1 Overview

In the last lecture we talked about communication complexity. We found that deterministic exact $F_0$ required $\Omega(n)$ space. We also proved the inequality $D(f) \geq R_{\delta}^{pri}(f) \geq R_{\delta}^{pub}(f) \geq D_{p,s}(f)$ that we use to provide lower bounds on randomized and approximate algorithms. Our main tool was reductions from INDEX which has a communication complexity of roughly $\Omega(n)$.

In this lecture we use INDEX to give a lower bound on the space usage of randomized exact $F_0$. We then present a lower bound for randomized approximate $F_0$, something that we have thus far been unable to do. We then provide a lower bound on $F_p$ via the disjointness problem. Then we move on to dimensionality reduction, distortion, and distributional Johnson-Lindenstrauss and the fact that it implies Johnson-Lindenstrauss.

2 Randomized Exact Bound $F_0$

We prove that randomized exact $F_0$ requires $\Omega(n)$ space with failure probability $\frac{1}{3}$.

Proof. We perform the following reduction from INDEX. Let Alice receive $x \in \{0,1\}^n$ and Bob receive $j \in [n]$. It is then Bob’s job to find the $j$’th index of $x$. They proceed in the following manner. Alice runs our $F_0$ algorithm on $x$ and sends both the memory contents of the algorithm and the support of $x$ to Bob. Bob then appends $j$ to the stream and queries $F_0$ from the the memory contents of the algorithm. If $F_0$ increases, Bob outputs 0, else he outputs 1. We conclude that for $S$ equal to the space usage of the algorithm

$$S + \log n \geq c \cdot n \implies S = \Omega(n)$$

Where $\log n$ factor comes from sending the support of $x$. 

3 Randomized Approximate Bound $F_0$

To prove randomized approximate $F_0$ has space lower bound $\Omega(\log n)$ we first state this theorem that can be found in Kushilevitz and Nisan. Roughly speaking, it lower bounds the private communication bound by the log of the deterministic communication bound.

**Theorem 1.** $\forall f : \{0,1\}^n x \{0,1\}^n \rightarrow \{0,1\}$, where $f$ is a communication problem, then

$$R_{\frac{1}{3}}^{pri} \geq \Omega(\log(D(f)))$$
Proof. If we view $f$ as a two player game between Alice and Bob on a binary tree of height $s$ and total leaves $2^s$, than Alice and Bob could deterministically simulate the private randomized procedure on this tree. For instance, for any path from root to leaf, Alice can compute the probability she would stay on the path given that Bob does as well. She can then send these probabilities to Bob for every single leaf. Bob can then compute the probabilities he stays on the same paths and can output the final result accordingly.

Now we prove that randomized approximate $F_0$ has space lower bound $\Omega(\log n)$

Proof. Let $C$ be a subset of $\{0,1\}^n$ such that $\forall c \in C$ the support of $c$ is $\frac{n}{100}$. Also $\forall c \neq c' \in C$, we have $|c \cap c'| \leq \frac{n}{2000}$. Finally $|C| \geq 2^{\Omega(n)}$. We have constructed this set in previous lectures. In essence, it is a collection of subsets that are largely disjoint but very numerous. We know deterministic equality, $EQ$, on $C$ requires $\Omega(n)$ communication. Then using Kushilevitz Nisan we have

$$R_{\frac{\delta}{3}}^{opt}(EQ_C) \geq \Omega(\log n)$$

Now we notice there is a natural reduction from $EQ_C$ to randomized approximate $F_0$. Namely, Alice runs $F_0$ on her set $c$ and sends the memory contents to Bob. Bob then runs $c'$ on the memory contents and determines whether the output for $F_0$ has roughly doubled. If it has, then $c \neq c'$, if not, than $c = c'$.

4 Disjointness Problem

We now move on to the $t$-player disjointness problem, useful for proving lower bounds for the $F_p$. We have $t$-players $p_1, p_2, ..., p_t$. We assign an $n$ bit string $x_i \in \{0,1\}^n$ to player $p_i$. We are then promised that either of the following conditions hold.

1. $\forall i \neq j$ we have $x_i \cap x_j \neq \emptyset$

2. $\exists k \in [n]$ such that $\forall i \neq j$ we have $x_i \cap x_j = \{k\}$

The problem is then to find $k$ with the least communication possible where communication occurs from player 1 passing on to player 2 and so on and so forth until player player $t$ gives the final result.

Jelani mentions this theorem but does not prove it because it takes too much time. The proof also uses an information theoretic approach, known as information complexity [4]. The idea is the following chain of inequalities, where $\Pi$ is the optimal $\delta$-error communication protocol for some function $f$: $R_{\delta}^{pub}(f) = |\Pi| \geq H(\Pi(X)) \geq I(X;\Pi(X))$, where $X$ is the set of inputs given to the $t$ players, and $\Pi(X)$ is the transcript of the communication protocol (or the “communication log”) when the input is $X$ (note that it is a random variable since $\Pi$ uses randomness). Then we define the information complexity $IC_{\mu,\delta}(f)$ as the minimum value $I(X,\Pi(X))$ achievable by any $\delta$-error protocol $\Pi$ when $X$ is drawn from distribution $\mu$. Then we have that $R_{\delta}^{pub}(f) \geq IC_{\mu,\delta}(f)$ for all $\mu$. A variant of this approach was used by [2] to obtain lower bounds for $t$-player disjointness, with improvements in [3]. The sharp bound was shown in [5], with a later work showing how the arguments in [2] could be strengthened to also get the sharp bound [6].
Theorem 2. $R_{\text{pub}}^{\text{pub}}(\text{DISJ}_t) = \Omega\left(\frac{n}{t^2}\right)$

Remark: Although we do not prove the theorem we know that it implies some player sends $\Omega\left(\frac{n}{t^2}\right)$ bits which is what we'll need to prove the following claim.

Claim 3. For $p > 2$ the randomized $1.1$ approximation to $F_p$ requires $\Omega(n^{1-\frac{2}{p}})$ bits of space.

Proof. Set $t = \lceil(2n)^{\frac{1}{p}}\rceil$ for the disjoint players problem. Each player creates a virtual stream containing $j$ if and only if $j \in x_i$. We then compute $F_p$ on these virtual streams. If all $x_i$ are disjoint then $F_p \leq n$. Otherwise, $F_p \geq t^p \geq 2n$ because some element $k$ must appear at least $t$ times. Then since our $F_p$ algorithm is a $1.1$ approximation, we can discern between the two cases. This implies the space usage of our algorithm, $S$ satisfies

$$S \geq \frac{n}{t^2} = \Omega\left(n^{1-\frac{2}{p}}\right)$$

as desired. \qed

5 Dimensionality Reduction

Dimensionality reduction is useful for solving problems involving high dimensional vectors as input. Typically we are asked to preserve certain structures such as norms and angles. Some of the problems include

1. nearest neighbor search
2. large scale regression problems
3. minimum enclosing ball
4. numerical linear algebra on large matrices
5. various clustering applications

Our goals run in the same vein as streaming. That is to say fast runtime, low storage, and low communication. Certain geometric properties that we would like to preserve upon lowering the dimension of the input data include

1. distances
2. angles
3. volumes of subsets of inputs
4. optimal solution to geometric optimization problem

First and foremost, we would like to preserve distances, and to do so we must first define distortion.
5.1 Distortion

Definition 4. Suppose we have two metric spaces, \((X, d_X)\) and \((Y, d_Y)\), and a function \(f : X \to Y\). Then \(f\) has distortion \(D_f\) if \(\forall x, x' \in X, C_1 \cdot d_X(x, x') \leq d_Y(f(x), f(x')) \leq C_2 \cdot d_X(x, x')\), where \(\frac{C_2}{C_1} = D_f\).

We will focus on spaces in which \(d_X(x, x') = \|x - x'\|_X\) (ie. normed spaces).

5.2 Limitations of Dimensionality Reduction

If \(\|\cdot\|_X\) is the \(l_1\) norm, then \(D_f \leq C = \Rightarrow\) in worst case, target dimension is \(n \Omega\left(\frac{1}{C^2}\right)\). That is, there exists a set of \(n\) points \(X\), such that for all functions \(f : (X, l_1) \to (X', l_1')\), with distortion \(\leq C\), then \(m\) must be at least \(n \Omega\left(\frac{1}{C^2}\right)\) \cite{10}.

More recently in 2010, we have the following theorem by Johnson and Naor \cite{12}

Theorem 5. Suppose \((X, \|\cdot\|_X)\) is a complete normed vector space or "Banach Space" such that for any \(N\) point subset of \(X\), we can map to \(O(\log n)\) dimension subspace of \(X\) with \(O(1)\) distortion, then every \(n\)-dimensional linear subspace of \(X\) embeds into \(l_2\) with distortion \(\leq 2^{O(\log^* n)}\).

5.3 Johnson Lindenstrauss

Theorem 6. The Johnson-Lindenstrauss (JL) lemma \cite{11} states that for all \(\epsilon \in (0, \frac{1}{2})\), \(\forall x_1, ..., x_n \in l_2\), there exists \(\Pi \in R^{m \times n}, m = O\left(\frac{1}{\epsilon^2} \log(n)\right)\) such that for all \(i, j\), \((1-\epsilon)\|x_i - x_j\|_2 \leq \|\Pi x_i - \Pi x_j\|_2 \leq (1 + \epsilon)\|x_i - x_j\|_2\).

\(f : (x, l_2) \to (x, l_2^m), f(x) = \Pi x\)

5.4 Distributional Johnson Lindenstrauss

Theorem 7. for all \(0 < \epsilon, \delta < \frac{1}{2}\), there exists a distribution \(D_{\epsilon, \delta}\) on matrices \(\Pi \in R^{m \times n}, m = O\left(\frac{1}{\epsilon^2} \log\left(\frac{1}{\delta}\right)\right)\) such that for all \(x \in R^n\), and \(\Pi\) drawn from the distribution \(D_{\epsilon, \delta}\),

\[P(\|\Pi x\|_2 \notin [(1-\epsilon)\|x\|_2, (1+\epsilon)\|x\|_2]) < \delta\]

Now we prove that the distributional Johnson Lindenstrauss proves Johnson Lindenstrauss.

Claim 8. \(DJL \implies JL\)

Proof. Set \(\delta < \frac{1}{(\frac{1}{\epsilon})}\) and look at \(T = \frac{x_i - x_j}{\|x_i - x_j\|_2}\) for \(i < j\). Also note that \(|T| = \binom{N}{2}\). Then

\[P(\Pi \text{ doesn’t have distortion } (1 + \epsilon) \text{ on } X) = P(\exists z \in T \text{ such that } \|\Pi z\|^2_2 - 1 \geq \epsilon)\]

and so by union bound this probability is \(\leq |T| \delta < 1\)

Bibliography.
References


