Lecture 9: Probabilistic methods

Rajat Mittal

IIT Kanpur

This lecture will focus on probabilistic methods. This is used to prove the existence of a *good* structure using probability. We will define a probability distribution over the set of structures. Then we prove that the good event happens with positive probability, which implies that a good structure exists.

These ideas are best illustrated with the help of applications.

1 Ramsey numbers

Previously in class we proved that if we color the edges of K_6 using blue or red color, then either there is a blue K_3 or a red K_3 as a subgraph. Here K_n is the complete graph (every pair of vertices are connected) on n vertices.

We can generalize the above concept and ask, are there complete graphs for which any 2-coloring gives rise to either a blue K_k or a red K_l. It has been shown that there always exists n, s.t., any two coloring of K_n will have a monochromatic blue K_k or monochromatic red K_l. The smallest such number n is called the *Ramsey number* $R(k, l)$.

It has been a big open question to find out the bounds on $R(k, l)$. We will use probabilistic method to give a lower bound on the diagonal Ramsey number $R(k, k)$.

Call an edge coloring of K_n *good*, if there are no monochromatic K_k's.

The idea would be to randomly color the edges of the graph K_n. If there is a positive probability (over the random coloring) that none of the K_k subgraphs are monochromatic red or blue, then there exist a coloring which is good.

We color every edge either red or blue independently with probability $1/2$. There are in total $\binom{n}{k}$ subgraphs K_k for a K_n.

Exercise 1. A particular subgraph K_k is monochromatic with probability $2^{1-\binom{k}{2}}$.

Hint: K_k could be completely red or completely blue.

We have already proved that,

$$Pr(\cup_{i=1}^{n} C_i) \leq \sum_{i=1}^{n} Pr(C_i).$$

So the total probability that any K_k is monochromatic is at most $\binom{n}{k}2^{1-\binom{k}{2}}$. If this probability is less than 1, then there is a positive probability that none of the K_k's are monochromatic.

Since the probability was over random coloring, there exist a good coloring (such that no K_k's are monochromatic).

Theorem 1. *If $2^{1-\binom{k}{2}}\binom{n}{k}$ is less than 1, then $R(k, k)$ is lower bound by n.*

To get an explicit lower bound, you can check that $n = \lceil 2^{k/2} \rceil$ will satisfy the above equation.

The essential argument in the above proof is that the number of colorings are much higher than the total number of graphs which have monochromatic K_k.

A counting argument for the above theorem can also be constructed (assignment problem). Actually, in all our applications, a counting argument can always be given. But the probabilistic argument in general is much simpler and easier to construct.

1.1 Probabilistic algorithm

One of the important thing to notice in a probabilistic method of proofs is that the proofs are non-constructive. For the previous example, it means that we were only able to show existence of a coloring. This proof does not construct the required coloring and hence is called non-constructive.

But suppose we choose n to be $\frac{1}{2}\left\lceil 2^{k/2}\right\rceil$. Then the probability of having a monochromatic K_k is very small. This shows that most of the random colorings will be good colorings.

This suggests a randomized algorithm. We take K_n and color the edges randomly. Because of the argument above, with high probability we will get a good coloring.

2 Sum-free subsets

Let's take another example. Given a set of integers S, $S+S$ is defined as the subset of integers which contain all possible sums of pair of elements in S.

$$S + S = \{t : t = s_1 + s_2,\ s_1, s_2 \in S\}$$

A set S is called *sum-free* if S does not contain any element of $S+S$.

Exercise 2. Construct a set of 10 elements which is sum-free. Construct a set of n elements which is sum-free.

Using probabilistic method, we will show that every large subset of integers contain a big enough subset which is sum-free.

Theorem 2. *For any subset S of n non-zero integers, There exist a subset of S which is sum-free and has size more than $n/3$.*

Proof. Suppose $S = \{s_1, s_2, \cdots, s_n\}$. The idea would be to map S to $rS = \{rs_1, rs_2, \cdots, rs_n\}$ for a random r. If some subset of rS is sum-free then the corresponding set in S will also be sum-free.

But taking r to be uniformly at random from $\mathbb{Z}$ is not feasible. First pick a prime p of the form $3k+2$, such that, p is at least 3 times bigger than the absolute value of any element of S.

You will show in the assignment that there are infinite primes of the form $3k+2$. We will do the calculations modulo p.

Notice that the set $T = \{k+1, k+2, \cdots, 2k+1\}$ is a sum-free subset when we do addition modulo p.

For applying the probabilistic method, pick a random x and consider the set $xS \mod p = \{xs_1 \mod p, xs_2 \mod p, \cdots, xs_n \mod p\}$.

Exercise 3. Show that if we pick an x at random from $0, 1, \cdots, p-1$ then xs_1 is also random with uniform probability.

Define a random variable Y which is the intersection size of $xS \mod p$ and T.

Using linearity of expectation,

$$E[Y] = \sum_i E[xs_i \mod p \in T].$$

Exercise 4. Show that $E[Y] = \frac{|S|}{3}$.

This implies that there exist at least one x for which $xS \mod p \cap T$ is of size at least $|S|/3$. Call that particular x, x_0. Then $T' = x_0 S \mod p \cap T$ is sum-free when addition is considered modulo p (T is sum-free). This implies that the pre-image in S which maps to T' is sum-free.

Exercise 5. Show that $x_0^{-1}T'$ is sum-free with respect to addition over integers.

□

3 Using linearity of expectation

We have already discussed linearity of expectation. It is a simple result to prove, but has profound implications. Again, the importance of linearity lies in the fact that we can even take dependent random variables and still decompose the expectation into components.

$$E[X + Y] = E[X] + E[Y].$$

for any two random variables X and Y.

Notice that we used linearity of expectation for the proof in the previous section. We will take some more examples now.

First let us look at the example of Ramsey number in the light of expectation.

Suppose we color each edge of K_n uniformly at random with blue or red. Define T to be the random variable which counts the number of monochromatic K_k in the coloring. We are interested in the expectation of T.

Define T_i (for i from 1 to $\binom{n}{k}$) to be the random variable which assigns 1 if a particular K_k is monochromatic otherwise 0. Convince yourself that $T = \sum_i T_i$.

Note 1. The random variables T_i are dependent on each other.

Then,

$$E[T] = \sum_i E[T_i] = \sum_i 2^{1-\binom{k}{2}} = \binom{n}{k} 2^{1-\binom{k}{2}}.$$

If $E[T] < 1$ then there exist a coloring which has less than or equal to $E[T]$ number of monochromatic K_k's. Since number of monochromatic K_k's is an integer, there exist a coloring for which number of monochromatic K_k's is zero.

Let's take another example of probabilistic method which utilizes linearity of expectation.

Theorem 3. *Given n unit vectors $v_i \in \mathbb{R}^n$, $i \in [n]$, there always exists a bit string $b \in \{-1, 1\}^n$, such that,*

$$\left\| \sum_i b_i v_i \right\| \leq \sqrt{n}.$$

Proof. Again, we will pick b_i's uniformly at random from $\{-1, 1\}$ and calculate the expected value of $N = \|\sum_i b_i v_i\|^2$.

From the definition of the length of a vector.

$$N = (\sum_i b_i v_i)^T (\sum_i b_i v_i) = \sum_{i,j} b_i b_j v_i^T v_j.$$

Notice that $v_i^T v_j$, the dot product between v_i and v_j, is a fixed number and the random variable are b_i's. Hence,

$$E[N] = \sum_{i,j} E[b_i b_j] v_i^T v_j.$$

By definition, we picked b_i and b_j independently. So b_i and b_j are independent if $i \neq j$. This implies that $E[b_i b_j] = E[b_i]E[b_j]$.

Exercise 6. Show that $E[b_i b_j] = 1$, if $i = j$ otherwise it is zero.

$$E[N] = \sum_i v_i^T v_i = n.$$

This implies that there is a choice of b_i's for which length of $\sum_i b_i v_i$ is less than or equal to $\sqrt{n}$.

□

Exercise 7. Given n unit vectors $v_i \in \mathbb{R}^n,\ i \in [n]$, there always exists a bit string $b \in \{-1, 1\}^n$, such that,

$$\left| \sum_i b_i v_i \right| \geq \sqrt{n}.$$

4 Super concentrators

This construction is taken from [4].

A *super concentrator* is a directed acyclic graph $G = (V, E)$, with n special input nodes $I \subset V$ and n output nodes $O \subset V$. It satisfies the property that for any $1 \leq k \leq n$, any k subset of I is connected with any other k subset of O with k disjoint paths.

This is a very strong connectivity property and can be used to design robust networks. It is very easy to design a super-concentrator with $O(n^2)$ edges (exercise). We will see how can linear size (number of edges) super-concentrators can be constructed.

We will do it recursively. But we first need *concentrators.*

A (n_1, n_2, u) concentrator has n_1 input nodes, n_2 output nodes. For any $1 \leq k \leq u$ input subset S, *there exists* a k size output subset which is connected to S using k disjoint paths.

Exercise 8. In concentrators, number of input and output nodes could be different and there is an upper limit on k size subset which we are allowed to choose. What is the other essential difference between concentrator and super-concentrator?

For all j, we will first construct $(6j, 4j, 3j)$ concentrator using probabilistic methods. Again the important thing is that we can construct such concentrators in linear size. Though since concentrators have weaker properties, it is less surprising than the linear construction of super-concentrators.

The concentrator is going to be a simple bipartite graph with $6j$ input nodes as one part and $4j$ output nodes as other part. There will be no extra vertices. We will show that if we pick linear number of edges randomly, with non-zero probability, there will be a graph which satisfies the properties of the concentrator.

Suppose every input node has out-degree 6 and every output node has in-degree 9 in our random graph. Every edge is labelled from one side by the vertex and a number between 1 and 6, from other side by an output vertex and a number between 1 and 9.

The random way to pick a graph is, for the first vertex v of input nodes and label 1 pick a partner ($36j$ choices). Then pick the partner for $v, 2$ and so on. Clearly there are $(36j)!$ ways of doing that.

Say a *non-concentrator* is a graph where there exists a subset S of size $\leq 3j$ of input nodes whose neighborhood is $\leq |S|$. We need to find the number of non-concentrators in the random process described above. We will show that it is less than $(36j)!$.

Exercise 9. Prove that if a graph is not a non-concentrator then it is a concentrator.

Hint: k disjoint paths in this case means a matching between k input and output nodes. Remember Hall's marriage theorem.

Note 2. We have taken a more stricter definition of non-concentrator then required. It is to help us in calculations.

Theorem 4. *The number of non-concentrators is less than* $(36j)!$.

Proof. A non-concentrator has a subset S of input nodes with size $k \leq 3j$, s.t., $|N(S)| \leq |S|$. If this is the case for an S,then $N(S)$ is a subset of some k size subset of output nodes T.

For a particular k, there are $\binom{6j}{k}$ ways to choose S and $\binom{4j}{k}$ ways to choose T. Fixing S and T, there are at most $\frac{(9k)!}{(3k)!}(36j - 6k)!$ ways for $N(S) \subseteq T$.

Since every vertex has at least 6 vertices in the neighborhood, we need to show,

$$\sum_{k=3}^{3j} \binom{6j}{k}\binom{4j}{k}\frac{(9k)!}{(3k)!}(36j-6k)! \leq (36j)!.$$

It is enough to show,

$$\sum_{k=3}^{3j} \binom{6j}{k}\binom{4j}{k}\binom{9k}{6k} \leq \binom{36j}{6k}. \tag{1}$$

This will be proved in the assignment.

□

Given a concentrator, the task to construct a super concentrator is easy. We will first connect all input nodes with one output node each creating a perfect matching of n edges. The remaining construction will use the concentrators. From $6j$ input nodes we will use a concentrator to connect them to new $4j$ nodes. Then put a super concentrator from these $4j$ nodes to next new $4j$ nodes. The final step is to connect the final $4j$ nodes to $6j$ output nodes using a reverse concentrator.

Exercise 10. There are two ways to go to output nodes. Using the perfect matching or the concentrators. For getting a super concentrator, prove that it is enough to show, for all $1 \leq k \leq 3j$, any k input set is connected to any k output set using concentrators with k disjoint paths.

Exercise 11. Show that for $k \leq 3j$, any k input nodes are connected to any k output nodes with k disjoint paths using concentrators.

This should convince you that we have a super-concentrator.

Exercise 12. Show that it has linear number of edges.

The construction of concentrators can be made deterministic. This is done using *expander graphs*. Interested students are advised to read more about expander graphs.

5 Assignment

Exercise 13. Give a counting argument for Thm. 1.

Exercise 14. Read about Stirling's bound on factorial and binomial coefficients.

Exercise 15. Show that there are infinite primes of the form $3k+2$.

Exercise 16. Suppose a vertex v has $deg(v)$ neighbors. Prove that the probability in a random permutation, v comes before any of its neighbors is, $\frac{1}{deg(v)+1}$.

Exercise 17. Consider a graph $G = (V, E)$. Show that G contains some independent set of size at least $\sum_{v \in V} \frac{1}{deg(v)+1}$.

Hint: Consider all permutations of $v_1, v_2, \cdots, v_n$. Take the independent set by considering vertices in the order of the permutation and once taken in the independent set, delete all its neighbors from the permutation.

Exercise 18. Let X be the random variable which counts the number of fixed points (i maps to i) in a random permutation. What is the expected value of X.

Exercise 19. Prove Eq. 1.

Hint: First prove that $\binom{36j}{6k} \geq \binom{6j}{k}\binom{4j}{k}\binom{26j}{4k}$ using a simple combinatorial argument.

References

1. N. Alon and J. H. Spencer. The Probabilistic Method. *Wiley*, 2008.
2. H. Tijms Understanding Probability. *Cambridge University Press*, 2012.
3. D. Stirzaker. Elementary Probability. *Cambridge University Press*, 2003.
4. U. Schöning. Gems of Theoretical Computer Science. *Springer-Verlag*, 1998.