



Estimation robuste dans un modèle paramétrique avec rupture

Gabriela Ciuperca

► To cite this version:

Gabriela Ciuperca. Estimation robuste dans un modèle paramétrique avec rupture. 41èmes Journées de Statistique, SFdS, Bordeaux, 2009, Bordeaux, France, France. inria-00386645

HAL Id: inria-00386645

<https://hal.inria.fr/inria-00386645>

Submitted on 22 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ESTIMATION ROBUSTE DANS UN MODÈLE PARAMÉTRIQUE AVEC RUPTURE

Gabriela Ciuperca

*Université de Lyon, Université Lyon 1,
CNRS, UMR 5208, Institut Camille Jordan,
Bat. Braconnier, 43, blvd du 11 novembre 1918,
F - 69622 Villeurbanne Cedex, France,
email: Gabriela.Ciuperca@univ-lyon1.fr
tel: 33(0)4.72.43.16.90, fax: 33(0)4.72.43.16.87*

Résumé

Le papier considère un modèle paramétrique avec et sans points de rupture. Les paramètres du modèle sont estimés en utilisant la méthode LAD. L'intérêt de cette méthode par rapport aux méthodes "traditionnelles" est qu'elle permet d'obtenir des estimateurs robusts, surtout quand le modèle contient des outliers. On considère également un estimateur pénalisé pour pallier la non-différentiabilité de la fonction objectif. Pour ces estimateurs la vitesse de convergence, la loi asymptotique et un critère pour trouver le nombre de points de rupture sont donnés. Des simulations numériques confirment que ces estimateurs sont plus efficaces que l'estimateur LSQ.

Abstract

The paper considers a nonlinear parametric model with and without change-points. The parameters of model are estimated using the LAD method. The interest of this method is its robustness with respect to other traditional methods when the errors of model contain outliers. Since the objective function is not differentiable, a penalized estimator can be considered. For these estimators, the convergence rate, asymptotic distribution and a criterion to found the number of the change-points are given. Monte Carlo simulations show that these estimators are more efficient than the LSQ estimator.

1 Introduction

Considérons le modèle:

$$Y_i = g_\theta(x_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad (1)$$

la fonction g comportant K ($K \geq 0$) points de rupture (change-points):

$$g_\theta(x_i) = h_{\beta_1}(x_i)\mathbb{1}_{i \leq l_1} + h_{\beta_2}(x_i)\mathbb{1}_{l_1 < i \leq l_2} + \dots + h_{\beta_{K+1}}(x_i)\mathbb{1}_{l_K < i}.$$

Les paramètres de régression sont notés $\theta_1 = (\beta_0, \beta_1, \dots, \beta_K)$ et les points de rupture $\theta_2 = (l_1, \dots, l_K)$ avec $l_1 < l_2 < \dots < l_K$. On note $\theta = (\theta_1, \theta_2)$. Si $K = 0$, le modèle

est sans rupture. Les applications de ces modèles sont très nombreuses. Pour ne donner que quelques unes, on peut citer les processus de signal sismique, l'analyse des électrocardiogrammes, finance, épidémiologie.

Il est bien connu que la présence des *ouliers* peut provoquer que les estimateurs des paramètres obtenus par la méthode des moindres carrés aient une grande variance. Dans ce cas, il est préférable de considérer l'estimateur LAD pour les paramètres:

$$\hat{\theta}_n = \arg \min_{\theta} \sum_{i=1}^n |Y_i - g_{\theta}(x_i)|.$$

Par rapport à d'autres estimateurs, l'étude des propriétés du LAD-estimateur est plus difficile puisque la fonction critère est non-différentiable.

2 Résultats en absence des points de rupture

Dans cette section considérons le modèle de régression:

$$Y_i = h_{\beta}(x_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad (2)$$

sous l'hypothèse que ε_i sont des variables aléatoires i.i.d. et $\beta \in \Gamma \subset \mathbb{R}^p$, avec Γ un ensemble compact, l'expression analytique de la fonction $h_{\beta}(x)$ étant connue. Dans ce modèle, Y_i est une variable aléatoire pendant que x_i est déterministe. L'ensemble $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ est compact sous la métrique $d(x, y) = |\arctan x - \arctan y|$. Alors, on peut considérer que x_i prend des valeurs dans un compact Υ de \mathbb{R} . Les paramètres de régression β sont inconnus.

Pour la variable aléatoire ε , les suppositions suivantes sont faites:

- La fonction de répartition F de ε , f la densité, satisfont les conditions: $f(0) > 0$, $F(0) = 1/2$
- il existe un $c > 0$ tel que pour tout y dans un voisinage de 0: $|f(y) - f(0)| \leq c|y|^{1/2}$
- il existe $c^{(0)} > 0$ tel que $|F(y) - F(x)| \leq c^{(0)}|y - x|$, $\forall x, y \in \mathbb{R}$.

Le but est d'estimer les paramètres de régression quand n observations $(Y_i, x_i)_{1 \leq i \leq n}$ sont disponibles, en utilisant le principe des "*least absolute deviations*" (LAD). Par définition, l'estimateur LAD est:

$$\hat{\beta}_n = \arg \min_{\beta \in \Gamma} \sum_{i=1}^n |Y_i - h_{\beta}(x_i)|. \quad (3)$$

Parce que la fonction objectif n'est pas différentiable, l'inférence statistique basée sur cette méthode est plus compliquée.

Soit β^0 la vraie valeur du paramètre β . On suppose que β^0 est un point intérieur de l'ensemble Γ et que la fonction $h : \Upsilon \times \Gamma \rightarrow \mathbb{R}$ satisfait les conditions:

- (H1) pour tout $x \in \Upsilon$, la fonction $h_{\beta}(x)$ est deux fois différentiable en β
- (H2) pour tout $x \in \Upsilon$, $h_{\beta}(x)$, $\|\partial h_{\beta}(x)/\partial \beta\|$, $\|\partial^2 h(x, \beta)/\partial \beta^2\|$ sont bornées dans un

voisinage de β^0 .

En plus, on suppose que les (x_i) satisfont les conditions:

(H3) pour n suffisamment grand, il existe un $c^{(1)} > 0$ tel que:

$$\frac{1}{n} \sum_{i=1}^n \sup_{\beta \in \Gamma} \left\| \frac{\partial h_{\beta}(x_i)}{\partial \beta} \right\| \leq c^{(1)} < \infty. \quad (4)$$

(H4) La limite suivante existe:

$$M := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left\| \frac{\partial h_{\beta^0}(x_i)}{\partial \beta} \right\|^2.$$

Considérons la fonction: $\nu_{(\beta_1, \beta_2)}(x) = h_{\beta_1}(x) - h_{\beta_2}(x)$. Pour l'étude de la fonction objectif: $\sum_{i=1}^n |Y_i - h_{\beta}(x_i)| = \sum_{i=1}^n |\varepsilon_i + h_{\beta^0}(x_i) - h_{\beta}(x_i)|$, les processus suivants sont également considérés: $\eta_i(\beta) = |\varepsilon_i - \nu_{(\beta, \beta^0)}(x_i)| - |\varepsilon_i|$, $\xi_i(\beta) = \eta_i(\beta) - \mathbb{E}[\eta_i(\beta)]$.

L'estimateur $\hat{\beta}_n$ est consistant si la condition suivante est satisfaite Oberhofer [3]: pour $\Gamma_0 \subset \Gamma$ un fermé ne contenant pas β^0 , il existe $\epsilon > 0$ et $n_0 \in \mathbb{N}$ tels que, pour tout $n \geq n_0$:

$$\inf_{\beta \in \Gamma_0} \frac{1}{n} \sum_{i=1}^n |\nu_{(\beta, \beta^0)}(x_i)| \min \left\{ F \left(\frac{|\nu_{(\beta, \beta^0)}(x_i)|}{2} \right) - \frac{1}{2}; \frac{1}{2} - F \left(-\frac{|\nu_{(\beta, \beta^0)}(x_i)|}{2} \right) \right\} \geq \epsilon. \quad (5)$$

Il est bien connu dans la littérature Weiss [4] que l'estimateur LAD de β est asymptotiquement normal, avec la matrice de covariance dépendant de $f(0)$. D'autre part, il n'y a pas de résultat connu pour la vitesse de convergence. On prouve alors:

Théorème 2.1 *Sous les hypothèses (H1)-(H4) et la condition (5), pour toute suite positive monotone (v_n) telle que :*

$$v_n \rightarrow 0, \quad n v_n^2 \rightarrow \infty \quad \text{pour } n \rightarrow \infty, \quad (6)$$

on a: $\|\hat{\beta}_n - \beta^0\| = O_{\mathbb{P}}(v_n)$.

Le résultat du Théorème 2.1 est plus fort que le résultat obtenu par Babu [1] pour une régression linéaire: $v_n = (\log n/n)^{1/2}$.

3 Plusieurs points de rupture

Pour le modèle (1) on considère que la fonction $g_{\theta}(x)$ a K ($K \geq 1$) points de rupture. Aucune hypothèse de continuité pour g n'est pas imposée dans les points de rupture. Pour tout $r = 1, \dots, K$, $K+1$ supposons que $\beta_k \in \Gamma \subset \mathbb{R}^p$ avec Γ compact et $\theta_2 \in \mathbb{R}^K$, où $\theta \in \Omega = \Gamma^{K+1} \times \Upsilon^K$. Soit $\theta^0 = (\theta_1^0, \theta_2^0)$ la vraie valeur (inconnue) de θ . Les suppositions

suivantes sont faites pour le modèle:

(A1) On suppose que les points de rupture sont assez "éloignés": pour un $u \geq 3/4$ il existe $c > 0$ tel que $l_{r+1}^0 - l_r^0 \geq cn^u$, $\forall r = 1, \dots, K$ avec $l_0^0 = 1$ et $l_{K+1}^0 = n$.

(A2) A gauche et à droite d'un point de rupture l_r^0 les fonctions de régression sont différentes: $\beta_r^0 \neq \beta_{r+1}^0$, $\forall r = 1, \dots, K$

La construction des estimateurs se fait dans deux étapes: d'abord les paramètres de régression sont trouvés et ensuite les points de rupture sont localisés. Pour les K points de rupture l_1, \dots, l_K , notons $\hat{\theta}_1(\theta_2) = \hat{\theta}_1(l_1, \dots, l_K) = \arg \min_{\theta_1} \sum_{r=1}^{K+1} \sum_{i=l_{r-1}+1}^{l_r} |Y_i - g_{\theta}(x_i)|$ l'estimateur LAD de θ_1^0 pour un paramètre θ_2 . Pour l_1, \dots, l_K fixés, notons la somme des moindres déviations par:

$$S(l_1, \dots, l_K) := \sum_{r=1}^{K+1} \inf_{\theta_1} \sum_{i=l_{r-1}+1}^{l_r} |Y_i - g_{\theta}(x_i)|. \quad (7)$$

On définit l'estimateur LAD des points de rupture:

$$\hat{\theta}_{2n} = (\hat{l}_1, \dots, \hat{l}_K) := \arg \min_{l_1 < \dots < l_K} S(l_1, \dots, l_K). \quad (8)$$

Enfin, les estimateurs des paramètres de régression sont obtenus en utilisant les estimateurs des points de rupture: $\hat{\theta}_{1n} := (\hat{\beta}_{1,n}, \dots, \hat{\beta}_{K+1,n}) = \hat{\theta}_1(\hat{\theta}_{2n})$.

La vitesse de convergence de l'estimateur des points de rupture est donnée par le résultat:

Théorème 3.1 *Pour tout $\rho > 1/2$, avec les suppositions (A1), (A2), (H1), (H2), (H3), (H4) et pour toute suite (v_n) comme dans le Théorème 2.1, on a pour $n \rightarrow \infty$:*

$$\mathbb{P} \left[|\hat{l}_r - l_r^0| > \inf \{n^\rho, n^2 v_n^4\} \right] \rightarrow 0, \quad r = 1, \dots, K.$$

Si on considère la condition supplémentaire:

(H5) pour n assez grand, il existe $c^{(2)} > 0$ tel que:

$$\frac{1}{n} \sum_{i=1}^n \sup_{\beta \in \Gamma} \left\| \frac{\partial h_{\beta}(X_i)}{\partial \beta} \right\|^2 \leq c^{(2)} < \infty. \quad (9)$$

le résultat du Théorème 3.1 peut être amélioré.

Théorème 3.2 *Sous les conditions (A1), (A2), (H1), (H2), (H3), (H4) et (H5), pour tout $r = 1, \dots, K$ on a: $\hat{l}_r - l_r^0 = O_{\mathbb{P}}(1)$.*

Les lois limite des estimateurs sont données par les deux théorèmes suivantes:

Théorème 3.3 *Sous les mêmes conditions que pour le Théorème 3.2, pour chaque $r = 1, 2, \dots, K$, la différence $\hat{l}_r - l_r^0$ converge en distribution pour $n \rightarrow \infty$ vers L_r , le point de minimum du processus aléatoire $\{\dots, Z_{-1}^{(r)}, Z_0^{(r)}, Z_1^{(r)}, \dots\}$, avec $Z_0^{(r)} = 0$.*

Pour $j = 1, 2, \dots$

$$Z_j^{(r)} = \sum_{i=l_r^0+1}^{l_r^0+j} \left\{ \left| \varepsilon_i^{(r)} - \nu_{(\beta_{r-1}^0, \beta_r^0)}(x_i) \right| - \left| \varepsilon_i^{(r)} \right| \right\},$$

et pour $j = -1, -2, \dots$:

$$Z_j^{(r)} = \sum_{i=l_r^0+j}^{l_r^0} \left\{ \left| \varepsilon_i^{(r)} + \nu_{(\beta_{r-1}^0, \beta_r^0)}(x_i) \right| - \left| \varepsilon_i^{(r)} \right| \right\},$$

où $\varepsilon_i^{(r)}$ est une copie indépendante de ε_i .

Théorème 3.4 *Sous les mêmes conditions que pour le Théorème 3.2 et sous la condition que la densité f des erreurs ε est Lipschitz, on a pour chaque $r = 1, \dots, K + 1$*

$$2f(0) \left(\hat{l}_r - \hat{l}_{r-1} \right) \left(\hat{\beta}_{r,n} - \beta_r^0 \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, V_r), \quad (10)$$

où $\hat{l}_{K+1} = l_{K+1}^0 = n$ et

$$V_r = \frac{1}{l_r^0 - l_{r-1}^0} \sum_{i=l_{r-1}^0+1}^{l_r^0} \left(\frac{\partial h_{\beta_r^0}(x_i)}{\partial \beta} \right) \left(\frac{\partial h_{\beta_r^0}(x_i)}{\partial \beta} \right)^t.$$

Remarque 3.1 *Par la méthode des moindres carrés, si $h_\beta(x)$ est constante: $h_\beta(x) = \beta$, Yao et Au [6] ont montré que l'estimateur du point de rupture converge en distribution vers le point de minimum d'une marche aléatoire. Pour une régression linéaire: $h_\beta(x) = \beta x$, Bai et Perron [2] ont montré que la loi asymptotique de l'estimateur des points de rupture est le maximum d'un processus Wiener process avec drift.*

Pour estimer le nombre K de points de rupture, on adapte le critère de Schwarz proposé par Yao [5]. Soit K_0 la vraie valeur de K . Pour un nombre K fixé, considérons $(\hat{l}_{1,K}, \dots, \hat{l}_{K,K}) = \arg \min S(l_1, \dots, l_K)$ and $\hat{s}_K = S(\hat{l}_{1,K}, \dots, \hat{l}_{K,K})/n$, où la fonction S est définie par (7).

Théorème 3.5 *Sous les conditions $\mathbb{E}[|\varepsilon|] < \infty$, (A1), (A2), (H1)-(H4), soit \hat{K}_n la valeur de K qui minimise $B(K) := n \log \hat{s}_K + K C_n$, avec (C_n) une suite satisfaisant $C_n \rightarrow \infty$, $C_n n^{-3/4} \rightarrow 0$ et $C_n n^{-1/2} \rightarrow \infty$ pour $n \rightarrow \infty$. Alors:*

$$\mathbb{P} \left[\hat{K}_n = K_0 \right] \rightarrow 1, \quad \text{pour } n \rightarrow \infty.$$

4 Estimateur LAD pénalisé

Pour que la fonction objectif soit plus "régulière", on considère une suite $d_n > 0$ telle que $d_n \rightarrow 0$ pour $n \rightarrow \infty$. D'abord, si le modèle n'a pas de points de rupture, modèle (2), l'estimateur LAD pénalisé (noté SLAD) est:

$$\hat{\beta}_n^s = \arg \min_{\beta \in \Gamma} Q_n^s(\beta)$$

avec $Q_n^s(\beta) := n^{-1} \sum_{i=1}^n \left\{ [\varepsilon_i - \nu_{(\beta, \beta^0)}(x_i)]^2 + d_n^2 \right\}^{1/2}$. On peut montrer que dans ce cas, l'estimateur SLAD a la même vitesse de convergence que l'estimateur non pénalisé. Si le modèle contient des points de rupture, modèle (1), considérons la fonction objectif:

$$S^s(l_1, \dots, l_K) := \sum_{r=1}^{K+1} \inf_{\theta_1} \sum_{i=l_{r-1}+1}^{l_r} \left\{ |Y_i - g_{\theta}(x_i)|^2 + d_n^2 \right\}^{1/2} \quad (11)$$

L'estimateur SLAD des points de rupture est défini par:

$$\hat{\theta}_{2n}^s = (\hat{l}_1^s, \dots, \hat{l}_K^s) := \arg \min_{l_1 < \dots < l_K} S^s(l_1, \dots, l_K) \quad (12)$$

Les estimateurs SLAD des paramètres de régression sont: $\hat{\theta}_{1n}^s := (\hat{\beta}_{1n}^s, \dots, \hat{\beta}_{K+1,n}^s) = \hat{\theta}_1^s(\hat{\theta}_{2n}^s)$. Alors, l'estimateur SLAD pour les paramètres du modèle (1) est $\hat{\theta}_n^s := (\hat{\theta}_{1n}^s, \hat{\theta}_{2n}^s)$.

Théorème 4.1 *Sous les conditions (C1)-(C3), (H1), (H2), (H3), (H4), (A1), (A2), considérons les suites positives (a_n) , (w_n) telles que $a_n \rightarrow \infty$, $w_n \ll a_n \ll n^{3/4}$ et $w_n := \max(n^\alpha, nd_n)$, $\alpha \in (1/2, 3/4)$. Alors:*

$$\mathbb{P} \left[|\hat{l}_r^s - l_r^0| > a_n \right] \rightarrow 0, \quad r = 1, \dots, K$$

La vitesse de convergence pour les estimateurs SLAD des paramètres de régression est:

Théorème 4.2 *Sous les mêmes conditions que pour le Théorème 4.1 et sous (H5), nous avons: $\|\hat{\theta}_{1n}^s - \theta_1^0\| = O_{\mathbb{P}}(n^{\alpha-\rho+\delta})$, pour $\forall \delta \in (0, \rho - \alpha)$, $\rho > 3/4$ et $\alpha \in (1/2, 3/4)$.*

Le critère pour trouver le nombre de points de rupture peut être adapté dans ce cas, en considérant pour tout nombre K fixé, l'estimateur $(\hat{l}_{1,K}^s, \dots, \hat{l}_{K,K}^s) = \arg \min_{l_1 < \dots < l_K} S^s(l_1, \dots, l_K)$ et $\hat{s}_K^s = S^s(\hat{l}_{1,K}^s, \dots, \hat{l}_{K,K}^s)/n$, où la fonction S^s est définie par (11).

Théorème 4.3 *Sous les mêmes conditions que pour le Théorème 4.1 et $\mathbb{E}[|\varepsilon|] < \infty$, soit \hat{K}_n la valeur de K qui minimise $B(K) := n \log \hat{s}_K^s + KC_n$, avec (C_n) toute suite satisfaisant $C_n \rightarrow \infty$ et $w_n \ll C_n \ll a_n$, pour $n \rightarrow \infty$. Alors:*

$$\mathbb{P} \left[\hat{K}_n = K_0 \right] \rightarrow 1, \quad \text{pour } n \rightarrow \infty$$

Table 1: Modèle nonlinéar, un point de rupture, erreur Laplace.

parameter	$a_1(a_1^0 = 10)$	$b_1(b_1^0 = 2)$	$a_2(a_2^0 = 7)$	$b_2(b_2^0 = 1.75)$	$l_1(l_1^0 = 60)$
mean by LAD	9.6377	2.0214	6.9069	1.7550	60
mean by SLAD	9.6369	2.0215	6.9054	1.7551	60
mean by LSQ	-15.7310	3.5832	6.5418	1.9697	63
std by LAD	0.2067	0.0192	0.0329	0.0021	0.23
std by SLAD	0.2003	0.0186	0.0317	0.0020	0.23
std by SLQ	373.5897	6.2752	0.9266	1.9763	31.22
median by LAD	9.6728	2.0199	6.9076	1.7550	60
median by SLAD	9.6611	2.0201	6.9066	1.7550	60
median by LSQ	11.1088	1.8515	6.5268	1.7768	60

5 Comparaison d'estimateurs par simulation

On donne les résultats obtenus pour l'estimation des paramètres dans un modèle avec rupture par trois méthodes: LSQ, LAD et SLAD. Pour la fonction $h_\beta(x) = a(1 - x^b)/b$, $\beta = (a, b) \in [-100, 100] \times [0.1, 20]$, un seul point de rupture, soit le modèle:

$$Y_i = \left(a_1 \frac{1 - x_i^{b_1}}{b_1}\right) \mathbb{1}_{i \leq l_1} + \left(a_2 \frac{1 - x_i^{b_2}}{b_2}\right) \mathbb{1}_{l_1 < i} + \varepsilon_i, \quad i = 1, \dots, n \quad (13)$$

avec $x_i = i/10$, $n = 200$ et les vraies valeurs $a_1^0 = 10$, $b_1^0 = 2$, $a_2^0 = 7$, $b_2^0 = 1.75$, $l_1^0 = 60$. (voir la Figure 1). Les moyennes, les écarts-type, les médianes empiriques des estimations obtenues par les trois méthodes pour 200 répliques Monte Carlo de (13) sont donnés dans le Tableau 1 (on a considéré $d_n^2 = n^{-1}$). La méthode des moindres carrés donne des mauvais résultats s'il y a des outliers dans le modèle.

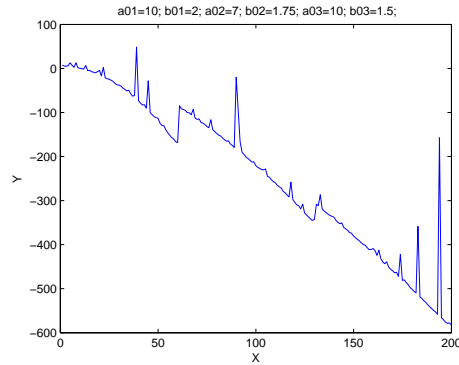
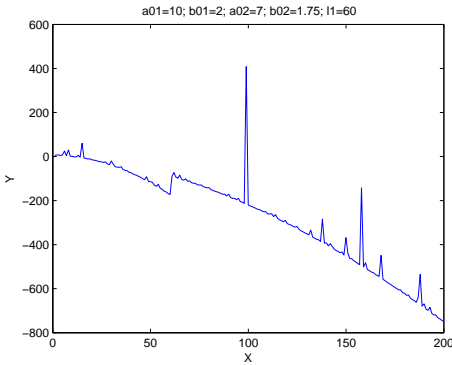


Figure 1: Nonlinear model with Laplace errors and with one change-point.

Figure 2: Nonlinear model with Laplace errors and with two change-points.

Table 2: Modèle nonlinear, deux points de rupture, erreur Laplace.

parameter	a_1	b_1	a_2	b_2	a_3	b_3	l_1	l_2
mean by LAD	9.5	2	6.8	1.76	9.88	1.5	60	130
mean by SLAD	9.5	2	6.8	1.76	9.88	1.5	60	130
mean by LSQ	11.02	1.9	4.98	10.19	10.5	1.48	66	124
std by LAD	0.2	0.01	0.12	0.01	0.08	0.004	0	0.2
std by SLAD	0.2	0.01	0.12	0.01	0.09	0.004	0	0.2
std by SLQ	5.4	0.43	2.63	26.14	1.95	0.078	16	15
median by LAD	9.5	2.034	6.83	1.76	9.88	1.5	60	130
median by SLAD	9.5	2.036	6.83	1.76	9.9	1.5	60	130
median by LSQ	8.74	2.1	6.36	1.79	9.67	1.5	60	130

Soit le modèle:

$$Y_i = \left(a_1 \frac{1 - x_i^{b_1}}{b_1}\right) \mathbb{1}_{i \leq l_1} + \left(a_2 \frac{1 - x_i^{b_2}}{b_2}\right) \mathbb{1}_{l_1 < i} + \left(a_3 \frac{1 - x_i^{b_3}}{b_3}\right) \mathbb{1}_{l_1 < i} + \varepsilon_i, \quad i = 1, \dots, n \quad (14)$$

avec $x_i = i/10$, $n = 200$ et les vraies valeurs $a_1^0 = 10$, $b_1^0 = 2$, $a_2^0 = 7$, $b_2^0 = 1.75$, $a_3^0 = 10$, $b_3^0 = 1.5$, $l_1^0 = 60$, $l_2^0 = 130$. (voir Figure 2). Les moyennes, les écarts-type, les médianes empiriques des estimations obtenues par les trois méthodes pour 30 répliques Monte Carlo de (14) sont donnés dans le Tableau 2. Les estimateurs SLAD et LAD sont non-biaisés et plus efficaces que le LSQ estimateur.

References

- [1] Babu, G.J., (1989), Strong representations for LAD estimators in linear models. *Probability Theory and Related Fields*, **83**, pp. 547-558.
- [2] Bai, J., Perron, P., (1998), Estimating and testing linear models with multiple structural changes. *Econometrica*, **66**, pp. 47-78.
- [3] Oberhofer, W., (1982), The consistency of nonlinear regression minimizing the L_1 -norm. *The Annals of Statistics*, **10**, No. 1, pp. 316-319.
- [4] Weiss, A.A., (1991), Estimating nonlinear dynamic models using least absolute error estimation. *Econometric Theory*, **7**, pp. 46-68.
- [5] Yao, Y.C., (1988), Estimating the number of change-points via Schwarz's criterion. *Statistics and Probability Letters*, **6**, pp. 181-189.
- [6] Yao, Y.C., Au, S.T., (1988), Least-squares estimation of a step function. *Sankhya*, **51**, pp. 370-381.