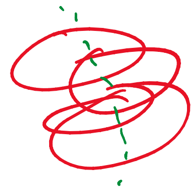
May 5, 2021 8:56 AM

Derandomization Techniques

- brute force
- k-wise indep.
- method of conditional probabilities ←
- or use known derand. constructions!
  (ε-nets, expanders, ...)

Ex 1 Discrepancy



Given set X of n elements $\{x_1, \dots, x_n\}$
& m subsets $S_1, \dots, S_n \subseteq X$.

Want to color the elems as red/blue s.t.

$$\Delta = \max_i \left| (\#\text{ red in } S_i) - (\#\text{ blue in } S_i) \right| \text{ small}$$

Thm $\exists$ coloring s.t. $\Delta \le c_0 \sqrt{n \log n}$

We proved this by taking a rand. coloring
& using Chernoff & union bd

But how to construct such a coloring det'ly?

idea 1 - k-wise indep
$n^{O(k)}$ time but weak bd $n^{\frac{1}{2} + O(\frac{1}{k})}$ ...

new idea - color elems one by one ←
& keep track of prob values ...

let E = event of error
$= \left[\exists i, \left| (\#\text{reds in } S_i) - (\#\text{blues in } S_i) \right| > c_0\sqrt{n \log n} \right]$.

at beginning, know $\Pr(E) < \frac{1}{2}$. ←

Know $\Pr(E) = \frac{1}{2} \Pr(E \mid x_1 \text{ red}) + \frac{1}{2} \Pr(E \mid x_1 \text{ blue})$.

Know $Pr(E) = \frac{1}{2} Pr(E \mid x_1 \text{ red}) +$
$\frac{1}{2} Pr(E \mid x_1 \text{ blue})$.

So, $Pr(E \mid x_1 \text{ red}) < \frac{1}{2}$ ← set x_1 red
or $Pr(E \mid x_1 \text{ blue}) < \frac{1}{2}$. ← set x_1 blue

repeat!

Need to compute cond. prob.
score $= Pr(E \mid \text{given partial coloring})$

But not easy...

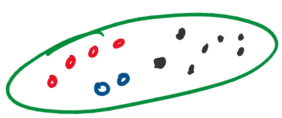
modified idea - maintain an upper bd on cond. prob.

"pessimistic estimator"

let $E_i = [\, |\#\text{reds in } S_i - \#\text{blues in } S_i| > c_0\sqrt{n \log n}\,]$

At beginning, Chernoff yields $\sum_{i=1}^{n} Pr(E_i) < \frac{1}{2}$.

Work with $\widehat{\text{score}} = \sum_{i=1}^{n} \underbrace{Pr(E_i \mid \text{given partial coloring})}_{\text{straightforward to compute}}$



S_i

$\Longrightarrow$ polytime.

Ex2 MAX-SAT

Given CNF formula F with n vars $x_1, \ldots, x_n$ & m clauses $C_1, \ldots, C_m$,

find assignment that maximizes # clauses satisfied.

Approx alg'm?

Trivial Rand. Alg'm:

Trivial Rand. Alg'm:

pick rand. assignment:

$$x_i = \begin{cases} 1 & \text{w. prob } 1/2 \\ 0 & \text{else} \end{cases} \quad \text{(indep'ly)}$$

$$\Pr[C_j \text{ satisfied}] = \begin{cases} \frac{1}{2} & \text{if } C_j \text{ has length } 1 \\ \frac{3}{4} & \text{" " " } 2 \\ \frac{7}{8} & \text{" " } 3 \\ \vdots & \end{cases}$$

$\bar{x}_1 \vee x_2$, $\bar{x}_3$, $x_1 \vee x_2 \vee x_3$, ...

$$\geq \frac{1}{2}$$

$$\Rightarrow E[\#\text{ clauses satisfied}] \geq \frac{1}{2} m$$

$$\geq \frac{1}{2} OPT$$

$\Rightarrow$ expected approx factor $\geq \frac{1}{2}$.

can be derand. by method of cond. probabilities (or cond. expectation)

assign values to vars one by one

keep track of score $= E[\#\text{ clauses satisfied} \mid \text{given partial assignment}]$.

$$= \sum_{j=1}^{m} \Pr[C_j \text{ satisfied} \mid \text{given partial assignment}]$$

$x_1 \vee \bar{x}_2 \vee x_3 \vee x_4$

(but there are more direct det. $\frac{1}{2}$-approx alg'ms)

Better Approx Alg'm:

Solve LP relaxation:

$z_j = 1$ "means" C_j satisfied

$y_i = 1$ "means" x_i true

$$Z_{LP} = \max \sum_{j=1}^{m} z_j$$

$y_i = 1$ means ...

$$z_{LP} = \max \underbrace{\sum_{j=1} z_j}$$

$$\text{s.t. } y_{i_1} + \dots + y_{i_k} \geq z_j$$

if $C_i = x_{i_1} \vee \dots \vee x_{i_k}$

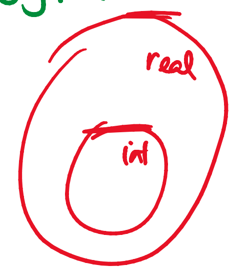
(if C_i contains $\overline{x_{i_1}}$, use $1 - y_{i_1} \dots$)

$$0 \leq y_1, \dots, y_n \leq 1$$
$$0 \leq z_1, \dots, z_m \leq 1$$

real numbers!

$\Downarrow$

linear program



Then $z_{LP} \geq OPT$

max real sol'n val.   max int sol'n val.

But y_i's not integers.

idea - rand. rounding

$$x_i = \begin{cases} 1 & \text{w. prob } y_i \\ 0 & \text{else} \end{cases} \quad \text{(indep.)}$$

$$\Rightarrow \Pr(C_j \text{ not satisfied}) = (1 - y_{i_1}) \cdots (1 - y_{i_k})$$

(say $C_j = x_{i_1} \vee \dots \vee x_{i_k}$)

$$\underset{\text{by GM-AM ineq}}{\leq} \left(1 - \frac{y_{i_1} + \dots + y_{i_k}}{k}\right)^k$$

$$\leq \left(1 - \frac{z_j}{k}\right)^k$$

$$\Rightarrow \Pr[C_j \text{ satisfied}] \geq \underline{1 - \left(1 - \frac{z_j}{k}\right)^k}$$

$$\geq \underline{\left(1 - \boxed{\left(1 - \frac{1}{k}\right)^k}\right)} z_j$$

$f(1 - (1 - \frac{x}{k})^k)$ by Calculus

$\frac{1}{x}\left(1-\left(1-\frac{x}{k}\right)\right)$

by Calculus $(0 \le z_j \le 1)$

$$\geq \left(1-\frac{1}{e}\right) z_j$$

$$\Rightarrow \quad E[\#\text{ clauses sat.}] \geq \left(1-\frac{1}{e}\right)\sum_{j=1}^{m} z_j$$

$$\geq \left(1-\frac{1}{e}\right) z_{LP}$$

$$\geq \left(1-\frac{1}{e}\right) OPT$$

$\Rightarrow$ expected approx factor $1-\frac{1}{e} > \boxed{0.632}$

can be derand. in same way.

Rmk - can be improved

		LP relax. factor	triv. rand.	avg
clause length	1 :	1	$\frac{1}{2}$	$\to$ 3/4
	2 :	3/4	3/4	$\to$ 3/4
	3 :	19/27 ~ 0.704	7/8	> 3/4
	⋮	⋮	⋮	⋮
		$\downarrow$	$\downarrow$	
		1 − 1/e	1	

take max ≥ avg of 2 alg'ms

$\Rightarrow$ approx factor $\geq \boxed{3/4}$

(current record ... 0.7968)