Explicit Forms for Ergodicity Coefficients of Stochastic Matrices

Eugene Seneta
School of Mathematics and Statistics, F07
University of Sydney
N.S.W. 2006, Australia

Submitted by Richard A. Brualdi

ABSTRACT

Motivated by explicit expressions appearing in the work of A. Rhodius (1993) for $n \times n$ stochastic matrices P, it is shown that ordinary matrix norms on $\mathbb{R}^{n-1}$ for $(n-1) \times (n-1)$ matrices of the form APB can be used to generate results of this kind.

1. ERGODICITY COEFFICIENTS FOR STOCHASTIC MATRICES

Suppose $\|\cdot\|$ is any vector norm on row vectors constituting $\mathbb{R}^n$, and S_n $(n \geqslant 2)$ the set of all $n \times n$ stochastic matrices. The definition of an ergodicity coefficient τ for $P = \{p_{ij}\}$ with respect to the norm $\|\cdot\|$, given in Seneta (1979) [see also Seneta (1981)], is

$$\tau(P) = \sup\{\|x^T P\| : \mathbf{x}^T \in \mathbb{H}^n, \quad \|\mathbf{x}^T\| = 1\}, \tag{1.1}$$

where

$$\mathbb{H}^n = \left\{\mathbf{x}^T : \mathbf{x}^T \in \mathbb{R}^n, \quad \mathbf{x}^T \mathbf{1}_n = 0\right\}.$$

Such coefficients have the following properties:

(A) $\tau(P_1 P_2) \leqslant \tau(P_1)\tau(P_2)$, $P_1 P_2 \in S_n$;
(B) $\tau(P) = 0$ iff rank $P = 1$ (i.e. iff $P = \mathbf{1}\mathbf{v}^T$, $\mathbf{v} \geqslant 0$, $\mathbf{v}^T\mathbf{1} = 1$);
(C) $|\lambda| \leqslant \tau(P)$ for any eigenvalue λ of P, $\lambda \neq 1$.

655 Avenue of the Americas, New York, NY 10010 0024-3795/93/$6.00

(A) and (B) follow easily from the definition. Since $\|\cdot\|$ is a norm on $\mathbb{R}^n$ rather than $\mathbb{C}^n$, the validity of (C) is less obvious. It has been proved in general by Rothblum and Tan (1985, Theorem 3.1, purportedly only for irreducible P, but in fact for any $P \in S_n$), although its validity in special cases of $\|\cdot\|$ (the l_p norms where $p = 1, \infty$) was known earlier.

Properties (A) and (B) are useful in consideration of ergodicity problems of inhomogeneous Markov chains in that the scalar submultiplicative functional τ provides a measure of divergence from equal row sums for a stochastic P. Property (C) is useful for homogeneous Markov chains for which P is a primitive matrix provided $\tau(P) < 1$, since for such a matrix all eigenvalues, except one unit eigenvalue, have moduli less than unity. In the setting (C) in particular, an explicit form in terms of the elements of P is desirable.

Explicit forms for τ in terms of the entries of P are known for the l_p norms when $p = 1$, $p = \infty$. In the case $p = 1$ we denote the coefficient by δ:

$$\delta(P) = \max_{i,j} \frac{1}{2} \sum_{s=1}^{n} |p_{is} - p_{js}|, \tag{1.2}$$

so that $\delta(P) \leqslant 1$ for all $P \in S_n$.

We note that in effect $\|\cdot\|$ may be regarded as a norm on $\mathbb{H}^n$, and then $\tau(P)$ is an ordinary norm of the operator P on $\mathbb{H}^n$, the vector subspace of $\mathbb{R}^n$ which is the orthogonal complement of $\mathbf{1}^T$. Very recently, Rhodius (1993) has focused on other norms on $\mathbb{H}^n$ and used the relation between norms and convex bodies to produce the explicit forms

$$\tau^k(P) = \max_{\substack{i=1,\dots,n \\ i \neq k}} \sum_{\substack{j=1 \\ j \neq k}}^{n} |p_{kj} - p_{ij}| \tag{1.3}$$

and

$$\tau_q(P) = \max_{i=1,\dots,n-1} \sum_{j=1}^{n-1} \left| \sum_{k=1}^{j} (p_{ik} - p_{i+1,k}) \right|. \tag{1.4}$$

He has also shown that there are $P \in S_n$ such that $\tau^k(P) < \delta(P) < 1$, which is enough to illustrate the occasional usefulness of such τ.

The purpose of the present note is to provide a possibly simpler approach to the generation of explicit forms such as (1.3) and (1.4). This amounts, in the end, to the use of ordinary norms, such as l_1 and l_∞ but on $\mathbb{R}^{n-1}$

(without any linear constraint such as $\mathbf{x}^T \mathbf{1} = 0$, which introduces the complication of a subspace), as suggested by the forms of (1.3) and (1.4), which are standard matrix norms on $\mathbb{R}^{n-1}$. In particular, this paper shows that the results of Rhodius (1993) fall within the framework of Rothblum and Tan 0985).

2. UNDERLYING RESULTS

The underlying idea of this paper is, as with Rhodius (1993), the consideration of norms on the subspace $\mathbb{H}^n$. Note again that P may be regarded as an operator on $\mathbb{H}^n$, since for $\mathbf{x}^T \in \mathbb{H}^n$, $\mathbf{x}^T P \mathbf{1} = \mathbf{x}^T \mathbf{1} = 0$. Let $\|\cdot\|_{\mathbb{H}^n}$ be any vector norm on $\mathbb{H}^n$, so that $\|P\|_{\mathbb{H}^n}$ denotes the corresponding operator norm.

THEOREM 1. ***For each*** $\|\cdot\|_{\mathbb{H}^n}$ ***there is a vector norm on*** $\mathbb{R}^n$ ***such that the corresponding*** $\tau(\cdot)$ *defined* ***by*** (1.1) ***satisfies***

$$\tau(P) = \|P\|_{\mathbb{H}^n}.$$

In particular, properties (A)-(C) ***hold for the functional*** $\|\cdot\|_{\mathbb{H}^n}$.

Proof. As noted, $\mathbf{1}^T$ spans the orthogonal complement of $\mathbb{H}^n$ in $\mathbb{R}^n$. Thus we may write uniquely, for any $\mathbf{v}^T \in \mathbb{R}^n$,

$$\mathbf{v}^T = \mathbf{h}^T + \alpha \mathbf{1}^T,$$

where $\alpha \in \mathbb{R}^1$, $\mathbf{h}^T \in \mathbb{H}^n$.

It is readily verified that $|||\cdot|||$ defined for $\mathbf{v}^T \in \mathbb{R}^n$ by

$$|||\mathbf{v}^T||| = \|\mathbf{h}^T\|_{\mathbb{H}^n} + |\alpha|$$

is a norm on $\mathbb{R}^n$ (thus, an extension of the norm $\|\cdot\|_{\mathbb{H}^n}$ to $\mathbb{R}^n$). Further, for $\mathbf{x}^T \in \mathbb{H}^n$ we have $|||\mathbf{x}^T||| = \|\mathbf{x}^T\|_{\mathbb{H}^n}$, so from (1.1), using $|||\cdot|||$ for the norm on $\mathbb{R}^n$,

$$\tau(P) = \sup\{\|\mathbf{x}^T P\|_{\mathbb{H}^n} : \mathbf{x}^T \in \mathbb{H}^n, \|\mathbf{x}^T\|_{\mathbb{H}^n} = 1\} = \|P\|_{\mathbb{H}^n},$$

■

$\mathbb{H}^n$.

THEOREM 2. ***Let*** A ***be a real*** $(n-1) \times n$ *matrix of rank* $n-1$ ***and such that*** $A\mathbf{1}_n = \mathbf{0}_{n-1}$. ***Let B be a real,*** $n \times (n-1)$, ***and such that AB =***

I_{n-1}. *If* $\|\cdot\|^{(n-1)}$ *is any vector norm on* $\mathbb{R}^{n-1}$, *then for* $\mathbf{y}^T \in \mathbb{H}^n$, $\|\mathbf{y}^T B\|^{(n-1)}$ *defines a norm on the linear subspace* $\mathbb{H}^n$ *of* $\mathbb{R}^n$.

Proof. The only property of norms which doesn't follow immediately from the fact that $\|\cdot\|^{(n-1)}$ is a norm on $\mathbb{R}^{n-1}$ is

$$\|\mathbf{y}^T B\|^{(n-1)} = 0 \quad \Rightarrow \quad \mathbf{y}^T = \mathbf{0}^T, \quad \mathbf{y}^T \in \mathbb{H}^n.$$

Now, $\mathbb{H}^n$ is the subspace of $\mathbb{R}^n$ orthogonal to the subspace spanned by $\mathbf{1}_n^T$. Since $A\mathbf{1}_n = \mathbf{0}$, and the rows of A are linearly independent, they span $\mathbb{H}^n$, and so for some $\mathbf{x}^T \in \mathbb{R}^{n-1}$ we have $\mathbf{y}^T = \mathbf{x}^T A$ and $\mathbf{y}^T \neq \mathbf{0}_n^T \Leftrightarrow \mathbf{x}^T \neq 0_{n-1}^T$. Further, $\mathbf{y}^T B = \mathbf{x}^T AB = \mathbf{x}^T$, since $AB = I_{n-1}$. Thus $\|\mathbf{y}^T B\|^{(n-1)} = 0 \Rightarrow \mathbf{y}^T B = \mathbf{0}_{n-1}^T \Rightarrow \mathbf{x}^T = \mathbf{0}_{n-1}^T \Rightarrow \mathbf{y}^T = \mathbf{0}_n^T$, as required. ■

Thus we may use, for any norm $\|\cdot\|^{(n-1)}$ on $\mathbb{R}^{n-1}$ and fixed B satisfying the conditions of Theorem 2, the class of ergodicity coefficients

$$\tau(P) = \sup\{\|\mathbf{y}^T PB\|^{(n-1)} : \mathbf{y}^T \in \mathbb{H}^n, \|\mathbf{y}^T B\|^{(n-1)} = 1\}$$

to have available all the properties deriving from (1.1). Further, since $\mathbf{y}^T \in \mathbb{H}^n \Leftrightarrow \mathbf{y}^T = \mathbf{x}^T A$, $\mathbf{x}^T \in \mathbb{R}^{n-1}$, it follows that

$$\tau(P) = \sup\{\|\mathbf{x}^T APB\|^{(n-1)}, \|\mathbf{x}^T\|^{(n-1)} = 1\}$$

since $AB = I_{n-1}$, so that this class of coefficients is given simply by

$$\tau(P) = \|APB\|^{(n-1)}, \qquad P \in S_n, \tag{2.1}$$

for fixed B and A described in Theorem 2.

3. CONSTRUCTION OF A AND B

A general construction of appropriate A and B is as follows. Let V be an $(n-1) \times (n-1)$ nonsingular real matrix, and write $\mathbf{a} = -V\mathbf{1}_{n-1}$. Then put

$$A = (\mathbf{a}, V), \qquad B = \begin{bmatrix} \mathbf{0}^T \\ V^{-1} \end{bmatrix}. \tag{3.1}$$

These satisfy the conditions of Theorem 2 and may be used in (2.1). Then by writing P in partitioned form:

$$P = \begin{bmatrix} \mathbf{p}_1^T \\ \tilde{P} \end{bmatrix},$$

we find

$$AP = \mathbf{a}\mathbf{p}_1^T + V\tilde{P} = -V\mathbf{1}_{n-1}\mathbf{p}_1^T + V\tilde{P} = V\left(\tilde{P} - \mathbf{1}_{n-1}\mathbf{p}_1^T\right),$$

so that $APB = VP^{(1)}V^{-1}$, where

$$P^{(1)} = \{p_{ij} - p_{1j}\}, \qquad i,j = 2,\dots,n,$$

whence with the choice (3.1),

$$\tau(P) = \|VP^{(1)}V^{-1}\|^{(n-1)}. \tag{3.2}$$

Other possible constructions of A and B lead to the coefficients

$$\tau(P) = \|VP^{(n)}V^{-1}\|^{(n-1)}, \qquad \tau(P) = \|VP^{(k)}V^{-1}\|^{(n-1)}, \tag{3.3}$$

where $2 \leqslant k \leqslant n$, and

$$P^{(n)} = \{p_{ij} - p_{nj}\}, \quad i,j = 1,\dots,n-1;$$

$$P^{(k)} = \{p_{ij} - p_{kj}\}, \quad i,j = 1,\dots,n \quad (i,j \neq k).$$

4. EXAMPLES

EXAMPLE 1. Take $V = I_{n-1}$ in (3.2) and the l_1 norm for $\|\cdot\|^{(n-1)}$: this yields

$$\tau^1(P) = \max_{i=2,\dots,n} \sum_{j=2}^{n} |p_{ij} - p_{1j}|$$

in the notation of (1.3). More generally, using (3.3) with the same V and norm produces (1.3). Alternatively, if J is a permutation matrix which moves the kth entry to first position in the manner $(1,\ldots,k,\ldots,n) \to (k,1,\ldots,k-1,k+1,\ldots,n)$, then $\bar{\tau}_J(P) = \tau^1(JPJ^{-1}) = \tau^k(P)$.

EXAMPLE 2. Using (3.3) and

$$V = \begin{bmatrix} 1 & -1 & & & 0 \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ 0 & & & & 1 \end{bmatrix}, \qquad V^{-1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ & 1 & \cdots & 1 \\ & & \ddots & \vdots \\ 0 & & & 1 \end{bmatrix},$$

the (i,j) entry of $VP^{(n)}V^{-1}$ is

$$\sum_{k=1}^{j} (p_{i,k} - p_{i+1,k}), \qquad i,j = 1,2,\ldots,n-1,$$

and using the l_1 norm for $\|\cdot\|^{(n-1)}$ results in (1.4).

EXAMPLE 3. In Examples 1 and 2 use the l_∞ norm for $\|\cdot\|^{(n-1)}$ to obtain

$$\max_{j=2,\ldots,n} \sum_{i=2}^{n} |p_{ij} - p_{1j}|, \qquad \max_{j=1,\ldots,n-1} \sum_{i=1}^{n-1} \left| \sum_{k=1}^{j} (p_{i,k} - p_{i+1,k}) \right|.$$

EXAMPLE 4. In Example 2 replace V by V^T. Then $V^T P^{(n)} (V^{-1})^T$ has (i,j) entry

$$\begin{cases} \displaystyle\sum_{s=j}^{n-1} (p_{1s} - p_{ns}) & \text{at} \quad i = 1, \quad j = 1,\ldots,n-1, \\ \displaystyle\sum_{s=j}^{n-1} (p_{i,s} - p_{i-1,s}) & \text{at} \quad i = 2,\ldots,n-1, \quad j = 1,\ldots,n-1. \end{cases}$$

EXAMPLE 5. In Example 2 replace V by V^{-1}. Then $V^{-1}P^{(n)}(V^{-1})^T$ has (i,j) entry

$$\begin{cases} \sum_{k=i}^{n-1} (p_{k1} - p_{n1}), & i = 1,\dots,n-1, \quad j = 1, \\ \sum_{k=i}^{n-1} \{p_{kj} - p_{nj} - (p_{k,j-1} - p_{n,j-1})\}, & i = 1,\dots,n-1, \\ & j = 2,\dots,n-1. \end{cases}$$

5. EXTENSIONS

Theorem 1 can be reformulated for $n \times n$ irreducible nonnegative matrices T with Perron-Frobenius eigenvalue ρ and corresponding right eigenvector $\mathbf{w}$. With stochastic matrices P the irreducibility assumption is unnecessary, and we also have the probabilistically interesting property (A), which does not extend to irreducible T_1, T_2 which have different right eigenvectors. Define $\tau(T)$ by (1.1) again, where now, however,

$$\mathbb{H}^n = \{\mathbf{x}^T : \mathbf{x}^T \in \mathbb{R}^n, \quad \mathbf{x}^T\mathbf{w} = 0\}.$$

Since $\mathbf{x}^T T\mathbf{w} = \rho\mathbf{x}^T\mathbf{w} = 0$ for any $\mathbf{x}^T \in \mathbb{H}^n$, we may consider T as an operator on $\mathbb{H}^n$.

THEOREM 1'. *For each vector norm $\|\cdot\|_{\mathbb{H}^n}$, there is a vector norm on $\mathbb{R}^n$ such that*

$$|\lambda| \leqslant \tau(T) = \|T\|_{\mathbb{H}^n}$$

for any eigenvalue $\lambda \neq \rho$, where T is any irreducible nonnegative matrix.

Proof. Since $\mathbf{w}$ spans the orthogonal complement of $\mathbb{H}^n$ in $\mathbb{R}^n$, we may write uniquely for any $\mathbf{v}^T \in \mathbb{R}^n$

$$\mathbf{v}^T = \mathbf{h}^T + \alpha\mathbf{w}^T,$$

where $\alpha \in \mathbb{R}^1$, $\mathbf{h}^T \in \mathbb{H}^n$. Define the norm $|||\cdot|||$ for $\mathbf{v}^T \in \mathbb{R}^n$ in exactly the same way as in Theorem 1, and proceed as in that theorem with T replacing P, to obtain $\tau(T) = \|T\|_{\mathbb{H}^n}$. The result $|\lambda| \leqslant \tau(T)$ now follows directly from the fundamental Theorem 3.1 of Rothblum and Tan (1985). ■

Likewise we could state and prove a Theorem 2′ in precisely the same way Theorem 2, with $\mathbf{w}$ replacing $\mathbf{1}_n$. Equation (2.1) holds with T replacing P.

The specific extension used of a norm on $\mathbb{H}^n$ to that on $\mathbb{R}^n$ is due to a referee.

REFERENCES

Lesanovsky, A. 1990. Coefficients of ergodicity generated by non-symmetrical vector norms, *Czechoslovak Math. J.* 40:284–294.

Rhodius, A. 1993. On explicit forms for ergodicity coefficients, *Linear Algebra Appl.*, to appear.

Rothblum, U. G. and Tan, C. P. 1985. Upper bounds on the maximum modulus of subdominant eigenvalues of nonnegative matrices, *Linear Algebra Appl.* 66:45–86.

Seneta, E. 1979. Coefficients of ergodicity: Structure and applications. *Adv. Appl. Probab.* 11:576–590.

Seneta, E. 1981. *Nonnegative Matrices and Markov Chains*, 2nd ed., Springer-Verlag, New York.

Received 27 May 1992; final manuscript accepted 10 August 1992