6.S897: Algebra and Computation April 10, 2015

Lecture 17: Lower bounds for multi-point polynomial evaluation

Lecturer: Madhu Sudan *Scribe: Saeed Mehraban*

1 Introduction

In this lecture we look into the lower bound problem for the multipoint polynomial evaluation problem. We specially focus on the following problem:

$$f(x_1, x_2, \ldots, x_n) = (x_1^r, x_2^r, \ldots, x_n^r)$$

And we show that any algebraic circuit for the above map requires size $\Omega(n \log(r))$. There are two proofs for this fact. The first one, due to Strassen, uses techniques in algebraic geometry, namely the Bezout's theorem and the second one, according to Smolensky uses simple combinatorics and elementary algebra. In this lecture we mainly focus on Strassen's proof.

2 Strassen's Proof

We work with and algebraic closed field $\mathbb{K}$, and thereby, any non-constant polynomial over the field contains a root in the field. Strassen views an algebraic circuit as a collection of polynomial equations of degree at most 2. A multiplication gate $(x, y) \to x * y = z$ corresponds to the polynomial equation $z - x * y = 0$, and an addition gate $(x, y) \to x + y = z$ is the degree one polynomial equation $z - x - y = 0$. Therefore, we work with three types of variables, the inputs $(x_1, x_2, \ldots, x_n)$, intermediate variables $(y_1, y_2, \ldots, y_n)$ and the outputs $(z_1, z_2, \ldots, z_n)$. Each gate introduces an new variable in this setting. Therefore, a circuit of size s corresponds to a set of polynomial equation:

$$\begin{gathered} p_1(\tilde{x}, \tilde{y}, \tilde{z}) \\ p_1(\tilde{x}, \tilde{y}, \tilde{z}) \\ \vdots \\ p_s(\tilde{x}, \tilde{y}, \tilde{z}) \end{gathered}$$

Thereby, computation in the circuit corresponds to finding the set of common roots in these polynomials. In order to prove a circuit lower bound for the discussed map $(x_1, x_2, \ldots, x_n) \to (x_1^r, x_2^r, \ldots, x_n^r)$, let us consider the special case where all the outputs are ones.

Claim 1. *Define $V = \{(\alpha, \beta) \in \mathbb{K}^{n+m} | p_j(\alpha, \beta, 1^n) = 0, j \in [s]\}$, as the set of common zeros $(\tilde{x}, \tilde{y}) = (\alpha, beta)$, when $\tilde{z} = (1, 1, \ldots, 1)$. If the polynomials are according to the circuit computing f, then $|V| = r^n$*

Proof. In order to see this, let X be the set of inputs $(x_1, x_2, \ldots, x_n)$ mapped to $(x_1^r, x_2^r, \ldots, x_n^r) = (1, 1, \ldots, 1)$. If ω is the principal r'th root of unity, then $X = \{\omega, \omega^2, \ldots, \omega^r = 1\}$. So we restricted the variables to a case where $(x_1, x_2, \ldots, x_n) \in X^n$, where X_n has cardinality r^n. □

Now consider the following fact from Algebraic Geometry:

Theorem 1. *(Classical Bezout's Theorem) if $p_1, p_2, \ldots, p_s$ are polynomials in $\mathbb{K}(x_1, x_2, \ldots, x_n)$, if $\mathbb{K}$ is an algebraically closed field, then the number of common zeros for these polynomials is upper-bounded by the product of the degrees $\prod_{j\in[s]} deg(P_j)$.*

So given the Bezout's theorem, for the case of the discussed problem the number of common zeros, or equivalently $|V|$ is bounded by 2^s. So any circuit computing f should satisfy $r^n \leq 2^s$ and thereby $s = \Omega(n \log(r))$.

3 Remarks on the Bezout's theorem

Although the statement of the Bezout's theorem is a simple one, the rigorous proof for the statement is not as simple. In this section we give a primitive sketch of the proof and further remarks on the classical Bezout's theorem.

In order to approach this problem, we need a couple of definitions. Here we introduce a stronger version of the Bezout's theorem in which we deal with essential elements and ideas in algebraic geometry. Then we prove that the stronger version of the statement readily implies the classical Bezout's theorem.

Definition 1. *For a set of polynomials in $\mathbb{K}[x_1, x_2, \ldots, x_n]$, we define the associated veriety, as a subset of points $\mathbb{K}^n$ along which the set of polynomials simultaneously vanish.*

Also remember that for a commutative ring R, and ideal is a subset of the ring which is closed under addition and multiplication by the elements of the ring. More precisely, an ideal I of the ring $\mathbb{K}[x_1, x_2, \ldots, x_n]$, is a subset $I \subset \mathbb{K}[x_1, x_2, \ldots, x_n]$ such that if $p, q \in I$ and $r \in \mathbb{K}[x_1, x_2, \ldots, x_n]$, then $p + q \in \mathbb{K}[x_1, x_2, \ldots, x_n]$ and $p.r \in I$. Now for polynomials $p_1, p_2, \ldots, p_s$ define the following ideal.

Definition 2. *The ideal of the set of polynomials $p_1, p_2, \ldots, p_s$ in $\mathbb{K}[x_1, x_2, \ldots, x_n]$ is the set:*

$$I(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{s} q_i p_i | q_i \in \mathbb{K}[x_1, x_2, \ldots, x_n]\}$$

From these two definitions we can also define veriety of an ideal $V(I)$ for ideal I, to be the set of point in $\mathbb{K}^n$ along which all the members of the ideal vanish. And also define an ideal of a variety $I(V)$, to be the set of polynomials, vanishing on V. Clearly $I \subset I(V(I))$. Since $I(V(I))$ at least contains I but it might contain more elements. Now we can define define dimension and degree for a variety.

Definition 3. *The dimension of V is:*

$dim(V) := \min(k)$: *there exists and affine subspace* A *of dimension* $n-k$ *such that* $0 < A \cap V < \infty$

From this definition one can conclude that for varieties V and W, the union satisfies:

$$dim(V \cup W) = \max\{dim(V), dim(W)\}$$

Also define the degree of a veriety to be:

Definition 4. *The degree of a variety* V *is:*

$$deg(V) = \max_{Affine\ subspsce A: dim(A)+dim(V)=n; |A\cap V|<\infty} \{|A \cap V|\}$$

Therefore, from this definition, if the size of V is finite the degree $deg(V)$ is equal to the number of elements $|V|$ in V. Now that we have defined the essential definitions, we can phrase the strong Bezout's.

Theorem 2. *(Strong Bezout's Theorem) Let* $\mathbb{K}$ *to be an algebraically closed field, and* V *and* W *to be varieties over the field. Then:*

$$deg(V \cap W) \leq deg(V).deg(W)$$

Given this abstract form of the Bezout's theorem, the major step is to show:

Claim 2. *Strong Bezout's Theorem* $\implies$ *Classical Bezout's Theorem*

Proof. Consider the polynomials $p_1, p_2, \ldots, p_s$ to be a polynomials in $\mathbb{K}[x_1, x_2, \ldots, x_n]$, where $\mathbb{K}$ is some algebraically closed field. Then for each polynomial we consider the variety $V(p_i)$ as the set of points in $\mathbb{K}^n$ where the polynomial vanishes. Also let $V = \cap_{i\in[s]} V(p_i)$. From the strong Bezout's theorem:

$$deg(V) \leq \prod_{j\in s} deg(V_i)$$

Now it suffices to translate the meaning of degree in the classical Bezout's settings. By the definition, if the dimension of V is nonzero then the size of V is infinite. But we know that the size of V is finite, and there fore using the definition for the degree we get that $deg(V) = |V|$. Next we must show that $deg(V(p_i)) = deg(p_i)$. If this is true then we are done. First notice that the dimension of $V(p_i)$ is $n-1$. Because V can be infinite and according to the definition $dim(V_i) \leq n-1$. If we choose A to be a line, i.e., an affine subspace of co-dimension $n-1$, then because of the fundamental theorem of algebra $|A \cap V(p_i)|$ is at most $deg(p_i)$. Therefore V_i is finite and $deg(V_i) = |V_i| = deg(p_i)$.

□

4 Smolensky's Proof

Smolensky considers an equivalent version of the circuit where the addition and multiplication gates have fanin 2 and fanout 1, and he adds a duplicator gate which copies a wire into two wires. Namely the following gates:

$$\text{addition} : (x, y) \mapsto (x + y)$$

$$\text{multiplication} : (x, y) \mapsto (x * y)$$

$$\text{duplicator} : (x) \mapsto (x, x)$$

So for example the polynomial equations corresponding to such circuit looks like $y_1 \leftarrow x_1 + x_2$, $y_2 \leftarrow y_1 * x_3$ and $(y_3, y_4) \leftarrow y_2$.

Strangely, Smolensky proves a lower-bound on the number of duplicators used in a circuit of the map $(x_1, x_2, \ldots, x_n) \mapsto (x_1^r, x_2^r, \ldots, x_n^r)$ and recovers the $\Omega(n \log(r))$ lower bound. This imposes a lower-bound on the general circuits with arbitrary fanout. That is because any general circuit of size s can be converted to a restricted circuit with duplicators with $O(s)$ extra number of duplicators.

Here we phrase the main lemma:

Lemma 1. *Consider any restricted circuit, with inputs $\tilde{x} = x_1, x_2, \ldots, x_n$, outputs $z_1(\tilde{x}), z_2(\tilde{x}), \ldots, z_m(\tilde{x})$, with duplicator gates $D_1, D_2, \ldots, D_s$ which duplicated the polynomials $d_1(\tilde{x}), d_2(\tilde{x}), \ldots, d_s(\tilde{x})$. Consider any polynomial $T(x_1, \ldots, x_n, z_1, \ldots, z_n)$, with degree $\leq k$, then there exists a polynomial τ in $n + s$ variables such that:*

$$T(x_1, x_2, \ldots, x_n, z_1(\tilde{x}), z_2(\tilde{x}), \ldots, z_m(\tilde{x})) = \tau(x_1, \ldots, x_n, d_1^k(\tilde{x}), \ldots, d_s^k(\tilde{x}))$$

Where τ has degree $\leq k$ over the first n variables and degree at most s in the last s variables.

In the next lecture we will talk about how this lemma leads to a lower bound for the function f.