Hidden Markov Linear Regression Model and Its Parameter Estimation

HEFEI LIU[1], KUNQJNU WANG[1], AND YONG LI[ID][2]
[1]School of Mathematics and Statistics, Qujing Normal University, Qujing 655011, China
[2]School of Information and Engineering, Qujing Normal University, Qujing 655011, China

Corresponding author: Yong Li (qjsfxyly@163.com)

ABSTRACT This article first defines a hidden Markov linear regression model for the purpose of further studying the mutual transformation between different states in the linear regression model, and the regression relationship between the dependent variable and the independent variable in each state. And then, K-means clustering analysis methods are used to identify the hidden states of observed data, and the maximum likelihood estimation of the hidden state transition probability matrix elements is obtained by using the maximum likelihood estimation method, and parameter estimation of unknown parameters in linear regression model is also presented by using the least squares method. Finally, the observation vector set is generated according to the defined model, and the empirical simulation demonstrates that the parameter estimation method shown in this work is reliable.

INDEX TERMS Hidden markov model, linear regression model, cluster analysis, parameter estimation.

I. INTRODUCTION

In the 19th century, when the well-known British biologist and statistician Galton studied the genetic laws of parent height and the height of their children, he established an empirical straight-line equation for the height of an adult child about the average height of the parent, and named it as regression equation. After more than 100 years of development and evolution, regression analysis has become to be an important mathematical method and has been widely studied and applied in many disciplines, such as biological signal analysis [1], marine biological optical relationship reasoning [2], etc. The essence of regression analysis is a statistical analysis method of the quantitative change rule between multiple variables with related relationships, and then according to a mathematical model, the value of the independent variable is used to estimate or predict the possible value of the dependent variable. According to different mathematical models, regression analysis can be divided into linear regression models and nonlinear regression models.

Hidden Markov Model (HMM) is used to describe a Markov process. Although the state of the hidden Markov

The associate editor coordinating the review of this manuscript and approving it for publication was Fanbiao Li[ID].

model cannot be directly observed, it can be observed through the sequence of observation vectors, and each observation vector is represented by various probability density distributions in various states, and each observation vector is generated by a state sequence of the corresponding probability density distributions. Therefore, the Hidden Markov Model is a double random process, it has a certain number of hidden Markov chains and a set of display random functions, its research goal is to infer unobservable state transition information and distribution information in each state based on the information of observed variables [3], which in recent years has been widely used in wearable device data identification [4], SDN network early data stream matching [5], speech recognition [6], malfunction diagnosis [7], gene recognition [8], etc., and thereby providing us a series of research results.

Due to the good application effect of hidden Markov model in many fields mentioned above, many researches have paid their attention to Hidden Markov model. Yu used EM algorithm and forward-backward recursive algorithm to infer hidden Markov model [9]. Song et al studied the semi parametric hidden Markov model with latent variables by Bayesian method [10]. Liu *et al.* used reversible jump MCMC algorithm to choose unknown number of hidden states in

hidden Markov structural equation model [11]. Du investigates an adaptive sliding-mode controlled design problem for a class of Markov jump system with actuator faults [12]. Li *et al.* use homogenous polynomial approach to investigate Markovian jump system subjected to time-varying delays and infinite distributed delays [13].

Based on the existing research results of linear regression model and hidden Markov model, this article further studies the hidden Markov linear regression model with a fixed number of hidden states. For example, it is well known that family income has a linear correlation with various consumption expenditures, and a multiple linear regression model of income and various consumption expenditures can be established. However, the macro economic situation is divided into inflation and deflation, and the consumer market is divided into consumption upgrade and consumption degradation. When the macro economy is in two different states with inflation and deflation, and the consumer market is in two different states with consumption upgrade and consumption degradation, the regression relationship between the dependent and independent variables in the linear regression model is different, and these two states are often changed. Therefore, it is necessary to further study the mutual transformation law between different states, and the regression relationship between the dependent variable and independent variable in each state. In this article, a model capable of correctly expressing the rules of mutual transformation between different states and the regression relationship between the dependent variable and the independent variable in each state, named a hidden Markov multiple linear regression model, is introduced and is then committed to the model inference and its parameters estimation research.

Last but not least, the key difficulty in hidden Markov model inference is how to determine the hidden state of observation variables. At present, the most commonly used method to determine the hidden state of observation variables is forward backward recursive algorithm. However, the calculation of this algorithm is too complex to be realized, which brings difficulties to the applications of hidden Markov model. In this article, K-means cluster method is employed to address this problem. The hidden Markov model reduces the threshold of using hidden Markov model, because of its simplicity and fundamentality. This is the important contribution and value of this article, since it brings benefits to the application of hidden Markov model.

II. HIDDEN MARKOV LINEAR REGRESSION MODEL

Let Z_t be the hidden state at the t-th observation time point, and its value range is $\{1, 2, \ldots, K\}$, and vector $Dt = (y_t, x_{t1}, x_{t2}, \ldots, x_{tr})^T$ is the observation vector of the model at the t-th moment, where, y_t is the value of the dependent variable at the t-th moment, $x_{t1}, x_{t2}, \ldots, x_{tr}$ is the value of r independent variables at the t-th moment, t is the observation time point, $t = 1, 2, \cdots, T$.

Assume that the transition process of the hidden state satisfies the conditions of the Markov chain given below.

$$P\left(Z_t = s | Z_1, Z_2, \cdots, Z_{t-1} = u\right) = P\left(Z_t = s | Z_{t-1} = u\right) = a_{us}.$$

(1)

where $u = 1, 2, \cdots, K; s = 1, 2, \cdots, K; t = 2, 3, \cdots, T$; a_{us} is the transition probability from the hidden state u of the previous time point to the hidden state s of the next time point, the matrix of all possible hidden state transition probabilities is called the hidden state transition probability matrix, which is written as shown in the following formula.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{K1} & \cdots & a_{KK} \end{bmatrix}.$$

(2)

In summary, the triad $[D_t, Z_t, A]$ is called a hidden Markov model with K hidden states.

When the hidden state $Z_t = k$, a multiple linear regression model describing the relationship between the independent and dependent variables can be defined as follows.

$$[Y_k | Z_t = k] = X_k \beta_k + \varepsilon_k,$$

(3)

where β_k is the r-dimensional linear regression coefficient vector, ε_k is the deviation of the model, $[\varepsilon_k | Z_t = k] \sim N\left(0, \sigma_k^2 I\right)$, and

$$Y_k = \left[y_1, y_2, y_3, \cdots, y_{n_k}\right]^T,$$

$$X_k = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1r} \\ 1 & x_{21} & x_{22} & \cdots & x_{2r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n_k 1} & x_{n_k 2} & \cdots & x_{n_k r} \end{bmatrix},$$

$$\beta_k = [\beta_{k0}, \beta_{k1}, \beta_{k2}, \cdots, \beta_{kr}]^T,$$

$$\varepsilon_k = \left[\varepsilon_1, \varepsilon_2, \varepsilon_3, \cdots, \varepsilon_{n_k}\right]^T, \quad \varepsilon_i \sim N\left(0, \sigma^2\right).$$

Therefore, formula (3) is the hidden Markov linear regression model studied in this article.

III. THE PRINCIPLE OF PARAMETER ESTIMATION

A. THE DETERMINATION OF HIDDEN STATE

The hidden state determining of observation vectors is an important issue in the study of hidden Markov statistical models. In previous studies, Chen *et al.* used forward-backward algorithms to determine hidden states [14], and Liu *et al.* selected appropriate prior distributions for model parameters, and used Bayesian methods to infer the number of hidden states [15], [16]. However, both the forward and backward algorithms are very complicated not only in theoretical ideas but also in execution processes. Bayesian methods have many limitations in applications owing to the involved selection of prior distributions. Therefore, this article uses the cluster analysis method in multivariate statistical analysis to determine the hidden state of the observed variables.

Cluster analysis is a modern statistical analysis method that divides research objects into several categories according to

certain rules [17]. Common cluster analysis methods include systematic clustering and K-means clustering. The systematic clustering method first treats each case as a class, and then continuously merges classes according to the distance among classes until all cases are classified into one class, then a pedigree is obtained, and cases are classified reference to the pedigree. Choosing different distances will lead to different clustering results, each step of the systematic clustering algorithm needs to calculate the inter-class distance, so, the calculation amount of systematic clustering method is very large, and it takes up a lot of computer memory space, which thus requires high computing power. In order to improve this deficiency, Macqueen proposed a fast clustering method in 1965, which is the so-called K-means clustering method [18]. The K-means clustering method first roughly divides n objects into K categories, and then modifies the unreasonable classification according to some optimal criterion until the criterion function converges, and then obtains the final classification.

For the hidden Markov model with a fixed number of hidden states, this article uses K-means clustering to divide observation points into K classes. The class corresponding to each observation point can be used as the hidden state of the observation point at that time.

The principle and steps of k-means clustering method to determine the hidden state are as follows:

Step I, When the number of hidden states is S, randomly select S observation values as the initial cluster center, denoted as $L^{(0)} = \{X_1^{(0)}, X_2^{(0)}, \ldots, X_S^{(0)}\}$, and perform initial classification to obtain $G^{(0)} = \{G_1^{(0)}, G_2^{(0)}, \ldots, G_S^{(0)}\}$, where $G_i^{(0)} = \{X \in \Omega | d(X, X_i^{(0)}) \leq d(X, X_j^{(0)}), j = 1, 2, \ldots, S, i = 1, 2, \ldots, S, j \neq i\}$.

Step II, Firstly, starting from $G^{(0)}$, taking the center of gravity of $G_i^{(0)}$ as a new polynucleus, a new polynucleus set $L^{(1)} = \{X_1^{(1)}, X_2^{(1)}, \ldots, X_S^{(1)}\}$ is obtained. Then, starting from $L^{(1)}$, classify each observed variable to get a new classification $G^{(1)} = \{G_1^{(1)}, G_2^{(1)}, \ldots, G_S^{(1)}\}$, where, $G_i^{(1)} = \{X \in \Omega | d(X, X_i^{(1)}) \leq d(X, X_j^{(1)}), j = 1, 2, \ldots, S, i = 1, 2, \ldots, S, j \neq i\}$. Lastly, repeating it in order.

Step m, similar to step 2 to get the classification $G^{(m)} = \{G_1^{(m)}, G_2^{(m)}, \ldots, G_S^{(m)}\}$, here, $X_i^{(m)}$ is the center of gravity of the class $G_i^{(m-1)}$, but not necessarily the center of gravity of the class $G_i^{(m)}$. When m gradually increases, the classification tends to be stable, at this time, $X_i^{(m)}$ will be approximated to the center of gravity of the class. Therefore, when $X_i^{(m)} \approx X_i^{(m+1)}, G_i^{(m)} \approx G_i^{(m+1)}$, the clustering is complete. At this time, the class of each observation is the hidden state of the observation.

Because K-means clustering algorithm is convergent, the clustering result of the observed variable is convergent, and the hidden state judgment result is convergent, so the parameter estimation result is convergent.

B. THE ESTIMATION OF HIDDEN STATE TRANSITION PROBABILITY MATRIX

The hidden state transition probability matrix is an important part of the hidden Markov model, and its estimation has always been one of the core research problems of the model. Common estimation methods include the moment estimation method and the Baum-Welch algorithm based on the EM algorithm [19]. These two algorithms require researchers to have strong programming skills. This article uses the traditional classic maximum likelihood estimation method [20].

Let N_{ij} be the number of samples transferred from the previous hidden state z_i to the next hidden state z_j during the hidden state transition. In the hidden state transition probability matrix, the transition probability between the rows does not affect each other. For simplicity, the row index can be ignored and the maximum likelihood estimate of the transition probability can be derived using the transition probability of any row as an example.

Since the number of hidden states K is determined, there are K possibilities for the transition from the previous hidden state to the next hidden state, and the sum of the probability of these K possibilities $\sum_{i=1}^{K} a_i = 1$, there are $a_K = 1 - a_1 - \cdots - a_{K-1}$.

Then the likelihood function of the transition probability of any row can be written as

$$L(a_j) = a_1^{N_1} a_2^{N_2} \cdots a_{K-1}^{N_{K-1}} (1 - a_1 - \cdots - a_{K-1})^{N_K}.$$

Taking the natural logarithm on both sides of the above formula, we get the log-likelihood function as follows.

$$lnL(a_j) = N_1 lna_1 + N_2 lna_2 + \cdots + N_{K-1} lna_{K-1} + N_K ln(1 - a_1 - \cdots - a_{K-1})$$

Using the method of finding the maximum value of a function in analysis, in what follows we calculate the maximum point of a log-likelihood function. Calculating the partial derivative of the log-likelihood function $lnL(a_j)$ with respect to a_1, we get

$$\frac{\partial lnL(a_j)}{\partial a_1} = \frac{N_1}{a_1} - \frac{N_K}{1 - a_1 - \cdots - a_{K-1}}.$$

Finally, by making the partial derivative zero, we have

$$\frac{N_1}{a_1} = \frac{N_K}{1 - a_1 - \cdots - a_{K-1}}.$$

Then,

$$\frac{a_1}{a_K} = \frac{N_1}{N_K}.$$

Similarly, for any a_j ($j = 1, 2, \cdots, K - 1$) whose partial derivative is zero, we can obtain

$$\frac{a_j}{a_K} = \frac{N_j}{N_K}.$$

That is,

$$a_1 : a_2 : \cdots : a_{K-1} : a_K = N_1 : N_2 : \cdots : N_{K-1} : N_K.$$

And $\sum_{i=1}^{K} a_i = 1$, so the maximum likelihood of a_i can be estimated as,

$$\hat{a}_j = \frac{N_j}{\sum_{j=1}^{K} N_j}.$$

From the arbitrariness of the row mark i, the maximum likelihood estimate of any element a_{ij} in the hidden state transition probability matrix can be obtained as shown in formula (4):

$$\hat{a}_{ij} = \frac{N_{ij}}{\sum_{j=1}^{K} N_{ij}}, \tag{4}$$

where $i = 1, 2, \cdots, K; j = 1, 2, \cdots, K$。

C. THE ESTIMATION OF LINEAR REGRESSION COEFFICIENTS

This section mainly studies the estimation of the linear regression coefficient $\boldsymbol{\beta}_k$ under different hidden states. Since the parameter estimation of regression coefficients in traditional linear regression models mostly uses the least squares method or maximum likelihood estimation, and the least squares estimation has good characteristics such as optimality and unbiasedness, this article tends to use the least squares method to estimate the regression coefficient $\boldsymbol{\beta}_k$.

As mentioned before, in the k-th hidden state, the model can be written as follows

$$\boldsymbol{Y}_k = \boldsymbol{X}_k \boldsymbol{\beta}_k + \boldsymbol{\varepsilon}_k.$$

The so-called least squares estimation is to find $\boldsymbol{\beta}_k$, and to estimate when $(\boldsymbol{Y}_k - \boldsymbol{X}_k \boldsymbol{\beta}_k)^T (\boldsymbol{Y}_k - \boldsymbol{X}_k \boldsymbol{\beta}_k)$ taking the minimum value. Recording the sum of squared errors $Q(\boldsymbol{\beta}_k) = (\boldsymbol{Y}_k - \boldsymbol{X}_k \boldsymbol{\beta}_k)^T (\boldsymbol{Y}_k - \boldsymbol{X}_k \boldsymbol{\beta}_k)$, and finding the least square estimate of $\boldsymbol{\beta}_k$ being equivalent to finding the minimum value of $Q(\boldsymbol{\beta}_k)$. Using the matrix derivative formula, we get

$$\frac{\partial (\boldsymbol{Y}_k^T \boldsymbol{X}_k \boldsymbol{\beta}_k)}{\partial \boldsymbol{\beta}_k} = \boldsymbol{X}_k^T \boldsymbol{Y}_k, \quad \frac{\partial (\boldsymbol{\beta}_k^T \boldsymbol{X}_k^T \boldsymbol{X}_k \boldsymbol{\beta}_k)}{\partial \boldsymbol{\beta}_k} = 2\boldsymbol{X}_k^T \boldsymbol{X}_k \boldsymbol{\beta}_k,$$

from which we conclude that

$$\frac{\partial Q(\boldsymbol{\beta}_k)}{\partial \boldsymbol{\beta}_k} = -2\boldsymbol{X}_k^T \boldsymbol{Y}_k + 2\boldsymbol{X}_k^T \boldsymbol{X}_k \boldsymbol{\beta}_k = 0.$$

Let $\boldsymbol{X}_k$ be a full rank matrix, and the least square estimate of the regression coefficient vector $\boldsymbol{\beta}_k$ is

$$\boldsymbol{\beta}_k = \left(\boldsymbol{X}_k^T \boldsymbol{X}_k\right)^{-1} \boldsymbol{X}_k^T \boldsymbol{Y}_k \tag{5}$$

IV. EMPIRICAL SIMULATION

In order to test the reliability of the inference method of the hidden Markov multiple linear regression model introduced in this article, this section will give the number of hidden states K, the hidden state transition probability matrix $\boldsymbol{A}$, and the linear regression model in each hidden state Coefficient $\boldsymbol{\beta}_k$. First, a hidden state sequence set is generated according to the transition probability matrix $\boldsymbol{A}$, and then the observation vector at each observation time point is generated according to the hidden state value of each observation point and

the value of the multivariate linear regression model corresponding to the hidden state. Then, according to the above-mentioned method introduced in this article, the K-means clustering analysis method is used to cluster and identify the hidden states of the observation vector set, and the least square estimation of the coefficient $\boldsymbol{\beta}_k$ in the linear regression model with the number of hidden states is fixed. Finally, the results of parameter estimation are compared with real models to verify the reliability of the method.

A. SIMULATION I

First, taking the number of hidden states $K = 2$, then the hidden state probability transition matrix $\boldsymbol{A}$ is a second-order square matrix.

Letting,

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

For simplicity, suppose there is a ternary linear regression model in each hidden state.

Specifically,

$$\begin{cases} Y_1 = -2 - 2X_{11} - 2X_{12} - 2X_{13} \\ Y_2 = 2 + 2X_{21} + 2X_{22} + 2X_{23}. \end{cases}$$

In true simulation [21], two hidden states are generated first, and then 200 observation points are randomly generated according to the setting of the two hidden states, recorded as $D_t = [y_t, x_{t1}, x_{t2}, x_{t3}], t = 1, 2, 3, \cdots, 200$, where, $x_{t1} \sim N(0, 1), x_{t2} \sim U(0, 1), , x_{t3} \sim Exp(1)$.

Note that the frequency of the two hidden states obtained by computer simulation is z_1 and z_2, and the frequency of the two hidden states after cluster analysis is z_1' and z_2'.

In simulation I, the effects of identifying the two hidden states using K-means clustering are shown respectively in Table 1 and figure 1, and the experimental results are shown in Table 2, 3.

TABLE 1. Contingency table comparison of the number of hidden states before and after clustering.

hidden state	z_1'	z_2'	total
z_1	88	3	91
z_2	1	108	109
Total	89	111	200

B. SIMULATION II

Letting,

$$\boldsymbol{A} = \begin{bmatrix} 0.4 & 0.6 \\ 0.5 & 0.5 \end{bmatrix}.$$

$$\begin{cases} Y_1 = -2 - 1.5X_{11} - 0.5X_{12} - 1.2X_{13} \\ Y_2 = 2 + 1.2X_{21} + 1.5X_{22} + 1.5X_{23}. \end{cases}$$

In simulation II, the effects of identifying the two hidden states using K-means clustering are shown respectively
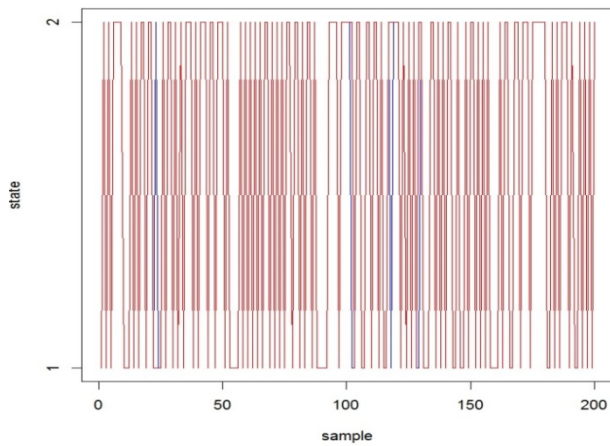
FIGURE 1. Hidden state path diagram.

TABLE 2. Estimation results of the hidden state transition probability matrix.

parameter	true value	estimated value
a_{11}	0.3	0.270
a_{12}	0.7	0.730
a_{21}	0.6	0.582
a_{22}	0.4	0.418

TABLE 3. True and estimated values of the parameters of the linear regression model.

parameter	actual value	estimated value	parameter	actual value	estimated value
β_{10}	-2	-2.036	β_{20}	2	2.023
β_{11}	-2	-1.989	β_{21}	2	1.990
β_{12}	-2	-1.962	β_{22}	2	1.981
β_{13}	-2	-1.990	β_{23}	2	1.993

in Table 4 and figure 2, and the experimental results are shown in Table 5, 6.

TABLE 4. Contingency table comparison of the number of hidden states before and after clustering.

hidden state	z_1'	z_2'	total
z_1	93	3	96
z_2	1	103	104
Total	94	106	200

C. SIMULATION ANALYSIS

In Figure 1 and 2, we use the green line to represent the real hidden state, while the red line to represent the hidden state result of clustering analysis. Table 1 and 4, figure 1 and 2 show that 196 clusters of 200 observation points are clustered correctly and 4 clusters are clustered incorrectly.
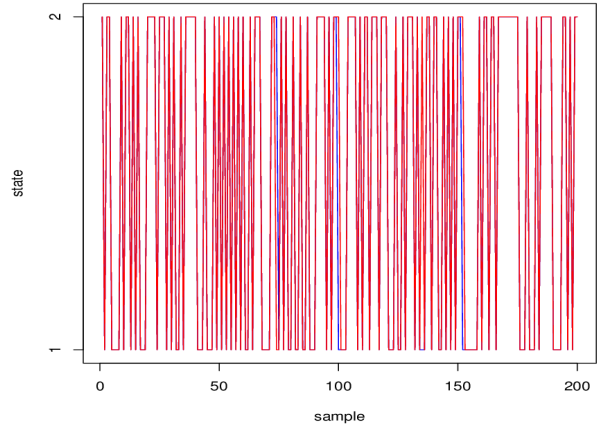
FIGURE 2. Hidden state path diagram.

TABLE 5. Estimation results of the hidden state transition probability matrix.

parameter	true value	estimated value
a_{11}	0.4	0.404
a_{12}	0.6	0.596
a_{21}	0.5	0.533
a_{22}	0.5	0.4

TABLE 6. True and estimated values of the parameters of the linear regression model.

parameter	actual value	estimated value	parameter	actual value	estimated value
β_{10}	-2	-2.008	β_{20}	2	2.289
β_{11}	-1.5	-1.459	β_{21}	1.2	1.063
β_{12}	-0.5	-0.541	β_{22}	1.5	1.182
β_{13}	-1.2	-1.187	β_{23}	1.5	1.452

After clustering analysis, this article then uses Eq. (4) to estimate the maximum likelihood of the hidden state transition matrix. The experimental results are shown in Table 2 and 4. Finally, in order to obtain a linear regression model, this article continues to use the explanatory variables and response variables after clustering to perform a least squares estimation. The experimental results are recorded in Table 3 and 6.

The results in Tables 1 to 6, Figure 1 and 2 show that the K-means clustering method is effective for the hidden data identification, hidden state transition probability matrix, and linear regression model parameter estimation of observation data.

V. CONCLUSION

This article combines the hidden Markov model and the linear regression model to give the definition of the hidden Markov linear regression model. The hidden state determination, state transition probability matrix and parameter estimation problems involved in the hidden Markov linear

regression model are introduced. K-means clustering method is used to determine hidden states. This is the innovation of this article. We use maximum likelihood method to estimate the element of transition probability matrix. And we also use the least square method to estimate the parameter of linear regression model. Simulation results demonstrate that the estimate effect is good. However, this article studies a hidden Markov linear regression model with a fixed number of hidden states, and the use of the model has certain limitations. Therefore, the next research direction will be the reasoning and application research of more complex models, such as hidden Markov logistic regression model, hidden Markov quartile regression model, and hidden Markov logarithmic linear model, and nonlinear stochastic semi-Markov model [22], [23], or hidden Markov model with an unknown number of hidden states [24].

REFERENCES

[1] W. Yubo and V. Kalyana, "Time-frequency analysis of non-stationary biological signals with sparse linear regression based Fourier linear combiner," *Sensors*, vol. 17, no. 6, pp. 1386–1399, 2017.

[2] M. Bellacicco, V. Vellucci, M. Scardi, M. Barbieux, S. Marullo, and F. D'Ortenzio, "Quantifying the impact of linear regression model in deriving bio-optical relationships: The implications on ocean carbon Estimations," *Sensors*, vol. 19, no. 13, pp. 3032–3046, 2019.

[3] P. Tiawongsombat, M.-H. Jeong, A. Pirayawaraporn, J.-J. Lee, and J.-S. Yun, "Vision-based attentiveness determination using scalable HMM based on relevance theory," *Sensors*, vol. 19, no. 13, pp. 5331–5351, 2019.

[4] F.-C. Martindale, S. Sebastijan, and M.-E. Bjoern, "Hidden Markov model-based smart annotation for benchmark cyclic activity recognition database using wearables," *Sensors*, vol. 19, no. 8, pp. 1820–1841, 2019.

[5] C. Wang and H.-Y. Youn, "Entry aggregation and early match using hidden Markov model of flow table in SDN," *Sensors*, vol. 19, no. 10, pp. 2341–2360, 2019.

[6] M.-Z. Wei and M.-Z. Tang, "Hidden Markov models for speech recognition technology based on classification and identification," in *Proc. 2nd Int. Conf. Inf. Commun. Technol. Educ.*, ChangSha, China, 2015, pp. 26–29.

[7] Y.-Z. Jia, M.-Q. Xu, and R.-X. Wang, "Symbolic important point perceptually and hidden Markov model based hydraulic pump fault diagnosis method," *Sensors*, vol. 18, no. 13, pp. 4460–4479, 2018.

[8] L. Qing and X.-D. Yan, "Design and implementation of gene recognition system based on hidden Markov model," *Comput. Eng. Appl.*, vol. 39, no. 24, pp. 69–71, 2003.

[9] S.-Z. Yu, "Hidden semi-Markov models," *Artif. Intell.*, vol. 174, no. 2, pp. 215–243, Feb. 2010.

[10] X. Song, K. Kang, M. Ouyang, X. Jiang, and J. Cai, "Bayesian analysis of semiparametric hidden Markov models with latent variables," *Struct. Equation Model., Multidisciplinary J.*, vol. 25, no. 1, pp. 1–20, Jan. 2018.

[11] H.-F. Liu and X.-Y. Song, "Bayesian analysis of hidden Markov structural equation models with an unknown number of hidden states," *Econometrics Statist.*, May 2020, doi: 10.1016/j.ecosta.2020.03.003.

[12] C. Du, F. Li, and C. Yang, "An improved homogeneous polynomial approach for adaptive sliding-mode control of Markov jump systems with actuator faults," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 955–969, Mar. 2020.

[13] F. Li, X. Li, X. Zhang, and C. Yang, "Asynchronous filtering for delayed Markovian jump systems via homogeneous polynomial approach," *IEEE Trans. Autom. Control*, vol. 65, no. 5, pp. 2163–2170, May 2020.

[14] H.-Y. Chen, X.-G. Gao, and J.-F. Mei, "Fast forward and backward algorithm of generalized hidden Markov model," *Syst. Eng. Electron.*, vol. 34, no. 10, pp. 2175–2179, 2012.

[15] H.-F. Liu, K. Wang, and C.-F. Jiang, "Bayesian inference of hidden Markov multivariate normal distribution with unknown hidden states," *Stat. Res.*, vol. 34, no. 12, pp. 119–125, 2017.

[16] H.-F. Liu, K. Wang, and C.-F. Jiang, "Hidden Markov structural equation model and its Bayesian estimation," *Math. Statist. Manage.*, vol. 37, no. 2, pp. 272–279, 2018.

[17] B.-H. Wang, *Multivariate Statistical Analysis and R Language Modeling*. Guangzhou, China: Jinan Univ. Press, 2010, sec. 6, pp. 133–150.

[18] J.-B. Macqueen, "On convergence of k-means and partitions with minimum average variance," *Ann. Math. Statist.*, vol. 34, no. 8, p. 36, 1965.

[19] H. Li, *Statistical Learning Methods*. Beijing, China: Tsinghua Univ. Press, 2012, sec. 2, pp. 183–185.

[20] S.-S. Yun, J.-L. Wang, and X.-L. Yun, *Advanced Mathematical Statistics*. Beijing, China: Higher Education Press, 2012, sec. 2, pp. 145–149.

[21] Y.-C. Tang, *R Language and Statistical Analysis*. Beijing, China: Higher Education Press, 2008, pp. 56–67.

[22] W. Qi, G. Zong, and H. R. Karimi, "Sliding mode control for nonlinear stochastic singular semi-Markov jump systems," *IEEE Trans. Autom. Control*, vol. 65, no. 1, pp. 361–368, Jan. 2020.

[23] W. Qi, G. Zong, and H. R. Karimi, "Sliding mode control for nonlinear stochastic semi-Markov switching systems with application to SRMM," *IEEE Trans. Ind. Electron.*, vol. 67, no. 5, pp. 3955–3966, May 2020.

[24] H. Liu and X. Y. Song, "Bayesian analysis of mixture structural equation models with an unknown number of components," *Struct. Equation Model., Multidisciplinary J.*, vol. 25, no. 1, pp. 41–55, Jan. 2018.
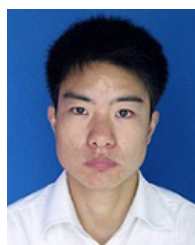
HEFEI LIU received the M.S. degree from the School of Probability Theory and Mathematical Statistics, Yunnan University, Kunming, China, in 2011. He is currently an Assistant Professor with the School of Mathematics and Statistics, Qujing Normal University, China. His research interests include mathematical statistics and Bayesian inference.

KUNQJNU WANG received the M.S. degree from the School of Probability Theory and Mathematical Statistics, Yunnan University, Kunming, China, in 2009. He is currently a Lecturer with the School of Mathematics and Statistics, Qujing Normal University, China. His research interests include statistics and data processing.

YONG LI received the M.S. degree from the School of Computer Science, Wuyi University, Jiangmeng, China, in 2010. He is currently an Assistant Professor with the School of Information Engineering, Qujing Normal University of China, Qujing, China. His research interests include computer networks, information security, and cloud computing.

• • •