepsilon_n)$ sont iid, centrés ($\mathbb{E}(\varepsilon_i)=0$), de variance finie σ^2, de densité h.
- $(X_1,\dots,X_n)$ sont iid de densité f.
- (X_i) et (ε_i) sont indépendants.

On veut estimer la fonction g sur l'intervalle A où $f \geq f_0$ par projection. Pour cela, on considère la suite de sous-espaces vectoriels $V_m = \text{Vect}\{\varphi_{j,m}, 1\leq j\leq D_m\}$ vérifiant :
- $\exists c, \forall m, D_m = c2^m$
- $\forall m, D_m \subset D_{m+1}$
- $\forall m$, les fonctions $(\varphi_{j,m})$ forment une BON de V_m et $\exists c, \forall m, \forall x, \sum \varphi_{j,m}^2(x) \leq cD_m$.
- $\exists c$, $\forall m$, $\forall \alpha \leq l$, $\forall t \in \mathcal{B}_{2,\infty}^{\alpha}$, $\|t-t_m\|_{L^2}^2 \leq cD_m^{-2\alpha}$.

Exercice 1 : Densité uniforme

Pour cet exercice, on suppose que $X_i \sim \mathcal{U}([0,1])$.

On considére l'estimateur par projection

$$\hat{g}_m(x) = \frac{\sum_{j=0}^{D_m-1} \hat{a}_{j,m}\varphi_{j,m}(x)}{f(x)} \quad \text{où} \quad \hat{a}_{j,m} = \frac{1}{n}\sum_{i=1}^{n} Y_i\varphi_{j,m}(X_i)$$

1. Majorer le risque L^2 de l'estimateur.
2. Comment se comporte $\hat{g}_m$ quand la dimension m est petite ? Grande ? Faire un dessin.
3. Trouver $D_{m_{opt}}$ qui minimise le risque si $g \in B_{2,\infty}^{\alpha}$ avec $\alpha \leq l$. Que vaut le risque dans ce cas ?
4. Quelle est la vitesse de convergence de $\hat{g}_m$?
5. Expliquer comment on peut construire un estimateur adaptatif de g.

Exercice 2 : Densité connue

On veut estimer la fonction g sur l'intervalle $A = [0,a]$ (a est un entier). On suppose maintenant que les X_i sont de densité f connue et que $\forall x \in A, 0 < f_0 \leq f(x) \leq f_1$.

On considère les sous-espaces V_m engendrés par la base de Haar.

1. Reprendre l'estimateur de la question précédente et calculer son espérance. Qu'est-ce qu'estime cet estimateur ?
2. Modifier l'estimateur précédent pour obtenir un estimateur de la fonction g.
3. Majorer le risque de cet estimateur.
4. Comment se comporte $\hat{g}_m$ quand la dimension m est petite ? Grande ? Faire un dessin.
5. Quelles sont les valeurs de j pour lesquelles φ_j et $\psi_{j,m}$ sont non nulles sur A ?

Exercice 3 : Un peu de simulation
On voudrait construire un programme permettant de reconstituer g. On sait que les variables X_i suivent une loi exponentielle $\mathcal{N}(0,\sigma^2)$ avec $\sigma = 2$.

1. On considère qu'il est raisonnable d'estimer g sur l'intervalle où on retrouve 86% des X_i. Quel est cet intervalle ?

   Pour des raisons de commodité, on va prendre l'intervalle entier le plus proche (bien qu'il soit possible en modifiant légèrement les bases d'estimer une fonction sur n'importe quel intervalle). On note I cet intervalle. Pour estimer g, on souhaite utiliser une méthode de projection sur la base de Haar. Haar.
2. Quelles sont les fonctions de cette base qui sont non identiquement nulles sur l'intervalle I ?
3. Expliquer comment calculer, pour j fixé, les estimateurs $\hat{a}_{k,j}$ et $\hat{b}_{k,j}$ avec un logiciel de calcul (type Scilab ou R). On n'utilisera pas de boucles, mais des matrices. On ne demande ni le code `R` : on écrira proprement les matrices, en précisant bien leur dimension, et on précisera les manipulations nécessaires pour calculer les coefficients.
4. Expliquer comment, une fois qu'on a les coefficients, on peut calculer le vecteur $(\hat{g}_0(x_1),\ldots,\hat{g}_0(x_n))$
5. Expliquer comment, une fois qu'on a les coefficients, on peut calculer le vecteur $((\hat{g}_m - \hat{g}_{m-1})(x_1),\ldots,(\hat{g}_m - \hat{g}_{m-1}(x_n))$ pour x un vecteur déterministe sur l'intervalle I avec un logiciel de calcul. On ne demande pas de code, mais il ne faut pas faire de boucles.
6. Expliquer comment on peut tracer la fonction $\hat{g}_m$ aux points donnés par le vecteur x.

Exercice 4.
On fixe $n = 1000$. Simuler 1000 variables Y_i où

$$Y_i = f\left(\frac{i}{n}\right) + \varepsilon_i$$

où les variables ε_i sont i.i.d de loi $\mathcal{N}(0,1)$ et où

$$f(x) = \frac{1}{2}\mathbf{1}_{[0,1/2[} + \frac{3}{2}\mathbf{1}_{[1/2,5/8[} + \frac{1}{8}\mathbf{1}_{[5/8,7/8[} + \mathbf{1}_{[7/8,1[}$$

Tracer les points $(i/n, Y_i)$ ainsi que la fonction $f(x)$ sur le même graphique. Peut-on reconstruire f en voyant seulement les points (Y_i) ? On oublie maintenant f et on veut l'estimer par une méthode par projection.

1. Construire les fonctions ϕ et ψ de la base de Haar.
2. Pour estimer f, on va estimer les coefficients $\hat{a}_k$ (correspondant aux fonctions φ_k et $\hat{b}_{k,j}$ (correspondant aux fonctions $psi_{k,j}$) pour j allant de 0 à $m-1$. Quelles sont les valeurs de k pour lesquelles on doit estimer les coefficients ?

On va maintenant estimer f. Comme on utilise une décomposition sur une base d'ondelettes, on pourra facilement obtenir $\hat{f}_m$ à partir de $\hat{f}_{m-1}$ (ce ne serait pas le cas avec la base de polynômes par morceaux).

3. Construire une fonction `coef` qui étant donné Y et m, renvoie :
   – le vecteur des coefficients $\hat{a}_{k_{\min}},\ldots,\hat{a}_{k_{\max}}$ si $m = 0$.
   – le vecteur des coefficients $\hat{b}_{k_{\min(m-1)}},\ldots,\hat{b}_{k_{\max(m-1)}}$ si $m > 0$
   construire les matrices

$$\begin{pmatrix} Y_1 & \ldots & Y_n \\ \ldots & \ldots & \ldots \\ Y_1 & \ldots & Y_n \end{pmatrix}, \quad \begin{pmatrix} 1/n & \ldots & 1 \\ \ldots & \ldots & \ldots \\ 1/n & \ldots & 1 \end{pmatrix}, \quad \begin{pmatrix} k_{\min} & \ldots & k_{\min} \\ \ldots & \ldots & \ldots \\ k_{\max} & \ldots & k_{\max} \end{pmatrix}$$

   puis calculer $\hat{a}_k$ ou $\hat{b}_{k,m-1}$.
   Tester le programme pour $m = 0, 1, 2$
4. Construire un vecteur x qui varie doucement de 0 à 1.

5. Construire une fonction *const* qui étant Y, x, m, reconstruise la fonction $\hat{f}_m - \hat{f}_{m-1}$.
   On utilisera la fonction précédente on considérera les matrices :

$$\begin{pmatrix} x_1 & \dots & x_1 \\ \dots & \dots & \dots \\ x_N & \dots & x_N \end{pmatrix}, \quad \begin{pmatrix} k_{\min} & \dots & k_{\max} \\ \dots & \dots & \dots \\ k_{\min} & \dots & k_{\max} \end{pmatrix}, \quad \begin{pmatrix} a_{k_{\min}} & \dots & a_{k\max} \\ \dots & \dots & \dots \\ a_{k_{\min}} & \dots & a_{k_{\max}} \end{pmatrix}$$

6. Tracer le graphe de $f(x)$. Superposer au graphe existant les fonctions $\hat{f}_0, \hat{f}_1, \hat{f}_2, \hat{f}_3, \hat{f}_4$ et $\hat{f}_5$. Que remarque-t-on ?

On va maintenant construire un estimateur adaptatif

7. Modifier la fonction `coef` : cette fonction attribue maintenant 0 au coefficient $\hat{a}_k$ ou $\hat{b}_{j,k}$ si sa valeur absolue est plus petite que $e\sqrt{\ln(n)/n}$. Le paramètre e est en entrée de la fonction. Modifier la fonction `const` en conséquence.
8. Tracer le graphe de $f(x)$. Superposer au graphe existant les fonctions $\hat{f}_0, \hat{f}_1, \hat{f}_2, \hat{f}_3, \hat{f}_4$ et $\hat{f}_5$ pour e=1. Que remarque-t-on ?