Numerical Evaluation of Derivatives

The numerical evaluation of derivatives is one of the topics less investigated in numerical methods. Whenever possible, it should be avoided since the process can involve large errors. For example, the polynomials with high degree tend to oscillate between points of constraints.

Using the definition of the derivative

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}, \tag{1}$$

one can also encounter high numerical errors.

The Mathematica command for derivation is **Derivative**$[n_1, n_2, \ldots][f][x, y, \ldots]$, where n_1, n_2, $\ldots$ refer to the number of derivations with respect to the first variable x, to the second y, and so on and f is the function to be derived.

1 Classical Difference Formulas (Derivation through Interpolation)

This method is based on interpolation and therefore, in order to determine the n-th derivative of a function, one needs the value of the function in $n+1$ points.

We define

$$\Delta f(x_i) = f(x_i) - f(x_{i-1}) \tag{2}$$

$$\Delta x = x_i - x_{i-1}. \tag{3}$$

The difference is linear and

$$\Delta^n f(x) = \Delta(\Delta^{n-1} f(x)). \tag{4}$$

One can approximate

$$\frac{df(x)}{dx} \approx \frac{\Delta f(x)}{\Delta x}. \tag{5}$$

Fundamental theorem of finite differences calculus

The n-th difference of a polynomial of degree n is a constant and the $n+1$ difference is zero.

Derivation using interpolation

One can approximate the function to be derived using a polynomial

$$f(x) \approx P_n(x). \tag{6}$$

For example, using Lagrange interpolation

$$f(x) \approx P(x) = \sum_{i=0}^{n} \left(y_i \prod_{j=0}^{i-1} \frac{x - x_j}{x_i - x_j} \prod_{j=i+1}^{n} \frac{x - x_j}{x_i - x_j} \right), \tag{7}$$

where $y_i = f(x_i)$.

The polynomial of degree n that approximates the function can be generally written as:

$$P(x) = \sum_{i=0}^{n} a_i x^i. \tag{8}$$

Then, from equation (7), it follows that[1]

$$f'(x) \approx \sum_{i=1}^{n} i a_i x^{i-1} \tag{9}$$

$$f''(x) \approx \sum_{i=2}^{n} i(i-1) a_i x^{i-2} \tag{10}$$

$$\ldots\ldots\ldots \tag{11}$$

$$f^{(j)}(x) \approx \sum_{i=j}^{n} i(i-1)\ldots(i-j+1) a_i x^{i-j} \tag{12}$$

$$\ldots\ldots\ldots \tag{13}$$

$$f^{(n)}(x) \approx n! a_n. \tag{14}$$

As it will be noticed in practical examples, the approximation becomes problematic at end points and it can not function properly for higher-order derivatives.

2 Richardson Extrapolation

Richardson extrapolation is based on setting the spacing between different types of differences close to zero. The method greatly improves the formulas discussed before.

Consider $f(x)$ a function that can be represented by a Taylor series. If for a small h we define

$$x = x_0 + kh, \quad k = 1, 2, \ldots \tag{15}$$

then

$$f(x_0 + kh) = f(x_0) + khf'(x_0) + \frac{1}{2!}(kh)^2 f''(x_0) + \frac{1}{3!}(kh)^3 f^{(3)}(x_0) + \cdots + \frac{1}{n!}(kh)^n f^{(n)}(x_0) \tag{16}$$

$$f(x_0 - kh) = f(x_0) - khf'(x_0) + \frac{1}{2!}(kh)^2 f''(x_0) - \frac{1}{3!}(kh)^3 f^{(3)}(x_0) + \ldots. \tag{17}$$

Subtracting the two expressions, one obtains:

$$f(x_0+kh) - f(x_0-kh) = 2khf'(x_0) + \frac{2}{3!}(kh)^3 f^{(3)}(x_0) + \frac{2}{5!}(kh)^5 f^{(5)}(x_0) + \cdots + \frac{2}{(2n+1)!}(kh)^{2n+1} f^{(2n+1)}(x_0). \tag{18}$$

Analogously,

$$f(x_0 + kh) + f(x_0 - kh) = 2f(x_0) + \frac{2}{2!}(kh)^2 f''(x_0) + \frac{2}{4!}(kh)^4 f^{(4)}(x_0) + \cdots + \frac{2}{(2n)!}(kh)^{2n} f^{(2n)}(x_0). \tag{19}$$

Depending on the values of k, one has different rules for determining the derivatives of $f(x)$.

[1]When selecting the coefficients in Mathematica using the command **CoefficientList**$[P(x), x]$, the output will be a list $a = \{a_1, a_2, \ldots, a_n\}$. The polynomial in equation (8) will have its first coefficient a_1 and not a_0. Therefore, in the following formulas, a_i will be replaced by a_{i+1}. Consequently, the sums in (7) will start from 1.

2.1 Three point rule

For $k = 1$, equation (18) is equivalent to

$$f(x_0 + h) - f(x_0 - h) = 2hf'(x_0) + \frac{2}{6}h^3 f^{(3)}(x_0) + \mathcal{O}(h^5). \tag{20}$$

Up to order h^2, one can determine the first order derivative of the function as:

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0 - h)}{2h} + \mathcal{O}(h^2) \equiv D_1^{three}(x_0, h) + \mathcal{O}(h^2). \tag{21}$$

The truncation error bound for $x_0 \in [a, b]$ is

$$E_1^{three}(h) = \left| -\frac{f^{(3)}(x_0)}{6} h^2 \right| \leq \frac{M_3}{6} h^2, \tag{22}$$

with $M_3 = \max_{a \leq x \leq b} |f^{(3)}(x)|$.

Analogously, from (19) one can deduce the formula for the second order derivative.

$$f(x_0 - h) - 2f(x_0) + f(x_0 + h) = h^2 f''(x_0) + \frac{1}{12} h^4 f^{(4)}(x_0) + \mathcal{O}(h^6) \tag{23}$$

and

$$f''(x_0) = \frac{f(x_0 - h) - 2f(x_0) + f(x_0 + h)}{h^2} + \mathcal{O}(h^2) \equiv D_2^{three}(x_0, h) + \mathcal{O}(h^2). \tag{24}$$

The truncation error bound for $x_0 \in [a, b]$ is

$$E_2^{three}(h) = \left| -\frac{f^{(4)}(x_0)}{12} h^2 \right| \leq \frac{M_4}{12} h^2, \tag{25}$$

with $M_4 = \max_{a \leq x \leq b} |f^{(4)}(x)|$.

2.2 Five point rule

In order to derive the five point rule, we write equation (18) for both $k = 2$ and $k = 1$ up to order 5 in h.

$$f(x_0 + h) - f(x_0 - h) = 2hf'(x_0) + \frac{2}{6}h^3 f^{(3)}(x_0) + \frac{2}{5!}h^5 f^{(5)}(x_0) + \ldots \tag{26}$$

$$f(x_0 + 2h) - f(x_0 - 2h) = 4hf'(x_0) + \frac{16}{6}h^3 f^{(3)}(x_0) + \frac{2}{5!}(2h)^5 f^{(5)}(x_0) + \ldots . \tag{27}$$

By multiplying the first equation with 8 and then subtracting the second one, we can eliminate the terms with $f^{(3)}(x_0)$ and write the five point rule for the first order derivative of $f(x)$:

$$f'(x_0) = \frac{f(x_0 - 2h) - 8f(x_0 - h) + 8f(x_0 + h) - f(x_0 + 2h)}{12h} + \mathcal{O}(h^4) \equiv D_1^{five}(x_0, h) + \mathcal{O}(h^2). \tag{28}$$

The truncation error bound for $x_0 \in [a, b]$ is

$$E_1^{five}(h) = \left| \frac{f^{(5)}(x_0)}{30} h^4 \right| \leq \frac{M_5}{30} h^4, \tag{29}$$

with $M_5 = \max_{a \leq x \leq b} |f^{(5)}(x)|$.

Formula (28) has a surprisingly good convergence rate with the decrease of h and it is exact for cubic polynomials as it involves 4 points of the function.

Theorem: Richardson extrapolation

The central difference formula for the first order derivative based on five points (equation (28)) is a linear combination of $D_1^{three}(x,h)$ and $D_1^{three}(x,2h)$

$$f'(x) \approx D_1^{five}(x,h) = \frac{4D_1^{three}(x,h) - D_1^{three}(x,2h)}{3} + \mathcal{O}(h^4). \tag{30}$$

2.3 Generalization of the method

For a function that can be written as

$$f(x) = D(x,\alpha h) + Ch^n + \mathcal{O}(h^m), \tag{31}$$

with $m > n$, $\alpha > 0$ and $\alpha \neq 1$, the first order derivative can be expressed as

$$f'(x) \approx \frac{\alpha^n D(x,h) - D(x,\alpha h)}{\alpha^n - 1} + \mathcal{O}(h^m). \tag{32}$$