A Note to Panos:

Algebraic Connectivity of Perfect Dominance Graphs

Thomas R. Cameron

June 11, 2019

Abstract

This note is an introduction to a wonderful connection between the algebraic connectivity of a graph, the field of values of a matrix, and the rankability of data.

1 Introduction

Recently, the problem of rankability was posed, which refers to a dataset's inherent ability to produce a meaningful ranking of its items [1]. Given data that can be modeled as a binary directed graph, the idea is to measure how far that graph is from a *perfect dominance graph*, i.e., an acyclic tournament graph. A perfect dominance graph represents an ideal situation for ranking as it is associated with a clear and unique ranking, e.g., the perfect dominance graph associated with the ranking $[1, 2, 3, 4, 5, 6]$ is shown in Figure 1.

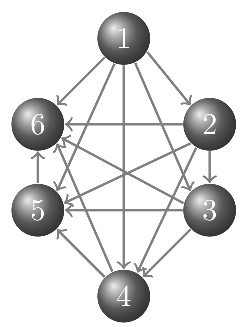

Figure 1: Perfect dominance graph associated with ranking $[1, 2, 3, 4, 5, 6]$.

Now, in theory, the rankability measure proposed in [1] is simple. Let k denote the minimum number of edge changes(additions or deletions) needed to obtain a perfect dominance graph. Given k edge changes, denote by p the number of perfect dominance graphs that can be obtained. Then, the rankability measure of the given dataset is defined by

$$\text{IPR} = 1 - \frac{kp}{k_{max}p_{max}}, \tag{1}$$

where $k_{max} = (n^2 - n)/2$ and $p_{max} = n!$. We denote this rankability measure by IPR since an integer program is used to compute k and p.

This rankability measure is clearly expensive to compute; in fact, it is NP Hard. Therefore, even more recently, we proposed a cheaper rankability measure, which is motivated by a spectral-degree characterization of perfect dominance graphs.

Theorem 1.1 (See Corollary 2.9 of Pre-Print)**.** *Let $\Gamma \in \mathbb{G}$ have binary weights and let L be the graph Laplacian of Γ. Then, Γ is a perfect dominance graph if and only if*

$$\sigma(L) = \left\{d^+(1), \ldots, d^+(n)\right\}$$

and there exists a re-ordering of the vertices such that $d^+(i) = n - i$ for $i = 1, \ldots, n$.

Armed with this spectral-degree characterization, we propose the following measure of the rankability of data. Note that $\mathrm{hd}(A, \tilde{A})$ denotes the Hausdorff distance between the eigenvalues of A and the eigenvalues of $\tilde{A}$.

denotes the Hausdorff distance and the maximum upper bound for both $\mathrm{hd}(D, S)$ and $\mathrm{hd}(L, S)$ is $(n-1)$. Therefore, the division by $2(n-1)$ is a normalization factor. Finally, note that r ranges between 0 and 1, which indicates a transition between ill-rankable and well-rankable data.

Algorithm 1 Spectral Rankability of Graph Data Γ.

function $[r] = \mathrm{SpecR}\,(\Gamma)$:
- $n \leftarrow$ the number of vertices in Γ
- $D \leftarrow$ the out-degree matrix of Γ
- $L \leftarrow$ graph Laplacian of Γ
- $S = \mathrm{diag}\,(n-1, n-2, \ldots, 0)$
- $r = 1 - \frac{\mathrm{hd}(D,S)+\mathrm{hd}(L,S)}{2(n-1)}$
- **return**

In Figure 2, we consider several examples from [1], which tests whether a rankability measure aligns with our intuitive classification of certain structured datasets as well-rankable or ill-rankable. For each dataset, we compare SpecR with the measure IPR as defined in (1).

Note that the graphs are displayed from most rankable to least rankable as determined by the measure IPR. The measures SpecR and IPR have a strong correlation, the Pearson correlation coefficient between them is 0.92. In addition, the measures SpecR and IPR have exact agreement on the extreme cases, i.e., the perfect dominance graph and empty graph. Finally, if the matching distance is used to measure the variation, then there is also exact agreement on the completely connected graph, and the Pearson coefficient is 0.94.

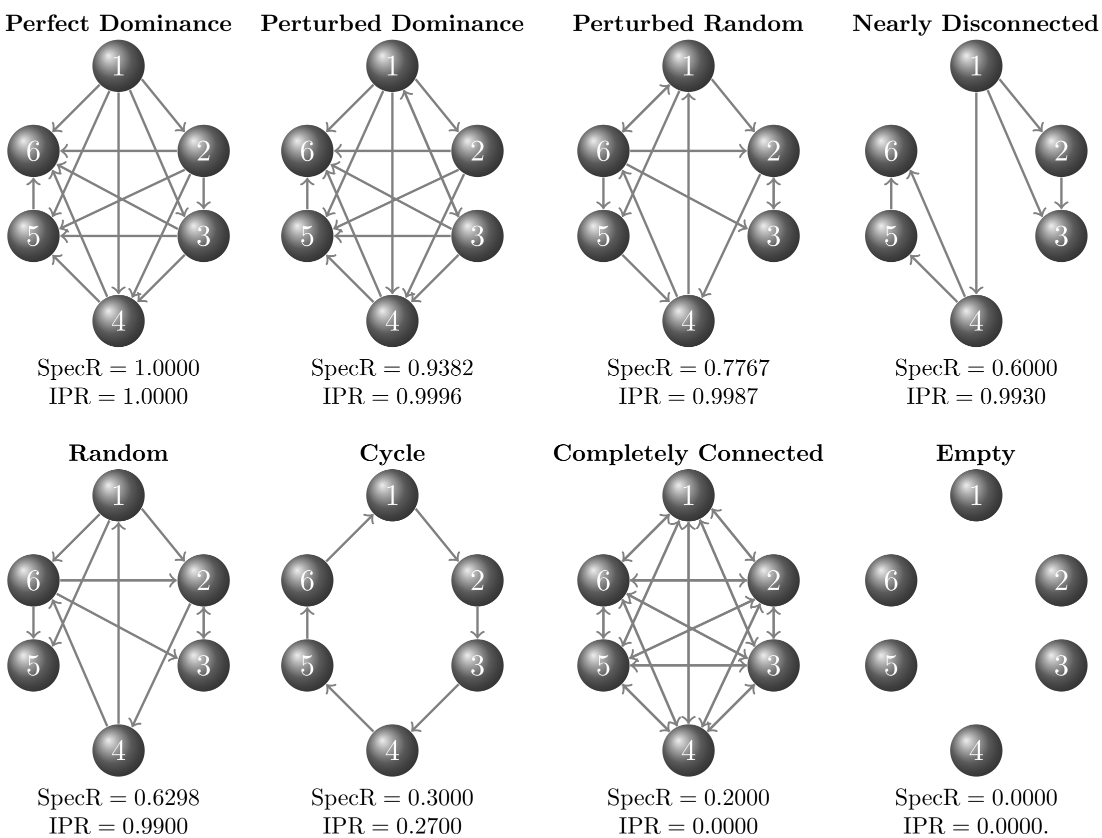

Figure 2: Structured graph data from [1].

In Table 1, we record the rankability measure for each year of the NCAA Big East Football Conference

Year	SpecR	Massey	Colley	Cycles
1995	0.857	0.893	0.926	7
1996	0.857	0.857	0.929	1
1997	0.815	0.679	0.704	62
1998	0.817	0.750	0.917	30
1999	0.857	0.821	0.857	23
2000	0.857	0.929	0.929	3
2001	0.857	0.857	0.963	1
2002	0.857	0.893	0.962	3
2003	0.857	0.786	0.893	18
2004	0.662	0.762	0.750	48
2005	0.838	0.821	0.880	16
2006	0.805	0.750	0.778	45
2007	0.684	0.643	0.692	205
2008	0.805	0.714	0.846	43
2009	0.857	0.786	0.888	26
2010	0.708	0.750	0.808	174
2011	0.714	0.643	0.792	125
2012	0.714	0.643	0.720	111

Table 1: Big East Football Rankability

from 1995 to 2012. For comparison purposes, we also provide the predictability of Colley and Massey rankings [2, 7] and the number of cycles in the corresponding graph. We note that the Pearson correlation coefficients between SpecR and Massey is 0.719, between SpecR and Colley is 0.781, and between SpecR and the number of cycles is -0.826. Hence, the SpecR rankability measure is giving a good indication of the inherent ability of this data to produce a meaningful ranking of its items.

However, it is also clear from these measures that SpecR has trouble differentiating between certain years of the Big East data. This motivates our investigation of other graph properties that can be used to measure the rankability of data. In particular, we have had success using the algebraic connectivity of a graph to differentiate between years of Big East data and identify a unique least and most rankable years. Now, our goal is to provide a rigorous theoretical justification for the use of the algebraic connectivity, just as was done with the spectral-degree characteristic in the Pre-Print.

2 Algebraic Connectivity

Let Γ be a directed graph with non-negative weights and let L be the graph Laplacian of Γ. Furthermore, let e denote the all ones vector. Then, the algebraic connectivity of L is defined as follows [8]:

$$\alpha(\Gamma) = \min_{x \in S} x^T L x,$$

where

$$S = \{x \in \mathbb{R}^n \colon x \perp e, \|x\| = 1\}.$$

Another related and useful quantity is the following:

$$\beta(\Gamma) = \max_{x \in S} x^T L x.$$

It is important to note that both $\alpha(\Gamma)$ and $\beta(\Gamma)$ are invariant under re-ordering of vertices of Γ since the space S is invariant under permutation.

Let Q be an orthonormal matrix whose columns span S. Then, we have

$$\alpha(\Gamma) = \min_{\|Qx\|=1} x^T Q^T L Q x$$

and

$$\beta(\Gamma) = \max_{\|Qx\|=1} x^T Q^T L Q x.$$

Furthermore, let $F(A)$ denote the field of values of a complex matrix A, as defined in [4]. Then, the Hermitian part of A, denoted $H(A) = \frac{1}{2}(A + A^*)$, satisfies the following:

$$F(H(A)) = \mathrm{Re}\,(F(A)).$$

Moreover, the end points of $\mathrm{Re}\,(F(A))$ are well-known to be the minimum and maximum eigenvalues of $H(A)$. Therefore, we have

$$\alpha(\Gamma) = \lambda_{\min}\left(\frac{1}{2}Q^T(L + L^T)Q\right)$$

and

$$\beta(\Gamma) = \lambda_{\max}\left(\frac{1}{2}Q^T(L + L^T)Q\right).$$

2.1 Basic Properties

One of the motivations for the definition of $\alpha(\Gamma)$ is the several properties of Fiedler's definition of the algebraic connectivity for undirected graphs that remain valid [3]. For instance, we have the following super and sub additivity property.

Proposition 2.1 (See Lemma 1 of [8])**.** *If two graphs Γ_1 and Γ_2 have the same vertex set and disjoint edge sets, then*

$$\alpha(\Gamma_1) + \alpha(\Gamma_2) \leq \alpha(\Gamma_1 \cup \Gamma_2) \leq \beta_(\Gamma_1 \cup \Gamma_2) \leq \beta(\Gamma_1) + \beta(\Gamma_2).$$

We also have the following bounds that follow readily from the Courant-Fisher min-max theorem.

Proposition 2.2. *Denote the eigenvalues of the Hermitian part of L by*

$$\lambda_1 \leq \ldots \leq \lambda_n.$$

Then, we have

$$\lambda_1 \leq \alpha(\Gamma) \leq \lambda_2$$

and

$$\lambda_{n-1} \leq \beta(\Gamma) \leq \lambda_n.$$

This leads to our first result on the algebraic connectivity of perfect dominance graphs.

Theorem 2.3. *Let Γ be a perfect dominance graph. Then,*

$$\beta(\Gamma) < n.$$

Proof. Without loss of generality, we can assume that L is in Frobenius normal form. Then, the Hermitian part of L satisfies

$$H(L) = \begin{bmatrix} n-1 & -1/2 & \cdots & -1/2 \\ -1/2 & n-2 & \cdots & -1/2 \\ \vdots & & \ddots & \vdots \\ -1/2 & -1/2 & \cdots & 0 \end{bmatrix}.$$

Note that

$$\lim_{k\to\infty} \left(\frac{1}{n}H(A)\right)^k < 1.$$

Therefore, by Theorem 5.6.12 of [5], the spectral norm satisfies

$$\rho(H(A)) < n,$$

and the result follows from Proposition 2.2. □

The following proposition aligns our definition of $\beta(\Gamma)$ for directed graphs with Fiedler's original definition for undirected graphs.

Proposition 2.4 (See Lemma 4 of [8])**.** *Let $\overline{\Gamma}$ denote the complement of Γ, which is obtained by reversing the direction of all the edges. Then,*

$$\alpha(\Gamma) + \beta(\bar{(\Gamma)}) = n.$$

We also have the following connection between our definition of $\alpha(\Gamma)$ and Fiedler's original definition for undirected graphs as the second smallest eigenvalue of the graph Laplacian.

Proposition 2.5 (See Lemma 5 of [8])**.** *If λ is an eigenvalue of L not corresponding to the eigenvector e, then*

$$\alpha(\Gamma) \leq \mathrm{Re}\,(\lambda)\,.$$

This lead to our next result on the algebraic connectivity of perfect dominance graphs.

Theorem 2.6. *Let Γ be a perfect dominance graph. Then,*

$$\alpha(\Gamma) \leq 1.$$

Proof. Let L be the graph Laplacian of Γ. By Theorem 3.2 of the attached pre-print, the eigenvalues of L satisfy

$$\sigma(L) = \{n-1, \ldots, 1, 0\}\,,$$

where 0 is the only eigenvalue corresponding to the eigenvector e. The result follows from Proposition 2.5. ☐

We conclude this section with the following corollary, which provides bounds on the algebraic connectivity of perfect dominance graphs.

Corollary 2.7. *Let Γ be a perfect dominance graph. Then,*

$$0 < \alpha(\Gamma) \leq 1 \quad \textit{and} \quad n-1 \leq \beta(\Gamma) < n.$$

Proof. Theorem 2.3 and Theorem 2.6 imply that

$$\alpha(\Gamma) \leq 1 \quad \text{and} \quad \beta(\Gamma) < n.$$

Furthermore, the reversal of a perfect dominance graph is also a perfect dominance graph. In fact, $\overline{\Gamma}$ is a perfect dominance graph associated with the reverse ranking of Γ. Hence $\overline{\Gamma}$ and Γ are isomorphic and it follows that

$$\alpha(\Gamma) = \alpha(\overline{\Gamma}) \quad \text{and} \quad \beta(\Gamma) = \beta(\overline{\Gamma}).$$

Therefore, Proposition 2.4 implies that

$$\alpha(\Gamma) + \beta(\Gamma) = n.$$

Hence,

$$\alpha(\Gamma) = n - \beta(\Gamma) > 0$$

since $\beta(\Gamma) < n$. and

$$\beta(\Gamma) = n - \alpha(\Gamma) \geq n-1$$

since $\alpha(\Gamma) \leq 1$. ☐

2.2 The Field of Values

We aspire to give a more specific description of $\alpha(\Gamma)$ and $\beta(\Gamma)$ when Γ is a perfect dominance graph. Ideally, we would have a characterization result analogous to the spectral-degree characterization in Theorem 1.1. To this end, we started an investigation on the field of values of the matrix Q^TLQ, where Q is defined as in the beginning of Section 2. After running several numerical experiments, we noticed that $F(Q^TLQ)$ is an ellipse whenever Γ is a perfect dominance graph. In Figure 3, we display the field of values of Q^TLQ for a perfect dominance graph on 3, 6, 9, and 12 vertices.

Note that for a perfect dominance graph on 3 vertices, the matrix Q^TLQ is 2×2. In particular, we have

$$\begin{aligned} Q^TLQ &= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/\sqrt{6} \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 1/\sqrt{3} & 1 \end{bmatrix}. \end{aligned}$$

The eigenvalues of Q^TLQ are

$$\lambda_1 = 1 \quad \text{and} \quad \lambda_2 = 2.$$

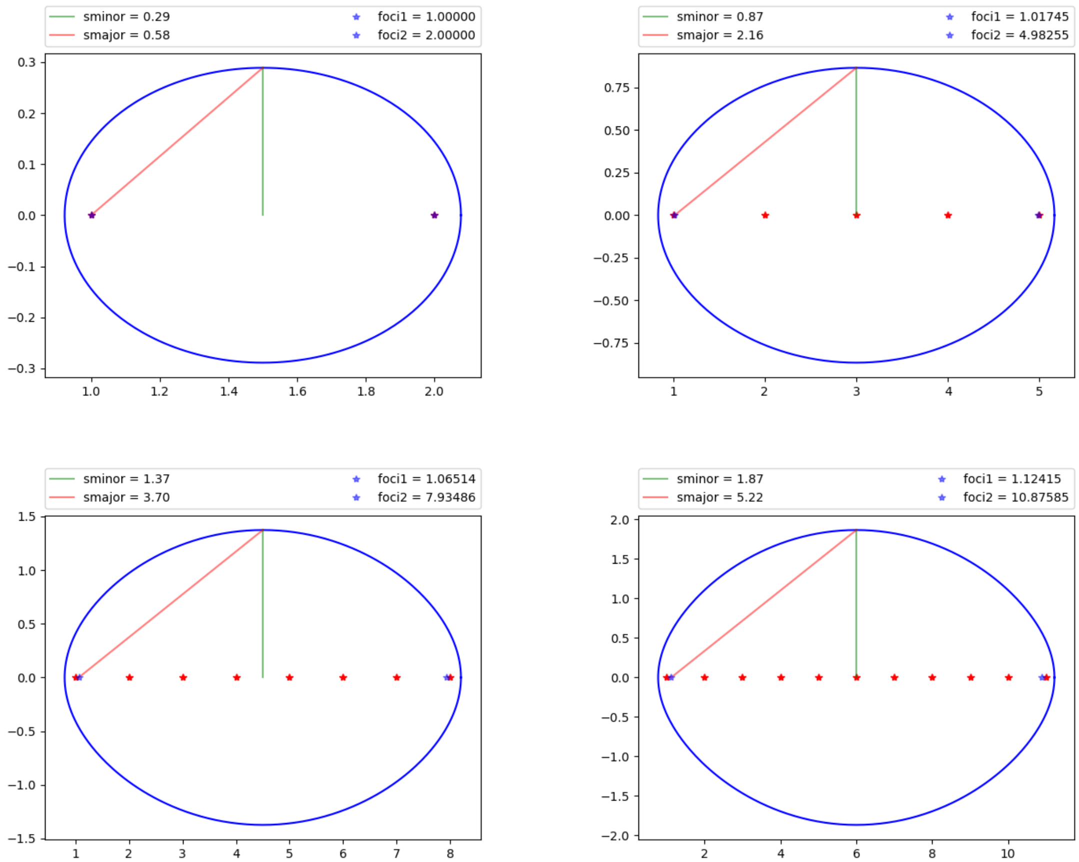

Figure 3: $F(Q^T LQ)$ for Perfect Dominance Graphs.

Furthermore, the singular values of $Q^T LQ$ are

$$\sigma_1 = \sqrt{\frac{2}{3}\left(4+\sqrt{7}\right)} \quad \text{and} \quad \sigma_2 = \sqrt{\frac{2}{3}\left(4-\sqrt{7}\right)}$$

Therefore, by the elliptical range theorem [6], $F(Q^T LQ)$ is an ellipse with foci at 1 and 2, and with a major axis of length $\sqrt{16/3 - 1 - 4} \approx 0.58$.

2.3 Future Research

We are interested in proving that the field of values of $Q^T LQ$ is always an ellipse when Γ is a perfect dominance graph. The first step is to prove it when Γ is a perfect dominance on 4 vertices and, hence, $Q^T LQ$ is 3×3. To generalize this result, we may find it useful to look into the many proofs of the elliptical range theorem for ideas.

We hope that a proof of this result will lead to additional information about the algebraic connectivity of perfect dominance graphs. Ideally, this information will lead to another characterization and further justification of its use as a measure of the rankability of data.

References

[1] P. E. Anderson, T. P. Chartier, and A. N. Langville, *The rankability of data*, SIMODS, (2018).

[2] W. N. Colley, *Colley's Bias Free Matrix Rankings*. `www.colleyrankings.com`, 2002.

[3] M. Fiedler, *Algebraic connectivity of graphs*, Czechoslovak Mathematical Journal, 23 (1973), pp. 298–305.

[4] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, NY, 1991.

[5] ———, *Matrix Analysis*, Cambridge University Press, New York, NY, 2nd ed., 2013.

[6] C.-K. Li, *A simple proof of the elliptical range theorem*, Proc. Amer. Math. Soc., 127 (1996), pp. 1985–1986.

[7] K. Massey, *Statistical modelis applied to the rating of sports teams.* Bachelor's thesis, Bluefield College, 1997.

[8] C. W. Wu, *Algebraic connectivity of directed graphs*, Linear Multilinear Algebra, 53 (2005), pp. 203–223.