

Eigenvalues of Exponentiated Adjacency Matrices

Tracy Powell

Lesley A. Ward, Advisor

Estelle Basor, Reader

November 21, 2005

HARVEY MUDD
COLLEGE

Department of Mathematics

Problem

A graph G is made up of vertices, or nodes, and edges connecting them. The corresponding adjacency matrix $A = A(G)$ is an $n \times n$ matrix (n being the number of nodes) where the a_{ij}^{th} entry is 1 if there is a directed edge, or link, from node i to node j and 0 otherwise. We will consider the matrix B defined by

$$B = e^A - I = A + A^2/2! + A^3/3! + \dots$$

In Farahat et al. (2005), it is proven that the matrices BB^T and B^TB have simple, or nonrepeated, dominant eigenvalues if the graph G is weakly connected. As n goes to infinity for a particular family of graphs (which will be discussed in further detail later on), it can be observed that the second leading eigenvalue of B , $\lambda_2 = \lambda_2(B)$, becomes increasingly close to the leading eigenvalue of B , $\lambda_1 = \lambda_1(B)$. If the ratio of λ_2 to λ_1 is not bounded above by a number less than one, then the simpleness of the dominant eigenvalue no longer holds, that is, “in the limit” the leading eigenvalue will be repeated.

There have been several papers written trying to find a “good” upper bound for leading eigenvalues of certain trees and graphs. Hofmeister finds upper bounds for the leading and second leading eigenvalues of an adjacency matrix based on the number of vertices of a tree in Hofmeister (1997). A lower bound on the leading eigenvalue and an upper bound on the second leading eigenvalue will put an upper bound on the ratio of the two. There are also several articles on improved bounds, bounds on the eigenvalues of other types of graphs (other than trees), Laplacian eigenvalues, etc. (See bibliography below.)

The particular family of trees that we consider in this paper are directed trees made up of one node of degree h that branches out to h nodes of degree 2. These nodes branch out to h more nodes of degree 2 l times then terminate with the last set of h nodes having b branches coming out of them. (A definition of the degree of a node can be found below.) An example from this family of trees is shown in Figure 1. The parameters h , l , and b are taken directly from the structure of the graph. Respectively, these parameters correspond to the number of handles, the length, and the number of bristles of the graph. This family of graphs will be denoted $G_{h,l,b}$. When no third parameter is given, that is, when $G_{h,l}$ is written, it is assumed that

$b = 1$.

For this family of trees, $G_{h,l,b}$, as $l \rightarrow \infty$, the leading and second leading eigenvalues of $B^T B$ seem to become increasingly close, that is, $\lambda_2/\lambda_1 \rightsquigarrow 1$. Specifically, numerical experiments Farahat et al. (2005) suggested that $\lambda_2/\lambda_1 \rightarrow 1$ for $b = 1, 2$, and 3 . The initial goal of this project is to determine whether λ_2/λ_1 does in fact have limit 1 , as $l \rightarrow \infty$, in these examples.

Applications

Internet web page search engines, such as Google and Teoma, use particular algorithms to rank web pages based on their relevancy to the search query. One specific search algorithm is Kleinberg's HITS (Hypertext Induced Topic Search) algorithm (used by Teoma and Ask Jeeves). This algorithm uses a form of the adjacency matrix of the graph of web pages, $A^T A$, to calculate the rankings of the pages in the graph. When the dominant eigenvalue of $A^T A$ is nonsingular, the ranking of these web pages is not unique. This implies that the same search query entered at different times, or on a different computer, could potentially produce web pages in different orders. To prevent the nonsingularity of the dominant eigenvalue, the matrix $B^T B$ can be used where B is as defined above.

The following is a more specific explanation of the HITS algorithm. A search term, or query, is fed into the algorithm by the user. HITS will get a set S of the pages that include the text of the query. This set is then enlarged to the set T by adding in pages that link to or are linked to by a page in S . Kleinberg (1999) calls S the root set and T the base set. Typically, T has 3000–5000 pages. (This is different from Google's PageRank algorithm in that PageRank first ranks all of the pages in its database, about 8 billion pages, instead of looking at a subset of the pages.) HITS will now rank the pages of T in the following way. The pages of T and the links between them are represented by a graph, G , where the pages in T are the nodes of G and the hyperlinks between the pages are the edges of G . The nodes are numbered and the authority and hub vectors are initialized uniformly. That is, the initial authority vector, \vec{a}_0 , and the initial hub vector, \vec{h}_0 , look

like

$$\vec{a}_0 = \vec{h}_0 = \begin{bmatrix} 1/\sqrt{n} \\ 1/\sqrt{n} \\ \vdots \\ 1/\sqrt{n} \end{bmatrix}$$

where n is the number of nodes of G . Then the HITS algorithm updates the authority vector such that, for node i ,

$$\tilde{a}_1(i) = \sum_{j:j \rightarrow i} \vec{h}_0(j)$$

where $j : j \rightarrow i$ means that node j points directly to node i . The hub vector is updated similarly,

$$\tilde{h}_1(i) = \sum_{j:i \rightarrow j} \tilde{a}_1(j).$$

Both vectors are normalized so that

$$\vec{a}_1 = \frac{\tilde{a}_1}{\|\tilde{a}_1\|} \quad \text{and} \quad \vec{h}_1 = \frac{\tilde{h}_1}{\|\tilde{h}_1\|}.$$

This iteration is repeated until $\vec{h}_n = \vec{h}_{n+1}$ and $\vec{a}_n = \vec{a}_{n+1}$. This sequence of computation is then $a_0 = h_0, a_1, h_1, a_2, h_2, a_3, \dots$. This implies that in general, at the k^{th} iteration,

$$\vec{a}_k = \phi_k A^T \vec{h}_{k-1} \tag{1}$$

$$\vec{h}_k = \psi_k A \vec{a}_k \tag{2}$$

where $\phi_k, \psi_k \in \mathbb{R}^{>0}$ are normalization constants and

$$\sum_{i=1}^n \vec{a}_k(i)^2 = \sum_{i=1}^n \vec{h}_k(i)^2 = 1.$$

Focusing on the authority vector, we can rewrite Equation (1) and Equation (2) to show

$$\vec{a}_k = \phi_k \psi_{k-1} A^T A \vec{a}_{k-1}. \tag{3}$$

Since $A^T A$ is a symmetric matrix ($(A^T A)^T = A^T (A^T)^T = A^T A$), the following theorems apply. (The proofs of these theorems can be found in Poole (2003).)

Theorem 1 (The Spectral Theorem). *Let A be an $n \times n$ matrix. Then A is symmetric if and only if it is orthogonally diagonalizable. In particular, a symmetric matrix A is diagonalizable.*

Theorem 2. *Let A be an $n \times n$ diagonalizable matrix with dominant eigenvalue λ_1 . Then there exists a nonzero vector \vec{x}_0 such that the sequence of vectors \vec{x}_k defined by*

$$\vec{x}_1 = A\vec{x}_0, \vec{x}_2 = A\vec{x}_1, \vec{x}_3 = A\vec{x}_2, \dots, \vec{x}_k = A\vec{x}_{k-1}, \dots$$

approaches a dominant eigenvector of A .

Thus, the sequence determined by Equation (3) converges to a dominant eigenvector of $A^T A$. A similar conclusion can be made for the hub vectors such that the sequence of hub vectors will converge to a dominant eigenvector of AA^T . This iterative method is called the Power Method, and this form of the power method uses Rayleigh quotients (Poole, 2003: page 312).

Notice that Equation (3) converges to an eigenvector of the dominant eigenvalue of $A^T A$. If the dominant eigenvalue is repeated, there can be multiple eigenvectors associated with this eigenvalue. The Power Method will then converge to one of these dominant eigenvectors, but the one it converges to depends on the initial vector \vec{a}_0 . Therefore, for internet search rankings to be unique, it is imperative that the dominant eigenvalue of the matrix used in the iteration is simple.

Previous Research

Professor Estelle Basor from California State Polytechnic San Luis Obispo has proven that, for a graph with two handles and one bristle, the ratio of the second largest eigenvalue to the largest eigenvalue is bounded above by 0.94... She uses theory and techniques outlined in Böttcher and Grudsky (2000).

In Hofmeister (1997), Hofmeister proves the following theorem.

Theorem 3. *Let T be a tree with n vertices with $n \geq 4$ and $T \neq S_n$, the star with n vertices. Then*

$$\lambda_1 \leq \sqrt{\frac{1}{2} \left(n - 1 + \sqrt{n^2 - 6n + 13} \right)}.$$

This theorem, however, is applicable only to the dominant eigenvalue of the adjacency matrix A and not to that of $B = e^A - I$. Note that this bound is dependent upon the number of vertices n . As we let $n \rightarrow \infty$, this bound will also increase to infinity and be of little use.

Definitions

Here are some of the key definitions we will need for our project. The following are from graph theory.

Definition 1. A *directed graph*, or *digraph*, is a graph with directed edges. A *simple directed graph* is a directed graph having no repeated edges.

Definition 2. In a directed graph, the *in-degree* of a vertex v is the number of edges pointing to v . The *out-degree* of a vertex v is the number of edges pointing away from v . The *degree* (or *total degree*) of a vertex v is the sum of the in-degree and out-degree of v .

Definition 3. A directed graph is said to be *weakly connected* if it is possible to reach any node starting from any other node by traversing edges in some direction (not necessarily in the direction they point). The nodes in a weakly connected digraph (or directed graph) therefore must all have either outdegree or indegree of at least 1.

Definition 4. The *adjacency matrix* of a simple directed graph is a matrix A where the a_{ij} th entry is 1 if there is a path from vertex i to vertex j and 0 otherwise.

The following definitions are from linear algebra.

Definition 5. A matrix A is *symmetric* if $A = A^T$.

Definition 6. The *dominant* or *leading eigenvalue* of a matrix is the largest eigenvalue. Similarly, the *dominant eigenvector* is the eigenvector corresponding to the dominant eigenvalue.

Definition 7. A *simple eigenvalue* is one which is not repeated.

Definition 8. A *nilpotent* matrix is a square matrix A such that A^n is the zero matrix for some positive integer n .

Definition 9. A square matrix A is called *reducible* if, subject to some permutation of the rows and the same permutation of columns, A can be written in block form as

$$\begin{bmatrix} B & C \\ O & D \end{bmatrix}$$

where B and D are square. A square matrix that is not reducible is called *irreducible*.

Definition 10. Given $2n-1$ numbers a_k , where $k = -n+1, \dots, -1, 0, 1, \dots, n-1$, a *Toeplitz matrix* is a square matrix which has constant values along negative-sloping diagonals, i.e., a matrix of the form

$$\begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & a_{-n+1} \\ a_1 & a_0 & a_{-1} & & \\ a_2 & a_1 & a_0 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{-1} \\ a_{n-1} & \cdots & a_1 & a_0 & \end{bmatrix}.$$

Definition 11. An $n \times n$ matrix A is *primitive* if there exists some k such that $(A^k)_{i,j} > 0$ for all i, j .

Examples

The following is an example of a tree within the family of trees under consideration. This example gives an adjacency matrix that is 9×9 , since there are $n = 9$ nodes.

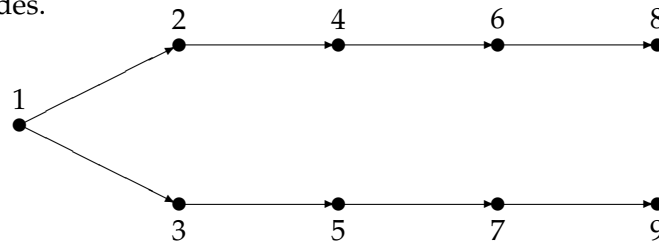
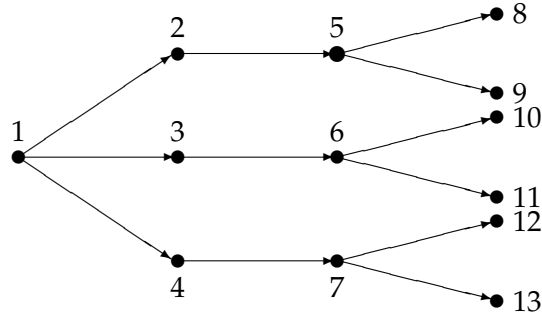


Figure 1: $G_{2,3}$

This example is the graph with two handles, $h = 2$, length $l = 3$, and one bristle, $b = 1$. Notice that the graph is denoted $G_{2,3}$ as opposed to $G_{2,3,1}$ as described above. The associated adjacency and exponentiated adjacency matrices are

Figure 2: $G_{3,2,2}$

$$B = \begin{bmatrix} 0 & 1 & 1 & 1 & 1/2 & 1/2 & 1/2 & 1/3! & 1/3! & 1/3! & 1/3! & 1/3! & 1/3! \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Here, the eigenvalues of $B^T B$ are 4.7014, 2.7808, 2.7808, 2.1105, 0.7192, 0.7192, 0.6047, and 0 (with algebraic multiplicity six). The authority and hub vectors are

$$\vec{a} = \vec{v}_1(B^T B) = \begin{bmatrix} 0 \\ 0.3915 \\ 0.3915 \\ 0.3915 \\ 0.3050 \\ 0.3050 \\ 0.3050 \\ 0.2086 \\ 0.2086 \\ 0.2086 \\ 0.2086 \\ 0.2086 \\ 0.2086 \\ 0.2086 \end{bmatrix} \quad \vec{h} = \vec{v}_1(BB^T) = \begin{bmatrix} 0.8489 \\ 0.2369 \\ 0.2369 \\ 0.2369 \\ 0.1924 \\ 0.1924 \\ 0.1924 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Again, the best authorities are nodes 2, 3, and 4 and the best hub is node 1. The ratio of the second largest eigenvalue to the largest eigenvalue, denoted $\alpha(h, l, b)$, is $\alpha(3, 2, 2) = 0.5915$.

The following table shows the ratios of $\lambda_2(B^T B)$ to $\lambda_1(B^T B)$ for the graphs $G_{h,l}$ where $b = 1$. It is interesting to note that the ratio tends to

decrease along a row and increase along a column. However, from $l = 2$ to $l = 9$, this is not the case as h increases from 1 to 2.

	$h = 1$	2	3	4	5	6
$l = 1$	0.37162	0.35961	0.25000	0.19098	0.15436	0.12948
2	0.46565	0.53828	0.38680	0.29919	0.24334	0.20488
3	0.56377	0.65641	0.48033	0.37338	0.30431	0.25648
4	0.64665	0.73118	0.53991	0.42036	0.34276	0.28895
5	0.71296	0.77936	0.57824	0.45040	0.36430	0.30965
6	0.76434	0.81223	0.60371	0.47030	0.38353	0.32333
7	0.80427	0.83501	0.62128	0.48399	0.39470	0.33275
8	0.83549	0.85146	0.63383	0.49377	0.40267	0.33947
9	0.86020	0.86368	0.64309	0.50099	0.40856	0.34443
10	0.87994	0.87292	0.65005	0.50641	0.41298	0.34816
11	0.89595	0.88013	0.65546	0.51062	0.41641	0.35106
12	0.90905	0.88580	0.65970	0.51393	0.41911	0.35333

Table 1: The ratios $\alpha(h, l) = \lambda_2/\lambda_1$ of $B^T B$ for graphs $G_{h,l}$.

Bibliography

Howard Anton and Robert Busby. *Contemporary Linear Algebra*. Wiley, Indianapolis, IN, 2002. ISBN 0471163627.

Section 5.4 goes into depth about the power method and its specific application to internet search algorithms.

Albrecht Böttcher and Sergei M. Grudsky. *Toeplitz matrices, asymptotic linear algebra, and functional analysis*. Birkhäuser Verlag, Basel, 2000. ISBN 3-7643-6290-1.

This book contains the substance of the proof by Professors Basor and Morrison that the ratio of the second largest to the largest eigenvalues of $B^T B$ is bounded below 1.

Wai-Shun Cheung, Chi-Kwong Li, D. D. Olesky, and P. van den Driessche. Optimizing quadratic forms of adjacency matrices of trees and related eigenvalue problems. *Linear Algebra Appl.*, 325(1-3):191–207, 2001. ISSN 0024-3795.

Proposition 1.5 seems to be the only item that has much to do with eigenvalues. The section on optimal labeling of trees might be interesting to consider since it is different from our current labeling.

Chris H. Q. Ding, Hongyuan Zha, Xiaofeng He, Parry Husbands, and Horst D. Simon. Link analysis: hubs and authorities on the World Wide Web. *SIAM Rev.*, 46(2):256–268 (electronic), 2004. ISSN 0036-1445.

Ayman Farahat, Thomas LoFaro, Joel C. Miller, Gregory Rae, and Lesley A. Ward. Authority rankings from HITS, PageRank, and SALSA: Existence, uniqueness, and effect of initialization. *To appear in the Journal of Scientific Computing, SIAM.*, 2005.

Provides a useful theorem about the simpleness of the dominant eigenvalue for the exponentiated adjacency matrix.

C. D. Godsil. Eigenvalues of graphs and digraphs. *Linear Algebra Appl.*, 46: 43–50, 1982. ISSN 0024-3795.

From the abstract, it does not look like this article gives any kind of bounds for eigenvalues of digraphs.

Ji-Ming Guo and Shang-Wang Tan. A note on the second largest eigenvalue of a tree with perfect matchings. *Linear Algebra Appl.*, 380:125–134, 2004. ISSN 0024-3795.

This paper finds an upper bound for the second largest eigenvalue of trees with $n = 2k = 4t$ ($t \geq 2$) nodes. This would only apply to particular trees within our family of trees.

M. Hofmeister. On the two largest eigenvalues of trees. *Linear Algebra Appl.*, 260:43–59, 1997. ISSN 0024-3795.

The introduction and the last theorem might be particularly helpful. The proofs and bit about partial eigenvectors might be of less use.

Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *J. ACM*, 46(5):604–632, 1999. ISSN 0004-5411.

The primary source for information on the HITS algorithm.

Mikhail Klin, Akihiro Munemasa, Mikhail Muzychuk, and Paul-Hermann Zieschang. Directed strongly regular graphs obtained from coherent algebras. *Linear Algebra Appl.*, 377:83–109, 2004. ISSN 0024-3795.

As can be seen from the title, this article has nothing to do with bounds for eigenvalues. However, it might make an interesting read some summer day.

Bo Lian Liu, Brendan D. McKay, Nicholas C. Wormald, and Ke Min Zhang. The exponent set of symmetric primitive $(0, 1)$ matrices with zero trace. *Linear Algebra Appl.*, 133:121–131, 1990. ISSN 0024-3795.

This article might contain information about the primitiveness of the adjacency matrices of our graphs.

David Poole. *Linear Algebra: A Modern Introduction*. Brooks/Cole, Pacific Grove, CA, 2003. ISBN 0-534-34174-8.

A linear algebra text book with useful definitions and theorems.

Dragan Stevanović. Bounding the largest eigenvalue of trees in terms of the largest vertex degree. *Linear Algebra Appl.*, 360:35–42, 2003. ISSN 0024-3795.

This article gives upper and lower bounds for the leading eigenvalue of the adjacency matrix and the Laplacian matrix of a graph G .

Hong Yuan. The k th largest eigenvalue of a tree. *Linear Algebra Appl.*, 73: 151–155, 1986. ISSN 0024-3795.

Yuan shows that, for the k th eigenvalue of a tree T with n vertices, $\lambda_k(T) \leq \sqrt{\lfloor (n-2)/k \rfloor}$ for $2 \leq k \leq \lfloor n/2 \rfloor$, where $\lfloor x \rfloor$ is the largest integer not greater than x .

Shuqin Zhao and Yuan Hong. On the bounds of maximal entries in the principal eigenvector of symmetric nonnegative matrix. *Linear Algebra Appl.*, 340:245–252, 2002. ISSN 0024-3795.

This article is not directly related to eigenvalues as much as it is related to the entries of dominant eigenvectors.