CENTRALITY

degree

eigenvector

katz

PageRank

Hubs & Authorities

betweenness

closeness

SYSM 6302

CLASS 7

Closeness Centrality

Let d_{ij} = length of the shortest path from i to j

$\Rightarrow$ Recall diameter $= \max_{i,j} d_{ij}$

Mean distance of other nodes from i: $\ell_i = \frac{1}{n} \sum_{j=1}^{n} d_{ij}$

small is more important

for unweighted networks the range of ℓ_i is not very broad

Closeness centrality: $C_i = \frac{1}{\ell_i} = \frac{n}{\sum_j d_{ij}}$

A less standard method: $\tilde{C}_i = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d_{ij}}$ (called a harmonic mean)

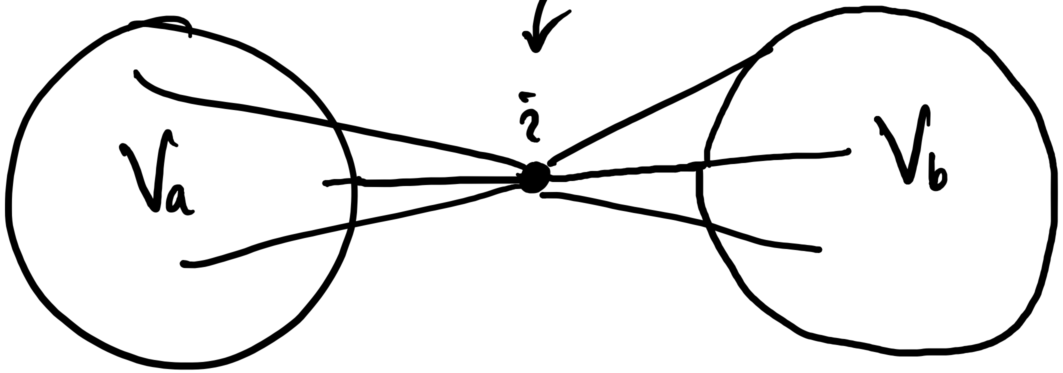
This allows for elegant inclusion of nodes in other components

Contribution to centrality is inversely proportional to distance, so local connectivity dominates.

Betweenness Centrality

$$x_i = \sum_s \sum_t \frac{\text{\# of shortest paths from } s \text{ to } t \text{ that include } i}{\text{\# of shortest paths from } s \text{ to } t}$$

on *all* the shortest paths from nodes in V_a to nodes in V_b



Degree Centrality

Nodes with higher number of connections are ranked as more important

⇒ "popularity"

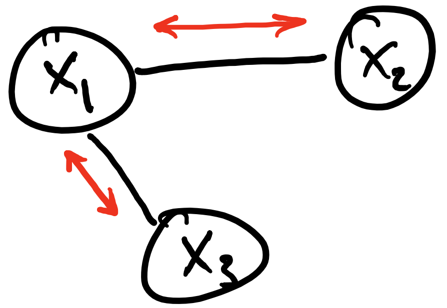
This is a local property of the node

Eigenvector Centrality

⇒ Degree assigns importance equally in terms of the # of neighbors.

⇒ Eigenvector Centrality assigns importance proportional to the importance of the node's neighbors

We can view this measure at the "steady-state" of each node "passing" its importance to all of its neighbors



$x(t+1) = Ax(t) \longrightarrow$ steady state when $x(t+1) = x(t) = v$ $\longrightarrow v = Av$

Do this more carefully:

$$x(t+1) = Ax(t) \implies x(t) = A^t x(0)$$

If $A \in \mathbb{R}^{n \times n}$ is symmetric $\implies$ its eigenvectors span $\mathbb{R}^n$ & are linearly independent

$\implies$ Any $u \in \mathbb{R}^n$ can be written as a linear combination of these eigenvectors:

$$u = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

constants

Since $x(0) \in \mathbb{R}^n$: $\quad x(0) = \sum_{i=1}^{n} c_i v_i \implies x(t) = A^t \sum_{i=1}^{n} c_i v_i$

$$x(t) = \sum_{i=1}^{n} c_i A^t v_i = \sum_{i=1}^{n} c_i \lambda_i^t v_i = c_1 \lambda_1^t v_1 + \lambda_1^t \sum_{i=2}^{n} c_i \underbrace{\left(\frac{\lambda_i}{\lambda_1}\right)^t}_{<1} v_i \xrightarrow{t \to \infty} \underbrace{c_1 \lambda_1^t}_{\text{a number}} \underbrace{v_1}_{\text{Vector}}$$

largest eigenvalue

$$x(t) \to c_1 \lambda_1^t v_1 \quad \text{as } t \to \infty$$

But since absolute values of centrality are not important,

$$x \sim v_1$$ ← eigenvector corresponding to the largest eigenvalue of A

$$Av_1 = \lambda_1 v_1 \implies x = \frac{1}{\lambda_1} A x$$

j points to i

$$\Rightarrow x_i = \frac{1}{\lambda_1} \sum_{j=1}^{n} A_{ij} x_j$$

centrality of node i

centrality of node j

x_i can be large because it is connected to many nodes OR because it is connected to nodes with large x_j (large centralities)

Perron-Frobenius Theorm

For a positive matrix A (ie. $A_{ij} > 0 \quad \forall i,j = 1,\dots,n$):

① There exists a real, positive eigenvalue λ_1 such that the magnitude of all other eigenvalues is ~~less~~ less than λ_1^{+}: $|\lambda_i| < |\lambda_1| \quad \forall i = 2,3,\dots,n$.

② The eigenspace of λ_1 is dimension 1 (it has only 1 corresponding eigenvector)

③ The corresponding eigenvector, v_1, of λ_1 has entries that are the same sign $(+/-)^{*}$. All other eigenvectors have at least 1 element that is positive and 1 that is negative.

For non-negative A ($A_{ij} \geq 0$)

*or zero $\quad$ $^{+}$or the same $|\lambda_i| \leq |\lambda_1|$

Eigenvector Centrality for Directed Graphs

① A is not symmetric $\rightarrow$ eigenvectors do not always span $\mathbb{R}^n$

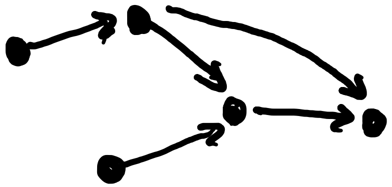
$\rightarrow$ left $\neq$ right eigenvectors

left: $u^T A = \lambda u^T$ right: $Av = \lambda v$

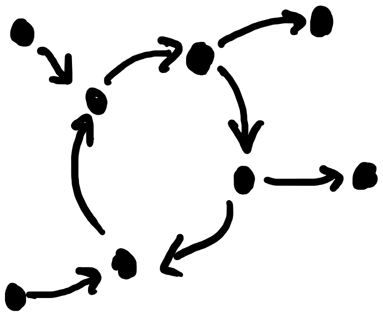
$$\begin{bmatrix} x_i \\ \\ \\ \end{bmatrix} = \begin{array}{c} \begin{array}{cc} & j \end{array} \\ \begin{array}{c} i \end{array}\begin{bmatrix} & A_{ij} \\ & \\ & \end{bmatrix} \end{array} \begin{bmatrix} \\ \\ x_j \end{bmatrix}$$

Centrality is received

② Acyclic Graphs:



what is: $A^{10}x = \,?$



The graph requires strongly connected components to have nonzero centralities

Computation of Eigenvector Centrality

As n gets large computing eigenvectors can become costly.

A cheaper alternative is to use our iterative approach:

$$x(t+1) = A\,x(t) \implies x(t) = A^t x(0)$$

for some large enough t, the <u>direction</u> of $x(t)$ will stabilize,

but the amplitude will grow.

↳ Solution! Normalize at each iteration step

Katz Centrality

"Everyone should be at least a little important"

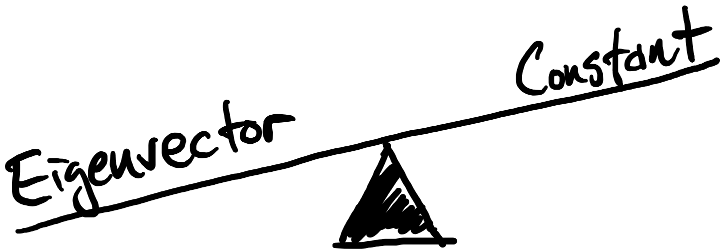
$$x(t+1) = A\,x(t) \longrightarrow x(t+1) = \alpha A\,x(t) + \beta \vec{1}$$

Eigenvector

Katz

$\Downarrow$ $x(t+1) = x(t)$ "steady state"

Why α?



$$x = \underbrace{(I - \alpha A)}^{-1} \cdot \beta \vec{1}$$

blows up when $(I - \alpha A)$ is singular!

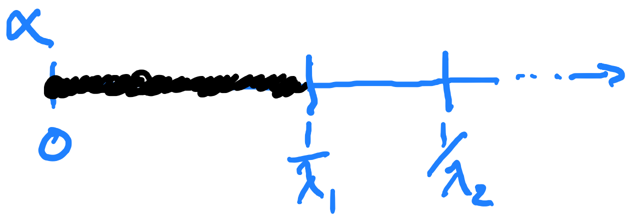
$$\det(I - \alpha A) = 0 \Rightarrow \det(A - \alpha^{-1} I) = 0$$

$\frac{1}{\alpha} = \lambda$

if $\alpha \to 0 \Rightarrow x = \beta \cdot \vec{1}$

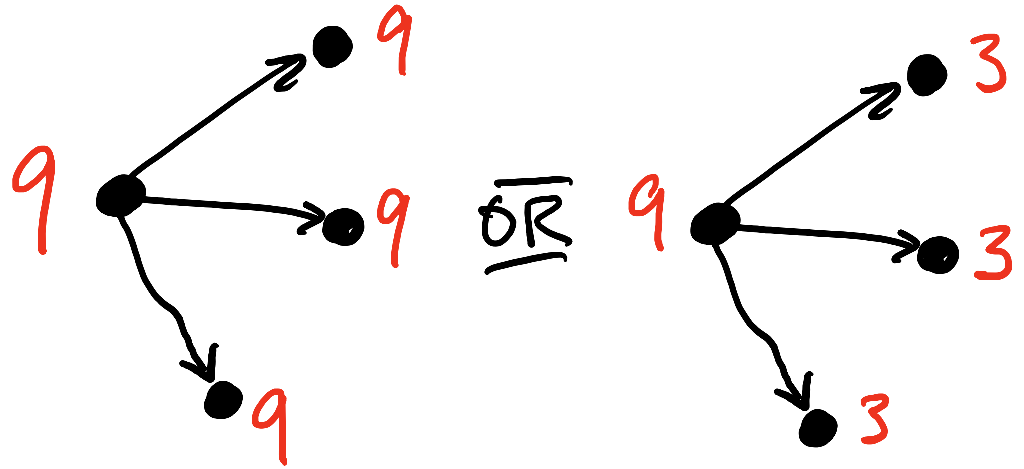
all nodes have equal centrality

$$0 \le \alpha < \frac{1}{\lambda_1}$$



PageRank Centrality

"Exploring conservation of centrality"



$$x_i = \alpha \sum_{j=1}^{n} A_{ij} x_j + \beta$$ Katz

$$x_i = \alpha \sum_{j=1}^{n} A_{ij} \frac{x_j}{k_j^{OUT}} + \beta$$ PageRank

out degree of node j

$$x = \alpha A D^{-1} x + \beta \cdot \vec{1}$$

$D_{ii} = \max(k_i^{OUT}, 1)$

$$= (I - \alpha A D^{-1})^{-1} \cdot \beta \vec{1}$$

for undirected: $0 \leq \alpha < 1$

in lab!

Google uses $\alpha = 0.85$

Hubs & Authorities

$$y_i \qquad x_i$$

"I don't necessarily have to know SOMETHING, I could know SOMEONE."

Two Centralities:

$$x_i = \alpha \sum_j A_{ij} y_j \qquad \Rightarrow \qquad x = \alpha A y \longrightarrow AA^T x = \frac{1}{\alpha\beta} x$$

Cocitation!

$$y_i = \beta \sum_j A_{ij} x_j \qquad \qquad y = \beta A x \longrightarrow A^T A y = \frac{1}{\alpha\beta} y$$

bibliographic coupling!

Authorities = eigenvector centrality of the cocitation graph

Hubs = eigenvector centrality of the bibliographic coupling graph