Electronic Journal of Linear Algebra

Volume 23 *ELA Volume 23 (2012)* Article 69

2012

On the approximation of ergodic projections and stationary distributions of stochastic matrices

Bernd Schomburg
bernd.schomburg@me.com

Recommended Citation
Schomburg, Bernd. (2012), "On the approximation of ergodic projections and stationary distributions of stochastic matrices", *Electronic Journal of Linear Algebra*, Volume 23.
DOI: https://doi.org/10.13001/1081-3810.1570

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012
ELA
http://math.technion.ac.il/iic/ela

ON THE APPROXIMATION OF ERGODIC PROJECTIONS AND STATIONARY DISTRIBUTIONS OF STOCHASTIC MATRICES*

BERND SCHOMBURG†

Abstract. Let A be a stochastic $n \times n$ matrix and P_A be the ergodic projection of A, i.e., the projection onto $N(I-A)$ along $R(I-A)$. This paper considers approximations of P_A and of stationary distributions of A by using appropriate families of stochastic matrices induced by A and derives error estimates in terms of the Dobrushin ergodicity coefficient of $(I-A)^{\#}$, the group inverse of $I-A$.

Key words. Stochastic matrices, Ergodic projections, Stationary distributions, Ergodicity condition number, Approximation.

AMS subject classifications. 15B51, 15A12, 41A25, 41A35.

1. Introduction.

Let A be a stochastic $n \times n$ matrix, i.e., A is real, $A \geq 0$ and $A\mathbf{e} = \mathbf{e}$ with $\mathbf{e} = [1, \ldots, 1]^T$, and P_A be its ergodic projection, i.e., the projection onto $N(I-A)$ along $R(I-A)$. In this note, we will consider approximations of P_A by appropriate families of stochastic matrices induced by A and derive error estimates in terms of the Dobrushin ergodicity coefficient of $(I-A)^{\#}$, the group inverse of $I-A$. Since every stationary distribution of A is of the form μP_A for a suitable probability distribution μ, these will in turn yield approximations of stationary distributions. We will also treat approximations using powers and forward products of stochastic matrices of the form $A_t = tA + (1-t)(\mathbf{e} \otimes \mu)$, $0 \leq t \leq 1$.

Notation. With $\mathbb{R}^{n\times n}$ we denote the $\mathbb{R}$-algebra of all real $n \times n$ matrices and with $\mathcal{M}_n^1 = \{\mu \in (\mathbb{R}^n)^* \mid \mu \geq 0, \sum_{j=1}^n \mu_j = 1\}$ the set of all n-dimensional probability distributions. For $A \in \mathbb{R}^{n\times n}$, $p_A(t) = \det(tI - A) \in \mathbb{R}[t]$ is the characteristic polynomial of A and $\sigma(A) = \{\lambda \in \mathbb{C} \mid p_A(\lambda) = 0\}$ its spectrum. We equip $\mathbb{R}^n$ with the ℓ_∞-norm $\|x\|_\infty = \max_{1\leq i\leq n} |x_i|$ and $(\mathbb{R}^n)^*$ with the (dual) ℓ_1-norm $\|\mu\|_1 = \sum_{i=1}^n |\mu_i|$. For $S = [s_{ij}]_{1\leq i,j\leq n} \in \mathbb{R}^{n\times n}$, we use the matrix norm $\|S\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^n |s_{ij}|$, so that

$$\|Sx\|_\infty \leq \|S\|_\infty \|x\|_\infty \quad \forall x \in \mathbb{R}^n, \quad \|\mu S\|_1 \leq \|\mu\|_1 \|S\|_\infty \quad \forall \mu \in (\mathbb{R}^n)^*.$$

Note that for stochastic A, we have $\|A\|_\infty = 1$ and $\sigma(A) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$. The unit circle $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$ will be denoted by S^1.

*Received by the editors on August 24, 2012. Accepted for publication on December 8, 2012. Handling Editor: Stephen J. Kirkland.

†PO Box 1644, 53734 Sankt Augustin, Germany (bernd.schomburg@me.com).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

ELA
http://math.technion.ac.il/iic/ela

2. Tools. We recall some facts from linear algebra and the theory of stochastic matrices:

For a given matrix $S \in \mathbb{R}^{n\times n}$, there is at most one $X \in \mathbb{R}^{n\times n}$ such that

$$SXS = S, \quad XSX = X \quad \textit{and} \quad SX = XS.$$

If such an X exists, it is called the *group (generalized) inverse of* S and denoted by $S^\#$. One can show that $S^\#$ exists if and only if $\mathbb{R}^n = N(S) \oplus R(S)$. Note that $N(S^\#) = N(S)$ and that $SS^\#$ is the projection onto $R(S)$ along $N(S)$; thus, if $T \in \mathbb{R}^{n\times n}$ and $N(S) \subseteq N(T)$, then

$$T = TSS^\#. \tag{2.1}$$

If $\alpha \neq 0$, then $(\alpha S)^\# = \alpha^{-1} S^\#$. For the spectra of S and $S^\#$, we have the relation $\sigma(S^\#) \setminus \{0\} = \left\{\frac{1}{\lambda} \mid \lambda \in \sigma(S) \setminus \{0\}\right\}$. If S is a projection, i.e., if $S^2 = S$, then $S^\# = S$.

In the following lemma, we summarize some well-known results about stochastic matrices.

Lemma 2.1. *Let $A \in \mathbb{R}^{n\times n}$ be stochastic.*

(i) *The limit $P_A = \lim_{k\to\infty} \frac{1}{k}\sum_{j=0}^{k-1} A^j$ exists; P_A is the projection onto $N(I-A)$ along $R(I-A)$, in particular, $P_A^2 = P_A = AP_A = P_A A$. ($P_A$ is called the ergodic projection of A.)*
(ii) *A is semiconvergent (i.e., $\lim_{k\to\infty} A^k$ exists) if and only if $\sigma(A) \cap S^1 = \{1\}$, in which case $P_A = \lim_{k\to\infty} A^k$.*
(iii) *$I - A$ has a group inverse for which $(I-A)(I-A)^\# = I - P_A$.*
(iv) *A probability distribution $\nu \in \mathcal{M}_n^1$ is* stationary *with respect to A (i.e., $\nu A = \nu$) if and only if $\nu = \mu P_A$ for some $\mu \in \mathcal{M}_n^1$. A has a unique stationary distribution ν if and only if* $\dim N(I-A) = 1$*; in this case $N(I-A) = \langle \boldsymbol{e} \rangle$ and $P_A = \boldsymbol{e} \otimes \nu$.*

Note that (i) is a finite-dimensional version of the mean ergodic theorem, see e.g. Meyer [13, p. 697], and that (iv) is an easy consequence of (i). For (ii), cf. Berman and Plemmons [2, p. 152]. Part (iii) is contained in Meyer [11, Theorem 2.2].

We recall the following definition of the Dobrushin ergodicity coefficient.

Definition 2.2. Let $A \in \mathbb{R}^{n\times n}$ be a matrix with equal row sums, i.e., $\mathbf{e}$ is an eigenvector of A. Then the Dobrushin *ergodicity coefficient* $\tau_1(A)$ is defined by:

$$\tau_1(A) = \sup\{\|\nu A\|_1 \mid \nu \in (\mathbb{R}^n)^*,\ \|\nu\|_1 = 1,\ \nu \cdot \mathbf{e} = 0\}.$$

Definition 2.3. Let $A \in \mathbb{R}^{n\times n}$ be a stochastic matrix. Then we call

$$\kappa(A) = \tau_1((I-A)^\#)$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

the *ergodicity condition number* of A. ($\kappa(A)$ is well-defined since $(I-A)^{\#}\mathbf{e} = 0$. Note that, following [5], this condition number is also denoted by κ_6.)

The next result is a useful tool to calculate ergodicity coefficients, cf. Seneta [15, p. 584].

LEMMA 2.4. *Let $A = [a_{ik}]_{1\leq i,k\leq n}$ be a real matrix with equal row sums. Then*

$$\tau_1(A) = \sup\left\{ \frac{\|\nu A - \nu' A\|_1}{\|\nu - \nu'\|_1} \mid \nu, \nu' \in \mathcal{M}_n^1, \nu \neq \nu' \right\} = \frac{1}{2} \max_{1\leq i,j\leq n} \sum_{k=1}^{n} |a_{ik} - a_{jk}|.$$

In particular, if A is stochastic, $0 \leq \tau_1(A) \leq 1$.

We note the following well-known consequence of Definition 2.2.

LEMMA 2.5. *Let $S, T \in \mathbb{R}^{n\times n}$ be matrices with eigenvector $\boldsymbol{e}$. Then*

$$\tau_1(ST) \leq \tau_1(S)\tau_1(T),$$

and, if in addition $S\boldsymbol{e} = 0$,

$$\|ST\|_\infty \leq \|S\|_\infty \tau_1(T).$$

REMARK 2.6. Let $A, \hat{A} \in \mathbb{R}^{n\times n}$ be stochastic with stationary distributions π and $\hat{\pi}$, respectively. If $\dim N(I - A) = 1$ (in which case π is uniquely determined), then

$$\pi - \hat{\pi} = \hat{\pi}(A - \hat{A})(I - A)^{\#},$$

a result which is essentially due to Meyer [12], after previous work by Schweitzer [14]. Lemma 2.5 then immediately leads to (cf. Seneta [18])

$$\|\pi - \hat{\pi}\|_1 \leq \|A - \hat{A}\|_\infty \kappa(A).$$

Let $A \in \mathbb{R}^{n\times n}$ be stochastic. By Lemma 2.5,

$$\tau_1(A^j - P_A) \leq \|A^j - P_A\|_\infty \leq \|I - P_A\|_\infty \tau_1(A^j - P_A) \leq 2\tau_1(A^j - P_A),$$

and thus, A is semiconvergent if and only if $\lim_{k\to\infty} \tau_1(A^k - P_A) = 0$. If A is semiconvergent and $k \in \mathbb{N}$ such that $\tau_1(A^k - P_A) < 1$, then

$$\tau_1(A^j - P_A) \leq \tau_1(I - P_A)\tau_1(A^k - P_A)^{\lfloor j/k\rfloor} \leq 2\tau_1(A^k - P_A)^{\lfloor j/k\rfloor} \tag{2.2}$$

for all $j \geq 0$, where we have used that for general r

$$\tau_1(A^r - P_A) = \tau_1(A^r(I - P_A)) \leq \tau_1(I - P_A).$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012
ELA
http://math.technion.ac.il/iic/ela

We call A *ergodic* if there is a $\psi \in \mathcal{M}_n^1$ such that $\psi = \lim_{k\to\infty} \nu A^k$ for all $\nu \in \mathcal{M}_n^1$. In this case $P_A = \mathbf{e} \otimes \psi$ and ψ is called the *limiting distribution* of A. A is ergodic if and only if $\lim_{k\to\infty} \tau_1(A^k) = 0$. If $A \in \mathbb{R}^{n\times n}$ is ergodic with limiting distribution ψ and if k is such that $\tau_1(A^k) < 1$, then

$$\|\psi - \nu A^j\|_1 \le 2\tau_1(A^k)^{\lfloor j/k \rfloor} \tag{2.3}$$

for all $\nu \in \mathcal{M}_n^1$ and $j \ge 0$, cf. Seneta [17, pp. 191–193].

Let $\mathcal{S} = (S_j)_{j\in\mathbb{N}}$ be a sequence of stochastic matrices. The family $\mathcal{T}(\mathcal{S}) = (T_{p,r})_{p\ge 0, r\in\mathbb{N}}$ of *forward products* associated with $\mathcal{S}$ is given by

$$T_{p,r} = S_{p+1} \cdots S_{p+r}.$$

It is called *weakly ergodic* if $\lim_{r\to\infty} \tau_1(T_{p,r}) = 0$ for all $p \ge 0$, and *strongly ergodic* if there is a $\psi \in \mathcal{M}_n^1$ such that $\lim_{r\to\infty} \nu T_{p,r} = \psi$ for all $\nu \in \mathcal{M}_n^1$ and $p \ge 0$; in the latter case ψ is called the limiting distribution of $\mathcal{T}(\mathcal{S})$. We recall the following theorem of Isaacson and Madsen [8, Theorem V.4.3] which we supplement by an error estimate.

THEOREM 2.7. *Let $\mathcal{S} = (S_j)_{j\in\mathbb{N}}$ be a sequence of stochastic matrices such that $\mathcal{T}(\mathcal{S})$ is weakly ergodic and assume that there is a sequence $(\psi_j)_{j\in\mathbb{N}}$ such that $\psi_j = \psi_j S_j$ for all $j \in \mathbb{N}$ and*

$$\sum_{j=1}^{\infty} \|\psi_j - \psi_{j+1}\|_1 < \infty. \tag{2.4}$$

Then $\psi = \lim_{j\to\infty} \psi_j$ exists and $\mathcal{T}(\mathcal{S})$ is strongly ergodic with limiting distribution ψ. Furthermore

$$\|\psi - \nu T_{p,r}\|_1 \le \alpha_{p+r} + 2 \min_{1\le s<r} (\alpha_{p+s+1} + \tau_1(T_{p+s,r-s})) \tag{2.5}$$

for all $\nu \in \mathcal{M}_n^1$, $p \ge 0$ and $r \in \mathbb{N}$ where

$$\alpha_k = \sum_{j=k}^{\infty} \|\psi_j - \psi_{j+1}\|_1.$$

Proof. Cf. [1, pp. 215–216]. First note that (2.4) implies that $(\psi_j)_{j\in\mathbb{N}}$ is a Cauchy sequence in the complete metric space $\mathcal{M}_n^1$ (with the metric induced by the ℓ_1-norm), thus there is a $\psi \in \mathcal{M}_n^1$ such that $\psi = \lim_{j\to\infty} \psi_j$ and

$$\|\psi - \psi_k\|_1 \le \alpha_k \quad \forall k \in \mathbb{N}.$$

Next, since $\psi_j S_j = \psi_j$,

$$\psi - \psi T_{p,r} = (\psi - \psi_{p+r}) + \sum_{l=1}^{r-1} (\psi_{p+l+1} - \psi_{p+l}) T_{p+l,r-l} + (\psi_{p+1} - \psi) T_{p,r}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

ELA
http://math.technion.ac.il/iic/ela

Thus,

$$\|\psi - \psi T_{p,r}\|_1 \leq \alpha_{p+r} + 2\alpha_{p+1}.$$

Finally, if $1 \leq s < r$,

$$\begin{aligned}\|\psi - \nu T_{p,r}\|_1 &\leq \|\psi - \psi T_{p+s,r-s}\|_1 + \|(\psi - \nu T_{p,s})T_{p+s,r-s}\|_1 \\ &\leq \alpha_{p+r} + 2\alpha_{p+s+1} + 2\tau_1(T_{p+s,r-s}),\end{aligned}$$

which proves (2.5). Since $\mathcal{T}(\mathcal{S})$ is weakly ergodic and $(\alpha_k)_{k\in\mathbb{N}}$ is a monotone non-increasing null sequence, the assertion follows. ◻

3. Approximations of P_A. For all theorems in Sections 3, 4 and 5, let $A \in \mathbb{R}^{n\times n}$ be stochastic and P_A be its ergodic projection. Note that the case of semiconvergent matrices is already covered by (2.2) although this error estimate is primarily of theoretical interest. We thus focus on the general case and start with a quantitative version of the mean ergodic theorem.

THEOREM 3.1. *We have*

$$\|P_A - \frac{1}{k}\sum_{j=0}^{k-1} A^j\|_\infty \leq \frac{2}{k}\kappa(A) \quad \forall k \in \mathbb{N}. \tag{3.1}$$

Proof. Let $T_k = \frac{1}{k}\sum_{j=0}^{k-1} A^j$. If we use (2.1) with $T = P_A - T_k$ and $S = I - A$, we get

$$P_A - T_k = \frac{1}{k}(A^k - I)(I - A)^{\#}.$$

By Lemma 2.5 we obtain

$$\|P_A - T_k\|_\infty \leq \frac{1}{k}\|A^k - I\|_\infty \kappa(A) \leq \frac{2}{k}\kappa(A). \quad ◻$$

EXAMPLE 3.2. Let A be the irreducible permutation matrix

$$A = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 \end{bmatrix}.$$

Then $(I - A)^{\#}$ is the circulant matrix

$$\begin{bmatrix} c_1 & c_2 & \cdots & c_n \\ c_n & c_1 & \cdots & c_{n-1} \\ \vdots & & & \vdots \\ c_2 & \cdots & c_n & c_1 \end{bmatrix}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

with $c_j = \frac{n+1-2j}{2n}$. Thus,

$$\kappa(A) = \begin{cases} \frac{n}{4}, & \text{if } n \text{ is even,} \\ \frac{n}{4} - \frac{1}{4n}, & \text{if } n \text{ is odd.} \end{cases}$$

Finally, $P_A = \mathbf{e} \otimes \gamma$ with $\gamma = [1/n, \ldots, 1/n]$ and

$$\|P_A - \frac{1}{k}\sum_{j=0}^{k-1} A^j\|_\infty = \frac{2}{k}\kappa(A)$$

for all k such that $k \equiv \lfloor \frac{n}{2} \rfloor$ *mod* n. Thus, the bound (3.1) is optimal.

For the next theorem, we note that since $\|A\|_\infty = 1$,

$$P_t = (1-t)(I - tA)^{-1}$$

exists for $t \in [0, 1)$ and is stochastic, too.

THEOREM 3.3. *$P_A = \lim_{t\to 1-} P_t$ with*

$$\|P_A - P_t\|_\infty \leq 2\kappa(A)(1-t) \quad \forall t \in [0,1) \tag{3.2}$$

and $(I - A)^\# = -\lim_{t\to 1-} \frac{d}{dt}P_t$ where

$$\|\frac{d}{dt}P_t\|_\infty \leq 4\kappa(A) \quad \forall t \in [0,1). \tag{3.3}$$

Proof. Using (2.1) with $T = P_t - P_A$ and $S = I - A$ we get

$$P_t - P_A = (P_t - P_A)(I - A)(I - A)^\# = (1-t)(I - P_tA)(I - A)^\#.$$

Thus, Lemma 2.5 yields

$$\|P_t - P_A\|_\infty \leq (1-t)\|I - P_tA\|_\infty \kappa(A) \leq 2\kappa(A)(1-t)$$

for all $t \in [0, 1)$, i.e., (3.2). Furthermore, it is easily verified that

$$S_t = -\frac{d}{dt}P_t = (I - A)(I - tA)^{-2},$$

$S_t(I - A) = (I - AP_t)^2$ and (using again (2.1))

$$S_t = (I - AP_t)^2(I - A)^\#.$$

Thus,

$$\lim_{t\to 1-} S_t = (I - AP_A)^2(I - A)^\# = (I - P_A)(I - A)^\# = (I - A)^\#.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

ELA
http://math.technion.ac.il/iic/ela

Finally we have by Lemma 2.5

$$\|S_t\|_\infty \leq \|(I - AP_t)^2(I - A)^\#\|_\infty \leq 4\kappa(A). \quad \square$$

REMARK 3.4. Let $A = \mathbf{e} \otimes \pi$ with $\pi \in \mathcal{M}_n^1$. Then $I - A$ is a projection, so

$$(I - A)^\# = I - A = I - \mathbf{e} \otimes \pi.$$

Thus, $\kappa(A) = 1$. Furthermore, $P_A = A$ and $P_t = (1 - t)I + tA$, so

$$\|P_t - P_A\|_\infty = (1 - t)\|I - A\|_\infty = 2(1 - \min_{1\leq i\leq n} \pi_i)(1 - t).$$

Thus, the bound (3.2) is optimal.

REMARK 3.5. The fact that $P_A = \lim_{t\to 1-} P_t$ is well-known, cf. Campbell and Meyer [4, Corollary 7.6.4] or, for a recent account with applications, Horn and Serra-Capizzano [6]. With regard to $(I - A)^\# = -\lim_{t\to 1-} \frac{d}{dt}P_t$, cf. Langville and Meyer [10, Theorem 6.1.3].

4. Approximations using perturbations of A. We continue with approximations using rank-one perturbations of A. To be more specific, let $\mu = [\mu_1, \cdots, \mu_n] \in \mathcal{M}_n^1$ and consider for $0 \leq t \leq 1$

$$A_t = tA + (1 - t)(\mathbf{e} \otimes \mu), \tag{4.1}$$

where the tensor product $\mathbf{e} \otimes \mu$ denotes the rank-one stochastic matrix

$$\begin{bmatrix} \mu_1 & \cdots & \mu_n \\ \vdots & & \vdots \\ \mu_1 & \cdots & \mu_n \end{bmatrix}.$$

First note, that

$$\tau_1(A_t) = t\tau_1(A) \leq t \tag{4.2}$$

(see e.g. [7, p. 154]), thus A_t is ergodic for $0 \leq t < 1$, and for its limiting distribution of π_t, we have by (2.3)

$$\|\pi_t - \nu A_t^k\|_1 \leq 2t^k \tau_1(A)^k$$

for all $\nu \in \mathcal{M}_n^1$ and $k \in \mathbb{N}$. Next note that $\pi_t = \mu P_t$ (cf. [20, p. 14]) since $\pi_t = \pi_t A_t = \pi_t tA + (1 - t)\mu$ if and only if $\pi_t(I - tA) = (1 - t)\mu$. Thus, by (3.2), we have the following theorem.

THEOREM 4.1. *We have*

$$\|\mu P_A - \nu A_t^k\|_1 \leq 2\kappa(A)(1 - t) + 2t^k\tau_1(A)^k \quad \forall \nu \in \mathcal{M}_n^1,\ t \in [0, 1),\ k \in \mathbb{N}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

ELA
http://math.technion.ac.il/iic/ela

REMARK 4.2. Matrices of the form (4.1) and their stationary distributions play a crucial role in ranking algorithms for search engines, cf. Langville and Meyer [10] for details and Bryan and Leise [3] for a short and instructive introduction to this topic.

THEOREM 4.3. *Let $\mu \in \mathcal{M}_n^1$ and $(t_j)_{j\in\mathbb{N}} \subset [0,1)$ be a monotone non-decreasing sequence converging to* 1 *such that $\sum_{j=1}^{\infty}(1-t_j) = \infty$. Then,*

$$\mu P_A = \lim_{r\to\infty} \nu A_{t_1} A_{t_2} \cdots A_{t_r} \quad \forall \nu \in \mathcal{M}_n^1.$$

Furthermore, if $\delta_j = 1 - t_j$, $\epsilon_j = \exp(-\sum_{l=1}^{j}\delta_l)$,

$$\|\mu P_A - \nu A_{t_1} A_{t_2} \cdots A_{t_r}\|_1 \leq 4\kappa(A)\delta_r + 2 \min_{1\leq s<r}\left(4\kappa(A)\delta_{s+1} + \frac{\epsilon_r}{\epsilon_s}\right) \tag{4.3}$$

for all $\nu \in \mathcal{M}_n^1$ and $r \in \mathbb{N}$.

Proof. Let $\mathcal{S} = (S_j)_{j\in\mathbb{N}}$ with $S_j = A_{t_j}$ and $\psi_j = \mu P_{t_j}$ so that ψ_j is stationary with respect to S_j and $\lim_{j\to\infty}\psi_j = \mu P_A$. Using Lemma 2.5, (4.2) and the assumption that $\sum_{j=1}^{\infty}(1-t_j) = \infty$ we get

$$\tau_1(A_{t_{p+1}} A_{t_{p+2}} \cdots A_{t_{p+r}}) \leq \prod_{k=p+1}^{p+r} t_k \leq \frac{\epsilon_{p+r}}{\epsilon_p} \to 0, \quad p \geq 0,$$

as $r \to \infty$, thus $\mathcal{T}(\mathcal{S})$ is weakly ergodic. Furthermore, by (3.3)

$$\begin{aligned}\|\psi_j - \psi_{j+1}\|_1 &\leq \|P_{t_j} - P_{t_{j+1}}\|_\infty \leq \sup_{t_j\leq t\leq t_{j+1}} \|\tfrac{d}{dt}P_t\|_\infty (t_{j+1}-t_j)\\ &\leq 4\kappa(A)(t_{j+1}-t_j).\end{aligned}$$

Thus,

$$\sum_{j=1}^{\infty} \|\psi_j - \psi_{j+1}\|_1 \leq 4\kappa(A)(1-t_1) < \infty.$$

By Theorem 2.7, $\mathcal{T}(\mathcal{S})$ is strongly ergodic with limiting distribution μP_A and the first part of the assertion follows. Furthermore, for $\alpha_k = \sum_{j=k}^{\infty}\|\psi_j - \psi_{j+1}\|_1$, we have

$$\alpha_k \leq 4\kappa(A)\delta_k.$$

Thus, (2.5) shows that

$$\begin{aligned}\|\mu P_A - \nu A_{t_1} A_{t_2} \cdots A_{t_r}\|_1 &\leq \alpha_r + 2\min_{1\leq s<r}(\alpha_{s+1} + \tfrac{\epsilon_r}{\epsilon_s})\\ &\leq 4\kappa(A)\delta_r + 2\min_{1\leq s<r}(4\kappa(A)\delta_{s+1} + \tfrac{\epsilon_r}{\epsilon_s}). \quad \square\end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

EXAMPLE 4.4. In order to illustrate the use of (4.3), let $t_j = 1 - \frac{1}{j}$. Then $\delta_j = \frac{1}{j}$ and $e^{-1}\frac{1}{j} \le \epsilon_j \le \frac{1}{j}$. Thus,

$$\|\mu P_A - \nu A_{t_1} A_{t_2} \cdots A_{t_r}\|_1 \le 4\kappa(A)\frac{1}{r} + 2e\left(\frac{4\kappa(A)e^{-1}}{s+1} + \frac{s}{r}\right) \tag{4.4}$$

for $1 \le s < r$. Now, if $r > 4\kappa(A)e^{-1}$, then we can choose $s = \lfloor (4\kappa(A)e^{-1}r)^{1/2} \rfloor$ in (4.4), and finally, obtain

$$\|\mu P_A - \nu A_{t_1} A_{t_2} \cdots A_{t_r}\|_1 \le 4\kappa(A)\frac{1}{r} + 8(\kappa(A)e)^{1/2}\frac{1}{r^{1/2}} < 10(\kappa(A)e)^{1/2}\frac{1}{r^{1/2}}.$$

REMARK 4.5. We mention another method of approximating P_A. For $t \in (0,1)$, consider

$$A(t) = tA + (1-t)I.$$

Then $I - A(t) = t(I-A)$ and, since $N(I-A(t)) = N(I-A)$ and $R(I-A(t)) = R(I-A)$, $P_{A(t)} = P_A$. Furthermore,

$$\sigma(A(t)) \cap S^1 = \{1\}.$$

Thus, $A(t)$ is semiconvergent (cf. Lemma 2.1 (ii)) with

$$P_A = \lim_{k\to\infty} A(t)^k.$$

Note that $(I - A(t))^\# = t^{-1}(I-A)^\#$ and $\kappa(A(t)) = t^{-1}\kappa(A)$.

5. Bounds for the ergodicity condition number. The results in Sections 3 and 4 show that effective bounds for $\kappa(A)$ are essential for estimating the respective approximation errors for P_A and μP_A. We briefly review existing results.

5.1. Lower bounds. We start with the observation that

$$\sigma((I-A)^\#) = \left\{\frac{1}{1-\lambda} \mid \lambda \in \sigma(A) \setminus \{1\}\right\} \cup \{0\}.$$

Thus, the spectral localization theorem [19, Theorem 2.10] shows that

$$\kappa(A) \ge \max\left\{\frac{1}{|1-\lambda|} \mid \lambda \in \sigma(A) \setminus \{1\}\right\}, \tag{5.1}$$

where the right-hand side is the spectral radius of $(I-A)^\#$. Inequality (5.1) indicates the difficulty in finding general upper bounds. The following result due to Kirkland, Neumann and Sze [9, Theorem 3.2] gives a sharp lower bound.

PROPOSITION 5.1. *If* $\dim N(I-A) = 1$, *then*

$$\kappa(A) \ge 1 - \frac{1}{n}. \tag{5.2}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

ELA
http://math.technion.ac.il/iic/ela

REMARK 5.2. Let $A = \frac{n}{n-1}\mathbf{e}\otimes\gamma - \frac{1}{n-1}I$ with $\gamma = [1/n, \ldots, 1/n]$. Then $(I-A)^{\#} = \frac{n-1}{n}(I - \mathbf{e}\otimes\gamma)$, and therefore,

$$\kappa(A) = \frac{n-1}{n}\tau_1(I - \mathbf{e}\otimes\gamma) = 1 - \frac{1}{n}.$$

Furthermore, if $n = 3$ and $A \in \mathbb{R}^{n\times n}$ as in Example 3.2, then $\kappa(A) = \kappa(A^2) = \frac{2}{3} = 1 - \frac{1}{n}$. One can show (cf. [9]) that these are the only examples where equality holds in (5.2).

5.2. Upper bounds. We continue with a theorem which generalizes a result of Seneta [16], [17] originally proved for ergodic A.

THEOREM 5.3. *Let A be stochastic and semiconvergent and $k \in \mathbb{N}$ such that $\tau_1(A^k - P_A) < 1$. Then,*

$$\kappa(A) \le \tau_1(I - P_A)\frac{k}{1 - \tau_1(A^k - P_A)}. \tag{5.3}$$

Proof. First note that (2.2) implies that $\sum_{j=0}^{\infty}\tau_1(A^j - P_A)$ converges. As a consequence, $\sum_{j=0}^{\infty}(A^j - P_A)$ converges, too, and

$$(I - A)^{\#} = \sum_{j=0}^{\infty}(A^j - P_A).$$

Finally, again by (2.2),

$$\begin{aligned} \tau_1((I-A)^{\#}) &\le \textstyle\sum_{j=0}^{\infty}\tau_1(A^j - P_A) \\ &\le \tau_1(I - P_A)\textstyle\sum_{j=0}^{\infty}\tau_1(A^k - P_A)^{\lfloor j/k\rfloor} \\ &= \tau_1(I - P_A)k\textstyle\sum_{l=0}^{\infty}\tau_1(A^k - P_A)^l \\ &= \tau_1(I - P_A)\tfrac{k}{1-\tau_1(A^k - P_A)}. \quad \square \end{aligned}$$

REMARK 5.4. (i) If A is ergodic, then $\tau_1(I - P_A) = 1$ and $\tau_1(A^k - P_A) = \tau_1(A^k)$. Thus, (5.3) can be written in the form

$$\kappa(A) \le \frac{k}{1 - \tau_1(A^k)}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

(ii) Let A be stochastic, $B = A(t)$ for some $t \in (0,1)$ and $k \in \mathbb{N}$ such that $\tau_1(B^k - P_A) < 1$ (see Remark 4.5). Then,

$$\kappa(A) = t\kappa(B) \leq 2t \frac{k}{1 - \tau_1(B^k - P_A)}.$$

(iii) In [18], Seneta proved that for general stochastic A,

$$\kappa(A) \leq \operatorname{tr}((I - A)^{\#}). \tag{5.4}$$

Examples given in [18] show that in general (5.4) will not provide satisfying estimates. We mention that $\operatorname{tr}((I - A)^{\#})$ can be calculated using the relation

$$\operatorname{tr}((I - A)^{\#}) = \frac{p_A^{(l+1)}(1)}{(l+1)p_A^{(l)}(1)},$$

where $l = \dim N(I - A)$ and $p_A^{(r)}$, $r \geq 1$, denotes the r-th order derivative of p_A.

Acknowledgment. The author would like to thank the anonymous referee for many helpful suggestions.

REFERENCES

[1] E. Behrends. *Introduction to Markov Chains.* Friedr. Vieweg & Sohn, Braunschweig/Wiesbaden, 2000.
[2] A. Berman und R.J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994.
[3] K. Bryan and T. Leise. The $25,000,000,000 eigenvector: The linear algebra behind Google. *SIAM Rev.*, 48:569–581, 2006.
[4] S.L. Campbell and C.D. Meyer. *Generalized Inverses of Linear Transformations.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2009.
[5] G.E. Cho and C.D. Meyer. Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra Appl.*, 335:137–150, 2001.
[6] R.A. Horn and S. Serra-Capizzano. A general setting for the parametric Google matrix. *Internet Math.*, 3:385–411, 2008.
[7] I.C.F. Ipsen and T.M. Selee. Ergodicity coefficients defined by vector norms. *SIAM J. Matrix Anal. Appl.*, 32:153–200, 2011.
[8] D.L. Isaacson and R.W. Madsen. *Markov Chains: Theory and Applications.* John Wiley & Sons, New York, 1976.
[9] S.J. Kirkland, M. Neumann, and N.-S. Sze. On optimal condition numbers for Markov chains. *Numer. Math.*, 110:521–537, 2008.
[10] A.N. Langville and C.D. Meyer. *Google's PageRank and Beyond: The Science of Search Engine Rankings.* Princeton University Press, Princeton, NJ, 2006.
[11] C.D. Meyer. The role of the group generalized inverse in the theory of finite Markov chains. *SIAM Rev.*, 17:443–464, 1975.
[12] C.D. Meyer. The condition of a finite Markov chain and perturbation bounds for the limiting probabilities. *SIAM J. Algebraic Discrete Meth.*, 1:273–283, 1980.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 989-1000, December 2012

ELA
http://math.technion.ac.il/iic/ela

[13] C.D. Meyer. *Matrix Analysis and Applied Linear Algebra.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[14] P.J. Schweitzer. Perturbation theory and finite Markov chains. *J. Appl. Probability*, 5:410–413, 1968.

[15] E. Seneta. Coefficients of ergodicity: structure and applications. *Adv. in Appl. Probab.*, 11:576–590, 1979.

[16] E. Seneta. Perturbation of the stationary distribution measured by ergodicity coefficients. *Adv. in Appl. Probab.*, 20:228–230, 1988.

[17] E. Seneta. Applications of ergodicity coefficients to homogeneous Markov chains. *Contemp. Math.*, 149:189–199, 1993.

[18] E. Seneta. Sensitivity of finite Markov chains under perturbation. *Statist. Probab. Lett.*, 17:163–168, 1993.

[19] E. Seneta. *Non-negative Matrices and Markov Chains*, revised reprint of second edition. Springer, New York, 2006.

[20] E. Seneta. Markov and the creation of Markov chains. In A.N. Langville and W.J. Stewart (editors), *MAM 2006: Markov Anniversary Meeting*, Boson Books, Raleigh, NC, 1–20, 2006.