数理解析研究所講究録
1145 巻 2000 年 105-112

Sufficient Conditions for the Second Largest Characteristic Value of a Non-Negative Matrix

Kazuo Kishimoto (岸本 一男)
Institute of Policy and Planning Sciences, University of Tsukuba
1-1-1 Tenno-dai, Tsukuba, Ibaraki 305-8573 JAPAN
e-mail address: kishimot@sk.tsukuba.ac.jp

Abstract

It is shown that if a matrix with real components maps any "monotone non-decreasing vector" to a "monotone increasing vector," the matrix has a "monotone increasing characteristic vector," and the modulus of corresponding characteristic value is the second largest among the moduli of characteristic values of the matrix. This proposition is proved as a corollary to more general propositions. Some other corollaries to these general propositions are also remarked.

1 Introduction

Let $A = (a_{ij})$ be an $n \times n$ matrix with real components. Let us call a characteristic value λ the *mth characteristic value* of A if $|\lambda|$ is the mth largest among the moduli of the characteristic values of A. In the analysis of asymptotic behavior of $A^k \boldsymbol{u}_0$ as $k \to \infty$ for an intial vector $\boldsymbol{u}_0$, the first characteristic value plays the main role.

In the case of stochastic matrix, i.e., if $a_{ij} \geq 0$ $(i, j = 1, \ldots n)$ and $\sum_{j=1}^{n} a_{ij} = 1$ $(i = 1, 2, \ldots n)$ hold, $\lambda_1 = 1$ is always the first characteristic value. Under suitable additional assumptions (for details, see, Lemma 5), the corresponding *left* characteristic vector $\boldsymbol{\psi}_1^{\mathrm{T}} = (\psi_{11}, \psi_{12}, \ldots, \psi_{1n})$ becomes the unique asymptotic stationary distribution of the finite state Markov chain with transition probability matrix A if $\sum_{j=1}^{n} \psi_{1j}$ is normalized to one. Here, $^{\mathrm{T}}$ indicates the transposition of a vector. In this case our next concern will be the second characteristic value, which characterizes the rate of convergence of $A^k \boldsymbol{u}_0$ to $\boldsymbol{\psi_1}^{\mathrm{T}}$ as $k \to \infty$. So long as the author knows, however, not enough efforts have been made at getting necessary *or* sufficient conditions for the second characteristic value. (For a few related results, see, [1], [2].)

This note gives, in Propositions 1 and 2, simple sufficient conditions for finding the second characteristic values for a nonnegative matrix. We will prove them as corollaries to more general propositions (Propositions 6 and 7).

2 Simple Practical Forms

We say that a matrix A is a *nonnegative* (resp. *positive*) *matrix* if all components of A are nonnegative (resp. positive). Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)^{\mathrm{T}}$ be a vector with real components. Let us call $\boldsymbol{u}$ a *horizontal vector* (resp. a *non-horizontal vector*) if $u_1 = u_2 = \ldots = u_n$ hold (resp. do not hold). We say that $\boldsymbol{u}$ is a *monotone non-decreasing* (resp. *monotone increasing*) *vector* if $i > j$ implies $u_i \geq u_j$ (resp. if $i > j$ implies $u_i > u_j$). For two vectors $\boldsymbol{u}$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)^{\mathrm{T}}$ $(\neq \boldsymbol{u})$ with real components, we write $\boldsymbol{u} \geq \boldsymbol{v}$ (resp. $\boldsymbol{u} > \boldsymbol{v}$) if $u_j \geq v_j$ (resp. $u_j > v_j$) $(j = 1, 2, \ldots, n)$ hold.

The following proposition is a simplest practical form of our results:

Proposition 1 *Let A be a nonnegative matrix. Suppose that there exists a natural number r such that A^r maps any non-horizontal monotone non-decreasing vector to a monotone increasing vector. Then, there exist a positive number ρ and a stochastic matrix S such that $A = \rho S$. The matrix A has a monotone increasing right characteristic vector $\boldsymbol{\phi}_2$, which corresponds to a real characteristic value λ_2 satisfying the following relations:*

$$\rho \equiv \lambda_1 > \lambda_2 > |\lambda_3| \geq \ldots \geq |\lambda_n|. \tag{1}$$

Here, λ_1 is the real characteristic value to which a horizontal right characteristic vector corresponds.

As we will see in section 4, we can use the following proposition to check the premise in this proposition:

Proposition 2 *An $n \times n$ matrix $F = (f_{ij})$ maps any non-horizontal monotone non-decreasing vector to a monotone increasing vector if and only if $g < h$ implies the following inequalities:*

$$\sum_{j=k}^{n} f_{gj} < \sum_{j=k}^{n} f_{hj}, \quad k = 2, 3, \ldots, n. \tag{2}$$

An illustrative numerical example of these propositions are as follows:

Example 3 *A stochastic matrix*

$$F = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{3}{20} & \frac{13}{60} & \frac{13}{60} & \frac{3}{20} & \frac{4}{15} \\ \frac{7}{60} & \frac{7}{30} & \frac{1}{6} & \frac{1}{6} & \frac{19}{60} \\ \frac{1}{10} & \frac{11}{60} & \frac{11}{60} & \frac{11}{60} & \frac{7}{20} \\ \frac{1}{30} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{11}{30} \end{pmatrix}$$

satisfies inequalities (2). We can check that F maps any non-horizontal monotone non-decreasing (column) vector to a monotone increasing vector, and $(-6, -\frac{7}{2}, -1, \frac{3}{2}, 4)^{\mathrm{T}}$ is the (right) characteristic vector corresponding to the characteristic value $\frac{1}{6}$. We see that it is the second largest among the characteristic values $1, \frac{1}{6}, \frac{\sqrt{2}}{30}, -\frac{1}{30}, -\frac{\sqrt{2}}{30}$.

From here on, we always deal with right characteristic vectors when we consider characteristic vectors, so that we simply say "characteristic vectors," suppressing *right.*

3 General Framework

Let us consider characteristic values of an $n \times n$ matrix A. We deal with the case where we can choose an $n \times n$ nonsingular matrix P such that

$$A = P^{-1}BP \tag{3}$$

and

$$B = \left(\begin{array}{ccc|ccc} b_{11} & \cdots & b_{1m} & b_{1,m+1} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \cdots \\ b_{m1} & \cdots & b_{mm} & b_{m,m+1} & \cdots & b_{mn} \\ \hline 0 & \cdots & 0 & b_{m+1,m+1} & \cdots & b_{m+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{n,m+1} & \cdots & b_{nn} \end{array}\right) = \left(\begin{array}{c|c} \tilde{B}_{11} & \tilde{B}_{12} \\ \hline \tilde{O}_{21} & \tilde{B}_{22} \end{array}\right) \tag{4}$$

hold for some m. In this case, we define submatrices $\tilde{P}_1$ and $\tilde{P}_2$ of P as follows:

$$P = \left(\begin{array}{ccc} p_{1,1} & \cdots & p_{1n} \\ \vdots & \ddots & \cdots \\ p_{m,1} & \cdots & p_{mn} \\ \hline p_{m+1,1} & \cdots & p_{m+1,n} \\ \vdots & \ddots & \cdots \\ p_{n,1} & \cdots & p_{nn} \end{array}\right) = \left(\begin{array}{c} \tilde{P}_1 \\ \hline \tilde{P}_2 \end{array}\right). \tag{5}$$

Let us begin with the following elementary observation:

Lemma 4 *If an $n \times n$ matrix A is of the form (3) and (4), the characteristic values of A are the union of characteristic values of $\tilde{B}_{11}$ and $\tilde{B}_{22}$, where $\tilde{B}_{11}$ and $\tilde{B}_{22}$ are defined in (4).*

If $\tilde{B}_{22}$ is irreducible and satisfies

$$b_{ij} \geq 0 \quad i, j = m+1, m+2, \ldots, n, \tag{6}$$

we can use the Perron-Frobenius theorem for identifying the first characteristic value(s) of $\tilde{B}_{22}$. The following expression is taken from [3, p.53]:

Lemma 5 (Perron-Frobenius) *An irreducible (non-zero) nonnegative matrix $A = (a_{ij})$ always has a positive characteristic value ρ that is a simple root of the characteristic equation. The moduli of all the other characteristic values do not exceed ρ. To the maximal characteristic value ρ there corresponds a characteristic vector with positive coordinates. Moreover, if A has h characteristic values $\lambda_1 = \rho, \lambda_2, \ldots, \lambda_h$ of modulus ρ, then these numbers are all distinct and are roots of the equation*

$$\lambda^h - \rho^h = 0. \tag{7}$$

More generally: The whole spectrum $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A, regarded as a system of points in the complex λ-plane, goes over into itself under a rotation of the plane by the angle $\frac{2\pi}{h}$. If $h > 1$, then A can be put by means of a permutation into the following 'cyclic' form:

$$A = \begin{pmatrix} O & A_{12} & O & \ldots & O \\ O & O & A_{23} & \ldots & O \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ O & O & O & \ldots & A_{h-1,h} \\ A_{h1} & O & O & \ldots & O \end{pmatrix}, \tag{8}$$

where there are square blocks along the main diagonal.

By expressing the non-negativity and irreducibility of $\tilde{B}_{22}$ in (4) in terms of A, we have the following proposition:

Proposition 6 *Let $\tilde{P}_2$ be an $(n-m) \times n$ matrix defined by (5) in terms of a nonsingular matrix P. Suppose that an $n \times n$ matrix A satisfies the following conditions for any n-dimensional real vector $\boldsymbol{u}$:*

1. $\tilde{P}_2\boldsymbol{u} \geq 0$ *implies* $\tilde{P}_2A\boldsymbol{u} \geq 0$.

2. *Suppose that* $\tilde{P}_2\boldsymbol{u} \geq 0$ *and* $\tilde{P}_2\boldsymbol{u} \neq 0$ *hold. Then, for any* j $(j = 1, 2, \ldots, n)$*, there is a natural number* r *(depending on* j*) such that the* j*th component of* $\tilde{P}_2A^r\boldsymbol{u}$ *is positive.*

Then, the characteristic values of A *are composed of the following three types.*

1. *There are* m *characteristic values* λ_j $(j = 1, 2, \ldots, m)$*, any vector* $\boldsymbol{u}$ *in the corresponding root subspaces of which satisfies the following:*

$$\tilde{P}_2\boldsymbol{u} = 0. \tag{9}$$

2. *There are* h *characteristic values* $\lambda_j = \exp\left(2\pi i\frac{j-m-1}{h}\right)\lambda_{m+1}$ $(j = m + 1, m+2, \ldots, m+h)$*, where* λ_{m+1} *is a real positive and the characteristic vector* $\boldsymbol{\phi}_{m+1}$ *corresponding to* λ_{m+1} *satisfies* $\tilde{P}_2\boldsymbol{\phi}_{m+1} > 0$*. If* $h > 1$*, then* PAP^{-1} *can be put by means of a permutation into the following 'cyclic' form:*

$$PAP^{-1} = \begin{pmatrix} O & B_{12} & O & \ldots & O \\ O & O & B_{23} & \ldots & O \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ O & O & O & \ldots & B_{h-1,h} \\ B_{h1} & O & O & \ldots & O \end{pmatrix}, \tag{10}$$

   where there are square blocks along the main diagonal.

3. *There are* $n - m - h$ *characteristic values* λ_j $(j = m + h + 1, m + h + 2, \ldots, n)$*, which satisfy* $|\lambda_j| < \lambda_{m+1}$.

In the case of $m = 1$, the nonsingularity of P assures that equation (9) has just one nontrivial solution up to constant factor, which we denote by $\boldsymbol{\phi}_1$. The characteristic value λ_1 is the component of 1×1 matrix $\tilde{B}_{11}$. In this sense, fixing an appropriate $\tilde{P}_2$ is equivalent to finding a characteristic vector of A if $m = 1$.

If A itself is an irreducible nonnegative matrix, and $\boldsymbol{\phi}_1 > 0$ holds, Lemma 5 assures that λ_1 in Proposition 6 is the first characteristic value of A. In this case, the first characteristic value of $\tilde{B}_{22}$ is just the second characteristic value of A:

Proposition 7 *In addition to the same assumptions as in Theorem 6, we further assume the followings:*

1. $m = 1$.

2. $a_{ij} \geq 0$ $(i, j = 1, 2, \ldots, n)$.

3. Equation (9) has a nontrivial solution $\boldsymbol{\phi}_1$ *satisfying* $\boldsymbol{\phi}_1 > 0$.

Then, $\lambda_{m+1} (= \lambda_2)$ *defined in Proposition 6 is the second characteristic value.*

4 Propositions 1, 2 and Other Corollaries

We will have any number of corollaries to Propositions 6 and 7 by taking suitable P in (3). A nonnegative matrix A is a stochastic matrix multiplied by a constant factor $c(> 0)$ if and only if $\boldsymbol{\phi}_1 = (1, 1, \ldots, 1)^{\mathrm{T}}$ is a characteristic vector corresponding to a characteristic value c. In this case, we can take any nonsigular matrix P satisfying

$$\sum_{j=1}^{n} p_{ij} = 0, \quad i = 2, 3, \ldots, n, \tag{11}$$

for obtaining a sufficient condition for the second characteristic value.

As a simplest example, let us take P such that each row except for the first row contains just one "1" and one "-1" as non-zero components, one of which occupies the diagonal location. We see that (11) is naturally satisfied. The condition $\tilde{P}_2 \boldsymbol{u} \geq 0$ defines a partial ordering among the components of $\boldsymbol{u}$. Premise 1 in Proposition 6 means that this partial ordering is preserved through the linear transformation defined by A. Let us consider the following $(n-1) \times n$ matrix:

$$\Delta_n \equiv \begin{pmatrix} -1 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & -1 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -1 & 1 & 0 \\ 0 & \ldots & \ldots & 0 & -1 & 1 \end{pmatrix}.$$

Now, we are ready to prove Propositions 1 and 2 as corollaries to Propositions 6 and 7.

(*Proof of Proposition 1*)
For $\tilde{P}_2 = \Delta_n$, premise 3 in Proposition 7 is assured by setting $\boldsymbol{\phi}_1 = (1, 1, \ldots, 1)^{\mathrm{T}}$. From the assumption of the existence of r satisfying $PA^r\boldsymbol{u} > 0$ for any non-horizontal vector $\boldsymbol{u}$, we see that $B^r = PA^rP^{-1}$ is a positive matrix. Thus, we have $h = 1$ from Lemma 5. Proposition 1 follows from Proposition 7. □

(Proof of Proposition 2)
Setting $\tilde{P}_2 = \Delta_n$ and $h = 1$, Proposition 2 follows from the non-negativity of PAP^{-1}. □

For other possible corollaries, if we take $(n-1) \times n$ submatrix $\tilde{P}_2$ as

$$\tilde{P}_2 = \begin{pmatrix} -1 & 1 & 0 & \dots & \dots & 0 \\ -1 & 0 & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ -1 & 0 & \dots & 0 & 1 & 0 \\ -1 & 0 & \dots & 0 & 0 & 1 \end{pmatrix},$$

we obtain another sufficient condition for the second characteristic value λ_2 of A. In this case, any n-dimensional vector with the minimum first component is mapped to a vector with the minimum first component. We need the following proposition instead of Proposition 2:

Proposition 8 *An $n \times n$ matrix $F = (f_{ij})$ maps any n-dimensional vector with the minimum first component to a vector with the minimum first component if and only if the following inequalities hold:*

$$f_{gk} < f_{hk}, \quad k = 2, 3, \dots, n. \tag{12}$$

The characteristic vector ϕ_2 corresponding to λ_2 has the minimum first component.

If $m \geq 2$, we must be satisfied with weaker results. Propositione 6 assures that the first characteristic value of B_{22} is equal to or larger than the $(m+1)$th characteristic value of A. For example, let us take an $(n-m) \times n$ matrix $\tilde{P}_2 = \Delta_{n-m+1}\Delta_{n-m+2}\dots\Delta_n$. In the case of $m = 2$, in particular, we may interpret the property $\tilde{P}_2\phi_2 > 0$ as the "convexity" of ϕ_2. We see that the characteristic value to which ϕ_2 corresponds is equal to or larger than the third characteristic value.

5 Conclusion

We have given sufficient conditions for the second characteristic value of a nonnegative matrix (Propositions 1, 2 and remarks in section 4), which we have proved as corollaries to more general propositions (Propositions 6 and 7). Some other applications of these general propositions are also remarked.

We would like to remark that we can generalize these results to those in more abstract spaces by similar but slightly more careful reasoning. It will be discussed it in anoter paper with application to a special type of time series model.

6 Acknowledgement

The author is partially supported by the Grant-in-Aid for the Scientific Research of the Ministry of Education, Science, Sports and Culture of the Government of Japan and by the Statistical Data Bank Project Research Fund of the Ministry of Education, Science, Sports and Culture of the Government of Japan.

References

[1] Friedland, S. and Gurvits, L.: Upper bound for the real parat of nonmaximal eigenvalues of nonnegative irrducible matarices, *SIAM J. Matrix Anal. Appl.* 15:1015-1017(1994)

[2] Friedland, S. and Nabben, R.: On the second real eigenvalue of nonnegative and Z-matrices, *Linear Algebra and Its Appl.*, 255:303-313(1997).

[3] Gantmacher, F.R.: *Matrix Theory, Vol. 2*, Chelsea: New York, 1957.