Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao, 11-14 July 2010

MARKOV PROPERTY OF MARKOV CHAINS AND ITS TEST

YU-FEN ZHANG[1], QUN-FENG ZHANG[1], RUI-HUA YU[2]

[1]Key Lab. for Machine Learning and Computational Intelligence, College of Mathematics and Computer, Hebei University, Baoding 071002, China
[2]Department of Information Technology, Baoding University, Baoding 071000, China
E-Mail：happy6002_cn@sina.com

Abstract:

Markov chains, with Markov property as its essence, are widely used in the fields such as information theory, automatic control, communication techniques, genetics, computer sciences, economic administration, education administration, and market forecasts. While using Markov chains to predict the future events, we must test the Markov property of random variable sequences of the past statistic data. Only when the random variable sequences satisfy the Markov property, can the prediction could be precise. This paper discusses the concept of Markov property and its features, studies its test method, and by example demonstrates the effectiveness of this prediction method.

Keywords:

Stochastic process; Random variable sequences; Markov property; χ^2 test

1. Introduction

Stochastic process, born in the early twentieth century, is an important branch of modern probability, and deals with the statistic rules in the random phenomena concerned with time parameter. Among a large number of stochastic processes, the Markov process is the most important and popular one with its wide application to modern physics, biology, popular business, geology, water sources science and atmospherics. As the simplest Markov process, the Markov chain has discrete time and status parameters. The Markov chains are widely used in the fields of automatic control, communication techniques, genetics, computer science, information theory, physics, ideology, economic administration, education administration, business managements, marketing prediction and so on. Markov chains have three characteristics: (1) dissertation of the process. The development of the system may be discredited into finite or countable states from time. (2) Randomness of the process. In the system, a state can transmit into another randomly with a probability of its history. (3) The Markov property of the process. The transition probability in the system relates only to the present states, not to the past states. That is to say, .when the present state of an object is known, its future state can be predicted. This is called the Markov property. This property makes us can avoid the difficulties in searching history data like in other prediction methods, so it is of importance in theory and application. Being of the Markov property is the premise of the sequences of random variables to be dealt with the Markov chain model. Many papers apply this property directly, but make little research about the essence and test of the Markov property. This paper makes a detained study about the Markov property and its test.

2. The concept and features of Markov property

Definition 1 (stochastic process)

Let (Ω, F, P) be a probability space, T a set (countable or not) of parameter in real line.

$\{\xi_t(\omega), t \in T\}$ is called a stochastic process if $\xi(\omega, t) = \xi_t(\omega)$ is a random variable for every $t \in T$.

Definition2 (Markov process)

Given a stochastic process $\{\xi(t), t \in T\}$. If the observation values $x_1, x_2, x_3, \cdots, x_m, x_{m+1}$ of $\xi(t)$ in time $t_1, t_2, t_3, \cdots, t_m, t_{m+1}$ ($t_1 < t_2 < t_3 < \cdots < t_m < t_{m+1} \in T$) satisfy the following condition:

$$P\{\xi(t_{m+1}) \in A | \xi(t_1) = x_1, \xi(t_2) = x_2, \cdots, \xi(t_m) = x_m\}$$
$$= P\{\xi(t_{m+1}) \in A | \xi(t_m) = x_m\} \quad \text{………（1）}$$

then this stochastic process is called to be of Markov property or a Markov process.

Definition3 (Markov chain)

Let $\{\xi(t),\ t = 0,1,2,\cdots\}$ be a stochastic process of discrete state with state space T and of

978-1-4244-6527-9/10/$26.00 ©2010 IEEE

non-negative integer parameters. If $\xi(t)$ satisfies:

$$P\{\xi(t+1)=j|\xi(0)=i_0,\xi(1)=i_1,\cdots\xi(t-1)=i_{t-1},\xi(t)=i\}$$
$$=P\{\xi(t+1)=j|\xi(t)=i\}\cdots\cdots\ (2)$$

then this stochastic process is called a Markov chain, and the condition (2) is called Markov property.

The condition in the definition can be paraphrased as follows. When $\xi(t)$ has the value i at time t, its future state is unrelated to the past states. In other words, the future of $\xi(t)$ relates to the past only via "present" state. Once the present state is decided, the future has no relation to the past states. This condition is called no-after-effectiveness or Markov property. So, the Markov property is the essence of the definition of the Markov chain.

From the definition of the Markov property, the following conclusion can be obtained.

Theorem1 Let $\{\xi(t), t=0,1,2,\cdots\}$ be a sequence of random variables, and E a countable real number set, $\xi(t)\in E$, $t=0,1,2,\cdots$.The following propositions are equivalent:.

(1) $\{\xi(t), t=0,1,2,\cdots\}$ is a Markov chain of countable states;

(2) For any $n\geq 1$ and any $i_0,i_1,i_2,\cdots i_{t-1},i\in E$, the formulae (2) hold;

(3) For any $n\geq 1$, any rigorously increased non-negative integer sequence $t_0,t_1,t_2,\cdots t_n$ and for any $i_0,i_1,i_2,\cdots i_n\in E$,

$$P\{X_{t_n}=i_n|X_{t_{n-1}}=i_{n-1},\cdots X_{t_1}=i_1,X_{t_0}=i_0\}=P\{X_{t_n}=i_n|X_{t_{n-1}}=i_{n-1}\}$$

(4) For any $n\geq 1$, any rigorously increased non-negative integer sequence $t_0,t_1,t_2,\cdots t_n$ and for any $i_0,i_1,i_2,\cdots i_n\in E$

$$P\{X_{t_0}=i_0,X_{t_1}=i_1,\cdots X_{t_n}=i_n\}$$
$$=\sum_{k\in E}P\{X_0=k\}P\{X_{t_0}=i_0|X_0=k\}P\{X_{t_1}=i_1|X_{t_0}=i_0\}\cdots P\{X_{t_n}=i_n|X_{t_{n-1}}=i_{n-1}\}$$

holds.

(5) For any $n\geq 1$, $m\geq 1$, any rigorously increased non-negative integer sequence $t_0,t_1,t_2,\cdots t_n,t_{n+1},\cdots t_{n+m}$ and $i_0,i_1,i_2,\cdots i_n,i_{n+1},\cdots,i_{n+m}\in E$,

$$P\{X_{t_0}=i_0,X_{t_1}=i_1,\cdots X_{t_{n-1}}=i_{n-1},X_{t_{n+1}}=i_{n+1}\cdots,X_{t_{n+m}}=i_{n+m}|X_{t_n}=i_n\}$$
$$=P\{X_{t_0}=i_0,\cdots,X_{t_{n-1}}=i_{n-1}|X_{t_n}=i_n\}\bullet P\{X_{t_{n+1}}=i_{n+1},\cdots,X_{t_{n+m}}=i_{n+m}|X_{t_n}=i_n\}$$

holds.

(6) For any $n\geq 1$, $m\geq 1$, any rigorously increased non-negative integer sequence $t_0,t_1,t_2,\cdots t_n,t_{n+1},\cdots t_{n+m}$ and $i_0,i_1,i_2,\cdots i_n,i_{n+1},\cdots,i_{n+m}\in E$

$$P\{X_{t_{n+1}}=i_{n+1}\cdots,X_{t_{n+m}}=i_{n+m}|X_{t_n}=i_n,\ X_{t_{n-1}}=i_{n-1},\cdots X_{t_0}=i_0\}$$
$$=P\{X_{t_{n+1}}=i_{n+1},\cdots,X_{t_{n+m}}=i_{n+m}|X_{t_n}=i_n\},$$

holds.

3. The statistical test of the Markov property

Generally, when predicting the trend of a phenomenon, we must take into account both its past and present states. For example, when we use the point estimation, the interval estimation, the variance analysis and the recurrence analysis to predict the development trend of some random variable, we must construct mathematical model based on its past and present data. However, when some stochastic process has the Markov property, we can predict its future state only by its present states. So, the Markov prediction method can avoid some difficulties in searching the past data and plays an important role in both theory and application. But before using this prediction method we must assure that the stochastic process has the Markov property.

3.1 The test of Markov property of the sequences of random variables

Having the Markov property is the premise for a sequence to apply the Markov chain model. Numerous papers use this prediction method directly without testing whether the stochastic process has the Markov property, which is contrary to the scientific spirit. In the following, we give the test theorem of the Markov property.

*Theorem 2***.** Suppose the index sequence has m possible states, denoting f_{ij} the frequency from the i state to j state through one step in the index sequence x_1, x_2:, …, x_n, and $i,j\in E$. If the quotient of the sum of the j column of transition frequency matrix divided by the total of all columns and rows of the matrix is called marginal probability, denoted by $p_{\cdot j}$,

i.e. $p_{\bullet j}=\sum_{i=1}^{m}f_{ij}\Big/\sum_{i=1}^{m}\sum_{j=1}^{m}f_{ij}$, then the statistic

$$\chi^2=2\sum_{i=1}^{m}\sum_{j=1}^{m}f_{ij}\left|\log\frac{p_{ij}}{p_{\bullet j}}\right|$$ subjects to the χ^2

distribution with free degree $(m-1)^2$ as its limitation distribution, where $p_{ij} = f_{ij} \Big/ \sum_{j=1}^{m} f_{ij}$. And, when $\chi^2 > \chi_\alpha{}^2((m-1)^2)$, here α denoting the significance level, then $\{x_i\}$ has the Markov property; otherwise, $\{x_i\}$ has not the Markov property, and can not be viewed as the Markov chain.

3.2 Example for the test of Markov property

Whether the water resource of a region is abundant is decided by its rainfall. To make a prediction about the coming water resources, we must first predict the rainfall of the region. However, owing to the complexity, variance and versatility of water conditions, the process of raining consists of a huge imprecision and vagueness and we can not make a precise prediction about the rainfall of the coming period (a year, a season or a month). In practice, a prediction interval is enough for many purposes. Therefore, the liability of prediction increases as the range of prediction expands (i.e. from point prediction to interval prediction).

The next is the annual rainfall data of the past 47 years (from 1962 — 2008) , in a region of Hebei province. We take it for example to make an analysis. First, we order the rainfall sequence increasingly. Then, we use the ordering clustering method to sort the annual rainfall to 5 grades, i.e. 5 states. Based on our experience, we call the year of annual rainfall less than 280mm a drought year, the year of 280~380mm rainfall a dry year, the year of 380~450mm rainfall a normal year, the year of 450~550mm rainfall a slightly flood year, and the year of more than 550mm a flood year. The above grades of rainfall are called states, described in table 1.

In the following, we make a Markov property test on the sequence of rainfall of 47 years. From the data in table 1, we derive the following matrices.

$$(f_{ij})_{5\times5} = \begin{pmatrix} 0 & 3 & 3 & 1 & 1 \\ 0 & 3 & 4 & 3 & 2 \\ 2 & 4 & 1 & 2 & 2 \\ 2 & 2 & 0 & 2 & 2 \\ 3 & 1 & 3 & 0 & 0 \end{pmatrix}$$

$$(p_{ij})_{5\times5} = \begin{pmatrix} 0 & 3/8 & 3/8 & 1/8 & 1/8 \\ 0 & 3/12 & 4/12 & 3/12 & 2/12 \\ 2/11 & 4/11 & 1/11 & 2/11 & 2/11 \\ 2/8 & 2/8 & 0 & 2/8 & 2/8 \\ 3/7 & 7 & 3/7 & 0 & 0 \end{pmatrix}$$

The marginal probability

$p_{\bullet1} = 7/46$ $p_{\bullet2} = 13/46$ $p_{\bullet3} = 11/46$

$p_{\bullet4} = 8/46$ $p_{\bullet5} = 7/46$

The values of statistic

$\chi^2 = 2\sum_{i=1}^{5}\sum_{j=1}^{5} f_{ij} \left|\log \frac{p_{ij}}{p_{\bullet j}}\right|$ are shown in table 2.

Given the significance level $\alpha = 0.05$, we can obtain the upper tail point $\chi_\alpha^2\left((m-1)^2\right) = \chi_\alpha^2(16) = 26.296$. Since $\chi^2 = 33.8216 > \chi_\alpha^2\left((m-1)^2\right)$, we know the annual rainfall has the Markov property from theorem 2.

4. Conclusions

Markov chains have the Markov property, which means that the future states are dependent on the present state but independent of the past states. Therefore we can make predictions about the future by use of the Markov chains model. We must bear in mind that the Markov property is the premise. Before we make predictions using the Markov chains model, we must assure that the stochastic process has the Markov property. Only by this can we assure the precision of the prediction. That is to say, the test of the Markov property is necessary in the application of the Markov chains model. Then we must decide the intervals of indices in the sequences of random variables using classification method. Especially, we should point out that the results of this prediction method are a range instead of a concrete value. For some problems, this kind of results is more reliable and practicable. In a word, this paper researches the concept of Markov property and its features and test methods, and makes a detailed analysis on the test method by means of examples.

References

[1] Luo ji-yu, Xing ying. Statistical analysis methods in economics and prediction. Beijing, Tsinghua University Press.1987.

[2] Wang Zi-kun, Yang Xin-gu. Birth-and-death process

and Markov chains (2ed edition).Beijing,. Science press, 2005.
[3] Lei Qin-li. Multi-statistical analysis in economic administration. Beijing, China statistical press.2002.
[4] Xia le-tian. Markov chain prediction and its application in hydrology. PhD thesis in Hehai University.2005.
[5] A.T. Bharucha-Reid. The theory of Markov chains and its application (Chinese). Shanghai, Shanghai science and technology press, 1979.
[6] Qian min-ping, Gong guang-lu. The application stochastic process. Beijing. Beijing University press. 1998.
[7] Han D.Hu X J. On the relative balanced growth of price of the closed dynamic model with variable coefficients. Lecture Notes of OR，1998(3):412-422.
[8] Han .Analysis of the Markova chain on the stoke Prieenad stokes Preparation. Proceedings of ICOTA. Singoreorld Silicified，1995，810-814.
[9] Chen M.F. From Markova chains to non-equilibrium particle systems. World Scientific Singapore，1992.
[10] S.M. Ross. Stochastic Process [M].New York: John Wiley&Sons, Inc. 1991
[11] ZEKA I SEN. Critical Drought Analysis by Second orders Markov Chain. Journal of Hydrology，1990，120(1-4):183-202.
[12] Financial Accounting Standards Board (1998).Statement of inancial Accounting Standards No.133.

TABLE 1. THE PAST 47 YEAR (1962---2008) RAINFALL DATA OF A REGION OF HEBEI PROVINCE

year	**1962**	**1963**	**1964**	**1965**	**1966**	**1967**	**1968**	**1969**	**1970**	**1971**	**1972**	**1973**
rainfall（mm）	261.9	486.4	631.5	299.0	568.0	398.2	479.6	697.6	397.7	640.4	247.1	387.7
states	1	4	5	2	5	3	4	5	3	5	1	3
year	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985
rainfall（mm）	694.2	211.4	322.6	656.6	325.3	603.8	624.8	383.3	238.8	423.0	237.1	330.7
states	5	1	2	5	2	5	3	3	1	3	1	2
year	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
rainfall（mm）	445.9	518.9	492.6	490.3	257.0	400.6	347.5	363.8	411.5	356.2	381.2	317.0
states	3	4	4	4	1	3	2	3	3	2	3	2
year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	
rainfall（mm）	473.0	373.7	369.0	348.3	469.4	228.1	338.8	546.1	358.9	412.0	372.3	
states	4	3	3	2	4	1	2	4	2	3	3	

TABLE 2. THE VALUES OF STATISTIC $\chi^2 = 2\sum_{i=1}^{5}\sum_{j=1}^{5} f_{ij}\left|\log\frac{p_{ij}}{p_{\bullet j}}\right|$

states	$f_{i1}\left\|\log\frac{p_{i1}}{p_{\bullet 1}}\right\|$	$f_{i2}\left\|\log\frac{p_{i2}}{p_{\bullet 2}}\right\|$	$f_{i3}\left\|\log\frac{p_{i3}}{p_{\bullet 3}}\right\|$	$f_{i4}\left\|\log\frac{p_{i4}}{p_{\bullet 4}}\right\|$	$f_{i5}\left\|\log\frac{p_{i5}}{p_{\bullet 5}}\right\|$	total
1	0	0.8486	1.3498	0.3302	0.1967	2.7253
2	0	0.3678	1.3285	1.0887	0.1819	2.9669
3	0.356	1.0084	0.9671	0.0889	0.356	2.7764
4	0.9929	0.2452	0	0.7258	0.9929	2.9568
5	3.1063	0.6288	1.7503	0	0	5.4854
Total	4.4552	3.0988	5.3957	2.2336	1.7275	33.8216