MTH742P: Advanced Combinatorics 14/11/2016

Lecture 7: The Probabilistic Method

Lecturer: Heng Guo

1 The probabilistic method

We now move on to a very interesting and powerful technique in combinatorics, called the probabilistic method. There is an excellent (really excellent!) textbook written by Noga Alon and Joel Spencer [AS16]. In fact, most materials that we will cover are from [AS16].

The method works as follows. Our goal is to prove the existence of certain object or structure with desired properties. We start by define an appropriate probability space and then show that the desired properties hold with strictly positive probability.

We will illustrate the method by a result due to Erdős which is regarded as the initiation of the probabilistic method.

Definition 1. *The* Ramsey number $R(k, \ell)$ *is the smallest integer* n *such that in any two-colouring of the edges of a complete graph* K_n *by red and blue, either there is a red* K_k *(whose edges are all red) or there is a blue* K_ℓ*.*

Ramsey showed in 1929 that $R(k, \ell)$ is finite for any two integers k and ℓ. Let us start with a concrete example that $R(3, 3) = 6$.

To show that $R(3, 3) \leq 6$, consider a K_6. There are $\binom{6}{3} = \frac{6\cdot5\cdot4}{1\cdot2\cdot3} = 20$ triangles in K_6. For an arbitrary edge-colouring, let us count the number of triples uvw such that uv is red and vw is blue. Let $r(v)$ be the number of red edges adjacent to v, and $6 - r(v)$ is the number of blue edges adjacent to v. Then if $r(v) = 0$ or 5, then there is no such triples. If $r(v) = 1$ or 4, there are at most $1 \times 4 = 4 < 6$ such triples. If $r(v) = 2$ or 3, there are at most $2 \times 3 = 6$ such triples. Thus, there are at most $6 \times 6 = 36$ many such triples. On the other hand, every non-monochromatic triangle contributes 2 to these triples. Hence, the number of monochromatic triangles is at least $20 - 36/2 = 2$.

To show that $R(3, 3) > 5$, we only need to give a two edge-colouring of K_5 such that there is neither red K_3 nor blue K_3. This is shown in Figure 1.

Erdős's result is a general lower bound for the diagonal Ramsey numbers $R(k, k)$.

Theorem 1 (Erdős 1947)**.** *If* $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$*, then* $R(k, k) > n$*. In particular,* $R(k, k) > \lfloor 2^{k/2} \rfloor$ *for all* $k \geq 4$*.*

Proof. Consider a random colouring of edges of K_n by colouring each edge independently and uniformly; that is, the two colours for every edge are equally likely. For every fixed set

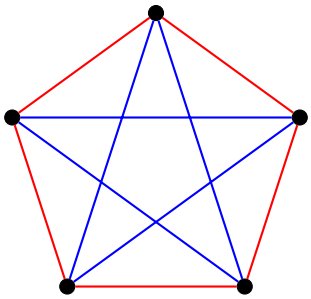

Figure 1: A 2-edge-colouring of K_5 without monochromatic triangles

R of k vertices, let A_R be the event that the induced subgraph K_R is monochromatic. For any R, there are $\binom{k}{2}$ edges and two monochromatic colourings. It implies that

$$\Pr(A_R) = \frac{2}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}.$$

There are $\binom{n}{k}$ possible choices of R. The probability that at least one of A_R occurs is at most $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ by the condition of the theorem. Thus, there is positive probability $1 - \binom{n}{k}2^{1-\binom{k}{2}} > 0$ that none of A_R occurs. This means there exists at least one colouring such that there is no monochromatic K_R.

For the second part, we pick $n = \lfloor 2^{k/2} \rfloor$ where $k \geq 4$, and verify that

$$\begin{aligned}\binom{n}{k}2^{1-\binom{k}{2}} &= \frac{n(n-1)\cdots(n-k+1)}{k!} \cdot 2^{1-k^2/2+k/2} < \frac{n^k}{k!} \cdot \frac{2^{1+k/2}}{2^{k^2/2}} \\ &= \left(\frac{n}{2^{k/2}}\right)^k \cdot \frac{2^{k/2+1}}{k!} < \frac{2^{k/2+1}}{2^{k-1}} \leq 1.\end{aligned}$$

This finishes the proof. □

In the proof, we used the union bound, which states that for any events $A_1, \cdots, A_n$,

$$\Pr\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \Pr(A_i). \tag{1}$$

An easy way to see this is by induction on n. Notice that

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \leq \Pr(A) + \Pr(B),$$

which implies the base case. For the induction step, we have that

$$\begin{aligned}\Pr\left(\bigcup_{i=1}^{n} A_i\right) &\leq \Pr\left(\bigcup_{i=1}^{n-1} A_i\right) + \Pr\left(A_n\right) \\ &\leq \sum_{i=1}^{n-1} \Pr\left(A_i\right) + \Pr\left(A_n\right) = \sum_{i=1}^{n} \Pr(A_i). \qquad \text{(by IH)}\end{aligned}$$

This simple example demonstrates the essence of the probabilistic method. In order to prove the existence of a good colouring, we do not represent one explicitly, but rather, in a nonconstructive way, show that it exists. This argument appeared in a paper of Erdős in 1947. Although the first appearance of the probabilistic method is due to Szele in 1943, Erdős was certainly the first to understand its power.

The basic paradigm is the following:

1. Define a suitable probability space.

2. Define the "bad events" A_i for $i \in [N]$ where N is the total number of bad events. Calculate $\Pr(A_i)$. Say $p \geq \Pr(A_i)$ for every $i \in [N]$.

3. Use the union bound:

$$\Pr\left(\bigwedge_{i\in[N]} \overline{A_i}\right) = 1 - \Pr\left(\bigvee_{i\in[N]} A_i\right) \geq 1 - \sum_{i\in[N]} \Pr(A_i) \geq 1 - Np.$$

4. Conclude that if $Np < 1$, then there exists a point in the probability space that avoids all bad events.

5. If necessary, show that $Np < 1$ is possible.

Of course, one may think that this argument is no different from a counting argument. However, we emphasize that there are many non-trivial tools in the probability theory, that are not easily translated to counting arguments, even though the probability space is finite and discrete, such as the second moment method, the Lovász Local Lemma, the concentration via martingales, etc.

Also, there is an algorithmic aspect of the probabilistic method. When we know some good object exists, of course we can enumerate all objects in the probability space. However typically doing so would take enormous time. For example, to find the good colouring in Theorem 1, there are $2^{\binom{n}{2}}$ many colourings to check. The probabilistic method is non-constructive in the sense that it does not directly provide an efficient way to find the good object either. However, we can often turn it into an efficient algorithm, randomized or deterministically.

For example, in the settings of Theorem 1. Let $k = 20$, and $n = \left\lfloor 2^{k/2} \right\rfloor = 2^{10} = 1024$. We randomly colour every edge of K_{1024} independently and uniformly. Using the calculation above, the probability of at least one K_{20} to be present (namely, the colouring is "bad") is at most

$$\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{2^{k/2+1}}{2^{k-1}} = \frac{2^{11}}{2^{19}} = 2^{-8} < 0.004.$$

In other words, a random colouring is good with probability at least 99.6%.

2 More examples

We will see a few more examples of the basic method.

2.1 Tournaments

Definition 2. *A* tournament *is an orientation of the complete graph K_n, such that every edge is $u \to v$ or $v \to u$.*

The name tournament is intuitive. It is a representation of a round-robin tournament in which every player encounters every other player exactly once, and in which no draws occur. Say we have n players, and if player a beats b, then we draw a directed edge from a to b.

We say a tournament has property S_k if for every k players, there exists another player v who defeats all of them. For example, a triangle $u \to v \to w \to u$ has property S_1 but not S_2. On the other hand, if we have $u \to v$, $u \to w$, and $v \to w$, then it does not have property S_1, as we can pick u.

Schüte raised the question of whether there exists a tournament with property S_k for every finite k. Erdős showed in 1963 that this can be answered almost trivially using the probabilistic method. The idea is that for sufficiently large n, a random tournament is very likely to have the property S_k.

Let $V = \{1, 2, \cdots, n\} = [n]$ be the set of vertices. For each pair $\{i, j\}$, we add the edge (i, j) or (j, i) uniformly at random. Thus, all $2^{\binom{n}{2}}$ many possible tournaments on V are equally likely.

Theorem 2 (Erdős 1963)**.** *If $\binom{n}{k}(1-2^{-k})^{n-k} < 1$, then there is a tournament on n vertices that has the property S_k.*

Proof. Consider the random tournament as described above. For every fixed subset $K \subset V$ of size K, let A_k be the event that there is no vertex that beats all members of K. Then we have that

$$\Pr(A_k) = (1 - 2^{-k})^{n-k}.$$

This is because for every fixed $v \in V \setminus K$, the probability that v beats all members of K is 2^{-k}, implying that the converse has probability $1 - 2^{-k}$. Moreover, all of these events are independent of each other.

Observe that the tournament does not have property K if all events A_k hold. By the union bound (1),

$$\Pr\left(\bigvee_{K \subset V,\, |K|=k} A_K\right) \leq \sum_{K \subset V,\, |K|=k} \Pr(A_K) = \binom{n}{k}(1-2^{-k})^{n-k} < 1.$$

As a consequence, the probability of none of A_k occurring is positive. Thus, there exists at least one tournament with property S_k. □

Let $s(k)$ be the minimum possible n such that there exists a tournament on $[n]$ with property S_k. Thus Theorem 2 gives an upper bound of $s(k)$. However, to compute this upper bound, we need to use some estimate of $\binom{n}{k}$. We will need the following bounds often:

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \frac{n^k}{k!} < \left(\frac{ne}{k}\right)^k. \tag{2}$$

The first one is because

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)\cdots(n-k+1)}{1\cdot 2\cdot 3\cdots k} \\ &= \frac{n}{k}\cdot\frac{n-1}{k-2}\cdots\frac{n-k+1}{1} \\ &\geq \left(\frac{n}{k}\right)^k,\end{aligned}$$

where we use the fact that $\frac{n-t}{k-t} \geq \frac{n}{k}$ for any $k \leq n$. The second one is because

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!}.$$

The last inequality is due to the Stirling approximation:

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

To be more precise, we have that

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

In particular,

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

The bounds in (2) are quite loose, but they are often useful to get the correct asymptotic.

Let us go back to Theorem 2. Notice that

$$\begin{aligned}(1-2^{-k})^{n-k} &= \left((1-2^{-k})^{2^k}\right)^{(n-k)/2^k} \\ &\leq e^{-(n-k)/2^k}.\end{aligned}$$

Together with (2), we have that

$$\binom{n}{k}(1-2^{-k})^{n-k} < \left(\frac{ne}{k}\right)^k e^{-(n-k)/2^k}.$$

Let $n = c \cdot k^2 2^k$ for some constant c. Then the right hand side simplifies to

$$\begin{aligned}\left(\frac{ne}{k}\right)^k e^{-(n-k)/2^k} &= \left(ck2^k e\right)^k e^{-ck^2+k/2^k} \\ &= \exp\left(k + \log 2 \cdot k^2 + k \log k + c \log k - ck^2 + k/2^k\right).\end{aligned}$$

The leading term of the exponent is k^2, and its coefficient is $\log 2 - c$. Thus if $c > \log 2$ is large enough, then $\binom{n}{k}(1-2^{-k})^{n-k} < 1$. In other words, we see that

$$s(k) \leq ck^2 2^k$$

for some constant c. It has also been shown that $s(k) \geq c_1 k 2^k$ for some constant c_1 by Szekeres.

On the other hand, to explicitly construct a tournament with property S_k and $n = ck^2 2^k$ is rather non-trivial. We will not cover it here.

2.2 2-colourable hypergraphs

Hypergraphs are generalizations of graphs in that each edge contains not necessarily two vertices.

Definition 3. *A* hypergraph *is $H = (V, E)$, where V is the set of* vertices*, and E is the set of* vertices*. Every $e \in E$ is a subset of vertices; namely $e \subset V$. If $|e| = k$ for every $e \in E$, then H is called* k-uniform*.*

Thus, the normal graph is just 2-uniform.

We say H is 2-*colourable* if we can colour all vertices so that none of the edges are monochromatic.

Theorem 3 (Erdős 1963)**.** *Every k-uniform hypergraph with less than 2^{k-1} edges is 2-colourable.*

Proof. Let H be a k-uniform hypergraph with $m < 2^{k-1}$ edges. Colour V uniformly at random by two colours. For each edge $e \in E$, let A_e be the event that e is monochromatic. Thus $\Pr(A_e) = 2^{1-k}$. By the union bound,

$$\Pr\left(\bigvee_{e \in E} A_e\right) \leq \sum_{e \in E} \Pr(A_e) = m2^{1-k} < 1.$$

Thus there exists a proper 2-colouring. □

Let $m(k)$ be the minimum number of edges such that a k-uniform hypergraph is not 2-colourable. Then we have that $m(k) \geq 2^{k-1}$ by Theorem 3. This bound has been improved into

$$m(k) \geq \Omega\left(\left(\frac{k}{\log k}\right)^{1/2} 2^k\right),$$

by Cherkashin and Kozik (2015).

To get an upper bound on $m(k)$, we need to change the probability setting. Now our goal is to construct a k-uniform hypergraph, with as many edges as possible, such that it cannot be 2-coloured. We will construct it by fixing n vertices, and randomly choosing m hyperedges (as subsets of $[n]$).

Let $V = [n]$, and e be a uniformly at random subset of V of size k. Fixing χ a 2-colouring of V with a red vertices and b blue vertices ($a + b = n$), we have that

$$\Pr(e \text{ is monochomatic under } \chi) = \frac{\binom{a}{k} + \binom{b}{k}}{\binom{n}{k}}.$$

We have shown that the function $\binom{n}{k}$ is convex. Thus $\binom{a}{k} + \binom{b}{k}$ is minimized when $a = b = \frac{n}{2}$. (Let n be even.) It implies that

$$\Pr(e \text{ is monochomatic under } \chi) \geq p,$$

where

$$p := \frac{2\binom{n/2}{k}}{\binom{n}{k}}.$$

Now let $e_1, \cdots, e_m$ be uniformly and independently chosen hyperedges of size k. For each colouring χ, let A_χ be the event that none of e_i is monochromatic. In other words, A_χ is the event that χ is proper. Since e_i is chosen independently,

$$\Pr(A_\chi) \leq (1-p)^m.$$

There are 2^n many colourings, implying that

$$\Pr\left(\bigvee_\chi A_\chi\right) \leq 2^n(1-p)^m.$$

As usual, if $2^n(1-p)^m < 1$, then there exists a collection of m hyperedges such that none of A_χ holds. In other words, there exist $e_1, \cdots, e_m$ so that all of χ are not proper.

To solve $2^n(1-p)^m < 1$, we will use the inequality $1 - p \leq e^{-p}$. This is valid for all $p > 0$, and this estimate is rather tight when p is small. Thus

$$2^n(1-p)^m \leq 2^n e^{-pm} = \exp(\log 2 \cdot n - pm).$$

We have that $m > \frac{\log 2 \cdot n}{p}$ implies $2^n(1-p)^m < 1$.

The next task is to minimize $\frac{n}{p}$. (Recall that p depends on n!) We have that

$$\begin{aligned} p = \frac{2\binom{n/2}{k}}{\binom{n}{k}} &= 2 \cdot \frac{n/2(n/2-1)\cdots(n/2-k+1)}{n(n-1)\cdots(n-k+1)} \\ &= 2^{1-k} \prod_{i=0}^{k-1} \frac{n-2i}{n-i}. \end{aligned}$$

Note that

$$\frac{n-2i}{n-i} = 1 - \frac{i}{n-i} \le e^{-i/(n-i)} = e^{-i/n+O(i^2/n^2)},$$

if $n \gg k > i$. Thus we have that

$$\begin{aligned}\frac{n}{p} &\sim n2^{k-1}e^{k^2/2n} \\ &= \exp(\log n + (k-1)\log 2 + k^2/2n).\end{aligned}$$

To optimize the exponent, take the derivative with respect to n and we get $\frac{1}{n} - \frac{k^2}{2n^2} = 0$. Thus we should pick $n = \frac{k^2}{2}$, in which case $n/p = (1+o(1))e/4 \cdot k^2 2^k$.

Theorem 4 (Erdős 1964)**.** $m(k) < (1+o(1))\frac{e\log 2}{4} \cdot k^2 2^k$.

2.3 Sum-free sequences

A set S of integers is said to be *sum-free* if for any $a_1, a_2 \in S$, $a_1 + a_2 \notin S$. So $\{2,3,7\}$ is sum-free whereas $\{2,3,4\}$ is not.

Theorem 5 (Erdős 1965)**.** *Every set $B = \{b_1, \cdots, b_n\}$ of n non-zero integers contains a sum-free subset A so that $|A| > \frac{n}{3}$.*

Proof. Let $p = 3k+2$ be a prime such that $p > 2\max_{i\in[n]} b_i$, and let $C = \{k+1, k+2, \cdots, 2k+1\}$. It is not hard to see that C is sum-free, and

$$\frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}.$$

Our goal is to randomly "project" C down to B.

Let us choose uniformly at random an integer x such that $1 \le x \le p$. Define $d_i \equiv xb_i \mod p$. For every fixed $i \in [n]$, as x ranges over $1, 2, \cdots, n-1$, d_i ranges over $1, 2, \cdots, n-1$ as well. Thus $\Pr(d_i \in C) = |C|/(p-1) > 1/3$.

Let X_i be the indicator variable that $d_i \in C$. Thus $\mathbb{E}\, X_i > 1/3$. By the linearity of expectations, the expected number of elements b_i such that $d_i \in C$ is

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}\, X_i > n/3.$$

Therefore, there exists an x such that a subset A of B satisfies that $|A| > n/3$ and $xa \in C \mod p$ for all $a \in A$.

We claim that this A is sum-free. Suppose otherwise, we have that $a_1 + a_2 = a_3$ for some $a_1, a_2, a_3 \in A$. This implies that $xa_1 + xa_2 \equiv xa_3 \mod p$. However, $xa_i \in C$ for $i = 1, 2, 3$. This contradicts to the fact that C is sum-free. □

Eberhard, Green, and Manners (2013) showed that the constant in Theorem 5 is tight. However, in the setting of a general abelian group (instead of integers), the optimal constant is 2/7 (Alon and Kleitman 1990).

References

[AS16] Noga Alon and Joel Spencer. *The Probabilistic Method*. John Wiley, fourth edition, 2016.