Developing a Strategy for Buying and Selling Stocks Based on Semi-Parametric Markov Switching Time Series Models

Hossein Naderi

Ph.D. Candidate, Department of Accounting, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. (Email: hossein.acc@gmail.com)

Mehrdad Ghanbari*

*Corresponding Author, Assistant Prof., Department of Accounting, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. (Email: ghanbari@iauksh.ac.ir)

Babak Jamshidi Navid

Assistant Prof., Department of Accounting, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. (Email: jamshidinavid@iauksh.ac.ir)

Arash Nademi

Assistant Prof., Department of Statistics, Ilam Branch, Islamic Azad University, Ilam, Iran. (Email: Arash_nademi@ilam-iau.ac.ir)

Abstract

The modeling of strategies for buying and selling in Stock Market Investment has been the object of numerous advances and uses in economic studies, both theoretically and empirically. One of the popular models in economic studies is applying the Markov Switching models for forecasting the time series observations based on stock prices. The semi-parametric estimators for these models are a class of popular methods that have been used extensively by researchers to increase the accuracy of estimation. The main part of these estimators is based on kernel functions. Despite the existence of many kernel

functions that are capable in applications for forecasting the stock prices, there is a widely use of Gaussian kernel in these estimators. But there is a question if other types of kernel function can be used in these estimators. This paper tries to introduce the other kernel functions that can be a good replacement for this kernel function to increase the ability of Markov Switching models. We first test six popular kernel functions to find the best one based on simulation studies and then offer the new strategy of buying and selling stocks by the best kernel function selection on real data.

.

Introduction

The many investigations in economic and financial mathematics focused on what makes an investor profitable in the stock market. These studies can aid the researchers to decrease the investment risks and increase opportunities for a high return of gaining. One of the important questions in the stock market is when the investors can buy the stocks and when they can sell their stocks. In research economic papers, there are two aspects of analysis: fundamental and technical analysis. In fundamental aspects, the researchers find the reasons for changing stock prices, in response to reasons of changing prices caused by exogenous geopolitical events, supply disruptions or financial operation of the companies, etc. But technical analysis noted more the statistical and probabilities rules governed by processes of the data. In an aspect of technical analysis, there are a lot of models in time series to capture the stock prices. The Markov Switching models are the popular models in time series that are applied most widely in financial and economic data. These models exhibit abrupt changes in the behavior of time series data, called switches of regimes, where the switching between the regimes is controlled by a hidden Markov Chain process. (See Chang, Yongok, & Joon 2017; Von Ganske, 2016; Billio, Casarin, Ravazzolo, & Van Dijk, 2016; Di Persio, & Frigo, 2016; Neale, Clark, Dolan, & Hunter, 2016; Nademi, & Nademi, 2019).

Recently, Markov switching AR–ARCH[1] models have repeatedly applied for making switching regimes processes and every one of them offers an algorithm for estimating the parameters. The finding of the best algorithms for parameters estimation of these models has been the object of many expansions and usages over the last decade. Numerous algorithms based on parametric and nonparametric methods have been offered to capture the parameters of the modeling process. In this respect, the combination of parametric and nonparametric methods, called semi-parametric algorithms, are popular and most broadly applied. (See Chan, & Wang, 2017; Chang, Tang, & Wu, 2016; Chen, Shen, Wei, & Lin, 2017; Gupta, Cobre, Polpo, & Sinha, 2016; Gu, Ma, & Balasubramanian, 2015; Nademi, & Farnoosh, 2014).

In a semi-parametric class of algorithms, a special function, called kernel function, is used. The selection of proper kernel function is an important item for estimating the parameters. Such that, if we apply the proper type, we can have a fast and unbiased estimating process. So, offering the best kernel functions for estimation algorithms can be essential for the modelling process. In this paper, we first focus on selecting the best kernel function in a special class of Markov switching models called semi-parametric Markov switching offered by Nademi & Farnoosh (2014) for modeling the time series data and then offer the new strategy of buying and selling stocks by the best-selected kernel function of this model on real data.

In the next section, the model and its algorithm will be introduced. Section 3 discuss the best selection of kernel function by a simulation study and offer the buying and selling strategy by comparing the different kernel functions and the feasibility of these kernels will be shown. Finally, section 4 conclude and introduce the opportunities for future studies.

The model and EM algorithm

This section consists of two subsections. In the first subsection, we introduce the Markov switching model introduced by Nademi & Farnoosh (2014) and in the second subsection, their algorithm for estimating the parameters will be reviewed. Note that, their semi-parametric algorithm is a part of a more general algorithm as EM algorithm that applies to the class of Markov switching models.

---

[1] Autoregressive–Autoregressive Conditionally Heteroscedasticity models.

1. The model

Suppose $Y_1, \ldots, Y_N$ are part of a strictly stationary time series that are generated by the following semi-parametric switching model

$$Y_t = \sum_{k=1}^{M} Z_{t,k}\big(\mu(Y_{t-1}, Y_{t-2}; \rho_k) + \sigma(Y_{t-1}, Y_{t-2}; \omega_k, \alpha_k, \beta_k)u_{t,k}\big), \qquad (1)$$

such that,

$$\mu(Y_{t-1}, Y_{t-2}; \rho_k) = f_k(Y_{t-1}) + \rho_k\big(Y_{t-1} - f_k(Y_{t-2})\big), \qquad (2)$$

and

$$\sigma^2(Y_{t-1}, Y_{t-2}; \omega_k, \alpha_k, \beta_k) = \omega_k + \alpha_k Y_{t-1}^2 + \beta_k Y_{t-2}^2,$$

with

$$Z_{tk} = \begin{cases} 1 & Q_t = k \\ 0 & otherwise, \end{cases}$$

where switching between the regimes is controlled by a hidden Markov chain Q_t, with values in $\{1, \ldots, M\}$, and the residuals $u_{t,k}, t = 1, \ldots, N, k = 1, \ldots, M$ are i.i.d. random variables with mean 0 and variance 1. $Z_t = (Z_{t1}, Z_{t2}, \ldots, Z_{tM})^T$ are random variables which assume as values of the unit vectors $e_1, e_2, \ldots, e_M \in R^M$, i.e. exactly one of the Z_{tk} is 1, and the others are 0. The stationary distribution of the hidden regime process is given by the $M \times M$ transition probability matrix A, i.e, $A_{jk} = pr(Q_t = k | Q_{t-1} = j)$. We get the stationary probabilities $\pi_k = pr(Z_t = e_k) = pr(Q_t = k)$ by equation $\boldsymbol{\pi A = \pi}$.

$f_k(.)$ is considered as a semi-parametric function such that $f_k(x) \in \{g(x, \theta_k)\xi_k(x); \theta_k \in V\}$, where $\xi_k(x)$ is a nonparametric adjustment factor and $g(x, \theta_k)$ is a known function of x and θ_k and $V \subseteq R^p$ is the parametric space. So, we can refer to mean function (2) and rewrite it by the following form:

$$\mu(Y_{t-1}, Y_{t-2}; \theta_k, \rho_k) = g(Y_{t-1}, \theta_k)\xi_k(Y_{t-1}) + \rho_k\big(Y_{t-1} - g(Y_{t-2}, \theta_k)\xi_k(Y_{t-2})\big).$$

2. The EM algorithm based on the semi-parametric method

Supposing the definition of $Y^{(N)} = (Y_1, Y_2, \ldots, Y_N)$ and $Z^{(N)} = (Z_1, Z_2, \ldots, Z_N)$, Nademi & Farnoosh (2014) applied a special class of log-likelihood function, called "complete" log-likelihood function, by the following form

$$l_c\left(v, A \mid Y^{(N)}, Z^{(N)}\right)$$

$$= log\pi_{q_1} \sum_{t=2}^{N} logA_{q_{t-1}, q_t}$$

$$+ \sum_{t=2}^{N} \sum_{k=1}^{M} Z_{tk} log \frac{1}{\sigma(Y_{t-1}, Y_{t-2}; \omega_k, \alpha_k, \beta_k)} \varphi\left(\frac{Y_t - \mu(Y_{t-1}, Y_{t-2}; \theta_k, \rho_k)}{\sigma(Y_{t-1}, Y_{t-2}; \omega_k, \alpha_k, \beta_k)}\right)$$

Where $v = (\theta_1, \dots, \theta_M, \rho_1, \dots, \rho_M, \omega_1, \dots, \omega_M, \alpha_1, \dots, \alpha_M, \beta_1, \dots, \beta_M)^T \in V$ and $\varphi(.)$ is the normal density with mean $\mu(.)$ and standard deviation $\sigma(.)$ The word "complete" refers to this definition that if we would have observed the complete data $\left(Y^{(N)}, Z^{(N)}\right)$, instead of just $Y^{(N)}$, we could maximize the complete data log-likelihood (see Franke et al. [8]), instead of the ordinary log-likelihood.

Applying this method leads to the use of the Expectation and Maximization algorithm known as the EM algorithm. The EM algorithm repeats between drawing the unseen variables Z_{tk} by conditional expectations Z_{tk} given the seen data $Y^{(N)}$ and using an elementary estimate of the parameters on the one phase (E-step), and by maximizing $l_c\left(v, A \mid Y^{(N)}, Z^{(N)}\right)$ to get an update of approximations of A or v on the other phase (M-step). These two phases until assuring certain stopping criteria are iterated. The algorithm offered with the EM algorithm can be summarized in the following steps.

E-step: Suppose $\hat{\pi}_1, \dots, \hat{\pi}_M, \hat{\rho}_1, \dots, \hat{\rho}_M, \hat{\theta}_1, \dots, \hat{\theta}_M, \hat{\omega}_1, \dots, \hat{\omega}_M, \hat{\alpha}_1, \dots, \hat{\alpha}_M$ and $, \hat{\beta}_1, \dots, \hat{\beta}_M$ are given. S0, the conditional expectation of the unseen variables Z_{tk} given $Y^{(N)}$ are calculated by

$$C_{tk} = E[Z_{tk} \mid Y^{(N)}] = \frac{\alpha_k^t \beta_k^t}{\sum_{i=1}^{M} \alpha_i^t \beta_i^t} \quad k = 1, \dots, M \quad t = 1, \dots, N,$$

where α_i^t and β_i^t are estimated by following recursive relations

$$\alpha_j^{t+1} = \varphi\left(Y_{t+1}; \mu(Y_t, Y_{t-1}; \theta_j, \rho_j), \sigma(Y_t, Y_{t-1}; \omega_j, \alpha_j, \beta_j)\right) \sum_{k=1}^{M} A_{kj} \alpha_k^t,$$

and

$$\beta_j^t = \sum_{k=1}^{M} \beta_k^{t+1} \varphi\big(Y_{t+1}; \mu(Y_t, Y_{t-1}; \theta_k, \rho_k), \sigma(Y_t, Y_{t-1}; \omega_k, \alpha_k, \beta_k)\big)A_{jk},$$

where $\varphi\big(Y_{t+1}; \mu(Y_t, Y_{t-1}; \theta_k, \rho_k), \sigma(Y_t, Y_{t-1}; \omega_k, \alpha_k, \beta_k)\big)$ is the normal density with mean $\mu(.)$ And standard deviation $\sigma(.)$.

M-step: Suppose the approximations C_{tk} for the unseen variables Z_{tk} are given. Then, the transition probabilities are calculated by

$$\hat{A}_{ij} = \frac{\sum_{t=1}^{N} \delta_{ij}^{t,t+1}}{\sum_{t=1}^{N} C_{ti}},$$

where $\delta_{ij}^{t,t+1}$ are the joint conditional probability of $Q_t = i$ and $Q_{t+1} = j$ given the entire sequence of observations $(Y^{(N)})$ estimated by

$$\delta_{ij}^{t,t+1} = p\big(Q_t = i, Q_{t+1} = j \big| Y^{(N)}\big)$$
$$= \frac{\beta_j^{t+1} \varphi\left(Y_{t+1}; \mu(Y_t, Y_{t-1}; \theta_j, \rho_j), \sigma(Y_t, Y_{t-1}; \omega_j, \alpha_j, \beta_j)\right) A_{ij} \alpha_i^t}{\sum_{k=1}^{M} \alpha_k^t \beta_k^t}.$$

The probabilities $\pi_1, \dots, \pi_M$ are approximated by

$$\widehat{\pi_k} = \frac{1}{N}\sum_{t=1}^{N} C_{tk}, \quad k = 1, \dots, M,$$

and the functions $\mu(x, y; \theta_k, \rho_k)$ are estimated by

$$\mu\big(x, y; \hat{\theta}_k, \hat{\rho}_k\big) = g\big(x, \hat{\theta}_k\big)\hat{\xi}_k(x) + \hat{\rho}_k\left(x - g(y, \hat{\theta}_k)\hat{\xi}_k(y)\right),$$

such that, $(\hat{\theta}_k, \hat{\rho}_k)$ gets from $\big(\hat{\theta}_k, \hat{\rho}_k\big) = argminQ_n(\theta_k, \rho_k), \theta_k, \rho_k \in V, |\rho_k| < 1$ for $k = 1, \dots, M$, where $Q_n(\theta_k, \rho_k)$ is

$$Q_n(\theta_k, \rho_k) = \sum_{t=2}^{N} C_{tk}(Y_t - g(Y_{t-1}, \theta_k) - \rho_k(Y_{t-1} - g(Y_{t-2}, \theta_k)))^2,$$

and $\hat{\xi}_k(x)$ is

$$\hat{\xi}_k(x)$$
$$= \frac{\sum_{t=2}^{N} C_{tk}\left[k\left(\frac{Y_{t-1} - x}{h_k}\right) g(Y_{t-1}, \hat{\theta}_k)Y_t + k\left(\frac{Y_{t-2} - x}{h_k}\right) g(Y_{t-2}, \hat{\theta}_k)Y_{t-1}\right]}{\sum_{t=2}^{N} C_{tk}\left[k\left(\frac{Y_{t-1} - x}{h_k}\right) g^2(Y_{t-1}, \hat{\theta}_k) + k\left(\frac{Y_{t-2} - x}{h_k}\right) g^2(Y_{t-2}, \hat{\theta}_k)\right]}, \quad (3)$$

where $k(.)$ is a Gaussian Kernel function and $(\omega_k, \alpha_k, \beta_k)$ are estimated by $(\hat{\omega}_k, \hat{\alpha}_k, \hat{\beta}_k)$

$$= Arg\,max \sum_{t=2}^{N} C_{tk}\,log\,\frac{1}{\sigma(Y_{t-1}, Y_{t-2}; \omega_k, \alpha_k, \beta_k)}\,\varphi\left(\frac{\hat{e}_{tk}}{\sigma(Y_{t-1}, Y_{t-2}; \omega_k, \alpha_k, \beta_k)}\right),$$

for $k = 1, \dots, M$, where $\hat{e}_{tk} = Y_t - \mu(x, y; \hat{\theta}_k, \hat{\rho}_k)$ denote the sample residuals. The optimal selection of the bandwidth h_k are also estimated by

$$\hat{h}_k = arg\,max_{h_k} \sum_{t=2}^{N} C_{tk}\left[Y_t - g(Y_{t-1}, \hat{\theta}_k)\hat{\xi}_k(Y_{t-1}) - \hat{\rho}_k\left(Y_{t-1} - g(Y_{t-2}, \hat{\theta}_k)\hat{\xi}_k(Y_{t-2})\right)\right]^2.$$

The estimate of the parameters are obtained by iterating these two steps (E-step and M-step) until convergence.

In relation (3), they applied the Gaussian Kernel function for estimating $\hat{\xi}_k(x)$. But there is a question that if other types of kernel function can improve performance of the semi-parametric algorithm. By the definition of kernel function, a function $k(h)$, with compact support, is a kernel function satisfying

$$k(u) \geq 0, k(-u) = k(u), \int k(u)du = 1.$$

So, by this definition, we can find other functions that have these conditions. We want to try some other types of kernel functions that are popular in the mathematics field. Table 1 shows several types of kernel functions that are commonly used. These functions consist of Uniform, Triangle, Epanechnikov, Bisquare and Triweight. We also apply the Gaussian kernel to compare this function with the candidate kernel functions. Figure 1 shows the plots of these functions. We apply these functions to compare their ability to the improvement of the EM algorithm.

Table 1. The popular kernel functions.

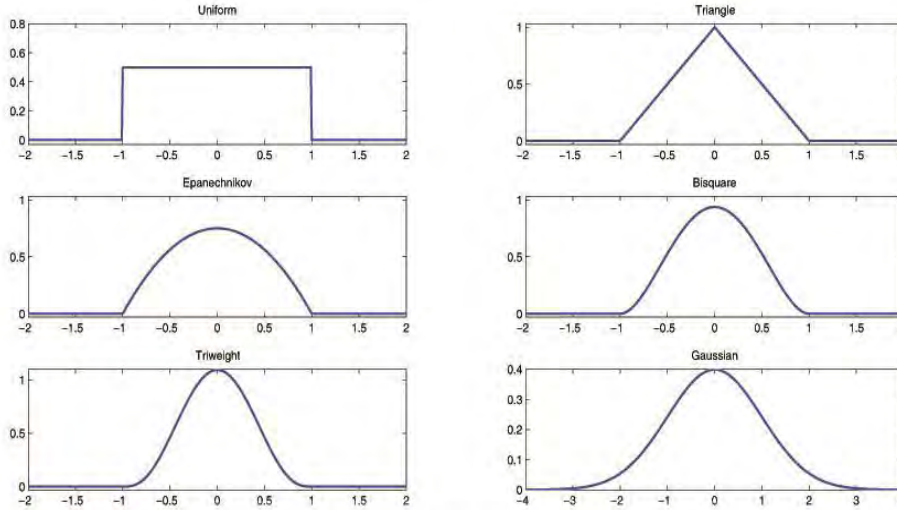| Name	Kernel
Uniform	$\frac{1}{2}I_{[-1,1]}(u)$
Triangle	$(1 - |u|)I_{[-1,1]}(u)$
Epanechnikov	$\frac{3}{4}(1 - u^2)I_{[-1,1]}(u)$
Bisquare	$\frac{15}{16}(1 - 2u^2 + u^4)I_{[-1,1]}(u)$
Triweight	$\frac{35}{32}(1 - 3u^2 + 3u^4 - u^6)I_{[-1,1]}(u)$
Gaussian	$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}u^2\right)$

Figure 1. The six popular Kernel functions

Research Background

In this section, we first carry out a simulation study to examine the finite sample performance of the proposed kernel functions, and then the Semi-parametric Markov Switching model with the selected kernel function will apply to financial observations, including the Automotive Industry Index of the Tehran Stock market data, to find the strategy of buying and selling in the stocks of industrial companies. Because there is a high correlation between the prices of stocks for automotive companies and the Automotive Industry Index, we focus on the Automotive Industry Index data to offer the strategy of buying and selling.

1. Simulation study

We intend to examine the feasibility of various kernel functions by estimating parameters in the semi-parametric model (1). We generated two sets of data with sample size N=500 based on nonlinear functions $f_k^{(1)}(x)$ and $f_k^{(2)}(x)$ by the following forms:

$$f_k^{(1)}(x) = \theta_k x \exp(\theta_k x), k = 1,2$$

$$f_k^{(2)}(x) = \sin(\theta_k x), k = 1,2,$$

and it was supposed that $M = 2, \rho_1 = 0.4, \rho_2 = 0.6, \theta_1 = -2, \theta_2 = -6, h_1 = 0.02$ and $h_2 = 0.03$. The transition probability matrix was considered by

$$A = \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}.$$

We also chose the variance parameters as $\omega_1 = 0.001, \alpha_1 = 0.01, \beta_1 = 0.01, \omega_2 = 0.004, \alpha_2 = 0.03$ and $\beta_2 = 0.05$. Figures 2(a) and (d) show the generated observations based on $f_k^{(1)}(x)$ and $f_k^{(2)}(x)$ respectively. Figure 1(b, c) and (e, f) show the corresponding scatter plot of Y_t with Y_{t-1} and Y_{t-2} for two simulated data sets. These plots also indicate the degree of dependency of the data is 2. Such that, this degree of dependency is because of the structure of the model (1).

For comparing the models, we apply two indices the square Root of Mean Squared Error (RMSE) by the following form:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (Y_t - \hat{Y}_t)^2},$$

and classifying index "$\text{Max}C_{tk}, k = 1, \dots, M$". The classifying index "$\text{Max}C_{tk}, k = 1, \dots, M$" is defined by the following definition:

"Y_t is belonging to regime k if and only if $C_{tk} = \text{Max}_{i=1,\dots,M} C_{ti}$."

This index is suitable for the evaluation of the models, such that the proper model is powerful in classifying the observations in the right regimes.

Table 2 shows the estimated parameters of the simulated model based on $f_k^{(1)}(x)$. Comparing the measures of RMSE for the six kernel functions, we can find that the Triangle kernel with RMSE 0.0710 is more efficient than the others kernel functions. After that Triweight kernel with RMSE 0.0781 can be selected as the proper kernel. Figure 3 shows $\text{Max}(C_{t1}, C_{t2})$ for six kernel functions in which for Triangle kernel, except for a few cases, almost all are at least greater than 0.8, i.e. there is a clear decision for one of the two phases in the large majority of cases and we find that high percentage of the data is correctly classified. With making change the nonlinear function $f_k^{(1)}(x)$ by $f_k^{(2)}(x)$ for the second simulated data, we founded the different order of the kernel functions (table 3). Such that, the Uniform kernel function (RMSE=0.0521) had the best performance among other kernel functions. Also, after the Uniform kernel function, The Gaussian kernel with RMSE 0.0584 is more proper than the others. Figure 4 shows $\text{Max}(C_{t1}, C_{t2})$ for the second simulation data set which indicates the ability of the Uniform kernel in classifying the data comparing the others. These results indicate the different

efficiency of kernel functions in various problems. This demonstrates in applying the semi-parametric algorithms based on selecting kernel functions, trying different kernel functions can help to get proper estimating.

Table 2. The estimated parameters for the simulated data based on $f_k^{(1)}(x)$.

| The Parameters	Uniform	Triangle	Epanechnikov	Bisquare	Triweight	Gaussian
θ_1	-3.0148	-2.0212	-2.9824	-2.3684	-2.0084	-2.0461
θ_2	-4.9842	-5.9810	-5.2617	-6.8410	-6.0316	-4.1586
ρ_1	0.3245	0.4215	0.3874	0.3841	0.3848	0.4167
ρ_2	0.5120	0.5984	0.5361	0.4835	0.5549	0.5684
ω_1	0.0104	0.0012	0.0022	0.0101	0.0010	0.0025
ω_2	0.0032	0.0039	0.0024	0.0213	0.0026	0.0038
α_1	0.0201	0.0101	0.0191	0.0318	0.0013	0.0198
α_2	0.0215	0.0312	0.0234	0.0227	0.0301	0.0279
β_1	0.0110	0.0101	0.0150	0.0012	0.0027	0.0012
β_2	0.0318	0.0491	0.0284	0.0046	0.0394	0.0394
π_1	0.4462	0.4374	0.3976	0.4284	0.4504	0.4598
π_2	0.5538	0.5626	0.6024	0.5716	0.5496	0.5402
A_{12}	0.5107	0.4021	0.4462	0.4872	0.4138	0.4463
A_{21}	0.4115	0.3126	0.2945	0.3651	0.3391	0.3798
h_1	0.0262	0.0215	0.0371	0.0315	0.0259	0.0297
h_2	0.0334	0.0316	0.0502	0.0078	0.0298	0.0467
$RMSE$	0.0889	0.0710	0.0885	0.0881	0.0781	0.0789
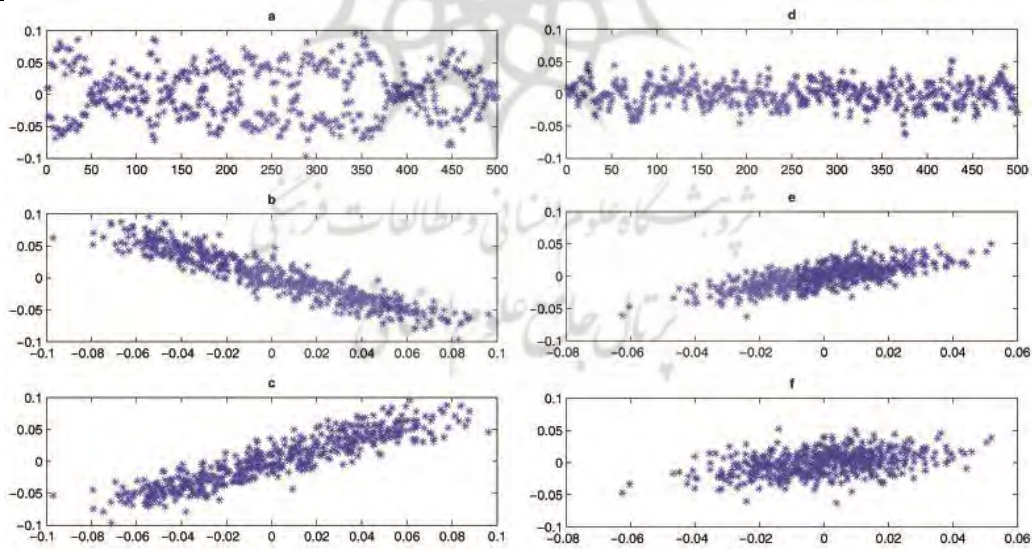
Figure 2. For nonlinear function $f_k^{(1)}(x)$: {(a). Simulated data, (b). Scater plot of (Y_t, Y_{t-1}), (c). Scater plot of (Y_t, Y_{t-2})}, For nonlinear function $f_k^{(2)}(x)$: {(d). Simulated data, (e). Scater plot of (Y_t, Y_{t-1}), (f). Scater plot of (Y_t, Y_{t-2})}.

Table 3. The estimated parameters for the simulated data based on $f_k^{(2)}(x)$.

| The Parameters	Uniform	Triangle	Epanechnikov	Bisquare	Triweight	Gaussian
θ_1	-2.1480	-3.5942	-3.4618	-4.1128	-2.9916	-2.1384
θ_2	-5.9257	-4.3302	-4.5280	-5.2648	-3.5218	-6.7681
ρ_1	0.4635	0.6891	0.6457	0.5561	0.4691	0.4954
ρ_2	0.5549	0.4697	0.4630	0.6021	0.3559	0.6894
ω_1	0.0017	0.0052	0.0036	0.0021	0.0649	0.0016
ω_2	0.0029	0.0063	0.0074	0.0034	0.0529	0.0048
α_1	0.0138	0.0627	0.0108	0.0251	0.0024	0.0113
α_2	0.0204	0.0104	0.0371	0.0319	0.0031	0.0287
β_1	0.0157	0.0264	0.0349	0.0108	0.0013	0.0015
β_2	0.0313	0.0129	0.0137	0.0008	0.0062	0.0292
π_1	0.4462	0.4374	0.3976	0.4284	0.4504	0.4598
π_2	0.5538	0.5626	0.6024	0.5716	0.5496	0.5402
A_{12}	0.4410	0.5649	0.5410	0.4026	0.4952	0.4137
A_{21}	0.3217	0.2149	0.3619	0.2237	0.2237	0.4679
h_1	0.0149	0.0338	0.0246	0.0213	0.0150	0.0346
h_2	0.0226	0.0108	0.0117	0.0346	0.0357	0.0243
$RMSE$	0.0521	0.0621	0.0602	0.0703	0.0648	0.0584
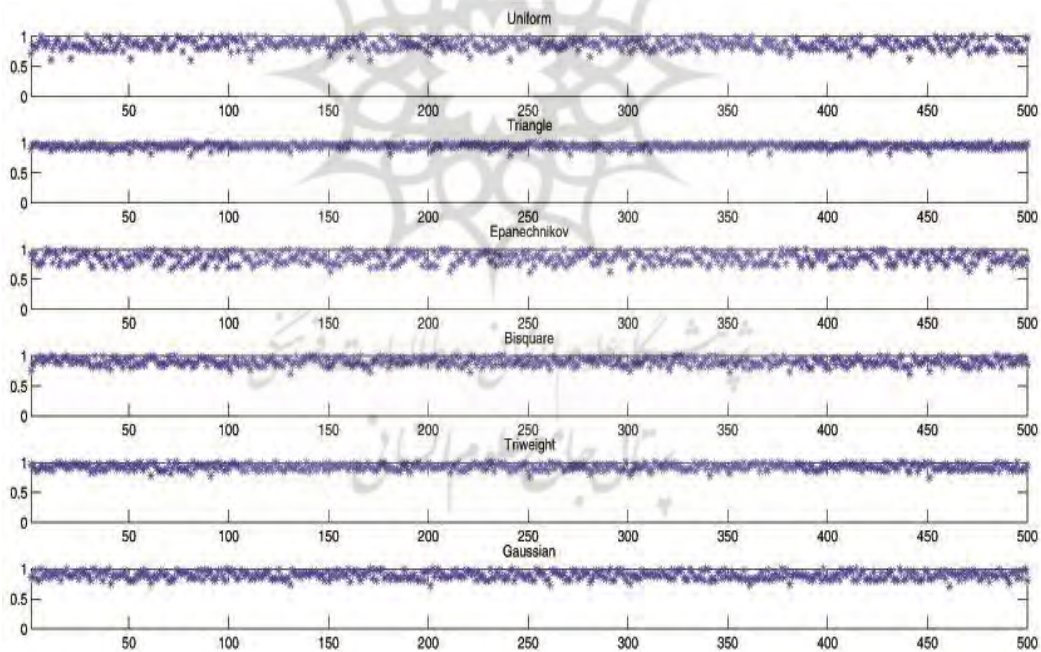
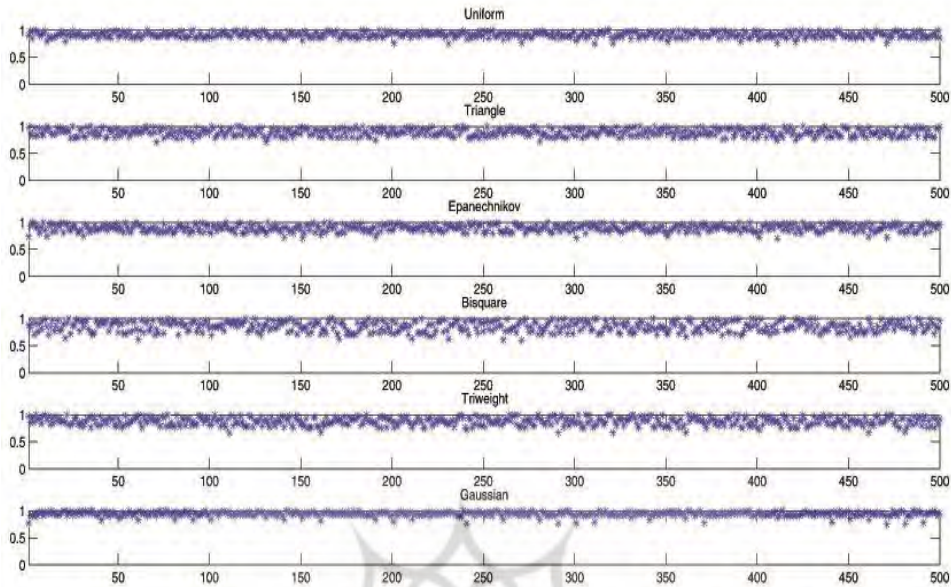Figure 3. $Max(C_{t1}, C_{t2})$ for six kernel functions based on $f_k^{(1)}(x)$.

Figure 4. Max(C_{t1}, C_{t2}) for six kernel functions based on $f_k^{(2)}(x)$.

Results

We consider a data set of Automotive Industry Index of Tehran Stock market for the period March 25, 2018, to May 3, 2021, downloaded from "http://tse.ir/archive.html", where the sample size is 744. In the first step, we must determine the number of regimes in observations. This can determine by drawing the sample path of data and observing changes trends as increasing and decreasing function or by classifying the data in two classes of high and low volatility (see Nademi, & Farnoosh, 2014; Nademi, 2019). For sake of simplicity and good showing the regimes, we draw 200 observations of the data set. Figure 5 (blue line) shows the sample path of the data. According to this plot, we applied the step function (red line), the down step (regime=1) and the upper step (regime=2), to indicate the regimes such that we named increasing trends and decreasing trends by regimes =1 and regimes=2, respectively. So, we considered six Semi-parametric Markov Switching models (called MS-SEMI-K (i), i=1,…,6) based on six kernel functions and two regimes (M=2).
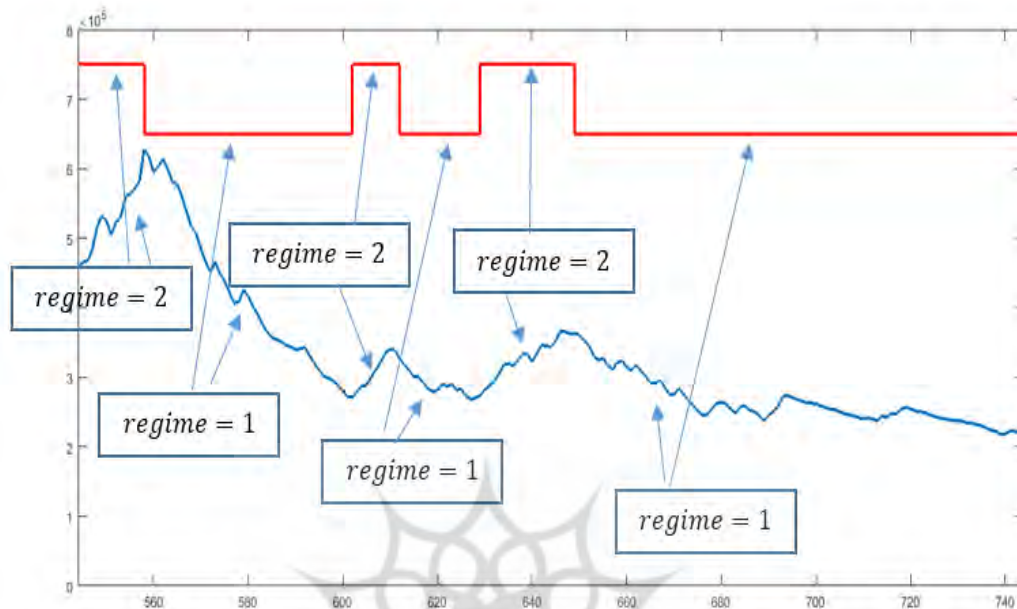
Figure 5. Automotive Industry Index data and the regimes.

In the second step, we define Y_t as the stationary automotive industry index data in time t (input variables) and apply the semiparametric model (1) to fit the observations in the following relation:

$$Y_t = \sum_{k=1}^{M} Z_{t,k}\big(\mu(Y_{t-1}, Y_{t-2}; \rho_k) + \sigma(Y_{t-1}, Y_{t-2}; \omega_k, \alpha_k, \beta_k)u_{t,k}\big),$$

where $Z_{t,k}$, $\rho_k, \mu(Y_{t-1}, Y_{t-2}; \rho_k), \sigma(Y_{t-1}, Y_{t-2}; \omega_k, \alpha_k, \beta_k)$ and $u_{t,k}$ are defined by section 2.1. In the third step, we apply the EM algorithm described in section 2.2 to estimate the parameters of the model. The EM algorithm is a numeric procedure and start by initial parameters and then iterating the E-step and M-step until the convergence of parameters. Note that, in one part of M-step, we must select a kernel function to estimate $k\left(\frac{Y_{t-1}-x}{h_k}\right)$ and $k\left(\frac{Y_{t-2}-x}{h_k}\right)$ in function $\hat{\xi}_k(x)$ (relation 3). So, we trial six kernel functions Gaussian, Uniform, Triangle, Epanechnikov, Bisquare and Triweight to get a proper result. This proper result can compare by the index RMSE. Such that, the best model has minimum RMSE. We also compare the ability of models in classifying the observation to regimes by the index of $Max(C_{tk})$. Finally, in

the fourth step, the proper model will be selected for creating the strategy of buying and selling stocks by applying the estimated joint conditional probabilities $\delta_{ij}^{t,t+1}$ defined in M-step in section 2.2.

Table 4 lists the estimated parameters for six different specifications of the semi-parametric Markov switching models in the Automotive Industry Index. The results of RMSE criteria for the forecasting indicate that in all of the six different specifications of the semi-parametric Markov switching models, the MS-SEMI-K(1) model is more accurate for the prediction of the Automotive Industry Index than other semi-parametric Markov switching models. After this model, the MS-SEMI-K(6), MS-SEMI-K(2), MS-SEMI-K(5), MS-SEMI-K(4) and MS-SEMI-K(3) models have more accurate forecasting for Automotive Industry Index, respectively. Therefore, the proper kernel function for forecasting the data is Uniform. Figure 6 shows the transition probabilities based on index $Max(C_{tk})$, The values of $Max(C_{t1}, C_{t2})$ for six models, which are all greater than 0.5, show the ability of semi-parametric models in classifying the data. On the other hand, in the MS-SEMI-K (1) model the belonging probabilities greater than 0.73 indicate that this model is more powerful than the other models in classifying the observations.

Table 4. The estimated parameters for the Automotive Industry Index data

| The Parameters	Uniform	Triangle	Epanechnikov	Bisquare	Triweight	Gaussian
θ_1	-0.0158	-0.0269	-2.2139	-0.3814	-0.4283	-0.1123
θ_2	-1.3541	-1.2285	-1.3025	-5.1281	-3.2891	-1.0293
ρ_1	0.4128	0.4838	0.4123	0.5528	0.7138	0.5112
ρ_2	0.3984	0.3564	0.8236	0.6349	0.6318	0.4631
ω_1	0.0001	0.0004	0.0102	0.0021	0.0124	0.0049
ω_2	0.0002	0.0001	0.0219	0.0038	0.0104	0.0012
α_1	0.0031	0.0026	0.0113	0.0394	0.0129	0.0028
α_2	0.0101	0.0098	0.0518	0.0354	0.0163	0.0037
β_1	0.0128	0.0109	0.0321	0.0194	0.0142	0.0020
β_2	0.0142	0.0113	0.0584	0.0101	0.0028	0.0721
π_1	0.2665	0.1771	0.5079	0.4285	0.2158	0.3491
π_2	0.7335	0.8229	0.4921	0.5715	0.7842	0.6509
A_{12}	0.5214	0.5329	0.5023	0.4358	0.4138	0.5563
A_{21}	0.1894	0.1147	0.5184	0.3268	0.1139	0.2984
h_1	0.0001	0.0002	0.0049	0.0059	0.0091	0.0056
h_2	0.0003	0.0001	0.0108	0.0018	0.0065	0.0061
$RMSE$	0.0520	0.0692	0.0918	0.0925	0.0711	0.0601
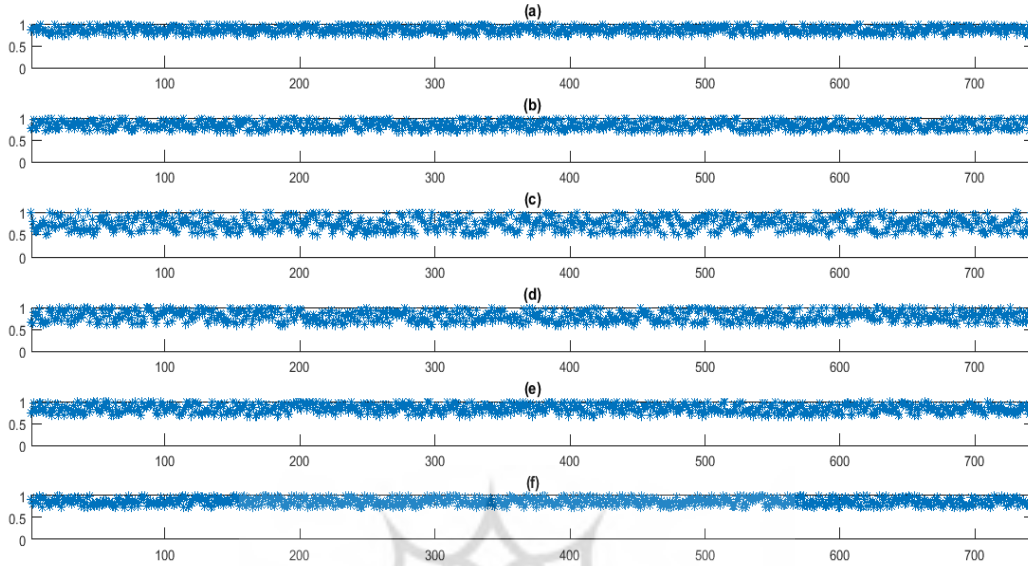
Figure 6. Max(C_{t1}, C_{t2}) for (a). MS-SEMI-K (1) model, (b). MS-SEMI-K (2) model, (c). MS-SEMI-K (3) model, (d). MS-SEMI-K (4) model, (e). MS-SEMI-K (5) model, (f). MS-SEMI-K (6) model.

Now that, the appropriate model has been identified (MS-SEMI-K (1) model), we apply the estimated joint conditional probabilities to introduce the strategy of buying and selling stocks. For the sake of simplicity in showing the estimated joint conditional probabilities in the graphs, we chose 5 observations (T=739 to T=744) at the end of the Automotive Industry Index, including April 29, 2021, to May 3, 2021. Table 5 lists the estimated joint conditional probabilities based on $\delta_{ij}^{t,t+1} = p(Q_t = i, Q_{t+1} = j|Y^{(N)})$, such that, we can write the joint conditional probability matrix

$$\delta_{ij}^T = \begin{pmatrix} p(Q_t = 1, Q_{t+1} = 1|Y^{(N)}) & p(Q_t = 1, Q_{t+1} = 2|Y^{(N)}) \\ p(Q_t = 2, Q_{t+1} = 1|Y^{(N)}) & p(Q_t = 2, Q_{t+1} = 2|Y^{(N)}) \end{pmatrix},$$

For the best model (MS-SEMI-K(1) model). This matrix can offer the strategy of buying and selling stock in financial markets. Such that, the probability elements of the matrix indicate the behavior of the data in passing time "t" to "t+1". Note that, we define regime 1 as decreasing trend and regime 2 as an increasing trend.

Table 5. The estimated joint conditional probability matrix based on the six models

The model	Time period				
	t=739, t+1=740	t=740, t+1=741	t=741, t+1=742	t=742, t+1=743	t=743, t+1=744
MS-SEMI-K(1)	$\begin{pmatrix} 0.1184 & \mathbf{0.53} \\ 0.0156 & 0.33 \end{pmatrix}$	$\begin{pmatrix} 0.1647 & \mathbf{0.55} \\ 0.0423 & 0.23 \end{pmatrix}$	$\begin{pmatrix} 0.0112 & 0.26 \\ 0.2457 & \mathbf{0.48} \end{pmatrix}$	$\begin{pmatrix} 0.3676 & 0.01 \\ \mathbf{0.5321} & 0.08 \end{pmatrix}$	$\begin{pmatrix} \mathbf{0.5549} & 0.01 \\ 0.4172 & 0.01 \end{pmatrix}$

Table 5 shows the estimated joint conditional probability matrix for the observations of the Automotive Industry Index for the selected time period based on the MS-SEMI-K(1) model. According to the result of this table for the best-selected model (MS-SEMI-K(1)), we can see the maximum probabilities among the elements of the matrices $\delta_{ij}^{t,t+1}$ for period (t=739, t+1=740) and (t=740, t+1=741) are $\mathbf{0.5318}$ and $\mathbf{0.5581}$, respectively, that are belong to the switching between the regimes from 1 to 2 ($Q_t = 1$, $Q_{t+1} = 2$). So, we have an increasing trend. This offers that the strategy of buying the stocks in a period of time t=739.

t=739, t+1=740	t=740, t+1=741
$\begin{pmatrix} 0.1184 & \mathbf{0.5318} \\ 0.0156 & 0.3342 \end{pmatrix}$	$\begin{pmatrix} 0.1647 & \mathbf{0.5581} \\ 0.0423 & 0.2349 \end{pmatrix}$

On the other hand, the results indicate that the maximum probabilities among the elements of the matrices $\delta_{ij}^{t,t+1}$ for period (t=741, t+1=742) are $\mathbf{0.4821}$. This shows the switching between the regimes is from 2 to 2 ($Q_t = 2$, $Q_{t+1} = 2$) that indicates the stock may go to a decreasing trend that offers the strategy of selling the stocks in a period of time t=741 to t= 742.

t=741, t+1=742
$\begin{pmatrix} 0.0112 & 0.2610 \\ 0.2457 & \mathbf{0.4821} \end{pmatrix}$

Also, the maximum probabilities of matrices $\delta_{ij}^{t,t+1}$ for period (t=742, t+1=743) and (t=743, t+1=744) are $\mathbf{0.5321}$ and $\mathbf{0.5549}$, respectively, that are belong to the switching between the regimes from 2 to 1 ($Q_t = 2, Q_{t+1} = 1$) and 1to 1 ($Q_t = 2, Q_{t+1} = 1$). So, we have a decreasing trend. This offers that the strategy of buying the stocks at the end of a period of t= 744.

t=742, t+1=743	t=743, t+1=744
$\begin{pmatrix} 0.3676 & 0.0132 \\ \mathbf{0.5321} & 0.0871 \end{pmatrix}$	$\begin{pmatrix} \mathbf{0.5549} & 0.0173 \\ 0.4172 & 0.0106 \end{pmatrix}$

Figure 7 demonstrates the estimated elements of the joint conditional probability matrix for the best-selected model (MS-SEMI-K (1)). Figure 7(a) shows the sample path of the Automotive Industry Index, including April 29, 2021, to May 3, 2021, which shows an increasing trend until lag 3 and decreasing trend to lag 6. Figure 7(b) indicates the estimated elements of the joint conditional probability matrix. Such that, the red line shows the switching between regimes from decreasing to increasing trend ($Q_t = 1$, $Q_{t+1} = 2$) that was explained for a period (t=739, t+1=740) and (t=740, t+1=741). The blue line indicates the switching between regimes from an increasing trend to increasing one ($Q_t = 2$, $Q_{t+1} = 2$) or staying in an increasing trend for two lags (t=741, t+1=742). The green line indicates the switching between regimes from increasing to decreasing trend ($Q_t = 2$, $Q_{t+1} = 1$) that was explained for a period (t=742, t+1=743) and finally, the yellow line demonstrates the switching between regimes from a decreasing trend to decreasing one ($Q_t = 1$, $Q_{t+1} = 1$) or staying in a decreasing trend that was explained for a period (t=743, t+1=744). We can find the maximum probability in every of lags (1-6).
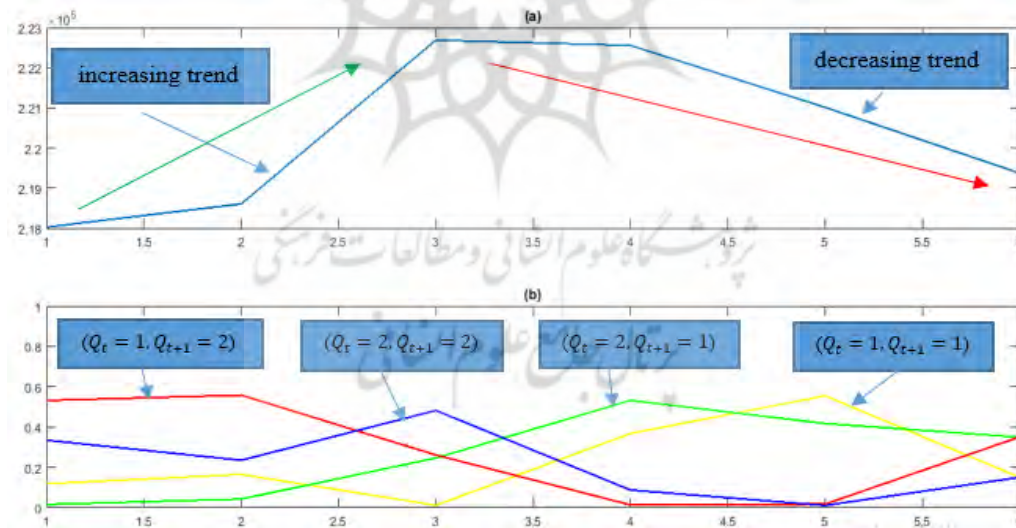

Figure 7. (a). The sample path of the Automotive Industry Index for the period of April 29, 2021, to May 3, 2021, (b). the estimated elements of joint conditional probability matrix for the best selected model (MS-SEMI-K(1)), the colors of red, blue, green, and yellow are the joint conditional probability of ($Q_t = 1, Q_{t+1} = 2$), ($Q_t = 2, Q_{t+1} = 2$), ($Q_t = 2, Q_{t+1} = 1$) and ($Q_t = 1, Q_{t+1} = 1$), respectively.

Conclusion

We have offered the strategy of buying and selling stock in financial markets by a special class of Markov switching models based on the joint conditional probability matrix. This strategy can be capable of selecting the various kernel functions. Such that, the best selection of this function, increase the accuracy of the joint conditional probability matrix. We have compared a set of kernel functions with the semi-parametric Markov switching models in terms of their ability to capturing the parameters in simulation studies. The estimation results of the models have indicated that the proper kernel function for estimating the parameters refer to the nature of data and there is not a firm decision in selecting these functions. Such that we found the proper performance of the Triweight kernel in the first simulation data set. While we found Uniform kernel as the best kernel function in the second simulation one. This reveals that the Gaussian kernel function is not the best selection in every data modeling. This reality shows that the researchers in estimating processes must be trial some kernel functions to find the best performance of their algorithms and the use of Gaussian kernel function in every algorithm can be deceptive.

We also suggest to academics that they compare these kernel functions with other semi-parametric and nonparametric models to improve the current knowledge about the better model for forecasting time series data.

References

Billio, M., Casarin, R., Ravazzolo, F., & Van Dijk, H. K. (2016). Interconnections between Eurozone and US booms and busts using a Bayesian panel Markov switching VAR model. Journal of Applied Econometrics.

Chan, K. C. G., & Wang, M. C. (2017). Semiparametric modeling and estimation of the terminal behavior of recurrent marker processes before failure events. Journal of the American Statistical Association, 112(517), 351-362.

Chang, Yoosoon., Yongok, Choi., & Joon Y. Park. (2017). A New Approach to Model Regime Switching. Journal of Econometrics, 196(1), 127-143.

Chang, J., Tang, C. Y., & Wu, Y., (2016). Local Independence Feature Screening for Nonparametric and Semiparametric Models by Marginal Empirical Likelihood. Annals of Statistics, 44(2), 515-539.

Chen, C. M., Shen, P. S., Wei, J. C., & Lin, L. (2017). A semiparametric mixture cure survival model for left-truncated and right-censored data. Biometrical Journal, 59(2), 270-290.

Di Persio, L., & Frigo, M. (2016). Gibbs sampling approach to the regime-

switching analysis of financial time series. Journal of Computational and Applied Mathematics.

Di Persio, L., & Frigo, M. (2015). Maximum Likelihood Approach to Markov Switching Models. WSEAS Transactions on Business and Economics, 12, 239-242.

Franke, J., Stockis, J.P., Kamgaing, J.T., & Li, W.K. (2011). Mixtures of nonparametric autoregressions. J. Nonparametric statistics, 23, 287-303.

Gu, X., Ma, Y., & Balasubramanian, R. (2015). Semiparametric Time to Event Models in the Presence of Error-Prone, Self-Reported Outcomes-with Application to the Women's Health Initiative. The annals of applied statistics, 9(2), 714-730.

Gupta, C., Cobre, J., Polpo, A., & Sinha, D. (2016). Semiparametric Bayesian estimation of quantile function for breast cancer survival data with cured fraction. Biometrical journal. Biometrische Zeitschrift, 58(5), 1164-77.

Nademi, A., & Farnoosh, R. (2014). Mixtures of autoregressive-autoregressive conditionally heteroscedastic models: semi-parametric approach. Journal of Applied Statistics, 41(2), 275-293.

Nademi, A., & Nademi, Y. (2019). Forecasting crude oil prices by a semiparametric Markov switching model: OPEC, WTI, and Brent cases. Energy Economics, 74, 757-766.

Nademi, A. (2019). The single-index panel data models with heterogeneous link function: mixture approach. Communications in Statistics-Simulation and Computation, 1-14.

Neale, M. C., Clark, S. L., Dolan, C. V., & Hunter, M. D. (2016). Regime Switching Modeling of Substance Use: Time-Varying and Second-Order Markov Models and Individual Probability Plots. Structural equation modeling: a multidisciplinary journal, 23(2), 221-233.

Von Ganske, J. (2016). A Regime Switching Partial Least Squares Approach to Forecasting Industry Stock Returns. Mimeo Edhec-Risk Institute.

Bibliographic information of this paper for citing:

Naderi, Hossein; Ghanbari, Mehrdad; Jamshidi Navid, Babak & Nademi, Arash (2021). Developing a Strategy for Buying and Selling Stocks Based on Semi-Parametric Markov Switching Time Series Models. *Iranian Journal of Finance*, 5(3), 108-126.