Analysis on the Slice

Yuval Filmus

8 March 2016

1 The Slice

Analysis of Boolean functions, by now a respectable field at the intersection of probability theory, functional analysis, combinatorics, and theoretical computer science, has focused mostly on functions on the Boolean cube $\{0,1\}^n$, or more generally on product domains (corresponding to abelian groups). Some attention has also been given to non-abelian groups, mostly the symmetric group. However, in some situations one is interested in other domains. In this talk, our domain of choice is going to be the slice:

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

(We will usually assume that $k \leq n/2$.) Why the slice? It is the natural home of the original Erdős–Ko–Rado theorem, and of $G(n,M)$ random graphs. It has been used to prove a stability version of the Kruskal–Katona theorem (work of O'Donnell and Wimmer). And it is also the simplest domain which is not a group. On the other hand, it is a very canonical *association scheme* arising from a *Gelfand pair*.

2 Functions on the Slice

When discussing functions on $\{0,1\}^n$, we usually tacitly identify a function $\{0,1\}^n \to \mathbb{R}$ with its unique multilinear representation. Since each function $\binom{[n]}{k} \to \mathbb{R}$ can certainly be extended to a function $\{0,1\}^n \to \mathbb{R}$, this shows that every function $\binom{[n]}{k} \to \mathbb{R}$ can be represented as a multilinear polynomial in the x_i, in many different ways. We thus need to add more constraints in order to get a unique representation.

A hint comes up when we consider the *invariance principle* for the slice. The invariance principle of Mossel, O'Donnell and Oleszkiewicz on $\{0,1\}^n$ states that a low-degree, low-influence function on the cube has a similar distribution if each of the x_i is replaced by a Gaussian random variable with the same mean and variance. (The invariance principle can be thought of as a generalization of the central limit theorem to low-degree polynomials.) It's not clear what the correct notion of "low degree" is for functions on the slice, so let's agree to use the usual notion of degree. We then run into a problem when considering the polynomial

$$x_1 + \cdots + x_n,$$

whose distribution on the slice (constant) is pretty different from its distribution on the cube (not constant). More generally, functions of the form

$$(x_1 + \cdots + x_n - k) \cdot P \pmod{x_i^2 = x_i}$$

are problematic, since they vanish on the slice but have a non-trivial distribution on the cube. We would like to "mod" them out. The remainders, it turns out, are exactly the polynomials Q such that

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

We call these polynomials *harmonic*. Dunkl showed that *every function on the slice has a unique representation as a harmonic multilinear polynomial of degree at most* $\min(k, n-k)$, and we proved that the invariance principle indeed holds for such functions which also have low degree and low influences.

3 Degrees on the Slice

What is the correct notion of *degree* for functions on the slice? Let us start by exploring the notion of degree for functions on the cube. For these functions, the degree is defined as the degree of the largest monomial in the multilinear expansion of the function. Put differently, we can write any function f as a sum of its homogeneous parts $f = \sum_d f^{=d}$, and the degree is the largest d such that $f^{=d} \neq 0$. The homogeneous parts arise naturally in many key formulas in analysis of Boolean functions. For example,

$$\operatorname{Inf}[f] = \sum_d d\|f^{=d}\|^2,$$
$$\langle f, T_\rho g\rangle = \sum_d \rho^d \langle f^{=d}, g^{=d}\rangle.$$

What is common for these statistics is that they can be written as $f'Af$ or $f'Ag$ for matrices A on which the (x, y) entry depends only on the Hamming distance $d(x, y)$. The space of these matrices form (the Bose–Mesner algebra of) the *Hamming association scheme*, and they are commuting and so have common eigenspaces, the dth eigenspace spanned by the degree-d monomials. This is how the homogeneous parts arise naturally from the geometry of the Boolean cube.

Something very similar happens for the slice. In this case the matrices we consider are labelled by elements of $\binom{[n]}{k}$ rather than by elements of $\{0, 1\}^n$, and the notion of distance is $d(x, y) = k - |x \cap y|$. Matrices in which the (x, y) entry depends only on $d(x, y)$ belong to the *Johnson association scheme*. Again, these matrices are commuting and so have common eigenspaces, the dth eigenspace consisting of all homogeneous degree-d harmonic multilinear polynomials.

At this point it is natural to conclude that the degree of a function on the slice should be defined as the degree of its unique harmonic multilinear representation. But theoretically speaking, this relies on the arbitrary ordering of the homogeneous parts in order of increasing degree. This ordering, however, is dictated by the following property of polynomials, which holds also for the slice: the product of a degree $\leq a$ polynomial and a degree $\leq b$ polynomial has degree $\leq a + b$. (This property is known as the *Q-polynomial* or *cometric* property.)

4 Gelfand–Tsetlin Basis for the Slice

The decomposition into homogeneous parts is already good enough for many applications, just as in the case of the cube. But sometimes we need an actual orthogonal basis. Here there is a slight problem, inherited from the representation theory of S_n: there is no canonical basis.

Indeed, the space of linear homogeneous harmonic multilinear polynomials has rank $n - 1$ whereas the number of variables is n, so we will need to make some arbitrary choices. It turns out that it's enough to fix an ordering of $[n]$ — say the standard ordering — and then we magically obtain a canonical basis, the Gelfand–Tsetlin basis! It's actually a basis for harmonic multilinear polynomials, and is essentially determined (up to scalar multiples) by the following two properties:

1. The basis for degree d homogeneous multilinear polynomials on $n + 1$ variables extends the basis for degree d homogeneous multilinear polynomials on n variables.

2. The basis functions are orthogonal *with respect to* any *exchangeable (symmetric) distribution*!

More of its amazing properties are listed below. But first, let me tell you what this basis looks like. For some d, let $A = (a_1, \ldots, a_d)$, $B = (b_1, \ldots, b_d)$ be two lists of distinct numbers in $[n]$. We say that $A < B$ if:

1. $\{a_1, \ldots, a_n\} \cap \{b_1, \ldots, b_n\} = \emptyset$.

2. $b_1 < \cdots < b_n$.

3. $a_1 < b_1, \ldots, a_n < b_n$.

Call B a *top-set* if $A < B$ for some A. Counting ballot sequences shows that the number of top-sets of size d is $\binom{n}{d} - \binom{n}{d-1}$. The basis is given by χ_B for all top-sets B, where χ_B is

$$\chi_B = \sum_{A<B} \prod_{i=1}^{|B|} (x_{a_i} - x_{b_i}).$$

For example, here is the basis for $\binom{[4]}{2}$:

$$\begin{aligned}
\chi_\emptyset &= 1, \\
\chi_{\{2\}} &= x_1 - x_2, \\
\chi_{\{3\}} &= x_1 + x_2 - 2x_3, \\
\chi_{\{4\}} &= x_1 + x_2 + x_3 - 3x_4, \\
\chi_{\{2,4\}} &= (x_1 - x_2)(x_3 - x_4), \\
\chi_{\{3,4\}} &= (x_1 - x_3)(x_2 - x_4) + (x_2 - x_3)(x_1 - x_4).
\end{aligned}$$

One very useful property of the standard Fourier expansion states that if a function depends only on the variables in B, then its Fourier expansion is supported on subsets of B. The analog in the case of the slice is somewhat less convenient: if a function depends only on $x_{m+1}, \ldots, x_n$ then its Fourier expansion is supported on subsets of $\{m + 1, \ldots, n\}$. While this doesn't catch all possible dependency structures, it is enough to prove a junta theorem such as Friedgut's junta theorem for the slice, first proved by Wimmer.

As promised, these functions are orthogonal for all exchangeable measures, but their norms of course depend on the measure. It turns out that they only depend on the norms of the special elements $\chi_d = \chi_{\{2,4,\ldots,2d\}}$:

$$\|\chi_{\{b_1,\ldots,b_d\}}\|^2 = \binom{b_1}{2}\binom{b_2 - 2}{2} \cdots \binom{b_d - 2(d-1)}{2} \|\chi_d\|^2.$$

If the measure in question is the uniform distribution on $\binom{[n]}{n/2}$, then a simple calculation shows that $\|\chi_d\|^2 \approx 2^{-d}$, which is approximately the same as the corresponding value for the uniform measure on $\{0,1\}^n$! This shows that the norm of a low-degree harmonic multilinear function on the slice is close to its norm on the cube. The invariance principle on the slice states that the same is true not only for the norm, but for distribution itself.

5 Terwilliger Algebra

Surprisingly enough, understanding the slice from a spectral perspective helps us understand the cube itself from a spectral perspective! We mentioned earlier that the concept of "degree" emerges naturally from considering matrices in which rows and columns are labelled with elements of $\{0,1\}^n$, and the contents of the (x,y) entry depends only on $d(x,y)$. While many interesting matrices, such as the Laplacian, are of this form, this is not quite the guarantee that we get in other cases.

In many cases in which a matrix is involved, the names of the coordinates are not important, and so the matrix is invariant under renaming coordinates, that is, permutations of $[n]$. In a matrix having these symmetries, the value of the (x,y) entry depends not only on $d(x,y)$ but also on $|x|$ and $|y|$ (the triples $(|x|,|y|,d(x,y))$ are the orbits of the relevant action of the symmetric group). These matrices belong to the so-called *Terwilliger algebra* of the Hamming association scheme (the earlier matrices belong to the *Bose–Mesner* algebra).

General theory shows that it is possible to "condense" these matrices into a space of much lower dimension: it is possible to map matrices in the Terwilliger algebra to a tuple of matrices $(B_0,\ldots,B_{\lfloor n/2\rfloor})$ of dimensions $n+1, n-1, \ldots$ in a linear way. Moreover, this mapping is actually a mapping of C^* algebras, so that (for example) the spectral norm of the original matrix equals the maximum of the spectral norms of the smaller matrices.

In some applications it is important to be able to construct this linear map explicity, one reason being that we want to impose certain properties of the original matrix, such as zeroes in certain places. What we are looking for is a change of basis under which each matrix in the Terwilliger algebra becomes a block diagonal matrix with $\binom{n}{d} - \binom{n}{d-1}$ copies of B_d, for each d. (This multiplicity comes from representation theory.) Amazingly, it turns out that an appropriate basis is formed by putting together the Gelfand–Tsetlin bases for all different slices! The ith row and column in matrix B_d corresponds to the dth level of the ith slice $\binom{[n]}{i}$, and $\binom{n}{d} - \binom{n}{d-1}$ is just the number of different basis vectors.

This representation has found uses in coding theory, and we are now at the process of trying to apply it to extremal combinatorics. Will let you know if it works out!