NumericalIntegrationWithPython

October 30, 2019

0.0.1 Solving integrals numerically

Introduction Many observables in nature can be predicted with the solution an integral of a certain function. In this section I will make a short excursion on how to solve integrals numerically.

From my experience, many students and researches excellently trained to define the set of equations describing their problems mathematically. Seconds, the evaluation of the individual equations with multiple variables imposes (in most cases) no problem. Third, it is also part of the common knowledge, that the integral of any function is equivalent to the area below the curve. A college of mine once told me, that she learn calculating the integral in school by: 1. drawing the function for multiple points on a paper 2. combine them with a line 3. count the squares below the curve

Nothing else is done, when solving integrals numerically.

On the other side solving an integral analytically requires in many cases advanced mathematical skills and often approximation/simplifications are introduces to be able to solve the problem. Those tasks are often hard to master for many people (I continuously fail at solving integral) and the simplifications often reduce the solution to specific boundary conditions (as for example the Boltzmann approximation).

As mentioned solving integrals numerically is fairly easy, even if one might not feel very comfortable with counting squares. But even if counting is not an option, there are modern tools to solve this task very efficiently! If haven't been introduced yet, here they come!

So if the elements to be integrated can evaluated for each point between the boundaries of the integral, not much is in the way to solve the integral numerically. Here a simple example of solving: $\int_3^5 x^5 dx$

The analytical solution solution is: $[1/6x\ ]_3^5 = 1/6 * 5\ \ - 1/6 * 3\ \backsimeq 14896.17$

The `quad` function The `quad` function from the `scipy.integrate` package will be used to integrate the given function. The help file of `quad` says: >Integrate func from `a` to `b` (possibly infinite interval) using a technique from the Fortran library QUADPACK.

This description reveals, that the Fortran library QUADPACK is used in the background. So nothing new is shown here from the "scientific" point of view. I'd rather like to point out, how easy this can be applied in a Jupyter notebook. From discussion with colleagues I know, that the biggest challenge is how to technically implement the numerical solving algorithm in Python. So here it comes:

[37]:
```
%pylab
```

```
Using matplotlib backend: Qt5Agg
Populating the interactive namespace from numpy and matplotlib
```

[38]:
```
from scipy.integrate import quad
num_sol, num_error = quad(lambda x:x**5,3,5)
ana_sol = 1/6*(5**6-3**6)

print(f'Numerical solution: {num_sol:.2f} +- {num_error:.2f}')
print(f'Analytical solution: {ana_sol:.2f}')
```

```
Numerical solution: 2482.67 +- 0.00
Analytical solution: 2482.67
```

In this case, the numerical and the analytical solution result in the same results. How about a more complex problem? Let's look at the "Normal distribution":

$$f(x) = \frac{1}{\sqrt{2\Pi}} * e^{-\frac{x^2}{2}} \tag{1}$$

Since the probability distribution is normalized the integral from $-\infty$ to ∞ is 1:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{2}$$

The analytical solution of this integral is already a rather advanced task, but still doable. The numerical results are obtained in the following cell.

[39]:
```
def f(x):
    return 1/(2*pi)**0.5*e**(-(x**2/2))

num_sol, num_error = quad(f, -np.inf,np.inf)
ana_sol = 1

num_sol, num_error = quad(f, -inf,inf)
ana_sol = 1

print(f'Numerical solution: {num_sol:.12f} +- {num_error:.12f}')
print(f'Analytical solution: {ana_sol:.12f}')
```

```
Numerical solution: 1.000000000000 +- 0.000000010178
Analytical solution: 1.000000000000
```

Also here a numerical solution is in line with the expected result from the analytic solution.

There two examples cover functions, where a analytical solution is known. In the following parts of this thesis, many analytical solutions are not known. In this case the shortcut of using a numerical solution instead of relaying on the exact solution is reasonable.

Side-Note: The `quad` function does not only give back the numerical solution, but also an estimate of the absolute error in the result.

The `print()` statement is used to add the results in the output of the notebook. The `print` command requires a text *string* between its parenthesis. In Python a *string* consists of multiple characters between quotation marks: e.g. `'L3TT3R5'`. Additionally another rather new feature of Python is used here. This feature is called *formatted strings*. *Formatted strings* are constructed with an `f` in front of the *string*: `f'L3TT3R5'`.

For formatted strings variables inside curly parenthesis are then replaced with their string representation. The formatting of the representation may be given after `:`. For example `.12f` tells the formatter to represent the variable as a float with 12 digits after the decimal separator. When reading this as an interactive notebook, feels free to modify the formatting statement and check the result.