TTIC/CMSC 31150 Mathematical Toolkit **Spring 2013**

Lecture 12: May 13, 2013

Madhur Tulsiani Scribe: David Hong Kyun Kim

We will be starting on a new topic, the Fourier analysis of boolean functions. A boolean function $f : \{0,1\}^n \to \mathbb{R}$ may arise in various contexts: f may be an algorithm which computes a real valued output, or a function which assigns concept classes or labels to some data. For example, one may have points in the plane which need to be partitioned by a line or hyperplane. Another example where boolean functions are studied is in social choice theory, when f is a voting function and an input bit represents a voter. One may study the influence of a single bit or a subset of the input bits on the output of f. Boolean functions are also used in combinatorics, where we have $\mathcal{F}$, a family of subsets of $[n] = \{1, 2, ..., n\}$, and an input $x \in \{0,1\}^n$ which represents a subset S, with $x_i = 1$ iff $i \in S$. Then $f(x) = 1$ iff $S \in \mathcal{F}$.

Example: Let $f : \{0,1\}^3 \to \{0,1\}$ and $\mathcal{F} = \{\emptyset, \{2\}, \{1\}, \{1,2,3\}\}$. Then $f(000) = f(010) = f(100) = f(111) = 1$, and $f(x) = 0$ otherwise.

1 The Fourier Expansion of Boolean Functions

1.1 The Fourier Basis

We first we define an inner product on the space of real valued boolean functions.

Definition 1.1 (Inner product of two boolean functions) *Given two boolean functions $f, g : \{0,1\}^n \to \mathbb{R}$, the inner product of f and g is defined as follows:*

$$< f, g >= \mathop{\mathbb{E}}_{x \in \{0,1\}^n}[f(x)g(x)] = \frac{1}{2^n}\sum_x f(x)g(x)$$

You can also think of the inner product as taking the function tables for f and g, then normalizing the inner product of the two columns (containing the function values for the 2^n inputs) with $\frac{1}{2^n}$. With this definition, the square of the norm of a boolean function is:

$$< f, f >= ||f||_2^2 = \mathop{\mathbb{E}}_{x \in \{0,1\}^n}[(f(x))^2]$$

An example of an orthonormal basis is as follows:

- Let f_i be the function which has 1 non-zero entry on the i'th element, and 0 everywhere else. Then $\{f_1, ..., f_{2^n}\}$ is a basis for the space of boolean functions on $\{0,1\}^n$, which is a vector space of dimension 2^n.

Now let's define another orthonormal basis, which we will use.

Definition 1.2 (The Fourier Basis) *Let* $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$, *where* $x \in \{0,1\}^n$ *and* $S \in [n]$.

(example) $\chi_{\{1,2\}}(x) = (-1)^{x_1 + x_2}$

Claim 1.3 *The Fourier basis is an orthnormal basis.*

Note that there are 2^n subsets of $[n]$, and hence 2^n such functions. Corresponding to the empty set is $\chi_\emptyset(x) = (-1)^0 = 1$. Here we check that these 2^n functions are orthonormal.

- Unit vectors. The norm of each of these functions is 1, as shown below.

$$||\chi_S||^2 = \mathop{\mathbb{E}}_{x \in \{0,1\}^n} \left[((-1)^{\sum_{i \in S} x_i})^2 \right] = 1$$

- Orthogonality. Let $S \neq T$ be two subsets of $[n]$.

$$\mathop{\mathbb{E}}_{x \in \{0,1\}^n} [\chi_S(x) \chi_T(x)] = \mathop{\mathbb{E}}_{x \in \{0,1\}^n} \left[(-1)^{\sum_{i \in S} x_i} (-1)^{\sum_{j \in T} x_j} \right] = \mathop{\mathbb{E}}_{x \in \{0,1\}^n} \left[(-1)^{\sum_{i \in S} x_i + \sum_{j \in T} x_j} \right]$$

  Let's look at each element of $S \subset [n]$.

  - $i \in S \cap T \Rightarrow x_i$ appears twice $\Rightarrow (-1)^{x_i + x_i} = 1$
  - $i \notin S, i \notin T \Rightarrow i \in \overline{(S \cup T)} \Rightarrow x_i$ does not appear.

  Therefore, only the elements in the symmetric difference of S and T, $S \Delta T = (S \backslash T) \cup (T \backslash S)$, remain.

$$\mathop{\mathbb{E}}_{x \in \{0,1\}^n} \left[(-1)^{\sum_{i \in S \Delta T} x_i} \right] = \mathop{\mathbb{E}}_{x \in \{0,1\}^n} \left[\prod_{x_i \in S \Delta T} (-1)^{x_i} \right]$$

$$(\textit{since the } x_i\textit{'s are independent})$$

$$= \prod_{i \in S \Delta T} \left(\mathop{\mathbb{E}}_{x_i \in \{0,1\}} [(-1)^{x_i}] \right) = 0$$

Any boolean function $f : \{0,1\}^n \to \mathbb{R}$ can be written as a linear combination of the Fourier basis vectors, where $f(x) = \sum_{S \subset [n]} c_S \chi_S(x)$. Note that c_S is just a coefficient for the vector $\chi_S(x)$. c_S is usually denoted as $\hat{f}(S)$, called the Fourier coefficient of f corresponding to the set S.

1.2 Parseval's Identity

Consider two boolean functions mapping $\{0,1\}^n \to \mathbb{R}$, where $f = \sum_S \hat{f}(S)\chi_S$ and $g = \sum_S \hat{g}(S)\chi_S$. There are two ways to look at the inner product, $< f, g >$.

1. $\mathop{\mathbb{E}}_x[f(x)g(x)]$
2. $< \sum_S \hat{f}(S)\chi_S, \sum_T \hat{g}(T)\chi_T >$

$$= \sum_{S,T} \hat{f}(S)\hat{g}(S) < \chi_S, \chi_T >= \sum_S \hat{f}(S)\hat{g}(S) < \chi_S, \chi_S >= \sum_S \hat{f}(S)\hat{g}(S)$$

(*By the orthonormality of the Fourier basis vectors*)

Theorem 1.4 (Plancherel's Theorem) *The two expressions for the inner product above give us the following theorem.*

$$< f, g >= \mathop{\mathbb{E}}_x[f(x)g(x)] = \sum_S \hat{f}(S)\hat{g}(S)$$

Theorem 1.5 (Parseval's Identity) *Let $f = g$ to get the square of the norm,*

$$||f||^2 = \mathop{\mathbb{E}}_x[f(x)^2] = \sum_S (\hat{f}(S))^2$$

2 Applications

2.1 Linearity Testing

Definition 2.1 *A boolean function $f : \{0,1\}^n \to \{0,1\}$ is linear if $\forall x, y \in \{0,1\}^n, f(x+y) = f(x) + f(y)$ (mod 2).*

Given $f : \{0,1\}^n \to \{0,1\}$, we want to check if f is linear. f is a huge object, as we need 2^n bits to describe it, and verifying linearity by testing on each pair of the inputs would be very inefficient. Instead, we do *property testing* - checking a property (in our case linearity) of f with a small number of tests such that

- if the property holds, the tests always return "yes"
- if the property does not hold, the tests return "no" with some probability

We want to test our function f such that if the test passes, f will be a linear function or very close to a linear function. The test, decribed below, will be the most obvious thing one can do.

TEST: Pick random $x, y \in \{0,1\}^n$

- return "yes" if $f(x) + f(y) = f(x+y) \pmod 2$
- return "no" otherwise

If f is linear, the test will always return "yes". We want to show that if the test returns "yes", then f is very close to linear.

ANALYSIS:

Suppose $\mathop{\mathbb{P}}_{x,y}[\text{Test says "yes"}] \geq 1 - \varepsilon$. Let $g(x) = (-1)^{f(x)}$.

$$f(x) + f(y) = f(x+y) \bmod 2 \iff g(x) \cdot g(y) = g(x+y) \iff \frac{1 + g(x)g(y)g(x+y)}{2} = 1$$

Note that the last expression on the right becomes 0 when the test fails and is a random variable depending on our choice of x and y. Since it takes values 0 and 1, the probability that it has the value 1 is equal to its expectation.

$$\mathop{\mathbb{P}}_{x,y}[\text{Test says "yes"}] = \mathop{\mathbb{P}}_{x,y}[f(x) + f(y) = f(x+y) \bmod 2] = \mathop{\mathbb{P}}_{x,y}\left[\frac{1 + g(x)g(y)g(x+y)}{2} = 1\right]$$

$$= \mathop{\mathbb{E}}_{x,y}\left[\frac{1 + g(x)g(y)g(x+y)}{2}\right] \geq 1 - \varepsilon$$

$$\Rightarrow \mathop{\mathbb{E}}_{x,y}[g(x)g(y)g(x+y)] \geq 1 - 2\varepsilon.$$

g can be written as $\sum_S \hat{g}(S)\chi_S(x)$ using the Fourier basis, and

$$\mathop{\mathbb{E}}_{x,y}[g(x)g(y)g(x+y)] = \mathop{\mathbb{E}}_{x,y}\left[\sum_{S,T,W} \hat{g}(S)\hat{g}(T)\hat{g}(W)\chi_S(x)\chi_T(x)\chi_W(x+y)\right]$$

claim: $\forall W, \chi_W(x+y) = \chi_W(x) \cdot \chi_W(y)$
$\rightarrow \chi_W(x+y) = (-1)^{\sum_{i \in W} x_i + y_i} = (-1)^{\sum_{i \in W} x_i} \cdot (-1)^{\sum_{i \in W} y_i} = \chi_W(x) \cdot \chi_W(y)$

So the above expression evaluates to

$$\mathop{\mathbb{E}}_{x,y}\left[\sum_{S,T,W} \hat{g}(S)\hat{g}(T)\hat{g}(W)\chi_S(x)\chi_T(x)\chi_W(x+y)\right]$$

$$= \mathop{\mathbb{E}}_{x,y}\left[\sum_{S,T,W} \hat{g}(S)\hat{g}(T)\hat{g}(W)\chi_S(x)\chi_T(y)\chi_W(x)\chi_W(y)\right]$$

$$= \sum_{S,T,W} \hat{g}(S)\hat{g}(T)\hat{g}(W)\left(\mathbb{E}[\chi_S(x)\chi_W(x)]\right)\left(\mathbb{E}[\chi_T(y)\chi_W(y)]\right)$$

The inner product of χ_S, χ_W and χ_T, χ_W are both non-zero only when $S = T = W$, and is equal to 1. So the whole expression becomes

$$\mathop{\mathbb{E}}_{x,y}[g(x)g(y)g(x+y)] = \sum_S (\hat{g}(S))^3 \text{ and}$$

$$\mathbb{P}[\text{Test returns "yes"}] = \tfrac{1}{2} + \tfrac{1}{2}\mathop{\mathbb{E}}_{x,y}[g(x)g(y)g(x+y)] = \tfrac{1}{2} + \tfrac{1}{2}\sum_S (\hat{g}(S))^3 \geq 1 - \varepsilon$$

This gives us

$$\sum_S (\hat{g}(S))^3 \geq 1 - 2\varepsilon$$

Let's put a bound on this quantity:

$$\Rightarrow \max_T(\hat{g}(T))(\sum_S (\hat{g}(S))^2) \geq \sum_S (\hat{g}(S))^3 \geq 1 - 2\varepsilon$$

Since $\sum_S (\hat{g}(S))^2 = \mathop{\mathbb{E}}_x[(g(x))^2]$ by Parseval's identity, and since $\mathop{\mathbb{E}}_x[(g(x)^2)] = 1$, we have

$$\Rightarrow \max_T(\hat{g}(T)) \geq 1 - 2\varepsilon$$

We want to use this bound to say that f is close to being a linear function. We said $\exists T$ such that $\hat{g}(T) \geq 1 - 2\varepsilon$. Another way of computing $\hat{g}(T)$ is

$$\hat{g}(T) = < g, \chi_T > = \mathop{\mathbb{E}}_x[g(x)\chi_T(x)] = \mathop{\mathbb{E}}_x\left[(-1)^{f(x)} \cdot (-1)^{\sum_{i\in T} x_i}\right]$$

$(-1)^{f(x)} \cdot (-1)^{\sum_{i\in S} x_i}$ is equal to either 1 or -1. Let p be the probability that it is 1. Then

$$\mathop{\mathbb{E}}_x\left[(-1)^{f(x)} \cdot (-1)^{\sum_{i\in T} x_i}\right] = p \cdot 1 + (1-p) \cdot (-1) \geq 1 - 2\varepsilon.$$

Since the random variable in the expression evaluates to 1 exactly when $f(x) = \sum_{i\in T} x_i \pmod 2$, we have

$$p = \mathop{\mathbb{P}}_x\left[f(x) = \sum_{i\in T} x_i (mod 2)\right] \geq 1 - \varepsilon.$$

But $\sum_{i\in T} \chi_i \pmod 2$ is a linear function, and we have that our function is linear with high probability when the test returns "yes". Note that these are the only linear functions, as if f is linear, we would have $\exists T$ such that $f(x) = \sum_{i\in T} x_i \pmod 2$ with $\varepsilon = 0$ and $\mathbb{P}[\text{test returns "yes"}] = 1$.