Annotated slides

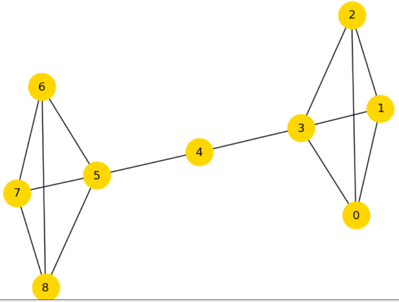
CS4423: Networks

Week 7, Part 1: Closeness and Betweenness Centrality

Dr Niall Madden

School of Maths, University of Galway

26 + 27 February 2025)



Reminders

- **Assignment 1** Due 5pm Friday, 27th February.
- **Class Test** 14:00, Thursday 6th March (Week 8)

Outline

Today's notes are split between these slides, and a Jupyter Notebook.

1. Centrality Measures (again)
2. Eigenvector Centrality (again)
3. Closeness Centrality
   - Normalised
   - Distance Matrix
4. Betweenness Centrality
   - Normalised
   - Examples

Slides are at:
https://www.niallmadden.ie/2425-CS4423

Centrality Measures (again)

Last week we learned about some centrality measures:

Measures of centrality include:

- The **degree centrality**, c_i^D of Node i in $G = (X, E)$ is the degree of i (i.e., the number of neighbours it has). So $c_i^D = \deg(i)$.
- The **normalised degree centrality**, C_i^D of Node i is $C_i^D = \deg(i)/(n-1)$ where n is the order of the network.
- **Eigenvector Centrality**, which we'll recap now.

Then we'll look at:

- **Closeness Centrality**, and
- **Betweenness Centrality**.

Note: $0 \le \deg(i) \le n-1$ for all i.

Eigenvector Centrality (again)

Eigenvector Centrality

1. Let A be the adjacency matrix of a network. G.
2. We know, thanks to Perron-Frobenius, that A has a positive eigenvalue, λ, which is equal to the spectral radius of A.
3. There is a positive eigenvector, v associated with λ.
4. Choose v so that $v^T v = v_1^2 + v_2^2 + \cdots + v_n^2 = 1$.
5. v_i is the **eigenvector centrality** of Node i.

Closeness Centrality

A node x in a network can be regarded as being central, if it is **close** to (many) other nodes, as it can then quickly interact with them.

Recalling the $d(i,j)$ is the distance between Nodes i and j (i.e., the length of the shortest path between them). The we can use $1/d(i,j)$ as a measure of "closeness".

Definition (Closeness Centrality)

In a simple, *connected* graph $G = (X, E)$ of order n, the **closeness centrality**, c_i^C, of Node i is defined as

Closeness

$$c_i^C = \frac{1}{\sum_{j \in X} d(i,j)} = \frac{1}{s(i)},$$

Centrality

where $s(i)$ is the **distance sum** for node i.

As is usually the case, there is a **normalised** version of this measure.

Normalised closeness centrality

The **normalised closeness centrality** of Node i, defined as

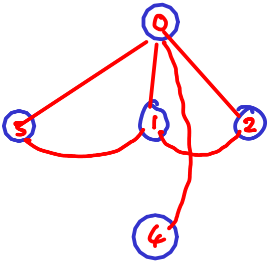
$$C_i^C = (n-1)c_i^C = \frac{n-1}{\sum_{j \in X} d(i,j)} = \frac{n-1}{s(i)}.$$

Note: $0 \leq C_i^C \leq 1$. (*Why?*)

If node i is a distance 1 from all of the other (n-1) nodes, then s(i) = n-1 so the max possible value of c_i^C is n-1.

Example

Compute the normalised closeness centrality of all nodes in the graph on nodes $\{0, 1, 2, 3, 4\}$, with edges $0 - 1$, $0 - 2$, $0 - 3$, $0 - 4$, $1 - 2$, $1 - 3$.



Compute distance between all pairs

	0	1	2	3	4	$S(i)$	C_i^c
0	0	1	1	1	1	4	$4/4 = 1$
1	1	0	1	1	2	5	$4/5$
2	1	1	0	2	2	6	$4/6$
3	1	1	2	0	2	6	$4/6$
4	1	2	2	2	0	7.	$4/7$

In that example we effectively computed the **distance matrix** of the graph.

Distance Matrix

The **distance matrix** of a graph, G, of order n is the $n \times n$ matrix, $\mathcal{D} = (d_{ij})$ such that

$$d_{ij} = d(i, j).$$

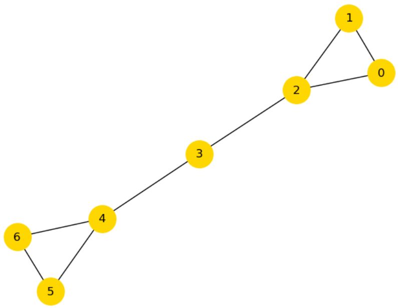
Note : $D = D^T$ for an undirected Graph.

We'll return to how to compute $\mathcal{D}$ tomorrow, but for now we note:

- $s(i)$ is the sum of row i of $\mathcal{D}$;
- If $\boldsymbol{s}$ is the vector of of distance sums, then $\boldsymbol{s} = \mathcal{D}\boldsymbol{e}$, where $\boldsymbol{e} = (1, 1, \ldots, 1)^T$.

Betweenness Centrality

Consider the following graph (as the $3-1$ Barbell Graph):



We can, I hope, convince ourselves, that, in a sense:

- Node 3 is the most central, in the sense that belongs to the most shortest paths.
- Node 0 (for example), is very much not central in that sense.

Betweenness Centrality

Definition (Betweenness Centrality)

In a simple, connected graph G, the **betweenness centrality** c_i^B of node i is defined as

$$c_i^B = \sum_j \sum_k \frac{n_i(j,k)}{n(j,k)}, \qquad j \neq k \neq i$$

where $n(j,k)$ denotes the *number* of shortest paths from node j to node k, and where $n_i(j,k)$ denotes the number of those shortest paths *passing through* node i.

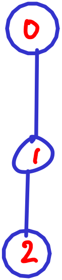
Definition (Normalised Betweenness Centrality)

In a simple, connected graph G, the **normalised betweenness centrality** C_i^B of node i is defined as

$$C_i^B = \frac{c_i^B}{(n-1)(n-2)}$$

why $(n-1)(n-2)$?

Example 1 : P_3. Find C_0^B, C_1^B. $\left(C_2^B = C_0^B\right)$



Node 1 :

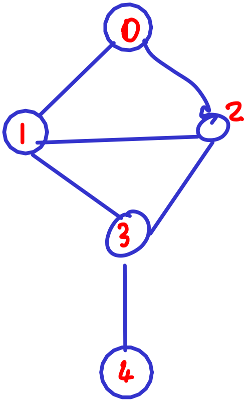
Pairs j,k	$n_1(j,k)$	$n(j,k)$	$\frac{n_0(j,k)}{n(j,k)}$
0, 2	1	1	1
2, 0	1	1	1

$\sum_j \sum_i = 2$.

$$C_1^B = \frac{2}{(2)(1)} = 1$$

Note $C_0^B = 0$ since $n_0(1,2) = 0$ & $n_0(2,1) = 0$.

Example : Find C_3^B



pairs j	k	$n_3(j,k)$	$n(j,k)$	The sum!
0	1	0	–	
0	2	0	–	
0	4	1	2	
1	0	0	–	
1	2	0	–	
1	4	1	1	
2	0	0	–	
2	1	0	–	
2	4	1	1	
4	0	2	2	
4	1	1	1	
4	2	1	1	