

PURDUE UNIVERSITY

CS 699

SPRING 2016

Research Thesis

Author:

Hai NGUYEN

Instructor:

Prof. Hemanta MAJI

February 15, 2016

Chapter 1

Fourier Basics

1.1 Vector Space of Functions on Boolean Hyper-cube

Definition 1.1 (Inner Product). Consider the 2^n -dimensional vector space of all functions $f : \{0, 1\}^n \rightarrow \mathbb{R}$. We define an inner product on this space by

$$\langle f, g \rangle := \mathbb{E}[f \cdot g] = \frac{1}{2^n} \sum_{x \in \{0, 1\}^n} f(x)g(x)$$

.

1.2 Characteristic Functions

Definition 1.2 (Characteristic function). For each $S \subseteq [n] = \{1, 2, \dots, n\}$, we define the characteristic function of S as

$$\chi_S(x) = (-1)^{S \cdot x}, \text{ where } S \cdot x = \sum_{i=1}^n S_i \cdot x_i = \sum_{i \in S} x_i$$

.

Lemma 1.3. For every $S \subseteq [n]$,

$$\sum_{x \in \{0, 1\}^n} \chi_S(x) = \begin{cases} 2^n & \text{if } S = \emptyset \\ 0 & \text{if } S \neq \emptyset \end{cases}$$

Proof. If $S = \emptyset$, then $S \cdot x = 0$. So $\sum_{x \in \{0, 1\}^n} \chi_S(x) = \sum_{x \in \{0, 1\}^n} 1 = 2^n$.

If $S \neq \emptyset$, then there exists k such that $S_k \neq 0$. Hence,

$$\begin{aligned}
\sum_{x \in \{0,1\}^n} \chi_S(x) &= \sum_{x \in \{0,1\}^n} (-1)^{\sum_{i \in S} x_i} \\
&= \sum_{x \in \{0,1\}^n} [(-1)^{x_k} \cdot (-1)^{\sum_{i \in S \setminus \{k\}} x_i}] \\
&= \sum_{x_k \in \{0,1\}} (-1)^{x_k} \cdot \sum_{x \setminus x_k \in \{0,1\}^{n-1}} (-1)^{\sum_{i \in S \setminus \{k\}} x_i} \\
&= [(-1)^0 + (-1)^1] \sum_{x \setminus x_k \in \{0,1\}^{n-1}} (-1)^{\sum_{i \in S \setminus \{k\}} x_i} \\
&= 0
\end{aligned}$$

□

Theorem 1.4. For every $S, T \subseteq [n]$,

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

Proof.

$$\langle \chi_S, \chi_T \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{S \cdot x + T \cdot x} = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{(S \Delta T) \cdot x}$$

where Δ is the symmetric different between two sets S and T .

$S \Delta T = \emptyset$ if and only if $S = T$. Hence, our goal follows immediately from Lemma 1.3. □

1.3 Fourier Basis

Theorem 1.5. The set of all χ_S defines an orthonormal basis for the space of all real-valued function on $\{0, 1\}^n$

Proof. From Theorem 1.4, the set of all χ_S is an orthonormal set. Also, there are 2^n different χ_S . Hence, the set of all χ_S must be an orthonormal basis for the space of all real-valued functions on $\{0, 1\}^n$. □

The set of all χ_S is called the *Fourier basis*.

1.4 Fourier Transform

Definition 1.6 (Fourier transform function). For each $S \subseteq [n]$, we define the Fourier transform of f as following:

$$\widehat{f}(S) := \mathbb{E}[f \cdot \chi_S] = \langle f, \chi_S \rangle$$

Theorem 1.7. *The mapping $\mathcal{F} : f \rightarrow \hat{f}$ is linear.*

Proof. This follows from the properties of inner product.

$$\langle af + bg, \chi_S \rangle = \langle af, \chi_S \rangle + \langle bg, \chi_S \rangle = a\langle f, \chi_S \rangle + b\langle g, \chi_S \rangle = a\hat{f} + b\hat{g} \quad \square$$

The linear map $\mathcal{F} : f \rightarrow \hat{f}$ is called the *Fourier transform*.

Theorem 1.8. *The linear map \mathcal{F} is a bijection.*

Proof. Since the set of χ_S forms an orthonormal basis,

$$f = \sum_S \hat{f}(S) \chi_S. \quad (1.1)$$

Suppose $\mathcal{F}(f_1) = \mathcal{F}(f_2)$, i.e. $\hat{f}_1(S) = \hat{f}_2(S)$ for every S , then it is followed from equation (1.1) that $f_1 = f_2$. So \mathcal{F} is injective.

Also, equation (1.1) implies that for every \hat{f} , there exists a function $f = \sum_S \hat{f}(S) \chi_S$ such

that $\mathcal{F}(f) = \hat{f}$, which means that \mathcal{F} is surjective.

Thus, \mathcal{F} is a bijection as derived \square

1.5 Dual of a Vector Space

Definition 1.9 (Dual space). Let V be a subspace of $\{0, 1\}^n$. We define the dual of V as $V^\perp = \{x \in \{0, 1\}^n \mid x \cdot v = 0 \ \forall v \in V\}$.

Theorem 1.10. V^\perp is a subspace of $\{0, 1\}^n$.

Proof. For any $x, y \in V^\perp, a \in \{0, 1\}, (a \cdot x + y) \cdot v = a \cdot (x \cdot v) + y \cdot v = 0 + 0 = 0$. \square

Lemma 1.11. $\sum_{i:\text{even}} \binom{n}{i} = \sum_{i:\text{odd}} \binom{n}{i} = 2^{n-1}$.

Theorem 1.12. *For any subspace V of dimension k of $\{0, 1\}^n$, there exists a unique dual space V^\perp of dimension $(n - k)$.*

Proof. We will show that $|V^\perp| = 2^{n-k}$ by induction on k .

If $k = 0$, then $V = \{\mathbf{0}\}$. Clearly, $V^\perp = \{0, 1\}^n$.

If $k = 1$, let $V = \{\vec{0}, v\}$. Suppose the number of $v_i = 1$ is t , then the number of x such that $x \cdot v = 0$ is $\sum_{i: 2|t-i} \binom{n}{i} 2^{n-t} = 2^{t-1} \cdot 2^{n-t} = 2^{n-1}$ by Lemma 1.11.

Suppose that there exists a unique orthogonal subspace V^\perp of dimension $(n - k + 1)$ for any subspace V of dimension $k - 1$ of $\{0, 1\}^n$, where $k \geq 2$.

Let $V = \langle v_1, v_2, \dots, v_k \rangle$, $S_1 = \langle v_1, v_2, \dots, v_{k-1} \rangle$, and $S_2 = \langle v_k \rangle$. Then, $V^\perp = S_1^\perp \cap S_2^\perp$.

Suppose $\dim(V^\perp) = t$. We want to show $t = n - k$.

By induction hypothesis, $\dim(S_1^\perp) = n - k + 1$ and $\dim(S_2^\perp) = n - 1$.

If $t \leq n - k - 1$, then we need $[(n - k + 1) - t]$ independent vectors to cover S_1^\perp from extending V^\perp , and we need $[(n - 1) - t]$ independent vectors to cover S_2^\perp from extending V^\perp . Since

$S_1^\perp \cup S_2^\perp \subseteq \{0, 1\}^n$, we must have $[(n - k + 1) - t] + [(n - 1) - t] + t \leq n$, which is equivalent to $t \geq n - k$, contradiction.

If $t \geq n - k + 1$, then $S_1^\perp \subseteq S_2^\perp$, this is impossible since v_k is independent from v_1, v_2, \dots, v_{k-1} . Thus, $t = n - k$. So $|V^\perp| = 2^{n-k}$. \square

1.6 Convolution

Definition 1.13. Given any two function f and $g : \{0, 1\}^n \rightarrow \mathbb{R}$, the convolution of $f * g : \{0, 1\}^n \rightarrow \mathbb{R}$ is defined as

$$(f * g)(x) := \frac{1}{2^n} \sum_{y \in \{0, 1\}^n} f(x \oplus y)g(y)$$

Theorem 1.14. If X and Y are n -bits random independent variables with probability distributions f and g , respectively, then $2^n(f * g)$ is the distribution of the random variable $Z = X \oplus Y$.

Proof.

$$\begin{aligned} Pr[Z = z] &= Pr[X = z \oplus Y] \\ &= \sum_{y \in \{0, 1\}^n} Pr[X = z \oplus y | Y = y] \\ &= \sum_{y \in \{0, 1\}^n} Pr[X = z \oplus y] \cdot Pr[Y = y] \\ &= \sum_{y \in \{0, 1\}^n} f((z \oplus y)) \cdot g(y) \\ &= 2^n(f * g)(z) \end{aligned}$$

\square

Theorem 1.15. For every $S \subseteq [n]$,

$$\widehat{f * g}(S) = \widehat{f}(S) \cdot \widehat{g}(S)$$

Proof.

$$\begin{aligned} \widehat{f * g}(S) &= \frac{1}{2^n} \sum_x (f * g)(x) \chi_S(x) \\ &= \frac{1}{2^n} \sum_x \left(\frac{1}{2^n} \sum_y f(x \oplus y)g(y) \right) \chi_S(x) \\ &= \frac{1}{2^{2n}} \sum_x \sum_y f(x \oplus y)g(y) \chi_S(x \oplus y) \chi_S(y) \\ &= \frac{1}{2^n} \sum_x f(x \oplus y) \chi_S(x \oplus y) \left(\frac{1}{2^n} \sum_y g(y) \chi_S(y) \right) \\ &= \widehat{f}(S) \cdot \widehat{g}(S) \end{aligned}$$

□

Intuitively, the convolution $f * g$ is the product of the Fourier transforms of f and g .

Theorem 1.16. *Let V be a subspace of dimension k of $\{0, 1\}^n$ and let V^\perp be the dual of V . Define*

$$f(x) = \begin{cases} \frac{1}{2^k} & \text{if } x \in V \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\widehat{f}(S) = \begin{cases} \frac{1}{N} & \text{if } S \in V^\perp \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Suppose v_1, v_2, \dots, v_k is a basis of V .

Claim 1.17. $V = \langle v_1 \rangle \oplus \langle v_2 \rangle \oplus \dots \oplus \langle v_k \rangle$

Claim 1.18. $\langle v_1 \rangle^\perp \cap \dots \cap \langle v_k \rangle^\perp = \langle v_1, \dots, v_k \rangle^\perp = V^\perp$

Let $f_i = \begin{cases} \frac{1}{2} & \text{if } S \in \langle v_i \rangle \\ 0 & \text{otherwise} \end{cases}$, then $\widehat{f}_i = \begin{cases} \frac{1}{N} & \text{if } S \in \langle v_i \rangle^\perp \\ 0 & \text{otherwise.} \end{cases}$

From above claims, we immediately obtain following result.

Claim 1.19. $f = f_1 \oplus f_2 \oplus \dots \oplus f_k$

Hence,

$$\widehat{f}(S) = N^{k-1} \widehat{f}_1(S) \dots \widehat{f}_k(S)$$

If $S \in V^\perp$, then $S \in \langle v_i \rangle^\perp$ for every i , so $\widehat{f}(S) = N^{k-1} \cdot (\frac{1}{N})^k = \frac{1}{N}$.

If $S \notin V^\perp$, then there exists some i such that $S \notin \langle v_i \rangle^\perp$, which implies $\widehat{f}_i(S) = 0$. Hence, $\widehat{f}(S) = 0$ □

1.7 Parseval's Identity

Because the χ_S form an orthonormal basis, we have the following equality:

$$\langle f, g \rangle = \sum_S \widehat{f}(S) \widehat{g}(S) \quad (1.2)$$

In particular, when $f = g$ we get Parseval's identity:

$$\|f\|_2^2 = \sum_S \widehat{f}(S)^2 \quad (1.3)$$

This also implies:

$$\|f - g\|_2^2 = \sum_S (\widehat{f}(S) - \widehat{g}(S))^2 \quad (1.4)$$

Chapter 2

Min Entropy

Let $X = (x_0, x_1, \dots, x_{N-1})$ be a distribution function of a random variable over $\{0, 1\}^n$, where $N := 2^n$.

Definition 2.1 (Min entropy). We define the min entropy of X as follow.

$$H_\infty(X) := \max_i (-\log x_i)$$

This implies that if $H_\infty(X) \geq k$ then $x_i \leq \frac{1}{2^k}$ for every $0 \leq i \leq N - 1$

Theorem 2.2 (Collision Probability). *If we sample X twice, then the probability we get the same result, denoted $Col(X)$, is $\sum_{i=0}^{N-1} x_i^2 = N \cdot \|X\|_2^2$.*

Definition 2.3. A probability distribution function $f: \{0, 1\}^n \rightarrow (0, 1)$ is a T -flat if there $\exists S \subseteq \{0, 1\}^n$ such that $|S| = T$ and $f(x) = \begin{cases} \frac{1}{T} & \text{if } x \in S \\ 0 & \text{otherwise} \end{cases}$

Theorem 2.4. *For every integer $k \geq 0$, if $H_\infty(X) \geq k$, then $X = \sum \alpha_i X_i$, where each X_i is a 2^k -flat, $\sum \alpha_i = 1$, and $\alpha_i \in [0, 1]$ for every i .*

Proof. □

Theorem 2.5. *If $H_\infty(X) \geq k$, then $Col(X) \leq \frac{1}{2^k}$.*

Proof. By theorem 2.5, we can write X as $X = \sum \alpha_i X_i$, where each X_i is a 2^k -flat $\sum \alpha_i = 1$, and $\alpha_i \in [0, 1]$ for every i . It is obvious that $Col(X_i) = \|X_i\|_2^2 = \frac{1}{2^k}$.

By Cauchy-Sachwarz inequality,

$$\|X\|_2^2 = \|\sum \alpha_i X_i\|_2^2 \leq |\sum \alpha_i| \cdot |\sum \alpha_i X_i^2| = \sum \alpha_i \|X_i\|_2^2 = \sum \alpha_i \frac{1}{2^k} = \frac{1}{2^k}$$
□

Theorem 2.6. *If $H_\infty(X) \geq k$, then $\sum_S \hat{X}(S)^2 \leq \frac{1}{2^k}$.*

Proof. This follow immediately from the Parseval's identity. □

Definition 2.7. Let \mathcal{D} be a probability distribution function over $\{0, 1\}^n$. We say that \mathcal{D} is α^* -bias if it fools all linear tests.

Fools all linear test means that if for any test s , a sample of \mathcal{D} , it returns 0 with probability $\frac{1}{2} + \alpha$ and returns 1 with probability $\frac{1}{2} - \alpha$, where $\alpha \leq \alpha^*$

Theorem 2.8. *If \mathcal{D} is a α^* -bias, then $\widehat{\mathcal{D}}(S) \leq \frac{2\alpha^*}{N}$ for all S .*

Proof. $\widehat{\mathcal{D}}(S) = \frac{1}{N} \sum_x \mathcal{D}(x)(-1)^{Sx} = \frac{1}{N}((\frac{1}{2} + \alpha) - (\frac{1}{2} - \alpha)) = \frac{2\alpha}{N} \leq \frac{2\alpha^*}{N}$ \square

Definition 2.9. *The statistical different between two distributions A and B is defined as follow.*

$$SD(A, B) = \frac{1}{2} \sum_i |a_i - b_i|$$

Theorem 2.10. *Let \mathcal{D} be a small bias distribution with $\widehat{\mathcal{D}}(S) \leq \frac{\alpha}{N}$ for all S , let M be a min entropy source such that $H_\infty(M) \geq k$, and let \mathcal{U} be the uniform distribution over n -bits string. Then*

$$SD(\mathcal{D} \oplus M, \mathcal{U}) \leq \frac{\alpha}{2}$$

Proof. \square