No. 808

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

by

Kazuo Kishimoto

February 1999

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

Kazuo Kishimoto

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 vector with "non-zero non-negative m-th difference" to a vector with the "positive m-th difference," the matrix has a characteristic vector with the "positive m-th difference." The corresponding characteristic value is positive, and is equal to or larger than the m-th largest modulus of the characteristic values of the matrix. If $m=2$ and the matrix is non-negative, it is just the second largest characteristic value. A simple sufficient condition for this property is also given.

keywords: Perron-Frobenius theorem, positive matrix, monotone increasing, second largest characteristic value

The correspondence should be made to the address specified in the title.

Use the followin fax number if necessary: +81-298-55-3849(office of inst.)

1 Introduction

Let $A = (a_{jk})$ be an $n \times n$ matrix with real components. Let us call a characteristic value λ the *m-th characteristic value* of A if $|\lambda|$ is the m-th largest among the moduli of the characteristic values of A. In the analysis of the asymptotic behavior of A^n as $n \to \infty$, the first characteristic value plays the main role.

In the case of stochastic matrix, i.e., if $a_{jk} \geq 0$ $(j, k = 1, \ldots n)$ and $\sum_{k=1}^{n} a_{jk} = 1$ $(j = 1, 2, \ldots n)$ hold, the first characteristic value λ_1 is always 1, and the corresponding *left* characteristic vector $\psi_1^{\mathrm{T}} = (\psi_{11}, \psi_{12}, \ldots, \psi_{1n})$ becomes the stationary distribution of the finite state Markov chain with the 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 λ_2, which characterizes the rate of convergence to $\psi_1{}^{\mathrm{T}}$ if $|\lambda_2| \neq \lambda_1$. So long as the author knows, however, not enough efforts have been made at getting necessary *or* sufficient conditions for λ_2.

This note gives, in Proposition 2, a simple sufficient condition for finding λ_2 for a non-negative matrix. We will prove it as a corollary to a more general proposition (Proposition 3) for which neither $m = 2$ nor non-negativity of A is assumed. We will also give an illustrutive example (Propostion 4) in which Proposition 2 is conveniently used.

2 Results

The results of this note are based on the Perron-Frobenius theorem (cf., e.g., [1, p.53]):

Lemma 1 (Perron-Frobenius) *An irreducible non-negative matrix* $A = (a_{jk})$ *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 ad are roots of the equation*

$$\lambda^h - \rho^h = 0. \tag{1}$$

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:*

$$A = \begin{pmatrix} O & A_{12} & O & \ldots & O \\ O & O & A_{23} & \ldots & O \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ O & O & O & \ldots & A_{h-1,h} \\ A_{h1} & O & O & \ldots & O \end{pmatrix}, \tag{2}$$

where there are square blocks along the main diagonal.

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

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

Proposition 2 *Let A be a non-negative 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 ϕ_2, and its corresponding characteristic value λ_2 satisfies the following relation:*

$$\rho \equiv \lambda_1 > \lambda_2 > |\lambda_3| \geq \cdots \geq |\lambda_n|, \tag{3}$$

where λ_1 is the characteristic value corresponding to a positive horizontal right characteristic vector.

Notice that the irreducibility of A is not assumed. This note proves this proposition as a corollary to a more general proposition. From here on, we always deal with right characteristic vectors when we consider characteristic vectors, so that we simply say "characteristic vectors," suppressing *right*.

Let us call a linear map Δ a *difference map* if Δ maps an n-dimensional vector $\boldsymbol{u}$ to $(n-1)-$ dimensional vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_{n-1})^{\mathrm{T}} = (u_2 - u_1, u_3 - u_2, \ldots, u_n - u_{n-1})$. If we apply Δ m times $(m < n)$, we have Δ^m, which maps an n-dimensional vector to an $(n-m)-$ dimensional vector. We have the following proposition:

Proposition 3 *Suppose that there is an integer m for which a matrix A satisfies the following conditions:*

1. $\Delta^m \boldsymbol{u} \geq 0$ *implies* $\Delta^m(A\boldsymbol{u}) \geq 0$.

2. *Suppose that* $\Delta^m \boldsymbol{u} \neq 0$ *holds. 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* $\Delta^m(A^r \boldsymbol{u})$ *is positive.*

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

1. m *characteristic values* λ_j $(j = 1, 2, \ldots, m)$, *any vector* $\boldsymbol{u}$ *in the corresponding root subspaces of which satisfies* $\Delta^m \boldsymbol{u} = 0$.

2. 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* $\lambda_{m+1} > 0$ *and the characteristic vector* $\Delta^m \phi_{m+1}$ *corresponding to* λ_{m+1} *satisfies* $\Delta^m \phi_{m+1} > 0$.

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

Proof. Define a matrix P by

$$P = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -1 & 1 & \ddots & & & \vdots \\ 0 & -1 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -1 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

Since we have

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

P is non-singular. Notice that $u \geq 0$ implies $\Delta^m(Au) \geq 0$ if and only if A is of the form:

$$A = (P^{-1})^m B P^m, \tag{4}$$

where

$$B = \left(\begin{array}{ccc|ccc} b_{11} & \dots & b_{1m} & b_{1,m+1} & \dots & b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \dots \\ b_{m1} & \dots & b_{mm} & b_{m,m+1} & \dots & b_{mn} \\ \hline 0 & \dots & 0 & b_{m+1,m+1} & \dots & b_{m+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & b_{n,m+1} & \dots & b_{nn} \end{array}\right) = \left(\begin{array}{c|c} \tilde{B}_{11} & \tilde{B}_{12} \\ \hline \tilde{O}_{21} & \tilde{B}_{22} \end{array}\right) \tag{5}$$

$$b_{jk} \geq 0 \quad (j, k = m+1, m+2, \dots, n). \tag{6}$$

Clearly, A and B have the same characteristic values.

The characteristic values of B are the union of the characteristic values of $\tilde{B}_{11}$ and those of $\tilde{B}_{22}$. Vectors $\tilde{\phi}_k = (\phi_{k1}, \phi_{k2}, \dots, \phi_{km}, 0, 0, \dots, 0)^{\mathrm{T}}$ ($k = 1, 2, \dots, m$) form root subspaces of $\tilde{B}_{22}$ corresponding to $\lambda_1, \lambda_2, \dots, \lambda_m$. Since $\tilde{B}_{22}$ is an irreducible non-negative matrix from the second assumption of in the theorem under consideration, Lemma 1 assures the required results. □

Proposition 2 directly follows from Proposition 3, because $\lambda_1 > |\lambda_k|$ $(k \geq 2)$ follow from Lemma 1 when $m = 1$.

The following proposition is a typical application of Proposition 2:

Proposition 4 *Let (a_{jk}) be an $n \times n$ stochastic matrix. Suppose that $r < m$ implies*

$$\sum_{j=1}^{k} a_{jr} \leq \sum_{j=1}^{k} a_{jm}, \quad k = 1, 2, \ldots, n. \tag{7}$$

Then, (a_{jk}) has an monotone increasing characteristic vector, and its corresponding characteristic values is the second characteristic value.

The following example shows that (7) in Proposition 4 is not a necessary condition:

Example 5 *The matrix*

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

maps any non-horizontal monotone non-decreasing vector to a monotone increasing vector, and $(-6, -\frac{7}{2}, -1, \frac{3}{1}, 4)^{\mathrm{T}}$ is the characteristic vector corresponding to the characteristic value $\frac{1}{6}$. We see that it is the second largest among the characteristic values $1, \frac{1}{6}, \frac{\sqrt{2}}{30}, -\frac{1}{30}, -\frac{\sqrt{2}}{30}$. However, (7) does not hold.

3 Conclusion

We have given a sufficient condition for the second maximality of the characteristic value of a non-negative matrix (Proposition 2), which we have proved as a corollary to a more general proposition (Proposition 3). Shown was an illustrative example how to use it in an actual matrix (Proposition 4).

We would like to remark that we can generalize this result 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.

4 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 Cuture of the Government of Japan.

References

[1] Gantmacher, F.R.: *Matrix Theory, Vols. II*, Chelsea: New York, 1957.