Advanced Algorithms	**Lecture 7: April 5, 18 and 19, 1995**
Lecturer: Johan Håstad	*Scribe: Lars Arvestad*

7.1 Factoring polynomials over the integers

Given is a polynomial $g(x) = \sum_{i=0}^{n} g_i x^i, g_i \in \mathbb{Z}$. What is the factorization of G?

A question to start with is how such a factorization looks like. We assume that $\gcd(g_0, \ldots, g_n) = 1$, because if that is not the case, we can always cancel that factor out, getting the polynomial in the desired form. Can the factors of g have rational coefficients for instance? The problem is settled with the following lemma.

Lemma 1 *If $g(x)$ can be written as a product $h_1(x)h_2(x)$ then h_1 and h_2 have integer coefficients.*

This is disregarding such "cheating" as writing the factors as $\frac{1}{7}h_1(x) \cdot 7h_2(x)$, of course.

Proof: Suppose that h_1 has integer coefficients with greatest common divisor 1. If not, multiply h_1 with a suitable rational and divide h_2 with the same. We now claim that h_2 also has integer coefficients. For contradiction, assume h_2 does not have integer coefficients and that the least common denominator of the coefficients in h_2 is M. Let p be a prime factor in M and consider the following equation.

$$M \cdot g(x) = h_1(x) \cdot Mh_2(x)$$

The part $Mh_2(x)$ is now an integer polynomial. If we look at this modulo p, we can say that $Mg(x) \equiv 0$, but $h_1(x) \not\equiv 0$ since the greatest common divider of the coefficients of h_1 is 1. This also applies for Mh_2, so we have reached a contradiction. ■

Another question is how large the coefficients of h_1 and h_2 can be. They can considerable larger than the coefficients in g! We can estimate how large the coefficients of a polynomial can be in terms of the size of the complex roots of the polynomial. Let $M = \max\{|g_0|, \ldots, |g_n|\}$ and write g as $g(x) = g_n \prod_{i=1}^{n}(x - x_i)$ where x_i is the complex roots of $g(x) = 0$.

Let us estimate the magnitude of the roots.

Lemma 2 $|x_i| \leq 2M/|g_n|$

Proof: The idea is to show that if $|x_i|$ is larger than the given bound, then $g_n x^n$ is the dominant term and cannot be cancelled. Suppose x is a root for g.

$$|g(x)| = \left|\sum_{i=0}^{n} g_i x^i\right| = |g_n x^n| - \left|\sum_{i=0}^{n-1} g_i x^i\right| \geq |g_n||x^n| - M\sum_{i=0}^{n-1} |x|^i$$

The sum in the last term is a geometric series with quotient > 2 if we would assume that $|x| > 2M/|g_n|$. Therefore,

$$|g_n||x^n| - M\sum_{i=0}^{n-1} |x|^i > |g_n||x^n| - 2M|x|^{n-1} \geq |x|^{n-1}(|g_n||x| - 2M) \geq 0$$

which is a contradiction and the lemma holds. ■

We can use this estimate for giving an upper bound on the sizes of the coefficients of the factors of g. Let $h_1(x) = h_1^{led}\prod_{i\in S}(x - x_i)$ where h_1^{led} is the leading coefficient in h_1 and S is the set of roots of g that also solve $h_1(x) = 0$. We know that h_1^{led} divides g_n, so $|h_1^{led}| \leq |g_n|$.

If h_1 has degree m then the coefficient for degree i is bounded by

$$h_1^{led}\binom{m}{i}\left(\frac{2M}{|g_n|}\right)^{m-i} \leq |g_n| \cdot 2^m \cdot \left(\frac{2M}{|g_n|}\right)^m \leq (4M)^m \leq (4M)^n.$$

One idea would be to factorize modulo the upper bound we have. In other words, find $h_1, h_2 \in \mathbb{Z}[x]$ such that $g(x) = h_1(x)h_2(x)$ and with coefficients less than $(4M)^n$. We could do the factorization modulo a prime $p > (4M)^n$ so we study $g(x) = \prod_{i=1}^{s} q_i(x) \pmod{p}$ and $h_1(x) = \prod_{i\in S} q_i(x) \pmod{p}$, $h_2(x) = \prod_{i\notin S} q_i(x) \pmod{p}$.

A straightforward algorithm would be to try $h_1 = \prod_{i\in S} q_i(x)$ for every S and see if h_1 divides g. This actually works reasonable in practice, but in the worst case we will get in the order of n linear factors of g modulo p and this would give us 2^n ways of choosing h_1. We will give an algorithm that is polynomial even in the worst case.

7.1.1 The algorithm

A high level description of our algorithm is as follows.

1. Find factors $g(x) = h_1(x)h_2(x)$ modulo p^k for some p and k. Think of p as small and k as large. Here h_1 is an irreducibel polynomial and h_2 is "the rest".

2. Find a polynomial $h(x) = h_1(x)z(x) \pmod{p^k}$ where $deg(h) < deg(g)$. The coefficients of h shall be small, say less than M'.

3. Compute $gcd(g(x), h(x))$.

Now, why would this work? We want the *gcd* to be non-trivial and there is a simple condition for that. Say $deg(f) = n$ and $deg(g) = m$.

Lemma 3 *$gcd(f(x), g(x))$ is non-trivial if and only if there is a non-trivial solution to $f(x)h_1(x) + g(x)h_2(x) = 0$ with $deg(h_1) < m$ and $deg(h_2) < n$.*

Proof: Say $gcd(f(x), g(x)) = d(x)$. Then $h_1(x) = g(x)/d(x)$ and $h_2(x) = -f(x)/d(x)$. Inversly, if $gcd(f(x), g(x)) = 1$ then every irreducible factor in g divides $h_1(x)$ which in turn means that $g(x)|h_1(x)$. This is a contradiction because $deg(h_1) \geq deg(g)$. ∎

This condition is therefore a matter of solvability of a homogenous linear equation system and we can check this efficiently by looking at the *Sylvester resultant*.

The Sylvester resultant

Define the matrix S as

$$\underbrace{\begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{n-1} & f_n & 0 & \dots & 0 \\ 0 & f_0 & f_1 & \dots & f_{n-2} & f_{n-1} & f_n & \dots & 0 \\ \vdots & & & & & & & & \\ 0 & 0 & 0 & \dots & f_0 & f_1 & f_2 & \dots & f_n \\ g_0 & g_1 & g_2 & \dots & g_m & 0 & 0 & \dots & 0 \\ 0 & g_0 & g_1 & \dots & g_{m-1} & g_m & 0 & \dots & 0 \\ \vdots & & & & & & & & \\ 0 & 0 & 0 & \dots & g_0 & g_1 & g_2 & \dots & g_m \end{pmatrix}}_{n+m \text{ elements}}$$

What we want to check is wether the system

$$S^T \begin{pmatrix} h_1^{(0)} \\ h_1^{(1)} \\ \vdots \\ h_1^{(m-1)} \\ h_2^{(0)} \\ h_2^{(1)} \\ \vdots \\ h_2^{(n-1)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

is solvable. The Sylvester resultant $R(f,g)$ is defined as the determinant of S and the system is solvable by a non-trivial solution if $R(f,g) = 0$. So by the above lemma, $gcd(f(x), g(x))$ is non-trivial if and only if $R(f,g) = 0$ and this applies both over the integers, modulo p and modulo p^k.

Example 1 *We are given polynomials $f(x) = 1 + 2x + x^3$ and $g(x) = 5 + x^2$. Does there exist polynomials $h_1(x) = h_1^0 + h_1^1 x$ and $h_2(x) = h_2^0 + h_2^1 x + h_2^2 x^2$ such that $h_1(x)f(x) + h_2(x)g(x) = 0$? We get the linear system*

$$\begin{pmatrix} 1 & 0 & 5 & 0 & 0 \\ 2 & 1 & 0 & 5 & 0 \\ 0 & 2 & 1 & 0 & 5 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h_1^0 \\ h_1^1 \\ h_2^0 \\ h_2^1 \\ h_2^2 \end{pmatrix} = \bar{0}$$

which is solvable if

$$R(f,g) = \begin{vmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 1 \\ 5 & 0 & 1 & 0 & 0 \\ 0 & 5 & 0 & 1 & 0 \\ 0 & 0 & 5 & 0 & 1 \end{vmatrix} = 0.$$

In this case, $R(f,g) = 46$ and the system has therefore no solution except the zero vector.

We can now see that the outlined algorithm should work, because with $h_1(x)$ dividing both $h(x)$ and $g(x)$ modulo p^k, the Sylvester resultant will be 0 modulo p^k and thereby announcing that $gcd(h(x), g(x))$ modulo p^k is non-trivial.

We would wish that having $R(h,g) = 0 \pmod{p^k}$ would imply that $R(h,g) = 0$ over the integers. This can be achieved if we choose p and k appropriately. There is a rough upper bound on $R(h,g)$ that we can use. Say $m = deg(h)$ and $n = deg(g)$. We are computing a determinant with m rows of numbers that are less than M and have n rows with elements smaller than M'. This give us that $|R(h,g)| \leq (n+m)! M^m M'^n$. We should choose p^k to be larger than this number, because $R(h,g)$ in the integers would then be the same as modulo p^k.

7.1.2 Finding a factorization of $g(x)$ modulo p^k

The first task in the algorithm is to find factors $h_1(x)$ and $h_2(x)$ such that $g(x) = h_1(x)h_2(x)$ modulo some p^k. We can do that as follows.

(i) Choose p such that $g(x)$ is square free modulo p.

(ii) Factorize $g(x)$ modulo p.

(iii) For $i := 1$ to $k-1$, "lift" the factorization modulo p^i to modulo p^{i+1}.

We know from previous lectures how to do item (ii). Item (iii) is simplified by having $g(x)$ square free, which explains why we do item (i). Having $g(x)$ square free is equivalent to $gcd(g, g')$ being trivial, which in it's turn is equivalent to $R(g, g') \neq 0$, so we want to choose a p that does not divide $R(g, g')$. We can use that $R(g, g') \leq M^{2n}(2n)!$ to find such a prime number. If the resultant would be divisible by every prime up to some x, then $R(g, g')$ would be in the order of e^x, so there must be a $p < \log(M^{2n}(2n)!) = 2n \log M + 2n \log n$ we can use, and we can find it by simply compute the resultant over the integers and then try all small primes. Note that we can assume $g(x)$ to be square free over the integers, because if that was not the case, we can easily get a factor of $g(x)$ by computing $gcd(g, g')$.

Hensel lifting

The third task can be done with a procedure called Hensel lifting. We have $g(x) = h_1^{(i)}(x)h_2^{(i)}(x)$ modulo p^i and want $g(x) = h_1^{(i+1)}(x)h_2^{(i+1)}(x)$ modulo p^{i+1}. Consider $h_1^{(i+1)}(x) = h_1^{(i)}(x) + p^i h_1'(x)$ and $h_2^{(i+1)}(x) = h_2^{(i)}(x) + p^i h_2'(x)$. Then

$$h_1^{(i+1)}(x)h_2^{(i+1)}(x) = h_1^{(i)}(x)h_2^{(i)}(x) + h_1^{(i)}(x)p^i h_2'(x) + h_2^{(i)}(x)p^i h_1'(x) + p^{2i} h_1'(x)h_2'(x)$$

would hopefully equal $g(x)$ modulo p^{i+1}. This can be written as the equation

$$p^i g'(x) = h_1^{(i)}(x) p^i h_2'(x) + h_2^{(i)}(x) p^i h_1'(x) \bmod p^{i+1}$$

if we let $p^i g'(x) = g(x) - h_1^{(i)}(x) h_2^{(i)}(x)$. Then we can divide by p^i everywhere and get

$$h_1^{(i)}(x) h_2'(x) + h_1'(x) h_2^{(i)}(x) = g'(x) \pmod{p}$$

This is a linear equation system modulo p and if it is solvable, then it is computationally easy to do that. It turns out that this system is always solvable when $g(x)$ is square free. There are $n+2$ unknowns in h_1' and h_2' so we can fix one parameter, say the leading coefficient of h_1', and solve.

The whole procedure of finding a factorization of $g(x)$ modulo p^k involves k equation systems modulo p of size $n+1 \times n+1$.

7.1.3 Finding a small polynomial with a factor in common with $g(x)$

We now want to do task 2 in the algorithm, finding a polynomial $h(x) = h_1(x)z(x)$ modulo p^k, $deg(h) \leq n-1$, with small coefficients. The tool here is Lovász' algorithm for finding short vectors in a lattice. Given a set of vectors b_i, the corresponding lattice is defined as all integer combinatitions of these vectors i.e. all vectors of the form $\sum a_i b_i, a_i \in \mathbb{Z}$. To use this notation, we view our polynomials as vectors.

A polynomial $f(x) = \sum_{i=0}^{m} f_i x^i$ has a corresponding vector $(f_0, f_1, \ldots, f_m, 0, 0, \ldots, 0)$ in an n-dimensional space. Some consequences:

$$\begin{aligned} xf &\leftrightarrow (0, f_0, f_1, \ldots) \\ x^{n-m-1} f &\leftrightarrow (0, 0, \ldots, 0, f_0, \ldots, f_m) \end{aligned}$$

In our application, L_{h_1} is the lattice spanned by the vectors corresponding to $x^i h_1, 0 \leq i \leq n-m-1$ and $p^k e_i$, where e_i are the unit vectors. The vectors in this lattice corresponds exactly to the multiples of h modulo p^k. If we can find a short vector in L_{h_1}, then we have found the polynomial h with small coefficients.

Lovász' algorithm finds a vector of length less than $2^{(n-1)/2}\lambda$ in a lattice where λ is the length of the shortest vector. With input vectors of length B, the running time comes to $O(n^5 \log B)$. (See also notes from lecture 8.)

Having the length of the shortest vector bounded by $\sqrt{n}(4M)^n$ since the coefficients are smaller than $(4M)^n$, the length of the output vector is at worst

$2^{n/2}\sqrt{n}(4M)^n \leq (8M)^n$. This output vector can be translated into a polynomial h that is a multiple of h_1, but with coefficients bounded by $(8M)^n$. Thus if we choose $p^k > (n+m)!M^m M'^n \sim (8M)^{n^2}$ we are guaranteed that $gcd(g,h)$ will be non-trivial and this greatest common divisor can of course be calculated efficiently.

The length of the input vectors is in the order of $p^k \sim (8M)^{n^2}$. Therefore, the running time of Lovász' algorithm, which also dominates the whole factorization, comes to $O(n^5 \cdot \log((8M)^{n^2})) = O(n^7 \log M)$. This running time can be improved with a more careful implementation and analysis.

References

[1] Lenstra, Lenstra, Lovász: Factoring Polynomials with Rational Coefficients Math. Ann. 261, pp 515–534, 1982