NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET

On Comparison of the Perron-Frobenius Eigenvalues of Two ML-Matrices

by

Bo Henry Lindqvist

PREPRINT
STATISTICS NO. 16/2001

NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
http://www.math.ntnu.no/preprint/statistics/2001/S16-2001.ps
Bo Henry Lindqvist has homepage http://www.math.ntnu.no/~bo/
Email: bo@math.ntnu.no

On Comparison of the Perron-Frobenius Eigenvalues of Two ML-Matrices

Bo Henry Lindqvist

Department of Mathematical Sciences
Norwegian University of Science and Technology
N-7491 Trondheim, Norway

Abstract

A simple inequality relating the Perron-Frobenius eigenvalues of two ML-matrices is obtained by performing a limiting operation on an inequality by Bapat [1].

The problem of bounding the Perron-Frobenius eigenvalue (PF-value) of an $n \times n$ nonnegative matrix B, $r(B)$, is a classical one. Sometimes one is able to approximate B by an $n \times n$ nonnegative matrix A for which the PF-value is known, and it is then of interest to have explicit results for the relation between $r(B)$ and $r(A)$. A nice result of this kind was given by Bapat [1]. The purpose of the present paper is to show how one by an elementary limiting operation on Bapat's result obtains a new relation between $r(B)$ and $r(A)$, which is sometimes sharper than the original one. An application is given to a problem concerning absorbing Markov chains.

Bapat [1] proved by elementary methods that if $r(A) > 0$ and if (x_i) and (y_i) are, respectively, left and right eigenvectors at the eigenvalue $r(A)$ (called PF-vectors) satisfying $\sum_{i=1}^{n} x_i y_i = 1$, then

$$\frac{r(B)}{r(A)} \geq \prod_{i,j=1}^{n} \left(\frac{b_{ij}}{a_{ij}} \right)^{a_{ij} x_i y_j / r(A)} \tag{1}$$

Here we may take $(z/0)^0 = 1$ for any real number z. The inequality (1) gives a lower bound for $r(B)$ and is useful when B can be approximated by a matrix A for which one knows $r(A)$ as well as the PF-vectors (x_i) and (y_i). Note that the property $\sum_{i=1}^{n} x_i y_i = 1$ is possible at least when A is irreducible. We refer to Remark 1 for a further discussion of this property.

To be slightly more general we shall in the following assume that A and B are ML-matrices. Recall from Seneta [5] that a square matrix M is called an ML-matrix if $m_{ij} \geq 0$ for all $i \neq j$. We denote by $r(M)$ the PF-value, which by [5] has the property that $r(M) > \mathrm{Re}(\lambda)$ for any other eigenvalue of M and coincides with the PF-value for nonnegative matrices.

Let $c = \min\{a_{ii}, b_{ii} : i = 1, \dots n\}$, i.e. c is the minimal diagonal element of A and B. Then $A + sI$ and $B + sI$ are nonnegative for $s \geq -c$, and clearly

$$\begin{aligned} r(B + sI) &= r(B) + s \\ r(A + sI) &= r(A) + s \end{aligned}$$

Let (x_i) and (y_i) be left and right, respectively, PF-vectors of $A + sI$ for $s \geq -c$ (these are independent of s and will be called PF-vectors of A). Assume, furthermore, that we have $\sum_{i=1}^{n} x_i y_i = 1$.

Applying inequality (1) to the matrices $B + sI$ and $A + sI$ for $s \geq -c$ (or $s > -c$ if r(A)=c), we get

$$\begin{aligned} r(B) \geq & [r(A) + s] \prod_{i=1}^{n} \left(\frac{b_{ii} + s}{a_{ii} + s} \right)^{(a_{ii}+s)x_i y_i/[r(A)+s]} \prod_{i \neq j} \left(\frac{b_{ij}}{a_{ij}} \right)^{a_{ij} x_i y_j/[r(A)+s]} - s \\ \equiv & \gamma(s) \end{aligned} \tag{2}$$

Thus we obtain a family of lower bounds $\gamma(s)$ for $r(B)$ by means of varying s. Bapat's bound (1) is $\gamma(0)$, valid for nonnegative matrices A and B.

In particular, by letting $s \to \infty$ in the inequality $r(B) \geq \gamma(s)$ we get:

Theorem 1 *Let A and B be ML-matrices of dimension $n \times n$, where A admits left and right PF-vectors $(x_i), (y_i)$, respectively, with $\sum_{i=1}^{n} x_i y_i = 1$. Suppose further that $b_{ij} = 0$ implies $a_{ij} = 0$ when $x_i y_j > 0$. Then*

$$r(B) \geq r(A) + \sum_{i=1}^{n} x_i y_i (b_{ii} - a_{ii}) + \sum_{i \neq j} a_{ij} x_i y_j \ln \frac{b_{ij}}{a_{ij}} \quad (\equiv \gamma(\infty)) \tag{3}$$

where $0/0 \equiv 1, 0 \ln \infty = 0$.

The example below shows that none of the two inequalities (1) and (3) is generally better than the other for nonnegative matrices A and B.

Example

Let

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & b \\ 4 & 3 \end{pmatrix}$$

where $b > 0$. Now $r(A) = 4$ and we may take $x_1 = x_2 = \frac{1}{2}$, $y_1 = y_2 = 1$. Direct substitution in (1) and (3) show that Bapat's inequality (1) gives the best lower bound for $r(B)$ if $b \geq 4.74$ or $b \leq .247$, while (3) is the best one for b in the interval $(.247, 4.74)$. There exist, however, values of s for which the lower bound $\gamma(s)$ is better than both $\gamma(0)$ and $\gamma(\infty)$. As an example, take $b = 6$. Then $\gamma(0) = 7.2288$ and $\gamma(\infty) = 6.9713$, while the maximum value of $\gamma(s)$ is 7.3667 which occurs for $s = 1.47$. This is close to the correct value 7.4244 for $r(B)$ when $b = 6$.

The following proposition provides two simple criteria for comparing the lower bounds $\gamma(s)$ in particular cases. Part (i) implies that if A and B coincide on the diagonal, then the best lower bound of $r(B)$ is obtained by (1), while part (ii) implies that if A and B are equal outside the diagonal, then Theorem 1 gives the best bound.

Proposition 1 *Let A and B be as in Theorem 1 .*

(i) If $b_{ii} = a_{ii}$ for all i, then $\gamma(s)$ is either constant or strictly decreasing (and convex) in s for $s > -c$.

(ii) If $b_{ij} = a_{ij}$ for all $i \neq j$, then $\gamma(\infty) > \gamma(s)$ for all $s > -c$, provided $\sum_{i \neq j} a_{ij} x_i y_j > 0$ (which holds in particular if A is irreducible) and $b_{ii} \neq a_{ii}$ for at least one i.

Proof: To prove (i), note that in this case we can write

$$\gamma(s) = [r(A) + s] k^{1/[r(A)+s]} - s$$

where

$$k = \prod_{i \neq j} \left(\frac{b_{ij}}{a_{ij}} \right)^{a_{ij} x_i y_j}$$

is independent of s. If $k = 1$, then $\gamma(s) = r(A)$ for all s. Suppose therefore $k \neq 1$. Differentiating γ twice it is seen that $\gamma''(s) > 0$ and $\lim_{s \to \infty} \gamma'(s) = 0$. This proves (i).

To prove (ii) we use the generalized arithmetic mean-geometric-mean inequality, which states that for $\alpha_k \geq 0$, $w_k \geq 0$, $\sum_k w_k = 1$, we have $\prod_k \alpha_k^{w_k} \leq \sum_k w_k \alpha_k$, with equality if and only if the α_k are all equal.

Using the fact that $\sum_{i,j} a_{ij} x_i y_j = r(A)$, we get

$$\begin{aligned}
\gamma(s) &= [r(A)+s] \prod_{i=1}^{n} \left(\frac{b_{ii}+s}{a_{ii}+s} \right)^{(a_{ii}+s)x_i y_i/[r(A)+s]} \prod_{i \neq j} 1^{a_{ij}x_i y_j/[r(A)+s]} - s \\
&\leq [r(A)+s] \left(\sum_{i=1}^{n} \frac{(a_{ii}+s)x_i y_i}{r(A)+s} \cdot \frac{b_{ii}+s}{a_{ii}+s} + \sum_{i \neq j} \frac{a_{ij} x_i y_j}{r(A)+s} \right) - s \\
&= \sum_{i=1}^{n} x_i y_i (b_{ii}+s) + \sum_{i \neq j} a_{ij} x_i y_j - s \\
&= \gamma(\infty)
\end{aligned}$$

If the exponents of the $1's$ are not all 0, then equality holds for any particular s if and only if $b_{ii} = a_{ii}$ for all i. This completes the proof. □

Application: Absorbing Markov chain

A continuous time Markov chain $\{X(t);\ t \geq 0\}$ with state space $\{1, 2, \ldots, n+1\}$ determines an ML-matrix Q of instantaneous transition rates, with row sums all equal to 0 (Ross [4, p. 326]). Suppose that state $n+1$ is absorbing. Then $q_{n+1,j} = 0$ for $j = 1, 2, \ldots, n+1$ and hence we can write

$$Q = \begin{pmatrix} B & r \\ 0 & 0 \end{pmatrix}$$

where $r \neq 0$ is a nonnegative column vector of dimension n and B is an $n \times n$ ML-matrix. We assume for simplicity that B is irreducible, which in the Markov chain terminology means that the states in $S \equiv \{1, 2, \ldots, n\}$ intercommunicate. Suppose that the Markov chain is started in state $i \in S$ with probability μ_i and let $\mu = (\mu_i)$. Then if $T = \inf\{t > 0 : X(t) = n+1\}$ is the absorption time we might be interested in the probability $P_\mu(T > t)$ that absorption takes place after time t. By [4, p. 350] we can write

$$P_\mu(T > t) = \mu \exp(tB) 1 \tag{4}$$

where 1 is a column vector of $1's$. Let the PF-value of B be $-b$ $(b > 0)$ and let $\rho = (\rho_i)$ be the left PF-vector of B with $\sum_{i=1}^{n} \rho_i = 1$. Then it follows from (4) that

$$P_\rho(T > t) = e^{-bt}$$

In Markov chain theory ρ is called the quasi-stationary distribution of the Markov chain. It is the limiting probability distribution of $X(t)$ as $t \to \infty$, conditionally on the event that absorption has not yet taken place at time t (Keilson [2]).

Lower bounds for the PF-value $-b$ of B can be found by approximating B by ML-matrices A with known PF-value and PF-vectors, and then using for example the bounds $\gamma(s)$ of (2). In the present application a natural candidate for A is the matrix $A = B + \text{diag}(r)$, where $\text{diag}(r)$ is the $n \times n$ diagonal matrix with diagonal r. Now A is an ML-matrix with all row sums equal to 0, and hence $r(A) = 0$. Moreover, $a_{ij} = b_{ij}$ for all $i \neq j$ so Proposition 1 (ii) applies and suggests the lower bound of Theorem 1. Note that A corresponds to an irreducible Markov chain on S. Hence (x_i) can be taken to be its stationary distribution while (y_i) is a vector of 1's.

In applications to reliability of systems, the Markov chain $X(t)$ is typically stochastically monotone ([2]), with state "1" corresponding to a perfectly working system. Then, letting the lower bound from Theorem 1 be denoted $-a$ (> 0), we obtain the inequality $P_1(T > t) \geq P_\rho(T > t) \geq e^{-at}$ which gives an often useful and easily computable lower bound on system reliability.

Remark 1

We return to the assumption that for A there exist left and right PF-vectors (x_i) and (y_i), respectively, with $\sum_i x_i y_i = 1$. As noted already, the assumption holds whenever A is irreducible. The following simple example shows, however, that it may not hold when A is reducible, even if $r(A) > 0$. Let namely

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Then $r(A) = 1$, while any left PF-vector is of the form $(x, 0)$ for $x > 0$ and any right PF-vector is $(0, y)$ for $y > 0$.

In order to give a necessary and sufficient condition for the existence of PF-vectors as required, we need to consider the structure of a general (reducible) nonnegative matrix A. Without loss of generality we can assume that A is written on so-called Frobenius normal form

$$A = \begin{pmatrix} A_{11} & 0 & 0 & \cdots & 0 \\ A_{21} & A_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ A_{s1} & A_{s2} & A_{s3} & \cdots & A_{ss} \end{pmatrix}$$

where $A_{11}, A_{22}, \ldots, A_{ss}$ are irreducible with spectral radii, respectively, $r_1, \ldots, r_s$. Now $r(A) = \max_i r_i$. For $j, k \in \{1, 2, \ldots, s\}$, $j > k$, we write $j \to k$ if there are indices $j \equiv k_1 > k_2 > \cdots > k_p \equiv k$ with $A_{k_i k_{i+1}} \neq 0$ for $i = 1, \ldots, p-1$. Then it can be shown from Theorem 6.1 of Lindqvist [3] that PF-vectors with the desired property exist if and only if for at least one block A_{jj} with $r_j = r(A)$ we have $r_k < r(A)$ for all k such that $j \to k$ or $k \to j$.

Remark 2

Song [6] presents a class of lower bounds on the PF-value which turn out to be consequences of (1) and in fact can be improved using (1). Let B be as before and let (x_i) be a nonnegative n-dimensional vector with $\sum_{i=1}^n x_i = 1$. Then let A be the matrix with n rows all equal to (x_i). It is clear that $r(A) = 1$ and that (x_i) is a left PF-vector. Also, $(y_i) = (1, 1, \ldots, 1)$ is a right PF-vector and we have $\sum_{i=1}^n x_i y_i = 1$. Thus (1) gives

$$r(B) \geq \prod_{i,j=1}^{n} \left(\frac{b_{ij}}{x_j} \right)^{x_i x_j} \tag{5}$$

where as before we can take $(z/0)^0 = 1$.

Now by the generalized arithmetic mean-geometric-mean inequality we have

$$\prod_{i,j=1}^{n} x_j^{x_i x_j} \leq \sum_{i,j=1}^{n} x_i x_j^2 = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} x_j^2 \right) = \sum_{j=1}^{n} x_j^2$$

so that (5) implies

$$r(B) \geq \frac{\prod_{i,j=1}^{n} b_{ij}^{x_i x_j}}{\sum_{j=1}^{n} x_j^2} \tag{6}$$

which is Theorem 1 of Song [6].

References

[1] R.B. Bapat, Comparing the spectral radii of two nonnegative matrices, Amer. Math. Monthly 96:137-139 (1989).

[2] J. Keilson, Markov chain models – rarity and exponentiality, Springer, New York (1979).

[3] B. H. Lindqvist, Asymptotic properties of powers of nonnegative matrices, with applications, Linear Algebra Appl. 114:555-588 (1989).

[4] S.M.Ross, Introduction to probability models (7th edition), Harcourt/Academic Press, San Diego, (2000).

[5] E. Seneta, Nonnegative Matrices, Allen & Unwin, London, 1973.

[6] Y. Song, Lower bounds for the Perron root of a nonnegative matrix, Linear Algebra Appl. 169:269-278 (1992).