*ICIC Express Letters*
Volume **10**, Number **6**, June **2016**

ICIC International ©2016 ISSN 1881-803X
pp. 1427–1433

# MULTIVARIATE TIME SERIES SEGMENTATION APPROACH BASED ON HIDDEN MARKOV MODELS

Hongyue Guo[1,2], Zhubin Sun[2] and Xiaodong Liu[2]

[1]School of Mathematical Sciences
[2]School of Control Science and Engineering
Faculty of Electronic Information and Electrical Engineering
Dalian University of Technology
No. 2, Linggong Road, Ganjingzi District, Dalian 116024, P. R. China
xdliuros@dlut.edu.cn

Received November 2015; accepted February 2016

Abstract. *In this paper, an approach based on hidden Markov models (HMM) is applied to segmenting multivariate time series. In this algorithm, each observation in the underlying time series is dependent on a corresponding hidden state, and the segmentation problem comes down to finding the hidden states. When related parameters are given, the states can be obtained under a maximum likelihood framework using the Viterbi algorithm. Some meaningful methods including vector autoregression (VAR) models are utilized to estimate the parameters. In this way, the state estimation step and the parameter estimation step will perform repeatedly until the convergence condition is satisfied. The segmentation procedure is evaluated by a hydrometeorological time series.*

**Keywords:** Hidden Markov model, Segmentation, Multivariate time series, Vector autoregression

1. **Introduction.** The hidden Markov model (HMM) is a simple dynamic Bayesian network. In an HMM, the (hidden) states are considered to be a Markov process. In recent years, different types of hidden Markov models have been used in pattern recognition [8, 10], and time series segmentation [5, 6].

As a particular type of clustering, time series segmentation aims to partition a given time series into segments with similar characteristics. Due to its importance for hydrometeorological time series, several segmentation approaches have been proposed. For example, [5] proposed an HMM segmentation procedure to partition univariate hydrological and environmental time series. In [6], shifting means hidden Markov models have been applied to segmenting hydrological time series. Besides, [7] employed dynamic programming (DP) algorithm to partition time series into segments, and Bayesian information criterion (BIC) [9] is applied to determining the global optimal segmentation for every segmentation order. [4] generalized the DP segmentation approach to segment multivariate time series. [3] incorporated the remaining cost concept into the DP algorithm to segment long time series.

As discussed by [4, 11], it is helpful to detect sudden changes and segment multivariate time series. In fields such as climate change studies and wind resources, time series segmentation is useful for understanding their characteristics in order to model them, or to predict possibilities of some extreme cases. This paper aims to generalize the HMM segmentation approach proposed by [5] to partition multivariate time series. The segmentation problem is transformed into finding states. To estimate related parameters, some meaningful approaches are presented. The vector autoregression (VAR) model is incorporated into the HMM segmentation algorithm as an alternative. Then the states can be calculated by the Viterbi algorithm to maximize log-likelihood function.

The rest of this paper is organized as follows. In Section 2, formulation of the segmentation problem is introduced. In Section 3, the segmentation approach is given including the state and parameter estimation. Experimental results of the HMM segmentation algorithm are presented in Section 4. Section 5 shows some conclusions.

2. **Definitions and Formulation of the Segmentation Problem.** Let $\boldsymbol{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T]$ be a $d$-dimensional multiple time series, with $\boldsymbol{x}_t = [x_{1t}, \ldots, x_{dt}]^T$. Time series segmentation aims to find a segmentation $\boldsymbol{t} = [t_0, t_1, \ldots, t_{N-1}, t_N]$, which satisfies $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$. Here, $t_0, t_1, \ldots, t_{N-1}, t_N$ are called change points, and $[t_{i-1} + 1, t_i]$ forms the $i$-th segment and the number of segments $N$ is called the segmentation order.

In what follows, the segmentation problem is explained under the framework of HMM. Let $\boldsymbol{X}$ be a $d$-dimensional random vector, at time $t$, the observation of $\boldsymbol{X}_t$ depends on its corresponding state $Z_t$, where $Z_t$ is not directly visible. Suppose $\boldsymbol{X}_t$ is a multivariate normal random vector with mean $\boldsymbol{\mu}_{Z_t}$ and covariance matrix $\boldsymbol{\Sigma}_{Z_t}$. In other words, if $Z_t = z_t$ is given, $\boldsymbol{X}_t \sim N(\boldsymbol{\mu}_{z_t}, \boldsymbol{\Sigma}_{z_t})$, that is

$$f(\boldsymbol{x}_t|z_t) = (2\pi)^{-\frac{d}{2}}|\boldsymbol{\Sigma}_{z_t}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\boldsymbol{x}_t - \boldsymbol{\mu}_{z_t}\right)^T \boldsymbol{\Sigma}_{z_t}^{-1}\left(\boldsymbol{x}_t - \boldsymbol{\mu}_{z_t}\right)\right) \tag{1}$$

Further, if $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, …, $\boldsymbol{X}_T$ are mutually independent under the conditions that $\boldsymbol{Z}_1$, $\boldsymbol{Z}_2$, …, $\boldsymbol{Z}_T$ are given, we have

$$f(\boldsymbol{x}|\boldsymbol{z}) = (2\pi)^{-\frac{Td}{2}}\prod_{t=1}^{T}|\boldsymbol{\Sigma}_{z_t}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\boldsymbol{x}_t - \boldsymbol{\mu}_{z_t}\right)^T \boldsymbol{\Sigma}_{z_t}^{-1}\left(\boldsymbol{x}_t - \boldsymbol{\mu}_{z_t}\right)\right) \tag{2}$$

Let $\boldsymbol{M}$ and $\boldsymbol{\Sigma}$ be the sets of means and covariance matrices separately for all segments, that is, $\boldsymbol{M} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_N]$ and $\boldsymbol{\Sigma} = [\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \ldots, \boldsymbol{\Sigma}_N]$.

Denote the hidden state sequence of $\boldsymbol{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T]$ as $\boldsymbol{z} = [z_1, z_2, \ldots, z_T]$, which can be calculated when $\boldsymbol{x}$ is given (see Section 3.1). Next, the one-to-one correspondence between the states $\boldsymbol{z} = [z_1, z_2, \ldots, z_T]$ and the segmentation $\boldsymbol{t} = [t_0, t_1, \ldots, t_N]$ is explained. Given a state set $\boldsymbol{z}$, the segmentation $\boldsymbol{t}$ is a set of the locations where $z_{t_i} \neq z_{t_i+1}$, $i = 1, 2, \ldots, N-1$. When a segmentation $\boldsymbol{t}$ is given, the state sequence is obtained as $z_{t_{i-1}+1} = \cdots = z_{t_i} = i$, $i = 1, 2, \ldots, N$.

In an HMM, the (hidden) state process $Z_t$ is a Markov process, that is

$$P(Z_t = z_t|Z_1 = z_1, \ldots, Z_{t-1} = z_{t-1}) = P(Z_t = z_t|Z_{t-1} = z_{t-1}) \tag{3}$$

Let the transition probability be $P_{i,j} = P(Z_{t+1} = j|Z_t = i)$, and then we have the joint probability of $Z_1, Z_2, \ldots, Z_T$ as

$$P(Z_1 = z_1, Z_2 = z_2, \ldots, Z_T = z_T) = \pi_{z_0} \cdot P_{z_0,z_1} \cdot P_{z_1,z_2} \cdot \ldots \cdot P_{z_{T-1},z_T} \tag{4}$$

where $\pi$ is the initial probability, and $\pi_1 = 1$, $\pi_i = 0$, for $i = 2, 3, \ldots, N$. Following [5], we employ "left-to-right" continuous HMM [8], namely, the transition probability matrix $\boldsymbol{P}$ is forward and $P_{i,j} = 0$ for $i = 1, 2, \ldots, N$ and $j \neq i, i+1$ and $P_{N,N} = 1$. Then we have the transition probability matrix as

$$\boldsymbol{P} = \begin{bmatrix} P_{1,1} & 1 - P_{1,1} & & & & \\ & P_{2,2} & 1 - P_{2,2} & & & \\ & & & \ddots & \ddots & \\ & & & & P_{N-1,N-1} & 1 - P_{N-1,N-1} \\ & & & & & 1 \end{bmatrix} \tag{5}$$

## 3. Segmentation for Multivariate Time Series.

This section describes an HMM based segmentation approach for multivariate time series in detail. The HMM segmentation algorithm needs to repeat the state and parameter estimation step until the convergence condition is satisfied. Section 3.1 reviews the Viterbi algorithm used to estimate states. Parameter estimation methods are discussed in Section 3.2. Section 3.3 combines the state and parameter estimation and shows the HMM segmentation procedure.

### 3.1. Viterbi algorithm.

For a given time series $\boldsymbol{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T]$, where $\boldsymbol{x}_i$ is a $d$-dimensional vector. Suppose the parameters $\boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{P}$ are known, then the Viterbi algorithm [2] can find an optimal state $\boldsymbol{z} = [z_1, z_2, \ldots, z_T]$ by maximizing $f(\boldsymbol{z}|\boldsymbol{x})$, that is

$$\hat{\boldsymbol{z}} = \arg\max_{\boldsymbol{z}} f(\boldsymbol{z}|\boldsymbol{x}) \tag{6}$$

According to condition probability formula, the conditional probability of state $\boldsymbol{z}$ can be obtained by

$$f(\boldsymbol{z}|\boldsymbol{x}) = \frac{f(\boldsymbol{x}, \boldsymbol{z})}{f(\boldsymbol{x})} = \frac{f(\boldsymbol{x}|\boldsymbol{z}) f(\boldsymbol{z})}{f(\boldsymbol{x})} \tag{7}$$

As $f(\boldsymbol{x})$ is independent with $\boldsymbol{z}$, we have

$$\hat{\boldsymbol{z}} = \arg\max_{\boldsymbol{z}} f(\boldsymbol{z}|\boldsymbol{x}) = \arg\max_{\boldsymbol{z}} f(\boldsymbol{x}, \boldsymbol{z}) \tag{8}$$

Based on Equations (2) and (4), the joint distribution of $\boldsymbol{x}$ and $\boldsymbol{z}$ can be calculated as follows.

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{z}) &= f(\boldsymbol{x}|\boldsymbol{z}) f(\boldsymbol{z}) \\ &= (2\pi)^{-\frac{Td}{2}} \prod_{t=1}^{T} P_{z_{t-1}, z_t} |\boldsymbol{\Sigma}_{z_t}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (\boldsymbol{x}_t - \boldsymbol{\mu}_{z_t})^T \boldsymbol{\Sigma}_{z_t}^{-1} (\boldsymbol{x}_t - \boldsymbol{\mu}_{z_t})\right) \end{aligned} \tag{9}$$

Let $L$ be the log-likelihood function of $\boldsymbol{x}$ and $\boldsymbol{z}$, that is

$$\begin{aligned} L &= \ln f(\boldsymbol{x}, \boldsymbol{z}) \\ &= -\frac{Td}{2} \ln(2\pi) + \sum_{t=1}^{T} \left(\ln P_{z_{t-1}, z_t} - \frac{1}{2} \ln |\boldsymbol{\Sigma}_{z_t}| - \frac{1}{2} (\boldsymbol{x}_t - \boldsymbol{\mu}_{z_t})^T \boldsymbol{\Sigma}_{z_t}^{-1} (\boldsymbol{x}_t - \boldsymbol{\mu}_{z_t})\right) \end{aligned} \tag{10}$$

Then dynamic programming [1] approach is applied to computing the state set. The pseudocode of the Viterbi algorithm is shown in Algorithm 1.

**Algorithm 1** Viterbi Algorithm

**Forward Recursion:**

Set $q_{1,0} = 1$, $q_{2,0} = q_{3,0} = \cdots = q_{N,0} = 0$;
**for** $t = 1, 2, \ldots, T$ **do**
  **for** $i = 1, 2, \ldots, N$ **do**
    $q_{i,t} = \max_{\max(1,i-1) \le j \le i} (q_{j,t-1} \cdot P_{j,i} \cdot f(\boldsymbol{x}_t; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i))$
    $r_{i,t} = \arg\max_{\max(1,i-1) \le j \le i} (q_{j,t-1} \cdot P_{j,i} \cdot f(\boldsymbol{x}_t; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i))$
  **end for**
**end for**

**Backward Recursion:**

$\hat{L}_{N,T} = \max_{1 \le i \le N}(q_{i,T})$;
$\hat{z}_T = \arg\max_{1 \le i \le N}(q_{i,T})$;
**for** $t = T, T-1, \ldots, 1$ **do**
  $\hat{z}_{t-1} = r_{\hat{z}_t, t}$
**end for**

3.2. **Parameter estimation.** This section presents the methods to estimate the parameters $\boldsymbol{M}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{P}$. Suppose a state set $\boldsymbol{z} = [z_1, z_2, \ldots, z_T]$ is given, and the corresponding segmentation is $\boldsymbol{t} = [t_0, t_1, \ldots, t_N]$. Then, to estimate $\boldsymbol{M} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_N]$, a feasible method is

$$\hat{\boldsymbol{\mu}}_i = \frac{\sum_{t=t_{i-1}+1}^{t_i} \boldsymbol{x}_t}{T_i}, \qquad i = 1, 2, \ldots, N \tag{11}$$

where $T_i = t_i - t_{i-1}$ is the length of the $i$-th segment.

Besides, vector autoregressive (VAR) model can be another alternative. Suppose $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T$ are generated by a VAR($p$) process, that is

$$\boldsymbol{X}_t = \Phi_0 + \Phi_1 \boldsymbol{X}_{t-1} + \cdots + \Phi_p \boldsymbol{X}_{t-p} + \boldsymbol{a}_t \tag{12}$$

where $\boldsymbol{a}_t$ is a $d$-dimensional white noise, and $p$ is the order of autoregression. $\Phi_0$ is a vector parameter of $d$ dimension and $\Phi_1, \ldots, \Phi_p$ are all $d \times d$ parameter matrices.

Suppose the data in the $i$-th segment satisfy

$$\boldsymbol{x}_t = \Phi_0^{(i)} + \Phi_1^{(i)} \boldsymbol{x}_{t-1} + \cdots + \Phi_p^{(i)} \boldsymbol{x}_{t-p} + \boldsymbol{a}_t^{(i)} \tag{13}$$

then the estimations $\hat{\Phi}_0^{(i)}$, $\hat{\Phi}_1^{(i)}$, $\ldots$, $\hat{\Phi}_p^{(i)}$ can be obtained based on the data $[\boldsymbol{x}_{t_{i-1}+1}, \ldots, \boldsymbol{x}_{t_i}]$ and least-square (LS) estimation approach. Then, the conditional expectation of each point $\boldsymbol{X}_t$ is

$$E(\boldsymbol{X}_t | \boldsymbol{x}_{t-1}, \ldots, \boldsymbol{x}_{t-p}, Z_t = z_t) = \hat{\Phi}_0^{(z_t)} + \hat{\Phi}_1^{(z_t)} \boldsymbol{x}_{t-1} + \cdots + \hat{\Phi}_p^{(z_t)} \boldsymbol{x}_{t-p} \tag{14}$$

In this way, each segment gets a set of regressive parameters, that is to say, we can obtain $N$ sets of regressive parameters. At each iteration, this method aims to reestimate $\Phi_0^{(i)}$, $\Phi_1^{(i)}$, $\ldots$, $\Phi_p^{(i)}$, $i = 1, 2, \ldots, N$, rather than reestimate $\boldsymbol{M}$.

To estimate the covariance matrix $\boldsymbol{\Sigma}_i$, $i = 1, 2, \ldots, N$, one could use

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{T_i - 1} \sum_{t=t_{i-1}+1}^{t_i} (\boldsymbol{x}_t - \hat{\boldsymbol{\mu}}_i)(\boldsymbol{x}_t - \hat{\boldsymbol{\mu}}_i)^T \tag{15}$$

Also, a simpler estimation is to set all $\boldsymbol{\Sigma}_i$, $i = 1, 2, \ldots, N$, to be equal. That is,

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_N = \boldsymbol{\Sigma} \tag{16}$$

The estimation of $\boldsymbol{\Sigma}$ can be calculated by

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T - 1} \sum_{t=1}^{T} (\boldsymbol{x}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{x}_t - \hat{\boldsymbol{\mu}})^T \tag{17}$$

where

$$\hat{\boldsymbol{\mu}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{x}_t \tag{18}$$

It should be noted that the estimation Equation (17) is independent with segmentation.

As to the estimation of transition probability matrix $\boldsymbol{P}$, [5] shows three methods, one of which is segmentation-independent. The experiments in [5] indicated that the segmentation results obtained by using different methods are the same. Accordingly, this paper selects a simpler method. As discussed in Section 2, we only need to estimate $P_{i,i}$, $i = 1, 2, \ldots, N - 1$. The estimation approach we used is

$$\hat{P}_{i,i} = \frac{T_i - 1}{T_i} \tag{19}$$

where $T_i$ is the length of the $i$-th segment.

3.3. **HMM segmentation algorithm.** In this section, we integrate the state and parameter estimation step into the HMM segmentation algorithm. The inputs are the time series $\boldsymbol{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T]$, the number of segments and a termination variable $\varepsilon$. First, a segmentation is randomly generated and the parameters $\boldsymbol{M}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{P}$ are initialized. Then, at each iteration, the states set $\boldsymbol{z}$ is updated by the Viterbi algorithm, and the parameters $\boldsymbol{M}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{P}$ are calculated based on the segmentation obtained from the states. The aforementioned two steps will perform until the convergence of the log-likelihood function $L$ (Equation (10)). In this way, the optimal states $\hat{\boldsymbol{z}}$ can be solved, and the corresponding segmentation $\hat{\boldsymbol{t}}$ can be obtained. Algorithm 2 gives the pseudocode of the HMM segmentation algorithm.

**Algorithm 2** HMM Segmentation Algorithm

**Input:**
Time series $\boldsymbol{x}$;
The number of segments $N$;
Termination value $\varepsilon$;
**Output:**
The segmentation $\hat{\boldsymbol{t}}$;
**main:**
Randomly generate a segmentation $\boldsymbol{t}^{(0)}$;
Compute $\boldsymbol{z}^{(0)}$ based on $\boldsymbol{t}^{(0)}$;
Compute $\boldsymbol{M}^{(0)}$ by Equation (11) (or Equation (14));
Compute $\boldsymbol{\Sigma}^{(0)}$ by Equation (15);
Compute $\boldsymbol{P}^{(0)}$ by Equation (19);
**for** $t = 1, 2, \ldots, N_{sim}$ **do**
  **State Estimation**
  Update $\boldsymbol{z}^{(t)}$ by the Viterbi algorithm, and calculate the log-likelihood function $L^{(t)}$;
  Compute $\boldsymbol{t}^{(t)}$ from $\boldsymbol{z}^{(t)}$;
  **Parameter Estimation**
  Compute $\boldsymbol{M}^{(t)}$ by Equation (11) (or Equation (14));
  Compute $\boldsymbol{\Sigma}^{(t)}$ by Equation (15);
  Compute $\boldsymbol{P}^{(t)}$ by Equation (19);
  **Termination Criterion**
  **if** $\left|L^{(t)} - L^{(t-1)}\right| < \varepsilon$ **then**
    $\hat{\boldsymbol{z}} = \boldsymbol{z}^{(t)}$; $\hat{\boldsymbol{M}} = \boldsymbol{M}^{(t)}$; $\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{(t)}$;
  **end if**
  Compute $\hat{\boldsymbol{t}}$ from $\hat{\boldsymbol{z}}$.
**end for**

For univariate time series, [5] discussed the convergence of the HMM segmentation algorithm, which is also viable for multivariate case. Suppose $L^{(0)}$, $L^{(1)}$,… converge to $\hat{L}$, which is a local maximum of $L$. That is, the obtained states $\hat{\boldsymbol{z}}$ (the corresponding segmentation $\hat{\boldsymbol{t}}$) is a local optimal solution.

4. **Experiments.** In this section, the HMM segmentation approach is applied to segmenting a hydrometeorological time series plotted in Figure 1, which has been previously used by [4, 11] (with a different temporal interval).

The HMM segmentation algorithms based on homogeneity of segment means and vector autoregression (VAR) model separately are both run 10 trials for $N = 2, 3, 4, 5$. For each $N$, the segmentation maximizing log-likelihood function $L$ is selected to be the most

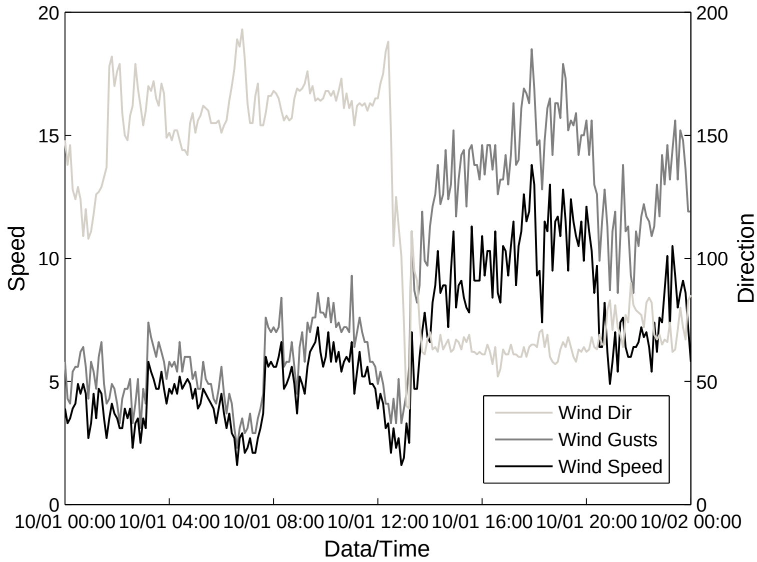


FIGURE 1. Plots of wind direction, wind gusts and wind speed courtesy of Arecibo, PR on 2013/10/01

TABLE 1. Optimal segmentation results of the series plotted in Figure 1, with $N = 2, 3, 4, 5$. The means are estimated by Equation (11).

| $N$ | Segmentation | Log-likelihood |
|---|---|---|
| 2 | 0 134 241 | $-1642.9$ |
| 3 | 0 126 137 241 | $-1571.8$ |
| 4 | 0 77 123 133 241 | $-1502.5$ |
| 5 | 0 77 125 141 205 241 | $-1442.6$ |

TABLE 2. Optimal segmentation results of the series plotted in Figure 1, with $N = 2, 3, 4, 5$. The means are estimated by VAR model.

| $N$ | Segmentation | Log-likelihood |
|---|---|---|
| 2 | 0 134 241 | $-1491.4$ |
| 3 | 0 125 134 241 | $-1432.6$ |
| 4 | 0 29 125 134 241 | $-1375.4$ |
| 5 | 0 3 77 125 137 241 | $-1308.8$ |

proper segmentation. Tables 1 and 2 show the segmentation results, and the maximum log-likelihood for each $N$ is listed in the last column.

When $N$ is set to be 2, the segmentation obtained is 0, 134, 241, which is the same as the global optimal segmentation given by [4] based on Bayesian information criterion (BIC). In this paper, log-likelihood function $L$ depends on the parameters $\boldsymbol{M}$, $\boldsymbol{\Sigma}$, $\boldsymbol{P}$ and $N$. Due to the difference in optimal functions, for $N = 3, 4, 5$, the optimal segmentation results are not equal.

Now, we turn to the selection of segmentation order. Several methods have been proposed to select the optimal number of segments, see [4, 5, 7, 11]. For univariate time series, [5] determined the number of segments based on log-likelihood function $L$, and obtained satisfactory segmentation results. According to the experiments in our paper, $L$ increases with $N$ for $N = 2, 3, 4, 5$. For multivariate time series, with the increasing of dimension, the penalization in $L$ on $N$ may not be enough. However, we have calculated the optimal segmentation for $N = 2, 3, 4, 5$, and users can select a proper segmentation based on their own judgments.

5. **Conclusion.** In this paper, a hidden Markov model (HMM) segmentation approach is proposed to segment multivariate time series. This algorithm repeatedly performs two steps: the state estimation and the parameter estimation. The Viterbi algorithm is applied to calculating the states, and the parameters are estimated using some meaningful methods. The experiments show that the algorithm is effective with good convergence. In the HMM segmentation approach, the selection of segmentation order for multivariate time series needs to be studied in the future.

**Acknowledgment.** This work is supported by the National Natural Science Foundation of China under Grant 61175041. The authors also thank for reviewers' useful comments, which have improved the presentation.

## REFERENCES

[1] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific Belmont, MA, 1995.
[2] G. D. Forney Jr, The Viterbi algorithm, *Proc. of the IEEE*, vol.61, no.3, pp.268-278, 1973.
[3] A. Gedikli, H. Aksoy, N. E. Unal and A. Kehagias, Modified dynamic programming approach for offline segmentation of long hydrometeorological time series, *Stochastic Environmental Research and Risk Assessment*, vol.24, no.5, pp.547-557, 2010.
[4] H. Guo, X. Liu and L. Song, Dynamic programming approach for segmentation of multivariate time series, *Stochastic Environmental Research and Risk Assessment*, vol.29, no.1, pp.265-273, 2015.
[5] A. Kehagias, A hidden Markov model segmentation procedure for hydrological and environmental time series, *Stochastic Environmental Research and Risk Assessment*, vol.18, no.2, pp.117-130, 2004.
[6] A. Kehagias and V. Fortin, Time series segmentation with shifting means hidden Markov models, *Nonlinear Processes in Geophysics*, vol.13, no.3, pp.339-352, 2006.
[7] A. Kehagias, E. Nidelkou and V. Petridis, A dynamic programming segmentation procedure for hydrological and environmental time series, *Stochastic Environmental Research and Risk Assessment*, vol.20, nos.1-2, pp.77-94, 2006.
[8] L. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, *Proc. of the IEEE*, vol.77, no.2, pp.257-286, 1989.
[9] G. Schwarz et al., Estimating the dimension of a model, *The Annals of Statistics*, vol.6, no.2, pp.461-464, 1978.
[10] A. Vinciarelli, A survey on off-line cursive word recognition, *Pattern Recognition*, vol.35, no.7, pp.1433-1446, 2002.
[11] N. Wang, X. Liu and J. Yin, Improved Gath-Geva clustering for fuzzy segmentation of hydrometeorological time series, *Stochastic Environmental Research and Risk Assessment*, vol.26, no.1, pp.139-155, 2012.