Lecture 5: Graphs

Rajat Mittal

IIT Kanpur

Combinatorial graphs provide a natural way to model connections between different objects. They are very useful in depicting communication networks, social networks and many other kind of networks. For the purpose of this course, graphs will mean combinatorial graphs as opposed to graphs of functions etc. which you might have learnt previously.

Graphs have become such an important tool that a complete field, *Graph Theory*, is devoted to learning about the properties of graphs. In the next few lectures we will learn learn about graphs, associated concepts and and various properties of graphs.

1 Definition

A *graph*, $G = (V, E)$, is described by a set of vertices V and edges E which represent the connection between the vertices ($E \subseteq V \times V$). If (u, v) is an element in E then we say that vertex u is connected to vertex v. We also say that u is adjacent to v. The vertex set of a graph G is represented by $V(G)$ and the edge set by $E(G)$.

The graph is called *simple* if,

- there are no loops, i.e., no vertex is connected to itself,
- there is only one edge between any two vertices,
- if there is no direction assigned to the edges.

If there is a possibility of multiple edges between vertices then it is called a *multigraph*. If the edges are assigned directions, i.e., (u, v) is ordered then it is called a *directed* graph as opposed to an *undirected* graph.

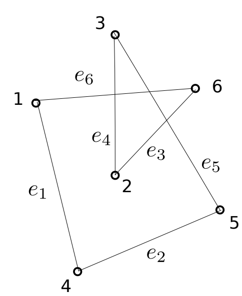

Fig. 1. A simple graph

Graphs are used extensively in modelling social networks. The members of the social group are represented by vertices and their relationship is modelled by the edges between them. If the graph depicts the friendships between people, like Facebook, it is an undirected graph. If it has a relation which has direction (say a is elder to b in a family), then this will be a directed graph.

Another example of a directed graph is how different processes are run on a computer. Suppose a process a can be run only if process b is completed. This is called the *precedence graph* and a directed edge (u, v) shows that u should run before v.

We will study undirected graphs in this course. Note that many of the concepts described below can be studied for the directed version too.

A graph $H = (V, E')$ is called the *subgraph* of a graph $G = (V, E)$ if the edge set E' is a subset of the edges E in G. In other words, some edges of G are not present in H, but every edge of H is present in G.

The *degree* of a vertex v in a graph G is the number of edges from v in G. The degree of a vertex can be greater than $|V| - 1$. By double counting, we can easily prove the following theorem.

Theorem 1. *The sum of the degrees of all the vertices in a graph is equal to twice the number of edges.*

Proof. Left as an exercise. □

One of the major problems in theoretical computer science is to figure out if two given graphs are isomorphic. Two graphs, G and H, are called isomorphic if there exist a bijection $\pi : V(G) \to V(H)$, s.t., $(u, v) \in E(G)$ if and only if $(\pi(u), \pi(v)) \in E(H)$.

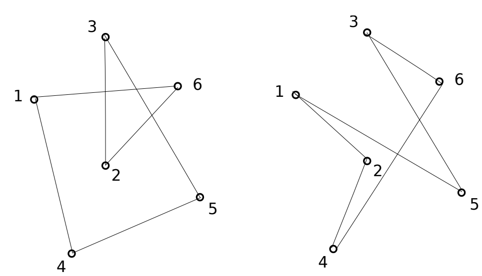

Fig. 2. Are these two graphs isomorphic?

Exercise 1. Construct all possible non-isomorphic connected graphs on four vertices with at most 4 edges.

Exercise 2. Construct two graphs which have same degree set (set of all degrees) but are not isomorphic.

2 Paths and cycles

An important question in graph theory is of connectivity. Given a graph, can we reach from a source vertex s to target vertex t using the edges of the graph?

A *walk* of length k in a graph is a sequence of vertices $x_0, x_1, \cdots, x_k$, where any two consecutive vertices are connected by an edge in the graph. In a walk, it is allowed to take an edge or a vertex multiple times.

If all the vertices in a walk are distinct, except possibly the first and the last vertex, then it is called a *path*. In many places, paths are called simple paths and walks are called paths.

The path of length greater than two is called a cycle if the first and the last vertex are same.

Lemma 1. *If there is a walk between two vertices in a graph then there is a path between them.*

Proof. Consider the walk of least length, say $P = \{x_0, x_1, \cdots, x_k\}$. If all vertices are distinct then it is a path, otherwise there is a vertex v which occurs twice in the walk. If we delete the portion of the walk between the two occurrences of the vertex v, we still get a walk P' from x_0 to x_k.

But P was the smallest walk from x_0 to x_k, this is a contradiction. □

Exercise 3. Show that if there is a walk with the same first and last vertex and no two consecutive edges are the same, then there is a cycle in the graph.

A graph is called *connected* if there is a path between every pair of vertices of the graph. A graph can always be divided into disjoint parts which are connected within themselves but not connected to each other. These are called the *connected components* of the graph.

Exercise 4. Show that if there is path between a particular vertex v and every other vertex of the graph, then the graph is connected.

3 Trees

A graph is called a *tree* if,

- It is connected,
- There are no cycles in the graph.

Tree is a special graph where every pair of vertices have a unique path between them. Let's prove this formally.

Theorem 2. *A graph G is a tree iff there is a unique path between every pair of vertices in G.*

Proof. If there is a unique path between every pair of vertices then the graph is connected. Suppose such a graph has a cycle $x_0, x_1, \cdots, x_k$. Then there are two distinct paths between x_0 and x_1 (why?). So a graph with unique path between every pair is a tree.

For the converse, In a tree, suppose there are two paths between vertices u and v. Say $P = \{u = x_0, x_1, \cdots, x_k = v\}$ and $P' = \{u = y_0, y_1, \cdots, y_k = v\}$. Let i be the first index, s.t., $x_{i+1} \neq y_{i+1}$. Similarly j be the last index, s.t., $x_{j-1} \neq y_{j-1}$.

Exercise 5. Show that i and j exist and $j - i \geq 2$.

Then consider the walk $x_i, x_{i+1}, \cdots, x_j = y_j, y_{j-1}, \cdots, y_{i+1}, y_i$. This is a walk with first and last vertex being same and no two consecutive edges are same. So there is a cycle in this tree. Contradiction. □

Exercise 6. Show that if you remove any edge of a tree, you will get two disjoint connected components of the graph.

You will prove in the assignment that a tree always has a vertex of degree one. Then using induction on number of vertices, we can show that,

Theorem 3. *A tree on n vertices has $n - 1$ edges.*

If a subgraph T of a graph G is a tree then T is called a *spanning tree* of the graph G. If the graph G is connected, we can always construct a spanning tree. Define S_0 to be the initial set containing a particular vertex v. At every stage, construct S_{i+1} by including an edge which connects S_i to some vertex not in S_i.

Exercise 7. Why does this not create a cycle?

The process ends when S_i has all the vertices of the graph. We can proceed with each stage (find a vertex to add to S_i) because the graph is connected.

Spanning trees are important in many applications. They are the smallest structure which preserve the connectivity. If we assign weights/cost to every edge of the graph then we are mostly interested in minimum weight spanning tree. You will study algorithms to find minimum weight spanning tree in a graph in future courses.

4 Eulerian circuit and Hamiltonian cycle

A *simple circuit* is a walk where first and last point are the same and no edges are repeated. An *Euler's circuit* is a simple circuit which uses all possible edges of the graph.

Note 1. The vertices can be repeated

The definition of Euler circuits arose from the problem of Königsberg bridges.

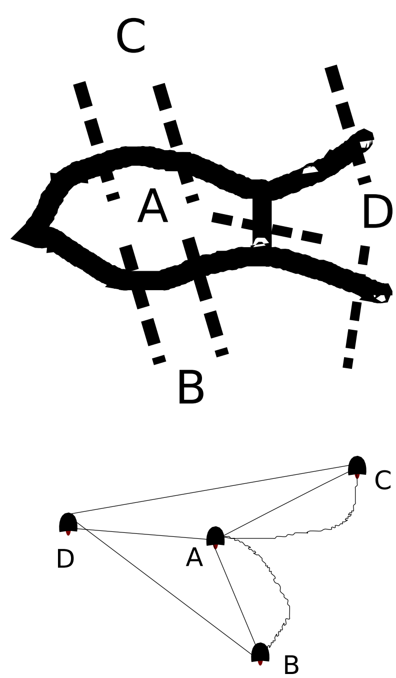

Fig. 3. Königsberg bridges and their graph representation. Dashed lines represent the bridges.

There are four regions connected with seven bridges. Can you go through all the bridges without revisiting any bridge? Converting it into graph as shown in Fig. 1, the question is equivalent to finding a Eulerian circuit in the graph.

The following theorem gives the answer.

Theorem 4. *A connected graph has a Eulerian circuit iff all vertices have even degree.*

Proof. If the graph has a Eulerian circuit than every vertex has even degree (show it as an exercise). Let's prove that if every vertex has an even degree then there is a Eulerian circuit.

We will construct the Eulerian circuit recursively. Start with an arbitrary vertex, say w, find an edge and move in that direction. Since the degree is even, when ever we arrive at a particular vertex, we have an edge to leave that vertex too.

Since the number of edges are finite, we will arrive back at the starting vertex (call this cycle C). If all edges are covered then we have found the Euler circuit. If not, call the subgraph with edges of C removed as H. All the connected components of H will be connected to cycle C (since the graph is connected). Suppose the connected components are $H_1, H_2, \cdots, H_k$ and connection points are $w_1, w_2, \cdots, w_k$ when traversing the cycle C from w in a particular direction (say anticlockwise).

Construct the Euler circuit in all connected components of H. We can use induction on number of edges, since H_i has less number of edges then the original graph. The base case, total 4 edges in graph, is easy.

Then the Euler circuit for G will be, go from w to w_1 using edges of C, take the Euler circuit of H_1, go from w_1 to w_2, take the euler circuit for H_2, $\cdots$, reach w_k and take the Euler circuit for H_k, come back to w using remaining edges of C.

Exercise 8. Write the pseudocode for finding a Eulerian circuit in a graph G.

□

Exercise 9. Show that it is not possible to take a Eulerian walk on the Königsberg bridges.

Similar to Euler circuits, a path $x_0, x_1 \cdots, x_k$ is called a *Hamiltonian path* if it goes through all the vertices of the graph. Remember that in a path we are not allowed to repeat the vertices.

It might seem like an easy task to give a necessary and sufficient condition for the existence of Hamiltonian path like Euler's circuit, but it turns out to be a really hard problem. You will study later that finding a Hamiltonian path in a graph is an NP-complete problem (the list of problems which are assumed to be hard for computers to solve efficiently).

There are many sufficient conditions known for existence and non-existence of Hamiltonian path.

Exercise 10. Show that a complete graph has a Hamiltonian path for $n \geq 3$.

Exercise 11. If the first and the last vertex of a Hamiltonian path is same then it is called a Hamiltonian cycle. Show that if a graph has a vertex of degree one then it can't have a Hamiltonian cycle.

5 Adjacency matrix

Adjacency matrix is a representation of a graph in matrix format. Given a graph G, its adjacency matrix A_G is a $|V| \times |V|$ matrix. Its rows and columns are indexed by the vertices of the graph. The (i,j)-th entry of the matrix is one if and only if vertex i is adjacent to vertex j, otherwise it is zero.

The adjacency matrix for the graph in Fig. 1 is,

$$A_G = \begin{pmatrix} 0\,0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,0\,1 \\ 0\,1\,0\,0\,1\,0 \\ 1\,0\,0\,0\,1\,0 \\ 0\,0\,1\,1\,0\,0 \\ 1\,1\,0\,0\,0\,0 \end{pmatrix}$$

Note 2. The adjacency matrix for an undirected graph is always symmetric.

Another matrix representation of a graph is called the *incidence matrix*. Here the rows are indexed by vertices and columns are indexed by edges. An entry (i,e) is one if and only if vertex i is part of edge e, otherwise it is zero. The incidence matrix for the graph in Fig. 1 is,

	e_1	e_2	e_3	e_4	e_5	e_6
v_1	1	0	0	0	0	1
v_2	0	0	1	1	0	0
v_3	0	0	0	0	1	1
v_4	1	1	0	0	0	0
v_5	0	1	0	0	1	0
v_6	0	0	1	0	0	1

Note 3. Here v_i represents the i-th vertex.

The eigenvalues and eigenvectors of the adjacency matrix provide us with lot of information about the graphs. We will see a few examples below. First, a lemma which will help us in proving results about eigenvalues.

Lemma 2. *Suppose $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ are the eigenvalues of M. Then the eigenvalues of $\lambda I + M$ are $\{\lambda + \lambda_1, \lambda + \lambda_2, \cdots, \lambda + \lambda_n\}$, where I is the identity matrix. The converse is also true.*

Proof. Exercise □

A graph is called *regular* if every vertex has the same degree d. In this case, d is called the degree of the graph.

Theorem 5. *The maximum eigenvalue of the adjacency matrix of a regular graph G is d, the degree of the graph.*

Proof. By using Lemma 2, it is sufficient to prove that all eigenvalues of $dI - A_G$ are greater than zero.

For the sake of contradiction, assume that $M = dI - A_G$ has a negative eigenvalue μ with eigenvector u. Then $u^T M u = \mu u^T u < 0$. We will prove that for every vector v, $v^T M v \geq 0$, hence get a contradiction.

Let v_i be the i-th entry of v and M_{ij} be the (i, j)-th entry. Then,

$$\begin{aligned} v^T M v &= \sum_{i,j} M_{ij} v_i v_j \\ &= \sum_i d_i (v_i)^2 - \sum_{(i,j) \in E_G} v_i v_j \\ &= \sum_{(i,j) \in E_G} (v_i - v_j)^2 \geq 0 \end{aligned} \tag{1}$$

Note 4. The term $v^T M v$ is known as the *quadratic form* of M.

To prove that there exist an eigenvalue d. Notice that every row of the matrix A_G sums up to d.

Exercise 12. What is the eigenvector corresponding to the eigenvalue d.

□

A graph is called *bipartite* if the vertex set can be divided into two parts A and B, s.t., there are no edges inside A and no edges inside B. A regular bipartite graph can be characterized by its eigenvalues.

Theorem 6. *The minimum eigenvalue of the adjacency matrix of a connected regular graph is greater than or equal to $-d$. It is bipartite if and only if the minimum eigenvalue is $-d$.*

Proof. By using Lemma 2, it is sufficient to prove that all eigenvalues of $dI + A_G$ are greater than or equal to zero. There will be an eigenvalue 0 iff the graph is bipartite.

Again we will notice the quadratic form $v^T(dI + A_G)v$,

$$\begin{aligned} v^T M v &= \sum_{i,j} M_{ij} v_i v_j \\ &= \sum_i d_i (v_i)^2 + \sum_{(i,j)\in E_G} v_i v_j \\ &= \sum_{(i,j)\in E_G} (v_i + v_j)^2 \geq 0 \end{aligned} \tag{2}$$

Suppose equality holds in the above equation. Notice that v_i corresponds to the i-th vertex. Let's say S_1 is the set of vertices for which v_i is positive and S_2 is the set of vertices for which v_i is negative. Then there are no edges inside S_1 and inside S_2. This implies that the graph is bipartite.

Exercise 13. Why do we need the condition that graph is connected.

If the graph is bipartite, then assign 1 to one part and -1 to other part. This will show that the least eigenvalue of $dI + A_G$ is zero (why?). ☐

Exercise 14. Show that the bipartite graph can't have an odd cycle.

A *complete graph* is a simple graph where every possible edge is present. A complete graph on n vertices is called K_n.

Exercise 15. How many edges are there in a complete graph?

The adjacency matrix of the complete graph K_n is $J - I$, where J is all 1's matrix and I is the identity matrix, both of dimension $n \times n$. You will find all the eigenvalues of K_n as part of the exercise.

6 Assignment

Exercise 16. There always exist a vertex of degree one in a tree.

Exercise 17. How many non-isomorphic graphs are there of degree 3 with 7 vertices?

Exercise 18. When will K_n, the complete graph, will have a Eulerian circuit? Can you give an explicit description of such a circuit when n is a prime?

Exercise 19. How many subgraphs are there for a complete graph.

Exercise 20. What are the eigenvalues of the adjacency matrix of the complete graph.

Exercise 21. Show that the maximum eigenvalue of an adjacency matrix is less than the maximum degree of the graph corresponding to it.

Exercise 22. For a graph G, show that the number of walks of length m between vertex i and vertex j is the (i, j)-th entry of A_G^m.

References

1. K. H. Rosen. Discrete Mathematics and Its Applications. *McGraw-Hill*, 1999.