First Order Difference-Based Error Variance Estimator in Nonparametric Regression with a Single Outlier

Chun Gun Park[1,a]

[a]Department of Mathematics, Kyonggi University

Abstract

We consider some statistical properties of the first order difference-based error variance estimator in non-parametric regression models with a single outlier. So far under an outlier(s) such difference-based estimators has been rarely discussed. We propose the first order difference-based estimator using the leave-one-out method to detect a single outlier and simulate the outlier detection in a nonparametric regression model with the single outlier. Moreover, the outlier detection works well. The results are promising even in nonparametric regression models with many outliers using some difference based estimators.

Keywords: First order difference-based error variance estimator, leave-one-out, Lipschitz condition, Single outlier.

1. Introduction

There are two types of estimating an error variance in nonparametric regression models. One is that the residual sum of squares method first estimates a regression curve (Wahba, 1990; Hall and Carroll, 1989; Carter and Eagleson, 1992; Neumann, 1994). The other is that the estimation of the error variance which uses differences to remove trend in the regression curve has attracted a great deal of attention; see for example Rice (1984), Gasser *et al.* (1986), Hall *et al.* (1990), Dette *et al.* (1998), and Tong and Wang (2005), among others.

Robust nonparametric regression estimators are widely used when these models with a single outlier(s) are fitted (Boente and Fraiman, 1989). Meanwhile, such difference-based estimators which have been rarely discussed under the outlier problems have the advantages of no fitting a regression curve, fast convergence rate and computational convenience (Dette *et al.*, 1998). However, they do not perform well in the presence of outliers.

We investigate some statistical properties of the first order difference-based error variance in a nonparametric regression model with a single outlier. To detect the outlier, we propose the first order difference-based estimators using the leave-one-out method and their statistical properties are derived.

The rest of the paper is organized as follows. In Section 2, we introduce the nonparametric regression models with a single outlier and the first order difference-based error variance estimator. In Section 3, we propose the first order difference-based estimator using the leave-one-out method and derive some statistical properties on the outlier detection problem. In Section 4, we conduct a simulation study to observe the behavior of the first order difference-based estimator in some unknown regression curve with a single outlier under finite sample performance. Section 5 provides conclusion and further working. And finally, technical proofs are collected in the Appendix.

---

[1] Assistant Professor, Department of Mathematics, Kyonggi University, Gyeonggi-do 443-760, Korea.
  E-mail: cgpark@kgu.ac.kr

2. First Order Difference-Based Error Variance Estimator

2.1. Nonparametric regression models with a single outlier

In this paper we consider a particular aspect of some statistical properties for the first order difference-based error variance estimator in nonparametric regression models with a single outlier under a finite sample size.

The most basic form of our model without the outlier is

$$W_i = g(x_i) + \epsilon_i, \quad i = 1, \ldots, n, \tag{2.1}$$

where g is an unknown regression curve and the error ϵ_i's are *i.i.d.* random noise with distribution $F(\cdot)$.

The most basic form of our model with the single outlier is

$$Y_i = \begin{cases} g(x_i) + \epsilon_i, & i = 1, \ldots, t-1, \\ g(x_i) + \epsilon_i + \delta, & i = t, \\ g(x_i) + \epsilon_i, & i = t+1, \ldots, n, \end{cases} \tag{2.2}$$

where δ is a constant.

2.2. First order difference-based variance estimator

We investigate some statistical properties of the first order difference-based estimator under the following assumptions:

(A1) The regression curve $g : [0, 1] \to R$ is Lipschitz continuous, that is, there exits a constant c such that $|g(x) - g(y)| \leq c|x - y|$ for any $x, y \in [0, 1]$.

(A2) The fixed design points are $x_i = i/n, i = 1, \ldots, n$.

Rice (1984) proposed the first order difference-based estimator

$$\hat{\sigma}_R^2 = \frac{1}{2(n-1)} \sum_{i=2}^{n} (Y_i - Y_{i-1})^2.$$

All results are summarized under $|g(x_j) - g(x_i)| = c_{i(j-i)}(i/n)$ for $1 \leq i < j \leq n$. In details, see Park (2011).

Theorem 1. *Some statistical properties of the first order difference-based estimator are the following;*

(i) from (2.1), that is, without a single outlier

$$\centerdot \; E\left(\hat{\sigma}_R^2\right) = \sigma^2 + \frac{1}{2n^2(n-1)} \sum_{i=2}^{n} c_{i(1)}^2$$

$$\centerdot \; Var\left(\hat{\sigma}_R^2\right) = \frac{2\sigma^2}{n^2(n-1)^2} \left(\sum_{i=2}^{n} c_{i(1)}^2 - \sum_{i=2}^{n-1} c_{i(1)} c_{i+1(1)} \right) + 2\sigma^4 \frac{tr(D^2)}{tr(D)^2}$$

(ii) from (2.2), that is, with a single outlier for $1 < t < n - 1$

$$\cdot \left(\hat{\sigma}_O^2\right) = \sigma^2 + \frac{1}{2n^2(n-1)} \sum_{i=2}^{n} c_{i(1)}^2 + R_t(\delta)$$

$$\cdot \ Var\left(\hat{\sigma}_O^2\right) = \frac{2\sigma^2}{n^2(n-1)^2} \left(\sum_{i=2}^{n} d_{i(1)}^2 - \sum_{i=2}^{n-1} d_{i(1)}d_{i+1(1)} \right) + 2\sigma^4 \frac{tr(D^2)}{tr(D)^2},$$

where $R_t(\delta) = \delta/\{n(n-1)\} (c_{t(1)} - c_{t+1(1)}) + \delta^2/(n-1)$,

$$d_{i(1)} = \begin{cases} c_{i(1)} + n\delta, & i = t, \\ c_{i(1)} - n\delta, & i = t+1, \\ c_{i(1)}, & i = 2,\ldots,t-1,t+2,\ldots,n, \end{cases}$$

and $D = \tilde{D}^T \tilde{D}$

$$\tilde{D} = \begin{pmatrix} -1 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -1 & 0 & 1 \end{pmatrix} \in \mathrm{R}^{(n-1)\times n}.$$

If $c_{t(1)} = c_{t+1(1)}$, that is, $g(x)$ is a linear function, then $R_t(\delta) = \delta^2/(n-1)$. In the cases of a regression curve which is nonlinear, the expectation of the estimator consists of three parts (1) the error variance (2) the biased term and (3) the effect term of the outlier. In particular, the error variance and the biased term should be affected by the shape of the regression curve exit under the model (2.1), however, the effect of the outlier which exits under the model (2.2) depends on the sign and the magnitude of the outlier.

3. First Order Difference-Based Variance Estimator Using a the leave-one-out method

From (ii) of Theorem 1 the estimator has a biased term and the effect of the outlier. To estimate the error variance, in advance the outlier detection problem should be dealt with. To do this, we propose the first order difference-based estimator using the leave-one-out method.

Theorem 2. *From (2.2) and assuming that the single outlier is the tth observation, Y_t for $1 < t < n$, some statistical properties of the first order difference-based estimator using the leave-one-out method are as follows;*

(i) *If the tth observation which is an outlier is removed, then*

$$\cdot \ E\left({}_{(-t)}\hat{\sigma}_O^2\right) = \sigma^2 + \frac{1}{2n^2(n-2)} \sum_{i=2}^{n} c_{i(1)}^2 + \frac{c_{t(1)}c_{t+1(1)}}{n^2(n-2)},$$

$$\cdot \ Var\left({}_{(-t)}\hat{\sigma}_O^2\right) = \frac{2\sigma^2}{n^2(n-1)^2} \left(\sum_{i=2}^{n} d_{i(1)}^2 - \sum_{i=2}^{n-1} d_{i(1)}d_{i+1(1)} \right) + 2\sigma^4 \frac{tr(B^2)}{tr(B)^2},$$

where

$$d_{i(1)} = \begin{cases} c_{i(1)} + c_{i+1(1)}, & i = t, \\ c_{i(1)} + c_{i-1(1)}, & i = t+1, \\ c_{i(1)}, & i = 2,\ldots,t-1,t+2,\ldots,n. \end{cases}$$

(ii) If the kth $(1 < k < t)$ observation removed is not the outlier, then

- $E\left((_{-k)}\hat{\sigma}_O^2\right) = \sigma^2 + \dfrac{1}{2n^2(n-2)} \displaystyle\sum_{i=2}^{n} c_{i(1)}^2 + \dfrac{c_{k(1)}c_{k+1(1)}}{n^2(n-2)} + R_k(\delta),$

- $Var\left((_{-k)}\hat{\sigma}_O^2\right) = \dfrac{2\sigma^2}{n^2(n-1)^2}\left(\displaystyle\sum_{i=2}^{n} d_{i(1)}^2 - \sum_{i=2}^{n-1} d_{i(1)}d_{i+1(1)}\right) + 2\sigma^4\dfrac{tr(B^2)}{tr(B)^2},$

where

$$R_k(\delta) = \frac{\delta}{n(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-2},$$

$$d_{i(1)} = \begin{cases} c_{i(1)}, & i = 2,\ldots,k-1,k+2,\ldots,n, \\ c_{i(1)} + c_{i+1(1)}, & i = k, \\ c_{i-1(1)} - c_{i(1)}, & i = k+1, \\ c_{i(1)} + n\delta, & i = t, t+1 \end{cases}$$

and $B = \tilde{B}^T\tilde{B}$

$$\tilde{B} = \begin{pmatrix} -1 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -1 & 0 & 1 \end{pmatrix} \in \mathrm{R}^{(n-2)\times(n-1)}.$$

(iii) From (ii) of Theorem 1, let compare the estimators. Then

- $E\left(\hat{\sigma}_O^2\right) - E\left((_{-t)}\hat{\sigma}_O^2\right) \approx R_t(\delta)$

$$= \frac{\delta}{n(n-1)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-1},$$

- $E\left(\hat{\sigma}_O^2\right) - E\left((_{-k)}\hat{\sigma}_O^2\right) \approx -\dfrac{c_{k(1)}c_{k+1(1)}}{n^2(n-2)},$

- $E\left((_{-k)}\hat{\sigma}_O^2\right) - E\left((_{-t)}\hat{\sigma}_O^2\right) \approx R_k(\delta)$

$$= \frac{\delta}{n(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-2}.$$

From Theorem 2 (iii), we can figure out the effect of the outlier at the tth observation regardless of the variation of the estimators. However, as sample sizes increases, the effect of the outlier decreases. Therefore, to reduce the effect of sample sizes in the outlier detection problem, we modify the difference between the estimators for the leave-one-out method from (iii) of Theorem 2 as

$$E\left(\mathrm{DFF}_{(t)}\right) = E\left[(n-1)\left(\hat{\sigma}_O^2 - {}_{(-t)}\hat{\sigma}_O^2\right)\right] \approx \frac{\delta}{n}\left(c_{t(1)} - c_{t+1(1)}\right) + \delta^2$$
$$\approx \delta^2 \tag{3.1}$$

and

$$E\left(\mathrm{DFF}_{(k)}\right) = E\left[(n-1)\left(\hat{\sigma}_O^2 - {}_{(-k)}\hat{\sigma}_O^2\right)\right] \approx -\frac{c_{k(1)}c_{k+1(1)}}{n^2}. \tag{3.2}$$

From (3.1) and (3.2), we propose the single outlier detection procedure:

<Step 1> Compute $DFF_{(k)}$, $k = 2, \ldots, n - 1$.

<Step 2> Find the tth observation, Y_t, such that

$$t = \arg \max DFF_{(k)}, \ k = 2, \ldots, n - 1.$$

<Step 3> If $DFF_{(t)} > cv$, then

the tth observation is an outlier,

else

the tth observation is not an outlier.

where a cv(critical value) is obtained from the concepts of the outlier in normal distribution with mean zero and variance one.

4. Simulation Study

We perform a small simulation on the outlier detection that is developed from the first order difference-based estimator for leaving-one-out in nonparametric regression models with the single outlier. Our simulations are based on the followings, including the previous two assumptions:

(A3) $\epsilon_i \sim iid\ N(0, \sigma^2)$, $i = 1, \ldots, n$.

(A4) Three standard deviations from a normal distribution with mean zero are $\sigma = 0.01, 0.1, 1$.

(A5) The single outlier is defined as $\delta = 5, 10$ of mild and extreme levels (Barbato *et al.*, 2011).

(A6) The regression curve is $g(x) = 5 \sin(4\pi x)$.

(A7) Four sample sizes, $n = 15, 50, 100, 500$.

For each simulation setting, we generate observations and detect the single outlier. We repeat this process 100 times and the results are summarized in Figure 1 to Figure 6.

In this simulation study we investigate the behavior of DFFs with the outliers, $\delta = 5, 10$, under the variances varying. Here we do not discuss the critical values which are on the outlier detection.

Figure 1 and Figure 4 are the scatter plots for the regression curve and observations with the single outlier which are extreme and mild levels ($\delta = 10, 5$), respectively. Figure 2 and Figure 5 are the scatter plot of DFFs for 100 replications. Figure 3 and Figure 6 show the positions of a possible outlier such that $\arg \max DFF_{(k)}$, $k = 2, \ldots, n - 1$.

From Figure 3, since the magnitudes of the outlier is extreme regardless of the error variances given, the maximum of DFFs are always around δ^2. However, from Figure 6, the outlier detection does not work well under the error variance, $\sigma = 1$, which the magnitudes are not extreme.

5. Conclusion and Further Work

This article proposes a difference-based variance estimator of detecting a outlier in nonparametric regression models with a single outlier. The outlier detection problem has been rarely discussed in the nonparametric regression model using the difference. The proposed outlier detection method is simple. In addition, the results are promising even in nonparametric regression models with many outliers
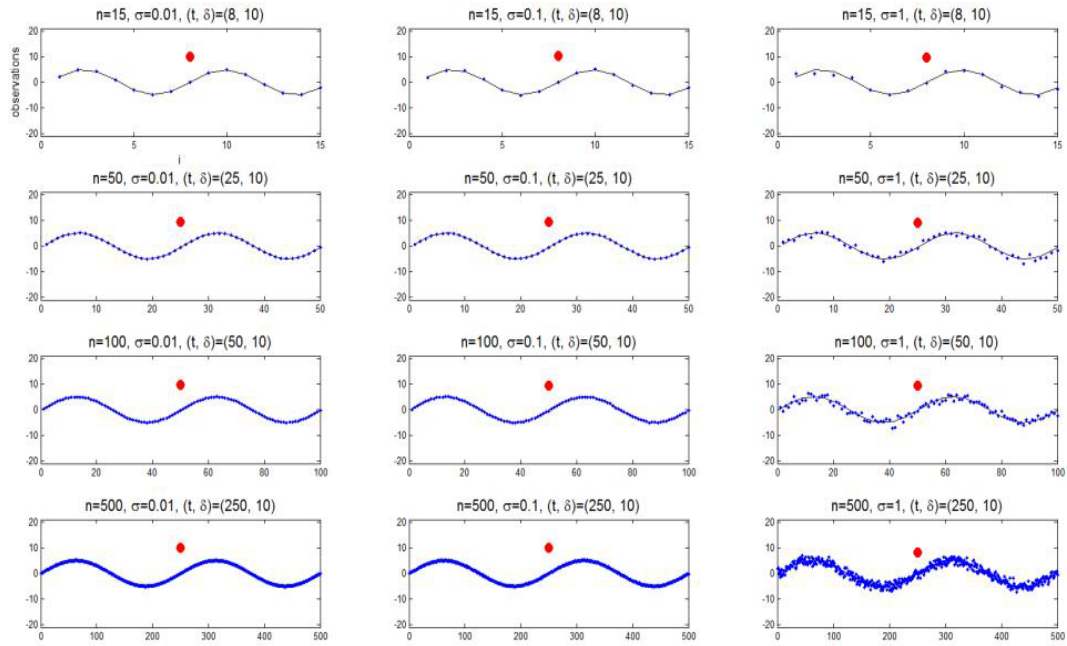
Figure 1: *The regression curve with errors and a single outlier (n = sample sizes, σ = standard deviation, t = position of the outlier and δ = magnitude of the outlier)*
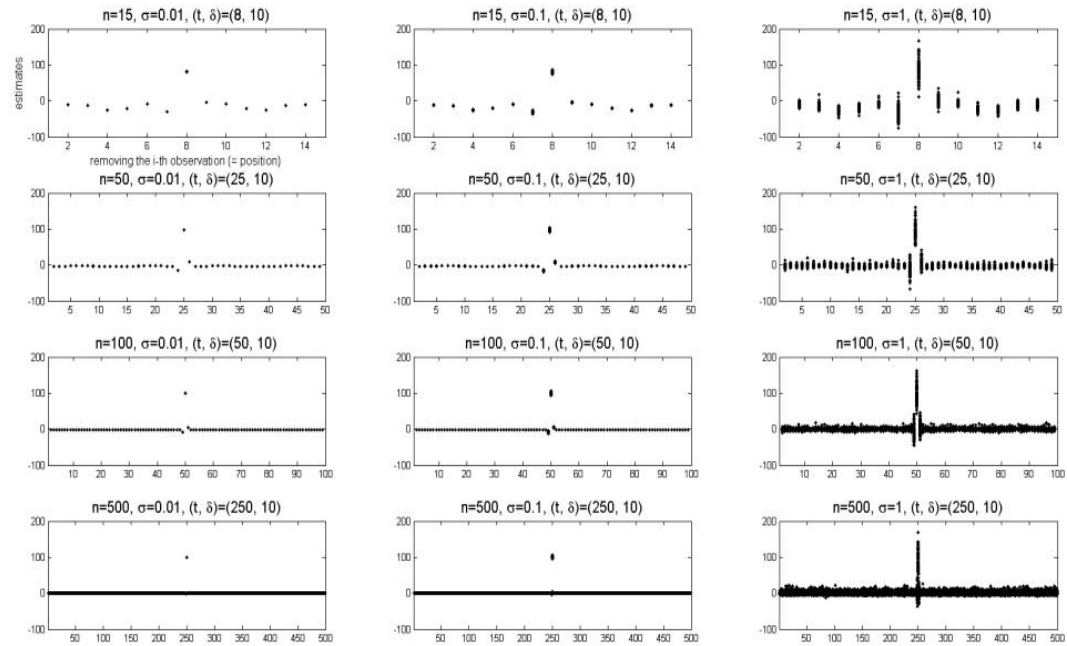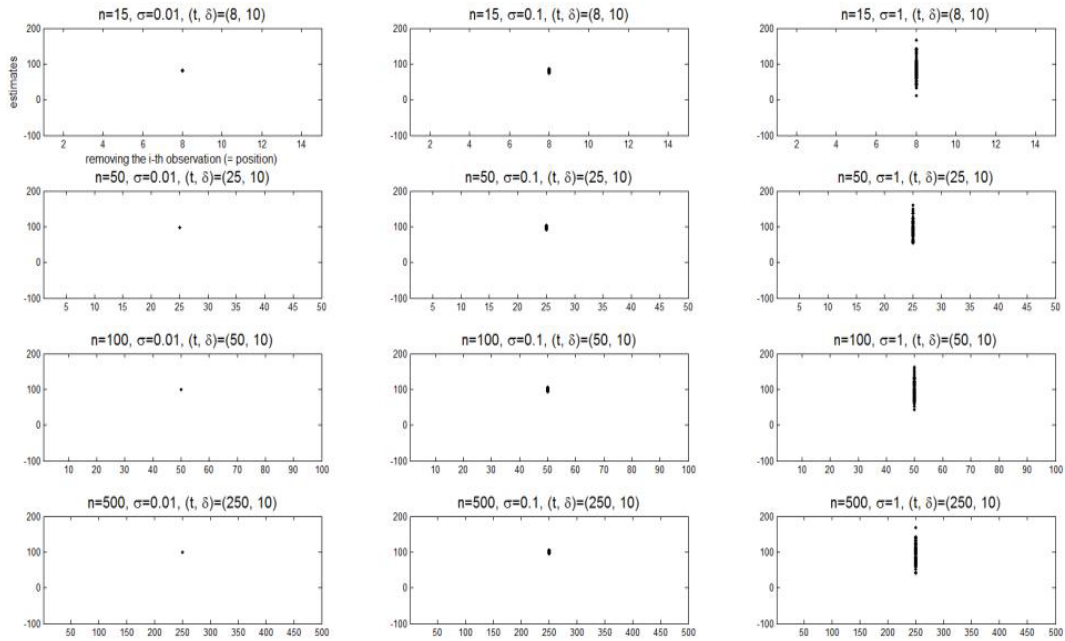

Figure 2: *For 100 replications, DFF's for δ = 10*

Figure 3: $DFF_{(t)}$ such that $t = \arg\max DFF_{(k)}, k = 2, \ldots, n-1$ for $\delta = 10$
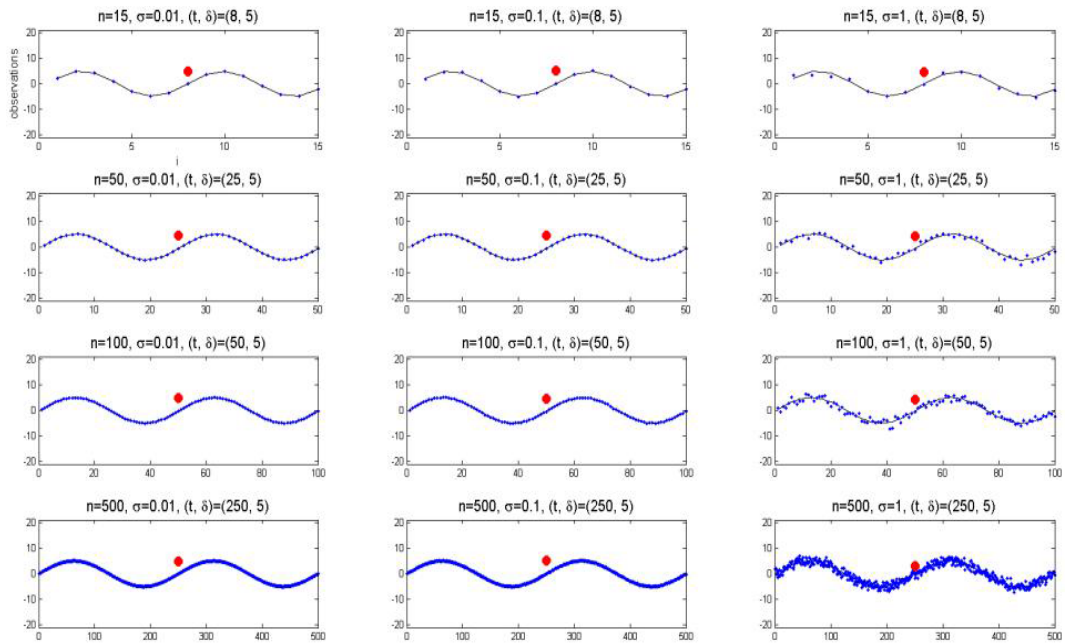

Figure 4: *The regression curve with errors and a single outlier (n = sample sizes, σ = standard deviation, t = position of the outlier and δ = magnitude of the outlier)*
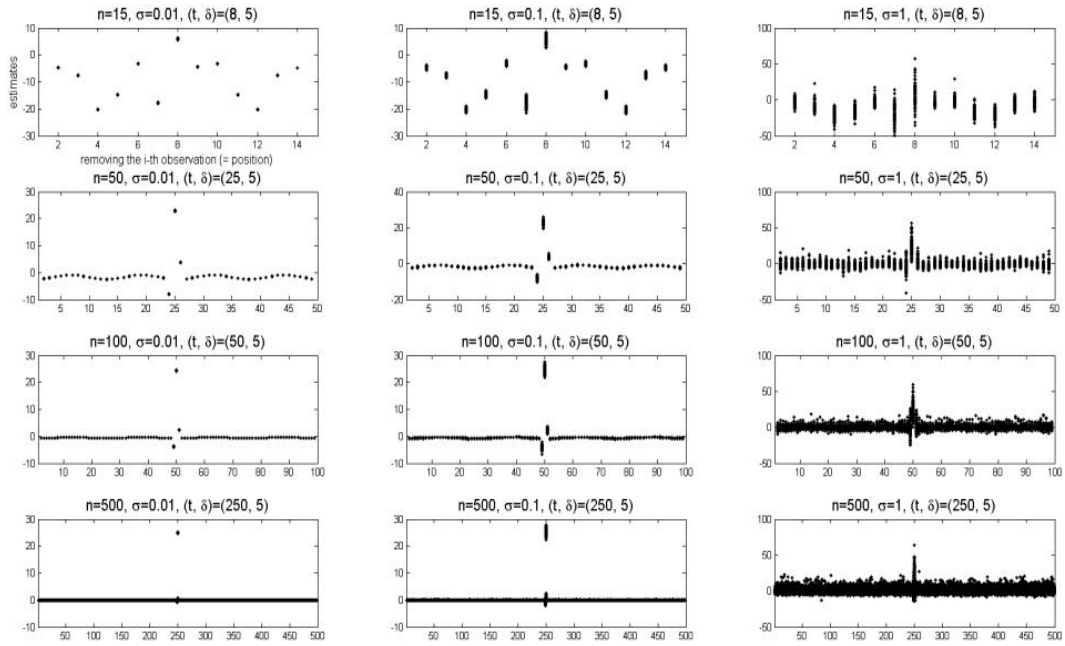
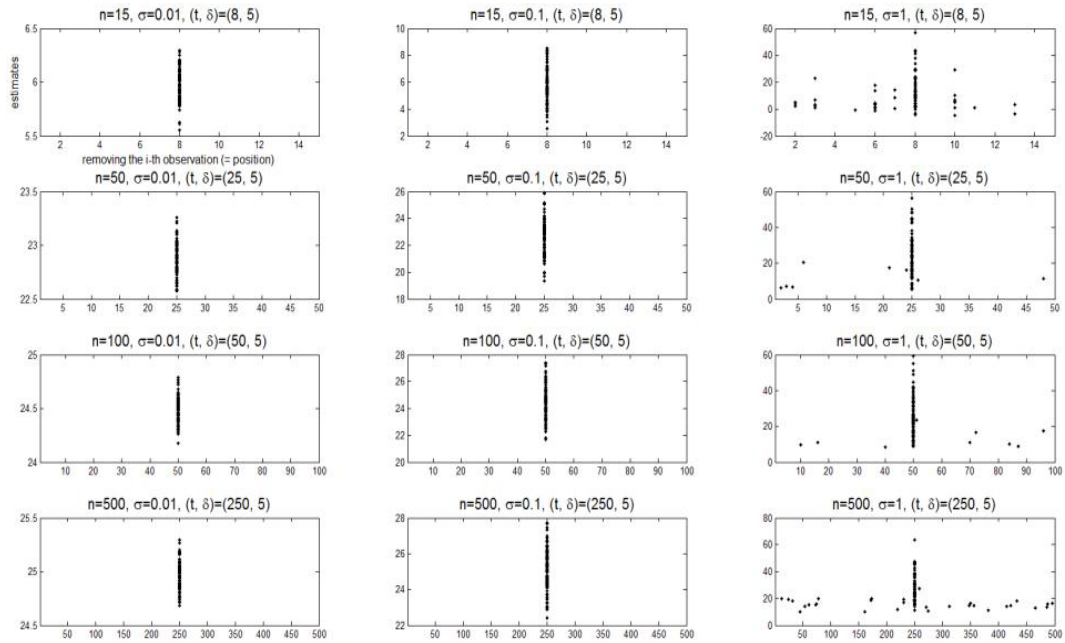Figure 5: *For 100 replication, DFF's for δ = 5*

Figure 6: *DFF$_{(t)}$ such that $t = \arg\max DFF_{(k)}$, $k = 2, \ldots, n-1$ for $\delta = 5$*

using some difference based estimators. To use the difference based estimators for outlier detection problems in the cases of many outliers, first the statistical properties of them and the estimators for the leave-one-out method should be derived. Second, some cut-off value or a test statistic should be devised

Appendix:

Proof of Theorem 1

The expectation of the first order difference-based estimator is the form as

(ii) If a single outlier is $1 < t < n$, then

$$
\hat{\sigma}_O^2 = \frac{1}{2(n-1)} \sum_{i=2}^{n} (Y_i - Y_{i-1})^2
$$

$$
= \frac{1}{2(n-1)} \left(\sum_{i=2}^{t-1} (Y_i - Y_{i-1})^2 + (Y_t - Y_{t-1})^2 + (Y_{t+1} - Y_t)^2 + \sum_{i=t+1}^{n} (Y_i - Y_{i-1})^2 \right)
$$

$$
= \frac{1}{2(n-1)} \left(\sum_{i=2}^{t-1} ((g_i - g_{i-1}) + (\epsilon_i - \epsilon_{i-1}))^2 + ((g_t - g_{t-1}) + (\epsilon_t - \epsilon_{t-1}) + \delta)^2 \right.
$$

$$
\left. + ((g_{t+1} - g_t) + (\epsilon_{t+1} - \epsilon_t) - \delta)^2 + \sum_{i=t+1}^{n} ((g_i - g_{i-1}) + (\epsilon_i - \epsilon_{i-1}))^2 \right)
$$

$$
= \frac{1}{2n^2(n-1)} \sum_{i=2}^{n} c_{i(1)}^2 + \frac{1}{n(n-1)} \sum_{i=2}^{n} c_{i(1)} (\epsilon_i - \epsilon_{i-1}) + \frac{1}{2(n-1)} \sum_{i=2}^{n} (\epsilon_i - \epsilon_{i-1})^2
$$

$$
+ \frac{\delta}{n(n-1)} (c_{t(1)} - c_{t+1(1)}) + \frac{\delta}{n-1} [(\epsilon_t - \epsilon_{t-1}) - (\epsilon_{t+1} - \epsilon_t)] + \frac{\delta^2}{n-1}
$$

$$
= \frac{1}{2n^2(n-1)} \sum_{i=2}^{n} d_{i(1)}^2 + \frac{1}{n(n-1)} \sum_{i=2}^{n} d_{i(1)} (\epsilon_i - \epsilon_{i-1}) + \frac{1}{2(n-1)} \sum_{i=2}^{n} (\epsilon_i - \epsilon_{i-1})^2
$$

$$
+ \frac{\delta}{n(n-1)} (d_{t(1)} - d_{t+1(1)} - 2n\delta) + \frac{\delta^2}{n-1}
$$

$$
\boldsymbol{.} \, E\left(\hat{\sigma}_O^2\right) = \sigma^2 + \frac{1}{2n^2(n-1)} \sum_{i=2}^{n} c_{i(1)}^2 + R_t(\delta),
$$

where

$$
R_t(\delta) = \frac{\delta}{n(n-1)} (c_{t(1)} - c_{t+1(1)}) + \frac{\delta^2}{n-1}
$$

and

$$
d_{i(1)} = \begin{cases} c_{i(1)} + n\delta, & i = t, \\ c_{i(1)} - n\delta, & i = t + 1, \\ c_{i(1)}, & i = 2, \ldots, t-1, t+2, \ldots, n. \end{cases}
$$

The proofs of (i) and variance of (ii) are omitted as it is straightforward in details, see Park (2011) in the Appendix.

Proof of Theorem 2

Assume that a single outlier is the tth observation, Y_t for $1 < t < n - 1$. The expectation of the first order difference-based estimator for the leave-one-out method are the form as

(i) If the tth observation removed is the outlier, then

$$
{}_{(-t)}\hat{\sigma}_O^2 = \frac{1}{2(n-2)} \left(\sum_{i=2}^{t-1} (Y_i - Y_{i-1})^2 + (Y_{t+1} - Y_{t-1})^2 + \sum_{i=t+2}^{n} (Y_i - Y_{i-1})^2 \right)
$$

$$
= \frac{1}{2n^2(n-2)} \left(\sum_{i=2}^{n} c_{i(1)}^2 + 2c_{t(1)}c_{t+1(1)} \right) + \frac{1}{n(n-2)} \sum_{i=2}^{n} c_{i(1)} (\epsilon_i - \epsilon_{i-1})
$$

$$
+ c_{t(1)}(\epsilon_{t+1} - \epsilon_t) + c_{t+1(1)} (\epsilon_t - \epsilon_{t-1})
$$

$$
+ \frac{1}{2(n-2)} \left(\sum_{i=2}^{n} (\epsilon_i - \epsilon_{i-1})^2 + 2 (\epsilon_t - \epsilon_{t-1}) (\epsilon_{t+1} - \epsilon_t) \right),
$$

where

$$
g_{t+1} - g_{t-1} = (g_{t+1} - g_t) + (g_t - g_{t-1}) \Leftrightarrow c_{t+1(2)} = c_{t+1(1)} + c_{t(1)}.
$$

$$
\bullet \, E\left({}_{(-t)}\hat{\sigma}_O^2 \right) = \sigma^2 + \frac{1}{2n^2(n-2)} \left(\sum_{i=2}^{t-1} c_{i(1)}^2 + c_{t+1(2)}^2 + \sum_{i=t+2}^{n} c_{i(1)}^2 \right)
$$

$$
= \sigma^2 + \frac{1}{2n^2(n-2)} \left(\sum_{i=2}^{n} c_{i(1)}^2 + 2c_{t(1)}c_{t+1(1)} \right)
$$

$$
= \sigma^2 + \frac{1}{2n^2(n-2)} \sum_{i=2}^{n} c_{i(1)}^2 + \frac{c_{t(1)}c_{t+1(1)}}{n^2(n-2)}
$$

$$
\bullet \, \mathrm{Var}\left(\hat{\sigma}_O^2 \right) = \frac{2\sigma^2}{n^2(n-1)^2} \left(\sum_{i=2}^{n} d_{i(1)}^2 - \sum_{i=2}^{n-1} d_{i(1)}d_{i+1(1)} \right) + 2\sigma^4 \frac{\mathrm{tr}\left(B^2 \right)}{\mathrm{tr}\,(B)^2},
$$

where

$$
d_{i(1)} = \begin{cases} c_{i(1)} + c_{i+1(1)}, & i = t, \\ c_{i(1)} + c_{i-1(1)}, & i = t+1, \\ c_{i(1)}, & i = 2, \ldots, t-1, t+2, \ldots, n. \end{cases}
$$

(ii) Let the tth observation be the outlier, If the kth $(k < t)$ observation removed is not the outlier, then

$$
{}_{(-k)}\hat{\sigma}_O^2 = \frac{1}{2(n-2)} \left(\sum_{i=2}^{k-1} (Y_i - Y_{i-1})^2 + (Y_{k+1} - Y_{k-1})^2 + \sum_{i=k+2}^{t-1} (Y_i - Y_{i-1})^2 \right.
$$

$$
\left. + (Y_t - Y_{t-1})^2 + (Y_{t+1} - Y_t)^2 + \sum_{i=t+2}^{n} (Y_i - Y_{i-1})^2 \right)
$$

$$
= \frac{1}{2n^2(n-2)} \sum_{i=2}^{n} d_{i(1)}^2 + \frac{1}{n(n-2)} \sum_{i=2}^{n} d_{i(1)} (\epsilon_i - \epsilon_{i-1}) + \frac{1}{2(n-2)} \sum_{i=2}^{n} (\epsilon_i - \epsilon_{i-1})^2
$$

$$
+ \frac{(\epsilon_k - \epsilon_{k-1})(\epsilon_{k+1} - \epsilon_k)}{(n-2)} + \frac{c_{k(1)}c_{k+1(1)}}{n^2(n-2)} + R_k(\delta)
$$

$$\centerdot E\left((-k)\hat{\sigma}_O^2\right) = \sigma^2 + \frac{1}{2n^2(n-2)}\left(\sum_{i=2}^{t-1} c_{i(1)}^2 + c_{k+1(2)}^2 + \sum_{i=t+2}^{n} c_{i(1)}^2\right) + \frac{\delta}{n(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-2}$$

$$= \sigma^2 + \frac{1}{2n^2(n-2)}\left(\sum_{i=2}^{n} c_{i(1)}^2 + 2c_{k(1)}c_{k+1(1)}\right) + \frac{\delta}{n(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-2}$$

$$= \sigma^2 + \frac{1}{2n^2(n-2)}\sum_{i=2}^{n} c_{i(1)}^2 + \frac{c_{k(1)}c_{k+1(1)}}{n^2(n-2)} + \frac{\delta}{n(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-2}$$

$$= \sigma^2 + \frac{1}{2n^2(n-2)}\sum_{i=2}^{n} c_{i(1)}^2 + R_k(\delta) + \frac{c_{k(1)}c_{k+1(1)}}{n^2}$$

$$\centerdot \mathrm{Var}\left((-k)\hat{\sigma}_O^2\right) = \frac{2\sigma^2}{n^2(n-1)^2}\left(\sum_{i=2}^{n} d_{i(1)}^2 - \sum_{i=2}^{n-1} d_{i(1)}d_{i+1(1)}\right) + 2\sigma^4\frac{\mathrm{tr}(B^2)}{\mathrm{tr}(B)^2},$$

where

$$R_k(\delta) = \frac{\delta}{n(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-2}$$

and

$$d_{i(1)} = \begin{cases} c_{i(1)}, & i = 2,\ldots,k-1,k+2,\ldots,n, \\ c_{i(1)} + c_{i+1(1)}, & i = k, \\ c_{i-1(1)} - c_{i(1)}, & i = k+1, \\ c_{i(1)} + n\delta, & i = t, t+1. \end{cases}$$

The proofs of variance of (i) and (ii) are omitted as it is straightforward in details, see Park (2011) in the Appendix.

(iii) From (i), (ii) and Theorem 1, let compare the estimators. Then

$$\centerdot E\left(\hat{\sigma}_O^2\right) - E\left((-t)\hat{\sigma}_O^2\right) = R_t(\delta) - \frac{1}{2n^2(n-1)(n-2)}\sum_{i=2}^{n} c_{i(1)}^2 - \frac{c_{t(1)}c_{t+1(1)}}{n^2(n-2)}$$

$$\approx R_t(\delta) - \frac{c_{t(1)}c_{t+1(1)}}{n^2(n-2)}$$

$$\approx \frac{\delta}{n(n-1)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-1}$$

$$\centerdot E\left(\hat{\sigma}_O^2\right) - E\left((-k)\hat{\sigma}_O^2\right) = R_t(\delta) - R_k(\delta) - \frac{1}{2n^2(n-1)(n-2)}\sum_{i=2}^{n} c_{i(1)}^2 - \frac{c_{k(1)}c_{k+1(1)}}{n^2(n-2)}$$

$$\approx R_t(\delta) - R_k(\delta) - \frac{c_{k(1)}c_{k+1(1)}}{n^2(n-2)}$$

$$\approx -\frac{\delta^2}{(n-1)(n-2)} - \frac{\delta}{n(n-1)(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right) - \frac{c_{k(1)}c_{k+1(1)}}{n^2(n-2)}$$

$$\approx -\frac{\delta^2}{(n-1)(n-2)} - \frac{\delta}{n(n-1)(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right)$$

$$\bullet E\left({}_{(-k)}\hat{\sigma}_O^2\right) - E\left({}_{(-t)}\hat{\sigma}_O^2\right) = R_k(\delta) + \frac{c_{k(1)}c_{k+1(1)}}{n^2(n-2)} - \frac{c_{t(1)}c_{t+1(1)}}{n^2(n-2)}$$

$$\approx R_k(\delta)$$

$$= \frac{\delta}{n(n-2)}\left(c_{t(1)} - c_{t+1(1)}\right) + \frac{\delta^2}{n-2}.$$

References

Barbato, G., Barini, E. M., Genta, G., and Levi, R. (2011). Features and performance of some outlier detection methods, em Journal of Applied Statistics, **38**, 2133-2149.

Boente, G. and Fraiman, R. (1989). Robust nonparametric regression estimation for dependent observations, *The Annals of Statistics*, **17**, 1242–1256.

Carter, C. K. and Eagleson, G. K. (1992). A comparison of variance estimators in nonparametric regression, *Journal of the Royal Statistical Society: Series B*, **54**, 773–780.

Dette, H., Munk, A. and Wagner, T. (1998). Estimating the variance in nonparametric regression? What is a reasonable choice?, *Journal of the Royal Statistical Society: Series B*, **60**, 751–764.

Gasser, T., Sroka, L. and Jennen-Steinmetz, C. (1986). Residual variance and residual pattern in nonlinear regression, *Biometrika*, **73**, 625–633.

Hall, P. and Carroll, R. J. (1989). Variance function estimation in regression: the effect of estimating the mean, *Journal of the Royal Statistical Society: Series B*, **51**, 3–14.

Hall, P., Kay, J. W. and Titterinton, D. M. (1990). Asymptotically optimal difference-based estimator of variance in nonparametric regression, *Biometrika*, **77**, 521–528.

Neumann, M. H. (1994). Fully data-driven nonparametric variance estimators, *Statistics*, **25**, 189–212.

Park, C. G. (2011). Estimation of error variance in nonparametric regression under a finite sample using ridge regression, *Journal of Korean Data & Information Science Society*, **22**, 1223–1232.

Rice, J. A. (1984). Bandwidth choice for nonparametric regression. *Annals of statistics*, **12**, 1215–1220.

Tong, T. and Wang, Y. (2005). Estimating residual variance in nonparametric regression using least squares, *Biometrika*, **92**, 821–830.

Wahba, G. (1990). Spline models for observational data, *CBMS-NSF Regional Conference Series in Applied Mathematics*, **59**, SIAM, Philadelphia.