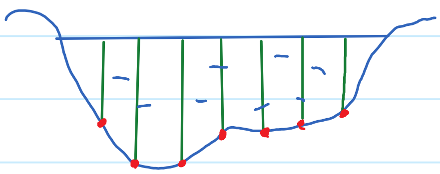
Lecture 17

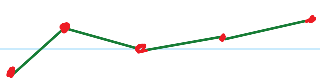
Monday, February 17, 2020

We have discussed an important topic in Numerical Analysis – the rootfinding problem. This can be done by the bisection method, Newton method, or fixed point method. There are many more methods which one can learn later when needed.

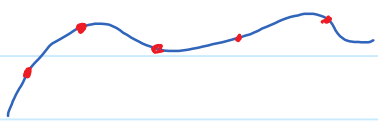
We will now discuss another classical topic in Numerical Analysis – the interpolation problem. An interpolation problem usually involves finding a curve that passes a set of data points. In real life, these data points are the data obtained by sampling or discretization. For example, to find the depth of a lake as a function in position, one can only sample to depth of the lake a finitely many locations (the red points). Then one needs to reconstruct the depth function from those points.



This is the problem of find a curve that fits a data set.

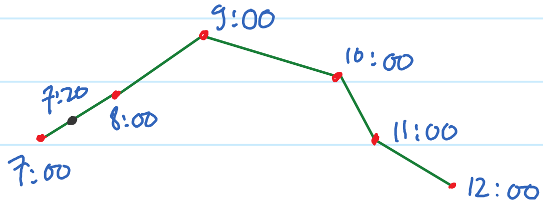


polygon curve



polynomial curve

There are of course many curves that can fit a data set. The simplest way is connect two consecutive points by a straight segment. This gives us a polygon curve. It is also called a linear spline curve. Such linear interpolation method was used anciently by the Babylonian and Greek to study the positions of stars. For example, at certain times of a day, people observe that a star is at certain positions on the sky (the red dots on the next picture).



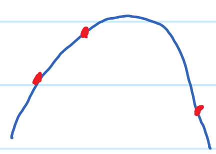
It is natural to ask where the star is at 7:20 PM. By connecting the two dots at 7:00 and 8:00 by a straight line, one can approximate the position (x, y) of the star at 7:20 by

$$\begin{cases} x = \dfrac{2}{3} x_1 + \dfrac{1}{3} x_2 \\[2ex] y = \dfrac{2}{3} y_1 + \dfrac{1}{3} y_2 \end{cases}$$

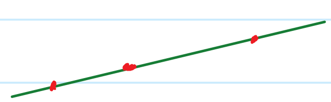
where (x_1, y_1) and (x_2, y_2) are the position of the star at 7:00 and 8:00 respectively.

The advantage of linear interpolation is that it is simple to calculate. A disadvantage is that it doesn't look natural: there are corners at each data point. In reality, a star doesn't abruptedly change the slope of its trajectory at any point. Rather, the slope should change smoothly. Polynomial interpolation can fix this issue. Instead of connecting two consecutive points by straight lines, we find a polynomial curve that goes through the given points.

One of the first person who used polynomial curve for interpolation was Chinese mathematician Liu Zhuo (~660 AD). He found a parabola that passes through 3 points.



Generally, one can't fit 3 points by a line. In other words, a polynomial of degree ≤ 1 is in general unable to fit 3 given points. Exception happens when the points happen to lie on the same line. Generally, one needs a polynomial of degree at least 2 to fit a set of 3 data points.

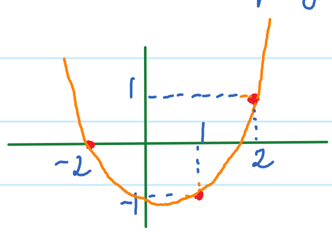


One can naturally ask: what is the polynomial with smallest degree that can fit a set of n data points, namely $(x_1,y_1), (x_2,y_2), \ldots, (x_n,y_n)$? If we call such a polynomial $P(x)$, then it has to satisfy a system of n equations

$$\begin{cases} P(x_1) = y_1, \\ P(x_2) = y_2, \\ \quad \ldots \\ P(x_n) = y_n. \end{cases}$$

Each coefficient of the polynomial P is an unknown. Since we have n equation, generally the number of unknowns should be at least n for the system to have a solution. Thus, P should have a least n coefficients. This implies that the degree of P should be at least $n-1$.

Ex: Find a polynomial that fits three points $(1,-1), (-2,0), (2,1)$.



We search for a polynomial of the form

$$P(x) = ax^2 + bx + c.$$

Because $P(1) = -1$, $P(-2) = 0$, $P(2) = 1$, we have a linear system

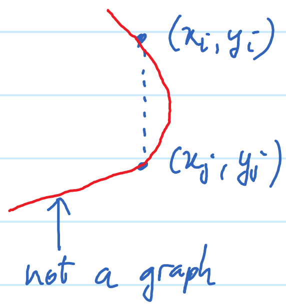
$$\begin{cases} a + b + c = -1, \\ 4a - 2b + c = 0, \\ 4a + 2b + c = 1. \end{cases}$$

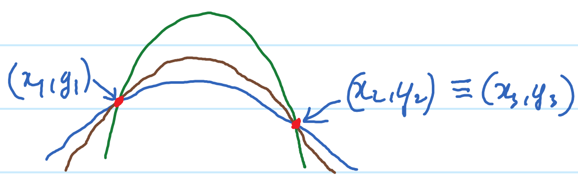
One can solve this system easily using the techniques learned from Linear Algebra I.

The method of finding P by solving a system of n equations for n unknowns (the coefficients of P) is called a direct method. This method is quite easy to understand, but the amount of calculation grows quickly if n (the number of data points) is large.

In theory, a linear system of n equations and n unknowns may have no solutions or infinitely many solutions.

Intuitively, we can see that if two of the given points are lined up vertically, then there is no polynomial curve can pass through these two points. In this case, the system has no solutions. On the other hand, if two of n given points happen to coincide each other then there are infinitely many polynomials of degree $n-1$ that fit the data points.





<u>Theorem</u>: given n points $(x_1, y_1), \ldots, (x_n, y_n)$ such that no two points are vertically lined up or coincide each other. Then there is a unique polynomial curve of degree $\le n-1$ that passes through these points.

Due to the uniqueness of the polynomial, it is called the interpolation polynomial of the given points. It is the smallest degree polynomial that fit those points. Note that there are infinitely many polynomials of degree $\geq n$ that can fit n points. If the degree is restricted to be $\le n-1$, then there is only one such polynomial. It is the interpolation formula.

In 1795, Lagrange introduced a method to find interpolation polynomials without solving any equations. His idea is as follows.

Given a set of n data points $(x_1, y_1), \ldots, (x_n, y_n)$, we first form the so-called Lagrange basis polynomials $L_1, L_2, \ldots, L_n$.

L_1 is a polynomial of degree $n-1$ that is equal to 1 at x_1 and 0 at $x_2, x_3, \ldots, x_n$. There is an incredibly simply formula for L_1:

$$L_1(x) = \frac{(x-x_2)(x-x_3)\cdots(x-x_n)}{(x_1-x_2)(x_1-x_3)\cdots(x_1-x_n)}$$

One can verify easily that $L_1(x_1)=1$, $L_1(x_2) = L_1(x_3) = \cdots = L_1(x_n) = 0$.
More generally, L_i is a polynomial of degree $n-1$ that is equal to 1 at x_i and 0 at x_j for any $j \neq i$. The formula for L_i is

$$L_i(x) = \frac{(x-x_1)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_1)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)}$$

Then the interpolation polynomial P is given as a linear combination of these basis polynomials:

$$P(x) = y_1 L_1(x) + y_2 L_2(x) + \cdots + y_n L_n(x).$$

One can check easily that $P(x_1)=y_1$, $P(x_2)=y_2$, ..., $P(x_n)=y_n$.
We will consider some examples next time.