Application of Markov Chains

Qianfan Luo
International Curriculum Center, Shenzhen Foreign Languages School, Guangdong Province, 518000, China
Corresponding author’s e-mail: angela@cas-harbour.org

Abstract. This paper introduces the theory of Markov Chains and illustrates the deduction of its transition probability matrix. A mathematical model for the growth of a bacteria colony and a mathematical model of the artificial intelligence in the venous transfusion are built. They are based on Markov chains and accord with the properties of transition matrix. The limitation of the model is discussed. The material is cited from authoritative websites, research done by academic institutions, and scientific essays.

1. Introduction
Markov chains are widely used in mathematics, especially in the field of statistics and information theory. They are widely used in economics, game theory, communication, genetics and finance. [4] When it comes to real-world problems, they are used to hypothesize solutions to study cruise-control systems in motor vehicles, queues or queues of customers arriving at airports, exchange rates and so on [2]. However, there are still some technical problems. For example, venous transfusion is a common treatment in people’s daily lives. However, irregular use of venous transfusion can cause some dangers. Therefore, it's necessary to design a cheap and efficient intelligent intravenous infusion management system to prevent medical accidents caused by improper venous transfusion. By using the transition matrix theory, this paper draws a picture of the application of Markov Chain in this field and gives a proof of the nth matrix equation.

2. Analysis
2.1. Application of Markov chains in the prediction of the growth of bacteria
2.1.1. Assumption
Firstly, only one bacteria cell can duplicate in a time interval. The remaining bacteria are assumed not changed during that time. Secondly, during the duplication process, the probability for the bacteria to replicate or die is 50%. Thirdly, the population of the bacteria is limited to N and whether the bacteria can survive is 50%. Fourthly, the bacteria can only decay for limited generations. Under these circumstances, assume that the total number of the bacteria is N=4, and the probability of the bacteria to decay is P (decay) and the probability of the bacteria to die is P (die), suppose that the initial state of the bacteria is i and the final state of the bacteria is j. We find that the bacteria decay from state i to state j in a particular k generations, where k is an integral. Assume that the maximal capacity of the bacteria is 4, ie k=1, 2, 3, or 4, and the problem is how to find the possibility of passing from a specific state to another state within particular generations.
2.1.2. Formalism
To display the information in a more compact way, we can put the possible states into a matrix. Because there is only one possibility in the initial state, \( P(00)=1 \). Similarly, we can depict a diagram of the bacterial decay where the numbers represent the total number of bacteria and the rows refer to the generation of bacteria.

\[
\begin{array}{cccccc}
0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 \\
2 & 2 & 2 & 2 & 2 \\
3 & 3 & 3 & 3 & 3 \\
4 & 4 & 4 & 4 & 4 \\
\end{array}
\]

Each number stands for the state of the bacteria, while the arrows represent the transition of the state. Therefore, we can get a sequence of probabilities:

\[
\{ P_{00}, P_{01}, P_{02}, \ldots, P_{44} \}
\]

Because when a bacterium decays, it only has one possibility to decay into two bacteria, so the matrix which is given from state \( i \) to state \( j \) by one step can be shown as:

\[
\begin{pmatrix}
P_{00} & P_{01} & P_{02} & P_{03} & P_{04} \\
P_{10} & P_{11} & P_{12} & P_{13} & P_{14} \\
P_{20} & P_{21} & P_{22} & P_{23} & P_{24} \\
P_{30} & P_{31} & P_{32} & P_{33} & P_{34} \\
P_{40} & P_{41} & P_{42} & P_{43} & P_{44} \\
\end{pmatrix}
\]

where the number serve as the probability of bacteria in each state.

2.1.3. Solutions
According to the second assumption, we can set up a model where \( P(\text{decay}) = P(\text{die}) = 1/2 \). Based on the assumption above, we can conclude that when bacteria decay, it has two options, decay or die, and the probability of each option is equal to 50%. In our assumption, the maximal generation is equal to 4.

We have already known that from state \( i \) to state \( j \), the probability of each state is given in the matrix above. Therefore, according to the Multiplication Principle, we can infer that in order to calculate the probability of bacteria decay in each state in two steps, we need to calculate the square of the matrix, that is, multiply the matrix by itself.

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

According to the Matrix Multiplication Rule we have discussed in the third section, we can get the results of the probability of each state where bacteria decay:
Similarly, when \( k=3 \), which means that we have to go three steps to get the result from state \( i \) to state \( j \), the result is the cube of the original matrix, or we can calculate the product of the square sum of the matrices:

\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{4} & 0 \\
\frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\
0 & \frac{1}{4} & 0 & 0 \\
\end{pmatrix}
\]

In the same way, we can calculate the result when \( k \) is equal to 4, that is, we have to go four steps to get from state \( i \) to state \( j \) by identifying the biquadratic of the matrix. By multiplying the cube of the matrix, which we have already calculated above, with the original matrix, we can get the following results:

\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{8} \\
\frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\
\frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

We can use the same method to calculate the rest of the results.

Thus, we can figure out the general term formula when there are \( n \) steps. To make the calculation easier, we can find the relationship between two numbers in the same position in the two matrices and conclude that when \( k=n \), the result will be

\[P^{(n)} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
\frac{3}{4} + \alpha_n & \beta_n & \gamma_n & \epsilon_n \\
\frac{1}{2} + \alpha_{n-1} & \beta_{n-1} & \gamma_{n-1} & \epsilon_{n-1} \\
\frac{3}{4} + \alpha_n & \beta_n & \gamma_n & \epsilon_{n-1} \\
0 & 0 & 0 & 1 \\
\end{pmatrix}\]

Define that:

\[\alpha_n = \left(\frac{\sqrt{2}}{2}\right)^n\left[-\frac{1}{4} - \frac{\sqrt{2}}{8} + (-1)^n\left(-\frac{1}{4} + \frac{\sqrt{2}}{8}\right)\right]\]

\[\alpha_{n-1} = \left(\frac{\sqrt{2}}{2}\right)^{n-1}\left[-\frac{1}{4} - \frac{\sqrt{2}}{8} + (-1)^{n-1}\left(-\frac{1}{4} + \frac{\sqrt{2}}{8}\right)\right]\]

\[\beta_n = \frac{1}{4}\left(-\frac{\sqrt{2}}{2}\right)^n + \left(\frac{\sqrt{2}}{2}\right)^n\]

\[\beta_{n-1} = \frac{1}{4}\left(-\frac{\sqrt{2}}{2}\right)^{n-1} + \left(\frac{\sqrt{2}}{2}\right)^{n-1}\]

\[\gamma_n = \frac{\sqrt{2}}{4}\left(\frac{\sqrt{2}}{2}\right)^n - \left(\frac{\sqrt{2}}{2}\right)^n\]

\[\gamma_{n-1} = \frac{\sqrt{2}}{4}\left(\frac{\sqrt{2}}{2}\right)^{n-1} - \left(\frac{\sqrt{2}}{2}\right)^{n-1}\]

\[\epsilon_n = \frac{1}{4}\left(-\frac{\sqrt{2}}{2}\right)^n + \left(\frac{\sqrt{2}}{2}\right)^n\]

\[\epsilon_{n-1} = \frac{1}{4}\left(-\frac{\sqrt{2}}{2}\right)^{n-1} + \left(\frac{\sqrt{2}}{2}\right)^{n-1}\]

What’s more, to explain the concepts in a more specific way in the result matrix, \( P(12)=\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8} \) due to the Matrix Multiplication Rule.

We assume that each decay is only influenced by the decay before it, which can explain the theory
of Markov Chains. As a result, when there are two steps for the bacteria to decay into another state, we have to identify the probability of each step, and according to the Multiplication Principle, the two possibilities should be multiplied. Similarly, in the case of n steps, we find that the result of the probability will be the product of P and the power of n. So, when we go n steps to get from state i to state j, the probability is $P^n$.

2.2. Application of Markov chains in artificial intelligence using in the venous transfusion

2.2.1. Situation
We use a trial module to simulate the real situation of an intravenous injection (programmed with Arduino electronic blocks) that is supposed to be connected to an operator controlled by a nurse. The nurse need to operate the computer to adjust Arduino electronic blocks that are connected with Bluetooth, and to regulate the speed of the intravenous drip bottle. We also designed a smart bracelet to monitor the heart rate of the patient. However, we cannot avoid accidents. For example, the sudden breakdown of the speed controller, network signal problems, and information transmission errors.

2.2.2. Formalism
Under this condition, we can only consider the error probability of the speed controlling machine, the smart bracelet as well as the internet connection. Assume that $P(x)$ is the probability of getting wrong information of each state, $P(m)$ is the probability of the transfusion speed controlling machine making mistakes, $P(b)$ is the error probability of smart bracelet, and $k$ is the number of times of information transmitting. Consequently, for each time the information transmitted, the previous non-mistaken message will not influence whether the information will be transmitted successfully at this time, that is,

$$P(X_{t+1},\ldots,X_{t-k},X_t|X_t) = P(X_{t+1}|X_t)$$

At the same time, any event is independent of the other events, so the only thing that can affect an event is the probability of the previous event. Consequently, we can apply Markov Chains to this system. The process of transmitting information starts from the smart bracelet detecting the patient’s heart rate. When it finds something unusual, it will send a message to the computer which is monitored by a nurse. Then the nurse can manipulate the computer to control the speed of transfusion. When the Arduino electronic block receives the message, it will adjust the infusion speed artificially. Therefore, information needs to be transmitted four times by the machines in the system. In our experiment, we found out that the Arduino electronic block failed every twenty times we test it, that is, the error probability is roughly equal to 5%. Approximately, we can ignore the errors when we render the error probability from 0 to 10%.

According to the test on the smart bracelet, the probability of the system’s stability is 93.27%, that is, 6.73% of the data detected by the smart bracelet is inaccurate. The computer, which is operated by the nurse, only has 10.6% probability to get wrong information or miss the signal from the Bluetooth. As above, we consider the data of which error being less than 5% is a normal experimental data.

2.2.3. Solutions
Above all these assumptions, each information transfer choice has a Markov chain when $X_i$ is the number of the exchange in the ith state. Assume that at first there is only one piece of information, which means the initial state is 1, then the possible routes through different states are as follows:
The information can only be transferred four times, so \( k \) can only be equal to 0, 1, 2, 3, or 4. Under the conditions described above, we first need to determine the expected value of the error probability of the whole system to better clarify the expected values of our experiment.

\[
EV = \frac{1}{4} \cdot 5\% + \frac{1}{4} \cdot 6.73\% + \frac{1}{2} \cdot 10.6\% = 8.2325\%
\]

If we want to figure out the probability of state \( i \) to state \( j \), we can first draw a matrix which might be related to the bacterial culture problem:

\[
\begin{pmatrix}
P_{00} & P_{01} & P_{02} & P_{03} & P_{04} \\
P_{10} & P_{11} & P_{12} & P_{13} & P_{14} \\
P_{20} & P_{21} & P_{22} & P_{23} & P_{24} \\
P_{30} & P_{31} & P_{32} & P_{33} & P_{34} \\
P_{40} & P_{41} & P_{42} & P_{43} & P_{44}
\end{pmatrix}
\]

which is equivalent to:

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0.082325 & 0 & 0.082325 & 0 & 0 \\
0 & 0.082325 & 0 & 0.082325 & 0 \\
0 & 0 & 0.082325 & 0 & 0.082325 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\]

Naturally, if a matrix satisfies the two properties above and \( k \) is given as a constant, then it is a transition matrix. Because if we have such a matrix and know every element of it, we can construct Markov chains and let

\[
P^{(k)}_{ij} = a_{ij}^{(k)}
\]

Thus, this matrix has to be a transition matrix.

Depending on the bacterial culture instance, we can easily find the solution to the venous transfusion problem. When \( k=2 \), the result will be the previous matrix multiplied by itself because of the Multiplication Principle, so we can get the result:

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0.082325 & 0.006777 & 0 & 0.006777 & 0 \\
0.006777 & 0 & 0.013555 & 0 & 0.006777 \\
0 & 0.006777 & 0 & 0.006777 & 0.082325 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\]

Similarly, we are able to calculate the error probability of state \( i \) to state \( j \), where \( k \) equals to 3, or 4. Therefore, the theory \( P^{(n)} = P^{(n)} \) can be proved. [3]

**Proof.** We need to prove that each element in the two matrices is equal. Assume that \( a_{ij}^{(n)} \) is the element of the \( i \)th row, \( j \)th column in \( P^{(n)} \), and \( b_{ij}^{(n)} \) is the element of the \( i \)th row, \( j \)th column in \( P^{(n)} \). Now we assume that \( P^{(n-1)} = P^{(n-1)} \) is true.

Obviously,
To calculate $a_{ij}^{(n)}$, firstly, it has to travel $(n-1)$ steps. In fact, we already knew the $(n-1)$-step transition matrix because we have assumed that $P^{n-1} = P^{(n-1)}$. According to the Markov property, we can assume that it will first travel to the $k$ state in the $(n-1)$th step. For different $k$, these events are disjoint. So,

$$\sum_{k=1}^{M} a_{ik} a_{kj} = a_{ij}^{(2)}$$

Therefore, the two matrices are equal.

3. Comment and Evaluation

3.1. Limitations of Bacteria Decay
Firstly, there are about “1,000,000 bacteria per milliliter (1/5 of a teaspoon) of water in the coastal ocean” according to the Cryomics Lab [3]. Hence, the number of bacteria cannot be four. Secondly, bacteria can double in a short period of time. Therefore, bacteria do not decay at a constant speed. Thirdly, bacteria need specific environment to reproduce. Thus, the probability of the bacteria decay or die cannot always be equal to 1/2.

3.2. Limitations of factors
In our assumption, there are only four factors that influence the efficiency of messages transmission; however, it is obviously not true in real life. Factors that may affect the final result include external signal interference, signal connection, human interference and others.

3.3. Suggestions and Improvements
In order to make the experiment more precise, we need to expand our sample space. Additionally, we can consider some uncertainties in our data or set some random variables in the number, such as setting the probability of the bacteria decay from 40% to 60%, so that we are able to reduce the error.

For the venous transfusion part, we can simulate the real world by adding numerical constraints, but human behavior is the hardest to predict, so we can’t make accurate predictions, only rough data values and a reasonable uncertainty range. As a result, we can only set the random variable of the machine or human performance possibility to the veracity of our data.

4. Conclusion
The conclusion leads to the basic application of Markov chain in daily life, and a derivation to its transition probability is maken by providing the detailed calculation process. However, there are still several problems that this paper does not solve. Firstly, there is no general way to calculate all the $P^{(n)}$ if we rarely know that $P^{(n)}$ is a transition matrix. The main reason is that not every transition matrix can be diagonalized. However, if we know that the state $i$ can only move to state $(i+1)$ or state $(i-1)$ in one step and the number of states is limited, we can solve this problem as we have done before. Besides, if the probability for a bacterium to duplicate itself successfully is not 1/2 but $p$, the transition matrix can still be diagonalized. In this case, we can divide the whole fluctuation process into several short periods, during which we consider the transition probability as a constant value. Through the step-by-step calculation process, we can eventually apply the Markov chain to solve the problem.
Future papers would focus on the unsolved problems in this paper and try to improve the accuracy of data collection and calculate the probability of eventual outcomes.

Reference
[1] R Eric Collins (2009), How many bacteria are in the ocean? And how far is it between bacteria in the ocean? http://www.reric.org/wordpress/archives/648
[2] Paul A. Gagniuc (2017), Markov Chains: From Theory to Implementation and Experimentation (1st edition). Wiley publishing, New York City.
[3] Joseph K. Blitzstein (2008), Markov Chain, Introduction to Probability (2nd edition), Athena Scientific publishing, Belmont, MA. pp. 406-409
[4] Sejal Jaiswal (2018), Markov Chains in Python: Beginner Tutorial. https://www.datacamp.com/community/tutorials/markov-chains-python-tutorial