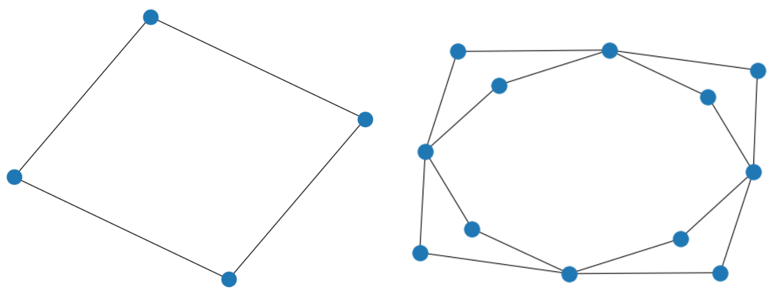
Network Science
Spring 2024
Problem sheet 2

---

1. Consider the following model for a growing undirected graph. Initial state: 2 nodes connected by 1 link. Iteration 1: Replace the link between the node pair with two new nodes and four new links arranged so that the two original nodes are only connected by 2 "fully distinct" length-2 paths. Here, two paths are "fully distinct" if they have zero links in common. Iteration $i+1$: Apply the process for iteration 2 to each linked node pair in the graph at iteration i. So, for each linked pair of nodes in the graph at iteration i, remove the link and replace it with two new nodes and four new links so that the two 'old' nodes are connected by 2 new fully distinct length 2 paths.

   (a) Draw the graph after iteration 1 (it should have 4 nodes and 4 links) and iteration 2.
   Solution:

   (b) Find expressions for the number of links and nodes after i iterations
   Solution: The number of links after the ith iteration is, $L_i = 4^i$. The number of nodes is:

   $$N_i = 2 + 2 + 2 * 4 + 2 * 4^2 + ... + 2 * 4^{i-1}.$$

   This can be rewritten as,

   $$N_i = 2 + 2\sum_{l=0}^{i-1} 4^l = 2 + \frac{2}{3}\left(4^i - 1\right)$$

   (c) How does the diameter depend on the number of nodes in the graph? Compare your result to rectangular lattice graphs and Cayley trees.
   Solution: The diameter is equal to the distance between the two original nodes in the graph and is given by, $D_i = 2^i$. So,

   $$N_i = 2 + \frac{2}{3}\left(D^2 - 1\right)$$

   or:

   $$D^2 = \frac{3}{2}\left(N_i - 2\right) + 1.$$

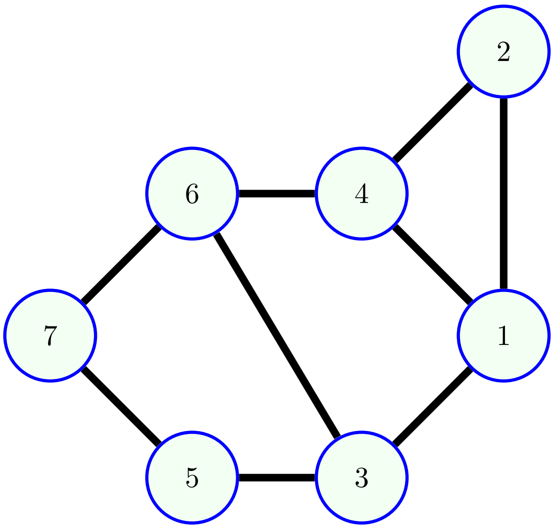
   We can see that for large N_i, the diameter will vary with the square root of the number of nodes which is the behavior we see for 2-d rectangular lattices.

2. Consider a simple connected graph where each node has the same degree, k. Show that the eigenvector centrality is the same for each node. Is the Katz centrality for each node the same?
Solution: The eigenvector centrality for a node is linearly proportional to the sum of centralities of its neighbors. Since each node has k neighbors, we can set the proportionality constant to $1/k$ and then each node can have the same (arbitrary) centrality. The Katz centrality for node i is, $x_i = \alpha \sum_{j=1}^{N} A_{ij}x_j + 1$. Say that each node has centrality β. Then, β must satisfy, $\beta = \alpha k \beta + 1$, or $\beta = 1/(1 - \alpha k)$. From our discussion of the eigenvector centrality and the Perron-Frobenius theorem, we know that the leading eigenvalue is k, so choosing α to be less than $1/k$ gives us the unique Katz centrality for node i which is the same for all i in the graph.

3. Consider the matrix $\mathbf{G}$ used to compute the PageRank centrality for graphs where each node has at least one out-link. Let $\mathbf{x}$ be an eigenvector of $\mathbf{G}$ corresponding to eigenvalue $\lambda = 1$ and let $\mathbf{x}$ contain at least two elements with differing sign.

   (a) Show that $\sum_{i=1}^{N} |x_i| < \sum_{i=1}^{N} \sum_{j=1}^{N} G_{ij}|x_j|$, and explain why this implies that such an eigenvector cannot exist
   Solution: The triangle inequality tells us that $|x_i| \leq \sum_{j=1}^{N} |G_{ij}x_j|$, and when $\mathbf{x}$ contains elements of different sign, this becomes a strict inequality. Then, since all elements of G are positive, we have, $|x_i| < \sum_{j=1}^{N} G_{ij}|x_j|$. Summing both sides over all i gives the desired expression, $\sum_{i=1}^{N} |x_i| < \sum_{i=1}^{N} \sum_{j=1}^{N} G_{ij}|x_j|$. Exchanging the order of the summations on the right-hand side and using the result from lecture that the sum of each column of G is one, this expression simplifies to, $\sum_{i=1}^{N} |x_i| < \sum_{j=1}^{N} |x_j|$ which is a contradiction, so all elements of this eigenvector must have the same sign.

   (b) Show that $x_i \geq 0$ for all i implies that $x_i > 0$ for all i (assuming $\mathbf{x}$ is non-trivial).
   Solution: $x_i = \sum_{j=1}^{N} G_{ij}x_j$ and $\mathbf{G}$ is positive, so for non-trivial $\mathbf{x}$ with each element non-negative, the RHS of this equation must be positive for each i.



4. Using the cosine similarity, determine which pairs of nodes in the graph above are the most similar
Solution: Four node pairs have 2 common neighbors: (3,7), (5,6), (3,4), and (1,6). The degrees of nodes 3 and 7, are 3 and 2 as are the degrees of nodes 6 and 5, so both pairs have the same similarity, $\sigma_{37} = \sigma_{56} = 2/\sqrt{6}$. The degrees of node 4 is 3, so, $\sigma_{34} = \sigma_{16} = 2/\sqrt{9}$

All other node pairs have one or zero common neighbors and smaller similarities. The degrees of the pairs with two common neighbors have a geometric mean of $\sqrt{6}$ or 3, so $2/\sqrt{6}$ is the largest cosine similarity in the graph.