Lec 24-25 Extremal Set Theory

Problem: Given a family of k-sets of $[n]$, when are the t-sets contained in those k-sets minimized?

Definition: a *k-uniform* family is a family of k-sets. The *t-shadow* of a set system $\mathcal{F}$ is the family of all t-sets contained in members of $\mathcal{F}$. The *shadow* $\partial\mathcal{F}$ of a k-uniform family $\mathcal{F}$ is its $(k-1)$-shadow. The *shade* is the family of all $(k+1)$-sets that contain members of $\mathcal{F}$.

In the language of shadow, we want to find the family with the smallest shadow, among all k-uniform families of size m.

Lem: k-sets can be indexed, and can also be bijectively mapped to binary k-words.

Colex ordering: a colex ordering on a family of k-sets is obtained by putting $x < y$ if $x_i < y_i$ in the highest coordinate where their binary incidence vector differ.

Example: the lexicographic order of $\binom{[N]}{3}$ is $123, 124, 125, 126, \ldots, 134, 135, 136, \ldots, 234, 235, \ldots$; the colex ordering for $\binom{[5]}{3}$ is: 123, 124, 134, 234, 125, 135, 145, 235, 245, 345.

Lemma: If the vector with index m, where $m \geq 1$, in the colex ordering on $\binom{[n]}{k}$ has 1s in position $m_1, m_2, \ldots, m_k$, then

$$m = \binom{m_k - 1}{k} + \binom{m_{k-1} - 1}{k-1} + \ldots + \binom{m_1 - 1}{1} + 1.$$

Proof: Let σ be the vector with index m. To reach σ, we must skip all vectors whose kth 1 appear before position m_k, and there are $\binom{m_k-1}{k}$ of these. In addition, some vectors with last 1 in position m_k precede σ, and their first $k-1$ 1s precede position m_{k-1}, and there are $\binom{m_{k-1}-1}{k-1}$ of these. Continuing this procedure.

Definition: (*k-binary expansion of m*) For given k, each position integer m can be expressed in the form $\binom{m_k}{k} + \binom{m_{k-1}}{k-1} + \ldots + \binom{m_i}{i}$ with $m_k > m_{k-1} > \ldots > m_i \geq i$.

Lemma: The shadow of the first m vectors in the colex order on $\binom{[n]}{k}$ consists of the first $\partial_k(m) = \binom{m_k}{k-1} + \binom{m_{k-1}}{k-2} + \ldots + \binom{m_i}{i-1}$ vectors in the colex order on $\binom{[n]}{k-1}$.

The Kruskal-Katona Theorem: The shadow of a family of m elements of $\binom{[n]}{k}$ is minimized by the family consisting of the first m elements in the colex ordering on $\binom{[n]}{k}$. Furthermore, the size of the shadow is $\partial_k(m)$.

Proof: let $\mathcal{F}$ be a set of m elements in $\binom{[n]}{k}$. The *compression* of $\mathcal{F}$ is the set $C\mathcal{F}$ consisting of the first $|\mathcal{F}|$ elements in the colex ordering on $\binom{[n]}{k}$. The idea is to show that $|\partial(C\mathcal{F})| \leq |\partial\mathcal{F}|$ when $\mathcal{F} \subset \binom{[n]}{k}$.

Problems: what is the maximum size of a family of sets in which no member contains another (antichain)?

Definition: an *antichain* of sets is a family of sets in which no member contains another.

Theorem (LYM inequality) Let $\mathcal{F}$ be an antichain on $[n]$. Let $\mathcal{F}_k = \mathcal{F} \cap \binom{[n]}{k}$ and $a_k = |\mathcal{F}_k|$. Then $\sum_k \frac{a_k}{\binom{n}{k}} \leq 1$.

Proof: Counts the permutations of X in two different ways. First, by counting all permutations of X directly ($n!$). But secondly, one can generate a permutation (i.e., an order) of the elements of X by selecting a set S in A and concatenating a permutation of the elements of S with a permutation of the nonmembers (elements of $X - S$). If $|S| = k$, it will be associated in this way with $k!(n-k)!$ permutations, and in each of them the first k elements will be just the elements of S. Each permutation can only be associated with a single set in A, for if two prefixes of a permutation both formed sets in A then one would be a subset of the other. Therefore, the number of permutations that can be generated by this procedure is $\sum_{S \in A} |S|!(n - |S|)! = \sum_k a_k k!(n-k)! \leq n!$. It follows that $\sum_k \frac{a_k}{\binom{n}{k}} \leq 1$.

Proof: by using probabilistic method. Choose a maximal chain C uniformly random.....

Theorem: (Sperner) The maximum size of an antichain of subsets of $[n]$ is $\binom{n}{\lfloor n/2 \rfloor}$, achieved only by antichains whose sets all have the same size.

Proof (using LYM inequality): By LYM inequality, $1 \geq \sum_k \frac{a_k}{\binom{n}{k}} \geq \sum_k \frac{a_k}{\binom{n}{\lfloor n/2 \rfloor}} = \frac{|F|}{\binom{n}{\lfloor n/2 \rfloor}}$.

Problems: what is the maximum size of a family of sets in which no member contains another (antichain) and is also required to be pairwise intersecting?

Definition: An *t-intersecting family* is a family in which every two sets have at least t common elements. A *star* is a family of sets having a universal common element; a *t-star* is a family sharing t universal common elements.

Example: an intersecting family of subsets of [n] has size at most 2^{n-1}.

An other maximum intersecting family consists of all sets with more than half the elements, plus (when n is even) the sets of size $n/2$ containing a particular element.

Definition: An $EKR(k,t)$-family is an antichain $\mathcal{F}$ that is also a t-intersecting family in which the size of each member is at most k.

Theorem: (Erdos-Ko-Rado, t=1) For $n \geq 2k$, the maximum size of an $EKR(k,1)$-family is $\binom{n-1}{k-1}$, achieved by a star in $\binom{[n]}{k}$.

Proof. Let $\mathcal{F}$ be such a family. We may assume that $\mathcal{F} \subseteq \binom{[n]}{k}$.

Given a circular arrangement σ of $[n]$, we ask how many members of $\mathcal{F}$ can occur in σ as a consecutive string of elements. For such a string x, every consecutive k-set that intersects x has a boundary at one of the $k-1$ locations between elements of x. Hence at most $k-1$ members of $\mathcal{F}$ other than x occur consecutively in σ.

Summing this over all $(n-1)!$ circular permutations yields at most $(n-1)!k$ appearances of members of $\mathcal{F}$. Each members appears consecutively in $k!(n-k)!$ circular permutations. Thus $|\mathcal{F}| \leq \frac{(n-1)!k}{k!(n-k)!} = \binom{n-1}{k-1}$.

Theorem: (Erdos-Ko-Rado) For n sufficiently large, a t-star of k-sets forms a maximum $EKR(k,t)$-family.

Sketch of the proof: We assume that $\mathcal{F}$ is a t-intersecting family of k-sets. We push members of $\mathcal{F}$ toward sets containing $[t]$ by using "shift operator" $\tau_{i,j}$. For $i < j$ and $x \in \mathcal{F}$, define $\tau_{i,j}(x)$ by

$$\tau_{i,j}(x) = \begin{cases} x - j + i, \text{if } j \in x \text{ and } i \notin x \text{ and } x - j + i \notin \mathcal{F} \\ x, \text{otherwise.} \end{cases}$$

Let $\tau_{i,j}(\mathcal{F}) = \{\tau_{i,j}(x) : x \in \mathcal{F}\}$. Note that $|\tau_{i,j}(\mathcal{F})| = |\mathcal{F}|$. We can verify that $\tau_{i,j}$ preserves the t-intersection property and study the form of a family unchanged by these operators.

Remark: Frankl and Wilson showed that the t-star of k-sets is optimal when $n \geq (t+1)(k-t+1)$. For smaller n, other families are larger.