CSE 541 - Differentiation

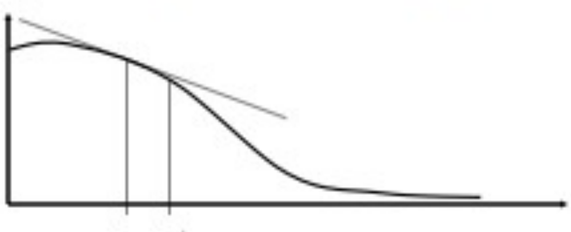
Roger Crawfis

Numerical Differentiation

- The mathematical definition:

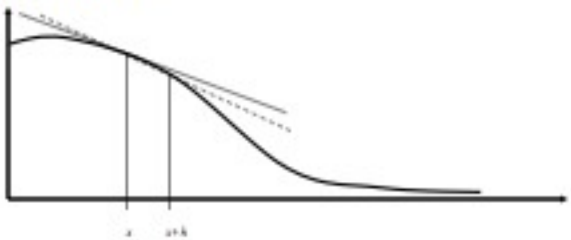
$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

- Can also be thought of as the tangent line.



Numerical Differentiation

- We can not calculate the limit as h goes to zero, so we need to approximate it.
- Apply directly for a non-zero h leads to the slope of the secant curve.



Numerical Differentiation

- This is called **Forward Differences** and can be derived using Taylor's Series:

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

$$\therefore f(x+h) - f(x) = f'(x)h + f''(\xi)\frac{h^2}{2!}$$

$$\therefore f'(x) = \frac{f(x+h) - f(x)}{h} - f''(\xi)\frac{h}{2!}$$

$$\therefore \frac{f(x+h) - f(x)}{h} \to f'(x) \text{ as } h \to 0$$

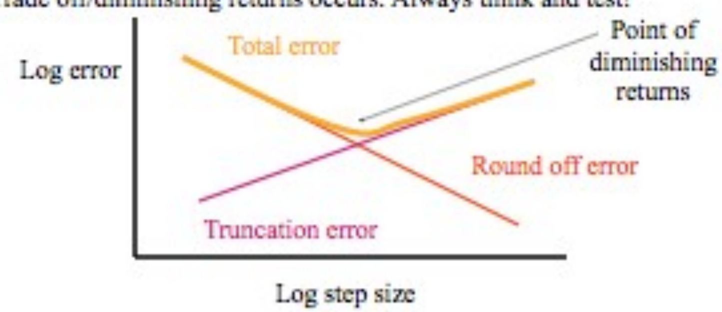
Theoretically speaking

Truncation Errors

- Let $f(x)$ = a+e, and $f(x+h)$ = a+f.
- Then, as h approaches zero, e<<a and f<<a.
- With limited precision on our computer, our representation of $f(x) \approx a \approx f(x+h)$.
- We can easily get a random round-off bit as the most significant digit in the subtraction.
- Dividing by h, leads to a very wrong answer for $f'(x)$.

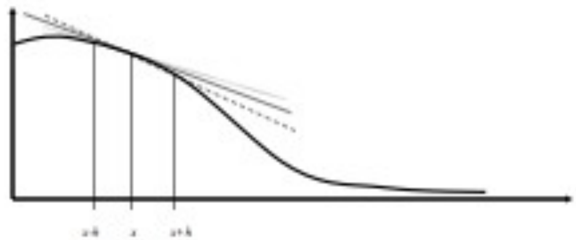
Error Tradeoff

- Using a smaller step size reduces truncation error.
- However, it increases the round-off error.
- Trade off/diminishing returns occurs: Always think and test!



Numerical Differentiation

- This formula favors (or biases towards) the right-hand side of the curve.
- Why not use the left?



Numerical Differentiation

- This leads to the **Backward Differences** formula.

$$f(x-h) = f(x) - f'(x)h + f''(\xi)\frac{h^2}{2!}$$

$$\therefore f'(x) = \frac{f(x) - f(x-h)}{h} + f''(\xi)\frac{h}{2!}$$

$$\therefore \frac{f(x) - f(x-h)}{h} \rightarrow f'(x) \ as \ h \rightarrow 0$$

Numerical Differentiation

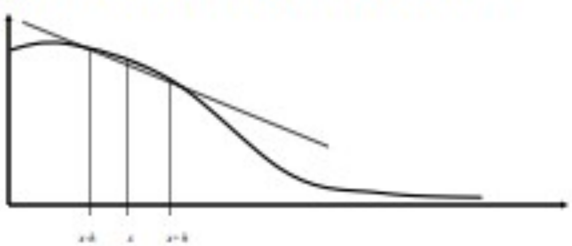
- Can we do better?
- Let's average the two:

$$f'(x) = \frac{1}{2}\left(\underbrace{\frac{f(x+h)-f(x)}{h}}_{\textit{Forward difference}} + \underbrace{\frac{f(x)-f(x-h)}{h}}_{\textit{Backward difference}}\right) = \frac{f(x+h)-f(x-h)}{2h}$$

- This is called the **Central Difference** formula.

Central Differences

- This formula does not *seem* very good.
  - It does not follow the calculus formula.
  - It takes the slope of the secant with width 2h.
  - The actual point we are interested in is not even evaluated.



Numerical Differentiation

- Is this any better?
- Let's use Taylor's Series to examine the error:

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2} + f'''(\xi)\frac{h^3}{3!}$$

$$f(x-h) = f(x) - f'(x)h + f''(x)\frac{h^2}{2} - f'''(\varsigma)\frac{h^3}{3!}$$

subtracting

$$f(x+h) - f(x-h) = 2f'(x)h + \left(f'''(\xi)\frac{h^3}{3!} + f'''(\varsigma)\frac{h^3}{3!}\right)$$

Central Differences

- The central differences formula has much better convergence.

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{1}{6}f'''(\zeta)h^2, \zeta \in [x-h, x+h]$$

- Approaches the derivative as h^2 goes to zero!!

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

Warning

- Still have truncation error problem.
- Consider the case of:
- Build a table with smaller values of *h*.
- What about large values of *h* for this function?

$$f(x)=\frac{x}{100}$$

$$f'(x);\ \frac{\left\lfloor\frac{x+h}{100}\right\rfloor-\left\lfloor\frac{x-h}{100}\right\rfloor}{2h}$$

at $x=1, h=0.000333$, *with 6 significant digits*

$$f'(x);\ \frac{0.0100033-0.0099966}{0.000666666}=0.010050$$

Relative error:

$$\frac{|0.01-0.010050|}{0.01}=0.5\%$$

Richardson Extrapolation

- Can we do better?
- Is my choice of *h* a good one?
- Let's subtract the two Taylor Series expansions again:

$$f(x+h)=f(x)+f'(x)h+f''(x)\frac{h^2}{2}+f'''(x)\frac{h^3}{3!}+f^{(4)}(x)\frac{h^4}{4!}+f^{(5)}(x)\frac{h^5}{5!}+\cdots$$

$$f(x-h)=f(x)-f'(x)h+f''(x)\frac{h^2}{2}-f'''(\varsigma)\frac{h^3}{3!}+f^{(4)}(x)\frac{h^4}{4!}-f^{(5)}(x)\frac{h^5}{5!}+\cdots$$

subtracting

$$f(x+h)-f(x-h)=2f'(x)h+2\frac{f'''(x)}{3!}h^3+2\frac{f'''(x)}{3!}h^3+2f^{(5)}(x)\frac{h^5}{5!}+\cdots$$

Richardson Extrapolation

- Assuming the higher derivatives exist, we can hold x fixed (which also fixes the values of $f(x)$), to obtain the following formula.

$$f'(x) = \frac{1}{2h}\left[f(x+h) - f(x-h)\right] + a_2h^2 + a_4h^4 + a_6h^6 + \cdots$$

- Richardson Extrapolation examines the operator below as a function of h.

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

Richardson Extrapolation

- This function approximates $f'(x)$ to $O(h^2)$ as we saw earlier.
- Let's look at the operator as h goes to zero.

$$\varphi(h) = f'(x) - a_2h^2 - a_4h^4 - a_6h^6 - \cdots$$

$$\varphi\left(\frac{h}{2}\right) = f'(x) - a_2\left(\frac{h}{2}\right)^2 - a_4\left(\frac{h}{2}\right)^4 - a_6\left(\frac{h}{2}\right)^6 - \cdots$$

Same leading constants

Richardson Extrapolation

- Using these two formula's, we can come up with another estimate for the derivative that cancels out the h^2 terms.

$$\varphi(h) - 4\varphi\left(\frac{h}{2}\right) = -3f'(x) - \frac{3}{4}a_4h^4 - \frac{15}{16}a_6h^6 - \cdots$$

or

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

new estimate — *difference between old and new estimates*

Extrapolates by assuming the new estimate undershot.

Richardson Extrapolation

- If h is small (h<<1), then h^4 goes to zero much faster than h^2.
- Cool!!!
- Can we cancel out the h^6 term?
  - Yes, by using h/4 to estimate the derivative.

Richardson Extrapolation

- Consider the following *property*:

$$\varphi(h) = f'(x) - \sum_{k=1}^{\infty} a_{2k} h^{2k}$$
$$= L - \sum_{k=1}^{\infty} a_{2k} h^{2k}$$

- where L is unknown,

$$L = \lim_{h \to 0} \varphi(h) = f'(x)$$

- as are the coefficients, a_{2k}.

Richardson Extrapolation

- Do not forget the formal definition is simply the central-differences formula:

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

- New symbology *(is this a word?)*:

$$D(n,0) = \varphi\left(\frac{h}{2^n}\right)$$
$$= L + \sum_{k=1}^{\infty} A(k,0)\left(\frac{h}{2^n}\right)^{2k}$$

From previous slide

Richardson Extrapolation

- D(n,0) is just the central differences operator for different values of *h*.
- Okay, so we proceed by computing D(*n*,0) for several values of *n*.
- Recalling our cancellation of the h^2 term.

$$f'(x) = \varphi(\frac{h}{2}) + \frac{1}{3}\left[\varphi(\frac{h}{2}) - \varphi(h)\right] + O\left(h^4\right)$$
$$= D(1,0) + \frac{1}{4-1}\left[D(1,0) - D(0,0)\right] + O\left(h^4\right)$$

Richardson Extrapolation

- If we let $h \to h/2$, then in general, we can write:

$$f'(x) = D(n,0) + \frac{1}{4-1}\left[D(n,0) - D(n-1,0)\right] + O\left(\left(\frac{h}{2^n}\right)^4\right)$$

- Let's denote this operator as:

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

Richardson Extrapolation

- Now, we can formally define Richardson's extrapolation operator as:

$$D(n,m) = \frac{4^m}{4^m-1} D(n,m-1) - \frac{1}{4^m-1} D(n-1,m-1), \quad (1 \le m \le n)$$

new estimate — *old estimate*

- or

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

Richardson Extrapolation

- Now, we can formally define Richardson's extrapolation operator as:

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

Memorize me!!!!

Richardson Extrapolation Theorem

- These terms approach *f'(x)* very quickly.

$$D(n,m) = L + \sum_{k=m+1}^{\infty} A(k,m)\left(\frac{h}{2^n}\right)^{2k}$$

Order starts much higher!!!!

Richardson Extrapolation

- Since $m \leq n$, this leads to a two-dimensional triangular array of values as follows:

$$\begin{array}{lllll} D(0,0) & & & & \\ D(1,0) & D(1,1) & & & \\ D(2,0) & D(2,1) & D(2,2) & & \\ \vdots & \vdots & \vdots & \ddots & \\ D(N,0) & D(N,1) & D(N,2) & \cdots & D(N,N) \end{array}$$

- We must pick an initial value of *h* and a max iteration value *N*.

Example

$$f(x)=\frac{\left(\cos(100x^2)\right)^5}{x^3}$$

$$x=1.3,\ h=\frac{1}{128}$$

$$D(0,0)=16.696386$$
$$D(1,0)=40.583393$$
$$D(2,0)=109.322528$$
$$D(3,0)=135.031747$$
$$D(4,0)=142.068615$$
$$D(5,0)=143.866937$$

$$h=\frac{1}{4096}=0.00244$$

Example

$$D(0,0)=16.696386$$
$$D(1,0)=40.583393 \quad D(1,1)=48.583393$$
$$D(2,0)=109.322528 \quad D(2,1)=132.235574$$
$$D(3,0)=135.031747 \quad D(3,1)=143.601487$$
$$D(4,0)=142.068615 \quad D(4,1)=144.414238$$
$$D(5,0)=143.866937 \quad D(5,1)=144.466377$$

$$\textit{extrapolate}\ \frac{1}{3}$$

Example

$D(0,0)=16.696386$

$D(1,0)=40.583393 \quad D(1,1)=48.583393$

$D(2,0)=109.322528 \quad D(2,1)=132.235574 \quad D(2,2)=137.814897$

$D(3,0)=135.031747 \quad D(3,1)=143.601487 \quad D(3,2)=144.359214$

$D(4,0)=142.068615 \quad D(4,1)=144.414238 \quad D(4,2)=144.468421$

$D(5,0)=143.866937 \quad D(5,1)=144.466377 \quad D(5,2)=144.469853$

$\text{extrapolate}\,\frac{1}{15}$

Example

16.696386					
40.583393	48.583393				
109.322528	132.235574	137.814897			
135.031747	143.601487	144.359214	144.463092		
142.068615	144.414238	144.468421	144.470154	144.470182	
143.866937	144.466377	144.469853	144.469876	144.469875	$D(5,5)=144.469875$
	$\text{extrapolate}\,\frac{1}{3}$	$\text{extrapolate}\,\frac{1}{15}$	$\text{extrapolate}\,\frac{1}{63}$	$\text{extrapolate}\,\frac{1}{255}$	$\text{extrapolate}\,\frac{1}{1023}$

- Which converges up to eight decimal places.
- Is it accurate?

Example

- We can look at the (theoretical) error term on this example.

$$D(5,5) = L + \sum_{k=5+1}^{\infty} A(k,5)\left(\frac{h}{2^5}\right)^{2k}$$
$$= f'(1.3) + A(6,5)\left(\frac{1}{4096}\right)^{12} + \sum_{k=7}^{\infty} A(k,5)\left(\frac{h}{2^5}\right)^{2k}$$

2^{-144}

- Taking the derivative:

$$f'(1.3) = 144.469874253\ldots$$

Round-off error

Second Derivatives

- What if we need the second derivative?

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2} + f'''(x)\frac{h^3}{3!} + f^{(4)}(x)\frac{h^4}{4!} + f^{(5)}(x)\frac{h^5}{5!} + \cdots$$
$$f(x-h) = f(x) - f'(x)h + f''(x)\frac{h^2}{2} - f'''(\xi)\frac{h^3}{3!} + f^{(4)}(x)\frac{h^4}{4!} - f^{(5)}(x)\frac{h^5}{5!} + \cdots$$

- Any guesses?

Second Derivatives

- Let's cancel out the odd derivatives and double up the even ones:
  - Implies adding the terms together.

$$f(x+h)+f(x-h)=2f(x)+2f''(x)\frac{h^2}{2}+2f^{(4)}(x)\frac{h^4}{4!}+\cdots$$

Second Derivatives

- Isolating the second derivative term yields:

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

- With an error term of:

$$E=-\frac{1}{12}h^2 f^{(4)}(\xi)$$

Partial Derivatives

- Remember: Nothing special about partial derivatives:

$$\frac{\partial f}{\partial x}(x,y) \approx \frac{f(x+h,y)-f(x-h,y)}{2h}$$
$$\frac{\partial f}{\partial y}(x,y) \approx \frac{f(x,y+h)-f(x,y-h)}{2h}$$

Calculating the Gradient

- For lab 2, you need to calculate the gradient.
- Just use central differences for each partial derivative.
- Remember to normalize it (divide by its length).