

Research Article

A New Improved Parsimonious Multivariate Markov Chain Model

Chao Wang and Ting-Zhu Huang

School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China

Correspondence should be addressed to Chao Wang; wangchao1321654@163.com

Received 12 November 2012; Revised 17 December 2012; Accepted 2 January 2013

Academic Editor: Marco H. Terra

Copyright © 2013 C. Wang and T.-Z. Huang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a new improved parsimonious multivariate Markov chain model. Moreover, we find a new convergence condition with a new variability to improve the prediction accuracy and minimize the scale of the convergence condition. Numerical experiments illustrate that the new improved parsimonious multivariate Markov chain model with the new convergence condition of the new variability performs better than the improved parsimonious multivariate Markov chain model in prediction.

1. Introduction

The probability distribution of Markov chains plays an important role in a wide range of applications such as telecommunication systems, manufacturing systems, and inventory systems, see, for instance, [1] and the references therein. In recent years, the prediction of multivariate Markov chain models [2] has become more and more useful in many real-world applications: sales demand predictions [2, 3], DNA sequencing [4], and credit and financial data modeling [5]. The major merit of the multivariate Markov chain model is to detect the relations among the sequences and to predict more precisely.

Different models for multiple categorical data sequences are introduced in the following part. A multivariate Markov chain model has been presented in [2]. Ching et al. constructed a new matrix by means of the transition probability matrices among different sequences. To improve prediction accuracy, Ching et al. incorporated positive and negative parts in an improved parsimonious multivariate Markov chain model [5]. Miao and Hambly presented recursive formulas for the default probability distribution which is feasible for computation in this simple version [6]. A more advanced model, namely, higher-order multivariate Markov chain model has been exhibited in [7]. To reduce the number of parameters of the model, a parsimonious higher-order

multivariate Markov chain model has been proposed in [8], where the number of parameters is $\mathcal{O}((n+s)sm^2)$. Certainly, there are many other papers contributing to the multivariate Markov chain models, for example, [1, 9, 10].

With the development of science technologies with their applications, the number of data in sequences become larger, and the results need to be more precise. It is inevitable that a large categorical data sequence group will cause high computational costs, especially using the convergence condition as in [5]. In this paper, a new improved parsimonious multivariate Markov chain model and a new convergence condition with a new variability are presented to enhance the precision of the prediction and save the computational costs.

The rest of the paper is organized as follows. In Section 2, we briefly review several multivariate Markov chain models. In Section 3, a new improved parsimonious multivariate Markov chain model and a new convergence condition with a new variability are presented. Section 4 gives the estimation methods for the parameters of the new improved parsimonious multivariate Markov chain model with different convergence conditions. Numerical experiments with three examples are presented to demonstrate the effectiveness of our proposed model with the new convergence condition of the new variability in Section 5. Finally, concluding remarks are given in Section 6.

2. Review of the Multivariate Markov Chain Models

In this section, we briefly introduce several multivariate Markov chain models, for example, the Markov chain model [3], the multivariate Markov chain model [2], and the improved parsimonious multivariate Markov chain model [5].

2.1. The Markov Chain Model. First, we introduce some definitions of the Markov chain [2, 11]. Let the state set of the categorical data sequences be $\mathcal{M} = \{1, 2, \dots, m\}$. The discrete-time Markov chain with m states satisfies the following relations:

$$\begin{aligned} \text{Prob}(x_{t+1} = \theta_{t+1} \mid x_0 = \theta_0, x_1 = \theta_1, \dots, x_t = \theta_t) \\ = \text{Prob}(x_{t+1} = \theta_{t+1} \mid x_t = \theta_t), \end{aligned} \quad (1)$$

where $\theta_t \in \mathcal{M}$, $t \in \{0, 1, 2, \dots\}$. The conditional probability $\text{Prob}(x_{t+1} = \theta_{t+1} \mid x_t = \theta_t)$ is called the one-step transition probability of the Markov chain. The transition probability is

$$P^{(i,j)} = \text{Prob}(x_{t+1} = i \mid x_t = j), \quad \forall i, j \in \mathcal{M}, \quad (2)$$

and the transition probability matrix is $P^{(i,j)}$. The Markov chain model can be represented as follows:

$$X_{t+1} = PX_t, \quad (3)$$

where

$$\begin{aligned} P = [P^{(i,j)}], \quad 0 \leq P^{(i,j)} \leq 1, \quad \forall i, j \in \mathcal{M}, \\ \sum_{i=1}^m P^{(i,j)} = 1, \quad \forall j \in \mathcal{M}, \end{aligned} \quad (4)$$

X_0 is the initial probability distribution, and $X_t = (x_t^1, x_t^2, \dots, x_t^m)^T$ is the state probability distribution at time t .

2.2. The Multivariate Markov Chain Model. Suppose the number of categorical data sequences $s > 1$. The multivariate Markov chain model [2] is represented as follows:

$$x_{t+1}^{(j)} = \sum_{k=1}^s \lambda_{j,k} P^{(j,k)} x_t^{(k)}, \quad \forall j = 1, 2, \dots, s, \quad t = 0, 1, \dots, \quad (5)$$

where

$$\begin{aligned} \lambda_{j,k} \geq 0, \quad \forall j, k = 1, 2, \dots, s, \\ \sum_{k=1}^s \lambda_{j,k} = 1, \quad \forall j = 1, 2, \dots, s. \end{aligned} \quad (6)$$

Here, $x_0^{(j)}$ is the initial probability distribution of the j th sequence, $x_t^{(k)}$ is the state probability distribution of the k th sequence at time t , and $x_{t+1}^{(j)}$ is the state probability distribution of the j th sequence at time $t + 1$. Here, $P^{(j,k)}$ is

the one-step transition probability from the state in the k th sequence at time t to the state in the j th sequence at time $t + 1$. In matrix form, we have

$$X_{t+1} = \begin{pmatrix} x_{t+1}^{(1)} \\ x_{t+1}^{(2)} \\ \vdots \\ x_{t+1}^{(s)} \end{pmatrix} = \begin{pmatrix} \lambda_{1,1}P^{(1,1)} & \lambda_{1,2}P^{(1,2)} & \dots & \lambda_{1,s}P^{(1,s)} \\ \lambda_{2,1}P^{(2,1)} & \lambda_{2,2}P^{(2,2)} & \dots & \lambda_{2,s}P^{(2,s)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{s,1}P^{(s,1)} & \lambda_{s,2}P^{(s,2)} & \dots & \lambda_{s,s}P^{(s,s)} \end{pmatrix} \times \begin{pmatrix} x_t^{(1)} \\ x_t^{(2)} \\ \vdots \\ x_t^{(s)} \end{pmatrix}. \quad (7)$$

Here, $P^{(j,k)}$ and $\lambda_{j,k}$ can be obtained from the s categorical data sequences and the corresponding linear programming, for details, refer to [2].

2.3. The Improved Parsimonious Multivariate Markov Chain Model. With the same notations as introduced in Section 2.2, we introduce the improved parsimonious multivariate Markov chain model originating from multivariate Markov chain model. Consider

$$Z_{t+1} = \frac{1}{m-1} (\mathbf{e} - X_t), \quad (8)$$

where the factor $1/(m-1)$ is a constant for normalization. Here, \mathbf{e} is the vector of all ones and $m \geq 2$. In matrix form, the improved parsimonious multivariate Markov chain model can be represented as

$$\begin{pmatrix} x_{t+1}^{(1)} \\ x_{t+1}^{(2)} \\ \vdots \\ x_{t+1}^{(s)} \end{pmatrix} = \Lambda^+ \begin{pmatrix} x_t^{(1)} \\ x_t^{(2)} \\ \vdots \\ x_t^{(s)} \end{pmatrix} + \frac{1}{m-1} \Lambda^- \begin{pmatrix} \mathbf{e} - x_t^{(1)} \\ \mathbf{e} - x_t^{(2)} \\ \vdots \\ \mathbf{e} - x_t^{(s)} \end{pmatrix}, \quad (9)$$

where

$$\begin{aligned} \Lambda^+ &= \begin{pmatrix} \lambda_{1,1}P^{(1,1)} & \lambda_{1,2}P^{(1,2)} & \dots & \lambda_{1,s}P^{(1,s)} \\ \lambda_{2,1}P^{(2,1)} & \lambda_{2,2}P^{(2,2)} & \dots & \lambda_{2,s}P^{(2,s)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{s,1}P^{(s,1)} & \lambda_{s,2}P^{(s,2)} & \dots & \lambda_{s,s}P^{(s,s)} \end{pmatrix}, \\ \Lambda^- &= \begin{pmatrix} \lambda_{1,-1}P^{(1,1)} & \lambda_{1,-2}P^{(1,2)} & \dots & \lambda_{1,-s}P^{(1,s)} \\ \lambda_{2,-1}P^{(2,1)} & \lambda_{2,-2}P^{(2,2)} & \dots & \lambda_{2,-s}P^{(2,s)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{s,-1}P^{(s,1)} & \lambda_{s,-2}P^{(s,2)} & \dots & \lambda_{s,-s}P^{(s,s)} \end{pmatrix}, \end{aligned} \quad (10)$$

$$\sum_{j=-s}^s \lambda_{j,k} = 1, \quad \lambda_{j,k} \geq 0, \quad \text{for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s.$$

$\Lambda^+ X_t$ and $\Lambda^- Z_t$ are the positive and negative parts of the transition probability matrices in (9). The above model has two directions to approach the steady solution X .

Let $A = [a_{i,j}] \in \mathbb{R}^{m \times m}$, $B = [b_{i,j}] \in \mathbb{R}^{m \times m}$. If $|a_{i,j}| \geq b_{i,j}$ for all $i, j = \{1, 2, \dots, m\}$ and there exist $i, j \in \{1, 2, \dots, m\}$ satisfying $|a_{i,j}| > b_{i,j}$, then we note that $|A| > B$.

Lemma 1 (see [12]). Let $A \in \mathbb{R}^{m \times m}$ be a nonnegative and irreducible matrix, $B \in \mathbb{C}^{m \times m}$ a complex matrix, and λ an eigenvalue of B . If $|A| > B$, then $\rho(A) > |\lambda|$.

3. A New Improved Parsimonious Multivariate Markov Chain Model

In this section, we propose a new improved parsimonious multivariate Markov chain model and a new convergence condition with a new variability.

In the new model, the state probability distribution of the j th sequence at time $t + 1$ depends on the state probability distribution of all the sequences at time t . Let the number of categorical data sequences be $s > 1$, m the number of states in every sequences, $\beta < 1/m$, then the new improved parsimonious multivariate Markov chain model can be represented as follows:

$$X_{t+1}^{(j)} = \sum_{k=1}^s \lambda_{j,k} P^{(j,k)} X_t^{(k)} + \frac{1}{m\beta - 1} \sum_{k=1}^s \lambda_{j,-k} P^{(j,k)} (\beta \mathbf{e} - X_t^{(k)}), \quad (11)$$

where $X_0^{(k)}$ is the initial probability distributions of the k th sequence and

$$\sum_{k=-s}^s \lambda_{j,k} = 1, \quad \lambda_{j,k} \geq 0, \text{ for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s. \quad (12)$$

Here, $X_t^{(k)}$ is the state probability distribution of the k th sequence at time t , $P^{(j,k)}$ is the one-step transition probability matrix from the state in the k th sequence at time t to the state in the j th sequence at time $t + 1$. Here, $X_{t+1}^{(j)}$ is the state probability distribution of the j th sequence at time $t + 1$. Let

$$X_{t+1} = \left((X_{t+1}^{(1)})^T, (X_{t+1}^{(2)})^T, \dots, (X_{t+1}^{(s)})^T \right)^T \in \mathbb{R}^{ms \times 1}, \quad (13)$$

then the new improved parsimonious multivariate Markov chain model in matrix form is

$$X_{t+1} = \Lambda^+ X_t + \frac{1}{m\beta - 1} \Lambda^- (\beta \mathbf{e} - X_t), \quad (14)$$

which also can be represented as

$$\begin{pmatrix} X_{t+1}^{(1)} \\ X_{t+1}^{(2)} \\ \vdots \\ X_{t+1}^{(s)} \end{pmatrix} = \Lambda^+ \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \vdots \\ X_t^{(s)} \end{pmatrix} + \frac{1}{m\beta - 1} \Lambda^- \begin{pmatrix} \beta \mathbf{e} - X_t^{(1)} \\ \beta \mathbf{e} - X_t^{(2)} \\ \vdots \\ \beta \mathbf{e} - X_t^{(s)} \end{pmatrix}, \quad (15)$$

where

$$\Lambda^+ = \begin{pmatrix} \lambda_{1,1} P^{(1,1)} & \lambda_{1,2} P^{(1,2)} & \dots & \lambda_{1,s} P^{(1,s)} \\ \lambda_{2,1} P^{(2,1)} & \lambda_{2,2} P^{(2,2)} & \dots & \lambda_{2,s} P^{(2,s)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{s,1} P^{(s,1)} & \lambda_{s,2} P^{(s,2)} & \dots & \lambda_{s,s} P^{(s,s)} \end{pmatrix},$$

$$\Lambda^- = \begin{pmatrix} \lambda_{1,-1} P^{(1,1)} & \lambda_{1,-2} P^{(1,2)} & \dots & \lambda_{1,-s} P^{(1,s)} \\ \lambda_{2,-1} P^{(2,1)} & \lambda_{2,-2} P^{(2,2)} & \dots & \lambda_{2,-s} P^{(2,s)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{s,-1} P^{(s,1)} & \lambda_{s,-2} P^{(s,2)} & \dots & \lambda_{s,-s} P^{(s,s)} \end{pmatrix}, \quad (16)$$

$$\sum_{k=-s}^s \lambda_{j,k} = 1, \quad \lambda_{j,k} \geq 0, \text{ for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s.$$

$\Lambda^+ X_t$ and $(1/(m\beta - 1))\Lambda^-(\beta \mathbf{e} - X_t)$ are, respectively, the positive and negative parts of the transition probability matrices in (14) where $\|X_t\|_1 = \|(1/(m\beta - 1))(\beta \mathbf{e} - X_t)\|_1 = 1$. Each column sum of $P^{(j,k)}$ is equal to one.

From (14), after t times iterations, it has

$$\begin{aligned} X_{t+1} &= \Lambda^+ X_t + \frac{1}{m\beta - 1} \Lambda^- (\beta \mathbf{e} - X_t) \\ &= \left(\Lambda^+ - \frac{1}{m\beta - 1} \Lambda^- \right) X_t + \frac{\beta}{m\beta - 1} \Lambda^- \mathbf{e} \\ &= \left(\Lambda^+ - \frac{1}{m\beta - 1} \Lambda^- \right) \left(\left(\Lambda^+ - \frac{1}{m\beta - 1} \Lambda^- \right) X_{t-1} \right. \\ &\quad \left. + \frac{\beta}{m\beta - 1} \Lambda^- \mathbf{e} \right) + \frac{\beta}{m\beta - 1} \Lambda^- \mathbf{e} \\ &= \left(\Lambda^+ - \frac{1}{m\beta - 1} \Lambda^- \right)^2 X_{t-1} \\ &\quad + \left(\Lambda^+ - \frac{1}{m\beta - 1} \Lambda^- \right) \frac{\beta}{m\beta - 1} \Lambda^- \mathbf{e} + \frac{\beta}{m\beta - 1} \Lambda^- \mathbf{e} \\ &= \left(\Lambda^+ - \frac{1}{m\beta - 1} \Lambda^- \right)^{t+1} X_0 \\ &\quad + \sum_{k=0}^t \left(\Lambda^+ - \frac{1}{m\beta - 1} \Lambda^- \right)^k \frac{\beta}{m\beta - 1} \Lambda^- \mathbf{e}. \end{aligned} \quad (17)$$

If $\rho(\Lambda^+ - (1/(m\beta - 1))\Lambda^-) < 1$, the iteration of the new improved parsimonious multivariate Markov chain model is convergent. For finding a more simple and efficient convergence condition from the point of view of properties of special matrices, we get the following theorem.

Theorem 2. In the new improved parsimonious multivariate Markov chain model, if $\Lambda^+ > (\gamma/(m\beta - 1))\Lambda^-$ where $\gamma \geq 1$ and $\beta < 1/m$, then the iteration of the new improved parsimonious multivariate Markov chain model is convergent.

Proof. Because $\Lambda^+ > (\gamma/(m\beta - 1))\Lambda^-$, we obtain

$$\Lambda^+ - \frac{\gamma}{m\beta - 1}\Lambda^- > 0. \quad (18)$$

For $\Lambda^+, \Lambda^- > 0$, it has

$$-\Lambda^+ < 0 < \Lambda^+ - \frac{\gamma}{m\beta - 1}\Lambda^- < \Lambda^+. \quad (19)$$

Then, we obtain

$$\left| \Lambda^+ - \frac{\gamma}{m\beta - 1}\Lambda^- \right| < \Lambda^+. \quad (20)$$

From Lemma 1, it has

$$\rho\left(\Lambda^+ - \frac{\gamma}{m\beta - 1}\Lambda^-\right) < \rho(\Lambda^+). \quad (21)$$

In the new model, $\sum_{k=-s}^s \lambda_{j,k} = 1, 0 \leq \lambda_{j,k} \leq 1$, and $0 \leq P^{(j,k)} \leq 1$ for all $1 \leq j \leq s, 1 \leq |k| \leq s$. It suffices to prove that $\|\Lambda^+\|_1 + \|\Lambda^-\| \leq 1$. Then, we obtain

$$\rho\left(\Lambda^+ - \frac{1}{m\beta - 1}\Lambda^-\right) < \rho(\Lambda^+) \leq 1. \quad (22)$$

The new improved parsimonious multivariate Markov chain model is convergent. \square

4. Estimation of the Parameters of the New Improved Parsimonious Multivariate Markov Chain Model

In this section, we estimate the parameters of the new improved parsimonious multivariate Markov chain model in the new convergence condition with the new variability which has been proved in Theorem 2. The transition probability matrices $P^{(j,k)}$ are estimated at first. If the data sequences are given and the state set is $\mathcal{M} = \{1, 2, \dots, m\}$, $f_{i_j, i_k}^{(j,k)}$ is the frequency from the i_k state in the k th sequence at time t to the i_j state in the j th sequence at time $t + 1$ with $i_j, i_k \in \mathcal{M}$, then the transition frequency matrix $F^{(j,k)}$ can be constructed as follows:

$$F^{(j,k)} = \begin{pmatrix} f_{1,1}^{(j,k)} & f_{1,2}^{(j,k)} & \cdots & f_{1,m}^{(j,k)} \\ f_{2,1}^{(j,k)} & f_{2,2}^{(j,k)} & \cdots & f_{2,m}^{(j,k)} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m,1}^{(j,k)} & f_{m,2}^{(j,k)} & \cdots & f_{m,m}^{(j,k)} \end{pmatrix}_{m \times m}. \quad (23)$$

Here, $P^{(j,k)}$ can be obtained by normalizing the frequency transition probability matrix as

$$P^{(j,k)} = \begin{pmatrix} p_{1,1}^{(j,k)} & p_{1,2}^{(j,k)} & \cdots & p_{1,m}^{(j,k)} \\ p_{2,1}^{(j,k)} & p_{2,2}^{(j,k)} & \cdots & p_{2,m}^{(j,k)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m,1}^{(j,k)} & p_{m,2}^{(j,k)} & \cdots & p_{m,m}^{(j,k)} \end{pmatrix}_{m \times m}, \quad (24)$$

where

$$p_{i_j, i_k}^{(j,k)} = \begin{cases} \frac{f_{i_j, i_k}^{(j,k)}}{\sum_{i_j=1}^m f_{i_j, i_k}^{(j,k)}} & \text{if } \sum_{i_j=1}^m f_{i_j, i_k}^{(j,k)} \neq 0, \\ \frac{1}{m} & \text{otherwise.} \end{cases} \quad (25)$$

Subsequently, the way of estimating the parameter $\lambda_{j,k}$ is introduced. X_t is a joint state probability distribution of the new improved parsimonious multivariate Markov chain model at t and can be represented as

$$X_t = \left((X_t^{(1)})^T, (X_t^{(2)})^T, \dots, (X_t^{(s)})^T \right)^T \in \mathbb{R}^{ms \times 1}, \quad (26)$$

which satisfies

$$X_{t+1} = \Lambda^+ X_t + \frac{1}{m\beta - 1} \Lambda^- (\beta \mathbf{e} - X_t), \quad (27)$$

where Λ^+, Λ^- have been denoted in Section 3 satisfying

$$\sum_{j=-s}^s \lambda_{j,k} = 1, \quad \lambda_{j,k} \geq 0, \text{ for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s, \quad (28)$$

and $\beta < 1/m$. Based on the idea of the convergence condition in [5], the iteration matrix of the new improved parsimonious multivariate Markov chain model M_s satisfies

$$\|M_s\|_\infty \leq \max_{1 \leq k \leq s} \left\{ m \left| \lambda_{k,k} - \frac{\lambda_{k,-k}}{(m\beta - 1)} \right| + \sum_{k \neq i} \left| \lambda_{k,i} - \frac{\lambda_{k,-i}}{(m\beta - 1)} \right| \right\}. \quad (29)$$

By imposing an upper bound $\alpha < 1$, the convergence condition of the new improved parsimonious multivariate Markov chain model is

$$m \left| \lambda_{k,k} - \frac{\lambda_{k,-k}}{(m\beta - 1)} \right| + \sum_{k \neq i} \left| \lambda_{k,i} - \frac{\lambda_{k,-i}}{(m\beta - 1)} \right| \leq \alpha \quad (30)$$

for $i = 1, 2, \dots, s$.

Then the new improved parsimonious multivariate Markov chain model in this convergence condition can be represented as a set of s linear programming problems

$$\min_{\lambda_{j,k}} \sum_j \omega_j \quad (31)$$

subject to

$$\begin{aligned}
 b_{j,k} &= \sum_{k=1}^s \left(\left(\lambda_{j,k} - \frac{\lambda_{j,-k}}{m-1} \right) \Delta_{j,k} X_t + \frac{1}{m-1} \lambda_{j,k} \Delta_{j,k} \mathbf{e} \right), \\
 \omega_j &\geq [b_{j,k} - X^{(j)}], \\
 \omega_j &\geq -[b_{j,k} - X^{(j)}], \\
 \sum_{k=-s}^s \lambda_{j,k} &= 1 \quad \text{for } \forall 1 \leq j \leq s, \\
 \lambda_{j,k} &\geq 0 \quad \text{for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s, \\
 A_j \Lambda_j &\leq \alpha \mathbf{e} \quad \text{for } \forall 1 \leq j \leq s, \beta < \frac{1}{m},
 \end{aligned} \tag{32}$$

where

$$\begin{aligned}
 \Delta_{j,k} &= \begin{cases} P^{(j,j)}, & \text{if } j = k, \\ I, & \text{if } j \neq k, \end{cases} \\
 \Lambda_j &= (\lambda_{j,1}, \dots, \lambda_{j,s}, \lambda_{j,-1}, \dots, \lambda_{j,-s})^T, \\
 A_j &= [A_{1j}, A_{2j}], \\
 A_{1j} &= \begin{pmatrix} 1 & \cdots & m & \cdots & 1 & 1 \\ 1 & \cdots & m & \cdots & 1 & -1 \\ 1 & \cdots & m & \cdots & -1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & \cdots & -m & \cdots & -1 & 1 \\ -1 & \cdots & -m & \cdots & -1 & -1 \end{pmatrix}, \\
 A_{2j} &= -\frac{1}{m\beta - 1} A_{1j}.
 \end{aligned} \tag{33}$$

Here, A_{1j} covers all of the rows of each component taking one of the two possible values, 1 and -1 , particularly, in j th column taking m and $-m$. Then $A = [A_j]$ has $s \times 2^s$ rows. This convergence condition is viable only when the sequence group is not large.

To speed up the convergence of the new improved parsimonious multivariate Markov chain model and minimize the scale of the convergence condition, we apply a new convergence condition $B > (\gamma/(m\beta - 1))C$ with $\gamma \geq 1$ and $\beta < 1/m$ for the new model. Then, the form of the new improved parsimonious multivariate Markov chain model can be represented as

$$\min_{\lambda_{j,k}} |f^{(j)} - X^{(j)}|_j \tag{34}$$

subject to:

$$\begin{aligned}
 f^{(j)} &= \sum_{k=1}^s \lambda_{j,k} P^{(j,k)} X^{(k)} + \frac{1}{m\beta - 1} \sum_{k=1}^s \lambda_{j,k} P^{(j,k)} (\beta \mathbf{e} - X^{(k)}), \\
 \sum_{k=-s}^s \lambda_{j,k} &= 1 \quad \text{for } \forall 1 \leq |k| \leq s, \\
 \lambda_{j,k} &\geq 0 \quad \text{for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s, \\
 \lambda_{j,k} &> \frac{\gamma}{m\beta - 1} \lambda_{j,-k} \quad \text{for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s, \\
 \gamma &\geq 1, \beta < \frac{1}{m},
 \end{aligned} \tag{35}$$

where $[\cdot]_i$ is the i th entry of the vector. Certainly, the optimization problem can also be represented as a linear programming problem:

$$\min_{\lambda_{j,k}} \sum_j \omega_j \tag{36}$$

subject to

$$\begin{aligned}
 f^{(j)} &= \sum_{k=1}^s \lambda_{j,k} P^{(j,k)} X^{(k)} + \frac{1}{m\beta - 1} \sum_{k=1}^s \lambda_{j,k} P^{(j,k)} (\beta \mathbf{e} - X^{(k)}), \\
 \omega_j &\geq [f^{(j)} - X^{(j)}], \\
 \omega_j &\geq -[f^{(j)} - X^{(j)}], \\
 \sum_{k=-s}^s \lambda_{j,k} &= 1 \quad \text{for } \forall 1 \leq |k| \leq s, \\
 \lambda_{j,k} &\geq 0 \quad \text{for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s, \\
 \lambda_{j,k} &> \frac{\gamma}{m\beta - 1} \lambda_{j,-k} \quad \text{for } \forall 1 \leq j \leq s, 1 \leq |k| \leq s, \\
 \gamma &\geq 1, \beta < \frac{1}{m}.
 \end{aligned} \tag{37}$$

5. Numerical Experiments

In this section, numerical experiments with three examples of different improved parsimonious multivariate Markov chain models with different convergence conditions are reported.

Noting that the new improved parsimonious multivariate Markov chain model with the original convergence condition [5] is “IPM1,” (especially, the new improved parsimonious multivariate Markov chain model is an improved multivariate Markov chain model when we choose $\beta = 1$), the new improved parsimonious multivariate Markov chain model with the new convergence condition of the new variability is “IPM2,” the convergence factor of the original convergence condition is “ α ,” and the variabilities of the new convergence condition are “ β ” and “ γ .” The stopping criterion can be found

in Matlab order of *linprog*. We add a notation “*” at the back of data in Tables 1, 2, 3, and 4 when the stopping criterion is satisfied but the accuracy is not reached.

5.1. *Example 1.* There are three categorical data sequences [13]:

$$\begin{aligned} S_1 &= \{1, 2, 2, 1, 3, 3, 2, 3, 1, 3, 3, 2, 2, 3, 2, 1, 2, 1, 2, 2\}, \\ S_2 &= \{2, 2, 1, 1, 2, 1, 2, 1, 2, 2, 2, 3, 2, 2, 1, 2, 1, 2, 2\}, \\ S_3 &= \{3, 1, 1, 2, 2, 1, 2, 3, 1, 1, 3, 3, 3, 2, 3, 2, 3, 3, 1, 2\}. \end{aligned} \quad (38)$$

IPM1 with $\alpha = 0.7$ [5] can be represented as

$$\begin{aligned} x_{r+1}^{(1)} &= 0.3000P^{(1,2)}x_r^{(2)} + 0.4737P^{(1,3)}(\mathbf{e} - x_r^{(3)}), \\ x_{r+1}^{(2)} &= 0.0088P^{(2,1)}x_r^{(1)} + 0.1912P^{(2,1)}(\mathbf{e} - x_r^{(1)}) \\ &\quad + 0.8000P^{(2,3)}x_r^{(3)}, \\ x_{r+1}^{(3)} &= 1.0000P^{(3,3)}x_r^{(3)}. \end{aligned} \quad (39)$$

IPM2 with $\beta = 1, \gamma = 3$ is

$$\begin{aligned} x_{r+1}^{(1)} &= 1.0000P^{(1,1)}x_r^{(1)}, \\ x_{r+1}^{(2)} &= 0.0054P^{(2,1)}x_r^{(1)} + 0.0385P^{(2,1)}(\mathbf{e} - x_r^{(1)}) \\ &\quad + 0.8653P^{(2,2)}x_r^{(2)} \\ &\quad + 0.0908P^{(2,2)}(\mathbf{e} - x_r^{(2)}), \\ x_{r+1}^{(3)} &= 0.0771P^{(3,3)}x_r^{(3)} + 0.9229P^{(3,3)}(\mathbf{e} - x_r^{(3)}). \end{aligned} \quad (40)$$

In IPM2 with $\beta = 1.1, \gamma = 3$, it has

$$\begin{aligned} x_{r+1}^{(1)} &= 1.0000P^{(1,1)}x_r^{(1)}, \\ x_{r+1}^{(2)} &= 0.0147P^{(2,1)}(\mathbf{e} - x_r^{(1)}) + 0.8783P^{(2,2)}x_r^{(2)} \\ &\quad + 0.1070P^{(2,2)}(\mathbf{e} - x_r^{(2)}), \\ x_{r+1}^{(3)} &= 0.0350P^{(3,3)}x_r^{(3)} + 0.9650P^{(3,3)}(\mathbf{e} - x_r^{(3)}). \end{aligned} \quad (41)$$

5.2. *Example 2.* Let the three categorical data sequences be

$$\begin{aligned} S_1 &= \{1, 4, 3, 3, 1, 1, 3, 4, 4, 3, 2, 3\}, \\ S_2 &= \{2, 3, 4, 4, 4, 2, 3, 3, 1, 4, 3, 2\}, \\ S_3 &= \{3, 3, 1, 1, 3, 1, 4, 3, 2, 4, 3, 2\}. \end{aligned} \quad (42)$$

Suppose that $\bar{X}_t = [(\bar{X}_t^{(1)})^T, (\bar{X}_t^{(2)})^T, \dots, (\bar{X}_t^{(s)})^T]^T$ is the prediction probability at time t and $X_t = [(X_t^{(1)})^T, (X_t^{(2)})^T, \dots, (X_t^{(s)})^T]^T$ is the fact value at time t where $X_t^{(i)} = e_{(m_i)} = \{0, \dots, 0, 1, 0, \dots, 0\}^T \in \mathbb{R}^{1 \times m}$, $i \in \{1, \dots, s\}$, m_i is the fact state at time t in the i th categorical

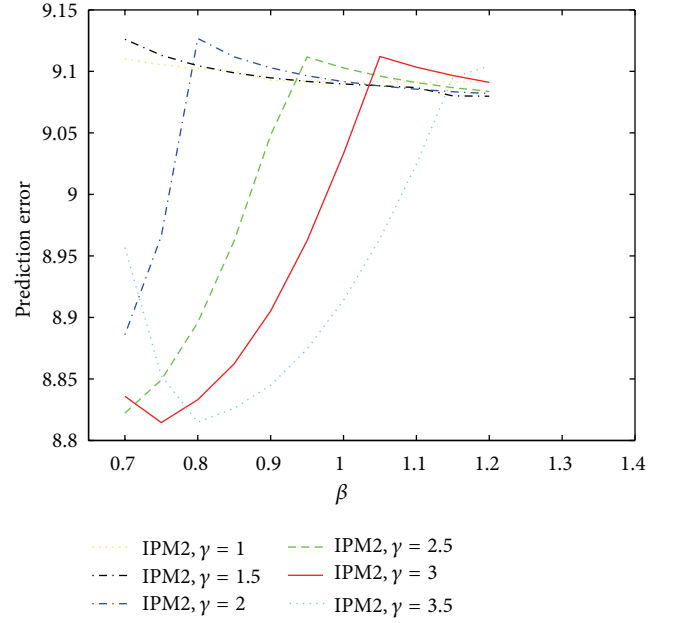


FIGURE 1: The prediction errors of the new improved parsimonious multivariate Markov chain model with our new convergence condition of the new variability in Example 2.

data sequence. “nA” is the number of the categorical data in one sequence and “pe” is the prediction error of the models which can be estimated by the equation:

$$pe = \sum_{t=4}^{nA} \|\bar{X}_t - X_t\|_2. \quad (43)$$

In Table 1, the prediction errors of the new improved parsimonious multivariate Markov chain model when $\beta = 1.2$ is better than the prediction errors of other values of β . Table 1 illustrates the efficiency of the new improved parsimonious multivariate Markov chain model when $\beta = 1.2$ in the original convergence condition [5].

In Table 2, numerical experiments on the prediction errors of the new improved parsimonious multivariate Markov chain model of the new convergence condition with the new variability “ γ ” are reported. The best performance of the prediction errors of the new model with the new convergence condition of the new variability is 8.6790 when we choose $\beta = 0.8$ and $\gamma = 3$. In different cases of β , the best prediction errors of the new model with the new convergence condition of the new variability are in the diagonal line of the result matrix between $\gamma = 2.5$ and $\gamma = 5$.

For comparing the performances of the new improved parsimonious multivariate Markov chain model with the new convergence condition more clearly, we present Figure 1.

To compare the performances of the new improve parsimonious multivariate Markov chain model in different convergence conditions, we show Figure 2 with the data of Tables 1 and 2. Obviously, the results of the new model in the new convergence condition with the new variability are

TABLE 1: Prediction errors of the new improved parsimonious multivariate Markov chain model with original convergence condition in Example 2.

	$\beta = 0.7$	$\beta = 0.8$	$\beta = 0.9$	$\beta = 1$	$\beta = 1.1$	$\beta = 1.2$
IPM1 when $\alpha = 0.1$	31.1845*	31.0831*	29.8349*	29.8349*	31.0906*	31.0853*
IPM1 when $\alpha = 0.2$	27.0472*	26.9434*	26.8745*	26.8745*	26.9486*	26.8462*
IPM1 when $\alpha = 0.3$	22.8309*	22.7298*	22.7296*	22.7302*	22.7321*	22.7310*
IPM1 when $\alpha = 0.4$	18.6215*	18.5266*	18.5256*	18.5256*	18.5282*	18.5211*
IPM1 when $\alpha = 0.5$	14.3298*	14.2594*	14.2614*	14.2614*	14.2586*	14.2705*
IPM1 when $\alpha = 0.6$	10.4531*	10.2961*	10.3231*	10.2962*	10.2795*	10.2971*
IPM1 when $\alpha = 0.7$	9.9017	9.8994	9.8872	9.8879	9.8896	9.8552
IPM1 when $\alpha = 0.8$	10.1173	10.1440	10.1255	10.1271	10.1283	10.0536
IPM1 when $\alpha = 0.9$	10.3910	10.4131	10.3945	10.3886	10.3803	10.2726
IPM1 when $\alpha = 1$	10.4059	10.4268	10.4059	10.4059	10.4059	10.2590

TABLE 2: Prediction errors of the new improved parsimonious multivariate Markov chain model with the new convergence condition of the new variability in Example 2.

	$\beta = 0.7$	$\beta = 0.8$	$\beta = 0.9$	$\beta = 1$	$\beta = 1.1$	$\beta = 1.2$
IPM2 when $\gamma = 1$	8.9588	9.1114	9.0964	9.0894	9.0860	9.0778
IPM2 when $\gamma = 1.5$	8.8126	8.8788	9.0326	9.1028	9.0904	9.0832
IPM2 when $\gamma = 2$	8.9746	8.8142	8.8544	8.9447	9.0892	9.0974
IPM2 when $\gamma = 2.5$	8.6957	8.9355	8.8167	8.8449	8.9054	9.0014
IPM2 when $\gamma = 3$	8.9006	8.6790	8.7050	8.8189	8.8410	8.8849
IPM2 when $\gamma = 3.5$	9.4994	8.7730	8.6838	8.7168	8.8209	8.8393
IPM2 when $\gamma = 4$	9.1024*	9.1909	8.7041	8.6911	8.7275	8.8228
IPM2 when $\gamma = 4.5$	9.3544*	9.5649*	9.0116	8.6994	8.6993	8.7371
IPM2 when $\gamma = 5$	12.3474*	9.0286*	9.3882	8.8994	8.7029	8.7076
IPM2 when $\gamma = 5.5$	13.5788*	9.4195*	9.4562*	9.1952	8.8252	8.7046

much better than those of the new model in the original convergence condition.

5.3. An Application to Sales Demand Predictions. In this part, the sales demand sequences are presented to show the effectiveness of the new improved parsimonious multivariate Markov chain model of the new convergence condition with the new variability. Since the requirement of the market fluctuates heavily, the production planning and the inventory control directly affect the estate cost. Thus, studying the interplay between the storage space requirement and the overall growing sales demand is a pressing issue for the company. Suppose that the products are classified into six possible states (1, 2, 3, 4, 5, 6), for example, 1 = no sales volume, 2 = very low sales volume, 3 = low sales volume, 4 = standard sales volume, 5 = fast sales volume, and 6 = very fast sales volume. The data of customer's sales demand of five important products can be found in [3].

By computing the proportion of the occurrence of each state in the sequences, the initial probability distributions of the five categorical data sequences are

$$x_0^{(1)} = (0.0818, 0.4052, 0.0483, 0.0335, 0.0037, 0.4275)^T,$$

$$x_0^{(2)} = (0.3680, 0.1970, 0.0335, 0.0000, 0.0037, 0.3978)^T,$$

$$x_0^{(3)} = (0.1450, 0.2045, 0.0186, 0.0000, 0.0037, 0.6283)^T,$$

$$x_0^{(4)} = (0.0000, 0.3569, 0.1338, 0.1896, 0.0632, 0.2565)^T,$$

$$x_0^{(5)} = (0.0000, 0.3569, 0.1227, 0.2268, 0.0520, 0.2416)^T.$$

(44)

The transition probability matrix $P^{(j,k)}$ can be obtained after normalizing the transition frequency matrix. By solving the linear programming problem corresponding to the new improved parsimonious multivariate Markov chain model with the new convergence condition where $\beta = 0.7$ and $\gamma = 4$, $\lambda_{i,j}$ is obtained and the model is presented as follows:

$$\begin{aligned} x_{r+1}^{(1)} = & 0.4359P^{(1,1)}x_r^{(1)} + 0.0193P^{(1,2)}(\mathbf{e} * 0.7 - x_r^{(2)}) \\ & + 0.1210P^{(1,3)}(\mathbf{e} * 0.7 - x_r^{(3)}) \\ & + 0.0359P^{(1,5)}(\mathbf{e} * 0.7 - x_r^{(5)}), \end{aligned}$$

$$\begin{aligned} x_{r+1}^{(2)} = & 0.4950P^{(2,1)}x_r^{(1)} + 0.2416P^{(2,3)}x_r^{(3)} + 0.1189P^{(2,4)}x_r^{(4)} \\ & + 0.1445P^{(2,5)}(\mathbf{e} * 0.7 - x_r^{(5)}), \end{aligned}$$

$$x_{r+1}^{(3)} = 0.7748P^{(3,1)}x_r^{(1)} + 0.2252P^{(3,2)}x_r^{(2)},$$

TABLE 3: Prediction errors of the new improved parsimonious multivariate Markov chain model with the original convergence condition in sales demand prediction.

	$\beta = 0.7$	$\beta = 0.8$	$\beta = 0.9$	$\beta = 1$	$\beta = 1.1$	$\beta = 1.2$
IPM1 when $\alpha = 0.1$	1624.2*	1628.5*	1631.7*	1634.1*	1636.1*	1637.7*
IPM1 when $\alpha = 0.2$	1354.0*	1355.8*	1359.3*	1361.1*	1363.8*	1365.1*
IPM1 when $\alpha = 0.3$	1083.7*	1086.4*	1088.4*	1089.9*	1091.1*	1092.1*
IPM1 when $\alpha = 0.4$	809.2156*	811.0515*	812.4306*	813.5030*	814.3598*	815.0606*
IPM1 when $\alpha = 0.5$	543.7497*	545.0841*	546.0720*	546.8347*	547.4426*	547.9362*
IPM1 when $\alpha = 0.6$	366.9729	366.9684	366.9580	366.9577	366.9533	366.9589
IPM1 when $\alpha = 0.7$	367.3149	367.2101	367.1579	367.1367	367.1448	367.1498
IPM1 when $\alpha = 0.8$	367.5954	367.5916	367.5874	367.5769	367.5771	367.5766
IPM1 when $\alpha = 0.9$	368.2437	368.2422	368.2464	368.2530	368.2550	368.2639
IPM1 when $\alpha = 1$	369.1255	369.1145	369.1161	369.1119	369.1088	369.1062

TABLE 4: Prediction errors of the new improved parsimonious multivariate Markov chain model with the new convergence condition of the new variability in sales demand prediction.

	$\beta = 0.7$	$\beta = 0.8$	$\beta = 0.9$	$\beta = 1$	$\beta = 1.1$	$\beta = 1.2$
IPM2 when $\gamma = 1$	348.9387	348.8616	348.6692	348.4583	346.8533	347.9374
IPM2 when $\gamma = 2$	341.3333	348.4608	348.5667	348.6407	348.6676	348.6457
IPM2 when $\gamma = 3$	333.1122	334.4701	339.0412	343.5102	348.3476	348.3801
IPM2 when $\gamma = 4$	334.2837	331.1873	333.7044	334.5513	337.5047	340.6662
IPM2 when $\gamma = 5$	333.1028*	332.8355	330.7546	332.4308	333.8498	334.6122
IPM2 when $\gamma = 6$	327.5181*	333.2228*	335.7616	331.0009	331.4727	332.7485
IPM2 when $\gamma = 7$	481.5991*	327.9535*	333.2418*	335.7036	334.4605	331.1581
IPM2 when $\gamma = 8$	550.2992*	457.7242*	329.7488*	333.3142*	335.7388	333.6250
IPM2 when $\gamma = 9$	618.8261*	521.8871*	452.3191*	331.4777*	333.4008*	335.0657
IPM2 when $\gamma = 10$	527.6297*	579.7398*	501.1963*	442.6577*	331.6313*	333.5236*

$$\begin{aligned}
x_{r+1}^{(4)} &= 0.3554P^{(4,1)}x_r^{(1)} + 0.1354P^{(4,3)}(\mathbf{e} * 0.7 - x_r^{(3)}) \\
&\quad + 0.0661P^{(4,5)}(\mathbf{e} * 0.7 - x_r^{(5)}), \\
x_{r+1}^{(5)} &= 0.0288P^{(5,1)}(\mathbf{e} * 0.7 - x_r^{(1)}) + 0.1121P^{(5,3)}x_r^{(3)} \\
&\quad + 0.1927P^{(5,3)}(\mathbf{e} * 0.7 - x_r^{(3)}) \\
&\quad + 0.0560P^{(5,4)}(\mathbf{e} * 0.7 - x_r^{(4)}).
\end{aligned} \tag{45}$$

Here, the prediction error equation are represented as

$$pe = \sum_{t=9}^{nA} \|\bar{X}_t - X_t\|_2. \tag{46}$$

In Table 3, we present the prediction errors of the new improved parsimonious multivariate Markov chain model with the original convergence condition [5]. From the data of Table 3, we get Figure 3. In Figure 3, it is obvious that the new improved parsimonious multivariate Markov chain model of the original convergence condition performs the best when we choose $\alpha = 0.6$. The smaller α we choose, the better error prediction we get.

In Table 4, the best performance of the prediction errors of the new improved parsimonious multivariate model with

the new convergence condition in sales demand prediction is 331.1873 when $\beta = 0.8$ and $\gamma = 4$. In different cases of β , the best prediction errors are in the diagonal line of the result matrix between $\gamma = 3$ and $\gamma = 8$.

For comparing the performances of the new improved parsimonious multivariate model with the new convergence condition of the new variability more clearly, we present Figure 4 where the data are extracted from Table 4. As the value of β increases, the parameter γ of the best performances of the prediction results increases. The best performances of the IPM2 are almost the same in different cases of β .

With the data of Tables 3 and 4, we get Figure 5. It illustrates the benefits of the new parsimonious multivariate Markov chain model of the new convergence condition with the new variability in prediction accuracy.

6. Conclusion

In this paper, we present a new improved parsimonious multivariate Markov chain model and a new convergence condition with a new variability, which can enhance the prediction accuracy of the models and save the computational estate. Numerical experiments with three examples illustrate that the new improved parsimonious multivariate Markov chain model of the new convergence condition with the new variability performs better than the improved parsimonious

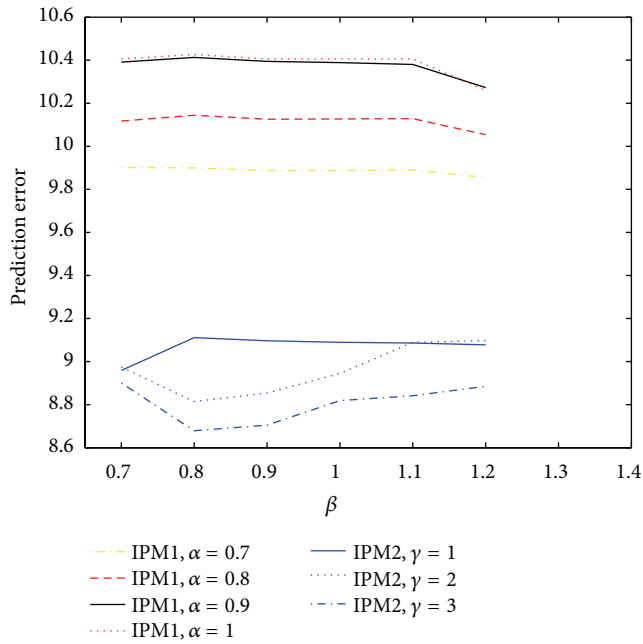


FIGURE 2: The prediction errors of the new improved parsimonious multivariate Markov chain model with the original convergence condition and the new convergence condition of the new variability in Example 2.

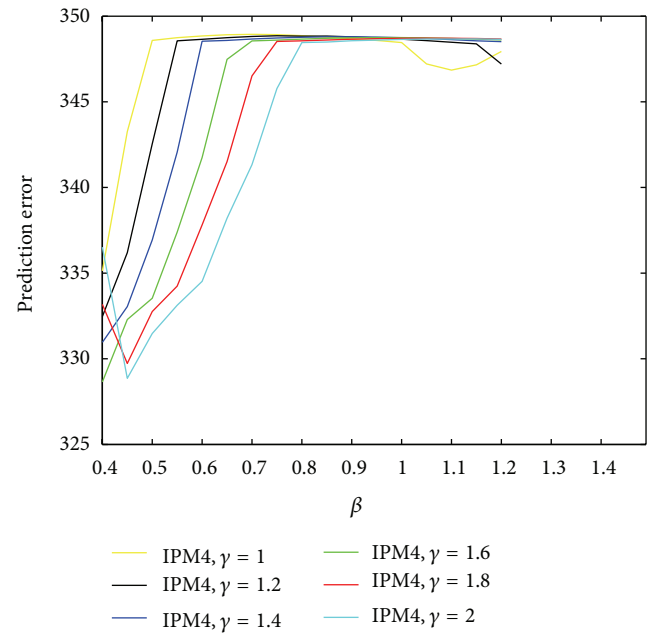


FIGURE 4: The prediction errors of the new improved parsimonious multivariate Markov chain model with the new convergence condition of the new variability in sales demand prediction.

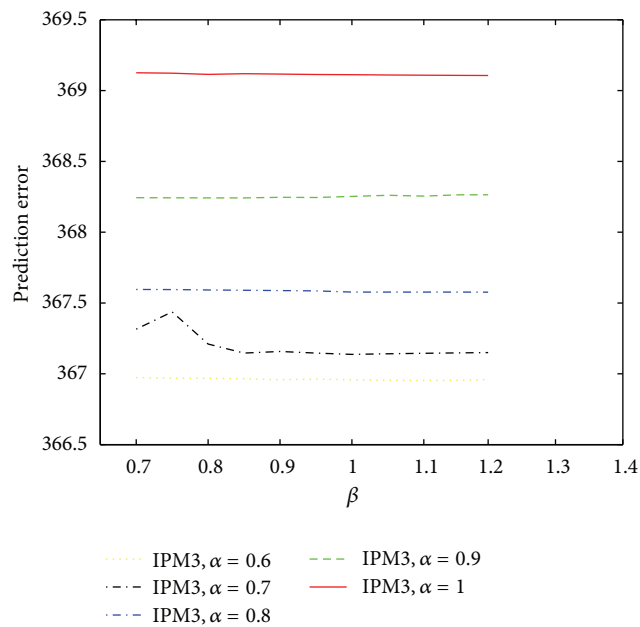


FIGURE 3: Prediction errors of the new improved parsimonious multivariate Markov chain model with the original convergence condition in sales demand prediction.

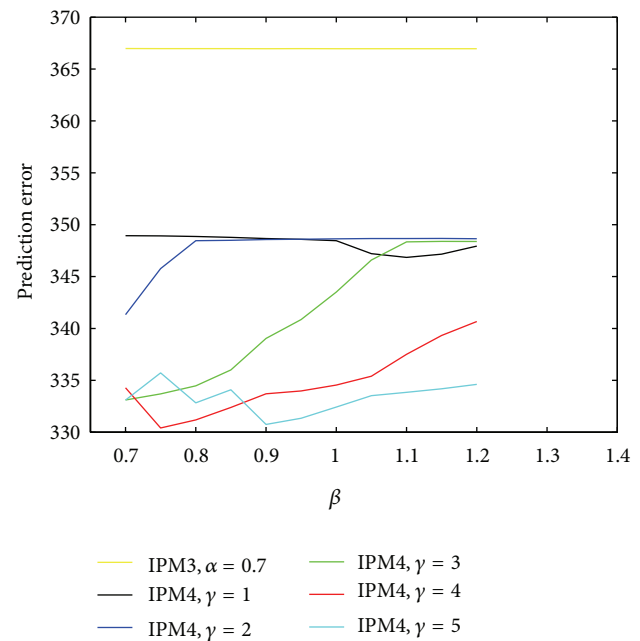


FIGURE 5: The prediction errors of the new improved parsimonious multivariate Markov chain model with the original convergence condition and the new convergence condition of the new variability in sales demand prediction.

multivariate Markov chain model with the original convergence condition. Certainly, our new model can also be applied into credit risk and other research areas.

Acknowledgments

This research is supported by the Chinese Universities Specialized Research Fund for the Doctoral Program (20110185110020) and Sichuan Province Sci. & Tech. Research Project (12ZC1802).

References

- [1] W.-K. Ching and M. K. Ng, *Markov Chains: Models, Algorithms and Applications*, International Series on Operations Research and Management Science, Springer, New York, NY, USA, 2006.
- [2] W.-K. Ching, E. S. Fung, and M. K. Ng, "A multivariate Markov chain model for categorical data sequences and its applications in demand predictions," *IMA Journal of Management Mathematics*, vol. 13, no. 3, pp. 187–199, 2002.
- [3] W. K. Ching, E. S. Fung, and M. K. Ng, "Higher-order Markov chain models for categorical data sequences," *Naval Research Logistics*, vol. 51, no. 4, pp. 557–574, 2004.
- [4] W. K. Ching, M. M. Ng, E. S. Fung, and T. Akutsu, "On construction of stochastic genetic networks based on gene expression sequences," *International Journal of Neural Systems*, vol. 15, no. 4, pp. 297–310, 2005.
- [5] W. Ching, T. Siu, and L. Li, "An improved parsimonious multivariate Markov chain model for credit risk," *Journal of Credit Risk*, vol. 5, pp. 1–25, 2009.
- [6] D. W. C. Miao and B. M. Hambly, "Recursive formulas for the default probability distribution of a heterogeneous group of defaultable entities," 2012.
- [7] W.-K. Ching, M. K. Ng, and E. S. Fung, "Higher-order multivariate Markov chains and their applications," *Linear Algebra and Its Applications*, vol. 428, no. 2–3, pp. 492–507, 2008.
- [8] C. Wang, T. Z. Huang, and C. Wen, "A simplified higher-order multivariate Markov chains model," submitted.
- [9] C. Wang and T. Z. Huang, "Improved multivariate Markov chain model with the new convergence condition," submitted.
- [10] T.-K. Siu, W.-K. Ching, E. S. Fung, and M. K. Ng, "On a multivariate Markov chain model for credit risk measurement," *Quantitative Finance*, vol. 5, no. 6, pp. 543–556, 2005.
- [11] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [12] Z.-Y. You and C.-L. Wang, "A concept of nonlinear block diagonal dominance," *Journal of Computational and Applied Mathematics*, vol. 83, no. 1, pp. 1–10, 1997.
- [13] C. Wang, T. Z. Huang, and W. K. Ching, "On simplified parsimonious models for higher-order multivariate Markov chains," submitted.