CDM

Polynomials

Klaus Sutner
Carnegie Mellon University

Polynomials, Informally

Intuition says that a (univariate) polynomial is an expression of the form

$$p(x) = x^3 - 2x^2 + 3x - 1$$

There is an unknown or variable x and we

- form powers x^i of the variable (monomials),
- multiply them by an element in some ground ring (coefficients),
- add several such terms.

In the example, the ground ring is presumably $\mathbb{Z}$.

We can add and multiply polynomials, but division in general does not work (at least not without a remainder).

Coefficient Representation

Clearly, we can represent a polynomial by a list of its coefficients:

$$\boldsymbol{a} = (a_0, a_1, \ldots, a_d)$$

represents

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_d x^d$$

Typically we want $a_d \neq 0$, so d is the degree of p.

Note that this representation has no unknown, a minor cause for concern.

A typical "definition," found in may textbooks.

Definition

Let R be a commutative, unital ring. A univariate polynomial (over R) is an expression of the form

$$a_d x^d + a_{d-1} x^{d-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

where the coefficients a_i are in R and x is the unknown or variable. If $a_d \neq 0$, d is the degree of the polynomial. The terms x^i are monomials[†].

This suggests that a polynomial is a R-linear combination of monomials. Suggests, but it does not say that.

[†] The whole terms $a_i x^i$ is also called a monomial.

Implicit vs Explicit

As stated, the last definition is next to useless, certainly

$$(x-1)^{10}$$

"is" also a polynomial.

We can fix the problem by admitting all algebraic expressions that can be turned into a polynomial by expansion, and perhaps some simplifications.

That's OK, but from a computational perspective, there is a major issue: expanding out an implicit polynomial like the one above is not polynomial time.

The product $\prod_{i\leq 5}(x-a_i)$ expands to

$$x^5 - x^4a_1 - x^4a_2 + x^3a_1a_2 - x^4a_3 + x^3a_1a_3 + x^3a_2a_3 - x^2a_1a_2a_3 - x^4a_4 + x^3a_1a_4 + x^3a_2a_4 - x^2a_1a_2a_4 + x^3a_3a_4 - x^2a_1a_3a_4 - x^2a_2a_3a_4 + xa_1a_2a_3a_4 - x^4a_5 + x^3a_1a_5 + x^3a_2a_5 - x^2a_1a_2a_5 + x^3a_3a_5 - x^2a_1a_3a_5 - x^2a_2a_3a_5 + xa_1a_2a_3a_5 + x^3a_4a_5 - x^2a_1a_4a_5 - x^2a_2a_4a_5 + xa_1a_2a_4a_5 - x^2a_3a_4a_5 + xa_1a_3a_4a_5 + xa_2a_3a_4a_5 - a_1a_2a_3a_4a_5$$

In fact, expansion is connected to computational hardness, see below.

Suppose we have a univariate polynomial $p(x)$ over ring R.

We can replace the unknown x by elements in R and evaluate to obtain another element in the ring: For example,

$$p(x) = x^3 - 2x^2 + 3x - 1$$

produces

$$p(2) = 2^3 - 2 \cdot 2^2 + 3 \cdot 2 - 1 = 5.$$

Hence we can associate the polynomial p with a polynomial function

$$\widehat{p} : R \to R, \quad a \mapsto p(a)$$

This may seem like splitting hairs, but sometimes it is important to keep the two notions apart.

The polynomial and the associated polynomial function really are two different objects. Consider the ground ring $\mathbb{Z}_2$. The polynomial

$$p(x) = x + x^2$$

has the associated function

$$\widehat{p}(a) = 0$$

for all $a \in \mathbb{Z}_2$.

In fact, any polynomial $p(x) = \sum_{i \in I} x^i$ produces the identically 0 map as long as $I \subseteq \mathbb{N}^+$ has even cardinality.

Exercise

Describe all polynomial functions over ground rings $\mathbb{Z}_m$.
Start with m prime, then see what changes when m is composite.

Assume we have a polynomial in coefficient form $\boldsymbol{a} = (a_0, a_1, \ldots, a_d)$ we can easily evaluate at any point $b \in R$:

$$f(b) = ((\ldots(a_d\, b + a_{d-1})\, b + a_{d-2})\, b + \ldots)\, b + a_0$$

Actually, this is a perfect example of a polynomial represented in implicit form. This turns out to be very helpful for numerical evaluation.

Proposition

A polynomial of degree d can be evaluated in d ring multiplications and d ring additions.

One of the uses of polynomials is to fit data. Say we want $f(a_i) = b_i$ for $i = 0, \ldots, n$ where $a_i < a_{i+1}$. Define the Lagrange interpolant

$$L_i^n(x) = \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}$$

a polynomial of degree n that has a single root at a_i.

Proposition

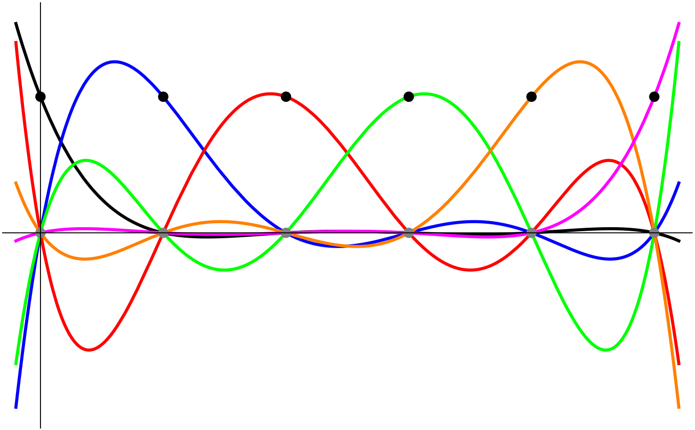
$L_i^n(a_i) = 1$ and $L_i^n(a_j) = 0$ for $i \neq j$.

Hence we can choose

$$f(x) = \sum_{i<n} b_i L_i^n(x)$$

Note that f has degree bound n.

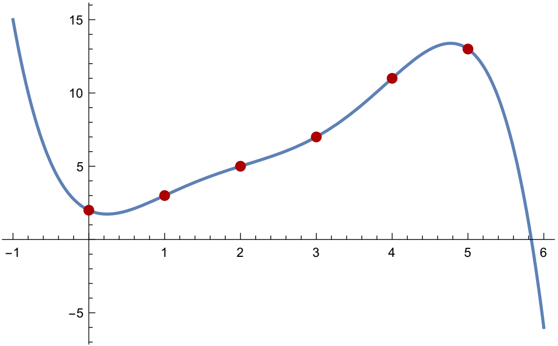
Lagrange Functions



Example: Primes

Suppose we want $f(i) =$ the ith prime for $i = 0, \ldots, 5$. The Lagrange interpolation looks like

$$f(x) = 2L_0^6 + 3L_1^6 + 5L_2^6 + 7L_3^6 + 11L_4^6 + 13L_5^6$$
$$= \frac{1}{120}(240 - 286\,x + 735\,x^2 - 425\,x^3 + 105\,x^4 - 9\,x^5)$$



Application: Secret Sharing

Suppose you have a “secret” a, a natural number, that you want to distribute over n people in such a way that no proper subgroup of the n persons can access the secret but the whole group can.

More generally may want to distribute the secret so that k out of n persons can access it, but not any subgroup of size $k-1$.

Here is a simple idea for the $n = k$ problem:

We may safely assume that a is a m-bit number. Generate $n-1$ m-bit numbers a_i and give number a_i to person i, $i = 1, \ldots, n-1$. Person n receives

$$a_n = a \oplus a_1 \oplus a_2 \oplus \ldots \oplus a_{n-1}$$

where $\oplus$ is bit-wise xor.

Clearly all n secret sharers can compute a, but if one is missing they are stuck with a random number.

A more flexible approach is built on the following idea. Construct a polynomial f that has a as its lowest coefficient, so $f(0) = a$. All other coefficients are chosen at random.

The secret sharers are given not the coefficients of f but pieces of the point-value description of f, which are obtained by repeated evaluation.

If all n agree, they can use their information to reconstruct the coefficient representation of f by interpolation.

Once f is reconstructed a can be found be a simple evaluation.

But for any proper subset f will be underdetermined, and a cannot be recovered.

It is not hard to generalize this method to $k < n$.

Shamir's Method

More precisely, pick a prime $p > a, n$. We will use the ground ring $\mathbb{Z}_p$.

Generate random numbers $0 < a_i < p$ for $i = 1, \ldots, n-1$.

Define the polynomial

$$f(x) = a + a_1 x + \ldots + a_{n-1} x^{n-1}$$

f is completely determined by the n point-value pairs (i, b_i), $i = 1, \ldots, n$.

By interpolation we can retrieve f from the point-value pairs, hence we can determine $a = a_0$.

An important point is that $n-1$ persons can obtain no information about the zero coefficient; every coefficient is equally likely.

The informal "definition" of a polynomial looks something like this.

> A polynomial is an expression of the form
>
> $$a_d x^d + a_{d-1} x^{d-1} + \ldots + a_1 x + a_0$$

No problem, right?

x is the mysterious "unknown," the x^i are monomials, the a_i are coefficients (from some ring), the $a_i x^i$ are terms and d is the degree (assuming $a_n \neq 0$).

If you come across an alleged definition of the kind

> A foobag is an expression of the form blah-blah-blah.

run for the hills.

Most of the time, all you get is one example of the expressions in question, often containing at least one ellipsis.

There is no explanation of the underlying language, no explanation of what exactly an expression is, certainly no formal grammar that defines everything.

Here is the "definition" of a real number from an otherwise great introductory analysis text.

Now we begin formally. What is a real number? It will be an expression of the form

$$\pm a_1 a_2 a_3 \ldots a_m . b_1 b_2 b_3 \ldots .$$

Here the $\pm$ represents a choice between plus and minus. The digit a_1 is an integer between 0 and 9 inclusive (unless m is different from 1 in which case it is restricted to being between 1 and 9, since it is the leading digit.) All other digits are integers between 0 and 9 inclusive.

Alas, there is nothing formal about this alleged definition. According to the author, a real number literally is an *expression*, an infinite string in this case. On that understanding, the real number π is just the string

$$+3.14159265358979323846264338327950288 4197\ldots$$

where we have omitted a few digits at the end to save paper.

Rant, Continued

Sorry, this is just malpractice. The reals are not a bunch of "expressions" like the ones in the definition. No way, never, ever.

The only justification for this line of attack is that the author does not want to get involved with a real definition, based on Cauchy sequences or Dedekind cuts. That's fine, the text is about analysis after all, but it should be clearly stated, with links to an actual definition.

And, for crying out loud, don't call it "formal," it's anything but.

Rant, Amended

Just to be clear, intuition is the power that drives everything. And, the StringWorld^{TM} approach can be helpful to get one's intuition going. In fact, it is often the right place to start. To paraphrase Knuth:

> Premature formalization is the root of all evil.

But, but, but . . .

One absolutely, totally cannot stop there. Wishy-washy land is where algorithms go to die–to compute, we need to build data types, and those need to refer to an actual definition, not just some vague appeal to intuition and analogy.

We start with a definition that tries to home in on the critical algebraic properties of a polynomial.

Let R be a commutative ring with 1 throughout.

Definition

Given a ring R, the $\mathbb{N}$-coproduct of R is defined by

$$\coprod_{\mathbb{N}} R = \{\, (a_n) \in R^{\mathbb{N}} \mid \text{only finitely many } a_n \neq 0 \,\}$$

An element of the coproduct is a sequence $(a_i)_{i \geq 0}$ of elements of R, where $a_n = 0$ for all $n \geq m$, for some threshold m.

So, in a way, we are still dealing with finite sequences.

Fudging it

Since almost all the terms a_n are 0, it makes sense to write

$$a_0 + a_1 x + \ldots + a_d x^d$$

instead of $(a_0, a_1, \ldots, a_d, 0, 0, 0, \ldots)$.

In fact, one often insists that a_d is the last non-zero element in the sequence (or $d = 0$ if they all are 0).

Note that the "unknown" x is nothing but syntactic sugar, all we really have is a sequence with finite support. We might as well use X, y, z, `fred`, whatever.

Again, there is no “unknown” in the definition of the coproduct. That makes it easier to give clean definitions of the algebraic structure. Addition is easy:

$$(a_n) + (b_n) = (a_n + b_n)$$

The sum $(a_n) + (b_n)$ is again an element of the coproduct and it is not too hard to check that this operation is associative and commutative.

But multiplication is somewhat more complicated (Cauchy product):

$$(a_n) \cdot (b_n) = \left(\sum_{i+j=n} a_i \cdot b_j \right)$$

Proposition

The product $(a_n) \cdot (b_n)$ is an element of the coproduct.

Write $\mathbf{0}$ and $\mathbf{1}$ for the sequences $(0, 0, 0, \ldots)$ and $(1, 0, 0, \ldots)$.

We have $\boldsymbol{a} + \mathbf{0} = \boldsymbol{a}$ so that

$$\left\langle \coprod R; +, \mathbf{0} \right\rangle$$

is a commutative monoid and even a group.

Likewise $\mathbf{1} \cdot \boldsymbol{a} = \boldsymbol{a} \cdot \mathbf{1} = \boldsymbol{a}$ and

$$\left\langle \coprod R; \cdot, \mathbf{1} \right\rangle$$

is also a commutative monoid (assuming that R is commutative).

The Unknown

Here is a much more interesting element: let

$$x = (0, 1, 0, 0, 0, \ldots)$$

Then $x^2 = (0, 0, 1, 0, 0, \ldots)$, $x^3 = (0, 0, 0, 1, 0, \ldots)$ and so forth.

This justifies the all the syntactic sugar in the notation

$$a_0 + a_1 x + \ldots + a_d x^d$$

instead of the actual coproduct element

$$(a_0, a_1, \ldots, a_d, 0, 0, \ldots)$$

We have quietly used the fact that we can embed all of R in the coproduct:

$$a \mapsto (a, 0, 0, \ldots)$$

Moreover, this map is trivially a ring monomorphism.

Lemma

$\langle \coprod R; +, \cdot, \mathbf{0}, \mathbf{1} \rangle$ is a ring. This ring is commutative whenever R is.

This is unsurprising, but note that a proof requires a bit of work: we have to verify e.g. that multiplication as defined above really is associative.

We ignore the details.

Definition

The ring $\langle \coprod R; +, \cdot, \mathbf{0}, \mathbf{1} \rangle$ is the polynomial ring with coefficients in R and is usually written $R[x]$.

In calculus one studies $\mathbb{R}[x]$.

For our purposes, $\mathbb{Q}[x]$, $\mathbb{Z}[x]$, $\mathbb{Z}_m[x]$ or $\mathbb{F}[x]$ where $\mathbb{F}$ is a finite field will be more important.

More Abstraction

The last definition may seem fairly abstract, but one can push even further ahead.

Note that $\mathbb{N}$ was not just used to define the coproduct (a sequence is a map $\mathbb{N} \to R$) but that addition on $\mathbb{N}$ was used in the definition of multiplication:

$$(a_i) \cdot (b_i) = (\sum_{r+s=i} a_r \cdot b_s)$$

What we are really using here is not just any old countable set but the monoid

$$\langle \mathbb{N}, +, 0 \rangle$$

The fact that the equation $r + s = i$ has only finitely many solutions in this monoid was crucial.

So? Everybody knows kindergarten arithmetic. Why make a fuss about it?

Suppose we have a commutative monoid

$$\langle M; +, 0\rangle$$

where equations $x + y = m$ have only finitely many solutions.

Then we can define a ring $R[M]$ on the carrier set

$$\coprod_M R = \{\, \boldsymbol{a} : M \to R \mid \text{only finitely many non-zeros} \,\}$$

in the exact same way as before.

We need at least one example for M (other then $\mathbb{N}$) that makes this construction interesting. It is probably good to stay close to $\mathbb{N}$. $\mathbb{Z}$ won't work since it violates the finiteness conditions.

How about $M = \mathbb{N} \times \mathbb{N}$?

Let $p, q \in \coprod_M R$. For any $k \in M$ we have

$$(p+q)(k) = p(k) + q(k)$$
$$(p \cdot q)(k) = \sum_{i+j=k} p(i)q(j)$$

For a generator r of M we get a "variable"

$$x_r(k) = \begin{cases} 1 & \text{if } r = k \\ 0 & \text{otherwise.} \end{cases}$$

OK, but we need at least one example for M (other then $\mathbb{N}$) that makes this construction interesting. It is probably good to stay close to $\mathbb{N}$. $\mathbb{Z}$ won't work since it violates the finiteness conditions.

Instead of sequences we now have a grid of coefficients (a two-dimensional array instead of a one-dimensional one).

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & \dots & a_{0j} & \dots \\ a_{10} & a_{11} & a_{12} & \dots & a_{1j} & \dots \\ a_{20} & a_{21} & a_{22} & \dots & a_{2j} & \dots \\ & \vdots & & \ddots & & \\ a_{i0} & a_{i1} & a_{i2} & \dots & a_{ij} & \dots \\ & \vdots & & & & \end{pmatrix}$$

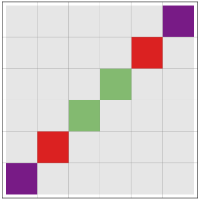
Addition in $R[\mathbb{N} \times \mathbb{N}]$ is essentially the same as before (since it does not depend on the algebraic structure of M, just the carrier set).

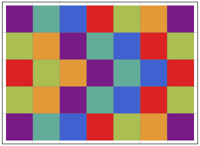
But multiplication becomes more interesting.

Multiplication

Let $x = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$ and $y = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$.

Then for $R = \mathbb{Z}_7$ the products $(x+y)^5$ and $(1+x)^5(2+y)^5$ look like this (all other entries are 0).





A Finite Monoid

Another benefit of the monoid approach: We could also use $M = \mathbb{Z}_m$ to build a polynomial ring.

Here there is one generator $1 \in \mathbb{Z}_m$ and the corresponding variable is

$$x(k) = \begin{cases} 1 & \text{if } k = 1 \\ 0 & \text{otherwise.} \end{cases}$$

One can check that

$$x^e(k) = 1 \iff k = e \bmod m$$

so that $x^m = x^0$: the powers wrap around.

We can get the same effect by factoring the ring $R[x]$ by the ideal $(x^m - 1)$.

Multivariate Polynomials

One can easily show that any element in $R[\mathbb{N} \times \mathbb{N}]$ has the form

$$\boldsymbol{a} = \sum_{i,j \geq 0} a_{i,j} x^i y^j$$

where only finitely many coefficients are non-zero and addition and multiplication work as one would expect. So we really have a rock-solid definition of bivariate polynomials. And, of course, using the monoid

$$M = \langle \mathbb{N}^n; +, 0 \rangle$$

we can get multivariate polynomials in general.

Definition

The ring of polynomials over R in n variables is defined by

$$R[x_1, \ldots, x_n] = R[\mathbb{N}^n]$$

Normal Form

There is a natural way to write down a multivariate polynomial analogous to the univariate and bivariate case. A monomial is a term of the form

$$\boldsymbol{x}^{\boldsymbol{e}} = x_1^{e_1} x_2^{e_2} \dots x_n^{e_n}.$$

where $\boldsymbol{e} = (e_1, \dots, e_n) \in \mathbb{N}^n$. Then any multivariate polynomial can be written as a sum of monomials, multiplied by the appropriate coefficients.

$$p(\boldsymbol{x}) = \sum a_{\boldsymbol{e}} \; \boldsymbol{x}^{\boldsymbol{e}}$$

And, arithmetic works as expected.

Note that the sum is naturally ordered (use the natural lex order on the exponents).

Note: Some authors consider the coefficient to be part of the monomial.

Definition

The degree of a monomial $\boldsymbol{x}^{\boldsymbol{e}}$ is $\sum e_i$. The degree of a polynomial is the largest degree of any of its monomials. If the polynomial is zero its degree is $-\infty$.

Lemma

Let R be an integral domain and $p, q \in R[x]$. Then $\deg(pq) = \deg(p) + \deg(q)$.

Note that if R be an integral domain then by the lemma $R[x]$ is again an integral domain.

The lemma fails without the integral domain assumption: Let $R = \mathbb{Z}_4$ and consider $p(x) = q(x) = 2x$.

Why not simply define, say, bivariate polynomials by saying they are expressions like

$$p(x, y) = x^2y^2 + 3x - y + 1$$

Or, if you want more precision, since the operation $R \mapsto R[x]$ works for any ring R, we can simply repeat it to get multivariate polynomials:

$$R[x_1, \ldots, x_n] = R[x_1, \ldots, x_{n-1}][x_n]$$

All true, but these "definitions" obscure a lot of things.

- What is $R[x][x]$, what $R[x, x]$?
- What is the difference between $R[x]$, $R[X]$ and $R[y]$?
- What is the difference between $R[x, y]$, $R[y, x]$?
- What is the role of evaluation?

Clarity

One might argue that $R[x][x]$ and $R[x, x]$ simply make no sense.

$R[x]$ and $R[y]$ are isomorphic, so in a sense they are the same.

For a human being (a mathematician) these difficulties are usually irrelevant: one can determine from context what is meant, or even correct the author if need be.

But if one wants to implement polynomial algebra things are more problematic. For example, if evaluation is done by rewrite rules, the name of the variable does play a huge role. Here $R[x]$ and $R[y]$ can behave very differently.

One Ring To Bind Them All

We can get around these question by sidestepping the "a polynomial is an expression . . . " model, and focusing on the abstract meaning of a polynomial ring in multiple variables.

To wit, we can implement $R[\mathbb{N}^n]$, and we need to do this just once: to connect this to the usual variables approach, one has to implement $R[\mathbb{N}^n]$ only once, and keep a list of variable names such as $x_1, \ldots, x_n$, $X_1, \ldots, X_n$, $z_1, \ldots, z_n$ and so on.

Again, this is just syntactic sugar.

We have avoided one important topic so far: What is evaluation?

The answer is clear in the term model: Replace the variables by concrete ring elements and then evaluate the corresponding ground term.

It is tempting to ask whether one can express this idea within our purely algebraic framework, without mentioning terms and substitutions.

An elegant way to do this is to think of evaluation as a ring homomorphism:

$$\begin{array}{rrcl} \mathsf{eval}: & R[x] & \longrightarrow & (R \to R) \\ & p & \longmapsto & \lambda r.p(r) \end{array}$$

Fine and good, but the critical point is that once we fix the value of the unknown(s), everything else is determined, too.

Uniqueness

More precisely, suppose $f : R \to S$ is some ring homomorphism. Consider the set of all ring homomorphisms from $R[x]$ to S that agree with f on R:

$$H = \{\, h : R[x] \to S \mid h \restriction R = f \,\}$$

Then the map

$$H \to S \qquad h \mapsto h(x)$$

is a bijection.

Universal Property

Another way of saying the same thing:

For any homomorphism $f : R \to S$ and any element $a \in S$ there is exactly one homomorphism $h : R[x] \to S$ such that

- $h \restriction R = f$ and
- $h(x) = a$.

Exercise

Give a similar result for multivariate polynomials.

Exercise

Show that this property in fact characterizes polynomial rings.

Definition

The degree of a monomial $\boldsymbol{x}^{\boldsymbol{e}}$ is $\sum e_i$. The degree of a polynomial is the largest degree of any of its monomials. If the polynomial is zero its degree is $-\infty$.

Lemma

Let R be an integral domain and $p, q \in R[x]$. Then $\deg(pq) = \deg(p) + \deg(q)$.

Note that if R be an integral domain then by the lemma $R[x]$ is again an integral domain.

The lemma fails without the integral domain assumption: Let $R = \mathbb{Z}_4$ and consider $p(x) = q(x) = 2x$.

Definition

A ring element $a \in R$ is a root of $p(x) \in R[x]$ if $p(a) = 0$.

In other words, a root is any solution of the equation $p(x) = 0$.

Finding roots of polynomial equations is often very difficult, in particular when several variables are involved.

For univariate polynomials over the reals good numerical methods exist, but over other rings things are problematic.

For example, computing square roots, i.e. solving $x^2 - a = 0$, over $\mathbb{Z}_m$ is surprisingly difficult. Of course there is a brute-force algorithm, but think of modulus m having thousands of digits.

And for $\mathbb{Z}[x_1, x_2, \ldots, x_n]$ it is even undecidable whether a root exists.

Definition

Given two polynomials f and g, g divides f if for some polynomial q: $q \cdot g = f$.

For the integers, the most important algorithm associated with the notion of divisibility is the Division Algorithm: we can compute quotient q and remainder r such that $a = qb + r$, $0 \leq r < b$.

The situation for polynomials is very similar.

Theorem (Division Algorithm)

Assume that F is a field. Let f and g be two univariate polynomials over F, $g \neq 0$. Then there exist polynomials q and r such that

$$f = q \cdot g + r \qquad \textit{where } \deg(r) < \deg(g).$$

Moreover, q and r are uniquely determined.

Proof Sketch

For existence consider the set of possible remainders

$$S = \{ f - qg \mid q \in F[x] \}.$$

If $\mathbf{0} \in S$ we are done, so suppose otherwise.

Trick: let $r \in S$ be any element of minimal degree, say $r = f - qg$.

Write $m = \deg(r)$ and $n = \deg(g)$, so we need $m < n$.

Assume $m \geq n$ and define

$$r' = r - a_m/b_n \, x^{m-n} \, g$$

where a_m and b_n are the leading coefficients of r and g, respectively.

But then $\deg(r') < \deg(r)$ and $r' \in S$, contradicting minimality.

Uniqueness is left as an exercise.

The Actual Algorithm

Though the theorem is often referred to as “Division Algorithm” it’s just an existence and uniqueness result. However, with a little work one can turn the proof into an algorithm.

Write $\mathsf{lc}(h)$ for the leading coefficient of a polynomial h. Suppose g is monic, so that $\mathsf{lc}(g) = 1$.

Here is an abstract version, to be called on f.

```
remainder( f, g ) {

      r = f;
      while( deg(r) >= deg(g) ) {
             c = lc( r );
             k = deg(r) - deg(g);
             r = r - c * x^k * g;
      }
      return r;
}
```

Another Version

Often one needs to compute both q and r, here is an array-based version which does that. Assume $\deg(f) = n$ and $\deg(g) = m$.

```
r = f;

for i = n-m downto 0 do
      if( deg r = m + i )
                q[i] = lc(r);
                r = r - q[i] * x^i * g;
      else
                q[i] = 0;

return  q[];
```

Note that sparseness is a tricky issue here: divide $x^{n+1} - 1$ by $x - 1$.

Application: GCD

An important application of the Division Algorithm for integers is the Euclidean algorithm for the GCD.

Likewise we can obtain a polynomial GCD algorithm from the Division Algorithm for polynomials.

In fact, essentially the same algorithm works, just replace $\mathbb{Z}$ by $\mathbb{Z}[x]$.

For example, we can obtain cofactors s and t such that

$$\gcd(f, g) = sf + tg.$$

Consider the polynomial

$$p(x) = x^4 - 8x^3 + 23x^2 - 28x + 12$$

We can read off immediately that $p(0) = 12$ and that the tangent at that point has slope -28.

But what about the polynomial around $x = 1$? Here it is convenient to use Taylor expansion to rewrite p in the form

$$p(x) = (x-1)^4 - 4(x-1)^3 + 5(x-1)^2 - 2(x-1)$$

We can see that $x = 1$ is a root and the tangent has slope -2.

More Local Views

Similarly

$$p(x) = (x-2)^4 - (x-2)^2$$
$$p(x) = (x-3)^4 + 4(x-3)^3 + 5(x-3)^2 + 2(x-3)$$

so it is clearly useful to consider polynomials in non-expanded form.

Exercise

What conclusions can you draw from the various representations of $p(x)$?

On rare occasions one can also easily establish bounds given the “right” representation of a polynomial. For instance,

$$p(x) = x^4 - 4x^3 + 7x^2 - 10x + 10$$

appears to be positive over $\mathbb{R}$.

One can prove this by tediously checking the behavior over the intervals with endpoints $-\infty$, -1, 1, $3/2$, 2, ∞. Alternatively, one can use differentiation to find the global minimum of p.

Or one can note that $p(x) = (x^2 - 1)^2 + (x - 3)^2$.

A Strange Automorphism

Consider some polynomial ring $R[x]$.

The discrete shift E is defined on $R[x]$ by

$$\mathsf{E}^n\, p(x) = p(x+1)$$

Claim: E is an automorphism of $R[x]$ that leaves R fixed.

As a consequence, p is irreducible iff $\mathsf{E}\, p$ is irreducible.

More generally, E preserves divisibility properties: $p \mid q$ iff $\mathsf{E}\, p \mid \mathsf{E}\, q$. We will use this fact countless times.

Proof

E is obviously compatible with the ring operations and leaves R fixed.

Let $p = \sum_{i \leq d} a_i x^i$. Then

$$\mathsf{E}\, p = \sum_{i \leq d} \left(\sum_{i \leq j \leq d} \binom{j}{i} a_j \right) x^i$$

Now suppose $\mathsf{E}\, p = \mathsf{E}\, q$, so $\deg(p) = \deg(q)$. Comparing the two sides starting with the leading coefficient, one can see that $p = q$.

$a_d = b_d$, $a_{d-1} + \binom{d}{1} a_d = b_{d-1} + \binom{d}{1} b_d$ and so on.

And Back . . .

For degree 4, the coefficient transformation is given by the matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and the inverse matrix is

$$\begin{pmatrix} 1 & -1 & 1 & -1 & 1 \\ 0 & 1 & -2 & 3 & -4 \\ 0 & 0 & 1 & -3 & 6 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Assorted Definitions

Let $p = (a_i)$ and $q = (b_i)$ be two polynomials over R.

The degree $\deg(p)$ is the maximum d for which $a_d \neq 0$.
For convenience, let $\deg(0) = -\infty$.

The leading coefficient $\mathsf{lc}(p)$ is a_d, $d = \deg(p)$.

p is monic if $\mathsf{lc}(p) = 1$.

q divides p, $q \mid p$, if $p = q \cdot s$ for some $s \in R[x]$.

p is irreducible if no polynomial of degree d, $1 < d < \deg(p)$, divides p.

Polynomials

Informally, a (univariate) polynomial is an expression of the form

$$x^3 - 2x^2 + 3x - 1$$

There is an unknown or variable x and we

- form powers x^i of the variable (monomials),
- multiply them by an element in some ground ring (coefficients),
- add several such terms.

Of course, division is not allowed.

The ground ring supplies the coefficients of the polynomials.

In the example, it is presumably $\mathbb{Z}$.

Hence we can represent a polynomial by a vector of its coefficients:

$$\boldsymbol{a} = (a_0, a_1, \dots, a_{n-1})$$

represents

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1}$$

If $a_{n-1} \neq 0$ then $n-1$ is the degree of p.

In general, $p(x)$ has degree bound n (note that this is a strict bound; this notion will be useful later in several algorithms).

Evaluation

Suppose we have a univariate polynomial $p(x)$ over ring R.

We can replace the unknown x by elements in R and evaluate to obtain another element in the ring: For example,

$$p(x) = x^3 - 2x^2 + 3x - 1$$

produces

$$p(2) = 2^3 - 2 \cdot 2^2 + 3 \cdot 2 - 1 = 5.$$

Hence we can associate the polynomial p with a polynomial function

$$\widehat{p} : R \to R, \quad a \mapsto p(a)$$

This may seem like splitting hairs, but sometimes it is important to keep the two notions apart.

The polynomial and the associated polynomial function really are two different objects. Consider the ground ring $\mathbb{Z}_2$. The polynomial

$$p(x) = x + x^2$$

has the associated function

$$\widehat{p}(a) = 0$$

for all $a \in \mathbb{Z}_2$.

In fact, any polynomial $p(x) = \sum_{i \in I} x^i$ produces the identically 0 map as long as $I \subseteq \mathbb{N}^+$ has even cardinality.

Exercise

Describe all polynomial functions over ground rings $\mathbb{Z}_2$ and $\mathbb{Z}_3$.

Horner's Rule

Assume we have a polynomial in coefficient form $\boldsymbol{a} = (a_0, a_1, \ldots, a_d)$ we can easily evaluate at any point $b \in R$:

$$f(b) = ((\ldots(a_d\, b + a_{d-1})\, b + a_{d-2})\, b + \ldots)\, b + a_0$$

Actually, this is a perfect example of a polynomial represented in compressed form.

Proposition

A polynomial of degree d can be evaluated in d ring multiplications and d ring additions.

Suppose we wish to construct a polynomial f that evaluates to given target values at certain points. Say we want $f(a_i) = b_i$ for $i = 0, \ldots, n-1$. Define the Lagrange interpolant

$$L_i^n(x) = \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}$$

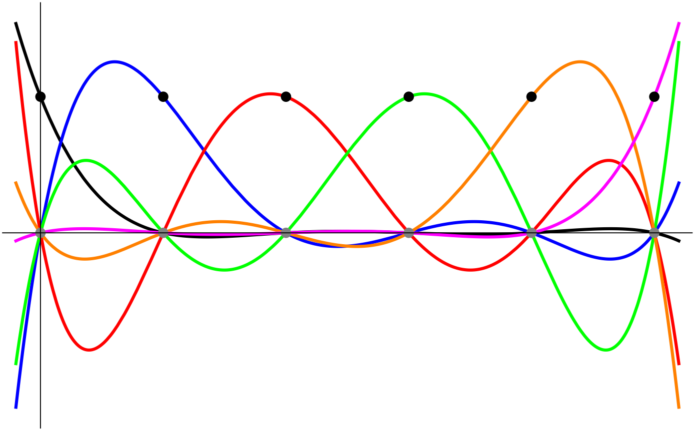
Proposition

$L_i^n(a_i) = 1$ *and* $L_i^n(a_j) = 0$ *for* $i \neq j$.

Hence we can choose

$$f(x) = \sum_{i<n} b_i L_i^n(x)$$

Note that f has degree bound n.



Example

Suppose we want $f(i) =$ the ith prime for $i = 0, \ldots, 5$. The Lagrange interpolation looks like

$$f(x) = 2L_0^6 + 3L_1^6 + 5L_2^6 + 7L_3^6 + 11L_4^6 + 13L_5^6$$

which, after expansion and simplification, produces

$$\frac{1}{120}(240 - 286\,x + 735\,x^2 - 425\,x^3 + 105\,x^4 - 9\,x^5)$$

Secret Sharing

Suppose you have a “secret” a, a natural number, that you want to distribute over n people in such a way that no proper subgroup of the n persons can access the secret but the whole group can.

More generally may want to distribute the secret so that k out of n persons can access it, but not any subgroup of size $k-1$.

Here is a simple idea for the $n = k$ problem:

We may safely assume that a is a m-bit number. Generate $n-1$ m-bit numbers a_i and give number a_i to person i, $i = 1, \ldots, n-1$. Person n receives

$$a_n = a \oplus a_1 \oplus a_2 \oplus \ldots \oplus a_{n-1}$$

where $\oplus$ is bit-wise xor.

Clearly all n secret sharers can compute a, but if one is missing they are stuck with a random number.

A more flexible approach is built on the following idea. Construct a polynomial f that has a as its lowest coefficient, so $f(0) = a$. All other coefficients are chosen at random.

The secret sharers are given not the coefficients of f but pieces of the point-value description of f, which are obtained by repeated evaluation.

If all n agree, they can use their information to reconstruct the coefficient representation of f by interpolation.

Once f is reconstructed a can be found be a simple evaluation.

But for any proper subset f will be underdetermined, and a cannot be recovered.

It is not hard to generalize this method to $k < n$.

More precisely, pick a prime $p > a, n$. We will use the ground ring $\mathbb{Z}_p$.

Generate random numbers $0 < a_i < p$ for $i = 1, \ldots, n-1$.

Define the polynomial

$$f(x) = a + a_1 x + \ldots + a_{n-1} x^{n-1}$$

f is completely determined by the n point-value pairs (i, b_i), $i = 1, \ldots, n$.

By interpolation we can retrieve f from the point-value pairs, hence we can determine $a = a_0$.

An important point is that $n-1$ persons can obtain no information about the zero coefficient; every coefficient is equally likely.

Lemma

Let $\mathbb{F}$ be a field and $f \in F[x]$.

If a is a root of f then $(x - a)$ divides $f(x)$.

If f is non-zero, then it has at most $\deg(f)$ many roots.

Proof.

Write
$$f = q(x - a) + r$$
where $\deg(r) < 1$. But then r must be 0, done.

Use the last result and induction on the degree.

$\square$

Decomposition

So if $\deg(f) = n$ and f has n roots we decompose f completely into linear terms:

$$f = c(x - a_1)(x - a_2) \dots (x - a_n)$$

Of course, there may be fewer roots, even over a rich field such as $\mathbb{R}$: $f = x^2 + 2$ has no roots.

This problem can be fixed by enlarging $\mathbb{R}$ to the field of complex numbers $\mathbb{C}$ (the so-called algebraic completion of $\mathbb{R}$).

Suppose $p(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0 \in \mathbb{C}[z]$ has complex roots $x_1, \ldots, x_n$, not necessarily distinct. Then

$$\sum_{\substack{I \subseteq [n] \\ |I| = k}} \prod_{i \in I} x_i = (-1)^k \frac{a_{n-k}}{a_n}$$

In particular

$$x_1 + \ldots + x_n = -a_{n-1}/a_n$$

$$x_1 \cdot \ldots \cdot x_n = (-1)^n a_0 / a_n$$

Proof. Expand the linear decomposition. □

Rational Root Test

For polynomials $p(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0$ with integral coefficients it is often necessarty to determing whether any of the roots are rational.

Plugging in a term p/q (in lowest common terms) one finds that

$$p \text{ divides } a_0$$
$$q \text{ divides } a_n$$

Thus, there are only finitely many cases to check. Unfortunately, this works only for univariate polynomials (it is not known whether the existence of rational solutions to a Diophantine equation is decidable).

More Roots

Note that over arbitrary rings more roots may well exist.

For example over $R = \mathbb{Z}_{15}$ the equation $x^2 - 4 = 0$ has four roots: $\{2, 7, 8, 13\}$.

But of course

$$(x-2)(x-7)(x-8)(x-13) = 1 + 7x^2 + x^4 \not\equiv x^2 - 4$$

Exercise

Using the Chinese Remainder theorem explain why there are four roots in the example above. Can you generalize?

Point-Value Representation

The fact that a non-zero polynomial of degree n can have at most n roots can be used to show that the interpolating polynomial

$$f(x) = \sum_i b_i \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}$$

is unique: suppose g is another interpolating polynomial so that $g(a_i) = b_i$.

Then $f - g$ has $n + 1$ roots and so is identically zero.

Hence we have an alternative representation for polynomials: we can give a list of point-value pairs rather than a list of coefficients.

To the naked eye this proposal may seem absurd: why bother with a representation that is clearly more complicated? As we well see, there are occasions when point-value is computationally superior to coefficient list.

Multiplying Polynomials

Suppose we have two univariate polynomials f and g of degree bound n.

Using the brute force algorithm (i.e., literally implementing the definition of multiplication in $\coprod R$) we can compute the product fg in $\Theta(n^2)$ ring operations.

Now suppose we are dealing with real polynomials. There is a bizarre way to speed up multiplication:

- Convert f and g into point-value representation where the support points are carefully chosen.
- Multiply the values pointwise to get h.
- Convert h back to coefficient representation.

It may seem absurd to spend all the effort to convert between coefficient representation and point-value representation. Surprisingly, it turns out that the conversions can be handled in $\Theta(n \log n)$ steps using a technique called Fast Fourier Transform.

But the pointwise multiplication is linear in n, so the whole algorithm is just $\Theta(n \log n)$.

Theorem

Two real polynomials of degree bound n can be multiplied in $\Theta(n \log n)$ steps.

Take a look at CLR for details.

Vandermonde Matrices

Here is another look at conversions between coefficient and point-value representation, i.e., between evaluation and interpolating.

Definition

Define the n by n Vandermonde matrix by

$$\mathsf{VM}(x_0, x_1, \ldots, x_{n-1}) = \begin{pmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^{n-1} \\ 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \ldots & x_{n-1}^{n-1} \end{pmatrix}$$

Lemma

$$|\mathsf{VM}(\boldsymbol{x})| = \prod_{i<j} x_j - x_i$$

It follows that the Vandermonde matrix is invertible iff all the x_i are distinct. Now consider a polynomial

$$f(x) = c_0 + c_1 x + \ldots + c_{n-1} x^{n-1}$$

To evaluate f at points $\boldsymbol{a} = (a_0, \ldots, a_{n-1})$ we can use matrix-by-vector multiplication:

$$\boldsymbol{b} = \mathsf{VM}(\boldsymbol{a}) \cdot \boldsymbol{c}$$

But given the values $\boldsymbol{b}$ we can obtain the coefficient vector by

$$\boldsymbol{c} = \mathsf{VM}(\boldsymbol{a})^{-1} \cdot \boldsymbol{b}$$

Polynomial Identities

A task one encounters frequently in symbolic computation is to check whether two polynomials are equivalent, i.e., whether an equation between polynomials

$$f(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_n)$$

holds in the sense that for all $x_1, x_2, \ldots, x_n \in R$.

Definition

Two polynomials $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are equivalent if for all $\boldsymbol{a} \in R$: $f(\boldsymbol{a}) = g(\boldsymbol{a})$. In particular f is identically zero if $f(\boldsymbol{x})$ and 0 are equivalent.

In other words, two polynomials $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are equivalent iff $\widehat{f}(\boldsymbol{x}) = \widehat{g}(\boldsymbol{x})$.

Notation: $f(\boldsymbol{x}) \equiv g(\boldsymbol{x})$.

Zero Testing

Note that the polynomial identity $f \equiv g$ can be rewritten as $f - g \equiv 0$.

Problem:

How can we check whether a multivariate polynomial $f \in F[\boldsymbol{x}]$ is identically zero?

Note: The polynomial may not be given in normal form, but as in the example in a much shorter, parenthesized form. We want a method that is reasonably fast without having to expand out the polynomial first.

Using Roots

Assume that the ground ring is a field F.

Suppose f has degree d. As we have seen, in the univariate case there are at most d roots and d roots determine the polynomial (except for a multiplicative constant).

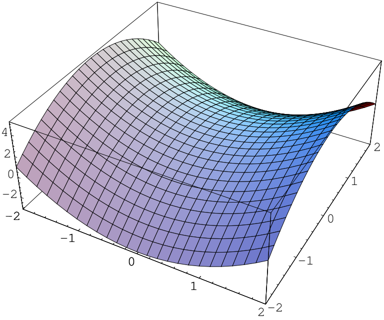
So we could simply check if the polynomial vanishes at, say, $a = 0, 1, 2, \ldots, d$: f will vanish on all $d + 1$ points iff it is identically zero.

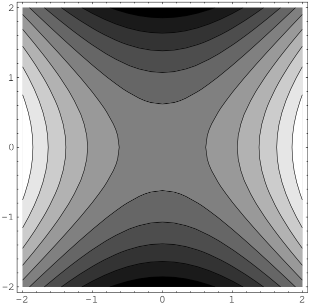
Requires only $d + 1$ evaluations of the polynomial (in any form).

Unfortunately, the roots for multivariate polynomials are a bit more complicated.

A Quadratic Surface

$$x^2 - y^2 + 1$$





Lemma

Let $f \in F[x_1, \dots, x_n]$ be of degree d and $S \subseteq F$ a set of cardinality s. Then f is not identically zero then f has at most ds^{n-1} roots in S^n.

Proof.

The proof is by induction on n.

$\square$

The set S here could be anything. For example, over $\mathbb{Q}$ we might choose $S = \{0, 1, 2, \dots, s-1\}$.

The main application of the lemma is to give a probabilistic algorithm to check whether a polynomial is zero.

Suppose f is not identically zero and has degree d.

Choose a point $\boldsymbol{a} \in S^n$ uniformly at random and evaluate $f(\boldsymbol{a})$. Then

$$\mathsf{Pr}[\, f(\boldsymbol{a}) = 0\,] \leq \frac{d}{s}$$

So by selecting S of cardinality $2d$ the error probability is $1/2$.

Note that the number of variables plays no role in the error bound.

With more work we can make sure the f really vanishes.

Corollary

Let $f \in F[x_1, \ldots, x_n]$ be of degree d and $S \subseteq F$ a set of cardinality $s > d$. If f vanishes on S^n then f is identically zero.

But note that for finite fields we may not be able to select a set of cardinality higher than d.

Recall the example over $\mathbb{Z}_2$: in this case we can essentially only choose $S = \mathbb{Z}_2$, so only degree 1 polynomials can be tackled by the lemma.

That's fine for univariate polynomials (since any monomial x^i simplifies to x) but useless for multivariate polynomials.

Application

Recall that a perfect matching in a bipartite graph is a subset M of the edges such that the edges do not overlap and every vertex is incident upon one edge in M.

There is a polynomial time algorithm to check whether a perfect matching exists, but using Schwartz's lemma one obtains a faster and less complicated algorithm.

Suppose the vertices are partitioned into u_i and v_i, $i = 1, \ldots, n$.

Define a $n \times n$ matrix A by

$$A(i,j) = \begin{cases} x_{ij} & \text{if } (u_i, v_j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the determinant of A is a polynomial whose non-zero terms look like

$$\pm x_{1\pi(1)} x_{2\pi(2)} \cdots x_{n\pi(n)}$$

Testing Perfect Matchings

Proposition

The graph has a perfect matching iff the determinant of A is not identically zero.

Proof.

If the graph has no perfect matching then all the terms in the determinant are 0 (they all involve at least one non-edge).

But if the graph has a perfect matching it must have the form $M = \{ (u_i, v_{\pi(i)}) \mid i \in [n] \}$ where π is a permutation.

But then the determinant cannot be 0 since the corresponding monomial cannot be canceled out.

□