A Note to Panos:

Algebraic Connectivity of Perfect Dominance Graphs

Thomas R. Cameron

June 10, 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 of 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.
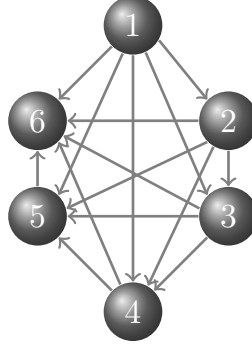

Figure 1: Perfect dominance graph associated with ranking $[1, 2, 3, 4, 5, 6]$.

Now, in theory, the rankability measure proposed in [1] is simple. Let k denote the minimum number of edge deletions or additions made to the given graph in order 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, I believe 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*

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

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

1

Armed with this spectral-degree characterization, we propose the following measure of the rankability of data. Note that hd(,d)enotes the Hausdorff distance and the maximum upper bound for both $\text{hd}(D, S)$ and $\text{hd}(L, S)$ 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 $[r] = \text{SpecR}\,(\Gamma)$:
   $n \leftarrow$ the number of vertices in Γ
   $D \leftarrow$ the out-degree matrix of Γ
   $L \leftarrow$ graph Laplacian of Γ
   $S = \text{diag}\,(n-1, n-2, \ldots, 0)$
   $r = 1 - \frac{\text{hd}(D,S)+\text{hd}(L,S)}{2(n-1)}$
   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.
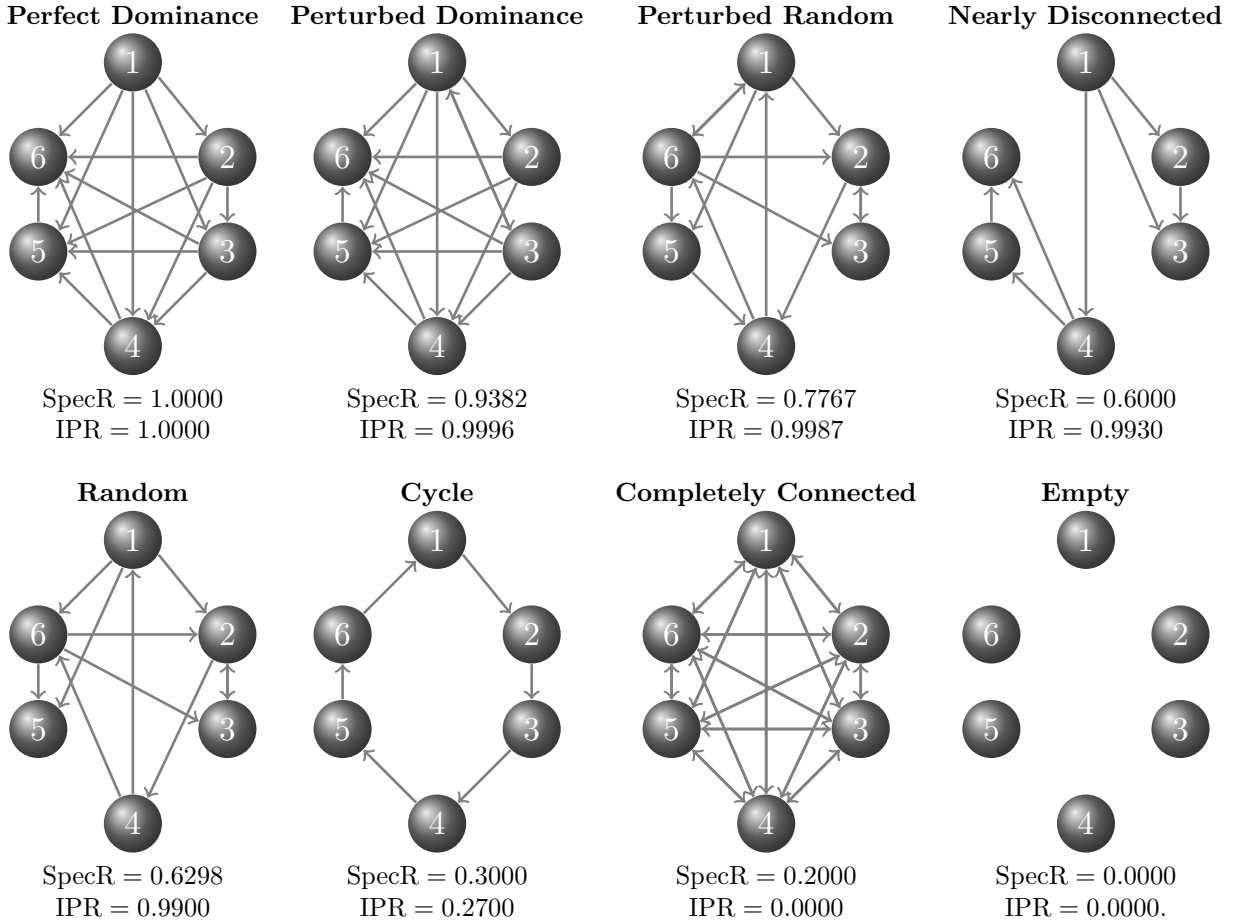

Perfect Dominance
SpecR = 1.0000
IPR = 1.0000

Perturbed Dominance
SpecR = 0.9382
IPR = 0.9996

Perturbed Random
SpecR = 0.7767
IPR = 0.9987

Nearly Disconnected
SpecR = 0.6000
IPR = 0.9930

Random
SpecR = 0.6298
IPR = 0.9900

Cycle
SpecR = 0.3000
IPR = 0.2700

Completely Connected
SpecR = 0.2000
IPR = 0.0000

Empty
SpecR = 0.0000
IPR = 0.0000.

Figure 2: Structured graph data from [1].

Year	SpecR	Year	SpecR
1995	0.8571	2004	0.6615
1996	0.8571	2005	0.8375
1997	0.8149	2006	0.8049
1998	0.8169	2007	0.6841
1999	0.8571	2008	0.8049
2000	0.8571	2009	0.8571
2001	0.8571	2010	0.7082
2002	0.8571	2011	0.7143
2003	0.8571	2012	0.7143

Table 1: Big East Football Rankability

In Table 1, we record the rankability measure for each year of the NCAA Big East Football Conference from 1995 to 2012. It is important to note that there is still a strong correlation to the rankability measure IPR. Moreover, there is a strong correlation to the predictability of Colley and Massey rankings [2, 6]. Hence, these rankability measures give a good indication of the inherent ability of the this data to produce a meaningful ranking of its items.

However, it is also clear from these measures that SpecR has trouble differentiation 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 used the algebraic connectivity of a graph to differentiate between years of the big east data and identify unique years as the least and most rankable. 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 [7]:

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

where

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

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:

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

Moreover, the end points of $\operatorname{Re}(F(A))$ 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

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

3

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 [7]). *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

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

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. $\qquad\square$

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 [7]). *Let $\overline{\Gamma}$ denote the complement of Γ, which is obtained by reversing the direction of all the edges. Then,*

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

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 [7]). *If λ is an eigenvalue of L not corresponding to the eigenvector e, then*

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

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. $\qquad\square$

We conclude this section with the following corollary.

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

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

Proof. Theorem 2.3 and Theorem 2.6 imply that

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

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

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

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. To this end, we have started an investigation of the field of values of the matrix $Q^T L Q$, where Q is defined as in the beginning of Section 2.

After running some numerical experiments, the first thing we noticed is that $F(Q^T L Q)$ is an ellipse whenever Γ is a perfect dominance graph.

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

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