6.842 Randomness and Computation February 6, 2022

Lectures on Probabilistic Methods

Lecturer: Ronitt Rubinfield *Scribe: Yuchong Pan*

The plan of probabilistic methods is to show that some certain "object" exists by showing that the probability that it exists is greater than 0. Since the probability that an object exists is either 0 or 1, then being greater than 0 implies that it has to be 1.

1 Examples

1.1 Hypergraph Coloring

We consider the following problem: Given a set X of elements and $S_1, \ldots, S_m \subset X$ each of size ℓ, output whether we can 2-color objects in X such that each set S_i is not monochromatic.

Theorem 1. *If $m < 2^{\ell-1}$, then there exists a proper 2-coloring.*

Proof. We randomly color elements of X red or blue, independently with probability 1/2. Then for all $i \in [m]$,

$$\mathbb{P}\left[S_i \text{ is monochromatic}\right] = \frac{1}{2^\ell} + \frac{1}{2^\ell} = \frac{1}{2^{\ell-1}}.$$

If $m < 2^{\ell-1}$, then by the union bound,

$$\mathbb{P}\left[\exists i \in [m] \text{ s.t. } S_i \text{ is monochromatic}\right] \leq \sum_{i=1}^{m} \mathbb{P}\left[S_i \text{ is monochromatic}\right] = \frac{m}{2^{\ell-1}} < \frac{2^{\ell-1}}{2^{\ell-1}} = 1.$$

Hence, $\mathbb{P}$[all S_i's are 2-colored] > 0, so there exists a setting of colors which is a good 2-coloring. □

The above proof does not give an explicit proper 2-coloring. However, if we require m to be even smaller, we can explicitly and quickly output a proper 2-coloring.

Theorem 2. *If $m < 2^{\ell-2}$, then there exists a proper 2-coloring, and we can find it quickly.*

Proof. We prove the second part of the theorem only. As in the previous proof, since $m < 2^{\ell-2}$,

$$\mathbb{P}\left[\exists i \in [m] \text{ s.t. } S_i \text{ is monochromatic}\right] = \frac{m}{2^{\ell-1}} < \frac{1}{2}.$$

Therefore, a random coloring of X is good with probability at least 1/2. (If not good, re-color until you find a good one.) The expected number of times to re-color is 2.[1] □

[1] Recall that given a coin with bias (i.e., probability of "heads") p, the expected number of tosses until we see heads is $1/p$.

1.2 Dominating Sets

Definition 3. Given a graph $G = (V, E)$, a subset $U \subset V$ is a *dominating set* if v has at least one neighbor in U for all $v \in V \setminus U$.

Theorem 4. *Let G be a graph. If G has minimum degree Δ, then G has a dominating set of size at most $\frac{4n\ln(4n)}{\Delta+1}$.*

Proof. We construct a subset $\hat{U} \subset V$ by putting each $v \in V$ into $\hat{U}$ independently at random with probability $p = \frac{\ln(4n)}{\Delta+1}$. For each $w \in V$, the probability that w has no neighbor in $\hat{U}$ and is not in $\hat{U}$ is $(1-p)^{\Delta+1}$. Recall that $\lim_{x\to\infty}(1-1/x)^x = 1/e$. Hence,

$$\begin{aligned}\mathbb{P}\left[\hat{U} \text{ is not a dominating set}\right] &= \mathbb{P}\left[\exists w \in V \text{ s.t. } w \text{ has no neighbor in } \hat{U} \text{ and is not in } \hat{U}\right] \\ &\leq n(1-p)^{\Delta+1} = n\left(1 - \frac{\ln(4n)}{\Delta+1}\right)^{\frac{\Delta+1}{\ln(4n)}\cdot\ln(4n)} \leq ne^{-\ln(4n)} = \frac{1}{4}.\end{aligned}$$

For each $w \in V$, let

$$\sigma_w = \begin{cases} 1, & \text{if } w \in \hat{U}, \\ 0, & \text{otherwise.} \end{cases}$$

Then $\mathbb{E}[\sigma_w] = \mathbb{P}[\sigma_w = 1] = p$. Since $|\hat{U}| = \sum_{w\in V}\sigma_w$,

$$\mathbb{E}\left[\left|\hat{U}\right|\right] = \sum_{w\in V}\mathbb{E}\left[\sigma_w\right] = np.$$

By Markov's inequality,

$$\mathbb{P}\left[\left|\hat{U}\right| > 4np\right] \leq \frac{\mathbb{E}\left[\left|\hat{U}\right|\right]}{4np} = \frac{np}{4np} = \frac{1}{4}.$$

Hence,

$$\mathbb{P}\left[\hat{U} \text{ is a dominating set of size at most } \frac{4n\ln(4n)}{\Delta+1}\right] \geq 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2} > 0.$$

This completes the proof. □

1.3 Sum-Free Subsets

Definition 5. A subset $A \subset \mathbb{N}$ is *sum-free* if there do not exist $a_1, a_2, a_3 \in A$ such that $a_1 + a_2 = a_3$.

Theorem 6 (Erdős '65)**.** *For any set B of size n, there exists a sum-free subset $A \subset B$ such that $|A| > n/3$.*

For example, if $B = [n]$, then $A = \{\lfloor n/2\rfloor, \ldots, n\}$ is sum-free and $|A| > n/3$.

Proof. Without loss of generality b_n is the maximum element in B. Pick a prime $p > 2b_n$ such that $p \equiv 2 \pmod 3$, i.e., $p = 3k+2$ for some $k \in \mathbb{Z}$. Let $C = \{k+1, \ldots, 2k+1\}$ be the "middle third." Note that

(i) $C \subset \mathbb{Z}_p$;[2]

[2]We write $\mathbb{Z}_p = \{0, \ldots, p-1\}$ and $\mathbb{Z}_p^* = \{1, \ldots, p-1\}$. Recall that $\mathbb{Z}_p$ has unique multiplicative inverses modulo p since p is a prime.

(ii) C is sum-free even in $\mathbb{Z}_p$ (why? the sum of any two elements is at least $2k+2$ and at most $4k+2 \equiv k \pmod{3k+2}$);

(iii) $\frac{|C|}{p-1} > 1/3$.

We construct A as follows: pick $x \in_R [p-1]$, and use x to define a linear map $f_x(a) = x \cdot a \mod p$.[3] Let $A_x = \{b \in B : f_x(b) = x \cdot b \mod p \in C\}$, i.e., x maps the elements of A_x to the middle third.

We claim that A_x is sum-free. If not, then there exist $b_1, b_2, b_3 \in A_x$ such that $b_1 + b_2 = b_3$. Hence, $xb_1 + xb_2 \equiv xb_3 \pmod{p}$, a contradiction.

We claim that there exists $x \in [p-1]$ such that $|A_x| > n/3$. The following fact follows from the unique inverse property when p is a prime:

Fact 7. *For all $y \in \mathbb{Z}_p^*$ and for all $b \in B$, exactly one $x \in \mathbb{Z}_p^*$ satisfies $y \equiv x \cdot b \pmod{p}$.*

It follows from Fact 7 that for all $y \in \mathbb{Z}_p^*$ and for all $b \in B$, $\mathbb{P}_x[y \text{ is mapped from } b] = \frac{1}{p-1}$. Moreover, Fact 7 implies that for all $b \in B$, there exist $|C|$ choices of x such that $x \cdot b \pmod{p} \in C$.

For $b \in B$ and $x \in \mathbb{Z}_p^*$, define

$$\sigma_b^{(x)} = \begin{cases} 1, & \text{if } x \cdot b \pmod{p} \in C \text{ (i.e., } b_i \text{ maps to } C \text{ under } x), \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\mathbb{E}_x\left[\sigma_b^{(x)}\right] = \mathbb{P}_x\left[\sigma_b^{(x)} = 1\right] = \frac{|C|}{p-1} > \frac{1}{3}.$$

Therefore,

$$\mathbb{E}_x\left[|A_x|\right] = \mathbb{E}_x\left[\sum_{b \in B} \sigma_b^{(x)}\right] = \sum_{b \in B} \mathbb{E}_x\left[\sigma_b^{(x)}\right] > \frac{n}{3}.$$

This shows that at least one $x \in \mathbb{Z}_p^*$ has $|A_x| > n/3$. □

2 The Lovász Local Lemma

Let $A_1, \ldots, A_n$ be "bad" events. In probabilistic methods, we would like to show that there is a positive probability that none of the bad events occur. We have the following cases:

- If the events A_i are independent and non-trivial (i.e., $\mathbb{P}[A_i] \neq 1$),

$$\mathbb{P}\left[\bigcup A_i\right] = 1 - \mathbb{P}\left[\bigcap \overline{A_i}\right] = 1 - \prod \mathbb{P}\left[\overline{A_i}\right] < 1.$$

- By the union bound, if $\mathbb{P}[A_i] < 1/n$ for all $i \in [n]$,

$$\mathbb{P}\left[\bigcup A_i\right] \leq \sum \mathbb{P}[[A_i] < 1.$$

Is there anything between the union bound and total independence?

Definition 8. An event A is *independent* of events $B_1, \ldots, B_k$ for all $J \subset [k]$ with $J \neq \emptyset$,

$$\mathbb{P}\left[A \cap \bigcap_{j \in J} B_j\right] = \mathbb{P}[A] \cdot \mathbb{P}\left[\bigcup_{j \in J} B_j\right].$$

[3] We use "$\in_R$" to indicate randomly choosing from a set.

Definition 9. Let $A_1, \ldots, A_n$ be events. The digraph $D = (V, E)$ with $V = [n]$ is the *dependency digraph* of $A_1, \ldots, A_n$ if each A_i is independent of all events A_j that are not neighbors in D.

Theorem 10 (Lovász local lemma)**.** *Let $A_1, \ldots, A_n$ be events such that $\mathbb{P}[A_i] \leq p$ for all $i \in [n]$ with dependency digraph D such that D has maximum degree at most d. If $ep(d+1) \leq 1$,*

$$\mathbb{P}\left[\bigcap_{i=1}^{n} \overline{A_i}\right] > 0.$$

2.1 Revisiting Hypergraph Coloring

Theorem 11. *Let X be a set. Let $S_1, \ldots, S_m \subset X$ such that $|S_i| = \ell$ for all $i \in [m]$ and that each S_i intersects at most d other S_j's. If $e(d+1) \leq 2^{\ell-1}$, then one can 2-color X such that each S_i is not monochromatic.*

Proof. Color each element of X red or blue independently with probability $1/2$. For each $i \in [m]$, let A_i be the event that S_i is monochromatic. As before, $\mathbb{P}[A_i] = 1/2^{\ell-1}$ for all $i \in [m]$. Note that A_i is independent of all A_j's such that $A_i \cap A_j = \emptyset$, so it depends on at most d other A_j's (this gives the degree bound). Since $ep(d+1) = e \cdot 1/2^{\ell-1} \cdot (d+1) \leq 1$ by assumption, then the Lovász local lemma implies that $\mathbb{P}[\bigcap_{i=1}^{m} \overline{A_i}] > 0$. □

2.2 History

Sometimes we are interested in finding a good solution. The partial history of the algorithmic Lovász local lemma is summarized in Table 1.

Beck	$d \leq 2^{\ell/1000}$
Alon	$d \leq 2^{\ell/8}$
⋮	
Moser	$d \leq 2^{\ell-1}/e$
Moser & Tardos	

Table 1: The partial history of the algorithmic Loász local lemma.

2.3 The Moser-Tardos Algorithm

We present the Moser-Tardos algorithm for hypergraph coloring in Algorithm 1.

```
1 2-color X randomly
2 repeat
3     pick monochromatic S_i
4     randomly re-assign colors
5 until proper coloring
```

Algorithm 1: The Moser-Tardos algorithm for hypergraph coloring.

```
// first pass
color each j ∈ X red or blue randomly
S_i is bad if it has at most αℓ red elements or at most αℓ blue elements
B ← {S_i : S_i is bad}
first pass is successful if all connected components of B are at most d^2 log m; if not, retry
  (an edge between A_i and A_j if A_i ∩ A_j ≠ ∅)
// second pass
brute force to fix each connected component without making neighbors monochromatic
```

Algorithm 2: A Beck-like algorithm for hypergraph coloring. The algorithm is given $S_1, \ldots, S_m \subset X$ with $|S_i| = \ell$ for all $i \in [m]$, and assumes that ℓ, d are constant.

2.4 A Beck-like Algorithm

We present an algorithm similar to Beck's and Alon's with stronger assumptions (to be formulated in red) in Algorithm 2.

There are some questions that need answering:

- *Does a good solution exist?*

  We have the following cases:

  - If S_i is not bad and has fewer than $\alpha\ell$ vertices in bad neighbors, then S_i will still be bichromatic after the second pass recoloring.
  - If S_i is not bad and has at least $\alpha\ell$ vertices in bad neighbors, then at least $\alpha\ell$ vertices get recolored. If the vertices are recolored randomly, then $\mathbb{P}[S_i \text{ is monochromatic}] = 2^{-\alpha\ell}$. Using the Lovász local lemma and assuming $2e(d+1) < 2^{\alpha\ell}$, a solution exists.

- *For how many times one has to retry in the first pass?*

  Fact 12. *For each $i \in [m]$, $\mathbb{P}[S_i \text{ is bad}] \leq 2 \cdot 2^{(H(\alpha)-1)\ell}$, where $H(x) = -x \log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function.*

  Let $p = 2 \cdot 2^{(H(\alpha)-1)\ell}$. Then

  $$\mathbb{P}\left[\text{independent sets } S_{i_1}, \ldots, S_{i_k} \text{ are all in } B\right] \leq p^k.$$

  First try. We have

  $$\begin{aligned} &\mathbb{P}[\text{a specific big component of size } s \text{ survives}] \\ \leq\ &\mathbb{P}[\text{a big independent set of size } s' < s \text{ in the big component survives}] \\ \leq\ &p^{s'}. \end{aligned}$$

  Hence,

  $$\mathbb{P}[\text{any big component survives}] \leq (\#\text{ big components}) \cdot p^{s'}.$$

  The problem is that the number of big components can be at most $\binom{n}{s}$, and that $s' = 1$ if the component is a clique. We shall use the following known fact from graph theory:

  Fact 13. *Given a subgraph H of size s in a graph of degree at most Δ, there exists an independent set of size at least $\frac{|H|}{\Delta+1}$.*

```
1 I ← ∅
2 repeat
3     I ← I ∪ {an arbitrary vertex u in H}
4     remove u and all its neighbors from H
5 until H is empty
```

Algorithm 3: A greedy algorithm that finds an independent set in a subgraph H.

Proof. We generate an independent set by Algorithm 3. In each round, the size of I increases by 1, and the number of vertices in H decreases by $\Delta + 1$. Therefore, the number of rounds is at least $\frac{|H|}{\Delta+1}$, so the size of the independent set is at least $\frac{|H|}{\Delta+1}$.

□

Therefore, we can use $s' = \frac{s}{\Delta+1}$.

- *How much time does the brute force in the second pass need?*

  Recall that the size of surviving components (i.e., components in B) is $O(d^2 \log m)$, and that the number of settings to the variables in each surviving component is $(2^\ell)^{O(d^2 \log m)} = m^{O(\ell d^2)}$. Assuming ℓ and d are constant, $O(\ell d^2) = O(1)$.