

(FOURIER) ANALYSIS OF BOOLEAN FUNCTIONS

Main reference

- Analysis of Boolean Functions by Ryan O'Donnell
- Lecture 8 of Ryan's CS Theory Toolkit.

Boolean Function: $f: \{0,1\}^n \rightarrow \{0,1\}$

Many many applications of Fourier analysis of these functions:

- Quantum computing
- Communication complexity (Module 5)
- Learning theory
- Hardness of approximation (Module 6)
- Pseudorandomness (Module 5)
- Random Graph theory
- Social choice theory etc...

Notations (flexible)

$$\begin{array}{ccc} f: \{0,1\}^n \rightarrow \{0,1\} & \text{or} & f: \{-1,1\}^n \rightarrow \{-1,1\} \\ \downarrow & & \downarrow \\ \text{domain: } \mathbb{F}_2^n \text{ with XOR operation} & \equiv & \text{domain: for } \pm 1 \in \mathbb{R} \text{ with } \times \text{ operation (multiply)} \\ & \uparrow & \\ & \text{(means equivalent)} & \end{array}$$

- $f: \{-1,1\}^n \rightarrow \{-1,1\}$
- $f: \mathbb{F}_2^n \rightarrow \{-1,1\}$

Goal: look at different representations of a boolean function, that helps analyze it.

Walsh/Hadamard Fourier Transform

(Interpolating boolean fn f with a polynomial)

Evaluation Representation $\Rightarrow$ Polynomial (Multivariate)

Recall: In DFT from last class, the evaluation points were roots of unity and we got univariate polynomials.

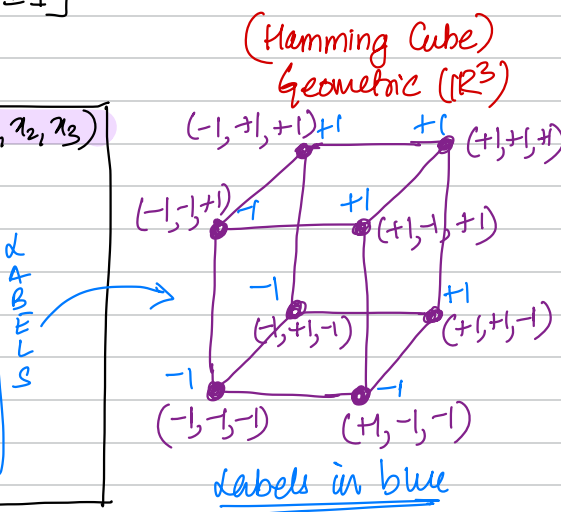
Now, we have evaluation points of f to be $\{\pm 1\}^n$ and the polynomial is multivariate.

EXAMPLE 8: INTERPOLATING MAJORITY FUNCTION

$$f = \text{Maj}_3: \{\pm 1\}^3 \rightarrow \{\pm 1\}$$

Truth table Representation

x_1	x_2	x_3	$\text{Maj}_3(x_1, x_2, x_3)$
+1	+1	+1	+1
+1	+1	-1	+1
+1	-1	+1	+1
-1	+1	+1	+1
+1	-1	-1	-1
-1	+1	-1	-1
-1	-1	+1	-1
-1	-1	-1	-1



Now, we want a polynomial that interpolates these function values.

(Bring in "indicator" polynomials)

$\text{Maj}_3(x_1, x_2, x_3)$: to get this polynomial, find an indicator polynomial for each vertex.
 $x = (x_1, x_2, x_3)$

i.e. for each vertex (i, j, k) , find polynomial $P_{i,j,k}(x) := \begin{cases} 1, & \text{if } x = (i, j, k) \\ 0, & \text{if } x \neq (i, j, k) \end{cases} \rightarrow \text{called indicator poly}$

$$\text{Maj}_3(x) = \sum_{(i,j,k) \in V} P_{i,j,k}(x) \cdot \text{label}(i,j,k)$$

$\uparrow$ vertex set of the Hamming cube $\downarrow$ indicator polynomial $\downarrow$ this is $\text{Maj}_3(i,j,k)$

$$P_{i,j,k}(x) = \left(\frac{1}{2} + i \cdot \frac{1}{2} x_1\right) \left(\frac{1}{2} + j \cdot \frac{1}{2} x_2\right) \left(\frac{1}{2} + k \cdot \frac{1}{2} x_3\right)$$

Note

$$P_{i,j,k}(i,j,k) = \left(\frac{1}{2} + \frac{i^2}{2}\right) \left(\frac{1}{2} + \frac{j^2}{2}\right) \left(\frac{1}{2} + \frac{k^2}{2}\right) = 1 \cdot 1 \cdot 1 = 1$$

& $\neq 0$ ($x \neq (i,j,k)$), either $(x_1 = -i)$ or $(x_2 = -j)$ or $(x_3 = -k)$
 $\Rightarrow$ one of the three terms in the product is 0

$$\text{i.e. } P_{i,j,k}(x) = 0.$$

If you simplify, $\text{Maj}_3(x) = \frac{1}{2} x_1 + \frac{1}{2} x_2 + \frac{1}{2} x_3 - \frac{1}{2} x_1 x_2 x_3$

*** How much time did it take to compute this polynomial?

If $N = 2^n$ then it will take time $O(N^2)$ by above method to get polynomial for Maj_n .

(Similar to last class, we will see how to do this)
 (in time $O(N \log N)$)

Practice Exercise: Try interpolating $\text{Max}_2(x)$ which computes maximum of 2 bits (in ± 1 notation)

Observation

① Since the indicator polynomials are "multilinear",^{why?} meaning that highest power/degree of any x_i is 1. (no x_i^2, x_i^3 , etc...), $\therefore$ the polynomial obtained by interpolating is multilinear.

② This interpolation process also works if you take codomain as $\mathbb{R}$, i.e. $f: \{+1, -1\}^n \rightarrow \mathbb{R}$

Theorem 7.1 FOURIER EXPANSION OF f

Every function $f: \{+1, -1\}^n \rightarrow \mathbb{R}$ can be uniquely expressed as a multilinear polynomial (over $\mathbb{R}$):

$$f(x) = f(x_1, x_2, \dots, x_n)$$

$$= \sum_{S \subseteq [n]} \hat{f}(S) \cdot \prod_{i \in S} x_i$$

called the
fourier coefficient of f
on S . (can $\in \mathbb{R}$)

meaning coefficients
in $\mathbb{R}$

(one monomial
corresponding to
every subset $S \subseteq [n]$)
 $[n] = \{1, 2, \dots, n\}$

This monomial is
denoted by: $\chi_S(x) = \prod_{i \in S} x_i$

($\chi_S: \{+1, -1\}^n \rightarrow \{+1, -1\}$)

(Here $\prod_{i \in \emptyset} x_i = 1$)
the empty
set

χ_S represents the "parity" of x . (i.e. $+1$ or -1 if and only if an odd number of x_i 's are -1) in S .
(for domain $\{0, 1\}^n$, parity is XOR of all bits)

How many different χ_S 's? 2^n

Why is this representation unique? Think & prove!

(Intuition: #evaluations in truth table of $f = 2^n$
#fourier coefficients in the expansion = 2^n)

Example: Majority function

$$\hat{\text{maj}}_3(113) = \frac{1}{2}, \quad \hat{\text{maj}}_3(123) = \frac{1}{2}, \quad \hat{\text{maj}}_3(133) = \frac{1}{2}$$
$$\hat{\text{maj}}_3(1,2,3) = -\frac{1}{2}, \quad \hat{\text{maj}}_3(S) = 0, \text{ for all other } S.$$

∴ we had

$$\text{maj}_3(x) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3 \rightarrow \text{parity fns.}$$

What if the function $f: \mathbb{F}_2^n \rightarrow \mathbb{R}$?

Then $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S(x)$, where now

$$\chi_S(x) = \prod_{i \in S} (-1)^{x_i}$$

Evaluations of function f $\xleftrightarrow{\text{interpolation}}$ Fourier coefficients
 $\xleftarrow{\text{evaluation}}$

Analogous to what we did last time (DFT)

Q: Similar to last time, say given $f: \mathbb{F}_2^n \rightarrow \mathbb{R}$
Given: $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \cdot \chi_S(x)$

What is the evaluation representation of f ?

$$\begin{bmatrix} f(00\dots 0) \\ f(00\dots 1) \\ \vdots \\ f(x) \\ \vdots \\ f(11\dots 1) \end{bmatrix} = \begin{bmatrix} x_0(00\dots 0) & x_{1,y}(00\dots 0) & \dots & x_{n-1}(00\dots 0) \\ \vdots & \vdots & & \vdots \\ \text{xth row} & \text{s.th col.} & x_s(x) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ x_0(11\dots 1) & x_{1,y}(11\dots 1) & \dots & x_{n-1}(11\dots 1) \end{bmatrix} \begin{bmatrix} \hat{f}(\phi) \\ \hat{f}(\phi_1) \\ \hat{f}(\phi_2) \\ \vdots \\ \hat{f}(s) \\ \vdots \\ \hat{f}(\epsilon n) \end{bmatrix}$$

$\underbrace{\hspace{10em}}_{2^n \text{ evaluations}} \quad \underbrace{\hspace{10em}}_{\text{Walsh-Hadamard Fourier Transform}} \quad \underbrace{\hspace{10em}}_{2^n \text{ Fourier coefficients}}$

This matrix called Hadamard matrix, denoted by H_n with x th row & s th col. = $\prod_{i \in S} (-1)^{x_i}$

Let's try to see how this H_{2^n} looks like?

Example $n=2$

$$H_{2^2} = H_4 = \begin{matrix} & \begin{matrix} 00 & 01 & 10 & 11 \end{matrix} \\ \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix} & \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix} \end{matrix}$$

$\begin{matrix} \nearrow & \uparrow \\ \text{take} & \text{take} \\ \text{as } x_2 & \text{as } x_1 \end{matrix}$

These Hadamard matrices have several applications!
Some properties: (has much simpler recursive structure than DFT_{2ⁿ)}

①

Let

$$H_2 = \begin{pmatrix} +1 & +1 \\ +1 & -1 \end{pmatrix} \rightsquigarrow \text{this is also DFT}_2$$

Then, $H_{2^n} = H_2 \otimes H_2 \otimes \dots \otimes H_2 \rightarrow \underline{\text{Exercise prove this}}$

this operation called
Kronecker product

For example
(prove it in general)

$$H_2 \otimes H_2 = \begin{pmatrix} +1 \cdot H_2 & +1 \cdot H_2 \\ +1 \cdot H_2 & -1 \cdot H_2 \end{pmatrix} = \begin{pmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{pmatrix} = H_4$$

② "Fast Walsh-Hadamard Transform" (similar to FFT)
is a divide & conquer algorithm that computes

$$H_{2^n} \cdot \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{2^n-1} \end{bmatrix} \quad \text{in time } O(N \cdot \log N) \text{ operations.}$$

(Here, $N = 2^n$)

? coefficients

③ Similar to DFT's properties, we have

A) Proposition: Inner product of two distinct columns of H_N is 0 & of two same columns is N .
(proof on next page)

B) $\frac{1}{\sqrt{N}} H_N$ is a unitary matrix

(it is like a rotation or a reflection; it preserves vector lengths.)

C) $\frac{1}{N} H_N^T \cdot H_N = I$ (the identity matrix)
i.e. $H_N^{-1} = \frac{1}{N} H_N^T$

D) Also H_N is symmetric, i.e. $H_N^T = H_N$ (Exercise $\Rightarrow$ prove it)
 $\therefore H_N^{-1} = \frac{1}{N} \cdot H_N$
 interpolation evaluation

But, now (2) D) gives a formula for the Fourier coeffs.

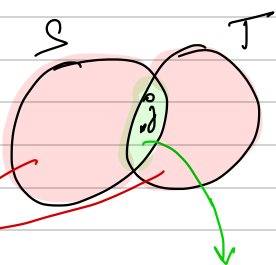
$$\begin{aligned}\hat{f}(\xi) &= \frac{1}{N} \cdot \sum_{x \in \mathbb{F}_2^n} f(x) \cdot \chi_\xi(x) \\ &= \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} f(x) \cdot \chi_\xi(x) \\ &= \mathbb{E}_{x \sim \mathbb{F}_2^n} [f(x) \cdot \chi_\xi(x)]\end{aligned}$$

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

(expectation of $f(x) \cdot \chi_\xi(x)$ for x picked uniformly at random from $\mathbb{F}_2^n$.)

PROOF OF PROPOSITION (2) A)

Inner product of S & T columns = $\sum_{x \in \mathbb{F}_2^n} \chi_S(x) \cdot \chi_T(x)$
 of H_N



$S \Delta T$
(symmetric difference)

• for $j \in S \cap T$
 $(-1)^{x_j} (-1)^{x_j} = (-1)^{2x_j} = 1$

$$\begin{aligned}&= 2^n \cdot \left(\frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} \chi_S(x) \cdot \chi_T(x) \right) \\ &= 2^n \mathbb{E}_{x \sim \mathbb{F}_2^n} [\chi_S(x) \cdot \chi_T(x)] \\ &= 2^n \mathbb{E}_{x \sim \mathbb{F}_2^n} \left[\prod_{i \in S} (-1)^{x_i} \prod_{i \in T} (-1)^{x_i} \right] \\ &= 2^n \mathbb{E}_{x \sim \mathbb{F}_2^n} \left[\prod_{i \in S \Delta T} (-1)^{x_i} \right] \rightarrow \textcircled{1}\end{aligned}$$

Since $S = T \Leftrightarrow S \Delta T = \emptyset$, $\therefore \prod_{i \in S \Delta T} (-1)^{x_i} = 1$

$\Rightarrow$ Inner product of same columns = 2^n

If $S \neq T$, by independence (of picking each x_i randomly)

$$\begin{aligned} \textcircled{1} &= 2^n \prod_{i \in S \Delta T} \mathbb{E}_{x_i \sim \mathcal{U}_2} [(-1)^{x_i}] \\ &= 2^n \prod_{i \in S \Delta T} \left(-1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} \right) \quad \left(\because x_i = \begin{cases} 0 & \text{w.p. } 1/2 \\ 1 & \text{w.p. } 1/2 \end{cases} \right) \\ &= 0 \end{aligned}$$

Definition/Notation

Inner product $\langle \cdot, \cdot \rangle$ on pairs of functions $f, g: \{\pm 1\}^n \rightarrow \mathbb{R}$ is defined as

$$\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} f(x) \cdot g(x) = \mathbb{E}_{x \sim \mathcal{U}_{\pm 1}^n} [f(x) \cdot g(x)]$$

what we proved above is: $\langle \chi_S, \chi_T \rangle = \begin{cases} 1 & \text{if } S = T \\ 0 & \text{if } S \neq T \end{cases}$

& in this notation, the fourier coefficients are:

$$\hat{f}(S) = \langle f, \chi_S \rangle$$

Remark: These fourier coefficients tell you nice properties of the boolean function f

For example, $\hat{f}(\emptyset) = \langle f, \chi_\emptyset \rangle = \langle f, 1 \rangle = \mathbb{E}_x [f(x)]$

↓
the constant
coefficient of fourier expansion gives the
average value of the boolean function f .

Parseval's Theorem : $f: \{-1, +1\}^n \rightarrow \mathbb{R}$

$$\langle f, f \rangle = \mathbb{E}_{x \sim \{-1, +1\}^n} [f(x)^2] = \sum_{S \subseteq [n]} (\hat{f}(S))^2$$

Plancherel's Theorem : $f, g: \{-1, +1\}^n \rightarrow \mathbb{R}$

$$\langle f, g \rangle = \mathbb{E}_{x \sim \{-1, +1\}^n} [f(x) \cdot g(x)] = \sum_{S \subseteq [n]} \hat{f}(S) \cdot \hat{g}(S)$$

Proofs are simple :

$$\langle f, g \rangle = \left\langle \sum_S \hat{f}(S) \cdot \chi_S, \sum_T \hat{g}(T) \cdot \chi_T \right\rangle$$

$$= \sum_{S, T} \hat{f}(S) \cdot \hat{g}(T) \langle \chi_S, \chi_T \rangle$$

$$= \sum_{S \subseteq [n]} \hat{f}(S) \cdot \hat{g}(S) \quad (\text{By Prop. ③A})$$

measures their
correlation

(how similar they are)

If $f, g: \{-1, +1\}^n \rightarrow \{-1, +1\}$ (Boolean-valued)
then

$$\langle f, g \rangle = \mathbb{E}_{x \sim \{-1, +1\}^n} [f(x) \cdot g(x)]$$

$$f(x) \cdot g(x) = \begin{cases} +1 & \text{if } f(x) = g(x) \\ -1 & \text{if } f(x) \neq g(x) \end{cases}$$

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

$$\text{Specifically, } \langle f, f \rangle = 1 \Rightarrow \sum_{S \subseteq [n]} (\hat{f}(S))^2 = 1$$

NEXT CLASS : Some cool applications

EXTRA NOTES (THE LINEAR ALGEBRA VIEW)

$V = \{f : f: [-1, 1]^n \rightarrow \mathbb{R}\}$ forms a vector space over $\mathbb{R}$, with the two operations:

$f+g$: addition of two functions, i.e.
 $(f+g)(x) = f(x) + g(x) \quad \forall x \in [-1, 1]^n$

$a \cdot f$: scalar multiplication by $a \in \mathbb{R}$
 $(a \cdot f)(x) = a \cdot f(x) \quad \forall x \in [-1, 1]^n$

dimension of $V = 2^n$

what fourier expansion says: $f = \sum_{s \in \{0,1\}^n} \hat{f}(s) \chi_s$

$\Rightarrow \mathcal{B} = \{\chi_s : s \in \{0,1\}^n\}$ is a spanning set for V . Moreover it has $2^n (= \dim V)$ parity functions

$\Rightarrow \mathcal{B}$ is a basis for V .

$\downarrow$
Hence, the uniqueness of the fourier expansion!

Now, the definition of inner product on V is as before (with scaling of 2^n)

$$\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in [-1, 1]^n} f(x) g(x) = \mathbb{E}_{x \sim \mathcal{U}_{[-1, 1]^n}} [f(x) \cdot g(x)]$$

& common notation:

$$\|f\|_2 := \sqrt{\langle f, f \rangle} \quad \text{& } \|f\|_p = \mathbb{E}[|f(x)|^p]^{1/p}$$

$\swarrow$ ℓ_2 -norm $\quad \quad \quad \swarrow$ ℓ_p -norm

Proposition 3A $\Rightarrow \mathcal{B}$ is an orthonormal basis