DISSERTATION

IMPROVED INFERENCE IN HETEROSKEDASTIC REGRESSION MODELS WITH

MONOTONE VARIANCE FUNCTION ESTIMATION

Submitted by

Soo Young Kim

Department of Statistics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2018

Doctoral Committee:

    Advisor: Haonan Wang
    Co-Advisor: Mary C. Meyer

    Bailey K. Fosdick
    Jean D. Opsomer
    J. Rockey Luo

ABSTRACT

IMPROVED INFERENCE IN HETEROSKEDASTIC REGRESSION MODELS WITH

MONOTONE VARIANCE FUNCTION ESTIMATION

The problems associated with heteroskedasticity often lead to incorrect inferences in a regression model, especially when the form of the heteroskedasticity is obscure. In this dissertation, I present methods to estimate a variance function in a heteroskedastic regression model where the variance function is assumed to be smooth and monotone in a predictor variable. Maximum likelihood estimation of the variance function is derived under normal or double-exponential error distribution assumptions based on regression splines and the cone projection algorithm. A penalized spline estimator is also introduced, and the estimator performs well when there exists a spiking problem at a boundary of domain. The convergence rates of the estimated variance functions are derived, and simulations show that it tends to be closer to the true variance function in a variety of scenarios compared to the existing method. The estimated variance functions from the proposed methods provide improved inference about the mean function, in terms of a coverage probability and an average length for an interval estimate. The utility of the method is illustrated through the analysis of real datasets such as LIDAR data, abalone data, California air pollution data, and U.S. temperature data. The methodology is implemented in the `R` package `cgam`. In addition to the variance function estimation method, the hypothesis test procedure of a smooth and monotone variance function is discussed. The likelihood ratio test is introduced under normal or double-exponential error distribution assumptions. Comparisons of the proposed test with existing tests are conducted through simulations.

ACKNOWLEDGEMENTS

DEDICATION

To those who have kindly encouraged me and believed in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Review of Regression Splines

Regression splines have been widely used for approximating functions, and B-splines are a standard choice for regression basis function because of its flexibility and smoothness. Basic properties of B-splines are illustrated in de Boor (2001) and Schumaker (2007). Given the sample size n and K_n distinct knot points $t_1, \ldots, t_{K_n}$ where $0 = t_1 < \cdots < t_{K_n} = 1$, let $B_{k,p}(x)$ be the kth B-spline basis function of degree p in a variable x for any $0 \le x \le 1$. The B-spline basis function of degree p is a piecewise polynomial function of degree p on an interval $[0, 1]$. The number of interior knots K_n increases as n increases, and the number of basis functions is $L_n = K_n + p - 1$. Based on the recurrence relation introduced in de Boor (2001), a B-spline basis function is defined as

$$B_{k,p}(x) = \frac{x - t_k}{t_{k+p} - t_k} B_{k,p-1}(x) + \frac{t_{k+p+1} - x}{t_{k+p+1} - t_{k+1}} B_{k+1,p-1}(x),$$

where

$$B_{k,0}(x) = \begin{cases} 1 & \text{if } t_k \le x < t_k + 1 \\ 0 & \text{otherwise.} \end{cases}$$

The first derivative of a B-spline basis function of degree p is defined as

$$B'_{k,p}(x) = p \left(\frac{-B_{k+1,p-1}(x)}{t_{k+p+1} - t_{k+1}} + \frac{B_{k,p-1}(x)}{t_{k+p} - t_k} \right),$$

which is a function of B-spline basis functions of degree $p - 1$. We can define a B-spline function of degree p on the interval $[0, 1]$ given by

$$S(x) = \sum_{k=1}^{L_n} \alpha_k B_{k,p}(x),$$

where α_k's are coefficients. In this dissertation, we choose to use quadratic splines that correspond to $p = 2$ because the quadratic B-spline basis functions are widely used to approximate monotone functions. The kth quadratic B-spline basis, $B_{k,2}$, is consists of three nontrivial quadratic pieces and vanishes outside the interval $[t_{k-2}, t_{k+1})$. The connection between the shape of a spline function S and the behavior of its coefficients are illustrated in the Section 4.9 of Schumaker (2007). The function S is said to be monotone increasing (decreasing) if the first derivatives are positive (negative) at all knots. It is because the first derivative of the spline function $S(x)$ can be written as

$$S'(x) = \frac{d}{dx}\left(\sum_{k=1}^{L_n} \alpha_k B_{k,2}(x)\right) = 2\sum_{k=1}^{L_n-1} \frac{\alpha_{k+1} - \alpha_k}{t_{k+3} - t_{k+1}} B_{k,1}(x), \tag{1.1}$$

if $t_{k+2} - t_k > 0$. From (1.1), non-increasing (non-decreasing) quadratic B-spline function can be obtained if $\alpha_{k+1} \leq \alpha_k$ ($\alpha_{k+1} \geq \alpha_k$) because $t_{k+3} - t_{k+1}$ is non-negative for every $k \in [1, K_n - 2]$. It is also possible to construct monotone function using higher order splines, but it requires more complex constraints which is unnecessary. Figure 1.1 shows quadratic B-spline basis functions with five knots and $n = 100$ and their first derivatives.
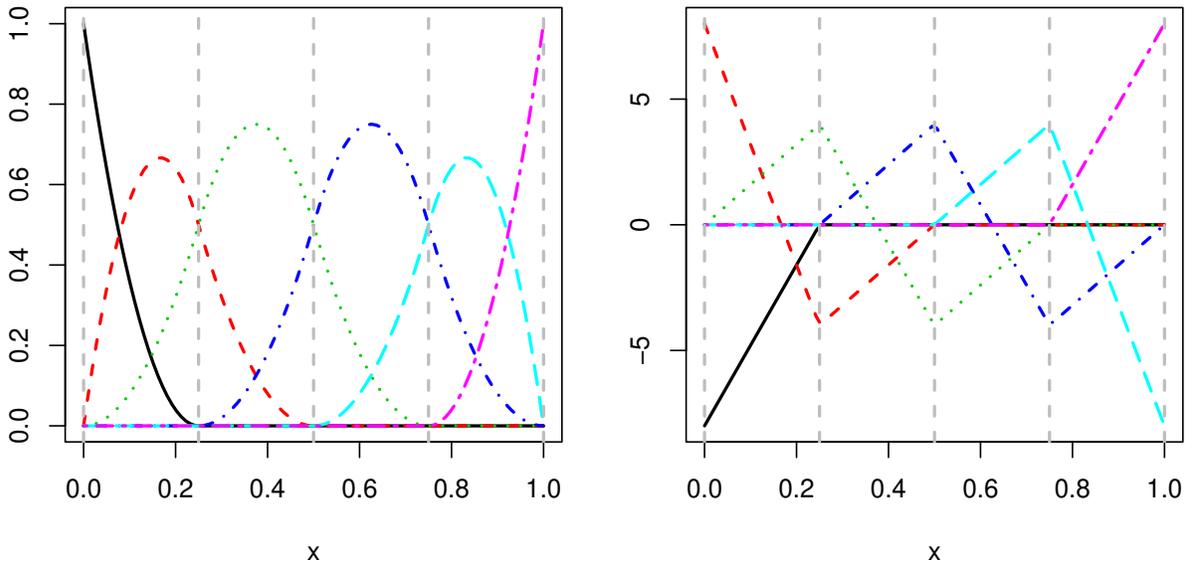
Figure 1.1: Quadratic B-spline basis functions and their first derivatives. Left: Quadratic B-spline basis functions with five knots and $n = 100$. Right: The first derivatives of quadratic B-spline basis functions in the left panel. Gray dashed vertical lines mark knots.

1.2   Review of Cone Projection

The cone projection algorithm is a special case of the quadratic programming algorithm which is used in the process of optimizing a quadratic objective function subject to one or more constraints. The quadratic programming algorithm solves the problem of finding $\hat{\boldsymbol{\theta}}$ that minimizes the quadratic objective function

$$\boldsymbol{\theta}^T \boldsymbol{Q} \boldsymbol{\theta} - 2\boldsymbol{c}^T \boldsymbol{\theta} \tag{1.2}$$

subject to $\boldsymbol{A_0}\boldsymbol{\theta} \geq \boldsymbol{d}$ where $\boldsymbol{Q}$ is an $n \times n$ positive definite matrix, $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{A_0}$ is an $m \times n$ matrix, and $\boldsymbol{d} \in \mathbb{R}^m$.

Let $\boldsymbol{Q} = \boldsymbol{U}^T \boldsymbol{U}$ by the Cholesky decomposition. For $\boldsymbol{\theta_0}$ such that $\boldsymbol{A_0}\boldsymbol{\theta_0} = \boldsymbol{d}$, define a vector $\boldsymbol{z} = (\boldsymbol{U}^{-1})^T(\boldsymbol{c} - \boldsymbol{Q}\boldsymbol{\theta_0})$ and an irreducible matrix $\boldsymbol{A} = \boldsymbol{A_0}\boldsymbol{U}^{-1}$. The matrix $\boldsymbol{A}$ is called irreducible if any row of $\boldsymbol{A}$ is not a positive linear combination of other rows of $\boldsymbol{A}$, and the origin is not a positive linear combination of rows of $\boldsymbol{A}$. The quadratic programming algorithm is equivalent to solve a problem of finding $\hat{\boldsymbol{\phi}}$ that minimizes the objective function

$$||\boldsymbol{z} - \boldsymbol{\phi}||^2$$

subject to $\boldsymbol{A}\boldsymbol{\phi} \geq \boldsymbol{0}$. A solution $\hat{\boldsymbol{\phi}}$ is the projection of $\boldsymbol{z}$ onto the set

$$\mathcal{C} = \{\boldsymbol{\phi} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{\phi} \geq \boldsymbol{0}\}. \tag{1.3}$$

It can be seen that the set $\mathcal{C}$ is a convex cone. The set $\mathcal{C}$ is a cone because $c_1\boldsymbol{\phi} \in \mathcal{C}$ for any $\boldsymbol{\phi} \in \mathcal{C}$ and a non-negative real number c_1. Further, the set $\mathcal{C}$ is a convex because for any $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2 \in \mathcal{C}$ and $c_1, c_2 \geq 0$, we have $c_1\boldsymbol{\phi}_1 + c_2\boldsymbol{\phi}_2 \in \mathcal{C}$. Then, we can find a unique solution $\hat{\boldsymbol{\phi}}$ that optimizes a convex function over a convex set.

Figure 1.2 shows the two dimensional example of the projection of $\boldsymbol{z}$ onto the cone $\mathcal{C}$ and the linear space $\mathcal{L}$. The linear space $\mathcal{L}$ is a plane containing the $EFGH$. The cone $\mathcal{C}$ is included in $\mathcal{L}$,

and contains AOB. The point D is the projection of $\boldsymbol{z}$ onto $\mathcal{L}$, and the point C is the projection of $\boldsymbol{z}$ onto $\mathcal{C}$. The point C can be represent as the projection of D onto $\mathcal{C}$.
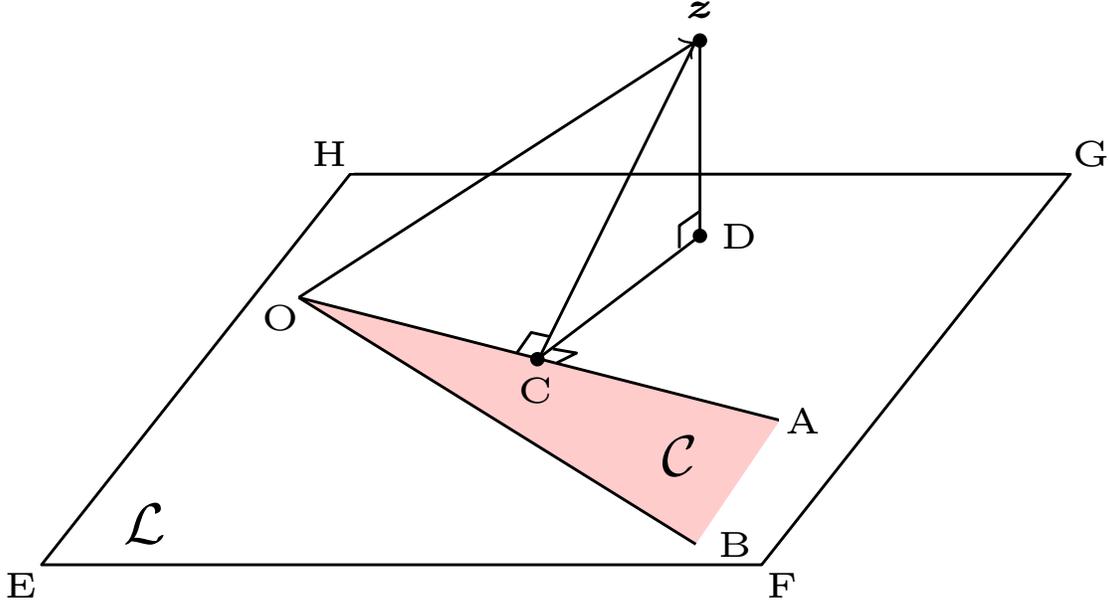
Figure 1.2: The projection of $\boldsymbol{z}$ onto the linear space $\mathcal{L}$ and the cone $\mathcal{C} \subset \mathcal{L}$.

Let the linear space $\boldsymbol{V} \in \mathcal{C}$ be the null space of $\boldsymbol{A}$; i.e, $\boldsymbol{V}$ is orthogonal to the space spanned by the rows of $\boldsymbol{A}$. Let the edges of the cone be $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_M$ for $M \geq m$. The edges are orthogonal to $\boldsymbol{V}$, and they can be derived from the method introduced by Meyer (1999).

The set $\mathcal{C}$ consists of elements that can be written as the sum of a vector in $\boldsymbol{V}$ and a linear combination of the edges of $\mathcal{C}$ with non-negative coefficients. Therefore, the cone defined in (1.3) can be rewritten as

$$\mathcal{C} = \left\{ \boldsymbol{\phi} \in \mathbb{R}^n : \boldsymbol{\phi} = \boldsymbol{v} + \sum_{j=1}^{M} b_j \boldsymbol{\delta}_j, \text{ where } \boldsymbol{v} \in \boldsymbol{V}, \text{ and } b_1, \ldots, b_M \geq 0 \right\}.$$

We consider the projection of $\boldsymbol{z}$ onto a face of the cone defined as

$$\mathcal{F}_J = \left\{ \boldsymbol{\phi} \in \mathbb{R}^n : \boldsymbol{\phi} = \boldsymbol{v} + \sum_{j \in J} b_j \boldsymbol{\delta}_j, \ \boldsymbol{v} \in \boldsymbol{V}, b_j > 0 \text{ for } j \in J \right\}$$

where J is the subset of $\{1, \ldots, M\}$.

The cone projection algorithm to find $\hat{\phi} \in \mathcal{C}$ introduced in Meyer (2013) can be summarized in three steps. Let J_0 be the initial guess of indices of edges: i.e. subset of $\{1, \ldots, M\}$. At the kth iteration,

1 Project $\boldsymbol{z}$ onto the linear space spanned by $\{\boldsymbol{\delta}^j, j \in J_k\}$, to get $\boldsymbol{\phi}^k = \sum_{j \in J_k} b_j^{(k)} \boldsymbol{\delta}_j$.

2 Check to see if $\boldsymbol{\phi}^k$ satisfies the constraints, i.e. if all $b_j^{(k)}$ are non-negative:

    · If yes, go to step 3.

    · If no, choose j for which $b_j^{(k)}$ is smallest, and remove it from the set; go to step 1.

3 Compute $< \boldsymbol{z} - \boldsymbol{\phi}^k, \boldsymbol{\delta}^j >$ for each $j \notin J_k$. If these are all nonpositive, then stop. If not, choose j for which this inner product is largest, add it to the set, and go to step 1.

The proof of the convergence of the algorithm is provided in Meyer (2013). The computation of the algorithm is efficiently coded in the `qprog` function available in an R package `coneproj` (Liao and Meyer, 2014). The `qprog` function yields $\hat{\phi}$ that optimizes the objective function (1.2), given $\boldsymbol{Q}$, $\boldsymbol{c}$, $\boldsymbol{A}_0$, and $\boldsymbol{d}$.

1.3 Overview

The convergence rates for monotone regression splines has been established in Chapter 2. It is shown that the constrained spline estimator attains the optimal rate of convergence when the true function satisfies the monotonicity constraints. In Chapter 3, I propose the methods for estimating a smooth monotone variance function in a heteroskedastic regression model. The method introduced in Chapter 3 is based on the maximum likelihood principle and regression splines, and its computation is carried out through the convex programming. The convergence rates of the monotone spline approximant are derived, and the same optimal rate is preserved as an unconstrained spline approximant. Further, the application of the proposed method is illustrated through the analysis of several real datasets such as LIDAR data, abalone data, California air pollution data, and U.S.

temperature data. In Chapter 4, I introduce an alternative variance function estimation method by including a penalty term to handle a problem at a boundary of domain. In Chapter 5, our focus shifts from the estimators of variance functions to hypothesis testing. I present a hypothesis test of a constant variance versus a smooth monotone variance function. The test is based on the likelihood ratio and the proposed test is comparable or outperforms existing methods under various settings. Chapter 2 is based on the published paper by Meyer et al. (2018), and Chapter 3 is based on the submitted paper (Kim et al., 2018).

Chapter 2

Convergence Rates for Constrained Regression Splines[1]

2.1 Introduction and Background

We consider maximum likelihood estimation of a real-valued function f with domain $[0, 1]$, where f is assumed to have sufficient smoothness and be monotone. Given observed data, f is estimated using regression splines. Huang (2001) showed that, under mild conditions the unconstrained regression spline estimators in concave extended linear models attained optimal rates as established by Stone (1980). Monotone regression splines were developed by Ramsay (1988), and Meyer (2008) extended the idea to convex function estimation. The set of shape-constrained spline functions is a convex polyhedral cone that is a subset of the linear space of spline functions. The purpose of this chapter is to show that the constrained splines also attain optimal rates.

We will begin with the notation similar to that of Huang (2001). The regression function f will be estimated from i.i.d. random observations $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_n$ with probability density $p(f, \boldsymbol{w})$. Let $l(f; \boldsymbol{w}) = \log p(f, \boldsymbol{w})$ and $\Lambda(f) = \mathrm{E}(l(f; \boldsymbol{W}))$ be the log-likelihood and expected log-likelihood functions, respectively, and let

$$\ell(f) = \frac{1}{n} \sum_{i=1}^{n} l(f; \boldsymbol{W}_i)$$

be the scaled log-likelihood. The linear space $\mathbb{G}$ of regression splines is determined by the spline order and the knots. Let $\| \cdot \|$ be the L_2 norm. Throughout Chapter 2, the following assumptions hold.

(A1) The true function f lies in $\mathbb{H}$, the space of real-valued functions on $[0, 1]$ with the third continuous derivative.

---

[1]Chapter 2 is based on the published paper (Meyer et al., 2018).

(A2) The number K_n of knots grows with n; the knots $0 = t_1 < t_2 < \cdots < t_{K_n} = 1$ have bounded mesh ratio; that is, there is an $M_0 \in (1, \infty)$ not depending on n, such that for all n,

$$M_0^{-1} K_n^{-1} \le t_i - t_{i-1} \le M_0 K_n^{-1}, \quad \text{for all } i = 2, \ldots, K_n.$$

(A3) The true function satisfies the constraints: $f'(x) \le 0$ for $x \in [0, 1]$.

(A4) The assumptions about the model are similar to those for the concave extended linear model as defined in Huang (2001).

(i) The expected log-likelihood function Λ is smooth; that is, for any $K_1 > 0$ there are $M_1, M_2 \in (0, \infty)$, such that for any h_1 and h_2 in $\mathbb{H}$ with $\max_{x \in [0,1]} |h_1(x)| \le K_1$ and $\max_{x \in [0,1]} |h_2(x)| \le K_1$,

$$-M_1 \|h_2 - h_1\|^2 \le \frac{d^2}{d\alpha^2} \Lambda(h_1 + \alpha(h_2 - h_1)) \le -M_2 \|h_2 - h_1\|^2$$

(ii) The maximizer $\bar{f}$ of Λ in $\mathbb{G}$ exists and $\max_{x \in [0,1]} |\bar{f}(x)| \le K_0$ for some $K_0 \in (0, \infty)$.

(iii) The log-likelihood is sufficiently flat at $\bar{f}$:

$$\sup_{g \in \mathbb{G}} \frac{|\frac{d}{d\alpha} \ell(\bar{f} + \alpha g)|_{\alpha=0}|}{\|g\|} = O_p\left((K_n/n)^{1/2}\right).$$

(iv) For any positive constant K_2, there exist $M_3 \in (0, \infty)$ and $M_4 \in (0, \infty)$ such that for any $g_1, g_2 \in \mathbb{G}$ with $\max_{x \in [0,1]} |g_1(x)| \le K_2$ and $\max_{x \in [0,1]} |g_2(x)| \le K_2$,

$$-M_3 \|g_1 - g_2\|^2 \le \frac{d^2}{d\alpha^2} \ell(g_1 + \alpha(g_2 - g_1)) \le -M_4 \|g_1 - g_2\|^2$$

except on an event with probability tending toward zero as n increases without bound.

8

2.2 Convergence Rates

Let $\widetilde{f}$ maximize the likelihood over $\mathbb{G}$; under the above conditions, Huang (2001) established rates for the approximation error $\bar{f} - f$ and estimation error $\widetilde{f} - \bar{f}$, based on the number of knots and spline order. Because increasing the number of knots tends to improve the approximation error and decreasing the number of knots improves the estimation error rate, the optimal rate for the estimator is achieved when the rates for each type of error are equal. This occurs when $K_n \asymp n^{1/7}$.

To obtain the estimation error rate, Huang (2001) determined a constant a so that $\|g - \bar{f}\| = a(K_n/n)^{1/2}$ implies $\ell(g) < \ell(\bar{f})$ under the conditions given. Because $\ell(\widetilde{f}) \geq \ell(\bar{f})$ by definition of $\widetilde{f}$, and by concavity of ℓ, we must have $\|\widetilde{f} - \bar{f}\| \leq a(K/n)^{1/2}$, and the estimation error rate is established for the unconstrained spline estimator. For the constrained estimator, we let $\widehat{f}$ maximize ℓ over the subset $\mathbb{C}$ of $\mathbb{G}$ that contains the constrained spline functions. If the constraints hold for $\bar{f}$, then $\ell(\widehat{f}) \geq \ell(\bar{f})$ using the same argument as for the unconstrained case, and thus $\|\widehat{f} - \bar{f}\| \leq a(K_n/n)^{1/2}$.

If the constraints hold strictly for f, the approximant $\bar{f}$ will be in $\mathbb{C}$ for a fine enough mesh, but if f has intervals over which the constraints do not hold strictly, i.e., flat spots for the increasing assumption or linear stretches for the convex assumption, there is no guarantee that $\bar{f} \in \mathbb{C}$ as n gets large. To establish the convergence rate for the constrained splines, we will find an $f^* \in \mathbb{C}$ such that for all $x \in [0,1]$, $|f^*(x) - \bar{f}(x)| \leq cK_n^{-3} = c(K_n/n)^{1/2}$ for $K_n \asymp n^{1/7}$, where c does not depend on K_n. Then, using concavity of ℓ, we can show that the constrained splines attain the same optimal rate as the unconstrained splines.

Theorem 6.25 of Schumaker (2007) states that, under (A1) and (A2), there is an $\widetilde{g} \in \mathbb{G}$ and constants C_0, C_1, and C_2 not depending on K_n, such that

$$\max_{x \in [0,1]} |\widetilde{g}(x) - f(x)| \leq C_0 K^{-3}, \tag{2.1}$$

$$\max_{x \in [0,1]} |\widetilde{g}'(x) - f'(x)| \leq C_1 K^{-2}, \tag{2.2}$$

$$\max_{x \in [0,1]} |\widetilde{g}''(x) - f''(x)| \leq C_2 K^{-1}. \tag{2.3}$$

9

In the next section, we will prove the following theorem.

Theorem 1. *Suppose that $f \in \mathbb{H}$ and satisfies the constraints. There exists a function $f^* \in \mathbb{C}$ such that, for all $x \in [0, 1]$, $\|f^*(x) - f(x)\| \leq cK_n^{-3} = O(K_n/n)^{1/2}$ where $K_n = O(n^{1/7})$, and c does not depend on K_n.*

The above theorem quantifies the approximation error of f with smoothness and shape constraints. In addition, because f is close to $\bar{f}$, $\tilde{g}$, and $\tilde{f}$, by the triangle inequality the L_2 distance between any pair of $\tilde{g}$, f^*, $\bar{f}$, f, and $\tilde{f}$ is on the order of $(K_n/n)^{1/2}$. Therefore, there is an $A > 0$ not depending on n, so that $\|f^* - \tilde{f}\| \leq AK_n^{-3}$.

By Taylor's expansion, we have for g_1, $g_2 \in \mathbb{G}$,

$$\ell(g_1) = \ell(g_2) + \frac{d}{d\alpha}\ell(g_2 + \alpha(g_1 - g_2))|_{\alpha=0} + \int_0^1 (1 - \alpha)\frac{d^2}{d\alpha^2}\ell(g_2 + \alpha(g_1 - g_2))d\alpha.$$

Because $\frac{d}{d\alpha}\ell(\tilde{f} + \alpha(g - \tilde{f}))|_{\alpha=0} = 0$ for all $g \in \mathbb{G}$, by (A4(iv)) we have

$$\frac{M_4}{2}\|\tilde{f} - g\|^2 \leq \ell(\tilde{f}) - \ell(g) \leq \frac{M_3}{2}\|\tilde{f} - g\|^2.$$

Therefore

$$\|\hat{f} - \tilde{f}\|^2 \leq \frac{2}{M_4}[\ell(\tilde{f}) - \ell(\hat{f})] \leq \frac{2}{M_4}[\ell(\tilde{f}) - \ell(f^*)] \leq \frac{M_3}{M_4}\|\tilde{f} - f^*\|^2 \leq \frac{A^2 M_3}{M_4}K^{-6}.$$

Finally, we have the following result on the estimation error.

Theorem 2. *Suppose that $f \in \mathbb{H}$. The maximizer $\hat{f}$ of ℓ over $\mathbb{C}$ satisfies that, for all $x \in [0, 1]$, $\|f^*(x) - \hat{f}(x)\| = O_p(K_n/n)^{1/2}$, where $K = O(n^{1/7})$.*

2.3 Technical Details

2.3.1 Useful Lemmas

For the proofs of the Theorem 1 in Section 2.3.2, Lemmas 1 through 4 are used with the assumption that f is a monotone function. For all $x \in [0, 1]$, there exists coefficients $\widetilde{\alpha}_1, \cdots, \widetilde{\alpha}_{L_n}$ so that $\widetilde{g}(x) = \sum_{k=1}^{L_n} \widetilde{\alpha}_k B_{k,2}(x) \in \mathbb{G}$. Let $B_{k,2} = B_k$ for simplification.

Lemma 1. *Let $\widetilde{g} \in \mathbb{G}$ be the approximant satisfying (2.1). If $0 \leq a < b \leq 1$, then $\widetilde{g}(b) - \widetilde{g}(a) \leq 2C_0 K_n^{-3}$.*

Proof. Since the function f is non-increasing, we have

$$0 \leq f(a) - f(b) = f(a) - \widetilde{g}(a) + \widetilde{g}(a) - \widetilde{g}(b) + \widetilde{g}(b) - f(b).$$

As a consequence of (2.1), we can obtain

$$\widetilde{g}(b) - \widetilde{g}(a) \leq f(a) - \widetilde{g}(a) + \widetilde{g}(b) - f(b) \leq |f(a) - \widetilde{g}(a)| + |\widetilde{g}(b) - f(b)| \leq 2C_0 K_n^{-3}.$$

$\diamondsuit$

Lemma 2. *Let $\widetilde{g} \in \mathbb{G}$ be the approximant satisfying (2.1) and (2.2). For any $x \in [0, 1]$, we have $\widetilde{g}'(x) \leq C_1 K_n^{-2}$.*

Proof. If $\widetilde{g}(x)$ is decreasing at any $x \in [0, 1]$, we have $\widetilde{g}'(x) < 0$. Suppose $\widetilde{g}$ is non-decreasing on $[a, b]$ where $0 \leq a \leq b \leq 1$. Since f is a non-increasing function, $f'(x) \leq 0$ for all x. Thus, if $\widetilde{g}'(x) \geq 0$ for some $x \in [a, b]$, then the desired result is a direct consequence of (2.2).

$\diamondsuit$

Lemma 3. *Let $\widetilde{g} \in \mathbb{G}$ be the approximant satisfying (2.1) and (2.2). Then, we have $\widetilde{\alpha}_{k+1} - \widetilde{\alpha}_k \leq M_0 C_1 K_n^{-3}$ for $k = 1, \ldots, L_n - 1$.*

Proof. First suppose that $\widetilde{g}$ is decreasing at t_j for any $j \in [1, K_n]$. Then, we have $\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_j < 0$ from the properties of B-splines and (1.1). Next, suppose that $\widetilde{g}$ is non-decreasing at t_j for any

11

$j \in [1, K_n]$. Then, from (1.1) and Lemma 2, we have

$$0 \le \widetilde{g}'(t_j) = \frac{2(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_j)}{t_{j+3} - t_{j+1}} B_{j,1}(t_j) \le C_1 K_n^{-2}. \tag{2.4}$$

From (A2) and (2.4), we conclude that $\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_j \le M_0 C_1 K_n^{-3}$. Note that $0 \le B_{k,p}(x) \le 1$ by the property of B-spline basis functions for any p and k for $x \in [0, 1]$. Therefore, we obtained required results.

$$\diamondsuit$$

Lemma 4. *Let $\widetilde{g} \in \mathbb{G}$ be the approximant satisfying (2.1) and (2.2). Let ℓ be an index such that $\widetilde{g}'(t_\ell) > 0$. Suppose there exists $m > \ell$ which is the smallest index such that $\widetilde{\alpha}_m < \widetilde{\alpha}_\ell$. Then, we have $0 \le \widetilde{\alpha}_k - \widetilde{\alpha}_\ell \le 2(C_0 + M_0 C_1) K_n^{-3}$ for all $k = \ell + 2, \ldots, m - 1$.*

Proof. From Lemma 1, for any $j = \ell + 1, \cdots, m - 2$,

$$
\begin{aligned}
2C_0 K_n^{-3} &\ge \widetilde{g}(t_j) - \widetilde{g}(t_\ell) \\
&= \widetilde{\alpha}_j B_j(t_j) + \widetilde{\alpha}_{j+1} B_{j+1}(t_j) - \widetilde{\alpha}_\ell B_\ell(t_\ell) - \widetilde{\alpha}_{\ell+1} B_{\ell+1}(t_\ell) \\
&= \widetilde{\alpha}_j B_j(t_j) + \widetilde{\alpha}_{j+1} B_{j+1}(t_j) - \widetilde{\alpha}_\ell \big(B_\ell(t_\ell) + B_{\ell+1}(t_\ell)\big) - \big(\widetilde{\alpha}_{\ell+1} - \widetilde{\alpha}_\ell\big) B_{\ell+1}(t_\ell) \\
&= \big(\widetilde{\alpha}_j - \widetilde{\alpha}_\ell\big) B_j(t_j) + \big(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell\big) B_{j+1}(t_j) - \big(\widetilde{\alpha}_{\ell+1} - \widetilde{\alpha}_\ell\big) B_{\ell+1}(t_\ell) \tag{2.5}
\end{aligned}
$$

The last equality in (2.5) holds from the property of quadratic B-splines

$$1 = B_\ell(t_\ell) + B_{\ell+1}(t_\ell) = B_j(t_j) + B_{j+1}(t_j).$$

If $\widetilde{\alpha}_j \ge \widetilde{\alpha}_{j+1}$, then from (2.5) and Lemma 3,

$$
\begin{aligned}
2C_0 K_n^{-3} + M_0 C_1 K_n^{-3} &\geq \left(\widetilde{\alpha}_j - \widetilde{\alpha}_\ell\right) B_j(t_j) + \left(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell\right) B_{j+1}(t_j) \\
&\geq \left(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell\right)\left(B_j(t_j) + B_{j+1}(t_j)\right) \\
&= \widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell \\
&\geq 0.
\end{aligned}
\tag{2.6}
$$

Similarly, for $\widetilde{\alpha}_j < \widetilde{\alpha}_{j+1}$, we have

$$
\begin{aligned}
2C_0 K_n^{-3} + M_0 C_1 K_n^{-3} &\geq \left(\widetilde{\alpha}_j - \widetilde{\alpha}_\ell\right) B_j(t_j) + \left(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell\right) B_{j+1}(t_j) \\
&= \left(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell\right) B_{j+1}(t_j) + \left(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell\right) B_j(t_j) \\
&\quad - \left(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_j\right) B_j(t_j) \\
&= \widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell - \left(\widetilde{\alpha}_{j+1} - \widetilde{\alpha}_j\right) B_j(t_j).
\end{aligned}
\tag{2.7}
$$

Then, from Lemma 3 and (2.7), we have

$$
2(C_0 + M_0 C_1) K_n^{-3} \geq \widetilde{\alpha}_{j+1} - \widetilde{\alpha}_\ell \geq 0,
\tag{2.8}
$$

when $\widetilde{\alpha}_j < \widetilde{\alpha}_{j+1}$. From Lemma 3, (2.6), and (2.8), we obtained the required results.

$$\diamondsuit$$

2.3.2 Proof of Theorem 1

In this section, we will given the intuition and ideas of the proof for the existence of f^* for monotone splines. The alternative proof of Theorem 1 when the true function f is non-decreasing is available in Meyer et al. (2018).

Recall that we use quadratic splines for monotone constraints. Such a choice is due to the fact that a quadratic spline function is non-increasing if and only if the slopes at the knots are non-positive. Thus, we have linear inequality constraints that are necessary and sufficient conditions

for monotonicity. Without loss of generality, we will treat explicitly the case in which f is non-increasing. It does not follow that its approximant $\widetilde{g}$ is non-increasing.

We construct a function f^* that is non-increasing and $|f^*(x) - \widetilde{g}(x)| \leq CK_n^{-3}$ for all $x \in [0, 1]$ for a constant $C > 0$ not depending on K_n. A quadratic spline function g is non-increasing in $[0, 1]$ if and only if $g'(t_j) \leq 0$, for all $j = 1, \ldots, K_n$.

There are three scenarios to construct function f^* by modifying $\widetilde{g}$ based on the sign of $\widetilde{g}'$ at each knot. The three scenarios are described as below.

Scenario 1. *If $\widetilde{g}'(t_j) \geq 0$ for all $j = 1, \ldots, K_n$, define $f^*(x) = \widetilde{g}(0)$ for $x \in [0, 1]$; then $|f^*(x) - \widetilde{g}(x)| \leq 2C_0 K_n^{-3} = O(K_n^{-3})$ for all $x \in [0, 1]$.*

Proof. Because $\widetilde{g}$ is non-decreasing over $[0, 1]$, for any $x \in [0, 1]$, we have

$$0 \leq \widetilde{g}(x) - f^*(x) = \widetilde{g}(x) - \widetilde{g}(0) \leq 2C_0 K_n^{-3}$$

by Lemma 1.

$\diamondsuit$

Scenario 2. *If $\widetilde{g}'(t_j) > 0$ for $j = 1, \ldots, \ell - 1$ but $\widetilde{g}'(t_\ell) \leq 0$ for some $\ell \in \{2, \ldots, K_n\}$, there is a quadratic spline function g such that $|g(x) - \widetilde{g}(x)| \leq (C_1 M_0 + 2C_0) K_n^{-3} = O(K_n^{-3})$ for all $x \in [0, 1]$ and $g'(x) \leq 0$ on $[0, t_\ell]$.*

Proof. Define a spline function $g(x) = \sum_{k=1}^{L_n} \alpha_k B_k(x)$ on $[0, 1]$ such that $g(x) = \widetilde{g}(x)$ on $[t_\ell, 1]$ and $g'(t_j) = 0$ for $j = 1, \ldots, t_{\ell-1}$ (as depicted on the left in Figure 2.1). The coefficients are defined as follows:

$$\alpha_j = \begin{cases} \widetilde{\alpha}_\ell & \text{for } j = 1, \ldots, \ell - 1 \\ \widetilde{\alpha}_j & \text{for } j = \ell, \ldots, L_n. \end{cases} \tag{2.9}$$

Then for $x \in [0, t_{\ell-1}]$, $g(x) = g(t_{\ell-1})$. Because $\widetilde{g}(x)$ is increasing on $[0, t_{\ell-1}]$, we have on this range $g(x) - \widetilde{g}(x) \geq 0$ and

14

$$
\begin{aligned}
g(x) - \widetilde{g}(x) &= g(x) - \widetilde{g}(t_{\ell-1}) + \widetilde{g}(t_{\ell-1}) - \widetilde{g}(x) \\
&= g(t_{\ell-1}) - \widetilde{g}(t_{\ell-1}) + \widetilde{g}(t_{\ell-1}) - \widetilde{g}(x) \\
&= (\alpha_{\ell-1} - \widetilde{\alpha}_{\ell-1})B_{\ell-1}(t_{\ell-1}) + (\alpha_\ell - \widetilde{\alpha}_\ell)B_\ell(t_{\ell-1}) + \widetilde{g}(t_{\ell-1}) - \widetilde{g}(x) \\
&= (\widetilde{\alpha}_\ell - \widetilde{\alpha}_{\ell-1})B_{\ell-1}(t_{\ell-1}) + \widetilde{g}(t_{\ell-1}) - \widetilde{g}(x) \\
&\leq (C_1 M_0 + 2C_0)K_n^{-3}
\end{aligned}
$$

by Lemmas 1 and 3. For $x \in [t_{\ell-1}, t_\ell]$,

$$
\begin{aligned}
g(x) - \widetilde{g}(x) &= (\alpha_{\ell-1} - \widetilde{\alpha}_{\ell-1})B_{\ell-1}(x) + (\alpha_\ell - \widetilde{\alpha}_\ell)B_\ell(x) + (\alpha_{\ell+1} - \widetilde{\alpha}_{\ell+1})B_{\ell+1}(x) \\
&= (\widetilde{\alpha}_\ell - \widetilde{\alpha}_{\ell-1})B_{\ell-1}(x) \\
&\leq C_1 M_0 K_n^{-3},
\end{aligned}
$$

by Lemma 3.

$$\diamondsuit$$

In Scenario 2, if $\widetilde{g}$ is non-decreasing on $[t_\ell, 1]$, set $f^* = g$, as g is a spline function that is sufficiently close to $\widetilde{g}$ and non-increasing on $[0, 1]$. This is shown on the left in Figure 2.1. Otherwise, replace $\widetilde{g}$ with g, and continue. That is, we now assume that $\widetilde{g}'(0) \leq 0$.

Scenario 3. *If $\widetilde{g}'(t_j) \leq 0$ for $j = 1, \ldots, \ell-1$, but $\widetilde{g}'(t_\ell) > 0$, find the smallest index $m \geq \ell+1$ such that $\widetilde{\alpha}_m \leq \widetilde{\alpha}_\ell$. If there is such an m, then there is a quadratic spline function g where $g'(x) \leq 0$ for $x \in [0, t_m]$ and $|g(x) - \widetilde{g}(x)| \leq 4(C_0 + C_1 M_0)K_n^{-3}$ for $x \in [0, 1]$. Otherwise, we can find a spline function g such that $g'(x) \geq 0$ and $|g(x) - \widetilde{g}(x)| \leq (2C_0 + C_1 M_0)K_n^{-3}$ for $x \in [0, 1]$.*

Proof. For the first part, where such an m exists, set $g(x) = \widetilde{g}(x)$ for $x \in [0, t_{\ell-1}] \cup [t_m, 1]$. Further set $g'(t_\ell) = \cdots = g'(t_{m-2}) = 0$; for $g(x) = \sum_{k=1}^{L_n} \alpha_k B_k(x)$, this is equivalent to set

$$
\alpha_j = \begin{cases} \widetilde{\alpha}_j & \text{for } j = 1, \ldots, \ell-1, m, \ldots, L_n \\ \widetilde{\alpha}_\ell & \text{for } j = \ell, \ldots, m-1. \end{cases} \tag{2.10}
$$
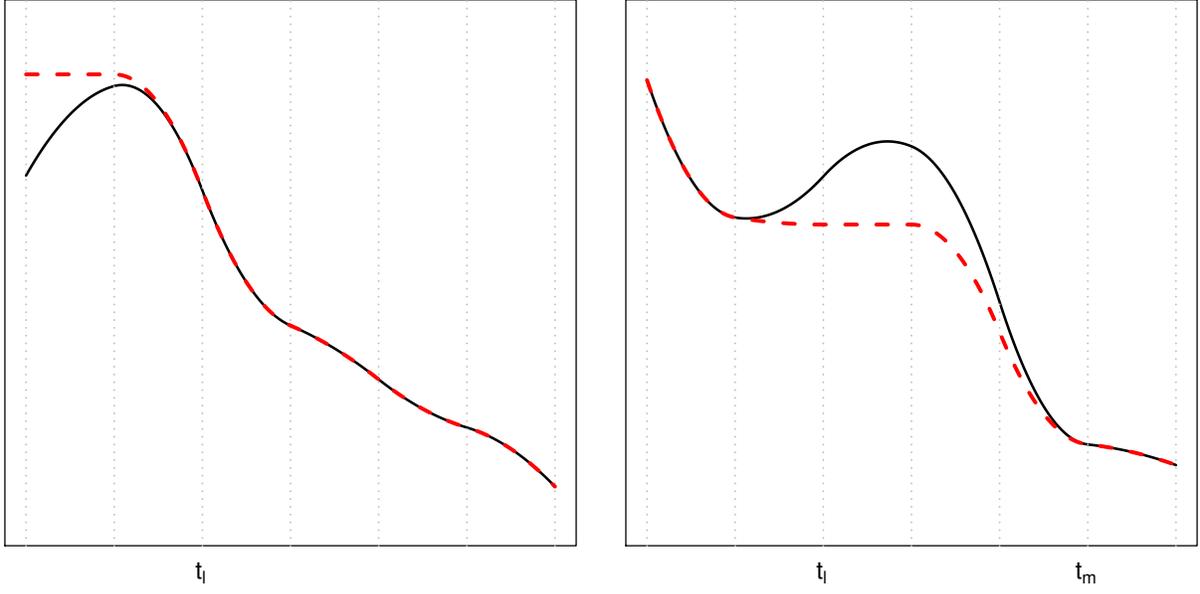
Figure 2.1: On the left is the Illustration of Scenario 2 ; on the right is the illustration of Scenario 3. For each, the solid curve is $\widetilde{g}$ and the dashed curve is g.

Then, g is non-increasing on $[0, t_m]$, and $g(x) = g(t_\ell)$ for $x \in [t_\ell, t_{m-2}]$. See the right-hand plot of

Figure 2.1 for an illustration. For $x \in [t_{\ell-1}, t_\ell]$, we have

$$
\begin{aligned}
0 \leq \widetilde{g}(x) - g(x) &= (\widetilde{\alpha}_{\ell-1} - \alpha_{\ell-1}) B_{\ell-1}(x) + (\widetilde{\alpha}_\ell - \alpha_\ell) B_\ell(x) + (\widetilde{\alpha}_{\ell+1} - \alpha_{\ell+1}) B_{\ell+1}(x) \\
&= (\widetilde{\alpha}_{\ell+1} - \alpha_{\ell+1}) B_{\ell+1}(x) \\
&= (\widetilde{\alpha}_{\ell+1} - \widetilde{\alpha}_\ell) B_{\ell+1}(x) \\
&\leq M_0 C_1 K_n^{-3},
\end{aligned}
$$

by Lemma 3 and (2.10). Next, denote $b = \operatorname{argmax}_{x \in [t_\ell, t_m]} \widetilde{g}(x)$. For $x \in [t_\ell, t_{m-2}]$, we have

16

$$
\begin{aligned}
0 \leq \widetilde{g}(x) - g(x) &= \widetilde{g}(x) - g(t_\ell) \\
&\leq \widetilde{g}(b) - g(t_\ell) \\
&= \widetilde{g}(b) - \widetilde{g}(t_\ell) + \widetilde{g}(t_\ell) - g(t_\ell) \\
&= \widetilde{g}(b) - \widetilde{g}(t_\ell) + (\widetilde{\alpha}_\ell - \alpha_\ell)B_\ell(t_\ell) + (\widetilde{\alpha}_{\ell+1} - \alpha_{\ell+1})B_{\ell+1}(t_\ell) \\
&= \widetilde{g}(b) - \widetilde{g}(t_\ell) + (\widetilde{\alpha}_{\ell+1} - \widetilde{\alpha}_\ell)B_{\ell+1}(t_\ell) \\
&\leq (2C_0 + M_0C_1)K_n^{-3},
\end{aligned}
$$

by (2.10), Lemmas 1 and 3. For $x \in [t_{m-2}, t_{m-1}]$, from (2.10), and Lemma 4:

$$
\begin{aligned}
\widetilde{g}(x) - g(x) &= (\widetilde{\alpha}_{m-2} - \alpha_{m-2})B_{m-2}(x) + (\widetilde{\alpha}_{m-1} - \alpha_{m-1})B_{m-1}(x) \\
&= (\widetilde{\alpha}_{m-2} - \widetilde{\alpha}_\ell)B_{m-2}(x) + (\widetilde{\alpha}_{m-1} - \widetilde{\alpha}_\ell)B_{m-1}(x) \\
&\leq 4(C_0 + M_0C_1)K_n^{-3}.
\end{aligned}
$$

Lastly, for $x \in [t_{m-1}, t_m]$, we have

$$
\begin{aligned}
\widetilde{g}(x) - g(x) &\leq (\widetilde{\alpha}_{m-1} - \alpha_{m-1})B_{m-1}(x) + (\widetilde{\alpha}_m - \alpha_m)B_m(x) + (\widetilde{\alpha}_{m+1} - \alpha_{m+1})B_{m+1}(x) \\
&= (\widetilde{\alpha}_{m-1} - \widetilde{\alpha}_\ell)B_{m-1}(x) \\
&\leq 2(C_0 + M_0C_1)K_n^{-3},
\end{aligned}
$$

by (2.10), and Lemma 4. Therefore, we have $0 \leq \widetilde{g}(x) - g(x) \leq 4(C_0 + M_0C_1)K_n^{-3}$ for $x \in [0, t_m]$, and we have the required result.

For the second Scenario, where such an m does not exist, define $g(x) = \widetilde{g}(x)$ for $x \in [0, t_{\ell-1}]$ and $g'(t_j) = 0$ for $j = \ell, \ldots, K_n$. That is equivalent to set:

$$
\alpha_j = \begin{cases} \widetilde{\alpha}_j & \text{for } j = 1, \ldots, \ell-1 \\ \widetilde{\alpha}_\ell & \text{for } j = \ell, \ldots, L_n. \end{cases}
$$

17

Then g is non-decreasing in $[0, 1]$, and for $x \in [t_{\ell-1}, t_\ell]$, the same argument as for the first Scenario will show g and $\widetilde{g}$ are sufficiently close. For $x > t_\ell$, $g(x) = g(t_\ell)$. For $x \in [t_j, t_{j+1}]$, $j = \ell, \ldots, K_n - 1$, again the same argument as for the first Scenario will show g and $\widetilde{g}$ are sufficiently close.

$\diamondsuit$

In the case where such an m does not exist, we are finished because the function g is non-decreasing on $[0, 1]$ and sufficiently close to $\widetilde{g}$. For the case where such an m exists (illustrated on the right in Figure 2.1 where the solid curve is $\widetilde{g}$ and the dashed curve is g), we have a g that is nondecreasing on $[0, t_m]$. If there is an index $m' > m$ where $\widetilde{g}(t_{m'}) > 0$, we can repeat Scenario 3 as many times as necessary to produce a g that is non-decreasing on $[0, 1]$.

Chapter 3

Monotone Variance Function Estimation

3.1 Literature Review

Various methods for estimating variance functions in heteroskedastic regression models have been developed over the years because ignoring heteroskedasticity in the regression analysis may lead to substantial loss of efficiency and incorrect inference. Parametric variance function estimation methods based on absolute residuals or squared residuals were introduced by Davidian and Carroll (1987). Carroll and Ruppert (1988) summarized the methods to fit the heteroskedastic regression model including transformation and weighting techniques. These methods are useful; however, in general, a parametric form of the variance function is unknown.

Both kernel and spline methods have been used for nonparametric variance function estimation. Carroll (1982) proposed a nonparametric method using kernel smoothing to estimate the variance, when it is modeled as a function of the mean response. Müller and Stadtmuller (1987) introduced local variance estimators using kernel smoothers without the assumption of a parametric form of the mean function. Hall and Carroll (1989) and Wang et al. (2008) studied the effect of the smoothness of the unknown mean function on the rate of convergence for kernel estimators of the variance functions. Ruppert et al. (1997) proposed the local polynomial estimate of the mean and variance functions using a standard bandwidth select method. For example, one could use the bandwidth introduced in Ruppert et al. (1995) or Fan and Gijbels (1996). Liu et al. (2007) proposed smoothing splines to estimate variance functions for chi-squared distributed response variables. Staudenmayer et al. (2008) estimated the density of a random variable in the presence of heteroskedastic measurement errors, using a spline method based on the Metropolis-Hastings algorithm and Gibbs sampling. Gijbels et al. (2010) considered the extended double-exponential family models and P-splines to estimate the mean function and the dispersion function simulta-

neously. Lian et al. (2015) developed an estimation method for the mean and variance functions where the mean function is in a partially linear single-indexed model.

Monotonicity of the variance function is often a more reasonable assumption, compared to homogeneity or linearity assumptions. Dette and Pilz (2009) implemented a two-step kernel-based method to estimate a monotone variance function, where in step one, an unconstrained estimate of the variance function is a kernel smooth (bandwidth k_r) of the pseudo residuals or the residuals from an estimate of the mean function. Step two monotonizes the estimate using another kernel with bandwidth k_d. The estimate seems to be sensitive to the choice of the two bandwidths, and sometimes yields negative values of the variance function. The monotone variance function estimation method is conveniently coded in an R package `monreg` by Pilz and Titoff (2015).

In this chapter, we obtain maximum likelihood estimators of the variance function using quadratic regression splines, when either normal or double-exponential errors are assumed. Constraining the slope of the spline function to be positive (negative) at the knots provides a necessary and sufficient condition for an increasing (decreasing) function over the range of knots. (Unconstrained regression splines can be sensitive to the number and placement of knots, but constrained splines are robust to knot choices.) Our focus is on improving estimation and inference about the regression function through weighting provided by the estimated variance function. We first compute the unweighted regression function using least squares for the normal errors assumption or least absolute deviations for the double-exponential errors assumption. The residuals from this fit are used to estimate the variance function, which in turn is used to obtain an improved weighted regression function estimate, the residuals from which can be used to incrementally improve the variance function estimate. We show formally that this procedure of alternate improvements of the mean and variance functions produces consistent estimators; in practice two or three iterations are sufficient. As shown in the simulations, if the errors are "heavy-tailed" the double-exponential assumptions provide better inference about the mean function. The method has been incorporated within the R package `cgam`.

3.2 Model Setup and Monotone Variance Function Estimator

Consider a random sample $\{(x_i, y_i, \boldsymbol{z}_i), i = 1, ..., n\}$, from the model

$$y_i = \mu(\boldsymbol{z}_i) + \sigma(x_i)\epsilon_i, \quad i = 1, \ldots, n \tag{3.1}$$

where the ϵ_i's are i.i.d. mean-zero random errors, independent of the $(x_i, \boldsymbol{z}_i)$s. We assume $x_i \in [a, b]$, $i = 1, \ldots, n$ for $-\infty < a < b < \infty$ and without loss of generality, $a = 0$ and $b = 1$. The design vector $\boldsymbol{z}_i \in \mathbb{R}^{q+1}$ might include x_i but not necessarily where q is the number of predictors for the mean function. To begin with, we assume the mean function is known, and for simplicity, assume $\mu(\boldsymbol{z}_i) \equiv 0$ in (3.1). Throughout the paper, we assume that the variance function $\sigma^2(x)$ is non-decreasing on $[0, 1]$. Denote $f(x) = 1/\sigma^2(x)$ for the normal errors assumption and $f(x) = 1/\sigma(x)$ for double-exponential errors. It will be seen that the log-likelihood function (for either error distribution) is concave in f; thus the maximum likelihood variance function is obtained through estimating f, a non-increasing function.

To obtain a spline estimate of f, define knots $0 = t_1 < \cdots < t_{K_n} = 1$ and quadratic spline basis functions $B_1(x), \ldots, B_{L_n}(x)$ where $L_n = K_n + 1$. Let $\mathbb{G}$ be the linear space spanned by $B_1, \ldots, B_{L_n}$, and let $\mathbb{C} \subset \mathbb{G}$ be the collection of non-increasing, positive spline functions. The spline function $f(x) = \sum_{j=1}^{L_n} \alpha_j B_j(x)$ is non-increasing and positive if and only if $\boldsymbol{A}\boldsymbol{\alpha} \geq 0$, where $\boldsymbol{A}$ is an $L_n \times L_n$ constraint matrix and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{L_n})^\top$. The constraint matrix is defined as follows:

$$\boldsymbol{A} = - \begin{bmatrix} B_1'(t_1) & B_2'(t_1) & \ldots & B_{L_n}'(t_1) \\ B_1'(t_2) & B_2'(t_2) & \ldots & B_{L_n}'(t_2) \\ \vdots & \vdots & & \vdots \\ B_1'(t_{K_n}) & B_2'(t_{K_n}) & \ldots & B_{L_n}'(t_{K_n}) \\ B_1(1) & B_2(1) & \ldots & B_{L_n}(1) \end{bmatrix}$$

Given observed $x_1, \ldots, x_n$, let $\boldsymbol{B}$ be an $n \times L_n$ matrix where $B_{ij} = B_j(x_i)$, and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)^\top$ with $\theta_i = f(x_i)$ for $i = 1, \ldots, n$. The following subset of $\mathbb{R}^n$ has a one-to-one mapping with the

set $\mathbb{C}$ of constrained spline functions:

$$\mathcal{C} = \{\boldsymbol{\theta} \in \mathbb{R}^n : \boldsymbol{\theta} = \boldsymbol{B\alpha}, \text{ where } \boldsymbol{A\alpha} \geq 0\}. \tag{3.2}$$

The set $\mathcal{C}$ is a convex cone because for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathcal{C}$ and $c_1, c_2 \geq 0$, we have $c_1\boldsymbol{\theta}_1 + c_2\boldsymbol{\theta}_2 \in \mathcal{C}$.

Adding parametrically-modeled covariates to the inverse variance function is straight forward. Suppose the vector $\boldsymbol{u}_i \in \mathbb{R}^r$ contains the covariate values for the ith observation, and $\boldsymbol{\beta} \in \mathbb{R}^r$ is an unknown parameter vector. If we define $\theta_i = f(x_i) + \boldsymbol{u}_i^\top \boldsymbol{\beta}$, $i = 1, \ldots, n$, then $\boldsymbol{\theta} = \boldsymbol{B\alpha} + \boldsymbol{U\beta}$ where the rows of the $n \times r$ matrix $\boldsymbol{U}$ are $\boldsymbol{u}_1^\top, \ldots, \boldsymbol{u}_n^\top$. In addition to the monotonicity constraint for $\boldsymbol{\alpha}$, we add linear inequality constraints to ensure the variance function is positive for all values of $(x, \boldsymbol{u})$. Let

$$\boldsymbol{B}_c = [\boldsymbol{B}|\boldsymbol{U}] = \left[\begin{array}{c|c} & \boldsymbol{u}_1^\top \\ \boldsymbol{B} & \vdots \\ & \boldsymbol{u}_n^\top \end{array} \right] \text{ and } \boldsymbol{\gamma} = (\alpha_1, \ldots, \alpha_{L_n}, \beta_1, \ldots, \beta_r)^\top$$

Then the cone can be written as

$$\mathcal{C} = \{\boldsymbol{\theta} \in \mathbb{R}^n : \boldsymbol{\theta} = \boldsymbol{B}_c\boldsymbol{\gamma}, \text{ where } \boldsymbol{A}_c\boldsymbol{\gamma} \geq 0\}, \tag{3.3}$$

where $\boldsymbol{A}_c$ is a constraint matrix which constrains the estimated function f to be non-increasing and the estimated reciprocal variance function to be positive.

For a simple example (which will be seen in the analysis of the abalone data in Chapter 3.6), let the covariate be a group indicator that correspond to $r = 1$, so that $\boldsymbol{U}$ is $n \times 1$ with elements in $\{0, 1\}$ and $\boldsymbol{\beta} \in \mathbb{R}^1$. The coefficient vector is $\boldsymbol{\gamma}^\top = [\boldsymbol{\alpha}^\top|\boldsymbol{\beta}]$ and the $(L_n + 1) \times (L_n + 1)$ matrix $\boldsymbol{A}_c$ is defined as $\boldsymbol{A}_{c(i,j)} = \boldsymbol{A}_{ij}$ for $i = 1, \ldots, L_n$ and $j = 1, \ldots, L_n$; further, $\boldsymbol{A}_{c(L_n+1,j)} = B_j(1)$ for $j = 1, \ldots, L_n$, $\boldsymbol{A}_{c(L_n+1,L_n+1)} = 1$, and $\boldsymbol{A}_{c(j,L_n+1)} = 0$ for $j = 1; , \ldots, L_n$. That is,

$$
\boldsymbol{A}_c =
\begin{bmatrix}
& & & 0 \\
& \boldsymbol{A} & & 0 \\
& & & 0 \\
\hline
B_1(1) & \dots & B_{L_n}(1) & 1
\end{bmatrix}
$$

3.3 Maximum Likelihood Estimation

For maximum likelihood estimation of the variance function, it is necessary to assume a distribution for the ε_i. We consider both normal and double-exponential errors, as the methods and large sample theory are straight forward and remarkably similar. In either case, the log-likelihood function is concave in f, the reciprocal of the variance function for normal errors or the reciprocal of the standard deviation function for double-exponential errors.

3.3.1 Likelihood Functions

For the normal errors assumption, $f = 1/\sigma^2$. For observed $\{(x_i, \boldsymbol{u}_i, y_i),\ i = 1, \dots, n\}$, with knots $0 = t_1 < \cdots < t_{K_n} = 1$, construct the spline basis functions and the cone $\mathbb{C}$ as in (3.2) or (3.3), with $\theta_i = f(x_i)$ or $\theta_i = f(x_i, \boldsymbol{u}_i)$, respectively. The log-likelihood function with $\mu \equiv 0$ is defined by

$$
\ell(\boldsymbol{\theta}; \boldsymbol{y}) = \frac{1}{2n} \sum_{i=1}^{n} \left[\log(\theta_i) - y_i^2 \theta_i \right]. \tag{3.4}
$$

The maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ is the maximizer of $\ell(\boldsymbol{\theta}; \boldsymbol{y})$ in (3.4) over $\mathcal{C}$.

If we assume the ϵ_i's are independent random errors from the double-exponential distribution with mean zero and variance one, we let f be $1/\sigma$ instead of $1/\sigma^2$. The log-likelihood function with $\mu \equiv 0$ is

$$
\ell(\boldsymbol{\theta}; \boldsymbol{y}) = \frac{1}{n} \sum_{i=1}^{n} \left[\log(\theta_i) - \sqrt{2}|y_i|\theta_i \right], \tag{3.5}
$$

and $\widehat{\boldsymbol{\theta}}$ is the maximizer of $\ell(\boldsymbol{\theta}; \boldsymbol{y})$ in (3.5) over $\mathcal{C}$.

3.3.2 Estimation Procedure: Iterative Quadratic Programming

Let $\psi(\boldsymbol{\theta}) = -\ell(\boldsymbol{\theta}; \boldsymbol{y})$ for ℓ in either (3.4) or (3.5). In Appendix A.1, we prove that the sequence $\boldsymbol{\theta}_{(k)}$ generated by following algorithm will converge to $\widehat{\boldsymbol{\theta}}$.

Step 0. Initialize $\boldsymbol{\theta}_{(0)} \in \mathcal{C}$.

Step 1. At kth iteration, evaluate the gradient and the Hessian of $\psi(\boldsymbol{\theta}; \boldsymbol{y})$ at $\boldsymbol{\theta}_{(k)}$, denoted by $\nabla \psi(\boldsymbol{\theta}_{(k)})$ and $\boldsymbol{H}(\boldsymbol{\theta}_{(k)})$, respectively.

Step 2. Find the minimizer $\boldsymbol{a}$ of $\psi_k(\boldsymbol{\theta})$ over $\mathcal{C}$, where

$$\psi_k(\boldsymbol{\theta}) = \psi(\boldsymbol{\theta}_{(k)}) + \nabla \psi(\boldsymbol{\theta}_{(k)})^\top (\boldsymbol{\theta} - \boldsymbol{\theta}_{(k)}) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_{(k)})^\top \boldsymbol{H}(\boldsymbol{\theta}_{(k)})(\boldsymbol{\theta} - \boldsymbol{\theta}_{(k)}). \qquad (3.6)$$

Evaluate $\psi(\boldsymbol{a})$. If $\psi(\boldsymbol{a}) < \psi(\boldsymbol{\theta}_{(k)})$, let $\boldsymbol{\theta}_{(k+1)} = \boldsymbol{a}$. Otherwise, find $\boldsymbol{\theta}_{(k+1)}$ which minimizes $\psi(\boldsymbol{\theta})$ over the line segment connecting $\boldsymbol{\theta}_{(k)}$ and $\boldsymbol{a}$.

Step 3. Repeat Step 1 and Step 2 until the convergence criterion below is satisfied.

In Step 2, minimizing the quadratic function in (3.6) with linear constraints can be efficiently carried out through quadratic programming conveniently coded in the `qprog` function available in an R package `coneproj` (Liao and Meyer, 2014). For the normal errors model the convergence criterion is $\left(\sum_{i=1}^{n} |\widehat{\sigma}_{(k+1),i}^2 - \widehat{\sigma}_{(k),i}^2|\right) / \left(\sum_{i=1}^{n} |\widehat{\sigma}_{(k),i}^2|\right) \leq \delta$ where $\widehat{\sigma}_{(k+1),i}^2 = \boldsymbol{\theta}_{(k+1),i}^{-1}$, and for double-exponential errors we use the criterion $\left(\sum_{i=1}^{n} |\widehat{\sigma}_{(k+1),i} - \widehat{\sigma}_{(k),i}|\right) / \left(\sum_{i=1}^{n} |\widehat{\sigma}_{(k),i}|\right) \leq \delta$ where $\widehat{\sigma}_{(k+1),i} = \boldsymbol{\theta}_{(k+1),i}^{-1}$.

In practice, it is more common that the mean function is unknown. In this case, we can estimate the mean function and the variance function alternately for several iterations. First, the unknown mean function is estimated ignoring heteroskedasticity, using least-squares for the normal errors assumption and least absolute deviations under the assumption of double-exponential errors. Let the estimated mean function be $\widehat{\mu}(x)$, then the log-likelihood function with normal errors in (3.4)

can now be replaced by

$$\ell_{\hat{\mu}}(\boldsymbol{\theta}; \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \frac{1}{2n} \sum_{i=1}^{n} \left[\log(\theta_i) - \left(y_i - \hat{\mu}(\boldsymbol{z}_i) \right)^2 \theta_i \right].$$

The variance function is estimated using the above algorithm with $y_i - \hat{\mu}(\boldsymbol{z}_i)$ in place of y_i in (3.4) or (3.5). Next, the mean function can be updated using the weighted regression model, using the estimated variance function, and the mean function and the variance function are updated iteratively. For the simulations and examples in this paper, the mean function is updated twice.

3.4 Large Sample Properties

The following assumptions are used for the theorems in this section.

(A1) Assume f is a positive and non-increasing function in $\mathbb{H}$, the space of all three times differentiable, real-valued functions on $[0, 1]$, and there exists $M_0 > 1$ such that for all $x \in [0, 1]$

$$\frac{1}{M_0} \leq f(x) \leq M_0.$$

(A2) Assume that the number of distinct knots K_n grows as $\mathcal{O}(n^{1/7})$ with bounded mesh ratio. That is, there exists $M_1 > 1$ such that for all n, and any $j \in \{1, \ldots, K_n - 1\}$,

$$\frac{1}{M_1} K_n^{-1} \leq t_{j+1} - t_j \leq M_1 K_n^{-1}.$$

(A3) The observed x_i follow a design density $p_X(x)$ that is bounded away from zero: $p_X(x) \geq p_0 > 0$ for all $x \in [0, 1]$.

For any $h \in \mathbb{H}$, we define the norm as $||h||^2 = \langle h, h \rangle$ where $\langle h_1, h_2 \rangle = \int h_1(x) h_2(x) p_X(x) dx$. Let the monotone maximum likelihood estimator of f be $\hat{f}$, which maximizes the log-likelihood $\ell(g; \boldsymbol{y})$ in (3.4) over $g \in \mathcal{C}$. The following theorems are presented for the convergence rate for the estimated monotone variance functions.

Theorem 3. *Suppose that Assumptions (A1) through (A3) hold. In the known-mean case, if $\widehat{f}$ is the MLE of f in $\mathbb{C}$, then $||f - \widehat{f}|| = \mathcal{O}_p(n^{-3/7})$.*

Proof. For a positive function $h \in \mathbb{H}$ and normal errors, the log-likelihood is a constant plus

$$l(h; x, y) = \frac{1}{2} \log\big(h(x)\big) - \frac{1}{2}y^2 h(x), \tag{3.7}$$

which is concave in h for each value x and y. The expected log-likelihood can be expressed as

$$\Lambda(h) = E[l(h; X, Y)] = \frac{1}{2} \int_0^1 \Big[\log\big(h(x)\big) - \frac{h(x)}{f(x)}\Big] p_X(x) \, dx, \tag{3.8}$$

where $p_X(x)$ is a probability density function, and $\Lambda(h)$ is concave in h as well. The true reciprocal variance function f maximizes the expected log-likelihood function $\Lambda(h)$ over $h \in \mathbb{H}$.

Let the unconstrained maximum likelihood estimator of f in $\mathbb{G}$ be $\widetilde{f}$, which maximizes $\ell(\mathrm{g})$ over $\mathrm{g} \in \mathbb{G}$, and let $\bar{f} = \mathrm{argmax}_{g \in \mathbb{G}} \Lambda(g)$. Recall that $\mathbb{G} = \mathbb{G}(n)$ has increasing dimension as n increases. Following ideas from Huang (2001), the total error can be decomposed as $|\widetilde{f} - f| \leq |\bar{f} - f| + |\widetilde{f} - \bar{f}|$, and the rates of convergence for the approximation error $|\bar{f} - f|$ and the estimation error $|\widetilde{f} - \bar{f}|$ can be determined separately. With increasing numbers of knots, the approximation error decreases while the estimation error increases, so we obtain the optimal rate of convergence by setting equal the approximation and estimation errors. In the proofs, let $||h||_\infty = \sup_{x \in [0,1]} |h(x)|$. For vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, the vector norm is defined as $||\boldsymbol{u}||_n^2 = \langle \boldsymbol{u}, \boldsymbol{u} \rangle_n$, where $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_n = \frac{1}{n} \sum_{i=1}^n u_i v_i$.

We start with the approximation error rate. From Theorem 6.25 of Schumaker (2007) and (A2), there exists a sequence of functions $g^* \in \mathbb{G}$ such that $\rho_n := ||f - g^*||_\infty = \mathcal{O}_p(K_n^{-3})$. Since $\mathbb{G} \subset \mathbb{H}$, we have $\Lambda(g^*) \leq \Lambda(\bar{f}) \leq \Lambda(f)$. From (3.8), we have for $h \in \mathbb{H}$,

$$\Lambda((1-\alpha)f + \alpha h) = \frac{1}{2} \int_0^1 \Big[\log\big((1-\alpha)f(x) + \alpha h(x)\big) - \frac{(1-\alpha)f(x) + \alpha h(x)}{f(x)}\Big] p_X(x) dx,$$

and

26

$$\frac{d^2}{d\alpha^2}\Lambda((1-\alpha)f+\alpha h) = -\frac{1}{2}\int_0^1 \left[\frac{\big[h(x)-f(x)\big]^2}{\big[f(x)+\alpha\big(h(x)-f(x)\big)\big]^2}\right]p_X(x)dx. \tag{3.9}$$

Consider $h \in \mathbb{H}$ such that $M_0^{-1} \le h(x) \le M_0$ for $x \in (0,1)$. From (A1), we have

$$\frac{1}{M_0^2} \le \big[f(x)+\alpha\big(h(x)-f(x)\big)\big]^2 \le M_0^2,$$

so we can bound (3.9) as

$$-\frac{1}{2}M_0^2||h-f||^2 \le \frac{d^2}{d\alpha^2}\Lambda((1-\alpha)f+\alpha h) \le -\frac{1}{2M_0^2}||h-f||^2. \tag{3.10}$$

Integrating by parts, the expected log-likelihood can be expanded about f as

$$\Lambda(h) = \Lambda(f) + \frac{d}{d\alpha}\Lambda((1-\alpha)f+\alpha h)\big|_{\alpha=0} + \int_0^1\left[(1-\alpha)\frac{d^2}{d\alpha^2}\Lambda((1-\alpha)f+\alpha h)\right]d\alpha. \tag{3.11}$$

Since f maximizes $\Lambda(h)$, $(d/d\alpha)\Lambda((1-\alpha)f+\alpha h)|_{\alpha=0} = 0$. Therefore, from (3.10) and (3.11)

$$\frac{1}{2M_0^2}||h-f||^2 \le \Lambda(f) - \Lambda(h) \le \frac{1}{2}M_0^2||h-f||^2. \tag{3.12}$$

Fix $a > 4M_0^2$. For $g \in \mathbb{G}$ such that $||g-f|| = a\rho_n$,

$$\Lambda(f) - \Lambda(g) \ge \frac{1}{2M_0^2}||g-f||^2 = \frac{1}{2M_0^2}a^2\rho_n^2 > 8M_0^2\rho_n^2.$$

Since we have $\Lambda(f)-\Lambda(\bar{f}) \le \Lambda(f)-\Lambda(g^*) \le 8M_0\rho_n^2$ by (3.12), we must have $||f-\bar{f}|| \le a_1\rho_n$ by concavity of Λ. Hence, we have the required rate of convergence as $||f-\bar{f}|| = \mathcal{O}_p(K_n^{-3})$.

Next, we tackle the estimation error. The goal is to show that for g "far" from $\bar{f}$, $\ell(g;\boldsymbol{x},\boldsymbol{y})$ must be smaller than $\ell(\bar{f};\boldsymbol{x},\boldsymbol{y})$; then since we know ℓ is larger at $\widetilde{f}$, we must have that $\widetilde{f}$ is "close" to $\bar{f}$. Using the same ideas as in the proof of approximation error, we expand ℓ about $\bar{f}$:

27

$$\ell(g; \boldsymbol{x}, \boldsymbol{y}) = \ell(\bar{f}; \boldsymbol{x}, \boldsymbol{y}) + \frac{d}{d\alpha}\ell(\bar{f} + \alpha(g - \bar{f}); \boldsymbol{x}, \boldsymbol{y})\Big|_{\alpha=0} + \int_0^1 (1 - \alpha)\frac{d^2}{d\alpha^2}\ell(\bar{f} + \alpha(g - \bar{f}); \boldsymbol{x}, \boldsymbol{y})d\alpha.$$

$$\tag{3.13}$$

We consider $g_0 \in \mathbb{G}$ such that $M_0^{-1} \le g_0(x) \le M_0$ for $x \in [0, 1]$; then

$$-\frac{1}{2}M_0^2||g_0 - \bar{f}||^2 \le \frac{d^2}{d\alpha^2}\ell(\bar{f} + \alpha(g_0 - \bar{f})) \le -\frac{1}{2M_0^2}||g_0 - \bar{f}||^2 \tag{3.14}$$

using the same reasoning as for the approximation error. For the middle term of the right side of (3.13),

$$\frac{d}{d\alpha}\ell(\bar{f} + \alpha(g_0 - \bar{f}); \boldsymbol{x}, \boldsymbol{y})\Big|_{\alpha=0} = \frac{1}{n}\sum_{i=1}^n (g_0(x_i) - \bar{f}(x_i))\left[\frac{1}{\bar{f}(x_i)} - y_i^2\right].$$

Let $w_i = \bar{f}(x_i)^{-1} - y_i^2$, then

$$\frac{\sum_{i=1}^n (g_0(x_i) - \bar{f}(x_i))w_i}{\left[\sum_{i=1}^n (g_0(x_i) - \bar{f}(x_i))^2\right]^{1/2}} \le \sup_{g \in \mathbb{G}} \frac{\langle \boldsymbol{g}, \boldsymbol{w}\rangle_n}{||\boldsymbol{g}||} = ||\Pi(\boldsymbol{w}|\mathbb{G})||$$

where $\Pi(\boldsymbol{w}|\mathbb{G})$ is the projection of $\boldsymbol{w}$ onto the vector space $\mathbb{G} \subset \mathbb{R}^n$. We have $||\mathrm{E}(\boldsymbol{w})||_n = \mathcal{O}(K_n^{-3})$ and there is an M not depending on n such that $\mathrm{var}(w_i) \le M$. If $\boldsymbol{P}$ is the projection matrix for $\mathbb{G}$,

$$||\boldsymbol{P}\boldsymbol{w}||_n \le ||\boldsymbol{P}(\boldsymbol{w} - \mathrm{E}(\boldsymbol{w}))||_n + ||\boldsymbol{P}\mathrm{E}(\boldsymbol{w})||_n = \mathcal{O}_p((K_n/n)^{1/2}) + \mathcal{O}(K_n^{-3}).$$

Because K_n increases as $n^{1/7}$, the two rates are $n^{-3/7}$; therefore, there is an a_0 not depending on n so that

$$\left|\frac{d}{d\alpha}\ell(\bar{f} + \alpha(g - \bar{f}); \boldsymbol{x}, \boldsymbol{y})\Big|_{\alpha=0}\right| \le a_0 n^{-3/7}||g - \bar{f}||. \tag{3.15}$$

Choose $a_1 > 2a_0 M_0^2$, and consider $g \in \mathbb{G}$ such that $||g - \bar{f}|| = a_1 n^{-3/7}$. Then combining (3.13), (3.14), and (3.15), we have $\ell(g; \boldsymbol{x}, \boldsymbol{y}) \le \ell(\bar{f}; \boldsymbol{x}, \boldsymbol{y})$. Because $\ell(\widetilde{f}; \boldsymbol{x}, \boldsymbol{y}) \ge \ell(\bar{f}; \boldsymbol{x}, \boldsymbol{y})$, and by concavity of ℓ, we have $||\widetilde{f} - \bar{f}|| \le a_1 n^{-3/7}$.

The estimation error rate of convergence for the constrained estimator $\widehat{f}$ can be derived using the rate of convergence for $\widetilde{f}$. If $\bar{f} \in \mathbb{C}$, then we can use the same argument, because $\ell(\widehat{f}; \boldsymbol{x}, \boldsymbol{y}) \ge$

$\ell(\bar{f}; \boldsymbol{x}, \boldsymbol{y})$. For the case where $\bar{f}$ is not in $\mathbb{C}$, from Theorem 1 in Meyer et al. (2018), there is a function $f^* \in \mathbb{C}$ such that $||\bar{f} - f^*||_\infty = \mathcal{O}(K_n^{-6})$, and $\ell(\widehat{f}; \boldsymbol{x}, \boldsymbol{y}) \geq \ell(f^*; \boldsymbol{x}, \boldsymbol{y})$. Expanding ℓ around $\widetilde{f}$, we have

$$\ell(g) = \ell(\widetilde{f}) + \frac{d}{d\alpha}\ell(\widetilde{f} + \alpha(g - \widetilde{f}))\Big|_{\alpha=0} + \int_0^1 (1 - \alpha)\frac{d^2}{d\alpha^2}\ell(\widetilde{f} + \alpha(g - \widetilde{f}))d\alpha, \ g \in \mathbb{G}. \quad (3.16)$$

Since $\widetilde{f}$ maximizes $\ell(\cdot)$ in $\mathbb{G}$, we have $(d/d\alpha)\ell(\widetilde{f} + \alpha(g - \widetilde{f}))|_{\alpha=0} = 0$ and for g such that $M_0^{-1} \leq g(x) \leq M_0$ for $x \in [0, 1]$, we have

$$\frac{2}{M_0^2}||\widetilde{f} - g||^2 \leq \ell(\widetilde{f}) - \ell(g) \leq 2M_0^2||\widetilde{f} - g||^2.$$

Therefore, we have

$$||\widetilde{f} - \widehat{f}||^2 \leq \frac{1}{2}M_0^2[\ell(\widetilde{f}) - \ell(\widehat{f})] \leq \frac{1}{2}M_0^2[\ell(\widetilde{f}) - \ell(f^*)] \leq M_0^4||\widetilde{f} - f^*||^2.$$

Finally, $||\widetilde{f} - f^*|| \leq ||\widetilde{f} - \bar{f}|| + ||\bar{f} - f^*|| = \mathcal{O}_p(K_n^{-3})$, and again by the triangle inequality we have $||f - \widehat{f}|| = \mathcal{O}_p(K_n^{-3})$.

Similarly, the rate of convergence for the model with double-exponential errors can be derived, and it is the same as that for the model with normal errors. For the double-exponential errors, the likelihood function in (3.7) can be replaced by

$$l(h; x, y) = \log\big(h(x)\big) - \sqrt{2}|y|h(x),$$

which is concave in h for each value x and y. The expected log-likelihood defined in (3.8) is replaced by

$$\Lambda(h) = E[l(h; X, Y)] = \int_0^1 \left[\log\big(h(x)\big) - \frac{h(x)}{2f(x)}\right]p_X(x) \, dx, \quad (3.17)$$

where $p_X(x)$ is a probability density function. From (3.17), we have

$$\Lambda((1-\alpha)f + \alpha h) = \int_0^1 \left[\log\left((1-\alpha)f(x) + \alpha h(x)\right) - \frac{(1-\alpha)f(x) + \alpha h(x)}{2f(x)} \right] p_X(x) dx,$$

the second derivative of $\Lambda((1-\alpha)f + \alpha h)$ is the same as (3.9).

By following the same steps as mentioned above, we have the same optimal rate of convergence for approximation error.

For the estimation error, we have

$$\frac{d}{d\alpha} \ell(\bar{f} + \alpha(g_0 - \bar{f}); \boldsymbol{x}, \boldsymbol{y})\Big|_{\alpha=0} = \frac{1}{n} \sum_{i=1}^{n} (g_0(x_i) - \bar{f}(x_i)) \left[\frac{1}{\bar{f}(x_i)} - \sqrt{2}|y_i| \right].$$

Let $w_i = \bar{f}(x_i)^{-1} - \sqrt{2}|y_i|$, then we have $||E(\boldsymbol{w})||_n = \mathcal{O}(K_n^{-3})$ and there exists a positive constant M such that $\text{var}(w_i) \leq M$. The remainder of the proof follows the same steps as that of the normal error model, and we have the same rate of convergence for estimation error.

$\diamondsuit$

The maximum likelihood estimator is $\widehat{\sigma}^2(x) = \widehat{f}(x)^{-1}$ for the normal errors assumption, and is $\widehat{\sigma}^2(x) = \widehat{f}(x)^{-2}$ for the double-exponential errors assumption; either is one-to-one function on $[0,1]$. The rate of convergence of the variance function is $||\sigma^2 - \widehat{\sigma}^2|| = \mathcal{O}_p(n^{-3/7})$, from Remark 2.2 in Huang (2001) and Assumption (A1). The optimal rate of convergence for the proposed monotone estimator is maintained when the mean function is unknown, if the rate of convergence of the estimated mean function is fast enough.

Theorem 4. *Suppose that Assumptions (A1) through (A3) hold. If $\widehat{f}_{\hat{\mu}}$ is the MLE of f in $\mathbb{C}$ when μ is unknown and estimated with $\widehat{\mu}$, for which $||\widehat{\mu} - \mu|| = \mathcal{O}_p(n^{-3/7})$ then $||f - \widehat{f}_{\hat{\mu}}|| = \mathcal{O}_p(n^{-3/7})$.*

Proof. When the mean function is unknown, the proof of the optimal rate of convergence for the variance function is similar to the proof of Theorem 3.

For the normal errors, we can rewrite the normalized log-likelihood as

$$\ell_{\hat{\mu}}(f; \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2n} \sum_{i=1}^{n} \left(\log f(x_i) - (y_i - \widehat{\mu}_i)^2 f(x_i) \right),$$

and

$$\frac{d}{d\alpha}\ell_{\hat{\mu}}(\bar{f}+\alpha g;\boldsymbol{x},\boldsymbol{y})\Big|_{\alpha=0} = \frac{1}{2n}\sum_{i=1}^{n}\left[\frac{g(x_i)}{\bar{f}(x_i)}-\left(y_i-\widehat{\mu}_i\right)^2 g(x_i)\right]=\frac{1}{2}\Big\langle\tilde{\boldsymbol{w}},\boldsymbol{g}\Big\rangle_n,$$

where $\tilde{w}_i=\bar{f}(x_i)^{-1}-(y_i-\widehat{\mu}_i)^2$. For $w_i=\bar{f}(x_i)^{-1}-(y_i-\mu_i)^2$, we have

$$
\begin{aligned}
||E(\tilde{\boldsymbol{w}})||_n^2 &= n^{-1}\sum_{i=1}^{n}\left[E(\tilde{w}_i)\right]^2 \\
&= n^{-1}\sum_{i=1}^{n}\left[E\big(\bar{f}(x_i)^{-1}-(y_i-\widehat{\mu}_i)^2\big)\right]^2 \\
&= n^{-1}\sum_{i=1}^{n}\left[E\big(\bar{f}(x_i)^{-1}-(y_i-\mu_i)^2\big)-E\big(\mu_i-\widehat{\mu}_i\big)^2\right]^2 \\
&= n^{-1}\sum_{i=1}^{n}\left[E(w_i)-E(\mu_i-\widehat{\mu}_i)^2\right]^2 \\
&= n^{-1}\sum_{i=1}^{n}\left[E(w_i)\right]^2 + n^{-1}\sum_{i=1}^{n}\left[E\big(\mu_i-\widehat{\mu}_i\big)^2\right]^2 - 2n^{-1}\sum_{i=1}^{n}E(w_i)E\big(\mu_i-\widehat{\mu}_i\big)^2 \\
&= ||E(\boldsymbol{w})||_n^2 + ||E(\boldsymbol{\mu}-\widehat{\boldsymbol{\mu}})||_n^2 - 2n^{-1}\sum_{i=1}^{n}E(w_i)E\big(\mu_i-\widehat{\mu}_i\big)^2, \quad\quad (3.18)
\end{aligned}
$$

and $||E(\boldsymbol{w})||_n = \mathcal{O}(K_n^{-3})$ from the proof of Theorem 3. For $\widehat{\boldsymbol{\mu}}$ which satisfies $||E(\boldsymbol{\mu}-\widehat{\boldsymbol{\mu}})||_n = \mathcal{O}(K_n^{-3})$, we have $\left[\sum_{i=1}^{n}E(w_i)E\big(\mu_i-\widehat{\mu}_i\big)^2\right]^2 \leq \sum_{i=1}^{n}\left[E(w_i)\right]^2\sum_{i=1}^{n}E\left[\big(\mu_i-\widehat{\mu}_i\big)^2\right]^2$ by Cauchy-Schwarz inequality and $\left|n^{-1}\sum_{i=1}^{n}E(w_i)E\big(\mu_i-\widehat{\mu}_i\big)^2\right| = \mathcal{O}(K_n^{-6})$. From (3.18), we have $E(\tilde{\boldsymbol{w}}) = \mathcal{O}(K_n^{-3})$. Using the proof of Theorem 3 with y_i replaced by $y_i-\mu_i$, we find

$$\left|\frac{d}{d\alpha}\ell(\bar{f}+\alpha(g-\bar{f});\boldsymbol{x},\boldsymbol{y})\Big|_{\alpha=0}\right| \leq a_0 n^{-3/7}||g-\bar{f}||,$$

and the rest of the proof is identical to the proof of Theorem 3.

The same argument as for the model with normal errors will show the rate of convergence for the model with double-exponential errors. The normalized likelihood function can be written as

$$\ell_{\hat{\mu}}(f;\boldsymbol{x},\boldsymbol{y})=\frac{1}{n}\sum_{i=1}^{n}\Big(\log f(x_i)-\sqrt{2}|y_i-\widehat{\mu}_i|f(x_i)\Big),$$

31

and we have

$$\frac{d}{d\alpha}\ell_{\hat\mu}(\bar f + \alpha g; \boldsymbol{x}, \boldsymbol{y})\Big|_{\alpha=0} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{g(x_i)}{\bar f(x_i)} - \sqrt{2}|y_i - \widehat\mu_i|g(x_i)\right] = \frac{1}{2}\Big\langle \tilde{\boldsymbol{w}}, \boldsymbol{g}\Big\rangle_n,$$

where $\tilde w_i = \bar f(x_i)^{-1} - \sqrt{2}|y_i - \widehat\mu_i|$ and $||E(\tilde{\boldsymbol{w}})||_n = \mathcal{O}(K_n^{-3})$. From the proof of Theorem 3, we find

$$\left|\frac{d}{d\alpha}\ell(\bar f + \alpha(g - \bar f); \boldsymbol{x}, \boldsymbol{y})\Big|_{\alpha=0}\right| \le a_0 n^{-3/7}\|g - \bar f\|,$$

and the rest of proof is identical to proof of Theorem 3. $\diamondsuit$

3.5   Simulation Study

We compare the finite sample accuracy of the spline variance function estimators with the kernel estimators introduced by Ruppert et al. (1997) and Dette and Pilz (2009). Ruppert et al. (1997) proposed the variance function estimation method without shape constraints by smoothing of squared residuals using local polynomial smoother matrices. Dette and Pilz (2009) implemented the two-step kernel based method for estimating monotone variance function. Datasets were generated with different scenarios of standard deviation function, mean trend and sample size, using the model $y_i = \mu(i/n) + \sigma(i/n)\epsilon_i$ for $i = 1, \cdots, n$. For the standard deviation function $\sigma(\cdot)$, we consider following functions.

1. Convex: $\sigma_1(x) = .1 + x^2$

2. Sigmoid: $\sigma_2(x) = .1 + \frac{\exp(15x-8)}{1+\exp(15x-8)}$

The first standard deviation function is convex and increasing. The second standard deviation function is a sigmoid shape which increases rapidly in the middle, but slowly at the beginning and the end. For each standard deviation function, we used three error distributions, four sample sizes, and three mean functions. The errors are generated from following distributions:

1. $\epsilon^{[1]} \sim N(0, 1)$: Normal distribution with mean 0 and variance 1.

2. $\epsilon^{[2]} \sim DE(0, 2^{-1/2})$: Double-exponential distribution with mean 0 and variance 1.

3. $\epsilon^{[3]} \sim$ Contaminated normal distribution: 90% of errors are from $N(0, 0.65^2)$ and 10% of errors are from $N(0, 2.5^2)$.

Three mean trends were considered as below:

1. For $\mu_1(x) = 0$, we assume the mean trend is known.

2. For $\mu_2(x) = 3x + 5$, we assume the trend is a linear function.

3. For $\mu_3(x) = 50(x - .5)^4 I(x \geq 0.5)$, we assume the trend is smooth and increasing.

The sample sizes are n = 50, 100, 200, and 500, and for each scenario, M=10,000 data sets were generated. For the proposed spline methods, the numbers of knots for estimating the variance functions are 4, 5, 6 and 8 when n = 50, 100, 200, and 500 respectively.

For the data sets generated with known mean μ_1, the variance function is estimated directly for both spline and kernel methods. For μ_2 and μ_3 and the spline method, the mean and variance functions are estimated iteratively as described in Section 3.3.2. However, for the kernel methods, the mean function is estimated ignoring the heteroskedasticity, then the variance function is estimated using this estimated mean function. It is not feasible to iterate the estimation as in spline method, because the kernel method frequently yields negative values for the estimated variance function.

The slope and intercept for μ_2 are estimated using weighted least squares when normal errors are assumed and the weighted least absolute deviations method when double-exponential errors are assumed. The least-squares estimates of the smooth increasing μ_3 are obtained using the R package cgam (Liao and Meyer, 2017) when normal errors are assumed, and the function optim is used to obtain constrained spline least absolute deviations estimators when the errors are assumed to be double-exponential. In either case, the total number of knots for estimating the spline mean functions of μ_3 are 4, 5, 6, and 8 for n = 50, 100, 200, and 500, respectively. For both Ruppert and Dette methods, the simple linear regression method is used to estimate μ_2 for comparing the performance of estimated variance function. The kernel based method with the monotonicity constraint introduced by Dette et al. (2006) estimates μ_3 and the monotone variance function through an R package monreg.

To estimated mean and variance functions suggested in Ruppert et al. (1997), the matrices are defined as

$$X_p(x) = \begin{bmatrix} 1 & x_1 - x & \ldots & (x_1 - x)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x & \ldots & (x_n - x)^p \end{bmatrix} \text{ and } W_h(x) = \operatorname*{diag}_{1 \le i \le n} K\left(\frac{x_i - x}{h}\right),$$

where K is a pdf, and $\operatorname{diag}_{1 \le i \le n} a_i$ denotes the n by n diagonal matrix with $a_1, \ldots, a_n$ on the diagonal. Then, the (i, j) element of the local polynomial smoother matrix $S_{p,h}$ with degree p and bandwidth h is defined as

$$(S_{p,h})_{i,j} = e_1^T \{X_p(x_i)^T W_h(x_i) X_p(x_i)\}^{-1} X_p(x_i)^T W_h(x_i) e_j,$$

where $e_j \in \mathbb{R}^n$ is the column vector with one in the j the position and zeros elsewhere and $e_1 \in \mathbb{R}^p$ is the column vector with one in the the first position and zeros elsewhere. The mean function can be estimated as $\hat{\mu} = S_{p_1,h_1} y$ with the degree of polynomial p_1 and the bandwidth h_1. The estimated variance function at x_i with the degree p_2 and the bandwidth h_2 is obtained as

$$\widehat{v}(x_i) = \frac{e_1^T \{X_{p_2}(x_i)^T W_{h_2}(x_i) X_{p_2}(x_i)\}^{-1} X_{p_2}(x_i)^T W_{h_2}(x_i) r^2}{1 + e_1^T \{X_{p_2}(x_i)^T W_{h_2}(x_i) X_{p_2}(x_i)\}^{-1} X_{p_2}(x_i)^T W_{h_2}(x_i) \Delta} = \frac{S^*_{p_2,h_2}(x_i) r^2}{1 + S^*_{p_2,h_2}(x_i) \Delta},$$

where $r = y - \hat{\mu}$, $\Delta = \operatorname{diag}(S_{p_1,h_1} S^T_{p_1,h_1} - 2 S_{p_1,h_1})$, and

$$S^*_{p_2,h_2}(x_i) = e_1^T \{X_{p_2}(x_i)^T W_{h_2}(x_i) X_{p_2}(x_i)\}^{-1} X_{p_2}(x_i)^T W_{h_2}(x_i).$$

Let S_v be a $n \times b$ matrix where the ith row is $S^*_{p_2,h_2}(x_i)$, then the estimated variance can be written as $\widehat{v} = S_v r^2$. For the data sets with μ_2, the hat matrix of simple linear regression model is used for S_{p_1,h_1}. Further, the kernel method without the monotonicity constraints suggested in Ruppert et al. (1997) is used to estimate both μ_3 and the variance function. As suggested in Ruppert et al.

(1997), we set $p_1 = 3$ and $p_2 = 1$. The bandwidths are chosen based on Fan and Gijbels (1996) using the function `pluginBw` in an R package `locpol` (Cabrera, 2018).

To compare the performance of different variance etimation methods, the root mean squared error (RMSE) and the mean absolute error (MAE) are computed as follows:

$$\text{RMSE} = \left[\frac{1}{nM} \sum_{i=1}^{n} \sum_{j=1}^{M} \left(\widehat{\sigma}_j^2(x_i) - \sigma^2(x_i) \right)^2 \right]^{1/2},$$

and

$$\text{MAE} = \frac{1}{nM} \sum_{i=1}^{n} \sum_{j=1}^{M} \left| \widehat{\sigma}_j^2(x_i) - \sigma^2(x_i) \right|,$$

where $\widehat{\sigma}_j^2(x_i)$ is the estimated variance at x_i for the jth data set. Simulation results of RMSE and MAE are shown in Table 3.1. Ruppert and Dette columns show results for the kernel estimator without the shape constraints and the kernel estimator with the shape constraints, respectively. Even though the kernel estimator with the shape constraint tends to have smaller RMSE for smaller n, its MAE is always comparable or larger than for the spline estimator. It happens that a few data sets estimate the reciprocal of the variance function to be close to zero at $x = 1$, and for these few data sets $(\hat{\sigma}^2(1) - \sigma^2(1))^2$ can be very large. In addition, if we compare $\hat{f}$ with f, the spline estimator always has smaller RMSE and MAE. Moreover, the kernel based methods have the serious disadvantage of yielding negative values of the variance.

Next, we turn to the more interesting question of the effect of the variance function on inference for the mean function. First, we simulated datasets using the linear mean trend μ_2 and computed confidence intervals for the slope. With the normal errors assumption, let $\widehat{\beta}_1$ be the estimated (least-squares) slope and let $\widehat{\sigma}^2(x_i)$ be the estimated variance at x_i. The 95% confidence intervals are computed as

$$\widehat{\beta}_1 \pm t_{(0.975)}(n - 2 - L_n)\sqrt{(X^\top W X)^{-1}_{[2,2]}}, \tag{3.19}$$

where the matrix $X = (\mathbf{1}, \boldsymbol{x})$ and $W = \text{diag}(\widehat{\sigma}^{-2}(x_1), \ldots, \widehat{\sigma}^{-2}(x_n))$ where $\mathbf{1} = (1, \ldots, 1)^\top \in \mathbb{R}^n$. Under the double-exponential error assumption, let $\widehat{\beta}_{D,1}$ be the estimated (least absolute devia-

Table 3.1: RMSE (MAE) from 10,000 estimated variance functions of the proposed spline estimators and the kernel estimators.

Error	n	Spline N(0,1)	Spline DE	Kernel Ruppert	Kernel Dette	Spline N(0,1)	Spline DE	Kernel Ruppert	Kernel Dette
		σ1 : Convex				σ2: Sigmoid			
		$\mu_1 = 0$							
NM ($\epsilon^{[1]}$)	50	.27 (.10)	.45 (.15)	.31 (.13)	.23 (.12)	.31 (.16)	.52 (.23)	.40 (.20)	.31 (.22)
	100	.19 (.07)	.32 (.12)	.21 (.09)	.16 (.09)	.23 (.12)	.41 (.20)	.28 (.15)	.24 (.17)
	200	.13 (.06)	.23 (.10)	.15 (.06)	.11 (.07)	.16 (.09)	.31 (.16)	.20 (.11)	.18 (.14)
	500	.09 (.04)	.18 (.08)	.10 (.04)	.08 (.04)	.11 (.06)	.25 (.15)	.13 (.07)	.13 (.10)
DE ($\epsilon^{[2]}$)	50	.43 (.14)	.55 (.14)	.45 (.17)	.32 (.16)	.50 (.22)	.55 (.22)	.59 (.28)	.44 (.28)
	100	.36 (.11)	.35 (.11)	.36 (.13)	.25 (.13)	.40 (.18)	.37 (.17)	.45 (.21)	.34 (.23)
	200	.24 (.08)	.22 (.08)	.24 (.09)	.17 (.10)	.27 (.13)	.25 (.12)	.31 (.16)	.25 (.18)
	500	.14 (.06)	.13 (.05)	.15 (.06)	.11 (.06)	.17 (.09)	.16 (.08)	.19 (.10)	.18 (.13)
CN ($\epsilon^{[3]}$)	50	.72 (.18)	.80 (.15)	.70 (.22)	.48 (.21)	.82 (.30)	.70 (.23)	.89 (.36)	.63 (.36)
	100	.52 (.14)	.41 (.11)	.50 (.17)	.33 (.17)	.59 (.24)	.39 (.17)	.64 (.28)	.45 (.29)
	200	.39 (.11)	.26 (.08)	.36 (.13)	.24 (.13)	.42 (.17)	.27 (.13)	.46 (.22)	.34 (.23)
	500	.22 (.08)	.14 (.06)	.22 (.09)	.16 (.09)	.25 (.12)	.17 (.09)	.29 (.15)	.23 (.17)
		$\mu_2 = 3x + 5$							
NM ($\epsilon^{[1]}$)	50	.28 (.10)	.45 (.15)	.30 (.13)	.25 (.18)	.31 (.16)	.51 (.23)	.39 (.20)	.36 (.27)
	100	.19 (.07)	.31 (.12)	.21 (.09)	.19 (.15)	.23 (.12)	.40 (.20)	.28 (.15)	.29 (.22)
	200	.13 (.06)	.23 (.10)	.15 (.06)	.16 (.12)	.16 (.09)	.31 (.16)	.20 (.11)	.24 (.19)
	500	.09 (.04)	.18 (.08)	.10 (.04)	.12 (.10)	.11 (.06)	.25 (.15)	.13 (.07)	.19 (.16)
DE ($\epsilon^{[2]}$)	50	.43 (.14)	.56 (.15)	.44 (.17)	.30 (.20)	.50 (.22)	.55 (.22)	.58 (.28)	.44 (.31)
	100	.36 (.11)	.35 (.11)	.35 (.13)	.24 (.17)	.40 (.18)	.37 (.17)	.45 (.21)	.35 (.26)
	200	.24 (.08)	.22 (.08)	.24 (.09)	.19 (.14)	.27 (.13)	.25 (.12)	.31 (.16)	.28 (.22)
	500	.14 (.06)	.13 (.05)	.15 (.06)	.14 (.11)	.17 (.09)	.16 (.08)	.19 (.10)	.22 (.17)
CN ($\epsilon^{[3]}$)	50	.73 (.18)	.81 (.15)	.67 (.21)	.40 (.23)	.83 (.30)	.71 (.23)	.86 (.38)	.58 (.37)
	100	.52 (.14)	.41 (.11)	.49 (.17)	.30 (.19)	.59 (.24)	.39 (.17)	.63 (.28)	.44 (.30)
	200	.39 (.11)	.26 (.08)	.36 (.13)	.23 (.16)	.42 (.17)	.27 (.13)	.45 (.21)	.34 (.25)
	500	.22 (.08)	.14 (.06)	.22 (.09)	.17 (.12)	.25 (.12)	.17 (.09)	.28 (.15)	.26 (.20)
		$\mu_3 = 50(x - .5)^4 I(x \geq 0.5)$							
NM ($\epsilon^{[1]}$)	50	.36 (.11)	.36 (.12)	.37 (.14)	.24 (.16)	.35 (.16)	.40 (.18)	.48 (.23)	.33 (.24)
	100	.20 (.08)	.25 (.10)	.23 (.09)	.18 (.14)	.22 (.12)	.32 (.16)	.29 (.15)	.26 (.20)
	200	.13 (.06)	.19 (.08)	.15 (.07)	.14 (.11)	.16 (.09)	.26 (.14)	.20 (.11)	.21 (.17)
	500	.09 (.04)	.15 (.08)	.10 (.04)	.10 (.08)	.11 (.06)	.23 (.13)	.13 (.07)	.16 (.13)
DE ($\epsilon^{[2]}$)	50	.49 (.14)	.52 (.13)	.50 (.18)	.30 (.18)	.52 (.22)	.47 (.20)	.65 (.29)	.44 (.29)
	100	.35 (.11)	.31 (.10)	.36 (.13)	.24 (.16)	.38 (.17)	.32 (.16)	.46 (.21)	.34 (.24)
	200	.23 (.08)	.20 (.08)	.24 (.09)	.18 (.13)	.26 (.13)	.23 (.11)	.31 (.16)	.27 (.20)
	500	.14 (.06)	.12 (.05)	.15 (.06)	.13 (.09)	.17 (.09)	.15 (.08)	.19 (.10)	.19 (.15)
CN ($\epsilon^{[3]}$)	50	.75 (.18)	.77 (.15)	.75 (.23)	.41 (.22)	.82 (.29)	.66 (.22)	.97 (.37)	.59 (.36)
	100	.50 (.14)	.39 (.11)	.49 (.17)	.31 (.18)	.56 (.23)	.37 (.17)	.63 (.28)	.45 (.30)
	200	.38 (.11)	.24 (.08)	.35 (.13)	.24 (.15)	.40 (.17)	.26 (.13)	.45 (.22)	.34 (.24)
	500	.22 (.08)	.14 (.06)	.22 (.09)	.16 (.11)	.25 (.12)	.17 (.10)	.28 (.15)	.24 (.18)

tions) slope. Bassett and Koenker (1978) showed the asymptotic normality of the estimator that minimizes the sum of absolute errors for the linear model. Based on this, the confidence interval is obtained as

$$\widehat{\beta}_{D,1} \pm 1.96 \sqrt{(2\widehat{f}_\epsilon(0))^{-2}(X^\top W X)^{-1}_{[2,2]}}$$

where $\widehat{f}_\epsilon(0)$ is the kernel density estimation of the error distribution at $x = 0$ using a Gaussian kernel.

For the kernel method (Ruppert) without the monotonicity constraint, the estimated mean function of μ_2 is updated only once using weighted least squares method because of the negative values in the estimated variance function. The 95% confidence intervals are also computed as (3.19) with a trace of S_v as a replacement of L_n, after removing the data sets with negative estimated variance values. Table 3.2 shows the percentage of removed data sets out of 10,000 data sets due to the negative variance values for computing the confidence interval of the slope parameter. The kernel method (Dette) with the monotonicity constraint can not be used to compare the performance of the mean function because all the estimators of variance functions contain at least one negative value.

Table 3.2: The percentage of removed data sets to compare the estimated slope parameter of the linear mean function ($\mu_2 = 3x + 5$) for the kernel method without the shape constraint (Ruppert et al., 1997).

σ	n	NM ($\epsilon^{[1]}$)	DE ($\epsilon^{[2]}$)	CN($\epsilon^{[3]}$)
σ_1	50	11.15	18.20	21.43
	100	5.28	12.66	17.53
	200	1.54	5.61	11.82
	500	0.06	0.89	3.64
σ_2	50	9.92	17.59	22.60
	100	3.74	10.51	17.06
	200	0.79	3.90	10.22
	500	0.01	0.54	2.29

The coverage probabilities and widths of the confidence intervals are shown in Table 3.3 for the two standard deviation functions, three error distributions, and four sample sizes described for

Table 3.3: 95% coverage probability (width of CI) of the estimated slope parameter of the linear mean function ($\mu_2 = 3x + 5$). The OLS and LAE estimators and confidence intervals (incorrectly) assume homoskedasticity. The Kernel (Ruppert) method obtains estimators and confidence intervals without shape restrictions.

σ	Error	n	OLS	Spline N(0,1)	LAE	Spline DE	Kernel Ruppert
σ_1	NM ($\epsilon^{[1]}$)	50	.890 (.98)	.931 (.63)	.824 (.69)	.921 (.77)	.880 (.64)
		100	.891 (.71)	.937 (.45)	.820 (.48)	.933 (.57)	.919 (.46)
		200	.895 (.50)	.940 (.32)	.814 (.34)	.942 (.41)	.936 (.33)
		500	.896 (.32)	.948 (.21)	.805 (.21)	.953 (.26)	.950 (.21)
	DE ($\epsilon^{[2]}$)	50	.892 (.97)	.938 (.61)	.883 (.52)	.952 (.57)	.906 (.60)
		100	.895 (.70)	.940 (.44)	.895 (.36)	.965 (.40)	.930 (.44)
		200	.900 (.50)	.945 (.32)	.891 (.25)	.969 (.28)	.944 (.32)
		500	.894 (.32)	.947 (.20)	.870 (.15)	.969 (.17)	.951 (.21)
	CM ($\epsilon^{[3]}$)	50	.909 (.96)	.944 (.58)	.835 (.50)	.925 (.55)	.904 (.56)
		100	.903 (.69)	.942 (.42)	.837 (.35)	.935 (.40)	.926 (.42)
		200	.895 (.49)	.944 (.30)	.822 (.24)	.942 (.29)	.935 (.31)
		500	.890 (.32)	.945 (.20)	.813 (.15)	.946 (.18)	.949 (.20)
σ_2	NM ($\epsilon^{[1]}$)	50	.927 (1.32)	.935 (.73)	.738 (.72)	.922 (.84)	.892 (.77)
		100	.926 (.95)	.944 (.52)	.727 (.49)	.937 (.62)	.931 (.55)
		200	.927 (.68)	.941 (.36)	.708 (.34)	.943 (.44)	.948 (.39)
		500	.932 (.42)	.948 (.23)	.688 (.20)	.951 (.28)	.956 (.24)
	DE ($\epsilon^{[2]}$)	50	.928 (1.30)	.941 (.70)	.820 (.55)	.946 (.62)	.909 (.72)
		100	.931 (.94)	.946 (.50)	.819 (.38)	.962 (.44)	.935 (.52)
		200	.930 (.67)	.945 (.35)	.816 (.25)	.967 (.31)	.952 (.38)
		500	.929 (.43)	.948 (.23)	.780 (.15)	.967 (.19)	.959 (.24)
	CM ($\epsilon^{[3]}$)	50	.938 (1.29)	.946 (.67)	.765 (.53)	.924 (.59)	.914 (.68)
		100	.930 (.94)	.946 (.48)	.750 (.36)	.935 (.44)	.933 (.50)
		200	.932 (.67)	.946 (.34)	.729 (.25)	.941 (.31)	.943 (.37)
		500	.926 (.43)	.944 (.22)	.698 (.15)	.946 (.20)	.955 (.24)

the previous simulations. For each scenario, 10,000 data sets were generated. For each data set, confidence intervals were computed: the OLS and LAE columns show results for standard least-squares and least absolute deviations, where homoskedasticity is assumed. For normal errors, the confidence interval computed using the spline variance estimator as weights, assuming normality, performs best. With the heavier-tailed errors, the proposed method with the double-exponential

assumption outperforms the others. Although the method using splines with the assumption of normal errors has good coverage probability, the interval widths are larger.

The histograms of the estimated slope of β_1 are shown in Figures 3.1, 3.2, and 3.3 for normal errors, double-exponential errors, and contaminated normal errors, respectively. From the histograms, it can be seen that the range of estimate slope ignoring heteroskedasiticity is wider compare to the range of estimate slope incorportation heteroskedasiticity. Also, the distribution of estimated slope parameters from the kernel method based on Ruppert et al. (1997) is skewed.

We also simulate data sets with mean trend μ_3 and normal errors, to consider the effect of the variance function for inference for a non-parametrically estimated mean function. For each standard deviation function, we simulated 10,000 data sets with equally spaced x-values. For each data set, the mean and variance functions were estimated using monotone regression splines, and the point-wise confidence intervals at each x-value were computed using the method introduced by Meyer (2018). The coverage probabilities (top curves, left scale) and average interval widths (bottom curves, right scale) are shown in Figure 3.4. Black dashed lines and red dash-dot lines show the coverage probability and the mean interval length using the splines estimator and the kernel estimator (Ruppert), respectively. For the kernel estimator, 315 data sets for σ_1 and 154 data sets for σ_2 are removed before calculating the coverage probability and the mean interval length due to the negative values of estimated variances. At boundaries, the spline estimator has better coverage probabilities with a narrower interval length. Gray solid lines show the coverage probability and mean interval length when we ignore heteroskedasticity. It shows that the coverage probability of the mean function is substantially improved by accounting for the estimated variance function.
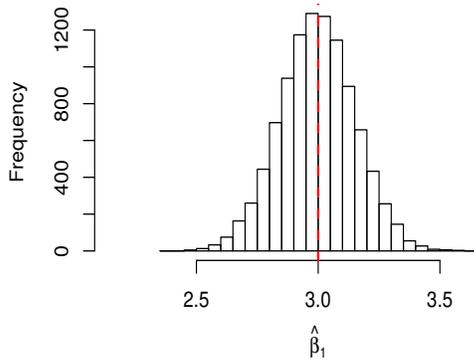
(a) OLS with σ_1 (convex).
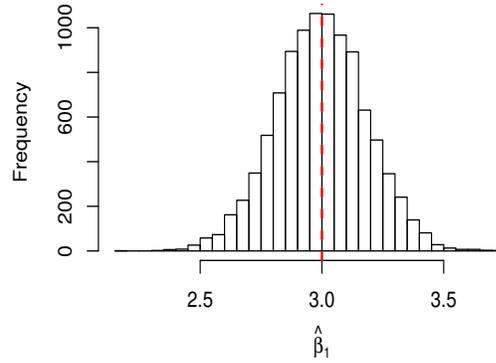
(b) OLS with σ_2 (sigmoid).
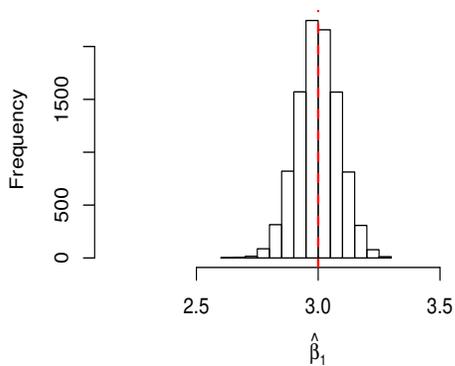
(c) Spline N(0,1) with σ_1 (convex).
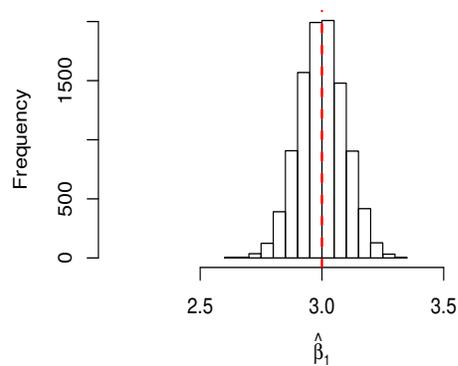
(d) Spline N(0,1) with σ_2 (sigmoid).
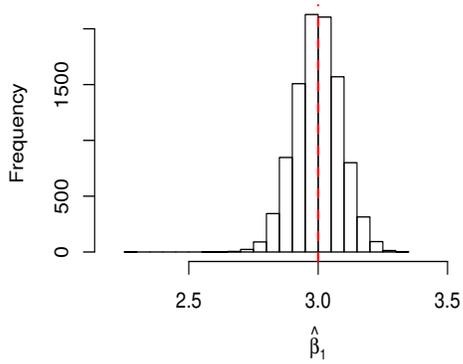
(e) Kernel (Ruppert) with σ_1 (convex).
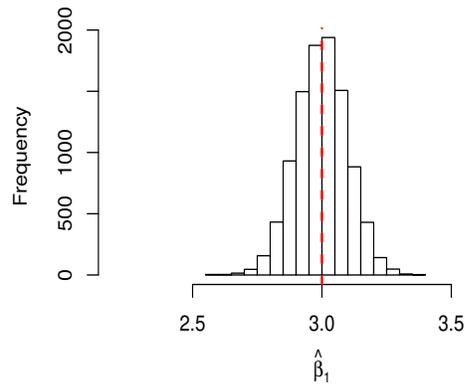
(f) Kernel (Ruppert) with σ_2 (sigmoid).

Figure 3.1: Histogram of estimated slope parameters of β_1 simulated with normal errors ($\epsilon^{[1]}$) and $n = 200$.
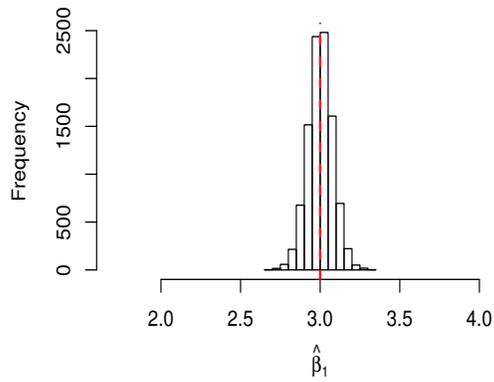
(a) LAE with σ_1 (convex).
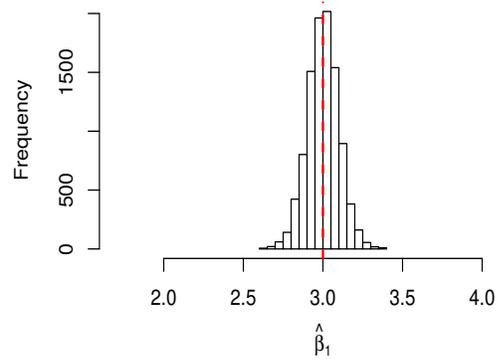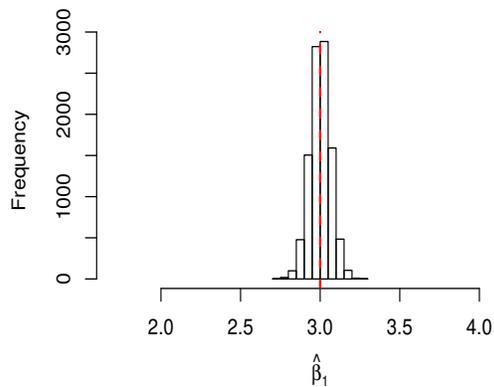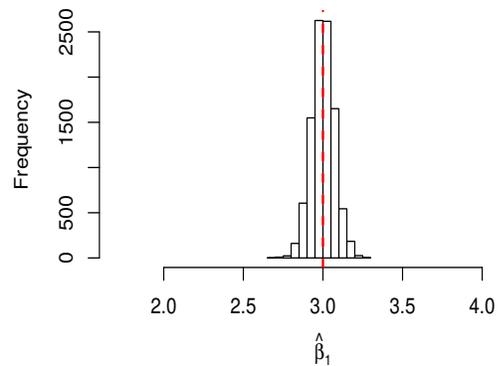
(b) LAE with σ_2 (sigmoid).

(c) Spline DE with σ_1 (convex).

(d) Spline DE with σ_2 (sigmoid).
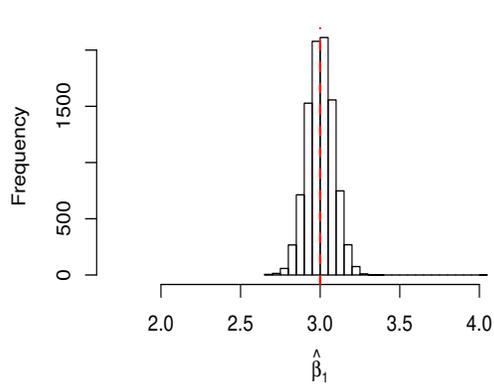
(e) Kernel (Ruppert) with σ_1 (convex).
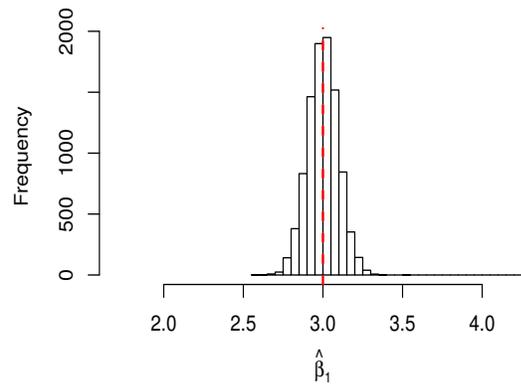
(f) Kernel (Ruppert) with σ_2 (sigmoid).

Figure 3.2: Histogram of estimated slope parameters of β_1 simulated with double-exponential errors ($\epsilon^{[2]}$) and $n = 200$.
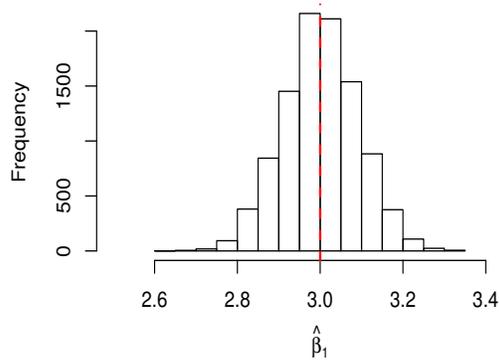
(a) LAE with σ_1 (convex).
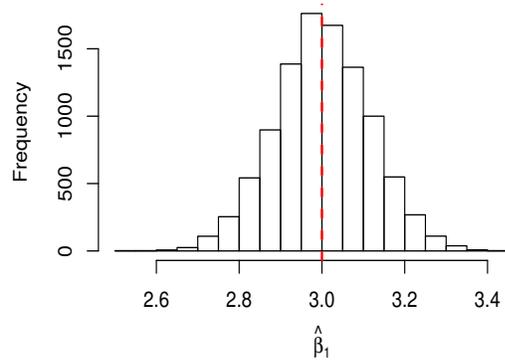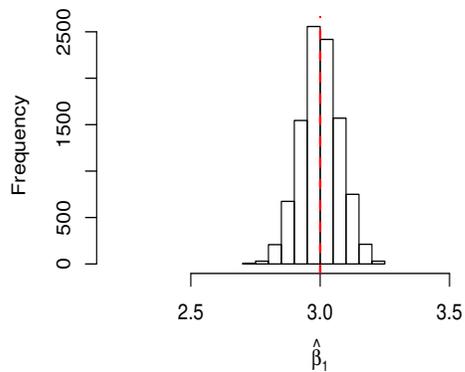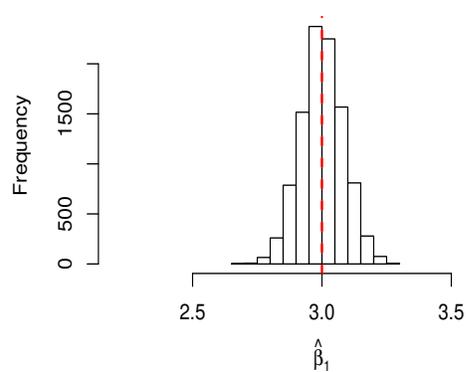
(b) LAE with σ_2 (sigmoid).

(c) Spline DE with σ_1 (convex).

(d) Spline DE with σ_2 (sigmoid).
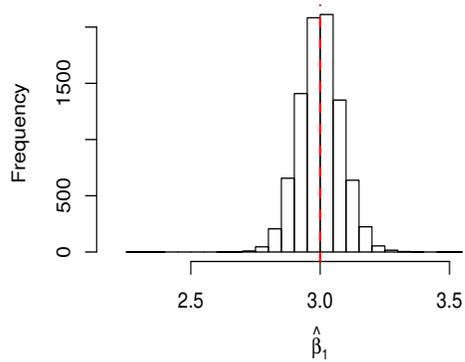
(e) Kernel (Ruppert) with σ_1 (convex).
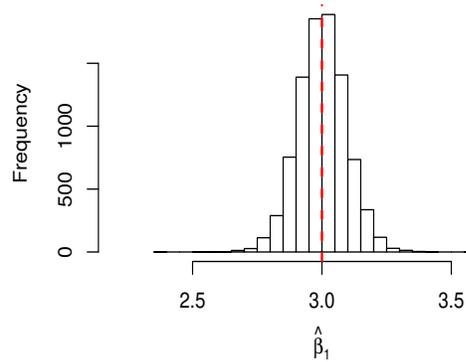
(f) Kernel (Ruppert) with σ_2 (sigmoid).

Figure 3.3: Histogram of estimated slope parameters of β_1 simulated with contaminated normal errors ($\epsilon^{[3]}$) and $n = 200$.
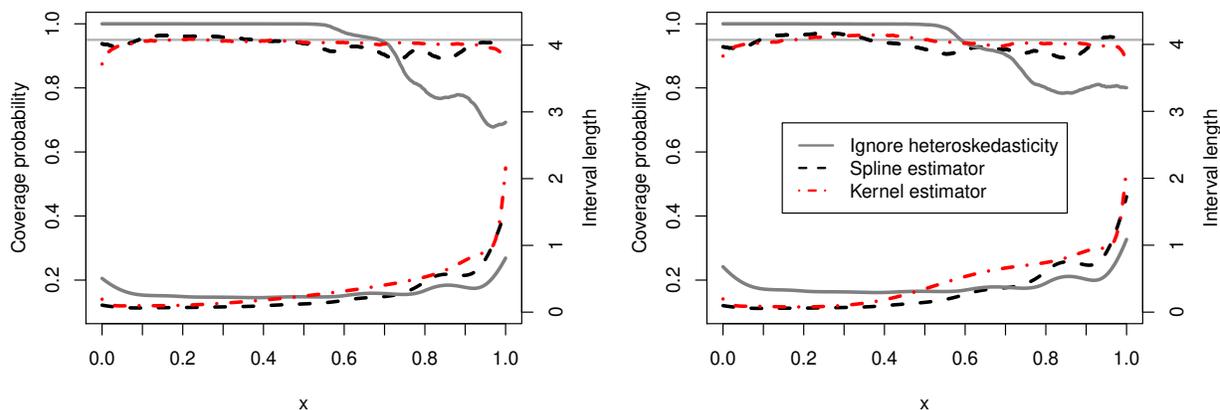
Figure 3.4: Simulation with $\mu_2 = 50(x - .5)^4 I(x \geq 0.5)$, $n = 200$, and normal errors. The results in the left use the true variance functions σ_1^2(convex), and for the right panel, σ_2^2(sigmoid) is used. Coverage probability (top curve, left scale) and average interval length (bottom curves, right scale) using spline variance estimator (dashed lines) and kernel variance estimator (dash-dot lines). Gray solid lines show the coverage probability and mean interval length ignoring heteroskedasticity.

3.6 Real Data Analysis

3.6.1 LIDAR Data

The LIDAR data set described in Ruppert et al. (2003) contains 221 observations from a light detection and ranging (LIDAR) experiment. The LIDAR experiment provides information about the shape of the surface of the source. The response variable is a logarithm of the ratio of received light from two laser sources, and the predictor variable is the distance traveled before the light is reflected back to its source. It is reasonable to model the mean function as smooth and decreasing, and the variance function as smooth and increasing. Figure 3.5 shows the estimated mean function (solid line) and the 95% confidence bands (dashed lines) from (a) the proposed spline method when normal errors are assumed, and from (b) the kernel method (Ruppert) without the monotonicity constraint. For the spline method, the R package `cgam` is used for the mean and confidence bands with six evenly spaced knots marked as 'x'. The suggested bandwidth for the kernel method (Ruppert) is computed through an R function `pluginBw` in `locpol` package with the Gaussian kernel. There are spurious bumps and wiggles in the estimated mean and bands from the kernel

method (Ruppert) after an range of 600. The squared residuals from the spline fit in (a) are shown in Panel (c). The squared residuals should be randomly and evenly scattered if the homoskedasticity assumption is satisfied. However, panel (c) apparently shows that the squared residual increases as the predictor variable increases, and the homoskedasticity assumption is violated. Panel (d) shows the estimated variance functions from the kernel method (Ruppert) without the monotonicity constraint, and the solid line shows the estimated variance function with suggested bandwidth. The curves are shown to have wiggles after an range of 600 even with a large bandwidth. The estimated variance functions from the proposed splines method under the normal errors assumption and the kernel method (Dette) with the monotonicity constraint are shown in Panels (e) and (f), respectively, showing that the spline estimator is more robust to tuning parameters. Since the kernel method (Dette) yields the negative values for estimated variance function when the range is small, it is not able to construct the confidence intervals of the mean function.

To confirm our assumption about the error distribution, we look at the normal quantile-quantile plot using the weighted residuals, shown at the left in Figure 3.6. When the mean and variance functions are estimated using the double-exponential errors assumption, the quantile plot of the weighted residuals (on the right) shows deviations from this assumption. Therefore, it is appropriate to assume that errors follow the normal distribution.

3.6.2   Abalone Data

The Abalone data set, available in an R package `PivotalR`, was used by Dette and Pilz (2009) to illustrate the kernel method, assuming that that interest is in estimating physical measurements as a function of age (indicated by the number of rings). The physical measurements of each abalone in the data set consist of the length, diameter, height, whole weight, shucked weight, viscera weight, and shell weight. There are 4177 observations, and we remove 21 obvious erroneous entries (samples which have the smaller whole weight compared to the sum of shell weight and shucked weight or the smaller length than diameter or with zero height). The first principal component of physical measurements accounts for 97.53% of the variability in the physical measurements. The log trans-

(a)

(b)

(c)

(d)

(e)

(f)
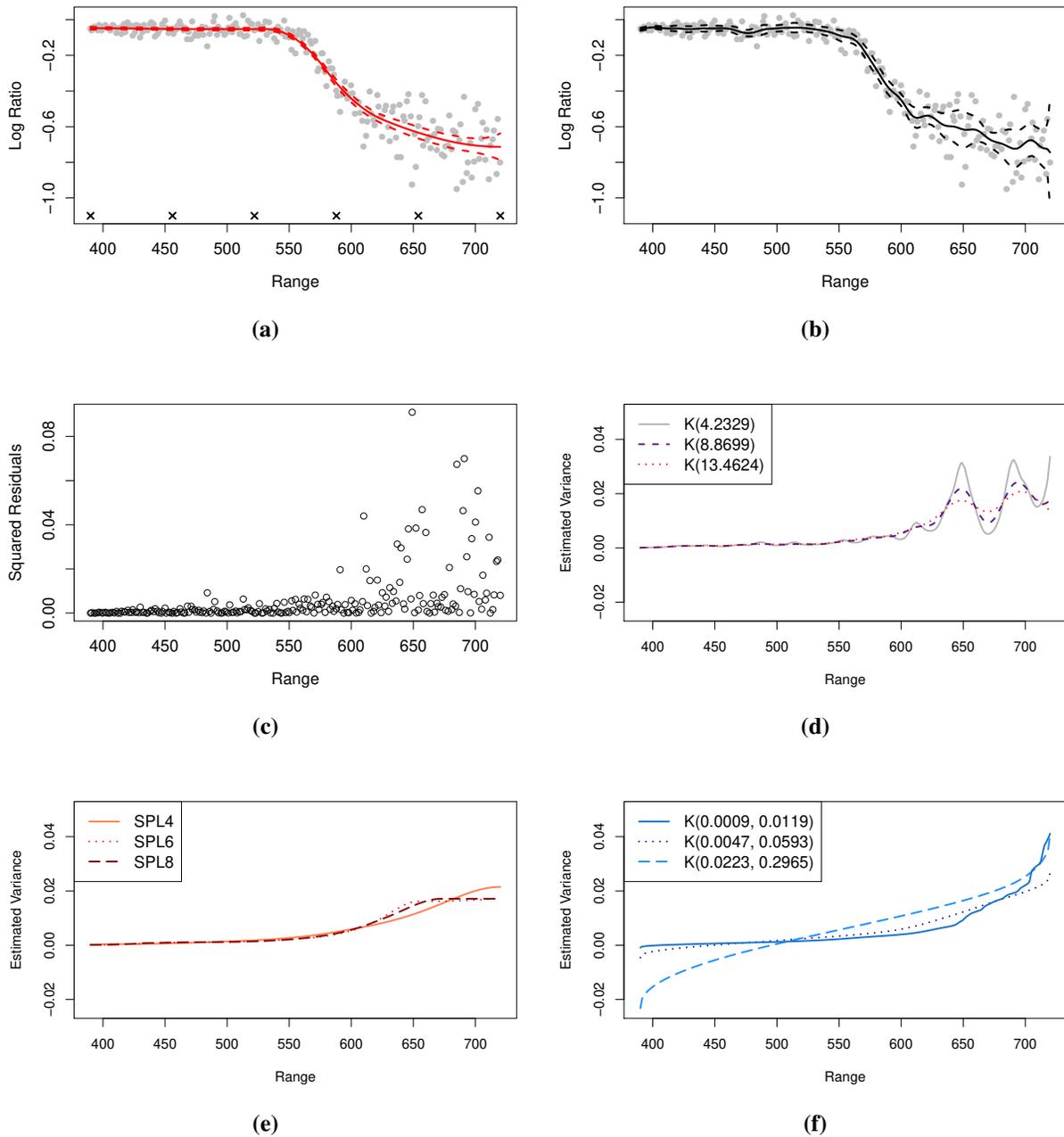
Figure 3.5: LIDAR data. (a) The scatter plot of LIDAR data with the estimated mean function (solid line) and point-wise confidence bands (dashed line) from the proposed spline method. Knots are marked as 'x'. (b) The fit (solid line) with point-wise confidence bands (dashed line) from the kernel method (Ruppert et al., 1997). (c) Squared residuals; squared value of difference between observed value and estimated mean value in (a). (d) The estimated variance using the kernel method (Ruppert et al., 1997) with different bandwidths. (e) Estimated variance functions using the proposed splines method with different number of knots. (f) The estimated variance using the kernel method (Dette and Pilz, 2009) with three different combinations of bandwidths.
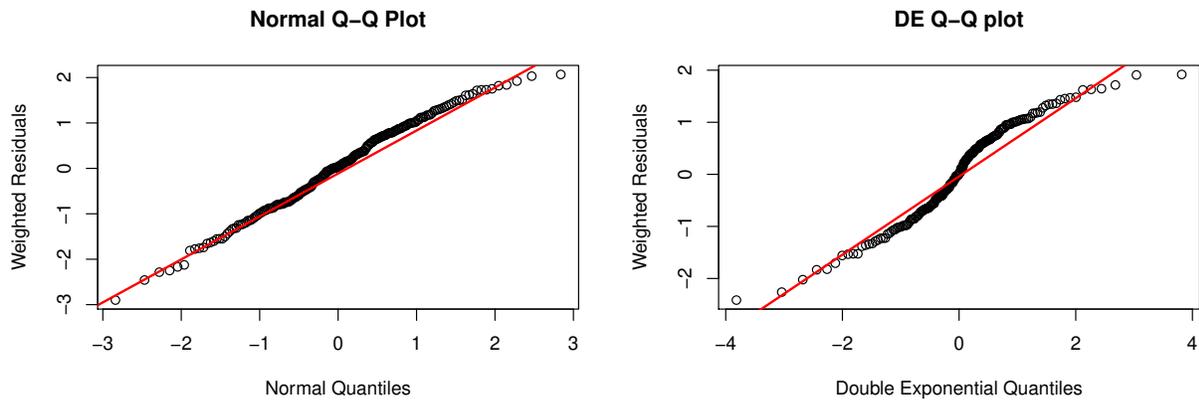
Figure 3.6: Quantile plots of the weighted residuals from the spline fit to the LIDAR data. The residuals follow a normal distribution (left) much more closely than a double-exponential distribution (right).

formation of the first principal components after adding a constant is used as a predictor variable. The coefficients are all positive and approximately equal, indicating that this linear combination of measurements can be labeled "overall size" of the abalone. It seems reasonable to assume that overall size is increasing in age, and that the variance of the size tends to increase as well. Panel (a) in Figure 3.7 shows the scatter plot and the estimated mean function (solid line) with the point-wise 95% confidence bands (dashed line) based on the splines method. Again, cgam is used to obtain the mean function and the bands with evenly spaced eleven knots. Knots are marked as 'x'. Panel (b) shows the fit and bands form the kernel method (Ruppert). The confidence interval in (b) is wider than the confidence interval in (a) at the boundaries. Panel (c) shows the squared residuals, squared values of difference between the first principal components of physical measurements and its estimated value, calculated by using the estimated mean values in (a). It shows that the error assumption of homoskedasticity is violated. Panel (d) shows the estimated variance functions from the kernel method (Ruppert) without the monotonicity constraint, and the dashed line shows the estimated variance function with suggested bandwidth. The estimated variance function varies after the age of 25 based on the bandwidths. With the bandwidths 0.6480 and 1.1255, the kernel method (Ruppert) yields the negative values for the estimated variance function. Therefore, the bandwidth 0.5815 is used to construct the confidence bands in (b) even though it is not the suggested bandwidth. The estimated variance functions using the proposed spline method and the

kernel method from Dette and Pilz (2009) are shown in Panel (e) and Panel (f), respectively. Three different number of equally spaced knots are used as knots for the proposed spline method and the different choices of bandwidths are used for the kernel method (Dette). The variance functions estimated by the kernel method (Dette) vary more with the choice of bandwidths, and they violate the assumption that the variances are positive. The solid curve with bandwidths (.003, .0045) uses the recommended bandwidths from Dette and Pilz (2009).

3.6.3 Abalone Data with an Indicator Variable

The abalone dataset includes a categorical covariate `Sex` that has three levels: Male, Female and Infant. An abalone is labeled as Infant if it is not mature enough to categorize into either male or female. The scatter plot in Panel (a) of Figure 3.8 shows that infant abalones tend to have smaller mean and variance compared to adults (male and female) abalones if they are the same age. Let z be the indicator variable for adult abalones, that is, $z_i = 0$ for infants and $z_i = 1$ for adults for $i = 1, \ldots, n$. We model the reciprocal of the variance as $\theta_i = f(x_i) + \beta z_i$, where β is the coefficient for the indicator variable and $f(x)$ will be approximated by a non-increasing spline function. Along with the monotonicity constraints, we need an extra constraint that $f(1) + \beta > 0$ to satisfy the assumption of the positive variance. The estimated mean function is shown in Panel (a) of Figure 3.8, along with the 95% point-wise confidence bands from the proposed spline method under the normal errors assumption. Nine evenly spaced points over the range of age, a minimum age of adults abalones, and a maximum age of infants abalones are used as knots, and they are marked as 'x'. Both kernel methods do not have the option for incorporating the parametric covariate in the variance function. Panel (b) of Figure 3.8 shows the estimated variance functions with three different number of knot choices. It shows that the estimated variance function is robust to the tuning parameter. Panel (c) of Figure 3.8 shows the normal Q-Q Plot of standardized residuals, confirming that the normal errors assumption is reasonable compared to the double-exponential errors assumptions.
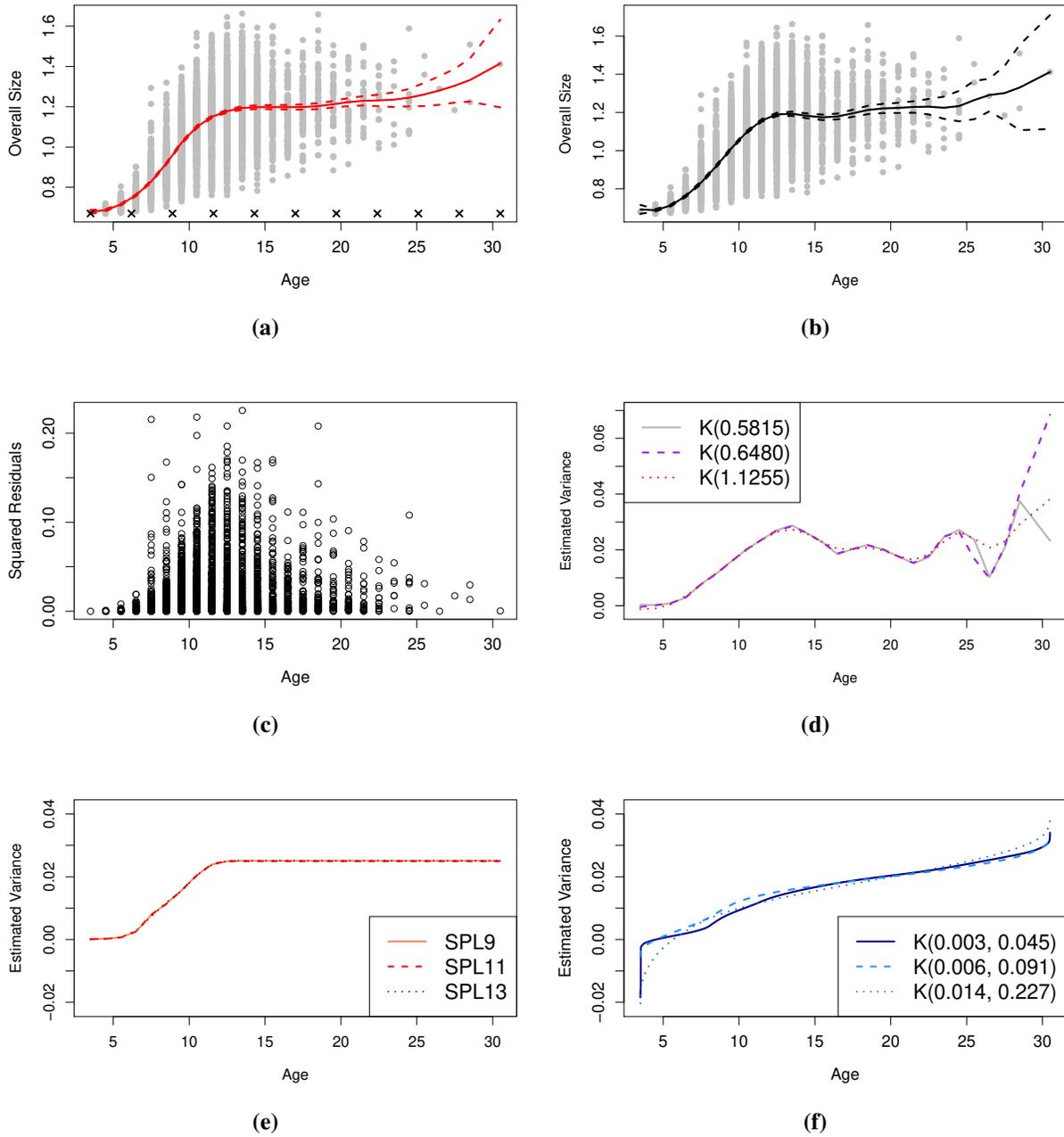
Figure 3.7: Abalone data. (a) The scatter plot of Abalone data with the estimated mean function (solid line) and point-wise confidence bands (dashed line) using `cgam` function in an R package. Knots are marked as 'x'. (b) The fit (solid line) with point-wise confidence bands (dashed line) from the kernel method (Ruppert et al., 1997). (c) Squared residuals; squared value of difference between observed value and estimated mean value in (a). (d) The estimated variance using the kernel method (Ruppert et al., 1997) with different bandwidths. (e) Estimated variance functions from the proposed splines method with 9, 11, 13 number of knots. (f) The estimated variance from the kernel method (Dette and Pilz, 2009) with three combinations of bandwidths.

48
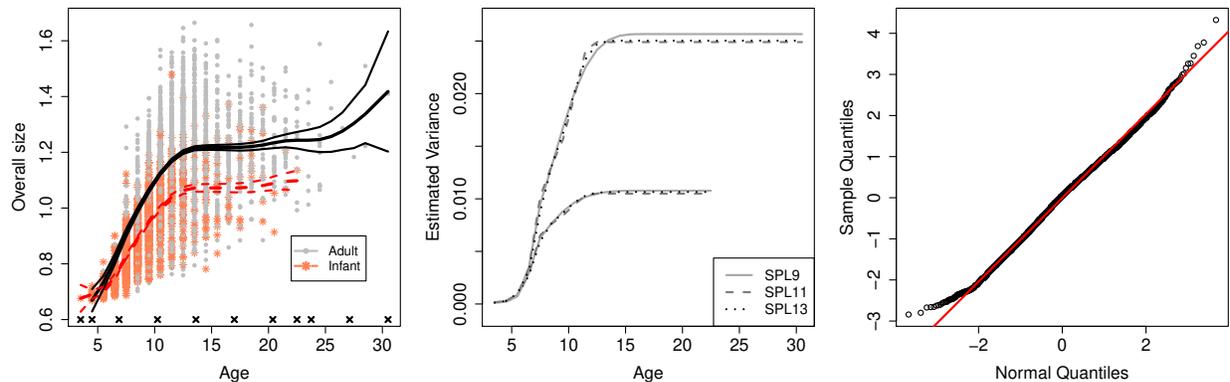
Figure 3.8: Abalone data with an infant indicator variable. (a) Scatter plot of Abalone data with the estimated mean and 95% confidence bands from the proposed spline method under the normal errors assumption. Knots are marked as 'x'. (b) The estimated variance functions from the proposed splines method with the different number of knots for adults and infants separately. (c) Normal Q-Q plot of standardized residuals.

3.6.4 California Air Pollution Data

California Air Pollution Data (Breiman and Friedman, 1985), available in a `SemiPar` package, consists of daily ozone level and other meteorological measurements for 345 days in California in 1976. We are interested in the relationship between air pollution (ozone) and the meteorological quantity as a function of temperature. The response variable is a Ozone concentration which records the daily maximum one hour average ozone concentration (ppm) at Sandburg Air Force Base. The predictor variable is the inversion base temperature in degrees Fahrenheit. Figure 3.9 shows the scatter plot with estimated mean function (solid line) with point-wise confidence bands (dashed line) form (a) the proposed spline method under the normal errors assumption, and (b) the kernel method (Ruppert et al., 1997) without the monotonicity constraints. Form the scatter plot, we can see that both mean and variance are increases as the inversion of temperature increases. The panels (a) and (b) show that the estimated mean and bands wiggle more for the kernel method (Ruppert) when the inversion base temperature is greater than 80. Panel (c) shows squared residuals using the fit in (a). It is clear that the homoskedasticity assumption for variance is violated because the squared residual tends to increase as the predictor variable increases. Figure 3.9 shows the estimated variance functions using the kernel methods ((d) Ruppert et al. (1997) and (f) Dette

and Pilz (2009)), and (e) our proposed spline method under the normal errors assumption. The solid line shows in Figure 3.9 (d) uses the suggested bandwidth introduced in Fan and Gijbels (1996). The estimated variance function with the largest bandwidth in Figure 3.9 (d) yields a negative value even though it is much smoother than the estimated variance functions with larger bandwidth. In Figure 3.9 (f), the estimated variance function using the kernel method (Dette) with the suggested bandwidths from is represented in the solid line. It is clear that the variance is not linearly increasing, so the estimated variance with the smaller bandwidths might be more convincing. Similar to the previous examples, Figure 3.9 (e) depicts that the estimated variance function from the proposed spline method is quite stable even though the number of knots is different. The estimated variances using four number of knots has lager estimates compare to the others when the inversion of temperature is greater than 80, but it became stable when we use the large enough number of knots. Moreover, the ambiguity might arises because of the small number of data when Temperature is greater than 80.

We also confirm that the error distribution follows the normal distribution from quantile-quantile plots of weighted residuals in Figure 3.10. The quantile plot of the weighted residuals (on the right) under the double-exponential errors assumption shows deviations from the linear line while the quantile plot of the weighted residuals (on the left) under the normal errors assumption shows the approximately linear line.
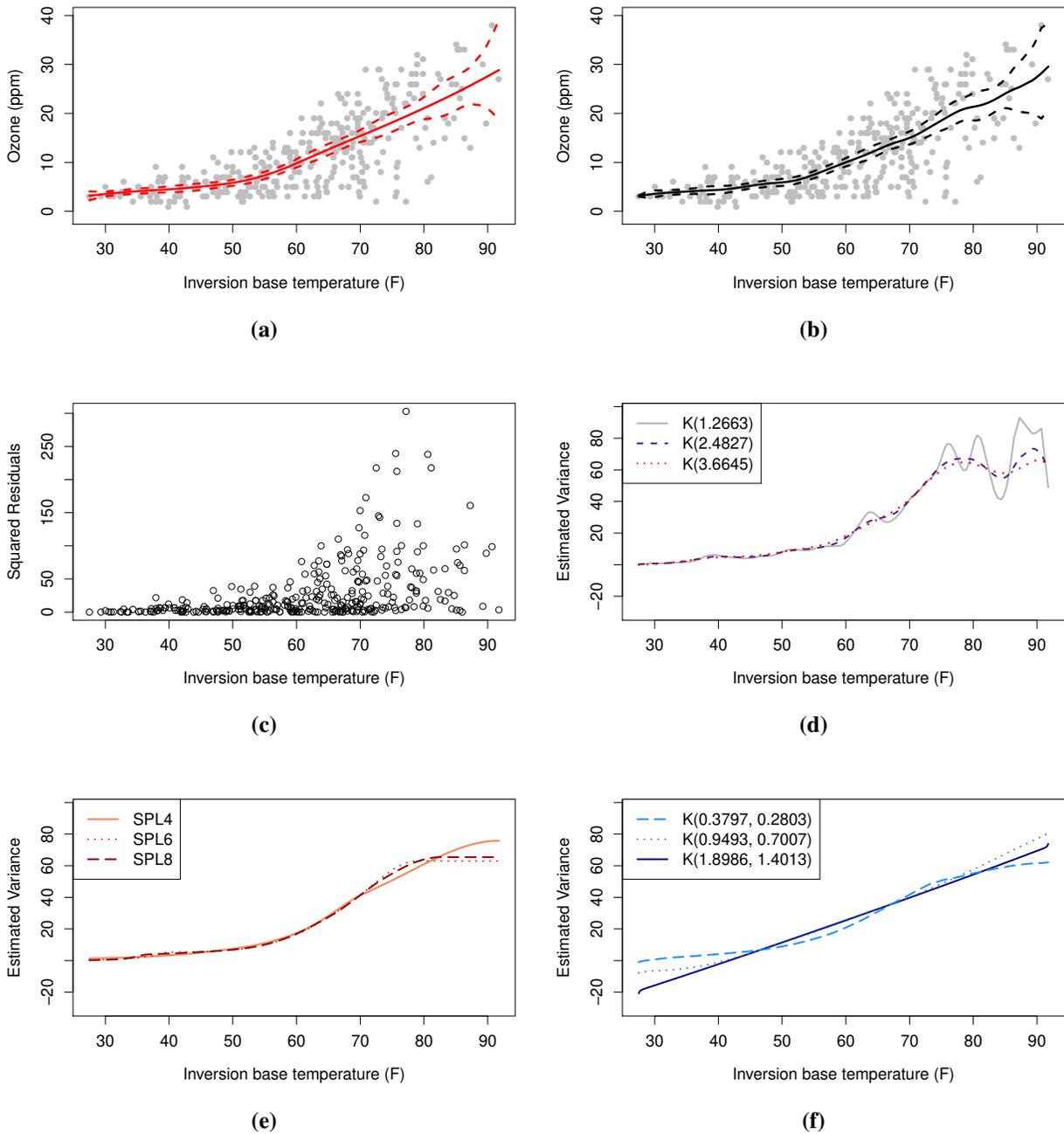
Figure 3.9: California air pollution data. (a) The scatter plot of California air quality data with estimated mean function (solid line) and point-wise confidence intervals (dashed line) from the proposed spline method under the normal errors assumption. (b) The estimated fit (solid line) with point-wise confidence bands (dashed line) from the kernel method (Ruppert et al., 1997). (c) Squared residuals; squared value of difference between observed value and estimated mean value in (a). (d) The estimated variance using local polynomial kernel method Ruppert et al. (1997) with three bandwidths. (e) Estimated variance functions using the proposed method with different number of knots. (f) The estimated variance using kernel method (Dette and Pilz, 2009) with three different combinations of bandwidths.
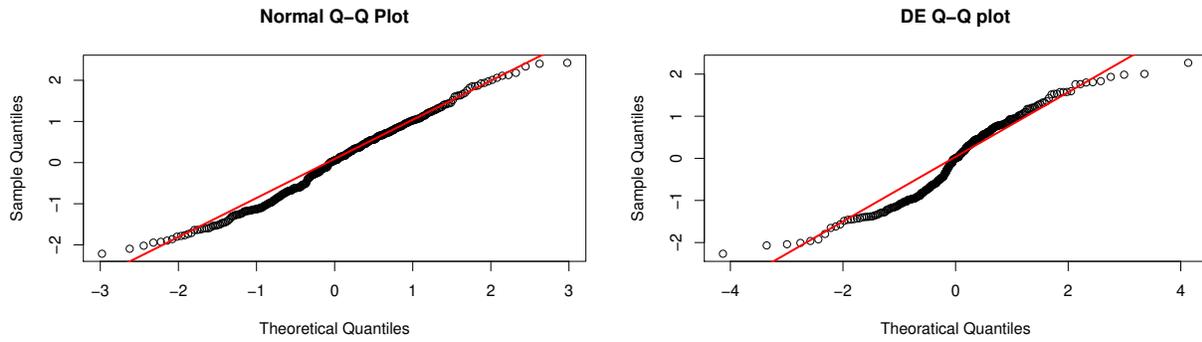
Figure 3.10: California air pollution data: Quantile plots of the weighted residuals from the spline fit. The residuals follow a normal distribution (left) much more closely than a double-exponential distribution (right).

Chapter 4

Monotone Variance Function Estimation with a Penalty Term

4.1 Motivation

The proposed splines estimators are consistent and converge to the true variance function as mentioned in Chapter 3. However, a problem arises at a boundary of domain for a few data sets in the simulation studies. The estimated non-decreasing variance function spikes when $\widehat{f}(1)$ is close to zero for a couple of unlucky data sets. If $\widehat{f}(1)$ is close to zero, the estimated variance at 1 (which is the reciprocal of $\widehat{f}(1)$) become larger. The left panel of Figure 4.1 shows the example of the unlucky data set generated with $\mu_1 \equiv 0$, $\sigma_1 = 0.1 + x^2$, $n = 200$, and normal errors. The right panel of Figure 4.1 shows the true variance function (solid line) and estimated variance function (Spline(N), dotted line) from the proposed spline method in Chapter 3 under the normal error assumption. The estimate variance function is much larger than the true variance function around $x = 1$. Even though there are only a few data sets out of 10,000 simulated data sets, RMSE values are much higher for these data sets compared to the other data sets. In this chapter, the penalized maximum likelihood estimator of a monotone variance function is considered as an alternative to ameliorate this problem at the boundary. The dashed line in the right panel of Figure 4.1 shows the penalized maximum likelihood estimator, and it shows that the estimated variance function around $x = 1$ is close to the true variance function compare to the maximum likelihood estimator without the penalty parameter. The penalized likelihood estimation method is described in the following section.
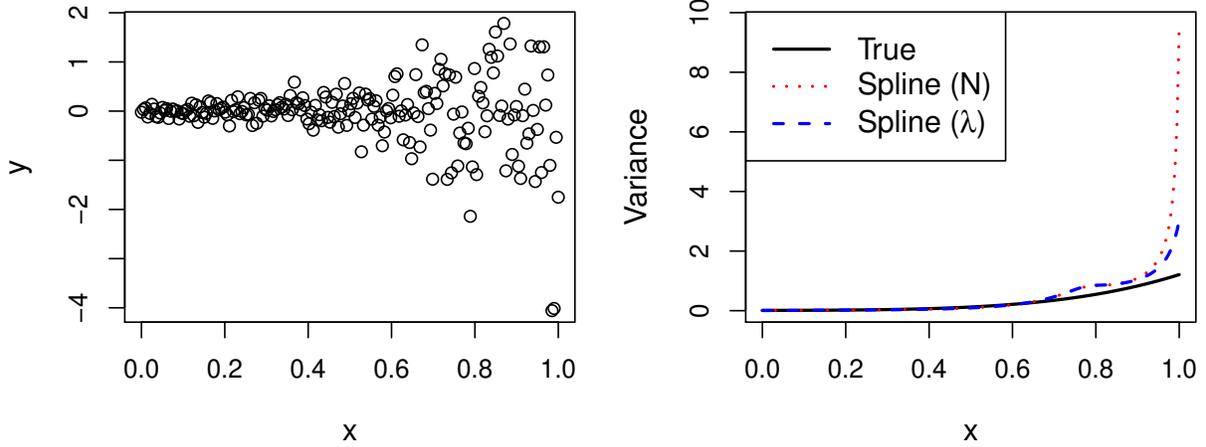
Figure 4.1: Example of the simulated data with $\mu_1 = 0$, $\sigma_1 = .1 + x^2$, and $n = 200$ when the estimated variance function spike at the boundary of domain. Left: Scatter plot, Right: True variance function (solid line) and estimated variance functions from the proposed spline methods without a penalty them (dotted line) and with a penalty term (dashed line).

4.2 Penalized Maximum Likelihood Estimation

Under the normal errors assumption, we suggest to maximize the penalized likelihood function of the form

$$\ell_\lambda(\boldsymbol{\theta}; \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2n} \left\{ \sum_{i=1}^{n} \left[\log(\theta_i) - (y_i - \widehat{\mu}(x_i))^2 \theta_i \right] + \sum_{i=n-(k-1)}^{n} \lambda_i \theta_i \right\} \tag{4.1}$$

where k is about 20% of the number of observations in the last knot interval, and λ_i's are penalty parameters. The penalty parameter λ_i is defined as

$$\lambda_i = \frac{1}{2}(y_i - \widehat{\mu}(x_i))^2.$$

The same algorithm proposed in Section 3.3.2 is used to find the maximizer of (4.1). We suggest the penalty parameter because it reduces the influence of only k data points around the boundary. Further, the penalized estimator attains the same optimal rate of convergence as unpenalized estimator, and it is proved in the Appendix A.2.

Similarly, the method can be easily adjusted for the double-exponential errors assumption. We can obtain the penalized maximum likelihood estimator with double-exponential errors by

maximizing

$$\ell_\lambda(\boldsymbol{\theta}; \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{n}\left\{ \sum_{i=1}^{n} \left[\log(\theta_i) - \sqrt{2}|y_i - \widehat{\mu}(x_i)|\theta_i \right] + \sum_{i=n-(k-1)}^{n} \sqrt{2}\lambda_i\theta_i \right\}, \qquad (4.2)$$

where

$$\lambda_i = \frac{1}{2}|y_i - \widehat{\mu}(x_i)|.$$

Again, we can use the same procedure suggested in Section 3.3.2 to optimize the penalized like-lihood functions in (4.2). The proposed penalty term is easy to set and does not affect the rate of convergence. The proof of the rate of convergence for the penalized maximum likelihood estimator is provided in the Appendix A.2.

The performances of the proposed penalized estimator are demonstrated through simulations with the same data sets generated in Chapter 3.5. Table 4.1 shows RMSE and MAE from the penalized splines estimators with the normal error assumption and the double-exponential error assumption. Both RMSE and MAE from the penalized estimation methods are substantially im-proved for both error assumptions.

Table 4.2 shows the 95% coverage probability and the average width of confidence intervals of the estimated slope parameter of the linear mean function (μ_2) from the penalized spline estimator. The average width of confidence intervals from the penalized spline estimator is narrower than that from the spline estimators without the penalty term.

The histograms of estimated slope parameter in the linear mean trend μ_2 from the penalized spline method are presented in Figures 4.2, 4.3, and 4.4. The range of histograms from the penal-ized spline method is similar to the those from the spline method without the penalty term.

Table 4.1: RMSE (MAE) from 10,000 estimated variance functions using the proposed penalty parameters.

Error	σ	n	$\mu_1(x) = 0$		$\mu_2(x)$: Linear		$\mu_3(x)$: Quartic	
			Spline(λ) N(0,1)	Spline(λ) DE	Spline(λ) N(0,1)	Spline(λ) DE	Spline(λ) N(0,1)	Spline(λ) DE
NM ($\epsilon^{[1]}$)	σ_1	50	.18 (.09)	.20 (.10)	.19 (.09)	.20 (.10)	.20 (.10)	.20 (.10)
		100	.14 (.07)	.16 (.09)	.15 (.07)	.16 (.09)	.15 (.08)	.16 (.09)
		200	.11 (.06)	.13 (.07)	.11 (.06)	.13 (.07)	.12 (.06)	.13 (.07)
		500	.09 (.04)	.11 (.07)	.09 (.04)	.11 (.06)	.09 (.04)	.11 (.06)
	σ_2	50	.24 (.14)	.31 (.16)	.25 (.14)	.30 (.16)	.25 (.14)	.27 (.14)
		100	.19 (.11)	.26 (.14)	.19 (.11)	.25 (.14)	.19 (.11)	.22 (.12)
		200	.14 (.08)	.20 (.11)	.14 (.08)	.20 (.11)	.14 (.08)	.18 (.10)
		500	.10 (.06)	.18 (.11)	.10 (.06)	.17 (.11)	.10 (.06)	.16 (.10)
DE ($\epsilon^{[2]}$)	σ_1	50	.28 (.12)	.23 (.12)	.28 (.13)	.24 (.12)	.29 (.13)	.24 (.12)
		100	.22 (.10)	.18 (.09)	.22 (.10)	.18 (.09)	.22 (.10)	.19 (.10)
		200	.16 (.08)	.15 (.07)	.16 (.08)	.15 (.07)	.16 (.08)	.15 (.08)
		500	.11 (.06)	.11 (.06)	.11 (.06)	.11 (.06)	.11 (.06)	.12 (.06)
	σ_2	50	.39 (.20)	.33 (.18)	.39 (.21)	.34 (.18)	.38 (.20)	.32 (.18)
		100	.30 (.16)	.24 (.14)	.30 (.16)	.25 (.14)	.29 (.16)	.24 (.14)
		200	.21 (.12)	.18 (.10)	.21 (.12)	.18 (.11)	.21 (.12)	.18 (.11)
		500	.14 (.08)	.12 (.07)	.14 (.08)	.12 (.07)	.14 (.08)	.13 (.07)
CN ($\epsilon^{[3]}$)	σ_1	50	.45 (.16)	.27 (.12)	.45 (.16)	.27 (.13)	43. (.16)	.27 (.13)
		100	.32 (.13)	.20 (.10)	.32 (.13)	.20 (.10)	.31 (.13)	.21 (.11)
		200	.24 (.10)	.17 (.08)	.24 (.10)	.17 (.09)	.23 (.10)	.18 (.09)
		500	.15 (.07)	.14 (.07)	.15 (.07)	.14 (.07)	.15 (.07)	.14 (.07)
	σ_2	50	.61 (.28)	.36 (.19)	.61 (.28)	.37 (.20)	.58 (.27)	.36 (.20)
		100	.44 (.22)	.26 (.16)	.44 (.22)	.27 (.16)	.43 (.22)	.27 (.17)
		200	.31 (.16)	.21 (.12)	.31 (.16)	.21 (.13)	.30 (.16)	.22 (.13)
		500	.21 (.11)	.16 (.10)	.21 (.11)	.16 (.10)	.20 (.11)	.17 (.10)

Table 4.2: 95% coverage probability (width of CI) of the estimated slope parameter of the linear mean function using the penalized spline variance estimator.

ϵ	n	$\sigma_1(x)$: Convex		$\sigma_2(x)$: Sigmoid	
		N(0,1)	DE	N(0,1)	DE
NM $(\epsilon^{[1]})$	50	.925 (.62)	.923 (.78)	.933 (.72)	.924 (.85)
	100	.933 (.45)	.933 (.57)	.943 (.51)	.940 (.63)
	200	.937 (.32)	.940 (.41)	.939 (.36)	.944 (.45)
	500	.946 (.21)	.954 (.26)	.947 (.23)	.953 (.29)
DE $(\epsilon^{[2]})$	50	.934 (.60)	.954 (.57)	.939 (.70)	.949 (.63)
	100	.937 (.43)	.965 (.40)	.944 (.50)	.964 (.45)
	200	.944 (.31)	.969 (.28)	.945 (.35)	.969 (.31)
	500	.945 (.20)	.969 (.17)	.947 (.22)	.968 (.19)
CN $(\epsilon^{[3]})$	50	.941 (.57)	.925 (.55)	.944 (.67)	.927 (.60)
	100	.940 (.41)	.936 (.40)	.945 (.48)	.938 (.44)
	200	.941 (.30)	.942 (.29)	.944 (.34)	.942 (.31)
	500	.943 (.20)	.946 (.18)	.944 (.22)	.947 (.20)
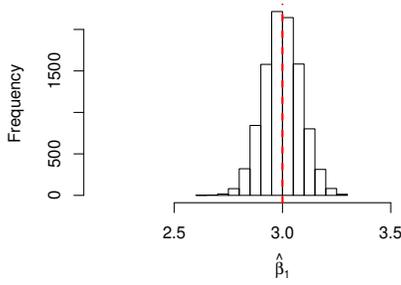
(a) Spline N(0,1) with σ_1 (convex).
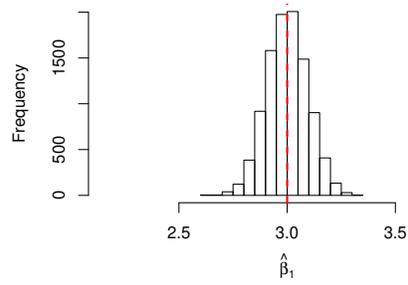
(b) Spline N(0,1) with σ_2 (sigmoid).

Figure 4.2: Histogram of estimated slope parameters of β_1 with normal errors ($\epsilon^{[1]}$) and $n = 200$.
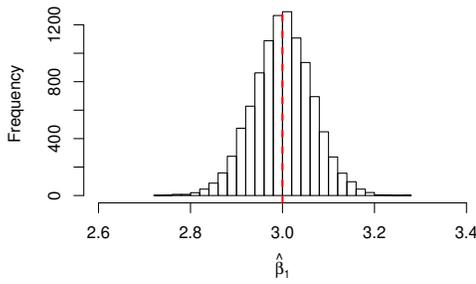
(a) Spline DE with σ_1 (convex).
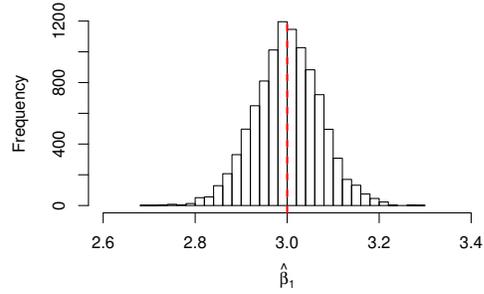
(b) Spline DE with σ_2 (sigmoid).

Figure 4.3: Histogram of estimated slope parameters of β_1 with double-exponential errors ($\epsilon^{[2]}$) and $n = 200$.
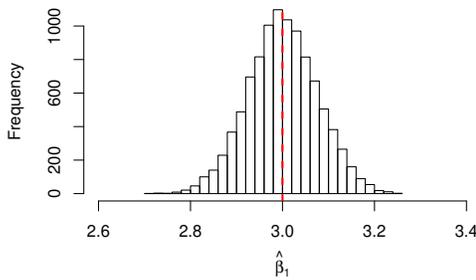
(a) Spline DE with σ_1 (convex).
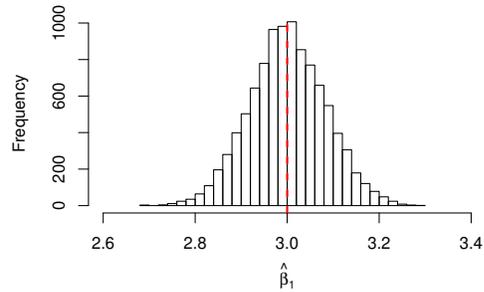
(b) Spline DE with σ_2 (sigmoid).

Figure 4.4: Histogram of estimated slope parameters of β_1 with contaminated normal errors ($\epsilon^{[3]}$) and $n = 200$.

4.3 Real Data Analysis

The penalized splines method can be applied to the U.S. temperature data (Peixoto, 1990) available in a `SemiPar` package within `R`. The data set contains average January minimum temperature, latitude, and longitude of 56 U.S. cities. The average January minimum temperature (Temperature) is obtained by adding the average minimum temperatures for January 1931 through January 1960 and dividing the total by 30. In Peixoto (1990), the response variable is the average January minimum temperature and the predictor variables are the latitude and the longitude in degrees. The ordinary regression models including the higher order terms, such as quadratic and cubic terms of predictor variables, are considered under the homoskedasitic errors assumption in Peixoto (1990). However, the average January minimum temperature can be estimated in terms of the latitude by accounting for the estimated variance function, and the model interpretability can be improved. The scatter plot in Figure 4.5 shows the estimated mean function (solid line) and the 95% confidence bands (dashed lines) from (a) the spline method with the penalized estimator of the variance function and (b) the kernel method (Ruppert) without the monotonicity constraints. Knots are marked as 'x' in (a). The bands in the panel (b) are wiggly compare to the bands in the panel (a). The fit in the panel (b) shows the temperature increases as the latitude the increase when the latitude is greater 45, and it is contradictory to common knowledge that the temperature decrease is mainly due to the increase in latitude. Panels (c) and (d) in Figure 4.5 show the estimated variance functions from the proposed spline method without the penalty term and the proposed spline method with the penalty term, respectively. The penalty term reduce the influence of the data values at the boundary of domain, and handle the spiking problem. The variance functions estimated from the kernel methods introduced by Ruppert and Dette are shown in panels (e) and (f), respectively. Solid lines in panels (e) and (f) are the estimate variance functions with the suggested bandwidths, and both curves yield negative values. The estimated variance functions from the kernel method (Ruppert) with larger bandwidths yield positive variances, but the estimated variance function form kernel method (Dette) yields the negative values even with the smaller bandwidths.
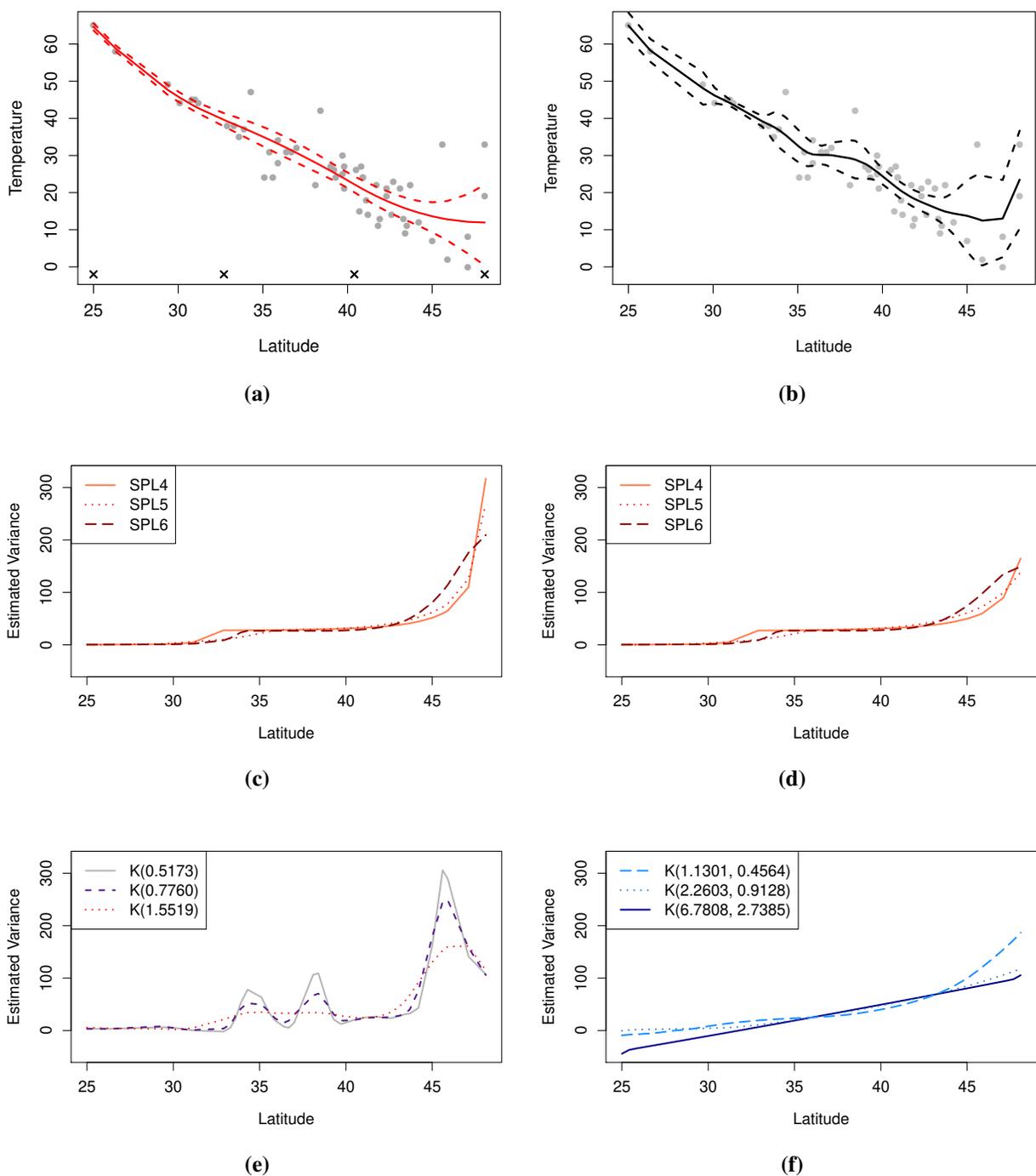
Figure 4.5: U.S. temperature data: (a) Scatter plot of average January minimum temperature (Temperature) and latitude with the estimated mean function and point-wise confidence bands from the proposed penalized spline method. Knots are marked as 'x'. (b) Fit and bands from the kernel method (Ruppert). Estimated variance functions using the proposed spline method (c) without the penalty term and (d) with the penalty term. (e) Estimated variance functions using local polynomial kernel method (Ruppert et al., 1997) with three different bandwidths. (f) Estimated variance functions using kernel method (Dette and Pilz, 2009) with three different combinations of bandwidths.

60

Chapter 5

Hypothesis testing

5.1 Literature Review

In this chapter, our focus is on the hypothesis test for the heteroskedastic regression model. Several test statistics have been introduced for testing the homoskedasticity against monotone variance functions under the normal errors distribution assumption by Goldfeld and Quandt (1965) and Szroeter (1978). Goldfeld and Quandt (1965) proposed the parametric and the nonparametric test (GQ test) for testing homoskedasticity against the alternative that the variance function is monotonic in an independent variable. The parametric GQ test is widely used when the errors are assumed to be normally distributed. The test is constructed by following process. To begin with, the observations are ordered by the value of the independent variable x which is called the potential deflator. Then, choose the number of central observations, k for the sample size n, to be omitted. After omitting the k observations in the middle, we fit the linear regression model separately to the first $(n-k)/2$ observations and the last $(n-k)/2$ observations. Denote S_1 and S_2 to be the sum of squares of the residuals from the regression models based on the relatively small and relatively large value of x, respectively. The test statistic can be defined as $R = S_2/S_1$, and it follows F-distribution with

$$\left(\frac{n-k-2q-2}{2}, \frac{n-k-2q-2}{2}\right)$$

degrees of freedom under the null hypothesis where q is the number of covariates to fit the mean trend. The nonparametric GQ test is constructed based on the number of peaks in the ordered sequence of absolute residuals. It assume that the number of observed peaks will tend to be large if the variance increases with a deflated parameter. For the residuals $\hat{e}_1, \cdots, \hat{e}_n$, define a peak at j-th observation where $|\hat{e}_j| \geq |\hat{e}_i|$ for all $i < j$. The cumulated probabilities of yielding $0, 1, \cdots, n-1$ peaks are computed under the null hypothesis, and are used to compare with the confidence level.

Szroeter (1978) introduced the parametric bound tests as an alternative to the parametric test proposed by Goldfeld and Quandt (1965). It is assumed that the variances have been ordered such as $\sigma(x_i) \leq \sigma(x_{i+1})$ for $i = 1, \ldots, n-1$. The variance functions are no need to be a smooth monotonic function of an independent variable. Let $\hat{e}_i$ be the ith residual. The Szroeter test statistic is defined by

$$Q = \sqrt{\frac{6n}{n^2-1}} \left(h - \frac{n+1}{2} \right) \text{ where } h = \frac{\sum_{i=1}^{n} i\hat{e}_i^2}{\sum_{i=1}^{n} \hat{e}_i^2},$$

and it follows the approximate standard normal distribution.

Fujino (1979) introduced tests for testing the homogeneity of a set of variances of normal populations against ordered alternatives based on the normal distribution. In this paper, we are interested in testing a constant variance against alternatives of heteroskedasticity in the form of monotone variance function in regression models. We propose a likelihood ratio test of the null hypothesis that the variance is constant, versus the alternative that it is a smooth non-decreasing function.

5.2 The Proposed Test

We assume that the sample data $\{(x_i, y_i, \boldsymbol{z}_i), i = 1, ..., n\}$ are generated from the model

$$y_i = \mu(\boldsymbol{z}_i) + \sigma(x_i)\epsilon_i, \tag{5.1}$$

where $\sigma(\cdot)$ is a smooth non-decreasing function. Assume x_i's are equally spaced in $[0, 1]$, and ordered from the smallest to the largest. For the error term ϵ, we consider the normal error distribution and the double-exponential error distribution with mean zero and variance one. The design vector $\boldsymbol{z}_i \in \mathbb{R}^{q+1}$ might include x_i but not necessarily where q is the number of predictors for the mean function. We proposed the likelihood ratio test (LRT) for homogeneous variance against the heteroskedastic variance with the monotonicity constraints in the regression model. That is

$$H_0 : \sigma^2(x_i) \equiv \sigma^2 \text{ versus } H_1 : \sigma^2(x) \text{ is monotone in } x,$$

where σ is a positive constant. The proposed test statistics are depend on the error distribution as follows.

5.2.1 Normal Errors Case

Assume that the errors follow the standard normal distribution. For the non-decreasing variances, a vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ is defined where $\theta_i = \sigma^{-2}(x_i)$ for $i = 1, \ldots, n$. The variance function estimation method introduced in Chapter 3 under the normal errors assumption is used to compute the likelihood functions. Let the sets $\Theta_0 = \{\boldsymbol{\theta} : \theta_i \equiv \sigma^{-2} \text{ for } i = 1, \ldots, n\}$ and $\Theta_1 = \{\boldsymbol{\theta} \in \mathbb{R}^n : \boldsymbol{\theta} = \boldsymbol{B}\boldsymbol{\alpha}, \text{ where } \boldsymbol{A}\boldsymbol{\alpha} \geq 0\}$ from (3.2). From the model assumption (5.1) with the assumption of normal errors, the likelihood function of $\boldsymbol{\theta}$ can be written as

$$L(\boldsymbol{\theta}; \boldsymbol{y}) = 2\pi^{-n/2} \left(\prod_{i=1}^{n} \sqrt{\theta_i} \right) \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \left(y_i - \hat{\mu}(\boldsymbol{z}_i) \right)^2 \theta_i \right). \tag{5.2}$$

The LRT statistic is defined as

$$T(\boldsymbol{y}) = -2 \ln \Lambda(\boldsymbol{\theta}, \boldsymbol{y}), \tag{5.3}$$

where the likelihood ratio for the hypothesis test can be written as

$$\Lambda(\boldsymbol{\theta}, \boldsymbol{y}) = \frac{\sup_{\boldsymbol{\theta} \in \Theta_0} L(\boldsymbol{\theta}; \boldsymbol{y})}{\sup_{\boldsymbol{\theta} \in \Theta_1} L(\boldsymbol{\theta}; \boldsymbol{y})}. \tag{5.4}$$

Note that the likelihood ratio $\Lambda(\boldsymbol{\theta}, \boldsymbol{y})$ is smaller than or equal to 1 because $\Theta_0 \subset \Theta_1$. The proof that the LRT statistic is scale invariant is found in Appendix A.3.

5.2.2 Heavy-Tailed Errors Case

When the true error distribution is heavy-tailed, the double-exponential distribution is considered. Here, let $\theta_i = \sigma(x_i)^{-1}$ for $i = 1, \ldots, n$. The variance function estimation method introduced in Chapter 3 under the double-exponential errors assumption is used to compute the likelihood functions. Suppose the sets $\Theta_0 = \{\boldsymbol{\theta} : \theta_i \equiv \sigma^{-1} \text{ for } i = 1, \ldots, n\}$ and $\Theta_1 = \{\boldsymbol{\theta} \in \mathbb{R}^n : \boldsymbol{\theta} = \boldsymbol{B}\boldsymbol{\alpha}, \text{ where } \boldsymbol{A}\boldsymbol{\alpha} \geq 0\}$. The likelihood function under the double-exponential error assumption

can be written as

$$L(\boldsymbol{\theta}; \boldsymbol{y}) = 2^{-n/2} \left(\prod_{i=1}^{n} \theta_i \right) \exp \left(-\sqrt{2} \sum_{i=1}^{n} |y_i - \hat{\mu}(\boldsymbol{z}_i)| \theta_i \right). \tag{5.5}$$

Similar to the normal error case, the LRT statistic for the heavy-tailed errors case, $T(\boldsymbol{y})$, can be obtained as (5.3) using the double-exponential likelihood function and the ratio in (5.4) and (5.5).

5.3   Simulation study

In this section, we explore the performances of the proposed LRT test under both normal and double-exponential error assumptions through a series of simulations.

For the critical values of the proposed test, we generate $M_0 = 100,000$ samples of size n under the null hypothesis assuming $\mu \equiv 0$; i.e., $y_i = \sigma \epsilon_i$ where σ is a positive constant. Note that the choice of a constant value of σ does not affect the critical values of the proposed methods because of the scale-invariant property which is shown in Appendix A.3. For each of the sample size $n = 50, 100, 200$, and 500, the LRT statistics, $T_1, \ldots, T_{M_0}$, were computed under both normal and double-exponential error assumptions. For $\alpha = .01, .05$, and $.10$, the critical value $T_{(1-\alpha)}$ for each sample size is obtained as the $(1-\alpha)100$th percentile of M_0 simulated test statistics and shown in Table 5.1. Further, simulation studies were carried out to determine the test size and the power, and

Table 5.1: Empirical critical values for the test statistics of the propose LRT when the errors follow the normal distribution and the double-exponential distribution.

n \ α	Normal errors			Double-exponential errors		
	0.10	0.05	0.01	0.10	0.05	0.01
50	3.3873	4.7339	7.8953	3.3609	4.7246	7.9882
100	3.7431	5.2088	8.5491	3.7753	5.2320	8.6702
200	4.0623	5.5717	9.0177	4.0473	5.5234	8.9622
500	4.5564	6.1213	9.5640	4.5521	6.0900	9.6896

to compare it with existing hypothesis tests; the parametric and the nonparametric Goldfeld and

Quandt (GQ) test and Szroeter's test. For the nonparametric GQ test, we obtained the empirical cumulative probabilities of the distribution of peaks from $100,000$ samples for each sample size to compare it with the confidence level. For the parametric GQ test, we consider $k = 0$ and k is the

Table 5.2: Empirical cumulative probabilities of the distribution of peaks of absolute residuals for the nonparametric GQ test.

x\n	P(number of peaks $\leq x$)												
	0	1	2	3	4	5	6	7	8	9	10	11	12
50	0.019	0.108	0.295	0.530	0.739	0.879	0.953	0.984	0.995	0.999	1.000	1.000	1.000
100	0.010	0.061	0.187	0.379	0.591	0.769	0.887	0.951	0.981	0.994	0.998	0.999	1.000
200	0.005	0.035	0.117	0.263	0.453	0.642	0.795	0.895	0.952	0.981	0.993	0.998	0.999
500	0.002	0.016	0.061	0.156	0.302	0.475	0.644	0.783	0.881	0.940	0.972	0.989	0.996

30% of the sample size. For $k = 0$, the first half of data points are used for the first group and the last half of data points are used for the second group. When k is the 30% of the sample size, the first 35% of data values are used for the first group and the last 35% of the data values are used for the second group after removing 30% of the data values in the middle.

The data sets are generated with following error distributions:

1. $\epsilon^{[1]} \sim N(0,1)$: Normal distribution with mean 0 and variance 1.

2. $\epsilon^{[2]} \sim DE(0, 2^{-1/2})$: Double-exponential distribution with mean 0 and variance 1.

3. $\epsilon^{[3]} \sim$ Contaminated normal distribution : 90% from $N(0, .65^2)$ and 10% from $N(0, 2.5^2)$.

The mean trends considered are:

1. For $\mu_1(x) = 0$, we assume the mean trend is known.

2. For $\mu_2(x) = 3x + 5$, we assume the trend is a linear function.

3. For $\mu_3(x) = 50(x - .5)^4 I(x \geq 0.5)$, we assume the trend is smooth and increasing.

For the parametric GQ test and the Szroter test, the unknown mean function (μ_2 or μ_3) is estimated using the linear regression.

For the power curves, the data sets are generated using following standard deviation functions with a slope a:

1. Convex : $\sigma_1(x) = .1 + ax^2$,

2. Sigmoid : $\sigma_2(x) = .1 + a\exp(15x - 8)/(1 + \exp(15x - 8))$

3. Quartic : $\sigma_3(x) = 2 - a(x - .5)^4 I(x \leq .5)$.

5.3.1 Test Size

The test size is the probability of rejecting the true null hypothesis. To describe the superiority in test size of the proposed test, data sets are generated from the model $y_i = \mu(x_i) + \sigma \epsilon_i$ where $\sigma = 0.1$. Test size with 95% confidence level are shown in Table 5.3. For the normal errors, the proposed LRT under the normal errors assumption is comparable to the GQ test or the Szroter test. The test sizes of nonparametric GQ test are always larger than 0.05. For the heavy tailed errors, the proposed LRT based on the double-exponential assumption performs consistently well compared to GQ test and the Szroter test. The test sizes for the parametric GQ test and the Szroter's test are inflated even with the large sample sizes, especially for the data sets generated with the contaminated normal errors ($\epsilon^{[3]}$).

5.3.2 Power Study

For each standard deviation function, we choose 10 values for slope a to demonstrate the performances of our proposed tests. Figures 5.1, 5.2, and 5.3 show the power curves along with the slope a when we generate the data with $\mu_1 = 0$ and $n = 200$. The increasing patterns of the power are similar with different mean functions or different sample sizes. Figure 5.1 shows the power curves for the data sets generated with normal errors. The proposed LRT based on the normal likelihood function (circle points, solid curve) is comparable to other method when the true variance function is (a) convex (σ_1^2) or (b) sigmoid (σ_2^2). When the true variance function is quartic (σ_3^2) in Figure 5.1 (c), the power of the proposed test with the normal errors assumption (circle points, solid curve) increases faster than others as the slope a increases. It means that the proposed test

66

Table 5.3: Test size: the probability of rejecting the true null hypothesis with 95% confidence level.

μ	ϵ	n	LR N(0,1)	LR DE	GQ $k=0$	GQ $k=.3n$	Szroter	GQ NP
$\mu_1 = 0$	NM ($\epsilon^{[1]}$)	50	0.047	0.009	0.046	0.042	0.046	0.124
		100	0.051	0.008	0.050	0.049	0.051	0.111
		200	0.048	0.006	0.047	0.050	0.049	0.104
		500	0.053	0.007	0.052	0.052	0.051	0.065
	DE ($\epsilon^{[2]}$)	50	0.177	0.051	0.133	0.125	0.126	0.122
		100	0.202	0.054	0.143	0.135	0.140	0.123
		200	0.219	0.052	0.154	0.146	0.148	0.102
		500	0.234	0.044	0.144	0.147	0.139	0.060
	CN ($\epsilon^{[3]}$)	50	0.300	0.073	0.227	0.208	0.198	0.125
		100	0.361	0.072	0.243	0.234	0.220	0.115
		200	0.404	0.081	0.251	0.245	0.238	0.099
		500	0.457	0.083	0.238	0.240	0.235	0.056
μ_2 : Linear	NM($\epsilon^{[1]}$)	50	0.060	0.018	0.050	0.049	0.046	0.130
		100	0.057	0.011	0.051	0.050	0.051	0.120
		200	0.051	0.008	0.048	0.052	0.050	0.105
		500	0.052	0.008	0.052	0.053	0.051	0.066
	DE ($\epsilon^{[2]}$)	50	0.183	0.064	0.128	0.127	0.123	0.124
		100	0.205	0.058	0.143	0.136	0.136	0.123
		200	0.222	0.054	0.152	0.144	0.146	0.106
		500	0.235	0.044	0.144	0.144	0.140	0.060
	CN ($\epsilon^{[3]}$)	50	0.306	0.091	0.223	0.206	0.194	0.130
		100	0.366	0.081	0.242	0.233	0.217	0.114
		200	0.405	0.086	0.250	0.246	0.237	0.102
		500	0.458	0.084	0.237	0.238	0.235	0.057
μ_3: Quartic	NM ($\epsilon^{[1]}$)	50	0.077	0.018	0.050	0.050	0.043	0.140
		100	0.050	0.014	0.050	0.052	0.041	0.122
		200	0.045	0.007	0.049	0.049	0.044	0.113
		500	0.045	0.006	0.050	0.049	0.043	0.059
	DE ($\epsilon^{[2]}$)	50	0.189	0.055	0.123	0.115	0.096	0.135
		100	0.177	0.057	0.136	0.127	0.109	0.120
		200	0.203	0.047	0.146	0.139	0.118	0.110
		500	0.229	0.049	0.146	0.150	0.126	0.060
	CN ($\epsilon^{[3]}$)	50	0.315	0.078	0.210	0.191	0.161	0.130
		100	0.337	0.090	0.225	0.219	0.186	0.118
		200	0.396	0.079	0.241	0.237	0.209	0.106
		500	0.441	0.080	0.244	0.244	0.221	0.062

outperforms the other test. In Figures 5.2 and 5.3, both power and test size for the proposed LRT based on the double-exponential likelihood function (triangle points, dashed curve) outperforms other existing methods. Even though the test sizes for the contaminated normal errors with sample size 500 from the nonparametric GQ test are also close to 0.05 in Table 5.3, the proposed LRT under the double-exponential errors assumption is more favorable because of the power curve.
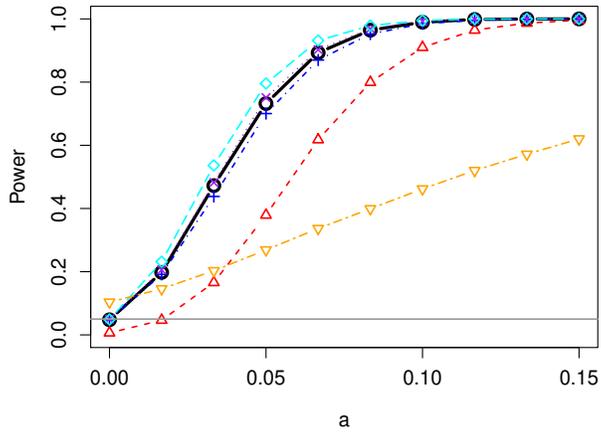
(a) Data sets generated with σ_1(convex) and $\epsilon^{[1]}$(NM)
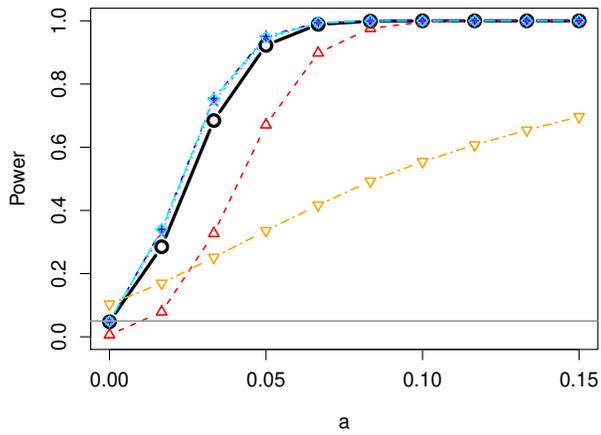
(b) Data sets generated with σ_2(sigmoid) and $\epsilon^{[1]}$(NM)
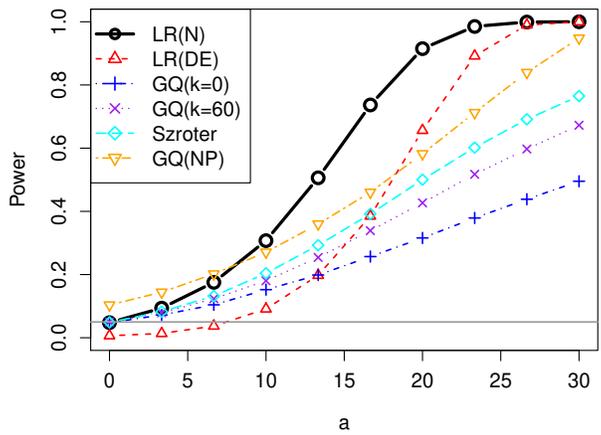
(c) Data sets generated with σ_3(quartic) and $\epsilon^{[1]}$(NM)

Figure 5.1: Power curves for normal errors ($\epsilon^{[1]}$). Data sets are generated with $\mu_1 = 0$ and $n = 200$. The proposed LRT based on the normal error assumption (circle points, solid curves) is comparable or outperforms other tests.
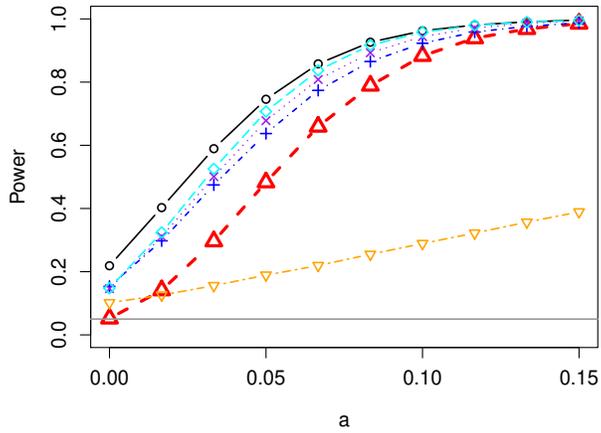
(a) Data sets generated with σ_1(convex) and $\epsilon^{[2]}$(DE)
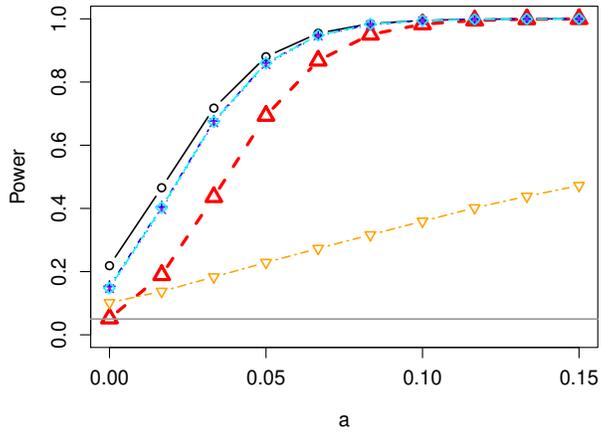
(b) Data sets generated with σ_2(sigmoid) and $\epsilon^{[2]}$(DE)
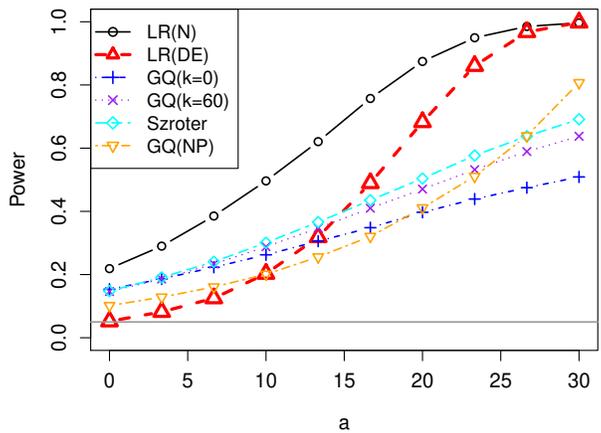
(c) Data sets generated with σ_3(quartic) and $\epsilon^{[2]}$(DE)

Figure 5.2: Power curves for double-exponential errors ($\epsilon^{[2]}$). Data sets are generated with $\mu_1 = 0$ and $n = 200$. The proposed LRT based on the double-exponential error assumption (triangle points, dashed curves) performs consistently well and outperforms other tests.
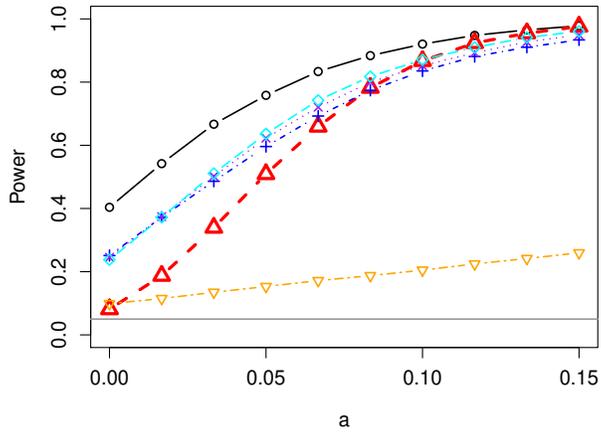
(a) Data sets generated with σ_1(convex) and $\epsilon^{[3]}$(CN)
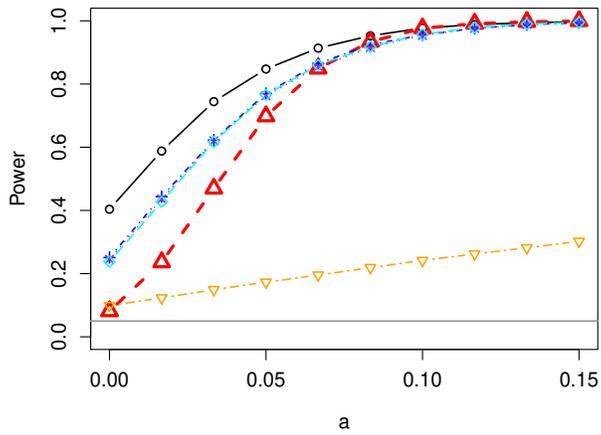
(b) Data sets generated with σ_2(sigmoid) and $\epsilon^{[3]}$(CN)
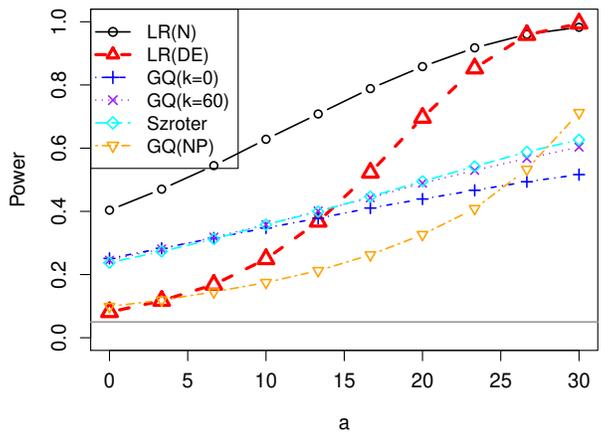
(c) Data sets generated with σ_3(quartic) and $\epsilon^{[3]}$(CN)

Figure 5.3: Power curves for contaminated normal errors ($\epsilon^{[3]}$). Data sets are generated with $\mu_1 = 0$ and $n = 200$. The proposed LRT based on the double-exponential error assumption (triangle points, dashed curves) performs consistently well and outperforms other tests.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this dissertation, the convergence rate of the constrained spline estimator has been established when the true function is a monotone function. In addition, we develop the constrained spline estimation methods and the hypothesis test for a smooth monotone variance function in a regression model. The convergence rates of proposed estimators are established, and they attained the optimal rate. Simulation results show that our proposed methods are comparable or outperforms existing methods under various settings. The proposed estimators have a bigger advantage when the variance function is close to zero compare to the existing kernel estimators. The proposed estimators always satisfy the assumption of positive variances, and provide the substantially improved inference about the mean trend in a heteroskedastic regression model. Application to real data sets, such as LIDAR data, abalone data, California air pollution data, and US temperature data, demonstrates that the proposed method provide more accurate and robust results compared with existing ones. Further, the proposed methods allow to incorporate the parametrically modeled covariate in the variance structure. The penalized estimator is also presented for improving the performance when there exists a spike at the boundary, by reducing the effect of the observation near the boundary. Hypothesis test for the monotone variance function is also discussed, and the performance of the proposed test is favorable in terms of the test size and the power.

6.2 Future Work

In Chapter 3.6, our proposed method for monotone variance function estimation is applied to various data sets. There are some other interesting shapes of the variance function that have not been discussed. For example, we could develop the method to consider convex, convex and increasing, or V-shaped variance functions. Other types of splines could be considered, such as

I-splines (Ramsay, 1988) for monotone functions or C-splines (Meyer, 2008) for convex functions. We will add more shape options for the variance function estimation method and the hypothesis testing procedure in an `R` package.

Moreover, the method could be extended to include the construction of confidence bands under a heavy-tailed error distribution in a heteroskedastic regression model.

Further development could include the application of the proposed method to more complex models, such as quantile regression models and generalized linear models. In addition, investigation of overdispersion for discrete data sets and the theoretical properties of estimators could be pursued. Finally, the estimation problem could also be reframed to consider the effect of a heteroskedastic regression model with correlated errors.

Bibliography

Bassett, G. and Koenker, R. (1978). Asymptotic theory of least absolute error regression. *Journal of the American Statistical Association*, 73(363):618–622.

Breiman, L. and Friedman, J. H. (1985). Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80(391):580–598.

Cabrera, J. L. O. (2018). *locpol: Kernel Local Polynomial Regression*. R package version 0.7-0.

Carroll, R. J. (1982). Adapting for heteroscedasticity in linear models. *Ann. Statist.*, 10(4):1224–1233.

Carroll, R. J. and Ruppert, D. (1988). *Transformation and Weighting in Regression*. Chapman & Hall, Ltd., London, UK, UK.

Davidian, M. and Carroll, R. J. (1987). Variance function estimation. *J. Amer. Statist. Assoc.*, 82(400):1079–1091.

de Boor, C. (2001). *A practical guide to splines*, volume 27 of *Applied Mathematical Sciences*. Springer-Verlag, New York, revised edition.

Dette, H., Neumeyer, N., and Pilz, K. F. (2006). A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli*, 12(3):469–490.

Dette, H. and Pilz, K. (2009). On the estimation of a monotone conditional variance in nonparametric regression. *Ann. Inst. Statist. Math.*, 61(1):111–141.

Fan, J. and Gijbels, I. I. (1996). *Local polynomial modelling and its applications*. London ; New York : Chapman & Hall, 1st ed edition.

Fujino, Y. (1979). Tests for the homogeneity of a set of variances against ordered alternatives. *Biometrika*, 66(1):133–139.

Gijbels, I., Prosdocimi, I., and Claeskens, G. (2010). Nonparametric estimation of mean and dispersion functions in extended generalized linear models. *Test*, 19(3):580–608.

Goldfeld, S. M. and Quandt, R. E. (1965). Some tests for homoscedasticity. *Journal of the American Statistical Association*, 60:539–547.

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 (Methodological)*, 51(1):3–14.

Huang, J. Z. (2001). Concave extended linear modeling: A theoretical synthesis. *Statistica Sinica*, 11:173–197.

Kim, S.-Y., Meyer, M. C., and Wang, H. (2018). Improved inference in heteroskedastic regression models with monotone variance function estimation. Submitted for publication.

Lian, H., Liang, H., and Carroll, R. J. (2015). Variance function partially linear single-index models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(1):171–194.

Liao, X. and Meyer, M. C. (2014). coneproj: An r package for the primal or dual cone projections with routines for constrained regression. *Journal of Statistical Software*, 61(12):1–22.

Liao, X. and Meyer, M. C. (2017). *cgam: Constrained Generalized Additive Model*. R package version 1.7.

Liu, A., Tong, T., and Wang, Y. (2007). Smoothing spline estimation of variance functions. *Journal of Computational and Graphical Statistics*, 16(2):312–329.

Meyer, M. C. (1999). An extension of the mixed primal-ĂŞdual bases algorithm to the case of more constraints than dimensions. *Journal of Statistical Planning and Inference*, 81(1):13 – 31.

Meyer, M. C. (2008). Inference using shape-restricted regression splines. *Ann. Appl. Stat.*, 2(3):1013–1033.

Meyer, M. C. (2013). A simple new algorithm for quadratic programming with applications in statistics. *Comm. Statist. Simulation Comput.*, 42(5):1126–1139.

Meyer, M. C. (2018). A framework for estimation and inference in generalized additive models with shape and order restrictions. Submitted for publication.

Meyer, M. C., Kim, S.-Y., and Wang, H. (2018). Convergence rates for constrained regression splines. *Journal of Statistical Planning and Inference*, 193:179–188.

Müller, H.-G. and Stadtmuller, U. (1987). Estimation of heteroscedasticity in regression analysis. *Ann. Statist.*, 15(2):610–625.

Peixoto, J. L. (1990). A property of well-formulated polynomial regression models. *The American Statistician*, 44(1):26–30.

Pilz, K. and Titoff, S. (2015). *monreg: Nonparametric Monotone Regression*. R package version 0.1.3.

Ramsay, J. O. (1988). Monotone regression splines in action. *Statist. Sci.*, 3(4):425–441.

Ruppert, D., Sheather, S. J., and Wand, M. P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, 90(432):1257–1270.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric regression*, volume 12 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.

Ruppert, D., Wand, Matt, P., Holst, U., and Hössjer, O. (1997). Local polynomial variance-function estimation. 39(3):262–273.

Schumaker, L. L. (2007). *Spline functions: basic theory*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition.

Silvapulle, M. and Sen, P. (2005). *Constrained Statistical Inference: Order, Inequality, and Shape Constraints*. Wiley Series in Probability and Statistics. Wiley.

Staudenmayer, J., Ruppert, D., and Buonaccorsi, J. P. (2008). Density estimation in the presence of heteroscedastic measurement error. *Journal of the American Statistical Association*, 103(482):726–736.

Stone, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Annals of Statistics*, 8:1348–1360.

Szroeter, J. (1978). A class of parametric tests for heteroscedasticity in linear econometric models. *Econometrica*, 46(6):1311–27.

Wang, L., Brown, L. D., Cai, T. T., and Levine, M. (2008). Effect of mean on variance function estimation in nonparametric regression. *The Annals of Statistics*, 36(2):646–664.

Appendix A

Technical Details

A.1   Proof of the Convergence of the Algorithm.

For both normal and double-exponential error distributions, the minimizer $\widehat{\boldsymbol{\theta}}$ of the convex function ψ over the closed convex set $\mathcal{C}$ is unique, and the necessary and sufficient conditions for $\widehat{\boldsymbol{\theta}}$ are

$$\nabla\psi(\widehat{\boldsymbol{\theta}})^{\top}\widehat{\boldsymbol{\theta}} = 0 \text{ and } \nabla\psi(\widehat{\boldsymbol{\theta}})^{\top}\boldsymbol{\theta} \geq 0, \tag{A.1}$$

for all $\boldsymbol{\theta} \in \mathbb{C}$ (proposition 3.12.3 of Silvapulle and Sen (2005)). We first show that $\widehat{\boldsymbol{\theta}}$ is a fixed point of the algorithm. That is, if $\boldsymbol{\theta}_{(k)} = \widehat{\boldsymbol{\theta}}$, then $\boldsymbol{\theta}_{(k+1)} = \widehat{\boldsymbol{\theta}}$ where $\boldsymbol{\theta}_{(k)}$ is the kth solutions from the algorithm. If $\boldsymbol{\theta}_{(k)} = \widehat{\boldsymbol{\theta}}$, then, from (3.6), we have

$$\psi_k(\boldsymbol{\theta}) = \psi(\widehat{\boldsymbol{\theta}}) + \nabla\psi(\widehat{\boldsymbol{\theta}})^{\top}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}) + \frac{1}{2}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})^{\top}\boldsymbol{H}(\widehat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}). \tag{A.2}$$

Since the Hessian matrix $\boldsymbol{H}$ is positive definite, the last term of (A.2) is positive. From (A.1), $\nabla\psi(\widehat{\boldsymbol{\theta}})^{\top}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}) \geq 0$. Therefore, $\boldsymbol{\theta}_{(k+1)} = \widehat{\boldsymbol{\theta}}$, and $\widehat{\boldsymbol{\theta}}$ is the fixed point of an algorithm.

Next, assume $\boldsymbol{\theta}_{(k)} \neq \widehat{\boldsymbol{\theta}}$. Consider $\boldsymbol{\theta} = p\widehat{\boldsymbol{\theta}} + (1-p)\boldsymbol{\theta}_{(k)}$ where $p \in (0,1]$. Then, from $\boldsymbol{\theta} - \boldsymbol{\theta}_{(k)} = p(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)})$ and (3.6), we have

$$\begin{aligned}\psi_k(\boldsymbol{\theta}) &= \psi(\boldsymbol{\theta}_{(k)}) + \nabla\psi(\boldsymbol{\theta}_{(k)})^{\top}(\boldsymbol{\theta} - \boldsymbol{\theta}_{(k)}) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_{(k)})^{\top}\boldsymbol{H}(\boldsymbol{\theta}_{(k)})(\boldsymbol{\theta} - \boldsymbol{\theta}_{(k)}) \\ &= \psi(\boldsymbol{\theta}_{(k)}) + p\nabla\psi(\boldsymbol{\theta}_{(k)})^{\top}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)}) + \frac{1}{2}p^2(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)})^{\top}\boldsymbol{H}(\boldsymbol{\theta}_{(k)})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)}). \end{aligned} \tag{A.3}$$

By convexity, we have $\psi(\widehat{\boldsymbol{\theta}}) > \psi(\boldsymbol{\theta}_{(k)}) + \nabla\psi(\boldsymbol{\theta}_{(k)})^{\top}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)})$, and it follows that $0 > \psi(\widehat{\boldsymbol{\theta}}) - \psi(\boldsymbol{\theta}_{(k)}) > \nabla\psi(\boldsymbol{\theta}_{(k)})^{\top}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)})$. We can choose p in (A.3) such that

$$0 < p < 2\left|\nabla\psi(\boldsymbol{\theta}_{(k)})^{\top}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)})\right|\left[(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)})^{\top}\boldsymbol{H}(\boldsymbol{\theta}_{(k)})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{(k)})\right]^{-1}.$$

There exists a $\boldsymbol{\theta} \neq \boldsymbol{\theta}_{(k)}$ in $\mathcal{C}$ such that $\psi_k(\boldsymbol{\theta}) < \psi(\boldsymbol{\theta}_{(k)})$, and $\boldsymbol{\theta}_{(k)}$ is not a fixed point of the algorithm.

In step 2 of the algorithm, the vector $\boldsymbol{a}$ is a minimizer of $\psi_k(\boldsymbol{\theta})$ over $\mathcal{C}$. We know that $\psi(\boldsymbol{\theta})$ is strictly decreasing at $\boldsymbol{\theta}_{(k)}$ in the direction towards $\boldsymbol{a}$. Therefore, there is a solution $\boldsymbol{\theta}_{(k+1)}$ that minimizes ψ over the line segment between $\boldsymbol{\theta}_{(k)}$ and $\boldsymbol{a}$. Hence, $\psi(\boldsymbol{\theta}_{(k+1)}) < \psi(\boldsymbol{\theta}_{(k)})$.

Now, we have a sequence $\boldsymbol{\theta}_{(k)}$ of distinct points in $\mathbb{R}^n$ (unless $\boldsymbol{\theta}_{(k)} = \widehat{\boldsymbol{\theta}}$). If the points are in a compact set in $\mathbb{R}^n$, there must be a limit point for the sequence by the Bolzano-Weierstrass theorem, and this must be $\widehat{\boldsymbol{\theta}}$, the only fixed point of the algorithm. To show compactness, we show that the set $\mathcal{S} = \{\boldsymbol{\theta} : \psi(\boldsymbol{\theta}) \leq \psi(\boldsymbol{\theta}_{(0)})\}$ is bounded, or equivalently if $\psi(\boldsymbol{\theta}) \leq \psi(\boldsymbol{\theta}_{(0)})$, there exists a positive constant M_2 such that $0 < \theta_n \leq \cdots \leq \theta_1 \leq M_2$. For the normal error assumption, $\psi(\boldsymbol{\theta})$ is minimized when $\theta_i = 1/y_i^2$ for $i = 1, \ldots, n$, and $\psi(\boldsymbol{\theta}) \geq n + \sum_{i=1}^{n} \log(y_i^2)$. For fixed y, the function $y^2\theta - \log\theta$ is above the line tangent at $\theta = 2/y^2$, and we have

$$y^2\theta - \log\theta \geq 1 + \log\left(\frac{y^2}{2}\right) + \frac{y^2}{2}\theta. \tag{A.4}$$

From (A.4), we have

$$\begin{aligned} \psi(\boldsymbol{\theta}_{(0)}) \;\geq\; \psi(\boldsymbol{\theta}) &= \sum_{i=1}^{n}(y_i^2\theta_i - \log(\theta_i)) \\ &\geq\; \sum_{i=1}^{n}\left(1 + \log(y_i^2) - \log 2 + \frac{1}{2}y_i^2\theta_i\right) \\ &\geq\; n + \sum_{i=1}^{n}\log(y_i^2) - n\log 2 + \frac{1}{2}y_1^2\theta_1, \end{aligned} \tag{A.5}$$

for all $\boldsymbol{\theta} \in \mathcal{S}$. The upper bound for θ_1 can be obtained as

$$\theta_1 \leq \frac{2\left[\psi(\boldsymbol{\theta}_{(0)}) - \left(n + \sum_{i=1}^{n}\log(y_i^2)\right) + n\log 2\right]}{y_1^2}. \tag{A.6}$$

Therefore, for all $\boldsymbol{\theta} \in \mathcal{S}$, we have

$$0 < \theta_n \leq \cdots \leq \theta_1 \leq \frac{2\big[\psi(\boldsymbol{\theta}_{(0)}) - \big(n + \sum_{i=1}^{n} \log(y_i^2)\big) + n \log 2\big]}{y_1^2},$$

which is finite by the assumption that the observed $y_1^2, \ldots, y_n^2$ are strictly positive.

Similarly, we can also find the upper bound for the double-exponential distribution. Note that $\psi(\boldsymbol{\theta})$ is minimized when $\theta_i = 1/(\sqrt{2}|y_i|)$ for $i = 1, \ldots, n$, we have $\psi(\boldsymbol{\theta}) \geq n + \sum_{i=1}^{n} \log(\sqrt{2}|y_i|)$. For fixed y, the function $\sqrt{2}|y|\theta - \log \theta$ is above the line tangent at $\theta = \sqrt{2}/|y|$, and we have

$$\sqrt{2}|y|\theta - \log \theta \geq 1 + \log \left(\frac{|y|}{\sqrt{2}}\right) + \frac{|y|}{\sqrt{2}}\theta. \tag{A.7}$$

From (A.7), we have

$$
\begin{aligned}
\psi(\boldsymbol{\theta}_{(0)}) \;\geq\; \psi(\boldsymbol{\theta}) &= \sum_{i=1}^{n}(\sqrt{2}|y_i|\theta_i - \log(\theta_i)) \\
&\geq \sum_{i=1}^{n}(1 + \log(|y_i|) - \log\sqrt{2} + \frac{1}{\sqrt{2}}|y_i|\theta_i) \\
&\geq n + \sum_{i=1}^{n}\log(|y_i|) - n\log\sqrt{2} + \frac{1}{\sqrt{2}}|y_1|\theta_1,
\end{aligned}
\tag{A.8}
$$

for all $\boldsymbol{\theta} \in \mathcal{S}$. From (A.8), the upper bound for θ_1 can be obtained as

$$\theta_1 \leq \frac{\sqrt{2}\big[\psi(\boldsymbol{\theta}_{(0)}) - \big(n + \sum_{i=1}^{n} \log(|y_i|)\big) + n \log\sqrt{2}\big]}{|y_1|}.$$

Therefore, for all $\boldsymbol{\theta} \in \mathcal{S}$, we have

$$0 < \theta_n \leq \cdots \leq \theta_1 \leq \frac{\sqrt{2}\big[\psi(\boldsymbol{\theta}_{(0)}) - \big(n + \sum_{i=1}^{n} \log(|y_i|)\big) + n \log\sqrt{2}\big]}{|y_1|},$$

which is finite by the assumption that the observed $|y_1|, \ldots, |y_n|$ are strictly positive.

A.2 Proof of the Rate of Convergence for the Penalized Estimator in Section 4

From (4.1), for normal erros, we have

$$\left| \frac{d}{d\alpha} \ell_\lambda(\bar{f} + \alpha g; \boldsymbol{x}, \boldsymbol{y}) \Big|_{\alpha=0} \right| \leq a_0 n^{-3/7} \|g - \bar{f}\| + \left| \frac{1}{4n} \sum_{i=n-k+1}^{n} (y_i - \widehat{\mu}_i)^2 g(x_i) \right|.$$

For g such that $M_0^{-1} \leq g(x) \leq M_0$, we have

$$\begin{aligned}
E\left[\frac{1}{4n} \sum_{i=n-k+1}^{n} (y_i - \widehat{\mu}_i)^2 g(x_i) \right] &\leq E\left[\frac{1}{4n} \sum_{i=n-k+1}^{n} (y_i - \mu_i)^2 g(x_i) \right] \\
&\leq E\left[\frac{1}{4n} \sum_{i=1}^{n} (y_i - \mu_i)^2 g(x_i) \right] \\
&= \frac{1}{4} \left\langle \boldsymbol{f}^{-1}, \boldsymbol{g} \right\rangle_n \\
&\leq \frac{1}{4} M_0^2,
\end{aligned}$$

from Assumptions (A1). Then, we have

$$\left| \frac{d}{d\alpha} \ell_\lambda(\bar{f} + \alpha g; \boldsymbol{x}, \boldsymbol{y}) \Big|_{\alpha=0} \right| \leq a_0 n^{-3/7} \|g - \bar{f}\| + \mathcal{O}_p(1),$$

and the remainder of the proof is identical to the proof of Theorem 3.

For double-exponential errors, we have

$$\left| \frac{d}{d\alpha} \ell_\lambda(\bar{f} + \alpha g; \boldsymbol{x}, \boldsymbol{y}) \Big|_{\alpha=0} \right| \leq a_0 n^{-3/7} \|g - \bar{f}\| + \left| \frac{1}{\sqrt{2}n} \sum_{i=n-k+1}^{n} |y_i - \widehat{\mu}_i| g(x_i) \right|.$$

Similarly, from Assumptions (A1),

$$E\left[\frac{1}{\sqrt{2n}}\sum_{i=n-k+1}^{n}\left|y_i - \widehat{\mu}_i\right|g(x_i)\right] \;\leq\; E\left[\frac{1}{\sqrt{2n}}\sum_{i=n-k+1}^{n}\left|y_i - \mu_i\right|g(x_i)\right]$$

$$\leq\; E\left[\frac{1}{\sqrt{2n}}\sum_{i=1}^{n}\left|y_i - \mu_i\right|g(x_i)\right]$$

$$=\; \left\langle \boldsymbol{f}^{-1}, \boldsymbol{g}\right\rangle_n$$

$$\leq\; M_0^2,$$

and the same rate of convergence is attained from the proof of Theorem 3.

A.3   Proof of Invariance Property of the LRT Statistics

In this section, we will show the scale invariant property of LRT statistics. Under the null hypothesis, we assume that $\mu_i \equiv 0$ without loss of generality. The normal likelihood function can be written as

$$L(\theta_0; \boldsymbol{y}) = \left(\frac{\theta_0}{2\pi}\right)^{n/2} \exp\left(-\frac{\theta_0}{2}\sum_{i=1}^{n} y_i^2\right), \tag{A.9}$$

and the maximum likelihood estimator of θ_0 is

$$\widehat{\theta}_0 = \frac{n}{\sum_{i=1}^{n} y_i^2}. \tag{A.10}$$

Let the maximum likelihood estimator of $\boldsymbol{\theta}$ under the alternative hypothesis be $\widehat{\boldsymbol{\theta}}_1 = (\hat{\theta}_{11}, \cdots, \hat{\theta}_{1n})$. This estimator $\widehat{\boldsymbol{\theta}}$ can be obtained through the convex programming using the constrained quadratic splines. From (5.4), (A.9), and (A.10), the likelihood ratio has the form

$$\Lambda(\boldsymbol{\theta}, \boldsymbol{y}) \;=\; \frac{L(\widehat{\theta}_0; \boldsymbol{y})}{L(\widehat{\boldsymbol{\theta}}_1; \boldsymbol{y})}$$

$$=\; \widehat{\theta}_0^{\frac{n}{2}}\left(\prod_{i=1}^{n}\widehat{\theta}_{1i}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\widehat{\theta}_0\sum_{i=1}^{n} y_i^2 + \frac{1}{2}\sum_{i=1}^{n} y_i^2\widehat{\theta}_{1i}\right)$$

$$=\; \left(\frac{n}{\sum_{i=1}^{n} y_i^2}\right)^{\frac{n}{2}}\left(\prod_{i=1}^{n}\widehat{\theta}_{1i}\right)^{-\frac{1}{2}} \exp\left(-\frac{n}{2} + \frac{1}{2}\sum_{i=1}^{n} y_i^2\widehat{\theta}_{1i}\right) \tag{A.11}$$

Then, the LR statistics has the form

$$T(\boldsymbol{y}) = -2 \ln \Lambda(\boldsymbol{\theta}, \boldsymbol{y}) = -n \ln n + n \ln \left(\sum_{i=1}^{n} y_i^2 \right) + \sum_{i=1}^{n} \ln \widehat{\theta}_{1i} + n - \sum_{i=1}^{n} y_i^2 \widehat{\theta}_{1i}. \qquad (A.12)$$

Now, we will show that the LR statistic (A.12) is independent from the unknown scale parameter θ_0 under the null hypothesis and $\boldsymbol{\theta}$ under the alternative hypothesis . To show the property of scale-invariance, suppose there are two sets of sample $\boldsymbol{y}_1 = (y_{11}, \cdots, y_{1n})$ and $\boldsymbol{y}_2 = (y_{21}, \cdots, y_{2n})$ where $y_{1i} \sim N(0, \sigma_1^2)$ and $y_{2i} = k y_{1i} \sim N(0, k^2 \sigma_1^2)$ for $i \in 1, \ldots, n$. Let $\theta_1 = 1/\sigma_1^2$ and $\theta_2 = 1/(k^2 \sigma_1^2)$. Under the null hypothesis, the ML estimators of θ_1 can be defined as

$$\widehat{\theta}_{01} = \frac{n}{\sum_{i=1}^{n} y_{1i}^2}, \qquad (A.13)$$

and the ML estimators of θ_2 can be written as

$$\widehat{\theta}_{02} = \frac{n}{\sum_{i=1}^{n} y_{2i}^2} = \frac{n}{\sum_{i=1}^{n} k^2 y_{1i}^2} = \frac{1}{k^2} \widehat{\theta}_{01}. \qquad (A.14)$$

Similarly, suppose the ML estimates of the reciprocal of variances for $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ under the alternative hypothesis are $\widehat{\boldsymbol{\theta}}_{11}$ and $\widehat{\boldsymbol{\theta}}_{12}$, respectively. Then, we can show that $\widehat{\boldsymbol{\theta}}_{11} = k^2 \widehat{\boldsymbol{\theta}}_{12}$. We have $\widehat{\boldsymbol{\theta}}_{11}$ which maximizes the log-likelihood function

$$\ell(\boldsymbol{\theta}; \boldsymbol{y}_1) = -\frac{n}{2} \ln(2\pi) + \frac{1}{2} \sum_{i=1}^{n} \left[\ln(\theta_{1i}) - y_{1i}^2 \theta_{1i} \right], \qquad (A.15)$$

while satisfies the monotonic constraints. Similarly, $\widehat{\boldsymbol{\theta}}_{12}$ maximizes the function

$$
\begin{aligned}
\ell(\boldsymbol{\theta}; \boldsymbol{y}_2) &= -\frac{n}{2} \ln(2\pi) + \frac{1}{2} \sum_{i=1}^{n} \left[\ln(\theta_{2i}) - y_{2i}^2 \theta_{2i} \right] \\
&= -\frac{n}{2} \ln(2\pi) + \frac{1}{2} \sum_{i=1}^{n} \left[\ln(k^2 \theta_{2i}) - \ln(k^2) - k^2 y_{1i}^2 \theta_{2i} \right] \\
&= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(k^2) + \frac{1}{2} \sum_{i=1}^{n} \left[\ln(k^2 \theta_{2i}) - y_{1i}^2 (k^2 \theta_{2i}) \right]. \qquad (A.16)
\end{aligned}
$$

Since k is a constant, we can conclude that $k^2\widehat{\boldsymbol{\theta}}_{12} = \widehat{\boldsymbol{\theta}}_{11}$. Then, from (A.22), (A.23), (A.15) and (A.24),

$$
\begin{aligned}
T(\boldsymbol{y}_2) &= -n\ln n + n\ln\Big(\sum_{i=1}^{n} y_{2i}^2\Big) + \sum_{i=1}^{n}\ln\widehat{\theta}_{2i} + n - \sum_{i=1}^{n} y_{2i}^2\widehat{\theta}_{2i} \\
&= -n\ln n + n\ln\Big(\sum_{i=1}^{n} k^2 y_{1i}^2\Big) + \sum_{i=1}^{n}\ln(\frac{1}{k^2}\widehat{\theta}_{1i}) + n - \sum_{i=1}^{n}(k^2 y_{1i}^2)\frac{1}{k^2}\widehat{\theta}_{1i} \\
&= -n\ln n + n\ln\Big(\sum_{i=1}^{n} y_{1i}^2\Big) + \sum_{i=1}^{n}\ln(\widehat{\theta}_{1i}) + n - \sum_{i=1}^{n} y_{1i}^2\widehat{\theta}_{1i} \\
&= T(\boldsymbol{y}_1).
\end{aligned}
\tag{A.17}
$$

Therefore, the LR test statistic is scale-invariant. From the scale-invariant property, the critical value for the hypothesis test is not depends on the scale parameter θ_0 and can be obtained from the simulations in section 5.3.

Similarly, the double-exponential likelihood function under the null hypothesis can be written as

$$
L(\theta_0; \boldsymbol{y}) = 2^{-n/2}\theta_0^n \exp\Big(-\sqrt{2}\theta_0\sum_{i=1}^{n} y_i^2\Big),
\tag{A.18}
$$

and the maximum likelihood estimator of θ_0 is

$$
\widehat{\theta}_0 = \frac{n}{\sum_{i=1}^{n}\sqrt{2}|y_i|}.
\tag{A.19}
$$

$$
\begin{aligned}
\Lambda(\boldsymbol{\theta}, \boldsymbol{y}) &= \frac{L(\widehat{\theta}_0; \boldsymbol{y})}{L(\widehat{\boldsymbol{\theta}}_1; \boldsymbol{y})} \\
&= \Big(\frac{n}{\sum_{i=1}^{n}\sqrt{2}|y_i|}\Big)^n\Big(\prod_{i=1}^{n}\widehat{\theta}_{1i}\Big)^{-1}\exp\Big(-n + \sqrt{2}\sum_{i=1}^{n}|y_i|\widehat{\theta}_{1i}\Big).
\end{aligned}
\tag{A.20}
$$

Then, the LR statistics has the form

84

$$T(\boldsymbol{y}) = -2\ln \Lambda(\boldsymbol{\theta}, \boldsymbol{y}) = -2\left(n\ln n + n\ln\left(\sum_{i=1}^{n}\sqrt{2}|y_i|\right) + \sum_{i=1}^{n}\ln\widehat{\theta}_{1i} + n - \sqrt{2}\sum_{i=1}^{n}|y_i|\widehat{\theta}_{1i}\right). \tag{A.21}$$

Assume y_{1i}'s are generated from the double-exponential distribution with mean zero and standard deviation σ_1, and $y_{2i} = ky_{1i}$. Let $\theta_1 = 1/\sigma_1$ and $\theta_2 = 1/\sigma_2 = \theta_1/|k|$. Then,

Under the null hypothesis, the ML estimators of θ_1 can be defined as

$$\widehat{\theta}_{01} = \frac{n}{\sum_{i=1}^{n}\sqrt{2}|y_{1i}|}, \tag{A.22}$$

and the ML estimators of θ_2 can be written as

$$\widehat{\theta}_{02} = \frac{n}{\sum_{i=1}^{n}\sqrt{2}|y_{2i}|} = \frac{n}{\sum_{i=1}^{n}\sqrt{2}|ky_{1i}|} = \frac{n}{\sum_{i=1}^{n}\sqrt{2}|k||y_{1i}|} = \frac{\widehat{\theta}_{01}}{|k|}. \tag{A.23}$$

Similarly, suppose the ML estimates of the reciprocal of standard deviation for $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ under the alternative hypothesis are $\widehat{\boldsymbol{\theta}}_{11}$ and $\widehat{\boldsymbol{\theta}}_{12}$, respectively. Then, we can show that $\widehat{\boldsymbol{\theta}}_{11} = |k|\widehat{\boldsymbol{\theta}}_{12}$.

From (A.18), $\widehat{\boldsymbol{\theta}}_{11}$ maximizes the log-likelihood function

$$\ell(\boldsymbol{\theta}; \boldsymbol{y}_1) = -n\ln(n) + \sum_{i=1}^{n}\left[\ln(\theta_{1i}) - \sqrt{2}|y_{1i}|\theta_{1i}\right], \tag{A.24}$$

while satisfies the monotonic constraints. Similarly, from (A.24), $\widehat{\boldsymbol{\theta}}_{12}$ maximizes the function

$$\begin{aligned}
\ell(\boldsymbol{\theta}; \boldsymbol{y}_2) &= -n\ln(n) + \sum_{i=1}^{n}\left[\ln(\theta_{2i}) - \sqrt{2}|y_{2i}|\theta_{2i}\right] \\
&= -n\ln(n) - n\ln|k| + \sum_{i=1}^{n}\left[\ln(|k|\theta_{2i}) - \sqrt{2}|y_{1i}||k|\theta_{2i}\right]
\end{aligned} \tag{A.25}$$

Since k is a constant, we can conclude that $|k|\widehat{\boldsymbol{\theta}}_{12} = \widehat{\boldsymbol{\theta}}_{11}$. Therefore, we have

$$
\begin{aligned}
T(\boldsymbol{y}_2) &= -2\left(n\ln n + n\ln\Big(\sum_{i=1}^{n}\sqrt{2}|y_{2i}|\Big) + \sum_{i=1}^{n}\ln\widehat{\theta}_{2i} + n - \sqrt{2}\sum_{i=1}^{n}|y_{2i}|\widehat{\theta}_{2i}\right) \\
&= -2\left(n\ln n + n\ln\Big(\sum_{i=1}^{n}\sqrt{2}|y_{1i}|\Big) + \sum_{i=1}^{n}\ln(\widehat{\theta}_{1i}) + n - \sqrt{2}\sum_{i=1}^{n}|y_{1i}|\widehat{\theta}_{1i}\right) \\
&= T(\boldsymbol{y}_1).
\end{aligned}
\tag{A.26}
$$