Kaicheng Luo

Assignment 3

Econ C103
December 6, 2019

Problem 1

a. Prove that $WB_k > 0$ can happen only if k has more than one link and some of k's neighbours aren't linked to one another.

Proof by Contradiction:
a.1) If k has only one link, then k cannot be the intermediary node in a shortest path. Hence $P(ij) = 0$, resulting in $WB_k = 0$. That contradicts with our assumption.
a.2) If all of k's neighbours are inter-connnected, and k is on at least one of the shortest path, then WLOG, suppose i-k-j is on one of the original shortest path. It directly follows that the detour to k seems unnecessary as we can shift to i-j excluding k and get an even shorter path. Contradiction!

b. $WB_k = 1$ for the center node of a star network.

The formula for WB_k is

$$WB_k = \sum_{path \in \mathcal{Q}^v} \frac{\frac{W_k(ij)}{P(ij)}}{\binom{n-1}{2}} \tag{1}$$

where

$$W_k(ij) = \frac{P_k(ij)}{l(ij) - 1} \tag{2}$$

The star network is symmetric, so we can arbitrary take 2 nodes i and j without loss of generality. Altogether, there're $\binom{n-1}{2}$ choices, all of which passes through the center. The length of the shortest path is always 2, so

$$W_k(ij) = \frac{P(ij)}{l(ij) - 1} = \frac{1}{1} = 1 \tag{3}$$

and the final result is

$$WB_k = \sum_{path \in \mathcal{Q}^v} \frac{1}{\binom{n-1}{2}} = \frac{\binom{n-1}{2}}{\binom{n-1}{2}} = 1 \tag{4}$$

c. Prove the uniqueness

1. By definition, $\frac{P_k(ij)}{P(ij)} \leq 1$, because a shortest path that passes though k is itself at least a shortest path.
2. Meanwhile, $l(ij) - 1 \geq 1$, because any path passing through contains at least 3 nodes.
3. The number of shortest paths shall be less or equal to $\binom{n-1}{2}$, which is the total number of possible pairs excluding k. Combining all of them, we have

$$WB_k = \sum_{path \in \mathcal{Q}^v} \frac{\frac{1}{l(ij)-1} \frac{P_k(ij)}{P(ij)}}{\binom{n-1}{2}} \leq 1 \tag{5}$$

The equal sign holds only if $l(ij) = 2$, $P(ij) = P_k(ij)$ for all k, and $|\mathcal{Q}^v| = \binom{n-1}{2}$, which itself implies a star network with k in the center containing all nodes.

d. Weighted Betweeness Measure compared with Standard Betweeness Measure

$$\begin{aligned}
WB_4 &= \frac{1}{21}\left\{(\frac{1}{3}+\frac{1}{4}+\frac{2}{5})+(\frac{1}{2}+\frac{1}{3}+\frac{2}{4})+(1+\frac{1}{2}+\frac{2}{3})\right\} = \frac{269}{1260} = 0.21 \\
B_4 &= \frac{12}{21} = \frac{4}{7} \\
WB_5 &= \frac{1}{21}\left\{(\frac{1}{4}+\frac{2}{5})+(\frac{1}{3}+\frac{2}{4})+(\frac{1}{2}+\frac{2}{3})+(1+\frac{2}{2})\right\} = \frac{31}{140} = 0.22 \\
W_5 &= \frac{12}{21} = \frac{4}{7}
\end{aligned} \tag{6}$$

Note that the weighted betweeness is different but the standard betweeness measure offers us identical results. Intuitively, the latter measure explains the general structure of the network yet the former takes the relative position into account.

Problem 2

Show that if $L(g) \leq M$, where M is a definite constant. then $\forall\, K > 0$, $\exists\, N > 0$ s.t. as long as $n > N$, $max\{d_i(g)\} > K$

Proof: If we can find such n_0 such that it defines the maximum number of vertices given a bound in dimension M and a bound in degree K, then let $n = n_0 + 1$, given the diameter is still bounded, $max\{d_i(g)\} > K$
So the original problem can be converted to finding n_0, or at least the upper bound for n_0. The way to construct this is to imitate a breadth first search (BFS), where we begin with one random node in the network, and try to add as many nodes as we can step by step (i.e. At each step, add the largest number of nodes, given the degree bound, if we increase diameter by 2, or radius by 1)
In the first step, the original node can have K neighbours. For each step that follows, the node on the outside layer can have $K-1$ neighbours. So the total vertices of the specific graph is bounded by

$$n_0 \leq 1 + \sum_{i=1}^{M/2} K(K-1)^i = 1 + \frac{K(K-1)^{M/2} - 1}{K-2} \tag{7}$$

$n_0 = M + 1$ when $K = 2$ Note that this process does not necessarily offer us the maximum network given $\{M, K\}$ (when M is odd). However in our case, an upper bound is enough. This finishes our proof.

Problem 3

$$g = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \hat{g} = \begin{bmatrix} 0 & 0.5 & 0 \\ 1 & 0 & 1 \\ 0 & 0.5 & 0 \end{bmatrix} \tag{8}$$

Katz Prestige Centrality

Note that $p^{(k)}(g)$ is given by $(I - \hat{g})p^{(k)}(g) = 0$. Hence $p^{(k)}(g)$ is the eigenvector of $\hat{g}$ corresponding to eigenvalue 1

$$\begin{bmatrix} 1 & -0.5 & 0 \\ -1 & 1 & -1 \\ 0 & -0.5 & 1 \end{bmatrix} p^{(k)}(g) = 0 \tag{9}$$

So the Katz prestige centrality is

$$p^{(k)}(g) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \tag{10}$$

Degree Centrality

$$d(g) = \begin{bmatrix} 0.5 \\ 1 \\ 0.5 \end{bmatrix} \tag{11}$$

Betweeness Centrality

$$Ce^B(g) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{12}$$

The Katz Prestige centrality is identical to degree centrality in this case, up to a constant multiplier. But they're different from betweeness centrality.

Some Numerical Examples

$$Ce^B(g, \frac{1}{2}, \frac{1}{4}) = (I - \frac{1}{2}g)^{-1}g\mathbf{1} = \begin{bmatrix} 0.86 \\ 1.43 \\ 0.86 \end{bmatrix} \tag{13}$$

$$Ce^B(g, \frac{1}{2}, \frac{1}{2}) = (I - \frac{1}{2}g)^{-1}g\mathbf{1} = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} \tag{14}$$

$$Ce^B(g, \frac{1}{2}, \frac{3}{4}) = (I - \frac{1}{2}g)^{-1}g\mathbf{1} = \begin{bmatrix} -10 \\ -14 \\ -10 \end{bmatrix} \tag{15}$$

Problem 4

The overall clustering is measured by

$$Cl(g) = \frac{\sum_{i=1}^{n} \#\{jk \in g | k \neq j, j \in N_i(g), k \in N_i(g)\}}{\sum_{i=1}^{n} \#\{jk | k \neq j, j \in N_i(g), k \in N_i(g)\}} \tag{16}$$

Note that

$$Cl_i(g) = \frac{\#\{jk \in g | k \neq j, j \in N_i(g), k \in N_i(g)\}}{\frac{1}{2}d_i(g)[d_i(g) - 1]} \tag{17}$$

and

$$\sum_{i=1}^{n} \#\{jk | k \neq j, j \in N_i(g), k \in N_i(g)\} = \frac{1}{2}\sum_{i=1}^{n} d_i(g)[d_i(g) - 1] \tag{18}$$

So we can rewrite $Ci(g)$ as

$$Cl(g) = \sum_{i=1}^{n} Cl_i(g) \frac{d_i(g)[d_i(g) - 1]}{\sum_{i=1}^{n} d_i(g)[d_i(g) - 1]} \equiv \sum_{i=1}^{n} Cl_i(g)\omega_i \tag{19}$$

where $\sum \omega_i = 1$, denoting the weights
Compare with the average cluster weight

$$Cl^{Avg}(g) = \sum_{i=1}^{n} Cl_i(g)\frac{1}{n} \tag{20}$$

Case 1: If $Cl_i(g) \geq Cl_j(g)$ when $d_i(g) \geq d_j(g)$, more weights are put on larger values, so $Cl(g) \geq Cl^{Avg}(g)$
Case 2: If $Cl_i(g) \leq Cl_j(g)$ when $d_i(g) \leq d_j(g)$, more weights are put on larger values, so $Cl(g) \leq Cl^{Avg}(g)$

Problem 5

Conditional degree distribution given two nodes are linked

The problem can be converted to the following conditional distribution

$$Pr(d_j = d \,|\, ij\, are\, linked) = \frac{Pr(d_j = d)Pr(ij\, are\, linked\, |d_j = d)}{Pr(ij\, are\, linked)} \tag{21}$$

Note that $Pr(ij\, are\, linked\,|d_j = d) = \frac{d}{n-1}$, because j with degree d is randomly connected among (n-1) choices. Meanwhile, $Pr(ij\, are\, linked) = \mathbb{E}(d) = (n-1)p$, by the property of Poisson distribution. The above quantity can be simplified as.

$$Pr(d_j = d\,|\,ij\, are\, linked) = \frac{P(d)d}{(n-1)p} \tag{22}$$

where

$$P(d) = \frac{[(n-1)p]^d}{d!}e^{-(n-1)p} \tag{23}$$

is the cononical Poisson distribution.
Remark: The result will be more intuitive if we use binomial distribution to depict the original degree distribution, assuming that each pair in the network is connected by a link with probability p, i.i.d.. The conditional distribution can be viewed as the degree distribution of a given node given that one of its neighbour is connected to it. Note that the links are independently drawn, so it simplifies to a sub problem where

$$Pr(d_j = d\,|\,ij\, are\, linked) = \tbinom{n-2}{d-1}p^{d-1}(1-p)^{n-1-d} \tag{24}$$

It is consistent with our former discussion in the following sense.

$$\begin{aligned} Pr(d_j = d\,|\,ij\, are\, linked) &= \frac{P(d)d}{< d >} = \frac{\binom{n-1}{d}p^d(1-p)^{n-1-d}d}{(n-1)p} \\ &= \tbinom{n-2}{d-1}p^{d-1}(1-p)^{n-1-d} \end{aligned} \tag{25}$$

Conditional degree distribution for the neighbour of a given node

The result is different due to the correlation of the degree distribution. Consider a simple case where n=4, A is connected to B, C and D. We'll use the binomial distribution as the original degree distribution for illustration purpose. Using Bayesian Rule, the conditional degree distribution can be represented as

$$\hat{P}(d) : \begin{cases} d=1 & Pr = (1-p)^3 + p(1-p)^2 \\ d=2 & Pr = \frac{2}{3}3p(1-p)^2 + \frac{2}{3}3p^2(1-p) \\ d=3 & Pr = p^3 + p^2(1-p) \end{cases} \tag{26}$$

where all the possible number of links are depicted by a binomial distribution. The conditional probability is a simple application of Bayesian rule. That's apparently different from the discussion in the first part of the problem where

$$\tilde{P}(d) = \tbinom{2}{d-1}p^{d-1}(1-p)^{3-d} \tag{27}$$

The asymptotic behaviour of the two degree distributions are identical, namely,

$$\lim_{n\to\infty,(n-1)p=m} \hat{P}(d) = \lim_{n\to\infty,(n-1)p=m} \tilde{P}(d) \tag{28}$$

because the properties of Poisson distribution tells us that when n goes to infinity, the degree distribution of nodes are (asymptotically) uncorrelated.

Explanation about the difference

The underlying reason for the difference is the correlation of the degree distribution of neighbour's given a randomly chosen node. When n is small, the correlation is significant. So when we're applying Bayes theorem updating our belief about our neighbour's distribution, not only are we concerned about the degree distribution of a given neighbour, but also the joint distribution of all of them. Due to c So the two probabilities will converge on each other.

Problem 6

Show that $q = 1 \iff P(0) = 0$

If $q = 1$, then LHS=0, All exponential terms of 0 is 0, except $0^0 = 1$. So the only term left is

$$1 * P(0) = 0 \tag{29}$$

If $P(0) = 0$, on the other hand, then the equation can be simplified as

$$(1-q)(1-P(1)-(1-q)P(2)-(1-q)^2P(3)-...) = 0 \tag{30}$$

It directly follows that $P(0) = 0$
Apparently $q = 1$ is a solution. Note that it's indeed the only solution, because $\sum P(i) = 1$, the right bracket can never be zero unless q=1.

Find a non-zero solution of q

Note that $P(0) = \frac{1}{3}$, $P(2) = \frac{2}{3}$

$$1-q = (1-q)^0P(0) + (1-q)^2P(2) \tag{31}$$

$$1-q = \frac{1}{3} + \frac{2}{3} + \frac{2}{3}q^2 - \frac{4}{3}q \tag{32}$$

$$\frac{2}{3}q^2 - \frac{1}{3}q = 0 \tag{33}$$

The solutions are $q_1 = 0$, $q_2 = \frac{1}{2}$. Where $q = \frac{1}{2}$ is the non-zero solution we want.

Problem 7

The probability that an efficient network forms

We've already proved in class that only star networks are efficient in the distance-based utility model. When $n = 4$, denote the nodes as $\{a, b, c, d\}$

Step 1: Randomly form a link, without loss of generality, say ab is formed. $Pr(Efficient) = 1$
Step 2: To ensure that it's a star network, one of the end of the former link must be an end of the new link. $Pr(Efficient) = \frac{4}{5}$, WLOG, say ac is formed
Step 3: ad must be formed $Pr(Efficient) = \frac{1}{4}$
The final probability is the multiplication of the three

$$Pr(Efficient) = 1 * \frac{4}{5} * \frac{1}{4} = \frac{1}{5} \tag{34}$$

Problem 8

The utility function in the coauthor model is

$$u_i(g) = \sum_{j:i,j \in g} \left(\frac{1}{d_i(g)} + \frac{1}{d_j(g)} + \frac{1}{d_i(g)d_j(g)}\right) \tag{35}$$

As is proposed by Jackson and Wolinsky (Jackson P.167), the efficient network consists of $\frac{n}{2}$ separate pairs. The pairwise stable network, on the other hand, is inefficient and can be partitioned into fully intraconnected components.
In the efficient network, $d_i(g) = 1$ for all i in the network

$$u_i(g) = 1 + 1 + 1 = 3 \tag{36}$$

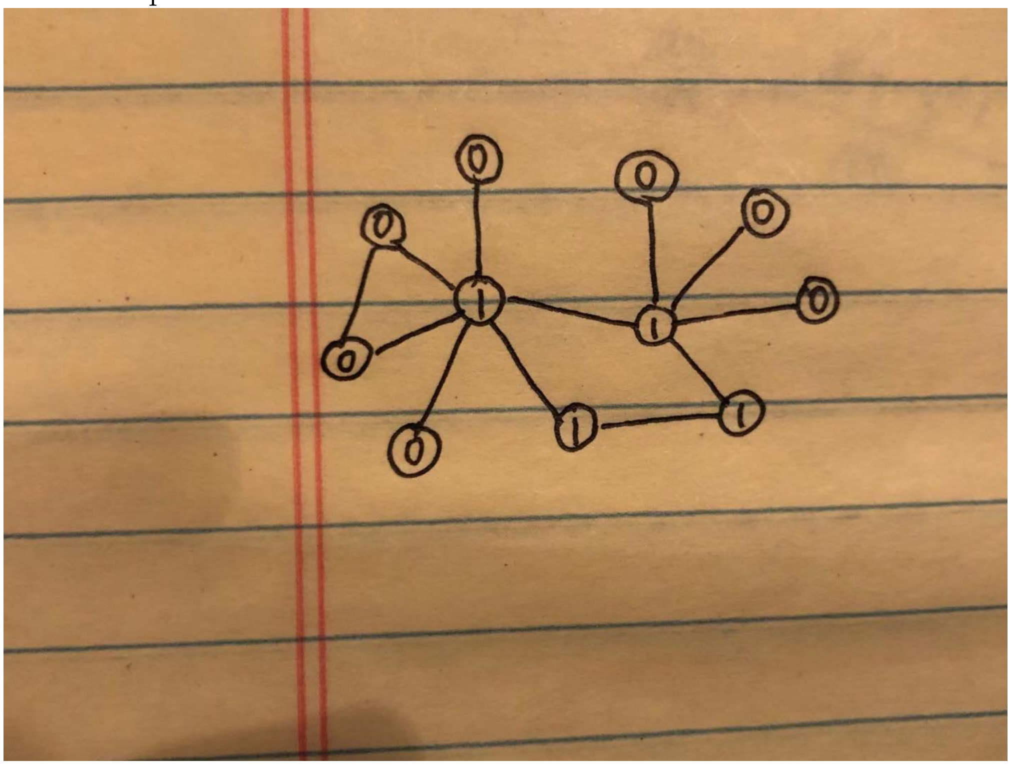
In a pairwise stable network, we have all the components of the network are intraconnected. So $d_i(g) = d_j(g)$ if $link(i, j) \in g$

$$\tilde{u}_i(g) = \sum_{j:i,j \in g} \left(\frac{2}{d_i(g)} + \frac{1}{d_i^2(g)}\right) = d_i(g)\left(\frac{2}{d_i(g)} + \frac{1}{d_i^2(g)}\right) \leq u_i(g) \tag{37}$$

Hence for each node in the network, one can improve his utility without any negative effect on others. The pairwise stable network is hence Pareto dominated by the efficient network.

Problem 9

Another equilibrium is constructed as follows:



Problem 10

The utility function in a local public goods model is

$$u_i(x_i, x_{N_i(g)}) = f(x_i + \sum_{j \in N_i(g)} x_j) - c(x_i) \tag{38}$$

where f is continuously differentiable and strictly concave. c is continuously differentiable and strictly convex.

Complete network

In the equilibrium of a complete network, all player's actions shall be symmetric, because

$$F.O.C. \quad \frac{d}{dt} u_i(x_i, x_{N_i(g)}) = f^{'}(\sum_{i=1}^{n} x_i) - c^{'}(x_i) = 0 \tag{39}$$

$$\therefore c^{'}(x_i) = c^{'}(x_j) \quad \forall i, j \in \{1, 2, ..., n\} \tag{40}$$

Assume there exists $\tilde{x} > 0$ such that $f^{'}(n\tilde{x}) = c^{'}(\tilde{x})$. Then we can propose that the action set that everyone contributes $\tilde{x}$ is the only pure strategy Nash equilibrium.
$Proof$: Let $g(x_i) = f^{'}(x_i + (n-1)\tilde{x}) - c^{'}(x_i)$. As is stated in the assumptions, $f^{''}(x) < 0$ and $c^{''}(x) > 0$. So $g(.)$ is monotonously decreasing w.r.t. x_i. Therefore the solution of $g(x_i) = 0$ is unique. Note that $f^{'}(n\tilde{x}) = c^{'}(\tilde{x})$. If and only $x_i = \tilde{x}$ for all i Will we satisfy all the first order condition. There's no profitable deviation from the Nash equilibrium.

Circle Network

A specialized equilibrium in such a network consists of half of the players contributing 0 and another half contributing x^* where $f^{'}(x^*) = c^{'}(x^*)$.
To prove that this is a Nash equilibrium, we have to show that for each player there's no profitable deviation given other player's action.
For those who contribute x^*, his neighbour's do not contribute. Therefore his best response correspondence is identically x^*

$$F.O.C. \quad g(x_i) = \frac{d}{dt}u_i(x_i, x_{N_i(g)}) = f^{'}(x_i) - c^{'}(x_i) = 0 \tag{41}$$

It follows similarly as the last subsection. $f^{''}(x) < 0$ and $c^{''}(x) > 0$. So $g(.)$ is monotonously decreasing w.r.t. x_i. Hence the solution of $g(x_i) = 0$ is unique ($x_i = x^*$)
For those who do not contribute, their two neighbours contributes x^*

$$\begin{aligned} u_i(x_i, x_{N_i(g)}) &= f(2x^*) \\ \frac{d}{dx_i}u_i(x_i, x_{N_i(g)})\Big|_{x_i=0} &= f^{'}(2x^*) - c^{'}(x^*) < 0 \end{aligned} \tag{42}$$

There's also no profitable deviation from contributing zero. Summing up, the scenario we proposed is a pure strategy Nash equilibrium.

Problem 11

The correct proposition shall be stated as follows:
A is the original perturbation where we set $a_i^0 = 1$, $B \cup A$ is the eventual set of nodes playing 1 if and only if C is more than (1-q)-cohesive and for any nonempty subset D of B, $D \cup C$ has a cohesiveness of no more than 1-q.
$Proof$: First show that it's a necessary condition.
If $B \cup A$ is the eventual set of nodes playing 1, then for any node in C, the fraction of its neighbors in set $B \cup A$ must be less than q, otherwise the best response of it will be contributing 1 in the coordination game. It directly follows that the fraction of any of its neighbors in the set $(B \cup A)^c$ is more than 1-q. That's equivalent to say that $(B \cup A)^c$ is more than (1-q)-cohesive.

Moreover, if $D \subseteq B$ and $D \cup C$ is more than (1-q)-cohesive. Then less than a fraction of q of its neighbors will be in the set $B \cup A$, the nodes in D won't be affected at the very beginning, hence contradiction. $(D \cup C)$ must be less or equal to (1-q)-cohesive
Next let's establish sufficiency.
if C is more than (1-q)-cohesive and for any nonempty subset D of B, $D \cup C$ has a cohesiveness of no more than 1-q. First consider $D = B$. $(B \cup C)^c = A$. The cohesiveness of A is no more than 1-q implies that there exists at least one node i with at least q neighbors in A. Hence $a_i^1 = 1$. Let $A_1 = A \cup \{i\}$ denote the set of players playing 1 (containing i). Then by the same procedure we can let $D = B \backslash \{i\}$. $(D \cup C)^c = A \cup \{i\}$. The cohesiveness of that is no more than 1-q implies that there exists at least one node j with at least q neighbors in $A \cup \{i\}$. Hence $a_j^2 = 1$. Let $A_2 = A \cup \{i\} \cup \{j\}$ denote the set of players playing 1 (containing i and j) and repeat the same procedure. By induction, we construct a strictly growing sequence $\{A_i\}$ until all of the set B is contained. Due to the fact that C is more than (1-q)-cohesive. The contagion will not have enough influence on C, so it ends when $A \cup B$ is affected.
$\square$