Application of Discrete-Time Markov Models

Narela Bajram[a], Daria Ler[b]

[a] Faculty of Business and Administration, International University of Sarajevo,
Hrasnicka Cesta 15, Ilidza, 71200 Sarajevo BiH, narela7@yahoo.com.

[b] Faculty of Engineering and Natural Science, International University of Sarajevo,
Hrasnicka Cesta 15, Ilidza, 71200 Sarajevo BiH, dariamail@gmail.com

Abstract— This paper introduced a general class of mathematical models, Markov chain models, which are appropriate for modeling of phenomena in the physical life, medicine, engineering and social sciences. Application of Markov chains are quite common and have become a standard tool of decision making. What matters in predicting the future of the system is its present state, and not the path by which the system got to its present state. Two methods are presented that exemplify the flexibility of this approach: the regular Markov chain and absorbing Markov chain. The long-term trend in absorbing Markov chains depends on the initial state. In addition, changing the initial state can change the final result. This property distinguishes absorbing Markov chains from regular Markov chains, where the final result is independent of the initial state. The problems are formulated by using the Wolfram Mathematical Programming System.

Keywords— regular and absorbing Markov chains, prediction the future and long-term trend.

1. INTRODUCTION

Many decision makers use a certain type of random process to make decisions. The main feature of this type of process is that it's "memoryless" of the past. Such a process is called a Markov Process, first proposed by the Russian mathematician Andrei Andreevich Markov in 1906 [1].

The work on the central limit theorem and the law of large numbers [2] by Chebyshev, Lyapunov, and Markov created the basis for the modernization of probability theory.

A stochastic process is a mathematical model that evolves over time in a probabilistic manner.

Since the early studies, there have been many applications of Markov processes to the modeling of phenomena in the physical life, medicine, engineering and social sciences. Paul and Tatyana Ehrenfast investigated a Markov chain model for diffusion in a 1907 paper about the same time that Einstein and Smoluchowski were using Markov processes to study Brownian motion.

On other hand, some of the areas of investigation in the social sciences that have been pursued through the use of Markov chains include voting behavior, geographical mobility within a country, growth and decline of town, the size of economic firms, prediction of enrollments in colleges and universities and etc.

In this paper we study the special kind of stochastic process, called a Markov chain.

According to Hogben L. (1987), a "Markov chain" is a random process described by a physical system that at any given time (t = 1, 2, 3...) occupies one of a finite number of states. At each time t the system moves from state j to state i with probability pij that does not depend on t. The numbers p_{ij} are called "transition probabilities." [3]. A notable feature of Markov chains is that they are historyless – the next state of the system depends only on the current state, not any prior states [4]. Andrei Nikolaevich Kolmogorov in his seminal work [5] remarked that "Historically, the independence of experiments and random variables represents the very mathematical concept that has given probability its peculiar stamp".

II REGULAR MARKOV CHAIN

First of all it was reviewed some of the basic properties of regular Markov chains.

DEFINITION: A Markov process is a *regular chain* if some power of the transition matrix has only positive entries.

In particular, the Markov process is regular if all entries in the transition matrix $P = P1$ are positive. A Markov process is a regular one if there is some positive integer n, so that the process may be in anyone of the possible states n steps after starting, regardless of the initial state. The smallest n for which this is possible is the smallest positive integer n for which Pn has no zero entries.

If P is the transition matrix of a regular Markov chain, then it turns out that the powers of P approach a matrix V, all of whose rows are the same. If $\mathbf{v}$ denotes the row vector formed from any of the rows of V, then it also happens that $\mathbf{v}P = \mathbf{V}$.

DEFINITION: A vector v is a *fixed point vector* of the matrix P if P = V. A Markov chain is said to be in *equilibrium* if the probability distribution at some step is given by a fixed point vector of the transition matrix.

As an application model it was considered an example from the biomedical research studying the risk of heart attack in mean [6]. By studying the male ancestors, sons and grandsons of these men, the researcher comes up with the following transition matrix.

	Thin	Normal	Overweight
Thin	0.3	0.5	0.2
Normal	0.2	0.6	0.2
Overweight	0.1	0.5	0.4

The transition matrix P shows the probability of change in weight from one generation to the next. For example, if a parent is in state 3 (the overweight), what is the probability that a grandchild will be in state 2. To find out, we will start with a tree diagram, as shown in Fig. 1. The various probabilities come from transition matrix P. The arrows point to the outcomes "grandchild in state 2". The probability that a parent in state 3 will have a grandchild in state 2 is given by sum of the probabilities indicated with arrows, or expressed as **0.05+0.3+0.2=0.55**
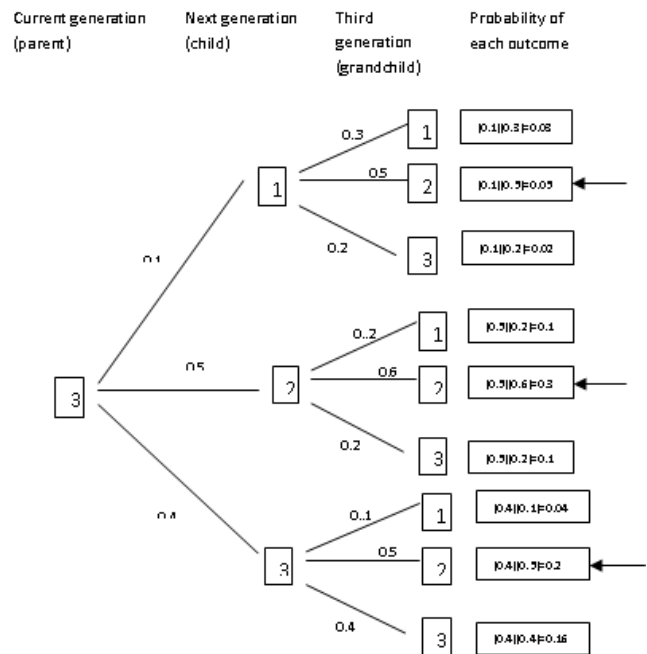

Figure 1: Decision tree

We used p_{ij} to represent the probability of changing from state i to state j in one generation. This notation can be used to write

the probability that a parent in state 3 will have a grandchild in state 2:

$$p_{31}*p_{12}+p_{32}*p_{22}+p_{33}*p_{32}$$

This sum of products of probabilities is nothing more than one step in the process of multiplying matrix P by itself. Thus,

MatrixForm[P2]

$$P^2 = \begin{pmatrix} 0.21 & 0.55 & 0.24 \\ 0.2 & 0.56 & 0.24 \\ 0.17 & 0.55 & 0.28 \end{pmatrix}$$

Note: The entry in row 3, column 2 in P^2 gives the probability that a person in state 3 will have a grandchild in state 2;that is, that an overweighed person will have normal grandchild. This number, 0.55, is the result found through using the tree diagram.

In the same way the matrix P^2 gives the probability of changes after two generations, the matrix $P^3=P*P^2$ gives the probability of changes after three generations.
For matrix P,

MatrixForm[P3]

$$P^3 = \begin{pmatrix} 0.197 & 0.555 & 0.248 \\ 0.196 & 0.556 & 0.248 \\ 0.189 & 0.555 & 0.256 \end{pmatrix}$$

Matrix P^3 gives a probability of 0.555 that a person in state 3 will have great-grandchild in state 2. Also the probability that the person in state 3 will have great-grandchild in state 3 is 0.256. In order to develop a long-range prediction for the proportion of the population in each weight group, we should assign the initial distribution of people which can be written as probability vectors. If we suppose that the initial probability vector is Xo= [0.2   0.5   0.25], the distribution after one generations is

$$Xo*P= [0.195 \quad 0.555 \quad 0.25]$$

Using this information, we can compute the distribution of weight group for two and more generations as illustrated in Table 1.

Table 1. Distribution of weight group, initial probability vector [0.2 0.5 0.25]

After Generation n	Thin	Normal	Overweight
0	0.2	0.55	0.25
1	0.195	0.555	0.25
2	0.195	0.555	0.25

But what will happened if the initial probability vector is different from [0.2   0.5   0.25] ? If we suppose that [0.65   0.25 0.1] is used, then the results are given in Table 2.

Table 2. Distribution of weight group, initial probability vector [0.65 0.25 0.1]

After Generation n	Thin	Normal	Overweight
0	0.65	0.25	0.1
1	0.255	0.525	0.22
2	0.2035	0.525	0.244
3	0.195	0.555	0.25
4	0.195	0.555	0.25

The results again are approaching the numbers in the probability vector [0.19 0.55 0.25], the same numbers approached with the initial probability vector [0.2   0.55   0.25]. In either case, the long-range trend is for about 55% of the people to be classified as normal weighted. This example suggests that this long-range trend does not depend on the initial distribution.

III ABSORBING MARKOV CHAIN
Not all Markov Chains are regular. In fact, some of the most important life science applications of Markov chains do not involve transition matrix that are regular. One type is known as an absorbing Markov chain. A Markov chain is an *absorbing chain* if it has at least one absorbing state and from every state it is possible to reach some absorbing state in a finite number of steps. If a state is not an absorbing state, it is called a *transient state*. We will call a state Si absorbin if $pii = 1$, $pij = 0$ for $i \neq j$. Once in Si you can never leave. The transition matrix P can be written as

$$P = \begin{bmatrix} I_1 & O \\ R & Q \end{bmatrix}$$

Where I_1 represents the 1 x 1 identity matrix, O represents the matrix of zeros, R represent the matrix in the lower left and Q represent the matrix in the lower right. The fundamental matrix for an absorbing Markov chain is defined as matrix F, where

$$F = (In-Q)^{-1}$$

Here In is the n x n identity matrix corresponding in size to matrix Q, so that the difference In-Q exists.

As an application mode it was considered an example from the study by Beck, Robert and Stephen in medical prognoses [7]. The study estimates a patient's prognosis for improving under various treatment strategies whereby the researchers come up with the following transition matrix.

	well (1)	ill (2)	dead(3)
well(1)	0.3	0.5	0.2
ill(2)	0	0.5	0.5
dead(3)	0	0	1

The transition matrix shows that the probability of going from state 1 to state 3 is 0.2, and that the probability of staying in state 3, is 1. Thus, once state 3 is entered, it is impossible to leave. For this reason, state 3 is called and absorbing state. For the long-term trend we should find various powers of the transition matrix. Using Wolfram Mathematical Programming System we generate the following results:

```
P3 = N[MatrixPower[P, 3]] // MatrixForm
```

$$\begin{pmatrix} 0.027 & 0.245 & 0.728 \\ 0. & 0.125 & 0.875 \\ 0. & 0. & 1. \end{pmatrix}$$

```
P10 = N[MatrixPower[P, 10]] // MatrixForm
```

$$\begin{pmatrix} 5.9049 \times 10^{-6} & 0.00242664 & 0.997567 \\ 0. & 0.000976563 & 0.999023 \\ 0. & 0. & 1. \end{pmatrix}$$

It can be noticed that when *P* is raised to higher and higher powers, the system will tend toward the absorbing state, so that the probability is 1 that the patient will eventually die. In addition, the long-term trend in absorbing Markov chains depends on the initial state. Changing the initial state can change the final result. This property distinguishes absorbing Markov chains from regular Markov chains, where the final result is independent of the initial state. In order to find the final probabilities of entering an absorbing state without finding all the powers of the transition matrix, it is necessary only to work with the non-absorbing states. As we already noted above the fundamental matrix is defined as matrix F, where

$$F = (In\text{-}Q)^{-1}$$

and In is the identity matrix corresponding in size to matrix Q.

```
ID = IdentityMatrix[2]
```

{{1, 0}, {0, 1}}

```
Q = {{0.3, 0.5}, {0, 0.5}}
```

{{0.3, 0.5}, {0, 0.5}}

```
F = Inverse[ID - Q]
```

{{1.42857, 1.42857}, {0., 2.}}

```
MatrixForm[F]
```

$$\begin{pmatrix} 1.42857 & 1.42857 \\ 0. & 2. \end{pmatrix}$$

The fundamental matrix (F) gives the expected number of visits to each state before absorption occurs. For example, if the patient is currently well, the first row just computed says that it is expected number of cycles that a well patient will continue to be well before dying is 1.428 and the expected number of cycles that a well patient will be ill before dying is also 1.428. Finally, to verify our answer that the probability is 1 that if the patient was originally well will ended up dying we use the product matrix FR. The product FR gives the matrix of the probability that if the system was originally in a particular non-absorbing state; it ended up in the absorbing state.

4. CONCLUSION

Markov chains are a relatively simple but very interesting and useful class of random processes. It is usual to think of Markov Chains as describing the trajectories of dynamic objects. The changes are not completely predictable, but rather are governed by probability distributions. These probability distributions incorporate a simple sort of dependence structure, where the conditional distribution of the next state of the system depends only on the present state not on preceding states.

That is, what matters in predicting the future of the system is its present state, and not the path by which the system got to its present state.

Application of Markov chains are quite common and have become a standard tool of decision making. Two methods are presented that exemplify the flexibility of its application: the regular Markov chain and absorbing Markov chain.

The long-term trend in absorbing Markov chains depends on the initial state. In addition changing the initial state can change the final result. This property distinguishes absorbing Markov chains from regular Markov chains, where the final result is independent of the initial state.

5. SELECTED EXTENSIONS

This classical subject is still very much alive, with important developments in recent decades in both theory and applications. As interesting extensions I briefly described two of them which are of high importance in modeling complex system.

1.So-called adaptive Markov chains. These are systems in which the transition matrix is adjusted depending upon the entire history of the system or some statistical summary of that history.

2. Non-linear Markov Chain in which the distribution of X_n depends upon both X_{n-1} and its distribution , η_{n-1}. This is the evaluation of a Feynman-Kac system. Moreover, an excellent monograph on Feyman-Kac formulae and their mean field approximations has recently been written [8].

6. REFERENCES

[1] A.A. Markov, (1908) ,Rasprostranenie predel'nyh teorem ischisleniya veroyatnostej na summu velichin svyazannyh v cep', Zapiski Akademii Nauk po Fiziko-matematicheskomu otdeleniyu, VIII seriya 25 (3). Translated into German, Ausdehnung der Satze uber die Grenzwerte in der Wahrscheinlichkeitsrechnung auf eine Summe verketteter Grossen, in: A.A. Markoff (Ed.), Wahrscheinlichkeitsrechnung (translated by H. Liebmann), B.G. Teuber, Leipzig, 1912, pp. 272–298. Translated into English, Extension of the limit theorems of probability theory to a sum of variables connected in a chain (translated by S. Petelin) in: R.A. Howard (Ed.), Dynamic Probabilities Systems, vol. 1, Wiley, New York, 1971, pp. 552–576.

[2] P.L. Chebyshev (1846) "An experience in the elementary analysis of the probability theory", Crelle's J. Angew. Math.

[3] Hogben, Leslie. *Elementary Linear Algebra*. West Publishing Company, St. Paul, MN, 1987: 81-92 .

[4] Hefferon, Jim. (2003). "Linear Algebra", chapter 3. Retrieved 4/20/05 from
http://joshua.smcvt.edu/linearalgebra.

[5] Kolmogorov A.N. (1956), *"Foundations of the Theory of Probability"*, Chelsea Publishing Company, second English edition.

[6] Lay,David C.(2003), "Linear Algebra and its Applications" 3rd ed. Addison Wesley, Pearson Education, Inc.

[7] Beck, J.Robert, and Stephen G. Paukeer, (1983),"The Markov Process in Medical Prognosis,"Medical Decision Making, Vol.4, No3, pp. 419-458.

[8] P.Del Mora (2004), " Feynman-Kac formulae: genealogical and interacting particle systems with applications", Probability and Its Application. Springer Verlag, New York.