6.889 Sublinear Time Algorithms	April 3, 2019
Lecture 15	
Lecturer: Ronitt Rubinfeld	*Scribe: Allen Liu*

1 Testing Monotonicity of Distributions

Definition 1 *A probability distribution P over $[n]$ is monotone decreasing if $P(i) \geq P(i+1)$ for all $1 \leq i \leq n-1$.*

For the remainder of these notes, we use the term monotone to refer to a distribution that is monotone decreasing. Our goal will be to develop a monotonicity tester that satisfies the following properties.

- If P is monotone, the tester outputs pass with probability at least $\frac{3}{4}$
- If P is ϵ-far from any monotone distribution, the tester outputs fail with probability at least $\frac{3}{4}$

1.1 Birge Decomposition

We now introduce the idea of a Birge Decomposition which will be central to developing a monotonicity tester. Decompose the domain into $l = \theta\left(\frac{\log(\epsilon n)}{\epsilon}\right) \sim \theta\left(\frac{\log n}{\epsilon}\right)$ intervals $I_1^\epsilon, \dots, I_l^\epsilon$, in order, such that I_k^ϵ has length $\lfloor (1+\epsilon)^k \rfloor$. Note that the Birge Decomposition is oblivious in the sense that it does not depend on the actual distribution being studied.

Definition 2 *Given a probability distribution q on $[n]$, we define the flattened distribution $\tilde{q}$ as $\tilde{q}(i) = \frac{q(I_j)}{|I_j|}$ where I_j is the interval in the Birge decomposition that contains i.*

The distribution $\tilde{q}$ is constant on each interval of the Birge decomposition. The following result implies that $\tilde{q}$ is actually a good approximation of q if q is monotone or even just close to monotone.

Theorem 3 (Birge's Theorem) *If q is monotone then $||q - \tilde{q}||_1 \leq \epsilon$.*

Proof We will give a sketch of a proof that gives a bound of $||q - \tilde{q}||_1 \leq O(\epsilon)$ as this will suffice for our purposes. For an interval I_j in the Birge decomposition, let x_j denote the smallest element and y_j denote the largest element. Note that $x_{j+1} = y_j + 1$. Also, the L^1 error between q and $\tilde{q}$ on an interval I_j is at most

$$(q(x_j) - q(y_j))|I_j| \leq (q(x_j) - q(x_{j+1}))|I_j|$$

We now consider three types of intervals.

- Length 1 intervals
- Short intervals of length $|I_j| \leq \frac{1}{\epsilon}$
- Long intervals of length $|I_j| > \frac{1}{\epsilon}$

Let $S \subset [l]$ be the subset of indices such that I_j is a short interval and $L \subset [l]$ be the subset of indices such that I_j is a long interval. Note that q and $\tilde{q}$ agree exactly on any length 1 interval. Also, there are $\Omega(\frac{1}{\epsilon})$ intervals of length 1 so for any I_j with $|I_j| > 1$, the probability that q assigns to any element in I_j is at most $O(\epsilon)$. Next

$$||q - \tilde{q}||_1 \leq \sum_{j \in S} |I_j|(q(x_j) - q(x_{j+1})) + \sum_{j \in L} |I_j|(q(x_j) - q(x_{j+1}))$$

We now deal with the first term. Let $c = \lfloor \frac{1}{\epsilon} \rfloor$. For each $2 \leq i \leq c$, let z_i be the smallest element in the first interval of length i. Consider combining all intervals of length 2, all intervals of length 3, and so on. We can rearrange

$$\begin{aligned}\sum_{j \in S} |I_j|(q(x_j) - q(x_{j+1})) &= 2(q(z_2) - q(z_3)) + \cdots + c(q(z_c) - q(z_{c+1})) \\ &= 2q(z_2) + q(z_3) + \cdots + q(z_c) - cq(z_{c+1})\end{aligned}$$

Now to deal with the second term, let l_0 be the lowest index such that $|I_{l_0}| > \frac{1}{\epsilon}$. Note we must have $|I_{l_0}| = c + 1$ and $x_{l_0} = z_{c+1}$. We rearrange

$$\sum_{j \in L} |I_j|(q(x_j) - q(x_{j+1})) = |I_{l_0}|q(x_{l_0}) + \sum_{j=l_0}^{l} q(x_{j+1})(|I_{j+1}| - |I_j|)$$

Combining this with the previous expression,

$$||q - \tilde{q}||_1 \leq 2q(z_2) + q(z_3) + \cdots + q(z_c) + q(z_{c+1}) + \sum_{j=l_0}^{l} q(x_{j+1})(|I_{j+1}| - |I_j|)$$

Let s_i be the total length of all intervals of length i for $2 \leq i \leq c$. Note $s_i \sim \Omega(\frac{1}{\epsilon})$. This means there is a constant C such that $s_i \geq \frac{C}{\epsilon}$ for all i. Rearranging the previous expression we get $\frac{s_i \epsilon}{C} \geq 1$. Also $|I_{j+1}| - |I_j| \sim \epsilon |I_j|$. Since the total probability mass of q is 1, we have

$$q(z_3)s_2 + \cdots + q(z_{c+1})s_c + \sum_{j=l_0}^{l} q(x_{j+1})|I_j| \leq 1$$

Thus,

$$q(z_3) + \cdots + q(z_c) + q(z_{c+1}) + \sum_{j=l_0}^{l} q(x_{j+1})(|I_{j+1}| - |I_j|) \leq$$

$$\frac{\epsilon}{C}(q(z_3)s_2 + \cdots + q(z_{c+1})s_c) + O(\epsilon)\left(\sum_{j=l_0}^{l} q(x_{j+1})|I_j|\right) \leq O(\epsilon)$$

and we immediately get

$$||q - \tilde{q}||_1 \leq 2q_2 + O(\epsilon) = O(\epsilon)$$

■

Corollary 4 *If q is ϵ-close to monotone then* $||q - \tilde{q}||_1 \leq O(\epsilon)$

1.2 Monotonicity Tester

We are now ready to give an algorithm for testing monotonicity of a distribution q on $[n]$. Below, $\epsilon' = \frac{1}{poly(\frac{1}{\epsilon})}$ is a parameter that will be set in terms of the desired precision ϵ.

Algorithm 1 Monotonicity Tester

Take $m \sim O\left(\sqrt{n} \cdot poly\left(\log n, \frac{1}{\epsilon'}\right)\right)$ samples from q
for $j \in \{1, 2, \ldots, l\}$ **do**
  Uniformity test on $I_j^{\epsilon'}$ using samples that fall in the interval
end for
if Greater than ϵ'-fraction of samples are in a failing interval **then**
  Output fail
end if
Compute weights w_j to be the fraction of samples in $I_j^{\epsilon'}$
if There exists some monotone distribution P such that $\{w_1, \ldots, w_l\}$ and $\{P(I_1^{\epsilon'}), \ldots, P(I_l^{\epsilon'})\}$ are ϵ'-close **then**
  Output pass
end if
Output fail

Note that the last step involves solving a linear program on $O(\log n)$ variables, which can be done efficiently. The sample complexity comes from the fact that we need $\sqrt{I_i}$ samples in each interval and there are a total of $O\left(\frac{\log n}{\epsilon'}\right)$ intervals.

1.3 Analysis of Monotonicity Tester

We will give an outline of the proof that our tester indeed has the desired behavior. Let ϵ be the desired precision parameter.

1.3.1 Distributions that are monotone

- Consider the flattened distribution $\tilde{q}$ with parameter ϵ'. By Birge's Theorem, $||q - \tilde{q}|| \leq \epsilon'$.
- In general the uniformity test is only guaranteed to pass a distribution that is exactly uniform. However, we can show that for distributions where the ratio of the maximum and minimum probabilities is at most 2, the uniformity test is likely to pass distributions that are $\frac{\epsilon'}{2}$-close to uniform.
- Note we can simply disregard elements whose weight is less than $\frac{\epsilon}{2n}$. After eliminating such elements, by monotonicity, there are at most $\sim \log n$ intervals in the partition that do not satisfy the above property. We can then show that the total weight in such intervals is small.

1.3.2 Distributions that are ϵ-far from monotone

- If q is likely to pass, then the weights of the Birge intervals $\{w_1, \ldots, w_l\}$ must be close to the weights of some monotone distribution q'
- Since almost all of the weight of q must be on Birge intervals where it is close to uniform, we can correct q to the flattened distribution q' with small error

1.4 Learning Monotone Distributions

We now approach monotone distributions from a slightly different perspective. The question we ask is given a distribution q that is promised to be monotone, how many samples do we need to construct an estimate q' such that $||q - q'||_1 \leq \epsilon$.

It turns out that $O(\frac{1}{\epsilon^3} \log n)$ samples suffice as by Birge's theorem, it suffices to estimate the flattened distribution $\tilde{q}$ which we can do by simply counting the fraction of samples in each Birge interval.