1. Basic idea of integration

Integration of functions is an important calculation in computational physics, both as a fundamental task and as a component of larger problems. Many integrals do not have closed forms and require numerical computation.

What is the definition of an integral?

An integral is defined by the limit:

$$\int_a^b \mathrm{d}x f(x) = \lim_{\mathrm{d}x\to 0}\left[\mathrm{d}x \sum_{i=1}^{(b-a)/\mathrm{d}x} f(x_i)\right] \tag{1}$$

where x_i are spaced between a and b with separations $\mathrm{d}x$.

What is a simple numerical estimate of an integral?

Just to perform this sum with some finite $\mathrm{d}x$:

$$\int_a^b \mathrm{d}x f(x) = \left[\mathrm{d}x \sum_{i=1}^{(b-a)/\mathrm{d}x} f(x_i)\right] \tag{2}$$

where x_i are spaced between a and b with separations $\mathrm{d}x$.

This is just a particular case of the more general form that most integration methods take, which is that it can be approximated as some linear combination of evaluations of the function:

$$\int_a^b \mathrm{d}x f(x) = \sum_{i=1}^{N} f(x_i) w_i \tag{3}$$

2. Trapezoid rule

The simple estimate above can be thought of as approximating the function as piecewise constant. Obviously there are better approximations that can be made! Better algorithms for integration generally boil down to better models of the function. In this respect, integration is closely allied to interpolation of functions.

The trapezoid rule is the result of integrating a linear interpolation of the function. Each term in the integral will become:

$$\frac{1}{2}\mathrm{d}x\left(f_i + f_{i+1}\right) \tag{4}$$

The next term is:

$$\frac{1}{2}\mathrm{d}x\left(f_{i+1} + f_{i+2}\right) \tag{5}$$

For equally spaced points, then $w_i = \mathrm{d}x$, except for $w_1 = w_N = \mathrm{d}x/2$.

For what sort of function is the trapezoid rule exactly correct?

For a linear function. Of course, this property is not very useful!!

3. Simpson's rule

Simpson's rule represents the next level of sophistication in interpolation. Here, the function is approximated locally around the points $i-1$, i, $i+1$, as a quadratic:

$$f(x) = \alpha' + \beta' x + \gamma' x^2 \tag{6}$$

This is not a very convenient form. Let us instead use:

$$f(x) = \alpha + \beta\left(\frac{x - x_i}{\mathrm{d}x}\right) + \gamma\left(\frac{x - x_i}{\mathrm{d}x}\right)^2 = \alpha + \beta y + \gamma y^2 \tag{7}$$

with a change of variable to $y = (x - x_i)/\mathrm{d}x$. For a set of three points, $i-1$, i, and $i+1$, you can fit the parabola using the fact:

$$\begin{aligned} f_{i-1} &= \alpha - \beta + \gamma \\ f_i &= \alpha \\ f_{i+1} &= \alpha + \beta + \gamma \end{aligned} \tag{8}$$

This can be easily solved:

$$\begin{aligned} \alpha &= f_i \\ \gamma &= \frac{f_{i+1} + f_{i-1}}{2} - f_i \\ \beta &= \frac{f_{i+1} - f_{i-1}}{2} \end{aligned} \tag{9}$$

What is the integral over the region defined by these three points?

The integral over the region defined by these three points :

$$\begin{aligned} \int_{x_{i-1}}^{x_{i+1}} \mathrm{d}x f(x) &= \mathrm{d}x \int_{-1}^{1} \mathrm{d}y \left(\alpha + \beta y + \gamma y^2\right) \\ &= \mathrm{d}x \left[\alpha y + \frac{\beta}{2} y^2 + \frac{\gamma}{3} y^3\right]_{-1}^{1} \\ &= \mathrm{d}x \left[2\alpha + \frac{2\gamma}{3}\right] \end{aligned} \tag{10}$$

Plugging in α and γ:

$$\int_{x_{i-1}}^{x_{i+1}} \mathrm{d}x f(x) = \mathrm{d}x \left(2f_i + \frac{f_{i+1} + f_{i-1}}{3} - \frac{2}{3} f_i\right) = \mathrm{d}x \left(\frac{1}{3} f_{i-1} + \frac{4}{3} f_i + \frac{1}{3} f_{i+1}\right) \tag{11}$$

Simpson's rule comes from using this approximation across the length from a to b, by dividing the interval into an even number of segments, and integrating each separately. This yields a full summation:

$$\int_a^b \mathrm{d}x f(x) = \sum_{i=1}^{N} = \mathrm{d}x \left[\frac{1}{3}f_1 + \frac{4}{3}f_2 + \frac{2}{3}f_3 + \frac{4}{3}f_4 + \ldots + \frac{2}{3}f_{N-2} + \frac{4}{3}f_{N-1} + \frac{1}{3}f_N\right] \tag{12}$$

Because this is applied to two segments at a time, it requires an even number of segments, which means N must be odd.

The weights for the three points used in in Simpson's rule are set to exactly integral a quadratic function — a second degree polynomial. What must N be to exactly integrate an M-degree polynomial?

$N = M + 1$. We can show this as follows. Each of the N points x_k yields a linear equality:

$$f_k = \sum_{j=0}^{M} \alpha_j x^j \tag{13}$$

that can be used to determine the coefficients of the function, and thus its integral. So this yields a system of N linear equations, with $M + 1$ unknowns. So to guarantee a solution, you need $N = M + 1$.

These methods are good methods, but it turns out we can be even cleverer. But before we do so, we have a little bit of work to do.

4. Rescaling of integrals

It may appear trivial, but just as in differentiation, there are rescaling of integrals that can be performed for various reasons of convenience or otherwise.

The simplest rescaling is linear, which just rescales the limits of the integral:

$$I = \int_a^b \mathrm{d}x f(x) = \frac{b-a}{b'-a'} \int_{a'}^{b'} \mathrm{d}x' f(x(x')) \tag{14}$$

which simply follows from the tranformation:

$$\begin{aligned} x' &= (x-a)\left(\frac{b'-a'}{b-a}\right) + a' \\ \mathrm{d}x' &= \mathrm{d}x\left(\frac{b'-a'}{b-a}\right) \end{aligned} \tag{15}$$

or:

$$\begin{aligned} x &= (x' - a')\left(\frac{b-a}{b'-a'}\right) + a' \\ \mathrm{d}x &= \mathrm{d}x'\left(\frac{b-a}{b'-a'}\right) \end{aligned} \tag{16}$$

This is a pretty trival rescaling, but it can be useful if you can rescale an integral to a previously calculated integral. We will use this below in the specific case: $a' = -1$, $b' = 1$:

$$I = \int_a^b \mathrm{d}x f(x) = \frac{b-a}{2}\int_{-1}^1 \mathrm{d}x' f(x(x')) \tag{17}$$

This will allow us to develop some useful algorithms for the specific range -1 to 1, which can then be generalized to any finite range.

If we want to alter $[-1, 1]$ to an infinite range that is possible too. For example:

$$\begin{aligned} x &= a\frac{1+x'}{1-x'} \\ dx &= a\left(\frac{1}{1-x'} + \frac{1+x'}{(1-x')^2}\right) \\ &= a\mathrm{d}x'\frac{1-x'+1+x'}{(1-x')^2} \\ &= \mathrm{d}x'\frac{2a}{(1-x')^2} \end{aligned} \tag{18}$$

which lets us rewrite an infinite range:

$$I = \int_0^\infty \mathrm{d}x f(x) = \int_{-1}^1 \mathrm{d}x' \frac{2ax'}{(1-x')^2} f(x(x')) \tag{19}$$

In this case, a is a choice to be made, and $x = a$ when $x' = 0$. So there are better choices for a than others – you want it to be somewhere near where the integral is expected to reach about half its total.

The other forms of weighting are given in the book, and may be derived similarly.

5. Gaussian quadrature

Now we have all the tools to derive one of the workhorse algorithms for integrating function, which is Gaussian quadrature. Gaussian quadrature has the advantage that it yields a systematic way to write an algorithm for integration which utilizes N points, that is *exact* for any polynomial of order $2N - 1$ or less. Note that this is much better than we found before, the path we were on for Simpson's rule, which utilized $N+1$ points to exactly integrate a polynomial of N points. It turns out that the improvement is gained by choosing the points carefully.

We will show how to do this for the integral:

$$\int_{-1}^{1} \mathrm{d}x f(x) \tag{20}$$

where $f(x)$ is a $2N-1$ degree polynomial (or less). Clearly we can rescale the limits as necessary above for the problem at hand.

We are seeking an exact formula for the integral of this function which is:

$$\int_{-1}^{1} \mathrm{d}x f(x) = \sum_{i=1}^{N} w_i f(x_i) \tag{21}$$

The derivation of this is neat. Note that the derivation in the book is extremely confusing and contains at least one error.

We will use the Legendre polynomials to aid us. In fact, it will be the roots of the Legendre polynomials (where they are zero) that turn out to be the locations of the integration points.

What are the Legendre Polynomials? Where have you seen them before.

They usually arise in the physics curriculum because they are the solutions to Laplace's Equation ($\nabla^2\Phi = 0$) under cylindrical symmetry. They also have some interesting properties. We will refer to them here as $P_n(x)$, where $-1 < x < 1$, and n is the order of the Legendre Polynomial. They have the property that each Legendre Polynomial is a polynomial of order n:

$$P_n(x) = \sum_{i=0}^{n} a_n x^n \tag{22}$$

Specifically, they are:

$$P_n = \frac{1}{2^n n!} \frac{\mathrm{d}^n (x^2-1)^n}{\mathrm{d}x^n} \tag{23}$$

Or:

$$\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= x \\
P_2(x) &= \frac{1}{2}(3x^2-1) \\
P_3(x) &= \frac{1}{2}(5x^3-3x) \\
P_4(x) &= \frac{1}{8}(35^4-30x^2+3) \\
&\ldots
\end{aligned} \tag{24}$$

They form a complete set in function space (any function can be expressed as a sum of a sufficient number of Legendre Polynomials). A related property is that any polynomial of order n can be expressed as a sum of Legendre Polynomials with orders $\leq n$.

All the Legendre Polynomials are orthogonal to each other. What does that mean?

It means that their dot products are zero. Functions live in a linear vector space. E.g. it is infinite-dimensional, and one set of basis functions are Dirac δ-functions. You can define a dot product in that space as:

$$q(x) \cdot r(x) = \int_{-1}^{1} \mathrm{d}x q(x) r(x) \tag{25}$$

As we will see below, this choice is not unique!

In any case, it means that the statement that Legendre Polynomials are orthogonal means the following:

$$\int_{-1}^{1} \mathrm{d}x P_n(x) P_m(x) = \delta_{nm} \frac{2}{2n+1} \tag{26}$$

If all the Legendre Polynomials are orthogonal to each other, and any polynomial of order n can be expressed as a sum of Legendre Polynomials with order $\leq n$, then what is this integral:

$$\int_{-1}^{1} \mathrm{d}x P_{n+1}(x) x^m \tag{27}$$

0

We start by noting that $f(x)$ can be in general factored in the following way:

$$f(x) = q(x) P_N(x) + r(x) \tag{28}$$

where we choose $q(x)$ to be an $N-1$-degree polynomial. The first term is therefore a polynomial of order $2N-1$, or less. Since we can choose the coefficients of the $q(x)$ polynomial to be whatever we want, we can always match the coefficients of all the polynomial terms of order N or greater in $f(x)$. This leaves a remainder $r(x)$ which is an $N-1$-degree polynomial (or less).

The integral:

$$\int_{-1}^{1} \mathrm{d}x q(x) P_N(x) = 0 \tag{29}$$

because the Legendre polynomials are always orthogonal to lower-order polynomials.

So:

$$\int_{-1}^{1} \mathrm{d}x f(x) = \int_{-1}^{1} \mathrm{d}x r(x) \tag{30}$$

Since $r(x)$ is an $N-1$-degree polynomial or less, there is a way to integrate the function with N points or less, as we found above.

The locations of these points can be found as follows. The integral is now known to be writable as:

$$\int_{-1}^{1} \mathrm{d}x f(x) = \sum_{i=1}^{N} w_i f(x_i) = \sum_{i+1}^{N} w_i q(x_i) P_N(x_i) + \sum_{i+1}^{N} w_i r(x_i) \tag{31}$$

If we choose the points x_i to be the N roots of the Legendre polynomial of order N, then:

$$\int_{-1}^{1} \mathrm{d}x f(x) = \sum_{i=1}^{N} w_i r(x_i) \tag{32}$$

where:

$$r(x) = \sum_{j=0}^{N-1} \alpha_j x_j \tag{33}$$

This is great, and we also know how to set the w_i. These are derived in the same way as for Simpson's rule. Using the appropriate linear set of equations analogous to Equation 13, you can express the solution for the coefficients α_j in terms of values of the function $f(x_i)$, which by design is equal to $r(x_i)$ at the chosen points, and then given the closed form of the integral of the polynomial, this translates into a form for w_i.

The notebook shows an implementation given the weights and locations determined for $N = 4$, and demonstrates performance up to $2N - 1$.

Rather than determine weights and locations yourself to higher order, the SciPy routines in its `integrate` module already contain this information. In particular, `fixed_quad` performs the fixed-order Gaussian quadrature that we find here.

6. Generalizations of Gaussian quadrature

The Gaussian quadrature method is good for smooth functions. However, if it is not smooth, or in particular has singularities, it will be an issue. In addition, there turns out to be plenty of scope in generalizing the method to handle certain non-polynomial functions exactly.

Specifically, it turns out that you can generally find *exact* expressions for integrals of the following form:

$$\int_{-1}^{1} \mathrm{d}x W(x) f(x) \tag{34}$$

where $W(x)$ is a known function and $f(x)$ is a polynomial.

The proof involves redefining the dot product between two functions $q(x)$ and $r(x)$:

$$q(x) \cdot r(x) = \int_{-1}^{1} \mathrm{d}x W(x) q(x) r(x) \tag{35}$$

Then it turns out we can find a complete basis set of polynomals that are orthogonal under this definition. Legendre Polynomials are just one case of this, for $W(x) = 1$. The locations are determined by the roots of these new polynomials, and the weights are determined in an analogous manner.

As an example, look at the problem:

$$\int_{-1}^{1} \mathrm{d}x \frac{1}{\sqrt{1-x^2}} = \pi \tag{36}$$

If we use regular Gaussian quadrature, our answers are very bad.

But we can define $W(x) = 1/\sqrt{1-x^2}$ and $f(x) = 1$. This is called *Gauss-Chebyshev* quadrature. SciPy doesn't have this directly, but it does have a routine that gives you the weights. This works very well for integrating over this singularity.

Another generally useful form has $W(x) = \exp(-x^2)$. This yields Gauss-Hermite polynomials, and Gauss-Hermite quadrature. A slightly altered Gaussian is a common form for a number of real-world distributions.

7. A physical example: nuclear reaction rates

One example of a process that involves an integral that needs to be estimated numerically is that of nuclear reactions in stars.

Nuclear fusion reactions in stars are driven by the following process. The center of the star consists of very hot ionized gas. The nuclei of the atoms in the gas have high enough energies that two of them can get close enough to one another to tunnel through their Coulomb repulsion into their energetically preferred bound state (or to otherwise interact). The rates of these reactions are driven by the number densities of the nuclear species, their temperatures, and their Coulomb charges.

Specifically, the cross-section for the reactions can be written:

$$\sigma(E) = \frac{S(E)}{E} \exp\left[-(E_c/E)^{1/2}\right] \tag{37}$$

where we will set the slowly varying function $S(E) = 1$ for simplicity and:

$$E_c = \frac{2\pi^2 Z_1^2 Z_2^2 e^4 \mu}{\hbar^2} \tag{38}$$

Then the reaction rate per unit volume, which ultimately determines in the case of our Sun how much energy is produced, and thus how brightly it shines, is:

$$r_{12} = \frac{n_1 n_2}{(\pi\mu)^{1/2}} \left(\frac{2}{kT}\right)^{3/2} \int_0^\infty \mathrm{d}E S(E) \exp\left[-(E_c/E)^{1/2}\right] \exp\left(-E/kT\right) \tag{39}$$

Basically, this is an integral over all of the relative kinetic energies the particles can have, weighted by the cross section of interaction. This governs the rates for any two species, but of course in reality the Sun's core contains many different simultaneous reactions. We will just calculate a small part of this problem.

We may be interested in a very important dependence, which is the dependence of this reaction rate on temperature. This dependence is part of what sets the overall structure of stars.

For two protons, the value:

$$E_c = \frac{2\pi^2 e^4 m_p}{2\hbar^2} \approx 7.9 \times 10^{-14} \quad \mathrm{kg\ m^2\ s^{-2}} \tag{40}$$

The temperatures are of order 10^7 K, so:

$$\begin{aligned} kT &\sim 1.3806 \times 10^{-23} \times 10^7 \quad \mathrm{kg\ m^2\ s^{-2}} \\ &\sim 1.3806 \times 10^{-16} \quad \mathrm{kg\ m^2\ s^{-2}} \end{aligned} \tag{41}$$

So typically $E_c/kT \sim 1000$.

For convenience and to keep us as safe as possible from underflows and overflows, we want to integrate over variables that are closer to unity, not $10-15$. Also, if we look at the integrand it is clear that E_c and kT do not matter individually (except as overall scale factors), just their ratio; we can avoid doing some integrals if we cast results just in terms of the ratio. So it makes sense to perform the transformation:

$$\begin{aligned} x &= E/kT \\ \mathrm{d}x &= \mathrm{d}E/kT \end{aligned} \tag{42}$$

and then define $R = E_c/kT$, which then gives us:

$$\begin{aligned} r_{12} &= \frac{2n_1 n_2}{(\pi\mu)^{1/2}} \left(\frac{2}{kT}\right)^{1/2} \int_0^\infty \mathrm{d}x S(E(x)) \exp\left[-(R/x)^{1/2}\right] \exp(-x) \\ &= \frac{2n_1 n_2}{(\pi\mu)^{1/2}} \left(\frac{2}{kT}\right)^{1/2} I(R) \end{aligned} \tag{43}$$

where

$$I(R) = \int_0^\infty \mathrm{d}x S(E(x)) \exp\left[-(R/x)^{1/2}\right] \exp(-x) \tag{44}$$

Now, the integration involves just R, and the rest of the scaling of the reaction rate is just multiplication.

The integral has the form:

$$I(R) = \int_0^\infty \mathrm{d}x f(x;R) \exp(-x) \tag{45}$$

where

$$f(x;R) = \exp\left[-(R/x)^{1/2}\right] \tag{46}$$

This is amenable to a form of Gaussian quadrature, specifically Gauss-Laguerre quadrature. This means:

$$I(R) \approx \sum_{i=1}^{N} w_i f(x_i;R) \tag{47}$$

If we want to know how r_{12} scales with temperature:

$$r_{12} \propto R^{1/2} I(R) \tag{48}$$

8. Monte Carlo

9. Multidimensional integrals