CAS CS 599 B: Mathematical Methods for TCS

Lecturer: Mark Bun | Spring 2022

Lecture Notes 2:

BLR Test, Influence

Reading.

- O'Donnell, Analysis of Boolean Functions §1.4-1.5, 2.1-2.2

1 Using the Fourier representation

The Fourier coefficients of a function f encode interesting information about its "average case" behavior. For example:

Proposition 1.

$$\mathop{\mathbb{E}}_{x\sim\{-1,1\}^n}[f(x)] = \hat{f}(\emptyset).$$

Proof. This follows from the fact that $\mathbb{E}[f] = \langle f, 1\rangle = \langle f, \chi_\emptyset\rangle = \hat{f}(\emptyset)$. □

Similarly, we can compute the variance of f, i.e., the variance of the random variable $f(x)$ where x is uniformly random.

Proposition 2.

$$\mathop{\mathrm{Var}}_{x\sim\{-1,1\}^n}[f(x)] = \sum_{S\neq\emptyset}\hat{f}(S)^2.$$

Proof. We calculate

$$\begin{aligned}
\mathop{\mathrm{Var}}_{x\sim\{-1,1\}^n}[f(x)] &= \mathbb{E}[(f-\mathbb{E}[f])^2] \\
&= \langle f-\mathbb{E}[f], f-\mathbb{E}[f]\rangle \\
&= \mathbb{E}[f^2] - (\mathbb{E}[f])^2 \\
&= \left(\sum_{S\subseteq[n]}\hat{f}(S)^2\right) - \hat{f}(\emptyset)^2
\end{aligned}$$

□

where the last equality follows from Parseval and Proposition 1. The same calculation using Plancharel in place of Parseval shows that the *covariance* between two functions f, g is $\sum_{S\neq\emptyset}\hat{f}(S)\hat{g}(S)$.

Finally, an operation on two functions that comes up from time to time is the *convolution*.

Definition 3. For $f, g : \{-1,1\}^n \to \mathbb{R}$, define the convolution

$$(f * g)(x) = \mathop{\mathbb{E}}_{y \sim \{-1,1\}^n} [f(y) g(x \circ y)] = \mathop{\mathbb{E}}_{y \sim \{-1,1\}^n} [f(x \circ y) g(y)]$$

where $\circ$ denotes the entrywise product.

If $f, g : \{-1,1\}^n \to [0,1]$ represent the probability density functions of random variables X, Y, then $f * g$ is the probability density of $X \circ Y$.

Theorem 4.

$$\widehat{f * g}(S) = \hat{f}(S) \hat{g}(S).$$

Proof.

$$\begin{aligned}
\widehat{f * g}(S) &= \mathop{\mathbb{E}}_{x \sim \{-1,1\}^n} [(f * g)(x) \chi_S(x)] \\
&= \mathop{\mathbb{E}}_{x \sim \{-1,1\}^n} \left[\mathop{\mathbb{E}}_{y \sim \{-1,1\}^n} [f(y) g(x \circ y)] \, \chi_S(x) \right] \\
&= \mathop{\mathbb{E}}_{z \sim \{-1,1\}^n} \left[\mathop{\mathbb{E}}_{y \sim \{-1,1\}^n} [f(y) g(z)] \, \chi_S(y \circ z) \right] \\
&= \mathop{\mathbb{E}}_{y, z \sim \{-1,1\}^n} [f(y) g(z) \chi_S(y) \chi_S(z)] \\
&= \hat{f}(S) \hat{g}(S).
\end{aligned}$$

$\square$

2 BLR Test

Suppose your friend hands you a Boolean function $f : \{-1,1\}^n \to \{-1,1\}$ and tells you it's supposed to be a parity function. The way the function is given to you is as a "black box", meaning you can evaluate it at points $x \in \{-1,1\}^n$ of your choice, but you otherwise can't see the internal structure of how f is constructed or computed. How can you verify that f is indeed a parity function?

If you want to do this perfectly, you'll need to query f at all 2^n locations, since your friend could have given you a function that agrees with a parity on all but one point. But if you are willing to settle for an approximation, you can be convinced that f is *close* to a parity using only three queries!

This is the kind of problem studied in the field of *property testing*. A property testing problem is a set $\mathcal{P}$ of Boolean functions (e.g., the set of all parity functions), called a "property". An instance of such a problem is a Boolean function f. A tester makes a small number of (possibly random) queries to f and has the following guarantees.

Completeness: If $f \in \mathcal{P}$, the test *accepts* with high probability.

Soundness: If f is "far from" $\mathcal{P}$, the test *rejects* with high probability. For a parameter $\varepsilon > 0$, we say that f is ε-far from $\mathcal{P}$ if $\operatorname{dist}(f, g) \geq \varepsilon$ for all $g \in \mathcal{P}$, where $\operatorname{dist}(f, g) = \Pr_{x \sim \{-1,1\}^n}[f(x) \neq g(x)]$.

Property testing (in particular, the parity tester) was originally motivated by self-testing/correcting programs and probabilistically checkable proofs. Since then, it has blossomed into a vibrant area of research in its own right. See, e.g., Oded Goldreich's book `https://www.wisdom.weizmann.ac.il/~oded/pt-intro.html` or Sofya Raskhodnikova's BU class `http://cs-people.bu.edu/sofya/sublinear-course/`.

The three-query BLR parity tester is motivated by the following observation. A parity function χ_S respects multiplication in the sense that for every x, y,

$$\chi_S(x \circ y) = \prod_{i \in S} x_i y_i = \left(\prod_{i \in S} x_i\right)\left(\prod_{i \in S} y_i\right) = \chi_S(x)\chi_S(y).$$

And in fact, the converse is true: Every Boolean function that respects multiplication is a parity on some subset of variables. The BLR tester is based on the idea that this property holds *robustly*. That is, a Boolean function respects multiplication for "most" inputs if and only if its is close to a parity function. Here is the tester:

BLR Tester Given query access to a function $f : \{-1,1\}^n \to \{-1,1\}$:

1. Sample $x, y \sim \{-1,1\}^n$ independently.
2. Query f at x, y, and $x \circ y$.
3. *Accept* if $f(x \circ y) = f(x)f(y)$. *Reject* otherwise.

2.1 Analysis of the BLR Test

Completeness: If f is a parity function, then the tester accepts with probability 1, since $f(x \circ y) = f(x)f(y)$ for every pair x, y.

Soundness: This follows from the contrapositive of the following theorem. Strictly speaking, this only shows that if f is ε-far from every parity, then the test rejects with small probability ε. But the soundness error can be amplified by repeating the test k times independently and rejecting if any of the k runs rejects. Thus, using $3k$ queries, the BLR test is guaranteed to reject any ε-far function f with probability at least $1-(1-\varepsilon)^k$.

Theorem 5. *If the BLR test accepts f with probability $1-\varepsilon$, then f is ε-close to a parity function.*

Proof. Observe that

$$\frac{1}{2} + \frac{1}{2}f(x)f(y)f(x \circ y) = \begin{cases} 1 & \text{if } f(x)f(y) = f(x \circ y) \\ 0 & \text{otherwise,} \end{cases}$$

so

$$\Pr[\text{test accepts } f] = \mathop{\mathbb{E}}_{x,y}\left[\frac{1}{2} + \frac{1}{2}f(x)f(y)f(x \circ y)\right] \implies 2\Pr[\text{test accepts } f] - 1 = \mathop{\mathbb{E}}_{x,y}\left[f(x)f(y)f(x \circ y)\right].$$

Estimating this latter quantity, we get

$$\begin{aligned}
\mathop{\mathbb{E}}_{x,y}[f(x)f(y)f(x\circ y)] &= \mathop{\mathbb{E}}_{x}\left[f(x)\mathop{\mathbb{E}}_{y}[f(y)f(x\circ y)]\right] \\
&= \mathop{\mathbb{E}}_{x}[f(x)\cdot(f*f)(x)] \\
&= \sum_{S\subseteq[n]} \hat{f}(S)\widehat{f*f}(S) && \text{by Plancharel} \\
&= \sum_{S\subseteq[n]} \hat{f}(S)^3 && \text{by Thm 4} \\
&\le \max_{S\subseteq[n]} \hat{f}(S)\cdot \sum_{S\subseteq[n]} \hat{f}(S)^2 \\
&= \max_{S\subseteq[n]} \hat{f}(S) && \text{by Parseval.}
\end{aligned}$$

Hence if the BLR test accepts f with probability at least $1-\varepsilon$, there exists an S such that

$$2(1-\varepsilon)-1 \le \hat{f}(S) = \langle f, \chi_S\rangle = 1 - 2\,\mathrm{dist}(f,\chi_S).$$

Rearranging shows that there exists an S such that $\mathrm{dist}(f,\chi_S)\le\varepsilon$. □

Historical notes: The BLR test was discovered around 1990 by Blum, Luby, and Rubinfeld, who gave a combinatorial proof. The Fourier analytic proof here is due to Bellare, Coppersmith, Håstad, Kiwi, and Sudan. The parity test usually goes by the name "linearity test" because the parity functions are exactly the linear functions over $\mathbb{F}_2^n$.

3 Influence

We'll be studying a number of "proto-complexity" measures of Boolean functions including influence, noise stability, and spectral concentration using Fourier analysis. These quantities are interesting in their own right, as well as useful for downstream applications in circuit complexity, pseudorandomness, learning, and more.

For a string $x\in\{-1,1\}^n$, an index $i\in[n]$, and $b\in\{-1,1\}$.

- $x^{\oplus i} = (x_1, x_2, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n)$.
- $x^{(i\to b)} = (x_1, x_2, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_n)$

Definition 6 (Pivotal coordinate)**.** A coordinate $i\in[n]$ is pivotal for f at input x if $f(x)\neq f(x^{\oplus i})$

Definition 7 (Influence)**.** The influence of coordinate i on f is the probability that i is pivotal for a random input:

$$\mathrm{Inf}_i[f] = \Pr_{x\sim\{-1,1\}^n}[f(x)\neq f(x^{\oplus i})]$$

Influence has a nice interpretation in terms of social choice. In social choice theory, one thinks of the function f as a voting rule mapping n votes for candidates "+1" and "-1" to an election outcome. Under what's called the "impartial culture assumption" that votes are uniformly random, the influence of coordinate i captures the probability that i's vote swings the election.

Example 8. Let's estimate the influences of a few natural voting rules.

- The function AND_n elects candidate -1 if and only if all voters vote for -1. For every i, its influence is $\mathrm{Inf}_i(\mathrm{AND}_n) = \Pr_{x\sim\{-1,1\}^n}[x_1 = x_2 = \cdots = x_{i-1} = x_{i+1} = \cdots = x_n = -1] = 2^{-n+1}$.

- For odd n, the function MAJ_n elects the candidate that wins the majority vote. For every i, its influence is
$$\mathrm{Inf}_i(\mathrm{MAJ}_n) = \binom{n-1}{\frac{n-1}{2}} \cdot 2^{-n+1} \approx \frac{\sqrt{2/\pi}}{\sqrt{n}}.$$
This follows from Stirling's Formula, which says that $m! = (m/e)^m(\sqrt{2\pi m} + O(1/\sqrt{m}))$.

- For a coordinate j, define the j'th dictator function by $\chi_j(x) = x_j$. Then we have
$$\mathrm{Inf}_i(\chi_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$