6.895 Randomness and Computation	February 13, 2008
Lecture 3	
Lecturer: Ronitt Rubinfeld	*Scribe: Megumi Ando*

Last Time

In the previous lecture, we made a notational switch from using Boolean functions of the form $f : \{0,1\}^n \to \{0,1\}$ to functions of the form $f : \{-1,1\}^n \to \{-1,1\}$. We defined linearity over this form, and what it meant to be ϵ-close to linear.

Definition 1 *$f : \{-1,1\}^n \to \{-1,1\}$ is* linear *if $\forall x, y \in \{-1,1\}^n$, $f(x)f(y)f(x \cdot y) = 1$, where $x \cdot y = (x_1 ... x_n) \cdot (y_1 ... y_n) = (x_1 y_1, x_2 y_2, ..., x_n y_n)$.*

There are 2^n linear functions over $\{-1,1\}^n \to \{-1,1\}$. Each of them can be written as $\chi_S(x) = \prod_{i \in S} x_i$, where $S \subseteq \{1, ..., n\}$.

Definition 2 *A function f is* ϵ-close to linear *if $\exists$ linear g such that $Pr_x[f(x) \neq g(x)] \leq \epsilon$. Otherwise, f is* ϵ-far*.*

Finally, we proposed the following linearity tester:

- Repeat $O\left(\frac{1}{\rho} \log \frac{1}{\beta}\right)$ times:
  - Pick $x, y \in \{-1,1\}^n$.
  - If $f(x)f(y)f(x \cdot y) \neq 1$, output "FAIL" and halt.
- Output "PASS."

The rejection probability of one pass through the loop is $\delta \equiv E_{x,y}\left[\frac{1 - f(x)f(y)f(x \cdot y)}{2}\right]$.

1 Fourier Analysis (Basics)

We will use a few times the following simple fact about linear functions.

Fact 3

$$\chi_S(x) \cdot \chi_T(x) = \prod_{i \in S} x_i \cdot \prod_{j \in R} x_j = \prod_{i \in S \triangle T} x_i,$$

where $S \triangle T$ is the symmetric difference of S and T, i.e., the set of elements that appear in exactly one of the sets S and T.

1.1 Vector Space of Functions $g : \{-1,1\}^n \to \mathbb{R}$

The set $G = \{g | g : \{-1,1\}^n \to \mathbb{R}\}$ is a vector space of dim 2^n.

Definition 4 Indicator functions *are functions of the form: If $x = a$, then $e_a(x) = 1$. Otherwise, $e_a(x) = 0$.*

Note that the indicator functions are basis functions of G. However, we will not be using them. Instead we will be using the parity functions, $\{\chi_S\}_{S \subseteq [n]}$, described in the previous lecture.

Definition 5 *For $f, g : \{-1,1\} \to \{-1,1\}$, the "inner product"*

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

where the sum $\sum f(x)g(x)$ is the "correlation," a measure of how often f and g agree.

Note that:

1. $< \chi_S, \chi_S > = \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} \chi_S^2(x) = 1$. (Absolute correlation.)
2. If $S \neq T$,

$$\begin{aligned} < \chi_S, \chi_T > &= \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} \chi_S(x)\chi_T(x) \\ &= \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} \prod_{i \in S} x_i \prod_{j \in T} x_j \qquad \text{(by definition)} \\ &= \frac{1}{2^n} \sum \prod_{i \in S \triangle T} x_i \qquad \text{(by Fact 3)} \end{aligned}$$

where $S \triangle T$ is non-empty. Therefore, there exists a $j \in S \triangle T$. Let $x^{\oplus j}$ equals x with the j-th bit flipped.

$$\begin{aligned} &= \frac{1}{2^n} \sum_{\text{pairs } (x, x^{\oplus j})} \left(x_j \prod_{i \in S \triangle T \setminus \{j\}} x_i + \overline{x_j} \prod_{i \in S \triangle T \setminus \{j\}} x_i \right) \\ &= 0 \end{aligned}$$

From Notes 1 and 2 above, we see that every parity function X_S is normal to the others, and thus, the parity functions form an orthonormal basis.

1.2 Fourier Coefficients

The following corollary follows.

Corollary 6

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

where $\hat{f}(z)$ is the Fourier coefficient, which can be calculated as follows:

$$\begin{aligned} \hat{f}(S) &= \langle f, \chi_S \rangle \\ &= \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)\chi_S(x) \end{aligned}$$

In particular, a parity function has all but one coefficients equal zero.

Fact 7 (Fourier Coefficients of Parity Functions χ_T)

$$f = \chi_T \iff \hat{f}(T) = 1.$$

Furthermore, $\forall S \neq T, \hat{f}(S) = 0.$

A few more examples of Fourier coefficients:

Function	Fourier Representation
$f(x) = 1$	$1 \cdot \chi_\emptyset$
$f(x) = x_i$	$1 \cdot \chi_{\{i\}}$
$\text{and}(x_1, x_2)$	$\frac{1}{2}\chi_\emptyset + \frac{1}{2}\chi_{\{1\}} + \frac{1}{2}\chi_{\{2\}} - \frac{1}{2}\chi_{\{1,2\}}$
$\text{maj}(x_1, x_2, x_3)$	$\frac{1}{2}\chi_{\{1\}} + \frac{1}{2}\chi_{\{2\}} + \frac{1}{2}\chi_{\{3\}} - \frac{1}{2}\chi_{\{1,2,3\}}$

1.3 Fourier Coefficients and Distance to Linearity

Let $\text{dist}(f, g)$ denote the fraction of inputs on which two Boolean functions $f, g : \{-1, 1\}^n \to \{-1, 1\}$ disagree. That is, $\text{dist}(f, g) = \Pr_{x \in \{-1,1\}^n}[f(x) \neq g(x)]$. For instance, the distance between two different parity functions is 1/2.

Fact 8 *For $S \neq T$,* $\text{dist}(\chi_S, \chi_T) = \frac{1}{2}$.

It turns out that Fourier coefficients can be used to express the distance of a function to a given linear function.

Fact 9 (Agreement of f with Linear Functions) *For $f : \{-1, 1\}^n \to \{-1, 1\}$,*

$$\hat{f}(S) = 1 - 2\,\text{dist}(f_1, \chi_S).$$

Proof

$$\begin{aligned}
2^n \hat{f}(s) &= \sum_x f(x)\chi_S(x) \\
&= \sum_{x \text{ s.t. } f(x) = \chi_S(x)} f(x)\chi_S(x) + \sum_{x \text{ s.t. } f(x) \neq \chi_S(x)} f(x)\chi_S(x) \\
&= 2^n - 2|\{x | f(x) \neq \chi_S(x)\}| \\
&= 2^n \left(1 - 2\frac{|\{x | f(x) \neq \chi_S(x)\}|}{2^n}\right) \\
\hat{f}(s) &= 1 - 2\,\text{dist}(f_1 \cdot \chi_S)
\end{aligned}$$

■

1.4 Plancherel's Theorem

The following simple theorem holds.

Theorem 10 (Plancherel's Theorem) *For $f, g : \{-1, 1\} \to \mathbb{R}$,*

$$\langle f, g \rangle = E_x[f(x) \cdot g(x)] = \sum_{S \subseteq [n]} \hat{f}(S)\hat{g}(S).$$

Proof

$$\begin{aligned}
\langle f, g \rangle &= \left\langle \sum_S \hat{f}(S)\chi_S(x), \sum_T \hat{g}(T)\chi_T(x) \right\rangle \\
&= \sum_S \sum_T \hat{f}(S)\hat{g}(T) \langle \chi_S(x), \chi_T(x) \rangle \\
&= \sum_{S=T} \hat{f}(S)\hat{g}(T) \cdot 1 = \sum_S \hat{f}(S)\hat{g}(S)
\end{aligned}$$

■

The theorem yields multiple useful properties.

Corollary 11 (Parseval's identity) *For* $f : \{-1,1\}^n \to R$, $\langle f, f \rangle = \sum \hat{f}^2(S)$.

Corollary 12 *For* $f : \{-1,1\}^n \to \{-1,1\}$, $\sum \hat{f}^2(S) = \langle f, f \rangle = 1$.

Corollary 13

$$E_x[\chi_S(x)] = \begin{cases} 1 & \textit{if } S = \emptyset, \\ 0, & \textit{otherwise.} \end{cases}$$

2 Analysis of the Proposed Linearity Tester

Recall that δ is the probability that a single pass through the loop detects that the input function f is not linear, and it can be expressed as

$$\delta = E_{x,y}\left[\frac{1 - f(x)f(y)f(x \cdot y)}{2}\right].$$

Lemma 14 (Main Lemma) $1 - \delta = \frac{1}{2} + \frac{1}{2}\sum_{S \subseteq [n]} \hat{f}^3(S)$

Proof

$$\begin{aligned}
1 - \delta &= E_{x,y}\left[\frac{1 + f(x)f(y)f(xy)}{2}\right] \\
&= \frac{1}{2} + \frac{1}{2}E_{x,y}[f(x)f(y)f(xy)] \\
E_{x,y}[f(x)f(y)f(xy)] &= E_{x,y}[(\sum_S \hat{f}(S)\chi_S(x))(\sum_T \hat{f}(T)\chi_T(y))(\sum_U \hat{f}(U)\chi_U(x \cdot y)] \\
&= \sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U)E_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x \cdot y)] \\
E_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x \cdot y)] &= E_{x,y}[\prod_{i \in S} x_i \prod_{j \in T} y_j \prod_{k \in U} x_k y_k] \\
&= E_{x,y}[\prod_{i \in S \triangle U} x_i \prod_{j \in T \triangle U} y_j] \\
&= E_x[\chi_{S \triangle U}(x)]E_y[\chi_{T \triangle U}(y)] \\
E_x[\chi_{S \triangle U}(x)] &= \begin{cases} 1 & \text{if } S = U, \\ 0, & \text{otherwise} \end{cases} \\
E_y[\chi_{T \triangle U}(y)] &= \begin{cases} 1 & \text{if } T = U, \\ 0, & \text{otherwise} \end{cases}
\end{aligned}$$

So, $E_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x\cdot y)]$ is non-zero if and only if $S = T = U$. If $S = T = U$, then the expectation is 1. Hence,

$$E_{x,y}[f(x)f(y)f(xy)] = \sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U)E_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x\cdot y)] = \sum_S \hat{f}^3(S)$$

and

$$1-\delta = \frac{1}{2} + \frac{1}{2}E_{x,y}[f(x)f(y)f(xy)] = \frac{1}{2} + \frac{1}{2}\sum_S \hat{f}^3(S).$$

■

Theorem 15 *If f is ϵ-far from linear, then $\delta = \Pr_{x,y}[f(x)f(y)f(x\cdot y) \neq 1] \geq \epsilon$.*

Proof
We will prove Theorem 10 by proving its contrapositive; we will assume that $\delta < \epsilon$, and demonstrate that this assumption implies that f is ϵ-close.

The Main Lemma implies that

$$\begin{aligned}
1-\delta &\leq \frac{1}{2} + \frac{1}{2}\sum_S \hat{f}^3(S) \\
1-2\delta &\leq \sum_S \hat{f}^3(S) \\
&\leq \left(\max_S \hat{f}(S)\right)\sum_S \hat{f}^2(S) = \max_S \hat{f}(S),
\end{aligned}$$

Let $T = \arg\max_S \hat{f}(S)$. We have

$$1-2\delta \leq \hat{f}(T),$$

and by Fact 9,

$$\mathrm{dist}(f,\chi_T) = \frac{1}{2} - \frac{1}{2}\hat{f}(T) < \frac{1}{2} - \frac{1}{2}(1-2\delta) = \delta < \epsilon.$$

Therefore, f is ϵ-close to a linear function; an impossibility. ■