Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

ELA

A COMBINATORIAL APPROACH TO THE CONDITIONING OF A SINGLE ENTRY IN THE STATIONARY DISTRIBUTION FOR A MARKOV CHAIN*

S. KIRKLAND†

Abstract. For an irreducible stochastic matrix T of order n, a certain condition number $\kappa_j(T)$ that measures the sensitivity of the $j-$th entry of the corresponding stationary distribution under perturbation of T is considered. A lower bound on κ_j is produced in terms of the directed graph of T, and the case of equality is characterized in that lower bound. Also all of the directed graphs D are characterized such that $\kappa_j(T)$ is bounded from above as T ranges over the set of irreducible stochastic matrices having directed graph D. For those D for which κ_j is bounded, a tight upper bound is given on κ_j in terms of information contained in D.

Key words. Stochastic matrix, Markov chain, Stationary vector, Condition number, Directed graph.

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

1. Introduction.

Suppose that we have an $n \times n$ irreducible stochastic matrix T, which we can think of as the transition matrix of a Markov chain. It is well known that T possesses a unique *stationary distribution*, that is, an entrywise positive vector π^T such that $\pi^T T = \pi^T$ and $\pi^T \mathbf{1} = 1$, where $\mathbf{1}$ denotes the all ones vector of the appropriate order. In the case that T is primitive (i.e. some power of T has all positive entries) a standard result in the area asserts that the iterates of the Markov chain converge to π^T, independently of the initial distribution for the chain (see [9]). Thus the stationary distribution is one of the central quantities of interest in the study of Markov chains.

Given that interest in the stationary distribution, the following question arises naturally: how sensitive are the entries in the stationary distribution to perturbations in the entries of the transition matrix? Specifically, if T is perturbed to yield another irreducible stochastic matrix $\tilde{T} \equiv T + E$, say with corresponding stationary distribution $\tilde{\pi}^T$, can we bound the moduli of the entries of $\tilde{\pi}^T - \pi^T$ in terms of the norm of the perturbing matrix E? For instance, if the transition probabilities arising as entries in T have been generated from some data set, one may wish to quantify the extent to which entries in the stationary distribution will be affected by errors in the data.

One technique for discussing this kind of problem is through condition numbers for the chain. A *condition number* for the stationary distribution is a real valued function $\kappa(T)$ such that for each irreducible stochastic matrix $\tilde{T} = T + E$, we have $||\tilde{\pi}^T - \pi^T||_a \le \kappa(T)||E||_b$ for some suitable norms $|| \bullet ||_a$ and $|| \bullet ||_b$. Several of these condition numbers for Markov chains are surveyed and compared in [3], and most of

*Received by the editors 12 March 2004. Accepted for publication 17 June 2004. Handling Editor: Hans Schneider.

†Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, S4S 0A2, CANADA (kirkland@math.uregina.ca). Research supported in part by an NSERC Research Grant.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

ELA

those surveyed involve the group generalized inverse $(I-T)^{\#}$ of the matrix $I-T$ (for background on group inverses, see [2]).

In this paper, we consider the conditioning of a single entry in the stationary distribution. In order to measure that conditioning, we use the following quantity, which appears in [7]. Let T and $\tilde{T} = T + E$ be as above, and for each $j = 1, \ldots, n$, define $\kappa_j(T)$ as

$$\kappa_j(T) \equiv \frac{1}{2} max\{(I-T)^{\#}_{jj} - (I-T)^{\#}_{ij} | i = 1, \ldots, n\}. \tag{1.1}$$

Letting $|| \bullet ||_\infty$ denote the absolute row sum norm, it is shown in [7] that

$$|\tilde{\pi}_j - \pi_j| \le \kappa_j(T) ||E||_\infty, \tag{1.2}$$

and further it is shown in [5] that there is a family of perturbing matrices $\mathcal{E}$ such for each $\epsilon > 0, \exists E \in \mathcal{E}$ such that $||E||_\infty < \epsilon, \tilde{T} = T + E$ is irreducible and stochastic, and

$$|\tilde{\pi}_j - \pi_j| > \frac{1}{2} \kappa_j(T) ||E||_\infty. \tag{1.3}$$

Thus the quantity κ_j provides a reasonable measure of the sensitivity of the $j-$th entry in the stationary distribution to perturbations in the transition matrix.

In this paper we consider the role played by the directed graph associated with $T, \Delta(T)$, in determining the value of $\kappa_j(T)$. Our results parallel those of [6], which adopts a similar perspective on the interplay between $\Delta(T)$ and a condition number arising in [7]. In Section 2 we provide a lower bound on $\kappa_j(T)$ in terms of $\Delta(T)$, and characterize the case of equality. In Section 3 we characterize the strongly connected directed graphs D such that $\kappa_j(T)$ is bounded as T ranges over the set of stochastic matrices such that $\Delta(T) = D$. Further, for those D such that κ_j is bounded, we find the supremum of κ_j over that set, expressing it in terms of D. Our results thus help to illuminate the influence of $\Delta(T)$ on the sensitivity of a single entry in the stationary distribution.

Throughout this paper we will rely on standard results from the theory of stochastic matrices, and on basic notions regarding directed graphs and on generalized inverses. We refer the reader to [9], [1] and [2], respectively, for background in those areas.

2. A lower bound on $\kappa_j(T)$ based on $\Delta(T)$. In this section, we establish a lower bound on the condition number κ_j in terms of the length of the shortest cycle on at least two vertices in $\Delta(T)$ that goes through vertex j. Our technique for obtaining that lower bound involves the following slight recasting of κ_j, the proof of which can be found in [4].

LEMMA 2.1. *Let T be an irreducible stochastic matrix of order n, partitioned as* $T = \left[\begin{array}{c|c} T_n & \mathbf{1} - T_n\mathbf{1} \\ \hline x^T & 1 - x^T\mathbf{1} \end{array}\right]$. *Then* $\kappa_n(T) = \dfrac{||(I-T_n)^{-1}||_\infty}{2(1 + x^T(I-T_n)^{-1}\mathbf{1})}$.

The following result uses Lemma 2.1 in order to establish a generalization of Theorem 2.10 in [4]. Our result below refers to the cyclic normal form for an irreducible periodic stochastic matrix - see Section 1.3 of [9] for background on that normal form.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

THEOREM 2.2. *Let T be an irreducible stochastic matrix of order n, and fix an index j between 1 and n. Suppose that g is the length of the shortest cycle in $\Delta(T)$ that goes through vertex j and involves at least two vertices. Then $\kappa_j(T) \geq \frac{g-1}{2g}$, with equality holding if and only if T is a periodic matrix with period g, and $\{j\}$ is one of the classes in the cyclic normal form for T.*

Proof. We suppose without loss of generality that $j = n$, and we take T to be partitioned as $T = \left[\begin{array}{c|c} T_n & \mathbf{1} - T_n\mathbf{1} \\ \hline x^T & 1 - x^T\mathbf{1} \end{array}\right]$. It then follows from Lemma 2.1 that $\kappa_n(T) \geq \frac{||(I-T_n)^{-1}||_\infty}{2(1+||(I-T_n)^{-1}||_\infty)}$, so that the inequality will follow once we show that $||(I- \ T_n)^{-1}||_\infty \geq g-1$.

For each pair of vertices $l_1, l_2 \in \Delta(T)$, let $d(l_1, l_2)$ be the distance from vertex l_1 to vertex l_2 - i.e. the length of the shortest path from l_1 to l_2 (we take this distance to be 0 if $l_1 = l_2$). Let i be a vertex on a $g-$cycle through vertex n such that $d(i, n) = g-1$ (there is such a vertex, since the length of the shortest cycle through n that is not a loop has length g). Note that $e_i^T(I-T_n)^{-1}\mathbf{1} = 1 + \sum_{l=1}^{\infty} e_i^T T_n^l \mathbf{1}$. For each $l = 1, \ldots, g-2$, there is no path from i to n of length l; thus for each such l, every walk of length l starting at vertex i passes only through vertices in $\{1, \ldots, n-1\}$. It now follows that $e_i^T T_n^l \mathbf{1} = 1$ for $l = 1, \ldots, g-2$. Consequently, $||(I-T_n)^{-1}||_\infty \geq e_i^T(I-T_n)^{-1}\mathbf{1} \geq 1+g-2 = g-1$, from which we readily find that $\kappa_n(T) \geq \frac{g-1}{2g}$.

Suppose now that $\kappa_n(T) = \frac{g-1}{2g}$. Inspecting the proof above, we find that necessarily $x^T\mathbf{1} = 1$, that $||(I-T_n)^{-1}||_\infty = g-1$, and that $x_j > 0$ only if $e_j^T(I-T_n)^{-1}\mathbf{1} = g-1$. Arguing as above, we find that for each vertex $v \neq n, e_v^T(I-T_n)^{-1}\mathbf{1} \geq d(v, n)$ so that $g-1 = ||(I-T_n)^{-1}||_\infty \geq e_v^T(I-T_n)^{-1}\mathbf{1} \geq d(v, n)$. Thus for each $v \in \{1, \ldots, n-1\}, d(v, n) \leq g-1$. Finally, we also note from the argument above that for any vertex i such that $d(i, n) = g-1$, we must have $e_i^T T_n^{g-1}\mathbf{1} = 0$, so that there is no walk from vertex i through vertices in $\{1, \ldots, n-1\}$ that has length $g-1$ or longer.

Next, we claim that each vertex in $\{1, \ldots, n-1\}$ is on a $g-$cycle that passes through vertex n. To see the claim, let $v \in \{1, \ldots, n\}$, let $d(n, v) = q$ and let $d(v, n) = p$. Suppose that the arc $n \to w$ is the first arc on a shortest path from n to v and that the arc $u \to n$ is the last arc on a shortest path from v to n. We find that there is a walk from w to u of length $p+q-2$ that passes only through vertices $\{1, \ldots, n-1\}$, from which it follows that $p+q-2 \leq g-2$. Further, there is a closed walk of length $p+q$ that passes through vertex n, so that $p+q \geq g$. Thus we find that $p+q = g$, so that v is on a closed walk of length g that passes through vertex n. It now follows that in fact v is on a $g-$cycle that passes through n, as claimed.

Let $A_0 = \{n\}$, and for each $1 \leq i \leq g-1$, let $A_i = \{u | d(u, n) = i\}$. Observe that these sets partition the vertices of the directed graph associated with T. Note that if $n \to u$, then $u \in A_{g-1}$, and that $u \in A_i, v \in A_j$ and $u \to v$, then necessarily $i \leq j+1$. We claim that in fact if this is the case, we must have $i = j+1$. To see the claim, suppose to the contrary that $i \leq j$. From our claim above, u is on some $g-$cycle C_1

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

involving vertex n; let w be the vertex on C such that $n \to w$ is an arc of C. Note that since $n \to w$, we must have $d(w, n) = g-1$, otherwise n is on a cycle (that is not a loop) of length less than g. In particular, we see that $w \in A_{g-1}$. Similarly, there is a $g-$cycle C_2 involving both v and n; let x be the vertex such that $x \to n$ is an arc of C_2. By considering C_1 and C_2, we find that there is a path from w to u of length $g-1-i$, and a path from v to x of length $j-1$, and that neither of these paths involves vertex n. Thus there is a walk from w to x passing only through vertices in $\{1, \ldots, n-1\}$ having length $g+j-i-2+1 \geq g-1$, a contradiction. We conclude that necessarily $i = j+1$, so that the only arcs in the directed graph associated with T are those from a vertex in A_{j+1} to a vertex in A_j (where the subscripts are taken modulo g). From the fact that T is irreducible and that the shortest cycle through vertex n on at least two vertices has length g, it follows that T is periodic with period g, and from the construction above, we see that $\{n\}$ is one of the classes in the periodic normal form for T.

Finally, suppose that T is periodic with period g, and that $\{n\}$ is one of the classes in the periodic normal form for T. Without loss of generality, T can be written as a $g \times g$ block matrix of the form

$$T = \left[\begin{array}{c|c|c|c|c|c} 0 & S_1 & 0 & 0 & \ldots & 0 \\ \hline 0 & 0 & S_2 & 0 & \ldots & 0 \\ \hline \vdots & & \ddots & & & \vdots \\ \hline 0 & 0 & \ldots & & 0 & \mathbf{1} \\ \hline \sigma^T & 0^T & \ldots & & 0^T & 0 \end{array}\right],$$

where the last diagonal block is 1×1, $\sigma^T > 0$ and $\sigma^T \mathbf{1} = 1$. Applying the notation above, we find readily that $||(I-T_n)^{-1}||_\infty = g-1$ with the rows corresponding to the first block in the partitioning of T yielding the maximum row sums for $(I-T_n)^{-1}$. It now follows that $\kappa_n(T) = \dfrac{g-1}{2g}$. ☐

Our first corollary recasts the case of equality in Theorem 2.2 in a more graph-theoretic manner.

Corollary 2.3. *Let T be as in Theorem 2.2. Then $\kappa_j(T) = \dfrac{g-1}{2g}$ if and only if every cycle of $\Delta(T)$ has length g, and vertex j is on each cycle of $\Delta(T)$.*

Proof. Suppose that $\kappa_j(T) = \dfrac{g-1}{2g}$. From Theorem 2.2, we find that the greatest common divisor of the cycle lengths in $\Delta(T)$ is g, and that $\{j\}$ is one of the classes in the cyclic normal form for T. In particular, every cycle in $\Delta(T)$ goes through vertex j. Further, there is no vertex $i \neq j$ such that $d(j,i) \equiv 0 \ mod \ g$ (otherwise such an i belongs to the same class as j in the cyclic normal form). It now follows that $\Delta(T)$ contains no cycle of length greater than g; thus every cycle must have length g. The converse is straightforward. ☐

Recalling that there is a connection between κ_j and certain mean first passage times for the Markov chain, (see [3], for example) we have the following result.

Corollary 2.4. *Let T be an irreducible stochastic matrix of order n. Fix an index j, and let g be the length of the shortest cycle in $\Delta(T)$ that goes through vertex*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

j and has at least two vertices. Let M be the mean first passage matrix associated with T. Then there is some index $i \neq j$ such that $m_{i,j} \geq \frac{g-1}{2g} m_{j,j}$.

The following is an immediate consequence of (1.3).

Corollary 2.5. *Let T be an irreducible stochastic matrix of order n, and suppose that g is the length of the shortest cycle in $\Delta(T)$ that goes through vertex j and has at least two vertices. Then there is a family of perturbation matrices $\mathcal{E}$ such that for any $\epsilon > 0, \exists E \in \mathcal{E}$ such that $||E||_\infty < \epsilon, \tilde{T} = T + E$ is irreducible and stochastic, and $|\tilde{\pi}_j - \pi_j| \geq \frac{g-1}{4g}||E||_\infty$.*

Remark 2.6. A minor variation on the proof of Theorem 2.2 shows that (keeping the same notation as that result) if g is the length of the shortest cycle in $\Delta(T)$ that goes through vertex n and is not a loop, and if $T_{nn} = a > 0$, then $\kappa_n(T) \geq \frac{g-1}{2(1+(1-a)(g-1))}$, with equality holding if and only if T can be written as a $g \times g$ block matrix of the form

$$T = \left[\begin{array}{c|c|c|c|c|c} 0 & S_1 & 0 & 0 & \ldots & 0 \\ \hline 0 & 0 & S_2 & 0 & \ldots & 0 \\ \hline \vdots & & \ddots & & & \vdots \\ \hline 0 & 0 & \ldots & & 0 & \mathbf{1} \\ \hline (1-a)\sigma^T & 0^T & \ldots & & 0^T & a \end{array}\right],$$

where $\sigma^T > 0$ and $\sigma^T \mathbf{1} = 1$.

3. Digraphs for which κ_j is bounded. In this section, we investigate the extent to which the information in $\Delta(T)$ can be used to produce an upper bound on $\kappa_j(T)$. (For notational convenience, we restrict ourselves henceforth to the case that $j = n$, without loss of generality.) To this end, we consider the following problem: given a strongly connected graph D, we seek conditions on D so that $\kappa_n(T)$ is bounded from above as T ranges over $\mathcal{S}_D$, the set of (necessarily irreducible) stochastic matrices T with $\Delta(T) = D$, and for those D such that κ_n *is* bounded, we seek a tight graph theoretic upper bound on κ_n. A similar question is dealt with in [6] for the quantity $c(T) \equiv max\{\kappa_j(T)|j = 1, \ldots, n\}$, which functions as a condition number for the stationary distribution (see Section 1). Despite the similarity with the problem considered in [6], here some of the details are slightly more technical. A key lemma in [6] ensures that if D is a strongly connected digraph such that $c(T)$ is bounded on $\mathcal{S}_D$, then in fact for any stochastic matrix M in the closure of $\mathcal{S}_D$, M has 1 as an algebraically simple eigenvalue, so that a group generalized inverse for $I - M$ exists. In the present setting, it turns out that there are directed graphs D such that $\kappa_n(T)$ is bounded over $\mathcal{S}_D$, but where the closure of $\mathcal{S}_D$ contains stochastic matrices M having 1 as an eigenvalue of algebraic multiplicity 2 or more, so that $I - M$ does not possess a group inverse. Indeed Example 3.1 discusses just such a digraph.

Consequently, as part of our investigation, we will make use of the set $\Sigma_D \equiv \{T|T$ is $n \times n$, stochastic, has a single essential class of indices, and $\Delta(T) \subseteq D\}$ (see [9] for a discussion of essential classes). We remark that $\mathcal{S}_D$ is easily seen to be a

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

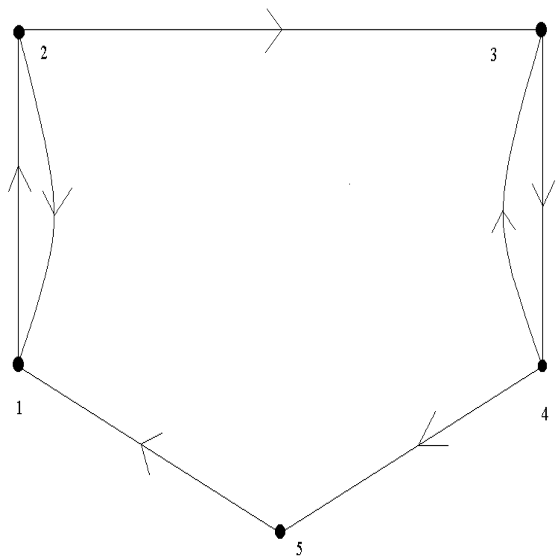

Figure 1

dense subset of Σ_D. Further, if $A \in \Sigma_D$, then $I - A$ has 1 as an algebraically simple eigenvalue, so that $(I-A)^{\#}$ exists. We claim that if $\kappa_n(A)$ is bounded from above for $A \in \Sigma_D$, then $sup\{\kappa_n(T)|T \in \mathcal{S}_D\} = sup\{\kappa_n(A)|A \in \Sigma_D\}$. To see the claim, note that certainly for any stochastic $T \in \mathcal{S}_D$, we have $\kappa_n(T) \le sup\{\kappa_n(A)|A \in \Sigma_D\}$. Now let A_m denote a sequence of matrices in Σ_D such that $\kappa_n(A_m) \to sup\{\kappa_n(A)|A \in \Sigma_D\}$ as $m \to \infty$. Note that for each $m \in \mathbb{N}$, $I - A_m$ has a group inverse. Further, for each $m \in \mathbb{N}$, the group inverse is continuous at $I - A_m$ (see [6]); since $\mathcal{S}_D$ is dense in Σ_D, it follows that there is a sequence of such matrices in $\mathcal{S}_D$, say $T_m \in \mathcal{S}_D$, such that $\kappa_n(T_m) \to sup\{\kappa_n(A)|A \in \Sigma_D\}$ as $m \to \infty$. The claim now follows. A similar argument shows that if $\kappa_n(A)$ is not bounded from above for $A \in \Sigma_D$, then there is a sequence of matrices $T_m \in \mathcal{S}_D$ such that $\kappa_n(T_m) \to \infty$ as $m \to \infty$. Consequently, we see that κ_n is bounded over $\mathcal{S}_D$ if and only if it is bounded over Σ_D, and that in the case that κ_n is bounded over both sets, $sup\{\kappa_n(T)|T \in \mathcal{S}_D\} = sup\{\kappa_n(A)|A \in \Sigma_D\}$.

We begin with an illuminating example.

Example 3.1. Consider digraph D_0 given in Figure 1, and note that a typical matrix $T \in \mathcal{S}_{D_0}$ is of the form $T = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1-a & 0 & a & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1-b & 0 & b \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$ for some $0 < a, b < 1$.
We note that the main result in [6] asserts that for a strongly connected directed graph D, $c(T)$ is bounded above as T ranges over $\mathcal{S}_D$ if and only if any pair of cycles

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

in D share at least one vertex. Noticing that D_0 has at least two vertex disjoint cycles, we see that $c(T)$ is not bounded as T ranges over $\mathcal{S}_{D_0}$.

However, letting the leading principle submatrix of $I - T$ of order 4 be U, we see from Lemma 2.1 that $\kappa_5(T) = \frac{||U^{-1}||_\infty}{2(1+e_1^T U^{-1}\mathbf{1})}$. We find readily that $U^{-1}\mathbf{1} = \begin{bmatrix} \frac{2}{a}+\frac{2}{b} \\ \frac{2}{a}+\frac{2}{b}-1 \\ \frac{2}{b} \\ \frac{2}{b}-1 \end{bmatrix}$, so that $\kappa_5(T) = \dfrac{\frac{1}{a}+\frac{1}{b}}{1+\frac{2}{a}+\frac{2}{b}} < \dfrac{1}{2}$. In particular, we see that $\kappa_5(T)$ is bounded above by $\dfrac{1}{2}$ for each $T \in \mathcal{S}_{D_0}$. Thus our example illustrates that it is possible for $\kappa_n(T)$ to be bounded over $\mathcal{S}_D$ even if $c(T)$ is unbounded over that set.

Next, we establish a necessary condition for κ_n to be bounded over $\mathcal{S}_D$.

LEMMA 3.2. *Let D be a strongly connected directed graph on vertices $1,\ldots,n$. Suppose that D has two disjoint cycles, one of which goes through vertex n. Then $\kappa_n(T)$ is unbounded from above as T ranges over $\mathcal{S}_D$.*

Proof. Without loss of generality, suppose that the two disjoint cycles in D are C_1 given by $1 \to 2 \to \ldots \to k \to 1$ and C_2 given by $n-j+1 \to n-j+2 \to \ldots \to n \to n-j+1$, where $k < n-j+1$. For each $1 > \epsilon > 0$ and $1 > \delta > 0$, construct a matrix $T_{\epsilon,\delta} \in \mathcal{S}_{\mathcal{D}}$ as follows: let the entries of $T_{\epsilon,\delta}$ corresponding to arcs on C_1 be $1-\epsilon$, let the entries of $T_{\epsilon,\delta}$ corresponding to arcs on C_2 be $1-\delta$, and fill in the remaining entries of $T_{\epsilon,\delta}$ so that it is stochastic, irreducible, and has $\Delta(T_{\epsilon,\delta}) = D$. Write $T_{\epsilon,\delta}$ as

$$T_{\epsilon,\delta} = \left[\begin{array}{c|c} A_{\epsilon,\delta} & \mathbf{1} - A_{\epsilon,\delta}\mathbf{1} \\ \hline x_{\epsilon,\delta}^T & 1 - x_{\epsilon,\delta}^T\mathbf{1} \end{array}\right].$$

From the fact that the principal submatrix of $A_{\epsilon,\delta}$ on its first k rows and columns has constant row sums $1-\epsilon$, it follows that for each $i = 1,\ldots,k, e_i^T(I - A_{\epsilon,\delta})^{-1}\mathbf{1} \geq \dfrac{1}{\epsilon}$. Also, note that as $\delta \to 0^+$, the principal submatrix of $A_{\epsilon,\delta}$ on its last $j-1$ rows and columns converges to a nilpotent matrix while the submatrix of $A_{\epsilon,\delta}$ on rows $n-j+1,\ldots,n-1$ and columns $1,\ldots,n-j$ converges to the zero matrix. Hence, $\forall \epsilon \in (0,1), \exists \delta(\epsilon) > 0$ such that $x_{\epsilon,\delta(\epsilon)}^T(I - A_{\epsilon,\delta(\epsilon)})^{-1}\mathbf{1} < j$. It now follows that as $\epsilon \to 0^+$, $\kappa_n(T_{\epsilon,\delta(\epsilon)})$ diverges to ∞. ☐

The following result will be useful in the sequel. Its proof is essentially the same as that in [8], which proves the result in the case that both T and $T+E$ are irreducible.

LEMMA 3.3. *Suppose that T and $T+E$ are stochastic matrices, each of which has 1 as an algebraically simple eigenvalue. Let π^T be the stationary distribution for T, and let $Q = I - T$, so that $Q - E = I - (T+E)$. If $I - EQ^\#$ is invertible, then $(Q-E)^\# = Q^\#(I - EQ^\#)^{-1} - \mathbf{1}\pi^T(I - EQ^\#)^{-1}Q^\#(I - EQ^\#)^{-1}$.*

We now apply the lemma above in order to establish the following.

LEMMA 3.4. *Suppose that T is an $n \times n$ stochastic matrix, and that for some $1 \leq i \leq n$, the $i-$th row of T has at least two positive entries, say $T_{i,a}, T_{i,b} > 0$. Suppose further that for each $\epsilon \in [-T_{i,a}, T_{i,b}]$, the matrix $T(\epsilon) \equiv T + \epsilon e_i(e_a - e_b)^T$ has 1 as an algebraically simple eigenvalue. Then considered as a function of ϵ on $[-T_{i,a}, T_{i,b}]$, $\kappa_n(T(\epsilon))$ attains its maximum at either $\epsilon = -T_{i,a}$ or $\epsilon = T_{i,b}$.*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

Proof. Set $Q(\epsilon) \equiv I - T(\epsilon)$, and note that from Lemma 3.3, it follows that $Q(\epsilon)^{\#}$ is a continuous function of ϵ on $[-T_{i,a}, T_{i,b}]$; in particular, $\kappa_n(T(\epsilon))$ certainly attains its maximum on that interval. Suppose now that the maximum is attained at some $\epsilon_0 \in (-T_{i,a}, T_{i,b})$, say with $max_{\epsilon\in[-T_{i,a},T_{i,b}]}\kappa_n(T(\epsilon)) = (Q(\epsilon_0)^{\#}_{n,n} - Q(\epsilon_0)^{\#}_{k,n})/2$. Note that for all $t \in \mathbb{R}$ with $|t|$ sufficiently small, we have

$$Q(\epsilon_0 + t)^{\#} = Q(\epsilon_0)^{\#}(I + te_i(e_a - e_b)^t Q(\epsilon_0)^{\#})^{-1} + \mathbf{1}z^T,$$

for some vector z^T (depending on t). It follows that for such t,

$$Q(\epsilon_0 + t)^{\#}_{n,n} - Q(\epsilon_0 + t)^{\#}_{k,n} = Q(\epsilon_0)^{\#}_{n,n} - Q(\epsilon_0)^{\#}_{k,n}$$

$$-\frac{t}{1 + t(Q(\epsilon_0)^{\#}_{a,i} - Q(\epsilon_0)^{\#}_{b,i})}(Q(\epsilon_0)^{\#}_{n,i} - Q(\epsilon_0)^{\#}_{k,i})(Q(\epsilon_0)^{\#}_{a,n} - Q(\epsilon_0)^{\#}_{b,n}).$$

Since $\kappa_n(T(\epsilon)$ is maximized at ϵ_0, we find that necessarily

$$(Q(\epsilon_0)^{\#}_{a,n} - Q(\epsilon_0)^{\#}_{b,n})(Q(\epsilon_0)^{\#}_{n,i} - Q(\epsilon_0)^{\#}_{k,i}) = 0.$$

But in that case, we find that for any t such that $\epsilon_0 + t \in [-T_{i,a}, T_{i,b}]$ and $1 + t(Q(\epsilon_0)^{\#}_{a,i} - Q(\epsilon_0)^{\#}_{b,i}) \neq 0$, we have $Q(\epsilon_0 + t)^{\#}_{n,n} - Q(\epsilon_0 + t)^{\#}_{k,n} = Q(\epsilon_0)^{\#}_{n,n} - Q(\epsilon_0)^{\#}_{k,n}$. Thus, selecting t so that $\epsilon_0 + t \in \{-T_{i,a}, T_{i,b}\}$ and $t(Q(\epsilon_0)^{\#}_{a,i} - Q(\epsilon_0)^{\#}_{b,i}) \geq 0$, we find that for either $\epsilon = -T_{i,a}$ or $\epsilon = T_{i,b}$, $(Q(\epsilon)^{\#}_{n,n} - Q(\epsilon)^{\#}_{k,n})/2 = max_{\epsilon\in[-T_{i,a},T_{i,b}]}\kappa_n(T(\epsilon))$. Thus the maximum for $\kappa_n(T(\epsilon))$ is attained at an end point of $[-T_{i,a}, T_{i,b}]$. □

Next, we apply Lemma 3.4 to help to identify a useful subset of sparse matrices in Σ_D.

PROPOSITION 3.5. *Let D be a strongly connected directed graph on vertices $1, \ldots, n$, and suppose that each cycle that goes through vertex n intersects every other cycle in D. Suppose that $T \in \Sigma_D$ and that vertex n of $\Delta(T)$ is on a cycle. Then there is a matrix $M \in \Sigma_D$ such that $\kappa_n(T) \leq \kappa_n(M)$ and such that $\Delta(M)$ has the following properties:*
i) vertex n is on exactly one cycle C in $\Delta(M)$;
ii) if vertex i is not on C, then its outdegree in $\Delta(M)$ is 1;
iii) if vertex $i \neq n$ is on C, then its outdegree in $\Delta(M)$ is at most 2;
iv) vertex n has outdegree 1 in $\Delta(M)$.

Proof. First, let A be any matrix in Σ_D with the property that in $\Delta(A)$, vertex n is on at least one cycle. We claim that 1 is an algebraically simple eigenvalue, so that $(I - A)^{\#}$ exists. The see the claim, note that from the fact that n is on a cycle in $\Delta(A)$, there is a nontrivial strongly connected component of $\Delta(A)$ that contains n, say on vertices $n - j + 1, \ldots, n$. From the hypothesis on D, we see that there are no cycles in the subgraph induced by vertices $1, \ldots, n - j$, from which we deduce that A has exactly one essential class of indices, namely the class containing vertex n. The claim now follows.

Suppose that vertex n of $\Delta(T)$ is on at least two cycles, say C_1 and C_2. Observe that from the hypothesis on D, those two cycles must intersect, say with vertex i on

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

ELA

both C_1 and C_2. But then row i of T has two positive entries, say $T_{i,a}, T_{i,b} > 0$, where the arcs $i \to a$ and $i \to b$ are on C_1 and C_2, respectively. Further, observe that from the claim above, both $T - T_{i,a}e_i(e_a - e_b)^T$ and $T + T_{i,b}e_i(e_a - e_b)^T$ are in Σ_D, and note that in the directed graph of each, vertex n is on at least one cycle, but is on fewer cycles than in $\Delta(T)$. It follows that T satisfies the hypotheses of Lemma 3.4, and we conclude that there is a matrix $\hat{T} \in \Sigma_D$ such that vertex n is on a cycle in $\Delta(\hat{T})$, $\kappa_n(T) \leq \kappa_n(\hat{T})$, and vertex n is on fewer cycles in $\Delta(\hat{T})$ than in $\Delta(T)$. Thus we may iterate the argument above on $\hat{T}$, and so produce a matrix $\tilde{T} \in \Sigma_D$ such that vertex n is on just one cycle C in $\Delta(\tilde{T})$, each vertex on C has outdegree at most 2 in $\Delta(\tilde{T})$, and $\kappa_n(T) \leq \kappa_n(\tilde{T})$. Note that necessarily, vertex n of $\Delta(\tilde{T})$ has outdegree 1, for if not, the hypothesis on D implies that there is a path in $\Delta(\tilde{T})$ of the form $1 \to u_1 \to \ldots \to u_l \to u_{l+1}$ such that $u_1, \ldots, u_l \notin C$, but $u_{l+1} \in C$, contradicting the fact that n is on just one cycle in $\Delta(\tilde{T})$.

Observe that if v is a vertex of $\Delta(\tilde{T})$ that is not on C and that has outdegree at least 2, we may again iteratively apply Lemma 3.4 to $\tilde{T}$ to produce a matrix $\bar{T}$ so that vertex v of $\Delta(\bar{T})$ has outdegree 1, all other vertices of $\Delta(\bar{T})$ have the same outarcs as in $\Delta(\tilde{T})$, and $\kappa_n(\tilde{T}) \leq \kappa_n(\bar{T})$. It now follows that we can construct a matrix $M \in \Sigma_D$ such that $\kappa_n(T) \leq \kappa_n(M)$ and such that $\Delta(M)$ satisfies properties i)-iv). ∎

Remark 3.6. Suppose that M is a $(0,1)$ matrix that satisfies properties i)-iv) of Proposition 3.5, and suppose that the cycle C has length g. Observe that for each $i = 1, \ldots, n-1$, there is a unique path from vertex i to vertex n in $\Delta(M)$, say of length $d(i,n)$. It follows from Theorem 2.3 of [6] that $\kappa_n(M) = max_{1 \leq i \leq n-1} \dfrac{d(i,n)}{2g}$.

Further, if M is a $(0,1)$ matrix such that $\Delta(M)$ has exactly one cycle, and that cycle does not go through vertex n, we find from Theorem 2.3 of [6] that $\kappa_n(M) = 1/2$.

The information in Remark 3.6 is helpful in the proof of the following result.

Proposition 3.7. *Let D be a strongly connected graph on n vertices. Suppose that for the matrix $M \in \Sigma_D$, $\Delta(M)$ satisfies i)-iv) of Proposition 3.5, and that the cycle C in $\Delta(M)$ going through vertex n has length g. If $\Delta(M)$ has at least one vertex of degree 2, then $\kappa_n(M) \leq max \left\{1/2, \dfrac{d(1,n)}{2g}, \ldots, \dfrac{d(n-1,n)}{2g}\right\}$. Further, there is a matrix $\hat{M} \in \Sigma_D$ such that $\kappa_n(\hat{M}) = max \left\{1/2, \dfrac{d(1,n)}{2g}, \ldots, \dfrac{d(n-1,n)}{2g}\right\}$.*

Proof. We proceed by induction on the number of vertices of degree 2 in $\Delta(M)$. Suppose that there is just one vertex of degree 2, say vertex j, necessarily on C. We have $j \to j+1$, where $j+1 \in C$ and $j \to l$, say, where $l \notin C$. Note that $\Delta(M) \setminus \{j \to l\}$ has just one essential class of indices (since n is still on a cycle in that graph), and that $\Delta(M) \setminus \{j \to j+1\}$ also has just one essential class of indices, namely the class containing vertex j, since every vertex in that graph has a path to vertex j. Consequently, we see that for each $\epsilon \in [-M_{j,l}, M_{j,j+1}]$, $M + \epsilon e_j(e_l - e_{j+1})^T$ has 1 as an algebraically simple eigenvalue. Letting $M_1 = M - M_{j,l}e_j(e_l - e_{j+1})^T$ and $M_2 = M + M_{j,j+1}e_j(e_l - e_{j+1})^T$, we find from Lemma 3.4, we see that $\kappa_n(M) \leq max\{\kappa_n(M_1), \kappa_n(M_2)\}$. Note that M_1 is a $(0,1)$ matrix, so

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

that $\kappa_n(M_1) = max_{1\le i\le n}\left\{\dfrac{d(i,n)}{2g}\right\}$, while since vertex n of $\Delta(M_2)$ is not any cycle, $\kappa_n(M_2) = 1/2$. Finally, note that we may select $\hat{M}$ to be M_1 or M_2, according as $1 \le max_{1\le i\le n}\left\{\dfrac{d(i,n)}{g}\right\}$ or $1 > max_{1\le i\le n}\left\{\dfrac{d(i,n)}{g}\right\}$, respectively.

Next we suppose that statements hold whenever there are p vertices of degree 2 on the cycle in $\Delta(M)$ through vertex n , and that in $\Delta(M)$, the cycle C through vertex n has $p+1$ vertices of degree 2. Let j be the vertex on C of degree 2 whose distance to n is smallest. Suppose that $j \to j+1$, where $j+1 \in C$ and $j \to l$, say, where $l \notin C$. Arguing as above, we find that $\forall \epsilon \in [-M_{j,l}, M_{j,j+1}], M + \epsilon e_j(e_l - e_{j+1})^T$ has 1 as an algebraically simple eigenvalue. Applying Lemma 3.4, we find that $\kappa_n(M) \le max\{\kappa_n(M_1), \kappa_n(M_2)\}$, where $M_1 = M - M_{j,l}e_j(e_l - e_{j+1})^T$ and $M_2 = M + M_{j,j+1}e_j(e_l - e_{j+1})^T$. Now $\Delta(M_1)$ satisfies i)-iv) and has p vertices of degree 2 on the cycle through vertex n, so by the induction hypothesis, $\kappa_n(M_1) \le max\left\{1/2, \dfrac{d(1,n)}{2g}, \ldots, \dfrac{d(n-1,n)}{2g}\right\}$. Note also that vertex n of M_2 is not on any cycle, so that $\kappa_n(M_2) = 1/2$. The inequality on $\kappa_n(M)$ now follows. Also, from the induction step, there is a matrix $\hat{M} \in \Sigma_D$ such that $\kappa_n(\hat{M}) = max\left\{1/2, \dfrac{d(1,n)}{2g}, \ldots, \dfrac{d(n-1,n)}{2g}\right\}$. This completes the induction. □

Here is our main result.

THEOREM 3.8. *Suppose that D is a strongly connected directed graph on n vertices with the property that each cycle that goes through vertex n intersects every other cycle in D. Suppose further that at least one cycle of D does not go through vertex n. Let G be a subgraph of D having the properties that*
a) G has a unique cycle, say of length g, and that cycle goes through vertex n, and
b) each vertex of G has outdegree 1.
For each such graph G, let $\delta_n(G) = max_{\{1\le i\le n\}}\left\{\dfrac{d(i,n)}{2g}\right\}$. Then $max\{\kappa_n(T)|T \in \Sigma_D\} = max\{1/2, \delta_n(G)\}$, where the latter maximum is taken over all subgraphs G of D satisfying a) and b).

Proof. From Proposition 3.5 we see that if $T \in \Sigma_D$, then there is a matrix $M \in \Sigma_D$ such that $\kappa_n(T) \le \kappa_n(M)$ and such that $\Delta(M)$ satisfies properties i)-iv) of Proposition 3.5. From Proposition 3.7, we see that necessarily $\kappa_n(M) \le max\{1/2, \delta_n(G)\}$, where the maximum is taken over all subgraphs G of D satisfying a) and b).

Let G be a graph satisfying a) and b) such that $\delta_n(G)$ is maximal (where the maximum is taken over all $G \subseteq D$ satisfying a) and b)). Letting A be the $(0,1)$ adjacency matrix of G, we see that $A \in \Sigma_D$ and, from Remark 3.6, that $\kappa_n(A) = max\{\delta_n(G)\}$. Thus we see that $max\{\delta_n(G)\}$ is attainable as $\kappa_n(M)$ for some $M \in \Sigma_D$.

Next we show that $1/2$ is also attainable. Note that since D contains a cycle C_1 that does not go through vertex n, we claim that there is a graph $\hat{G} \subseteq D$ such that every vertex has degree 1, and $\hat{G}$ has just one cycle, namely C_1. To see the claim, let C_2 be a cycle of D that goes through vertex n, and observe that C_1 and C_2 intersect in at least one vertex. For each vertex i off of $C_1 \cup C_2$, note that there is a path in D from i to a vertex on $C_1 \cup C_2$. We may then readily construct a graph $G \subseteq D$ such

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

that every vertex off of $C_1 \cup C_2$ has outdegree 1, and from which there is a (unique) path to $C_1 \cup C_2$, while the remaining arcs of G are those on $C_1 \cup C_2$. Now construct $\hat{G}$ from G as follows: for each vertex i on $C_1 \cap C_2$ delete the arc from i to the vertex on C_1. Then in $\hat{G}$ there is just one essential class of indices, and that class does not contain n. Letting $\hat{A}$ be the adjacency matrix for $\hat{G}$, it now follows that $\kappa_n(\hat{A}) = 1/2$. In particular $\hat{A} \in \Sigma_D$ and $\kappa_n(\hat{A}) = 1/2$; the result now follows. □

REMARK 3.9. Note that Theorem 3.8 applies only to directed graphs having at least one cycle that does not go through vertex n. Suppose now that D is a strongly connected directed graph such that every cycle goes through vertex n. It follows from a minor modification of the proof of Theorem 3.8 that $max\{\kappa_n(T)|T \in \Sigma_D\} = max\{\delta_n(G)\}$, where the latter maximum is taken over all subgraphs G of D satisfying a) and b).

We close with an example dealing with a particularly well-structured class of directed graphs.

EXAMPLE 3.10. In this example, we consider the strongly connected directed graphs D that correspond to stochastic Hessenberg matrices. Observe that a stochastic Hessenberg matrix can be taken to have the form

$$T = \begin{bmatrix} 1-a_0 & a_0 & 0 & 0 & \dots & 0 \\ b_1 & 1-b_0-b_1 & b_0 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ c_{n-1} & c_{n-2} & \dots & & c_1 & 1-\sum_{i=1}^{n-1} c_i \end{bmatrix};$$

such matrices arise in the study of $M \setminus G \setminus 1$ queues, and can be thought of as corresponding to a single server queue with a maximum of $n-1$ customers in line (arriving customers are lost to the system if there are already $n-1$ customers in line). Here, the chain is in state i if there are $n-i$ customers in the queue, so that state n corresponds to an empty queue.

The directed graph D for such a matrix has the following form: for each $i = 1, \dots, n-1$, D contains the arc $i \to i+1$, while all remaining arcs (if any) are of the form $i \to j$, where $1 \le j \le i \le n$. Suppose that D is such a graph, and that vertex n of D has outarcs $n \to i_0, \dots, n \to i_k$, where $1 \le i_0 < i_1 < \dots < i_k \le n$. Note that in order for D to be strongly connected, it must also contain an arc of the form $j \to 1$, where $i_0 \le j$. Observe that the cycles in D are in one-to-one correspondence with the positive elements of T on and below the diagonal, or equivalently, with the arcs $i \to j$, where $1 \le j \le i \le n$. Evidently such an arc corresponds to the cycle $i \to j \to j+1 \to \dots \to i \to j$.

Thus we find that every cycle going through vertex n also goes through each of vertices $i_k, i_{k+1}, \dots, n-1$, from which it follows that a particular cycle C intersects every cycle through vertex n if and only if C goes through some vertex i with $i_k \le i \le n$. From Theorem 3.7, we find that $\kappa_n(T)$ is bounded as T ranges over Σ_D if and only if for each arc of D of the form $i \to j$, where $1 \le j \le i \le n$, we must have $i_k \le i$. (In particular, that condition implies that if $i \to 1$ then $i_k \le i$.) Note that when D satisfies that condition, we find from Theorem 3.7 that if $1 < i_k$, then

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 168-179, July 2004

$max\{\kappa_n(T)|T \in \Sigma_D\} = \dfrac{n-1}{2(n-i_k+1)}$, while if $i_k = 1$, then $max\{\kappa_n(T)|T \in \Sigma_D\}$ is either $\dfrac{n-1}{2n}$ or 1/2 according as D is an $n-$cycle or not, respectively.

Acknowledgment. The author is grateful to Professor Michael Neumann for helpful conversations during the early phases of this research. In particular, Remark 2.6 is due to Professor Neumann.

REFERENCES

[1] R. Brualdi and H. Ryser. *Combinatorial Matrix Theory.* Cambridge University Press, Cambridge, 1991.
[2] S. Campbell and C. Meyer. *Generalized Inverses of Linear Transformations.* Dover, New York, 1991.
[3] 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.
[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. Digraph-based conditioning for Markov chains. *Linear Algebra and its Applications*, 385:81–93, 2004.
[7] 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.
[8] 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.
[9] E. Seneta. *Non-negative Matrices and Markov Chains.* Springer-Verlag, New York, 1981.