6.842 Randomness and Computation March 28, 2022

Lecture 15: Basics of Fourier Analysis on the Boolean Cube

Lecturer: Ronitt Rubinfeld *Scribe: Hristo Papazov*

1 Introduction.

In today's lecture, we cover

- Basics of Fourier Analysis on the Boolean Cube;
- Analysis of Linearity Testing.

Recall that, strictly speaking, a Boolean function on the $n-$dimensional Boolean cube is defined as a mapping $f : \{0,1\}^n \to \{0,1\}$. However, using the group homomorphism

$$\varphi : \{0, 1, +\} \to \{1, -1, \times\}$$
$$x \mapsto (-1)^x,$$

we can alternatively consider the often more convenient multiplicative analog of a Boolean function $f' = \varphi \circ f \circ \varphi^{-\otimes n} : \{\pm 1\}^n \to \{\pm 1\}$.

2 Fourier Analysis on the Boolean Cube

Without further ado, we jump straight into finding a suitable basis for the space of all Boolean Functions $f : \{\pm 1\}^n \to \{\pm 1\}$.

1^{st} **Idea:** We could consider the indicator functions $e_a : \{\pm 1\}^n \to \mathbb{R}$ for every $a \in \{\pm 1\}^n$ such that

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

Notice that these indicator functions actually form the standard orthonormal basis of $\mathbb{R}^{2^n}$. Clearly, all Boolean functions g can be expressed as the sum

$$g(x) = \sum_{a \in \{\pm 1\}^n} g(a) e_a(x).$$

All in all, this could be a useful representation of Boolean functions, but since we would like to apply Fourier Analysis to Linearity Testing, we would like to have a basis of linear functions.

2^{nd} **Idea:** So, we could instead consider the parity functions $\{\chi_S\}_{S \subseteq [n]}$ such that for every $x \in \{\pm 1\}^n$,

$$\chi_S(x) = \prod_{i \in S} x_i,$$

whenever $S \neq \emptyset$ and $\chi_\emptyset \equiv 1$. Clearly, the range of each χ_S is $\{\pm 1\}$, so the parity functions are Boolean functions. Moreover, they are also linear. Indeed, let $S \subseteq [n]$ and let $x, y \in \{\pm 1\}^n$. Then,

$$\begin{aligned} \chi_S(x \odot y) &= \prod_{i \in S} x_i y_i \\ &= \prod_{i \in S} x_i \prod_{i \in S} y_i \\ &= \chi_S(x) \chi_S(y). \end{aligned}$$

Definition 1 *For $f, g : \{\pm 1\}^n \to \mathbb{R}$, we define their* ***inner product*** *as* $\langle f, g \rangle \stackrel{\text{def}}{=} \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} f(x)g(x)$.

With this definition of inner product on the function space $\mathbb{R}^{\{\pm 1\}^n}$, we are ready to state our first fact about the parity functions.

Fact 2 *The parity functions χ_S form an orthonormal basis of $\mathbb{R}^{\{\pm 1\}^n}$ with respect to the aforementioned inner product.*

Proof

- **Normality:** Notice that $\langle \chi_S, \chi_S \rangle = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \chi_S(x)^2 = 1, \forall S \subseteq [n]$.
- **Orthogonality:** Let $S \neq T$ be two subsets of $[n]$ and fix $x \in \{\pm 1\}^n$. Then,

$$\chi_S(x)\chi_T(x) = \prod_{i \in S} x_i \prod_{j \in T} x_j = \prod_{i \in S \cap T} x_i^2 \prod_{j \in S \triangle T} x_j = \prod_{j \in S \triangle T} x_j = \chi_{S \triangle T}(x).$$

 Now, $S \triangle T \neq \emptyset$, so there exists a number $k \in S \triangle T$. For any bitstring $x \in \{\pm 1\}^n$, we define $x^{\oplus k}$ as x with the k^{th} bit flipped. Now,

$$\begin{aligned}
\langle \chi_S, \chi_T \rangle &= \frac{1}{2^n} \prod_{x \in \{\pm 1\}^n} \chi_S(x)\chi_T(x) \\
&= \frac{1}{2^n} \prod_{x \in \{\pm 1\}^n} \chi_{S \triangle T}(x) \\
&= \frac{1}{2^{n+1}} \prod_{x, x^{\oplus k} \in \{\pm 1\}^n} \chi_{S \triangle T}(x) + \chi_{S \triangle T}(x^{\oplus k}) \\
&= \frac{1}{2^{n+1}} \prod_{x, x^{\oplus k} \in \{\pm 1\}^n} (x_k + x_k^{\oplus k}) \prod_{i \notin S \triangle T \setminus \{k\}} x_i \\
&= 0.
\end{aligned}$$

Hence, we proved that the family of parity functions $\{\chi_S\}_{S \subseteq [n]}$ is orthonormal with respect to $\langle \cdot \rangle$. Since $\mathbb{R}^{\{\pm 1\}^n}$ is of dimension 2^n and there are 2^n parity functions in total, it follows that $\{\chi_S\}_{S \subseteq [n]}$ is an orthonormal basis of $\mathbb{R}^{\{\pm 1\}^n}$. ∎

Corollary 3 *Every Boolean function $f : \{\pm 1\}^n \to \{\pm 1\}$ is uniquely expressible as a linear combination of parity functions.*

If the Boolean function f has the following representation in the parity basis:

$$f(x) = \sum_{S \subseteq [n]} \alpha_S \chi_S(x),$$

then $\forall S \subseteq [n], \alpha_S = \langle f, \chi_S \rangle$ by orthonormality.

Definition 4 *For every $S \subseteq [n]$, we define the $S-$**Fourier coefficient** of $f \in \mathbb{R}^{\{\pm 1\}^n}$ to be*

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

Therefore, $\forall f \in \mathbb{R}^{\{\pm 1\}^n}$,

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

3 Fourier Coefficients of Linear Functions

We begin to build the machinery that will be applied to Linearity Testing.

Fact 5 *A Boolean function $f : \{\pm 1\}^n \to \{\pm 1\}$ is linear $\iff \exists S \subseteq [n]$ such that $\hat{f}(S) = 1$ and $\forall T \subseteq [n]$ such that $T \neq S$, $\hat{f}(T) = 0$. Equivalently, the parity functions χ_S are the only linear Boolean functions.*

Proof The proof of this fact rests on a simple counting argument. Recall that there is a bijection between linear "multiplicative" and linear "additive" Boolean functions as discussed in the introduction. Each "additive" linear function $f : \{0,1\}^n \to \{0,1\}$ is uniquely determined by its values at $\mathbb{1}_i$ for $i = 1, 2, \ldots, n$. Thus, there are a total of 2^n linear Boolean functions. We already saw that the 2^n parity functions are linear, which proves that all linear functions are parity functions. ■

The next lemma shows the connection between Fourier coefficients and distance to linearity. Recall that for two functions $f, g \in \mathbb{R}^{\{\pm 1\}^n}$,

$$\text{dist}(f, g) = \mathbb{P}(f \neq g) = \frac{\#x \text{ s.t. } f(x) \neq g(x)}{2^n}.$$

Lemma 6 *For every $S \subseteq [n], f : \{\pm 1\}^n \to \{\pm 1\}$, $\hat{f}(S) = 1 - 2\mathit{dist}(f, \chi_S)$.*

Proof Fix $S \subseteq [n]$. Then,

$$\begin{aligned}
2^n \hat{f}(S) &= \sum_{x \in \{\pm 1\}^n} f(x)\chi_S(x) \\
&= \sum_{x : f(x) = \chi_S(x)} 1 - \sum_{x : f(x) \neq \chi_S(x)} 1 \\
&= \left(\sum_{x \in \{\pm 1\}^n} 1 \right) - 2 \left(\sum_{x : f(x) \neq \chi_S(x)} 1 \right) \\
&= 2^n - 2^{n+1}\text{dist}(f, \chi_S).
\end{aligned}$$

■

The lemma we just proved constitutes one of the reasons we consider the "multiplicative" version of Boolean functions.

Observation 7 *Any two distinct linear functions χ_S and χ_T differ on exactly half of the input.*

Proof Since the parity functions form an orthonormal basis,

$$0 = \langle \chi_S, \chi_T \rangle = 1 - 2\text{dist}(\chi_S, \chi_T).$$

Hence, $\text{dist}(\chi_S, \chi_T) = 1/2$. ■

So, every linear function χ_S where $S \neq \emptyset$ differs from $\chi_\emptyset \equiv 1$ on exactly half of the input. Hence, $\mathbb{E}[\chi_S] = 0, \forall S \neq \emptyset$.

We continue our discussion with two useful identities.

Theorem 8 (Plancharel's Indentity) *Let $f, g \in \mathbb{R}^{\{\pm 1\}^n}$. Then,*

$$\langle f, g \rangle = \sum_{S \subseteq [n]} \hat{f}(S)\hat{g}(S).$$

Proof The proof is immediate from orthonormality:

$$\begin{aligned}\langle f, g\rangle &= \langle \sum_{S\subseteq[n]} \hat{f}(S)\chi_S, \sum_{T\subseteq[n]} \hat{g}(T)\chi_T\rangle \\ &= \sum_{S,T\subseteq[n]} \hat{f}(S)\hat{g}(T)\langle\chi_S, \chi_T\rangle \\ &= \sum_{S\subseteq[n]} \hat{f}(S)\hat{g}(S).\end{aligned}$$

■

When we consider the case $f = g$, we get the following corollary.

Corollary 9 (Parseval's Indentity) *Let $f \in \mathbb{R}^{\{\pm 1\}^n}$. Then,*

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

And if we apply Parseval's Identity to the Boolean function case, we obtain "Boolean Parseval's":

$$\boxed{\sum_{S\subseteq[n]} \hat{f}(S)^2 = 1.}$$

Indeed, when f ranges over $\{\pm 1\}$, $\langle f, f\rangle = \frac{1}{2^n}\sum_{x\in\{\pm1\}^n} f(x)^2 = 1$.

4 Linearity Testing

Throughout this section, we will only work with "multiplicative" Boolean functions.

Definition 10 *We will say that the Boolean function f is $\varepsilon-$**linear** if there exists a linear Boolean function g such that $\mathbb{P}_{x\in\{\pm1\}^n}(f = g) \geq 1 - \varepsilon$. Equivalently, f is $\varepsilon-$**linear** when dist$(f, \chi_S) \leq \varepsilon$ for some $S \subseteq [n]$, which is in turn equivalent to $\hat{f}(S) \geq 1 - 2\varepsilon$.*

We will use the following linearity test for a Boolean function f.

Linearity Test: Pick uniformly at random and independently $x, y \in \{\pm1\}^n$. Test if $f(x)f(y) \stackrel{?}{=} f(x \odot y)$.

The rejection probability of this test is defined as

$$\begin{aligned}\delta_f &\stackrel{\text{def}}{=} \mathbb{P}_{x,y}(f(x)f(y) \neq f(x\odot y)) \\ &= \mathbb{E}_{x,y}[\frac{1 - f(x)f(y)f(x\odot y)}{2}].\end{aligned}$$

We proceed to state and prove the main result of this lecture.

Theorem 11 *Every Boolean function $f : \{\pm1\}^n \to \{\pm1\}$ is δ_f-linear.*

Proof The proof strategy is to evaluate the rejection probability δ_f in terms of the Fourier coefficients of f, which are tied to the linearity distance as we saw. Without further ado,

$$\begin{aligned}
\mathbb{E}_{x,y}\left[f(x)f(y)f(x\odot y)\right] &= \mathbb{E}_{x,y}\left[(\sum_S \hat{f}(S)\chi_S(x))(\sum_T \hat{f}(T)\chi_T(y))(\sum_U \hat{f}(U)\chi_U(x\odot y))\right] \\
&= \mathbb{E}_{x,y}\left[\sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U)\chi_S(x)\chi_T(y)\chi_U(x\odot y)\right] \\
&= \mathbb{E}_{x,y}\left[\sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U)\chi_S(x)\chi_T(y)\chi_U(x)\chi_U(y)\right] \\
&= \sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbb{E}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x)\chi_U(y)] \\
&= \sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbb{E}_x[\chi_S(x)\chi_U(x)]\mathbb{E}_x[\chi_S(y)\chi_U(y)] \\
&= \sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U)\langle\chi_S(x),\chi_U(x)\rangle\langle\chi_S(y),\chi_U(y)\rangle \\
&= \sum_S \hat{f}(S)^3,
\end{aligned}$$

where we used the linearity of expectation, the independence of x and y, and the orthonormality of the parity functions. Hence,

$$\mathbb{E}_{x,y}\left[f(x)f(y)f(x\odot y)\right] \leq \left(\max_S \hat{f}(S)\right)\sum_S \hat{f}(S)^2 = \max_S \hat{f}(S)$$

by Boolean Parseval's. We conclude the proof by noting that

$$\begin{aligned}
\delta_f &= \frac{1-\mathbb{E}_{x,y}\left[f(x)f(y)f(x\odot y)\right]}{2} \\
&\geq \frac{1-\max_S \hat{f}(S)}{2} \\
&= \frac{1-(1-2\min_S \operatorname{dist}(f,\chi_S))}{2} \\
&= \min_S \operatorname{dist}(f,\chi_S).
\end{aligned}$$

Thus, $\mathbb{P}(f=\chi_s) \geq 1-\delta_f$ for some $S \subseteq [n]$.
■