Markov Chain

11

Hao Zhang

Learning Goals

1. Understand the Markov process and the basic concept of Markov chain.
2. Understand the classification of Markov models, the connotation of continuous-time Markov chain and hidden Markov models.
3. Clarify the application scope and application methods of Markov model, and use Markov model to make predictions.

Introduction
Study Notes from a Project Manager

Due to the needs of the project, I checked the literature about marketing application of Markov chain and made some attempts.

Markov chain is a predictive tool. Given that the choice space faced by customers can be divided into n mutually exclusive states, the long-term trend of customers can be described by their transitions between different states. Transferring has two important characteristics: randomness and being of no aftereffect (or being memoryless). It's like a walk, metaphorically speaking, without a set goal. Each step depends only on where the previous step went, and there are several possibilities. This feature enables the prediction of customers' long-term behavior to be divided into several independent units, and the state of each time point is determined by the state of the previous moment and the transition probability matrix representing all possibilities, which is very flexible.

H. Zhang (✉)
School of E-commerce and Logistics, Beijing Technology and Business University, Beijing, China
e-mail: zhhaozhhao@126.com

Markov chain fits the description of many economic phenomena. The most typical example is the stock market. When you are buying wealth management products, funds for example, you will notice the company's statement that the past performance is not necessarily indicative of future results, which is a live embodiment of Markov chain. The stock market is irregular. Studies have proved that the accuracy of using historical data to predict stocks or the stock market trends is not better than that of flipping a coin. In the field of marketing, such as online analysis, we often use path analysis to find out how customers use the website. However, path analysis assumes that the customer's browsing process has certain rules to follow, and the browsing behavior of the customer is actually more consistent with the Markov process. From one page to another, it is completely stochastic, so it is more accurate to describe it with the Markov model. Are all problems applicable to Markov process? A Markov chain requires tracking an object (a customer, for example) as it moves between different states over time, and its behavior is repetitive. The customer can only choose one state at a time: either to stay in the current state, or to enter another state. Some researchers use Markov models to predict the direction of the real estate market. However, most property purchases are one-off without continuity; the customers who have purchased multiple estates are more likely to own more than one estates, and their purchase behaviors are highly related with previous purchase experiences. Thus it is appropriately to describe the latter by other models.

The establishment of Markov chain model itself is not complicated, with three steps necessarily: ① Set the state; ② Calculate the transition probability matrix; ③ Calculate the result of the transition. The status may be given, such as different brands and web pages; or it needs to be divided according to the results of data analysis. In theory, the more states, the more accurate the prediction results. However, too many classifications will lose marketing significance and cause difficulties in use. RFM (Recency Frequency Monetary) is a common state division method to predict customer lifetime value using Markov chain. RFM can express customer's transaction status and reflect the relationship between customers and companies. There is a problem of optimization when using RFM to distinguish customer status. If there are many variables, cluster analysis or decision trees can be considered. The transition probability matrix can be constructed by using observations directly or by assigning values based on expert opinions. It is available to use models such as multi-state logistic regression, decision trees, neural networks, or stochastic function models. The advantage of models is that the noise in the observed data can be eliminated, and thus the transition probability can be refined to the individual. Whether the transition probability is stable is a problem that needs special attention. Markov model assumes that the transition probability of customers between different states is constant and does not change with time. Therefore, the estimation of the customer's state at each time point is carried out through the iteration of a single transition matrix. But for a longer time frame, especially in industries with long-term relationships with customers, such as banking and insurance, this assumption needs to be revised. Some important events,

such as marriage, childbirth, retirement, etc., will inevitably change the transition probability of customers.

Whether the application of Markov chain can properly predict the market and the value of customers, like other models, needs to be tested by practice. The performance of the model after landing is the final judgement. However, validation with data is a necessary step in the modeling process.

Source   http://blog.sina.com.cn/s/blog_6520908501017qv9.html   (accessed 09/01/2013).

11.1   Markov Process

Markov process is a kind of stochastic process. Its original model is the Markov chain, which was proposed by the Russian mathematician A. A. Markov in 1906.

Markov process is a typical stochastic process. The theory studies the state of a system and its transfer. It determines the change trend of the state by studying the initial probability of different states and the transition probability between states, so as to predict the future.

Markov process has two basic characteristics (Markov property). One is "no aftereffect", that is, the future state of a thing and the probability of its occurrence only depends on the state of the thing now, and has nothing to do with the state of the previous time. In other words, it does not depend on its past evolution. The other is "ergodicity", that is, no matter what state things are in, the Markov process gradually tends to be stable over a long period of time, and it has nothing to do with the initial state. In the practice, many processes are Markov processes, such as Brownian motion caused by particles in liquids, the number of people infected with infectious diseases, inventory problems in stores, and the queues at banks, etc.

Markov processes are expressed mathematical as:

Definition 11.1 Let $X(t), t \in T$ be a stochastic process. If $X(t)$ is observed at the time of $t_1, t_2, \cdots, t_{n-1}, t_n (t_1 < t_2 < \cdots < t_{n-1} < t_n \in T)$, the corresponding observed value $x_1, x_2, \cdots, x_{n-1}, x_n$ satisfied the condition.

$$
\begin{aligned}
&P\{X(t_n) \leq x_n \,|X(t_{n-1}) = x_{n-1}, X(t_{n-2}) = x_{n-2}, \ldots, X(t_1) = x_1\} \\
&= P\{X(t_n) \leq x_n \,|X(t_{n-1}) = x_{n-1}\}
\end{aligned} \tag{11.1}
$$

or

$$
F_X(x_n; t_n|x_{n-1}, x_{n-2}, \ldots, x_2, x_1; t_{n-1}, t_{n-2}, \ldots, t_2, t_1) = F_X(x_n; t_n|x_{n-1}; t_{n-1}) \tag{11.2}
$$

such a process is called the process with Markov properties or Markov process.

where

$$F_X(x_n; t_n | x_{n-1}, x_{n-2}, \ldots, x_2, x_1; t_{n-1}, t_{n-2}, \ldots, t_2, t_1)$$

represents the conditional distribution function when the value of time $X(t_n)$ is x_n under the condition of $X(t_{n-1}) = x_{n-1}, X(t_{n-2}) = x_{n-2}, \cdots, X(t_1) = x_1$.

If time t_{n-1} is regarded as "present", because $t_1 < t_2 < \cdots < t_{n-1} < t_n$, then t_n can be regarded as "future", and $t_1, t_2, \cdots, t_{n-2}$ as "past". Therefore, the above definition can be expressed as that the value of $X(t_n)$ in the future is independent of the value of $X(t_1), X(t_2), \cdots, X(t_{n-2})$ in the past, given the value of $X(t_{n-1})$ in the present state is x_{n-1}.

11.2   Markov Chain

11.2.1 Definition

Markov chain refers to a Markov process in which time and state parameters are discrete. It is the simplest Markov process.

The time studied in general Markov process is infinite, and it is a continuous variable with continuous numerical values. Two adjacent values can be divided infinitely, and the states of study are infinite. The time parameters of Markov chain are discrete values. In economic forecast, the general dates are days, months, seasons and years. At the same time, the state of the Markov chain is finite. For example, the market sales state can be "salable" and "unsalable". The future state of the market is only related to the current state, and not to the previous state (no aftereffect is established).

It is described mathematical as:

Definition 11.2 If the stochastic process $X(n), n \in T$ satisfies the following conditions:

(1) The time set is taken as a non-negative integer set $T = \{0, 1, 2, \cdots\}$ corresponding to each moment. The state space is a discrete set, denoted as $E = \{E_0, E_1, E_2, \cdots\}$, that is, $X(n)$ is a discrete time state.
(2) For any integer $n \in T$, the conditional probability satisfies:

$$\begin{aligned} P\{X(n+1) = &E_{n+1} | X(n) = E_n, X(n-1) = E_{n-1}, \ldots, X(0) = E_0\} \\ = &P\{X(n+1) = E_n | X(n) = E_n\} \end{aligned} \tag{11.3}$$

then call $X(n), n \in T$ a Markov chain, and denote that

$$P_{ij}^{(k)} = P\{X(m+k) = E_j | X(m) = E_i\}, E_i, E_j \in E \qquad (11.4)$$

represent the probability that the system is in state E_i at time m and the system is in the state E_j at time $m + k$.

The conditional probability equation, that is, the probability of the state $X(m+k) = E_j$ of $X(n)$ at time $X(m+k) = E_j$ is only related to the state $X(m) = E_i$ at time m, and is independent of the state before m. It is one of the mathematical expressions of Markov property (no aftereffect). No aftereffect means that once the state of a certain stage is determined, the evolution of the subsequent process will no longer be affected by various previous states and decisions.

11.2.2 Relevant Concepts

1. **State and state variables**

State: A condition in which an objective thing may appear or exist. For example, goods may be salable or may be unsalable; the machine may ran normally or may not work properly.

Different states of the same thing must be mutually independent: two states cannot exist at the same time. The state of objective things is not fixed. When conditions change, the state often changes. For example, a product is originally unsalable in the market, but by reason of promotion and other factors, it may become a best-selling one.

State variables are generally used to represent the state: $X_t = {}_i\binom{i = 1, 2, \cdots, N}{t = 1, 2, \cdots}$, which represents a stochastic motion system, and at time $t(t = 1, 2, \cdots)$, the state is $i(i = 1, 2, \cdots N)$.

2. **State transition probability and its transition probability matrix**

(1) One-step transition probability matrix. Suppose the state space of the system is $E = (E_1, E_2, \ldots, E_n)$, and each time the system can only be in one of these states, so each state has n turns (including turns to itself), that is

$$E_i \to E_1, E_i \to E_2, \ldots, E_i \to E_i, \ldots, E_i \to E_n$$

Under the condition that the system is in state E_i at time m, the conditional probability of the system in state E_j at time $m + k$ can be expressed as:

$$P_{ij}^{(k)} = P\{X(m+k) = E_j | X(m) = E_i\}, E_i, E_j \in E \qquad (11.5)$$

In particular, when $k = 1$,

$$p_{ij} = P\{X(m+1) = E_j | X(m) = E_i\}, E_i, E_j \in E$$

that is, when the system is in state E_i at time m, the conditional probability that the system is in state E_j at time $m + 1$ is called the transition probability from state E_i to state E_j through one transition. The matrix formed by the set of one-step transition probabilities of all the states of the system is called one-step state transition probability matrix. Its form is as follows:

$$P = \begin{matrix} & E_1 \ E_2 \ \cdots \ E_n \\ \begin{matrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{matrix} & \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \end{matrix} \tag{11.6}$$

This matrix has the following two properties:

Non-negative: $p_{ij} \geq 0, i, j = 1, 2, \cdots, n$

The sum of row elements is 1, that is $\sum_{j=1}^{n} p_{ij} = 1, i = 1, 2, \cdots, n$

Example 11.1 There are three garment factories A, B and C producing the same kind of clothing, and there are 1000 customers. It is assumed that during the study period, no new users join and no old users quit, only some customers are transferred. It is known that in April, 500 are customers to factory A; 400 customers to B; 100 customers to C. In May, A had 400 original customers left, with 50 transferring to B and 50 to C. B had 300 original customers left, with 20 transferring to A and 80 to C. C had 80 original customers left, with 10 transferring to A and 10 to B.

Calculate its state transition probability.

Solution:
The customer transfer in May is shown in Table 11.1.

Table. 11.1 Customer transfer in May

	A	B	C	Total
A	300	50	50	400
B	120	300	80	500
C	20	20	60	100
Total	440	370	190	1000

$$P_{11} = 400/500 = 0.8 \quad P_{12} = 50/500 = 0.1 \quad P_{13} = 50/500 = 0.1$$
$$P_{21} = 20/400 = 0.05 \quad P_{22} = 300/400 = 0.75 \quad P_{23} = 80/400 = 0.2$$
$$P_{31} = 10/100 = 0.1 \quad P_{32} = 10/100 = 0.1 \quad P_{33} = 80/100 = 0.8$$

State transition probability matrix:

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.05 & 0.75 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

(2) k-step transition probability matrix. According to the definition of one-step transition probability, k-step transition probability is the probability of the system's transition from state E_i to state E_j through k times, which can be expressed as

$$P_{ij}^{(k)} = P\{X(m+k) = E_j | X(m) = E_i\}, E_i, E_j \in E$$

Therefore, the k-step transition probability matrix of the system is a matrix composed of the k-step transition probability sets of all states. The form is as follows:

$$P^{(k)} = \begin{matrix} & \begin{matrix} E_1 & E_2 & \cdots & E_n \end{matrix} \\ \begin{matrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{matrix} & \begin{pmatrix} p_{11}^{(k)} & p_{12}^{(k)} & \cdots & p_{1n}^{(k)} \\ p_{21}^{(k)} & p_{22}^{(k)} & \cdots & p_{2n}^{(k)} \\ \vdots & \vdots & & \vdots \\ p_{n1}^{(k)} & p_{n2}^{(k)} & \cdots & p_{nn}^{(k)} \end{pmatrix} \end{matrix} \qquad (11.8)$$

This matrix has the following three properties:

Non-negative: $p_{ij}^{(k)} \geq 0, i, j = 1, 2, \ldots, n$

The sum of row elements is 1, that is $\sum_{j=1}^{n} p_{ij}^{(k)} = 1, i = 1, 2, \ldots, n$

$P^{(n)} = P^{(n-1)} P = P^n$

Example 11.2 The market of Borui Company has three states: E1, E2 and E3 (that is salable, ordinary and unsalable). The market transfer of the company is shown in Table 11.2. Try to find the two-step state transition probability matrix of the company's market.

Table. 11.2 Company market states transition

The current state of the company's market	The next state of the company's market		
	E1	E2	E3
E1	21	7	14
E2	16	8	12
E3	10	8	2

Solution:

First write down the one-step transition probability matrix

$$P^{(1)} = \begin{bmatrix} 0.500 & 0.167 & 0.333 \\ 0.444 & 0.222 & 0.334 \\ 0.500 & 0.400 & 0.100 \end{bmatrix}$$

The two-step state transition probability matrix can be calculated from the one-step transition probability matrix by the formula $P^{(n)} = P^n$:

$$P^{(2)} = P^2 = \begin{bmatrix} 0.500 & 0.167 & 0.333 \\ 0.444 & 0.222 & 0.334 \\ 0.500 & 0.400 & 0.100 \end{bmatrix}^2 = \begin{bmatrix} 0.491 & 0.254 & 0.255 \\ 0.488 & 0.257 & 0.255 \\ 0.478 & 0.212 & 0.310 \end{bmatrix}$$

(3) **Steady-State Probability**

The state probability is the steady-state probability when the Markov chain reaches a steady state. Under certain conditions, the Markov chain will reach a stable state after k-step transfer.

(1) Conditions of stable state. If the one-step transition probability matrix is a normal probability matrix, the Markov chain can reach a stable state.
(2) Solving the steady-state probability. According to the definition of steady state of Markov chain, when in a stable state, there is $S^{(k+1)} = S^{(k)}$, that is $S^{(k+1)} = S^{(k)} P = S^{(k)}$.

Assume $\begin{cases} S^{(k)} = (x_1, x_2, \cdots, x_n) \\ S^{(k+1)} = S^{(k)} \cdot P = S^{(k)} \end{cases}$, and $\sum_{i=1}^{n} x_i = 1$ is the state vector after k-step transition. The one-step transition probability matrix is

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix}$$

According to $S^{(k+1)} = S^{(k)} \cdot P = S^{(k)}$, it is expanded to

$$(x_1, x_2, \ldots, x_n) \begin{bmatrix} P_{11} & \ldots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \ldots & P_{nn} \end{bmatrix} = S^{(k)} = (x_1, x_2, \ldots, x_n) \qquad (11.9)$$

By calculation, the following equation set is obtained:

$$\begin{cases} P_{11}x_1 + P_{21}x_2 + \ldots + P_{n1}x_n = x_1 \\ P_{12}x_1 + P_{22}x_2 + \ldots + P_{n2}x_n = x_2 \\ \qquad\qquad \vdots \\ P_{1n}x_1 + P_{2n}x_2 + \ldots + P_{nn}x_n = x_n \\ \qquad x_1 + x_2 + \ldots + x_n = 1 \end{cases} \qquad (11.10)$$

The shift term is transformed into

$$\begin{cases} (P_{11} - 1)x_1 + P_{21}x_2 + \ldots + P_{n1}x_n = 0 \\ P_{12}x_1 + (P_{22} - 1)x_2 + \ldots + P_{n2}x_n = 0 \\ \qquad\qquad \vdots \\ P_{1n}x_1 + P_{2n}x_2 + \ldots + (P_{nn} - 1)x_n = 0 \\ \qquad x_1 + x_2 + \ldots + x_n = 1 \end{cases} \qquad (11.11)$$

There are n variables in Eq. (11.11), but there are $n + 1$ equations, indicating that one of the equations is not independent and the nth equation needs to be eliminated:

$$\begin{bmatrix} (P_{11} - 1) & P_{21} & \ldots & P_{n1} \\ P_{12} & (P_{22} - 1) & \ldots & P_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \qquad (11.12)$$

let

$$P_1 = \begin{bmatrix} (P_{11} - 1) & P_{21} & \ldots & P_{n1} \\ P_{12} & (P_{22} - 1) & \ldots & P_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \ldots & 1 \end{bmatrix}, X^{((n))} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

then

$$P_1 X^{(n)} = B$$

$$X^{(n)} = P_1^{-1} B$$

that is, $X^{(n)}$ is the steady-state probability of the Markov chain.

11.3   Classification of Markov Chain Models

11.3.1 Continuous-Time Markov Chains

Definition 11.3 Suppose the stochastic process $\{X(t), t \geqslant 0\}$, the state space $I = \{i_n, n \geqslant 0\}$, if for any and $0 \leqslant t_1 < t_2 < \ldots < t_{n+1}$ $i_1, i_2, \ldots, i_{n+1} \in I$, there is.

$$P\{X(t_{n+1}) = i_{n+1} \,|\, X(t_1) = i_1, X(t_2) = i_2, \ldots, X(t_n) = i_n\}$$
$$= P\{X(t_{n+1}) = i_{n+1} \,|\, X(t_n) = i_n\} \tag{11.13}$$

then call $\{X(t), t \geqslant 0\}$ as a continuous-time Markov chain.

In the above formula, the conditional probability is expressed as

$$P\{X(s + t) = j \,|\, X(s) = i\} = p_{ij}(s, t)$$

Definition: If the transition probability of $p_{ij}(s, t)$ is independent of s, then it is said that the continuous-time Markov chain has a stationary or homogeneous transition probability, and then the transition probability is abbreviated as

$$p_{ij}(s, t) = p_{ij}(t)$$

its transition probability matrix is abbreviated as

$$P(t) = \big(p_{ij}(t)\big)$$

A continuous-time Markov chain, whenever it enters state i, has the following properties:

(1) The time in state i before moving to another state follows an exponential distribution with parameter v_i;
(2) When the process leaves state i, it then enters state j with probability p_{ij}, $\sum_{j \neq i} p_{ij} = 1$.

When $v_i = \infty$, state i is called an instantaneous state;

When $v_i = 0$, state i is called the absorbed state.

A continuous-time Markov chain transfers from one state to another according to a discrete time Markov chain. But before transferring to the next state, the time it stays in each state obeys an exponential distribution. In addition, the staying time in the state i process and the next arrival state must be independent stochastic variables.

11.3.2 Hidden Markov Model

Hidden Markov model (HMM) is a kind of Markov chain, whose states cannot be observed directly, but can be observed through a sequence of observation vectors. Each observation vector is expressed in various states through some probability density distributions. Each observation vector is generated by a state sequence with corresponding probability density distribution. Therefore, HMM is a double stochastic process, a hidden Markov chain with a certain number of states and a set of display stochastic functions. Since the twentieth century, HMM has been applied to speech recognition, computer character recognition, mobile communication core technology "multi-user detection", bioinformatics science, fault diagnosis and other fields.

HMM can be described by five elements, including two state sets and three probability matrices.

(1) The hidden state S. These states satisfy the Markov property among them and are the actually implied in the Markov model. These states are usually not available by direct observation (such as S_1, S_2, S_3, etc.).
(2) The observable state O. It is associated with the hidden state in the model and can be obtained by direct observation. (such as O_1, O_2, O_3, etc., the number of observable states is necessarily not the same as the number of hidden states.)
(3) The initial state probability matrix π. Represents the probability matrix of the hidden state at the initial time $t = 1$. For example, when $t = 1, P(S_1) = p_1$, $P(S_1) = p_2$, $P(S_1) = p_3$, then the initial state probability matrix $\pi = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix}$.
(4) Hidden state transition probability matrix A. It describes the transition probability between states in the hidden Markov model. Where $A_{ij} = P(S_j|S_i)$, $1 \leqslant i, j \leqslant N$ is the probability that the state is S_j at time $t + 1$ under the condition that the state is S_i at time t.
(5) Observation state transition probability matrix B. Let N represent the number of hidden states and M represent the number of hidden states, then

$$B_{ij} = P(O_i|S_j), 1 \leq i \leq M, 1 \leq j \leq N$$

represents the probability that the observation state is O_i at time t and the hidden condition state is S_j.

HMM can be succinctly represented by $\lambda = (A, B, \pi)$ triples. HMM adds the set of observable states and the probability relationship between these states, which is actually an extension of the standard Markov model.

11.4   Application of Markov Chain Models

Markov analysis, also known as Markov transition matrix method, refers to a prediction method that predicts future changes of stochastic variables by analyzing the current changes of these variables under the assumption of Markov process.

The simplest type of Markov chain prediction method is to predict the most likely state in the next period. Here are the steps:

Step 1: Divide the states of the predicted object. Starting from the prediction purposes, consider the decision making needs to classify the state of the phenomenon.

Step 2: Calculate the initial probability. The state probability obtained by analyzing historical data of practical problems is called the initial probability.

Step 3: Calculate the state transition probability.

Step 4: Make prediction according to transition probability.

From the state transition probability matrix P, if the prediction object is currently in state E_i, then P_{ij} describes the possibility that the current state E_i will change to state $E_j (j = 1, 2, …, N)$ in the future. According to the maximum possibility as the selection principle: choose the largest of P_{j1}, P_{j2}, …, P_{jN} as the prediction result.

1. **Calculate market share**

Example 11.3 Guangzhou, Shenzhen and Macau Special Administrative Region of the People's Republic of China produce and sell certain food ingredients. It is necessary to predict the market share in the next few months. The specific steps are as follows:

Step 1: Conduct market survey

(1) Current market share (the proportion of customers purchasing food ingredients from Guangzhou, Shenzhen and Macau).

Results: the customers who bought Guangzhou ingredients accounted for 40%, those to Shenzhen and Macau accounted for 30% respectively, and (40%, 30%, 30%) is called the current market share distribution or the initial distribution.

(2) Investigate the flow of customers.

The flow condition is:

① 40% of the customers who bought food ingredients from Guangzhou last month remain this month, and 30% of them transferred to Shenzhen and Macau respectively.

Table. 11.3 The flow of ingredients purchased by customers

	Guangzhou (%)	Shenzhen (%)	Macau (%)
Guangzhou	40	30	30
Shenzhen	60	30	10
Macau	60	10	30

② 60% of the customers who bought the ingredients from Shenzhen last month transferred to Guangzhou this month, 30% remain, and 10% to Macau.

③ 60% of the customers who bought food ingredients from Macau transferred to Guangzhou, 10% to Shenzhen, and 30% remain.

Step 2: Establish a mathematical model.

For the convenience of calculation, 1, 2 and 3 represent the food ingredients in Guangzhou, Shenzhen and Macau respectively. According to the results of market survey, the flow of customers' purchase of food ingredients is shown in Table 11.3.

$$P = \begin{pmatrix} P11 & P12 & P13 \\ P21 & P22 & P23 \\ P31 & P32 & P33 \end{pmatrix} = \begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.6 & 0.3 & 0.1 \\ 0.6 & 0.1 & 0.3 \end{pmatrix}$$

Step 3: Make market forecasts.

Suppose the initial market share distribution is $(P_1, P_2, P_3) = (0.4, 0.3, 0.3)$, and the market share distribution after three months is $(P_1(3), P_2(3), P_3(3))$.

If the trend of customer flow is stable for a long time, the market share will reach a stable equilibrium after a period of time.

$$(P_1(n), P_2(n), P_3(n)) = (P_1, P_2, P_3) \begin{pmatrix} P_{11}(n) & P_{12}(n) & P_{13}(n) \\ P_{21}(n) & P_{22}(n) & P_{23}(n) \\ P_{31}(n) & P_{32}(n) & P_{33}(n) \end{pmatrix}$$

$$= (P_1, P_2, P_3) \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}$$

A stable market equilibrium means that the number of customers lost to each product is offset by the number of new customers gained during the flow of customers.

Step 4: Forecast long-term market share.

Since the one-step transition probability matrix P is a normal probability matrix, the long-term market share is the market share under equilibrium state, that is, the stationary distribution of the Markov chain.

Let the long-term market share be

$$X = (x_1, x_2, x_3)$$

then

$$\begin{cases} (x_1, x_2, x_3) \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.6 & 0.3 & 0.1 \\ 0.6 & 0.1 & 0.3 \end{bmatrix} = (x_1, x_2, x_3) \\ x_1 + x_2 + x_3 = 1 \end{cases}$$

such that

$$X = (x_1, x_2, x_3) = (0.5, 0.25, 0.25)$$

2. Human resources forecast

Example 11.4 The employees of Borui Company are divided into five categories: intern, ordinary staff, director, general manager, and former staff. The current status (550 employees) is expressed as:

$$P(0) = (135, 240, 115, 60, 0)$$

The Company's previous record is

$$P = \begin{pmatrix} 0.6 & 0.4 & 0 & 0 & 0 \\ 0 & 0.6 & 0.25 & 0 & 0.15 \\ 0 & 0 & 0.55 & 0.21 & 0.24 \\ 0 & 0 & 0 & 0.8 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Try to analyze the structure of employees after three years and how many new employees should be recruited into the workforce while keeping the distribution of employees unchanged (550) in three years.
Solution:
Distribution of employees after one year:

$$(1) = P(0) \cdot P$$

$$= (135, 240, 115, 60, 0) \begin{pmatrix} 0.6 & 0.4 & 0 & 0 & 0 \\ 0 & 0.6 & 0.25 & 0 & 0.15 \\ 0 & 0 & 0.55 & 0.21 & 0.24 \\ 0 & 0 & 0 & 0.8 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= (81, 198, 123, 72, 76)$$

To maintain the total number of 550, 76 have left, so 76 new employees should be recruited in the first year:

$$P'(1) = (81 + 76, 198, 123, 72, 0)$$

Distribution of employees after the second year:

$$P(2) = P'(1) \cdot P$$

$$= (157, 198, 123, 72, 0) \begin{pmatrix} 0.6 & 0.4 & 0 & 0 & 0 \\ 0 & 0.6 & 0.25 & 0 & 0.15 \\ 0 & 0 & 0.55 & 0.21 & 0.24 \\ 0 & 0 & 0 & 0.8 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= (94, 182, 117, 83, 74)$$

To keep the total number of people unchanged, 74 employees should be added:

$$P'(2) = (94 + 74, 182, 117, 83, 0)$$

Distribution of employees after the third year:

$$P(3) = P'(2) \cdot P$$

$$= (168, 182, 117, 83, 0) \begin{pmatrix} 0.6 & 0.4 & 0 & 0 & 0 \\ 0 & 0.6 & 0.25 & 0 & 0.15 \\ 0 & 0 & 0.55 & 0.21 & 0.24 \\ 0 & 0 & 0 & 0.8 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= (101, 176, 111, 91, 72)$$

72 employees should be added. At the end of the third year, the staff structure is

$$P'(3) = (173, 176, 111, 91, 0)$$

3. Profit forecast

The state in the nth period is represented by X_n:

$$X_n = \begin{cases} 1, & products\ of\ the\ n-th\ period\ are\ salable \\ 2, & products\ of\ the\ n-th\ period\ are\ unsalable \end{cases}$$

Let $\{X_n\}$ be a homogeneous Markov chain with a state space of $S = \{1, 2, \ldots, N\}$, and its transition matrix is $P = (P_{ij})_{N \times N}$.

Let $r(i)$ denote the profit obtained when the system is in state i during a certain period. Such a Markov chain is called profitable.

When $r(i) > 0$, it is called profit; when $r(i) < 0$, it is called expense.

(1) Total expected profit for a limited period. Let $vk(i)$ denote the expected total profit $(k \geqslant 1, i \in S)$ obtained before the state transition at of the kth step under the condition that the initial state is:

$$vk(i) = \sum_{n=0}^{k-1} \text{expected profit for the nth period}$$

$$= \sum_{n=0}^{k-1} E\{r(X_n)|X_0 = i\}$$

$$= \sum_{n=0}^{k-1} (\sum_{j=1}^{N} p_{ij}^{(n)} r(j)) \tag{11.14}$$

Example 11.5 shows that $k = 4$, the current month is recorded as the first month, and find the expected profit $v_4(1)$ obtained before the fourth step of the state (that is, the first four months).

Let $r(i)$ denote the profit obtained when the system is in state i in a certain period, let 1 denote "salable", and 2 denote "unsalable". Then

$$v_4(1) = r(1) + \sum_{n=1}^{4-1} \left[p_{11}^{(n)} r(1) + p_{12}^{(n)} r(2) \right]$$

$$v_4(1) = r(2) + \sum_{n=1}^{4-1} \left[p_{21}^{(n)} r(1) + p_{22}^{(n)} r(2) \right]$$

$$v_4 = (v_4(1), v_4(2))^T$$

$$P^{(n)} = \begin{bmatrix} p_{11}^{(n)} & p_{12}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} \end{bmatrix}$$

$$r = (r(1), r(2))^T$$

$$v_4 = r + \sum_{n=1}^{4-1} p^{(n)} r = \sum_{n=0}^{4-1} p^{(n)} r$$

$$= \left(\sum_{n=0}^{4-1} p^{(n)} \right) r = (E + P + P^2 + P^3) r \tag{11.15}$$

Table. 11.4   Sales records for the past 24 months

Month	1	2	3	4	5	6
Sales status	Salable	Salable	Unsalable	Salable	Unsalable	Unsalable
Month	7	8	9	10	11	12
Sales status	Salable	Salable	Unsalable	Salable	Salable	Unsalable
Month	13	14	15	16	17	18
Sales status	Salable	Salable	Unsalable	Unsalable	Salable	Salable
Month	19	20	21	22	23	24
Sales status	Unsalable	Salable	Salable	Unsalable	Salable	Salable

Example 11.5 The electronic products produced by Bo Rui Company have two kinds of monthly market conditions: salable and unsalable. If the product is salable, it will make a profit of 500,000 yuan; if the product is unsalable, it will result in a loss of 300,000 yuan. The survey recorded sales over the past 24 months as shown in Table 11.4.

Question: If the products are salable in the current month, take the current month as the first month, and find the total expected profit before the fourth step of the state transition (that is, the first four months).

Solution:
Let 1 for "salable" and 2 for "unsalable". Given r

$$r = \begin{bmatrix} r(1) \\ r(2) \end{bmatrix} = \begin{bmatrix} 50 \\ -30 \end{bmatrix}$$

$i = 1$, V_4 has three forms of formula:

$$v_4(1) = r(1) + \sum_{n=1}^{4-1} \left[p_{11}^{(n)} r(1) + p_{12}^{(n)} r(2) \right]$$

$$v_4 = \left(\sum_{n=0}^{4-1} p^{(n)} \right) r = (E + P + P^2 + P^3) r$$

$$v_4(i) = r(i) + \sum_{j=1}^{2} p_{ij} v_3(j), i = 1, 2$$

$$v_0(i) = 0, i = 1, 2, \ldots, N$$

Find the state transition probability matrix P.

Estimate the state transition matrix P, and estimate the probability of continuously salable with statistical frequency.

$$p_{11} = \frac{7}{15 - 1} = 50\%$$

The numerator 7 is the number of continuously "salable" occurrences in the Table 11.4. The denominator 15 is the number of times "salable" appears in the Table 11.4. Since the 24th month was "salable" and there was no subsequent record, so it is reduced by 1.

$$p_{12} = \frac{7}{15 - 1} = 50\%, \ p_{21} = \frac{7}{9} = 78\%, \ p_{22} = \frac{2}{9} \approx 22\%$$

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.78 & 0.22 \end{bmatrix}$$

$$r = \begin{bmatrix} r(1) \\ r(2) \end{bmatrix} = \begin{bmatrix} 50 \\ -30 \end{bmatrix}, \ P = \begin{bmatrix} 0.5 & 0.5 \\ 0.78 & 0.22 \end{bmatrix}$$

$$v_4 = \left(\sum_{n=0}^{4-1} p^{(n)} \right) r = (E + P + P^2 + P^3) r$$

$$v_4 = \begin{bmatrix} 1.875 & 0.875 \\ 1.86295 & 0.27905 \end{bmatrix} \begin{bmatrix} 50 \\ -30 \end{bmatrix} = \begin{bmatrix} 67.5 \\ 54.776 \end{bmatrix}$$

$$v_4(1) = 67.5$$

The result is: If the products are salable in the current month, the total expected profit obtained in the first four months is 675,000 yuan.

(2) Average profit per unit time in unlimited period

For $i \in S$, the average profit per unit time for an unlimited period with the initial state of i is defined as

$$v(i) = \lim_{k \to \infty} \frac{V_k(i)}{k} \tag{11.16}$$

Denote

$$v = [v(1) v(2) \dots v(N)]^T, \ V_k = [v_k(1), v_k(2), \dots, v_k(N)]^T$$

$$V_k = \left(\sum_{n=0}^{k-1} P^n \right) r = (E + P + P^2 + \dots + P^{k-1}) r$$

then

$$V = \lim_{k \to \infty} \frac{V_k}{k} = \lim_{k \to \infty} \frac{(E + P + P^2 + \dots + P^{k-1}) r}{K}$$

If the Markov chain considered has a stationary distribution:

$$
P^m = \begin{bmatrix} p_{11}^{(m)} & p_{12}^{(m)} & \cdots & p_{1N}^{(m)} \\ p_{21}^{(m)} & p_{22}^{(m)} & \cdots & p_{2N}^{(m)} \\ \vdots & \vdots & & \vdots \\ p_{N1}^{(m)} & p_{N2}^{(m)} & \cdots & p_{NN}^{(m)} \end{bmatrix} \rightarrow \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_N \\ \pi_1 & \pi_2 & \cdots & \pi_N \\ \vdots & \vdots & & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_N \end{bmatrix}
$$

It can be proved that:

$$
v = \lim_{k \to \infty} \frac{V_k}{k} = \lim_{k \to \infty} \frac{\left(E + P + P^2 + \cdots + P^{k-1}\right)r}{K} = \lim_{k \to \infty} P^k r
$$

$$
= \begin{bmatrix} \pi_1 & \pi_2 & \ldots & \pi_N \\ \pi_1 & \pi_2 & \ldots & \pi_N \\ \vdots & \vdots & \ddots & \vdots \\ \pi_1 & \pi_2 & \ldots & \pi_N \end{bmatrix} \begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(N) \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{N} \pi_j r(j) \\ \sum_{j=1}^{N} \pi_j r(j) \\ \vdots \\ \sum_{j=1}^{N} \pi_j r(j) \end{bmatrix} \tag{11.17}
$$

The average profit per unit time of unlimited period has nothing to do with the initial state, which is expressed as

$$
v(i) = \sum_{j=1}^{N} \pi_j r(j)
$$

Chapter Summary

Markov chain is often used in the prediction or evaluation of queuing problems, and coding technology, bioinformatics, hydrological resources and other fields. The premise of its application is that the research object should be discrete events with Markov properties. When it is applied to the study management science, we should first understand the characteristics of the research object. The main points of this chapter include the basic concepts and application methods of Markov process, Markov chain, continuous-time Markov and hidden Markov model.

Key Concepts and Terms

Markov process
Markov chain
Stochastic process
State transition probability
Steady-state probability
Continuous-time Markov chain, CTMC

Table. 11.5 Air states in the past month

Date	1	2	3	4	5	6	7	8	9	10	11	12
Air condition	2	2	3	1	1	2	5	5	4	5	5	4
Date	13	14	15	16	17	18	19	20	21	22	23	24
Air condition	1	2	2	4	3	4	2	2	5	5	0	1
Date	25	26	27	28	29	30	31					
Air condition	1	3	2	4	4	1	1					

Hidden Markov model, HMM

Questions and Exercises

(1) Suppose there are six states of air quality: non-polluted, excellent, good, lightly polluted, heavy polluted, and severely polluted, which are represented by state variables $X_n = 0, 1, 2, 3, 4, 5$. The air conditions for the past month are shown in Table 11.5. Try to find the state transition probability.

(2) Suppose that an institution's investment income in a stock has three states, namely 1, 2, and 3. When the market is in state 1, the annual return is -4%. When the market is in state 2, the annual return is 30%. When the market is in state 3, the annual return is 10%. Suppose the state transition probability matrix P is applied to the weekly state transition of this stock:

$$P = \begin{pmatrix} 0.8 & 0.04 & 0.16 \\ 0.05 & 0.8 & 0.15 \\ 0.1 & 0.15 & 0.75 \end{pmatrix}$$

(1) Try to find the steady-state distribution of this stock in the market.
(2) Assume that 1 million yuan is invested in this stock for 6 years, and find the expected total profit.

3. Briefly describe the properties of no aftereffect and state transition probability of Markov chains.

Further Readings

1. Sun Qinghua & Sun Hao. *Content, Method and Technique of Stochastic Process* [M]. Wuhan: Huazhong University of Science and Technology Press, 2004.
2. William J. Stewart. *Probability Theory, Markov Chains, Queuing and Simulation* [M]. Beijing: World Publishing Corporation, 2013.
3. Cheng Weiqi, Huang Ximin *et al*. *Markov Chains: Models, Algorithms, and Applications* [M]. (Translated by Chen Xi). Beijing: Tsinghua University Press, 2015.