Simpson's Rule

Saad Mehmood

September 2024

1 Introduction

Simpson's Rule is a numerical integration technique that approximates the area under a curve by fitting parabolas to segments of the curve. Although the rule is often attributed to Thomas Simpson, it was originally developed by Johannes Kepler in 1671, referred to as *Keplersche Fassregel* in German. Simpson further refined this method in the early 1700s.

There are two primary variants of Simpson's Rule: the $1/3^{\text{rd}}$ rule and the $3/8^{\text{th}}$ rule. The 3/8 rule offers improved error bounds but maintains the same order of error. The key advantage of Simpson's Rule lies in its parabolic approximation, which can more accurately model nonlinear curves compared to linear methods like the Trapezoidal Rule.

2 Problem Setup

Consider a function $f(x)$ defined over the domain $[x_0, x_n]$.

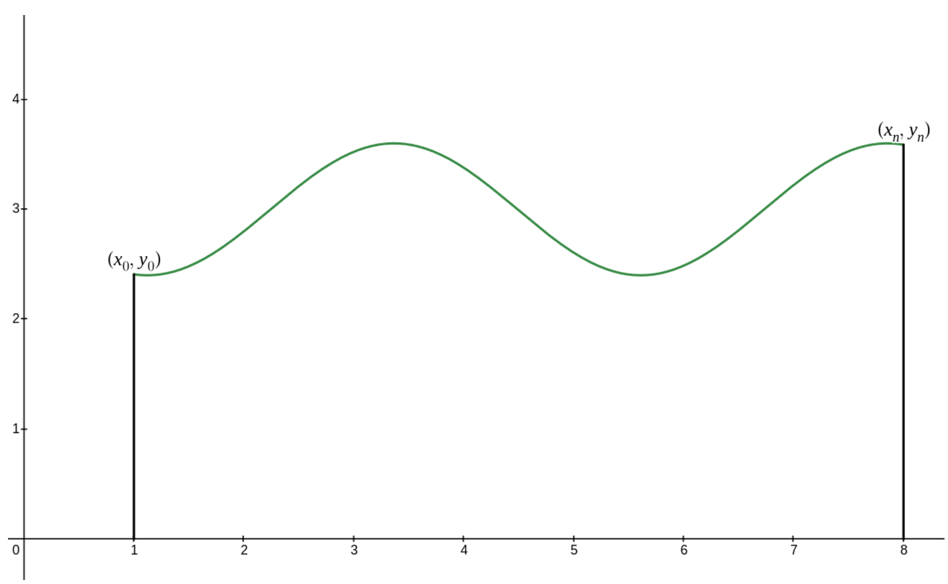

Figure 1: An arbitrary curve.

We discretize the domain into N intervals. For each subinterval, three points, P_0, P_1, and P_2, are selected. Using these points, we fit a parabola described by the general equation:

$$y = ax^2 + bx + c \tag{2.1}$$

The coefficients a, b, and c are determined by solving the system of equations based on the selected points. A visualization of this process is shown in Fig. 2.

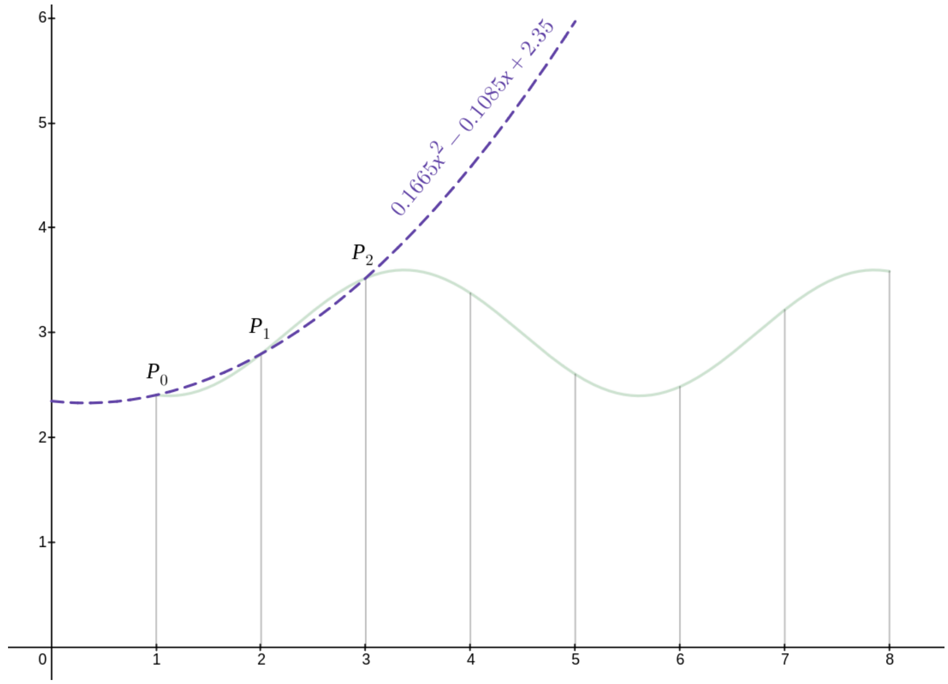

Figure 2: Discretization of the domain.

For the first three points, the equation of the fitted parabola is:

$$y = 0.1665x^2 - 0.1085x + 2.35 \tag{2.2}$$

We then compute the area under this parabola for the first three intervals. Fig. 3 highlights the approximation error, where the purple area represents the approximate area under the parabola. The red-shaded region corresponds to the overestimation, and the green-shaded region to the underestimation. These errors tend to cancel out, providing a more accurate result than linear approximations.

3 Generalized Expression

Before we move on we first have to find an expression to find the area of a parabola.

$$f(x) = ax^2 + bx + c \tag{3.1}$$

For the area we compute the definite integral of Eq. (3.1) over the domain $[-h, h]$,

$$Area = \int_{-h}^{h} ax^2 + bx + cdx \tag{3.2}$$

$$Area = \frac{ax^3}{3} + \frac{bx^2}{2} + cx\Big|_{-h}^{h} \tag{3.3}$$

$$Area = \frac{2ah^3}{3} + 2ch \tag{3.4}$$

Eq. (3.4) is the general expression for the area under a parabola. Now if we choose 3 neighbouring points in the domain $[-h, h]$ using Eq. (3.1),

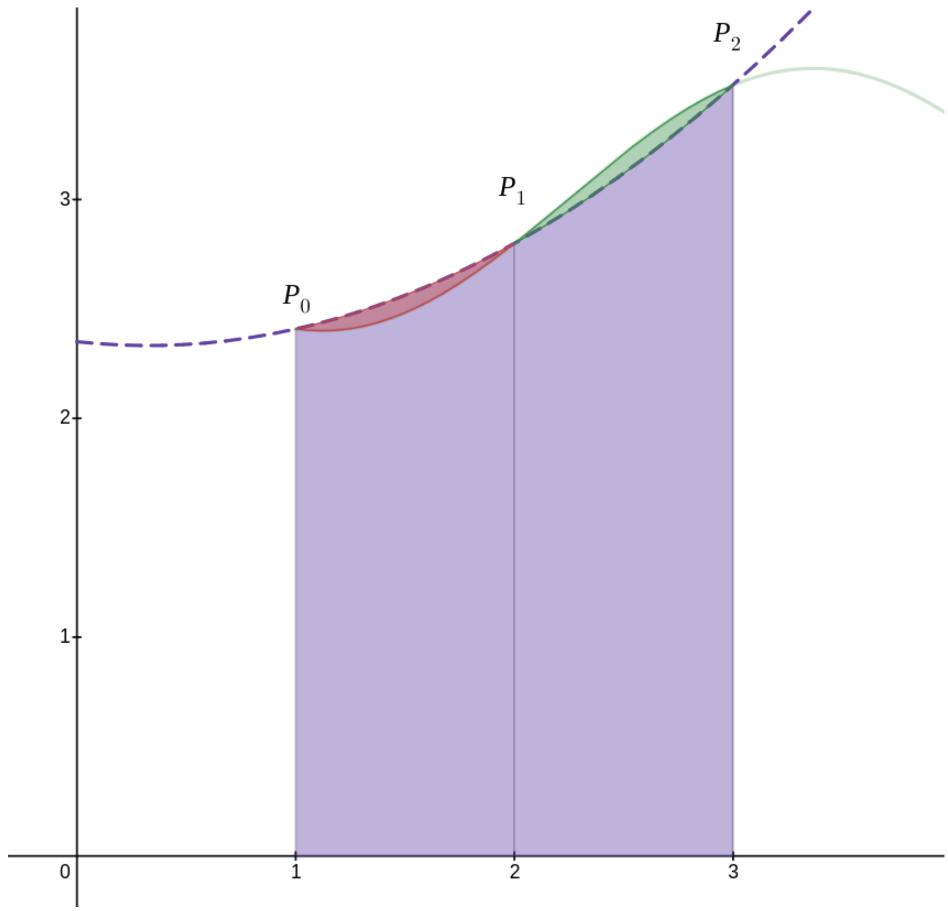

Figure 3: Approximation error for the first three intervals.

$$P_0 = y_0 = ah^2 - bh + c \tag{3.5}$$
$$P_1 = y_1 = c \tag{3.6}$$
$$P_2 = y_2 = ah^2 + bh + c \tag{3.7}$$

and we compute the expression,

$$y_0 + 4y_1 + y_2 = 2ah^2 + 6c \tag{3.8}$$

We can see a very interesting observation. Factoring out $\frac{h}{3}$ from Eq. (3.4) we can see it takes the form of Eq. (3.8).

$$Area = \frac{h}{3}(2ah^2 + 6c) \tag{3.9}$$

Hence we can write the area under a parabola in terms of the y values as;

$$Area = \frac{h}{3}(y_0 + 4y_1 + y_2) \tag{3.10}$$

Now as long as the domain is symmetric around some point, the area under the parabola can be computed using Eq. (3.10).

Using this expression for our case we can compute the under our curve by going through our domain 3 points at a time,

$$Area = \frac{h}{3}(P_0 + 4P_1 + P_2) + \frac{h}{3}(P_2 + 4P_3 + P_4) + \frac{h}{3}(P_4 + 4P_5 + P_6) + \cdots + \frac{h}{3}(P_{n-2} + 4P_{n-1} + P_n) \quad (3.11)$$

Looking at Eq. (3.11) we can make some every important observations,

1. The first and last points of our domain appear only once.

2. All points have an odd index have a coefficient of 4.

3. All even points appear twice and hence have a coefficient of 2.

Using these observations we can conclude that the general expression for the area under a curve using Simpson's method,

$$Area = \frac{h}{3}(y_0 + 4(y_{odd}) + 2(y_{even}) + y_n) \quad (3.12)$$

4 Algorithm

Using the formula derived above, we calculate the area as follows:

1. Define the domain and discretize it into N intervals.

2. Ensure that N is even. If not, adjust N by adding 1.

3. For each interval, assign weights to the function values:

- **i.** Multiply the function value by 4 if the index is odd.
- **ii.** Multiply the function value by 2 if the index is even.

4. Use the formula to compute the total area.

5 Example Problem

Consider the function:

$$f(x) = 0.4\sin(0.8x^2) + 2 \tag{5.1}$$

We compute the area under the curve for the domain $[0, 5]$. Fig. 4 shows the plot of this function.

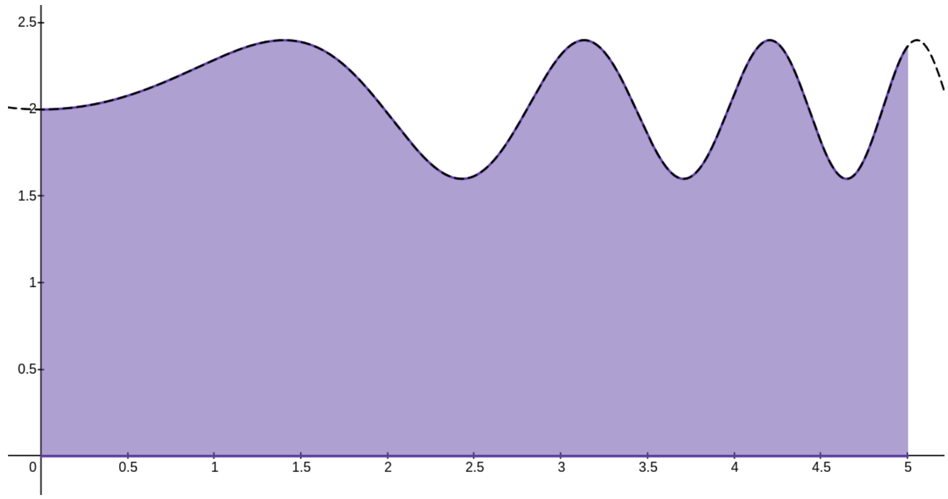

Figure 4: Plot of $f(x) = 0.4\sin(0.8x^2) + 2$.

Using Python, the results for different numbers of intervals are presented below:

No. of Intervals	Actual Area	Computed Area	Error	Runtime
10	10.2588	10.1187	1.37%	0.00006s
100	10.2588	10.2587	0.0005%	0.0004s
1000	10.2588	10.2588	0.00%	0.0024s
10,000	10.2588	10.2588	0.00%	0.0286s

Table 1: Computed areas and errors for different intervals.

6 Computation problems

Simpson's Rule is highly effective for approximating the integral of nonlinear curves due to its use of parabolas to estimate the area under the curve. It also performs well for linear functions, as a straight line can be viewed as a special case of a parabola. A critical requirement for applying Simpson's Rule is that the domain must be divided into an even number of intervals, as the method fits parabolas to groups of three points.

However, Simpson's Rule struggles with highly oscillatory functions and may not accurately approximate non-polynomial functions, potentially leading to errors in the final result.