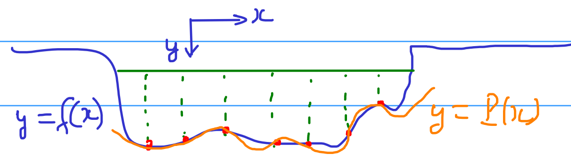
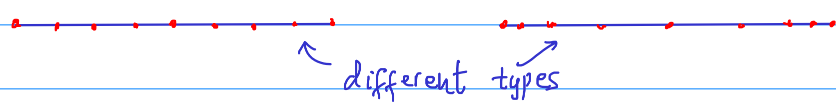
Lecture 20 (11/15/2019)

Error estimate of polynomial interpolation.

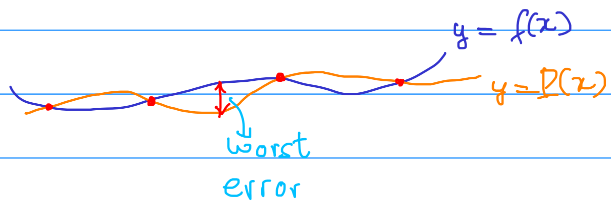


In many situations, we don't know an exact function f, for example the depth of a lake as a function of position. However, we do know the value of f at some sample points $x_1, \dots, x_n$. We can also construct a polynomial P fitting the points $(x_1, y_1), \dots, (x_n, y_n)$. A natural question is: will the error $|f(x) - P(x)|$ go to zero as $n \to \infty$?

This question has to be formulated more carefully. First, the polynomial P does not only depend on n, but also on the sample points $(x_1, y_1), \dots, (x_n, y_n)$.



Secondly, the error $|f(x) - P(x)|$ depends on x (the position). The quantity $\max_{x \in [a,b]} |f(x) - P(x)|$ can be seen as the "worst error".



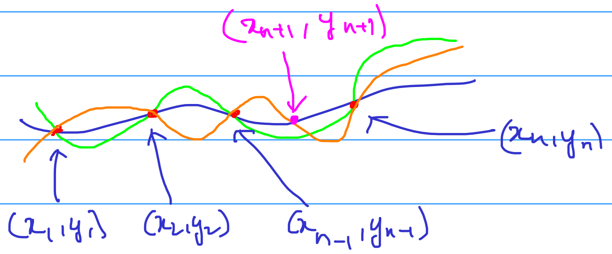
We will try to estimate the worst error.

Let $\alpha \in [a,b]$ and $\alpha \neq x_1, x_2, \dots, x_n$. How big is the error

$$|f(\alpha) - P_n(\alpha)|$$

where P_n is the polynomial fitting $(x_1, y_1), \dots, (x_n, y_n)$?

Put $x_{n+1} = \alpha$ and $y_{n+1} = f(\alpha)$.



Let P_{n+1} be the polynomial fitting $(x_1, y_1), \ldots, (x_{n+1}, y_{n+1})$. This is the orange curve. P_n is the green curve.

By Newton formula,

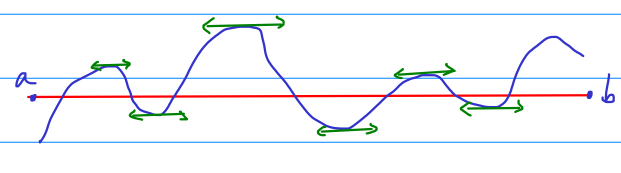
$$P_{n+1}(x) = P_n(x) + \underbrace{c}_{= f[x_1, x_2, \ldots, x_{n+1}]}(x - x_1)\cdots(x - x_n)$$

To see how close $f(\alpha)$ is to $P_{n+1}(\alpha)$, we consider the error function

$$g(x) = f(x) - P_{n+1}(x)$$

We know that $g(x_1) = g(x_2) = \cdots = g(x_{n+1}) = 0$.

Because g vanishes at $n+1$ different points, g' must vanish at n points. Then g'' must vanish at $n-1$ points, etc. Then $g^{(n)}$ must vanish at one point. This is an application of Fermat's theorem.



Therefore, there exists $r = r_\alpha$ (depending on α) such that $g^{(n)}(r) = 0$. This means

$$f^{(n)}(r) - P^{(n)}_{n+1}(r) = 0 \qquad (*)$$

Recall that

$$P_{n+1}(x) = \underbrace{P_n(x)}_{\text{degree} \le n-1} + c\underbrace{(x-x_1)\cdots(x-x_n)}_{\text{degree } n}$$

Then $P_{n+1}^{(n)}(x) = c \cdot n!$ (Constant function)

Thus, $P_{n+1}^{(n)}(r) = c \cdot n!$ Then $(*)$ implies

$$c = \frac{f^{(n)}(r)}{n!}$$

We have

$$f(\alpha) = P_{n+1}(\alpha) = P_n(\alpha) + c(\alpha - x_1)\cdots(\alpha - x_n)$$

$$= P_n(\alpha) + \frac{f^{(n)}(r_\alpha)}{n!}(\alpha - x_1)\cdots(\alpha - x_n)$$

Since α was taken arbitrarily at the beginning, we can replace α by a general name x:

$$f(x) = P_n(x) + \frac{f^{(n)}(r_x)}{n!}(x-x_1)\cdots(x-x_n)$$

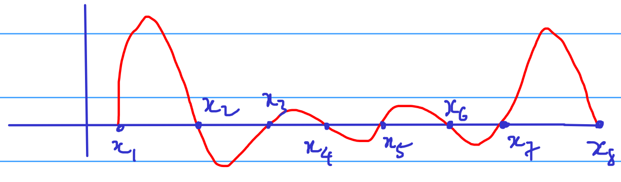
here r_x is some number in (a,b).

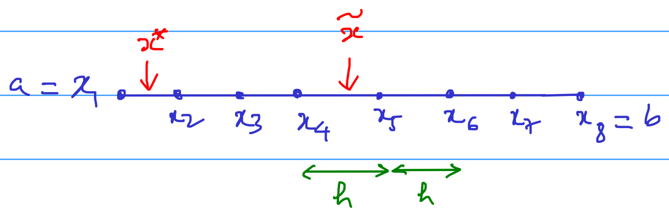
We get

$$\underbrace{|f(x) - P_n(x)|}_{\substack{\text{error at} \\ \text{position } x}} = \frac{|f^{(n)}(r_x)|}{n!}|x-x_1|\cdots|x-x_n|$$

The size of this error depends on the size of the n'th derivative of f and the size of the function $(x-x_1)\cdots(x-x_n)$.

By experimenting on Matlab, we see that $(x-x_1)\cdots(x-x_n)$ is large when x is near $a = \min\{x_1,\dots,x_n\}$ or near $b = \max\{x_1,\dots,x_n\}$.





A rough explanation to this phenomenon is as follows. If $x = \tilde{x}$ is near the middle of the interval then it has two neighbors within a distance h, four neighbors within a distance $2h$, ... If $x = x^*$ is near $a = x_1$ then it has two neighbors within distance h, three within distance $2h$, four within distance $3h$, ... Thus, x^* is "further" from the cluster $\{x_1, x_2, \dots, x_n\}$. Thus,

$$|x^* - x_1|\,|x^* - x_2| \cdots |x^* - x_n| > |\tilde{x} - x_1|\,|\tilde{x} - x_2| \cdots |\tilde{x} - x_n|$$

We will do more experiments on the size of $|f(x) - P_n(x)|$ next time.