Mar 15 | 1 min checkin

REMINDERS

(•) HW 3 due by 11:59pm tonight
(•) Mini project 2-pg. report due 11:59pm in 2 weeks
(•) No in-person class on Friday

RECAP

(•) "Degree mantra" Every non-zero poly of deg t has $\leq t$ roots
(•) "Claim 3": $f(\alpha)=0 \Rightarrow (X-\alpha)$ divides $f(X)$ i.e. $\exists\, g(X)$ s.t $f(X)=(X-\alpha)\,g(X)$

Plan for today (•) Prove Claim 3
(•) Testing for the next pandemic (aka group testing)

Pf of Claim 3: By fundamental rule of division: (for $f(x)$ & $h(x)$ s.t $\deg(h) \leq \deg f$)
$f(X) = q(X)\cdot h(X) + r(X)$, $\deg(r) < \deg(h)$

$$f(X) = (X-\alpha)\cdot q(X) + r(X)$$
$$\deg(r) < \deg(X-\alpha) = 1$$
$$\Rightarrow \deg(r) = 0$$
$$\Rightarrow r(X) = r \text{ for some } r \in \mathbb{F}_q$$
$$0 = f(\alpha) = (\alpha-\alpha)\cdot q(\alpha) + r(\alpha)$$
$$= 0\cdot q(\alpha) + r = 0 + r = r$$
$$\Rightarrow r = 0 \Rightarrow f(X) = (X-\alpha)\cdot q(X) \quad \blacksquare$$

take $g(X) = q(X)$

Group testing

Consider this scenario
(•) You're in charge of UB testing a new pandemic disease
(•) You're short of testing kits
(•) Say there are N students sample (# kits < N)
(•) Estimate $\leq d$ ~~will~~ have the disease ($d \ll N$)
(•) Goal: Figure out which of the n samples are +ve
while using as few test kits as possible.

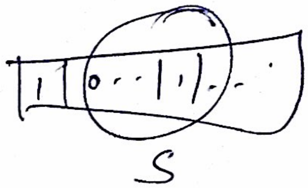
$\leftrightarrow$ Simpler Q: Is there atleast one +ve sample

Q: Can you make this work with 1 test?
A: Yes, "pool" the samples together & test?
(~~test~~ 1 answer if at least one positive sample &
\- 0 o/w).

Formalize problem: Unknown $\bar{x} \in \{0,1\}^N$

$$\text{s.t } \forall i \in [n] \quad x_i = \begin{cases} 1 & \text{if } i^{th} \text{ sample is +ve} \\ 0 & \text{o/w} \end{cases}$$

$wt(\bar{x}) \leq d$



→ Only access $\bar{x}$ via a test/query
→ Each test/query $S \subseteq [N]$

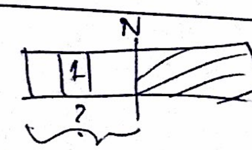
Answer to query S, $A(S) = \begin{cases} 1 & \text{if } \sum_{i \in S} x_i \geq 1 \\ 0 & \text{o/w} \end{cases}$

$\equiv A^{\bar{x}}(S) = \bigvee_{i \in S} x_i$

Problem: figure out tests/queries $S_1, \ldots, S_t$
(Algo) s.t given $A^{\bar{x}}(S_1), \ldots, A^{\bar{x}}(S_t)$, can figure out $\bar{x}$ exactly

---

Q: Solve $d=1$ with $O(\log N)$ queries.



⇒ Binary search $\Theta(\log N)$ adaptive queries.

Adaptive tests: The choice of S_i can depends on
$S_1, A^{\bar{x}}(S_1), S_2, A^{\bar{x}}(S_2), \ldots, S_{i-1}, A^{\bar{x}}(S_{i-1})$

Non-adaptive tests: The choice of $S_1, \ldots, S_t$ has to be fixed without looking at $A(S_1), \ldots, A(S_t)$

& Note: non-adaptive is also adaptive

~~Note~~, Def: $t(d,N)$ min # of non-adaptive tests needed to identify any $\bar{x} \in \{0,1\}^N$ s.t $wt(\bar{x}) \leq d$.

Def: $t^a(d,N)$ ——— adaptive ———

$t^a(1,N) \leq O(\log N)$

Q1: Asymptotic tight bounds on $t(d,N)$ and $t^a(d,N)$?

Q2: ~~Efficient~~ Efficient recovery: $A(S_1), \ldots, A(S_t)$ unknown $\rightarrow$ compute $\bar{X}$ known

Claim: $1 \le t^a(d,N) \le t(d,N) \le N$

since non-adaptive $\subseteq$ adaptive

fix $\forall i \in [N]$

~~Tight~~ Bounds on $t^a(d,N)$ $\leftarrow O(d \log N)$ by running d binary searchs

Thm 1: $t^a(d,N) \le O\left(d \log \frac{N}{d}\right)$

Lemma 1: $t^a(d,N) \ge \Omega\left(d \log \frac{N}{d}\right)$

Pf: $r(\bar{X}) \in \{0,1\}^t$ be the "answer" vector $wt(\bar{X}) \le d$

Claim $\forall\, \bar{X} \ne \bar{y} \in \{0,1\}^N$ s.t $wt(\bar{X}),\ wt(\bar{y}) \le d$

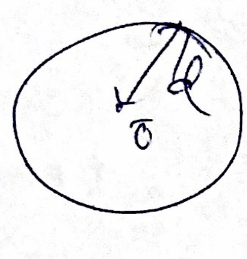
$r(\bar{X}) \ne r(\bar{y})$.

Pf: If $\bar{r} = r(\bar{X}) = r(\bar{y}) \Rightarrow$ given $\bar{r}$ don't know if unknown is $\bar{X}$ or $\bar{y}$.

---

ever $\bar{X}$ s.t $wt(\bar{X}) \le d \Rightarrow$ gives a distinct $r(\bar{X})$

$\Rightarrow$ how many distinct $\bar{r}$ can we have $\in \{0,1\}^t$



$\ge$ as many $\bar{X}$ s.t $wt(\bar{X}) \le d$

$= Vol_2(d,n)$

O.t.o.h # distinct $\bar{r} \le 2^t$

$\Rightarrow 2^t \ge$ ~~log~~ $Vol_2(d,n)$

$\Rightarrow t \ge \log_2\left(Vol_2(d,n)\right) \ldots\ldots \ge d \log \frac{N}{d}$