Computer Science 294 Lecture 5 Notes

Daniel Raban

January 31, 2023

1 Isoperimetric Inequality, Noise Stability, and Arrow's Theorem

1.1 Recap: Influence of voters and total influence

For a boolean function $f : \{\pm 1\}^n \to \{\pm 1\}^n$ and $i \in [n]$, we defined the **influence** of voter i to be

$$\mathrm{Inf}_i(f) = \mathbb{P}_{X \sim \{\pm 1\}^n}(f(X) \neq f(X^{\oplus i})).$$

More generally, for $f : \{\pm 1\}^n \to \mathbb{R}$, we have

$$\mathrm{Inf}_i(f) = \mathbb{E}_X \left[\left(\frac{f(X^{i \mapsto 1}) - f(X^{i \mapsto -1})}{2} \right)^2 \right]$$

We saw that $\mathrm{Inf}_i(f) = \sum_{S \ni i} \widehat{f}(X)^2$ and multiple interpretations of the total influence:

$$\mathbb{I}(f) = \sum_{i=1}^n \mathrm{Inf}_i(f) = \sum_{i=1}^n \sum_{S \ni i} \widehat{f}(S)^2 = \sum_{S \subseteq [n]} \widehat{f}(S)^2 \cdot |S|,$$

$$\mathbb{I}(f) = \frac{\#\text{ of sensitive edges}}{2^{n-1}}.$$

Exercise 1.1. If $f : \{\pm 1\}^n \to \{\pm 1\}$ and for all sets S of size > 1 $\widehat{f}(S) = 0$, then f is either constant, dictator, or anti-dictator.

1.2 Poincare inequality and isoperimetric inequalities

Theorem 1.1 (Poincaré Inequality). *For any* $f : \{\pm 1\}^n \to \mathbb{R}$,

$$\mathbb{I}(f) \geq \mathrm{Var}(f).$$

Proof.

$$
\begin{aligned}
\mathbb{I}(f) &= \sum_{S\subseteq[n]} \widehat{f}(S)^2|S| \\
&= \sum \varnothing \neq S \subseteq [n]\widehat{f}(S)^2 \cdot |S| \\
&\geq \sum_{\varnothing\neq S\subseteq[n]} \widehat{f}(S)^2 \cdot 1 \\
&= \mathrm{Var}(f).
\end{aligned}
$$

□

Remark 1.1. Equality occurs if and only if for all sets of size > 1, $\widehat{f}(S) = 0$. So by the exercise, this is exactly when f is a constant, dictator, or anti-dictator function.

Here is an interpretation of this as an *isoperimetric inequality*. Let $A = \{x \in \{\pm 1\}^n : f(x) = 1\}$ and $B = \{x \in \{\pm 1\}^n : f(x) = -1\}$. The set of sensitive edges is the cut

$$E(A, \overline{A}) = \{(u, v) \in E : u \in A, v \in \overline{A}\}.$$

An isoperimetric inequality relates the size of A to the size of the cut $E(A, \overline{A})$.

Corollary 1.1 (isoperimetric inequality)**.** *If* $|A| = \alpha \cdot 2^n$*, then*

$$|E(A, \overline{A})| \geq 2\alpha(1-\alpha)2^n.$$

Proof. Let

$$f_A(x) = \begin{cases} -1 & x \in A \\ +1 & x \notin A. \end{cases}$$

Then

$$
\begin{aligned}
\mathrm{Var}(f_A) &= \mathbb{E}[f_A^2] - (\mathbb{E}[f_A])^2 \\
&= 1 - (\alpha \cdot (-1) + (1-\alpha) \cdot 1)^2 \\
&= 4\alpha(1-\alpha).
\end{aligned}
$$

On the other hand,

$$\mathbb{I}(f) = \frac{|E(A, \overline{A})|}{2^{n-1}},$$

so we have

$$|E(A, \overline{A})| \geq 2^{n-1} \cdot 4\alpha(1-\alpha).$$

□

Here is a much sharper (for $\alpha \ll 1/2$) isoperimetric inequality that we will not prove.

Theorem 1.2 (Harper)**.** *If* $|A| = \alpha \cdot 2^n$*. Then*

$$|E(A, \overline{A})| \geq \alpha \cdot \log_2(1/\alpha) \cdot 2^n.$$

1.3 Noise stability and sensitivity

We will apply noise independently to all coordinates according to a parameter that we will vary.

Definition 1.1. Let $\rho \in [-1,1]$. For a fixed $x \in \{\pm 1\}^n$, we write $Y \sim N_p(x)$ to be the random variables such that for each $i \in [n]$, independently,

$$Y_i = \begin{cases} x_i & \text{with probability } \frac{1+\rho}{2} \\ -x_i & \text{with probability } \frac{1-\rho}{2}. \end{cases}$$

When $\rho = 1$, there is no noise. When $\rho = -1$, we flip every bit. And when $\rho = 0$, we ignore the input altogether.

Definition 1.2. The random variables (X, Y) are a **ρ-correlated pair** if we can sample them as follows:

1. Pick $X \sim \{\pm 1\}^n$ uniformly at random.
2. Pick $Y \sim N_\rho(X)$.

These have correlation ρ because for each x_i,

$$\mathbb{E}[x_i Y_i] = \left(\frac{1+\rho}{2}\right) - \left(\frac{1-\rho}{2}\right) = \rho.$$

Remark 1.2. Alternatively, for any i independently, sample $(X_i, Y_i) \in \{\pm 1\}^2$ with

$$\mathbb{E}[X_i] = 0, \qquad \mathbb{E}[Y_i] = 0, \qquad \mathbb{E}[X_i Y_i] = \rho.$$

This shows that this definition is symmetric in X and Y.

Definition 1.3. The **noise stability** with parameter ρ is

$$\mathrm{Stab}_\rho(f) = \mathbb{E}_{(X,Y)\ \rho\ \text{coord.}}[f(X)f(Y)].$$

If f is boolean,

$$\begin{aligned} \mathrm{Stab}_\rho(f) &= \mathbb{P}(f(X) = f(Y)) - \mathbb{P}(f(X) \neq f(Y)) \\ &= 2\mathbb{P}(f(X) = f(Y)) - 1 \\ &= 1 - 2\mathbb{P}(f(X) \neq f(Y)). \end{aligned}$$

Example 1.1. The noise stability of a dictator function is

$$\begin{aligned} \mathrm{Stab}_\rho(\chi_i) &= \mathbb{E}_{(X,Y)\ \rho\ \text{coord.}}[\chi(X)\chi(Y)] \\ &= \mathbb{E}[X_i Y_i] \\ &= \rho. \end{aligned}$$

Example 1.2. The noise stability of the parity function is

$$\mathrm{Stab}_\rho(\mathrm{Parity}_n) = \mathbb{E}_{(X,Y)\ \rho\ \text{coord.}}\left[\prod_{i=1}^n X_i \cdot \prod_{i=1}^n Y_i\right]$$

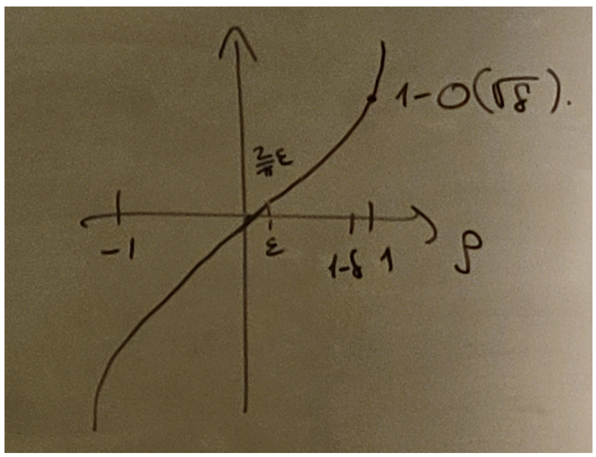
Since (X_i, Y_i) is independent of $\{(X_j, Y_j)\}_{j\neq i}$,

$$= \prod_{i=1}^n \mathbb{E}_{(X,Y)\ \rho\ \text{coord.}}[X_i Y_i]$$
$$= \rho^n.$$

Here is a theorem we will prove later.

Theorem 1.3.

$$\lim_{n\to\infty} \mathrm{Stab}_\rho(\mathrm{MAJ}_n) = \frac{2}{\pi}\arcsin(\rho).$$



It may be easier to think of this in terms of noise sensitivity.

Definition 1.4. For $f : \{\pm 1\}^n \to \{\pm 1\}$ and $\delta \in [0,1]$, pick $X \sim \{\pm 1\}^n$, and pick Y by flippiing each bit of X with probability δ, independently. Then the **noise sensitivity** of f is

$$\mathrm{NS}_\delta(f) = \mathbb{P}(f(X) \neq f(Y)).$$

This is just a reparametrization of noise stability, and such a pair (X,Y) is $1-2\delta$ correlated:

Proposition 1.1.

$$\mathrm{NS}_\delta(f) = \frac{1}{2} - \frac{1}{2}\mathrm{Stab}_{1-2\delta}(f).$$

Rephrasing the theorem, we get

Theorem 1.4.

$$\mathrm{NS}_\delta(\mathrm{MAJ}_n) = O(\sqrt{\delta}).$$

Remark 1.3. This is also true for any linear threshold function. We will prove this later.

1.4 The noise operator and Fourier representation for stability

Definition 1.5. The **noise operator** T_ρ with parameter $\rho \in [-1, 1]$ maps $f : \{\pm 1\}^n \to \mathbb{R}$ to $T_\rho f : \{\pm 1\}^n \to \mathbb{R}$ by

$$T_\rho f(x) = \mathbb{E}_{Y \sim N_\rho(x)}[f(Y)].$$

This is a way to "smooth out" a boolean function. Let's find a Fourier representation for $T_\rho f$:

Proposition 1.2.

$$T_\rho f = \sum_{S \subseteq [n]} \widehat{f}(S) \cdot \rho^{|}S| \cdot \chi_S.$$

Compare this to

$$f = \sum_{S \subseteq [n]} \widehat{f}(S) \cdot \chi_S.$$

Proof. Observe that T_ρ is a linear operator, so it suffices to prove this claim for $f = \chi_S$. Indeed,

$$\begin{aligned}
T_\rho \chi_S(x) &= \mathbb{R}_{Y \sim N_\rho(x)}[\chi_S(Y)] \\
&= \mathbb{E}_{Y \sim N_\rho(x)}\left[\prod_{i \in S} Y_i\right] \\
&= \prod_{i \in S} \mathbb{E}[Y_i] \\
&= \prod_{i \in S} \underbrace{\mathbb{E}[Y_i x_i]}_{\rho} x_i \\
&= \rho^{|S|} \chi_S(x).
\end{aligned}$$

□

This gives us a Fourier representation for stability.

Corollary 1.2.

$$\mathrm{Stab}_\rho(f) = \sum_{S \subseteq [n]} \widehat{f}(S)^2 \rho^{|S|}.$$

Proof.

$$\begin{aligned}\mathrm{Stab}_\rho(f) &= \mathbb{E}_{(X,Y)\ \rho\ \text{coord.}}[f(X)f(Y)] \\ &= \mathbb{E}_{X\sim\{\pm1\}^n}[f(X)\,\mathbb{E}_{Y\sim N_\rho(X)}[f(Y)]] \\ &= \mathbb{E}_{X\sim\{\pm1\}^n}[f(X)\cdot T_\rho f(X)]\end{aligned}$$

Using Plancherel's identity,

$$\begin{aligned}&= \sum_{S\subseteq[n]} \widehat{f}(S)\cdot \widehat{T_\rho f}(S) \\ &= \sum_{S\subseteq[n]} \widehat{f}(S)^2\rho^{|S|}.\end{aligned}$$

□

Remark 1.4. Which functions maximize $\mathrm{Stab}_\rho(f)$? This are the constant functions. But if we want to maximize Stability under "balanced" functions, we look to dictator functions: Assuming $\widehat{f}(\varnothing) = 0$ and $\rho \geq 0$,

$$\begin{aligned}\mathrm{Stab}_\rho(f) &= \sum_{\varnothing\neq S\subseteq[n]} \widehat{f}(S)^2\rho^{|S|} \\ &\leq \sum_{\varnothing\neq S\subseteq[n]} \widehat{f}(S)^2\cdot\rho \\ &= \rho \underbrace{\sum_{\varnothing\neq S\subseteq[n]} \widehat{f}(S)^2}_{=1} \\ &= \rho.\end{aligned}$$

Equality is only when the Fourier coefficients are 0 for sets of size > 1, so f must be a dictator function.

1.5 Arrow's theorem

Suppose we want to pick between three alternatives, so everyone will have a preference list. We want the following properties:

1. **Independence of irrelevant alternatives (IIA)**: Changing our opinion about c doesn't affect our ranking of a vs b.
2. **Unanimous**: If everyone prefers $a > b$, then society prefers $a > b$.
3. **Rationality/no cycles**: No $a > b, b > c, c > a$.

From property 1, we can model this as $\binom{3}{2}$ pairwise elections.

Example 1.3.

	Voter 1 $a > b > c$	Voter 2 $b > c > a$	Voter 3 $c > a > b$	Society
a (+1) vs b (-1)	$+1$	-1	$+1$	$f(x)$
b (+1) vs c (-1)	$+1$	$+1$	-1	$g(y)$
c (+1) vs a (-1)	-1	$+1$	$+1$	$h(z)$

We can see already that majority rule can't satisfy these three properties. Also note that each column cannot have all $+1$ or all -1

Theorem 1.5 (Arrow)**.** *The only rule that satisfies 1, 2, and 3 is when* $f = g = h = \chi_i$*.*

Arrow did a reduction from the 3 function version to the 1 function version. We will just treat this case. Kalai proved a robust version of Arrow's theorem.

Theorem 1.6 (Kalai)**.**

$$\mathbb{P}_{X,Y,Z}(f(X), f(Y), f(Z) \text{ is rational}) = \frac{3}{4} - \frac{3}{4}\mathrm{Stab}_{-1/3}(f).$$

We will prove Kalai's theorem next time. Here is how this implies Arrow's theorem:

Corollary 1.3. *If* $\mathbb{P}_{X,Y,Z}(f(X), f(Y), f(Z) \text{ is rational}) = 1$*, then* $f = \chi_i$ *or* $f = -\chi_i$*.*

Proof. Kalai's theorem implies that

$$\mathrm{Stab}_{-1/3}(f) = -1/3.$$

Now, in general,

$$-\mathrm{Stab}_{-1/3}(f) = -\sum_S \widehat{f}(S)^2\,(-1/3)^{|S|} \leq \sum_{|S| \text{ odd}} \widehat{f}(S)^2 \cdot \left(\frac{1}{3}\right)^{|S|} \leq \frac{1}{3},$$

with equality if and only if $\sum_{|S|=1} \widehat{f}(S)^2 = 1$. So f must be a dictator function. □