Sublinear Algorithms

Lecture 26

Sofya Raskhodnikova

Penn State University

Thanks to Madhav Jha (Penn State) for help with creating these slides.

Testing Linearity

Linear Functions Over Finite Field $\mathbb{F}_2$

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

$$f(x_1, \dots, x_n) = a_1 x_1 + \cdots + a_n x_n \text{ for some } a_1, \dots, a_n \in \{0,1\}$$

no free term

- Work in finite field $\mathbb{F}_2$
  - Other accepted notation for $\mathbb{F}_2$: GF_2 and $\mathbb{Z}_2$
  - Addition and multiplication is mod 2
  - $\boldsymbol{x}=(x_1, \dots, x_n)$, $\boldsymbol{y}=(y_1, \dots, y_n)$, that is, $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$
    $\boldsymbol{x} + \boldsymbol{y}=(x_1 + y_1, \dots, x_n + y_n)$

example

$$\begin{array}{r} 001001 \\ +\ 011001 \\ \hline 010000 \end{array}$$

Testing if a Boolean function is Linear

Input: Boolean function $f: \{0,1\}^n \to \{0,1\}$

Question:

Is the function linear or ε-far from linear

($\geq \varepsilon 2^n$ values need to be changed to make it linear)?

Today: can answer in $O\left(\frac{1}{\varepsilon}\right)$ time

Motivation

- Linearity test is one of the most celebrated testing algorithms
  - A special case of many important property tests
  - Computations over finite fields are used in
    - Cryptography
    - Coding Theory
  - Originally designed for program checkers and self-correctors
  - Low-degree testing is needed in constructions of Probabilistically Checkable Proofs (PCPs)
    - Used for proving inapproximability
- Main tool in the correctness proof: Fourier analysis of Boolean functions
  - Powerful and widely used technique in understanding the structure of Boolean functions

Equivalent Definitions of Linear Functions

Definition. f is *linear* if $f(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n$ for some $a_1, \ldots, a_n \in \mathbb{F}_2$

$$\Updownarrow$$

$[n]$ is a shorthand for $\{1, \ldots n\}$

$$f(x_1, \ldots, x_n) = \sum_{i \in S} x_i \text{ for some } S \subseteq [n].$$

Definition′. f is *linear* if $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$.

- Definition ⇒ Definition′
  $f(\boldsymbol{x} + \boldsymbol{y}) = \sum_{i \in S} (\boldsymbol{x} + \boldsymbol{y})_i = \sum_{i \in S} x_i + \sum_{i \in S} y_i = f(\boldsymbol{x}) + f(\boldsymbol{y}).$
- Definition′ ⇒ Definition

Let $\alpha_i = f((\overbrace{0, \ldots, 0,1,0, \ldots, 0}^{e_i}))$

Repeatedly apply Definition′:

$$f\big((x_1, \ldots, x_n)\big) = f\left(\sum x_i e_i\right) = \sum x_i f(e_i) = \sum \alpha_i x_i.$$

Based on Ryan O'Donell's lecture notes: http://www.cs.cmu.edu/~odonnell/boolean-analysis/

Linearity Test [Blum Luby Rubinfeld 90]

BLR Test (f, ε)

1. Pick $\boldsymbol{x}$ and $\boldsymbol{y}$ independently and uniformly at random from $\{0,1\}^n$.
2. Set $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$ and query f on $\boldsymbol{x}, \boldsymbol{y}$, and $\boldsymbol{z}$. **Accept** iff $f(\boldsymbol{z}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.

Analysis

If f is linear, BLR always accepts.

Correctness Theorem [Bellare Coppersmith Hastad Kiwi Sudan 95]

If f is ε-far from linear then $> \varepsilon$ fraction of pairs $\boldsymbol{x}$ and $\boldsymbol{y}$ fail BLR test.

- Then, by Witness Lemma (Lecture 1), $2/\varepsilon$ iterations suffice.

Analysis Technique: Fourier Expansion

Representing Functions as Vectors

Stack the 2^n values of $f(\boldsymbol{x})$ and treat it as a vector in $\{0,1\}^{2^n}$.

$$f = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} f(0000) \\ f(0001) \\ f(0010) \\ f(0011) \\ f(0100) \\ \cdot \\ \cdot \\ \cdot \\ f(1101) \\ f(1110) \\ f(1111) \end{bmatrix}$$

Linear functions

There are 2^n linear functions: one for each subset $S \subseteq [n]$.

$$\chi_\emptyset = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \chi_{\{1\}} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \cdots \cdots, \qquad \chi_{[n]} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Parity on the positions indexed by set S is $\chi_S(x_1, \dots, x_n) = \sum_{i \in S} x_i$

Great Notational Switch

Idea: Change notation, so that we work over reals instead of a finite field.

- Vectors in $\{0,1\}^{2^n}$ $\longrightarrow$ Vectors in $\mathbb{R}^{2^n}$.
- 0/False $\longrightarrow$ 1 1/True $\longrightarrow$ -1.
- Addition (mod 2) $\longrightarrow$ Multiplication in $\mathbb{R}$.
- Boolean function: $f : \{-1, 1\}^n \rightarrow \{-1,1\}$.
- Linear function χ_S: $\{-1, 1\}^n \rightarrow \{-1,1\}$ is given by $\chi_S(\boldsymbol{x}) = \prod_{i \in S} x_i$.

Benefit 1 of New Notation

- The dot product of f and g as vectors in $\{-1,1\}^{2^n}$:

$$(\#\ \boldsymbol{x}\text{'s such that } f(\boldsymbol{x}) = g(\boldsymbol{x})) - (\#\ \boldsymbol{x}\text{'s such that } f(\boldsymbol{x}) \neq g(\boldsymbol{x}))$$

$$= 2^n - 2 \cdot \underbrace{(\#\ \boldsymbol{x}\text{'s such that } f(\boldsymbol{x}) \neq g(\boldsymbol{x}))}_{\textit{disagreements between } f \textit{ and } g}$$

Inner product of functions $f, g : \{-1,1\} \to \{-1,1\}$

$$\langle f, g\rangle = \frac{1}{2^n}(\text{dot product of } f \text{ and } g \text{ as vectors})$$

$$= \operatorname*{avg}_{\boldsymbol{x}\in\{-1,1\}^n}[f(\boldsymbol{x})g(\boldsymbol{x})] = \mathop{\mathrm{E}}_{\boldsymbol{x}\in\{-1,1\}^n}[f(\boldsymbol{x})g(\boldsymbol{x})].$$

$\langle f, g\rangle = 1 - 2 \cdot$ (fraction of *disagreements* between f and g)

Benefit 2 of New Notation

Claim. The functions $(\chi_S)_{S\subseteq[n]}$ form an orthonormal basis for $\mathbb{R}^{2^n}$.

- If $S \neq T$ then χ_S and χ_T are orthogonal: $\langle \chi_S, \chi_T \rangle = 0$.
  - Let i be an element on which S and T differ (w.l.o.g. $i \in S \setminus T$)
  - Pair up all n-bit strings: $(\boldsymbol{x}, \boldsymbol{x}^{(i)})$ where $\boldsymbol{x}^{(i)}$ is $\boldsymbol{x}$ with the i^{th} bit flipped.
  - Each such pair contributes $ab - ab = 0$ to $\langle \chi_S, \chi_T \rangle$.
  - Since all $\boldsymbol{x}$'s are paired up, $\langle \chi_S, \chi_T \rangle = 0$.
- Recall that there are 2^n linear functions χ_S .
- $\langle \chi_S, \chi_S \rangle = 1$
  - In fact, $\langle f, f \rangle = 1$ for all $f : \{-1,1\}^n \to \{-1,1\}$.
  - (The norm of f, denoted $|f|$, is $\sqrt{\langle f, f \rangle}$)

$$\begin{array}{c|c|c}
 & +1 & -1\\
 & -1 & +1\\
 & +1 & +1\\
\boldsymbol{x} & +a & b\\
 & +1 & +1\\
 & \vdots & \vdots\\
\boldsymbol{x}^{(i)} & -a & b\\
 & +1 & -1\\
 & -1 & +1\\
 & -1 & +1\\
\hline
 & \chi_S & \chi_T
\end{array}$$

Fourier Expansion Theorem

Idea: Work in the basis $(\chi_S)_{S\subseteq[n]}$, so it is easy to see how close a specific function f is to each of the linear functions.

Fourier Expansion Theorem

Every function $f:\{-1,1\}^n \to \mathbb{R}$ is uniquely expressible as a linear combination (over $\mathbb{R}$) of the 2^n linear functions:

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

where $\hat{f}(S) = \langle f, \chi_S\rangle$ is the Fourier Coefficient of f on set S.

Proof: f can be written uniquely as a linear combination of basis vectors:

$$f = \sum_{S\subseteq[n]} c_S \cdot \chi_S$$

It remains to prove that $c_S=\hat{f}(S)$ for all S.

$$\hat{f}(S) = \langle f,\chi_S\rangle = \left\langle \sum_{T\subseteq[n]} c_T \cdot \chi_T , \chi_S \right\rangle = \sum_{T\subseteq[n]} c_T \cdot \langle \chi_T, \chi_S\rangle = c_S$$

Definition of Fourier coefficients

Linearity of $\langle\cdot,\cdot\rangle$

$\langle \chi_T, \chi_S\rangle = \begin{cases} 1 & \text{if } T = S \\ 0 & \text{otherwise} \end{cases}$

Examples: Fourier Expansion

$\boldsymbol{f}$	**Fourier transform**
$f(\boldsymbol{x}) = 1$	1
$f(\boldsymbol{x}) = x_i$	x_i
$\text{AND}(x_1, x_2)$	$\frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1x_2$
$\text{MAJORITY}(x_1, x_2, x_3)$	$\frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$

Parseval Equality

Parseval Equality

Let $f:\{-1,1\}^n \to \mathbb{R}$. Then

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

Proof:

By Fourier Expansion Theorem

$$\langle f, f \rangle = \left\langle \sum_{S \subseteq [n]} \hat{f}(S)\chi_S \,,\, \sum_{T \subseteq [n]} \hat{f}(T)\chi_T \right\rangle$$

By linearity of inner product

$$= \sum_S \sum_T \hat{f}(S)\,\hat{f}(T)\langle \chi_S, \chi_T \rangle$$

By orthonormality of χ_S's

$$= \sum_S \hat{f}(S)^2$$

Parseval Equality

Parseval Equality for Boolean Functions

Let $f: \{-1,1\}^n \to \{-1,1\}$. Then

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

Proof:

By definition of inner product

$$\langle f, f \rangle = \mathop{\mathrm{E}}_{\boldsymbol{x} \in \{-1,1\}^n}[f(\boldsymbol{x})^2]$$

Since f is Boolean

$$= 1$$

BLR Test in {-1,1} notation

BLR Test (f, ε)

1. Pick $\boldsymbol{x}$ and $\boldsymbol{y}$ independently and uniformly at random from $\{-1,1\}^n$.
2. Set $\boldsymbol{z} = \boldsymbol{x} \circ \boldsymbol{y}$ and query f on $\boldsymbol{x}, \boldsymbol{y}$, and $\boldsymbol{z}$. **Accept** iff $f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{z}) = 1$.

Vector product notation: $\boldsymbol{x} \circ \boldsymbol{y} = (x_1 y_1, x_2 y_2, \dots, x_n y_n)$

Sum-Of-Cubes Lemma. $\Pr_{\mathbf{x},\mathbf{y} \in \{-1,1\}^n}[\mathrm{BLR}(f)\text{accepts}] = \frac{1}{2} + \frac{1}{2} \sum_{S \subseteq [n]} \hat{f}(S)^3$

Proof: Indicator variable

$$\mathbb{1}_{BLR} = \begin{cases} 1 & \text{if BLR accepts} \\ 0 & \text{otherwise} \end{cases} \Rightarrow \mathbb{1}_{BLR} = \tfrac{1}{2} + \tfrac{1}{2} f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{z}).$$

$$\Pr_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^n}[\mathrm{BLR}(f)\text{accepts}] = \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y} \in \{-1,1\}^n}[\mathbb{1}_{BLR}] = \frac{1}{2} + \frac{1}{2} \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y} \in \{-1,1\}^n}[f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{z})]$$

By linearity of expectation

Proof of Sum-Of-Cubes Lemma

So far: $\Pr_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\text{BLR}(f)\text{accepts}] = \frac{1}{2} + \frac{1}{2}\mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{z})]$

Next:

$$\mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{z})]$$

By Fourier Expansion Theorem

$$= \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}\left[\left(\sum_{S\subseteq[n]}\hat{f}(S)\chi_S(\boldsymbol{x})\right)\left(\sum_{T\subseteq[n]}\hat{f}(T)\chi_T(\boldsymbol{y})\right)\left(\sum_{U\subseteq[n]}\hat{f}(U)\chi_U(\boldsymbol{z})\right)\right]$$

Distributing out the product of sums

$$= \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}\left[\left(\sum_{S,T,U\subseteq[n]}\hat{f}(S)\hat{f}(T)\hat{f}(U)\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})\right)\right]$$

By linearity of expectation

$$= \sum_{S,T,U\subseteq[n]}\hat{f}(S)\hat{f}(T)\hat{f}(U)\mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})]$$

Proof of Sum-Of-Cubes Lemma (Continued)

$$\Pr_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\mathrm{BLR}(f)\text{accepts}] = \frac{1}{2} + \frac{1}{2}\sum_{S,T,U\subseteq[n]} \hat{f}(S)\hat{f}(T)\hat{f}(U) \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\mathbf{x})\chi_T(\mathbf{y})\chi_U(\mathbf{z})]$$

Claim. $\mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\mathbf{x})\chi_T(\mathbf{y})\chi_U(\mathbf{z})]$ is 1 if $S = T = U$ and 0 otherwise. ✓

- Let $S\Delta T$ denote symmetric difference of sets S and T

$$\mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\mathbf{x})\chi_T(\mathbf{y})\chi_U(\mathbf{z})] = \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}\left[\prod_{i\in S} x_i \prod_{i\in T} y_i \prod_{i\in U} z_i\right]$$

$$= \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}\left[\prod_{i\in S} x_i \prod_{i\in T} y_i \prod_{i\in U} x_i y_i\right]$$

Since $\mathbf{z} = \mathbf{x} \circ \mathbf{y}$

$$= \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}\left[\prod_{i\in S\Delta U} x_i \prod_{i\in T\Delta U} y_i\right]$$

Since $x_i^2 = y_i^2 = 1$

$$= \mathop{\mathrm{E}}_{\mathbf{x}\in\{-1,1\}^n}\left[\prod_{i\in S\Delta U} x_i\right] \cdot \mathop{\mathrm{E}}_{\mathbf{y}\in\{-1,1\}^n}\left[\prod_{i\in S\Delta U} y_i\right]$$

Since $\mathbf{x}$ and $\mathbf{y}$ are independent

$$= \prod_{i\in S\Delta U} \mathop{\mathrm{E}}_{\mathbf{x}\in\{-1,1\}^n}[x_i] \cdot \prod_{i\in T\Delta U} \mathop{\mathrm{E}}_{\mathbf{y}\in\{-1,1\}^n}[y_i]$$

Since $\mathbf{x}$ and $\mathbf{y}$'s coordinates are independent

$$= \prod_{i\in S\Delta U} \mathop{\mathrm{E}}_{x_i\in\{-1,1\}}[x_i] \cdot \prod_{i\in T\Delta U} \mathop{\mathrm{E}}_{y_i\in\{-1,1\}}[y_i]$$

$$= \begin{cases} 1 & \text{when } S\Delta U = \emptyset \text{ and } T\Delta U = \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Proof of Sum-Of-Cubes Lemma (Done)

$$\Pr_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\mathrm{BLR}(f)\text{accepts}] = \frac{1}{2} + \frac{1}{2}\sum_{S,T,U\subseteq[n]} \hat{f}(S)\hat{f}(T)\hat{f}(U) \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{z})]$$

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

Sum-Of-Cubes Lemma. $\Pr_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\mathrm{BLR}(f)\text{accepts}] = \frac{1}{2} + \frac{1}{2}\sum_{S\subseteq[n]} \hat{f}(S)^3$

Proof of Correctness Theorem

Correctness Theorem (restated)

If f is ε-far from linear then $\Pr[\mathrm{BLR}(f) \text{ accepts}] \leq 1-\varepsilon$.

Proof: Suppose to the contrary that

$$1-\varepsilon < \Pr_{\mathbf{x},\mathbf{y}\in\{-1,1\}^n}[\mathrm{BLR}(f)\text{accepts}]$$

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

By Sum-Of-Cubes Lemma

$$\leq \frac{1}{2} + \frac{1}{2}\cdot\left(\max_{S\subseteq[n]} \hat{f}(S)\right)\cdot \sum_{S\subseteq[n]} \hat{f}(S)^2$$

Since $\hat{f}(S)^2 \geq 0$

$$= \frac{1}{2} + \frac{1}{2}\cdot\left(\max_{S\subseteq[n]} \hat{f}(S)\right)$$

Parseval Equality

- Then $\max_{S\subseteq[n]} \hat{f}(S) > 1-2\varepsilon$. That is, $\hat{f}(T) > 1-2\varepsilon$ for some $T\subseteq[n]$.
- But $\hat{f}(T) = \langle f, \chi_T\rangle = 1 - 2\cdot$(fraction of *disagreements* between f and χ_T)
- f disagrees with a linear function χ_T on $<\varepsilon$ fraction of values. ※

Summary

BLR tests whether a function $f: \{0,1\}^n \to \{0,1\}$ is

linear or ε-far from linear

($\geq \varepsilon 2^n$ values need to be changed to make it linear)

in $O\left(\frac{1}{\varepsilon}\right)$ time.