Class 10

Derandomization Techniques

Warm Up

1. Let $G = (V, E)$ have no self-loops. Show there is a cut with at least $|E|/2$ edges crossing it.

2. Let ϕ be a 3-CNF formula.
   - E.g., $\phi = (x_1 \vee \overline{x_2} \vee x_3) \wedge (x_2 \vee \overline{x_4} \vee x_5) \wedge \cdots$
   - An assignment (e.g., $x_1 \leftarrow 1, x_2 \leftarrow 0, \ldots$) **satisfies** ϕ if $\phi = 1$ when you plug in the assignment.

   Show that ϕ has some assignment σ so that at least $7/8$ of the clauses in ϕ are satisfied.

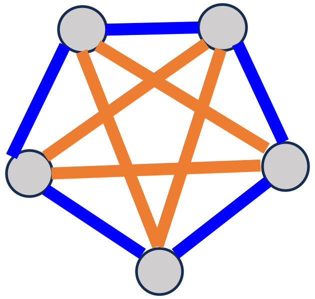
Announcements

- HW4 Due Friday!

Recap

- The probabilistic method!
  - If $\Pr[X = C] > 0$ for some random variable X, then C exists!

- Examples:
  - **Ramsey numbers R_k**: How big does n have to be before a complete graph with n vertices **must** contain a monochromatic k-clique?
    - Showed $R_k > 2^{k/2}$ by showing that a random coloring of $K_{2^{k/2}}$ has a nonzero chance of not having a k-clique!
  - **Independent set:**
    - Any graph with $m > \frac{n}{2}$ edges has an I.S. of size at least $\frac{n^2}{4m}$
    - Remove a bunch of vertices at random; then for each surviving edge remove one of its endpoints.



$R_3 > 5$

Questions?
Warm-up/minilectures/quiz?

Warm Up

1. Let $G = (V, E)$ have no self-loops. Show there is a cut with at least $|E|/2$ edges crossing it.

2. Let ϕ be a 3-CNF formula.
   - E.g., $\phi = (x_1 \vee \overline{x_2} \vee x_3) \wedge (x_2 \vee \overline{x_4} \vee x_5) \wedge \cdots$
   - An assignment (e.g., $x_1 \leftarrow 1, x_2 \leftarrow 0, \ldots$) **satisfies** ϕ if $\phi = 1$ when you plug in the assignment.

   Show that ϕ has some assignment σ so that at least $7/8$ of the clauses in ϕ are satisfied.

Solutions to Warm-up 1

1. Let $G = (V, E)$ have no self-loops. Show there is a cut with at least $|E|/2$ edges crossing it.

Let $G = (V, E)$ be a graph.

- Choose a random $S \subseteq V$
- $$\mathbb{E}\left[\#\text{edges that cross } (S, \bar{S})\right] = \sum_{\{u,v\} \in E} \underbrace{\mathbb{P}\left\{ \begin{matrix} \{u,v\} \text{ crosses} \\ S, \bar{S} \end{matrix} \right\}}_{1/2} = \frac{|E|}{2}$$

$$\Rightarrow \exists \text{ cut so that } \left(\begin{matrix} \#\text{edges} \\ \text{crossing cut} \end{matrix} \right) \geq \frac{|E|}{2}$$

Solutions to Warm-up 2

2. Let ϕ be a 3-CNF formula. Show that ϕ has some assignment σ so that at least 7/8 of the clauses in ϕ are satisfied

Let φ be a 3-CNF formula.

Choose a random assignment σ

$$\mathbb{E}\left[\begin{matrix}\#\text{satisfied clauses}\\ \text{in } \varphi(\sigma)\end{matrix}\right] = \sum_{\text{clauses } C} \underbrace{\mathbb{P}\left[C(\sigma) = \text{TRUE}\right]} = \frac{7}{8} \cdot \left(\begin{matrix}\#\text{ of}\\ \text{clauses}\end{matrix}\right)$$

e.g., $\mathbb{P}\{x_1 \vee x_2 \vee \bar{x}_3 = \text{TRUE}\} = 7/8$

since there's only one out of eight assignments so that it is false.

$\Rightarrow \exists\, \sigma$ s.t. $\varphi(\sigma)$ has $\geq 7/8$ of the clauses satisfied.

Today: Derandomization!

- In those warm-ups (and in the minilectures) you saw ways to prove that something nice exists…but not how to find those nice things efficiently.

- Sometimes, it's possible to turn a probabilistic proof into a deterministic algorithm!

- Today, we'll see one technique called "Derandomization via conditional expectation."

Group Work!

- We will derandomize Warm-Up 1.
  - Find a cut in a graph G=(V,E) that has at least |E|/2 edges crossing it!
  - Note: there is an easy greedy algorithm for this, but the goal here is to illustrate the method of derandomization via conditional expectation.

Group Work!

This is a summary: see the handout for more details and hints!!

1. Say that $G = (V, E)$, and $V = \{v_1, \ldots, v_n\}$.
   - Choose $S \subseteq V$ by including each v_i independently with probability ½.
   - Let X be number of edges crossing the cut $(S, \bar{S})$
   - Show $\mathbb{E}[X \mid v_1 \in S] = |E|/2$
2. Say you've chosen S-membership for $v_1, v_2, \ldots, v_{t-1}$, so that
   $$\mathbb{E}[\, X \mid \text{S\_membership for } v_1, \ldots, v_{t-1} \,] \geq |E|/2$$
   - Show there's a way to assign v_t so that $\mathbb{E}[\, X \mid \text{S\_membership for } v_1, \ldots, v_t \,] \geq |E|/2$
3. Show how to deterministically, efficiently, make that assignment.
4. Design an algorithm to find a set S with at least |E|/2 edges crossing the cut.

Group Work
Problem 1

- Choose $S \subseteq V$ by including each v_i independently with probability ½.
- Let X be number of edges crossing the cut $(S, \bar{S})$
- Show $\mathbb{E}[X \mid v_1 \in S] = |E|/2$

- $\mathbf{E}[X \mid v_1 \in S] = \mathbf{E}[X \mid v_1 \in \bar{S}]$ by symmetry.
- $\frac{1}{2}\mathbf{E}[X \mid v_1 \in S] + \frac{1}{2}\mathbf{E}[X \mid v_1 \in \bar{S}] = \mathbf{E}[X] = \frac{|E|}{2}$
- So both must be equal to $\frac{|E|}{2}$

Let X be the #edges crossing $(S, \bar{S})$

Group Work

Problem 2: Having picked "good" $v_1, \ldots, v_{t-1}$, pick a "good" v_t

$$\frac{|E|}{2} \leq \mathbb{E}\left[X \,\middle|\, \begin{array}{c}\text{choices for} \\ v_1, \ldots, v_{t-1}\end{array}\right] = \frac{1}{2} \cdot \mathbb{E}\left[X \,\middle|\, \begin{array}{c}\text{choices for} \\ v_1, \ldots, v_{t-1}\end{array}, v_t \in S\right] + \frac{1}{2} \cdot \mathbb{E}\left[X \,\middle|\, \begin{array}{c}\text{choices for} \\ v_1, \ldots, v_{t-1}\end{array}, v_t \notin S\right]$$

One of these must be $\geq |E|/2$

Group Work
Problem 3

$$\frac{1}{2}\cdot \mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for} \\ v_1,\dots,v_{t-1}\end{matrix}, v_t \in S\right] + \frac{1}{2}\cdot \mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for} \\ v_1,\dots,v_{t-1}\end{matrix}, v_t \notin S\right]$$

One of these must be $\geq |E|/2$

So we just want to see which of those is the case…

Consider: $$\mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for} \\ v_1,\dots,v_{t-1}\end{matrix}, v_t \in S\right] - \mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for} \\ v_1,\dots,v_{t-1}\end{matrix}, v_t \notin S\right]$$

- If this is positive, then we should put $v_t \in S$.
- Otherwise, put $v_t \in \overline{S}$

$$\mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for}\\ v_1,\dots,v_{t-1}\end{matrix},\ v_t \in S\right] - \mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for}\\ v_1,\dots,v_{t-1}\end{matrix},\ v_t \notin S\right]$$

$$= \sum_{\{u,v\}\in E} \left(\mathbb{P}\left\{ \begin{matrix}\{u,v\} \text{ crosses}\\ (S,\bar{S})\end{matrix} \,\middle|\, \begin{matrix}\text{choices for}\\ v_1,\dots,v_{t-1}\end{matrix},\ v_t \in S \right\} - \mathbb{P}\left\{ \begin{matrix}\{u,v\} \text{ crosses}\\ (S,\bar{S})\end{matrix} \,\middle|\, \begin{matrix}\text{choices for}\\ v_1,\dots,v_{t-1}\end{matrix},\ v_t \notin S \right\} \right)$$

- If $\{u,v\}$ doesn't include v_t, this is 0
- If $\{u,v\} = \{v_i, v_t\}$ for $i > t$, this is $\frac{1}{2} - \frac{1}{2} = 0$
- If $\{u,v\} = \{v_i, v_t\}$ for $i < t$ this is $\begin{cases} +1 & v_i \notin S \\ -1 & v_i \in S \end{cases}$

$$= \left(\begin{matrix}\# i \le t \text{ s.t.}\\ v_i \notin S\end{matrix} \right) - \left(\begin{matrix}\# i \le t \text{ s.t.}\\ v_i \in S\end{matrix} \right)$$

Group Work
Problem 3

$$\frac{1}{2} \cdot \mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for} \\ v_1, \ldots, v_{t-1}\end{matrix}, v_t \in S\right] + \frac{1}{2} \cdot \mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for} \\ v_1, \ldots, v_{t-1}\end{matrix}, v_t \notin S\right]$$

One of these must be $\geq |E|/2$

So we just want to see which of those is the case…

Consider:

$$\mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for} \\ v_1, \ldots, v_{t-1}\end{matrix}, v_t \in S\right] - \mathbb{E}\left[X \,\middle|\, \begin{matrix}\text{choices for} \\ v_1, \ldots, v_{t-1}\end{matrix}, v_t \notin S\right]$$

$$= \left(\begin{matrix}\# i < t \text{ s.t.} \\ v_i \notin S\end{matrix}\right) - \left(\begin{matrix}\# i < t \text{ s.t.} \\ v_i \in S\end{matrix}\right)$$

- If this is positive, then we should put $v_t \in S$.
- Otherwise, put $v_t \in \overline{S}$

Aka, if there are more edges from v_t to $\overline{S}$ than there are to S among the choices we've already made.

Group Work

Problem 4

ALGORITHM

$S = \emptyset$

$T = \emptyset$ // This will become $\bar{S}$

For $t = 1, 2, 3, \ldots, n$:

If v_t has more edges to S than T:

add v_t to T

Else:

add v_t to S

Return S

It's the greedy algorithm!

General Paradigm

Derandomization via conditional expectation

- Suppose you know that $\mathbf{E}[\text{something}]$ is good

- Suppose you can build [something] one choice at a time

- Then assuming that

$$\mathbf{E}[\text{something} \mid \text{choices } 1,2\ \ldots, t-1\,] \text{ is good,}$$

there is a way to make t^{th} choice so that

$$\mathbf{E}[\text{something} \mid \text{choices } 1,2\ \ldots, t\,] \text{ is good.}$$

- If you can find that way to make the t^{th} choice efficiently, you have an algorithm!

Let's try another example!

Derandomizing the second warm-up!

1. Let φ be a 3-CNF formula with n variables and m clauses, and 3 distinct variables in each clause. Use the method of derandomization via conditional expectation to give an efficient (polynomial in n, m) deterministic algorithm to find an assignment to φ so that at least a 7/8-fraction of the clauses are satisfied.
2. (If time) There is also a natural greedy algorithm for this problem:
   - For $i = 1, 2, \ldots, n$:
     - Assign x_i to be whichever value makes the most currently unsatisfied clauses true (breaking ties arbitrarily).

   In the previous example (maximizing the size of a cut), the algorithm we came up with was secretly the natural greedy algorithm. Is your algorithm from the previous part the same as this natural greedy algorithm? Is it better or worse?

Solutions to Group Work

Choose values (TRUE/FALSE) for $x_1, x_2, x_3, \ldots, x_n$ one at a time.

At each step, make sure that $\mathbb{E}\left[\#\text{Sat. clauses} \;\middle|\; \text{choices for } x_1, \ldots, x_t\right] \geq \frac{7m}{8}$

The choice exists by induction:

- base case = warmup exercise
- $\frac{7m}{8} \leq \mathbb{E}\left[\#\text{sat clauses} \;\middle|\; \text{choices for } x_1, \ldots, x_{t-1}\right]$

$$= \frac{1}{2}\mathbb{E}\left[\#\text{sat Clauses} \;\middle|\; \text{choices for } x_1, \ldots, x_{t-1}, x_t = \text{TRUE}\right] + \frac{1}{2}\mathbb{E}\left[\#\text{sat Clauses} \;\middle|\; \text{choices for } x_1, \ldots, x_{t-1}, x_t = \text{FALSE}\right]$$

one of these is $\geq 7m/8$

How to make the choice efficiently?

Want to know when this is larger than $\frac{7m}{8}$

$$\mathbb{E}\left[\begin{array}{c}\#\text{sat.}\\ \text{clauses}\end{array}\middle|\ \begin{array}{c}\text{choices for}\\ x_1,\ldots,x_{t-1}\end{array},\ x_t = \text{TRUE}\right] = \sum_{\text{clauses } C} \mathbb{P}\left\{C = \text{TRUE}\ \middle|\ \begin{array}{c}\text{choices for}\\ x_1,\ldots,x_{t-1}\end{array},\ x_t = \text{TRUE}\right\}$$

This is 1 if the choices have already made C true.
Otherwise it's $1 - 1/2^k$, where $k \in \{0,1,2,3\}$ is the # of free variables left in C.

In particular, we can compute this efficiently.

Time $O(m)$!

Algorithm

For $t = 1, \ldots, n$:

Compute $\mathbb{E}\left[\begin{matrix}\#\text{sat.}\\ \text{clauses}\end{matrix} \;\middle|\; \begin{matrix}\text{choices for}\\ x_1, \ldots, x_{t-1}\end{matrix}, x_t = \text{TRUE}\right]$ Time $O(m)$

If it is at least $7m/8$, set $x_t = \text{TRUE}$

Otherwise, set $x_t = \text{FALSE}$

Time $O(nm)$ total!

Recap: General Paradigm

Derandomization via conditional expectation

- Suppose you know that $\mathbf{E}[\text{something}]$ is good
- Suppose you can build [something] one choice at a time
- Then assuming that

$$\mathbf{E}[\text{something} \mid \text{choices } 1,2 \ \ldots, t-1\,] \text{ is good,}$$

  there is a way to make tth choice so that

$$\mathbf{E}[\text{something} \mid \text{choices } 1,2 \ \ldots, t\,] \text{ is good.}$$

- If you can find that way to make the tth choice efficiently, you have an algorithm!

Next time: More Probabilistic Method!