Math2411 - Calculus II
Guided Lecture Notes
Numerical Integration

University of Colorado Denver / College of Liberal Arts and Sciences

Department of Mathematics

Numerical Integration Introduction:

Our objective is to estimate integrals rather than evaluate them directly using antiderivatives. The antiderivatives of many functions either cannot be expressed or cannot be expressed easily in closed form (that is, in terms of known functions). Consequently, rather than evaluate definite integrals of these functions directly, we can use certain techniques to approximate their values. This is known as ***numerical integration***. In this section we explore several of these techniques.

The Midpoint Rule:

When we first studied definite integrals we used Riemann sums as an initial estimation for the area between the graph of a function and the x-axis. One type of Riemann sum we used is a ***Midpoint Riemann Sum***.

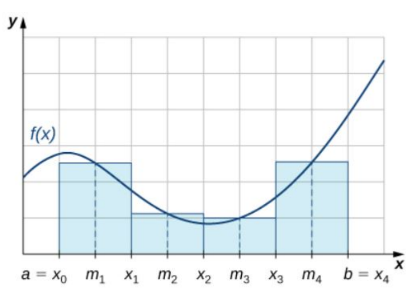

Figure 1: Midpoint Riemann Sum Estimation of Area

Then we would have

$$\int_{x=a}^{x=b} f(x)\,dx \approx f(m_1)\cdot\triangle x + f(m_2)\cdot\triangle x f(m_3)\cdot\triangle x + f(m_4)\cdot\triangle x, \text{ where } \triangle x = (b-a)/4.$$

Theorem 3.3: The Midpoint Rule

Assume that $f(x)$ is continuous on $[a, b]$. Let n be a positive integer and $\Delta x = \frac{b-a}{n}$. If $[a, b]$ is divided into n subintervals, each of length Δx, and m_i is the midpoint of the ith subinterval, set

$$M_n = \sum_{i=1}^{n} f(m_i)\Delta x. \tag{3.10}$$

Then $\lim_{n\to\infty} M_n = \int_a^b f(x)dx.$

Let's consider an example together.

Example 1. Estimate $\displaystyle\int_{x=1}^{\infty} x^2\,dx$ using M_4.

Workspace:

Solution:

Each subinterval has length $\Delta x = \frac{1-0}{4} = \frac{1}{4}$. Therefore, the subintervals consist of

$$\left[0, \frac{1}{4}\right], \left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{1}{2}, \frac{3}{4}\right], \text{ and } \left[\frac{3}{4}, 1\right].$$

The midpoints of these subintervals are $\left\{\frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}\right\}$. Thus,

$$M_4 = \frac{1}{4}f\left(\frac{1}{8}\right) + \frac{1}{4}f\left(\frac{3}{8}\right) + \frac{1}{4}f\left(\frac{5}{8}\right) + \frac{1}{4}f\left(\frac{7}{8}\right) = \frac{1}{4}\cdot\frac{1}{64} + \frac{1}{4}\cdot\frac{9}{64} + \frac{1}{4}\cdot\frac{25}{64} + \frac{1}{4}\cdot\frac{21}{64} = \frac{21}{64}.$$

Since

$$\int_0^1 x^2\,dx = \frac{1}{3} \text{ and } \left|\frac{1}{3} - \frac{21}{64}\right| = \frac{1}{192} \approx 0.0052,$$

we see that the midpoint rule produces an estimate that is somewhat close to the actual value of the definite integral.

Let's have you try an example on your own.

Example 2. Estimate the arclength of $y = x^2/2$ when $1 \leq x \leq 4$.

Workspace:

Solution:

The length of $y = \frac{1}{2}x^2$ on $[1, 4]$ is

$$\int_1^4 \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx.$$

Since $\frac{dy}{dx} = x$, this integral becomes $\int_1^4 \sqrt{1 + x^2}\, dx.$

If $[1, 4]$ is divided into six subintervals, then each subinterval has length $\Delta x = \frac{4-1}{6} = \frac{1}{2}$ and the midpoints of the subintervals are $\left\{\frac{5}{4}, \frac{7}{4}, \frac{9}{4}, \frac{11}{4}, \frac{13}{4}, \frac{15}{4}\right\}$. If we set $f(x) = \sqrt{1 + x^2}$,

$$\begin{aligned} M_6 &= \tfrac{1}{2}f\left(\tfrac{5}{4}\right) + \tfrac{1}{2}f\left(\tfrac{7}{4}\right) + \tfrac{1}{2}f\left(\tfrac{9}{4}\right) + \tfrac{1}{2}f\left(\tfrac{11}{4}\right) + \tfrac{1}{2}f\left(\tfrac{13}{4}\right) + \tfrac{1}{2}f\left(\tfrac{15}{4}\right) \\ &\approx \tfrac{1}{2}(1.6008 + 2.0156 + 2.4622 + 2.9262 + 3.4004 + 3.8810) = 8.1431. \end{aligned}$$

There are other techniques besides the midpoint rule.

The Trapezoid Rule:

One observation about the midpoint rule is that the top of the rectangles do not match the graph of the function very well which might create large errors. One attempt to fix this is to use trapezoids instead of rectangles.

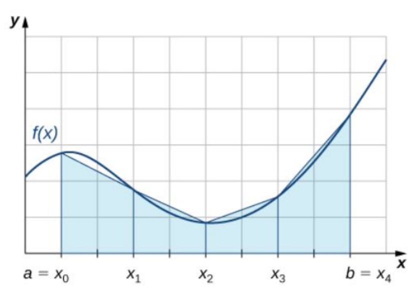

Figure 2: Trapezoid Estimation of Area

So we have $\displaystyle\int_{x=a}^{x=b} f(x)\, dx \approx \sum_{k=1}^{n} Area\big(Trapezoid_k\big).$

Fortunately, the area of a trapezoid is very easy to compute. If a trapezoid has two heights h_1 and h_2 a base with length b, then

$$Area\big(Trapezoid\big) = \frac{b}{2}\big(h_1 + h_2\big).$$

This means we have

$$\begin{aligned}\sum_{k=1}^{n} Area\big(Trapezoid_k\big) &= \frac{\triangle x}{2}\Big(f(x_0) + f(x_1)\Big) + \frac{\triangle x}{2}\Big(f(x_1) + f(x_2)\Big) + \cdots \frac{\triangle x}{2}\Big(f(x_{n-1}) + f(x_n)\Big) \\ &= \frac{\triangle x}{2}\Big(f(x_0) + \underbrace{f(x_1) + f(x_1)}_{2f(x_1)} + \underbrace{f(x_2) + f(x_2)}_{2f(x_2)} + \cdots + \underbrace{f(x_{n-1} + f(x_{n-1})}_{2f(x_{n-1})} + f(x_n)\Big)\end{aligned}$$

Theorem 3.4: The Trapezoidal Rule

Assume that $f(x)$ is continuous over $[a, b]$. Let n be a positive integer and $\Delta x = \frac{b-a}{n}$. Let $[a, b]$ be divided into n subintervals, each of length Δx, with endpoints at $P = \{x_0, x_1, x_2 \ldots, x_n\}$. Set

$$T_n = \frac{1}{2}\Delta x\big(f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)\big). \qquad (3.11)$$

Then, $\lim_{n \to +\infty} T_n = \int_a^b f(x)dx.$

Discussion:

Before we work any examples, we should observe that

$$T_n = \frac{1}{2}\Big(L_n + R_n\Big)$$

where L_n is a left Riemann sum and R_n is a right Riemann sum. Moreover, the trapezoidal rule tends to overestimate the value of a definite integral over intervals where the function is concave up and to underestimate the value of a definite integral over intervals where the function is concave down. On the other hand, the midpoint rule tends to average out these errors somewhat by partially overestimating and partially underestimating the value of the definite integral over these same types of intervals. This leads us to hypothesize that, in general, the midpoint rule tends to be more accurate than the trapezoidal rule. In fact, as we will see later, the theoretical upper bound for error with the midpoint rule M_n is one half the theoretical upper bound for the error with a trapezoid rule T_n.

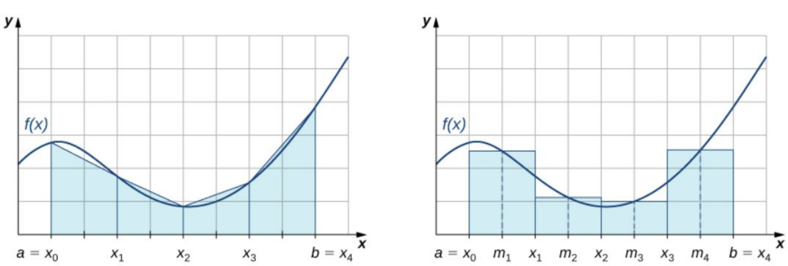

Figure 3: Trapezoid vs. Midpoint estimation of Area

Let's consider an example together.

Example 3. Estimate $\int_{x=1}^{\infty} x^2 \, dx$ using T_4.

Workspace:

Solution:

The endpoints of the subintervals consist of elements of the set $P = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$ and $\Delta x = \frac{1-0}{4} = \frac{1}{4}$. Thus,

$$\begin{aligned}\int_0^1 x^2\,dx &\approx \frac{1}{2}\cdot\frac{1}{4}\left(f(0) + 2f\left(\frac{1}{4}\right) + 2f\left(\frac{1}{2}\right) + 2f\left(\frac{3}{4}\right) + f(1)\right)\\ &= \frac{1}{8}\left(0 + 2\cdot\frac{1}{16} + 2\cdot\frac{1}{4} + 2\cdot\frac{9}{16} + 1\right)\\ &= \frac{11}{32}.\end{aligned}$$

So what about the error for these approximations? Do we have any tools to analyze the error? We define our error as follows in two ways.

Definition

If B is our estimate of some quantity having an actual value of A, then the absolute error is given by $|A - B|$. The relative error is the error as a percentage of the absolute value and is given by $\left|\frac{A-B}{A}\right| = \left|\frac{A-B}{A}\right| \cdot 100\%$.

Here is a result about analyzing the magnitude of our error.

Theorem 3.5: Error Bounds for the Midpoint and Trapezoidal Rules

Let $f(x)$ be a continuous function over $[a, b]$, having a second derivative $f''(x)$ over this interval. If M is the maximum value of $|f''(x)|$ over $[a, b]$, then the upper bounds for the error in using M_n and T_n to estimate $\int_a^b f(x)dx$ are

$$\text{Error in } M_n \le \frac{M(b-a)^3}{24n^2} \tag{3.12}$$

and

$$\text{Error in } T_n \le \frac{M(b-a)^3}{12n^2}. \tag{3.13}$$

Example 4. Estimate $\displaystyle\int_{x=1}^{\infty} e^{x^2}\,dx$ to within 0.01 using the midpoint rule..

Workspace:

Solution:

We begin by determining the value of M, the maximum value of $|f''(x)|$ over $[0, 1]$ for $f(x) = e^{x^2}$. Since $f'(x) = 2xe^{x^2}$, we have

$$f''(x) = 2e^{x^2} + 4x^2 e^{x^2}.$$

Thus,

$$|f''(x)| = 2e^{x^2}\left(1 + 2x^2\right) \le 2 \cdot e \cdot 3 = 6e.$$

From the error-bound **Equation 3.12**, we have

$$\text{Error in } M_n \le \frac{M(b-a)^3}{24n^2} \le \frac{6e(1-0)^3}{24n^2} = \frac{6e}{24n^2}.$$

Now we solve the following inequality for n:

$$\frac{6e}{24n^2} \le 0.01.$$

Thus, $n \ge \sqrt{\frac{600e}{24}} \approx 8.24$. Since n must be an integer satisfying this inequality, a choice of $n = 9$ would guarantee that $\left|\int_0^1 e^{x^2}\,dx - M_n\right| < 0.01$.

Discussion:

We might have been tempted to round 8.24 down and choose $n = 8$, but this would be incorrect because we must have an integer greater than or equal to 8.24. We need to keep in mind that the error estimates provide an upper bound only for the error. The actual estimate may, in fact, be a much better approximation than is indicated by the error bound.

Question: Can you think of an issue that we have with both the Midpoint Rule and Trapezoid Rule approximations?

Answer: We are approximating curved graphs using straight lines having zero curvature. With the midpoint rule, we estimated areas of regions under curves by using rectangles. In a sense, we approximated the curve with piecewise constant functions. With the trapezoidal rule, we approximated the curve by using piecewise linear functions.

Simpson's Rule:

What if we used piecewise quadratic functions. That is, on subintervals of $[a, b]$ use the graph of a quadratic function (meaning a parabola) to approximate the graph $y = f(x)$. This technique is known as ***Simpson's Rule***.

***Note*:** In order to make this process work we need an even number of subintervals and we use the endpoints of two adjacent intervals to build our parabolas as seen in Figure 4 provided below.

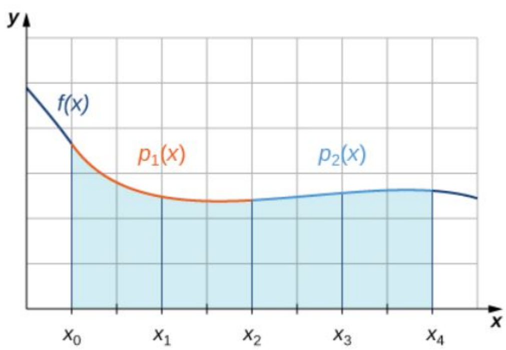

Figure 4: Simpson's Estimation of Area

Theorem 3.6: Simpson's Rule

Assume that $f(x)$ is continuous over $[a, b]$. Let n be a positive even integer and $\Delta x = \frac{b-a}{n}$. Let $[a, b]$ be divided into n subintervals, each of length Δx, with endpoints at $P = \{x_0, x_1, x_2, \ldots, x_n\}$. Set

$$S_n = \frac{\Delta x}{3}(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)). \qquad (3.14)$$

Then,

$$\lim_{n \to +\infty} S_n = \int_a^b f(x)dx.$$

The derivation of this formula is not too difficult but is very tedious. You can check the text for a derivation if you are interested. Let's now consider an example together.

Example 5. Estimate the arclength of the graph $y = x^2/2$ where $1 \le x \le 4$ using S_6.

Workspace:

Solution:

The length of $y = \frac{1}{2}x^2$ over $[1, 4]$ is $\int_1^4 \sqrt{1 + x^2}\,dx$. If we divide $[1, 4]$ into six subintervals, then each subinterval has length $\Delta x = \frac{4-1}{6} = \frac{1}{2}$, and the endpoints of the subintervals are $\left\{1, \frac{3}{2}, 2, \frac{5}{2}, 3, \frac{7}{2}, 4\right\}$. Setting $f(x) = \sqrt{1 + x^2}$,

$$S_6 = \frac{1}{3} \cdot \frac{1}{2}\left(f(1) + 4f\left(\frac{3}{2}\right) + 2f(2) + 4f\left(\frac{5}{2}\right) + 2f(3) + 4f\left(\frac{7}{2}\right) + f(4)\right).$$

After substituting, we have

$$\begin{aligned} S_6 &= \frac{1}{6}(1.4142 + 4 \cdot 1.80278 + 2 \cdot 2.23607 + 4 \cdot 2.69258 + 2 \cdot 3.16228 + 4 \cdot 3.64005 + 4.12311) \\ &\approx 8.14594. \end{aligned}$$

So what about the theoretical error for Simpson's Rule? Our geometric intuition suggests it should be much smaller for the same number of subintervals.

Rule: Error Bound for Simpson's Rule

Let $f(x)$ be a continuous function over $[a, b]$ having a fourth derivative, $f^{(4)}(x)$, over this interval. If M is the maximum value of $\left|f^{(4)}(x)\right|$ over $[a, b]$, then the upper bound for the error in using S_n to estimate $\int_a^b f(x)dx$ is given by

$$\text{Error in } S_n \leq \frac{M(b-a)^5}{180n^4}. \qquad \textbf{(3.15)}$$

Example 6. Give the theoretical error bound for $\int_{x=1}^{x=4} \sqrt{1 + x^2}\,dx$ using S_6.

Workspace:

Solution: In this problem we have $n = 6$, $n^4 = 1296$, and $(b-a)^5 = 3^5 = 243$. Next we have $f'(x) = x/\sqrt{1+x^2}$ which has a maximum value $M = 4/\sqrt{17} \approx 0.970143$ at $x = 4$. Then we have

$$\text{Error in } S_6 \leq \frac{M(b-a)^5}{180n^4} = \frac{4}{\sqrt{17}} \cdot \frac{243}{180 \cdot 1296} = 0.00101057.$$

Our approximation using S_6 is within two decimal places of accuracy. If we were to use M_6 our theoretical error bound would be

$$\text{Error in } M_6 = \frac{M(b-a)^3}{24n^2} = \frac{4}{\sqrt{17}} \cdot \frac{27}{24 \cdot 36} = 0.0058678.$$

As you can see, the theoretical error bound is improved with Simpson's Rule.

Note*:** Just because the error bound for S_6 is smaller than the error bound for M_6 ***DOES NOT mean that S_6 give a better approximation for this integral

Let's have you try an example on your own.

Example 7. Use S_2 to estimate $\displaystyle\int_{x=0}^{x=1} x^3\,dx$ and give a bound for the error.

<u>Workspace:</u>

Solution:

Since $[0,\ 1]$ is divided into two intervals, each subinterval has length $\Delta x = \frac{1-0}{2} = \frac{1}{2}$. The endpoints of these subintervals are $\left\{0,\ \frac{1}{2},\ 1\right\}$. If we set $f(x) = x^3$, then

$S_4 = \frac{1}{3} \cdot \frac{1}{2}\left(f(0) + 4f\left(\frac{1}{2}\right) + f(1)\right) = \frac{1}{6}\left(0 + 4 \cdot \frac{1}{8} + 1\right) = \frac{1}{4}$. Since $f^{(4)}(x) = 0$ and consequently $M = 0$, we see that

$$\text{Error in } S_2 \leq \frac{0(1)^5}{180 \cdot 2^4} = 0.$$

This bound indicates that the value obtained through Simpson's rule is exact. A quick check will verify that, in fact, $\int_0^1 x^3\, dx = \frac{1}{4}$.

Discussion:

In later chapters we will see there are other ways to estimate integrals using power series. But one advantage of the techniques in these notes is the explicit theoretical error bounds for each of the three approximation methods. For even more on numerical integration the interested student should consider a course in numerical analysis.

Please let me know if you have any questions, comments, or corrections!