PURDUE UNIVERSITY

CS 699

SPRING 2016

Research Thesis

Author:
Hai NGUYEN

Instructor:
Prof. Hemanta MAJI

February 14, 2016

Chapter 1

Fourier Basics

1.1 The vector space of functions on boolean cube

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

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

1.2 The characteristic function

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

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

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

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

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,

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

□

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

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

Proof.

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

where ∇ is the symmetric different between two sets S and T.
If $S = T$, then $S\nabla T = \emptyset$, otherwise $S\nabla T \neq \emptyset$. Hence, our goals follow immediately from lemma 1.3. □

1.3 The Fourier basis

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

Proof. From theorem 1.4, it suffices to show that the set of all χ_S is a basis. Suppose $\sum\limits_{S\subset[n]} a_S \chi_S = \mathbf{0}$. We will show that X_S are linearly independent by proving $a_S = 0$ for every S. Let S be a subset of $[n]$, then

$$
\begin{aligned}
0 = \chi_S \sum_{T\subset[n]} a_T \chi_T &= \sum_{T\subset[n]} a_T \langle \chi_S, \chi_T \rangle \\
&= a_S \langle \chi_S, \chi_S \rangle + \sum_{T\subset[n], T\neq S} a_T \langle \chi_S, \chi_T \rangle = a_S
\end{aligned}
$$

Also, there are 2^n different X_S. Hence, the set of all χ_S is a basis for the space of all real-valued function on $\{0,1\}^n$ □

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

1.4 The Fourier transform

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

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

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

Proof. This follows from the properties of inner product.
$\langle af + bg, \chi_S \rangle = \langle af, \chi_S \rangle + \langle bg, \chi_S \rangle = a\langle f, \chi_S \rangle + b\langle g, \chi_s \rangle = a\hat{f} + b\hat{g}$ □

The linear map $\mathcal{F} : f \to \hat{f}$ 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,

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

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 $f = \sum\limits_S \hat{f}(S)\chi_S$ such that $\mathcal{F}(f) = \hat{f}$, which means that $\mathcal{F}$ is surjective.
Thus, $\mathcal{F}$ is a bijection as derived □

1.5 The dual space

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

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

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

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

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 $\sum\limits_{t-i \text{ even}} \binom{n}{i} 2^{n-t} = 2^{t-1} \cdot 2^{n-t} = 2^{n-1}$ by lemma 1.11.
Suppose that there exists a unique orthogonal subspace $V \perp$ of dimension $n-k+1$ for any subspace V of $\{0,1\}^n$, where $k \geq 2$.
Let $V = \langle v_1, v_2, ..., v_k \rangle$, $S_1 = \langle v_1, v_2, ..., v_{k-1} \rangle$, and $S_2 = \langle v_k \rangle$. By induction hypothesis, $|S_1^\perp| = 2^{n-k+1}$ and $|S_2^\perp| = 2^{n-1}$.

Claim 1.13. $V^\perp = S_1^\perp \cap S_2^\perp$ has a basis of size $n-k$.

Suppose $V^\perp$ has a basis of size t, then $|S_1^\perp \cap S_2^\perp| = 2^t$.
If $t \le n-k-1$, then there are at least $2^{n-k+1} - 2^t \ge 2^{n-k}$ vectors in $S_1^\perp \setminus V^\perp$, which implies that we need at least $n-k-t$ additional vectors, which are independent with t vectors above, for $S_1^\perp - V^\perp$. Similarly, and there are at least $2^{n-1} - 2^t \ge 2^{n-2}$ vectors in $S_1^\perp - V^\perp$, which implies that we need at least $n-2-t$ for $S_2^\perp - V^\perp$. Hence, there are at least $t + (n-k-t) + (n-2-t) = 2n-t-2 > n$ independent vectors for every $t \le n-3$, contradiction.
If $t \ge n-k+1$, then $S_1^\perp \subset S_2^\perp$, this is impossible since v_k is independent from $v_1, v_2, ..., v_{k-1}$. Thus, $t = n-k$. So $|V^\perp| = 2^{n-k}$. □

Theorem 1.14. *Let V be a subspace of dimension k of $\{0,1\}^n$ and let $V^\perp$ be the dual of V. Define*
$f(x) = \begin{cases} \frac{1}{2^k} & \textit{if } x \in V \\ 0 & \textit{otherwise.} \end{cases}$.
Then $\hat{f}(S) = \begin{cases} \frac{1}{N} & \textit{if } S \in V^\perp \\ 0 & \textit{otherwise.} \end{cases}$.

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

Claim 1.15. $V = \langle v_1 \rangle \oplus \langle v_2 \rangle \oplus ... \oplus \langle v_k \rangle$

Claim 1.16. $\langle v_1 \rangle^\perp \cap ... \cap \langle v_k \rangle^\perp = \langle v_1, ..., v_k \rangle^\perp = V^\perp$

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

Claim 1.17. $f = f_1 \oplus f_2 \oplus ... \oplus f_k$

Proof. This is immediate from claims 1.15 and 1.16. □

Hence,

$$\begin{aligned} \hat{f}(S) &= f_1 \widehat{\oplus ... \oplus} f_k(S) \\ &= N \hat{f}_1(S) \cdot f_2 \widehat{\oplus ... \oplus} f_k(S) \\ &\cdots \\ &= N^{k-1} \hat{f}_1(S) ... \hat{f}_k(S) \end{aligned}$$

If $S \in V^\perp$, then $S \in \langle v_i \rangle^\perp$ for every i, so $\hat{f}(S) = N^{k-1} \cdot (\frac{1}{N})^k = \frac{1}{N}$.
If $S \notin V^\perp$, then there exits some i such that $S \notin \langle v_i \rangle^\perp$, which implies $\hat{f}_i(S) = 0$. Hence, $\hat{f}(S) = 0$ □

1.6 Parseval's Identity

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

$$\langle f, g \rangle = \sum_S \hat{f}(S)\hat{g}(S) \tag{1.2}$$

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

$$\|f\|_2^2 = \sum_S \hat{f}(S)^2 \tag{1.3}$$

This also implies:

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

Chapter 2

Min Entropy

Let $X = (x_0, x_1, ..., x_{N-1})$ 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.

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

Theorem 2.2. *For every k, $H_\infty(X) \geq k$ if and only if $x_i \leq \frac{1}{2^k}$ for every $0 \leq i \leq N-1$*

Theorem 2.3 (Collision Probability)**.** *If we sample X twice, then the probability we get the same result, denoted $Col(X)$, is $\sum\limits_{i=0}^{N-1} {x_i}^2 = \|X\|_2^2$.*

Definition 2.4. A probability distribution function $f\colon \{0,1\}^n \to (0,1)$ is a *T-flat* if there $\exists\ S \subset \{0,1\}^n$ such that $|S| = T$ and $f(x) = \begin{cases} \frac{1}{T} & \text{if } x \in S \\ 0 & \text{otherwise} \end{cases}$

Theorem 2.5. *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 2.6. *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 2.7. *If $H_\infty(X) \geq k$, then $\sum\limits_S \widehat{X}(S)^2 \leq \frac{1}{2^k}$.*

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

Definition 2.8. 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 2.9. *If $\mathcal{D}$ is a α^*-bias, then $\hat{\mathcal{D}}(S) \leq \frac{2\alpha^*}{N}$ for all S.*

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

Definition 2.10. We define *the statistical different* between A and B as follow.

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

Theorem 2.11. *Let $\mathcal{D}$ be a small bias distribution with $\hat{\mathcal{D}} \leq \frac{\alpha}{N}$ for all S, let M be a min entropy source, and let $\mathcal{U}$ be the uniform distribution. Then $SD(\mathcal{D} \oplus M, \mathcal{U}) \leq \frac{\alpha}{2}$*

Proof. □