HEART: Statistics and Data Science With Networks

Joshua Agterberg

Johns Hopkins University

Fall 2021

Outline

Outline

1. Basic Graph Theory
   - Directed and Undirected Graphs
   - Adjacency and Lacplacian Matrices
   - Basic Statistical Properties of Graphs

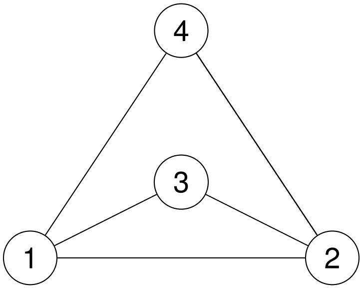
Graphs

A graph (or network) is a set of vertices V and edges E

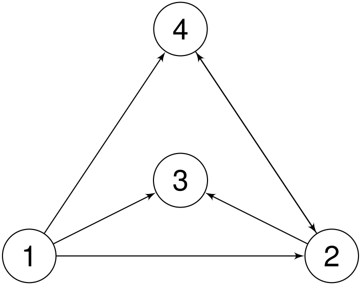
Graphs

A graph (or network) is a set of vertices V and edges E
Can be *undirected* or *directed*

Undirected Graph



Directed Graph



Examples

Examples of *Undirected* Graphs:

Examples

Examples of *Undirected* Graphs:

- Facebook
- Friendships (typically)
- Coauthors on papers

Examples

Examples of *Undirected* Graphs:

- Facebook
- Friendships (typically)
- Coauthors on papers

Examples of *Directed* Graphs:

Examples

Examples of *Undirected* Graphs:

- Facebook
- Friendships (typically)
- Coauthors on papers

Examples of *Directed* Graphs:

- Twitter
- Instagram
- COVID transmission graph

Degrees: Undirected Graphs

- For undirected graphs, the degree for a vertex is just the number of edges for that vertex.

Degrees: Undirected Graphs

- For undirected graphs, the degree for a vertex is just the number of edges for that vertex.
- E.g. You are friends with 25 people, so your degree is 25, but Joshua is only friends with his mom, so his degree is 1.

Degrees: Undirected Graphs

- For undirected graphs, the degree for a vertex is just the number of edges for that vertex.
- E.g. You are friends with 25 people, so your degree is 25, but Joshua is only friends with his mom, so his degree is 1.
- *Self-Loops*: If an edge is allowed from one vertex to itself, then that is called a self-loop

Degrees: Undirected Graphs

- For undirected graphs, the degree for a vertex is just the number of edges for that vertex.
- E.g. You are friends with 25 people, so your degree is 25, but Joshua is only friends with his mom, so his degree is 1.
- *Self-Loops*: If an edge is allowed from one vertex to itself, then that is called a self-loop
- E.g. Joshua is friends with himself and his mom, so his degree is 2 allowing self-loops.

Degrees: Undirected Graphs

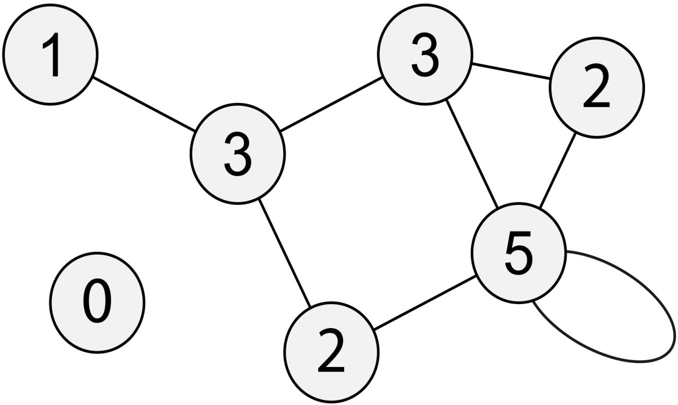


Figure: : Source

Degrees: Directed Graphs

- *In* Degree:
- *Out* Degree:

Degrees: Directed Graphs

- *In* Degree: Counts the number of *incoming* edges into a vertex
- *Out* Degree: Counts the number of *outgoing* edges from a vertex

Degrees: Directed Graphs

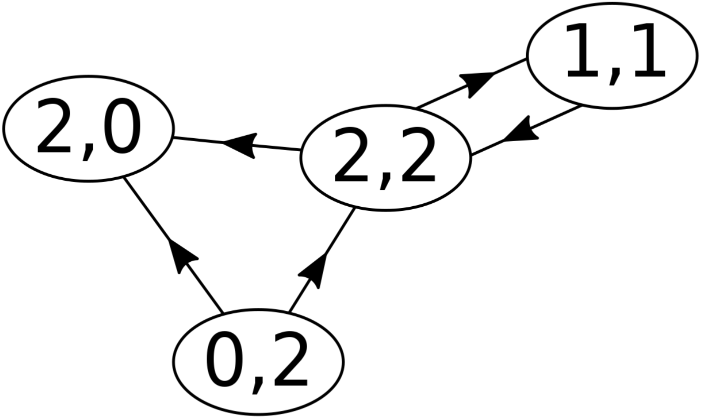


Figure: : source

Adjacency Matrix

- A matrix is a square ordering of numbers (learn about them in linear algebra)

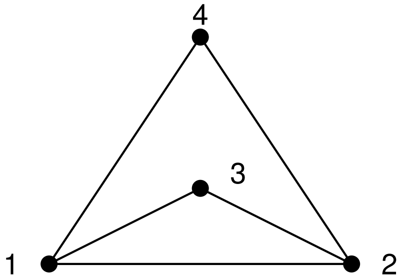
Adjacency Matrix

- A matrix is a square ordering of numbers (learn about them in linear algebra)
- Adjacency matrix encodes graph information – entry (i, j) is one if there is an edge between vertex i and vertex j

Adjacency Matrix

- A matrix is a square ordering of numbers (learn about them in linear algebra)
- Adjacency matrix encodes graph information – entry (i, j) is one if there is an edge between vertex i and vertex j

Graph:



Adjacency Matrix:

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

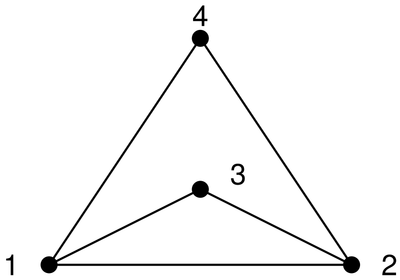
Laplacian Matrix

- Combinatorial Laplacian: $D - A$, where D_{ii} is the degree of the i'th vertex
- Normalized Laplacian: $I - D^{-1/2}AD^{-1/2}$ (complicated)

Laplacian Matrix

- Combinatorial Laplacian: $D - A$, where D_{ii} is the degree of the i'th vertex
- Normalized Laplacian: $I - D^{-1/2} A D^{-1/2}$ (complicated)

Graph:



Combinatorial Laplacian Matrix:

$$\begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{pmatrix}$$

Some simple “first-pass” ways to look a a network

- Edge density:
  - $\frac{2|E|}{n(n-1)}$ for undirected graphs
  - $\frac{|E|}{n(n-1)}$ for directed graphs

Some simple “first-pass” ways to look a a network

- Edge density:
  - $\frac{2|E|}{n(n-1)}$ for undirected graphs
  - $\frac{|E|}{n(n-1)}$ for directed graphs
- Average degree:
  - Let $d_1, \ldots, d_n$ be the degrees of each vertex
  - Average degree is just $\frac{1}{n}\sum_{i=1}^{n} d_i$

Some simple “first-pass” ways to look a a network

- Edge density:
  - $\frac{2|E|}{n(n-1)}$ for undirected graphs
  - $\frac{|E|}{n(n-1)}$ for directed graphs
- Average degree:
  - Let $d_1, \ldots, d_n$ be the degrees of each vertex
  - Average degree is just $\frac{1}{n}\sum_{i=1}^{n} d_i$
- Going further: degree distributions
  - Look at distribution of degrees (e.g. histogram)

Some simple “first-pass” ways to look a a network

- Edge density:
  - $\frac{2|E|}{n(n-1)}$ for undirected graphs
  - $\frac{|E|}{n(n-1)}$ for directed graphs
- Average degree:
  - Let $d_1, \ldots, d_n$ be the degrees of each vertex
  - Average degree is just $\frac{1}{n}\sum_{i=1}^{n} d_i$
- Going further: degree distributions
  - Look at distribution of degrees (e.g. histogram)

Karate Club Python Example

Some simple “first-pass” ways to look a a network

- Edge density:
  - $\frac{2|E|}{n(n-1)}$ for undirected graphs
  - $\frac{|E|}{n(n-1)}$ for directed graphs
- Average degree:
  - Let $d_1, \ldots, d_n$ be the degrees of each vertex
  - Average degree is just $\frac{1}{n}\sum_{i=1}^{n} d_i$
- Going further: degree distributions
  - Look at distribution of degrees (e.g. histogram)

Karate Club Python Example
Random Graph R Example