6.885 Algebra and Computation November 30, 2005

Lecture 23

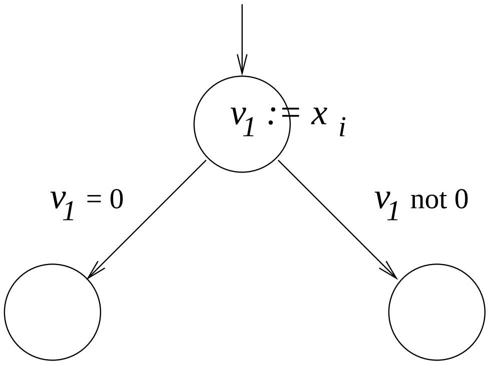
Lecturer: Madhu Sudan *Scribe: Karola Mészáros*

Strassen's lower bound for polynomial evaluation and Bezout's theorem

Recall Strassen's algorithm from the previous lecture:
Given: $(a_0, \ldots, a_{n-1}), (x_1, \ldots, x_n) \in K$, and polynomial $p(x) = \sum_{i=0}^{n-1} a_i x^i$
Task: find $(z_1, \ldots, z_n)$, $z_i = p(x_i)$

How many steps do we need to accomplish this task? Using the Fast Fourier Transform (FFT) we need $O(n \log^2 n)$ steps. Strassen was interested whether it can be done faster, and he showed that $\Omega(n \log n)$ steps are needed in algebraic computation trees.

Reminder of Algebraic Computation Trees



Perform computations of the form $v_i := x_j$, or $v_i := v_j \cdot v_k$, where $j, k < i$ and $\cdot \in \{+, -, \times, /\}$ and then branch depending on whether v_i is 0 or not. The complexity of the computation is the depth of the tree. By the time we reach the leaf level we want to know the values $z_1, \ldots, z_n$.

Idea behind the lower bound:

- paths in the computation tree form a variety
- varieties are not "too complex" in small depth trees
- variety represented by FT is complex

Paths are varieties: We have an input $(x_1, \ldots, x_n)$, and as we go down the algebraic computation tree we compute $v_1, v_2, \ldots, v_i$, and at a given stage the state variables are $(x_1, \ldots, x_n, v_1, v_2, \ldots, v_i)$. Supposing that at node v_i we branched left, i.e. $v_i = 0$, we have our variety defined by the equations coming from each node, which are equations of the form $v_k - x_j = 0$ or $v_i - (v_j + v_k) = 0$, or, the $+$ substituted by one of the operations, plus, if we branched left we have in addition the equation $v_i = 0$. On the other hand, if we bracnhed right from v_i, that is if $v_i \neq 0$, then instead of the equation $v_i = 0$, we have another one, thanks to Rabinowich's trick, namely, $1 - v_i {v_i}^{-1} = 0$. Thus, we see that paths are determined by equations, and so are varieties. Also, each node adds two equations in total, and clearly the degree of both equation is not more than 2, and so things are not getting "very complex."

Complexity measure

In order to talk about the complexity of computation we must have established a complexity measure to start with. Since the paths in the algebraic computation trees correspond to varieties, it is natural to take the "degree of the variety" to be the complexity measure. We are not going to give a formal definition of this concept from algebraic geometry, rather we are going to give a more intuitive/elementary definition. So, for us, the degree of the variety will roughly mean, the number of points that are in the variety. Of course, there could be infinitely many points in the variety, in which case our definition is not the most beneficial. But, let's start with a simple case.

Consider polynomials $f_1, \dots, f_n \in k[x_1, \dots, x_n]$. Suppose that the number of points in the variety $V(f_1, \dots, f_n)$ is finite. In this case it may be reasonable to impose that the complexity measure is the number of points in the variety. For example, if the polynomials $f_1, \dots, f_n$ are all linear, the number of points in the variety is at most 1. Also, if the polynomials $f_1, \dots, f_n$ are defined by $f_i = p(x_i)$, where p is a polynomial of degree d, then the complexity is at most d^n.

Given polynomials $Q_1(x, y)$ and $Q_2(x, y)$ of degrees d_1 and d_2 respectively, then either Q_1 and Q_2 have at most $d_1 \cdot d_2$ common zeroes, or Q_1 and Q_2 have a common factor. For example, if $Q_1(x, y) = (x+y)(x^2+y^2-1)$ and $Q_2(x, y) = (x+y)(x^3+y^2-3)$. Then, Q_1 and Q_2 have infinitely many common zeroes, since all the zeroes of $x + y = 0$ are their zeroes too. By our complexity measure the varieties defined by $Q_1(x, y)$ and $Q_2(x, y)$ would be the same. However, we would like some complexity measure to deal with this kind of case too, which would be intrinsic to the variety and not the way it is specified. Thus, we modify our definition, and set the complexity of a variety to be the number of isolated points of the variety.

We have a geometric understanding of isolated points in real space, indeed, a point of a point set is isolated if there is a neighborhood of it that contains no other point from the point set. Inspired by this geometric notion, we define isolated points in general, as follows.

Definition. Given $f_1, \dots, f_n \in k[x_1, \dots, x_n]$, and $\alpha \in k^n$, we say that α is an isolated point of $V(f_1, \dots, f_n)$ if $f_i(\alpha) = 0$ for all $i \in [n]$, and the matrix M with entries $M_{ij} = \frac{\partial f_i}{\partial x_j}(\alpha)$ is nonsingular.

Consider the variety defined by polynomials $f_1, \dots, f_m \in k[x_1, \dots, x_n]$, with $m < n$, then the complexity of the variety is equal to max_L(number of isolated points of $(V \cap U)$), where L is a linear subspace. On the other hand, if $m \geq n$, then we need to generalize our definition of an isolated point, as follows:

Definition. Given $f_1, \dots, f_m \in k[x_1, \dots, x_n]$, $m \geq n$, and $\alpha \in k^n$, we say that α is an isolated point of $V(f_1, \dots, f_n)$ if $f_i(\alpha) = 0$ for all $i \in [n]$, and the matrix M with entries $M_{ij} = \frac{\partial f_i}{\partial x_j}(\alpha)$ has rank n.

Claim 1. If $(f_1, \dots, f_m)$ and $(g_1, \dots, g_{m'})$ generate same ideal, then α is isolated with respect to $(f_1, \dots, f_m)$ if and only if α is isolated with respect to $(g_1, \dots, g_{m'})$.

From this we see that although an isolated point is defined through the basis of the ideal, it does not depend in it, and is intrinsic to the ideal itself.

Claim 2. If α is isolated with respect to I it follows that α is isolated with respect to $I(V(I))$.

One more notion of complexity

Fix a polynomial $p(x) = 0$ with n distinct zeroes. Consider $x_1, \dots, x_n \in B = \{b_1, \dots, b_n\}$, and an algebraic computation tree. There are n^n possible inputs from B^n (corresponding to the root of the tree). If the depth of the tree is D, then there exists a leaf node reached by $\frac{|B^n|}{2^D}$ inputs, since at every node we branch to exactly two other nodes. At this leaf we have a point of form $(x_1, \dots, x_n, p(x_1), \dots, p(x_n), \dots)$ and project onto points $(x_1, \dots, x_n, p(x_1), \dots, p(x_n))$. We intersect this variety by $\cap_{i=1}^n z_i$.

Claim. Every $(x_1, \dots, x_n) \in B^n$ that leads to the above leaf is an isolated point in the above variety.

Indeed, $\{z_i = 0\}_{i \in [n]} \cap V \subset V(I(z_1 = p(x_1) = 0, \dots, z_n = p(x_n) = 0))$, and the complexity of projected variety is at least $\frac{|B^n|}{2^D}$.

Bezout's Theorem. (Simplified version) The complexity of $V(f_1, \dots, f_m)$ is $\prod_{i=1}^m \deg f_i$.

Theorem. The complexity of a projection of a variety is less than or equal to the complexity of the variety.

Backup: why is the projection onto points $(x_1, \dots, x_n, p(x_1), \dots, p(x_n))$ from $(x_1, \dots, x_n, p(x_1), \dots, p(x_n), \dots)$ a variety? Because of quantifier elimination. $V_\pi = \{(\alpha_1, \dots, \alpha_{n/2}) |$there exist $\alpha_{n/2+1}, \dots, \alpha_n$ such that $(\alpha_1, \dots, \alpha_n) \in V\}$.

Now, assuming both theorems, we get that every leaf has complexity at most 4^D. So, $|B|^n/2^D \leq 4^D$, thus, $D = \Omega(n \log n)$.

We will not prove the theorem about the complexity of the projection of the variety, but will give an outline of the proof of Bezout's theorem. The proof that we will follow is due to Wooley from 1996. The idea of the proof is to suppose the opposite of what is stated and derive a contradiction. Indeed, consider $f_1, \dots, f_m \in k[x_1, \dots, x_n]$, $m = n$, and let $S = \{\alpha_1, \dots, \alpha_N\}$ be a subset of the isolated zeroes, and suppose $N > \prod_{i=1}^m d_i$, where $d_i = \deg f_i$. From here we wish to derive a contradiction.

We would like to construct a polynomial $P(x_1)$ such that $P(\alpha^{(1)}) = 0$ for every $\alpha = (\alpha^{(1)}, \dots, \alpha^{(n)}) \in S$, P is not identically zero, and $\deg P \leq \prod_{i=1}^m d_i$. We would like to eliminate $x_2, \dots, x_n$, so we should do quantifier elimination. This, however, does not quite work.

First construction. Let $Q(y_1, \dots, y_m, x_1) \in k'[y_1, \dots, y_m, x_1]$ (where k' is a slight modification of k, see paper by Wooley) such that $\deg_{x_1} Q \leq \prod_{i=1}^m d_i$, $Q(\widetilde{f_1}, \dots, \widetilde{f_m}, x_1) = 0$ (where $\widetilde{f_i}$ is a slight modification of f_i, see paper by Wooley), and $Q \notin k'[y_1, \dots, y_m]$.

Now, set $P_0(x_1) = Q(0, \dots, 0, x_1)$. Then $\deg P_0 \leq \prod_{i=1}^m d_i$, and $P(\alpha^{(1)}) = Q(0, \dots, 0, \alpha^{(1)}) = Q(f_1(\alpha), \dots, f_m(\alpha), \alpha^{(m)}) = Q(f_1(x), \dots, f_m(x), x_1)|_{x=\alpha} = 0$. However, this construction fails, since we cannot assert that P_0 is not identically zero.

A new idea: Consider ring $k[z]$. There exist $\gamma_1, \dots, \gamma_m \in k[z]$ such that $Q(\gamma_1 z, \dots, \gamma_m z, x_1) \neq 0$. Set $P_1(x_1) = Q(\gamma_1 z, \dots, \gamma_m z, x_1)$. Then $\deg P_1 \leq \prod_{i=1}^m d_i$, and P_1 is not identically zero. Also, $P_1(\alpha^{(1)}) = Q(\gamma_1 z, \dots, \gamma_m z, x_1) = Q(0, \dots, 0, \alpha^{(1)}) \bmod z = 0 \bmod z$. Now the proof uses essentially Hensel's lifting and linear algebra. Indeed, consider $P(z)$, and $\alpha_1, \dots, \alpha_k$ such that $P(\alpha_i) = 0 \bmod N$. If $(\alpha_i - \alpha_j, N) = 1$, then $k \leq degP$. Also, to the initial set of conditions $Q(y_1, \dots, y_m, x_1) \in k'[y_1, \dots, y_m, x_1]$ (where k' is a slight modification of k, see paper by Wooley) such that $\deg_{x_1} Q \leq \prod_{i=1}^m d_i$, $Q(\widetilde{f_1}, \dots, \widetilde{f_m}, x_1) = 0$ (where $\widetilde{f_i}$ is a slight modification of f_i, see paper by Wooley), and $Q \notin k'[y_1, \dots, y_m]$, condition $\widetilde{f_i} = f_i \bmod z$ is added. For details see [Wooley, '96].