KTH Engineering Sciences

Order of Accuracy

1 Terminology

We consider a numerical approximation of an exact value u. The approximation depends on a small parameter h, which can be for instance the grid size or time step in a numerical method. We denote the approximation by $\tilde{u}_h$. The numerical method has *order of accuracy* p if there is a number C independent of h such that

$$|\tilde{u}_h - u| \leq Ch^p, \tag{1}$$

at least for sufficiently small h. Hence, the larger the order of accuracy, the faster the error is reduced as h decreases. We say that the *convergence rate* of the method is h^p. The number C typically depends on the exact solution u and possibly on other parameters in the numerical scheme. What is important is that it does not depend on h.

Often the error $\tilde{u}_h - u$ depends smoothly on h. Then there is an error coefficient D such that

$$\tilde{u}_h - u = Dh^p + O\left(h^{p+1}\right). \tag{2}$$

Note that this is not equivalent to (1) since the error may be a non-smooth function of h. We will get back to this issue in Section 4 below. For now, however, we will assume (2) holds.

Example 1 *In numerical differentiation we approximate $u = f'(0)$ by the forward difference*

$$\tilde{u}_h = \frac{f(h) - f(0)}{h}.$$

After Taylor expansion we get

$$\tilde{u}_h - u = \frac{f(0) + hf'(0) + \frac{h^2}{2}f''(\xi) - f(0)}{h} - f'(0) = \frac{h}{2}f''(\xi)$$

for some $\xi \in [0, h]$. Hence, for sufficiently small h, say $0 < h < 1$ we can write as in (1),

$$|\tilde{u}_h - u| \leq Ch, \qquad C = \frac{1}{2}\max_{\xi\in[0,1]} |f''(\xi)|,$$

where C does not depend on h. (But it does depend on f!) The order of accuracy is therefore one. Moreover, if f is three times continuously differentiable we can continue the Taylor expansion one more step to get

$$\tilde{u}_h - u = \frac{h}{2}f''(0) + \frac{h^2}{6}f'''(\xi), \qquad \xi \in [0, h].$$

We thus get also (2) with $D = f''(0)/2$, provided f is sufficiently smooth.

Example 2 *In piecewise linear interpolation of a function $u(x)$ on equidistant nodes in the interval $[a, b]$ we let $\tilde{u}_h$ be the piecewise linear interpolant when the distance between nodes is h. The error can then be bounded as (see Lecture notes 5)*

$$\max_{a \le x \le b} |\tilde{u}_h(x) - u(x)| \le \max_{a \le \xi \le b} \frac{|u''(\xi)|}{8} h^2.$$

If $u(x)$ is two times continuously differentiable, this is thus a second order method. If u is three times continuously differentiable the error also depends smoothly on h such that $\max_x |\tilde{u}_h(x) - u(x)| = Dh^2 + O(h^3)$ for some number D.

Example 3 *In the* trapezoidal *rule we approximate the exact integral*

$$u = \int_a^b f(x)dx,$$

by a sum

$$\tilde{u}_h = \frac{h}{2} f(x_0) + h \sum_{j=1}^{N-1} f(x_j) + \frac{h}{2} f(x_N), \qquad h = \frac{b-a}{N}, \qquad x_j = a + jh.$$

For sufficiently smooth functions $f(x)$ this is a second order method and $\tilde{u}_h - u = Dh^2 + O(h^3)$.

2 Determining the order of accuracy empirically

We are often faced with the problem of how to determine the order of accuracy p given a sequence of approximations $\tilde{u}_{h_1}, \tilde{u}_{h_2}, \ldots$ This is can be a good check that a method is correctly implemented (if p is known) and also a way to get a feeling for the trustworthiness of an approximation $\tilde{u}_h$ (high p means high trustworthiness). We can either be in the situation that the exact value u is known, or, more commonly, that u is unknown.

2.1 Known u

If the exact value u is known, it is straightforward to determine the order of accuracy. Then we can check the sequence

$$\log |\tilde{u}_h - u| = \log |Dh^p(1 + O(h))| = \log |Dh^p| + \log |1 + O(h)| = \log |D| + p \log h + O(h),$$

for h_1, h_2, ... and fit it to a linear function of $\log h$ to approximate p. A quick way to do this is to plot $|\tilde{u}_h - u|$ as a function of h in a `loglog` plot in MATLAB and determine the slope of the line that appears. The standard way to get a precise number for p is to halve the parameter h and look at the ratio of the errors $u - \tilde{u}_h$ and $u - \tilde{u}_{h/2}$,

$$\frac{\tilde{u}_h - u}{\tilde{u}_{h/2} - u} = \frac{Dh^p + O(h^{p+1})}{D(h/2)^p + O((h/2)^{p+1})} = \frac{D + O(h)}{D2^{-p} + O(h)} = 2^p + O(h).$$

Hence

$$\log_2 \left(\frac{\tilde{u}_h - u}{\tilde{u}_{h/2} - u} \right) = p + O(h).$$

Example 4 *The exact integral of $\sin(x)$ over $[0, \pi]$ equals two. Computing $\tilde{u}_h$ with the trapezoidal rule and plotting $|\tilde{u}_h - 2|$ in a `loglog` plot we get the result shown in Figure 1.*

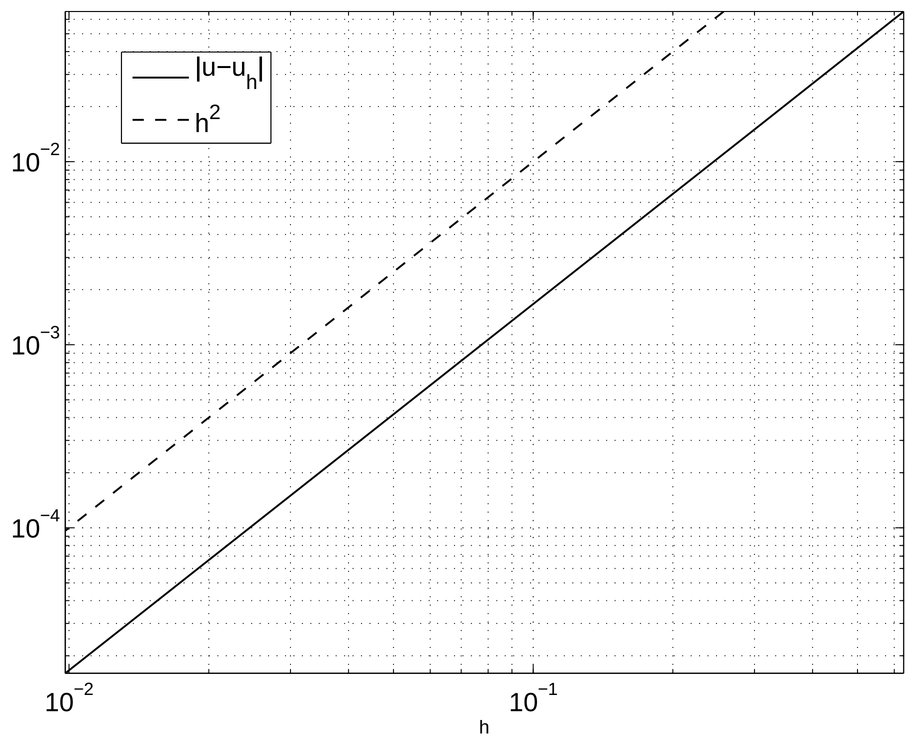

Figure 1. Error in trapezoidal rule for $f(x) = \sin(x)$ as a function of h. The dashed line is h^2 as a function of h which has precisely slope two. It thus indicates the slope for a second order method, for comparison.

h	$\tilde{u}_h$	$\tilde{u}_h - \tilde{u}_{h/2}$	$\frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}}$	$\log_2 \frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}}$
$\pi/5$	1.933765598092805	-0.049757939416650	4.024930251575880	2.008963782835339
$\pi/10$	1.983523537509455	-0.012362435199260	4.006184396966857	2.002228827158397
$\pi/20$	1.995885972708715	-0.003085837788350	4.001543117204195	2.000556454557076
$\pi/40$	1.998971810497066	-0.000771161948770	4.000385593360853	2.000139066704584
$\pi/80$	1.999742972445836	-0.000192771904301	4.000096386716427	2.000034763740606
$\pi/160$	1.999935744350136	-0.000048191814813		
$\pi/320$	1.999983936164949			

Table 1. Table of values for the trapezoidal rule for $f(x) = \sin(x)$. The last column is the final approximation of the order of accuracy p.

2.2 Unknown u

When u is not known there are two main approaches. The first one is to compute a numerical reference solution with a very small h and then proceed as in the case of a known u. This can be quite an expensive strategy if $\tilde{u}_h$ is costly to compute. Using the resulting p to gauge the trustworthiness of $\tilde{u}_h$ is also less relevant when we already have a good reference solution.

The second approach is to look at ratios of differences between $\tilde{u}_h$ computed for different h. Most commonly we compare solutions where h is halved successively. When p is large this gives a fairly good approximation of the error $\tilde{u}_h - u$ since

$$\tilde{u}_h - \tilde{u}_{h/2} = Dh^p - D(h/2)^p + O(h^{p+1}) = Dh^p(1-2^{-p}) + O(h^{p+1}) = (\tilde{u}_h - u)(1-2^{-p}) + O(h^{p+1}).$$

What is more important, however, is that, regardless of p, this difference decays to zero with the same speed as the actual error $\tilde{u}_h - u$. Therefore one can do the same trick as when u is known and consider the ratio of the differences. We get

$$\frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}} = \frac{Dh^p - D(h/2)^p + O(h^{p+1})}{D(h/2)^p - D(h/4)^p + O(h^{p+1})} = \frac{1 - 2^{-p} + O(h)}{2^{-p} - 2^{-2p} + O(h)} = 2^p + O(h). \tag{3}$$

Hence, after computing $\tilde{u}_h$ for h, $h/2$ and $h/4$ we can evaluate the expression above and get an estimate of p, as before

$$\log_2\left(\frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}}\right) = p + O(h).$$

Example 5 *Consider again Example 4. If the exact integral value was not known we would look at the values computed by the trapezoidal rule and check the ratios of differences as above. The result is summarized in Table 1.*

3 Asymptotic region

We note that the estimates of p in all the methods above gets better as $h \to 0$ because of the $O(h)$ term. (The precise value is only given in the limit $h \to 0$.) We say that the method is in its asymptotic region (or range) of accuracy when h is small enough to give a good estimate of p — then the $O(h^{p+1})$ term in (2) is significantly smaller than Dh^p. This required size of h can, however, be quite different for different problems. To verify that we are indeed in the asymptotic region, it can be valuable to make the estimate of p for several different h and check that we get approximately the same value. Usually one therefore computes $\tilde{u}_h$ not just for three values of h, but for a longer sequence, $h, h/2, h/4, h/8, h/16, \ldots$ and compares the corresponding ratios,

$$\frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}}, \quad \frac{\tilde{u}_{h/2} - \tilde{u}_{h/4}}{\tilde{u}_{h/4} - \tilde{u}_{h/8}}, \quad \frac{\tilde{u}_{h/4} - \tilde{u}_{h/8}}{\tilde{u}_{h/8} - \tilde{u}_{h/16}}, \quad \ldots$$

Similarly, if u is known one considers $\tilde{u}_h - u$ for several decreasing values of h when fitting the line.

Example 6 *If we perform the same experiments as in Example 4 and Example 5 above, but with $f(x) = \sin(31x)$ the constant D will be much bigger, meaning that the asymptotic region is shifted to smaller h. The results are shown in Figure 2 and Table 2. It is not until $h < \pi/40 \approx 10^{-1}$ that the numbers start to look reasonable. The general size of the error is also much larger than in Figure 2 because of the bigger D.*

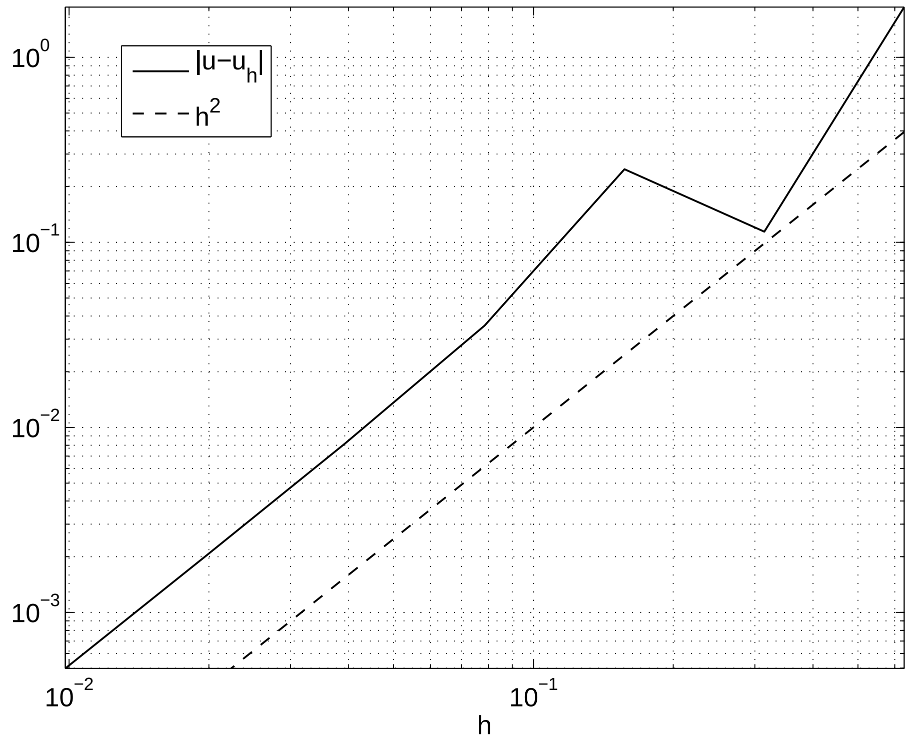

Figure 2. Error in trapezoidal rule for $f(x) = \sin(31x)$. The dashed line is h^2 which indicates the slope for a second order method.

h	$\tilde{u}_h$	$\tilde{u}_h - \tilde{u}_{h/2}$	$\frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}}$	$\log_2 \frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}}$
$\pi/5$	1.933765598092808	1.983523537509458	14.784906442999516	3.886053209184444
$\pi/10$	-0.049757939416650	0.134158680351247	-0.630173999781565	—
$\pi/20$	-0.183916619767896	-0.212891487744257	7.778391902691306	2.959471924644287
$\pi/40$	0.028974867976361	-0.027369601635860	4.437830912882666	2.149854700028653
$\pi/80$	0.056344469612220	-0.006167337641551	4.096338487974619	2.034334932805155
$\pi/160$	0.062511807253771	-0.001505573247830		
$\pi/320$	0.064017380501601			

Table 2. Table of values for the trapezoidal rule for $f(x) = \sin(31x)$. The last column is the final approximation of the order of accuracy p.

h	$\tilde{u}_h$	$\tilde{u}_h - \tilde{u}_{h/2}$	$\frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}}$	$\log_2 \frac{\tilde{u}_h - \tilde{u}_{h/2}}{\tilde{u}_{h/2} - \tilde{u}_{h/4}}$
0.2	0.302842712474619	0.009289321881345	26.142135623725615	4.708305098603142
0.1	0.293553390593274	0.000355339059327	1.999999999999688	0.999999999999775
0.05	0.293198051533946	0.000177669529664	2.000000000001875	1.000000000001352
0.025	0.293020382004283	0.000088834764832	2.635450714080436	1.398049712285012
0.0125	0.292931547239451	0.000033707617584	12.589489353884787	3.654147861537719
0.00625	0.292897839621867	0.000002677441208		
0.003125	0.292895162180659			

Table 3. Table of values for the trapezoidal rule for $f(x) = |x-\alpha|$ with $\alpha = 1/\sqrt{2}$. The last column is the final approximation of the order of accuracy p, which fails for this case.

4 Non-smooth error

So far we have assumed that the error depends smoothly on the parameter h. Then the error is of the form in (2). This is, however not always the case. The error can, for instance, depend discontinuously on h, eventhough it is bounded as in (1). The reason for this can be discontinuities in the method itself (e.g. case switches) or non-smooth functions in the problem (e.g. solutions, sources, integrands). When the error is non-smooth one cannot check convergence rates by looking at ratios of differences as in Section 2.2. Other methods must be used.

Example 7 *Consider the trapezoidal rule applied to the integral*

$$\int_0^1 |x-\alpha| dx,$$

for some value $0 < \alpha < 1$. Here the integrand is not smooth at $x = \alpha$ so the standard second order of accuracy of the trapezoidal rule is not guaranteed. However, since the integrand is linear away from $x = \alpha$, the trapezoidal rule is exact everywhere except in the node interval which contains α. The error there depends crucially on the distance between α and the nearest node. More precisely, if $x_j \le \alpha < x_{j+1}$ and $x_{j+1} - x_j = h$,

$$\begin{aligned} u - \tilde{u}_h &= \int_{x_j}^{x_{j+1}} |x-\alpha| dx - h\frac{|x_j - \alpha| + |x_{j+1} - \alpha|}{2} \\ &= \int_{x_j}^{\alpha} (\alpha - x) dx + \int_{\alpha}^{x_{j+1}} (x-\alpha) dx - h\frac{x_{j+1} - x_j}{2} = \beta(\beta-1)\frac{h^2}{2}, \end{aligned}$$

where $\beta = \beta(h) = (\alpha - x_j)/h$, i.e. the fractional part of α/h, which is a discontinuous function of h. The method is still second order accurate since $|\beta(h)| \le 1$ and (1) therefore holds with $C = 1/8$. However, the results presented in Figure 3 and Table 3 clearly shows the non-smoothness of the error and the failure of the ratios of the differences to predict the order of convergence.

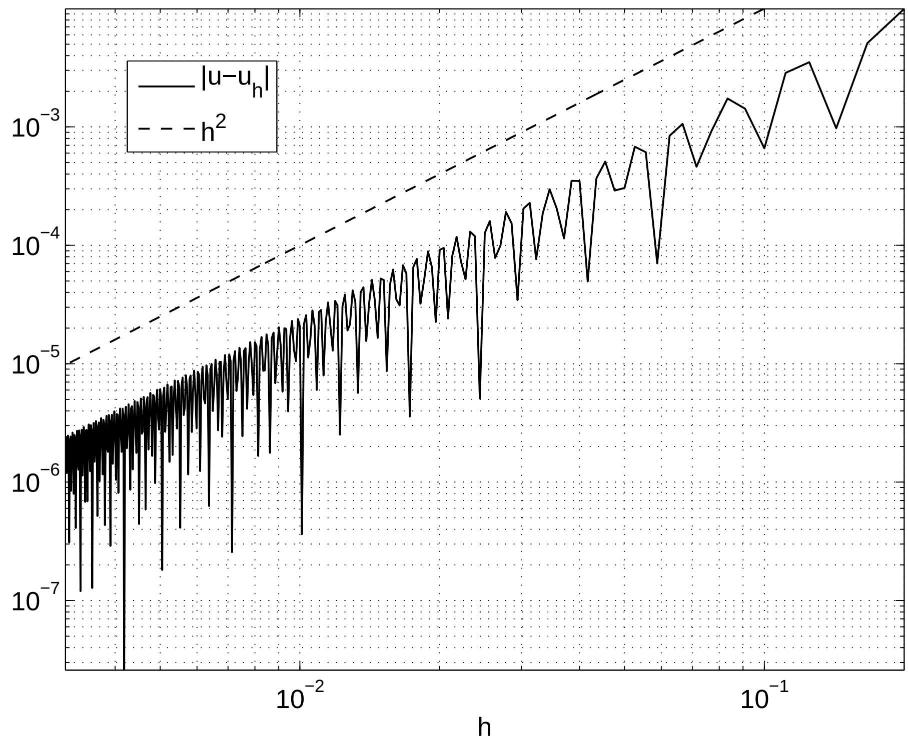

Figure 3. Error in trapezoidal rule for $f(x) = |x - \alpha|$ with $\alpha = 1/\sqrt{2}$. The dashed line is h^2 which indicates the slope for a second order method.