LECTURE 16

Errors in Polynomial Interpolation

As with any approximate method, the utility of polynomial interpolation can not be stretched too far. In this lecture we shall quantify the errors that can occur in polynomial interpolation and develop techniques to minimize such errors.

We shall begin with an easy theorem.

THEOREM 16.1. *Let f be a function in $C^{n+1}[a,b]$, and let P be a polynomial of degree $\leq n$ that interpolates the function f at $n+1$ distinct points $x_0, x_1, \ldots, x_n \in [a,b]$. Then to each $x \in [a,b]$ there exists a point $\xi_x \in [a,b]$ such that*

$$f(x) - p(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i)$$

REMARK 16.2. Although this formula for the error is somewhat reminiscent of the error term associated with an n^{th} order Taylor expansion, this theorem has little to do with Taylor expansions.

Proof. If $x = x_i$, one of the nodes of the interpolation, then the statement is certainly true since both sides vanish identically. Suppose now that $x \neq x_i$, $i = 0, 1, \ldots, n$. Put

$$\begin{aligned} w(t) &= \prod_{i=0}^{n} (t - x_i) \\ \phi_x(t) &= f(t) - P(t) - \frac{f(x) - P(x)}{w(x)} w(t) \end{aligned}$$

Then $\phi_x(t) \in C^{n+1}[a,b]$, and $\phi_x(t)$ vanishes at $n+2$ distinct points: i.e., when $t = x_0, x_1, \ldots, x_n$, or x. Now from Calculus I, we have Rolle's Theorem which states that if a differentiable function $f(x)$ has n distinct zeros, then its derivative must have at least $n-1$ zeros (these being the points where the graph of the function $f(x)$ turns around to re-cross the x-axis). Hence, $\phi_x'(t)$ has at least $n+1$ distinct zeros, $\phi_x''(t)$ has at least n distinct zeros, and so on until we can conclude that $\phi_x^{(n+1)}(t)$ has at least one distinct zero in $[a,b]$; call it ξ_x. Now

$$\begin{aligned} \phi_x^{(n+1)}(t) &= \frac{d^{(n+1}}{dt^{(n+1)}} \left(f(t) - P(t) - \frac{f(x) - P(x)}{w(x)} w(t) \right) \\ &= f^{(n+1)}(t) - P^{(n+1)}(t) - \left(\frac{f(x) - P(x)}{w(x)} \right) \frac{d^{(n+1}}{dt^{(n+1)}} (w(t)) \\ &= f^{(n+1)}(t) - P^{(n+1)}(t) - \left(\frac{f(x) - P(x)}{w(x)} \right) \frac{d^{(n+1}}{dt^{(n+1)}} ((t - x_0)(t - x_1) \cdots (t - x_n)) \\ &= f^{(n+1)}(t) - P^{(n+1)}(t) - \left(\frac{f(x) - P(x)}{w(x)} \right) (n+1)! \end{aligned}$$

Hence,

$$0 = \phi_x^{(n+1)}(\xi_x) = f^{(n+1)}(\xi_x) - P^{(n+1)}(\xi_x) - \left(\frac{f(x) - P(x)}{w(x)} \right) (n+1)!$$

Now, because $P(x)$ is a polynomial of degree n, $P^{(n+1)}(x) = 0$. Hence we have

$$0 = f^{(n+1)}(\xi_x) - \left(\frac{f(x) - P(x)}{w(x)}\right)(n+1)!$$

or

$$f(x) - P(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) w(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i) .$$

EXAMPLE 16.3. If $P(x)$ is the polynomial that interpolates the function $f(x) = \sin(x)$ at 10 points on the interval $[0, 1]$, what is the greatest possible error?

- In this example, we have $n + 1 = 10$ and
$$f^{(n+1)}(x) = f^{(10)}(x) = -\sin(x)$$
so the largest possible error would be the maximal value of
$$\left| \frac{1}{10!} f^{(10)}(\xi_x) \prod_{i=0}^{n} (x - x_i) \right| \quad ,$$
for $x, x_0, x_1, \ldots, x_n, \xi_x \in [0, 1]$. Clearly, on the interval $[0, 1]$
$$\begin{aligned} \max |x - x_i| &= 1 \\ \max \left| f^{(n+1)}(\xi_x) \right| &= \max |-\sin(\xi_x)| = 1 \end{aligned}$$
so the maximal error would be
$$\frac{1}{10!}(1)(1)^{n+1} \approx 2.8 \times 10^{-7}$$

1. Chebyshev Polynomials and the Minimalization of Error

The theorem in the preceding system not only tells us how large the error could be when a given function is replaced by an interpolating polynomial; it also gives us a clue as to how we might arrnge things to make the error as small as possible.

To see this, let me write down again the expression for the error term:

$$E(x) \equiv f(x) - P(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i) \quad , \quad \text{for some } \xi_x \in [a, b]$$

Now we don't even know what ξ_x is except that it's some point in the interval $[a, b]$ that depends on x; so there's not much we can do with the term $f^{(n+1)}(\xi_x)$ (particularly, because, in physical applications, we don't even know what f is). However, we can try to make the term

$$\prod_{i=0}^{n} (x - x_i)$$

as small as possible by picking a suitable choice of nodes $\{x_i\}$.

Now one simple choice for the x_i would be to set

$$\begin{aligned} \Delta x &= \frac{a + b}{n} \\ x_i &= a + i(\Delta x) \end{aligned}$$

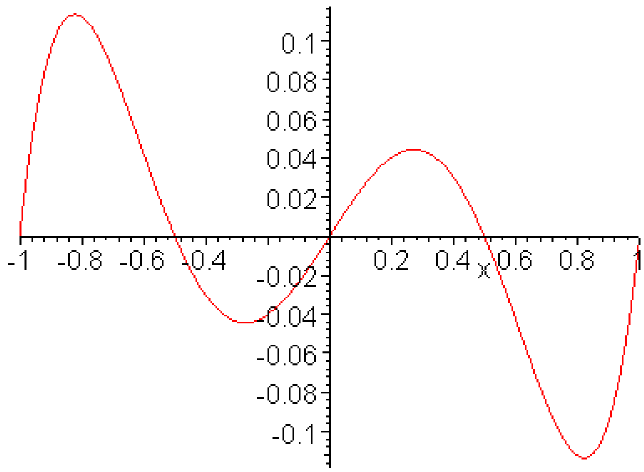
However, this is a case where the simplest choice of the x_i turns out not to be the best choice. We can see this with a simple example. Consider the case where $a = -1$, $b = 1$, and

$$x_i = -1 + (0.5)i \quad , \quad i = 0, 1, 2, 3, 4$$

Then if we set

$$w(x) \equiv \prod_{i=0}^{n}(x - x_i) = (x-1)(x-0.5)(x-0)(x-0.5)(x-1)$$

and plot it



we find that has a maximum value of about 0.11.

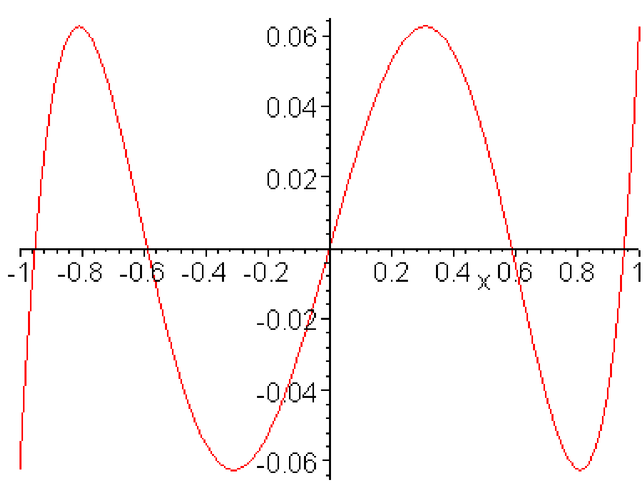
Suppose instead we, for some strange reason, choose the points

$$\begin{aligned} x_0 &= -0.9510565160 \\ x_1 &= -0.5877852520 \\ x_2 &= 0.0 \\ x_3 &= 0.5877852520 \\ x_4 &= 0.9510565160 \end{aligned}$$

and plot

$$w_1(x) = (x+0.9510565160)(x+0.5877852520)(x-0)(x-0.5877852520)(x-0.9510565160)$$

We then find



which has a maximum value of about 0.06, which is about half the value that we obtained for the simpler choice of points x_i.

Thus, by choosing a special set of points x_i it is possible to reduce the contribution of the factor

$$\prod_{i=0}^{n}(x-x_i)$$

to the error term, and thus minimize the overall error of the interpolation polynomial.

So now the question becomes: **how to choose a good set of points to sample data, so that a polynomial interpolation is as accurate as possible?** This is where Chebyshev polynomials will come into play.

1.1. Chebyshev Polynomials. The Chebyshev polynomials are defined recursively, via the formula

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x) \quad , \quad n = 1,2,3,4,\ldots \end{aligned}$$

The first six Chebyshev polynomials are thus

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \\ T_5(x) &= 16x^5 - 20x^3 + 5x \\ T_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1 \\ &\vdots \end{aligned}$$

Note that the leading term of the Chebyshev polynomial $T_n(x)$ is $2^{n-1}x^n$.

THEOREM 16.4. *For $x \in [-1,1]$ we have*

$$T_n(x) = \cos\left(n\cos^{-1}(x)\right)$$

This theorem is proved by showing that $f_n(x) \equiv \cos(n\cos^{-1}(x))$, then

$$\begin{aligned} f_0(x) &= 1 \\ f_1(x) &= x \end{aligned}$$

and then using the trig identity

$$\cos(A+B) = \cos(A)\cos(B) - \sin(A)\sin(B)$$

to demonstrate that

$$f_{n+1}(x) = 2xf_n(x) - f_{n-1}(x)$$

Hence, $f_n(x)$ satisfies the defining properties of the Chebyshev polynomials.

COROLLARY 16.5. *We have*

$$\begin{aligned} |T_n(x)| &\leq 1 \quad , \quad x \in [-1,1] \\ T_n\left(\cos\left(\frac{j\pi}{n}\right)\right) &= (-1)^j \quad , \quad j = 0,\ldots,n \\ T_n\left(\cos\left(\frac{2j-1}{2n}\pi\right)\right) &= 0 \quad , \quad j = 1,\ldots,n \end{aligned}$$

First, however we shall give a negative result.

DEFINITION 16.6. *A polynomial P of degree n is called **monic** if the coefficient x^n is 1.*

Note that expressions of the form

$$w(x) = \prod_{i=0}^{n}(x - x_i)$$

are monic polynomials, as are the polynomials obtained from the Chebyshev polynomials by dividing through by the leading coefficient.

$$Q_n(x) = \frac{1}{2^{n-1}}T_n(x)$$

Our first application of Chebyshev polynomials will be to prove a lower bound for maximum value of a monic polynomial on the interval $[-1, 1]$.

THEOREM 16.7. *If P is a monic polynomial of degree n, then*

$$\|P(x)\|_\infty \equiv \max_{-1\leq x\leq 1} |P(x)| \geq 2^{1-n}$$

Proof. Suppose that $P(x)$ is a monic polynomial of degree n and that

$$|P(x)| < 2^{1-n} \quad , \quad \forall\, x \in [-1, 1]$$

Set

$$\begin{aligned} Q_n(x) &= 2^{1-n}T_n(x) \\ x_i &= \cos\left(\frac{i\pi}{n}\right) \quad , \quad i = 0, 1, \ldots, n \end{aligned}$$

Then by construction $Q_n(x)$ is a monic polynomial of degree n, and we'll have

$$(-1)^i Q_n(x_i) = (-1)^i 2^{1-n} T_n\left(\cos\left(\frac{i\pi}{n}\right)\right) = 2^{1-n}(-1)^i(-1)^i = 2^{1-n}$$

Since $P(x)$ and $Q_n(x)$ both have leading coefficient 1, their difference $Q_n(x) - P(x)$ will be a polynomial of degree $\leq n - 1$. On the other hand,

$$(-1)^i P(x_i) \leq |P(x_i)| < 2^{1-n} = (-1)^i Q_n(x_i) \quad , \quad i = 0, 1, 2, \ldots, n$$

Hence,

$$(-1)^i [Q_n(x_i) - P(x_i)] > 0 \quad , \quad i = 0, 1, 2, \ldots, n$$

Thus, the function $Q_n(x) - P_n(x)$ must oscillate in signs at least $n + 1$ times over the interval $[-1, 1]$. But this is not possible since $Q_n(x) - P(x)$ is a polynomial of degree at most $n - 1$. Hence, we have a contradiction if $|P(x)| < 2^{1-n}$, $\forall\, x \in [-1, 1]$. Thus, the opposite inequality must hold.

LEMMA 16.8. *If $Q_n(x)$ is the monic polynomial defined by*

$$Q_n(x) = 2^{1-n}T_n(x)$$

then the maximal value of $|Q_n(x)|$ on the interval $[-1, 1]$ is 2^{-1-n}.

Proof. This is easy since on the interval $[-1, 1]$

$$Q_n(x) = 2^{1-n}T_n(x) = 2^{1-n}\cos\left(n\cos^{-1}(x)\right)$$

and $|\cos(\theta)| \leq \cos(0) = 1$ for all θ.

LEMMA 16.9. *Let*

$$x_i = \cos\left(\frac{2i+1}{2n+2}\pi\right) \quad , \quad i = 1, \dots, n$$

Then

$$(x - x_1)\cdots(x - x_n) = Q_n(x) \equiv 2^{1-n}T_n(x) \tag{16.1}$$

Proof. This follows from the Fundamental Theorem of Algebra. By construction each of the x_i is a distinct root of the monic polynomial $Q_n(x)$, which is of degree n. The Fundamental Theorem of Algebra tells us that $Q_n(x)$ must therefore factorize as$Q_n(x) = (x - x_1)\cdots(x - x_n)$.

We now have

THEOREM 16.10. *If the nodes x_i are chosen as the roots of the Chebyshev polynomial $T_{n+1}(x)$*

$$x_i = \cos\left(\frac{2i+1}{2n+2}\pi\right) \quad , \quad i = 0, 1, \dots, n$$

then the error term for polynomial interpolation using the nodes x_i is

$$E(x) = |f(x) - P(x)| \leq \frac{1}{2^n(n+1)!} \max_{-1 \leq t \leq 1} \left| f^{(n+1)}(t) \right|$$

Moreover, this is the best upper bound we can achieve by varying the choice of the x_i.

1.2. Picking Optimal Nodes on More General Intervals. The results of the preceding section can be summarized as follows: if we want a polynomial interpolating a function f at $n+1$ points x_i in the interval $[-1, 1]$ to be as accurate as possible, then we should choose the data points x_i so that they are the zeros of the Chebyshev polynomial $T_{n+1}(x)$.

Put more practically: suppose we have an experiment that measures a quantity Q that depends on a parameter $x \in [-1, 1]$. If we are to find the polynomial $P(x)$ of degree n that most accurately represents the actual function $Q(x)$, by interpolating the data taken at $n+1$ points x_i, then we should choose the x_i so that they are the zeros of the Chebyshev polynomial $T_{n+1}(x)$.

What do we do if an experimental parameter x is allowed to range through some other intevarl $[a, b] \neq [-1, 1]$?

The answer is quite easy. To find an optimal set of $n+1$ data points x_i in an interval $[a, b]$ we simply rescale the $n+1$ zeros of $T_{n+1}(x)$ to points in $[a, b]$. More precisely, let s be the linear map that maps a point $x \in [-1, 1]$ to a point $s(x) \in [a, b]$, such that $s(-1) = a$ and $s(1) = b$. These properties actually fix s uniquely,

$$s(x) = a + \frac{(b-a)}{2}(x+1) = \frac{b+a}{2} + \frac{b-a}{2}x$$

The optimal set of $n+1$ data points x_i for interpolating a function $Q(x)$ on the interval $[a, b]$ will then be the image under s of the $n+1$ zeros of $T_{n+1}(x)$:

$$x_i = \frac{(b+a)}{2} + \frac{b-a}{2}\cos\left(\frac{2i+1}{2n+2}\pi\right) \quad , \quad i = 0, \dots, n$$