L2: Integration

1 Brief Introduction

Integrals in one or more dimensions appear in a large number of applications. In practice, they are often of such a nature that they cannot be solved analytically. This may be because it is very difficult to find such a solution (it may not even exist) or the integrand is given as discrete points, for example from measurement data. In those cases, numerical methods are used for calculating integrals.

- The principle of numerical integration is that the interval is divided into subintervals and on each such subinterval the integrand is approximated with a simpler function, for example a polynomial. The simpler function is chosen such that it can be integrated exactly with an analytic expression. The value of the integral is then approximated as the sum of the integrals of the simpler functions over all subintervals.

- There are many different methods of numerical integration. Two simple but common methods are the *Trapezoidal rule* and the *Simpson's rule*. With two demo programs, we will look at how these two methods work.

Start by downloading the files for Section 2 of this lab: `quadDemo.py` and `adaptDemo.py`. Sections 3-5 are suggested to be done in scripts that you create yourself.

Note! If you are using an IDE, you may need to set it up so that interactive shapes can be created. For example in Spyder: Preferences → IPython Console → Graphics → Backend → Automatic.

2 Trapezoidal Rule and Simpson's Rule

We will start by studying the integral

$$\int_{-\pi}^{2\pi} f(x)\,dx, \tag{1}$$

where the integrand is

$$f(x) = \cos(2x) + \sin(x). \tag{2}$$

1. Run the program `quadDemo`. The program draws the integrand and solves the integral using the trapezoidal rule and Simpson's rule. Use a small number of subintervals (n) to make it easier to study the figures, for example 10, 20, and 30. Note that the number of subintervals is linked to how accurate the solution will be. The trapezoidal rule is based on approximating the integrand as a straight line on each subinterval, thus forming trapezoids. The integral value is approximated as the total area of all trapezoids. Try to see it in the figure.

2. Comparing trapezoidal rule and Simpson's rule, which seems most accurate? Study the errors that are printed. One can theoretically show that an increase in the number of subintervals by a factor q leads to a decrease in the error by a factor q^p, where p is a positive integer called the *order of accuracy* of the method. Investigate numerically what the order of accuracy is for the trapezoidal rule and Simpson's

rule. For example, compare the errors with 40 and 80 subintervals respectively, i.e., with $q = 2$.

In quadDemo you had to specify how many subintervals the area should be divided into, i.e., which discretization would be used. Often, when using ready-made programs for numerical integration, you do not need to specify the number of intervals. Instead, the functions themselves determine intervals such that the accuracy of the solution meets a certain tolerance; such methods are called adaptive. You will now look at how Simpson's rule works when the intervals are chosen adaptively.

3. Run the program adaptDemo, initially use the tolerance 10^{-5} (can be written as 1e-5 or 0.00001). Note that here you cannot vary the number of intervals, instead you vary the accuracy (tolerance). Experiment with a few different choices of accuracy and see what happens. Be aware that the length of the intervals varies depending on how the integrand looks. Is it possible to see any principle for how? When is the interval shorter and when is it longer?

To think about: How can a numerical method know that the calculation meets a certain accuracy without knowing the exact integral's value? The method must estimate the error without knowing the exact solution. How this works is explained in the lecture.

3 SciPy functions for integration

When calculating integrals in Python, you can use functions from *SciPy*. This part of the lab is an introduction to two of the most common integral functions in SciPy: quad and trapezoid.

When the integrand is given as a continuous function, the SciPy function quad is used. It is called as follows:

```
from scipy.integrate import quad
(I, err) = quad(f,xl,xr)
```

Here xl and xr are the limits of integration and f is a Python function that defines the integrand. The integral value is stored in the variable I, and the estimate of the error is stored in the variable err. For additional information and examples of how quad is used, see the SciPy documentation: https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.quad.html.

4. Calculate the integral (1) again, this time using quad. Start by defining the integrand (2) as a Python function, the structure of the function should be:

```
def f(x):
  return ...
```

You can use functions and constants from math or numpy. Then solve the integral by calling quad.

5. quad can also be used when one or both of the integration limits go to infinity, then inf from math or numpy is used as an argument. Calculate the integral

$$\int_1^\infty \frac{4}{x^2+1}\,dx \tag{3}$$

using quad. Does the result seem reasonable?

Often the integrand is not available as a continuous function but as discrete points, for example measured values. Then you cannot use `quad`; an alternative is the SciPy function `trapezoid`. As the name suggests, that function uses the trapezoidal rule.

6. Calculate the integral of $S(t)$ on the interval $[0, 0.8]$, where

t	0	0.2	0.4	0.6	0.8
$S(t)$	0.2	0.368	0.381	0.228	0.049

Start by creating two vectors, `t` and `S`, which store the data. The vectors can be stored as lists or NumPy vectors. Plot the values to understand what to integrate. Finally, calculate the integral with the call:

```
from scipy.integrate import trapezoid
I = trapezoid(S,t)
```

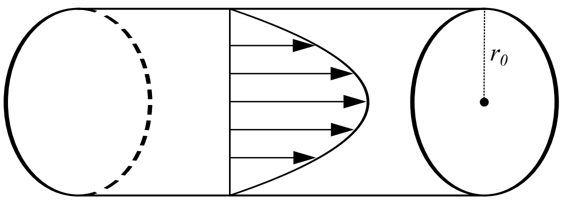
Does the answer seem reasonable? Study the graph. Note that the accuracy is determined by the number of discrete points, in this case the number of measurements. It is therefore not possible to get better accuracy with the same method without getting more measurement values, and thus there is no room for adaptivity.

4 Calculation of water flow through a pipe

Given that the speed of a flowing liquid in a pipe is known, one can calculate the amount of water, i.e. the volume that flows through the pipe per unit of time. This is conveniently done using numerical integration. The total volume of liquid, I, flowing through a pipe of radius r_0 per unit time can be described as

$$I = 2\pi \int_0^{r_0} r v(r)\, dr, \quad \text{where} \quad v(r) = 2(1 - \frac{r}{r_0})^{1/6},$$

$v(r)$ is the velocity distribution and r is the distance from the center points, see Figure 1.



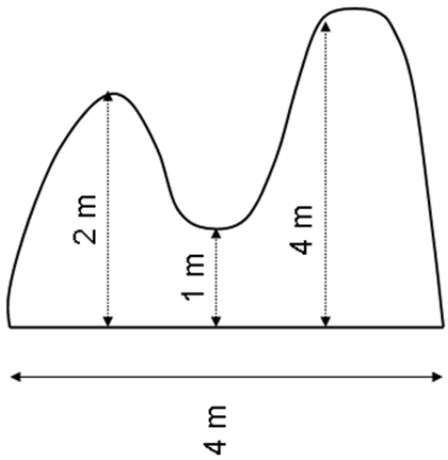
Figur 1: Velocity distribution in cylindrical pipe with radius r_0.

7. Write a function that defines the integrand. Verify that the function works by calling it for a few different choices of the input argument.

8. Calculate the integral using SciPy's function `quad`, for example use $r_0 = 3$. Note that r_0 appears both in the integrand function and as the limit of integration. Does the answer seem reasonable? Since the amount of liquid is proportional to the area of the tube, an increase in the radius r_0 by a factor p should lead to an increase in the volume by a factor p^2. Check if this is correct.

5 Calculation of the amount of soil in a flower bed

You need to cover a flower bed with a 5 cm thick layer of new topsoil. The topsoil is sold in bags of 0.05 m^3 (50 liters) and to buy the right number of bags you need to calculate the volume. Unfortunately, soil is irregularly distributed because it is next to a rock, see Figure 2.



Figur 2: The distribution of topsoil from above.

The topsoil volume is given by the integral

$$I = 0.05 \int_0^4 f(x)\, dx, \tag{4}$$

where $f(x)$ is the width of the flower bed at position x along the edge. Here there is no formula for the integrand $f(x)$, but instead you have measured the width of the flower bed at different points on the edge:

x (meters)	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
$f(x)$ (meters)	0.0	1.2	2.0	1.4	1.0	1.5	4.0	4.1	0.0

9. Calculate the volume of topsoil and the number of bags needed using the above measurements. Since the integrand $f(x)$ is not given in the form of a continuous function, but as measurement points, `trapezoid` is used.