The Second Eigenvalue of the Google Matrix

Taher H. Haveliwala and Sepandar D. Kamvar

Stanford University
{taherh,sdkamvar}@cs.stanford.edu

Abstract. We determine analytically the modulus of the second eigenvalue for the web hyperlink matrix used by Google for computing PageRank. Specifically, we prove the following statement:
"For any matrix $A = [cP + (1-c)E]^T$, where P is an $n \times n$ row-stochastic matrix, E is a nonnegative $n \times n$ rank-one row-stochastic matrix, and $0 \leq c \leq 1$, the second eigenvalue of A has modulus $|\lambda_2| \leq c$. Furthermore, if P has at least two irreducible closed subsets, the second eigenvalue $\lambda_2 = c$."
This statement has implications for the convergence rate of the standard PageRank algorithm as the web scales, for the stability of PageRank to perturbations to the link structure of the web, for the detection of Google spammers, and for the design of algorithms to speed up PageRank.

1 Theorem

Theorem 1. *Let P be an $n \times n$ row-stochastic matrix. Let c be a real number such that $0 \leq c \leq 1$. Let E be the $n \times n$ rank-one row-stochastic matrix $E = \boldsymbol{e}\boldsymbol{v}^T$, where $\boldsymbol{e}$ is the n-vector whose elements are all $e_i = 1$, and $\boldsymbol{v}$ is an n-vector that represents a probability distribution*[1].
Define the matrix $A = [cP + (1-c)E]^T$. Its second eigenvalue $|\lambda_2| \leq c$.

Theorem 2. *Further, if P has at least two irreducible closed subsets (which is the case for the web hyperlink matrix), then the second eigenvalue of A is given by $\lambda_2 = c$.*

2 Notation and Preliminaries

P is an $n \times n$ row-stochastic matrix. E is the $n \times n$ rank-one row-stochastic matrix $E = \boldsymbol{e}\boldsymbol{v}^T$, where $\boldsymbol{e}$ to be the n-vector whose elements are all $e_i = 1$. A is the $n \times n$ column-stochastic matrix:

$$A = [cP + (1-c)E]^T \tag{1}$$

We denote the ith eigenvalue of A as λ_i, and the corresponding eigenvector as $\boldsymbol{x_i}$.

$$A\boldsymbol{x_i} = \lambda_i \boldsymbol{x_i} \tag{2}$$

By convention, we choose eigenvectors $\boldsymbol{x_i}$ such that $||\boldsymbol{x_i}||_1 = 1$. Since A is column-stochastic, $\lambda_1 = 1, 1 \geq |\lambda_2| \geq \ldots \geq |\lambda_n| \geq 0$.

[1] i.e., a vector whose elements are nonnegative and whose L_1 norm is 1.

We denote the ith eigenvalue of P^T as γ_i, and its corresponding eigenvector as $\boldsymbol{y_i}$: $P^T\boldsymbol{y_i} = \gamma_i\boldsymbol{y_i}$. Since P^T is column-stochastic, $\gamma_1 = 1, 1 \geq |\gamma_2| \geq \ldots \geq |\gamma_n| \geq 0$.

We denote the ith eigenvalue of E^T as μ_i, and its corresponding eigenvector as $\boldsymbol{z_i}$: $E^T\boldsymbol{z_i} = \mu_i\boldsymbol{z_i}$. Since E^T is rank-one and column-stochastic, $\mu_1 = 1, \mu_2 = \ldots = \mu_n = 0$.

An $n \times n$ row-stochastic matrix M can be viewed as the transition matrix for a Markov chain with n states.
For any row-stochastic matrix M, $M\boldsymbol{e} = \boldsymbol{e}$.
A set of states S is a *closed subset* of the Markov chain corresponding to M if and only if $i \in S$ and $j \notin S$ implies that $M_{ij} = 0$.
A set of states S is an *irreducible closed subset* of the Markov chain corresponding to M if and only if S is a closed subset, and no proper subset of S is a closed subset.
Intuitively speaking, each irreducible closed subset of a Markov chain corresponds to a leaf node in the strongly connected component (SCC) graph of the directed graph induced by the nonzero transitions in the chain.
Note that E, P, and A^T are row stochastic, and can thus be viewed as transition matrices of Markov chains.

3 Proof of Theorem 1

We first show that Theorem 1 is true for $c = 0$ and $c = 1$.

CASE 1: $c = 0$
If $c = 0$, then, from equation 1, $A = E^T$. Since E is a rank-one matrix, $\lambda_2 = 0$. Thus, Theorem 1 is proved for c=0.

CASE 2: $c = 1$
If $c = 1$, then, from equation 1, $A = P^T$. Since P^T is a column-stochastic matrix, $|\lambda_2| \leq 1$. Thus, Theorem 1 is proved for c=1.

CASE 3: $0 < c < 1$
We prove this case via a series of lemmas.

Lemma 1. The second eigenvalue of A has modulus $|\lambda_2| < 1$.

Proof. Consider the Markov chain corresponding to A^T. If the Markov chain corresponding to A^T has only one irreducible closed subchain S, and if S is aperiodic, then the chain corresponding to A^T must have a unique eigenvector with eigenvalue 1, by the Ergodic Theorem [3]. So we simply must show that the Markov chain corresponding to A^T has a single irreducible closed subchain S, and that this subchain is aperiodic. Lemma 1.1 shows that A^T has a single irreducible closed subchain S, and Lemma 1.2 shows this subchain is aperiodic.

Lemma 1.1 There exists a unique irreducible closed subset S of the Markov chain corresponding to A^T.
Proof. We split this proof into a proof of existence and a proof of uniqueness.
Existence. Let the set U be the states with nonzero components in $\boldsymbol{v}$. Let S consist of the set of all states reachable from U along nonzero transitions in the chain. S trivially forms a closed subset. Further, since every state has a transition to U, no subset of S can be closed. Therefore, S forms an irreducible closed subset.
Uniqueness. Every closed subset must contain U, and every closed subset containing U must contain S. Therefore, S must be the unique irreducible closed subset of the chain.

Lemma 1.2 The unique irreducible closed subset S is an aperiodic subchain.
Proof. From Theorem 5 in the Appendix, all members in an irreducible closed subset have the same period. Therefore, if at least one state in S has a self-transition, then the subset S is aperiodic. Let u be any state in U. By construction, there exists a self-transition from u to itself. Therefore, S must be aperiodic.

From Lemmas 1.1 and 1.2, and the Ergodic Theorem, $|\lambda_2| < 1$ and Lemma 1 is proved.

Lemma 2. *The second eigenvector $\boldsymbol{x_2}$ of A is orthogonal to e: $\boldsymbol{e}^T\boldsymbol{x_2} = 0$.*

Proof. Since $|\lambda_2| < |\lambda_1|$ (by Lemma 1), the second eigenvector $\boldsymbol{x_2}$ of A is orthogonal to the first eigenvector of A^T by Theorem 3 in the Appendix. From Section 2, the first eigenvector of A^T is $\boldsymbol{e}$. Therefore, x_2 is orthogonal to $\boldsymbol{e}$.

Lemma 3. $E^T\boldsymbol{x_2} = 0$

Proof. By definition, $E = \boldsymbol{e}\boldsymbol{v}^T$, and $E^T = \boldsymbol{v}\boldsymbol{e}^T$. Thus, $E^T\boldsymbol{x_2} = \boldsymbol{v}\boldsymbol{e}^T\boldsymbol{x_2}$. From Lemma 2, $\boldsymbol{e}^T\boldsymbol{x_2} = 0$. Therefore, $E^T\boldsymbol{x_2} = 0$.

Lemma 4. The second eigenvector $\boldsymbol{x_2}$ of A must be an eigenvector $\boldsymbol{y_i}$ of P^T, and the corresponding eigenvalue is $\gamma_i = \lambda_2/c$.

Proof. From equation 1 and equation 2:

$$cP^T\boldsymbol{x_2} + (1-c)E^T\boldsymbol{x_2} = \lambda_2\boldsymbol{x_2} \tag{3}$$

From Lemma 3 and equation 3, we have:

$$cP^T\boldsymbol{x_2} = \lambda_2\boldsymbol{x_2} \tag{4}$$

We can divide through by c to get:

$$P^T\boldsymbol{x_2} = \frac{\lambda_2}{c}\boldsymbol{x_2} \tag{5}$$

If we let $\boldsymbol{y_i} = \boldsymbol{x_2}$ and $\gamma_i = \lambda_2/c$, we can rewrite equation 4.

$$P^T\boldsymbol{y_i} = \gamma_i\boldsymbol{y_i} \tag{6}$$

Therefore, $\boldsymbol{x_2}$ is also an eigenvector of P^T, and the relationship between the eigenvalues of A and P^T that correspond to $\boldsymbol{x_2}$ is given by:

$$\lambda_2 = c\gamma_i \tag{7}$$

Lemma 5. $|\lambda_2| \leq c$

Proof. We know from Lemma 4 that $\lambda_2 = c\gamma_i$. Because P is stochastic, we have that $|\gamma_i| \leq 1$. Therefore, $|\lambda_2| \leq c$, and Theorem 1 is proved.

4 Proof of Theorem 2

Recall that Theorem 2 states: If P has at least two irreducible closed subsets, $\lambda_2 = c$.

Proof.
CASE 1: $c = 0$
This is proven in Case 1 of Section 3.

CASE 2: $c = 1$
This is proven trivially from Theorem 3 in the Appendix.

CASE 3: $0 < c < 1$
We prove this as follows. We assume P has at least two irreducible closed subsets. We then construct a vector x_i that is an eigenvector of A and whose corresponding eigenvalue is $\lambda_i = c$. Therefore, $|\lambda_2| \geq c$, and there exists a $\lambda_i = c$. From Theorem 1, $\lambda_2 \leq c$. Therefore, if P has at least two irreducible closed subsets, $\lambda_2 = c$.

Lemma 1. *Any eigenvector y_i of P^T that is orthogonal to e is an eigenvector x_i of A. The relationship between eigenvalues is $\lambda_i = c\gamma_i$.*

Proof. It is given that $\boldsymbol{e}^T\boldsymbol{y_i} = 0$. Therefore,

$$E^T\boldsymbol{y_i} = \boldsymbol{v}\boldsymbol{e}^T\boldsymbol{y_i} = 0 \tag{8}$$

By definition,

$$P^T\boldsymbol{y_i} = \gamma_i\boldsymbol{y_i} \tag{9}$$

Therefore, from equations 1, 8, and 9,

$$A\boldsymbol{y_i} = cP^T\boldsymbol{y_i} + (1-c)E^T\boldsymbol{y_i} = cP^T\boldsymbol{y_i} = c\gamma_i\boldsymbol{y_i} \tag{10}$$

Therefore, $A\boldsymbol{y_i} = c\gamma_i\boldsymbol{y_i}$ and Lemma 1 is proved.

Lemma 2. There exists a $\lambda_i = c$.

Proof. We construct a vector $\boldsymbol{x_i}$ that is an eigenvector of P and is orthogonal to $\boldsymbol{e}$.

From Theorem 3 in the Appendix, the multiplicity of the eigenvalue 1 for P is equal to the number of irreducible closed subsets of P. Thus we can find two linearly independent eigenvectors $\boldsymbol{y_1}$ and $\boldsymbol{y_2}$ of P^T corresponding to the dominant eigenvalue 1. Let

$$k_1 = \boldsymbol{y_1}^T \boldsymbol{e} \tag{11}$$

$$k_2 = \boldsymbol{y_2}^T \boldsymbol{e} \tag{12}$$

If $k_1 = 0$, let $\boldsymbol{x_i} = \boldsymbol{y_1}$, else if $k_2 = 0$, let $\boldsymbol{x_i} = \boldsymbol{y_2}$. If $k_1, k_2 > 0$, then let $\boldsymbol{x_i} = \boldsymbol{y_1}/k_1 - \boldsymbol{y_2}/k_2$. Note that $\boldsymbol{x_i}$ is an eigenvector of P^T with eigenvalue exactly 1 and that x_i is orthogonal to e. From Lemma 1, $\boldsymbol{x_2}$ is an eigenvector of A corresponding to eigenvalue c. Therefore, the eigenvalue λ_i of A corresponding to eigenvector $\boldsymbol{x_i}$ is $\lambda_i = c$.

Therefore, $|\lambda_2| \geq c$, and there exists a $\lambda_i = c$. However, from Theorem 1, $\lambda_2 \leq c$. Therefore, $\lambda_2 = c$ and Theorem 2 is proved.[2]

5 Implications

The matrix A is used by Google to compute PageRank, an estimate of web-page importance used for ranking search results [11]. PageRank is defined as the stationary distribution of the Markov chain corresponding to the $n \times n$ stochastic transition matrix A^T. The matrix P corresponds to the web link graph; in making P stochastic, there are standard techniques for dealing with web pages with no outgoing links [6]. Furthermore, the web graph has been empirically shown to contain many irreducible closed subsets [1], so that Theorem 2 holds for the matrix A used by Google.
Theorem 1 has implications for the rate of convergence of PageRank, for the stability of PageRank to perturbations to the link structure, and for the design of algorithms to speed up PageRank computations. Furthermore, it has broader implications in areas ranging from graph partitioning to reputation schemes in peer-to-peer networks. We briefly discuss these implications in this section.

Convergence of PageRank. The PageRank algorithm uses the power method to compute the principal eigenvector of A. The rate of convergence of the power method is given by $\frac{|\lambda_2|}{|\lambda_1|}$ $[13, 2]$. For PageRank, the typical value of c has been given as 0.85; for this value of c, Theorem 2 thus implies that the convergence rate of the power method $|\lambda_2/\lambda_1|$ for any web link matrix A is 0.85. Therefore, the convergence rate of PageRank will be fast, even as the web scales.

Stability of PageRank to Perturbations in the Link Structure. The modulus of the nonprincipal eigenvalues also determines whether the corresponding Markov chain is well-conditioned. As shown by Meyer in [9], the greater the eigengap $|\lambda_1| - |\lambda_2|$, the more stable the stationary distribution is to perturbations in the Markov chain. Our analysis

[2] Note that there may be additional eigenvalues with modulus c, such as $-c$.

provides an alternate explanation for the stability of PageRank shown by Ng et al. [10].

Accelerating PageRank Computations. Previous work on accelerating PageRank computations assumed λ_2 was unknown [6]. By directly using the equality $\lambda_2 = c$, improved extrapolation techniques may be developed as in [6].

Spam Detection. The eigenvectors corresponding to the second eigenvalue $\lambda_2 = c$ are an artifact of certain structures in the web graph. In particular, each pair of leaf nodes in the SCC graph for the chain P corresponds to an eigenvector of A with eigenvalue c. These leaf nodes in the SCC are those subgraphs in the web link graph which may have incoming edges, but have no edges to other components. Link spammers often generate such structures in attempts to hoard rank. Analysis of the nonprincipal eigenvectors of A may lead to strategies for combating link spam.

Broader Implications. This proof has implication for spectral methods beyond web search. For example, in the field of peer-to-peer networks, the EigenTrust reputation algorithm given in [7] computes the principal eigenvector of a matrix of the form defined in equation 1. This result shows that EigenTrust will converge quickly, minimizing network overhead. In the field of image segmentation, Perona and Freeman [12] present an algorithm that segments an image by thresholding the first eigenvector of the affinity matrix of the image. One may normalize the affinity matrix to be stochastic as in [8] and introduce a regularization parameter as in [11] to define a matrix of the form given in equation 1. The benefit of this is that one can choose the regularization parameter c to be large enough so that the computation of the dominant eigenvector is very fast, allowing the Perona-Freeman algorithm to work for very large scale images.

Acknowledgments

We would like to thank Gene Golub and Rajeev Motwani for useful conversations.

This paper is based on work supported in part by the National Science Foundation under Grant No. IIS-0085896 and Grant No. CCR-9971010, and in part by the Research Collaboration between NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation and CSLI, Stanford University (research project on Concept Bases for Lexical Acquisition and Intelligently Reasoning with Meaning).

References

1. A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. In *Proceedings of the Ninth International World Wide Web Conference*, 2000.
2. G. H. Golub and C. F. V. Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 1996.
3. G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Oxford University Press, 1989.
4. M. Iosifescu. *Finite Markov Processes and Their Applications*. John Wiley and Sons, Inc., 1980.

5. D. L. Isaacson and R. W. Madsen. *Markov Chains: Theory and Applications*, chapter IV, pages 126–127. John Wiley and Sons, Inc., New York, 1976.
6. S. D. Kamvar, T. H. Haveliwala, C. D. Manning, and G. H. Golub. Extrapolation methods for accelerating PageRank computations. In *Proceedings of the Twelfth International World Wide Web Conference*, 2003.
7. S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The EigenTrust Algorithm for Reputation Management in P2P Networks. In *Proceedings of the Twelfth International World Wide Web Conference*, 2003.
8. M. Meila and J. Shi. A random walks view of spectral segmentation. In *AI and Statistics (AISTATS)*, 2001.
9. C. D. Meyer. Sensitivity of the stationary distribution of a Markov chain. *SIAM Journal on Matrix Analysis and Applications*, 15(3):715–728, 1994.
10. A. Y. Ng, A. X. Zheng, and M. I. Jordan. Link analysis, eigenvectors and stability. In *IJCAI*, pages 903–910, 2001.
11. L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: Bringing order to the web. *Stanford Digital Libraries Working Paper*, 1998.
12. P. Perona and W. Freeman. A factorization approach to grouping. In *Proc. EECV*, 1998.
13. J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Oxford University Press, Oxford, 1965.

Appendix

This appendix contains theorems that are proven elsewhere and are used in proving Theorems 1 and 2 of this paper.

Theorem 3. *(from page 126 of [5]) If P is the transition matrix for a finite Markov chain, then the multiplicity of the eigenvalue 1 is equal to the number of irreducible closed subsets of the chain.*

Theorem 4. *(from page 4 of [13]) If $\boldsymbol{x_i}$ is an eigenvector of A corresponding to the eigenvalue λ_i, and $\boldsymbol{y_j}$ is an eigenvector of A^T corresponding to λ_j, then $\boldsymbol{x_i}^T \boldsymbol{y_j} = 0$ (if $\lambda_i \neq \lambda_j$).*

Theorem 5. *(from page 82 of [4]) Two distinct states belonging to the same class (irreducible closed subset) have the same period. In other words, the property of having period d is a class property.*