Feb 9

REMINDERS:

(•) HW 2 due by 11:59pm tonight
(•) HW 3 out by 11:59pm tonight
(•) Yunus OH have changed: M T W 11 - 11:50am
(•) 2 pg project report due 11:59pm on Mar 30 (Wed week after spring break)

PROOF READER for today: Any volunteers? David

GROUP TESTING

Consider scenario:
(•) In charge UB's testing service
(•) You're short of testing kits
(•) Say n student samples
(•) Estimate $\leq d$ COVID +ve samples.
↳ Cannot test all n samples individually
(•) Goal: figure out which of the n samples are COVID +ve while using as few tests as possible.
⇔ Simpler Q: If there is atleast one +ve sample?
Q: Can you do this with 1 test?
A: Yes! "pool" the samples & test!

Formalize the problem: Unknown $\bar{X} \in \{0,1\}^n$ s.t $wt(\bar{X}) \leq d$.

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

→ Only access $\bar{X}$ via test
→ Each test/query $S \subseteq [n]$
Answer to query S, $A(S) = \begin{cases} 1 & \text{if } \sum_{i \in S} X_i \geq 1 \\ 0 & o/w \end{cases}$

or $A(S) = \bigvee_{i \in S} X_i$

(Algorithmic) problem: figure tests/queries $S_1, \ldots, S_t$ s.t given $A(S_1), \ldots, A(S_t)$ can figure of $\bar{X}$ exactly.

Q: Think of set of n tests that are always fine?

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

Non-adaptive tests: The choices of $S_1, \ldots, S_t$ have to be made without looking at any answer $A(S_1) \ldots, A(S_t)$

Notation: $t(d,N)$: min # of non-adaptive tests need to identify any $\overline{X} \in \{0,1\}^n$, $wt(\overline{X}) \le d$
$t^a(d,N)$: ——— adaptive ———

Q1: Asymptotic tight bounds on $t(d,N)$, $t^a(d,N)$
Q2: Efficient recovery $A(S_1), \ldots, A(S_t) \to \overline{X}$

unknown

know this

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

0 tests doesn't we any info about $\overline{X}$.

every non-adaptive set of tests is trivially "adaptive"

pick $\{1\}, \{2\}, \ldots \{N\}$

Change of notation:
$\overline{X} \in \{0,1\}^N$

Bounds on $t^a(d,N)$

Claim: $\forall \; 1 \le d \le N$,
$t^a(d,N) \ge d \cdot \log_2 \frac{N}{d}$

Notation:
Let $r(\overline{X}) = (A(S_1), \ldots, A(S_t))$
vector of answer for $\overline{X}$

Pf: $\overline{X} \ne \overline{y} \in \{0,1\}^N$ $wt(\overline{X}), wt(\overline{y}) \le d$

Assume $\exists$ a valid adaptive set of tests $S_1, \ldots, S_t$

Claim: $r(\overline{X}) \ne r(\overline{y})$

$\hookrightarrow$ If not $r(\overline{X}) = r(\overline{y}) \Rightarrow$ you cannot figure out whether unknown vector is $\overline{X}$ or $\overline{y}$.
(no side-information)

$\to$ Since there at t tests $\Rightarrow$ 2^t distinct answer vectors are possible.

over all vectors $\overline{X} \in \{0,1\}^N$ s.t. $wt(\overline{X}) \le d \Longrightarrow \le 2^t$ possible $r(\overline{X}) \in \{0,1\}^t$
$= Vol_2(d,n)$

$\Rightarrow 2^t \ge Vol_2(d,n) \ge \binom{N}{d} \ge \left(\frac{N}{d}\right)^d$

$$\Rightarrow 2^t \geq \left(\frac{N}{d}\right)^d \Rightarrow t \geq \log_2\left(\frac{N}{d}\right)^d$$

$$= d\log_2\left(\frac{N}{d}\right) \quad \blacksquare$$

Ex! $t^a(d,N) \leq O\left(d \log \frac{N}{d}\right)$