The Estimating Nonlinear Autoregressive Model and Conditional Heteroscedastic Nonlinear Autoregressive model by Using Smoothing Spline and Penalized Spline Methods

Autcha Araveeporn [1]

Last revised on: October 5, 2015

[1]Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand. *E-mail : kaautcha@kmitl.ac.th*

Abstract

The objective of this research is to compare the estimation of the NonLinear AutoRegressive (NLAR) model and Conditional Heteroscedastic Nonlinear Autoregressive (CHNLAR)model with Smoothing Spline (SS) and Penalized Spline (PS) methods in a class of nonparametric regression method. NLAR model consists of a response variable and a function of predictor variable as a past of response variable. CHNLAR model consists of the function of trend and heteroscedastic (volatility) as a past of response variable at lag 1. Moreover, the nonparametric regression method has been developed the smoothing technique which produces a smoother based on NLAR and CHNALR model. The SS and PS methods are computed to fit NLAR and CHNLAR model with stationary and nonstationary time series data.

For simulation study of NLAR model, the data is generated by the autoregressive process with several coefficient autocorrelations and sample sizes. The performance of SS and PS methods is used the criterion by minimizing the average Mean Square Error (MSE) values. The SS method exhibits a good power estimation in all cases of stationary and nonstationary data. For economic data, the gold price is an important factor for pretty much all of the world market. The gold price (US Dollars per Troy Ounce) is then applied by using SS and PS methods that collected in term of the monthly volume from January, 1984 to December 2013. The result is founded that the SS method performs better than PS method which is similar the result in case of simulation study.

For simulation study of CHNLAR model, the data is simulated by the trend and heteroscedastic functions. The performance of SS and PS methods is used the hypothesis testing by the bias of the trend and heteroscedastic estimators. The SS and PS methods exhibit a good power estimation in most cases of generated data. For real data, the gold price (US Dollars per Troy Ounce) is then applied by using SS and PS methods. The results show that the SS method performs similar the PS method which is similar the result in case of simulation study.

Keywords: Conditional Heteroscedastic Nonlinear Autoregressive model, NonLinear AutoRegressive model; Smoothing Spline Method; Penalized Spline Method.

Chapter 1

Nonlinear Autoregressive Model

1.1 Introduction

Parametric regression method is widely used for estimating regression function when dependent variable and independent variable are focus on the relationship. More specifically, parametric regression method requires classical assumptions such as, the relationship between dependent variable and independent variable is linearity and additivity, the error terms are independent, the variance of the error terms are homoscedastity, and the error terms are normality.

To overcome these assumptions, nonparametric regression method is a choice for estimating function when the data may not be available to use. The basic idea of nonparametric regression method is to let the dependent variable produce a smoothing function from independent variable. The smoothing function is depended on the smoothing parameter which is controlled the trade-off between fidelity to data and roughness of function.

There exists a vast on nonparametric regression methods. Popular methods can be used by smoothing spline method (Wahba (1990), Green and Silverman (1994)) and penalized spline method (Ruppert et al. (2003)). The smoothing spline method base on the natural cubic spline that used each data points to fit smoothing function and minimize the penalized sum of squares to fit smoothing function. The penalized spline method has found a lot application in recent year, which is used for fitting and flexible choice of knots and smoothing parameter in nonparametric regression model. Ruppert et al.(2003) described penalized spline models, based on reduce-knot truncated power function basis with penalties on the untransformed coefficients, fitted as a mixed model, and motivated as a simple low-rank smoothing spline.

Normally, the nonparametric regression model consists of dependent variable and smoothing function or called mean function in terms of independent variable which may be acted a parametric and nonparametric variables. Autoregressive model is an important class of time series model that specified the present value based on the its over previous values. NonLinear AutoRegressive (NLAR) Model is developed from the autoregressive model and mixed the mean function of nonparametric regression model in terms over the previous values. Some of the classical nonlinear autoregressive models were proposed by Tong (1983), Haggan and Ozaki (1981), Chan and Tong (1986), and Granger and Teräsvirta (1993).

In this case, we define the observations of present values as the dependent variables, and the mean function corresponds the first of previous values as the smoothing function based on nonparametric regression model. We propose the smoothing spline method and penalized spline method to approximate smoothing function.

1.2 NonLinear AutoRegressive (NLAR) Model

The nonlinear autoregressive model is written as

$$y_t = \mu(y_{t-1}) + \varepsilon_t, \qquad t = 2, 3, \ldots, n, \tag{1.1}$$

where $y_t, t = 2, 3, \ldots, n$ are know the dependent variables, $y_{t-1}, t = 2, 3, \ldots, n$ are the past of independent variables at lag 1, $\mu(y_{t-1})$ are the mean function of nonlinear autoregressive model, and $\varepsilon_t,\ t = 2, 3, \ldots, n$ denote the error terms.

1.3 Nonparametric Regression Method

The popular nonparametric regression methods include smoothing spline method and penalized spline method. The concept of these methods is to let the observed data interpolate the most suitable form the fitting function by smoothing parameter.

1.3.1 Smoothing Spline Method

The smoothing spline was studied by Wahba (1990) that the smoothing spline estimator is estimated the natural polynomial spline ($S_\lambda^{(K)}(\mu)$) which is depended on the smoothing parameter (λ) following;

$$S_\lambda^{(K)}(\mu) = \sum_{t=1}^{n} \{y_t - \mu(x_t)\}^2 + \lambda \int_a^b \left\{\mu^{(m)}(x_t)\right\}^2 dx_t, \tag{1.2}$$

where K is the number of knots on mean function with domain $[a, b]$, m is the mth derivative of $\mu(x_t)$, y_t is the dependent variables, and $\mu(x_t)$ is mean function in a class of nonparametric regression function with independent variables.

Green and Silverman (1994) emphasized $m = 2$ so-called the natural cubic spline to fit the nonparametric regression function by minimizing

$$S_\lambda^{(K)}(\mu) = \sum_{t=1}^{n} \{y_t - \mu(x_t)\}^2 + \lambda \int_a^b \{\mu''(x_t)\}^2 dx_t. \tag{1.3}$$

In this case, we propose the NLAR model via smoothing spline method, and the natural cubic spline can be written as

$$S_\lambda^{(K)}(\mu) = \sum_{t=2}^{n} \{y_t - \mu(y_{t-1})\}^2 + \lambda \int_a^b \{\mu''(y_{t-1})\}^2 dy_t. \tag{1.4}$$

The natural cubic spline is given the value and second derivatives at each knots y_t as

$$\begin{aligned} \boldsymbol{\mu} &= \mu(y_{t-1}), \\ \boldsymbol{\gamma} &= \mu''(y_{t-1}), \quad t = 2, 3, \ldots, n \end{aligned}$$

Let $\boldsymbol{\mu}$ be the vector $(\mu_1, \ldots, \mu_{n-1})^\top$ and let $\boldsymbol{\gamma}$ be the vector $(\gamma_1, \ldots, \gamma_{n-1})^\top$.

The condition of natural cubic spline depends on two matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ below

$$\boldsymbol{Q} = \begin{pmatrix} h_1^{-1} & 0 & \dots & 0 \\ -h_1^{-1} - h_2^{-1} & h_2^{-1} & \dots & 0 \\ h_2^{-1} & -h_2^{-1} - h_3^{-1} & \dots & 0 \\ 0 & h_3^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h_{n-1}^{-1} \end{pmatrix}_{(n-1)\times(n-3)},$$

where $h_t = y_{t+1} - y_t$, for $t = 2, \dots, n-1$, then $\boldsymbol{Q}$ is an $(n-1) \times (n-3)$ matrix.

$\boldsymbol{R}$ is a symmetric $(n-3) \times (n-3)$ matrix with elements below

$$\boldsymbol{R} = \begin{pmatrix} \frac{1}{3}(h_2 + h_3) & \frac{1}{6}h_3 & \dots & 0 \\ \frac{1}{6}h_3 & \frac{1}{3}(h_3 + h_4) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{3}(h_{n-2} + h_{n-1}) \end{pmatrix}_{(n-3)\times(n-3)}.$$

The matrix $\boldsymbol{K}$ can be decomposed by

$$\boldsymbol{K} = \boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Q}^\top. \tag{1.5}$$

The vector $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ specify a natural cubic spline $\mu(y_t)$ if and only if the condition

$$\boldsymbol{Q}^\top \boldsymbol{\mu} = \boldsymbol{R}\,\boldsymbol{\gamma} \tag{1.6}$$

is satisfied. If (1.6) is satisfied the roughness penalty will satisfy

$$\int_b^a \{\mu''(y_{t-1})\}^2 dy_t = \boldsymbol{\gamma}^\top \boldsymbol{R}\,\boldsymbol{\gamma} = \boldsymbol{\mu}^\top \boldsymbol{K}\,\boldsymbol{\mu}. \tag{1.7}$$

To illustrate, it can be written in matrix form introduced by Green and Silverman (1994) as

$$\text{RSS} = \sum_{t=2}^{n} \{y_t - \mu(y_{t-1})\}^2 = (\boldsymbol{y} - \boldsymbol{\mu})^\top (\boldsymbol{y} - \boldsymbol{\mu}), \tag{1.8}$$

where $\boldsymbol{y} = (y_2, \dots, y_n)^\top$ with y_t corresponding value to y_{t-1} and

$$\boldsymbol{\mu} = (\mu(y_1), \dots, \mu(y_{n-1}))^\top. \tag{1.9}$$

The roughness penalty term $\int \mu''^2$ as $\boldsymbol{\mu}^\top \boldsymbol{K}\boldsymbol{\mu}$ in (1.7) to obtain

$$\begin{aligned} S_\lambda(\mu) &= (\boldsymbol{y}-\boldsymbol{\mu})^\top(\boldsymbol{y}-\boldsymbol{\mu}) + \lambda\, \boldsymbol{\mu}^\top \boldsymbol{K}\boldsymbol{\mu} \\ &= \boldsymbol{\mu}^\top(\boldsymbol{I}+\lambda\boldsymbol{K})\boldsymbol{\mu} - 2\boldsymbol{y}^\top\boldsymbol{\mu} + \boldsymbol{y}^\top\boldsymbol{y}, \end{aligned} \tag{1.10}$$

since $\lambda\boldsymbol{K}$ is non-negative definite, the matrix $\boldsymbol{I}+\lambda\boldsymbol{K}$ is strictly positive definite. It therefore follows that (1.10) has a unique minimum, other smoothing spline estimator is obtained by

$$\hat{\boldsymbol{\mu}}_\lambda = (\boldsymbol{I}+\lambda\boldsymbol{K})^{-1}\boldsymbol{y}, \tag{1.11}$$

where $\boldsymbol{I}$ denote the n-dimensional identity matrix.

The smoothing spline estimator depends on the smoothing parameter λ, Allen (1974) and Stone (1974) proposed the cross-validation(CV) method to select the smoothing parameter(λ) by the cross-validation criterion.

At first sight, it appears from (1.11) that the smoothing spline estimator $\hat{\boldsymbol{\mu}}_\lambda$ depends linearly on the data y_t through the equation

$$\begin{aligned} \hat{\boldsymbol{\mu}}_\lambda &= (\boldsymbol{I}+\lambda\boldsymbol{K})^{-1}\boldsymbol{y} \\ &= A(\lambda)\boldsymbol{y}. \end{aligned} \tag{1.12}$$

The matrix $A(\lambda)$ is defines by

$$A(\lambda) = (\boldsymbol{I}+\lambda\boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Q}^\top)^{-1}. \tag{1.13}$$

Wahba (1977) and Craven and Wahba (1979) suggested to replace the ordinary cross-validation to Generalized Cross-Validation(GCV) for choosing the smoothing parameter and it can be written as

$$GCV(\lambda) = n^{-1}\frac{\sum_{t=1}^n \{y_t - \hat{\mu}_\lambda(y_{t-1})\}^2}{\{1-n^{-1}\ \mathrm{tr}A(\lambda)\}^2}. \tag{1.14}$$

Just as in ordinary cross-validation, the GCV choice of smoothing parameter is then carried out by minimizing the function $GCV(\lambda)$ over λ.

1.3.2 Penalized Spline Method

Penalized spline smoother is estimated using the truncated power function (Ruppert and Carroll (2000)), and the penalized spline model is written as

$$\mu(x_t) = \sum_{j=0}^{m-1} \alpha_j x_t^j + \sum_{k=1}^{K} \beta_k (x_t - \tau_k)^{2m-1}, t = 1, 2, \ldots, n \tag{1.15}$$

The natural cubic spline is denoted $m = 2$ or called low-rank thin-plate spline which tend to have very good numerical properties. The low-rank thin-plate spline representation of $\mu(.)$ is

$$\mu(x_t, \boldsymbol{\theta}) = \alpha_0 + \alpha_1 x_t + \sum_{k=1}^{K} \beta_k |x_t - \tau_k|^3, t = 1, 2, \ldots, n \tag{1.16}$$

where $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1, \ldots, \beta_K)^\top$ is the vector of regression coefficients, and $\tau_1 < \tau_2 < \ldots < \tau_K$ are fixed knots.

In this case, we focus the NLAR model based on penalized spline method, then the natural cubic spline can be written as

$$\mu(y_{t-1}, \boldsymbol{\theta}) = \alpha_0 + \alpha_1 y_{t-1} + \sum_{k=1}^{K} \beta_k |y_{t-1} - \tau_k|^3, t = 2, 3, \ldots, n \tag{1.17}$$

To avoid overfitting, we minimize

$$\sum_{t=1}^{n} \{y_t - \mu(y_{t-1}, \boldsymbol{\theta})\}^2 + \frac{1}{\lambda} \boldsymbol{\theta}^\top \boldsymbol{D} \boldsymbol{\theta} \tag{1.18}$$

where λ is the smoothing parameter and $\boldsymbol{D}$ is a known positive semi-definite penalty matrix. The thin-plate spline penalty matrix is

$$\boldsymbol{D} = \begin{bmatrix} \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times K} \\ \mathbf{0}_{K\times 2} & \boldsymbol{\Omega}_K \end{bmatrix} \tag{1.19}$$

where the (l, k)th entry of $\boldsymbol{\Omega}$ is $|\tau_l - \tau_k|^3$ and penalized only coefficient of $|y_{t-1} - \tau_k|^3$.

Just as with the linear model, we can generalize penalized spline in general linear mixed model(Brumback et al. (1999)) is

$$\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{\alpha} + \boldsymbol{Z}_K \boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1.20}$$

where $\boldsymbol{y} = (y_2, \ldots, y_n)^\top$, $\boldsymbol{Y}$ be the matrix with the $t-1$th row $\boldsymbol{Y}_t = (1, y_{t-1})$, $\boldsymbol{Z}_K$ be the matrix with the tth row $\boldsymbol{Z}_{Kt} = \{|y_{t-1}-\tau_1|^3, \ldots, |y_{t-1}-\tau_K|^3\}$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^\top$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)^\top$, and $\boldsymbol{\epsilon}$ is $N(0, \sigma_\epsilon^2 I)$. Consider the vector $\boldsymbol{\alpha}$ as a fixed parameters and the vector $\boldsymbol{\beta}$ as a set of random parameters with $\mathrm{E}(\boldsymbol{\beta}) = 0$ and $\mathrm{cov}(\boldsymbol{\beta}) = \sigma_\beta^2$. This class of penalized spline smoothers ($\hat{\mu}(.)$) may also be expressed as

$$\hat{\boldsymbol{\mu}} = \boldsymbol{C}(\boldsymbol{C}^\top \boldsymbol{C} + \lambda^3 \boldsymbol{D})^{-1} \boldsymbol{C}^\top \boldsymbol{y} \tag{1.21}$$

where

$$\boldsymbol{C} = \begin{bmatrix} 1 & y_{t-1} & |y_{t-1}-\tau_k|^3_{1\le k\le K} \end{bmatrix}_{1\le t\le n}, \boldsymbol{D} = \begin{bmatrix} \boldsymbol{0}_{2\times 2} & \boldsymbol{0}_{2\times K} \\ \boldsymbol{0}_{K\times 2} & (\boldsymbol{\Omega}_K^{1/2})^\top \boldsymbol{\Omega}_K^{1/2} \end{bmatrix}, \tag{1.22}$$

and $\lambda = \sigma_\beta^2/\sigma_\varepsilon^2$ is a smoothing parameter.

Restricted maximum likelihood (REML) is a method for smoothing parameter selection in penalized spline that have become the most common strategies for estimating the parameter in covariance matrices.

Recall from (1.20), for general linear mixed model

$$\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{\alpha} + \boldsymbol{Z}_K \boldsymbol{\beta} + \boldsymbol{\epsilon}$$

the REML criterion for estimating covariance matrix parameters is

$$\begin{aligned} \ell_R(\boldsymbol{V}) &= -\frac{1}{2}[n\log(2\pi) + \log|\boldsymbol{V}| + \log|\boldsymbol{Y}^\top \boldsymbol{V}^{-1}\boldsymbol{Y}| \\ &\quad + \boldsymbol{y}^\top \boldsymbol{V}^{-1}\{\boldsymbol{I} - \boldsymbol{Y}(\boldsymbol{Y}^\top \boldsymbol{V}^{-1}\boldsymbol{Y})^{-1}\boldsymbol{Y}^\top \boldsymbol{V}^{-1}\}\boldsymbol{y}], \end{aligned} \tag{1.23}$$

where $\boldsymbol{V} = \mathrm{Cov}(\boldsymbol{y}) = \sigma_\beta^2 \boldsymbol{Z}_K \boldsymbol{Z}_K^\top + \sigma_\epsilon^2 I$. Therefore, if the REML criterion is minimized over $\sigma_\beta^2, \sigma_\epsilon^2 \ge 0$ to give estimates $(\sigma^2_{\beta,\mathrm{REML}}, \sigma^2_{\epsilon,\mathrm{REML}})$, then the smoothing parameter for REML is

$$\hat{\lambda}_{REML} = \sigma^2_{\beta,\mathrm{REML}}/\sigma^2_{\epsilon,\mathrm{REML}} \tag{1.24}$$

1.4 Simulation Study

The data is generated by R program for estimating the parameter of smoothing spline and penalized spline methods based on NLAR model. At the beginning, we generate data $y_t, t = 1, 2, \ldots, n$ from AutoRegressive (AR) process at lag 1 by taking the coefficients $\rho = 0.1, 0.5, 0.7$, and 0.99 in the equation following;

$$y_t = \rho y_{t-1} + \varepsilon_t, t = 1, 2, \ldots, n, \tag{1.25}$$

where ρ is the AR coefficients, and ε_t is the error in term of normal distribution with $\mu = 0$, and $\sigma^2 = 1$. We simulate data with sample sizes $n = 50, 100, 200$, and 400, and repeat the generation at 500 times. To illustrate the application, Figure 1.1 shows the 100 sample sizes with four AR coefficients. It should be noted that

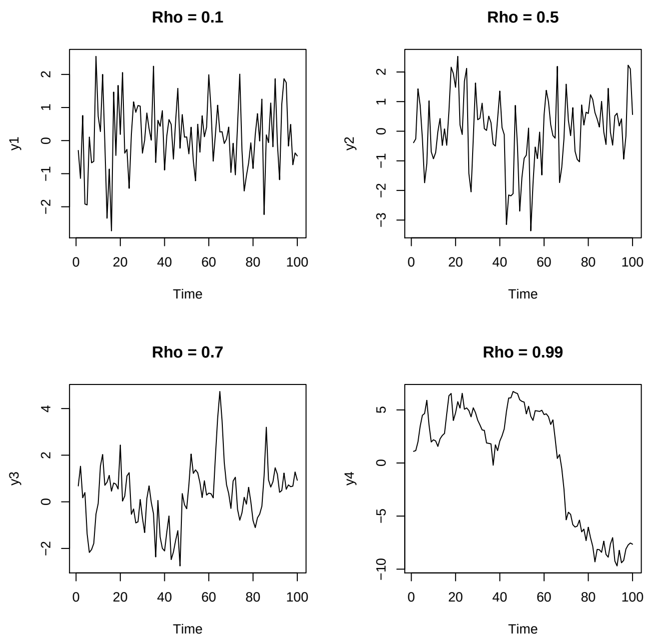

Figure 1.1: The time series plot of NLAR model with ρ = 0.1, 0.5, 0.7, and 0.99.

ρ = 0.1 and 0.5 are closed to ba a stationary process, and ρ = 0.7 and 0.99 is closed to ba a nonstationary process. The criterion to perform the efficiency of smoothing spline and penalized spline methods is Mean Square Errors (MSE) and given by

$$MSE = \frac{1}{n-1} \sum_{t=2}^{n} (y_t - \hat{y}_t)^2 \tag{1.26}$$

where y_t denoted the simulated values and $\hat{y}_t$ denoted the fitting values.

Table 1.1: The average MSE of Smoothing Spline (SS) method and Penalized Spline (PS) method based on NLAR model.

AR Coefficients	Sample Sizes	SS Method	PS Method
$\rho = 0.1$	n=50	3.0353	3.8311
	n=100	3.7652	3.8414
	n=200	3.8829	3.9291
	n=400	3.9451	3.9684
$\rho = 0.5$	n=50	2.9404	3.8084
	n=100	3.7804	3.8934
	n=200	3.8596	3.9130
	n=400	3.9684	3.9905
$\rho = 0.7$	n=50	2.9606	3.7417
	n=100	3.7402	3.8404
	n=200	3.8608	3.9163
	n=400	3.9544	3.9809
$\rho = 0.99$	n=50	2.7581	3.6084
	n=100	3.6566	3.7827
	n=200	3.8906	3.9378
	n=400	3.9254	3.9553

After fitting NLAR model, the average MSE's are shown at Table 1.1 for each AR coefficients and sample sizes with 500 replications. From Table 1.1, the average MSE of smoothing spline method is less than penalized spline method for all cases. However, when sample sizes (n) is large, the average MSE of smoothing spline method and penalized spline method is slightly difference.

1.5 Applications of Real Data

In this section, we consider the application of NLAR model using smoothing spline and penalized spline methods that we developed in the previous section. As financial data, we use the monthly volume of gold price (US Dollars per Troy Ounce) from January, 1984 to December 2013, which consisted of 360 records and shown on Figure 1.2. These data are obtained from http://www.indexmundi.com.

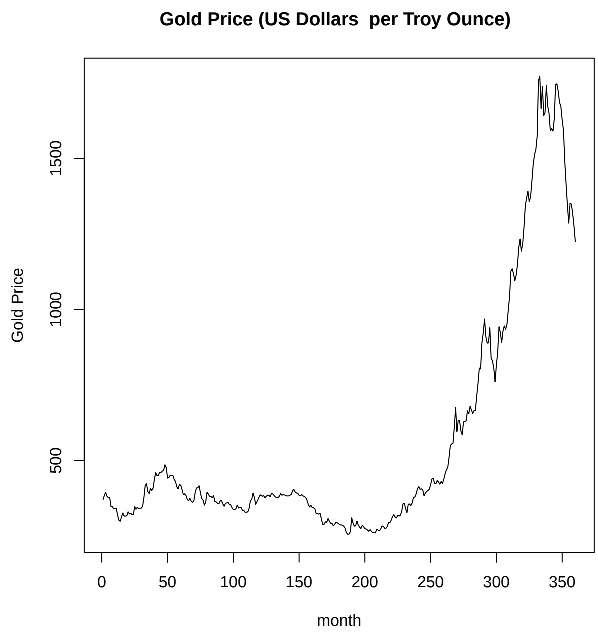

Figure 1.2: The time series plot for gold price consists of 360 records.

Let $y_t, t = 2, 3, \ldots, n-1$ denote the gold price to be a dependent variables and $y_{t-1}, t = 1, 2, \ldots, n-1$ define a independent variables depended on NLAR model. The smoothing spline and penalized spline methods are used to fit model, then it can be seen at Figure 1.3. From Figure 1.3, it's hard to decide the performance of

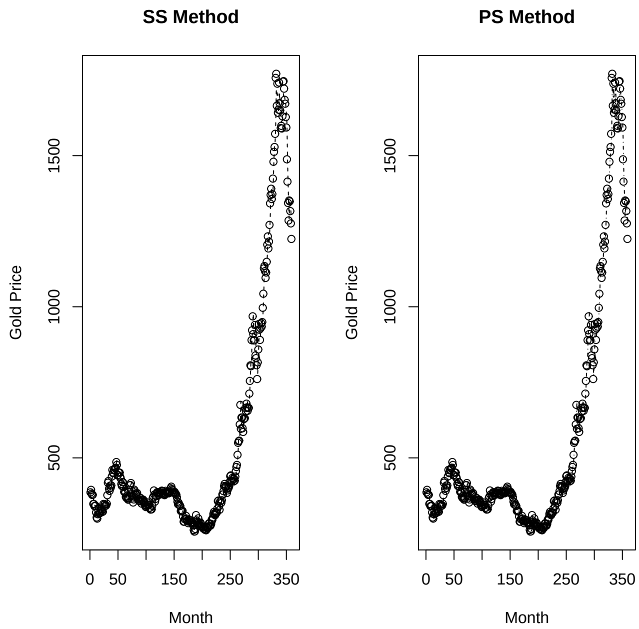

Figure 1.3: The true values of monthly in gold price, the fitting values of the Smoothing Spline (SS) method and Penalized Spline (PS) method.

smoothing spline and penalized spline methods. The MSE values show that the MSE of smoothing spline is 685.0235 and penalized spline is 733.9173. Therefore it can be concluded that the smoothing spline method outperforms penalized spline method.

1.6 Conclusion

We have focused on smoothing spline method and penalized spline method for fitting NLAR model. Depending on the simulated data, and real data, the smoothing spline method works more efficient than the penalized spline method. Furthermore the smoothing parameter could be sensitive to fit mean function because the smoothing parameter with GCV can be converged to interpolating spline in a class of smoothing spline estimator. This finding suggests that the user can choose smoothing spline method that provided a good approximation for fitting mean function. Future work will examine the estimating mean function ability of the different nonparametric regression methods, as well as issues, such as stationary, nonstationary, random work for fitting NLAR model.

Chapter 2

Conditional Heteroscedastic Nonlinear Autoregressive (CHNLAR) Model

2.1 Introduction

Currently the economics growth is a part for developing countries. These data are mostly stored in the form of time series data, whether it is a daily, a monthly, a quarterly, a yearly, like the unemployment rate, an economics growth rate, the gold price, the exchange rate. These information are sensitive and rapid fluctuations by the external factors such as natural disasters, wars, epidemics which these factors can not be controlled and predictable. For this reason, it is difficult to estimate or predict the precise business information.

Heteroscedastic or volatility happens quickly changes values in a class of economics time series data which the mean or the variance of the data is not fixed. Heteroscedastic or volatility model is useful to study for estimating or forecasting time series data where the right approach is to solve this problem.

There are several methods to model volatility in time series, such as the autoregressive conditional heteroscedastic model (ARCH) by Engle (1982), who was the first to introduce the ARCH model. The mean-corrected asset return is serially uncorrelated, but condition heteroscedastic asr changing over time. Bollerslev (1986) extended the Generalized ARCH (GARCH), and assumed that the mean equation can be adequatedly described by ARMA model. Gouriéroux and Monfort (1992) and Masry and Tjøstheim (1999a) have proposed the Conditional Heteroscedastic Nonlinear Autoregressive (CHNLAR) model in financial time series. For simplicity, the case is one lag of the CHNLAR model were studied to model the foreign exchange rates (Bossaerts et al., 1996).

The modeling of available explanatory variables has a variety of applications in regression model. The parametric and nonparametric method are the choice for estimating regression function between two sets of variables that consist of a vector of predictors and a response variable. A parametric regression model requires an assumption that the form of the underlying regression function. The selection of parametric model depends much on the problem and may be too restrictive in some applications. To overcome the difficulty caused by the restrictive assumption of the parametric form of the regression function, one may remove the restriction that the regression function belongs to a parametric family. This approach leads to so-called nonparametric regression.

Typically, the nonparametric regression methods are based on a smoothing technique which produces a smoother. A smoother is a tool for summarizing the trend of a response variable as a function of one or more predictor variables. The single predictor case is called scatterplot smoothing that can be used to enhance the visual appearance of the scatterplot of response versus predictor variable. There are many smoothing techniques, E.g., smoothing splines (Wahba, 1990, Green and Silverman, 1994), and penalized splines (Ruppert, Wand and Carroll, 2003). These smoothing techniques are generally based on the assumption of homoscedastic variance model which may not be suitable when the data involves high volatility.

For these various reasons, we interest to develop of the CHNLAR model using smoothing spline and penalized spline methods based on the simulated data, and real data.

2.2 CHNLAR model

The conditional heteroscedastic nonlinear autoregressive model is written as

$$y_t = \mu(y_{t-1}) + \sigma(y_{t-1})\varepsilon_t, \qquad t = 2, 3, \ldots, n, \tag{2.1}$$

where $y_t, t = 2, 3, \ldots, n$ are know the dependent variables, $y_{t-1}, t = 2, 3, \ldots, n$ are the past of independent variables at lag 1, $\mu(y_{t-1})$ are the mean function of conditional heteroscedastic autoregressive nonlinear model, $\sigma(y_{t-1})$ are the volatility function of conditional heteroscedastic autoregressive nonlinear model, and $\varepsilon_t,\ t = 2, 3, \ldots, n$ denote the error terms.

2.3 Nonparametric Regression Methods

2.3.1 Smoothing Spline (SS) Method

Wahba (1990) defined the natural polynomial spline $s(x) = s_K^m(x)$ is a real-valued function on $[a, b]$ with the aid of K so-called knots $-\infty \leq a < x_1 < x_2 < \ldots < x_K < b \leq \infty$. The class of m-order splines with domain $[a, b]$ will be denoted by $W^m[a, b]$.

The natural measure associated with the function $f \in W^m[a, b]$ that used to measure the roughness of curve which is called the quadratic penalty function given by

$$\int_b^a \{f^{(m)}(x)\}^2 dx \tag{2.2}$$

where $f^{(m)}(x)$ is the mth derivative of $f(x)$ with respect to x.

Consider the simple nonparametric regression model where the observation y_t at design points $x_t, t = 1, 2, \ldots, n$ assumed to satisfy

$$y_t = f(x_t) + \varepsilon_t, \qquad t = 1, 2, \ldots, n \tag{2.3}$$

where $f(\cdot)$ denotes a smooth function. To estimate $\hat{f}(\cdot)$ minimizes $s_K^{(m)}(f)$ over the class of function $f(\cdot)$ following

$$s_K^{(m)}(f) = \sum_{t=1}^{n} \{y_t - f(x_t)\}^2 + \lambda \int_a^b \left\{f^{(m)}(x)\right\}^2 dx \tag{2.4}$$

where $\lambda > 0$ denotes a smoothing parameter to be determined by a suitable cross-validation criteria or information criteria. The smoothing parameter controls the trade-off between fidelity to the data and roughness of function, if $\lambda \to 0$, the $\hat{f}_\lambda(\cdot)$ converges to linear function, if $\lambda \to \infty$, the $\hat{f}_\lambda(\cdot)$ converges to interpolating spline.

In this study, we emphasize $m = 2$ so-called the natural cubic spline which is commonly considered in the statistical literature (see Green and Silverman ,1994). We use this class of cubic smoothing splines to fit $\hat{f}(\cdot)$ by starting with the simple nonparametric regression model.

The first procedure of the smoothing spline method is considered a least square problem to fit a function $f(\cdot)$ that minimizes the residuals sum of squares

$$RSS = \sum_{t=1}^{n} \{y_t - f(x_t)\}^2 \tag{2.5}$$

Assuming the range of $f(\cdot)$ in (2.3) is finite interval, $[a, b] = [x_1, x_K]$, where x_1 denotes the ith order statistic and the roughness penalty of $f(\cdot)$ is measured by

$$\int_a^b \{f''(x_t)\}^2 \, dx \tag{2.6}$$

This leads of the following penalized least squares regression to find $\hat{f}_\lambda(x_t)$ called the smoothing spline estimator by minimizing

$$\hat{f}_\lambda(x_t) = \arg\min_f S_\lambda(f) \tag{2.7}$$

and

$$S_\lambda(f) = \sum_{t=1}^{n} \{y_t - f(x_t)\}^2 + \lambda \int_a^b \{f''(x)\}^2 \, dx \tag{2.8}$$

where $\lambda > 0$ is a smoothing parameter controlling the size of the roughness penalty and used to trade-off the goodness of fit.

In this paper, we also select the smoothing parameter using the method of generalized cross-validation (GCV) suggested by Wahba (1990) and Craven and Wahba (1979) . In practice, this step can be implemented by using the function of smooth.spline in the software R.

2.3.2 Penalized Spline (PS) Method

Eubank(1988, 1999) introduced the regression spline that the local neighborhoods are specified by a group of locations:

$$\tau_0, \tau_1, \tau_2, \ldots, \tau_K, \tau_{K+1} \tag{2.9}$$

in the range of interval $[a, b]$, where $a = \tau_0 < \tau_1 < \ldots < \tau_K < \tau_{K+1} < b$. These locations are known as knots, and $\tau_r, r = 1, 2, \ldots, K$ are called interior knots.

A regression spline can be constructed using the k-th degree truncated power basis or called the B-spline basis with K knots $\tau_1, \tau_2, \ldots, \tau_K$:

$$1, x, \ldots, x^k, (x - \tau_1)_+^k, \ldots, (x - \tau_K)_+^k, \tag{2.10}$$

where w_+^k denotes k-th power of the positive part of w where $w_+ = max(0, w)$. The first $(k+1)$ basis functions of the truncated power basis (2.10) are polynomials of degree up to k, and the others are all the truncated power functions of degree k. Conventionally, the truncated power basis of degree k, "$k = 0, 1, 2,$ and 3", is called "constant, linear, quadratic, and cubic" truncated power basis, respectively.

Using the truncated power basis in (2.10), a regression spline can be expressed as

$$f(x) = \sum_{s=0}^{k} \beta_s x^s + \sum_{r=1}^{K} \beta_{k+r}(x - \tau_r)_+^k \tag{2.11}$$

where $\beta_0, \beta_1, \ldots, \beta_{k+K}$ are the unknown coefficients to be estimated by a suitable loss minization.

The penalized spline is a method to estimate a unknown smooth function using the truncated power function (Ruppert and Carroll, 2000), and the penalized spline can be expressed as

$$f(x) = \sum_{j=0}^{m-1} \alpha_j x^j + \sum_{k=1}^{K} \beta_k (x - \tau_k)^{2m-1} \tag{2.12}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)^T \ N(0, \sigma_\beta^2 \boldsymbol{\Omega}^{-1/2}(\boldsymbol{\Omega}^{1/2})^T)$, and the (l, k)th entry of Ω is $|\tau_1 - \tau_k|^{2m-1}$ and only the coefficient of $|x_t - \tau_k|^{2m-1}$ are penalized so that a reasonably large order K can be used.

In this case, we focus m=2, as the natural cubic spline, or called low-rank thin-plate spline which tend to have very good numerical properties. The low-rank thin-plate spline representation of $f(.)$ is

$$f(x, \boldsymbol{\theta}) = \alpha_0 + \alpha_1 x + \sum_{k=1}^{K} \beta_k |x - \tau_k|^3 \tag{2.13}$$

where $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1, \ldots, \beta_K)^\top$ is the vector of regression coefficients, and $\tau_1 < \tau_2 < \ldots < \tau_K$ are fixed knots. The number of knots, K can be selected using a cross validation method or information theoretic methods (e.g., BIC or AIC).

This class of penalized spline smoothers ($\hat{f}(.)$) may also be expressed as

$$\hat{f} = \boldsymbol{C}(\boldsymbol{C}^\top \boldsymbol{C} + \lambda^3 \boldsymbol{D})^{-1} \boldsymbol{C}^\top \boldsymbol{y} \tag{2.14}$$

where

$$\begin{aligned} \boldsymbol{C} &= \begin{bmatrix} 1 & x_t & |x_t - \tau_k|^3_{1 \le k \le K} \end{bmatrix}_{1 \le t \le n}, \\ \boldsymbol{D} &= \begin{bmatrix} \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times K} \\ \mathbf{0}_{K\times 2} & (\boldsymbol{\Omega}_K^{1/2})^\top \boldsymbol{\Omega}_K^{1/2} \end{bmatrix}, \end{aligned}$$

and $\lambda = \sigma^2_\beta / \sigma^2_\varepsilon$ is a smoothing parameter. The penalized spline smoothers is estimated by using the SemiPar package in the software R.

2.4 Proposed Trend and Volatility Estimator

The trend $\mu(y_{t-1})$ and volatility $\sigma^2(y_{t-1})$ can also be considered in CHNLAR model. As an initial step, we start by estimating the trend $\mu(y_{t-1})$ using the concept of NLAR model written as

$$y_t = \mu(y_{t-1}) + \delta_t, \qquad t = 2, 3, \ldots \tag{2.15}$$

where $\delta_t = \sigma(y_{t-1})\, \varepsilon_t$. Next, we obtain $\hat{\mu}(y_{t-1})$ from SS and PS in R Program and the residuals can be estimated as

$$\hat{\delta}_t = y_t - \hat{\mu}(y_{t-1}) \tag{2.16}$$

$$(\hat{\delta}_t)^2 = (\sigma(y_{t-1})\, \varepsilon_t)^2 \tag{2.17}$$

We transform $\sigma(y_{t-1}) = \exp\left\{\frac{h(y_{t-1})}{2}\right\}$, and take log with residuals in (2.17)

$$\log\hat{\delta}_t^2 = h(y_{t-1}) + \log\varepsilon_t^2, \tag{2.18}$$

$$\log\hat{\delta}_t^2 - E[\log\varepsilon_t^2] = h(y_{t-1}) + \log\varepsilon_t^2 - E[\log\varepsilon_t^2] \tag{2.19}$$

If we require ε_t to be normally distributed with mean 0 and variance 1, then $E[\log\varepsilon_t^2] = -1.2704$ and hence we can apply in SS and PS to obtain

$$\log\hat{\delta}_t^2 + 1.2704 = h(y_{t-1}) + \log\varepsilon_t^2 + 1.2704 \tag{2.20}$$

$$\tilde{y}_t = h(y_{t-1}) + \tilde{\epsilon}_t \tag{2.21}$$

where $\tilde{y}_t = \log\hat{\delta}_t^2 + 1.2704$ and $\tilde{\epsilon}_t = \log\varepsilon_t^2 + 1.2704$. Next, we obtain a smooth estimate $\hat{h}(y_{t-1})$ using SS and PS by using (2.21) and update the volatility estimate to be

$$\hat{\sigma}(y_{t-1}) = \exp\left\{\frac{\hat{h}(y_{t-1})}{2}\right\} \tag{2.22}$$

At the second stage of estimation we update the trend estimate by using the following model

$$y_t = \mu(y_{t-1}) + \exp\left\{\frac{\hat{h}(y_{t-1})}{2}\right\}\varepsilon_t \tag{2.23}$$

$$\exp\left\{-\frac{\hat{h}(y_{t-1})}{2}\right\} y_t = \exp\left\{-\frac{\hat{h}(y_{t-1})}{2}\right\}\mu(y_{t-1}) + \varepsilon_t \tag{2.24}$$

$$\breve{y}_t = g(y_{t-1}) + \varepsilon_t \tag{2.25}$$

where $\breve{y}_t = \exp\left\{-\frac{\hat{h}(y_{t-1})}{2}\right\} y_t$ and $g(y_{t-1}) = \exp\left\{-\frac{\hat{h}(y_{t-1})}{2}\right\}\mu(y_{t-1})$. If $\hat{g}(y_{t-1})$ is the estimate obtained by using SS and PS, the second stage estimate of $\mu(y_{t-1})$ is given by

$$\hat{\mu}(y_{t-1}) = \exp\left\{\frac{\hat{h}(y_{t-1})}{2}\right\}\hat{g}(y_{t-1}) \tag{2.26}$$

Finally, when the estimates of $\mu(y_{t-1})$ and $\sigma(y_{t-1})$ converge we obtain

$$\hat{\varepsilon}_t = \frac{y_t - \hat{\mu}(y_{t-1})}{\hat{\sigma}(y_{t-1})}, \qquad t = 2, 3, \ldots \tag{2.27}$$

as the standardized residuals based on the converged values of $\hat{\mu}(.)$ and $\hat{\sigma}(.)$.

2.5 Simulation Study

In simulation study to estimate the performance of smoothing spline (SS) method and penalized spline (PS) are divided into two parts. The first part consists of the study in CHNLAR model

$$y_t = \mu(y_{t-1}) + \sigma(y_{t-1})\ \varepsilon_t, \qquad t = 2, 3, \ldots, n \tag{2.28}$$

where $\mu(y_{t-1})$ and $\sigma(y_{t-1})$ are generated following

$$\begin{aligned} \mu(y_{t-1}) &= 0.1(y_{t-1}) \\ \sigma(y_{t-1}) &= \exp\{0.1 \times y_{t-1}/2\} \end{aligned}$$

where $y_1 \sim \text{Normal}(0, 1)$.

In Figure 2.1, we present y_t in CHNLAR model. The error process $\varepsilon_t, t = 1, 2, \ldots, n$

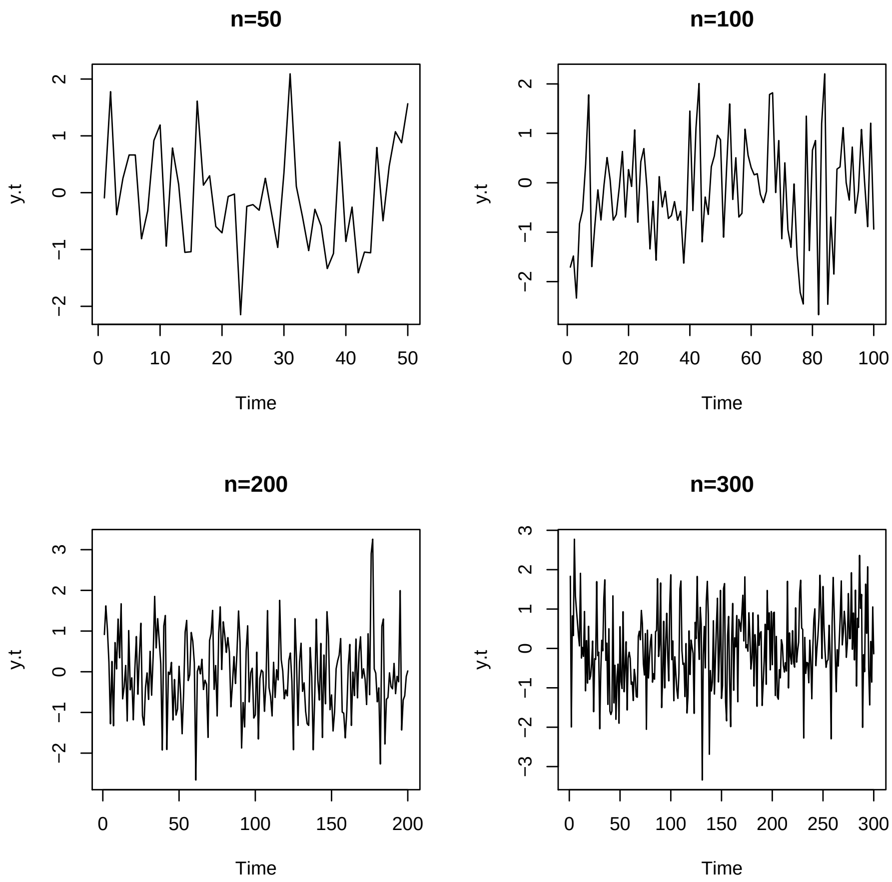

Figure 2.1: The time series data in CHNLAR Model

from in (2.28) which is assumed to follow the normal distribution with mean 0 and variance 1. We consider the sample sizes $n = 50, 100, 200$, and 300.

The second part, the estimates of $\hat{\mu}(y_{t-1})$ and $\hat{\sigma}(y_{t-1})$, we compute the bias of

$\mu(\cdot)$ and $\sigma(\cdot)$, the Mean Square Error (MSE) of $\mu(\cdot)$ and $\sigma(\cdot)$ by

$$
\begin{aligned}
\mu_{bias} &= \frac{1}{n}\sum_{t=2}^{n}\frac{\hat{\mu}(y_{t-1})-\mu(y_{t-1})}{\mu(y_{t-1})} \\
\sigma_{bias} &= \frac{1}{n}\sum_{t=2}^{n}\frac{\hat{\sigma}(y_{t-1})-\sigma(y_{t-1})}{\sigma(y_{t-1})} \\
MSE(\mu) &= \frac{1}{n}\sum_{t=2}^{n}(\hat{\mu}(y_{t-1})-\mu(y_{t-1}))^2 \\
MSE(\sigma) &= \frac{1}{n}\sum_{t=2}^{n}(\hat{\sigma}(y_{t-1})-\sigma(y_{t-1}))^2
\end{aligned}
$$

We simulated data with the sample sizes $n = 50, 100, 200$, and 300, and repeated the data generation and model fitting 500 times.

Tables 2.1 presents the average MSE of SS and PS methods for all sample sizes. For all sample sizes, the average MSE of $\mu(\cdot)$ and $\sigma(\cdot)$ are reduced with increasing sample sizes. For $\mu(\cdot)$, the average MSE of PS is less than SS, but the average MSE of SS are larger than the PS at $\sigma(\cdot)$.

Tables 2.2 and 2.3 show various Monte Carlo(MC) summary statistics of the estimates obtained by the SS and PS methods. The third and the fourth columns of these tables represent the MC sample mean and standard deviation of biases. The sample mean for the lower and upper bounds of 95% confidence interval are given in next two columns. The last two columns of these tables list the t-statistic, and p-values for hypothesis testing ($H_0:$ bias $= 0$ versus $H_1:$ bias $\neq 0$).

By observing the p-values, the results appear following:

From Tables 2.2 and 2.3, the SS and PS methods provides asymptotically unbiased estimates of $\sigma(\cdot)$ nearly for all two methods.

The histogram of the biases of all parameter estimates are presented in Figures 2.2-2.5. From the histogram is apparent that the distribution of $\sigma(\cdot)$ the biases appear to be more normally distributed for larger sample sizes.

Table 2.1: The average of MSE of SS and PS methods for different sample (500 replications)

	SS method		PS method	
sample size	$\mu(\cdot)$	$\sigma(\cdot)$	$\mu(\cdot)$	$\sigma(\cdot)$
n=50	0.300	0.0251	0.0249	0.1427
n=100	0.0105	0.0114	0.0095	0.0622
n=200	0.0054	0.0061	0.0033	0.0193
n=300	0.00372	0.00413	0.0020	0.0100

2.5.1 Smoothing Spline Method

Table 2.2: The simulation of smoothing spline method for different sample (500 replications)

bias	sample size	mean	s.d.	lci	uci	t-stat	p-value
μ_{bias}	n=50	6.507	157.106	-7.296	20.3116	0.9261	0.3548
	n=100	-1.172	23.367	-3.225	0.8810	-1.1216	0.2626
	n=200	-0.2977	10.2154	-1.1953	0.5998	-0.6517	0.5149
	n=300	3.5053	49.927	-0.8815	7.8921	1.5699	0.1171
σ_{bias}	n=50	-0.0114	0.1577	-0.025	0.0024	-1.6176	0.1064
	n=100	0.0020	0.1072	-0.0073	0.0114	0.4283	0.6686
	n=200	-0.0074	0.0776	-0.0142	-0.0005	-2.1335	0.0333*
	n=300	0.00184	0.0644	-0.0038	0.0075	0.6412	0.5217

*indicates significance at 5% level

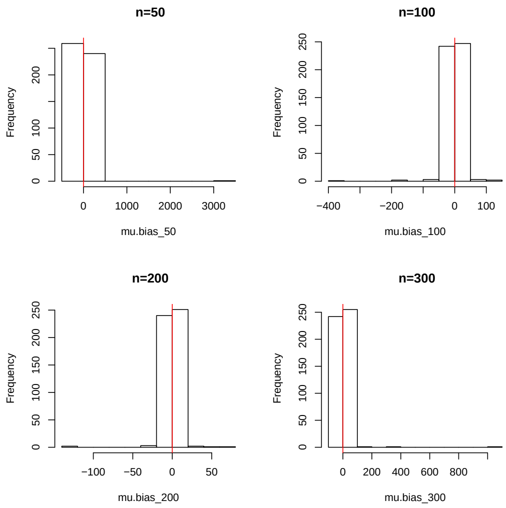

Figure 2.2: Histogram of Bias for $\mu(.)$ with Smoothing Spline Method

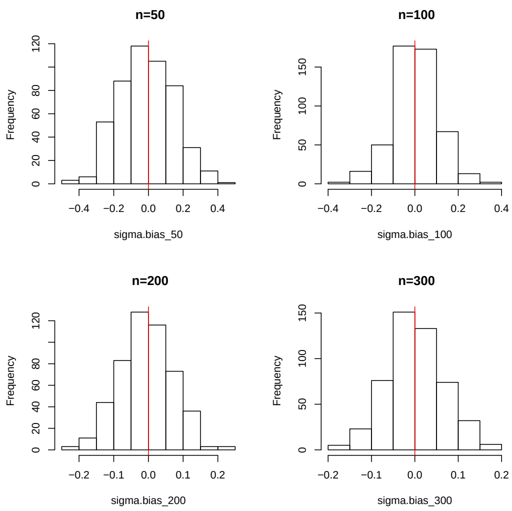

Figure 2.3: Histogram of Bias for $\sigma(.)$ with Smoothing Spline Method

2.5.2 Penalized Spline Method

Table 2.3: The simulation of penalized spline method for different sample (500 replications)

bias	sample size	mean	s.d.	lci	uci	t-stat	p-value
μ_{bias}	n=50	-4.734	79.833	-11.749	2.279	-1.326	0.1854
	n=100	1.485	47.561	-2.693	5.664	0.698	0.485
	n=200	0.0454	32.9606	-2.8506	2.9415	0.0308	0.9754
	n=300	-0.5388	10.4931	-1.4608	0.3831	-1.1483	0.2514
σ_{bias}	n=50	0.00429	0.172	-0.010	0.019	0.556	0.578
	n=100	-0.0001	0.1120	-0.009	0.009	-0.026	0.9792
	n=200	-0.0018	0.0790	-0.0087	0.0051	-0.5170	0.6054
	n=300	0.0035	0.0617	-0.0018	0.0090	1.3005	0.194

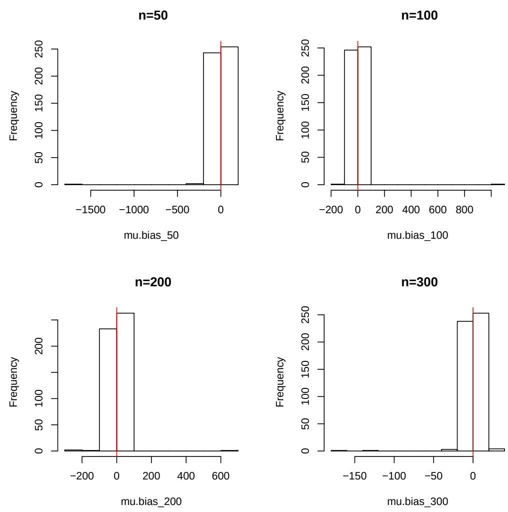

Figure 2.4: Histogram of Bias for $\mu(.)$ with Smoothing Spline Method

2.6 Applications for Real Data

In this section, we will consider the application of CHNLAR model using the smoothing spline and penalized spline methods that we developed in the previous chapter. The first data sets, we use the monthly volume of gold price (US Dollars per Troy Ounce) from January, 1984 to December 2013, which con-

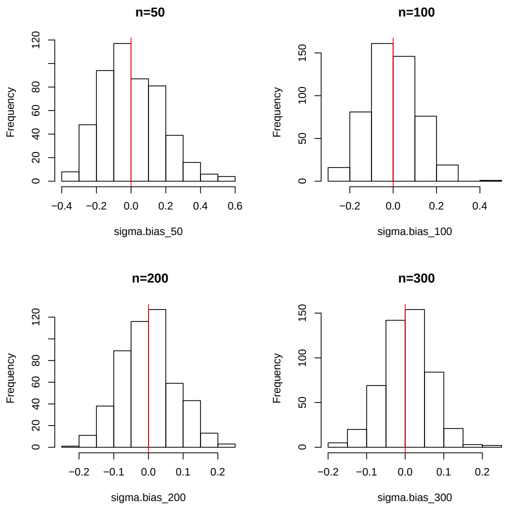

Figure 2.5: Histogram of Bias for $\sigma(.)$ with Smoothing Spline Method

sisted of 360 records and shown on Figure 2.6. These data are obtained from http://www.indexmundi.com. The process for predictive future values following:

At the first, we considered the CHNLAR model following

$$y_t = \mu(y_{t-1}) + \sigma(y_{t-1})\ \varepsilon_t \tag{2.29}$$

where ε_t's are independently and identically distributed with mean 0 and variance 1.

The second for parameter estimation, we fitted the CHNLAR model to obtain smooth estimate of the trend function, $\mu(\cdot)$ and the volatility function, $\sigma(\cdot)$. We obtained $\hat{\mu}(y_{t-1}), \hat{\sigma}(y_{t-1}), t = 2, 3, \ldots, n$ using smoothing spline and penalized spline methods.

Let $\hat{\mu}(y_{t-1})$ and $\hat{\sigma}(y_{t-1})$ denote the converged estimates of $\mu(\cdot)$ and $\sigma(\cdot)$ and let

$$\hat{\varepsilon}_t = \frac{y_t - \hat{\mu}(y_{t-1})}{\hat{\sigma}(y_{t-1})}, \qquad t = 2, 3, \ldots, n \tag{2.30}$$

denote the standardized residuals based on the converged values of $\hat{\mu}(y_{t-1})$ and $\hat{\sigma}(y_{t-1})$.

Finally, we obtain forecast values of $\hat{y}_2, \ldots, \hat{y}_n$ using the forecast trend and volatility based on the CHNLAR model:

$$\hat{y}_t = \hat{\mu}(y_{t-1}) + \hat{\sigma}(y_{t-1})\ \hat{\varepsilon}_t, \qquad t = 2, 3, \ldots, n \tag{2.31}$$

where the forecast trend $\hat{\mu}(y_{t-1})$ and the forecast volatility $\hat{\sigma}(y_{t-1})$ from smoothing spline and penalized spline methods.

The gold price is an important in our society and the economics world price. The data consisted of 360 records of the monthly volume of of gold price (US Dollars per Troy Ounce) from January, 1984 to December 2013.

Let y_t denote the gold price of month t where $t = 1$ represents January of 1984 and $t = 360$ represents December of 2013 shown in Figure 2.6. The method for

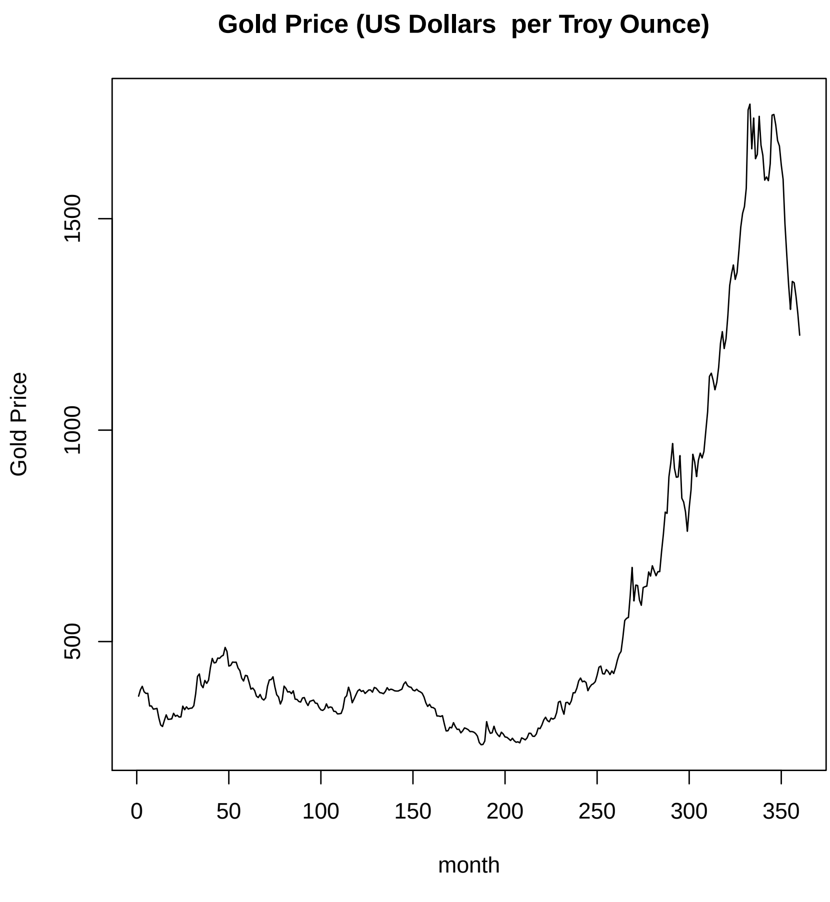

Figure 2.6: The time series plot of the monthly gold price values from January, 1984 to December, 2013

estimating values from Febuary, 1984 to December, 2013 of the smoothing spline and penalized spline methods are also computed the MSE, mean, and standard deviation of $\mu(\cdot)$ and $\sigma(\cdot)$ that shown in Table 2.4. From the Table 2.4, the MSE of 2 methods are equal values, but the $\mu(\cdot)$ and $\sigma(\cdot)$ is shown the different slightly values. In comparison, we also proposed the different of true values, forecast values of smoothing spline method in Figure 2.7 and penalized spine method in Figure 2.8.

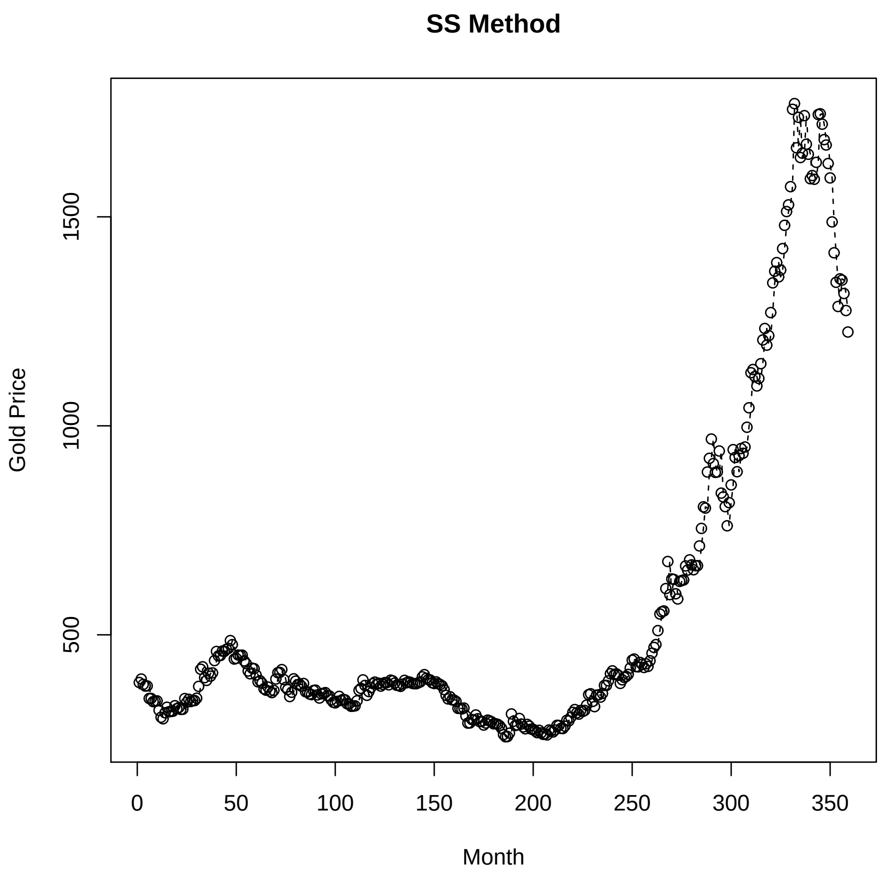

Figure 2.7: The actual gold price and estimated values for smoothing spline method

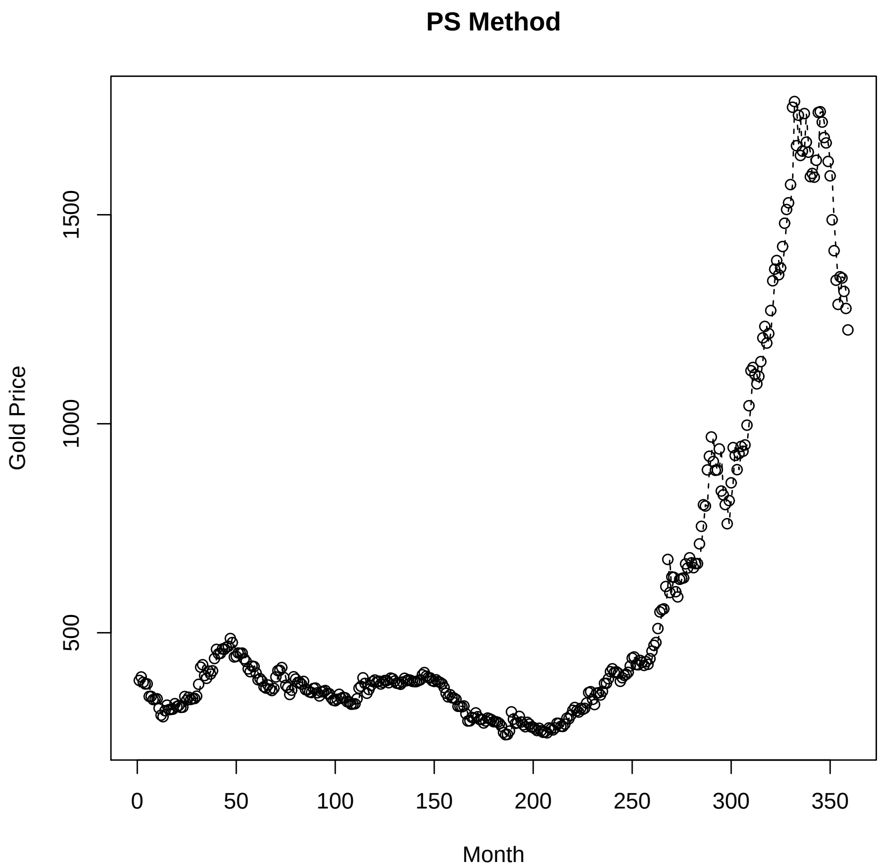

Figure 2.8: The actual gold price and estimated values for penalized spline

Table 2.4: The mean (standard deviation) and mean square error of smoothing spline (SS) and penalized spline (PS) method for different sample

Estimator	SS method	PS method
μ	564.4387 (396.7862)	564.4463 (396.755)
σ	16.92417 (15.54389)	20.4015 (19.66588)
MSE	5.653	5.653

2.7 Conclusion

In this section, we have developed the nonparametric regression method such as the smoothing spline method and the penalized spline method to estimate smooth unknown trend and smooth unknown volatility with CHNLAR model. Through a Monte Carlo simulation study, we evaluated the performance of smoothing spline method procedure and showed that the trend estimator ($\hat{\mu}(\cdot)$) and volatility estimator ($\hat{\sigma}(\cdot)$) work reasonably well for most data of all sample sizes except in one case when volatility estimator is statistical significant at $n = 200$. The point volatility estimators approach their corresponding true values as the sample sizes increase.

Taking a Monte Carlo study into consideration, we shown that the trend estimator of penalized spline works well for all small sample sizes, because the smoothing parameter is the high enough, indicating that the small sample sizes can obtain reliable conclusion from interpolating these models.

For application in actual data, we are also interested in comparing the power of estimating values by considering the Mean Square Error (MSE). The MSEs of a smoothing spline method and penalized spline method are equal values. However, we consider the mean of the trend and volatility estimator, we can see that the means of trend and volatility estimator of smoothing spline method performs slightly different in penalized spline method.

references

Allen, D. M. (1974). The relationship between variable and data augmentation and a method of prediction. *Technometrics*, **16**, 125 – 127.

Bollerslev, T (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **31**, 307 – 327.

Bossaerts, P., Härdle, W. and Hafner, C. (1996). Foreign exchange-rates have surprising volatility. In *Athens Conference on Applied Probability and Time Series*, Vol. 2 (ed. P.M. Robinson). Lecture Notes in Statistics 115, New York : Springer, 55 – 72.

Brumback, B., Ruppert, D., and Wand, M.P. (1999). Comment on variable selection and function estimation in additive nonparametric regression using data-based prior by shively. *Journal of American Statistics Association*, **94**, 794 – 797.

Chan, K.S. and Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, **7**, 179 – 190.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, **31**, 377 – 403.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987 – 1007.

Eubank, R.L. (1988). Spline Smoothing and Nonparametric Regression. Marcel Dekker, New York.

Eubank, R.L. (1999). Nonparametric Regression and Spline Smoothing. Marcel Dekker, New York.

Gouriéroux, C. and Monfort, A. (1992). Qualitative threshold ARCH models. *Journal of Econometrics*, **52**, 159 – 199.

Granger, C. W. J. and Teräsvirta, T. (1993). *Modelling Nonlinear Economic Relationships.* Oxford University Press, Oxford.

Green, P.J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach.* Chapman and Hall, London.

Haggan, V. and Ozaki, T. (1981). Modelling nonlinear vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*, **68**, 189 – 196.

Marsy, E. and Tjøstheim, D. (1995a). Non-parametric estimation and identification of ARCH nonlinear time series: strong convergence and asymptotic normality. *Econometric Theory*, **11**, 258 – 289.

Ruppert, D. and Carroll, R.J. (2000). Spatial-adaptive penalties for spline fitting. *Australian and New Zealand Journal of Statistics*, **42**, 205 – 224.

Ruppert, D., Wand, M.P., and Carroll. R.J. (2003). *Semiparametric Regression*, Cambridge University Press.

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

Tong, H. (1983). *Threshold models in nonlinear time series analysis*, Lecture Notes in Statistics 21, Heidelberg: Springer.

Wahba, G. (1977). A survey of some smoothing problems and the method of generalized cross-validation for solving them. *In Application of Statistics*, (P. R. Krisnaiah, ed.) North Holland, Amsterdam, 507 – 523.

Wahba, G. (1990). *Spline Models for Observational Data.* SIAM, Philadelphia, PA.