FKN theorem for balanced functions on S_n

Yuval Filmus

January 8, 2019

1 Introduction

1.1 Background

An easy argument shows that if $f\colon \{0,1\}^n \to \{0,1\}$ has degree 1 then f is a dictator, a result which we call the *dictator theorem*. The Friedgut–Kalai–Naor (FKN) theorem is a stability version of this result: if f is merely close to degree 1 (or merely close to Boolean) then f is close to a dictator. Quantitatively, if $\|f^{>1}\|^2 = \epsilon$ then f is $O(\epsilon)$-close to a dictator.

It is interesting to consider what happens in other domains. The simplest domain one can consider is the biased hypercube, $(\{0,1\}^n, \mu_p)$ for small p. The dictator theorem still holds, since the degree of a function doesn't depend on p. However, the FKN theorem must necessarily take a somewhat different form, since narrow positive clauses (disjunctions) are also close to degree 1. The correct version of the FKN theorem states that if $\|f^{>1}\|^2 = \epsilon$ (where the norm is with respect to μ_p) then either f or $1-f$ is $O(\epsilon)$-close to a positive clause of width $O(\sqrt{\epsilon}/p)$. This also implies that f is $O(\sqrt{\epsilon})$-close to a constant, so in some sense there are no "non-trivial" Boolean almost degree 1 functions for small p.

The next step is to consider the situation on the slice. The dictator theorem becomes only slightly more challenging to prove, and the FKN theorems generalize as well. In other words, from this perspective, the slice behaves exactly as the hypercube.

1.2 Symmetric group

The situation becomes more interesting when considering the symmetric group. We first have to explain what we mean by *degree*. We can represent the input permutation as a permutation matrix. Every function on S_n can then be represented as a polynomial over the n^2 entries of the permutation matrix. The degree of a function is the minimal degree of a polynomial representing it. Equivalently, a function has degree d if it is a linear combination of "strict" d-juntas, which correspond to degree d monomials (a more relaxed notion of junta will appear implicitly below, when we consider the dictator theorem). This definition is equivalent to the natural spectral definition, in which a function has degree d if it is supported on the isotypical component corresponding to Young diagrams with at most d boxes beyond the first row.

In the particular case of $d = 1$, we get that a function has degree 1 if it can be written as follows:

$$f(x) = \sum_{i,j} a_{ij} x_{ij},$$

where x_{ij} is the indicator of i mapping to j.

A crucial difference between the slice and the symmetric group is that some of the x_{ij}'s are mutually exclusive. For this reason, we get more Boolean degree 1 functions than in the preceding cases (a similar phenomenon appears on the Grassmann scheme). Ellis, Friedgut and Pilpel show that a Boolean degree 1 function takes one of the following forms:

$$\sum_{j \in J} x_{ij} \quad \text{or} \quad \sum_{i \in I} x_{ij}.$$

In other words, the answer depends either on the image of some point i or on the inverse image of some point j. The two options correspond to the anti-isomorphism $\pi \mapsto \pi^{-1}$ of S_n.

The symmetric group corresponds, in some sense, to μ_p for $p = 1/n$. For this reason, we expect the FKN theorem to exhibit behavior similar to the very biased hypercube. Indeed, in EFF1 we (Ellis, Friedgut, and myself) prove that if f is a Boolean function of density t/n close to degree 1 then f is close to a positive clause of width t. This result works for relatively small t, and relies on moment calculations. The sparseness of f translates to its being close to pure degree 1.

When f is balanced, the argument of EFF1 breaks down. A totally different argument (EFF2), which is the subject of this talk, proves a stronger result: f is close to a Boolean degree 1 function, that is, the variables appearing in the positive clause are mutually exclusive. Intuitively, the reason that this stronger property holds is that the approximation

$$x_{i_1 j_1} \vee \cdots \vee x_{i_\ell j_\ell} \approx x_{i_1 j_1} + \cdots + x_{i_\ell j_\ell}$$

is only valid when ℓ is small, whereas ℓ is linear in n for balanced f.

2 Overview of the argument

We will focus on the case in which f is exactly balanced. It will be a bit nicer to think of f as ± 1-valued, since in this case $\mathbb{E}[f] = 0$. We will show that if f is ± 1-valued, has expectation 0, and is ϵ-close to a degree 1 function (which we can take to be the projection $f^{\leq 1}$ to the space of functions of degree 1), then f is $\epsilon^{O(1)} + (1/n)^{O(1)}$-close to a "dictator", that is, a Boolean degree 1 function.

We stated above that a function on S_n has degree 1 if it can be written in the form

$$f^{\leq 1}(x) = \sum_{i,j} a_{ij} x_{ij}.$$

This expression has n^2 parameters, but the dimension of the space of degree 1 functions is only $(n-1)^2 + 1$. This implies that there are many different representations of $f^{\leq 1}$. The only canonical representation is

$$a_{ij} = (n-1)\langle f^{\leq 1}, x_{ij} \rangle = (n-1)\langle f, x_{ij} \rangle.$$

It is instructive to see what a dictatorship looks like in this representation. The matrix of coefficients a_{ij} corresponding to the function $2(x_{11} + \cdots + x_{1(n/2)}) - 1$ is

$$\begin{bmatrix} 1-\frac{1}{n} & \cdots & 1-\frac{1}{n} & -1+\frac{1}{n} & \cdots & -1+\frac{1}{n} \\ -\frac{1}{n} & \cdots & -\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & \cdots & -\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \approx \begin{bmatrix} 1 & \cdots & 1 & -1 & \cdots & -1 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Our goal is to show that the matrix must always be approximately of this form.

The formula for a_{ij} makes it apparent that the rows and columns sum to zero. Another simple property is

$$f^{\leq 1}(\pi) = \sum_i a_{i\pi(i)}.$$

In words, the values of $f^{\leq 1}$ correspond to sums of "generalized diagonals" of the matrix.

The proof has two major parts. In the first part, we show that for most permutations π, the generalized diagonal $\{a_{i\pi(i)} : i \in [n]\}$ contains one "large" entry (close to ± 1) and $n-1$ "small" entries (close to zero). In the second part, we show that the large values are aligned: they are essentially all on a single row or on a single column. The main result is then easy to deduce.

3 First part

Let's say that a coefficient a_{ij} is "large" if it is close to ± 1. Our goal in this part of the proof is to show that most generalized diagonals in $A = (a_{ij})$ contain exactly one large element. This will happen via a reduction to the FKN theorem on the (unbiased) hypercube.

For a typical permutation π, the sum $\sum_i a_{i\pi(i)}$ is close to ± 1, say to 1. We will show that typically, the following function on the hypercube is close to $\{0, 1\}$:

$$h_\pi(y) = \sum_i y_i a_{i\pi(i)}.$$

The FKN theorem on the hypercube then implies that exactly one of the elements in $\{a_{i\pi(i)} : i \in [n]\}$ is large (in this case, close to 1).

In the argument above, we are implicitly considering the experiment of first choosing a random permutation $\pi \in S_n$, and then choosing a random subset $A \subseteq [n]$ and looking at $\pi|_A$. We can also reverse the experiment, first choosing two random subsets $A, B \subseteq [n]$ of the same size, and then choosing a random permutation among those sending A to B. It is the reverse experiment which we will be able to analyze directly.

Let $T_{A,B}$ consist of those permutations sending A to B, and consider the restriction of $f^{\leq 1}$ to $T_{A,B}$, which we denote by g. We can write g as a sum of two functions g_1, g_2, where g_1 depends on $\pi|_A$ and g_2 depends on $\pi|_{\overline{A}}$. For typical A, B, these functions will satisfy the following properties:

1. g is close to ± 1 with high probability and in L_2. This is because $f^{\leq 1}$ is close to the ± 1-valued function f.

2. $\mathbb{E}[g_1]$ and $\mathbb{E}[g_2]$ are close to zero. This is because the expected value of $\mathbb{E}[g_1]$ is exactly zero, and its variance can be calculated to be roughly $\frac{1}{4n}$.

The idea now is that $g = g_1 + g_2$ is a roughly ± 1-valued function which is the sum of two functions on independent inputs. If g were exactly ± 1-valued, then this can only happen if one of g_1, g_2 is constant, and the other attains exactly two values. We can prove an approximate version of this property by considering two pairs of inputs α_1, β_1 and α_2, β_2 for g_1, g_2 (respectively), and the corresponding four values of g,

$$\begin{matrix} g_1(\alpha_1) + g_2(\alpha_2) & g_1(\alpha_1) + g_2(\beta_2) \\ g_1(\alpha_2) + g_2(\alpha_2) & g_1(\alpha_2) + g_2(\beta_2) \end{matrix}$$

A simple case analysis shows that if all of these are close to ± 1 then one of the following cases must happen:

(a) $g_1(\alpha_1) \approx g_1(\beta_1)$ and $g_2(\alpha_2) \approx g_2(\beta_2)$.

(b) $g_1(\alpha_1) \approx g_1(\beta_1)$ and $|g_2(\alpha_2) - g_2(\beta_2)| \approx 2$.

(c) $|g_1(\alpha_1) - g_1(\beta_1)| \approx 2$ and $g_2(\alpha_2) \approx g_2(\beta_2)$.

In particular, it must be that either $|g_1(\alpha_1) - g_1(\beta_1)| \approx 2$ or $|g_2(\alpha_2) - g_2(\beta_2)| \approx 2$ is unlikely. Let's assume the first event is unlikely. Fixing a typical α_1, this shows that g_1 is approximately constant. In fact, since $\mathbb{E}[g_1] \approx 0$ and g is not too wild, the function g_1 is approximately zero. Fixing a typical α_2, a similar argument shows that g_2 is concentrated around the two values ± 1.

So far we have shown that for typical A, B and typical $\pi \in T_{A,B}$, either $\sum_{i \in A} a_{i\pi(i)} \approx 0$ and $\sum_{i \in \overline{A}} a_{i\pi(i)} \approx \pm 1$ or vice versa. Switching the order of random selection, this shows that for typical π, the function h_π defined above is close to $\{0, 1\}$, and so $\{a_{i\pi(i)} : i \in [n]\}$ contains exactly one large element.

4 Second part

In the first part we have classified elements into large (close to ± 1) and small (close to zero), showing that most generalized diagonals in A consist of one large element and $n-1$ small elements. Other elements of A might be neither large nor small. In this part, we show that in fact most of the large elements lie on a single "line", either a row or a column, which consists almost exclusively of large elements.

Suppose we knew that *all* generalized diagonals of A contained exactly one large element. No two large elements can be compatible, and so any two must be either on the same row or on the same column. This can only happen if all large elements lie on on a single row or on a single column. In fact, simple arguments show that this is the case even if we only knew that a $1 - 4/n$ fraction of generalized diagonals contain exactly one large element.

Our goal in this part is to show a stability version of this property: if a $1 - \delta$ fraction of generalized diagonals are *good* (contain exactly one large element) then there is a row or a column that consists almost entirely of large elements (all but an $O(\delta)$ fraction). We will prove this by induction, using the following terminology:

- A matrix is δ-good if a $1 - \delta$ fraction of generalized diagonals are good.
- A γ-strong line (row or column) in a matrix is one which contains a $1 - \gamma$ fraction of large elements.

Let $X \subset [n]$ be a set of size $n/2$. An averaging argument shows that there exists a set Y of size $n/2$ such that a $1 - \delta$ fraction of generalized diagonals in $T_{X,Y}$ are good. A simple argument shows that this can only happen if either $A|_{X \times Y}$ is δ-good and $A|_{\overline{X} \times \overline{Y}}$ consists almost entirely of non-large elements, or vice versa. In the first case, $A|_{X \times Y}$ contains a γ-strong line by induction (for $\gamma = O(\delta)$), and in the second, $A|_{\overline{X} \times \overline{Y}}$ contains a γ-strong line.

We want to patch all these strong lines together. We start by showing that any two strong lines must be the same, that is, they must both correspond to some row i or to some column j. Indeed, otherwise, a simple inclusion-exclusion argument shows that the probability that a random generalized diagonal contains two large elements is too big. For definiteness, let this common strong line be the first row.

For every $Y \subset [n]$ of size $n/2$, the argument above shows that either the Y-coordinates or the $\overline{Y}$-coordinates of the first row contain at most $\gamma(n/2)$ non-large elements, and so the first row contains at most γn non-large elements, that is, it is γ-strong. (I'm slightly cheating here.)

Putting both parts of the proof together, we conclude that there is a line of A, say the first row, which consists almost exclusively of large entries. In order to complete the proof, let $F = \{\pi : f(\pi) = 1\}$, and notice that

$$\frac{|F \cap T_{ij}|}{|T_{ij}|} \approx \frac{1 + a_{ij}}{2}.$$

Recall that an entry a_{ij} is large if it is close to ± 1. This translates to F either containing almost all of T_{ij} or almost none of it. In particular, for most j, the set F either contains almost all of T_{1j} or almost none of it. Since $|F| = n!/2$, the first case holds for roughly half the j, and the second for the other half. We conclude that F is roughly the union of the cosets of the first type.

5 Unbalanced functions

When the function is not exactly balanced, we need to modify the argument in certain ways. Suppose that $|F|/n! = c \le 1/2$, which we think of as constant. It turns out that the correct choice of a_{ij} is now

$$a_{ij} = (n-1)\langle f, x_{ij} \rangle - \frac{n-2}{n}(2c-1).$$

In the first part of the argument, we considered a typical $T_{A,B}$, the function $g = f^{\le 1}|_{T_{A,B}}$, and its decomposition $g = g_1 + g_2$. When $c = 1/2$, one of g_1, g_2 was almost zero, and the other almost ± 1-valued. For general c, one of the functions is close to $c - 1/2$, and the other is close to being $1/2 - c \pm 1$-valued. As before, this implies that a generalized diagonal typically consists of one large entry and $n-1$ small entries, where this time an entry is large if it is close to $\pm 2c$ or to $\pm 2(1-c)$.

The second part shows that there is a line, say the first row, consisting mostly of large entries. When $c = 1/2$, this implied that F contains either almost all of T_{1j} or almost none of it, for every j. For general c, there is an additional option that we have to rule out, namely that F contains a $2c$ fraction of T_{ij}. An exchange argument shows that any two large entries on the first row are either both good (correspond to containing all or none of T_{1j}) or both bad. Hence either all large entries on the first row are good, or all are bad. The latter case cannot happen, however, since in that case we would have $|F| \approx 2cn!$. Hence F is the roughly the union of cn cosets T_{1j}.

The argument also works for sub-constant c, as long as $\epsilon \le c^{O(1)}$ (for a specific power!). When ϵ is too large compared to c, what breaks down is the FKN theorem (applied to h_π), which no longer holds. Indeed, in general the correct structure is genuinely different: a union of not-necessarily-disjoint cosets T_{ij}.