Chapter 3

Quadrature Formulas

There are several different methods for obtaining the area under an unknown curve $f(x)$ based on just values of that function at given points. During our investigations in this class we will look at the following main categories for numerical integration:

1. **Newton-Cotes formulas**
   In this case, we obtain methods for numerical integration which can be derived from the Lagrange interpolating method. Alternatively the formulas can also be derived from Taylor expansion. The idea is similar to the way we obtain numerical differentiation schemes. We can easily derive not just integration formulas but also their errors using this technique. The schemes which we develop here will be based on the assumption of equidistant points.

2. **Composite, Newton - Cotes formulas (open and closed)**
   These methods are *composite* since they repeatedly apply the simple formulas derived previously to cover longer intervals. This idea allows for piecewise estimates of the integral thus improving the error of our integration. (we will also assume equidistant nodes in our presentation).

3. **Romberg Integration**
   This method allows us to improve the error of our integration methods by doing minimal extra work. The idea is really based on the Richardson extrapolation which we saw earlier in the numerical differentiation section.

4. **Adaptive Integration**
   Here we are free to choose the points over which we calculate the numerical integral of $f(x)$ so as to minimize our error. Adaptive integration does not therefore require equidistant nodes. Thus if the function is not very smooth at some interval the step size h of the numerical integration method decreases to make sure we do not accumulate too much error in our calculation.

5. **Gaussian Integration**
   We explore methods which can achieve optimal error reduction provided we place the nodes at specific locations. Computing the best weights for our numerical quadratures guarantees optimal approximation of our integral.

3.1 Simple Newton-Cotes methods

Following the ideas developed earlier on numerical differentiation, we use once again Lagrange interpolating polynomials as a starting point in obtaining numerical integration methods. In fact in this section we learn how to derive the well-known Trapezoidal, midpoint, and Simpson formulas among others, which you have seen before in calculus classes.

As usual we start with the Lagrange interpolating formula including the error term,

$$P_n(x) = \sum_{i=0}^{n} f(x_i)L_i(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}\prod_{i=0}^{n}(x - x_i)$$

Simply integrating the above will produce a variety of numerical integration methods based on the number of nodes used. Let us look at a simple example of how exactly we can obtain our first simple formula for integration.

3.1.1 Trapezoidal Rule

For this rule all we need is to start with the Lagrange interpolating polynomial for just two points, $a = x_0$ and $b = x_1$. Thus the corresponding Lagrange interpolating polynomial with truncation error is,

$$P_1(x) = f(x_0)\frac{x - x_1}{x_0 - x_1} + f(x_1)\frac{x - x_0}{x_1 - x_0} + \frac{f^{(2)}(\xi)}{2}(x - x_0)(x - x_1)$$

where as usual $\xi \in (a, b)$. Integrating the above between $x_0 = a$ and $x_1 = b$ we obtain,

$$\int_a^b P_1(x)\,dx = f(x_0)\int_a^b \frac{x - x_1}{x_0 - x_1}\,dx + f(x_1)\int_a^b \frac{x - x_0}{x_1 - x_0}\,dx + \frac{f^{(2)}(\xi}{2!}\int_a^b (x - x_0)(x - x_1)\,dx$$

Note here that the integral applies only on the x-variables while we bring outside all known values such as $f(x_0), f(x_1), f^{(2)}(\xi)$ etc... Let us now integrate each piece on the right hand side,

$$\left[f(x_0)\frac{(x - x_1)^2}{2(x_0 - x_1)} + f(x_1)\frac{(x - x_0)^2}{2(x_1 - x_0)} + \frac{f^{(2)}(\xi)}{2}\left(\frac{x^3}{3} - \frac{(x_1 + x_0)}{2} + x_0x_1x\right)\right]_{x=x_0}^{x=x_1} \tag{3.1}$$

We can in fact write the error term above in a simpler form,

$$\frac{f^{(2)}(\xi)}{2}\left(\frac{x^3}{3} - \frac{(x_1 + x_0)}{2} + x_0x_1x\right)\Bigg]_{x_0}^{x_1} = f^{(2)}(\xi)\frac{h^3}{6}$$

assuming first that $h = x_1 - x_0$ and also using a weighted version of the mean value theorem for integrals. Therefore evaluating expression (3.1) at both end points we obtain our first formula for numerical integration for the unknown function $f(x)$,

$$\int_{x_0}^{x_1} f(x)\,dx = \frac{x_1 - x_0}{2}(f(x_0) + f(x_1)) - \frac{h^3}{6}f^{(2)}(\xi) = \frac{h}{2}(f(x_0) + f(x_1)) - \frac{h^3}{6}f^{(2)}(\xi) \tag{3.2}$$

This is the well-known Trapezoidal rule for numerical integration. Note that we have an exact description of the error for this approximation. When will there be zero error involved in numerically obtaining the integral of $f(x)$? That would be possible if the $f^{(2)}$ would be zero. That implies

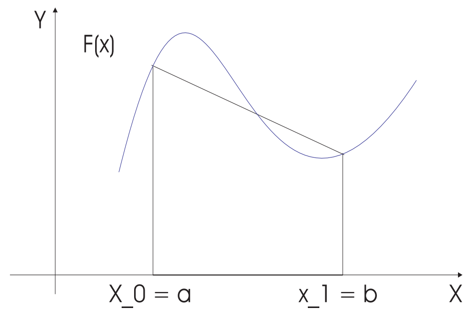

Figure 3.1: Trapezoidal rule. The integral is found by estimating the area under the curve $f(x)$ using the trapezoidal rule. Assuming that $h = x_1 - x_0$ then $A = \frac{h}{2}(f(b) + f(a))$.

that we have no error if we approximate the integral of functions which are of degree 1 or smaller. Remember that this is the famous trapezoidal rule. In other words it calculates the derivative based on the area of the trapezoid with heights $f(x_0$ and $f(x_1$. It is clear, geometrically at least based on Figure 3.1, that if the function is linear then clearly there would be no error involved in using this rule.

3.1.2 Simpson's Rule

Similarly we can derive a higher order integration scheme based on three points x_0, x_1 and x_2 of the Lagrange interpolating polynomial. That will be the well-known Simpson rule. However, we will instead present a more accurate method for obtaining Simpson's rule based on Taylor expansion around the middle point x_1 written with the error term up to forth order,

$$f(x) = f(x_1) + f^{'}(x_1)(x - x_1) + \frac{f^{''}(x_1)}{2}(x - x_1)^2 + \frac{f^{(3)}(x_1)}{3!}(x - x_1)^3 + \frac{f^{(4)}(\xi)}{4!}(x - x_4)^4$$

Thus starting with the above and integrating from x_0 to x_2 we obtain,

$$\begin{aligned}\int_{x_0=a}^{x_2=b} f(x)\,dx &= f(x_1)\int_{x_0}^{x_2} 1\,dx + f(x_1)\int_{x_0}^{x_2} x - x_1\,dx + \frac{f^{''}(x_1)}{2}\int_{x_0}^{x_2}(x - x_1)^2\,dx \\ &+ \frac{f^{(3)}(x_1)}{6}\int_{x_0}^{x_2}(x - x_1)^3\,dx + \frac{f^{(4)}(\xi)}{24}\int_{x_0}^{x_2}(x - x_1)^4\,dx\end{aligned}$$

where again using the weighted mean value theorem for integrals and $h = x_2 - x_1 = x_1 - x_0 = (b - a)/2$ we obtain,

$$\int_{x_0}^{x_2} f(x)\,dx = 2hf(x_1) + \frac{h^3}{3}f^{''}(x_1) + \frac{f^{(4)}(\xi)}{60}h^5$$

Note however that we have an approximation for the second derivative

$$f^{''}(x_1) = \frac{f(x_0) - 2f(x_1) + f(x_2)}{2h} - \frac{h^2}{12}f^{(4)}(\xi_1)$$

based on the numerical differentiation section. Thus putting this together our Simpson rule emerges:

$$\begin{aligned}\int_{x_0}^{x_2} f(x)\,dx &= 2hf(x_1) + \frac{h^3}{3}\left(\frac{f(x_0) - 2f(x_1) + f(x_2)}{2h} - \frac{h^2}{12}f^{(4)}(\xi_1)\right) + \frac{f^{(4)}(\xi)}{60}h^5 \\ &= \frac{h}{3}[f(x_0) + 4f(x_1) + f(x_2)] - \frac{h^5}{12}\left(\frac{1}{3}f^{(4)}(\xi_1) - \frac{1}{5}f^{(4)}(\xi)\right)\end{aligned}$$

In fact it can be shown (using alternate techniques) that the above formula for Simpson's rule can be improved in terms of writing down a more concise error term as follows,

$$\int_{x_0}^{x_2} f(x)\,dx = \frac{h}{3}[f(x_0)+4f(x_1)+f(x_2)] - \frac{h^5}{90}f^{(4)}(\xi)$$

where $\xi \in (x_0, x_2)$ as usual.

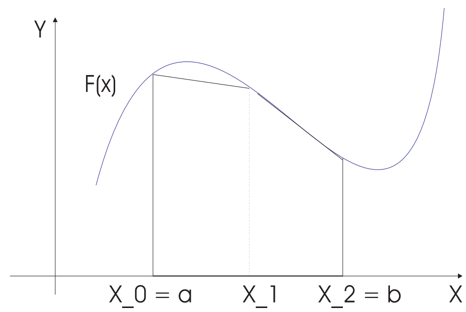

Figure 3.2: Simpson rule. The integral is found by estimating the area under the curve $f(x)$ using the Simpson rule. To do this we use three different points $x_0 = a, x_1$ and $x_2 = b$ and define $h = x_2 - x_1 = x_1 - x_0$. Then $A = \frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2))$.

Example

Apply both trapezoidal and Simpson's rule in order to approximate the area of the function $f(x) = \sqrt{x}$ between $1 \le x \le 1.30$. Also obtain estimates of the error you are committing from using either approximation.

Solution

Starting with the trapezoidal rule,

$$A_{\text{trap}} = \int_1^{1.3} \sqrt{x}\,dx \approx \frac{h}{2}(\sqrt{1.3}+\sqrt{1})$$

where $h = 1.3 - 1$. Thus the area using the trapezoidal is,

$$A_{\text{trap}} = \frac{.3}{2}(\sqrt{1.3}+\sqrt{1}) = .32102631$$

Using Simpson's rule on the other hand implies that we use three points $x_0 = 1, x_1 = 1.15$ and $x_2 = 1.3$ and the following formulation

$$A_{\text{Simpson}} = \int_1^{1.3} \sqrt{x}\,dx = \frac{.15}{3}(\sqrt{1} - 4\sqrt{1.15} + \sqrt{1.3}) = .32148417$$

Note that in fact the true area is,

$$A = \int_1^{1.5} \sqrt{x}\,dx = .32149$$

To obtain the error due to the trapezoidal rule we first need to find an upper bound for the second derivative of f in the interval $[1, 1.3]$ as follows,

$$f^{(2)}(\xi) = \frac{-1}{4\sqrt{\xi^3}} \le \frac{1}{4\sqrt{1^3}} = \frac{1}{4}$$

Therefore the error due to trapezoidal rule is found by obtaining an upper bound, in absolute value, for the following term

$$\left|\frac{h^3}{6}f^{(2)}(\xi)\right| \leq \frac{.3^3}{6}\frac{1}{4} = .001125.$$

Thus the error in our trapezoidal approximations is at most .0012 which agrees with our findings since the difference from the true area and the approximate area is $.32102631 - .32149 = .000443$. Thus the error we really made was .00044 instead of the maximum we could have made of .0012.

Similarly to find the error using Simpson's formula we first need to approximate the forth derivative

$$f^{(4)}(\xi) \leq \frac{15}{16\sqrt{\xi^7}} = \frac{15}{16\sqrt{1^7}} = \frac{15}{16}$$

Therefore the error of our approximation is

$$\left|-\frac{h^5}{90}\right| f^{(4)}(\xi) \leq \frac{.15^5}{90}\frac{15}{16} = .000000791$$

Note that indeed our true error $|.32148417 - .32149| = .00001$ is less than the possible maximum error we could have made of .0000008.

3.2 Some theoretical results

Let us now look as usual at error formulas for general integration formulas. We will refer from now on to integration formulas as **quadratures**. The formulas presented thus far are called **closed Newton-Cotes** quadratures. They are closed because the end points of the interval of integration are included in the formula. Otherwise, if the end points are not included in the formula then we have an **open Newton-Cotes** quadrature.

Before examining the general error for our quadrature formulas we give a definition which will help in understanding the accuracy of our methods:

Definition 3.2.1. *The algebraic degree of accuracy of a quadrature formula is given by the power of the polynomial $P_n(x)$ for which the quadrature is exact. That is has no error at all.*

Note for example that if we integrate the polynomial $f(x) = 3 - x$ in the interval $[1, 2]$ using the trapezoidal rule we obtain the following approximation,

$$\int_1^2 3 - x\, dx \approx \frac{h}{2}(f(1) + f(2)) = \frac{1}{2}(2 + 1) = \frac{3}{2}$$

Note that in fact the true integral is,

$$\int_1^2 3 - x\, dx = \left[3x - \frac{x^2}{2}\right]_1^2 = (6 - 2) - (3 - 1/2) = 4 - 2.5 = 3/2$$

Note that in fact the trapezoidal rule seems to be exact for the polynomial $f(x) = 3 - x$! This is true in general for any linear function and the trapezoidal rule. It is exact for all linear functions. If we try to do the same thing for a quadratic function however we will soon discover that the trapezoidal rule is not exact in that case. Therefore for the trapezoidal rule has degree of accuracy 2. See if you can find out what is the degree of accuracy for Simpson's rule.

Let us now display the error formulas for a general Newton-Cotes quadrature without proof.

Theorem 3.2.2. *Suppose the following quadrature formula for $n+1$ points*

$$\int_{x_0}^{x_n} f(x)\,dx \approx \sum_{i=0}^{n} a_i f(x_i)$$

Then the **error** *involved is given by:*

- *For n even:*

$$\begin{aligned}\underline{\textit{Closed}}\ \textit{Newton-Cotes}: \textit{error} &= \frac{h^{(n+3)} f^{(n+2)}(\xi)}{(n+2)!}\int_0^n t^2(t-2)\dots(t-n)\,dt \\ \underline{\textit{Open}}\ \textit{Newton-Cotes}: \textit{error} &= \frac{h^{(n+3)} f^{(n+2)}(\xi)}{(n+2)!}\int_{-1}^{n+1} t^2(t-2)\dots(t-n)\,dt\end{aligned}$$

- *For n odd:*

$$\begin{aligned}\underline{\textit{Closed}}\ \textit{Newton-Cotes}: \textit{error} &= \frac{h^{(n+2)} f^{(n+1)}(\xi)}{(n+1)!}\int_0^n t(t-1)\dots(t-n)\,dt \\ \underline{\textit{Open}}\ \textit{Newton-Cotes}: \textit{error} &= \frac{h^{(n+2)} f^{(n+1)}(\xi)}{(n+1)!}\int_{-1}^{n+1} t(t-1)\dots(t-n)\,dt\end{aligned}$$

where $\xi \in (x_0, x_n)$ as usual. Here for closed quadratures we use $h = (x_n - x_0)/n$ while for open we use $h = (x_{n+1} - x_{-1})/(n+2)$.

All in all we present here a few of the most commonly used Newton - Cotes formulas. Note that some are *closed* while others are *open*:

- **Midpoint**:

$$\int_{x_{-1}}^{x_1} f(x)\,dx = 2hf(x_0) + \frac{h^3}{3}f''(\xi)$$

- **Trapezoidal**:

$$\int_{x_0}^{x_1} f(x)\,dx \frac{h}{2}(f(x_0) + f(x_1)) - \frac{h^3}{12}f''(\xi)$$

- **Simpson's**:

$$\int_{x_0}^{x_2} f(x)\,dx = \frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2)) - \frac{h^5}{90}(\xi)$$

- **Simpson's 3/8 rule**:

$$\int_{x_0}^{x_3} f(x)\,dx = \frac{3h}{8}(f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)) - \frac{3h^5}{80}(\xi)$$

- **Higher order rule**:

$$\int_{x_{-1}}^{x_4} f(x)\,dx = \frac{5h}{4}(11f(x_0) + f(x_1) + f(x_2) + 11f(x_3)) + \frac{95h^5}{144}f^{(4)}(\xi)$$

3.3 Composite Quadratures

The problem of the Newton-Cotes type quadratures is similar to the problem which we encountered with Lagrange polynomials over large intervals or several nodes. Lagrange polynomials tend to display high variation under these conditions. Instead we used a piecewise method to counter this problem (the piecewise spline interpolation). Similarly here the Newton-Cotes formulas, which are essentially derived from Lagrange polynomials, are not suitable for large intervals or several nodes. Instead we can try a piecewise approach in order to reduce our errors.

In essence we will still apply the same Newton-Cotes quadratures but over several smaller intervals instead of a large one. Since finding the integral of a function is equivalent to adding the areas over several smaller intervals then this procedure should work.

Composite Trapezoidal Rule

Let us start by applying the trapezoidal rule in each of those subintervals. We can prove the following results,

Composite Trapezoidal Rule

Theorem 3.3.1. *Suppose $f \in C^2[a,b]$. Then the composite trapezoidal rule for $n+1$ points, $a = x_0, x_1, \dots, x_n = b$ is given by,*

$$\int_a^b f(x)\,dx = \frac{h}{2}\left[f(a) + 2\sum_{j=1}^{n-1} f(x_j) + f(b)\right] - \frac{(b-a)h^2}{12} f''(\xi)$$

where $h = (b-a)/n, x_j = a + jh$ and $\xi \in (a,b)$ as usual.

Proof: Starting from the integral we will split it into n subintervals,

$$\int_a^b f(x)\,dx = \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} f(x)\,dx$$

Now using the trapezoidal rule in the subinterval (x_j, x_{j+1}) we can replace the integration,

$$\sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} f(x)\,dx = \sum_{j=0}^{n-1}\left[\frac{(x_{j+1}-x_j)}{2}\big(f(x_j) + f(x_{j+1})\big) - \frac{(x_{j+1}-x_j)^3}{12} f''(\xi)\right]$$

Let us first define $h = x_{j+1} - x_j$ and then start summing up this expression, collecting like terms when applicable,

$$\begin{aligned}
\sum_{j=0}^{n-1} \quad & \left[\frac{(x_{j+1}-x_j)}{2}\big(f(x_j) + f(x_{j+1})\big) - \frac{(x_{j+1}-x_j)^3}{12} f''(\xi)\right] \\
&= \frac{h}{2}\left[f(x_0) + 2\sum_{j=0}^{n-1} f(x_j) + f(x_n)\right] - \sum_{j=0}^{n-1} h^3 \frac{f''(\xi)}{12} \\
&= \frac{h}{2}\left[f(a) + 2\sum_{j=0}^{n-1} f(x_j) + f(b)\right] - h^3 n \frac{f''(\xi)}{12} \\
&= \frac{h}{2}\left[f(a) + 2\sum_{j=0}^{n-1} f(x_j) + f(b)\right] - (b-a)h^2 \frac{f''(\xi)}{12}
\end{aligned}$$

Using similar methods we can prove the following two theorems:
Composite Midpoint Rule:

Theorem 3.3.2. *Suppose $f \in C^2[a,b]$. Then the composite Simpson's rule for $n+1$ points, $a = x_0, x_1, \ldots, x_n = b$ is given by,*

$$\int_a^b f(x)\,dx = 2h\sum_{j=0}^{n/2} f(x_{2j}) - \frac{(b-a)h^2}{6} f''(\xi)$$

where $h = (b-a)/(n+2)$, $x_j = a + (j+1)h$ for $j = -1, 0, \ldots, n, n+1$ and $\xi \in (a,b)$ as usual.

Composite Simpson's Rule

Theorem 3.3.3. *Suppose $f \in C^4[a,b]$. Then the composite Simpson's rule for $n+1$ points, $a = x_0, x_1, \ldots, x_n = b$ is given by,*

$$\int_a^b f(x)\,dx = \frac{h}{3}\left[f(a) + 2\sum_{j=1}^{n/2-1} f(x_{2j}) + 4\sum_{j=1}^{n/2} f(x_{2j-1}) + f(b)\right] - \frac{(b-a)h^4}{180} f^{(4)}(\xi)$$

where $h = (b-a)/n$, $x_j = a + jh$ and $\xi \in (a,b)$ as usual.

A simple example is in order here.

Example
Find the area under the curve $f(x) = \cos(x)$ between 0 and $\pi/2$ using the midpoint rule with an error not to exceed .001.
Solution
Based on the midpoint rule,

$$\int_0^{\pi/2} f(x)\, dx = 2h \sum_{j=0}^{n/2} f(x_{2j}) + \frac{x_n - x_0}{6} h^2 f''(\xi) \tag{3.3}$$

where $x_0 = 0, x_n = \pi/2$ and

$$h = \frac{x_n - x_0}{n+2} = \frac{\pi/2 - 0}{n+2}$$

The real question here is how big should we take n so that our error will be less than .001. Note that the error using the midpoint rule is

$$|\text{Error}| = \left(\frac{\pi/2 - 0}{6}\right)\left(\frac{\pi/2 - 0}{n+2}\right)^2 |f''(\xi)| \leq \frac{(\pi/2 - 0)^3}{6(n+2)^2} M$$

where M denotes the maximum value of,

$$|f''(\xi)| = cos(\xi) \leq 1 \equiv M \quad \text{for} \quad \xi \in (0, \pi/2)$$

As a result we need the error to be less than .001

$$\frac{(\pi/2)^3}{6(n+2)^2} = \frac{\pi^3}{48(n+2)^2} \leq .001$$

Solving the above for n we obtain,

$$n \geq \sqrt{\frac{\pi^3}{48.001}} - 2 = 23.41$$

Since n must be an integer we take it to be $n = 24$. Let us check whether this is really correct by evaluating the numerical integral from formula (3.3) with this value of n and comparing it with the exact value of the integral.

$$\begin{aligned}
\int_0^{\pi/2} \cos(x)\, dx &\approx 2\frac{\pi}{2}\left[\sum_{j=0}^{12} \cos(x_{2j})\right] \\
&= \pi[\cos(x_0) + \cos(x_2) + \cos(x_4) + \cdots + \cos(x_{20}) + \cos(x_{22})] \\
&= 1.0006
\end{aligned}$$

Note that the exact value of the integral is

$$\int_0^{\pi/2} \cos(x)\, dx = 1$$

Thus the actual error we are committing is $|1.0006 - 1.0| = .0006$. This is indeed less than .001 as we wished! Just to emphasize how accurate this is let us also calculate the midpoint approximate integral for $n = 23$ which is one less than what we should use! In that case we find using the midpoint formula that the approximate integral is .99868. This gives an error of $|.99868 - 1.0| = .002$. Which is not enough to put us in the range of .001 maximum error. Thus indeed $n = 24$ was just right in order to produce an error less than .001.

Pseudo-code The most common composite integration rule used in practice is actually Simpson's. We therefore presented below the pseudo-code for composite Simpson's rule on n subintervals: Suppose that we wish to approximate the following integral

$$\int_a^b f(x)\, dx$$

1. Let $h = (b - a)/n$
2. Initialize the following variables:

$$\begin{aligned} I_0 &= f(a) + f(b) \\ I_1 &= 0 \\ I_2 &= 0 \end{aligned}$$

3. For $i = 1, \ldots, n - 1$ do the following
   - Let X = a + ih
   - If i is even then $I_2 = I_2 + f(x)$
   - If i is odd then $I_1 = I_1 + f(x)$
4. Let
$$I = \frac{h}{3}(I_0 + 2I_2 + I_1)$$
5. Quadrature is finished. Result is I.