COMPARISON OF PERTURBATION BOUNDS FOR THE STATIONARY DISTRIBUTION OF A MARKOV CHAIN

GRACE E. CHO * AND CARL D. MEYER †

November 14, 2000

Abstract. The purpose of this paper is to review and compare the existing perturbation bounds for the stationary distribution of a finite, irreducible, homogeneous Markov chain.

Key words. Markov chains, stationary distribution, stochastic matrix, group inversion, sensitivity analysis, perturbation theory, condition numbers

AMS subject classifications. 65F35, 60J10, 15A51, 15A12, 15A18

1. Introduction.

Let P be the transition probability matrix of an n state finite, irreducible, homogeneous Markov chain. The stationary distribution vector of P is the unique positive vector π^T satisfying

$$\pi^T P = \pi^T, \qquad \sum_{j=1}^{n} \pi_j = 1.$$

Suppose P is perturbed to a matrix $\widetilde{P}$, that is the transition probability matrix of an n state finite, irreducible, homogeneous Markov chain as well. Denoting the stationary distribution vector of $\widetilde{P}$ by $\widetilde{\pi}$, the goal is to describe the change $\pi - \widetilde{\pi}$ in the stationary distribution in terms of the change $E \equiv P - \widetilde{P}$ in the transition probability matrix. For suitable norms,

$$\|\pi - \widetilde{\pi}\| \leq \kappa \|E\|$$

for various different condition numbers κ. We review eight existing condition numbers $\kappa_1, \cdots, \kappa_8$. Most of the condition numbers we consider are expressed in terms of either the fundamental matrix of the underlying Markov chain or the group inverse of $I - P$. The condition number κ_8 is expressed in terms of *mean first passage times*, providing a qualitative interpretation of error bound. In § 4, we compare the condition numbers.

2. Notation.

Throughout the article the matrix P denotes the transition probability matrix of an n state finite, irreducible, homogeneous Markov chain $\mathcal{C}$ and π denotes the stationary distribution vector. Then

$$\pi^T P = \pi^T, \qquad \pi > 0, \qquad \pi^T e = 1,$$

where e is the column vector of all ones. The perturbed matrix $\widetilde{P} = P - E$ is the transition probability matrix of another n state finite, irreducible, homogeneous Markov chain $\widetilde{\mathcal{C}}$ with the stationary distribution vector $\widetilde{\pi}$:

$$\widetilde{\pi}^T \widetilde{P} = \widetilde{\pi}^T, \qquad \widetilde{\pi} > 0, \qquad \widetilde{\pi}^T e = 1.$$

*Mathematics & Engineering Analysis, The Boeing Company, P.O. Box 3707, MC 7L-22, Seattle, WA 98124-2207 (grace.e.cho@boeing.com, http://www4.ncsu.edu/~ehcho/home.html).

†Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205, USA (meyer@ncsu.edu, http://meyer.math.ncsu.edu/). The work of this author was supported in part by National Science Foundation grants CCR-9731856 and DMS-9714811.

The identity matrix of size n is denoted by I. For a matrix B, the (i, j) component of B is denoted by b_{ij}.

The 1-norm $\|v\|_1$ of a vector v is the absolute entry sum, and the ∞-norm $\|B\|_\infty$ of a matrix B is its maximum absolute row sum.

3. Condition Numbers of a Markov Chain. The norm-wise perturbation bounds we review in this section are of the following form:

$$\|\pi - \widetilde{\pi}\|_p \leq \kappa_l \|E\|_q,$$

where $(p, q) = (\infty, \infty)$ or $(1, \infty)$, depending on l.

Most of the perturbation bounds we will consider are in terms of one the two matrices related to the chain $\mathcal{C}$: the *fundamental matrix* and the *group inverse* of $A \equiv I - P$. The **fundamental matrix** of the chain $\mathcal{C}$ is defined by

$$Z \equiv (A + e\pi^T)^{-1}.$$

The **group inverse** of A is the unique square matrix $A^\#$ satisfying

$$AA^\# A = A, \; A^\# AA^\# = A^\#, \text{ and } AA^\# = A^\# A.$$

The lists of the condition numbers κ_l and the references are as follows:

Schweitzer 1968[17]	$\kappa_1 = \|Z\|_\infty$
Meyer 1980[12]	$\kappa_2 = \|A^\#\|_\infty$
Haviv & van Heyden 1984[5] Kirkland et al. 1998[9]	$\kappa_3 = \dfrac{\max_j (a^\#_{jj} - \min_i a^\#_{ij})}{2}$
Funderlic & Meyer 1986[4]	$\kappa_4 = \max_{i,j} \lvert a^\#_{ij}\rvert$
Seneta 1988[21]	$\kappa_5 = \dfrac{1}{1 - \tau_1(P)}$
Seneta 1991[22]	$\kappa_6 = \tau_1(A^\#) \;\; = \;\; \tau_1(Z)$
Ipsen & Meyer 1994[6] Kirkland et al. 1998[9]	$\kappa_7 = \dfrac{\min_j \|A^{-1}_{(j)}\|_\infty}{2}$
Cho & Meyer 1999[1]	$\kappa_8 = \dfrac{1}{2} \max_j \left[\dfrac{\max_{i \neq j} m_{ij}}{m_{jj}} \right],$

where $m_{ij}, \; i \neq j$, is the *mean first passage time* from state i to state j, and m_{jj} is the the *mean return time* state j.

3.1. Schweitzer 1968. Kemeny and Snell [8] call Z the fundamental matrix of the chain because most of the questions concerning the chain can be answered in terms of Z. For instance, the stationary distribution vector of the perturbed matrix $\widetilde{P}$ can be expressed in terms of π and Z:

$$\begin{aligned} \widetilde{\pi}^T &= \pi^T (I + EZ)^{-1}, \qquad \text{and} \\ \pi^T - \widetilde{\pi}^T &= \widetilde{\pi}^T EZ. \end{aligned} \tag{3.1}$$

(3.1) The equation (3.1), by Schweitzer [17], gives the first perturbation bound:

$$\|\pi - \widetilde{\pi}\|_1 \leq \|Z\|_\infty \|E\|_\infty.$$

We define

$$\kappa_1 \equiv \|Z\|_\infty.$$

3.2. Meyer 1980. The second matrix related to $\mathcal{C}$ is the group inverse of A. In his papers [11, 12], Meyer showed that the group inverse $A^\#$ can be used in a similar way Z is used, and, since

$$Z = A^\# + e\pi^T,$$

'all relevant information is contained in $A^\#$, and the term $e\pi^T$ is redundant' (Meyer [14]). In fact, in the place of (3.1), we have

$$\begin{aligned} \widetilde{\pi}^T &= \pi^T (I + EA^\#)^{-1}, \qquad \text{and} \\ \pi^T - \widetilde{\pi}^T &= \widetilde{\pi}^T EA^\#, \end{aligned} \tag{3.2}$$

and the resulting perturbation bound is (Meyer [12])

$$\|\pi - \widetilde{\pi}\|_1 \leq \|A^\#\|_\infty \, \|E\|_\infty.$$

We define the second condition number

$$\kappa_2 \equiv \|A^\#\|_\infty.$$

3.3. Haviv and van Heyden 1984 & Kirkland, Neumann, and Shader 1998. The perturbation bound in this section is derived from (3.2) with a use of the following lemma:

LEMMA 3.1. *For any vector d and for any vector c such that $c^T e = 0$,*

$$|c^T d| \leq \|c\|_1 \, \frac{max_{i,j} \, |d_i - d_j|}{2}.$$

The resulting perturbation bound is

$$\|\pi - \widetilde{\pi}\|_\infty \leq \frac{\max_j \, (a^\#_{jj} - \min_i \, a^\#_{ij})}{2} \, \|E\|_\infty.$$

(Haviv and van Heyden [5] & Kirkland, Neumann, and Shader [9]). We define

$$\kappa_3 \equiv \frac{\max_j \, (a^\#_{jj} - \min_i \, a^\#_{ij})}{2}.$$

3.4. Funderlic and Meyer 1986. The equation (3.2) provides a componentwise bound for the stationary distribution vector:

$$|\pi_j - \widetilde{\pi}_j| \leq \max_i \ |a_{ij}^{\#}| \ \|E\|_\infty,$$

which leads us to

$$\|\pi - \widetilde{\pi}\|_\infty \leq \max_{i,j} \ |a_{ij}^{\#}| \ \|E\|_\infty.$$

Funderlic and Meyer [4] called the number

$$\kappa_4 \equiv \max_{i,j} \ |a_{ij}^{\#}|$$

the **chain condition number**.

The behaviour of κ_4 is provided by Meyer [13, 14]. He showed that the size of κ_4 is primarily governed by how close the subdominant eigenvalues of the chain are to 1. (This is not true for arbitrary matrices. For an example, see [14, p.716].) To be precise, denoting the eigenvalues of P by $1, \ \lambda_2, \cdots, \lambda_n$, the lower bound and the upper bound of κ_4 are given by

$$\frac{1}{n \ \min_i |1-\lambda_i|} \leq \max_{i,j} \ |a_{ij}^{\#}| < \frac{2(n-1)}{\prod_i (1-\lambda_i)}. \tag{3.3}$$

Hence, if the chain is well-conditioned, then all subdominant eigenvalues must be well separated from 1, and if all subdominant eigenvalues are well separated from 1, then the chain must be well-conditioned.

In [13] and [14], it is indicated that the upper bound $2(n-1)/\prod_i (1-\lambda_i)$ in (3.3) is a rather conservative estimate of κ_4. If no single eigenvalue of P is extremely close to 1, but enough eigenvalues are within range of 1, then $2(n-1)/\prod_i (1-\lambda_i)$ is large, even if the chain is not too badly condition. Seneta [23] provides a condition number and its bounds to overcome this problem. (See § 3.6.)

3.5. Seneta 1988. The condition numbers κ_1, κ_2, and κ_4 are in terms of matrix norms. Seneta [21], however, proposed the *ergodicity coefficient* instead of the matrix norm. The **ergodicity coefficient** $\tau_1(B)$ of a matrix B with equal row sums b is defined by

$$\tau_1(B) \equiv \sup_{\substack{\|v\|_1=1 \\ v^T e=0}} \|v^T B\|_1.$$

Note that $\tau_1(B)$ is an ordinary norm of B on the hyperspace $\mathbf{H}^n \equiv \{v : v \in \mathbf{R}^n, \ v^T e = 0\}$ of $\mathbf{R}^n$. eigenvalues are to b. For more discussion and study, we refer readers to Dobrushin [2, 3], Issacson and Madsen [7], Seneta [18, 19, 20, 21, 22, 23], Tan [25, 26], Rothblum and Tan [16], Lešanovský [10], and Rhodius [15].

For the stochastic matrix P, the ergodicity coefficient satisfies $0 \leq \tau_1(P) \leq 1$. In case of $\tau_1(P) < 1$, we have a perturbation bound in terms of the ergodicity coefficient of P (Seneta [21] 1988):

$$\|\pi - \widetilde{\pi}\|_1 \leq \frac{1}{1-\tau_1(P)} \|E\|_\infty. \tag{3.4}$$

(For the case $\tau_1(P) = 1$, see Seneta [22].) We denote

$$\kappa_5 \equiv \frac{1}{1-\tau_1(P)}.$$

3.6. Seneta 1991. In the previous parts we noted that the group inverse $A^{\#}$ of $A = I - P$ can be used in place of Kemeny and Snell's fundamental matrix Z. In fact, if we use ergodicity coefficients as a measure of sensitivity of the stationary distribution, then Z and $A^{\#}$ give the exactly same information:

$$\kappa_6 \equiv \tau_1(A^{\#}) = \tau_1(Z),$$

which is the condition number in the perturbation bound given by Seneta [22]:

$$\|\pi - \widetilde{\pi}\|_1 \leq \tau_1(A^{\#})\,\|E\|_\infty \; (= \tau_1(Z)\,\|E\|_\infty).$$

In § 3.4, we observed that the size of the condition number $\kappa_4 = \max_{i,j}\ |a^{\#}_{ij}|$ is governed by the closeness of the subdominant eigenvalues of P to 1, giving the lower and upper bound for κ_4. However, the problem of overestimating the upper bound for κ_4 occurs if enough eigenvalues of P are within range of 1, even if no single eigenvalue λ_i of P is close to 1. The following bounds for κ_5 overcome this problem:

$$\frac{1}{\min_i |1-\lambda_i|} \leq \tau_1(A^{\#}) \leq \sum_i \frac{1}{1-\lambda_i} \leq \frac{n}{\min_i |1-\lambda_i|}.$$

Unlike the upper bound (3.3) for κ_4, the far left upper bound for κ_6 takes only the closest eigenvalue to 1 in to account. Hence, it shows that as long as the closest eigenvalue of P is not close to 1, the chain is well-conditioned.

3.7. Ipsen and Meyer 1994, & Kirkland, Neumann, and Shader 1998. Ipsen and Meyer [6] derived a set of perturbation bounds and showed that all stationary probabilities react in an uniform manner to perturbations in the transition probabilities. The main result of their paper is based on the following component-wise perturbation bounds:

$$\begin{aligned} \left|\frac{\pi_j - \widetilde{\pi}_j}{\pi_j}\right| &\leq \|A_{(j)}^{-1}\|_\infty\,\|E\|_\infty, && \text{for all } j = 1, 2, \cdots, n,\\ |\pi_k - \widetilde{\pi}_k| &\leq \min_j\ \|A_{(j)}^{-1}\|_\infty\,\|E\|_\infty, && \text{for all } k = 1, 2, \cdots, n, \end{aligned} \tag{3.5}$$

where $A_{(j)}$ is the principal submatrix of A obtained by deleting the j-th row and column from A.

Kirkland, Neumann, and Shader [9] improved the perturbation bound in (3.5) by a factor of 2:

$$\|\pi - \widetilde{\pi}\|_\infty < \frac{\min_j\ \|A_{(j)}^{-1}\|}{2}\|E\|_\infty,$$

giving the condition number

$$\kappa_7 \equiv \frac{\min_j\ \|A_{(j)}^{-1}\|_\infty}{2}.$$

3.8. Cho and Meyer 1999. In the previous sections, we saw a number of perturbation bounds for the stationary distribution vector of an irreducible Markov chain. Unfortunately, they provide little qualitative information about the sensitivity of the underlying Markov chain. Moreover, the actual computation of the corresponding condition number is usually expensive relative to computation of the stationary distribution vector itself.

In this section a perturbation bound is presented in terms of the structure of the underlying Markov chain. To be more precise, the condition number κ is in terms of *mean first passage times.*

For an n-state irreducible Markov chain $\mathcal{C}$, the **mean first passage time** m_{ij} from state i to state j $(j \neq i)$ is defined to be the expected number of steps to enter in state j for the first time, starting in state i. The **mean return time** m_{jj} of state j is the expected number of steps to return to state j for the first time, starting in state j.

Using the relationship

$$m_{jj} = \frac{1}{\pi_j}, \qquad \text{and}$$
$$a^{\#}_{ij} = a^{\#}_{jj} - \pi_j m_{ij}, i \neq j$$

we obtain the perturbation bound

$$\|\pi - \widetilde{\pi}\|_\infty \leq \frac{1}{2} \max_j \left[\frac{\max_{i \neq j} m_{ij}}{m_{jj}} \right] \|E\|_\infty$$

(Cho and Meyer [1]). We define

$$\kappa_8 \equiv \frac{1}{2} \max_j \left[\frac{\max_{i \neq j} m_{ij}}{m_{jj}} \right].$$

Viewing sensitivity in terms of mean first passage times can sometimes help practitioners decide whether or not to expect sensitivity in their Markov chain models merely by observing the structure of the chain without computing or estimating condition numbers. For example, consider chains consisting of a dominant central state with strong connections to and from all other states. Physical systems of this type have historically been called *mammillary systems* (see Sheppard and Householder [24] and Whittaker [27]). The simplest example of a mammillary Markov chain is one whose transition probability matrix has the form

$$P = \begin{pmatrix} 1-p_1 & 0 & \cdots & 0 & p_1 \\ 0 & 1-p_2 & \cdots & 0 & p_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-p_k & p_k \\ q_1 & q_2 & \cdots & q_k & 1-\sum_j q_j \end{pmatrix} \tag{3.6}$$

in which the p_i's and q_i's are not unduly small (say, $p_i > .5$ and $q_i \approx 1/k$). Mean first passage times m_{ij}, $i \neq j$, in mammillary structures are never very large. For example, in the simple mammillary chain defined by (3.6) it can be demonstrated that

$$m_{ij} = \begin{cases} \dfrac{1}{p_i} & \text{when } j = k+1 \\ \dfrac{1+\sigma}{q_j} - \dfrac{1}{p_j} & \text{when } i = k+1 \\ \dfrac{1+\sigma}{q_j} + \dfrac{1}{p_i} - \dfrac{1}{p_j} & \text{when } i, j \neq k+1 \end{cases} \qquad \text{where} \quad \sigma = \sum_{h=1}^{k} \frac{q_h}{p_h}.$$

Since each $m_{jj} \geq 1$, it's clear that κ_8 cannot be large, and consequently no stationary probability can be unduly sensitive to perturbations in P. It's apparent that similar remarks hold for more general mammillary structures as well.

4. Comparison of Condition Numbers. In this section, we compare the condition numbers κ_l. The norm-wise perturbation bounds in § 3 are of the following form:

$$\|\pi - \tilde{\pi}\|_p \leq \kappa_l \|E\|_q,$$

where $(p,q) = (\infty, \infty)$, or $(1, \infty)$, depending on l. (For κ_4, a strict inequality holds.) The purpose of this section is simply to compare the condition numbers appearing in those norm-wise bounds. Therefore we use the form

$$\frac{\|\pi - \tilde{\pi}\|_\infty}{\|E\|_\infty} \leq \kappa_l$$

for all condition numbers although some of the bounds are tighter than in this form. (See Remark 4.1.)

LEMMA 4.1.

$$(a) \quad \max_j \max_{i,k} |a^\#_{ij} - a^\#_{kj}| = \max_j (a^\#_{jj} - \min_i a^\#_{ij})$$

$$(b) \quad \frac{\max_j (a^\#_{jj} - \min_i a^\#_{ij})}{2} \leq \max_{i,j} |a^\#_{ij}| < \max_j (a^\#_{jj} - \min_i a^\#_{ij})$$

$$(c) \quad \max_{i,j} |a^\#_{ij}| \leq \|A^\#\|_\infty$$

$$(d) \quad \max_j (a^\#_{jj} - \min_i a^\#_{ij}) \leq \tau_1(A^\#)$$

$$(e) \quad \tau_1(A^\#) \leq n \, \frac{\max_j (a^\#_{jj} - \max_i a^\#_{ij})}{2}$$

$$(f) \quad \begin{aligned} &\tau_1(A^\#) \leq \|A^\#\|_\infty \\ &\tau_1(A^\#) \leq \|Z\|_\infty \\ &\tau_1(A^\#) \leq \frac{1}{1 - \tau_1(P)} \end{aligned}$$

$$(g) \quad \|A^\#\|_\infty - 1 \leq \|Z\|_\infty \leq \|A^\#\|_\infty + 1$$

Proof.

(a) Using a symmetric permutation, we may assume that a particular probability occurs in the last position of π. Partition A and π as follows:

$$A = \begin{pmatrix} A_{(n)} & c \\ d^T & a_{nn} \end{pmatrix}, \qquad \pi = \begin{pmatrix} \pi_{(n)} \\ \pi_n \end{pmatrix},$$

where $A_{(n)}$ is the principal matrix of A obtained by deleting the n-th row and column from A. Since *rank* $A = n - 1$, the relationship $a_{nn} = d^T A^{-1}_{(n)} c$ holds, so that

$$A^\# = \begin{pmatrix} (I - e\pi^T_{(n)}) A^{-1}_{(n)} (I - e\pi^T_{(n)}) & -\pi_n (I - e\pi^T_{(n)}) A^{-1}_{(n)} e \\ -\pi^T_{(n)} A^{-1}_{(n)} (I - e\pi^T_{(n)}) & \pi_n \pi^T_{(n)} A^{-1}_{(n)} e \end{pmatrix}.$$

Hence

$$(4.1) \qquad a^\#_{in} = \begin{cases} -\pi_n e_i^T A^{-1}_{(n)} e + \pi_n \pi^T_{(n)} A^{-1}_{(n)} e, & i \neq n, \\ \pi_n \pi^T_{(n)} A^{-1}_{(n)} e, & i = n, \end{cases}$$

where e_i is the i-th column of I. By M-matrix properties, it follows that $A_{(n)}^{-1} > 0$ so that $\pi_n e_i^T A_{(n)}^{-1} e$, $\pi_n \pi_{(n)}^T A_{(n)}^{-1} e > 0$. Thus

$$\max_i a_{in}^{\#} = a_{nn}^{\#},$$

and the result follows.

(b) For each j, $a_{jj}^{\#} > 0$, by (4.1). Since $\pi^T A^{\#} = \mathbf{0}^T$, it follows that there exists k_0 such that $a_{k_0 j}^{\#} < 0$. Furthermore, let i_0 be such that $\max_i |a_{ij}^{\#}| = |a_{i_0 j}^{\#}|$. Then

$$|a_{i_0 j}^{\#}| < \begin{cases} |a_{i_0 j}^{\#} - a_{k_0 j}^{\#}|, & \text{if } a_{i_0 j}^{\#} \geq 0, \\ |a_{i_0 j}^{\#} - a_{jj}^{\#}|, & \text{otherwise.} \end{cases}$$

Thus, for all j,

$$\max_i |a_{ij}^{\#}| < \max_{i,k} |a_{ij}^{\#} - a_{kj}^{\#}|$$

so that

$$\max_{i,j} |a_{ij}^{\#}| < \max_j \max_{i,k} |a_{ij}^{\#} - a_{kj}^{\#}|.$$

On the other hand,

$$\frac{\max_j \max_{i,k} |a_{ij}^{\#} - a_{kj}^{\#}|}{2} \leq \frac{\max_{i,j} |a_{ij}^{\#}| + \max_{k,j} |a_{kj}^{\#}|}{2} = \max_{i,j} |a_{ij}^{\#}|.$$

Now the assertion follows by (a).

(c) follows directly from the definitions of 1-, and ∞- norm, and their relationship,

(d) For a real number a, define $a^+ \equiv \max\{a, 0\}$. Then

$$\begin{aligned} \max_j \max_{i,k} |a_{ij}^{\#} - a_{kj}^{\#}| &= \max_{i,k} \max_j |a_{ij}^{\#} - a_{kj}^{\#}| \\ &\leq \max_{i,k} \sum_j (a_{ij}^{\#} - a_{kj}^{\#})^+ \\ &= \tau_1(A^{\#}), \qquad \text{by Seneta [19, p.139]}, \end{aligned}$$

and the assertion follows by (a).

(e) For any vector $x \in \mathcal{R}^n$,

$$\|x\|_1 \leq n \|x\|_\infty,$$

and the result follows, since

$$\tau_1(A^{\#}) = \frac{1}{2} \max_{i,k} \|A_{i*}^{\#} - A_{k*}^{\#}\|_1,$$

and

$$\frac{\max_j \max_{i,k} |a_{ij}^{\#} - a_{kj}^{\#}|}{2} = \frac{1}{2} \max_{i,k} \|A_{i*}^{\#} - A_{k*}^{\#}\|_\infty,$$

where B_{i*} denotes the i^{th} row of a matrix B.

(f) Since for any matrix B with equal row sums,

$$\tau_1(B) = \sup_{\substack{\|v\|_1=1 \\ v^T e=0}} \|v^T B\|_1,$$

and

$$\|y^T B\|_1 \leq \|y\|_1 \, \|B\|_\infty,$$

for any vector $y \in \mathcal{R}^n$, we have

$$\tau_1(B) \leq \|B\|_\infty.$$

The inequalities follow by this relationship together with the following facts, proven in Seneta [22]:

$$\tau_1(A^{\#}) = \tau_1(Z),$$

and if $\tau_1(P) < 1$, then

$$\tau_1(A^{\#}) \leq \frac{1}{1 - \tau_1(P)}.$$

(g) follows by applying the triangle inequality to

$$Z = A^{\#} + e\pi^T \quad \text{and} \quad A^{\#} = Z - e\pi^T. \quad \square$$

The following list summarizes the relationship between the condition numbers:
Relation between condition numbers.

$$\begin{aligned} &\kappa_8 = \kappa_3 \leq \kappa_4 < 2\kappa_3 \leq \kappa_6 \leq \kappa_l, \qquad \text{for } l = 1, 2, 5, \\ &\kappa_6 \leq n\kappa_3, \\ &\kappa_2 - 1 \leq \kappa_1 \leq \kappa_2 + 1. \end{aligned}$$

REMARK 4.1. *As remarked at the beginning of this section, some of the condition numbers provide tighter bounds than the form used in this section. To be more precise, for $l = 1, 2, 5, 6$,*

$$\frac{\|\pi - \widetilde{\pi}\|_\infty}{\|E\|_\infty} \leq \frac{\|\pi - \widetilde{\pi}\|_1}{\|E\|_\infty} \leq \kappa_l.$$

To give a 'fairer' comparison to these condition numbers, note that $(\pi - \widetilde{\pi})^T e = 0$ implies $\|\pi - \widetilde{\pi}\|_\infty \leq (1/2)\|\pi - \widetilde{\pi}\|_1$ so that

$$\frac{\|\pi - \widetilde{\pi}\|_\infty}{\|E\|_\infty} \leq \kappa_l',$$

where $\kappa_l' = (1/2)\kappa_l$, for $l = 1, 2, 5, 6$.[1] *The comparison of these 'new' condtion numbers κ_l' with κ_3 is as given above:*

$$\kappa_3 \leq \kappa_l',$$

for $l = 1, 2, 5, 6$.

[1]The authors thank the referee for bringing this point to our attention.

5. Concluding Remarks. We reviewed eight existing perturbation bounds, and the condition numbers are compared. The list at the end of the last section clarifies the relationships between seven condition numbers. Among these seven condition numbers, the smallest condition number is κ_3

$$\frac{\max_j (a^{\#}_{jj} - \min_i a^{\#}_{ij})}{2}$$

by Haviv and van Heyden, and Kirkland et al., or equivalently, κ_8

$$\frac{1}{2} \max_j \left[\frac{\max_{i \neq j} m_{ij}}{m_{jj}} \right]$$

by Cho and Meyer.

The only condition number not included in the comparison is

$$\kappa_7 = \frac{\min_j \|A^{-1}_{(j)}\|_\infty}{2}.$$

Is κ_3 a smaller condition number than κ_7? Since by (4.1),

$$a^{\#}_{jj} - \min_i a^{\#}_{ij} = \pi_j \|A^{-1}_{(j)}\|_\infty,$$

for each j, the question is whether

$$\max_j \pi_j \|A^{-1}_{(j)}\|_\infty \leq \min_j \|A^{-1}_{(j)}\|_\infty \tag{5.1}$$

holds. In other words, is π_j always small enough so that $\max_j \pi_j \|A^{-1}_{(j)}\|_\infty$ is never larger than $\min_j \|A^{-1}_{(j)}\|_\infty$? This question leads us to the relationship between a stationary probability π_j and the corresponding $A^{-1}_{(j)}$.

In a special case, the relationship is clear.

LEMMA 5.1. *If P is of rank 1, then*

$$\|A^{-1}_{(j)}\|_\infty = \frac{1}{\pi_j}, \qquad \textit{for all } j = 1, 2, \cdots, n.$$

The proof appears in Kirkland et al. 1998[9, Observation 3.3].

This relationship for an arbitrary irreducible Markov chain does not hold. In general, however, $\|A^{-1}_{(j)}\|_\infty$ tends to increase as π_j decreases, and vice versa.

Notice that κ_8 is equal to κ_3. However, viewing sensitivity in terms of mean first passage times can sometimes help practitioner decide whether or not to expect sensitivity in their Markov chain models merely by observing the structure of the chain, thus obviating the need for computing or estimating condition numbers.

REFERENCES

[1] G. CHO AND C. MEYER, *Markov chain sensitivity measured by mean first passage times*, Tech. Rep. Technical Report #112242-0199, Series 3.25.984, March, 1999, North Carolina State University, 1999. To appear in *Lin. Alg. Appl.*

[2] R. Dobrushin, *Central limit theorem for nonstationary Markov chains. I*, Theory of Probability and its Applications, 1 (1956), pp. 65–80.

[3] ———, *Central limit theorem for nonstationary Markov chains. II*, Theory of Probability and its Applications, 1 (1956), pp. 329–383.

[4] R. Funderlic and C. Meyer, *Sensitivity of the stationary distribution vector for an ergodic Markov chain*, Lin. Alg. Appl, 76 (1986), pp. 1–17.

[5] M. Haviv and L. Van der Heyden, *Perturbation bounds for the stationary probabilities of a finite Markov chain*, Adv. Appl. Prob., 16 (1984), pp. 804–818.

[6] I. Ipsen and C. Meyer, *Uniform stability of Markov chains*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1061–1074.

[7] D. Isaacson and R. Madsen, *Markov Chains: Theory and Application*, John Wiley & Sons, Inc., New York, 1976.

[8] J. Kemeny and J. Snell, *Finite Markov Chains*, Van Nostrand, New York, 1960.

[9] S. Kirkland, M. Neumann, and B. Shader, *Applications of Paz's inequality to perturbation bounds for Markov chains*, Lin. Alg. Appl, 268 (1998), pp. 183–196.

[10] A. Lešanovský, *Coefficients of ergodicity generated by non-symmetrical vector norms*, Czechoslovak Mathematical Journal, 40 (1990), pp. 284–294.

[11] C. Meyer, *The role of the group generalized inverse in the theory of finite Markov chains*, SIAM Rev., 17 (1975), pp. 443–464.

[12] ———, *The condition of a finite Markov chain and perturbation bounds for the limiting probabilities*, SIAM J.Alg.Disc.Meth., 1 (1980), pp. 273–283.

[13] ———, *The character of a finite Markov chain*, in Proceedings of the IMA Workshop on Linear Algebra, Markov Chains, and Queuing Models, New York, 1993, C.D. Meyer and R.J. Plemmons,eds.,Springer-Verlag, pp. 47–58.

[14] ———, *Sensitivity of the stationary distribution of a Markov chain*, SIAM J. Matrix Appl., 15 (1994), pp. 715–728.

[15] A. Rhodius, *On explicit forms for ergodic coefficients*, Lin. Alg. Appl., 194 (1993), pp. 71–83.

[16] U. Rothblum and C. Tan, *Upper bounds on the maximum modulus of subdominant eigenvlues of nonnegative matrices*, Lin. Alg. Appl., 66 (1985), pp. 45–86.

[17] P. Schweitzer, *Perturbation theory and finite Markov chains*, Journal of Applied Probability, 5 (1968), pp. 410–413.

[18] E. Seneta, *Coefficients of ergodicity: Structure and applications*, Adv. Appl. Pro., 11 (1979), pp. 576–590.

[19] ———, *Non-Negative Matrices and Markov Chains*, Springer, New York, 1981.

[20] ———, *Explicit forms for ergodicity coefficients and spectrum localization*, Lin. Alg. Appl, 60 (1984), pp. 187–197.

[21] ———, *Perturbation of the stationary distribution measured by ergodicity coefficients*, Adv.Appl.Pro., 20 (1988), pp. 228–230.

[22] ———, *Sensitivity analysis, ergodicity coefficients, and rank-one updates for finite Markov chains*, in W.J. Stewart(ed.) Numerical Solution of Markov Chains, Marcel Dekker, NY, 1991, pp. 121–129.

[23] ———, *Sensitivity of finite Markov chains under perturbation*, Statist. and Probab. Lett., 17 (1993), pp. 163–168.

[24] C. Sheppard and A. Householder, *The mathematical basis of the interpretation of tracer experiments in closed steady-systems*, J. Appl. Phys., 22 (1951), pp. 510–520.

[25] C. Tan, *A functional form for a coefficient of ergodicity*, Journal of Applied Probability, 19 (1982), pp. 858–863.

[26] ———, *Coefficients of ergodicity with respect to vector norms*, Journal of Applied Probability, 20 (1983), pp. 277–287.

[27] R. Whittaker, *Experiments with radisphosphorus tracer in aquarium microcosms*, Ecological Monographs, 31 (1961), pp. 157–188.