SIAM J. MATRIX ANAL. APPL.
Vol. 25, No. 2, pp. 507–516

THE LOCAL COEFFICIENT OF ERGODICITY OF A NONNEGATIVE MATRIX*

MARC ARTZROUNI†

Abstract. The local coefficient of ergodicity $\tau(T, Y', w)$ of a nonnegative column-allowable matrix T at a fixed positive vector Y is defined as the supremum of $d(X'T, Y'T)/d(X', Y')$ for X not colinear to Y and $d(X', Y') \leq w$ (d is the projective distance in the positive quadrant). A near-closed-form expression is given for $\tau(T, Y', w)$. If T' is scrambling (i.e., no two rows of T' are orthogonal), then for any $Y > 0$, $w < \infty$ we have $\tau(T, Y', w) < 1$. When Y is a positive left eigenvector of T and $X_o > 0$, these results can be used to prove the convergence in direction of $X_o^{'}T^p$ to Y'. Results are illustrated with a numerical example.

Key words. nonnegative matrix, coefficient of ergodicity, eigenvector, dynamical system

AMS subject classifications. 15A18, 15A48, 39A11

DOI. 10.1137/S0895479801391801

1. Definitions and first properties.

In the theory of nonnegative matrices the coefficient of ergodicity $\tau(T)$ of a column-allowable $n \times n$ matrix $T = (t_{ij})$ (i.e., a matrix having no zero column) 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 $d(X', Y') = \max_{i,j} \ln(x_i y_j / x_j y_i)$ is the projective distance between the positive vectors $X = (x_i)$ and $Y = (y_i)$ [4, p. 83].

The quantity $\tau(T)$ (which is between 0 and 1) is a contraction coefficient for the linear operator T since $d(X'T, Y'T) \leqslant \tau(T) d(X', Y')$; $\tau(T)$ takes its full usefulness when it is < 1 since T is then a contracting operator.

For an initial vector $X_o > 0$ we may be interested in the dynamical system $X_p' = X_o' T^p$, $p = 0, 1, \ldots$. Suppose Y is a left Perron vector of T. The corresponding eigenvalue is the spectral radius $\rho(T)$ of T [3, p. 493] and $Y'T^p = \rho(T)^p Y'$ for any integer $p \geq 1$. The projective distance between Y (the fixed point of T) and X_p then satisfies

$$d(X_p', Y') = d(X_o' T^p, Y') = d(X_o' T^p, Y' T^p) \leqslant \tau(T)^p d(X_o', Y'), \qquad p = 0, 1, \ldots. \tag{1.2}$$

If $\tau(T) < 1$, then (1.2) shows that the vectors $X_o' T^p$ approach Y in direction when $p \to \infty$.

The problem is that $\tau(T) < 1$ only for T positive, which is a rather strong condition. With a single zero element in the matrix, the coefficient $\tau(T)$ is 1; T is then not a contraction and (1.2) can no longer be used to easily conclude that $d(X_o' T^p, Y') \to 0$.

If T is primitive, then a power of T is positive and the same result holds. However, there are cases where T is imprimitive or even reducible and one would still like to use a simple contraction-type argument to see whether $d(X_o' T^p, Y') \to 0$ for $p \to \infty$.

*Received by the editors July 4, 2001; accepted for publication (in revised form) by H. J. Werner March 4, 2003; published electronically September 17, 2003.
http://www.siam.org/journals/simax/25-2/39180.html

†Department of Applied Mathematics, University of Pau, 64000 Pau, France (marc.artzrouni@univ-pau.fr).

The approach used here will hinge upon the fact that one of the two vectors appearing in the projective distances of (1.2) (i.e., Y' or $Y'T^p$) is always a scalar multiple of Y. Therefore in the definition (1.1) of $\tau(T)$ the vector Y could be fixed. Furthermore, it may not be necessary to find a supremum for arbitrarily large values of $d(X', Y')$ in the denominator. Indeed, in the example of (1.2) it would suffice for $\tau(T)$ to be a supremum over all $d(X', Y')$ bounded by some positive $w > 0$ since the projective distances between the iterates $X_o'T^p$ and Y' are nonincreasing.

In order to address these issues we define $B(Y, w)$ as the ball of center Y and radius $w > 0$ for the projective distance, i.e., $X \in B(Y, w) \iff d(X', Y') \le w$. We now consider the following definition (in which the vector Y is an arbitrary positive vector; Y is not assumed to be a Perron vector of T).

DEFINITION 1.1. *The local coefficient of ergodicity (LCE) of a nonnegative matrix T in a neighborhood $B(Y, w)$ of a vector $Y > 0$ is defined as*

$$\tau(T, Y', w) \overset{\text{def}}{=} \sup_{\substack{X>0; X\neq\lambda Y \\ X\in B(Y,w)}} \frac{d(X'T, Y'T)}{d(X', Y')}. \tag{1.3}$$

If $X \in B(Y, w)$, then

$$d(X'T, Y'T) \le \tau(T, Y', w) d(X', Y'), \tag{1.4}$$

and if $\tau(T, Y', w) < 1$, we say that T is a local contraction *with respect to Y.*

The definition and properties of $\tau(T)$ insure that $\tau(T, Y', w)$ is defined and is ≤ 1. The LCE has a submultiplicative property similar to the one that holds for τ (i.e., $\tau(T_1T_2) \le \tau(T_1)\tau(T_2)$). Indeed, let $\{T_i\}_{i=1,2,\ldots}$ be a sequence of column-allowable matrices and define the forward product $U_p = T_1T_2\ldots T_p$. The submultiplicative property is then given in the following proposition.

PROPOSITION 1.2. *For positive vectors X, Y we define $w_o \overset{\text{def}}{=} d(X', Y')$. With the notation given above, we then have*

$$\tau(T_1T_2, Y', w_o) \le \tau(T_2, Y'T_1, w_o)\tau(T_1, Y', w_o) \tag{1.5}$$

and more generally

$$\tau(U_k, Y', w_o) \le \tau(T_k, Y'U_{k-1}, w_o)\tau(T_{k-1}, Y'U_{k-2}, w_o)\ldots\tau(T_1, Y', w_o). \tag{1.6}$$

Proof. Because $d(X'T_1, Y'T_1) \le d(X', Y')$ we have

$$\begin{aligned}\tau(T_1T_2, Y', w_o) &= \sup_{\substack{X>0; X\neq\lambda Y \\ X\in B(Y,w)}} \frac{d(X'T_1T_2, Y'T_1T_2)}{d(X', Y')} \\ &\le \sup_{\substack{X'T_1\neq\lambda Y'T_1 \\ X\in B(Y,w)}} \frac{d((X'T_1)T_2, (Y'T_1)T_2)}{d(X'T_1, Y'T_1)} \sup_{\substack{X>0; X\neq\lambda Y \\ X\in B(Y,w)}} \frac{d(X'T_1, Y'T_1)}{d(X', Y')} \\ &\le \tau(T_2, Y'T_1, w_o)\tau(T_1, Y', w_o),\end{aligned} \tag{1.7}$$

which is the desired result of (1.5), from which (1.6) follows by induction.

When all the matrices T_i are equal to some T, and $Y > 0$ is a left positive eigenvector of T, then $Y'U_p = Y'T^p$; $Y'T^p$ and Y' are colinear for all p and (1.6) yields

$$d(X_o'T^p, Y') = d(X_o'T^p, Y'T^p) \leq \tau(T^p, Y', w_o)w_o \leq \tau(T, Y', w_o)^p w_o, \quad p = 0, 1, \ldots, \tag{1.8}$$

which shows that $d(X_o'T^p, Y')$ approaches 0 exponentially fast if $\tau(T, Y', w_o) < 1$.

We will see that $\tau(T, Y', w_o)$ is < 1 under conditions that are much weaker than the positivity assumption needed for $\tau(T) < 1$. In fact we will show that for any $Y > 0$ (not necessarily a positive eigenvector) the LCE $\tau(T, Y', w)$ is < 1 for any finite w as soon as T' is scrambling (i.e., any two rows of T' have at least one positive entry in a coincident position, which means that no two rows are orthogonal). This result is not entirely surprising because when T' is scrambling, then $d(X'T, Y'T) < d(X', Y')$. However, this inequality alone is not sufficient to make a contraction-type argument as in (1.8) when $\tau(T, Y', w_o) < 1$.

The scrambling condition is obviously much weaker than the condition $T > 0$, which must be satisfied in order to have $\tau(T) < 1$. For these reasons the LCE is useful not only here but also in other similar situations when we are interested in the ratio

$$R(X, Y) \stackrel{\text{def}}{=} d(X'T, Y'T)/d(X', Y') \tag{1.9}$$

with X and Y not colinear. The remainder of this paper is devoted to the study of $\tau(T, Y', w)$, with an emphasis on the conditions under which $\tau(T, Y', w) < 1$.

2. Preliminary definitions and results. We first define the subset Ω of the nonnegative quadrant $\mathbb{R}_+^n$ as the set of all vectors with components between 0 and 1, with at least one component equal to 0 and one equal to 1:

$$\Omega = \{E = (e_i) \in \mathbb{R}_+^n : 0 \leq e_i \leq 1; \text{ at least one } e_i = 0, \text{ one } e_i = 1\}. \tag{2.1}$$

With T column-allowable, we define for a fixed vector $Y = (y_i) > 0$ the row-stochastic matrix $P(T, Y)$ as

$$P(T, Y)_{ij} \stackrel{\text{def}}{=} \frac{y_j t_{ji}}{\sum_{q=1}^{q=n} y_q t_{qi}}, \quad i, j = 1, 2, \ldots, n. \tag{2.2}$$

The matrices $P(T, Y)$ and T have "transposed incidences": the incidence of $P(T, Y)$ is that of T'. If we let $P(T, Y)_i$ denote the ith row of $P(T, Y)$, and if $E = (e_i)$ is a vector of Ω, we note that $P(T, Y)_i E$ is the scalar product $\sum_{k=1}^n P(T, Y)_{ik} e_k$.

We next let $A \circ B$ be the Hadamard (componentwise) product of two matrices or vectors. As in [2] we will express any vector $X > 0$ in a way that will simplify the expressions for $d(X', Y')$ and $d(X'T, Y'T)$.

PROPOSITION 2.1. *For any $X = (x_i) > 0$ not colinear to $Y > 0$ there exists a unique $s > 0$ and a vector $E \in \Omega$ such that $Y + sY \circ E$ and X are colinear. Then*

$$d(X', Y') = d([Y + sY \circ E]', Y') = \ln(1 + s), \tag{2.3}$$

$$d(X'T, Y'T) = \max_{i,j} \ln \left(\frac{1 + s \times P(T, Y)_i E}{1 + s \times P(T, Y)_j E} \right). \tag{2.4}$$

Proof. The components x_i of X can be written as $x_i = y_i(r + \sigma e_i)$, where $r \stackrel{\text{def}}{=} \min_i x_i/y_i$, $\sigma \stackrel{\text{def}}{=} \max_i x_i/y_i - \min_i x_i/y$, and $E = (e_i) = (x_i/y_i - r)/\sigma$ ($\sigma > 0$ because X and Y are not colinear). Then

$$X = Yr + \sigma Y \circ E, \tag{2.5}$$

which after setting $s = \sigma/r$ shows that $Y + sY \circ E$ and X are colinear. This proves (2.3). Also

$$d(X'T, Y'T) = d([Y + sY \circ E]'T, Y'T) = \max_{i,j} \ln \left[\frac{1 + s\sum_k e_k P(T,Y)_{ik}}{1 + s\sum_k e_k P(T,Y)_{jk}} \right], \tag{2.6}$$

which is the desired result of (2.4).

If we define $w^* \stackrel{\text{def}}{=} \exp(w) - 1$, we then have

$$\tau(T, Y, w) = \max_{i,j} \sup_{\substack{0<s\le w^* \\ E\in\Omega}} \frac{\ln\left[\frac{1+sP(T,Y)_i E}{1+sP(T,Y)_j E}\right]}{\ln(1+s)}. \tag{2.7}$$

We will now proceed in two steps: First we will find for fixed s, i, j the sup over E of the bracketed expression in the numerator. Then we will seek the supremum over $0 < s \le w^*$ of the ratio of the two logarithms.

2.1. Supremum over *E*. We define for two probability-normed vectors $a = (a_i)$, $b = (b_i)$ the function

$$Z(s, a, b, E) = \frac{1 + sa'E}{1 + sb'E}, \tag{2.8}$$

of which we seek the supremum for $E = (e_i) \in \Omega$. ($a'E = \sum_{k=1}^n a_k e_k$, $b'E = \sum_{k=1}^n b_k e_k$, so that a and b represent the ith and jth rows of $P(T, Y)$.)

A supremum is necessarily reached for each e_i equal to either 0 or 1. In this context we define the finite subset Ω' of Ω consisting of vectors having their last k components equal to 1 ($k = 2, 3, \ldots, n$) and the others equal to 0:

$$\Omega' = \{E(k) = (0, 0, \ldots, 0, 1, 1, , , 1),\ \text{first “1” in } k\text{th position},\ k = 2, 3, \ldots, n\}. \tag{2.9}$$

We now reorder the components (a_i, b_i) in the following way. We first have those components that are both 0. Then we have in increasing order of the ratios $r_i = a_i/b_i$ the components for which $b_i > 0$. Finally we have in increasing order of the a_i's those components for which $b_i = 0$. For example, if $a = (0.3\ 0\ 0\ 0.2\ 0.5)$ and $b = (0\ 0\ 0.2\ 0.1\ 0.7)$, then the vectors with reordered components are $a = (0\ 0\ 0.5\ 0.2\ 0.3)$ and $b = (0\ 0.2\ 0.7\ 0.1\ 0)$. The corresponding vector of increasing ratios is $r = (0/0\ 0/0.2\ 0.5/0.7\ 0.2/0.1\ 0.3/0)$.

In what follows we will assume that the components of a and b have been reordered in this manner, and we say that the pair (a, b) has the *increasing ratio property (IRP)*. (Note that in general the reordering of (a, b) and of (b, a) are not the same. In fact the two orderings are mirror images of one another.)

We first dispose of two trivial cases:

i. If $a = b$, then $Z(s, a, b, E)$ is 1 for any E.

ii. If a and b are orthogonal (i.e., do not have a positive term in a coincident position, which means $a_i b_i = 0$ for all i), then there exists an $E(k)$ such that $Z(s,a,b,E(k))$ is equal to its maximum possible value $1+s$.

We now assume that $a \neq b$ and that a and b are not orthogonal. Then there is necessarily at least one ratio $r_m = a_m/b_m$ that is strictly less than 1 and one that is strictly larger than 1 (and possibly $+\infty$). We thus define

$$m_1 = \{m/\ r_m \leq 1 < r_{m+1}\}, \tag{2.10}$$

$$m_2 = \min\{m/\ r_m = r_{m+1} = \cdots = r_n\}. \tag{2.11}$$

In words, r_{m_1} is the largest ratio in the list $\{r_k\}_{k=1,2,\ldots,n}$ to be ≤ 1; r_{m_2} is the first in the list to be equal to the last (and largest) ratio r_n (where r_n may be $+\infty$). With the example $a = (0\ 0\ 0.5\ 0.2\ 0.3)$, $b = (0\ 0.2\ 0.7\ 0.1\ 0)$, $r = (0/0\ 0/0.2\ 0.5/0.7\ 0.2/0.1\ 0.3/0)$, we have $m_1 = 3, m_2 = 5$. If $a = (0.56\ 0.4\ 0.04)$, $b = (0.67\ 0.3\ 0.03)$, then $r = (0.56/0.67\ 0.4/0.3\ 0.04/0.03)$ so that $m_1 = 1, m_2 = 2$ (m_2 is strictly larger than m_1 because $a \neq b$).

We will now use this reordering to partition the set of all positive real numbers into intervals $I(k)$ within which the sup of $Z(s,a,b,E)$ over E is attained for $E(k)$. We define the $m_2 - m_1$ half-open intervals

$$I(k) \stackrel{\text{def}}{=} [S(k-1), S(k)\ ), \quad k = m_1+1, m_1+2, \ldots, m_2,$$

where the quantities $S(k)$ are given by

$$S(m_1) = 0; \quad S(k) = \frac{a_k - b_k}{b_k \sum_{p=k}^n a_p - a_k \sum_{p=k}^n b_p}, \qquad k = m_1+1, m_1+2, \ldots, m_2. \tag{2.12}$$

The $S(k)$'s are nonnegative numbers that satisfy

$$S(m_1) = 0 \leq S(m_1+1) \leq \cdots \leq S(m_2-1) < S(m_2) = +\infty, \tag{2.13}$$

$$Z(S(k), a, b, E(k)) = r_k, \quad k = m_1+1, m_1+2, \ldots, m_2. \tag{2.14}$$

In short the $m_2 - m_1$ intervals $I(k)$ constitute a partition of the set of real positive numbers such that

$$s \in I(k) \quad \Longleftrightarrow r_{k-1} \leq Z(s,a,b,E(k)) < r_k. \tag{2.15}$$

We will now show that when (a,b) has the IRP, then for a fixed s in any $I(k)$, the supremum of $Z(s,a,b,E)$ over $E = (e_i) \in \Omega$ is attained at $Z(s,a,b,E(k))$.

PROPOSITION 2.2. *For two probability-normed vectors (a,b) having the IRP, with m_1, m_2 given in (2.10)–(2.11), we have*

$$\begin{aligned} s \in I(k) \Rightarrow & \sup_{E\in\Omega} Z(s,a,b,E) = Z(s,a,b,E(k)) \\ & = \frac{1 + sa'E(k)}{1 + sb'E(k)}, \quad k = m_1+1, m_1+2, \ldots, m_2, \end{aligned} \tag{2.16}$$

$$b'E(k) \leq a'E(k) \leq 1. \tag{2.17}$$

Proof. The proof will hinge upon the following elementary numerical results concerning four nonnegative numbers u, u', v, v':

$$\frac{u+u'}{v+v'} \lesseqgtr \frac{u'}{v'} \Longleftrightarrow \frac{u}{v} \lesseqgtr \frac{u'}{v'} \Longleftrightarrow \frac{u}{v} \lesseqgtr \frac{u+u'}{v+v'}. \tag{2.18}$$

We recall that a sup is necessarily reached with each e_k equal to either 0 or 1. Let us assume that the sup is reached for some $E^* = (e^*_k)$ that has a "1" to the left of a "0":

$$\sup_{E\in\Omega} Z(s,a,b,E) = Z(s,a,b,E^*) = \frac{1+s\sum_{i=1}^n a_i e^*_i}{1+s\sum_{i=1}^n b_i e^*_i},\ \exists\ w<q \text{ with } e^*_w = 1, e^*_q = 0. \tag{2.19}$$

We will show that a contradiction will follow, which will leave only the elements of Ω' as candidates for the optimum E. Indeed if $Z(s,a,b,E^*) \le r_q$, then (2.18) shows that $Z(s,a,b,E^*)$ can be increased by changing e^*_q to 1. If $Z(s,a,b,E^*) > r_q$, then (2.18) shows that $Z(s,a,b,E^*)$ can be increased by changing e^*_w to 0. This contradicts the assumption that E^* is a supremum and shows that the sup is necessarily reached for some element of Ω'. (We note in particular that e_n, the last component of E, is necessarily 1 at the optimal value. This insures that one component is indeed 1, which was an early requirement.)

We now show that $E(k)$ is the element of Ω' at which the supremum in reached. To simplify the writing we define $h(k) = Z(s,a,b,E(k))$. We will show that if $s \in I(k)$, then $h(k)$ is the maximum value of $h(m)$ for all m. First $h(k+1) \le h(k)$ because we obtain $h(k+1)$ by removing (a_k, b_k) from the numerator and the denominator of $h(k)$, which from (2.18) decreases $h(k)$ since $h(k) < r_k \le r_{k+1}$; $h(k+2)$ is obtained by removing (a_{k+1}, b_{k+1}) from $h(k+1)$ but $h(k+1) \le h(k) < r_{k+1}$ so $h(k+2) \le h(k+1)$. Hence with each increase in the index $j, h(k+j)$ becomes smaller because pairs (a_{k+j}, b_{k+j}) with increasing ratios r_{k+j} are removed while $h(k+j)$ decreases. A similar reasoning holds if j decreases: $h(k-1) \le h(k)$ because we obtain $h(k-1)$ by adding (a_{k-1}, b_{k-1}) to the numerator and the denominator of $h(k)$, which from (2.18) implies $r_{k-1} \le h(k-1) \le h(k)$; $h(k-2)$ is obtained by adding (a_{k-2}, b_{k-2}) to $h(k-1)$; however, $r_{k-2} \le r_{k-1} \le h(k-1)$, which insures that $r_{k-2} \le h(k-2) \le h(k-1)$. Hence with each decrease in the index $j, h(k-j)$ becomes smaller while staying larger than r_{k-j}. This shows that for $s \in I(k)$, the maximum value of $h(m)$ is $h(k)$.

2.2. Supremum over *s*. Now that we have $Z(s,a,b,E(k))$ as the supremum of $Z(s,a,b,E)$ over E in each $I(k)$, we seek the supremum of $\ln[Z(s,a,b,E(k)]/\ln(1+s)$ for $s \in I(k)$. We thus consider the function

$$Q(s,\alpha,\beta) \stackrel{\text{def}}{=} \frac{\ln\frac{1+s\alpha}{1+s\beta}}{\ln(1+s)},\ s>0; \quad 0 \le \beta < \alpha \le 1, \tag{2.20}$$

whose derivative $Q'(s,\alpha,\beta)$ with respect to s is

$$Q'(s,\alpha,\beta) = \frac{\dfrac{\frac{(1+s)(\alpha-\beta)}{1+\beta s}}{1+\alpha s} - \dfrac{\ln\left(\frac{1+\alpha s}{1+\beta s}\right)}{\ln(1+s)}}{(1+s)\ln(1+s)}. \tag{2.21}$$

Two examples of the function $Q(s,\alpha,\beta)$ are given in Figure 2.1, one with $\alpha+\beta = 1.3$, the other with $\alpha+\beta = 0.35$. As we will see below, the function is monotone decreasing when $\alpha+\beta \ge 1$ and has one maximum when $\alpha+\beta < 1$.

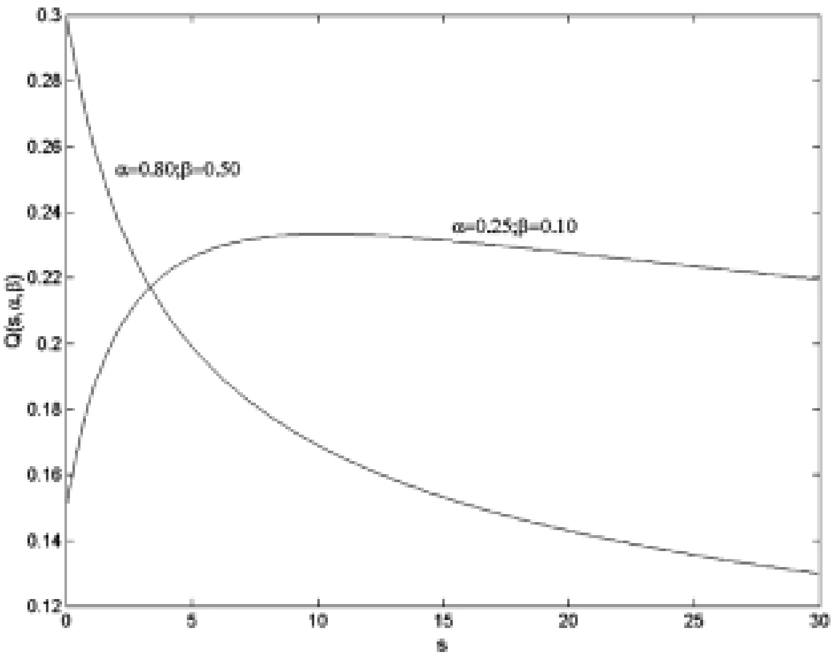

FIG. 2.1. *Two examples of the function* $Q(s, \alpha, \beta)$.

PROPOSITION 2.3. *When* $\alpha > 0$ *the function* $Q(s, \alpha, 0)$ *increases monotonically from* α *to* 1 *as* s *grows from* 0^+ *to* $+\infty$.

i. *If* $\alpha + \beta \geq 1, \beta > 0$, *then* $Q(s, \alpha, \beta)$ *decreases monotonically from* $\alpha - \beta$ *to* 0 *with* s.

ii. *If* $\alpha + \beta < 1, \beta > 0$, *then* $Q(s, \alpha, \beta)$ *first increases from* $\alpha - \beta$ *to a maximum* $Q^*(\alpha, \beta)$, *then decreases to* 0 *as* $s \to +\infty$. *The maximum* $Q^*(\alpha, \beta)$ *is reached at a value* $s^* = s^*(\alpha, \beta)$ *of* s *that is the unique positive root of the equation (in* s*)*

$$\left(\frac{1 + s\alpha}{1 + s\beta}\right) = (1 + s)^{\frac{(1+s)(\alpha-\beta)}{(1+\alpha s)(1+\beta s)}}. \tag{2.22}$$

Then

$$Q^*(\alpha, \beta) = \frac{\ln \frac{1+s^*\alpha}{1+s^*\beta}}{\ln(1 + s^*)} = \frac{(1 + s^*)(\alpha - \beta)}{(1 + \alpha s^*)(1 + \beta s^*)} \leq (\alpha - \beta)/(\alpha + \beta) < 1. \tag{2.23}$$

Proof. By setting $Q'(s, \alpha, \beta)$ of (2.21) equal to 0, one obtains (2.23) and thus (2.22). The upper bound $(\alpha - \beta)/(\alpha + \beta)$ for $Q^*(\alpha, \beta)$ is the maximum (over s^*) of the third term in (2.23). Other elementary details are omitted.

There is no closed-form expression for $s^*(\alpha, \beta)$, the root of (2.22). However, calculating $s^*(\alpha, \beta)$ is an elementary numerical problem because $Q(s, \alpha, \beta)$ is then a simple function that increases then decreases.

Bearing in mind the expression for $\tau(T,Y,w)$ given in (2.7), we now have

$$\tau(T,Y,w) = \max_{i,j} \sup_{\substack{0<s<w^* \\ E\in\Omega}} Q[s, P(T,Y)_i E, P(T,Y)_j E]. \tag{2.24}$$

We will partition $[0,w^*)$ using the intervals $I(k)$ in which we now know that the supremum over E is $E(k)$. If for any $z>0$ we let $k(z)$ denote the index of the interval that contains z, then $w^* \in I(k(w^*)) = [S(k(w^*)-1), S(k(w^*)))$. Then the interval $[0,w^*)$ is the union of the intervals $I(k) = [S(k-1), S(k))$ for k going from m_1+1 to $k(w^*)-1$, to which we add the interval $[S(k(w^*)-1), w^*)$ (which for simplicity of notation we will call $I(k(w^*))$ below, even though strictly speaking it is contained in $I(k(w^*))$ and not equal to $I(k(w^*))$. The next section gives the main result, which hinges upon this particular partitioning of the interval $[0,w^*)$.

3. Main result and applications. With the notation given above, the following theorem provides a near-closed-form expression for the LCE $\tau(T,Y,w)$.

THEOREM 3.1. *For a column-allowable matrix T and a positive vector Y, the LCE $\tau(T,Y,w)$ is equal to*

$$\tau(T,Y,w) = \max_{i,j} \max_{p=m_1+1,m_1+2,\dots,k(w^*)-1,k(w^*)} \sup_{s\in I(p)} Q[s, P(T,Y)_i E(p), P(T,Y)_j E(p)], \tag{3.1}$$

where the vectors $(P(T,Y)_i, P(T,Y)_j)$ have the IRP, i.e., the ratios

$$r_k = P(T,Y)_{ik}/P(T,Y)_{jk} \tag{3.2}$$

are increasing for $k = m_1, m_1+1, \dots, m_2$, with m_1, m_2 defined in (2.10), (2.11). *(m_1, m_2, and $k(w^*)$ depend on the particular pair of indices (i,j).)*

If T' is not scrambling (i.e., T' has two orthogonal rows), then two vectors $P(T,Y)_i$ and $P(T,Y)_j$ are orthogonal and $\tau(T,Y,w)=1$. If the columns of T are multiples of a common nonzero vector Z (T of rank 1), then all the rows of $P(T,Y)$ are identical and $\tau(T,Y,w)=0$.

In the general case (T scrambling and of rank >1), we recall the notation

$$S(m_1) = 0, \tag{3.3}$$

$$S(k) = \frac{P(T,Y)_{ik} - P(T,Y)_{jk}}{P(T,Y)_{jk}\sum_{p=k}^{n} P(T,Y)_{ip} - P(T,Y)_{ik}\sum_{p=k}^{n} P(T,Y)_{jp}}; \quad k = m_1+1, m_1+2, \dots, m_2, \tag{3.4}$$

and $w^* = \exp(w)-1$. For $p = m_1+1, m_1+2, \dots, k(w^*)-1$ the intervals $I(p)$ in (3.1) *are $I(p) = [S(p-1), S(p))$ as in* (2.12). *The last interval $I(k(w^*))$ stops at w^* and is $[S(k(w^*)-1), w^*)$ rather than $[S(k(w^*)-1), S(k(w^*)))$.*

We next define the values $Q_{le}^{(p)}$ and $Q_{re}^{(p)}$ of the function $Q[s, P(T,Y)_i E(p), P(T,Y)_j E(p)]$ at the left end (le) and right end (re) of the corresponding interval $I(p)$, i.e.,

$$p = m_1+1 \Rightarrow \begin{cases} Q_{le}^{(p)} = P(T,Y)_i E(m_1+1) - P(T,Y)_j E(m_1+1), \\ Q_{re}^{(p)} = \ln(r_{m_1+1})/\ln(1+S(m_1+1)), \end{cases} \tag{3.5}$$

$$p = m_1+2, m_1+3, \ldots, k(w^*)-1 \Rightarrow \quad \begin{cases} Q_{le}^{(p)} = \ln(r_{p-1})/\ln(1+S(p-1)), \\ Q_{re}^{(p)} = \ln(r_p)/\ln(1+S(p)), \end{cases} \tag{3.6}$$

$$p = k(w^*) \Rightarrow \quad \begin{cases} Q_{le}^{(p)} = \ln(r_{k(w*)-1})/\ln(1+S(k(w^*)-1)), \\ Q_{re}^{(p)} = \ln\left(\dfrac{1+w^*P(T,Y)_iE(k(w^*))}{1+w^*P(T,Y)_jE(k(w^*))}\right)/\ln(1+w^*). \end{cases} \tag{3.7}$$

When $P(T,Y)_iE(p) + P(T,Y)_jE(p) < 1$*, we let* s^* *be the value at which* $Q[s, P(T,Y)_iE(p), P(T,Y)_jE(p)]$ *reaches its maximum (see Proposition* 2.3*).*

The suprema of (3.1) *are now as follows:*

i. *If either* $(P(T,Y)_iE(p) + P(T,Y)_jE(p) \geq 1$ *and* $P(T,Y)_jE(p) > 0)$ *or* s^**is to the* left *of the interval* $I(p)$*, then the supremum is reached at the left end of* $I(p)$*:*

$$\sup_{s\in I(p)} Q[s, P(T,Y)_iE(p), P(T,Y)_jE(p)] = Q_{le}^{(p)}. \tag{3.8}$$

ii. *If either* $(P(T,Y)_iE(p) + P(T,Y)_jE(p) < 1$ *and* $P(T,Y)_jE(p) = 0)$ *or* s^* *is to the* right *of* $I(p)$*, then the supremum is reached at the right end of* $I(p)$*:*

$$\sup_{s\in I(p)} Q[s, P(T,Y)_iE(p), P(T,Y)_jE(p)] = Q_{re}^{(p)}. \tag{3.9}$$

iii. *If* s^* *is inside* $I(p)$*, then*

$$\sup_{s\in I(p)} Q[s, P(T,Y)_iE(p), P(T,Y)_jE(p)] = Q[s^*, P(T,Y)_iE(p), P(T,Y)_jE(p)]. \tag{3.10}$$

Proof. The expressions obtained in this theorem are direct consequences of previous results stemming from Proposition 2.3. Equation (3.1) reflects the partitioning of the interval $[0, w^*)$ into intervals $I(p)$ over which the sup over E is E_p. The supremum over each $I(p)$ is at the left end or right end of $I(p)$, or at the value s^* at which $Q[s, P(T,Y)_iE(p), P(T,Y)_jE(p)]$ reaches a maximum, depending on the value of $P(T,Y)_iE(p)+P(T,Y)_jE(p)$ relative to 1 (see Proposition 2.3). The value $Q_{le}^{(p)}$ given in (3.5) for the left end of the first interval $I(m_1+1) = [S(m_1), S(m_1+1)) = [0, S(m_1+1))$ is obtained by taking the limit of $Q[s, P(T,Y)_iE(m_1+1), P(T,Y)_jE(m_1+1)]$ for $s \to 0$.

We now give a simple condition for $\tau(T,Y,w)$ to be strictly less than 1.

Corollary 3.2. *For a column-allowable matrix* T*, a positive vector* Y*, and any* $w > 0$*, the LCE* $\tau(T,Y,w)$ *is strictly less than* 1 *if and only if* T' *is scrambling.*

Proof. We showed above that T' is not scrambling $\Longrightarrow \tau(T,Y,w) = 1$. Conversely, let us assume that $\tau(T,Y,w) = 1$. From Proposition 2.3 a supremum of 1 can be reached only if there is a $P(T,Y)_jE(p) = 0$ and $P(T,Y)_jE(p) = 1$, which means that T' has two orthogonal rows and is not scrambling.

A simple numerical example is provided by the triangular matrix

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}$$

with positive left eigenvector $Y' = \begin{pmatrix} 2 & 1 & 1 \end{pmatrix}$ and corresponding eigenvalue $\rho(T) = 2$. No conclusion can be drawn on the iterates $X_o'T^p$ by considering the powers of this reducible matrix which has a coefficient of ergodicity equal to 1 and whose powers remain triangular. However, T' is scrambling, and the results proved here immediately yield the desired conclusion, namely, exponential convergence to 0 of the projective distance $d(X_o'T^p, Y')$. Indeed

$$d(X_o'T^p, Y') \leq \tau(T^p, Y', w_o)w_o \leq \tau(T, Y', w_o)^p w_o \underset{p\longrightarrow\infty}{\longrightarrow} 0. \tag{3.11}$$

With $w_o = 2$ a MATLAB program (available from the author) yields $\tau(T, Y, 2) = 0.88$. As $X_o'T^p$ approaches Y' in direction, the quantity $\tau(T, Y, 0) \overset{\text{def}}{=} \lim_{w\rightarrow 0} \tau(T, Y, w)$ is an asymptotic rate of convergence equal in the present case to 0.75.

If in order to emphasize the dependence on (i, j) we write $m_1(i, j)$ for m_1 of (3.1), then

$$\tau(T, Y, 0) = \lim_{w\rightarrow 0} \tau(T, Y, w) = \max_{i,j} [P(T, Y)_i E(m_1(i, j)+1) - P(T, Y)_j E(m_1(i, j)+1)]. \tag{3.12}$$

It can easily be seen that this asymptotic rate of convergence $\tau(T, Y, 0)$ is equal to $\tau_1(P(T, Y))$, where τ_1 is the classical coefficient of ergodicity defined on row-stochastic matrices [4], i.e., for any row-stochastic matrix $Q = (q_{ij})$,

$$\tau_1(Q) = 0.5 \max_{i,j} \sum_k |q_{ik} - q_{jk}| = \max_{i,j} \sum_{\substack{k\in\Delta(i,j) \\ \Delta(i,j)\overset{\text{def}}{=}\{k:q_{ik}-q_{jk}>0\}}} (q_{ik} - q_{jk}). \tag{3.13}$$

We showed in [1] that $\tau(T) = \sup_{Y>0} \tau_1(P(T, Y))$, and we thus come full circle with

$$\sup_{Y>0;\sigma>0} \tau(T, Y, \sigma) = \tau(T) = \sup_{Y>0} \tau_1(P(T, Y)) = \sup_{Y>0} \tau(T, Y, 0). \tag{3.14}$$

Acknowledgment. The author wishes to thank an anonymous referee, whose comments substantially improved the paper.

REFERENCES

[1] M. ARTZROUNI AND X. LI, *A note on the coefficient of ergodicity of a column-allowable nonnegative matrix*, Linear Algebra Appl., 214 (1995), pp. 93–101.
[2] M. ARTZROUNI AND O. GAVART, *Nonlinear matrix iterative processes and generalized coefficients of ergodicity*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1343–1353.
[3] R. A. HORN AND C.R. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
[4] E. SENETA, *Non-Negative Matrices and Markov Chains*, 2nd ed., Springer-Verlag, Berlin, 1981.