Kernel estimation of discontinuous regression functions

Kee-Hoon Kang[a], Ja-Yong Koo[b], Cheol-Woo Park[a]

[a] *Department of Statistics, Seoul National University, Seoul, 151-742, Korea*

[b] *Department of Statistics, Hallym University, Chunchon, Kangwon-Do 200-702, Korea*

Abstract

A kernel regression estimator is proposed wherein the regression function is smooth, except possibly for a finite number of points of discontinuity. The proposed estimator uses preliminary estimators for the location and size of discontinuities or change-points in an otherwise smooth regression model and then uses an ordinary kernel regression estimator based on suitably adjusted data. Global L_2 rates of convergence of curve estimates are derived. It is shown that these rates of convergence are the same as those for ordinary kernel regression estimators of smooth curves. Moreover, pointwise asymptotic normality is also obtained. The proposed method is tested on simulated examples and performs well under our consideration.

Keywords: Boundary kernel, change-points, jump location, jump size, L_2 convergence, rate of convergence, weak convergence.

1. Introduction

We consider the estimation of a regression function which is smooth except possibly for a finite number of points of discontinuity or change-points, which describe sudden localized changes. An ordinary kernel regression method assumes that the regression function is smooth over the entire range of a covariate. Sometimes a generally smooth curve might contain isolated discontinuities or change-points in the curve, in which case an ordinary kernel estimator may show a Gibbs phenomenon, see Fig. 3 in Section 4 for an example of this phenomenon. In such cases we need a kernel regression method which can alleviates such problems.

The main focus of this paper is on the estimation of the entire regression function rather than on the estimation of change-points and jump sizes. In this context, Müller (1992), Donoho (1994) and Koo (1997) use a kernel method, a segmented wavelet method and a spline method, respectively.

Suppose the response measurements Y_i at fixed design points $t_i = i/n,\ i = 1, \cdots, n$ satisfy the following regression model:

$$Y_i = g(t_i) + \varepsilon_i, \quad 1 \le i \le n, \tag{1}$$

where g is the regression function and ε_i are independent errors with $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2 < \infty$. Now, assume that g has a jump at a change-point τ, $0 < \tau < 1$, and is smooth otherwise. It can be deduced that there exists a smooth function $f \in \mathcal{C}^k([0,1])$ for some $k \ge 2$, such that

$$g(t) = f(t) + \Delta 1_{[\tau,1]}(t), \quad 0 \le t \le 1. \tag{2}$$

Define $g_+(\tau) = \lim_{t\downarrow\tau} g(t)$ and $g_-(\tau) = \lim_{t\uparrow\tau} g(t)$. Then the jump size at the change-point τ is given by

$$\Delta = g_+(\tau) - g_-(\tau). \tag{3}$$

The basic idea of the proposed estimator is as follows. First, we adjust data in some way using estimators for the location of the change-point and the corresponding jump size. Define $Z_i = Y_i - \Delta 1_{[\tau,1]}(t_i)$ for $i = 1, \ldots, n$. If Z_i's are available, then one can use an ordinary kernel regression estimator such as (4) to estimate f. Suppose that we have estimators of τ and Δ with good performance. Then, If we subtract our jump size estimate from the observations after the change-point estimate, the adjusted data may be considered as observations with a smooth regression function. Using the adjusted data, we can estimate the smooth function f appearing in g by an ordinary kernel regression estimator. After that, a final estimate can be obtained by adding the jump size estimate to the estimate of g.

Müller (1992) estimated the regression function g in the right and left regions of the estimate of τ separately with the boundary kernels. Accordingly, at most half of the data set can be used to construct an estimate near the change-point. However, the proposed estimator uses the whole data set to construct estimates and does not use boundary kernels in the interior part containing the change-point.

For the asymptotic study we assume that there is a single jump in an otherwise smooth regression function having approximately the same amount of smoothness over the entire range of the covariate except at the change-point. However, the result can be directly extended for the case when there are a finite number of change-points.

Section 2 describes the model studied in this paper and summarizes the estimators for the location of the change-point and the corresponding jump size which are developed in Müller (1992). Section 3 proposes a kernel regression estimator with a change-point and states asymptotic results. The performance in practice of our approach is investigated in Section 4 through some simulated examples. Section 5 gives some concluding remarks. The proofs of asymptotic results are given in Section 6.

2. Jump detection

For the setting (1), one of the most well known estimators of g has the the following form which was discussed in Gasser and Müller (1979):

$$\hat{g}(t) = \frac{1}{b}\sum_{i=1}^{n} Y_i \int_{s_{i-1}}^{s_i} K\left(\frac{t-u}{b}\right) du, \tag{4}$$

where $t \in [0,1]$, $s_i = (t_i + t_{i+1})/2$, $i = 1, \ldots, n-1$, $s_0 = 0$, $s_n = 1$, and $b = b(n)$ is a sequence of bandwidths which is required to satisfy $b \to 0$, $nb \to \infty$ as $n \to \infty$, and K is the kernel function of order k with compact support $[-1, 1]$.

For the setting (2), the regression function g has a change-point and the estimator (3) must be modified. Fig. 3 in Section 4 shows the inadequacy of the unmodified estimator for this setting.

The curve estimation problem considered in this paper is highly dependent on the estimation of the location of τ and the jump-size Δ based on the data (t_i, Y_i), $i = 1, \ldots, n$.

Müller (1992) estimates τ and Δ based on the difference of one-sided kernel estimates:

$$\hat{\Delta}(t) = \hat{g}_+(t) - \hat{g}_-(t), \tag{5}$$

where

$$\hat{g}_\pm(t) = \frac{1}{b}\sum_{i=1}^{n} Y_i \int_{s_{i-1}}^{s_i} K_\pm\left(\frac{t-u}{b}\right) du. \tag{6}$$

Here K_+ and K_- are one-sided kernel functions of order k with support $[-1, 0]$, $[0, 1]$, respectively and $b \to 0$, $nb \to \infty$ as $n \to \infty$. Assumptions for the kernel function K, K_+ and K_- are exactly the same with (K1) and (K2) in Müller (1992), which correspond to the case $\nu = 0, \mu = 1$. For the simplicity, We do not include them here.

Now, $\hat{\tau}$, the estimator of τ, and $\hat{\Delta}$, the estimator of Δ, are defined as follows:

$$\hat{\tau} = \inf\{\rho \in Q \ : \ \hat{\Delta}(\rho) = \sup_{x \in Q} |\hat{\Delta}(x)|\}$$

and

$$\hat{\Delta} \equiv \hat{\Delta}(\hat{\tau}) = \hat{g}_+(\hat{\tau}) - \hat{g}_-(\hat{\tau}), \tag{7}$$

where $Q \subset (0, 1)$ is a closed interval such that $\tau \in Q$. For the change-point estimator $\hat{\tau}$ with bandwidth b_1, Müller (1992) showed the following holds:

$$|\hat{\tau} - \tau| = O_p(n^{-1/2} b_1^{1/2}). \tag{8}$$

For the jump size estimator $\hat{\Delta}$, we get the following lemma.

Lemma 1 *If b_2 is the bandwidth used for the estimation of Δ with $b_1 \le b_2$, then*

$$|\hat{\Delta} - \Delta| = O_p(b_2^{k+1} + n^{-1/2} b_2^{-1/2}).$$

For two positive sequences $\{a_n\}$ and $\{b_n\}$, let $a_n \sim b_n$ mean that $a_n = O(b_n)$ and $b_n = O(a_n)$ as $n \to \infty$. Then, if we choose $b_1 \sim n^{-(2k-1)/(2k+1)}$ and $b_2 \sim n^{-1/(2k+3)}$, the rates of convergence of $|\hat{\tau} - \tau|$ and $|\hat{\Delta} - \Delta|$ become $O_p(n^{-2k/(2k+1)})$ and $O_p(n^{-(k+1)/(2k+3)})$, respectively.

3. Curve estimation and asymptotic results

Using the preliminary estimators of τ and Δ defined in Section 2, we compute the adjusted data as

$$Y_i^* = Y_i - \hat{\Delta} 1_{[\hat{\tau},1]}(t_i), \quad 1 \le i \le n. \tag{9}$$

Given Y_i^*'s, the proposed regression function estimator has the following form:

$$\hat{g}^*(t) = \frac{1}{b} \sum_{i=1}^{n} Y_i^* \int_{s_{i-1}}^{s_i} K^* \left(\frac{t-u}{b} \right) du + \hat{\Delta} 1_{[\hat{\tau},1]}(t), \tag{10}$$

where K^* is defined by

$$K^* \left(\frac{t-u}{b} \right) = \begin{cases} K_+ \left(\frac{t-u}{b}, q \right) & \text{for } 0 \le t \le b \text{ with } q = \frac{t}{b}, \\ K \left(\frac{t-u}{b} \right) & \text{for } b \le t \le 1-b, \\ K_- \left(\frac{t-u}{b}, q \right) & \text{for } 1-b \le t \le 1 \text{ with } q = \frac{1-t}{b}. \end{cases} \tag{11}$$

Here, $K_+(\cdot, q)$ and $K_-(\cdot, q)$ are kernel functions with more general asymmetric supports $[-1, q]$, $[-q, 1]$, respectively. The assumptions on $K_+(\cdot, q)$ and $K_-(\cdot, q)$ are described in a similar way as (2.8) and (2.9) in Müller (1992).

The proposed estimator can apply the ordinary method and needs no more new idea except translation. It is natural and very simple in computational aspect.

The convergence rate in the L_2 sense of the proposed regression function estimator is given by the following theorem, whose proof is provided in Section 6. Theorem 1(b) states that our estimator results in the same convergence rate for the case when the regression function g is k-times differentiable.

Theorem 1 *Suppose that assumptions for $K, K_+, K_-, K_+(\cdot, q)$ and $K_-(\cdot, q)$ hold.*

(a) If b_1 and b_2 are the bandwidths used for the estimation of τ and Δ, respectively and b is the bandwidth for $\hat{g}^*(t)$ in (10) with $b_1 \le b \le b_2$, then

$$\int_0^1 |\hat{g}^*(t) - g(t)|^2 dt = O_p \left(\frac{1}{nb} \right) + O_p(b^{2k}) + O_p \left(\left(\frac{b_1}{n} \right)^{1/2} \right) + O_p \left(\frac{b_1^{1/4}}{n^{3/4} b_2^{1/2}} \right).$$

(b) If $b_1 \sim n^{-(2k-1)/(2k+1)}$, $b_2 \sim n^{-1/(2k+3)}$ and $b \sim n^{-1/(2k+1)}$, then

$$\int_0^1 |\hat{g}^*(t) - g(t)|^2 dt = O_p \left(n^{-2k/(2k+1)} \right).$$

Furthermore, we get pointwise asymptotic normality of the proposed estimator, as described in the following theorem.

Theorem 2 *Under the assumptions of Theorem 1,*

$$(nb)^{1/2}(\hat{g}^*(t) - g(t)) \xrightarrow{\mathcal{D}} \mathcal{N}(f^{(k)}(t)B_k,\ \sigma^2 V),$$

where

$$B_k = \frac{1}{k!}\int x^k K^*(x)dx \text{ and } V = \int K^{*^2}(x)dx.$$

Using this result, one can obtain asymptotic $100(1-\alpha)\%$ confidence intervals for $g(t)$ with consistent estimator $\hat{\sigma}^2$ for σ^2. A general class of such estimators was considered in Hall, Kay and Titterington (1990).

4. Numerical results

A program for implementing our procedure as it applies to data with a possible discontinuity has been written in FORTRAN. As in Müller (1992), one-sided kernel estimates $\hat{g}_+$ and $\hat{g}_-$, which are defined in (6), employ one-sided kernels $K_+(x) = 6(1+x)(1+2x)$, $-1 \le x \le 0$, $K_-(x) = 6(1-x)(1-2x)$, $0 \le x \le 1$, respectively, for all examples considered in this section.

To investigate the practical performance of the proposed estimator defined in Section 3, a simulation study is carried out. For the simulation study, n response-predictor pairs (t_i, Y_i) are generated according to the prescription

$$Y_i = g(t_i) + \varepsilon_i, \quad i = 1, \ldots, n,$$

where $t_i = i/n$ and ε_i are drawn from a normal distribution with standard deviation σ. The sample size for the simulated examples is chosen by $n = 512$.

Fig. 1 about here.

Fig. 1 displays the true regression function and the scatterplot of the simulated data. The regression function for this example is given by

$$g_1(t) = \begin{cases} 4t^2(3-4t) & \text{for } 0 \le t \le \frac{1}{2}, \\ \frac{4}{3}t(4t^2 - 10t + 7) - \frac{3}{2} & \text{for } \frac{1}{2} < t \le \frac{3}{4}, \\ \frac{16}{3}t(t-1)^2 & \text{for } \frac{3}{4} < t \le 1. \end{cases}$$

Gaussian white noise with $\sigma = 0.15$ is added to produce the simulated data. This regression function was contrived by Nason and Silverman (1994).

Fig. 2 about here.

In Fig. 2, the regression function is

$$g_2(t) = 2 - 2|t - 0.26|^{1/5}1(t \le 0.26) - 2|t - 0.26|^{3/5}1(t > 0.26) + 1(t \ge 0.78)$$

which appears in Wang (1995). The standard deviation of the Gaussian noise is $\sigma = 0.2$. In this case g_2 has a jump and an unbalanced cusp.

Fig. 3 about here.

For the first example, the unmodified kernel estimate $\hat{g}(\cdot)$ defined in (4) (modifications made at the boundaries, but not at a possible change-point) is shown in Fig. 3. The kernel function K used for this estimate is $K(x) = 3(1 - x^2)/4$, and the bandwidth is determined visually and chosen as $b = 0.1$. The strong evidence for the existence of a change point is smeared out. It is interesting to compare it with Fig. 6 and 7 in Nason and Silverman (1994): their method also blurs discontinuity in the data because they do not consider the detection of a jump.

With a choice of bandwidth $b = 0.1$, the estimator $\hat{\tau}$ of τ is found as the maximizer of the function $|\hat{\Delta}(t)|$ and is given by $\hat{\tau} = 0.5$. For this data set the estimator $\hat{\tau}$ gives the true value 0.5 exactly. To check whether this is common or not, we have done the experiment described above 10 times holding all factors constant, except the noise that is generated. In most cases (9 out of 10) the estimates of the change-point coincide with the true value and even in the exceptional case the estimate is 0.499 which is extremely close to the change-point. This may be explained by the faster convergence of the estimator for the change-point. For estimating Δ we use the same bandwidth as used in $\hat{\tau}$ which results in $\hat{\Delta}(\hat{\tau}) = -0.513$. The coupling of bandwidth choice with the problem of finding the number of change-points is discussed in Braun and Müller (1998).

Fig. 4 about here.

By applying our method defined in (10) with the same bandwidth $b = 0.1$, we get the final estimate of the regression curve which is presented in Fig. 4. Comparing Fig. 4 with Fig. 3 displayed before, and Fig. 7 in Nason and Silverman (1994), we believe that our method appears to perform reasonably well in that it can alleviate the Gibbs phenomenon or peak shrinkage around the change-point.

Fig. 5 about here.

Fig. 5 shows the final estimate of regression function g_2. The same kernel functions as in the first example with a bandwidth $b = 0.1$ are used. We obtain the estimated change-point $\hat{\tau} = 0.781$ with jump size $\hat{\Delta}(\hat{\tau}) = 0.976$. Although detection of the sharp cusp is not satisfactory, global features of the estimated regression function look quite good.

5. Concluding remarks

We have established theoretical properties of a kernel estimator of a regression function with a change-point and considered practical aspects of the proposed estimator via numerical study. Regarding the rate of convergence given in Theorem 1(b), it is the same as the usual rate of convergence for the case when the regression function g is k-times differentiable. It is anticipated that this rate of convergence is actually optimal in the sense of Stone (1982) or Yatracos (1988). According to a limited simulation study, we believe that our method can actually alleviate the Gibbs phenomenon or peak shrinkage.

We have described a kernel regression estimator when the regression function has a single change-point, which appears to be easier to understand and computationally simpler to implement than the method of Müller (1992). In cases where the regression function has multiple discontinuity points, first, the number of discontinuity points should be obtained from prior knowledge about the process under study or should be estimated by the method of Yin (1988). Then our estimator can be easily extended as in the case of single change-point.

The basic idea of this paper might be applied to other methods of estimating a regression function including wavelet methods. Considering the wavelet estimates in Nason and Silverman (1995), which show peak erosion even if they are known to have local adaptivity for spatially varying functions, an algorithm taking care of discontinuity should be involved. However, due to the local adaptivity of wavelets for various functions, it seems to us that a method using wavelets is a promising tool, for example, near the cusp point in Fig. 2. For jump detection, we can use the results of Wang (1995) or Raimondo (1998). Once we have wavelet estimators for change-point and size, we can apply the usual wavelet method to the adjusted data.

6. Proofs

6.1. Proof of Theorem 1

To prove Theorem 1, we first write

$$\begin{aligned}\hat{g}^*(t) - g(t) &= \sum_{i=1}^{n} W_i \int_{s_{i-1}}^{s_i} K_b^*(t-u)du - f(t) \\ &\quad + \sum_{i=1}^{n} (\Delta 1_{[\tau,1]}(t_i) - \hat{\Delta} 1_{[\hat{\tau},1]}(t_i)) \int_{s_{i-1}}^{s_i} K_b^*(t-u)du + \hat{\Delta} 1_{[\hat{\tau},1]}(t) - \Delta 1_{[\tau,1]}(t),\end{aligned}$$

where $K_b^*(\cdot) = K^*(\cdot/b)/b$ and $W_i = f(t_i) + \varepsilon_i$. So,

$$\begin{aligned}\int_0^1 |\hat{g}^*(t) - g(t)|^2 dt &\leq 3\int_0^1 \left(\sum_{i=1}^{n} W_i \int_{s_{i-1}}^{s_i} K_b^*(t-u)du - f(t) \right)^2 dt \\ &\quad + 3\int_0^1 \left(\sum_{i=1}^{n} (\hat{\Delta} 1_{[\hat{\tau},1]}(t_i) - \Delta 1_{[\tau,1]}(t_i)) \int_{s_{i-1}}^{s_i} K_b^*(t-u)du \right)^2 dt \\ &\quad + 3\int_0^1 |\hat{\Delta} 1_{[\hat{\tau},1]}(t) - \Delta 1_{[\tau,1]}(t)|^2 dt.\end{aligned}$$

The first term in (12) is the integrated squared error of the smooth regression function estimate without any discontinuity. Since K_b^* is a boundary kernel in the boundary region, we obtain

$$\int_0^1 \left(\sum_{i=1}^n W_i \int_{s_{i-1}}^{s_i} K_b^*(t-u)du - f(t) \right)^2 dt = O_p\left(b^{2k} + \frac{1}{nb} \right). \quad (12)$$

¿From (8) and Lemma 1, the second term in (12) can be shown as

$$\int_0^1 \left(\sum_{i=1}^n (\hat{\Delta} 1_{[\hat{\tau},1]}(t_i) - \Delta 1_{[\tau,1]}(t_i)) \int_{s_{i-1}}^{s_i} K_b^*(t-u)du \right)^2 dt = O_p\left(\frac{1}{nb} \right). \quad (13)$$

The third term in (12) can be written as

$$\int_0^1 (\hat{\Delta} - \Delta)^2 1_{[\hat{\tau},1]}(t)dt + \Delta^2 \int_0^1 (1_{[\hat{\tau},1]}(t) - 1_{[\tau,1]}(t))^2 dt$$
$$+ 2\Delta \int_0^1 |\hat{\Delta} - \Delta| 1_{[\hat{\tau},1]}(t) |1_{[\hat{\tau},1]}(t) - 1_{[\tau,1]}(t)| dt.$$

Let us observe that

$$\int_0^1 |\hat{\Delta} - \Delta| 1_{[\hat{\tau},1]}(t) |1_{[\hat{\tau},1]}(t) - 1_{[\tau,1]}(t)| dt = O_p\left(\frac{b_1^{1/4}}{n^{3/4} b_2^{1/2}} \right).$$

Then from (8) and Lemma 1,

$$\int_0^1 |\hat{\Delta} 1_{[\hat{\tau},1]}(t) - \Delta 1_{[\tau,1]}(t)|^2 dt$$
$$= O_p\left(\frac{1}{nb} \right) + O_p\left(\left(\frac{b_1}{n} \right)^{1/2} \right) + O_p\left(\frac{b_1^{1/4}}{n^{3/4} b_2^{1/2}} \right). \quad (14)$$

Now from (12), (13) and (14), Theorem 1(a) holds. Theorem 1(b) is a direct result. □

6.2 Proof of Theorem 2

$$\begin{aligned} (nb)^{1/2}(\hat{g}^*(t) - g(t)) &= (nb)^{1/2} \left(\sum_{i=1}^n W_i \int_{s_{i-1}}^{s_i} K_b^*(t-u)du - f(t) \right) \\ &\quad + (nb)^{1/2} \left(\sum_{i=1}^n (\Delta 1_{[\tau,1]}(t_i) - \hat{\Delta} 1_{[\hat{\tau},1]}(t_i)) \int_{s_{i-1}}^{s_i} K_b^*(t-u)du \right) \\ &\quad + (nb)^{1/2} \left(\hat{\Delta} 1_{[\hat{\tau},1]}(t) - \Delta 1_{[\tau,1]}(t) \right) \\ &= A + B + C. \end{aligned}$$

Some calculations lead to

$$E[A] = \frac{1}{k!} f^{(k)}(t) \int v^k K_+(v)dv + o(1), \;\; Var[A] = \sigma^2 \int K^{*^2} + o(1).$$

¿From (4.15) and (4.16) in Müller (1988),

$$A \xrightarrow{\mathcal{D}} \mathcal{N}(f^{(k)}(t)B_k, \sigma^2 V).$$

It is easy to show that $B \to 0$, $C \to 0$ in probability, which imply the result holds. □

Acknowledgement

We would like to thank to Hans-Georg Müller for helpful discussions. This research was supported by KOSEF (980701-3).

References

Braun, J. V., Müller, H. G., 1998. Statistical methods for DNA sequence segmentation. Statist. Sci. 13, 142–62.

Carlstein, E., Müller, H. G., Siegmund, D., 1994. Change-Point Problems, Lecture Notes. Vol. 23. Hayward: Institute of Mathematical Statistics.

Cobb, G. W., 1978. The problem of the Nile: Conditional solution to a change-point problem. Biometrika 65, 243–51.

Donoho, D. L., 1994. On minimum entropy segmentation. In Wavelets: Theory, Algorithms, and Applications, Ed. Chui, C. K., Montefusco, L., and Puccio, L., pp. 233–69. Academic Press, San Diego.

Eubank, R. L., Speckman, P. L., 1994. Nonparametric estimation of functions with jump discontinuities. In Change-Point Problems, Lecture Notes-Monograph Series, Vol. 23, Ed. Carlstein, E., Müller, H. G., Siegmund, D., pp. 130–44. Hayward: Institute of Mathematical Statistics.

Gasser, T., Müller, H. G., 1979. Kernel estimation of regression functions. In Smoothing Techniques for Curve Estimation, Lecture Notes in Mathematics 757, Ed. Gasser, T. and Rosenblatt, M., pp. 23–68. Springer-Verlag, Heidelberg.

Gordon, K., Smith, A. F., 1990. Modeling and monitoring biomedical time series. J. Amer. Statist. Assoc. 85, 328–37.

Hall, P., Kay, J. W., Titterington, D. W., 1990. Asymptotically optimal difference-based estimation of variance in nonparametric regression. Biometrika 77, 521–28.

Koo, J. -Y., 1997. Spline estimation of discontinuous regression functions. J. Comput. Graphical. Statist. 6, 266–84.

Lombard, F., 1988. Detecting change points by Fourier analysis. Technometrics 30, 305–10.

McDonald, J. A., Owen, A. B., 1986. Smoothing with split linear fits. Technomtrics 28, 195–208.

Müller, H. G., 1988. Nonparametric Regression Analysis of Longitudinal Data. Springer, New York.

Müller, H. G., 1992. Change-points in nonparametric regression analysis. Ann. Statist. 20, 737–61.

Nason, G. P., Silverman, B. W., 1994. The discrete wavelet transform in S. J. Comput. Graphical. Statist. 3, 163–91.

Raimondo, M., 1998., Minimax estimation of sharp change points. Ann. Statist., to appear.

Stone, C. J., 1982. Optimal global rates of convergence for nonparametric regression. Ann. Statist. 10, 1040–53.

Wang, Y., 1995. Jump and sharp cusp detection by wavelets. Biometrika 82, 385–97.

Yatracos, Y. G., 1988. A lower bound on the error in nonparametric regression type problems. Ann. Statist. 16, 1180–87.

Yin, Y. Q., 1988. Detection of the number, locations and magnitudes of jumps. Commun. Statist. Stochastic Models 4, 445–55.

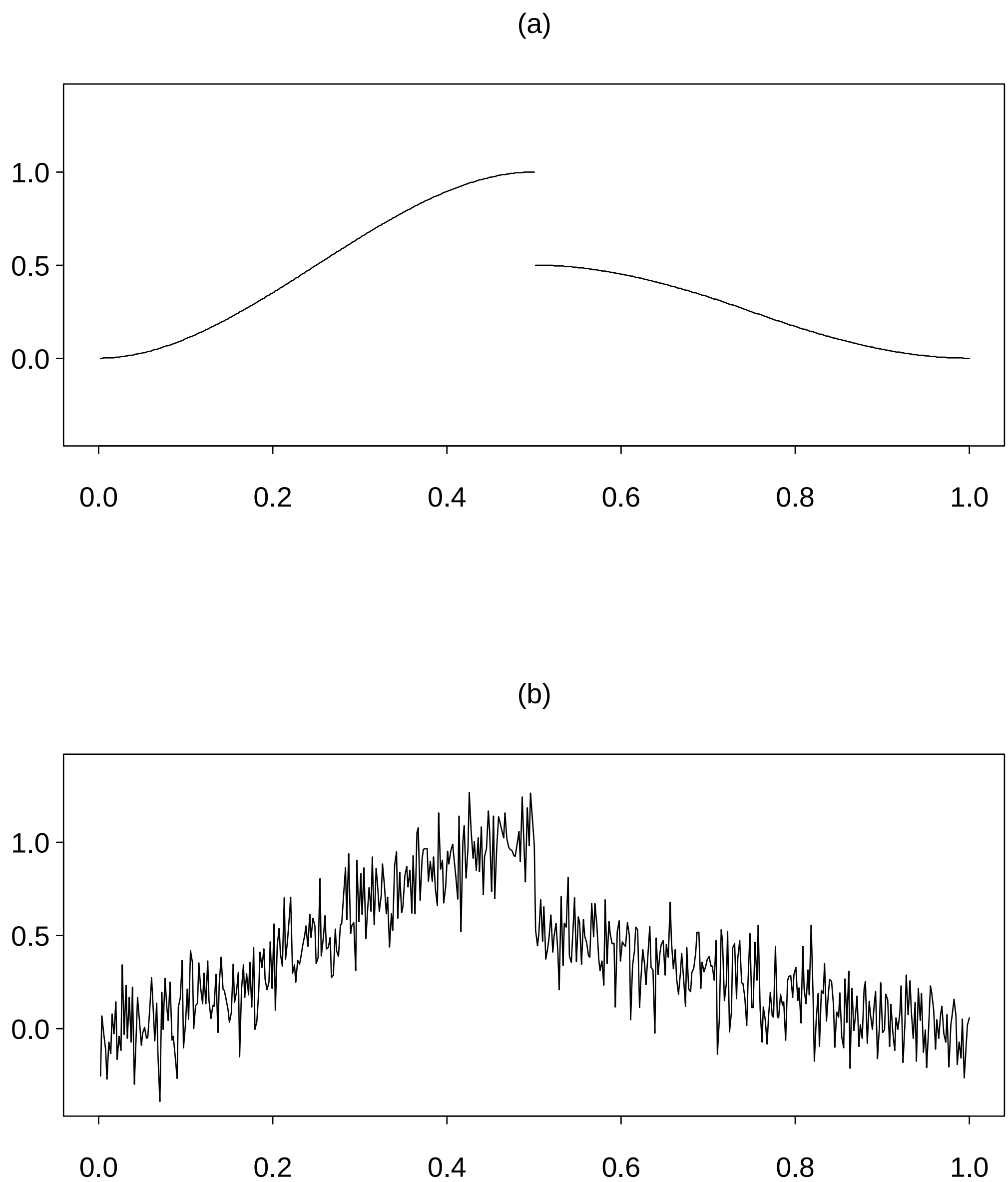

Fig. 1. The case of g_1. (a) displays g_1 and (b) noisy data with gaussian white noise.

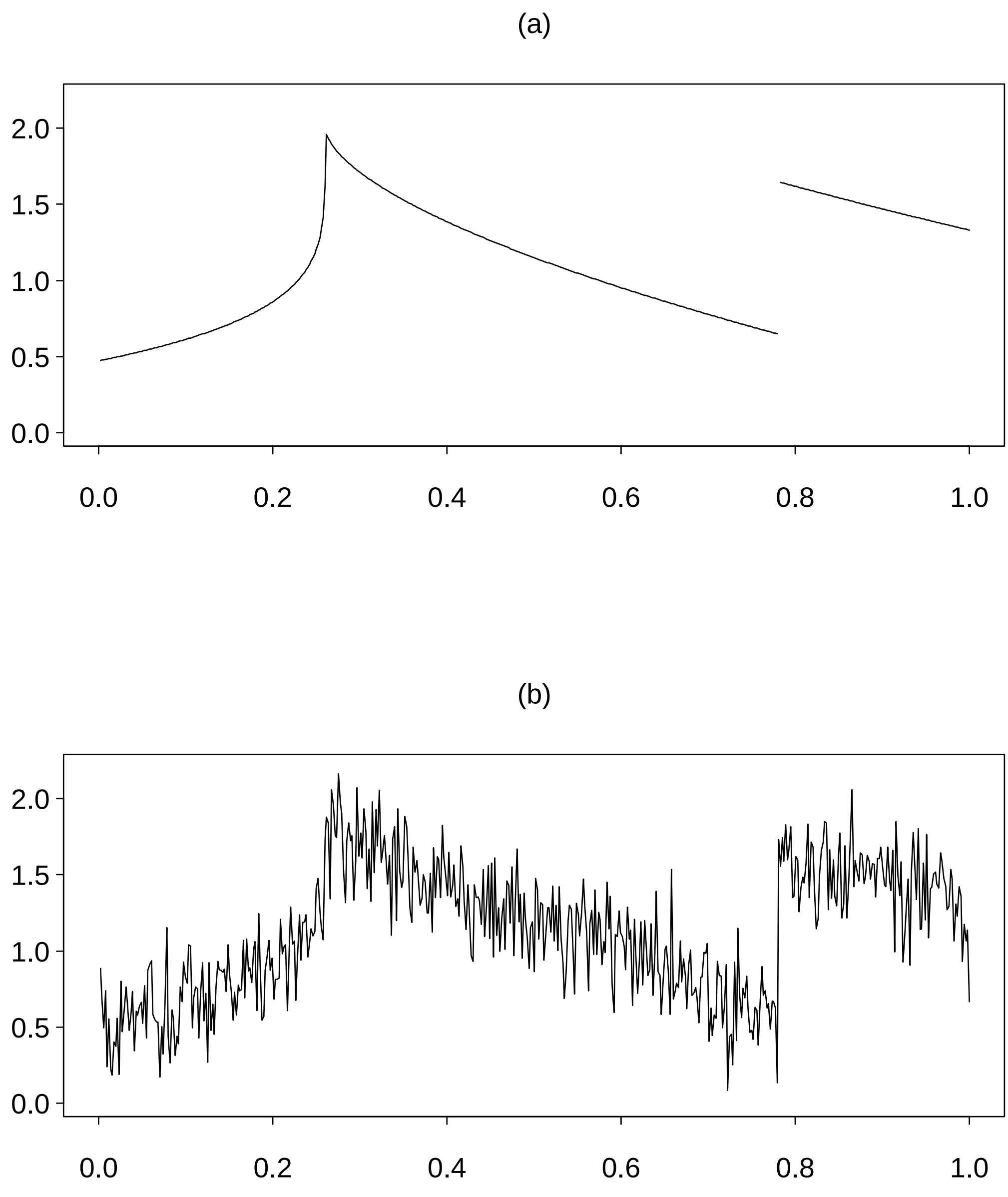

Fig. 2. The case of g_2. (a) displays g_2 and (b) noisy data with gaussian white noise.

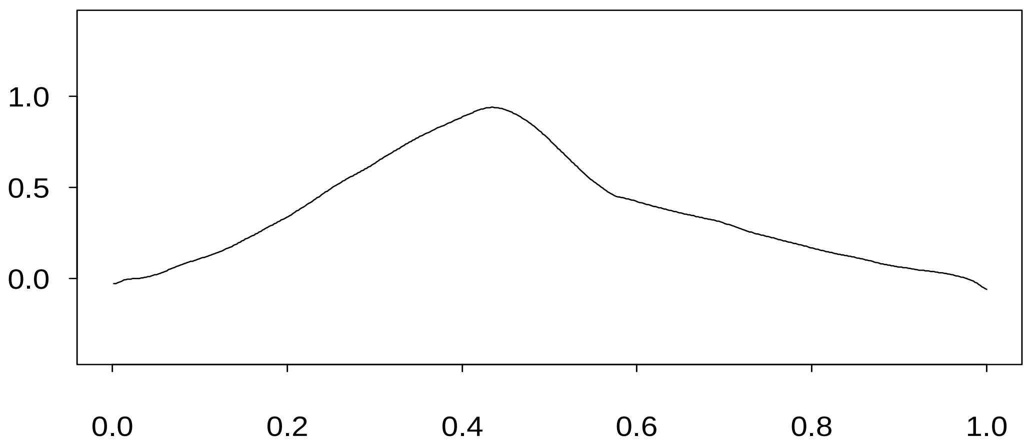

Fig. 3. Unadjusted kernel estimate of g_1.

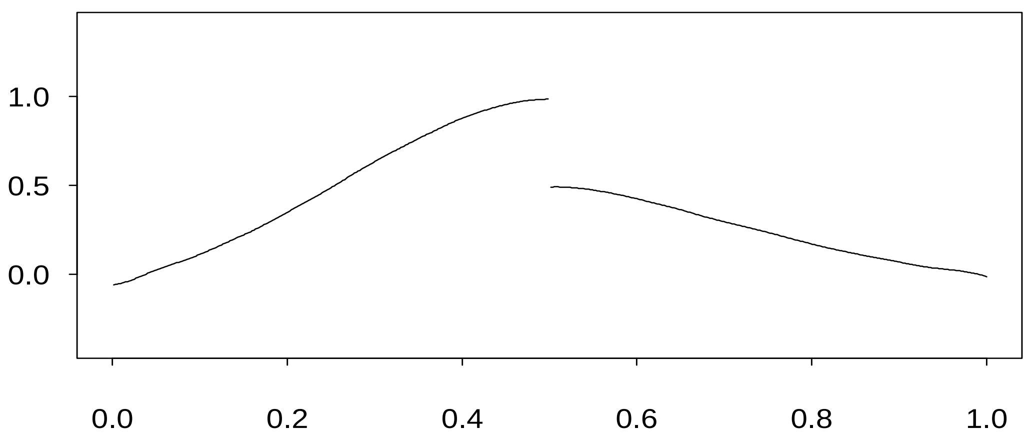

Fig. 4. Kernel estimate $\hat{g}_1^*$ of g_1 adjusted to change-point.

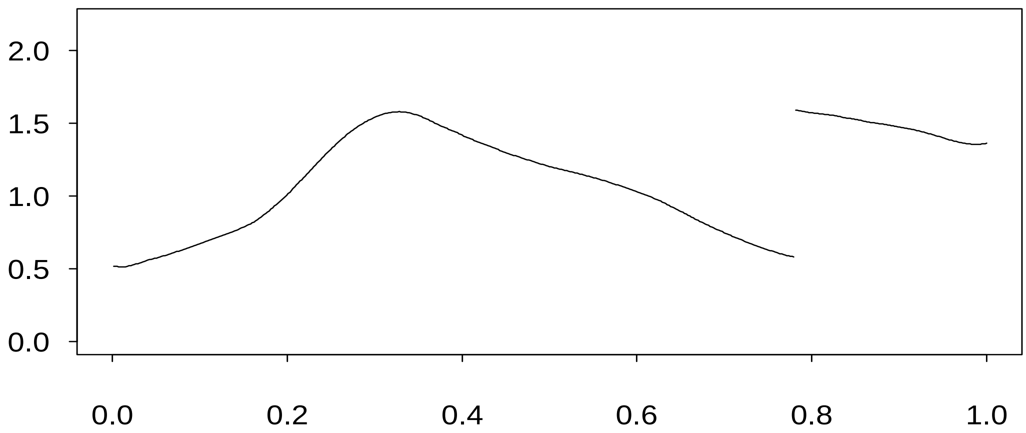

Fig. 5. Kernel estimate $\hat{g}_2^*$ of g_2 adjusted to change-point.