3 Centrality Measures

A centrality measure, or index, on a network is a nonnegative function defined on each vertex, and that ranks the nodes according to their importance (hence the word "centrality") within the network itself. Many classical centrality indices were introduced in the 1950s, but research in this area is still very active at the time when these lecture notes are being written (early 2020s).

Of course, a problem with this idea is that what "importance" means is not clearly defined at all. As a consequence, mathematically it is difficult, when not impossible, to decide whether one centrality meaasure is better worse than another at quantifying importance of nodes. Thus, in practice choices are made by looking either at the application at hand, or at computational issues, or at a compromise between these two criteria.

Chances are that, indeed, intuitions of what importance should mean exists for many applications, but are application-dependent. As a consequence, a centrality measure may appear entirely appropriate and justified for one application, and not at all satisfactory for another. Regarding computation, one can compare the complexity of computing centrality measures; a class of centrality indices that is particularly interesting from this point of view is that of centrality based on the combinatorics of walks. The reason is that the latter can be computed by means of adjacency matrices. This implies that such centralities can be computed using numerical linear algebraic libraries, that are typically extremely reliable and computationally efficients.

While centralities can be defined and studied also for weighted or directed graphs, in the course we focus on simple graphs. Students interested in the more general scenario can investigate it as a project.

3.1 Degree centrality

One of the simplest possible centralities can be defined on its vertex simply by taking its degrees. The underlying intuition is the degree is a (naive) measure of how well connected a vertex is to other vertices. For a simple graph, we mention two ways of computing the degree centrality starting from the adjacency matrix:

1. As the degree of node i is the number of edges starting from node i, we can compute it as the sum of all the nonzero elements in the ith row of A. Therefore the vector of degree centralities is $d = Ae$.

2. The degree of node i is also equal to to the number of closed paths of length 2 starting from and ending on that vertex (why?). Hence, the vector of degree centralities is $d = \mathrm{diag}(A^2)$.

3.2 Katz centrality

The first centrality truly based on the combinatorics of walk that we analyse is called Katz centrality. The name is to honour its discoverer, as this measure was first introduced by the social scientist L. Katz in the 1950s. It is based on the idea to measure the centrality of a node as the weighted sum of all the possible walks departing by it, with a weight that decreases exponentially with the length of the walk. Intuitively, this means that the easier it is to explore the network starting from a given node the more important the node is. Moreover, the downweight based on the walk length can be interpreted by arguing that longer walks are more expensive means of exploration than shorter walks, and therefore have lesser importance for the node centrality.

These principles can be translated into formulae. Let us consider a network represented by a simple graph G having adjacency matrix A. Recall that the number of walks from node i to node j having length k is A^k_{ij}. We consider the sum of all walks of length k, downweighted by α^k where $\alpha \in (0, 1)$ is called the Katz parameter; moreover, to compute the centrality of node i, we sum over all possible endvertices j. As a result, we obtain that the centrality of node i is

$$v_i = \sum_{j=1}^{n} \sum_{k=0}^{\infty} \alpha^k A^k_{ij} \tag{1}$$

and therefore the vector of Katz centralities can be expressed as

$$v = \left(\sum_{k=0}^{\infty} \alpha^k A^k \right) e \tag{2}$$

where $e \in \mathbb{R}^n$ is the vector whose components are all equal to 1.

It is known in matrix theory that the matrix power series between brackets, equal to $I + \alpha A + \alpha^2 A^2 + \ldots$, and often called the *resolvent*, converges to $(I - \alpha A)^{-1}$ for all (possibly complex) values of α of modulus less than ρ^{-1} where ρ is the *spectral radius* of A, i.e., the modulus of the largest eigenvalue of A.

This is a general result: in the special case of our interest, where A is an adjacency matrix, we know (by Perron-Frobenius theory) that ρ is actually an eigenvalue. We give below a proof in the case where A is symmetric.

Theorem 3.1. *Let A be the adjacency matrix of an undirected graph such that its spectral radius is ρ. Then, for all (possibly complex) $|z| < \rho^{-1}$,*

$$\sum_{k=0}^{\infty} z^k A^k = (I - zA)^{-1}.$$

Proof. As A is a symmetric matrix, by the spectral theorem there exist an orthogonal matrix Q and a diagonal eigenvalue matrix Λ such that $A = Q\Lambda Q^T$. Moreover, all the diagonal entries of Λ satisfy $|\lambda_i| \leq \rho$. It is then easy to prove by induction on k that indeed $A^k = Q\Lambda^k Q^T$ for all $k \geq 0$. Hence,

$$\sum_{k=0}^{\infty} z^k A^k = Q \sum_{k=0}^{\infty} z^k \Lambda^k Q^T,$$

and therefore we may equivalenty study the equivalent infinite sum with the matrix Λ as an argument. Clearly, each diagonal entry of that sum has the form

$$\sum_{k=0}^{\infty} z^k \lambda_i^k = \frac{1}{1 - z\lambda_i}$$

as long as $|z\lambda_i| < |z|\rho < 1$. We conclude that

$$\sum_{k=0}^{\infty} z^k A^k = Q \begin{bmatrix} \frac{1}{1-z\rho} & & & \\ & \frac{1}{1-z\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{1-z\lambda_n} \end{bmatrix} Q^T.$$

Moreover

$$Q \begin{bmatrix} \frac{1}{1-z\rho} & & & \\ & \frac{1}{1-z\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{1-z\lambda_n} \end{bmatrix} Q^T (I - zA) =$$

$$Q \begin{bmatrix} \frac{1}{1-z\rho} & & & \\ & \frac{1}{1-z\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{1-z\lambda_n} \end{bmatrix} Q^T - Q \begin{bmatrix} \frac{z\rho}{1-z\rho} & & & \\ & \frac{z\lambda_2}{1-z\lambda_2} & & \\ & & \ddots & \\ & & & \frac{z\lambda_n}{1-z\lambda_n} \end{bmatrix} Q^T = I,$$

proving (by the uniqueness of the matrix inverse) that

$$Q \begin{bmatrix} \frac{1}{1-z\rho} & & & \\ & \frac{1}{1-z\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{1-z\lambda_n} \end{bmatrix} Q^T = (I - zA)^{-1}.$$

□

This analysis ytelds a restriction on the possibly allowed values of the Katz parameter α: indeed, Katz centrality is equal to

$$v = (I - \alpha A)^{-1} e \tag{3}$$

as long as $0 < \alpha < \rho^{-1}$.

From the computational viewpoint, (3) is very attractive. Indeed, Katz centrality is the solution of the linear system of equations $(I - \alpha A)v = e$. There exist extremely efficient software for the solution of linear systems of equations. If the cofficient matrix is sparse, the numerical methods that are employed in the available linear algebraic libraries are competitive even when the dimension of the system is as large as billions. This precisely the scenario that happens in practice for networks. Indeed, real life networks are often characterized by having a number of edges which is typically linear in the number of nodes, thus making the matrix $I - \alpha A$ very sparse (to see this, compare a number of nonzero entries equal to the n elements of the diagonal plus the number of edges, which is assumed to be linear in n, with the potential total number of n^2 nonzero elements: in the limit where n goes to infinity, only a square root of the worst-case number of elements are nonzero!). This fact, together with other "tricks" that exploit the special structure of the linear system (3) (features that can lead to speed-ups with respect to method for general linear systems include, for example, sparsity, the fact that the right hand side vector has all equal entries, the fact that the coefficient matrix is the adjacency matrix of a graph) can be exploited to efficiently compute Katz centrality even for extremely large networks.

This provides some evidence of the stament previously made that walk-based centrality measure are computationally attractive.

3.2.1 Variants of Katz centrality

Katz's idea has been extended in various way. To discuss how, let us think back about the proof of Theorem 3.1. Essentially, the basic idea was to exploit the spectral theorem to diagonalize A, and then work individually on the geometric power series of each eigenvalue. Then, we exploited the result in basic analysis that states that, for all $|z| < 1$, the geometric power series in z converges:

$$|z| < 1 \Rightarrow \sum_{k=0}^{\infty} z^k = \frac{1}{1-z}.$$

This led to the conclusion that, as long as $\alpha\rho < 1$, the matrix power series $\sum_{k=0}^{\infty} \alpha^k A^k$ also converges.

It is clear that this trick can be generalized, leading to modification of Katz's original idea. Indeed, suppose that $(c_k)_k$ is a sequence of nonnegative coefficients such that $\sum_{k=0} c_k z^k$ converges for $|z| < r$, r being the *convergence radius* of the power series. Then, similarly to Theorem 3.1, we can argue the convergence of the matrix power series

$$0 < \alpha < \frac{r}{\rho} \Rightarrow \sum_{k=0}^{\infty} c_k \alpha^k A^k < \infty. \tag{4}$$

Suppose that $f(z)$ is a function, analytic in $(0, r)$ such that $\sum_{k=0} c_k z^k = f(z)$. We can then *define* $f(\alpha A)$ as the limit of the matrix power series (4). This leads to the so-called *total f-total communicability centrality measure*: similarly to Katz centrality, f-total communicability of node i is defined as the weighted sum of all the walks that start from node i, with the weight for walks of length k being $c_k \alpha^k$. By slightly modifying the argument we gave for Katz centrality, we see that (for $0 < \alpha < r\rho^{-1}$), the f-total communicability can be computed as the vector

$$v = f(\alpha A)e.$$

(Note that, although it has got a special name for historical reasons, Katz centrality can now be seen as a particular case of f-total communicability, corresponding to the function $f(z) = (1-z)^{-1}$.) For certain common choices

of f, for example the exponential function ($c_k = 1/k!$), very efficient numerical methods exist so that this is not just a mathematical curiosity, but an effectively computable centrality. We observe, however, that f-total communicability can (at least mathematically, and ignoring for a moment the question of whether it can be computed efficiently) be defined for any function f analytic on open subinterval of $(0, 1)$ of positive measure, and such that the coefficients of its power series centered at 0 are nonnegative. It is possible to prove that the set of such functions is infinite and indeed even uncountable: therefore we now have constructed a fairly large zoo of family of centrality measures, each of them being defined for an (also uncountable) set of parameters α: definitely a large pool from which we can pick a centrality measure, and indeed, chossing a good one is an interesting problem. Similary, choosing the best value of the parameter α is an interesting and nontrivial question, which we will discuss in the next subsection.

We conclude this subsection by mentioning that other variants of Katz centrality are possible. For example, some authors have proposed that rather than counting every walk starting from node i, one could count only closed walks, i.e., walk starting from and ending on node i. This centrality measure is known as *f-subgraph centrality*. For a general function f, this leads to the formula

$$v_i = \sum_{k=0}^{\infty} c_k \alpha^k A_{ii}^k = f(\alpha A)_{ii},$$

that is, the vector of f-subgraph centralities is the vector containing the diagonal elements of the matrix $f(\alpha A)$ (formally defined as above). The problem of computing subgraph centrality is apparently more computationally demanding than that of computing total communicability. For example, for the subgraph analougue of Katz, one would need to compute all the diagonal entries of $(I - \alpha A)^{-1}$, which with a naive approach involves solving n linear systems. In practice, this problem has been carefully studied, and efficient methods (mainly based on so-called low-rank approximation of large matrices) have been devised so that even subgraph centralities can be approximated with very competitive computational costs, at least in the case of A being sufficiently sparse. A full discussion is beyond the goals of this course, but students interested in these aspects can ask for a project in this direction.

3.2.2 Limiting behaviour of Katz centrality

Katz centrality is defined for a whole interval of possible values for α, which generates the question of which one to pick. Generally, the ranking of the nodes do depend on α. Again, this is one question whose answer is likely to be application dependent, and particular choices may easily be excellent for one kind of network but poor for another. Let us prove the following theorem, which describes what the behaviour of the Katz centrality ranking in the limit of small and large values of the parameter.

Definition 3.1. *Let G be a simple connected graph with adjacency matrix A. The associated eigenvector centrality is the centrality measure induced by any positive Perron-Frobenius eigenvector of A.*

Theorem 3.2. *Consider a simple connected graph and suppose that there are no ties in the ranking of either degree centrality of eigenvector centrality. The limit of Katz centrality for $\alpha \to 0^+$ gives the same ranking of nodes as degree centrality. The limit of Katz centrality for $\alpha \to \rho^-$ gives the same ranking of nodes as eigenvector centrality.*

Proof. Consider the Taylor expansion in α

$$v = e + \alpha Ae + o(\alpha);$$

since e is a constant vector, it does not influence the ranking. Similarly, dividing by the scalar factor α does not change the ranking. Therefore, for sufficiently small value of α, v ranks the nodes in the same way as $Ae = \text{diag}(\Delta)$.

For the second part, observe that, by the very same arguments as in Theorem 3.1, we have that

$$(I - \alpha A)^{-1}e = \sum_{i=1}^{n} q_i \lambda_i(\alpha) q_i^T e \to \lambda_1(\alpha) q_1 q_1^T e + l.o.t.$$

and note that

$$\lim_{\alpha \to \rho^{-1}} q_1$$

must be the eigenvector corresponding to the null space of $(I - \rho^{-1}A)$, which is the Perron-Frobenius eigenvector of A. □

For the more general case of f-total communicability and f-subgraph centrality, a recent (2015) result by M. Benzi and C. Klymko shows that the very same limiting behaviour holds: the raking produced in the limit of $\alpha \to 0$ is the same as degree centrality, and the raking produced in the limit of α approaching the largest possible value (which depends both on f and A) is eigevenctor centrality, defined as above as the centrality measure prescribed by a positive Perron-Frobenius eigenvector of A. This elegant result shows that there is a qualitative universality (at least in the sense of the ranking produced by extremal values of α). A possible project that can be done for the course exam is to read in detail the research paper proving the Benzi-Klymko theorem, and expose it.