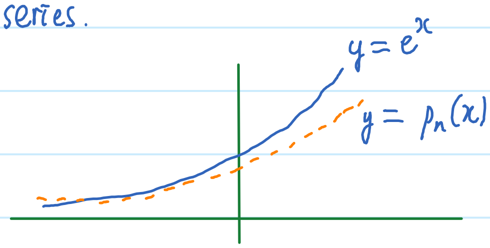
Lecture 21

Wednesday, February 26, 2020

Polynomials are simple functions to work with. For example, to compute the value of a polynomial at a given value of x, one only needs to invoke the basic arithmetic operations: addition, subtraction, multiplication.

Working with polynomials is usually desirable. There are many ways to approximate a non-polynomial function by a polynomial. Taylor approximation is an example. The idea of Taylor approximation is to truncate the Taylor series.

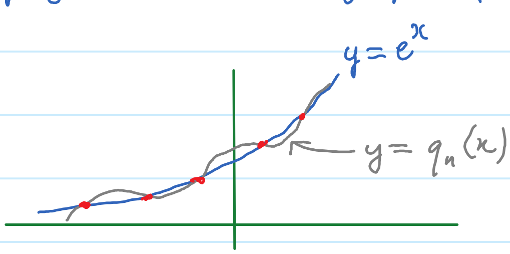


For example, we know that

$$e^x = \underbrace{1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}}_{p_n(x)} + \cdots$$

On a finite interval, say $[0,2]$, the polynomial $p_n(x)$ approximates well e^x when n is large enough. One can estimate the error between e^x and $p_n(x)$ using Lagrange's theorem. Although $p_n(x)$ approximates well e^x, the two curves hardly intersect at any points.

Interpolation gives an alternative way to approximate a function by a polynomial. On the graph of $f(x) = e^x$, we take n points and find an interpolation polynomial of these points. (The curve $y = q_n(x)$ on the picture.)



We are now interested in the error between the true function $f(x)$ and the approximate function $q_n(x)$.

There is a theorem analogous to Lagrange's theorem for Taylor approximation.

* **Theorem:** let $f : [a,b] \to \mathbb{R}$ and $a \le x_1 < x_2 < \dots < x_n \le b$. Put $y_1 = f(x_1), y_2 = f(x_2), \dots, y_n = f(x_n)$. Let P be the interpolation polynomial of n points $(x_1, y_1), \dots, (x_n, y_n)$. Then

$$f(x) - P(x) = \frac{f^{(n)}(r)}{n!}(x-x_1)(x-x_2)\dots(x-x_n). \qquad (*)$$

for some $r \in (a,b)$ depending on x.

We won't discuss the proof of this theorem. It suffices to say that the formula $(*)$ comes from a clever way of using Newton's formula. The error is obtained by taking the absolute value of both sides:

$$|f(x) - P(x)| = \frac{|f^{(n)}(r)|}{n!}|x-x_1||x-x_2|\dots|x-x_n|. \qquad (**)$$

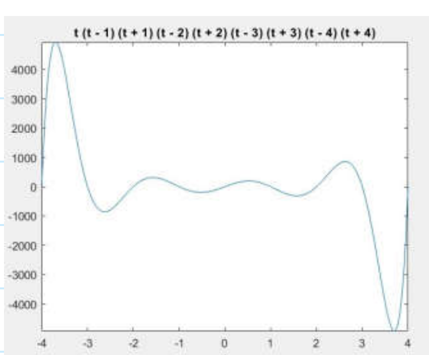
Since the error depends on x, we are only interested in the worst error. That is $\max_{[a,b]} |f(x) - P(x)|$. We want to know how large the worst error can be.

From $(**)$, we see that the size of the error depends on the size of two functions:

(1) The n'th order derivative $f^{(n)}$.

(2) The polynomial $(x-x_1)(x-x_2)\dots(x-x_n)$.

The first function depends on f. The second function doesn't. One can plot this polynomial to see where it attains maximum or minimum.



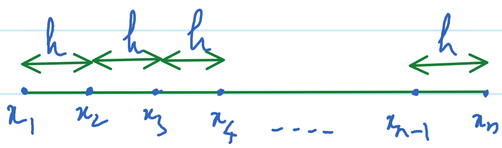
On the picture is the polynomial

$$Q(x) = (x-x_1)\dots(x-x_n)$$

with $x_1 = -4, x_2 = -3, \dots, x_8 = 3, x_9 = 4$.

We see that $Q(x)$ is largest (in size) when x is in between x_1 and x_2, or between x_8 and x_9

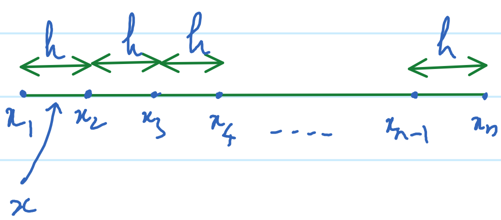
Let us consider the general case when $x_1 < x_2 < \dots < x_n$ are equally spaced.



Why is it that $Q(x)$ is large when x is near the ending points of the interval?

Here is a hand-waving argument. We know that $Q(x)$ is small (i.e. close to zero) when x is close to one of the points $x_1, x_2, \dots, x_n$. To make $Q(x)$ large, x should be relatively far from the cluster $x_1, x_2, \dots, x_n$. If x is somewhere in the middle (about $x_{n/2}$), it will have two neighbors in the neighborhood of h, four neighbors in the neighborhood of $2h$, six neighbors in the neighborhood of $3h$, ... If x is near the endpoints (either x_1 or x_n) then it will have two neighbors in the neighborhood of h, three neighbors in the neighborhood of $2h$, four neighbors in the neighborhood of $3h$, ...

In this manner, x is furthest from the cluster $x_1, x_2, \dots, x_n$ if it is near the endpoints of the whole interval.



To see how large the product
$|x-x_1|\,|x-x_2| \dots |x-x_n|$
could be, we only need to consider x near the endpoints, say between x_1 and x_2. We have

$$\left.\begin{aligned} |x-x_1| &\le h \\ |x-x_2| &\le h \\ |x-x_3| &\le 2h \\ |x-x_4| &\le 3h \\ &\dots \\ |x-x_n| &\le (n-1)h \end{aligned}\right\}$$

Thus, $|x-x_1| \dots |x-x_n| \le h(h)(2h)\dots(n-1)h$
$= h^n (n-1)!$

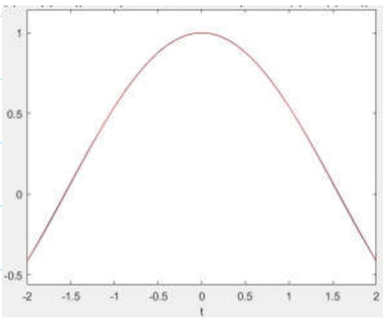
Therefore, the worst error is bounded by

$$\max_{[a,b]} |f(x) - P(x)| \le \frac{1}{n!} h^n (n-1)! \max_{[a,b]} |f^{(n)}|$$

$$= \frac{1}{n} \left(\frac{b-a}{n-1}\right)^n \max |f^{(n)}|$$

Ex: $f(x) = \cos x$, $[a,b] = [-2,2]$.

Divide the interval $[a,b]$ into 20 equal intervals



The interpolation polynomial (the blue curve) is almost indistinguishable from the cosine curve (the red curve).

The difference between two curves is

$$|f(x) - P(x)| = \frac{|f^{(n)}(r)|}{n!} |x - x_1| \dots |x - x_n|$$

$$\le \frac{1}{n} \left(\frac{b-a}{n-1}\right)^n \max_{[a,b]} |f^{(n)}|.$$

Note that $f^{(n)}$ can only be $\pm \sin x$ or $\pm \cos x$ Thus, $|f^{(n)}(x)| \le 1$. We get

$$|f(x) - P(x)| \le \frac{1}{n} \left(\frac{4}{n-1}\right)^n.$$

This number goes to 0 very quickly. The cosine curve is therefore approximated very well by the interpolation polynomial curve.

Ex: $f(x) = \dfrac{1}{4x^2+1}$ on $[a,b] = [-3,3]$.

Divide the interval $[a,b]$ into equal subintervals of length 0.2. We see that the interpolation polynomial approximates f well when x is near the middle of the interval The polynomial oscilates wildly near the endpoints. This is known as Runge phenominon. The reason is that the higher derivatives of f grows too quickly when n increases.

We will discuss the reason in detail next time.