A Brief Introduction to Fourier Analysis on the Boolean Cube: Advances and Challenges

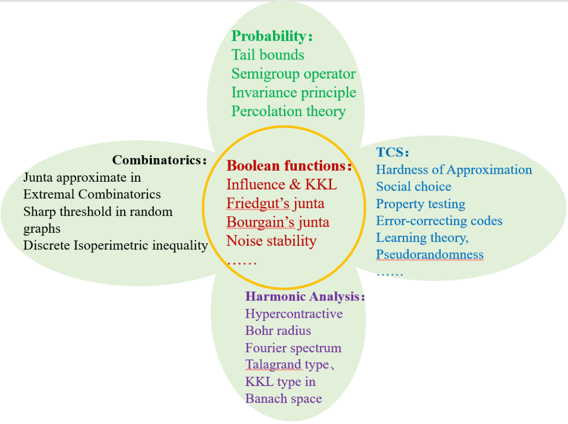
Fan Chang

Shandong University

Contents

Classical Boolean function analysis

- Central object: $f : \{\pm 1\}^n \to \{\pm 1\}$ or real-valued functions on the Boolean cube $\{\pm 1\}^n$.
- Boolean-valued functions appear frequently in theoretical computer science and mathematics.



The Fourier–Walsh Basis

- Consider the coordinates x_i of a vector x as functions on $\{\pm 1\}^n$.
  For every $S \subset \{1, \ldots, n\}$ define
  $$\chi_S := \prod_{i \in S} x_i, \chi_\emptyset \equiv 1.$$
- This set, of 2^n monomials, forms an orthonormal basis of the space of real functions on $\{\pm 1\}^n$.

Fundamental theorem of Boolean function analysis

Every function $f : \{\pm 1\}^n \to \mathbb{R}$ has unique expansion as multilinear polynomial, the *Fourier expansion*:

$$f(x_1, \ldots, x_n) = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S.$$

Definitions and Basic Properties

Consider the 2^n-dimensional vector space of all functions $f : \{\pm 1\}^n \to \mathbb{R}$.

- (Inner product) $\langle f, g\rangle = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} f(x)g(x) = \mathop{\mathbb{E}}_{x \sim \{\pm 1\}^n} [f \cdot g]$.
- (ℓ_p-norm) $\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$.
- (Fourier coefficients) $\hat{f}(S) := \langle f, \chi_S \rangle = \mathbb{E}[f \cdot \chi_S]$.
- The *degree* of f is $\deg(f) = \max\{|S| : \hat{f}(S) \neq 0\}$.
- (Plancherel) $\langle f, g\rangle = \sum_{S \subseteq [n]} \hat{f}(S)\hat{g}(S)$.
- (Parseval's identity) $\|f\|_2^2 = \sum_{S \subseteq [n]} \hat{f}(S)^2$.

If f is Boolean, then

$$1 = \mathbb{E}[|f|^2] = \|f\|_2^2 = \sum_{S \subseteq [n]} \hat{f}(S)^2.$$

Boolean degree 1 functions

Question

Suppose $f : \{\pm 1\}^n \to \{\pm 1\}$ has degree 1. What does f look like?

- $\deg f \leq 1 \Leftrightarrow f(x_1, \ldots, x_n) = c_0 + \sum_{i=1}^{n} c_i x_i$.
- Dictator: function depending on one coordinate.

Dictator theorem

If $f : \{\pm 1\}^n \to \{\pm 1\}$ has degree 1, then

$$f \in \{\pm 1, \pm x_1, \ldots, \pm x_n\}.$$

Boolean almost degree 1 functions

Stability problem

Suppose $f : \{\pm 1\}^n \to \{\pm 1\}$ satisfies

$$\mathop{\mathbb{E}}_{x \sim \{\pm 1\}^n} [|f(x) - g(x)|^2] = \varepsilon$$

for some $g : \{\pm 1\}^n \to \mathbb{R}$ of degree 1.
What does f look like?

Friedgut–Kalai–Naor (FKN) theorem

Suppose $f : \{\pm 1\}^n \to \{\pm 1\}$ satisfies $\|f^{>1}\|_2^2 = \varepsilon$. Then

$$\mathbf{Pr}[f \neq h] = O(\varepsilon) \text{ for some } h \in \{\pm 1, \pm x_1, \dots, \pm x_n\}.$$

FKN theorem on the slice

The *slice* or *Johnson scheme* is

$$\binom{[n]}{k} = \{(x_1, \ldots, x_n) \in \{0,1\}^n : \sum_{i=1}^{n} x_i = k\}.$$

- (Dunkl) Every function $f : \binom{[n]}{k} \to \mathbb{R}$ has unique expansion as *harmonic* multilinear polynomial.

Dictator theorem on the slice

If $f : \binom{[n]}{k} \to \{0,1\}$ has degree 1 and $k \neq 1, n-1$, then

$$f \in \{0, 1, x_1, 1-x_1, \ldots, x_n, 1-x_n\}.$$

FKN theorem on the slice (Filmus 16)

Suppose $f : \binom{[n]}{k} \to \{0,1\}$ is ε-close to an affine function, where $2 \leq k \leq n-2$. Define $p := \min(k/n, 1-k/n)$. Then either f or $1-f$ is $O(\varepsilon)$-close to $\max_{i \in S} x_i$ (when $p \leq 1/2$) or to $\min_{i \in S} x_i$ (when $p \geq 1/2$) for some set S of size at most $\max(1, O(\sqrt{\varepsilon}/p))$.

Challenge 1. Boolean functions beyond the hypercube

Other domains:

- p-biased cube: important in random graph theory (Kahn–Kalai Conjecture & Park–Pham theorem).
- Johnson scheme/slice (all k-subsets of $[n]$): important in Erdős–Ko–Rado type theorem (Das–Tuan: "set family" version of removal lemma).
- Grassmann scheme (all k-dimensional subspaces of $\mathbb{F}_q^n$): used to prove 2-to-1 conjecture and 2-to-2 conjecture (Dinur–Khot–Kindler–Minzer–Safra).
- Symmetric group $S_n = \{\pi : [n] \to [n] | \pi$ is a permutation$\}$: representation theory (Ellis–Filmus–Friedgut).
- Other association schemes, Gaussian space, Cayley graphs of codes, high-dimensional expanders, ...

Contents

Influences

Given a voting rule $f : \{\pm 1\}^n \to \{\pm 1\}$ it's natural to try to measure the "influence" or "power" of the ith voter.

- (Probabilistic) $\mathrm{Inf}_i[f] = \mathbf{Pr}_{x \sim \{\pm 1\}^n}[f(x) \neq f(x^{(i)})]$, where $x^{(i)} := f(x \oplus e_i)$ is obtained from x by flipping the ith coordinate.
- (Analytically) The *ith (discrete) derivative operator* ∂_i maps the function $f : \{\pm 1\}^n \to \mathbb{R}$:

$$\partial_i f(x) = \frac{f(x^{(i\mapsto 1)}) - f(x^{(i\mapsto -1)})}{2}.$$

  $\mathrm{Inf}_i[f] = \mathbb{E}_x[\partial_i f(x)^2] = \|\partial_i f\|^2$.
- (Spectral) $\mathrm{Inf}_i[f] = \sum_{i \in S} \hat{f}(S)^2$.
- (Geometric) $\mathrm{Inf}_i[f]$ measures the number of edges of the cube crossing from A to A^c in direction i.

Total influence

The *total influence* :

$$I[f] = \sum_{i=1}^{n} \mathsf{Inf}_i[f].$$

- (Spectral) $I[f] = \sum_{i=1}^{n} \mathsf{Inf}_i[f] = \sum_{i=1}^{n} \sum_{i \in S} \hat{f}(S)^2 = \sum_S |S|\hat{f}(S)^2 = \sum_{d=0}^{n} d\|f^{=d}\|^2$.
- (Geometric: *edge boundary*) $I[f]$ measures the total number of edges crossing from A to its complement.
- (TCS) The *sensitivity* of f at a point x is the number of coordinates i such that $f(x) \neq f(x^{(i)})$, denoted by $\mathsf{sens}_f(x)$.

$$I[f] = \mathbb{E}_x[\mathsf{sens}_f(x)].$$

- (Analytically) The *(discrete) gradient operator* ∇ maps the function $f : \{\pm 1\}^n \to \mathbb{R}$:

$$\nabla f(x) = (\partial_1 f(x), \partial_2 f(x), \dots, \partial_n f(x)).$$

$I[f] = \mathbb{E}_x[\|\nabla f(x)\|^2]$.

Simple Example

Example 1.

- Let $X = (\{0,1\}, \mu_{1/2})$ be the uniform distribution on $\{0,1\}$.
- Let $f(x_1, \ldots, x_n) = x_1 + \cdots + x_n (\text{mod} 2)$.
- Total influence of f is $n/2$.

Example 2 (p-biased case).

- Let $X = (\{0,1\}, \mu_p)$ be the Bernoulli distribution with parameter $p = 1/n$.
- Let $f : (\{0,1\}^n, \mu_p^{\otimes n}) \to \{0,1\}$ be any function.
- Total influence of $I^p[f] \leq 2(n-1)/n \approx 2$.

$$I_i[f] := \mathbf{Pr}[f(x_1, \ldots, x_i, \ldots, x_n) \neq f(x_1, \ldots, x_i^{(i)}, \ldots, x_n)]$$
$$\leq \mathbf{Pr}[x_i \neq y_i] = 2p(1-p) = \frac{2(n-1)}{n^2}.$$

Challenge 2. Structure v.s. Pseudorandom

Main Question

What can we say about the structure of functions $f : (\Omega^n, \nu^{\otimes n}) \to \{0, 1\}$ with $I^\nu[f] \leq K$?

Research Progress

- Russo (1982): First systematic study of functions with low total influences.
- Kahn-Kalai-Linial (1988): $\exists i \in [n], \mathrm{Inf}_i[f] \geq c\frac{\log n}{n}\mathbb{V}[f]$.
- Talagrand (1993): Talagrand isoperimetric inequality (both discrete and continuous, and product and non-product models).
- Friedgut (1998): Every low total influence Boolean function is close to a *junta* (uniform case).
- Friedgut (JAMS1999): A complete characterization of graph properties with $I[\mathcal{G}] = O(1)$ on $G(n,p)$.
- Bourgain (JAMS1999): Partially extended this to general setting $f : X^n \to \{0,1\}$.
- Hatami (Annals2012): If the total influence of $f : X^n \to \{0,1\}$ is $O(1)$, then f is essentially a *pseudo-junta*.

Junta

Definition (Junta)

The value of $f(x_1, \ldots, x_n)$ depends on a small set of variables $\{x_{i_1}, \ldots, x_{i_k}\}$:

$$f(x_1, \ldots, x_n) := g(x_{i_1}, \ldots, x_{i_k}).$$

- Every variable outside the junta has influence 0.
- Juntas have total influence $O(1)$.

Friedgut's junta theorem 98

Let $f : \{\pm 1\}^n \to \{\pm 1\}$ with $I[f] \leq K$. Then for any $\varepsilon > 0$, there is a Boolean junta g, depending on $2^{O(K/\varepsilon)}$ coordinates, such that $\Pr[f \neq g] = \varepsilon$.

Challenge 3. Bourgain's Junta theorem

Bourgain's junta theorem 02

Let $f : \{\pm 1\}^n \to \{\pm 1\}$. Suppose that $\sum_{|S|>k} \hat{f}(S)^2 \leq (\frac{\varepsilon}{k})^{1/2+o(1)}$. Then for any $\varepsilon > 0$, there is a $2^{O(k)}/\varepsilon^{O(1)}$-junta g such that $\|f - g\|_2^2 \leq \varepsilon$.

Kindler–Kirshner–O'Donnell 18

Let $f : \{\pm 1\}^n \to \{\pm 1\}$. Suppose that $\sum_{|S|>k} \hat{f}(S)^2 \leq \frac{c\varepsilon}{\sqrt{k}}$. Then for any $\varepsilon > 0$, there is a $2^{O(k)}/\varepsilon^4$-junta g such that $\|f - g\|_2^2 \leq \varepsilon$.

- The constant $c = \frac{1}{3\pi} \approx 0.1061$. NOT optimal! (take $f(x) = \text{sgn}(x)$ and we will see $c = (\frac{2}{\pi})^{3/2} \approx 0.5079$.
- **Open.** Obtain a $2^{O(k)}/\varepsilon$-junta!

Challenge 4. Aaronson–Ambainis Conjecture

Aaronson–Ambainis Conjecture

Let $f : \{\pm 1\}^n \to [-1, 1]$ have degree at most k. Then there exists $i \in [n]$ such that $\mathrm{Inf}_i[f] \geq \left(\frac{\mathbb{V}[f]}{k}\right)^{O(1)}$.

- True for $f : \{\pm 1\}^n \to \{\pm 1\}$; this follows from a result of O'Donnell, Schramm, Saks, and Servedio.
- The weaker lower bound $(\mathbb{V}[f]/2^k)^{O(1)}$ follows from a result of Dinur, Kindler, Friedgut, and O'Donnell.
- True for many special cases (Montanaro 12,O'Donnell and Zhao 16, Defant, Mastylo and Perez 18, Bansal, Sinha and Wolf 22).

Dinur, Kindler, Friedgut, and O'Donnell 06

Let $f : \{\pm 1\}^n \to [-1, 1]$. Suppose that $\sum_{|S|>k} \hat{f}(S)^2 \leq \exp(-O(k^2 \log k)/\varepsilon)$. Then for any $\varepsilon > 0$, there is a $2^{O(k)}/\varepsilon^2$-junta g such that $\|f - g\|_2^2 \leq \varepsilon$.

Challenge 5. Generalized influences

- (Geometric influences) Keller, Mossel and Sen: for the Gaussian measure on $\mathbb{R}^n$, the geometric influences satisfy the KKL-type theorem.
- (Convex influences) De, Nadimpalli and Servedio introduce a new notion of influence for symmetric convex sets over Gaussian space.

Conjecture (Friedgut's Junta Theorem for convex influences).

Let $K \subseteq \mathbb{R}^n$ be a convex symmetric set with $I[K] \leq I$. Then there are $J \leq 2^{O(I/\varepsilon)}$ orthonormal directions $v^1, \ldots, v^J \in \mathbb{S}^{n-1}$ and a symmetric convex set $L \subseteq \mathbb{R}^n$, such that
(a) $L(x)$ depends only on the values of $v^1 \cdot x, \ldots, v^J \cdot x$, and
(b) $\mathbf{Pr}_{\mathbf{x} \sim \mathcal{N}(0,1)}[K(\mathbf{x}) \neq L(\mathbf{x})] \leq \varepsilon$.

Challenge 6. Junta approximation method in E.C.

Keller and Lifshitz develop the work of Dinur and Friedgut to present a general approach to such problems:

- Erdős *matching conjecture*
- Erdős-Sós *forbidding one intersection* problem
- Frankl-Füredi *special simplex* problem
- Ramsey-type problems (?)

Structural theorem (Keller–Lifshitz 2019)

Fixed hypergraph $\mathcal{H}$. Let $n \in \mathbb{N}, C < k < n/C$. Suppose that $\mathcal{F} \subseteq \binom{[n]}{k}$ is free of $\mathcal{H}^+$. Then there exists an $\mathcal{H}^+$-free junta $\mathcal{J} \subseteq \binom{[n]}{k}$ which depends on at most j coordinates, such that

$$|\mathcal{F} \setminus \mathcal{J}| \leq \max\left(e^{-k/C}, C\frac{k}{n}\right) \cdot |\mathcal{J}|.$$

Challenge 6. Junta approximation method in E.C.

The set family $\mathcal{F}$ is said to be a J-*junta* if it depends only upon the coordinates in J–formally, if $\exists \mathcal{G} \subset \mathcal{P}(J)$ s.t. $S \in \mathcal{F} \Leftrightarrow S \cap J \in \mathcal{G}$, for all $S \subset [n]$.

Dinur–Friedgut 08

For all $\eta > 0, \varepsilon > 0$ there exists $J \in \mathbb{N}$ such that the following holds. If $\mathcal{F} \subseteq \{0,1\}^n$ is an intersecting family, and $\eta < p < \frac{1}{2} - \eta$, then there exists an intersecting J-junta $\mathcal{J} \subseteq \{0,1\}^n$ such that $\mu_p(\mathcal{F} \setminus \mathcal{J}) \leq \varepsilon$.

- Define "monotonicity" and assume $\mathcal{F}$ to be monotone increasing.
- Define "pseudo-randomness", and show that any set family $=$ "Pseudo-random" sub-families $+$ junta sub-families.
- Pseudo-random families contain intersections of any constant size.

Contents

Noise operator

The tensor definition of the operator:

- Let $\rho \in [0,1]$. Let $f : \{-1,1\} \to \mathbb{R}, f(x) = ax + b$.
- Define $T_\rho(f)(x) := \rho a x + b$.
  Then $T_\rho := T_\rho^{\otimes n}$ is a linear operator acting on real functions on $\{-1,1\}^n$.

The spectral definition of the operator:

- Define $T_\rho x_i := \rho x_i$.
- $T_\rho \chi_S = \prod_{i \in S} T_\rho x_i = \rho^{|S|} \chi_S$.
- $T_\rho f = \sum_{S \subseteq [n]} \hat{f}(S) \rho^{|S|} \chi_S$.

Noise operator

The noise/averaging definition of the operator:

- Let $\rho \in [0,1]$. Let X be chosen from any distribution on $\{\pm 1\}^n$.
- Let Y be such that for every $1 \leq i \leq n$, the coordinate Y_i is chosen independently so that $\mathbf{Pr}[Y_i = X_i] = \frac{1+\rho}{2}$, or, in other words, $\mathbb{E}[X_i Y_i] = \rho$.
- X and Y are called an ρ-correlated pair.
- Define for any f and fixed X,

$$T_\rho(f)(X) = \mathbb{E}[f(Y)],$$

where X and Y are ρ-correlated pair.

Hypercontractivity

Hypercontractivity is the secret spice behind much of Boolean function analysis.

Bonami[68,70],Gross[75],Beckner[75]

Let $f : \{\pm 1\}^n \to \mathbb{R}$, and $\rho \in [0, 1]$. Then

$$\|T_\rho f\|_2 \leq \|f\|_{1+\rho^2}.$$

Dual version

Let $f : \{\pm 1\}^n \to \mathbb{R}$ be a polynomial of degree d, and $q \geq 2$. Then

$$\|f\|_q \leq (\sqrt{q-1})^d \|f\|_2.$$

Challenge 7. Global hypercontractivity and its applications

Keevash, Lifshitz, Long and Minzer (2019+, 2021+) establish an effective hypercontractive inequality for general p that applies to "global functions", i.e. functions that are not significantly affected by a restriction of a small set of coordinates.

- Strengthen Bourgain's sharp threshold theorem (quantitively tight and applicable in the sparse regime)
- A sharp threshold result for global monotone functions
- p-biased generalisation of the seminal invariance principle of Mossel, O'Donnell and Oleszkiewicz.
- Turán numbers for the family of bounded degree expanded hypergraphs.
- (Kaufan–Minzer 22+) Quantitative version of optimal tester for the Reed–Muller code.

Unfortunately, applications in a very large number of areas have to be completely left out, including in learning theory, pseudorandomness, arithmetic combinatorics, random graphs and percolation, communication complexity, coding theory, metric and Banach spaces,. . .

Thanks for your attention!