A COMPARISON OF NONLINEAR AND NONPARAMETRIC REGRESSION METHODS

MIN CHEN

A Comparison of Nonlinear and Nonparametric Regression Methods

by

Min Chen
B.Science (Memorial University of Newfoundland)

A practicum submitted to the
School of Graduate Studies
in partial fulllment of the
requirements for the degree of
Master of Applied Statistics.

Department of Mathematics and Statistics
Memorial University of Newfoundland

March 12, 2010

St. John's Newfoundland

Contents

Abstract

In this report, we investigate the performance of nonlinear regression and nonparametric regression with data set simulated under a nonlinear parametric model. First, we consider the nonlinear least squares estimation method for the model. Then, we apply various nonparametric regression methods such as kernel methods, spline smoothing, and wavelet version of estimators with the same model. The nonlinear least squares estimation method produces the best estimation in terms of MSE among all the regression methods. Both kernel methods and wavelet version of estimation methods produce reasonably small values of MSE. Moreover, the wavelet regression method performances best among all the nonparametric methods. The spline method produces unacceptablly large MSE due to large variance of estimation. The boundary issues do exist in all the nonparametric regression methods due to less density of data points.

Acknowledgements

I would like to show my deepest gratitude to my research supervisor Dr. A. J. Oyet for his valuable suggestions and patient guidance from the start to the end of this study. This project would not possible to be completed without his kind supports during my program.

I would also like to thank Dr G. Sneddon for his generous support through my program. I appreciate all the helps provided by members in Department of Mathematics and Statistics.

Finally, I appreciate Fang Fang for her accompanying through my whole academic career at Memorial University of Newfoundland.

List of Tables

List of Figures

Chapter 1

Introduction

One of the common tasks in statistics is to determine the functional relationship between an observed response and corresponding inputs. In regression analysis, the observed response variable is often referred to as dependent variable, and denoted by y. The inputs are usually called regressors or independent variables, and denoted by $\mathbf{x}$ which is a $k \times 1$ vector. The functional relation is usually expressed in the mathematical form,

$$y = f(x_1, x_2, \cdots, x_k; \theta_1, \theta_2, \cdots, \theta_p),$$

where $\theta_1, \theta_2, \cdots, \theta_p$ are unknown coefficients, usually called parameters, that have to be estimated from data set with observed responses, y, and associated inputs $x_1, x_2, \cdots, x_k$.

In linear regression analysis (see Seber [2003]), the model is expressed in the form

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1} + \varepsilon,$$

where ε is the error that is due to either unknown factors in the experiment or error in measurement. The x_i's can also be in the form of squares, cross products, higher powers, and even transformations (e.g. logarithms). The essential requirement of a linear model is that the mean response function should be linear in the unknown parameters β_j. That is, if $\frac{\partial f(x; \theta_1, \cdots, \theta_p)}{\partial \theta_j} = g(x)$, then $g(x)$ is independent of θ_j. However, there are several practical situations in which the mean response function is nonlinear in the parameters (see Seber and Wild [1989]). For instance,

$$y = \beta_0 + \beta_1 e^{-\beta_2 x} + \varepsilon$$

is a nonlinear model since it is nonlinear in β_2.

Linear and nonlinear models have been studied in great details by Gallan (1975), Bates and Watts(1988), and Seber and Wild (1989) among others. Various nonlinear models have been proposed and applied to different fields of study, such as in physics, biology, chemistry and agriculture. All nonlinear models can be classified into two distinct categories based on their properties (Seber and Wild, 1989), namely transformably linear and intrinsically nonlinear models. The nonlinear relationship between x and y can be transformed into a linear relationship for some nonlinear models. Those models are referred to as transformable nonlinear models. Consider the following model,

$$y = e^{\alpha+\beta x} + \varepsilon,$$

where ε is the associated error with $E(\varepsilon) = 0$ and $\mathrm{var}(\varepsilon) = \sigma$. If we take the logarithms of both sides of the above model, then

$$\log_{\mathrm{e}}^{y} = \alpha + \beta x + \varepsilon^*.$$

If we set $y^* = \log_{\mathrm{e}}^{y}$, $x^* = x$, $\beta_1^* = \beta$, and $\beta_0^* = \alpha$, then

$$y^* = \beta_0^* + \beta_1^* x^* + \varepsilon^*.$$

The transformed model is now a linear model. However, some nonlinear models can not be linearized by means of a suitable transformation. Those models are referred to as intrinsically nonlinear. An example of an intrinsically nonlinear model is

$$f(t) = \frac{\theta_1}{\theta_1 - \theta_2}\{\exp(-\theta_2 t) - \exp(-\theta_1 t)\}$$

Another example of a transformably linear model discussed in Bates and Watts (1988) is Michaelis-Menten model, commonly used to describe the relationship between the utilization of nitrite in bush beans and light intensity. The mathematical formula of the model is given by

$$f(x,\ \theta) = E(Y|x) = \frac{\theta_1 x}{\theta_2 + x + \theta_3 x^2}. \tag{1.1}$$

where x is the level of light intensity with values ranging from 0 to 170 with measurement unit $\mu\mathrm{E/m^2\,s}$, and Y represents the nitrite utilization of plant leaves with measurement unit nmol/g hr. It is apparently a nonlinear model. The Michaelis-Menten model (1.1) will be used in our current study for the purpose of comparing nonlinear and nonparametric estimation methods.

There are two fundamental ways, parametric approach and nonparametrical approach, to approximate the mean function. The parametric approach is usually used when the mathematical formula of the mean function is known. The nonparametric approach is adopted when there is no specific form for mean function. According to Tapia and Thompson (1978), there is an extensive debate between Pearson and Fisher in the twenties regarding which approach should be taken in data analysis. Pearson noted that parametric fitting had the issue of misspecification, which would lead to high model bias. On the other hand, Fisher pointed out that nonparametric fitting would create more variable estimates, especially for small sample size of observations.

Many researche work has been done to compare parametric and nonparametric fitting when the form of the mean response function is unknown. Azzalini, Bowman, and Härdle (1989) proposed a pseudo likelihood ratios test based on constructing an estimate of the Kullback-Leibler distance between the parametric model and the nonparametric model. The test was used to check the validity of the parametric model. Moreover, Härdle and Mammen (1993) proposed a bootstrapped statistic to compare nonparametric and parametric regression fits. The bootstrapped statistic, T_n, is defined in terms of squared deviation $\int(\hat{m}_h(x) - m_{\hat{\theta}}(x))^2dx$. In this statistic, the nonparametric estimate, $\hat{m}_h(x)$, is compared with the parametric estimate, $m_{\hat{\theta}}(x)$. Härdle and Mammen used T_n as a test statistic to check the appropriateness of the parametric model that is proposed to fit paired observations.

There are limited literatures on comparing the perfromance of parametric and nonparametric estimates when proposed models are appropriate. In the current study, we compare the performance of parametric and nonparametric smoothing under different methods of estimation.This report is organised as follows. In Chapter 2 we introduce the nonlinear least squares estimation method. The Chapter 3 and Chapter 4 will introduce various nonparametric estimation methods such as, the kernel methods, spline smoothing, and wavelet modeling. Simulation studies will be conducted in Chapter 5.

The mean squared error (MSE) is used as the criterion for comparing different methods of estimations since the MSE is a reasonable criterion that takes into account both the variance and the bias of an estimator (Rice, 2007). If T is an estimator of $\tau(\theta)$, then MSE is defined as,

$$\begin{aligned} MSE(T) &= E[T - \tau(\theta)]^2 \\ &= E[T - E(T) + E(t) - \tau(\theta)]^2 \\ &= Var(T) + [b(T)]^2. \end{aligned}$$

where $b(T)$ is the bias of the estimate given by

$$b(T) = E(T) - \tau(\theta).$$

Chapter 2

Nonlinear Regression

In this chapter we give a brief introduction to the theory of nonlinear regression. To be specific, we shall discuss the least squares method of estimation. Additional details can be found in Hartley (1961), Hartley and Booker (1965), Gallant (1975), Seber and Wild (1989).

2.1 Nonlinear Least Squares Estimation

Suppose that we are given n observations (x_i, y_i), $i = 1, 2 \cdots, n$, from a nonlinear model,

$$y_i = f(x_i, \theta) + \epsilon_i \qquad (i = 1, 2 \cdots, n), \tag{2.1}$$

where $f(x_i, \theta)$ is an already known nonlinear functional relationship, ϵ_i is the error term with expected value 0 and constant variance, θ is a p dimensional unknown parameter that need to be estimated, and x_i is an $n \times 1$ vector.

The nonlinear least-squares estimate of θ, denoted by $\hat{\theta}$, is obtained by the minimization of

$$S(\theta) = \sum_{i=1}^{n} [y_i - f(x_i, \theta)]^2. \tag{2.2}$$

The solution to $S(\theta)$ may have several relative minima in addition to the absolute minimum $\hat{\theta}$ since the mean response function, $f(x,\theta)$, is a nonlinear function. We shall assume the following:

A(1): The ϵ_i are i.i.d. with mean zero and variance $\sigma^2(\sigma^2 > 0)$.

A(2): For each i, $f(\mathbf{x}_i;\ \theta)$ is a continuous function of θ for $\theta \in \Theta$.

A(3): Θ is a closed, bounded (i.e. compact) subset of $\mathbb{R}^p$.

Any value $\theta \in \Theta$ that minimizes equation (2.2) is a least-squares estimator of θ. Under assumptions A(1) to A(3), Jennrich (1969) showed that there exists a measurable function $\hat{\theta} = \hat{\theta}(y_1, y_2, \cdots, y_n)$ of the observations which minimizes $S(\theta)$. Moreover, $\hat{\theta}$ is also the maximum-likelihood estimator if we assume that the ϵ_i's are normally distributed.

If we let $y = (y_1,\ y_2,\ \cdots,\ y_n)'$, and $f(\theta) = (f(x_1,\theta),\ f(x_2,\theta),\ \cdots,\ f(x_n,\theta))'$, the equation (2.2) could be expressed as,

$$S(\theta) = [y - f(\theta)]'\ [y - f(\theta)] \tag{2.3}$$

$$= \|y - f(\theta)\|^2 \tag{2.4}$$

In order to minimize $S(\theta)$ as a function of θ, it is common to take the derivative of the function with respect to θ, and let it equals to 0,

$$\frac{\partial\, S(\theta)}{\partial\, \theta}|_{\theta} = 0 \tag{2.5}$$

$$\Rightarrow \quad \sum_{i=1}^{n}[y_i - f(x_i,\theta)]\frac{\partial f(x_i,\theta)}{\partial\theta_r}|_{\theta=\hat{\theta}} = 0 \quad (r = 1,\ 2,\ \cdots,\ p). \tag{2.6}$$

The equation (2.6) is called the normal equation for the nonlinear model. In nonlinear estimation, numerical iterative methods are necessarily adopted to solve the normal equation.

The following notations will be used for convenience. Define

$$\mathbf{f}(\theta) = (f_1(\theta),\, f_2(\theta),\, \cdots,\, f_n(\theta))'$$

and

$$F_.(\theta) = \frac{\partial f(\theta)}{\partial \theta'} = \left[\left(\frac{\partial f_i(\theta)}{\partial \theta_j}\right)\right],$$

where $f_i(\theta) = f(x_i, \theta)$. For brevity, let

$$F_. = F_.(\theta) \;\; and \;\; \hat{F}_. = F_.(\hat{\theta}).$$

Therefore, the dot subscript denotes first derivatives. The normal equation (2.6) can then be expressed as,

$$\begin{aligned} 0 &= \sum_{i=1}^{n}[y_i - f_i(\theta)]\frac{\partial f_i(\theta)}{\partial \theta_r}|_{\theta=\hat{\theta}} \quad (r = 1,\, 2,\, \cdots,\, p) && (2.7) \\ &= \hat{F}_.'[y - f(\hat{\theta})] && (2.8) \\ &= \hat{F}_.'\hat{\varepsilon}. && (2.9) \end{aligned}$$

2.2 Statistical Properties of Least Squares Estimators

In the next few paragraphs, we demonstrate the linear approximation of nonlinear response functions. Asymptotical statistical properties of least squares estimator, $\hat{\theta}$, are obtained using linear approximation under certain regularity assumptions. See Seber and Wild (1989) for more details

Let θ^* be in a small neighborhood of the true value θ, then the linear Taylor series expansion of $f_i(\theta)$ is given by

$$f_i(\theta) \approx f_i(\theta^*) + \sum_{r=1}^{p}\frac{\partial f_i}{\partial \theta_r}|_{\theta_*}(\theta_r - \theta_r^*) \qquad (2.10)$$

or

$$f(\theta) \approx f(\theta^*) + F_.(\theta - \theta^*). \tag{2.11}$$

Let $z = y - f(\theta^*) = \varepsilon$ and $\beta = \theta - \theta^*$. Substitue equation (2.11) into equation (2.4), to obtain

$$\begin{aligned} S(\theta) &= \|y - f(\theta)\|^2 \\ &\approx \|y - f(\theta^*) - F_.(\theta - \theta^*)\|^2 \\ &= \|z - F_.\beta\|^2 \end{aligned} \tag{2.12}$$

The equation (2.12) is similar to the problem of minimizing $S(\beta) = \|y - X\beta\|^2$ in the linear regression model $y = X\beta + \varepsilon$. Therefore, from the properties of linear models, $S(\theta)$ is approximatly minimized when β is given by

$$\hat{\beta} = (F_.^{'} F_.)^{-1} F_.^{'} z.$$

When n is large, it has been shown that under certain regularity conditions $\hat{\theta}$ is almost certain to be within a small neighborhood of θ^* (Seber and Wild, 1989). Hence $\hat{\beta} \approx \hat{\theta} - \theta^*$ and

$$\hat{\theta} - \theta^* \approx (F_.^{'} F_.)^{-1} F_.^{'} \varepsilon \tag{2.13}$$

Moreover, with $\theta = \hat{\theta}$, in (2.11)

$$\begin{aligned} f(\hat{\theta}) - f(\theta^*) &\approx F_.(\hat{\theta} - \theta^*) \\ &\approx F_.(F_.^{'} F_.)^{-1} F_.^{'} \varepsilon \\ &= P_F \varepsilon \end{aligned} \tag{2.14}$$

Where $P_F = F_.(F_.^{'} F_.)^{-1} F_.^{'}$ is a symmetric and idempotent matrix.

Assuming that $\varepsilon \sim N(0,\ \sigma^2 I_n)$, and given the following regularity conditions in addition to A(1), A(2), and A(3), A(4): $n^{-1}B_n(\theta, \theta_1)$ converges uniformly for all θ, θ_1 in Θ to a function $B(\theta, \theta_1)$, and $D(\theta, \theta^*) = 0$ if and only if $\theta = \theta^*$. Where,

$$B_n(\theta, \theta_1) = \sum_{i=1}^{n} f_i(\theta) f_i(\theta_1), \qquad D_n(\theta, \theta^*) = \sum_{i=1}^{n} [f_i(\theta) - f_i(\theta^*)]^2.$$

A(5): θ^* is an interior point of Θ.

A(6): The first and second derivative, $\partial f_i(\theta)/\partial\theta_r$ and $\partial^2 f_i(\theta)/\partial\theta_r\partial\theta_s$ $(r, s = 1, 2, \cdots, p)$, exist and are continuous for all $\theta \in \Theta^*$.

A(7): $\frac{1}{n}\sum_{i=1}^{n}(\partial f_i/\partial\theta)(\partial f_i/\partial\theta')$ converges to some matrix $\Phi(\theta)$ uniformly in θ for $\theta \in \Theta^*$.

A(8): $\frac{1}{n}\sum_{i=1}^{n}[\partial^2 f_i(\theta)/\partial\theta_r\partial\theta_s]^2$ converges uniformly in θ for $\theta \in \Theta^*$ $(r, s = 1, 2, \cdots, p)$.

A(9): $\Phi \in \Phi(\theta^*)$ is nonsingular.

we have the following four properties regrading $\hat{\theta}$ for large n.

Theorem 1 *Let the $p \times 1$ random vector y be $N_p(\mu, \Sigma)$, let a be any $p \times 1$ vector of constants, and let A be any $k \times p$ matrix of constants with rank $k \leq p$. Then*

$$z = a'y \sim N(a'\mu,\ a'\Sigma a) \quad \text{and} \quad Z = Ay \sim N_k(A\mu,\ A\Sigma A').$$

Property 1 $\hat{\theta} - \theta^* \sim N_p(0,\ \sigma^2 C^{-1})$, *where* $C = F_.'F_. = F_.'(\theta^*)F_.(\theta^*)$.

Proof: Since $\hat{\theta} - \theta^* \approx (F_.'F_.)^{-1}F_.'\varepsilon$ (2.13) and $\varepsilon \sim N(0,\ \sigma^2 I_n)$, $(F_.'F_.)^{-1}F_.'\varepsilon$ will also follow normal distribution, according to Theorem 1, with,

$$\begin{aligned} E(\hat{\theta} - \theta^*) &\approx E[(F_.'F_.)^{-1}F_.'\varepsilon] \\ &\approx (F_.'F_.)^{-1}F_.'E(\varepsilon) \\ &\approx 0 \end{aligned}$$

$$\begin{aligned}\text{var}(\hat{\theta} - \theta^*) &\approx \text{var}[(F_.'F_.)^{-1}F_.'\varepsilon] \\ &\approx (F_.'F_.)^{-1}F_.'var(\varepsilon)F_.(F_.'F_.)^{-1} \\ &\approx \sigma^2(F_.'F_.)^{-1}\end{aligned}$$

Therefore, $\hat{\theta} - \theta^*$ approximatly follows a normal distribution with mean 0, and variance $\sigma^2(F_.'F_.)^{-1}$.

Theorem 2 *Let y be distributed as $N_p(\mu, \Sigma)$, let A be a $p \times p$ symmetric matrix of constants of rank r, and let $\lambda = \frac{1}{2}\mu' A\mu$. Then $y'Ay \sim \chi^2(r, \lambda)$ if and only if $A\Sigma$ is idempotent.*

Property 2 $\dfrac{(n-p)s^2}{\sigma^2} \approx \dfrac{\varepsilon'(I_n - P_F)\varepsilon}{\sigma^2} \sim \chi^2_{n-p}$.

Proof: The estimate of the variance of the errors, ε, corresponding to the least squares estimator $\hat{\theta}$ is

$$s^2 = \frac{1}{n-p}S(\hat{\theta}). \tag{2.15}$$

Moreover,

$$\begin{aligned}S(\hat{\theta}) &= \|y - f(\hat{\theta})\|^2 \\ &\approx \|y - f(\theta^*) - F_.(\hat{\theta} - \theta^*)\|^2 \qquad \text{substitute equation (2.11)} \\ &= \|\varepsilon - P_F\varepsilon\|^2 \\ &= \varepsilon'(I_n - P_F)\varepsilon, \end{aligned} \tag{2.16}$$

where $I_n - P_F$ is a symmetric and idempotent matrix of rank $n - p$. Now, substitute equation (2.16) into equation (2.15), and we have

$$s^2 \approx \frac{\varepsilon'(I_n - P_F)\varepsilon}{n-p} \tag{2.17}$$

or

$$\frac{(n-p)s^2}{\sigma^2} \approx \frac{\varepsilon'(I_n - P_F)\varepsilon}{\sigma^2} \tag{2.18}$$

Since both ε' and ε follow normal distributions, with result in Theorem 2, $\varepsilon'(I_n - P_F)\varepsilon$ follows chi-square distribution with $n-p$ degrees of freedom that applies to property 2.

Property 3 *$\hat{\theta}$ is statistically independent of s^2.*

Proof: Since

$$\begin{aligned} S^2 &= \frac{S(\hat{\theta})}{n-p} \\ &= \frac{\|z - F_.\hat{\beta}\|^2}{n-p}, \end{aligned}$$

it is sufficicent to show that $\hat{\beta}$ and $z - F_.\hat{\beta}$ are statistically independent.

$$\begin{aligned} & \operatorname{cov}(\hat{\beta},\ z - F_.\hat{\beta}) \\ = \ & \operatorname{cov}((F_.'F_.)^{-1}F_.'z,\ z - F_.(F_.'F_.)^{-1}F_.'z) \\ = \ & \operatorname{cov}((F_.'F_.)^{-1}F_.'z,\ (I - P_F)z) \\ = \ & (F_.'F_.)^{-1}F_.'cov(z)(I - P_F)' \\ = \ & \sigma^2(F_.'F_.)^{-1}F_.'(I - P_F)' \\ = \ & 0 \end{aligned}$$

Moreover, $\hat{\beta}$ and $z - F_.\hat{\beta}$ follow multivariate normal distributions. Therefore, they are statistically independent which in turn suggests that $\hat{\theta}$ and s^2 are independent as well.

Property 4

$$\frac{[S(\theta^*) - S(\hat{\theta})]/p}{S(\hat{\theta})/(n-p)} \approx \frac{\epsilon' P_F \epsilon}{\epsilon'(I_n - P_F)\epsilon} \cdot \frac{n-p}{p} \sim F_{p,n-p}$$

Proof: Using equation (1.17), we have

$$\begin{aligned} &= S(\theta^*) - S(\hat{\theta}) \\ &= \|y - f(\theta^*)\|^2 - \epsilon'(I_n - P_F)\epsilon \\ &= \|\epsilon\|^2 - \epsilon'(I_n - P_F)\epsilon \\ &= \epsilon' P_F \epsilon \end{aligned} \tag{2.19}$$

According to property 2, we have

$$\frac{\epsilon' P_F \epsilon}{\sigma^2} \sim \chi_p^2$$

and

$$\frac{\epsilon'(I_n - P_F)\epsilon}{\sigma^2} \sim \chi_{n-p}^2$$

Since p and $n-p$ are constant, now, we have,

$$\frac{\frac{\epsilon' P_F \epsilon}{\sigma^2}/p}{\frac{\epsilon'(I_n - P_F)\epsilon}{\sigma^2}/n-p} = \frac{\epsilon' P_F \epsilon}{\epsilon'(I_n - P_F)\epsilon} \cdot \frac{n-p}{p} \sim F_{p,n-p}$$

or

$$\frac{[S(\theta^* - S(\hat{\theta})]/p}{S(\hat{\theta})/(n-p)} \sim F_{p,n-p}.$$

In summary, the random variable $\hat{\theta}$ has a p-dimensional multivariate normal distribution with mean θ and covariance matrix $\sigma^2(F_.' F_.)^{-1}$, and that $(n-p)s^2/\sigma^2$ has a chi-square distribution with $n-p$ degrees of freedom. As in linear regression, the practical importance of these properties is in their use to construct confidence intervals for the unknown parameter θ.

2.3 Gauss-Newton Method

The Gauss-Newton method is widely used for computing nonlinear least squares estimators. It is based on the substitution of the first-order Taylor series expansion in the least squares estimation formula. Suppose $\theta^{(a)}$ is an approximation to the least-squares estimate $\hat{\theta}$. For θ close to $\theta^{(a)}$, we have

$$f(\theta) \approx f(\theta^{(a)}) + F_{.}^{(a)}(\theta - \theta^{(a)}) \tag{2.20}$$

where $F_{.}^{(a)} = F_{.}(\theta^{(a)})$. Let $r(\theta)$ be the residual vector given by

$$\begin{aligned} r(\theta) &= y - f(\theta) \\ &\approx y - f(\theta^{(a)}) - F_{.}^{(a)}(\theta - \theta^{(a)}) \\ &= r(\theta^{(a)}) - F_{.}^{(a)}(\theta - \theta^{(a)}). \end{aligned} \tag{2.21}$$

Now, substitute equation (2.21) in the nonlinear least squares estimating equation (2.3) to obtain

$$\begin{aligned} S(\theta) &= r'(\theta)r(\theta) \\ &= \left[r(\theta^{(a)}) - F_{.}^{(a)}(\theta - \theta^{(a)})\right]' \left[r(\theta^{(a)}) - F_{.}^{(a)}(\theta - \theta^{(a)})\right] \\ &= r'(\theta^{(a)})r(\theta^{(a)}) - 2r'(\theta^{(a)})F_{.}^{(a)}(\theta - \theta^{(a)}) + (\theta - \theta^{(a)})'F_{.}^{'(a)}F_{.}^{(a)}(\theta - \theta^{(a)}). \end{aligned}$$

To minimize $S(\theta)$, we take the derivative of the above equation with respect to $\theta - \theta^{(a)}$, equal to 0, and simply to obtain

$$\begin{aligned} \theta - \theta^{(a)} &= (F_{.}^{'(a)}F_{.}^{(a)})^{-1}F_{.}^{'(a)}r(\theta^{(a)}) \\ &= \delta^{(a)}. \end{aligned}$$

This suggests that , with current approximation $\theta^{(a)}$, the next approximation is given by:

$$\theta^{(a+1)} = \theta^{(a)} + \delta^{(a)}. \tag{2.22}$$

The equation (2.22) provides an iterative algorithm,which is called Gauss-Newton method, for obtaining $\hat{\theta}$. The Gauss-Newton algorithm is convergent if the initial value $\theta^{(1)}$ is close to the true value θ^{*} and n is large enough.

Chapter 3

Nonparametric Regression

Nonlinear regression is frequently refered to as parametric regression in which the specific mathematical relation between x and y is already known to us. However, there are situations where we have little or no prior knowledge about the regression curve. Nonparametric regression methods are best suited for this situation.

Let the design points X_i be fixed. Let $\{(X_i,\, Y_i)\}_{i=1}^n$ be n pairs of observations. The regression relationship between X_i and Y_i can be expressed as:

$$Y_i = m(X_i) + \varepsilon_i, \qquad i = 1, \cdots, n, \tag{3.1}$$

where the regression function m is unknown, and $E(\varepsilon_i) = 0$.

Estimating the mean response curve m involves smoothing of the dataset $\{(X_i,\, Y_i)\}_{i=1}^n$. The basic idea of smoothing is local averaging procedure which chooses the mean of the response variables near a point x. The procedure is formally defined as

$$\hat{m}(x) = \sum_{i=1}^{n} W_i(x) Y_i \tag{3.2}$$

where $\hat{m}(x)$ denotes the estimated regression mean function, and $W_i(x)$ are weights which may depend on the whole vector $\{X_i\}_{i=1}^n$.

The following sections are devoted to a general introduction to nonparametric regression. To be specific, it includes Kernel estimation and Natural Spline Estimation.

Additional details can be found in Härdle (1989), Wahba and Wold (1975), and Stone (1974).

3.1 Kernel Estimation

A conceptually simple representation of the weights is to express the shape of the weight function $W_i(x)$ by a density function with a scale parameter that adjusts the size and the form of the weights near x. The density function is often referred to as a kernel K. The kernel function is choosen such that it is a continuous, bounded and symmetric real function. Since it is a density function,

$$K(u) \geq 0, \qquad \int K(u)du = 1. \tag{3.3}$$

These properties ensure that $\hat{m}(x)$ will always be a probability density. Moreover,

$$\int uK(u)du = 0, \tag{3.4}$$

$$\int u^2 K(u)du < \infty. \tag{3.5}$$

Gaussian kernel is a popular choice of univariate kernels with unbounded support. It is given by the density function

$$K(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}, \quad x \in \mathbb{R}. \tag{3.6}$$

Other choices of univariate kernels are the compactly supported "polynomial" kernels,

$$K(x) = k_{rs}(1 - |x|^r)^s I(|x| \leq 1), \tag{3.7}$$

where

$$k_{rx} = \frac{r}{2\text{Beta}(s+1,\ 1/r)}, \quad r > 0,\ s \geq 0,$$

and, $I(|x| \leq 1)$ is an indicator function. The "polynomial" kernels have support on $[-1,\ 1]$. The Epanechnikov kernel is obtained if $r = 2,\ s = 1$ ($k_{21} = 0.75$),

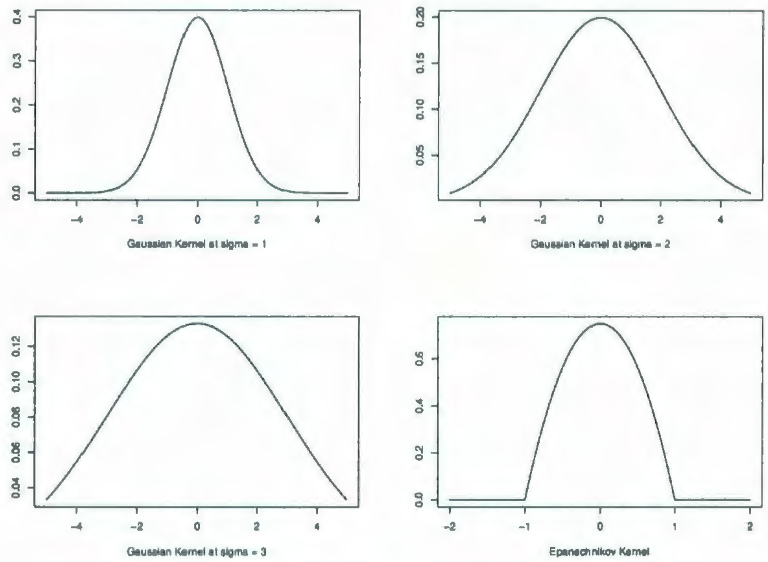

Figure 3.1: Plot of Gaussian Kernel at different sigmas and Epanechnikov Kernel

$$K(x) = 0.75(1 - x^2)I(|x| \leq 1). \tag{3.8}$$

Plots of Epanechnikov Kernel, and Gaussian Kernel at different values of σ are given in Figure 3.1. Both of the kernels are symmetric unimodal densities.

For one-dimensional x, the kernel density estimator of the marginal density of X is defined by,

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i) \tag{3.9}$$

where

$$K_h(u) = \frac{K(u/h)}{h},$$

and h is a scale factor which is usually referred to as bandwidth. See Takezawa (2006) for more details regarding the derivations in following paragraphs.

Let $f(x)$, $f(y)$ be probability density functions (pdf) of X and Y, and $f(y|x)$ is the conditional pdf of Y given X, and $f(x,y)$ the joint pdf of X and Y. Then,

$$\begin{aligned} m(x) &= \mathrm{E}(Y|X=x) \\ &= \int_{-\infty}^{\infty} y f(y|x) dy \\ &= \int_{-\infty}^{\infty} y \frac{f(x,y)}{f(x)} dy \\ &= \frac{1}{f(x)} \int_{-\infty}^{\infty} y f(x,y) dy \end{aligned} \tag{3.10}$$

The kernel density estimate, $\hat{f}(x,y)$ and $\hat{f}(x)$, are given by equation (3.9),

$$\hat{f}(x,y) = \frac{1}{nh_x h_y} \sum_{i=1}^{n} K\left(\frac{x-X_i}{h_x}\right) K\left(\frac{y-Y_i}{h_y}\right) \tag{3.11}$$

$$\hat{f}(x) = \frac{1}{nh_x} \sum_{1}^{n} K\left(\frac{x-X_i}{h_x}\right) \tag{3.12}$$

Now, substitute equation (3.11) and (3.12) into (3.10), we have,

$$\begin{aligned} \hat{m}(x) &= \int_{-\infty}^{\infty} y \frac{\hat{f}(x,y)}{\hat{f}(x)} dy \\ &= \int_{-\infty}^{\infty} y \frac{\frac{1}{nh_x h_y} \sum_{i=1}^{n} K\left(\frac{x-X_i}{h_x}\right) K\left(\frac{x-Y_i}{h_y}\right)}{\frac{1}{nh_x} \sum_{i=1}^{n} K\left(\frac{x-X_i}{h_x}\right)} dy \\ &= \frac{\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_x}\right) \int_{-\infty}^{\infty} yK\left(\frac{y-Y_i}{h_y}\right) dy}{h_y \sum_{i=1}^{n} K\left(\frac{x-X_i}{h_x}\right)}. \end{aligned} \tag{3.13}$$

Let $\eta = \frac{y-Y_i}{h_y}$, then

$$\begin{aligned} &\frac{1}{h_y} \int_{-\infty}^{\infty} yK\left(\frac{y-Y_i}{h_y}\right) dy \\ &= \int_{-\infty}^{\infty} (h_y \eta + Y_i) K(\eta) d\eta \\ &= Y_i. \end{aligned} \tag{3.14}$$

Now, substitute equation (3.14) into (3.13), we finally have,

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right) Y_i}{\sum_{k=1}^{n} K\left(\frac{x-X_k}{h}\right)} = \sum_{i=1}^{n} W_i(x) Y_i. \tag{3.15}$$

Where,

$$W_i(x) = \frac{K\left(\frac{x-X_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{x-X_k}{h}\right)}.$$

The kernel estimate (3.15) is frequently referred to as Nadaraya-Watson estimator, and was proposed by Nadaraya (1964) and Watson (1964).

Gasser and Müller (1979) suggested a weight sequence

$$W_h(x) = \int_{S_{i-1}}^{S_i} K_h(x-u)du \tag{3.16}$$

Where S_i is the mid-point of ordered X-data, X_{i-1} and X_i. Let $X_1 \leq X_2 \leq, \ldots, \leq X_n$, then

$$S_i = \frac{X_i + X_{i+1}}{2}, \qquad i = 1, 2, \cdots, n-1$$

With these weights, the Gasser-Müller smoother is given by

$$\hat{m}_h(x) = \sum_{i=1}^{n} Y_i \int_{S_{i-1}}^{S_i} K_h(x-u)du. \tag{3.17}$$

Two approaches, L_1 and L_2, are usually used to access the consistency of non-parametric estimators (Izenman, 1991). The L_1 approach is based on the integrated absolute error (IAE) which is defined as

$$\text{IAE} = \int_{-\infty}^{\infty} |\hat{m}(x) - m(x)|dx.$$

If IAE $\rightarrow 0$ in probability as $n \rightarrow \infty$, then $\hat{m}(x)$ is said to be a consistent estimator of $m(x)$; strong consistency of $\hat{m}(x)$ occurs when convergence holds almost surely. Under this approach, Devroye (1983) proved that if the kernel function, K, is a probability density function satisfying (3.3), then the kernel estimator is a strongly consistent estimate of $m(x)$ if and only if $h \rightarrow 0$ and $nh^k \rightarrow \infty$, as $n \rightarrow \infty$.

The L_2 approach is based on mean squared error. If $m(x)$ is square integrable, then MSE of $\hat{m}(x)$ is

$$\begin{aligned} \text{MSE} &= \text{E}[\hat{m}(x) - m(x)]^2 \\ &= \text{var}[\hat{m}(x)] + \{\text{bias}[\hat{m}(x)]\}^2 \end{aligned}$$

If MSE$\to$ 0 as $n \to \infty$, then $\hat{m}(x)$ is a pointwise consistent estimator of $m(x)$. Under this approach, and the regularity conditions, if bandwidth (h) is a function of sample size (n) satisfying $\lim_{n\to\infty} nh^2 = \infty$, and if $m(x)$ is uniformly continuous, Parzen (1962) showed that if $h \to 0$ as $n \to \infty$, the univeriate kernel estimator was asymptotically unbiased and asymptotically normal.

3.2 Spline Smoothing

A spline function is a curve constructed from piecewise polynomials that are subject to conditions of continuity at their joints. The joints are known as knots or nodes. The spline smoothing approach provides a balance between the aim to produce a good fit to the data and the aim to produce a curve without too much rapid local variation.The goodness of fit is assessed by the residual sum of squares

$$\sum_{i=1}^{n}(Y_i - m(x_i))^2.$$

Whereas the local variation is measured by the integrated squared second derivative whch is called the roughness penalty

$$\int (m''(x))^2 dx.$$

Using those measures, the weighted sum is defined as

$$S_\lambda(m) = \sum_{i=1}^{n}(Y_i - m(x_i))^2 + \lambda \int (m''(x))^2 dx \tag{3.18}$$

where λ denotes a smoothing parameter which represents the rate of exchange between residual error and roughness of the curve m. The equation (3.18) is also frequently referred to as penalized least squares method.

The fitted spline is the solution to the optimization problem that minimizes $S_\lambda(\cdot)$ over the class of all twice differentiable functions on the interval $[a, b] = [X_{(1)}, X_{(n)}]$.

There is a unique solution $\hat{m}_\lambda(x)$ that is defined as a cubic spline. The estimated curve $\hat{m}_\lambda(\cdot)$ has the following three properties:

1. $\hat{m}_\lambda(x)$ is a cubic polynomial between any two successive X values (X_i, X_{i+1}).

2. $\hat{m}_\lambda(x)$, $\hat{m}'_\lambda(x)$, $\hat{m}''_\lambda(x)$ are continuous at the observation points X_i.

3. $\hat{m}_\lambda(x)$ is zero at boundary points $X_{(1)}$ and $X_{(n)}$.

Since $m_\lambda(x)$ is a cubic polynomial between (X_i, X_{i+1}), it can be expressed as

$$m_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \tag{3.19}$$

where x ranges from x_i to x_{i+1}.

The first and second derivatives of this function are

$$m'_i(x) = 3a_i(x - x_i)^2 + 2b_i(x - x_i) + c_i \tag{3.20}$$

and

$$m''_i(x) = 6a_i(x - x_i) + 2b_i. \tag{3.21}$$

Since $m(x)$ is a continuous function, the adjacent functions S_i and S_{i+1} $(i = 1, \cdots, n)$ should meet the following five conditions (3.22) through (3.26).

$$m_i(x) = Y_i, \qquad \text{where } i = 1, \ldots, n-1 \tag{3.22}$$

$$m_{n-1}(x_n) = Y_n. \tag{3.23}$$

$$m_i(x_{i+1}) = m_{i+1}(x_{i+1}), \qquad \text{where } i = 1, \ldots, n-2 \tag{3.24}$$

The condition that the first and second derivaties should be equal at the knots is expressed in equation (3.20) and (3.21) as

$$m'_i(x_{i+1}) = m'_{i+1}(x_{i+1}), \qquad \text{where } i = 1, \ldots, n-2 \tag{3.25}$$

$$m_i''(x_{i+1}) = m_{i+1}''(x_{i+1}), \qquad \text{where } i = 1, \ldots, n-2 \tag{3.26}$$

Now, let the bandwidth of the ith interval be $h_i = x_{i+1} - x_i$. Note that there are $n-1$ intervals for n points. Therefore, there are $(n-1)$ pieces of $m_i(x)$ which give $(n-1)$ unknown a_i, b_i, c_i, and d_i that need be computed.

From (3.22) and (3.23),

$$m_i(x) = Y_i = d_i \tag{3.27}$$

and

$$\begin{aligned} m_{i+1}(x_{i+1}) &= Y_{i+1} \\ &= a_{i+1}(x_{i+1} - x_i)^3 + b_{i+1}(x_{i+1} - x_i)^2 + c_{i+1}(x_{i+1} - x_i) + Y_i \\ &= a_{i+1}h_i^3 + b_{i+1}h_i^2 + c_{i+1}h_i + Y_i. \end{aligned} \tag{3.28}$$

From (3.24),

$$Y_{i+1} = a_i h_i^3 + b_i h_i^2 + c_i h_i + Y_i. \tag{3.29}$$

The following notations will be used for convenience. Let $S_{i-1} = m_i''(x_{i-1})$ and $S_n = m_{n-1}''(x_n)$, then from equation (3.20 and 3.21), we have

$$\begin{aligned} S_i &= m_i''(x_i) \\ &= 6a_i(x - x_i) + 2b_i \\ &= 2b_i \end{aligned}$$

and

$$\begin{aligned} S_{i+1} &= m_i''(x_{i+1}) \\ &= 6a_i h_i + 2b_i \\ &= 6a_i h_i + S_i. \end{aligned}$$

It follows that

$$b_i = \frac{S_i}{2} \tag{3.30}$$

and

$$a_i = \frac{S_{i+1} - S_i}{6h_i}. \tag{3.31}$$

Substitute a_i and b_i in equation (3.29) to obtain

$$Y_{i+1} = \left(\frac{S_{i+1} - S_i}{6h_i}\right) h_i^3 + \frac{S_i}{2} h_i^2 + c_i h_i + Y_i.$$

Solving the above equation for c_i we have

$$c_i = \frac{Y_{i+1} - Y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6}. \tag{3.32}$$

It is clear that (3.25) also implies that,

$$m'_{i-1}(x_i) = m'_i(x_i).$$

Therefore, from equation (3.20) and (3.21).

$$m'_i(x_i) = c_i \tag{3.33}$$

$$m'_{i-1}(x_i) = 3a_{i-1}h_{i-1}^2 + 2b_{i-1}h_{i-1} + c_{i-1}. \tag{3.34}$$

From equation (3.32) and (3.33), we have

$$m'_i(x_i) = \frac{Y_{i+1} - Y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6} \tag{3.35}$$

and from equation (3.30), (3.31), (3.32), and (3.34) we have,

$$m'_{i-1}(x_i) = 3\left(\frac{S_i - S_{i-1}}{6h_{i-1}}\right) h_{i-1}^2 + 2\frac{S_{i-1}}{2} h_{i-1} + \frac{Y_i - Y_{i-1}}{h_{i-1}} - \frac{2h_{i-1}S_{i-1} + h_{i-1}S_i}{6}. \tag{3.36}$$

Combining equation (3.35) and (3.36), we obtain,

$$h_{i-1}S_{i-1} + (2h_{i-1} + 2h_i) + h_i S_{i+1} = 6\left(\frac{Y_{i+1} - Y_i}{h_i} - \frac{Y_i - Y_{i-1}}{h_{i-1}}\right) \tag{3.37}$$

Clearly, equation (3.37) gives $n-2$ equations for unknown S_i for $i = 2, \ldots, n-1$. There are two additional equations that are required for S_1 and S_n. The most frequently used choice is to let both S_1 and S_n equal to 0. This gives the natural spline. Using this choice we obtain a bidiagonal system,

$$\begin{bmatrix} 2(h_1+h_2) & h_2 & 0 & \ldots & 0 & 0 \\ h_2 & 2(h_2+h_3) & h_3 & 0 & \ldots & 0 \\ 0 & h_3 & 2(h_3+h_4) & h_4 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \end{bmatrix} \begin{bmatrix} S_2 \\ S_3 \\ \vdots \\ S_{n-1} \end{bmatrix} =$$

$$\begin{bmatrix} 6\left(\frac{Y_3-Y_2}{h_2} - \frac{Y_2-Y_1}{h_1}\right) \\ 6\left(\frac{Y_4-Y_3}{h_3} - \frac{Y_3-Y_2}{h_2}\right) \\ \vdots \\ 6\left(\frac{Y_n-Y_{n-1}}{h_{n-1}} - \frac{Y_{n-1}-Y_{n-2}}{h_{n-2}}\right) \end{bmatrix}$$

which may be solved by backsubstitution to obtain the values of a_i, b_i, c_i, and d_i.

3.3 Cross Validation

The accuracy of nonparametric smoothers as estimators of the function $m(x)$ depend on the smoothing parameter, h. Therefore, choice of the smoothing parameter is very important. One frequently used criteria is to choose the h to minimize the average mean squared error (AMSE),

$$R(h) = \frac{1}{n} E\left(\sum_{i=1}^{n} (\hat{m}(x_i) - m(x_i))^2\right).$$

$R(h)$ is a function of $m(x)$ which is unknown. A popular method is to estimate $R(h)$ by the leave-one out method. It is based on regression smoothers in which one observation is left out. The estimation is given by the function,

$$CV(h) = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \hat{m}_{(-i)}(x_i)\right)^2 \tag{3.38}$$

where,

$$\hat{m}_{(-1)}(x_i) = \sum_{j \neq i}^{n} W_j(x_i) Y_j$$

which is the regression smoother that omits the ith observation (x_i, y_i). The function CV is commonly called a cross-validation function since it validates the ability to predict $\{Y_j\}_{i=1}^{n}$ across the subsamples $\{(X_j,\ Y_j)\}_{j \neq i}$. For the best accuracy of the smoothing, h is chosen to minimize cross-validation function.

Wahba and Wold (1975) proposed a similar technique in the context of spline smoothing. Let $\hat{S}_\lambda(x_i)$ be a spline estimator of y_i. The cross-validation function of λ is given by the average squared error,

$$CV(\lambda) = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \hat{S}_\lambda^{(-i)}(x_i) \right)^2 \qquad (3.39)$$

where, $S_\lambda^{(-i)}(x_i)$ is the solution to the penalized least squares equation (see [3.18]) with the omission of the ith observation $(x_i,\ y_i)$. The smoothing parameter λ is chosen to minimize $CV(\lambda)$.

Chapter 4

Wavelet Methods for Regression

4.1 Some Backgrounds on Wavelets

This section is devoted to a brief introduction to the definition and theory of wavelets that will be used in our study. For more information, including proofs of the theorems in full generality and more extensive discussions, see Oyet (2002), Antoniadis, Gregoire, and McKeague (1994), Härdle, Kerkyacharian, and Tsubakov (1998), and Vidakovic (1999).

Two basic functions are required to define wavelets. The first one is the scaling function $\varphi(x)$ given by

$$\varphi(x) = \sum_{k \in \mathbb{Z}} c_k \varphi(2x - k) \tag{4.1}$$

with normalization

$$\int_{-\infty}^{\infty} \varphi(x) dx = 1.$$

The primary wavelet $\psi(x)$ is defined in terms of the scaling function as

$$\psi(x) = \sum_{k \in \mathbb{Z}} (-1)^k c_{k+1} \psi(2x + k) \tag{4.2}$$

where $\psi(x)$ satisfies $\int_{-\infty}^{\infty} \psi(x) dx = 0$. The coefficients $\{c_k,\ k \in \mathbb{Z}\}$ are called filter coefficients.

A wavelet system is defined as the infinite collection of translated and scaled versions $\{\varphi_{j,k}(x),\ \psi_{j,k}(x),\ j,\ k \in \mathbb{Z}\}$ of a scaling function $\varphi(x)$ and a primary wavelet $\psi(x)$ where

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k), \qquad j, k \in \mathbb{Z}, \tag{4.3}$$

and

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), \qquad j, k \in \mathbb{Z} \tag{4.4}$$

Two essential properties of the filter coefficients are normalization and orthogonality. The existence of a unique $L^1(R)$ solution to (4.1) is ensured by the normalization condition

$$\sum_{k\in\mathbb{Z}} c_k = 2.$$

With conditions

$$\sum_k c_k c_{k+2l} = 2 \quad \text{if } l = 0,$$

and

$$\sum_k c_k c_{k+2l} = 0 \quad \text{if } l \in \mathbb{Z},\ l \neq 0,$$

it ensures that $\{\psi_{j,k}(x),\ j,\ k \in \mathbb{Z}\}$ is an orthonormal basis of $L^2(R)$, and $\{\varphi_{j,k}(x),\ j,\ k \in \mathbb{Z}\}$ is an orthonormal system in $L^2(R)$ for each $j \in \mathbb{Z}$.

Therefore, any $f(x) \in L^2(R)$ can be written as

$$f(x) = \sum_{j\in Z}\sum_{k\in Z} d_{j,k}\psi_{j,k}(x). \tag{4.5}$$

The relation (4.5) is called homogeneous wavelet expansion.

Let the spaces spanned by $\varphi_{j,k}(x)$ and $\psi_{j,k}(x)$ over the parameter k, with j fixed, be denoted by V_j and W_j respectively,

$$V_j = span_{k\in Z}\varphi_{j,k}(x),$$

$$W_j = span_{k\in Z}\psi_{j,k}(x).$$

It can be shown that the spaces V_j and W_j are related by

$$\cdots\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots\cdots$$

The nested spaces V_j have the following properties,

(a) An intersection that is trivial.

$$\bigcap_{j\in Z} V_j = \{0\}$$

.

(b) A union that is dense in $L^2(R)$.

$$\overline{\bigcup_{j\in Z} V_j} = L^2(R).$$

The spaces V_j and W_j have the following relationship,

$$V_{j+1} = V_j \oplus W_j.$$

Within the larger space V_{j+1}, W_j is the orthogonal complement of V_j. In other word, any function $f(x) \in V_{j+1}$ could be expressed as a linear combinaton or direct sums of functions in V_j and W_j. It can be verified by iteration,

$$V_{j+1} = V_0 \oplus_{i=0}^{j} W_i$$

.

Therefore, for any fixed j_0 the decomposition $L^2(R) = V_{j0} \oplus \bigoplus_{j=j0}^{\infty} W_j$ corresponds to the representation

$$f(x) = \sum_{k\in Z} c_{j0,k}\varphi_{j0,k}(x) + \sum_{j\geq j0}\sum_{k\in Z} d_{j,k}\psi_{j,k}(x), \tag{4.6}$$

Where the coefficients are given by

$$c_{j0,k} = \int_{-\infty}^{\infty} f(x)\varphi_{j0,k}(x)dx, \qquad d_{j,k} = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx$$

The relation (4.6) is called inhomogeneous wavelet expansion.

Mallat (1989) introduced the notion of a multiresolution analysis, which is one of the most important concepts in discrete wavelet theory.

Definition 1 *A multiresolution analysis of $L^2(R)$ consists of an increasing sequence of closed subspaces $V_j, j \in Z$, of $L^2(R)$ such that*

(a) $\cap V_j = \{0\}$;

(b) $\overline{\cup V_j} = L^2$;

(c) there exists a scaling function $\varphi(x) \in V_0$ such that $\{\varphi(x-k),\ k \in Z)\}$ is an orthonormal basis of V_0

(d) for all $k \in Z$, $f(2^j x) \in V_j \Leftrightarrow f(2^j x - k) \in V_j$, and

(e) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$.

The intuitive meaning of (e) is that in passing from V_j to V_{j+1}, the resolution of the approximation is doubled. Mallat (1989) has shown that given any multiresolution analysis, it is possible to derive any function $\varphi(x)$ such that the family $\{\varphi_{j,k}(x), k \in Z\}$ is an orthonormal basis of the orthogonal complement W_j of V_j in V_{j+1}, so that $\{\varphi_{j,k}(x), j, k \in Z\}$ is an orthonormal basis of $L^2(R)$. The equation (4.6) is called a multiresolution expansion of $f(x)$, if $\{\varphi_{j,k}(x), k \in Z\}$ is a general basis for W_j.

4.2 Some Important Wavelet Bases

Several versions of wavelets have already been introduced in wavelet literatures. This section will only introduce two important versions of wavelets that will be used in this study. To be specific, they are the Haar and Daubechies' wavelets.

4.2.1 Haar Wavelet

The simplest example of a wavelet system is the Haar system. The Haar scaling function is defined by

$$\varphi(x) = \begin{cases} 1, & 0 \leq x < 1; \\ 0, & \text{otherwise.} \end{cases}$$

and the primary wavelet can also be described by a step function,

$$\psi(x) = \begin{cases} 1, & 0 \leq x < 1/2; \\ -1, & 1/2 \leq x < 1; \\ 0, & \text{otherwise.} \end{cases}$$

The disadvantage of the Haar wavelet is that it is discontinuous, and not differentiable. Therefore, the Haar system is not very convenient for approximation of smooth functions.

4.2.2 Daubechies' Wavelet

Daubechies was a pioneer in the construction of compactly supported orthogonal wavelets. The collection of the scaling functions $\{\varphi_{j,k}(x),\ j,k \in Z\}$, which is derived from Daubechies' wavelet, is an orthonormal system. Moreover, the collection of the primary wavelet functions $\{\psi_{j,k}(x),\ j,k \in Z\}$, which is derived from Daubechies' compactly supported dilated and translated versions of the wavelet function, is also an orthonormal basis in $L^2(R)$. The primary wavelet $\psi(x)$ has N vanishing moments that determines the accuracy of approximations based on the wavelet. That is,

$$\int x^n \psi(x) dx = 0, \quad n = 0, 1, \cdots, N-1.$$

The number of filter coefficients is determined by the number of vanishing moments for all different versions of the Daubechies' wavelet. For instance, if number of the filter coefficient is denoted by L. Then, $L = 2N$, and the scaling function and the primary wavelet have compact support $\text{supp}\varphi = [0,\ 2N-1], \text{supp}\psi = [-N+1,\ N]$.

Daubechies' wavelets are commonly represented by DAUBN, where N is the number of vanishing moments. It is interesting to note that DAUB1 wavelet is equivalent to the Haar wavelet. Figure 4.1 shows the plots of scaling and primary wavelet functions with vanishing moments, $N = 2, 4, 8$. Table 4.1 is the list of the filter coefficients for DAUB2-DAUB10 wavelets.

4.3 Wavelet Version of Estimators

Consider the nonparametric regression model (3.1),

$$Y_i = m(X_i) + \varepsilon_i, \qquad i = 1, 2, \cdots, n.$$

which contains an unknown regressor m. Using the expansion (4.6), it is clear that we can write $m(x)$ in (3.1) with the wavelet expansion,

$$m(x) = \sum c_k \varphi_{0,k}(x) + \sum_j \sum_k d_{j,k} \psi_{j,k}(x)$$

where

$$c_k = \int_s m(x)\varphi_{0,k}(x)dx$$
$$d_{j,k} = \int_s m(x)\psi_{j,k}(x)dx.$$

Given the observed responses $y_1, y_2, \cdots, y_n$ in $S = [a, b]$, we partition S into mutually exclusive but collectively exhaustive subspaces A_j such that $y_j \in A_j$ and $\cup_{i=1}^{n} A_i = S$. It follows that

$$d_{j,k} = \int_s m(x)\psi_{j,k}(x)dx = \sum_{i=1}^{n} \int_{A_i} m(x)\psi_{j,k}(x)dx.$$

Since $m(x)$ is unknown, we use Y_i to approximate $m(x) \in A_i$. Thus,

$$d_{j,k} = \sum_{i=1}^{n} \int_{A_i} m(x)\psi_{j,k}(x)dx \approx \sum_{i=1}^{n} Y_i \int_{A_i} \psi_{j,k}(s)ds = \hat{d}_{j,k}$$

the estimator of $d_{j,k}$.

Similarly, we can write

$$\hat{c}_k = \sum_{i=1}^{n} Y_i \int \varphi_{o,k}(s)ds$$

It follows that

$$\hat{m}(x) = \sum_k \sum_{i=1}^{n} Y_i \int \varphi_{0,k}(s)ds\varphi_{0,k}(x) + \sum_j \sum_k Y_i \int \psi_{j,k}(s)ds\psi_{j,k}(x)$$

Therefore, we can write

$$\hat{m}(x) = \sum_j \sum_k Y_i \int K_m(x,s)ds$$

, the well-known wavelet version of the Gasser-Müller Estimator.

Let x_i's be nonrandom design points within the design space $S = [0,1]$. There are total n design points, and have the order $0 \le x_1 \le \cdots \le x_n \le 1$. Let $[A_{i-1}, A_i]$ be n subintervals of the design space [0,1], and each subinterval contains a design point x. So, $x_i \in [A_{i-1}, A_i]$, and $\cup_{i=1}^{n} = [0,1]$. The wavelet version of Gasser-Müller kernel estimator is given by,

$$\hat{\eta}_M(x) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} Y_{ij} \int_{A_{i-1}}^{A_i} K_m(x,s)ds. \tag{4.7}$$

In (4.7), n_i denotes the number of observations at x_i, $K_m(x,s)$ is the wavelet kernel which can be represented as

$$K_m(x,s) = nq_m'(x)q_m(s)$$

where $q_m(x)$ is a $2^{m+1} \times 1$ vector consisting of a set of dilated and translated versions of primary wavelets, $\psi^{-j,k}(x) = 2^{j/2}\psi(2^j x - k), (j = 0,1,\cdots,m;\ k = 0,\cdots,2^j - 1)$, and a scaling function $\varphi(x)$. For instance, if $m = 2$, then,

$$q_2'(x) = \{\varphi(x), \psi(x), \psi^{-1,0}(x), \psi^{-2,0}(x), \psi^{-2,1}(x), \psi^{-2,2}(x), \psi^{-2,3}(x)\}.$$

In (4.7), m has a similar role as the bandwidth h in the standard kernel smoothers. Therefore, the choice of value for m need be considered. We note that as m is increased by 1, the number of functions in the vector $q_m(x)$ is doubled. The value for m can not be too large because unreasonable large value of m will produce a rougher plot including wigles and spikes of the observations. Antoniadis, Gregoire & McKeague (1994) suggested that it is sufficient to consider a maximum value of 5 for m when the sample size is between 100 and 200. Oyet (2002) suggested that the choice of m should depend on the structure of the response curve instead of sample size since the value of m determines the number of observations required to uniquely estimate the response.

Oyet (2002) suggested a modidfied kernel in which

$$K_m(x,s) = q_m'(x)B^{-1}q_m(s)$$

where

$$B = Z'PQ = \sum_{i=1}^{n} p_i z_i q_m'(x_i) \quad \text{and} \quad z_i = \int_{A_{i-1}}^{A_i} q_m(s)ds.$$

The matrix Q contains vectors $q_m'(x_i)$, $(i = 1, 2, \cdots, n)$, with the dimension $n \times r'$, Z is a $n \times r$ matrix with rows $z'(x_i)$, $i = 1, 2, \cdots, n$, and $P = diag(p_1, p_2, \cdots, p_n)$. The modification "is motivated by a desire to obtain an exact expression for the bias of $\hat{\beta}_M$" in (4.8).

The estimator in (4.7) can be expressed in terms of the collection of scaling and primary wavelet functions as

$$\hat{\eta}_M(x) = q_m'(x)\hat{\beta}_M \tag{4.8}$$

where

$$\hat{\beta}_M = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n_i} Y_{ij} B^{-1} z_i.$$

According to Oyet (2002), the regression weights have a positive impact on the accuracy of wavelet estimation. Therefore, the least squares estimation of β could be a better option. It is defined as

$$\hat{\beta}_{WLS} = \left\{ \frac{1}{n} \sum_{i=1}^{n} n_i w_i q_m(x_i) q_m'(x_i) \right\}^{-1} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} w_i q_m(x_i) Y_{ij}, \tag{4.9}$$

where

$$w_i = w(x_i) = \int_0^1 \frac{||q_m(s)||ds}{||q_m(x_i)||}.$$

We will have wavelet version of weighted least squares regression once replacing $\hat{\beta}_M$ in equation (4.8) by $\hat{\beta}_{WLS}$,

$$\hat{\eta}_{WLS}(x) = q_m'(x) \hat{\beta}_{WSLS} \tag{4.10}$$

Oyet and Wiens (2002) "have demonstrated that w_i are minimum variance unbiased weights for weighted least squares estimation of the parameters in wavelet regression models".

Table 4.1: The h filters for Daubechies' wavelets for $N = 2, \cdots, 10$ vanishing moments.

k	DAUB2	DAUB3	DAUB4
0	0.482962913145	0.332670552950	0.230377813309
1	0.836516303738	0.806891509311	0.714846570553
2	0.224143868042	0.459877502118	0.630880767930
3	-0.129409522551	-0.135011020010	-0.027983769417
4		-0.085441273882	-0.187034811719
5		0.035226291882	0.030841381836
6			0.032883011667
7			-0.010597401785

k	DAUB5	DAUB6	DAUB7
0	0.160102397974	0.111540743350	0.077852054085
1	0.603829269797	0.494623890398	0.396539319482
2	0.724308528438	0.751133908021	0.729132090846
3	0.138428145901	0.315250351709	0.469782287405
4	-0.242294887066	-0.226264693965	-0.143906003929
5	-0.032244869585	-0.129766867567	-0.224036184994
6	0.077571493840	0.097501605587	0.071309219267
7	-0.006241490213	0.027522865530	0.080612609151
8	-0.012580751999	-0.031582039317	-0.038029936935
9	0.003335725285	0.000553842201	-0.016574541631
10		0.004777257511	0.012550998556
11		-0.001077301085	0.000429577973
12			-0.001801640704
13			0.000353713800

k	DAUB8	DAUB9	DAUB10
0	0.054415842243	0.038077947364	0.026670057901
1	0.312871590914	0.243834674613	0.188176800078
2	0.675630736297	0.604823123690	0.527201188932
3	0.585354683654	0.657288078051	0.688459039454
4	-0.015829105256	0.133197385825	0.281172343661
5	-0.284015542962	-0.293273783279	-0.249846424327
6	0.000472484574	-0.096840783223	-0.195946274377
7	0.128747426620	0.148540749338	0.127369340336
8	-0.017369301002	0.030725681479	0.093057364604
9	-0.044088253931	-0.067632829061	-0.071394147166
10	0.013981027917	0.000250947115	-0.029457536822
11	0.008746094047	0.022361662124	0.033212674059
12	-0.004870352993	-0.004723204758	0.003606553567
13	-0.000391740373	-0.004281503682	-0.010733175483
14	0.000675449406	0.001847646883	0.001395351747
15	-0.000117476784	0.000230385764	0.001992405295
16		-0.000251963189	-0.000685856695
17		0.000039347320	-0.000116466855
18			0.000093588670

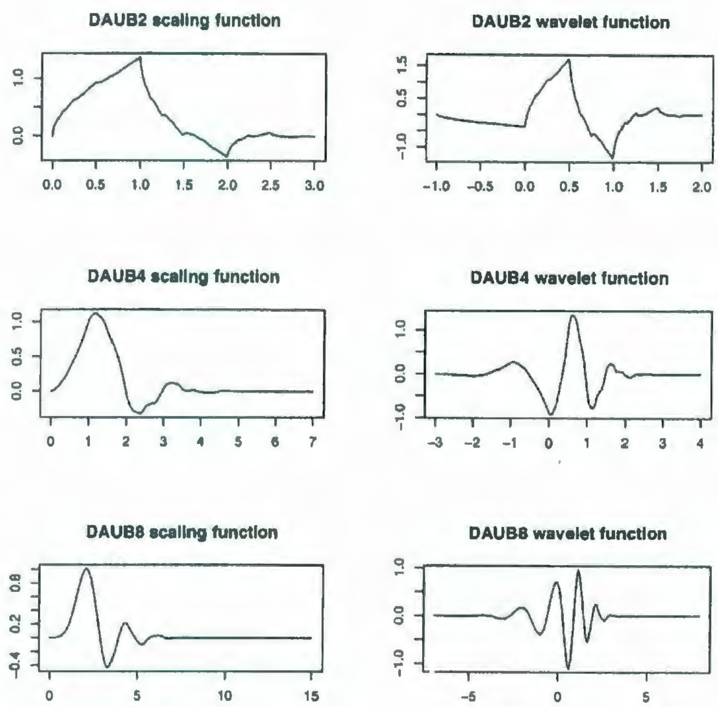

Figure 4.1: Graph of scaling and wavelets wavelets functions from Daubechies' family, $N = 2, 4,$ and 8

Chapter 5

Simulation Study

In this chapter, we conduct numerical simulation studies to compare the performance of different regression methods that are discussed in the first four chapters. The parameters of the Michaelis-Menten model in (1.1) are assigned with true value $\theta_1 = 129,\ \theta_2 = 3,\ \theta_3 = 0.06$. Then, the true model is

$$f(x,\ \theta) = \frac{\theta_1 x}{\theta_2 + x + \theta_3 x^2} = \frac{129x}{3 + x + 0.06x^2} \tag{5.1}$$

The data is simulated based on the model,

$$y_i = f(x_i;\ \theta) + \varepsilon_i, \tag{5.2}$$

with the error term ε. In this study, the errors are assumed to follow the normal distribution, $\varepsilon_i \sim N(\mu = 1,\ \sigma^2 = 2)$. The independent variable, x, is designed as 100 fixed points by the sequence from 1 to 100. Therefore, the sample size n is 100, and $\{\ x_i\ \}_{i=1}^{100} = \{\ 1, 2, 3, \cdots, 100\ \}$. For each simulation, we have 100 pairs of x_i and $f(x_i)$, $\{x_i, f(x_i)\}_{i=1}^{100}$.

For each simulation, different smoothing methods that was introduced in previous chapters were applied to evaluate $\hat{f}(x_i)$, which is the estimator of $f(x_i)$. The simulation processes were repeated 1000 times. Therefore, there are 1000 estimates of $f(x_i)$ at each corresponding fixed design point x_i. Let $\hat{f}_j(x_i), (j = 1, 2, \cdots, 1000)$, denote the 1000 repeated estimations at x_i. Moreover, let $\bar{f}(x_i)$ denotes the mean value of the estimations at each point x_i. Then,

$$\bar{f}(x_i) = \frac{f_1(x_i) + f_2(x_i) + \cdots + f_{1000}(x_i)}{1000}$$

If f_i denotes the true value at each corresponding point of x_i, then, the bias, variance, and MSE of the estimator at each design point are given by,

$$\text{bias} = \bar{f}(x_i) - f_i \tag{5.3}$$

$$\text{variance} = \frac{1}{1000-1}\sum_{j=1}^{1000}\left(\hat{f}_j(x_i) - \bar{f}(x_i)\right)^2 \tag{5.4}$$

$$\text{MSE} = \text{bias}^2 + \text{variance} \tag{5.5}$$

where,

$$f_i = \frac{129x_i}{3 + x_i + 0.06x_i^2}.$$

5.1 Nonlinear Least Squares Estimation

Since the model (5.2) is a nonlinear model with the relation function already known, the nonlinear least squares estimation is the first regression method to be considered. To estimate $f(x_i)$, we need to estimate $\theta = (\theta_1, \theta_2, \theta_3)'$ in advance. The estimator of θ was obtained by Gauss-Newton method with (2.22). The $F.$ matrix in (2.22) is,

$$\begin{aligned} F. &= \frac{\partial f(\theta)}{\partial \theta'} \\ &= \left(\frac{\partial f_i(\theta)}{\partial \theta_1}, \frac{\partial f_i(\theta)}{\partial \theta_2}, \frac{\partial f_i(\theta)}{\partial \theta_3}\right) \\ &= \begin{bmatrix} x_i(\theta_2 + x_i + \theta_3 x_i^2)^{-1} \\ -\theta_1 x_i(\theta_2 + x_i + \theta_3 x_i^2)^{-2} \\ -\theta_1 x_i^3(\theta_2 + x_i + \theta_3 x_i^2)^{-2} \end{bmatrix}' \end{aligned}$$

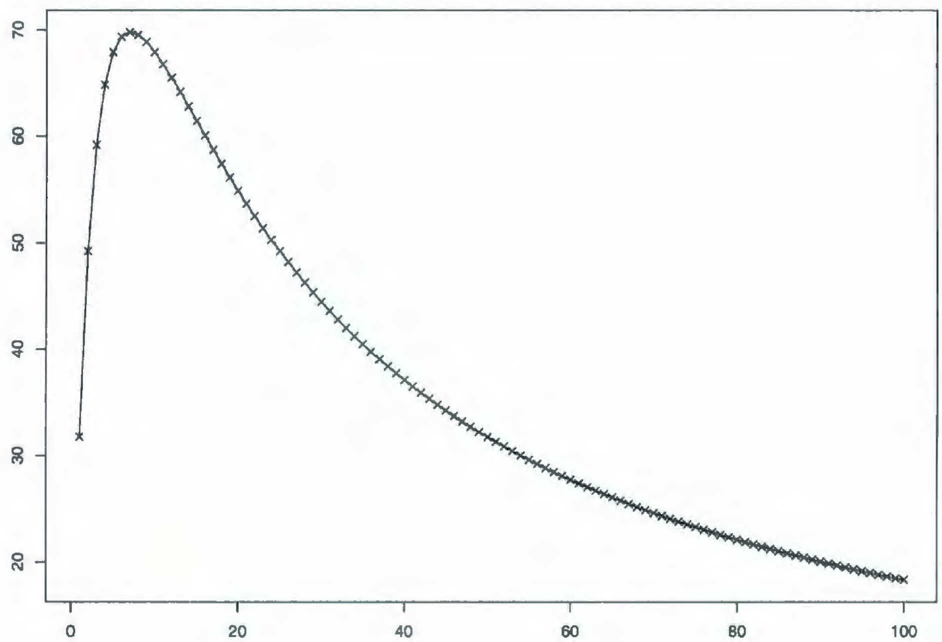

Figure 5.1: nonlinear least squares regression

Figure (5.1) is the plot of nonlinear least squares regression result. The dots are the true values, f_i. The line is the mean value of 1000 repeated estimations. According to the graph, the smoothing method gives a very good fit to the true values. The mean values, bias, variance, and MSE of the estimations are listed in Appendix A. The value of bias square and variance of the estimations suggest that the nonlinear least squares method, overall, gives accuracy estimations of $f(x_i)$ with constant small variance. Variance is the major contribution to the MSE.

5.2 Nonparametric Smoothing

If we assume the regression function is unknown to us, the nonparametric methods could be adapted to estimate $f(x_i)$. In this section, we will use different nonparametric

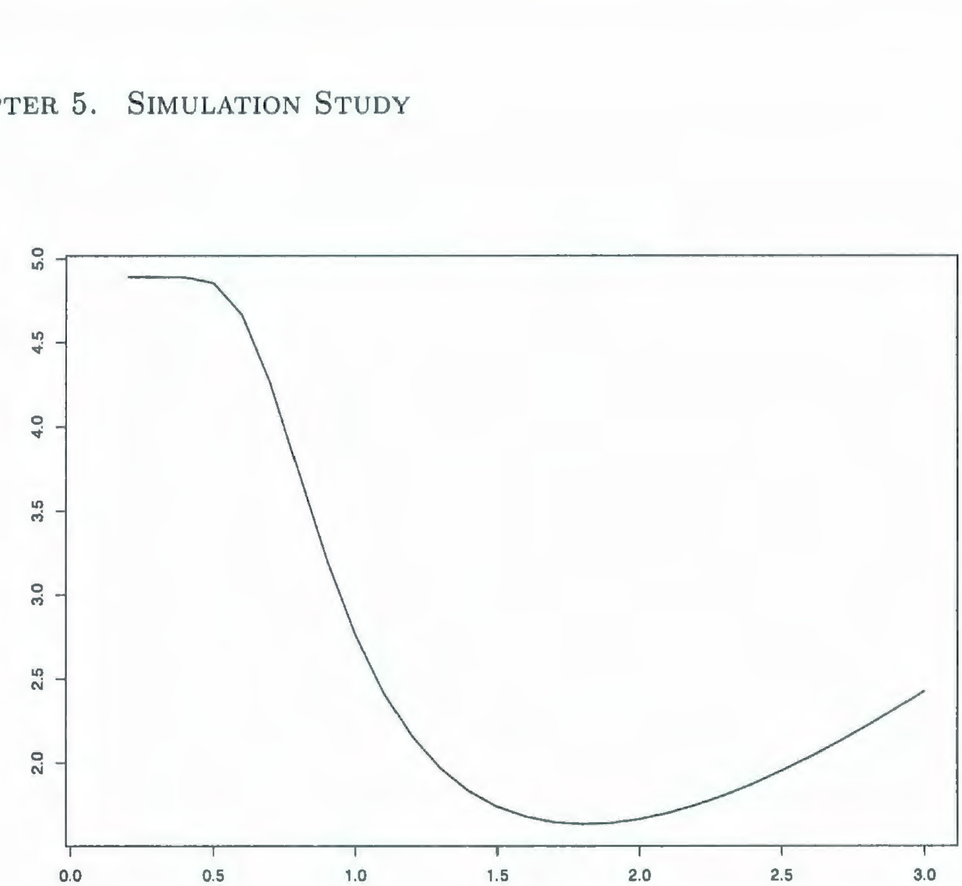

Figure 5.2: Cross Validation plot of Nadaraya-Watson Estimator

methods to fit the simulated data set, $\{x_i,\ f(x_i)\}_{i=1}^{100}$. The Gaussian kernel (3.6) is used for kernel estimators.

5.2.1 Nadaraya-Watson Estimator

The Nadaraya-Watson estimator is given by (3.15). The choice of bandwidth, h, is determined by the cross-validatoin function (3.38). The h value is chosen to minimize the function, $CV(h)$. Figure (5.2) is the plot of the function at different h values. The minimum value of the function is obtained when $h = 1.8$.

Figure (5.3) is the plot of Naradaya-watson smoothing result. The dots are the true values, f_i. The line is the mean value of 1000 repeated estimations. According to the graph, the curve is well fitted except at the first point, x_1, and the several points around the bend located at the top of the curve. Appendix B lists the mean

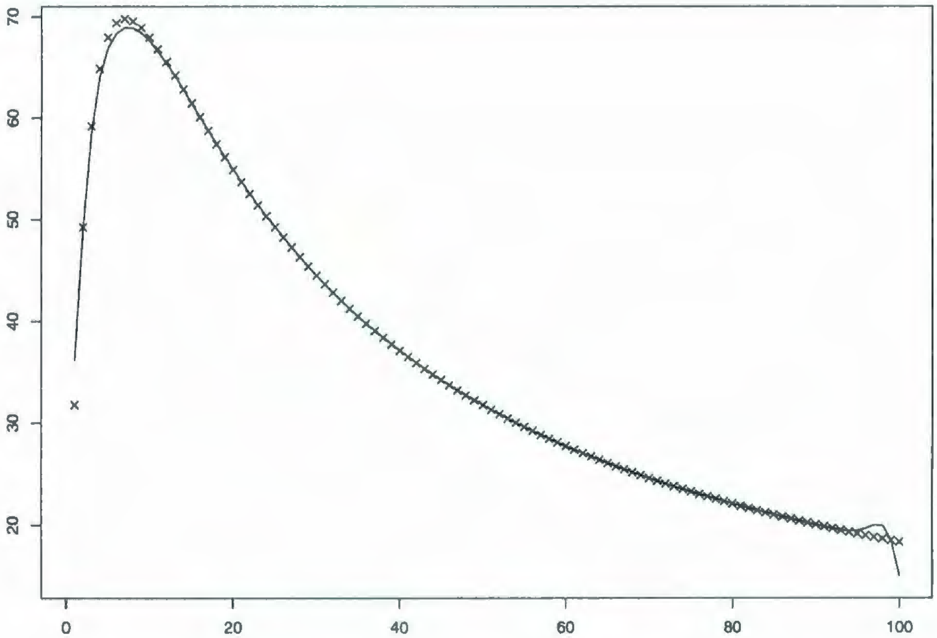

Figure 5.3: Nadaraya-Watson Regression Plot

values, bias, variance, and MSE of the estimations of $f(x_i)$ by Nadaraya-Watson method. The MSE of the estimation at the first point, x_1, and the last point, x_{100}, are unusually large than the rest values of MSE. Moreover, MSE are relatively larger at design points x_5, x_6, x_7, x_{97}, andx_{98} compared to the MSE at all other design points. The larger MSE at these points are contributed by the bias.

5.2.2 Gasser-Müller Estimator

The second kernel estimator is Gasser-Müller method which is given by (3.17). The choice of bandwidth, h, follows the same methods as in the Nadaraya-Watson estimator. Figure (5.4) shows the plot of cross-validation function regarding different h values. The minimum value of the CV function is achieved when h takes value 4.1.

Figure (5.5) is the plot of Gasser-Müller smoothing. The curve is well fitted by the

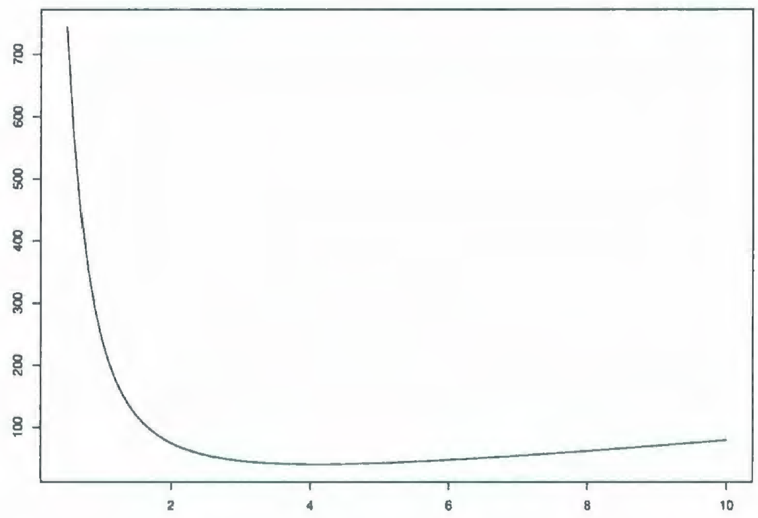

Figure 5.4: Cross Validation Plot of Gasser-Müller Estimator

Gasser-Müller smoothing method except at two tails of the curve. Appendix C lists the mean values, bias, variance, and MSE of the estimations. MSE's are extermely large at points from x_2 to x_{12}, and points from x_{94} and x_{100}.

5.2.3 Natural Spline Smoothing

In this section, the natural spline method is adapted to smooth simulated data $\{x_i,\ f(x_i)\}_{i=1}^{100}$. The estimation of $f(x_i)$ is given by equation (3.19) in which the coefficients, a_i, b_i, c_i, and d_i, could be solved through relation (3.37).

Figure (5.6) is the plot of Natural Spline smoothing of the simulated data. The curve is well fitted except at the first few design points. Appendix D lists the mean values, bias, variance, and MSE of the estimations. The MSE are extremely large at all estimation points. The large MSE are contributed by the large estimation variance. However, the bias values are very small which suggests a good accuracy of the estimation.

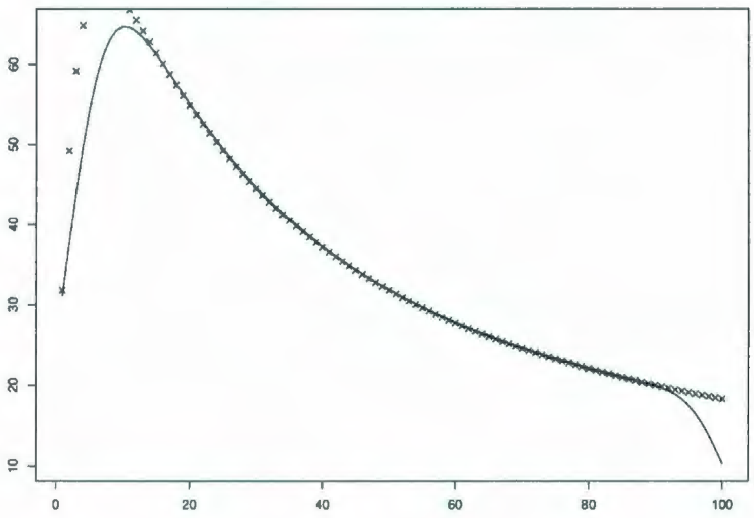

Figure 5.5: Gasser-Müller Regression Plot

5.2.4 Wavelet Version of Gasser-Müller Smoothing

In this section the wavelet version of Gasser-Müller method are applied with two different wavelet systems, Haar Wavelet and Daubechies' Wavelet. The estimator is given by equation (4.8). Figure (5.7) is the plot of the regression with Harr Wavelet and Daubechies's Wavelet. The smoother with Harr Wavelt shows a steps plot. The reason is that the scaling function of Harr Wavelet is a discontinuous step function. Appendix E lists the mean values, bias, variance, and MSE of the estimations with Haar Wavelet.

The regressor with Haar Wavelet is unacceptable since our curve is continuous. Therefore, Daubechies' Wavelet should be more favorable than Haar Wavelet in this study. The best estimation result is obtained at $m = 3$. Plot (Figure 5.7) shows that the regressor with Daubechies' Wavelet fitted the curve pefectly except a spike at the top of the curve. Appendix F lists the mean values, bias, variance, and MSE of the estimations with Daubechies' Wavelet. The MSE of the estimation are consid-

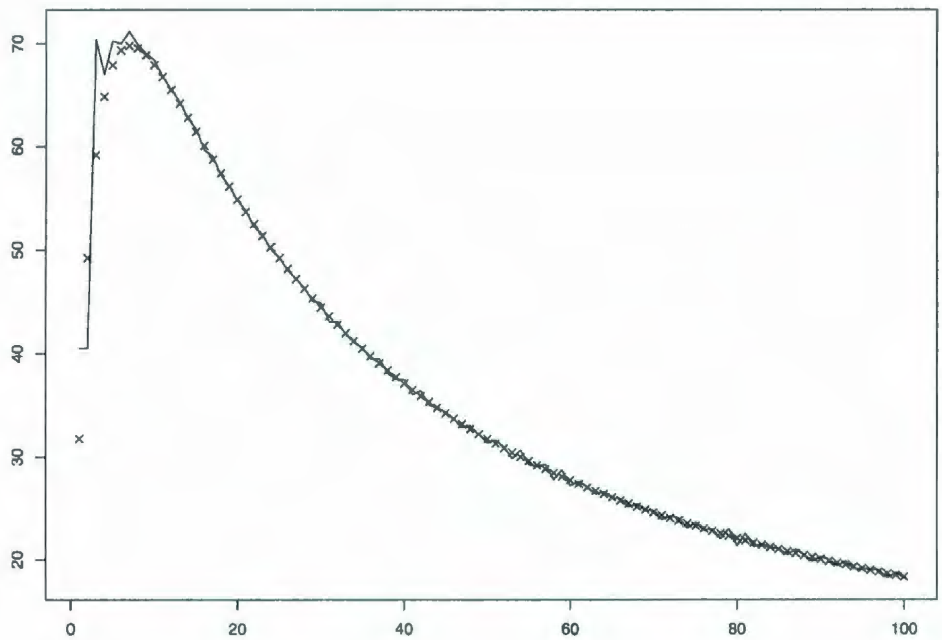

Figure 5.6: Natural Spline Smoothing Plot

erably large at the first 13 design points. It is because of large bias produced by the estimation.

5.2.5 Wavelet Weighted Least Squares Smoothing

In this section, the simulated data is estimated by wavelet Wegihted Least Squares (WLS) methods with Daubechies' Wavelet system. The estimator is given by equation (4.10). The wavelet WLS methods produces a similar result as the the wavelet Gasser-Müller methods does. The best estimating result is obtained at $m = 3$. Figure (5.8) is the plot of the fitted curve. The curve is perfectly fitted except several points around the bend part of the curve. Appendix G lists the mean values, bias, variance, and MSE of the 1000 repeated estimations. The MSE of the first 11 of 12 points are vary large. The rest of estimations give acceptable small MSE. Moreover, the MSE given

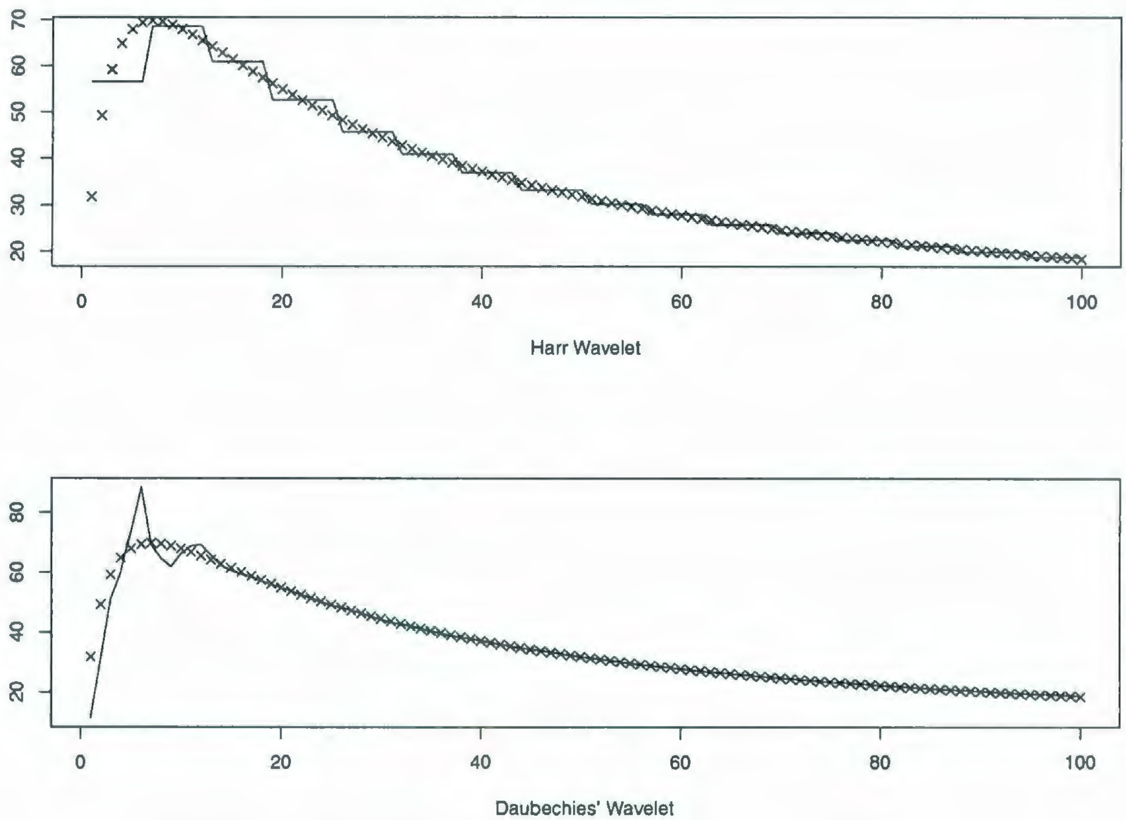

Figure 5.7: Wavelet Version of Gasser Muller Regression with Different Wavelet Systems

by WLS is generally smaller than MSE given by wavelet Gasser-Müller method. It confirms with Oyet's (2002) suggestion that WLS gave a better weight.

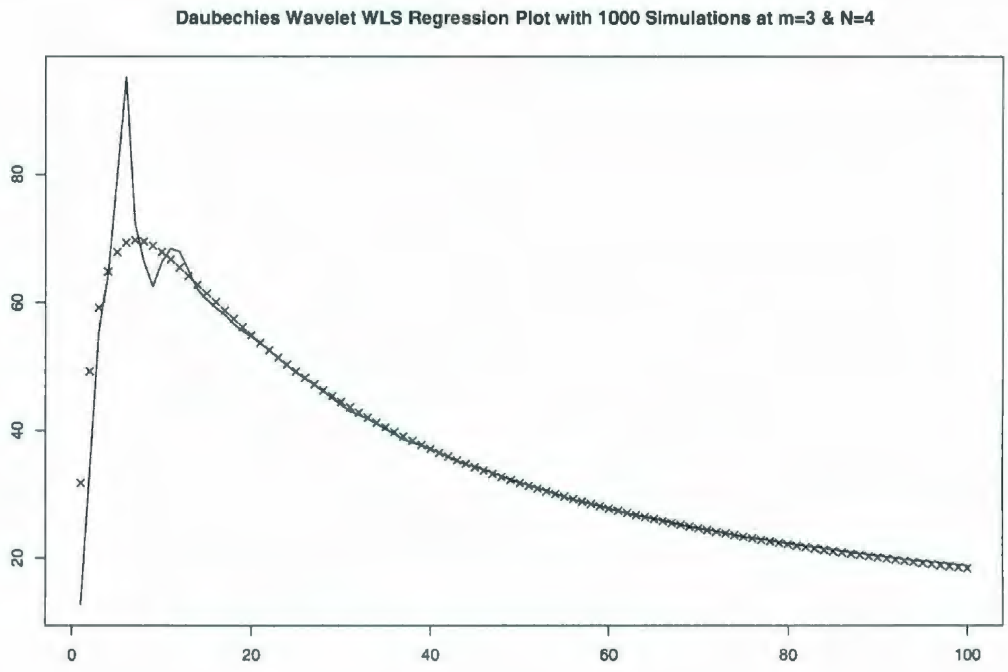

Figure 5.8: Wavelet Least Squares estimation with Daubechies' Wavelet Systems

Chapter 6

Concluding Remarks

The result of this study demonstrated different smoothing methods regarding the model with that the relation function is known to us. Three terms, bias, variance and MSE of estimation of $f(x_i)$, are adopted to access the goodness of fit regarding different smoothing methods.

Since the regression function is kown to us, the nonlinear least squares estimation is the favorite choice of smoothing methods. Moreover, nonparametric and wavelet smooting methods can be applied since the relation function is assumed unknown. The first smoothing method gives the best estimations inclusively compared to all other regressors including nonparametric and wavelet modelling methods. Among all the methods, the nonlinear least squares estimation produces the smallest bias and variance which in turn produces the smallest MSE.

The natural spline smoother is the worst regressor in the study since it produces unusually large MSE at each individual point of estimation. All the other regressors, that assume the regression model is unknown, fit the simulated data adequately. It is of interest to note that the estimation given by Daubechies' wavelet version of Gasser-Müller smoother could be a competitor of nonlinear least squares estimator in terms of bias, variance, and MSE. The only deficiency with the Daubechies' wavelet regressor is the unacceptable large MSE with the estimations at the first 13 design points. It is probabily due to the much less density of data points located around first 6 design points.

In future studies, it might be interesting to consider the performance of these regressors with models consisting multiple dimensions of dependent variable x, and to improve the wavelet modelling methods with situations that data points are less concentrated.

References

Antoniadis, A., Gregoire, G., and McKeague, W. (1994). Wavelet Methods for Curve Estimation. *Journal of the American Statistical Association*, 89, 1340-1352.

Azzalini, A., Bowman, A.W., and Härdle, W. (1989). On the use of nonparametric regression for model checking. *Biometrika*, 76, 1-11.

Bates, D.M., and Watts, D.G. (1988). *Nonlinear Regression Analysis and its Applicatoins.* Wiley, New York.

Devroye, L. (1983). The Equivalence of Weak, Strong, and Complete convergence in L_1 for Kernel Density Estimates. *The Annals of Statistics*, 11, 896-904.

Gallant, A.R. (1975). Nonlinear Regression. *The American Statistician*, 29, 73-81.

Härdle, W. (1989). *Applied Nonparametric Regression.* Cambridge, Cambridge.

Härdle, W., Kerkyacharian, G., and Tsubakov, A. (1998). *Wavelets, Approximation, and Statistical Applications.* Springer, New York.

Härdle, W., and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *The Annals of Statistics.* 21, 1926-1947.

Hartley, O.H. (1961). The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares. *American Statistical Association and American Society for Quality*, 3, 269-280.

Hartley, O.H., and Booker, A. (1965). Nonlinear Least Squares Estimation. *The Annals of Mathematical Statistics*, 36, 638-650.

Izenman, A.J. (1991). Recent Developments in Nonparametric Density Estimation. *Journal of the American Statistical Association*, 86, No. 314, Review Paper.

Oyet, A.J. (2002). Minimax A- and D-optimal integer-valued wavelet designs for estimation. *The Canadian Journal of Statistics*, 30, 301-306.

Oyet, A.J. and Wiens, D.P. (2000). On exact minimax wavelet designs obtained by simulated annealing. Unpublished manuscript.

Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, 33, 1065-1076.

Rice, J.A. (1997). *Mathematical statistics and data analysis.* Duxbury, New York.

Seber, G.A.F, and Wild, C.J. (1989). *Nonlinear Regression.* New York: Wiley.

Seber, G.A.F., and Lee, A.J. (2003). *Linear Regression Analysis*, 2nd ed. New York: Wiley.

Stone, M. (1974). Cross-Validatory choice and assessment of statistical predictions (with discussion). *Journal of the Royal Statistical Society*, Series B, 36, 111-147.

Takezawa, K. (2006). *Introduction to nonparametric regression*, New York: Wiley.

Wahba, G., and Wold, S. (1975). A completely automatic French curve: fitting spline functions by cross-validation. *Communications in Statistics*, Series A, 4, 1-17.

Vidakovic, B. (1999), *Statistical Modeling by Wavelets*, Wiley, New York.

Appendix A

	y_tru	y_exp	bias2	y_var	MSE
[1,]	31.77340	31.75675	2.772600e-04	0.25904094	0.25931820
[2,]	49.23664	49.22404	1.588043e-04	0.26996845	0.27012725
[3,]	59.17431	59.17062	1.361427e-05	0.19164838	0.19166200
[4,]	64.82412	64.82868	2.078193e-05	0.13374098	0.13376176
[5,]	67.89474	67.90549	1.156538e-04	0.10350226	0.10361791
[6,]	69.35484	69.36973	2.218605e-04	0.08949957	0.08972143
[7,]	69.78362	69.80098	3.013628e-04	0.08266179	0.08296316
[8,]	69.54178	69.56035	3.447260e-04	0.07817628	0.07852100
[9,]	68.86121	68.88008	3.561603e-04	0.07399542	0.07435158
[10,]	67.89474	67.91329	3.442055e-04	0.06948054	0.06982474
[11,]	66.74506	66.76288	3.173510e-04	0.06460263	0.06491998
[12,]	65.48223	65.49904	2.824560e-04	0.05954146	0.05982392
[13,]	64.15455	64.17019	2.444830e-04	0.05450720	0.05475168
[14,]	62.79555	62.80993	2.067407e-04	0.04967311	0.04987985
[15,]	61.42857	61.44166	1.712633e-04	0.04515884	0.04533011
[16,]	60.06985	60.08165	1.391714e-04	0.04103438	0.04117356
[17,]	58.73058	58.74112	1.109646e-04	0.03733051	0.03744147
[18,]	57.41840	57.42771	8.674028e-05	0.03404989	0.03413663
[19,]	56.13834	56.14649	6.634991e-05	0.03117658	0.03124293
[20,]	54.89362	54.90065	4.950600e-05	0.02868329	0.02873280
[21,]	53.68609	53.69208	3.585307e-05	0.02653671	0.02657257
[22,]	52.51665	52.52166	2.501325e-05	0.02470115	0.02472617
[23,]	51.38552	51.38960	1.661431e-05	0.02314096	0.02315757
[24,]	50.29240	50.29561	1.030592e-05	0.02182202	0.02183233
[25,]	49.23664	49.23904	5.768046e-06	0.02071265	0.02071841
[26,]	48.21737	48.21901	2.714480e-06	0.01978397	0.01978668
[27,]	47.23352	47.23447	8.931804e-07	0.01901012	0.01901101
[28,]	46.28396	46.28425	8.483180e-08	0.01836815	0.01836824
[29,]	45.36745	45.36713	1.003987e-07	0.01783795	0.01783805
[30,]	44.48276	44.48188	7.782107e-07	0.01740193	0.01740271
[31,]	43.62863	43.62722	1.980921e-06	0.01704490	0.01704688
[32,]	42.80382	42.80192	3.592534e-06	0.01675374	0.01675733
[33,]	42.00710	42.00476	5.515628e-06	0.01651717	0.01652269
[34,]	41.23731	41.23454	7.668805e-06	0.01632559	0.01633326
[35,]	40.49327	40.49011	9.984424e-06	0.01617078	0.01618077
[36,]	39.77390	39.77037	1.240659e-05	0.01604578	0.01605819
[37,]	39.07811	39.07425	1.488937e-05	0.01594470	0.01595959
[38,]	38.40489	38.40072	1.739534e-05	0.01586255	0.01587995
[39,]	37.75326	37.74880	1.989415e-05	0.01579517	0.01581506
[40,]	37.12230	37.11757	2.236153e-05	0.01573904	0.01576141
[41,]	36.51111	36.50614	2.477821e-05	0.01569126	0.01571604
[42,]	35.91885	35.91365	2.712917e-05	0.01564940	0.01567653
[43,]	35.34472	35.33930	2.940288e-05	0.01561145	0.01564086
[44,]	34.78794	34.78232	3.159078e-05	0.01557578	0.01560737
[45,]	34.24779	34.24198	3.368670e-05	0.01554105	0.01557473
[46,]	33.72357	33.71760	3.568650e-05	0.01550617	0.01554185
[47,]	33.21464	33.20851	3.758767e-05	0.01547028	0.01550787
[48,]	32.72036	32.71408	3.938903e-05	0.01543271	0.01547210
[49,]	32.24013	32.23372	4.109051e-05	0.01539292	0.01543401
[50,]	31.77340	31.76687	4.269289e-05	0.01535053	0.01539322
[51,]	31.31962	31.31297	4.419768e-05	0.01530524	0.01534943
[52,]	30.87829	30.87154	4.560692e-05	0.01525685	0.01530245
[53,]	30.44892	30.44207	4.692306e-05	0.01520524	0.01525216
[54,]	30.03104	30.02410	4.814889e-05	0.01515035	0.01519850
[55,]	29.62422	29.61720	4.928743e-05	0.01509216	0.01514145
[56,]	29.22803	29.22094	5.034187e-05	0.01503070	0.01508104
[57,]	28.84208	28.83492	5.131549e-05	0.01496603	0.01501735
[58,]	28.46599	28.45876	5.221164e-05	0.01489824	0.01495045
[59,]	28.09939	28.09210	5.303368e-05	0.01482743	0.01488046
[60,]	27.74194	27.73460	5.378497e-05	0.01475373	0.01480751
[61,]	27.39330	27.38592	5.446883e-05	0.01467726	0.01473173
[62,]	27.05317	27.04575	5.508851e-05	0.01459818	0.01465327
[63,]	26.72125	26.71379	5.564719e-05	0.01451663	0.01457228
[64,]	26.39724	26.38974	5.614799e-05	0.01443276	0.01448890
[65,]	26.08087	26.07335	5.659390e-05	0.01434671	0.01440331
[66,]	25.77189	25.76434	5.698783e-05	0.01425866	0.01431564
[67,]	25.47003	25.46246	5.733259e-05	0.01416873	0.01422606
[68,]	25.17507	25.16747	5.763086e-05	0.01407709	0.01413472

```
 [70,] 24.60490 24.59728 5.809817e-05 0.01388923 0.01394733
 [71,] 24.32928 24.32164 5.827206e-05 0.01379328 0.01385155
 [72,] 24.05968 24.05204 5.840915e-05 0.01369617 0.01375458
 [73,] 23.79593 23.78828 5.851160e-05 0.01359802 0.01365653
 [74,] 23.53782 23.53017 5.858147e-05 0.01349896 0.01355754
 [75,] 23.28520 23.27754 5.862070e-05 0.01339909 0.01345771
 [76,] 23.03788 23.03022 5.863117e-05 0.01329853 0.01335716
 [77,] 22.79570 22.78805 5.861462e-05 0.01319738 0.01325600
 [78,] 22.55851 22.55086 5.857275e-05 0.01309575 0.01315432
 [79,] 22.32616 22.31851 5.850714e-05 0.01299372 0.01305223
 [80,] 22.09850 22.09086 5.841929e-05 0.01289140 0.01294982
 [81,] 21.87539 21.86776 5.831065e-05 0.01278886 0.01284717
 [82,] 21.65670 21.64908 5.818256e-05 0.01268618 0.01274436
 [83,] 21.44230 21.43469 5.803631e-05 0.01258343 0.01264147
 [84,] 21.23207 21.22446 5.787310e-05 0.01248070 0.01253858
 [85,] 21.02589 21.01829 5.769409e-05 0.01237805 0.01243574
 [86,] 20.82364 20.81605 5.750037e-05 0.01227553 0.01233303
 [87,] 20.62521 20.61764 5.729295e-05 0.01217320 0.01223050
 [88,] 20.43049 20.42294 5.707281e-05 0.01207113 0.01212820
 [89,] 20.23940 20.23186 5.684086e-05 0.01196936 0.01202620
 [90,] 20.05181 20.04429 5.659797e-05 0.01186794 0.01192453
 [91,] 19.86765 19.86014 5.634496e-05 0.01176691 0.01182325
 [92,] 19.68682 19.67933 5.608260e-05 0.01166631 0.01172240
 [93,] 19.50922 19.50175 5.581162e-05 0.01156619 0.01162200
 [94,] 19.33478 19.32733 5.553270e-05 0.01146658 0.01152211
 [95,] 19.16341 19.15598 5.524649e-05 0.01136750 0.01142275
 [96,] 18.99503 18.98762 5.495360e-05 0.01126900 0.01132395
 [97,] 18.82957 18.82217 5.465462e-05 0.01117109 0.01122574
 [98,] 18.66694 18.65957 5.435008e-05 0.01107380 0.01112815
 [99,] 18.50709 18.49974 5.404049e-05 0.01097716 0.01103120
[100,] 18.34993 18.34260 5.372634e-05 0.01088119 0.01093492
```

Appendix B

	y_tru	y_expNW	bias2_NW	y_varNW	MSE_NW
[1,]	31.77340	36.22639	1.982916e+01	0.2032081	20.0323658
[2,]	49.23664	49.12078	1.342375e-02	0.2320732	0.2454970
[3,]	59.17431	58.46012	5.100696e-01	0.2099256	0.7199952
[4,]	64.82412	63.97802	7.158778e-01	0.1826502	0.8985281
[5,]	67.89474	66.86161	1.067346e+00	0.1658619	1.2332079
[6,]	69.35484	68.27968	1.155962e+00	0.1573186	1.3132808
[7,]	69.78362	68.85257	8.668506e-01	0.1584904	1.0253410
[8,]	69.54178	68.83628	4.977221e-01	0.1641707	0.6618928
[9,]	68.86121	68.36917	2.421056e-01	0.1653837	0.4074894
[10,]	67.89474	67.56929	1.059128e-01	0.1626049	0.2685177
[11,]	66.74506	66.54170	4.135568e-02	0.1577133	0.1990690
[12,]	65.48223	65.36679	1.332673e-02	0.1506609	0.1639876
[13,]	64.15455	64.09945	3.036403e-03	0.1473855	0.1504219
[14,]	62.79555	62.77845	2.922576e-04	0.1521218	0.1524140
[15,]	61.42857	61.43432	3.298986e-05	0.1611401	0.1611731
[16,]	60.06985	60.09090	4.433539e-04	0.1665845	0.1670278
[17,]	58.73058	58.76453	1.152589e-03	0.1638415	0.1649941
[18,]	57.41840	57.46375	2.056943e-03	0.1589453	0.1610023
[19,]	56.13834	56.19188	2.866806e-03	0.1580046	0.1608714
[20,]	54.89362	54.95150	3.350361e-03	0.1607610	0.1641114
[21,]	53.68609	53.74603	3.592947e-03	0.1656308	0.1692238
[22,]	52.51665	52.57763	3.717490e-03	0.1685612	0.1722787
[23,]	51.38552	51.44645	3.712516e-03	0.1669156	0.1706281
[24,]	50.29240	50.35301	3.674022e-03	0.1631261	0.1668001
[25,]	49.23664	49.29842	3.817192e-03	0.1608110	0.1646282
[26,]	48.21737	48.28105	4.054985e-03	0.1609478	0.1650028
[27,]	47.23352	47.29560	3.853476e-03	0.1603791	0.1642326
[28,]	46.28396	46.33809	2.929925e-03	0.1569280	0.1598579
[29,]	45.36745	45.40988	1.799817e-03	0.1530531	0.1548529
[30,]	44.48276	44.51514	1.048804e-03	0.1500476	0.1510964
[31,]	43.62863	43.65582	7.392734e-04	0.1490936	0.1498329
[32,]	42.80382	42.83035	7.038485e-04	0.1526752	0.1533791
[33,]	42.00710	42.03560	8.118713e-04	0.1580898	0.1589017
[34,]	41.23731	41.26862	9.806015e-04	0.1602829	0.1612635
[35,]	40.49327	40.52760	1.178091e-03	0.1611291	0.1623072
[36,]	39.77390	39.81145	1.410201e-03	0.1634506	0.1648608
[37,]	39.07811	39.11880	1.655657e-03	0.1637832	0.1654389
[38,]	38.40489	38.44879	1.926910e-03	0.1603689	0.1622958
[39,]	37.75326	37.80175	2.350445e-03	0.1551426	0.1574930
[40,]	37.12230	37.17634	2.920197e-03	0.1490426	0.1519628
[41,]	36.51111	36.56767	3.198307e-03	0.1437620	0.1469603
[42,]	35.91885	35.97100	2.718687e-03	0.1420182	0.1447368
[43,]	35.34472	35.38615	1.716777e-03	0.1422030	0.1439197
[44,]	34.78794	34.81684	8.355525e-04	0.1426138	0.1434494
[45,]	34.24779	34.26733	3.818971e-04	0.1471867	0.1475686
[46,]	33.72357	33.74020	2.763999e-04	0.1553083	0.1555847
[47,]	33.21464	33.23456	3.967765e-04	0.1606175	0.1610143
[48,]	32.72036	32.74512	6.132583e-04	0.1596425	0.1602557
[49,]	32.24013	32.26544	6.404545e-04	0.1536680	0.1543084
[50,]	31.77340	31.79329	3.958334e-04	0.1464696	0.1468655
[51,]	31.31962	31.33191	1.509417e-04	0.1430751	0.1432260
[52,]	30.87829	30.88584	5.698340e-05	0.1473920	0.1474490
[53,]	30.44892	30.45656	5.834774e-05	0.1555374	0.1555957
[54,]	30.03104	30.04194	1.187457e-04	0.1597360	0.1598547
[55,]	29.62422	29.63893	2.163427e-04	0.1594630	0.1596793
[56,]	29.22803	29.24538	3.008693e-04	0.1588588	0.1591597
[57,]	28.84208	28.86014	3.262359e-04	0.1577220	0.1580482
[58,]	28.46599	28.48317	2.952892e-04	0.1563949	0.1566902
[59,]	28.09939	28.11519	2.496793e-04	0.1577812	0.1580309
[60,]	27.74194	27.75702	2.275219e-04	0.1582007	0.1584282
[61,]	27.39330	27.40938	2.584532e-04	0.1537398	0.1539982
[62,]	27.05317	27.07173	3.443021e-04	0.1499304	0.1502747
[63,]	26.72125	26.74126	4.004126e-04	0.1512355	0.1516360
[64,]	26.39724	26.41531	3.265530e-04	0.1528193	0.1531459
[65,]	26.08087	26.09450	1.858078e-04	0.1510707	0.1512565
[66,]	25.77189	25.78137	8.988191e-05	0.1492966	0.1493865
[67,]	25.47003	25.47679	4.572178e-05	0.1509544	0.1510002
[68,]	25.17507	25.17950	1.962244e-05	0.1547229	0.1547425
[69,]	24.88676	24.88766	7.946409e-07	0.1562766	0.1562774

[71,]	24.32928	24.31761	1.362081e-04	0.1544819	0.1546181
[72,]	24.05968	24.04274	2.868991e-04	0.1568903	0.1571772
[73,]	23.79593	23.77849	3.039027e-04	0.1605234	0.1608273
[74,]	23.53782	23.52510	1.619682e-04	0.1619950	0.1621569
[75,]	23.28520	23.27976	2.956081e-05	0.1598800	0.1599096
[76,]	23.03788	23.03935	2.167872e-06	0.1550745	0.1550767
[77,]	22.79570	22.80242	4.516241e-05	0.1510063	0.1510514
[78,]	22.55851	22.56892	1.083323e-04	0.1507477	0.1508560
[79,]	22.32616	22.33955	1.791499e-04	0.1542585	0.1544377
[80,]	22.09850	22.11522	2.795647e-04	0.1585638	0.1588434
[81,]	21.87539	21.89617	4.314995e-04	0.1612360	0.1616675
[82,]	21.65670	21.68181	6.304882e-04	0.1631678	0.1637983
[83,]	21.44230	21.47126	8.382377e-04	0.1640232	0.1648614
[84,]	21.23207	21.26250	9.257636e-04	0.1633643	0.1642900
[85,]	21.02589	21.05243	7.045734e-04	0.1632410	0.1639456
[86,]	20.82364	20.84012	2.718955e-04	0.1641491	0.1644210
[87,]	20.62521	20.62987	2.170027e-05	0.1652096	0.1652313
[88,]	20.43049	20.42824	5.082209e-06	0.1662462	0.1662513
[89,]	20.23940	20.23729	4.425237e-06	0.1676769	0.1676813
[90,]	20.05181	20.05247	4.350351e-07	0.1702598	0.1702602
[91,]	19.86765	19.86919	2.357556e-06	0.1705343	0.1705366
[92,]	19.68682	19.69133	2.035919e-05	0.1643657	0.1643861
[93,]	19.50922	19.53514	6.719019e-04	0.1561032	0.1567751
[94,]	19.33478	19.43480	1.000481e-02	0.1512991	0.1613039
[95,]	19.16341	19.45551	8.532420e-02	0.1504641	0.2357883
[96,]	18.99503	19.67220	4.585583e-01	0.1561536	0.6147119
[97,]	18.82957	20.01016	1.393795e+00	0.1731847	1.5669794
[98,]	18.66694	19.95041	1.647281e+00	0.2016577	1.8489386
[99,]	18.50709	18.51037	1.078711e-05	0.2241076	0.2241184
[100,]	18.34993	15.01008	1.115456e+01	0.1991217	11.3536811

Appendix C

	y_tru	y_expGM	bias2_GM	y_varGM	MSE_GM
[1,]	31.77340	31.15036	3.881714e-01	0.03872510	0.42689654
[2,]	49.23664	37.53848	1.368470e+02	0.04698000	136.89396584
[3,]	59.17431	43.72302	2.387424e+02	0.05374248	238.79616336
[4,]	64.82412	49.37605	2.386428e+02	0.05878074	238.70155176
[5,]	67.89474	54.24240	1.863862e+02	0.06232638	186.44857423
[6,]	69.35484	58.16760	1.251544e+02	0.06481528	125.21919367
[7,]	69.78362	61.10326	7.534862e+01	0.06661360	75.41523047
[8,]	69.54178	63.09260	4.159189e+01	0.06788373	41.65976923
[9,]	68.86121	64.24358	2.132253e+01	0.06862251	21.39114848
[10,]	67.89474	64.69854	1.021569e+01	0.06879671	10.28448838
[11,]	66.74506	64.60766	4.568494e+00	0.06846426	4.63695821
[12,]	65.48223	64.11014	1.882643e+00	0.06780427	1.95044736
[13,]	64.15455	63.32392	6.899454e-01	0.06704965	0.75699501
[14,]	62.79555	62.34246	2.052926e-01	0.06638261	0.27167521
[15,]	61.42857	61.23610	3.704424e-02	0.06587262	0.10291686
[16,]	60.06985	60.05586	1.957272e-04	0.06549729	0.06569302
[17,]	58.73058	58.83770	1.147380e-02	0.06521979	0.07669358
[18,]	57.41840	57.60654	3.539937e-02	0.06505856	0.10045793
[19,]	56.13834	56.37945	5.813142e-02	0.06510059	0.12323201
[20,]	54.89362	55.16796	7.526487e-02	0.06545699	0.14072186
[21,]	53.68609	53.97983	8.628550e-02	0.06619625	0.15248175
[22,]	52.51665	52.82014	9.210501e-02	0.06729594	0.15940095
[23,]	51.38552	51.69211	9.399706e-02	0.06863564	0.16263270
[24,]	50.29240	50.59762	9.316263e-02	0.07002937	0.16319200
[25,]	49.23664	49.53760	9.057453e-02	0.07127558	0.16185011
[26,]	48.21737	48.51223	8.694524e-02	0.07219685	0.15914209
[27,]	47.23352	47.52119	8.275063e-02	0.07265614	0.15540678
[28,]	46.28396	46.56374	7.828003e-02	0.07256181	0.15084184
[29,]	45.36745	45.63892	7.369605e-02	0.07188512	0.14558117
[30,]	44.48276	44.74561	6.909127e-02	0.07069314	0.13978441
[31,]	43.62863	43.88266	6.453049e-02	0.06916550	0.13369599
[32,]	42.80382	43.04891	6.007223e-02	0.06755694	0.12762917
[33,]	42.00710	42.24326	5.577022e-02	0.06610745	0.12187768
[34,]	41.23731	41.46460	5.166293e-02	0.06495277	0.11661571
[35,]	40.49327	40.71182	4.776206e-02	0.06409799	0.11186005
[36,]	39.77390	39.98377	4.404930e-02	0.06346969	0.10751899
[37,]	39.07811	39.27932	4.048489e-02	0.06300098	0.10348587
[38,]	38.40489	38.59731	3.702542e-02	0.06268358	0.09970900
[39,]	37.75326	37.93669	3.364409e-02	0.06255753	0.09620162
[40,]	37.12230	37.29650	3.034564e-02	0.06266148	0.09300712
[41,]	36.51111	36.67595	2.716995e-02	0.06298802	0.09015796
[42,]	35.91885	36.07436	2.418304e-02	0.06346869	0.08765173
[43,]	35.34472	35.49121	2.145891e-02	0.06398698	0.08544589
[44,]	34.78794	34.92599	1.905933e-02	0.06441120	0.08347054
[45,]	34.24779	34.37824	1.701886e-02	0.06464381	0.08166267
[46,]	33.72357	33.84743	1.533937e-02	0.06467435	0.08001373
[47,]	33.21464	33.33293	1.399417e-02	0.06460345	0.07859762
[48,]	32.72036	32.83410	1.293810e-02	0.06460629	0.07754438
[49,]	32.24013	32.35022	1.211919e-02	0.06484201	0.07696119
[50,]	31.77340	31.88058	1.148775e-02	0.06536376	0.07685151
[51,]	31.31962	31.42451	1.100072e-02	0.06609540	0.07709612
[52,]	30.87829	30.98135	1.062095e-02	0.06689517	0.07751611
[53,]	30.44892	30.55047	1.031325e-02	0.06765850	0.07797176
[54,]	30.03104	30.13124	1.003991e-02	0.06837979	0.07841970
[55,]	29.62422	29.72300	9.758507e-03	0.06912456	0.07888307
[56,]	29.22803	29.32511	9.424079e-03	0.06993373	0.07935781
[57,]	28.84208	28.93693	8.995694e-03	0.07073360	0.07972929
[58,]	28.46599	28.55789	8.445939e-03	0.07132339	0.07976933
[59,]	28.09939	28.18753	7.770036e-03	0.07145858	0.07922861
[60,]	27.74194	27.82555	6.990717e-03	0.07098242	0.07797313
[61,]	27.39330	27.47176	6.155915e-03	0.06992562	0.07608153
[62,]	27.05317	27.12617	5.328912e-03	0.06851776	0.07384667
[63,]	26.72125	26.78888	4.573903e-03	0.06711111	0.07168501
[64,]	26.39724	26.46002	3.942230e-03	0.06606130	0.07000353
[65,]	26.08087	26.13973	3.464156e-03	0.06561422	0.06907838
[66,]	25.77189	25.82799	3.147889e-03	0.06583113	0.06897902
[67,]	25.47003	25.52465	2.983653e-03	0.06656953	0.06955319
[68,]	25.17507	25.22937	2.948756e-03	0.06752807	0.07047683

[70,]	24.60490	24.66087	3.131723e-03	0.06872193	0.07185365
[71,]	24.32928	24.38645	3.268686e-03	0.06851419	0.07178288
[72,]	24.05968	24.11785	3.383189e-03	0.06777548	0.07115867
[73,]	23.79593	23.85463	3.446286e-03	0.06672071	0.07016699
[74,]	23.53782	23.59650	3.443097e-03	0.06563581	0.06907890
[75,]	23.28520	23.34327	3.372703e-03	0.06477020	0.06814291
[76,]	23.03788	23.09484	3.244323e-03	0.06426311	0.06750743
[77,]	22.79570	22.85113	3.072048e-03	0.06413402	0.06720607
[78,]	22.55851	22.61209	2.870372e-03	0.06432626	0.06719663
[79,]	22.32616	22.37766	2.651651e-03	0.06476167	0.06741332
[80,]	22.09850	22.14775	2.425376e-03	0.06536918	0.06779455
[81,]	21.87539	21.92228	2.198409e-03	0.06608145	0.06827986
[82,]	21.65670	21.70114	1.975082e-03	0.06682107	0.06879616
[83,]	21.44230	21.48421	1.756102e-03	0.06749942	0.06925552
[84,]	21.23207	21.27125	1.535273e-03	0.06803578	0.06957105
[85,]	21.02589	21.06185	1.293317e-03	0.06838714	0.06968046
[86,]	20.82364	20.85511	9.907761e-04	0.06857109	0.06956187
[87,]	20.62521	20.64924	5.774677e-04	0.06866480	0.06924227
[88,]	20.43049	20.44070	1.042146e-04	0.06877398	0.06887819
[89,]	20.23940	20.22302	2.682257e-04	0.06898635	0.06925458
[90,]	20.05181	19.98500	4.464488e-03	0.06933870	0.07380319
[91,]	19.86765	19.70874	2.525413e-02	0.06981324	0.09506736
[92,]	19.68682	19.36786	1.017350e-01	0.07034376	0.17207873
[93,]	19.50922	18.92694	3.390541e-01	0.07079415	0.40984826
[94,]	19.33478	18.34329	9.830499e-01	0.07090147	1.05395138
[95,]	19.16341	17.57206	2.532377e+00	0.07023045	2.60260724
[96,]	18.99503	16.57477	5.857672e+00	0.06820893	5.92588063
[97,]	18.82957	15.32997	1.224717e+01	0.06427865	12.31145180
[98,]	18.66694	13.84348	2.326580e+01	0.05812434	23.32392008
[99,]	18.50709	12.15448	4.035562e+01	0.04988309	40.40550795
[100,]	18.34993	10.33464	6.424491e+01	0.04021763	64.28512595

Appendix D

	y_tru	y_expSP	bias2_SP	y_varSP	MSE_SP
[1,]	31.77340	40.51862	7.647897e+01	11.31884	87.79781
[2,]	49.23664	40.51862	7.600381e+01	11.31884	87.32265
[3,]	59.17431	70.39456	1.258939e+02	110.54483	236.43874
[4,]	64.82412	67.02208	4.831047e+00	139.96766	144.79871
[5,]	67.89474	70.23779	5.489879e+00	140.26201	145.75189
[6,]	69.35484	70.01089	4.304003e-01	137.95184	138.38224
[7,]	69.78362	71.17167	1.926699e+00	139.01832	140.94501
[8,]	69.54178	69.92414	1.461965e-01	145.04715	145.19334
[9,]	68.86121	69.03201	2.917391e-02	148.93092	148.96010
[10,]	67.89474	68.37797	2.335148e-01	155.02154	155.25506
[11,]	66.74506	66.82594	6.541379e-03	150.24510	150.25164
[12,]	65.48223	65.56118	6.232264e-03	139.90695	139.91318
[13,]	64.15455	64.30224	2.181057e-02	137.34611	137.36792
[14,]	62.79555	62.75563	1.593582e-03	136.74785	136.74944
[15,]	61.42857	61.67008	5.832791e-02	138.14969	138.20802
[16,]	60.06985	59.65256	1.741278e-01	137.64166	137.81578
[17,]	58.73058	59.05415	1.046931e-01	139.75989	139.86458
[18,]	57.41840	57.24583	2.977927e-02	134.93746	134.96724
[19,]	56.13834	56.16796	8.770191e-04	132.58887	132.58975
[20,]	54.89362	54.80810	7.313305e-03	132.89440	132.90171
[21,]	53.68609	53.80515	1.417670e-02	135.67695	135.69112
[22,]	52.51665	52.27096	6.036552e-02	137.13789	137.19826
[23,]	51.38552	51.55406	2.840634e-02	138.93363	138.96204
[24,]	50.29240	50.12495	2.804022e-02	139.41202	139.44006
[25,]	49.23664	49.28099	1.967092e-03	139.83823	139.84020
[26,]	48.21737	48.06969	2.180803e-02	143.89482	143.91663
[27,]	47.23352	47.22624	5.302931e-05	136.68858	136.68863
[28,]	46.28396	46.37755	8.759523e-03	128.10639	128.11515
[29,]	45.36745	45.09142	7.619187e-02	132.99265	133.06884
[30,]	44.48276	44.91489	1.867352e-01	133.52996	133.71669
[31,]	43.62863	43.11058	2.683684e-01	128.98836	129.25672
[32,]	42.80382	43.13110	1.071147e-01	134.51448	134.62160
[33,]	42.00710	41.87436	1.762145e-02	137.91708	137.93470
[34,]	41.23731	41.11068	1.603476e-02	135.20459	135.22063
[35,]	40.49327	40.61573	1.499613e-02	137.58015	137.59515
[36,]	39.77390	39.53358	5.775240e-02	132.69363	132.75138
[37,]	39.07811	39.50763	1.844901e-01	132.64274	132.82723
[38,]	38.40489	38.13627	7.215479e-02	143.90644	143.97860
[39,]	37.75326	37.56489	3.548511e-02	148.61754	148.65303
[40,]	37.12230	37.51344	1.529879e-01	146.34098	146.49397
[41,]	36.51111	36.07435	1.907654e-01	141.90122	142.09198
[42,]	35.91885	36.43731	2.687976e-01	137.00344	137.27224
[43,]	35.34472	35.05824	8.207062e-02	133.01973	133.10180
[44,]	34.78794	34.78511	7.997099e-06	137.90064	137.90065
[45,]	34.24779	34.35466	1.142187e-02	143.07697	143.08839
[46,]	33.72357	33.72199	2.507095e-06	148.19715	148.19716
[47,]	33.21464	32.88928	1.058550e-01	146.98221	147.08806
[48,]	32.72036	33.04529	1.055808e-01	138.21186	138.31744
[49,]	32.24013	32.21962	4.204817e-04	133.78174	133.78216
[50,]	31.77340	31.41750	1.266633e-01	136.10432	136.23098
[51,]	31.31962	31.66696	1.206446e-01	140.20290	140.32355
[52,]	30.87829	31.00977	1.728607e-02	145.33298	145.35027
[53,]	30.44892	29.87566	3.286192e-01	139.83436	140.16298
[54,]	30.03104	30.68106	4.225304e-01	133.14168	133.56421
[55,]	29.62422	29.31252	9.715738e-02	132.19246	132.28962
[56,]	29.22803	29.09372	1.803899e-02	134.80414	134.82218
[57,]	28.84208	29.32696	2.351043e-01	142.85284	143.08794
[58,]	28.46599	27.79963	4.440304e-01	139.48834	139.93237
[59,]	28.09939	28.76090	4.376013e-01	139.74665	140.18425
[60,]	27.74194	27.22890	2.632087e-01	143.91364	144.17684
[61,]	27.39330	27.63424	5.805113e-02	143.43885	143.49690
[62,]	27.05317	27.07234	3.674226e-04	143.74535	143.74572
[63,]	26.72125	26.44720	7.509893e-02	144.03920	144.11430
[64,]	26.39724	26.62573	5.221087e-02	139.73843	139.79064
[65,]	26.08087	26.00650	5.530419e-03	136.79722	136.80275
[66,]	25.77189	25.95052	3.190946e-02	134.82185	134.85376
[67,]	25.47003	25.15630	9.842770e-02	135.10260	135.20102
[68,]	25.17507	25.41412	5.714714e-02	135.60026	135.65741

[70,]	24.60490	24.79991	3.802752e-02	142.59007	142.62810
[71,]	24.32928	23.94711	1.460497e-01	147.00363	147.14968
[72,]	24.05968	24.28285	4.980217e-02	149.18220	149.23200
[73,]	23.79593	24.06937	7.477347e-02	142.94436	143.01914
[74,]	23.53782	23.08783	2.024990e-01	136.92687	137.12937
[75,]	23.28520	23.63519	1.224970e-01	137.09954	137.22203
[76,]	23.03788	22.88153	2.444484e-02	136.37667	136.40112
[77,]	22.79570	22.99213	3.858248e-02	147.62952	147.66810
[78,]	22.55851	22.06118	2.473442e-01	154.04730	154.29465
[79,]	22.32616	23.00317	4.583433e-01	153.79714	154.25548
[80,]	22.09850	21.35223	5.569163e-01	147.21559	147.77251
[81,]	21.87539	22.54525	4.487064e-01	140.95894	141.40765
[82,]	21.65670	21.27885	1.427742e-01	139.85949	140.00226
[83,]	21.44230	21.50797	4.311448e-03	136.05713	136.06144
[84,]	21.23207	21.15593	5.797942e-03	139.36777	139.37357
[85,]	21.02589	21.10318	5.973582e-03	141.42993	141.43590
[86,]	20.82364	20.48473	1.148579e-01	143.14619	143.26105
[87,]	20.62521	21.02531	1.600852e-01	142.27884	142.43892
[88,]	20.43049	20.45887	8.050256e-04	144.99446	144.99526
[89,]	20.23940	19.86031	1.437048e-01	137.99528	138.13898
[90,]	20.05181	20.28717	5.539046e-02	133.34564	133.40103
[91,]	19.86765	20.00411	1.862137e-02	136.86381	136.88243
[92,]	19.68682	19.40411	7.992174e-02	139.91166	139.99158
[93,]	19.50922	19.85309	1.182455e-01	140.94482	141.06307
[94,]	19.33478	19.09901	5.558933e-02	143.15241	143.20800
[95,]	19.16341	19.13567	7.692042e-04	144.52121	144.52198
[96,]	18.99503	19.05069	3.097610e-03	142.96373	142.96683
[97,]	18.82957	18.98852	2.526699e-02	140.39779	140.42306
[98,]	18.66694	18.25962	1.659087e-01	140.36994	140.53585
[99,]	18.50709	18.74733	5.771668e-02	122.13085	122.18857
[100,]	18.34993	18.25873	8.316427e-03	40.07961	40.08793

Appendix E

	y_tru	y_expHARR3	bias2_HARR3	y_varHARR3	MSE_HARR3
[1,]	31.77340	56.56517	6.146321e+02	0.1709790	614.8030688
[2,]	49.23664	56.56517	5.370739e+01	0.1709790	53.8783666
[3,]	59.17431	56.56517	6.807603e+00	0.1709790	6.9785815
[4,]	64.82412	56.56517	6.821020e+01	0.1709790	68.3811836
[5,]	67.89474	56.56517	1.283590e+02	0.1709790	128.5299809
[6,]	69.35484	56.56517	1.635755e+02	0.1709790	163.7465111
[7,]	69.78362	68.54133	1.543272e+00	0.1789890	1.7222614
[8,]	69.54178	68.54133	1.000895e+00	0.1789890	1.1798841
[9,]	68.86121	68.54133	1.023222e-01	0.1789890	0.2813112
[10,]	67.89474	68.54133	4.180847e-01	0.1789890	0.5970737
[11,]	66.74506	68.54133	3.226587e+00	0.1789890	3.4055763
[12,]	65.48223	68.54133	9.358081e+00	0.1789890	9.5370697
[13,]	64.15455	60.91862	1.047123e+01	0.1783814	10.6496159
[14,]	62.79555	60.91862	3.522850e+00	0.1783814	3.7012313
[15,]	61.42857	60.91862	2.600467e-01	0.1783814	0.4384281
[16,]	60.06985	60.91862	7.204191e-01	0.1783814	0.8988005
[17,]	58.73058	60.91862	4.787519e+00	0.1783814	4.9659000
[18,]	57.41840	60.91862	1.225158e+01	0.1783814	12.4299641
[19,]	56.13834	52.61171	1.243713e+01	0.1540266	12.5911530
[20,]	54.89362	52.61171	5.207096e+00	0.1540266	5.3611231
[21,]	53.68609	52.61171	1.154287e+00	0.1540266	1.3083132
[22,]	52.51665	52.61171	9.035714e-03	0.1540266	0.1630623
[23,]	51.38552	52.61171	1.503540e+00	0.1540266	1.6575671
[24,]	50.29240	52.61171	5.379213e+00	0.1540266	5.5332396
[25,]	49.23664	52.61171	1.139109e+01	0.1540266	11.5451206
[26,]	48.21737	45.70822	6.295798e+00	0.1825839	6.4783823
[27,]	47.23352	45.70822	2.326540e+00	0.1825839	2.5091235
[28,]	46.28396	45.70822	3.314692e-01	0.1825839	0.5140531
[29,]	45.36745	45.70822	1.161259e-01	0.1825839	0.2987098
[30,]	44.48276	45.70822	1.501764e+00	0.1825839	1.6843477
[31,]	43.62863	45.70822	4.324719e+00	0.1825839	4.5073026
[32,]	42.80382	40.92617	3.525541e+00	0.1660584	3.6915990
[33,]	42.00710	40.92617	1.168412e+00	0.1660584	1.3344707
[34,]	41.23731	40.92617	9.680419e-02	0.1660584	0.2628626
[35,]	40.49327	40.92617	1.874024e-01	0.1660584	0.3534608
[36,]	39.77390	40.92617	1.327745e+00	0.1660584	1.4938038
[37,]	39.07811	40.92617	3.415350e+00	0.1660584	3.5814080
[38,]	38.40489	36.89999	2.264729e+00	0.1625099	2.4272388
[39,]	37.75326	36.89999	7.280819e-01	0.1625099	0.8905918
[40,]	37.12230	36.89999	4.942396e-02	0.1625099	0.2119339
[41,]	36.51111	36.89999	1.512222e-01	0.1625099	0.3137321
[42,]	35.91885	36.89999	9.626214e-01	0.1625099	1.1251314
[43,]	35.34472	36.89999	2.418863e+00	0.1625099	2.5813729
[44,]	34.78794	33.19851	2.526298e+00	0.1460601	2.6723577
[45,]	34.24779	33.19851	1.100994e+00	0.1460601	1.2470539
[46,]	33.72357	33.19851	2.756969e-01	0.1460601	0.4217569
[47,]	33.21464	33.19851	2.602664e-04	0.1460601	0.1463204
[48,]	32.72036	33.19851	2.286274e-01	0.1460601	0.3746875
[49,]	32.24013	33.19851	9.184818e-01	0.1460601	1.0645419
[50,]	31.77340	33.19851	2.030927e+00	0.1460601	2.1769875
[51,]	31.31962	30.19539	1.263907e+00	0.1786215	1.4425280
[52,]	30.87829	30.19539	4.663583e-01	0.1786215	0.6449798
[53,]	30.44892	30.19539	6.427786e-02	0.1786215	0.2428994
[54,]	30.03104	30.19539	2.700999e-02	0.1786215	0.2056315
[55,]	29.62422	30.19539	3.262350e-01	0.1786215	0.5048566
[56,]	29.22803	30.19539	9.357787e-01	0.1786215	1.1144003
[57,]	28.84208	27.97266	7.558929e-01	0.1665407	0.9224337
[58,]	28.46599	27.97266	2.433720e-01	0.1665407	0.4099127
[59,]	28.09939	27.97266	1.605991e-02	0.1665407	0.1826006
[60,]	27.74194	27.97266	5.323355e-02	0.1665407	0.2197743
[61,]	27.39330	27.97266	3.356548e-01	0.1665407	0.5021955
[62,]	27.05317	27.97266	8.454558e-01	0.1665407	1.0119965
[63,]	26.72125	25.78835	8.703033e-01	0.1545561	1.0248594
[64,]	26.39724	25.78835	3.707485e-01	0.1545561	0.5253046
[65,]	26.08087	25.78835	8.557065e-02	0.1545561	0.2401267
[66,]	25.77189	25.78835	2.709672e-04	0.1545561	0.1548270
[67,]	25.47003	25.78835	1.013252e-01	0.1545561	0.2558813
[68,]	25.17507	25.78835	3.761127e-01	0.1545561	0.5306688

[70,]	24.60490	23.90385	4.914733e-01	0.1754880	0.6669613
[71,]	24.32928	23.90385	1.809864e-01	0.1754880	0.3564743
[72,]	24.05968	23.90385	2.428295e-02	0.1754880	0.1997709
[73,]	23.79593	23.90385	1.164811e-02	0.1754880	0.1871361
[74,]	23.53782	23.90385	1.339771e-01	0.1754880	0.3094650
[75,]	23.28520	23.90385	3.827334e-01	0.1754880	0.5582213
[76,]	23.03788	22.44564	3.507528e-01	0.1557531	0.5065060
[77,]	22.79570	22.44564	1.225478e-01	0.1557531	0.2783009
[78,]	22.55851	22.44564	1.274174e-02	0.1557531	0.1684949
[79,]	22.32616	22.44564	1.427389e-02	0.1557531	0.1700270
[80,]	22.09850	22.44564	1.205024e-01	0.1557531	0.2762555
[81,]	21.87539	22.44564	3.251772e-01	0.1557531	0.4809303
[82,]	21.65670	21.14970	2.570515e-01	0.1783561	0.4354076
[83,]	21.44230	21.14970	8.561669e-02	0.1783561	0.2639728
[84,]	21.23207	21.14970	6.784972e-03	0.1783561	0.1851411
[85,]	21.02589	21.14970	1.532983e-02	0.1783561	0.1936860
[86,]	20.82364	21.14970	1.063185e-01	0.1783561	0.2846746
[87,]	20.62521	21.14970	2.750937e-01	0.1783561	0.4534499
[88,]	20.43049	19.85144	3.353065e-01	0.1458534	0.4811600
[89,]	20.23940	19.85144	1.505117e-01	0.1458534	0.2963652
[90,]	20.05181	19.85144	4.015034e-02	0.1458534	0.1860038
[91,]	19.86765	19.85144	2.628477e-04	0.1458534	0.1461163
[92,]	19.68682	19.85144	2.710048e-02	0.1458534	0.1729539
[93,]	19.50922	19.85144	1.171129e-01	0.1458534	0.2629663
[94,]	19.33478	19.85144	2.669365e-01	0.1458534	0.4127899
[95,]	19.16341	18.74744	1.730320e-01	0.1769207	0.3499527
[96,]	18.99503	18.74744	6.130212e-02	0.1769207	0.2382228
[97,]	18.82957	18.74744	6.745106e-03	0.1769207	0.1836658
[98,]	18.66694	18.74744	6.479514e-03	0.1769207	0.1834002
[99,]	18.50709	18.74744	5.776879e-02	0.1769207	0.2346895
[100,]	18.34993	18.74744	1.580133e-01	0.1769207	0.3349340

Appendix F

	y_tru	y_expDAU34	bias2_DAU34	y_varDAU34	MSE_DAU34
[1,]	31.77340	11.31468	4.185591e+02	0.01456763	418.57363589
[2,]	49.23664	31.18661	3.258037e+02	0.05382402	325.85754160
[3,]	59.17431	51.04958	6.601125e+01	0.11989234	66.13114646
[4,]	64.82412	59.81697	2.507157e+01	0.15687627	25.22844789
[5,]	67.89474	73.97201	3.693320e+01	0.22818391	37.16138748
[6,]	69.35484	88.43444	3.640310e+02	0.31499531	364.34602598
[7,]	69.78362	69.10524	4.602009e-01	0.09769224	0.55789313
[8,]	69.54178	64.58645	2.455531e+01	0.08273671	24.63804175
[9,]	68.86121	61.97427	4.742990e+01	0.11186860	47.54177210
[10,]	67.89474	66.16306	2.998721e+00	0.14696016	3.14568124
[11,]	66.74506	68.81692	4.292585e+00	0.19640299	4.48898827
[12,]	65.48223	69.18951	1.374388e+01	0.27285505	14.01673274
[13,]	64.15455	65.71744	2.442632e+00	0.12471963	2.56735157
[14,]	62.79555	62.22715	3.230790e-01	0.08253558	0.40561463
[15,]	61.42857	60.60545	6.775369e-01	0.10986952	0.78740639
[16,]	60.06985	59.38899	4.635732e-01	0.15389782	0.61747101
[17,]	58.73058	58.50146	5.249754e-02	0.20441389	0.25691143
[18,]	57.41840	57.08221	1.130213e-01	0.28694393	0.39996528
[19,]	56.13834	55.93189	4.262165e-02	0.21400981	0.25663146
[20,]	54.89362	54.86101	1.063112e-03	0.09060719	0.09167030
[21,]	53.68609	53.70600	3.963051e-04	0.08899275	0.08938905
[22,]	52.51665	52.54999	1.111056e-03	0.12336517	0.12447622
[23,]	51.38552	51.32681	3.447023e-03	0.15243237	0.15587940
[24,]	50.29240	50.14020	2.316401e-02	0.20824327	0.23140728
[25,]	49.23664	48.95342	8.021257e-02	0.28002410	0.36023667
[26,]	48.21737	48.23642	3.631481e-04	0.09326772	0.09363086
[27,]	47.23352	47.29094	3.297160e-03	0.09338736	0.09668452
[28,]	46.28396	46.33412	2.516285e-03	0.13204449	0.13456078
[29,]	45.36745	45.26171	1.118108e-02	0.16329501	0.17447608
[30,]	44.48276	44.24427	5.687623e-02	0.22049399	0.27737022
[31,]	43.62863	43.23138	1.578080e-01	0.29504034	0.45284835
[32,]	42.80382	42.68311	1.457060e-02	0.11207623	0.12664683
[33,]	42.00710	42.02553	3.396164e-04	0.08425243	0.08459204
[34,]	41.23731	41.25515	3.184989e-04	0.11246743	0.11278593
[35,]	40.49327	40.41843	5.601533e-03	0.14978220	0.15538373
[36,]	39.77390	39.58363	3.619979e-02	0.19890878	0.23510857
[37,]	39.07811	38.79059	8.266883e-02	0.27907011	0.36173894
[38,]	38.40489	38.26009	2.096630e-02	0.15498384	0.17595014
[39,]	37.75326	37.79116	1.436250e-03	0.08378211	0.08521836
[40,]	37.12230	37.18729	4.223343e-03	0.09742866	0.10165200
[41,]	36.51111	36.57256	3.774999e-03	0.13443527	0.13821027
[42,]	35.91885	35.90786	1.208411e-04	0.17517209	0.17529293
[43,]	35.34472	35.27906	4.310733e-03	0.24076304	0.24507377
[44,]	34.78794	34.69366	8.889130e-03	0.25810898	0.26699811
[45,]	34.24779	34.33942	8.397277e-03	0.09187619	0.10027347
[46,]	33.72357	33.85375	1.694476e-02	0.09147597	0.10842073
[47,]	33.21464	33.36961	2.401546e-02	0.12849034	0.15250580
[48,]	32.72036	32.80518	7.194693e-03	0.15811901	0.16531371
[49,]	32.24013	32.28070	1.646223e-03	0.21394481	0.21559103
[50,]	31.77340	31.75753	2.517038e-04	0.28586070	0.28611240
[51,]	31.31962	31.40237	6.847407e-03	0.09151675	0.09836415
[52,]	30.87829	30.96617	7.723146e-03	0.08534980	0.09307295
[53,]	30.44892	30.52650	6.018998e-03	0.11800509	0.12402409
[54,]	30.03104	30.05043	3.757967e-04	0.15051112	0.15088692
[55,]	29.62422	29.58550	1.499047e-03	0.20167288	0.20317193
[56,]	29.22803	29.12847	9.911658e-03	0.27734262	0.28725427
[57,]	28.84208	28.81824	5.683142e-04	0.12266441	0.12323272
[58,]	28.46599	28.50313	1.379331e-03	0.08171336	0.08309269
[59,]	28.09939	28.12981	9.257788e-04	0.10826938	0.10919516
[60,]	27.74194	27.74210	2.758246e-08	0.14940599	0.14940601
[61,]	27.39330	27.34818	2.035930e-03	0.19936586	0.20140179
[62,]	27.05317	26.96997	6.923137e-03	0.28030726	0.28723040
[63,]	26.72125	26.66959	2.668340e-03	0.19795058	0.20061892
[64,]	26.39724	26.42443	7.395964e-04	0.09650274	0.09724234
[65,]	26.08087	26.12775	2.197193e-03	0.09630205	0.09849924
[66,]	25.77189	25.82839	3.193279e-03	0.12463214	0.12782542
[67,]	25.47003	25.49370	5.603052e-04	0.15399061	0.15455091
[68,]	25.17507	25.17901	1.557002e-05	0.20408053	0.20409610

[70,]	24.60490	24.65019	2.050733e-03	0.09023786	0.09228859
[71,]	24.32928	24.38520	3.127692e-03	0.09274798	0.09587567
[72,]	24.05968	24.11928	3.551290e-03	0.13274816	0.13629945
[73,]	23.79593	23.82766	1.007045e-03	0.16415793	0.16516497
[74,]	23.53782	23.54902	1.253668e-04	0.22381336	0.22393873
[75,]	23.28520	23.27055	2.146034e-04	0.29983403	0.30004864
[76,]	23.03788	23.09179	2.906834e-03	0.11179918	0.11470602
[77,]	22.79570	22.88296	7.612890e-03	0.09042504	0.09803793
[78,]	22.55851	22.65639	9.580362e-03	0.12149001	0.13107037
[79,]	22.32616	22.41368	7.659529e-03	0.16099527	0.16865480
[80,]	22.09850	22.17205	5.409439e-03	0.21304213	0.21845157
[81,]	21.87539	21.93920	4.071574e-03	0.29661359	0.30068516
[82,]	21.65670	21.74345	7.525426e-03	0.15349323	0.16101865
[83,]	21.44230	21.55214	1.206338e-02	0.08685876	0.09892214
[84,]	21.23207	21.34320	1.234913e-02	0.10205820	0.11440733
[85,]	21.02589	21.13292	1.145669e-02	0.14041829	0.15187498
[86,]	20.82364	20.91644	8.613010e-03	0.18303976	0.19165277
[87,]	20.62521	20.70500	6.366287e-03	0.25295505	0.25932133
[88,]	20.43049	20.51008	6.334604e-03	0.24168047	0.24801508
[89,]	20.23940	20.36339	1.537320e-02	0.08880886	0.10418206
[90,]	20.05181	20.18761	1.844112e-02	0.08919506	0.10763618
[91,]	19.86765	20.01250	2.098211e-02	0.12923471	0.15021683
[92,]	19.68682	19.81694	1.693355e-02	0.15766257	0.17459612
[93,]	19.50922	19.63184	1.503468e-02	0.21591333	0.23094801
[94,]	19.33478	19.44684	1.255763e-02	0.29055874	0.30311637
[95,]	19.16341	19.34851	3.426410e-02	0.09495114	0.12921524
[96,]	18.99503	19.20591	4.447073e-02	0.09498029	0.13945102
[97,]	18.82957	19.06280	5.439774e-02	0.13229641	0.18669415
[98,]	18.66694	18.90098	5.477544e-02	0.16738737	0.22216281
[99,]	18.50709	18.74592	5.704232e-02	0.22379720	0.28083952
[100,]	18.34993	18.59393	5.953522e-02	0.30324519	0.36278041

Appendix G

	y_tru	y_expDAU34_WLS	bias2_DAU34_WLS	y_varDAU34_WLS	MSE_DAU34_WLS
[1,]	31.77340	12.72683	3.627718e+02	0.01569225	362.78750093
[2,]	49.23664	33.98444	2.326296e+02	0.05950288	232.68908821
[3,]	59.17431	55.23248	1.553802e+01	0.13379980	15.67182326
[4,]	64.82412	64.60073	4.990453e-02	0.17548638	0.22539091
[5,]	67.89474	79.73741	1.402490e+02	0.25597537	140.50495859
[6,]	69.35484	95.20320	6.681378e+02	0.35407884	668.49185668
[7,]	69.78362	72.40093	6.850304e+00	0.10935136	6.95965534
[8,]	69.54178	66.37461	1.003094e+01	0.09036463	10.12130667
[9,]	68.86121	62.50789	4.036470e+01	0.11993113	40.48463039
[10,]	67.89474	66.34471	2.402596e+00	0.15731133	2.55990699
[11,]	66.74506	68.44287	2.882549e+00	0.20989124	3.09244039
[12,]	65.48223	67.95704	6.124689e+00	0.29130668	6.41599525
[13,]	64.15455	65.01286	7.366887e-01	0.13789091	0.87457962
[14,]	62.79555	62.05462	5.489746e-01	0.08985933	0.63883397
[15,]	61.42857	60.45586	9.461579e-01	0.11380041	1.05995836
[16,]	60.06985	59.15191	8.426085e-01	0.15436709	0.99697563
[17,]	58.73058	58.08738	4.137119e-01	0.20140711	0.61511904
[18,]	57.41840	56.63592	6.122761e-01	0.27858227	0.89085838
[19,]	56.13834	55.53888	3.593563e-01	0.20442038	0.56377668
[20,]	54.89362	54.65096	5.888093e-02	0.08881086	0.14769180
[21,]	53.68609	53.58318	1.058978e-02	0.09332485	0.10391463
[22,]	52.51665	52.51317	1.213520e-05	0.13627917	0.13629130
[23,]	51.38552	51.29967	7.371157e-03	0.16785972	0.17523088
[24,]	50.29240	50.16434	1.639957e-02	0.23168931	0.24808888
[25,]	49.23664	49.02866	4.325602e-02	0.31292368	0.35617969
[26,]	48.21737	48.15288	4.158370e-03	0.09885375	0.10301212
[27,]	47.23352	47.15061	6.875200e-03	0.09487389	0.10174909
[28,]	46.28396	46.14207	2.013067e-02	0.13298728	0.15311795
[29,]	45.36745	45.06957	8.873171e-02	0.16476781	0.25349952
[30,]	44.48276	44.02747	2.072834e-01	0.22344633	0.43072970
[31,]	43.62863	42.98791	4.105227e-01	0.30058043	0.71110310
[32,]	42.80382	42.48788	9.981510e-02	0.11985790	0.21967300
[33,]	42.00710	41.86146	2.121203e-02	0.09122786	0.11243989
[34,]	41.23731	41.10442	1.765811e-02	0.11865077	0.13630888
[35,]	40.49327	40.27055	4.960701e-02	0.15613371	0.20574072
[36,]	39.77390	39.43891	1.122153e-01	0.20515899	0.31737432
[37,]	39.07811	38.65562	1.784976e-01	0.28333603	0.46183367
[38,]	38.40489	38.14895	6.550552e-02	0.15684055	0.22234607
[39,]	37.75326	37.70749	2.094838e-03	0.08676657	0.08886141
[40,]	37.12230	37.12318	7.634208e-07	0.10245214	0.10245290
[41,]	36.51111	36.52736	2.638030e-04	0.14266299	0.14292679
[42,]	35.91885	35.87864	1.616968e-03	0.18614100	0.18775797
[43,]	35.34472	35.26793	5.896327e-03	0.25660408	0.26250041
[44,]	34.78794	34.69849	8.000388e-03	0.27530867	0.28330906
[45,]	34.24779	34.35113	1.067981e-02	0.09977979	0.11045961
[46,]	33.72357	33.87754	2.370538e-02	0.10115573	0.12486111
[47,]	33.21464	33.40542	3.639884e-02	0.14303400	0.17943284
[48,]	32.72036	32.85620	1.845480e-02	0.17422546	0.19268026
[49,]	32.24013	32.34536	1.107246e-02	0.23474488	0.24581734
[50,]	31.77340	31.83576	3.888680e-03	0.31248231	0.31637099
[51,]	31.31962	31.44575	1.590851e-02	0.10165696	0.11756547
[52,]	30.87829	30.99799	1.432819e-02	0.09356729	0.10789549
[53,]	30.44892	30.54778	9.773892e-03	0.12692444	0.13669833
[54,]	30.03104	30.07161	1.645976e-03	0.16165015	0.16329612
[55,]	29.62422	29.60340	4.332611e-04	0.21587739	0.21631066
[56,]	29.22803	29.14081	7.607510e-03	0.29593541	0.30354292
[57,]	28.84208	28.83360	7.195009e-05	0.13251106	0.13258301
[58,]	28.46599	28.52125	3.054175e-03	0.08578676	0.08884093
[59,]	28.09939	28.14770	2.333712e-03	0.10979559	0.11212931
[60,]	27.74194	27.75900	2.910613e-04	0.14864188	0.14893294
[61,]	27.39330	27.36377	8.724391e-04	0.19621806	0.19709050
[62,]	27.05317	26.98506	4.640027e-03	0.27435829	0.27899831
[63,]	26.72125	26.68465	1.338998e-03	0.18704140	0.18838040
[64,]	26.39724	26.44001	1.829753e-03	0.08838334	0.09021310
[65,]	26.08087	26.14337	3.906733e-03	0.09369272	0.09759945
[66,]	25.77189	25.84405	5.207768e-03	0.13091276	0.13612052
[67,]	25.47003	25.50906	1.523592e-03	0.16425575	0.16577934
[68,]	25.17507	25.19426	3.684727e-04	0.22440915	0.22477762

[70,]	24.60490	24.65930	2.958743e-03	0.09489953	0.09785827
[71,]	24.32928	24.39127	3.843255e-03	0.09318149	0.09702475
[72,]	24.05968	24.12236	3.928325e-03	0.13123495	0.13516328
[73,]	23.79593	23.82926	1.111125e-03	0.16199419	0.16310532
[74,]	23.53782	23.54838	1.113969e-04	0.22070703	0.22081842
[75,]	23.28520	23.26766	3.077476e-04	0.29604371	0.29635146
[76,]	23.03788	23.11116	5.369406e-03	0.11666125	0.12203065
[77,]	22.79570	22.91719	1.475770e-02	0.09671945	0.11147715
[78,]	22.55851	22.70118	2.035210e-02	0.12703765	0.14738975
[79,]	22.32616	22.46507	1.929596e-02	0.16589797	0.18519393
[80,]	22.09850	22.23031	1.737328e-02	0.21669059	0.23406387
[81,]	21.87539	22.00648	1.718495e-02	0.29700744	0.31419239
[82,]	21.65670	21.82493	2.830106e-02	0.15943217	0.18773323
[83,]	21.44230	21.64848	4.250700e-02	0.09357084	0.13607784
[84,]	21.23207	21.45180	4.828155e-02	0.10657186	0.15485340
[85,]	21.02589	21.25359	5.184913e-02	0.14148140	0.19333053
[86,]	20.82364	21.04826	5.045743e-02	0.18093655	0.23139398
[87,]	20.62521	20.84871	4.995419e-02	0.24558432	0.29553851
[88,]	20.43049	20.65996	5.265453e-02	0.23497439	0.28762892
[89,]	20.23940	20.50393	6.997877e-02	0.09256397	0.16254274
[90,]	20.05181	20.32822	7.639795e-02	0.09289652	0.16929447
[91,]	19.86765	20.15294	8.139183e-02	0.13019435	0.21158618
[92,]	19.68682	19.96382	7.673165e-02	0.15733278	0.23406443
[93,]	19.50922	19.78177	7.428317e-02	0.21215459	0.28643776
[94,]	19.33478	19.59980	7.023380e-02	0.28240163	0.35263543
[95,]	19.16341	19.51529	1.238211e-01	0.09809907	0.22192018
[96,]	18.99503	19.38099	1.489684e-01	0.10253885	0.25150723
[97,]	18.82957	19.24612	1.735200e-01	0.14494409	0.31846410
[98,]	18.66694	19.09022	1.791671e-01	0.18391405	0.36308117
[99,]	18.50709	18.94192	1.890785e-01	0.24573286	0.43481137
[100,]	18.34993	18.79706	1.999247e-01	0.33178944	0.53171412