Level 5 Computational Physics - Lecture 2

Dr. Simon Hanna

October 25, 2012

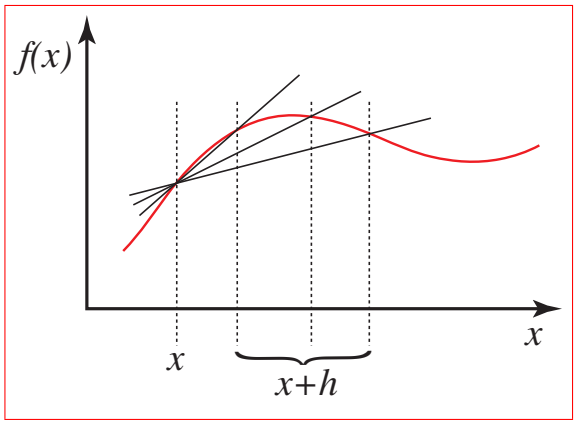
1 Introduction

In this lecture we consider the techniques of numerical differentiation and integration, and then begin examining techniques for solving differential equations.

2 Numerical Differentiation

- It is often the case that we wish to find the differential of a given function, either as a problem in its own right, or more usually as part of a more complex algorithm (e.g. minimisation of a function or solution of a differential equation).
- **Wherever possible, it is best to take the analytic derivative of the function** in question and use that directly within your code, as this improves efficiency and avoids numerical errors. However, this is not always possible, and numerical evaluation of differentials must sometimes be used.

The basic technique is illustrated by the definition of the derivative:



$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

For finite h, have the finite difference formula:

$$f'(x) = \frac{f(x+h)-f(x)}{h} \tag{1}$$

- For finite h this is an approximation;
- $f'(x)$ becomes more accurate as h gets smaller;
- BUT, if h is very small, we have floating point precision problems;
- Generally, h limited to $h = x\epsilon^{1/2}$ where ϵ is the relative precision of the floating-point format ($\epsilon \simeq 10^{-7}$ for single precision and $\epsilon \simeq 10^{-16}$ for double precision variables).

The size of error in Eq. (1) is found from the Taylor expansion:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \ldots$$

from which:

$$f'(x) = \frac{1}{h}\left[f(x+h) - f(x) - \frac{h^2}{2!}f''(x) + \ldots\right] \tag{2}$$

Comparing Eq. (1) and (2) we see the difference, or error, is:

$$E(h) = \frac{h}{2}f''(x) + \ldots = \mathcal{O}(h)$$

so Eq. (1) will never be very accurate.

Eq. (1) was an example of a 'forward difference' formula. A 'backward difference' formula could also be used, but with similar accuracy:

$$f'(x) = \frac{f(x)-f(x-h)}{h} + \mathcal{O}(h) \tag{3}$$

Consider the Taylor expansions for Eq. (1) and (3):

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \ldots \quad (4)$$

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2!}f''(x) - \frac{h^3}{3!}f'''(x) + \ldots \quad (5)$$

Subtracting Eq. (5) from (4) gives a 'central difference' formula:

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + E(h) \quad (6)$$

where the even-order derivatives cancel, to give an error of:

$$E(h) = -\frac{h^2}{3!}f'''(x) - \ldots \simeq \mathcal{O}(h^2)$$

Symmetric expressions are generally more accurate than non-symmatric ones, and Eq. (6) is recommended for accurate work.

2.1 Higher order derivatives

Adding Eq. (4) and (5) gives:

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + E(h) \quad (7)$$

with an error of:

$$E(h) = -\frac{h^2}{12}f^{iv}(x) - \ldots \simeq \mathcal{O}(h^2)$$

2.2 Higher order formulae

- Eq. (6) and (7) are examples of 'three-point' formulae (even though one of the points has zero contribution in Eq. (6)).
- Five point formulae may also be developed and are more accurate:

$$f'(x) = \frac{f(x-2h) - 8f(x-h) + 8(f+h) - f(x+2h)}{12h} + \mathcal{O}(h^4)$$

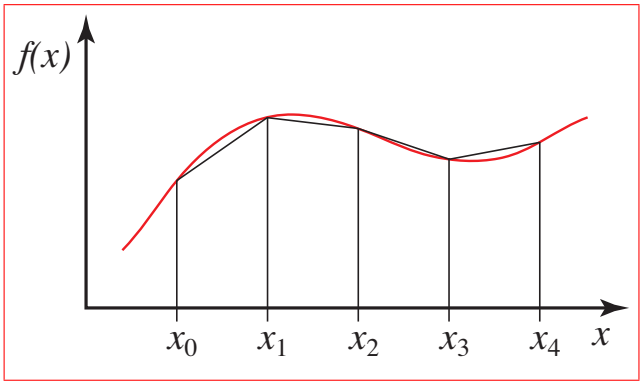
$$f''(x) = \frac{-f(x-2h) + 16f(x-h) - 30f(x) + 16(f+h) - f(x+2h)}{12h^2} + \mathcal{O}(h^4)$$

3 Numerical Integration

- The problem of numerical integration is a common one, because there are many well-behaved functions, e.g. $f(x) = e^{-x^2}$, which have *no* analytic integral.
- There are a range of techniques for numerical integration which trade complexity and computer time for accuracy, or which avoid evaluation of the function at inconvenient points (e.g. singularities).
- We will examine first the case where we can evaluate the function at a set of equally-spaced points in its domain (this is not always the case).

3.1 Trapezium Rule

The simplest numerical integration method is the two-point trapezium rule. We divide the integration range into subranges of width h , and approximating the area of each by a trapezium.



Summing to cover the entire integration range, gives:

$$\int_{x_0}^{x_N} f(x)dx \approx \frac{h}{2}(f_0+f_1) + \frac{h}{2}(f_1+f_2) + \ldots + \frac{h}{2}(f_{N-1}+f_N)$$

where $f_n = f(x_n)$. i.e.:

$$\int_{x_0}^{x_N} f(x)dx = h\left(\frac{f_0}{2} + f_1 + f_2 + \ldots + f_{N-1} + \frac{f_N}{2}\right) \tag{8}$$

(Strictly speaking this is the 'composite' trapezium rule).

The estimated error in the integral is:

$$E(h) = \frac{h^2}{12}\left(f_0' - f_N'\right) - \mathcal{O}(h^4)$$

In practice, it is normal to evaluate the integral repeatedly with decreasing values of h until you are satisfied that the calculation has converged to within the desired accuracy.

3.2 Simpson's Rule

The trapezium method is equivalent to approximating the integrand by a series of straight line segments, and integrating. A more accurate approach is to approximate the integrand by a series of quadratics, each involving three adjacent points. This results in the well-known composite Simpson's Rule:

$$\int_{x_0}^{x_N} f(x)dx \approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \ldots + 2f_{N-2} + 4f_{N-1} + f_N) + \mathcal{O}(h^5) \tag{9}$$

Things to watch out for:

- Trapezium rule and Simpson's rule work best when $f(x)$ is slowly varying over the region of integration;
- $f(x)$ must be defined over the whole region of integration i.e. no singularities;
- A function may be well behaved but difficult to evaluate e.g. $\sin(x)/x$ for $x \to 0$;
- Sometimes a change of variable helps e.g.

$$\int_0^a \frac{dx}{(1+x)\sqrt{x}} = \int_0^{2\sqrt{a}} \frac{1}{1+y^2/4}dy$$

  by putting $y = 2\sqrt{x}$;

- Other, more sophisticated, methods are available (Romberg integration, Gaussian quadrature) and may be found in textbooks.

3.3 Multidimensional Integration

- Numerical integration by finite differences is fine for low-dimensional functions, but can be inefficient for multi-dimensional problems, especially where the integration region is complex in shape, i.e. when the ranges of the 'inner' integrals depend on parameters.
- But, after all, this is what computers are for!
- Monte Carlo integration is an alternative, and will be covered later.

Example of a two dimensional integral:

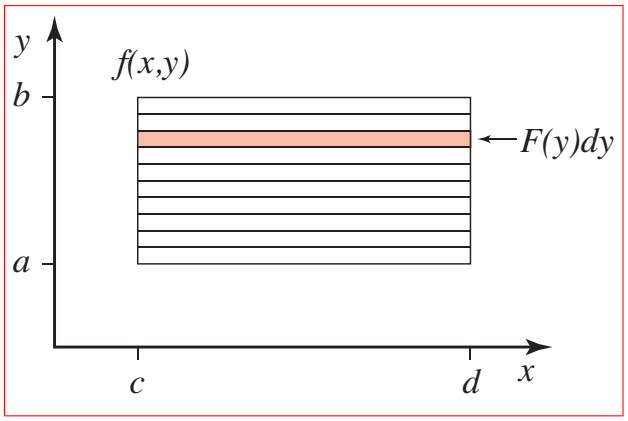
$$I = \int_a^b \int_c^d f(x,y)\,dx\,dy$$

If a, b, c and d are constants, we have a rectangular region of integration and we can rewrite the integral as:

$$I = \int_a^b F(y)\,dy$$

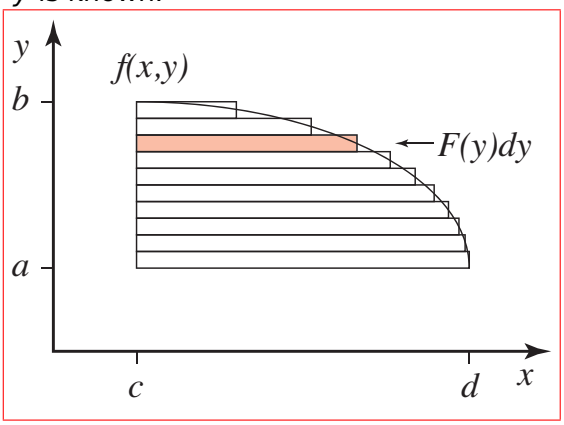
where:

$$F(y) = \int_c^d f(x,y)\,dx$$

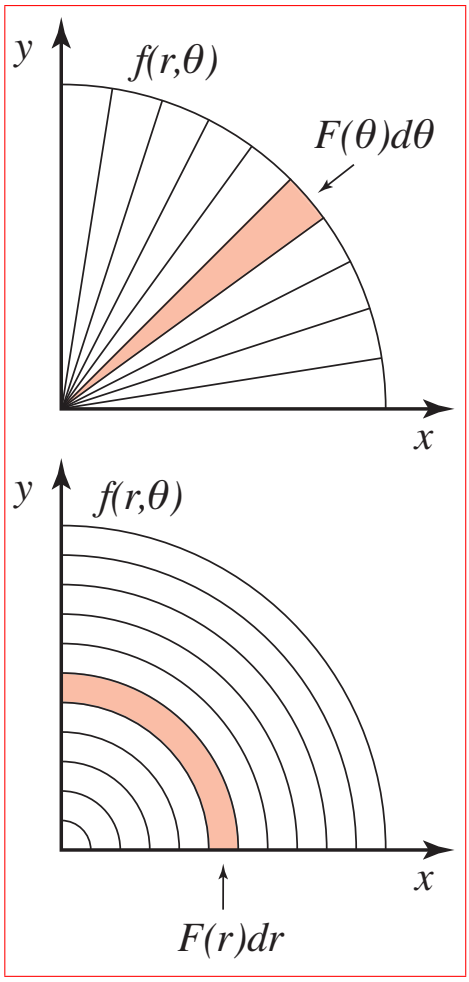


- The integration could be performed in various ways:
  - Perform x-integrals at fixed values of y (as shown) and store the values in an array. Then perform the y-integral.
  - Perform y-integrals at fixed x-values, store the results in an array, and then perform the x-integral.
  - Perform both integrals in a single summation using two nested loops.

- Irregular integration regions can be handled, as long as the dependence on x and y is known.



- Sometimes, a change of coordinate system might be helpful, e.g. to polar coordinates, but be aware that the sample points will be further apart at larger values of r:



3.4 Monte Carlo Integration

- Rather than evaluating the function at regular intervals, as in the trapezium and Simpson method, it is possible to evaluate the function at random points (hence the link with the casinos of Monte Carlo).
- The integral of the function will be given by the average value of the function, multiplied by the range i.e.:

$$\int_a^b f(x)dx \approx (b-a)\frac{1}{N}\sum_{i=1}^{N} f(x_i) \qquad (10)$$

- Clearly, as $N \to \infty$, the estimate of the average tends to the true value, and the integral becomes more correct.
- Convergence can be slow, with the accuracy only improving with $\sqrt{N}$.
- But, for integrals over more than 4 dimensions, the Monte Carlo convergence is actually better than the trapezium method.

- Monte Carlo integration is also useful for irregularly shaped integration regions: points are generated at random, and a simple check is used to decide whether to include each point in the average: