6.842 Randomness and Computation	April 11, 2012
Lecture 17	
Lecturer: Ronitt Rubinfeld	*Scribe: Jiayang Jiang*

Today we will discuss the following:

- Learning parity functions with "noise" (continued)

1 Review

Given a black-box function $f : \{\pm 1\}^n \to \{\pm 1\}$, we want to find the heavy coefficients. More specifically, we want to output all $S \subseteq [n]$ such that $\hat{f}(S) > \theta$ and no S such that $\hat{f}(S) < \frac{\theta}{2}$.

Last time, we looked at these cases:

- Warmup 2: $f = \chi_S$ for some S

  The algorithm works as follows:

  - For each i, put i in S if $f(1, \dots, 1) \neq f(1, \dots, 1, -1, 1, \dots, 1)$, where the -1 is in the ith position.

- Warmup 3: $\Pr[f(x) = \chi_S(x)] \geq \frac{3}{4} + \frac{\theta}{2}$

  The algorithm works as follows:

  - Choose $r_1, \dots, r_t \in \{\pm 1\}^n$, with $t = \Theta(\frac{\log n}{\theta^2})$.
  - For all $i \in [n]$, put i in S if the majority of $f(r_j) \neq f(r_j \oplus e_i)$.

  For any $r \in \{\pm 1\}^n$, we get a wrong answer if either $f(r) \neq \chi_S(r)$ or $f(r \oplus e_i) \neq \chi_S(r \oplus e_i)$. Using a union bound, Pr[test does not work] $\leq \frac{1}{2} - \theta$. In particular, Pr[test works] $> \frac{1}{2}$. So by repeating this for $r_1, \dots, r_t$, we can put i in S correctly with high probability.

2 Warmup 4

We want to output all S such that f agrees with χ_S on $\geq \frac{1}{2} + \frac{\theta}{2}$ fraction of inputs. We consider θ as a constant for now.

First note that the algorithm for Warmup 3 cannot be extended in a straightforward way to Warmup 4, since now we have $\Pr[f(x) = \chi_S] \geq \frac{1}{2} + \frac{\theta}{2}$. So on each r, the test $f(r) \neq f(r \oplus e_i)$ could fail with high probability, by union bound.

The algorithm works as follows:

- Choose $r_1, \dots, r_t \in \{\pm 1\}^n$, with $t = \Theta(\log n)$.
- For all possible settings of $\sigma_1, \dots, \sigma_t \in \{\pm 1\}$,
  - For all $i \in [n]$, put i in $S_{\sigma_1 \dots \sigma_t}$ if the majority of $\sigma_j \neq f(r_j \oplus e_i)$
  - Sample to see if $S_{\sigma_1 \dots \sigma_t}$ agrees with f on $\geq \frac{1}{2} + \frac{3}{8}\theta$ of the inputs. If yes, output $\chi_{S_{\sigma_1 \dots \sigma_t}}$.

Intuitively, the algorithm uses $\sigma_1, \dots, \sigma_t$ as guesses for $\chi_S(r_1), \dots, \chi_S(r_t)$. And for these guesses, it generates a candidate set $S_{\sigma_1 \dots \sigma_t}$, which we then sample and test the candidate set to eliminate cases where $\hat{f}(S_{\sigma_1 \dots \sigma_t}) < \frac{\theta}{2}$ (since the guesses may be totally wrong).

Behavior:

- Since we picked $t = \Theta(\log n)$, enumerating all the possible settings of $\sigma_1, \ldots, \sigma_t$ will take $2^t = \text{poly}(n)$ trials.
- We already have an algorithm to estimate any Fourier coefficient, and the last step of the algorithm for each setting $\sigma_1, \ldots, \sigma_t$ uses this algorithm to estimate $\hat{f}(S_{\sigma_1,\ldots,\sigma_t})$. Using Chernoff bounds, we can determine whether $S_{\sigma_1,\ldots,\sigma_t}$ satisfies the condition $\hat{f}(S_{\sigma_1,\ldots,\sigma_t}) > \frac{\theta}{2}$ with high probability.
- For each S that should be output, some setting of $\sigma_1, \ldots, \sigma_t$ agrees with χ_{S} for all j. In other words, for this setting, $\chi_{\text{S}}(r_j) = \sigma_j$ for all j.

  For this setting,

$$\begin{aligned}\Pr[\text{wrong answer on } r_j \text{ for } i] &= \Pr[\sigma_j f(r_j \oplus e_i)(-1)^{1_{i\in S}} = -1] \\ &\leq \Pr[f(r_j \oplus e_i) \neq \chi_{\text{S}}(r_j \oplus e_i)] \\ &\leq \frac{1}{2} - \frac{\theta}{2}\end{aligned}$$

  The wrong answer is according to whether $i \in S$. The second line follows from the fact that $\sigma_j = \chi_{\text{S}}(r_j)$ and $f(r_j \oplus e_i)$ is uniformly distributed.

  Using Chernoff bounds and the fact that $t = \Theta(\log n)$, we can show that $\Pr[\text{wrong answer on } i] \leq \frac{1}{cn}$ for some constant c. Finally, by the union bound, $\Pr[\text{wrong answer on any } i] \leq \frac{1}{c}$. Therefore, S is output with probability at least $1 - \frac{1}{c}$.

3 Algorithm for the General Case

We want to output all S such that f and χ_{S} agree on $\geq \frac{1}{2} + \frac{\theta}{2}$ fraction of inputs, where θ can be $1/\text{poly}(n)$.

Note that now θ is not necessarily constant. In Warmup 4, t actually has order $\Theta(\frac{\log n}{\theta^2})$, so to enumerate over all possible settings $\sigma_1, \ldots, \sigma_t$, the running time of the could be exponential. To solve this problem, we use pairwise independence.

The algorithm works as follows:

- Choose $s_1, \ldots, s_k \in \{\pm 1\}^n$, where the number of guesses is $k = \log(t+1)$, and the number of r_i's generated is $t \geq \frac{cn}{\theta^2}$.
- For all possible settings $\sigma_1, \ldots, \sigma_k \in \{\pm 1\}^n$ (guesses of $\chi_{\text{S}}(s_i)$):
  - For every $W \subseteq \{1, \ldots, k\}, W \neq \emptyset$,
    $r_{\text{W}} \leftarrow \bigoplus_{j\in W} s_j$
    $p_{\text{W}} \leftarrow \prod_{j\in W} \sigma_j$
  - For all $i \in [n]$, put i in $S_{\sigma_1 \ldots \sigma_k}$ if the majority of $p_{\text{W}} \neq f(r_{\text{W}} \oplus e_i)$.
  - Test $S_{\sigma_1 \ldots \sigma_k}$ to see if it agrees with f on at least $\frac{1}{2} + \frac{3}{8}\theta$ of the inputs. If yes, output $\chi_{\text{S}_{\sigma_1 \ldots \sigma_k}}$.

<u>Behavior</u>:

- For S such that f agrees with χ_{S} on at least $\frac{1}{2} + \frac{\theta}{2}$ fraction of the inputs, if for all j, $\sigma_j = \chi_{\text{S}}(s_j)$, the setting $\sigma_1, \ldots, \sigma_k$ is correct. Then,

$$\begin{aligned}p_{\text{W}} &= \prod_{j\in W} \chi_{\text{S}}(s_j) \\ &= \chi_{\text{S}}(\bigoplus_{j\in W} s_j) \\ &= \chi_{\text{S}}(r_{\text{W}})\end{aligned}$$

- By construction, the r_{W}'s are pairwise independent.
- Let $X_W = 1$ if $p_{\mathrm{W}} f(r_{\mathrm{W}} \bigoplus e_i)(-1)^{1_{i \in S}}$ and $X_W = 0$ otherwise. Thus, $X_W = 1$ iff W gives the right answer for χ_{S}. Note that the X_W's are pairwise independent.

  We use Chebyshev's inequality to show the probability S is not chosen is low.

  Theorem 1 (Chebyshev's inequality) *Let $X_1, \ldots, X_n$ be pairwise independent random variables, each with mean $\mu = E[X_i]$ and variance $\sigma^2 = \mathrm{Var}(X_i)$, then for any $\epsilon > 0$, $\Pr[|\frac{\sum_{i=1}^n X_i}{n} - \mu| > \epsilon] \leq \frac{\sigma^2}{\epsilon^2 n}$*

  Here, $E[X_W] \geq \frac{1}{2} + \frac{\theta}{2}$ and $\mathrm{Var}(X_W) \geq \frac{1}{4} - \frac{\theta^2}{4}$.
  Let $X = \sum_{W \subseteq [k]} X_W$. Then $E[\frac{X}{t}] \geq \frac{1}{2} + \frac{\theta}{2}$. Using Chebyshev's inequality and the fact that $t \geq \frac{cn}{\theta^2}$,

$$\begin{aligned} \Pr[i \text{ is not placed correctly in } S] &= \Pr\left[\frac{X}{t} < \frac{1}{2}\right] \\ &\leq \frac{1}{cn} \end{aligned}$$

  Using union bound, $\Pr[S \text{ is not chosen}] \leq \frac{1}{c}$.