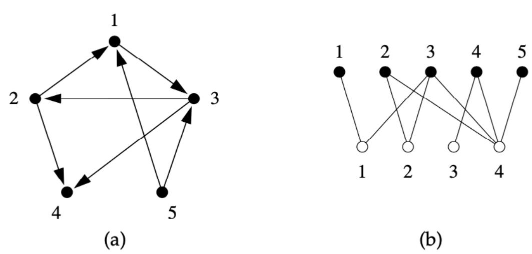
Complex Networks and Statistical Learning
Homework 3

Chenghua Liu
liuch18@mails.tsinghua.edu.cn
Department of Computer Science
Tsinghua University

1 Adjacency Matrix

1.1 ”Networks”, Exercise 6.3

Consider the following two networks:



Network (a) is directed. Network (b) is undirected but bipartite. Write down:

a) The adjacency matrix of network (a);

b) The incidence matrix of network (b);

c) The projection matrix (Eq. (6.10)) for the projection of network (b) onto its black nodes.

Solution:

(a)

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

(b)

$B_{ij} = 1$ if and only if ith vertex of the first kind (white) is connected to jth vertex of the second kind(black) .

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

(c)

We calculate $B^T B$:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 \\ 1 & 2 & 3 & 1 & 1 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

1.2 "Networks", Exercise 6.4

Let $\mathbf{A}$ be the adjacency matrix of an undirected network and 1 be the column vector whose elements are all 1. In terms of these quantities write expressions for:

a) The vector $\mathbf{k}$ whose elements are the degrees k_i of the nodes;

b) The number m of edges in the network;

c) The matrix $\mathbf{N}$ whose element N_{ij} is equal to the number of common neighbors of nodes i and j;

d) The total number of triangles in the network, where a triangle means three nodes, each connected by edges to both of the others.

Solution:
(a) $\mathbf{A} \times \mathbf{1}$
(b) $\frac{\mathbf{1}^T \times \mathbf{A} \times \mathbf{1}}{2}$
(c) $\mathbf{A}^2$
(d) $\frac{trace(\mathbf{A}^3)}{3!} = \frac{trace(\mathbf{A}^3)}{6}$

2 Directed Acyclic Graph

2.1 "Networks", Exercise 6.7

Consider an acyclic directed network of n nodes, labeled $i = 1 \ldots n$, and suppose that the labels are assigned in the manner of Fig. 6.3 on page 111, such that all edges run from nodes with higher labels to nodes with lower.

a) Write down an expression for the total number of ingoing edges at nodes $1 \dots r$ and another for the total number outgoing outgoing at nodes $1 \dots r$, in terms of the in- and out-degrees k_i^{in} and k_i^{out} of the nodes.

b) Hence find an expression for the total number of edges running to nodes $1 \dots r$ from nodes $r+1 \dots n$

c) Hence or otherwise show that in any acyclic network the in-and out-degrees must satisfy

$$k_r^{\text{in}} \leq \sum_{r+1}^{n} \left(k_i^{\text{out}} - k_i^{\text{in}} \right), \quad k_{r+1}^{\text{out}} \leq \sum_{i=1}^{r} \left(k_i^{\text{in}} - k_i^{\text{out}} \right) ,$$

for all r.

Solution:

(a) (in , out) at nodes 1 to r:

(2,0) (2,0) (1,2) (2,0) (2,2) (1,2) (2,2) (0,2) (0,2)

(b) The total number of edges running to nodes $1 \dots r$ from nodes $r+1 \dots n$ in going=$\sum_{i=0}^{i=r} k_i^{in}$, and out going =$\sum_{i=0}^{i=r} k_i^{out}$

(c) This conclusion is extremely obvious bacause that sum of out going equal to sum of in going.

3 Path

3.1 "Networks", Exercise 6.12(a)

Consider the set of all paths from node s to node t on an undirected network with adjacency matrix $\mathbf{A}$. Let us give each path a weight equal to α^r, where α is a constant and r is the length of the path.

a) Show that the sum of the weights of all the paths from s to t is given by the st element of the matrix $\mathbf{Z} = (\mathbf{I} - \alpha\mathbf{A})^{-1}$, where $\mathbf{I}$ is the identity matrix.

b) What condition must α satisfy for the sum to converge?

c) Hence, or otherwise, show that the length ℓ_{st} of the shortest path from s to t, if there is one, is

$$\ell_{st} = \lim_{\alpha \to 0} \frac{\partial \log Z_{st}}{\partial \log \alpha}$$

Solution:

(a)

Let N_{st} be the number of shortest path from t to s. i.e., $N_{st}^r = \left[A^T\right]_{st}$. Then we have the sum of the weight

$$\sum_{r=0}^{\infty} \alpha^r N_{st}^r = \left[\sum_{r=0}^{\infty} (A\alpha)^r\right]_{st}$$

If $Z = \sum_{r=0}^{\infty} (A\alpha)^r$ the proof is over. So we next to proof it.

$$(I - \alpha A) \sum_{r=0}^{\infty} (A\alpha)^r = I - (A\alpha)^{n+1}$$

Notice that $\lim_{n\to\infty} (A\alpha)^n = 0$,which means $(I - \alpha A)$ is invertible.Then $(I - \alpha A)^{-1} = \lim_{n\to\infty} \sum_{r=0}^{\infty} (A\alpha)^r$ .We get $Z = \sum_{r=0}^{\infty} (A\alpha)^r$.

(b) $|\alpha| < \frac{1}{\rho(A)}, \rho(A)$ is the spectral radius of A.

(c)

By chain rule, we have

$$\frac{\partial \log Z_{st}}{\partial \log \alpha} = \frac{\alpha}{Z_{st}} \frac{\partial Z_{st}}{\partial \alpha}$$

Note that

$$Z_{st} = Z = \sum_{\gamma=0}^{\infty} (A\alpha)_{st}^r$$

And the shortest path $l_{st} = \min\left\{r : [A^r]_{st} > 0\right\}$

$$\begin{aligned}
&Z_{st} = \alpha^{l_{st}} \textstyle\sum_{r=0}^{\infty} \alpha^r \left(A^{r+l_{st}}\right)_{st} \\
&\Rightarrow \tfrac{\partial Z_{st}}{\partial \alpha} = l_{st} \left(\alpha^{l_{st}-1}\right) \textstyle\sum_{r=0}^{\infty} \alpha^r \left(A^{r+l_{st}}\right)_{st} + \alpha^{l_{st}} \textstyle\sum_{r=0}^{\infty} r\alpha^{r-1} \left(A^{r+l_{st}}\right)_{st}
\end{aligned}$$

Then

$$\begin{aligned}
\frac{\partial \log Z_{st}}{\partial \log \alpha} &= \frac{l_{st} \left(\alpha^{l_{st}}\right) \sum_{r=0}^{\infty} \alpha^r \left(A^{r+l_{st}}\right)_{st} + \alpha^{l_{st}} \alpha \sum_{r=0}^{\infty} r\alpha^{r-1} \left(A^{r+l_{st}}\right)_{st}}{\alpha^{l_{st}} \sum_{r=0}^{\infty} \alpha^r \left(A^{r+l_{st}}\right)_{st}} \\
&= l_{st} + \frac{\alpha \sum_{r=0}^{\infty} r\alpha^{r-1} \left(A^{r+l_{st}}\right)_{st}}{\sum_{r=0}^{\infty} \alpha^r \left(A^{r+l_{st}}\right)_{st}}
\end{aligned}$$

the second term tends to zero, hence

$$l_{st} = \lim_{a\to 0} \frac{\partial \log Z_{st}}{\partial \log \alpha}$$

4 Katz Centrality

4.1 "Networks", Exercise 7.1

Consider a connected k -regular undirected network (i.e., a network in which every node has degree k and there is only one component).

a) Show that the uniform vector $1 = (1, 1, 1, \ldots)$ is an eigenvector of the adjacency matrix with eigenvalue k. In a connected network there is only one eigenvector with all elements positive and hence the eigenvector 1 gives, by definition, the eigenvector centrality of our k -regular network and the centralities are the same for every vertex.

b) Find the Katz centralities of all nodes in the network as a function of k.

c) You should find that, like the eigenvector centralities, the Katz centralities of all nodes are the same. Name a centrality measure that could give different centrality values for different nodes in a regular network.

Solution:

(a) We just need to confirm $A1 = k1$,which is obviously right(every node has degree k).

(b) Let $C_{\text{Katz}}(i)$ denotes Katz centrality of a node i then mathematically:

$$C_{\text{Katz}}(i) = \sum_{k=1}^{\infty} \sum_{j=1}^{n} \alpha^k \left(A^k\right)_{ji}$$

Note that the above definition uses the fact that the element at location (i, j) of A^k reflects the total number of k degree connections between nodes i and j. The value of the attenuation factor α has to be chosen such that it is smaller than the reciprocal of the absolute value of the largest eigenvalue of A. In this case the following expression can be used to calculate Katz centrality:

$$\vec{C}_{\text{Katz}} = \left(\left(I - \alpha A^T\right)^{-1} - I\right) \vec{I}$$

Note that our network is k-regular undirected,which means $A^T = A$.Then

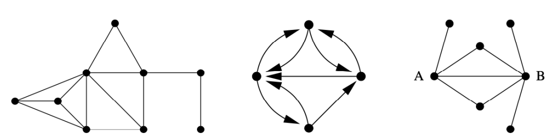
$$\vec{C}_{\text{Katz}} = \frac{k-1}{\alpha + 1 - k} \vec{I}$$

(c) Fine.

5 Measures

5.1 "Networks", Exercise 7.8

Consider these three networks:



a) Find a 3-core in the first network.

b) What is the reciprocity of the second network?

c) What is the cosine similarity of nodes A and B in the third network?

Solution:

(a) Let named nodes from left to right first and from below to top "1,2,$\cdots$,9" .A 3-core is {3,4,6}.

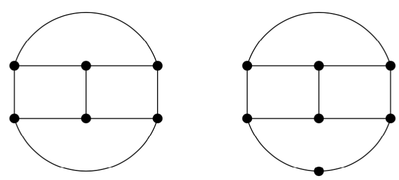
(b) $r = \frac{1}{m}\sum_{ij} A_{ij}A_{ji} = \frac{1}{m}\,\mathrm{Tr}\,\mathbf{A}^2 = \frac{6}{8} = \frac{3}{4}$

(c) $\sigma_{ij} = \cos\theta = \frac{\sum_k A_{ik}A_{kj}}{\sqrt{\sum_k A_{ik}^2}\sqrt{\sum_k A_{jk}^2}} = \frac{2}{\sqrt{20}} = \frac{1}{\sqrt{5}}$

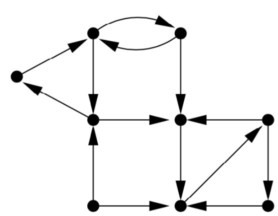
5.2 "Networks", Exercise 7.9

Consider the following networks.

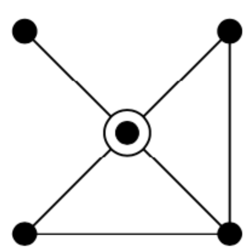
a) Find a 3-core in these two networks or state that there is none:



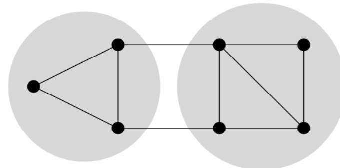
b) Find all the strongly connected components in this graph:



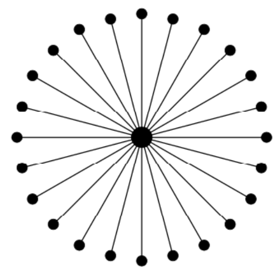
c) Calculate the local clustering coefficient of each node in this network:



d) The two circled groups of nodes in the following network represent people from Mars (on the left) and people from Venus (on the right). What is the modularity Q of the network with respect to planet of origin?



e) A "star graph"consists of a single central node and $n-1$ other nodes connected to it thus:



What is the (unnormalized) betweenness centrality, Eq. (7.24), of the central node as a function of n?

Solution:

(a) Let named nodes from left to right first and from below to top "1,2,$\cdots$" .The first graph is 3-cores.The second graph is 2-cores. It doesn't have 3 degree vertex after removing lesser degree vertices one by one.

(b) Let named nodes from left to right first and from below to top "1,2,$\cdots$" . The all strongly connected connected components in this graph are {1,3,4}, {4,7}, {5,6,9}, {5,8,9}

(c) Let named nodes from left to right first and from below to top "1,2,$\cdots$" .

$$C_i = \frac{(\text{ number of pairs of neighbors of } i \text{ that are connected })}{(\text{ number of pairs of neighbors of } i)}$$

$C_1 = 1 \quad C_2 = \frac{0}{0} \quad C_3 = \frac{1}{3} \quad C_4 = \frac{2}{3} \quad C_5 = 1$

(d) For partitioning a network into c communities,

$$Q = \frac{1}{(2m)} \sum_{vw} \left[A_{vw} - \frac{k_v k_w}{(2m)} \right] \delta(c_v, c_w) = \sum_{i=1}^{c} \left(e_{ii} - a_i^2 \right)$$

where e_{ij} is the fraction of edges with one end vertices in community i and the other in community j :

$$e_{ij} = \sum_{vw} \frac{A_{vw}}{2m} 1_{v \in c_i} 1_{w \in c_j}$$

and a_i is the fraction of ends of edges that are attached to vertices in community i

$$a_i = \frac{k_i}{2m}$$

So $Q = \frac{19}{50}$.

(e) The betweenness centrality of a node v is given by the expression:

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of those paths that pass through v. Obviously, $g(central\ node) = 1$.