6.5420 Randomness and Computation March 31, 2026

Lecture 13

Lecturer: Ronitt Rubinfeld *Scribe: Axel Adjei*

1 Topics

- Fourier analysis on boolean hypercube
- Linearity test analysis

2 Review

2.1 Linear Functions

Recall the definition of linear functions from last lecture.

Definition 1. *A function* $f : \{\pm 1\}^n \to \{\pm 1\}$ *is said to be* ***linear*** *if for all* $x, y \in \{\pm 1\}^n$,

$$f(x)f(y) = f(x \odot y)$$

where "$\odot$*" denotes element-wise multiplication.*

From this definition, the following fact can readily be seen.

Fact 2. *All linear functions on* $\{\pm 1\}^n$ *are of the form*

$$\chi_S(x) = \prod_{i \in S} x_i$$

for some subset $S \subseteq [n]$. *Moreover, each function of this form is linear.*

2.2 Linear Testing

We are interested in efficiently testing whether a function is nearly linear, we formalize a notion of near linearity.

Definition 3. *A function* $f : \{\pm 1\}^n \to \{\pm 1\}$ *is said to be* ϵ***-linear*** *if there exists a linear function* g *such that*

$$\Pr_{x \sim \mathcal{U}}[f(x) = g(x)] \geq 1 - \epsilon$$

where $\mathcal{U}$ *denotes the uniform distribution on* $\{\pm 1\}^n$.

Definition 4. *A* ***linearity tester*** *is an algorithm that takes as input query access to a function* f *and a parameter* $\epsilon > 0$ *and satisfies the following requirements:*

- *If* f *is linear,* $\Pr[\textit{tester outputs "pass"}] = 1$.
- *If* f *is not* ϵ*-linear,* $\Pr[\textit{tester outputs "fail"}] \geq \frac{3}{4}$.

Here, $\frac{3}{4}$ *could be replaced by any number in* $(0, 1)$ *since we can always repeat run a tester multiple times to reduce the error probability.*

Consider the following proposed linearity tester.

Algorithm 1 Proposed Linearity Tester

Require: $\epsilon \geq 0, n \in \mathbb{N}$, query access to function $f : \{\pm 1\}^n \to \{\pm 1\}$.
  for $i = 1, \ldots, k$ **do**
    Sample $x, y \overset{i.i.d.}{\sim} \mathcal{U}$
    Reject if $f(x)f(y) \neq f(x \odot y)$.
  end for
  Accept.

Clearly, the tester always passes when f is linear, but we want to figure out how large k needs to be in order for the tester to fail with high probability when f is not ϵ-linear. In order to do so, we will need to use tools from Fourier analysis on the boolean cube.

3 Fourier Analysis

Definition 5. *Let $\mathcal{F}_n$ be the vector space of functions from $\{\pm 1\}^n$ to $\mathbb{R}$. For any $f, g \in \mathcal{F}_n$, we define the **inner product** of f and g as*

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

Theorem 6. *The set of linear functions $\{\chi_S\}_{S \subseteq [n]}$ is an orthonormal basis with respect to the inner product.*

Proof. We first prove that the χ_S each have norm 1. Let $S \subseteq [n]$.

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

We now prove that the linear functions are orthogonal. Let $S, T \subseteq [n]$ where $S \neq T$. Let $k \in S \triangle T$. Consider the set

$$A = \{x \in \{\pm 1\}^n \mid x_k = 1\}.$$

For any $x = (x_1, \ldots, x_{k-1}, 1, x_{k+1}, \ldots, x_n) \in A$, let $x^{\oplus k} = (x_1, \ldots, x_{k-1}, -1, x_{k+1}, \ldots, x_n)$. Clearly, the map $x \mapsto x^{\oplus k}$ is a bijection between A and $\{\pm 1\}^n \setminus A$. We use this to compute the inner product of χ_S and χ_T.

$$\begin{aligned}
\langle \chi_S, \chi_T \rangle &= \frac{1}{2^n} \sum_{x \in S} \chi_S(x)\chi_T(x) \\
&= \frac{1}{2^n} \sum_{x \in S} \prod_{i \in S} x_i \prod_{j \in T} x_j \\
&= \frac{1}{2^n} \sum_{x \in S} \left(\prod_{i \in S \cap T} x_i \prod_{i \in S \setminus T} x_i \right) \left(\prod_{j \in S \cap T} x_j \prod_{i \in T \setminus S} x_j \right) \\
&= \frac{1}{2^n} \sum_{x \in S} \prod_{i \in S \cap T} x_i^2 \prod_{j \in S \triangle T} x_j \\
&= \frac{1}{2^n} \sum_{x \in S} \prod_{j \in S \triangle T} x_j
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2^n}\left(\sum_{x\in A}\prod_{j\in S\triangle T} x_j + \sum_{x\in\{\pm1\}^n\setminus A}\prod_{j\in S\triangle T} x_j\right)\\
&= \frac{1}{2^n}\left(\sum_{x\in A}\prod_{j\in S\triangle T} x_j + \sum_{x\in A}\prod_{j\in S\triangle T} x_j^{\oplus k}\right)\\
&= \frac{1}{2^n}\left(\sum_{x\in A}\prod_{j\in (S\triangle T)\setminus\{k\}} x_j + \sum_{x\in A} -\prod_{j\in (S\triangle T)\setminus\{k\}} x_j\right)\\
&= \frac{1}{2^n}(0)\\
&= 0.
\end{aligned}
$$

Therefore, the χ_S are orthonormal. Since there are 2^n of them and the dimension of $\mathcal{F}_n$ is 2^n, they must span all of $\mathcal{F}_n$ and, thus, form an orthonormal basis. □

This implies that any function in $\mathcal{F}_n$ can be written as a linear combination of linear functions, with coefficient obtained by taking inner products with the χ_S. This is formalized by the following definition and corollary.

Definition 7. *For $f : \{\pm1\}^n \to \mathbb{R}$ and $S \subseteq [n]$, define*

$$\hat{f}(S) = \langle f, \chi_S \rangle.$$

Corollary 8. *For all $f : \{\pm1\}^n \to \mathbb{R}$,*

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

Fact 9. *A function $f : \{\pm1\}^n \to \{\pm1\}$ is linear if and only if there exists a set S such that $\hat{f}(S) = 1$ and $\hat{f}(T) = 0$ for all $T \neq S$.*

Next, we show that when $\hat{f}(S)$ is large, f is close to χ_S.

Lemma 10. *For all $f : \{\pm1\}^n \to \{\pm1\}$ and $S \subseteq [n]$,*

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

where $\mathrm{dist}(f, \chi_S) = \Pr_{x\sim\mathcal{U}}[f(x) = \chi_S(x)]$.

Proof. From definitions 5 and 7,

$$
\begin{aligned}
\hat{f}(S) &= \langle f, \chi_S \rangle\\
&= \frac{1}{2^n}\sum_{x\in\{\pm1\}^n} f(x)\chi_S(x)\\
&= \frac{1}{2^n}\sum_{x:f(x)=\chi_S(x)} f(x)\chi_S(x) + \frac{1}{2^n}\sum_{x:f(x)\neq\chi_S(x)} f(x)\chi_S(x)\\
&= \frac{1}{2^n}\sum_{x:f(x)=\chi_S(x)} (1) + \frac{1}{2^n}\sum_{x:f(x)\neq\chi_S(x)} (-1)\\
&= \Pr_{x\sim\mathcal{U}}[f(x) = \chi_S(x)] - \Pr_{x\sim\mathcal{U}}[f(x) \neq \chi_S(x)]
\end{aligned}
$$

$$= 1 - \Pr_{x\sim\mathcal{U}}[f(x) \neq \chi_S(x)] \Pr_{x\sim\mathcal{U}}[f(x) \neq \chi_S(x)]$$
$$= 1 - 2\Pr_{x\sim\mathcal{U}}[f(x) \neq \chi_S(x)].$$

□

From this, we observe the following.

Observation 11. *For all $S, T \subseteq [n]$ where $S \neq T$,*

$$\Pr_{x\sim\mathcal{U}}[\chi_S(x) = \chi_T(x)] = \frac{1}{2}.$$

Proof. We can prove the observation by combining the fact that χ_S and χ_T are orthogonal with lemma 10.

$$1 - 2\Pr_{x\sim\mathcal{U}}[\chi_S(x) = \chi_T(x)] = \langle \chi_S, \chi_T \rangle = 0.$$

Rearranging terms yields the statement of the observation. □

3.1 Identities

We are nearly ready to analyze the linearity test proposed in the previous section. We just need a few more identities that allow us to express inner products in terms of Fourier coefficients.

Theorem 12 (Plancherel's Identity)**.** . *For any pair of functions $f, g : \{\pm 1\}^n \to \mathbb{R}$,*

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

Proof. Using Corollary 8,

$$\begin{aligned}
\langle f, g \rangle &= \left\langle \sum_{S\subseteq[n]} \hat{f}(S)\chi_S, \sum_{T\subseteq[n]} \hat{g}(T)\chi_T \right\rangle \\
&= \sum_{S,T\subseteq[n]} \hat{f}(S)\hat{g}(T)\langle \chi_S, \chi_T \rangle \\
&= \sum_{S,T\subseteq[n]} \hat{f}(S)\hat{g}(T)\mathbb{1}\{S = T\} \\
&= \sum_{S\subseteq[n]} \hat{f}(S)\hat{g}(S)
\end{aligned}$$

.

where we use Theorem 6 for the penultimate equality. □

The special case where $f = g$ is of particular importance.

Corollary 13 (Parseval's Identity)**.** *For any function $f : \{\pm 1\}^n \to \mathbb{R}$,*

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

When the image of f is $\{\pm 1\}$, we get an even stronger result.

Corollary 14 (Boolean Parseval's Identity)**.** *For any function $f : \{\pm 1\}^n \to \{\pm 1\}$,*

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

Proof. From Parseval's Identity,

$$\sum_{S \subseteq [n]} \hat{f}(S)^2 = \langle f, f \rangle = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} f(x)^2 = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} 1 = 1.$$

□

4 Linearity Test Analysis

We are finally ready to analyze our proposed linearity tester. Fix a function $f : \{\pm 1\}^n \to \{\pm 1\}$. For any $x, y \in \{\pm 1\}^n$,

$$\frac{1 - f(x)f(y)f(x \odot y)}{2} = \mathbb{1}\{f(x)f(y) \neq f(x \odot y)\}$$

Thus, this quotient is the indicator for the pair x, y causing the tester to reject. Let δ_f be the probability that the tester fails on a particular iteration. Then,

$$\delta_f = \Pr_{x,y \overset{i.i.d.}{\sim} \mathcal{U}} [f(x)f(y) \neq f(x \odot y)] = \mathbf{E}_{x,y}\left[\frac{1 - f(x)f(y)f(x \odot y)}{2}\right]$$

Theorem 15. *Any function $f : \{\pm 1\}^n \to \{\pm 1\}$ is δ_f-linear*

Before we prove the theorem, we state and prove a useful lemma.

Lemma 16. *For any $S, T, U \subseteq [n]$,*

$$\mathbf{E}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x \odot y)] = \mathbb{1}\{S = T = U\}$$

Proof. For any x, y,

$$\chi_U(x \odot y) = \prod_{k \in U} x_k y_k = \prod_{k \in U} x_k \prod_{k \in U} y_k = \chi_U(x)\chi_U(y).$$

Thus,

$$\begin{aligned}
\mathbf{E}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x \odot y)] &= \mathbf{E}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x)\chi_U(y)] \\
&= \mathbf{E}_x[\chi_S(x)\chi_U(x)]\mathbf{E}_y[\chi_T(y)\chi_U(y)] \\
&= \left(\frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \chi_S(x)\chi_U(x)\right)\left(\frac{1}{2^n} \sum_{y \in \{\pm 1\}^n} \chi_T(y)\chi_U(y)\right) \\
&= \langle \chi_S, \chi_U \rangle \langle \chi_T, \chi_U \rangle \\
&= \mathbb{1}\{S = U\}\mathbb{1}\{T = U\} \\
&= \mathbb{1}\{S = T = U\}.
\end{aligned}$$

where we use the fact that x and y are independent for the second equality and the orthonormality of the linear functions for the fifth. □

Proof of Theorem 15. We begin by calculating the expectation of $f(x)f(y)f(x \odot y)$.

$$\begin{aligned}
\mathbf{E}_{x,y}[f(x)f(y)f(x \odot y)] &= \mathbf{E}_{x,y}\left[\left(\sum_{S\subseteq[n]} \hat{f}(S)\chi_S(x)\right)\left(\sum_{T\subseteq[n]} \hat{f}(T)\chi_T(y)\right)\left(\sum_{U\subseteq[n]} \hat{f}(U)\chi_U(x\odot y)\right)\right] \\
&= \mathbf{E}_{x,y}\left[\sum_{S,T,U\subseteq[n]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\chi_S(x)\chi_T(y)\chi_U(x\odot y)\right] \\
&= \sum_{S,T,U\subseteq[n]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x\odot y)] \\
&= \sum_{S,T,U\subseteq[n]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbb{1}\{S=T=U\} \\
&= \sum_{S\subseteq[n]} \hat{f}(S)^3 \\
&\leq \max_{S\subseteq[n]} \hat{f}(S) \cdot \sum_{S\subseteq[n]} \hat{f}(S)^2 \\
&= \max_{S\subseteq[n]} \hat{f}(S) \\
&= \max_{S\subseteq[n]}(1 - 2\Pr[f(x) \neq \chi_S(x)]) \\
&= 1 - 2\min_{S\subseteq[n]} \Pr[f(x) \neq \chi_S(x)])
\end{aligned}$$

Therefore,

$$\begin{aligned}
\delta_f &= \mathbf{E}_{x,y}\left[\frac{1 - f(x)f(y)f(x\odot y)}{2}\right] \\
&= \frac{1}{2} - \frac{1}{2}\mathbf{E}_{x,y}[f(x)f(y)f(x\odot y)] \\
&\geq \frac{1}{2} - \frac{1}{2}\left(1 - 2\min_{S\subseteq[n]} \Pr[f(x) \neq \chi_S(x)])\right) \\
&= \min_{S\subseteq[n]} \Pr[f(x) \neq \chi_S(x)])
\end{aligned}$$

Thus, there exists a linear function χ_S such that $\Pr[f(x) = \chi_S(x)] \geq 1 - \delta_f$ so f is δ_f-linear. ☐

Corollary 17. *If we run the linearity testing algorithm for* $k = \left\lceil \frac{\log 4}{\epsilon} \right\rceil$ *iterations, if* f *is not* ϵ*-linear, the tester rejects it with probability at least* $\frac{3}{4}$.

Proof. Suppose f is not ϵ-linear. From Theorem 15, we know that f is δ_f-linear, so $\delta_f > \epsilon$. The probability that the tester accepts f is

$$\Pr[\text{tester outputs "pass"}] = (1-\delta_f)^k \leq e^{-\delta_f k} \leq e^{-\delta_f \frac{\log 4}{\epsilon}} \leq e^{-\log 4} = \frac{1}{4}.$$

Therefore, the tester rejects with probability at least $\frac{3}{4}$. ☐

What's surprising about this result is that the number of iterations does not depend on the dimension n.