University of California, Los Angeles
CS 289A Communication Complexity

Instructor: Alexander Sherstov
Scribe: Mohan Yang
Date: January 18, 2012

LECTURE
3

Matrix Rank in Communication Complexity

This lecture focuses on proving communication lower bounds using matrix rank. Similar to fooling sets and rectangle size bounds, the matrix rank technique also gives a lower bound on the number of monochromatic rectangles in any partition of $X \times Y$ but it does so in an algebraic way[1]. This makes algebraic tools available for proving communication lower bounds. We begin by solving the problem about matrix rank from last lecture. We then present the rank technique and reprove tight lower bounds on the communication complexity of equality, greater-than, disjointness, and inner product. Finally, we look into the comparative strength of fooling sets, rectangle size bounds, and matrix rank as techniques for proving communication lower bounds.

3.1 Rank of Boolean Matrix over Different Fields

CLAIM 3.1. *Given fields $\mathbb{K} \supset \mathbb{F}$ and a matrix $M \in \mathbb{F}^{m \times n}$, one has $rank_{\mathbb{F}}(M) = rank_{\mathbb{K}}(M)$.*

Proof. Since rows linearly dependent over $\mathbb{F}$ remain linearly dependent over $\mathbb{K}$, we have $\text{rank}_{\mathbb{K}}(M) \leq \text{rank}_{\mathbb{F}}(M)$. For the other direction, consider the matrix M' (shown in Figure 3.1) in row echelon form obtained by performing elementary row and column operations over $\mathbb{F}$ on M.

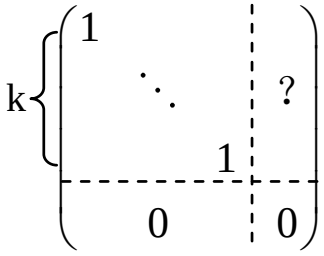

FIGURE 3.1: The matrix M'.

It is clear that $\text{rank}_{\mathbb{K}}(M') = \text{rank}_{\mathbb{K}}(M)$, and $\text{rank}_{\mathbb{F}}(M') = \text{rank}_{\mathbb{F}}(M)$. But $\text{rank}_{\mathbb{F}}(M') = \text{rank}_{\mathbb{K}}(M') = k$, so $\text{rank}_{\mathbb{F}}(M) = \text{rank}_{\mathbb{K}}(M)$. □

We can now show that the rank of a Boolean matrix over the reals is at least as large as its rank of any other field.

LEMMA 3.2. *Given a matrix $M \in \{0,1\}^{m\times n}$, $rank_{\mathbb{R}}(M) \geq rank_{\mathbb{F}}(M)$ for any field $\mathbb{F}$.*

In Remark 3.8 below, we will see a Boolean matrix with an exponential gap between its rank over the reals and over $\mathbb{F}$ =GF(2).

Proof of the lemma. By Claim 3.1, $\text{rank}_{\mathbb{Q}}(M) = \text{rank}_{\mathbb{R}}(M)$. It remains to show that if rows $v_1, \cdots, v_k$ are linearly dependent over $\mathbb{Q}$, they are also linearly dependent over $\mathbb{F}$.

Let $\sum_{i=1}^{k} \lambda_i v_i = 0$ be a linear dependence, where $\lambda_1, \cdots, \lambda_k \in \mathbb{Q}$ are not all zero. W.l.o.g, $\lambda_1, \cdots, \lambda_k \in \mathbb{Z}$ and $\gcd(\lambda_1, \cdots, \lambda_k) = 1$. Then $\sum_{i=1}^{k} \lambda_i v_i = 0$ also holds over $\mathbb{F}$, where we identify $\lambda_i \in \mathbb{Z}$ with the field element $1 + 1 + \cdots + 1 \in \mathbb{F}$ (λ_i times). It remains to show that $\lambda_1, \ldots, \lambda_k$ are not all zero as elements of $\mathbb{F}$. If $\mathbb{F}$ has characteristic of 0, then this is immediate. For fields of nonzero characteristic, $\lambda_1 = \cdots = \lambda_k = 0$ means that $\lambda_1, \cdots, \lambda_k$ are all integer multiples of ch F, which contradicts the assumption that $\gcd(\lambda_1, \cdots, \lambda_k) = 1$. □

3.2 Rank Lower Bound

THEOREM 3.3 (Mehlhorn and Schmidt [2]). *For any function $f : X \times Y \to \{0,1\}$, $D(f) \geq \log_2(2 rank_{\mathbb{F}}(M_f) - 1)$ over any field $\mathbb{F}$.*

Proof. Fix a partition $X \times Y = \bigcup_{i=1}^{2^c} R_i$, where $c = D(f)$. Let $\mathcal{R}_0$ and $\mathcal{R}_1$ be the sets of all 0-monochromatic rectangles and 1-monochromatic rectangles, respectively, among $R_1, R_2, \ldots, R_{2^c}$. For a rectangle R, define

$$M_R(x,y) = \begin{cases} 1 & \text{if } (x,y) \in R; \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$M_f = \sum_{R\in\mathcal{R}_1} M_R, \qquad M_{\text{not}(f)} = \sum_{R\in\mathcal{R}_0} M_R. \tag{3.1}$$

By the properties of the rank, we have

$$\text{rank}_{\mathbb{F}}(M_f) \leq \sum_{R\in\mathcal{R}_1} \text{rank}_{\mathbb{F}}(M_R), \qquad \text{rank}_{\mathbb{F}}(M_{\text{not}(f)}) \leq \sum_{R\in\mathcal{R}_0} \text{rank}_{\mathbb{F}}(M_R). \tag{3.2}$$

The rank of each M_R is at most 1 over $\mathbb{F}$, so $\text{rank}_{\mathbb{F}}(M_f) \leq |\mathcal{R}_1|$ and $\text{rank}_{\mathbb{F}}(M_{\text{not}(f)}) \leq |\mathcal{R}_0|$. Adding the two inequalities, we have

$$\text{rank}_{\mathbb{F}}(M_f) + \text{rank}_{\mathbb{F}}(M_{\text{not}(f)}) \leq |\mathcal{R}_1| + |\mathcal{R}_0| = 2^c. \tag{3.3}$$

Observe that the ranks of M_f and $M_{\text{not}(f)}$ differ by at most 1, since $M_{\text{not}(f)} = J - M_f$ where J is the all-ones matrix (whose rank is 1 over $\mathbb{F}$). Therefore, we have $2^c \geq 2\text{rank}_{\mathbb{F}}(M_f) - 1$, or equivalently $D(f) \geq \log_2(2\text{rank}_{\mathbb{F}}(M_f) - 1)$ for any field $\mathbb{F}$. □

In previous lectures, we proved (using fooling sets and rectangle size bounds) that the deterministic communication complexity of each of the functions EQ_n, GT_n, DISJ_n, and IP_n is $n + 1$. We now reprove these facts using the rank technique.

EXAMPLE 3.4. The communication complexity of the equality function $\mathrm{EQ}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is $n+1$. Indeed, $M_{\mathrm{EQ}_n} = I_{2^n}$, the identity matrix of order 2^n, and thus $\mathrm{rank}(M_{\mathrm{EQ}_n}) = 2^n$ and $D(\mathrm{EQ}_n) \geq \lceil \log_2(2 \cdot 2^n - 1) \rceil = n+1$.

EXAMPLE 3.5. The communication complexity of the greater-than function $\mathrm{GT}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is $n+1$. To see this, note that the characteristic matrix M_{GT_n} is a $2^n \times 2^n$ lower-triangular matrix, thus $\mathrm{rank}(M_{\mathrm{GT}_n}) = 2^n$, and $D(\mathrm{GT}_n) \geq \lceil \log_2(2 \cdot 2^n - 1) \rceil = n+1$.

EXAMPLE 3.6. We now show that the communication complexity of the disjointness function $\mathrm{DISJ}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is $n+1$. It can be proved by induction that

$$M_{\mathrm{DISJ}_1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad M_{\mathrm{DISJ}_n} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{\otimes n}$$

where $\otimes$ is the tensor product defined as

$$A \otimes B = \begin{pmatrix} A_{11}B & \cdots & A_{1n}B \\ \vdots & \ddots & \vdots \\ A_{n1}B & \cdots & A_{nn}B \end{pmatrix}$$

where n is the order of A. By the property of the tensor product, we have $\mathrm{rank}(A \otimes B) = \mathrm{rank}(A) \cdot \mathrm{rank}(B)$ and thus $\mathrm{rank}(M_{\mathrm{DISJ}_n}) = (\mathrm{rank}(M_{\mathrm{DISJ}_1}))^n = 2^n$. As a result, $D(\mathrm{DISJ}_n) \geq \lceil \log_2(2 \cdot 2^n - 1) \rceil = n+1$.

EXAMPLE 3.7. We now prove that the communication complexity of the inner product function $\mathrm{IP}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is $n+1$. Define $\widetilde{\mathrm{IP}}_n(x,y) = (-1)^{\langle x,y \rangle}$. Then we have

$$M_{\widetilde{\mathrm{IP}}_1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad M_{\widetilde{\mathrm{IP}}_n} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{\otimes n}$$

by induction. In particular, $M_{\widetilde{\mathrm{IP}}_n}$ has full rank. Since $M_{\widetilde{\mathrm{IP}}_n} = 2M_{\mathrm{IP}_n} - J$, where J is the all-ones matrix, we have $\mathrm{rank}(M_{\mathrm{IP}_n}) \geq \mathrm{rank}(M_{\widetilde{\mathrm{IP}}_n}) - 1 = 2^n - 1$. Thus $D(\mathrm{IP}_n) \geq \lceil \log_2(2 \cdot (2^n - 1) - 1) \rceil = n+1$.

REMARK 3.8. Although M_{IP_n} has almost full rank over $\mathbb{R}$, its rank is at most n over $\mathbb{GF}(2)$. Inded, letting $\langle x,y \rangle$ denote the inner product over GF(2),

$$M_{\mathrm{IP}_n} = [\langle x,y \rangle]_{x,y \in \{0,1\}^n} = \left[\sum_{i=1}^{n} x_i y_i \right] = \sum_{i=1}^{n} [x_i y_i], \tag{3.4}$$

so that M_{IP_n} is the sum of n matrices of rank 1 over GF(2). Hence, $\mathrm{rank}_{\mathbb{GF}(2)}(M_{\mathrm{IP}_n}) \leq n$.

3.3 Other Uses of Matrix Rank

PROPOSITION 3.9. *Let $f : X \times Y \to \{0,1\}$ be a Boolean function. If all the rows of M_f are distinct, then $D(f) \geq \log\log|X|$.*

Proof. Let $r = \text{rank}_{\mathbb{GF}(2)}(M_f)$. Pick r rows $v_1, \cdots, v_r$ that span the row space of M_f. As all the rows of M_f are distinct, we have

$$|X| \leq 2^r = 2^{\text{rank}_{\mathbb{GF}(2)}(M_f)}. \tag{3.5}$$

Thus,

$$D(f) \geq \log_2 \text{rank}_{\mathbb{GF}(2)}(M_f) \geq \log_2 \log_2 |X|. \tag{3.6}$$

□

PROPOSITION 3.10. *For any $f : X \times Y \to \{0,1\}$, one has $D(f) \leq rank(M_f) + 1$.*

Proof. Let $r = \text{rank}_{\mathbb{GF}(2)}(M_f)$. Pick r rows $v_1, \cdots, v_r$ that span the row space of M_f. Consider the following communication protocol:

1) Alice writes her row as a linear combination of the rows, and sends these r bits to Bob;

2) Bob responses with 1 bit.

The above protocol requires $(r+1)$ bits in total. Thus,

$$D(f) \leq \text{rank}_{\mathbb{GF}(2)}(M_f) + 1 \leq \text{rank}(M_f) + 1, \tag{3.7}$$

where we used Lemma 3.2 in the last inequality. □

3.4 Comparison with Fooling Sets and the Rectangle Size Bound

The scopes of applicability of these lower bound techniques are shown in Figure 3.2. That fooling sets are a special case of rectangle size bounds was shown in previous lectures. Proposition 3.11 shows the inclusion relation between the rank technique and the fooling sets technique.

PROPOSITION 3.11. *Let $f : X \times Y \to \{0,1\}$ be a Boolean function. If f has a fooling set S, then $rank_{\mathbb{F}}(M_f) \geq \sqrt{|S|}$ over any field $\mathbb{F}$. In particular, if the fooling set technique gives a lower bound of b bits for $D(f)$, then the rank technique will give a lower bound of at least $b/2$ bits for $D(f)$.*

Proof. It follows from the definition of a fooling set that if $(x,y),(x',y') \in S$, then $x' \neq x$ and $y' \neq y$. W.l.o.g, assume $f(S) = 1$. Let $S = \{(x_1,y_1), \cdots, (x_n,y_n)\}$ be the elements of the fooling set. Consider the submatrix of M_f that corresponds to the rows $x_1, \ldots, x_n$ and columns $y_1, \ldots, y_n$:

$$A = \begin{array}{c} \\ x_1 \\ \vdots \\ x_n \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & \cdots & y_n \end{array} \\ \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \end{array}.$$

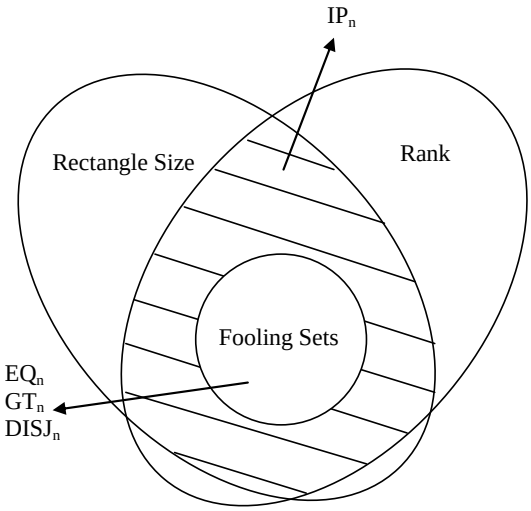

FIGURE 3.2: Scope of applicability of the three lower bound methods.

From the definition of fooling set, we have $A \odot A^T = I$ where $\odot$ is the entrywise product, defined as

$$B \odot C = \begin{pmatrix} B_{11}C_{11} & \cdots & B_{1n}C_{1n} \\ \vdots & \ddots & \vdots \\ B_{n1}C_{n1} & \cdots & B_{nn}C_{nn} \end{pmatrix}.$$

As $B \odot C$ is a submatrix of $B \otimes C$, we have $\text{rank}(B \odot C) \leq \text{rank}(B \otimes C)$. Thus,

$$|S| = \text{rank}_{\mathbb{F}}(I) = \text{rank}_{\mathbb{F}}(A \odot A^T) \leq \text{rank}_{\mathbb{F}}(A \otimes A^T) = (\text{rank}_{\mathbb{F}}(A))^2 \leq (\text{rank}_{\mathbb{F}}(M_f))^2. \quad (3.8)$$

□

EXAMPLE 3.12. For function $\text{IP}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, Remark 3.8 shows that $\text{rank}_{\mathbb{GF}(2)}(M_{\text{IP}_n}) \leq n$. Thus, the size of any fooling set for IP_n is at most n^2, yielding a communication complexity lower bound of only $\Omega(\log n)$. However, as we have seen, the rank technique and rectangle size technique both give a lower bound of $\Omega(n)$. So IP_n lies in the shaded area of Figure 3.2.

The other three functions, EQ_n, GT_n and DISJ_n, are all in the area covered by the "Fooling Sets" technique. We will prove that most matrices lie in the shaded area of Figure 3.2.

THEOREM 3.13. *Let $M \in \{0,1\}^{2^n \times 2^n}$ be a random matrix, then w.v.h.p, M has no fooling sets of size $10n$.*

Proof. Let $r = 10n$. For an ordered collection of columns $S = (s_1, s_2, \ldots, s_r)$ and an ordered collection of rows $T = (t_1, t_2, \ldots, t_r)$, let $X_{S,T} \in \{0,1\}$ be defined by

$$X_{S,T} = \begin{cases} 1 & \text{if } \{(s_1,t_1), \ldots, (s_r,t_r)\} \text{ is a fooling set;} \\ 0 & \text{otherwise.} \end{cases}$$

Then, for $|S| = |T| = r$ and a random choice of M,

$$\Pr[X_{S,T} = 1] = 2^{-r} \left(\frac{3}{4}\right)^{\binom{r}{2}} \cdot 2. \tag{3.9}$$

Thus,

$$\begin{aligned} \Pr[\exists \text{ a fooling set of size } r] &\leq \tbinom{2^n}{r}\tbinom{2^n}{r} \cdot \Pr[X_{S,T} = 1] \\ &\leq 2^{2nr - r + 1 - \log_2\left(\frac{4}{3}\right)\binom{r}{2}}. \end{aligned} \tag{3.10}$$

For $r = 10n$, this probability is extremely small: $\Pr[\exists \text{ a fooling set of size } 10n] \leq 2^{-\Omega(n^2)}$. □

In preparation for next lecture, think about the following problem.

PROBLEM 3.14. Prove that the rank technique and rectangle size technique work well for random matrices.

References

[1] E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, 2nd edition, 2006.

[2] K. Mehlhorn and E. Schmidt. Las vegas is better than determinism in vlsi and distributed computing. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing*, pages 330–337. ACM, 1982.