NBTW Centrality

October 13, 2022

Let A be the $n \times n$ adjacency matrix for an undirected network on n vertices. Let $D = \mathrm{diag}(A^2)$, the diagonal matrix of vertex degrees.

We have seen that Katz centrality counts the number of walks between vertex i and vertex j, down-weighting walks by their length. Let us re-cast that calculation so it may be more easily generalised. Let P_r be the matrix that counts the number of walks between all pairs of vertices, of length r (so $(P_r)_{ij}$ is exactly the number of walks of length r between vertex i and vertex j).

Then

$$P_r = AP_{r-1}$$

for $r \geq 1$ and $P_0 = I$. We wish to find (and evaluate)

$$Q = \sum_{r=0} \alpha^r P_r,$$

usually known as the "Generating Function". Here $\alpha > 0$ must be small enough so that Q converges. Multiplying the recurrence relationship by α^r and summing, we obtain

$$Q - P_0 = \alpha AQ.$$

Thus,

$$(I - \alpha A)Q = P_0 = I.$$

This is Katz centrality: $Q = (I - \alpha A)^{-1}$ will be non-negative and converging, provided $0 < \alpha < 1/\rho(A)$.

One of the problems with this is that these walks contain backtracking subsequences of edges.

A walk is backtracking if it contains a vertex sequence of the form uvu Otherwise it is a **nonbacktracking walk** (NBTW).

Suppose now that we wish to calculate the non-backtracking walk (NBTW) centrality for a large undirected network. Let A denote adjacency matrix for the network, as before, which is symmetric (since the network is undirected), and let us assume it has zeros along its diagonal (meaning that there are no self-loops in the network). Let D denote the usual nodal-degree diagonal matrix.

We define the matrix polynomial

$$M(\alpha) = I - \alpha A + \alpha^2(D - I).$$

Then we claim that we have, Q, denoting the non-backtracking walk version of the centrality matrix, given by

$$Q = (1 - \alpha^2)M(\alpha)^{-1}.$$

Here $\alpha \geq 0$ must be chosen to be small enough so that the spectrum of $M(\alpha)$ lies to the right of zero. This is certainly true for $\alpha < \rho(A)^{-1}$. There is a limiting value $\alpha^* \in (\rho(A)^{-1}, 1)$ at which $M(\alpha^*)$ has a zero eigenvalue, and its spectrum lies to the right of zero for all $\alpha \in [0, \alpha^*)$ [].

In fact Q is derived exactly as before (in the backtracking inclusive case), by writing it as a generating function, as follows. We follow (P. Grindrod, D. Higham, V. Noferini, The Deformed Graph Laplacian and Its Applications to Network Centrality Analysis March 2018, SIAM J. Matrix Anal. Appl.)

Let P_n be the matrix such that its (i, j)th term counts the total number of non-backtracking walks of length k from node i to node j within the network.

Consider an NBTW version of centrality. The following undirected network case was known (*Zeta functions of finite graphs and coverings*, Stark & Terras, Advances in Mathematics, 1996).

Suppose A. is as above Let P_r be the matrix that counts the number of NBTWs between all pairs of vertices, of length r (so $(P_r)_{ij}$ is exactly the number of NBTWs of length r between vertex i and vertex j).

Exercise. Show that we must have $P_0 = I,\ P_1 = A,\ P_2 = A^2 - D$; and the recursion:

$$P_r = AP_{r-1} - (D - I)P_{r-2}, \quad r \geq 3.$$

Solution.
The idea here is that the term AP_{r-1} over-counts, and has constructed and

included some backtracking walks. For each NBTW between vertex i and vertex j of length $r-2$, counted in P_{r-2}, starting out from vertex i, going to vertex k, say, first, there are exactly $D_i - 1$ walks that are non-backtracking except that they start by going *there and back* to each of the other $D_i - 1$ vertices connected to i (other than vertex k). That is, for each v_l adjacent to v_i but not equal to v_k (there are $D_i - 1$ of them), the (i, j)th term of AP_{r-1} has counted walks that begin

$$v_i v_l v_i v_k \ldots\ldots v_j.$$

So we have to subtract these by employing the $-(D-I)P_{r-2}$ term.
QED

Again we wish to find (and evaluate) the "Generating Function",

$$Q = \sum_{r=0} \alpha^r P_r.$$

Here, as before, $\alpha > 0$ must be small enough so that Q converges. Notice that $P_r \leq A^r$, since the latter includes a count of some backtracking walks. So Q certainly converges if $\sum_{r=0} \alpha^r A^r$ converges. This latter is a geometric progression just as in Katz, and converges, given by $(I - \alpha A)^{-1}$, iff $\alpha < 1/\rho(A)$, as above.

Exercise. Solve the above difference equation for Q. Show that

$$Q = (1-\alpha^2)(I - \alpha A + \alpha^2(D-I))^{-1}.$$

Note, there is no extra cost involved in making these calculations above that for simple Katz; since $I - \alpha A + \alpha^2(D-I)$ is just as sparse as $I - \alpha A$.

Solution.
Multiplying through by α^r. So

$$\alpha^r P_r = \alpha A \alpha^{r-1} P_{r-1} - \alpha^2(D-I)\alpha^{r-2}P_{r-2}, \quad r \geq 3.$$

Summing over $r \geq 3$, we have

$$Q - \alpha^2 P_2 - \alpha P_1 - P_0 = \alpha A(Q - \alpha P_1 - P_0) - \alpha^2(D-I)(Q - P_0)$$

$$(I - \alpha A + \alpha^2(D-I))Q = \alpha^2 P_2 + \alpha P_1 + P_0 - \alpha A(\alpha P_1 + P_0) + \alpha^2(D-I)P_0$$

$$= \alpha^2(A^2 - D) + \alpha A + I - \alpha A(\alpha A + I) + \alpha^2(D - I)$$
$$= \alpha^2 A^2 - \alpha^2 D + \alpha A + I - \alpha^2 A^2 - \alpha A + \alpha^2 D - \alpha^2 I.$$

So that
$$(I - \alpha A + \alpha^2(D - I))Q = (1 - \alpha^2)I.$$

This is valid providing α is small enough so that $Q > 0$ exists (the series converges).

There exists $\alpha^* \geq 1/\rho(A)$ such that $Q = \sum_{r=0} \alpha^r P_r$ converges for all $0 \leq \alpha < \alpha^*$.