Note on Numerical Integration: Newton-Coates Quadrature

Makoto Nakajima, UIUC
February 2006

1 Introduction to Numerical Integration

We want to integrate a function numerically when there is no analytical formula for the integral. So our problem is:

Problem 1
We want to obtain numerically the integral of a function $f(x)$ over $[\underline{x}, \overline{x}]$

Usually it is costly to evaluate the function $f(x)$, so we are facing a trade-off between the efficiency (the computational time) and the precision.

There are two big categories in numerical integration. Furthermore, there are two big subcategories in the quadrature methods:

1. **Quadrature method:** applies an approximation method to the integral.
   (a) **Newton-Coates quadrature**
   (b) **Gaussian quadrature**
2. **Monte-Carlo method:** uses sampling method to obtain integral numerically.

This note overviews the Newton-Coates quadrature.

2 Overview of Newton-Coates Quadrature

The basic idea of the Newton-Coates quadrature is to approximate the function $f(x)$ by a collection of simple functions, and approximate the integral by computing the sum of the integral of the simple functions. In this sense, the method is similar to finite element method for functional approximation.

We are going to overview the following methods which are classified as Newton-Coates quadrature.

1. **Midpoint rule**
2. **Trapezoid rule**
3. **Simpson's rule**
4. **Romberg quadrature**
5. **Cubic spline quadrature**
6. **Adaptive quadrature**

3 Midpoint Rule

Algorithm 1 (Midpoint Rule)

1. *Divide* $[\underline{x}, \overline{x}]$ *into equally spaced intervals* $i = 1, 2, ..., N$.
2. *The size of each interval is* $h \equiv \frac{\overline{x}-\underline{x}}{N}$.
3. *Denote the midpoint of each interval as* $\{x_i\}_{i=1}^N$.
4. *The numerical integral is:*

$$\int_{\underline{x}}^{\overline{x}} f(x)dx \simeq h \sum_{i=1}^{N} f(x_i)$$

A couple of comments below:

1. Basically the integral is approximated by the sum of the area of rectangles.
2. In other words, midpoint method is approximating the function $f(x)$ using order zero Taylor approximation at the midpoint of each interval.
3. Using the Taylor expansion formula, the following can be shown:

$$\int_{\underline{x}}^{\overline{x}} f(x)dx = h \sum_{i=1}^{N} f(x_i) + \frac{h^2(\overline{x} - \underline{x})}{24} f''(\zeta)$$

   for some $\zeta \in [\underline{x}, \overline{x}]$. It implies that the error declines quadratically with the size of h.
4. This method does not use end points of $[\underline{x}, \overline{x}]$. So it can be used to approximate the derivative over an open interval. This kind of method is generally called an *open rule*.

4 Trapezoid Rule

Algorithm 2 (Trapezoid Rule)

1. *Divide* $[\underline{x}, \overline{x}]$ *into equally spaced intervals* $i = 1, 2, ..., N-1$. *The points that define the intervals are denoted as* $\{x_i\}_{i=1}^N$.
2. *The size of each interval is* $h \equiv \frac{\overline{x}-\underline{x}}{N-1}$.
3. *For an interval* i, *the integral of* $f(x)$ *over* $[x_i, x_{i+1}]$ *can be approximated by the trapezoid which consists of the four points:* $(x_i, 0)$, $(x_{i+1}, 0)$, $(x_i, f(x_i))$, $(x_{i+1}, f(x_{i+1}))$. *In other words, the integral for the interval* i *is approximated by approximating the function* $f(x)$ *using a linear function which goes through* $(x_i, f(x_i))$ *and* $(x_{i+1}, f(x_{i+1}))$.
4. *Therefor, the numerical integral is:*

$$\int_{\underline{x}}^{\overline{x}} f(x)dx \simeq \sum_{i=1}^{N-1} \frac{h(f(x_i) + f(x_{i+1}))}{2} = \frac{h}{2} \sum_{i=1}^{N-1} (f(x_i) + f(x_{i+1}))$$

A couple of comments below:

1. The method uses the end points. Therefore, the method is called a *closed rule*.
2. Define $f_i \equiv f(x_i)$. It can be shown that:

$$\int_{\underline{x}}^{\overline{x}} f(x)dx = \frac{h}{2}(f_1 + 2f_2 + ... + 2f_{N-1} + f_N) - \frac{h^2(\overline{x} - \underline{x})}{12} f''(\zeta)$$

   for some $\zeta \in [\underline{x}, ubx]$.
3. In other words, the approximation error again decreases quadratically.

5 Simpson's Rule

Let's see why first order Taylor expansion gives us the Trapezoid rule. Pick up one interval i. Use Taylor first order expansion around x_i gives us:

$$f(x) \simeq f(x_i) + f'(x_i)(x - x_i)$$

If we replace $f'(x_i)$ by $\frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$, then:

$$f(x) \simeq f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}(x - x_i)$$

Using the notation: $f_i = f(x_i)$:

$$f(x) \simeq f_i + \frac{f_{i+1} - f_i}{h}(x - x_i)$$

If we integrate the expression over $[x_i, x_{i+1}]$:

$$\begin{aligned}
\int_{x_i}^{x_{i+1}} f(x)dx &\simeq \int_{x_i}^{x_{i+1}} f_i + \frac{f_{i+1} - f_i}{h}(x - x_i)dx \\
&\simeq \left[x f_i + \frac{f_{i+1} - f_i}{h} \left(\frac{x^2}{2} - x_i x \right) \right]_{x_i}^{x_{i+1}} \\
&\simeq h f_i + \left(\frac{f_{i+1} - f_i}{h} \right) \frac{h^2}{2} \\
&\simeq \frac{h}{2} (f_{i+1} + f_i)
\end{aligned}$$

If we sum up the approximated integral across $i = 1, 2, ..., N-1$, then we have (notice that f_i with $i = 2, 3, ..., N-1$ appear twice):

$$\int_{\underline{x}}^{\overline{x}} f(x)dx \simeq \frac{h}{2}(f_1 + 2f_2 + ... + 2f_{N-1} + f_N)$$

Can we extend this into second order Taylor approximation? Notice that, in order to use the similar trick that we used to approximate the derivative of $f(x)$, we need now three points. Another way

of saying this is that we need three points in an interval to approximate the function $f(x)$ using quadratic function. For notational simplicity, let's look at the interval 1. This interval contains three points: x_1, x_2, and x_3. Since we locate the points with equal space, $x_1 = x_1 - h$, $x_3 = x_2 + h$. Obviously, the width of the interval is $2h$.

Now, the derivative of $f(x)$ at x_2 can be approximated as follows:

$$f'(x_2) \simeq \frac{f(x_2 + h) - f(x_2)}{h}$$

Similarly,

$$f'(x_1) \simeq \frac{f(x_2) - f(x_2 - h)}{h}$$

Using the equations above, we can approximate the second derivative of $f(x)$ as follows:

$$f''(x_2) \simeq \frac{f'(x_2) - f'(x_1)}{h}$$

Equivalently:

$$\begin{aligned} f''(x_2) &\simeq \frac{\frac{f(x_2+h)-f(x_2)}{h} - \frac{f(x_2)-f(x_2-h)}{h}}{h} \\ &\simeq \frac{f(x_2 + h) - 2f(x_2) + f(x_2 - h)}{h^2} \end{aligned}$$

Now, let's simplify the integration formula using second order Taylor expansion:

$$\begin{aligned} \int_{x_1}^{x_3} f(x)dx &\simeq \int_{x_1}^{x_3} f(x_2) + f'(x_2)(x - x_2) + \frac{1}{2}f''(x_2)(x - x_2)^2 dx \\ &\simeq \left[xf(x_2) + f'(x_2)\left(\frac{x^2}{2} - x_2 x\right) + \frac{1}{2}f''(x_2)\left(\frac{x^3}{3} - x^2 x_2 + x_2^2 x\right)\right]_{x_2-h}^{x_2+h} \\ &\simeq 2hf(x_2) + f'(x_2)\left(2x_2 h - 2x_2 h\right) + \frac{1}{2}f''(x_2)\left(\frac{6x_2^2 h + 2h^3}{3} - 4x_2^2 h + 2x_2^2 h\right) \\ &\simeq 2hf(x_2) + \frac{h^3}{3}f''(x_2) \end{aligned}$$

If we plug-in the approximation formula for the second derivative of $f''(x)$, we get:

$$\begin{aligned} \int_{x_1}^{x_3} f(x)dx &\simeq 2hf(x_2) + \frac{h^3}{3}\frac{f(x_2 + h) - 2f(x_2) + f(x_2 - h)}{h^2} \\ &\simeq \frac{h}{3}(f(x_1) + 4f(x_2) + f(x_3)) \end{aligned}$$

This is what we call the Simpson's rule.

If we have n intervals each of which contain 3 points, we have $N = 2n + 1$ points, and the approximated integral according to the Simpson's rule is:

$$\int_{\underline{x}}^{\overline{x}} f(x)dx \simeq \frac{h}{3}(f_1 + 4f_2 + 2f_3 + 3f_4 + ... + 2f_{N-2} + 4f_{N-1} + f_N)$$

Clearly, other than the value at end points, the coefficients attached to the value of the function at internal points oscillate between 4 and 2.

Algorithm 3 (Simpson's Rule)

1. *Divide* $[\underline{x}, \overline{x}]$ *into equally spaced* $N-1 = 2n$ *intervals* $i = 1, 2, ..., N-1$. *The points that define the intervals are denoted as* $\{x_i\}_{i=1}^{N}$.

2. *The size of each interval is* $h \equiv \frac{\overline{x}-\underline{x}}{N-1}$.

3. *According to the Simpson's rule, the integral of* $f(x)$ *over* $[\underline{x}, \overline{x}]$ *can be approximated as follows:*

$$\int_{\underline{x}}^{\overline{x}} f(x)dx \simeq h\left(\frac{1}{3}f_1 + \frac{4}{3}f_2 + \frac{2}{3}f_3 + \frac{4}{3}f_4 + ... + \frac{2}{3}f_{N-2} + \frac{4}{3}f_{N-1} + \frac{1}{3}f_N\right)$$

where $f_i = f(x_i)$.

A couple of comments:

1. It can be shown that:

$$\int_{\underline{x}}^{\overline{x}} f(x)dx = h\left(\frac{1}{3}f_1 + \frac{4}{3}f_2 + \frac{2}{3}f_3 + \frac{4}{3}f_4 + ... + \frac{2}{3}f_{N-2} + \frac{4}{3}f_{N-1} + \frac{1}{3}f_N\right) - \frac{h^4(\overline{x}-\underline{x})}{180}f''''(\zeta)$$

   for some $\zeta \in [\underline{x}, \overline{x}]$ In other words, the approximation error decreases proportion to h^4.

2. This is again a closed rule. You can construct a similar open rule using quadratic approximation.

6 Romberg Integration

6.1 Motivating Examples

Consider approximating the integral of $f(x)$ over $[\underline{x}, \overline{x}]$. Let's put three equally-spaced points $\{x_1, x_2, x_3\}$ on the interval. Denote the value of $f(x)$ at the three points as $\{f_1, f_2, f_3\}$. Let's also denote the distance between the points as h. In other words, $\overline{x} - \underline{x} = 2h$

First of all, let's use only two points $\{x_1, x_3\}$ and approximate the integral using the trapezoid rule. Let's call the approximation as $T_{1,1}$ (we'll see why shortly). We have:

$$\int_{\underline{x}}^{\overline{x}} f(x)dx \simeq T_{1,1} = (\overline{x} - \underline{x})\left(\frac{f_1}{2} + \frac{f_3}{2}\right)$$

Remember that the approximation error is proportional to $4h^2$.

How we can improve the precision of approximation? One way is to divide the interval to finer ones. For our example, suppose we use all of the three points $\{x_1, x_2, x_3\}$ to obtain the approximation. Let's call the approximation as $T_{1,2}$. Then we get:

$$\int_{\underline{x}}^{\overline{x}} f(x)dx \simeq T_{1,2} = \frac{\overline{x} - \underline{x}}{2}\left(\frac{f_1}{2} + f_2 + \frac{f_3}{2}\right)$$

What is the size of the error? It is h^2. The approximation error shrinks by $\frac{1}{4}$.

Another way to improve the approximation is to use Simpson's method. Since we need at least three points for implementing Simpson's method, let's use all of $\{x_1, x_2, x_3\}$ again. Let's call the approximation as $T_{2,1}$. The approximation takes the following form:

$$\int_{\underline{x}}^{\overline{x}} f(x)dx \simeq T_{2,1} = \frac{\overline{x} - \underline{x}}{2}\left(\frac{f_1}{3} + \frac{4f_2}{3} + \frac{f_3}{3}\right)$$

What the gain in terms of the approximation error? The error is proportional to h^4. If h is sufficiently small (equivalently, there are sufficiently large number of points on the interval), the error declines quadratically by moving from the trapezoid rule to the Simpson's rule. The gain is still substantial if we compare the trapezoid rule using three points, and the Simpson's rule.

What is the cost of moving from the 3 points trapezoid rule to the Simpson's rule? None. Both methods use exactly the same number of evaluation of $f(x)$ (three times). The only difference is the *weight* attached to each of $\{f_1, f_2, f_3\}$.

The example tells that there is a substantial gain by moving to the higher order method (like moving from the trapezoid rule to the Simpson's rule) in terms of the precision of approximation, without any additional cost (other than the complexity of the method). However, in general, it is not trivial to obtain the approximation formula for even higher order method. We can imagine from our experiences deriving the formula for the Simpson's method.

However, there is a general recursive formula to move to higher order approximation. This is the Romberg integration. Romber integration allows us to move to higher order approximation method and thus to achieve substantially smaller approximation error, without any cost in terms of the number of functional evaluations.

6.2 Constructing General Formula

Let ℓ be the order of approximation method. Since the trapezoid method uses 1st order polynomial to approximate the function $f(x)$, we denote the trapezoid method as $\ell = 1$. Similarly, the Simpson's method is $\ell = 2$.

Next, let k be an integer which represents the number of intervals n in the following way:

$$n = 2^{k-1}$$

$k = 1$ implies that there is one interval used for approximation. $k = 2$ implies that there are two intervals used for approximating the integral. Notice that the trapezoid method uses 2 points for each interval, whereas the Simpson's method requires 3 points for each interval. In our previous example, we used the trapezoid rule of $k = 1$ and $k = 2$, and the Simpson's rule of $k = 1$.

Now it should be clear why we used $T_{1,1}$ to represent the approximated integral using the trapezoid method with 1 interval, $T_{1,2}$ to represent the approximated integral using the trapezoid method with 2 intervals, and $T_{2,1}$ to represent the approximated integral using the Simpson's method with 1 interval. In general, the approximated integral is represented by $T_{\ell,k}$ where ℓ represents the order of polynomials used, and k represents (indirectly) the number of intervals used for approximation.

Remember the formula again. We have:

$$T_{1,1} = (\overline{x} - \underline{x})\left(\frac{f_1}{2} + \frac{f_3}{2}\right)$$

$$T_{1,2} = \frac{\overline{x} - \underline{x}}{2}\left(\frac{f_1}{2} + f_2 + \frac{f_3}{2}\right)$$
$$T_{2,1} = \frac{\overline{x} - \underline{x}}{2}\left(\frac{f_1}{3} + \frac{4f_2}{3} + \frac{f_3}{3}\right)$$

For these tree formula, there is the following relationship:

$$T_{2,1} = \frac{1}{3}(4T_{1,2} - T_{1,1})$$

We omit how to derive, but, in general, we can show the following recursive formula:

$$T_{\ell,k} = \frac{1}{4^{\ell-1} - 1}(4^{\ell-1}T_{\ell-1,k+1} - T_{\ell-1,k})$$

This formula implies that we can compute the approximated integral using order ℓ polynomials and $n = 2^{k-1}$ intervals, if we have both $T_{\ell-1,k+1}$, and $T_{\ell-1,k}$. In other words, we can go up in terms of the order of approximating function recursively, because we only need T associated with order $\ell - 1$ to obtain T associated with order ℓ.

Let's consider an example of how to compute $T_{5,2}$.

1. In order to obtain $T_{5,2}$, we need $T_{4,2}$ and $T_{4,3}$.
2. In order to obtain $T_{4,2}$ and $T_{4,3}$, we need $T_{3,2}$, $T_{3,3}$, and $T_{3,4}$.
3. In order to obtain $T_{3,2}$, $T_{3,3}$, and $T_{3,4}$, we need $T_{2,2}$, $T_{2,3}$, $T_{2,4}$, $T_{2,5}$.
4. In order to obtain $T_{2,2}$, $T_{2,3}$, $T_{2,4}$, $T_{2,5}$, we need $T_{1,2}$, $T_{1,3}$, $T_{1,4}$, $T_{1,5}$, $T_{1,6}$.
5. Notice that $T_{1,k}$ is just the trapezoid rule. So we can compute easily with an arbitrary k, using the following formula:
$$T_{1,k} = \frac{\overline{x} - \underline{x}}{2^{k-1}}\left(\frac{f_1}{2} + f_2 + f_3 + ... + f_{2^{k-1}} + \frac{f_{2^{k-1}+1}}{2}\right)$$
6. Now we can go backward to obtain $T_{\ell,k}$ with arbitrary ℓ and k recursively.

Remember increasing ℓ gives a substantial gain in terms of the approximation error. In general, Romberg integration can achieve the same size of error with substantially small number of function evaluation compared with other methods like the trapezoid method or the Simpson's method. There is no computational cost associated with the Romberg integration, except for the complexity of the method.

In one example that I found on the web, Romberg integration requires only 17 evaluations of the function to achieve a certain error, whereas the trapezoid and the Simpson's rule require 4097 evaluations and 129 evaluations, respectively, to achieve the same size of error.

Now, let's summarize the algorithm:

Algorithm 4 (Romberg Integration)

1. *Pick ℓ, which is the order of polynomials used for approximating the function $f(x)$.*

2. *Pick k, which determines indirectly n, which is the number of intervals used, with the following formula:*

$$n = 2^{k-1}$$

3. *Let's denote the approximation as $T_{\ell,k}$.*

4. *Compute $T_{1,j}$ where $j = k, k+1, ..., k+\ell-1$, using the following trapezoid formula:*

$$T_{1,j} = \frac{\overline{x} - \underline{x}}{2^{j-1}} \left(\frac{f_1}{2} + f_2 + f_3 + ... + f_{2^{j-1}} + \frac{f_{2^{j-1}+1}}{2} \right)$$

5. *Apply recursively the following formula until obtaining $T_{\ell,k}$:*

$$T_{\ell,k} = \frac{1}{4^{\ell-1} - 1}(4^{\ell-1} T_{\ell-1,k+1} - T_{\ell-1,k})$$

7 Other Methods

7.1 Cubic Spline Quadrature

The basic idea is to use finite element methods to approximate the function $f(x)$. Once $f(x)$ is approximated by some finite element method, it's easy to compute the integral the approximated function. The only difference from the methods that we have seen is to use finite element methods instead of polynomial approximation to approximate $f(x)$.

7.2 Adaptive Quadrature

All the Newton-Coates methods that we have seen use equally spaced intervals when approximating the function $f(x)$. But there is no reason that we need to do this. Alternative is to look at each interval, and to keep an interval finer and finer as long as there is a substantial gain in doing this. We can stop making the interval finer when the gain of making the interval finer falls below some predetermined level. This is the basic idea of the adaptive quadrature methods.

By using this strategy, we can concentrate on the part of interval where there is a larger gain in making the interval finer, and thus can reduce the number of function evaluations.

Below is a sample algorithm, using Simpson's method. Of course, there are a large variety of adaptive methods.

Algorithm 5 (Adaptive Quadrature)

1. *Choose the tolerance parameter ϵ.*

2. *We start from the entire interval $[\underline{x}, \overline{x}]$.*

3. *Put the midpoint. Denote it as x_1.*

4. *Apply Simpson's method, using $\{\underline{x}, x_1, \overline{x}\}$. Call the approximated integral S_1*

5. *Look at the two sub-intervals* $[\underline{x}, x_1]$ *and* $[x_1, \overline{x}]$ *separately. In particular, apply the steps 2-4 separately to the two intervals. Call the resulting integrals as* S_{21}, S_{22}. *The approximated integral over the entire interval* $S_2 = S_{21} + S_{22}$.

6. *Compare* S_1 *and* S_2 *If* $|S_1 - S_2| < \epsilon$, *then take* S_2 *as an acceptable approximation of the integral.*

7. *If not, we are going to dig further. First look at the interval* $[\underline{x}, x_1]$. *Apply the same steps 3-7 for the interval. Do the same thing for the other interval* $[\underline{x}, x_1]$.

8. *Notice that the procedure has a recursive structure. You keep separating the interval until the difference in approximation of the integral becomes sufficiently small* ($< \epsilon$).