Mitri Kitti

Axioms for Centrality Scoring with Principal Eigenvectors

Aboa Centre for Economics

Discussion paper No. 79

Turku 2012

The Aboa Centre for Economics is a joint initiative of the economics departments of the University of Turku and Åbo Akademi University.

ISSN 1796-3133

Printed in Uniprint
Turku 2012

Mitri Kitti

Axioms for Centrality Scoring with Principal Eigenvectors

Aboa Centre for Economics
Discussion paper No. 79
October 2013
(First draft November 2012)

ABSTRACT

Techniques based on using principal eigenvectors of matrices representing binary relations of sets of alternatives are commonly used in social sciences, bibliometrics, and web search engines. In most applications the binary relations can be represented by a directed graph and the question of ranking or scoring the alternatives can be turned into the relevant question of how to score the nodes of the graph. This paper characterizes the principal eigenvector as a scoring function with a set of axioms. A zero-sum scoring function based on the difference of principal right and left eigenvectors is introduced and axiomatized. Furthermore, a method of assessing individual and group centralities simultaneously is characterized by a set of axioms. The specific case of this method is the Hyperlink-Induced Topic Search (HITS) used in ranking websites.

JEL Classification: C60, C70, D70

Keywords: principal eigenvector, graph, centrality, scoring, axiom

Contact information

Mitri Kitti
Department of Economics
University of Turku
FI-20014, Finland

E-mail: mitri.kitti (at) utu.fi

Acknowledgements

I thank Hannu Salonen, Olli Lappalainen, and Matti Pihlava for discussions and comments.

1. Introduction

The question of assigning scores to a set of alternatives by using information on their bilateral relationships arises in many fields. In sports the players or teams need to be ranked according to outcomes of their games, the quality of scientific publications is assessed using their mutual citations, Internet search engines organize websites according to their link structure, in social choice theory alternatives are ranked based on voters' preferences, in the analysis of social networks the importance of individuals is evaluated according to their ties. These situations can be represented by directed graphs and the question then becomes on scoring the nodes of the graph to assess their centrality. For an overview on various scoring and ranking methods see David (1988) and Laslier (1997).

The link structure of a graph can be presented by a matrix which indicates whether there is a link between any pair of nodes and how strong this link is. The idea of using the eigenvectors of this matrix, the adjacency matrix of the graph, first appears in the works of Seeley (1949); Wei (1952) and Kendall (1955). See also Daniels (1969), Moon and Pullman (1970), and Pinski and Narin (1976) for early developments of methods based on principal eigenvectors.

The first step in centrality scoring is to form a matrix representing the relations of nodes. The scores of the nodes can then be obtained by using the principal eigenvector; either the right or left eigenvector. There are numerous ways to form adjacency matrices from underlying data, which may contain for example the tournament outcomes or the link structure of websites. One way is to take the components of the matrix directly from the data describing the binary relations of the nodes in the graph. The other way is to normalize the matrix, which usually leads to a stochastic matrix.

The normalization can be done in several ways. The two most commonly used methods, the invariant method and the fair bets method have been axiomatized by Palacios-Huerta and Volij (2004) and Slutzki and Volij (2005, 2006). The celebrated PageRank method of Google (Brin and Page, 1998) is closely related to the invariant method, and it has been axiomatized by Altman and Tennenholtz (2005). All these methods rely on forming a stochastic matrix from the original binary relations, and computing its stationary distribution.

For the non-normalized method—also known as the Kendall-Wei method or the long path method (Laslier, 1997)—there is no characterization results in the prior literature. This paper remedies this matter. In essence, the principal eigenvector of an irreducible matrix is axiomatized: the principal eigenvector is the unique scoring function that satisfies a set of axioms introduced in this paper. This result relies on the Perron-Frobenius theorem.

The normalized principal right eigenvector represents the asymptotic proportions of directed walks in the graph that begin from each node of the graph (Laslier, 1997; Cvetković et al., 1997). A walk between two nodes in a graph means that it is possible to get from one node to another by following the edges of the graph. The left eigenvector, on the other hand, tells the same information on walks terminating to each node. A new zero-sum measure is defined by taking the difference of the two principal eigenvectors. This measure gives the net flow of walks to each node. Zero-sum scoring functions have been previously introduced by David (1987) and Herings et al. (2005). The Copeland scoring function measuring the differences between number of successors and predecessors of nodes is another example of a zero-sum scoring function. It has been axiomatized by Rubinstein (1980); Henriet (1985), and van den Brink and Gilles (2000).

The zero-sum scoring function defined as the difference of the right and left principal eigenvectors is axiomatized. Moreover, it is shown that the zero sum scoring function is obtained as a limit of a weighted index similarly as the usual eigenvectors scores are the limiting cases of the index by Katz (1953) and Hubbell (1965).

One of the most widely known methods for discovering relevant webpages for a particular topic is the HITS (Hyperlink-Induced Topic Search) algorithm of Kleinberg (1999). The method is based on computing iteratively the principal eigenvector of a particular matrix obtained from the one that describes the link structure of websites. As a scoring function the HITS is essentially a specific case of the method of Bonacich (1991) for assessing the centrality of individuals and groups simultaneously. The Bonacich-Kleinberg method, including the HITS as a specific case, will be axiomatized in this paper.

The paper is structured as follows. The properties of the usual principal eigenvector scoring function are studied in Section 2. Section 4 introduces the zero-sum scoring method. The Bonacich-Kleinberg method for scoring of groups and individuals is characterized in Section 5.

2. Preliminaries

Let us assume that the purpose is to score the nodes of directed graphs. Such graphs can be used for example in presenting networks of people or results of tournaments. In a tournament the nodes of the graph could represent players and the edges could indicate that a player has won its opponent. The weight given for an edge could tell the number of times a player has beaten an opponent.

Let us assume that there are n nodes in the graph. Let $\mathbf{A}$ be an $n \times n$-matrix with non-negative components. This matrix is the weighted adjacency matrix of the graph. For a tournament each row represents the wins of a player over the other players. The component of $\mathbf{A}$ in its i'th row and j'th column, denoted by a_{ij}, gives the number of times that i has won j. For a social network $a_{ij} = 1$ would indicate that the individuals i and j are connected. In that case $\mathbf{A}$ becomes a symmetric matrix. In general, a_{ij} can be any non-negative weight assigned to the relation between two nodes i and j in the graph. In the following $\mathbf{I}$ will denote the identity matrix.

The below example illustrates how a social choice problem can be formulated as a graph.

Example 1. Let us assume that there are five alternatives (or candidates): a, b, c, d, and e. Furthermore, let us assume that we have social preference relations for some of them expressed, e.g. in a vote between two alternatives. The graph in Figure 1 expresses these relations; the alternatives are the nodes and the edges between them express the preferences. For example a is preferred over b. The edges between c and e can be interpreted as the alternatives being equally good.

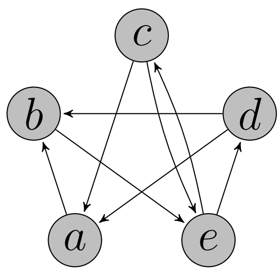

Figure 1: The graph in Example 1.

The adjacency matrix expressing the social preferences, or connection between the nodes, is

$$A = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \end{array} \begin{array}{c} \begin{array}{ccccc} a & b & c & d & e \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

If the preferences over the alternatives were expressed as votes, we could attach to each edge the number of votes by which an alternative beats another. This would lead to a weighted adjacency matrix.

Example 1 demonstrates that we can associate an adjacency matrix to a graph. On the other hand, for a given positive matrix we can find a corresponding graph in which there is an edge between nodes i and j if $a_{ij} > 0$. Without loss of generality we may assume that the weight of the edge is a_{ij}. Let $\mathcal{G}(\mathbf{A})$ denote the directed graph corresponding to $\mathbf{A}$.

It will be assumed that $\mathbf{A}$ is irreducible; nodes that cannot be connected to each other by some walk in the graph are not compared. A walk of length k from node i to node j means that there is a sequence of directed edges leading from i to j in k steps, i.e., there are $i_0, i_1, \ldots, i_k$ with $i_0 = i$ and $i_k = j$ such that $a_{i_t,i_{t+1}} > 0$ for all $t = 0, \ldots, k-1$. Formally, matrix $\mathbf{A}$ is irreducible if $\mathcal{G}(\mathbf{A})$ is strongly connected: for each pair of nodes i, j there is a walk connecting them. For example, in a voting setup this condition holds when we have the number of voters who prefer an alternative to another for all pairs of alternatives, and the resulting matrix is formed from these numbers.

We can also assign the weight $a_{i_0}a_{i_1} \times \cdots \times a_{i_k}$ to the directed walk between nodes i and j. In this case the walk is said to be weighted. If there are several walks of length k between two nodes then we add the weights corresponding to these walks to obtain a^k_{ij} which is the total weight of all walks of length k between the two nodes. Note that a^k_{ij}, $i, j = 1, \ldots, n$, are the components of $\mathbf{A}^k$, i.e., $\mathbf{A}$ to power k.

If the components of $\mathbf{A}$ tell how many edges there are between two nodes, then a^k_{ij} is the number of length k walks between i and j. For instance, let us assume that a_{ij}, $i, j = 1, \ldots, n$, are the numbers of voters who prefer one alternative to another. The number a^k_{ij} tells how many ways there are to get i indirectly preferred to j by making k pairwise comparisons of alternatives where in each comparison $t = 0, \ldots, k-1$ some voter prefers alternative i_t to alternative i_{t+1}. When the components of $\mathbf{A}$ can be any non-negative numbers, multiplying the weights corresponding to the nodes appearing in a walk connecting them gives the weight of the length k connection. In particular, if $\mathbf{A}$ is a stochastic matrix, i.e., the rows sum are one, then a^k_{ij} is the probability of reaching node j in k steps from node i. For general positive weights, the strength of connections between nodes is multiplicative along walks between them, and a^k_{ij} is the sum of the strengths of all length k walks joining the nodes.

In addition to irreducibility it will be assumed that $\mathbf{A}$ is primitive, i.e., $\mathbf{A}^k$ has strictly positive components for some k. Such matrices are irreducible. It turns out that assuming that $\mathbf{A}$ is primitive can be done without loss of generality. In the following $\mathcal{M}$ denotes the set of primitive matrices.

A scoring function $\mathbf{F} = (F_1, \ldots, F_n)$ gives for each $i = 1, \ldots, n$ a score $F_i(\mathbf{A})$, i.e., $\mathbf{F} : \mathcal{M} \mapsto \mathbb{R}^n$. In principle, the score of a node can be any real number, positive or negative. As explained earlier, scoring functions have numerous applications in situations where we need to rank a set of alternatives and are possible interested in choosing the best alternative. For instance, when the underlying data presents voters preferences over a set alternatives, the scoring function can be interpreted as a social welfare function which ranks the different alternatives (Laslier, 1997).

Let $\mathbf{R}(\mathbf{A}) \in \mathbb{R}^n$ denote the principal right eigenvector of $\mathbf{A}$, in brief the principal eigenvector, which is normalized such that $\sum_i R_i(\mathbf{A}) = 1$. The principal right eigenvector is also known as the Kendall-Wei scoring function. By definition we have

$$\mathbf{A}\mathbf{R}(\mathbf{A}) = \lambda(\mathbf{A})\mathbf{R}(\mathbf{A}),$$

where $\lambda(\mathbf{A}) \in \mathbb{R}$ is the principal eigenvalue, i.e., the largest eigenvalue of $\mathbf{A}$. Moreover, let $\mathbf{L}(\mathbf{A})$ denote the corresponding eigenvector of $\mathbf{A}^\top$, i.e., $\mathbf{L}(\mathbf{A})$ is the principal left eigenvector of $\mathbf{A}$. All vectors in this paper are column vectors. The vector $\mathbf{1}$ denotes the n-dimensional vector with all components equal to one.

The main results of this paper rely on the Perron-Frobenius theorem, see, e.g., Meyer (2000). For the purposes of axiomatizing eigenvector measures we need the following results from the Perron-Frobenius theorem.

Theorem 1. *For any $\mathbf{A} \in \mathcal{M}$ it holds that*

1. $\lambda(\mathbf{A}) > 0$*,*
2. *the normalized principal eigenvector $\mathbf{R}(\mathbf{A})$ is unique and strictly positive, i.e., $\mathbf{R}(\mathbf{A}) \gg \mathbf{0}$,*
3. *there are no other non-negative eigenvectors than the positive multiples of $\mathbf{R}(\mathbf{A})$,*
4. $$\lim_{k\to\infty} \frac{\mathbf{A}^k}{\lambda^k(\mathbf{A})} = \frac{\mathbf{R}(\mathbf{A})\mathbf{L}(\mathbf{A})^\top}{\rho(\mathbf{A})}, \tag{1}$$

   where $\rho(\mathbf{A}) = \mathbf{R}(\mathbf{A})^\top\mathbf{L}(\mathbf{A})$.

Before proceeding to the axiomatization of the principal eigenvector let us briefly discuss the restriction that the underlying graph is assumed to be connected, i.e., $\mathbf{A} \in \mathcal{M}$. This restriction is common for all methods relying on principal eigenvectors, because otherwise the principal eigenvector may not be unique. It is, however, possible to extend the eigenvector based methods for graphs having non-connected components. Slutzki and Volij (2005) describe how to obtain an ordinal extension of a method defined for reducible matrices. Similar approaches are also discussed by Borm et al (2002) and

Hu and Shapley (2003). Echenique and Fryer (2007) introduce an explicit formula for extending the principal eigenvector to irreducible matrices by partitioning the graph into connected components and utilizing the principal eigenvectors for each component.

Let us assume that that there are m strongly connected components in the graph $\mathcal{G}(\mathbf{A})$. Then the adjacency matrix $\mathbf{A}$ corresponds to the adjacency matrices $\mathbf{A}_1, \ldots, \mathbf{A}_m$ of these components. Note that each node belongs to exactly one component. If node i belongs to the k'th component, the score of the node is

$$F_i(\mathbf{A}) = \lambda(\mathbf{A}_k) R_i(\mathbf{A}_k) |\mathbf{A}_k|, \tag{2}$$

where $R_i(\mathbf{A}_k)$ is the eigenvector score of i corresponding to the adjacency matrix of the k'th strongly connected component, and $|\mathbf{A}|_k$ is the number of nodes in the k'th component. Echenique and Fryer (2007) axiomatize a segregation measure which involves the scores of the individuals, and they show that the resulting scores are given by Equation (2). One of their axioms is that each connected component is scored by an eigenvector.

3. The eigenvector scoring function

3.1. Axiomatization of the principal eigenvector

In this section it is assumed that the purpose is to score the nodes of a directed graph, and the adjacency matrix of the graph is irreducible and has non-negative components, i.e., belongs to $\mathcal{M}$. In particular, all matrices mentioned in this section belong to $\mathcal{M}$.

Let us now introduce a set of axioms for a scoring function $\mathbf{F} : \mathcal{M} \mapsto \mathbb{R}^n$.

(A1) $\mathbf{F}(\mathbf{A}) \gg \mathbf{0}$,
(A2) $\mathbf{F}(\beta\mathbf{A}) = \mathbf{F}(\mathbf{A})$ for all $\beta > 0$.
(A3) $\mathbf{F}(\mathbf{A} + \beta\mathbf{I}) = \mathbf{F}(\mathbf{A})$ for all $\beta \geq 0$.
(A4) $\mathbf{F}(\mathbf{A}^k) = \mathbf{F}(\mathbf{A})$ for all $k = 1, 2, \ldots$.
(A5) If $\mathbf{F}(\mathbf{A}) = \mathbf{F}(\mathbf{B})$, then $\mathbf{AF}(\mathbf{A}) = \beta\mathbf{BF}(\mathbf{B})$ for some $\beta > 0$.
(A6) $\sum_i F_i(\mathbf{A}) = 1$.

The first axiom means that all the scores are positive. When dealing with an irreducible matrix $\mathbf{A}$ all nodes are connected to at least one other node. Hence, it makes sense that all nodes get positive scores.

The second axiom, the homogeneity of degree zero, means that $\mathbf{F}$ is invariant for multiplying all the weights of the edges of the graph by the same constant. This property holds for all commonly used scoring methods. The axiom can be interpreted as the invariance to the unit in which the strength of connections in the network are measured. It is a reasonable assumption that the unit or scale does not matter. For instance, in a social choice problem it should not matter whether we express the comparisons of two alternatives as absolute values, e.g., numbers of votes, or percentages of voters preferring one alternative to another.

It should be noticed that scaling the matrix with a positive constant should not be confused with normalizing it. A normalization means that we manipulate the relative weights given by the nodes. For example, we may turn the comparison of i and j into a probability of i beating j. Let us consider a graph with nodes $1, 2, 3$ and $a_{12} = 2$, $a_{21} = 1$, and $a_{23} = 4$ and $a_{32} = 2$. We can normalize the matrix by giving the edge from i to j the weight $\hat{a}_{ij} = a_{ij}/(a_{ij} + a_{ji})$. This gives $\hat{a}_{12} = \hat{a}_{23} = 2/3$. This normalization cannot be expressed as a change of scale, i.e., as multiplication of $\mathbf{A}$ by a scalar. Note, however, that when there is $c > 0$ such that $a_{ij} + a_{ji} = c$ for all $i \neq j$, and $a_{ii} = 0$ for all $i = 1, \ldots, n$, such as in voting or in tournaments, then the above normalization is equivalent with multiplying A with $1/c$.

The third axiom means that the score remains the same if we add each node an edge to itself with equal weight. This means that the results of the scoring function cannot be manipulated by multiplying the self-connections by the same number. However, this does not mean that self-connections were not relevant at all. What matter are the relative strengths of self-connections, not their absolute values. This property holds for all methods were self-connections are removed, which is rather common, e.g., in bibliometrics or in assessing the authoritativeness of webpages. The motive for removing the self-connections such as self-citations or self-links is to prevent manipulation of the scores.

The fourth axiom means that if we consider the k-length weighted walks from each node to other nodes, the scores remain the same. Hence, $\mathbf{F}$ is invariant to the strength of connections corresponding to any length of walks between the nodes. For instance in a voting situation a walk of length k between two nodes can be interpreted as an alternative being indirectly preferred to another. Recall, however, that the walks of any given length are given the weights corresponding to $\mathbf{A}^k$. The axiom (A4) then means that it does not matter whether we use the information on votes between all pairs of alternatives or information on indirect preferences with weights $\mathbf{A}^k$ of any fixed length k of walks. Hence, the axiom can be interpreted as the invariance to weighted indirect comparisons of nodes. If $\mathbf{F}$ is a continuous scoring function that satisfies (A3), we can define another scoring function $\mathbf{F}^*(\mathbf{A})$ as the limit of $\mathbf{F}(\mathbf{A}^k)$, which by Equation 1 is

$$\mathbf{F}\left(\mathbf{R}(\mathbf{A})\mathbf{L}(\mathbf{A})^\top\right),$$

and the resulting function $\mathbf{F}^*$ satisfies (A4).

In the fifth axiom we multiply the scores of two matrices $\mathbf{A}$ and $\mathbf{B}$ with these matrices. This gives us $\mathbf{v}^1 = \mathbf{AF}(\mathbf{A})$ and $\mathbf{v}^2 = \mathbf{BF}(\mathbf{B})$. The axiom says that if the scores of two matrices are the same, the vectors $\mathbf{v}^1$ and $\mathbf{v}^2$ are scalar multiples of each other. To clarify the axiom let us assume that we give each node $i = 1, \ldots, n$ in a network an initial mass of power $F_i(\mathbf{A})$. The vector $\mathbf{AF}(\mathbf{A})$ gives the distribution of power after we share the initial vector of power $\mathbf{F}(\mathbf{A})$ by computing the net outflow of power from each node to the whole network. The outflow from node i to node j is assumed to be a_{ij} multiplied with the power of node j.

The axiom (A5) says that if the initial power vectors are the scores and they are the same for two matrices, then the distributions on net flows of power are scalar multiples of each other. Hence, the relative flows, i.e., flows from each node compared to any other node or the total flow of the network, remain the same for matrices with equal scores. This property, the equality of relative flows for matrices with the same scores, is the most distinguishing feature of principal eigenvectors. No other commonly used scoring function satisfies it for all vectors having the same score. However, the usual points method where the scores are equal to the out-degrees of the nodes has this property for any two matrices for which all the nodes get equal scores.

The last axiom is only the normalization of scores. Note that the normalization of scores is a different operation than the normalization of the original data, e.g., by forming a stochastic matrix from the initial adjacency matrix describing the binary relations of the nodes.

We can observe that $\mathbf{R}$ satisfies the aforementioned axioms. This follows directly from the standard results for eigenvectors. Adding a matrix $\beta\mathbf{I}$ changes only the principal eigenvalue, while the normalized principal eigenvector remains the same. Same holds when the matrix is multiplied by a scalar. Moreover, eigenvectors are unaffected by taking any positive power of the matrix.

Remark 1. $\mathbf{R}(\mathbf{A})$ satisfies axioms (A1)–(A6).

An important implication of (A3) is that we can assume that $a_{ii} > 0$ for all i without the loss of generality. Recall that a non-negative matrix with a positive component on its diagonal is primitive. Hence, we can always turn an irreducible matrix into a primitive matrix by adding components on the diagonal without affecting the scores.

The main result of this paper is that $\mathbf{R}$ is the only scoring function that satisfies the axioms (A1)–(A6).

Proposition 1. *The unique scoring function* $\mathbf{F}$ *that satisfies (A1)–(A6) for any primitive matrix is* $\mathbf{F} = \mathbf{R}$.

Proof. Let $\mathbf{F} : \mathcal{M} \mapsto \mathbb{R}^n$ be a function that satisfies the assumptions of the proposition. By (A2) the vector $\mathbf{F}^k = \mathbf{F}(\mathbf{A}^k/\lambda^k(\mathbf{A}))$ equals $\mathbf{F}(\mathbf{A})$ for all $k = 1, 2, \ldots$. Hence, the limit of the sequence $\{\mathbf{F}^k\}_k$ is $\mathbf{F}(\mathbf{A})$. By (A5) we have

$$\mathbf{A}^k\mathbf{F}(\mathbf{A})/\lambda^k(\mathbf{A}) = \beta_k\mathbf{A}\mathbf{F}(\mathbf{A}), \tag{3}$$

where $\beta_k > 0$ for all k. According to the Perron-Frobenius theorem, i.e., Equation (1)

$$\mathbf{A}^k/\lambda^k(\mathbf{A}) \to \mathbf{R}\,(\mathbf{A})\mathbf{L}(\mathbf{A})^\top/\alpha(\mathbf{A}),$$

where $\alpha(\mathbf{A}) = \mathbf{R}(\mathbf{A})^\top\mathbf{L}(\mathbf{A}) > 0$. Moreover, both sides of (3) are bounded, which implies that the sequence $\{\beta_k\}_k$ is a bounded sequence of positive numbers. Hence, it has a convergent subsequence with a limit $\bar{\beta} \geq 0$. By taking the limit of both sides of (3) corresponding to the convergent subsequence of $\{\beta_k\}_k$ gives

$$\mathbf{R}\,(\mathbf{A})\,\mathbf{L}\,(\mathbf{A})^\top\,\mathbf{F}(\mathbf{A})/\alpha(\mathbf{A}) = \bar{\beta}\mathbf{A}\mathbf{F}(\mathbf{A}).$$

We can write this as

$$\rho(\mathbf{A})\mathbf{R}\,(\mathbf{A})\,/\alpha(\mathbf{A}) = \bar{\beta}\mathbf{A}\mathbf{F}(\mathbf{A}), \tag{4}$$

where $\rho(\mathbf{A}) = \mathbf{L}\,(\mathbf{A})^\top\,\mathbf{F}(\mathbf{A})$.

If $\rho(\mathbf{A}) = 0$ or $\bar{\beta} = 0$, then $\mathbf{F}(\mathbf{A})$ would be orthogonal to $\mathbf{L}(\mathbf{A})$, which cannot be the case since $\mathbf{L}(\mathbf{A}), \mathbf{F}(\mathbf{A}) \gg \mathbf{0}$. Hence, $\rho(\mathbf{A}), \bar{\beta} > 0$.

By (A3) we can assume that $\mathbf{A}$ is invertible. Namely, if it was not, then we could add $\beta\mathbf{I}$ to $\mathbf{A}$ to get a diagonally dominant matrix with the same scores as $\mathbf{A}$. Diagonally dominant matrix is non-singular by the Levy-Desplanques theorem. Because $\mathbf{R}(\mathbf{A})$ is an eigenvector we have $\mathbf{R}(\mathbf{A}) = \mathbf{A}\mathbf{R}(\mathbf{A})/\lambda(\mathbf{A})$. If we plug in $\mathbf{A}\mathbf{R}(\mathbf{A})/\lambda(\mathbf{A})$ in place of $\mathbf{R}(\mathbf{A})$ and multiply both sides of Equation (4) with $\mathbf{A}^{-1}$ we observe that $\mathbf{F}(\mathbf{A})$ is parallel to $\mathbf{R}(\mathbf{A})$. The axiom (A6) guarantees that $\mathbf{F}(\mathbf{A}) = \mathbf{R}(\mathbf{A})$. □

Let us next discuss the independence of the axioms characterizing the principal eigenvector. First, if $\mathbf{F}(\mathbf{A})$ is allowed to be non-positive, then any eigenvector, e.g., the eigenvectors corresponding to the second largest eigenvalue would satisfy the axioms.

To see that (A2) is independent of other axioms, let us define a scoring function $\mathbf{F}(\mathbf{A})$ for stochastic matrices such that it satisfies all other axioms except for (A2). We can extend the function outside of the domain of stochastic matrices (or matrices for which row sums are equal) by defining it as the usual principal eigenvector. Let $\mathbf{R}^1(\mathbf{A})$ denote the principal eigenvector and let $\mathbf{R}^2(\mathbf{A})$ denote the eigenvector corresponding to the second largest eigenvalue. Let us also denote $\gamma(\mathbf{A}) = \max_i |R_i^2(\mathbf{A})|/\min_i R_i^1(\mathbf{A})$, and let $\beta(\mathbf{A})$

be the largest column sum of $\mathbf{A}$. The function $\mathbf{F}(\mathbf{A})$ is now obtained by normalizing $\beta(\mathbf{A})\gamma(\mathbf{A})\mathbf{R}^1(\mathbf{A}) + \mathbf{R}^2(\mathbf{A})$ such that (A6) holds. By assuming that $\mathbf{A}$ and $\mathbf{B}$ in the axioms are stochastic matrices such that $\mathbf{R}^i(\mathbf{A}) = \mathbf{R}^i(\mathbf{B})$, $i = 1, 2$, we can observe that (A3)–(A6) hold. However, (A2) need not hold.

To show that the third axiom is independent of the other axioms, let us define $\mathbf{F}(\mathbf{A})$ for matrices with non-empty null-space by normalizing the vector $\mathbf{R}(\mathbf{A}) - \rho(\mathbf{A})\mathbf{v}(\mathbf{A})$, such that (A6) holds. Here $\mathbf{v}(\mathbf{A})$ is a vector in the null space of $\mathbf{A}$, and $\rho(\mathbf{A}) \geq 0$ is chosen such that $\mathbf{R}(\mathbf{A}) - \rho(\mathbf{A})\mathbf{v}(\mathbf{A}) \gg \mathbf{0}$. Then for $\mathbf{A}$ and $\mathbf{B}$ having the same null-space with positive dimension (there is $\mathbf{v}(\mathbf{A}) \neq \mathbf{0}$) all the other axioms except for (A3) hold. We can extend the function outside of the domain of matrices with non-empty null-space by defining at the usual principal eigenvector if the null-space is empty.

If we dropped (A4), then any $\mathbf{F}(\mathbf{A}) = \mathbf{R}(\mathbf{A}^k)$, $k \geq 1$, would satisfy the rest of the axioms. If the axiom (A5) was not assumed, then the rest of the axioms hold for any $\mathbf{F}(\mathbf{A}) = \mathbf{g}(\mathbf{R}(\mathbf{A}))$ where $\mathbf{g}(\mathbf{x}) \geq \mathbf{0}$ for all $\mathbf{x} \geq \mathbf{0}$ and $\sum g_i(\mathbf{x}) = 1$. Finally, it can directly be seen that (A6) cannot be omitted. Otherwise, any scalar multiple of $\mathbf{R}(\mathbf{A})$ would satisfy the other axioms.

3.2. Interpretation and properties

Before proceeding to other scoring functions that can be derived from the principal right and left eigenvectors, let us briefly go through some properties of the eigenvector scoring function. Let us begin from the intuitive interpretation of the meaning of $\mathbf{R}(\mathbf{A})$. First of all, we can interpret each component of $\mathbf{A}$ as the number of ways to get from one node of a graph to another. Recall also that if we want to know the how many ways, i.e., total sum of weighted walks, there are to get from one node into another by going through k nodes, we get this information from the components of $\mathbf{A}^k$. The total number of walks originating from each node is given by $\mathbf{A}^k\mathbf{1}$. The asymptotic growth rate of the total number of walks is equal to $\lambda(\mathbf{A})$, and the asymptotic proportions of walks from each node are given by the components of $\mathbf{R}(\mathbf{A})$, see Cvetković et al. (1997). More specifically, if $\mathbf{R}^*$ is a (non-normalized) principal eigenvector, then $R_i(\mathbf{A}) = R_i^*(\mathbf{A})/\sum_i R_i^*(\mathbf{A})$ tells the proportion of all walks that begin from node i in the limit when k goes to infinity.

For stochastic matrices and tournament matrices the principal eigenvector has additional interpretations. For a stochastic matrix $\mathbf{R}(\mathbf{A})$ gives the stationary distribution of the corresponding Markov chain. When $\mathbf{A}$ is a tournament matrix with $a_{ij} + a_{ji} = 1$ for all $i \neq j$, the principal eigenvector expresses the probabilities of playing against each team in a sequential tournament where the winner gets to play against against randomly chosen opponent. Namely, when the prior probability of playing against each team is given by $\mathbf{R}(\mathbf{A})$, then the probabilities of being a winner are $\mathbf{AR}(\mathbf{A})$. When the probability of playing against any team is proportional to the winning probability of the team we get $\mathbf{R}(\mathbf{A})$ from $\mathbf{AR}(\mathbf{A})$ by scaling the latter probabilities.

One desirable property of $\mathbf{R}(\mathbf{A})$ is that the scores of nodes are unaffected by the way they are labeled. This property is known as anonymity. Formally, for any permutation π of nodes with matrix $\mathbf{P}_\pi$ it holds that

$$F_i(\mathbf{A}) = F_{\pi(i)}\left(\mathbf{P}_\pi \mathbf{A} \mathbf{P}_\pi^\top\right) \text{ for all } i = 1, \ldots, n.$$

The second feature of $\mathbf{R}(\mathbf{A})$, which is common for many other scoring functions, is that if nodes are in a symmetric position their scores are equal. Nodes i and j are symmetric when there is a permutation π such that $\pi(i) = j$ and $\pi(j) = i$, and $\mathbf{A} = \mathbf{P}_\pi \mathbf{A} \mathbf{P}_\pi^\top$. The property that symmetric nodes get the same score will be called symmetry.

The third property is that adding a node linked to itself with any non-negative weight and to other nodes with the number of edges to both directions being $\mathbf{F}(\mathbf{A})$, does not affect the relative scores of the original nodes. Formally, if we create a new matrix

$$\mathbf{A}^* = \begin{pmatrix} \mathbf{A} & \mathbf{R}(\mathbf{A}) \\ \mathbf{R}(\mathbf{A})^\top & \alpha \end{pmatrix},$$

where $\alpha > 0$, i.e., we add an "average" node into the graph, then $F_i(\mathbf{A}^*)/F_j(\mathbf{A}^*)$, $i, j = 1, \ldots, n$ remain the same. This property is called as invariance to adding an average node.

Proposition 2. *The scoring functions* $\mathbf{R}(\mathbf{A})$ *and* $\mathbf{L}(\mathbf{A})$ *satisfy anonymity, symmetry, and invariance to adding an average node.*

Proof. Permuting the order of players affects only the order of components of eigenvectors. Hence, anonymity holds for $\mathbf{R}$. Symmetry nodes is evident. The last property follows by observing that there is an eigenvector of the form $(\mathbf{R}(\mathbf{A}), f(\alpha)) \gg \mathbf{0}$ for

$$\begin{pmatrix} \mathbf{A} & \mathbf{R}(\mathbf{A}) \\ \mathbf{R}(\mathbf{A})^\top & \alpha \end{pmatrix}$$

corresponding to the eigenvalue which is the positive root of

$$(x - \alpha)[x - \lambda(\mathbf{A})] - \|\mathbf{R}(\mathbf{A})\|^2 = 0. \tag{5}$$

This quadratic equation is obtained from the eigenvalue equation

$$\begin{pmatrix} \mathbf{A} & \mathbf{R}(\mathbf{A}) \\ \mathbf{R}(\mathbf{A})^\top & \alpha \end{pmatrix} \begin{pmatrix} \mathbf{R}(\mathbf{A}) \\ f(\alpha) \end{pmatrix} = x \begin{pmatrix} \mathbf{R}(\mathbf{A}) \\ f(\alpha) \end{pmatrix}$$

by utilizing the fact that $\mathbf{AR}(\mathbf{A}) = \lambda(\mathbf{A})\mathbf{R}(\mathbf{A})$. Namely, this condition gives $(\lambda(\mathbf{A})+f(\alpha))\mathbf{R}(\mathbf{A}) = x\mathbf{R}(\mathbf{A})$, i.e., $f(\alpha) = x-\lambda(\mathbf{A})$, and $\|\mathbf{R}(\mathbf{A})\|^2+\alpha f(\alpha) = xf(\alpha)$.

It can be seen that Equation (5) has a positive root for any $\alpha \geq 0$. Let

$$\beta(\alpha) = \frac{\lambda(\mathbf{A}) + \alpha + \sqrt{[\lambda(\mathbf{A}) - \alpha]^2 + 4\|\mathbf{R}(\mathbf{A})\|^2}}{2}$$

denote this root. Then we have $f(\alpha) = \beta(\alpha) - \lambda(\mathbf{A})$. Because $\beta(\alpha) > (\lambda(\mathbf{A}) + \alpha + |\lambda(\mathbf{A}) - \alpha|)/2$, it holds that $\beta(\alpha) > \lambda(\mathbf{A})$ for all $\alpha \geq 0$, i.e., $f(\alpha) > 0$. By the Perron-Frobenius theorem there are no other positive eigenvectors than the principal one. Since all the components of $(\mathbf{R}(\mathbf{A}), f(\alpha))$ are positive it is the principal eigenvector. The deduction goes through for $\mathbf{L}(\mathbf{A})$ the same way. □

4. Zero-sum scoring with principal eigenvectors

In this section it will be assumed that the graph presenting the binary relations is directed. Consequently, **A** is not necessarily a symmetric matrix. This is typically the case for tournament matrices, where a link between two nodes reflects the wins (or losses). Especially, in tournaments we may want to utilize the information on both wins and losses to assess the players. For example, a player may have a significant number of both wins and losses. Hence, a centrality measure that is based only on wins may give a biased view on the actual strength of a player. The following example demonstrates this.

Example 2. Let us assume that there are four players. Player a has beaten players b and d but is defeated by player c, who has lost to the other two players. Player b has beaten player d. Hence, the tournament matrix is

$$\mathbf{A} = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}.$$

The corresponding graph is illustrated in Figure 2.

The usual points method would given equal score to players a, b, and c, and ranks them best. The eigenvector ranking is a, b, c, and d. Both players a and b have one loss; a has been beaten by c, and b has been beaten by a. It is worth noticing that a has lost to a relatively weaker player than c.

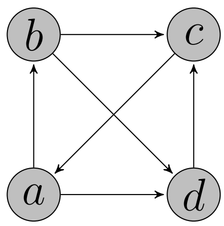

Figure 2: The graph in Example 2.

Hence, if we take into account also the losses in the scoring we would expect b to be higher ranked than a. This is also what will happen when ranking the players according to the zero-sum method; the ranking is b, a, c, and d.

4.1. Definition and interpretation

By utilizing the left and right eigenvectors it is possible to define a zero-sum scoring function as their difference. It will be called zero-sum spectral scoring function.

Definition 1. Let $\mathbf{A}$ be non-negative and irreducible. Zero-sum spectral scoring function is

$$\mathbf{F}(\mathbf{A}) = \mathbf{R}(\mathbf{A}) - \mathbf{L}(\mathbf{A}).$$

Let us now derive the zero-sum spectral scoring function by considering walks originating from a node and terminating to the same node. Again we can consider the components of $\mathbf{A}$ as numbers of ways to get from one node to another. Recall that the number of walks of length k that start from i and terminate to j are obtained by taking the component a_{ij}^k of $\mathbf{A}^k$. Moreover, the number of walks originating from j and terminating to i are given by the component $a_{ij}^{(k\top)}$ of $\mathbf{A}^\top$ to power k, i.e., $\mathbf{A}^{k\top}$. The difference

$$a_{ij}^k - a_{ij}^{k\top}$$

reflects the net flow of walks from i to j. Observe that the corresponding difference for node j is simply

$$a_{ij}^{k\top} - a_{ij}^k.$$

To assess the relative importance of node i we can simply sum over the differences

$$a_{ij}^k - a_{ij}^{k\top}.$$

In matrix notation we get the scores

$$\left[\mathbf{A}^k - \mathbf{A}^{k\top}\right]\mathbf{1}.$$

The following result is a direct consequence of the Perron-Frobenius theorem, see, e.g., Meyer (2000).

Proposition 3. *Let* $\mathbf{A}$ *be a primitive matrix and let*

$$\alpha(\mathbf{A}) = \left[\mathbf{R}(\mathbf{A})^\top \mathbf{L}(\mathbf{A})\right]^{-1/2}.$$

Then

$$\frac{\left[\mathbf{A}^k - \mathbf{A}^{k\top}\right] \mathbf{1}}{\lambda^k(\mathbf{A})} \to \alpha^2(\mathbf{A})\mathbf{F}(\mathbf{A}) \text{ as } k \to \infty.$$

Proof. The matrices $\mathbf{A}$ and $\mathbf{A}^\top$ have the same principal eigenvalue $\lambda(\mathbf{A})$. By the Perron-Frobenius theorem

$$\frac{\mathbf{A}^k}{\lambda^k(\mathbf{A})} \to R^*(\mathbf{A})L^*(\mathbf{A})^\top \text{ and } \frac{\mathbf{A}^{k\top}}{\lambda^k(\mathbf{A})} \to L^*(\mathbf{A})R^*(\mathbf{A})^\top$$

as $k \to \infty$. It is assumed that $R^*(\mathbf{A})^\top L^*(\mathbf{A}) = 1$, and $R^*(\mathbf{A})$ is parallel to $\mathbf{R}(\mathbf{A})$ and $L^*(\mathbf{A})$ is parallel to $\mathbf{L}(\mathbf{A})$. If we take $R^*(\mathbf{A}) = \alpha(\mathbf{A})\mathbf{R}(\mathbf{A})$ and $L^*(\mathbf{A}) = \alpha(\mathbf{A})\mathbf{L}(\mathbf{A})$ we get the result because then $R^*(\mathbf{A})L^*(\mathbf{A})^\top\mathbf{1} = R^*(\mathbf{A})\alpha(\mathbf{A}) = \alpha^2(\mathbf{A})\mathbf{R}(\mathbf{A})$ and $L^*(\mathbf{A})R^*(\mathbf{A})^\top\mathbf{1} = \alpha^2(\mathbf{A})\mathbf{L}(\mathbf{A})$. □

If $\mathbf{A}$ is not primitive, it is still possible to find the eigenvectors but the above explanation in terms of walks no longer works.

Recall from the previous section that $\mathbf{R}(\mathbf{A})$ gives the asymptotic proportion of length k walks that start from each node of the graph. The vector $\mathbf{L}(\mathbf{A})$ on the other hand gives the asymptotic proportions of walks that end to each node. Hence, the scores are simply differences of these proportions.

If $\mathbf{A}$ is a matrix corresponding to results of a tournament, then its components could represent points that a player has gained from other players. In case of tournaments the components of $\mathbf{F}$ are proportional to the asymptotic growth rates of the net of incoming and outgoing points for each player.

4.2. Properties of the zero-sum scoring function

Let us consider an arbitrary zero-sum scoring function $\mathbf{F}$. It is additive if $\mathbf{F}(\mathbf{A}) = \mathbf{F}^i(\mathbf{A}) - \mathbf{F}^o(\mathbf{A})$. Intuitively, $\mathbf{F}^i(\mathbf{A})$ represents the inflows of connections to each node while $\mathbf{F}^o$ represents the outflows. For instance, when considering a tournament $\mathbf{F}^i$ contains the inflows of points due to victories while $\mathbf{F}^o$ is composed of the outflows of points due losses. Recall that if $\mathbf{A}$ contains the data on wins for each pair, i.e., $a_{ij} > 0$ indicates that i has beaten j, then $\mathbf{A}^T$ contains the losses. Hence, it is reasonable to assume that $\mathbf{F}^i(\mathbf{A}) = \mathbf{F}^o(\mathbf{A}^\top)$. Consequently, nodes are equally scored if and only if their scores are all zeros.

It follows immediately from the definition of the zero-sum spectral scoring function that it is additive and the inflow and outflow components are related to each other as described above, i.e., by setting $\mathbf{F}^i(\mathbf{A}) = \mathbf{R}(\mathbf{A})$ and $\mathbf{F}^o(\mathbf{A}) = \mathbf{L}(\mathbf{A})$ it holds that $\mathbf{F}^i(\mathbf{A}) = \mathbf{F}^o(\mathbf{A}^\top)$. As a consequence we can make the following observation.

Remark 2. For any irreducible $\mathbf{A}$ it holds that $\sum_j F_j(\mathbf{A}) = 0$. The players scores are equal if and only if $F_j(\mathbf{A}) = 0$ for all $j = 1, \ldots, n$.

The zero-sum scoring function obviously inherits most of the properties of the principal eigenvector. In particular, it has the properties (A2)–(A4). Moreover, due to its zero-sum nature it gives zero scores, when $\mathbf{A}$ is symmetric. For a tournament matrix this would mean that each player has equally many wins and losses from each player. This result is a direct corollary of the above remark. Moreover we can add any symmetric matrix to $\mathbf{A}$ without affecting the scores, which follows from Remark 1. Namely, left and right eigenvectors are the same for symmetric matrices. More generally, when we add another symmetric matrix to $\mathbf{A}$ the scores remain the same. For a tournament matrix this would mean that players cannot improve their positions by agreeing to play pairs of matches where wins and losses are equal for each player.

Remark 3. If $\mathbf{A}$ is symmetric, then it holds that $F_i(\mathbf{A}+\mathbf{B}) = F_j(\mathbf{A})$ for all $i, j = 1, \ldots, n$, and for all symmetric non-negative matrices $\mathbf{B}$.

By utilizing Proposition 1 for $\mathbf{R}$ it is possible to characterize the zero-sum spectral scoring function. By definition $\mathbf{F}(\mathbf{A})$ is $\mathbf{F}(\mathbf{A}) = \mathbf{R}(\mathbf{A}) - \mathbf{L}(\mathbf{A})$. The functions $\mathbf{R}(\mathbf{A})$ and $\mathbf{L}(\mathbf{A})$ are vectors with strictly positive components for all irreducible non-negative matrices $\mathbf{A}$, which follows from the Perron-Frobenius theorem. As seen in the previous section $\mathbf{R}$ satisfies the axioms (A1)–(A6). Consequently, these properties fully characterize the zero-sum spectral scoring function.

Proposition 4. *The zero-sum spectral scoring function is the unique scoring function which satisfies the following properties for any non-negative irreducible matrix* $\mathbf{A}$*.*

1. $\mathbf{F}(\mathbf{A}) = \mathbf{F}^i(\mathbf{A}) - \mathbf{F}^o(\mathbf{A})$*,*
2. $\mathbf{F}^i(\mathbf{A}) = \mathbf{F}^o(\mathbf{A}^\top)$*,*
3. $\mathbf{F}^i(\mathbf{A})$ *satisfies (A1)–(A6).*

Proof. The last assumption implies that $\mathbf{F}^i(\mathbf{A})$ is uniquely defined as $\mathbf{R}(\mathbf{A})$. The second assumption on the other hand assures that $\mathbf{F}^o(\mathbf{A})$ is the left eigenvector $\mathbf{L}(\mathbf{A})$. Hence, the unique scoring function having additive structure and components that satisfy the second and third assumptions is the zero-sum spectral scoring function. □

4.3. Other zero-sum scoring functions

The idea of zero-sum scoring functions has been previously utilized in defining scoring functions. For instance, the point method, or the Copeland scoring function, gives each player scores equal to the out degree minus the in degree, i.e., the scores are given by

$$\left[\mathbf{A} - \mathbf{A}^\top\right] \mathbf{1}.$$

Rubinstein (1980) characterizes this scoring function for tournaments. In the David's method the scores are given by

$$\left[\mathbf{A}^2 - \mathbf{A}^{2\top} + \mathbf{A} - \mathbf{A}^\top\right] \mathbf{1}.$$

More generally we can define a whole family $\mathbf{F}^k$, $k = 1, 2, \ldots$, of scoring functions of this kind. Let us first denote

$$\mathbf{G}^k(\mathbf{A}) = \sum_{i=1}^{k} \left[\mathbf{A}^i - \mathbf{A}^{i\top}\right] \mathbf{1}.$$

For each $k \geq 1$ we can define a scoring function

$$\mathbf{F}^k(\mathbf{A}) = \frac{\mathbf{G}^k(\mathbf{A})}{\max_i G_i^k(\mathbf{A})},$$

when $\max_i G_i^k(\mathbf{A}) \neq 0$. This normalization guarantees that $\max_i F_i^k(\mathbf{A}) = 1$. Moreover, multiplying $\mathbf{A}$ with a scalar does not affect the scoring. If $\max_i G_i^k(\mathbf{A}) = 0$ we can set $\mathbf{F}^k(\mathbf{A}) = \mathbf{0}$.

A zero-sum scoring function called the net power function (Herings et al., 2005) is defined as

$$\mathbf{F}(\mathbf{A}) = \mathbf{F}^p(\mathbf{A}) - \mathbf{F}^p\left(\mathbf{A}^\top\right),$$

where the power function $\mathbf{F}^p$ is

$$\mathbf{F}^p(\mathbf{A}) = \frac{1}{n}\left[\mathbf{I} - \frac{1}{n}\mathbf{A}\right]^{-1} \mathbf{A}\mathbf{1}.$$

The zero-sum spectral scoring function can be related to a weighted scoring function in a similar fashion as principal eigenvector relates to the Katz-Hubbell index: the principal eigenvector is the limit of Katz-Hubbell index. Analogously it can be shown that the zero-sum spectral scores are obtained as a limit of a weighted zero-sum scoring function defined below.

Definition 2. Let $w < 1/|\lambda(\mathbf{A})|$. The weighted zero-sum scoring function is defined as

$$\mathbf{F}^w(\mathbf{A}) = \left[\mathbf{I} - w\mathbf{A}\right]^{-1} \mathbf{1} - \left[\mathbf{I} - w\mathbf{A}^\top\right]^{-1} \mathbf{1}. \tag{6}$$

Another way to define the weighted zero-sum scoring function is to set

$$\mathbf{F}^w(\mathbf{A}) = \sum_{k=0}^{\infty} \left[w^k \left(\mathbf{A}^k - \mathbf{A}^{k\top} \right) \right]. \tag{7}$$

The weight w is called the damping factor, and when $1/|\lambda(\mathbf{A})|$ the infinite sum is well defined. Moreover, it holds that

$$\sum_{k=0}^{\infty} \left(w^k \mathbf{A}^k \right) = [\mathbf{I} - w\mathbf{A}]^{-1},$$

when $w < 1/|\lambda(\mathbf{A})|$. Hence, Equation (7) leads to exactly the same scores as Equation (6).

The weighted scoring function

$$\mathbf{F}(\mathbf{A}) = \sum_{k=0}^{\infty} \left(w^k \mathbf{A}^k \right) = [\mathbf{I} - w\mathbf{A}]^{-1} \mathbf{1}$$

is known as the Katz-Hubbell index (Katz, 1953; Hubbell, 1965), see also Bonacich and Lloyd (2001) for a related measure where the vector $\mathbf{1}$ is replaced with a vector reflecting the exogenous sources of status. Bonacich (1987) defines a family of related measures where the damping factor may be negative.

Proposition 5. *When w converges to $1/\lambda(\mathbf{A})$ from below, $\mathbf{F}^w(\mathbf{A})$ converges to $\mathbf{F}(\mathbf{A})$.*

Proof. It is well-known that $\mathbf{R}(\mathbf{A})$ is the limit of

$$[\mathbf{I} - w\mathbf{A}]^{-1}\mathbf{1},$$

when w goes to $1/\lambda(\mathbf{A})$, see, e.g., Thompson (1958) and Bonacich (1987, 1991). The analogous result holds for $\mathbf{A}^\top$. Hence, the result follows. $\square$

5. Axiomatization of the Bonacich-Kleinberg scoring function

5.1. The HITS method and simultaneous scoring of groups and individuals

Assessing the authoritativeness of webpages is an important task of Internet search engines. Kleinberg (1999) has proposed the notion of authoritativeness of a webpage based on the relationship between the webpages and the hub pages that join them together in the network structure. To clarify the idea of the HITS method let us consider the graph in example 1 and reinterpret the nodes as websites.

Example 3. In this case we have five websites; a, b, c, d, e. Each node corresponds to a website while an edge corresponds to hyperlink between two websites. The question is now how to assess the authoritativeness of a website. Each website serves also as a hub that connects other websites. Hence, we may also want to assess the importance of a website as a hub that connects links to other webpages. The more a website has links the more relevant it is as a hub. In principle we can define five groups corresponding to each hub. For example, website b belongs to the group corresponding to a because it is a neighbor of a. Respectively, each row of the adjacency matrix corresponds to a hub while the columns correspond to websites.

Let $\mathbf{A}$ be the adjacency matrix of the network. The authoritativeness of nodes and hubs are obtained in a consistent manner from the principal eigenvectors of $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$, respectively. Observe that the authoritativeness of a webpage is related to how relevant hubs it is connected to. This method of assessing the authoritativeness of webpages is known as the HITS (hypertext induced topic search) and it has become rather popular in ranking web-pages among another eigenvector based technique known as the Page Rank (Brin and Page, 1998).

The HITS method can be seen as a special case of the method introduced by Bonacich (1991) for assessing individual and group rankings simultaneously. Let us assume that there are m groups and n individuals who are the possible members of the groups. The $m \times n$ matrix $\mathbf{A}$ contains the weights associated with each individual in each group. In the Bonacich's method the scores of individuals are given by $\mathbf{R}(\mathbf{A}^\top\mathbf{A})$ and the scores of groups are given by $\mathbf{R}(\mathbf{A}\mathbf{A}^\top)$. Evidently, HITS corresponds to the specific case where the groups are simply the neighbors of each node. In the following we call the method as the Bonacich-Kleinberg method. Its main feature is that the group scores are a weighted sum of the scores of the members of the group.

The Bonacich-Kleinberg method can be used for preference aggregation in a social choice setup. In this case a number of agents have preferences over a set of alternatives and the question is on aggregating these preferences into a social welfare order. Now the rows of the matrix $\mathbf{A}$ represent the agents and columns represent the alternatives. The elements a_{ij} describes the preference of agent i for alternative j. Another application of the Bonacich-Kleinberg method in the social choice context is to use it to assess a set of alternatives together with a set of characteristics related to them. Let us demonstrate this with a simple example.

Example 4. Let us assume that in Example 1, the vertices represent alternatives, and the edges represent binary relations over them. Moreover, let us assume that alternatives a and b involve a certain decision D_1 to be made, while b and d require decision D_2 and the alternative c requires decision D_3 and D_1, the alternative e requires decision D_3. The question is now on ranking both the alternatives and the decisions simultaneously. Basically, we can divide the alternatives into groups corresponding to D_1, D_2, and D_3. An alternative belongs to a group if it involves the decision corresponding to a group.

Let us now create a matrix $\mathbf{M}$ where each row corresponds to a decision and each column corresponds to an action. Let us also give a weight to each alternative as follows. An alternative gets one point from each edge to another node. For instance, a gets one point from b, b gets one point from e, and so on. Since an alternative may belong to several groups, i.e., require several decisions, we give an alternative the weight as a member of one group by dividing the points by the number of groups it belongs to. This gives us

$$\mathbf{M} = \begin{array}{c} \\ D_1 \\ D_2 \\ D_3 \end{array} \begin{array}{c} \begin{array}{ccccc} a & b & c & d & e \end{array} \\ \begin{pmatrix} 1 & 1/2 & 1.5/2 & 0 & 0 \\ 0 & 1/2 & 0 & 2 & 0 \\ 0 & 0 & 1.5/2 & 0 & 1.5 \end{pmatrix} \end{array}.$$

Now we are ready to score both the decisions and the alternatives simultaneously. The Bonacich-Kleinberg method would give the ranking of decisions D_2, D_1, D_3, and the ranking of alternatives d, b, c, a, e.

5.2. Axiomatization

To define the axioms for the Bonacich-Kleinberg scoring function we need the operations of adding groups and the power of $m \times n$ matrix. Recall that the rows of $\mathbf{A}$ represent groups and the columns represent individuals.

Let $\mathbf{A}$ be a matrix corresponding to m groups and n individuals. If $\mathbf{B}$ is a matrix corresponding to k groups and n individuals we can append $\mathbf{A}$ with $\mathbf{B}$ to obtain

$$\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}.$$

We denote this operation as $\mathbf{A} \boxplus \mathbf{B}$.

The powers of $\mathbf{A}$ are defined as follows: $\mathbf{A}^{(1)} = \mathbf{A}$, $\mathbf{A}^{(2)} = \mathbf{A}^\top\mathbf{A}$, $\mathbf{A}^{(3)} = \mathbf{A}\mathbf{A}^\top\mathbf{A}, \ldots$. Note that for k even $\mathbf{A}^{(k)}$ is $n \times n$ matrix and for k odd $\mathbf{A}^{(k)}$ is $m \times n$ matrix.

We are now ready to describe the axioms needed in characterizing the Bonacich-Kleinberg scoring function. Below a scoring function $\mathbf{F}$ maps any $m \times n$ matrix for which $\mathbf{A}^\top\mathbf{A}$ belongs to $\mathcal{M}$ into $\mathbb{R}^n$.

(A1') $\mathbf{F}(\mathbf{A} \boxplus \beta\mathbf{P}) = \mathbf{F}(\mathbf{A})$ for any $\beta \geq 0$ and $n \times n$ permutation matrix $\mathbf{P}$,
(A2') $\mathbf{F}(\mathbf{A}^{(2k)}) = \mathbf{F}(\mathbf{A})$ for all $k = 1, 2, \ldots$.

(A3') Let $\mathbf{A}$ and $\mathbf{B}$ have the same dimensions. If $\mathbf{F}(\mathbf{A}) = \mathbf{F}(\mathbf{B})$ then $\mathbf{AF}(\mathbf{A}) = \beta\mathbf{BF}(\mathbf{B})$ for some $\beta > 0$.

The first axiom means that the scores are invariant for adding dummy groups, i.e., adding groups corresponding to each individual with only that individual as the member. This means that the scores cannot be manipulated by adding identical groups for each individual. The second axiom says that we get the same score when we form an adjacency matrix $\mathbf{A}^\top\mathbf{A}$ to individuals and consider the resulting walks of any fixed length. The product $\mathbf{A}^\top\mathbf{A}$ means that the strength of the connection between two individuals is set equal to the sum of products of the weights in which the two individuals belong to each group. The last axiom is analogous to (A5) in Section 2. The difference to (A5) is that $\mathbf{A}$ and $\mathbf{B}$ should have the same dimensions. The interpretation of is that if the scores of individuals corresponding to two matrices with same dimensions are the same, then we get the same scores for the groups if we compute them by summing the weight-score products of all individuals belonging to each group and normalize the resulting vectors. As will be seen later, this is indeed how the scores of the groups and individuals are related in the Bonacich-Kleinberg method.

The main result for the Bonacich-Kleinberg scoring function is that it is the unique scoring function that satisfies the above axioms together with (A1), (A2), and (A6).

Proposition 6. $\mathbf{F}(\mathbf{A}) = \mathbf{R}(\mathbf{A}^\top\mathbf{A})$ *is the unique function that satisfies axioms (A1')–(A3'), (A1), (A2), and (A6).*

Proof. The proof is similar to the proof of Proposition 1. By (A2) and (A2') we have

$$\mathbf{F}^k = \mathbf{F}\left(\mathbf{A}^{(2k)}/\lambda\left(\mathbf{A}^\top\mathbf{A}\right)\right) = \mathbf{F}(\mathbf{A}),$$

for all $k = 1, 2, \ldots$, which implies that the limit of the sequence $\{\mathbf{F}^k\}_k$ is $\mathbf{F}(\mathbf{A})$. By (A3') we have

$$\mathbf{A}^{(2k)}\mathbf{F}(\mathbf{A})/\lambda(\mathbf{A}^\top\mathbf{A}) = \beta_k\mathbf{A}^\top\mathbf{A}\mathbf{F}(\mathbf{A}) \tag{8}$$

for $\beta_k > 0$ for all $k = 1, 2, \ldots$. The Perron-Frobenius theorem, i.e., Equation (1) gives

$$\mathbf{A}^{(2k)}/\lambda\left(\mathbf{A}^\top\mathbf{A}\right) \to \alpha^2\left(\mathbf{A}^\top\mathbf{A}\right)\mathbf{R}\left(\mathbf{A}^\top\mathbf{A}\right)\mathbf{L}\left(\mathbf{A}^\top\mathbf{A}\right)^\top$$

when k goes to infinity. In particular both sides of Equation (8) are bounded which guarantees that $\{\beta_k\}_k$ is a bounded sequence that has a convergent subsequence. Taking the limit of both sides of Equation (8) corresponding to this subsequence we get

$$\alpha^2 \left(\mathbf{A}^\top\mathbf{A}\right) \mathbf{R} \left(\mathbf{A}^\top\mathbf{A}\right) \mathbf{L} \left(\mathbf{A}^\top\mathbf{A}\right)^\top \mathbf{F}(\mathbf{A}) = \bar{\beta}\mathbf{A}^\top\mathbf{A}\mathbf{F}(\mathbf{A}),$$

for $\bar{\beta} \geq 0$. If either $\bar{\beta} = 0$ or $\mathbf{L}\left(\mathbf{A}^\top\mathbf{A}\right)^\top \mathbf{F}(\mathbf{A}) = 0$, then $\mathbf{L}\left(\mathbf{A}^\top\mathbf{A}\right)$ and $\mathbf{F}(\mathbf{A})$ would be orthogonal which cannot be the case since $\mathbf{L}\left(\mathbf{A}^\top\mathbf{A}\right), \mathbf{F}(\mathbf{A}) \gg \mathbf{0}$. Hence, there is $\gamma > 0$ such that $\mathbf{R}\left(\mathbf{A}^\top\mathbf{A}\right) = \gamma\mathbf{A}^\top\mathbf{A}\mathbf{F}(\mathbf{A})$, which can also be written as

$$\mathbf{A}^\top\mathbf{A}\mathbf{R}\left(\mathbf{A}^\top\mathbf{A}\right)/\lambda\left(\mathbf{A}^\top\mathbf{A}\right) = \gamma\mathbf{A}^\top\mathbf{A}\mathbf{F}(\mathbf{A}),$$

because $\mathbf{R}\left(\mathbf{A}^\top\mathbf{A}\right)$ is an eigenvector of $\mathbf{A}^\top\mathbf{A}$. By (A1') we can assume that $\mathbf{A}^\top\mathbf{A}$ is invertible. If it was not we could add $\beta\mathbf{P}$, to $\mathbf{A}$ to make $(\mathbf{A} \boxplus \beta\mathbf{P})^\top(\mathbf{A} \boxplus \beta\mathbf{P}) = \mathbf{A}^\top\mathbf{A} + \beta^2\mathbf{I}$ an invertible matrix. Because $\beta^2\mathbf{I}$ is a diagonal matrix, it follows that $(\mathbf{A} \boxplus \beta\mathbf{P})^\top(\mathbf{A} \boxplus \beta\mathbf{P})$ can be turned into a diagonally dominant matrix by choosing sufficiently large β. Such matrices are invertible. Hence, we obtain that $\mathbf{R}\left(\mathbf{A}^\top\mathbf{A}\right)$ is parallel to $\mathbf{F}(\mathbf{A})$. Axiom (A6) gives $\mathbf{F}(\mathbf{A}) = \mathbf{R}\left(\mathbf{A}^\top\mathbf{A}\right)$, which concludes the proof. □

The Bonacich-Kleinberg scoring function has some additional properties which are worth mentioning. First, the Bonacich-Kleinberg function is anonymous in the sense that the order of individuals, i.e., permutation of the rows only change the order of the scores. Moreover, changing the order of groups does not affect the scores of individuals. It is also symmetric in the sense that if a permutation of rows keeps the matrix the same, the scores remain the same. In other words, symmetric individuals have the same scores. Anonymity and symmetry follow directly from the properties of the principal eigenvector.

A natural consistency requirement for scoring groups and individuals is that the scores of groups are obtained from the scores of individuals by summing the scores with weights given by $\mathbf{A}$. This is property holds for the Bonacich-Kleinberg function. If we append the matrix $\mathbf{A}$ with $\mathbf{F}(\mathbf{A})$, i.e., we add a group where each individuals weight is $\mathbf{F}(\mathbf{A})$, the scores remain unchanged. On the other hand, if we add an individual with weight $\mathbf{A}\mathbf{F}(\mathbf{A})$ to each group, the relative scores of the original individuals remain unchanged. Intuitively, adding an "average group" or an "average individual" does not affect the scores. The Bonacich-Kleinberg function is not, however, the only one having this property. For example the function that assigns a vector $(1/n, \ldots, 1/n)$ to a matrix has this property as well. The latter two properties are shown below.

Proposition 7. *The Bonacich-Kleinberg scoring function has the following properties:*

1. $\mathbf{AF}(\mathbf{A}) = \rho(\mathbf{A})\mathbf{F}(\mathbf{A}^\top)$ *for some* $\rho(\mathbf{A}) > 0$,
2. $\mathbf{F}\left(\mathbf{A} \boxplus \mathbf{F}(\mathbf{A})\right) = \mathbf{F}(\mathbf{A})$, *and*
3. $F_i\left(\left(\mathbf{A} \quad \mathbf{AF}(\mathbf{A})\right)\right)/F_j\left(\left(\mathbf{A} \quad \mathbf{AF}(\mathbf{A})\right)\right) = F_i(\mathbf{A})/F_j(\mathbf{A})$ *for all* $i, j = 1, \ldots, n$.

Proof. The principal eigenvector of $\mathbf{AA}^\top$ satisfies

$$\mathbf{AA}^\top\mathbf{R}\left(\mathbf{AA}^\top\right) = \lambda\left(\mathbf{AA}^\top\right)\mathbf{R}\left(\mathbf{AA}^\top\right). \tag{9}$$

If we replace $\mathbf{R}(\mathbf{AA}^\top)$ with $\mathbf{AR}(\mathbf{A}^\top\mathbf{A})$ and use the fact that $\mathbf{R}(\mathbf{A}^\top\mathbf{A})$ is an eigenvector of $\mathbf{A}^\top\mathbf{A}$ we see that the right hand side of Equation (9) becomes $\lambda(\mathbf{A}^\top\mathbf{A})\mathbf{AR}(\mathbf{A}^\top\mathbf{A})$, which implies that $\mathbf{AR}(\mathbf{A}^\top\mathbf{A})$ is an eigenvector of $\mathbf{AA}^\top$ corresponding to the eigenvalue $\lambda(\mathbf{A}^\top\mathbf{A})$. Because $\mathbf{A}$ has only non-negative components and $\mathbf{R}(\mathbf{A}^\top\mathbf{A}) \gg \mathbf{0}$ we have $\mathbf{AR}(\mathbf{A}^\top\mathbf{A}) \gg \mathbf{0}$. Perron-Frobenius theorem assures that $\mathbf{AR}(\mathbf{A}^\top\mathbf{A})$ is the principal eigenvector of $\mathbf{AA}^\top$. Hence, the scores of the groups are obtained by scaling $\mathbf{AR}(\mathbf{A}^\top\mathbf{A})$ with a scalar such that the resulting vector satisfies (A6). This proves the first result.

The second result follows by verifying that $\mathbf{F}(\mathbf{A})$ is the eigenvector of $\left(\mathbf{A} \boxplus \mathbf{F}(\mathbf{A})\right)^\top\left(\mathbf{A} \boxplus \mathbf{F}(\mathbf{A})\right)$ corresponding to the eigenvalue $(\lambda(\mathbf{A}^\top\mathbf{A}) + \|\mathbf{F}(\mathbf{A})\|^2)$. By the Perron-Frobenius theorem there are no other positive eigenvectors for $\mathbf{A} \boxplus \mathbf{F}(\mathbf{A})$ than the principal eigenvector. Hence, $\mathbf{F}(\mathbf{A})$ is the principal eigenvector.

The third result follows by verifying that the vector

$$\left(\mathbf{F}(\mathbf{A}), \|\mathbf{AF}(\mathbf{A})\|^2/\lambda(\mathbf{A}^\top\mathbf{A})\right)$$

is an eigenvector of the matrix $\left(\mathbf{A} \quad \mathbf{AF}(\mathbf{A})\right)$ corresponding to the eigenvalue $\|\mathbf{AF}(\mathbf{A})\|^2 + \lambda(\mathbf{A}^\top\mathbf{A})$. Again, we can argue that the aforementioned vector is the principal eigenvector. Normalizing the vector does not change the relative scores of first n individuals, i.e., the original individuals. □

References

Altman, A., Tennenholtz, M., 2005. Ranking systems: The PageRank axioms. In: Proceedings of the 6th ACM conference on Electronic commerce (EC-05). ACM Press, New York, pp. 1–8.

Bonacich, P., 1987. Power and centrality: a family of measures. American Journal of Sociology 92, 1170–1182.

Bonacich, P., 1991. Simultaneous group and individual centralities. Social Networks 13, 155–168.

Bonacich, P., and P. Lloyd, 2001. Eigenvector-like measures of centrality for asymmetric relations. Social Networks 23, 191-201.

Borm, P., R. van den Brink, and M. Slikker, 2002. An iterative procedure for evaluating digraph competitions. Annals of Operations Research 109, 61-75.

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems 30, 107–117.

Cvetković, D., Rowlinson, P., Simić, 1997. Eigenspaces of graphs. Cambridge University Press.

Daniels, H. E., 1969. Round-robin tournament scores. Biometrika 56, 295–299.

David, H. A., 1987. Ranking from unbalanced paired-comparison data. Biometrika 74, 432–436.

David, H. A., 1988. The method of paired comparisons, 2nd Edition. Charles Griffin and Company, London.

Echenique, F., Fryer, R. G., 2007. A measure of segregation based on social interactions. Quarterly Journal of Economics 122(2), 441–485.

Henriet, D., 1985. The Copeland choice function: an axiomatic characterization. Social Choice and Welfare 2, 49–63.

Herings, J. J., van der Laan, G., Talman, D., 2005. The positional power of nodes in digraphs. Social Choice and Welfare 24, 439–454.

Hu, X., and L.S. Shapley, 2003. On authority distributions in organizations: equilibrium. Games and Economic Behavior 45, 132-152.

Hubbell, C. H., 1965. An input-output approach to clique identification. Sociometry 28 (4), 377–399.

Katz, L., 1953. A new status index derived from sociometric analysis. Psychometrika 18 (1), 39–43.

Kendall, M. G., 1955. Further contributions to the theory of paired comparisons. Biometrics 11 (1), 43–62.

Kleinberg, J. M., 1999. Authoritative sources in a hyperlinked environment. Journal of the ACM 46 (5), 604–632.

Laslier, J.-F., 1997. Tournament solutions and majority voting. Springer-Verlag, Berlin, Heidelberg, New York.

Meyer, C., 2000. Matrix analysis and applied linear algebra. SIAM.

Moon, J. W., Pullman, N. J., 1970. On generalized tournament matrices. SIAM Review 12, 384–399.

Palacios-Huerta, I., Volij, O., 2004. The measurement of intellectual influence. Econometrica 72 (3), 963–977.

Pinski, G., Narin, F., 1976. Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. Information Processing and Management 12 (5), 297–312.

Rubinstein, A., 1980. Ranking the participants in a tournament. SIAM Journal on Applied Mathematics 38, 108–111.

Seeley, J. R., 1949. The net of reciprocal influence: A problem in treating sociometric data. Canadian Journal of Psychology 3, 234–240.

Slutzki, G., Volij, O., 2005. Ranking participants in generalized tournaments. International Journal of Game Theory 33 (2), 255–270.

Slutzki, G., Volij, O., 2006. Scoring of webpages and tournaments — axiomatizations. Social Choice and Welfare 26 (1), 75–92.

Thompson, G. L., 1958. Lectures on game theory, Markov chains and related topics. Sandia Corporation Monograph SCR-11, Washington, D. C.

van den Brink, J., Gilles, R., 2000. Measuring domination in directed networks. Social Networks 22, 141–157.

Wei, T. H., 1952. The algebraic foundations of ranking theory. Cambridge University Press, London.

The **Aboa Centre for Economics (ACE)** is a joint initiative of the economics departments of the Turku School of Economics at the University of Turku and the School of Business and Economics at Åbo Akademi University. ACE was founded in 1998. The aim of the Centre is to coordinate research and education related to economics.

Contact information: Aboa Centre for Economics, Department of Economics, Rehtorinpellonkatu 3, FI-20500 Turku, Finland.

www.ace-economics.fi

ISSN 1796-3133