Lecture 21

- Self-correcting for linear fctns
- testing linearity

Linear Functions:

Today: every group is commutative!!

$f: G \to H$ $\qquad$ G, H $\underbrace{\text{finite groups}}_{\text{closure, associative, identity, inverse}}$ with operations $+_G, +_H$

f is "linear" (homomorphism) if

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

Examples of finite groups: $\qquad G = \mathbb{Z}_m$ with operation "+ mod m"

$\mathbb{Z}_m = \{0, 1, \dots m-1\}$

$G = \mathbb{Z}_m^k$ with coordinatewise "+ mod m"

$(x_1 \dots x_k)$ s.t. $x_i \in \mathbb{Z}_m$

Examples of homomorphisms:

$f(x) = x$

$f(x) = 0$

$f(x) = ax \bmod q$ for $G = \mathbb{Z}_q$

$$f_{\bar{a}}(\bar{x}) = \sum a_i x_i \bmod 2 = (x_1 \dots x_n) \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

$G = \mathbb{Z}_2^n \qquad H = \mathbb{Z}_2$

def. f is "linear" (homomorphism) if $\forall\, x, y \in G$ $\quad f(x) +_H f(y) = f(x +_G y)$

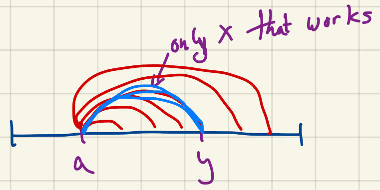
def f is "ε-linear" if $\exists$ linear fctn g s.t. $\leftarrow$ "ε-close to linear"

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

else, f is "ε-far" from linear

A useful observation:

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



since only $x = a - y$ satisfies equation

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

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

example:

If $G = \mathbb{Z}_2^n$ with operation $(a_1 \cdots a_n) + (b_1 \cdots b_n)$

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

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 uniform

• b_i's indep $\Rightarrow$ $a_i \oplus b_i$'s indep

Self-Correcting: also known as "random self-reducibility"

Given f s.t. $\exists$ linear g s.t. $\Pr_x[f(x)=g(x)] \geq 7/8$ ← not given g, just f!!!

Can compute $g(x)$ $\forall x$: (using f)

- for $i = 1 \ldots c\log\frac{1}{\beta}$
  - Pick $y \in_R G$
  - $\text{answer}_i \leftarrow f(y) + f(x-y)$ ← y & $x-y$ are unif in domain
- Output most common value for answer_i

hope: if f agrees with g everywhere

$g(y)+g(x-y) = g(x)$

$\Rightarrow$ $\text{answer}_i = g(x)$

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

Pf.

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

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

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

so each $\text{answer}_i = g(x)$ with prob $\geq 3/4$

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

Linearity Testing

Goal: Given f

- if f linear, pass
- if f ε-far from linear, fail with prob $\geq 2/3$

need to change value of f on $\geq \varepsilon$ fraction of domain

equivalently, $\forall g$ linear $\Pr_{x \in D}[f(x) \neq g(x)] > \varepsilon$

Proposed Test

do ? times:

Pick $x, y \in_u G$

if $f(x) + f(y) \neq f(x+y)$ output "FAIL" + halt

Output PASS

Behavior of Test

f linear $\Rightarrow$ always passes ✓

if f ε-far from linear?

to show (contrapositive):

if f likely to pass then f is ε-linear

(equivalent: f ε-far from linear $\Rightarrow$ f likely to fail)

Plan

- if f ε-close to linear then fctn g you get from self-correcting f:

$$g(x) = \underset{y}{\text{majority}} \left[\underbrace{f(x+y) - f(y)}_{\text{y's vote for value of } g(x)} \right]$$

will be
(1) linear
(2) close to f

- if f not close to linear, then no guarantees on $g(x)$
but if test fails rarely, then you do get guarantees

e.g. • most x satisfy $f(x) = \underset{y}{\text{majority}} \left[f(x+y) - f(y) \right]$

• if x satisfies does $x+y$?

Thm Suppose $\delta \equiv \Pr_{x,y}[f(x)+f(y) \neq f(x+y)] < \frac{1}{\cancel{16}}$. Then f is $\overset{\varepsilon}{\overbrace{2\delta}}$-close to linear

\# times we need to repeat lin test is $\Omega(\frac{1}{\delta})$ so $\gg 16$
$$= \Omega(\tfrac{1}{\varepsilon})$$

Proof let g be the self-correction of f:

def $g(x) = \underset{y}{\text{plurality}}\,[\underbrace{f(x+y) - f(y)}_{y\text{'s vote for } f(x)}]$ ⟸ break ties arbitrarily will show: no ties

def x is $\overset{<\frac{1}{2}}{p}$-good if $\Pr_y[\underbrace{g(x) = f(x+y) - f(y)}_{\text{how many votes did } g(x) \text{ disagree with?}}] > 1-p$

suppose $1-p > \frac{1}{2}$, $g(x)$ defined via clear majority

$\frac{1}{2}$-good x: clear winner

$\delta = \Pr_{x,y}[f(x)+f(y) \neq f(x+y)] < \frac{1}{16}$

def $g(x) = \underset{y}{\text{plurality}}\,[f(x+y)-f(y)]$

def x is $\overset{<1/2}{\rho}$-good if $\Pr_y[g(x) = f(x+y) - f(y)] > \rho$

First: g & f usually agree

Claim 1: for $\rho < 1/2$

$$\Pr_x[\,x \text{ is } \rho\text{-good } \& \ g(x) = f(x)] > 1 - \frac{\delta}{\rho}$$

$\Rightarrow$ fraction of x for which $f+g$ agree is $> 1 - \underset{\rho<1/2}{2\delta} > 7/8$ since $\delta < 1/16$

Pf of Claim 1

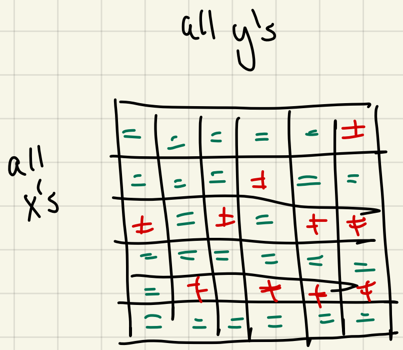
let $\alpha_x = \Pr_y[f(x) \neq f(x+y) - f(y)]$ $\longleftrightarrow$ fraction of "$\neq$" in a row

if $\alpha_x < \rho < 1/2$ then x is ρ-good & $g(x) = f(x)$

$$E_x[\alpha_x] = \frac{1}{|G|} \cdot \sum_{x\in G} \Pr_y[f(x) \neq f(x+y) - f(y)]$$

$$= \Pr_{x,y\in G}[f(x) \neq f(x+y) - f(y)] = \delta$$

So $\Pr[\alpha_x > \rho] \le \delta/\rho$ $\leftarrow$ Markov's $\neq$ $= \left(\frac{\rho}{\delta}\right)\cdot\delta$



$\neq$ if $f(x)+f(y) \neq f(x+y)$

$=$ o.w.

Fraction of $\neq$ in matrix $= \delta$

$E\{$fraction of $\neq$ in row$\} = \delta$

Fraction of rows with $> c\cdot\delta$ is at most $\frac{1}{c}$ Markov's $\neq$

Second: Show g "is a homomorphism" (at least, where it is defined)

$\delta = \Pr_{x,y}[f(x)+f(y) \neq f(x+y)] < \frac{1}{16}$

def $g(x) = \text{plurality}_y [f(x+y) - f(y)]$

def x is ρ-good if $\Pr_y [g(x) = f(x+y) - f(y)] > 1-\rho$ ($\rho < \frac{1}{2}$)

Claim 1: for $\rho < \frac{1}{2}$

$\Pr_x[x \text{ is } \rho\text{-good} \;\&\; g(x) = f(x)] \geq 1 - \frac{\delta}{\rho}$

$\Rightarrow$ fraction of x for which f & g agree is $> 1 - 2\delta > 7/8$

Claim 2 $\rho < 1/4$. If x, y both ρ-good then

(1) $x+y$ is 2ρ-good

(2) $g(x+y) = g(x) + g(y)$

Pf of Claim 2

let $h(x+y) = g(x) + g(y)$

bad events:

$\Pr_z[g(y) \neq f(y+z) - f(z)] < \rho$ since y is ρ good

$\Pr_z[g(x) \neq f(x+(y+z)) - f(y+z)] < \rho$ since x ρ good, $y+z$ unif dist

so $\Pr_z[h(x+y) = g(x) + g(y)$

$= f(\cancel{y}+z) - f(z) + f(x+(y+z)) - f(\cancel{y}+z) = f(x+y+z) - f(z)] \geq 1 - 2\rho$

$> \frac{1}{2}$ since $\rho < 1/4$

union bnd over bad events

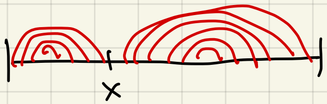
$\Rightarrow g(x+y) = h(x+y)$ by def of g & since $f(x+y+z) - f(z) = h(x+y)$ for $> \frac{1}{2}$ z's

$= g(x) + g(y)$ by def of h ☺

Third: Show that g is actually defined for all x.

Claim 3 $\delta < 1/16$, $\forall x$, x is 4δ-good & $g(x)$ is defined via majority element

Pf of Claim 3



if $\exists\, y$ s.t. y & $(x-y)$ both 2δ-good

then claim 2 $\Rightarrow$ x is 4δ-good

& $g(x) \equiv g(y) + g(x-y)$

To show y exists:

$$Pr_y[y \;\&\; (x-y) \text{ both } 2\delta\text{-good}] > 1 - 2\cdot\frac{\delta}{2\delta} \geq 0$$

both uniform

$$\delta = Pr_{x,y}[f(x)+f(y) \neq f(x+y)] < \frac{1}{16}$$

def $g(x) = \text{plurality}_y\, [f(x+y) - f(y)]$

def x is ρ-good if $Pr_y[g(x) = f(x+y) - f(y)] > 1-\rho$ ($\rho < 1/2$)

Claim 1: for $\rho < 1/2$

$$Pr_x[x \text{ is } \rho\text{-good} \;\&\; g(x) = f(x)] > 1 - \frac{\delta}{\rho}$$

$\Rightarrow$ fraction of x for which f & g agree is $> 1 - 2\delta > 7/8$

Claim 2 $\rho < 1/4$. If x, y both ρ-good then

(1) $x+y$ is 2ρ-good

(2) $g(x+y) = g(x) + g(y)$

claim 1 $\Rightarrow$ $> 1 - \delta/\rho \geq 1 - \frac{1}{16}\cdot 4 = \frac{3}{4}$ of xs are ρ-good

since $Pr > 0$

$\exists y$ s.t. y & $x-y$ both 2δ-good

■

Claim 3 $\Rightarrow$

$\forall x$, $g(x)$ is defined via majority

$\Rightarrow$ for $\rho = 4\delta$, x is ρ-good

claim 2 $\Rightarrow$ g is homomorphism

$\forall x, y$ $g(x)+g(y) = g(x+y)$

claim 1 $\Rightarrow$ f + g agree on $\geq 1-2\delta$ fraction of domain G

So f is 2δ-close to homomorphism $\blacksquare$

$\delta = \Pr_{x,y}[f(x)+f(y) \neq f(x+y)] < \frac{1}{16}$

def $g(x) = \text{plurality}_y\, [f(x+y) - f(y)]$

def x is $\underset{<1/2}{\rho}$-good if $\Pr_y[g(x) = f(x+y) - f(y)] > 1-\rho$

Claim 1: for $\rho < 1/2$

$\Pr_x[\, x$ is ρ-good + $g(x) = f(x)] \geq 1 - \frac{\delta}{\rho}$

$\Rightarrow$ fraction of x for which f + g agree is $> 1-2\delta > 7/8$

g + f are close

Claim 2 $\rho < 1/4$. If x, y both ρ-good then

(1) $x+y$ is 2ρ-good

(2) $g(x+y) = g(x) + g(y)$

claim 1 $\Rightarrow$ $> 1 - \delta/\rho \geq 1 - \frac{1}{16}4 = \frac{3}{4}$ of xs are ρ-good

g is a homomorphism

Claim 3 $\delta < 1/16$. $\forall x$, x is $\underset{1/4}{4\delta}$-good + $g(x)$ is defined via majority element

g is defined everywhere as majority

Improvements: only need $\delta < 2/9$

$\Rightarrow O(9/2)$ tests give const prob of failure instead of $O(16)$

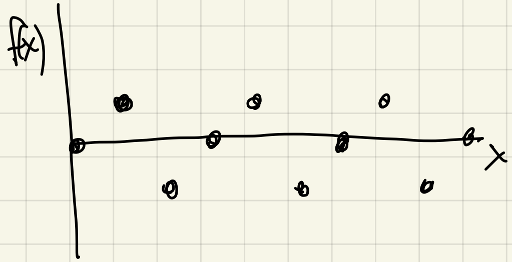
big deal? can lead to improvements in exponents of hardness of approximation results.

over $GF(2)$, can get better δ

in general $2/9$ is tight: (Coppersmith's example) $\leftarrow$ $\frac{2}{9}$ is a "threshold"

$$f(x) = \begin{cases} 1 & \text{if } x \equiv 1 \mod 3 \\ 0 & \text{if } x \equiv 0 \mod 3 \\ -1 & \text{if } x \equiv 2 \mod 3 \end{cases}$$

integers over $\mathbb{Z}$



closest linear fctn: $g(x) \equiv 0$

$\underbrace{\Pr[f(x) = g(x)] = \frac{1}{3}}_{\frac{2}{3}\text{-far}}$

$f(x)+f(y) = 2$
$f(x+y) = -1$

f fails when $\left.\begin{array}{l} x = y = 1 \mod 3 \\ x = y = 2 \mod 3 \end{array}\right\}$ prob $= 2/9$ $\Big\}$ pass with prob $7/9$

else passes