2006 Mathematics Clinic

HARVEY MUDD COLLEGE

Singular Values of the Exponentiated Adjacency Matrices of Broom-Tree Graphs

mathematics harvey mudd college

1 Broom-tree graphs

We begin with a broom-tree graph as in Figure 1. We denote these trees by $G_{h,\ell,b}$. Here h, ℓ, and b are positive integers. One node in $G_{h,\ell,b}$ has out-degree h and in-degree zero. Each of the h nodes it points to begins a chain of $\ell+1$ nodes each of in-degree and out-degree one. The ℓth node in each of these chains has out-degree b.

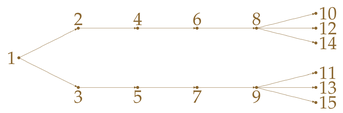


Figure 1: Graph $G_{h,\ell,b} = G_{2,4,3}$.

When viewing these graphs with the first node at the top and the h handles extending down, they look like brooms with multiple handles. This is where the name broom-tree graph comes from. The parameters h, ℓ, and b that are taken directly from the structure of the graph correspond to the number of handles, the length, and the number of bristles at the end of each handle of the graph, respectively. The above graph is denoted $G_{2,4,3}$ since there are $h = 2$ handles of length $\ell = 4$ including the $b = 3$ bristles at the end of each handle.

The adjacency matrix A for this graph is the $n \times n$ matrix, where $n = h(\ell+1)+b$ is the number of nodes of $G_{h,\ell,b}$, whose a_{ij}^{th} entry is one if node i points to node j and zero otherwise. For the above graph, the adjacency matrix is

$$A = \begin{bmatrix}
0&1&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&1&0&1&0\\
0&0&0&0&0&0&0&0&0&0&1&0&1&0&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{bmatrix}.$$

We define the exponentiated adjacency matrix B by

$$B = e^A - I = A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

For the graph in Figure 1, the exponentiated adjacency matrix is

$$B = \begin{bmatrix}
0&1&1&\frac{1}{2!}&\frac{1}{2!}&\frac{1}{3!}&\frac{1}{3!}&\frac{1}{3!}&\frac{1}{4!}&\frac{1}{4!}&\frac{1}{5!}&\frac{1}{5!}&\frac{1}{6!}&\frac{1}{6!}&\frac{1}{6!}&\frac{1}{6!}\\
0&0&0&1&0&\frac{1}{2!}&0&\frac{1}{3!}&0&\frac{1}{4!}&0&\frac{1}{4!}&0&\frac{1}{4!}&0&\frac{1}{4!}\\
0&0&0&0&1&0&\frac{1}{2!}&0&\frac{1}{3!}&0&\frac{1}{4!}&0&\frac{1}{4!}&0&\frac{1}{4!}&0\\
0&0&0&0&0&1&0&\frac{1}{2!}&0&\frac{1}{3!}&0&\frac{1}{3!}&0&\frac{1}{3!}&0&\frac{1}{3!}\\
0&0&0&0&0&0&1&0&\frac{1}{2!}&0&\frac{1}{3!}&0&\frac{1}{3!}&0&\frac{1}{3!}&0\\
0&0&0&0&0&0&0&1&0&\frac{1}{2!}&0&\frac{1}{2!}&0&\frac{1}{2!}&0&\frac{1}{2!}\\
0&0&0&0&0&0&0&0&1&0&\frac{1}{2!}&0&\frac{1}{2!}&0&\frac{1}{2!}&0\\
0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1\\
0&0&0&0&0&0&0&0&0&0&1&0&1&0&1&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{bmatrix}.$$

The singular values of an $n \times n$ matrix M are the square roots of the eigenvalues of M^TM. A number related to the i^{th} singular value of M is the $(i-1)^{\text{th}}$ approximation number of M. The j^{th} approximation number of M is the minimum distance between M and any $n \times n$ matrix with rank at most j. The operator norm of a linear operator $T : V \to W$ where V and W are normed vector spaces, is the largest value by which T stretches and element of V. When T is given by a matrix M, the operator norm $\|T\|$ of T is the largest singular value of M. It is then easy to see that

$$\sqrt{\lambda_1(M^TM)} = \sigma_1(M) = a_0(M) = \|M\| \tag{1}$$

where σ_1 is the largest singular value of M and a_0 is the largest approximation number of M.

For this family of trees $G_{h,\ell,b}$, as $\ell \to \infty$, the leading and second leading eigenvalues of B^TB seem to become increasingly close. Specifically, numerical experiments using Matlab [1] suggest that $\lambda_2/\lambda_1 \to 1$ for $b = 1$, 2, and 3. The results can be seen in Figure 2. The question addressed here is whether λ_2/λ_1 does in fact have limit 1, as $\ell \to \infty$, for subfamilies of these graphs.

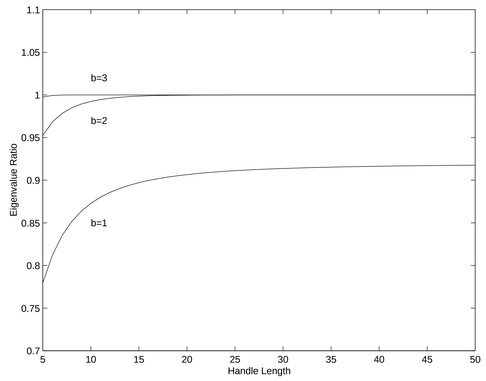
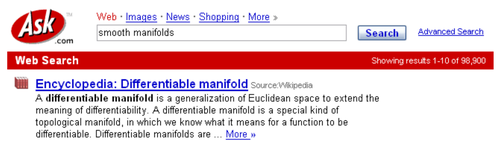


Figure 2: The ratio λ_2/λ_1 of B^TB for different values of b [1].

2 Application to Internet search



Internet search engines such as Google and Teoma use particular algorithms to rank web pages based on their relevance to a given search query. Teoma for instance, which has now merged with Ask.com, uses the Hypertext Induced Topic Search, or HITS, algorithm developed by Kleinberg [2]. The algorithm views the Internet as a graph where the web pages are nodes and the directed edges are links from one node to the other. Briefly stated, the HITS algorithm is an iteration that uses the symmetrized adjacency matrix A^TA of the graph of web pages to calculate the rankings of the pages in the graph. When the largest eigenvalue of A^TA is repeated, the ranking of these web pages is not unique [1]. 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 repetition of the dominant eigenvalue, the matrix B^TB can be used in place of A^TA, where B is as defined above.

To measure how close the second largest eigenvalue λ_2 of B^TB is to the largest eigenvalue λ_1 of B^TB, we look at the ratio $\lambda_2(B^TB)/\lambda_1(B^TB)$. The HITS algorithm ranks the pages included in the search based on the entries of the eigenvector of B^TB associated with the largest eigenvalue of B^TB. When $\lambda_1(B^TB)$ is repeated, there can be multiple linearly independent eigenvectors associated with this eigenvalue. The HITS algorithm will converge to an eigenvector in the eigenspace of $\lambda_1(B^TB)$. This eigenvector will be different depending on the initial vector used in the iteration. The rate of convergence of the HITS algorithm also depends on the ratio λ_2/λ_1 of A^TA or B^TB. The closer this ratio is to one, the slower the algorithm will converge, and vice versa. So, to have a fast algorithm, it would be good for this ratio to always be less than one.

3 Results on λ_2/λ_1 of B^TB

Figure 3: The graph $G_{h,\ell,1}$.

The following theorem was proven for broom-tree graphs with two handles, $h = 2$, by Estelle Basor and Kent Morrison of California Polytechnic State University, San Luis Obispo. I have modified the proof to extend to any $h \geqslant 2$.

Theorem 1. *Let $G_{h,\ell,1}$ be the directed graph shown in Figure 3. Also let $\lambda_j(M)$ denote the j^{th} largest eigenvalue of an $n \times n$ matrix M and let $B = e^A - I$ where A is the adjacency matrix corresponding to $G_{h,\ell,1}$. Then, for n sufficiently large,*

$$\frac{\lambda_2(B^TB)}{\lambda_1(B^TB)} < 0.94430808$$

for all natural numbers h and ℓ. In particular, for fixed h as $\ell \to \infty$, the ratio $\lambda_2(B^TB)/\lambda_1(B^TB)$ remains bounded above by a constant less than one.

Here is a sketch of the proof of this theorem. We begin by finding a lower bound for the largest singular value of B. Using Matlab, we find that $\|B\| = \sigma_1(B) > 1.77$ for all $G_{h,\ell,1}$ where $h \geqslant 2$ and $\ell \geqslant 3$. Therefore,

$$\lambda_1(B^TB) > \sqrt{1.77}.$$

To find an upper bound on the second largest singular value of B, we first note that we can write B as the sum of a rank one matrix and an augmented Toeplitz matrix C. A Toeplitz matrix T is a matrix with the same entries on negative sloping diagonals. Then

$$B = \begin{bmatrix}\text{rank one}\\ \text{matrix}\end{bmatrix} - \left[\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T & \\ 0 & & & \end{array}\right]$$

and so

$$B - \begin{bmatrix}\text{rank one}\\ \text{matrix}\end{bmatrix} = C.$$

So $\sigma_2(B) = a_1(B) \leqslant \|C\|$ by equation (1). Since the matrices C and T only differ by a row and column of zeros, $\|T\| = \|C\|$. From theory contained in [3], $\|T\| \leqslant e - 1$. Therefore,

$$\lambda_2(B^TB) < \sqrt{1.72}$$

and

$$\frac{\lambda_2(B^TB)}{\lambda_1(B^TB)} = \left(\frac{\sigma_2(B)}{\sigma_1(B)}\right)^2 < 0.944300808.$$

This completes the sketch of the proof.

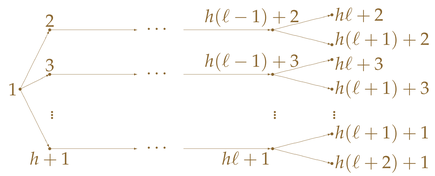


Figure 4: The graph $G_{h,\ell,2}$.

Another result comes from an analogous proof where we let $b = 2$.

Theorem 2. *Let $G_{h,\ell,2}$ be the directed graph shown in Figure 4. Also let $\lambda_i(M)$ be the ith largest eigenvalue of an $n \times n$ matrix M and let $B = e^A - I$ where A is the adjacency matrix of $G_{h,\ell,2}$. Then, for n sufficiently large,*

$$\frac{\lambda_{h+2}(B^TB)}{\lambda_1(B^TB)} < 0.85512776.$$

4 Open question

We still do not have an upper bound on the ratio of the second largest to the largest singular value of B for $b \geqslant 2$. The technique to find a lower bound for the largest singular value of B will be the same as before, using numerical calculations. However, the problem with the second part of the proof is that there is no way to write B as the sum of a rank one matrix and a Toeplitz matrix when the number of bristles is larger than one. Numerical methods or other linear algebra theory might help to put an upper bound on $\sigma_2(B)$ for $b \geqslant 2$.

5 Summary

It can be inferred from Figure 2 that the ratio of the largest two eigenvalues of B^TB should be bounded below one when there is only one bristle at the end of each handle. It is hard to tell, however, if that is the case for $b \geqslant 2$. We have shown the following results. First, for the broom-tree graphs $G_{h,\ell,1}$, with h handles, one bristle at the end of each handle, and length ℓ increasing to infinity, the ratio $\lambda_2(B^TB)/\lambda_1(B^TB)$ of the second to the first eigenvalues of the symmetrized exponentiated adjacency matrix B^TB remains bounded above by a constant less than one. Second, for the broom-tree graphs $G_{h,\ell,2}$, with h handles, two bristles at the end of each handle, and length ℓ increasing to infinity, the ratio $\lambda_{h+2}(B^TB)/\lambda_1(B^TB)$ of the $(h+2)$nd to the first eigenvalues of B^TB remains bounded below one. The first result guarantees a unique ranking of the nodes of $G_{h,\ell,1}$ by the HITS algorithm.

References

[1] 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, SIAM J. Sci. Comput. 27 (2006), no. 4, 1181-1201 (electronic), MR 2199745.

[2] Jon M. Kleinberg, *Authoritative sources in a hyperlinked environment*, J.ACM46 (1999), no. 5, 604-632, MR 1747649.

[3] Albrecht Böttcher and Sergei M. Grudsky, Toeplitz matrices, asymptotic linear algebra, and functional analysis, Birkhäuser Verlag, Basel, 2000, MR 1772773 (2001g:47043).

6 Acknowledgements

I would like to express my gratitude to my thesis advisor, Lesley Ward, for her constructive and motivational comments and my second reader, Estelle Basor, for writing the proof that contributed to the basis of my thesis. Many thanks to Claire Connelley for her support with LaTeX.

Copyright © 2005, Harvey Mudd College Department of Mathematics.