The Estimation for the Eigenvalues of Stochastic Matrices

Yuan Lu (Corresponding author)

Department of Fundamental Sciences, Yancheng Institute of Technology

Yancheng 224051, China

E-mail: ycgxyluyuan@yahoo.com.cn

Abstract

The purpose of this paper is to locate and estimate the eigenvalues of stochastic matrices. We present several estimation theorems about the eigenvalues of stochastic matrices. Meanwhile, we obtain the distribution theorem for the eigenvalues of tensor product of two stochastic matrices. We will conclude the paper with the distribution for the eigenvalues of generalized stochastic matrices.

Keywords: Stochastic matrices, Eigenvalues, Tensor product

1. Introduction

In the past two decades, due to study on matrix theory and some engineering background problems, many scholars dedicated to special matrix , and obtained some important and valuable results (TingZhu, 2007 - Yigeng Huang, 1994). But in combination matrix theory, combinatorics, probability theory (especially Markov chain), mathematical economics and reliability theory etc. areas there is a special class of non-negative stochastic matrix, which in recent years becomes concerned. This article discusses location, distribution and estimate of the eigenvalue for stochastic matrix . Section 2 introduces the concepts of stochastic matrix and generalized stochastic matrix. Section 3 gives a few estimation theorems of stochastic matrix eigenvalue, also the eigenvalue distribution for tensor product of two stochastic matrices is obtained. In section 4, we discuss the eigenvalue distribution for generalized stochastic matrices.

2. Basic concept

Definition 1. If the sum of elements in every row in the n order non-negative matrix A is 1, A is row stochastic matrix; If the sum of elements in every column in the n order non-negative matrix A is 1, A is column stochastic matrix; If both A and A^T are row stochastic matrices, A is double stochastic matrix; Row stochastic matrix, column stochastic matrix and double stochastic matrix are called stochastic matrix, denoted by $S(n)$.

Definition 2. If the sum of elements in every row in the n order non-negative matrix A is s, A is called the first generalized row stochastic matrix; If the sum of elements in every column in the n order non-negative matrix A is s, A is called the first generalized column stochastic matrix; If both A and A^T are the first generalized row stochastic matrices, A is called the first generalized double stochastic matrix; The first generalized row stochastic matrix, the first generalized column stochastic matrix and the first generalized double stochastic matrix are called the first generalized stochastic matrix, denoted by $S_I(n)$.

Definition 3. If the absolute value sum of elements in every row in the n order matrix A is 1, A is called the second generalized row stochastic matrix; If the sum of elements in every column in the n order matrix A is 1, A is called the second generalized column stochastic matrix; If both A and A^T are the second generalized row stochastic matrices, A is called the second generalized double stochastic matrix; The second generalized row stochastic matrix, the second generalized column stochastic matrix and the second generalized double stochastic matrix are called the second generalized stochastic matrix, denoted by $S_{II}(n)$.

Definition 4. If the absolute value sum of elements in every row in the n order matrix A is s, A is called the third generalized row stochastic matrix; If the sum of elements in every column in the n order matrix A is s, A is called the third generalized column stochastic matrix; If both A and A^T are the third generalized row stochastic matrices, A is called the third generalized double stochastic matrix; The third generalized row stochastic matrix, the third generalized column stochastic matrix and the third generalized double stochastic matrix are called the third generalized stochastic matrix, denoted by $S_{III}(n)$.

$S_I(n), S_{II}(n)$ and $S_{III}(n)$ are called generalized stochastic matrices. Obviously, for $S(n), S_I(n), S_{II}(n)$ and $S_{III}(n)$, we have the following simple conclusions: (1).$S(n) \subset S_I(n) \subset S_{III}(n)$; (2).$S(n) \subset S_{II}(n)$; (3).$S(n) \subset S_{II}(n) \subset S_{III}(n)$.

3. Eigenvalue estimate of stochastic matrix

Theorem 1. Suppose $A = (a_{ij})_{n\times n}$ is a row stochastic matrix and $m = min\{a_{ii}, i = 1, 2, \cdots, n\}$, then

$$\lambda(A) \subset G(A) = \{z : |z - m| \leq 1 - m\},$$

where $\lambda(A)$ is denoted the whole eigenvalues of matrix A, $G(A)$ is Gerschgorin disc of matrix A.

Proof: Since λ is an arbitrary eigenvalue of matrix $A = (a_{ij})_{n\times n}$ and $X = (x_1, x_2, \cdots, x_n)^T \in R^{n\times 1}$ is the corresponding column eigenvector, let

$$y_i = \frac{x_i}{t_i}, \text{where} \quad t_i(i = 1, 2, \cdots, n) \text{is positive number,}$$

and

$$|y_m| = max|y_i|(i = 1, 2, \cdots, n),$$

and from $AX = \lambda X$, we get

$$\lambda t_i y_i = \sum_{j=1}^{n} a_{ij} t_j y_j,$$

$$\lambda t_m y_m = \sum_{j=1}^{n} a_{mj} t_j y_j = t_m a_{mm} y_m + \sum_{j=1, j\neq m}^{n} t_j a_{mj} y_j.$$

Multiply right each item of the above equation with y_m^*, then

$$\lambda t_m y_m y_m^* = t_m a_{mm} y_m y_m^* + \sum_{j=1, j\neq m}^{n} t_j y_j y_m^*,$$

i.e.

$$\lambda t_m - t_m a_{mm} = \frac{\sum\limits_{j\neq m}^{n} t_j a_{mj} y_j y_m^*}{|y_m|^2}.$$

By using trigonal inequality, we get

$$|\lambda t_m - t_m a_{mm}| \leq \sum_{j=1, j\neq m}^{n} |t_j a_{mj}|,$$

i.e.

$$|\lambda - a_{mm}| \leq P_m = \sum_{j=1, j\neq m}^{n} |a_{mj}| = 1 - a_{mm}.$$

Therefore,

$$|\lambda - m| = |\lambda - a_{mm} + a_{mm} - m| \leq |\lambda - a_{mm}| + |a_{mm} - m| \leq 1 - a_{mm} + a_{mm} - m = 1 - m.$$

Since λ is an arbitrary eigenvalue of matrix $A = (a_{ij})_{n\times n}$, then we have

$$\lambda(A) \subset G(A) = \{z : |z - m| \leq 1 - m\},$$

so the eigenvalues of A are located in the Gerschgorin disc whose center is $m = min\{a_{ii}, i = 1, 2, \cdots, n\}$ and radius is $1 - m$.

Theorem 2. Suppose $A = (a_{ij})_{n\times n}$ is a row stochastic matrix and $M_i = \max\limits_{i}\{a_{ij}, j = 1, 2, \cdots, n\}$, then

$$\lambda(A) \subset G(A) = \{z : |z - \frac{Tr(A)}{n}| \leq \sqrt{\frac{n-1}{n}(\sum_{i=1}^{n} M_i - \frac{(Tr(A))^2}{n})}\},$$

where $\lambda(A)$ is denoted the whole eigenvalues of matrix A, $G(A)$ is denoted disc whose center is $\frac{Tr(A)}{n}$ and radius is $\sqrt{\frac{n-1}{n}(\sum\limits_{i=1}^{n} M_i - \frac{(Tr(A))^2}{n})}$.

Proof: From paper (Yixi Gu, 1994) and for arbitrary matrix A, we have

$$|\lambda - \frac{Tr(A)}{n}| \leq \sqrt{\frac{n-1}{n}(||A||_F^2 - \frac{(Tr(A))^2}{n})}.$$

And because $A = (a_{ij})_{n\times n} \in S(n)$, $||A||_F^2 \leq \sum\limits_{i=1}^{n} M_i$. So we have

$$\lambda(A) \subset G(A) = \{z : |z - \frac{Tr(A)}{n}| \leq \sqrt{\frac{n-1}{n}(\sum_{i=1}^{n} M_i - \frac{(Tr(A))^2}{n})}\},$$

Similarly, we get

$$\lambda(A) \subset G(A) = \{z : |z - \frac{Tr(A)}{n}| \leq \sqrt{(n-1)(1-(\frac{Tr(A)}{n})^2}\}.$$

Theorem 3. Suppose $A = (a_{ij})_{n\times n}$ and $B = (b_{ij})_{m\times m}$ are row stochastic matrices, $m_1 = min\{a_{ii}, i = 1, 2, \cdots, n\}$ and $m_2 = min\{b_{jj}, j = 1, 2, \cdots, m\}$, then

$$\lambda(A \otimes B) \subset G(A \otimes B) = \{z : |z - m_1| \leq 1 - m_1\} \cdot \{z : |z - m_2| \leq 1 - m_2\},$$

where $\lambda(A \otimes B)$ is denoted the whole eigenvalues of tensor product for matrix A and matrix B, $G(A \otimes B)$ is the oval region of the product for elements of Gerschgorin disc whose center is $m_1 = min\{a_{ii}, i = 1, 2, \cdots, n\}$ and radius is $1 - m_1$ and Gerschgorin disc whose center is $m_2 = min\{b_{jj}, j = 1, 2, \cdots, m\}$ and radius is $1 - m_2$.

Proof: Let $\lambda(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ and $\lambda(B) = \{\mu_1, \mu_2, \cdots, \mu_m\}$. From theorem 1, we have

$$\lambda(A) \subset \{z : |z - m_1| \leq 1 - m_1\},$$

$$\lambda(B) \subset \{z : |z - m_2| \leq 1 - m_2\}.$$

And since $\lambda(A \otimes B) = \{\lambda_i \mu_j | i = 1, 2, \cdots, n, j = 1, 2, \cdots, m\}$, we get

$$\lambda(A \otimes B) \subset G(A \otimes B) = \{z : |z - m_1| \leq 1 - m_1\} \cdot \{z : |z - m_2| \leq 1 - m_2\}.$$

Therefore, the eigenvalues of tensor product for matrix A and matrix B are located in the oval region $G(A \otimes B)$.

Theorem 4. Suppose $A = (a_{ij})_{n\times n}$ is a row stochastic matrix and $M_i = \underset{i}{max}\{a_{ij}, j = 1, 2, \cdots, n\}$, then

$$\lambda(A) \subset G(A) = \bigcup_{i=1}^{n} \{z : |z - a_{ii}| \leq \sqrt{(n-1)M_i(1 - a_{ii})}\},$$

where $\lambda(A)$ is denoted the whole eigenvalues of matrix A, $G(A)$ is denoted generalized Gerschgorin disc of matrix A.

Proof: Because λ is an arbitrary eigenvalue of matrix $A = (a_{ij})_{n\times n}$ and $X = (x_1, x_2, \cdots, x_n)^T \in R^{n\times 1}$ is the corresponding column eigenvector. For $AX = \lambda X$, we get

$$\sum_{j=1}^{n} a_{mj} x_j = \lambda x_m.$$

So

$$(\lambda - a_{mm}) x_m = \sum_{j=1, j\neq m}^{n} a_{mj} x_j.$$

From Schwarz inequality and trigonal inequality, we have the following result:

$$|\lambda - a_{mm}| = |\frac{\sum_{j\neq m} a_{mj} x_j x_m^*}{|x_m|^2}| \leq \sqrt{\sum_{j\neq m} |a_{mj}|^2} \cdot \sqrt{\sum_{j\neq m} |\frac{x_j}{|x_m|}|^2 |\frac{x_m^*}{|x_m|}|^2} \leq \sqrt{(n-1) \sum_{j\neq m} |a_{mj}|^2} = \sqrt{n-1} R_m,$$

where $R_m = \sqrt{\sum_{j\neq m} |a_{mj}|^2}, m = 1, 2, \cdots, n$. And since

$$R_m = \sqrt{\sum_{j\neq m} |a_{mj}|^2} \leq \sqrt{M_m \sum_{j\neq m} |a_{mj}|} = \sqrt{M_m(1 - a_{mm})}, m = 1, 2, \cdots, n,$$

$$|\lambda - a_{mm}| \leq \sqrt{(n-1)M_m(1 - a_{mm})}$$

holds.

Because λ is an arbitrary eigenvalue of matrix A,

$$\lambda(A) \subset G(A) = \bigcup_{i=1}^{n} \{z : |z - a_{ii}| \leq \sqrt{(n-1)M_i(1 - a_{ii})}\}.$$

The theorem is proven.

4. Eigenvalue Estimate for Generalized Stochastic Matrix

Theorem 5. (Brauer. A., 1964) Suppose $A = (a_{ij})_{n\times n} \in S_I(n)$, a_{ii} and a_{jj} are the most small diagonal elements in A, then

$$\lambda(A) \subset G(A) = \{z : |z - a_{ii}||z - a_{jj}| \leq (s - a_{ii})(s - a_{jj})\},$$

where $\lambda(A)$ is denoted the whole eigenvalues of matrix A, $G(A)$ is denoted Cassini oval region of matrix A.

Theorem 6. Suppose $A = (a_{ij})_{n\times n} \in S_I(n)$ and $B = (a_{ij})_{m\times m} \in S_I(m)$ are row stochastic matrices, then

$$\lambda(A \otimes B) \subset G(A \otimes B) = \{z : |z - a_{ii}||z - a_{jj}| \leq (s - a_{ii})(s - a_{jj})\} \cdot \{z : |z - bii||z - b_{jj}| \leq (s - b_{ii})(s - b_{jj})\},$$

where $\lambda(A \otimes B)$ is denoted the whole eigenvalues of tensor product for matrix A and matrix B, $G(A \otimes B)$ is the oval region of the product for Cassini oval region elements of matrix A and Cassini oval region elements of matrix B.

Proof: The method is same to theorem 3, which is leaven for readers.

Theorem 7. Suppose $A = (a_{ij})_{n\times n} \in S_{III}(n)$ and $m = min\{|a_{ii}|, i = 1, 2, \cdots, n\}$, then

$$\lambda(A) \subset G(A) = \{z : |z - m| \leq s + m\},$$

where $\lambda(A)$ is denoted the whole eigenvalues of matrix A, $G(A)$ is the disc whose center is $m = min\{|a_{ii}|, i = 1, 2, \cdots, n\}$ and radius is $s + m$.

Proof: From Gerschgori disc theorem, we have

$$|\lambda - a_{mm}| \leq P_m = \sum_{j=1, j\neq m}^{n} |a_{mj}| = s - |a_{mm}|.$$

Therefore

$$|\lambda - m| = |\lambda - a_{mm} + a_{mm} - m| \leq |\lambda - a_{mm}| + |a_{mm} - m| \leq s - |a_{mm}| + |a_{mm}| + m = s + m.$$

Because λ is an arbitrary eigenvalue of matrix $A = (a_{ij})_{n\times n}$,

$$\lambda(A) \subset G(A) = \{z : |z - m| \leq 1 + m\}.$$

So the eigenvalues of matrix A are located in the disc whose center is $m = min\{|a_{ii}|, i = 1, 2, \cdots, n\}$ and radius is $s + m$.

Theorem 8. Suppose $A = (a_{ij})_{n\times n} \in S_{III}(n)$, $B = (a_{ij})_{m\times m} \in S_{III}(m)$ and $m_1 = min\{|a_{ii}|, i = 1, 2, \cdots, n\}$, $m_2 = min\{|b_{jj}|, j = 1, 2, \cdots, m\}$, then

$$\lambda(A \otimes B) \subset G(A \otimes B) = \{z : |z - m_1| \leq s + m_1\} \cdot \{z : |z - m_2| \leq s + m_2\},$$

where $\lambda(A \otimes B)$ is denoted the whole eigenvalues of tensor for matrix A and matrix B, $G(A \otimes B)$ is the oval region of the product for elements of disc whose center is $m_1 = min\{|a_{ii}|, i = 1, 2, \cdots, n\}$ and radius is $s + m_1$ and disc whose center is $m_2 = min\{|b_{jj}|, j = 1, 2, \cdots, m\}$ and radius is $s + m_2$.

Proof: The method is same to theorem 3, which is omitted.

Theorem 9. Suppose $A = (a_{ij})_{n\times n} \in S_{III}(n)$, a_{ii} and a_{jj} are the most small module diagonal cross elements in A, then

$$\lambda(A) \subset G(A) = \{z : |z - a_{ii}||z - a_{jj}| \leq (s + |a_{ii}|)(s + |a_{jj}|)\},$$

where $\lambda(A)$ is denoted the whole eigenvalues of matrix A, $G(A)$ is denoted Cassini oval region of matrix A.

Theorem 10. Suppose $A = (a_{ij})_{n\times n} \in S_I(n)$ and $B = (a_{ij})_{m\times m} \in S_I(m)$ are row stochastic matrices, then

$$\lambda(A \otimes B) \subset G(A \otimes B) = \{z : |z - a_{ii}||z - a_{jj}| \leq (s + |a_{ii}|)(s + |a_{jj}|)\} \cdot \{z : |z - bii||z - b_{jj}| \leq (s + |b_{ii}|)(s + |b_{jj}|)\},$$

where $\lambda(A \otimes B)$ is denoted the whole eigenvalues of tensor product for matrix A and matrix B, $G(A \otimes B)$ is the oval region of the product for Cassini oval region elements of matrix A and Cassini oval region elements of matrix B.

References

Brauer. A. (1964). On the characteristic roots nonnegative matrices, *Recent Advances in Matrix Theory*. Wiles Press, 3-38.

Gu, Yixi. (1994). The distribution of eigenvalues of matrix. *Acta Mathematics Applicatae Sinica*, (04), 501-511.

Guo, Zhong. (1987). The existence of real double stochastic with positive eigenvalue. *Mathematics in practice and theory*, (01).

Huang, Tingzhu, Yang, Chuansheng. (2007). *Special Matrix Analysis and Applications*. Beijing: Science Press.

Li, Jiansheng. (1981). Roots of real eigen equation for double stochastic matrix. *Mathematics in practice and theory*,(02).

Li, Jinyu. (2006). Inequalities of the Spectral Radius of Kronecker Products for Random Matrices. *College Mathematics*, (02), 85-88.

Lin, Lvduan. (1984). Estimates of determinant and eigvenvalues for two types matrices of row modules equaling. *Mathematics in practice and theory*, (04).

Shen, Guangxing. (1991). The Distribution of Eigenvalue of Generalized Stochastic Matrix. *Chinese Journal of Engineering Mathematics*, (03), 111-115.

Tong, Wenting. (1977). Eigenvalue distribution on several types of matrices. *Acta Mathematics Sinica*, (04).

Zhang, Zhinan. (2001). The Jordan Cannonical Form of a Real Random Matrix. *Numerical Mathematics A Journal of Chinese Universities*, (04), 363-367.