Applications and Applied Mathematics: An International Journal (AAM)

Volume 12 | Issue 1 | Article 6

6-2017

A Semiparametric Estimation for the Nonlinear Vector Autoregressive Time Series Model

Rahman Farnoosh
Iran University of Science and Technology

Mahtab Hajebi
Iran University of Science and Technology

Seyed J. Mortazavi
Islamic Azad University

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Applications and Applied Mathematics:
An International Journal
(AAM)

Vol. 12, Issue 1 (June 2017), pp. 78 - 97

A Semiparametric Estimation for the Nonlinear Vector Autoregressive Time Series Model

Rahman Farnoosh[*1], Mahtab Hajebi[2] and Seyed Javad Mortazavi[3]

[1]School of Mathematics
Iran University of Science and Technology
Tehran, Iran
rfarnoosh@iust.ac.ir;
[2]School of Mathematics
Iran University of Science and Technology
Tehran, Iran
mhajebi@iust.ac.ir;
[3]Department of Mathematics Sciences
Islamic Azad University
Chalous Branch, Iran
jmortazavi.amiri@gmail.com;

[*]Corresponding author

Submitted January 15, 2017; Accepted May 1, 2017

Abstract

In this paper, the nonlinear vector autoregressive model is considered and a semiparametric method is proposed to estimate the nonlinear vector regression function. We use Taylor series expansion up to the second order which has a parametric framework as a representation of the nonlinear vector regression function. After the parameters are estimated through the least squares method, the obtained nonlinear vector regression function is adjusted by a nonparametric diagonal matrix, and the proposed diagonal matrix is also estimated through the nonparametric smooth-kernel approach. Estimating the parameters can yield the desired estimate of the vector regression function based on the data. Under some conditions, the asymptotic consistency properties of the proposed semiparametric method are established. In this case, some simulated results for the semiparametric estimators in a nonlinear vector autoregressive function are presented. Mean Squares Error (MSE) criterion is also applied to verify the accuracy and the efficiency of the suggested model. The results of the study indicate the accuracy of the suggested model. Furthermore, the method is applied for the Retail Trade Survey to provide short-term economic indicators of the retail trade sector. The hypothesis of nonlinearity in the vector autoregression function model is also considered by the use of principal components. We use this test for the Retail Trade Survey (RTS) data which is considered to provide short-term economic indicators of the retail trade sector. Here, we use a nonlinear vector autoregression function model to forecast the sales of fresh, takeaway, supermarket, and café restaurant food in New Zealand during 2000 to 2010 on yearly basis. For our empirical example, the test of nonlinearity clearly indicates our model is nonlinear.

Keywords: Semiparametric Estimation; Least Squares Estimation; Vector Autoregressive Model; Taylor Series Expansion; Kernel approach; Nonparametric Adjustment; Simulation

MSC 2010 No.: 62G08, 62H12

1. Introduction

Time series analysis has been classified into univariate and multivariate time series analysis. Multivariate (vector) time series analysis was pioneered during the 1980s. Although in many situations we are only interested in predicting one variable in the future, we need to consider all of these variables as a vector time series. In vector time series analysis, a framework is needed for describing not only the properties of the individual series but also the possible cross-relationships among the series. Forecasting is one of the main objective of vector time series analysis. One of the most important models for prediction of vector time series is the vector generalization of the univariate autoregressive (AR) model which is called a vector AR (VAR) model. In order to make a more efficient model in various branches of sciences, particularly in applied statistics, econometric and financial studies, a combination of parametric forms and nonlinear functions has been used. In this case, the nonlinear VAR (NVAR) models are the most popular models for the nonlinear vector time series analysis.

There are, of course, many other nonlinear time series models shown to be useful in some applications. The exponential autoregressive model to apply in the modeling of sound vibration is proposed by Haggen and Ozaki (1981). Chen and Tsay (1993) considered a class of nonlinear additive autoregressive models with exogenous variable for nonlinear time series analysis. An additive nonlinear ARX time series model to consider the estimation and identification of components, both endogenous and exogenous is proposed by Cai and Masry (2000). Modern parametric and nonparametric methods for analysing nonlinear models are discussed by Fan and Yao (2003). Huang and Yang (2004) proposed a lag selection method for nonlinear additive autoregressive models that is based on spline estimation and the Bayes' information criterion. Yu et al. (2009) proposed a semiparametric method for an autoregressive model by combining a parametric regression estimator with a nonparametric adjustment. Farnoosh and Mortazavi (2011) proposed the first-order nonlinear autoregressive model with dependent errors to estimate the yearly amount of deposits in an Iranian Bank. A semiparametric method for the nonlinear autoregressive model is investigated by Mortazavi and Farnoosh (2013). Nademi and Farnoosh (2014) considered the mixtures of autoregressive-autoregressive conditionally heteroscedastic models with semiparametric approach. Farnoosh et al. (2014) also proposed a semiparametric method for estimating regression function in the partially linear autoregressive time series model. All these methods were proposed for the univariate time series.

Several authors also considered nonlinear vector time series models. For example, Hardle et al.

(1998) modelled the autoregression as a high dimensional nonparametric function estimated with a local linear estimator. Li and Genton (2009) proposed a new class of nonlinear autoregressive models for vector time series models. Shao and Yang (2011) considered the Yule-Walker estimator of the autoregressive coefficient based on the observed time series that contains an unknown trend function and an autoregressive error term. Generalised partial linear single-index mixed models for analysing repeated measures data are proposed by Chen et al. (2014). The extension of nonparametric additive quantile regression models to a partially linear additive quantile regression model is also described by Hoshino (2014).

The goal of this paper is to extend the work of Farnoosh et al. (2014) in the semiparametric estimation for the NVAR model with independent errors to estimate the retail trade sector in New Zealand. They suppose that $f(.)$ has a parametric framework as

$$f(x) \in \{g(x,\theta); \theta \in \mathrm{R}^m\},$$

and the regression function $f(.)$ is estimated by

$$\hat{f} = g(x,\hat{\theta}), \tag{1}$$

where $\hat{\theta}$ is an estimator of θ. In fact, the parametric regression estimator (1) is a crude guess of $f(x)$. When this initial parametric approximation is adjusted by a nonparametric multiplier $\xi(x)$, the semiparametric for $g(x,\hat{\theta})\xi(x)$ can be obtained. In their work, they could obtain the estimator $\hat{f}(x) = g(x,\hat{\theta})\hat{\xi}(x)$.

In this paper, a new class of semiparametric NVAR time series model is developed. Consider the following NVAR model

$$\mathrm{Y}_t = F(\mathrm{Y}_{t-1}) + \boldsymbol{\varepsilon}_t, \quad t \in Z, \tag{2}$$

where $\mathrm{Y}_t = [Y_{1t}, Y_{2t}, ..., Y_{mt}]^T$ is a vector time series, and $\boldsymbol{\varepsilon}_t = [\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{mt}]^T$ is a vector white noise with zero mean and covariance matrix $\sum$, and $F(.)$ is an unknown NVAR function and has continuous derivatives.

Our interest is to estimate the unknown vector function $F(\mathrm{Y}_{t-1})$ that can be formed as $G(\mathrm{Y}_{t-1}, \boldsymbol{\Theta})$, where $G(\mathrm{Y}_{t-1}, \boldsymbol{\Theta})$ is a vector function of the Taylor series of the function $F(\mathrm{Y}_{t-1})$ up to the second order about the point $A_0 = [\alpha_1, \alpha_2, ..., \alpha_m]^T$ and also, is a known function of Y_{t-1} and $\boldsymbol{\Theta}$.

Hence, we can ultimately obtain the following estimator

$$\hat{F}(Y_{t-1}) = \hat{\xi}(Y_{t-1})G(Y_{t-1}, \hat{\mathbf{\Theta}}),$$

where $\hat{\xi}(Y_{t-1})$ is a nonparametric adjustment diagonal matrix estimator of dimension $m \times m$.

Therefore, we use a combination of parametric method and nonparametric adjustment. The parameters and nonparametric adjustment are estimated by using least squares method and smooth-kernel method respectively.

The contents of the manuscript are organized as follows. A least squares estimation is considered to estimate the parameters matrix $\mathbf{\Theta}$ in Section 2. Also, in this section the semiparametric regression estimator is introduced by a natural consideration of the local L2-fitting criterion. The asymptotic behaviors of the estimator is investigated in Section 3. A simulation study to confirm the advantage of this method is presented in Section 4. In Section 5, this model is used to estimate the retail trade sector in New Zealand and the test of nonlinearity for the model is also considered. In Section 6, a test of nonlinearity in the vector autoregression function model is proposed.
Finally, some conclusions and summary are noted in Section 7.

2. Semiparametric method in the vector autoregressive model

We consider the following model

$$Y_t = F(Y_{t-1}) + \boldsymbol{\varepsilon}_t, \quad t \in Z, \tag{3}$$

Here, $F(.)$ is an unknown nonlinear vector autoregressive function and has continuous derivatives. Hence, we can write $F(.)$ as follows

$$F(Y_{t-1}) = \begin{bmatrix} f_1(Y_{t-1}) \\ f_2(Y_{t-1}) \\ \vdots \\ f_m(Y_{t-1}) \end{bmatrix}.$$

We want to estimate the unknown vector function $F(Y_{t-1})$ that can be formed as $G(Y_{t-1}, \mathbf{\Theta})$, where $G(Y_{t-1}, \mathbf{\Theta})$ is a vector function of the Taylor series of the function $F(Y_{t-1})$ up to the second order about the point $A_0 = [\alpha_1, \alpha_2, ..., \alpha_m]^T$ that we usually consider this point as an average of data.

Therefore, $G(Y_{t-1}, \mathbf{\Theta})$ is given by

$$G(\mathrm{Y}_{t-1},\mathbf{\Theta})=\begin{pmatrix} g_1(\mathrm{Y}_{t-1},\Theta_1) \\ g_2(\mathrm{Y}_{t-1},\Theta_2) \\ \vdots \\ g_m(\mathrm{Y}_{t-1},\Theta_m)\end{pmatrix} = \begin{pmatrix} B_{01}+B_{11}^T(\mathrm{Y}_{t-1}-A_0)+B_{21}^T vech\{(\mathrm{Y}_{t-1}-A_0)(\mathrm{Y}_{t-1}-A_0)^T\} \\ B_{02}+B_{12}^T(\mathrm{Y}_{t-1}-A_0)+B_{22}^T vech\{(\mathrm{Y}_{t-1}-A_0)(\mathrm{Y}_{t-1}-A_0)^T\} \\ \vdots \\ B_{0m}+B_{1m}^T(\mathrm{Y}_{t-1}-A_0)+B_{2m}^T vech\{(\mathrm{Y}_{t-1}-A_0)(\mathrm{Y}_{t-1}-A_0)^T\} \end{pmatrix},$$

where

$$\mathbf{\Theta}=(\Theta_1,\Theta_2,\ldots,\Theta_m)^T,$$

such that

$$\Theta_i\in\mathrm{R}^{n_i},\quad n_i\in\mathrm{N},\qquad i=1,2,\ldots,m,$$

and

$$B_{0i}=f_i(A_0),\qquad i=1,2,...,m,$$

$$B_{1i}=\{Df_i(A_0)\}=[\frac{\partial f_i(A_0)}{\partial\alpha_1},\frac{\partial f_i(A_0)}{\partial\alpha_2},...,\frac{\partial f_i(A_0)}{\partial\alpha_m}]^T,$$

$$B_{2i}=\frac{1}{2}[\frac{\partial^2 f_i(A_0)}{\partial\alpha_1^2},\frac{2\partial^2 f_i(A_0)}{\partial\alpha_1\partial\alpha_2},\ldots,\frac{2\partial^2 f_i(A_0)}{\partial\alpha_1\partial\alpha_m},\frac{\partial^2 f_i(A_0)}{\partial\alpha_2^2},\frac{2\partial^2 f_i(A_0)}{\partial\alpha_2\partial\alpha_3},\ldots,\frac{2\partial^2 f_i(A_0)}{\partial\alpha_2\partial\alpha_m},\ldots,\frac{\partial^2 f_i(A_0)}{\partial\alpha_m^2}]^T.$$

Note that the operator $vech(.)$ with dimension $(\frac{m(m+1)}{2}\times 1)$, is the half-vectorization operator of the lower triangular portion of a symmetric matrix with dimension $(m\times m)$.

Hence, $\mathbf{\Theta}$ is a matrix of unknown parameters with dimension $(m\times(\frac{(m+1)(m+2)}{2}))$, defined by

$$\mathbf{\Theta}=\begin{pmatrix} B_{01} & B_{11}^T & B_{21}^T \\ B_{02} & B_{12}^T & B_{22}^T \\ \vdots & \vdots & \vdots \\ B_{0m} & B_{1m}^T & B_{2m}^T \end{pmatrix}.$$

Here, $\Theta_i=(B_{0i},B_{1i}^T,B_{2i}^T),\quad i=1,2,\ldots,m$.

For the model (3), $\mathbf{\Theta}$ should be well estimated with the least squares method as follows.

$$\hat{\boldsymbol{\Theta}} = argmin_{\Theta} \sum_{t=1}^{n} [Y_t - G(\mathrm{Y}_{t-1}, \boldsymbol{\Theta})]^T [Y_t - G(\mathrm{Y}_{t-1}, \boldsymbol{\Theta})]. \tag{4}$$

By using the Taylor series up to the second order, (4) can be written as follows

$$\hat{\boldsymbol{\Theta}} \approx argmin_{\Theta} \sum_{t=1}^{n} \sum_{i=1}^{m} [Y_{it} - B_{0i} - B_{1i}^T (\mathrm{Y}_{t-1} - A_0) - B_{2i}^T vech\{(\mathrm{Y}_{t-1} - A_0)(\mathrm{Y}_{t-1} - A_0)^T\}]^2.$$

In order to simplify the notation and put all observations in one compact model, we define a vector time series model as follows

$$\mathbf{Y} := (\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n)_{m \times n}, \tag{5}$$

$$\boldsymbol{\Theta} := (\mu, \mathbf{A}, \mathbf{B})_{m \times k}, \tag{6}$$

where $k = (m+1)(m+2)/2$, $\mu = [B_{01}, B_{02}, \ldots, B_{0m}]_{m \times 1}^T$, $\mathbf{A} = [B_{11}^T, B_{12}^T, \ldots, B_{1m}^T]_{m \times m}^T$

and

$$\mathbf{B} = [B_{21}^T, B_{22}^T, \ldots, B_{2m}^T]^T_{m \times (\frac{m(m+1)}{2})}.$$

Also,

$$X_t = \begin{pmatrix} 1 \\ (\mathrm{Y}_t - A_0) \\ vech\{(\mathrm{Y}_t - A_0)(\mathrm{Y}_t - A_0)^T\} \end{pmatrix}_{k \times 1}, \tag{7}$$

$$\mathbf{X} := (X_0, X_1, \ldots, X_{n-1})_{k \times n}, \tag{8}$$

$$\mathbf{E} := (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_n)_{m \times n}, \tag{9}$$

$$y := vec(\mathbf{Y})_{mn \times 1}, \tag{10}$$

$$\beta := vec(\boldsymbol{\Theta})_{mk \times 1}, \tag{11}$$

$$\omega := vec(\mathbf{E})_{mn \times 1}. \tag{12}$$

Notice, $vec(.)$ is the vectorization operator, which orders a matrix into a column-stack vector. Specifically, the vectorization of an $(m \times n)$ matrix A, denoted by $vec(A)$, is the $(mn \times 1)$ column vector obtained by stacking the columns of the matrix A on top of one another.

By using the notations of Equations (5) - (12), the VAR model in Equation (3) for $t = 1,2,\ldots,n$, can be written as the following compact model

$$\mathbf{Y} = \mathbf{\Theta X} + \mathbf{E}. \tag{13}$$

By taking $vec(.)$ on both sides of Equation (13), and using the $vec(.)$ operation rule, we have

$$\begin{aligned} vec(\mathbf{Y}) &= vec(\mathbf{\Theta X}) + vec(\mathbf{E}) \\ &= (\mathbf{X}^T \otimes \mathbf{I}_m) vec(\mathbf{\Theta}) + vec(\mathbf{E}). \end{aligned} \tag{14}$$

Therefore, considering the notations in Equations (10) - (12), Equation (14) can be written as

$$y = (\mathbf{X}^T \otimes \mathbf{I}_m)\beta + \omega. \tag{15}$$

This model will be used to find the least squares estimator of the parameters matrix $\mathbf{\Theta}$ by obtaining the least squares estimator for the parameter vector β.

Let $S(\beta)$ be the function of the sum of squares of the errors of the model (15) as follows

$$\begin{aligned} S(\beta) &= \omega^T \omega \\ &= (y - (\mathbf{X}^T \otimes \mathbf{I}_m)\beta)^T (y - (\mathbf{X}^T \otimes \mathbf{I}_m)\beta) \\ &= y^T y - y^T (\mathbf{X}^T \otimes \mathbf{I}_m)\beta - \beta^T (\mathbf{X} \otimes \mathbf{I}_m) y + \beta^T (\mathbf{X}\mathbf{X}^T \otimes \mathbf{I}_m)\beta. \end{aligned} \tag{16}$$

By taking the derivative on both sides of Equation (16) relative to the vector parameter β and using the matrix derivative rules, we have

$$\frac{\partial S(\beta)}{\partial \beta} = -2(\mathbf{X} \otimes \mathbf{I}_m) y + 2(\mathbf{X}\mathbf{X}^T \otimes \mathbf{I}_m)\beta. \tag{17}$$

Then, by considering Equation (17) equal to zero, the subsequent normal equation can be obtained

$$(\mathbf{X}\mathbf{X}^T \otimes \mathbf{I}_m)\hat{\beta} = (\mathbf{X} \otimes \mathbf{I}_m) y. \tag{18}$$

Let us not forget from Equation (10) that $y := vec(\mathbf{Y})$ and Equation (11) that $\beta := vec(\mathbf{\Theta})$. Equation (18) can be shown

$$vec(\hat{\mathbf{\Theta}}) = \hat{\beta} = [(\mathbf{XX}^T \otimes \mathbf{I}_m)]^{-1}(\mathbf{X} \otimes \mathbf{I}_m)y$$
$$= (\mathbf{XX}^T)^{-1} \otimes \mathbf{I}_m)(\mathbf{X} \otimes \mathbf{I}_m)y$$
$$= vec(\mathbf{YX}^T(\mathbf{XX}^T)^{-1}). \tag{19}$$

Hence, the least squares estimator of $\mathbf{\Theta}$ is equal to

$$\hat{\mathbf{\Theta}} = \mathbf{YX}^T(\mathbf{XX}^T)^{-1}. \tag{20}$$

The Hessian matrix of $S(\beta)$ is the second derivative of the function $S(\beta)$, and is given by

$$\frac{\partial^2 S(\beta)}{\partial \beta \partial \beta^T} = 2(\mathbf{XX}^T \otimes \mathbf{I}_m). \tag{21}$$

This is clearly a positive definite matrix, and confirms that $\hat{\beta}$ is the minimizer of the sum of squares of the errors function $S(\beta)$ defined in Equation (16).

Now, we estimate $\xi(\mathrm{Y}_{t-1})$ in $F(\mathrm{Y}_{t-1}) = \xi(\mathrm{Y}_{t-1})G(\mathrm{Y}_{t-1}, \mathbf{\Theta})$ by using a similar idea of Hjort and Jones (1996), Naito (2004), Yu et al. (2009), Farnoosh and Mortazaviand (2011) and Farnoosh et al. (2014) , we define the local L2-fitting criterion as follows

$$q(X,\xi) = \sum_{t=1}^{n}[F(\mathrm{Y}_{t-1}) - \xi(\mathrm{Y}_{t-1})G(\mathrm{Y}_{t-1}, \hat{\mathbf{\Theta}})]^T[F(\mathrm{Y}_{t-1}) - \xi(\mathrm{Y}_{t-1})G(\mathrm{Y}_{t-1}, \hat{\mathbf{\Theta}})]K_H(\mathrm{Y}_{t-1} - X)$$
$$= \sum_{i=1}^{m}\sum_{t=1}^{n}[f_i(\mathrm{Y}_{t-1}) - g_i(\mathrm{Y}_{t-1}, \hat{\Theta}_i)\xi_i(\mathrm{Y}_{t-1})]^2 K_H(\mathrm{Y}_{t-1} - X), \tag{22}$$

where $F(.)$ is an unknown vector autoregression function. Here, $K_H(\mathrm{Y}_{t-1} - X)$ is a gaussian kernel function as follows

$$K_H(\mathrm{Y}_{t-1} - X) = \frac{1}{(2\pi)^{m/2}\sqrt{\det(H^T H)}} exp\{-1/2(\mathrm{Y}_{t-1} - X)^T (H^T H)^{-1}(\mathrm{Y}_{t-1} - X)\}, \tag{23}$$

and H is the bandwidth diagonal matrix of dimension $m \times m$ which depends on the sample size (n).

We can obtain the estimator $\hat{\xi}(\mathrm{Y}_{t-1})$ of $\xi(\mathrm{Y}_{t-1})$ by minimizing the criterion in Equation (22) with respect to $\xi_i(\mathrm{Y}_{t-1})$ for $i = 1,2,\ldots,m$. Therefore, we can obtain nonparametric adjustment diagonal matrix estimator of $\xi(\mathrm{Y}_{t-1})$ as

$$\hat{\xi}_i(\mathrm{Y}_{t-1}) = \frac{\sum_{t=1}^{n} K_H(\mathrm{Y}_{t-1} - X) f_i(\mathrm{Y}_{t-1}) g_i(\mathrm{Y}_{t-1}, \hat{\Theta}_i)}{\sum_{t=1}^{n} K_H(\mathrm{Y}_{t-1} - X) g_i^2(\mathrm{Y}_{t-1}, \hat{\Theta}_i)}, \qquad i = 1, 2, \ldots, m. \tag{24}$$

Then, the estimator of $F(\mathrm{Y}_{t-1})$ could be

$$\hat{F}(\mathrm{Y}_{t-1}) = \hat{\xi}(\mathrm{Y}_{t-1}) G(\mathrm{Y}_{t-1}, \hat{\boldsymbol{\Theta}}). \tag{25}$$

Unfortunately, the formula in Equation (24), contains the unknown function $f_i(\mathrm{Y}_{t-1})$, therefore by using

$$\varepsilon_{it} = Y_{it} - f_i(\mathrm{Y}_{t-1}), \qquad i = 1, 2, \ldots, m,$$

and with regard to the fact the errors of model are small values, we have

$$Y_{it} \approx f_i(\mathrm{Y}_{t-1}), \qquad i = 1, 2, \ldots, m.$$

Thus, we get nonparametric adjustment diagonal matrix estimator of $\xi(\mathrm{Y}_{t-1})$ as

$$\tilde{\xi}_i(\mathrm{Y}_{t-1}) = \frac{\sum_{t=1}^{n} K_H(\mathrm{Y}_{t-1} - X) Y_{it} g_i(\mathrm{Y}_{t-1}, \hat{\Theta}_i)}{\sum_{t=1}^{n} K_H(\mathrm{Y}_{t-1} - X) g_i^2(\mathrm{Y}_{t-1}, \hat{\Theta}_i)}, \qquad i = 1, 2, \ldots, m. \tag{26}$$

Finally, the vector autoregression function estimator is given by

$$\tilde{F}(\mathrm{Y}_{t-1}) = \tilde{\xi}(\mathrm{Y}_{t-1}) G(\mathrm{Y}_{t-1}, \hat{\boldsymbol{\Theta}}). \tag{27}$$

We concentrated on $\tilde{F}(\mathrm{Y}_{t-1})$ to estimate the unknown nonlinear vector autoregression function. It can be calculated with the sample and simulation data.

3. The consistency properties of semiparametric estimators

In this section, some consistency properties of semiparametric estimators in a nonlinear vector autoregressive function, with independent errors, are investigated.

For the sake of consistency in the vector autoregressive function estimator, we accept the assumptions (A1-A11) of Farnoosh and Mortazavi (2011). If these conditions hold, we have Theorems 3.1 and 3.2; to prove the Theorems, we need the following Lemma.

Lemma 3.1.

Under the conditions of (A1-A11), we have as $n \to \infty$,

$$n^{-4/5} \sum_{t=1}^{n} K_H(\mathrm{Y}_{t-1} - X) f_i(\mathrm{Y}_{t-1}) g_i(\mathrm{Y}_{t-1}, \hat{\Theta}_i) \xrightarrow{p} C\mu_i(X) f_i(X) g_{0i}(X)$$
$$n^{-4/5} \sum_{t=1}^{n} K_H(\mathrm{Y}_{t-1} - X) g_i^2(\mathrm{Y}_{t-1}, \hat{\Theta}_i) \xrightarrow{p} C\mu_i(X) g_{0i}^2(X)$$

where C is a constant value and $g_{0i}(X) = g_i(X, \Theta_0)$ for $i = 1,2,...,m$.

To prove Lemma 3.1, we can refer to Yu et al. (2009).

Theorem 3.1.

Let $\hat{F}(\mathrm{Y}_{t-1})$ be the introduced estimator in Eq (25). Then $\hat{F}(\mathrm{Y}_{t-1}) \xrightarrow{p} F(\mathrm{Y}_{t-1})$, as $n \to \infty$.

Proof:

With the application of strong consistency of $\hat{\Theta}_n$ (Lai et al. (1978); Mingzhong and Xiru (1999)) and Lemma 3.1, we can prove Theorem 3.1.

Theorem 3.2.

Let $\tilde{F}(\mathrm{Y}_{t-1})$ be the defined autoregression function estimator in Equation (27). Then,

$$\| \hat{F}(\mathrm{Y}_{t-1}) - \tilde{F}(\mathrm{Y}_{t-1}) \| \xrightarrow{p} 0, \text{ as } n \to \infty.$$

Proof:

We have

$$\tilde{f}_i(\mathrm{Y}_{t-1}) - \hat{f}_i(\mathrm{Y}_{t-1}) = \frac{g_i(\mathrm{Y}_{t-1}, \hat{\Theta}_n) \sum_{t=1}^{n} K_H(\mathrm{Y}_{t-1} - X) \hat{\varepsilon}_{it} g_i(\mathrm{Y}_{t-1}, \hat{\Theta}_n)}{\sum_{t=1}^{n} K_H(\mathrm{Y}_{t-1} - X) g_i^2(\mathrm{Y}_{t-1}, \hat{\Theta}_n)}, \qquad i = 1, 2, ..., m.$$

Also,

$$n^{-4/5}\sum_{t=1}^{n}K_H(\mathrm{Y}_{t-1}-X)\varepsilon_{it}g_i(\mathrm{Y}_{t-1},\hat{\Theta}_n)=n^{-4/5}\sum_{t=1}^{n}K_H(\mathrm{Y}_{t-1}-X)\varepsilon_{it}(g_i(\mathrm{Y}_{t-1},\hat{\Theta}_n)-g_i(\mathrm{Y}_{t-1},\Theta_0))$$
$$+n^{-4/5}\sum_{t=1}^{n}K_H(\mathrm{Y}_{t-1}-X)\varepsilon_{it}g_i(\mathrm{Y}_{t-1},\Theta_0):=A_n+B_n.$$

It is known that $max_{1\le i\le n}|\varepsilon_{it}|=O((\log n)^{1/2})$ a.s. as $n\to\infty$ (Yu et al. (2009)).

It follows that

$$\begin{aligned}|A_n|&\le n^{-4/5}\sum_{t=1}^{n}K_0|\varepsilon_{it}|\left|g_i(\mathrm{Y}_{t-1},\hat{\Theta}_n)-g_i(\mathrm{Y}_{t-1},\Theta_0)\right|\\&\le n^{-4/5}nO((\log n)^{1/2})O((\log_2 n/n)^{1/2})\\&=O((\log n\log_2 n)^{1/2}/n^{3/10}),\end{aligned}$$

where K_0 is an upper bound of kernel. Therefore, $A_n\to 0$ a.s. as $n\to\infty$.

Since

$$E(B_n)=n^{-4/5}E\left\{\sum_{t=1}^{n}K_H(\mathrm{Y}_{t-1}-X)\varepsilon_{it}g_i(\mathrm{Y}_{t-1},\Theta_0)\right\}=0,$$

and

$$E\left\{n^{-4/5}\sum_{t=1}^{n}K_H(\mathrm{Y}_{t-1}-X)\varepsilon_{it}g_i(\mathrm{Y}_{t-1},\Theta_0)\right\}^2=n^{-8/5}E\left\{\sum_{t=1}^{n}K_H(\mathrm{Y}_{t-1}-X)\varepsilon_{it}g_i(\mathrm{Y}_{t-1},\Theta_0)\right\}^2$$

$$=n^{-8/5}\Big\{\sum_{t=1}^{n}E\left(K_H^2(\mathrm{Y}_{t-1}-X)\varepsilon_{it}^2g_i^2(\mathrm{Y}_{t-1},\Theta_0)\right)$$

$$+2\sum_{1\le t<t'\le n}E\big[K_H(\mathrm{Y}_{t-1}-X)\varepsilon_{it}g_i(\mathrm{Y}_{t-1},\Theta_0)$$

$$K_H(\mathrm{Y}_{t'-1}-X)\varepsilon_{it'}g_i(\mathrm{Y}_{t'-1},\Theta_0)\big]\Big\}$$

$$\le n^{-8/5}nK^*\sigma^2=O(1/n^{3/5}),$$

where $K^*>0$ is a constant, $B_n\xrightarrow{p}0$ as $n\to\infty$. Hence, we have

$$\left|\tilde{f}_i(\mathrm{Y}_{t-1}) - \hat{f}_i(\mathrm{Y}_{t-1})\right| \xrightarrow{p} 0, \qquad i = 1, 2, \ldots, m.$$

Since

$$\left\|\tilde{F}(\mathrm{Y}_{t-1}) - \hat{F}(\mathrm{Y}_{t-1})\right\| \leq \left|\tilde{f}_1(\mathrm{Y}_{t-1}) - \hat{f}_1(\mathrm{Y}_{t-1})\right| + \left|\tilde{f}_2(\mathrm{Y}_{t-1}) - \hat{f}_2(\mathrm{Y}_{t-1})\right| + \ldots + \left|\tilde{f}_m(\mathrm{Y}_{t-1}) - \hat{f}_m(\mathrm{Y}_{t-1})\right|,$$

we get

$$\left\|\tilde{F}(\mathrm{Y}_{t-1}) - \hat{F}(\mathrm{Y}_{t-1})\right\| \to 0.$$

Therefore, by the strong consistency of $\hat{\boldsymbol{\Theta}}_n$ (Lai et al. (1978); Mingzhong and Xiru (1999)) and Lemma 3.1, Theorem 3.2 is proved.

4. Simulation study

In this section, we conduct a simulation study to evaluate the performance of the proposed semiparametric estimation method for the nonlinear vector time series data. For convenience, let us consider vector autoregressive function processes with dimension m =2. The two-dimensional vector time series $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n$ for different choices of n are generated from the following NVAR model with independent errors

$$\mathrm{Y}_t = F(\mathrm{Y}_{t-1}) + \boldsymbol{\varepsilon}_t, \quad t \in Z, \tag{28}$$

where the elements of the nonlinear vector function $F(\mathrm{Y}_{t-1}) = \left[f_1(\mathrm{Y}_{t-1}), f_2(\mathrm{Y}_{t-1})\right]^T$ are given as

$$f_1(\mathrm{Y}_{t-1}) = 1 + 0.3e^{-Y_{1(t-1)}Y_{2(t-1)}}, \qquad \mathrm{Y}_{t-1} = [Y_{1(t-1)}, Y_{2(t-1)}]^T, \tag{29}$$

$$f_2(\mathrm{Y}_{t-1}) = 1 \quad 0Y2_{2(t-1)}\, e^{-Y_{1(t-1)}^2} \tag{30}$$

For the model (28), we assumed that $\{\boldsymbol{\varepsilon}_t\}$ are independent, Gaussian, two-dimensional random vectors with zero mean and $Var(\boldsymbol{\varepsilon}_t) = (0.125)^2 I$.

Monte Carlo simulations are carried out for assessing the performance of the proposed method. For each simulation, we consider sample sizes of n = 100, 300, 500, 700, 1000, 1500 and 2000 with 500 replications. For each simulated data set, the matrix of the parameters, $\boldsymbol{\Theta}$, is estimated by using its least square estimator given in Equation (20).

For example, with the sample size n = 2000, the estimated values of the parameters for the model (28) is as follows

$$\hat{\boldsymbol{\Theta}} = \begin{pmatrix} 1.0933 & -0.0991 & -0.1027 & 0.0553 & 0.0171 & 0.0509 \\ 1.0638 & -0.1377 & 0.0609 & 0.0942 & -0.1123 & 0.0118 \\ & & & & & \end{pmatrix}. \tag{31}$$

Here, the true parameters of the model is considered as

$$\boldsymbol{\Theta}_0 = \begin{pmatrix} 1.0933 & -0.0995 & -0.1022 & 0.0530 & 0.0157 & 0.0560 \\ 1.0642 & -0.1407 & 0.0602 & 0.0899 & -0.1320 & 0 \\ & & & & & \end{pmatrix}. \tag{32}$$

It can be seen that

$$\| \hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0 \| \to 0, \quad \text{as} \quad n \to \infty. \tag{33}$$

After estimating the parameters, the semiparametric estimator function of $F(\mathrm{Y}_{t-1})$ is obtained by

$$\begin{aligned} \tilde{F}(\mathrm{Y}_{t-1}) &= \begin{pmatrix} \tilde{f}_1(\mathrm{Y}_{t-1}) \\ \tilde{f}_2(\mathrm{Y}_{t-1}) \\ \end{pmatrix} \\ &= \begin{pmatrix} \tilde{\xi}_1(\mathrm{Y}_{t-1}) & 0 \\ 0 & \tilde{\xi}_2(\mathrm{Y}_{t-1}) \end{pmatrix} \begin{pmatrix} g_1(\mathrm{Y}_{t-1}, \hat{\Theta}_1) \\ g_2(\mathrm{Y}_{t-1}, \hat{\Theta}_2) \\ \end{pmatrix}, \end{aligned} \tag{34}$$

where $\tilde{\xi}_i(\mathrm{Y}_{t-1}), \ \ i = 1, 2$, can be obtained from Equation (26), and the bandwidth diagonal matrix of dimension (2×2) in the kernel function is best to be chosen proportional to $n^{(-0.2)} I_2$ (see Li et al. (2015)). Finally, we compute the mean square error (MSE) for the efficiency of the proposed estimation method by using

$$\begin{aligned} MSE(\tilde{F}(\mathrm{Y}_{t-1})) &= tr[E[(\tilde{F}(\mathrm{Y}_{t-1}) - F(\mathrm{Y}_{t-1}))(\tilde{F}(\mathrm{Y}_{t-1}) - F(\mathrm{Y}_{t-1}))^T]] \\ &\approx tr[\frac{1}{n} \sum_{t=1}^{n} [(\tilde{F}(\mathrm{Y}_t) - F(\mathrm{Y}_t))(\tilde{F}(\mathrm{Y}_t) - F(\mathrm{Y}_t))^T]]. \end{aligned} \tag{35}$$

Note that, $tr(A) = \sum_{i=1}^{n} a_{ii}$, if $A = [a_{ij}]_{n \times n}$.

For different *n*, MSE is approximately obtained in Table 1. Also, Figure 1 shows the norm of the difference between the estimated vector autoregressive function and the true vector autoregressive function, i.e., $\left\|\tilde{F}(\mathrm{Y}_{t-1})-F(\mathrm{Y}_{t-1})\right\|$. As it can be seen from Figure 1, the norm of the difference converges to zero which indicates our theoretical results in Theorem 3.1 and Theorem 3.2. Thus, the simulation results show that the semiparametric estimator of the vector autoregressive function performs well.

Table 1. *MSE* for the proposed estimation method with $H = n^{-0.2} \times I_2$.

n	100	300	500	700	1000	1500	2000
MSE	3×10^{-2}	9.04×10^{-4}	5.4×10^{-4}	2.74×10^{-4}	1.14×10^{-4}	1.11×10^{-4}	9.19×10^{-5}
$H(\times I_2)$	0.3981	0.3195	0.2885	0.2697	0.2511	0.2316	0.2186

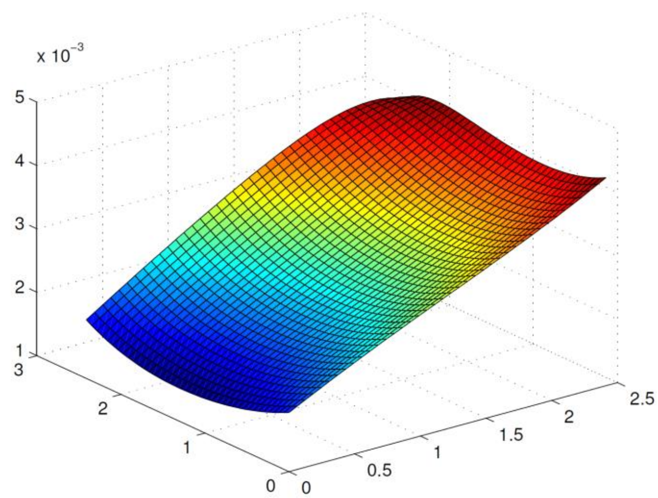

Figure 1. $\|\tilde{F}(\mathrm{Y}_{t-1})-F(\mathrm{Y}_{t-1})\| \to 0,\ MSE = 9.19\times10^{-5}$

5. Empirical application

In this research, the Retail Trade Survey (RTS) is considered to provide short-term economic indicators of the retail trade sector. In order to illustrate the suitability of our methodology to economical data, in preparing family cost and expenses (per person), we used a NVAR model to forecast the sales of fresh, takeaway, supermarket, and café restaurant food in New Zealand during 2000 to 2010 on yearly basis. We consider $\mathrm{Y}_t^* = \dfrac{\mathrm{Y}_t - \mathrm{Y}_{t-1}}{\mathrm{Y}_{t-1}}$, where Y_t is a vector with dimension *m*=4 at time *t*. Using the presented semiparametric method, we estimated nonlinear vector regression function for the yearly sales of fresh, takeaways, supermarkets, and café restaurants food variables at time *t*. We computed the mean squares error (MSE) for the efficiency of the proposed estimation method. We see that the presented semiparametric method

for a NVAR model with independent errors is proved to be efficient.
The estimated values of the parameters for the NVAR model is as follows

$$\hat{\boldsymbol{\Theta}} = \begin{pmatrix} 0.08 & 0.2 & -0.22 & 0.38 & 0.16 & -4.27 & 2.57 & -5.88 & -5.16 & -1.55 & -6.67 & 0.14 & -2.71 & 14.01 & -3.24 \\ 0.07 & -0.32 & -0.3 & 0.73 & 0.45 & -11.5 & 8 & -13.54 & 7.22 & -1.93 & -2.62 & -4.17 & 0 & 23.03 & -8.1 \\ 0.01 & 0 & -0.18 & 0.37 & 0.11 & 0.9 & -2.68 & -23.2 & 2.32 & 2.2 & 14.9 & -7.55 & -1.7 & 14.28 & 1.64 \\ 0.07 & 1 & -0.27 & 0.51 & -0.33 & -4.3 & -1.75 & 0 & 0.17 & -0.71 & -1.36 & 0.63 & -7.84 & 3.72 & -4.4 \end{pmatrix}.$$

Figure 2 shows the observed yearly sales of fresh food series and the fitted yearly sales of fresh food series obtained from the NVAR model. The solid blue line represents the actually observed yearly sales of fresh food series, the dash red line represents its fitted series based on the NVAR model. Figure 3 shows the observed yearly sales of takeaways food series and the fitted yearly sales of takeaways food series with the same description as in Figure 2. Figure 4 shows the observed yearly sales of supermarkets food series and the fitted yearly sales of supermarkets food series with the same description as in Figure 2. Similarly, Figure 5 shows the observed yearly sales of caférestaurants food series and the fitted yearly sales of caférestaurants food series with the same description as in Figure 2.

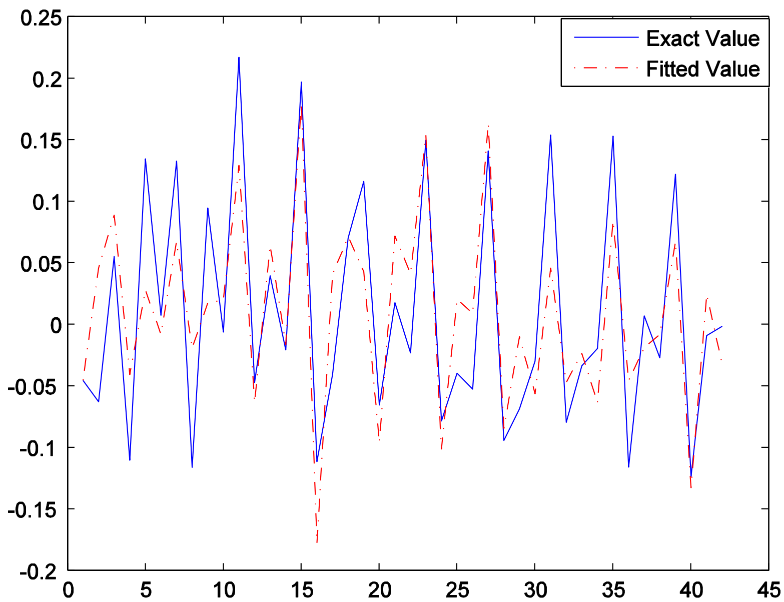

Figure 2. Observed and fitted values of the yearly sales of fresh food in New Zealand, *MSE*=0.006

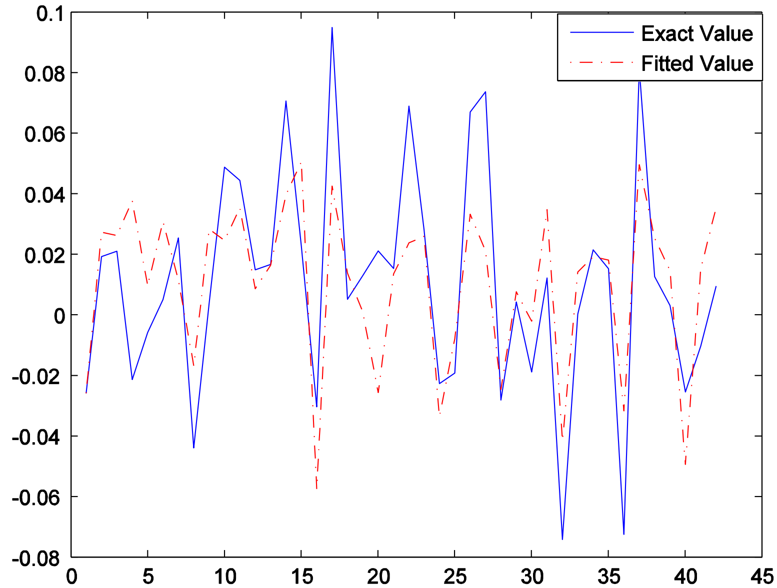

Figure 3. Observed and fitted values of the yearly sales of takeaways food in New Zealand, *MSE*=0.006

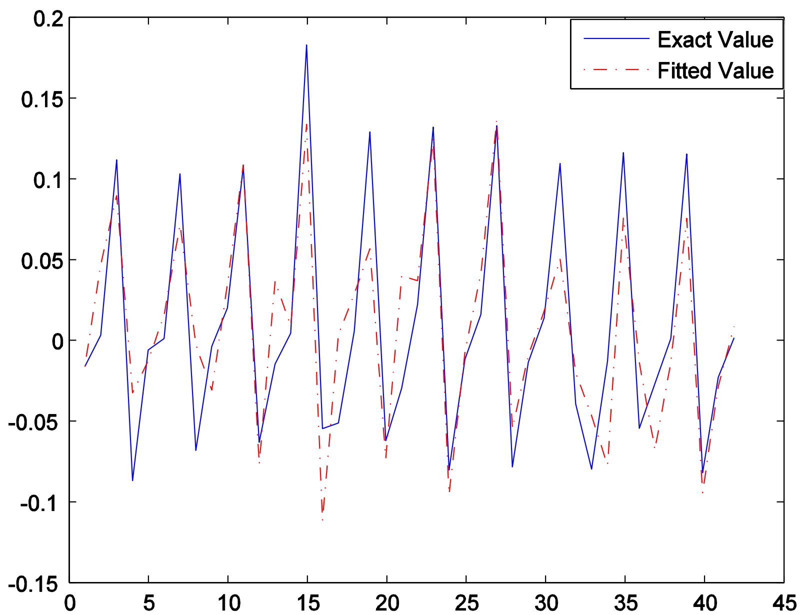

Figure 4. Observed and fitted values of the yearly sales of supermarkets food in New Zealand, *MSE*=0.006

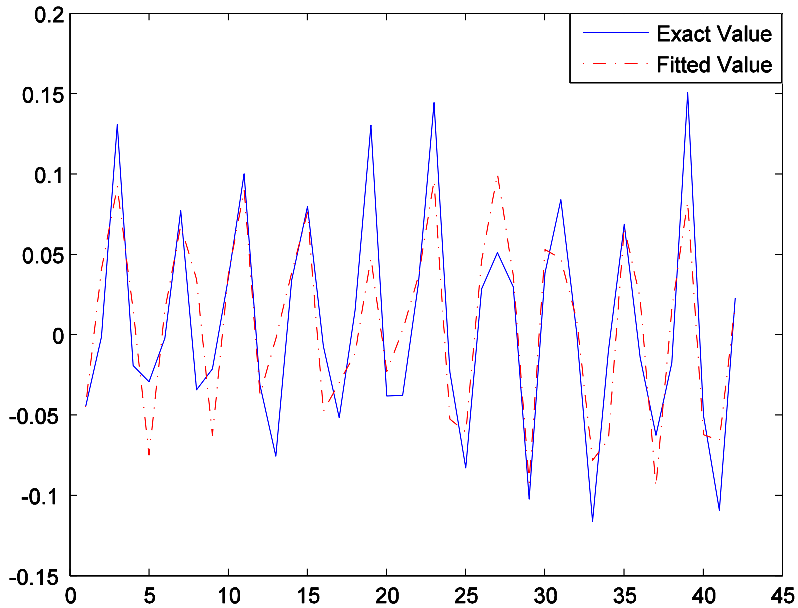

Figure 5. Observed and fitted values of the yearly sales of caférestaurants food in New Zealand, *MSE*=0.006

6. Test of nonlinearity

Here, we would like to test the hypothesis of nonlinearity in a vector time series model. It is worth mentioning that all the content in the following subsection is extracted from Psaradakis and Vavra (2014).

For this, one shall consider the VAR model for a m -variate time series $\{Y_t\}$ given by

$$Y_t = \mu + \sum_{i=1}^{k} \mathbf{A}_i Y_{t-i} + \varepsilon_t, \quad t \in Z, \tag{36}$$

where $k \geq 1$ is a fixed integer, μ is a $(m \times 1)$ vector of real constants and $\mathbf{A}_i$, $i = 1,2\ldots,k$, are $(m \times m)$ matrices of real constants.

A test for nonlinearity by the use of principal components could be applied as a test for the hypothesis $B_2 = 0$ in the multivariate regression as follows

$$\hat{\varepsilon}_t = b_0 + \mathbf{B}_1 v_t + \mathbf{B}_2 y_t + \eta_t, \qquad t = 1,2\ldots,T, \tag{37}$$

where $\hat{\varepsilon}_t$ is the $(m\times 1)$ vector of the least squares estimators from (35), $(b_0,\mathbf{B}_1,\mathbf{B}_2)$ are parameters, v_t is the $(mk\times 1)$ vector defined as $v_t=(Y'_{t-1},\ldots,Y'_{t-k})'$, and $y_t=(Y_{1t},\ldots,Y_{nt})'$, $1\le n\le m$, consisting of the first n sample principal components of w_t which is the $(\frac{1}{2}mk(mk+1)\times 1)$ vector defined as $w_t=vech(v_tv'_t)$. The i th principal component is calculated as $Y_{it}=\xi'_i w^*_t$, $i=1,\ldots,n$, where ξ_i is the normalized eigenvector and w^*_t is the standardized version of w_t, also η_t is an error term. It is to be considered that, for large values of the likelihoodratio statistic (Λ), linearity is rejected

$$\Lambda=(T-\tau)(\ln\det S_0-\ln\det S_1), \tag{38}$$

where S_1 and S_0 are the least squares residual sum of squares matrices from Eq. (36) with B_2 unrestricted and $B_2=0$, respectively, and $\tau=mk+\frac{1}{2}(m+n+3)$ is Bartlett's correction factor (see Anderson (2003)). For large T, Λ could be approximately considered as χ^2_{mn} under the null hypothesis that $\{Y_t\}$ satisfies the linear model in Equation (35).
For our empirical example, with n=3, Λ is 47.79 which clearly indicates, this model is nonlinear with respect to Y_{t-1}. In fact, n is the smallest integer in a way $\lambda_{n+1}\le\tilde{\lambda}$; we set $\tilde{\lambda}=0.7$ based on the suggestion made by Jollie (2002) and bearing in mind average-root rule.

7. Summary and Conclusion

In this research, a new semiparametric method to the first-order nonlinear vector autoregressive time series model is introduced. The first-order nonlinear vector autoregressive model is currently used in a variety of fields, including econometrics, financial studies, engineering and biology. This paper proposed the first-order nonlinear vector autoregressive in which the errors are independent and also, the errors and observations are independent for each *t*. Since, parametric methods are not very efficient to estimate vector regression function, semiparametric methods are used. MSE criterion is also applied to verify the accuracy of the suggested model. The results of the study indicate the accuracy of the suggested model.

In this work, instead of making a crude guess of true density function $F(.)$, we use a multivariate Taylor series expansion approximation to estimate the parametric function $G(Y_{t-1},\mathbf{\Theta})$. That is, we assume $G(Y_{t-1},\mathbf{\Theta})$ is the Taylor series expansion of the vector autoregression function $F(X)$ about the point $A_0=[\alpha_1,\alpha_2,\ldots,\alpha_m]^T$ up to the second order.

In fact, $\mathbf{\Theta}$ is a matrix of unknown parameters with dimension $(m\times(\frac{(m+1)(m+2)}{2}))$ which can be obtained by using the least squares estimation method. Then, we can adjust the initial Taylor series approximation by a diagonal adjustment factor matrix $\xi(\mathrm{Y}_{t-1})$ where this $\xi(\mathrm{Y}_{t-1})$ can be estimated nonparametrically. By minimizing the local L2-fitting criterion with respect to $\xi(\mathrm{Y}_{t-1})$, in the equation $F(\mathrm{Y}_{t-1}) = \xi(\mathrm{Y}_{t-1})G(\mathrm{Y}_{t-1}, \hat{\mathbf{\Theta}})$, one can obtain the estimator $\hat{\xi}(\mathrm{Y}_{t-1})$ of $\xi(\mathrm{Y}_{t-1})$. Therefore, we could suggest a semiparametric form as $\hat{\xi}(\mathrm{Y}_{t-1})G(\mathrm{Y}_{t-1}, \hat{\mathbf{\Theta}})$ for the unknown nonlinear vector autoregression function $F(.)$. In practice, we are not able to use the estimator $\hat{\xi}(\mathrm{Y}_{t-1})$ since it contains the unknown function $f(\mathrm{Y}_{t-1})$. By using the fact that errors of the model are small values, we could get the nonparametric adjustment diagonal matrix estimator of $\xi(\mathrm{Y}_{t-1})$ as $\tilde{\xi}\,(\mathrm{Y}_{t-1})$. Therefore, the nonlinear vector autoregression function estimator could be estimated by $\tilde{F}(\mathrm{Y}_{t-1})$ which is equal to $\tilde{\xi}(\mathrm{Y}_{t-1})G(\mathrm{Y}_{t-1}, \hat{\mathbf{\Theta}})$. One can concentrate on $\tilde{F}(\mathrm{Y}_{t-1})$ to estimate the unknown nonlinear vector autoregression function $F(.)$. Also, under some conditions, we could establish the asymptotic consistency properties of the proposed semiparametric method. The simulation results show that the semiparametric estiamator of the nonlinear vector autoregression function with independent errors performs well. The mean squares error (MSE) criterion is also applied to verify the efficiency of the suggested model. Moreover, the performance of the semiparametric method in our model for a real data, is considered with an empirical application. At the end, by the use of principal components the test of nonlinearity in the vector autoregression function model is applied. For our empirical example, this test clearly indicates the model is nonlinear with respect to Y_{t-1}.

***Acknowledgment**:*

The authors would like to thank the Editor and the referees for careful reading and for their comments which highly improved the paper.

REFERENCES

Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, third ed, Wiley, New York.

Cai, Z. and Masry, E. (2000). Nonparametric estimation of additive nonlinear ARX time series: Local linear fitting and projection, Vol. 16, No. 4, *Econometric Theory*, 465-501.

Chen, J., Kim, I., Terrel, G.R. and Liu, L. (2014). Generalised partial linear single-index mixed models for repeated measures data, Vol. 26, No. 2, *Journal of Nonparametric Statistics*, 291-303.

Chen, R. and Tsay, R. S. (1993). Functional-coefficient autoregressive models, Vol. 88, No. 421, *Journal of the American Statistical Association*, 298-308.

Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer, NewYork.

Farnoosh, R. and Mortazavi, S.J. (2011). A semiparametric method for estimating nonlinear autoregressive model with dependent errors, *Nonlinear Analysis*, 6358-6370.

Farnoosh, R., Hajebi, M. and Mortazavi, S.J. (2014). A semiparametric estimation for regression functions in the partially linear autoregressive time series model, *Applications and Applied Mathematics: An International Journal* (*AAM*), 573-591.

Haggen, V. and Ozaki, T. (1981). Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model, Vol. 68, No. 1, *Biometrika*, 189-196.

Hardle, W., Tsybakov, A. B. and Yang, L. (1998). Nonparametric vector autoregression, 68 (2), *Journal of Statistical Planning and Inference*, 221-245.

Harville, D. A. (1997). *Matrix Algebra from a Statisticians Perspective*. New York, Springer-Verlag.

Hjort, N.L. and Jones, M.C. (1996). Locally parametric nonparametric density estimation, Vol. 24, *Ann. Statist*, 1619-1647.

Hoshino, T. (2014). Quantile regression estimation of partially linear additive models, Vol. 26, No. 3, *Journal of Nonparametric Statistics*, 509-536.

Huang, J. Z. and Yang, L. (2004). Nonparametric time series analysis, Vol. 66, *J. R. Statist. Soc. B*, 463-477.

Jolliffe, I.T. (2002). *Principal Component Analysis*. second ed, Springer, New York, 2002.

Lai, T. L., Robbins, Herbert. and Wei, C. Z. (1978). Strong consistency of least squares estimates in multiple regression, Vol. 75, No. 7, *Proc. Natl. Acad. Sd. USA*, 3034-3036.

Li, Y. and Genton, M. G. (2009). Single-Index additive vector autoregressive time series models, Vol. 36, *Scandinavian Journal of Statistics*, 369-388.

Li, D., Phillips, P. C. B. and Gao, J. (2015). Uniform consistency of nonstationary kernel weighted sample covariances for nonparametric regression, Vol. 32, *Econometric Theory*,1-31.

Mingzhong, J. and Xiru, C. (1999). Strong consistency of least squares estimate in multiple regression when the error variance is infinite, Vol. 9, *Statistica Sinica*, 289-296.

Mortazavi, S.J. and Farnoosh, R. (2013). The prediction nonlinear-autoregressive model for annual ring width of pinus eldarica with semi-parametric approach, Vol. 26, No. 6, World Applied Sciences Journal, 783-787.

Nademi, A. and Farnoosh, R. (2014). Mixtures of autoregressiveautoregressive conditionally heteroscedastic models: semiparametric approach, Vol. 41, No. 2, *Journal of Applied Statistics*, 275-293.

Naito, K. (2004). Semiparametric density estimation by local l2-fitting, Vol. 32, *Ann. Statist*, 1162-1191.

Psaradakis, Z. and Vávra, M. (2014). On testing for nonlinearity in multivariate time series, Vol. 125, *Economics Letters*, 1-4.

Seber, G. A. F. (2008). *A Matrix Handbook for Statisticians*. Hoboken, Wiley.

Shao, Q. and Yang, L. J. (2011). Autoregressive coefficient estimation in nonparametric analysis, Vol. 32, *Journal of Time Series Analysis*, 587-597.

Yu, Z., Wang, D. and Shi, N. (2009). Semiparametric estimation of regression functions in autoregressive models, *Statistics and Probability Letters*, 165-172.