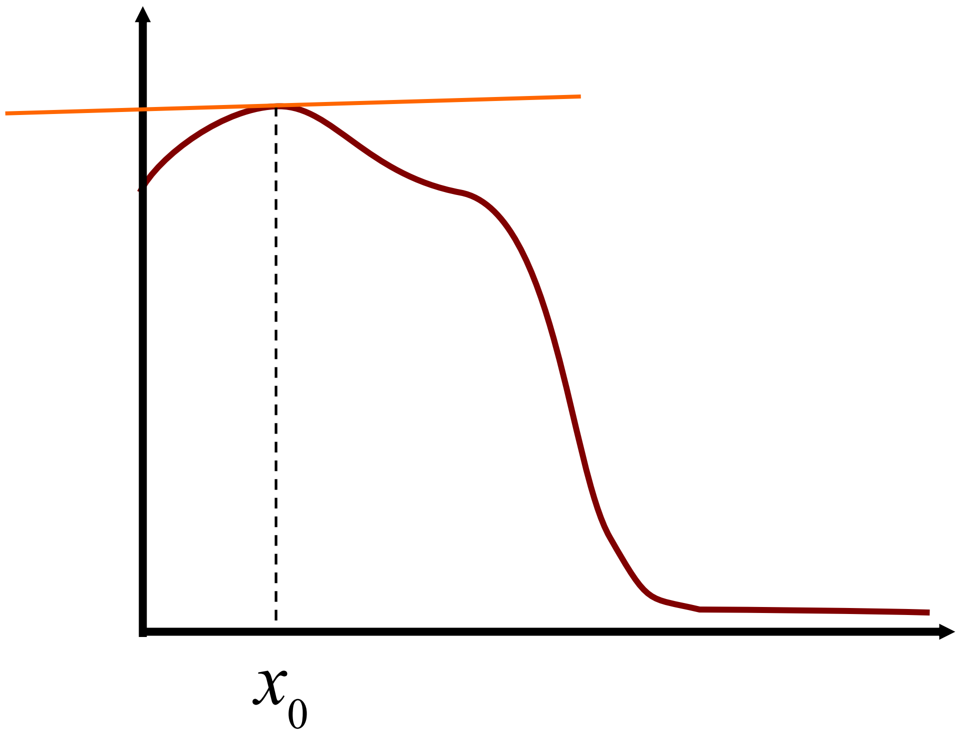
Numerical differentiation

Recall: derivatives

- Function *f : R --> R*, $x_0 \in R$

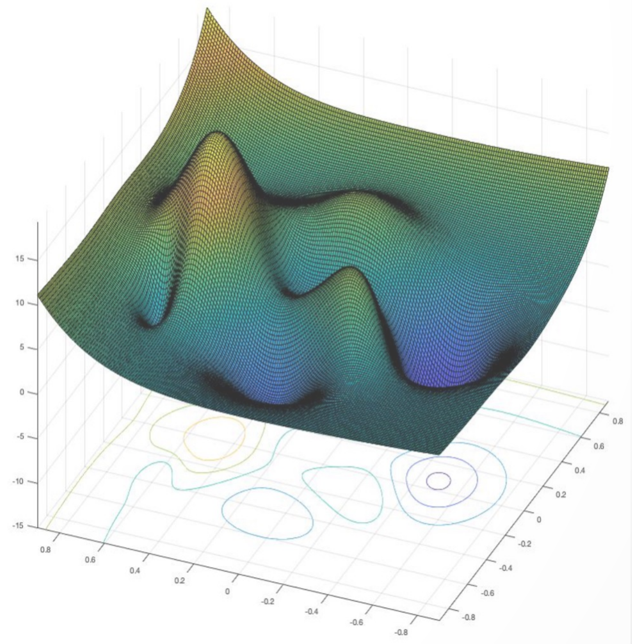
$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

- Interpretation: it is the slope of the tangent line at x_0



Why do we need numerical derivatives?

- Basic operation in other numerical problems:
  - Solving differential equations (both ODEs and PDEs)
  - Numerical optimization
- We need to compute
  - Gradient of f : $\quad g_k = \dfrac{\partial f}{\partial x_k}$
  - Jacobian of $\boldsymbol{f}$: $\quad J_{jk} = \dfrac{\partial f_j}{\partial x_k}$
  - Hessian of f : $\quad H_{jk} = \dfrac{\partial^2 f}{\partial x_j \partial x_k}$



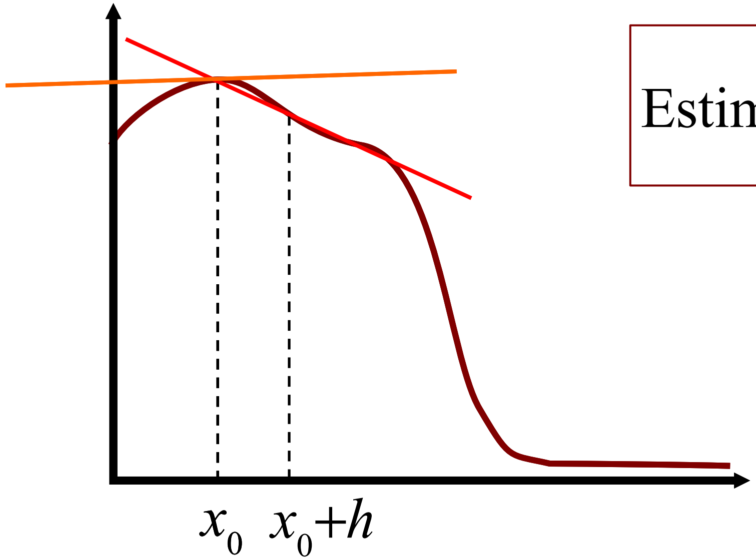
What we care about when computing derivatives numerically?

- Accuracy of our numerical estimate, and how we can improve it
- Computational complexity (how many times our function f needs to be evaluated to obtain the derivative with a desired accuracy)

Forward differences (FDF)

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h}$$

- Numerical differentiation: the limit cannot be achieved, so we have to approximate: *h>0* is fixed (and small)



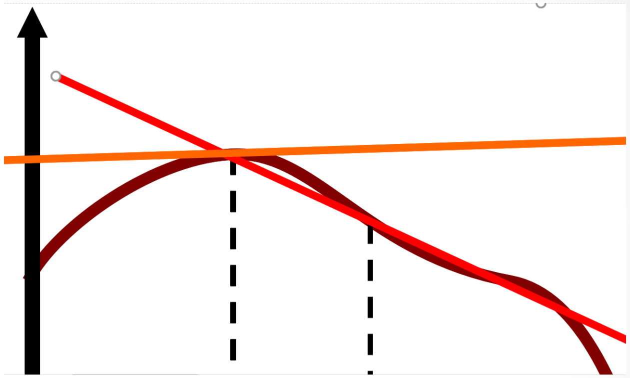
$$\text{Estimate } f'(x_0) \approx \frac{f(x_0+h) - f(x_0)}{h}$$

- This natural approximation is called Forward Differences (FDF)

Interpretations of finite differences

- Mean-value theorem: if a function $f : R \to R$ is continuous on $[a, b]$, and differentiable on (a, b), then there exists $c \in (a, b)$ such that:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$



- Another interpretation: since

$$f(b) - f(a) = \int_a^b f'(x)\,dx$$

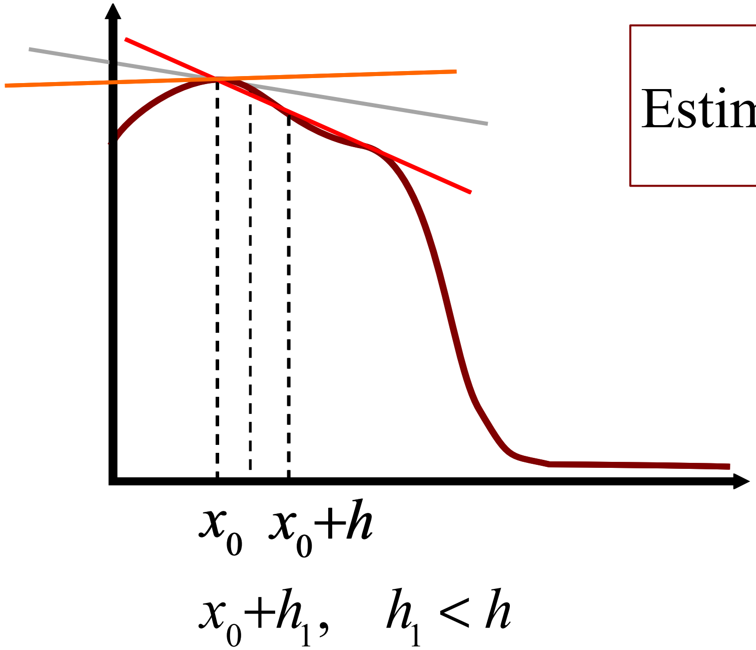
  - in the case of FDF we have $b=x_0+h$, $a=x_0$, and if we divide by h

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{1}{h} \int_{x_0}^{x_0+h} f'(x)\,dx$$

  - i.e. our estimate is the average value of f' over (x_0, x_0+h)

Forward differences (FDF)

- Smaller *h* ---> better approximation



$$\text{Estimate } f'(x_0) \approx \frac{f(x_0+h) - f(x_0)}{h}$$

Forward differences: example

- Let us test this: for $f(x)=e^{(5*x)}$ compute numerically the value of the first derivative at $x_0 = 0.2$ using different values of h:

h	df/dx	Error
1.0e-01	17.63407	4.04266328
1.0e-02	13.93693	0.34551982
1.0e-03	13.62544	0.03403522
1.0e-04	13.59481	0.00339842
1.0e-05	13.59175	0.00033979
1.0e-06	13.59144	0.00003398
1.0e-07	13.59141	0.00000340
1.0e-08	13.59141	0.00000037
1.0e-09	13.59141	0.00000026
1.0e-10	13.59141	0.00000114
1.0e-11	13.59139	0.00001447
1.0e-12	13.59091	0.00050296

How fast FDF approaches $f'(x_0)$?

Truncation error →

$$\left| f'(x_0) - \frac{f(x_0+h) - f(x_0)}{h} \right| = O(?\,h)$$

Reminder: big O notation

For *f, g : R --> R* we can write:

$$f(x) = O\big(g(x)\big) \quad \text{as} \quad \mathrm{x} \to \infty$$

if and only if (iff) there exist positive *C* and x_0 such that

$$|f(x)| \le C|g(x)| \quad \text{for} \quad \forall \mathrm{x} \ge \mathrm{x}_0$$

Similarly, we can write

$$f(x) = O\big(g(x)\big) \quad \text{as} \quad \mathrm{x} \to a$$

iff there exist there exist positive *C* and *e* such that

$$|f(x)| \le C|g(x)| \quad \text{for} \quad |x-a| < e$$

Taylor series:

$$f(x_0+h) = f(x_0) + f'(x_0)h + \frac{f^{(2)}(x_0)h^2}{2!} + \frac{f^{(3)}(x_0)h^3}{3!} + \cdots + O(h^n)$$

$$\Omega = O(h^n) \Rightarrow \exists C \in R, : |\Omega| \le C|h|^n$$

How fast FDF approaches $f'(x_0)$?

Truncation error

$$\left| f'(x_0) - \frac{f(x_0+h) - f(x_0)}{h} \right| = O(?\,h)$$

- We assume that $f : R \dashrightarrow R$ is C^2, i.e., 2x continuously differentiable function, $x_0 \in R$, $h>0$, and $h<C_h$
  - using Taylor expansion we get

$$f(x_0+h) = f(x_0) + f'(x_0)h + f''(\xi)\frac{h^2}{2!},\ x_0 \le \xi \le x_0 + h$$

How fast FDF approaches $f'(x_0)$?

Truncation error →

$$\left| f'(x_0) - \frac{f(x_0+h)-f(x_0)}{h} \right| = O(?h)$$

- We assume that $f : R \dashrightarrow R$ is C^2, i.e., 2x continuously differentiable function, $x_0 \in R$, $h>0$, and $h<C_h$
  - using Taylor expansion we get

$$f(x_0+h) = f(x_0) + f'(x_0)h + f''(\xi)\frac{h^2}{2!}, x_0 \le \xi \le x_0 + h$$

  - so

$$\left| f'(x_0) - \frac{f(x_0+h)-f(x_0)}{h} \right| = \frac{h}{2}\left| f''(\xi) \right|$$

  - and

$$\left| f'(x_0) - \frac{f(x_0+h)-f(x_0)}{h} \right| \le \frac{h}{2} \max_{x_0 \le x \le x_0 + C_h} \left| f''(x) \right|$$

  - i.e.

$$\left| f'(x_0) - \frac{f(x_0+h)-f(x_0)}{h} \right| \le Ch$$

Forward differences: example

- Therefore:

$$\left| f'(x_0) - \frac{f(x_0+h)-f(x_0)}{h} \right| = O(h)$$

10x smaller *h* -> the FDF error 10x smaller

- Let us test this: for *f(x)=e^(5*x)* compute numerically the value of the first derivative at $x_0 = 0.2$ using different values of *h*:

h	df/dx	Error
1.0e-01	17.63407	4.04266328
1.0e-02	13.93693	0.34551982
1.0e-03	13.62544	0.03403522
1.0e-04	13.59481	0.00339842
1.0e-05	13.59175	0.00033979
1.0e-06	13.59144	0.00003398
1.0e-07	13.59141	0.00000340
1.0e-08	13.59141	0.00000037
1.0e-09	13.59141	0.00000026
1.0e-10	13.59141	0.00000114
1.0e-11	13.59139	0.00001447
1.0e-12	13.59091	0.00050296

Round-off error

Round-off error

- Round-off error occurs due to a finite machine precision
- Matlab uses IEEE double precision standard -> the numbers are represented by 52 floating binary digits (or 15 floating decimal digits)
- Test this in Matlab:
  - >> (2^52+1) - 2^52
  - >> (2^53+1) - 2^53

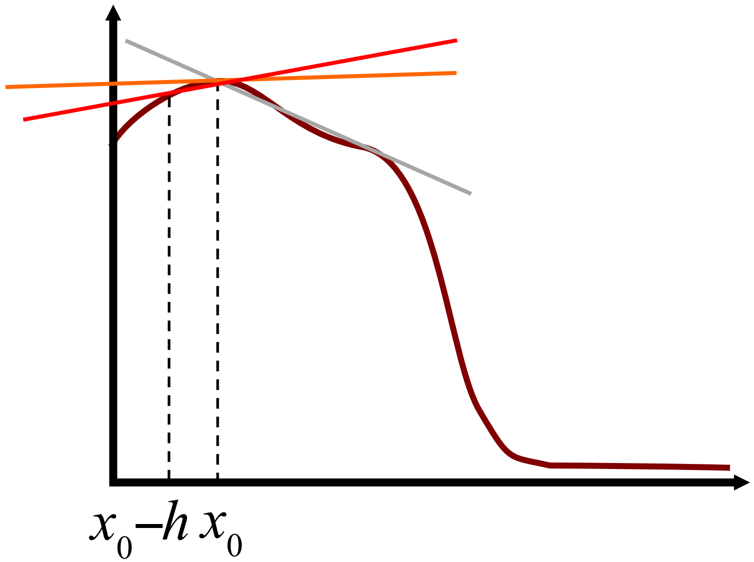
h	Error
1.0e-01	4.04266328
1.0e-02	0.34551982
1.0e-03	0.03403522
1.0e-04	0.00339842
1.0e-05	0.00033979
1.0e-06	0.00003398
1.0e-07	0.00000340
1.0e-08	0.00000037
1.0e-09	0.00000026
1.0e-10	0.00000114
1.0e-11	0.00001447
1.0e-12	0.00050296

- Using a finer grid (smaller *h*) reduces the truncation error
- It also increases the round-off error
- There is a trade-off between these two errors !

Backward differences (BDF)

- Use instead the point on the left-hand side

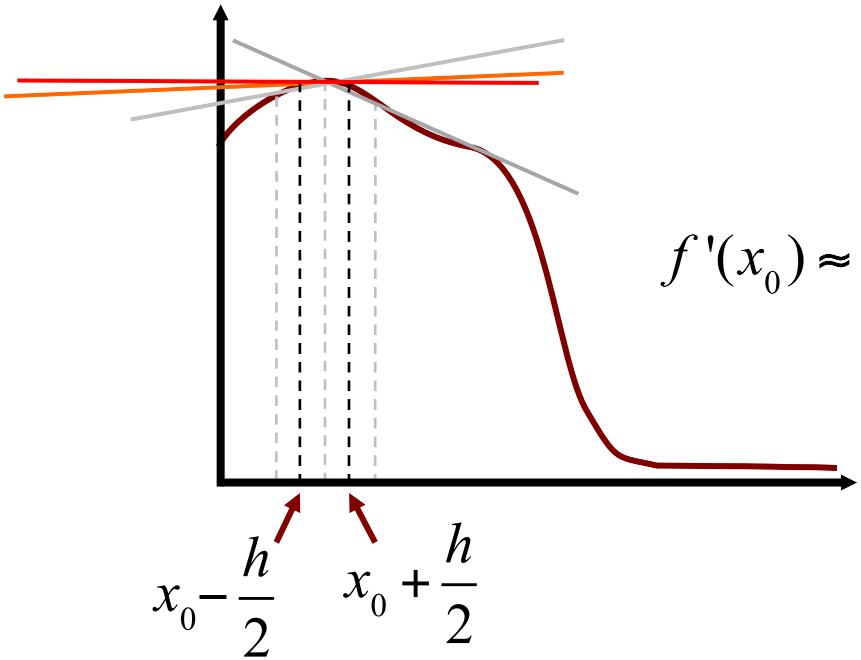
$$f'(x_0) \approx \frac{f(x_0) - f(x_0 - h)}{h}$$



- BDF has similar properties to FDF as *h* gets smaller, i.e.

$$\left| f'(x_0) - \frac{f(x_0) - f(x_0 - h)}{h} \right| = O(h)$$

Central differences (CDF)



$$f'(x_0) \approx \frac{f(x_0 + \frac{1}{2}h) - f(x_0 - \frac{1}{2}h)}{h}$$

- How fast CDF approaches the actual derivative as *h* gets smaller?

$$\left| f'(x_0) - \frac{f(x_0 + \frac{1}{2}h) - f(x_0 - \frac{1}{2}h)}{h} \right| = O(?h)$$

Central differences (CDF): truncation error

- We assume that $f : R \to R$ is C^3, $x_0 \in R$, $0<h<C_h$, and we apply

  Taylor's expansion:

$$f(x_0+\frac{h}{2}) = f(x_0)+f'(x_0)\frac{h}{2}+f''(x_0)\frac{h^2}{8}+f'''(\xi)\frac{h^3}{8\cdot 3!}, \quad x_0 \le \xi \le x_0+\frac{h}{2}$$

$$f(x_0-\frac{h}{2}) = f(x_0)-f'(x_0)\frac{h}{2}+f''(x_0)\frac{h^2}{8}-f'''(\theta)\frac{h^3}{8\cdot 3!}, \quad x_0-\frac{h}{2} \le \theta \le x_0$$

- subtract two expressions:

$$f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2}) = f'(x_0)h+\left(f'''(\xi)\frac{h^3}{8\cdot 3!}+f'''(\theta)\frac{h^3}{8\cdot 3!}\right)$$

- Subsequently, we get

$$\left|f'(x_0)-\left(f(x_0+\frac{1}{2}h)-f(x_0-\frac{1}{2}h)\right)\Big/h\right| = f'''(\xi)\frac{h^2}{8\cdot 3!}+f'''(\theta)\frac{h^2}{8\cdot 3!}$$

$$\Downarrow$$

$$\left|f'(x_0)-\left(f(x_0+\frac{1}{2}h)-f(x_0-\frac{1}{2}h)\right)\Big/h\right| \le \frac{h^2}{8\cdot 3}\max_{x_0-C_h/2\le x\le x_0+C_h/2}\left|f'''(x)\right|$$

Central differences (CDF): example

$$\left| f'(x_0) - \frac{f(x_0 + \frac{1}{2}h) - f(x_0 - \frac{1}{2}h)}{h} \right| = O(h^2)$$

- 10x smaller *h* -> the error of CDF becomes 100x smaller
- Numerical test for $f(x)=e^{(5*x)}$ at $x_0 = 0.2$

h	df/dx	Error
1.0e-01	13.73343	0.14202027
1.0e-02	13.59282	0.00141582
1.0e-03	13.59142	0.00001416
1.0e-04	13.59141	0.00000014
1.0e-05	13.59141	0.00000000
1.0e-06	13.59141	0.00000000
1.0e-07	13.59141	0.00000000
1.0e-08	13.59141	0.00000003
1.0e-09	13.59141	0.00000063
1.0e-10	13.59141	0.00000330
1.0e-11	13.59139	0.00001447
1.0e-12	13.59091	0.00050296

Second derivatives

- Function *f : R --> R* is C^2 and $x_0 \in R$

$$f''(x_0) = \lim_{h \to 0} \frac{f'(x_0 + h/2) - f'(x_0 - h/2)}{h}$$

- using CDF to estimate the 1st derivatives, i.e.,

$$f'(x_0 + h/2) = \frac{f(x_0 + h/2 + h/2) - f(x_0 + h/2 - h/2)}{h} = \frac{f(x_0 + h) - f(x_0)}{h}$$

$$f'(x_0 - h/2) = \frac{f(x_0 - h/2 + h/2) - f(x_0 - h/2 - h/2)}{h} = \frac{f(x_0) - f(x_0 - h)}{h}$$

- We finally get:

$$f''(x_0) \approx \frac{f(x_0 + h) - 2f(x_0) + f(x_0 - h)}{h^2}$$

Partial derivatives

- Function $f : R^2 \text{-->} R$ is C^1 and $x_0, y_0 \in R$

$$\frac{\partial f}{\partial x}(x_0, y_0) \approx \frac{f\left(x_0 + \frac{1}{2}h, y_0\right) - f\left(x_0 - \frac{1}{2}h, y_0\right)}{h}$$

$$\frac{\partial f}{\partial y}(x_0, y_0) \approx \frac{f\left(x_0, y_0 + \frac{1}{2}h\right) - f\left(x_0, y_0 - \frac{1}{2}h\right)}{h}$$

- Exactly the same properties as CDF, i.e., no particularities

Improving estimates

- Several ways to improve the accuracy of the estimates:
  - Decrease the step size
  - Use higher-order expressions employing more points
  - Use FDF, BDF, CDF or other derivative estimates for two different *h* to compute a more accurate approximation
    - Richardson extrapolation

Richardson extrapolation

- The Taylor expansions (assuming that higher derivatives exist):

$$f(x_0+\frac{h}{2})=f(x_0)+f'(x_0)\frac{h}{2}+f''(x_0)\frac{h^2}{8}+f'''(x_0)\frac{h^3}{48}+f^{(4)}(x_0)\frac{h^4}{384}+f^{(5)}(x_0)\frac{h^5}{3840}+\cdots$$

$$f(x_0-\frac{h}{2})=f(x_0)-f'(x_0)\frac{h}{2}+f''(x_0)\frac{h^2}{8}-f'''(x_0)\frac{h^3}{48}+f^{(4)}(x_0)\frac{h^4}{384}-f^{(5)}(x_0)\frac{h^5}{3840}+\cdots$$

- if we subtract these two expressions, we get

$$f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2})=f'(x_0)h+\frac{f'''(x_0)}{24}h^3+f^{(5)}(x_0)\frac{h^5}{1920}+\cdots$$

- Therefore

$$f'(x_0)=\frac{f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2})}{h}-\frac{f'''(x_0)}{24}h^2-f^{(5)}(x_0)\frac{h^4}{1920}-\cdots$$

$$\Downarrow$$

$$f'(x_0)=\underbrace{\frac{f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2})}{h}}_{\varphi(h)}-a_2h^2-a_4h^4-\boxed{R}$$

$\varphi(h)$ ← CDF

Remainder

Richardson extrapolation

- Idea: get rid of

$$f'(x_0) = \varphi(h) - a_2 h^2 - a_4 h^4 - R(h) \qquad \mathbf{+}$$

- For *h/2* we can write

$$f'(x_0) = \varphi\left(\frac{h}{2}\right) - a_2\left(\frac{h}{2}\right)^2 - a_4\left(\frac{h}{2}\right)^4 - R\left(\frac{h}{2}\right) \qquad \mathbf{X\ -4}$$

- We obtain

$$-3f'(x_0) = \varphi\left(h\right) - 4\varphi\left(\frac{h}{2}\right) - \frac{3}{4}a_4 h^4 + 4R\left(\frac{h}{2}\right) - R(h)$$

- wherefrom we get

$$f'(x_0) = \frac{4}{3}\varphi\left(\frac{h}{2}\right) - \frac{1}{3}\varphi\left(h\right) + O(h^4)$$

- So 2x smaller *h* -> the error becomes 16x smaller

Richardson extrapolation:example

- Numerical test for $f(x)=e^{(5*x)}$ at $x_0 = 0.2$

h	Error (FDF)	Error (CDF)	Error (Richardson)
1.0e-01	4.04266328	0.14202027	-1.1081e-04
1.0e-02	0.34551982	0.00141582	-1.1061e-08
1.0e-03	0.03403522	0.00001416	-7.5318e-13
1.0e-04	0.00339842	0.00000014	1.1678e-11
1.0e-05	0.00033979	0.00000000	-2.1037e-10
1.0e-06	0.00003398	0.00000000	1.1811e-09
1.0e-07	0.00000340	0.00000000	-2.1171e-08
1.0e-08	0.00000037	0.00000003	1.4462e-07
1.0e-09	0.00000026	0.00000063	-5.5112e-07
1.0e-10	0.00000114	0.00000330	3.2977e-06
1.0e-11	0.00001447	0.00001447	4.4746e-05
1.0e-12	0.00050296	0.00050296	-5.0296e-04

Richardson extrapolation

- Idea: get rid of

$$f'(x_0) = \varphi(h) - a_2 h^2 - a_4 h^4 - R(h) \quad \mathbf{+}$$

- For *h/2* we can write

$$f'(x_0) = \varphi\left(\frac{h}{2}\right) - a_2\left(\frac{h}{2}\right)^2 - a_4\left(\frac{h}{2}\right)^4 - R\left(\frac{h}{2}\right) \quad \mathbf{X\ {-4}}$$

- We obtain

$$-3f'(x_0) = \varphi(h) - 4\varphi\left(\frac{h}{2}\right) - \frac{3}{4}a_4 h^4 + 4R\left(\frac{h}{2}\right) - R(h)$$

- i.e.

$$f'(x_0) = \frac{4}{3}\varphi\left(\frac{h}{2}\right) - \frac{1}{3}\varphi(h) + O(h^4)$$

- So 10x smaller *h* -> the error becomes 10000x smaller

- Is it possible to cancel out the *h^4* term?

Richardson extrapolation table

$D(0,0)=\varphi(h)$	
$D(1,0)=\varphi(h/2)$	$D(1,1)$

- Recall that

$$f'(x_0)=\frac{4}{3}\varphi\left(\frac{h}{2}\right)-\frac{1}{3}\varphi\left(h\right)+O(h^4)\Rightarrow f'(x_0)=\underbrace{\varphi\left(\frac{h}{2}\right)+\frac{1}{3}\left(\varphi\left(\frac{h}{2}\right)-\varphi\left(h\right)\right)}+O(h^4)$$

$$\varphi(h)=\frac{f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2})}{h}$$

$$D(1,1)=D(1,0)+\frac{1}{4^1-1}\left(D(1,0)-D(0,0)\right)$$

Richardson extrapolation table

$D(0,0)=\varphi(h)$	
$D(1,0)=\varphi(h/2)$	$D(1,1)$
$D(2,0)=\varphi(h/4)$	$D(2,1)$

- Recall that

$$f'(x_0)=\frac{4}{3}\varphi\left(\frac{h}{2}\right)-\frac{1}{3}\varphi\left(h\right)+O(h^4)\Rightarrow f'(x_0)=\underbrace{\varphi\left(\frac{h}{2}\right)+\frac{1}{3}\left(\varphi\left(\frac{h}{2}\right)-\varphi\left(h\right)\right)}+O(h^4)$$

$$\varphi(h)=\frac{f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2})}{h}$$

$$D(1,1)=D(1,0)+\frac{1}{4^1-1}\left(D(1,0)-D(0,0)\right)$$

Richardson extrapolation table

$D(0,0)=\varphi(h)$			
$D(1,0)=\varphi(h/2)$	$D(1,1)$		
$D(2,0)=\varphi(h/4)$	$D(2,1)$	$D(2,2)$	

- Recall that

$$f'(x_0)=\frac{4}{3}\varphi\left(\frac{h}{2}\right)-\frac{1}{3}\varphi\left(h\right)+O(h^4)\Rightarrow f'(x_0)=\varphi\left(\frac{h}{2}\right)+\frac{1}{3}\left(\varphi\left(\frac{h}{2}\right)-\varphi\left(h\right)\right)+O(h^4)$$

$$\varphi(h)=\frac{f(x_0+\frac{h}{2})-f(x_0-\frac{h}{2})}{h}$$

$$D(1,1)=D(1,0)+\frac{1}{4^1-1}\left(D(1,0)-D(0,0)\right)$$

Richardson extrapolation table

$m \downarrow$

$D(0,0)=\varphi(h)$			
$D(1,0)=\varphi(h/2)$	$D(1,1)$		
$D(2,0)=\varphi(h/4)$	$D(2,1)$	$D(2,2)$	
$D(3,0)=\varphi(h/8)$	$D(3,1)$	$D(3,2)$	$D(3,3)$

$n \longrightarrow$

$$\boxed{D\left(n,m\right)=D(n,m-1)+\frac{1}{4^m-1}\left[D\left(n,m-1\right)-D\left(n-1,m-1\right)\right]}$$