Massachusetts Institute of Technology
Departments of Electrical Engineering & Computer Science and Economics
6.207/14.15 Networks
Spring 2018

---

## Problem Set 2
Solutions

---

**Problem 2.1**
Let the directed graph $G$ be a ring: node $i$ is connected to $i+1$ if $i < m$ and $m$ is connected to 1. Compute both eigenvalue centrality and Katz centrality (with $\ \ = \mathbf{1}$). Comment on your result. Do the same for a $k$-regular undirected network (i.e., an undirected network in which every vertex has degree $k$). You may find the steps outlined in Newman Problem 7.1 helpful. Comment on your result.

**Solution:**

(a) Eigenvector Centrality:
For an adjacency matrix A, the eigenvector with the highest eigenvalue represents the eigenvector centrality of each node in A. G is an orthogonal matrix with the highest eigenvalue 1 and corresponding eigenvector

$$\alpha = \begin{bmatrix} 1 \\ . \\ . \\ 1 \end{bmatrix}$$

Here we prove that G can not have an $eigenvalue > 1$

$Ax = \ \ x$
$\Rightarrow \sum a_{ij} x_i \leq x_{max} \forall i \rightarrow if \ \ > 1$
$\Rightarrow \ \ x_i \leq x_{max}$
$\Rightarrow Let\ i = i'\ and\ x_i = x_{max}$
$\Rightarrow \ \ x_{i'} = \ \ x_{i'max} \leq x_{max}$
$\Rightarrow But \ \ > 1 \Rightarrow Contradiction$

Here we show that 1 is an eigenvalue of A because its an orthogonal matrix with unit vectors

$$A * \begin{bmatrix} 1 \\ . \\ . \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ . \\ . \\ 1 \end{bmatrix}$$

This means that all eigenvalues of $G \leq 1$. But since G is orthogonal, all of its eigenvalues lie on the complex plane with modulus 1 which means one of them is real and equal to 1 and $\alpha = \begin{bmatrix} 1 \\ . \\ . \\ 1 \end{bmatrix}$ is the corresponding eigenvector.

Katz Centrality:

$$\begin{matrix} x_i = \alpha x_{i \quad 1} + 1 \\ x_i = x_{i \quad 1} \end{matrix} \quad \rightarrow x_i = \frac{1}{1 \quad \alpha}, \forall i, \alpha \: \epsilon \: (0, 1)$$

Naturally, all nodes have equal centrality (both eigenvector and katz) due to symmetry.

(b) Eigenvector Centrality

The vector $\alpha = \begin{bmatrix} 1 \\ . \\ . \\ 1 \end{bmatrix}$

is an eigenvector of the adjacency matrix A with eigenvalue $k$, because

$$\sum_{j}^{n} A_{ij} = k$$

Using the same contradiction proof from (a), we can see that $max(Ax) = k * x_{max}$ because k edges are contributing towards the vector multiplication of $A_i$ and $x$. Hence, $k$ is the leading eigenvalue with corresponding eigenvector $\alpha = \begin{bmatrix} 1 \\ . \\ . \\ 1 \end{bmatrix}$

Katz Centrality

Setting $x_1 = ... = x_n = x$ and $c = 1$ and $\alpha$ as the attenuating factor, we can write

$$x = \alpha k x + 1 \rightarrow x = \frac{1}{1 \quad \alpha k}$$

Thus, the centralities increase as k increases. All nodes have equal centrality (both eigenvector and katz) due to symmetry.

**Problem 2.2**

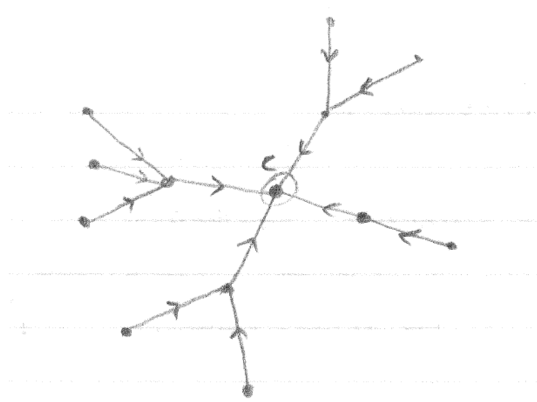
[Problem 7.2 from Newman] Suppose a directed network takes the form of a tree with all edges pointing inward towards a central vertex: (see figure in Newman). What is the PageRank centrality of the central vertex in terms of the single parameter $\alpha$ appearing in the definition of PageRank and the geodesic distances $d_i$ from each vertex $i$ to the central vertex?

**Solution:**

We want to computer the Pagerank centrality of $x_c$ of the central vertex and we assume that C = 1. Notice that 3 nodes have distance 3 from c, 4 nodes have distance 2 from c and 2 nodes have distance 1 from c. Generalizing from the pattern any node that is distance d from c will have its centrality scaled with $\alpha$ d number of times, towards the computation of $x_c$.

Therefore,

$$x_c = \left( \sum_{d=1}^{\infty} |i : d_i = d| \, \alpha^d \right) + 1]$$



**Problem 2.3**

Let the adjacency matrix $A$ of a directed graph be nilpotent (i.e., $A^r = 0$ for some $r$).

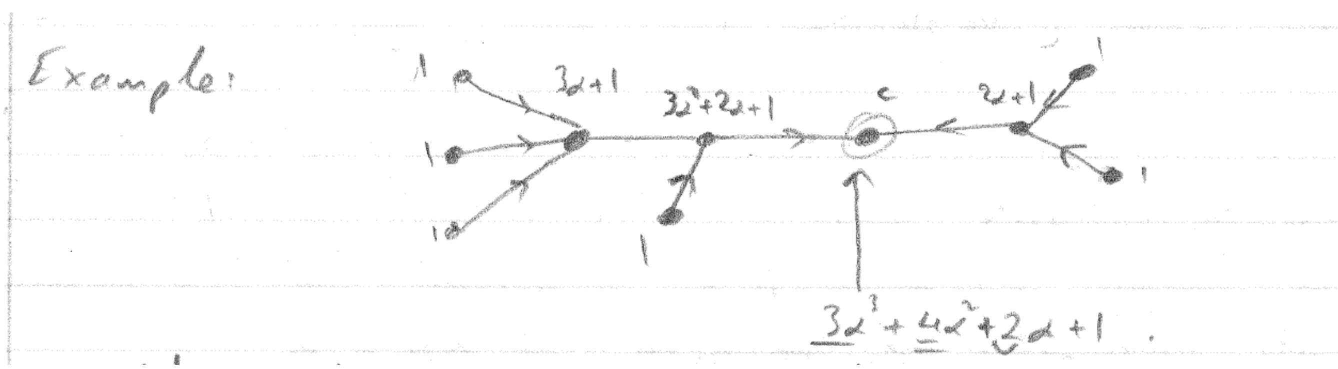
(a) Give an example of such a graph with 5 nodes.

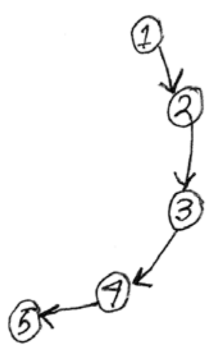
(b) Compute the eigenvalue centrality. Explain your answer.

(c) Compute the Katz centrality with $\quad = \mathbf{1}$. Explain your answer.

**Solution:**

(a) Any directed acyclic graph would do. Here is an example given below




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

$$A^5 = 0$$

(b) Since A is nilpotent, the characteristic polynomial is $t^n = 0$ and all the eigenvalues $_i$ are zero. So it does not make sense to look at the eigenvector centrality, defined as the eigenvector associated with the largest eigenvalue, $_{max}$.

(c) The katz centrality is computed iteratively using

$$x[k+1] = \alpha A x[k] + \beta \; for$$

$$\alpha \in [0, \frac{1}{\lambda_{max}}]$$

. Picking $\alpha = 0.3$, for instance, and $= 1$ we get (after 4 iterations)

$$c_{katz} = [1 \; 1.3 \; 1.39 \; 1.417 \; 1.4251]$$

which provides a sensible ranking if we take incoming edges to indicate importance.

**Problem 2.4**

As flu season is upon us, we wish to have a Markov chain that models the spread of a flu virus. Assume a population of $n$ individuals. At the beginning of each day, each individual is either infected or susceptible (capable of contracting the flu). Suppose that each pair (i, j), $i \neq j$, independently comes into contact with one another during the daytime with probability $p$. Whenever an infected individual comes into contact with a susceptible individual, he/she infects him/her. In addition, assume that overnight, any individual who has been infected for at least 24 hours will recover with probability $0 < q < 1$ and return to being susceptible, independently of everything else (i.e., assume that a newly infected individual will spend at least one restless night battling the flu)

(a) Suppose that there are $m$ infected individuals at daybreak. What is the distribution of the number of new infections by day end?

(b) Draw a Markov chain with as few states as possible to model the spread of the flu for $n = 2$. In epidemiology, this is called an $SIS$ (Susceptible-Infected-Susceptible) model.

(c) Identify all recurrent states.

Due to the nature of the flu virus, individuals almost always develop immunity after contracting the virus. Consequently, we improve our model and assume that individuals become infected at most one time. Thus, we consider individuals as either infected, susceptible, or recovered.

(d) Draw a Markov chain to model the spread of the flu for $n = 2$. In epidemiology, this is called an $SIR$ (Susceptible-Infected-Recovered) model.

(e) Identify all recurrent states.

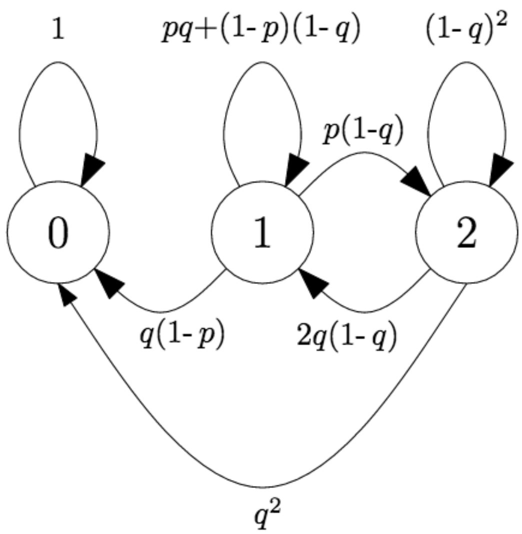
**Solution:**

(a) If $m$ out of $n$ individuals are infected, then there must be $n \quad m$ susceptible individuals. Each one of these individuals will be independently infected over the course of the day with probability $p$ = 1 - (1 - $p$ )$m$.

Thus the number of new infections, $I$, will be a binomial random variable with parameters $n \quad m$ and $p$. That is,

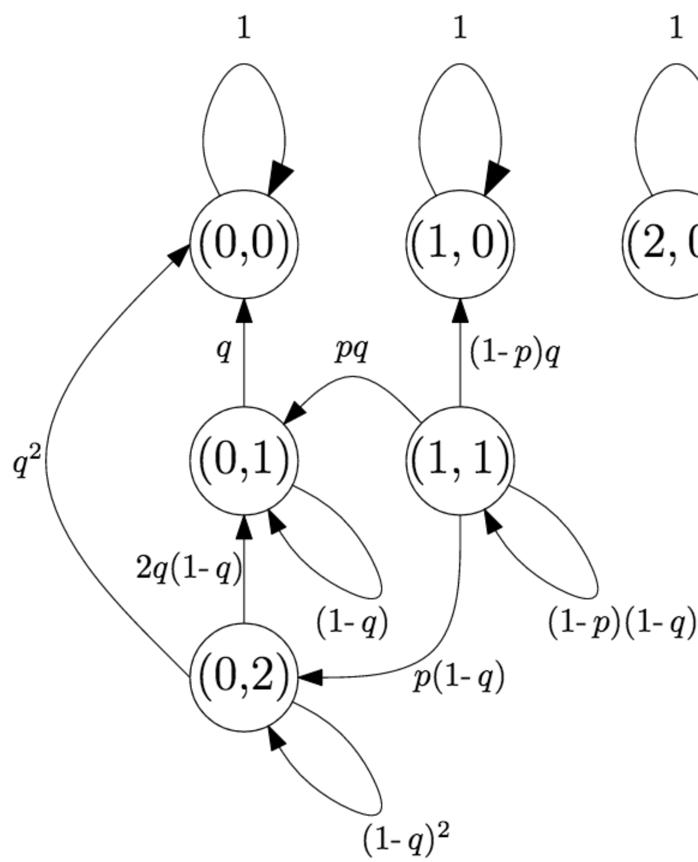
$$pI(k) = \left( \binom{n \quad m}{k} \right) (p^k)(1 \quad p)^{n \ \ m \ \ k} \qquad k = 0, 1, ..., n \quad m$$

(b) Let the state of the SIS model be the number of infected individuals. For n = 2, the corresponding Markov chain is illustrated below.



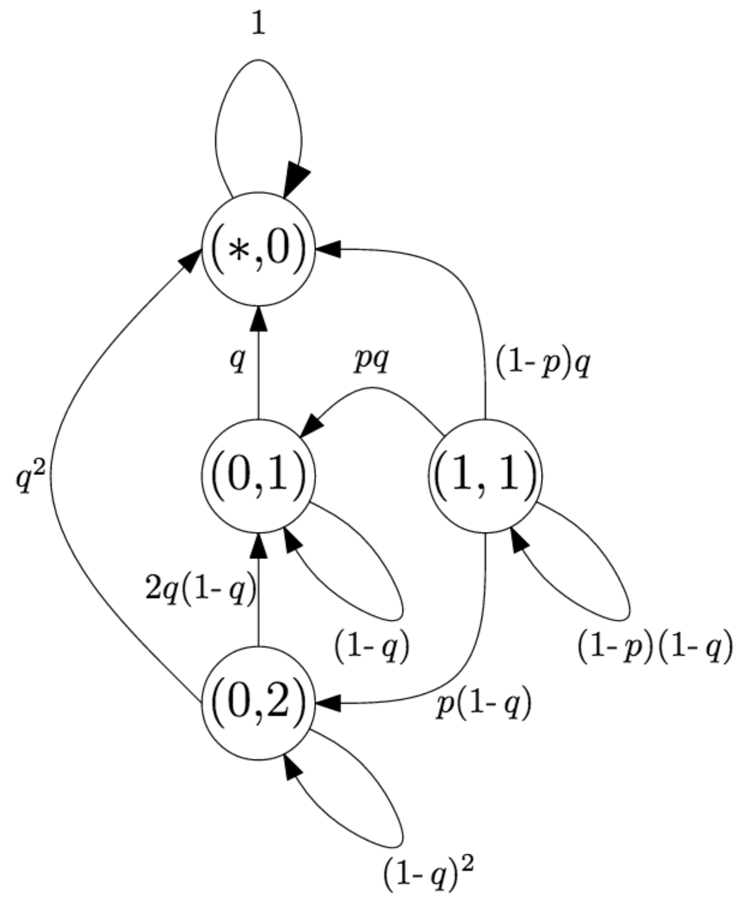
(c) The only recurrent state is the one with 0 infected individuals

(d) Let the state of the SIR model be (S,I), where $S$ is the number of susceptible individuals and $I$ is the number of infected individuals. For n = 2, the corresponding Markov chain is illustrated below



If one did not wish to keep track of the breakdown of susceptible and recovered individuals when no one was infected, the three states free of infections could be consolidated into a single state as illustrated below

(e) Any state where the number of infected individuals equals 0 is a recurrent state. For n = 2,there are either one or three recurrent states, depending on the Markov chain drawn in part(d).



**Problem 2.5**

There are $n$ fish in a lake, some of which are green and the rest blue. Each day, Helen catches 1 fish. She is equally likely to catch any one of the $n$ fish in the lake. She throws back all the fish, but paints each green fish blue before throwing it back in. Let $G_i$ denote the event that there are $i$ green fish left in the lake.

(a) Show how to model this fishing exercise as a Markov chain, where $G_i$ are the states. Explain why your model satisfies the Markov property.

(b) Find the transition probabilities $p_{ij}$

(c) List the transient and recurrent states

**Solution:**

(a) The number of remaining green fish at time $n$ completely determines all the relevant information of the systems entire history (relevant to predicting the future state). Therefore it is immediate that the number of green fish is the state of the system and the process has the Markov property:

$$P(X_{m+1} = j | X_m = i, X_{m\ 1} = i_{m\ 1}, ..., X_1 = i_1) = P(X_{m+1} = j | X_m = i)$$

(b) For j > i clearly $p_{ij} = 0$, since a blue fish will never be painted green. For $0 \le i, j \le k$, we have the following

$$p_{ij} = P(i \to j\ green\ fish\ are\ caught | current\ state = i) = \begin{cases} \frac{n-i}{n} & j = i \\ \frac{i}{n} & j = i - 1 \\ 0 & otherwise \end{cases}$$

(c) The state 0 is an absorbing state since there is a positive probability that the system will enter it, and once it does, it will remain there forever. Therefore the state with 0 green fish is the only recurrent state, and all other states are then transient.