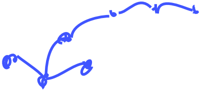
MITx:
Statistics, Computation & Applications

Criminal Networks Module

Lecture 2: Centrality Measures

$$C_i = \left(\frac{1}{n-1} \sum_{k \neq i} d_{ik} \right)^{-1}$$

$$B_i = \sum_{s,t} \frac{n^i_{st}}{g_{st}}$$



$$x = Ax$$

MITx: Statistics, Computation & Applications

Criminal Networks Module

Lecture 2: Centrality Measures

Find important nodes

- **Centrality measure**: A measure that captures importance of a node's position in the network
- There are many different centrality measures
  - degree centrality (indegree / outdegree)
  - "propagated" degree centrality (score that is proportional to the sum of the score of all neighbors)
  - closeness centrality
  - betweenness centrality

Which centrality measure to use

Choice of centrality measure depends on application!

In a friendship network:

- high degree centrality: most popular person
- high eigenvector centrality: most popular person that is friends with popular people
- high closeness centrality: person that could best inform the group
- high betweenness centrality: person whose removal could best break the network apart

Small network in which distinct nodes maximize degree, eigenvector, closeness and betweenness centralities?

Degree centrality

- For undirected graphs the degree k_i of node i is the number of edges connected to i, i.e. $k_i = \sum_j A_{ij}$
- For directed graphs the indegree of node i is $k_i^{\mathrm{in}} = \sum_j A_{ij}$ and the outdegree is $k_i^{\mathrm{out}} = \sum_j A_{ji}$
- Simple, but intuitive: individuals with more connections have more influence and more access to information.

Degree centrality

- For undirected graphs the degree k_i of node i is the number of edges connected to i, i.e. $k_i = \sum_j A_{ij}$
- For directed graphs the indegree of node i is $k_i^{\text{in}} = \sum_j A_{ij}$ and the outdegree is $k_i^{\text{out}} = \sum_j A_{ji}$
- Simple, but intuitive: individuals with more connections have more influence and more access to information.
- Does not capture "cascade of effects": importance better captured by having connections to important nodes

Closeness and betweenness centrality

- **Closeness centrality:** Tracks how close a node is to any other node:

Closeness and betweenness centrality

- **Closeness centrality:** Tracks how close a node is to any other node:

$$C_i = \left(\frac{1}{n-1} \sum_{j \neq i} d_{ij} \right)^{-1},$$

where d_{ij} is the distance between nodes i and j

Closeness and betweenness centrality

- **Closeness centrality:** Tracks how close a node is to any other node:

$$C_i = \left(\frac{1}{n-1} \sum_{j \neq i} d_{ij} \right)^{-1},$$

where d_{ij} is the distance between nodes i and j

  - In disconnected networks: average over nodes in same component as i or use **harmonic centrality**: $H_i = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d_{ij}}$

Closeness and betweenness centrality

- **Closeness centrality:** Tracks how close a node is to any other node:

$$C_i = \left(\frac{1}{n-1} \sum_{j \neq i} d_{ij} \right)^{-1},$$

where d_{ij} is the distance between nodes i and j

  - In disconnected networks: average over nodes in same component as i or use **harmonic centrality**: $H_i = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d_{ij}}$

- **Betweenness centrality:** Measures the extent to which a node lies on paths between other nodes:

Closeness and betweenness centrality

- **Closeness centrality:** Tracks how close a node is to any other node:

$$C_i = \left(\frac{1}{n-1} \sum_{j \neq i} d_{ij} \right)^{-1},$$

where d_{ij} is the distance between nodes i and j

  - In disconnected networks: average over nodes in same component as i or use **harmonic centrality**: $H_i = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d_{ij}}$

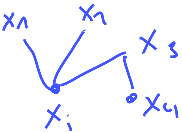
- **Betweenness centrality:** Measures the extent to which a node lies on paths between other nodes:

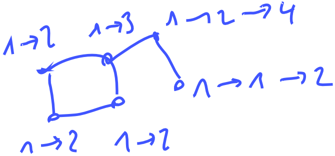
$$B_i = \frac{1}{n^2} \sum_{s,t} \frac{n_{st}^i}{g_{st}},$$

where n_{st}^i is number of shortest paths between s and t that pass through i, and g_{st} is total number of shortest paths between s and t

Eigenvector centrality

- gives each node a score that is proportional to the sum of the scores of all its neighbors
- need to know scores of all neighbors, which we don't know





Eigenvector centrality

- gives each node a score that is proportional to the sum of the scores of all its neighbors
- need to know scores of all neighbors, which we don't know
- start with equal centrality: $x_i^{(0)} = 1$ for all nodes $i = 1, \ldots, n$
- update each centrality by the centrality of the neighbors:

$$x_i^{(1)} = \sum_{j=1}^{n} A_{ij} x_j^{(0)}$$

$$\begin{pmatrix} 1\ 0 & \cdot & 1\ 0 \end{pmatrix} \qquad x_1,\ x_2,\ \ldots,\ x_n$$

- iterate this process: $\quad x^{(k)} = A^k x^{(0)}$

$$x^{(1)} = A x^{(0)}$$

$$= \sum_{i=1}^{n} \lambda_i^k v_i v_i^T x^{(0)} \qquad \lambda_1 \geq \lambda_2 \ldots$$

$$= \lambda_1^k v_1 \underbrace{v_1^T x^{(0)}}_{\alpha_1} + \lambda_2^k v_2 v_2^T x^{(0)} + \ldots$$

$$= \lambda_1^k \left(\alpha_1 v_1 + \frac{\lambda_2^k}{\lambda_1^k} \alpha_2 v_2 + \ldots \right)$$

$$\xrightarrow{k \to \infty} \lambda_1^k \alpha_1 v_1 \qquad \text{if } \lambda_1 > \lambda_2$$

Eigenvector centrality

- gives each node a score that is proportional to the sum of the scores of all its neighbors
- need to know scores of all neighbors, which we don't know
- start with equal centrality: $x_i^{(0)} = 1$ for all nodes $i = 1, \ldots, n$
- update each centrality by the centrality of the neighbors:

$$x_i^{(1)} = \sum_{j=1}^{n} A_{ij} x_j^{(0)}$$

$$x^{(k)} = \frac{1}{\lambda_{\max}} A x^{(k-1)}$$

- iterate this process: $x^{(k)} = A^k x^{(0)}$
- if there exists $m > 0$ such that $A^m > 0$, then one can show that

$$x^{(k)} \xrightarrow{k \to \infty} \alpha \cancel{\lambda_{\max}^k} v,$$

where $\lambda_{\max}$ is the largest eigenvalue and $v \geq 0$ the corresponding eigenvector; α depends on choice of $x^{(0)}$ (Perron-Frobenius theorem)

Eigenvector centrality

Interpretation: $v_i = \frac{1}{\lambda_{\max}} \sum_{j=1}^{n} A_{ij} v_j$

- node is important if it has important neighbors
- node is important if it has many neighbors
- eigenvector corresponding to largest eigenvalue of A provides a ranking of all nodes

Eigenvector centrality

Interpretation: $v_i = \frac{1}{\lambda_{\max}} \sum_{j=1}^{n} A_{ij} v_j$

- node is important if it has important neighbors
- node is important if it has many neighbors
- eigenvector corresponding to largest eigenvalue of A provides a ranking of all nodes

What happens when G is directed?

Eigenvector centrality

Interpretation: $v_i = \frac{1}{\lambda_{\max}} \sum_{j=1}^{n} A_{ij} v_j$

- node is important if it has important neighbors
- node is important if it has many neighbors
- eigenvector corresponding to largest eigenvalue of A provides a ranking of all nodes

What happens when G is directed?

- right eigenvector: $v_i = \frac{1}{\lambda_{\max}} \sum_{j=1}^{n} A_{ij} v_j$
  - importance comes from nodes i points to
  - Example: determining malfunctioning genes
- left eigenvector: $w_i = \frac{1}{\lambda_{\max}} \sum_{j=1}^{n} w_j A_{ji}$
  - importance comes from nodes pointing to i
  - Example: ranking websites

Katz centrality

- Can the eigenvector centrality be applied to any directed network?

Katz centrality

- Can the eigenvector centrality be applied to any directed network?
- Any directed graph with a source / sink node gives zero eigenvector centrality (since the zeros are propagated through)

Katz centrality

- Can the eigenvector centrality be applied to any directed network?
- Any directed graph with a source / sink node gives zero eigenvector centrality (since the zeros are propagated through)
- **Remedy:** Give every node some fixed (but small) centrality for free:

$$x_i^{(k+1)} = \alpha \sum_{j=1}^{n} A_{ij} x_j^{(k)} + \beta_i$$

or equivalently,

$$x^{(k+1)} = \alpha A x^{(k)} + \beta$$

Katz centrality

- Can the eigenvector centrality be applied to any directed network?
- Any directed graph with a source / sink node gives zero eigenvector centrality (since the zeros are propagated through)
- **Remedy:** Give every node some fixed (but small) centrality for free:

$$x_i^{(k+1)} = \alpha \sum_{j=1}^{n} A_{ij} x_j^{(k)} + \beta_i$$

or equivalently,

$$x^{(k+1)} = \alpha A x^{(k)} + \beta$$

- If α is chosen in the interval $(0, 1/\lambda_{\max}(A))$, then one can show that

$$x^{(k)} \overset{k\to\infty}{\longrightarrow} v,$$

where $v = (I - \alpha A)^{-1}\beta \geq 0$ (for example: for DAGs it holds that $\lambda_{\max} = 0$, hence no constraints on α; take e.g. $\alpha = 1$)

Page rank

- Drawback of Katz centrality: A node of high centrality pointing to many nodes gives them all high centrality.

Page rank

- Drawback of Katz centrality: A node of high centrality pointing to many nodes gives them all high centrality.
- **Remedy:** Scale by the degree of a node:

$$x_j^{(k+1)} = \alpha \sum_{i=1}^{n} A_{ij} \frac{x_i^{(k)}}{k_i^{\text{out}}} + \beta_j,$$

or equivalently,

$$x^{(k+1)} = \alpha D^{-1} A x^{(k)} + \beta, \quad \text{where} \quad D = \text{diag}(k_1^{\text{out}}, \ldots, k_n^{\text{out}})$$

Page rank

- Drawback of Katz centrality: A node of high centrality pointing to many nodes gives them all high centrality.

- **Remedy:** Scale by the degree of a node:

$$x_j^{(k+1)} = \alpha \sum_{i=1}^{n} A_{ij} \frac{x_i^{(k)}}{k_i^{\text{out}}} + \beta_j,$$

  or equivalently,

$$x^{(k+1)} = \alpha D^{-1} A x^{(k)} + \beta, \quad \text{where} \quad D = \text{diag}(k_1^{\text{out}}, \ldots, k_n^{\text{out}})$$

- If α is chosen in the interval $(0, 1/\lambda_{\max}(D^{-1}A))$, then one can show that

$$x^{(k)} \xrightarrow{k\to\infty} v,$$

  where $v = (I - \alpha D^{-1} A)^{-1} \beta \geq 0$

Hubs and authorities

- Example: Paper can be important because

Hubs and authorities

- Example: Paper can be important because
  - it contains important information itself (authority)
  - it points to important papers (hub)

Hubs and authorities

- Example: Paper can be important because
  - it contains important information itself (authority)
  - it points to important papers (hub)
- **Approach:** Define 2 centrality measures x (hub, is high if it points to many authorities) and y (authority, is high if many hubs point to it)

$$x_i^{(k+1)} = \alpha \sum_{j=1}^{n} A_{ij}\, y_j^{(k)}, \qquad \text{i.e.,} \qquad x^{(k+1)} = \alpha A y^{(k)}$$

$$y_i^{(k)} = \beta \sum_{j=1}^{n} A_{ji}\, x_j^{(k)}, \qquad \text{i.e.,} \qquad y^{(k)} = \beta A^T x^{(k)}$$

$$AA^T v = \lambda v$$

$$(A^T A)\underbrace{(A^T v)}_{w} = \lambda \underbrace{(A^T v)}_{w}$$

$$x^{(k+2)} = \alpha\beta\, A A^T x^{(k)}$$

$$y^{(k+2)} = \alpha\beta\, A^T A\, y^{(k)}$$

Hubs and authorities

- Example: Paper can be important because
  - it contains important information itself (authority)
  - it points to important papers (hub)

- **Approach:** Define 2 centrality measures x (hub, is high if it points to many authorities) and y (authority, is high if many hubs point to it)

$$x_i^{(k+1)} = \alpha \sum_{j=1}^{n} A_{ij}\, y_j^{(k)}, \qquad \text{i.e.,} \qquad x^{(k+1)} = \alpha A y^{(k)}$$

$$y_i^{(k)} = \beta \sum_{j=1}^{n} A_{ji}\, x_j^{(k)}, \qquad \text{i.e.,} \qquad y^{(k)} = \beta A^T x^{(k)}$$

- Choosing $\alpha\beta = 1/\lambda_{\max}(AA^T)$ $= \frac{1}{\lambda_{\max}}(A^TA)$, then

$$x^{(k)} \xrightarrow{k\to\infty} v \quad \text{and} \quad y^{(k)} \xrightarrow{k\to\infty} w$$

such that $AA^T v = \lambda v$ and $A^T A w = \lambda w$ (in fact $w = A^T v$)

References

- Chapters 6 - 10 (but mostly Chapter 7) in

  M. E. J. Newman. *Networks: An Introduction*. 2010.