Hypergraphs, Extremal Set Theory: Math 454 Lecture 18 (7/27/2017)

Cole Franks

July 31, 2017

Keller and Trotter Extremal Set Theory Chapter:

http://www.rellek.net/book/s_kitchensink_extremal-set-theory.html

Contents

1 The Boolean Lattice

1.1 Sperner's Theorem

Question 1. Suppose the UN had 10 official languages (which would be true if the four proposed new languages became official). Say two interpreters are redundant if one can speak all the languages the other speaks. How many interpreters can the UN have if no two interpreters are redundant?

We can state this language in terms of families of sets. Let A_i be the set of languages spoken by interpreter i; where $i \in [m]$ if there are m interpreters. Then a pair of redundant interpreters is a pair $A_i \subset A_j$. So the question becomes:

Question 2. What is the largest family of subsets of [10] such that no set in the family is contained within another? Equivalently, what is the size of the largest antichain in the Boolean lattice?

One idea is just to let each interpreter speak some 5-subset of the languages. These certainly don't overlap, and this way we can have $\binom{10}{5} = 252$ interpreters. Is this the largest number? It turns out to be so. The "middle level" of $\mathcal{B}_n$ is the largest antichain.

Theorem 1.1 (Sperner's Lemma)**.** *If $\mathcal{F}$ is a family of subsets of $[n]$ such that there are no distinct subsets $A, B \in \mathcal{F}$ satisfying $A \subset B$, then*

$$|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

In other words, the largest antichain of $\mathcal{B}_n$ is of size

$$\binom{n}{\lfloor n/2 \rfloor},$$

formed by all subsets of $[n]$ of size $\lfloor n/2 \rfloor$.

Though the statement has nothing to do with probability, there is a beautiful probabilistic proof.

Proof. Suppose $\mathcal{F}$ is an antichain in $\mathcal{B}_n$. It is convenient to break $\mathcal{F}$ up by cardinality. If $i \in \{0, \ldots n\}$, define

$$\mathcal{F}_i = \{F \in \mathcal{F} : |F| = i\}.$$

Choose a random permutation σ of $[n]$.

Next, we cook up a random variable X. Pick a random permutation σ, and use it to define a chain

$$C = \emptyset \subset \{\sigma_1\} \subset \{\sigma_1, \sigma_2\} \subset \cdots \subset \{\sigma_1, \ldots, \sigma_{n-1}\} \subset \{\sigma_1, \ldots, \sigma_{n-1}\}.$$

C is the chain of numbers contained in prefixes of σ. Now let

$$X = |C \cap \mathcal{F}| = |\{i : \sigma_1, \ldots, \sigma_i\} \in \mathcal{F}|.$$

In other words, it is the number of "prefixes" of the permutation whose entries form a member of $\mathcal{F}$. $\mathcal{F}$ is an antichain and C is a chain, so by Fact **??**,

$$X \leq 1 \textit{ always.}$$

In particular,

$$\mathbb{E}[X] \leq 1.$$

On the other hand, $X = \sum_{i=0}^{n} X_i$, where X_i is the random variable that is 1 if $\{\sigma_1, \ldots, \sigma_i\} \in \mathcal{F}$ and 0 otherwise, so

$$\mathbb{E}[X] = \sum_{i=0}^{n} \mathbb{E}[X_i].$$

Since $\{\sigma_1, \ldots, \sigma_i\}$ is just a random i-subset of $[n]$,

$$\mathbb{E}[X_i] = \frac{|\mathcal{F}_i|}{\binom{n}{i}}.$$

By linearity of expectation,

$$\sum_{i=0}^{n} \frac{|\mathcal{F}_i|}{\binom{n}{i}} = \mathbb{E}[X] \leq 1.$$

Since $\binom{n}{i} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$,

$$\frac{|\mathcal{F}|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = \sum_{i=0}^{n} \frac{|\mathcal{F}_i|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq \sum_{i=0}^{n} \frac{|\mathcal{F}_i|}{\binom{n}{i}} \leq 1,$$

which implies

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

□

Corollary 1.1. *$\mathcal{B}_n$ can be partitioned into $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ chains.*

2 Hypergraphs

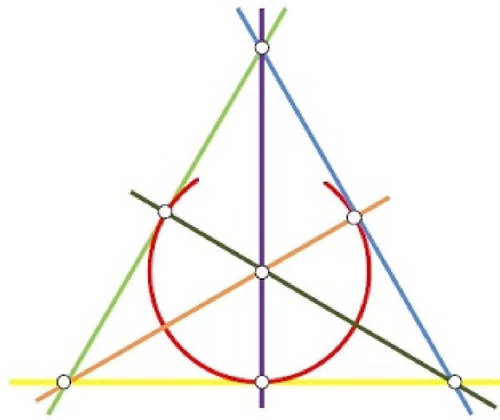
Hypergraphs are really just families of sets, which we have already talked about a great deal.

2.1 Definitions

Definition 2.1. A *hypergraph* $H = (V, E)$ is a pair of sets V, called the *vertex set*, and E, called the *edge set*. E is a family of subsets of V. Alternately, E is a subset of the power set of V, so $E \subset 2^V$. The elements of the edge set are called the *edges* of H. The *degree* $d(v)$ of a vertex v is the number of edges containing v. The *size* of H is $|E|$, the number of edges, and the *order* of H is $|V|$, the number of vertices.

Definition 2.2. A hypergraph $H = (V, E)$ is said to be k-uniform if every edge in E has the same cardinality k. H is said to be r-regular if every vertex has degree r.

Example 1. Consider the following hypergraph H, where the vertices are the white circles and each colored line is an edge (including the curved one).



This is a 3-uniform hypergraph.

Definition 2.3. We can represent a hypergraph H as an $m \times n$ matrix M with 0-1 entries, called the *incidence matrix*. If V is numbered $\{1, \ldots, n\}$ and E is numbered $\{1, \ldots, m\}$, then the incidence matrix is given by

$$M_{ij} = \begin{cases} 1: & \text{edge } i \text{ contains vertex } j \\ 0: & \text{otherwise.} \end{cases} \quad \text{for } i \in [m], j \in [n].$$

That is, edges label the rows and vertices label the columns, and there is a 1 in the i^{th} row and j^{th} column of M if the i^{th} edge contains the j^{th} vertex.

Example 2. If $H = ([4], \{\{1,2,3\}, \{2,3,4\}, \{3,4,1\}, \{4,1,2\}\})$, then

$$M = \begin{array}{cc} & \text{vertices} \\ \text{edges} & \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \end{array}.$$

Example 3. Every graph G is a 2-uniform hypergraph. The notion of degree in G as a hypergraph is exactly the same as the notion of degree in G as a graph.

Theorem 2.1. *If $H = (V, E)$ is a finite, k-uniform hypergraph, then*

$$\frac{\sum_{v \in V} d(v)}{k} = |E|.$$

Proof. There are multiple ways to prove this; one is by counting the sum of the entries of the incidence matrix in two different ways by summing either each column first or by summing each row first. □

Remark 2.1. Why not just call them set families? One uses the term hypergraph when trying to answer graph-like questions such as connectivity or talk about degrees, and set family when one wants to answer set theory type questions. We'll abuse notation and talk about hypergraphs even though we are answering set theory questions.

2.2 Extremal set theory: How large can a hypergraph be that does...

Theorem 2.2 (Sperner)**.** *A hypergraph on n vertices with the property that no edge is contained in another has at most $\binom{n}{\lfloor n/2 \rfloor}$ vertices.*

Just as a sanity check, make sure you know the size of the largest hypergraph with n vertices.

An *intersecting* hypergraph $H = (V, E)$ is one such that if $A \cap B \neq \emptyset$ for all $A, B \in E$. Suppose $H = ([n], E)$ is an intersecting hypergraph. How large can the size of H be? One thing we could try is letting E be all edges containing 1. Since there are 2^{n-1} such edges, we can find an hypergraph of size 2^{n-1} on n vertices.

Theorem 2.3. *An intersecting hypergraph on n vertices has at most 2^{n-1} edges.*

Proof. Suppose $H = ([n], E)$ is an intersecting hypergraph. Split $2^{[n]}$, the set of all subsets of $[n]$, into pairs $\{A, [n] \setminus A\}$. No subset appears in more than one pair. Since the union of these pairs is 2^n and they have no subsets in common, there must be 2^{n-1} pairs. However, A and $[n] \setminus A$ are always disjoint, so each of the pairs can contain at most one element of E. Thus, $|E| \leq 2^{n-1}$. □

The following proof wasn't covered in class. I may cover it Monday. If you like reading ahead, here's a rough draft.

Less than $2k$ case, at least $2k$ case. Lower bounds for each.

Theorem 2.4 (Erdős, Ko, Rado)**.** *If $n \geq 2k$, a k-uniform intersecting hypergraph on n vertices has at most*

$$\binom{n-1}{k-1}$$

edges.

This beautiful proof is due to Katona. First we need a little Lemma.

Lemma 2.1. *Suppose $H = ([n], E)$ is an intersecting, k-uniform hypergraph, and define*

$$A_s = \{s, \dots, s+k-1\}$$

for $0 \leq s \leq n-1$, where addition is understood to be $\pmod n$ that is, these "windows" wrap around. Then at most k of the sets A_s can be edges of H.

Proof of Lemma. Suppose $A_t \in E$. Then the rest of the A_s that intersect A_s can be split up into $k-1$ disjoint pairs; namely

$$A_{i-k}, A_i$$

for $i \in \{s+1, \dots, s+k-1\}$. Thus, at most one of each of these $k-1$ disjoint pairs can be an edge of H, and so at most k of the A_s are edges of H. □

Proof of EKR. Arrange the elements of $[n]$ on a circle at random, and let A be the set starting at σ_1. Consider the random variable X that is the expected number of windows that are edges of H. We know that the number is *always* at most k. $\mathbb{E}[X] \leq k$. On the other hand, each window is a random k-subset of $[n]$. So, if we let X_i be the indicator random variable that the i^{th} window is an edge of H (that is, 1 if it is an edge of H, 0 otherwise), then

$$\mathbb{E}[X_i] = \Pr[X_i = 1] = \frac{|E|}{\binom{n}{k}}.$$

Now

$$k \geq \mathbb{E}[X] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = n\frac{|E|}{\binom{n}{k}},$$

and so

$$|E| \leq \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}.$$

□