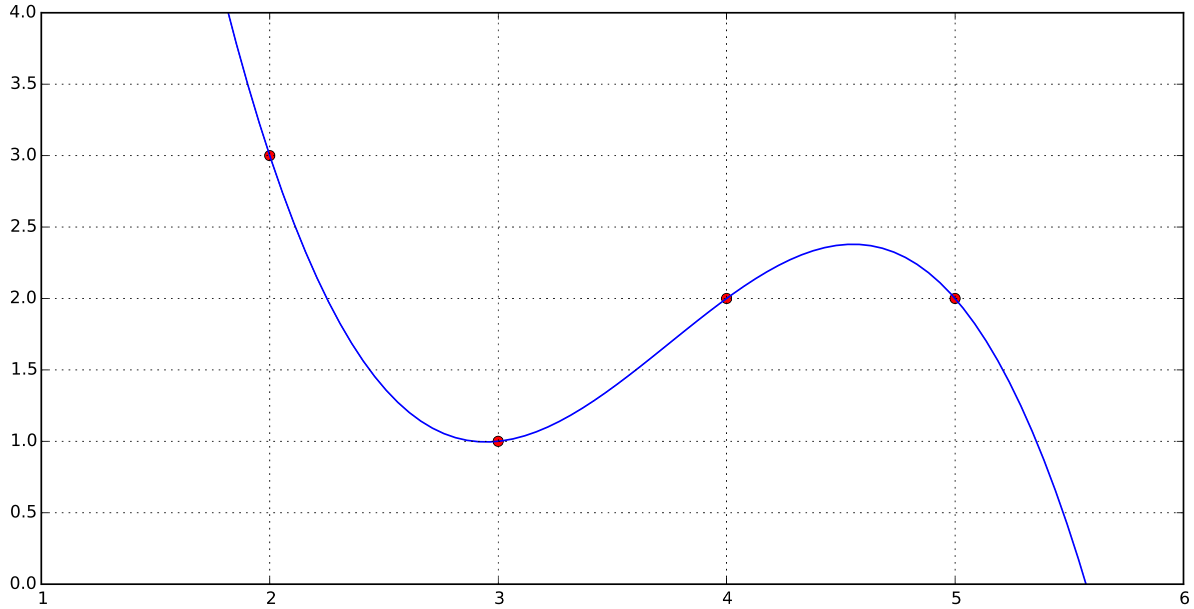
Interpolation and Integration

Polynomial Interpolation

Let us consider the problem of finding a polynomial that passes through all these points

$$(x_0, y_0)(x_1, y_1)(x_2, y_2)(x_3, y_3) = (2, 3)(3, 1)(4, 2)(5, 2)$$

as shown in the picture below



One way to solve this problem is to notice the function

$$W_0(x) = \frac{(x-3)(x-4)(x-5)}{(2-3)(2-4)(2-5)}$$

W_0 is a polynomial of degree 3 (number of data points -1). Furthermore, the function W_0 is 0 at $x \in 3, 4, 5$ and is 1 when $x = 2$. Note that 1 denotes the index of the data point. W_0 is 0 if x is from other datapoints and is 1 if $x = x_0$

We can construct similar function W_k by

$$W_k = \prod_{i=0, i \neq k}^{n} \frac{(x - x_i)}{(x_k - x_i)}$$

In particular,

$$W_1(x) = \frac{(x-2)(x-4)(x-5)}{(3-2)(3-4)(3-5)}$$
$$W_2(x) = \frac{(x-2)(x-3)(x-5)}{(4-2)(4-3)(4-5)}$$
$$W_3(x) = \frac{(x-2)(x-3)(x-4)}{(5-2)(5-3)(5-4)}$$

Now the polynomial that passes through all these points is given by

$$P(x) = 3 \times W_0(x) + 1 \times W_1(x) + 2 \times W_2(x) + 2 \times W_3(x)$$

It is easy to verified that this pass through all the points. The reason that this works is that when $x = 2$, W_1, W_2, W_3 terms are all zero and W_0 is just 1. So you are left with

$$P(x) = 3 \times W_0(2) + 0 + 0 + 0 = 3 \times 1 = 3$$

The same thing happens when x is of other data points. This polynomial is called Legendre polynomail. There are more efficient way to do this one is the coolest and the most intuititive.

Let us recap. Given the points

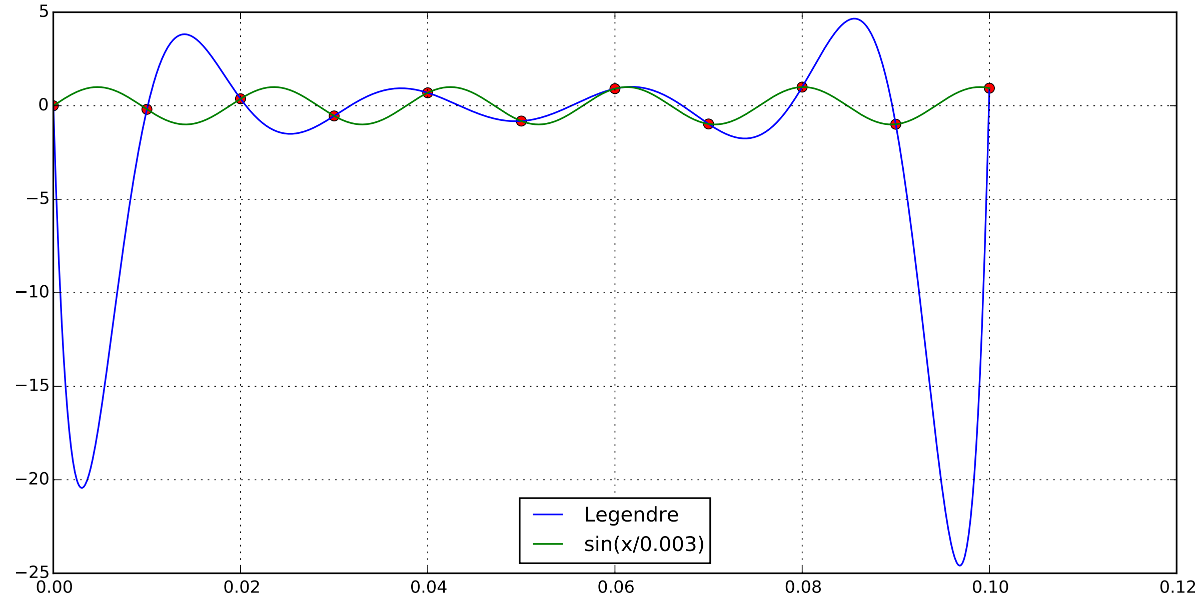
$$(x_0, y_0)(x_1, y_1)(x_2, y_2)(x_3, y_3)\dots(x_n, y_n)$$

then the polynomial of degree n that passes through all these points is given by

$$\begin{aligned} P(x) =& \sum_{k=0}^{k=n} \left(y_k \prod_{i=0, i\neq k}^{i=n} \frac{(x-x_i)}{(x_k-x_i)} \right) \\ =& y_1 \frac{(x-x_2)}{(x_1-x_2)} \cdots \frac{(x-x_n)}{(x_1-x_n)} + y_2 \frac{(x-x_1)}{(x_2-x_1)} \frac{(x-x_3)}{(x_2-x_3)} \cdots \frac{(x-x_n)}{(x_2-x_n)} + \ldots + \\ & y_n \frac{(x-x_1)}{(x_n-x_1)} \cdots \frac{(x-x_{n-1})}{(x_n-x_{n-1})} \end{aligned}$$

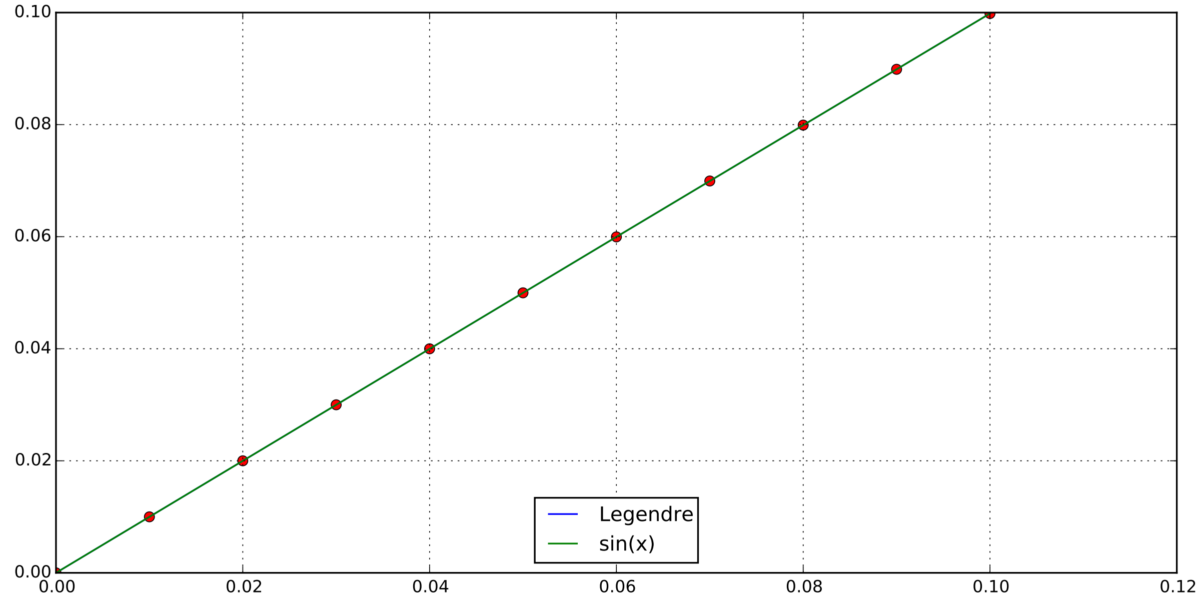
The expression looks pretty mouthful. You should try solving a couple specific problem to se that this expression make sense.

Even though the function is guarantee to pass through all the point. The function does not guarantee to be a good representation of the underlying function we draw the data points from. This can be illustrated in the picture below where we draw data points from $sin(x/0.003)$ for couple values of x from 0 to 0.1 and perform legendre polynomial interpolation.



This is indeed a lousy interpolation. Even though function pass through all the points we ask the function is not a close representation of $sin(x/0.003)$. The function doesn't even look natural. This is a characteristic of polynomial interpolation; the large swing. We will learn a way to smooth this out later.

This doesnot mean it will perform lousily in all cases. If the function is just $sin(x)$ and sample it at the same points you will see that it works really well



Interpolation error

Let us talk about the interpolation error. How far off our interpolation is from the real function. This would depend on the original function and where we are talking about. For example

$$E_n(0.5) = f(0.5) - P_n(0.5)$$

measures how far off the real function is from the polynomial we use for interpolation. The n subscript for P and E indicate the degree of the polynomial used for interpolation.

So, we are interested in the value of

$$E_n(x) = f(x) - P_n(x)$$

for a fixed value of x. There is a neat trick to find out how far this is from the real function. Watch carefully. First we define and auxilary function (t is the argument, while x is fixed)

$$Y(t) = E_n(t) - \frac{E_n(x)}{S(x)} S(t)$$

where

$$S(u) = \prod_{i=0}^{i=n} (u - x_i) = (u - x_0)(u - x_1)\dots(u - x_n)$$

There are couple things to note there

1. $S(x_0) = S(x_1) = \dots = S(x_n) = 0$ by construction.
2. $E_n(x_0) = E_n(x_1) = \dots = E_n(x_n) = 0$ since $P(n)$ pass through all the point we need.
3. When $t = x$, the two terms are equal making $Y(x) = 0$.
4. This means that $Y(t) = 0$ has $n+2$ roots. $n+1$ roots from $x_0, x_1 \dots x_k$ and another root when $t = x$.
5. By Rolle's theorem $Y'(t)$ has $n + 1$ roots.
6. Iteratively apply Rolle's theorem and we know that $Y^{(n+1)}$ as 1 root, ξ where $\xi \in [x_0, x_n]$.

This means

$$Y^{(n+1)}(t) = E_n^{n+1}(t) - \frac{E_n(x)}{S(x)} S^{n+1}(t).$$

There are couple things to notice here

1. Since E_n is a function subtract off by a polynomial of degree n. When we take $n+1$ derivative the polynomial term is gone. So,

$$E_n^{n+1}(t) = f^{(n+1)}(t) - 0 = f^{(n+1)}(t)$$

2. S is a polynomial of degreen $n + 1$ and the coefficient in front of x^{n+1} is 1. Thus,

$$S^{n+1}(t) = (n + 1)!$$

Plugging in all these back into $Y^{n+1}(t)$ we have

$$Y^{(n+1)}(t) = f^{(n+1)}(t) - \frac{E_n(x)}{S(x)}(n+1)!$$

Since we know that $\exists\xi \in [x_0, x_n]$ such that $Y^{(n+1)}(\xi) = 0$

$$Y^{(n+1)}(\xi) = f^{(n+1)}(\xi) - \frac{E_n(x)}{S(x)}(n+1)! = 0 \tag{1}$$

Rearranging the terms gives

$$E_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \times (t - x_0)(t - x_1)\ldots(t - x_n) \tag{2}$$

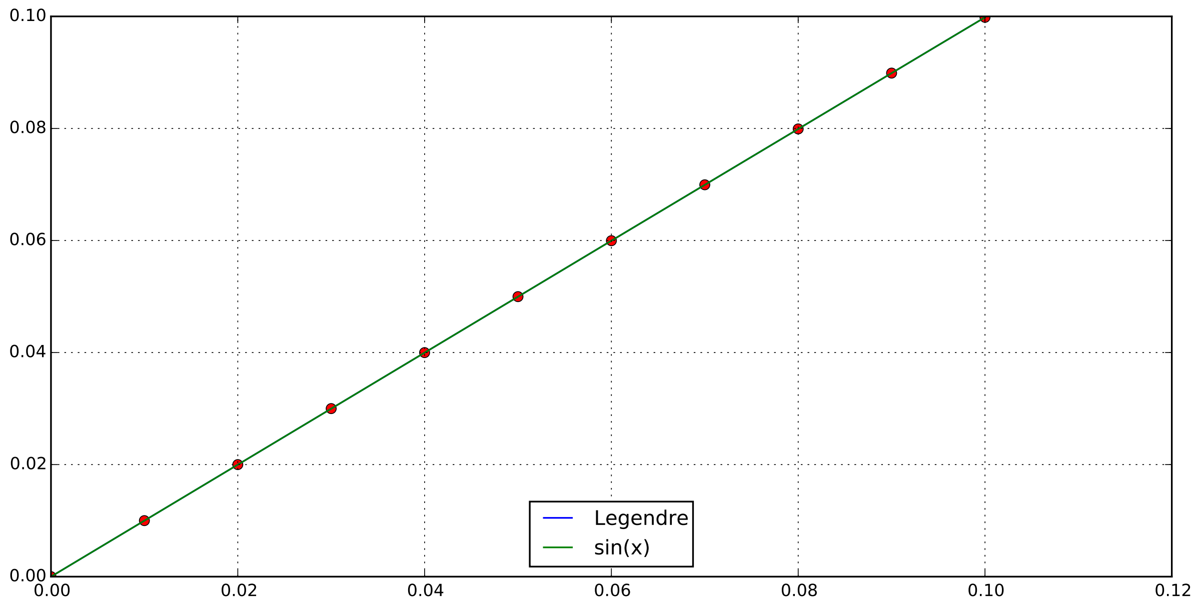
Let us write it in a more familiar form by rearranging the terms in the definition of E_n

$$f(x) = P_n(x) + E_n(x) \tag{3}$$
$$= P_n(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!} \times (t - x_0)(t - x_1)\ldots(t - x_n) \tag{4}$$

You may recognize that this looks almost exactly like Taylor's theorem. Yes this is more general form of Taylor's theorem. The Taylor's theorem is just a special case for Legendre interpolation where all the x_i are equal.

Example1

For the 10 data points we sample the $f(x) = \sin(x)$ function. We can bound the error between the $\sin(0.015)$ and $P_n(0.015)$.



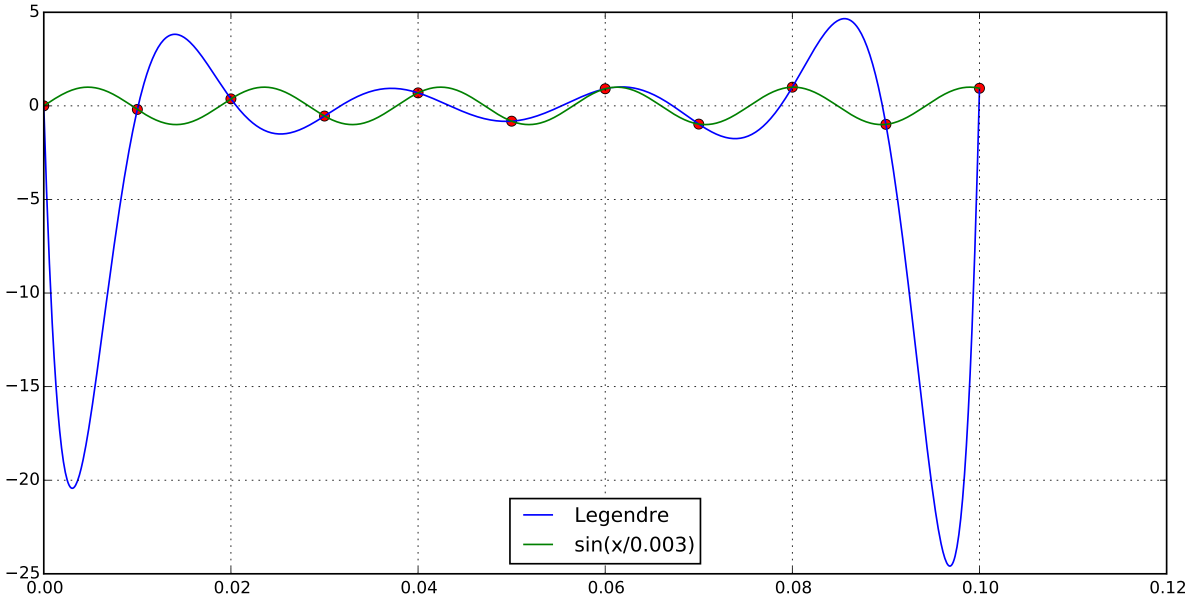
1. There are 11 points on the graph $(x_0 \dots x_{10})$. So $n = 10$.
2. To find the bound, again we use the maximum of $f^{(n+1)}(\xi)$ for $\xi \in [0, 0.1]$. The maximum of $f^{(10+1)}(\xi)$ is 1.

Thus, the bound on error is given by

$$E_n(x) \leq \frac{1}{11!} \times (0.015 - 0)(0.015 - 0.01)(0.015 - 0.02) \dots (0.015 - 0.1) \leq \text{Small number.}$$

Example2

For the 10 data points we sample the $f(x) = \sin(x/0.003)$ function. We can bound the error between the $\sin(0.015)$ and $P_n(0.015)$.



1. There are 11 points on the graph $(x_0 \dots x_{10})$. So $n = 10$.
2. To find the bound, again we use the maximum of $f^{(n+1)}(\xi)$ for $\xi \in [0, 0.1]$. The maximum of $f^{(10+1)}(\xi)$ is $\left(\frac{1}{0.003}\right)^{11} \times 1$.

Thus, the bound on error is given by

$$\begin{aligned} E_n(x) \leq & \frac{\left(\frac{1}{0.003}\right)^{11}}{11!} \times (0.015 - 0)(0.015 - 0.01)(0.015 - 0.02) \dots (0.015 - 0.1) \\ \leq & \text{a big number} \end{aligned}$$

You can see that the reason that the interpolation doesn't work too well since the function moves too fast compare the the sample points. There is a way to select the sample points to minimize the error but we will not go into that.

Numerical Integral

From calculus, most will agree that the most tricky part is how to do integral. Analytically evaluate the integral usually requires an ad-hoc solution. We learn that not all integral can be evaluated analytically. For example,

$$\int_{x=0}^{x=1} xe^{-x}\sin(x^2)\,dx$$

If you try this on Wolfram Alpha [1] asking for it to do analytical integration, it will give you an answer that you have no idea what they are.

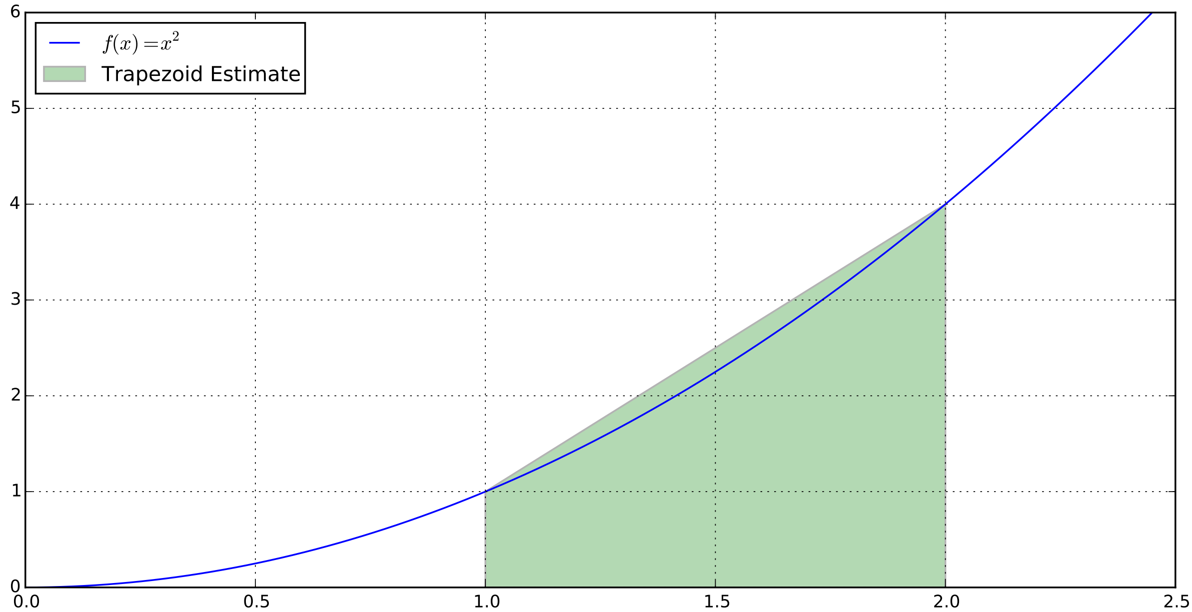
So the idea here is to avoid doing all that and evaluate the integral numerically.Integration is nothing but finding the area under the curve. We will try to approximate the area under the curve using various methods.

Trapezoid's Rule

Let us first, consider a very simple way to approximate the area under the curve. Just doing

$$\int_{x=1}^{x=2} x^2\,dx$$

Let us first plot the function along with trapezoid that assimilate the area.



The area of the trapezoid is given by

$$\int_{x=1}^{x=2} x^2\,dx \approx \text{Area} = \frac{(f(1)+f(2))}{2}\times(2-1) = \frac{1^2+2^2}{2} = \frac{5}{2} = 2.5$$

[1] http://www.wolframalpha.com/input/?i=Integrate+x*exp%28-x%29*sin%28x%5E2%29

This is a bit off from the actual answer is $7/3 \approx 2.3$. Yes, some of you may say that we can do better by subdividing it and and make smaller and smaller trapezoid. We will get to that in a minute. Let us go through how confident we are in this answer first.

Let us recap, using trapezoid rule to estimate the area is

$$\int\limits_{x=a}^{x=a+h} f(x)\, dx \approx (f(a) + f(a+h)) \times \frac{h}{2} \tag{5}$$

Trapezoid Rule's Error Analysis

Understanding the how far off Equation 5 is from the real answer is quite easy using what we found in Equation 4. We just let ($x_0 = a$ and $x_1 = a + h$). Then we have

$$\begin{aligned} f(x) =& f(a)\frac{(x-(a+h))}{(a-(a+h))} + f(a+h)\frac{(x-a)}{(a+h-a)} + \frac{f''(\xi)}{2}(x-a)(x-a-h) && (6)\\ =& f(a)\frac{a+h-x}{h} + f(a+h)\frac{(x-a)}{h} + \frac{f''(\xi)}{2}(x-a)(x-a-h) && (7)\end{aligned}$$

Integrate both sides from $x = a$ to $x = a + h$ gives

$$\begin{aligned} \int\limits_{x=a}^{x=a+h} f(x)\, dx =& \int\limits_{x=a}^{x=a+h} f(a)\frac{a+h-x}{h}\, dx + \int\limits_{x=a}^{x=a+h} f(a+h)\frac{(x-a)}{h}\, dx && (8)\\ &+ \int\limits_{x=a}^{x=a+h} \frac{f''(\xi)}{2}(x-a)(x-a-h)\, dx && (9)\\ =& f(a)\int\limits_{x=a}^{x=a+h} \frac{a+h-x}{h}\, dx + f(a+h)\int\limits_{x=a}^{x=a+h} \frac{(x-a)}{h}\, dx && (10)\\ &+ \frac{f''(\xi)}{2}\int\limits_{x=a}^{x=a+h} (x-a)(x-a-h)\, dx && (11)\\ =& f(a)\frac{h}{2} + f(a+h)\frac{h}{2} - \frac{f''(\xi)}{2}\frac{h^3}{6} && (12)\\ =& (f(a) + f(a+h)) \times \frac{h}{2} - f''(\xi)\frac{h^3}{12} && (13)\end{aligned}$$

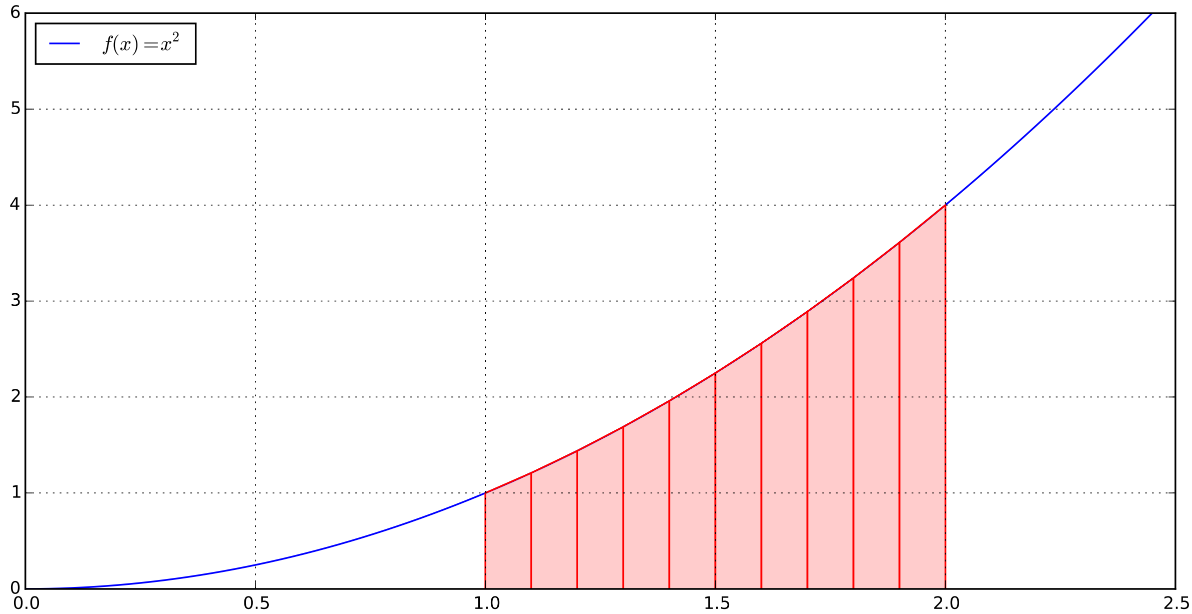
The most important thing here is the h^3. We normally say

$$\begin{aligned} \int\limits_{x=a}^{x=a+h} f(x)\, dx =& (f(a) + f(a+h)) \times \frac{h}{2} + f''(\xi)\frac{h^3}{12} && (14)\\ =& (f(a) + f(a+h)) \times \frac{h}{2} + \mathcal{O}(h^3). && (15)\end{aligned}$$

The single piece trapezoid accuracy is of order h^3. Don't be too excited yet for h^3. If we use a single trapezoid to estimate the area, our h is as huge as the bound.

Subdivision

A better way to estimate the area under the curve is to use multiple trapezoid.



Let us estimate the aread under the curve for

$$\int_{x=a}^{x=b} f(x)\, dx$$

We can see that if we sum up all these trapezoid we will have a much better estimate for the area. Let the x value for the subvision be $a, a+h, a+2h \dots a+nh$ where $a+nh = b$.

The area of the each trapezoid is given by

$$\begin{aligned}
A_0 &= (f(a) + f(a+h)) \times \frac{h}{2} \\
A_1 &= (f(a+h) + f(a+2h)) \times \frac{h}{2} \\
&\vdots \\
A_{n-1} &= (f(a+(n-1)h) + f(a+nh)) \times \frac{h}{2}
\end{aligned}$$

What's left for us to do is to sum those up

$$
\begin{aligned}
A =& A_0 + A_1 + \ldots + A_{n-1} \\
=& (f(a) + f(a+h)) \times \frac{h}{2} + (f(a+h) + f(a+2h)) \times \frac{h}{2} + (f(a+2h) + f(a+3h)) \times \frac{h}{2} + \ldots + \\
& (f(a+(n-2)h) + f(a+(n-1)h)) \times \frac{h}{2} + (f(a+(n-1)h) + f(a+nh)) \times \frac{h}{2} \\
=& \frac{h}{2} f(a) + hf(a+h) + hf(a+2h) + \ldots + hf(a+(n-1)h) + \frac{h}{2} f(a+nh) \\
=& \frac{h}{2} f(a) + hf(a+h) + hf(a+2h) + \ldots + hf(a+(n-1)h) + \frac{h}{2} f(a+nh) \\
=& \frac{h}{2} (f(a) + f(b)) + h \left[f(a+h) + f(a+2h) + \ldots + f(a+(n-1)) \right]
\end{aligned}
$$

The result is called trapezoid rule.

Error Analysis

This is pretty easy since each piece has the error of

$$E_i = \frac{1}{12} f''(\xi_i) h^3$$

where $\xi_i \in [x_i, x_{i+1}]$

So all we need to do is to sum the error up for all the pieces to get the total error

$$E_T = \frac{1}{12} f''(\xi_0) h^3 + \frac{1}{12} f''(\xi_1) h^3 + \ldots \frac{1}{12} f''(\xi_{n-1}) h^3 = \quad \frac{1}{12} h^3 \left(f''(\xi_0) + f''(\xi_1) + \ldots f''(\xi_{n-1}) \right)$$

Mean value theorem allow us the write the sum of $f''(\xi_i)$ as $nf''(\xi)$ $\exists \xi \in [a, b]$

$$E_T = \frac{1}{12} h^3 n f''(\xi)$$

and since $h = \frac{(b-a)}{n}$

$$E_T = \frac{1}{12} \frac{(b-a)^3}{n^2} f''(\xi) = \mathcal{O}(1/n^2) \tag{16}$$

This is an important result telling us that if we double the number of subdivision our answer would be 4 times more accurate. This is a desired behavior since we have multiplication for our extra effort.

Example1

How accurate is using 10 subdivision in estimation of

$$\int_{x=1}^{x=2} x^3 \, dx$$

To use Equation 16 we need to bound the second derivative of $f(x) = x^3$ for all $x \in [1, 2]$.

$$M = \max_{x\in[1,2]} |6x| = 12 \tag{17}$$

Now we can use 16

$$E_T = \frac{1}{12}\frac{(2-1)^3}{10^2} f''(\xi) \leq \frac{1}{12}\frac{(2-1)^3}{10^2} \times 12 = 0.01 \tag{18}$$

Example2

How many subdivisions we need for approximating

$$\int_{x=1}^{x=2} x^3 \, dx$$

with trapezoid rule such that the our answer is within 10^{-6} from the real answer.

With information from previous sample about the maximum of $f''(x)$ we have

$$E_T \leq \frac{1}{12}\frac{(2-1)^3}{n^2} 12 \leq 10^{-6} \tag{19}$$

Solve for n

$$n \geq \sqrt{10^6} = 1000$$

So we need at least 1000 pieces to make the answer accurate to at least 10^{-6}.