Analysis of Boolean functions

We consider functions of $\overset{\text{iid}}{\text{BERNOULLI}}\left(\frac{1}{2}\right)$ random variables
$f(X_1, \dots, X_n)$

By CLT, we will also derive results about functions of iid $N(0,1)$ variables.

It is more convenient to use $\{-1, 1\}$ instead of $\{0, 1\}$.

So we look at functions $f : \{-1, 1\}^n \to \{-1, 1\}$ (Boolean function)
or $\{-1, 1\}^n \to \mathbb{R}$.

- Fourier-Walsh expansion

Any BOOLEAN function is a multilinear polynomial.

Ex $\max_2 : \{-1, 1\}^2 \to \{-1, 1\}$ $\quad \max_2(x_1, x_2) = \frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1x_2$

$\mathrm{maj}_3 : \{-1, 1\}^3 \to \{-1, 1\}$ $\quad \mathrm{maj}_3(x_1, x_2, x_3) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$

Notation For $S \subset [n]$ define for $x = (x_i) \in \{-1, 1\}^n$

$$x^S = \prod_{i \in S} x_i \qquad \text{(WALSH functions)} \quad w_S(x) = x^S$$

Theorem Any $f : \{-1, 1\}^n \to \mathbb{R}$ can be uniquely expressed as a multilinear polynomial

$$f(x) = \sum_{S \subset [n]} \hat{f}(S)\, x^S$$

$$f = \sum_{S \subset [n]} \hat{f}(S)\, w_S$$

$\hat{f}(S)$: WALSH-FOURIER coefficients of f.

Proof $f(x) = \sum_{y \in \{-1, 1\}^n} f(y)\, \mathbb{1}_{\{y\}}(x)$

and $\mathbb{1}_{\{y\}}(x) = \left(\frac{1 + x_1 y_1}{2}\right) \cdots \left(\frac{1 + y_n x_n}{2}\right)$

so the result follows; uniqueness by dimension counting

Consider the space $L^2(\{-1,1\}^n, \mathbb{P})$

uniform probability measure

<u>theorem</u> $(w_S)_{S\subset[n]}$ is an orthonormal basis

<u>Proof</u> $\mathbb{E}[w_S w_T] = \mathbb{E} \prod_{i\in S} x_i \prod_{j\in T} x_j = \begin{cases} 1 & \text{if } S=T \\ 0 & \text{by independance if } S\neq T. \end{cases}$

(x_i) iid
unbiased random bits

OR : $w_S w_T = w_{S\Delta T}$ and $\mathbb{E} w_S = \begin{cases} 0 & S\neq\emptyset \\ 1 & S=\emptyset \end{cases}$

<u>Proposition</u> $\forall f : \{-1,1\}^n \to \mathbb{R}$

$$\hat{f}(S) = \mathbb{E}[f w_S]$$

<u>PARSEVAL formula</u> $\mathbb{E}[f^2] = \sum_{S\subset[n]} \hat{f}(S)^2$

we have $\hat{f}(\emptyset) = \mathbb{E} f$. therefore:

$$\mathrm{Var}(f) = \sum_{\substack{S\subset[n] \\ S\neq\emptyset}} \hat{f}(S)^2$$

<u>Social choice</u> BOOLEAN functions $f : \{-1,1\}^n \to \{-1,1\}$

can be interpreted as social choice functions, such that

majority (well defined if n ~~are~~ is odd)
dictatorship, antidictatorship
iterated majority

The i^{th} inpluence of a ~~fn~~ BOOLEAN function $f : \{-1,1\}^n \to \{-1,1\}$

is $\mathrm{Inf}_i(f) = \mathbb{P}\left(f(x) \neq f(x^{\oplus i})\right)$

where $x^{\oplus i} = (x_1, \dots, x_{i-1}, -x_i, x_{i+1}, \dots x_n)$

We also have

$$\mathrm{Inf}_i = \mathop{\mathbb{E}}_{x \in \{-1,1\}^n} D_i f(x)^2$$

where $D_i f(x) = \dfrac{f(x^{(i \to 1)}) - f(x^{(i \to -1)})}{2}$

and this definition makes sense for any function $\{-1,1\}^n \to \mathbb{R}$, not necessarily BOOLEAN.

The operation D_i is a discrete derivation operator

If $f = \sum_{S \subset [n]} \hat{f}(S) w_S$ then $D_i f = \sum_{\substack{S \subset [n] \\ i \in S}} \hat{f}(S) w_{S \setminus \{i\}}$

Indeed $D_i w_S(x) = \begin{cases} 0 & \text{if } i \notin S \\ w_{S \setminus \{i\}} & \text{if } i \in S \end{cases}$

PARSEVAL's identity implies that

$$\mathrm{Inf}_i f = \|D_i f\|_{L^2}^2 = \sum_{S \ni i} \hat{f}(S)^2$$

~~and Inf f~~

the total influence of a function is

$$\mathrm{Inf}\, f = \sum_i \mathrm{Inf}_i(f)$$

and we have therefore $\mathrm{Inf}(f) = \sum_i \sum_{S \ni i} \hat{f}(S)^2$

so $\mathrm{Inf}(f) = \sum_S |S|\, \hat{f}(S)^2$.

In particular

Poincaré inequality $\quad \mathrm{Var}(f) \leq \mathrm{Inf}(f)$

Corollary: ~~this is~~ For every function $f : \{-1,1\}^n \to \{-1,1\}$ then $\exists i$ s.t. $\mathrm{Inf}_i f \geq \frac{1}{n} \mathrm{Var}(f)$

(If $\mathbb{E} f = 0$ then $\mathrm{Var}(f) = 1$ since $\mathbb{E} f^2 = 1$).

Example majority $\quad \mathrm{Inf}_i f = O\left(\frac{1}{\sqrt{n}}\right) \quad \forall i$

i_0 dictatorship $\quad \mathrm{Inf}_{i_0} f = 1 \quad \mathrm{Inf}_i f = 0$ if $i \neq i_0$.

More sophisticated example: "tribes"

split $[n]$ into p groups (tribes) of size k, and

let $f = \begin{cases} 1 & \text{if unanimity for } 1 \text{ ain at least one group} \\ -1 & \text{otherwise} \end{cases}$

$$\mathbb{P}(f = -1) = \left(1 - \frac{1}{2^k}\right)^p$$

If we set k and p such that $\left(1 - \frac{1}{2^k}\right)^p \approx \frac{1}{2}$

More sophisticated example : <u>tribes</u>

We describe a function $f : \{-1,1\}^n \to \{-1,1\}$ as a voting system

Assume $n = k2^k$ and split $[n]$ into 2^k groups (tribes) of size k.

Define f by
- $f = 1$ if there is a tribe ~~with~~ with unanimity for 1
- $f = -1$ otherwise

$$\mathbb{P}(f = -1) = \left(1 - \frac{1}{2^k}\right)^{2^k} \approx \frac{1}{e} \qquad \text{so} \quad \mathrm{Var}(f) \approx 1 - \left(1 - \frac{2}{e}\right)^2$$

$$\mathbb{E} f = \left(1 - \frac{1}{e}\right) - \frac{1}{e} = 1 - \frac{2}{e}$$

$\mathrm{Inf}_i f$ does not depend on i and equals

$$\mathrm{Inf}_i f = \left(1 - \frac{1}{2^k}\right)^{2^{k-1}} \cdot \frac{1}{2^{k-1}} \sim \frac{c}{2^k} \sim \frac{c \log n}{n}$$

<u>This $\log n$ is sharp</u>

<u>Theorem</u> (Kahn-Kalai-Linial)

$$\boxed{\begin{array}{l} \forall f : \{-1,1\}^n \to \{-1,1\} \\ \exists i \quad \text{s.t.} \quad \mathrm{Inf}_i f \geq \dfrac{\log n}{n} \mathrm{Var}(f) \end{array}}$$

discrete analogue of 0/1 law.

We will prove this next time using hypercontractivity

The Poincaré inequality can be seen as an isoperimetric inequality on the discrete cube $\{-1,1\}^n$ equipped with HAMMING distance $d(x,y) = \#\{i \; x_i \neq y_i\}$.

Let $A \subset \{-1,1\}^n$ and ~~$f = 1_A$. then~~ $f = \begin{cases} 1 & \text{on } A \\ -1 & \text{on } \bar{A} \end{cases}$ $\qquad \alpha = \dfrac{|A|}{2^n}$

$\mathbb{E} f = \alpha - (1-\alpha) = 2\alpha - 1$ and $\mathrm{Var} f = 1 - (1-2\alpha)^2 = 4\alpha(1-\alpha)$

$\mathrm{Inf} f = \#$ edges adjacent to A

Noise operator

~~Fix~~ Fix $\rho \in [0,1]$

For $x \in \{-1,1\}^n$, we write $y \sim N_\rho(x)$ to denote a random string chosen as follows: $\forall i \in [n]$, independently,

$$y_i = \begin{cases} x_i & \text{proba } \rho \\ \text{uniform at random} & \text{proba } 1-\rho \end{cases} \qquad \mathbb{E}[y_i] = \rho x_i$$

Equivalently
$$y_i = \begin{cases} x_i & \text{proba } \frac{1}{2}+\frac{\rho}{2} \\ -x_i & \text{proba } \frac{1}{2}-\frac{\rho}{2} \end{cases}$$
and this makes sense for $\rho \in [-1,1]$.

We say that ~~y~~ ~~as~~ is ρ-correlated to x

We say that (x,y) is a ρ-correlated pair if

x is uniform on $\{-1,1\}^n$

y is ρ-correlated to x $\quad \Leftrightarrow \quad (y,x)$ ρ-correlated

($\Leftrightarrow$ $\mathbb{E}[x_i] = \mathbb{E}[y_i] = 0$ $\mathbb{E}[x_i y_i] = \rho$ + independence between indices)

The Noise stability of f is

$$\mathrm{Stab}_\rho[f] = \mathbb{E}_{\substack{(x,y) \\ \rho\text{-correlated}}} [f(x) f(y)].$$

Theorem For $f : \{-1,1\}^n \to \mathbb{R}$

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

Proof $\mathrm{Stab}_\rho[f] = \mathbb{E}[f \, T_\rho f]$ with $T_\rho f(x) = \mathbb{E}_{y \sim N_\rho(x)} f(y)$

$T_\rho w_S = \rho^{|S|} w_S$ since $T_\rho w_S(x) = \mathbb{E}_{y \sim N_\rho(x)} \prod_{i \in S} y_i$
$$= \rho^{|S|} \prod_{i \in S} x_i.$$

T_ε is the noise operator

$$T_\varepsilon f = \mathbb{E}_{y \sim N_\varepsilon(x)} f(y).$$

Digression ARROW's theorem

Consider non-binary social choice

E.g. an election with candidates A, B, C

Each voter has its preference list

$A > B > C$
$A > C > B$
$B > A > C$
$B > C > A$
$C > A > B$
$C > B > A$

Consider $f : \{-1,1\}^n \to \{-1,1\}$

~~A~~ Candidate A is a CONDORCET winner for f if

$$f\left(\left(\mathbb{1}_{A>B} \text{ for } i\text{-th voter}\right)_i\right) = 1$$
$$f\left(\left(\mathbb{1}_{A>C} \text{ ——— }\right)\right) = 1$$

CONDORCET paradox : $f = \mathrm{Maj}$ may not have a CONDORCET winner

Theorem : If $f : \{-1,1\}^n \to \{-1,1\}$ is a social choice function which always produces a CONDORCET winner, then f = dictatorship or f = antidictatorship

Proof Suppose voters randomly choose their preference list

Let $X = (x_i)$ $\quad x_i = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ voter says } A > B \\ -1 & \text{———} \quad A < B \end{cases}$

$Y = (y_i)$ $\quad y_i = \begin{cases} 1 & \text{———} \quad B > C \\ -1 & \text{———} \quad B < C \end{cases}$

$Z = (z_i)$ $\quad z_i = \begin{cases} 1 & \text{———} \quad C > A \\ -1 & \text{———} \quad C < A \end{cases}$

~~Each~~ X, Y, Z are iid on $\{-1,1\}^n$, but not independent

$$\mathbb{E}[x_i y_i] = \frac{2-4}{6} = -\frac{1}{3}$$

So x and y are ρ-correlated for $\rho = -\frac{1}{3}$.

Consider the function $NAE : \{-1,1\}^3 \to \{-1,1\}$ "not all equal"

$$NAE(w_1,w_2,w_3) = \begin{cases} -1 & \text{if } w_1 = w_2 = w_3 \\ 1 & \text{otherwise} \end{cases}$$

$$NAE(w_1,w_2,w_3) = \frac{3}{4} - \frac{1}{4} w_1 w_2 - \frac{1}{4} w_1 w_3 - \frac{1}{4} w_2 w_3$$

By assumption

$$\mathbb{E}\, NAE(f(x), f(y), f(z)) = 1$$

So $\frac{3}{4} - \frac{3}{4}\, Stab_\rho[f] = 1$

$\Updownarrow$

$$Stab_\rho[f] = -\frac{1}{3}$$

$$= \sum_{S \subseteq [n]} \left(-\frac{1}{3}\right)^{|S|} \hat{f}(S)$$

Since $\hat{f}\left(-\frac{1}{3}\right)^k \geq -\frac{1}{3} \quad \forall k \in \mathbb{N}$

This is only possible if f is a linear function

$\hat{f}(S) = 0$ if $|S| \neq 1$

So $f = \sum_{i=1}^{n} \alpha_i w_{\{i\}}$, since f takes values in $\{-1,1\}$

it is only possible if $f = \pm w_{\{i\}}$ ∎