Numerical Integration

October 18, 2015

As one learns from analysis, most functions are integrable. Moreover, integration is a *smoothing* operation, so the integral of a function is "more differentiable" than the original function.

We begin with rules for integrating well-behaved functions.

1 Integration using polynomial interpolation

We can approximate our given function by a polynomial interpolation. The integral of this polynomial is easy to find.

For example, the linear interpolate of a function f between x_0 and $x_1 = x_0 + h$ is given by

$$l(x) = (1/h)\left(f(x_0)(x_1 - x) + f(x_1)(x - x_0)\right)$$

Thus, we calculate

$$\int_{x_0}^{x_1} l(x)dx = (h/2)\left(f(x_0) + f(x_1)\right)$$

We can also see this as the area of the trapezoid with base of length h and perpendicular sides of length $f(x_0)$ and $f(x_1)$ respectively. For this reason, this numerical integration procedure is called the Trapezoidal rule.

Going to the next level, we write the quadratic interpolate of f at the points x_0, $x_1 = x_0 + h$ and $x_2 = x_0 + 2h$ as

$$q(x) = (1/2h^2)\left(f(x_0)(x - x_2)(x - x_1) + 2f(x_1)(x - x_0)(x_2 - x)f(x_2)(x - x_0)(x - x_1)\right)$$

Now, we calculate term by term using the substitution $x = x_0 + sh$,

$$\int_{x_0}^{x_2} (x - x_0)(x - x_1)dx = h^3 \int_0 2s(s - 1)ds = \frac{2h^3}{3}$$

Similarly,

$$\int_{x_0}^{x_2} (x - x_2)(x - x_1)dx = h^3 \int_0 2(s-2)(s-1)ds = \frac{2h^3}{3}$$

and finally

$$\int_{x_0}^{x_2} (x_2 - x)(x - x_0)dx = h^3 \int_0 2(2-s)sds = \frac{4h^3}{3}$$

Combining these, we obtain the formula, (note that $x_2 - x_0 = 2h$)

$$\int_{x_0}^{x_2} q(x)dx = (h/3)\left(f(x_0) + 4f(x_1) + f(x_2)\right)$$

This is called Simpson's rule for (numerical) integration.

We can naturally take this furher by interpolating the function by a polynomial of degree n using its values at x_0, $x_1 = x_0 + h$, $\ldots$, $x_n = x_0 + nh$. We write the polynomial in terms of $x = x_0 + sh$ as the formula is easier to read,

$$p(x_0 + sh) = \sum_{k=0}^{n} f(x_k) \frac{\prod_{j=0 j\neq k}^{n}(s-j)}{\prod_{j=0 j\neq k}^{n}(k-j)}$$

One can calculate the relevant integrals to obtain some general rules for numerical integration. We have some rational numbers which are *universal* in the sense that they do not depend on the function f.

$$C_{k,n} = \int_0^n \frac{\prod_{j=0 j\neq k}^{n}(s-j)}{\prod_{j=0 j\neq k}^{n}(k-j)} ds$$

It follows from the above calculation that

$$\int_{x_0}^{x_n} p(x)dx = h\int_0^n p(x_0 + sh)ds = h\sum_{k=0}^{n} f(x_k)C_{k,n}$$

This gives the general rule for numerical integration using polynomial interpolation.

2 Error estimation

To estimate the error involved in the above integration process, we assume that the function f is differentiable to high (at least 5) orders.

First of all, we note that Simpson's rule is based on quadratic interpolation. Hence, if $f(x) = Ax^2 + Bx + C$ is a quadratic polynomial functio, then Simpson's rule gives an *exact* answer; in other words, it gives the same answer as actually integrating the function.

Now consider Simpson's Rule applied to the integral of x^3 from x_0 to $x_0 + 2h$. We have $x^3 = (x - x_0 - h)^3 + q(x)$ for some quadratic polynomial in x. Since Simpson's rule is additive (i.e. gives a linear functional on the space of functions) we see that we need to understand Simpson's Rule on $(x - x_0 - h)^3$. This function has value 0 at $x_0 + h$ and has values of opposite signs at x_0 and $x_0 + 2h$. It follows that Simpson's Rule gives the same result as integration for all *cubic* polynomials as well!

If f is differentiable upto order 4, then we can use the mean value theorem to write

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + f^{(iv)}(\xi)\frac{x^4}{4!}$$

for suitable constants a_i and for some point (ξ) in the interval containing x and the points x_0, x_1 and x_2. If $f^{(iv)}$ is continuous in this interval, then we see that the last term is bound by a constant multiple of x^4 in such an interval. Applying Simpson's rule to the function thus gives an number that differs from the integral of f over (x_0, x_2) by at most a constant multiple of h^5. Hence, upon halving the interval, the error should get divided by 32.

This leads us to the adaptive method for applying Simpson's rule which is described briefly as follows. To calculate the integral of f over (a, b) upto an error e we first take $x_0 = a$, $x_1 = (a + b)/2$ and $x_2 = b$ and apply Simpson's Rule to obtain w. We then take the left interval $(a, (a + b)/2)$ and the right interval $((a + b)/2, b)$ and apply Simpson's Rule to each half to obtain the values l and r respectively. We check if $|l + r - w| < e$, and if that works we accept the value w. Otherwise, we repeat the process with the two intervals *separately* but with the error replaced by $e/2$ in each case. The idea is that if $f^{(iv)}$ is roughly constant over the interval (a, b), then each subdivision should improve our accuracy as seen above. Hence, after enough subdivisions, a small enough error is indicative of an accurate answer. (However, note that this is *not* a proof of accuracy.)