NTNU

TMA4215 Numerical Mathematics — Fall 2016

Problem Set 1

1 Assume that a function f is differentiable at a point x.

a) Prove that the derivative can be calculated as a *symmetric/central limit*:

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x-h)}{2h}.$$

The discrete analogue of the symmetric limit is the central difference formula.

Solution. Write

$$\frac{f(x+h)-f(x-h)}{2h} = \frac{1}{2}\left[\frac{f(x+h)-f(x)}{h} + \frac{f(x)-f(x-h)}{h}\right]$$

and let $h \to 0$. Make sure that you understand why the second expression also equals $f'(x)$ (hint: as $h \to 0$, we must have $-h \to 0$ as well.)

b) Does the existence of the symmetric limit at x imply that f is differentiable there?

Solution. The absolute value function $f(x) = |x|$ is not differentiable at 0, but the symmetric limit exists and equals 0. Indeed,

$$\lim_{h\to 0} \frac{|h|-|-h|}{2h} = 0$$

in the symmetric case, while the limit in the ordinary derivative approaches -1 if $h \to 0^-$ (from below) and $+1$ if $h \to 0^+$ (from above), and so does not exist at 0.

2 Assume that f is twice differentiable at x. Show that the second derivative can be calculated as the following symmetric limit:

$$f''(x) = \lim_{h\to 0} \frac{f(x+h)-2f(x)+f(x-h)}{h^2}.$$

Hint: Use l'Hôpital's rule and Exercise 1.

Solution. Let

$$A(h) = f(x+h)-2f(x)+f(x-h) \qquad \text{and} \qquad B(h) = h^2.$$

Since $f''(x)$ exists, both A and B are (continuously) differentiable in an interval around 0. L'Hôpital's rule gives that

$$\lim_{h\to 0} \frac{A(h)}{B(h)} = \lim_{h\to 0} \frac{A'(h)}{B'(h)} = \lim_{h\to 0} \frac{f'(x+h)-f'(x-h)}{2h},$$

and the latter expression equals $f''(x)$ by Exercise 1.

3 Let $h > 0$. Then, if

i) $f \in C^1([x, x+h])$ and f'' exists and is bounded in $(x, x+h)$, we know that

$$\frac{f(x+h)-f(x)}{h} - f'(x) = \mathcal{O}(h) \qquad \text{as} \qquad h \to 0^+;$$

ii) $f \in C^1([x-h, x])$ and f'' exists and is bounded in $(x-h, x)$, we know that

$$\frac{f(x)-f(x-h)}{h} - f'(x) = \mathcal{O}(h) \qquad \text{as} \qquad h \to 0^+;$$

iii) $f \in C^2([x-h, x+h])$ and $f^{(3)}$ exists and is bounded in $(x-h, x+h)$, we know that

$$\frac{f(x+h)-f(x-h)}{2h} - f'(x) = \mathcal{O}(h^2) \qquad \text{as} \qquad h \to 0^+;$$

The definition of the $\mathcal{O}$ (bigoh) notation is as follows:

$$f(x) = \mathcal{O}(g(x)) \quad \text{as} \quad x \to \xi$$

if there are constants $M > 0$ and $\delta > 0$ such that $|f(x)| \le M|g(x)|$ whenever $|x-\xi| \le \delta$. The interpretation of this is that the growth of f is bounded above (within a factor M) by that of g when we are sufficiently close to ξ.

a) The boundedness-assumptions in i)–iii) cannot be dropped if we want to use the $\mathcal{O}$ (bigoh) notation. Why?

Solution. We consider case i); the other two are similar. Since $f \in C^1([x, x+h])$ and f'' exists in $(x, x+h)$, we can apply Taylor's theorem to get that

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(\eta_h)$$

for some $\eta_h \in (x, x+h)$. The subscript is a reminder that η_h depends on h. Hence,

$$\frac{f(x+h)-f(x)}{h} - f'(x) = \frac{h}{2}f''(\eta_h).$$

It is tempting to interpret the right-hand side as $\mathcal{O}(h)$ as $h \to 0^+$, but this means that

$$\frac{h}{2}\left|f''(\eta_h)\right| \le Mh$$

for some $M > 0$ provided that h is sufficiently small. In other words, we implicitly assume that f'' is bounded.

An example where the assumptions fail is

$$f(x) = \begin{cases} x^3 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0; \\ 0 & \text{if } x = 0, \end{cases}$$

which is C^1 on all of $\mathbb{R}$, but with unbounded second derivative as $x \to 0$.

b) Verify i)–iii) by numerical experiments. Use the following test examples:

1) $f(x) = 1/x$ in $x_0 = 1$.

2) $f(x) = \sin x - \cos x$ in $x_0 = 1$ and $x_0 = \pi/4$.

3) $f(x) = \begin{cases} x^2 - 2x & \text{for } x > 1; \\ -x^2 + 2x - 2 & \text{for } x \leq 1 \end{cases}$ in $x_0 = 1$.

In each case, use stepsizes $h = 0.1 \times 2^{-i}$, where $i = 0, 1, \ldots, 10$. Present the results both as a table and as a convergence plot. Comment on the results: are they as expected? If not, explain why.

Solution. In Table 1 we have computed the numerical results of example 1); the two other examples are omitted. As is seen, the order of convergence for the forward (i)) and backward (ii)) schemes is linear, while the central (iii)) scheme is quadratic. These results can also be seen in Figure 1, where we in addition have included the results of examples 2) and 3).

Table 1: Order of convergence of finite difference schemes for example 1).

h	Forward scheme i) $\lvert e(h)\rvert$	Forward scheme i) $\frac{e(h)}{e(h/2)}$	Backward scheme ii) $\lvert e(h)\rvert$	Backward scheme ii) $\frac{e(h)}{e(h/2)}$	Central scheme iii) $\lvert e(h)\rvert$	Central scheme iii) $\frac{e(h)}{e(h/2)}$
0.1	9.091×10^{-2}	1.909	1.111×10^{-1}	2.111	1.010×10^{-2}	4.030
0.1×2^{-1}	4.762×10^{-2}	1.952	5.263×10^{-2}	2.053	2.506×10^{-3}	4.008
0.1×2^{-2}	2.439×10^{-2}	1.976	2.564×10^{-2}	2.026	6.254×10^{-4}	4.002
0.1×2^{-3}	1.235×10^{-2}	1.988	1.266×10^{-2}	2.013	1.563×10^{-4}	4.000
0.1×2^{-4}	6.211×10^{-3}	1.994	6.289×10^{-3}	2.006	3.906×10^{-5}	4.000
0.1×2^{-5}	3.115×10^{-3}	1.997	3.135×10^{-3}	2.003	9.766×10^{-6}	4.000
0.1×2^{-6}	1.560×10^{-3}	1.998	1.565×10^{-3}	2.002	2.441×10^{-6}	4.000
0.1×2^{-7}	7.806×10^{-4}	1.999	7.819×10^{-4}	2.001	6.104×10^{-7}	4.000
0.1×2^{-8}	3.905×10^{-4}	2.000	3.908×10^{-4}	2.000	1.526×10^{-7}	4.000
0.1×2^{-9}	1.953×10^{-4}	2.000	1.954×10^{-4}	2.000	3.815×10^{-8}	4.000
0.1×2^{-10}	9.765×10^{-5}		9.767×10^{-5}		9.536×10^{-9}	

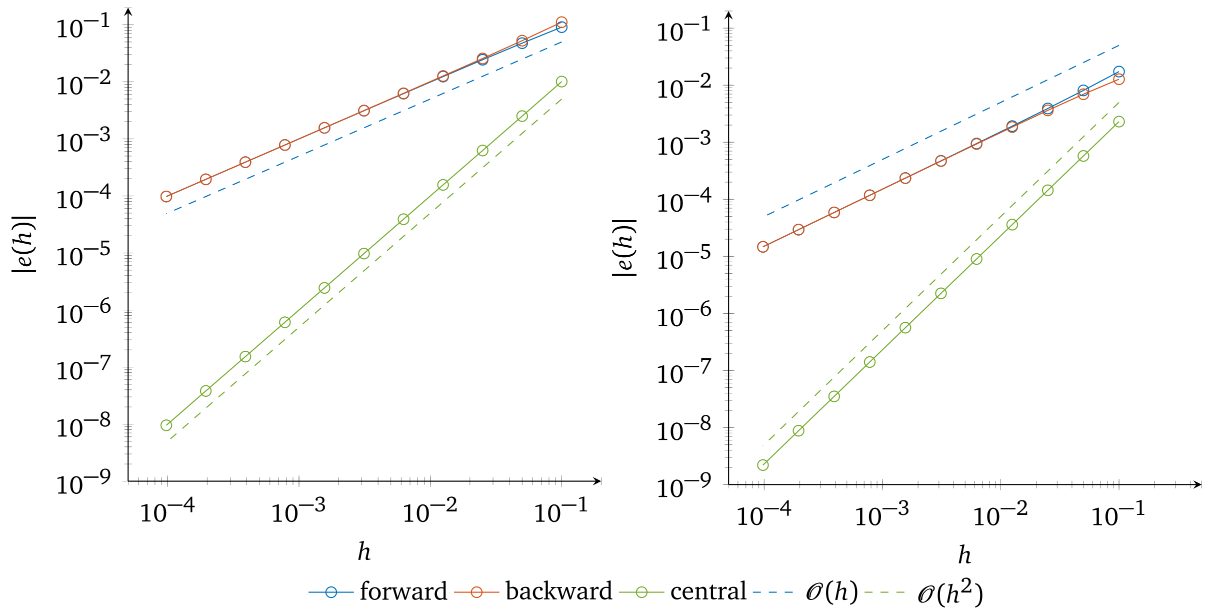
In example 2) with $x_0 = \pi/4$ we see that both the forward and backward schemes are quadratic. Is this in agreement with theory? Yes, first of all, the $\mathcal{O}$-notation indicates just an *upper* bound; the actual convergence speed can be much faster! The quadratic convergence occurs since $f''(\pi/4) = 0$ and f is smooth, so that in particular $f^{(3)}$ exists and is bounded in an interval around x_0.

In example 3) we first note that

$$f'(x) = \begin{cases} 2x - 2 & \text{for } x > 1; \\ -2x + 2 & \text{for } x \leq 1 \end{cases} \quad \text{and} \quad f''(x) = \begin{cases} 2 & \text{for } x > 1; \\ -2 & \text{for } x < 1, \end{cases}$$

while $f''(1)$ does not exist. The assumptions to i) and ii) are fullfilled, giving linear convergence. For the central scheme, however, the nonexistence of $f''(1)$ violates the assumptions, and we lose the guaranteed quadratic convergence[1].

[1]...from this result. But in principle, the fact that assumptions are not satisfied does not imply that the associated conclusion is false; see example 2).



(a) Example 1).

(b) Example 2) with $x_0 = 1$.

(c) Example 2) with $x_0 = \pi/4$.

(d) Example 3).

Figure 1: Order of convergence-plots. In (c) and (d) all the three solid lines lie on top of each other.

$\boxed{4}$ Let

$$y = \varphi_1(x) = \frac{1}{1+2x} - \frac{1-x}{1+x} \qquad \text{and} \qquad y = \varphi_2(x) = \frac{2x^2}{(1+2x)(1+x)}$$

be given.

a) Show that the two expressions are the same.

Solution. Introducing a common denominator in φ_1 gives φ_2.

b) Set $x = 10^{-8}$ and compute y by the two formulas (in MATLAB). Do you get the same result? If not, which of the results will you trust, and why?

Solution. The exact value of y is

$$\frac{1}{5\,000\,000\,150\,000\,001} \approx 1.999\,999\,940\,000\,001 \times 10^{-16},$$

and MATLAB yields (in double precision)

$$\varphi_1(x) = 1.110\,223\,024\,625\,16 \times 10^{-16}$$
$$\text{and} \qquad \varphi_2(x) = 1.999\,999\,94 \times 10^{-16}.$$

From the relative errors

$$|\delta\varphi_1(x)| \approx 0.444\,888 \gg \epsilon_{\mathrm{mach}} \qquad \text{and} \qquad |\delta\varphi_2(x)| \approx 7 \times 10^{-16} \approx 3\epsilon_{\mathrm{mach}}$$

we infer that φ_1 is a dangerous formula, while φ_2 is trustworthy.

Why is this the case? Note that formula φ_1 is well-conditioned (see the solution to the next exercise for the definition) since its condition number equals

$$\frac{3x+2}{(x+1)(1+2x)} \approx 2 \qquad \text{at} \qquad x = 10^{-8}.$$

Computationally, however, the process can be as follows in the computer:

$$\begin{array}{lll} a \leftarrow 2 \cdot x & b \leftarrow 1 - x & \\ a \leftarrow 1 + a & c \leftarrow 1 + x & y \leftarrow a - b \\ a \leftarrow 1/a & b \leftarrow b/c & \end{array}$$

In the last operation we subtract two almost equal numbers ($a \approx b$), which causes the substantial error. All the other arithmetic operations are well-conditioned.

5 Find the condition numbers for the following functions:

a) x^α with α constant; b) $\ln x$; c) $x^{-1}e^x$.

For which x are the problems ill-conditioned?

Solution. The condition number $\kappa(f,x)$ of a function f at a point x is a measure for the sensitivity of the function/problem. If we make an error in the input, is the corresponding output error small or large? More precisely, $\kappa(f,x)$ is defined as the infinitesimal rate of relative error in f at x (output/forward error) divided by the infinitesimal rate of relative error in x (input/backward error). Thus we consider the ratio

$$\left| \frac{[f(x+\Delta x) - f(x)]/f(x)}{\Delta x/x} \right| = \left| \frac{x}{f(x)} \cdot \frac{f(x+\Delta x) - f(x)}{(x+\Delta x) - x} \right|,$$

which becomes

$$\kappa(f,x) = \left|\frac{xf'(x)}{f(x)}\right|$$

as $\Delta x \to 0$. A problem is ill-conditioned at x if $\kappa(f,x) \gg 1$ (much greater than 1), and well-conditioned otherwise.

Routine calculations now yield

a) $\kappa = |\alpha|$, so well-conditioned for all x if $|\alpha|$ is not too big, and ill-cond. for all x when $|\alpha| \gg 1$. For example, if $\alpha = 1000$ and somehow we managed to measure $x = 1.01$ instead of $x = 1$, then the output is suddenly $\approx 21\,000 \gg 1$.

b) $\kappa = |\ln x|^{-1}$, so ill-conditioned whenever $x \approx 1$.

c) $\kappa = |x-1|$, so ill-conditioned when $|x| \gg 1$.

$\boxed{6}$ Define

$$x_n = \int_0^1 t^n (t+5)^{-1}\,\mathrm{d}t.$$

It is given that $x_{20} = 7.9975230282321638314521017213803814812635139208\ldots \times 10^{-3}$.

a) Show that $x_0 = \ln 1.2$, and that $x_n = n^{-1} - 5x_{n-1}$ for $n \geq 1$. Compute x_n for $n = 1, 2, \ldots, 20$ using this reccurence formula. Is x_{20} correct?

b) Now use the reccurence formula backwards to find x_n for $n = 19, 18, \ldots, 0$. Is x_0 correct?

In both cases, explain the behaviour of the reccurence.

Solution. a) Integration by parts gives $x_0 = \ln(t+5)\big|_0^1 = \ln 1.2$, and as regards the reccurence formula, notice that

$$x_n + 5x_{n-1} = \int_0^1 \left(t^n + 5t^{n-1}\right)(t+5)^{-1}\,\mathrm{d}t = \int_0^1 t^n\,\mathrm{d}t = \frac{1}{n}.$$

Computing with yields the incorrect value $x_{20} = 4.242\,637\,044\,921\,560 \times 10^{-3}$ (double precision). The reason for this is that the factor 5 amplifies the round-off error at each step.

b) The backwards reccurence formula is $x_n = \frac{1}{5}(n^{-1} - x_n)$, and we find that the computed x_0 is exactly $\ln 1.2$ in double precision. The problem has now disappeared due to a damping effect from the $\frac{1}{5}$ term.