Combinatorial Analysis

Raymond Bian

February 19, 2024

Contents

Lecture 1: Syllabus and Review

1 Introduction

This course is basically just a second course in Combinatorics, and will cover a range of topics.

Definition 1. **Matroids** are the structures that capture whether or not the greedy algorithm works. They will be covered later in the course.

Now, for some examples and review:

Definition 2. We say points are in **convex position** if no point is inside a triangle made by 3 other points.

Example. Given a finite set of points on the plane, what is the maxmimum number of points such that no 3 are on a line, and no 4 are in convex position.

Explanation. Informally, we know that the "outside" of our points has at most 3 points in the shape of a triangle. We can then place a point in the middle. However, if we try to add another point, then we find that 4 points are in convex position, which is a contradiction. Therefore, 4 points is the maximum size of such a set.

This example is actually part of a more general problem, shown below.

Theorem 1. (ES, 1935) The maxmimum number of points such that no 3 are on a line and no n are in convex position is $\leq 4^n$ and $\geq 2^{n-2}$.

Theorem 2. (Suk, 2017) This number is actually $\leq 2^{n+o(1)}$

Notation. Think of $o(1)$ as standing for a function $f(n)$ such that $\lim_{n\to\infty} f(n) = 0$. In other words, for every $\varepsilon > 0$, there exists n_0 such that $|f(n)| < \varepsilon$ for every $n \geq n_0$.

Example. How many distinct 5-letter words are there on the 26-letter english alphabet?

Explanation. There are 26 options for each of the 5 slots, so there are 26^5 words.

Example. What if repetitions aren't allowed?

Explanation. Each slot you lose an option, so there are $26 \cdot 25 \cdot 24 \cdot 23 \cdot 22 = \frac{26!}{21!}$ words.

Example. How many ways are there to choose 5 students out of 35 to present?

Explanation. There are $\binom{35}{5} = \frac{35!}{5! \cdot 30!}$ ways.

Lecture 2: Review of Proofs

We will now review the types of proofs covered in Math-3012, as well as guidelines for writing them in this class.

Notation. If F is a mapping from N to M, we write $F : N \to M$.

Notation. Sometimes, $N \backslash \{a\}$ will be instead written as $N - \{a\}$.

Proposition 1. Let N be an n-element set and M be an m-element set. Then, there are m^n mappings (or functions) from N to M.

Proof. (Inductive) We go by induction on n.

Base case.: For the base case $n = 0$, we consider the empty set $\varnothing$ to be a mapping from the empty set to M. So $m^0 = 1$ and the base case holds.

Inductive step.: Now, let $n \geq 1$ and assume that the proposition holds for $n-1$ by induction. So, let $a \in N$. There are m^{n-1} mappings $F' : N \backslash \{a\} \to M$. For each such F', we have m choices for where to send a. These mappings are all distinct, and every $F : N \to M$ can be obtained in this way. So, the number of mappings $F : N \to M$ is $m^{n-1} \cdot m = m^n$, as desired.

■

Definition 3. A **bijection** is a function $f : X \to Y$ such that f is one-to-one and onto.

Corollary. An n-element set X has 2^n many subsets.

Proof. (Bijective) For each $A \subseteq X$, let $F_A : X \to \{0,1\}$ such that for each $x \in X$,

$$F_A(x) = \begin{cases} 0 & \text{if } x \notin A \\ 1 & \text{if } x \in A \end{cases}.$$

These mappings $F_A, F_{A'}$ are distinct for distinct subsets $A, A' \subseteq X$, and every mapping $F : X \to \{0,1\}$ is equal to F_A for some $A \subseteq X$. So by proposition 1, the corollary holds. ■

Lemma 1. For any non-negative integers n, k ($n, k \in Z_{\geq 0}$), we have $\binom{n}{k} = \binom{n}{n-k}$.

Proof. (Algebraic) We have

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{k!(n-k)!} \\ &= \frac{n!}{(n-(n-k))!(n-k)!} \\ &= \binom{n}{n-k}, \end{aligned}$$

as desired. ■

Theorem 3. (Binomial Theorem) Let $n \in \mathbb{Z}_{\geq 0}$. Then

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Proof. Consider

$$\underbrace{(x+y)(x+y)\ldots(x+y)}_{n \text{ times}}.$$

For each $(x+y)$ term, we select either the x or the y, and there are $\binom{n}{k}$ ways to select k x's and $n-k$ y's. The formula follows. ■

Corollary. For any $n \in \mathbb{Z}_{\geq 0}$, we have

$$2^n = \sum_{k=0}^{n} \binom{n}{k} \text{ and } 0 = \sum_{k=0}^{n} \binom{n}{k} (-1)^k.$$

Proof. Apply the binomial theorem with $x = y = 1$ to yield the first result, and with $x = -1, y = 1$ to yield the second. ■

1.1 Counting Review

Definition 4. A **permutation** is a bijection from a finite set to itself.

Example. One such bijection could be $1 \mapsto 2, 2 \mapsto 1, 3 \mapsto 4, 4 \mapsto 5, 5 \mapsto 3$.

Lemma 2. The number of such bijections is $n!$.

Proof. Exercise to the student! ■

Lecture 3: Permutations and Cycles

Notation. $\tau : X \to X$ is a permutation **on** X. Can also be denoted by $\sigma : X \to X$.

We will show that all permutations τ can be "decomposed" into "cycles".

Example. From the example earlier, $(1,2)$ is a cycle, and $(3,4,5)$ is another cycle.

For the following, let $\tau : X \to X$.

Definition 5. A **cycle** of τ is a tuple (ordered set of elements) $(x_1, x_2, \ldots, x_k)$ such that $x_1, x_2, \ldots, x_k$ are distinct elements of X, and $\tau(x_1) = x_2, \tau(x_2) = x_3, \ldots, \tau(x_{k-1}) = x_k, \tau(x_k) = x_1$. We call $x_1, x_2, \ldots x_k$ the **elements** of the cycle.

Lemma 3. If $(x_1, x_2 \ldots, x_k)$ and $(y_1, y_2, \ldots y_r)$ have an element in common, then $\{x_1, x_2, \ldots x_k\} = \{y_1, y_2, \ldots y_r\}$.

Proof. Note that since $(x_1, x_2, \ldots, x_k)$ is a cycle, $(x_2, x_3, \ldots, x_k, x_1)$ is also a cycle. Because of this, we can assume that $x_1 = y_1$. So $x_2 = \tau(x_1) = \tau(y_1) = y_2$. Then, we have that $x_2 = y_2$. We can repeat this process until $x_k = y_k$ (swap x, y if $k > r$). Then, we have $x_1 = \tau(x_k) = \tau(y_k) = y_1$, which means that $r = k$. Therefore, all cycles are pairwise disjoint. ■

Lemma 4. For every $x \in X$, there exists a cycle of τ which has x as an element.

Proof. Consider visiting each element $x, \tau(x), \tau(\tau(x)), \ldots$, until the first time we re-visit any element. This will eventually happen, because X is finite. Then, let's suppoose that we have visited elements $x_1, x_2, \ldots x_k$ so far, such that $x_1, x_2, \ldots, x_k$ are distinct, and that $\tau(x_k) = x_i$ for some $i \in \{1, 2, \ldots, k\}$. We cannot have $i \geq 2$ because then both x_{i-1} and x_k would both map to x_i, which is a contradiction because a permutation is a bijection. Therefore, $i = 1$ and we have established our cycle. ■

Corollary. There exists cycles $C_1, C_2, \ldots, C_t$, so that every element of X is an element in exactly one such cycle.

Definition 6. The **cycle notation** for τ is written as
$$\tau = C_1 C_2 \ldots C_t.$$

Example. Find the cycle notation for the permutation τ of $\{1, 2, 3, 4, 5, 6\}$ where

$$\begin{aligned}
\tau(1) &= 4 \\
\tau(2) &= 6 \\
\tau(3) &= 2 \\
\tau(4) &= 5 \\
\tau(5) &= 1 \\
\tau(6) &= 3.
\end{aligned}$$

Explanation. By inspection, we have a cycle $(1, 4, 5)$ and another cycle $(2, 3, 6)$. Therefore, $\tau = (1, 4, 5)(2, 3, 6)$.

Definition 7. A **transposition** is a cycle with exactly two elements.

Problem. How quickly does $n!$ grow as n gets large?

Lecture 4: Estimates for $n!$

Lemma 5. (Simplest) For any positive integer $n \in \mathbb{Z}_{>0}$,
$$2^{n-1} \leq n! \leq n^{n-1}.$$

Proof. We have for the lower bound
$$n! = \prod_{i=2}^{n} i \geq \prod_{i=2}^{n} 2 = 2^{n-1}.$$
And for the upper bound,
$$n! = \prod_{i=2}^{n} i \leq \prod_{i=2}^{n} n = n^{n-1}.$$
■

Note that these bounds are very far off. Here is a motivating example.

Example. Suppose n students draw a card from a deck of n cards, replacing the card afterwards. What is the likelihood that all n cards drawn are distinct?

Explanation. The probability is the number of desireable outcomes over the total number of outcomes. This is just
$$\frac{n!}{n^n}.$$
Note that if we use the upper bound from this lemma, we would get that the probability is at most $\frac{1}{n}$. In reality however, the true probability is much, much smaller.

Lemma 6. A better set of bounds are the following:
$$\left(\frac{n}{2}\right)^{\frac{n}{2}} \leq n! \leq \frac{(n+1)^n}{2^{\frac{n}{2}}}.$$

Proof. Left as an exercise! ■

Lemma 7. For any two $a, b \geq 2$, we have $a \cdot b \geq a + b$.

Lemma 8. (Arithmetic-Geometric Mean Inequality) For any two $a, b \geq 0$, we have
$$\sqrt{ab} \leq \frac{a+b}{2}.$$

With these last two lemmas, we can show the following:

Theorem 4. (Gauss) For any $n \in \mathbb{Z}_{>0}$,

$$n^{\frac{n}{2}} \leq n! \leq \frac{(n+1)^n}{2^n}.$$

Proof. We instead look at $(n!)^2$. Pairing 1 with n, 2 with $n-1$, etc, we have for the lower bound

$$\begin{aligned} n! &= \left(\prod_{i=1}^{n} i\right)\left(\prod_{i=1}^{n} i\right) \\ &= \prod_{i=1}^{n} i(n+1-i) \\ &= \prod_{i=1}^{n} \sqrt{i(n+1-i)} \\ &\geq \prod_{i=1}^{n} \sqrt{n} && \text{(Lemma 7)} \\ &\geq n^{\frac{n}{2}}. \end{aligned}$$

And for the upper bound, we have

$$\begin{aligned} n! &= \left(\prod_{i=1}^{n} i\right)\left(\prod_{i=1}^{n} i\right) \\ &= \prod_{i=1}^{n} i(n+1-i) \\ &= \prod_{i=1}^{n} \sqrt{i(n+1-i)} \\ &\leq \prod_{i=1}^{n} \frac{i+n+1-i}{2} \\ &= \frac{(n+1)^n}{2^n}. \end{aligned}$$

■

Theorem 5. (Even better bound) For any $n \in \mathbb{Z}_{>0}$, we have

$$e\left(\frac{n}{e}\right)^n \leq n! \leq en\left(\frac{n}{e}\right)^n.$$

Proof. The lower bound will be given as a homework problem. The upper bound is as follows. Note that $\ln(n!) = \sum_{i=1}^{n} \ln(i)$. Then, we can take the integral of $\ln(x)$, which is greater than this sum.

$$\begin{aligned} \sum_{i=1}^{n} \ln(i) &\leq \int_{1}^{n+1} \ln(x)dx \\ &= (n+1)\ln(n+1) - n. \end{aligned}$$

Thus

$$\begin{aligned} n! &\leq e^{(n+1)\ln(n+1)-n} \\ &= \frac{e^{(n+1)\ln(n+1)}}{e^n} \\ &= \frac{\left(e^{\ln(n+1)}\right)^{n+1}}{e^n} \\ &= \frac{(n+1)^{n+1}}{e^n}. \end{aligned}$$

Applying this for $n(n-1)!$ gives the bound. ■

Lecture 5: Asymptotic Analysis

Theorem 6. Stirling's Formula says that

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

Definition 8. For two functions $f, g : \mathbb{Z}_{>0} \to \mathbb{R}$, we write $f \sim g$ and say f is asymptotic to g if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1.$$

Also note that $f \sim g \Leftrightarrow g \sim f$.

Example. $2n + \sqrt{n} \sim 2n$.

Explanation. This is because

$$\lim_{n\to\infty} \frac{2n+\sqrt{n}}{2n} = 1.$$

Example. (Informal) How many digits are in 100!?

Explanation. Using Stirling's Formula, we have that

$$100! \sim \sqrt{2\pi 100}\left(\frac{100}{e}\right)^{100} = 9.324\ldots \times 10^{157}.$$

, whereas $100! = 9.332\ldots \times 10^{157}$ (very close approximation).

Definition 9. The n-th harmonic number

$$H_n = \sum_{i=1}^{n} \frac{1}{i}.$$

Theorem 7. (Euler-Mascheroni) $H_n \sim \ln(n)$.

Proof. Omitted. ■

Lemma 9. For any positive integer $n \in \mathbb{Z}_{>0}$, we have

$$\frac{\lfloor \log_2(n) \rfloor}{2} \le H_n \le \lfloor \log_2(n) \rfloor + 1.$$

Proof. We can break up the proof into parts of size $2, 4, 8, 16 \ldots$. Let $S_k = \{i \in \mathbb{Z}_{>0} : 2^{k-1} \le i \le 2^k - 1\}$ for any $k \in \mathbb{Z}_{>0}$. Note that $|S_k| = 2^{k-1}$. Also, for every $x \in S_k$, we have

$$\frac{1}{2^k} < \frac{1}{x} \le \frac{1}{2^{k-1}}.$$

Therefore, we have

$$\begin{aligned} H_n = \sum_{i=1}^{n} \frac{1}{i} &= \sum_{k=1}^{\lfloor \log_2(n) \rfloor} \sum_{x \in S_k} \frac{1}{x} \\ &\ge \sum_{k=1}^{\lfloor \log_2(n) \rfloor} \sum_{x \in S_k} \frac{1}{2^k} \\ &= \sum_{k=1}^{\lfloor \log_2(n) \rfloor} 2^{k-1}/2^k \\ &= \sum_{k=1}^{\lfloor \log_2 n \rfloor} \frac{1}{2} \\ &= \frac{\lfloor \log_2(n) \rfloor}{2}. \end{aligned}$$

In the other direction, we have

$$\begin{aligned} H_n &\le \sum_{k=1}^{\lfloor \log_2(n) \rfloor + 1} \sum_{x \in S_k} \frac{1}{x} \\ &\le \sum_{k=1}^{\lfloor \log_2(n) \rfloor + 1} \frac{|S_k|}{2^{k-1}} \\ &= \lfloor \log_2(n) \rfloor + 1. \end{aligned}$$

■

Definition 10. Let $f, g : \mathbb{Z}_{>0} \to \mathbb{R}$. We say $f = O(g)$ or f is big-O of g if there exists n_0, C, such that

$$|f(n)| \le C \cdot g(n) \quad \forall n \ge n_0.$$

Note. If $f, g : \mathbb{Z}_{>0} \to \mathbb{R}$ and $f \sim g$, we have

$$f = O(g) \quad \text{and} \quad g = O(f).$$

If $\varepsilon = 1$ for all significantly large n, $\frac{f(n)}{g(n)} \le 2$.

Example. $\sum_{i=1}^{n} \frac{1}{i} = O(\log n)$.

Lemma 10. Let $a, \alpha, \beta > 0$ be fixed. Then as $n \to \infty$,

- $n^\alpha = O(n^\beta)$ if $\alpha < \beta$.
- $n^\alpha = O(a^n)$ if $a > 1$.
- $(\ln(n))^\alpha = O(n^\beta)$.

Lecture 6: Binomial Coefficients and Counting Primes

Note. Note that we can also write, for functions f, g, h $f = g + O(h)$, which means that $|f - g| + O(h)$.

Example.

$$\binom{n}{2} = \frac{n(n-1)}{2} = \frac{n^2}{2} - \frac{n}{2} = \frac{n^2}{2} + O(n).$$

Definition 11. $f(n) = \Theta(g(n))$ if $f = O(g)$ and $g = O(f)$.

Definition 12. $f = o(g)$ if $\lim_{n \to \infty} \frac{f}{g} = 0$.

Example. What are all primes less than 20?

Explanation. 2, 3, 5, 7, 11, 13, 17, 19

Definition 13. Let $\pi(n)$ be the number of primes that are $\le n$.

Theorem 8. The **prime number theorem** states that

$$\pi(n) \sim \frac{n}{\ln(n)}.$$

The **Riemann Hypothesis** states that

$$\pi(n) = \int_1^n \frac{1}{\ln(x)} dx + O(\sqrt{n} \ln(n)).$$

It's called a hypothesis because it is often used in other mathematical proofs, even if not proved yet. For example, determining whether a knot could be un-knotted is in NP if the RH is true.

Lemma 11. For any $k \ge 1$, we have that

$$\binom{2k+1}{k} \le 4^k.$$

Proof. Later. ■

Lemma 12. For any $n \geq 2$, the product of all primes $\leq n$ is at most 16^n.

Proof. We wish to prove that

$$\prod_{i=1}^{\pi(n)} p_i \leq 16^n.$$

where p_i denotes the i-th prime. We proceed with induction on n.

Base case: $n = 2, 3$. Holds trivially.

Step case 1: n is even. Note that n cannot be prime, such that by induction,

$$\prod_{i=1}^{\pi(n)} p_i = \prod_{i=1}^{\pi(n-1)} p_i \leq 16^{n-1} \leq 16^n.$$

Step case 2: n is odd. We write $n = 2k+1$ for some $k \geq 1$. Note that every prime p such that $k+2 \leq p \leq 2k+1$ divides $\binom{2k+1}{k}$. This is because

$$\binom{2k+1}{k} = \frac{(2k+1)!}{k!(k+1)!}.$$

such that p divides the numberator but not the denominator.

By induction the product of primes p such that $0 \leq p \leq k+1$ is bound by

$$\prod_{i=1}^{\pi(k+1)} p_i \leq 16^{k+1}.$$

Combining our bounds and lemma, we have

$$\begin{aligned}
\prod_{i=1}^{\pi(n)} p_i &= \left(\prod_{i=1}^{\pi(k+1)} p_i\right)\left(\prod_{i=\pi(k+1)+1}^{\pi(n)} p_i\right) \\
&\leq 16^{k+1}\binom{2k+1}{k} \\
&\leq 16^{k+1} \cdot 4^k \\
&\leq 16^{2k+1}.
\end{aligned}$$

Note that bounding by 4^n may work here, but 16^n is presented due to a mistake in the lecture notes.

■

Theorem 9. The weak prime number theorem states that

$$\pi(n) = \Theta\left(\frac{n}{\ln n}\right).$$

Proof. We will show the upper bound, i.e.

$$\pi(n) = O\left(\frac{n}{\ln n}\right).$$

Let $p_1, p_2, \ldots$ be the sequence of primes. Then,

$$\pi(n)! \leq \prod_{i=1}^{\pi(n)} p_i \leq 16^n.$$

because $p_1 \geq 1, p_2 \geq 2$, etc. We have also shown that

$$e\left(\frac{\pi(n)}{e}\right)^{\pi(n)} \leq \pi(n)!.$$

As such,

$$e\left(\frac{\pi(n)}{e}\right)^{\pi(n)} \leq 16^n.$$

Taking the ln of both sides, we have

$$\begin{aligned}
\ln\left(e \to 1 \left(\frac{\pi(n)}{e}\right)^{\pi(n)}\right) &\leq \ln\left(16^n\right) \\
\pi(n) \cdot \ln\left(\frac{\pi(n)}{e}\right) &\leq n\ln(16)
\end{aligned}$$

.

Assume towards a contradiction that

$$\pi(n) \geq \frac{100n}{\ln n}$$

. Then,

$$\begin{aligned}
\frac{100n}{\ln n} \cdot \ln\left(\frac{100n}{e\ln n}\right) &\leq n\ln 16 \\
\frac{100}{\ln 16 \cdot \ln n}\ln\left(\frac{100n}{e\ln n}\right) &\leq 1 \\
\frac{100}{\ln 16 \cdot \ln n}(\ln(100n) - \ln(e\ln n)) &\leq 1
\end{aligned}$$

.

which is a contradiction (after many calculations). Therefore, there exists some $C < 100$ such that $\pi(n) = O(\frac{n}{\ln n})$. ■

Lecture 7: More Bounds, Inclusion Exclusion

The proof from previous lecture will not be given, as it is on the homework.

Lemma 13. For any $n \geq k \geq 1$,

$$\binom{n}{k} \leq \left(\frac{en}{k}\right)^k.$$

Observation. If $n = 2k$, then

$$\binom{2k}{k} \leq \left(\frac{e \cdot 2k}{k}\right)^k \leq (2e)^k.$$

Lemma 14. (Bernoulli-type inequality) For any real number $x \geq 0$, we have

$$1 + x \leq e^x.$$

Proof. This is true for $x = 0$ because $1 + 0 \leq e^0 = 1$. Therefore, we just need to show that

$$f(x) = e^x - (1 + x).$$

is increasing. This holds because $f'(x) = e^x - 1 \geq 0$. ■

Lemma 15. For any $n \geq k \geq 1$, we have

$$\sum_{i=0}^{k} \binom{n}{i} \leq \left(\frac{en}{k}\right)^k.$$

Proof. Recall that for any x, y, we have

$$(x + y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

When $y = 1$,

$$(1 + x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i.$$

Choosing $x = \frac{k}{n}$ with the intuition that $x < 1$ allows us to continue:

$$\sum_{i=0}^{n} \binom{n}{i} x^i \geq \sum_{i=0}^{k} \binom{n}{i} x^i.$$

Dividing everything by x^k,

$$\frac{(1+x)^n}{x^k} = \sum_{i=0}^{n} \binom{n}{i} x^{i-k} \geq \sum_{i=0}^{k} \binom{n}{i} x^{i-k}.$$

Since $x \leq 1$,

$$\frac{(e^x)^n}{x^k} \geq \frac{(1+x)^n}{x^k} \geq \sum_{i=0}^{k} \binom{n}{i} x^{i-k} \geq \sum_{i=0}^{k} \binom{n}{i}.$$

Plugging in $x = \frac{k}{n}$ gets our result. ■

Lemma 16. For any $n \geq k \geq 1$, we have

$$\binom{n}{k} \geq \frac{n^k}{k^k}.$$

Proof.

$$\begin{aligned}
\binom{n}{k} &= \frac{n!}{k!(n-k)!} \\
&= \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!} \\
&= \frac{n}{k} \cdot \frac{n-1}{k-1} \cdot \frac{n-2}{k-2} \cdot \ldots \cdot \frac{n-k+1}{1} \\
&\geq \frac{n^k}{k^k}.
\end{aligned}$$

■

1.2 Inclusion Exclusion

Example. Say there are 20 math majors, 15 CS majors, and 5 who are majoring in both in one class. How many people are in the class?

Explanation. $20 + 15 - 5 = 30$.

Lemma 17. If A and B are finite sets, then $|A \cup B| = |A| + |B| - |A \cap B|$.

Proof. Count. ■

Lemma 18. If A, B, C are finite sets, then

$$\begin{aligned}
|A \cup B \cup C| = |A| &+ |B| + |C| \\
&- |A \cap B| - |B \cap C| - |C \cap A| \\
&+ |A \cap B \cap C|.
\end{aligned}$$

Definition 14. Given a set S, and a positive integer $k \leq |S|$, we write

$$\binom{S}{k}.$$

to denote the set of subsets of S whose size is exactly k.

Note.

$$\left|\binom{S}{k}\right| = \binom{|S|}{k}.$$

Theorem 10. (Inclusion Exclusion Principle) If $A_1, A_2, \ldots, A_n$ are finite sets, then

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{k=1}^{n}(-1)^{k-1} \sum_{I \in \binom{\{1,2,\ldots,n\}}{k}} \left|\bigcap_{i \in I} A_i\right|.$$

Lecture 8: Derangements

Lecture 9: Extremal Combinatorics

We will start this unit by looking at graphs and hypergraphs.

Definition 15. A **hypergraph** is a pair $H = (V, E)$ such that V is a set, and E is a subset of the powerset of V ($E \subset 2^V$). Unless otherwise noted, V is a finite set.

Definition 16. Given a hypergraph $H = (V, E)$, then the elements $v \in V$ are called **vertices** and the elements $e \in E$ are called **edges**.

Definition 17. A hypergraph $H = (V, E)$ is **isomorphic** to another hypergraph $H' = (V', E')$ if there is a bijection $\phi : V \to V'$ between the vertex sets such that for any $S \subseteq V$, we have $S \in E$ iff $\phi(S) \in E'$.

Definition 18. A **graph** is a hypergraph in which every edge has size 2. In other words, $E \subseteq \binom{V}{2}$.

Definition 19. Given $H = (V, E)$, a vertex $v \in V$ is **incident** to an edge $e \in E$ if $v \in e$.

Definition 20. The **degree** of a vertex $v \in V$ is the number of edges it is incident to. We write this as

$$d_H(v) = |\{e \in E : v \in e\}|.$$

Definition 21. The **incidence mamtrix** M of a hypergraph $H = (V, E)$ is a $V \times E$ matrix (V rows and E columns) so that each entry

$$M_{v,e} : \begin{cases} 1, & \text{if } v \in e \\ 0 & \text{otherwise} \end{cases}.$$

Theorem 11. (Hypergraph handshaking). If $H = (V, E)$ is a hypergraph, then

$$\sum_{v \in V} d_H(V) = \sum_{e \in E} |e|.$$

Proof. Let M be the incidence matrix of H. Then $\sum_{v \in V} d_H(v)$ counts the number of 1's of M by summing along rows, and $\sum_{e \in E} |e|$ counts the number of 1's by summing along columns. As the number of 1's in the matrix is the same, these two values are the same. ∎

Corollary. For any graph, we have

$$\sum_{v \in V} d_G(v) = 2|E|.$$

Proof. This is because $|e| = 2$ for all $e \in E$ by definition of a graph. ∎

Lecture 10: Extremal Combinatorics Continued

Lemma 19. For any graph $G = (V, E)$, there exists a vertex v of small degree such that

$$\deg(v) \leq \frac{2|E|}{|V|}.$$

Proof. By the handshaking lemma,

$$\frac{\sum_{v \in V} \deg(v)}{|V|} = \frac{\sum_{e \in E} |e|}{|V|} = \frac{2|E|}{|V|}.$$

is the average degree of a vertex in the graph. Then, there exists $v \in V$ whose degree is at most the average. ∎

Definition 22. A **triangulation** is a sequence of triangles $T, T_1, T_2, \ldots, T_n$ such that

1. $T, T_1, T_2, \ldots, T_n \subseteq \mathbb{R}^2$,
2. $T = \bigcup_{i=1}^{n} T_i$,
3. For any distinct $i, j \in \{1, \ldots, n\}$, $T_i \cap T_j$ is either empty, consists of one vertex, or one edge.

Lemma 20. (Sperner) For any vertices such that

1. The 3 outer vertices get different colors,
2. Vertices on the edge of T are colored the

same as one of the edge's endpoints.

Then, there exists an inner triangle with 3 different colors. More formally, if $T, T_1, T_2, \ldots, T_n$, and v_1, v_2, v_3 are the vertices of T, and the vertices are colored by $\{1, 2, 3\}$, such that

1. v_i is color i for $i = 1, 2, 3$, and
2. vertices on the edge of T between v_i and v_j are colored either i or j for $i \neq j \in \{1, 2, 3\}$,

then there exists a triangle T_k, $k \in \{1, 2, \ldots, n\}$, such that the three vertices of T_k receive colors 1, 2, 3.

Proof. Define a graph $G = (V, E)$ such that $V = \{T, T_1, T_2, \ldots, T_n\}$. Connect two vertices if and only if the edge the two triangles share have colors 1 and 2.

Observation. It is enough to prove that a triangle has odd degree in G if and only if three vertices of the triangle receive distinct colors. Then, as the outer triangle has odd degree, some other smaller triangle must also have odd degree.

■

Theorem 12. (Brower's Fixed Point) Every continuous map from the unit disc $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ to itself has a fixed point (a point p such that $p \mapsto p$).

Theorem 13. (Product Structure Theorem) Also uses Sperner's Lemma.

Lecture 11: More on Hypergraphs

Continuing with the proof of Sperner's Lemma:

Definition 23. Two vertices are **adjacent** if they share a common edge.

Proof. Consider the outer triangle T. All smaller triangles T_i adjacent to T must lie on the edge of T which has colors $1, 2$. Note that there must be odd triangles adjacent to T, as the there must be odd number of changes from 1 to 2 on the edge.

Now, consider a smaller triangle's degree $d_G(T_i)$, where $i \in \{1, 2, \ldots, n\}$. If T_i has 3 different colors, then $d_G(T_i) = 1$. For the other direction, assume that $d_G(T_i)$ is odd. Then, there must be at least one edge that receives colors 1 and 2. There are then 3 cases for the third vertex:

Case 1: It is of color 1. Then, there are two edges that have colors 1, 2.

Case 2: It is of color 2. Then, there are again two edges that have colors 1, 2.

Case 3: It is of color 3. Then, there is only one edge.

Therefore, the degree of T_i is odd if and only if it is colored vertices of different colors. There must be one such T_i in the graph, because the number of odd degree vertices in the graph is even. ■

Definition 24. A hypergraph $H = (V, E)$ is **laminar** if for all pairs of its edges $A, B \in E$ A, B is either disjoint, or one is a subset of the other. More formally, $A \cap B = \varnothing$, $A \subseteq B$, or $B \subseteq A$.

Lemma 21. Every laminar hypergraph with n vertices has at most $2n - 1$ edges.

Proof. Suppose that this is not the case. Let $H = (V, E)$ be a counterexample such that $|V|$ is minimum, and $|E|$ is maximum. Because H is a counterexample, there must exist an edge $A \in E$ such that $A \neq V$. Choose A such that $|A|$ is maximum. By the choice of H, $\overline{A} = V \setminus A$ is an edge of H (because otherwise, we would have a counterexample with even more edges, but our counterexample is maximum). Consider

$$H_1 = (A, \{e \in E : e \subseteq A\})$$
$$H_2 = (\overline{A}, \{e \in E : e \subseteq \overline{A}\}).$$

Then, by the choice of H, neither H_1 nor H_2 is a counterexample. Therefore, the number of edges contained in A is at most $2|A| - 1$. Likewise, the number of edges contained in $\overline{A}$ is $2|\overline{A}| - 1$. Then, the only edge that is not a subset of A or $\overline{A}$ is V. Therefore,

$$|E| \leq (2|A| - 1) + (2|\overline{A}| - 1) - 1 = 2|V| - 1.$$

as desired. ■

Definition 25. A hypergraph is **Sperner** if there are no distinct edges A and B with $A \subset B$.

Corollary. For any integer $n \geq 1$, there exists a Sperner hypergraph on n vertices and

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

edges.

Theorem 14. This bound is tight.

Lecture 12: Sperner Graphs

Note. All normal graphs are Sperner, as all edges have the same size.

Lemma 22. (LMY-Inequality) If $H = (V, E)$ is a Sperner hypergraph, and $n = |V|$, then

$$\sum_{A \in E} \frac{1}{\binom{n}{|A|}} \leq 1.$$

Proof. Without loss of generality, let $V = \{1, 2, \ldots, n\}$. We proceed with a double counting argument. We let $A \subseteq \{1, 2, \ldots, n\}$ be **initial** for τ if $\tau(\{1, 2, \ldots, |A|\}) = A$. In other words, the first $|A|$ elements map to the set A. Note that $\tau(\varnothing) \subseteq \tau(\{1\} \subseteq \tau(\{1,2\}) \subseteq \ldots \subseteq \tau(\{1, 2, \ldots, n\}))$ are all initial sets.

Observation. For every permutation τ, there is at most 1 edge $A \in E$ such that A is initial for τ. This is because in a Sperner hypergraph, no edge is contained in the other.

Observation. Every $A \in E$ is initial for $|A|!(n-|A|)!$ permutations. This is because there are $|A|!$ many ways to map $\{1, 2, \ldots, |A|\}$ to A. There are then $(n-|A|)!$ ways to map the rest.

Let

$$\chi_{A,\tau} = \begin{cases} 1, & \text{if } A \text{ initial for } \tau \\ 0, & \text{otherwise} \end{cases}.$$

Then,

$$\underbrace{\sum_{\tau} \sum_{A \in E} \chi_{A,\tau}}_{\leq \sum_{\tau} = n!} = \underbrace{\sum_{A \in E} \sum_{\tau} \chi_{A,\tau}}_{\sum_{A \in E} |A|!(n-|A|)!}.$$

Together, we have that

$$n! \geq \sum_{A \in E} |A|!(n-|A|)!.$$

such that

$$1 \geq \sum_{A \in E} \frac{|A|!(n-|A|)!}{n!} = \frac{1}{\binom{n}{|A|}}.$$

■

And now with our proof for Sperner's theorem:

Theorem 15. Sperner's

Proof. Recall that

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} \geq \binom{n}{k}.$$

for any $1 \leq k \leq n$. Therefore,

$$\frac{|E|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = \sum_{A \in E} \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq \sum_{A \in E} \frac{1}{\binom{n}{|A|}}.$$

which must be at most 1 from LMY. Therefore, $|E| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. ■

Lecture 13: Extremal Graph Theory

Definition 26. If we have that G and G' are **isomorphic**, we write that

$$G \cong G'.$$

Definition 27. If $e \in E$, then the **ends** of e are the two vertices $v \in e$.

Definition 28. We say that a graph G' is a **subgraph** of G if $V' \subseteq V$ and $E' \subseteq E$.

If we fix a graph G', what is the maximum number of edges in an n-vertex graph such that G does not contain a subgraph $\cong G'$?

Example. If G' is the single edge, then the maximum number of edges in a graph that does not contain G' is 0.

Example. If G' is the line graph with 2 edges, then for every vertex v we have that $d_G(v) \leq 1$. By the handshaking lemma, we have that

$$2|E| \leq \sum_{v \in V} d_G(v) \leq |V|.$$

Therefore,

$$|E| \leq \left\lfloor \frac{n}{2} \right\rfloor.$$

Observation. Note that if $G'' \subseteq G'$, and $G' \subseteq G$, then $G'' \subseteq G$.

Definition 29. For any integer $r \geq 1$, the **complete graph** on r vertices, denoted K_r, has r vertices and $\binom{r}{2}$ edges.

We now look at the case where $G' = K_3$ is the triangle graph. One such example is the complete bipartite graph with $\lfloor \frac{n}{2} \rfloor$ vertices on one side and $\lceil \frac{n}{2} \rceil$ vertices on the other. this graph has $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil$ edges, which is actually the unique graph to the extremal problem.

Theorem 16. (Turan) Every n-vertex graph with no subgraph isomorphic to K_3 has at most

$$\left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil .$$

edges.

Proof. (Erdos) Let $G = (V, E)$ be an n-vertex graph with no subgraph $\cong K_3$. Let $v \in V$ be a vertex of maximum degree. Let $N_G(v) = \{x \in V : (v, x) \in E\}$. Let this set be called the **neighborhood** of v in G.

Let $\hat{G} = (\hat{V}, \hat{E})$ be the graph with vertex set $\hat{V} = V$ and edge set

$$\hat{E} = \{(x, y) : x \in N_G(v) \text{ and } y \notin N_G(v)\}.$$

This graph $\hat{G}$ has no subgraph $\cong K_3$. We claim that

$$|\hat{E}| \geq |E|.$$

Continuing,

$$\begin{aligned} |E| &\leq \sum_{y \notin N_G(v)} d_G(y) \\ &\leq \sum_{y \notin N_G(v)} d_G(v) \\ &= |\hat{E}|. \end{aligned}$$

Remember to maximize ab such that $a + b = n$, we set $a = \lfloor \frac{n}{2} \rfloor$ and $b = \lceil \frac{n}{2} \rceil$. Therefore, the maximum number of edges in our graph is $|\hat{E}| = ab \leq \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil$. ■

Lecture 14: Turan's Theorem

This graph is called the **Turan Graph** $T(n, 2)$.

Definition 30. $T(n, r)$ is the unique (up to isomorphism) n-vertex graph whose vertex set V can be partitioned into r disjoint parts $V_1, V_2, \ldots, V_r$ such that each part has roughly equal size $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$, and whose edge set is defined by

$$E = \{(u, v) : u \in V_i \text{ and } v \in V_j \text{ for some } i \neq j.\}.$$

Theorem 17. For any $n, r \in \mathbb{Z}_{>0}$, any n-vertex graph $G = (V, E)$ with no subgraph isomorphic to K_{r+1} has at most as many edges as the Turan graph $T(n, r)$. Furthermore, if this graph has the same amount of edges as the Turan graph, $T(n, r)$, then it is isomorphic to the Turan graph.

Proof. On HW, but can use the following: ■

Definition 31. A **complete multipartite graph** is a graph $G = (V, E)$ such that V has a partition into sets of vertices $V_1, \ldots V_n$ such that

$$E = \{(u, v) : u \in V_i \text{ and } v \in V_j \text{ for } i \neq j\}.$$

Denote each set $V_1, \ldots V_n$ as a **part**.

Lemma 23. Let $n, r \in Z_{>0}$, and let $G = (V, E)$ be an n-vertex complete multipartite graph with $\leq r$ parts. Then G has at most as many edges as $T(n, r)$, and is $T(n, r)$ if $=$ holds.

Proof. Suppose G has ≥ 2 vertices in its largest part than in its smallest part. We will show that $|E| <$ the number of edges in $T(n, r)$.

While there are ≥ 2 more vertices in the largest part of G_i than in its smallest part, let G_{i+1} be the complete multipartite graph formed from moving one such vertex from the smallest part to the largest part of G_i. ■

Lemma 24. This process will end, and when it does, that graph is isomorphic to $T(n, r)$.

Proof. HW. ■

Lemma 25. The number of edges in G_i is strictly less than the number of edges in G_{i+1}.

Proof. Let v be the vertex of G_i which was moved. Then the number of edges of G_i, which are not incident to v, is the number of edges in G_{i+1} that are not incident to v. Also, we have that

$$d_{G_i}(v) = n - \text{vertices in part with } v.$$

and

$$d_{G_{(i+1)}}(v) = n - \text{vertices in new part with } v.$$

Therefore, the number of edges in G_{i+1} is strictly greater than that of G_i. ■

Lecture 15: Kovori-Sos-Turan

Definition 32. A graph G is **bipartite** if $G \subseteq T(n,2)$ for some n. Equivalently, G is **bipartite** if its vertex set can be paritioned into two sets A and B such that every edge of G has one end in A and one end in B. The two sets A and B are called a **bipartition** of the graph.

Theorem 18. (Kovori-Sos-Turan) Let $G' = (V', E')$ be a bipartite graph. Then there exists $\varepsilon > 0$ (which depends on G') such that every n-vertex graph with no subgraph isomorphic $\cong G'$ has $O(n^{2-\varepsilon})$ edges. In fact, we can take

$$\varepsilon = \frac{1}{\text{size of larger part of bipartition}}.$$

Theorem 19. (Special case of KST) If G' is the complete bipartite graph with bipartition (A, B) such that $|A| = 2$ and $|B| = 2$, then every n-vertex graph with no subgraph $\cong G'$ has

$$O(n^{\frac{3}{2}}).$$

Note that G' is also the 4-cycle C_4.

Note. The same for $K_{3,3}$ is an open problem, known as the **Zarankeiwicz Problem**.

We will now look at a construction for the lower bound.

Definition 33. In its essence, a **projective plane** is a system of $q^2 + q + 1$ lines and ponts such that every pair of points determines a line, every pair of lines determines a unique point, and every line contains $q + 1$ points.

Explanation. Consider the (biparite) incidence graph of a projective plane such that each point is connected to the lines that it is in. Formally,

$$V = \{\text{points}\} \cup \{\text{lines}\}$$

and

$$E = \{\{x, y\} : x \text{ is a point in line } y\}.$$

Observe that G has no subgraph isomorphic to C_4. Also note that the number of edges is the number of lines times the number of points in each line, which is

$$(q^2 + q + 1)(q + 1) \geq q^3.$$

Also note that the number of vertices is $2(q^2 + q + 1) = O(q^2)$. Therefore,

$$\text{number edges} = \Omega(\text{number vertices}^{1.5}).$$

Lecture 16: Graphs with no 4-Cycles

Theorem 20. (Cauchy-Schwarz) Given $x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $y_1, y_2, \ldots, y_n \in \mathbb{R}$,

$$\sum_{i=1}^{n} x_i y_i \leq \sqrt{\sum_{i=1}^{n} x_i^2} \cdot \sqrt{\sum_{y=1}^{n} y_i^2}.$$

We will now prove the special case of the Kovori-Sos-Turan theorem.

Proof. We count the number of subgraphs isomorphic to cherries such that v is red and the two green star vertices are $\in N_G(v)$. One way to count such cherries is choose every two vertices in the neighborhood of G, which is $\sum_{v \in V} \binom{d_G(V)}{2}$.

The second way is to first choose the two green star vertices, which is just $\binom{n}{2}$. Then, there is at most one choice for the red vertex as there is no subgraph isomorphic to C_4. This is therefore at most $\binom{n}{2}$.

Putting this all together, we have

$$\sum_{v \in V} \binom{d_G(v)}{2} \leq \binom{n}{2} \leq \frac{n^2}{2}.$$

Let the vertex set be $\{1, 2, \ldots, n\}$. Apply the Cauchy-Schwarz inequality with all $y_i = 1$, and

$x_i = d_G(i) - 1$. Then,

$$\begin{aligned}\sum_{i=1}^n x_i &\le \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n 1} \\ &= \sqrt{\sum_{i=1}^n (d_G(i)-1)^2} \cdot \sqrt{n} \\ &\le \sqrt{2\sum_{v\in V}\binom{d_G(v)}{2}} \cdot \sqrt{n}.\end{aligned}$$

All in all,

$$\begin{aligned}2|E| - n &\le \sqrt{2\sum_{v\in V}\binom{d_G(2)}{2}} \cdot \sqrt{n} \\ \sum_{v\in V}\binom{d_G(v)}{2} &\le \frac{n^2}{2} \\ 2|E| - n &\le \sqrt{n^2}\cdot\sqrt{n} \\ &= n^{\frac{3}{2}}.\end{aligned}$$

Such that

$$|E| \le \frac{n^{\frac{3}{2}} + n}{2} = O(n^{\frac{3}{2}}).$$

■

Theorem 21. (Szemeredi-Trotter) For incidences for points on lines in the plane in $\mathbb{R}^2$, the number of incidences is $O(n^{\frac{4}{3}})$.

Lecture 17: Blanche Descartes

Definition 34. The **chromatic number** of a graph is the minimum integer k such that $G \subseteq$ a complete multipartite graph with $\le k$ parts.

Theorem 22. (Tutte) For every $k \in \mathbb{Z}_{>0}$, there exists a graph $G = (V, E)$ with no subgraph $\cong K_3$ such that G is not a subgraph of any complete multipartite graph with $\le k$ parts.

Lemma 26. The minimum **chromatic number** of G is the minimum integer k such that there exists $\phi : V \to \{1, 2, \dots, k\}$ such that the ends of every edge have different colors.

Definition 35. Another way to define the **chromatic number** is $\chi(G)$. Any function $\phi : V \to \{1, \dots, k\}$ such that for all $(u, v) \in E$, $\phi(u) \neq \phi(v)$ is called a **proper k-coloring**.

We now show the proof of the theorem:

Proof. We proceed with induction on k. Formally, set $G_{k+1} = (V_{k+1}, E_{k+1})$ where

$$V_{k+1} = V \cup \underbrace{\{\hat{a} : a \in V_k\}}_{\hat{V}_k:} \cup \{v\}$$

and

$$\begin{aligned}E_{k+1} = E_k &\cup \{\{a, \hat{b}\} : \{a, b\} \in E_k\} \\ &\cup \{\{\hat{a}, v\} : a \in V_k\}.\end{aligned}$$

We now show that there is no $k+1$ coloring of a graph. Suppose there is for the sake of contradiction. WLOG assume $\phi(v) = k+1$. Define $\phi' : G_k \to \{1, \dots, k\}$ by setting, for each $a \in V_k$,

$$\phi'(a) = \begin{cases}\phi(a) & \text{if } \phi(a) \in \{1, \dots, k\} \\ \phi(\hat{a}) & \text{if } \phi(a) = k+1\end{cases}.$$

Note that $\phi'(a) \in \{1, \dots, k\}$. This is a proper k coloring, a contradiction. ■

Lecture 18: Triangle-Free Graphs

Definition 36. The n-sphere S_n is $\{(x_1, x_2, \dots, x_{n+1}) \in \mathbb{R}^{n+1} : x_1^2 + x_2^2 + \dots + x_{n+1}^2 = 1\}$.

Note that S_2 is a ball, like a soccer ball.

Definition 37. A set $X \subseteq S_n$ is **open** if for every $x \in X$, there exists $\varepsilon > 0$ such that all points $y \in S_n$ with $dist(x, y) < \varepsilon$ are in X.

Definition 38. A set $X \subseteq S_n$ is **closed** if $S_n \setminus X$ is open.

Lemma 27. Any union of open sets is open.

Proof. Omitted. ■

Theorem 23. (Bosuk-Ulam) Let $X_1, X_2, \dots, X_{n+1} \subseteq S_n$ such that

$$\bigcup_{i=1}^{n+1} X_i = S_n.$$

and each of X_i is either open or closed. Then there exists $i \in \{1, 2, \dots, n+1\}$ and $x \in S_n$ such that X_i contains both x and $-x$. These points are called **antipodal**.

Definition 39. The **Kneser** graph $G_{n,k} = (V_{n,k}, E_{n,k})$ has a vertex set

$$V_{n,k} = \{I \subseteq \{1, \ldots, n\} : |I| = k\}.$$

and

$$E_{n,k} = \{\{A, B\} : A \cap B = \varnothing\}.$$

Lemma 28. The graph $G_{3k-1,k}$ has no subgraph $\cong K_3$.

Proof. There are no 3 disjoint sets $A, B, C \in \{1, 2, \ldots, 3k-1\}$ of size k. ∎

Theorem 24. (Lovasz-Kneser) The chromatic number of $G_{3k-1,k}$ is $> k$.

Proof. Suppose not. Then ϕ is a proper k-coloring of $G_{3k-1,k}$. Let $P \subseteq S_k$ be a set of $3k-1$ points such that no $k+1$ points in P lie on a hyperplane through the origin. (For any vector $x \in \mathbb{R}^{k+1}$), the set of $\{y \in \mathbb{R}^{k+1} : x \cdot y = 0\}$ is a hyperplane through the origin).

For each color i, let X_i be the set of all $x \in S_k$. such that there exists a vertex $A \in V_{3k-1,k}$ such that $\phi(A) = i$, and $f(A)$ lies in the hemisphere of S_k centered at x.

It is possible to show that $X_1, X_2, \ldots, X_k$ are open, and that $S_k \setminus \bigcup_{i=1}^{k} X_i$ is closed. So by the theorem, there exists antipodal points $x, -x$ which are contained in the same one of X_i. If the the last set contains these two points, then there are two open hemispheres, both containing less than k points, which means that there are $k+1$ points contained in a hyperplane. If some X_i has antipodal points, then two disjoint sets have the same color, which is a contradiction. ∎