MATH 7: Analyzing Network Data
Winter 2026

Lecturer: Professor Peter Mucha

Notes by: Farhan Sadeek

Last updated: February 9, 2026

Contents

1 January 5, 2026 **3**

2 January 7, 2026 **3**

3 January 9, 2026 **5**

4 January 12, 2026 **6**

1 In the thick of it . . . 6

2 When you smile, the world smiles with you . . . 7

5 January 14, 2026 **7**

6 January 16, 2026 **7**

7 January 19, 2026 **7**

8 January 21, 2026 **7**

9 January 23, 2026 **7**

10 January 26, 2026 **7**

11 January 28, 2026 **7**

12 January 30, 2026 **7**

3 Erdos-Renyi random graph . . . 7

4 Configuration model . . . 9

5 Watts-Strogatz small world graph . . . 9

6 Barabasi-Albert scale-free graph . . . 9

13 February 2, 2026 **9**

14 February 4, 2026 **9**

15 February 6, 2026 **9**

16 February 9, 2026 **9**

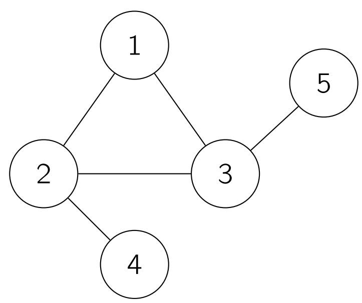
1 January 5, 2026

Today was the first day of classes, but I missed it because I was not there in Hanover, but I was able to talk to the professor and he said that the course will be a mix of lecture and discussion. We will have two textbooks one would be about the reading and the other would be about writing.

2 January 7, 2026

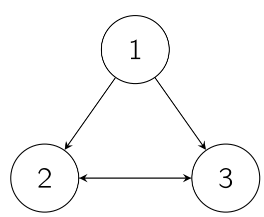
Today we are learning more about the Chapter 2 in the programming textbook, and professor was showing us how to code these graphs in R. But we started of the class with learning about graph. First of all, there are a few terms that we need to know about a graph. A **node** or **vertex** is a point in the graph, and **edge** is a line connecting two of these nodes. The representation of these nodes can be done with an **edge list**, **adjacency list** or **adjacency matrix**.

In this course, we will be using the `igraph` package in R to analyze graphs. There are two types of graphs that we will be learning about. The first one is an **undirected graph**, where the edges have no direction. The second one is a **directed graph**, where the edges have a direction.



In this diagram, each line is an **edge** between two vertices, and edges have no direction.

Now, consider a directed version where the arrows indicate the direction of travel:



Here, each edge has a direction, indicated by the arrow. For example, there is an edge from 1 to 2, and an edge from 2 to 3, etc.

We can also showcase a graph with an adjacency matrix, which is a square matrix where the rows and columns represent the vertices of the graph. The entries of the matrix are the edges between the vertices. For example, we can represent the graph above with the following adjacency matrix:

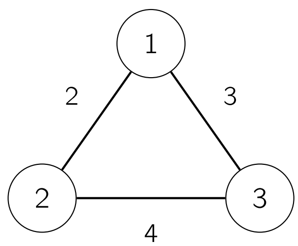
	1	2	3
1	0	1	1
2	1	0	1
3	1	1	0

Here, 0 represents no edge between the vertices, and 1 represents an edge between the vertices. Here, if we have an undirected graph, the number of 1s in a matrix is twice the number of edges in the graph, because if 1 is connected to 2, then 2 is also connected to 1. That means that the number of 1s in a matrix is twice the number of edges in the graph. Also this matrix is a symmetric matrix, because the matrix is the same as its transpose. For example, Facebook is under the hood a giant matrix with 2 billion rows and columns that represent the users and the edges between them.

We can also add weights to the edges of the graph. For example, we can represent the graph above with the following adjacency matrix:

	1	2	3
1	0	2	3
2	2	0	4
3	3	4	0

Here is a visualization of the weighted graph described by the adjacency matrix above:

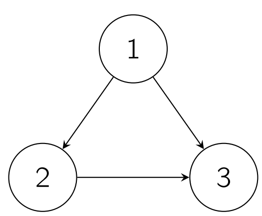


In this diagram, each edge label denotes the **weight** of that edge, corresponding to the entries in the weighted adjacency matrix above.

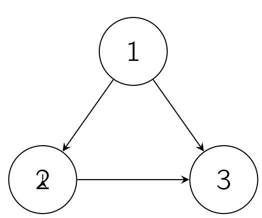
Definition (Simple Graph)

A **simple graph** is a graph when there are no self loops and no multiple edges between two of the same vertices.

So for example, the graph below is a simple directed graph.



and this graph is not a simple graph because there is a self loop.

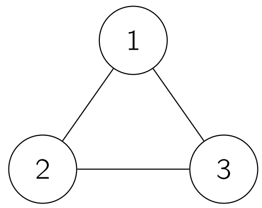


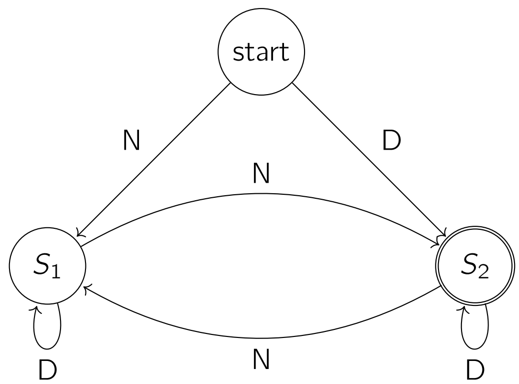
3 January 9, 2026

Today we are working on Chapter 3 of the book. Earlier this week, we talked about simple graphs, and we will define it today.

Definition (Simple Graph)
A **simple graph** is a graph when there are no self loops and no multiple edges between two of the same vertices. A **self loop** is an edge that connects a vertex to itself.

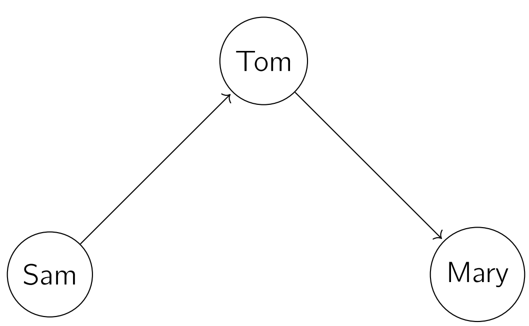
So for example, the graph below is a simple graph.





Definition (Degree of a Vertex)
The **degree** of a vertex is the number of edges connected to the vertex. For example, the degree of the vertex 1 in the graph above is 2, because it has 2 edges connected to it. So the degree of the vertex 1 is 2, and the degree of the vertex 2 is 3, and the degree of the vertex 3 is 2.

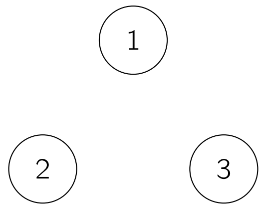
Definition (Degree Sequence)
The **degree sequence** of a graph is the sequence of the degrees of the vertices of the graph. For example, the degree sequence of the graph above is (2, 3, 2).

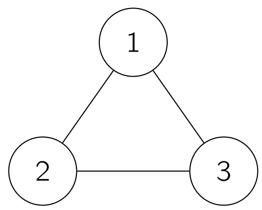


Definition (Bipartite Graph)

A **bipartite graph** is a graph whose vertices can be divided into two disjoint sets such that no two vertices within the same set are adjacent.

So for example, the graph below is a bipartite graph.





4 January 12, 2026

Today we did most of the discussion about the Chatper 1&2 of the reading book. First of all, we went to discuss about the Chapter 1.

1 In the thick of it

There are a few discussion and the first one was about the organ donors, this one is the example that I missed out on. But the second one was about the Norwegian soldier which say that the oldest child usually have the highest IQ. There are some flaws in the article and the interpretation was incredibly flawed. The study is good but the translation was not good. The third article we discussed are "The size of the network of a college graduate is twice the size of someone who didn't graduate from high school." The R value was an interesting topic of discussion, although it was not very high (0.35), since it's categorical and social science based. After that we discussed the Milgram's experiment[1] with the six degree connections and then the professor pulled up the article, and then discussed the last paper which was an extension of the Milgram's article

[1] Travers, J., & Milgram, S. (1969). An experimental study of the small world problem. *Sociometry*, 32(4), 425-443.

2 When you smile, the world smiles with you

5 January 14, 2026

6 January 16, 2026

7 January 19, 2026

8 January 21, 2026

9 January 23, 2026

10 January 26, 2026

11 January 28, 2026

12 January 30, 2026

Today we talked about the Chapter 5 of the KC book. First we talked about some graph structures such as

- Erdos-Renyi random graph
- Configuration model
- Watts-Strogatz small world graph
- Barabasi-Albert scale-free graph

3 Erdos-Renyi random graph

The Erdos-Renyi random graph[2] $G(n, p)$ is a graph that is generated by the following process:

- Start with n vertices.
- For each pair of vertices, add an edge with probability p.

[2] https://snap.stanford.edu/class/cs224w-readings/erdos60random.pdf

Definition (Connection Per Node)

The **Connection Per Node** is the average number of edges connected to a vertex in a graph. For example, the connection per node of the Erdos-Renyi random graph is $(n-1)p$, because for each pair of distinct vertices, there is a probability p of an edge between them. Now, since we are working with undirected graphs, it is the mean degree of the graph which is $2m/n$ where m is the number of edges and n is the number of vertices. Since we have the degree of nodes, are $< k >$ is the average degree of the graph.

Now, we want to know about the $< k >$ in the graph model $G(n, p)$. We can use the following formula because in the adjacency matrix (with no self-loops), we have

$$A_{ij} = \begin{cases} 1 & \text{with probability } p \text{ when } i \neq j \\ 0 & \text{otherwise (including } i = j\text{; no self-loops)} \end{cases}$$

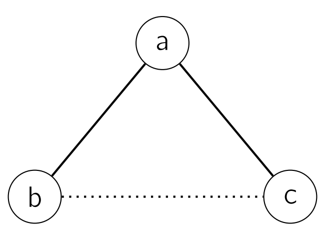
where A_{ij} is the element of the adjacency matrix at the i-th row and j-th column.

$$\begin{aligned} < k > &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} = \frac{1}{n} \sum_{i \neq j} A_{ij} \\ &= \frac{1}{n} \sum_{i \neq j} (1 \cdot p + 0 \cdot (1-p)) \\ &= \frac{1}{n} \cdot n(n-1) \cdot p \\ &= (n-1)p \end{aligned}$$

So, we have $< k >= (n-1)p$. Now since we have a binomial distribution the variance of this is $np(1-p)$, and the standard deviation is $\sqrt{np(1-p)}$.

Note

When n is large and p is small, $1 - p \approx 1$, so we can approximate the variance as np, so the standard deviation is $\sqrt{np} = \sqrt{< k >}$.



Since the edges are i.i.d, the probability of a vertex having an edge is p. But in reality that is almost impossible because your friend of friends are also your friends.

Fact

On average, your friend has more friends than you do.

4 Configuration model

The Configuration model is a graph that is generated by the following process:

- Start with n vertices.
- For each vertex, assign it d degrees.

5 Watts-Strogatz small world graph

The Watts-Strogatz small world graph[3] is a model designed to capture the "small world" phenomenon often seen in real networks, where most nodes can be reached from every other by a small number of steps, and clustering is high. The graph is generated by the following process:

- Start with n vertices arranged in a ring (a circle).
- Connect each vertex to its k nearest neighbors on each side, so that every vertex has $2k$ edges to begin with.
- For each edge in the ring, with probability p, "rewire" one end of the edge to another randomly chosen vertex (avoiding self-loops and duplicate edges).

This process interpolates between an ordered lattice ($p = 0$) and a random graph ($p = 1$), allowing us to study properties such as average path length and clustering coefficient as p varies.

6 Barabasi-Albert scale-free graph

The Barabasi-Albert scale-free graph[4] is a graph that is generated by the following process:

- Start with m vertices.
- For each new vertex, add m edges to it.

After that we talked about the **Giant Component** in a graph. The *Giant Component* is the largest connected component in a graph. It is the component that contains the most vertices.

13 February 2, 2026

14 February 4, 2026

15 February 6, 2026

16 February 9, 2026

[3] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440-442.
[4] Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509-512.