Lecture 14

- Linearity testing & self correcting

- Basics of Fourier Analysis
  on Boolean cube

Linearity Testing

$$f : G \to \cancel{G}\ H$$

G is finite group
H " " "

def. f is "linear" (homomorphism) if

$$\forall\ x, y \in G \qquad f(x) +_H f(y) = f(x +_G y)$$

$+_H$ is "plus" in group H

$+_G$ is "plus" in group G

e.g. $f(x) = x$

$f(x) = ax \bmod p$ for $G = \mathbb{Z}_p$

$f_{\bar{a}}(x) = \sum a_i x_i \bmod 2$ for $G = \mathbb{Z}_2^n$

def f is "ε-linear" if $\exists$ linear g s.t. f & g agree on $\geq 1-\varepsilon$ fraction of inputs.

Notation note that the following are equivalent statements:

- f & g agree on $\geq 1-\varepsilon$ fraction of inputs
- $$\frac{|\{x \mid f(x) = g(x),\ x \in G\}|}{|G|} \geq 1-\varepsilon$$
- $$\Pr_{x \in G}[f(x) = g(x)] \geq 1-\varepsilon$$

How hard is it to test linearity?

do we need to try all $x, y, x+y$ tuples?

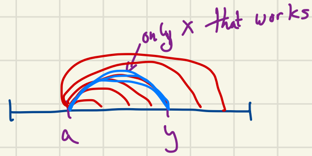
if domain is size n, this requires n^2 tests of $f(x)+f(y) = f(x+y)$

Proposed test: Pick random x, y

Test $f(x) + f(y) = f(x+y)$

repeat how many times?

First let's see some useful things:



A useful observation:

$$\forall\ a, y \in G \qquad \Pr_x[\, y = a + x \,] = \frac{1}{|G|}$$

since only $x = y - a$ satisfies equation

$\Rightarrow$ if pick $x \in_R G$

then $a + x$ is also unif dist in G $\quad (a + x \in_R G)$

(but <u>not</u> independent)

example:

If $G = \mathbb{Z}_2^n$ with operation

$$(a_1 \dots a_n) + (b_1 \dots b_n) = (a_1 \oplus b_1, a_2 \oplus b_2, \dots, a_n \oplus b_n)$$

then $(0110) + (b_1 b_2 b_3 b_4) = (0 \oplus b_1, 1 \oplus b_2, 1 \oplus b_3, 0 \oplus b_4)$

is distributed uniformly if b_i's are

why?
- each coord <u>uniform</u>
- b_i's indep $\Rightarrow$ $a_i \oplus b_i$'s <u>indep too</u>!

Self-Correcting:

also known as "random self-reducibility"

Given $f: G \to G$ s.t. $\exists$ linear $g: G \to G$

s.t. $\Pr_{x \in G}[f(x) = g(x)] \geq 7/8$ $\Leftarrow$ not given g, just f!!!

$\underbrace{}$ this just means f & g agree on $\geq 7/8$ of inputs

Can compute $g(x)$ $\forall x$!

for $i = 1 .. c \log \frac{1}{\beta}$

Pick $y \in_R G$

$answer_i \leftarrow f(y) + f(x-y)$ $\Leftarrow$ note: $x-y$ is unif dist over group by observation

Output most common value for $answer_i$

Claim: $\Pr[\text{output} = g(x)] \geq 1-\beta$

Pf.

$\Pr[f(y) \neq g(y)] \leq 1/8$

$\Pr[f(x-y) \neq g(x-y)] \leq 1/8$

since both y & $x-y$ are uniform over G & by assumption on f

$\therefore \Pr[\underbrace{f(y) + f(x-y)}_{answer_i} \neq \underbrace{g(y) + g(x-y)}_{= g(x) \text{ since } g \text{ is linear}}] \leq 1/4$

by union bound, both are equal with prob $\geq 3/4$

$\Rightarrow$ most common value $= g(x)$ with prob $\geq 1-\beta$ (Chernoff)

How do we test when domain is $\mathbb{Z}_p$?

Do $O(?)$ times

pick $x, y \in_u \mathbb{Z}_p$

If $f(x) + f(y) \neq f(x+y)$ output "fail" & halt

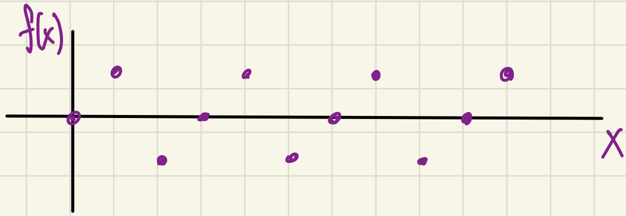
Output "pass"

<u>Possible difficulty:</u> (Coppersmith's example)

Tough function f

$f: \mathbb{Z}_p \to \mathbb{Z}_p$

$\forall x \in \mathbb{Z}_p \quad f(x) = \begin{cases} 1 & \text{if } x \equiv 1 \bmod 3 \\ 0 & 0 \\ -1 & 2 \end{cases}$



Closest linear fctn to f is $g(x) = 0 \ \forall x$

f is "far" from g: $\Pr_x[f(x) \neq g(x)] = 2/3$

but f does pretty well at linearity test:

f fails for $x \equiv y \equiv 1 \bmod 3$ $\quad x+y \equiv 2 \bmod 3$ $\quad 1+1 \neq -1$

$x \equiv y \equiv 2 \bmod 3$ $\quad x+y \equiv 1 \bmod 3$ $\quad -1+-1 \neq 1$

e.g. $x \equiv y \equiv 1 \bmod 3$

$$f(x) + f(y) \stackrel{?}{=} f(\overbrace{x+y}^{2 \bmod 3})$$
$$1 + 1 \neq -1$$

but f passes all other x, y!

$$\Rightarrow \delta_f \equiv \Pr_{x,y}[f(x) + f(y) \neq f(x+y)] = 2/9$$

← passes a lot

"failure probability of test"

& f is $2/3$-far from linear ← very far!

Good news:

$2/9$ is a "threshold"

if $\delta_f < 2/9$, f must be δ_f-close to linear
(known thm)

We will prove stronger bound
for Boolean fctns

need tools: Fourier analysis over Boolean cube

Characterizing linear fctns over Boolean cube

What are linear fctns mapping $\{0,1\}^n \to \{0,1\}$?

inner product $x \cdot y = \sum_{i=1}^{n} x_i y_i \bmod 2$ (XOR)

linear functions on $\{0,1\}^n$: $L_a(x) = a \cdot x$ for fixed $a \in \{0,1\}^n$

how many linear fctns? 2^n

alternate notation: $L_A(x) = \sum_{i \in A} x_i$

for $A \subseteq \{1..n\}$
set of indices that are 1 in $\bar{a}$

The great change of notation:

(less natural, but easier to work with)

$$f : \{\pm 1\}^n \to \{\pm 1\}$$

$$0 \rightsquigarrow +1$$

$$1 \rightsquigarrow -1$$

+	0	1
0	0	1
1	1	0

addition

$\longrightarrow$

×	1	−1
1	1	−1
−1	−1	1

multiplication

$$a \to (-1)^a$$

$$a+b \to (-1)^{a+b} = (-1)^a \cdot (-1)^b$$

now linearity corresponds to

$$f(a) + f(b) = f(a+b) \longrightarrow f(a) \cdot f(b) = f(a \odot b)$$

coordinatewise add $(x_1 \cdots x_n) + (y_1 \cdots y_n) = (x_1+y_1, \ldots, x_n+y_n)$

coordinatewise mult $(x_1 \cdots x_n) \cdot (y_1 \cdots y_n) = (x_1 y_1, \ldots, x_n y_n)$

Linear fctns are now:

$$\boxed{\begin{array}{c} S \subset \{1..n\} \\ \chi_S(x) = \prod_{i \in S} x_i \end{array}}$$

Parity fctns

Express event that test passes as algebraic fctn:

$f(x) \cdot f(y) = f(x \oplus y)$

$\Updownarrow$

$$f(x) \cdot f(y) \cdot f(x \oplus y) = \begin{cases} 1 & \text{if test accepts} \\ -1 & \text{" " rejects} \end{cases}$$

$\Updownarrow$

$f(x) \cdot f(y) \neq f(x \oplus y)$

$$\Updownarrow$$

indicator var

$$\frac{1 - f(x)\, f(y)\, f(x \oplus y)}{2} = \begin{cases} 0 & \text{if accepts} \\ 1 & \text{o.w.} \end{cases}$$

Now we have a new way to express rejection probability:

rejection probability

$$\boxed{\begin{aligned} \delta_f &\equiv \Pr_{x,y}\left[f(x) \odot f(y) \neq f(x \circ y) \right] \\ &= \mathop{E}_{x,y}\left[\frac{1 - f(x) f(y) f(x \odot y)}{2} \right] \end{aligned}}$$