Lecture 0: Fourier Analysis on the Boolean Hypercube

Lecturer: *Sanjeev Arora* Scribe:*Jonathan Ullman*

1 Introduction

We review the basic definitions and applications of the Fourier basis and its applications to functions defined on the boolean hypercube. These are the functions

$$f : \{-1, 1\}^n \longrightarrow \mathbb{R}$$

We will restrict our attention to boolean-valued functions

$$f : \{-1, 1\}^n \longrightarrow \{-1, 1\}$$

The Fourier basis forms an orthonormal basis for the set of functions on the boolean hypercube ($\mathbb{R}^{2^n}$). For all subsets $S \subseteq \{1, \ldots, n\}$ let $\chi_S(x) = \prod_{i \in S} x_i$.

The following are the three basic facts about Fourier analysis, their proofs can be found in standard textbooks:

1. For $S \neq T$, $\chi_S \perp \chi_T$. To see this, $\chi_S \cdot \chi_T = E_x[\chi_S(x)\chi_T(x)] = E_x[\chi_{S \triangle T})(x)] = 0$ if $S \triangle T \neq \emptyset$.

2. We can write any function f in terms of the Fourier basis $f(x) = \sum_S \hat{f}(S)\chi_S(x)$ where the function $\hat{f}$ is called the Fourier transform of f.

3. $\sum_S \hat{f}(S)^2 = 1$ This is also called Parseval's identity (To be more precise, it is a special case of Parseval identity for boolean functions on the boolean hypercube).

We will start by giving a few motivating examples for the theory and then formally analyze one notion of "autocorrelation" called *influence*.

2 Examples of What We Mean By Autocorrelation

Generally Fourier analysis is used in the following framework

1. Consider a function of interest and show that it has some interesting *autocorrelation* properties

2. Show that the Fourier coefficients of such functions satisfy certain properties

3. Show that functions with these Fourier coefficients must belong to some small family of functions.

(i) Fix a function f. Pick a random $i \in \{1, \ldots, N\}$, flip the i^{th} bit in the input, see if the value of f changes. If $\mathbf{Pr}_{x,i}[f(x) = f(x^{(i)})] \approx 1$ then f must depend on very few coordinates.

(ii) **Isoperimetry on the Hypercube**. Consider a graph on $\{-1,1\}^n$ where $x \sim y$ iff $\exists i$ s.t. $y = x^{(i)}$. Consider a cut $(S, \bar{S})$ and its expansion

$$\frac{|E(S,\bar{S})|}{n|S|} \tag{1}$$

Let $f : \{-1,1\}^n \longrightarrow \{-1,1\}$ be $f(x) = 1$ iff $x \in S$.

(iii) **PCP Theorems**. Suppose we have a function f and a test consisting of the OR of three bits. If this test succeeds (i.e. accepts w.h.p.) then f has some nontrivial autocorrelation, thus f has some special structure.

(iv) **Social Choice**. Suppose that n people have opinions $\{-1,1\}^n$, and we need to aggregates those opinions into a single value in $\{-1,1\}$. Then we need to consider a social choice function $f : \{-1,1\}^n \longrightarrow \{-1,1\}$. Generally start by defining some desirable properties, determine which autocorrelations these properties lead to, and show that f must belong to some small/unique/infeasible family of choice functions.

(v) **Phenomena in Random Systems**. For example, pick a random 3CNF formula on n variables and m clauses, empirically we know that at $m \approx 4.3$ there is a sharp transition from satisfiable to unsatisfiable, w.h.p.

Another example, take a random graph $G(n,p)$, around $p \approx \frac{c \log n}{n}$ there is a sharp transition from disconnected to connected.

Fredgut '97 shows formally that if properties do not have a "sharp threshold" then flipping a few bits of the input to f (similar to examples (i) and (ii)) does not change f too often, thus f has some nontrivial autocorrelations, thus f can only depend on a few coordinates. This theorem proves the existence of these threshold phenomena.

3 Influence

Let $f : \{-1,1\}^n \longrightarrow \{-1,1\}$. We define the *influence* of the i^{th} bit of f's input as follows.

Definition 1 $\mathtt{Inf}_i(f) = \mathbf{Pr}_x[f(x) \neq f(x^{(i)})]$. $\mathtt{INF}(f) = \sum_{i=1}^{n} \mathtt{Inf}_i(f)$

Some examples of the influence are

(i) Coordinate function. $f(x) = x_k$. $\mathtt{Inf}_i(f) = 1$ iff $i = k$, $\mathtt{INF}(f) = 1$.

(ii) Parity function. $f(x) = \prod_{i=1}^{n} x_i$. $\mathtt{Inf}_i(f) = 1$, $\mathtt{INF}(f) = n$.

(iii) Majority function. (Assume n odd). $f(x) =$ majority of $\{x_1, \ldots, x_n\}$.

$$\mathbf{Pr}_{x,i}[f(x) \neq f(x^{(i)})] = \mathbf{Pr}[x_{-i} \text{ is balanced}] = \binom{(n-1)}{(n-1)/2} 2^{-(n-1)} \sim c(n-1)^{-1/2}$$

so $\mathtt{Inf}_i(f) \sim \Theta(n^{-1/2})$. $\mathtt{INF}(f) \sim \Theta(n^{1/2})$.

We will use Fourier Analysis to prove the following theorem

Theorem 1
Let f be a balanced function with $\texttt{INF}(f) = 1$. Then f is a coordinate function.

Proof: We will start with the following two Lemmas that characterize the influence operator in terms of the Fourier coefficients of f.

Lemma 2
$\texttt{Inf}_i(f) = \sum_{S:i\in S} \hat{f}(S)^2$

Proof: We can write

$$\texttt{Inf}_i(f) = \mathbf{E}_x[\frac{1}{2} - \frac{1}{2}f(x)f(x^{(i)})] = \frac{1}{2} - \frac{1}{2}\mathbf{E}_x[f(x)f(x^{(i)})]$$

which reduces the Lemma to analyzing $\mathbf{E}[f(x)f(x^{(i)})]$.

$$\begin{aligned} f(x)f(x^{(i)}) &= \left(\sum_S \hat{f}(S)\chi_S(x)\right)\left(\sum_T \hat{f}(T)\chi_T(x^{(i)})\right) \\ &= \sum_{S,T} \hat{f}(S)\hat{f}(T)\chi_S(x)\chi_T(x^{(i)}) \end{aligned}$$

Letting $1^{(i)}$ be the 1 vector with a -1 in the i^{th} position, we can write $x^{(i)} = x * 1^{(i)}$ where $*$ represents component-wise multiplication.

$$= \sum_{S,T} \hat{f}(S)\hat{f}(T)\chi_{S\triangle T}(x)\chi_T(1^{(i)})$$

Taking expectation.

$$\begin{aligned} &\mathbf{E}_x\left[\sum_{S,T} \hat{f}(S)\hat{f}(T)\chi_{S\triangle T}(x)\chi_T(1^{(i)})\right] \\ &\sum_{S,T} \hat{f}(S)\hat{f}(T)\chi_T(1^{(i)})\mathbf{E}_x\left[\chi_{S\triangle T}(x)\right] \end{aligned}$$

where the expectation of $\chi_{S\triangle T}$ is 0 if $S \neq T$, 1 o.w.

$$\begin{aligned} &= \sum_S \hat{f}(S)^2\chi_S(1^{(i)}) \\ &= \sum_{S:i\notin S} \hat{f}(S)^2 - \sum_{S:i\in S} \hat{f}(S)^2 \end{aligned}$$

□

Lemma 3
$\texttt{INF}(f) = \sum_S |S|\hat{f}(S)^2$

Proof: This follows from Lemma 2 and the definition of $\texttt{INF}(f)$. $\square$

Proof of Theorem 1: Since f is balanced, $\hat{f}(\emptyset) = 0$. By assumption $\sum_{|S|\geq 1} |S|\hat{f}(S)^2 = 1$, but Parseval's Identity requires that $\sum_{|S|\geq 1} \hat{f}(S)^2 = 1$ as well. Thus it must be the case that $\hat{f}(S) = 0$ for any S with $|S| > 1$. Thus we can write f as follows

$$f = \sum_i \hat{f}(\{i\})\chi_{\{i\}}(x) = \sum_i \hat{f}(\{i\})x_i$$

Thus we see that f is a boolean hyperplane. Choose an x, clearly $f(x) \in \{-1, 1\}$, but observe that if we flip any bit, i, we get $f(x^{(i)}) = f(x) - x_i\hat{f}(\{i\}) \in \{-1, 1\}$. This forces $\hat{f}(\{i\}) \in \{-1, 0, 1\}$ for all i. However, since Parseval's Identity says that the sum of the square of all the Fourier coefficients is 1, we conclude that there exists a coordinate x_i s.t. $\hat{f}(\{i\}) = \pm 1$, and for every $j \neq i$, $\hat{f}(\{j\}) = 0$. Thus

$$f = x_i$$

or

$$f = -x_i$$

$\square$

Remark 1 This theorem can be generalized to show that if $\texttt{INF}(f) = 1 + \epsilon$ then f is very close to a coordinate function in some sense.