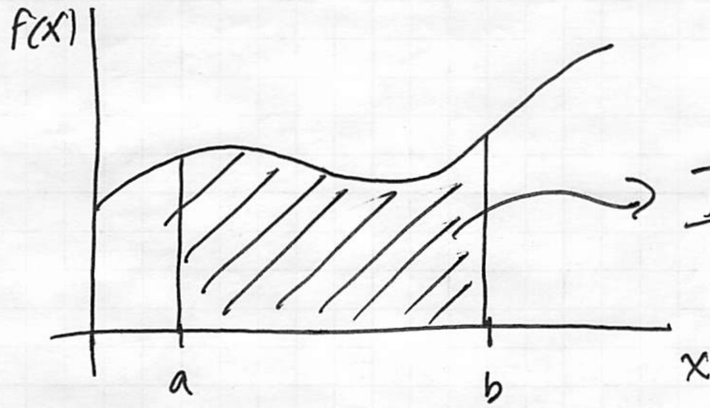
Up to now, you primary experience has been with integrals of functions with known anti-derivatives → that is, analytical integrals.

However, many functions don't have analytical anti-derivatives, but the concept of an integral is still there (i.e., the area under a curve).

Suppose we have some function $f(x)$ for which we want to compute it's integral between a & b,

$$I(a,b) = \int_a^b f(x)\,dx$$

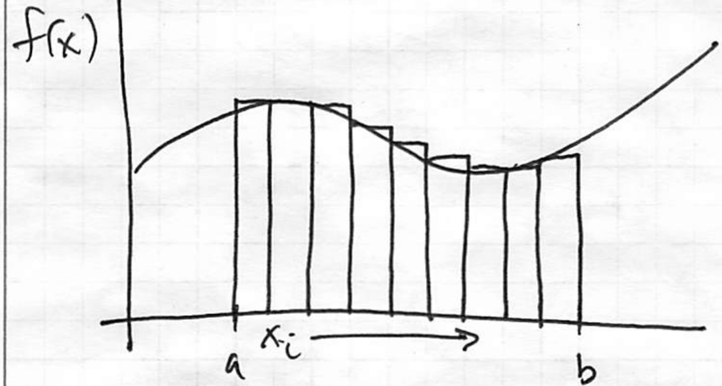


$I(a,b) \Rightarrow$ area under the curve

If we are unable to compute this integral because there's no analytic anti-derivative of the function*, $f(x)$, we can do it numerically by estimating the area under the curve.

* this technique also works for $f(x)$ where $\int f(x)dx$ is known.

Perhaps the simplest approach, which you've already thought of is using small ~~rectabl~~ rectangles



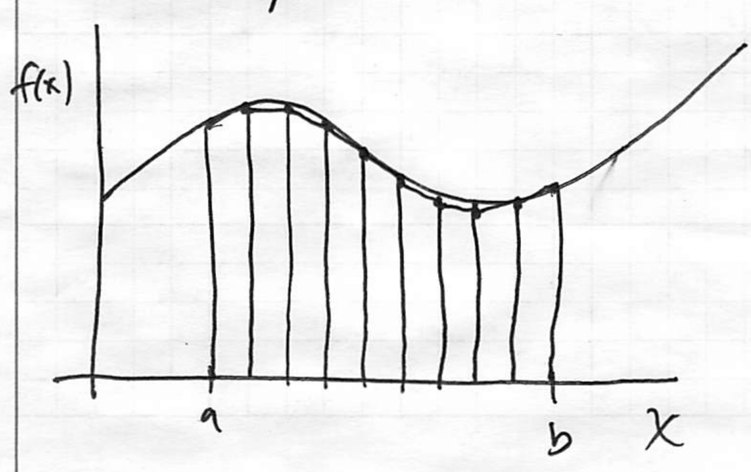
Use equally spaced rectangles with one edge (first or last) equal to $f(x_i)$ at each x_i.

We can then add up the total area using the sum of the areas of each rectangle.

⇒ This gives a poor approximation of the integral as it only takes into account the value of the function. We can do slightly better (often quite a bit better) by taking into account the value and the slope of $f(x)$.

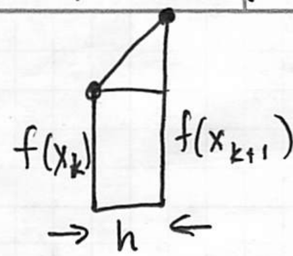
Trapezoidal Rule

If we instead take into account the approximate slope between neighboring points, we get (a much) better approximation (use Trapezoids)



Use equally spaced trapezoids instead of rectangles and add them up as before.

Area of a Trapezoid:



Area of rectangle + Area of triangle.

$$= f(x_k)h + \frac{1}{2}\left[f(x_{k+1}) - f(x_k)\right]h$$

$$A_{k+1} = \frac{1}{2}\left[f(x_k) + f(x_{k+1})\right]h$$

this makes things pretty straight-forward. Each slice contributes an amount equal to A_{k+1}. So let h be the width of the slices where, $h = (b-a)/N$ where N is the # of slices.

For the kth slice, the right hand side is at $x_k = a + kh$ and the left hand side is at $x_{k-1} = a + kh - h = a + (k-1)h$

Trapezoidal Rule: Area of k^{th} slice

$$A_k = \frac{1}{2}h\left[f(a+(k-1)h) + f(a+kh)\right]$$

Approximating our integral

So we just add up all the contributions,

$$I(a,b) \simeq \sum_{k=1}^{N} A_k = \sum_{k=1}^{N} \frac{1}{2}h\left[f(a+(k-1)h + f(a+kh)\right]$$

$$= h\left[\frac{1}{2}f(a) + f(a+h) + f(a+2h) + \ldots + f(a+(N-1)h) + \frac{1}{2}f(b)\right]$$

$$= h\left[\frac{1}{2}f(a) + \frac{1}{2}f(b) + \sum_{k=1}^{N-1} f(a+kh)\right] \leftarrow \text{algorithm ready!}$$

Example: known analytical integral

$f(x) = x^4 - 2x + 1$ from $x=0$, to $x=2$,

$$I(0,2) = \int_0^2 x^4 - 2x + 1 = \frac{1}{5}x^5 - x^2 + x \Big|_0^2 = 4.4$$

Let's use 10 slices, (Live Code this)

```
def f(x):
    return x**4 - 2*x + 1

N=10
a = 0
b = 2.0
h=(b-a)/N
s = 0.5*f(a) + 0.5*f(b)   # ½ f(a) + ½ f(b)
for k in range (1,N):
    s+= f(a+k*h)          # Σ_{k=1}^{N} f(a+kh)
print (h*s)
```

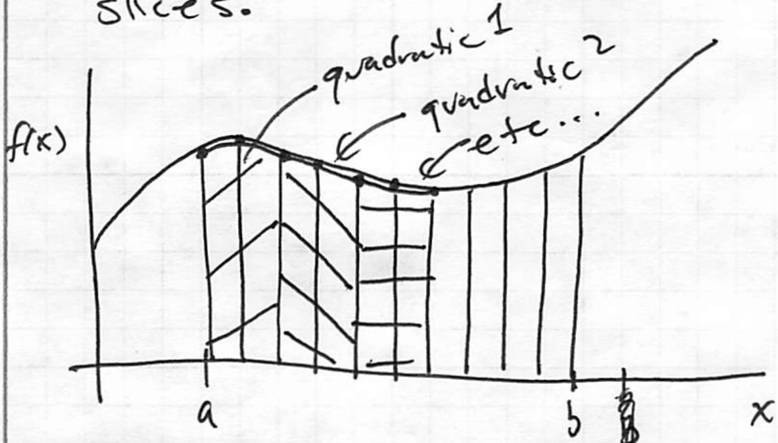
Result = 4.50656 increase steps $\Rightarrow$ 4.40001 ($N=1000$)

The trapezoidal rule work relatively well for many cases, but it can be slow (i.e. need alot of steps) and is inherently less accurate than more advanced approaches.

It only takes into account value and slope. It's a "first-order" integration method. It is only accurate up to terms proportional to h (step size). Errors are h^2 and higher.

Simpson's Rule

A better method will use value, slope, and approximate curvature. We can use quadratic functions to approximate the area of two adjacent slices.



Suppose we have three points $x = -h, 0, +h$ and we try to fit a quadratic to those points

$Ax^2 + Bx + C$ so,

$f(-h) = Ah^2 - Bh + C \quad f(0) = C \quad f(+h) = Ah^2 + Bh + C$

We can solve these for the unknown coeffs,

$C = f(0)$ (easy.)

$$A = \frac{1}{h^2}\left[\frac{1}{2}f(-h) - f(0) + \frac{1}{2}f(h)\right]$$

$$B = \frac{1}{2h}\left[f(h) - f(-h)\right]$$

the area under that quadratic approximation is,

$$\int_{-h}^{h}(Ax^2 + Bx + C)\,dx = \frac{2}{3}Ah^3 + 2Ch = \frac{1}{3}h\left[f(-h) + 4f(0) + f(h)\right]$$

This result is Simpson's rule and is very powerful b/c it only depends on the value of the function at 3 equally spaced points.

So for the pair of adjacent "bins," the area would be $\text{Area} \approx \frac{1}{3}h\left[f(x_k) + 4f(x_{k+1}) + f(x_{k+2})\right]$

The total integral is the sum of these pair of bins,

$$\begin{aligned}
I(a,b) \approx\ & \tfrac{1}{3}h\left[f(a) + 4f(a+h) + f(a+2h)\right] \\
& + \tfrac{1}{3}h\left[f(a+2h) + 4f(a+3h) + f(a+4h)\right] + \dots \\
& + \tfrac{1}{3}h\left[f(a+(N-2)h) + 4f(a+(N-1)h) + f(\overset{b}{\cancel{a+Nh}})\right]
\end{aligned}$$

We can clean this up by collecting terms,

$$I(a,b) \approx \tfrac{1}{3}h\Big[f(a) + \underbrace{4f(a+h)}_{\substack{\text{odd terms} \\ \times 4}} + \underbrace{2f(a+2h)}_{\substack{\text{even terms} \\ \times 2}} + 4f(a+3h) + \dots + f(b)\Big]$$

$$I(a,b) \approx \frac{1}{3}h\left[f(a) + f(b) + 4\sum_{k\,\text{odd}}^{N-1} f(a+kh) + 2\sum_{k\,\text{even}}^{N-2} f(a+kh)\right]$$

"trick" in python

odd terms: `for k in range(1, N, 2)` ← take 2 steps
even terms: `for k in range(2, N, 2)`

Typically Simpson's rule is much better (more efficient and more accurate) than the Trapezoidal rule. It's a third order method → accurate to h^3 with error terms of h^4 and higher.