Column Sums and the Conditioning of the Stationary Distribution for a Stochastic Matrix

Steve Kirkland*
Hamilton Institute
National University of Ireland Maynooth
Ireland
stephen.kirkland@nuim.ie

January 19, 2010

Abstract

For an irreducible stochastic matrix T, we consider a certain condition number $\kappa(T)$, which measures the sensitivity of the stationary distribution vector to perturbations in T, and study the extent to which the column sum vector for T provides information on $\kappa(T)$. Specifically, if c^T is the column sum vector for some stochastic matrix of order n, we define the set $\mathcal{S}(c) = \{A | A \text{ is an } n \times n \text{ stochastic matrix with column sum vector } c^T\}$. We then characterise those vectors c^T such that $\kappa(T)$ is bounded as T ranges over the irreducible matrices in $\mathcal{S}(c)$; for those column sum vectors c^T for which κ is bounded, we give an upper bound on κ in terms of the entries in c^T, and characterise the equality case.

Keywords: Stochastic matrix, Stationary distribution, Condition number.

AMS Classification Numbers: 15A51, 60J10, 65F35.

1 Introduction

An $n \times n$ entrywise nonnegative matrix T is *stochastic* if it has the property that $T\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ denotes an all ones vector of the appropriate order. Evidently such a matrix T has 1 as an eigenvalue, and we find from Perron-Frobenius theory

*This material is based upon works supported by the Science Foundation Ireland under Grant No. SFI/07/SK/I1216b.

(see [11]) that if λ is any eigenvalue of T then $|\lambda| \leq 1$. If the eigenvalue 1 of T is algebraically simple, then T has a unique *stationary distribution* - i.e., an entrywise nonnegative row vector π^T such that $\pi^T T = \pi^T$ and $\pi^T \mathbf{1} = 1$. In the special case that the stochastic matrix T is *irreducible*, that is, the directed graph associated with T is strongly connected, then T necessarily has a unique stationary distribution, all of whose entries are positive.

Stochastic matrices and their corresponding stationary distributions are the centrepiece of the theory of finite time homogeneous discrete-time Markov chains. In particular, if T is the transition matrix for the Markov chain, and 1 is an algebraically simple eigenvalue of T, then the iterates of the Markov chain converge to the stationary distribution of T, regardless of the initial distribution for the chain. We refer the reader to [11] for background on that subject.

Given the interest in stationary distributions for stochastic matrices, it is not surprising that there is a body of work on the conditioning properties of stationary distributions. Specifically, suppose that T is an irreducible stochastic matrix with stationary distribution π^T, and $\tilde{T} = T + E$ is a perturbation of T that is also irreducible and stochastic, with stationary distribution $\tilde{\pi}^T$; we say that a function $f(T)$ is a *condition number* for the stationary distribution if, for some pair of suitable norms $|| \, ||_a, || \, ||_b$ we have $||\pi^T - \tilde{\pi}^T||_a \leq f(T)||E||_b$ for all admissible perturbing matrices E. There are a number of condition numbers available for the stationary distribution, and the paper [2] surveys several of these and makes comparisons between them.

In this paper we focus on a particular condition number. Let $|| \, ||_\infty$ denote the maximum absolute row sum norm. For any irreducible stochastic matrix T of order n with stationary vector π^T, we let

$$\kappa(T) = \frac{1}{2} \max\{\pi_i ||(I - T_i)^{-1}||_\infty | i = 1, \ldots, n\},$$

where for each $i = 1, \ldots, n, T_i$ denotes the principal submatrix of T formed by deleting row i and column i. We note that $\kappa(T)$ can also be expressed in terms of the group generalised inverse $(I-T)^{\#}$ corresponding to the matrix $I-T$, namely

$$\kappa(T) = \frac{1}{2} \max_{i,j=1,\ldots,n} ((I-T)^{\#}_{jj} - (I-T)^{\#}_{ij})$$

(see [2]). According to results in [3] and [9], for any perturbing matrix E such that $\tilde{T} = T + E$ is irreducible and stochastic with stationary distribution $\tilde{\pi}^T$, we have $\max_i |\pi_i - \tilde{\pi}_i| \leq \kappa(T)||E||_{\infty}$. Indeed, from the results in [2] and [4], it follows that of the eight condition numbers surveyed in [2], $\kappa(T)$ is the smallest. Further properties of $\kappa(T)$ are developed in [5] and [7]; in particular, both of those papers provide bounds on $\kappa(T)$ in terms of parameters of the directed graph associated with T.

In [2], the authors comment briefly on Markov chains possessing 'a dominant central state with strong connections to and from all other states', asserting that the stationary distribution associated with such a chain cannot be unduly sensitive to perturbations in the entries of the transition matrix. The present paper deals with a related problem, by considering the vector of column sums of the transition matrix. Specifically, suppose that $n \in \mathbb{N}$ and that we have numbers $c_1 \geq c_2 \geq \ldots \geq c_n \geq 0$ such that $\sum_{i=1}^n c_i = n$. We refer to such a vector c^T as an *admissible column sum vector of order* n. Let $\mathcal{S}(c) = \{A | A \text{ is an } \ n \times n \text{ stochastic matrix with } \mathbf{1}^T A = c^T\}$; it straightforward to determine that $\mathcal{S}(c) \neq \emptyset$.

We focus on two key problems in the sequel:

a) characterise the admissible column sum vectors c^T for which $\kappa(T)$ is bounded from above as T ranges over the irreducible matrices in $\mathcal{S}(c)$;

b) for those admissible column sum vectors such that κ is bounded as in a), find an upper bound on the corresponding value of κ.

In this paper, we solve both of those problems.

Throughout, we will use standard ideas and terminology from the theory of stochastic matrices, and from the theory of generalised inverses. We refer the reader to [11] for the former and to [1] for the latter.

2 Preliminary results

In this section we address problem a) described in Section 1. Our first result will be useful in the sequel.

Lemma 2.1. *Let c^T be an admissible column sum vector of order n. Then $\kappa(A)$ is bounded as A ranges over the irreducible matrices in $\mathcal{S}(c)$ if and only if, for each matrix $A \in \mathcal{S}(c)$, there is just a single essential class of indices for A.*

Proof. Appealing to Corollary 2.6 and Theorem 2.8 of [4], we find that for any irreducible $A \in \mathcal{S}(c)$,

$$\frac{1}{2}\min_j ||(I - A_j)^{-1}||_\infty \geq \kappa(A) \geq \frac{1}{2n}\min_j ||(I - A_j)^{-1}||_\infty.$$

Thus we find that $\kappa(A)$ is bounded as A ranges over the irreducible matrices in $\mathcal{S}(c)$ if and only if $\min_j ||(I - A_j)^{-1}||_\infty$ is bounded as A ranges over the same set.

If $\mathcal{S}(c)$ admits a matrix M for which there are two essential classes, then note that for each $0 < \epsilon < 1$, the matrix $M(\epsilon) \equiv (1-\epsilon)M + \frac{\epsilon}{n}\mathbf{1}c^T$ lies in $\mathcal{S}(c)$ and is irreducible. It is not difficult to show that $\min_j ||(I - M(\epsilon)_j)^{-1}||_\infty$ is unbounded from above as $\epsilon \to 0^+$.

Now suppose that $\min_j ||(I - A_j)^{-1}||_\infty$ is unbounded as A ranges over the irreducible matrices in $\mathcal{S}(c)$. Then there is a sequence of matrices $\{A(m)\}_{m=1}^\infty \in \mathcal{S}(c)$ such that for each $j = 1, \ldots, n$, $||(I - A(m)_j)^{-1}||_\infty \to \infty$ as $m \to \infty$. Since $\mathcal{S}(c)$ is compact, $A(m)$ has a convergent subsequence, say with limiting matrix $T \in \mathcal{S}(c)$. Further, for each $j = 1, \ldots, n, I - T_j$ fails to be invertible. Hence for the matrix T, we see that for any index j there is an essential class that does not contain j, so that T must contain at least two essential classes.

$\square$

We now recast the key condition of Lemma 2.1 in terms of column sums.

Theorem 2.1. *Let c^T be an admissible column sum vector of order $n \geq 2$. Each $A \in \mathcal{S}(c)$ has a single essential class if and only if $c_2 < 1$.*

Proof. We proceed by proving the following equivalent statement: $\mathcal{S}(c)$ admits a matrix A with two or more essential classes if and only if $c_2 \geq 1$. This statement is straightforward to verify if $n = 2$, so we assume henceforth that $n \geq 3$.

Suppose first that $\mathcal{S}(c)$ admits a matrix A with two or more essential classes. Then there is a permutation matrix P such that PAP^T has the form $\left[\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right]$, were A_1 is $k \times k$, A_2 is $(n-k) \times (n-k)$, and where without loss of generality, the permutation P maps e_j to e_1 for some $j \geq k+1$, so that the first row and column of A corresponds to the j-th row and column of PAP^T (here, as usual, e_i denotes the i-th standard unit basis vector). Considering the sum of the entries in A_1, we find that there are indices $i_1, \ldots, i_k \geq 2$ such that $k = \mathbf{1}^T A_1 \mathbf{1} = \sum_{j=1}^{k} c_{i_j}$. Hence $\max\{c_{i_1}, \ldots, c_{i_k}\} \geq 1$, whence $c_2 \geq 1$.

Conversely, suppose that $c_2 \geq 1$, and let $\tilde{c}^T = \left[\begin{array}{ccc} c_3 & \ldots & c_n \end{array}\right]$. The matrix B given by

$$B = \left[\begin{array}{cc|c} 1 & 0 & 0^T \\ 0 & 1 & 0^T \\ \hline \frac{c_1-1}{n-2}\mathbf{1} & \frac{c_2-1}{n-2}\mathbf{1} & \frac{1}{n-2}\mathbf{1}\tilde{c}^T \end{array}\right]$$

is evidently in $\mathcal{S}(c)$, and has two essential classes of indices. $\square$

The following is immediate.

Corollary 2.1.1. *Let c^T be an admissible column sum vector of order n. Then $\kappa(A)$ is bounded as A ranges over the irreducible matrices in $\mathcal{S}(c)$ if and only if $c_2 < 1$.*

From Corollary 2.1.1 we see that the condition that κ is bounded above over $\mathcal{S}(c)$ is equivalent to the condition that $c_2 < 1$. Observe that in that case, necessarily we have $c_1 > 1$ and $c_i < 1$ for $i = 2, \ldots, n$, and so we can think of state 1 as being a dominant central state. Consequently, Corollary 2.1.1 serves to reinforce the comment of Cho and Meyer ([2]) quoted in Section 1.

Here is one of the main results of this section.

Theorem 2.2. *Let c^T be an admissible column sum vector of order n such that $c_2 < 1$. Then*

$$\sup\{\kappa(A)|A \text{ is irreducible }, A \in \mathcal{S}(c)\} =$$
$$\max\{\frac{1}{2}||(I-T)^{-1}||_\infty | T \text{ is } (n-1)\times(n-1), 0 \le T, T\mathbf{1} \le \mathbf{1}, \mathbf{1}^T T \le \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array}\right]\}. \quad (1)$$

Proof. Suppose that A is an irreducible matrix in $\mathcal{S}$ with stationary vector π^T. By Corollary 2.6 of [4], we have $2\kappa(A) \le \min_{1\le i\le n} ||(I-A_i)^{-1}||_\infty$, so in particular, $\kappa(A) \le \frac{1}{2}||(I-A_1)^{-1}||_\infty$. Since $A_1 \ge 0, A_\mathbf{1} \le \mathbf{1}$ and $\mathbf{1}^T A_1 \le \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array}\right]$, we conclude that $\sup\{\kappa(A)|A \text{ is irreducible}, A \in \mathcal{S}(c)\} \le \max\{\frac{1}{2}||(I-T)^{-1}||_\infty | 0 \le T, T\mathbf{1} \le \mathbf{1}, \mathbf{1}^T T \le \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array}\right]\}$.

Next, consider a matrix T that attains the maximum on the right hand side of (1). We claim that, without loss of generality, we may assume that $\mathbf{1}^T T = \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array}\right]$. To see the claim, suppose that for some $j = 2, \dots, n$, we have $\mathbf{1}^T T e_{j-1} < c_j$. We will construct a matrix S of order $n-1$ such that $S\mathbf{1} \le \mathbf{1}, \mathbf{1}^T S = \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array}\right]$, and $S \ge T$, from which the claim will follow. Since $\mathbf{1}^T T \mathbf{1} < n-1$, there is some row of T, say the i-th such that $e_i^T T\mathbf{1} < 1$. We may then increase the (i,j) entry of T to yield a substochastic matrix $\tilde{T}$ whose column sum vector is bounded above by $\left[\begin{array}{ccc} c_2 & \dots & c_n \end{array}\right]$, such that either $\mathbf{1}^T\tilde{T}e_{j-1} = c_j$, or $e_i^T\tilde{T}\mathbf{1} = 1$. In the case that $\mathbf{1}^T\tilde{T}e_{j-1} < c_j$, we may repeat the argument on $\tilde{T}$ and increase an entry in its $(j-1)$-st column. As $\sum_{j=2}^{n-1} c_j < n-1$, each matrix so constructed has at least one row sum strictly less than 1; it follows that this process must eventually construct a substochastic matrix $\overline{T}$ such that $\mathbf{1}^T\overline{T} \le \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array}\right]$, $\mathbf{1}^T\overline{T}e_{j-1} = c_j$, and $\overline{T} \ge T$. Iterating the argument now yields a matrix S with the desired properties. Henceforth we assume that $\mathbf{1}^T T = \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array}\right]$.

Let $A = \left[\begin{array}{c|c} 1 & 0^T \\ \hline \mathbf{1} - T\mathbf{1} & T \end{array}\right]$, which has stationary vector e_1^T. It is straightforward to see that $(I-A)^\# = \left[\begin{array}{c|c} 0 & 0^T \\ \hline -(I-T)^{-1}\mathbf{1} & (I-T)^{-1} \end{array}\right]$. In particular, $\kappa(A)$ coincides with the right hand side of (1).

For each $t \in [0,1]$, let $B(t) = (1-t)A + \frac{t}{n}\mathbf{1}c^T$. Then for each $t \in (0,1]$, $B(t)$ is an irreducible matrix in $\mathcal{S}(c)$. Further, from the approach taken in [10], it follows that $\kappa(B(t))$ is a continuous function of t on $[0,1]$. It now follows that as $t \to 0^+, \kappa(B(t)) \to \kappa(A)$, yielding the desired conclusion. $\square$

Let c^T be an admissible column sum vector of order n, and define $\overline{\kappa}(c)$ by

$$\overline{\kappa}(c) = \max\{\frac{1}{2}||(I-T)^{-1}||_\infty | T \text{ is } (n-1)\times(n-1), 0 \leq T, T\mathbf{1} \leq \mathbf{1}, \mathbf{1}^T T \leq \begin{bmatrix} c_2 & \ldots & c_n \end{bmatrix}\}.$$

Evidently for any irreducible stochastic matrix $A \in \mathcal{S}(c)$, and any perturbing matrix $E \neq 0$ such that $A+E$ is also irreducible and stochastic, we have the inequality

$$\frac{||\pi^T(A) - \pi^T(A+E)||_\infty}{||E||_\infty} \leq \overline{\kappa}(c), \tag{2}$$

where $\pi^T(A)$ and $\pi^T(A+E)$ denote the stationary distributions for A and $A+E$, respectively.

Remark 2.1. In this remark we show that the inequality (2) cannot be improved. Adopting the notation of the proof of Theorem 2.2, we let T be a matrix yielding the maximum on the right side of (1), let $A = \left[\begin{array}{c|c} 1 & 0^T \\ \hline \mathbf{1} - T\mathbf{1} & T \end{array}\right]$, and for each $t \in (0,1)$, let $B(t) = (1-t)A + \frac{t}{n}\mathbf{1}c^T$. For convenience, we suppose that $n \geq 3$.

For each such t, let $E_t = \frac{tc_2}{n}(e_1 - e_2)(e_1 - e_2)^T$, and note that since $n \geq 3$, $B(t) + E_t$ is an irreducible stochastic matrix in $\mathcal{S}(c)$. Observe that $||E_t||_\infty = \frac{2tc_2}{n}$. For each $t \in (0,1)$, denote the stationary distributions for $B(t)$ and $B(t)+E_t$ by $\pi^T(B(t))$ and $\pi^T(B(t)+E_t)$, respectively. For any such t, we have $\pi^T(B(t)+E_t)E_t(I-B(t))^\# = \pi^T(B(t)+E_t) - \pi^T(B(t))$. Hence $\frac{tc_2}{n}\pi_1(B(t)+E_t) - \pi_2(B(t)+E_t)((I-B(t))^\#_{11} - (I-B(t))^\#_{21}) = \pi_1(B(t)+E_t) - \pi_1(B(t))$. Consequently, for each $t \in (0,1)$, we have

$$\frac{|\pi_1(B(t)+E_t) - \pi_1(B(t))|}{||E_t||_\infty} = |\pi_1(B(t)+E_t) - \pi_2(B(t)+E_t)|\frac{1}{2}((I-B(t))^\#_{11} - (I-B(t))^\#_{21}).$$

Letting $t \to 0^+$, we see that the right side above converges to $\frac{1}{2}((I-A)^\#_{11} - (I-A)^\#_{21}) = \overline{\kappa}(c)$. Thus we see that for any $\delta > 0$, there is an irreducible matrix M in $\mathcal{S}(c)$ and a perturbing matrix E such that $M+E$ is also irreducible, and with the property that $\frac{||\pi^T(M) - \pi^T(M+E)||_\infty}{||E||_\infty} > \overline{\kappa}(c) - \delta$.

Our last result of this section presents an eigenvalue bound for matrices in $\mathcal{S}(c)$ when $c_2 < 1$.

Proposition 2.1. *Suppose that c^T is an admissible column sum vector of order n, and that $c_2 < 1$. Let $A \in \mathcal{S}(c)$, and suppose that $\lambda \neq 1$ is an eigenvalue of A. Then $|1-\lambda| \geq \frac{1}{n\overline{\kappa}(c)}$.*

Proof. Let M be any $n \times n$ matrix with constant row sums, say $M\mathbf{1} = \rho\mathbf{1}$. Define $\tau(M) = \frac{1}{2}\max_{i,j}||(e_i^T - e_j^T)M||_1$; from Theorem 2.10 of [11] we find that if $z \neq \rho$ is an eigenvalue of M, then $|z| \leq \tau(M)$.

Now consider the matrix $M = (I-A)^{\#}$, and note that $M\mathbf{1} = 0$. Further, we see that if $\lambda \neq 1$ is an eigenvalue of A, then $\frac{1}{|1-\lambda|}$ is an eigenvalue of M. Consequently $\frac{1}{|1-\lambda|} \leq \tau(M)$. Next, observe that $\tau(M) \leq \frac{n}{2}\max_{i,j}((I-A)^{\#}_{jj} - (I-A)^{\#}_{ij}) \leq n\overline{\kappa}(c)$. The conclusion now follows. □

3 Bounds on $\overline{\kappa}(c)$

In this section we produce upper bounds on $\overline{\kappa}(c)$. We begin with a useful lemma.

Lemma 3.1. *Suppose that $1 > c_2 \geq c_3 \geq \ldots \geq c_n$. Let T be a substochastic matrix of order $n-1$ whose column sum vector is $\tilde{c}^T = \begin{bmatrix} c_2 & \ldots & c_n \end{bmatrix}$. Fix an index $j = 1, \ldots, n-1$, and let $\hat{T}$ be the principal submatrix of T formed by deleting its j-th row and column. Let $\hat{c}$ be formed from $\tilde{c}$ by deleting the j-th entry (which is c_{j+1}), and let x denote the column vector formed from Te_j by deleting the j-th entry. Then $e_j(I-T)^{-1}\mathbf{1} = \frac{n-1-(\mathbf{1}^T-\hat{c}^T)(I-\hat{T})^{-1}\mathbf{1}}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)(I-\hat{T})^{-1}x}$.*

Proof. Evidently T is permutationally similar to the matrix

$$M = \left[\begin{array}{c|c} \hat{T} & x \\ \hline \hat{c}^T - \mathbf{1}^T\hat{T} & c_{j+1} - \mathbf{1}^T x \end{array}\right],$$

where the last row and column of M correspond, respectively to the j-th row and column of T. In particular we see that $e_j^T(I-T)^{-1}\mathbf{1} = e_{n-1}^T(I-M)^{-1}\mathbf{1}$. We observe in passing that $\hat{T}\mathbf{1} + x \leq \mathbf{1}, \mathbf{1}^T\hat{T} \leq \hat{c}^T, \mathbf{1}^T x \leq c_{j+1}$, and $\sum_{i=2}^n c_i - 1 \leq \mathbf{1}^T\hat{T}\mathbf{1} + \mathbf{1}^T x$.

Let $\Delta = 1 - c_{j+1} + \mathbf{1}^T x - (\hat{c}^T - \mathbf{1}^T \hat{T})(I - \hat{T})^{-1} x$. Then

$$(I - M)^{-1} = \left[\begin{array}{c|c} (I-\hat{T})^{-1} + \frac{1}{\Delta}(I-\hat{T})^{-1} x (\hat{c}^T - \mathbf{1}^T \hat{T})(I-\hat{T})^{-1} & \frac{1}{\Delta}(I-\hat{T})^{-1} x \\ \hline \frac{1}{\Delta}(\hat{c}^T - \mathbf{1}^T \hat{T})(I-\hat{T})^{-1} & \frac{1}{\Delta} \end{array} \right].$$

By using the relation $(I-\hat{T})^{-1} - \hat{T}(I-\hat{T})^{-1} = I$, we find that Δ can be rewritten as $1 - c_{j+1} + (\mathbf{1}^T - \hat{c}^T)(I-\hat{T})^{-1} x$. Similarly, we have $(\hat{c}^T - \mathbf{1}^T \hat{T})(I-\hat{T})^{-1}\mathbf{1} = n - 2 - (\mathbf{1}^T - \hat{c}^T)(I-\hat{T})^{-1}\mathbf{1}$. It now follows that $e_j^T (I-T)^{-1}\mathbf{1} = e_{n-1}^T (I-M)^{-1}\mathbf{1} = \frac{n-1-(\mathbf{1}^T - \hat{c}^T)(I-\hat{T})^{-1}\mathbf{1}}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)(I-\hat{T})^{-1}x}$, as desired. □

Next, we present a general bound on $\overline{\kappa}(c)$.

Theorem 3.1. *Let c^T be an admissible column sum vector of order n such that $c_2 < 1$. Then $\overline{\kappa}(c) \le \frac{1+\sum_{i=3}^n c_i}{2(1-c_2)}$. Equality holds if and only if $\sum_{i=2}^n c_i \le 1$.*

Proof. From the proof of Theorem 2.2, we see that the value for $\overline{\kappa}(c)$ corresponds to a substochastic matrix T of order $n-1$ with column sum vector $\left[\begin{array}{ccc} c_2 & \dots & c_n \end{array} \right]$. Consider such a T and suppose that for some $j = 1, \dots, n-1$, we have $||(I-T)^{-1}||_\infty = e_j (I-T)^{-1}\mathbf{1}$. Adopting the notation of Lemma 3.1, we find that $||(I-T)^{-1}||_\infty = \frac{n-1-(\mathbf{1}^T-\hat{c}^T)(I-\hat{T})^{-1}\mathbf{1}}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)(I-\hat{T})^{-1}x} \le \frac{n-1-(\mathbf{1}^T-\hat{c}^T)(I-\hat{T})^{-1}\mathbf{1}}{1-c_{j+1}} \le \frac{n-1-(\mathbf{1}^T-\hat{c}^T)\mathbf{1}}{1-c_{j+1}} = \frac{1+\sum_{i=2}^n c_i - c_{j+1}}{1-c_{j+1}}$, the last inequality following from the fact that $I \le (I-\hat{T})^{-1}$. It is now readily established that $\frac{1+\sum_{i=2}^n c_i - c_{j+1}}{1-c_{j+1}} \le \frac{1+\sum_{i=3}^n c_i}{1-c_2}$, which yields the desired upper bound on $||(I-T)^{-1}||_\infty$, and hence on $\overline{\kappa}(c)$.

Examining the argument above, we find that if $\overline{\kappa}(c) = \frac{1+\sum_{i=3}^n c_i}{1-c_2}$, then necessarily $x = 0$, $\hat{T} = 0$, and j can be taken to be 1, so that T can be taken as $T = e_1 \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array} \right]$. Since that matrix is substochastic, it must be the case that $\sum_{i=2}^n c_i \le 1$. Conversely, if $\sum_{i=2}^n c_i \le 1$, then a straightforward computation shows that $||(I - e_1 \left[\begin{array}{ccc} c_2 & \dots & c_n \end{array} \right])^{-1}||_\infty = \frac{1+\sum_{i=3}^n c_i}{1-c_2}$. □

Example 3.1. In this example, we illustrate the conclusion of Theorem 3.1, as well as the limiting process used in the proof of Theorem 2.2 to establish (1). Consider the admissible column sum vector $c^T = \left[\begin{array}{ccccc} 4 & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{array} \right]$; since $\sum_{i=2}^5 c_i = 1$, it follows from Theorem 3.1 that $\overline{\kappa}(c) = \frac{11}{6}$.

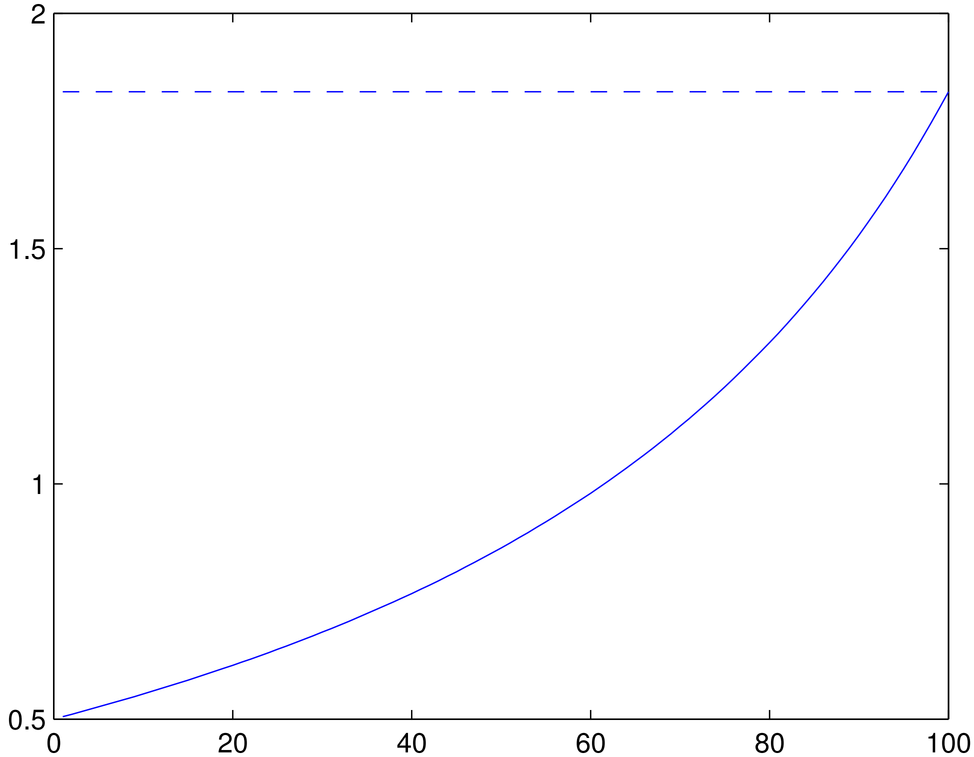

Figure 1: $\kappa(B(t))$ for $t \in [0,1]$

Let A be the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and for each $t \in [0,1]$, note that the convex combination $B(t) = (1-t)A + \frac{t}{5}\mathbf{1}c^T$ is in $\mathcal{S}(c)$. Figure 1 plots the computed values of $\kappa(B(\frac{i}{100}))$ for $i = 0, \ldots, 100$. Observe that as $t = \frac{i}{100}$ increases from 0 up to 1, the values of $\kappa(B(t))$ increase from $\frac{1}{2}$ when $t = 0$, up to the maximum possible value of $\frac{11}{6}$, which is denoted by the dashed horizontal line in Figure 1, when $t = 1$.

Before proceeding to refine the bound of Theorem 3.1, we require the following technical result.

Lemma 3.2. *Suppose that u and v^T are nonnegative vectors of orders p and q, respectively. There is a $p \times q$ nonnegative matrix S such that $S\mathbf{1} = u$ and $\mathbf{1}^T S \leq v^T$ if and only if $\mathbf{1}^T u \leq \mathbf{1}^T v$.*

Proof. Certainly if such a matrix S exists, then $\mathbf{1}^T u = \mathbf{1}^T S \mathbf{1} \leq \mathbf{1}^T v$.

To prove the converse, note that we need only consider the case that $\mathbf{1}^T v > 0$ (otherwise we select $S = 0$). We proceed by induction on $p + q$, and begin by observing that the cases $p = 1$ and $q = 1$ are easily established, since we can take $S = \frac{\mathbf{1}^T u}{\mathbf{1}^T v} v^T$ and $S = u$, respectively, in those cases.

Suppose that the result holds if $p+q = m$, and that we have vectors u and v^T, of orders p_0, q_0 respectively such that $p_0 + q_0 = m + 1$ and $\mathbf{1}^T u \leq \mathbf{1}^T v$. If $u_1 \leq v_1$, then since $u_2 + \ldots + u_{p_0} \leq (v_1 - u_1) + v_2 + \ldots + v_{q_0}$, we find from the induction hypothesis that there is a $(p_0 - 1) \times q_0$ nonnegative matrix $\overline{S}$ such that $\overline{S}\mathbf{1} = \begin{bmatrix} u_2 \\ \vdots \\ u_{p_0} \end{bmatrix}$ and $\mathbf{1}^T \overline{S} \leq \begin{bmatrix} (v_1 - u_1) & v_2 & \ldots & v_{q_0} \end{bmatrix}$. The matrix $S = \left[\begin{array}{c} u_1 e_1^T \\ \hline \overline{S} \end{array} \right]$ now has the desired properties. Similarly, if $v_1 \leq u_1$, then since $(u_1 - v_1) + u_2 + \ldots u_{p_0} \leq v_2 + \ldots + v_{q_0}$, we find from the induction hypothesis that there is a $p_0 \times (q_0 - 1)$ nonnegative matrix $\tilde{S}$ such that $\tilde{S}\mathbf{1} = \begin{bmatrix} u_1 - v_1 \\ u_2 \\ \vdots \\ u_{p_0} \end{bmatrix}$ and $\mathbf{1}^T \overline{S} \leq \begin{bmatrix} v_2 & \ldots & v_{q_0} \end{bmatrix}$. In this case, the matrix $S = \left[\begin{array}{c|c} v_1 e_1 & \tilde{S} \end{array} \right]$ has the desired properties. $\square$

Adapting the technique of Theorem 3.1, we establish our final result, which is one of the main results in this paper.

Theorem 3.2. *Let c^T be an admissible column sum vector of order $n \geq 3$ such that $c_2 < 1$. Suppose also that for some $k \in \mathbb{N}$, we have $k + 1 \geq \sum_{i=2}^n c_i > k$. Then*

$$\overline{\kappa}(c) \leq \frac{1}{2} \left(1 + \frac{1 + \sum_{i=3}^{k+1} c_i + c_{k+2} (\sum_{i=2}^n c_i - k)}{1 - c_2} \right). \tag{3}$$

Equality holds if and only if $\sum_{i=2}^{k+2} c_i \leq 1$.

Proof. We begin by remarking that for the case that $k = 1$, we interpret the quantity $\sum_{i=3}^{k+1} c_i$ as 0 in (3).

Referring to the proof of Theorem 2.2, we find that a matrix T yielding the maximum value in the definition of $\overline{\kappa}(c)$ can be taken to be substochastic, of order $n-1$, with column sum vector equal to $\begin{bmatrix} c_2 & \dots & c_n \end{bmatrix}$. Let T be such a matrix, and suppose that for some $j = 1, \dots, n-1$, we have $||(I-T)^{-1}||_\infty = e_j(I-T)^{-1}\mathbf{1}$. Adopting the notation of Lemma 3.1, we have $||(I-T)^{-1}||_\infty = \frac{n-1-(\mathbf{1}^T-\hat{c}^T)(I-\hat{T})^{-1}\mathbf{1}}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)(I-\hat{T})^{-1}x} \le \frac{n-1-(\mathbf{1}^T-\hat{c}^T)\mathbf{1}-(\mathbf{1}^T-\hat{c}^T)\hat{T}\mathbf{1}}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)x} = \frac{1-c_{j+1}+\sum_{i=2}^n c_i-(\mathbf{1}^T-\hat{c}^T)\hat{T}\mathbf{1}}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)x} = 1 + \frac{\sum_{i=2}^n c_i-(\mathbf{1}^T-\hat{c}^T)(\hat{T}\mathbf{1}+x)}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)x}$.

Recalling that $\hat{T}\mathbf{1} + x \le \mathbf{1}$ and that $\sum_{i=2}^n c_i - 1 \le \mathbf{1}^T\hat{T}\mathbf{1} + \mathbf{1}^T x$, we find that $||(I-T)^{-1}||_\infty \le 1 + \frac{\sum_{i=2}^n c_i-\min_u(\mathbf{1}^T-\hat{c}^T)u}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)x}$, where in the numerator, the minimum is taken over all vectors u of order $n-1$ such that $0 \le u \le \mathbf{1}$ and $\mathbf{1}^T u = \sum_{i=2}^n c_i - 1$.

Next, we select indices $i_1, \dots, i_k$ such that the entries $c_{i_1}, \dots, c_{i_k}$ are, respectively, the k largest entries in $\hat{c}^T$, listed in descending order. Evidently several cases arise:
i) if $j \ge k+1$, then $\begin{bmatrix} c_{i_1} & \dots & c_{i_k} \end{bmatrix} = \begin{bmatrix} c_2 & \dots & c_{k+1} \end{bmatrix}$;
ii) if $j = k$, then $\begin{bmatrix} c_{i_1} & \dots & c_{i_k} \end{bmatrix} = \begin{bmatrix} c_2 & \dots & c_k & c_{k+2} \end{bmatrix}$;
iii) if $2 \le j \le k-1$, then $\begin{bmatrix} c_{i_1} & \dots & c_{i_k} \end{bmatrix} = \begin{bmatrix} c_2 & \dots & c_j & c_{j+2} & \dots & c_{k+2} \end{bmatrix}$;
iv) if $j = 1$, then $\begin{bmatrix} c_{i_1} & \dots & c_{i_k} \end{bmatrix} = \begin{bmatrix} c_3 & \dots & c_{k+2} \end{bmatrix}$.
Using this notation, it is straightforward to determine that $\min_u(\mathbf{1}^T - \hat{c}^T)u = \sum_{l=1}^{k-1}(1-c_{i_l}) + (1-c_{i_k})(\sum_{i=2}^n c_i - k)$.

Thus we find that $||(I-T)^{-1}||_\infty \le 1+\frac{\sum_{i=2}^n c_i-\sum_{l=1}^{k-1}(1-c_{i_l})-(1-c_{i_k})(\sum_{i=2}^n c_i-k)}{1-c_{j+1}+(\mathbf{1}^T-\hat{c}^T)x}$. Further, since $\sum_{i=2}^n c_i > k \ge \sum_{l=1}^{k-1}(1-c_{i_l}) + (1-c_{i_k})(\sum_{i=2}^n c_i - k)$, it follows that our upper bound above on $||(I-T)^{-1}||_\infty$ is decreasing in each entry of x. Consequently, we have $||(I-T)^{-1}||_\infty \le 1 + \frac{\sum_{i=2}^n c_i-\sum_{l=1}^{k-1}(1-c_{i_l})-(1-c_{i_k})(\sum_{i=2}^n c_i-k)}{1-c_{j+1}} \equiv d_{j+1}$.

Next, we claim that for each $j = 1, \dots, n-1$,

$$d_{j+1} \le d_2 = 1 + \frac{\sum_{i=2}^n c_i - \sum_{l=3}^{k+1}(1-c_l) - (1-c_{k+2})(\sum_{i=2}^n c_i - k)}{1-c_2}.$$

To see the claim, we first note that if $j \ge k+1$, then

$$d_{j+1} = 1 + \frac{\sum_{i=2}^n c_i - \sum_{l=2}^{k}(1-c_l) - (1-c_{k+1})(\sum_{i=2}^n c_i - k)}{1-c_{j+1}} \le d_{k+2}.$$

An uninteresting computation shows that the inequality $d_{k+2} \leq d_2$ is equivalent to

$$(c_2 - c_{k+2}) \sum_{i=2}^{n} c_i + (1-c_2)^2 - (c_2 - c_{k+2}) \sum_{l=3}^{k} (1-c_l) - (1-c_{k+1})(1-c_{k+2}) + \\ (\sum_{i=2}^{n} c_i - k)((1-c_2)(1-c_{k+1}) - (1-c_{k+2})^2) \geq 0. \quad (4)$$

Note that (4) can be rewritten as

$$(\sum_{i=2}^{n} c_i - k)(c_{k+2} - c_{k+2}^2 - c_{k+1} + c_2 c_{k+1}) + (c_2 - c_{k+2})(\sum_{l=2}^{k} c_l + c_{k+2}) + \\ (1-c_{k+2})(c_{k+1} - c_{k+2}) \geq 0. \quad (5)$$

If $c_{k+2} - c_{k+2}^2 - c_{k+1} + c_2 c_{k+1} \geq 0$, then certainly (5) holds. On the other hand, if $c_{k+2} - c_{k+2}^2 - c_{k+1} + c_2 c_{k+1} < 0$, then since $\sum_{i=2}^{n} c_i - k \leq 1$, it follows that the left side of (5) is bounded below by $c_{k+2} - c_{k+2}^2 - c_{k+1} + c_2 c_{k+1} + (1-c_{k+2})(c_{k+1} - c_{k+2}) = c_{k+1}(c_2 - c_{k+1}) \geq 0$. In either case, we find that (4) holds, so that $d_{j+1} \leq d_{k+2} \leq d_2$ for each $j \geq k+1$.

Next, we note that another computation reveals that the inequality $d_{k+1} \leq d_2$ is equivalent to $(c_2 - c_{k+1})(c_{k+2}(\sum_{i=2}^{n} c_i - k) + \sum_{l=2}^{k+1} c_l) \geq 0$; as the latter clearly holds, we find that $d_{k+1} \leq d_2$. Similarly, for $j = 3, \ldots, k$, the inequality $d_j \leq d_2$ can be shown to be equivalent to the inequality $(c_2 - c_j)(c_{k+2}(\sum_{i=2}^{n} c_i - k) + \sum_{l=2}^{k+1} c_l) \geq 0$, which again clearly holds. Hence, $d_{j+1} \leq d_2$ for each $j = 2, \ldots, n-1$, establishing the claim.

From the considerations above, we have $||(I-T)^{-1}||_\infty \leq 1 + \frac{1 + \sum_{i=3}^{k+1} c_i + c_{k+2}(\sum_{i=2}^{n} c_i - k)}{1-c_2} = d_2$, which yields (3).

Next, we consider the case that equality holds in (3). Examining the argument above, we see that equality holds only if

(i) j can be taken to be 1, (ii) $x = 0$, (iii) the vector $\hat{T}\mathbf{1}$ has 1s in the positions $1, \ldots, k-1$, the entry $\sum_{i=2}^{n} c_i - k$ in position k, and 0s elsewhere, and (iv) $\hat{T}^2 = 0$.

Setting $y^T = \begin{bmatrix} c_3 & \ldots & c_n \end{bmatrix}$, we thus find that a matrix T that yields equality in (3) has the form

$$T = \left[\begin{array}{c|c} c_2 & y^T - \mathbf{1}^T \hat{T} \\ \hline 0 & \hat{T} \end{array}\right],$$

where $\hat{T}$ satisfies the conditions above. Applying the row sum condition (iii) on $\hat{T}$, we find that the first k rows of $\hat{T}$ are nonzero, while the remaining rows of $\hat{T}$ must be all zero. Next, applying the condition (iv) that $\hat{T}^2 = 0$, it follows that $\hat{T}$ is given by $\hat{T} = \left[\begin{array}{c|c} 0 & S \\ \hline 0 & 0 \end{array}\right]$, where S is a $k \times (n-2-k)$ matrix such that $S\mathbf{1} = \left[\begin{array}{c} 1 \\ \vdots \\ 1 \\ \sum_{i=2}^{n} c_i - k \end{array}\right]$ and $\mathbf{1}^T S \leq \left[\begin{array}{ccc} c_{k+3} & \dots & c_n \end{array}\right]$. By Lemma 3.2, such a matrix S exists only if $k - 1 + \sum_{i=2}^{n} c_i - k \leq \sum_{l=k+3}^{n} c_l$ - i.e. - only if $\sum_{i=2}^{k+2} c_i \leq 1$.

Conversely, if $\sum_{i=2}^{k+2} c_i \leq 1$, then using Lemma 3.2 we can find a $k \times (n-2-k)$ matrix S with $S\mathbf{1} = \left[\begin{array}{c} 1 \\ \vdots \\ 1 \\ \sum_{i=2}^{n} c_i - k \end{array}\right]$ and $\mathbf{1}^T S \leq \left[\begin{array}{ccc} c_{k+3} & \dots & c_n \end{array}\right]$. Then, setting $z^T = \left[\begin{array}{ccc} c_{k+3} & \dots & c_n \end{array}\right]$ we construct the matrix T given by

$$T = \left[\begin{array}{c|ccc|c} c_2 & c_3 & \dots & c_{k+2} & z^T - \mathbf{1}^T S \\ \hline 0 & & 0 & & S \\ \hline 0 & & 0 & & 0 \end{array}\right].$$

It is now straightforward to determine that $||(I-T)^{-1}||_\infty = d_2$, and that T has the desired row and column sum properties. $\square$

Example 3.2. In this example, we illustrate the conclusion of Theorem 3.2 in a manner similar to that in Example 3.1. Consider the admissible column sum vector $c^T = \left[\begin{array}{ccccccccccc} \frac{53}{6} & \frac{2}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{array}\right]$. We have $\sum_{i=2}^{11} c_i = \frac{13}{6}$, so that in the notation of Theorem 3.2, we have $k = 2$. We now find from Theorem 3.2 that $\overline{\kappa}(c) = \frac{55}{24}$.

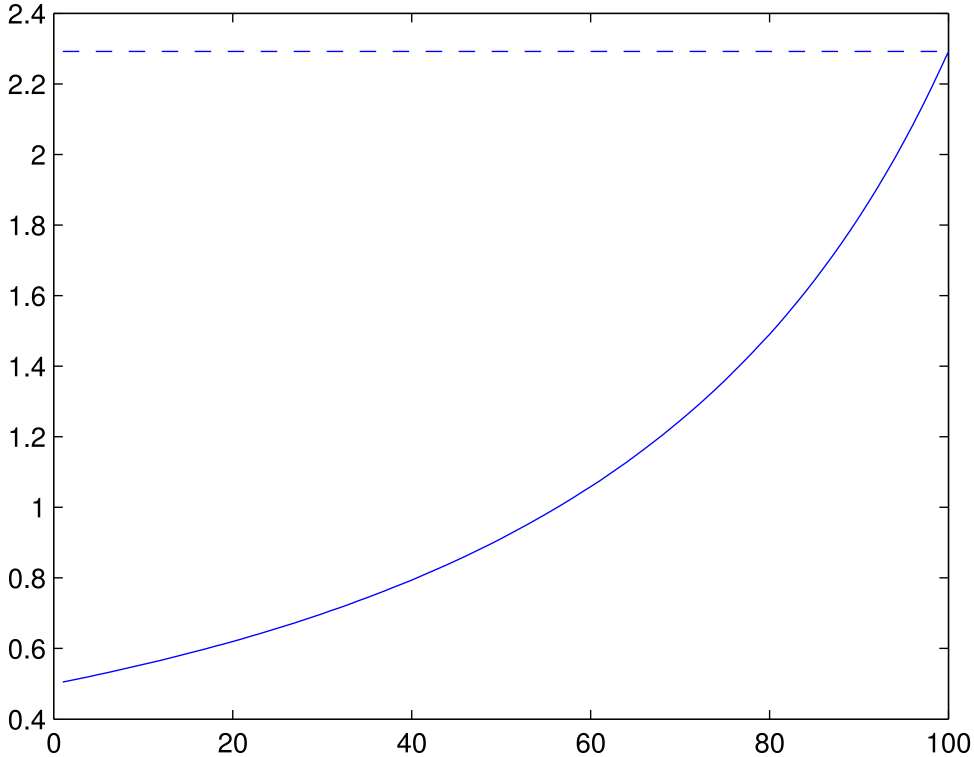

Figure 2: $\kappa(B(t))$ for $t \in [0,1]$

Let A be the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\ \frac{5}{6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{6} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

for each $t \in [0,1]$, note that the convex combination $B(t) = (1-t)A + \frac{t}{11}\mathbf{1}c^T$ is in $\mathcal{S}(c)$. Figure 2 plots the computed values of $\kappa(B(\frac{i}{100}))$ for $i = 0, \dots, 100$. We sse that as $t = \frac{i}{100}$ increases from 0 up to 1, $\kappa(B(t))$ increases from $\frac{1}{2}$ up to the maximum possible value of $\frac{55}{24}$, which is denoted by the dashed horizontal line in Figure 2.

4 Conclusion

In this paper, we have considered the condition number $\kappa(T)$, which measures the sensitivity of the stationary distribution of the irreducible stochastic matrix T. We have identified the admissible column sum vectors c^T for which $\kappa(T)$ is bounded as T ranges over the set $\mathcal{S}(c)$ of irreducible stochastic matrices with column sum vector c^T. For those admissible column vectors c^T for which $\kappa(T)$ is bounded as T ranges over the irreducible members of $\mathcal{S}(c)$, we provide sharp upper bounds on κ in terms of the entries of c^T.

The results in [6], [7] and [8] examine the conditioning of the stationary distribution of a stochastic matrix T by considering the structure of the directed graph of T. Specifically, those papers focus on how certain combinatorial parameters (i.e. qualitative information) associated with T can be used to construct bounds on $\kappa(T)$. The present paper provides some companion results to those works, by using readily available quantitative information – the column sums of T – in order to construct bounds on $\kappa(T)$.

References

[1] S. Campbell and C. Meyer, *Generalized Inverses of Linear Transformations*, Dover, New York, 1991.

[2] G. Cho and C. Meyer, Comparison of perturbation bounds for the stationary distribution of a Markov chain, *Linear Algebra and its Applications* 335: 137-150 (2001).

[3] M. Haviv and L. Van der Heyden, Perturbation bounds for the stationary probabilities of a finite Markov chain, *Advances in Applied Probability* 16: 804-818 (1984).

[4] S. Kirkland, On a question concerning condition numbers for Markov chains, *SIAM Journal on Matrix Analysis and Applications* 23: 1109-1119 (2002).

[5] S. Kirkland, Conditioning properties of the stationary distribution for a Markov chain, *Electronic Journal of Linear Algebra* 10: 1-15 (2003).

[6] S. Kirkland, A combinatorial approach to the conditioning of a single entry in the stationary distribution for a Markov chain, *Electronic Journal of Linear Algebra* 11: 168-179 (2004).

[7] S. Kirkland, Digraph-based conditioning for Markov chains, *Linear Algebra and its Applications* 385: 81-93 (2004).

[8] S. Kirkland, Conditioning of the entries in the stationary vector of a Google-type matrix, *Linear Algebra and its Applications* 418: 665-681 (2006).

[9] S. Kirkland, M. Neumann and B. Shader, Applications of Paz's inequality to perturbation bounds for Markov chains, *Linear Algebra and its Applications* 268: 183-196 (1998).

[10] C. Meyer, The condition of a finite Markov chain and perturbation bounds for the limiting probabilities, *SIAM Journal on Discrete and Algebraic Methods* 1: 273-283 (1980).

[11] E. Seneta, *Non-negative Matrices and Markov Chains*, Springer-Verlag, New York, 1981.