Problem Set V

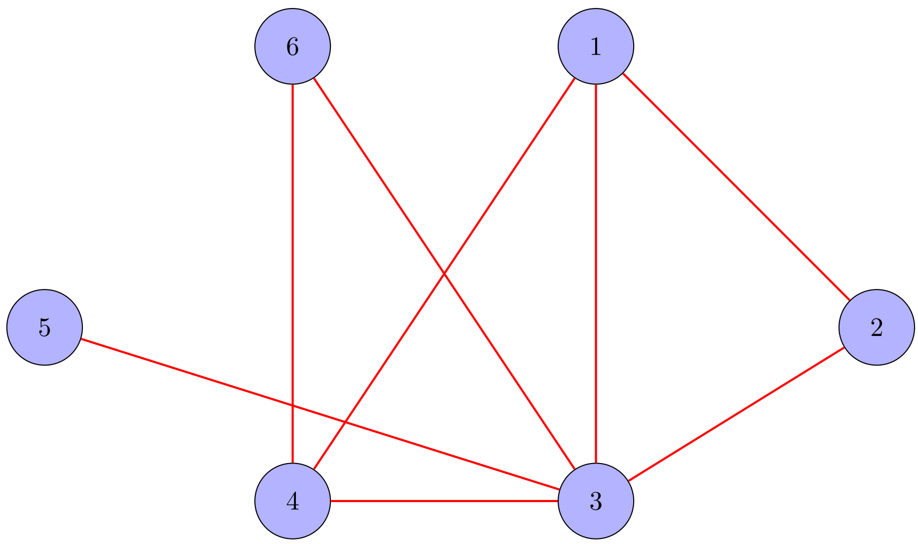
ECO 7427 - Econometric Theory II

Armand Kapllani
University of Florida Economics
Spring 2018

Social Networks

Let $N = \Big\{1, 2, 3, 4, 5, 6\Big\}$ and consider the following network (N, g).

Figure 1: Undirected and Unweighted Network



Construct the adjacency matrix:

$$g = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

To compute the degree of each vertex we sum across each row

$$d_i(g) = \sum_j g_{ij} \tag{1}$$

The neighborhood of each vertex:

Table 1: Neighborhood of each vertex

Vertex	$N_i(g)$
N = 1	$\{2, 3, 4\}$
N = 2	$\{1, 3\}$
N = 3	$\{1, 2, 4, 5, 6\}$
N = 4	$\{1, 3, 6\}$
N = 5	$\{3\}$
N = 6	$\{3, 4\}$

The degree of each node:

Table 2: Degree of each vertex in the network

Vertex	Degree
N = 1	3
N = 2	2
N = 3	5
N = 4	3
N = 5	1
N = 6	2

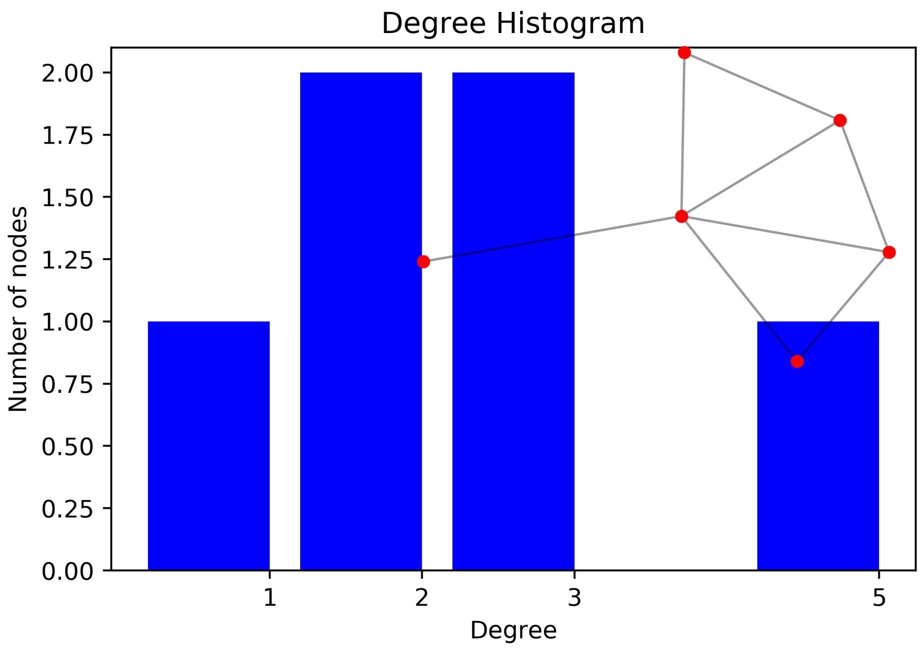
Average degree of the network $g \in G(N)$:

$$\frac{1}{N} \sum_{i \in N} d_i(g) = \frac{1}{6} \left[3 + 2 + 5 + 3 + 1 + 2 \right] = \frac{8}{3} \tag{2}$$

The density of the network $g \in G(N)$:

$$\frac{\sum_{i \in N} d_i(g)}{n(n-1)} = \frac{\frac{1}{n} \sum_{i \in N} d_i(g)}{n-1} = \frac{8/3}{5} = \frac{8}{15} \tag{3}$$

Figure 2: Degree Histogram



The individual clustering of the six vertices:

$$CI_i(g) = \frac{\left\{jk \in g | k \neq j, j \in N_i(g), k \in N_i(g)\right\}}{d_i(g)\left[d_i(g) - 1\right]} \tag{4}$$

Table 3: Individual clustering of each vertex

Vertex	$CI_i(g)$
N = 1	2/3
N = 2	1
N = 3	3/10
N = 4	2/3
N = 5	0
N = 6	1

The average clustering coefficient:

$$CI^{\text{Avg}}(g) = \frac{1}{n} \sum_{i \in N} CI_i(g) = \frac{1}{6}\left[\frac{2}{3} + 1 + \frac{3}{10} + \frac{2}{3} + 0 + 1\right] = \frac{109}{180} \tag{5}$$

Degree centrality for each vertex:

$$CI_i(g) = \frac{\Big\{jk \in g|k \neq j, j \in N_i(g), k \in N_i(g)\Big\}}{d_i(g)\big[d_i(g) - 1\big]} \tag{6}$$

Table 4: Degree centrality for each vertex

Vertex	$CI_i(g)$
N = 1	3/5
N = 2	2/5
N = 3	1
N = 4	3/5
N = 5	1/5
N = 6	2/5

Hence from Table 4 we can see that vertex 3 is well connected in terms of direct connections as it has the highest degree of centrality. However using this measure alone we cannot tell if vertex 3 is well located in the network (N, g).

Closeness centrality fo each vertex:

$$C_i(g) = \frac{1}{\frac{1}{n-1}\sum_{j\neq i} l(i,j)} \tag{7}$$

where $l(i,j)$ denotes the number of links in the **shortest path** between i and j.

Table 5: Closeness centrality fo each vertex

Vertex	$C_i(g)$
N = 1	5/7
N = 2	5/8
N = 3	1
N = 4	5/7
N = 5	5/9
N = 6	5/8

Decay centrality for each vertex:

$$D_i^{\delta}(g) = \sum_{j\neq i} \delta^{l(i,j)} \tag{8}$$

where $l(i,j)$ denotes the number of links in the **shortest path** between i and j.

Table 6: Decay centrality fo each vertex

Vertex	$D_i^\delta(g)$	$\delta = 1/2$
N = 1	$\delta(3+2\delta)$	2
N = 2	$\delta(2+3\delta)$	7/4
N = 3	5δ	5/2
N = 4	$\delta(3+2\delta)$	2
N = 5	$\delta(1+4\delta)$	3/2
N = 6	$\delta(1+4\delta)$	3/2

Betweenness centrality for each vertex:

$$Ce_i^B(g) = \sum_{k \neq j : i \notin \{k,j\}} \frac{P_i(kj)/P(kj)}{\frac{1}{2}(n-1)(n-2)} \tag{9}$$

Table 7: Betweenness centrality of a vertex

Vertex	$P_i(k,j)$	$Ce_i^B(g)$
N = 1	{23, 24, 25, 26, 34, 35, 36, 45, 46, 56}	1/20
N = 2	{13, 14, 15, 16, 34, 35, 36, 45, 46, 56}	0
N = 3	{12, 14, 15, 16, 24, 25, 26, 45, 46, 56}	3/5
N = 4	{12, 13, 15, 16, 23, 25, 26, 35, 36, 56}	1/20
N = 5	{12, 13, 14, 16, 23, 24, 26, 34, 36, 46}	0
N = 6	{12, 13, 14, 15, 23, 24, 25, 34, 35, 45}	0

Notice that for the first vertex ($N = 1$) the link 24 shown in red in the set $P_i(k,j)$ can either go through vertex 1 or through vertex 3 since both are geodesic.

Eigenvector centrality for each node:

We solve the following problem numerically to compute the eigencentrality of each vertex in the graph.

$$\lambda C^e(g) = g C^e(g) \tag{10}$$

where $C^e(g)$ is left-hand eigenvector of the adjacency matrix g and λ is the corresponding eigenvalue.

where $|\lambda_i| \leq \lambda_1 = 3.04$ ($\forall i \neq 1$) is the **Perron-Frobenius eigenvalue** of the adjacency matrix g also known as the **spectral radius**[1] $\rho(g)$. Eigencentrality is a measure of the influence of a node in a network.

[1]Theorem: Let $g \in C^{n\times n}$ with spectral radius $\rho(g)$, then $\rho_A < 1$ if and only if $\lim_{g\to\infty} g^k = 0$ and if $\rho(g) > 1$ then $\lim_{k\to\infty} \|g^k\| = \infty$

Table 8: Largest eigenvalue and its eigenvector

Vertices	$\lambda_{max} = \rho(g) =$ **3.04**
N = 1	0.45
N = 2	0.34
N = 3	**0.58**
N = 4	0.45
N = 5	0.19
N = 6	0.34

Katz-Bonacich Centrality:

$$C^B(g, a, b) = (I_n - bg)^{-1} ag\mathbf{1}_n \tag{11}$$

Table 9: Number of walks emanating from each vertex

Vertices	k = 1	...	k = ∞
N = 1	3	...	∞
N = 2	2	...	∞
N = 3	5	...	∞
N = 4	3	...	∞
N = 5	1	...	∞
N = 6	2	...	∞

where we would expect that the number of walks goes to infinity as $\lim_{k\to\infty} \left\|A^k\right\| = \infty$

Table 10: Katz-Bonacich centrality

Vertices	$\alpha = 1,\ \beta = 1/6$	$\alpha = 1,\ \beta = 1/3$
N = 1	6.32	4.99
N = 2	4.56	3.55
N = 3	**9.05**	**7.37**
N = 4	6.32	4.99
N = 5	2.50	1.92
N = 6	4.56	3.55

We also compute Katz-Bonacich centrality using an algorithmic approach in Python

$$x_i = \beta \sum_j A_{ij} x_j + \alpha \tag{12}$$

where equation (12) is iteratively computed and the parameter α controls the initial centrality and $\beta < \frac{1}{\lambda_{\max}}$. We use NetworkX package and function katz.centrality() and the results are

Table 11: Katz-Bonacich centrality using NetworkX in Python

Vertices	$\alpha = 1/6$, $\beta = 1$	$\alpha = 1/8$, $\beta = 1$
N = 1	0.43	0.42
N = 2	0.37	0.38
N = 3	**0.52**	**0.50**
N = 4	0.43	0.42
N = 5	0.30	0.32
N = 6	0.37	0.38

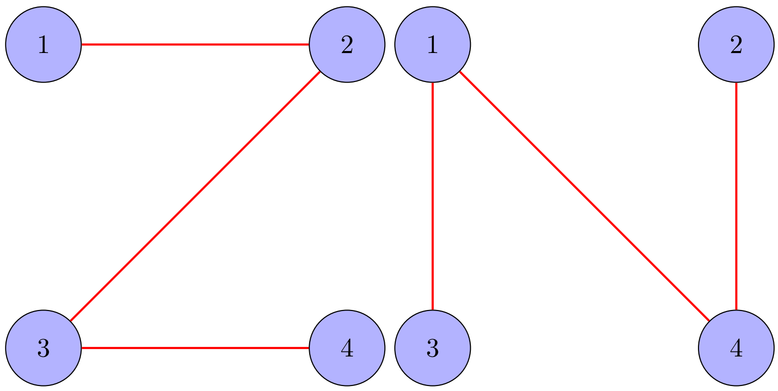
Graph Theory

Show that if a network is not connected, then its complement is.

Proof. Let $G = (V, E)$ be a graph comprising of a set of vertices V and a set of edges E. Let $< v_i, v_j >$ denote an edge between two arbitrary vertices v_i and v_j in the graph G where $< v_i, v_j > \in E$ and let $\tilde{G} = (V, \tilde{E})$ be the complement of the graph G. Since G is not a connected graph then we can partition the graph G into two disjoint sets of vertices V_1 and V_2 where $\forall\ v_1 \in V_1$ and $\forall\ v_2 \in V_2$ we have that $< v_1, v_2 > \notin E$. Hence for all $v_1 \in V_1$ and $v_2 \in V_2$ we have that $< v_1, v_2 > \in \tilde{E}$ which implies that $\tilde{G}$ is a connected graph. $\square$

Provide an example of a four-node network that is connected and its complement is also connected:

Figure 3: Left: $\big(\{1, 2, 3, 4\}, g\big)$, right: $\big(\{1, 2, 3, 4\}, g'\big)$



Correlation Scatter Plots

Figure 4: Scatter plot

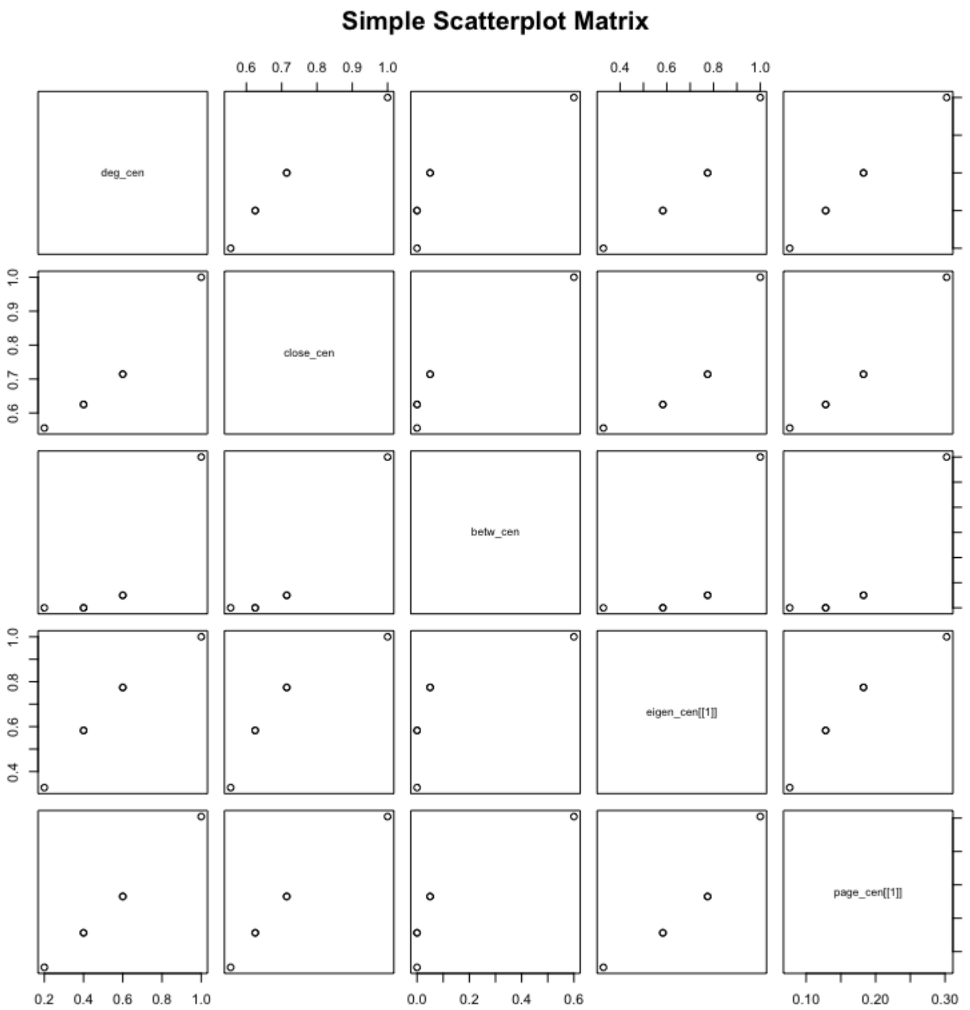


Figure 5: Scatter plot