Other finite difference formulas

Friday, September 25, 2020 1:39 PM

We saw $\dfrac{f(x_0+h)-f(x_0)}{h}$. Let's work w/ many nodes on an equispaced grid,

$$x_0,\ \underbrace{x_0+h}_{x_1},\ \underbrace{x_0+2h}_{x_2},\ \ldots,\ \underbrace{x_0+nh}_{x_n}$$

Generic finite-difference formula is

$$f'(x_0) \approx \frac{1}{h}\sum_{k=-p}^{q} a_k \cdot f(x_0 + k\cdot h)$$

(that's for x_0, but the weights are translation invariant, so same formula works for x_1, x_2, etc provided indices don't go out-of-bounds)

Ex: more forward finite diff. formulas

Accuracy	0	h	2h	3h	4h
$O(h)$	-1	1			
$O(h^2)$	-3/2	2	-1/2		
$O(h^3)$	-11/6	3	-3/2	1/3	
$O(h^4)$	-25/12	4	-3	4/3	-1/4

For example this means

$$f'(x_0) \approx \frac{-\frac{1}{2}f(x_0+2h) + 2f(x_0+h) - \frac{3}{2}f(x_0)}{h}$$

often written as $\frac{-25}{12}, \frac{48}{12}, \frac{-36}{12}, \frac{16}{12}, \frac{-3}{12}$

more terms = bad (more work, more chances for cancellation)

high-order = good

The forward formulas are good if you are at a left endpoint (and equivalent backward formulas are good at right endpoint x_n, just note that now h is negative)

But for interior nodes, we can do better by using symmetric "centered difference" formulas

Ex

Order		-4h	-3h	-2h	-h	0	h	2h	3h	4h
2	"3 pt."				-1/2	0	1/2			
4	"5 pt."			1/12	-2/3	0	2/3	-1/12		
6	"7 pt."		-1/60	3/20	-3/4	0	3/4	-3/20	1/60	
8		1/280	-4/105	1/5	-4/5	0	4/5	-1/5	4/105	-1/280

Nice! ⚠ $O(h^2)$ for just 2 function evaluations.

Ex: Show the 3-pt forward diff. method is $O(h^2)$

$$\frac{(-3f(x)\,{}^{A} + 4f(x+h)\,{}^{B} - f(x+2h)\,{}^{C})}{2h}\ {}^{D}$$

⚠ $O(h^2) - O(h^2)$ is NOT 0, it is $O(h^2)$

Sol'n: Taylor expand

$$\begin{aligned}
&-3f(x) \\
&+4\left(f(x) + h f'(x) + \tfrac{h^2}{2} f''(x) + O(h^3)\right) \\
&-\left(f(x) + 2h f'(x) + \tfrac{(2h)^2}{2} f''(x) + O(h^3)\right) \\
\hline
&(\cancel{-3+4-1})^{0} f(x) + (4-2)h f'(x) + \cancel{\left(\tfrac{4}{2} - \tfrac{2^2}{2}\right)h^2 f''(x)} + O(h^3)
\end{aligned}$$

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

Higher-order derivatives

We can also approximate $f''(x)$ (and you'll see a $\frac{1}{h^2}$ instead of $\frac{1}{h}$)

Ex. 2nd derivative midpoint formula is

$$\begin{array}{ccc} -h & 0 & h \\ \hline 1 & -2 & 1 \end{array} \Bigg\} \text{ the "stencil"}$$

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

Show this is $O(h^2)$:

when do you stop? Guess. It can't hurt to have extra terms

$$\begin{aligned}
f(x-h) &= \cancel{f(x)} - \cancel{h f'(x)} + \tfrac{h^2}{2} f''(x) - \cancel{\tfrac{h^3}{6} f'''(x)} + O(h^4) \\
-2f(x) &= -2\cancel{f(x)} \\
f(x+h) &= \cancel{f(x)} + \cancel{h f'(x)} + \tfrac{h^2}{2} f''(x) + \cancel{\tfrac{h^3}{6} f'''(x)} + O(h^4) \\
\hline
\end{aligned}$$

$$= \frac{h^2 f''(x) + O(h^4)}{h^2} = f''(x) + \underline{\underline{O(h^2)}}$$

Stability and roundoff

When we say the <u>forward finite-diff</u> (2 pts) is $O(h)$, this "$O(h)$" is the "truncation" / "approximation" error., often written $\tau_f(h)$

But we also have roundoff error due to finite precision.

Work w/ imprecise input $\tilde{x}$ aka $fl(x)$, and assume

$$\tilde{x} = (1+\varepsilon)x \quad \text{for } |\varepsilon| \le \varepsilon_{\text{machine precision}} \approx 10^{-16}$$

Can we write $x = (1+\varepsilon)\tilde{x}$ too?

Yes: Start w/ $\tilde{x} = (1+\varepsilon)x$

So $x = \frac{1}{1+\varepsilon}\tilde{x}$ $\underbrace{}_{\text{Taylor Exp.}}$

$= \tilde{x} - \varepsilon\tilde{x} + O(\varepsilon^2)$

$\approx (1-\varepsilon)\tilde{x}$

Let $\delta = \frac{f(x+h) - f(x)}{h}$ so $f'(x) = \delta + \underbrace{O(h)}_{\tau_f(h)}$

but we have roundoff error when computing δ,

let $\tilde{\delta}$ be actual floating point version.

Recall our relative condition number κ_f means (if ε is small) $\quad \left|\dfrac{f(\tilde{x}) - f(x)}{f(x)}\right| \approx \kappa_f(x)\cdot|\varepsilon|$ if $\tilde{x} = (1+\varepsilon)x$

ignore roundoff error via division.

Why? If $g(x) = \dfrac{c\cdot x}{(e.g., x/h)}$, $\kappa_g(x) = 1$

So

$$fl(\delta) = \frac{f(\,fl(x+h)\,) - f(\,fl(x)\,)}{h}$$

$$\approx \frac{\left(f(x+h) + \overset{\approx \kappa_f(x)}{\kappa_f(x+h)}\cdot f(x+h)|\varepsilon|\right) - \left(f(x) + \kappa_f(x)\cdot f(x)|\varepsilon|\right)}{h}$$

these don't cancel since it's really $O(|\varepsilon|)$

$$= \underbrace{\frac{f(x+h) - f(x)}{h}}_{\delta} + \frac{\kappa_f(x)\cdot f(x)}{h}\, O(|\varepsilon|)$$

So, $\underbrace{|f'(x) - fl(\delta)|}_{\text{what computer actually does}} \le \underbrace{|f'(x) - \delta|}_{\substack{\tau_f(h) \\ \text{truncation}}} + \underbrace{|fl(\delta) - \delta|}_{\approx \kappa_f(x)\cdot f(x)\cdot \frac{\varepsilon_{machine}}{h}}$

Usually $O(h)$, $O(h^2)$, etc.

So $|\text{error}| \le O(h^{k}) + \dfrac{\overbrace{\kappa_f(x) f(x)}^{c}\cdot \varepsilon_{machine}}{h}$ (ex: $k=1$)

We get to choose h, so choose h to minimize this bound (usually $\kappa_f(x)\cdot f(x)$ is unknown, so guess ≈ 1 ... one reason it's usually nice to scale things)

If $\text{error} = \underbrace{h + \dfrac{c\cdot\varepsilon}{h}}_{g(h)}$, what's best h?

$\min\limits_{h,\ h>0} g(h)$ by ① check endpoints $(h=0)$
② set $g'(h) \overset{\text{set}}{=} 0$

so $g'(h) = 1 - \dfrac{c\cdot\varepsilon}{h^2} = 0$, $\underline{\underline{h = \sqrt{c\cdot\varepsilon}}}$ so error $= \sqrt{c\cdot\varepsilon} + \dfrac{c\cdot\varepsilon}{\sqrt{c\cdot\varepsilon}}$

$= \underline{\underline{2\sqrt{c\cdot\varepsilon}}}$.

"Walker, Pernice" 1998 rule-of-thumb here is
$h = \sqrt{\varepsilon}\cdot(1 + |x|)$

If error $= \underbrace{h^{2} + \dfrac{c\cdot\varepsilon}{h}}_{g(h)}$, now $g'(h) = h - \dfrac{c\cdot\varepsilon}{h^2}$, $h^3 = c\cdot\varepsilon$,
$h = (c\cdot\varepsilon)^{1/3}$

and error is $(c\varepsilon)^{2/3} + \dfrac{c\cdot\varepsilon}{(c\varepsilon)^{1/3}} = \underline{\underline{2\cdot(c\varepsilon)^{2/3}}}$

To summarize:

2-pt forward diff ($O(h)$), total error (truncation + roundoff) $\approx \sqrt{\varepsilon_{mach}}$.

3pt centered diff ($O(h^2)$), total error (" ______ ") $\approx \varepsilon^{2/3}$

If $\varepsilon = 10^{-16}$, 2 pt. forward, error $\approx 10^{-8}$
3 pt. centered, error $\approx 10^{-10.67}$. Better

Rule of thumb

Higher order methods have overall (truncation + roundoff) better accuracy

Why not use 1000 pt. formulas then?

- must deal w/ many special cases due to boundaries and programming / book-keeping / speed isn't worth it
- more computation, and at some point improving from 10^{-15} to $10^{-15.5}$ rel. error isn't worth it

If we have

$f(x_0), f(x_1), f(x_2), f(x_3)$ $\quad (x_1 = x_0 + h,\ x_2 = x_0 + 2h, \dots)$

then I can use 3 pt. centered diff. here

but for points, use forward or backward.

I can't use a 5 pt. centered diff. formula for any of these points