Here we continue with our discussion of discrete integration. In the last class we learned how to estimate the value of the integral of a function on a given interval based only on the values of the function at the endpoints of the interval. Of the options we presented, the best estimate for the value of the integral came from the trapezoid rule, which is given by

$$\frac{f(x_1) + f(x_2)}{2} \Delta x = \int_{x_1}^{x_2} f(x) \, dx + \text{(error)}.$$  

(1)

However, generally we'll know more than just two discrete values of the function. What do we do then?

Let's set up our problem. Suppose that we have a function, $f$, that is discretized on $N$ points, $x_i \in (x_1, x_2, \ldots x_N)$. For simplicity, assume that the $x_i$ are evenly spaced, so that $x_{i+1} - x_i = \Delta x$ is a constant for all $i$. We want to estimate

$$\int_{x_1}^{x_N} f(x) \, dx$$

based on the discrete values $f(x_i)$. To proceed, we'll take advantage of what we learned from the last class; first, we'll divide the full set of discrete points into pairs of points, then we'll apply the trapezoid rule (1) on each pair of points.

So to be specific, we take the original interval, $[x_1, x_N]$, and divide it up into $N-1$ subintervals:

$N - 1$ subintervals: $[x_1, x_2], [x_2, x_3], \ldots [x_{N-1}, x_N]$.

Then, the trapezoid rule (1) applied on each interval gives us

First interval: $\frac{f(x_1) + f(x_2)}{2} \Delta x = \int_{x_1}^{x_2} f(x) \, dx + \text{(error)}$

Second interval: $\frac{f(x_2) + f(x_3)}{2} \Delta x = \int_{x_2}^{x_3} f(x) \, dx + \text{(error)}$

$\vdots$

$(N - 1)$st interval: $\frac{f(x_{N-1}) + f(x_N)}{2} \Delta x = \int_{x_{N-1}}^{x_N} f(x) \, dx + \text{(error)}$.

Adding all of these integrals together, we get

$$\int_{x_1}^{x_N} f(x) \, dx = \sum_{i=1}^{N-1} \int_{x_i}^{x_{i+1}} f(x) \, dx$$

$$= \sum_{i=1}^{N-1} \left[ \frac{\Delta x}{2} (f(x_i) + f(x_{i+1})) + \text{(error)} \right]$$

$$= \frac{\Delta x}{2} \left[ f(x_1) + f(x_2) + f(x_2) + f(x_3) + f(x_3) + \cdots ight.$$  

$$+ f(x_{N-1}) + f(x_{N-1}) + f(x_N) \right] + \text{(error)},$$

and, collecting terms, the result is

$$\Delta x \left[ \frac{1}{2} f(x_1) + \frac{1}{2} f(x_N) + \sum_{i=2}^{N-1} f(x_i) \right] = \int_{x_1}^{x_N} f(x) \, dx + \text{(error)}$$

(2)
which is known as the extended trapezoidal rule for discrete integration.

So, what is the error of the extended trapezoidal rule? We can determine this by applying the result for the error of the two-point trapezoid rule that we derived in the last class. The details are somewhat more complicated than we want to get into here, but suffice it to say that the result is a little surprising. Where the trapezoid rule from the last class was third-order accurate, as given by

\[
\frac{f(x_i) + f(x_{i+1})}{2} \Delta x = \int_{x_i}^{x_{i+1}} f(x) \, dx + \frac{f''(x_i)}{12} \Delta x^3 + \cdots = \int_{x_i}^{x_{i+1}} f(x) \, dx + O(\Delta x^3),
\]

the extended trapezoid rule turns out to be only second-order accurate, as given by

\[
\Delta x \left[ \frac{1}{2} f(x_1) + \frac{1}{2} f(x_N) + \sum_{i=2}^{N-1} f(x_i) \right] = \int_{x_1}^{x_N} f(x) \, dx + O(\Delta x^2).
\]

The reason is that since the extended trapezoidal rule is basically a summation of the two-point trapezoid rule, the errors in the two-point trapezoid rule get compounded, so that the extended trapezoidal rule is less accurate than you might expect. This does not mean, though, that the two-point rule is more accurate than the extended rule when computing the same integral, on the same interval (as we will see below). The key point when considering the accuracy of the extended trapezoidal rule is that the error term gets smaller as the subinterval size \( \Delta x \) gets smaller, which means that having more discrete points or subintervals involved in the discrete integration gives you more accurate results. The following example illustrates this.

**Example.** As a sample application of the extended trapezoidal rule, consider the integral of the sine function over the interval \( x = 0 \) to \( x = \pi \), as depicted in Fig. 1:

![Diagram](image)

Figure 1: The area of the shaded region is the integral of \( \sin x \) between \( x = 0 \) and \( x = \pi \).

The exact value of the integral is

\[
\int_0^\pi \sin x \, dx = \left[ -\cos x \right]_0^\pi = 2.
\]

We want to get a feel for the effect of different numbers of subintervals, and thus subinterval sizes, on the accuracy of the discrete integral. For starters, suppose that we know the value of the function only at the endpoints. That is, we have two discrete points, \( x_1 = 0 \) and \( x_2 = \pi \), so we can apply the regular trapezoid rule with interval size \( \Delta x = \pi \). The trapezoid rule estimate for the integral is thus

\[
\text{Trapezoid rule: } \frac{\Delta x}{2} (f(0) + f(\pi)) = \frac{\pi}{2} (\sin(0) + \sin(\pi)) = \frac{\pi}{2} (0 + 0) = 0,
\]

where \( \Delta x = \pi \), and \( f(0) = 0 \) and \( f(\pi) = 0 \).
which is obviously totally wrong.

Next, suppose that we divide the full interval into two equal subintervals, which means that we have three discrete points, \( x_1 = 0 \), \( x_2 = \pi/2 \), and \( x_3 = \pi \), and we can apply the extended trapezoid rule, with \( \Delta x = \pi/2 \), to estimate the integral as

Two subintervals:

\[
\frac{\pi}{2} \left( \frac{f(0)}{2} + \frac{f(\pi/2)}{2} + f(\pi) \right) = \frac{\pi}{2} \left( \frac{\sin(0)}{2} + \frac{\sin(\pi/2)}{2} + \sin(\pi) \right) \\
= \frac{\pi}{2} (0 + 0 + 1) \\
\approx 1.57,
\]

which is better, but still 20\% lower than the actual value.

Let’s keep going. If we divide the full interval into three subintervals, then we have four discrete points, \( x_1 = 0 \), \( x_2 = \pi/3 \), \( x_3 = 2\pi/3 \), and \( x_4 = \pi \), with \( \Delta x = \pi/3 \), and the extended trapezoidal rule gives

Three subintervals:

\[
\frac{\pi}{3} \left( \frac{f(0)}{2} + \frac{f(\pi/3)}{2} + f(\pi) + f(2\pi/3) \right) \\
= \frac{\pi}{3} \left( \frac{\sin(0)}{2} + \frac{\sin(\pi/3)}{2} + \sin(\pi/3) + \sin(2\pi/3) \right) \\
= \frac{\pi}{3} \left( 0 + 0 + \sqrt{3} + \sqrt{3} \right) \\
\approx 1.81,
\]

which is now within 10\% of the actual value of 2.

Finally, if we divide the full interval into four subintervals, then we have the five discrete points \( x_1 = 0 \), \( x_2 = \pi/4 \), \( x_3 = \pi/2 \), \( x_4 = 3\pi/4 \), and \( x_5 = \pi \), with \( \Delta x = \pi/4 \). The extended trapezoid rule then gives

Four subintervals:

\[
\frac{\pi}{4} \left( \frac{f(0)}{2} + \frac{f(\pi/4)}{2} + f(\pi/2) + f(3\pi/4) \right) \\
= \frac{\pi}{4} \left( \frac{\sin(0)}{2} + \frac{\sin(\pi/4)}{2} + \sin(\pi/4) + \sin(\pi/2) + \sin(3\pi/4) \right) \\
= \frac{\pi}{4} \left( 0 + 0 + \sqrt{2} + 1 + \sqrt{2} \right) \\
\approx 1.90,
\]

which is within 5\% of the actual value. Clearly, the discrete integral gets more accurate as the interval size \( \Delta x \) gets smaller, and for this integral at least, dividing the full interval into just four subintervals gives an estimate of the integral that’s within 5\% of the true value.