Lecture 40 Probabilistic Tests with Polynomials

In this lecture, we give a useful probabilistic technique for testing properties that are equivalent to the vanishing of some polynomial of low degree. This technique has many interesting applications, not only in algebraic algorithms, but also in graph theory and combinatorics. Several examples of its use will be given later.

Good deterministic algorithms sometimes require considerable effort to program, whereas "quick and dirty" methods involving random choices are often just as good in practice. For example, it is quite difficult to test deterministically whether a multivariate polynomial given by a straight-line program is identically zero; however, there is a fast probabilistic test: evaluate the polynomial on a randomly chosen input and check whether the result is 0. If not, the polynomial is certainly not identically 0; if so, chances are good that it is.

The technique is based on the following theorem due to Zippel [111] and independently to Schwartz [92]. It says essentially that the solutions of a multivariate polynomial equation of low degree are sparse. Intuitively, this theorem is true over the real numbers for any polynomial, regardless of degree: the set

$$\{(x_1, \ldots, x_n) \in \mathcal{R}^n \mid p(x_1, \ldots, x_n) = 0\}$$

is a surface of dimension $n-1$. For example, a linear equation

$$\sum_{i=1}^{n} a_i x_i \quad = \quad a$$

describes a hyperplane; in three dimensions, the quadratic equation

$$x^2 + y^2 - z^2 \quad = \quad 0$$

describes the surface of a cone. A randomly chosen point with respect to almost any reasonable probability distribution will almost certainly not lie on that surface. Under the degree restriction, the theorem is also true over other fields besides the reals, including finite fields.

Theorem 40.1 *Let k be a field and let $S \subseteq k$ be an arbitrary subset of k. Let $p(\overline{x})$ be a polynomial of n variables $\overline{x} = x_1, \ldots, x_n$ and total degree[5] d with coefficients in k. Then the equation $p(\overline{x}) = 0$ has at most $d \cdot |S|^{n-1}$ solutions in S^n.*

Proof. The proof is by induction on n and d. For $n = 1$, the result follows from the fact that a univariate polynomial of degree d can have no more than d roots in k. For $d = 1$, we need to show that a hyperplane

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \quad = \quad a \tag{64}$$

in k^n can intersect S^n in at most $|S|^{n-1}$ points. Pick some $a_i \neq 0$, say without loss of generality $a_1 \neq 0$. Then for all solutions $\overline{x}$ of (64),

$$x_1 \quad = \quad \frac{1}{a_1}(a - \sum_{i=2}^{n} a_i x_i) \ ,$$

therefore the value of x_1 is uniquely determined by the values of $x_2, \ldots, x_n$. There are exactly $|S|^{n-1}$ assignments to $x_2, \ldots, x_n$ from S, thus at most $|S|^{n-1}$ solutions to (64).

Now suppose we have a polynomial p of degree $d > 1$ with $n > 1$ variables. If p is not irreducible, *i.e.* if p has a nontrivial factorization $p = qr$ into two polynomials q and r of lower total degree, then by the induction hypothesis, q has no more than $\deg q \cdot |S|^{n-1}$ zeros in S^n and r has no more than $\deg r \cdot |S|^{n-1}$ zeros in S^n. But $p(\overline{a}) = 0$ iff $q(\overline{a})r(\overline{a}) = 0$ iff either $q(\overline{a}) = 0$ or $r(\overline{a}) = 0$, thus

$$\{\text{zeros of } p \text{ in } S^n\} \quad = \quad \{\text{zeros of } q \text{ in } S^n\} \cup \{\text{zeros of } r \text{ in } S^n\} \ .$$

It follows that

$$\begin{aligned}
&|\{\text{zeros of } p \text{ in } S^n\}| \\
&\quad = \quad |\{\text{zeros of } q \text{ in } S^n\} \cup \{\text{zeros of } r \text{ in } S^n\}| \\
&\quad \leq \quad |\{\text{zeros of } q \text{ in } S^n\}| + |\{\text{zeros of } r \text{ in } S^n\}| \\
&\quad \leq \quad \deg q \cdot |S|^{n-1} + \deg r \cdot |S|^{n-1} \\
&\quad = \quad (\deg q + \deg r) \cdot |S|^{n-1} \\
&\quad = \quad d \cdot |S|^{n-1} \ .
\end{aligned}$$

[5]Maximum degree of any term.

Finally, we are left with the case that p is irreducible of degree $d > 1$ with $n+1$ variables $x_1, \ldots, x_{n+1}$. Let $\overline{x} = x_1, \ldots, x_n$. Then $p = p(\overline{x}, x_{n+1})$. For each $s \in S$, consider the polynomial $p(\overline{x}, s) \in k[\overline{x}]$. By the induction hypothesis, $p(\overline{x}, s)$ has at most $d \cdot |S|^{n-1}$ zeros in S^n (unless $p(\overline{x}, s)$ is identically zero; but we show below that this cannot happen if p is irreducible). Since $p(\overline{x}, s)$ has at most $d \cdot |S|^{n-1}$ zeros in S^n, p has at most $|S| \cdot d \cdot |S|^{n-1} = d \cdot |S|^n$ zeros in S^{n+1}.

To show that $p(\overline{x}, s)$ is not identically zero, we show that if it were, then the polynomial $x_{n+1} - s$ would divide p, contradicting the irreducibility of p. Suppose then that $p(\overline{x}, s) = 0$. Collect terms of p with like powers of x_{n+1} so that p is expressed as a polynomial in x_{n+1} with coefficients in the polynomial ring $k[\overline{x}]$. Divide p by the polynomial $x_{n+1} - s$ using ordinary polynomial division with remainder. Then

$$p(\overline{x}, x_{n+1}) \quad = \quad q(\overline{x}, x_{n+1})(x_{n+1} - s) + r$$

where the degree of the remainder r is less than the degree of the divisor $x_{n+1} - s$, so r is a constant. Evaluating both sides of the equation at $x_{n+1} = s$, we get that $r = 0$. Thus

$$p(\overline{x}, x_{n+1}) \quad = \quad q(\overline{x}, x_{n+1})(x_{n+1} - s) \ ,$$

contradicting the irreducibility of p. □

The following corollary is immediate.

Corollary 40.2 *Let $p(x_1, x_2, \ldots, x_n)$ be a nonzero polynomial of degree d with coefficients in a field k, and let $S \subseteq k$. If p is evaluated on a random element $(s_1, \ldots, s_n) \in S^n$, then*

$$\Pr(p(s_1, \ldots, s_n) = 0) \quad \leq \quad \frac{d}{|S|} \ .$$

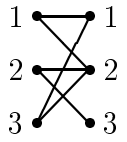
40.1 Applications

We give three applications of Theorem 40.1 and Corollary 40.2: finding perfect matchings, testing isomorphism of labeled trees, and computing the rank of a matrix over a finite field.

Perfect Matchings

We know how to test for the existence of a perfect matching in a bipartite graph G and find one if it exists in polynomial time. It is unknown whether this problem is in NC. However, the following approach, based on an observation of Lovász [74], gives a random NC algorithm.

Assign to each edge (i, j) of G an indeterminate x_{ij} and consider the $n \times n$ bipartite adjacency matrix X with these indeterminates instead of 1's. For example,



$$X = \begin{bmatrix} x_{11} & x_{12} & 0 \\ 0 & x_{22} & x_{23} \\ x_{31} & x_{32} & 0 \end{bmatrix}$$

The determinant $\det X$ is a polynomial of degree n in the indeterminates x_{ij} with one term for each perfect matching, and none of these terms cancel. For example, the graph above has two perfect matchings

corresponding to the two terms of the determinant

$$\det X = x_{12}x_{23}x_{31} - x_{11}x_{23}x_{32} .$$

Thus G has a perfect matching iff $\det X$ is not identically 0. This is difficult to test deterministically, since $\det X$ may be quite large. Chistov's or Berkowitz' algorithm gives a polylog-depth circuit with inputs x_{ij} that computes the value of $\det X$ for any specialization of the indeterminates x_{ij}, but it is difficult to test deterministically whether all such specializations give 0.

However, we can test this in *RNC* by assigning randomly chosen elements of a large enough finite field (say $\mathcal{Z}_p$, where p is some prime greater than $2n$) to the x_{ij}, and then asking whether the determinant evaluated at those random elements is 0. This will happen with probability 1 if $\det X$ is indeed identically 0, and with probability at most $\frac{n}{2n} = \frac{1}{2}$ if not, by Corollary 40.2.

Given the ability to test for the existence of a perfect matching, we can then find one by deleting edges one by one and testing for the existence of a perfect matching without that edge.

Isomorphism of Unordered Directed Trees

Here is an efficient probabilistic test for deciding whether two unordered[6] directed trees of height h and size n are isomorphic. Associate with each vertex v a polynomial f_v in the variables $x_0, x_1, \ldots, x_h$ inductively, as follows. For each leaf v, set $f_v = x_0$. For each internal node v of height k with children $v_1, \ldots, v_m$, set

$$f_v = (x_k - f_{v_1})(x_k - f_{v_2}) \cdots (x_k - f_{v_m}) .$$

The degree of f_v is equal to the number of leaves in the subtree rooted at v. Using the fact that polynomial factorization is unique, it can be shown that two trees are isomorphic iff the polynomials associated with the roots of the trees are equal. This gives an efficient probabilistic test for unordered tree isomorphism: test whether the difference of these two polynomials is identically zero by evaluating it on a random input.

[6] A directed tree is *ordered* if the left-to-right order of each node's children is given.

Matrix Rank

Mulmuley's algorithm computes the rank of a matrix over an arbitrary field k. Recall that for a square matrix A, if $\operatorname{rank} A = \operatorname{rank} A^2$, then $\operatorname{rank} A$ is given by the index of the last nonzero term in the characteristic polynomial of A. *I.e.*, if

$$\chi_A(\lambda) \;=\; \lambda^n - s_1\lambda^{n-1} + s_2\lambda^{n-2} - \cdots \pm s_r\lambda^{n-r}$$

where $s_r \neq 0$, then $\operatorname{rank} A = r$. If $\operatorname{rank} A \neq \operatorname{rank} A^2$ and we are working in the complex numbers, then we can take $\overline{A}^T A$, where $\overline{A}^T$ is the conjugate transpose of A. As shown in Lecture 33, this matrix has the same rank as A and the same rank as its square. Over finite fields, however, this does not work. Mulmuley's algorithm closes this gap, but his construction introduces an extra indeterminate, and dealing with the resulting symbolic expressions requires more processors.

Here is a probabilistic approach suggested in [15] that saves a factor of n in the processor bound over Mulmuley's deterministic algorithm. Multiply A on the left by a random matrix R. The elements of R are chosen uniformly at random from a sufficiently large set. By Corollary 40.2, R is nonsingular with high probability: R is singular if and only if its determinant vanishes, and this is a polynomial equation of low degree. Therefore, with high probability, RA has the same rank as A, since the rank of RA is the dimension of the image of RA as a linear map.

We argue also that with high probability, RA and $(RA)^2$ have the same rank, allowing us to compute the rank from the characteristic polynomial of RA as in Lemma 32.1.

Let $r = \operatorname{rank} A$. The condition

$$\operatorname{rank} (RA)^2 = \operatorname{rank} RA = \operatorname{rank} A \tag{65}$$

is equivalent to the condition that the subspaces $\operatorname{im} RA$ and $\ker A$ are of complementary dimension and intersect in the trivial subspace 0; in other words, that every vector in k^n can be represented uniquely as the sum of a vector in $\operatorname{im} RA$ and one in $\ker A$. In symbols,

$$k^n \;\cong\; \operatorname{im} RA \oplus \ker A \tag{66}$$

where $\oplus$ denotes direct sum and $\cong$ denotes isomorphism of vector spaces.

Now select a basis for $\operatorname{im} A$ among the columns of A. These columns will comprise an $n \times r$ matrix B. Let C be an $n \times (n-r)$ matrix whose columns span $\ker A$. Then condition (66) is equivalent to the condition that the $n \times n$ matrix $[RB|C]$ formed by juxtaposing the columns of RB and C is nonsingular; equivalently, $\det [RB|C] \neq 0$. By Corollary 40.2, this occurs with high probability.

The beauty of this approach is that we never need to compute B or C; we are happy enough just knowing that they exist.