Frontier estimation via kernel regression on high power-transformed data

Stéphane Girard, Pierre Jacob

Frontier estimation via kernel regression on high power-transformed data

Séphane Girard[1,⋆] & Pierre Jacob[2]

[1,⋆]Université Grenoble 1, LMC-IMAG, BP 53, 38041 Grenoble cedex 9, France,
Stephane.Girard@imag.fr (corresponding author)
[2]Université Montpellier 2, EPS-I3M, place Eugène Bataillon,
34095 Montpellier cedex 5, France, jacob@math.univ-montp2.fr

Abstract

We present a new method for estimating the frontier of a multidimensional sample. The estimator is based on a kernel regression on the power-transformed data. We assume that the exponent of the transformation goes to infinity while the bandwidth of the kernel goes to zero. We give conditions on these two parameters to obtain almost complete convergence and asymptotic normality. The good performance of the estimator is illustrated on some finite sample situations.

Keywords: kernel estimator, power-transform, frontier estimation.
AMS 2000 subject classification: 62G05, 62G07, 62G20.

1 Introduction

Let (X_i, Y_i), $i = 1, \ldots, n$ be independent copies of a random pair (X, Y) with support S defined by

$$S = \{(x, y) \in E \times \mathbb{R}; 0 \leq y \leq g(x)\}. \tag{1}$$

The unknown function $g : E \to \mathbb{R}$ is called the frontier. We address the problem of estimating g in the case $E = \mathbb{R}^d$. Our estimator of the frontier is based on a kernel regression on the power-transformed data. More precisely, the estimator of g is defined for all $x \in \mathbb{R}^d$ by

$$\hat{g}_n(x) = \left((p+1) \sum_{i=1}^{n} K_h(x - X_i) Y_i^p \Big/ \sum_{i=1}^{n} K_h(x - X_i) \right)^{1/p}, \tag{2}$$

where $p = p_n$ and $h = h_n$ are non random sequences such that $h \to 0$ and $p \to \infty$ as $n \to \infty$. This latter condition is the key so that the high power-transformed data "concentrate" along the frontier.

We have also introduced $K_h(t) = K(t/h)/h^d$ where K is a probability distribution function (pdf) on $\mathbb{R}^d$. In this context, h is called the window-width.

From the practical point of view, note that, compared to the extreme value based estimators [6, 7, 10, 11, 13, 12], projection estimators [18] or piecewise polynomial estimators [22, 21, 17], this estimator does not require a partition of S and is thus not limited to Bi-dimensional bounded supports. Moreover, it benefits from an explicit formulation which is not the case of estimators defined by optimization problems [9] such as local polynomial estimators [16, 15, 20]. From the theoretical point of view, this estimator reveals to be almost completely convergent to g without assumption neither on the of distribution X nor on the distribution of Y given $X = x$ (see Section 3). Note however that $(p+1)^{1/p} \to 1$ when $p \to \infty$. In fact, this correcting term is specially designed for the case where Y given $X = x$ is uniformly distributed on $[0, g(x)]$. In this latter situation, the estimator is asymptotically Gaussian with the rate of convergence $n^{-\alpha/(d+\alpha)}$ (see Section 4). This rate is proved to be minimax optimal for $\alpha-$ Lipschitzian $d-$ dimensional frontiers [22], Chapter 5. This result is generalized in [23] to boundaries of more general regions. Another extensions are provided in [14, 17] to densities of Y given $X = x$ decreasing as a power of the distance from the boundary. We refer to [3, 4, 8] for the estimation of frontier functions under monotonicity assumptions. We conclude this paper by an illustration of the behavior of our estimator on some finite sample situations in Section 5. Technical lemmas are postponed to the appendix.

2 Notations and assumptions

To motivate the estimator (2), consider the random variable $Z = (p+1)Y^p$ and the conditional expectation $r_n(x) = \mathbb{E}(Z|X = x)$. Estimating the frontier g is often related to estimating the regression function r_n. For instance, if Y given $X = x$ is uniformly distributed on $[0, g(x)]$, we have $r_n(x) = g^p(x)$. A similar remark is done in [19] where regression estimators are modified to build estimators of the frontier, but the profound difference here is that $p \to \infty$. We denote by f the pdf of the random vector X and we introduce

$$\hat{\varphi}_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - X_i)Z_i, \tag{3}$$

where $Z_i = (p+1)Y_i^p$. Note that $\hat{\varphi}_n(x)$ can be seen as a classical kernel estimator of $\varphi_n(x) = f(x)g^p(x)$ but keep in mind that $p \to \infty$. Similarly,

$$\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - X_i) \tag{4}$$

is an estimator of $f(x)$ and

$$\hat{r}_n(x) = \hat{\varphi}_n(x)/\hat{f}_n(x) \tag{5}$$

is an estimator of $r_n(x)$. Collecting (3), (4) and (5), our estimator (2) can be rewritten as

$$\hat{g}_n(x) = \hat{r}_n(x)^{1/p}.$$

To establish the asymptotic properties of $\hat{g}_n(x)$, the following assumptions are considered:

(A.0): Y given $X = x$ is uniformly distributed on $[0, g(x)]$.

(A.1): g is α-Lipschitz, f is β-Lipschitz, with $0 < \alpha \leq \beta \leq 1$,

(A.2): $0 < g_{\min} \leq g(x)$, $\forall x \in \mathbb{R}^d$,

(A.3): $f(x) \leq f_{\max} < \infty$, $\forall x \in \mathbb{R}^d$,

(A.4): K is a Lipschitzian pdf on $\mathbb{R}^d$, with support included in B, the unit ball of $\mathbb{R}^d$.

3 Almost complete convergence

First, the almost complete convergence of the frontier estimator toward the true frontier is established. Note that no assumption on the conditional distribution of Y given X is required. This is not so surprising since the L_p- norm of a bounded function converges to the L_∞- norm as $p \to \infty$.

Theorem 1 *Suppose* **(A.1)**–**(A.4)** *hold and* $nh^d/\log n \to \infty$. *Then* $\widehat{g}_n(x)$ *converges almost completely surely to* $g(x)$ *for all* $x \in \mathbb{R}^d$ *such that* $f(x) > 0$.

Proof : Let $x \in \mathbb{R}^d$ such that $f(x) > 0$ and let ε such that $0 < \varepsilon < g(x)$. Define $0 < \eta < 1/4$ by $\eta = \varepsilon/(4g(x))$. Then, from Lemma 5,

$$\begin{aligned}
\{|\widehat{g}_n(x) - g(x)| > \varepsilon\} &= \left\{ \left| \left(\frac{\widehat{r}_n(x)}{r_n(x)} \right)^{1/p} - 1 \right| > 4\eta \right\} \\
&\subseteq \left\{ \left| \left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)} \right)^{1/p} - 1 \right| > \eta \right\} \cup \left\{ \left| \left(\frac{\hat{f}_n(x)}{f(x)} \right)^{1/p} - 1 \right| > \eta \right\}.
\end{aligned}$$

Since $\hat{f}_n(x)$ converges almost completely surely to $f(x)$, see i.e. [1], Chapter 4, Theorem III.3, it follows that $(\hat{f}_n(x)/f(x))^{1/p}$ converges almost completely surely to 1. Therefore, writing

$$\left(\frac{\widehat{\varphi}_n(x)}{\varphi(x)} \right)^{1/p} = (p+1)^{1/p} T_n(x)$$

with

$$T_n(x) = \left[\frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i) \left(\frac{Y_i}{g(x)} \right)^p \frac{1}{f(x)} \right]^{1/p}$$

and remarking that $(p+1)^{1/p} \to 1$ as $n \to \infty$, it suffices to consider

$$\{|T_n(x) - 1| > \eta\} \subseteq \{T_n(x) > 1 + \eta\} \cup \{T_n(x) < 1 - \eta\}.$$

The two events are studied separately. First, let $0 < \delta < \eta$. Then, $\|x - X_i\| \leq h$ entails

$$Y_i - g(x)(1+\delta) \leq g(X_i) - g(x) - \delta g(x) \leq L_g h^\alpha - \delta g_{\min} < 0$$

for n large enough and where L_g is the Lipschitz constant associated to g. We thus have

$$\begin{aligned} T_n(x) &= \left[\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\left(\frac{Y_i}{g(x)}\right)^p \mathbf{1}\{Y_i < g(x)(1+\delta)\}\frac{1}{f(x)}\right]^{1/p} \\ &\leq (1+\delta)\left[\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i < g(x)(1+\delta)\}\frac{1}{f(x)}\right]^{1/p}, \end{aligned}$$

and consequently,

$$\begin{aligned} \{T_n(x) > 1+\eta\} &\subseteq \left\{\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i < g(x)(1+\delta)\}\frac{1}{f(x)} > \left(\frac{1+\eta}{1+\delta}\right)^p\right\} \\ &\subseteq \left\{\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i < g(x)(1+\delta)\}\frac{1}{f(x)} > 2\right\}, \end{aligned}$$

since, for n large enough, $((1+\eta)/(1+\delta))^p > 2$. From [1], Chapter 5, Corollary II.4, the following almost completely sure convergence holds:

$$\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i < g(x)(1+\delta)\}\frac{1}{f(x)} \stackrel{a.s.c.}{\longrightarrow} \mathbb{P}(Y < g(x)(1+\delta)|X=x) = 1,$$

and therefore

$$\sum_{n=1}^{\infty}\mathbb{P}(T_n(x) > 1+\eta) \leq \sum_{n=1}^{\infty}\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i < g(x)(1+\delta)\}\frac{1}{f(x)} - 1 > 1\right) < +\infty,$$

which concludes the first part of the proof. Second,

$$\begin{aligned} T_n(x) &\geq \left[\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\left(\frac{Y_i}{g(x)}\right)^p \mathbf{1}\{Y_i > g(x)(1-\delta)\}\frac{1}{f(x)}\right]^{1/p} \\ &\geq (1-\delta)\left[\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i > g(x)(1-\delta)\}\frac{1}{f(x)}\right]^{1/p}, \end{aligned}$$

and consequently,

$$\{T_n(x) < 1-\eta\} \subseteq \left\{\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i > g(x)(1-\delta)\}\frac{1}{f(x)} < \left(\frac{1-\eta}{1-\delta}\right)^p\right\}.$$

Now, since $\mathbb{P}(Y > g(x)(1-\delta)|X=x) > 0$, there exists $\gamma > 0$ such that, for n large enough,

$$\left(\frac{1-\eta}{1-\delta}\right)^p - \mathbb{P}(Y > g(x)(1-\delta)|X=x) < -\gamma,$$

entailing that, for n large enough,

$$\{T_n(x) < 1-\eta\} \subseteq \left\{\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i > g(x)(1-\delta)\}\frac{1}{f(x)} - \mathbb{P}(Y/g(x) > 1-\delta|X=x) < -\gamma\right\}.$$

Taking into account of the following almost completely sure convergence

$$\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i > g(x)(1-\delta)\}\frac{1}{f(x)} \stackrel{a.s.c.}{\longrightarrow} \mathbb{P}(Y > g(x)(1-\delta)|X=x),$$

it follows that

$$\begin{aligned}
& \sum_{n=1}^{\infty} \mathbb{P}(T_n(x) < 1-\eta) \\
\leq & \sum_{n=1}^{\infty} \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)\mathbf{1}\{Y_i > g(x)(1-\delta)\}\frac{1}{f(x)} - \mathbb{P}(Y > g(x)(1-\delta)|X=x) < -\gamma \right) \\
< & +\infty,
\end{aligned}$$

which concludes the second part of the proof. ■

4 Asymptotic normality

Second, the asymptotic normality of the frontier estimator centered on the true frontier is established. To this end, asymptotic expansions of the expectation and variance of $\hat{\varphi}_n(x)$ are needed. These calculations are done under the assumption that Y given $X = x$ is uniformly distributed on $[0, g(x)]$. The next two lemmas are similar to classical ones in kernel regression (see for instance [5], Theorem 6.11), but the dependence on n of the function $\varphi_n(x)$ to estimate induces technical difficulties.

Lemma 1 *Under* **(A.0)**–**(A.4)**, *if* $ph^\alpha \to 0$, *then for all* $x \in \mathbb{R}^d$

$$\mathbb{E}\widehat{\varphi}_n(x) = \varphi_n(x)\left[1 + O(ph^\alpha)\right].$$

Proof : From (3), it follows that

$$\mathbb{E}\widehat{\varphi}_n(x) = \mathbb{E}(K_h(x-X)Z) = \mathbb{E}(K_h(x-X)\mathbb{E}(Z|X)),$$

so that, by a straightforward calculation, and recalling that $\varphi_n(u) = g^p(u)f(u)$, we obtain

$$\begin{aligned}
\mathbb{E}\widehat{\varphi}_n(x) &= \mathbb{E}(K_h(x-X)g^p(X)) = \int_{\mathbb{R}^d} \frac{1}{h^d} K\left(\frac{x-u}{h}\right)\varphi_n(u)du \qquad (6)\\
&= \int_B K(y)\varphi_n(x-hy)dy,
\end{aligned}$$

with a classical change of variable, and since K has a compact support. We thus can write:

$$\mathbb{E}\widehat{\varphi}_n(x) - \varphi_n(x) = \int_B K(y)\left[\varphi_n(x-hy) - \varphi_n(x)\right]dy.$$

Consider now the decomposition below:

$$|\varphi_n(x-hy) - \varphi_n(x)| \leq f(x-hy)\,|g^p(x-hy) - g^p(x)| + g^p(x)\,|f(x-hy) - f(x)| := T_1 + T_2.$$

Following Lemma 4,

$$\begin{aligned}
T_1 &= f(x-hy)g^p(x)\left|\frac{g^p(x-hy)}{g^p(x)} - 1\right| \leq 2f_{\max}\frac{L_g}{g_{\min}}g^p(x)ph^\alpha = g^p(x)O(ph^\alpha),\\
T_2 &\leq g^p(x)L_f h^\beta = g^p(x)O(h^\beta) = g^p(x)o(ph^\alpha),
\end{aligned}$$

where L_f and L_g are the Lipschitz constants of the functions f and g. Finally,

$$\mathbb{E}\widehat{\varphi}_n(x) - \varphi_n(x) = g^p(x)O(ph^\alpha) = \varphi_n(x)O(ph^\alpha),$$

and the conclusion follows. ■

Lemma 2 *Under* **(A.0)**–**(A.4)**, *if* $ph^\alpha \to 0$ *then for all* $x \in \mathbb{R}^d$,

$$\mathbb{V}(\widehat{\varphi}_n(x)) = \frac{1}{nh^d}\frac{(p+1)^2}{2p+1}\int_B K^2(s)ds\frac{\varphi_n^2(x)}{f(x)}\left[1+o(1)\right].$$

Proof : We have

$$\begin{aligned}\mathbb{V}(\widehat{\varphi}_n(x)) &= \frac{1}{n^2}\sum_{i=1}^n \mathbb{V}\left(K_h(x-X_i)Z_i\right) = \frac{1}{n}\mathbb{V}\left(K_h(x-X)Z\right)\\ &= \frac{1}{nh^{2d}}\mathbb{E}\left(K^2\left(\frac{x-X}{h}\right)Z^2\right) - \frac{1}{n}\mathbb{E}^2\left(\widehat{\varphi}_n(x)\right) := T_3 + T_4.\end{aligned}$$

From Lemma 1, we immediately derive

$$T_4 = \frac{1}{n}\varphi_n^2(x)\left[1+o(1)\right].$$

We shall prove that

$$T_3 = \frac{1}{nh^d}\frac{(p+1)^2}{2p+1}\int_B K^2(s)ds\frac{\varphi_n^2(x)}{f(x)}\left[1+o(1)\right], \tag{7}$$

leading to $T_4/T_3 = O(h/p)$, and the announced result follows. To this end, remark that

$$\begin{aligned}T_3 &= \frac{1}{nh^{2d}}\mathbb{E}\left(K^2\left(\frac{x-X}{h}\right)\mathbb{E}(Z^2|X)\right)\\ &= \frac{1}{nh^{2d}}\frac{(p+1)^2}{2p+1}\mathbb{E}\left(K^2\left(\frac{x-X}{h}\right)g^{2p}(X)\right)\\ &= \frac{1}{nh^d}\frac{(p+1)^2}{2p+1}\int_B K^2(s)ds\int_{\mathbb{R}^d}\frac{1}{h^d}Q\left(\frac{x-u}{h}\right)g^{2p}(u)f(u)du,\end{aligned}$$

where we have introduced the kernel $Q = K^2/\int_B K^2(s)ds$. It is easily seen that the second integral is similar to this appearing in $\mathbb{E}\widehat{\varphi}_n(x)$, (see (6)), with K replaced by Q and p by $2p$. Thus, as in the proof of Lemma 1, we have

$$\int_{\mathbb{R}^d}\frac{1}{h^d}Q\left(\frac{x-u}{h}\right)g^{2p}(u)f(u)du = g^{2p}(x)f(x)\left[1+o(1)\right] = \frac{\varphi_n^2(x)}{f(x)}\left[1+o(1)\right],$$

and (7) is proved. ■

As a simple consequence of Lemma 1 and Lemma 2, we have

Corollary 1 *Under* **(A.0)**–**(A.4)**, *if* $ph^\alpha \to 0$, *then for all* $x \in \mathbb{R}^d$, $\widehat{\varphi}_n(x)/\varphi_n(x) \xrightarrow{P} 1$.

We can now turn to our main result.

Theorem 2 *Suppose that* $nph^{d+2\alpha} \to 0$ *and* $p/(nh^d) \to 0$. *Let us define*

$$\sigma_n^{-1}(x) = ((2p+1)nh^d)^{1/2} \left(\frac{f(x)}{\int_B K^2(t)dt} \right)^{1/2}.$$

Then, under **(A.0)**–**(A.4)**, *for all* $x \in \mathbb{R}^d$,

$$\sigma_n^{-1}(x) \left(\frac{\widehat{g}_n(x)}{g(x)} - 1 \right) \xrightarrow{d} N(0,1).$$

Proof : First, note that $nph^{d+2\alpha} \to 0$ and $p/(nh^d) \to 0$ imply $ph^\alpha \to 0$. From Lemma 8, it suffices to prove that

$$\xi_n := \frac{\sigma_n^{-1}(x)}{p} \left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)} - \frac{\mathbb{E}\widehat{\varphi}_n(x)}{\varphi_n(x)} \right) \xrightarrow{d} N(0,1).$$

To this end, define

$$W_{i,n} = \frac{\sigma_n^{-1}(x)}{np} \frac{1}{\varphi_n(x)} K_h(x - X_i) Z_i$$

so that we can write

$$\xi_n = \sum_{i=1}^n \left(W_{i,n} - \mathbb{E}W_{i,n} \right).$$

Following Lemma 2, we have

$$\begin{aligned} \mathbb{V}(\xi_n) &= n\mathbb{V}(W_{1,n}) = \frac{\sigma_n^{-2}(x)}{p^2} \frac{1}{\varphi_n^2(x)} \mathbb{V}(\widehat{\varphi}_n(x)) \\ &= \frac{(2p+1)\, nh^d f(x)}{p^2 \int_B K^2(s)ds} \frac{1}{\varphi_n^2(x)} \frac{1}{nh^d} \frac{(p+1)^2}{2p+1} \int_B K^2(s)ds \frac{\varphi_n^2(x)}{f(x)} \left[1 + o(1)\right] \\ &= 1 + o(1). \end{aligned} \tag{8}$$

Thus, the condition of Lyapounov reduces to

$$\sum_{i=1}^n \mathbb{E} \left| W_{i,n} - \mathbb{E}W_{i,n} \right|^3 = n\mathbb{E} \left| W_{1,n} - \mathbb{E}W_{1,n} \right|^3 \to 0. \tag{9}$$

In view of Lemma 9, we need the three first moments of $W = W_{1,n}$. From Lemma 1,

$$\begin{aligned} \mathbb{E}(W) &= \frac{\sigma_n^{-1}(x)}{np} \frac{1}{\varphi_n(x)} \mathbb{E}(K_h(x - X)Z) = \frac{\sigma_n^{-1}(x)}{np} \frac{1}{\varphi_n(x)} \mathbb{E}(\widehat{\varphi}_n(x)) \\ &= \frac{\sigma_n^{-1}(x)}{np}(1 + o(1)) = \left(\frac{2h^d}{np} \right)^{1/2} \left(\frac{f(x)}{\int_B K^2(s)ds} \right)^{1/2} (1 + o(1)). \end{aligned} \tag{10}$$

Remarking that $\mathbb{E}(W^2) = \mathbb{V}(W) + (\mathbb{E}W)^2$ and taking (8), (10) into account, it follows that

$$\mathbb{E}(W^2) = \frac{1}{n} \left[1 + o(1)\right] + O\left(\frac{h^d}{np} \right) = \frac{1}{n}(1 + o(1)). \tag{11}$$

Introducing the kernel $K^3 / \int_B K^3(s)ds$, and mimicking the proof of Lemma 2, we obtain

$$\mathbb{E}(W^3) = \frac{n^{-3/2} h^{-d/2} p^{1/2} 2^{3/2}}{3 f(x)^{3/2}} \frac{\int_B K^3(s)ds}{\left(\int_B K^2(s)ds \right)^{3/2}} (1 + o(1)) = \kappa n^{-3/2} h^{-d/2} p^{1/2} (1 + o(1)), \tag{12}$$

where κ is a positive constant. Collecting (10), (11) and (12), it follows that

$$\begin{aligned} \mathbb{E}(W)\mathbb{E}(W^2) &= o\left(\mathbb{E}(W^3) \right) \\ \mathbb{E}^3(W) &= o\left(\mathbb{E}(W^3) \right), \end{aligned}$$

so that, from Lemma 9,

$$\mathbb{E}\left|W-\mathbb{E}W\right|^3=\mathbb{E}(W^3)(1+o(1))=\kappa n^{-3/2}h^{-d/2}p^{1/2}(1+o(1)).$$

Returning to (9), we have

$$\sum_{i=1}^{n}E\left|W_{i,n}-\mathbb{E}W_{i,n}\right|^3=n\mathbb{E}\left|W_n-\mathbb{E}W_n\right|^3=\kappa\left(\frac{p}{nh^d}\right)^{1/2}(1+o(1))\to 0,$$

and the result is proved. ■

Remark 1 *Theorem 2 holds when $\sigma_n^{-1}(x)$ is replaced*

$$\widehat{\sigma}_n^{-1}(x)=((2p+1)nh^d)^{1/2}\left(\frac{\widehat{f}_n(x)}{\int_B K^2(t)dt}\right)^{1/2},$$

since in this context $\widehat{f}_n(x)\xrightarrow{P}f(x)$. This allows to produce pointwise confident intervals for the frontier.

Remark 2 *To fulfill the assumptions of Theorem 2, one can choose $h=n^{-1/(d+\alpha)}$ and $p=\varepsilon_n n^{\alpha/(d+\alpha)}$, where (ε_n) is a sequence tending to zero arbitrarily slowly. These choices yield*

$$\sigma_n^{-1}(x)=\varepsilon_n^{1/2}n^{\alpha/(d+\alpha)}\left(\frac{2f(x)}{\int_B K^2(t)dt}\right)^{1/2}(1+o(1)),$$

which is the optimal speed (up to the ε_n factor) for estimating $\alpha-$ Lipschitzian $d-$ dimensional frontiers, see [22], Chapter 5.

The good performances of $\hat{g}_n(x)$ on finite sample situations are illustrated in the next section. Remark 2 will be of great help to choose p and h sequences.

5 Numerical experiments

Here, we limit ourselves to unidimensional random variables X ($p=1$) with compact support $E=[0,1]$. Besides, Y given $X=x$ is uniformly distributed on $[0,g(x)]$. The behavior of the proposed frontier estimator is investigated on different situations:

- Two distributions are considered for X: a uniform distribution $U([0,1])$ and a beta distribution $B(2,2)$.
- Two frontiers are considered. The first one

$$g_1(x)=\left|\begin{array}{ll}1+\exp\left(-60(x-1/4)^2\right) & \text{if } 0\le x\le 1/3,\\ 1+\exp\left(-5/12\right) & \text{if } 1/3<u\le 2/3,\\ 1+5\exp\left(-5/12\right)-6\exp\left(-5/12\right)x & \text{if } 2/3<u\le 5/6,\\ 6x-4 & \text{if } 5/6<u\le 1.\end{array}\right.$$

is continuous but is not derivable at $x = 1/3$, $x = 2/3$ and $x = 5/6$. The second one

$$g_2(x) = (1/10 + \sin(\pi x)) \left(11/10 - \exp\left(-64(x - 1/2)^2\right)/2\right)$$

is C^∞.

- Two sample sizes are used $n_1 = 300$ and $n_2 = 150$.

The following kernel is chosen

$$K(t) = \cos^2(\pi t/2)\mathbf{1}\{t \in [-1, 1]\},$$

with associated window width $h = 4\hat{\sigma}(X)n^{-1/2}$ and with $p = n^{1/2}$. The dependence of these sequences with respect to n is chosen according to Remark 2. The multiplicative constants are chosen empirically. The experiment involves several steps:

- First, $m = 100$ replications of the sample are simulated.
- For each of the m previous set of points, the frontier estimator $\hat{g}_n$ is computed.
- The m associated L_1 distances to g are evaluated on a grid.
- Finally, the best situation (ie the estimation corresponding to the smallest L_1 error) and the worst situation (ie the estimation corresponding to the largest L_1 error) are represented.

Results are depicted on Figure 1–3. Note that, even in the worst situations, the empirical choices of sequences h and p seem satisfying for all the considered frontiers and densities of X. In fact, the worst situations are obtained when no points were simulated at the boundaries of the $[0, 1]$ interval. This is specially the case on Figure 2(b) since the density of the beta distribution decreases to 0 at the boundaries of this interval.

6 Appendix: Auxiliary lemmas

The following lemma provides convenient bounds obtained by a specific study of the functions $u \to |(1 + u)^p - 1| - 2p|u|$ and $u \to (1 + u)^{1/p} - 1 - \frac{1}{p}u$. The study is left to the reader. Note that these bounds could not be directly derived from the Taylor formulas $|(1 + u)^p - 1| = |pu + o(u)|$ and $\left|(1 + u)^{1/p} - 1 - \frac{1}{p}u\right| = \left|\frac{1}{2p}(\frac{1}{p} - 1)u^2 + o(u^2)\right|$ where the dependence on p of $o(u)$ and $o(u^2)$ is not precised.

Lemma 3 *Suppose $p \geq 1$.*

(i) *Then, $p|u| \leq \ln 2$ entails $|(1 + u)^p - 1| \leq 2p|u|$.*

(ii) *Let $C \geq 2$. Then, $|u| < 1/2$ entails $\left|(1 + u)^{1/p} - 1 - \frac{1}{p}u\right| \leq \frac{C}{p}u^2$.*

The next lemma is dedicated to the control of the local variations of the frontier on a neighborhood of size h.

Lemma 4 *Suppose* **(A.1)**, **(A.2)** *hold. If* $ph^\alpha \to 0$ *and* $\|x-y\| \le h$, *then for sufficiently large* n,

$$\left|\left(\frac{g(x)}{g(y)}\right)^p - 1\right| \le 2\frac{L_g}{g_{\min}} ph^\alpha,$$

where L_g *is the Lipschitz constant of the function* g.

Proof : Take $u = \frac{g(x)}{g(y)} - 1$ and observe that $p|u| \le p\frac{L_g}{g_{\min}}\|x-y\|^\alpha$, Thus, if $\|x-y\| \le h$, and $ph^\alpha \to 0$, we have $p|u| \le \ln 2$ for sufficiently large n. Then, following Lemma 3(i), for sufficiently large n, we obtain

$$|(1+u)^p - 1| = \left|\left(\frac{g(x)}{g(y)}\right)^p - 1\right| \le 2p|u| \le 2\frac{L_g}{g_{\min}} ph^\alpha,$$

and the result is proved. ■

Lemma 5 will reveal useful to establish the almost complete convergence of random variables ratio.

Lemma 5 *Let* S, T *be real random variables,* a, b *non zero real numbers, and* $0 < \eta < 1/2$. *Then,*

$$\left\{\left|\frac{S}{T} - \frac{a}{b}\right| > 4\eta\left|\frac{a}{b}\right|\right\} \subseteq \left\{\left|\frac{S}{a} - 1\right| > \eta\right\} \cup \left\{\left|\frac{T}{b} - 1\right| > \eta\right\}.$$

Proof : Consider the following obvious equality:

$$\left(\frac{S}{T} - \frac{a}{b}\right) = \frac{a}{b}\left(\frac{S}{a} - 1\right) + \frac{a}{b}\left(1 - \frac{T}{b}\right) + \left(\frac{S}{T} - \frac{a}{b}\right)\left(1 - \frac{T}{b}\right). \tag{13}$$

The triangular inequality yields for all $\eta > 0$:

$$\left\{\left|\frac{S}{a} - 1\right| \le \eta\right\} \cap \left\{\left|\frac{T}{b} - 1\right| \le \eta\right\} \subseteq \left\{\left|\frac{S}{T} - \frac{a}{b}\right| \le 2\eta\left|\frac{a}{b}\right| + \eta\left|\frac{S}{T} - \frac{a}{b}\right|\right\}.$$

Taking $0 < \eta < 1$, we obtain

$$\left\{\left|\frac{S}{a} - 1\right| \le \eta\right\} \cap \left\{\left|\frac{T}{b} - 1\right| \le \eta\right\} \subseteq \left\{\left|\frac{S}{T} - \frac{a}{b}\right| \le \frac{2\eta}{1-\eta}\left|\frac{a}{b}\right|\right\}.$$

Finally, note that $\frac{2\eta}{1-\eta} < 4\eta$ for $0 < \eta < 1/2$. ■

The next three lemmas are of great use to deduce successively the asymptotic normality of $\widehat{g}_n(x)$ from $\widehat{r}_n(x)$ and the asymptotic normality of $\widehat{r}_n(x)$ from $\widehat{\varphi}_n(x)$.

Lemma 6 *Let* $x \in \mathbb{R}^d$. *If* $\widehat{f}_n(x)/f(x) \xrightarrow{P} 1$ *and* $\widehat{\varphi}_n(x)/\varphi_n(x) \xrightarrow{P} 1$, *then*

$$\left(\frac{\widehat{r}_n(x)}{r_n(x)} - 1\right) = \left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)} - 1\right) - \left(\frac{\widehat{f}_n(x)}{f(x)} - 1\right)(1 + o_p(1)).$$

Proof : The hypotheses yield $\frac{\widehat{r}_n(x)}{r_n(x)} = \frac{\widehat{\varphi}_n(x)}{\widehat{f}_n(x)} / \frac{\varphi_n(x)}{f(x)} \xrightarrow{P} 1$. Thus it suffices to consider $S = \frac{\widehat{\varphi}_n(x)}{\varphi_n(x)}$, $T = \frac{\widehat{f}_n(x)}{f(x)}$, and $a = b = 1$ in the equality (13). ■

Lemma 7 *Let $x \in \mathbb{R}^d$. If $\widehat{f}_n(x)/f(x) \xrightarrow{P} 1$ and $\widehat{\varphi}_n(x)/\varphi_n(x) \xrightarrow{P} 1$, then*

$$\left(\frac{\widehat{g}_n(x)}{g(x)} - 1\right) = \frac{1}{p}\left(\frac{\widehat{r}_n(x)}{r_n(x)} - 1\right)(1 + o_p(1)).$$

Proof : From the hypotheses, $w_n(x) := \frac{\widehat{r}_n(x)}{r_n(x)} - 1 = o_p(1)$. Moreover, following Lemma 3(ii), on the event $\{|w_n(x)| < 1/2\}$ we have:

$$\Delta_n(x) := \left|\left(\frac{\widehat{g}_n(x)}{g(x)} - 1\right) - \frac{1}{p}\left(\frac{\widehat{r}_n(x)}{r_n(x)} - 1\right)\right| = \left|(1 + w_n(x))^{1/p} - 1 - \frac{w_n(x)}{p}\right| \leq C\frac{1}{p}w_n^2(x).$$

We thus have, on the one hand,

$$p\Delta_n(x)\mathbf{1}_{\{|w_n(x)|<1/2\}} = o_p(w_n(x)).$$

On the other hand, for all $\varepsilon > 0$,

$$\left\{p\frac{\Delta_n(x)}{w_n(x)}\mathbf{1}_{\{|w_n(x)|\geq 1/2\}} > \varepsilon\right\} \subseteq \{|w_n(x)| \geq 1/2\}$$

leading to

$$P\left\{p\frac{\Delta_n(x)}{w_n(x)}\mathbf{1}_{\{|w_n(x)|\geq 1/2\}} > \varepsilon\right\} \leq P\{|w_n(x)| \geq 1/2\} \to 0,$$

and thus

$$p\frac{\Delta_n(x)}{w_n(x)}\mathbf{1}_{\{|w_n(x)|\geq 1/2\}} = o_p(w_n(x)),$$

which completes the proof. ■

Lemma 8 *Suppose that $nph^{d+2\alpha} \to 0$ and $p/(nh^d) \to 0$. Let us define*

$$\sigma_n^{-1}(x) = ((2p+1)nh^d)^{1/2}\left(\frac{f(x)}{\int_B K^2(t)dt}\right)^{1/2},$$

and let Q be an arbitrary distribution. Then, under **(A.0)**–**(A.4)**,

$$\left\{\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)} - \frac{\mathbb{E}\widehat{\varphi}_n(x)}{\varphi_n(x)}\right) \xrightarrow{d} Q\right\} \Longrightarrow \left\{\sigma_n^{-1}(x)\left(\frac{\widehat{g}_n(x)}{g(x)} - 1\right) \xrightarrow{d} Q\right\}.$$

Proof : First, note that $nph^{d+2\alpha} \to 0$ and $p/(nh^d) \to 0$ imply $ph^\alpha \to 0$. Thus, from Corollary 1, $\widehat{\varphi}_n(x)/\varphi_n(x) \xrightarrow{P} 1$. Besides, $p/(nh^d) \to 0$ implies $nh^d \to \infty$, and thus, using a classical result on density estimation (see for instance [1], Chapter 4, Theorem II.1), we have $\widehat{f}_n(x)/f(x) \xrightarrow{P} 1$. Lemma 6 thus entails

$$\begin{aligned}
\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{r}_n(x)}{r_n(x)} - 1\right) &= \frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)} - 1\right) - \frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{f}_n(x)}{f(x)} - 1\right)(1 + o_p(1)) \\
&= \frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)} - \frac{\mathbb{E}\widehat{\varphi}_n(x)}{\varphi_n(x)}\right) - \frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{f}_n(x)}{f(x)} - \frac{\mathbb{E}\widehat{f}_n(x)}{f(x)}\right)(1 + o_p(1)) \\
&+ \frac{\sigma_n^{-1}(x)}{p}\left(\frac{\mathbb{E}\widehat{\varphi}_n(x)}{\varphi_n(x)} - 1\right) - \frac{\sigma_n^{-1}(x)}{p}\left(\frac{\mathbb{E}\widehat{f}_n(x)}{f(x)} - 1\right)(1 + o_p(1)).
\end{aligned}$$

Following Lemma 1, we have,

$$\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\mathbb{E}\widehat{\varphi}_n(x)}{\varphi_n(x)}-1\right)=O\left(\left(\frac{nh^d}{p}\right)^{1/2}\right)O\left(ph^{\alpha}\right)=O\left(\left(nph^{d+2\alpha}\right)^{1/2}\right)=o(1),$$

and from a classical result on density estimation $\mathbb{E}\widehat{f}_n(x)-f(x)=O(h^{\alpha})$, see [2], Proposition 2.1, we have

$$\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\mathbb{E}\widehat{f}_n(x)}{f(x)}-1\right)=O\left(\left(\frac{nh^d}{p}\right)^{1/2}\right)O\left(h^{\alpha}\right)=O\left(\left(np^{-1}h^{d+2\alpha}\right)^{1/2}\right)=o(1).$$

Consequently,

$$\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{r}_n(x)}{r_n(x)}-1\right)=\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)}-1\right)-\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{f}_n(x)}{f(x)}-1\right)(1+o_p(1))+o_p(1).$$

Again, using a classical result on density estimation, $\mathbb{V}(\widehat{f}_n(x))=O(1/(nh^d))$, see [2], Proposition 2.2, we have

$$\mathbb{V}\left(\frac{\sigma_n^{-1}(x)}{p}\frac{\widehat{f}_n(x)}{f(x)}\right)=O\left(\frac{nh^d}{p}\right)O\left(\frac{1}{nh^d}\right)=O(1/p)=o(1),$$

and thus

$$\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{r}_n(x)}{r_n(x)}-1\right)=\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)}-1\right)+o_p(1). \tag{14}$$

Suppose now that there exists a probability distribution Q such that

$$\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{\varphi}_n(x)}{\varphi_n(x)}-\frac{\mathbb{E}\widehat{\varphi}_n(x)}{\varphi_n(x)}\right)\xrightarrow{d}Q.$$

From (14), we deduce that

$$\frac{\sigma_n^{-1}(x)}{p}\left(\frac{\widehat{r}_n(x)}{r_n(x)}-1\right)\xrightarrow{d}Q.$$

Finally, from Lemma 7 we can conclude that

$$\sigma_n^{-1}(x)\left(\frac{\widehat{g}_n(x)}{g(x)}-1\right)\xrightarrow{d}Q,$$

and the result is proved. ■

Finally, Lemma 9 allows to build bounds for the centered third moment of a positive random variable basing on the non-centered moments.

Lemma 9 *Let W be a positive random variable such that $\mathbb{E}\left|W^3\right|<\infty$, and $m=\mathbb{E}(W)<\infty$. Then,*

$$\mathbb{E}\left(W^3\right)+3m\mathbb{E}\left(W^2\right)-2m^3\leq\mathbb{E}\left|W-m\right|^3\leq\mathbb{E}\left(W^3\right)+3m\mathbb{E}\left(W^2\right).$$

Proof : Remarking that

$$\mathbb{E}\left|W-m\right|^3=\mathbb{E}\left(W-m\right)^3-2\mathbb{E}\left[\left(W-m\right)^3 1_{\{W<m\}}\right],$$

it follows that

$$\mathbb{E}\left(W-m\right)^3\leq\mathbb{E}\left|W-m\right|^3\leq\mathbb{E}\left(W-m\right)^3+2m^3,$$

and the lemma is proved. ■

References

[1] D. Bosq and J.P. Lecoutre. *Théorie de l'estimation fonctionnelle*. Economie et Statistiques avancées. Economica, Paris, 1987.

[2] G. Collomb. *Estimation non paramétrique de la régression par la méthode du noyau*. PhD thesis, Université Paul Sabatier de Toulouse, 1976.

[3] D. Deprins, L. Simar, and H. Tulkens. Measuring labor efficiency in post offices. In P. Pestieau M. Marchand and H. Tulkens, editors, *The Performance of Public Enterprises: Concepts and Measurements*. North Holland ed, Amsterdam, 1984.

[4] M.J. Farrel. The measurement of productive efficency. *Journal of the Royal Statistical Society A*, 120:253–281, 1957.

[5] F. Ferraty and P. Vieu. *Nonparametric modelling for functional data*. Springer, 2005.

[6] L. Gardes. Estimating the support of a Poisson process via the Faber-Shauder basis and extreme values. *Publications de l'Institut de Statistique de l'Université de Paris*, XXXXVI:43–72, 2002.

[7] J. Geffroy. Sur un problème d'estimation géométrique. *Publications de l'Institut de Statistique de l'Université de Paris*, XIII:191–210, 1964.

[8] I. Gijbels, E. Mammen, B. U. Park, and L. Simar. On estimation of monotone and concave frontier functions. *Journal of the American Statistical Association*, 94(445):220–228, 1999.

[9] S. Girard, A. Iouditski, and A. Nazin. L_1-optimal nonparametric frontier estimation via linear programming. *Automation and Remote Control*, 66(12):2000–2018, 2005.

[10] S. Girard and P. Jacob. Extreme values and Haar series estimates of point process boundaries. *Scandinavian Journal of Statistics*, 30(2):369–384, 2003.

[11] S. Girard and P. Jacob. Projection estimates of point processes boundaries. *Journal of Statistical Planning and Inference*, 116(1):1–15, 2003.

[12] S. Girard and P. Jacob. Extreme values and kernel estimates of point processes boundaries. *ESAIM: Probability and Statistics*, 8:150–168, 2004.

[13] S. Girard and L. Menneteau. Central limit theorems for smoothed extreme value estimates of point processes boundaries. *Journal of Statistical Planning and Inference*, 135(2):433–460, 2005.

[14] P. Hall, M. Nussbaum, and S. Stern. On the estimation of a support curve of indeterminate sharpness. *Journal of Multivariate Analysis*, 62(2):204–232, 1997.

[15] P. Hall and B. U. Park. Bandwidth choice for local polynomial estimation of smooth boundaries. *Journal of Multivariate Analysis*, 91(2):240–261, 2004.

[16] P. Hall, B. U. Park, and S. E. Stern. On polynomial estimators of frontiers and boundaries. *Journal of Multivariate Analysis*, 66(1):71–98, 1998.

[17] W. Härdle, B. U. Park, and A. B. Tsybakov. Estimation of a non sharp support boundaries. *Journal of Multivariate Analysis*, 43:205–218, 1995.

[18] P. Jacob and P. Suquet. Estimating the edge of a Poisson process by orthogonal series. *Journal of Statistical Planning and Inference*, 46:215–234, 1995.

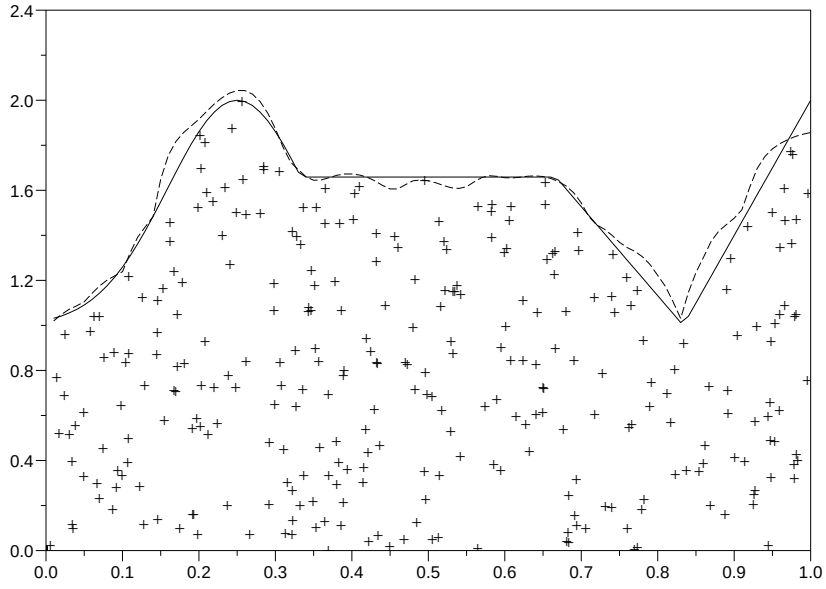
[19] P. Jacob and P. Suquet. Regression and edge estimation. *Statistic and Probability Letters*, 27:11–15, 1996.

[20] K. Knight. Limiting distributions of linear programming estimators. *Extremes*, 4(2):87–103, 2001.

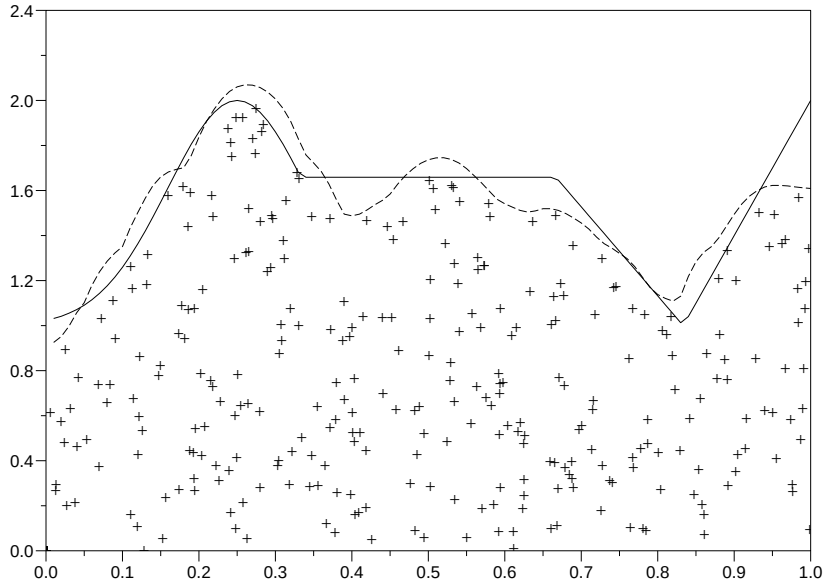
[21] A. Korostelev, L. Simar, and A. B. Tsybakov. Efficient estimation of monotone boundaries. *The Annals of Statistics*, 23:476–489, 1995.

[22] A.P. Korostelev and A.B. Tsybakov. *Minimax theory of image reconstruction*, volume 82 of *Lecture Notes in Statistics*. Springer-Verlag, New-York, 1993.

[23] E. Mammen and A. B. Tsybakov. Asymptotical minimax recovery of set with smooth boundaries. *The Annals of Statistics*, 23(2):502–524, 1995.

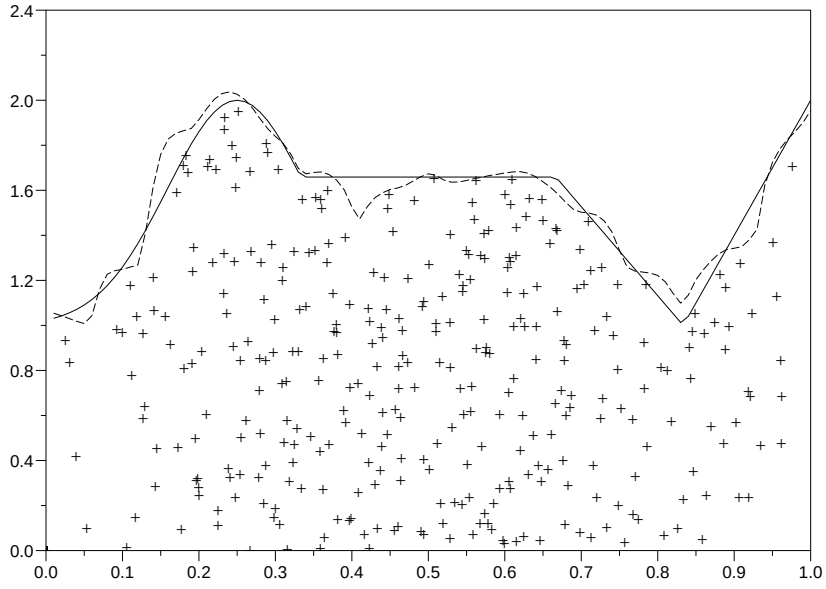


(a) Best situation

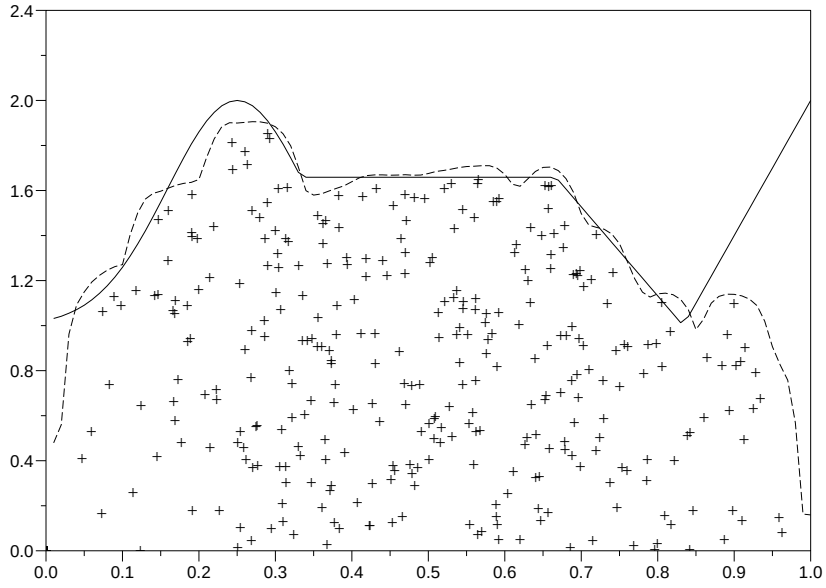


(b) Worst situation

Figure 1: The frontier g_1 (continuous line) and its estimation (dashed line). The sample size is $n_1 = 300$ and X is uniformly distributed on $[0, 1]$.

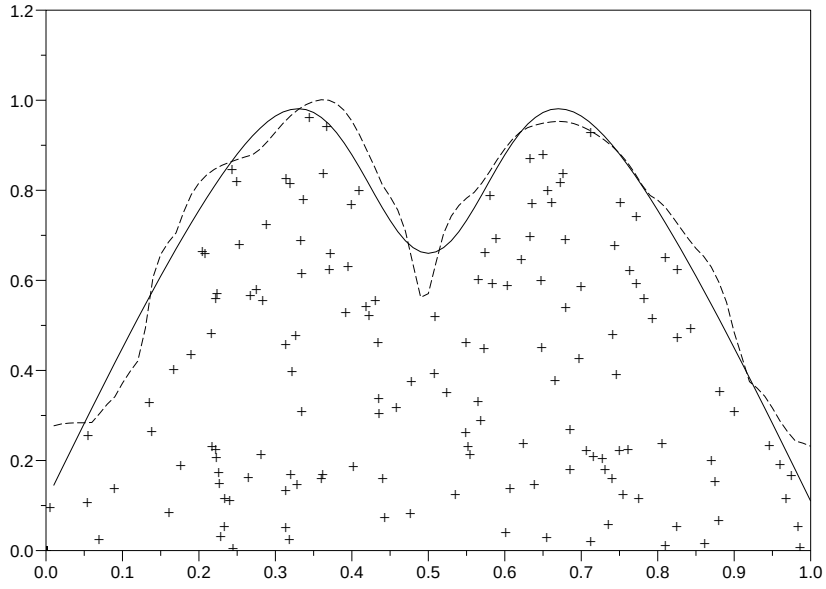


(a) Best situation

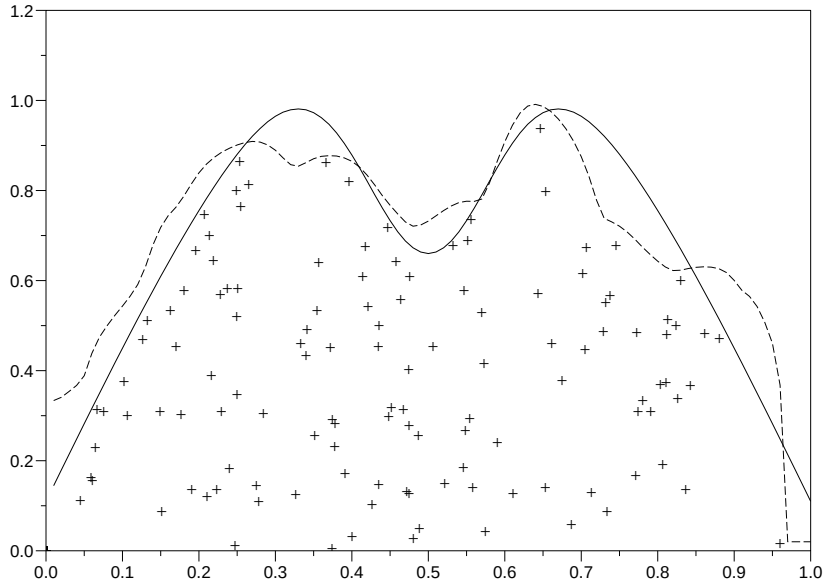


(b) Worst situation

Figure 2: The frontier g_1 (continuous line) and its estimation (dashed line). The sample size is $n_1 = 300$ and X is $B(2, 2)$ distributed $[0, 1]$.



(a) Best situation



(b) Worst situation

Figure 3: The frontier g_2 (continuous line) and its estimation (dashed line). The sample size is $n_2 = 150$ and X is uniformly distributed on $[0, 1]$.