1 Lecture: Approximate integration

- Approximating integrals by the trapezoid rule and Simpson's rule.
- Using the error estimates to compute integrals to within a specified absolute error.

1.1 Introduction

We have spent a good deal of time studying methods for finding anti-derivatives. Once we have the anti-derivative, the Fundamental theorem of calculus, part II, can be used to evaluate the definite integrals. However, there are functions for which anti-derivative cannot be written down in elementary terms. Two important examples are:

$$\int e^{-x^2}\,dx \text{ and } \int \frac{\sin x}{x}\,dx.$$

The first arises in probability theory since a multiple of e^{-x^2} is the "bell curve" or normal distribution. The second arises in electrical engineering, I think.

Thus, we would like an efficient way to find an approximate value for definite integrals such as

$$\int_0^1 e^{-t^2}\,dt.$$

Of course, we already know one way–write a Riemann sum and evaluate the sum. However, we will find that there are much better ways. These other methods are better because they require less arithmetic to obtain a specified accuracy.

We describe two methods. The trapezoid rule and Simpson's rule. For each of these rules, we approximate a function f by simpler functions that we can integrate exactly.

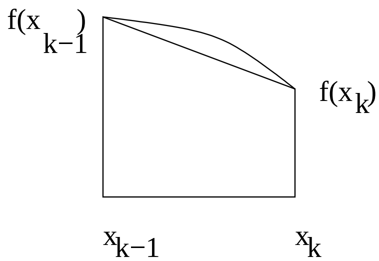
1.2 The trapezoid rule.

Suppose that we are trying to evaluate

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

As with a Riemann sum, we begin with a partition and for simplicity, we choose the regular partition $a = x_0 < x_1 < \ldots < x_n = b$ which divides the interval $[a, b]$ into n equal intervals. The points x_k are given by $x_k = a + \frac{k}{n}(b-a)$ and we let $h = (b-a)/n$ be the length of each interval. On each interval $[x_{k-1}, x_k]$, we approximate f by the linear function $L(x)$ that agrees with f at the endpoints. The area under the graph of this linear function is a trapezoid and we can compute its area easily as:

$$\frac{h}{2}(f(x_{x-1}) + f(x_k)).$$



Summing on k, gives that an approximate value for the integral is

$$\frac{h}{2}(f(x_0) + f(x_1) + f(x_1) + f(x_2) + \ldots + f(x_{n-1}) + f(x_n)). \tag{1}$$

Observe that each of the interior points of the partition occurs twice, so that we can simplify the above expression to

$$T_n = \frac{h}{2}(f(x_0) + 2f(x_1)2f(x_2) + \ldots + 2f(x_{n-1}) + f(x_n)). \tag{2}$$

This is the *trapezoid rule.*

Example. Use the trapezoid rule with $n = 5$ to approximate $\ln 3$ which is given by the integral

$$\int_1^3 \frac{1}{x}\, dx.$$

Compare the approximate value with the approximate value from your calculator.

Solution. If we divide $[1, 3]$ into n equal intervals, the length of each interval will be $(3-1)/5 = 2/5$. The regular partition is $1, 7/5, 9/5, 11/5, 13/5, 3$. The trapezoid rule gives us:

$$T_5 = \frac{1}{5}(1 + 2\frac{5}{7} + 2\frac{5}{9} + 2\frac{11}{9} + 2\frac{1}{3}9 + \frac{1}{3}) \approx 1.11027$$

My calculator tells me that $\ln 3 \approx 1.09861223$ so that the error, which we call E_T is

$$E_T = \int_1^3 \frac{1}{x}\, dx - T_5 \approx -0.01166.$$

Note that because the function $1/x$ is convex up, we know that the trapezoid rule will be larger than the integral. ∎

Remark. One could certainly complain that we need to use a calculator (or computer) to compute the trapezoid rule T_n so we might as well use the ln key on our calculator. But in this case, at least, the trapezoid rule only requires that we do arithmetic in order to compute the more complicated logarithm function. ∎

1.3 Error bound for the trapezoid rule.

The way we computed the error above is silly–if we know the logarithm accurrately then there is no need to compute the error. It is a remarkable fact that we can find a useful upper bound for the error without knowing the exact value of the integral. The error is estimated in the following theorem.

Theorem 1 *If f has two derivatives on $[a,b]$ and $|f''(x)| \leq M_2$ for all x in $[a,b]$, then the error for the trapezoid rule E_T satisfies*

$$|E_T| \leq \frac{M_2(b-a)^3}{12n^2}.$$

Notice that this theorem tells us that the error tends to zero as $n \to \infty$. Even better, it tells us that the error goes to zero at a certain rate.

Example. Use this theorem to estimate

$$T_5 - \int_1^3 \frac{1}{x}\, dx.$$

Solution. If $f(x) = 1/x$, then $f'(x) = -1/x^2$ and $f''(x) = +2/x^3$. Since $1/x^2$ is decreasing on $[1,3]$, the maximum value occurs at $x = 1$ and so we can let $M_2 = 2$. Thus

$$E_T = \frac{16}{12 \cdot 5^2} = .05\bar{3}$$

∎

We see that in this case, the error estimate is of the same order of magnitude as the error we found above.

A more interesting use of the error estimate is that we can use it find a value of n which guarantees that the error is less than a predetermined tolerance.

Example. Find a value of n so that T_n is within 10^{-4} of the integral $\int_1^3 \frac{1}{x}\, dx$.

Solution. We use the value of $M_2 = 2$ for the function $1/x$ on the interval $[1,3]$ which we found in the previous example. Thus the error T_n is less than $2 \cdot (2-1)^3/(12n^2)$. Thus we want to find n so that

$$\frac{16}{12n^2} < 10^{-4}.$$

Solving this inequality for n gives

$$\frac{4}{3} \cdot 10^4 \leq n^2$$

or that

$$115.47 < n.$$

Since n has to be a whole number we choose $n = 116$. ∎

1.4 Simpson's rule

In principle, the trapezoid rule is enough. With enough computing power, we can choose n large enough and approximate the integral as accurately as we like. This is not as easy as it sounds. To compute to 10^{-13} accuracy (as your calculator does), we would need n to be around 10^6.

A bit of thought lets us produce a rule that is much more accurate, and thus requires less work.

In the trapezoid rule, we replaced f by a function which was linear on each interval and computed the integral of the simpler function. An obvious thing to try is to approximate f by a polynomial on intervals and then integrate the polynomial. We consider the case of quadratic polynomials, or parabolas. Notice that a quadratic expression $Q(x) = Ax^2 + Bx + C$ has three coefficients, so it is natural to try and find a quadratic function which agrees with f at three points.

Suppose that we know $f(-h)$, $f(0)$ and $f(h)$. This is a bit of a mess, but if we ask that $Q(-h) = f(-h)$, $Q(0) = f(0)$ and $Q(h) = f(h)$, we obtain the system of equations

$$\begin{aligned} Ah^2 + Bh + C &= f(h) \\ C &= f(0) \\ Ah^2 - Bh + C &= f(-h) \end{aligned}$$

Solving this system, we obtain

$$Q(x) = \frac{1}{2h^2}(f(-h)(x^2 - xh) + 2f(0)(h^2 - x^2) + f(h)(x^2 + xh)).$$

If we are brave and compute we find

$$\int_{-h}^{h} Q(x)\, dx = \frac{h}{3}(f(-h) + 4f(0) + f(h)).$$

Now, we consider an interval $[a, b]$ and use the regular partition with n points. If n is even, we can apply the above argument on the interval $[x_0, x_2]$, $[x_2, x_4] \ldots [x_{2n-2}, X_{2n}]$, we obtain the following approximate expression for the integral $\int_a^b f(x)\, dx$,

$$\frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2) + f(x_2) + 4f(x_3) + f(x_4) + \ldots + f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)).$$

Again, we collect the repeated terms and obtain *Simpson's rule* for n even:

$$S_n = \frac{h}{3}(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3)2f(x_4) + \ldots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)).$$

To illustrate this rule, we approximate an integral.

Example. Use Simpson's rule rule with $n = 6$ to approximate

$$\int_1^3 \frac{1}{x}\, dx.$$

Use the value of $\ln 3$ from your calculator to find the error.

Solution. We use subintervals of length $h = 1/3$. We put $f(x) = 1/x$. Simpson's rule for $n = 6$ reads

$$S_6 = \frac{1}{6}(1 + 4 \cdot \frac{3}{4} + 2\frac{3}{5} + 4 \cdot \frac{1}{2} + 2 \cdot \frac{3}{7} + 4 \cdot \frac{3}{8} + \frac{1}{3}) \approx 1.0989.$$

We have that $\ln 3 - 1.0989 \approx 3 \cdot 10^{-4}$.

∎

It is amazing that the pattern of coefficients in Simpson's rule gives such an improvement.

As with the trapezoid rule, there is an error estimate which reads:

Theorem 2 *If f has four derivatives on $[a, b]$ and $|f^{(4)}(x)| \leq M_4$ for all x in $[a, b]$, then the error for the trapezoid rule E_S satisfies*

$$|E_S| \leq \frac{M_4(b-a)^5}{180n^4}.$$

Notice that the exponent here is 4, while it was 2 for the trapezoid rule. This tells us that for a fixed n, Simpson's rule should give us approximately twice as many decimal places correct.

Example. Find a value of n so that when we approximate $\int_1^3 \frac{1}{x}\,dx$, by S_n, we obtain an error of at most 10^{-4}.

Solution. If $f(x) = 1/x$, then $f^{(4)}(x) = 24/x^5$. We may use $M_4 = 24$ on the interval $[1, 3]$. We want the error to be less than 10^{-4}. Thus we want

$$\frac{24 \cdot 32}{180n^4} < 10^{-4}.$$

Solving this inequality for n gives

$$\frac{4.2\bar{6} \cdot 10^4}{180} < n^4$$

or $14.37 < n$. Since n must be even, we choose $n = 16$. ∎

Example. Explain how to compute π to within an error of 10^{-2}.

Solution. There are many ways to compute π. A nice one is:

$$\pi = 4\tan^{-1}(1) = 4\int_0^1 \frac{1}{1+x^2}\,dx.$$

Thus, we need to evaluate the integral $\int_0^1 \frac{1}{1+x^2}\,dx$ to within an error of $\frac{1}{4}\cdot 10^{-2}$.

To estimate the error in the trapezoid rule, we need an upper bound on the second derivative of $f(x) = 1/(1+x^2)$. Compute: $f'(x) = -2x/(1+x^2)^2$ and $f''(x) = (6x^2-2)/(1+x^2)^3$. A rough estimate is $f''(x) \leq 4$ for x in $[0,1]$. Thus

$$|E_T| \leq \frac{4\cdot 1^3}{12n^2} \leq \frac{1}{4}10^{-2}.$$

or $\frac{4}{3}10^2 \leq n^2$. or $n \approx 11.54$. Thus $n = 12$ will do.

If we are industrious and do the arithmetic, we obtain

$$\pi \approx 3.1404.$$

∎

Remark. One could also try to use Simpson's rule. But finding the fourth derivative of $f(x) = 1/(1+x^2)$ is quite a mess. ∎