Numerical Integration I

We know about numerical differentiation. Now we look at numerical integration. It is performed by approximating the function to integrate by interpolation. The degree k of the polynom chosen for the interpolation defines the following methods:

- Trapezoidal Rule ($k = 1$)
- Simpson's Rule ($k = 2$)
- Other: Romberg's method

Numerical Integration II

Other method:

- Simple Monte Carlo Integration (for multidimensional integral)

Numerical recipes- The Art of Scientific Computing , by William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, second edition available online `http://www.nrbook.com/a/bookcpdf.php`. (chp 4 & 7)

Numerical Integration I

Remember integrating f means finding F such that:

$$\int_a^b f(x)dx = F(x)|_a^b = F(b) - F(a)$$

Two possible reasons for integrating numerically are that

- the function F is expensive to compute
- the function F has no analytic form (e.g. $f = e^{-x^2}$).

In practise, we divide the interval $b - a$ by n to get slots of width h

$$h = \frac{b-a}{n} \qquad x_i = a + ih$$

We let f_i denote $f(x_i)$, and set $x_0 = a$ and $x_n = b$.

Numerical Integration II

In general we evaluate integrals piecewise, by taking k slots at a time ($k+1$ data points) and using an interpolating polynomial P_k of degree k to approximate f over that interval. We can then integrate the polynomial to get I_k, and evaluate this at the end-points to get an approximation to the integral. So for $\{x_0, \cdot, x_k\}$, the polynom P_k is interpolated:

$$\begin{aligned} P_k &= a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k \\ I_k &= a_0 x + \frac{a_1 x^2}{2} + \frac{a_2 x^3}{3} + \cdots + \frac{a_k x^{k+1}}{k+1} \\ \int_{x_0}^{x_k} P_k(x)dx &= I_k(x_0) - I_k(x_k) \end{aligned}$$

Having done this for $x_0..x_k$, we then repeat for $x_k..x_{2k}$, and then for $x_{2k}..x_{3k}$, an so on, until we reach x_n. The approximation of the integral of f in between a and b is then the sum of those approximations.

Best Choice of k for Integration I

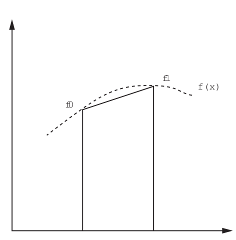
As with polynomial interpolation, we find that polynomials of high degree do not give better accuracy. Indeed, consider the function

$$f(x) = \frac{1}{1 + 25x^2}$$

we find that the error when integrating increases as k gets larger.
In practise we restrict k to 1 or 2.

Trapezoidal Rule ($k = 1$) I

We have one slot, with two data points, and we approximate the integral by the trapezoidal shape formed by those two points:



Trapezoidal Rule ($k = 1$) II

The area under the straight line is the area of the box of height f_0 plus that if the triangle on top of height $x_1 - x_0$, both of width h:

$$\begin{aligned} \text{area} &= h \cdot f_0 + \frac{1}{2} h \cdot (f_1 - f_0) \\ &= h f_0 + \frac{1}{2} h f_1 - \frac{1}{2} h f_0 \\ &= \frac{h}{2}(f_0 + f_1) \end{aligned}$$

So the trapezoidal rule states that

$$\int_{x_0}^{x_1} f(x) dx = \frac{h}{2}(f_1 + f_0) + R_T$$

where R_T is the truncation error:

$$R_T = \int_{x_0}^{x_k} (x - x_0)(x - x_1) \cdots (x - x_k) \frac{f^{(k+1)}(\xi)}{n+1!} dx$$

Trapezoidal Rule ($k = 1$) III

For the trapezoidal rule ($k = 1$) this gives us:

$$R_T = \int_{x_0}^{x_1} (x - x_0)(x - x_1) \frac{f''(\xi)}{2!} dx$$

By change of variable $x = x_0 + ph$, noting that $x_1 = x_0 + h$

$$\begin{aligned} \int_{x_0}^{x_1} (x - x_0)(x - x_1) \tfrac{f''(\xi)}{2!} dx &= h \int_0^1 ph(p-1)h \tfrac{f''(\xi)}{2} dp \\ &= \tfrac{h^3 f''(\xi)}{2} \int_0^1 p(p-1) dp \\ &= \tfrac{h^3 f''(\xi)}{2} \int_0^1 (p^2 - p) dp = \tfrac{h^3 f''(\eta)}{2} \left(\tfrac{p^3}{3} - \tfrac{p^2}{2}\right)\Big|_0^1 \\ &= \tfrac{h^3 f''(\xi)}{2} \left(\tfrac{1}{3} - \tfrac{1}{2}\right) = \tfrac{h^3 f''(\eta)}{12} \end{aligned}$$

So we see that the truncation error per trapezoidal step is $O(h^3)$.

Trapezoidal Rule ($k = 1$) IV

Trapezoidal rule.

We can apply the trapezoidal rule over the entire range $x_0..x_n$ in one go:

$$\int_{x_0}^{x_n} f(x) dx = h\left(\frac{1}{2} f_0 + f_1 + f_2 + \cdots + f_{n-1} + \frac{1}{2} f_n\right)$$

In fact given $n + 1$ data points, it is much easier to integrate the function they represent than to interpolate to find that function !

Simpson's Rule ($k = 2$) I

Simpson's Rule.

Simpson's rule uses a parabola (polynomial of degree 2) to fit three points. The resulting formula obtained on $\{x_0, x_1, x_2\}$ is:

$$\int_{x_0}^{x_2} f(x) dx = \frac{h}{3}(f_0 + 4f_1 + f_2) + R_T$$

Romberg's Method I

Romberg's Method is based on combining the previous methods with Richardson Extrapolation to improve the error bounds. It has the nice property that the calculation rounding error R_{XF} for the integral

$$\int_a^b f(x) dx$$

is bounded by $(b - a)\epsilon$:

$$|R_{XF}| \leq |b - a|\epsilon$$

where ϵ is the relative rounding error of the arithmetic system in use. This error bound is about as good as it is reasonably possible to expect.
We do not cover Romberg's methods here.

Exercise: Numerical Integration

Consider the following probability density function

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-x^2}{2\sigma^2}$$

1. Compute numerically the probability that $x \in [-\sigma; \sigma]$ using the trapezoidal and the Simpson methods for different values of h.
2. Compute numerically the probability that $x \in [-2\sigma; 2\sigma]$ using the trapezoidal and the Simpson methods for different values of h.
3. Compute numerically the probability that $x \in [-3\sigma; 3\sigma]$ using the trapezoidal and the Simpson methods for different values of h.

Simple Monte Carlo Integration I

Example

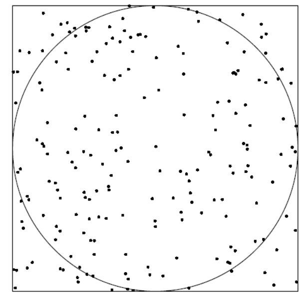
Consider a circle inscribed in a square. It is easy to find the area of the square than the circle (unless you remember your formula). You can numerically find the area of the circle by

1. compute the area A_s of the square.
2. throw darts to onto the square.
3. count how many darts n^* are inside the circle.
4. count how many darts n are inside the square.

Then the surface of the circle is:

$$A_c \approx \frac{n^*}{n} A_s$$

Simple Monte Carlo Integration II



A Monte Carlo method is *any method which solves a problem by generating suitable random numbers and observing that fraction of the numbers obeying some property or properties.*

Simple Monte Carlo Integration III

Consider the integral:

$$I = \int_a^b f(x)\ dx$$

with f continuous on $[a,b]$. Lets define $g(x)$ the uniform probability density function over $[a,b]$:

$$g(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

then I can be rewritten:

$$I = (b-a)\int_a^b g(x)\ f(x)\ dx = (b-a)\ \mathbb{E}[f(x)]$$

Simple Monte Carlo Integration IV

Numerically, it is solved by using n randomly selected points in $[a,b]$

$$\mathbb{E}[f(x)] \approx \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

and

$$I \approx \frac{(b-a)}{n}\sum_{i=1}^{n} f(x_i)$$

The uncertainty can also be computed. As $n \to \infty$, the MC estimate of the integral I converges toward the true value.

Simple Monte Carlo Integration V

Theorem (Monte Carlo Integration)

Lets pick N random points $(x_1 \cdots x_n)$, uniformaly distributed in a multidimensional volume V. The basic theorem of Monte-Carlo integration estimate the integral f over V by:

$$\int f\,dV \approx V\overline{f} \pm V\sqrt{\frac{\overline{f^2}-\overline{f}^2}{n}}$$

with $\overline{f} = \frac{1}{n}\sum_{i=1}^{n} f(x_i)$ and $\overline{f^2} = \frac{1}{n}\sum_{i=1}^{n} f^2(x_i)$

If you want to integrate a function g on a domain W that is not easy to sample randomly, just find a region V such that $W \subset V$ and such that V can easily be sampled. Then define f on V such that $f = g$ on W, and $f \equiv 0$ otherwise.

Simple Monte Carlo Integration VI

Exercise: Consider the surface domain defined by $D:\ x^2+y^2<R^2$ and $f(x,y) = \exp(-x^3+2y^2)$ for $(x,y) \in D$, 0 otherwise.

1. propose a way to randomly find n points in D.
2. write the Monte Carlo integration of f.