Nonlinear eigenvector centrality

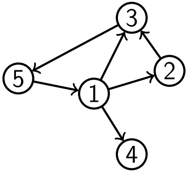
Francesco Tudisco
Gran Sasso Science Institute (Italy)

SIAM LA21 Minitutorial
Applied Nonlinear Perron-Frobenius Theory

May 18, 2021

Centrality

Given a graph $G = (V, E)$, define a function that quantifies the "importance" of nodes, based only on the topology of the graph.



$$A_{ij} = \begin{cases} 1 & i \to j \\ 0 & \text{otherwise} \end{cases}$$

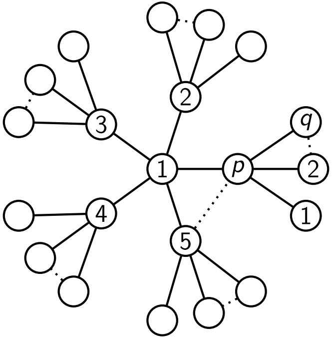
Basic example: count the degree.

$$d_i^{in} = \sum_{j:j\to i} 1 = (A^\top \mathbb{1}) \qquad \text{and} \qquad d_i^{out} = \sum_{j:i\to j} 1 = (A\mathbb{1})_i$$

If G is not directed $\boldsymbol{d}^{in} = \boldsymbol{d}^{out} = \boldsymbol{d}$.

Starred star example

Degree fails to capture global structure.



If $q > p$ degree assigns larger centrality to leaves than the center of the star.

Mutual reinforcement 1

Bonacich index: prototype of mutually reinforcing property

"Importance of nodes is proportional to the importance of their neighbors"

$$u_i \propto \sum_{j:j \to i} u_j = \sum_{j=1}^{n} A_{ji} u_j \qquad \text{that is} \qquad \boldsymbol{u} \propto A^\top \boldsymbol{u}$$

$\Longrightarrow$ $\boldsymbol{u}$ must be the **Perron eigenvector** of $A^\top$.

Mutual reinforcement 2

HITS: two mutual reinforcing indices for hubs (information spreaders) and authorities (information carriers)

"Important hubs point to relevant authorities; relevant authorities are pointed by important hubs"

$$\begin{cases} a_i \propto \sum_{j:j\to i} h_j = \sum_{j=1}^{n} A_{ji} h_j \\ h_i \propto \sum_{j:i\to j} a_j = \sum_{j=1}^{n} A_{ij} a_j \end{cases} \quad \text{that is} \quad \boxed{\begin{cases} \boldsymbol{a} \propto A^\top \boldsymbol{h} \\ \boldsymbol{h} \propto A\boldsymbol{a} \end{cases}}$$

$\Longrightarrow$ $\boldsymbol{a}$ and $\boldsymbol{h}$ must be the (left and right) **Perron singular vectors** of A.

Drawbacks

Centrality scores based on eigenvectors/singular vectors of matrices are among the most popular and useful.

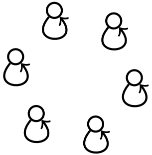
We identify two drawbacks:

- they are constrained to **linear proportionality relations**
- they **may be not well-defined**, even for simple graphs

Nonlinear Perron eigenvectors/singular vectors allow much greater flexibility and identify the “correct” centrality

Two illustrative examples

Illustrative example 1

$G =$  $\quad A = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$

Any $\boldsymbol{x} \succ 0$ is a Bonacich centrality $A\boldsymbol{x} = \boldsymbol{x}$.

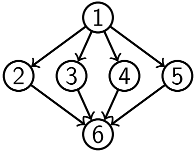
There is a unique nonnegative solution $\boldsymbol{x}^*$ of

$$Af(\boldsymbol{x}) = \lambda \boldsymbol{x}, \qquad f(\boldsymbol{x}) = \boldsymbol{x}^{1-\varepsilon}$$

for any $\varepsilon \in (0, 1)$, and $\boldsymbol{x}^* = (1, 1, \ldots, 1)$.

Proof: The mapping $F(\boldsymbol{x}) = Af(\boldsymbol{x})$ is homogeneous of degree $\theta = 1 - \varepsilon$.

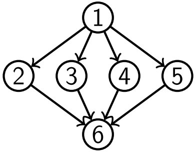
Illustrative example 2



$$A = \begin{bmatrix} & 1 & \cdots & 1 & \\ & & & & 1 \\ & & & & \vdots \\ & & & & 1 \\ & & & & \end{bmatrix}$$

Any $\boldsymbol{h} = (\alpha, \beta, \dots, \beta, 0)$ and $\boldsymbol{a} = (0, \gamma, \dots, \gamma, \delta)$ with $\alpha, \beta, \gamma, \delta > 0$ are hub and authority centralities $A^\top \boldsymbol{h} = \lambda \boldsymbol{a}$, $A\boldsymbol{a} = \lambda \boldsymbol{h}$.

Illustrative example 2



$$A = \begin{bmatrix} & 1 & \cdots & 1 & \\ & & & & 1 \\ & & & & \vdots \\ & & & & 1 \\ & & & & \end{bmatrix}$$

For any $\alpha\beta < 1$ there is a unique nonnegative pair of solutions $\boldsymbol{h}^*, \boldsymbol{a}^*$ to

$$A^\top \boldsymbol{h}^\alpha = \lambda \boldsymbol{a}, \qquad A\boldsymbol{a}^\beta = \lambda \boldsymbol{h}$$

		$\boldsymbol{h}^*$	$\boldsymbol{a}^*$	$\boldsymbol{h}^*$	$\boldsymbol{a}^*$	$\boldsymbol{h}^*$	$\boldsymbol{a}^*$	$\boldsymbol{h}^*$	$\boldsymbol{a}^*$
	(α, β)	0.5	0.5	0.9	0.9	0.5	0.9	0.9	0.5
nodes	1	0.39	0.0	0.34	0.0	0.47	0.0	0.24	0.0
	2-5	0.15	0.15	0.16	0.16	0.13	0.19	0.19	0.13
	6	0.0	0.39	0.0	0.34	0.0	0.24	0.0	0.47

Proof: the mapping $F(\boldsymbol{x}, \boldsymbol{y}) = (A\boldsymbol{y}^\beta, A^\top \boldsymbol{x}^\alpha)$ is multihom of degree $\Theta = \begin{bmatrix} & \alpha \\ \beta & \end{bmatrix}$

Higher-order network setting

Higher-order network models

In many applications we are confronted with “higher-order interaction data”.

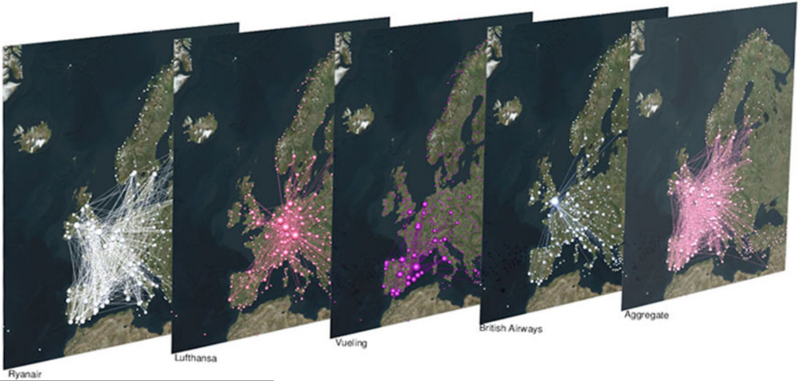
Relational data is full of interactions that happen in groups. For example, friendship relations very often happen in groups that are strictly larger than two individuals. Moreover, interactions naturally occur on multiple layers, for example work relations, sport relations, friendship relations, etc.

We consider the following two settings:

- **Multilayer**
- **Hypergraph**

Multi-layer network centrality

A multi-layer network $\mathbb{G}$ is a sequence $\mathbb{G} = (G_1, \ldots, G_\ell)$ where each $G_k = (V, E_k)$ is a network on the same set of nodes V. Each G_k is called a **layer** of $\mathbb{G}$.



[1]

[1] M. De Domenico et al., J. of Complex Networks, 2015

"Linearization" strategies

Aggregate layer connections
[Tsuda et al., 2005], [Zhou et al., 2007], [Solà et al., 2013]
For $\omega > 0$ define $A_{agg} = \sum_{k=1}^{\ell} \omega_k \, adj(G_k)$.
Centrality: x = Perron eigenvector of A_{agg}

Aggregate layer importances [Battiston et al., 2014]
Let x_k = Bonacich eigenvector of layer k. Centrality: $x = \sum_{k=1}^{\ell} \omega_k \, x_k$

Build a supra-adjacency matrix
[De Domenico et al., 2015], [Taylor et al, 2017] [Taylor et al., 2020]

$$\begin{bmatrix} adj(G_1) & I & \cdots & I \\ I & adj(G_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & I \\ I & \cdots & I & adj(G_\ell) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_\ell \end{bmatrix} = \rho \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_\ell \end{bmatrix}$$

Centrality: $x = \sum_{k=1}^{\ell} \omega_k \, x_k$

Multihomogeneous setting

$T = (T_{ijk})$ adjacency tensor of $\mathbb{G}$: $\boxed{T_{ijk} = 1 \text{ iff } i \to j \text{ on layer } k}$

$$\begin{cases} \lambda x_i = \sum_{jk} T_{ijk} x_j z_k \\ \mu z_k = \sum_{ij} T_{ijk} x_i x_j \end{cases} \qquad \lambda, \mu > 0$$

However:

- does this system of polynomial equations have a positive solution ?
- if yes, is that solution uniquely defined ?

Multihomogeneous setting

$T = (T_{ijk})$ adjacency tensor of $\mathbb{G}$: $\boxed{T_{ijk} = 1 \text{ iff } i \to j \text{ on layer } k}$

$$\begin{cases} \lambda x_i = \sum_{jk} T_{ijk} x_j z_k \\ \mu z_k = \sum_{ij} T_{ijk} x_i x_j \end{cases} \qquad \lambda, \mu > 0$$

However:

- does this system of polynomial equations have a positive solution ?
- if yes, is that solution uniquely defined ?

Example: $n = \ell = 2$

$$\begin{cases} T_{1,1,1} = 6 & T_{1,2,1} = 199/7 & T_{2,1,1} = 16/7 & T_{2,2,1} = 11 \\ T_{1,1,2} = 61/7 & T_{1,2,2} = 6 & T_{2,1,2} = 29 & T_{2,2,2} = 16/7 \end{cases}$$

$\boldsymbol{x} = \frac{1}{3}(2,1)$, $\boldsymbol{z} = \frac{1}{3}(1,2)$ and $\boldsymbol{x} = \frac{1}{4}(1,3)$, $\boldsymbol{z} = \frac{1}{4}(3,1)$ are solutions

Multihomogeneous centrality

$$\begin{cases} \lambda\, f(x_i) = \sum_{jk} T_{ijk} x_j z_k \\ \mu\, g(z_k) = \sum_{ij} T_{ijk} x_i x_j \end{cases} \qquad \lambda, \mu > 0$$

$f =$ hom of deg α
$g =$ hom of deg β

Multihomogeneous centrality

$$\begin{cases} \lambda\, f(x_i) = \sum_{jk} T_{ijk} x_j z_k \\ \mu\, g(z_k) = \sum_{ij} T_{ijk} x_i x_j \end{cases} \qquad \lambda, \mu > 0$$

$$\begin{aligned} f &= \text{ hom of deg } \alpha \\ g &= \text{ hom of deg } \beta \end{aligned}$$

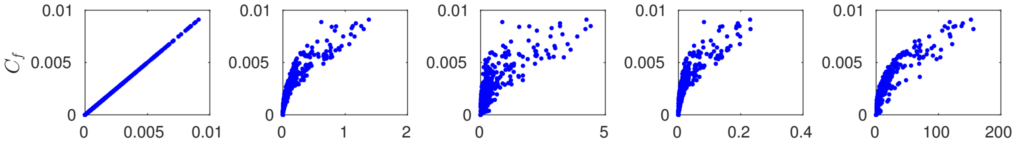
(x, z) is the Perron eigenvector of F multihomogeneous, with

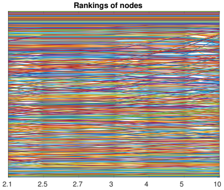
$$\Theta = \begin{bmatrix} 1/\alpha & 1/\alpha \\ 2/\beta & 0 \end{bmatrix}$$

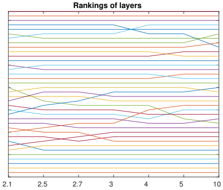
Thus:

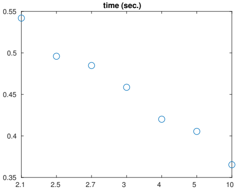
- For $\beta(\alpha - 1) > 2$: unique, positive and computable solution
- We define an importance for both nodes and layers
- We do not need any irreducibility assumption on $\mathbb{G}$

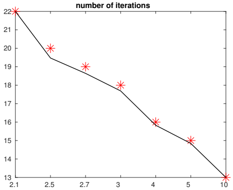
In practice (European Air Transports)



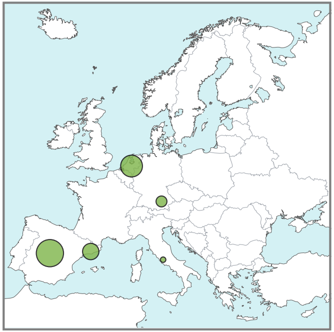








Centrality drawing

Geographical location of the top five European airports according to:

- [F. T. et al., SIAM J. Appl. Math.] (left)
- [De Domenico et al., Nature Comm.] (right)

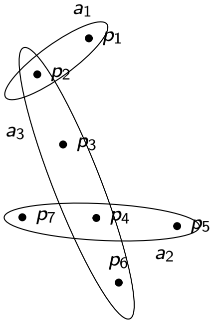
Hypergraphs

Hypergraph:

- $H = (V, \mathcal{E})$ where $e \in \mathcal{E}$ can contain an arbitrary number of nodes (in the graph case each e contains exactly two nodes)
- $w(e) > 0$ denotes the weight on each edge

Example

Journal papers $p_1, p_2, p_3, \ldots$ and authors $a_1, a_2, a_3, \ldots$



Coauthorship hypergraph

Graph projection approach

Transform H into a graph via some form of flattening (or projection)

Most famous example: clique-expansion graph

$$A_{ij} = \sum_{e:\, i,j \in e} w(e)$$

with $w(e)$ the weights of the original hypergraph.

Other approaches: clique averaging [Agarwal et al, 2005]; connectivity graph expansion [Banerjee, 2021], [Ferraz de Arruda et al, 2021]; star expansion [Zien et al., 1999].

Tensor eigenvector approach

Represent the hypergraph via a tensor

$$T_{i_1,\ldots,i_k} = \begin{cases} w(e) & e = (i_1, \ldots, i_k) \in \mathcal{E} \\ 0 & otherwise \end{cases}$$

This is a particularly natural approach in the case of **uniform hypergraphs**.

Centrality x_i of node i proportional to the product of the centralities of nodes on the hyperedge $e = (i, i_2, \ldots, i_k) \in \mathcal{E}$ [Benson, 2019]

$$\sum_{i_2,\ldots,i_k} T_{i,i_2,\ldots,i_m} x_{i_2} x_{i_3} \cdots x_{i_m} = \lambda \left|x_i\right|^p \mathrm{sign}(x_i)$$

Beyond matrices and tensors

$$\sum_{i_2,\ldots,i_k} T_{i,i_2,\ldots,i_m} x_{i_2} x_{i_3} \cdots x_{i_m} = \lambda \left|x_i\right|^p \mathrm{sign}(x_i)$$

- Limited to product of importances on each hyperedge e
- Use of T is limited to uniform hypergraphs

Beyond matrices and tensors

$$\sum_{e \in \mathcal{E}: i \in e} w(e) x_{i_2} x_{i_3} \cdots x_{i_m} = \lambda \, |x_i|^p \mathrm{sign}(x_i)$$

- Limited to product of importances on each hyperedge e
- Use of T is limited to uniform hypergraphs

Beyond matrices and tensors

$$\sum_{e \in \mathcal{E}: i \in e} w(e) x_{i_2} x_{i_3} \cdots x_{i_m} = \lambda \, |x_i|^p \operatorname{sign}(x_i)$$

- Limited to product of importances on each hyperedge e
- Use of T is limited to uniform hypergraphs

For a general hypergraph H:

$$\sum_{e \in \mathcal{E}: i \in e} w(e) f(x_{j_1}, \ldots, x_{j_{k_e}}) = \lambda \, g(x_i)$$

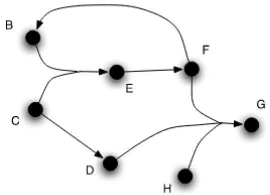
where $j_1, \ldots, j_{k_e}$ are the nodes different from i in the hyperedge $e \in \mathcal{E}$.

Thursday May 20, @11:55, Minisymposium on *Latest Advances in Spectral Linear Algebra in Network Science*

Machine learning explainability metrics

Exaplainability: provide a value that quantifies the effect of the i-th feature for the model prediction on each datapoint.

ML classifier f. Explainability (directed) hypergraph H with edge weight:



$$w(S, j) = \text{marginal impact that } j \text{ has on } f \text{ given the features in } S \cup \{j\}$$

Soon on arxiv: [Gautier, Tudisco et al., *Explainability hypergraphs*]

Thank you! Any question?

Antoine Gautier & Francesco Tudisco

https://ftudisco.github.io/siam-nonlinear-pf-tutorial/

(feedback very welcome)

See you soon at Part 2 @ 11:15 CT