| | |
|---|---|
| **6.842 Randomness and Computation** | March 12, 2012 |

# Lecture 10

*Lecturer: Ronitt Rubinfeld* — *Scribe: Hamidreza Jahanjou*

Today, we are going to discuss the following

- Linearity Testing,
- Fourier Analysis.

## 1 Linearity Testing

**Definition 1** *A function $f : \{0,1\}^n \to \{0,1\}$ is called linear if*

$$\forall x, y \in \{0,1\}^n : f(x) + f(y) = f(x+y) \pmod 2 \tag{1}$$

*where the plus sign represents mod 2 addition in vector space; specifically,*

$$(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1 \pmod 2,\ x_2 + y_2 \pmod 2, ...,\ x_n + y_n \pmod 2).$$

The property defined by equation (1) is also known as the homomorphism property.
Some examples of linear functions are

- $f(x) = 0$,
- $f(x) = x_i$, (projection functions),
- $f(x) = \bigoplus_{i=1}^{n} x_i$.

A useful relation is the following for which we give an informal proof.

**Claim 1** *A function $f : \{0,1\}^n \to \{0,1\}$ is linear iff $\forall x :\ f(x) = \bigoplus_{i \in S} x_i$ for some $S \subseteq [n]$.*

**Sketch of Proof** A linear function $f$ is uniquely determined by all $f(u_i)$ where $u_i = (0, ..., 0, 1, 0, ..., 0)$ is the $i$th unit vector. Clearly, $2^n$ possible settings of $f(u_i)$'s means $2^n$ possible linear functions. On the other hand, there are only $2^n$ sets $S \subseteq [n]$. Consequently, we're accounting for all linear functions. ■

How can we tell if a function $f$ is linear? Querying can be very inefficient since points of non-linearity may be very sparse in the whole space. This motivates the following definition.

**Definition 2** *A function $f$ is $\epsilon$-close to linear if there exists a linear function $g$ such that*

$$\Pr[f(x) = g(x)] = \frac{|\{x | f(x) = g(x)\}|}{2^n} \geq 1 - \epsilon.$$

*Otherwise $f$ is $\epsilon$-far to linear.*

Base on this idea, we propose a linearity test.

**function** TEST
  **repeat** $r \leftarrow 1$
    pick $x, y \in_R \{0,1\}$
    **if** $[f(x) + f(y) \neq f(x+y) \pmod 2]$ **then**
      Fail and Halt.
    **end if**
  **until** $r = O(\frac{1}{\epsilon})$

Output Pass.

**end function**

Regarding the behavior of our test, we observe the following.

- if $f$ is linear, then $\Pr[\text{Pass}] = 1$,
- if $f$ is $\epsilon$-far from linear, then $\Pr[\text{Fail}] \geq \frac{3}{4}$.

### A Notational Switch

From now on, the following changes are in effect:

| Previously | Now |
|---|---|
| $f_s : \{0,1\}^n \to \{0,1\}$ <br> $f_s(x) = \bigoplus_{i \in S} x_i = \sum_{i\in S} x_i \pmod 2$ | $f_s : \{\pm 1\}^n \to \{\pm 1\}$ <br> $f_s(x) = \prod_{i \in S} x_i$ |

We note that, with this notational change, $f(x)f(y) \neq f(x,y) \iff f(x)f(y)f(xy) = -1$. Therefore, it's possible to define the following indicator random variables

$$\frac{1 - f(x)f(y)f(xy)}{2} = \begin{cases} 0 & \text{if the test passes,} \\ 1 & \text{if the test fails.} \end{cases}$$

We also define their expected value:

$$\delta = \mathbb{E}_{x,y}\left[\frac{1 - f(x)f(y)f(xy)}{2}\right].$$

The expected value of the indicator random variable is the probability of rejection in one pass. The probability of being accepted is similarly define as

$$1 - \delta = \mathbb{E}_{x,y}\left[\frac{1 + f(x)f(y)f(xy)}{2}\right].$$

But, how can we calculate the value of these expressions? This is where Fourier Analysis comes in.

## 2 Fourier Analysis

Let's begin by considering the set $\mathbf{G} = \{g \mid g : \{\pm 1\}^n \to \mathrm{R}\}$. Note that $\mathbf{G}$ is a vector space of dimension $2^n$. In other words, all functions can be written as linear combinations of $2^n$ basis functions. For $f, g \in \mathbf{G}$, their inner product is defined as

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

Next comes the choice of a basis for our vector space. Here we consider two possibilities. First, let's consider indicator functions

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

In this case an arbitrary function $f \in \mathbf{G}$ may be written as $f(x) = \sum_a f(x)e_a(x)$. Even though indicator functions constitute an orthonormal basis, they're not very useful. Instead, we'll use character functions

$$\chi_{\mathrm{S}}(x) = \prod_{i \in S} x_i.$$

Some examples of functions that can be written in terms of character functions:

- $f(x) = 1$ can be written as 1 which is $\chi_{\emptyset}$.
- $f(x) = x_i$ can be written as $x_i$ which is $\chi_{\{i\}}$.
- $\mathrm{And}(x_1, x_2)$ can be written as $\frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1x_2$.
- $\mathrm{Maj}(x_1, x_2, x_3)$ can be written as $\frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$.

We haven't shown that character functions form an orthonormal basis yet. This is done in two steps. First we prove their orthonormality. Next, we show that they are indeed a basis.

**Lemma 2** *The set* $\{\chi_S \mid S \subseteq [n]\}$ *forms an orthonormal basis.*

**Proof** Normality:

$$\langle \chi_{\mathrm{S}}, \chi_{\mathrm{S}} \rangle = \frac{1}{2^n} \sum_x \underbrace{(\chi_{\mathrm{S}}(x))^2}_{1\text{'s}} = 1$$

where the second equality follows from the observation that the LHS expression is an average of 1's. Orthogonality: suppose $S \neq T$,

$$\begin{aligned}
\langle \chi_{\mathrm{S}}, \chi_{\mathrm{T}} \rangle &= \frac{1}{2^n} \sum_{x \in \{\pm 1\}} \chi_{\mathrm{S}}(x)\chi_{\mathrm{T}}(x) \\
&= \frac{1}{2^n} \sum_x \Big( \prod_{i \in S} x_i \prod_{j \in T} x_j \Big) \\
&= \frac{1}{2^n} \sum_x \Big( \prod_{x \in S \setminus T} x_i \prod_{j \in T \setminus S} x_j \overbrace{\prod_{k \in S \cap T} x_k^2}^{1} \Big) \\
&= \frac{1}{2^n} \sum_x \Big( \prod_{i \in S \Delta T} x_i \Big) \\
&= \frac{1}{2^n} \sum_{x, x^{\oplus j}} \Big( \prod_{i \in S \Delta T} x_i + \prod_{i \in S \Delta T} x_i^{\oplus j} \Big) \\
&= \frac{1}{2^n} \sum_{x, x^{\oplus j}} \prod_{i \in S \Delta T \setminus \{j\}} [x_j + \overline{x_j}] \\
&= 0
\end{aligned}$$

where $x^{\oplus j}$ is $x$ with the $j$-th entry flipped and the last equality holds because $x_j + \overline{x_j} = 0$. ∎

Now, we need to show that any $f \in \mathbf{G}$, has a unique representation as a linear combination of $\chi_{\mathrm{S}}$'s. The proof of the following theorem is left as an exercise.

**Theorem 3** *Suppose $f \in \mathbf{G}$, then*

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

*where*

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

Some nice properties:

1. $\chi_{\mathrm{S}}\chi_{\mathrm{T}} = \chi_{S\Delta T}$.
2. (parity functions:) if $f = \chi_{\mathrm{S}}$, then

$$\hat{f}(Z) = \begin{cases} 1 & \text{if } Z = S, \\ 0 & \text{if } Z \neq S. \end{cases}$$

3.

$$\hat{f}(S) = 1 - 2\Pr[\underbrace{f(x) \neq \chi_{\mathrm{S}}(x)}_{\mathrm{dist}(f,\chi_{\mathrm{S}})}].$$

**Proof**

$$\begin{aligned}
\hat{f}(S) &= \frac{1}{2^n} \sum_x f(x)\chi_{\mathrm{S}}(x) \\
&= \frac{1}{2^n}\Big[ \sum_{x: f(x)=\chi_{\mathrm{S}}(x)} \underbrace{f(x)\chi_{\mathrm{S}}(x)}_{1} + \sum_{x: f(x)\neq\chi_{\mathrm{S}}(x)} \underbrace{f(x)\chi_{\mathrm{S}}(x)}_{-1} \Big] \\
&= \frac{1}{2^n}\big[(2^n - |\{x \mid f(x) \neq \chi_{\mathrm{S}}(x)\}|) \times 1 + (|\{x \mid f(x) \neq \chi_{\mathrm{S}}(x)\}|) \times (-1)\big] \\
&= 1 - 2\Pr[f(x) \neq \chi_{\mathrm{S}}(x)].
\end{aligned}$$

■

4. if $S \neq T$, then $\mathrm{dist}(\chi_{\mathrm{S}}, \chi_{\mathrm{T}}) = \frac{1}{2}$.

   **Sketch of Proof** For $S \neq T$, consider the Fourier representation of $\chi_{\mathrm{S}}$, and in particular the $T$th Fourier coefficient. It is equal to 0 by the orthonormality of $\chi_{\mathrm{S}}$ and $\chi_{\mathrm{T}}$. Furthermore, by property 3, we have $0 = 1 - 2\Pr[\chi_{\mathrm{S}} \neq \chi_{\mathrm{T}}]$. Solving it gives us property 4. ■

*Comment: Hadamard codes* encode $a \in \{\pm 1\}^n$ by bit strings of length $2^n$ by writing the value of $\chi_a$ for all $n$-bit inputs. Now, for all $a \neq b$ we observe that $\mathrm{Had}(a)$ and $\mathrm{Had}(b)$ differ on $\frac{1}{2}$ of the bits.