NORTH-HOLLAND

A Note on the Coefficient of Ergodicity of a Column-Allowable Nonnegative Matrix

Marc Artzrouni* and Xuefeng Li
Department of Mathematics and Computer Science
Loyola University
New Orleans, Louisiana 70118

Submitted by Richard A. Brualdi

ABSTRACT

We give a simple proof of a closed-form expression for the coefficient of ergodicity of a column-allowable nonnegative matrix.

1. INTRODUCTION

The coefficient of ergodicity of a column-allowable nonnegative $n \times n$ matrix $T = (t_{ij})$ is defined as

$$\tau(T) = \sup_{X, Y > 0;\ X \neq \lambda Y} \frac{d(X'T, Y'T)}{d(X', Y')}, \tag{1.1}$$

where $d(X', Y')$ is the projective distance between the vectors $X = (x_i)$ and $Y = (y_i)$ in the positive quadrant. In general uppercase letters will denote vectors or matrices, and the corresponding lowercase letters will represent their entries; X' is the transpose of X, and the projective distance $d(X', Y')$ is defined as

$$\max_{i,j} \ln\left(\frac{x_i/y_i}{x_j/y_j}\right)$$

* Current address: Department of Applied Mathematics, University of Pau, France.

LINEAR ALGEBRA AND ITS APPLICATIONS 214:93–101 (1995)

655 Avenue of the Americas, New York, NY 10010

0024-3795/95/$9.50
SSDI 0024-3795(93)00058-8

(Seneta, 1981). A column-allowable matrix has at least one nonnegative entry in each column. This condition is necessary for (1.1) to be defined.

The following theorem provides a result on the coefficient of ergodicity of a column-allowable matrix T.

THEOREM 1.1 (Seneta, 1981, p. 108). *If T is positive, then*

$$\tau(T) = \frac{1 - \phi^{0.5}}{1 + \phi^{0.5}}, \tag{1.2}$$

where

$$\phi = \frac{1}{\max_{i,i',j,j'} t_{i'j} t_{j'i} / t_{j'j} t_{i'i}}. \tag{1.3}$$

If T has a row with at least one positive and one zero entry, then $\tau(T) = 1$. If the only zeros are in zero rows, then (1.2)–(1.3) still hold, except that the entries t_{ij} appearing in (1.3) must be replaced by the entries a_{ij} of the matrix A formed from T by deleting its zero rows.

It is easily seen that Theorem 1.1 can be paraphrased by saying that for any column-allowable matrix T, the result of (1.2) holds in all cases with

$$\phi = \min_{t_{j'j} t_{i'i} \neq 0} \frac{t_{i'j} t_{j'i}}{t_{j'j} t_{i'i}}. \tag{1.4}$$

In this paper we give a direct and relatively simple proof of this result. In the process we demonstrate a simple connection between $\tau(T)$ and the coefficient of ergodicity $\tau_1(P)$ of a row-stochastic matrix $P = (p_{ij})$ closely related to T. The coefficient $\tau_1(P)$ is defined as

$$\tau_1(P) = 0.5 \max_{i,j} \sum_k |p_{ik} - p_{jk}| = \max_{i,j} \sum_{k \in \Delta} (p_{ik} - p_{jk}),$$
$$\Delta = \{k \mid p_{ik} - p_{jk} > 0\}. \tag{1.5}$$

2. THEOREM

Given a column-allowable nonnegative matrix T, and a positive probability-normed vector $Y = (y_i)$ (i.e., $\sum y_i = 1$), we let $P(Y)$ denote the row-stochastic matrix with entry $P(Y)_{i,j} = t_{ji} y_j / \sum_s y_s t_{si}$ in its ith row, jth column.

THEOREM 2.1. *If $T = (t_{ij})$ is a column-allowable nonnegative matrix, then*

$$\tau(T) = \frac{1 - \phi^{0.5}}{1 + \phi^{0.5}} = \sup_{Y \in \mathfrak{I}} \tau_1[P(Y)],$$

$$\mathfrak{I} = \{Y = (y_i) | \Sigma_i y_i = 1,\ y_i > 0,\ i = 1, 2, \dots n\}, \tag{2.1}$$

where

$$\phi = \min_{t_{j'j} t_{i'i} \neq 0} \frac{t_{i'j} t_{j'i}}{t_{j'j} t_{i'i}}. \tag{2.2}$$

Proof. We first note that because T is column-allowable, ϕ is well defined (i.e., the set of indices i, i', j, j' such that $t_{j'j} t_{i'i} \neq 0$ is nonempty). Bearing in mind the definition (1.1) of $\tau(T)$, we consider for fixed $X = (x_i)$, $Y = (y_i)$ $(X, Y > 0)$ the ratios x_i/y_i, and we let C be the minimum of these ratios. If D is such that $C + D = \max_i x_i/y_i$ [i.e., $D = \max_i x_i/y_i - \min_i x_i/y_i$, and D is called osc(X/Y)], then each ratio x_i/y_i can be written as $x_i/y_i = C + De_i$, where C is the minimum and $C + D$ the maximum of these ratios. The vector $E = (e_i)$ has components between 0 and 1, with at least one of the e_i's equal to 0, and one equal to 1. We then have $x_i = y_i[C + De_i]$, $i = 1, 2, \dots, n$.

The distance $d(X', Y')$ is now simply $\ln[(C + D)/C] = \ln[1 + D/C]$. The ith component of $X'T$ is $\Sigma_k x_k t_{ki}$, and the jth component of $Y'T$ is $\Sigma_k y_k t_{kj}$. Therefore

$$d(X'T, Y'T) = \max_{i,j} \ln \frac{\sum_{k=1}^{n} x_k t_{ki} \Big/ \sum_{k=1}^{n} y_k t_{ki}}{\sum_{k=1}^{n} x_k t_{kj} \Big/ \sum_{k=1}^{n} y_k t_{kj}}$$

$$= \max_{i,j} \ln \frac{1 + (D/C) \sum_{k=1}^{n} e_k y_k t_{ki} \Big/ \sum_{k=1}^{n} y_k t_{ki}}{1 + (D/C) \sum_{k=1}^{n} e_k y_k t_{kj} \Big/ \sum_{k=1}^{n} y_k t_{kj}}. \tag{2.3}$$

We let $A^i = \Sigma_k e_k y_k t_{ki}/\Sigma_s y_s t_{si} = \Sigma_k e_k P(Y)_{i,k}$, $A^j = \Sigma_k e_k y_k t_{kj}/\Sigma_s y_s t_{sj} = \Sigma_k e_k P(Y)_{j,k}$, and $x = D/C$. If we let $G(E, Y, x, i, j)$ be the argument of the logarithm on the right-hand side of Equation (2.3), i.e.,

$$G(E,Y,x,i,j) = \frac{1 + (D/C) \sum_{k=1}^{n} e_k y_k t_{ki} \Big/ \sum_{k=1}^{n} y_k t_{ki}}{1 + (D/C) \sum_{k=1}^{n} e_k y_k t_{kj} \Big/ \sum_{k=1}^{n} y_k t_{kj}} = \frac{1 + xA^i}{1 + xA^j}, \quad (2.4)$$

we then have

$$d(X'T, Y'T) = \max_{i,j} \ln G(E,Y,x,i,j) = \max_{i,j} \ln \frac{1 + xA^i}{1 + xA^j}, \quad (2.5)$$

and therefore

$$\tau(T) = \sup_x \frac{\sup_{E,Y} \max_{i,j} \ln G(E,Y,x,i,j)}{\ln(1+x)} = \sup_x \frac{\sup_{E,Y} \max_{i,j} \ln \dfrac{1 + xA^i}{1 + xA^j}}{\ln(1+x)}, \quad (2.6)$$

where the supremum is for $x > 0$. The components of $E = (e_i)$ are between 0 and 1, with at least one $e_i = 0$ and at least one $e_i = 1$; Y is positive, and without loss of generality can be assumed probability-normed.

We will first dispose of two special cases:

(1) $\phi = 0$. This can occur if and only if there exist i^*, j^*, k^* $(i^* \neq j^*)$ such that $t_{k^*i^*} \neq 0$ and $t_{k^*j^*} = 0$ (i.e., there is a row k^* with one zero and one positive entry). If we let the component y_{k^*} of Y approach 1 (with Y remaining probability-normed and $e_{k^*} = 1$, $e_j = 0$ for $j \neq k^*$), then A^{i^*} approaches 1 and $A^{j^*} = 0$. The corresponding $G(E, Y, x, i^*, j^*)$ approaches $1 + x$ as $y_{k^*} \to 1$, and therefore $\tau(T) = 1$, by consideration of Equation (2.6). For the same vector Y with $y_{k^*} \to 1$, it can also be seen that $\tau_1[P(Y)] \to 1$, which proves that Equation (2.1) holds in this particular case with $\tau(T) = 1$.

(2) $\phi = 1$. For any i, i', j, j' such that $t_{j'j} t_{i'i} \neq 0$, we have $t_{i'j}/t_{i'i} = t_{j'j}/t_{j'i}$, and therefore the nonzero rows of T are proportional. (Any zero is necessarily in a row of zeros, since otherwise ϕ would be 0.) The coefficient

of ergodicity $\tau(T)$ is then trivially 0, and $\tau_1[P(Y)] = 0$, since the rows of $P(Y)$ are all equal.

In the sequel we will assume that $0 < \phi < 1$. The proof will proceed in two steps. First we will show that $\tau(T) \leqslant (1 - \phi^{0.5})/(1 + \phi^{0.5})$. [We define $K = (1 - \phi^{0.5})/(1 + \phi^{0.5})$.] Then we will show that for any $\varepsilon > 0$ there exist E, Y, x, i, and j such that

$$|[\ln G(E, Y, x, i, j)]/\ln(1 + x) - K| \leqslant \varepsilon. \tag{2.7}$$

We first consider the following proposition, which is easily proven by induction:

PROPOSITION 1. *Let $\{u_i\}$ and $\{v_i\}$ be two sets of n positive numbers. Suppose that u_1/v_1 is the smallest of the ratios u_i/v_i. Then $\Sigma u_i/\Sigma v_i \geqslant u_1/v_1$.*

We will now use the proposition to prove that for any i, j

$$A^j \geqslant \frac{A^i\phi}{A^i\phi + (1 - A^i)}. \tag{2.8}$$

Given that A^i and A^j are between 0 and 1, (2.8) is trivially true if $A^i = 0$ or $A^j = 1$. In other cases (2.8) is equivalent to $\phi \leqslant A^j(1 - A^i)/[A^i(1 - A^j)]$. We now let i^*, i'^*, j^*, j'^* be the indices for which $t_{i'j}t_{j'i}/t_{j'j}t_{i'i}$ (with $t_{j'j}t_{i'i} \neq 0$) reaches its minimum, i.e., $\phi = t_{i'^*j^*}t_{j'^*i^*}/t_{j'^*j^*}t_{i'^*i^*} \leqslant t_{i'j}t_{j'i}/t_{j'j}t_{i'i}$ for any i, i', j, j' such that $t_{j'j}t_{i'i} \neq 0$. By virtue of Proposition 1 we have

$$\frac{A^j(1 - A^i)}{A^i(1 - A^j)} = \frac{\left[\sum_{k=1}^{n} e_k t_{kj} y_k\right]\left[\sum_{k=1}^{n} (1 - e_k) t_{ki} y_k\right]}{\left[\sum_{k=1}^{n} e_k t_{ki} y_k\right]\left[\sum_{k=1}^{n} (1 - e_k) t_{kj} y_k\right]}$$

$$\geqslant \min_k \frac{t_{kj}}{t_{ki}} \times \min_{k'} \frac{t_{k'i}}{t_{k'j}} \geqslant \phi, \tag{2.9}$$

the minima are over the ratios for which both the numerator and the denominator are nonzero. If both are zero, the ratio contributes nothing, and the case of only the numerator (or only the denominator) being zero is excluded because it was dealt with earlier (i.e., ϕ is assumed > 0). Now that

(2.9) is established, we have

$$G(E,Y,x,i,j) = \frac{1+xA^i}{1+xA^j} \leqslant \frac{1+xA^i}{1+\dfrac{xA^i\phi}{A^i\phi+(1-A^i)}}$$

$$= \frac{(A^i)^2 x(\phi-1)+A^i(\phi-1+x)+1}{A^i(x\phi+\phi-1)+1}. \quad (2.10)$$

Elementary considerations show that the right-hand side of (2.10), considered as a function $M(A^i)$ of $A^i \in (0,1)$, reaches a maximum $M^* = 1 + (1-\phi)x/[1+(\phi\{1+x\})^{0.5}]^2$ for $A^i = 1/[1+(\phi\{1+x\})^{0.5}]^2$ [we note that $M^* > M(0)$ and $M^* > M(1)$ because $M^* > 1$ and $M(0) = 1 = M(1)$]. Therefore

$$\tau(T) = \sup_{x>0} k(x), \qquad k(x) = \frac{\ln\left[1+\dfrac{(1-\phi)x}{\left[1+(\phi\{1+x\})^{0.5}\right]^2}\right]}{\ln(1+x)}. \quad (2.11)$$

We will now show that $\sup_{x>0} k(x) = (1-\phi^{0.5})/(1+\phi^{0.5})$ (recall that this quantity is denoted K). We first note that $\lim_{x\to 0} k(x) = K$ and $\lim_{x\to\infty}, k(x) = 0$. We prove the result by contradiction and assume that $\sup_{x>0} k(x) > K$. The means that $\exists x_0 > 0$ such that $k(x_0) > K$. Given that $\lim_{x\to 0} k(x) = K$ and $\lim_{x\to\infty} k(x) = 0$, there is necessarily a point x_0 where the derivative $k'(x_0)$ equals 0 [since the function $k(x)$ is continuous and differentiable]. If $f(x)$ denotes the argument of the logarithm in the numerator of $k(x)$, and $g(x)$ the argument $1+x$ of the logarithm in the denominator, elementary considerations show that $k'(x_0) = 0$ is equivalent to

$$\frac{f'(x_0)}{f(x_0)} = \frac{k(x_0)}{g(x_0)}. \quad (2.12)$$

Given that $k(x_0)$ is assumed larger than K, Equation (2.12) yields

$$\frac{f'(x_0)}{f(x_0)} > \frac{K}{1+x_0}. \quad (2.13)$$

The derivative $f'(x_0)$ is equal to

$$f'(x_0) = \frac{(1-\phi)\left(\sqrt{1+x_0} + \sqrt{\phi}\right)}{\sqrt{1+x_0}\left(1 + \sqrt{\phi(1+x_0)}\right)^3}. \tag{2.14}$$

Bearing in mind the definition of K and Equation (2.14), it can be seen after elementary algebraic manipulations that (2.13) results in a contradiction. Therefore $\sup_{x>0} k(x) = K$.

For any ε we will now find E, Y, x, i, and j such that $|\ln G(E, Y, x, i, j)/\ln(1+x) - K| \leqslant \varepsilon$. Recalling Equation (2.6), this will ensure that $\tau(T) = K$. We let $H(E, Y, x, i, j) = \ln G(E, Y, x, i, j)/\ln(1+x)$ and note that

$$H(E,Y,x,i,j) = \frac{\ln G(E,Y,x,i,j)}{\ln(1+x)} = \frac{\ln \dfrac{1+xA^i}{1+xA^j}}{\ln(1+x)}$$

$$= \frac{\ln \dfrac{1 + x\Sigma_{k=1}^{n} e_k P(Y)_{i,k}}{1 + x\Sigma_{k=1}^{n} e_k P(Y)_{j,k}}}{\ln(1+x)}. \tag{2.15}$$

We then have

$$\lim_{x\to 0} H(E,Y,x,i,j) = \sum_{k=1}^{n} e_k\left[P(Y)_{i,k} - P(Y)_{j,k}\right] \stackrel{\text{def}}{=} H^*(E,Y,i,j). \tag{2.16}$$

Given i^*, j^*, i'^*, j'^* defined earlier, we set $i = i^*$, $j = j^*$ in (2.15) and (2.16). We also define

$$\Omega = \frac{t_{i'^* i^*} t_{i'^* j^*}}{t_{j'^* i^*} t_{j'^* j^*}}, \tag{2.17}$$

which is well defined and nonzero because $\phi > 0$. For any sufficiently small η, we can define, for $n \geqslant 3$, the probability-normed vector $Y_\eta = (y_i)$ ($i =$

$1, 2, \ldots, n)$ as

$$y_{i'^*} = \frac{1}{1 + \Omega^{0.5}} - \eta, \qquad y_{j'^*} = \frac{\Omega^{0.5}}{1 + \Omega^{0.5}} - \eta,$$

$$y_k = \frac{2\eta}{n-2} \quad \text{for } k \neq i'^*, j'^*. \tag{2.18}$$

If n (the dimension of T) is equal to 2, then set $\eta = 0$: $y_{i'^*}$ and $y_{j'^*}$ are the only components of Y_η. We also define the vector E as $E^* = (0, 0, \ldots, e_{i'^*} = 1, 0, \ldots, 0)$: E^* has a 1 in the i'^*th position and zeros elsewhere. Then

$$\begin{aligned} H^*(E^*, Y_\eta, i^*, j^*) &= P(Y_\eta)_{i^*, i'^*} - P(Y_\eta)_{j^*, i'^*} \\ &= \frac{y_{i'^*} t_{i'^* i^*}}{\frac{2\eta}{n-2}\left[\sum_{k \neq i'^*, j'^*} y_k t_{k i^*}\right] + y_{i'^*} t_{i'^* i^*} + y_{j'^*} t_{j'^* i^*}} \\ &\quad - \frac{y_{i'^*} t_{i'^* j^*}}{\frac{2\eta}{n-2}\left[\sum_{k \neq i'^*, j'^*} y_k t_{k j^*}\right] + y_{i'^*} t_{i'^* j^*} + y_{j'^*} t_{j'^* j^*}}. \end{aligned} \tag{2.19}$$

Now $H^*(E^*, Y_\eta, i^*, j^*) \to K$ when $\eta \to 0$ [if $n = 2$ then $H^*(E^*, Y_0, i^*, j^*) = K$]. Therefore, given ε, there is η (and Y_η) such that $|H^*(E^*, Y_\eta, i^*, j^*) - K| \leqslant \varepsilon/2$. Once i^*, j^*, Y_η, and E^* are fixed, there exists x^* such that $|H(E^*, Y_\eta, x^*, i^*, j^*) - H^*(E^*, Y_\eta, i^*, j^*)| \leqslant \varepsilon/2$. By the triangle inequality we finally have

$$\begin{aligned} &|H(E^*, Y_\eta, x^*, i^*, j^*) - K| \\ &\quad \leqslant |H(E^*, Y_\eta, x^*, i^*, j^*) - H^*(E^*, Y_\eta, i^*, j^*)| \\ &\quad + |H^*(E^*, Y_\eta, i^*, j^*) - K| \leqslant \varepsilon, \end{aligned} \tag{2.20}$$

which proves that $\tau(T) = K$.

To prove the result of (2.1) concerning $\tau_1[P(Y)]$ we first recall that for fixed Y

$$\tau_1[P(Y)] = \max_{i,j} \sum_{k \in \Delta} [P(Y)_{i,k} - P(Y)_{j,k}],$$

$$\Delta = \{k | [P(Y)_{i,k} - P(Y)_{j,k}] > 0\}. \qquad (2.21)$$

The fact that $H^*(E^*, Y_\eta, i^*, j^*)$ of Equation (2.19) approaches K as $\eta \to 0$ already shows that $\sup_Y \tau_1[P(Y)] \geqslant K$.

Let us see what happens if $\sup_Y \tau_1[P(Y)] > K$. We define $\mu = \sup_Y \tau_1[P(Y)] - K$ and the vector $E' = (e_k)$, where $e_k = 1$ if $k \in \Delta$ and $e_k = 0$ if $k \notin \Delta$; then $\tau_1[P(Y)] = \max_{i,j} H^*(E', Y, i, j)$. Therefore there exist i', j', Y' such that $H^*(E', Y', i', j') > K + 3\mu/4$. Recalling (2.16), it is then possible to find x such that $H(E', Y', x, i', j') > K + \mu/2$, which contradicts the fact that $\tau(T) = K$. Therefore necessarily $\sup_Y \tau_1[P(Y)] = K$, which completes the proof. ■

A result by Seneta (1981, p. 110) appears as a special case of Theorem 2.1. Indeed, for any probability-normed vector Y, Equation (2.1) shows that $\tau(T) \geqslant \tau_1[P(Y)]$. If T is then a column-stochastic matrix, and if we let Y be the n-dimensional vector with each entry equal to $1/n$, then Equation (2.1) shows that $\tau(T) \geqslant \tau_1(T')$.

The authors wish to acknowledge the useful comments of a referee on a previous version of this paper.

REFERENCES

Seneta, E. 1981. *Non-negative Matrices and Markov Chains*, 2nd ed., Springer-Verlag, Berlin.

Received 23 October 1991; final manuscript accepted 9 February 1993