MA 5109: Extremal Graph Theory

Amit Rajaraman

Last updated July 27, 2021

§0. Notation

We use $[n]$ to represent $\{1, 2, \ldots, n\}$.
For integers a and b, $[a, b]$ means $\{a, a+1, \ldots, b\}$.
A graph G_n is a graph with n vertices.
Given a graph G, $e(G)$ is the number of edges G has.
For a vertex v, denote by $N(v)$ the set of *neighbours* of v – all the vertices that have an edge to v.
For a vertex v, denote by $d(v) = |N(v)|$ the *degree* of v – the number of edges incident on it.
For $v \in V$ and $K \subseteq V$, $d(v, K)$ is the number of edges

$$\big|\{u \in K : uv \in E\}\big|$$

from v into K.

§1. Introduction

1.1. What are Graphs?

Definition 1.1. A (simple undirected) **graph** G is an ordered pair (V, E) where V is a finite set called the *vertex set* and E, called the *edge set*, is a subset of $\binom{V}{2}$, where $\binom{S}{k}$ represents the set of all k-element subsets of S.

We typically represent graphs pictorially, showing vertices as dots and edges as arcs joining the vertices present in the corresponding subset.
A few important graphs are:

- the *null graph* with vertex set V, where $E = \varnothing$.
- the *complete graph* K_n, where $V = [n]$ and $E = \binom{[n]}{2}$.
- the *complete bipartite graph* $K_{m,n}$, where $V = A \cup B$ with $|A| = m$, $|B| = n$, and A, B are disjoint, and $E = \{\{a, b\} : a \in A, b \in B\}$.
- the *path graph* of length n, where $V = [n+1]$ and $E = \{\{m, m+1\} : m \in [n]\}$.
- the *cycle* of length n, where $V = [n]$ and $E = \{\{l, m\} : l, m \in [n], (m - l) \equiv 1 \pmod{n}\}$.

Now, consider the graph G with vertex set $[4]$ and edge set $\{\{1, 3\}, \{3, 2\}, \{2, 4\}\}$. This graph appears to be the same as the path graph of length 3, but how do we make this correspondence more concrete?
Relabeling vertices doesn't create a "new" graph.

Definition 1.2 (Graph Isomorphism)**.** Two graphs $G = (V, E)$ and $G' = (V', E')$ are said to be **isomorphic** and we write $G \simeq G'$ if there exists a bijection $f : V \to V'$ such that there is an edge between two vertices u and v in G if and only if there is an edge between $f(u)$ and $f(v)$ in G'.

If two graphs are isomorphic, they are identical for our purposes (we only care about graphs up to isomorphism).

Definition 1.3 (Subgraph)**.** Given a graph $G = (V, E)$, a **subgraph** $H = (V', E')$ is a graph such that $V' \subseteq V$ and $E' \subseteq E$. Given $V' \subseteq V$, the subgraph *induced* by V' on G is that with vertex set V' and edge set $\binom{V'}{2} \cap E$.

1.2. The birth of Extremal Graph Theory

Extremal graph theory is motivated by the following simple problem:

> At most how many edges can a graph G_n have if it contains no triangles?

More precisely, what is

$$\max_{\substack{\text{no subgraph of } G_n \\ \text{is isomorphic to } K_3}} e(G_n)?$$

Clearly, this number is well-defined since a graph on n vertices cannot have more than $\binom{n}{2}$ vertices.
A simple observation is that any complete bipartite graph has no triangles: if there were a triangle, then two vertices would be in the same "part", which contradicts the existence of edges only between the two parts.
As a consequence, for any $1 \leq m \leq n$, it is possible to construct $m \times (n - m)$ edges (with this bound being attained for $K_{m,n-m}$). In particular, it is possible to construct a graph with $\lfloor n^2/4 \rfloor$ edges.

Theorem 1.1 (Mantel's Theorem)**.** If G_n has no triangle, then

$$e(G_n) \leq \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Further, equality is attained iff $G_n \simeq K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Proof. Suppose G_n has no triangles. Saying that G_n has no triangles is equivalent to saying that for distinct adjacent u, v, $N(u) \cap N(v) = \varnothing$.
So, $d(u) + d(v) \leq n$. Therefore,

$$\begin{aligned}
ne(G_n) &\overset{(1)}{\geq} \sum_{uv \in E} d(u) + d(v) \\
&= \sum_{uv \in E} |N(u) \cup N(v)| \\
&= \left|(e, w) : e = uv \in E, w \in N(u) \cup N(v)\right| \\
&= \sum_{u \in V} |\{(e, w) : w \in N(u), e = uv \in E\}| \\
&= \sum_{u \in V} |\{(v, w) : v, w \in N(u)\}| \\
&\overset{(2)}{=} \sum_{u \in V} d(u)^2 \\
&\overset{(3)}{\geq} \frac{1}{n} \left(\sum_{u \in V} d(u) \right)^2 \\
&\overset{(4)}{=} \frac{4e(G_n)^2}{n},
\end{aligned}$$

where (2) follows from the changing the main thing being summed over to u, the "middle" vertex in the L-like structure, (3) follows from the Cauchy-Schwarz inequality, and (3) follows from the handshaking lemma.

What happens when equality is attained? Let us look at the case where n is even.
(1) is only tight when $d(u) + d(v) = n$ for all edges uv and (3) is only tight when $d(u)$ is a constant (independent of u). This implies that $d(u) = \frac{n}{2}$ for every $u \in V$. Now, if uv is an edge, $N(u) \cap N(v) = \varnothing$ implies that $N(u) \cup N(v) = V$, and so $G_n = K_{\frac{n}{2},\frac{n}{2}}$.
The case where n is odd is analyzed similarly, with slight nuances in (3) since exact equality is not attained. ■

While the above is one of the early results in extremal graph theory, the subject was only really born due to Turán in the following result.

Theorem 1.2 (Turán's Thoerem)**.** If G_n has no K_{r+1} $(r \geq 2)$, then $e(G_n) \leq t_r(n)$, with equality attained iff $G \simeq T_r(n)$.

The version for $r = 2$ is just a triangle-free graph and is the same as Mantel's Theorem. In the proof of this, we split the vertex set into two parts and dumped all the edges between these parts.
If we want to avoid K_4 $(r = 3)$, then perhaps we could split the vertex set into three parts and dump all the edges between these parts.
In general, we want to partition V of size n into r "almost equal" parts and set only those edges between vertices in distinct parts – such a graph is known as the **Turán graph** $T_r(n)$ and the number of edges $e(T_r(n))$ is the **Turán number** $t_r(n)$.
In particular, when $r \mid n$,

$$t_r(n) = \binom{r}{2}\left(\frac{n}{r}\right)^2 = \frac{n^2}{2}\left(1 - \frac{1}{r}\right).$$

Here, we give three proofs of Turán's Theorem.

Proof 1 of Turán's Theorem. We perform strong induction on $n + r$. We have already proved the result for $r = 2$.
Suppose $e(G_n) \geq t_r(n)$ and G_n is K_{r+1}-free, where $r > 2$. We wish to prove that $G \simeq T_r(n)$.
Since $t_r(n) \geq t_{r-1}(n)$ (check this!), the inductive hypothesis implies that G has a copy $K \subseteq V$ of K_r. Observe that for $v \notin K$, $d(v, K) \leq r - 1$ – otherwise, there would be a copy of K_{r+1} in G.
As a result, $e(V \setminus K, K) \leq (r-1)(n-r)$. By the induction hypothesis, $e(V \setminus K, V \setminus K) \leq t_r(n-r)$. Therefore,

$$t_r(n) \leq e(G_n) \leq t_r(n-r) + (r-1)(n-r) + \binom{r}{2}.$$

However, as can be checked manually, $t_r(n-r) + (r-1)(n-r) + \binom{r}{2} = t_r(n)$!
It follows that equality holds everywhere – $e(G_n) = t_r(n)$, $e(V \setminus K) = t_r(n-r)$, and $d(v, K) = r - 1$ for all $v \in V \setminus K$.
This graph is then isomorphic to $T_r(n)$ – for each $v \in V \setminus K$, we can put the vertex in K that is not adjacent to v in the same bucket as v. Then, the only edges are those between distinct buckets (Why?). ■