Lecture 5

Lecture date: Sep 13 *Scribe: Allan Sly*

This lecture focuses on bounds of influence sums: $\sum_{i=1}^n I_i(f)$. But first a quick exercise.

Exercise 1 ($1\frac{1}{2}$ points) *1. Let $n = 3^k$ and let $f : \{-1,1\}_0^n \to \{-1,1\}$ be $f =$ Rec-Maj$_3^k$ the recursive majority function defined recursively by $\mathrm{Rec}-\mathrm{Maj}_3^k = \mathrm{Maj}(\mathrm{Rec}-\mathrm{Maj}_3^{k-1}, \mathrm{Rec}-\mathrm{Maj}_3^{k-1}, \mathrm{Rec}-\mathrm{Maj}_3^{k-1})$. Calculate $I_i(f)$.*

2. Let $n = r^k$ with r odd and let $f : \{-1,1\}_0^n \to \{-1,1\}$ be $f =$ Rec-Maj$_r^k$. Calculate asymptotics for $I_i(f)$ when r is large (and k is very large).

To begin with think about the example $f : \{-1,1\}_0^n \to \{-1,1\}$. Then $\sum_{i=1}^n I_i(f)$ is equal to the normalized edge boundary between the sets $\{x : f(x) = 1\}$ and $\{x : f(x) = -1\}$. An edge is a pair of points which differ on exactly one coordinate and the normalized edge boundary is the number of edges with points in different sets divided by 2^{n-1}. This follows from the fact that $I_i(f) = P[f(x) \neq f(x^{\oplus i})]$.

Lemma 2

$$\sum_{i=1}^n I_i(f) = \sum_S |S||f_S|^2 = \sum_J |J||\widehat{f}(J)|^2$$

Proof: By a lemma proved in the previous lecture

$$\sum_{i=1}^n I_i(f) = \sum_{i=1}^n \sum_{S:i\in S} |f_S|^2 = \sum_S |S||f_S|^2$$

and

$$\sum_{i=1}^n I_i(f) = \sum_{i=1}^n \sum_{J:J_i\neq 0} |\widehat{f}(J)|^2 = \sum_J |J||\widehat{f}(J)|^2.$$

□

Corollary 3 *Suppose that $\sum_{i=1}^n I_i(f) < a$ and $a < b$. Then*

$$\sum_{|S|>b} |f_S|^2 < \frac{a}{b}.$$

This corollary should be interpreted as saying that functions with low influence are well approximated by low coordinates. We can also make the trivial observation that if f depends on k variables then $\sum_{i=1}^n I_i(f) \leq k\mathbf{Var}[f]$. For any function we always have that $\sum_{i=1}^n I_i(f) \leq n\mathbf{Var}[f]$. This bound is achieved as $\sum_{i=1}^n I_i(\prod_{j=1}^n x_j) = n$. In the case of monotone functions $f : \{-1,1\}_0^n \to \{-1,1\}$ we can do better.

Lemma 4 *The monotone functions $f : \{-1,1\}_\theta^n \to \{-1,1\}$ which maximize $\sum_{i=1}^n I_i(f)$ are given by $f(x) = \mathrm{sgn}(\sum_{i=1}^n (x_i - \theta))$. The choice of ± 1 when $\sum_{i=1}^n (x_i - \theta) = 0$ does not affect $\sum_{i=1}^n I_i(f)$.*

Proof: For convenience of notation we will prove the result instead for $f : \{0,1\}_p^n \to \{0,1\}$. If $f : \{0,1\}_p^1 \to \{0,1\}$ is monotone then either

1. $f \equiv c$ and $\mathbf{Var}[f] = 0$,
2. $f \not\equiv c$ and $\mathbf{Var}[f] = p - p^2$.

Applying this to the definition of influence we get

$$\begin{aligned}
\sum_{i=1}^n I_i(f) &= (p-p^2)\sum_x p^{|x|-1}(1-p)^{n-|x|} \sum_{\substack{y \text{ below } x \\ |x-y|=1}} (f(x) - f(y)) \\
&= (p-p^2)\sum_{k=0}^n \sum_{x:|x|=k} p^{k-1}(1-p)^{n-k-1} f(x)(k(1-p) - (n-k)p) \\
&= (p-p^2)\sum_{k=0}^n \sum_{x:|x|=k} p^{k-1}(1-p)^{n-k-1} f(x)(k - np)
\end{aligned}$$

where y below x means that for all i, $y_i \leq x_i$ and the second and third equalities follow by rearranging. In this form it is easy to see that the sum of the influences is maximized by setting $f(x)$ to be 1 when $|x| - np > 0$ and $f(x)$ to be 0 when $|x| - np < 0$. Translating back to a range of $\{-1,1\}$ it follows that it is maximized by $f(x) = \mathrm{sgn}(\sum_{i=1}^n (x_i - \theta))$. $\Box$

Corollary 5 *For monotone functions $f : \{-1,1\}_\theta^n \to \{-1,1\}$,*

$$\sum_{i=1}^n I_i(f) \leq \sum_{i=1}^n I_i(\mathrm{Maj}) \leq \sqrt{\frac{2n}{\pi}}(1 + o(1)).$$

Exercise 6 (5 points) *The monotone functions $f : \{-1,1\}_{\theta_1} \times \{-1,1\}_{\theta_2} \times \ldots \times \{-1,1\}_{\theta_n} \to \{-1,1\}$ which maximize $\sum_{i=1}^n I_i(f)$ are given by $f(x) = \mathrm{sgn}(\sum_{i=1}^n (x_i - \theta_i))$.*

Definition 7 *(Lexicographical Order) For $x, y \in \{-1,1\}^n$ we say that $x <_L y$ in lexicographical order if $x_1 = y_1, \ldots, x_{i-1} = y_{i-1}$ and $x_i < y_i$. We denote $L(M)$ to be the first M vectors in lexicographical order in $\{-1,1\}^n$.*

For $A \subseteq \{-1,1\}^n$ we denote $\psi(A) = |\{e = (x,y) : x \in A, y \notin A\}|$ to be the size of the edge boundary of of A.

Theorem 8 *(Harper (1961)) For any set $A \subseteq \{-1,1\}^n$ of size m,*

$$\psi(A) \geq \psi(L(m)).$$

The set $L(m)$ is the unique minimum up to trivial transformations, permuting the coordinates $(x \mapsto (x_{\sigma(1)}, \ldots, x_{\sigma(n)}))$ and flipping of bits $(x \mapsto (x_1 y_1, \ldots, x_n y_n)), y \in \{-1,1\}^n$.

Proof: Let $A_{\pm 1}(i) = \{x \in A : x_i = \pm 1\}$ and let $n_{\min(i)} = \min\{|A_{-1}(i)|, |A_1(i)|\}$ and $n_{\max(i)} = \max\{|A_{-1}(i)|, |A_1(i)|\}$. Let C_i be the "compression operator" which maps

$$A = \begin{pmatrix} A_{-1} \\ A_{+1} \end{pmatrix} \mapsto \begin{pmatrix} \text{first } n_{\max}(i) \text{ elements with } x_i = -1 \text{ in lexicographical order,} \\ \text{first } n_{\min}(i) \text{ elements with } x_i = 1 \text{ in lexicographical order.} \end{pmatrix}$$

For any set A the compression operator reduces the number of edges, $\psi(C_i(A)) \leq \psi(A)$. This can be seen by looking at two types of edges separately. The number of edges in the i coordinate does not increase as the compression operator matches elements in $(C_i(A))_{-1}(i)$ and $(C_i(A))_{+1}(i)$ as much as possible. And by induction on n, $(C_i(A))_{-1}(i)$ and $(C_i(A))_{+1}(i)$ do not increase the number of boundary edges in the other coordinates.

In the next lecture we will see that repeated application of C_i over all i will stabilize to some set which, except in a special case, will be $L(m)$ which will complete the result. $\square$