Majority is Stablest Theorem

Lecturer: Prahladh Harsha Scribe: Uma Girish

The goal of these lecture notes will be to prove the Majority is Stablest theorem.
We set up some notation. Let $\mathcal{D}_n$ denote the random variable that is uniformly distributed on $\{-1,1\}^n$. Let $\mathcal{G}_n$ denote the n-dimensional Gaussian with mean 0 and variance Id. We refer to the Gaussian ρ-stability of a function f as $GStab_\rho(f)$ and the discrete stability as simply $Stab_\rho(f)$. We let $a \approx_\epsilon b$ denote that $a = b + O_\epsilon(1)$ and $a \leq_\epsilon b$ denote that $a \leq b + O_\epsilon(1)$.

Theorem 0.1. *Given any balanced boolean function $f : \{0,1\}^n \to \{0,1\}$ for which each variable has small attenuated influence*[1]*, that is $Inf_i^{1-\epsilon}(f) \leq \epsilon$. Then it's stability is at most* that of the majority. For all $\rho \in [0,1]$, we have the following:*

$$Stab_\rho(f) \leq_\epsilon Stab_\rho(Maj_n) = 1 - \frac{2}{\pi}\arccos\rho$$

The proof idea will be to use a classical result of Borell on such an inequality for Gaussian measures and then lift it to discrete measures using Invariance Principle.

Theorem 0.2 (Borell's Isoperimetric Inequality[2]). *Let $f : \mathbb{R}^n \to [-1,1]$ be a function with $\mathop{\mathbb{E}}_{x\sim\mathcal{G}_n} f(x) = 0$. Then, the Gaussian ρ-stability of f is at most that of a half space through the origin. In other words,*

$$GStab_\rho(f) \leq 1 - \frac{2}{\pi}\arccos\rho \tag{1}$$

Proof of Theorem 0.1:

1. Given f statisfying the conditions of Theorem 0.1. Let us consider the function as a multilinear polynomial so that we can define its values on arbitrary points in $\mathbb{R}^n$. Now we have a notion of Gaussian stability of f. Since f is multilinear, we see that

$$GStab(f) = f(\rho) = Stab(f) \tag{2}$$

   Ideally, we would be able to apply Theorem 0.2 to f and conclude that

$$Stab_\rho(f) = GStab_\rho(f) \leq 1 - \frac{2}{\pi}\arccos\rho$$

   f does satisfy $\mathop{\mathbb{E}}_{x\sim\mathcal{G}_n} f(x) = 0$, however, f as a multilinear polynomial does not satisfy the condition of the range being in $[-1,1]$. To set this right, we will consider the function $\tilde{f}$ defined by the truncation of f to $[-1,1]$.

2. Let $\phi : \mathbb{R} \to [-1,1]$ be defined as follows.

$$\phi(x) = \begin{cases} x & \text{if } x \in (-1,1) \\ 1 & \text{if } x \geq 1 \\ -1 & \text{if } x \leq -1 \end{cases}$$

   We define $\tilde{f}$ to be $\phi(f)$. We would like to use (1) and (2) to conclude the theorem. For that we require:

   (a) $\mathop{\mathbb{E}}_{\mathcal{G}_n} \tilde{f} \approx_\epsilon \mathop{\mathbb{E}}_{\mathcal{G}_n} f = 0$

[1] The δ-attenuated influence of f is defined as $\sum_{S \ni i} \delta^{|S|} \hat{f}_S^2$.

[2] In fact, something more general is true. The most* stable function of expectation α is the half space with Gaussian volume α.

(b) $GStab(\tilde{f}) \approx_\epsilon GStab(f)$.

Suppose we have the above two facts, the proof is as follows. We let $\tilde{\tilde{f}}$ be $\tilde{f}$ that is changed at some points so that $\tilde{\tilde{f}} : \mathbb{R}^n \to [-1,1]$ and $\mathbb{E}[\tilde{\tilde{f}}] = 0$. For instance, we could invert the sign of f at a fraction of points. We will prove the following chain of inequalitites.

$$Stab(f) = GStab(f) \leq_\epsilon GStab(\tilde{f}) \leq_\epsilon GStab(\tilde{\tilde{f}}) \leq_\epsilon 1 - \frac{2}{\pi}\arccos\rho$$

We will now justify each step of the above. The first one follows from (2), the second one follows from (b). It is not too difficult to observe that since we have changed the values of $\tilde{f}$ 'only a little', we have $\mathbb{E}_{\mathcal{G}_n}\left(\tilde{f}(x) - \tilde{\tilde{f}}(x)\right)^2 \leq O_\epsilon(1)$. From this, we will show that $GStab(\tilde{\tilde{f}}) \approx_\epsilon GStab(\tilde{f})$. Recall:

$$GStab(\tilde{f}) = \mathop{\mathbb{E}}_{\substack{x\sim\mathcal{G}_n\\ y\sim_\rho x}} \tilde{f}(x)\tilde{f}(y) \qquad\qquad GStab(\tilde{\tilde{f}}) = \mathop{\mathbb{E}}_{\substack{x\sim\mathcal{G}_n\\ y\sim_\rho x}} \tilde{\tilde{f}}(x)\tilde{\tilde{f}}(y)$$

$$|GStab(\tilde{f}) - GStab(\tilde{\tilde{f}})| \leq \mathop{\mathbb{E}}_{\substack{x\sim\mathcal{G}_n\\ y\sim_\rho x}} |\tilde{f}(x)|\left|\tilde{f}(y) - \tilde{\tilde{f}}(y)\right| + \mathop{\mathbb{E}}_{\substack{x\sim\mathcal{G}_n\\ y\sim_\rho x}} |\tilde{\tilde{f}}(y)|\left|\tilde{f}(x) - \tilde{\tilde{f}}(x)\right|$$

We bound the two terms on the right as follows.

$$\mathop{\mathbb{E}}_{\substack{x\sim\mathcal{G}_n\\ y\sim_\rho x}} |\tilde{f}(x)||\tilde{f}(y) - \tilde{\tilde{f}}(y)| \leq \sqrt{\mathop{\mathbb{E}}_{x\sim\mathcal{G}_n} \tilde{f}(x)^2}\sqrt{\mathop{\mathbb{E}}_{y\sim\mathcal{G}_n}\left(\tilde{f}(y) - \tilde{\tilde{f}}(y)\right)^2} \leq O_\epsilon(1)$$

This is because $\mathbb{E}\tilde{f}(x)^2 = 1$ since $\tilde{f}(x) \in [-1,1]$. Similarly, since $\tilde{\tilde{f}}(x) \in [-1,1]$, we have:

$$\mathop{\mathbb{E}}_{\substack{x\sim\mathcal{G}_n\\ y\sim_\rho x}} |\tilde{\tilde{f}}(y)||\tilde{f}(x) - \tilde{\tilde{f}}(x)| \leq \sqrt{\mathop{\mathbb{E}}_{y\sim\mathcal{G}_n} \tilde{\tilde{f}}(y)^2}\sqrt{\mathop{\mathbb{E}}_{x\sim\mathcal{G}_n}\left|\tilde{f}(x) - \tilde{\tilde{f}}(x)\right|^2} \leq O_\epsilon(1)$$

Therefore, we have concluded that $GStab(\tilde{\tilde{f}}) \approx_\epsilon GStab(\tilde{f})$. The last step of the inequality is simply Theorem 0.2 applied to $\tilde{\tilde{f}}$.

$$GStab(\tilde{\tilde{f}}) \leq 1 - \frac{2}{\pi}\arccos\rho$$

Thus it suffices to prove (a) and (b). We will instead prove the following:

$$\mathop{\mathbb{E}}_{x\sim\mathcal{G}_n} (f(x) - \tilde{f}(x))^2 \leq O_\epsilon(1) \tag{3}$$

Why does (a) and (b) follow from this?

(a) $\mathbb{E}_{\mathcal{G}_n}|f - \tilde{f}| \leq \sqrt{\mathbb{E}_{\mathcal{G}_n}|f - \tilde{f}|^2} \leq O_\epsilon(1)$. Hence, $\mathbb{E}_{\mathcal{G}_n}\tilde{f} \leq_\epsilon \mathbb{E}_{\mathcal{G}_n} f = 0$.

(b) This calculation is almost identical to the one above. The only point of difference is in bounding the first term. We can no longer say $f(x) \in [-1,1]$ to conclude $\mathop{\mathbb{E}}_{x\in\mathcal{G}_n} f(x)^2 \leq 1$. However, we anyway have $\mathop{\mathbb{E}}_{x\in\mathcal{G}_n} f(x)^2 = 1$ since f is a multilinear polynomial whose sum of squares of coefficients is 1. Thus (b) follows from (3).

We have shown that (3) implies (a) and (b). We will now focus on proving (3), that is

$$\mathop{\mathbb{E}}_{x\sim\mathcal{G}_n} (f(x) - \tilde{f}(x))^2 \leq O_\epsilon(1)$$

3. We consider the function $\psi : \mathbb{R} \to [-1, 1]$ which measures the distance of a number to $[-1, 1]$, i.e,

$$\psi(x) = \begin{cases} 0 & \text{if } x \in (-1, 1) \\ (x-1)^2 & \text{if } x \geq 1 \\ (x+1)^2 & \text{if } x \leq -1 \end{cases}$$

Observe that the quantity in (3) is precisely $\underset{x \sim \mathcal{G}_n}{\mathbb{E}} \psi f(x)$. Furthermore, $\underset{x \sim \mathcal{D}_n}{\mathbb{E}} \psi f(x) = 0$. We now state the Invariance Principle and see how it allows us to go between discrete measures and Gaussian measures. We will use the following version of it[3].

Theorem 0.3 (Invariance Principle)**.** *Let F be a n-variate multilinear polynomial of degree at most k. Further assume that* $\boldsymbol{Var[F]} \leq 1$ *and* $\boldsymbol{Inf_i(F) \leq \epsilon}$ *for all* $i \in [n]$*. Assume* $\psi : \mathbb{R} \to \mathbb{R}$ *is* $\mathcal{C}^4$ *with* $\boldsymbol{\|\psi''''\|_\infty \leq C}$*. Then,*

$$\left| \underset{x \sim \mathcal{D}_n}{\mathbb{E}} \psi F(x) - \underset{y \sim \mathcal{G}_n}{\mathbb{E}} \psi F(y) \right| \lesssim \boldsymbol{Ck9^k \epsilon}$$

This says that if a boolean function F has *small influence* and *small degree* (roughly $\ln \frac{1}{\epsilon}$), and ψ is a 'reasonable' function, then $\underset{\mathcal{G}_n}{\mathbb{E}}[\psi(F)] \approx_\epsilon \underset{\mathcal{D}_n}{\mathbb{E}} \psi(F)$. If we can apply this to ψ and f, we can prove (3) as follows.

$$\mathbb{E}_{x \sim \mathcal{G}_n}(f(x) - \tilde{f}(x))^2 = \underset{x \sim \mathcal{G}_n}{\mathbb{E}} \psi f(x) \approx_\epsilon \underset{y \sim \mathcal{D}_n}{\mathbb{E}} \psi f(y) = 0$$

However, there are issues with applying this.

- ψ as defined earlier is not $\mathcal{C}^4$, because of sudden changes in the second derivative at -1 and 1. However, it can be smoothened at these points. Furthermore, we can ensure that $\|\psi''''\|_\infty \leq B$ for some absolute constant B. We will not do this calculation.
- The main issue is that f could have large degree. We would like to damp out the large degree coefficients and this we do by first applying the noise operator $T_{1-\delta}$.

4.

$$\text{Let } g := T_{1-\delta} f \quad \text{for} \quad \delta \leq \frac{1}{\ln \ln \frac{1}{\epsilon}} \leq \epsilon$$

Now g may still have large degree coefficients, but they are small in value. Firstly, we would like to see that for this function g, we still have the implications (3) $\implies$ (a),(b).

For the proof of (a) from (3), we only required that $\underset{\mathcal{G}_n}{\mathbb{E}} g = 0$, which is still true, because the noise operator doesn't touch the zero-th level coefficient. We also have $g(x) \in [-1, 1]$ on the vertices of the hypercube since the noise operator averages f over vertices of the hypercube, on which f takes values in $\{-1, 1\}$. Hence $\underset{\mathcal{G}_n}{\mathbb{E}} g(x)^2 \leq 1$ and (b) follows. Furthermore, $\underset{\mathcal{D}_n}{\mathbb{E}} \psi g(x) = 0$. Also,

$$Inf_i(g) = \sum_{S \ni i} \hat{g}_S^2 = \sum_{S \ni i} (1-\delta)^{2|S|} \hat{f}_S^2 \leq \sum_{S \ni i} (1-\epsilon)^{|S|} \hat{f}_S^2 = Inf_i^{1-\epsilon}(f) \leq \epsilon$$

Suppose using all this and the invariance principle on ψ and g, we were able to conclude $Stab_\rho(g) \leq_\epsilon 1 - \frac{2}{\pi} \arccos \rho$. How do we transfer the same conclusion to f? Recall

$$Stab_\rho(g) = \sum_S \rho^{2|S|} (1-\delta)^{|S|} \hat{f}_S^2$$

$$\implies Stab_\rho(f) - Stab_\rho(g) = \sum_S \rho^{|S|} \left(1 - (1-\delta)^{2|S|}\right) \hat{f}_S^2$$

[3] See Ryan O'Donell's book [**?**] on Page 363.

We make the following observation.

$$\text{For } |S| \geq \frac{1}{\sqrt{\delta}} \qquad \rho^{|S|} \leq \rho^{1/\sqrt{\delta}} \leq \sqrt{\delta} \text{ for } \delta \text{ not too small}$$
$$\text{For } |S| \leq \frac{1}{\sqrt{\delta}} \qquad 1-(1-\delta)^{|S|} \leq \sqrt{\delta}$$

This will help us conclude that $Stab_\rho(f) \leq_\delta Stab_\rho(g)$ and thus obtain the theorem. However, we cannot apply invariance to g yet as we still have large degree coefficients in g. But their coefficients are small, so we truncate g as follows.

5. $$\text{Let } h := g^{\leq \frac{1}{\delta^2}}$$

Again, we would like to repeat the whole argument replacing the function with h. For (3) implies (a) we require $\mathop{\mathbb{E}}_{\mathcal{G}_n} h = 0$ which is still true, because we only truncated the large degree part. Also, $Inf_i(h) \leq \epsilon$ as we are only truncating terms in the expression for influence. We no longer have $h(x) \in [-1,1]$ or $\mathbb{E}(\psi(h(x))) = 0$. But we note that

$$\|g-h\|_2^2 = \sum_{|S| \geq \frac{1}{\delta^2}} \hat{g}_S^2 = \sum_{|S| \geq \frac{1}{\delta^2}} (1-\delta)^{2|S|} \hat{f}_S^2 \leq \sum_{|S| \geq \frac{1}{\delta^2}} e^{-1/\delta} \hat{f}_S^2 \leq e^{-1/\delta} \leq O_\epsilon(1)$$

From this, we can conclude that $\mathop{\mathbb{E}}_{\mathcal{G}_n} h(x)^2 \leq_\epsilon \mathop{\mathbb{E}}_{\mathcal{G}_n} g(x)^2 \leq_\epsilon 1$ and (b) follows with an additional ϵ loss. $Var(h) \leq_\epsilon 1$. Furthermore,

$$|\mathbb{E}\psi h(x)| \leq \mathbb{E}(h(x) - g(x))^2 \leq O_\epsilon(1)$$

Thus, we have all the necessary conditions satsfied for h to apply the Invariance principle and conclude that $Stab_\rho(h) \leq_\epsilon 1 - \frac{2}{\pi}\arccos\rho$. How do we transfer the same conclusion to g? Since h is the trunctation of g to $\leq \frac{1}{\delta^2}$ terms, the only difference in the stability of h and g comes from higher order terms, whose mass is less than $O_\epsilon(1)$, therefore $Stab_\rho(g) \leq_\epsilon Stab_\rho(h)$. From point 4, we conclude that $Stab_\rho(f) \leq_\epsilon Stab_\rho(g) \leq_\epsilon 1 - \frac{2}{\pi}\arccos\rho$. $\square$