Boolean function analysis beyond the Boolean cube

Yuval Filmus

March 6, 2018

Yesterday we have described classical Boolean function analysis on the Boolean cube. Today we will describe some other domains, focusing mainly on association schemes and related structures, but also discussing differential posets.

1 Association schemes

A graph G is *distance-regular* if for every i, j, the size of the set

$$\{z : d(x, z) = i \text{ and } d(y, z) = j\}$$

depends only on $d(x, y)$. Here $d(x, y)$ is the distance between x and y in G.

A *metric* or *P-polynomial* association scheme is the same thing as a connected distance-regular graph.

Here is a simple example: the Boolean cube. The distance between two vertices is the Hamming distance of the corresponding Boolean vectors. The group S_n acts transitively on the Boolean cube in the natural way. The orbits of its action on *pairs* of vertices are $\{(x, y) : d(x, y) = d\}$. This implies that the Boolean cube is distance-regular. We can also give an explicit formula for the number of z such that $d(x, z) = i$ and $d(y, z) = j$ if $d(x, y) = k$:

$$N_{ijk} := \binom{k}{\frac{k-i+j}{2}}\binom{n-k}{\frac{i-k+j}{2}}.$$

The *Bose–Mesner algebra* of an association scheme consists of all $V \times V$ matrices A in which $A(x, y)$ depends only on $d(x, y)$, say $A(x, y) = \alpha(d(x, y))$. If A, B are in the algebra then

$$(AB)(x, y) = \sum_z A(x, z)B(z, y) = \sum_{i,j} N_{ijd(x,y)}\alpha(i)\beta(j),$$

and so the algebra is closed under multiplication. Moreover, clearly $N_{ijk} = N_{jik}$ (since $d(y, x) = d(x, y)$), and so the algebra is commutative. Its dimension is $1 + \text{diam}(G)$.

Since the Bose–Mesner algebra is commutative, all matrices in the algebra have the same $1 + \text{diam}(G)$ eigenspaces. For example, let us examine the Boolean cube once again, which we identify with $\{\pm 1\}^n$. Let $\chi_S(x) = \prod_{i \in S} x_i$. If A is in the Bose–Mesner algebra of the cube, then

$$(A\chi_S)(x) = \sum_y A(x, y)\chi_S(y) = \sum_y A(xy, \mathbf{1})\chi_S(y) = \sum_y A(y, \mathbf{1})\chi_S(xy) = \chi_S(x) \sum_y A(y, \mathbf{1})\chi_S(y).$$

This shows that χ_S is an eigenvector of A. Furthermore, the symmetry of the Boolean cube under the action of S_n guarantees that the eigenvalue associated with χ_S depends only on $|S|$. Thus A has $n+1$ eigenspaces. The dth eigenspace is spanned by $\{\chi_S : |S| = d\}$, and consists of all homogeneous degree d multilinear polynomials.

Let us denote the dth common eigenspace by V_d, and the projection of a function f on the Boolean cube to V_d by $f^{=d}$. Notice that the degree of f is the maximal d such that $f^{=d} \neq 0$. This notion of degree satisfies the property of subadditivity:

$$\deg(fg) \leq \deg f + \deg g.$$

An association scheme which satisfies this property for some ordering of the common eigenspaces is known as *cometric* or *Q-polynomial*.

1.1 Some examples

There are four standard examples of association schemes which are both metric and cometric (there are other examples as well, such as polar spaces).

Hamming scheme The Hamming scheme $H(n, d)$ is the graph on $\mathbb{Z}_d^n$ in which two points (x, y) are connected if they differ in one coordinate. When $d = 2$, this is just the Boolean cube. When $d > 2$, we get the *multicube*. The Hamming scheme can be studied as an association scheme, but a more refined analysis follows from using Fourier analysis on the group $\mathbb{Z}_d^n$.

Johnson scheme The Johnson scheme $J(n, k)$ (also known as the *slice*) is the graph on $\binom{[n]}{k}$ in which two sets S, T are connected if $|S \triangle T| = 2$ or, equivalently, if $|S \cap T| = k - 1$. More generally, $d(S, T) = |S \triangle T|/2 = k - |S \cap T|$. Boolean function analysis on the Johnson scheme has been thoroughly studied, as we explain in some detail later on.

Grassmann scheme The Grassmann scheme $J_q(n, k)$, where q is a prime power, is the graph of all k-dimensional subspaces of an $\mathbb{F}_q$-vector space of dimension n. Two subspaces U, V are connected if $\dim(U \cap V) = k - 1$, and more generally $d(U, V) = k - \dim(U \cap V)$. The number of points in the scheme is given by the q-binomial coefficient $\begin{bmatrix} n \\ k \end{bmatrix}_q$.

The definition of the Grassmann scheme is very similar to that of the Johnson scheme, and it is known as the q-analog of the Grassmann scheme; the Johnson scheme is the case $q = 1$. For example, $\lim_{q \to 1} \begin{bmatrix} n \\ k \end{bmatrix}_q = \binom{n}{k}$. However, the two schemes have different qualitative behavior, for various reasons. One example is the dimensions of the eigenspaces: for the Johnson scheme they grow polynomially, while for the Grassmann scheme they grow exponentially.

Bilinear scheme The bilinear scheme $H_q(n, d)$, where q is a prime power, is the graph of all $n \times d$ matrices over $\mathbb{F}_q$. Two matrices A, B are connected if $\operatorname{rank}(A - B) = 1$. and more generally $d(A, B) = \operatorname{rank}(A - B)$. This is the q-analog of the Hamming scheme.

Relation between Grassmann and bilinear schemes It turns out that the bilinear scheme $H_q(n, d)$ can be realized as a subset of the Grassmann scheme $J_q(n + d, n)$ consisting of all subspaces having trivial intersection with a fixed subspace of degree d. This construction is called an *attenuated space*.

The recent work of Khot, Minzer and Safra uses a relation in the opposite direction: If $d \ll n$, then the rows of an $n \times d$ matrix have full rank with high probability, and so define a point of $J_q(n, d)$. This allows analyzing $J_q(n, k)$ through $H_q(n, k)$, which turns out to be simpler to handle.

Spatial degree We have described above how to define degree for a cometric association scheme. For the examples given above, this notion coincides with a spatial definition via juntas: a function has degree d if it is a linear combination of d-juntas. The definition of d-junta depends on the scheme:

- Hamming scheme: A function is a d-junta if it depends on d inputs.
- Johnson scheme: We think of the Johnson scheme as a subset of the Hamming scheme, and then use the same definition.

- Grassmann scheme: A function f is a d-junta if $f(V)$ depends on whether or not $x_1 \in V, \ldots, x_d \in V$ for some vectors $x_1, \ldots, x_d$.

(For the bilinear scheme the definition seems less natural.)

Other examples An interesting example of an association scheme which is neither metric not cometric is the *perfect matching scheme*, in which the points are perfect matchings on $2n$ points, and the notion of distance is the cycle structure of the union of two matchings.

1.2 More on the Johnson scheme

Let us think of the points of the Johnson scheme as $0,1$-vectors of length n containing exactly k many 1s. When considering functions on the scheme, we think of them as having n $0,1$ inputs $x_1, \ldots, x_n$.

Dunkl proved the following fundamental theorem concerning the Johnson scheme $J(n,k)$:

Theorem. Every function on $J(n,k)$ has a unique representation as a multilinear polynomial P of degree at most $\min(k, n-k)$ which satisfies

$$\sum_{i=1}^{n} \frac{\partial P}{\partial x_i} = 0.$$

We call such a polynomial *harmonic*.

Some examples of harmonic polynomials are $1, x_1 - x_2, (x_1 - x_2)(x_3 - x_4)$, and so on. In fact, these examples span the space of all harmonic multilinear polynomials.

This representation is not only unique, but also allows us to give a simple description of the common eigenspaces of the scheme: they consist of *homogeneous* degree d harmonic multilinear polynomials, where $0 \leq d \leq \min(k, n-k)$.

The first to consider the Johnson scheme from a Boolean function analysis perspective were O'Donnell and Wimmer, who generalized the KKL theorem to the scheme. They utilized a result of Lee and Yau, who proved hypercontractivity for the Johnson scheme when $0 \ll k/n \ll 1$ (what they actually did was calculate the log Sobolev constant for all values of k, n).

Here is a list of some results on the Johnson scheme:

- KKL (O'Donnell–Wimmer).
- Friedgut's junta theorem (Wimmer, F.).
- FKN (F.).
- Kindler–Safra (Keller–Klein).
- Explicit orthogonal basis (F.).
- Invariance principle and its consequences (F.–Kindler–Mossel–Wimmer, F.–Mossel).
- Tight bound on number of relevant variables in a bounded degree Boolean function, à la Nisan–Szegedy (F.–Ihringer).

Perhaps the most interesting result here is the invariance principle. It shows that when $0 \ll k/n \ll 1$, an $o(\sqrt{n})$-degree harmonic multilinear polynomial has similar distribution on $J(n,k)$ and on $\mu_{k/n}(\{0,1\}^n)$. In contrast to classical invariance principle, which only works when all influences are small, this invariance principle only requires the degree to be small.

The invariance principle suggests that the correct way to extend a function from a slice to the entire Boolean cube is via its unique harmonic multilinear representation.

1.3 More on the Grassmann scheme

Not much is known on Boolean function analysis on the Grassmann scheme. The analysis of the Grassmann agreement test, due to Khot, Minzer, and Safra, is a Kindler–Safra-like theorem. F. and Ihringer characterized all Boolean degree 1 functions. To the best of my knowledge, there are no other results on this scheme.

1.4 Beyond association schemes

Boolean function analysis has been considered on domains beyond association schemes. The most prominent example is the *symmetric group*.

We can define the concept of degree on the symmetric group in two different ways: spectrally and specially. Spectrally, a function on S_n has degree at most d if its Fourier transform is supported by irreps with at most d squares beyond the first line.

Spatially, a function on S_n has degree at most d if it is a linear combination of d-juntas, where a d-junta is a function that depends on d entries of the permutation matrix representation of the input. Ellis, F. and Friedgut proved FKN and Kindler–Safra theorems for S_n, and Ellis, Friedgut and Pilpel characterized all Boolean degree 1 functions.

The slice can be realized as the set of left cosets of $S_k \times S_{n-k}$ inside S_n. Alternatively, this is the permutation module $M^{n-k,k}$ corresponding to a two-rowed tableau. Every irrep occurs once in $M^{n-k,k}$ (we say that $(S_n, S_k \times S_{n-k})$ is a *Gelfand pair*), and this is behind the commutativity which makes the slice a scheme.

Another domain, which so far hasn't been studied, is the *multislice*, a generalization of the slice into a multicolored setting, in which we replace $S_k \times S_{n-k}$ with a product of more than two symmetric groups. Alternatively, the multislice consists of a c-coloring of $\{1, \ldots, n\}$ with given histogram, the case $c = 2$ corresponding to the slice. Not much is known about the multislice. Indeed, neither hypercontractivity nor a unique representation theorem are known at present.

1.5 How to study association schemes

One difficulty in studying association schemes beyond the Hamming schemes is apparent already in the Johnson scheme: it is the absence of a Fourier basis. While there does exist a canonical basis, the Gelfand–Tsetlin basis, it is not as useful as the usual Fourier basis, since it depends on an ordering of the coordinates.

Another difficulty arises in the Johnson scheme when k/n is small. This is a regime in which the classical FKN theorem doesn't hold. The classical FKN theorem states that if f is a Boolean function and $\|f^{>1}\|$ is small, then f is close to a dictator (a function depending on a single coordinate). This is no longer true when k/n is small: instead, f could depend on $O(n/k)$ coordinates. This particular difficulty is already apparent in the p-biased cube for small p, a domain which has classically been studied only in the context of sharp threshold theorems, which are the small p analogs of Friedgut's junta theorem. F. generalized the FKN theorem to this setting, and recently, Dinur, F. and Harsha generalized the Kindler–Safra theorem to this setting. All these results generalize to the slice with small k/n.

A new difficulty arises for the Grassmann scheme. The mapping $V \mapsto V^\perp$ maps $J_q(n, k)$ to the isomorphic $J_q(n, n-k)$. However, whereas the corresponding mapping $S \mapsto \overline{S}$ on the Johnson scheme is quite benign, in the Grassmann scheme it highlights a duality, which we illustrate by considering degree 1 functions. As stated above, a function has degree 1 if it is a linear combination of the functions $1_{x \in V}$. It turns out that instead of the collection $1_{x \in V}$, we can take the collection $1_{y \perp V}$ (on the Johnson scheme, this would be $1_{y \notin S}$, which is equivalent to $1_{x \in S}$).

In applications we may not know in advance which of the two bases to use, and indeed sometimes we need to use a combination of bases. This is what makes the proof of Khot–Minzer–Safra so challenging. They are able to move between the two bases using the Fourier transform on the Boolean cube.

2 Differential posets

We can think of the Boolean hypercube as the lattice of all subsets of a set $\{1, \ldots, n\}$. Denote by P_k the kth level of this lattice, which corresponds to the kth slice. The operator U maps functions on P_k to functions on P_{k+1}. It is defined by $(Uf)(S) = \sum_{T \lhd S} f(T)$. We can similarly defined an operator D mapping functions on P_k to functions on P_{k-1}. Notice that

$$(UDf)(S) = \sum_{|T \triangle S|=2} f(T) + kf(S),$$
$$(DUf)(S) = \sum_{|T \triangle S|=2} f(T) + (n-k)f(S).$$

Therefore $DU - UD = (n-2k)I$. A *differential lattice* is a lattice in which $DU - UD$ is a scalar for every starting level.

The original example was the Young lattice of all (unfilled) Young diagrams, in which the scalar doesn't depend on the level. Another example is the *q-cube*, consisting of all subspaces of an n-dimensional $\mathbb{F}_q$-vector space.

We can use the Boolean lattice to study the *slice*. Let $k \leq n/2$. Every element x of the poset at level $0 \leq \ell \leq k$ corresponds to a function on level k given by $f_x(y) = 1$ if $x \leq y$ (i.e., if $x \subseteq y$). A linear combination f of such functions, which we identify with a function on levels $0 \leq \ell \leq k$, is harmonic if $Df = 0$ (where D now maps $P_0 \cup \cdots \cup P_k$ to $P_0 \cup \cdots \cup P_{k-1}$, and maps P_0 to zero).

The unique representation of functions on the Grassmann scheme can be obtained in exactly the same way.

The unique representation theorem has the following abstract formulation.

Theorem. Let P be a (sequentially) differential posets, and let $DU - UD = r_k I$ at level k. Suppose further that the poset satisfies the unitary Peck property: $|P_k| = |P_{n-k}|$; $|P_0| < |P_1| < \cdots < |P_{\lfloor n/2 \rfloor}|$; and $r_i + \cdots + r_j \neq 0$ for $j \leq n - i$.

For $k \leq n/2$, every function f on P_k has a unique representation

$$f = \sum_{d=0}^{k} U^{k-d} f^{=d},$$

where $f^{=d}$ is a function on P_d satisfying $Df^{=d} = 0$. Moreover, the different components are orthogonal.

Here orthogonality is with respect to the uniform measure on the P_k. Let us briefly explain how orthogonality is proved. The operations U, D are adjoint, and so

$$\langle U^{k-d} f^{=d}, U^{k-e} f^{=e} \rangle = \langle f^{=d}, D^{k-d} U^{k-e} f^{=e} \rangle.$$

Suppose now that $k - d > k - e$. Using the relation $DU - UD = r_\ell I$, we can move the Ds to the right-hand side. Since $k - d > k - e$, we will get a bunch of terms having D at the right end, and so $Df^{=e} = 0$ implies that the entire inner product vanishes.

Similar ideas apply for analyzing *high-dimensional expanders*, which are high-dimensional analogs of expander graphs. The setup is more complicated since the correct measure on the slices is not uniform, and since $DU - UD$ is only close to a scalar, but some of the results can be recovered.

3 Other domains

Let us briefly mention a few other interesting domains.

Quantum Fourier analysis The Pauli matrices are four 2×2 Hermitian matrices which form a basis for the space of 2×2 Hermitian matrices over the reals. Tensorizing, this gives a Fourier basis for $2^n \times 2^n$ Hermitian matrices.

In quantum computing we are mainly interested in unitary operators. A unitary Hermitian matrix A satisfies the identity $A^2 = I$, which is analogous to the identity $f^2 = 1$ satisfied by $\{\pm 1\}$-valued functions on the Boolean cube.

Montanaro and Osborne initiated the study of the space of Hermitian matrices from the perspective of Boolean function analysis. They identified analogs of basic concepts such as the Fourier expansion (which is the Pauli expansion), degree, influences, the noise operator, and even hypercontractivity and (later) reverse hypercontractivity. They proved analogs of the FKN theorem. It would be interested to continue their study.

Continuous spaces Gaussian space is just one example of a continuous space of interest in Boolean function analysis (due to the invariance principle, relating it to the Boolean cube). The noise operator on Gaussian space is the Ornstein–Uhlenbeck semigroup, which corresponds to Brownian motion. Other spaces of potential interest include spheres and tori, the latter with the heat semigroup as a noise operator.

Raphael Bouyrie has shown that several results in Boolean function analysis can be proved in similar ways both in discrete settings and in continuous settings. One example is a generalization of Friedgut's junta theorem to bounded (rather than Boolean) functions.

Error-correcting codes Barak, Gopalan, Håstad, Meka, Raghavendra, and Steurer generalized Boolean function analysis to locally testable codes. They proved several intriguing results for the particular case of Reed–Muller codes, including an invariance principle and, as corollary, Majority is Stablest.