MATH2999 Directed Studies in Mathematics

Matrix Theory and Its Applications

Research Topic

Stationary Probability Vector

of a Higher-order Markov Chain

By

Zhang Shixiao

Supervisors: Prof. Li Chi-Kwong and Dr. Chan Jor-Ting

Contents

Abstract

In this project, our focus will be how to determine the stationary vector of a higher-order Markov Chain. Specifically, we mainly focus on the iterative method proposed by Li, Ng and Ye (2011). Rather than the computation method itself, our problem is what kind of conditions of the parameters will give out infinitely many solutions, multiple solutions or a unique solution to the higher-order Markov Chain, which means we want to study the number of stationary probability vectors in the solution set. This is a relatively new topic which may lead to more future analysis.

1. Introduction: Background

1.1 Fundamental Concepts and Results in the Theory of Non-negative Matrices

<u>Definition</u> A square matrix T is called *non-negative* if all its entries are non-negative real values. We write $T \geq 0$ to represent such matrices. Meanwhile T is a square matrix which is called *primitive* if there exists a positive integer k such that $T^k > 0$.

First note that a non-negative matrix is not sufficient to be a primitive matrix. A simple example is $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ are both primitive matrices since $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ while $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is non-negative but not primitive.

1.2 Discrete Time-Homogeneous Markov Chains

A discrete-time Markov chain is a stochastic process $\{X_t, t = 0,1,2\ldots\}$ with a discrete finite state space S such that with time independent probability $\Pr(X_{t+1} = j \mid X_t = i, X_{t-1} = i_{t-1}, X_{t-2} = i_{t-2}, \ldots,, X_1 = i_1, X_0 = i_0) = \Pr(X_{t+1} = j \mid X_t = i) = p_{ij}$ holds for all $i, j, i_0, \cdots, i_{t-1}$. Then a unit sum vector $\boldsymbol{X}$ is said to be a *stationary probability vector* or *distribution* of a finite Markov Chain if $\boldsymbol{X}P=\boldsymbol{X}$ where $P = (p_{ij}, i, j \in S)$. The following theorem guarantees the existence and uniqueness of the stationary probability vector of a discrete-time Markov Chain.

1.3 The Perron-Frobenius Theorem for Primitive Matrices (Without Proof)

Suppose T is an n by n non-negative primitive matrix. Then there exists an eigenvalue r such that r is a real positive simple root of the characteristic equation of T, r >|λ| for any eigenvalue λ≠r and the eigenvectors associated with r are unique to constant multiples. If 0≤B≤T and β is an eigenvalue of B, then |β|≤r. Moreover, |β|=r implies B=T.

We have to admit that the *Perron-Frobenius Theorem* is the most fundamental theorem for non-negative matrices, and it highlights and guarantees the so-called nature of finite Markov Chains, which is the convergence of an irreducible finite Markov Chain to its stationary probability distribution.

2. Higher-order Markov Chains

2.1 Definitions and Concepts

We consider a stochastic process with a sequence of random variables, $\{X_t, t=0,1,2\ldots\}$, which takes on a finite set $S=\{0,1,2,\ldots,n\}$ called the state set of the process.

Definition 2.1 Suppose the probability independent of time satisfying

$$\Pr\left(X_{t+1}=i \mid X_t=i_1, X_{t-1}=i_2, X_{t-2}=i_3,\ldots,, X_1=i_t, X_0=i_{t+1}\right)$$
$$=\Pr\left(X_{t+1}=i \mid X_t=i_1, X_{t-1}=i_2, X_{t-2}=i_3,\ldots, X_{t-m+1}=i_m\right)=p_{i,i_1,i_2,\cdots,i_m}$$

where $i,i_1,i_2,\cdots,i_m\in S$, then it is called a *m-th* order Markov Chain, in other words, the current state of the process depends on *m* past states. Observed that $\sum_{i=1}^{n} p_{i,i_1,i_2,\cdots,i_m}=1$. When *m=1*, it is just the regular standard Markov Chain.

Definition 2.2 Write $\mathrm{A}=\left(a_{i_1i_2i_3}\right)$ to be a three-order *n*-dimensional tensor, where $a_{i_1i_2i_3}\in\mathbb{R}$ and $1\le i_1,i_2,i_3\le n$, define an *n*-dimensional column vector

$$\mathrm{A}X^2=\left(\sum_{i_2,i_3=1}^{n} a_{ii_2i_3}x_{i_2}x_{i_3}\right)_{1\le i\le n} \quad \text{given } X=\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$$

Warning: this *'three-order'* has nothing to do with the *'m-th'* order in the previous Definition 2.1

Definition 2.3 A three-order *n*-dimensional tensor $\mathrm{A} = \left(a_{i_1 i_2 i_3}\right)$ is called *reducible* if there exists a non-empty proper index subset $I \subset \{1, 2, \ldots, n\}$ such that $a_{i_1 i_2 i_3} = 0, \forall i_1 \in I, \forall i_2, i_3 \notin I$, if $\mathrm{A} = \left(a_{i_1 i_2 i_3}\right)$ is not reducible, we call it *irreducible*.

In fact, if *P* is an irreducible non-negative three-order *n*-dimensional tensor of a high-order Markov Chain, Li, Ng and Ye (2011) has proved that in order to obtain the stationary probability vector ***X*** of a high-order Markov Chain, we just need to solve $\mathrm{P}X^2 = X$.

2.2 Conditions for Each Point in the Simplex Being a Stationary Vector

For simplicity, we rewrite the above equation $\mathrm{P}X^2 = X$ for a tensor as

$$x_1 A_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + x_2 A_2 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \cdots + x_n A_n \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

where A_i ’s are $n \times n$ column stochastic matrices with all entries are real numbers.

Theorem 2.1 Proposition about number of the stationary vectors for 2×2 case

Now we are considering, where all $a_1, a_2, b_1, b_2 \in [0,1]$, $x \in [0,1]$

$$x \begin{pmatrix} a_1 & b_1 \\ 1-a_1 & 1-b_1 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} + (1-x) \begin{pmatrix} a_2 & b_2 \\ 1-a_2 & 1-b_2 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} = \begin{pmatrix} x \\ 1-x \end{pmatrix}$$

Then one of the following holds

(1) If $a_1 = 1, b_2 = 0, a_2 + b_1 = 1$, then we must have infinitely many solutions, namely, every $\begin{pmatrix} x \\ 1-x \end{pmatrix}$ with $x \in [0,1]$ is a solution to the above equation.

(2) If $a_1 = 1, a_2 + b_1 < 1$, then we must have two solutions $x = 1$ or $x = \dfrac{b_2}{b_2 + 1 - a_2 - b_1}$ to the above equation.

(3) Otherwise, we must have a unique solution with the condition that

If $a_1 - a_2 - b_1 + b_2 = 0$, excluding the condition in (1), then $x = \dfrac{b_2}{2b_2 + 1 - a_2 - b_1}$

If $a_1 - a_2 - b_1 + b_2 \neq 0$, then $x = \dfrac{2b_2 + 1 - a_2 - b_1 - \sqrt{\Delta}}{2\left(a_1 - a_2 - b_1 + b_2\right)}$ given

$$\Delta = (a_2 + b_1 - 2b_2 - 1)^2 - 4b_2(a_1 - a_2 - b_1 + b_2) = (a_2 + b_1 - 1)^2 + 4b_2(1 - a_1) \geq 0$$

Proof: The setting is as above, write

$$f(x) = \left[x \begin{pmatrix} a_1 & b_1 \\ 1-a_1 & 1-b_1 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} + (1-x) \begin{pmatrix} a_2 & b_2 \\ 1-a_2 & 1-b_2 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} \right]_1$$

$$= (a_1 - a_2 - b_1 + b_2)x^2 + (a_2 + b_1 - 2b_2)x + b_2$$

We want to solve $f(x) = x$ to determine the solution.

Observed that if we set

$$g(x) = f(x) - x = (a_1 - a_2 - b_1 + b_2)x^2 + (a_2 + b_1 - 2b_2 - 1)x + b_2 = 0$$

$g(0) = b_2 \geq 0$ and $g(1) = a_1 - 1 \leq 0$, hence by the *Intermediate Value Theorem*, there must exist at least one $x_0 \in [0,1]$, such that $g(x_0) = 0$

Let $\Delta = (a_2 + b_1 - 2b_2 - 1)^2 - 4b_2(a_1 - a_2 - b_1 + b_2) = (a_2 + b_1 - 1)^2 + 4b_2(1 - a_1) \geq 0$

If $a_1 - a_2 - b_1 + b_2 = 0$, the quadratic equation reduced to $(a_2 + b_1 - 2b_2 - 1)x + b_2 = 0$, then if $a_2 + b_1 - 2b_2 - 1 = 0$, i.e. $a_1 = 1, b_2 = 0, a_2 + b_1 = 1$, there are infinitely many solutions; otherwise if $a_2 + b_1 - 2b_2 - 1 \neq 0$, then the *Intermediate Value Theorem* guarantees that the unique solution is $x = \dfrac{b_2}{2b_2 + 1 - a_2 - b_1}$

If $a_1 - a_2 - b_1 + b_2 \neq 0$, there are two solutions to the quadratic equation, which are

$$x_1 = \frac{2b_2 + 1 - a_2 - b_1 - \sqrt{\Delta}}{2(a_1 - a_2 - b_1 + b_2)} \text{ and } x_2 = \frac{2b_2 + 1 - a_2 - b_1 + \sqrt{\Delta}}{2(a_1 - a_2 - b_1 + b_2)}$$

When $\Delta = 0$, no matter $a_1 - a_2 - b_1 + b_2 < 0$ or $a_1 - a_2 - b_1 + b_2 > 0$, we get only one solution $x = \dfrac{2b_2 + 1 - a_2 - b_1}{2(a_1 - a_2 - b_1 + b_2)}$, it must be the unique solution we want by the *Intermediate Value Theorem*.

When $\Delta > 0$, note that $(2b_2 + 1 - a_2 - b_1)^2 - \Delta = 4b_2(a_1 - a_2 - b_1 + b_2)$

and

$$[2b_2 + 1 - a_2 - b_1 - 2(a_1 - a_2 - b_1 + b_2)]^2 - \Delta = (1 + a_2 + b_1 - 2a_1)^2 - \Delta = 4(a_1 - 1)(a_1 - a_2 - b_1 + b_2)$$

If $a_1 - a_2 - b_1 + b_2 < 0$, of course $b_2 \neq 0$ since we set $\Delta > 0$,

hence $(2b_2 + 1 - a_2 - b_1)^2 - \Delta < 0 \Leftrightarrow -\sqrt{\Delta} < 2b_2 + 1 - a_2 - b_1 < \sqrt{\Delta}$, then $x_2 = \frac{2b_2 + 1 - a_2 - b_1 + \sqrt{\Delta}}{2(a_1 - a_2 - b_1 + b_2)} < 0$, then we conclude there is only one satisfied solution which is $x_1 = \frac{2b_2 + 1 - a_2 - b_1 - \sqrt{\Delta}}{2(a_1 - a_2 - b_1 + b_2)}$

If $a_1 - a_2 - b_1 + b_2 > 0$, then $(2b_2 + 1 - a_2 - b_1)^2 - \Delta = 4b_2(a_1 - a_2 - b_1 + b_2) \ge 0$ and

$$\left[2b_2 + 1 - a_2 - b_1 - 2(a_1 - a_2 - b_1 + b_2)\right]^2 - \Delta = (1 + a_2 + b_1 - 2a_1)^2 - \Delta = 4(a_1 - 1)(a_1 - a_2 - b_1 + b_2) \le 0$$

(i) $a_1 \ne 1, b_2 \ne 0$, then

$$x_2 - 1 = \frac{2b_2 + 1 - a_2 - b_1 + \sqrt{\Delta}}{2(a_1 - a_2 - b_1 + b_2)} - 1 = \frac{1 + a_2 + b_1 - 2a_1 + \sqrt{\Delta}}{2(a_1 - a_2 - b_1 + b_2)} > 0 \Leftrightarrow x_2 > 1,$$

hence we conclude that $x_1 = \frac{2b_2 + 1 - a_2 - b_1 - \sqrt{\Delta}}{2(a_1 - a_2 - b_1 + b_2)}$ is the unique solution we require.

(ii) $a_1 \ne 1, b_2 = 0$ and $a_2 + b_1 \ne 1$, then the situation is the same as (i), just plug in $b_2 = 0$

(iii) $a_1 = 1, b_2 \ne 0$ and $a_2 + b_1 \ge 1$, then $\sqrt{\Delta} = 1 + a_2 + b_1 - 2a_1$, $x_1 = \frac{2b_2 + 1 - a_2 - b_1 - \sqrt{\Delta}}{2(a_1 - a_2 - b_1 + b_2)} = 1$ and $x_1 = \frac{b_2}{(1 - a_2 - b_1) + b_2} \ge 1$, hence we conclude the unique solution we require is $x_1 = 1$

(iv) $a_1 = 1, b_2 \ne 0$ and $a_2 + b_1 < 1$, then there are two solutions, which are $x = 1$ and $x = \frac{b_2}{b_2 + 1 - a_2 - b_1}$, just the same situation with (iii) except that $0 < x_1 = \frac{b_2}{(1 - a_2 - b_1) + b_2} < 1$

(v) $a_1 = 1, b_2 = 0$ and $a_2 + b_1 \ne 1$, directly plug in the value and solve the equation we get $x = 1$ and $x = 0$

Then we want to extend the condition for infinitely many solutions for $n \times n$ case

Theorem 2.2 For $x_1 A_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + x_2 A_2 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \cdots + x_n A_n \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, each element in the set

$\left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} | \sum_{i=1}^{n} x_i = 1, x_i \in [0,1] \; for \forall i = 1, 2, \cdots, n \right\}$ would be a solution to the equation

if and only if

$$A_1 = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ & 1-a_{12} & & \\ & & \ddots & \\ & & & 1-a_{1n} \end{pmatrix} \quad , \quad A_2 = \begin{pmatrix} 1-a_{12} & & & & \\ a_{12} & 1 & a_{23} & \cdots & a_{2n} \\ & & 1-a_{23} & & \\ & & & \ddots & \\ & & & & 1-a_{2n} \end{pmatrix} \quad , \quad \ldots\ldots,$$

$$A_i = \begin{pmatrix} 1-a_{1i} & & & & & & & \\ & 1-a_{2i} & & & & & & \\ & & \ddots & 1-a_{i-1,i} & & & & \\ a_{1i} & a_{2i} & \cdots & a_{i-1,i} & 1 & a_{i,i+1} & \cdots & a_{in} \\ & & & & & 1-a_{i,i+1} & & \\ & & & & & & \ddots & \\ & & & & & & & 1-a_{in} \end{pmatrix} \quad , \quad \ldots\ldots,$$

$$A_n = \begin{pmatrix} 1-a_{1n} & & & & & \\ & 1-a_{2n} & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & 1-a_{n-1,n} & \\ a_{1n} & a_{2n} & \cdots & \cdots & a_{n-1,n} & 1 \end{pmatrix}$$

Proof: "$\Leftarrow$", trivial, by directly checking row by row.

"$\Rightarrow$", first note that in Theorem 2.1, we have proved that for 2×2 case, if we have infinitely many solutions, the two matrices must be of the form

$A_1 = \begin{pmatrix} 1 & a_{12} \\ 0 & 1-a_{12} \end{pmatrix}, A_2 = \begin{pmatrix} 1-a_{12} & 0 \\ a_{12} & 1 \end{pmatrix}$, hence, for 3×3 case

$x_1\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}X + x_2\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & a_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}X + x_3\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}X = X$, if we set the third entry of the stationary vector X to be 0, then we can have infinitely many solutions as $X = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$ with $x_1 + x_2 = 1$, $x_1, x_2 \in [0,1]$ if and only if the sub-matrices $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ of A_1 and A_2 must be with the form of $\begin{pmatrix} 1 & a_{12} \\ 0 & 1-a_{12} \end{pmatrix}$ and $\begin{pmatrix} 1-a_{12} & 0 \\ a_{12} & 1 \end{pmatrix}$, which uniquely determine the entries of $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$. Similarly, we set the second entry of X to be 0, we can uniquely determine the sub-matrices $\begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix}$ and $\begin{pmatrix} c_{11} & c_{13} \\ c_{31} & c_{33} \end{pmatrix}$ of A_1 and A_3, which must be with the form of $\begin{pmatrix} 1 & a_{13} \\ 0 & 1-a_{13} \end{pmatrix}$ and $\begin{pmatrix} 1-a_{13} & 0 \\ a_{13} & 1 \end{pmatrix}$; also, we set the first entry of X to be 0, we can uniquely determine the sub-matrices $\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$ and $\begin{pmatrix} c_{22} & c_{23} \\ c_{32} & c_{33} \end{pmatrix}$ of A_2 and A_3, which must be with the form of $\begin{pmatrix} 1 & a_{23} \\ 0 & 1-a_{23} \end{pmatrix}$ and $\begin{pmatrix} 1-a_{23} & 0 \\ a_{23} & 1 \end{pmatrix}$, therefore, the three matrices are uniquely determined by

$$A_1 = \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & 1-a_{12} & 0 \\ 0 & 0 & 1-a_{13} \end{pmatrix}, A_2 = \begin{pmatrix} 1-a_{12} & 0 & 0 \\ a_{12} & 1 & a_{23} \\ 0 & 0 & 1-a_{23} \end{pmatrix}, A_3 = \begin{pmatrix} 1-a_{13} & 0 & 0 \\ 0 & 1-a_{23} & 0 \\ a_{13} & a_{23} & 1 \end{pmatrix}.$$

Similarly, we use this method and result for 3×3 case to determine the four matrices of 4×4 case. Inductively, we use the $(n-1)\times(n-1)$ with

$$A_1 = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1,n-1} \\ & 1-a_{12} & & \\ & & \ddots & \\ & & & 1-a_{1,n-1} \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1-a_{12} & & & & \\ a_{12} & 1 & a_{23} & \cdots & a_{2,n-1} \\ & & 1-a_{23} & & \\ & & & \ddots & \\ & & & & 1-a_{2,n-1} \end{pmatrix}, \quad \ldots\ldots,$$

$$A_i = \begin{pmatrix} 1-a_{1i} & & & & & & & \\ & 1-a_{2i} & & & & & & \\ & & \ddots & 1-a_{i-1,i} & & & & \\ a_{1i} & a_{2i} & \cdots & a_{i-1,i} & 1 & a_{i,i+1} & \cdots & a_{i,n-1} \\ & & & & & 1-a_{i,i+1} & & \\ & & & & & & \ddots & \\ & & & & & & & 1-a_{i,n-1} \end{pmatrix} \qquad , \qquad \ldots\ldots,$$

$$A_{n-1} = \begin{pmatrix} 1-a_{1,n-1} & & & & & \\ & 1-a_{2,n-1} & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & 1-a_{n-2,n-1} & \\ a_{1,n-1} & a_{2,n-1} & \cdots & \cdots & a_{n-2,n-1} & 1 \end{pmatrix}$$ for *i=1,2,…, (n-1)*

When we set the *j-th* entry of $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ $x_j = 0$ for *j=1,…,n* respectively, then we can determine the *(n-1)* sub-matrices which are exactly of the form

$$B_1 = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ & 1-a_{12} & & & & & \\ & & \ddots & & & & \\ & & & 1-a_{1,j-1} & & & \\ & & & & 1-a_{1,j+1} & & \\ & & & & & \ddots & \\ & & & & & & 1-a_{1n} \end{pmatrix}$$

$$B_2 = \begin{pmatrix} 1-a_{12} & & & & & & & \\ a_{12} & 1 & a_{23} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ & & 1-a_{23} & & & & & \\ & & & \ddots & & & & \\ & & & & 1-a_{2,j-1} & & & \\ & & & & & 1-a_{2,j+1} & & \\ & & & & & & \ddots & \\ & & & & & & & 1-a_{2n} \end{pmatrix}$$

……

$$B_{j-1}=\begin{pmatrix} 1-a_{1,j-1} & & & & & & & \\ & 1-a_{2,j-1} & & & & & & \\ & & \ddots & 1-a_{j-2,j-1} & & & & \\ a_{1,j-1} & a_{2,j-1} & \cdots & a_{j-2,j-1} & 1 & a_{j-1,j+1} & \cdots & a_{j-1,n} \\ & & & & & 1-a_{j-1,j+1} & & \\ & & & & & & \ddots & \\ & & & & & & & 1-a_{j-1,n} \end{pmatrix}$$

$$B_{j+1}=\begin{pmatrix} 1-a_{1,j+1} & & & & & & & \\ & 1-a_{2,j+1} & & & & & & \\ & & \ddots & 1-a_{j-1,j+1} & & & & \\ a_{1,j+1} & a_{2,j+1} & \cdots & a_{j-1,j+1} & 1 & a_{j+1,j+2} & \cdots & a_{j+1,n} \\ & & & & & 1-a_{j+1,j+2} & & \\ & & & & & & \ddots & \\ & & & & & & & 1-a_{j+1,n} \end{pmatrix}$$

……

$$B_n=\begin{pmatrix} 1-a_{1n} & & & & & & & \\ & 1-a_{2n} & & & & & & \\ & & \ddots & & & & & \\ & & & 1-a_{j-1,n} & & & & \\ & & & & 1-a_{j+1,n} & & & \\ & & & & & \ddots & & \\ & & & & & & 1-a_{n-1,n} & \\ a_{1n} & a_{2n} & \cdots & a_{j-1,n} & a_{j+1,n} & \cdots & a_{n-1,n} & 1 \end{pmatrix}$$

By combining all the sub-matrices obtained within n steps, use the same a_{ij}'s from above, hence, we get the matrices $A_1, \ldots, A_n$ being the form we require.

Remark: Observed that if we pick up the *i-th* row of each A_i to form a new matrix M, we can easily see that

$$M=\begin{pmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{12} & 1 & a_{22} & \cdots & a_{2n} \\ a_{13} & a_{22} & \ddots & & \vdots \\ \vdots & \vdots & & 1 & a_{n-1,n} \\ a_{1n} & a_{2n} & \cdots & a_{n-1,n} & 1 \end{pmatrix}$$, which is an $n\times n$ symmetric matrix with all entries of the diagonal equal to 1. Li, Ng and Ye (2011) state that given *P* is an irreducible non-negative tensor of order *p* and dimension *n*, if *1* is not the eigenvalue of *DT(x)*, the Jacobian matrix of *T*, for all $x\in\Omega\backslash\partial\Omega$, then ***X*** is unique where $T:\Omega\to\Omega,\left[T\left(X\right)\right]_i=\left[PX^{p-1}\right]_i$.

In fact, if the *i-th* column of A_i is e_i, which is *i-th* column of the identity matrix, and all the other entries equal to $\frac{1}{n}$, then there are *n* solutions which are e_1, e_2, ..., e_n. And there must be no common zero $n\times\left(n-k\right)$ blocks within these matrices. Finally, we state out an independent conclusion describing the nature of number of solutions.

<u>Theorem 2.3</u> Given any two solutions lying on the interior of 1-dimensional face of the boundary of the simplex, then the whole 1-dimensional face must be a set of collection of solutions to the above equation.

Proof: Observed that, for 2×2 case, if we are given two solutions of the form $X_1=\begin{pmatrix} x_1 \\ 1-x_2 \end{pmatrix}$ and $X_2=\begin{pmatrix} x_2 \\ 1-x_2 \end{pmatrix}$ where $x_1,x_2\in\left(0,1\right)$, then we must conclude that there are infinitely many solutions $X=\begin{pmatrix} x \\ 1-x \end{pmatrix},x\in\left[0,1\right]$ since if we have two and only two solutions, one of them must be $X=\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, which is a contradiction. Therefore, given two solution of the form

$$X_1=\begin{pmatrix} 0 & \cdots & 0 & \underset{i-th}{x_i} & 0 & \cdots & 0 & \underset{j-th}{x_j} & 0 & \cdots & 0 \end{pmatrix}^T$$

and $X_2=\begin{pmatrix} 0 & \cdots & 0 & \underset{i-th}{x_i'} & 0 & \cdots & 0 & \underset{j-th}{x_j'} & 0 & \cdots & 0 \end{pmatrix}^T$

with $x_i,x_j,x_i',x_j'\in\left(0,1\right),x_i+x_j=1,x_i'+x_j'=1$, then the two 2×2 sub-matrices, which are obtained by picking up the *i-th* and *j-th* rows and *i-th* and *j-th* columns from the A_i and A_j must be of the form $\begin{pmatrix} 1 & a \\ 0 & 1-a \end{pmatrix}$ and $\begin{pmatrix} 1-a & 0 \\ a & 1 \end{pmatrix}$, which will give out infinitely many solutions. i.e. all points lie on the whole 1-dimensional face will be a solution. From this perspective, we conclude that we could not observe three points which two of them lie on the same 1-dimensional face and a single point outside the face.

We conjecture that given any $k+1$ solutions lying in the interior of the k-dimensional face of the simplex, and any q of them ($q < k$) do not lie on the same $(q\text{-}1)$-dimensional face, then any point lying in the whole k-dimensional face, including the vertexes and boundaries, will be a solution to the equation. We leave it for readers to prove or disprove the result.

2.3 Applications in DNA Sequence Prediction

Higher-order Markov Chains are often used to describe the flow direction of sequences of random variables. One important application in predicting the DNA sequence rises up in recent years. In the book written by Ching and Ng (2006), they also highlight this aspect by considering the mouse αA-crystallin gene (Raftery and Tavare 1994). The main idea is to rewrite the model into the following mathematical form:

$$\mathbf{x}_{t+n} = \sum_{i=1}^{n} \lambda_i Q_i \mathbf{x}_{t+n-i}, \quad \text{where } \sum_{i=1}^{n} \lambda_i = 1 \text{ and } \lambda_i \geq 0$$

$\mathbf{x}_{t+i}$ is the state vector at time $(t + i)$ and $\mathbf{x}_{t+n}$ depends on $\mathbf{x}_{t+n-i}$ $(i = 1, 2, \ldots, n)$, then if Q_i is irreducible, $\lambda_i > 0$, then the model has a stationary distribution $\boldsymbol{x}$, where $\boldsymbol{x}$ is the unique solution of the linear system (Zhu and Ching 2011):

$$(I - \sum_{i=1}^{n} \lambda_i Q_i)\mathbf{x} = 0 \quad \text{and} \quad \mathbf{1}^T \mathbf{x} = 1$$

Here we ignore the details since we are considering the number of stationary probability vectors to our three-order n-dimensional tensor. But indeed, they have some underlying connections. For details, you may refer to (Ching and Ng 2006) and (Zhu and Ching 2011).

3. Conclusion

In this paper, we start with the results proposed by Li, Ng and Ye (2011) and try to figure out the assumption conjecture they raised in their paper. Originally, they are considering the general solution to a p-order n-dimensional tensor, but due to our understanding about the tensor itself, we are not considering many situations. But eventually we end up with some beautiful small theorems describing the nature of infinitely many solutions over the whole simplex for three-order case. Many other corollaries could be deduced from what we state in Theorem 2.2, we leave it for readers to reach out some more influential conclusions.

References

1. Li. Ng. and Ye. *Finding Stationary Probability Vector of a Transition Probability Tensor Arising from a Higher-order Markov Chain*. Department of Computer Science, Shenzhen Graduate School, Harbin Institute of Technology. Department of Mathematics, The Hong Kong Baptist University, Hong Kong. 8 February 2011.
2. Seki, T. *An Elementary Proof of Brouwer's Fixed Point Theorem.* Tokyo College of Science. Received December 25, 1956.
3. Morrison, C. and Stynes, M. *An Intuitive Proof of Brouwer's Fixed Point Theorem in* R^2. Mathematics Magazine, Vol. 56, No. 1, pp. 38-41. Mathematical Association of America. 1983. [Online]. Available: http://www.jstor.org/stable/2690266. Accessed: 20/03/2012 08:50.
4. Adke, R. and Deshmukh, R. *Limit Distribution of a High Order Markov Chain*. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 50, No.1, pp. 105-108. Blackwell Publishing for the Royal Statistical Society. 1988. [Online]. Available: http://www.jstor.org/stable/2345812. Accessed: 19/03/2012 08:56.
5. Zhu, D. and Ching, W. *A Note on the Stationary Property of High-dimensional Markov Chain Models*. International Journal of Pure and Applied Mathematics. Volume 66, No. 3, 321-330. 2011. [Online]. Available: http://www.ijpam.eu/contents/2011-66-3/6/6.pdf. Accessed: 10/05/2012 23:13.
6. Ching, W. and Ng, K. *Markov Chains: Models, Algorithms and Applications*. International Series in Operations Research & Management Science. Volume 83. 2006.
7. Seneta, E. *Non-negative Matrices and Markov Chains*. New York: Springer. [2nd ed]. 2006.