Math 168 – Intro to Networks

University of California, Los Angeles

Duc Vu

Spring 2022

This is math 168 – Introduction to Networks taught by Professor Chodrow. We meet weekly on MWF from 10:00 am to 10:50 am for lecture. The required textbook for the class is *Networks* 2^{nd} by *Newman*. Other course notes can be found at my blog site. Please let me know through my email if you spot any typos in the note.

Contents

List of Theorems

List of Definitions

§1 Lec 1: Mar 28, 2022

§1.1 Introduction

Some introduction and logistics stuffs of the class. Nothing mathy is discussed in this lecture.

§2 Lec 2: Mar 30, 2022

§2.1 Networks and Matrices

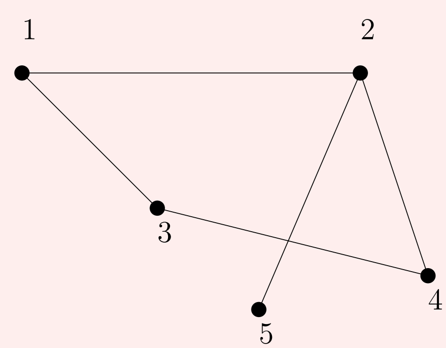
Definition 2.1 (Graph) — A (simple, undirected) graph is $G = (N, E)$, a node set N and an edge set $E \subseteq N \times N$ s.t. $i \neq j \forall (i, j) \in E$.

Definition 2.2 (Adjacency Matrix) — The adjacency matrix $\mathbf{A}$ of a graph $G = (N, E)$ is a matrix in $\mathbb{R}^{n \times n}$ where $n = |N|$ with entries

$$a_{ij} = \begin{cases} 1, & \text{if } (i, j) \in E \\ 0, & \text{otherwise} \end{cases}$$

Example 2.3

Consider the following graph



The adjacency matrix is

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

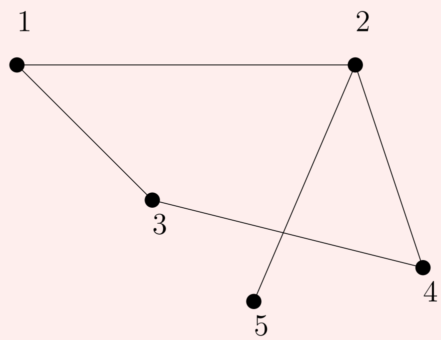
Definition 2.4 (Walk) — A walk in $G = (N, E)$ is a sequence of edges $(i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k)$ where

$$j_1 = i_2, j_2 = i_3, \ldots, j_{k-1} = i_k$$

This is a walk of length k (number of edges) from i_1 to j_k.

Example 2.5

Consider the above example



Walk $3 \to 2$ of length

- $1 : \emptyset$
- $2 : (3,1),(1,2);(3,4),(4,2)$
- $3 : \emptyset$
- $4 : (3,1),(1,2),(2,1),(1,2)$

Fact 2.1. The $ijth$ entry of $\mathbf{A}$ counts the number of walks of length 1 from node i to j.

Conjecture 2.1. The $ijth$ entry of $\mathbf{A}^k$ counts the number of walks of length k from i to j.

Proof. Suppose inductively that $W(k) \triangleq \mathbf{A}^k$ has entries $w_{ij}(k)$ counting k-walks from $i \to j$. Consider $W(k+1) = W(k)A$. Its entries are

$$w_{ij}(k+1) = \sum_{l \in N} w_{il}(k) a_{lj}$$

$\square$

§3 Lec 3: Apr 1, 2022

§3.1 Measures and Metrics

A walk of length 2, $i \leftrightarrow i$, is the number of edges attached to node $i \triangleq$ degree of node i, k_i

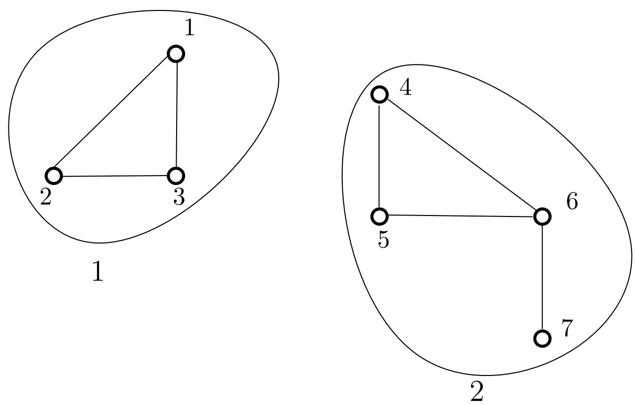
$$k_i = \sum_{j \in N} a_{ij} = \sum_{j \in N} a_{ji} = ii\text{th entry of } \mathbf{A}^2$$

Definition 3.1 (Degree) — The degree k_i of a node i is the number of edges attached to it

$$k_i = |\{j : (i,j) \in E\}|$$

Definition 3.2 (Path-connected) — Nodes i and j are path-connected if $\exists$ a walk $i \leftrightarrow j$ of any length. The connected component of i is the set of notes to which i is path-connected.
G is connected if it has 1 connected component.

Consider a disconnected graph G



Then, the adjacency graph is

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & 0 \\ 0 & \mathbf{A}_2 \end{bmatrix} \text{ up to permutations of node labels}$$

Question 3.1. How big is a graph?

- Number of nodes
- Number of edges
- Diameter

Definition 3.3 (Geodesic Path) — Geodesic (shortest) path between $i \leftrightarrow j$ is a walk s.t. no walk has shorter length.

Definition 3.4 (Diameter) — Diameter of G is

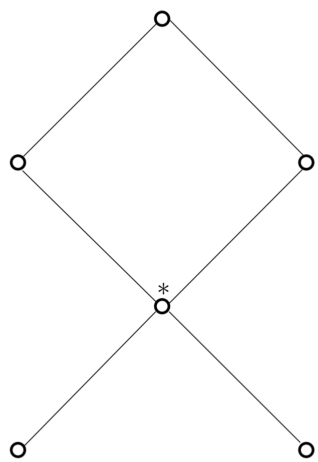
$$\max_{ij} \text{geodesic distance}(i, j)$$

which is undefined if i and j are not connected.

Example 3.5

"6 degrees of separation": in social networks, the diameter is usually about 6.

Node Importance:



- * has highest degree
- If * were removed, graph would be disconnected
- Short average distance to other nodes
- Betweeness: # of geodesic paths passing through $i \in N$.

§4 Dis 1: Mar 29, 2022

§4.1 Review of Linear Algebra

Networks can be represented as ordinary matrices and the related graph Laplacian so linear algebra will be very important for us.

Example 4.1
Spectral graph theory techniques.

Definition 4.2 (Matrix Kernel) — The kernel of a matrix $\mathbf{A}$ denoted $\ker(\mathbf{A})$ is the set of vectors $\vec{x}$ s.t. $\mathbf{A}\vec{x} = \mathbf{0}$.

Definition 4.3 (Range) — The range of a matrix $\mathbf{A}$ denoted $\operatorname{im}(\mathbf{A})$ is the set of vectors $\vec{x}$ s.t. there exists a vector $\vec{y}$ s.t. $\mathbf{A}\vec{y} = \vec{x}$.

Definition 4.4 (Eigenvalue) — λ is an eigenvalue of $n \times n$ matrix $\mathbf{A}$ with right-eigenvector (usually just called eigenvector if there's no chance of confusion) $\vec{x} \neq \mathbf{0}$ if $\mathbf{A}\vec{x} = \lambda\vec{x}$.
It has left-eigenvector $\vec{y} \neq \mathbf{0}$ if $\vec{y}^\top \mathbf{A} = \lambda \vec{y}^\top$ or equivalently $\mathbf{A}^\top \vec{y} = \lambda \vec{y}$.

<u>*Note*</u>: If $\vec{x}$ is a left or right eigenvector with eigenvalue λ, then so is $\alpha\vec{x}$ for any scalar $\alpha \neq 0$.

Definition 4.5 (Spectral Radius) — The spectral radius for an $n \times n$ matrix $\mathbf{A}$ denoted $\rho(\mathbf{A})$ is the maximum magnitude of its eigenvalues

$$\rho(\mathbf{A}) = \max\left\{|\lambda| : \lambda \text{ is an eigenvalue of } \mathbf{A}\right\}$$

Definition 4.6 (Span) — The span of a set of vectors is the set of all linear combinations of those vectors

$$\operatorname{span}\{v_1, v_2, v_3\} = \{a_1 v_1 + a_2 v_2 + a_3 v_3\}$$

where a_1, a_2, a_3 are scalars.

Question 4.1. How do we calculate eigenvalues/eigenvectors?

1. Calculate the characteristics polynomial

$$p_{\mathbf{A}}(\lambda) = \det\left(\mathbf{A} - \lambda \mathbf{I}_n\right)$$

2. The eigenvalues are the roots $\lambda_1, \ldots, \lambda_n$ of the characteristic polynomial
3. Calculate the eigenvectors, i.e., we want to solve for $\vec{x}$ s.t. $(\mathbf{A} - \lambda\mathbf{I})\vec{x} = \mathbf{0}$.

Theorem 4.7 (Spectral for Real Matrices)
Let $\mathbf{A}$ be a $n \times n$ symmetric matrix. Then $\mathbf{A}$ has real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (not necessarily distinct) and there exists corresponding eigenvectors $\{v_1, v_2, \ldots, v_n\}$ that form an orthonormal basis for $\mathbb{R}^n$.

Furthermore, $\mathbf{A}$ is diagonalizable here, i.e., $\mathbf{QDQ}^\top = \mathbf{A}$ where $\mathbf{D}$ is the diagonal matrix of eigenvalues and $\mathbf{Q}$ is its corresponding matrix of eigenvectors. This theorem also applies to the adjacency matrix of any undirected network since it will be symmetric.

Theorem 4.8 (Perron-Frobenius Version 1)
If $\mathbf{A}$ is a $n \times n$ real matrix with all non-negative entries, then $\mathbf{A}$ has a non-negative leading eigenvalue κ_1 with

$$\kappa_1 = \rho(\mathbf{A}) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } \mathbf{A}\}$$

So for this leading eigenvalue κ_1, all other eigenvalues λ of $\mathbf{A}$ satisfy $|\lambda| \le \kappa_1$.

In addition there exists left and right eigenvectors of κ_1 with all non-negative entries.

Proof. Hw 0 # 3a. □

Theorem 4.9 (Perron-Frobenius Version 2)
If $\mathbf{A}$ is the adjacency matrix (could be weighted) for a strongly-connected directed network (or any connected undirected network, then we call $\mathbf{A}$ irreducible). If $\mathbf{A}$ is a real $n \times n$ matrix that either

a) has all strictly positive entries or

b) is irreducible

then

i) The leading eigenvalue κ_1 with $\kappa_1 \ge |\lambda|$ is strictly positive ($\kappa_1 > 0$) and has one-dimensional eigenspace $\{v : \mathbf{A}v = \kappa_1 v\}$

ii) κ_1 has a leading left and right eigenvectors associated with it that have all strictly positive entries.

iii) The only non-negative eigenvectors of $\mathbf{A}$ are multiples of the leading eigenvectors. The eigenvectors for all other eigenvalues of $\mathbf{A}$ have at least one negative entry.

§4.2 Review of Probability

Definition 4.10 (Joint Probability) — The probability of events A and B occurring is denoted $P(A \cap B)$ or $P(A, B)$.

Definition 4.11 (Conditional Probability) — The probability of $\mathbf{A}$ given that $\mathbf{B}$ occurred is denoted $P(A|B) = \frac{P(A\cap B)}{P(B)}$.

Theorem 4.12
The probability of A or B (or both) occurring is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Law of Total Probability

Let X denote the sample space of all possible events. Let $\{B_i\}_{i=1}^{\infty}$ be a countable (it could be finite) partition of X, so $X = \bigcup_{i=1}^{\infty} B_i$, then

$$P(A) = \sum_{i=1}^{\infty} P(A \cap B_i) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i)$$

Bayes' Rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Definition 4.13 (Independent Events) — Events A and B are independent if an only if one of the following equivalent statements hold

i) $P(A \cap B) = P(A)P(B)$

ii) $P(A|B) = P(A)$

iii) $P(B|A) = P(B)$

Definition 4.14 (Expectation) — We have two cases

- Discrete: Let X be a discrete random variable with possible events $\{x_i\}_{i=1}^{\infty}$ and probability mass function (PMF) p_x. The expected value of $g(X)$ is

$$E[g(X)] = \langle g(X) \rangle = \sum_{i=1}^{\infty} g(x_i)p_x(x_i)$$

- Continuous: Let X be a continuous random variable with probability density function (PDF) f_X. Then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_x(x)dx$$

Note: Expectation is linear, i.e.,

$$E[X+Y] = E[X] + E[Y]$$
$$E[\alpha X] = \alpha E[X]$$

Definition 4.15 — A random variable X has

1. Mean $\mu_x = E[X]$
2. Variance $\text{var}(X) = \sigma_x^2 = E\left[(X - E[X])^2\right] = E\left[X^2\right] - E[X]^2$
3. n^{th} moment (about zero) $E[X^n]$

§4.3 Some Useful Inequalities

There are some useful inequalities to keep in mind

1. Cauchy-Schwartz for expectation
$$|E[XY]| \leq \sqrt{E[X^2]\,E[Y^2]}$$
2. Markov's Inequality: If X is a non-negative random variable and $a > 0$, then
$$P(X \geq a) \leq \frac{E[X]}{a}$$
3. Chebyshev's Inequality: Let X be a random variable with mean μ and variance σ^2. Then for $a > 0$,
$$P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}$$
4. Jensens's Inequality: If f is a convex function, then
$$E[f(X)] \geq f(E[X])$$
If f is a concave function, then the reverse inequality holds.

Note: $f : \Omega \to \mathbb{R}$ is convex if for all $0 \leq t \leq 1$ and $x_1, x_2 \in \Omega$,
$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$
f is concave if $-f$ is convex.