Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 502808, 8 pages
http://dx.doi.org/10.1155/2014/502808

Research Article

A New Multivariate Markov Chain Model for Adding a New Categorical Data Sequence

Chao Wang,[1] Ting-Zhu Huang,[1] and Wai-Ki Ching[2]

[1] *School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China*
[2] *Advanced Modeling and Applied Computing Laboratory, Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong*

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

Received 10 July 2013; Revised 26 January 2014; Accepted 10 February 2014; Published 23 March 2014

Academic Editor: Andrzej Swierniak

We propose a new multivariate Markov chain model for adding a new categorical data sequence. The number of the parameters in the new multivariate Markov chain model is only $\mathcal{O}(3s)$ less than $\mathcal{O}((s+1)^2)$ the number of the parameters in the former multivariate Markov chain model. Numerical experiments demonstrate the benefits of the new multivariate Markov chain model on saving computational resources.

1. Introduction

Markov chains are of interest in a wide range of applications, for example, telecommunication systems [1, 2], remanufacturing and inventory systems [3], speech recognition [4], PageRank [5–7], microbial gene [8], and AIDS [9]. In recent years, the predictions of data sequences have become more and more useful in other applications such as sales demand prediction [10], DNA sequencing [11], credit risk [12], and stock prices [13].

Different models have been proposed for multiple categorical data sequences prediction. A multivariate Markov chain model was proposed in [10] by Ching et al.; they constructed a new matrix by means of the transition probability matrices among different sequences. An improved multivariate Markov chain model had also been studied to speed up the convergent speed for computing the stationary or steady-state solutions [14]. In the improved multivariate Markov chain model, Ching et al. incorporated positive and negative association parts. The extensions of intensity-based models for pricing credit risk and derivative securities to the simulation and valuation of portfolios were discussed in [15]. Moreover, there are many other papers contributing to the multivariate Markov chain model, for example, [16–22] and so on.

With the developments of Markov chain models and their applications, the number of the sequences may be larger. It is inevitable that a large categorical data sequence group will cause high computational cost in multivariate Markov chain model. Thus, reducing the number of parameters in the models is useful in numerical computation. For the above reasons, we present a new multivariate Markov chain model for detecting the relations between the previous data sequences and the following data sequence.

The rest of the paper is organized as follows. In Section 2, we review two lemmas and several Markov chain models. In Section 3, the new multivariate Markov chain model is proposed for adding a new categorical sequence. Moreover, the convergence of the new multivariate Markov chain model is proved. Section 4 gives parameter estimation method for the new multivariate Markov chain model. Numerical experiments on sales demand prediction and stock prices prediction are presented to test the efficiency of the new multivariate Markov chain model in Section 5. Concluding remarks are given in Section 6. The data of the stocks' prices are provided in the Appendix.

2. A Review on Markov Chain Models

In this section, we briefly introduce two lemmas, the Markov chain model [23] and the multivariate Markov chain model [10].

Lemma 1 (see [24, Perron-Frobenius theorem]). *Let $A \in \mathbb{R}^{m\times m}$ be a nonnegative and irreducible matrix. Then,*

(1) *A has a positive real eigenvalue λ equal to its spectral radius; that is, $\lambda = \max_k|\lambda_k(A)|$ where $\lambda_k(A)$ denotes the kth eigenvalue of A;*

(2) *to λ there corresponds an eigenvector z of its entries being real and positive, such that $Az = \lambda z$;*

(3) *λ is a simple eigenvalue of A.*

Lemma 2 (see [22]). *Let B be the iterative matrix of multivariate Markov chain model and let X_t be the state distribution at time t. If B is irreducible and aperiodic, then there is a unique stationary distribution π satisfying $\pi = B\pi$ and $\lim_{t\to\infty}X_t = \pi$.*

2.1. The Markov Chain Model. Let the state set of the categorical data sequence be $\mathscr{M} = \{1, 2, \ldots, m\}$. The Markov chain satisfise the following relations:

$$\begin{aligned}&\operatorname{Prob}\left(x_{t+1} = \theta_{t+1} \mid x_0 = \theta_0, x_1 = \theta_1, \ldots, x_t = \theta_t\right)\\&\quad = \operatorname{Prob}\left(x_{t+1} = \theta_{t+1} \mid x_t = \theta_t\right),\end{aligned} \tag{1}$$

where $\theta_t \in \mathscr{M}$, $t \in \{0, 1, 2, \ldots\}$. The conditional probability $\operatorname{Prob}(x_{t+1} = \theta_{t+1} \mid x_t = \theta_t)$ is called one-step transition probability. If we rewrite the transition probability as

$$p_{i,j} = \operatorname{Prob}\left(x_{t+1} = i \mid x_t = j\right), \quad \forall i, j \in \mathscr{M}, \tag{2}$$

the Markov chain model can be presented as follows:

$$X_{t+1} = PX_t, \tag{3}$$

where

$$\begin{gathered}P = \left[p_{i,j}\right], \quad 0 \le p_{ij} \le 1, \quad \forall i, j \in \mathscr{M},\\ \sum_{i=1}^{m} p_{i,j} = 1, \quad \forall j \in \mathscr{M}.\end{gathered} \tag{4}$$

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

2.2. The Multivariate Markov Chain Model. The multivariate Markov chain model has the following form:

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

where

$$\begin{gathered}\lambda_{jk} \ge 0, \quad \forall j, k = 1, 2, \ldots, s,\\ \sum_{k=1}^{s} \lambda_{j,k} = 1, \quad \forall j = 1, 2, \ldots, s.\end{gathered} \tag{6}$$

$x_0^{(j)}$ is the initial probability distribution of the jth sequence, $x_t^{(k)}$ is the state probability distribution of the kth sequence at time t, $x_{t+1}^{(j)}$ is the state probability distribution of the jth sequence at time $t + 1$, and $P^{(j,k)}$ is the one-step transition probability matrix from the states in the kth sequence at time t to the states in the jth sequence at time $t + 1$. In the matrix form, (5) has

$$\begin{aligned}X_{t+1} &= \begin{pmatrix} x_{t+1}^{(1)} \\ x_{t+1}^{(2)} \\ \vdots \\ x_{t+1}^{(s)} \end{pmatrix}\\ &= \begin{pmatrix} \lambda_{1,1}P^{(1,1)} & \lambda_{1,2}P^{(1,2)} & \cdots & \lambda_{1,s}P^{(1,s)} \\ \lambda_{2,1}P^{(2,1)} & \lambda_{2,2}P^{(2,2)} & \cdots & \lambda_{1,1}P^{(2,s)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{s,1}P^{(s,1)} & \lambda_{s,2}P^{(s,2)} & \cdots & \lambda_{s,s}P^{(s,s)} \end{pmatrix} \begin{pmatrix} x_t^{(1)} \\ x_t^{(2)} \\ \vdots \\ x_t^{(s)} \end{pmatrix}.\end{aligned} \tag{7}$$

Entries $P^{(j,k)}$ can be obtained directly from the categorical data sequences, and $\lambda_{j,k}$ can be got by the linear programming [10].

3. A New Multivariate Markov Chain Model

In order to reduce the number of the parameters in multivariate Markov chain model, a new multivariate Markov chain model is proposed. Moreover, the convergent property of the new model is also analyzed.

Suppose that there are s categorical data sequences and each of the sequences has m possible states in $\mathscr{M}$. The multivariate Markov chain model for s categorical data sequences has the form

$$x_{r+1}^{(j)} = \sum_{k=1}^{s} \lambda'_{j,k} P^{(j,k)} x_r^{(k)} \quad \forall 1 \le j \le s, \ r \ge 0, \tag{8}$$

where

$$\begin{gathered}\lambda'_{j,k} \ge 0, \quad \forall 1 \le j, k \le s,\\ \sum_{k=1}^{s} \lambda'_{j,k} = 1, \quad \forall j = 1, 2, \ldots, s,\end{gathered} \tag{9}$$

where $x_0^{(j)}, x_r^{(k)}, x_{r+1}^{(j)}$, and $P^{(jk)}$ are defined the same as those in Section 2.2. In the matrix form, (8) is

$$X_{r+1} = \begin{pmatrix} x_{r+1}^{(1)} \\ x_{r+1}^{(2)} \\ \vdots \\ x_{r+1}^{(s)} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda'_{1,1}P^{(1,1)} & \lambda'_{1,2}P^{(1,2)} & \cdots & \lambda'_{1,s}P^{(1,s)} \\ \lambda'_{2,1}P^{(2,1)} & \lambda'_{2,2}P^{(2,2)} & \cdots & \lambda'_{1,1}P^{(2,1)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda'_{s,1}P^{(1,1)} & \lambda'_{s,2}P^{(s,2)} & \cdots & \lambda'_{s,s}P^{(s,s)} \end{pmatrix} \begin{pmatrix} x_r^{(1)} \\ x_r^{(2)} \\ \vdots \\ x_r^{(s)} \end{pmatrix}. \tag{10}$$

Transition probability matrix $P^{(j,k)}$ can be obtained directly by the s categorical data sequences. The parameters $\lambda'_{j,k}$ can be solved from the corresponding linear programming.

Assuming that the multivariate Markov chain model for previous sequences is obtained, we add a new sequence at the back of the previous sequences. For detecting the relations between the previous s categorical data sequences and the new categorical data sequence, a new multivariate Markov chain model is proposed and has the following form:

$$x_{r+1}^{(j)} = \begin{cases} l_j \sum_{k=1}^{s} \lambda'_{j,k} P^{(j,k)} x_r^{(k)} + \lambda_{j,s+1} P^{(j,s+1)} x_r^{(s+1)} & \forall j = 1, 2, \ldots, s, \\ \sum_{k=1}^{s+1} \lambda_{j,k} P^{(j,k)} x_r^{(k)} & \forall j = s+1, \end{cases} \tag{11}$$

where

$$\begin{gathered} l_j \geq 0, \quad l_j \lambda'_{j,k} = \lambda_{j,k}, \quad \lambda_{j,s+1} \geq 0, \quad \lambda_{s+1,k} \geq 0 \\ \forall 1 \leq j, k \leq s, \quad r = 0, 1, \ldots, \\ \sum_{k=1}^{s+1} \lambda_{j,k} = 1 \quad \forall j = 1, 2, \ldots, s+1. \end{gathered} \tag{12}$$

In the matrix form, (11) has

$$\begin{pmatrix} x_{r+1}^{(1)} \\ \vdots \\ x_{r+1}^{(s)} \\ x_{r+1}^{(s+1)} \end{pmatrix} = \begin{pmatrix} l_1\lambda'_{11}P^{(11)} & \cdots & l_1\lambda'_{1,s}P^{(1,s)} & \lambda_{1,s}P^{(1,s+1)} \\ \vdots & \vdots & \vdots & \vdots \\ l_s\lambda'_{s,1}P^{(s,1)} & \cdots & l_s\lambda'_{s,s}P^{(s,s)} & \lambda_{s,s+1}P^{(s,s+1)} \\ \lambda_{s+1,1}P^{(s+1,1)} & \cdots & \lambda_{s+1,s}P^{(s+1,s)} & \lambda_{s+1,s+1}P^{(s+1,s+1)} \end{pmatrix} \times \begin{pmatrix} x_r^{(1)} \\ \vdots \\ x_r^{(s)} \\ x_r^{(s+1)} \end{pmatrix}. \tag{13}$$

Let

$$X_{r+1} = \left(x_{r+1}^{(1)}, \ldots, x_{r+1}^{(s)}, x_{r+1}^{(s+1)}\right)^T,$$

$$B = \begin{pmatrix} l_1\lambda'_{11}P^{(11)} & \cdots & l_1\lambda'_{1,s}P^{(1,s)} & \lambda_{1,s}P^{(1,s+1)} \\ \vdots & \vdots & \vdots & \vdots \\ l_s\lambda'_{s,1}P^{(s,1)} & \cdots & l_s\lambda'_{s,s}P^{(s,s)} & \lambda_{s,s+1}P^{(s,s+1)} \\ \lambda_{s+1,1}P^{(s+1,1)} & \cdots & \lambda_{s+1,s}P^{(s+1,s)} & \lambda_{s+1,s+1}P^{(s+1,s+1)} \end{pmatrix}. \tag{14}$$

Equation (13) is abbreviated as

$$X_{r+1} = BX_r. \tag{15}$$

Theorem 3. *Let $l_j \geq 0$, $\lambda'_{j,k} \geq 0$, $l_j\lambda'_{j,k} = \lambda_{j,k}$. If $\lambda_{j,s+1} > 0$, $\lambda_{s+1,k} > 0$ $\forall 1 \leq j, k \leq s$, then the iterative matrix B has an eigenvalue equal to one and the modulus of all its eigenvalues are less than or equal to one.*

Proof. Suppose that

$$\Lambda = \begin{pmatrix} \lambda^{(1,1)} & \cdots & \lambda^{(1,s)} & \lambda^{(1,s+1)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda^{(s,1)} & \cdots & \lambda^{(s,s)} & \lambda^{(s,s+1)} \\ \lambda^{(s+1,1)} & \cdots & \lambda^{(s+1,s)} & \lambda^{(s+1,s+1)} \end{pmatrix} = \begin{pmatrix} l_1\lambda'^{(1,1)} & \cdots & l_1\lambda'^{(1,2)} & \lambda^{(1,s+1)} \\ \vdots & \vdots & \vdots & \vdots \\ l_2\lambda'^{(s,1)} & \cdots & l_s\lambda'^{(s,s)} & \lambda^{(s,s+1)} \\ \lambda^{(s+1,1)} & \cdots & \lambda^{(s+1,s)} & \lambda^{(s+1,s+1)} \end{pmatrix}. \tag{16}$$

From (9) and (12), each column sum of this matrix is equal to one and the matrix Λ is nonnegative and irreducible. According to Lemma 1, there exists a positive vector

$$y = [y_1, y_2, \ldots, y_s]^T \tag{17}$$

satisfying

$$\Lambda y = y, \qquad y^T \Lambda^T = y^T. \tag{18}$$

Let $1_m = [1, \ldots, 1] \in \mathbb{R}^{1\times m}$. Since $P^{(jk)}$ is a transition probability matrix, it has

$$1_m P^{(j,k)} = 1_m. \tag{19}$$

It is clear that

$$[y_1 1_m, y_2 1_m, \ldots, y_s 1_m]\, B = [y_1 1_m, y_2 1_m, \ldots, y_s 1_m], \tag{20}$$

with 1 an eigenvalue of B.

Now, our aim is to prove that the modulus of all the eigenvalues of B are less than or equal to one. Suppose that $D_v = \mathrm{Diag}(v)$, $v^T = y \otimes 1_m = [y_1 1_m, y_2 1_m, \ldots, y_s 1_m]$ satisfying $v^T B = v^T$. $\widehat{B} = D_v B D_v^{-1}$ is similar to B. From (20), it has $\|\widehat{B}\|_1 = 1$. Then

$$\rho(B) = \rho\left(\widehat{B}\right) \leq \left\|\widehat{B}\right\|_1 = 1. \tag{21}$$

□

Theorem 4. *Assume that* $\forall 1 \le j, k \le s+1$, $p^{(j,k)}$ *is irreducible,* $\forall 1 \le j, k \le s$, $l_j \lambda'_{j,k} = \lambda_{j,k} \ge 0$, *and* $\lambda^{(j,s+1)} > 0$, $\lambda^{(s+1,k)} > 0$. *Then there is a vector* $X = [x^{(1)}, x^{(2)}, \ldots, x^{(s+1)}]^T$ *satisfying* $X = BX$ *and*

$$\sum_{i=1}^{m} \left[x^{(i)}\right]_j = 1 \quad \forall i = 1, 2, \ldots, m. \tag{22}$$

Proof. The proof is similar to Proposition 2 in [1] and therefore it is omitted. □

To keep the irreducibility of B, we fill the column of $P^{(j,k)}$ with $1/m$ when the column sum of B is zero.

Theorem 5. *Let* X *be the stationary probability of the new multivariate Markov chain model. Then* $X = BX$ *and* $\lim_{t\to\infty} X_t = X$.

Proof. From Lemma 2, our goal is to prove that B is irreducible and aperiodic. Since B is connected, B is irreducible. Then we only need to prove that B is aperiodic. Let $S_1 = \{1, 2, \ldots, s-1\}$, $S = \{1, 2, \ldots, s\}$. There exists t_1 such that $(B^{(ij)})^{t_1} > 0$ $\forall i, j \in S_1$. From the form of B, we obtain $B^{(sj)} > 0$, $B^{(js)} > 0$. There exists t_2, t_3, t_4 satisfying $B_{ij}^{t_2} > 0$, $B_{ii}^{t_3} > 0$, $B_{ki}^{t_4} > 0$. Therefore, $B_{kj}^{t_2+t_3+t_4} \ge B_{ij}^{t_2} B_{ii}^{t_3} B_{ki}^{t_4} > 0$. Because $B_{kj}^{n} > 0$ $\forall n > t_2 + t_3 + t_4$, $(B)^n$ is a positive matrix. Then B is aperiodic. According to the above results, the conclusions of this theorem are obtained. □

4. Parameter Estimation Method of the New Multivariate Markov Chain Model

Let $\mathscr{M} = \{1, 2, \ldots, m\}$ be the states set and let $f_{i_j,i_k}^{(j,k)}$ be frequency from the i_k state in the kth sequence at time $t = r$ to the i_j state in the jth sequence at time $t = r+1$ $\forall 1 \le i_j, i_k \le m$. The transition frequency matrix $F^{(j,k)}$ is

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

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

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

where

$$p_{i_j,i_k}^{(j,k)} = \begin{cases} \dfrac{f_{i_j,i_k}^{(j,k)}}{\sum_{i_j=1}^{m} f_{i_j,i_k}^{(j,k)}}, & \text{if } \sum\limits_{i_j=1}^{m} f_{i_j,i_k}^{(j,k)} \ne 0, \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

Subsequently, the way of estimating the parameter $\lambda_{j,k}$ will be introduced. Consider X to be a joint stationary probability distribution. X can be presented as

$$X = \left(\left(X^{(1)}\right)^T, \left(X^{(2)}\right)^T, \ldots, \left(X^{(s)}\right)^T\right)^T_{ms\times 1} \tag{26}$$

satisfying

$$BX \equiv X. \tag{27}$$

One would expect that

$$\begin{pmatrix} B^{(1,1)} & B^{(1,2)} & \cdots & B^{(1,s)} \\ B^{(2,1)} & B^{(2,2)} & \cdots & B^{(2,s)} \\ \vdots & \vdots & \vdots & \vdots \\ B^{(s,1)} & B^{(s,2)} & \cdots & B^{(s,s)} \end{pmatrix} X \approx X. \tag{28}$$

Certainly, (28) can be interpreted as

$$|BX - X| \le \omega, \tag{29}$$

where ω is small enough.

One way of estimating $\lambda_{j,k}$ is to transform (29) into a minimization problem as the following form:

$$\begin{aligned} \min_{\lambda_{i,j}} \quad & \|BX - X\| \\ \text{subject to} \quad & \sum_{k=1}^{s+1} \lambda_{j,k} = 1, \quad \forall j \in \{1, 2, \ldots, s+1\}, \\ & \lambda_{j,k} \ge 0, \quad \forall j, k \in \{1, 2, \ldots, s+1\}. \end{aligned} \tag{30}$$

The minimization problem (30) is identical to the following form:

$$\begin{aligned} \min \lambda_{i,j} \max_i \quad & \left(\left\| \left[\sum_{k=1}^{s} l_j \lambda'_{j,k} P^{(j,k)} x^{(k)} + \lambda_{j,s+1}^{(h)} P^{(j,s+1)} x^{(s+1)} - x^{(j)} \right]_i \right\|\right) \\ \text{subject to} \quad & \sum_{k=1}^{s} l_j \lambda'_{j,k} + \lambda_{j,s+1} = 1, \quad \forall j \in \{1, 2, \ldots, s\}, \\ & \sum_{k=1}^{s+1} \lambda_{s+1,k} = 1, \\ & \lambda_{j,s+1}, \lambda_{s+1,k} \ge 0, \quad \forall j, k \in \{1, 2, \ldots, s+1\}, \end{aligned} \tag{31}$$

where $[\cdot]_i$ is the ith entry of the vector. Let the norm be $\|\cdot\|_2$. The above problem can be rewritten as a linear programming problem

$$\begin{aligned}
\min_{\lambda} \quad & \omega_j \\
\text{subject to} \quad & \begin{pmatrix} \omega_j \\ \omega_j \end{pmatrix} \geq x^{(j)} - B\begin{pmatrix} l_j \\ \lambda_{j,s+1} \end{pmatrix} \\
& \begin{pmatrix} \omega_j \\ \omega_j \end{pmatrix} \geq -x^{(j)} + B\begin{pmatrix} l_j \\ \lambda_{j,s+1} \end{pmatrix}, \quad \forall j \in \{1,2,\ldots,s\} \\
& \begin{pmatrix} \omega_j \\ \omega_j \\ \vdots \\ \omega_j \end{pmatrix} \geq x^{(j)} - B\begin{pmatrix} \lambda_{j,1} \\ \lambda_{j,2} \\ \vdots \\ \lambda_{j,s+1} \end{pmatrix} \\
& \begin{pmatrix} \omega_j \\ \omega_j \\ \vdots \\ \omega_j \end{pmatrix} \geq -x^{(j)} + B\begin{pmatrix} \lambda_{j,1} \\ \lambda_{j,2} \\ \vdots \\ \lambda_{j,s+1} \end{pmatrix}, \\
& \qquad \forall j = s+1,\ \omega_j \geq 0, \\
& l_j \sum_{k=1}^{s} \lambda_{j,k} + \lambda_{j,s+1} = 1, \quad \forall j \in \{1,2,\ldots,s\}, \\
& \sum_{k=1}^{s+1} \lambda_{s+1,k} = 1, \\
& \lambda_{j,s+1}, \lambda_{s+1,k} \geq 0, \quad \forall j,k \in \{1,2,\ldots,s+1\},
\end{aligned} \tag{32}$$

where

$$B = \begin{cases} \left[\lambda_{j,1} p^{(j,1)} x^{(1)} + \lambda_{j,2} p^{(j,2)} x^{(2)} + \cdots + \lambda_{j,s} p^{(j,s)} x^{(s)} \mid p^{(j,s+1)} x^{(s+1)}\right], & \text{if } 1 \leq j \leq s, \\ \left[p^{(j,1)} x^{(1)} \mid p^{(j,2)} x^{(2)} \mid \cdots \mid p^{(j,s)} x^{(s)} \mid p^{(j,s+1)} x^{(s+1)}\right], & \text{if } j = s+1. \end{cases} \tag{33}$$

5. Numerical Experiments

In this section, numerical experiments with different multivariate Markov chain models on sales demand prediction and stock prices prediction are given. We report on numerical results obtained with a Matlab 7.0.1 implementation on Windows XP with 2.93 GHz 64-bit processor and 1 GB memory.

5.1. Sales Demand Prediction. In this section, the sales demand sequences are presented to show the benefits of the new multivariate Markov chain model. 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. Here, our goal is to predict the sales demand of the market for minimizing the estate cost. Assume that products are classified into six possible states (1, 2, 3, 4, 5, 6); for example, 1 = no sale volume, 2 = very low sale volume, 3 = low sale volume, 4 = standard sale volume, 5 = high sale volume, and 6 = very high sale volume. The customers' sales demand data of five important products can be found in [10].

The multivariate Markov chain model of four categorical data sequences, *ProductA*, *ProductB*, *ProductC*, and *ProductD*, will be given. By computing the proportions of the occurrences of each state in each sequence, we formulate the initial probability distributions of four categorical data sequences

$$\begin{aligned}
x_0^{(1)} &= (0.0818, 0.4052, 0.0483, 0.0335, 0.0037, 0.4275)^T, \\
x_0^{(2)} &= (0.3680, 0.1970, 0.0335, 0.0000, 0.0037, 0.3978)^T, \\
x_0^{(3)} &= (0.1450, 0.2045, 0.0186, 0.0000, 0.0037, 0.6283)^T, \\
x_0^{(4)} &= (0.0000, 0.3569, 0.1338, 0.1896, 0.0632, 0.2565)^T.
\end{aligned} \tag{34}$$

The transition probability matrix $P^{(j,k)}$ can be obtained after normalizing the transition frequency matrix. By solving the corresponding linear programming problem, one can obtain $\lambda_{i,j}$. The multivariate Markov chain model is presented as follows:

$$\begin{aligned}
x_{r+1}^{(1)} &= P^{(1,2)} x_r^{(2)}, \\
x_{r+1}^{(2)} &= P^{(2,2)} x_r^{(2)}, \\
x_{r+1}^{(3)} &= P^{(3,4)} x_r^{(4)}, \\
x_{r+1}^{(4)} &= P^{(4,4)} x_r^{(4)}.
\end{aligned} \tag{35}$$

In order to uncover the relations of *ProductA*, *ProductB*, *ProductC*, *ProductD*, and *ProductE*, we add a new categorical data sequence *ProductE* at the back of the original categorical data sequences. With the data sequence of *ProductE*, the initial probability distribution of *ProductE* is obtained as follows:

$$x_0^{(5)} = (0.0000, 0.3569, 0.1227, 0.2268, 0.0520, 0.2416)^T. \tag{36}$$

In the multivariate Markov chain model, $\lambda_{j,k}$ can be calculated by a corresponding linear programming problem. The multivariate Markov chain model is

$$\begin{aligned}
x_{r+1}^{(1)} &= P^{(1,2)} x_r^{(2)}, \\
x_{r+1}^{(2)} &= P^{(2,2)} x_r^{(2)}, \\
x_{r+1}^{(3)} &= P^{(3,5)} x_r^{(5)},
\end{aligned}$$

Table 1: Numerical results of the multivariate Markov chain model and the new multivariate Markov chain model on sales demand prediction.

	A	B	C	D	E	*Time*	*np*	ω
Mmodel	0.4176	0.3985	0.6207	0.3525	0.3525	0.0938	25	0.0095
B, C, D, E add A	0.4176	0.3985	0.6207	0.3525	0.3525	0.0469	13	0.0095
A, C, D, E add B	0.4176	0.3985	0.6207	0.3525	0.3525	0.0469	13	0.0095
A, B, D, E add C	0.4176	0.3985	0.6207	0.3525	0.3525	0.0469	13	0.0095
A, B, C, E add D	0.4176	0.3985	0.6207	0.3525	0.3525	0.0469	13	0.0095
A, B, C, D add E	0.4176	0.3985	0.6207	0.3525	0.3525	0.0469	13	0.0095

$$x_{r+1}^{(4)} = 0.2783P^{(4,4)}x_r^{(4)} + 0.7217P^{(4,5)}x_r^{(5)},$$
$$x_{r+1}^{(5)} = P^{(5,4)}x_r^{(4)}. \tag{37}$$

After calculating $\lambda_{j,k}$ by a corresponding linear programming problem, the new multivariate Markov chain model can be presented as follows:

$$\begin{aligned} x_{r+1}^{(1)} &= P^{(1,2)}x_r^{(2)}, \\ x_{r+1}^{(2)} &= P^{(2,2)}x_r^{(2)}, \\ x_{r+1}^{(3)} &= 0.0007P^{(3,4)}x_r^{(4)} + 0.9993P^{(3,5)}x_r^{(5)}, \\ x_{r+1}^{(4)} &= 0.4706P^{(4,4)}x_r^{(4)} + 0.5294P^{(4,5)}x_r^{(5)}, \\ x_{r+1}^{(5)} &= P^{(5,4)}x_r^{(4)}. \end{aligned} \tag{38}$$

From the results of the new multivariate Markov model, *ProductA* and *ProductB* are closely related. Moreover, the sales demand of *ProductA* depends strongly on *ProductB*. The reason is that the chemical nature of *ProductA* and *ProductB* is the same, only used for different packaging of marketing purposes. *ProductC*, *ProductD*, and *ProductE* are closely related. The fact is that *ProductsC* and *ProductE* have the same product flavor, only different in packaging.

In the following, we use the new multivariate Markov chain model and the multivariate Markov chain model to predict the state of the kth sequence $x_t^{(k)}$ at time t. The maximum probability,

$$x_t^{(k)} = j, \quad \text{if } \left[x_t^{(k)}\right]_i \le \left[x_t^{(k)}\right]_j, \quad \forall 1 \le i \le m, t > 1, \tag{39}$$

is taken as the state at time t. For evaluating the effectiveness of the new multivariate Markov chain model, prediction results are measured by the prediction accuracy r defined as

$$r = \frac{1}{T-n} \times \sum_{t=n+1}^{T} \delta_t \times 100\%, \tag{40}$$

where T is the length of the data sequence and

$$\delta_t = \begin{cases} 1, & \text{if } x_t^{(k)} = \theta_t^{(k)}, \\ 0, & \text{otherwise.} \end{cases} \tag{41}$$

Note that "*time*" is CPU time, "ω" is the object function value of the corresponding linear programming problem, "*np*" is the number of the parameters in the models, and the prediction accuracies of *ProductA*, *ProductB*, *ProductC*, *ProductD*, and *ProductE* are "*A*," "*B*," "*C*," "*D*," and "*E*," respectively. Suppose that the results of the multivariate Markov chain model of *ProductA*, *ProductB*, *ProductC*, and *ProductD* are obtained. The new multivariate Markov chain model for *ProductA*, *ProductB*, *ProductC*, and *ProductD* adding *ProductE* is denoted as "A, B, C, D add E." The multivariate Markov chain model is denoted as "*Mmodel*." Stop criterion can be found in Matlab function *linprog*. The results are presented in Table 1.

Observing the results from Table 1, we find that the object function values of the new model and the multivariate Markov chain model are the same. The prediction accuracies of our new models are comparable to the prediction accuracy of the multivariate Markov chain model. The new multivariate Markov chain model performs better than the multivariate Markov chain model in time consumption and controlled parameters.

Table 2: Numerical results of the multivariate Markov chain model and the new multivariate Markov chain model on stock prices prediction.

	Time	*np*	ω
Mmodel	0.7354	144	0.3095
NMmodel	0.1406	34	0.3099

5.2. Stock Prices Prediction. The data of 12 American stocks' price from December 17, 2013, to January 16, 2014, are given in the Appendix. They are divided equally into 6 regions as 6 states between maximum price and minimum price of the stocks. The state set of 12 stocks is $\mathscr{M} = \{0, 1, 2, 3, 4, 5\}$. The data of 12 stocks in the Appendix are transformed into categorical data sequences.

"*time*," "ω," "*np*," and "*Mmodel*" are denoted the same as those in Section 5.2. Note that "*Mmodel* − 1" is the multivariate Markov chain model of all stocks except AMAP. Suppose that the results of "*Mmodel* − 1" are obtained. The new multivariate Markov chain model which is denoted as "*NMmodel*" can detect the relations of BIDU, CTRP, GA, EDU, SINA, SOHU, YOKU, XRS, QIHU, HTHT, HMIN, and AMAP. Stop criterion can be found in Matlab order *linprog*. The results are presented in Table 2.

From Table 2, the object function values of the new multivariate Markov chain model and the multivariate Markov chain model are nearly the same. The CPU time of the new

multivariate Markov chain model is the CPU time of the multivariate Markov chain model's 1/5. The number of the parameters in the new multivariate Markov chain model is one-third of those in the multivariate Markov chain model. The new multivariate Markov chain model is better than the multivariate Markov chain model in time consumption and controlled parameters.

6. Conclusions

In this paper, a new multivariate Markov chain model is proposed. The convergence of the new model is proved. With the results of the multivariate Markov chain model for s categorical data sequences, the relations of the s categorical data sequences and the new sequence can be detected by our new model. The new multivariate Markov chain model only needs $\mathcal{O}(3s)$ parameters less than $\mathcal{O}((s+1)^2)$ which is the number of the parameters in multivariate Markov chain model. Numerical experiments illustrate the benefits of our new model in saving computational resources. The performances of the new multivariate Markov chain model are nearly the same as the multivariate Markov chain model in prediction. Certainly, our model can also be applied in credit risk and other research areas.

Appendix

Consider

$$\begin{aligned}
\text{BIDU} = {} & \{168.33, 171.49, 170.39, 173.36, 172.30, \\
& 168.58, 167.28, 173.77, 173.99, \\
& 177.88, 179.99, 175.28, 176.63, 178.82, \\
& 181.79, 175.52, 179.66, 171.00, \\
& 172.87, 170.50, 173.00\}, \\
\text{CTRP} = {} & \{47.93, 48.39, 47.66, 48.68, 50.62, 50.95, 50.34, \\
& 52.55, 51.22, 49.62, 49.41, \\
& 45.53, 44.43, 46.02, 44.51, 40.49, 38.95, \\
& 39.14, 40.10, 40.34, 41.15\}, \\
\text{GA} = {} & \{11.08, 11.15, 11.12, 11.12, 11.26, 11.25, 11.27, \\
& 11.27, 11.28, 11.24, 11.19, \\
& 10.93, 10.93, 11.00, 10.95, 10.85, 10.84, 10.79, \\
& 10.84, 10.86, 11.10\}, \\
\text{EDU} = {} & \{28.93, 29.66, 29.70, 29.79, 29.96, 30.57, 30.02, \\
& 29.93, 31.00, 31.50, 30.62, \\
& 30.27, 30.38, 31.26, 32.41, 32.54, 32.63, \\
& 32.66, 32.77, 33.07, 33.00\}, \\
\text{SINA} = {} & \{77.32, 78.84, 79.71, 79.91, 80.09, 79.76, 79.22, \\
& 82.64, 82.21, 84.25, 84.77, \\
& 82.68, 84.35, 87.30, 88.96, 84.96, 85.72, \\
& 83.07, 84.99, 84.60, 80.58\}, \\
\text{SOHU} = {} & \{67.26, 68.53, 69.85, 70.04, 70.16, 72.58, 70.60, \\
& 71.07, 72.11, 72.93, 73.28, \\
& 71.96, 73.94, 76.61, 78.55, 75.89, 76.54, \\
& 74.23, 75.95, 75.87, 75.57\}, \\
\text{YOKU} = {} & \{28.91, 29.15, 29.14, 28.67, 30.13, 30.11, 30.72, \\
& 30.94, 30.28, 30.30, 31.80, 31.50, \\
& 33.92, 34.82, 33.55, 33.00, 33.69, \\
& 33.30, 34.51, 34.79, 35.55\}, \\
\text{XRS} = {} & \{19.78, 19.89, 19.89, 19.87, 20.00, 21.50, 21.98, \\
& 21.83, 21.99, 21.99, 21.15, 21.49, \\
& 21.58, 22.18, 23.49, 22.53, 22.91, \\
& 32.01, 23.36, 23.40, 24.20\}, \\
\text{QIHU} = {} & \{77.48, 76.40, 77.61, 78.72, 81.08, 80.81, 80.48, \\
& 81.24, 81.02, 82.05, 81.85, 79.78, \\
& 80.10, 81.40, 89.00, 81.17, 81.04, \\
& 83.48, 86.58, 89.08, 91.26\}, \\
\text{HTHT} = {} & \{26.25, 26.27, 26.07, 26.20, 28.85, 30.02, 30.92, \\
& 30.58, 30.37, 30.48, 31.02, 30.61, \\
& 29.37, 29.21, 30.08, 28.98, 29.02, \\
& 27.98, 28.37, 28.56, 28.8\}, \\
\text{HMIN} = {} & \{41.46, 41.21, 41.25, 40.73, 42.24, 43.70, 43.07, \\
& 43.69, 42.94, 43.64, 43.51, 41.94, \\
& 40.77, 41.37, 43.43, 40.07, 40.41, \\
& 39.04, 39.39, 40.88, 40.48\}, \\
\text{AMAP} = {} & \{13.97, 13.70, 13.80, 14.17, 14.10, 14.15, 14.56, \\
& 14.77, 14.57, 14.25, 14.98, 15.46, \\
& 15.55, 15.73, 15.92, 15.60, 15.96, \\
& 15.75, 15.57, 15.22, 15.14\}.
\end{aligned} \tag{A.1}$$

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This research is supported by the Chinese Universities Specialized Research Fund for the Doctoral Program (20110185110020).

References

[1] H. S. Wang and N. Moayeri, "Finite-state Markov channel—a useful model for radio communication channels," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 1, pp. 163–171, 1995.

[2] F. P. Kelly, "Stochastic models of computer communication systems," *Journal of the Royal Statistical Society B*, vol. 47, no. 3, pp. 379–395, 1985.

[3] E. van der Laan and M. Salomon, "Production planning and inventory control with remanufacturing and disposal," *European Journal of Operational Research*, vol. 102, no. 2, pp. 264–278, 1997.

[4] M. Gales and S. Young, "The application of hidden Markov models in speech recognition," *Foundations and Trends in Signal Processing*, vol. 1, no. 3, pp. 195–304, 2008.

[5] C. P. C. Lee, G. H. Golub, and S. A. Zenios, "A fast two-stage algorithm for computing PageRank and its extensions," Tech. Rep. SCCM-200315, Scientific Computation and Computational Mathematics, Stanford University, 2003.

[6] S. Kamvar, T. Haveliwala, and G. Golub, "Adaptive methods for the computation of PageRank," *Linear Algebra and Its Applications*, vol. 386, pp. 51–65, 2004.

[7] A. N. Langville and C. D. Meyer, *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, 2006.

[8] S. L. Salzberg, A. L. Deicher, and S. Kasif, "Microbial gene identification using interpolated Markov models," *Nucleic Acids Research*, vol. 26, no. 2, pp. 544–548, 1998.

[9] H. Frydman, "A nonparametric estimation procedure for a periodically observed three-state Markov process, with application to AIDS," *Journal of the Royal Statistical Society B*, vol. 54, no. 3, pp. 853–866, 1992.

[10] 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.

[11] W. Ching, M. M. Ng, and E. S. Fung, "On construction of stochastic genetic networks based on gene expression sequences," *International Journal of Neural Systems*, vol. 15, no. 4, pp. 297–310, 2005.

[12] G. D'Amico, J. Janssen, and R. Manca, "Initial and final backward and forward discrete time non-homogeneous semi-markov credit risk models," *Methodology and Computing in Applied Probability*, vol. 12, no. 2, pp. 215–225, 2010.

[13] Z. Psaradakis, M. Sola, and F. Spagnolo, "On markov error-correction models, with an application to stock prices and dividends," *Journal of Applied Econometrics*, vol. 19, no. 1, pp. 69–88, 2004.

[14] 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.

[15] N. J. Jobst and S. A. Zenios, "Extending credit risk (pricing) models for the simulation of portfolios of interest rate and credit risk sensitive securities," Wharton School Centre for Financial Institutions Working Papers 01-25, 2001.

[16] B.-Y. Pu, T.-Z. Huang, and C. Wen, "A new GMRES(m) method for Markov chains," *Mathematical Problems in Engineering*, vol. 2013, Article ID 206375, 7 pages, 2013.

[17] L. C. Lin and S. Yau, "Analyzing Taiwan IC assembly industry by Grey-Markov forecasting model," *Wu Mathematical Problems in Engineering*, vol. 2013, Article ID 658630, 6 pages, 2013.

[18] M. Davis and V. Lo, "Modeling default correlation in bond portfolios," in *Mastering Risk*, vol. 2, pp. 141–151, Financial Times Management, 2001.

[19] M. Kijima, K. Komoribayashi, and E. Suzuki, "A multivariate Markov model for simulating correlated defaults," *Journal of Risk*, vol. 4, pp. 1–32, 2002.

[20] A. E. Raftery, "A model for high-order Markov chains," *Journal of the Royal Statistical Society B*, vol. 47, no. 3, pp. 528–539, 1985.

[21] T.-K. Siu, W.-K. Ching, and E. S. Fung, "On a multivariate Markov chain model for credit risk measurement," *Quantitative Finance*, vol. 5, no. 6, pp. 543–556, 2005.

[22] D.-M. Zhu and W.-K. Ching, "A note on the stationary property of high-dimensional Markov chain models," *International Journal of Pure and Applied Mathematics*, vol. 66, no. 3, pp. 321–330, 2011.

[23] C. T. Haan, D. M. Allen, and J. O. Street, "A Markov chain model of daily rainfall," *Water Resources Research*, vol. 12, no. 3, pp. 443–449, 1976.

[24] E. Seneta, *Non-negative Matrices and Markov Chain*, Springer, New York, NY, USA, 1981.