Petr Veselý
Two contributions to the theory of coefficients of ergodicity

Czechoslovak Mathematical Journal, Vol. 42 (1992), No. 1, 73–88

Persistent URL: http://dml.cz/dmlcz/128308

Czechoslovak Mathematical Journal, 42 (117) 1992, Praha

TWO CONTRIBUTIONS TO THE THEORY OF COEFFICIENTS OF ERGODICITY

PETR VESELÝ, Praha

(Received November 14, 1990)

Coefficients of ergodicity play an important role in the theory of both homogeneous and inhomogeneous Markov chains, see [2, 8, 10, 11, 15]. They have also proved to be useful upper bounds of the maximum modulus of subdominant eigenvalues of stochastic and nonnegative irreducible matrices, see [4, 7–9, 12, 15–17].

This paper discusses two questions: the accuracy of numerical estimations of the ergodicity coefficients, and conditions for the existence of a coefficient of ergodicity which is equal to the maximum modulus of subdominant eingevalues of a given matrix.

1. PRELIMINARIES

We will use the following notation:

N – the set of all natural numbers;
R – the set of all real numbers;
C – the set of all complex numbers;
$\sigma(A)$ – the spectrum of a (square) matrix A;
$\varrho(A)$ – the spectral radius of a (square) matrix A;
$\operatorname{Lin} M$ – the linear span of a set $M \subset R^n$;
$\operatorname{conv} M$ – the convex hull of a set $M \subset R^n$;
$\| \, \|_p$ – the ℓ_p-norm ($p \in [1, \infty) \cup \{\infty\}$).

Let P be an $n \times n$ nonnegative, irreducible matrix and let $w \in R^n$ be a positive right eigenvector of P corresponding to the eigenvalue $\varrho(P)$. A coefficient of ergodicity $\tau_{\| \, \|}$ with respect to a vector norm $\| \, \|$ on R^n is defined as

$$\tau_{\| \, \|}(P) = \sup_{\substack{v \in R^n \\ v^T w = 0 \\ v \neq 0}} \frac{\|v^T P\|}{\|v\|}$$

(we put $||y^T|| = ||y||$ for all $y \in R^n$). Evidently,

$$\tau_{\|\ \|}(P) = \max_{\substack{v \in R^n \\ v^T w = 0 \\ \|v\| \leqslant 1}} \|v^T P\|$$

is an equivalent definition. For any ℓ_p-norm $\|\ \|_p$ we put $\tau_p(P) = \tau_{\|\ \|_p}(P)$.

The Perron-Frobenius theorem for square, nonnegative, irreducible matrices is fundamental for the theory of coefficients of ergodicity:

Suppose P is an $n \times n$ nonnegative irreducible matrix. Then

(a) $\varrho(P) > 0$ and $\varrho(P) \in \sigma(P)$.

(b) $\varrho(P)$ is a simple root of the characteristic equation of P.

(c) There exist positive left and right eigenvectors of P corresponding to the eigenvalue $\varrho(P)$.

(d) The eigenvectors corresponding to $\varrho(P)$ are unique up to a scalar multiple.

The proof can be found in [10].

The maximal modulus of a subdominant eigenvalue of an $n \times n$ nonnegative, irreducible matrix P is denoted by $\xi(P)$ and is defined as

$$\xi(P) = \max\{|\lambda|;\ \lambda \in \sigma(P),\ \lambda \neq \varrho(P)\} \text{ if } n > 1$$

and

$$\xi(P) = 0 \text{ if } n = 1.$$

The following propositions show well-known properties of norms on R^n.

Lemma 1.1. *Let V be a vector subspace of R^n and let ν be a norm on V. Then there exists a norm $\|\ \|$ on R^n such that $\|v\| = \nu(v)$ for each $v \in V$.*

Proof. Let $\{v^{(1)}, \ldots, v^{(k)}\}$ be a base of V and let

$$\left\{v^{(1)}, \ldots,\ v^{(k)},\ v^{(k+1)},\ \ldots,\ v^{(n)}\right\}$$

be a base of R^n. For each $x \in R^n$, $x = \sum_{i=1}^{n} \alpha_i v^{(i)}$ let

$$\|x\| = \nu\Big(\sum_{i=1}^{k} \alpha_i v^{(i)}\Big) + \sum_{i=k+1}^{n} |\alpha_i|.$$

It is easily seen that $\|\ \|$ is a norm on R^n and $\|v\| = \nu(v)$ for any $v \in V$. $\square$

Lemma 1.2. *Let V be a vector subspace of R^n and let $B \subset V$. Then B is a unit ball with respect to a certain norm on V if and only if B is a compact convex set, $\operatorname{Lin} B = V$ and $B = -B$.*

The proof can be found in [1].

Lemma 1.3. *Let V be a vector subspace of R^n and let μ, ν be norms on V. Then there exist positive numbers c_1, c_2 such that $c_2\nu(v) \leqslant \mu(v) \leqslant c_1\nu(v)$ holds for each $v \in V$.*

The proof can be found in [1].

2. Error estimation

For majority of norms explicit formulae are not known for the evaluation of the coefficient of ergodicity. Therefore coefficients of ergodicity are estimated numerically by a computer. Since

$$\tau_{\|\ \|}(P) = \max_{\substack{v \in R^n \\ v^T w = 0 \\ \|v\| \leqslant 1}} \|v^T P\|,$$

the problem is to find the maximum of a continuous function on a compact set. Usually it is performed in the following way: a finite set of points $v^{(1)}, \ldots, v^{(k)} \in R^n - \{0\}$, ${v^{(1)}}^T w = \ldots = {v^{(k)}}^T w = 0$ is successively found by any suitable algorithm and the value

$$\hat{\tau}_{\|\ \|}(P) = \max_{1 \leqslant i \leqslant k} \frac{\left\|{v^{(i)}}^T P\right\|}{\|v^{(i)}\|}$$

is taken as an estimate for $\tau_{\|\ \|}(P)$. Therefore it is useful to determine the upper bound of the error, i.e. we have to estimate the difference

$$\tau_{\|\ \|}(P) - \frac{\|v^T P\|}{\|v\|},$$

where $v \in R^n - \{0\}$, $v^T w = 0$.

This problem is solved in Theorem 2.2 and in Theorem 2.3. Tools for the solution of this problem are given in Theorem 2.1.

Theorem 2.1. *Let V be a vector subspace of R^n and let M be a family of norms on R^n. Let there exist positive numbers r_1 and r_2 such that for each norm $\nu \in M$ and for each vector $v \in V$,*

$$r_2 \|v\|_2 \leqslant \nu(v) \leqslant r_1 \|v\|_2$$

is true. Then

$$\kappa\left(\frac{v}{\mu(v)} - \frac{w}{\mu(w)}\right) \leqslant 3\left(\frac{r_1}{r_2}\right)^2 \left\|\frac{v}{\|v\|_2} - \frac{w}{\|w\|_2}\right\|_2$$

and

$$\kappa\left(\frac{v}{\mu(v)}-\frac{w}{\mu(w)}\right)\leqslant 3\left(\frac{r_1}{r_2}\right)^3\lambda\left(\frac{v}{\nu(v)}-\frac{w}{\nu(w)}\right)$$

holds for any norms κ, λ, μ, $\nu\in M$ and for any v, $w\in V-\{0\}$.

Proof. Let $\dim V\geqslant 2$ (the case $\dim V=1$ is easy). Let κ, λ, μ, ν be arbitrary norms from M. The inequality

$$\kappa\left(\frac{v}{\mu(v)}-\frac{w}{\mu(w)}\right)\leqslant r_1\left\|\frac{v}{\mu(v)}-\frac{w}{\mu(w)}\right\|_2 \tag{1}$$

holds for any v, $w\in V-\{0\}$.

Let $S(v)=\{s\in V-\{0\}|s^Tv=0\}$ for every $v\in V-\{0\}$. The assumption $\dim V\geqslant 2$ implies that $S(v)\neq\emptyset$ for every $v\in V$. If vectors v, $w\in V-\{0\}$ are linearly independent, then let $P(v,w)$ be the twodimensional subspace defined by v, w and 0, i.e. $P(v,w)=\{av+bw;\ a,b\in R\}$. Let B_μ be the unit ball with respect to the norm μ. Let $v\in V$, $\mu(v)=1$, $s\in S(v)$. Because B_μ is a convex set, the set $B_\mu\cap P(v,s)$ is convex, too. Furthermore, as v is a boundary point of $B_\mu\cap P(v,s)$, there exists at least one straight line in the plane $P(v,s)$, which intersects the point v and does not intersect the interior of a set $B_\mu\cap P(v,s)$. Let $t_\mu(v,s)$ be an arbitrary line with these properties. Let $\beta_\mu(v,s)$ be the angle between the straight line $t_\mu(v,s)$ and the straight line $\{av;\ a\in R\}$ (we have $\beta_\mu(v,s)\in[0,\frac{\pi}{2}]$). Let us put

$$\beta=\inf\{\beta_\mu(v,s)|v\in V,\mu(v)=1,s\in S(v)\}.$$

Let us suppose that $\sin\beta<r_2/r_1$. Then there exist a point $v^*\in V$, $\mu(v^*)=1$, and a vector $s^*\in S(v^*)$ such that $\sin\beta_\mu(v,s)<r_2/r_1$. If d is the Euclidean distance of the point 0 and the straight line $t_\mu(v^*,s^*)$, then

$$d=\|v^*\|_2\sin\beta_\mu(v,s)<\frac{\mu(v^*)}{r_2}\frac{r_2}{r_1}=\frac{1}{r_1}.$$

It follows that there exists a point $y^*\in t_\mu(v^*,s^*)$ such that $\|y^*\|_2<1/r_1$. The relation $y^*\in t_\mu(v^*,s^*)$ implies $\mu(y^*)\geqslant 1$. Altogether we have $r_1\|y^*\|_2<\mu(y^*)$, which is a contradiction to the supposition of the theorem. Thus $\sin\beta\geqslant r_2/r_1$.

Let $u(v,w)$ be the angle between the half-line $\{av;\ a\in[0,\infty)\}$ and the half-line $\{aw;\ a\in[0,\infty)\}$ for every $v,w\in V-\{0\}$ (we have $u(v,w)\in[0,\pi]$). Let δ be any number from the open interval $(0,1)$.

Let $v,w\in V$ and let $\mu(v)=\mu(w)=1$, $u(v,w)\in[0,\delta\beta)$. Let $\alpha=u(v,w)$. Let us consider two straight lines p_1, p_2 passing through the point v, which lie in the plane $P(v,w)$ and contain the angles $\beta,\beta-\alpha$ with the half-line $\{av;\ a\in[0,\infty)\}$. Let us put $g=p_1\cap\{aw;\ a\in[0,\infty)\}$, $h=p_2\cap\{aw;\ a\in[0,\infty)\}$, see Figure 1. From the

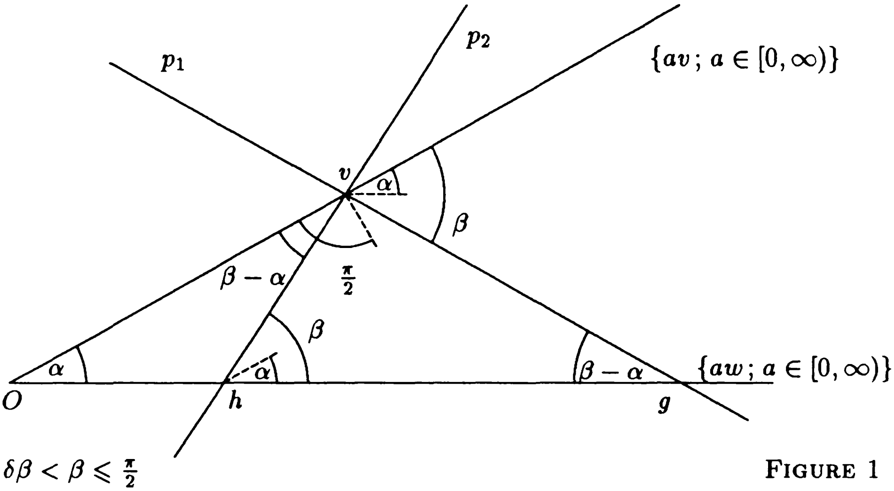

FIGURE 1

definition of the angle β we find that w is a member of the segment between the points g, h. Hence

$$\begin{aligned}
\|v-w\|_2 &\leqslant \max\{\|v-g\|_2, \|v-h\|_2\} \\
&= \max\left\{\frac{\|v\|_2}{\sin(\beta-\alpha)}\sin\alpha, \frac{\|v\|_2}{\sin(\pi-\beta)}\sin\alpha\right\} \\
&\leqslant \max\left\{\frac{\|v\|_2}{\sin(\beta-\delta\beta)}\sin\alpha, \frac{\|v\|_2}{\sin\beta}\sin\alpha\right\} \\
&= \frac{\|v\|_2}{\sin[(1-\delta)\beta]}\sin\alpha \\
&\leqslant \frac{\sin\alpha}{r_2\sin[(1-\delta)\beta]}.
\end{aligned}$$

If $v, w \in V$, $\mu(v) = \mu(w) = 1$, $u(v,w) \in [\delta\beta, \pi]$ is valid, then $\|v-w\|_2 \leqslant \|v\|_2 + \|w\|_2 \leqslant 2/r_2$ is true. We have proved that for all points $v, w \in V - \{0\}$ we have

$$\left\|\frac{v}{\mu(v)} - \frac{w}{\mu(w)}\right\|_2 \leqslant \frac{\sin[u(v,w)]}{r_2\sin[(1-\delta)\beta]} \tag{2a}$$

provided $u(v,w) \in [0, \delta\beta)$;

$$\left\|\frac{v}{\mu(v)} - \frac{w}{\mu(w)}\right\|_2 \leqslant \frac{2}{r_2} \tag{2b}$$

provided $u(v,w) \in [\delta\beta, \pi]$.

Let $v, w \in V$, $\nu(v) = \nu(w) = 1$ and let us put $\alpha = u(v,w)$. Let $\alpha \in [0, \delta\beta)$. Because $\|v\|_2 \sin\alpha$ is the Euclidean distance of the point v from the straight line $\{aw;\, w \in R\}$,

$$\|v-w\|_2 \geqslant \|v\|_2\sin\alpha \geqslant \frac{\nu(v)}{r_1}\sin\alpha = \frac{\sin\alpha}{r_1}$$

is true. For $\alpha \in [\delta\beta, \pi]$ we have

$$\begin{aligned} \|v-w\|_2 &\geqslant \inf\{\|x-y\|_2 | x,y \in V, \nu(x)=\nu(y)=1, u(x,y)=\alpha\} \\ &\geqslant \inf\{\|x-y\|_2 | x,y \in V, r_1\|x\|_2 = r_1\|y\|_2 = 1, u(x,y)=\alpha\} \\ &= \inf\{\|x-y\|_2 | x,y \in V, \|x\|_2 = \|y\|_2 = 1/r_1, u(x,y)=\alpha\} \\ &= \frac{2}{r_1}\sin\frac{\alpha}{2} \geqslant \frac{2}{r_1}\sin(\frac{1}{2}\delta\beta). \end{aligned}$$

We have proved that for all points $v, w \in V - \{0\}$

(3a) $$\left\| \frac{v}{\nu(v)} - \frac{w}{\nu(w)} \right\|_2 \geqslant \frac{\sin[u(v,w)]}{r_1}$$

provided $u(v,w) \in [0, \delta\beta)$;

(3b) $$\left\| \frac{v}{\nu(v)} - \frac{w}{\nu(w)} \right\|_2 \geqslant \frac{2}{r_1}\sin(\frac{1}{2}\delta\beta)$$

provided $u(v,w) \in [\delta\beta, \pi]$.

For all points $v, w \in V - \{0\}$ we have

(4) $$\left\| \frac{v}{\|v\|_2} - \frac{w}{\|w\|_2} \right\|_2 = 2\sin[\frac{1}{2}u(v,w)],$$

(5) $$\left\| \frac{v}{\nu(v)} - \frac{w}{\nu(w)} \right\|_2 \leqslant \frac{1}{r_2}\lambda\left(\frac{v}{\nu(v)} - \frac{w}{\nu(w)}\right).$$

From (1), (2a), (2b) and (4) we obtain that if $u(v,w) \in (0, \delta\beta)$ then

$$\begin{aligned} \kappa\left(\frac{v}{\mu(v)} - \frac{w}{\mu(w)}\right) &\leqslant r_1 \frac{\sin[u(v,w)]}{r_2\sin[(1-\delta)\beta]} \frac{1}{2\sin[\frac{1}{2}u(v,w)]} \left\| \frac{v}{\|v\|_2} - \frac{w}{\|w\|_2} \right\|_2 \\ &= \frac{r_1\cos[\frac{1}{2}u(v,w)]}{r_2\sin[(1-\delta)\beta]} \left\| \frac{v}{\|v\|_2} - \frac{w}{\|w\|_2} \right\|_2 \\ &\leqslant \frac{r_1}{r_2}\frac{1}{\sin[(1-\delta)\beta]} \left\| \frac{v}{\|v\|_2} - \frac{w}{\|w\|_2} \right\|_2, \end{aligned}$$

while if $u(v,w) \in [\delta\beta, \pi]$ then

$$\begin{aligned} \kappa\left(\frac{v}{\mu(v)} - \frac{w}{\mu(w)}\right) &\leqslant r_1 \frac{2}{r_2} \frac{1}{2\sin[\frac{1}{2}u(v,w)]} \left\| \frac{v}{\|v\|_2} - \frac{w}{\|w\|_2} \right\|_2 \\ &\leqslant \frac{r_1}{r_2}\frac{1}{\sin(\frac{1}{2}\delta\beta)} \left\| \frac{v}{\|v\|_2} - \frac{w}{\|w\|_2} \right\|_2 \end{aligned}$$

(the case $u(v,w) = 0$ is easy).

From (1), (2a), (2b), (3a), (3b) and (5) we obtain that if $u(v,w) \in (0, \delta\beta)$ then

$$\begin{aligned}\kappa\left(\frac{v}{\mu(v)} - \frac{w}{\mu(w)}\right) &\leqslant r_1 \frac{\sin[u(v,w)]}{r_2 \sin[(1-\delta)\beta]} \frac{r_1}{r_2 \sin[u(v,w)]} \lambda\left(\frac{v}{\nu(v)} - \frac{w}{\nu(w)}\right) \\ &= \left(\frac{r_1}{r_2}\right)^2 \frac{1}{\sin[(1-\delta)\beta]} \lambda\left(\frac{v}{\nu(v)} - \frac{w}{\nu(w)}\right),\end{aligned}$$

while if $u(v,w) \in [\delta\beta, \pi]$ then

$$\begin{aligned}\kappa\left(\frac{v}{\mu(v)} - \frac{w}{\mu(w)}\right) &\leqslant r_1 \frac{2}{r_2} \frac{r_1}{2 r_2 \sin(\frac{1}{2}\delta\beta)} \lambda\left(\frac{v}{\nu(v)} - \frac{w}{\nu(w)}\right) \\ &= \left(\frac{r_1}{r_2}\right)^2 \frac{1}{\sin(\frac{1}{2}\delta\beta)} \lambda\left(\frac{v}{\nu(v)} - \frac{w}{\nu(w)}\right)\end{aligned}$$

(the case $u(v,w) = 0$ is easy).

For $\delta = \frac{2}{3}$ we have $\sin[(1-\delta)\beta] = \sin(\frac{1}{2}\delta\beta) = \sin(\frac{1}{3}\beta)$ and

$$\frac{1}{\sin(\frac{1}{3}\beta)} \leqslant \frac{3}{\sin\beta} \leqslant 3\frac{r_1}{r_2}.$$

This completes the proof. □

Corollary 2.1. *Let $\|\ \|_p, \|\ \|_q, \|\ \|_r, \|\ \|_s$ be any ℓ_p-norms on R^n. Then*

$$\left\|\frac{v}{\|v\|_r} - \frac{w}{\|w\|_r}\right\|_p \leqslant 3n^2 \left\|\frac{v}{\|v\|_2} - \frac{w}{\|w\|_2}\right\|_2$$

and

$$\left\|\frac{v}{\|v\|_r} - \frac{w}{\|w\|_r}\right\|_p \leqslant 3n^3 \left\|\frac{v}{\|v\|_s} - \frac{w}{\|w\|_s}\right\|_q$$

holds for any $v, w \in R^n - \{0\}$.

Proof. For any ℓ_p-norm $\|\ \|_p$ and for any $v \in R^n$ we have

$$n^{-1/2}\|v\|_2 \leqslant \|v\|_\infty \leqslant \|v\|_p \leqslant \|v\|_1 \leqslant n^{1/2}\|v\|_2.$$

Hence it suffices to put $r_1 = n^{1/2}, r_2 = n^{-1/2}$. □

Theorem 2.2. *Let V be a vector subspace of R^n, let ν be a norm on R^n and let A be an $n \times n$ matrix such that $AV \subset V$. Let r_1, r_2 be positive numbers such that for each vector $v \in V$*

$$r_2\|v\|_2 \leqslant \nu(v) \leqslant r_1\|v\|_2.$$

Let $\operatorname{opt}(A,\nu)$ *be the set of all vectors* $y \in V - \{0\}$ *such that*

$$\sup_{\substack{v\in V\\ v\neq 0}} \frac{\nu(Av)}{\nu(v)} = \frac{\nu(Ay)}{\nu(y)}$$

(it is easy to see that $\operatorname{opt}(A,\nu) \neq \emptyset$*). Further, let* $x \in V - \{0\}$ *and let*

$$\varepsilon = \inf\left\{\left\|\frac{x}{\|x\|_2} - \frac{y}{\|y\|_2}\right\|_2 ; \, y \in \operatorname{opt}(A,\nu)\right\}.$$

Then

$$\begin{aligned}\sup_{\substack{v\in V\\ v\neq 0}} \frac{\nu(Av)}{\nu(v)} - \frac{\nu(Ax)}{\nu(x)} &\leqslant 3\left(\frac{r_1}{r_2}\right)^2 \sup_{\substack{v\in V\\ v\neq 0}} \frac{\nu(Av)}{\nu(v)}\varepsilon \\ &\leqslant 3\left(\frac{r_1}{r_2}\right)^3 [\varrho(A^TA)]^{1/2}\varepsilon.\end{aligned}$$

P r o o f. Let us put

$$T = \sup_{\substack{v\in V\\ v\neq 0}} \frac{\nu(Av)}{\nu(v)}.$$

For each $\delta > 0$ there exists a vector $y \in \operatorname{opt}(A,\nu)$ such that

$$\left\|\frac{x}{\|x\|_2} - \frac{y}{\|y\|_2}\right\|_2 \leqslant \varepsilon + \delta.$$

We have

$$\begin{aligned}T - \frac{\nu(Ax)}{\nu(x)} &= \frac{\nu(Ay)}{\nu(y)} - \frac{\nu(Ax)}{\nu(x)} \\ &= \nu\left[A\left(\frac{y}{\nu(y)} - \frac{x}{\nu(x)}\right) + \frac{Ax}{\nu(x)}\right] - \frac{\nu(Ax)}{\nu(x)} \\ &\leqslant \nu\left[A\left(\frac{y}{\nu(y)} - \frac{x}{\nu(x)}\right)\right] \\ &\leqslant \nu\left(\frac{y}{\nu(y)} - \frac{x}{\nu(x)}\right)T.\end{aligned}$$

According to Theorem 2.1 we obtain that

$$\begin{aligned}\nu\left(\frac{y}{\nu(y)} - \frac{x}{\nu(x)}\right)T &\leqslant 3\left(\frac{r_1}{r_2}\right)^2 \left\|\frac{y}{\|y\|_2} - \frac{x}{\|x\|_2}\right\|_2 T \\ &\leqslant 3\left(\frac{r_1}{r_2}\right)^2 T(\varepsilon + \delta).\end{aligned}$$

Hence

$$T - \frac{\nu(Ax)}{\nu(x)} \leqslant 3\left(\frac{r_1}{r_2}\right)^2 T\varepsilon.$$

Further,

$$T \leqslant \sup_{\substack{v\in V\\ v\neq 0}} \frac{r_1\|Av\|_2}{r_2\|v\|_2} \leqslant \frac{r_1}{r_2}\sup_{\substack{v\in R^n\\ v\neq 0}} \frac{\|Av\|_2}{\|v\|_2}$$

$$= \frac{r_1}{r_2}[\varrho(A^TA)]^{1/2}.$$

This completes the proof. □

Theorem 2.3. *Let P be an $n\times n$ stochastic matrix, let $V = \{v \in R^n \mid \sum_{i=1}^{n} v_i = 0\}$, let $\varepsilon > 0$ and let X_ε be a finite subset of $V - \{0\}$ such that for each $v \in V - \{0\}$*

$$\min_{x\in X_\varepsilon} \left\| \frac{x}{\|x\|_2} - \frac{v}{\|v\|_2} \right\|_2 \leqslant \varepsilon.$$

Then

$$\tau_p(P) - \max_{x\in X_\varepsilon} \frac{\|x^TP\|_p}{\|x\|_p} \leqslant \frac{3}{2}(2n)^{1/p}\varepsilon$$

for each $p \in [1,2)$,

$$\tau_p(P) - \max_{x\in X_\varepsilon} \frac{\|x^TP\|_p}{\|x\|_p} \leqslant \frac{3}{2}2^{1/p}n^{2-3/p}\varepsilon$$

for each $p \in [2,\infty)$, and

$$\tau_\infty(P) - \max_{x\in X_\varepsilon} \frac{\|x^TP\|_\infty}{\|x\|_\infty} \leqslant \frac{3}{2}n^2\varepsilon.$$

Proof. Let us put

$$r_1 = \sup_{\substack{v\in R^n\\ v\neq 0}} \frac{\|v\|_p}{\|v\|_2},$$

$$r_2 = \inf_{\substack{v\in R^n\\ v\neq 0}} \frac{\|v\|_p}{\|v\|_2}.$$

It is easily seen that

$$r_2\|v\|_2 \leqslant \|v\|_p \leqslant r_1\|v\|_2$$

is true for each $v \in V$. The reader can verify the following assertions: if $p \in [1,2)$ then

$$r_1 = \frac{\|(1,1,\ldots,1)\|_p}{\|(1,1,\ldots,1)\|_2} = n^{1/p-1/2},$$

$$r_2 = \frac{\|(1,0,\ldots,0)\|_p}{\|(1,0,\ldots,0)\|_2} = 1;$$

if $p \in [2, \infty)$ then

$$r_1 = \frac{\|(1,0,\ldots,0)\|_p}{\|(1,0,\ldots,0)\|_2} = 1,$$
$$r_2 = \frac{\|(1,1,\ldots,1)\|_p}{\|(1,1,\ldots,1)\|_2} = n^{1/p-1/2};$$

if $p = \infty$ then

$$r_1 = \frac{\|(1,0,\ldots,0)\|_\infty}{\|(1,0,\ldots,0)\|_2} = 1,$$
$$r_2 = \frac{\|(1,1,\ldots,1)\|_\infty}{\|(1,1,\ldots,1)\|_2} = n^{-1/2}.$$

By Theorem 2.2 (we put $A = P^T$) we obtain that

$$\tau_p(P) - \max_{x \in X_\varepsilon} \frac{\|x^T P\|_p}{\|x\|_p} \leqslant 3 \left(\frac{r_1}{r_2}\right)^2 \tau_p(P)\varepsilon,$$

where

$$\frac{r_1}{r_2} = n^{1/p-1/2}$$

for each $p \in [1, 2)$,

$$\frac{r_1}{r_2} = n^{1/2-1/p}$$

for each $p \in [2, \infty)$, and

$$\frac{r_1}{r_2} = n^{1/2}$$

for $p = \infty$.

Let S_n be the set of all $n \times n$ stochastic matrices. It is proved in [6] that if $p \in [1, \infty)$ and n is even then

$$\max_{P \in S_n} \tau_p(P) = \left(\frac{n}{2}\right)^{1-1/p},$$

if $p \in [1, \infty)$ and n is odd then

$$\max_{P \in S_n} \tau_p(P) = \left(\frac{1}{2}\right)^{1-1/p} \left(\frac{2}{(n+1)^{1-p} + (n-1)^{1-p}}\right)^{1/p},$$

if $p = \infty$ and n is even then

$$\max_{P \in S_n} \tau_\infty(P) = \frac{n}{2},$$

if $p = \infty$ and n is odd then

$$\max_{P \in S_n} \tau_\infty(P) = \frac{n-1}{2}.$$

The function $f(x) = x^{1-p}$ is convex on $(0,\infty)$ for each $p \in [1,\infty)$, hence $(n+1)^{1-p} + (n-1)^{1-p} \geqslant 2n^{1-p}$ for each $n \in N$, $n \geqslant 2$, $p \in [1,\infty)$. We obtain

$$\max_{P \in S_n} \tau_p(P) \leqslant \left(\frac{n}{2}\right)^{1-1/p}$$

for any $p \in [1,\infty)$ and any $n \in N$, and

$$\max_{P \in S_n} \tau_\infty(P) \leqslant \frac{n}{2}$$

for any $n \in N$. This completes the proof. $\square$

3. Conditions for the validity of $\xi(P) = \eta_{\|\ \|}(P)$

In [7] the following important theorems are proved (the symbol $\xi(P)$ is defined in Preliminaries):

[7, Theorem 3.1.] Let P be an $n \times n$ nonnegative irreducible matrix, and let $\|\ \|$ be a norm on R^n. Then $\xi(P) \leqslant \eta_{\|\ \|}(P)$.

[7, Theorem C.1.] Let P be an $n \times n$ nonnegative irreducible matrix, let w and v be positive right and left eigenvectors of P, respectively, corresponding to the eigenvalue $\varrho \equiv \varrho(P)$, and let $\varepsilon > 0$. Then there exists a norm $\|\ \|$ on R^n such that for $a = \varrho v / v^T w$,

$$\eta_{\|\ \|}(P) \leqslant \|P - wa^T\| \leqslant \xi(P) + \varepsilon,$$

where the matrix norm of a matrix $B \in R^{n \times n}$ is defined by

$$\|B\| = \max\{\|x^T B\| : \|x\| \leqslant 1,\ x \in R^n\}.$$

An interesting question is how to find whether there exists a norm $\|\ \|$ on R^n such that $\eta_{\|\ \|}(P) = \xi(P)$. This problem is solved in Theorem 3.1. Theorem 3.2 shows that there always exists a seminorm $\|\ \|$ on R^n such that $\eta_{\|\ \|}(P) = \xi(P)$.

Lemma 3.1. *Let V be a vector subspace of R^n, let $\|\ \|$ be a norm on R^n and let A be an $n \times n$ matrix such that $AV \subset V$. Then the following assertions are equivalent:*

(a) *There exists a norm ν on R^n such that*

$$\sup_{\substack{v \in V \\ v \neq 0}} \frac{\nu(Av)}{\nu(v)} \leqslant 1;$$

(b)

$$\sup_{k\in N}\sup_{\substack{v\in V\\ v\neq 0}}\frac{\|A^k v\|}{\|v\|}<\infty;$$

(c) $\sup\limits_{k\in N}\|A^k v\|<\infty$ *for each* $v\in V$.

Proof. (a) $\Rightarrow$ (b) The assertion (a) implies that the inequality $\nu(Av)\leqslant\nu(v)$ is true for each $v\in V$. It follows that $\nu(A^k v)\leqslant\nu(v)$ for each $v\in V$, $k\in N$. According to Lemma 1.3 there exist numbers $c_1,c_2>0$ such that $c_2\nu(v)\leqslant\|v\|\leqslant c_1\nu(v)$ for each $v\in V$. We conclude that for all $k\in N$

$$\sup_{\substack{v\in V\\ v\neq 0}}\frac{\|A^k v\|}{\|v\|}\leqslant\sup_{\substack{v\in V\\ v\neq 0}}\frac{c_1\nu(A^k v)}{c_2\nu(v)}\leqslant\frac{c_1}{c_2}.$$

(b) $\Rightarrow$ (c) For all $k\in N$, $v\in V$ the inequality

$$\|A^k v\|\leqslant\|v\|\sup_{\substack{x\in V\\ x\neq 0}}\frac{\|A^k x\|}{\|x\|}$$

is true. Hence

$$\sup_{k\in N}\|A^k v\|\leqslant\|v\|\sup_{k\in N}\sup_{\substack{x\in V\\ x\neq 0}}\frac{\|A^k x\|}{\|x\|}<\infty.$$

(c) $\Rightarrow$ (a) Let $\{v^{(1)},\ldots,v^{(m)}\}$ be a base of V. Let us put $G=\operatorname{conv}\{v^{(1)},\ldots,v^{(m)},-v^{(1)},\ldots,-v^{(m)}\}$. Then $\operatorname{Lin}G=V$ and $G=-G$ is true. Let us put

$$\gamma=\max_{1\leqslant i\leqslant m}\sup_{k\in N}\left\|A^k v^{(i)}\right\|.$$

The assertion (c) implies that $\gamma<\infty$. For all $\alpha_1,\ldots,\alpha_{2m}\geqslant 0$, $\sum\limits_{i=1}^{2m}\alpha_i=1$, we have

$$\sup_{k\in N}\left\|A^k\left(\sum_{i=1}^{m}(\alpha_i-\alpha_{m+i})v^{(i)}\right)\right\|\leqslant\sup_{k\in N}\sum_{i=1}^{m}|\alpha_i-\alpha_{m+i}|\left\|A^k v^{(i)}\right\|\leqslant\gamma.$$

Thus the set $G^*=\bigcup\limits_{k=1}^{\infty}A^kG$ is bounded. Let $H=\operatorname{conv}G^*$ and let H^* be the closure of the set H. The set H^* is a compact convex set, $\operatorname{Lin}H^*=V$ (since $G\subset G^*\subset H\subset H^*\subset V$ and $\operatorname{Lin}G=V$), and $H^*=-H^*$ (since $A^kG=-A^kG$ for each $k\in N$, hence $G^*=-G^*$, thus $H=-H$). According to Lemmas 1.1 and 1.2 we obtain that there exists a norm ν on R^n such that $\{v\in V|\nu(v)\leqslant 1\}=H^*$.

If $x\in G^*$, then there exists $k\in N$ fulfilling $x\in A^kG$. Hence $Ax\in A^{k+1}G$, thus $AG^*\subset G^*$. Let $v\in H$. There exist $s\in N$, $\alpha_1,\ldots,\alpha_s\geqslant 0$ and $x^{(1)},\ldots,x^{(s)}\in G^*$

such that $\sum_{i=1}^{s} \alpha_i = 1$ and $\sum_{i=1}^{s} \alpha_i x^{(i)} = v$. $Ax^{(i)} \in G^*$ is true for all $i = 1, \ldots, s$, because $AG^* \subset G^*$. We have

$$Av = A\left(\sum_{i=1}^{s} \alpha_i x^{(i)}\right) = \sum_{i=1}^{s} \alpha_i A x^{(i)} \in H.$$

Thus $AH \subset H$, hence $AH^* \subset H^*$. We have proved that if $v \in H^*$ (i.e. $\nu(v) \leqslant 1$) then $Av \in H^*$(i.e. $\nu(Av) \leqslant 1$). Hence

$$\sup_{\substack{v \in V \\ v \neq 0}} \frac{\nu(Av)}{\nu(v)} = \sup_{\substack{v \in V \\ \nu(v) \leqslant 1}} \nu(Av) \leqslant \sup_{\substack{v \in V \\ \nu(v) \leqslant 1}} 1 = 1.$$

□

Theorem 3.1. *Let P be an $n \times n$ nonnegative irreducible matrix such that $\xi(P) > 0$. Let w be a right eigenvector corresponding to the eigenvalue $\varrho(P)$, and let $\| \ \|$ be a norm on R^n. Then the following assertions are equivalent:*

(a) *There exists a norm ν on R^n such that $\tau_\nu(P) = \xi(P)$.*

(b) $\sup_{k \in N} [\xi(P)]^{-k} \tau_{\| \ \|}(P^k) < \infty$.

(c) $\sup_{k \in N} [\xi(P)]^{-k} \|v^T P^k\| < \infty$ *for each $v \in R^n$, $v^T w = 0$.*

(d) *If λ is an eigenvalue of P such that $|\lambda| = \xi(P)$, then its algebraic multiplicity is equal to its geometric multiplicity.*

Proof. The equivalence of the assertions (a), (b), (c), follows from [7, Theorem 3.1] and from Lemma 3.1 (we put $A = P^T/\xi(P), V = \{v \in R^n | v^T w = 0\}$). We show that (c) $\Leftrightarrow$ (d). Let $\lambda_0, \lambda_1, \ldots, \lambda_m$ be all eigenvalues of P, let $a_0, a_1, \ldots, a_m$ be their algebraic multiplicities, and let $g_0, g_1, \ldots, g_m$ be their geometric multiplicities. Let λ_0 be the Perron-Frobenius eigenvalue of P, i.e. $\lambda_0 = \varrho(P)$, $a_0 = 1$ and $Pw = \lambda_0 w$. Let $\Lambda = \mathrm{diag}\{\Lambda_0, \Lambda_1, \ldots, \Lambda_M\}$ be the Jordan matrix of P, where $\Lambda_0 = (\lambda_0)$ and Λ_1, $\ldots$, Λ_M are the Jordan cells of the eigenvalues $\lambda_1, \ldots, \lambda_m$. Let T be a regular matrix such that $P = T\Lambda T^{-1}$. Finally, let us put $V = \{v \in R^n | v^T w = 0\}$.

(c) $\Rightarrow$ (d) Assume that (c) is true and (d) is false. Then there exists $q \in \{1, \ldots, m\}$ such that $|\lambda_q| = \xi(P)$ and $g_q \neq a_q$. It follows that $g_q < a_q$(in accordance with the well-know inequality between the algebraic and the geometric multiplicity of eigenvalues). The number of Jordan cells of any eigenvalue is equal to its geometric multiplicity, while the sum of their matricial ranks is equal to its algebraic multiplicity (see for example [1]). Hence it follows that there exists a Jordan cell of λ_q such that its rank is at least 2. Denote this cell by Λ_Q $(1 \leqslant Q \leqslant M)$. Let $u^{(k)}$ be the second column of the matrix Λ_Q^k, where $k \in N$. It is easily seen that

$$u^{(k)} = (k\lambda_q^{k-1}, \lambda_q^k, 0, \ldots, 0)^T$$

for each $k \in N$. $u_2^{(k)} (\equiv \lambda_q^k)$ is a diagonal element of the matrix Λ^k. Let $u_2^{(k)}$ lie in the rth row and in the rth column of the matrix Λ^k. The inequality $r \geqslant 3$ is true since the rank of Λ_0 is 1 and the rank of Λ_Q is at least 2. Let $\Lambda_{(r)}^k$ be the rth column of the matrix Λ^k and let $T_{(r-1)}$ be the $(r-1)$ st column of T. The first column of T is equal to w (because $PT = T\Lambda$ and $\Lambda_0 = (\lambda_0)$) and $r - 1 \geqslant 2$, hence $T_{(r-1)} \neq cw$ for each $c \in R$. It follows that there exists a vector $z \in V$ such that $t_{r-1} \neq 0$, where $t = (t_1, \ldots, t_n) = z^T T$. We have

$$\begin{aligned}
[\xi(P)]^{-k} |t^T \Lambda_{(r)}^k| &= [\xi(P)]^{-k} |t_{r-1} k \lambda_q^{k-1} + t_r \lambda_q^k| \\
&\geqslant [\xi(P)]^{-k} (|t_{r-1} k \lambda_q^{k-1}| - |t_r \lambda_q^k|) \\
&= [\xi(P)]^{-1} |k t_{r-1}| - |t_r|,
\end{aligned}$$

thus

$$\lim_{k \to \infty} [\xi(P)]^{-k} |z^T T \Lambda_{(r)}^k| = \infty,$$

where $\Lambda_{(r)}^k$ is the rth column of Λ^k. Hence

$$\lim_{k \to \infty} [\xi(P)]^{-k} \| z^T T \Lambda^k \| = \infty,$$

whence

$$\sup_{k \in N} [\xi(P)]^{-k} \| z^T P^k \| = \sup_{k \in N} [\xi(P)]^{-k} \| z^T T \Lambda^k T^{-1} \| = \infty.$$

This contradicts (c).

(d) $\Rightarrow$ (c) Let (d) be true.

$$[\xi(P)]^{-k} P^k = T \operatorname{diag}\{[\xi(P)]^{-k} \Lambda_0^k, \ldots, [\xi(P)]^{-k} \Lambda_M^k\} T^{-1}$$

holds for each $k \in N$. The first column of T is equal to w, because $PT = T\Lambda$ and $\Lambda_0 = (\lambda_0)$. It follows that for each $k \in N$ and $v \in V$ the vector $[\xi(P)]^{-k} (v^T P^k) = [\xi(P)]^{-k} \left((v^T T) \Lambda^k T^{-1} \right)$ is independent of the Jordan cell $\Lambda_0^k = (\lambda_0^k)$. Let $e \in \{1, \ldots, m\}$ and let Λ_E be any Jordan cell of the eigenvalue λ_e. There are two possibilities:

1. Let $|\lambda_e| = \xi(P)$. Since the number of Jordan cells of any eigenvalue is equal to its geometric multiplicity and the sum of their matricial ranks is equal to its algebraic multiplicity (see for example [1]), and because (d) is true, we have $\Lambda_E = (\lambda_e)$. It follows that $[\xi(P)]^{-k} \Lambda_E^k = ([\lambda_e / \xi(P)]^k)$ and $|[\lambda_e / \xi(P)]^k| = 1$.
2. Let $|\lambda_e| \neq \xi(P)$. Then $|\lambda_e / \xi(P)| < 1$. It follows (see for example [1]) that any element of the matrix $[\xi(P)]^{-k} \Lambda_E^k$ tends to zero as $k \to \infty$. This completes the proof. □

Remark. Let P be an $n \times n$ nonnegative irreducile matrix such that $\xi(P) = 0$. Let w be a right eigenvector corresponding to the eigenvalue $\varrho(P)$. It is obvious that there exists a norm ν on R^n such that $\tau_\nu(P) = \xi(P) = 0$ if and only if there exists a vector $a \in R^n$ such that $P = w a^T$.

Theorem 3.2. *Let P be an $n \times n$ nonnegative irreducible matrix, let w be a right eigenvector of P corresponding to the eigenvalue $\varrho(P)$. Then there exists a seminorm $\|\ \|$ on R^n such that the set $\{v \in R^n | v^T w = 0, \|v\| \neq 0\}$ is non-empty and*

$$\sup_{\substack{v \in R^n \\ v^T w = 0 \\ \|v\| \neq 0}} \frac{\|v^T P\|}{\|v\|} = \xi(P).$$

Proof. Let $\lambda \in \varrho(P) - \{\varrho(P)\}$ be an eigenvalue of P such that $|\lambda| = \xi(P)$, and let $z \in C^n$ be a right eigenvector of P corresponding to the eigenvalue λ. For all $x \in R^n$ let us put $\|x\| = |x^T z|$. It is easily seen that $\|\ \|$ is a seminorm on R^n. Assume that $v^T z = 0$ for each $v \in R^n$, $v^T w = 0$. It follows that there exists a number $c \in C$, $c \neq 0$ such that $z = cw$. Hence $Pz = \varrho(P)z$, thus $\lambda = \varrho(P)$. This contradicts $\lambda \in \sigma(P) - \{\varrho(P)\}$.

Finally, $\|x^T P\| = |(x^T P)z| = |x^T(Pz)| = |x^T(\lambda z)| = |\lambda|\,|x^T z| = |\lambda|\,\|x\|$ is true for each $x \in R^n$. This completes the proof. □

Acknowledgment. I am very grateful to A. Lešanovský for his constant interest in my work, valuable suggestions and help during the preparation of this paper.

References

[1] *R. A. Horn, Ch. R. Johnson*: Matrix Analysis, Cambridge University Press, Cambridge, London, New York, New Rochelle, Melbourne and Sydney, 1985, (Russian translation *Р. Хорн, Ч. Джонсон*: Матричный анализ, Москва, Мир, 1989).

[2] *P. Kratochvíl, A. Lešanovský*: A contractive property in finite state Markov chains, Czechoslovak Math. J. *35(110)* (1985), 491–509.

[3] *T. S. Leóng*: A note on upper bounds on the maximum modulus of subdominant eigenvalues of nonnegative matrices, Linear Algebra Appl. *106* (1988), 1–4.

[4] *A. Lešanovský*: Coefficients of ergodicity generated by non-symetrical vector norms, Czechoslovak Math. J. *40(115)* (1990), 284–294.

[5] *A. Rhodius*: On almost scrambling stochastic matrices, Linear Algebra Appl. *126* (1989), 76–86.

[6] *A. Rhodius*: The maximal value for coefficients of ergodicity, Stochastic Process. Appl. *29* (1988), 141–145.

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

[8] *E. Seneta*: Coefficients of ergodicity: structure and applications, Adv. Appl. Prob. *11* (1979), 576–590.

[9] *E. Seneta*: Explicit forms for ergodicity coefficients and spectrum localization, Linear Algebra Appl. *60* (1984), 187–197.

[10] *E. Seneta*: Non-Negative Matrices and Markov Chains, Springer-Verlag, New York, Heidelberg and Berlin, 1981.

[11] *E. Seneta*: Perturbation of the stationary distribution measured by ergodicity coefficients, Adv. Appl. Prob. *20* (1988), 228–230.

[12] *E. Seneta*: Spectrum localization by ergodicity coefficients for stochastic matrices, Linear and Multilinear Algebra *14* (1983), 343–347.

[13] *E. Seneta, C. P. Tan*: The Euclidean and Frobenius ergodicity coefficients and spectrum localization, Bull. Malaysia Math. Soc. *(7)1* (1984), 1–7.
[14] *C. P. Tan*: A functional form for a particular coefficient of ergodicity, J. Appl. Probab. *19* (1982), 858–863.
[15] *C. P. Tan*: Coefficients of ergodicity with respect to vector norms, J. Appl Probab. *20* (1983), 277–287.
[16] *C. P. Tan*: Spectrum localization of an ergodic stochastic matrix, Bull. Inst. Math. Acad. Sinica *12* (1984), 147–151.
[17] *C. P. Tan*: Spectrum localization using Hőlder norms, Houston J. Math. *12* (1986), 441–449.

Author's address: 586 01 Jihlava, Matky Boží 11, Czechoslovakia.