CAS CS 591 B: Communication Complexity

Prof. Mark Bun | Fall 2019

Lecture Notes 13:

Unbounded error communication

Reading.

- Paturi-Simon, "Probabilistic Communication Complexity"

We'll begin studying a very strong model of probabilistic communication complexity called the "unbounded error model." In this model we are interested in computing a function with any advantage over random guessing. It will be convenient for us to change the range of our Boolean functions to $\{-1, 1\}$ instead of $\{0, 1\}$.

Definition 1. The unbounded error communication complexity of a function $f : X \times Y \to \{-1, 1\}$, denoted $\mathbf{UPP^{cc}}(f)$, is the minimum length of a private coin protocol Π such that for every (x, y)

$$\Pr[\Pi(x, y) = f(x, y)] > 1/2 \iff \operatorname{sgn} \mathbb{E}[\Pi(x, y)] = f(x, y).$$

This definition should be compared with that of the $\mathbf{PP^{cc}}$ model which we defined earlier when discussing discrepancy. Recall that the $\mathbf{PP^{cc}}$ cost of a protocol charges for both its length and for $\log(1/\delta)$, where δ is its advantage over random guessing on a worst-case input. One way to think about this additional charge of $\log(1/\delta)$ is as a proxy for charging for the amount of randomness used in the protocol. Moreover, by Newman's theorem, the definition of $\mathbf{PP^{cc}}$ changes by only additive log terms depending on whether we allow public randomness. The $\mathbf{UPP^{cc}}$ model, on the other hand, is only interesting in the absence of public randomness. If we were to permit the use of public randomness, then every function would have an $O(1)$ cost protocol. Alice could just send a bit indicating whether her string x is equal to the first n bits of the public random string. If so, then Bob can compute $f(x, y)$, and otherwise, he can output a random guess. This protocol has advantage 2^{-n}.

Even if we only allow for private coins, many functions have efficient $\mathbf{UPP^{cc}}$ protocols.

Example 2. $\mathbf{UPP^{cc}}(\mathrm{EQ}_n), \mathbf{UPP^{cc}}(\mathrm{GT}_n) = O(1)$ and $\mathbf{UPP^{cc}}(\mathrm{GH}_n), \mathbf{UPP^{cc}}(\mathrm{DISJ}_n) = O(\log n)$.

We've seen protocols for the latter two before. The easiest way to see that there are constant communication protocols for Equality and Greater-Than is by studying equivalent characterizations of $\mathbf{UPP^{cc}}$.

1 Sign Rank

$\mathbf{UPP^{cc}}$ has an extremely clean matrix analytic characterization. Recall the log rank conjecture, which asserts that deterministic communication complexity of f is characterized up to polynomial factors by $\log \operatorname{rank} M_f$. The analog of this conjecture for $\mathbf{UPP^{cc}}$ complexity is true in the the sharpest possible way.

Definition 3. The sign rank of a matrix M with entries in $\{-1,1\}$, denoted $\operatorname{rank}_\pm M$, is the minimum rank of a real-valued matrix R such that $\operatorname{sgn} R[x,y] = M[x,y]$ for every (x,y).

Theorem 4. *For a function $f : X \times Y \to \{-1,1\}$, let $S_f \in \mathbb{R}^{|X|\times|Y|}$ be its sign matrix $S_f[x,y] = f(x,y)$. Then*

$$\mathbf{UPP^{cc}}(f) = \log \operatorname{rank}_\pm(S_f) \pm O(1).$$

Proof. We first show that $\log \operatorname{rank}_\pm(S_f) \leq \mathbf{UPP^{cc}}(f) + O(1)$. Suppose Π is a $\mathbf{UPP^{cc}}$ protocol computing f with communication cost c. Define the matrix $R[x,y] = \mathbb{E}[\Pi(x,y)]$. Then $\operatorname{sgn} R[x,y] = S_f[x,y]$, so we just need to show that $\operatorname{rank} R \leq 2^c$. Write

$$\mathbb{E}[\Pi(x,y)] = \sum_{\ell:\Pi \text{ outputs } 1 \text{ at leaf } \ell} \Pr[\Pi(x,y) \text{ reaches } \ell] - \sum_{\ell:\Pi \text{ outputs } -1 \text{ at leaf } \ell} \Pr[\Pi(x,y) \text{ reaches } \ell],$$

and for each such leaf ℓ define the matrix $R_\ell[x,y] = \Pr[\Pi(x,y) \text{ reaches } \ell] = p_\ell(x) \cdot q_\ell(y)$ for some functions p_ℓ, q_ℓ, using the fact that Π is a private coin protocol. This implies that every R_ℓ is a rank-1 matrix, and hence

$$R = \sum_{\ell:\Pi \text{ outputs } 1 \text{ at leaf } \ell} R_\ell - \sum_{\ell:\Pi \text{ outputs } -1 \text{ at leaf } \ell} R_\ell$$

has rank at most 2^c.

For the other direction, we want to show that if $\operatorname{rank}_\pm(S_f) = r$, then we can construct a $\mathbf{UPP^{cc}}$ protocol for f with cost $\log r + O(1)$. The condition $\operatorname{rank}_\pm(S_f) = r$ is equivalent to the existence of vectors $\{u_x \in \mathbb{R}^r : x \in X\}$ and $\{v_y \in \mathbb{R}^r : y \in Y\}$ such that $\operatorname{sgn}\langle u_x, v_y\rangle = f(x,y)$ for every x,y. Assume without loss of generality that the vectors are normalized so that every $\|u_x\|_1 = 1$ and $\|v_y\|_1 = 1$. Hence each $|(u_x)_1|, \ldots, |(u_x)_r|$ represents a probability distribution over $[r]$ (and similarly for v_y). This suggests the following cost $\log r + O(1)$ protocol for computing f:

1. Alice samples $i \in [r]$ with probability $|(u_x)_i|$. She sends i together with $a = \operatorname{sgn}(u_x)_i$ to Bob.
2. Bob samples $b \in \{-1,1\}$ such that $\mathbb{E}[b] = (v_y)_i$ and outputs ab.

We compute the advantage of this protocol as follows:

$$\begin{aligned}
\mathbb{E}[\Pi(x,y)] &= \sum_{i=1}^r |(u_x)_i| \mathbb{E}[ab|i] \\
&= \sum_{i=1}^r (u_x)_i \mathbb{E}[b|i] \\
&= \sum_{i=1}^r (u_x)_i (v_y)_i \\
&= \langle u_x, v_y \rangle
\end{aligned}$$

whose sign agrees with $f(x,y)$. $\square$

2 Arrangements of Hyperplanes

A hyperplane through the origin in $\mathbb{R}^d$ is specified by its normal vector $v \in \mathbb{R}^d$ and consists of the points $u \in \mathbb{R}^d$ such that $\langle u, v\rangle = 0$. It divides $\mathbb{R}^d$ into two halfspaces: a "positive" side consisting of points u with $\langle u, v\rangle > 0$ and a "negative side" where $\langle u, v\rangle < 0$.

A collection of hyperplanes $V = \{v_y \in \mathbb{R}^d : y \in Y\}$ *realizes* a matrix $\{-1, 1\} \in \mathbb{R}^{|X|\times|Y|}$ if for every row vector $m_x \in \{-1, 1\}^{|Y|}$ of M there exists a point $u_x \in \mathbb{R}^d$ such that

$$(m_x)_y = \operatorname{sgn}\langle u_x, v_y\rangle$$

for every $y \in Y$. One way to think about what's going on is that a set of hyperplanes divides $\mathbb{R}^d$ into (at most) $2^{|Y|}$ regions. To each region, we can assign a $|Y|$-bit string indicating which side of each hyperplane that region lies in. The collection of hyperplanes realizes M if every row of M appears as one of these strings.

The sign rank of a matrix M is the minimum dimension d over which there is a collection of hyperplanes realizing M. In one direction, suppose M has sign rank d. Then there exist vectors $\{u_x \in \mathbb{R}^r : x \in X\}$ and $\{v_y \in \mathbb{R}^r : y \in Y\}$ such that $\operatorname{sgn}\langle u_x, v_y\rangle = M[x, y]$. So we can take the u_x's to be the points and the v_y's to the the hyperplanes. Conversely, a collection of hyperplanes realizing M induces the sets of of vectors required to show that M has sign rank d.

This "geometric" interpretation of sign rank is useful for several reasons. One is that it gives rise to machine learning algorithms; in fact, the fastest known algorithm for PAC learning DNF is based on a sign rank upper bound. To see the connection, suppose we have a collection of functions $\{f_y : X \to \{-1, 1\}\}$ indexed by $y \in Y$. Let $(x_1, f_{y^*}(x_1)), \ldots, (x_k, f_{y^*}(x_k))$ be a sequence of labeled samples from X. We'd like to be able to learn (an approximation to) the function f_{y*}. If the matrix $M[x, y] = f_y(x)$ has sign rank r, then the problem of finding f_{y*} reduces to the problem of finding an r-dimensional hyperplane v which separates the points u_{x_i} for which $f_{y^*}(x_i) < 0$ from the points u_{x_i} for which $f_{y^*}(x_i) > 0$. This can be done in $\operatorname{poly}(r)$ time using linear programming.

The other reason is that it lets us design the very efficient communication protocols mentioned at the beginning of the lecture. For instance, to design an arrangement of hyperplanes realizing the communication matrix of EQ_n, we can take the points u_x to be evenly spaced along the unit circle in $\mathbb{R}^2$ and the hyperplanes to be "just below" the tangents to the circle at these points. The dimension of this arrangement is 2, hence the sign rank of the communication matrix is 2 and there is an $O(1)$-communication protocol for this problem.

Next time, we'll talk about how to prove strong lower bounds on sign rank and $\mathbf{UPP^{cc}}$ communication.