

Lecture 28: Large Sets fool Large Linear Tests

Intuition

- Let $A \subseteq \{0, 1\}^n$ be a set of size 2^{n-t}
- We want to claim that the uniform distribution over the set A fools (most) large linear tests
- For example, consider A to be the set of n -bit strings that start with t 0s
- Consider any linear test S such that the support of S is restricted only to the first t indices. Then, the output of this linear test is completely biased (it always outputs 0)
- On the other hand, if S has support that is larger than t , then the output of the linear test is uniformly random bit. That is, the uniform distribution over A fools this linear test
- In general, we cannot expect to fool all large support linear tests. For example, we consider A to be the n -bit strings with even number of 1s. The uniform distribution over A does not fool the linear test corresponding to $S = N - 1$

- Let $A \subseteq \{0, 1\}^n$ such that $|A| = 2^{n-t}$
- Let $1_{\{A\}}$ be the indicator variable for the subset A
- Note that the uniform distribution over A is represented by the function

$$\frac{1}{|A|} 1_{\{A\}}$$

- Note that the bias of the output of the linear test S is

$$\text{Bias}_{\Delta}(S) := \frac{N}{|A|} \widehat{1_{\{A\}}}(S)$$

- Let us evaluate the sum of all the biases corresponding to linear tests S such that $|S| = k$

$$\sum_{S \in \{0,1\}^n : |S|=k} \text{Bias}_{\Delta}(S)^2 = \left(\frac{N}{|A|} \right)^2 \sum_{S \in \{0,1\}^n : |S|=k} \widehat{1_{\{A\}}}(S)^2$$

- Recall that the KKL Lemma states that, for any $\delta \in (0, 1)$ and $f: \{0, 1\}^n \rightarrow \{+1, 0, -1\}$, we have

$$\sum_{S \in \{0,1\}^n} \delta^{|S|} \widehat{f}(S)^2 \leq \mathbb{P} \left[f(x) \neq 0 : x \stackrel{s}{\leftarrow} \{0, 1\}^n \right]^{2/1+\delta}$$

- Note that, we have

$$LHS \geq \sum_{S \in \{0,1\}^n : |S|=k} \delta^k \widehat{f}(S)^2$$

- So, we conclude that

$$\sum_{S \in \{0,1\}^n : |S|=k} \widehat{f}(S)^2 \leq \frac{1}{\delta^k} \mathbb{P} \left[f(x) \neq 0 : x \stackrel{s}{\leftarrow} \{0, 1\}^n \right]^{2/1+\delta}$$

- Substituting $f = 1_{\{A\}}$, we get

$$\begin{aligned} \sum_{S \in \{0,1\}^n : |S|=k} \text{Bias}_{\mathbb{A}}(S)^2 &\leq \left(\frac{N}{|A|}\right)^2 \cdot \frac{1}{\delta^k} \cdot \left(\frac{|A|}{N}\right)^{2/1+\delta} \\ &= \frac{1}{\delta^k} \cdot \left(\frac{N}{|A|}\right)^{2\delta/1+\delta} \\ &\leq \frac{1}{\delta^k} \left(\frac{N}{|A|}\right)^{2\delta} = 2^{2t\delta - k \lg e \ln \delta} \end{aligned}$$

- Now, we choose δ that minimizes the RHS above. That value of δ is $\delta = k \lg e / 2t$
- Substituting this value of δ we get

$$\sum_{S \in \{0,1\}^n : |S|=k} \text{Bias}_{\mathbb{A}}(S)^2 \leq \left(\frac{2et}{k \lg e}\right)^k$$

- The average bias is

$$\binom{n}{k}^{-1} \sum_{S \in \{0,1\}^n: |S|=k} \text{Bias}_{\mathbb{A}}(S)^2 \leq \left(\frac{2e}{\lg e} \cdot \frac{t}{n} \right)^k = \left(O(t/n) \right)^k$$

Tightness of the Bound

- The bound we obtain above is essentially tight
- Consider A such that the first t bits of its elements are all 0
- Note that $\binom{t}{k}$ linear tests have bias 1
- The remaining linear tests have bias 0
- So, the average bias is

$$\binom{t}{k} \binom{n}{k}^{-1} \geq \left(\frac{1}{e} \cdot \frac{t}{n}\right)^k$$