19/10/23

Thm :- [Schwartz- Zippel - de Melo - Lipton]
→ multilinear.

Let $p : \{0,1\}^n \to \mathbb{R}$ of deg d.

Then $\Pr_{x \sim_u \{0,1\}^n} [p(x) \neq 0] \geq \frac{1}{2^d}$

Linearity testing

What is a linear function?

$f : \{0,1\}^n \to \{0,1\}$

$\{0,1\}^n \equiv \mathbb{F}_2^n$

$u, v \in \mathbb{F}_2^n$ $\qquad u \oplus v \in \mathbb{F}_2^n$

$0 \cdot 4 = 0$.

$\Rightarrow$ Defn:- $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ is said to be linear.

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

$\forall\, x, y \in \mathbb{F}_2^n$

$\mathbb{F}_2 = \{ \{0,1\}, +, * \}$

$$\left.\begin{array}{l} 0 \oplus 0 = 0 \\ 1 \oplus 1 = 0 \\ 1 \oplus 0 = 1 \end{array}\right\} \quad \begin{array}{l} 0 \cdot 0 = 0 \\ 1 \cdot 1 = 1. \end{array}$$

(2) $f(x) = \sum_{i \in S} x_i$ for some $S \subseteq [n]$.

$$= \bigoplus_{i \in S} x_i$$

Are these equivalent?

$(2) \Rightarrow (1)$

$$f(x+y) = \sum_{i \in S} (x+y)_i$$

$$= \sum_{i \in S} x_i + y_i$$

$$= \sum_{i \in S} x_i + \sum_{i \in S} y_i$$

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

$1 \Rightarrow 2$

$$x = (x_1, \ldots, x_n) = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

e_i = 0-1 vector of length n

s.t. 1 at i^{th} position

and 0 elsewhere

$$f(x) = f(x_1 e_1 + x_2 e_2 + \ldots + x_n e_n)$$

$$= \sum_{i=1}^{n} x_i f(e_i)$$

$$= \sum_{i \in T} x_i$$

$$T := \{ j \mid f(e_j) = 1 \}$$

⇒ Property testing.

Let $f : \{0,1\}^n \to \{0,1\}$
$g : \{0,1\}^n \to \{0,1\}$

$$\text{dist}(f, g) := \Pr_x [f(x) \neq g(x)]$$

↑ uniform over $\{0,1\}^n$

A property $\mathcal{P}$ is just a subset of functions.

e.g. the set of all linear functions.

$$f : \{0,1\}^n \to \{0,1\}$$

$$f \in ? \ \mathcal{P}$$

$$\mathrm{dist}(f, \mathcal{P}) = \min_{g \in \mathcal{P}} \mathrm{dist}(f, g)$$

when $f \in \mathcal{P}$

$$\mathrm{dist}(f, \mathcal{P}) = 0$$

$$\mathrm{dist}(f, \mathcal{P}) \geq 1 - \varepsilon$$

<u>Defn:</u> A property tester with $q(\varepsilon)$ queries is a randomized algorithm

that has the following proper

(1) if $f \in \mathcal{P}$ then
algorithm accept f
with prob $\geq 2/3$

(2) if $\mathrm{dist}(f, \mathcal{P}) \geq \varepsilon$
then algorithm accepts
f with prob $\leq 1/3$.

(1') $f(x+y) = f(x) + f(y)$
for "most" pairs of
x and y

2′ $f(x) = \sum_{i \in S} x_i$ for "most" inputs x.

2′ $\Rightarrow$ 1′ doable

1′ $\Rightarrow$ 2′ do now!

BLR – Linearity testing

→ Choose $x \sim \mathbb{F}_2^n$ uniformly at random

Choose $y \sim \mathbb{F}_2^n$ "

Define $z = x \oplus y$

Accept if $f(z) = f(x) + f(y)$.

Obs:- if f is linear.

Then $\Pr[\text{BLR accepts } f] = 1$.

Question:- What if $\text{dist}(f, \text{linear})$ is large?

$0 \equiv \text{False} \equiv 1$

$1 \equiv \text{True} \equiv -1$

$\oplus$

$1 \oplus 1 = 0 \qquad -1 \cdot -1 = 1.$

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

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

$$z = (x_1 y_1, x_2 y_2, \ldots, x_n y_n)$$

BLR accepts f

iff $f(x) \cdot f(y) \cdot f(z) = 1$.

$$\Pr_{x,y}[\text{BLR accepts } f]$$

$$= \Pr_{x,y}[f(x) \cdot f(y) \cdot f(z) = 1]$$

$$= \mathbb{E}_{x,y}\left[\frac{1 + f(x) \cdot f(y) \cdot f(z)}{2}\right]$$

$$= \frac{1}{2} + \frac{1}{2} \mathbb{E}_{x,y}[f(x) \cdot f(y) \cdot f(z)]$$

$$\mathbb{E}_{x,y}\left[f(x)\,f(y)\,f(z)\right]$$

$$= \mathbb{E}_{x,y}\left[\left(\sum_{S\subseteq[n]} \hat{f}(S) \prod_{i\in S} x_i\right)\cdot\left(\sum_{T\subseteq[n]} \hat{f}(T)\cdot \prod_{i\in T} y_i\right)\cdot\left(\sum_{U\subseteq[n]} \hat{f}(U)\cdot \prod_{i\in U} z_i\right)\right]$$

$$= \mathbb{E}_{x,y}\left[\sum_{S,T,U\subseteq[n]} \hat{f}(S)\cdot\hat{f}(T)\cdot\hat{f}(U)\cdot \prod_{i\in S} x_i \prod_{i\in T} y_i \cdot \underline{\prod_{i\in U} z_i}\right]$$

$$= \mathbb{E}_{x,y}\left[\sum_{S,T,U\subseteq[n]} \hat{f}(S)\cdot\hat{f}(T)\cdot\hat{f}(U)\cdot \underbrace{\prod_{i\in S} x_i \prod_{i\in T} y_i \prod_{i\in U} x_i y_i}\right]$$

$$= \mathbb{E}_{x,y}\left[\sum_{S,T,U\subseteq[n]} \hat{f}(S)\cdot\hat{f}(T)\cdot\hat{f}(U)\cdot \prod_{i\in \underline{S\Delta U}} x_i \prod_{i\in T\Delta U} y_i\right]$$

symmetric difference

$$= \sum_{S,T,U \subseteq [n]} \hat{f}(S) \cdot \hat{f}(T) \cdot \hat{f}(U) \cdot \mathbb{E}_{x,y}\left[\prod_{i \in S \Delta U} x_i \prod_{i \in T \Delta U} y_i\right]$$

$$\text{if } S \Delta U = \emptyset$$
$$T \Delta U = \emptyset$$
$$\Rightarrow S = T = U$$

$$\underbrace{\mathbb{E}_x\left[\prod_{i \in S \Delta U} x_i\right] \cdot \mathbb{E}_y\left[\prod_{i \in T \Delta U} y_i\right]}_{= 0}$$

$$\text{if } S \Delta U \neq \emptyset$$
$$\text{or } T \Delta U \neq \emptyset$$

$$= \sum_{S \subseteq [n]} \hat{f}(S)^3$$

So $\mathbb{E}_{x,y}\left[f(x)\cdot f(y)\cdot f(z)\right] = \sum_{S\subseteq[n]} \hat{f}(S)^3$

Recall

$\Pr[\text{BLR accepts } f] =$

$$\frac{1}{2} + \frac{1}{2}\,\mathbb{E}_{x,y}\left[f(x)\cdot f(y)\cdot f(z)\right]$$

$$= \frac{1}{2} + \frac{1}{2}\sum_{S\subseteq[n]} \hat{f}(S)^3$$

if $\Pr[\text{BLR accepts } f] \geq 1-\varepsilon$

then $\exists\ S \subseteq [n]$ s.t.

$\text{dist}(f, \chi_S) \leq \varepsilon$

$$\hat{f}(S) = \mathbb{E}_x[f(x) \cdot \chi_S(x)]$$

$$= \Pr_x[f(x) = \chi_S(x)]$$

$$- \Pr_x[f(x) \neq \chi_S(x)]$$

$$= 1 - \Pr_x[f(x) \neq \chi_S(x)]$$

$$- \Pr_x[f(x) \neq \chi_S(x)]$$

$$\geq 1 - 2\,\mathrm{dist}(f, \chi_S).$$

if $\mathrm{dist}(f, \chi_S) \leq \varepsilon$

$$\Rightarrow \hat{f}(S) \geq 1 - 2\varepsilon$$

$$\Pr[\text{BLR accepts}]$$

$$1 - \varepsilon \leq \quad\quad \leq \frac{1}{2} + \frac{1}{2} \sum_{S \subseteq [n]} \hat{f}(S)^3$$

$$2 - 2\varepsilon \leq 1 + \sum_{S \subseteq [n]} \hat{f}(S)^3$$

$$1 - 2\varepsilon \leq \sum_{S \subseteq [n]} \hat{f}(S)^3$$

$$\leq \left\{\max_S \hat{f}(S)\right\} \sum_{T \subseteq [n]} \hat{f}(T)^2$$

$$\leq \max_S \hat{f}(S)$$