MTH742P: Advanced Combinatorics 26/09/2016

Lecture 1: Basics of Graph Theory

Lecturer: Heng Guo

1 Graph Theory

Graphs are the central objects to study during our course. The concept of graphs is simple but it has found very wide application. It can be used to model many real-world networks, such as the internet.

This is a "pure mathematics" module, and hence we will focus on the mathematical theory of graphs, rather than on the applications.

Informally, a graph is a collection of points (called "vertices"), together with a collection of lines (called "edges") which join some pairs of points.

Definition 1. *A (labelled) graph is an ordered pair (V, E), where V is a set, and E is a set of (unordered) pairs of elements of V.*

If $G = (V, E)$ is a graph, then V is called the set of vertices of G, and E is the set of edges of G.

We also write $V(G)$ and $E(G)$ in place of V and E, if we want to emphasize the underlying graph.

Example: Let $V = \{1, 2, 3, 4, 5\}$ and $E = \{(1, 2), (2, 3), (4, 5)\}$.

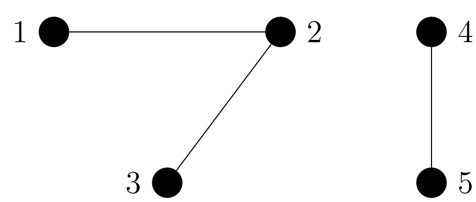


Figure 1: Graph G

Figure 1 is a picture of the graph. ⊠

For the most of the course, we will study *simple* and *finite* graphs. Here "finite" means that the vertex set V is finite, and "simple" means that there is no 'loops' and 'multiple edges' in G. A 'loop' is an edge joining a vertex to itself, such as $(1, 1)$. 'Multiple edges' are more than one edges joining the same pair of vertices, such as $(1, 2)$ and $(1, 2)$. When a graph contains loops or multiple edges, it is called a multigraph. In this course, all graphs are simple unless we explicitly mention multigraphs.

If we relabel all vertices, then we view it as the same graph.

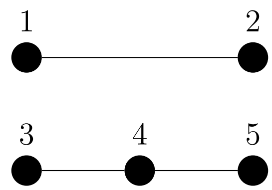


Figure 2: Graph H

For example, let H be a graph where $V(H) = \{1, 2, 3, 4, 5\}$ and $E(H) = \{12, 34, 45\}$. Note that here we write 12 instead of $(1, 2)$ for brevity. The graph H in Figure 2 and G in Figure 1 are the same graph. The relabelling is $1 \mapsto 3, 2 \mapsto 4, 3 \mapsto 5, 4 \mapsto 1, 5 \mapsto 2$.

To make it formal, we have the following definition.

Definition 2. *Let G and H be two labelled graphs. We say G and H are* isomorphic *if there exists a bijection $f : V(G) \to V(H)$ such that $(u, v) \in E(G)$ if and only if $(f(u), f(v)) \in E(H)$. The bijection f is called an* isomorphism *from G to H.*

Informally, an unlabelled graph is one where we omit the labelling of vertices. Thus G and H represent the same graph in Figure 3.

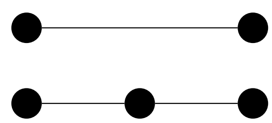


Figure 3: Unlabelled graph

To define unlabelled graph formally, define an equivalence relation $\sim$ by $G \sim H$ if and only if G and H are isomorphic. Thus an unlabelled graph is an equivalence class of $\sim$.

Real world examples:

- London tube map: vertices are stations, and two vertices are joined if and only if the two corresponding stations are adjacent stops on some line.
- The internet: webpages are vertices, and two webpages are joined if and only if one page links to the other. This is a very important application of graph theory. For example, the ranking algorithm of Google is based on this graph.
- The "Facebook" graph: vertices are people, and two vertices are joined if and only if they are friends.

  There is a famous "six degrees of separation" rule stating that for an arbitrary pairs of vertices (like 99%), there is a path of length at most 6 in this graph between them.

1.1 Basic definitions

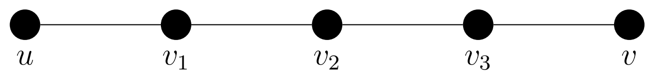
We now give some basic definitions of graph theory.

Definition 3. *Let G and H be two graphs. We say H is a subgraph of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$.*

Let G be a graph and $S \subseteq V$ is a subset of vertices. The induced subgraph of G on S is a graph H where $V(H) = S$ and $uv \in E(H)$ if and only if $uv \in E(G)$ and $u, v \in S$.

Definition 4. *If G is a graph and $u, v \in V(G)$, a* path *of length k in G from u to v is a subgraph P of G with k edges, where $E(P) = \{uv_1, v_1v_2, \cdots, v_{k-1}v\}$.*

Pictorially, a path (of length 4) has the form:



Definition 5. *A finite graph G is said to be* connected *if for any two vertices $u, v \in V(G)$, there is a path in G from u to v. If G is not connected it is said to be* disconnected.

For example, G in Figure 1 is disconnected.

Definition 6. *If G is a graph, a component of G is a maximal connected subgraph of G.*

In other words, a component of G is a subgraph H of G such that H is connected, and we cannot add any edges of G to H, nor any vertices of G into H, without H becoming disconnected. The components of G form a partition of $V(G)$.

Definition 7. *Let G be a graph and $u, v \in V(G)$. We say u is* adjacent *to v if $(u, v) \in E(G)$. The* neighbourhood *$\Gamma(v)$ of v is the set of all vertices of G that are adjacent to v; namely,*

$$\Gamma(v) = \{u \mid u \in V(G),\ (u, v) \in E(G)\}.$$

The degree *of v is the number of adjacent vertices of v, denoted by $d(v)$. In other words, $d(v) = |\Gamma(v)|$.*

Definition 8. *A graph G is said to be k-*regular *if $d(v) = k$ for all $v \in V(G)$.*

We write $A := B$ if we define A to be equal to B.

Definition 9. *Let $G = (V, E)$ be a graph. Define $|G| := |V(G)|$, the number of vertices of G. It is sometimes called the* order *of G.*

Define $e(G) := |E(G)|$, the number of edges of G.

Definition 10. *Let $G = (V, E)$ be a graph. We say G is* bipartite *if there is a partition $V_1 \cup V_2 = V$, $V_1 \cap V_2 = \emptyset$, and if $u, v \in V_i$, then $(u, v) \notin E$ for $i = 1, 2$.*

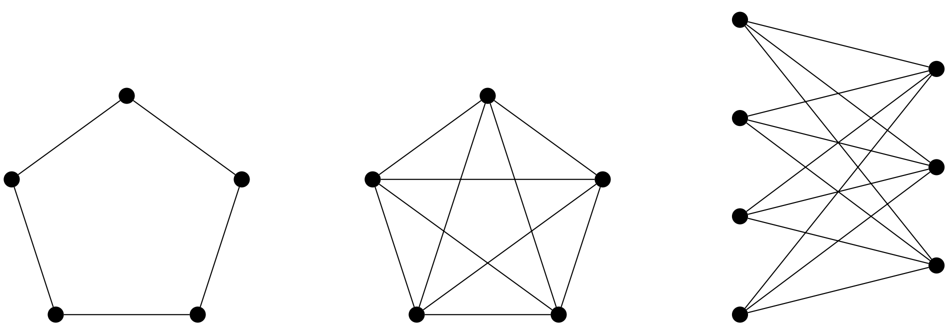
In other words, we can partition all vertices of a bipartite graph into two sides, where no edge joins two vertices of the same side.

Special Class of Graphs

- P_k = the path of length k. Note P_k has $k+1$ vertices and k edges.
- C_k = the path of length k. Note C_k has k vertices and k edges. For example, Figure 4a is C_5.
- K_k = the complete graph of order k. Here "complete" means that all possible edges are present. There is an edge between any pair of vertices. Hence $|K_k| = k$, $e(K_k) = \binom{k}{2}$, and K_k is $(k-1)$-regular.

  The complete graph is also called a "clique". For example, K_3 is simply a triangle. Another example can be found in Figure 4b, which is K_5.
- $K_{s,t}$ = the bipartite complete graph where the two sides have size s and t, respectively. Again, "complete" means that we add all possible edges to the graph. If two vertices u, v are on different sides of the graph, then (u, v) is an edge. Hence $|K_{s,t}| = s + t$ and $e(K_k) = st$.

  An example can be found in Figure 4c, which is $K_{4,3}$.



(a) C_5 (b) K_5 (c) $K_{4,3}$

Figure 4: Some special graphs

2 Extremal Graph Theory

Extremal graph theory is a branch of combinatorics to study questions like 'What is the maximum possible number of edges a graph can have, if it has a certain property P?' In general, we are interested in maximising or minimising ('extremizing') some parameter, over the collection of all graphs having some given property.

The first naive question is: What is the maximum possible number of edges in a graph with n vertices? It is easy to see that the answer is $\binom{n}{2}$, which is the number of edges of K_n.

One of the earliest questions in extremal graph theory was asked by Mantel:

Question (Mantel's Question): What is the maximum possible number of edges in a graph with n vertices that does not contain a triangle as a subgraph?

We call G *triangle-free* if G does not contain a triangle as a subgraph. Thus, we want to maximize a parameter (the number of edges) over all possible triangle-free graphs of order n.

How to construct a triangle-free graph with a lot of edges? First notice that a bipartite graph is triangle-free. Thus intuitively the bipartite complete graph $K_{n/2,n/2}$ is a candidate (assume n is even). If n is odd, we have to put $\lfloor n/2 \rfloor$ on one side and $\lceil n/2 \rceil$ on the other. Hence the candidate answer is $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ if n is odd. Mantel showed that these two cases are indeed the right answers.

Note that $e(K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}) = \lfloor n/2 \rfloor \lceil n/2 \rceil = \lfloor n^2/4 \rfloor$.

Theorem 1 (Mantel 1907)**.** *Let $G = (V, E)$ be a triangle-free graph on n vertices. Then*

$$e(G) \leq e(K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}) = \left\lfloor n^2/4 \right\rfloor,$$

where the equality holds if and only if $G = K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

We will first show Theorem 1 using Mantel's original method. An important ingredient is the Cauchy-Schwarz inequality.

Lemma 2 (The Cauchy-Schwarz inequality)**.** *If $x_1, \cdots, x_n \in \mathbb{R}$ and $y_1, \cdots, y_n \in \mathbb{R}$, then*

$$\left(\sum_{i=1}^{n} x_i y_i \right)^2 \leq \left(\sum_{i=1}^{n} x_i^2 \right) \left(\sum_{i=1}^{n} y_i^2 \right).$$

Proof. Consider the quadratic polynomial:

$$(x_1 z + y_1)^2 + (x_2 z + y_2)^2 + \cdots (x_n z + y_n)^2 \geq 0.$$

Note that

$$(x_1 z + y_1)^2 + (x_2 z + y_2)^2 + \cdots (x_n z + y_n)^2 = \left(\sum_{i=1}^{n} x_i^2 \right) z^2 + 2 \left(\sum_{i=1}^{n} x_i y_i \right) z + \left(\sum_{i=1}^{n} y_i^2 \right).$$

Since it is non-negative, it has at most one real root for z. Thus, its discriminant is less than or equal to 0. Hence

$$\left(2 \left(\sum_{i=1}^{n} x_i y_i \right) \right)^2 - 4 \left(\sum_{i=1}^{n} x_i^2 \right) \left(\sum_{i=1}^{n} y_i^2 \right) \leq 0.$$

Dividing the above equation by 4 yields the lemma. $\square$

Corollary 3. *Let $x_1, \cdots, x_n$ be n real numbers such that $\sum_{i=1}^{n} x_i = N$. Then*

$$\sum_{i=1}^{n} x_i^2 \geq \frac{N^2}{n}.$$

Proof. Take $y_i = 1$ in Lemma 2:

$$n\left(\sum_{i=1}^{n} x_i^2\right) \geq \left(\sum_{i=1}^{n} x_i\right)^2 = N^2 \qquad \square$$

Another ingredient is the famous handshaking lemma. Its proof utilizes an important trick called "double counting".

Lemma 4 (The Handshaking lemma)**.** *Let G be a graph. Then*

$$\sum_{v \in V(G)} d(v) = 2e(G).$$

Proof. We count the total number of ordered pairs of vertices that are adjacent. On one hand, each edge (u, v) is counted twice: (u, v) and (v, u). Hence the answer should be $2e(G)$.

On the other hand, for each vertex v, it contributes $d(v)$ to this sum. The answer should be $\sum_{v \in V(G)} d(v)$.

However these two answers are counting the same quantity. Hence they are equal. $\square$

Now we are ready to prove Theorem 1.

Proof of Theorem 1. For any edge $uv \in E(G)$, since G is triangle-free, then no vertex can be in the neighbourhood of both u and v. It implies that

$$\Gamma(u) \cap \Gamma(v) = \emptyset.$$

Thus,

$$d(u) + d(v) \leq n.$$

We sum over all edges $(u, v) \in E$:

$$\sum_{uv \in E} (d(u) + d(v)) \leq ne(G).$$

In the lefthand side of the equation above, each $d(v)$ appears $d(v)$ many times in the sum. Hence,

$$\begin{aligned} ne(G) &\geq \sum_{uv \in E} (d(u) + d(v)) \\ &= \sum_{v \in V} d(v)^2. \end{aligned}$$

Now we apply the Cauchy-Schwarz inequality:

$$\begin{aligned} ne(G) \geq \sum_{v \in V} d(v)^2 &\geq \frac{\left(\sum_{v \in V} d(v)\right)^2}{n} && \text{(By Corollary 3)} \\ &= \frac{4e(G)^2}{n}. && \text{(By Lemma 4)} \end{aligned}$$

Collect terms and we get

$$e(G) \leq \frac{n^2}{4}.$$

The theorem follows because $e(G)$ has to be an integer. $\square$

In the following we give an alternative proof of Theorem 1 using induction.

Alternative Proof of Theorem 1. We do an induction on n. The base cases of $n = 1, 2$ are trivial. Any graph with 1 vertex has no edge, and any graph with 2 vertices has at most 1 edge.

For the induction step, assume that Theorem 1 holds for any triangle-free graph G with $|G| \leq n - 1$, where $n \geq 3$. We will show that Theorem 1 holds for any triangle-free graph G with $|G| = n$.

Suppose for contradiction that the theorem does not hold, that is $e(G) \geq \left\lfloor \frac{n^2}{4} \right\rfloor + 1$. Take $uv \in E$. Let H be the induced subgraph of G on $V \backslash \{u, v\}$. Since G is triangle-free, so is H. Clearly $|H| = n - 2$ and by the induction hypothesis $e(H) \leq \left\lfloor \frac{(n-2)^2}{4} \right\rfloor$.

It implies that in G, the number of edges between H and $\{u, v\}$ is at least

$$\left\lfloor \frac{n^2}{4} \right\rfloor + 1 - \left\lfloor \frac{(n-2)^2}{4} \right\rfloor - 1 = \begin{cases} n - 1 & \text{if } n \text{ is even;} \\ \frac{n^2-1}{4} - \frac{(n-2)^2-1}{4} = n - 1 & \text{if } n \text{ is odd.} \end{cases}$$

Hence regardless of n being even or odd, there are at least $n-1$ many edges between H and $\{u, v\}$. Recall that $|H| = n - 2$, there must be one vertex w adjacent to both u and v. It means that $\{u, v, w\}$ is a triangle. Contradiction. $\square$

Theorem 1 also asserts when the maximum number of edges is achieved. It will be left as an exercise.