6.842 Randomness and Computation	March 14, 2012
Lecture 11	
Lecturer: Ronitt Rubinfeld	*Scribe: Ishaan Chugh*

Today, we are going to continue Linearity Testing and introduce two very important tools

- Plancheral's Theorem,
- Parseval's Theorem

Using these tools we will finish our calculation of the rejection probability δ of the Linearity Test. We will then proceed to see how we can save Randomness in Linearity Testing as well.

Linearity Testing Review

Recall from last lecture that for a function $f: \{0,1\}^n \rightarrow \{0,1\}$, its linearity test involved picking randomly $\{x, y\} \subseteq_R \{\pm 1\}^n$. The test rejects if $f(x)f(y)f(x \cdot y) \neq 1$. Where $x \cdot y$ is bitwise times between x and y. That is each bit i of x and y is multiplied to give i'th bit of $x \cdot y$.

Also recall that we defined the indicator random variable for the test failing as follows.

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

Since the rejection probability of the test δ is equal to the expected value of the indicator variable

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

Also recall that we chose our basis functions to be character functions

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

We also proved a result which stated

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

where $\mathrm{dist}(f, \chi_{\mathrm{S}})$ is the *distance* in the sense it tells us how far from linear f is. To calculate δ we need two develop some tools which we will in the rest of the lecture.

Plancheral's Theorem

This is the main tool we will be using.

Theorem 1 ***Plancheral's Theorem*** *Given* $f: \{0,1\}^n \rightarrow \mathbb{R}$*, then*

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

Proof

$$\begin{aligned}
\langle f, g\rangle &= \langle \sum_S \hat{f}\chi_S, \sum_T \hat{g}\chi_T\rangle \\
&= \sum_S \sum_T \hat{f}(S)\hat{g}(T)\langle \chi_S\, \chi_T\rangle \qquad (\langle \chi_S\, \chi_T\rangle = 0 \text{ if S = T and 1 if S} \neq T \text{ because of orthonormality}) \\
&= \sum_{S\subseteq[n]} \hat{f}(S)\hat{g}(S)
\end{aligned}$$

■

Corollary 2 ***Parseval's Theorem****: Given* $f: \{0,1\}^n \to \mathbb{R}$*, then*

$$\langle f, f\rangle \equiv E_x[f^2(x)] = \sum_{S\subseteq[n]} \hat{f}^2(S)$$

This follows directly from Plancheral's Theorem by replacing g with f.

Corollary 3 ***BOOLEAN Parseval's Theorem****: Given* $f: \{0,1\}^n \to \{\pm\}$*, then*

$$\langle f, f\rangle \equiv E_x[f^2(x)] = E_x[1] = 1$$

This follows because f^2 is always 1.

Some Applications

1. $$E_x[f(x)] = E_x[f(x)\cdot 1] = \sum_S \hat{f}(S)\hat{\chi_\phi}(S) = \hat{f}(\phi)$$

   This immediately follows from Parseval's Theorem by replacing $g(x)$ with 1 which is χ_ϕ and using the fact that $\hat{\chi_\phi}(S) = 1$ if S $= \chi$ else 0.

2. $$E_x[\chi_S(x)] = \begin{cases} 1 & \text{if S} = \phi \\ 0 & \text{Otherwise} \end{cases}$$

GOAL TO SHOW relationship between δ and "distance" of f from linearity, i.e. if "distance" is high so is rejection probability.

Return to Linearity Testing

Now that we have the right tools we can analyse our Linearity Testing algorithm and find the relationship between the rejection probability δ and ϵ which was the measure of how far a function f is from being linear. Recall a function f is ϵ-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 ϵ-far to linear.

Main Goal

Show $\delta \geq \min_S Pr_x[f(x) \neq \chi_S(x)]$

Main Lemma

$$\begin{aligned} 1 - \delta &= Pr[f(x)f(y)f(x \cdot y) = 1] \\ &= \frac{1}{2} + \frac{1}{2} \sum_{S \subseteq [n]} \hat{f}^3(S) \end{aligned}$$

Proof

$$\begin{aligned} \delta &= \mathbb{E}_{x,y}\left[\frac{1 - f(x)f(y)f(x \cdot y)}{2}\right] \\ 1 - \delta &= E_{xy}[\frac{1 + f(x)f(y)f(x \cdot y)}{2}] \\ &= \frac{1}{2} + \frac{1}{2}E_{xy}[f(x)f(y)f(x \cdot y)] \end{aligned}$$

Let us calculate $E_{xy}[f(x)f(y)f(x \cdot y)]$ so that we can plug it in the expression above.

$$\begin{aligned} E_{xy}\Big[f(x)f(y)f(x \cdot y)\Big] &= E_{xy}\Big[\sum_S \hat{f}(S)\chi_S(x) \sum_T \hat{f}(T)\chi_T(x) \sum_U \hat{f}(U)\chi_U(x \cdot y)\Big] \\ &= E_{xy}\Big[\sum_S \sum_T \sum_U \hat{f}(S)\hat{f}(T)\hat{f}(U)\chi_S(x)\chi_T(x)\chi_U(x \cdot y)\Big] \\ &= \sum_S \sum_T \sum_U \hat{f}(S)\hat{f}(T)\hat{f}(U)E_{xy}\Big[\chi_S(x)\chi_T(x)\chi_U(x \cdot y)\Big] \end{aligned}$$

Note that in the above S,T and U are different only for convenience. It avoids variable confusion. In the above expression we wish to evaluate $E_{xy}\Big[\chi_S(x)\chi_T(x)\chi_U(x \cdot y)\Big]$. We will show this expectation is almost always 0 and plug this value in the expression above which will be then plugged in the expression above that to finally prove the main lemma.

$$\begin{aligned} E_{xy}\Big[\chi_S(x)\chi_T(x)\chi_U(x \cdot y)\Big] &= E_{xy}\Big[\prod_{i \in S} x_i \prod_{j \in T} x_i \prod_{k \in U} x_u y_u\Big] \\ &= E_{xy}\Big[\prod_{i \in S \Delta U} x_i \prod_{j \in T \Delta U}\Big] \end{aligned}$$

The x's and y's in S and T respectively "square up" with x's and y's in U to evaluate to 1. So

$$\begin{aligned} E_{xy}\Big[\chi_S(x)\chi_T(x)\chi_U(x \cdot y)\Big] &= E_{xy}\Big[\prod_{i \in S \Delta U} x_i \prod_{j \in T \Delta U} y_j\Big] \\ &= E_x[\prod_{i \in S \Delta U} x_i]E_y[\prod_{j \in T \Delta U} y_j] \end{aligned}$$

The last step follows because of the independence of x and y. We have

$$E_x[\prod_{i \in S \Delta U} x_i] = \begin{cases} 1 & \text{if S } \Delta \text{ U} = \phi \\ 0 & \text{Otherwise} \end{cases}$$

$$E_x[\prod_{j\in T\Delta U} y_j] = \begin{cases} 1 & \text{if T } \Delta \text{ U} = \phi \\ 0 & \text{Otherwise} \end{cases}$$

Thus,

$$E_{xy}\Big[\chi_S(x)\chi_T(x)\chi_U(x\cdot y)\Big] = \begin{cases} 1 & \text{if S=T=U} \\ 0 & \text{Otherwise} \end{cases}$$

Putting this in

$$E_{xy}\Big[\chi_S(x)\chi_T(x)\chi_U(x\cdot y)\Big] = \sum_S\sum_T\sum_U \hat{f}(S)\hat{f}(T)\hat{f}(U)E_{xy}\Big[\chi_S(x)\chi_T(x)\chi_U(x\cdot y)\Big]$$

We get

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

Thus

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

■

We will use this lemma to prove the following Theorem.

Theorem 4

$$if\ f:\{\pm1\}^n \to \{\pm1\}\ is\ \epsilon - far\ from\ linear\ then\ \delta \geq \epsilon$$

Proof

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

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

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

$$1-2\delta \leq max_S\hat{f}(S)\cdot\sum_S\hat{f}^2(S)$$

$$\leq max_S\hat{f}(S)$$

Pick T to maximise $\hat{f}(T)$ i.e T is our closes linear function s.t. dist(f, χ_T) = ϵ
1-2$\delta \leq \hat{f}(T)$ = 1 - 2dist(f,χ_T) = 1 - 2ϵ

Thus, $\delta \geq \epsilon$

■

Coppersmith's Example

Note that the the result we proved only holds for Boolean functions. This is a counterexample due to Don Coppersmith.

$$f(x) = \begin{cases} -1 & \text{if x = 2} \mod 3 \\ 0 & \text{if x = 0} \mod 3 \\ 1 & \text{if x = 1} \mod 3 \end{cases}$$

Table 1: Coppersmith's Counterexample

x *mod* 3	y *mod* 3	f(x+y)	f(x) + f(y)
-1	-1	1	-2
-1	0	-1	-1
-1	1	0	0
0	-1	-1	-1
0	0	0	0
0	1	1	1
1	-1	0	0
1	0	1	1
1	1	-1	2

$\delta = \frac{2}{9}$. The closest linear function to f is $g(x) = 0 \; \forall x \; \epsilon = dist(f, 0) = \frac{2}{3}$.
$\delta \ngeq \epsilon$

Saving Randomness in Linearity Testing

Graph Test

Given graph G with k nodes and edges E
Pick $x_1....x_k \; \epsilon_R \{\pm 1\}^n$
$\forall (i, j) \epsilon$
test if $f(x_i)f(x_j) = f(x_i; x_j)$
Accept if it always passes.
Queries to f : K + $|E|$. K for each of the vertices(each of which represent a random n-length string and $|E|$ for each of the edges (which represent pairs of random strings)

Behaviour

If f is LINEAR , it ALWAYS PASSES.

If f is NOT LINEAR P[accept] $\leq 2^{-|E|} + \max_\alpha |\hat{f_\alpha}|$.

Recall, that previously we proved a bound where P[accept] $\leq \frac{1}{2} + \max_\alpha \frac{|\hat{f_\alpha}|}{2}$.

We accept if $\prod_{(i,j) \in E)} \frac{1 + f(x_i)f(y_j)f(x_i \cdot y_j)}{2} = 1$
This is an indicator variable for the Test Passing. Thus, to calculate the probability of the test passing we can calculate the expectation of the indicator random variable.

$$E\left[\prod_{(i,j) \in E)} \frac{1 + f(x_i)f(y_j)f(x_i \cdot y_j)}{2} \right] = E\left[\frac{\sum_{S \subseteq E} \prod_{(i,j) \in S} f(x_i)f(y_j)f(x_i \cdot y_j)}{2^{|E|}} \right]$$

Lemma 5

$$\forall S \neq \phi \ E\left[\prod_{(i,j)\in S} f(x_i)f(x_j)f(x_i \cdot x_j)\right] \leq max_\alpha |\hat{f_\alpha}|$$

We will assume this lemma holds (prove it in the homework) and use it to find the bound on P[accept].

Proof

$$\begin{aligned}
P[accept] &= E\left[\frac{\sum_{S\neq\phi}\prod f(x_i)f(x_j)f(x_i \cdot x_j)}{2^{|E|}} + \frac{1}{2^{|E|}}\right] \\
&= \frac{1}{2^{|E|}} + \frac{2^{|E|} max_\alpha \hat{f}_\alpha}{2^{|E|}} \\
&= \frac{1}{2^{|E|}} + max_\alpha \hat{f}_\alpha
\end{aligned}$$

■