

Antonín Lešanovský

Coefficients of ergodicity generated by non-symmetrical vector norms

Czechoslovak Mathematical Journal, Vol. 40 (1990), No. 2, 284–294

Persistent URL: <http://dml.cz/dmlcz/102380>

Terms of use:

© Institute of Mathematics AS CR, 1990

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://dml.cz>

COEFFICIENTS OF ERGODICITY GENERATED
BY NON-SYMMETRICAL VECTOR NORMS

ANTONÍN LEŠANOVSKÝ, Praha

(Received June 6, 1988)

Coefficients of ergodicity have proved to be a useful tool for the investigation of both homogeneous and inhomogeneous Markov chains. It has been also realized that they can be taken as upper bounds for the absolute value of the largest non-unit eigenvalue of stochastic matrices. There is an extensive literature in this field. The list of references below forms merely a certain sample in this respect.

Their applications to different problems caused, unfortunately, a certain ambiguity as to the notion coefficient of ergodicity itself. The following two definitions have appeared:

(i) Any scalar function $\tau(\cdot)$ continuous on the set \mathcal{S}_n of all $n \times n$ stochastic matrices (treated as points in R^{n^2}) and satisfying $0 \leq \tau(P) \leq 1$ for $P \in \mathcal{S}_n$ is called a *coefficient of ergodicity* – see [16], page 136.

(ii) Let $d(\cdot, \cdot)$ be any metric on the set $D_n = \{x; x \in R^n, x_i \geq 0, \sum_{i=1}^n x_i = 1\}$. Then the quantity

$$\tau(P) = \sup_{\substack{y, z \in D_n \\ y \neq z}} \frac{d(yP, zP)}{d(y, z)} \quad \text{for } P \in \mathcal{S}_n$$

is called a *coefficient of ergodicity* – see [15]. (All vectors in this paper are assumed to be row vectors.)

It is obvious that a function $\tau(P) = p_{11}$ fulfils the demands of (i) but cannot be generated on \mathcal{S}_n ($n \geq 2$) by any metric $d(\cdot, \cdot)$ on $D_n \times D_n$. On the other hand, a metric $d(y, z) = \|y - z\|_\infty$, where $\|\cdot\|_\infty$ is the l_∞ -norm, i.e. $\|x\|_\infty = \max\{|x_i|; i = 1, \dots, n\}$, generates by (ii) a coefficient of ergodicity denoted in this paper by $\tau_\infty(\cdot)$ which possesses by [17] and [10] the property

$$\sup_{P \in \mathcal{S}_n} \tau_\infty(P) = \begin{cases} n/2 & \text{for } n \text{ even,} \\ (n-1)/2 & \text{for } n \text{ odd, } n \geq 3. \end{cases}$$

Thus, for $n \geq 4$ there exists a $P \in \mathcal{S}_n$ such that $\tau_\infty(P) > 1$. This implies that $\tau_\infty(P)$ is not a coefficient of ergodicity in the sense of (i). These two examples show that none of the above definitions is fully compatible with the other.

In this paper, the following restricted version of (ii) will be considered:

Definition. Let $\|\cdot\|$ be any norm on R^n . Let $H_n = \{z; z \in R^n, z \neq 0, \sum_{i=1}^n z_i = 0\}$. The quantity

$$(1) \quad \tau(P) = \sup_{\substack{x \in H_n \\ \|x\|=1}} \|xP\| \quad \text{for } P \in \mathcal{S}_n$$

is called a *coefficient of ergodicity generated by the vector norm $\|\cdot\|$* .

Note that $d(yP, zP) = \|(y - z)P\|$ holds if the metric $d(\cdot, \cdot)$ corresponds to the norm $\|\cdot\|$, and that $y - z \in H_n$ whenever $y, z \in D_n, y \neq z$. Hence the definition is compatible with (ii).

It should be mentioned that all coefficients of ergodicity studied by now except for Birkhoff's one (for its definition see [16], page 83) are of the form (1). Attention has been paid to those generated by the l_p -norms

$$(2) \quad \|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \quad \text{for } x \in R^n, \quad p \in [1; \infty),$$

$$(3) \quad \|x\|_\infty = \max \{|x_i|; i = 1, 2, \dots, n\} \quad \text{for } x \in R^n,$$

in particular for $p = 1, 2$, or ∞ . Three main topics have been discussed – their functional form in terms of entries of a stochastic matrix (e.g. [3], [11], [15], [17]), their properties (e.g. [10], [11] or [15]) and applications. The applications concern the Markov chains [4], [8], [13] or the spectrum localization problem [1] and [11], etc.

The present paper deals with a class of coefficients of ergodicity generated by a non-symmetrical generalization of the l_1 -norm

$$(4) \quad \|x\|_{N_a} = \sum_{i=1}^n a_i |x_i| \quad \text{for } x \in R^n,$$

where $a = (a_1; \dots; a_n) \in R^n$ and $a_i > 0$ for all $i = 1, \dots, n$. It touches all the three items mentioned in the last paragraph. Moreover, we shall focus on the quantity $\sup_{P \in \mathcal{S}_n} \tau(P)$ as a function of τ (or, equivalently, as a function of the norm which generates τ). Theorem 3 provides a characterization of the l_1 -norm. It states that the only coefficient of ergodicity $\tau(\cdot)$ generated by a vector norm on R^n which fulfils the requirements of (i) is $\tau_1(\cdot)$.

1. NOTATION

n – an integer greater than 1

\mathcal{N} = $\{1; 2; \dots; n\}$

\mathfrak{R} = $\{(i, j); i, j \in \mathcal{N}, i \neq j\}$

\mathcal{S}_n – the set of all $n \times n$ stochastic matrices

R^n – the set of all row vectors $z = (z_1, \dots, z_n)$ of real numbers

C^n – the set of all row vectors $z = (z_1, \dots, z_n)$ of complex numbers
 $R_+^n = \{x; x \in R^n, x = (x_1, \dots, x_n), x_i > 0 \text{ for each } i \in \mathcal{N}\}$
 $H_n = \{x; x \in R^n, x \neq 0, \sum_{i=1}^n x_i = 0\}$
 $e^{(i)}$, for $i \in \mathcal{N}$ – an element of R^n the components of which are given by $e_i^{(i)} = 1$
 and $e_j^{(i)} = 0$ if $j \neq i$
 $u^{(i,j)} = e^{(i)} - e^{(j)}$ for $(i, j) \in \mathfrak{R}$
 $U_n = \{u^{(i,j)}; (i, j) \in \mathfrak{R}\}$
 $x^+ = \max\{x; 0\}$
 $x^- = \max\{-x; 0\}$
 I – identity matrix
 $\|\cdot\|_p$, for $p \in [1; \infty)$ or $p = \infty$ – the l_p – norm defined by (2) and (3)
 $\|\cdot\|_{N_a}$, for $a \in R_+^n$ – the norm defined by (4)
 $\tau_p(\cdot)$ – the coefficient of ergodicity generated by $\|\cdot\|_p$
 $\tau_{N_a}(\cdot)$ – the coefficient of ergodicity generated by $\|\cdot\|_{N_a}$
 N – the set of all positive integers

2. FORMULA FOR $\tau_{N_a}(P)$ IN TERMS OF THE ENTRIES OF THE MATRIX P

The well-known formula for $\tau_1(P)$ states that for calculation of $\tau_1(P)$ it is sufficient to consider (normalized) vectors $u^{(i,j)}$, for $i, j \in \mathcal{N}$, $i \neq j$. This important property is preserved by the coefficients of ergodicity generated by the generalized l_1 -norms $\|\cdot\|_{N_a}$, for each $a \in R_+^n$.

Theorem 1. *Let $n \in N - \{1\}$, $P \in \mathcal{S}_n$ and $a \in R_+^n$. Then*

$$(5) \quad \tau_{N_a}(P) = \max_{y \in U_n} \frac{\|yP\|_{N_a}}{\|y\|_{N_a}}.$$

This result has been proved in [11] – Theorem 4.1, where the coefficient of ergodicity τ_{N_a} has been denoted by τ_1^P .

3. CONTRIBUTION OF THE COEFFICIENTS OF ERGODICITY τ_{N_a} TO THE SPECTRUM LOCALIZATION

Let $P \in \mathcal{S}_n$ have eigenvalues $\lambda_1(P), \dots, \lambda_n(P)$ (repeated as many times as their multiplicities indicate). We put

$$\mu(P) = \max\{|\lambda_i(P)|; i \in \mathcal{N}, \lambda_i(P) \neq 1\}$$

if there exists an $i \in \mathcal{N}$ such that $\lambda_i(P) \neq 1$, and

$$(6) \quad \mu(P) = 1$$

if all eigenvalues of P are equal to 1, i.e. if $P = I$. The quantity $\mu(P)$ plays an im-

portant role in numerical mathematics, theory of Markov chains, etc. Its relation to the values of the coefficients of ergodicity $\tau_{N_a}(P)$ is given in

Proposition. *Let $n \in N - \{1\}$, $a \in R_+^n$, and $P \in \mathcal{S}_n$. Then*

$$(7) \quad \tau_{N_a}(P) \geq \mu(P),$$

i.e. the coefficients of ergodicity $\tau_{N_a}(P)$, $a \in R_+^n$, are upper bounds of $\mu(P)$ for each $P \in \mathcal{S}_n$.

Proof. If $P = I$ then obviously $\tau_{N_a}(P) = \mu(P) = 1$ holds for each $a \in R_+^n$. For $P \neq I$ define a function $f_a: C^n \rightarrow R$ by

$$(8) \quad f_a(z) = \max_{(i,j) \in \mathfrak{R}} \frac{|z_i - z_j|}{a_i + a_j}.$$

It is well-known that any vector $x \in H_n$ can be expressed in the form

$$(9) \quad x = \sum_{i=1}^n x_i^+ e^{(i)} - \sum_{j=1}^n x_j^- e^{(j)} = \frac{2}{\|x\|_1} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n x_i^+ x_j^- u^{(i,j)},$$

so that

$$(10) \quad \|x\|_{N_a} = \frac{2}{\|x\|_1} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n x_i^+ x_j^- \|u^{(i,j)}\|_{N_a}.$$

Thus

$$|xz'| \leq \frac{2}{\|x\|_1} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n x_i^+ x_j^- (a_i + a_j) \frac{|z_i - z_j|}{a_i + a_j} \leq f_a(z) \|x\|_{N_a}$$

for all $x \in H_n$ and $z \in C^n$. Further, by Theorem 1 we obtain

$$\begin{aligned} f_a(zP) &= \max_{(i,j) \in \mathfrak{R}} \frac{1}{a_i + a_j} \left| \sum_{k=1}^n (p_{ik} - p_{jk}) z_k \right| = \\ &= \max_{(i,j) \in \mathfrak{R}} \frac{1}{a_i + a_j} |u^{(i,j)} P z'| \leq f_a(z) \max_{(i,j) \in \mathfrak{R}} \frac{\|u^{(i,j)} P\|_{N_a}}{\|u^{(i,j)}\|_{N_a}} = f_a(z) \tau_{N_a}(P). \end{aligned}$$

It remains to take a right eigenvector \hat{z} of the matrix P corresponding to its eigenvalue λ which satisfies $|\lambda| = \mu(P)$ and $\lambda \neq 1$. We find that $f_a(\hat{z}) > 0$. Indeed, $f_a(z) \geq 0$ for all $a \in R_+^n$, $z \in C^n$, and $f_a(\hat{z}) = 0$ would imply that $\hat{z}_1 = \hat{z}_2 = \dots = \hat{z}_n$ so that $P\hat{z} = \hat{z} \neq \lambda\hat{z}$. Thus,

$$(11) \quad \tau_{N_a}(P) \geq \frac{f_a(P\hat{z})}{f_a(\hat{z})} = \frac{f_a(\lambda\hat{z})}{f_a(\hat{z})} = |\lambda| = \mu(P).$$

As a matter of fact, Theorem 3.1 of [11] states that each coefficient of ergodicity τ generated by a vector norm $\|\cdot\|$ on R^n fulfils $\tau(P) \geq \mu(P)$ for each stochastic matrix P . This general result was proved in [11] by constructing a norm $\|\|\cdot\|\|_a$ on C^n which

was an extension of the norm $\|\cdot\|$ on R^n . The proof of the Proposition above shows that considering the special case of the norms $\|\cdot\|_{N_a}$ it is sufficient to use much simpler technique.

The difference $\tau(P) - \mu(P)$ forms a suitable utility criterion (for a given stochastic matrix P) of the coefficient τ . A comparison of the quality of the estimates of $\mu(P)$ based on τ_{N_a} and τ_p is given in the examples below. It is worth-while to note in this setting that for any $\varepsilon > 0$ and for any stochastic matrix $P \in \mathcal{S}_n$ there exists a norm $\|\cdot\|$ on R^n such that $\tau_{\|\cdot\|}(P) - \mu(P) < \varepsilon$ — see Theorem C.1 of [11]. The form of the norm $\|\cdot\|$ depends on the matrix P in a rather complicated manner. A coefficient of ergodicity τ can, however, be used efficiently only if it is possible to express the value of $\tau(P)$ analytically in terms of entries of P . This condition is fulfilled for τ_1, τ_2 [15], τ_∞ [17] and for τ_{N_a} , where $a \in R_+^n$. (We have not mentioned Birkhoff's coefficient τ_B — see e.g. [16] because it cannot be generated by a vector norm on R^n — cf. Corollary 2 below.) The list of applicable coefficients of ergodicity is hence really short. An important advantage of the class $\{\tau_{N_a}; a \in R_+^n\}$ is that, although it keeps much of the simplicity of τ_1 , it exhibits a certain flexibility. This feature can be used when looking for an adequate coefficient of ergodicity to a given stochastic matrix. This is demonstrated by the following example.

Example 1. Let $n = 3$ and let

$$(12) \quad P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Easy calculations yield that $\lambda_1(P) = 1$, $\lambda_2(P) = \lambda_3(P) = 0$, i.e. $\mu(P) = 0$, and $u^{(1,2)}P = u^{(1,3)}P = (0, 1, -1)$, $u^{(2,3)}P = \mathbf{0}$, so that by Theorem 1 we have

$$\tau_{N_a}(P) = \max \left\{ \frac{a_2 + a_3}{a_1 + a_2}, \frac{a_2 + a_3}{a_1 + a_3} \right\} \quad \text{for } a \in R_+^n.$$

We put $a(k) = (k, 1, 1)$ for $k \in N$. Then

$$\lim_{k \rightarrow \infty} \tau_{N_{a(k)}}(P) = \lim_{k \rightarrow \infty} \frac{2}{k+1} = 0 = \mu(P).$$

On the other hand, $\|u^{(1,2)}\|_p = \|u^{(2,3)}\|_p$ obviously holds for each $p \in [1; \infty)$ and $p = \infty$, so that

$$\tau_p(P) \geq \frac{\|u^{(1,2)}P\|_p}{\|u^{(1,2)}\|_p} = \frac{\|u^{(2,3)}\|_p}{\|u^{(1,2)}\|_p} = 1$$

is true for each $p \in [1; \infty)$ and $p = \infty$. (Moreover, $\tau_B(P) = 1$ by the definition (see e.g. [16], page 83) of Birkhoff's coefficient.)

Thus, all classical coefficients of ergodicity provide poor estimates of $\mu(P)$ for P given by (12). For determining the value of $\mu(P)$ it is, however, possible to restrict

oneself to the class $\{\tau_{N_a}; a \in R_+^n\}$. Vaguely speaking, it makes it possible to apply an appropriately non-symmetrical tool to non-symmetrical stochastic matrices.

The next example shows, however, that the class $\{\tau_{N_a}; a \in R_+^n\}$ is not generally superior to $\{\tau_p; p \in [1; \infty) \text{ or } p = \infty\}$.

Example 2. Let $n = 3$ and let

$$(13) \quad P = \begin{pmatrix} 2/3 & 1/3 & 0 \\ 1/3 & 0 & 2/3 \\ 0 & 2/3 & 1/3 \end{pmatrix}.$$

We find that $\lambda_1(P) = 1$, $\lambda_2(P) = \frac{1}{3}\sqrt{3}$ and $\lambda_3(P) = -\frac{1}{3}\sqrt{3}$, i.e. $\mu(P) = \frac{1}{3}\sqrt{3}$. Further,

$$u^{(1,2)}P = \frac{1}{3}(1, 1, -2),$$

$$u^{(1,3)}P = \frac{1}{3}(2, -1, -1),$$

$$u^{(2,3)}P = \frac{1}{3}(1, -2, 1),$$

so that by Theorem 1 we obtain

$$\begin{aligned} \tau_{N_a}(P) &= \max \left\{ \frac{\frac{1}{3}(a_1 + a_2 + 2a_3)}{a_1 + a_2}; \frac{\frac{1}{3}(2a_1 + a_2 + a_3)}{a_1 + a_3}; \frac{\frac{1}{3}(a_1 + 2a_2 + a_3)}{a_2 + a_3} \right\} = \\ &= \frac{1}{3} + \frac{1}{3} \max \left\{ \frac{2a_3}{a_1 + a_2}; \frac{a_1 + a_2}{a_1 + a_3}; \frac{a_1 + a_2}{a_2 + a_3} \right\}. \end{aligned}$$

The system

$$2a_3 < a_1 + a_2, \quad a_1 + a_2 < a_1 + a_3, \quad a_1 + a_2 < a_2 + a_3$$

has no solution. It means that

$$\tau_{N_a}(P) \geq \frac{2}{3} \quad \text{for each } a \in R_+^n,$$

and

$$\tau_{N_a}(P) - \mu(P) \geq \frac{1}{3}(2 - \sqrt{3}) \quad \text{for each } a \in R_+^n.$$

Let us remark that the ideal coefficient of ergodicity for P given by (13) is τ_2 . A vector $x \in H_3$ such that $\|x\|_2 = 1$ can be written in the form $x = (y, z, -y - z)$, where $y, z \in R$ fulfil $2(y^2 + z^2 + yz) = 1$. Further,

$$xP = \frac{1}{3}(2y + z, -y - 2z, z - y)$$

and

$$\|xP\|_2 = \frac{1}{3}\sqrt{(6(y^2 + z^2 + yz))} = \frac{1}{3}\sqrt{3} = \mu(P).$$

4. THE RANGE OF COEFFICIENTS OF ERGODICITY

The convexity of the set \mathcal{S}_n implies that the range of a coefficient of ergodicity τ is the closed interval $[0; \max_{P \in \mathcal{S}_n} \tau(P)]$. Indeed, let $\bar{P} \in \mathcal{S}_n$ be such that $\tau(\bar{P}) = \max_{P \in \mathcal{S}_n} \tau(P)$.

(for details concerning the existence of this maximum see [10]) and let $C \in \mathcal{S}_n$ be the matrix with all its entries equal to $1/n$. Put $P^{(\alpha)} = \alpha \cdot \bar{P} + (1 - \alpha) \cdot C$ for all $\alpha \in [0; 1]$. We obtain $xP^{(\alpha)} = \alpha x\bar{P} + (1 - \alpha) \cdot xC = \alpha \cdot x\bar{P}$ for each $x \in H_n$, so that $\tau(P^{(\alpha)}) = \alpha \tau(\bar{P})$.

It is well-known that $\tau_1(P) \leq 1$ for all $P \in \mathcal{S}_n$. Taking into account that $\tau(I) = 1$ holds for each coefficient of ergodicity τ we conclude that

$$\max_{P \in \mathcal{S}_n} \tau_1(P) = 1.$$

Hence, the situation in case of τ_1 is very simple. On the other hand, there has been probably only one attempt [10] at investigating maximal possible values of other coefficients of ergodicity. The explicit results of [10] concern those generated by the l_p -norms and state that

$$(14) \quad \max_{P \in \mathcal{S}_n} \tau_p(P) = \begin{cases} \left(\frac{n}{2}\right)^{1-(1/p)} & \text{if } n \text{ is even,} \\ \left(\frac{1}{2}\right)^{1-(1/p)} \left[\frac{2}{(n+1)^{1-p} + (n-1)^{1-p}} \right]^{1/p} & \text{if } n \text{ is odd} \end{cases}$$

for $p \in [1; \infty)$, and

$$(15) \quad \max_{P \in \mathcal{S}_n} \tau_\infty(P) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even,} \\ \frac{n-1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

Observing the formulas (14) and (15) we find that for a given dimension n of stochastic matrices the values of $\max_{P \in \mathcal{S}_n} \tau_p(P)$ do not exceed $(n+1)/2$ for $p \in [1; \infty)$ or $p = \infty$.

This phenomenon of boundedness disappears if we take into consideration all the coefficients of ergodicity generated by vector norms. Moreover, we shall prove that for each real number $\omega \geq 1$ there exists such a coefficient of ergodicity τ generated by a vector norm on R^n that

$$\max_{P \in \mathcal{S}_n} \tau(P) = \omega.$$

We shall start with

Lemma. *Let $P \in \mathcal{S}_2$ and let $\lambda_1(P)$ and $\lambda_2(P)$ be its eigenvalues such that $1 = \lambda_1(P) \geq |\lambda_2(P)|$. Then the equality*

$$(16) \quad \frac{\|xP\|}{\|x\|} = |\lambda_2(P)|$$

holds for each $x \in H_2$ and for each vector norm $\|\cdot\|$ on R^2 .

Proof. A matrix $P \in \mathcal{S}_2$ and a vector $x \in H_2$ are obviously of the form

$$P = \begin{pmatrix} b & 1 - b \\ c & 1 - c \end{pmatrix},$$

$$x = d \cdot (1, -1)$$

where $b, c \in [0; 1]$ and $d \in \mathbb{R} - \{0\}$. Thus

$$xP = d(b - c)(1, -1),$$

so that for each vector norm $\|\cdot\|$ on \mathbb{R}^2 we obtain

$$\frac{\|xP\|}{\|x\|} = \frac{|d(b - c)| \cdot \|(1, -1)\|}{|d| \cdot \|(1, -1)\|} = |b - c|.$$

It remains to remark that the eigenvalues of P are $\lambda_1(P) = 1$ and $\lambda_2(P) = b - c$.

Theorem 2. Let $n \geq 3$. Then

$$(17) \quad \max_{P \in \mathcal{S}_n} \tau_{N_a}(P) = \frac{a_{j(a,1)} + a_{j(a,2)}}{a_{j(a,n)} + a_{j(a,n-1)}} \quad \text{for each } a \in \mathbb{R}_+^n,$$

where $j(a, \cdot)$ is a one-to-one mapping of the set \mathcal{N} onto \mathcal{N} such that

$$(18) \quad a_{j(a,1)} \geq a_{j(a,2)} \geq \dots \geq a_{j(a,n)}.$$

Proof. We apply Theorem 1 of [10] to the coefficient of ergodicity τ_{N_a} :

$$(19) \quad \max_{P \in \mathcal{S}_n} \tau_{N_a}(P) = \max_{x \in H_n} \frac{1}{\|x\|_{N_a}} \max_{(A_1, \dots, A_n) \in \mathcal{A}} \left\| \left(\sum_{i \in A_1} x_i; \dots; \sum_{i \in A_n} x_i \right) \right\|_{N_a},$$

where \mathcal{A} denotes the set of all decompositions of the set \mathcal{N} into n disjoint subsets A_1, \dots, A_n . We obtain

$$\begin{aligned} \frac{1}{\|x\|_{N_a}} \left\| \left(\sum_{i \in A_1} x_i; \dots; \sum_{i \in A_n} x_i \right) \right\|_{N_a} &= \frac{\sum_{j=1}^n a_j \left| \sum_{i \in A_j} x_i \right|}{\sum_{j=1}^n a_j |x_j|} \leq \\ &\leq \frac{a_{j(a,1)} \left| \sum_{i \in A_{j(a,1)}} x_i \right| + a_{j(a,2)} \sum_{\substack{j=1 \\ j \neq j(a,1)}}^n \left| \sum_{i \in A_j} x_i \right|}{a_{j(a,n)} |x_{j(a,n)}| + a_{j(a,n-1)} \sum_{\substack{j=1 \\ j \neq j(a,n)}}^n |x_j|} \leq \\ &\leq \frac{\frac{1}{2} [a_{j(a,1)} + a_{j(a,2)}] \sum_{j=1}^n \left| \sum_{i \in A_j} x_i \right|}{\frac{1}{2} [a_{j(a,n)} + a_{j(a,n-1)}] \sum_{j=1}^n |x_j|} \leq \frac{a_{j(a,1)} + a_{j(a,2)}}{a_{j(a,n)} + a_{j(a,n-1)}} \end{aligned}$$

because

$$\sum_{\substack{j=1 \\ j \neq j(a,1)}}^n \left| \sum_{i \in A_j} x_i \right| \geq \left| \sum_{\substack{i=1 \\ i \notin A_j(a,1)}}^n x_i \right| = \left| \sum_{i \in A_j(a,1)} x_i \right|$$

and

$$\sum_{\substack{j=1 \\ j \neq j(a,n)}}^n |x_j| \geq \left| \sum_{\substack{j=1 \\ j \neq j(a,n)}}^n x_j \right| = |x_{j(a,n)}|.$$

To complete the proof, it is sufficient to introduce $\hat{P} = (\hat{p}_{ij})_{i,j=1}^n \in \mathcal{S}_n$ and $\hat{x} \in H_n$ such that the fraction $\|\hat{x}\hat{P}\|_{N_a} / \|\hat{x}\|_{N_a}$ equals the right-hand side of (17). To this aim, we put

$$\begin{aligned} b &= j(a, 1), \\ c &= j(a, 2), \\ d &= j(a, n - 1), \\ e &= j(a, n), \\ \hat{p}_{ab} &= \hat{p}_{ic} = 1 \quad \text{for } i \in \mathcal{N} - \{d\}, \\ \hat{p}_{ij} &= 0 \quad \text{otherwise,} \end{aligned}$$

and

$$\hat{x} = u^{(d,e)}.$$

We find that

$$\hat{x}\hat{P} = u^{(b,c)},$$

so that

$$\frac{\|\hat{x}\hat{P}\|_{N_a}}{\|\hat{x}\|_{N_a}} = \frac{a_b + a_c}{a_d + a_e} = \frac{a_{j(a,1)} + a_{j(a,2)}}{a_{j(a,n-1)} + a_{j(a,n)}}.$$

Corollary 1. *The set of values of $\max_{P \in \mathcal{S}_n} \tau(P)$ when τ varies over the set of all coefficients of ergodicity generated by a vector norm on R^n is equal to $\{1\}$ if $n = 2$ and $[1; \infty)$ if $n \geq 3$.*

Take an arbitrary $n \geq 3$ and $\omega \in (1; \infty)$. Corollary 1 above guarantees that there exists a vector norm $\|\cdot\|$ such that the corresponding $\tau_{\|\cdot\|}$ fulfils

$$(20) \quad \max_{P \in \mathcal{S}_n} \tau_{\|\cdot\|}(P) = \omega.$$

Such a vector norm on R^n is no doubt far from unique. The non-uniqueness of vector norms can be regarded essential from the point of view of the coefficients of ergodicity only if they differ at an element of the set H_n . Examples of essentially different vector norms in the case considered are e.g. the norms $\|\cdot\|_{N_b}$ and $\|\cdot\|_{N_c}$ on R^3 , where $b = (2\omega - 1, 1, 1)$ and $c = (\omega^2, \omega, 1)$.

A completely different situation arises in the case of $\omega = 1$.

Theorem 3. *If a coefficient of ergodicity τ generated by a vector norm $\|\cdot\|$ on R^n*

fulfils

$$(21) \quad \sup_{P \in \mathcal{S}_n} \tau(P) = 1$$

then

(i) the norm $\|\cdot\|$ coincides on the set H_n (up to a multiplicative constant) with the l_1 -norm;

(ii) the relation $\tau(P) = \tau_1(P)$ holds for each $P \in \mathcal{S}_n$.

Proof. First, let $n = 2$. The one-dimensionality of the vector space H_2 implies that any norm on H_2 is determined by its value at $x = (1, -1)$, i.e. it equals a constant multiple of the l_1 -norm on H_2 .

Now, let $n \geq 3$. We know from (21) that

$$(22) \quad \frac{\|xP\|}{\|x\|} \leq 1 \quad \text{for all } x \in H_n \quad \text{and all } P \in \mathcal{S}_n.$$

Applying this fact to all elements of the set \mathcal{U} and to all permutation matrices $P \in \mathcal{S}_n$ we find that the values of $\|u^{(i,j)}\|$ are the same for all $(i, j) \in \mathfrak{N}$. We denote this value by φ . Further, any $x \in H_n$ can be expressed by (9) in the form

$$x = \frac{2}{\|x\|_1} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n x_i^+ x_j^- u^{(i,j)},$$

so that

$$(23) \quad \|x\| \leq \frac{2}{\|x\|_1} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n x_i^+ x_j^- \|u^{(i,j)}\| = \frac{\|x\|_1}{2} \varphi$$

holds for each $x \in H_n$. It remains to prove that $\|x\|$ equals $\frac{1}{2}\|x\|_1 \varphi$ on H_n . Let $P^{(x)} = (p_{ij}^{(x)})_{i,j=1}^n \in \mathcal{S}_n$ be given by

$$\begin{aligned} p_{i1}^{(x)} &= 1 & \text{if } x_i > 0, \\ p_{i2}^{(x)} &= 1 & \text{if } x_i \leq 0, \\ p_{ij}^{(x)} &= 0 & \text{otherwise.} \end{aligned}$$

Then we have

$$(24) \quad \|xP^{(x)}\| = \left\| \left(\sum_{\substack{i=1 \\ x_i > 0}}^n x_i; \sum_{\substack{i=1 \\ x_i \leq 0}}^n x_i; 0; \dots; 0 \right) \right\| = \frac{1}{2} \|x\|_1 \cdot \|u^{(1,2)}\| = \frac{1}{2} \|x\|_1 \cdot \varphi.$$

The relations (22), (23) and (24) imply

$$\|x\| = \frac{\varphi}{2} \|x\|_1 \quad \text{for all } x \in H_n.$$

This completes the proof of the theorem.

Corollary 2. Birkhoff's coefficient τ_B is not generated by a vector norm.

Proof. The coefficient τ_B fulfils (21) but τ_B and τ_1 do not coincide on \mathcal{S}_n .

Remark. An example of a vector norm defined on R^n which is different from $\|\cdot\|_1$ on R^n but coincides with $\|\cdot\|_1$ on H_n is the norm

$$\|x\| = \sum_{i=1}^n |x_i| + \left| \sum_{j=1}^n x_j \right| \quad \text{for } x \in R^n.$$

References

- [1] *F. L. Bauer, E. Deutsch, J. Stoer*: Abschätzungen für die Eigenwerte positiver linearen Operatoren. *Linear Algebra and Appl.* 2 (1969), 275–301.
- [2] *G. Birkhoff*: *Lattice Theory*. Amer. Math. Soc. Colloq. Publics., vol. XXV, Providence, R. I.-3rd edition (1967).
- [3] *R. L. Dobrushin*: Central limit theorem for non-stationary Markov chains I, II. *Theory Prob. Appl.* 1 (1956), 63–80, 329–383 (English translation).
- [4] *J. Hajnal*: Weak ergodicity in non-homogeneous Markov chains. *Proc. Camb. Phil. Soc.* 54 (1958), 233–246.
- [5] *R. A. Horn, Ch. A. Johnson*: *Matrix Analysis*. Cambridge University Press, Cambridge, London, New York, New Rochelle, Melbourne and Sydney (1985).
- [6] *S. Karlin*: *A First Course in Stochastic Processes*. Academic Press, New York and London (1968).
- [7] *D. G. Kendall*: Geometric ergodicity and the theory of queues. In: *Mathematical Methods in the Social Sciences*, K. J. Arrow, S. Karlin, P. Suppes (eds.), Stanford, California (1960).
- [8] *P. Kratochvíl, A. Lešanovský*: A contractive property in finite state Markov chains. *Czechoslovak Math. J.* 35 (110) (1985), 491–509.
- [9] *A. Paz*: *Introduction to Probabilistic Automata*. Academic Press, New York (1971).
- [10] *A. Rhodius*: The maximal value for coefficients of ergodicity. *Stochastic Processes Appl.* 29 (1988), 141–143.
- [11] *U. G. Rothblum and C. P. Tan*: Upper bounds on the maximum modulus of subdominant eigenvalues of nonnegative matrices. *Linear Algebra Appl.* 66 (1985), 45–86.
- [12] *T. A. Сарымсаков*: *Основы теории процессов Маркова*. Государственное издательство технико-теоретической литературы, Москва (1954).
- [13] *T. A. Сарымсаков*: К теории неоднородных цепей Маркова. *Докл. АН УзССР* 8 (1956), 3–7.
- [14] *E. Seneta*: On the historical development of the theory of finite inhomogeneous Markov chains. *Proc. Camb. Phil. Soc.* 74 (1973), 507–513.
- [15] *E. Seneta*: Coefficients of ergodicity: structure and applications. *Adv. Appl. Prob.* 11 (1979), 576–590.
- [16] *E. Seneta*: *Non-negative Matrices and Markov Chains*. Springer-Verlag, New York, Heidelberg and Berlin (1981).
- [17] *C. P. Tan*: A functional form for a particular coefficient of ergodicity. *J. Appl. Prob.* 19 (1982), 858–863.
- [18] *C. P. Tan*: Coefficients of ergodicity with respect to vector norms. *J. Appl. Prob.* 20 (1983), 277–287.
- [19] *D. Vere-Jones*: Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford* (2) 13 (1962), 7–28.

Author's address: 115 67 Praha 1, Žitná 25, Czechoslovakia (Matematický ústav ČSAV).