Lecture 15

- Basics of Fourier Analysis on Boolean cube (some review, some new)
- Analysis of linearity test
- learning Boolean functions - a model <

Fourier Analysis on Boolean Cube

want basis to describe all fctns. $f: \{\pm 1\}^n \to \{\pm 1\}$

1st idea: "input/output table"

indicator fctns

note: these are orthonormal!!

$$e_a(x) = \begin{cases} 1 & \text{if } x = a \\ 0 & \text{o.w.} \end{cases}$$

then $\forall$ fctns g: $g(x) = \sum_a \overbrace{\underbrace{g(a)}_{\text{scalar}} \underbrace{e_a(x)}_{\text{basis fctn evaluated at } x}}^{\text{can express } g \text{ as lin comb of basis vectors}}$

e.g.

x	$f(x)$
0	1
1	-1

$g(x) = +1 \cdot e_0(x) + (-1) \cdot e_1(x)$

2nd idea:

(Recall) Notation change:

$$\{0,1\} \to \{\pm 1\}$$

$$+ \to \times$$

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

coordinatwise mult

Linear fctns:

$$S \subseteq \{1..n\}$$

for $x \in \{\pm1\}^n$, $\boxed{\chi_S(x) = \prod_{i \in S} x_i}$ parity fctns

define $\boxed{\langle f,g\rangle = \frac{1}{2^n} \sum_{x\in\{\pm1\}^n} f(x)\, g(x)}$ inner product (but normalized)

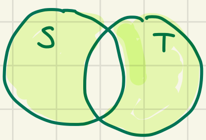
Fact parity (linear) fctns $\{\chi_S\}$ is orthonormal basis w.r.t. inner product!

Proof of fact:

- $\langle \chi_S, \chi_S \rangle = \frac{1}{2^n} \sum_x (\underbrace{\underbrace{\chi_S(x)}_{\pm 1})^2}_{+1} = \frac{2^n}{2^n} = 1$ normal

- if $S \neq T$:

$S \Delta T$:



$$\langle \chi_S, \chi_T \rangle = \frac{1}{2^n} \sum_x \chi_S(x) \cdot \chi_T(x)$$

if $i \in S \cap T$ then $x_i \cdot x_i = 1$ "drops out" so can ignore

$$= \frac{1}{2^n} \sum_x \chi_{\underbrace{S \Delta T}}(x)$$

nonempty since $S \neq T$ so pick $j \in S \Delta T$

$$= \frac{1}{2^n} \sum_{\substack{\text{pairs} \\ x, x^{\oplus j}}} \left(\chi_{S\Delta T}(x) + \chi_{S \Delta T}(x^{\oplus j}) \right)$$

$x^{\oplus j} = x$ with j^{th} bit flipped

$$= \frac{1}{2^n} \sum_{\substack{\text{pairs} \\ x, x^{\oplus j}}} x_j \cdot \underbrace{\prod_{i \in (S \Delta T) \setminus \{j\}} x_i} + \overline{x}_j \underbrace{\prod_{i \in (S\Delta T)\setminus\{j\}} x_i}$$

equal

One is +1 the other is -1 so sum to 0

$$= \frac{1}{2^n} \sum_{\text{pairs}} 0$$

$$= 0$$

Orthogonal!

So $\{\chi_S\}$ is an orthonormal basis

Thm f is uniquely expressible as linear comb. of χ_S.

$$\boxed{\underline{\text{Def.}}\ \hat{f}(S) \equiv \langle f, \chi_S \rangle = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} f(x)\chi_S(x)}$$

Fourier Coefficients of f

$$\boxed{\underline{\text{Thm}}\quad \forall f \quad f(x) = \sum \hat{f}(S)\chi_S(x)}$$

Fourier coeffs of linear fctns:

Fact f linear $\Leftrightarrow$ $\exists\ S \subseteq [n]$ s.t. $\hat{f}(S) = 1$ ← one is really big

& $\forall\ T \neq S$ $\hat{f}(T) = 0$ ← others are 0

e.g. if $f(x) = x_1 \cdot x_2$

$$f(x) = 0 \cdot \chi_\emptyset + 0 \cdot \chi_{\{1\}} + 0 \cdot \chi_{\{2\}} + 1 \cdot \chi_{\{1,2\}}$$

Fourier coeffs characterize distance to linear:

<u>Lemma</u> $\forall S \subseteq [n]$

$$\boxed{\hat{f}(S) = 1 - 2\,\mathrm{dist}(f, \chi_S)}$$

$$= 1 - 2 \Pr_{x \in \{\pm 1\}^n} [f(x) \neq \chi_S(x)]$$

<u>Pf</u> $2^n \cdot \hat{f}(S) = \sum_x f(x)\chi_S(x)$ def of Fourier coeff

$$= \sum_{\substack{x \text{ s.t.} \\ f(x) = \chi_S(x)}} 1 + \sum_{\substack{x \text{ s.t.} \\ f(x) \neq \chi_S(x)}} -1$$

$$= 2^n [1 - \mathrm{dist}(f, \chi_S)] - 2^n [\mathrm{dist}(f, \chi_S)]$$

$$= 2^n (1 - 2 \cdot \mathrm{dist}(f, \chi_S)) \blacksquare$$

example:

$f =$ all -1's

why? see below

$\forall S \neq \emptyset$ $\mathrm{dist}(f, \chi_S) = \frac{1}{2}$ so $\hat{f}(S) = 0$

for $S = \emptyset$: $\mathrm{dist}(f, \chi_\emptyset) = 1$ so $\hat{f}(\emptyset) = -1$

Observation: Any two distinct linear fcfns differ on exactly $\frac{1}{2}$ of inputs

pf. let $f = \chi_T$ for $T \neq S$
$g = \chi_S$

note $\hat{f}(S) = \hat{\chi}_T(S) = \langle \chi_T, \chi_S \rangle$

$$\underset{\substack{\uparrow \\ \text{since} \\ \text{orthonormal}}}{0 =} \langle \chi_T, \chi_S \rangle \underset{\substack{\uparrow \\ \text{lemma}}}{=} 1 - 2\,\mathrm{dist}(\chi_T, \chi_S)$$

$$\Rightarrow \mathrm{dist}(\chi_T, \chi_S) = \frac{1}{2}$$

$\blacksquare$

Very useful tools:

Plancherel's identity

$$\langle f, g \rangle = \left\langle \sum_{S \subseteq [n]} \hat{f}(S)\chi_S, \sum_{T \subseteq [n]} \hat{g}(T)\chi_T \right\rangle$$

$$= \sum_{S,T} \hat{f}(S)\,\hat{g}(T)\,\underbrace{\langle \chi_S, \chi_T \rangle}_{\substack{0 \text{ if } S \neq T \\ 1 \text{ if } S = T}}$$

bilinearity of $\langle \cdot, \cdot \rangle$

$$= \sum_S \hat{f}(S)\,\hat{g}(S)$$

Parseval's identity:

$$\boxed{\langle f, f \rangle = \sum_S \hat{f}(S)^2}$$

"Boolean Parseval's"

if $f: D \to \{\pm 1\}$ $\quad \langle f, f \rangle = \frac{1}{2^n} \sum_x \underbrace{f(x) \cdot f(x)}_{=+1 \text{ since } f \text{ is Boolean}} = \frac{1}{2^n} \cdot 2^n = 1$

$$\text{so} \quad \sum_S \hat{f}(S)^2 = 1$$

Back to linearity testing

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

f "linear" if $\forall x, y \quad f(x) \cdot f(y) = f(x \odot y)$

f "ε-linear" if $\exists$ linear g s.t.

$$\Pr_{x \in G}[f(x) = g(x)] \geq 1 - \varepsilon$$

Linearity test: Pick random x, y
Test $f(x) \cdot f(y) = f(x \odot y)$

$$\delta_f \equiv \Pr_{x,y \in G}[f(x) \cdot f(y) \neq f(x \odot y)]$$ "rejection probability"

Algebraic characterization of rejection probability

$$\delta_f \equiv \Pr_{x,y \in G}[f(x) \cdot f(y) \neq f(x \circ y)]$$

$$= E_{x,y}\left[\frac{1 - f(x) f(y) f(x \circ y)}{2}\right]$$

Note:

We saw this last time by defining indicator var:

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

Analysis of linearity test

if δ_f is small, can we conclude that f is close to linear?

doesn't this contradict Coppersmith's lower bound? this is only for $f: \{\pm1\}^n \to \{\pm1\}$

Yes! rejection probability gives upper bound on distance.

Thm. $f: \{\pm1\}^n \to \{\pm1\}$ is δ_f-close to some linear fctn

Pf.

$$E_{x,y}[f(x)f(y)f(x \circ y)]$$

$$= E_{xy}\left[\left(\sum_S \hat{f}(S)\chi_S(x)\right)\cdot\left(\sum_T \hat{f}(T)\chi_T(y)\right)\cdot\left(\sum_U \hat{f}(U)\chi_U(x\circ y)\right)\right]$$

$$= E_{xy}\Big[\sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U) \underbrace{\chi_S(x)\chi_T(y)\chi_U(x\circ y)}_{\text{what is this?}}\Big]$$

if $S=T=U$: $\chi_S(x)\chi_T(y)\chi_U(x\circ y)$

$$= \prod_{i\in S} x_i \cdot \prod_{i\in S} y_i \cdot \prod_{i\in S} \overbrace{(x\circ y)_i}^{x_i\cdot y_i}$$

$$= \prod_i x_i \cdot y_i \cdot (x_i\cdot y_i) = \prod_i x_i^2 y_i^2 = \prod_i (1\cdot 1) = 1$$

if $\neg(S=T=U)$;

$$E_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x\odot y)]$$

$$= E_{xy}\Big[\prod_{i\in S} x_i \prod_{j\in T} y_j \prod_{k\in U}(x_k\cdot y_k)\Big]$$

$$= E_{xy}\Big[\prod_{i\in S\Delta U} x_i \cdot \prod_{j\in T\Delta U} y_j\Big]$$

$$= E_{xy}\Big[\prod_{i\in S\Delta U} x_i\Big]\cdot E_{xy}\Big[\prod_{j\in T\Delta U} y_j\Big]$$ since x,y indep

assumption $\Rightarrow$ one of these holds so $\rightarrow$

if $S\neq U$ then $S\Delta U\neq\emptyset \Rightarrow E_{xy}\big[\prod_{i\in S\Delta U} x_i\big]=0$

analagously if $T\neq U$, $E_{xy}\big[\prod_{j\in T\Delta U} y_j\big]=0$

$= 0$ so all these terms drop out!

$$E_{x,y}[f(x)f(y)f(x\odot y)]$$

$$= E_{xy}\Big[\Big(\sum_S \hat f(S)\chi_S(x)\Big)\cdot\Big(\sum_T \hat f(T)\chi_T(y)\Big)\cdot\Big(\sum_U \hat f(U)\chi_U(x\odot y)\Big)\Big]$$

$$= E_{xy}\Big[\sum_{S,T,U} \hat f(S)\hat f(T)\hat f(U)\,\chi_S(x)\chi_T(y)\chi_U(x\odot y)\Big]$$

$$= \sum_{S=T=U} \hat f(S)^3$$

$$\leq \max_S \hat{f}(S) \cdot \underbrace{\sum_S \hat{f}(S)^2}_{=1 \text{ by "Boolean Parseval's"}}$$

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

$$= \max_S \left(1 - 2 \cdot \mathrm{dist}(f, \chi_S)\right)$$

$$= 1 - 2 \cdot \min_S \left(\mathrm{dist}(f, \chi_S)\right)$$

So $$\delta_f \geq \frac{1 - \left(1 - 2 \min_S \left(\mathrm{dist}(f, \chi_S)\right)\right)}{2}$$

$$= \min_S \mathrm{dist}(f, \chi_S)$$

$\Rightarrow \exists S$ s.t. f is δ_f-close to χ_S $\blacksquare$