MT1010: Topics in Mathematics: Polynomials

Nik Ruškuc

October 26, 2014

1. Definition, terminology and notation

A *polynomial* is an expression of the form

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where

- $n \geq 0$ is a non-negative integer;
- $a_0, a_1, \ldots, a_n$ are (fixed) numbers
- x is a variable (simply a letter, with no specific value assigned to it).

In fact, we could just as well define a polynomial as a tuple

$$(a_n, a_{n-1}, \ldots, a_1, a_0).$$

In addition, we require $a_n \neq 0$ if $n > 0$.

For a polynomial $p(x)$ as above:

- $a_0, \ldots, a_n$ are its *coefficients*;
- the coefficient $a_n \neq 0$ is the *leading coefficient*;
- a_0 is the *constant coefficient*;
- the number n is the *degree* of $p(x)$, denoted $\deg(p(x))$.

Polynomials of low degrees have special names:

deg	name
0	constant
1	linear
2	quadratic
3	cubic

2. Operations

At the moment a polynomial is just a formal expression or a sequence. If we want to do some mathematics with polynomials we need to perform some operations with or on them. In order to do this, we first need to be able to perform similar operations on coefficients. Mostly, we will assume that the coefficients come from the sets of rationals $\mathbb{Q}$, reals $\mathbb{R}$ or complex numbers $\mathbb{C}$. These are the most basic examples of what mathematicians call a *field* – a system in which basic arithmetic operations are possible, including division by arbitrary non-zero elements, and they obey the usual rules of arithmetic.

The basic arithmetic operations of addition, multiplication and subtraction are defined by treating x as if itself was a number, and using the basic arithmetic rules for numbers. Here are a few illustrative examples:

$$(x^2 - 3x + 1) + (2x^3 - 5x^2 - 7x + 3) = 2x^3 - 4x^2 - 10x + 4,$$

$$\begin{aligned} & (x^2 - 3x + 1)(2x^3 - 5x^2 - 7x + 3) \\ = \; & 2x^5 - 5x^4 - 7x^3 + 3x^2 - 6x^4 + 15x^3 + 21x^2 - 9x + 2x^3 - 5x^2 - 7x + 3 \\ = \; & 2x^5 - 11x^4 + 10x^3 + 19x^2 - 16x + 3, \end{aligned}$$

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

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

(Can you generalise the last two equalities?)

$$(x + 1)^2 = (x + 1)(x + 1) = x^2 + x + x + 1 = x^2 + 2x + 1,$$

$$\begin{aligned} (x + 1)^3 = (x + 1)(x + 1)^2 &= (x + 1)(x^2 + 2x + 1) = x^3 + 2x^2 + x + x^2 + 2x + 1 \\ &= x^3 + 3x^2 + 3x + 1. \end{aligned}$$

(Can you generalise these two?)

The ratio of two polynomials is very seldom again a polynomial. We will discuss the division of polynomials in more detail later.

A quicker, symbolic way for multiplying polynomials, which dispenses with the multitude of occurrences of x, is illustrated below, yet again for computing the product

$$(x^2 - 3x + 1)(2x^3 - 5x^2 - 7x + 3) = 2x^5 - 11x^4 + 10x^3 + 19x^2 - 16x + 3,$$

			1	−3	1
		2	−5	−7	3
		2	−5	−7	3
	−6	15	21	−9	
2	−5	−7	3		
2	−11	10	19	−16	3

The *composition* $p(x) \circ q(x)$ of two polynomials is defined by

$$p(x) \circ (q(x) = p(q(x)),$$

i.e. it is the result of substitution of every occurrence of x in $p(x)$ by $q(x)$ (and expanding).

For example, let us evaluate the composition of

$$p(x) = 2x^2 - 5x + 8 \text{ and } q(x) = 3x - 4.$$

We have

$$\begin{aligned} p(x) \circ q(x) &= p(q(x)) = 2(3x-4)^2 - 5(3x-4) + 8 \\ &= 18x^2 - 63x + 60. \end{aligned}$$

The operation of taking the *derivative* is borrowed from calculus: for a polynomial

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

its derivative is

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

It is a fortuitous, but very important fact that the derivative of a polynomial is again a polynomial. Note that the degree of the derivative is one less than the degree of the original polynomial.

3. Polynomial functions

Let us consider a polynomial

$$p(x) = a_n x^n + \cdots + a_1 x + a_0$$

with real coefficients. If $t \in \mathbb{R}$ is any number, we can substitute it for x and evaluate the resulting expression to obtain the *value of* $p(x)$ *at* t:

$$p(t) = a_n t^n + \cdots + a_1 t + a_0 \in \mathbb{R}.$$

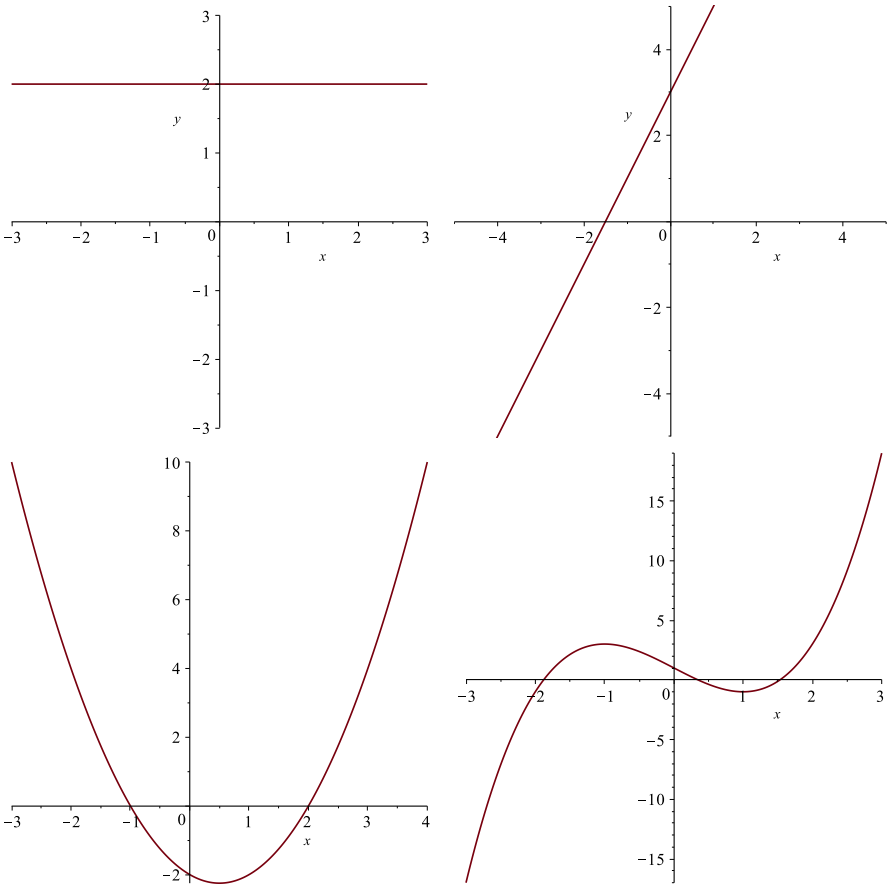

Figure 1: Graphs for typical constant (horizontal line), linear (sloped line), quadratic (parabola) and cubic polynomials

Since we can do this for *any* number t, it follows that our polynomial $p(x)$ gives rise to a function $\mathbb{R} \to \mathbb{R}$. This function is normally also denoted as $p(x)$, but one has to beware possible confusion this may cause at times.

In Figure 1 graphs of functions induced by a typical constant, linear , quadratic and cubic polynomials are plotted.

Theorem 3.1. *Let $p(x)$ be a real polynomial, and let (a, b) be any open interval of real numbers. If $p(t) = 0$ for all $t \in (a, b)$ then $p(x)$ is in fact the zero polynomial.*

Proof. Let us rephrase the assertion as follows: If $p(x)$ is any polynomial of degree $n \geq 1$ then there exists $t \in (a, b)$ such that $p(t) \neq 0$. We prove this assertion by induction on n. It is obvious for linear polynomials, since every such polynomial equals 0 for precisely *one* real number. Suppose the assertion is true for *all* polynomials of degree n. Let $p(x)$ be a polynomial of degree $n + 1$. Suppose $p(t) = 0$ for all $t \in (a, b)$. Then $p'(t) = 0$ for all $t \in (a, b)$. But $p'(x)$ is a polynomial of degree n, a contradiction. ■

Corollary 3.2. *Let $p(x)$, $q(x)$ be two real polynomials, and let (a, b) be an interval. If $p(t) = q(t)$ for all $t \in (a, b)$ then $p(x) = q(x)$ as polynomials.*

Proof. Let $r(x) = p(x) - q(x)$, and apply the previous theorem. ■

Let us also record the behaviour of polynomial functions at $\pm\infty$:

Theorem 3.3. *Let $p(x)$ be a polynomial of degree $n \geq 1$ with leading coefficient $a_n > 0$. Then*

$$\begin{aligned} \lim_{t \to +\infty} p(t) &= +\infty \\ \lim_{t \to -\infty} p(t) &= \begin{cases} -\infty & \text{if } n \text{ is odd} \\ +\infty & \text{if } n \text{ is even.} \end{cases} \end{aligned}$$

For $a_n < 0$ all the signs should be changed.

We will not prove this here. The proof is not hard, but it does require a reasonably detailed analysis of inequalities.

4. Division of polynomials

As we said before, the quotient of two polynomials is generally not a polynomial. It is a remarkable fact that, like integers, polynomials can be divided by quotient and reminder.

Theorem 4.1. *For any two polynomials $p(x)$, $u(x)$ there exists unique polynomials $q(x)$ (the quotient) and $r(x)$ (the remainder) such that*

$$p(x) = u(x)q(x) + r(x),$$
$$\deg(r(x)) < \deg(u(x)).$$

The proof of this is not actually hard, but it does involve lots of symbols (the coefficients for all the polynomials featuring in the statement). Working through an explicit example is just as instructive.

Let us attempt to divide *exactly* the polynomial

$$p(x) = 2x^4 - x^3 - 12x^2 + 9x + 5$$

by the polynomial

$$u(x) = 2x^2 + 5x - 1.$$

Thus we are looking for a polynomial $q(x)$ such that

$$p(x) = u(x)q(x). \tag{1}$$

Clearly, $q(x)$ will have to be quadratic, say

$$q(x) = ax^2 + bx + c.$$

We want to determine a, b, c. Rewrite (1) as

$$\begin{aligned} 2x^4 - x^3 - 12x^2 + 9x + 5 &= (2x^2 + 5x - 1)(ax^2 + bx + c) \\ &= 2ax^4 + (2b + 5a)x^3 + (2c + 5b - a)x^2 \\ &\quad +(5c - b)x + (-c + 7). \end{aligned}$$

Equating the corresponding coefficients yields

$$\begin{aligned} 2a &= 2 \\ 2b + 5a &= -1 \\ 2c + 5b - a &= -12 \\ 5c - b &= 9 \\ -c + 7 &= 5. \end{aligned}$$

The first three equations yield:

$$a = 1,\ b = -3,\ c = 2.$$

However, with these values we see that the final two equations fail to hold. Thus, at this point we have proved that the desired $q(x)$ does not exist. However, if we do let

$$q(x) = x^2 - 3x + 2$$

we know that $p(x)$ and $u(x)q(x)$ will agree on terms of degree $4, 3, 2$. Thus their difference $p(x) - u(x)q(x)$ will be linear; indeed:

$$p(x) - u(x)q(x) = -4x + 7.$$

Setting $r(x) = -4x + 7$ we have obtained

$$p(x) = u(x)q(x) + r(x),\ \deg(r(x)) = 1 < 2 = \deg(u(x)).$$

The division process can be performed much faster symbolically:

$$\begin{array}{cccc|cccc}
 & & & & & & 1 & -3 & 2 \\
\hline
2 & 5 & -1 & | & 2 & -1 & -12 & 9 & 5 \\
 & & & & 2 & 5 & -1 & & \\
\hline
 & & & & & -6 & -11 & 9 & 5 \\
 & & & & & -6 & -15 & 3 & \\
\hline
 & & & & & & 4 & 6 & 5 \\
 & & & & & & 4 & 10 & -2 \\
\hline
 & & & & & & & -4 & 7
\end{array}$$

Corollary 4.2. *The remainder of dividing a polynomial $p(x)$ with a linear polynomial of the form $x - a$ is $p(a)$, the value of p at a.*

Proof. Since we are dividing with a linear polynomial, the remainder must have degree 0, i.e. be constant. So we write

$$p(x) = (x - a)q(x) + r.$$

Substitute $x = a$, to obtain

$$p(a) = (a - a)q(a) + r \Rightarrow r = a,$$

as required. ■

5. Roots of polynomials

Definition 5.1. A number a is a *root* (or a *zero*) of a polynomial $p(x)$ if $p(a) = 0$.

Theorem 5.2. *The number a is a root of a polynomial $p(x)$ if and only if $x - a$ divides $p(x)$ exactly (i.e. without remainder).*

Proof. This follows immediately from Corollary 4.2. ■

Corollary 5.3. *If $a_1, \ldots, a_k$ are distinct roots of a polynomial $p(x)$ then*

$$p(x) = (x - a_1)(x - a_2) \ldots (x - a_k)q(x)$$

for some polynomial $q(x)$.

Proof. A repeated application of Theorem 5.2. ■

Corollary 5.4. *A polynomial of degree $n \geq 1$ cannot have more than n distinct roots.*

Proof. This follows from Corollary 5.3 and the fact that

$$\deg((x-a_1)\dots(x-a_k)q(x)) = k + \deg(q(x)) \geq k.$$

■

Corollary 5.5. *If two polynomials $p(x)$, $q(x)$ of degree n agree on $n+1$ distinct values, then they are identical as polynomials.*

Proof. Suppose that $t_1, \dots, t_{n+1}$ are distinct numbers such that

$$p(t_i) = q(t_i),\ i = 1, \dots, n+1.$$

Let

$$u(x) = p(x) - q(x),$$

a polynomial of degree $\leq n$. All $t_1, \dots, t_{n+1}$ are zeros of $u(x)$. It follows from Corollary 5.4 that $u(x)$ is the zero polynomial, and hence $p(x) = q(x)$. ■

How many roots a polynomial *actually* has depends on the field we are working in. For instance, the polynomial $x^2 - 2$ has no roots in $\mathbb{Q}$ (why?), but does have real roots $\pm\sqrt{2}$. Similarly, $x^2 + 1$ has no real roots, but does have two complex roots $\pm i$.

In fact, complex numbers have the following remarkable property:

Theorem 5.6 (The Fundamental theorem of Algebra) *Every non-constant polynomial with complex coefficients has a root in $\mathbb{C}$.*

The proof of this result is beyond the scope of this course, but it may be proved in a subsequent course.

Corollary 5.7. *Every complex polynomial factorises into a product of linear polynomials.*

This property is known as being *algebraically closed*.

We should clearly distinguish between the existence of roots and our ability to find them:

Theorem 5.8. *There is no expression involving the basic arithmetics operations and kth roots that would give the general solution to the equation*

$$a_n x^n + \dots + a_1 x + a_0 = 0$$

for $n \geq 5$ in terms of the coefficients $a_0, \dots, a_n$.

This is also beyond our scope, but will be proved in the Level 5 module on Galois Theory. Of course, for $n = 1, 2$ such formulae exist and are well known. They also exist, but are much more complicated, for $n = 3, 4$. On the other hand, there exist efficient numerical algorithms approximating roots of arbitrary polynomials to an arbitrary degree of precision; these are studied in Numerical Analysis.

It is well known that for any two distinct points in a plane there is a unique straight line passing through them. One way of re-interpreting this is that for any two $(x_1, y_1), (x_2, y_2)$ with $x_1 \neq x_2$ there exists a unique linear polynomial $p(x)$ such that

$$p(x_i) = y_i, \ i = 1, 2.$$

This can be generalised to arbitrary degrees as follows:

Theorem 5.9. *For any $(n+1)$ points*

$$(x_1, y_1), \ldots, (x_{n+1}, y_{n+1})$$

with $x_i \neq x_j$ $(i \neq j)$ there exists a unique polynomial $p(x)$ of degree at most n such that

$$p(x_i) = y_i, \ i = 1, \ldots, n+1.$$

Proof. [Sketch] We are looking for

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

such that

$$p(x_i) = y_i, \ i = 1, \ldots, n+1.$$

This can be written as a system of equations:

$$\begin{array}{rcccccc} a_n x_1^n & +a_{n-1} x_1^{n-1} & \ldots & +a_1 x_1 & +a_0 & = & y_1 \\ a_n x_2^n & +a_{n-1} x_2^{n-1} & \ldots & +a_1 x_2 & +a_0 & = & y_2 \\ \vdots & & & & & & \vdots \\ a_n x_{n+1}^n & +a_{n-1} x_{n+1}^{n-1} & \ldots & +a_1 x_{n+1} & +a_0 & = & y_{n+1} \end{array}$$

Note that we want to solve this system for $a_0, \ldots, a_{n+1}$, the coefficients of $p(x)$. So, there are $n+1$ variables, and $n+1$ equations. The determinant of the system is

$$d = \det \begin{pmatrix} x_1^n & x_1^{n-1} & \ldots & x_1 & 1 \\ x_2^n & x_2^{n-1} & \ldots & x_2 & 1 \\ \vdots & & & & \vdots \\ x_{n+1}^n & x_{n+1}^{n-1} & \ldots & x_{n+1} & 1 \end{pmatrix}$$

This is the so-called *Vandermonde* determinant, and can be shown to be equal

$$d = \prod_{1 \leq i < j \leq n+1} (x_j - x_i).$$

Since $x_i \neq x_j$ for $i \neq j$, it follows that $d \neq 0$, and hence the system has a unique solution. ■