Lecture 29: Discrete Fourier Analysis on the Boolean Hypercube (Introduction)

Focus of Study

- Functions with domain $\{0,1\}^n$ and range $\mathbb{R}$
- Let $f\colon \{0,1\}^n \to \mathbb{R}$
- We shall always use $N = 2^n$
- Any n-bit binary string shall be canonically interpreted as an integer in the range $\{0,1,\ldots,N-1\}$
- For any function $f\colon \{0,1\}^n \to \mathbb{R}$ we shall associate the following unique vector in $\mathbb{R}^N$

$$\big(f(0), f(1), \ldots, f(N-1)\big)$$

Kronecker Basis

- For $i \in \{0, 1, \ldots, N-1\}$, we define the function $\delta_i \colon \{0,1\}^n \to \mathbb{R}$ as follows

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

- Note that the functions $\{\delta_0, \delta_1, \ldots, \delta_{N-1}\}$ form a basis for $\mathbb{R}^N$
- Any function f can be expressed as a linear combination of these basis functions as follows

$$f = f(0)\delta_0 + f(1)\delta_1 + \cdots + f(N-1)\delta_{N-1}$$

- Our goal is to study the function f in a new basis, namely, the "Fourier Basis," that shall be introduced next. We emphasize that this basis need not be unique

- For $S = (S_1, S_2, \ldots, S_n) \in \{0,1\}^n$, we define the following function
$$\chi_S(x) := (-1)^{\sum_{i=1}^{n} S_i \cdot x_i}$$
- Several introductory materials on Fourier analysis interpret S as a subset of $\{1, 2, \ldots, n\}$. Although, the definition presented here is equivalent to this interpretation, I personally prefer this notation because it generalized to other domains.

An Example

- Suppose $n = 3$ and we are working with functions $f : \{0,1\}^n \to \mathbb{R}$
- Note that there are 8 different Fourier basis functions

$$\chi_{000}(x) = (-1)^0 = 1$$
$$\chi_{100}(x) = (-1)^{x_1}$$
$$\chi_{010}(x) = (-1)^{x_2}$$
$$\chi_{110}(x) = (-1)^{x_1+x_2}$$
$$\chi_{001}(x) = (-1)^{x_3}$$
$$\chi_{101}(x) = (-1)^{x_1+x_3}$$
$$\chi_{011}(x) = (-1)^{x_2+x_3}$$
$$\chi_{111}(x) = (-1)^{x_1+x_2+x_3}$$

Lemma

$$\sum_{x \in \{0,1\}^n} \chi_R(x) = \begin{cases} N, & \text{if } R = 0 \\ 0, & \text{otherwise} \end{cases}$$

Proof:

- Suppose $R = 0$, then we have

$$\sum_{x \in \{0,1\}^n} \chi_R(x) = \sum_{x \in \{0,1\}^n} 1 = N$$

- Suppose $R \neq 0$. Let $\{i_1, i_2, \ldots, i_r\}$ be the set of indices $\{i \colon R_i = 1\}$

$$\begin{aligned}
\sum_{x\in\{0,1\}^n} \chi_R(x) &= \sum_{x\in\{0,1\}^n} (-1)^{R_1x_1+\cdots+R_nx_n} \\
&= \sum_{x\in\{0,1\}^n} (-1)^{R_{i_1}x_{i_1}+\cdots+R_{i_r}x_{i_r}} \\
&= \sum_{x_{-i_1}\in\{0,1\}^{n-1}} (-1)^{R_{i_2}x_{i_2}+\cdots+R_{i_r}x_{i_r}} \sum_{x_{i_1}\in\{0,1\}} (-1)^{x_{i_1}} \\
&= \sum_{x_{-i_1}\in\{0,1\}^{n-1}} (-1)^{R_{i_2}x_{i_2}+\cdots+R_{i_r}x_{i_r}} \left((-1)^0 + (-1)^1\right) \\
&= \sum_{x_{-i_1}\in\{0,1\}^{n-1}} (-1)^{R_{i_2}x_{i_2}+\cdots+R_{i_r}x_{i_r}} \cdot 0 = 0
\end{aligned}$$

Inner Product

Definition (Inner Product)

The inner-product of two functions $f, g \colon \{0,1\}^n \to \mathbb{R}$ is defined as follows

$$\langle f, g \rangle := \frac{1}{N} \sum_{x \in \{0,1\}^n} f(x) g(x)$$

Lemma

$$\langle \chi_S, \chi_T \rangle = \begin{cases} 1, & \text{if } S = T \\ 0, & \text{otherwise} \end{cases}$$

Proof:

- $$\langle \chi_S, \chi_T \rangle = \frac{1}{N} \sum_{x \in \{0,1\}^n} \chi_S(x) \chi_T(x)$$
$$= \frac{1}{N} \sum_{x \in \{0,1\}^n} (-1)^{(S_1+T_1)x_1 + \ldots + (S_n+T_n)x_n}$$
- Note that if $S_i = T_i$ then $(-1)^{(S_i+T_i)x_i} = 1$; otherwise $(-1)^{(S_i+T_i)x_i} = (-1)^{x_i}$

Orthonormality of the Basis Functions II

- Define R such that $R_i = 1$ if $S_i \neq T_i$; otherwise $R_i = 0$
- Then, the right-hand side expression becomes

$$\begin{aligned}
\langle \chi_S, \chi_T \rangle &= \frac{1}{N} \sum_{x \in \{0,1\}^n} (-1)^{R_1 x_1 + \cdots + R_n x_n} \\
&= \frac{1}{N} \sum_{x \in \{0,1\}^n} \chi_R(x) \\
&= \begin{cases} \frac{1}{N} \cdot N, & \text{if } R = 0 \\ \frac{1}{N} \cdot 0, & \text{otherwise} \end{cases}
\end{aligned}$$

- Note that $R = 0$ if and only if $S = T$. This observation completes the proof

Summary

- Our objective is to study a function $f\colon \{0,1\}^n \to \mathbb{R}$
- Every function f is equivalently represented as the vector $\big(f(0), f(1), \ldots, f(N-1)\big) \in \mathbb{R}^N$, where $N = 2^n$
- For $S = S_1 S_2 \ldots S_n \in \{0,1\}^n$, define the following function

$$\chi_S(x) := (-1)^{S_1 x_1 + S_2 x_2 + \cdots + S_n x_n},$$

  where $x = x_1 x_2 \ldots x_n \in \{0,1\}^n$
- We defined an inner-product of functions

$$\langle f, g \rangle := \frac{1}{N} \sum_{x \in \{0,1\}^n} f(x) g(x)$$

- We showed that $\left\{\chi_S\colon S \in \{0,1\}^N\right\}$ is an orthonormal basis. That is,

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

Fourier Coefficients

- Since $\{\chi_S \colon S \in \{0,1\}^n\}$ is an orthonormal basis, we can express any f as follows

$$f = \widehat{f}(0)\chi_0 + \widehat{f}(1)\chi_1 + \cdots + \widehat{f}(N-1)\chi_{N-1},$$

where $\widehat{f}(S) \in \mathbb{R}$ and $S \in \{0,1\}^n$

- We interpret $\left(\widehat{f}(0), \widehat{f}(1), \ldots, \widehat{f}(N-1)\right)$ as a function $\widehat{f}$

- Fourier Transformation is a basis change that maps the function f to the function $\widehat{f}$
- We shall represent it as $f \stackrel{\mathcal{F}}{\mapsto} \widehat{f}$, where $\mathcal{F}$ is the Fourier Transformation

- Note that we have the following property. For any $S \in \{0,1\}^n$, we have $\langle f, \chi_S \rangle = \widehat{f}(S)$. So, we get

$$\left(f(0) f(1) \cdots f(N-1)\right) \cdot \frac{1}{N} \left(\chi_S(0) \chi_S(1) \cdots \chi_S(N-1)\right)^{\mathsf{T}} = \widehat{f}(S)$$

- Define the matrix

$$\mathcal{F} = \frac{1}{N} \begin{bmatrix} \chi_0(0) & \chi_1(0) & \cdots & \chi_{N-1}(0) \\ \chi_0(1) & \chi_1(1) & \cdots & \chi_{N-1}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_0(N-1) & \chi_1(N-1) & \cdots & \chi_{N-1}(N-1) \end{bmatrix}$$

- From the property mentioned above, note that we have the identity

$$f \cdot \mathcal{F} = \widehat{f}$$

Claim

For two function $f, g\colon \{0,1\}^n \to \mathbb{R}$*, we have*

$$\widehat{(f+g)} = \widehat{f} + \widehat{g}$$

Proof.

$$\widehat{(f+g)} = (f+g)\mathcal{F} = f\mathcal{F} + g\mathcal{F} = \widehat{f} + \widehat{g}$$

□

Claim

For a function $f\colon \{0,1\}^n \to \mathbb{R}$ *and* $c \in \mathbb{R}$*, we have*

$$\widehat{(cf)} = c\widehat{f}$$

Proof.

$$\widehat{(cf)} = (cf)\mathcal{F} = c(f\mathcal{F}) = c\widehat{f}$$

□

Theorem

Let $f\colon \{0,1\}^n \to \mathbb{R}$. Then, we have

$$\widehat{\left(\widehat{f}\right)} = \frac{1}{N} \cdot f$$

Proof.

- We shall prove that $\mathcal{F} \cdot \mathcal{F} = \frac{1}{N} I_{N\times N}$. This result shall directly imply that $\widehat{\left(\widehat{f}\right)} = (f\mathcal{F})\mathcal{F} = f\left(\frac{1}{N} I_{N\times N}\right) = \frac{1}{N} \cdot f$
- Let us compute the element $(\mathcal{F} \cdot \mathcal{F})_{i,j}$. This element is the product of the i-th row of $\mathcal{F}$ and the j-th colum of $\mathcal{F}$
- The j-th colum of $\mathcal{F}$ is $\left(\frac{1}{N}\chi_j\right)^{\mathsf{T}}$
- The i-th row of $\mathcal{F}$ is $\left(\chi_0(i)\chi_1(i)\cdots\chi_{N-1}(i)\right)$
- Note that $\chi_S(x) = \chi_x(S)$, i.e., the matrix $\mathcal{F}$ is symmetric

- So, the i-th row of $\mathcal{F}$ is $\frac{1}{N}\chi_i$
- Therefore, we have $(\mathcal{F}\mathcal{F})_{i,j} = \frac{1}{N^2} \cdot \chi_i \cdot \chi_j^{\intercal} = \frac{1}{N}\langle \chi_i, \chi_j \rangle$. The orthonormality of the Fourier basis completes the proof

Theorem (Plancherel)

Suppose $f, g \colon \{0,1\}^n \to \mathbb{R}$. Then, the following holds

$$\langle f, g \rangle = \sum_{S \in \{0,1\}^n} \widehat{f}(S)\widehat{g}(S)$$

Proof.

$$
\begin{aligned}
\langle f, g\rangle &= \left\langle \sum_{S\in\{0,1\}^n} \widehat{f}(S)\chi_S, \sum_{T\in\{0,1\}^n} \widehat{g}(T)\chi_T \right\rangle \\
&= \sum_{S\in\{0,1\}^n} \widehat{f}(S)\left\langle \chi_S, \sum_{T\in\{0,1\}^n} \widehat{g}(T) \right\rangle \\
&= \sum_{S\in\{0,1\}^n} \widehat{f}(S) \sum_{T\in\{0,1\}^n} \widehat{g}(T)\langle \chi_S, \chi_T\rangle \\
&= \sum_{S\in\{0,1\}^n} \widehat{f}(S)\widehat{g}(S)
\end{aligned}
$$

□

Note that, if $f, g\colon \{0,1\}^n \to \{+1,-1\}$ and we have $\langle f, g\rangle = 1-\varepsilon$, then f and g disagree at εN inputs. Intuitively, if $\left|\langle f, g\rangle\right|$ is close to 1 then the functions are highly correlated. On the other hand, if $\left|\langle f, g\rangle\right|$ is close to 0 then the functions are independent

Theorem (Parseval's Identity)

Suppose $f\colon \{0,1\}^n \to \mathbb{R}$. *Then*

$$\langle f, f\rangle = \sum_{S\in\{0,1\}^n} \widehat{f}(S)^2$$

Substitute $f = g$ in Plancherel's theorem.

Corollary

If $f\colon \{0,1\}^n \to \{+1,-1\}$, *then* $\sum_{S\in\{0,1\}^n} \widehat{f}(S)^2 = 1$

Follows from the fact that $\langle f, f \rangle = 1$ and the Parseval's identity