Czechoslovak Mathematical Journal

Petr Veselý

Majorants of matrix norms and spectrum localization

Czechoslovak Mathematical Journal, Vol. 44 (1994), No. 1, 141–161

Persistent URL: http://dml.cz/dmlcz/128445

MAJORANTS OF MATRIX NORMS AND SPECTRUM LOCALIZATION

PETR VESELÝ, Praha

(Received April 30, 1992)

Let $\mathbf{V}$ be a subspace of $\mathbb{R}^n$, let $\mathscr{A}$ be a set of $n\times n$ real matrices such that $Ax\in\mathbf{V}$ whenever $A\in\mathscr{A}$, $x\in\mathbf{V}$ and let $\tau\colon\mathscr{A}\to[0,\infty)$. The paper introduces conditions upon which there exists a vector norm $\|\ \|$ on $\mathbf{V}$ such that

$$(1)\qquad \sup_{\substack{x\in\mathbf{V}\\ x\neq 0}}\frac{\|Ax\|}{\|x\|}\leqslant\tau(A)\quad\text{for all } A\in\mathscr{A}.$$

Moreover, a constructive definition of norms satisfying (1) is presented.

The results are applied to spectrum localization of stochastic matrices in the second part of the paper. The maximum modulus of subdominant eigenvalues is an important characteristic of stochastic matrices. Determination of upper bounds for this value is one of the main objectives of the theory of coefficients of ergodicity, which investigates those of the form

$$(2)\qquad \sup_{\substack{x^T\mathbf{1}=0\\ x\neq 0}}\frac{\|x^TP\|}{\|x\|},\quad \sup_{\substack{\pi^Tx=0\\ x\neq 0}}\frac{\|Px\|}{\|x\|},$$

where P is a given stochastic matrix, $\|\ \|$ is a vector norm, π is the stationary distribution of P and $\mathbf{1}=(1,1,\ldots,1)^T$, see [2, 4, 6, 7, 9, 10]. Another upper bound is studied for example in [11].

In this paper, upper bounds of the maximum modulus of subdominant eigenvalues are approached in a more general setting. Namely, a broad subclass $\mathscr{C}$ of the class of all functions τ satisfying the inequality (1) for $\mathbf{V}=\{x\in\mathbb{R}^n\mid x^T\mathbf{1}=0\}$ and for at least one vector norm $\|\ \|$ is taken into consideration. The advantage of this approach is that an upper bound of the maximum modulus of subdominant eigenvalues can be chosen in such a way that both the verification of $\tau\in\mathscr{C}$ and the formula for $\tau(P)$

itself are simple in contrast to the calculation of (2) for most of the vector norms $\| \ \|$. Moreover, Theorem 2.4 shows that, in the case of an irreducible aperiodic stochastic matrix P, any $\tau \in \mathscr{C}$ can be taken to produce an arbitrary tight estimate for the second largest modulus of eigenvalues of the matrix P. Examples of suitable functions $\tau \in \mathscr{C}$ are presented as well.

Notation

$\mathbf{M}_n$ — the set of all $n \times n$ real matrices identified with the vector space $\mathbb{R}^{n^2}$;
$\langle , \rangle$ — the scalar product on $\mathbf{M}_n$ defined by

$$\langle A, B\rangle = \sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs}b_{rs}, \quad A, B \in \mathbf{M}_n;$$

$\langle\langle \ \rangle\rangle$ — the norm on $\mathbf{M}_n$ defined by

$$\langle\langle A\rangle\rangle = \langle A, A\rangle^{1/2}, \quad A \in \mathbf{M}_n;$$

$\mathscr{S}_n$ — the set of all $n \times n$ stochastic matrices;
Σ_n — the set of all $n \times n$ real matrices such that their column sums are equal to 1;
$\sigma(a)$ — the spectrum of a (square) matrix A;
$\varrho(A)$ — the spectral radius of a (square) matrix A;
$\mathrm{Lin}(M)$ — the linar span of a set $M \subseteq \mathbb{R}^n$;
I_n — the $n \times n$ identity matrix;
$\mathbf{0} = (0, 0, \ldots, 0)^T \in \mathbb{R}^n$;
$\mathbf{1} = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$.

1. Upper bounds for norms of matrix operators

Let $\mathbf{V}$ be a vector subspace of $\mathbb{R}^n$ and let $\dim \mathbf{V} \geqslant 1$. Troughout this section, $\mathscr{A}$ is a set of $n \times n$ real matrices such that

(I) if A, $B \in \mathscr{A}$ and $\alpha \in \mathbb{R}$ then $\alpha A + (1-\alpha)B \in \mathscr{A}$;
(II) if A, $B \in \mathscr{A}$ then $AB \in \mathscr{A}$;
(III) $\mathrm{Lin}(\mathscr{A}\mathbf{V}) = \mathbf{V}$;
(IV) there exists at least one matrix $A \in \mathscr{A}$ satisfying $A\mathbf{V} = \{\mathbf{0}\}$.

Let us give some examples. Let $\mathbf{V} = \mathbb{R}^n$. Then the set of all $n \times n$ real upper (lower) triangular matrices and the set of all $n \times n$ real matrices with a fixed eigenvector meet the demands of (I), …, (IV). The important set of matrices satisfying the conditions (I), …, (IV) is studied in Section 2.

This section is devoted to the study of funcions $\tau: \mathscr{A} \to [0,\infty)$ satisfying the conditions

(α) τ is a convex funcion on $\mathscr{A}$;

(β) $\tau(AB) \leqslant \tau(A)\tau(B)$ for all A, $B \in \mathscr{A}$;

(γ) there exists at least one matrix $Z \in \mathscr{A}$ such that $\tau(Z) = 0$;

(δ) if $Z \in \mathscr{A}$ and $\tau(Z) = 0$ then $Z\mathbf{V} = \{0\}$;

(ε) if A, $Z \in \mathscr{A}$ and $\tau(Z) = 0$ then $\tau(\alpha A + (1-\alpha)Z) = \alpha\tau(A)$ holds for any $\alpha > 0$.

Note that a set of functions $\tau: \mathscr{A} \to [0,\infty)$ satisfying these conditions is non-empty. Indeed, let $\|\ \|$ be a vector norm on $\mathbf{V}$ and let

$$\tau_{\|\ \|}(A) = \sup_{\substack{x\in\mathbf{V}\\ x\neq 0}} \frac{\|Ax\|}{\|x\|}$$

for every $A \in \mathscr{A}$. It is easily seen that the conditions (α), ..., (ε) hold for the function $\tau_{\|\ \|}$ (the condition (γ) holds by virtue of (IV)).

Definition. Let $\mathscr{V} = \{v^{(1)}, \dots, v^{(k)}\}$ be a finite subset of $\mathbf{V}$ satisfying $\mathrm{Lin}(\mathscr{A}\mathscr{V}) = \mathbf{V}$, let τ be a nonnegative function on $\mathscr{A}$ and let $x \in \mathbf{V}$. Define the quantity

$$[x]^{\tau}_{\mathscr{V}} = \inf\Big\{\sum_{i=1}^{2k} c_i\tau(Q_i) \mid (c,Q) \in \mathscr{K}_{\mathscr{V}}(x)\Big\},$$

where $\mathscr{K}_{\mathscr{V}}(x)$ is the set of all pairs $(c,Q) \in [0,\infty)^{2k} \times \mathscr{A}^{2k}$ such that

$$\sum_{i=1}^{k}(c_iQ_i - c_{k+i}Q_{k+i})v^{(i)} = x.$$

Remark. The condition (III) implies that there always exists a finite set $\mathscr{V} \subset \mathbf{V}$ such that $\mathrm{Lin}(\mathscr{A}\mathscr{V}) = \mathbf{V}$ (for example, any base of $\mathbf{V}$ has this property). The reader can verify that, for any such set $\mathscr{V}$ and for any $x \in \mathbf{V}$, the set $\mathscr{K}_{\mathscr{V}}(x)$ is non-empty. Thus the quantity $[x]^{\tau}_{\mathscr{V}}$ is well-defined.

Theorem 1.1. *Let τ be a nonnegative function on $\mathscr{A}$ satisfying the conditions (α) and (β), let $\mathscr{V} = \{v^{(1)}, \dots, v^{(k)}\}$ be a finite subset of $\mathbf{V}$ such that $\mathrm{Lin}(\mathscr{A}\mathscr{V}) = \mathbf{V}$. If $[x]^{\tau}_{\mathscr{V}} > 0$ for each $x \in \mathbf{V} - \{0\}$, then $[\ ]^{\tau}_{\mathscr{V}}$ is a norm on $\mathbf{V}$ and*

$$\sup_{\substack{x\in\mathbf{V}\\ x\neq 0}} \frac{[Ax]^{\tau}_{\mathscr{V}}}{[x]^{\tau}_{\mathscr{V}}} \leqslant \tau(A)$$

for all $A \in \mathscr{A}$.

P r o o f. It is obvious that $[\,]^{\tau}_{\mathscr{V}}$ is a nonnegative function on $\mathscr{A}$ and that $[x]^{\tau}_{\mathscr{V}} = 0$ if and only if $x = 0$ according to the assumptions.

Let $x \in \mathbf{V}$ and let $\lambda > 0$. The relation $(c, Q) \in \mathscr{K}_{\mathscr{V}}(x)$ holds if and only if $(\lambda c, Q) \in \mathscr{K}_{\mathscr{V}}(\lambda x)$. It implies $[\lambda x]^{\tau}_{\mathscr{V}} = \lambda [x]^{\tau}_{\mathscr{V}}$. The relation

$$\big((c_1, \dots, c_k, c_{k+1}, \dots, c_{2k}), (Q_1, \dots, Q_k, Q_{k+1}, \dots, Q_{2k})\big) \in \mathscr{K}_{\mathscr{V}}(x)$$

holds if and only if

$$\big((c_{k+1}, \dots, c_{2k}, c_1, \dots, c_k), (Q_{k+1}, \dots, Q_{2k}, Q_1, \dots, Q_k)\big) \in \mathscr{K}_{\mathscr{V}}(-x)$$

holds. Thus $[x]^{\tau}_{\mathscr{V}} = [-x]^{\tau}_{\mathscr{V}}$. We conclude that $[\lambda x]^{\tau}_{\mathscr{V}} = |\lambda| [x]^{\tau}_{\mathscr{V}}$ for each $\lambda \in \mathbb{R}$.

Let x, $y \in \mathbf{V}$. If $x = \mathbf{0}$ then the triangular inequality $[x]^{\tau}_{\mathscr{V}} + [y]^{\tau}_{\mathscr{V}} \geqslant [x + y]^{\tau}_{\mathscr{V}}$ is true. Let $x \neq \mathbf{0}$ and consider any $\varepsilon > 0$. We find from the definition of $[\,]^{\tau}_{\mathscr{V}}$ that there exist $(c, Q) \in \mathscr{K}_{\mathscr{V}}(x)$, $(d, R) \in \mathscr{K}_{\mathscr{V}}(y)$ satisfying

$$\sum_{i=1}^{2k} c_i \tau(Q_i) \leqslant [x]^{\tau}_{\mathscr{V}} + \frac{1}{2}\varepsilon, \quad \sum_{i=1}^{2k} d_i \tau(R_i) \leqslant [y]^{\tau}_{\mathscr{V}} + \frac{1}{2}\varepsilon.$$

Since $[x]^{\tau}_{\mathscr{V}} > 0$ the set $\mathscr{J} = \big\{ i \in \{1, \dots, 2k\} \mid c_i + d_i > 0 \big\}$ is non-empty. We have

$$\begin{aligned}
[x]^{\tau}_{\mathscr{V}} + [y]^{\tau}_{\mathscr{V}} &\geqslant \sum_{i=1}^{2k} c_i \tau(Q_i) + \sum_{i=1}^{2k} d_i \tau(R_i) - \varepsilon \\
&= \sum_{i \in \mathscr{J}} \big(c_i \tau(Q_i) + d_i \tau(R_i)\big) - \varepsilon \\
&= \sum_{i \in \mathscr{J}} (c_i + d_i) \Big(\frac{c_i}{c_i + d_i} \tau(Q_i) + \frac{d_i}{c_i + d_i} \tau(R_i) \Big) - \varepsilon \\
&\geqslant \sum_{i \in \mathscr{J}} (c_i + d_i)\, \tau \Big(\frac{c_i}{c_i + d_i} Q_i + \frac{d_i}{c_i + d_i} R_i \Big) - \varepsilon \\
&= \sum_{i=1}^{2k} e_i \tau(S_i) - \varepsilon,
\end{aligned}$$

where

$$e_i = \begin{cases} c_i + d_i & \text{if } i \in \mathscr{J}, \\ 0 & \text{if } i \in \{1, \dots, 2k\} - \mathscr{J}, \end{cases}$$

and

$$S_i = \begin{cases} \dfrac{c_i}{c_i + d_i} Q_i + \dfrac{d_i}{c_i + d_i} R_i & \text{if } i \in \mathscr{J}, \\ Q_1 & \text{if } i \in \{1, \dots, 2k\} - \mathscr{J}. \end{cases}$$

The reader can find that $(e, S) \in \mathscr{K}_{\mathscr{V}}(x+y)$ so that $[x]^\tau_{\mathscr{V}} + [y]^\tau_{\mathscr{V}} \geqslant [x+y]^\tau_{\mathscr{V}} - \varepsilon$ for any $\varepsilon > 0$. Thus, the triangular inequality is valid. We have proved that $[\ ]^\tau_{\mathscr{V}}$ is a norm on $\mathbf{V}$.

Let $A \in \mathscr{A}$ and consider any $x \in \mathbf{V}$. It is easy to see that if $(c, Q) \in \mathscr{K}_{\mathscr{V}}(x)$ then $(c, AQ) \in \mathscr{K}_{\mathscr{V}}(Ax)$, where

$$AQ = (AQ_1, \dots, AQ_{2k}).$$

Hence

$$[Ax]^\tau_{\mathscr{V}} \leqslant \inf\left\{ \sum_{i=1}^{2k} c_i \tau(AQ_i) \mid (c, Q) \in \mathscr{K}_{\mathscr{V}}(x) \right\} \leqslant \tau(A)[x]^\tau_{\mathscr{V}}$$

by (β). This completes the proof. □

Theorem 1.2. *Let τ be a nonnegative function on $\mathscr{A}$ satisfying the conditions (α) and (β), let $\| \ \|$ be a norm on $\mathbf{V}$. Then:*

(i) *If*

(3) $$\|x\| \leqslant (\max_{v \in \mathscr{V}} \|v\|)[x]^\tau_{\mathscr{V}}$$

for any $x \in \mathbf{V}$ and for any finite subset $\mathscr{V} \subset \mathbf{V}$ satisfying $\mathrm{Lin}(\mathscr{A}\mathscr{V}) = \mathbf{V}$, *then*

$$\sup_{\substack{x \in \mathbf{V} \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} \leqslant \tau(A)$$

for all $A \in \mathscr{A}$.

(ii) *If*

(4) $$\sup_{\substack{x \in \mathbf{V} \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} \leqslant \tau(A)$$

for all $A \in \mathscr{A}$, then

$$[\ ]^\tau_{\mathscr{V}} \text{ is a norm on } \mathbf{V},$$

$$\|x\| \leqslant (\max_{v \in \mathscr{V}} \|v\|)[x]^\tau_{\mathscr{V}}$$

and

$$\sup_{\substack{x \in \mathbf{V} \\ x \neq 0}} \frac{[Ax]^\tau_{\mathscr{V}}}{[x]^\tau_{\mathscr{V}}} \leqslant \tau(A)$$

for any finite subset $\mathscr{V} \subset \mathbf{V}$ satisfying $\mathrm{Lin}(\mathscr{A}\mathbf{V}) = \mathbf{V}$ *and for all $x \in \mathbf{V}$, $A \in \mathscr{A}$.*

P r o o f. (i) Let $A \in \mathscr{A}$ and let $x \in \mathbf{V} - \{\mathbf{0}\}$. Consider any finite set $\mathscr{V} = \{v^{(1)}, \ldots, v^{(k)}\} \subset \mathbf{V}$ satisfying $\operatorname{Lin}(\mathscr{A}\mathscr{V}) = \mathbf{V}$ and put

$$\vartheta = \|x\|(\max_{1 \leqslant i \leqslant k} \|v^{(i)}\|)^{-1}.$$

Since $\mathscr{V} \neq \{\mathbf{0}\}$ (because $\operatorname{Lin}(\mathscr{A}\mathscr{V}) = \mathbf{V}$ and $\dim \mathbf{V} \geqslant 1$) and $x \neq \mathbf{0}$ we have $\vartheta \in (0, \infty)$. Put $\mathscr{W} = \{w^{(1)}, \ldots, w^{(k)}, w^{(k+1)}\}$, where $w^{(1)} = \vartheta v^{(1)}$, ..., $w^{(k)} = \vartheta v^{(k)}$, $w^{(k+1)} = x$. It is obvious that $\operatorname{Lin}(\mathscr{A}\mathscr{W}) = \mathbf{V}$. Further, let $c_1 = 0$, ..., $c_k = 0$, $c_{k+1} = 1$, $c_{k+2} = 0$, ..., $c_{2(k+1)} = 0$ and let $Q_1 = A$, ..., $Q_{2(k+1)} = A$. We have

$$\sum_{i=1}^{k+1} (c_i Q_i - c_{k+1+i} Q_{k+1+i}) w^{(i)} = Q_{k+1} w^{(k+1)} = Ax,$$

so that $\big((c_1, \ldots, c_{2(k+1)}), (Q_1, \ldots, Q_{2(k+1)})\big) \in \mathscr{K}_{\mathscr{W}}(Ax)$. Thus

$$[Ax]^{\tau}_{\mathscr{W}} \leqslant \sum_{i=1}^{2(k+1)} c_i \tau(Q_i) = \tau(Q_{k+1}) = \tau(A).$$

We summarize

$$\frac{\|Ax\|}{\|x\|} \leqslant \frac{1}{\|x\|}(\max_{w \in \mathscr{W}} \|w\|)[Ax]^{\tau}_{\mathscr{W}} = \frac{1}{\|x\|}\|x\|[Ax]^{\tau}_{\mathscr{W}} \leqslant \tau(A)$$

(the first inequality holds by (3)).

(ii) Let $\mathscr{V} = \{v^{(1)}, \ldots, v^{(k)}\} \subset \mathbf{V}$, $\operatorname{Lin}(\mathscr{A}\mathscr{V}) = \mathbf{V}$. It is obvious that the set $\mathscr{J} = \big\{i \in \{1, \ldots, k\} \mid v^{(i)} \neq \mathbf{0}\big\}$ is non-empty. For any $x \in \mathbf{V}$, $(c, Q) \in \mathscr{K}_{\mathscr{V}}(x)$ we have

$$\begin{aligned}
\|x\| &= \Big\| \sum_{i=1}^{k} (c_i Q_i - c_{k+i} Q_{k+i}) v^{(i)} \Big\| \\
&= \Big\| \sum_{i \in \mathscr{J}} (c_i Q_i - c_{k+i} Q_{k+i}) v^{(i)} \Big\| \\
&\leqslant (\max_{i \in \mathscr{J}} \|v^{(i)}\|) \Big(\sum_{i \in \mathscr{J}} c_i \frac{\|Q_i v^{(i)}\|}{\|v^{(i)}\|} + \sum_{i \in \mathscr{J}} c_{k+i} \frac{\|Q_{k+i} v^{(i)}\|}{\|v^{(i)}\|} \Big) \\
&\leqslant (\max_{i \in \mathscr{J}} \|v^{(i)}\|) \Big(\sum_{i \in \mathscr{J}} c_i \sup_{\substack{z \in \mathbf{V} \\ z \neq \mathbf{0}}} \frac{\|Q_i z\|}{\|z\|} + \sum_{i \in \mathscr{J}} c_{k+i} \sup_{\substack{z \in \mathbf{V} \\ z \neq \mathbf{0}}} \frac{\|Q_{k+i} z\|}{\|z\|} \Big) \\
\text{(by (4))} &\leqslant (\max_{i \in \mathscr{J}} \|v^{(i)}\|) \Big(\sum_{i \in \mathscr{J}} c_i \tau(Q_i) + \sum_{i \in \mathscr{J}} c_{k+i} \tau(Q_{k+i}) \Big) \\
&\leqslant (\max_{1 \leqslant i \leqslant k} \|v^{(i)}\|) \Big(\sum_{i=1}^{2k} c_i \tau(Q_i) \Big).
\end{aligned}$$

Hence

$$\|x\| \leqslant \big(\max_{v\in\mathscr{V}} \|v\|\big)[x]^{\tau}_{\mathscr{V}}$$

for all $x \in \mathbf{V}$. This inequality implies $[x]^{\tau}_{\mathscr{V}} > 0$ for all $x \in \mathbf{V} - \{0\}$. To complete the proof of (ii), it is sufficient to apply Theorem 1.1. □

Remark. The condition (IV) is not essential for the validity of Theorems 1.1 and 1.2 as can be observed from their proofs.

The following characterization of the inverse matrix can be set as an example of the use of Theorem 1.2.

Corollary. *Let $\|\ \|$ be a norm on $\mathbf{R}^n$, let $\|\ \|_1$ be the l_1-norm on $\mathbf{R}^n$, let U be a real regular $n \times n$ matrix and let $u^{(1)}, \ldots, u^{(n)}$ be columns of U. Then*

$$\|x\| \leqslant \big(\max_{1\leqslant i\leqslant n} \|u^{(i)}\|\big)\|U^{-1}x\|_1$$

for each $x \in \mathbf{R}^n$.

Proof. Put $\mathbf{V} = \mathbf{R}^n$, $\mathscr{V} = \{u^{(1)}, \ldots, u^{(n)}\}$, $\mathscr{A} = \{\alpha I_n\,;\ \alpha \in \mathbf{R}\}$ and let $\tau(\alpha I_n) = |\alpha|$ for each $\alpha \in \mathbf{R}$. It is easy to see that the conditions (I), ..., (IV) as well as (α) and (β) are valid. The set $\mathscr{V}$ is a base of $\mathbf{R}^n$, because the matrix U is regular. It follows that $\mathrm{Lin}(A\mathscr{V}) = \mathbf{R}^n = \mathbf{V}$.

Let $x \in \mathbf{R}^n$. Put $a = (a_1, \ldots, a_n)^T = U^{-1}x$ and denote $a_i^+ = \frac{1}{2}\big(|a_i| + a_i\big)$ and $a_i^- = \frac{1}{2}\big(|a_i| - a_i\big)$. Since

$$\begin{aligned}
[x]^{\tau}_{\mathscr{V}} &= \inf\Big\{\sum_{i=1}^{2n} c_i\tau(Q_i) \mid (c, Q) \in \mathscr{K}_{\mathscr{V}}(x)\Big\}\\
&= \inf\Big\{\sum_{i=1}^{2n} c_i|\alpha_i| \mid c \in [0,\infty)^{2n},\ \alpha \in \mathbf{R}^{2n} \text{ and}\\
&\qquad c_i\alpha_i - c_{n+i}\alpha_{n+i} = a_i \text{ for all } i = 1,\ldots,n\Big\}\\
&\geqslant \inf\Big\{\sum_{i=1}^{n} |c_i\alpha_i - c_{n+i}\alpha_{n+i}| \mid c \in [0,\infty)^{2n},\ \alpha \in \mathbf{R}^{2n} \text{ and}\\
&\qquad c_i\alpha_i - c_{n+i}\alpha_{n+i} = a_i \text{ for all } i = 1,\ldots,n\Big\}\\
&= \sum_{i=1}^{n} |a_i| = \|U^{-1}x\|_1 = \sum_{i=1}^{n} a_i^+ + a_i^-
\end{aligned}$$

and

$$\big((a_1^+, \ldots, a_n^+, a_1^-, \ldots, a_n^-), (I_n, \ldots, I_n)\big) \in \mathscr{K}_{\mathscr{V}}(x)$$

we have $[x]^{\tau}_{\mathscr{V}} = \|U^{-1}x\|_1$. Theorem 1.2(ii) completes the proof. □

Lemma 1.1. *The set $\mathscr{A}$ is closed.*

Proof. Consider any $A \in \mathscr{A}$. The condition (I) implies that $\{B-A\,;\ B \in \mathscr{A}\}$ is a vector subspace of $\mathbf{M}_n$. Since $\mathbf{M}_n$ is a finite-dimensional vector space, the subspace $\{B - A\,;\ B \in \mathscr{A}\}$ is closed. Thus $\mathscr{A}$ is a closed set as well. □

Lemma 1.2. *Let τ be a nonnegative function on $\mathscr{A}$ satisfying the conditions (δ) and (ε). Let $\{a_q\}_{q=1}^{\infty}$ be a sequence of nonnegative real numbers and let $\{A_q\}_{q=1}^{\infty}$ be a convergent sequence of matrices of $\mathscr{A}$ such that*

$$\lim_{q\to\infty} a_q\tau(A_q) = 0,$$
$$\tau(\lim_{q\to\infty} A_q) = 0,$$
$$\tau(A_q) > 0 \quad \textit{for each } q \in \mathbf{N}.$$

Then there exist a sequence $\{b_r\}_{r=1}^{\infty}$ of nonnegative real numbers, a convergent sequence $\{B_r\}_{r=1}^{\infty}$ of matrices of $\mathscr{A}$ and an increasing sequence $\{q_r\}_{r=1}^{\infty}$ of positive integers such that

$$\lim_{r\to\infty} b_r\tau(B_r) = 0,$$
$$\tau(\lim_{r\to\infty} B_r) > 0,$$
$$\tau(B_r) > 0 \quad \textit{for each } r \in \mathbf{N},$$
$$b_r B_r x = a_{q_r} A_{q_r} x \quad \textit{for each } r \in \mathbf{N},\ x \in \mathbf{V}.$$

Proof. Denote $A = \lim_{q\to\infty} A_q$, $\mathscr{Z} = \{Z \in \mathscr{A} \mid \tau(Z) = 0\}$. By assumption, $A \in \mathscr{Z}$. First suppose that $\mathscr{Z} = \{A\}$. Put

$$\alpha_q = \frac{1}{\langle\langle A - A_q\rangle\rangle}$$

($\tau(A_q) > 0$ and $\tau(A) = 0$, thus $A_q \neq A$, hence $\alpha_q \in (0,\infty)$),

$$c_q = \frac{a_q}{\alpha_q},$$
$$C_q = \alpha_q A_q + (1-\alpha_q)A$$

for each $q \in \mathbf{N}$. We have $\langle\langle A - C_q\rangle\rangle = 1$ for each $q \in \mathbf{N}$ and $C_q \in \mathscr{A}$ for each $q \in \mathbf{N}$ by condition (I). Hence, by Lemma 1.1, there exists a convergent subsequence $\{C_{q_r}\}_{r=1}^{\infty}$ such that $\lim_{r\to\infty} C_{q_r} \in \mathscr{A}$. Denote $B = \lim_{r\to\infty} C_{q_r}$ and $b_r = c_{q_r}$, $B_r = C_{q_r}$ for each $r \in \mathbf{N}$. The sequence $\{B_r\}_{r=1}^{\infty}$ is convergent and $\{b_r\}_{r=1}^{\infty}$ is a sequence of nonnegative real numbers. Since $\langle\langle A - B_r\rangle\rangle = 1$ for each $r \in \mathbf{N}$, we have $\langle\langle A - B\rangle\rangle = 1$. Hence $B \notin \{A\} = \mathscr{X}$, so that $\tau(B) > 0$. The condition (ε) implies $\tau(B_r) = \alpha_{q_r}\tau(A_{q_r})$ and

$$(5) \qquad b_r\tau(B_r) = a_{q_r}\tau(A_{q_r})$$

for each $r \in \mathbf{N}$, hence $\tau(B_r) > 0$ for each $r \in \mathbf{N}$ and $\lim_{r\to\infty} b_r\tau(B_r) = 0$. The condition ($\delta$) implies $b_rB_rx = a_{q_r}A_{q_r}x$ for each $r \in \mathbf{N}$, $x \in \mathbf{V}$. This completes the proof of the case $\mathscr{X} = \{A\}$.

Now, let $\mathscr{X} \neq \{A\}$. Put $\mathscr{X}^* = \{Z - A;\ Z \in \mathscr{X}\}$. Then $\mathscr{X}^*$ is a vector subspace of $\mathbf{M}_n$ by condition (ε). Since $1 \leqslant \dim \mathscr{X}^* \leqslant \dim \mathbf{M}_n = n^2$, there exists a finite orthonormal base $\{Y_1^*, \ldots, Y_\nu^*\}$ of $\mathscr{X}^*$. Put

$$\begin{aligned}
U_q &= A + \sum_{t=1}^{\nu} \langle A_q - A, Y_t^*\rangle\, Y_t^*,\\
\alpha_q &= \frac{1}{\langle\langle U_q - A_q\rangle\rangle},\\
c_q &= \frac{a_q}{\alpha_q},\\
C_q &= \alpha_q A_q + (1 - \alpha_q)U_q
\end{aligned}$$

for each $q \in \mathbf{N}$. Since $\mathscr{X}^*$ is a vector space, we have $U_q - A \in \mathscr{X}^*$, i.e. $U_q \in \mathscr{X}$ for each $q \in \mathbf{N}$. It implies $A_q \neq U_q$ for each $q \in \mathbf{N}$, because $\tau(A_q) > 0$ for each $q \in \mathbf{N}$. Thus the numbers α_q are well-defined.

For each $q \in \mathbf{N}$, we have

$$\begin{aligned}
\langle\langle A - C_q\rangle\rangle &= \langle\langle \alpha_q U_q - \alpha_q A_q + A - U_q\rangle\rangle\\
&\leqslant \alpha_q \langle\langle A_q - U_q\rangle\rangle + \langle\langle A - U_q\rangle\rangle\\
&= 1 + \langle\langle A - U_q\rangle\rangle
\end{aligned}$$

and $\lim_{q\to\infty} \langle\langle A - U_q\rangle\rangle = 0$ by the definition of the matrices U_q and A. Thus, $\{C_q\}_{q=1}^{\infty}$ is a bounded sequence. It follows that there exists a convergent subsequence $\{C_{q_r}\}_{r=1}^{\infty}$. The condition (I) implies that $C_q \in \mathscr{A}$ for each $q \in \mathbf{N}$. Thus, $\lim_{r\to\infty} C_{q_r} \in \mathscr{A}$ by Lemma 1.1. Put $B = \lim_{r\to\infty} C_{q_r}$ and $b_r = c_{q_r}$, $B_r = C_{q_r}$ for each $r \in \mathbf{N}$. The sequence $\{B_r\}_{r=1}^{\infty}$ is convergent and $\{b_r\}_{r=1}^{\infty}$ is a sequence of nonnegative real numbers. Since for each $r \in \mathbb{N}$

$$\begin{aligned}
U_{q_r} - A &= \sum_{t=1}^{\nu} \langle A_{q_r} - A, Y_t^* \rangle Y_t^* \\
&= \alpha_{q_r} \sum_{t=1}^{\nu} \langle A_{q_r} - A, Y_t^* \rangle Y_t^* + (1 - \alpha_{q_r}) \sum_{t=1}^{\nu} \langle A_{q_r} - A, Y_t^* \rangle Y_t^* \\
&= \alpha_{q_r} \sum_{t=1}^{\nu} \langle A_{q_r} - A, Y_t^* \rangle Y_t^* + (1 - \alpha_{q_r}) \sum_{t=1}^{\nu} \langle U_{q_r} - A, Y_t^* \rangle Y_t^* \\
&= \sum_{t=1}^{\nu} \langle B_r - A, Y_t^* \rangle Y_t^*,
\end{aligned}$$

$U_{q_r} - A$ is an orthogonal projection of $B_r - A$ to $\mathscr{X}^*$ (see for example [5]). It follows that

$$\begin{aligned}
\inf_{Z \in \mathscr{X}} \langle\langle B_r - Z \rangle\rangle &= \inf_{Z^* \in \mathscr{X}^*} \langle\langle B_r - A - Z^* \rangle\rangle \\
&= \langle\langle B_r - A - (U_{q_r} - A) \rangle\rangle = 1.
\end{aligned}$$

Hence,

$$\inf_{Z \in \mathscr{X}} \langle\langle B - Z \rangle\rangle = 1,$$

thus $B \notin \mathscr{X}$, i.e. $\tau(B) > 0$. The condition (ε) implies $\tau(B_r) = \alpha_{q_r} \tau(A_{q_r})$ and

$$(5) \qquad b_r \tau(B_r) = a_{q_r} \tau(A_{q_r})$$

for each $r \in \mathbb{N}$, hence $\tau(B_r) > 0$ for each $r \in \mathbb{N}$ and $\lim_{r \to \infty} b_r \tau(B_r) = 0$. The condition ($\delta$) implies $b_r B_r x = a_{q_r} A_{q_r} x$ for each $r \in \mathbb{N}$, $x \in \mathbf{V}$. This completes the proof of the case $\mathscr{X} \neq \{A\}$. □

Now, we are able to formulate

Theorem 1.3. *Let τ be a nonnegative function on $\mathscr{A}$ satisfying the conditions (α), ..., (ε) and let $\mathscr{V}$ be a finite subset of* $\mathbf{V}$ *such that* $\mathrm{Lin}(\mathscr{A}\mathscr{V}) = \mathbf{V}$. *Then $[\ ]^\tau_{\mathscr{V}}$ is a norm on* $\mathbf{V}$ *and*

$$\sup_{\substack{x \in \mathbf{V} \\ x \neq \mathbf{0}}} \frac{[Ax]^\tau_{\mathscr{V}}}{[x]^\tau_{\mathscr{V}}} \leqslant \tau(A)$$

for all $A \in \mathscr{A}$.

Proof. By Theorem 1.1, it is sufficient to show that $[x]^\tau_{\mathscr{V}} > 0$ for each $x \in \mathbf{V} - \{\mathbf{0}\}$. Let us suppose that there exists $x \in \mathbf{V} - \{\mathbf{0}\}$ such that $[x]^\tau_{\mathscr{V}} = 0$. Hence, there exists a sequence $(c^{(q)}, Q^{(q)}) \in \mathscr{K}_{\mathscr{V}}(x)$, $q \in \mathbb{N}$, such that

$$\lim_{q\to\infty} \sum_{i=1}^{2k} c_i^{(q)} \tau(Q_i^{(q)}) = 0.$$

Consider a matrix $Z \in \mathscr{A}$ such that $\tau(Z) = 0$ (such a matrix does exist by condition (γ)). Put

$$\alpha_i^{(q)} = \begin{cases} \frac{q^{-1}}{\langle\langle Q_i^{(q)} - Z \rangle\rangle} & \text{if } Q_i^{(q)} \neq Z, \\ 1 & \text{if } Q_i^{(q)} = Z, \end{cases}$$

$$d_i^{(q,0)} = \frac{c_i^{(q)}}{\alpha_i^{(q)}},$$

$$R_i^{(q,0)} = \alpha_i^{(q)} Q_i^{(q)} + (1 - \alpha_i^{(q)}) Z$$

for each $q \in \mathbb{N}$, $i \in \{1, \dots, 2k\}$. We have

$$(d^{(q,0)}, R^{(q,0)}) \in \mathscr{K}_{\mathscr{V}}(x) \quad \text{for each } q \in \mathbb{N} \tag{6}$$

by condition (δ). Further, we have

$$\sum_{i=1}^{2k} d_i^{(q,0)} \tau(R_i^{(q,0)}) = \sum_{i=1}^{2k} c_i^{(q)} \tau(Q_i^{(q)}) \quad \text{for each } q \in \mathbb{N}$$

by condition (ε), hence

$$\lim_{q\to\infty} \sum_{i=1}^{2k} d_i^{(q,0)} \tau(R_i^{(q,0)}) = 0. \tag{7}$$

Since $\lim_{q\to\infty} \left\langle\left\langle R_i^{(q,0)} - Z \right\rangle\right\rangle = 0$, the equality

$$\lim_{q\to\infty} R_i^{(q,0)} = Z \tag{8}$$

is valid for each $i \in \{1, \dots, 2k\}$.

Let $\{(d^{(q,1)}, R^{(q,1)})\}_{q=1}^{\infty}$ be an arbitrary subsequence of $\{(d^{(q,0)}, R^{(q,0)})\}_{q=1}^{\infty}$ such that either

$$\tau(R_1^{(q,1)}) = 0 \quad \text{for all } q \in \mathbb{N}$$

or

$$\tau(R_1^{(q,1)}) > 0 \quad \text{for all } q \in \mathbb{N}$$

(recall that $R^{(q,0)} = (R_1^{(q,0)}, \ldots, R_{2k}^{(q,0)}) \in \mathscr{A}^{2k}$, $q \in \mathbf{N}$). Now, construct step by set sequences $\{(d^{(q,2)}, R^{(q,2)})\}_{q=1}^{\infty}, \ldots, \{(d^{(q,2k)}, R^{(q,2k)})\}_{q=1}^{\infty}$ in the following way: if sequences $\{(d^{(q,1)}, R^{(q,1)})\}_{q=1}^{\infty}, \ldots, \{(d^{(q,i)}, R^{(q,i)})\}_{q=1}^{\infty}$ have already been constructed ($1 \leqslant i \leqslant 2k-1$), then let $\{(d^{(q,i+1)}, R^{(q,i+1)})\}_{q=1}^{\infty}$ be an arbitrary subsequence of $\{(d^{(q,i)}, R^{(q,i)})\}_{q=1}^{\infty}$ such that either

$$\tau(R_{i+1}^{(q,i+1)}) = 0 \quad \text{for all } q \in \mathbf{N}$$

or

$$\tau(R_{i+1}^{(q,i+1)}) > 0 \quad \text{for all } q \in \mathbf{N}.$$

Thus, the sequence $\{(d^{(q,2k)}, R^{(q,2k)})\}_{q=1}^{\infty}$ is a subsequence of $\{(d^{(q,0)}, R^{(q,0)})\}_{q=1}^{\infty}$ and, for all $1 \leqslant i \leqslant 2k$, we have either

$$\tau(R_i^{(q,2k)}) = 0 \quad \text{for all } q \in \mathbf{N}$$

or

$$\tau(R_i^{(q,2k)}) > 0 \quad \text{for all } q \in \mathbf{N}.$$

Put

$$\mathscr{J} = \{i \in \{1, \ldots, 2k\} \mid \tau(R_i^{(q,2k)}) > 0 \text{ for all } q \in \mathbf{N}\}. \tag{9}$$

Suppose that $\mathscr{J} = \emptyset$. The condition (δ) implies that

$$\sum_{i=1}^{2k} (d_i^{(q,2k)} R_i^{(q,2k)} - d_{k+i}^{(q,2k)} R_{k+i}^{(q,2k)}) v^{(i)} = 0 \neq x$$

for all $q \in \mathbf{N}$. This contradicts the relation (6), because $\{(d^{(q,2k)}, R^{(q,2k)})\}_{q=1}^{\infty}$ is a subsequence of $\{(d^{(q,0)}, R^{(q,0)})\}_{q=1}^{\infty}$. Thus, $\mathscr{J} \neq \emptyset$.

Put

$$\{(e^{(q,0)}, S^{(q,0)})\}_{q=1}^{\infty} = \{(d^{(q,2k)}, R^{(q,2k)})\}_{q=1}^{\infty}.$$

Let $\{(e^{(q,1)}, S^{(q,1)})\}_{q=1}^{\infty}$ be a subsequence of the seqeuence $\{(e^{(q,0)}, S^{(q,0)})\}_{q=1}^{\infty}$ defined in the following way:

Put

$$\{(e^{(q,1)}, S^{(q,1)})\}_{q=1}^{\infty} = \{(e^{(q,0)}, S^{(q,0)})\}_{q=1}^{\infty}$$

if $1 \notin \mathscr{J}$. On the other hand, if $1 \in \mathscr{J}$, Lemma 1.2 applied for $a_q = e_1^{(q,0)}$ and $A_q = S_1^{(q,0)}$ generates the sequences $\{b_r\}_{r=1}^{\infty}$, $\{B_r\}_{r=1}^{\infty}$ and $\{q_r\}_{r=1}^{\infty}$ (the assumptions of Lemma 1.2 are valid by (7), (8) and (9)); put

$$e_j^{(r,1)} = \begin{cases} e_j^{(q_r,0)} & \text{if } j \neq 1, \\ b_r & \text{if } j = 1, \end{cases}$$

$$S_j^{(r,1)} = \begin{cases} S_j^{(q_r,0)} & \text{if } j \neq 1, \\ B_r & \text{if } j = 1, \end{cases}$$

for all $j = 1, \ldots, 2k$ and all $r \in \mathbf{N}$. Thus, if $\tau(S_1^{(q,0)}) = 0$ for all $q \in \mathbf{N}$, then

$$\{(e^{(q,1)}, S^{(q,1)})\}_{q=1}^{\infty} = \{(e^{(q,0)}, S^{(q,0)})\}_{q=1}^{\infty};$$

if $\tau(S_1^{(q,0)}) > 0$ for all $q \in \mathbf{N}$, then the entries $e_1^{(q_r,0)}$ and $S_1^{(q_r,0)}$ are redefined by

$$e_1^{(q_r,0)} = b_r, \quad S_1^{(q_r,0)} = B_r$$

and $\{(e^{(q,1)}, S^{(q,1)})\}_{q=1}^{\infty}$ is the subsequence of the just modified $\{(e^{(q,0)}, S^{(q,0)})\}_{q=1}^{\infty}$ determined by the sequence of indices $\{q_r\}_{r=1}^{\infty}$. Now construct step by step sequences $\{(e^{(q,2)}, S^{(q,2)})\}_{q=1}^{\infty}, \ldots, \{(e^{(q,2k)}, S^{(q,2k)})\}_{q=1}^{\infty}$ in the analogous way:

Suppose that $\{(e^{(q,1)}, S^{(q,1)})\}_{q=1}^{\infty}, \ldots, \{(e^{(q,i)}, S^{(q,i)})\}_{q=1}^{\infty}$ have already been constructed ($1 \leqslant i \leqslant 2k-1$). If $i+1 \notin \mathscr{J}$ then put

$$\{(e^{(q,i+1)}, S^{(q,i+1)})\}_{q=1}^{\infty} = \{(e^{(q,i)}, S^{(q,i)})\}_{q=1}^{\infty}.$$

On the other hand, if $i+1 \in \mathscr{J}$, Lemma 1.2 applied for $a_q = e_{i+1}^{(q,i)}$ and $A_q = S_{i+1}^{(q,i)}$ generates the sequences $\{b_r\}_{r=1}^{\infty}$, $\{B_r\}_{r=1}^{\infty}$ and $\{q_r\}_{r=1}^{\infty}$ (the assumptions of Lemma 1.2 are valid by (7), (8) and (9)); put

$$e_j^{(r,i+1)} = \begin{cases} e_j^{(q_r,i)} & \text{if } j \neq i+1, \\ b_r & \text{if } j = i+1, \end{cases}$$

$$S_j^{(r,i+1)} = \begin{cases} S_j^{(q_r,i)} & \text{if } j \neq i+1, \\ B_r & \text{if } j = i+1, \end{cases}$$

for all $j = 1, \ldots, 2k$ and all $r \in \mathbf{N}$.

Let $\{(e^{(q)}, S^{(q)})\}_{q=1}^{\infty} = \{(e^{(q,2k)}, S^{(q,2k)})\}_{q=1}^{\infty}$. By the construction of the sequence $\{(e^{(q)}, S^{(q)})\}_{q=1}^{\infty}$ (see Lemma 1.2) we have

(10) $$\lim_{q\to\infty} e_i^{(q)} \tau(S_i^{(q)}) = 0 \quad \text{for all } i \in \mathscr{J},$$

(11) $$\tau(\lim_{q\to\infty} S_i^{(q)}) > 0 \quad \text{for all } i \in \mathscr{J},$$

(12) $$\tau(S_i^{(q)}) = 0 \quad \text{for all } q \in \mathbf{N}, \quad i \in \{1, \ldots, 2k\} - \mathscr{J},$$

(13) $$(e^{(q)}, S^{(q)}) \in \mathscr{K}_{\mathscr{V}}(x) \quad \text{for each } q \in \mathbf{N}.$$

Put $S_i = \lim_{q\to\infty} S_i^{(q)}$ for each $i \in \mathcal{J}$. Let $\gamma = \frac{1}{2}\min_{i\in\mathcal{J}} \tau(S_i)$. It is clear from (11) that $\gamma > 0$. The conditions (I) and (α) imply that τ is a continuous function on $\mathcal{A}$. The continuity of τ and (11) imply the existence of a $q_0 \in \mathbf{N}$ such that

$$\tau(S_i^{(q)}) > \gamma > 0 \quad \text{for all } q \geqslant q_0,\ i \in \mathcal{J}. \tag{14}$$

The relations (10) and (14) imply

$$\lim_{q\to\infty} e_i^{(q)} = 0 \quad \text{for all } i \in \mathcal{J}. \tag{15}$$

The condition (δ) and the relation (12) imply

$$\sum_{\substack{i=1\\ i\notin\mathcal{J}}}^{k} e_i^{(q)} S_i^{(q)} v^{(i)} - \sum_{\substack{i=k+1\\ i\notin\mathcal{J}}}^{2k} e_i^{(q)} S_i^{(q)} v^{(i-k)} = 0 \tag{16}$$

for all $q \in \mathbf{N}$. Finally, (15) and (16) imply

$$\lim_{q\to\infty} \sum_{i=1}^{2k} (e_i^{(q)} S_i^{(q)} - e_{k+i}^{(q)} S_{k+i}^{(q)}) v^{(i)} =$$
$$= \lim_{q\to\infty} \left(\sum_{\substack{i=1\\ i\in\mathcal{J}}}^{2k} e_i^{(q)} S_i^{(q)} v^{(i)} - \sum_{\substack{i=k+1\\ i\in\mathcal{J}}}^{2k} e_i^{(q)} S_i^{(q)} v^{(i-k)} \right)$$
$$= \sum_{\substack{i=1\\ i\in\mathcal{J}}}^{k} 0 \cdot S_i v^{(i)} - \sum_{\substack{i=k+1\\ i\in\mathcal{J}}}^{2k} 0 \cdot S_i v^{(i-k)} = 0 \neq x,$$

which contradicts (13). Thus, $[x]_{\mathcal{V}}^{\tau} > 0$. □

Theorem 1.4. *Let $\| \ \|$ be a norm on $\mathbf{V}$, let $\tau_{\| \ \|}$ be a function on $\mathcal{A}$ defined as*

$$\tau_{\| \ \|}(A) = \sup_{\substack{x\in\mathbf{V}\\ x\neq\mathbf{0}}} \frac{\|Ax\|}{\|x\|}, \quad A \in \mathcal{A},$$

and let $\mathcal{V}$ be a finite subset of $\mathbf{V}$ such that $\mathrm{Lin}(\mathcal{A}\mathcal{V}) = \mathbf{V}$. Then $[\ ]_{\mathcal{V}}^{\tau_{\| \ \|}}$ is a norm on $\mathbf{V}$,

$$\|x\| \leqslant (\max_{v\in\mathcal{V}} \|v\|)[x]_{\mathcal{V}}^{\tau_{\| \ \|}}$$

for all $x \in \mathbf{V}$ *and*

$$\sup_{\substack{x \in \mathbf{V} \\ x \neq 0}} \frac{[Ax]_{\mathscr{V}}^{\tau_{\|\ \|}}}{[x]_{\mathscr{V}}^{\tau_{\|\ \|}}} \leqslant \tau_{\|\ \|}(A)$$

for all $A \in \mathscr{A}$.

Proof. Since the conditions (α), ..., (ε) hold for the function $\tau_{\|\ \|}$ ((γ) holds according to (IV)), it is sufficient to apply Theorem 1.2 and Theorem 1.3. □

2. Nonnegative functions on stochastic matrices

Troughout this section we assume $n \geqslant 2$. Define the vector space $\mathbf{V}$ by $\mathbf{V} = \{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0\}$ and put $\mathscr{A} = \Sigma_n$. We observe that the condtions (I), ..., (IV) are satisfied: the validity of (I) and (II) is obvious; (III) is true since $I_n \in \Sigma_n$; (IV) holds because $R\mathbf{V} = \{\mathbf{0}\}$ for each $R \in \Sigma_n$ such that all columns of R are identical.

Lemma 2.1. *Let* τ *be a nonnegative function on* Σ_n *satisfying the conditions* (α), ..., (ε) *and let* $S \in \Sigma_n$. *Then* $\tau(S) = 0$ *if and only if all columns of* S *are identical.*

Proof. Let $\mathscr{R}_n$ be the set of all matrices of Σ_n which have identical columns (i.e. $R \in \mathscr{R}_n$ iff $R \in \Sigma_n$ and there exists a vector $a \in \mathbf{R}^n$ such that $R = a\mathbf{1}^T$). It is easy to see that $R^2 = R$ for each $R \in \mathscr{R}_n$. Hence

$$\tau(R) \in \{0\} \cup [1, \infty) \quad \text{for each } R \in \mathscr{R}_n \tag{17}$$

by condition (β).

Consider a matrix $Z \in \Sigma_n$ such that $\tau(Z) = 0$ (the existence of such a matrix is guarenteed by condition (γ)). The form of the vector space $\mathbf{V}$ ($\mathbf{V} = \{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0\}$) and condition ($\delta$) imply that $Z \in \mathscr{R}_n$. Further, the function τ is continuous (by condition (α)) and $\mathscr{R}_n$ is a convex set. We summarize that $\tau(R) = 0$ for each $R \in \mathscr{R}_n$ by (17).

Finally, if $S \in \Sigma_n - \mathscr{R}_n$, then $\tau(S) > 0$ by condition (δ). This completes the proof. □

According to Lemma 2.1 we can replace the conditions (α), ..., (ε) by the following ones:

(S1) τ is a convex function on Σ_n;

(S2) $\tau(S_1 S_2) \leqslant \tau(S_1)\tau(S_2)$ for all $S_1, S_2 \in \Sigma_n$;

(S3) for each matrix $S \in \Sigma_n$ the equality $\tau(S) = 0$ holds if and only if the matrix S has identical columns;

(S4) if R, $S \in \Sigma_n$ and if the matrix R has identical columns then $\tau(\alpha S+(1-\alpha)R) = \alpha\tau(S)$ holds for any $\alpha > 0$.

Lemma 2.2. *The equality* $\Sigma_n \mathscr{V} = \{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0\}$ *holds for each finite set* $\mathscr{V} \subset \{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0,\ x \neq \mathbf{0}\}$.

Proof. It suffices to prove that

$$\Sigma_n\{v\} = \{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0\}$$

for all $v \in \{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0,\ x \neq \mathbf{0}\}$. Let

$$v, x \in \mathbf{R}^n, \quad \mathbf{1}^T v = 0, \quad \mathbf{1}^T x = 0, \quad v \neq \mathbf{0}.$$

Since $v \neq \mathbf{0}$, there exists an index k such that $v_k \neq 0$. For all i, $j \in \{1, \ldots, n\}$ put

$$s_{ij} = \begin{cases} \dfrac{1}{n} & \text{if } j \neq k, \\[2ex] \dfrac{1}{n} + \dfrac{x_i}{v_k} & \text{if } j = k. \end{cases}$$

It is easy to see that $S = (s_{ij})_{i,j=1}^n \in \Sigma_n$ and $Sv = x$. □

Theorem 2.1. *Let* τ *be a nonnegative function on* Σ_n *satisfying the conditions* (S1), ..., (S4) *and let* $\mathscr{V}$ *be a finite subset of* $\{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0,\ x \neq \mathbf{0}\}$. *Then* $[\ ]_{\mathscr{V}}^{\tau}$ *is a norm on the space* $\{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0\}$ *and*

$$\sup_{\substack{x \in \mathbf{R}^n - \{0\} \\ \mathbf{1}^T x = 0}} \frac{[P^T x]_{\mathscr{V}}^{\tau}}{[x]_{\mathscr{V}}^{\tau}} \leqslant \tau(P^T)$$

for each $n \times n$ *stochastic matrix* P.

Proof. If suffices to apply Theorem 1.3 and Lemmas 2.1 and 2.2. □

Theorem 2.2. *Let* τ *be a nonnegative function on* Σ_n *satisfying the conditions* (S1), ..., (S4) *such that* $\tau(P^T) \leqslant 1$ *for each* $n \times n$ *stochastic matrix* P. *Then*

$$\tau(P^T) \geqslant \frac{1}{2} \max_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \sum_{k=1}^{n} |p_{ik} - p_{jk}|$$

for each $n \times n$ *stochastic matrix* P. *Moreover, for any finite set* $\mathscr{V} \subset \{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0,\ x \neq \mathbf{0}\}$ *there exists* $a > 0$ *such that*

$$[x]_{\mathscr{V}}^{\tau} = \sum_{i=1}^{n} a|x_i|$$

for each $x \in \mathbf{R}^n$, $\mathbf{1}^T x = 0$.

Proof. This theorem follows from Theorem 2.1 and from [2, Theorem 3] (or from [3]). □

Theorem 2.3. *Let* τ *be a nonnegative function on* Σ_n *satisfying the conditions* (S1), ..., (S4) *and let* P *be an* $n \times n$ *irreducible stochastic matrix. Then*

$$\max\{|\lambda|; \lambda \in \sigma(P) - \{1\}\} \leqslant \tau(P^T).$$

Proof. This theorem follows from Theorem 2.1 and from [4, Theorem 3.1]. □

Lemma 2.3. *Let* τ_1, τ_2 *be nonnegative functions on* Σ_n *satisfying the conditions* (S1), ..., (S4). *Then there exists* $c > 0$ *such that* $\tau_1(S) \leqslant c\tau_2(S)$ *for all* $S \in \Sigma_n$.

Proof. Let this propostition be not true, i.e., for each $q \in \mathbf{N}$ let there exist a matrix $S_q \in \Sigma_n$ such that

$$\tau_1(S_q) > q\tau_2(S_q).$$

Lemma 2.1 implies that, for each $q \in \mathbf{N}$, $\tau_1(S_q) = 0$ if and only if $\tau_2(S_q) = 0$. Hence $\tau_2(S_q) > 0$ for each $q \in \mathbf{N}$.

Consider any matrix $R \in \Sigma_n$ such that R has identical columns. Let $\{\lambda_q\}_{q=1}^{\infty}$ be a sequence of positive real numbers such that

$$\lim_{q\to\infty} \lambda_q S_q + (1-\lambda_q)R = R.$$

For all $q \in \mathbf{N}$ put $a_q = 1$, $A_q = \lambda_q S_q + (1-\lambda_q)R$. By Lemma 1.2, there exist a sequence $\{b_r\}_{r=1}^{\infty}$ of positive real numbers (it is easy to see from the proof of Lemma 1.2 that if $a_q > 0$ for all $q \in \mathbf{N}$ then $b_r > 0$ for all $r \in \mathbf{N}$), a convergent sequence $\{B_r\}_{r=1}^{\infty}$ of matrices of Σ_n and an increasing sequence $\{q_r\}_{r=1}^{\infty}$ of positive integers such that

$$\begin{gathered}\lim_{r\to\infty} b_r\tau_i(B_r) = 0,\\ \tau_i(\lim_{r\to\infty} B_r) > 0,\\ \tau_i(B_r) > 0 \text{ for each } r \in \mathbf{N},\\ b_r B_r x = A_{q_r} x \text{ for each } r \in \mathbf{N},\ x \in \mathbf{V},\end{gathered} \tag{18}$$

where $i = 1, 2$ (it is easy to see from the proof of Lemma 1.2 that the sequences $\{b_r\}_{r=1}^{\infty}$, $\{B_r\}_{r=1}^{\infty}$ and $\{q_r\}_{r=1}^{\infty}$ can coincide for both τ_1 and τ_2). Moreover, the equality (5) implies that

$$b_r\tau_i(B_r) = a_{q_r}\tau_i(A_{q_r}) \quad \text{for all } r \in \mathbf{N},\ i \in \{1,2\}. \tag{19}$$

The condition (S4) implies

$$\tau_i(A_{q_r}) = \lambda_{q_r}\tau_i(S_{q_r}) \quad \text{for all } r \in \mathbf{N},\ i \in \{1,2\}. \tag{20}$$

We summarize that

$$\begin{aligned}\lim_{r\to\infty}\frac{\tau_1(B_r)}{\tau_2(B_r)} &= \lim_{r\to\infty}\frac{b_r\tau_1(B_r)}{b_r\tau_2(B_r)} = \lim_{r\to\infty}\frac{a_{q_r}\tau_1(A_{q_r})}{a_{q_r}\tau_2(A_{q_r})}\\ &= \lim_{r\to\infty}\frac{\tau_1(S_{q_r})}{\tau_2(S_{q_r})} \geqslant \lim_{r\to\infty}\frac{q_r\tau_2(S_{q_r})}{\tau_2(S_{q_r})} = \lim_{r\to\infty} q_r = \infty\end{aligned}$$

by (19) and (20). This contradicts (18), because the functions τ_1 and τ_2 are continuous by (S1). □

Theorem 2.4. *Let τ be a nonnegative function on Σ_n satisfying the conditions (S1), …, (S4) and let P be an $n\times n$ irreducible aperiodic stochastic matrix. Put $Q = P^T$. Then*

$$\begin{aligned}\max\{|\lambda|;\ \lambda\in\sigma(P)-\{1\}\} &= \lim_{k\to\infty}\left[\tau(Q^k)\right]^{1/k}\\ &= \inf_{k\in\mathbf{N}}\left[\tau(Q^k)\right]^{1/k}.\end{aligned}$$

Proof. Put $\xi(P) = \max\{|\lambda|;\lambda\in\sigma(P)-\{1\}\}$. Since P is an irreducible aperiodic stochastic matrix, we have

$$\sigma(P^k) = \{\lambda^k;\ \lambda\in\sigma(P)\},$$

thus $\xi(P^k) = [\xi(P)]^k$. Hence $\xi(P) \leqslant [\tau(Q^k)]^{1/k}$ for all $k\in\mathbf{N}$ by Theorem 2.3.

Let $\|\ \|$ be a norm on $\mathbf{R}^n$ and put

$$\tau_{\|\ \|}(S) = \sup_{\substack{x\in\mathbf{V}\\ x\neq\mathbf{0}}}\frac{\|Sx\|}{\|x\|},\quad S\in\Sigma_n,$$

where $\mathbf{V} = \{x\in\mathbf{R}^n \mid \mathbf{1}^Tx = 0\}$. By Lemma 2.3, there exists a number $c>0$ such that $\tau(S)\leqslant c\tau_{\|\ \|}(S)$ for all $S\in\Sigma_n$. It is proved in [4, Theorem 3.3] that

$$\xi(P) = \lim_{k\to\infty}\left[\tau_{\|\ \|}(Q^k)\right]^{1/k} = \inf_{k\in\mathbf{N}}\left[\tau_{\|\ \|}(Q^k)\right]^{1/k}.$$

Since $\lim_{k\to\infty} c^{1/k} = 1$, the proof is complete. □

Remark. Theorem 2.3 and Theorem 2.4 were proved in [4] for functions

$$\tau_{\|\ \|}(P) = \sup_{\substack{x\in\mathbf{V}\\ x\neq 0}} \frac{\|x^T P\|}{\|x\|}, \quad P \in \mathscr{S}_n,$$

where $\|\ \|$ is a norm on $\mathbf{R}^n$ and $\mathbf{V} = \{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 0\}$.

Examples. Put

$$\tau_F(S) = \left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left(s_{ij} - \frac{1}{n}\sum_{k=1}^{n} s_{ik}\right)^2\right)^{1/2}, \quad S \in \Sigma_n,$$

$$\tau_m(S) = (n-1) \max_{\substack{1\leqslant i\leqslant n-1\\ 1\leqslant j\leqslant n-1}} |s_{ij} - s_{in}|, \quad S \in \Sigma_n,$$

$$\tau_s(S) = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} |s_{ij} - s_{in}|, \quad S \in \Sigma_n.$$

These function meet the demands of the conditions (S1), ..., (S4). The function τ_F is called the Frobenius coefficient of ergodicity and is studied in [4] and [8].

The verification of (S1), (S3) and (S4) is easy. It is proved in [8] that (S2) holds true for τ_F. We prove (S2) in the case of τ_m and τ_s. Let

$$S_1 = (s_{ij}^{(1)})_{i,j=1}^n \in \Sigma_n, \quad S_2 = (s_{ij}^{(2)})_{i,j=1}^n \in \Sigma_n.$$

Since

$$\begin{aligned}
&\left|\sum_{k=1}^{n} s_{ik}^{(1)} s_{kj}^{(2)} - \sum_{k=1}^{n} s_{ik}^{(1)} s_{kn}^{(2)}\right| \\
&\quad = \left| s_{in}^{(1)}(s_{nj}^{(2)} - s_{nn}^{(2)}) + \sum_{k=1}^{n-1} s_{ik}^{(1)}(s_{kj}^{(2)} - s_{kn}^{(2)})\right| \\
&\quad = \left| s_{in}^{(1)} \sum_{k=1}^{n-1}(s_{kn}^{(2)} - s_{kj}^{(2)}) + \sum_{k=1}^{n-1} s_{ik}^{(1)}(s_{kj}^{(2)} - s_{kn}^{(2)})\right| \\
&\quad = \left|\sum_{k=1}^{n-1}(s_{ik}^{(1)} - s_{in}^{(1)})(s_{kj}^{(2)} - s_{kn}^{(2)})\right| \\
&\quad \leqslant \sum_{k=1}^{n-1} |s_{ik}^{(1)} - s_{in}^{(1)}||s_{kj}^{(2)} - s_{kn}^{(2)}|,
\end{aligned}$$

we have

$$
\begin{aligned}
\tau_m(S_1S_2) &= (n-1)\max_{\substack{1\leqslant i\leqslant n-1\\ 1\leqslant j\leqslant n-1}}\left|\sum_{k=1}^{n} s_{ik}^{(1)}s_{kj}^{(2)} - \sum_{k=1}^{n} s_{ik}^{(1)}s_{kn}^{(2)}\right|\\
&\leqslant (n-1)\max_{\substack{1\leqslant i\leqslant n-1\\ 1\leqslant j\leqslant n-1}}\sum_{k=1}^{n-1}|s_{ik}^{(1)} - s_{in}^{(1)}||s_{kj}^{2} - s_{kn}^{(2)}|\\
&\leqslant (n-1)^2\max_{\substack{1\leqslant i\leqslant n-1\\ 1\leqslant j\leqslant n-1\\ 1\leqslant k\leqslant n-1}}|s_{ik}^{(1)} - s_{in}^{(1)}||s_{kj}^{2} - s_{kn}^{(2)}|\\
&\leqslant (n-1)^2\max_{\substack{1\leqslant i\leqslant n-1\\ 1\leqslant k\leqslant n-1}}|s_{ik}^{(1)} - s_{in}^{(1)}|\max_{\substack{1\leqslant j\leqslant n-1\\ 1\leqslant k\leqslant n-1}}|s_{kj}^{(2)} - s_{kn}^{(2)}|\\
&= \tau_m(S_1)\tau_m(S_2)
\end{aligned}
$$

and

$$
\begin{aligned}
\tau_s(S_1S_2) &= \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left|\sum_{k=1}^{n} s_{ik}^{(1)}s_{kj}^{(2)} - \sum_{k=1}^{n} s_{ik}^{(1)}s_{kn}^{(2)}\right|\\
&\leqslant \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\sum_{k=1}^{n-1}|s_{ik}^{(1)} - s_{in}^{(1)}||s_{kj}^{(2)} - s_{kn}^{(2)}|\\
&= \sum_{j=1}^{n-1}\sum_{k=1}^{n-1}\left(|s_{kj}^{(2)} - s_{kn}^{(2)}|\sum_{i=1}^{n-1}|s_{ik}^{(1)} - s_{in}^{(1)}|\right)\\
&\leqslant \sum_{j=1}^{n-1}\sum_{k=1}^{n}\left(|s_{kj}^{(2)} - s_{kn}^{(2)}|\sum_{g=1}^{n-1}\sum_{i=1}^{n-1}|s_{ig}^{(1)} - s_{in}^{(1)}|\right)\\
&= \tau_s(S_1)\tau_s(S_2).
\end{aligned}
$$

References

[1] *R.A. Horn, Ch.R. Johnson*: Matrix Analysis. Cambridge University Press, Cambridge, London, New York, New Rochelle, Melbourne and Sydney, 1986.

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

[3] *R. Kühne, A. Rhodius*: A characterization of Dobrushin's coeffiecient of ergodicity. Zeitschrift für Analysis und ihre Anwendungen *9 (2)* (1990), 187–188.

[4] *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.

[5] *W. Rudin*: Real and Complex Analysis. McGraw-Hill, Inc. New York, 1974.

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

[7] *E. Seneta*: Spectrum localization by ergodicity coefficients for stochastic matrices. Linear and Multilinear Algebra *14* (1983), 343–347.
[8] *E. Seneta, C. P. Tan*: The Euclidean and Frobenius ergodicity coefficients and spectrum localization. Bull. Malaysia Math. Soc. *(7) 1* (1984), 1–7.
[9] *C.P. Tan*: Coefficients of ergodicity with respect to vector norms. J. Appl. Prob. *20* (1983), 277–287.
[10] *C.P. Tan*: Spectrum localization of an ergodic stochastic matrix. Bull. Inst. Math. Acad. Sinica *12* (1984), 147–151.
[11] *C.P. Tan*: Spectrum localization using Hölder norms. Houston J. Math. *12* (1986), 441–449.

Author's address: Department of Probability and Statistics, Charles University, Sokolovská 83, 186 00 Praha 8, Czech Republic.