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 $d_G(v) = |N(v)|$ the *degree* of v – the number of edges incident on it. If the graph G is clear from context, we write simply $d(V)$.
For $v \in V$ and $K \subseteq V$, $d(v, K)$ is the number of edges

$$\left|\{u \in K : uv \in E\}\right|$$

from v into K.

§1. Introduction

1.1. Basic Definitions

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). We now give a few more definitions that are useful.

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 $\binom{V'}{2} \cap E$.

Definition 1.4 (r-partite Graph)**.** A graph $G = (V, E)$ is said to be r**-partite** if there exists a partition $V_1, V_2, \ldots, V_r$ of V such that for any edge $e = uv \in E$, u and v are in distinct V_i. That is, there are no edges within any of the V_i. In particular, a 2-partite graph is said to be **bipartite**.

Definition 1.5 (Independent Set)**.** Given a graph $G = (V, E)$, $I \subseteq V$ is said to be **independent** if no two vertices of I are adjacent (the subgraph induced by I is null).
$\alpha(G)$, the *independence number* of G, denotes the size of the largest independent set in G.

Definition 1.6 (Clique)**.** Given a graph $G = (V, E)$, $K \subseteq V$ is said to be a **clique** if any two vertices of K are adjacent (the subgraph induced by I is complete). $\omega(G)$, the *clique number* of G, denotes the size of the largest clique in G.

Definition 1.7 (Complement Graph)**.** Given a graph $G = (V, E)$, the **complement graph** of G is $\bar{G} = (V, \binom{V}{2} \setminus E)$.

Observe that $S \subseteq V$ is independent in G if and only if S is a clique in $\bar{G}$. In particular, $\alpha(G) = \omega(\bar{G})$.

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 $1 \le m \le n$, 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,

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

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 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?), so $G_n \simeq T_r(n)$. ■

Erdős' Proof of Turán's Theorem. Erdős proves a slightly more general claim: given a K_{r+1}-free graph G_n, there exists an r-partite graph H on V such that $d_G(v) \leq d_H(v)$ for all $v \in V$.
It is then a simple task to check that among the r-partite graphs on n vertices, the Turán graph $T_r(n)$ has the most edges.

To prove our claim, we perform induction on r.
The claim is trivial for the base case $r = 1$.
Now, suppose the claim holds for values less than r. Let $v_0 \in V$ such that $d_G(v_0) = \max_{v \in V} d_G(v)$ (the vertex of maximum degree in G) and $W = N(z)$. Since G is K_{r+1}-free, W is K_r-free. Inductively, there is an $(r-1)$-partite graph H' on W such that for all $v \in W$, $d_{H'}(v) \geq d_W(v)$.
Let $U = V \setminus W$. For each $u \in U$, remove all its edges in G and set its new neighbour set as W.
Our desired graph H is that with these edges along with those in H' and the edges from v_0 to W. That is, the rth part is $U \cup \{v_0\}$ and the remaining $(r-1)$ parts are those formed by H'. The graph is clearly r-partite by definition.
What about the degree inequality?

- $d_G(v_0) = d_H(v_0)$ trivially.
- For $u \in U$, $d_H(u) = d_G(v_0) \geq d_G(u)$.
- For $w \in W$,

$$d_H(w) = |U| + 1 + d_{H'}(w) \geq |U| + 1 + d_W(w) \geq d_U(w) + 1 + d_W(w) = d_G(w).$$ ■

(Why does equality imply that the graph is isomorphic to $T_r(n)$?)

Theorem 1.3 (Turán's Theorem, reformulation)**.** If $d = e(G_n)/n$ is the average degree of the vertices of G_n, then G_n has an independent set of size at least $n/(d+1)$.

Proof. Why is this equivalent to Turán's Theorem?
If G_n has no K_{r+1}, then $\alpha(\bar{G}) \leq r$. If $\bar{G_n}$ has average degree d, the above result would imply that $r \geq n/(d+1)$, that is, $d \leq (n/r) - 1$. The total number of edges in G_n is then

$$\binom{n}{2} - \frac{nd}{2} \leq \binom{n}{2} - \frac{n}{2}\left(\frac{n}{r} - 1\right) = \frac{n^2}{2}\left(1 - \frac{1}{r}\right),$$

which gives Turán's bound!

Let us now get to the proof of the above reformulation. First, consider the following algorithm to come up with *some* independent set in G:

1. Order V to get $\{v_1, \dots, v_n\}$ and initialize $S = \varnothing$.
2. Add v_1 to S.
3. Having processed vertices v_1 through v_i, add v_{i+1} to S iff there is no vertex in S that is adjacent to v_{i+1}.

It is clear that this always produces an independent set, but the size of the independent set depends on the ordering we choose at the beginning.
For a given ordering σ, denote by $\mathcal{A}(\sigma)$ the independent set produced by the algorithm.
How do we choose a "good" ordering?
Enter the probabilistic method. Pick a random order, that is, a (uniformly) random permutation π of the vertices in V. Then,

$$\begin{aligned}
\mathbf{E}[|\mathcal{A}(\pi)|] &= \mathbf{E}\left[\sum_{v \in V} \mathbb{1}_{v \in \mathcal{A}(\pi)}\right] \\
&= \sum_{v \in V} \mathbf{E}\left[\mathbb{1}_{v \in \mathcal{A}(\pi)}\right] \\
&= \sum_{v \in V} \Pr\left[v \in \mathcal{A}(\pi)\right].
\end{aligned}$$

Fix some $v \in V$. What is the probability that $v \in \mathcal{A}(\pi)$?
If at the time of processing v for the ordering π, $N(v) \cap S \neq \varnothing$, then v is not picked. In particular, if v is the first element of $N(v) \cup \{v\}$ in the ordering π, then it is definitely chosen by the algorithm. The probability of this occurring is $\frac{1}{d(v)+1}$. So,

$$\begin{aligned}
\mathbf{E}[|\mathcal{A}(\pi)|] &= \sum_{v \in V} \Pr\left[\mathbb{1}_{v \in \mathcal{A}(\pi)}\right] \\
&\geq \sum_{v \in V} \frac{1}{d(v)+1} \\
&\overset{(*)}{\geq} \frac{n^2}{\sum_{v \in V}(d(v)+1)} = \frac{n}{d+1},
\end{aligned}$$

where $(*)$ follows from the AM-HM inequality.
Since the expectation of $|\mathcal{A}(\pi)|$ is at least $n/(d+1)$, there must exist some permutation σ such that $|\mathcal{A}(\sigma)| \geq n/(d+1)$, proving the result. ■