Communication Complexity 28th Sept 2011 (@ TIFR)

13. Lower bound for norm estimation

Lecturer: Prahladh Harsha *Scribe: Rakesh Venkat*

In this lecture, we will see an application of information theoretic methods to obtain a lower bound on the communication complexity of the Gap-L_∞ problem and generalizations of this problem. This problem results naturally from trying to prove lower bounds on estimating the L_∞ norm on data streams. The references for today's lecture are [BJKS04] and [AJP10].

13.1 The Gap-L_∞ problem

Problem 13.1 (the Gap-$L_\infty(m, n)$ problem)**.** *Parameters: n, m (with $n \gg m$ typically). The instances of the problem are pairs $(x, y) \in \{0, 1, \ldots, m\}^n \times \{0, 1, \ldots, m\}^n$.*

YES*: $\|x - y\|_\infty \geq m$, that is, $\exists i : |x_i - y_i| \geq m$.*

NO*: $\|x - y\|_\infty \leq 1$, that is, $\forall i : |x_i - y_i| \leq 1$.*

As usual, x goes to Alice, y goes to Bob, and they have to differentiate between the YES *and* NO *instances.*

Usually we will drop the (m, n) arguments, and simply refer to it as the Gap-L_∞ problem. Our goal for this lecture is to prove the following theorem:

Theorem 13.2 ([BJKS04])**.** $\mathrm{R}^{priv}_{1/3}(\text{Gap-}L_\infty) = \Omega\left(\frac{n}{m^2}\right)$.

As we observed in Lecture 5, this theorem implies lower bounds on the space requirements for streaming algorithms approximating the L_∞-norm, running on a stream of length n, where every stream element lies in $\{0, \ldots, m\}$.

Corollary 13.3. *Any streaming algorithm (even randomized) that approximates the ∞-norm to within a factor m requires space $\Omega(n/m^2)$.*

13.2 Hardness of approximating Gap-L_∞

We will express the Gap-L_∞ problem as a a disjunction of n copies of a smaller problem, called DIST. Each of these subproblems corresponds to a decision problem on one co-ordinate of the Gap-L_∞ problem. We then proceed by applying techniques similar to the previous lectures on disjointness. First, we define a distribution over the NO instances of the inputs that acts as a fooling set of sorts. If a private-coins protocol for Gap-L_∞ had communication δn for the original problem, then we will zoom-in on one co-ordinate, and show that it must have conveyed at most δ bits of information for DIST on this co-ordinate. But we know that at least some information must have been transmitted for this co-ordinate, since our inputs now come from a 'fooling set', and using randomized communication the protocol is able to distinguish between YES and NO instances with error at most $\leq \frac{1}{2} - \varepsilon$.

The main departure from the disjointness proof will come in the fact that the single co-ordinate DIST problem analysis will require a new property of Hellinger distance, different from the cut-and-paste property one used earlier. This is the Z-Lemma, and we will state and prove it along the way.

13.2.1 The DIST Problem

We first define the $\mathsf{DIST}(m)$ problem as follows:

Problem 13.4 (the $\mathsf{DIST}(m)$ problem)**.** *Instances are integer pairs* $(u, v) \in \{0 \ldots m\} \times \{0 \ldots m\}$*.*

YES*:* $|u - v| \geq m$*.*

NO*:* $|u - v| \leq 1$*.*

As before, we will drop the argument m, if is clear from context. Notice that given a $\mathsf{Gap}\text{-}L_\infty$ instance (x, y), we have

$$\mathsf{Gap}\text{-}L_\infty(x, y) = \bigvee_{i=1}^{n} \mathsf{DIST}(x_i, y_i)$$

13.2.2 The fooling set distribution for DIST and Gap-L_∞

We now define a distribution over questions for the players in the $\mathsf{Gap}\text{-}L_\infty$ problem. Let $\sigma \equiv (T_i, S_i)_{i=1}^n$ be a random variable, where each $\sigma_i = (T_i, S_i)$ taking values uniformly in $\{A, B\} \times \{0, \ldots, m-1\}$. For example, σ may look like $[(A, 0), (B, 4), (B, m-1), \ldots, (A, 7)]$. The distribution on (X, Y) is defined based on the value of σ drawn; we denote the distribution of X, Y conditioned on a given value of σ as (X^σ, Y^σ):

$$X^\sigma Y^\sigma: \text{ If } T_i = A \quad \text{then} \quad \begin{cases} X_i^\sigma \in_R \{S_i, S_i + 1\} \\ Y_i^\sigma = S_i + 1 \end{cases}$$

$$\text{If } T_i = B \quad \text{then} \quad \begin{cases} X_i^\sigma = S_i \\ Y_i^\sigma \in_R \{S_i, S_i + 1\} \end{cases}$$

Note that conditioned on a fixed value of σ, the questions are chosen independently across co-ordinates. Also, this distribution has support only on the NO instances, and acts like a fooling set for the problem.

13.2.3 Part 1: Reducing Gap-L_∞ to DIST

All probabilities in what follows are over the distribution defined above, and private randomness used by the players. Let us start off, as in the previous lecture, with $\Pi(X, Y)$ being the random variable corresponding to the transcript on the (random) questions X, Y (for the $\mathsf{Gap}\text{-}L_\infty$ problem). Suppose that this transcript succeeds with probability $\geq \frac{1}{2} + \varepsilon$, (this is over only the private randomness of the players, and holds for *every* input) and has length $\leq \delta n$. Let $(X^\sigma Y^\sigma)$ be the questions conditioned on σ (in general, superscripting with σ refers to conditioning with respect to σ) . We have the familiar inequality chain:

$$\begin{aligned}
\delta n &\geq H\left[\Pi(X,Y)\right] \\
&\geq H\left[\Pi(X^\sigma Y^\sigma)|\sigma\right] \\
&\geq I\left[X^\sigma Y^\sigma : \Pi(X^\sigma Y^\sigma)|\sigma\right] \\
&\geq \sum_{i\in[n]} I\left[X_i^\sigma Y_i^\sigma : \Pi(X^\sigma Y^\sigma)|\sigma\right]. && \text{[using the chain rule for mutual information]} \\
\implies \delta &\geq \mathop{\mathbb{E}}_{\sigma,k\in[n]}\left[I\left[X_k^\sigma Y_k^\sigma : \Pi|\sigma\right]\right] \\
&= \mathop{\mathbb{E}}_{k}\left[\mathop{\mathbb{E}}_{\sigma_{-k}}\left[\mathop{\mathbb{E}}_{\sigma_k}\left[I\left[X_k^\sigma Y_k^\sigma : \Pi|\sigma_{-k}\sigma_k\right]\right]\right]\right] \\
&\geq \mathop{\mathbb{E}}_{\sigma_k}\left[I\left[X_k^\sigma Y_k^\sigma : \Pi|\sigma_k\right]\right]. && \text{[for some fixing of } k,\sigma_{-k}]
\end{aligned}$$

We show a lower bound on this quantity. As all other co-ordinates are fixed to a NO instance of DIST, the protocol must compute DIST on co-ordinate k correctly with at least the same probability as it computes Gap-L_∞, which is $\frac{1}{2}+\varepsilon$.

Since $\sigma_k = (T_k, S_k)$, either Alice or Bob is active (depending on T_k, with probability 1/2 each. Further, the inactive party is set to a fixed value, which is uniform either over the space $\{0,\ldots,m-1\}$ or $\{1,\ldots,m\}$. Unrolling the above expectation over the values of σ_k gives:

$$\delta \geq \frac{1}{2m}\sum_{s=0}^{m-1} I\left[X_k^{(A,s)} : \Pi\left(X_k^{(A,s)}, s+1\right)\right] + I\left[Y_k^{(B,s)} : \Pi\left(s, Y_k^{(B,s)}\right)\right].$$

Note that the variables $X_k^{(A,s)}$ and $Y_k^{(B,s)}$ are uniform in $\{s, s+1\}$. For simplicity, denote $\Pi_{ab} \equiv \Pi(a,b)$. Applying the mutual-information to Hellinger distance property, we get:

$$\begin{aligned}
\delta &\geq \frac{1}{2m}\sum_{s=0}^{m-1} h^2(\Pi_{s,s+1},\Pi_{s+1,s+1}) + h^2(\Pi_{s,s},\Pi_{s,s+1}) \\
&\geq \frac{1}{4m^2}\left(\sum_{s=0}^{m-1} h(\Pi_{s,s+1},\Pi_{s+1,s+1}) + h(\Pi_{s,s},\Pi_{s,s+1})\right)^2 && \text{[By Cauchy Schwarz inequality]} \\
&\geq \frac{1}{4m^2}\left(\sum_{s=0}^{m-1} h(\Pi_{s,s},\Pi_{s+1,s+1})\right)^2 && \text{[By triangle inequality]} \\
&\geq \frac{1}{4m^2} h^2(\Pi_{00},\Pi_{mm}) && \text{[By triangle inequality]}
\end{aligned}$$

However, Π_{00}, Π_{mm} are possibly close in statistical distance, since both correspond to NO instances. Similarly, Π_{0m}, Π_{m0} could also be close since both are YES instances. So a routine cut and paste yields nothing, and we will need something more to proceed.

13.2.4 Part 2: The Z-lemma, finishing the proof

Lemma 13.5 (Z-Lemma for Hellinger distance)**.** *If* Π *is the transcript for a communication protocol, and let* $x, x^{'} \in \mathcal{X}$*,* $y, y^{'} \in \mathcal{Y}$*. Then we have the following property:*

$$h^2(\Pi_{xy}, \Pi_{x'y'}) \geq \frac{1}{2}\left(h^2(\Pi_{xy}, \Pi_{xy'}) + h^2(\Pi_{x'y}, \Pi_{x'y'})\right).$$

Before seeing the proof of this lemma, let us use it to finish the earlier proof. Applying the Z-Lemma, we have,

$$\begin{aligned}
\delta &\geq \frac{1}{8m^2}\left(h^2(\Pi_{00}, \Pi_{0m}) + h^2(\Pi_{m0}, \Pi_{mm})\right) \\
&\geq \frac{1}{16m^2}\left(\Delta^2(\Pi_{00}, \Pi_{0m}) + \Delta^2(\Pi_{m0}, \Pi_{mm})\right) \quad \text{[Moving from Hellinger to statistical distance]} \\
&\geq \frac{1}{16m^2} \cdot 8\varepsilon^2 \\
&= \frac{\varepsilon^2}{2m^2},
\end{aligned}$$

which gives us the final result that $\mathrm{R}^{\mathrm{priv}}_{1/3}(\mathsf{Gap}\text{-}L_\infty) = \Omega(\frac{n}{m^2})$ (since the total communication was δn).

Now, we prove the Z-lemma.

Proof of Z-Lemma 13.5. Let $\Pi(x, y)$ be the randomized transcript on questions $X = x, Y = y$. We know that there are functions q_A, q_B such that $\Pr[\Pi(x, y) = \tau] = q_A(\tau, x)q_B(\tau, y)$. Using this, we can write:

$$\begin{aligned}
&\frac{1}{2}(1 - h^2(\Pi_{xy}, \Pi_{x'y}) + \frac{1}{2}(1 - h^2(\Pi_{xy'}, \Pi_{x'y'})) \\
=&\frac{1}{2}\sum_\tau \sqrt{q_A(\tau, x)q_B(\tau, y)q_A(\tau, x')q_B(\tau, y)} + \sqrt{q_A(\tau, x)q_B(\tau, y')q_A(\tau, x')q_B(\tau, y')} \\
=&\sum_\tau \frac{q_B(\tau, y) + q_B(\tau, y')}{2}\sqrt{q_A(\tau, x)q_A(\tau, x')} \\
\geq&\sum_\tau \sqrt{q_B(\tau, y)q_B(\tau, y')}\sqrt{q_A(\tau, x)q_A(\tau, x')} \qquad \text{[AM-GM inequality]} \\
=&1 - h^2(\Pi_{xy}, \Pi_{x'y'}).
\end{aligned}$$

□

13.3 Generalization using Poincaré inequalities

Let g be a *distance function*, i.e. $g : \mathcal{X} \times \mathcal{X} \to \{0, 1\}$ satisfies $\forall x \in \mathcal{X} : g(x, x) = 0$ and $\forall x, y \in \mathcal{X} \times \mathcal{X} : g(x, y) = g(y, x)$. The DIST function is an example of such a distance function.

The problem we consider, is to lower bound the communication complexity of $f : \mathcal{X}^n \times \mathcal{X}^n \to \{0, 1\}$, the disjunction of n copies of g:

$$f(x,y) \triangleq \bigvee_{i=1}^{n} g(x_i, y_i)$$

Andoni, Jayram and Patrascu [AJP10] show that the proof method used for Gap-L_∞ can be generalized to find a lower bound on $\mathrm{R}^{\mathrm{priv}}_{1/3}(f)$, as long as g satisfies a *Poincaré inequality*.

A Poincaré inequality for g is stated with respect to two distributions η_0 and η_1 satisfying: $\mathrm{supp}(\eta_0) \subseteq g^{-1}(0)$ and $\mathrm{supp}(\eta_1) \subseteq g^{-1}(1)$. g is said to satisfy a Poincaré inequality with respect to these distributions, if for some $\alpha \in \mathbb{R}^+$ and $\forall \rho : \mathcal{X} \to \mathbb{S}_+$:

$$\mathop{\mathbb{E}}_{x,y\sim\eta_0} \|\rho(x) - \rho(y)\|_2^2 \geq \alpha \mathop{\mathbb{E}}_{x,y\sim\eta_1} \|\rho(x) - \rho(y)\|_2^2.$$

Poincaré inequalities of this form arise in many places. Notable examples are expanders and Boolean function analysis.

Example 13.6. *Consider the Boolean hypercube H on $\{0,1\}^n$, and define the function $g : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ as follows: $g(x,y) = 0$, if $\|x-y\|_1 \leq 1$ and $g(x,y) = 1$ if $\|x-y\|_1 \geq n$.*

Set η_0 to be uniform over the pairs (x,y) with $\|x-y\|_1 = 1$, and η_1 to be uniform over pairs $(x,\bar{x})$. Then we can show that g satisfies a Poincaré inequality for any mapping $\rho : H \to \mathbb{S}_+$:

$$\mathop{\mathbb{E}}_{(u,v)\sim\eta_0} [\|\rho(u) - \rho(v)\|_2^2] \geq \frac{1}{n} \mathop{\mathbb{E}}_{(u,v)\sim\eta_1} [\|\rho(u) - \rho(v)\|_2^2].$$

Example 13.7. *The DIST function on $\{0,\ldots,m\} \times \{0,\ldots,m\} \to \{0,1\}$ satisfies a Poincaré inequality, with η_0 be the uniform distribution over pairs $(s, s+1)$ with $s \in_R \{0,\ldots,m-1\}$, and η_1 supported completely on the single pair $(0,m)$:*

$$\mathop{\mathbb{E}}_{u\in_R\{0,\ldots,m-1\}} [\|\rho(u) - \rho(u+1)\|_2^2] \geq \frac{1}{m^2} \|\rho(0) - \rho(m)\|_2^2.$$

This has been implicitly shown in the proof of the previous section, using Cauchy-Schwarz and triangle inequalities.

Let us sketch the general proof method, that runs along the same lines as above. First, we define the distribution over the questions depending on $\sigma \in_R (\{A,B\} \times \eta_0)^n$. Note that the second component in every co-ordinate is a NO instance drawn from η_0. Let $\sigma_i = (T_i, (U,V))$. Then set the questions as follows:

$$\text{If } T_i = A \quad \text{then} \begin{cases} X_i^\sigma \in_R \{u,v\} \\ Y_i^\sigma = v \end{cases}$$

$$\text{If } T_i = B \quad \text{then} \begin{cases} X_i^\sigma = u \\ Y_i^\sigma \in_R \{u,v\}. \end{cases}$$

Let Π be the transcript of a private coins protocol for f that succeeds with probability $\frac{1}{2}+\varepsilon$ and has length $\leq \delta n$. Again, following chain of inequalities as in the Gap-L_∞ problem, we arrive at the point where, for some fixed σ_{-k} and k, we have that:

$$\begin{aligned}
\delta &\geq \mathop{\mathbb{E}}_{\sigma_k}\left[I[X_k^\sigma Y_k^\sigma : \Pi]\right] \\
&\geq \frac{1}{2}\mathop{\mathbb{E}}_{(u,v)\sim\eta_0}\left[I[X^{uv} : \Pi(X^{uv}, v)] + I[Y^{uv} : \Pi(u, Y^{uv})]\right] \\
&\geq \frac{1}{2}\mathop{\mathbb{E}}_{(u,v)\sim\eta_0}\left[h^2(\Pi_{uv}, \Pi_{vv}) + h^2(\Pi_{uu}, \Pi_{uv})\right] \\
&\geq \frac{1}{4}\mathop{\mathbb{E}}_{(u,v)\sim\eta_0}\left[h^2(\Pi_{uu}, \Pi_{vv})\right] && \text{[Cauchy Schwarz + $\triangle$ inequality]} \\
&\geq \frac{\alpha}{4}\mathop{\mathbb{E}}_{(u,v)\sim\eta_1}[h^2(\Pi_{uu}, \Pi_{vv})] && \text{[Poincaré inequality]} \\
&\geq \frac{\alpha}{8}\mathop{\mathbb{E}}_{(u,v)\sim\eta_1}[h^2(\Pi_{uv}, \Pi_{uu}) + h^2(\Pi_{vu}, \Pi_{vv})] && \text{[the Z-lemma]} \\
&\geq \frac{\alpha}{16}\mathop{\mathbb{E}}_{(u,v)\sim\eta_1}[\Delta^2(\Pi_{uv}, \Pi_{uu}) + \Delta^2(\Pi_{vu}, \Pi_{vv})] && \text{[moving to statistical distance]}
\end{aligned}$$

From reflexivity, we know that $g(u,u) = g(v,v) = 0$, but $g(u,v) = g(v,u) = 1$ since $(u,v) \sim \eta_1$. Thus, $\Delta(\Pi_{uv}, \Pi_{uu}) \geq 2\varepsilon$ and $\Delta(\Pi_{vu}, \Pi_{vv}) \geq 2\varepsilon$. Plugging this in gives us our bound:

$$\delta \geq \frac{\alpha\varepsilon^2}{2}.$$

This gives us that $\mathrm{R}_{1/3}^{\text{priv}}(f) = \Omega(\alpha n)$.

References

[AJP10] ALEXANDR ANDONI, T. S. JAYRAM, and MIHAI PATRASCU. *Lower bounds for edit distance and product metrics via Poincaré-type inequalities*. In *Proc. 21th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 184–192. 2010.

[BJKS04] ZIV BAR-YOSSEF, T. S. JAYRAM, RAVI KUMAR, and D. SIVAKUMAR. *An information statistics approach to data stream and communication complexity.* J. Computer and System Sciences, 68(4):702–732, June 2004. (Preliminary Version in *43rd FOCS*, 2002). doi:10.1016/j.jcss.2003.11.006.