Polynomials

Com S 229

Sep 28, 2009

Polynomials are perhaps the best understood and most applied functions. The foundation comes from algebra and calculus. Taylor's expansion says that a function can be locally expanded around a point into a polynomial whose coefficients depend on the derivative and higher order derivatives of the function at the point. Optimization of polynomial objective functions subject to linear and nonlinear constraints lies in the core of operations research which has impact on resource allocation, transportation, scheduling, economics, etc. Polynomials are used by scientists and engineers to interpolate their experimental data and model the behaviors of physical processes. Systems of multi-variate polynomial equations have received much attention in robotics (motion planning in particular), machine vision, etc. In computer graphics and geometric modeling, parametric curves and surfaces are based on polynomials to model objects in two and three dimensions. In computer vision, polynomials are often fit to image data to describe shape contours.

1 Polynomial Arithmetic

The most common form of a polynomial $p(x)$ is the *power form*:

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \qquad a_n \neq 0. \tag{1}$$

Here n is the *degree* of the polynomial, and $a_0, a_1, \ldots, a_n$ are its *coefficients*. The fundamental theorem of algebra says that $p(x)$ can be factorized over the complex domain into a product $a_n(x - r_1)(x - r_2) \cdots (x - r_n)$, where a_n is the leading coefficient and $r_1, r_2, \ldots, r_n$ are all of its n complex roots.

There are a number of operations that can be carried out on polynomials. Let us first start with an example to illustrate the four arithmetic operations.

Example 1. Consider the following cubic and quadratic polynomials:

$$\begin{aligned} p(x) &= -2x^3 + 4x^2 - 5x + 7, \\ q(x) &= 3x^2 - 5x - 6. \end{aligned}$$

Their sum and difference are determined as follows:

$$\begin{aligned} p(x) + q(x) &= (-2 + 0)x^3 + (4 + 3)x^2 + (-5 - 5)x + (7 - 6) \\ &= -2x^3 + 7x^2 - 10x + 1; \\ p(x) - q(x) &= (-2 - 0)x^3 + (4 - 3)x^2 + (-5 - (-5))x + (7 - (-6)) \\ &= -2x^3 + x^2 + 13. \end{aligned}$$

The product $p(x) \cdot q(x)$ has degree 5, which is the sum of the degree 3 of $p(x)$ and the degree 2 of $q(x)$. It can be represented as

$$p(x) \cdot q(x) = a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0.$$

The leading term a_5x^5 must be the product of the two respective leading terms of $p(x)$ and $q(x)$, namely, $-2x^3$ and $3x^2$. So it is $-6x^5$. What is the coefficient a_4 then? Two term products now contribute to a_4x^4: that of $-2x^3$ with $-5x$ and that of $4x^2$ with $3x^2$. The degrees of the terms in each pair, one from $p(x)$ and the other from $q(x)$, must add up to 4. Therefore,

$$a_4 = -2 \cdot (-5) + 4 \cdot 3 = 22.$$

Similarly, we obtain the remaining coefficients:

$$\begin{aligned} a_3 &= (-2) \cdot (-6) + 4 \cdot (-5) + (-5) \cdot 3 \\ &= -23, \\ a_2 &= 4 \cdot (-6) + (-5) \cdot (-5) + 7 \cdot 3 \\ &= 22, \\ a_1 &= (-5) \cdot (-6) + 7 \cdot (-5) \\ &= -5, \\ a_0 &= 7 \cdot (-6) \\ &= -42. \end{aligned}$$

The product polynomial is now completely determined:

$$p(x) \cdot q(x) = -6x^5 + 22x^4 - 23x^3 + 22x^2 - 5x - 42.$$

Finally, we divide $p(x)$ by $q(x)$ and determine the quotient $s(x)$ and remainder $r(x)$ such that

$$p(x) = s(x) \cdot q(x) + r(x).$$

It is easy to see that $s(x) = b_1x + b_0$ for some b_0, b_1. Dividing the leading coefficient of $p(x)$ by that of $q(x)$ gives us

$$b_1 = \frac{-2}{3} = -\frac{2}{3}.$$

Next, subtraction of $b_1x \cdot q(x)$ from $p(x)$ cancels the latter's leading term:

$$\begin{aligned} p(x) - b_1x \cdot q(x) &= \left(4 - \left(-\frac{2}{3}\right) \cdot (-5)\right) x^2 + \left(-5 - \left(-\frac{2}{3}\right) \cdot (-6)\right) x + 7 \\ &= \frac{2}{3}x^2 - 9x + 7 \\ &\equiv p_1(x). \end{aligned}$$

Next, we divide the remainder $p_1(x)$ by $q(x)$, obtaining $b_0 = \frac{2}{3}/3 = \frac{2}{9}$ and

$$\begin{aligned} p_2(x) &= p_1(x) - b_0 \cdot q(x) \\ &= \left(-9 - \frac{2}{9} \cdot (-5)\right) x + 7 - \frac{2}{9} \cdot (-6) \\ &= -\frac{71}{9}x + \frac{25}{3}. \end{aligned}$$

To summarize, we have obtained the quotient and remainder, respectively,

$$\begin{aligned} s(x) &= -\frac{2}{3}x + \frac{2}{9}, \\ r(x) &= -\frac{71}{9}x + \frac{25}{3}. \end{aligned}$$

Let $p(x)$ be the polynomial given in (1), and $q(x)$ a polynomial of degree m:

$$q(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0, \qquad b_m \neq 0.$$

Addition $p(x) + q(x)$ and subtraction $p(x) - q(x)$ are carried out pairwise on the coefficients associated with terms of the same degrees. Their product is defined by

$$p(x) \cdot q(x) = c_{n+m}x^{n+m} + c_{n+m-1}x^{n+m-1} + \cdots + c_1 x + c_0,$$

where

$$\begin{aligned}
c_{n+m} &= a_n b_m, \\
c_{n+m-1} &= a_{n-1}b_m + a_n b_{m-1}, \\
&\vdots \\
c_n &= a_{n-m}b_m + a_{n-m-1}b_{m-1} + \cdots + a_n b_0, \qquad (\text{ when } n \geq m) \\
&\vdots \\
c_m &= a_0 b_m + a_1 b_{m-1} + \cdots + a_m b_0, \qquad (\text{ when } n \geq m) \\
&\vdots \\
c_1 &= a_0 b_1 + a_1 b_0, \\
c_0 &= a_0 b_0.
\end{aligned}$$

In short, these coefficients are given as

$$c_i = \sum_{\max\{0, i-m\} \leq k \leq \min\{i, n\}} a_k b_{i-k}. \tag{2}$$

Computing all coefficients of the product polynomial as above takes quadratic numbers of arithmetic multiplications and additions. A much more efficient method, called the *fast Fourier transform* (FFT), is beyond the scope of this course and described in Com S 477/577.

Suppose the degree of $p(x)$ is higher than or equal to that of $q(x)$, that is, $n >= m$. We would like to divide $p(x)$ by $q(x)$, more specifically, to determine two polynomials $s(x)$ and $r(x)$ such that

$$p(x) = s(x) \cdot q(x) + r(x),$$

where the quotient $s(x)$ has degree $n - m$ and the remainder $r(x)$ has degree less than m. If $r(x) = 0$, the original polynomial $p(x)$ is said to be *divisible* by $q(x)$. We have the representation:

$$s(x) = d_{n-m}x^{n-m} + d_{n-m-1}x^{n-m-1} + \cdots + d_0.$$

Comparing $p(x)$ and $q(x)$, we immediately see that $d_{n-m} = \frac{a_n}{b_m}$. This is because subtracting $d_{n-m}x^{n-m} \cdot q(x)$ from $p(x)$ yields

$$\begin{aligned}
r_1(x) &= \left(a_n - b_m \cdot \frac{a_n}{b_m}\right) x^n + \left(a_{n-1} - b_{m-1} \cdot \frac{a_n}{b_m}\right) x^{n-1} + \cdots + \left(a_1 - b_1 \cdot \frac{a_n}{b_m}\right) x + \left(a_0 - b_0 \cdot \frac{a_n}{b_m}\right) \\
&= \left(a_{n-1} - b_{m-1} \cdot \frac{a_n}{b_m}\right) x^{n-1} + \cdots + \left(a_1 - b_1 \cdot \frac{a_n}{b_m}\right) x + \left(a_0 - b_0 \cdot \frac{a_n}{b_m}\right),
\end{aligned}$$

a polynomial of degree at least one less than that of $p(x)$. Rename the coefficients of the new polynomial such that

$$r_1(x) = a_{n-1}^{(1)} x^{n-1} + \cdots + a_1^{(1)} x + a_0^{(1)}.$$

If $a_{n-1}^{(1)} = \cdots = a_m^{(1)} = 0$, the division terminates with

$$s(x) = \frac{a_n}{b_m} \qquad \text{and} \qquad r(x) = a_{m-1}^{(1)} x^{m-1} + \cdots + a_1^{(1)} x + a_0^{(1)}.$$

Otherwise, we continue by dividing $r_1(x) = p(x) - d_{n-m} x^{n-m} \cdot q(x)$ by $q(x)$. Let $a_{l_1}^{(1)} \neq 0$, $l \leq n-1$, be its leading coefficient (i.e., $a_{n-1}^{(1)} = \cdots a_{l_1+1}^{(1)} = 0$). Then we obtain the second leading coefficient of the quotient $s(x)$:

$$d_{l_1-m} = \frac{a_{l_1}^{(1)}}{b_m},$$

and a new polynomial $r_2(x) = r_1(x) - d_{l_1-m} x^{l_1-m} \cdot q(x)$.

Since after each division the degree of $r_i(x)$ reduces by at least one, the procedure will stop at the kth step when $r_k(x)$ has degree less than m. In summary, we have carried out the following sequence of divisions:

$$\begin{aligned}
p(x) &= d_{n-m} x^{n-m} \cdot q(x) + r_1(x), \\
r_1(x) &= d_{l_1-m} x^{l_1-m} \cdot q(x) + r_2(x), \\
r_2(x) &= d_{l_2-m} x^{l_2-m} \cdot q(x) + r_3(x), \\
&\vdots \\
r_{k-1}(x) &= d_{l_{k-1}-m} x^{l_{k-1}-m} \cdot q(x) + r_k(x),
\end{aligned}$$

where $n > l_1 > \cdots > l_{k-1} \geq m$. The remainder of the original division is thus

$$r(x) = r_k(x).$$

2 Evaluation of Polynomials

Here we look at how to evaluate the polynomial $p(x)$ efficiently at a point t. A straightforward evaluation is coded in a C++ function `polyEval` below.

```
// polynomial of degree n and with coefficients a[0], ..., a[n]
double polyEval(int n, double a[], double t)
{
    double s = a[0];
    double power = t;

    for (int i=1; i<=n; i++)
    {
        s += a[i] * power;
        power *= t;
    }
    return s;
}
```

A total of $2n$ multiplications and n additions are needed.

Now, we make use of the nested form

$$p(x) = a_0 + x\Big(a_1 + x\Big(a_2 + \cdots + x(a_{n-1} + xa_n)\cdots\Big)\Big).$$

Below is an iterative procedure referred to as *Horner scheme* or *nested multiplication*:

$$\begin{array}{rcl} b_n & \longleftarrow & a_n \\ b_{n-1} & \longleftarrow & a_{n-1} + tb_n \\ & \vdots & \\ b_i & \longleftarrow & a_i + tb_{i+1} \\ & \vdots & \\ b_1 & \longleftarrow & a_1 + tb_2 \\ p(t) = b_0 & \longleftarrow & a_0 + tb_1 \end{array}$$

The above evaluation involves only n multiplications and n additions. Suppose an arithmetic operation is always done in constant time, that is, $\Theta(1)$. Then the evaluation takes time $\Theta(n)$ since the total number of arithmetic operations involved is on the order of n.

The intermediate quantities $b_0, \ldots, b_n$ computed above can serve another purpose. Note from the above iterative procedure that $a_n = b_n$ and

$$a_i = b_i - b_{i+1}t, \qquad \text{for } i = 0, \ldots n-1$$

Substituting these equations into $p(x)$ yields

$$\begin{aligned} p(x) &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \\ &= b_n x^n + (b_{n-1} - b_n t)x^{n-1} + \cdots + (b_1 - b_2 t)x + b_0 - b_1 t \\ &= b_0 + (x-t)b_n x^{n-1} + (x-t)b_{n-1}x^{n-2} + \cdots + (x-t)b_1 \\ &= b_0 + (x-t)q(x), \end{aligned} \tag{3}$$

where $q(x) = b_n x^{n-1} + \cdots + b_2 x + b_1$. In evaluating $p(t)$ as a number, we need to determine the coefficients $b_1, \ldots, b_n$ of a polynomial of degree $n-1$.

By differentiating (3) we get

$$p'(x) = q(x) + (x-t)q'(x).$$

In particular

$$p'(t) = q(t).$$

Because $p(x)$ is a polynomial, we have a very simple method for computing its derivative. Indeed, when evaluating $p(t)$ by Horner scheme, we can simultaneously evaluate $p'(t)$.

In case t is a root, or *zero*, of p, that is, $p(t) = b_0 = 0$. It follows immediately from (3) that

$$p(x) = (x-t)q(x), \qquad \text{or equivalently,} \qquad q(x) = \frac{p(x)}{x-t}.$$

Here $q(x)$ is the called the *deflated polynomial* after factorizing $x - t$ out of $p(x)$.

3 Roots of Low Order Polynomials

We now look at how to find roots, or *zeros*, of a polynomial. Closed-form roots exist only for polynomials of degrees up to four. For polynomials of higher degrees, we generally have to resort to numerical methods to find their roots.

3.1 Quadratics

For a quadratic equation

$$ax^2 + bx + c = 0, \qquad a \neq 0,$$

we know that

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Suppose the coefficients a, b, c are real. It follows that

if $b^2 - 4ac > 0$	the roots are real and unequal;
if $b^2 - 4ac = 0$	the roots are real and equal;
if $b^2 - 4ac < 0$	the roots are imaginary.

3.2 Cubics

The cubic equation

$$x^3 + px^2 + qx + r = 0$$

may be reduced by the substitution

$$x = y - \frac{p}{3}$$

to the normal form

$$y^3 + ay + b = 0, \tag{4}$$

where

$$a = \frac{1}{3}(3q - p^2), \tag{5}$$

$$b = \frac{1}{27}(2p^3 - 9pq + 27r). \tag{6}$$

Equation (4) has the solutions

$$y_1 = A + B, \tag{7}$$

$$y_2 = -\frac{1}{2}(A + B) + \frac{i\sqrt{3}}{2}(A - B), \tag{8}$$

$$y_3 = -\frac{1}{2}(A + B) - \frac{i\sqrt{3}}{2}(A - B), \tag{9}$$

where $i^2 = -1$ and

$$A = \sqrt[3]{-\frac{b}{2} + \sqrt{\frac{b^2}{4} + \frac{a^3}{27}}}, \tag{10}$$

$$B = \sqrt[3]{-\frac{b}{2} - \sqrt{\frac{b^2}{4} + \frac{a^3}{27}}}. \tag{11}$$

To verify the above roots, note that

$$\begin{aligned}(y-y_1)(y-y_2)(y-y_3) &= (y-A-B)\Big(y^2+(A+B)y+A^2-AB+B^2\Big)\\ &= y^3-3ABy-(A+B)(A^2-AB+B^2)\\ &= y^3-3ABy-A^3-B^3\\ &= y^3+ay+b.\end{aligned}$$

Suppose p,q,r are real (and hence a and b are real). Three cases exist:

- $\frac{b^2}{4}+\frac{a^3}{27}>0$. There are one real root $y=y_1$ and two conjugate imaginary roots.
- $\frac{b^2}{4}+\frac{a^3}{27}=0$. There are three real y roots of which at least two are equal:

$$\begin{array}{cccl} -2\sqrt{-\frac{a}{3}}, & \sqrt{-\frac{a}{3}}, & \sqrt{-\frac{a}{3}} & \text{if } b>0,\\ 2\sqrt{-\frac{a}{3}}, & -\sqrt{-\frac{a}{3}}, & -\sqrt{-\frac{a}{3}} & \text{if } b<0,\\ 0, & 0, & 0 & \text{if } b=0.\end{array}$$

- $\frac{b^2}{4}+\frac{a^3}{27}<0$. There are three real and unequal roots:

$$y_k=2\sqrt{-\frac{a}{3}}\cos\left(\frac{\phi}{3}+\frac{2k\pi}{3}\right), \qquad k=0,1,2,$$

  where

$$\cos\phi=\begin{cases}-\sqrt{\frac{b^2/4}{-a^3/27}} & \text{if } b>0;\\ \sqrt{\frac{b^2/4}{-a^3/27}} & \text{if } b<0.\end{cases}$$

EXAMPLE 2. Let us find the roots of the cubic equation

$$x^3+3x^2+2x-1=0.$$

We have the coefficients

$$p=3, \qquad q=2, \qquad \text{and} \quad r=-1.$$

Apply the substitution $x=y-\frac{p}{3}=y-1$ to obtain the normal form

$$y^3+ay+b=y^3-y-1=0.$$

The coefficients are obtained by substituting the values of p,q,r into equations (5) and (6):

$$\begin{aligned} a &= \frac{1}{3}(3\cdot 2-3^2)) = -1;\\ b &= \frac{1}{27}(2\cdot 3^3-9\cdot 3\cdot 2-27) = -1.\end{aligned}$$

Next, we calculate A and B according to (10) and (11):

$$\begin{aligned} A &= \sqrt[3]{\frac{1}{2}+\sqrt{\frac{1}{4}-\frac{1}{27}}}\\ &\approx 0.986991,\\ B &= \sqrt[3]{\frac{1}{2}-\sqrt{\frac{1}{4}-\frac{1}{27}}}\\ &\approx 0.337727.\end{aligned}$$

Plug the values of A and B into (7)–(9):

$$\begin{aligned} y_1 &= 1.32472, \\ y_2 &= -0.662539 + 0.56228i \\ y_3 &= -0.662539 - 0.56228i. \end{aligned}$$

Subtracting 1 from each above gives the three roots of the original cubic equation.

$$\begin{aligned} x_1 &= 0.32472, \\ x_2 &= -1.662539 + 0.56228i \\ x_3 &= -1.662539 - 0.56228i. \end{aligned}$$

3.3 Quartics

The quartic equation

$$x^4 + px^3 + qx^2 + rx + s = 0$$

may be reduced to the form

$$y^4 + ay^2 + by + c = 0 \tag{12}$$

by the substitution

$$x = y - \frac{p}{4},$$

where

$$\begin{aligned} a &= q - \frac{3p^2}{8}, \\ b &= r + \frac{p^3}{8} - \frac{pq}{2}, \\ c &= s - \frac{3p^4}{256} + \frac{p^2q}{16} - \frac{pr}{4}. \end{aligned}$$

The quartic (12) can be factorized under some condition. The equation that must be solved to make it factorizable is called the *resolvent cubic*:

$$z^3 - qz^2 + (pr - 4s)z + (4qs - r^2 - p^2s) = 0. \tag{13}$$

Let z_1 be a real root of the above cubic. Then the four roots of the original quartic are

$$\begin{aligned} x_1 &= -\frac{p}{4} + \frac{1}{2}(R + D), \\ x_2 &= -\frac{p}{4} + \frac{1}{2}(R - D), \\ x_3 &= -\frac{p}{4} - \frac{1}{2}(R - E), \\ x_4 &= -\frac{p}{4} - \frac{1}{2}(R + E), \end{aligned}$$

where

$$
\begin{aligned}
R &= \sqrt{\frac{1}{4}p^2 - q + z_1}, \\
D &= \begin{cases} \sqrt{\frac{3}{4}p^2 - R^2 - 2q + \frac{1}{4}(4pq - 8r - p^3)R^{-1}} & \text{if } R \neq 0, \\ \sqrt{\frac{3}{4}p^2 - 2q + 2\sqrt{z_1^2 - 4s}} & \text{if } R = 0, \end{cases} \\
E &= \begin{cases} \sqrt{\frac{3}{4}p^2 - R^2 - 2q - \frac{1}{4}(4pq - 8r - p^3)R^{-1}} & \text{if } R \neq 0; \\ \sqrt{\frac{3}{4}p^2 - 2q - 2\sqrt{z_1^2 - 4s}} & \text{if } R = 0. \end{cases}
\end{aligned}
$$

Example 3. Find the roots of the following quartic equation:

$$x^4 + 2x^2 - x - 1 = 0.$$

In this case,

$$p = 0, \qquad q = 2, \qquad r = -1, \qquad \text{and} \quad s = -1.$$

The resolvent cubic

$$z^3 - 2z^2 + 4z - 9 = 0$$

has a real root

$$z_1 = 2.11785.$$

Using this root, we calculate the following terms:

$$
\begin{aligned}
R &= 0.343293, \\
D &= 1.30694, \\
E &= 3.15338i.
\end{aligned}
$$

Hence the quartic equation has two real roots and two complex roots:

$$
\begin{aligned}
x_1 &= 0.82511, \\
x_2 &= -0.481816, \\
x_3 &= -0.171647 + 1.57669i, \\
x_4 &= -0.171647 - 1.57669i.
\end{aligned}
$$

4 Root Counting and Bounding

In the remaining sections, we look at how to find the roots of a polynomial of degree $n > 4$:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \qquad a_n \neq 0.$$

It is known that no root formulae exist in the general case. The fundamental theorem of algebra states that p has n real or complex roots, counting multiplicities. If the coefficients $a_0, a_1, \ldots, a_n$ are all real, then *the complex roots occur in conjugate pairs*, that is, in the form $c \pm di$, where $i^2 = -1$ and c and d are real numbers. If the coefficients are complex, the complex roots need not be related.

Using *Descartes' rules of sign*, we can count the number of real positive zeros that $p(x)$ has. More specifically, let v be the number of variations in the sign of the coefficients $a_n, a_{n-1}, \dots, a_0$ (ignoring coefficients that are zero). Let n_p be the number of real positive zeros. Then

(i) $n_p \le v$,

(ii) $v - n_p$ is an even integer.

Similarly, the number of real negative zeros of $p(x)$ is related to the number of sign changes in the coefficients of $p(-x)$.

Example 4. Consider the polynomial $p(x) = x^4 + 2x^2 - x - 1$. Then $v = 1$, so n_p is either 0 or 1 by rule (i). But by rule (ii) $v - n_p$ must be even. Hence $n_p = 1$.

Now look at $p(-x) = x^4 + 2x^2 + x - 1$. Again, the coefficients have one variation in sign, so $p(-x)$ has one positive zero. In other words, $p(x)$ has one negative zero.

To summarize, simply by looking at the coefficients, we conclude that $p(x)$ has one positive real root, one negative real root, and two complex roots as a conjugate pair.

As a matter of fact, $p(x)$ must have *at least one* root, real or complex, inside the circle of radius ρ_1 about the origin of the complex plane, where

$$\rho_1 = \min\left\{ n\left|\frac{a_0}{a_1}\right|, \sqrt[n]{\left|\frac{a_0}{a_n}\right|} \right\}. \tag{14}$$

In the case $a_1 = 0$, we have $\left|\frac{a_0}{a_1}\right| = \infty$ and $\rho_1 = \sqrt[n]{\left|\frac{a_0}{a_n}\right|}$. Meanwhile, if we let

$$\rho_2 = 1 + \max_{0 \le k \le n-1} \left|\frac{a_k}{a_n}\right|, \tag{15}$$

then *all* zeros of $p(x)$ lie inside the circle of radius ρ_2 about the origin.

Example 5. Let $p(x) = x^5 - 3.7x^4 + 7.4x^3 - 10.8x^2 + 10.8x - 6.8$. Applying bound (14), we know that at least one zero of p is inside the circle $|x| \le \rho_1$, where

$$\begin{aligned} \rho_1 &= \min\left\{ 5 \cdot \frac{6.8}{10.8}, \sqrt[5]{\frac{6.8}{1}} \right\} \\ &= \min\{3.14815, 1.46724\} \\ &= 1.46724. \end{aligned}$$

Applying bound (15), all zeros lie inside the circle $|x| \le \rho_2$, where

$$\rho_2 = 1 + \max\{3.7, 7.4, 10.8, 10.8, 6.8\} = 11.8.$$

How does one use the above root bounds? Use them as heuristics that give us a way of localizing the possible zeros of a polynomial. Such localization will provide guided initial guesses to our numerical root finders.

5 Müller's Method

Now we introduce a root finding technique — Müller's method, which can find any number of zeros, real or complex, often with global convergence. It is numerical in the sense that every root estimate is obtained iteratively. It is also global in the sense that it will almost always converge to a root starting with any initial estimate, no matter by how far the estimate seems to be off. Müller's method is also applicable to functions other than polynomials.

The method makes use of *quadratic interpolation*. Suppose the three prior estimates of a zero of $p(x)$ in the current iteration step are the points x_{k-2}, x_{k-1}, x_k. To compute the next estimate we will construct the polynomial of degree ≤ 2 that has the same value as $p(x)$ at x_{k-2}, x_{k-1}, x_k, then find one of its roots. This polynomial, called the *interpolating polynomial*, has the form

$$q(x) = p(x_k) + p[x_{k-1}, x_k](x - x_k) + p[x_{k-2}, x_{k-1}, x_k](x - x_k)(x - x_{k-1}),$$

where the coefficients are

$$p[x_{k-2}, x_{k-1}] = \frac{p(x_{k-2}) - p(x_{k-1})}{x_{k-2} - x_{k-1}}, \qquad (16)$$

$$p[x_{k-1}, x_k] = \frac{p(x_{k-1}) - p(x_k)}{x_{k-1} - x_k}, \qquad (17)$$

$$p[x_{k-2}, x_{k-1}, x_k] = \frac{p[x_{k-2}, x_{k-1}] - p[x_{k-1}, x_k]}{x_{k-2} - x_k}. \qquad (18)$$

Using the equality

$$(x - x_k)(x - x_{k-1}) = (x - x_k)^2 + (x - x_k)(x_k - x_{k-1}),$$

we can rewrite $q(x)$ as

$$q(x) = p(x_k) + b(x - x_k) + a(x - x_k)^2,$$

where

$$a = p[x_{k-2}, x_{k-1}, x_k], \qquad (19)$$

$$b = p[x_{k-1}, x_k] + p[x_{k-2}, x_{k-1}, x_k](x_k - x_{k-1}). \qquad (20)$$

Now let $q(x) = 0$ and solve for x as the next approximation

$$x_{k+1} = x_k - \frac{2p(x_k)}{b \pm \sqrt{b^2 - 4ap(x_k)}}. \qquad (21)$$

For numerical stability, the '$\pm$' sign in the denominator $b \pm \sqrt{b^2 - 4ap(x_k)}$, is chosen so as to maximize the magnitude of the denominator. Note that complex estimates are introduced automatically due to the square-root operation.

5.1 Deflation

The effort of root finding can be significantly reduced by the use of *deflation*. Once you have found a root t of a polynomial $p(x)$, consider next the deflated polynomial $q(x)$ which satisfies

$$p(x) = (x - t)q(x).$$

The computation of the coefficients of $q(x)$ is described in the end of Section 2. The roots of q are exactly the remaining roots of $p(x)$. Because the degree decreases, the effort of finding the remaining roots decreases. More importantly, with deflation we can avoid the blunder of having our iterative method converge twice to the same root instead of separately to two different roots.

In case the polynomial coefficients are real, complex roots must appear in conjugate pairs. So once you have found a complex root, say, $a + bi$; another root must be $a - bi$. You should compute the deflated polynomial

$$q(x) = p(x)/(x^2 + a^2 + b^2)$$

instead. This will require you to write a function to carry out the division.

Roots of deflated polynomials are really just "good suggestions" of the roots of p. Often we need to use the original polynomial $p(x)$ to polish the roots in the end.

5.2 Newton's Method and Root Polishing

A root found by Müller's method often needs to be polished to achieve high accuracy. For such purpose we will employ Newton's method[1] which is another iterative procedure. We start with a root, say, x_0 found by Müller's method. Suppose we have an estimate x_k of a root at iteration step k. Expand $p(x)$ into Taylor series at x_k:

$$p(x) = p(x_k) + p'(x_k)(x - x_k) + \frac{p''(x_k)}{2!}(x - x_k)^2 + \ldots + \frac{p^{(n)}(x_k)}{n!}(x - x_k)^n.$$

Now let $p(x)$ be zero and neglect all the terms containing powers of order higher than one:

$$p(x_k) + p'(x_k)(x - x_k) = 0.$$

Solving the above linear equation gives us the estimate at step $k + 1$:

$$x_{k+1} = x_k - \frac{p(x_k)}{p'(x_k)}. \tag{22}$$

Recall that we can evaluate $p(t)$ and $p'(t)$ together efficiently using Horner scheme, where $p'(t)$ is derived by differentiating the form

$$p(x) = p(t) + (x - t)q(x). \tag{23}$$

5.3 Termination Conditions

To find a zero of $p(x)$, Newton's method takes an initial guess x_0 of a root and iterates according to (22) until $|p(x_k)| < \epsilon$ for some small enough ϵ.

In general, a numerical routine terminates on one of several conditions: a) $x_{k+1} - x_k$ is "small"; b) $|f(x_k)|$ is "small"; or c) k is "large". One may wish to measure a) and b) as relative errors, say, respectively as

$$\begin{array}{lrcl} \text{a)} & |x_{k+1} - x_k| & \leq & \text{XTOL} \cdot |x_k|, \\ \text{b)} & |f(x_k)| & \leq & \text{FTOL} \cdot F, \end{array}$$

where XTOL and FTOL are some preset "tolerances" and F is an estimate of the magnitude.

[1] also referred to as the Newton-Raphson method.

5.4 Pitfalls

One situation that Newton's method does not work well is when the polynomial $p(x)$ has a *double root* r, that is, when

$$p(x) = (x - r)^2 h(x).$$

In this case $p(x)$ shares a factor with its derivative which is of the form

$$p'(x) = 2(x - r)h(x) + (x - r)^2 h'(x).$$

As x approaches r, both p and p' approach zero. In applying Newton's method, machine imprecision will dominate evaluation of the term $\frac{p(x)}{p'(x)}$ as x tends to r. In this case one tiny number is divided by one very small number.

One can avoid the double-root problem by seeking quadratic factors directly in deflation. This is also useful when looking for a pair of complex conjugate roots of a polynomial with real coefficients. For higher order roots, one can detect their possibilities to some extent and apply special techniques to either find or rule out them.

Polynomials are often sensitive to variations in their coefficients. Consequently, after several deflations, the remaining roots may be very inaccurate. One solution is to polish the roots using a very accurate method, once approximations to these roots have been found from deflated polynomials. Newton's method is generally good for polishing both real and complex roots. Just let x_0 take on the value of the root to be polished and run Newton's method.

6 A Root Finding Algorithm

Below is the pseudocode for finding the roots of a polynomial $p(x)$ using Müller's method:

```
if deg(p) ≤ 4 then
  use the appropriate closed forms in Section 3.
else
  p_0(x) ← p(x)
  for i = 0 to deg(p) − 5 do
    apply the rule in Section 4 to determine a circle guaranteed to contain one root of p_i.
    generate three root estimates within the circle.
    use Müller's method to find one root r_i of p_i(x) (i.e., (i + 1)-st of p(x)).
    if r_i is real
      p_{i+1}(x) ← p_i/(x − r_i)   (apply the Horner scheme in Section 2)
      i ← i + 1
    else
      r_i and its complex conjugate conj(r_i) are two roots.
      p_{i+2}(x) ← p_i/((x − r_i)(x − conj(r_i)))   where (x − r_i)(x − conj(r_i)) has real coefficients.
      i ← i + 2
  find the roots of the cubic or quartic p_i using the closed forms in Section 3.3.
  polish all the roots using Newton's method on p(x).
```

Example 6. Find the roots of $p(x) = x^3 - x - 1$ using Müller's method. For each root the iteration start with initial guesses $x_0 = 1, x_1 = 1.5$, and $x_2 = 2.0$. We terminate the iteration at step i when $|\Delta x_i - \Delta x_{i-1}| \le 5 \cdot 10^{-5}$.

The method estimates the first root to be $r_1 = 1.324718$ with an accuracy about $3.4 \cdot 10^{-7}$.

i	x_i	$p(x_i)$
-2	1.0	-1
-1	1.5	0.875
0	2.0	5
1	1.33333	0.037037
2	1.32447	-0.00105
3	1.324718	$1.44 \cdot 10^{-6}$
4	1.324718	$7.15 \cdot 10^{-13}$

Next, we work with the deflated polynomial $q(x) = p(x)/(x - r_1)$ and find the second root to be $r_2 = -0.66236 + 0.56228i$ with an accuracy of $2.2 \cdot 10^{-5}$.

i	x_i	$q(x_i)$
$\vdots$		
1	$-0.6623697 + 0.5622605i$	$2.1355 \cdot 10^{-5} - 1.2 \cdot 10^{-5}i$
2	$-0.66235898 + 0.5622795i$	10^{-11}

The third root must be a conjugate of the second root. So we have $r_3 = -0.66236 - 0.56228i$.

Let us do a more complete example.

Example 7. Find the roots of the polynomial

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

Earlier by checking the sign changes in the coefficient sequence we knew that $p(x)$ has one positive root, one negative root, and one pair of complex conjugate roots.

Now let us look at the bound heuristics:

$$\begin{aligned} \rho &= \min\left\{ n\frac{|a_0|}{|a_1|}, \sqrt[n]{\frac{|a_0|}{|a_n|}} \right\} \\ &= \min\{4, 1\} \\ &= 1. \end{aligned}$$

Thus there is at least one zero inside the complex circle of radius one about the origin. Furthermore, all zeros of $p(x)$ lie inside the circle of radius

$$\begin{aligned} r &= 1 + \max_{0 \le k \le n-1} \left| \frac{a_k}{a_n} \right| \\ &= 1 + \max\{1, 1, 2, 0\} \\ &= 3. \end{aligned}$$

Our search for the roots need only focus on the complex disk of radius 3. For each root we start the Müller's method with the initial guesses

$$x_{-2} = -\frac{1}{2}, \qquad x_{-1} = 0, \qquad x_0 = \frac{1}{2},$$

and terminate the search once $\Delta x_i \le 5 \cdot 10^{-5}$.

The first root is found to be 0.82511. We divide $p(x)$ by $x - 0.82511$ using the Horner scheme and generate the deflated polynomial:

$$x^3 + 0.82511x^2 + 2.68081x + 1.21196.$$

The Müller's method, now applied on the deflated polynomial, yields the second root -0.481816. Deflate the cubic polynomial above using this new root:

$$x^2 + 0.343294x + 2.51541.$$

We may apply Müller's method again to find the third and fourth roots which form a complex conjugate pair: $-0.171647 \pm 1.57669i$. Or we may simply use the close forms for quadratic polynomials.

The following table shows the iterations to find each root and its accuracy. Note that the two complex roots lie between the circle of radius 1 and the circle of radius 3.

roots	# iterations	accuracy
0.8251098	5	$2.4 \cdot 10^{-7}$
-0.4818156	4	$1.4 \cdot 10^{-6}$
$-0.171647 + 1.576686i$	2	$5.1 \cdot 10^{-6}$
$-0.171647 - 1.576686i$		

We should next polish the four root estimates using the original polynomial $p(x)$.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables.* New York: Dover, pp. 17-18, 1972.

[2] R. S. Burington. *Handbook of Mathematical Tables and Formulas.* McGraw-Hill, Inc., 5th edition, 1973.

[3] S. D. Conte and Carl de Boor. *Elementary Numerical Analysis.* McGraw-Hill, Inc., 2nd edition, 1972.

[4] M. Erdmann. Lecture notes for *16-811 Mathematical Fundamentals for Robotics.* The Robotics Institute, Carnegie Mellon University, 1998.

[5] C. C. MacDuffee. *Theory of Equations.* John Wiley & Sons, Inc., New York, 1954.

[6] W. H. Press, *et al.* *Numerical Recipes in C++: The Art of Scientific Computing.* Cambridge University Press, 2nd edition, 2002.

[7] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Springer-Verlag New York, Inc., 2nd edition, 1993.

[8] H. W. Turnbull. *Theory of Equations*, 5th edition. Oliver and Boyd Ltd., Edinburgh, 1952.

[9] Wolfram MathWorld. `http://mathworld.wolfram.com/QuarticEquation.html`.