Conditioning of the Entries in the Stationary Vector of a Google-Type Matrix

Steve Kirkland
University of Regina

June 15, 2006

Motivation: Google's PageRank algorithm finds the stationary vector of a stochastic matrix having a particular structure.

Start with a directed graph D on n vertices, with a directed arc from vertex i to vertex j if and only if page i has a link out to page j.

Next, a stochastic matrix A is constructed from the directed graph as follows. For each i, j, we have $a_{ij} = 1/d(i)$ if the out-degree of vertex $i, d(i)$ is positive and $i \rightarrow j$ in the directed graph D, and $a_{ij} = 0$ if $d(i) > 0$ but there is no arc from i to j in D. Finally, if vertex i has outdegree zero, we have $a_{ij} = 1/n$ for all j, where n is the order of the matrix.

Note that because of the disconnected nature of the web, A typically has several direct summands that are stochastic.

Next, a positive row vector v^T is selected, normalized so that $v^T\mathbf{1} = 1$. ($\mathbf{1}$ is the all ones vector here.)

Finally a parameter $c \in (0, 1)$ is chosen (Google reports that c is approximately 0.85), and the *Google matrix* G is constructed as follows:

$$G = cA + (1 - c)\mathbf{1}v^T. \tag{1}$$

It is the stationary distribution vector of G that is estimated, and the results are then used in Google's ranking of the pages on the web.

Motivated by the Google matrix, we consider the following class of *Google-type* stochastic matrices:

$$M = cA + (1 - c)\mathbf{1}v^T, \qquad (2)$$

where A is an $n \times n$ stochastic matrix, $c \in (0, 1)$ and v^T is a nonnegative row vector such that $v^T\mathbf{1} = 1$. Denote its stationary distribution vector by π^T.

Throughout, we impose the additional hypothesis that for index $1 \leq i \leq n$, the principal submatrix of $I - M$ formed by deleting row and column i is invertible.

Observe that in the special case that v^T is a positive vector and A is block triangular with at least two diagonal blocks that are stochastic, a matrix of the form (2) coincides with the Google matrix G of (1).

A General Question: Suppose that we have an $n \times n$ stochastic matrix S that has 1 as an algebraically simple eigenvalue, and stationary distribution vector σ^T. Given a row vector x^T whose entries sum to 1, how close is x^T to σ^T?

A Useful Approach: It turns out that $I - S$ has a unique *group generalized inverse*, $(I-S)^{\#}$, with the following properties:
$(I-S)^{\#}\mathbf{1} = 0, \sigma^T(I-S)^{\#} = 0^T$,
$(I-S)(I-S)^{\#} = (I-S)^{\#}(I-S) = I - \mathbf{1}\sigma^T$.

So, setting $y^T = x^T(I-S)$, we have
$y^T(I-S)^{\#} = x^T(I-S)(I-S)^{\#} =$
$x^T(I - \mathbf{1}\sigma^T) = x^T - \sigma^T$.

Objective: For a Google-type matrix M, want to discuss the conditioning of the stationary vector. That is, if we have an estimate p^T of the stationary vector for M, want to get a sense of the accuracy of that estimate.

Specifically, want to fix an index $j = 1, \ldots, n$, and consider the following questions:

Question 1. Given a vector p^T whose entries sum to 1, how close is p_j to π_j?

Question 2. If p^T is an estimate of π^T and we know that $p_i \geq p_j$, under what circumstances can we conclude that $\pi_i \geq \pi_j$?

Componentwise Error Bounds

Setup: Set $r^T = p^T(I-M)$. For each $j = 1, \ldots, n$, it turns out that $p_j - \pi_j = r^T(I-M)^{\#}e_j$. It follows that $|p_j - \pi_j| \leq \frac{||r^T||_1}{2} \max\{(I-M)^{\#}_{k,j} - (I-M)^{\#}_{i,j} | i,k = 1, \ldots, n\}$.

Handy Fact: For each $j = 1, \ldots, n$, we have $\frac{1}{2} \max\{(I-M)^{\#}_{k,j} - (I-M)^{\#}_{i,j} | i,k = 1, \ldots, n\} = \frac{1}{2}\pi_j ||(I-M_j)^{-1}||_\infty \equiv \kappa_j(M)$, where $|| \; ||_\infty$ denotes the maximum absolute row sum norm and $(I-M)_j$ is formed from $I-M$ by deleting the j^{th} row and column.

Theorem 1: a) Suppose that p^T is an n-vector whose entries sum to 1. Then for each $j = 1, \ldots, n$, we have $|p_j - \pi_j| \leq ||r^T||_1 \kappa_j(M)$.

b) Fix an index j between 1 and n. For each sufficiently small $\epsilon > 0$, there is a positive vector p^T whose entries sum to 1 such that $||r^T||_1 = \epsilon$ and $|p_j - \pi_j| = ||r^T||_1 \kappa_j(M)$.

Good news: $\kappa_j(M)$ provides a precise measure of the difference between p_j and π_j. *Bad news*: $\kappa_j(M)$ looks like it's tricky to compute.

Consider the case $j = n$. Write

$$A = \left[\begin{array}{c|c} A_n & \mathbf{1} - A_n\mathbf{1} \\ \hline a^T & 1 - a^T\mathbf{1} \end{array}\right], \pi^T = \left[\overline{\pi}^T | \pi_n\right], v^T = \left[\overline{v}^T | v_n\right]. \tag{3}$$

Lemma 1: Suppose that A, π^T and v^T are partitioned as in (3). We have the following.

a) $(I - M_n)^{-1}\mathbf{1} = \frac{(I-cA_n)^{-1}\mathbf{1}}{(1-(1-c)\overline{v}^T(I-cA_n)^{-1}\mathbf{1})}$.

b) $\pi_n = \frac{1-(1-c)\overline{v}^T(I-cA_n)^{-1}\mathbf{1}}{1+ca^T(I-cA_n)^{-1}\mathbf{1}}$.

Theorem 2: Suppose that the matrix A is partitioned as in (3). Then $\kappa_n(M) = max\{\frac{e_i^T(I-cA_n)^{-1}\mathbf{1}}{2(1+ca^T(I-cA_n)^{-1}\mathbf{1})} | i = 1, \ldots, n-1\}$.

Strategy: Want to use the directed graph associated with A, $\Delta(A)$, to yield information on the entries in $(I - cA_n)^{-1}\mathbf{1}$. Note that $\Delta(A)$ is formed from the original webgraph D by taking each vertex of outdegree 0 and adding all possible outarcs from it.

Useful Facts:
1. $(I - cA_n)^{-1} = \sum_{k=0}^{\infty} c^k A_n^k$.
2. $e_i^T A_n^k \mathbf{1} = 1$ iff every walk of length k in $\Delta(A)$ that starts at vertex i must avoid vertex n.
3. $||(I - cA_n)^{-1}||_\infty \leq \frac{1}{1-c}$, with equality iff there is a vertex i in $\Delta(A)$ having no path to vertex n.

Note that Useful Fact 3 allows us to bound the numerator of $\frac{e_i^T(I-cA_n)^{-1}\mathbf{1}}{2(1+ca^T(I-cA_n)^{-1}\mathbf{1})}$, so a bound on the denominator will be enough to yield a bound on $\kappa_n(M)$.

Lemma 2: Suppose that n is on a cycle of length at least 2 in $\Delta(A)$, and that g is the length of a shortest such cycle. Suppose that A is partitioned as in (3). Then $a^T(I-cA_n)^{-1}\mathbf{1} \geq a^T\mathbf{1}\frac{1-c^{g-1}}{1-c}$. Equality holds if and only if there is a stochastic principal submatrix of A having the form

$$S = \left[\begin{array}{c|c|c|c|c} 0 & S_{g-1} & \dots & 0 & 0 \\ \hline 0 & 0 & S_{g-2} & \dots & 0 \\ \hline & & & \ddots & \vdots \\ \hline 0 & 0 & \dots & 0 & \mathbf{1} \\ \hline b^T & 0 & \dots & 0 & 1-b^T\mathbf{1} \end{array}\right], \quad (4)$$

where the last row and column of S corresponds to vertex n in $\Delta(A)$.

Idea: Apply Useful Facts 1 and 2, and the definition of g.

Theorem 3: a) Suppose that vertex j is on a cycle of length at least 2 in $\Delta(A)$, and let g be the length of a shortest such cycle. Then $\kappa_j(M) \leq \frac{1}{2(1-c^g-ca_{jj}(1-c^{g-1}))}$. Equality holds if and only if there is some i such that there is no path from vertex i to vertex j in $\Delta(A)$, and there is a principal submatrix of A of the form (4), where the last row and column corresponds to index j.

b) If vertex j is on no cycle of length at least 2 in $\Delta(A)$ and $a_{jj} \neq 1$, then $\kappa_j(M) = \frac{1}{2(1-ca_{jj})}$.

c) If $a_{jj} = 1$, then $\kappa_j(M) \leq \frac{1}{2(1-c)}$, with equality if and only if there is a vertex i such that there is no path from vertex i to vertex j in $\Delta(A)$.

Upshot:

Corollary 1: a) If j is on a cycle of length at least 2 and g is the length of the shortest such cycle, then $|p_j - \pi_j| \leq \frac{||r^T||_1}{2(1-c^g-ca_{jj}(1-c^{g-1}))}$.

b) Suppose that vertex j is on no cycle of length 2 or more in $\Delta(A)$. Then $|p_j - \pi_j| \leq \frac{||r^T||_1}{2(1-ca_{jj})}$.

Notes:

1. Observe that the upper bound of Theorem 3 a) on κ_j is readily seen to be decreasing in g. We can interpret this bound as implying that if vertex j of $\Delta(A)$ is only on long cycles, then π_j will exhibit good conditioning properties.

2. The upper bounds of Theorem 3 a) and b) are increasing in a_{jj}. Note that in the context of the Google matrix, either $a_{jj} = 0$, or the j^{th} row of A is $\frac{1}{n}\mathbf{1}^T$.

3. Suppose that $c = .85$ and $a_{jj} = 0$. Then for $g = 2, 3, 4, 5$, the bounds in a) are $1.802, 1.296.1.046, 0.899$, respectively.

Question: What happens for an index corresponding to a row of M that is equal to $\frac{1}{n}\mathbf{1}^T$?

Note: There is evidence to suggest that the number of such rows may be large compared to n. A 2001 web crawl of 290 million pages produced roughly 220 million pages with no outlinks.

Corollary 2: Suppose that A has $m \geq 2$ rows equal to $\frac{1}{n}\mathbf{1}^T$, and that row j is one of those rows. Then $\kappa_j(M) \leq \frac{n-c(m-1)}{2((1-c^2)n-c(1-c)m)}$.

Idea: Partitioning out the $m-1$ rows of A_j equal to $\frac{1}{n}\mathbf{1}^T$, one can show that $\mathbf{1}^T(I - cA_j)^{-1}\mathbf{1} \geq \frac{n(n-1)}{n-c(m-1)}$. We then use that to get a bound on the denominator of the expression for $\kappa_j(M)$.

Notes: Suppose that A has m rows that are equal to $\frac{1}{n}\mathbf{1}^T$, and let $\mu = m/n$. For large values of n, we see that if $\mu > 0$, then the upper bound of Corollary 2 is roughly $\frac{1-c\mu}{2(1-c)(1+c-c\mu)}$, which is readily seen to be decreasing in μ. So, if the number of vertices of the original webgraph D having outdegree zero is large, the corresponding entries in π will exhibit good conditioning properties.

For instance if $c = .85$ and $\mu = \frac{22}{29}$, the bound of our Corollary 2 is approximately .9824.

We can apply the results above to address Question 2.

Corollary 3: a) Suppose that vertices i and j of $\Delta(A)$ are on cycles of length two or more, and let g_i and g_j denote the lengths of the shortest such cycles, respectively. If $p_i \geq p_j + ||r^T||_1 \left(\frac{1}{2(1-c^{g_i}-ca_{ii}(1-c^{g_i-1}))} + \frac{1}{2(1-c^{g_j}-ca_{jj}(1-c^{g_j-1}))}\right)$, then $\pi_i \geq \pi_j$.
b) Suppose that vertex i of $\Delta(A)$ is on a cycle of length two or more, and let g_i denote the length of the shortest such cycle. Suppose that vertex j is on no cycle of length two or more. If $p_i \geq p_j + ||r^T||_1 \left(\frac{1}{2(1-c^{g_i}-ca_{ii}(1-c^{g_i-1}))} + \frac{1}{2(1-ca_{jj})}\right)$, then $\pi_i \geq \pi_j$.
c) Suppose that neither of vertices i and j of $\Delta(A)$ are on a cycle of length two or more. If $p_i \geq p_j + ||r^T||_1 \left(\frac{1}{2(1-ca_{ii})} + \frac{1}{2(1-ca_{jj})}\right)$, then $\pi_i \geq \pi_j$.

Corollary 4: Suppose that A has $m \geq 2$ rows equal to $\frac{1}{n}\mathbf{1}^T$, one of which is row j.

a) Suppose that vertex i of $\Delta(A)$ is on a cycle of length two or more, and let g_i be the length of a shortest such cycle. If $p_i \geq p_j + ||r^T||_1 \left(\frac{1}{2(1-c^{g_i}-ca_{ii}(1-c^{g_i-1}))} + \frac{n-c(m-1)}{2((1-c^2)n-c(1-c)m)} \right)$, then $\pi_i \geq \pi_j$.

b) Suppose that vertex i is on no cycle of length two or more. If $p_i \geq p_j + ||r^T||_1 \left(\frac{1}{2(1-ca_{ii})} + \frac{n-c(m-1)}{2((1-c^2)n-c(1-c)m)} \right)$, then $\pi_i \geq \pi_j$.

c) Suppose that row i of A is equal to $\frac{1}{n}\mathbf{1}^T$. If $p_i \geq p_j + ||r^T||_1 \left(\frac{n-c(m-1)}{((1-c^2)n-c(1-c)m)} \right)$, then $\pi_i \geq \pi_j$.

Google has reported using the power method to estimate π^T. Suppose that $x(0)^T \geq 0^T$, with $x(0)^T \mathbf{1} = 1$, and that for each $k \in I\!N, x(k)^T$ is the $k-$th vector in the sequence of iterates generated by applying the power method to $x(0)^T$ with the matrix M.

Corollary 5: a) If vertex j is on no cycle of length at least 2 in $\Delta(A)$, then for each $k \in I\!N, |x(k)^T e_j - \pi_j| \leq \frac{c^k||\{x(1)^T - x(0)^T\}A^k||_1}{2(1-ca_{jj})} \leq \frac{c^k||x(1)^T - x(0)^T||_1}{2(1-ca_{jj})}$.
b) If vertex j is on a cycle of length at least 2 and g is the length of the shortest such cycle, then for each $k \in I\!N, |x(k)^T e_j - \pi_j| \leq \frac{c^k||\{x(1)^T - x(0)^T\}A^k||_1}{2(1-c^g-ca_{jj}(1-c^{g-1}))} \leq \frac{c^k||x(1)^T - x(0)^T||_1}{2(1-c^g-ca_{jj}(1-c^{g-1}))}$.
c) If row j of A is equal to $\frac{1}{n}\mathbf{1}^T$, and there are m such rows, then for each $k \in I\!N$,
$|x(k)^T e_j - \pi_j| \leq \frac{c^k(n-c(m-1))||\{x(1)^T - x(0)^T\}A^k||_1}{2((1-c^2)n-c(1-c)m)}$
$\leq \frac{c^k(n-c(m-1))||x(1)^T - x(0)^T||_1}{2((1-c^2)n-c(1-c)m)}$.

Relative Error Bounds

So far, we have considered the absolute error $|p_j - \pi_j|$, but how about the corresponding relative error $\frac{|p_j - \pi_j|}{\pi_j}$?

We have $\frac{|p_j - \pi_j|}{\pi_j} \le \frac{||r^T||_1}{2}||(I - M_j)^{-1}||_\infty$, so a bound on $||(I - M_j)^{-1}||_\infty$ will lead to a corresponding bound on the relative error.

Some Notation: Let $\widehat{S}$ be the set of vertices in $\Delta(A)$ for which there is no path to vertex n. For each vertex $j \notin \widehat{S}$, let $d(j,n)$ be the distance from vertex j to vertex n, and let $d = max\{d(j,n)|j \notin \widehat{S}\}$. For each $i = 0, \ldots, d$, let $S_i = \{j \notin \widehat{S}|d(j,n) = i\}$ (evidently $S_0 = \{n\}$ here). Suppose also that $\overline{v}^T$ is partitioned accordingly into subvectors $\overline{v_i}^T, i = 0, \ldots, d$, and $\widehat{v}^T$. Finally, for each $i = 1, \ldots, d$, let α_i be the minimum row sum of $A[S_i, S_{i-1}]$, the submatrix of A on rows S_i and columns S_{i-1}.

Theorem 4: We have

$$\kappa_n(M) \le \frac{\pi_n}{2(1-c)(v_n + \sum_{i=1}^d c^i \alpha_1 \dots \alpha_i \overline{v_i}^T \mathbf{1})},$$

so that in particular,

$$\frac{|p_n - \pi_n|}{\pi_n} \le \frac{||r^T||_1}{2(1-c)(v_n + \sum_{i=1}^d c^i \alpha_1 \dots \alpha_i \overline{v_i}^T \mathbf{1})}.$$

If $\widehat{S} \ne \emptyset$, then

$$\frac{\pi_n}{2(1-c)(v_n + \sum_{i=1}^d c^i \overline{v_i}^T \mathbf{1})} \le \kappa_n(M).$$

In particular, for each $\epsilon > 0$, there is a positive vector p^T whose entries sum to 1 such that $||r^T||_1 = \epsilon$ and

$$\frac{|p_n - \pi_n|}{\pi_n} \ge \frac{||r^T||_1}{2(1-c)(v_n + \sum_{i=1}^d c^i \overline{v_i}^T \mathbf{1})}.$$

Note: From the Theorem 4, we see that the vector v^T is influential on the relative conditioning of π_n. Specifically, if v^T places more weight on vertices in S_i for small values of i (i.e. on vertices whose distance to vertex n is short), then that has the effect of improving the relative conditioning properties of π_n.

We treat the situation of an index corresponding to a row of A that is equal to $\frac{1}{n}\mathbf{1}^T$ as a special case.

Notation: Suppose that row n of A is $\frac{1}{n}\mathbf{1}^T$. Let u_1^T be the subvector of v^T corresponding to rows of A not equal to $\frac{1}{n}\mathbf{1}^T$, and let u_2^T be the subvector of v^T corresponding to rows of A equal to $\frac{1}{n}\mathbf{1}^T$ and distinct from n.

Theorem 5: Suppose that A has m rows equal to $\frac{1}{n}\mathbf{1}^T$, one of which is row n. Then

$$\kappa_n(M) \leq \pi_n \frac{n - c(m-1)}{2(1-c)(v_n(n - c(m-1)) + c{u_2}^T\mathbf{1})}.$$

In particular,

$$\frac{|p_n - \pi_n|}{\pi_n} \leq \frac{(n - c(m-1))||r^T||_1}{2(1-c)(v_n(n - c(m-1)) + c{u_2}^T\mathbf{1})}.$$

Note: We note that in the case that $v^T = \frac{1}{n}\mathbf{1}^T$ and $\frac{m}{n} = \mu$, we find that the upper bound of the Theorem 5 on is roughly $\frac{|p_n - \pi_n|}{\pi_n} \leq \frac{n(1-c\mu)}{2(1-c)}$. Evidently the upper bound is decreasing in μ in this case.