Substitution of derivatives with finite differences

Finite differences: Approximations of derivatives $f'(\widehat{x}),\ f''(\widehat{x}),\ \ldots$ with function values $f(x_k)$ at some (finite) set of points $x_k,\ k = 1, \ldots\ K$.

Taylor theorem: Let the function $f : R \to R$ be $(n+1)$ times differentiable at some open interval $I \subset R$ and let the closed interval between points $\widehat{x}$ and x lie inside I. Let $h = x - \widehat{x}$. Then

$$f(\widehat{x} + h) = f(\widehat{x}) + f'(\widehat{x})\, h + \frac{f''(\widehat{x})}{2!}\, h^2 + \cdots + \frac{f^{(n)}(\widehat{x})}{n!}\, h^n + \mathcal{O}(h^{n+1})$$

Big $\mathcal{O}$ notation $g(h) = \mathcal{O}(h^k)$ describes the limiting behavior of a function $g(h)$ as $h \to 0$: Function g is said to be $\mathcal{O}(h^k)$, if there exists a constant M such that $|g(h)| < M\, |h|^k$ at some interval $0 < |h| < h_0$. (For more detailed explanation of this, see the next page.)

Approximation of the first derivative:

Forward difference: Let f be twice differentiable at I, let $h > 0$. Then

$$f'(\widehat{x}) \;=\; \frac{f(\widehat{x} + h) - f(\widehat{x})}{h} + \mathcal{O}(h)$$

Proof: $f(\widehat{x} + h) = f(\widehat{x}) + f'(\widehat{x})\, h + \mathcal{O}(h^2)$, then rearange and divide by h.

Backward difference: Let f be twice differentiable at I, let $h > 0$. Then

$$f'(\widehat{x}) \;=\; \frac{f(\widehat{x}) - f(\widehat{x} - h)}{h} + \mathcal{O}(h)$$

Proof: $f(\widehat{x} - h) = f(\widehat{x}) - f'(\widehat{x})\, h + \mathcal{O}(h^2)$, then rearange and divide by h.

Central difference: Let f be 3 times differentiable at I, let $h > 0$. Then

$$f'(\widehat{x}) \;=\; \frac{f(\widehat{x} + h) - f(\widehat{x} - h)}{2h} + \mathcal{O}(h^2)$$

Proof: $f(\widehat{x} + h) = f(\widehat{x}) + f'(\widehat{x})\, h + \frac{f''(\widehat{x})}{2!}\, h^2 + \mathcal{O}(h^3)$
$f(\widehat{x} - h) = f(\widehat{x}) - f'(\widehat{x})\, h + \frac{f''(\widehat{x})}{2!}\, h^2 + \mathcal{O}(h^3)$

After subtraction: $f(\widehat{x} + h) - f(\widehat{x} - h) = 2f'(\widehat{x})\, h + \mathcal{O}(h^3)$, then after division by $2h$, the desired result is obtained.

Approximation of the second derivative:

Second Central difference: Let f be 4 times differentiable at I, let $h > 0$. Then

$$f''(\widehat{x}) \;=\; \frac{f(\widehat{x} + h) - 2f(\widehat{x}) + f(\widehat{x} - h)}{h^2} + \mathcal{O}(h^2)$$

Proof: $f(\widehat{x} + h) = f(\widehat{x}) + f'(\widehat{x})\, h + \frac{f''(\widehat{x})}{2!}\, h^2 + \frac{f^{(3)}(\widehat{x})}{3!}\, h^3 + \mathcal{O}(h^4)$
$f(\widehat{x} - h) = f(\widehat{x}) - f'(\widehat{x})\, h + \frac{f''(\widehat{x})}{2!}\, h^2 - \frac{f^{(3)}(\widehat{x})}{3!}\, h^3 + \mathcal{O}(h^4)$

After addition: $f(\widehat{x} + h) + f(\widehat{x} - h) = 2f(\widehat{x}) + f''(\widehat{x})\, h^2 + \mathcal{O}(h^4)$, then after division by h^2, the desired result is obtained.

Big $\mathcal{O}$ notation

Notation $g(h) = \mathcal{O}(h^p)$ does not mean literally $g(h)$ is "equal" to $\mathcal{O}(h^p)$, rather it stands for a statement like *function* $g(h)$ *belongs to the class* $\mathcal{O}(h^p)$. This class, or a set of functions, consists of functions for which inequality $|g(h)| < M\,|h|^p$ holds for h close to zero, or strictly speaking for which there exists constants M and h_0 such that in the interval $0 < |h| < h_0$ the inequality holds. Note that the constants M and h_0 are generally different for different functions $g(h)$. Variable h usually represents the length of a step. We typically discuss the order that provides the tightest upper bound.

Taylor's remainder $R_n(h) = \mathcal{O}(h^{n+1})$:

Suppose $f^{(n+1)}$ is continuous in $< a,\, b >$ and $\widehat{x},\, \widehat{x}+h \in < a,\, b >$. Then

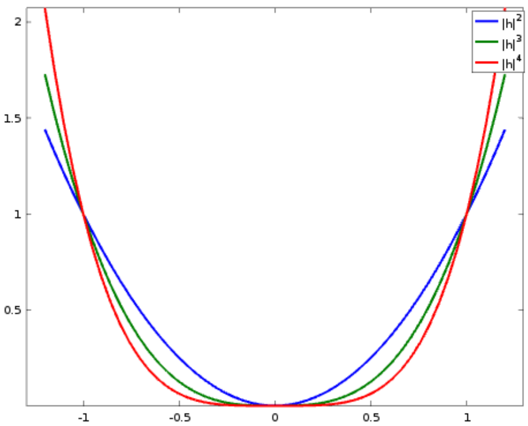
$$f(\widehat{x}+h) = f(\widehat{x})+f'(\widehat{x})\,h+\frac{f''(\widehat{x})}{2!}\,h^2+\cdots+\frac{f^{(n)}(\widehat{x})}{n!}\,h^n+\underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}\,h^{n+1}}_{R_n(h)},$$

where ξ lies between $\widehat{x}$ and $\widehat{x}+h$.

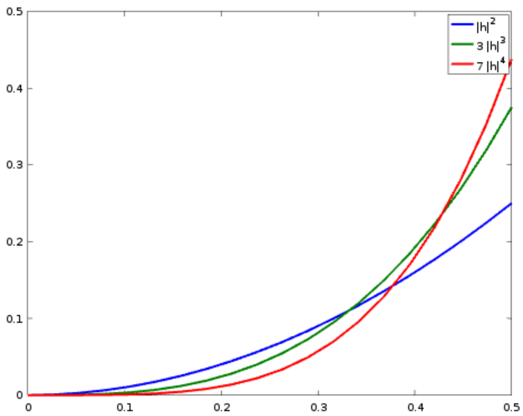
$$|R_n(h)| = |f^{(n+1)}(\xi)|\cdot\frac{1}{(n+1)!}\,|h|^{n+1} \leq \frac{M}{(n+1)!}\,|h|^{n+1}, \qquad \text{where } M = \max_{x\in<a,\,b>} |f^{(n+1)}(x)|\,.$$

(Such constant M exists as continuity of $f^{(n+1)}$ in $< a,\, b >$ is assumed.)

Behavior of $c\,|h|^n$:



Left: $|h|^2$ blue, $|h|^3$ green, $|h|^4$ red.



Right: $|h|^2$ blue, $3\,|h|^3$ green, $7\,|h|^4$ red.

Figure on the right illustrates that close to zero, any multiple of a higher power of $|h|^n$ always tends to zero more quickly than a multiple of its lower power.

Basic computations with $\mathcal{O}(h^p)$

- $h \cdot \mathcal{O}(h^p) = \mathcal{O}(h^{p+1})$ ... means that if $g(h) = \mathcal{O}(h^p)$, then $h \cdot g(h) = \mathcal{O}(h^{p+1})$

  proof: $|h\, g(h)| = |h|\, |g(h)| < |h|\, M\, |h|^p = M\, |h|^{p+1}$

- $$\frac{\mathcal{O}(h^p)}{h} = \mathcal{O}(h^{p-1}) \text{ ... means that if } g(h) = \mathcal{O}(h^p), \text{ then } g(h)/h = \mathcal{O}(h^{p-1})$$

  proof: similarly as above

- $a \cdot \mathcal{O}(h^p) = \mathcal{O}(h^p)$ ... means that if $g(h) = \mathcal{O}(h^p)$, then $a \cdot g(h) = \mathcal{O}(h^p)$

  proof: $|a\, g(h)| = |a|\, |g(h)| < |a|\, M\, |h|^p$

- if $p \le q$, then $\mathcal{O}(h^p) \pm \mathcal{O}(h^q) = \mathcal{O}(h^p)$ (specially $\mathcal{O}(h^p) - \mathcal{O}(h^p) = \mathcal{O}(h^p)$)

  ... means that if $g(h) = \mathcal{O}(h^p)$, $f(h) = \mathcal{O}(h^q)$, then $|g(h) \pm f(h)| = \mathcal{O}(h^p)$, proof:

  $|g(h) \pm f(h)| \le |g(h)| + |f(h)| < M\, |h|^p + N\, |h|^q \le \max(M, N)\, (|h|^p + |h|^q) < 2\, \max(M, N)\, |h|^p$

  (because $|h|^p > |h|^q$ for $|h| < 1$)

Behavior of errors for methods of different order

Question: What improvement in global error can we expect after halving the step?

Euler method – the first order method: norm of the global error $\|e(h)\| = \mathcal{O}(h)$
$\|e(h)\| < M\, |h|$, $\|e(\frac{h}{2})\| < M\, |\frac{h}{2}| = \frac{M}{2}\, |h|$ ... the error can be expected 2-times less

Midpoint (Collatz) method – the 2-nd order method: $\|e(h)\| = \mathcal{O}(h^2)$
$\|e(h)\| < M\, |h|^2$, $\|e(\frac{h}{2})\| < M\, |\frac{h}{2}|^2 = \frac{M}{4}\, |h|^2$ ... the error can be expected 4-times less

Comment: Not only in numerical analysis the *Big-$\mathcal{O}$* notation occurs. In computer science, when $\mathcal{O}(n^p)$ is used, *large* values of n are considered. It is used for measuring the complexity of algorithms (n typically represents a size of a problem and n^p is a number of algorithm operations). Therefore we should be aware of the type of the limit: either $h \to 0$, which is used in numerical analysis, or $n \to \infty$, which is used in computer science. Usually this is distinguished also by the letters used: h for limit to zero and n for limit to infinity.