

# Sales Demand Forecasting Using One of Multivariate Markov Chain Model Parameter

Annisa Martina 1\*

<sup>1</sup>Mathematics Department, Faculty of Science and Technology, UIN Sunan Gunung Djati Bandung Jalan A.H. Nasution 105, Bandung, Indonesia

\*annisamartina@uinsgd.ac.id

#### Abstract

The imbalance between demand and supply is frequently occurred in a market. This is due to the availability of goods that cannot match with the demand or the growth rate of customer. This is not preferable since the profit is not on the track. In contrast, the goods are probably over supplied so that company has to expense additional cost for extra storage. Both situations can be anticipated if the demand is precisely estimated. Therefore, in this study we will estimate demand in market situation by implementing multivariate Markov chain model. Multivariate Markov chain model is popular model for forecasting by observing current state in various applications. This model is compatible with 5 data sequences (product types) defined as product A, product B, product C, product D and product E, with 6 conditions (no sales volume, very slow-moving, slow-moving, standard, fast moving, and very fast moving). As the result, the highest transition probability value for the sales demand in a company is found at the transition probability matrix from product C to product C, from very fast moving to very fast-moving condition, which had the highest probability value 0.625 with the highest frequency 105 times.

Keywords: Multivariate Markov chain model, demand, forecasting, transition probability matrix

## I. INTRODUCTION

A developed company needs a professional management system, one of which is in the field of production. In general, the goal of a company to produce a product is to meet consumer needs and to achieve maximum company profits [1]. This really depends on how precisely the consumer's needs are met, or in other words the number of goods produced by the company must match the number of goods demanded by consumers. In fact, there are often conditions where the number of goods provided by the company does not match the number of requests from consumers. As a result, consumers find it difficult to get goods that sell well, and the company loses income from the amount of demand that cannot be fulfilled, so that the company does not get maximum profit. In addition, there are often conditions of accumulation of goods for unsold goods, so the company must provide a place to store these items (which requires a lot of money). If this condition persists at a certain time, the company will suffer a loss. Therefore, a method is needed to estimate the amount of goods produced by the company to match the demand for goods from consumers. In this study, the author will estimate the transition probability matrix which is one of multivariate Markov chain model parameter that can estimate future conditions by observing the present state.

#### II. LITERATURE REVIEW

The Markov chain is a stochastic process that has a state space in the form of a finite set [4]. For example, at time *n*, the process is in *k*th state, then it can be written  $X_n = k$ . What is meant by stochastic processes  $\{X_n, n = 0, 1, ...\}$  is a collection of random variables where *n* represents the time index. Formally the Markov chain can be written as follows:

for all  $k_0, k_1, \dots, k_{n-1}, k, j$  and all  $n \ge 0$ 

$$P\left\{\underbrace{X_{n+1}=j}_{\text{future events}}\middle|\underbrace{X_0=k_0, X_1=k_1, \dots, X_{n-1}=k_{n-1}}_{\text{past events}}, \underbrace{X_n=k}_{\text{current events}}\right\} = P\left\{X_{n+1}=j\middle|X_n=k\right\} = P_{jk}$$
(1)

Based on equation (1), the conditional probability of all future events  $X_{n+1}$ , given the past events  $X_{0,X_1,\ldots,X_{n-1}}$  and the current events  $X_n$ , are independent past events, and only depends on current events [2]. Probability  $P_{jk}$  is the probability of transition to state *j* given the current events, namely state *k*. The following are properties possessed by  $P_{ik}$ :

$$\sum_{k=1}^{m} P_{jk} = 1, P_{jk} \ge 0, j = 1, 2, ..., m$$
<sup>(2)</sup>

## **III. RESEARCH METHOD**

In this multivariate Markov chain model, [3] it is assumed that there are s categories of categorical data (product types), each of which has m states (for example: many, few, etc.).

$$M = \{1, 2..., m\}$$

Let  $\mathbf{x}_{n}^{(j)}$  be the state probability vector of the *j*th sequence (product types) at time *n*. If at the *n*th time, the *j*th sequence (product types) is in state *l*, then it can be written as  $\mathbf{x}_{n}^{(j)} = \mathbf{e}_{l} = (0,...,0,\underbrace{1,...,0}_{\text{state }l},0,...,0)^{T}$ .

In constructing the multivariate Markov chain model, the following equation is assumed:

$$\mathbf{x}_{n+1}^{(j)} = \sum_{k=1}^{s} \lambda_{jk} \mathbf{P}^{(jk)} \mathbf{x}_{n}^{(k)}, \text{ for } j = 1, 2, ..., s$$
(3)

where  $\lambda_{jk} \ge 0$ ,  $1 \le j$ ,  $k \le s$ , and  $\sum_{k=1}^{s} \lambda_{jk} = 1$ , for j = 1, 2, ..., s.

Thus, based on equation (3), the distribution of the probability states of the sequence (product type) *j* at time (n + 1) depends on the states of the sequence (product type) *j* and *k* at time *n* [1]. Here  $\lambda_{jk}$  is the probability weight which includes the effect of the sequence state (product type) *k* to *j*. As  $\mathbf{P}^{(jk)}$  is the probability of the sequence state (product type) *k* to *j*, and  $\mathbf{x}_n^{(k)}$  is the probability of the sequence state (product type) *k* at time *n*. The following is writing in matrix:

$$\mathbf{x}_{n+1} = \begin{pmatrix} \mathbf{x}_{n+1}^{(1)} \\ \mathbf{x}_{n+1}^{(2)} \\ \vdots \\ \mathbf{x}_{n+1}^{(s)} \end{pmatrix} = \begin{pmatrix} \lambda_{11} \mathbf{P}^{(11)} & \lambda_{12} \mathbf{P}^{(12)} & \dots & \lambda_{1s} \mathbf{P}^{(1s)} \\ \lambda_{21} \mathbf{P}^{(21)} & \lambda_{22} \mathbf{P}^{(22)} & \dots & \lambda_{2s} \mathbf{P}^{(2s)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{s1} \mathbf{P}^{(s1)} & \lambda_{s2} \mathbf{P}^{(s2)} & \dots & \lambda_{ss} \mathbf{P}^{(ss)} \end{pmatrix} \begin{pmatrix} \mathbf{x}_{n}^{(1)} \\ \mathbf{x}_{n}^{(2)} \\ \vdots \\ \mathbf{x}_{n}^{(s)} \end{pmatrix} \\ \equiv \mathbf{Q} \qquad \mathbf{x}_{n} \end{pmatrix}$$
(4)

In this paper, we estimate the transition probability matrix  $\mathbf{P}^{(jk)}$  for each data sequence. The steps for finding the estimated transition probability matrix parameters  $\mathbf{P}^{(jk)}$  by the following method [4]. The first, calculate the transition frequency  $f_{i_j i_k}^{(jk)}$  from a state both in the same data sequence, or in different data sequence, save as the transition frequency matrix. Then, calculate the number per column of the transition frequency matrix. After that, divide each element in the same column by the total number in that column. The last, save as the transition probability matrix.

# IV. RESULTS AND DISCUSSION

# A. The Transition Frequency Matrix

**F** is the transition frequency matrix for 5 types of products (s = 5), that is product A (1), product B (2), product C (3), product D (4), and product E (5), in six states (m = 6), no sales volume (1<sup>st</sup> state), very slow-moving (2<sup>nd</sup> state), slow-moving (3<sup>rd</sup> state), standard (4<sup>th</sup> state), fast moving (5<sup>th</sup> state), and very fast moving (6<sup>th</sup> state), here is the matrix:

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}^{(11)} & \mathbf{F}^{(12)} & \mathbf{F}^{(13)} & \mathbf{F}^{(14)} & \mathbf{F}^{(15)} \\ \mathbf{F}^{(21)} & \mathbf{F}^{(22)} & \mathbf{F}^{(23)} & \mathbf{F}^{(24)} & \mathbf{F}^{(25)} \\ \mathbf{F}^{(31)} & \mathbf{F}^{(32)} & \mathbf{F}^{(33)} & \mathbf{F}^{(34)} & \mathbf{F}^{(35)} \\ \mathbf{F}^{(41)} & \mathbf{F}^{(42)} & \mathbf{F}^{(43)} & \mathbf{F}^{(44)} & \mathbf{F}^{(45)} \\ \mathbf{F}^{(51)} & \mathbf{F}^{(52)} & \mathbf{F}^{(53)} & \mathbf{F}^{(54)} & \mathbf{F}^{(55)} \end{pmatrix} \\ \text{where } \mathbf{F}^{(jk)} = \begin{pmatrix} f_{11}^{(jk)} & f_{21}^{(jk)} & f_{31}^{(jk)} & f_{41}^{(jk)} & f_{51}^{(jk)} & f_{61}^{(jk)} \\ f_{12}^{(jk)} & f_{22}^{(jk)} & f_{32}^{(jk)} & f_{42}^{(jk)} & f_{52}^{(jk)} & f_{62}^{(jk)} \\ f_{13}^{(jk)} & f_{23}^{(jk)} & f_{33}^{(jk)} & f_{43}^{(jk)} & f_{53}^{(jk)} & f_{63}^{(jk)} \\ f_{14}^{(jk)} & f_{24}^{(jk)} & f_{34}^{(jk)} & f_{44}^{(jk)} & f_{54}^{(jk)} & f_{64}^{(jk)} \\ f_{15}^{(jk)} & f_{25}^{(jk)} & f_{35}^{(jk)} & f_{45}^{(jk)} & f_{55}^{(jk)} & f_{65}^{(jk)} \\ f_{16}^{(jk)} & f_{26}^{(jk)} & f_{36}^{(jk)} & f_{46}^{(jk)} & f_{56}^{(jk)} & f_{66}^{(jk)} \end{pmatrix}$$

for *j* = 1, 2...6 and *k* = 1, 2...,6

The following are the results of calculating the transition frequency using the MATLAB software:

$$\mathbf{F}^{(11)} = \begin{pmatrix} 0 & 5 & 2 & 1 & 0 & 14 \\ 7 & 51 & 3 & 0 & 0 & 48 \\ 1 & 6 & 4 & 1 & 0 & 1 \\ 0 & 1 & 3 & 4 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 14 & 46 & 1 & 2 & 1 & 50 \end{pmatrix}, \mathbf{F}^{(12)} = \begin{pmatrix} 6 & 8 & 0 & 0 & 0 & 8 \\ 44 & 25 & 4 & 0 & 1 & 35 \\ 1 & 7 & 2 & 0 & 0 & 3 \\ 1 & 4 & 2 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 46 & 9 & 1 & 0 & 0 & 58 \end{pmatrix}, \dots, \mathbf{F}^{(15)}$$
$$\mathbf{F}^{(21)} = \begin{pmatrix} 5 & 43 & 0 & 2 & 1 & 47 \\ 4 & 19 & 6 & 5 & 0 & 19 \\ 0 & 7 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 13 & 39 & 5 & 2 & 0 & 48 \end{pmatrix}, \mathbf{F}^{(22)} = \begin{pmatrix} 40 & 11 & 0 & 0 & 1 & 46 \\ 11 & 25 & 3 & 0 & 0 & 14 \\ 2 & 3 & 3 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 46 & 14 & 2 & 0 & 0 & 45 \end{pmatrix}, \dots, \mathbf{F}^{(25)}$$

$$\mathbf{F}^{(31)} = \begin{pmatrix} 2 & 13 & 4 & 0 & 0 & 20 \\ 5 & 23 & 1 & 4 & 0 & 22 \\ 0 & 2 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 15 & 70 & 8 & 5 & 1 & 69 \end{pmatrix}, \mathbf{F}^{(32)} = \begin{pmatrix} 11 & 7 & 0 & 0 & 0 & 21 \\ 25 & 11 & 1 & 0 & 0 & 18 \\ 1 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 61 & 35 & 7 & 0 & 1 & 64 \end{pmatrix}, \dots \mathbf{F}^{(35)}$$
$$\mathbf{F}^{(41)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 37 & 9 & 5 & 0 & 38 \\ 3 & 11 & 0 & 1 & 0 & 21 \\ 5 & 19 & 1 & 1 & 1 & 24 \\ 2 & 8 & 0 & 0 & 0 & 7 \\ 5 & 34 & 3 & 2 & 0 & 24 \end{pmatrix}, \mathbf{F}^{(42)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 27 & 26 & 5 & 0 & 1 & 37 \\ 16 & 2 & 0 & 0 & 0 & 18 \\ 21 & 8 & 0 & 0 & 0 & 22 \\ 6 & 4 & 0 & 0 & 0 & 7 \\ 29 & 13 & 4 & 0 & 0 & 22 \end{pmatrix}, \dots \mathbf{F}^{(45)}$$
$$\mathbf{F}^{(51)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 34 & 10 & 6 & 0 & 40 \\ 2 & 12 & 0 & 1 & 0 & 18 \\ 7 & 21 & 1 & 1 & 1 & 30 \\ 2 & 9 & 0 & 0 & 0 & 3 \\ 5 & 33 & 2 & 1 & 0 & 23 \end{pmatrix}, \mathbf{F}^{(52)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 28 & 26 & 5 & 0 & 1 & 36 \\ 15 & 3 & 0 & 0 & 0 & 15 \\ 23 & 10 & 0 & 0 & 0 & 28 \\ 6 & 2 & 0 & 0 & 0 & 6 \\ 27 & 12 & 4 & 0 & 0 & 21 \end{pmatrix}, \dots \mathbf{F}^{(55)}$$

On the transition frequency matrix  $\mathbf{F}$ , we can see the highest transition frequency value in each transition frequency matrix, as follows:

1) The first row: The first row in matrix **F** consist of:  $\mathbf{F}^{11}$ ,  $\mathbf{F}^{12}$ ,  $\mathbf{F}^{13}$ ,  $\mathbf{F}^{14}$ , and  $\mathbf{F}^{15}$ , each of which has the highest value  $f_{22}^{11} = 51$ ,  $f_{66}^{12} = 58$ ,  $f_{66}^{13} = 71$ ,  $f_{22}^{14} = 37$ , and  $f_{26}^{15} = 37$ .

2) Second row: Second row in matrix **F** consist of:  $\mathbf{F}^{21}$ ,  $\mathbf{F}^{22}$ ,  $\mathbf{F}^{23}$ ,  $\mathbf{F}^{24}$ , and  $\mathbf{F}^{25}$ , each of which has the highest value  $f_{66}^{21} = 48$ ,  $f_{16}^{22} = f_{61}^{22} = 46$ ,  $f_{66}^{23} = 73$ ,  $f_{26}^{24} = 33$ ,  $f_{26}^{25} = 36$ .

3) Third row: Third row in matrix **F** consist of:  $\mathbf{F}^{31}$ ,  $\mathbf{F}^{32}$ ,  $\mathbf{F}^{33}$ ,  $\mathbf{F}^{34}$ , and  $\mathbf{F}^{35}$ , each of which has the highest value  $f_{26}^{31} = 70$ ,  $f_{66}^{32} = 64$ ,  $f_{66}^{33} = 105$ ,  $f_{26}^{34} = 60$ ,  $f_{26}^{35} = 63$ .

4) Fourth row: Fourth row in matrix **F** consist of:  $\mathbf{F}^{41}$ ,  $\mathbf{F}^{42}$ ,  $\mathbf{F}^{43}$ ,  $\mathbf{F}^{44}$ , and  $\mathbf{F}^{45}$ , each of which has the highest value  $f_{62}^{41} = 38$ ,  $f_{26}^{42} = 37$ ,  $f_{62}^{43} = 62$ ,  $f_{22}^{44} = 47$ ,  $f_{22}^{45} = 45$ .

5) Fifth row: Fifth row in matrix **F** consist of:  $\mathbf{F}^{51}$ ,  $\mathbf{F}^{52}$ ,  $\mathbf{F}^{53}$ ,  $\mathbf{F}^{54}$ , and  $\mathbf{F}^{55}$ , each of which has the highest value  $f_{62}^{51} = 40$ ,  $f_{62}^{52} = 36$ ,  $f_{62}^{53} = 63$ ,  $f_{22}^{54} = 46$ ,  $f_{22}^{55} = 49$ .

Therefore, the highest frequency value on matrix **F** is at  $\mathbf{F}^{33}$  that is  $f_{66}^{33} = 105$ , it means the highest value is found at the transition frequency matrix from product C to product C, from very fast moving to very fast-moving condition. This interprets that the amount of consumer demand for product C in the following month is still in product C (not moving to another product) and with the same conditions (very fast-moving),

#### B. The Transition Probability Matrix

**P** is the transition probability matrix for 5 types of products (s = 5), that is product A (1), product B (2), product C (3), product D (4), and product E (5), in six states (m = 6), no sales volume (1<sup>st</sup> state), very slow-moving (2<sup>nd</sup> state), slow-moving (3<sup>rd</sup> state), standard (4<sup>th</sup> state), fast-moving (5<sup>th</sup> state), and very fast-moving (6<sup>th</sup> state), here is the matrix:

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}^{(11)} & \mathbf{P}^{(12)} & \mathbf{P}^{(13)} & \mathbf{P}^{(14)} & \mathbf{P}^{(15)} \\ \mathbf{P}^{(21)} & \mathbf{P}^{(22)} & \mathbf{P}^{(23)} & \mathbf{P}^{(24)} & \mathbf{P}^{(25)} \\ \mathbf{P}^{(31)} & \mathbf{P}^{(32)} & \mathbf{P}^{(33)} & \mathbf{P}^{(34)} & \mathbf{P}^{(35)} \\ \mathbf{P}^{(41)} & \mathbf{P}^{(42)} & \mathbf{P}^{(43)} & \mathbf{P}^{(44)} & \mathbf{P}^{(55)} \end{pmatrix}$$
where  $\mathbf{P}^{(jk)} = \begin{pmatrix} p_{11}^{(jk)} & p_{21}^{(jk)} & p_{31}^{(jk)} & p_{41}^{(jk)} & p_{51}^{(jk)} & p_{61}^{(jk)} \\ p_{12}^{(jk)} & p_{22}^{(jk)} & p_{32}^{(jk)} & p_{42}^{(jk)} & p_{52}^{(jk)} & p_{62}^{(jk)} \\ p_{13}^{(jk)} & p_{23}^{(jk)} & p_{33}^{(jk)} & p_{43}^{(jk)} & p_{53}^{(jk)} & p_{63}^{(jk)} \\ p_{13}^{(jk)} & p_{24}^{(jk)} & p_{34}^{(jk)} & p_{43}^{(jk)} & p_{53}^{(jk)} & p_{64}^{(jk)} \\ p_{15}^{(jk)} & p_{25}^{(jk)} & p_{35}^{(jk)} & p_{45}^{(jk)} & p_{55}^{(jk)} & p_{65}^{(jk)} \\ p_{16}^{(jk)} & p_{26}^{(jk)} & p_{36}^{(jk)} & p_{46}^{(jk)} & p_{56}^{(jk)} & p_{66}^{(jk)} \end{pmatrix}$ 

for *j* = 1, 2...6 and *k* = 1, 2...,6

The following are the results of calculating the transition frequency using the MATLAB software:

The number of probability value for each column is 1 in accordance with the nature of the transition opportunities described previously in equation (2).

Based on the value of the **P** matrix above, we can see the highest probability value with the highest probability value is in  $\mathbf{P}^{33}$  that is  $p_{66}^{33} = 0.625$  with frequency 105 times, it means the highest value is found at the transition probability matrix from product C to product C, from very fast-moving to very fast-moving

condition. This interprets that the probability for consumers to buy product C in the next month remains, don't move to another product.

## V. CONCLUSION

The transition probability matrix is one of the parameters in the multivariate Markov chain model. The transition probability matrix can be calculated by calculating the transition frequency first, then normalizing it to the number of columns of each transition frequency matrix. The highest transition probability value for the sales demand in a company is found at the transition probability matrix from product C to product C, from very fast moving to very fast-moving condition, which had the highest probability value 0.625 with the highest frequency 105 times.

### ACKNOWLEDGMENT

I would like to thank Dr. Agus Yodi Gunawan, S.Si., M.Si. for as the thesis supervisor, which is the idea of writing in this paper.

# REFERENCES

- [1] A. Martina, "Penggunaan Model Rantai Markov Multivariat Untuk Estimasi Permintaan Penjualan Pada Suatu Perusahaan", Thesis, *Indonesia: Bandung Institute of Technology*, 2015
- [2] L. Megasalindri, "Prediksi Permintaan Penjualan dengan Menggunakan Model Rantai Markov Multivariat" undergraduate final project, *Indonesia: Bandung Institute of Technology*,2013
- [3] W. Ching, Li, L., Li, T., Zhang, S., "A New Multivariate Markov Chain Model with Applications to Sales Demand Forecasting" in *China: International Conference on Industrial Engineering and* Systems Management, 2007.
- [4] W. Ching and Ng. Michael K. "Markov Chains: Models, Algorithms and Applications", *United States of America: Springer+Business Media, Inc.*, 2006.