CHAPTER TWO

THE TAYLOR EXPANSION

§1. HORNER'S METHOD

A polynomial

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

with its coefficents a_k drawn from a field $\mathbf{F}$ is evaluated at an element u in the field by replacing each occurrence of the indeterminate x by u to obtain

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

The obvious way to calculate $p(u)$ is to evaluate each term of the polynomial and then add the terms together. This is inefficient; a superior method, due to Horner, is suggested by writing the polynomial in nested form:

$$p(x) = (\cdots(((a_n x + a_{n-1})x + a_{n-2})x + a_{n-3})x + \cdots + a_1)x + a_0\ .$$

We can evaluate $p(u)$ recursively. Let $b_0 = a_n$, and $b_k = b_{k-1}u + a_{n-k}$ for $(1 \le k \le n)$. Then $p(u) = b_n$.

This process can be clarified through a table:

a_n	a_{n-1}	a_{n-2}	$\cdots$	a_2	a_1	a_0
	$b_0 u$	$b_1 u$	$\cdots$	$b_{n-3}u$	$b_{n-2}u$	$b_{n-1}u$
$a_n = b_0$	b_1	b_2	$\cdots$	b_{n-2}	b_{n-1}	b_n

Each element in the bottom row is the sum of the two above it.

For example, to evaluate the polynomial $4x^5 - 7x^4 + 6x^3 + 2x^2 - x + 3$ at 2, we can construct the table

4	−7	6	2	−1	3
	8	2	16	36	70
4	1	8	18	35	73

and moreover use it to obtain the representation

$$4x^5 - 7x^4 + 6x^3 + 2x^2 - x + 3 = (4x^4 + x^3 + 8x^2 + 18x + 35)(x - 2) + 73\ .$$

The last number, 73, of the bottom row is the value of the polynomial at 2.

More generally,

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = (b_0 x^{n-1} + b_1 x^{n-2} + \cdots + b_{n-2}x + b_{n-1})(x - u) + b_n\ ,$$

where $b_0 = a_n$, $b_1 = a_n u + a_{n-1}$, $b_2 = a_n u^2 + a_{n-1}u + a_{n-2}$, $\cdots$, $b_{n-1} = a_n u^{n-1} + a_{n-1}u^{n-2} + \cdots + a_2 u + a_1$, and $b_n = a_n u^n + a_{n-1}u^{n-1} + \cdots + a_1 u + a_0$, the value of the polynomial at u.

Thus, we have a constructive way of realizing the Remainder Theorem,

$$p(x) = p_1(x)(x - u) + p(u)\ ,$$

for a suitable polynomial $p_1(x)$ whose degree is $n - 1$, less by one than the degree of $p(x)$.

We can continue the Horner process to obtain

$$p_1(x) = p_2(x)(x-u) + p_1(u) \ ;$$

$$p_2(x) = p_3(x)(x-u) + p_2(u) \ ;$$

$$\cdots \ ;$$

$$p_{n-1}(x) = p_n(x)(x-u) + p_{n-1}(u) \ ,$$

where $p_k(x)$ is a polynomial of degree $n-k$, and in particular, $p_n(x)$ is a constant polynomial. This leads to the representation

$$p(x) = p_n(u)(x-u)^n + p_{n-1}(x-u)^{n-1} + p_{n-2}(x-u)^{n-2} + \cdots + p_1(u)(x-u) + p(u) \ .$$

In the case of the example $4x^5 - 7x^4 + 6x^3 + 2x^2 - x + 3$, we get the table

4	−7	6	2	−1	3
	8	2	16	36	70
4	1	8	18	35	73
	8	18	52	140	
4	9	26	70	175	
	8	34	120		
4	17	60	190		
	8	50			
4	25	110			
	8				
4	33				
4					

from which we can read off the identity

$$4x^5 - 7x^4 + 6x^3 + 2x^2 - x + 3 = 4(x-2)^5 + 33(x-2)^4 + 110(x-2)^3 + 190(x-2)^2 + 175(x-2) + 73 \ .$$

Return to the general case and the representation

$$p(x) = p_1(x)(x-u) + p(u)$$

where

$$\begin{aligned} p_1(x) &= a_n x^{n-1} + (a_n u + a_{n-1})x^{n-2} + (a_n u^2 + a_{n-1}u + a_{n-2})x^{n-3} + \cdots \\ &\quad + (a_n u^{n-2} + a_{n-1}u^{n-3} + \cdots + a_2)x + (a_n u^{n-1} + a_{n-1}u^{n-2} + \cdots + a_2 u + a_1) \\ &= a_n(x^{n-1} + x^{n-2}u + x^{n-3}u^2 + \cdots + u^{n-1}) + a_{n-1}(x^{n-2} + x^{n-3}u + \cdots + u^{n-2}) \\ &\quad + \cdots + a_2(x+u) + a_1 \ . \end{aligned}$$

Then $p_1(u) = na_n u^{n-1} + (n-1)u^{n-2} + \cdots + 2a_2 u + a_1$.

For any polynomial $p(x) = \sum_{k=0}^{n} a_k x^k$, we define the derivative $p'(x) = \sum_{k=1}^{n} k a_k x^{k-1}$. Note that this definition is made algebraically by identifying the coefficients of $p'(x)$ in terms of those of $p(x)$; there is no need to use the concept of limit. From this definition, the usual rules of the derivative with respect to derivatives of sums, products and composition of polynomials can be demonstrated. We can iterate differentiation to obtain derivatives of higher order. As differentiation reduces the degree of a polynomial by one, a polynomial is of degree n if and only if its nth derivative is a nonzero constant.

Since $p(x) = \sum_{k=0}^{n} p_k(u)(x-u)^k$, we have that

$$p'(x) = \sum_{k=1}^{n} k p_k(u)(x-u)^{k-1} \ ,$$

and, more generally,

$$p^{(r)}(x) = \sum_{k=r+1}^{n} k(k-1)\cdots(k-r+1)(x-u)^{k-r} + r!p_r(u) \ ,$$

for $1 \leq r \leq n$. Substituting u for x, we find that $p^{(r)}(u) = r!p_r(u)$, so that

$$p(x) = \sum_{k=0}^{n} \frac{p^{(k)}(u)}{k!}(x-u)^k \ .$$

A zero r of the polynomial $p(z)$ has multiplicity m if and only if $p(z)$ can be written in the form $p(z) = (z-r)^m q(z)$ where $q(r) \neq 0$. From the Taylor explansion, we see that this is equivalent to $p(r) = p'(r) = \cdots = p^{(m-1)}(r) = 0$ and $p^{(m)}(r) \neq 0$.

§2.NEWTON'S METHOD

The Taylor expansion of a polynomial about a point u can be used to motivate a standard method for approximating roots of a given polynomial equation $p(x) = 0$, *i.e.* a number r for which $p(r) = 0$. We have

$$0 = p(r) = p(u) + (r-u)p'(u) + \frac{1}{2}(r-u)^2 p''(u) + \cdots \ .$$

If we have a rough idea where r is located, and select u nearby, taking care to ensure that $p'(u)$ does not vanish, then we can approximate the value of r by the equation

$$0 = p(u) + (r-u)p'(u) \ .$$

This is equivalent to

$$r = u - \frac{p(u)}{p'(u)} \ .$$

This, of course, does not give the actual value of the root, but one might hope that it is a closer approximation.

Thus, we can define a sequence of approximants by taking $u_0 = u$ and, for each $n \geq 0$,

$$u_{n+1} = u_n - \frac{p(u_n)}{p'(u_n)} \ .$$

For real polynomials, this can be illustrated geometrically. Locate the point $(u_n, p(u_n))$ on the graph of the curve $y = p(x)$ in the cartesian plane. The tangent to the curve through this point meets the $x-$axis at the point $(u_{n+1}, 0)$. One can illustrate with a sketch that for certain shapes of the curve, the point $(u_{n+1}, 0)$ will be closer to the point $(r, 0)$ on the curve than $(u_n, 0)$.

We can illustrate how the sequence of approximants proceeds by looking at a quadratic polynomial. The behaviour in the complex plane for the quadratic $p(z) = z^2 - 1$ is typical. The recursion for the sequence $\{z_n\}$ is then $z_{n+1} = g(z_n)$ where

$$\begin{aligned} g(z) &= z - \frac{z^2-1}{2z} = \frac{z^2+1}{2z} \\ &= \frac{z}{2} + \frac{1}{2z} = \frac{z}{2} + \frac{\overline{z}}{2|z|^2} \ . \end{aligned}$$

Observe that

(1) Re $z > 0 \Rightarrow$ Re $g(z) = \frac{1}{2}(1 + 1/|z|^2)$Re $z > 0$;

(2) $|g(z)| \leq \frac{1}{2}(|z| + \frac{1}{|z|})$;

(3) arg $z = \theta \Rightarrow \tan\ \arg g(z) = (|z|^2 - 1)(|z|^2 + 1)^{-1} \tan\theta$;

(4) $|\arg\ z| = \theta < \frac{1}{2}\pi \Rightarrow$ Re $g(z) \geq \frac{r}{2}\cos\theta$.

If we start with any first approximant z_0 in the right half plane, then, by (2), the terms of the sequence of Newton approximants $\{z_n\}$ eventually satisfy $|z_n| \leq \frac{3}{2}$ when n exceeds some positive integer N. For such n,

$$|\tan \arg z_{n+1}| \leq \frac{5}{13} |\tan \arg z_n|$$

so that $\lim \arg z_n = 0$ and $\lim \mathrm{Re} z_n = 1$. Then, for some positive integer M, by (3), for $n \geq M$,

$$|z_n - 1| \leq \frac{1}{2} \ .$$

Thence, $|z_n| \geq \frac{1}{2}$ and

$$|z_{n+1} - 1| = |g(z_n) - 1| = \frac{|z_n - 1|^2}{2|z_n|} \leq |z_n - 1|^2 \ .$$

It follows that $\lim_{n\to\infty} z_n = 1$.

If we start Newton's approximating sequence anywhere in the right half plane, the limit of the sequence is the zero of $z^2 - 1$ in that half plane. Similarly, any sequence in the left half plane tends to the zero in that half plane. These half planes are known as the *basins of attraction* for their respective zeros. Any sequence that starts on the imaginary axis remains there, and its behaviour is quite variable depending on the initial term. In general, for any quadratic equation with distinct roots, the basin of attraction of each root with respect to Newton's sequence is the open half plane that contains that root, where the two half planes are separated by the right bisector of the segment joining the roots.

In the case of coincident roots, the situation is particularly simple. For the quadratic $(z - a)^2$, the function $g(z)$ is given by

$$g(z) = z - \frac{1}{2}(z - a) = a + \frac{z - a}{2}$$

and the basin of attraction is the entire complex plane.

To give a sense of how the Newtonian sequence behaves on the right bisector, consider the quadratic polynomial $z^2 + 1$ whose zeros are $\pm i$ and the bisector is the real axis. Then for $0 \neq x_0 \in \mathbf{R}$, we obtain the real sequence of Newton approximants defined by

$$x_{n+1} = x_n - \frac{x_n^2 + 1}{2x_n} = \frac{x_n^2 - 1}{2x_n}$$

for $n \geq 0$. If $x_0 = \cot\theta$, then it can be established by induction that $x_n = \cot 2^n\theta$. As there is no real root, there are no constant sequences. But there are periodic sequences of all orders exceeding 1, as well as other sequences exhibiting chaotic behaviour. For example, taking $\theta = \pi/3$, we obtain the period-2 sequence $\{3^{-1/2}, -3^{-1/2}, 3^{-1/2}, -3^{-1/2}, \cdots\}$.

For polynomials of degree exceeding 2, the behaviour of the Newton approximating sequence is unpredictable. For example, if one approximates the real zeros of $6x^3 - 4x + 1$ with the starting values 0.4500, 0.4600, 0.4700, 0.4800, the sequences of approximants have different limits. This can be understood from analyzing the situation graphically [**1**, pp. 169-170].

§3. A FIRST LOOK AT APOLARITY

For a polynomial of positive degree n, the equation

$$p(z) = p(u) + \frac{p'(u)}{1!}(z-u) + \frac{p''(u)}{2!}(z-u)^2 + \cdots + \frac{p^{(n)}(u)}{n!}(z-u)^n$$

is true for all values of z and u, and so it remains true when we interchange these symbols to obtain

$$p(u) = p(z) + \frac{p'(z)}{1!}(u-z) + \frac{p''(z)}{2!}(u-z)^2 + \cdots + \frac{p^{(n)}(z)}{n!}(u-z)^n \ .$$

Let $q(z) = (u-z)^n$, so that

$$q^{(j)}(z) = (-1)^j n(n-1)\cdots(n-j+1)(u-z)^{n-j} = (-1)^j \frac{n!}{(n-j)!}(u-z)^{n-j} \ ,$$

and we have the identity

$$p(u) = \frac{(-1)^n}{n!}\langle p,q\rangle \ ,$$

where

$$\langle p,q\rangle = p(z)q^{(n)}(z) - p'(z)q^{(n-1)}(z) + p''(z)q^{(n-2)}(z) + \cdots + (-1)^n p^{(n)}(z)q(z) \ .$$

While $\langle p,q\rangle$ is apparently dependent on z, we see that for $q(z) = (u-z)^n$, it takes the constant value $(-1)^n n! p(u)$.

In fact, as can be easily seen by taking the derivative, for any pair $(p(z), q(z))$ of polynomials of degrees not exceeding n, $\langle p,q\rangle$ is a constant. Can we identify the constant? Consider the case that both polynomials are exactly of degree n and have leading coefficient equal to 1.

Recall that, by the Fundamental Theorem of Algebra, each polynomials has exactly n zeros counting multiplicity. Let $u_1, u_2, \cdots, u_n$ be the zeros of $q(x)$ and $v_1, v_2, \cdots, v_n$ be the zeros of $p(x)$, and suppose that s_k is the $k-$th elementary symmetric function of the zeros of $q(x)$ and t_k the $k-$th elementary symmetric function of the zeros of $p(x)$. Thus, each of s_k and t_k is the sum of all $\binom{n}{k}$ products of k of the n zeros of the corresponding polynomial.

Then

$$\begin{aligned} p(z) &= (z-v_1)(z-v_2)\cdots(z-v_n) \\ p^{(k)}(0) &= (-1)^{n-k}k!t_{n-k} \\ q(z) &= (z-u_1)(z-u_2)\cdots(z-u_n) \\ q^{(k)}(0) &= (-1)^{n-k}k!s_{n-k} \end{aligned}$$

for $k = 0, 1, 2, \cdots, n$.

Since $\langle p,q\rangle$ is constant, we can evaluate it by taking $z = 0$. This yields that

$$\begin{aligned} \langle p,q\rangle &= (-1)^n[n!t_n - (-1)^{n-1}(n-1)!s_1(-1)1!t_{n-1} + (-1)^{n-2}(n-2)!s_2(-1)^2 2!t_{n-2} \\ &\quad + \cdots + (-1)^n n!s_n] \\ &= n!s_n - (n-1)!1!s_{n-1}t_1 + (n-2)!2!s_{n-2}t_2 + \cdots + (-1)^n n!t_n \ . \end{aligned}$$

By a straightforward but tedious computation, it can be seen that this is equal to

$$\sum_{\sigma}(u_1 - v_{\sigma(1)})(u_2 - v_{\sigma(2)})\cdots(u_n - v_{\sigma(n)})$$

where each term of the sum is an $n-$fold product and the sum is taken over all $n!$ permutations of $\{1, 2, \cdots, n\}$.

When p and q are linear, then $\langle p, q\rangle = 0$ occurs if and only both have the same zero. For a quadratic polynomial p, $\langle p, p\rangle$ is the negative of the discriminant, and so vanishes if and only if it has coincident roots. When p and q are distinct quadratics, then the condition $\langle p, q\rangle = 0$ implies that their roots satisfy the equation $(u_1 - v_1)(u_2 - v_2) + (u_1 - v_2)(u_2 - v_1) = 0$. When $u_1 = u_2 = u$, then this condition indicates that the zero u of $q(z)$ is one of the zeros of $p(z)$. Suppose that $u_1 \neq u_2$. To understand the significance of the condition $\langle p, q\rangle = 0$, we perform the transformation

$$\phi(z) = \frac{2z - (u_1 + u_2)}{u_1 - u_2}$$

which carries the zeros of $q(z)$ to 1 and -1. Thus,

$$\phi(u_1) = 1, \quad \phi(u_2) = -1,$$

$$w_1 \equiv \phi(v_1) = \frac{2v_1 - (u_1 + u_2)}{u_1 - u_2},$$

$$w_2 \equiv \phi(v_2) = \frac{2v_2 - (u_1 + u_2)}{u_1 - u_2} \ .$$

Then

$$\begin{aligned}\langle p, q\rangle &= \frac{1}{2}(u_1 - u_2)^2[(1 - w_1)(-1 - w_2) + (1 - w_2)(-1 - w_1)] \\ &= -1(u_1 - u_2)^2(1 - w_1 w_2) \ .\end{aligned}$$

When $\langle p, q\rangle = 0$, then $w_1 w_2 = 1$, so that, either w_1 and w_2 are real and interlace 1 and -1, or both lie on the unit circle on opposite sides of the real axis, or they lie on opposite sides of the real axis and one is inside and the other outside of the unit circle. Transforming back to the original pairs of roots, we see that one pair of roots correspond with respect to the composite of a reflection in the line joining the other pair and an inversion in the circle whose diameter is the segment joining them.

More generally, when $q(z) = (z - u_1)(z - u_2)$ and $p(z)$ has arbitrary degree, then the condition $\langle p, q\rangle = 0$ is equivalent to

$$\begin{aligned}0 &= n(n-1)u_1 u_2 - (n-1)(v_1 + v_2 + \cdots + v_n)(u_1 + u_2) + 2(v_1 v_2 + \cdots + v_{n-1} v_n) \\ &= \sum_{1 \leq i < j \leq n} (u_1 - v_i)(u_2 - v_j) + (u_1 - v_j)(u_2 - v_i) \ .\end{aligned}$$

§4. PROBLEMS AND INVESTIGATIONS

1. Determine all polynomials f of degree 2 for which $f(f(1)) = f(f(2)) = f(f(3))$.

2. Let n be a positive integer and let

$$P_n(x) = \sum_{j=0}^{n} \binom{n}{j}^2 x^{2j}(1 - x)^{2(n-j)} \ .$$

(a) Show that $P_n(1/2) \leq P_n(x)$ for $0 \leq x \leq 1$.

(b) Show that, if $P_n(x)$ is written as the sum of powers of $(x - 1/2)$,

$$P_n(x) = \sum_{k=0}^{2n} a_{nk} \left(x - \frac{1}{2} \right)^k ,$$

then $a_{nk} = 0$ when k is odd and $a_{nk} \geq 0$ when k is even.

Hints and Comments

1. Can $f(1)$, $f(2)$, $f(3)$ be all the same or all different? What is the form of a quadratic that takes the same value at u and v when $u \neq v$? Where does the quadratic assume its extreme value?

2. [AMM #11155: 112:5 (May, 2005), 467]

References

1. Edward J. Barbeau, *Polynomials*. Springer, 1989, 2003.