Agenda

Analysis of Boolean Functions

1) Fourier Analysis / Poly. over real numbers
2) Spectral graph theory (hypercube graph)
3) Applying Fourier analysis to property testing

Based on Ryan O'Donnell's lecture written by Sarah Allen

18.434

Analysis of Boolean Functions

Paul Hager
4.13.16

Today and Friday David and I will be discussing the class of functions called Boolean Functions.

$$\mathcal{F} = \{ f \mid f : \{0,1\}^n \to \{0,1\} \}$$

Any function that maps a string of n binary digits to a single bit. We tend to think of "0" as a symbol for False and 1 as a symbol for True. Then all of these functions are functions from vectors in $\mathbb{F}_2^n$ to $\{0,1\}$.

As long the alphabet of the domain is two elements we can use two arbitrary elements. It will be convenient to use

$$\mathcal{F} = \{ f \mid f : \{-1,1\}^n \to \{-1,1\} \}$$

or more generally

$$\mathcal{F} = \{ f \mid f : \{-1,1\}^n \to \mathbb{R} \}$$

One often considers one of the following common Boolean functions.

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

Majority$_n(x)$: the result is the decision made by a majority of the inputs. Assuming n is odd $\text{sgn}\left(\sum_{i=1}^{n} x_i\right)$

Parity: $\begin{cases} -1 & \text{odd number of } -1 \text{ inputs} \\ 1 & \text{even " " " "} \end{cases}$

Dictatorship$_i$: The result is the i^{th} input

The Fourier expansion of the majority$_3$ function is

$$\boxed{Maj_3(x) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3}$$

which can be verified by arithmatic or verifying each assignment.

The Fourier expansion of

$$Parity_3(x) = x_1x_2x_3 \qquad \text{Remember } \{-1,1\}^n \to \{-1,1\}$$

If asked, $Parity_3(x) = x_1 + x_2 + x_3$ for $\{0,1\}^n \to \{0,1\}$

These are just a few examples but

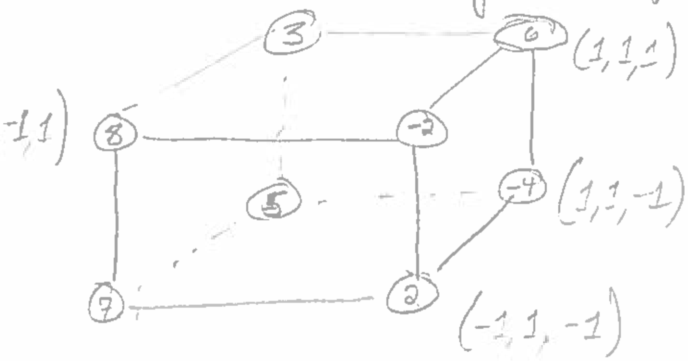
Claim Every function $f:\{-1,1\}^n \to \{-1,1\}$ can be uniquely computed as a multilinear polynomial.

Boolean Functions $\{-1,1\}^n \to \mathbb{R}$

Can be thought of as functions over the vertices of a hypercube. We will consider this running example in three dimensions.

Associating a value $a_v \in \mathbb{R}$ $\forall v \in \{-1,1\}^n$

$f: \{-1,1\}^3 \to \mathbb{R}$



Want to find a poly. that Langrangian interpolates the function, f. For each dimension, i, of the cube we want to consider a term that will be equal to 1 when $x_i = x_i$ is satisfied and 0 when x_j does not satisfied.

Specifically consider constructing a term for the $(1,-1,-1)$ vertex of the 3-dimensional hypercube.

$\left(\frac{1}{2} + \frac{x_1}{2}\right)$ for the first dimension

$\left(\frac{1}{2} - \frac{x_2}{2}\right)$ " " second "

$\left(\frac{1}{2} - \frac{x_3}{2}\right)$ " " third "

~~To ensure the term is no zero only when all three constraints~~ are satisfied, one uses the product. To ensure the function takes on the specified value of f one also has to multiply the product by $f(1,-1,-1)$

Ex.:

$$5 \cdot \left(\frac{1}{2} + \frac{x_1}{2}\right) \cdot \left(\frac{1}{2} - \frac{x_2}{2}\right) \cdot \left(\frac{1}{2} - \frac{x_3}{2}\right) = \text{The } (1,-1,-1) \text{ term of the polynomial}$$

If one sums the terms created for each possible assignment, one creates the polynomial.

One calls this the Fourier expansion.

Pf

Fourier expansion of f

$$f(x) = \sum_{S \subseteq [n]} c_S \prod_{i \in S} x_i$$

Notational Definition

$c_S \triangleq \hat{f}(S)$ This is called the S- Fourier Coefficient of f

There are 2^n subsets, S, $S \subseteq [n]$ which means there must be 2^n monomials. Therefore for every f there are 2^n parameters to select and exactly 2^n coefficients for each monomial. This implies the uniqueness of the fourier expansion.

Notational Definition

$$\chi_S = \prod_{i \in S} x_i \qquad \chi_\phi = 1$$

Note: $\chi_S = \text{Parity}_S(x)$

All Boolean functions are affine combinations of parity functions of the variables.

Spectral Graph Theory

$G \triangleq (V, E)$ Hypercube

n-regular graph

Consider function, f.

$\mathrm{Mean}(f) \triangleq \mathop{\mathbb{E}}_{x \sim \pi}[f(x)]$ where π is a uniform distribution over all vertices $\{-1,1\}^n$.

Definition

f is "Balanced" $\Leftrightarrow$ $\mathrm{Mean}(f) = 0$ Important for Friday's lectures.

$$\mathrm{Var}(f) = \mathop{\mathbb{E}}_{x \sim \pi}\left[f(x)^2\right] - \mathop{\mathbb{E}}_{x \sim \pi}[f(x)]^2$$

This term will always be 1 if $f : \{-1,1\}^n \rightarrow \{-1,1\}$.

Recall that for a vector space and functions f, g the inner product has been defined as

---

$$\langle f, g \rangle = \mathop{\mathbb{E}}_{x \sim \pi}[f(x)g(x)]$$

for Boolean Functions $\{-1,1\} \rightarrow \{-1,1\}$

$$= \left(1 - \mathop{P}_{x}(f(x) \neq g(x))\right) - \mathop{P}_{x}(f(x) \neq g(x))$$

if $f(x) = g(x)$ $\langle f, g \rangle = 1$

$$= 1 - 2 \mathop{P}_{x}(f(x) \neq g(x))$$

For use later in property testing

Parity functions form a basis for the vector space

Parity Functions — Eigenfunctions

From S.G.T.

K = normalized adjacency matrix "Markovian Operator"

$L = I - K$ "Laplacian Operator"

Both have the same eigenfunctions and

L's eigenvalues $l_i = 1 - k_i$ where k_i is an eigenvalue of K

$K f(x) = \underset{y \in N(x)}{\mathbb{E}}[f(y)]$ "The expected value of f on y, a random neighbor of x"

Claim 1 Eigenfunctions of K on the hypercube are the parity functions on every subset.

Pf Statement says $\forall S \subseteq [n] \quad \exists k$ s.t. $K \chi_S = k \chi_S$

for a given assignment x and a given vertex x_i ~~of x_i's n neighbors $n-|S|$ have the same value in f,~~ Because the parity$_S$ function does not include them.

$$K \chi_S = \left(\frac{n-2|S|}{n}\right)\chi_S = \left(1 - \frac{2|S|}{n}\right)\chi_S$$

$\left(1 - \frac{2|S|}{n}\right)$ is eigenvalue of χ_S

ranges from -1 to 1

Because this forms a complete basis every function can be a linear combo. of scaled eigenfunctions this is equivalent to what was shown in the Fourier analysis.

Claim 2 The Parity functions for all $S \subseteq [n]$

Pf Consider $S, T \subseteq [n]$

Statement says

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

$$\langle \chi_S, \chi_T \rangle = \mathop{\mathbb{E}}_{\pi} \left[\chi_S \chi_T \right]$$ π defined above

$$= \mathop{\mathbb{E}}_{\pi} \left[\prod_{i \in S} x_i \prod_{i \in T} x_i \right]$$

$$= \mathop{\mathbb{E}}_{\pi} \left[\prod_{i \in S \cap T} x_i^2 \prod_{i \in S \Delta T} x_i \right]$$ $S \Delta T$ is the symmetric difference of S and T

$$= \mathop{\mathbb{E}}_{\pi} \left[\prod_{i \in S \cap T} 1 \prod_{i \in S \Delta T} x_i \right]$$ Recall $f : \{-1, 1\}^n \to \{-1, 1\}$

$$= \mathop{\mathbb{E}}_{\pi} \left[\prod_{i \in S \Delta T} x_i \right]$$

Case I $S = T$

$$= \mathop{\mathbb{E}}_{\pi} [1] = 1$$ because $S \Delta T = \phi$

✓

Case II $S \neq T$

$$= \prod_{i \in S \Delta T} \mathop{\mathbb{E}}_{\pi} [x_i]$$ by independence of x_i

$$= \prod_{i \in S \Delta T} 0 = 0$$

✓

□

Now all functions can be uniquely written in this orthogonal basis.

Claim 3 The Fourier Coefficients are exactly the innerproducts of f onto the basis vectors. $\hat{f}(S) = \langle f, \chi_S \rangle \; \forall S \subseteq [n]$

Proof

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

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

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

$$= \hat{f}(S)$$ By $\langle A, A \rangle = 1$ and the orthogonality of the parity functions

□

The inner product is the correlation between f and parity_S

A Corrolary of Claim 3

Parseval **Thm.** for functions f, g

1799

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

Pf

$$\langle f, g \rangle = \mathbb{E}_X \left[f(x) g(x) \right]$$ $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S$

$$= \mathbb{E}_X \left[\sum_{S,T \subseteq [n]} \hat{f}(S) \chi_S \, \hat{g}(T) \chi_T \right]$$

$$= \mathbb{E}_X \left[\sum_{S \subseteq [n]} \hat{f}(S) \hat{g}(S) \right]$$ Orthogonality

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

Important Special Case

Note: $\langle f, f \rangle = 1 = \sum \hat{f}(S)^2$

Property Testing

Querying every possible input to a Boolean function is too costly. One can use a randomized algorithm to test whether a function has a property $\mathcal{P}$. One particularly interesting application is testing a property on a graph and are only able to query whether an edge exists between two vertices. This is made possible when the length of the input to the boolean function, n, is $\binom{|V|}{2}$ of the graph. Notation: If f has property $\mathcal{P}$, $f \in \mathcal{P}$.

When designing randomized algorithms we want the following two properties to be true of the algorithms.

1) If $f \in \mathcal{P}$, algo. outputs Yes with high probability. $(\mathbb{P} > \frac{1}{2})$.

   Algo. is <u>complete</u>.

2) If $f \notin \mathcal{P}$, algo. outputs No with high probability.

   Algo. is <u>sound</u>.

~~For the complexity analysis, we will assume an oracle~~ exists which given an assignment x, will return $f(x)$ in $O(1)$ time

Now there is a complication with the satisfying the soundness condition. Consider the property $\mathcal{P} = \{f \mid f \equiv 1\}$, to verify $f \notin \mathcal{P}$ one would have to query the oracle with all 2^n assignments of x. We introduce a relax the soundness condition in the following way.

<u>Definition</u> f is ϵ-far from $\mathcal{P}$ if $\forall g \in \mathcal{P}$, $\mathbb{P}_x[f(x) = g(x)] > \varepsilon$ or $\langle f, g \rangle < 1 - 2\varepsilon$.

2) If f is ε-far from $\mathcal{P}$, output No with high probability.

   Algo. is sound.

Theorem BLR 90

Property $\mathcal{P}$ where $f \in \mathcal{P}$ iff f is linear function. is testable in $O(\frac{1}{\varepsilon})$ queries

this means $f(x \cdot y) = f(x) f(y)$ where $(x \cdot y)_i = x_i \cdot y_i$. We will propose an algorithm and check that it satisfies both constraints.

Proposed Algorithm: Subroutine 1

1. Select $x, y \sim \{-1, 1\}^n$ independently and uniformly
2. Compute $z = x \cdot y$
3. Query $f(x), f(y), f(z)$
4. If $f(z) \neq f(x) f(y)$, return No

Run subroutine 1 $O(\frac{1}{\varepsilon})$ times.
If none return No, return Yes

If f is linear all of the subroutines will suceed so $\mathbb{P}[\text{Yes}] = 1$.
Algo. is complete.

If f is not linear, the $\mathbb{P}[\text{Yes}] = \mathop{\mathbb{E}}_{x,y}\left[\frac{1}{2} + \frac{1}{2} f(x) f(y) f(z)\right]$

$$\mathbb{P}[\text{Yes}] = \frac{1}{2} + \frac{1}{2} \mathop{\mathbb{E}}_{x,y}\left[\left(\sum_{S \subseteq [n]} \hat{f}(S) \prod_{i \in S} x_i\right)\left(\sum_{T \subseteq [n]} \hat{f}(T) \prod_{i \in T} y_i\right)\left(\sum_{U \subseteq [n]} \hat{f}(U) \prod_{i \in U} z_i\right)\right]$$

$$= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \subseteq [n]} \hat{f}(S) \hat{f}(T) \hat{f}(U) \mathop{\mathbb{E}}_{x,y}\left[\prod_{i \in S} x_i \prod_{i \in T} y_i \prod_{i \in U} x_i y_i\right]$$

$$= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \subseteq [n]} \hat{f}(S) \hat{f}(T) \hat{f}(U) \mathop{\mathbb{E}}_{x,y}\left[\prod_{i \in S \Delta U} x_i \prod_{i \in T \Delta U} y_i\right]$$

Using the symmetric difference trick used in Claim 2

$$= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \subseteq [n]} \hat{f}(S) \hat{f}(T) \hat{f}(U) \mathop{\mathbb{E}}_{x}\left[\prod_{i \in S \Delta U} x_i\right] \mathop{\mathbb{E}}_{y}\left[\prod_{i \in T \Delta U} y_i\right]$$

Due to the independent selection of x and y.

$\mathbb{P}[\text{Yes}] = \frac{1}{2} + \frac{1}{2} \sum_{S \subseteq [n]} (\hat{f}(S))^3$ Unless $S \triangle U$ and $T \triangle U$ are $\emptyset$ the terms zero.

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

$= \frac{1}{2} + \frac{1}{2} \max_S (\hat{f}(S))$ By Parseval's Theorem

$\Rightarrow P[\text{No}] = 1 - \mathbb{P}[\text{Yes}] \geq \frac{1}{2} + \frac{1}{2} \max_S (\hat{f}(S))$ which further implies the algorithm is sound.

The number of queries the algo. makes is $3O\left(\frac{1}{\varepsilon}\right) \in O\left(\frac{1}{\varepsilon}\right)$ $\square$