

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

**4,800**

Open access books available

**122,000**

International authors and editors

**135M**

Downloads

Our authors are among the

**154**

Countries delivered to

**TOP 1%**

most cited scientists

**12.2%**

Contributors from top 500 universities



**WEB OF SCIENCE™**

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.

For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Modeling Nonlinear Vector Time Series Data

---

Jiancheng Jiang and Sha Yu

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.70825>

---

## Abstract

In this chapter, we review nonlinear models for vector time series data and develop new nonparametric estimation and inference for them. Vector time series data exist widely in practice. In financial markets, multiple time series are usually correlated. When analyzing several interdependent time series, in general one should consider them as a single vector time series fitted by multivariate models, which provides a useful tool for modeling interdependencies among multiple time series and for simultaneously analyzing feedback and Granger causality effects. Since nonlinear features are widely observed in time series, we consider nonlinear methodology for modeling nonlinear vector time series data, which allows flexibility in the model structure and avoids the curse of dimensionality.

**Keywords:** cointegration, VAR, multivariate threshold autoregressive model, nonparametric smoothing, generalized likelihood ratio

---

## 1. Introduction

Multiple time series are of considerable interest in an array of domains, such as finance, economics, engineering and so on. The data are collected in time order and consist of several related variables of interest, for instance, the data of stock price indexes and the status data of important instruments such as shuttles. It is of much practical significance to model this kind of data well. Moreover, a lot of commonly seen multiple time series are correlated, which makes it reasonable to regard them as a single vector and to fit them using multivariate models. Multivariate models perform well in exploring the interdependencies among multiple time series and capturing the dynamic structure.

Plenty of contributions have been made in the field of parametric models for multivariate time series. For instance, Sims proposed vector autoregressive (VAR) models in 1980 [1], Engle and Kroner considered multivariate generalized autoregressive conditional heteroscedastic (GARCH) models in 1995 [2], and Tsay developed the multivariate threshold models in 1998 [3]. Compared

to parametric models, nonparametric models require less assumption about the model structure and are more flexible. Combined with the fact that nonlinearity widely exists in time series, it is ideal to model the multiple time series using nonparametric models. However, not much of achievements have been made about this. This is partly due to the complexity of nonparametric smoothing as well as the curse of dimensionality. With these objectives in mind, Jiang proposed the multivariate functional-coefficient model in 2014 [4], which provides a useful tool for modeling vector time series data.

In this chapter, we first review some vector time series models, next extend them to include an error-correction term by incorporating cointegration among integrated variables, then develop a single index model for choosing the smoothing variable and a variable selection method for the multivariate functional-coefficient models, and finally study multivariate time-varying coefficient models and related hypothesis testing problems.

The remainder of this chapter is organized as follows. In Section 2 we review vector autoregressive (VAR) models. In Section 3, we consider multivariate functional-coefficient regression models and their extensions, where a model selection rule is also proposed. In Section 4 we introduce multivariate time-varying coefficient models and propose a generalized likelihood ratio test. In Section 5 we make a conclusion and discuss some interesting research topics to be completed.

## 2. Review of VAR models

The vector autoregressive model is a generalization of the univariate autoregressive model for forecasting a vector of time series. This model was pioneered by Sims in Ref. [1] and it has acquired a prominent role in analyzing macroeconomic time series. Prior to 1980, large-scale statistical dynamic simultaneous equations model (DSEMs) was widely used in empirical macroeconomics, which often contained dozens or even hundreds of equations. As the economic environment has grown more complicated, the traditional simultaneous models have grown. Sims believed that since these models do not dichotomize variables into “endogenous” and “exogenous,” the exclusion restrictions used to identify the simultaneous equations models make little sense. Thus, he advocated the vector autoregressive model (VAR) to model the interrelationships among a set of macroeconomic variables. In the structure of VAR models, each variable is a linear function of past lags of itself and past lags of the other variables. Sims demonstrated that VARs provide a flexible and tractable framework for analyzing economic time series. While hardly relying on economic theorems, VAR models have proven efficient in capturing the dynamics of multivariate systems as well as forecasting [1]. Specifically, a vector autoregressive model of order  $p$  [VAR( $p$ )] has the following general form:

$$y_t = c + A_1 y_{t-1} + \dots + A_p y_{t-p} + e_t \quad (1)$$

where  $y_t = (y_{1t}, \dots, y_{Kt})'$  is a set of  $K$  time series variables,  $c$  is a  $K \times 1$  vector of constant,  $A_i$ 's are  $K \times K$  coefficient matrices, and  $e_t$  are error terms. Usually,  $e_t$  are assumed to be zero-mean

independent white noise with time-invariant and positive-definite covariance matrix  $\Sigma$ . For example, a VAR (1) model with two time series components can be written as:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix}$$

or the equation set

$$\begin{aligned} y_{1t} &= c_1 + A_{11}y_{1,t-1} + A_{12}y_{2,t-1} + e_{1,t} \\ y_{2t} &= c_2 + A_{21}y_{1,t-1} + A_{22}y_{2,t-1} + e_{2,t} \end{aligned}$$

Using lag-operator  $L$ , Eq. (1) can be written as the following form:

$$y_t = c + (A_1L + A_2L^2 + \dots + A_pL^p)y_t + e_t \quad (2)$$

Let  $A(z) = I - A_1z - A_2z^2 - \dots - A_pz^p$ , where  $z$  is a complex number. Then the VAR process is stable if

$$\det(A(z)) \neq 0 \text{ for } |z| \leq 1. \quad (3)$$

In other words, the determinant of the matrix polynomial has no roots in and on the complex unit circle. If the stability conditions are satisfied and the process can be extended to the infinite past, then the VAR process is stationary.

For model (1), since the right-hand side consists of only predetermined variables and the error terms are assumed to be independent white noise with time-invariant covariance, each equation can be estimated by ordinary least squares (OLS). Zellner proved that the OLS estimator coincides with the generalized LS (GLS) estimator [5].

The celebrated model (1) is easy to fit, and its autoregressive structure allows one to study the feedback effects and the Granger causality. However, model (1) employs only the lagged values of  $y_t$  for forecast and ignores other potentially important variables' effect. In addition, as time evolved, the coefficients remain constant, which may contrast the real situations where the dynamic structure of the relationship among different time series involves with time.

### 3. Multivariate functional-coefficient regression models and extensions

We briefly reviewed VAR models in the previous section. This parametric method has been significantly developed and widely applied to econometric dynamics as well as other domains. An alternative to modeling vector time series is the nonparametric method, which requires much fewer assumptions on the model structure and may shed light on the later parametric fitting. To illustrate the basic idea of this approach, let us begin with the multivariate threshold autoregressive model [3].

### 3.1. Multivariate threshold autoregressive model

The multivariate threshold autoregressive model is a generalization of the univariate threshold autoregressive model [6]. The idea is to partition one-dimensional variable into  $s$  regimes and impose an AR model with exogenous variables in each regime. Consider a  $k$ -dimensional time series  $y_t = (y_{1t}, \dots, y_{kt})'$  and a  $v$ -dimensional exogenous variable  $x_t = (x_{1t}, \dots, x_{vt})'$ , for  $t = 1, \dots, n$ . Let  $-\infty = r_0 < r_1 < \dots < r_s = \infty$ . The multivariate threshold model with threshold variable  $z_t$  and delay  $d$  has the following form:

$$y_t = c_j + \sum_{i=1}^n \phi_i^{(j)} y_{t-i} + \sum_{i=1}^q \beta_i^{(j)} x_{t-i} + \varepsilon_t^{(j)} \text{ if } r_{j-1} < z_{t-d} \leq r_j (j = 1, \dots, s), \quad (4)$$

where  $p$  and  $q$  are nonnegative integers and  $\varepsilon_t^{(j)} = \Sigma_j^{\frac{1}{2}} a_t$ , with  $\Sigma_j^{\frac{1}{2}}$  being a positive-definite matrix and  $\{a_t\}$  a sequence of serially uncorrelated random vectors with mean zero and covariance matrix  $I_k$ . The threshold variable  $z_t$  is assumed to be stationary and has a continuous distribution.

Model (4) is piecewise linear in the threshold space of  $z_{t-d}$ , but it is nonlinear when  $s > 1$  [3]. This model has proven to be useful in practice. Nevertheless, the assumption embedded in this model weakens the practicability, that is, the coefficients are assumed to be constants in the threshold space of  $z_{t-d}$  in model (4). This assumption is questionable since the economic conditions tend to change slowly over time and the coefficient functions may vary smoothly. Motivated by this, Jiang proposed the multivariate functional-coefficient model, in which the coefficients are functions of threshold variable  $z_{t-d}$  instead of constants [4].

### 3.2. Multivariate functional-coefficient models

The multivariate functional-coefficient model has the following form:

$$y_t = c(z_{t-d}) + \sum_{i=1}^p \phi_i(z_{t-d}) y_{t-i} + \sum_{i=1}^q \beta_i(z_{t-d}) x_{t-i} + \varepsilon_t, \quad (5)$$

where  $c(\cdot)$  is a  $k \times 1$  functional vector,  $\phi_i(\cdot)$  are  $k \times k$  functional matrices, and  $\beta_i(\cdot)$  are  $k \times v$  functional matrices. The innovation satisfies  $\varepsilon_t = \sigma_t^* a_t$ , where  $\sigma_t^*$  is a positive-definite matrix and  $\{a_t\}$  as in Eq. (4). Assume that  $\sigma_t^*$  is measurable with respect to the  $\sigma$ -field generated by the historical information  $\mathcal{F}_{t-1} = \{(w_j, z_{j-d}) : j \leq t\}$ , where  $w_j = (x_{j-1}, \dots, x_{j-q}, y_{j-1}, \dots, y_{j-p})$ . For model (5), we are interested in estimating the regression part. Once it is estimated, one may consider making simultaneous inference about parameters and using the residuals to study the structure of the volatility matrix. This model is a generalization of vector autoregressive models [1], threshold models [3] and functional-coefficient models [7–10]. Even for one-dimensional settings with  $k = 1$ , model (5) includes important predictive regression models in econometrics, such as the linear predictive models with nonstationary predictors [11–13] and functional-coefficient models for nonstationary time series data [14]. Model (5) can also be used to investigate the Granger Causality [15–17] and the feedback effect in engineering and finance [18, 19].

For model (5), a weighted local least squares estimation method was provided in [4]. Let  $X_t = \text{vec}(1, y_{t-1}, \dots, y_{t-p}, x_{t-1}, \dots, x_{t-p})$  and  $\Phi(z) = (c(z), \phi_1(z), \dots, \phi_p(z), \beta_1(z), \dots, \beta_q(z))$ . Then model (5) becomes

$$y_t = \Phi(z_{t-d})X_t + \varepsilon_t, \tag{6}$$

where  $\Phi(\cdot)$  is a  $k \times m$  matrix-valued function and  $X_t$  is an  $m \times 1$  vector with  $m = 1 + pk + qv$ . For any  $z_{t-d}$  in the neighborhood of  $z$ , by the Taylor expansion, we have

$$\Phi(z_{t-d}) \approx \Phi(z) + \Phi'(z)(z_{t-d} - z) \equiv A + B(z_{t-d} - z).$$

Let  $S$  and  $V$  be  $2 \times 2$  matrices whose  $(i, j)$ th elements are  $\mu_{i+j-2} = \int u^{i+j-2} K(u) du$  and  $\nu_{i+j-2} = \int u^{i+j-2} K^2(u) du$ , respectively, and let  $s = (\mu_2, \mu_3)'$ . Given any invertible working variance matrix  $\sigma_t^2$  of  $\sigma_t^{*2}$ , the estimator  $(\tilde{A}, \tilde{B})$  is achieved by minimizing

$$\sum_{t=s'+1}^n \|\sigma_t^{-1} [y_t - AX_t - BX_t(z_{t-d} - z)]\|^2 K_{h_n}(z_{t-d} - z),$$

where  $\|\cdot\|$  denotes the Euclidean norm,  $s' = \max(p, d, q)$ , and  $K_{h_n}(x) = h_n^{-1} K(x/h_n)$  for kernel function  $K(\cdot)$  with bandwidth  $h_n$  controlling the amount of smoothing. Let  $K_{h_n}^{(i)}(z_{t-d} - z) = h_n^{-i} (z_{t-d} - z) K_{h_n}(z_{t-d} - z)$  and  $\tilde{S}_{ni} = \sum_{t=s'+1}^n (X_t X_t^T) \otimes \sigma_t^{-2} K_{h_n}^{(i)}(z_{t-d} - z)$  for  $i = 0, 1, 2$ . Then the weighted estimators  $(\tilde{A}, \tilde{B})$  admit the closed form:

$$\begin{pmatrix} \text{vec}(\tilde{A}) \\ \text{vec}(h_n \tilde{B}) \end{pmatrix} = \begin{pmatrix} \tilde{S}_{n0} & \tilde{S}_{n1} \\ \tilde{S}_{n1} & \tilde{S}_{n2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=s'+1}^n (X_t \otimes \sigma_t^{-2}) y_t K_{h_n}(z_{t-d} - z) \\ \sum_{t=s'+1}^n (X_t \otimes \sigma_t^{-2}) y_t K_{h_n}^{(1)}(z_{t-d} - z) \end{pmatrix}.$$

Under certain conditions, the weighted estimators are asymptotically normal (see [4]).

Recall that, in model (5),  $\sigma_t^*$  is a positive-definite matrix measurable with respect to the sigma-algebra generated by historical information. If there is a parametric structure of  $\sigma_t^*$ , for example, the generalized autoregressive conditional heteroscedastic (GARCH) errors [4], then it helps to improve the efficiency of the weighted estimation. Example 3 in [4] exemplifies this point. Our intuition is that, if a parametric structure of  $\sigma_t^*$  is correctly specified, then the weighted estimation mimics the oracle estimation in the sense that  $\sigma_t^*$  is known. This intuition can be verified theoretically since  $\sigma_t^*$  can be estimated at rate of  $\sqrt{n}$  which is faster than what we can do for the regression function in model (5).

### 3.3. Extension of multivariate functional-coefficient models

Due to the fact that many economic factors are not stationary, classic regression analysis requiring the stationarity condition suffers from a great limitation. Cointegration analysis has

become a formidable toolkit in analyzing non-stationary economic time series. The concept of cointegration goes back to Granger [20] and initiated a literal research boom. Engle & Granger proposed the well-known Engel-Granger test to examine whether there is a cointegrating relationship among a set of first-order integrated variables [21].

Motivated by Granger and Engle & Granger, Jiang proposed an error-correction version of model (5) by incorporating the cointegrating relationship of first-order integrated variables [4]. This allows us to cope with the nonstationarity of vector time series and to improve the accuracy of forecasting.

Let  $s_t$  denote a  $k \times 1$  vector of first-order integrated variables and let  $y_t = s_t - s_{t-1}$ . Assume that there is a co-integrating relationship for  $s_t$ ; that is, there exists a unique  $k \times r$  ( $0 < r < k$ ) deterministic matrix  $\theta$  of rank  $r$  and a stationary process  $u_t$  such that  $\theta^T s_t = u_t$ . Then an error-correction form of model (5) is

$$y_t = c(z_{t-d}) + \gamma(z_{t-d})u_{t-1} + \sum_{i=1}^p \phi_i(z_{t-d})y_{t-i} + \sum_{i=1}^q \beta_i(z_{t-d})x_{t-i} + \varepsilon_t, \quad (7)$$

where  $\gamma(z_{t-d})$  is a  $k \times r$  coefficient matrix. This model simplifies to the Granger representation theorem if the coefficient functions are constant and there are no exogenous variables [4].

Due to the widespread presence of cointegrating variables in finance and economics, model (7) should improve the practicability of model (5). However, model (7) requires specification of variable  $z_t$ . This can be relaxed by using the idea of single index models. Recall that model (5) can be represented in succinct form (6). The similar operations can be applied to model (7). Now set  $z_t = \gamma^T X_t$  and let data decide the value of  $\gamma$ . Then model (7) can be extended as

$$y_t = \Phi(\gamma^T X_{t-d})X_t + \varepsilon_t, \quad (8)$$

where  $\gamma$  is a directional vector such that its first nonzero entry is positive. Model (8) is more flexible than model (7), it is key to estimate  $\gamma$ . We introduce the profile least squares method to estimate model (8). The estimation procedure consists of several steps:

**Step 1.** Given an initial value of  $\gamma$ , one obtains the weighted estimator  $\hat{\Phi}(\cdot; \gamma)$  of coefficient function in the same way as for model (6).

**Step 2.** Find the value  $\hat{\gamma}$  to minimize

$$\sum_{t=s'+1}^n \|y_t - \hat{\Phi}(\gamma^T X_{t-d}; \gamma)X_t\|^2. \quad (9)$$

**Step 3.** Update the value of  $\gamma$  by  $\hat{\gamma}$ , and repeat Step 1 and Step 2 many times until convergence. The coefficient function  $\Phi(\cdot)$  is estimated by  $\hat{\Phi}(\cdot; \hat{\gamma})$ .

It can be shown that  $\hat{\Phi}(\cdot; \hat{\gamma})$  shares the same asymptotic normality as the Oracle weighted estimator in the sense that it knows the true value of  $\gamma$ , since  $\hat{\gamma}$  is  $\sqrt{n}$ -consistent.

### 3.4. Variable selection of multivariate functional-coefficient models

In this section, we consider variable selection of model (6). Increasing the lags  $p$  and  $q$  will necessarily reduce the sum of squared errors. However, doing so will increase the burden of coefficient estimation and may also lead to overfitting. Hence, for the multivariate functional-coefficient model, order selection is of much importance.

Two widely used model selection criteria are Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). However, these stepwise methods yield heavy burden on computation and furthermore bring difficulty in establishing asymptotics for the estimation of selected models. The problems become more severe for high-dimensional data. Various regularization methods have been proposed to deal with these problems. Among them, a popular approach, called LASSO, proposed by Tibshirani, performs variable selection and parameter estimation simultaneously. See Ref. [22]. For univariate varying-coefficient regression models with i.i.d. data, Wang and Xia [23] developed a shrinkage estimation method by combining the idea of group LASSO [24] and kernel smoothing. In the following we develop a shrinkage estimation method for multivariate functional-coefficient model (6):

$$y_t = \Phi(z_{t-d})X_t + \varepsilon_t,$$

where the functional-coefficient matrix  $\Phi(z) = (c(z), \phi_1(z), \dots, \phi_p(z), \beta_1(z), \dots, \beta_q(z))$ . Since each column of  $\Phi(\cdot)$  corresponds to the effect of a component of  $X_t$ , for variable selection of  $X_t$  we should penalize each column of  $\Phi(\cdot)$  as a whole. This leads to minimizing

$$Q_\lambda(\Phi) = \sum_{i=s'+1}^n \sum_{t=s'+1}^n \|y_t - \Phi(z_{i-d})X_t\|^2 K(h^{-1}(z_{t-d} - z_{i-d})) + \sum_{j=1}^{p+q+1} \lambda_j \|\Phi_j\|, \quad (10)$$

where  $\Phi_j = (\Phi_j(z_{s'+1-d}), \dots, \Phi_j(z_{n-d}))$  with  $\Phi_j(\cdot)$  being the  $j$ th column of  $\Phi(\cdot)$ ,  $\lambda_j$ 's are tuning parameters, and for any matrix  $A$  we use  $\|A\|$  to denote the Hilbert-Schmidt norm of matrix. It is interesting to establish model selection consistency and the oracle property of the shrinkage estimation.

## 4. Multivariate time-varying coefficient models

Parallel to functional-coefficient model (5), it is natural to consider its alternative with time-varying coefficients [25]:

$$y_t = c(t/T) + \sum_{i=1}^p \phi_i(t/T)y_{t-i} + \sum_{i=1}^q \beta_i(t/T)x_{t-i} + \varepsilon_t, t = 1, \dots, T, \quad (11)$$

where  $y_t$  is a  $k \times 1$  vector,  $x_t$  is a  $v \times 1$  vector,  $c(s)$  a  $k \times 1$  vector,  $\phi_i(\cdot)$  are  $k \times k$  smooth matrices and  $\beta_i(\cdot)$  are  $k \times v$  smooth matrices. The innovation satisfies the same conditions as model (5). It is known that as time involves the economic conditions change slowly and smoothly. Model (11) reflects this smoothing change by allowing the coefficients being smoothing functions of time.



Let

$$\Phi(t/T) = \left( c(t/T), \phi_1(t/T), \dots, \phi_p(t/T), \beta_1(t/T), \dots, \beta_q(t/T) \right).$$

Using similar arguments to model (6), we can rewrite model (11) as

$$y_t = \Phi(t/T)X_t + \varepsilon_t, t = 1, \dots, T, \quad (12)$$

where  $\Phi(\cdot)$  is a  $k \times m$  matrix and  $X_t$  is the same as in model (6). By the Taylor expansion, for any  $t$  in the neighborhood of  $t_0 \in (0, T)$ , we have

$$\Phi(t/T) \approx \Phi(t_0/T) + \Phi'(t_0/T)((t - t_0)/T) \equiv P + Q((t - t_0)/T).$$

Running the local linear smoother for model (12), we minimize

$$\sum_{t=s+1}^T \|y_t - PX_t - QX_t((t - t_0)/T)\|^2 K_h(t - t_0) \quad (13)$$

over  $P$  and  $Q$ , where  $s = \max(p, q)$  and  $K_h(x) = h^{-1}K(x/hT)$ . Then it is straightforward to obtain an explicit form of the minimizer,  $(\hat{P}, \hat{Q})$ , for the above optimization problem,

$$\begin{pmatrix} \text{vec}(\hat{P}) \\ \text{vec}(h\hat{Q}) \end{pmatrix} = \begin{pmatrix} S_{T0} & S_{T1} \\ S_{T1} & S_{T2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=s+1}^T (X_t \otimes I_k) y_t K_h(t - t_0) \\ \sum_{t=s+1}^T (X_t \otimes I_k) y_t K_h^{(1)}(t - t_0) \end{pmatrix}, \quad (14)$$

where  $S_{Ti} = \sum_{t=s+1}^T (X_t X_t^T) \otimes I_k K_h^{(i)}(t - t_0)$  and  $K_h^{(i)}(t - t_0) = (Th)^{-i} (t - t_0)^i K_h(t - t_0)$ , for  $i = 0, 1, 2$ .

Define  $M = E[(X_t X_t^T) \otimes I_k]$  and  $N = E[(X_t X_t^T) \otimes (\sigma_t^*)^2]$ . Let  $\mu_i = \int u^i K(u) du$ ,  $v_i = \int u^i K^2(u) du$ ,

$$U = \begin{pmatrix} \mu_0 & \mu_1 \\ \mu_1 & \mu_2 \end{pmatrix}, V = \begin{pmatrix} v_0 & v_1 \\ v_1 & v_2 \end{pmatrix}$$

Using similar arguments to [4], we can show that this estimator is asymptotically normal with mean zero and variance  $\Sigma$ , where  $\Sigma = (U^{-1} V U^{-1}) \otimes (M^{-1} N M^{-1})$ .

#### 4.1. Generalized likelihood ratio tests

The multivariate time-varying coefficient regression model is flexible and powerful to estimate the dynamic changes of coefficients. After fitting a given dataset, some important questions arise, for example, whether the coefficient functions are actually constant or of some particular forms? This leads to statistical hypothesis testing. To answer these questions, we develop

generalized likelihood ratio statistics to test corresponding hypothesis testing problems about the coefficient functions [26].

For the multivariate time-varying coefficient model (12), assume  $\Sigma_0^{-1/2}\varepsilon_t$  has mean zero and covariance matrix  $I_k$  with  $\Sigma_0$  being a symmetric positive-definite constant matrix.

Consider the following hypothesis testing problem

$$H_0 : \Phi(t/T) \in \Theta_0(t/T) \leftrightarrow H_a : \Phi(t/T) \notin \Theta_0(t/T), \quad (15)$$

where  $\Theta_0(t/T)$  is some known constant matrix  $\Phi_0$  or a set of functional matrices. Let  $\hat{\Phi}(t/T)$  denote the nonparametric estimator of  $\Phi$ , and let  $\hat{\Phi}_0(t/T)$  denote the true or estimated value of coefficients under the null hypothesis. Following Fan et al. [26] and Fan and Jiang [27], we define a generalized likelihood ratio statistic for testing problem (15):

$$\lambda_T = \frac{T}{2} \log \left( \frac{RSS_0 - RSS_a}{RSS_a} \right), \quad (16)$$

where  $RSS_0 = \sum_{t=1}^T (y_t - \hat{\Phi}_0(t/T)X_t)^T \Sigma^{-1} (y_t - \hat{\Phi}_0(t/T)X_t)^T$ , and  $RSS_a = \sum_{t=1}^T (y_t - \hat{\Phi}(t/T)X_t)^T \Sigma^{-1} (y_t - \hat{\Phi}(t/T)X_t)^T$  with  $\Sigma$  being a known constant covariance matrix from a working model. It is meaningful to study the asymptotic distributions of the test statistic under the null and alternatives.

In the following example, we consider the case when  $\Theta_0(\cdot)$  is a known constant. For any  $u = t/T \in (0, 1)$ , if we rewrite matrix  $\Phi(u)$  as a vector,  $\Delta(u) \equiv \text{vec}(\Phi_1(u), \dots, \Phi_m(u))$ , and denote  $\Delta_0(u) \equiv \text{vec}(\Phi_{01}^*(u), \dots, \Phi_{0M}^*(u))$ , then the power of the test is evaluated against alternatives:

$$H_a : \Delta(u) = \Delta_0(u) + \frac{1}{\sqrt{Th}} G(u), \quad (17)$$

where  $G(u) = (g_1(u), \dots, g_m(u))^T$  is a vector of functions.

**Example 1.** To investigate the performance of the proposed generalized likelihood ratio test, 600 replications for each of sample sizes  $T=200$ ,  $T=400$  and  $T=800$  from the multivariate time-varying coefficient model were generated:

$$y_t = \Phi(t/T)X_t + \varepsilon_t, t = 1, \dots, T$$

where  $k=2$ ,  $v=p=q=1$ ,  $\Delta = \text{vec}(0.5, 0.0074, 0.08, 0.65, 0.25, 0.75)^T$ . We set the initial values  $x_1=0$  and  $y_1=(0.15, 0.2)$ . Accordingly,  $X_t = \text{vec}(y_{1,t-1}, y_{2,t-1}, x_{t-1})$  for  $t=2, \dots, T$ . Three distributions of the error term are considered: bivariate normal, bivariate log-normal, and bivariate  $t(5)$ , each with variance matrix  $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$ . According to alternative (17), the power of the test is evaluated for a sequence of alternatives index by  $\theta$ :

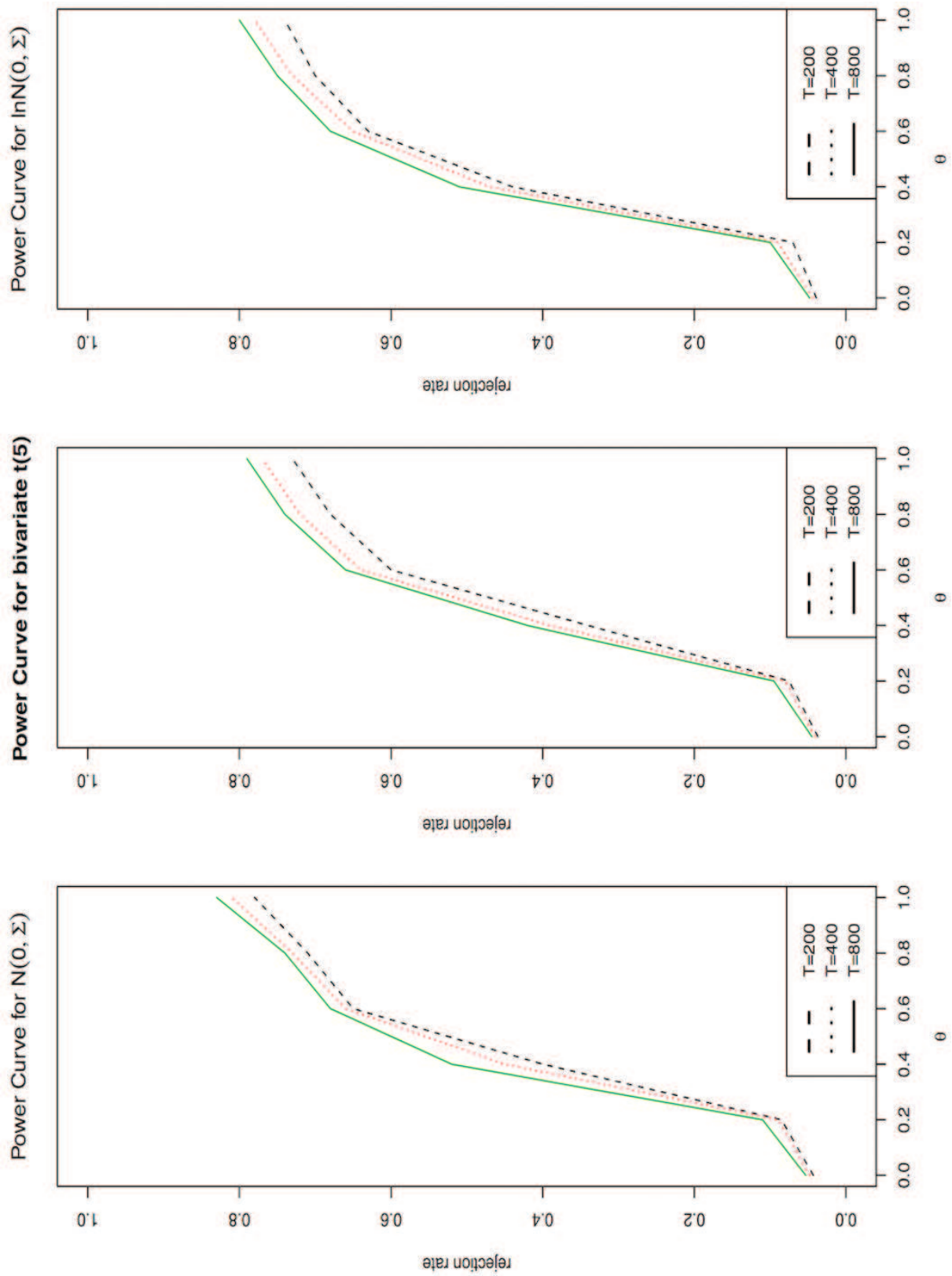


Figure 1. The power curves for Example 1. Significance level is 5%.

$$H_\theta : \Delta_\theta = (0.5, 0.0075, 0.08, 0.65, 0.25, 0.75)^T + \frac{\theta}{\sqrt{Th}} G(t/T), \quad (18)$$

where  $G(t/T) = ((\sin(\sqrt{2}\pi/T), -0.09 \cos(\pi t/T), 0.16 \sin(\sqrt{3}\pi/T), 0.8 \sin(\sqrt{2}\pi/T), 0.3 \sin(t/T), \cos(\sqrt{1.5}\pi t/T))^T$  and  $\theta = 0, 0.2, 0.4, 0.6, 0.8, 1$ . The power function is estimated by the relative rejection frequency of  $H_0$  in the above replicates.

The significance level is set to be 5%, and the critical values in simulations are calculated similarly by using the conditional bootstrap method in Ref. [26] for each given  $\theta$  value. Detail of this method is as follows:

**Step 1.** Compute the estimators of the coefficient  $\hat{\Phi}(t/T)$  under both the null and the alternative by setting the optimal bandwidth as the estimated value  $\hat{h}_{opt}$ .

**Step 2.** Compute the test statistic  $\lambda_T(H_0)$  and the residuals  $\{e_t\}$  from the alternative model.

**Step 3.** For each given  $X_t$ , draw a bootstrap residual  $e_t^*$  from the centered empirical distribution of  $e_t$  and compute  $y_t^* = \hat{\Phi}(t/T)X_t + e_t^*$ . This forms a conditional bootstrap sample  $\{X_t, y_t^*\}_{t=1}^T$ .

**Step 4.** Compute the test statistic  $\lambda_T^*(H_0)$  using the bootstrap sample constructed in Step 3.

**Step 5.** Repeat Step 3 and Step 4 to get a sample of the test statistic  $\lambda_T^*(H_0)$ . The critical values at significance level  $\alpha$  are calculated by the  $100(1 - \alpha)$ th percentile of the sample.

**Figure 1** displays the power curves in difference scenarios. We can tell from **Figure 1** that the patterns of power curves look like half of an inverted normal density. All the curves rise monotonically from a height equal to the significance level of 5% until eventually it reaches its maximum height of around 90%. It is evident from **Figure 1** that the test is powerful for all three different distributions of error terms. Moreover, the test becomes more powerful as sample size increases. These indicate that the proposed test keeps the size and is powerful for distinguishing the difference between the null and the alternative.

## 5. Conclusions

In this chapter, we have reviewed some parametric and nonparametric methods for modeling nonlinear vector time series data, which include the VAR model, the multivariate threshold autoregressive model, and the multivariate functional-coefficient regression model. These models have great significance in econometrical and statistical theory and application. Based on the weighted local least square estimation, we have proposed a variable selection method for the functional-coefficient model. This model selection procedure is applicable to the proposed multivariate single index models and multivariate time-varying coefficient models. We have also extended the generalized likelihood ratio test to the time-varying coefficient model

and demonstrated its performance through simulation. The proposed methodology is very useful for modeling nonlinear dynamic structures inherited in financial data. However, there are many problems remain unsolved for our procedure, such as the limiting theory about the proposed methodology. Future work includes, but not limited to, extending our models to nonstationary settings and exploring their performance in different applications.

## Acknowledgements

This research was supported in part by NSFC grant 71361010 and by funds provided by the University of North Carolina at Charlotte.

## Author details

Jiancheng Jiang\* and Sha Yu

\*Address all correspondence to: [jjiang1@uncc.edu](mailto:jjiang1@uncc.edu)

University of North Carolina at Charlotte, Charlotte, NC, USA

## References

- [1] Sims CA. Macroeconomics and reality. *Econometrica*. 1980;**48**(1):1-48. DOI: 10.2307/1912017
- [2] Engle RF, Kroner KF. Multivariate simultaneous generalized ARCH. *Econometric Theory*. 1995;**11**(1):122-150. DOI: 10.1017/S0266466600009063
- [3] Tsay RS. Testing and modeling multivariate threshold models. *Journal of the American Statistical Association*. 1998;**93**(443):1188-1202. DOI: 10.1080/01621459
- [4] Jiang J. Multivariate functional-coefficient regression models for nonlinear vector time series data. *Biometrika*. 2014;**101**(3):689-702. DOI: 10.1093/biomet/asu011
- [5] Zellner A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*. 1962;**57**(298):348-368. DOI: 10.2307/2281644
- [6] Tong H. On a threshold model. In: Chen C, editor. *Pattern Recognition and Signal Processing*. Netherlands: Sijthoff & Noordhoff; 1978. p. 575-586. DOI: 10.1007/978-94-009-9941-1\_24
- [7] Chen R, Tsay RS. Functional-coefficient autoregressive models. *Journal of the American Statistical Association*. 1993;**88**(421):298-308. DOI: 10.2307/2290725

- [8] Hastie T, Tibshirani R. Varying-coefficient models. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1993;**55**(4):757-796
- [9] Cai Z, Fan J, Yao Q. Functional-coefficient regression models for nonlinear time series. *Journal of the American Statistical Association*. 2000;**95**(451):941-956. DOI: 10.2307/2669476
- [10] Huang JZ, Shen H. Functional coefficient regression models for non-linear time series: A polynomial spline approach. *Scandinavian Journal of Statistics*. 2004;**31**(4):515-534. DOI: 10.1111/j.1467-9469.2004.00404.x
- [11] Campbell JY, Yogo M. Efficient tests of stock return predictability. *Journal of Financial Economics*. 2006;**81**:27-60
- [12] Paye BS, Timmermann A. Instability of return prediction models. *Journal of Empirical Finance*. 2006;**13**:274-315
- [13] Cai Z, Wang Y. Testing predictive regression models with nonstationary regressors. *Journal of Econometrics*. 2014;**178**:4-14
- [14] Cai Z, Li Q, Park JY. Functional-coefficient models for nonstationary time series data. *Journal of Econometrics*. 2009;**148**:101-113
- [15] Sims CA. Money, income, and causality. *American Economic Review*. 1972;**62**:540-552
- [16] Psaradakis Z, Ravn MO, Sola M. Markov switching causality and the money-output relationship. *Journal of Applied Econometrics*. 2005;**20**:665-683
- [17] Berger H, Österholm P. Does money still matter for U.S. output? *Economics Letters*. 2009;**102**:143-146
- [18] Subrahmanyam A, Titman S. Feedback from stock prices to cash flows. *Journal of Finance*. 2001;**56**:2389-2413
- [19] Åström KJ, Murray RM. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press: Princeton, NJ; 2010
- [20] Granger CWJ. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*. 1981;**16**(1):121-130. DOI: 10.1016/0304-4076(81)90079-8
- [21] Engle RF, Granger CWJ. Co-integration and error correction: Representation, estimation, and testing. *Econometrica*. 1987;**55**(2):251-276. DOI: 10.2307/1913236
- [22] Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1996;**58**(1):267-288. DOI: 10.1111/j.1467-9868.2011.00771.x
- [23] Wang H, Xia Y. Shrinkage estimation of the varying coefficient model. *Journal of the American Statistical Association*. 2009;**104**(486):747-757. DOI: 10.1198/jasa.2009.0138

- [24] Yuan M., Lin Y. Model Selection and Estimation in Regression With Grouped Variables. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*. 2006;**68**(1):49-67. DOI: 10.1111/j.1467-9868.2005.00532.x
- [25] Liu Y. Generalized quasi-likelihood ratio statistics for multivariate time-varying coefficient regression models [dissertation]. Charlotte: University of North Carolina at Charlotte; 2013. 47 p
- [26] Fan J, Zhang C, Zhang J. Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of Statistics*. 2001;**29**(1):153-193. DOI: 10.1214/aos/996986505
- [27] Fan J, Jiang J. Nonparametric inference for additive models. *Journal of the American Statistical Association*. 2005;**100**:890-907

IntechOpen