Topics In Combinatorics

Lecture 1

1 Averaging Arguments

If you've got a random variable that takes real variables, then with positive probability it will be at least as big as its average, and similarly at least as small:

Theorem 1.1: *Let X be a random variable. Then $\mathbb{P}[X \geq \mathbb{E}X] > 0$.*

When X is discrete, this result is almost trivial, but in the general (continuous) case it isn't *quite* as trivial.

Proof. Suppose that $\mathbb{P}[X \geq \mathbb{E}X] = 0$. The tempting idea here is to say that then X is always strictly less than the average, so when you take the average it's still strictly less than the average - we need to be careful about making this work:

Define $P_n = \mathbb{P}[\mathbb{E}X - \frac{1}{n} < X \leq \mathbb{E}X - \frac{1}{n+1}]$ - with P_0 denoting $-\infty$ on the left.

It is then the case that $\sum_{n=0}^{\infty} P_n = 1$, so $\exists n : P_n > 0$. But then

$$\mathbb{E}X \leq \sum_{n=0}^{\infty} P_n \left(\mathbb{E}X - \frac{1}{n+1} \right) = \mathbb{E}X - \sum_{n=0}^{\infty} \frac{P_n}{n+1} < \mathbb{E}X \quad \perp$$

Similarly with the other case. □

We won't use this case much, but it's fun to see!

The really surprising thing is that this extremely basic fact is also extremely useful. The way we use it (in the discrete case) is to simply deduce that such an event is possible.

Question. How many edges does an icosahedron have?

Perhaps this seems a little tedious - but there is a trick we can use.

We know the icosahedron has 20 faces, and that these faces are triangles. We then reason that each face has three edges, and each edge is part of two faces. That is to say, $2E = 3F$, so $E = 3F/2 = 30$.

The idea here is that both $2E$ and $3F$ are counting something, namely edge-face pairs.

Consider another example - let G be a bipartite graph, with vertex sets X and Y. Let's suppose that we have the regularity conditions:

- $(\forall x \in X) d(x) = d_1$
- $(\forall y \in Y) d(y) = d_2$

Then counting the edges from the perspectives of X and Y, we have that $|E(G)| = d_1|X| = d_2|Y|$. This is simply an abstraction of the above result.

Moreover, if we instead have

- $(\forall x \in X) d(x) \leq d_1$
- $(\forall y \in Y) d(y) \geq d_2$

Then $d_2|Y| \leq |E(G)| \leq d_1|X|$, so $|Y| \leq d_1|X|/d_2$. We can apply this in many ways, such as:

Let $[n] := \{1, 2, \cdots, n\}$, and $[n]^{(r)}$ be the collection of subsets of $[n]$ or size r.

Let $\mathcal{A} \subset [n]^{(r)}$ (this is an *r-uniform hypergraph*), and let $s > r$. The ***s-upper shadow*** of $\mathcal{A}$ is

$$\partial^s \mathcal{A} = \{B \in [n]^{(s)} : \exists A \in \mathcal{A} \text{ s.t. } A \subset B\}$$

Then join $A \in \mathcal{A}$ to $B \in \mathcal{B} = \partial^s \mathcal{A}$ iff $A \subset B$. We then have:

$$d(A) = \binom{n-r}{s-r} \quad \forall A \in \mathcal{A}$$
$$d(B) \leq \binom{s}{r} \quad \forall B \in \mathcal{B}$$

Then $\mathcal{A}\binom{n-r}{s-r} = E \leq \binom{s}{r}|\mathcal{B}|$. Hence $|\mathcal{B}| \leq |\mathcal{A}|\binom{n-r}{s-r}/\binom{s}{r} = |\mathcal{A}|\binom{n}{s}/\binom{n}{r}$.

We remark that this is not the tightest known bound - for an improvement see *Kruskal-Katona*.

We have a few more results on graphs.

Recall that a ***tree*** is a connected, acyclic graph. A basic fact about trees on n vertices is that they have $n-1$ edges. We will see why this is true now.

Firstly, we note that every tree has a vertex of degree one (a *leaf*); if not, then we start at a vertex v_1, find a neighbour v_2, then a neighbour of v_2 that is v_3, and keep going until we run out of new vertices. Then we must return to a previous v_i, forming a cycle.

Then to see that every tree has $n-1$ edges, we consider a tree on n vertices and remove a leaf. This cannot disconnect the graph and cannot create a cycle, so we are left with a tree of degree $n-1$, which by induction has $n-2$ edges. Thus the original tree has $n-1$ edges.

Theorem: (Euler's Formula) *If G is a connected planar graph with V vertices, E edges and F faces, then $V - E + F = 2$.*

Note that a *face* of a graph is just a component of the complement of the drawing of the graph in the plane. We prove this through a slightly unusual induction.

Proof. Base case: G is a tree. Then we have $V-1$ edges, and G is acyclic so we have only one face. Hence $V - E + F = V - (V-1) + 1 = 2$.

If G is not a tree, then G contains a cycle. We then remove an edge from a cycle. The graph remains connected, V stays the same, E decreases by 1, and F decreases by 1 since we have joined two components of the plane. Hence we obtain smaller G' with $V' - E' + F' = V - E + F$. Continue until we are left with a tree. □

Corollary: $E \leq 3V - 6$. This is since each edge is in ≤ 2 faces, and each face contains ≥ 3 edges. So $2E \geq 3F$, and hence $2 = V - E + F \leq V - E/3$. Hence $E \leq 3V - 6$.

2 Intersecting Families

Suppose we have some $\mathcal{A} \subset \mathcal{P}[n]$ with $A, B \in \mathcal{A} \implies A \cap B \neq \emptyset$. How big can $\mathcal{A}$ be?

Note that if $A \in \mathcal{A}$ then $A^c \notin \mathcal{A}$, hence $|\mathcal{A}| \leq 2^{n-1}$. Moreover, we can take $\mathcal{A} = \{A : 1 \in A\}$, so then $|\mathcal{A}| = 2^{n-1}$ and equality is achieved. So the problem looks pretty solved, but we can also ask what happens in the equality case - if we have an intersecting family of size 2^{n-1}, must it be of the above form? The answer is no:

If n is odd, we can take $\mathcal{A} = \{A \subset [n] : |A| > n/2\}$. Then for each $A \in [n]$, either $A \in \mathcal{A}$ or $A^c \in \mathcal{A}$, so $|\mathcal{A}| = 2^{n-1}$ and $\mathcal{A}$ is not of the above form.

If instead n is even, we can take all sets of size $> n/2$, and from the sets of size $n/2$ take exactly one from each complementary pair (A, A^c). This gives us exactly half of $\mathcal{P}[n]$, and it is still an intersecting family since the only way we could have an empty intersection among the size $n/2$ sets would be with a complementary pair.

Another example is $\mathcal{A} = \{A \subset [n] : |A \cap \{1, 2, 3\}| \geq 2\}$. Since any two pairs of size 2 subsets of $\{1, 2, 3\}$ intersect non-trivially, this is indeed an intersecting family. Moreover if $A \notin \mathcal{A}$ then A intersects $\{1, 2, 3\}$ at most once, so its complement does so at least twice and $A^c \in \mathcal{A}$ so this is indeed half of all the sets.

More generally, we could take some intersecting family that we want, in $\mathcal{P}[m] \subset [n]$, and extend all subsets in that family to get a maximal family that has half of the sets.

Hopefully it is now clear that this is no unique way in which $\mathcal{A}$ can be an intersecting family.

What if instead all sets have size k?

Let $\mathcal{A} \subset [n]^{(k)}$ be an intersecting family. How big can $\mathcal{A}$ be?

Boring case: when $k > n/2$, then any two sets intersect, so we can take all $\binom{n}{k}$ of them.

Mildly interesting case: if $k = n/2$, then since we can't pick any two from an intersecting pair we can only choose at most half of them - but as we have seen this is sufficient to create an intersecting family, so we can have $\frac{1}{2}\binom{n}{k}$.

Interesting case: what if $k < n/2$. First we will find a *lower* bound, then an upper bound.

If we take all sets containing x, then $|\mathcal{A}| = \binom{n-1}{k-1}$ - so this is certainly attainable.

Theorem: (Erdos-Ko-Rado Theorem) *IF $k < n/2$ and $\mathcal{A}$ is an intersecting family of k-subsets of $[n]$, then $|\mathcal{A}| \leq \binom{n-1}{k-1}$.*

Moreover, if $|\mathcal{A}| = \binom{n-1}{k-1}$, then $\exists x$ such that $\mathcal{A} = \{A \in [n]^{(k)} : x \in A\}$.

The proof given here is not the original one, but is a much simpler one due to Katona.

Proof. Let $\mathcal{A}$ be an intersecting family of k-sets in $[n]$. Pick a random k-set as follows.

First, choose a random cyclic order $x_1, \ldots, x_n$ of $\{1, \ldots, n\}$ - this is a permutation whereby the indices are given modulo n, *i.e.* $x_1 < x_2 < \cdots < x_n < x_1$.

We then define an interval or arc in this order to be a collection of elements that are next to each other. Pick a random interval of length k in this cyclic order.

The probability that this interval lies in $\mathcal{A}$ is simply $|\mathcal{A}|/\binom{n}{k}$, since the k-sets chosen in this way are uniformly distributed.

BUT How many intervals can be in $\mathcal{A}$? Consider one such interval $x_1|x_2|x_3|\dots|x_k \in \mathcal{A}$. Then any other interval in $\mathcal{A}$ must start before x_1 and end inside the interval, or start after x_k and end inside the interval. However, given a particular starting line as drawn above, we cannot have both the interval going left from it and going right from it as this pair is not intersecting. Since all the possible intersecting intervals arise in this way, we can have at most $k-1$ additional intervals, *i.e.* the number of dividing lines. So including the original interval we have in total $\leq k$ of these intervals.

Moreover, for the equality case, if we want all possible intervals we have to have, for each dividing line, either the interval ending on the left or starting on the right. In fact, if we have one going left from a given dividing line, then we can't have one going right from the next line, since $n > 2k$ (the intervals can't meet round the back).

So we have intervals to the right up to a certain point, and then left from that point onwards: $|\to|\to\dots|\to x\leftarrow|\leftarrow|\cdots\leftarrow|$. And we see that all the intervals contain the shown x where directions switch. This is what we want for this specific cyclic order, but we are not yet done.

So we can only have k intervals in any given cyclic order, and there are only n cyclic intervals of length k.

Note that we initially calculated the probability by first choosing a random cyclic order, and then choosing a random interval from it. If we do this the other way round, then given a cyclic order the probability we end up in the set is $\leq k/n$, but the cyclic orders are uniform across the k-sets.

So, given a set $A = \{x_1, \dots, x_k\}$, we have $\mathbb{P}(A \in \mathcal{A}) = \sum_\sigma \mathbb{P}(A \in \mathcal{A}|\sigma)\mathbb{P}(\sigma)$, where the sum is over all possible cyclic orders containing $x_1 < x_2 < \cdots < x_k$. But for any σ, $\mathbb{P}(A \in \mathcal{A}|\sigma) \leq k/n$ and hence $\mathbb{P}(A \in \mathcal{A}) \leq k/n$.

Thus $|\mathcal{A}| \leq \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}$.

For the uniqueness part, we remark that for every cyclic ordering of $[n]$, the intervals must be of the form all containing a single fixed element. Suppose we look at a cyclic ordering and have an interval in $\mathcal{A}$, say $x_1 \dots x_k$, and another interval $x_0 < x_1 < \cdots < x_{k-1}$ that is *not* in $\mathcal{A}$. Then the inflection point in the first interval must be x_k, and we contain all intervals beginning after x_1 and passing through x_k.

So if we take an arbitrary cyclic order, and choose $x_1 < x_2 < \cdots < x_{2k-1}$ such that all the intervals of length k within $\mathcal{A}$ are contained in $[x_1, x_{2k-1}]$ and we take $x_0 < x_1$, then $[x_0, x_{k-1}] \notin \mathcal{A}$, but $[x_j, x_{j+k-1}] \in \mathcal{A}$ for each $j = 1, \dots, k$.

Now what we want to prove is that if we have any set containing x_k, then it must be in $\mathcal{A}$, *i.e.* claim $B = \{y_1, \dots, y_{k-1}, x_k\} \in \mathcal{A}$.

wlog $y_1, \dots, y_r \in \{x_1, \dots, x_{k-1}\}$ and $y_{r+1}, \dots, y_{k-1} \notin \{x_1, \dots, x_{k-1}\}$. We construct a new cyclic order starting with x_0 as above: $x_0 < z_1 < z_2 < \cdots < z_s < y_1 < \cdots < y_r < x_k < y_{r+1} < \cdots < y_{k-1} < \dots$, where the elements z_i are precisely the elements $\{x_1, \dots, x_{k-1}\}\backslash\{y_1, \dots, y_r\}$.

Then in this new order the interval $[u, x_k]$ is the same interval as what we had before that did not belong to x_k, but $[z_1 \dots x_k]$ is simply a reordering of $[x_1, x_k]$ as before, so we have the same situation as before; two adjacent intervals, one not in $\mathcal{A}$ and one in $\mathcal{A}$. Hence all of the k-intervals containing x_k in this order are in $\mathcal{A}$ - and this includes the interval $[y_1, \dots, y_r, x_k, y_{r+1}, \dots, y_{k-1}]$. Hence the arbitrary $B \supset \{x_k\}$ is in $\mathcal{A}$. □

We now consider a different constraint on a set system - no set in the system is contained within any other; these are sometimes called *antichains*.

Before that though, we discuss the discrete cube.

We can think about the discrete cube in different ways; *e.g.* as $\mathcal{P}X$ where $|X| = n$, or $\{0,1\}^n$.

For the case $n = 3$, this is of course the standard 3D cube we are all familiar with. We can label the vertices of the cube with binary sequences of length 3, or with subsets of $\{1, 2, 3\}$, in layers (by drawing the cube appropriately) such that each layer correspond to the possible sizes of sets.

If we have $\mathcal{A} \subset \mathcal{P}[n]$, we define $\mathcal{A}_k = \{A \in \mathcal{A} : |A| = k\}$ to be the k^{th} *layer* of $\mathcal{A}$.

Theorem: (Sperner's Theorem) *Let $\mathcal{A} \subset \mathcal{P}[n]$. If we have $A, B \in \mathcal{A} \implies A \not\subseteq B$ (i.e. not a proper subset), then*

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

This proof has a remarkably short and simple proof, though it omits some further information about this situation.

Proof. Pick a random maximal chain $\mathcal{C} = \{\emptyset, \{x_1\}, \{x_1, x_2\}, \ldots, \{x_1, \ldots, x_n\}\}$. How many elements of $\mathcal{A}$ do we expect this chain to contain?

One answer is that we of course can have at most one, since for any pair in the chain one contains the other.

Let $Asubset[n]$ with $|A| = k$ and consider the probabilty that A belongs to this chain. Then if it belongs to the chain $\mathcal{C}$, then it is the unique set of size k within it. But the chains containing any k-set are distributed uniformly, so $\mathbb{P}[A \in \mathcal{C}] = 1/\binom{n}{k}$.

So the expected number of elements of $\mathcal{A}$ in $\mathcal{C}$ by linearity of expectation is

$$\sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} = \sum_{k=0}^{n} \frac{|\mathcal{A}_k|}{\binom{n}{k}}$$

As we noted before, this cannot exceed 1, hence

$$\sum_{k=0}^{n} \frac{|\mathcal{A}_k|}{\binom{n}{k}} \leq 1$$

$$\therefore \sum_{k=0}^{n} |\mathcal{A}_k| = |\mathcal{A}| \leq \binom{n}{\lfloor n/2 \rfloor}$$

We will deal with the equality case as well. If n is even, then when we multiplied through by $\binom{n}{n/2}$ we must have maintained equality, and so $|\mathcal{A}_k| = 0$ for $k \neq n/2$. Hence $\mathcal{A} \subset [n]^{(n/2)}$.

Similarly if n is odd, then setting $m = (n-1)/2$ we have

$$\mathcal{A} \subset [n]^{(m)} \cup [n]^{(m+1)}$$

We will show that we must simply have $\mathcal{A} = \mathcal{A}_m$ or $\mathcal{A}_{m+1}$.

We must have $\mathcal{A}_{m+1} \cap \partial\mathcal{A}_m = \emptyset$, so if $|\mathcal{A}_{m+1}| + |\mathcal{A}_m| = \binom{n}{m} = \binom{n}{m+1}$ then $|\partial^{m+1}\mathcal{A}_m| \leq |\mathcal{A}_m|$. Each $A \in \mathcal{A}_m$ is contained in $m+1$ sets in $[n]^{(m+1)}$ and hence in $\partial\mathcal{A}_m$. Each $B \in \partial A_m$ contains at most $m+1$ sets in $\mathcal{A}_m$, so $|\partial\mathcal{A}_m| \geq |\mathcal{A}_m|$ by our earlier double counting argument, with equality if and only if every set in $\partial\mathcal{A}_m$ contains $m+1$ sets in $\mathcal{A}_m$.

So if $A \in \mathcal{A}_m$, then adding any element and removing another gives another set in $\mathcal{A}_m$. Using this process we can turn any m-set into any other, hence either $\mathcal{A}_m = \emptyset$ or $\mathcal{A}_m = [n]^{(m)}$. □

Remark: The inequality

$$\sum_{k=0}^{n} \frac{|\mathcal{A}_k|}{\binom{n}{k}} \leq 1$$

is known as the LYM inequality, after Lubell, Yamamoto and Meshalkin, since it was discovered independently by these three people.

3 Szemeredi-Trotter Theorem

To describe this theorem, we first define an ***incidence***: if we have some points and lines in the plane, an incidence is simply a point on a line.

The proof we will give is a cuter one that is due to Szekely.

Definition: (Crossing Number) The crossing number of a graph is the smallest number of crossings (pairs of crossing edges) you can have when you draw the graph in the plane.

If G is planar then its crossing number is of course zero.

Recall that a planar graph with n vertices has at most $3n-6$ edges (for $n \geq 3$ - this breaks down for $n = 2$ in the obvious way).

We can say a bit about crossing numbers just using the above fact; if we have a graph G with $e(G) = m > 3n-6$ then G has a crossing. So we could find a crossing, and remove an edge from the pair. Then either there are no more crossings, or $m-1 > 3n-6$ and G has at least two crossings.

This simple argument gives us that G has at least $m-3n+6$ crossings. The important bit to remember is that it is $\geq m-3n$. However, we can say a whole lot more than that using an averaging argument.

Now let G be a graph with m edges and n vertices (we will have $m >> n$). Suppose we pick a random induced subgraph H of G by picking each vertex independently with probability p, and suppose G has t crossings. Then the expected number of vertices of H is pn, the expected number of edges is p^2m, and the expected number of crossings we count by considering how edge pairs survive; this requires all four vertices involved to be chosen, which happens with probability p^4. Hence we expect H to have p^4t.

It then follows that $p^4t \geq p^2m - 3pn$ (by expectation), otherwise there exists a graph where LHS - RHS ¡ 0, which is not possible. Then choose p to be $6n/m$ (for which we require $m \geq 6n$), so that $p^4t \geq 3pn$, and thus $t \geq 3n/p^3 = 3nm^3/216n^3 = m^3/72n^2$.

Hence if $m \geq 6n$ then # crossings $\geq m^3/72n^2$.

This is much better; if we have $m = O(n^2)$, then we have a lower bound of the form $O(n^4)$.

This crossing number inequality is precisely the ingredient that we need in order to prove the Szemeredi-Trotter Theoerem. Before that though, we need to turn a system of lines and points in the plane into a graph. The simple way in which we do this is to say that if we have two points that are joined by a line segment, then we regard that line segment as an edge. We note that it does not need to be planar.

Let the lines be $L_1, \ldots, L_m$, and the points $x_1, \ldots, x_n$. Define $r_i > 0 =$ # of points in line L_i. The # of incidences is then $\sum_{i=1}^{m} r_i = I$, the # of edges is $\sum_{i=1}^{m}(r_i - 1) = I - m$, since we do not use the parts of the line that stretch away to infinity. The # of vertices is n.

If $I - m \geq 6n$, then the crossing number inequality tells us that # crossings $\geq (I-m)^3/72n^2$. Moreover, we have the trivial upper bound that the number of crossings is at most the number of

pairs of lines, hence # crossings $\leq \binom{m}{2} \leq m^2/2$. Hence

$$\frac{(I-m)^3}{72n^2} \leq \frac{m^2}{2}$$
$$\therefore I \leq (36m^2n^2)^{1/3} + m$$

In particular, $I \leq C\max\{m, n, m^{2/3}n^{2/3}\}$ for some constant C. So we have

Theorem 3.1: (Szemeredi-Trotter Theorem)

$$I \leq C\max\{m, n, m^{2/3}n^{2/3}\}$$

There are a lot of occasions when working on a combinatorics problem that we require some kind of bounds on the parameters involved. We will now cover some fairly ubiquitous bounds that we will find very helpful.

The first bound we take a look at is a bound on $n!$. The main trick here is to consider

$$\log n! = \log 1 + \log 2 + \cdots + \log n$$

which we bound by sandwiching it between a pair of integrals. In particular, since $\log x$ is concave we have

$$\log n! \leq \int_1^{n+1} \log x \ \mathrm{d}x = [x\log x - x]_1^{n+1}$$

Hence

$$\log n! \leq (n+1)\log(n+1) - n$$
$$\implies n! \leq (n+1)^{n+1}\mathrm{e}^{-n} = n\cdot n^n \cdot \left(1+\frac{1}{n}\right)^{n+1}\mathrm{e}^{-n}$$
$$\leq (n+1)n^n\mathrm{e}^{-n} = (n+1)\left(\frac{n}{e}\right)^n$$

Similarly, if we take the limit to n instead of $n+1$, we have

$$\log n! \geq \int_1^n \log x \ \mathrm{d}x = [x\log x - x]_1^n = n\log n - n + 1$$
$$\therefore n! \geq n^n\mathrm{e}^{-n}\cdot \mathrm{e} \geq \left(\frac{n}{e}\right)^n$$

The takeaway here is that $(n/e)^n$ is often a good approximation for $n!$ - but not always. Consider

$$\binom{n}{\frac{n}{2}} = \frac{n!}{(\frac{n}{2})!\,(\frac{n}{2})!} \leq 2^n$$

and this isn't very good. But we can do better:

$$2^{-n}\binom{n}{n/2} = \frac{1\cdot 2\cdots n}{2\cdot 4\cdots n\cdot 2\cdot 4\cdots n} = \frac{1\cdot 3\cdots(n-1)}{2\cdot 4\cdots n}$$
$$\implies [\,\cdot\,]^2 = \left(\frac{1}{2}\right)^2\left(\frac{3}{4}\right)^2\cdots\left(\frac{n-1}{n}\right)^2$$
$$\leq \left(\frac{1}{2}\cdot\frac{2}{3}\right)\left(\frac{3}{4}\cdot\frac{5}{6}\right)\cdots\left(\frac{n-1}{n}\cdot\frac{n}{n+1}\right)$$
$$\leq \frac{1}{n+1}$$

But also,

$$\left(\frac{1}{2}\right)^2\left(\frac{3}{4}\right)^2\cdots\left(\frac{n-1}{n}\right)^2 \geq \left(\frac{1}{2}\cdot\frac{1}{2}\right)\left(\frac{2}{3}\cdot\frac{3}{4}\right)\cdots\left(\frac{n-2}{n-1}\cdot\frac{n-1}{n}\right)$$
$$= \frac{1}{2n}$$

Putting this all together, we have thus proved that

$$\frac{1}{\sqrt{2n}} \leq 2^{-n}\binom{n}{\frac{n}{2}} \leq \frac{1}{\sqrt{n}}$$

This isn't particularly surprising, if we consider an r.v. $X = X_1 + \cdots + X_n$, with $X_i = \pm 1$ with probability $1/2$. Then

$$2^{-n}\binom{n}{n/2} = \mathbb{P}[X = 0]$$

and the standard deviation of X is $\sqrt{n}$, so the distribution is clustered about 0 (roughly normally) contained mostly within a range $\sqrt{n}$, and the middle 'slice' of this is one of the discrete values. This isn't rigorous, but it is just a sanity check.

Another very simple estimate is

$$\binom{n}{m} \leq n^m$$

This is almost embarrassingly simple, but it is often very helpful. We can also improve the bound using our earlier work, to conclude

$$\binom{n}{m} \leq \left(\frac{\mathrm{e}n}{m}\right)^m$$

However, even the above bound stops being useful when $m \approx n/2$... we'll deal with this in a second. For now, consider

$$\frac{\binom{n}{k-1}}{\binom{n}{k}} = \frac{k}{n-k+1} \leq \frac{k}{n-k}$$

We can use this. Let $\theta > 0$ and assume everything that needs to be an integer is. Then

$$\binom{n}{\left(\frac{1}{2}-\theta\right)n} \leq \left(\frac{\frac{1}{2}-\frac{\theta}{2}}{\frac{1}{2}+\frac{\theta}{2}}\right)^{\frac{\theta n}{2}}\binom{n}{\left(\frac{1}{2}-\frac{\theta}{2}\right)n}$$

This looks like a complicated mess, but we get there simply by 'walking' up the $\theta n/2$ binomial coefficients between the two shown, picking up the upper bound on the ratio each time. Moreover, notice that

$$\left(\frac{1-\theta}{1+\theta}\right)^{\frac{\theta n}{2}} \leq (1-\theta)^{\frac{\theta n}{2}} \leq \mathrm{e}^{-\frac{\theta^2 n}{2}}$$

(in general, $(1-x)^t \leq \mathrm{e}^{-tx}$). Hence

$$\binom{n}{\left(\frac{1}{2}-\theta\right)n} \leq \mathrm{e}^{-\theta^2 n/2}\cdot 2^{-n}$$

We obtain a similar bound using a sum of independent random variables. Let $X_1 \ldots, X_n$ be independent taking values in $[-1, 1]$, all of mean zero, and take $X = \sum X_i$. We consider $\mathbb{P}[X \geq \varepsilon n]$, for some $\varepsilon > 0$. We will take ***exponential moments***, and then use ***Markov's Inequality***. Consider

$$\begin{aligned}\mathbb{E}\mathrm{e}^{\lambda X} &= \mathbb{E}\mathrm{e}^{\lambda \sum X_i} \\ &= \mathbb{E}\prod \mathrm{e}^{\lambda X_i} = \prod \mathbb{E}\mathrm{e}^{\lambda X_i}\end{aligned}$$

We can commute $\prod$ and $\mathbb{E}$ by the independence of the X_i. We have that

$$\begin{aligned}\mathbb{E}\mathrm{e}^{\lambda X_i} &= \mathbb{E}\left[1 + \lambda X_i + \frac{\lambda^2 X_i^2}{2} + \frac{\lambda^3 X_i^3}{3!} + \cdots\right] \\ &\leq 1 + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots \\ &\leq 1 + \lambda^2 \leq \mathrm{e}^{\lambda^2} \text{ for when } \lambda \leq 1\end{aligned}$$

Hence $\mathbb{E}\mathrm{e}^{\lambda X} \leq \mathrm{e}^{n\lambda^2}$. Thus

$$\mathbb{P}[X \geq \varepsilon n] = \mathbb{P}[\mathrm{e}^{\lambda X} \geq \mathrm{e}^{\lambda \varepsilon n}] \leq \mathrm{e}^{n\lambda^2 - \lambda\varepsilon n}$$

by Markov.

We are now free to choose $0 < \lambda < 1$; pick $\lambda = \varepsilon/2$, minimising the RHS, to get $\mathrm{e}^{-\varepsilon^2 n/4}$. We also note that by looking at $-X$ instead of X, we have

$$\mathbb{P}[X \leq -\varepsilon n] \leq \mathrm{e}^{-\varepsilon^2 n/4}$$

We now consider the following

$$2^{-n}\left(\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{\left(\frac{1}{2} - \theta\right) n}\right) = \mathbb{P}[X \leq -2\theta n]$$

which holds because, if the X_i were only to take values 0 or 1 with probability $1/2$, the LHS is the probability that their sum is not larger than $(1/2 - \theta)n$ by conditioning over the acceptable outcomes. Transforming to let the X_i take values $\{-1, 1\}$ then gives the desired equality. Then, by the above, we have:

$$2^{-n}\left(\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{\left(\frac{1}{2} - \theta\right) n}\right) \leq \mathrm{e}^{-\theta^2 n}$$

A final remark is that it is useful to have a better bound in some cases:

$$\binom{n}{\alpha n} \sim 2^{H(\alpha)n}$$

where

$$H(\alpha) = \alpha \log \frac{1}{\alpha} + (1 - \alpha) \log \frac{1}{1 - \alpha}$$

is the ***entropy function***.

4 Well-Separated Set Systems

We shall take a collection $\mathcal{A} \subset [n]^{(n/2)}$, with the condition that $A \neq B \in \mathcal{A} \implies |A \cap B| \leq \alpha n$.

Since we expect that $|A \cap B| = n/4$, it makes a huge difference how large α is relative to $1/4$.

If $\alpha > 1/4$, write $\alpha = \frac{1}{4} + \varepsilon$; *i*.e. the intersection sizes shouldn't be too much larger than if we just chose them at random.

Let $A \in [n]^{(n/2)}$, and choose $B \in [n]^{(n/2)}$ at random. What is $\mathbb{P}[|A \cap B| > (1/4 + \varepsilon)n]$?

Let $m = (1/4 - \varepsilon)n$; then $|A \cap B| \geq (1/4 + \varepsilon)n \iff |A^c \cap B| \leq m$.

The total number of such B is just $\binom{n}{n/2}$, and we want to count those which intersect as desired:

$$\begin{aligned}
&= \binom{n}{n/2}^{-1} \sum_{r=0}^{m} \binom{n/2}{r}\binom{n/2}{n/2-r} \\
&\leq \frac{2\sqrt{n}}{2^n} \frac{2^{n/2}}{\sqrt{n/2}} \sum_{r=0}^{m} \binom{n/2}{r} \\
&\leq 4 \cdot 2^{-n/2} \sum_{r=0}^{m} \binom{n/2}{r}
\end{aligned}$$

We then use the above bound derived earlier, to obtain ($n \mapsto n/2$, $\varepsilon \mapsto 2\varepsilon$):

$$\leq 4 \cdot \mathrm{e}^{-2\varepsilon^2 n}$$

Now pick $A_1, \ldots, A_N$ independently at random from $[n]^{(n/2)}$. Call a pair (i, j) *b*ad if $i \neq j$ and $|A_i \cap A_j| \geq (1/4 + \varepsilon)n$.

Then $\mathbb{E}\#$ bad pairs $\leq \binom{N}{2} \cdot 4 \cdot \mathrm{e}^{-2\varepsilon^2 n}$. So we can find $A_1, \ldots, A_N$ such that the number of bad pairs is also bounded by that amount.

If # bad pairs is $\leq N/2$, then we can throw away a set from each bad pair and still have $\geq N/2$ sets, which will then be well-separated as desired. For this to work, we have a sufficient condition

$$\begin{aligned}
\frac{N^2}{2} \cdot 4\mathrm{e}^{-2\varepsilon^2 n} &\leq \frac{N}{2} \\
\iff N &\leq \frac{1}{4}\mathrm{e}^{2\varepsilon^2 n}
\end{aligned}$$

This technique is good, but we could instead have simply chosen our sets greedily:

Pick A_1, find A_2 good, find A_3 good with both A_1, A_2 *e*tc..., until we get stuck at some A_N. Each A_i rules out at most $4\mathrm{e}^{-2\varepsilon^2 n} \cdot 2^n$ sets, so we will not get stuck unless $N \cdot 4\mathrm{e}^{-2\varepsilon^2 n} \geq 1$, so we can have $N = \frac{1}{4}\mathrm{e}^{2\varepsilon^2 n}$; hence the greedy algorithm works just as well here (though this is not always the case).

What if instead we have $\alpha < \frac{1}{4}$? We now turn to a ubiquitous technique in combinatorics:

Think of sets as vectors!

There are many ways in which we can do this; perhaps the most obvious way is to replace it by its indicator sequence.

In this case, we have $A \in [n]^{(n/2)}$; a rather nice way of converting it into a vector is to identify it with the function $f_A : [n] \to \mathbb{R}$, defined by

$$f_A(x) = \begin{cases} 1 & x \in A \\ -1 & x \notin A \end{cases}$$

If $|A \cap B| \leq (1/4 - \varepsilon)n$, then $\langle f_A, f_B \rangle \leq -4\varepsilon n$ by drawing a Venn diagram and bounding the various regions appropriately. If we then normalise by letting $g_A = n^{-1/2} f_A$, then $||g_A||_2 = 1$, and $\langle g_A, g_B \rangle \leq -4\varepsilon$.

So we can solve a new problem: how many unit vectors in $\mathbb{R}^n$ can there be if any two have inner product $\leq -\delta$, for some $\delta > 0$.

Let $u_1, \ldots, u_m$ be unit vectors, and $\langle u_i, u_j \rangle \leq -\delta$ when $i \neq j$. Consider:

$$\begin{aligned} 0 \leq ||\sum_{i=1}^{m} u_i||_2^2 &= \sum_{i,j=1}^{m} \langle u_i, u_j \rangle \\ &\leq m - m(m-1)\delta \\ \implies m &\leq 1 + \frac{1}{\delta} \end{aligned}$$

Surprisingly, this bound is independent of the dimension, which is a huge contrast from the exponentially-many sets that we had before. In particular, back in the original problem this gives us at most $1+1/4\varepsilon$ sets.

This argument generalises very easily. Suppose that each set has size θn, and $|A \cap B| \leq (\theta^2 - \varepsilon)n$. Then again we get a bound that is independent of n.

It is worth remarking that the $1 + 1/\delta$ bound is best possible when $\delta = 1/k$ for some $k \in \mathbb{N}$. To achieve this we take the $k + 1$ vertices of a regular simplex in k-dimensions. To prove this is rather tricky, unless you have a trick:

We are trying to find a simplex with $k+1$ vertices in $\mathbb{R}^k$; the trick is to look in the subsapce $\{x \in \mathbb{R}^{k+1} : x_1 + \cdots + x_{k+1} = 0\}$. We then find vertices of the form $(-1/k, -1/k, \ldots, -1/k, 1, -1/k, \ldots, -1/k) = v_i$ for the 1 in the i^{th} place.

Then for $i \neq j$, we have $\langle v_i, v_j \rangle = \frac{k-1}{k^2} - \frac{2}{k} = -\frac{1}{k^2} - \frac{1}{k}$. If $i = j$, then we have $||v_i||^2 = k/k^2 + 1 = 1 + 1/k$. So to normalise, we set $u_i = v_i/\sqrt{1 + 1/k}$, and this cancels out appropriately to give $\langle u_i, u_j \rangle = -1/k$, as desired.

Lecture 5

We now consider the case where $\alpha = 1/4$. We are quickly lead to the following question.

Suppose we have $u_1, \ldots, u_m$ non-zero vectors in $\mathbb{R}^n$, and $\langle u_i, j_j \rangle \leq 0$ for all $i \neq j$. How big can m be? (This corresponds to the question of asking how many sets we can have of size $n/2$ with $|A_i \cap A_j| \leq n/4$.)

We will prove by induction that if $u_1, \ldots, u_{2n+1}$ are non-zero vectors, then $\exists i \neq j$ such that $\langle u_i, u_j \rangle > 0$. Moreover if $u_1, \ldots, u_{2n}$ satisfy $\langle u_i, u_j \rangle \leq 0$ for all $i \neq j$, then they are of the form $a_i e_i, b_i e_i$ for $i = 1, \ldots, n$ with $a_i > 0$, $b_i < 0$ and $e_1, \ldots, e_n$ an orthonormal basis.

Suppose we have $u_1, \ldots, u_{2n} \in \mathbb{R}^n$ non-zero, and that $i \neq j \implies \langle u_i, u_j \rangle \leq 0$. For each $i \geq 2$, write $u_i = a_i u_1 + v_i$, with $\langle u_1, v_i \rangle = 0$. Since $\langle u_1, u_i \rangle \leq 0$, we get that $a_i \leq 0$. Moreover, $\langle u_i, u_j \rangle = a_i a_j + \langle v_i, v_j \rangle \leq 0$. Since the a_i are both non-positive, we conclude that $\langle v_i, v_j \rangle \leq 0$.

But $v_2, \ldots, v_{2n} \in \langle u_1 \rangle^{\perp}$, of dimension $n - 1$. And we have $2(n-1) + 1$ vectors, so by the induction hypothesis we deduce that v_i must be 0 for some i; let it be i.

So $u_i = a_i u_1$, so $a_i < 0$ since $u_i \neq 0$. So we have two of the vectors $u_1, a u_1$ for some $a < 0$, and this implies that all other vectors are orthogonal to u_1. This leaves a remaining $2(n-1)$ vectors, to which we apply induction and we are done.

Now suppose we have a set system $\mathcal{A} \subset [n]^{(n/2)}$ such that $A \neq B \in \mathcal{A} \implies |A \cap B| \leq n/4$. We then deduce that $|\mathcal{A}| \leq 2(n-1)$.

From this arises another question: can we find $n \times n$ matrices with ± 1 entries that are orthogonal? (By orthogonal, we mean that the rows and columns of the matrices are orthogonal - normalising them would give an orthogonal matrix in the normal linear algebra sense). These matrices are called ***Hadamard matrices***.

For example, observe that

$$\begin{pmatrix} A & A \\ A & -A \end{pmatrix}$$

is a Hadamard matrix. This construction gives us a way of building more Hadamard matrices, by letting A be itself an orthogonal matrix. For instance, we have (1), then $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, then

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

The k^{th} matrix in this sequence is known as W_k, the Walsh matrices.

We now look at a difference construction of a Hadamard matrix, called the ***Paley Matrices***. This requires a small amount of elementary number theory.

Take a prime p of the form $p = 4m + 3$. We define

$$A_{xy} = \begin{cases} 1 & x - y \text{ is a quad. res.} \\ -1 & \text{otherwise} \end{cases} = \left(\frac{x-y}{p}\right)$$

We note that we do not view 0 as a quadratic residue.

If we take the x^{th} row and the y^{th} rows of A, we get

$$\begin{aligned} \sum_z A_{xz}A_{yz} &= \sum_z \left(\frac{x-y}{p}\right)\left(\frac{y-z}{p}\right) \\ &= \sum_z \left(\frac{z}{p}\right)\left(\frac{z+d}{p}\right) \text{ for } d = y - x \\ &= -\left(\frac{d}{p}\right) \\ &= -\left(\frac{d}{p}\right) - \left(\frac{-d}{p}\right) + \sum_{z\neq 0,-d} \left(\frac{z}{p}\right)\left(\frac{z_d}{p}\right) \end{aligned}$$

Observe that if $-1 = x^2$ then $(-1)^{(p-1)/2} = x^{p-1} = 1$. So -1 is not a quadratic residue, so the first two terms cancel out.

We now take advantage of the multiplicative property of the Legendre symbol:

$$\begin{aligned} &= \sum_{z\neq 0,-d} \left(\frac{z^2}{p}\right)\left(\frac{1+dz^{-1}}{p}\right) \\ &= \sum_{z\neq 0,d} \left(\frac{1+dz^{-1}}{p}\right) \qquad\qquad = \sum_{x\neq 0,1} \left(\frac{x}{p}\right) \\ &= -1 \end{aligned}$$

We then take this matrix A, and surround it with 1s along the bottom row and final column. This contributes an additional 1 to each inner product between two rows, so the result is zero and all of the rows are clearly orthogonal to each other, with the exception of the final row. But this still works, because we get the sum of all the Legendre symbols off the diagonal and final column, a -1 on the

diagonal and a 1 on the final column. So the inner product is still zero, and all the rows are pairwise orthogonal.

This matrix is the Paley matrix, and it is another example of a Hadamard matrix.

Note: this is also a useful technique in combinatorics, which is to use algebra to come up with particularly nice constructions.

Consider now the question of for which n is there an $n \times n$ Hadamard matrix? We know that if $n \geq 4$, then n must be a multiple of 4. A current open problem is: can we do it for all multiples of 4? The least n for which we are unsure is $n = 668$.

5 The Sum-Product Problem

Suppose we have a set $A \subset \mathbb{R}_+$. We define the ***sum set*** $A + A = \{a + b : a, b \in A\}$ and the ***product set*** $A \cdot A = \{ab : a, b \in A\}$. We can do this for other operations too, like the ***quotient set*** $A/A = \{a/b : a, b \in A\}$.

What can we say about the sizes of these new sets?

For example, $|A| = n \implies |A + A| \leq \frac{n(n+1)}{2}$ by counting individual elements and distinct pairs. In the other direction, if we take $A = \{1, 2, \ldots, n\}$ then $|A + A| = 2n - 1$, and it is easy to see that this is best possible; if we have $a_1 < a_2 \cdots < a_n$, then we obtain $2n - 1$ distinct elements of $A + A$ by looking at $a_1 + a_1 < a_1 + a_2 < a_1 + a_3 < \cdots < a_1 + a_n < a_2 + a_n < \cdots < a_n + a_n$.

For the product set, a similar type of proof works. Suppose we have $A = \{1, 2, 4, \ldots, 2^{n-1}\}$. Then $|A \cdot A| = 2n - 1$, and this is again best possible.

Erdos Szemeredi sum-product problem:

This problem was first posed by Erdos and Szemeredi:

Is it true that $\forall c > 0\ \exists n_0$ such that $\forall n \geq n_0, \max\{|A + A|, |A \cdot A|\} \geq n^{2-c}$.

This is somewhat plausible, and if we were to find a counterexample it would need to be an intermediate case where both $A + A$ and $A \cdot A$ are small.

For the moment, you might ask what we can say about A if $|A + A| \leq n^{1.99}$, but unfortunately we don't yet know very much at all.

We haven't solved this problem yet, but we do have the following result:

Theorem: (Jozsef Solymosi) *There is a bound $\geq n^{4/3}$ for the sum-prodcut problem.*

Definition: (Additive Energy) Let $A \subset \mathbb{R}$. Define $\rho_A^+(x) = \#\{(a, b) \in A^2 : a + b = x\}$.

The ***additive energy*** of A is $\sum_{x \in A+A} \rho_A^+(x)^2$.

We similarly have $\rho_A^\times(x) = |\{(a, b) \in A^2 : ab = x\}|$, and $\rho_A^\div(x) = |\{(a, b) \in A^2 : a/b = x\}|$.

The ***multiplicative energy*** is then $\sum_{x \in A \cdot A} \rho_A^\times(x)^2$

Lemma:

$$\sum_{x \in A \cdot A} \rho_A^\times(x)^2 = \sum_{x \in A/A} \rho_A^\div(x)^2$$

Proof. Both sides count $|\{(a,b,c,d) \in A^4 : ab = cd\}|$. This is clearly true for the LHS; for each x we take all pairs multiplying to x, and square those to obtain quadruples *i.e.* paris of pairs.

But $ab = cd \iff a/c = d/b$, so it becomes clear that the RHS is exactly the same quantity for similar reasons. □

Lower bound

Note first that $\sum_{x \in A \cdot A} \rho_A^{\times}(x) = |A|^2$, since any pair of elements (a,b) contributes one to $|A|^2$ (clearly), and one to the LHS since it appears once in exactly one ρ-set. Moreover, by Cauchy-Schwarz, we have

$$\begin{aligned} |A|^2 &= \sum_{x \in A \cdot A} \rho_A^{\times}(x) \\ &\leq |A \cdot A|^{\frac{1}{2}} \left(\sum_{x \in A.A} \rho_A^{\times} 9x)^2 \right)^{\frac{1}{2}} \\ \therefore \text{ mult. energy } &\geq \frac{|A|^4}{|A.A|} \end{aligned}$$

Upper Bound

We'll begin by partitioning $A \times A$ according to "gradient", where we look at the plane split up by the gradients *i.e.* members of A/A.

Let $A/A = \{m_1, \ldots, m_t\}$, $m_i < m_j$ for $i < j$. For each i, let $B_i = \{(a,b) \in A^2 : b/a = m_i\}$ - these are the points in $A \times A$ lying on the line with gradient m_i.

Now it would be great if each B_i had the same size - unfortunately there is no reason for this to be the case. So we use a new strategy called ***dyadic decomposition***.

Each B_i has size between 1 and $|A|$. So we can partition A/A into at most $\lceil \log_2 |A| \rceil$ sets such that if m_i and m_j are in the same set, then $|B_i| \geq \frac{1}{2}|B_j|$.

We also know that $\sum \rho_A^{\div}(x)^2 = \sum_i |B_i|^2$. By averaging, we can find a collection of the B_is, which we shall rename as $B_1, \ldots, B_s$ such that

$$\sum_{i=1}^{s} |B_i|^2 \geq \frac{1}{\lceil \log_2 |A| \rceil} \sum \rho_A^{\div}(x)^2$$

We want to put an upper bound on the LHS. Consider $B_1 + B_2$; the directions of each set are linearly independent, we see that no two elements overlap and we get $|B_1 + B_2| = |B_1||B_2| \geq |B_1|^2/2$, and all the elements lie between the lines of B_1 and B_2. Moreover, $B_1 + B_2 \subset (A \times A) + (A \times A) = (A + A) \times (A + A)$ (this is the Cartesian product).

In general, we find that all $B_i + B_{i+1}$ are contained in this set, and in our lower bound we get all the $|B_i|^2$ except for $i = s$, so we need a bit more. We define a new set called B_{s+1} by drawing in a new vertical line through the first point in B_s, and mark off the points with the same height as those of B_s - these are still contained within $A \times A$, and then $B_s + B_{s+1} \subset (A + A) \times (A + A)$ also. So then

$$\begin{aligned} \sum_{i=1}^{s} |B_i|^2 &\leq 2|A + A|^2 \\ \therefore \frac{|A|^4}{|A.A|} &\leq \sum \rho_A^{\times}(x)^2 \leq 2|A + A|^2 \lceil \log_2 |A| \rceil \end{aligned}$$

Now we are done; we have that

$$|A+A|^2|A.A| \geq \frac{|A|^4}{2\lceil \log_2 |A| \rceil}$$
$$\implies \max\{|A+A|, |A.A|\} \geq \frac{|A|^{4/3}}{2^{1/3}\lceil \log_2 |A| \rceil^{1/3}}$$