Computational Methods Assignment 4 - Numerical Integration

Martik Aghajanian

1 Question 1

(a) To estimate the function within the domain $x_0 \leq x \leq x_2$, assuming that points x_0, , x_1, x_2 are equally spaced with spacing $h = x_2 - x_1 = x_1 - x_0$, the function $f(x)$ is written as a parabola:

$$f(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)^2$$

where $a_{0,1,2}$ are constants to be determined. The variable x is shifted to the right by the middle point x_1 so that the integration is easier to carry out. Integrating this over the interval $x_0 \leq x \leq x_2$ gives:

$$\begin{aligned}
I = \int_{x_0}^{x_2} f(x)\mathrm{d}x &= a_2 \int_{x_0}^{x_2} (x - x_1)^2 \mathrm{d}x + a_1 \int_{x_0}^{x_2} (x - x_1)\mathrm{d}x + a_0 \int_{x_0}^{x_2} \mathrm{d}x \\
&= a_2 \int_{-h}^{h} u^2 \mathrm{d}u + a_1 \int_{-h}^{h} u\mathrm{d}u + a_0 \int_{-h}^{h} \mathrm{d}u
\end{aligned}$$

where in each integral, the variable x has been substituted with $u = x - x_1 \implies \mathrm{d}u = \mathrm{d}x$. The limits hence go from $-h \to h$ since $h = x_2 - x_1 = x_1 - x_0$. The middle term, which integrates u, is odd over this interval and is hence evaluated as zero, leaving:

$$I = \int_{-h}^{h} \left(a_2 u^2 + a_0\right) \mathrm{d}u = \frac{2h^3}{3} a_2 + 2ha_0 = \frac{2h}{3}(a_2 h^2 + 3a_0)$$

The remaining problem is to evaluate the coefficients a_0, a_1, a_2, which can be done by imposing that the interpolating function passes through the points $(x_0, f(x_0))$, $(x_1, f(x_1))$, $(x_2, f(x_2))$, giving

$$\begin{aligned}
f(x_0) &= a_0 + a_1(x_0 - x_1) + a_2(x_0 - x_1)^2 = a_0 - a_1 h + a_2 h^2 \\
f(x_1) &= a_0 + a - 1(x_1 - x_1) + a_2(x_1 - x_1)^2 = a_0 \\
f(x_2) &= a_0 + a_1(x_2 - x_1) + a_2(x_2 - x_1)^2 = a_0 + a_1 h + a_2 h^2
\end{aligned}$$

Rearranging this gives:

$$\begin{aligned}
a_0 &= f(x_1) \\
a_1 &= \frac{1}{2h}(f(x_2) - f(x_0)) \\
a_2 &= \frac{1}{2h^2}(f(x_2) + f(x_0) - 2f(x_1))
\end{aligned}$$

Substituting this into the original equation for the area I:

$$\begin{aligned}
I &= \frac{2h}{3}\left(\frac{(f(x_2) + f(x_0) - 2f(x_1))}{2h^2} h^2 + 3f(x_1)\right) \\
&= \frac{2h}{3}\left(\frac{1}{2}f(x_0) + \frac{1}{2}f(x_2) - f(x_1) + 3f(x_1)\right) = h\left(\frac{1}{3}f(x_0) + \frac{4}{3}f(x_1) + \frac{1}{3}f(x_2)\right)
\end{aligned}$$

as required. For an arbitrary number N of points $x_0, x_1, \cdots, x_{N-1}$, this can be done for non-overlapping intervals and pieced together. In this extended form, the addition of areas interpolated through successive intervals $x_{2_i} \leq x \leq x_{2i+2}$ for $i = 0, 1, \cdots, \frac{N-3}{2}$, gives Simpson's rule for N points.

$$
\begin{aligned}
I &= \int_{x_0}^{x_{N-1}} f(x)\mathrm{d}x = \sum_{i=0}^{\frac{N-3}{2}} \int_{x_{2i}}^{x_{2i+2}} f(x)\mathrm{d}x \\
&= h\left(\frac{1}{3}f(x_0) + \frac{4}{3}f(x_1) + \frac{1}{3}f(x_2)\right) + h\left(\frac{1}{3}f(x_2) + \frac{4}{3}f(x_3) + \frac{1}{3}f(x_4)\right) + \cdots + h\left(\frac{1}{3}f(x_{N-3}) + \frac{4}{3}f(x_{N-2}) + \frac{1}{3}f(x_{N-1})\right)
\end{aligned}
$$

the edges of each evaluation of the parabolas between three points double up, except the 3-point intervals at the edges of the overall interval, and thus the total area evaluation by Simpson's rule is:

$$
I = h\left(\frac{1}{3}f(x_0) + \frac{4}{3}f(x_1) + \frac{2}{3}f(x_2) + \frac{4}{3}f(x_3) + \cdots + \frac{2}{3}f(x_{N-3}) + \frac{4}{3}f(x_{N-2}) + \frac{1}{3}f(x_{N-1})\right)
$$

Aside: The error term can be derived from taking the Taylor expansion of the function about the midpoint of the interval x_1, expanded with respect ot a variable δ which is small such that $x_0 \leq x_1 + \delta \leq x_2$ and truncating the expansion at quadratic order in δ. The integrals come out a similar way, and the result is obtained identicall but with a 4th order error term.

(b, c) To write the routine for the extended trapezoid rule and Simpson's rule, a C++ class *Integrate* was written which takes in an arbitrary function f, limits a, b, and specified precision ϵ upon instantiation. Within this class, updatable member variables contain the information of the integration, calling a member function *iterate()*, which increases the number of points within the interval by two and updates the corresponding integral of f over $[a, b]$. The previous value of the integral is kept to compare and assess convergence of the method. Since the rate of convergence differs for trapezoid rule and Simpson's rule, two separate member functions are used which call the *iterate()* member function until the integral *by that particular method* is within relative error ϵ of the previous iteration, signifying convergence.

(d) The results for the integral $I = \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2}\mathrm{d}z$, using both the extended trapezoid rule and extended Simpson's rule are:

$$
\begin{aligned}
\text{Trapezoid} : I = 0.995322,&\ \ \text{accuracy } \epsilon = 10^{-6} \implies \text{513 function evaluations} \\
\text{Simpson's} : I = 0.995322,&\ \ \text{accuracy } \epsilon = 10^{-6} \implies \text{65 function evaluations}
\end{aligned}
$$

This shows that the extended Simpson's rule uses far less function evaluations than the extended Trapezoid rule.