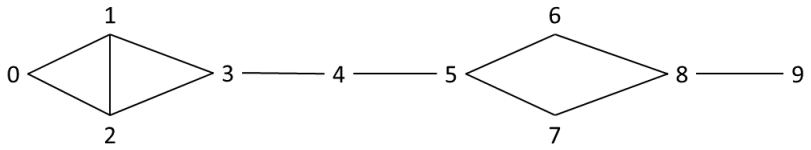
Math 345 - Notes
Centrality in network analysis
December 3, 2018

Centrality

What is a central data point in a network? Do we label a node "central" based on its location in the network, based on the number of neighbors, based on the quality of its neighbors? Let's consider the undirected graph below.



Degree centrality

The easiest way to assess importance or centrality is to define "central" nodes as those that have most neighbors (think of being important if you have many connections). This type of centrality is called **degree centrality**, where for each node i, we define its centrality by

$$c(i) = degree(i),$$

where $degree(i)$ counts the number of edges incident to i.
That means that for our example, $\max\limits_{0\leq i\leq 9} c(i) = 3$, and the most central nodes are $\{1, 2, 3, 5, 8\}$.

Betweenness centrality

Sometimes, having the most connections is not what it is thought of being central. We can define being "central" as frequently being on the shortest path between nodes. In that case, to compute the centrality of a node v, for all pairs i, j we find the proportion of shortest paths from i to j going through v. More precisely,

$$c(v) = \sum_{i\neq v} \sum_{\substack{j\neq i \\ j\neq v}} \frac{\text{number of shortest paths from } i \text{ to } j \text{ through } v}{\text{number of shortest paths from } i \text{ to } j}.$$

We call this type of centrality **betweenness centrality.** Being "central" means being in the middle of things.

In our example

$c(0)$	$= 0$	no shortest paths thru 0
$c(1)$	$= 7 \times \frac{1}{2} = 3.5$	shortest paths from 0, each being one of two possible paths
$c(2)$	$= 7 \times \frac{1}{2} = 3.5$	by symmetry with node 1
$c(3)$	$= 3 \times 6 = 18$	paths from $0, 1, 2$ to the remaining 6 nodes
$c(4)$	$= 4 \times 5 = 20$	paths from $0 \ldots 3$ to nodes $5 \ldots 9$
$c(5)$	$= 5 \times 4 + \frac{1}{2} = 20.5$	paths from $0 \ldots 4$ to nodes $6 \ldots 9$, and path 6 to 7
$c(6)$	$= 6 \times 2 \times \frac{1}{2} = 6$	paths from $0 \ldots 5$ to nodes $8, 9$
$c(7)$	$= 6 \times 2 \times \frac{1}{2} = 6$	symmetry with node 6
$c(8)$	$= 8 \times 1 + \frac{1}{2} = 8.5$	paths from $0 \ldots 7$ to node 9, and path 6 to 7
$c(9)$	$= 0$	no shortest paths thru 0

The most central node is node 5 with $c(5) = 20.5$.

Closeness centrality

A similar measure of being "central" is **closeness centrality**. First, let us define farness of a node as the sum of shortest path lengths from the node to all other nodes. More precisely, for node i,

$$\text{farness}(i) = \sum_{j \neq i} \text{length of shortest path from } i \text{ to } j.$$

Then closeness centrality is defined as

$$c(i) = \frac{1}{\text{farness}(i)}.$$

In our example, we compute the farness first, noting that for the largest centrality we need the smallest farness.

$$\begin{aligned}
\text{farness}(0) &= 1+1+2+3+4+5+5+6+7 = 34 \\
\text{farness}(1) &= 1+1+1+2+3+4+4+5+6 = 27 \\
\text{farness}(2) &= 1+1+1+2+3+4+4+5+6 = 27 \\
\text{farness}(3) &= 2+1+1+1+2+3+3+4+5 = 22 \\
\text{farness}(4) &= 3+2+2+1+1+2+2+3+4 = 20 \\
\text{farness}(5) &= 4+3+3+2+1+1+1+2+3 = 20 \\
\text{farness}(6) &= 5+4+4+3+2+1+2+1+2 = 24 \\
\text{farness}(7) &= 5+4+4+3+2+1+2+1+2 = 24 \\
\text{farness}(8) &= 6+5+5+4+3+2+1+1+1 = 28 \\
\text{farness}(7) &= 7+6+6+5+4+3+2+2+1 = 36
\end{aligned}$$

Thus, in this example, nodes 4 and 5 have the largest centrality, with $c(4) = c(5) = \frac{1}{20}$.

Eigenvector centrality

While the degree centrality is based on the *quantity* of connections, this last type of centrality is based on the *quality* of connections. For example, having a few central connections makes you more central than having many non-central connections, or in other words, it matters who your connections are.

Suppose node i centrality score is given by $c(i)$. Then we would like $c(i)$ to be directly proportional to the sum of centralities of its neighbors. More precisely, if $N(i)$ denotes the neighbors of i,

$$c(i) = \frac{1}{\lambda} \sum_{j \in N(i)} c(j),$$

where λ is the proportionality constant.

We recall from discrete math that a graph can be represented by an adjacency matrix, A, with $A_{ij} = 1$ if there is an edge between i and j, and 0 otherwise. Then we can express $c(i)$ in terms of the matrix A:

$$c(i) = \frac{1}{\lambda} \sum_{j} A_{ij} c(j),$$

and if we collect centrality scores in a vector $\mathbf{c}$, then we have

$$\mathbf{c} = \frac{1}{\lambda} A\mathbf{c} \quad \Leftrightarrow \quad A\mathbf{c} = \lambda \mathbf{c}.$$

Thus, $\mathbf{c}$ is the eigenvector of A corresponding to the eigenvalue λ (hence the name **eigenvector centrality**). The Perron-Frobenius Theorem ensures that the matrix has a unique largest real eigenvalue and that the corresponding eigenvector can be chosen to have strictly positive entries. Thus, we will take λ to be the largest eigenvalue of A, and $\mathbf{c}$ its corresponding eigenvector.

For our example, the adjacency matrix is

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Using a calculator or computational software, we find that the largest eigenvalue is $\lambda = 2.66$, with corresponding eigenvector

$$\mathbf{c}^T = [.38 \ .51 \ .51 \ .47 \ .23 \ .15 \ .08 \ .08 \ .07 \ .02].$$

How about much larger networks? How does one compute eigenvectors? We can use the Power Iteration algorithm to compute the eigenvector corresponding to the largest eigenvalue of a diagonalizable matrix A.

Power Iteration Algorithm

- Start with an initial vector $\mathbf{x}_0$, non-zero in the direction of the eigenvector corresponding to the largest eigenvalue.
- At time $t+1$, for $t \geq 0$, let
$$\mathbf{x}_{t+1} = \frac{A\mathbf{x}_t}{||A\mathbf{x}_t||}$$
- Iterate until $||\mathbf{x}_{t+1} - \mathbf{x}_t|| < \epsilon$, for some small error ϵ.

Remarks:

1. The rescaling in the update step is being done so that the resulting eigenvector is unique.
2. To see why the power iteration algorithm works, consider the following. Given that A is diagonalizable, we can find matrices P and diagonal matrix D so that
$$A = PDP^{-1}$$
From linear algebra, we recall that P has eigenvectors as columns and D is a diagonal matrix with corresponding eigenvalues on the diagonal. Then
$$A^k = (PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})(PDP^{-1}) = PD^kP^{-1} \;\Leftrightarrow\; A^kP = PD^k.$$
Let $\mathbf{x}_0 = P\mathbf{x}$ for some vector $\mathbf{x}$. Then, using that λ_1 is larger that all other eigenvalues,
$$A^k\mathbf{x}_0 = A^kP\mathbf{x} = PD^k\mathbf{x} = \sum_j \mathbf{p}_j\lambda_j^k x_j = \lambda_1^k \sum_j \mathbf{p}_j \left(\frac{\lambda_j}{\lambda_1}\right)^k x_j$$
As $k \to \infty$, $\mathbf{x}_k = A^k\mathbf{x}_0 \to \mathbf{p}_1\lambda_1^k x_1$ since all other terms converge to 0. That is, as we go though many iterations of the algorithm, we arrive at a vector parallel to $\mathbf{p}_1$, which is the eigenvector corresponding to the largest eigevalue. Normalizing the result will lead to a unit eigenvector parallel to $\mathbf{p}_1$, as claimed.