Linear functions:

actually these are homomorphisms

f is "linear" iff $\forall x,y \; f(x)+f(y)=f(x+y)$

will consider $f:\{0,1\}^d \to \{0,1\}$

here, "linear fncts" are the parity fctns

(XOR)

<u>observation</u> $\forall x,y \; f(x)+f(y)=f(x+y)$

iff

$$f(x)=\bigoplus_{i\in S} x_i \quad \text{for some } S\subseteq[d]$$

K-linear fctns:

f is "k-linear" if

(1) linear

(2) depends on $=k$ variables

i.e. $|S|=k$

also called "k-junta fctn

linearity testing:

given $f: \{0,1\}^n \to \{0,1\}$ is f linear?

i.e. $\forall x y$ $f(x) + f(y) = f(x+y)$?

e.g. $\forall x$ $f(x) = 0$ is linear

$\forall x$ $f(x) = 1$ is not

$\forall x$ $f(x) = x \cdot 6$ inner prod of $x \cdot 6$ is linear

Thm Can propertytest linearity in $O(1)$ queries:

linearity test:

Pick random x, y + fail if $f(x) + f(y) \neq f(x+y)$

Proof later lecture

lin ③

Consider functions $f: \{0,1\}^d \to \{0,1\}$ here, domain size $= 2^d \equiv n$

Testing k-linear functions: e.g. $f(x) = \bigoplus_{i \in S} x_i$ s.t. $|S| \le k$

related to testing if fctn is k-junta (depends only on k vars), low Fourier degree, computable by small depth decision trees, . . .

<u>First Algorithm</u>: ("learns" f) wlog assume $f(\bar{0})=0$

~~Given~~

Query f on all $e_i = (00..010..0)$ for $i = 1..d$

$\uparrow$ i^{th} locn

$\uparrow$ $\log n$

$O(d) = O(\log n)$

+ $(00....0)$

if $f(e_i) = 1$ for $\neq k$ i's then fail

else, test if

$$f(x) = \bigoplus_{i \text{ s.t. } f(e_i)=1} x_i$$

for most X via sampling

Can we do better?

What is Communication Complexity?

Setting:

Can talk over a channel

Alice has input $x = x_1 \cdots x_n$

Bob has input $y = y_1 \cdots y_n$

⋮

<u>Goal</u> Compute $f(x,y)$ ← how many bits, rounds of communication required?

<u>examples</u>:

1) $f(x,y) = \left(\bigoplus_i x_i\right) \oplus \left(\bigoplus_i y_i\right)$ • requires 2 bits/round of communication

$A \to B$ $\bigoplus_i x_i$

$B \to A$ $f(x,y)$ (or $\bigoplus y_i$)

2) $f(x,y) = \sum x_i + \sum y_i$ • requires $O(\log n)$ bits

$A \to B$ $\sum x_i$

$B \to A$ $\sum y_i$ (or $f(x,y)$)

can we do better

3) $f(x,y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{o.w.} \end{cases}$ • requires $O(\log n)$ bits

4) $f(x,y) =$ "do $x+y$ agree on any bit?" • requires $\Theta(n)$ bits

Communication Complexity (CC) lower bounds (we have these!)

$\Longrightarrow$ Property testing (PT) lower bounds

Idea give reduction from CC problem to PT problem

a lot of great work done in this area

$\Longrightarrow$ L.B. for C.C. problem yields L.B. for P.T. problem

so we get this almost for free!!

Example:

- A hard C.C problem

SET DISJOINTNESS

Alice

$x \in \{0,1\}^n$

Bob

$y \in \{0,1\}^n$

$$\mathrm{Disj}(x,y) = \bigvee_{i=1}^{n} (x_i \wedge y_i)$$

do A + B agree on any bit?

Known l.b.: $\Omega(n)$ bits of communication required to solve it.

even if allow many rounds, probabilistic protocols

Sparse Set disjointness: A + B have at most k 1's

needs $\Omega(k)$ bits communication

(even if guaranteed that intersect only once or not at all)

How can we use this to lower bound PT problems?

A reduction from sparse set disjointness to PT for k-linearity:

both Alice + Bob can query

$\boxed{\text{shared randomness}}$

Alice

Set A { n bit vector $\{0,1\}^n$ with exactly k 1's in it describing k-linear fctn f (i.e. f is XOR of bits with indices in A)

$\longrightarrow$
$\longleftarrow$

Bob

n bit vector $\{0,1\}^n$ with k 1's describing k-linear fctn g } Set B

Question:

does $h = f \oplus g$

have $2k$- linearity property?

note: if $A \cap B = \phi$ then h is $2k$-linear

if $A \cap B \neq \phi$ then h is j-linear

for $j \leq 2k-2$.

e.g. if $A = \{x_1, x_2\}$ + $B = \{x_3, x_4\}$

$A \cap B = \phi$.

$f = x_1 \oplus x_2$ $g = x_3 \oplus x_4$

$h = x_1 \oplus x_2 \oplus x_3 \oplus x_4$ ← 4 linear

if $A = \{x_1, x_2\}$ $B = \{x_2, x_3\}$

$A \cap B = \{x_2\}$

$f = x_1 \oplus x_2$ $g = x_2 \oplus x_3$

$h = x_1 \oplus \underbrace{x_2 \oplus x_2}_{=1} \oplus x_3$

$= x_1 \oplus x_3$ ← 2 linear

for all x_i in $A \cap B$,

<u>two</u> variables drop out of h

so h is $(k - 2|A \cap B|)$- linear

Fact if $h_1 \neq h_2$ are 2 linear fctns (for any k)

then $\frac{\# x \text{ s.t. } h_1(x) \neq h_2(x)}{2^n} = \frac{1}{2}$

we will prove this in h.w.

$\Rightarrow$ If $A \cap B = \emptyset$, h is $\frac{1}{2}$-far from $2k$-linear

Why is this interesting?

protocol for testing $2k$-linearity of h with q queries $\Rightarrow$ C.C. protocol for set disjointness of A, B

Shared random string which contains random bits for A's queries $\} R$

Bob simulates A's run on R. Bob computes x & then $g(x)$

A runs prop test alg. When needs $h(x) = f(x) \oplus g(x)$:

1) compute $f(x)$
2) ask Bob for $g(x)$
3) output $f(x) \oplus g(x)$ as $h(x)$

what is answer to my next question? $\longrightarrow$

$\longleftarrow$ $g(x)$

$\boxed{\text{Note: Alice doesn't need to sent } x\text{'s}}$

Total communication $= 2q$ bits

$\Rightarrow q = \Omega(k)$

Thm k-linearity testing requires $\Omega(k)$ queries!

Interesting, since linearity testing only needs $O(1)$!