MATH50003 Numerical Analysis

Sheehan Olver

November 6, 2023

Chapter I

Calculus on a Computer

In this first chapter we explore the basics of mathematical computing and numerical analysis. In particular we investigate the following mathematical problems which can not in general be solved exactly:

1. Integration. General integrals have no closed form expressions. Can we use a computer to approximate the values of definite integrals?
2. Differentiation. Differentiating a formula as in calculus is usually algorithmic, however, it is often needed to compute derivatives without access to an underlying formula, eg, of a function defined only in code. Can we use a computer to approximate derivatives? A very important application is in Machine Learning, where there is a need to compute gradients to determine the "right" weights in a neural network.
3. Root finding. There is no general formula for finding roots (zeros) of arbitrary functions, or even polynomials that are of degree 5 (quintics) or higher. Can we compute roots of general functions using a computer?

In this chapter we discuss:

1. I.1 Rectangular rule: We review the rectangular rule for integration and deduce the *converge rate* of the approximation. In the lab we investigate its implementation as well as extensions to the Trapezium rule.
2. I.2 Divided differences: We investigate approximating derivatives by a divided difference and again deduce the convergence rates. In the lab we extend the approach to the central differences formula and computing second derivatives. We also observe a mystery: the approximations may have significant errors in practice, and there is a limit to the accuracy.
3. I.3 Dual numbers: We introduce the algebraic notion of a *dual number* which allows implementing *forward-mode automatic differentiation*, a high accuracy alternative to divided differences for computing derivatives.
4. I.4 Newton's method: We review Newton's method for root finding, leveraging dual numbers for computing the derivatives.

I.1 Rectangular rule

One possible definition for an integral is the limit of a Riemann sum, for example:

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} h \sum_{k=1}^n f(x_k)$$

where $x_k = a + kh$ are evenly spaced points dividing up the interval $[a, b]$, that is wit the {\it step size} $h = (b-a)/n$. This suggests an algorithm known as the *(right-sided) rectangular rule* for approximating an integral: choose n large so that

$$\int_a^n f(x)\mathrm{d}x \approx h \sum_{k=1}^n f(x_k).$$

In the lab we explore practical implementation of this approximation, and observe that the error in approximation is bounded by C/n for some constant C. This can be expressed using ”Big-O” notation:

$$\int_a^b f(x)\mathrm{d}x = h \sum_{k=1}^n f(x_k) + O(1/n).$$

In these notes we consider the ”Analysis” part of ”Numerical Analysis”: we want to *prove* the convergence rate of the approximation, including finding an explicit expression for the constant C.

To tackle this question we consider the error incurred on a single ”rectangle”, then sum up the errors on rectangles.

Now for a secret. There are only so many tools available in analysis (especially at this stage of your career), and one can make a safe bet that the right tool in any analysis proof is either (1) integration-by-parts, (2) geometric series or (3) Taylor series. In this case we use (1):

Lemma 1 (Rectangular Rule error on one panel) *Assuming f is differentiable we have*

$$\int_a^b f(x)\mathrm{d}x = (b-a)f(a) + \delta$$

where $|\delta| \le M(b-a)^2$ *for* $M = \sup_{a\le x\le b} |f'(x)|$.

Proof We write

$$\int_a^b f(x)\mathrm{d}x = \int_a^b (x-a)' f(x)\mathrm{d}x = [(x-a)f(x)]_a^b - \int_a^b (x-a)f'(x)\mathrm{d}x = (b-a)f(b) + \underbrace{\left(-\int_a^b (x-a)f'(x)\mathrm{d}x\right)}_{\delta}.$$

Recall that we can bound the absolute value of an integral by the sepremum of the integrand times the width of the integration interval:

$$|\int_a^b g(x)\mathrm{d}x| \le (b-a) \sup_{a\le x\le b} |g(x)|.$$

The lemma thus follows since

$$|\int_a^b (x-a)f'(x)\mathrm{d}x| \le (b-a) \sup_{a\le x\le b} |(x-a)f'(x)| \le M(b-a)^2.$$

■

Now summing up the errors in each panel gives us the error of using the Rectangular rule:

Theorem 1 (Rectangular Rule error) *Assuming f is differentiable we have*

$$\int_a^b f(x)\mathrm{d}x = h\sum_{k=1}^n f(x_k) + \delta$$

where $|\delta| \leq M(b-a)h$ for $M = \sup_{a\leq x\leq b}|f'(x)|$, $h = (b-a)/n$ and $x_k = a + kh$.

Proof We split the integral into a sum of smaller integrals and :

$$\int_a^b f(x)\mathrm{d}x = \sum_{k=1}^n \int_{x_{k-1}}^{x_k} f(x)\mathrm{d}x = \sum_{k=1}^n [(x_k - x_{k-1})f(x_k) + \delta_k] = h\sum_{k=1}^n f(x_k) + \underbrace{\sum_{k=1}^n \delta_k}_{\delta}$$

where δ_k, the error on each panel as in the preceding lemma, satisfies

$$|\delta_k| \leq (x_k - x_{k-1})^2 \sup_{x_{k-1}\leq x\leq x_k} |f'(x)| \leq Mh^2.$$

Thus using the triangular inequality we have

$$|\delta| = \left|\sum_{k=1}^n \delta_k\right| \leq \sum_{k=1}^n |\delta_k| \leq Mnh^2 = M(b-a)h.$$

■

Note a consequence of this lemma is that the approximation converges as $n \to 0$. In the labs and problem sheets we will consider the left-sided rule:

$$\int_a^b f(x)\mathrm{d}x \approx h\sum_{k=0}^{n-1} f(x_k)$$

and the *Trapezium rule*, where we approximate the integral in each panel by an affine function:

$$\int_{x_{k-1}}^{x_k} f(x)\mathrm{d}x \approx \int_{x_{k-1}}^{x_k} [f(x_{k-1}) + (x - x_{k-1})/(x_k - x_{k-1})f(x_k)]\,\mathrm{d}x = \frac{h}{2}\,[f(x_{k-1}) + f(x_k)]$$

leading to the approximation

$$\int_a^b f(x)\mathrm{d}x \approx \frac{h}{2}f(a) + h\sum_{k=1}^{n-1} f(x_k) + \frac{h}{2}f(b)$$

We shall see both experimentally and provably that this approximation converges faster than the rectangular rule.