

Lecture 7

*Instructor: Sruthi Sekar**Scribe: Siddhi Pevekar*

1 (FOURIER) ANALYSIS OF BOOLEAN FUNCTIONS

1.1 References

- 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

Notations (flexible):

- $f : \{0, 1\}^n \rightarrow \{0, 1\}$ or $f : \{-1, +1\}^n \rightarrow \{-1, 1\}$
- Domain: $\mathbb{F}_2^n$ with XOR operations (equivalent to ± 1 in $\mathbb{R}$ with multiply operation)
- $f : \{\pm 1\}^n \rightarrow \{\pm 1\}$
- $f : \mathbb{F}_2^n \rightarrow \{\pm 1\}$

Goal: Look at different representations of a Boolean function that helps analyze it.

1.2 Walsh/Hadamard Fourier Transform

(Interpolating Boolean f with a polynomial.)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: Interpolating Majority Function

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

Truth table representation (Hamming Cube):

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 (using "indicator" polynomials).
 $\text{Maj}_3(X)$: to get this polynomial, find an indicator polynomial for each vertex that is for each vertex (i, j, k) , find polynomial

$$p_{i,j,k}(x) = \begin{cases} 1, & \text{if } x = (i, j, k) \\ 0, & \text{otherwise} \end{cases}$$

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

where V is vertex set of this Hamming cube.

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

If you simplify,

$$\text{Maj}_3(x) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$$

How much time did it take to compute this polynomial?

If $N = 2^n$, then it will take time $O(N^2)$ by the above method to get the 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 these indicator polynomials are "multilinear" (the highest power/degree of any x_i is 1), the polynomial obtained by interpolating is multilinear.
2. This interpolation process also works if you take codomain as $\mathbb{R}$, i.e., $f : \{+1, -1\}^n \rightarrow \mathbb{R}$.

1.3 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 $\mathbb{R}$):

$$f(x) = f(x_1, x_2, \dots, x_n) = \sum_{S \subseteq [n]} \hat{f}(S) \prod_{i \in S} x_i$$

where $\hat{f}(S)$ is called the Fourier coefficient of f on S . Let $X_S(x) = \prod_{i \in S} x_i$.

X_S represents the "parity" of x (it is -1 if and only if an odd number of x_i in S are -1).

How many different X_S ? 2^n

Why is this representation unique? Number of evaluations in truth table of $f = 2^n$; number of Fourier coefficients in the expansion $= 2^n$.

Example: Majority function

$$\hat{\text{Maj}}_3(\{1\}) = \frac{1}{2}, \quad \hat{\text{Maj}}_3(\{2\}) = \frac{1}{2}, \quad \hat{\text{Maj}}_3(\{3\}) = \frac{1}{2}$$

$$\text{Maj}_3(\{1, 2, 3\}) = -\frac{1}{2}, \quad \text{Maj}_3(S) = 0 \text{ for all other } S$$

Because we had

$$\text{Maj}_3(x) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$$

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

Then,

$$f(x) = \sum_{S \subseteq [n]} \hat{f}(S) X_S(x)$$

where now

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

Q. Similar to last time, say given f . What is evaluation representation of f ?

$$\begin{bmatrix} f(00 \cdots 0) \\ f(00 \cdots 1) \\ \vdots \\ f(x) \\ \vdots \\ f(11 \cdots 1) \end{bmatrix} = \begin{bmatrix} X_\emptyset(00 \cdots 0) & X_{\{1\}}(00 \cdots 0) & \cdots & X_{\{n\}}(00 \cdots 0) \\ X_\emptyset(00 \cdots 1) & X_{\{1\}}(00 \cdots 1) & \cdots & X_{\{n\}}(00 \cdots 1) \\ \vdots & \vdots & \ddots & \vdots \\ X_\emptyset(x) & X_{\{1\}}(x) & \cdots & X_{\{n\}}(x) \\ \vdots & \vdots & \ddots & \vdots \\ X_\emptyset(11 \cdots 1) & X_{\{1\}}(11 \cdots 1) & \cdots & X_{\{n\}}(11 \cdots 1) \end{bmatrix} \begin{bmatrix} \hat{f}(\emptyset) \\ \hat{f}(\{1\}) \\ \vdots \\ \hat{f}(S) \\ \vdots \\ \hat{f}(\{n\}) \end{bmatrix}$$

2^n evaluations

Walsh-Hadamard
Fourier Transform

2^n fourier
coefficients

Note: x th row, S th column of the middle matrix is $X_S(x)$.

This matrix is called the Hadamard matrix, denoted by H_{2^n} , with x th row, S th column = $\prod_{i \in S} (-1)^{x_i}$.

For example, for $n = 2$:

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

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

Fast Walsh-Hadamard Transform (similar to FFT) is a divide & conquer algorithm that computes $H_{2^n} \cdot [c_\emptyset, c_1, \dots]$ in time $O(N \log N)$ operations ($N = 2^n$).

Properties

1. Inner product of two distinct columns of H_N is 0; of two same columns is N .
2. $\frac{1}{\sqrt{N}} H_N$ is a unitary matrix (it preserves vector length).
3. $\frac{1}{N} H_N^\top H_N = I$ (the identity matrix), i.e., $H_N^{-1} = \frac{1}{N} H_N$.
4. H_N is symmetric, i.e., $H_N^\top = H_N$.

$$\hat{f}(S) = \frac{1}{N} \sum_{x \in \mathbb{F}_2^n} f(x) X_S(x)$$

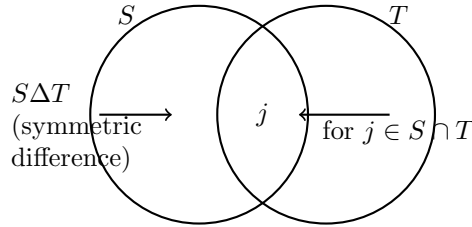
equivalently,

$$\hat{f}(S) = \mathbb{E}_{x \sim \mathbb{F}_2^n} [f(x) X_S(x)]$$

Proof of Proposition

Inner product of S and T columns =

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



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

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

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

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

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

We proved above:

$$\langle X_S, X_T \rangle = \begin{cases} 1 & \text{if } S = T \\ 0 & \text{if } S \neq T \end{cases}$$

And in this notation, the Fourier coefficients are:

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

Remarks:

These Fourier coefficients tell you nice properties of the Boolean function f .

For example,

$$\hat{f}(\emptyset) = \langle f, X_\emptyset \rangle = \langle f, 1 \rangle = \mathbb{E}[f(x)]$$

The constant coefficient of the Fourier expansion gives the average value of the Boolean function f .

Parseval's Theorem

For $f : \{\pm 1\}^n \rightarrow \mathbb{R}$:

$$\langle f, f \rangle = \mathbb{E}_x[f(x)^2] = \sum_{S \subseteq [n]} (\hat{f}(S))^2$$

Plancherel's Theorem

For $f, g : \{\pm 1\}^n \rightarrow \mathbb{R}$:

$$\langle f, g \rangle = \mathbb{E}_x[f(x)g(x)] = \sum_{S \subseteq [n]} \hat{f}(S)\hat{g}(S)$$

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

$$\langle f, g \rangle = \mathbb{E}_x[f(x)g(x)] = \Pr[f(x) = g(x)] - \Pr[f(x) \neq g(x)]$$

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

EXTRA NOTES (THE LINEAR ALGEBRA VIEW)

Let $V = \{f : \{\pm 1\}^n \rightarrow \mathbb{R}\}$ be a vector space over $\mathbb{R}$, with two operations:

- $f + g$: addition of two functions, i.e., $(f + g)(x) = f(x) + g(x)$ for all $x \in \{\pm 1\}^n$
- $a \cdot f$: scalar multiplication by $a \in \mathbb{R}$, $(a \cdot f)(x) = af(x)$ for all $x \in \{\pm 1\}^n$

Dimension of $V = 2^n$.

What Fourier expansion says: $f = \sum_{S \subseteq [n]} \hat{f}(S)X_S$.

$\beta = \{X_S : S \subseteq [n]\}$ is a spanning set for V . Moreover, it has 2^n parity functions, so β is a basis for V . Hence, the uniqueness of the Fourier expansion!

Now, the definition of inner product on V is as before (with scaling of $1/2^n$):

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

Common notation:

- ℓ_2 -norm: $\|f\|_2 := \sqrt{\langle f, f \rangle}$
- ℓ_p -norm: $\|f\|_p = \mathbb{E}[|f(x)|^p]^{1/p}$

Proposition (3A) $\implies \beta$ is an orthonormal basis.