



Geometric Interpretation and Manifold Structure of Markov Matrices

BÜLENT KARAKAŞ¹ , ŞENAY BAYDAŞ^{2,*} 

¹*Department of Mathematics, Faculty of Science, Bartın University, 74100, Bartın, Türkiye.*

²*Department of Mathematics, Faculty of Science, Van Yüzüncü Yıl University, 65080, Van, Türkiye.*

Received: 21-09-2020 • Accepted: 14-01-2021

ABSTRACT. In probability theory and statistics, the term Markov property refers to the memoryless property of a stochastic process. It is named after the Russian mathematician Andrey Andreyevich Markov. Every Markov matrix gives a linear equation system, and the solution of this equation system gives us a subset of \mathbb{R}_n^n . This paper presents the new manifold structure on the set of the Markov matrices. In addition, this paper presents the set of Markov matrices is drawable, and this gives geometrical interpretation to Markov matrices. For the proof, we use the one-to-one corresponding among $n \times n$ Markov matrices, the solution of linear equation system from derived Markov property, and the set of $(n - 1)$ -polytopes.

2010 AMS Classification: 60J99, 52B70

Keywords: Convex polytope, geometry, manifold, Markov matrices.

1. INTRODUCTION

A.A. Markov studied an important type of chance process. In this process, the outcome of a specified experiment can effect the outcome of the next experiment. This type of process is called a Markov chain.

Let $S = \{S_1, S_2, \dots, S_n\}$ be a set of states. The process begins in one of these states and moves successively from one state to another. Each move is called a step. If the chain is currently in state S_i , then it moves to state S_j at the next step with probability by P_{ij} , and this probability does not depend upon which states the chain was in before the current state. The probabilities P_{ij} are called transition probabilities. The transition probabilities of a Markov chain can be given in the Markov matrices.

A Markov (or stochastic) matrix A is defined to be a real $n \times n$ matrix with non-negative entries satisfying the equations $\sum_{j=1}^n A_{ij} = 1, 1 \leq i \leq n$.

The spectral properties of non-negative matrices and linear operators and in particular of Markov matrices have been studied in great detail, because of their great importance in finance, population dynamics, medical statistics, sociology and many other areas.

After the development of Markov chain, some of universal problems are studied by Laplace, D. Bernoulli and Ehrenfests [1].

In Markov's 1908 paper, for an irreducible chain satisfying an additional condition, Markov proved that 1 is a simple eigenvalue of transition matrix, and the moduli of the other eigenvalues are less than 1. This additional condition is stronger than the condition requiring a chain to be aperiodic. An interesting paper by Schneider examines Markov's

*Corresponding Author

Email addresses: bk@bartin.edu.tr (B. Karakaş), sbaydas@yyu.edu.tr (S. Baydaş)

irreducibility condition and his additional condition in detail. In the literature, there are a lot of article about Markov chain. Some of them are about the geometric structure of Markov matrices [7, 8].

Pullman (1965) and Bernhard (2013) studied geometry of finite Markov chains and geometry of Markov chain [2,6].

In this article, we construct a new manifold structure on Markov matrices. So we can interpret the Markov matrices, and we show that the Markov matrices can be drawn.

We will give the basic ideas needed for this article. Markov chains require to introduce q_i to be the probability that chain is in state i at time 0. In other words, $P(x_0 = i) = q_i$. The vector $\mathbf{q} = [q_1, \dots, q_n]$ is called the initial probability distribution for the Markov chain. The transition probabilities are displayed as an $n \times n$ transition probability matrix P . The transition probability matrix P may be written as

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

Given that the state at time t is i , the process must be somewhere at time $t + 1$. This means that for each i

$$\sum_{j=1}^n P(x_{t+1} = j | P(X_t = i)) = 1,$$

$$\sum_{j=1}^n P_{ij} = 1.$$

All entries in the transition probability matrix are nonnegative, and the sum of entries in each row must be 1.

Let P be the transition matrix of a Markov chain. The ij th entry $P_{ij}^{(n)}$ of the matrix P^n gives the the probability that the Markov chain, starting in states s_i , will be in state s_j after n steps. Let P be $n \times n$ matrix and $J = [1_j]$ is $n \times 1$ column matrix, all entries are 1. If

$$PJ = J, \tag{1.1}$$

then P is a Markov matrix and this is called Markov property [5].

The equation (1.3) is very useful and functional. Let A and B be two Markov matrices, $n \times n$. Then AB is a Markov matrix such that A and B are Markov matrices. So we can write

$$AJ = J, BJ = J$$

and we have

$$(AB)(J) = A(B(J)) = AJ = J$$

so AB is a Markov matrix.

Let v be a probability vector and P be a transition matrix of a Markov chain. Then the probability that the chain is in state s_i after n steps is the i th entry in the vector

$$v^{(n)} = vP^n$$

2. GEOMETRIC INTERPERATION AND MANIFOLD STRUCTURE OF MARKOV MATRICES

In this section we will give the geometric shape and manifold structure of Markov matrices.

Equation (2.1) is a property of Markov matrix.

$$PJ = J \tag{2.1}$$

where P is a Markov matrix, $n \times n$, and $J = [1_j]_{n \times 1}$. Equation (2.1) gives us a linear equation system. We will use x_{ij} Euclid coordinate function instead of P_{ij} . In this manner we can write the linear equation systems as follows

$$L_i : \sum_{j=1}^n x_{ij} = 1, \quad 1 \leq i \leq n.$$

The solution of L_i is a closed convex polytope which $x_{ij} \geq 0, \forall i, j, 1 \leq i, j \leq n$, and we show this solution set as $S(L_i)$.

Let

$$\begin{aligned} U_1 &= (1, 0, \dots, 0) \\ U_2 &= (0, 1, \dots, 0) \\ &\vdots \\ U_k &= (0, \dots, 0, 1, 0, \dots, 0) \\ &\vdots \\ U_n &= (0, 0, \dots, 0, 1) \end{aligned}$$

be given. $U_i U_{i+1}$ is a line segment. $\mathbf{U} = U_1 U_2 \cup U_2 U_3 \cup \dots \cup U_{n-1} U_n \cup U_n U_1$ is the collection of the line segments. Now we define a mapping φ , from \mathbf{U} to \mathbb{R}^n as follows.

$$\begin{aligned} \varphi(u_1, \dots, u_n) &= OA_1 + u_1 A_1 A_2 + u_1 u_2 A_2 A_3 + \dots \\ &\quad + u_{i-2} u_{i-1} A_{i-1} A_i + \dots + u_{n-2} u_{n-1} A_{n-1} A_n \end{aligned}$$

where $A_k = (0, \dots, 0, 1_k, 0, \dots, 0)$, $0 \leq u_1, \dots, u_n \leq 1$.

$\varphi(\mathbf{U})$ is a part of the hyperplane $x_1 + x_2 + \dots + x_n = 1$: HP . If we say $\mathbb{R}_+^n = \{(x_1, \dots, x_n) \mid x_i \geq 0, 1 \leq i \leq n\}$, then we have $\varphi(\mathbf{U}) = HP \cap \mathbb{R}_+^n$. For every point of $\varphi(\mathbf{U})$ is the solution any equation L_i .

In fact, if $Y \in \varphi(u_1, \dots, u_n)$, then

$$\begin{aligned} Y &= (1, 0, \dots, 0) + u_1(-1, 1, 0, \dots, 0) + \dots + \\ &\quad u_{i-2} u_{i-1} (0, \dots, 0, -1_{i-2}, 1_{i-1}, 0, \dots, 0) + \dots + (0, \dots, 0, u_{n-1}) \\ &= (1 - u_1, u_1 - u_2, \dots, u_{n-1} - u_{n-2}, u_{n-1}), \end{aligned}$$

that is

$$\sum_{i=1}^n y_i = 1.$$

For example, for $n = 3$, $\{S(L_i)\} = \varphi(\mathbf{U})$ is a triangle region with corners $A_1 = (1, 0, 0)$, $A_2 = (0, 1, 0)$ and $A_3 = (0, 0, 1)$. For every closed set A , we can find an open ball, $B_r(m)$, with center m and radius r , such that $A \subset B_r(m)$. For every, r, m , $B_r(m)$ is a (smooth) submanifold of \mathbb{R}^n [4]. Consequently, A is a closed submanifold of \mathbb{R}^n . P_{n-1} is a hyperplane in \mathbb{R}^n , with dimension $(n - 1)$. $\varphi(\mathbf{U}) \subset P_{n-1}$. So $\varphi(\mathbf{U})$ is a manifold with dimension $(n - 1)$.

Theorem 2.1 ([3]). *If M_1, M_2, \dots, M_r are differentiable manifolds with dimension, m_1, m_2, \dots, m_r , respectively, then $M_1 \times M_2 \times \dots \times M_r$ is a differentiable manifold with dimension $m_1 + m_2 + \dots + m_r$.*

Proof. Suppose that M_1, M_2, \dots, M_r are differentiable manifolds with dimension m_1, m_2, \dots, m_r . Let $\mathbf{A}_i = \{(U_{ij}, \varphi_{ij})\}$, $1 \leq i \leq r$ be differentiable structure of the manifold M_i . Then we can define an atlas as

$$\mathbf{A} = \{(U_{i1} \times U_{i2} \times \dots \times U_{ir}, \varphi_{i1} \times \varphi_{i2} \times \dots \times \varphi_{ir})\}.$$

The atlas \mathbf{A} is a differentiable atlas. The atlas \mathbf{A} defines a differentiable structure on $M = M_1 \times M_2 \times \dots \times M_r$ and M is a differentiable manifold with dimension $m = m_1 + m_2 + \dots + m_r$. \square

Theorem 2.2. *$S(L_1) \times S(L_2) \times \dots \times S(L_n)$ is a differentiable manifold with dimension $n(n - 1)$.*

Proof. We know that every $S(L_i)$ is a differentiable manifold with dimension $(n - 1)$. For every i , we take $S(L_i) = M_i$, and using Theorem 2.1, then we have $S = S(L_1) \times S(L_2) \times \dots \times S(L_n)$ is a product manifold with dimension $n(n - 1)$. \square

Theorem 2.3. *There are one-to-one corresponding among $MR(n \times n)$, $\{S(L_i)\}$ and $\varphi(\mathbf{U})^n$.*

Proof. Let $A \in MR(n \times n)$, $n \times n$ Markov matrices. Then every row A_i of A is the solution of a L_i and A goes to $S(L_i)$. Conversely, when $S(L_1), S(L_2), \dots, S(L_n)$ are given, we can write a Markov matrix which its rows are $A_i = S(L_i)$. We know that every point of $\varphi(\mathbf{U}) = Y$ has the property

$$\sum y_i = 1.$$

If we take a point $(Y_1, Y_2, \dots, Y_n) \in \varphi(\mathbf{U})^n$, we can write a Markov matrix which its rows are Y_i . Conversely, suppose that $A \in MR(n \times n)$. $A_i = (a_{i1}, a_{i2}, \dots, a_{in})$,

$$\sum a_{ij} = 1$$

we have $(A_1, \dots, A_n) \in \varphi(\mathbf{U})^n$. □

The case $n = 3$. Firstly we take the L_i equation.

$$L_i : x_{i1} + x_{i2} + x_{i3} = 1, \quad 1 \leq i \leq 3.$$

The solution of L_i is in the part of a plane at first part of space, $\forall x_i \geq 0$. For the sake of brevity, if we take $x = x_{i1}$, $y = x_{i2}$, $z = x_{i3}$, we have $P : x + y + z = 1$ in \mathbb{R}^3 . For the solution of the equation L_i we can write

$$S(L_i) \in P \cap \mathbb{R}_+^3$$

where $\mathbb{R}_+^3 = \{(x, y, z) \mid x, y, z \geq 0\}$.

In generally, let A, B and C be nonlinear points in \mathbb{R}^3 , then the 2-simplex $\Delta(ABC)$ can be given with the parametric form as follows.

$$\varphi(u, v) = OA + uAB + uvBC, \quad 0 \leq u, v \leq 1.$$

For $A = A_1 = (1, 0, 0)$, $B = A_2 = (0, 1, 0)$, $C = A_3 = (0, 0, 1)$, $0 \leq u_1, u_2 \leq 1$, we have

$$\varphi(u_1, u_2) = OA_1 + u_1A_1A_2 + u_1u_2A_2A_3.$$

If $X \in \varphi(u_1, u_2)$, then we write

$$\begin{aligned} X &= (1, 0, 0) + u_1(-1, 1, 0) + u_1u_2(0, -1, 1) \\ &= (1 - u_1, u_1 - u_1u_2, u_1u_2) \end{aligned}$$

and we have

$$\sum_{i=1}^3 x_i = 1,$$

where $x_1 = 1 - u$, $x_2 = u_1 - u_1u_2$ and $x_3 = u_1u_2$.

Every 3×3 Markov matrix gives us solution of a linear equation system $\{L_i, 1 \leq i \leq 3\}$. Every solution set $S(L_i)$ defines a 2-simplex $\Delta(A_1A_2A_3) = N$. So all 3×3 Markov matrices are represented by $N^3 = N \times N \times N$. Taking a point from N^3 is the same taking 3-points from N .

Theorem 2.4. *Non-collinear three points of N^3 defines a Markov matrix which determinant different zero.*

Proof. Let $A, B, C \in N^3$ be non-collinear. $[A \ B \ C]$ is a matrix which rows are linear independent so $\det[A \ B \ C] \neq 0$. □

Theorem 2.5. *3-fold points $(A, A, A) \in N^3$ define a Markov matrix that it is limit matrix.*

Proof. Let A be Markov matrix. For $n \in \mathbb{Z}^+$, every Markov matrix has the form that every rows are same. This matrix is called the limit matrix of given matrix. Every subset with three points of $\Delta(A_1A_2A_3)$ goes to a one point as limit. □

We show that there are a one-to-one corresponding between $MR(n \times n)$ and the n -simplices in the region which border are line segments $[A_iA_{i+1}]$, where $A_i = (0, \dots, \delta_{ij}, \dots, 0)$, $1 \leq i, j \leq n$. So all properties of n -simplices can be interpret in $MR(n \times n)$ and conversely.

For example, the centroid of the triangle $\Delta(A_1A_2A_3)$ is $C = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Point C defines a Markov matrix $C = [a_{ij}]$, $a_{ij} = \frac{1}{3}$, $1 \leq i, j \leq 3$. Matrix C takes the any Markov matrix to its limit matrix. That is, if A is any Markov matrix, CA is the limit matrix of A .

3. CONCLUSION

Every Markov matrix gives a linear equation system, and the solution of this equation system gives us a subset of $(\mathbb{R}^n)^n$. Every linear equation L_i has a solution on \mathbb{R}^n . This solution set lies on the $\sum x_i = 1$ hyperplane and from the restriction $x_i \geq 0$, the solution also lies on the first region of \mathbb{R}^n . Every $S(L_i)$ has a differentiable manifold structure and $(S(L_i))^n$ is a differentiable product manifold. $MR(n \times n)$, $n \times n$ Markov matrices, defines a differential manifold. This manifold is a product manifold and each multiplier is a region in first part of \mathbb{R}^n which border is a $(n - 1)$ -polytope and corner points are $A_i = (0, \dots, \delta_{ij}, \dots, 0)$. n -points from the region which corner points are A_i and border surface is $(n - 1)$ -polytope gives us a $n \times n$ Markov matrix. There is a one-to-one corresponding among $n \times n$ Markov matrices, the solution of linear equation system $\{L_i\}$ and $(n - 1)$ -polytope. So the Markov matrices can be drawn.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this article.

REFERENCES

- [1] Basharin, G.P., Langville, A.N., Naumov, V.A., *The life and work of A.A. Markov*, Linear Algebra Appl. **386**(2004), 3-26.
- [2] Bernhard, J., *The geometry of Markov chain limit theorems*, Markov Processes. Related Fields. **19**(1)(2003), 99-124.
- [3] Boothby, W.M., *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press California, 1975.
- [4] Helgason, S., *Differential Geometry, Lie Groups and Symmetric Space*. AMS Edition, Providence, 2001.
- [5] Privault, N., *Understanding Markov Chains*, Springer, Singapore, 2013.
- [6] Pullman, N., *The geometry of finite Markov chains*, Canad. Math. Bull. **8**(3)(1965), 345-358.
- [7] Sumner, J.G., Fernandez-Sanchez, J., Jarvis, P.D., *Lie Markov models*, J. Theoret. Biol. **2012**(2012), 298.
- [8] Sumner, J.G., *Lie geometry of 2×2 Markov matrices*, J. Theoret. Biol. **327**(2013), 88-90.