Lecture 21

- Self-correcting for linear fctns
- testing linearity

Linear Functions:

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

f is "linear" (homomorphism) if

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

Examples of finite groups:

$G = \underbrace{\mathbb{Z}_m}_{\{0,1,\ldots,m-1\}}$ with operation "+ mod m"

$G = \underbrace{\mathbb{Z}_m^k}_{(x_1 \cdots x_k)}$ with coordinatewise "+ mod m"

each $x_i \in \{0, \ldots, m-1\}$

Examples of homomorphisms:

$f(x) = x$

$f(x) = 0$

$f(x) = ax \bmod q$

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

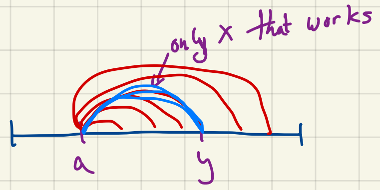
<u>def</u>. f is "linear" (homomorphism) if $\forall\, x, y \in G$ $\quad f(x) +_H f(y) = f(x +_G y)$

<u>def</u> f is "ε-linear" if $\exists$ linear fctn g s.t.

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

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 $x = y - a$ satisfies equation

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

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

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 <u>uniform</u>
- b_i's indep $\Rightarrow$ $a_i \oplus b_i$'s <u>indep</u>

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

Given f s.t. $\exists$ linear g s.t. $\underbrace{\Pr_x[f(x) = g(x)] \geq 7/8}_{\text{this just means } f \,\&\, g \text{ agree on } \geq 7/8 \text{ of inputs}}$ $\Leftarrow$ not given g, just f!!!

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

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

Pick $y \in_R G$

$\text{answer}_i \leftarrow f(y) + f(x-y)$ $\Leftarrow$ 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.

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

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

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

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

$\Rightarrow$ most common 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 $\Pr_{x \in D}[f(x) \neq g(x)] \geq \varepsilon$

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:

$$g(x) = \text{majority}_y \ [\underbrace{f(x+y) - f(y)}_{y\text{'s vote for } g(x)}]$$

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

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

  e.g.
  - most x satisfy $f(x) = \text{majority}_y \ [f(x+y) - f(y)]$
  - if x satisfies does $x+y$?

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

\# times we do lin test needs to be $\Omega(\frac{1}{\delta})$ so $\gg \frac{1}{16}$

$= \Omega(\frac{1}{\varepsilon})$

$\underline{\text{Proof}}$ let g be the self-correction of f:

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

$\underline{\text{def}}$ x is $\underbrace{\overset{<1/2}{p}}_{\text{measure of how much the vote won by}}$-good if $\Pr_y[\underbrace{g(x) = f(x+y) - f(y)}_{> 1-p > \frac{1}{2} \text{ fraction of } y\text{'s agree on vote}}] > 1-p$

$\Longrightarrow$ for $\frac{1}{2}$-good x, $g(x)$ defined via majority element

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

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

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

First: g & f usually agree

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

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

$\Longrightarrow$ fraction of x for which f & g agree

is $> 1 - \underbrace{2\delta}_{\rho < 1/2} > 7/8$

Pf of Claim 1

$$\alpha_x = \Pr_y[f(x) \neq f(x+y) - f(y)]$$

fraction of "≠" in a row

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

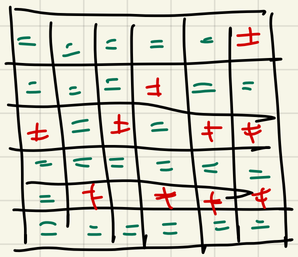
$$E_x[\alpha_x] = \frac{1}{|G|} \cdot \sum_{x \in G} \Pr_y[f(x) \neq f(x+y) - f(y)]$$

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

so $\Pr[\alpha_x > \rho] \leq \delta/\rho$ $= (\rho/\delta) \cdot \delta$

all y's

all x's



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

$=$ o.w.

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

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

Fraction of rows with $> c \cdot \delta$ is at most $1/c$ (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 $\Pr_y [g(x) = f(x+y) - f(y)] > 1-\rho$ ($<\frac{1}{2}$)

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

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

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 is ρ-good & $\underset{\text{fixed}}{y} + \underset{\text{uniform}}{z} \in_R G$ via observation

so $\Pr_z [\, h(x+y) = g(x) + g(y)$
$= f(y+z) - f(z) + f(x+(y+z)) - f(y+z)] = f(x+y+z) - f(z)] > 1 - 2\cdot\rho$ (union bound over bad events) $> \frac{1}{2}$

cancel

$\Rightarrow$ $g(x+y) = h(x+y)$ by def of g since $f((x+y)+z) - f(z)$ is same for $\geq \frac{1}{2}$ of z's & so 2ρ-good

$= g(x) + g(y)$ by def of h ∎

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

Claim 3 $\delta < 1/16$, $\forall x$, x is $\underset{1/4}{4\delta}$-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
& $g(x) = g(y) + g(x-y)$

To show y exists:

$$\Pr_y[\underbrace{y \text{ \& } (x-y)}_{\text{both uniform}} \text{ both } 2\delta\text{-good}] > 1 - \underbrace{\left(\frac{\delta}{2\delta}\right)}_{\text{claim 1}} \cdot \underbrace{2}_{\text{union bnd}} = 0$$

Since $\Pr > 0$, $\exists y$ s.t. x & $(x-y)$ both 2δ-good ■

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

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

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

Claim 1: for $p < 1/2$
$\Pr_x[\,x \text{ is } p\text{-good} \text{ \& } g(x) = f(x)] > 1 - \frac{\delta}{p}$

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

Claim 2 $p < 1/4$. If x, y both p-good then
(1) $x+y$ is $2p$-good
(2) $g(x+y) = g(x) + g(y)$

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

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

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 $\rho < 1/2$

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

$\Rightarrow$ fraction of x for which f & g agree

is $\geq 1-2\delta > 7/8$

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

g is a homomorphism

Claim 3 $\delta < 1/16$. $\forall x$, x is $\underbrace{4\delta}_{1/4}$-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.

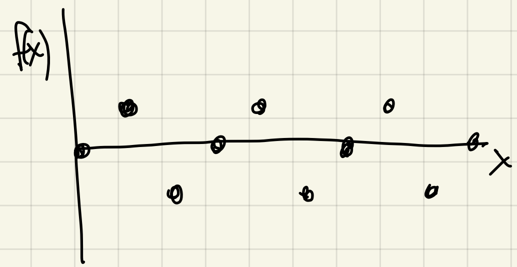
over $GF(2)$, can get better δ

in general $2/9$ is tight: (Coppersmith's example)

$\frac{2}{9}$ is a "threshhold"

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

integers over $\mathbb{Z}$



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

$\Pr[f(x) = g(x)] = \frac{1}{3}$

$\frac{2}{3}$ - 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 $= 2/9$

else passes

$\Big\{$ passes with prob $7/9$