Analysis of Boolean Functions **(CMU 18-859S, Spring 2007)**

Lecture 20: Noise Stability of Majority

Mar. 29, 2007

Lecturer: Ryan O'Donnell *Scribe: Yi Wu*

Today we are going to show $\mathbb{S}_\rho(Maj_n)$ will converge to $\frac{2}{\pi}\arcsin(\rho)$ when n becomes big.

1 Berry-Essen Theorem(CLT with error bound)

Theorem 1.1 *(simplified) Let $x_1, x_2, ...x_n$ be independent r.v.s, and we assume*

- $\mathbf{E}(x_i) = 0$
- $\sum \mathbf{E}(x_i^2) = \sigma^2$
- $\forall i, |x_i| \le \eta\sigma$

Then $\sum x_i$ is distributed like a gaussian N(0,σ^2) satisfies that

- $\forall$ *intervals* $I \subseteq \mathbb{R}, |\mathbf{Pr}[\sum x_i \in I] - \mathbf{Pr}[N(0,\sigma^2) \in I]| < O(\eta)$
- $|\mathbf{E}[\sum x_i|] - \mathbf{E}[N(0,\sigma^2)]| \le O(\eta)$

From the theorem we can see, if x_i is some random bits, then

- $|\mathbf{Pr}([\sum \frac{x_i}{\sqrt{n}} \in I] - \mathbf{Pr}[N(0,1) \in I] \le O(\frac{1}{\sqrt{n}})$
- $|E[|\frac{\sum x_i}{\sqrt{n}}|] = \sqrt{\frac{2}{\pi}} \pm O(\frac{1}{\sqrt{n}})$

If $\sum \alpha_i^2 = 1$, then $|\mathbf{Pr}([\sum x_i\alpha_i \in I]) - \mathbf{Pr}[N(0,1) \in I]| \le O(\max|\alpha_i|)$.

2 Calculating Majority's Noise Stability

We want to calculate following value when n goes into infinity.

$$\mathbb{S}_\rho(Maj_n) = \mathbf{E}[Maj_n(x)Maj_n(y)]$$

The expectation is taken over random bit x and y satisfying that $y = x$ w.p. $\frac{1}{2} + \frac{1}{2}p$ and $y = -x$ w.p. $\frac{1}{2} - \frac{1}{2}p$.

Essentially,

$$\mathbb{S}_\rho(Maj_n) = 1 - 2\mathbf{Pr}[\frac{\sum x_i}{\sqrt{n}}, \frac{\sum y_i}{\sqrt{n}} \text{ have different signs}].$$

We can view $\begin{bmatrix} \sum \frac{x_n}{\sqrt{n}} \\ \sum \frac{y_i}{\sqrt{n}} \end{bmatrix}$ as the sum of n two dimension vector

$$\sum_{i=1}^{n} \begin{bmatrix} \frac{x_i}{\sqrt{n}} \\ \frac{y_i}{\sqrt{n}} \end{bmatrix}$$

There is a two dimension Berry-Esseen Theorem. It is saying that $\begin{bmatrix} \frac{\sum x_n}{\sqrt{n}} \\ \frac{\sum y_i}{\sqrt{n}} \end{bmatrix}$ will converge to some two dimension gaussian with mean $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and covariance matrix $\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$. More specifically, the error bound will be as follows:

$$\forall K \subseteq \mathbb{R}^2, |\mathbf{Pr}(\begin{bmatrix} \frac{\sum x_i}{\sqrt{n}} \\ \frac{\sum y_i}{\sqrt{n}} \end{bmatrix} \in K) - \mathbf{Pr}(N(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}) \in K)| \leq O(\frac{1}{\sqrt{1-\rho}\sqrt{n}})$$

Here a random variable $\begin{bmatrix} x \\ y \end{bmatrix}$ following distribution $N(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix})$ can be viewed as generated by following process:

1. generating $x \sim N(0,1)$
2. generating $y \sim \rho x + \sqrt{1-\rho^2}N(0,1)$

We already have that

$$\begin{aligned} \mathbb{S}_\rho(Maj_n) =& 1 - 2\mathbf{Pr}[\sum x_i/\sqrt{n}, \sum y_i/\sqrt{n} \text{ has different sign}] \\ =& \mathbf{Pr}(\operatorname{sgn}(x) \neq \operatorname{sgn}(y) \text{ for } \begin{bmatrix} x \\ y \end{bmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix})) + O(\frac{1}{n\sqrt{1-\rho}}) \end{aligned}$$

We only to understand the ρ correlated two dimensional gaussian.

Let $u = (1,0)$ and $v = (\rho, \sqrt{1-\rho^2})$ Let $Z = (Z_1, Z_2)$ be two independent gaussian. Then $x = u \cdot Z_1$ and $y = v \cdot Z_2$. Let $Z' = (-Z_2, Z_1)$ which is orthogonal to Z.

Then we have

$$\mathbf{Pr}[x, y \text{ has the different sign}] = \mathbf{Pr}[\text{ Z' split } u, v\ ] = \frac{\arccos(u \cdot v)}{\pi} = \frac{\arccos(\rho)}{\pi}.$$

Here we notice the fact that direction of Z' can be uniformly random from $[0, 2\pi)$ and it only have chance $\frac{\arccos(\rho)}{\pi}$ to split u, v.

So overall, we have

$$\mathbb{S}_\rho(Maj_n) = 1 - 2\frac{\arccos(\rho)}{\pi} + O(\frac{1}{n\sqrt{1-\rho}}) = \frac{2}{\pi}\arcsin(\rho) + O(\frac{1}{n\sqrt{1-\rho}})$$

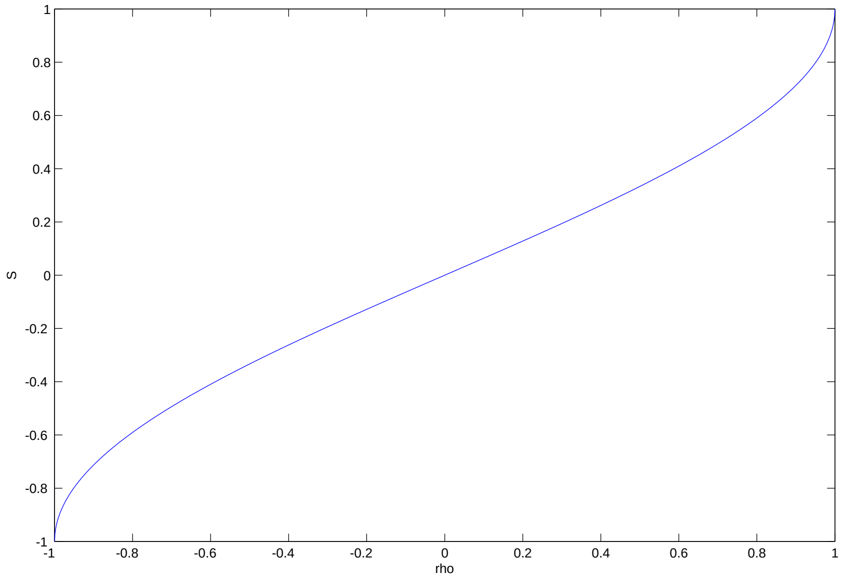

Figure 1: The curve

We plot the curve of $\frac{2}{\pi}\arcsin(\rho)$ as in Figure 1.

Further, when n is big, we would have

$$\begin{aligned}\mathbb{NS}_\delta(Maj) =& \tfrac{1}{2} - \tfrac{1}{2}\mathbb{S}_{(1-2\delta)}(Maj)\\ =& \frac{1}{\pi}\arccos(1-2\delta)\\ \sim& \frac{2}{\pi}\sqrt{\delta}\end{aligned}$$

Recall by the Peres's Theorem, $\mathbb{NS}(LTF) \leq (\sqrt{\frac{2}{\pi}} + O_\delta(1))\sqrt{\delta}$. So Majority function does not reach the bound.

It is a open question for odd n, whether any LTF f satisfies that

$$\mathbb{NS}_\delta(f) \leq \mathbb{NS}_\delta(Maj_n).$$

3 Majority is Stablest?

Theorem 3.1 *Suppose* $f = \operatorname{sgn}\sum \alpha_i x_i$. *Here* $\sum \alpha_i^2 = 1$. *Then*

$$\mathbb{S}_\rho = \frac{2}{\pi}(\arcsin\rho) \pm O(\frac{\max|\alpha_i|}{\sqrt{1-\rho}}).$$

Proof of the theorem is very similar to above.

Recall that it can be shown if

$$f = \mathrm{sgn}(\sum \alpha_i x_i), \sum \alpha_i^2 = 1.$$

Then

$$\max_i \mathrm{Inf}_i(f) = \theta(\max |\alpha_i|).$$

If f is LTF and $\mathrm{Inf}_i(f)$ is "small" for all i, then $\mathbb{S}_\rho(f) = \frac{2}{\pi} \arcsin(\rho) + \text{"}small\text{"}$.

We will show in later lecture that Majority function(or those small influence LTF) is the stablest for function with small influence. Let us state the theorem more formally as follows.

Theorem 3.2 ("Majority is stablest") *Fix $0 < \rho < 1$, Let $f : \{-1, 1\}^n \to [-1, 1]$ satisfies*

- $\mathbf{E}[f(y)] = 0$
- $\mathrm{Inf}_i f \leq \epsilon$*, for any* $i \in [n]$

Or

- f *is* $(\epsilon, 1/log(1/\epsilon))$ *quasirandom.*

Then $\mathbb{S}_\rho(f) \leq \frac{2}{\pi} \arcsin \rho + O(\frac{loglog\frac{1}{\epsilon}}{log1\epsilon})$