1. Consider an undirected network. Recall that Katz-Bonacich centrality with parameter α is defined by the equation $\mathbf{c} = \alpha\mathbf{g}\mathbf{c}+\mathbf{1}$, where $\mathbf{c}$ is a $n\times 1$ vector and $\mathbf{g}$ is an $n\times n$ adjacency matrix.

a. Show that Katz-Bonacich centrality can also be written in the series form $\mathbf{c} = \mathbf{1}+\alpha\mathbf{g}\mathbf{1}+\alpha^2\mathbf{g}^2\mathbf{1}+\ldots$. (You may assume that this series converges.) Give an intuitive explanation of this formula.

Note that $\mathbf{c} = \mathbf{\Lambda 1}$ where $\Lambda = (I-\alpha\mathbf{g})^{-1}$ is the Leontief inverse, which in class we showed to equal $\mathbf{I}+\alpha\mathbf{g}+\alpha^2\mathbf{g}^2+\ldots$. Hence, $\mathbf{\Lambda 1} = \mathbf{1}+\alpha\mathbf{g}\mathbf{1}+\alpha^2\mathbf{g}^2\mathbf{1}+\ldots$. The intuition is that this measures the direct and indirect influence of each node i on all other nodes, where the indirect influence through a walk of length ℓ is discounted by α^ℓ.

b. Suppose that two nodes i and j satisfy $\mathbf{c}_i > \mathbf{c}_j$ for all sufficiently small values of the parameter α. What, if anything, can we conclude about the degrees of nodes i and j?

When α is small, $\mathbf{c} \approx \mathbf{1}+\alpha\mathbf{g}\mathbf{1}$, so $\mathbf{c}_i > \mathbf{c}_j$ iff $(\mathbf{g1})_i > (\mathbf{g1})_j$. This says that i has a greater degree than j.

c. Draw a network and label two nodes i and j for which you would guess that $\mathbf{c}_i > \mathbf{c}_j$ for sufficiently small values of α but $\mathbf{c}_i < \mathbf{c}_j$ for sufficiently large values of α. Explain your guess. You do not need to do any calculations or verify that your guess is correct.

This would be a network where i has greater degree than j but j's neighbors have higher centrality than i's neighbors. For example, consider a network consisting of two disjoint components, where the first component is a star with n nodes with i at the center, and the second component is a clique with $n-1$ nodes that includes node j.

2. Consider the DeGroot learning model with N agents with initial belief vector $x(0) = (x_1(0),\ldots,x_N(0))$ and an $N\times N$, non-negative, row-stochastic matrix T such that, for every period t, we have

$$x(t) = Tx(t-1).$$

a. Suppose that $N = 3$ and

$$T = \begin{pmatrix} \frac{3}{5} & 0 & \frac{2}{5} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

Compute the limit belief $x^* = \lim_{t\to\infty} x(t)$ as a function of $x(0)$. Which of the three agents is the most "influential"?

The inluence vector $\mathbf{s}$ is given by

$$\begin{aligned} s_1 &= \frac{3}{5}s_1 + \frac{1}{3}s_2 + \frac{1}{4}s_3, \\ s_2 &= \frac{1}{3}s_2 + \frac{1}{4}s_3, \\ s_3 &= \frac{2}{5}s_1 + \frac{1}{3}s_2 + \frac{1}{2}s_3, \\ s_1 + s_2 + s_3 &= 1. \end{aligned}$$

Solving this system gives

$$s_1 = \frac{15}{37},\ s_2 = \frac{6}{37},\ s_3 = \frac{16}{37}.$$

Thus, $x^* = \frac{15}{37}x_1(0) + \frac{6}{37}x_2(0) + \frac{16}{37}x_3(0)$. Agent 3 is the most influential.

b. Suppose that the matrix T is strongly connected and aperiodic, and that there is an agent i and a number t such that $(T^t)_{ji} = 0$ for all j (including $j = i$), where $(T^t)_{ji}$ denotes the (j, i) component of the t^{th} power of the matrix T. Prove that, for any two vectors of initial beliefs $x(0)$ and $\hat{x}(0)$ such that $x_j(0) = \hat{x}_j(0)$ for all $j \neq i$, we have $\lim_{t\to\infty} x(t) = \lim_{t\to\infty} \hat{x}(t)$. (You may appeal to any results proved in class, provided you cite them correctly.)

Since T is strongly connected and aperiodic, the limit belief $\lim_{t\to\infty} x(t)$ and the influence vector s are well-defined. For every number t, the influence vector s satisfies

$$sT^t = s$$

Since the i^{th} colum of the matrix T^t equals 0, this implies that $s_i = 0$. Hence, we have

$$\lim_{t\to\infty} x(t) = \sum_j s_j x_j(0) = \sum_{j\neq i} s_j x_j(0) = \sum_{j\neq i} s_j \hat{x}_j(0) = \lim_{t\to\infty} \hat{x}(t).$$

3. Consider the standard SIR model with basic reproduction number $R_0 > 1$. Suppose that the government has the ability to implement a *lockdown*. Assume that a lockdown prevents all new infections until the currently infected share of the population falls to a small number ε, at which point the lockdown is lifted. Assume that the government can only implement a lockdown once (for example, this could be because society will not tolerate multiple lockdowns). The government's problem is to choose the timing of the lockdown so as to minimize the long-run share of the population that ever gets infected, $\lim_{t\to\infty} R(t)$.

a. When should the government implement the lockdown? Your answer should take the form of a rule that the the government can implement if at each point in time t it knows what fraction of the population is susceptible, infected, and recovered. Assuming that the government locks down at the optimal time, what is $\lim_{t\to\infty} R(t)$?

A lockdown cannot prevent herd immunity from being reached, so the timing of the lockdown should be chosen to minimize overshooting. This is done by imposing the lockdown as soon as herd immunity is reached, as in this case there is no overshooting. This occurs at the time t such that $S(t) = 1/R_0$. Since there is no overshooting, we have $\lim_{t\to\infty} R(t) = 1 - 1/R_0$.

b. Now consider the behavioral SIR model covered in Lecture 11, and assume that the critical infection level I^* defined in that lecture is small enough that the path of the epidemic is different in the standard and behavioral versions of the model. Suppose the government's objective remains the same. Again, when should the government implement the lockdown? Will this point in time be reached earlier or later than in the standard SIR model? What is $\lim_{t\to\infty} R(t)$?

Again, the lockdown should be imposed as soon as herd immunity reached, which occurs at the time t such that $S(t) = 1/R_0$. However, this time is later than it was in part (a), because some people are vigilant for some period of time before herd immunity is reached. We again have $\lim_{t\to\infty} R(t) = 1 - 1/R_0$.

c. How does the long-run share of the population that ever gets infected differ in parts (a) and (b)? Intuitively, would you say that society is better-off in part (a) or part (b)? Explain.

The long-run share of the population that ever gets infected is the same in parts (a) and (b). Intuitively, society is probably better-off in part (a), because the same number of people get infected in both cases, but in part (b) some people also incur the cost of vigilance. (However, one could also argue that this is ambiguous, because people get infected later in part (b), so time discounting may be a reason why society could be better off in part (a). We accepted a range of reasonable answers here.) An explanation is that since in either case herd immunity is reached and there is no overshooting, society can be better-off reaching herd immunity quickly, as this saves on the cost of vigilance.