MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

6.207/14.15: Networks
Spring 2018

Midterm

- This is a closed book exam, but two $8\frac{1}{2}'' \times 11''$ sheets of notes (4 sides total) are allowed.
- Calculators are **not** allowed.
- There are **3** problems, each carrying 10pts, on the exam.
- The problems are not necessarily in order of difficulty.
- Record all your solutions in the answer booklet provided. **NOTE: Only the answer booklet is to be handed in—no additional pages will be considered in the grading**. You may want to first work things through on the scratch paper provided and then neatly transfer to the answer sheet the work you would like us to look at. Let us know if you need additional scratch paper.
- A correct answer does not guarantee full credit, and a wrong answer does not guarantee loss of credit. You should clearly but concisely indicate your reasoning and **show all relevant work**. Your grade on each problem will be based on our best assessment of your level of understanding as reflected by what you have written in the answer booklet.
- Please be neat—we can't grade what we can't decipher!

Clustering

Consider the Erdos-Renyi random graph $G_1(n, p)$ with mean degree a.

(a) (*2pt*) Show that in the limit of large n, the expected number of triangles in the network is a constant. **Call $\Delta_{i,j,k}$ as the indicator random variable that denotes a triangle between i, j, k. Then**

$$\begin{aligned} E[\sum_{i<j<k} \Delta_{i,j,k}] &= \binom{n}{3} P(\Delta_{i,j,k} = 1) \\ &= \lim_{n\to\infty} \binom{n}{3} p^3 \\ &= a^3/6 \end{aligned}$$

(b) (*2pt*) Calculate the clustering coefficient C in the limit of large n.
Note: The clustering coefficient is defined as three times the number of triangles divided by the number of connected triplets. A "connected triplet" means three vertices uvw with edges (u, v) and (v, w). The edge (u, w) can be present or not.

Call $\Gamma_{i,j,k}$ as the indicator random variable that denotes a triple between i, j, k.

$$E[\sum_{i<j<k} \Gamma_{i,j,k}] = \binom{n}{3}\binom{3}{2} p^2$$

$$\textbf{Clustering Coefficient} = p$$

$\lim_{n\to\infty} p = 0$.

(c) (*2pt*) Calculate the clustering coefficient C for the Erdos-Renyi random graph $G_2(n, p)$ with $p(n) = a \log(n)/n$. Compare your answer with part (b) in the limit of large n.

As shown before, it is just p, and still 0. But since p is larger in the latter case, there is a higher chance that a triple is actually a triangle.

(d) (*2pt*) Compare the ratio of the diameters of G_1 and G_2 in the limit of large n.

For the value of p in G_1 the graph is almost surely disconnected. As a result it has infinite diameter. On the other hand, G_2 has a diameter that is $\Omega(log(n))$ (and is almost surely connected). Therefore, the ratio $G_2/G_1 = 0$.

(e) (*2pt*) Now, consider a different construction for a random graph model. We take n vertices and go through each distinct trio of three vertices and with independent probability $p = \dfrac{a}{(n-1)(n-2)}$ connect the trio using three edges to form a triangle. Compute the mean vertex degree and clustering coefficient for this network model.

Pick an arbitrary vertex i, let Δ_{ij} be the indicator random variable denoting a link between i, j. Then

$$\begin{aligned} E[\Delta_{ij}] &= 1 - (1-p)^{n-2} \\ &= 1 - e^{(-p(n-2))} \\ &= 1 - e^{(-a/(n-1))} \\ &= a/(n-1) \end{aligned}$$

Then the average degree is $E[\Delta_{ij}](n-1) = a$. Observe that in this case the number of triangles grow linearly as the number of nodes

$$\textbf{Number of Triangle} = \binom{n}{3} p \cong na/6$$

Also, the number of triples of uvw occurs when either uvw is a triangle or there are triangles uvj and vwk and no triangle uvw where $j, k \neq u, v, w$. Call the event of no triangle in uvw but there exists a triple between them as $\Gamma(uvw)$

$$\begin{aligned} \mathbb{P}(\Gamma(uvw)) &= 3(1 - (1-p)^{n-3})^2(1-p) \\ &\cong 3a^2/n^2 \end{aligned}$$

The ratio is roughly $3/(3a+1)$ and points have been awarded to show that the clustering coefficient is a constant (does not decay to 0 as $n \to \infty$).

2. Centrality in Infinite Graphs.

In this problem, you will demonstrate an example that shows that eigenvector centrality can be very sensitive to minimal changes in a network. The problem is broken into different components that finally lead to the conclusion.

Part I

First, consider the infinite ring network as in Figure 1.

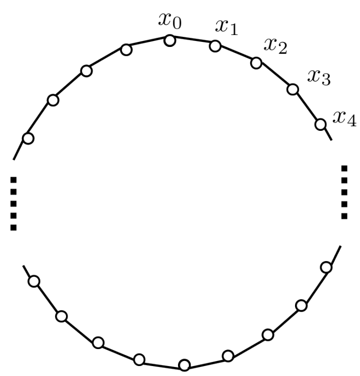

Figure 1:

Assume that x_i is the eigenvector centrality measure of node i.

(a) (*1pt*) Show that the x_i's are computed by finding the largest λ for which there exists a set of x_i for $i = 0, 1, 2, \ldots$ such that:

$$\lambda x_i = x_{i-1} + x_{i+1}, \quad \forall i = 1, 2, \ldots$$

Note that we can always normalize the eigenvector centrality by dividing x_i by x_0 for all i, so that $x_0 = 1$.

This follows by representing the ring network as $Ax = \lambda x$ where λ is the largest eigenvalue and A is the adjacency matrix.

(b) (*2pt*) Show that all the nodes have equal ranking. (*Hint: Show that the is $x_i = 1$ for all i. This should be a straightforward conclusion and can be proved by inspection.*).

One can argue by symmetry or show that $\lambda = 2$. If $\lambda = 2$ all $x_i = 1$ is a solution and hence the eigenvector. λ cannot be larger than 2. If so, and x_t was the largest then

$$\lambda x_t = x_{t-1} + x_{t+1} \leq 2x_t$$

$\lambda \leq 2$ **indeed.**

Part II

Next, as shown in Figure 2, we add an edge between two nodes so that the infinite ring is divided into two symmetric halves. We will examine the eigenvector centrality of this new network. By symmetry, we only need to find the eigenvector centrality

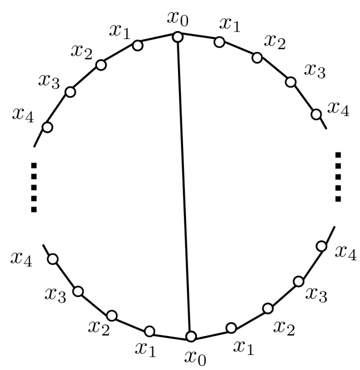

Figure 2:

measures indexed by x_0, x_1, x_2, ... As before, we always normalize it so that $x_0 = 1$.

(c) (*1pt*) Write down the system of equations which characterize the eigenvalue centrality.

$$\lambda x_0 = 2x_1 + x_0$$
$$\lambda x_t = x_{t-1} + x_{t+1} \;\; t > 0$$

(d) (*2pt*) Show that the eigenvector centrality must satisfy:

$$x_0 \geq x_1 \geq x_2 \geq \cdots \geq 0$$

(Hint: start with an initial guess $x_i = 1$ for all i, and try to iteratively compute the eigenvector. You can prove the inequalities by induction.) **Take initial guess as $x_i^t = 1$, where**

$$\lambda x_i^{t+1} = x_{i-1}^t + x_{i+1}^t$$

Clearly for $t = 0$ this is true (after many iterations this will converge to the true centrality). Assume this for $t = k$. Now observe for

$$\lambda x_i^{k+1} - \lambda x_{i+1}^{k+1} = x_{i-1}^k - x_i^k + x_{i+1}^k - x_{i+2}^k$$
$$\geq 0$$

The last inequality follows from inductive hypothesis.

(e) (*1pt*) Show that the largest eigenvalue λ must satisfy:

$$2 \leq \lambda \leq 3$$

(Hint: observe the system of equations you wrote down for (c).) $\lambda \geq 2$ **because if we sum up the equation in (c) we get**

$$\lambda(\sum_{i=0}^{n} x_i) = x_1 + 2(\sum_{i=0}^{n} x_i)$$

Since x_0 is the largest, we also have that

$$\lambda x_0 = 2x_1 + x_0 \leq 3x_0$$

(f) (*3pt*) We have shown in Part (e) that x_n is positive and decreasing in n. Please prove that $\lim_{n\to\infty} x_n = 0$.
(Hint: consider writing the system of equations in Q2 into the form of a linear dynamical system with state $y[n]$ given by:

$$y[n] = \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix},$$

write down the recursive equation $y[n+1] = Ay[n]$ that describes the evolution of the linear dynamics, and think about the equilibrium.) Now you have demonstrated that by adding a single edge one can change the relative centrality measure $\frac{x_0}{x_n}$ drastically.

Observe that

$$y[n+1] = \begin{bmatrix} \lambda & -1 \\ 1 & 0 \end{bmatrix} y[n]$$

for all $n \geq 1$. Then since $x_n \leq x_{n-1}$ we only need to show that there is an eigenvalue of A that is < 1. As it turns out that eigenvalue is

$$\lambda_A = \frac{\lambda - \sqrt{\lambda^2 - 4}}{2}$$

It can be shown that $\lambda_A < 1$ whenever $\lambda > 2$. This implies that $x_n \to 0$.

3. Synchronization.

An oscillator is a simple dynamical system that can be modeled by a first order differential equation. A network of n oscillators can be modeled by a system of differential equations of the form:

$$\frac{d\theta_i}{dt} = \omega + \sum_i A_{ij} g(\theta_i - \theta_j), \quad i = 1, \ldots, n$$

where θ_i represents the phase angle and is the state of the oscillator on vertex i, ω is a constant, and the function $g(x)$ has $g(0) = 0$ and respects the rotational symmetry of the phases, meaning that $g(x + 2\pi) = g(x)$ for all x.

(a) (*2pt*) Characterize all solutions of the form $\theta_i(t) = a_i t + b_i$ to the set of dynamical equations, *i.e.*, find a_i, b_i, $i = 1, \ldots, n$.

$a_i = \omega, b_i = b + 2k_i\pi$

(b) (*3pt*) Consider a small perturbation away from the state $\theta_i = \omega t + \epsilon_i$ and show that the vector $\epsilon = (\epsilon_1, \epsilon_2, \ldots,)$ satisfies

$$\frac{d\epsilon}{dt} = g^{'}(0)\mathbf{L}\epsilon$$

Your solution should specify $\mathbf{L}$ in terms of $[A_{ij}]$, the adjacency matrix of an undirected graph. (Hint: Using the Taylor series approximation of $g(\cdot)$ around x_0, *i.e.*, $g(x) = g(x_0) + (x - x_0)g^{'}(x_0) + \frac{(x-x_0)^2}{2} g^{''}(x_0) + \ldots$. maybe helpful)

Substitute $\theta_i = \omega t + \epsilon_i$. Then the Taylor approximation around 0 can be written as

$$\begin{aligned}\frac{d\epsilon_i}{dt} &= \sum_j g^{'}(0) A_{ij}(\epsilon_i - \epsilon_j) \\ &= g^{'}(0) d_i \epsilon_i - \sum_j A_{ij}\epsilon_j\end{aligned}$$

Now stacking the ϵ_i we get

$$\frac{d\epsilon}{dt} = g^{'}(0)(D - A)\epsilon$$

$(D - A) = L$.

$\mathbf{L}$ is also called the graph Laplacian. We will prove that $\mathbf{L}$ has only non–negative eigenvalues.

(c) (*2pt*) Show that $\mathbf{L} = M^T M$ where M is the incidence matrix, *i.e.*, the rows correspond to the edges and columns correspond to the vertices. Therefore, for every edge $e = (i, j)$ between i, j where $i < j$ we have that

$$M_{ev} = -1 \text{ if } v = i$$
$$M_{ev} = 1 \text{ if } v = j$$
$$M_{ev} = 0 \text{ otherwise}$$

The proof is here.

(d) (*2pt*) Argue that $\mathbf{L}$ is a symmetric matrix and that for any vector $\mathbf{x}$ we have that $\mathbf{x}^T\mathbf{L}\mathbf{x} \geq 0$. Conclude from this that all the eigenvalues of $\mathbf{L}$ are non–negative. (Hint: Use the fact that all eigenvalues, λ, of a symmetric matrix, P, are of the form

$$\frac{v^T P v}{v^T v} = \lambda$$

where v is the corresponding eigenvector.)

(e) (*1pt*) For what values of $g'(0)$ is the system stable to small perturbations around the origin?
(Hint: You can use a Lyapunov argument with quadratic Lyapunov function $V(x) = x^T x$ to examine stability of the linearized system in part (b)).

$g'(0) < 0$ **ensures system is stable.**