۴.۴ مشتق مرتبه‌ی دوم

بار دیگر از چندجمله‌ای درونیاب استفاده می‌کنیم: $f''(x) \approx p''(x)$. در فرمول درونیابی با تفاضلات پیشرو، چندجمله‌ایِ p به صورت تابعی برحسب s بیان می‌شود و $s = \frac{x-x_i}{h}$ نیز خود تابعی برحسب x است یعنی داریم $p = p(s(x))$. به یاد آورید که برطبق قاعده‌ی زنجیره‌ای برای مشتق دوم داریم[۱]:

$$\frac{d^2p}{dx^2} = \frac{d^2p}{ds^2}(\frac{ds}{dx})^2 + \frac{dp}{ds}\frac{d^2s}{dx^2}.$$

از سوی دیگر در (۱.۴) دیدیم که

$$\frac{dp}{ds} = \Delta f_i + \frac{2s-1}{2!}\Delta^2 f_i + \frac{3s^2-6s+2}{3!}\Delta^3 f_i + \cdots .$$

پس داریم:

$$\frac{d^2p}{ds^2} = \frac{2}{2!}\Delta^2 f_i + \frac{6s-6}{3!}\Delta^3 f_i + \frac{12s^2-36s+22}{4!}\Delta^4 f_i + \cdots .$$

همچنین واضح است که $(\frac{ds}{dx})^2 = \frac{1}{h^2}$ و $\frac{d^2s}{dx^2} = 0$. پس داریم:

$$f''(x) \approx \frac{d^2p}{dx^2} = \frac{1}{h^2}\Big(\Delta^2 f_i + (s-1)\Delta^3 f_i + (\frac{s^2}{2} - \frac{3}{2}s + \frac{11}{12})\Delta^4 f_i + \cdots\Big).$$

در نتیجه به ازای مقدار خاص $s = 0$ یعنی به ازای $x = x_i$ داریم:

$$f_i'' := f''(x_i) \approx \frac{1}{h^2}\Big(\Delta^2 f_i - \Delta^3 f_i + \frac{11}{12}\Delta^4 f_i + \cdots\Big). \quad (۷.۴)$$

[۱] برای $p = p(s(x))$ داریم $\frac{dp}{dx} = \frac{dp}{ds}\frac{ds}{dx}$. پس

$$\begin{aligned}\frac{d^2p}{dx^2} &= \frac{d}{dx}(\frac{dp}{dx}) = \frac{d}{dx}(\frac{dp}{ds}\frac{ds}{dx}) = \Big(\frac{d}{dx}(\frac{dp}{ds})\Big)\frac{ds}{dx} + \frac{dp}{ds}\Big(\frac{d}{dx}(\frac{ds}{dx})\Big)\\ &= \Big(\frac{d}{ds}(\frac{dp}{ds})\Big)\frac{ds}{dx}\frac{ds}{dx} + \frac{dp}{ds}(\frac{d^2s}{dx^2}) = \frac{d^2p}{ds^2}(\frac{ds}{dx})^2 + \frac{dp}{ds}\frac{d^2s}{dx^2}.\end{aligned}$$

با استفاده از تنها یک جمله از سمت راست فرمول بالا داریم:

$$f_i'' \approx \frac{\Delta^2 f_i}{h^2} = \frac{f_{i+2} - 2f_{i+1} + f_i}{h^2}$$

که فرمول سه نقطه‌ای (پیشروی) مشتق دوم نام دارد. همچنین با استفاده از دو جمله از (۷.۴) فرمول چهارنقطه‌ای (پیشروی) مشتق دوم به صورت زیر به دست می‌آید:

$$f_i'' \approx \frac{\Delta^2 f_i - \Delta^3 f_i}{h^2} = \frac{-f_{i+3} + 4f_{i+2} - 5f_{i+1} + 2f_i}{h^2}.$$

مثال ۱.۴.۴. به کمک دو فرمول قبل، مقدار $f''(1.5)$ را با استفاده از داده‌های زیر به دست آورید:

x_i	1.3	1.5	1.7	1.9	2.1
f_i	3.669	4.482	5.474	6.686	8.166

بر طبق فرمول سه‌نقطه‌ای داریم

$$f''(1.5) \approx \frac{f(1.9) - 2f(1.7) + f(1.5)}{(0.2)^2} = 5.500.$$

اما طبق فرمول چهارنقطه‌ای داریم

$$f''(1.5) \approx \frac{-f(2.1) + 4f(1.9) - 5f(1.7) + 2f(1.5)}{(0.2)^2} = 4.300.$$

با توجه به اینکه گفتیم داده‌های این مثال خاص مربوط به تابع نمایی است (توضیحات پایانی مثال ۱.۲.۴ را بینید) و طبق جدول داریم $f(1.5) = 4.482$ پس واضح است که پاسخ فرمول چهارنقطه‌ای درست‌تر از پاسخی می‌باشد که با فرمول سه نقطه‌ای به دست آمده است. □

۵.۴ برخی فرمول‌های دیگر به کمک بسط تیلور

تاکنون فرمول‌های مشتق‌گیری عددی را با استفاده از چندجمله‌ای درونیاب به دست آوردیم و از سری تیلور تنها برای تعیین میزان خطای هر فرمول استفاده کردیم. اما می‌توان از سری تیلور مستقیما برای ساختن فرمول‌های مشتق‌گیری عددی نیز استفاده کرد که در این صورت می‌توان میزان خطای فرمول را نیز همزمان

با ساخته‌شدن فرمول مشخص کرد. بعنوان نمونه چنانچه f به اندازه‌ی کافی هموار بوده، نقاط x_i برای $i = 0, 1, \cdots, n$ هم‌فاصله بوده و $0 < h = x_{i+1} - x_i$، طبق بسط تیلور وجود دارد $x_i < \zeta < x_{i+1}$ به طوری که

$$f_{i+1} = f(x_{i+1}) = f(x_i + h) = f_i + hf_i' + \frac{h^2}{2!}f_i'' + \frac{h^3}{3!}f'''(\zeta)$$

و یا

$$f_{i+1} = f(x_{i+1}) = f(x_i + h) = f_i + hf_i' + \frac{h^2}{2!}f_i'' + \mathcal{O}(h^3).$$

از سوی دیگر

$$f_{i-1} = f(x_{i-1}) = f(x_i - h) = f_i - hf_i' + \frac{h^2}{2!}f_i'' + \mathcal{O}(h^3).$$

از تفریق دو رابطه‌ی بالا داریم

$$f_{i+1} - f_{i-1} = 2hf_i' + \mathcal{O}(h^3).$$

پس

$$f_i' = \frac{f_{i+1} - f_{i-1}}{2h} + \mathcal{O}(h^2).$$

این بدان معناست که با استفاده از فرمول تقریبی

$$f_i' \approx \frac{f_{i+1} - f_{i-1}}{2h} \tag{۸.۴}$$

خطای برشی $\mathcal{O}(h^2)$ حاصل می‌شود. این، یک فرمول ((تفاضلات مرکزی)) است به این دلیل که برای یافتنِ مشتق در نقطه‌ی x_i از مقدار تابع در نقاطی در هر دو سمت x_i استفاده می‌کند.

بعنوان نمونه‌ای دیگر از کاربرد سری تیلور در ساخت فرمول‌های مشتق‌گیری عددی به همراه خطای برشی داریم

$$\begin{aligned} f_{i+1} &= f_i + hf_i' + \frac{h^2}{2!}f_i'' + \frac{h^3}{3!}f_i''' + \mathcal{O}(h^4), \\ f_{i-1} &= f_i - hf_i' + \frac{h^2}{2!}f_i'' - \frac{h^3}{3!}f_i''' + \mathcal{O}(h^4), \end{aligned}$$

پس

$$f_{i+1} + f_{i-1} = 2f_i + h^2 f_i'' + \mathcal{O}(h^4).$$

بنابراین فرمول تقریبی

$$f_i'' \approx \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2}$$

دارای خطای برشی $\mathcal{O}(h^2)$ می‌باشد. این رابطه نیز (مانند فرمول قبل و تمرین بعدی) یک فرمول ((تفاضلات مرکزی)) است.

تمرین ۰.۱۱ رابطه‌ی زیر را برای مشتق چهارم تابع f در نقطه‌ی x_i به دست آورده و

$$f_i^{(4)} = \frac{f_{i+2} - 4f_{i+1} + 6f_i - 4f_{i-1} + f_{i-2}}{h^4}.$$

ثابت کنید خطای تقریب آن از مرتبه‌ی h^2 است.

۶.۴ فرمول‌های مشتق‌گیری عددی در عمل

تاکنون با برخی از فرمول‌های تفاضلات متناهی برای مشتق‌گیری عددی و خطای برشی آن‌ها آشنا شده‌ایم. حال می‌خواهیم خصوصیات پایداری عددی این الگوریتم‌ها را به طور مختصر بررسی کنیم و ببینیم که آیا این فرمول‌ها در عمل با حساب ممیز شناور نیز کار می‌کنند یا خیر؟

برای این کار دو فرمول تفاضلات متناهی را درنظر می‌گیریم. یکی فرمول (۳.۴)

$$f_i' \approx \frac{f_{i+1} - f_i}{h}$$

که دیدیم از مرتبه‌ی اول بود یعنی خطای برشی آن $\mathcal{O}(h)$ بود و دیگری فرمول (۸.۴)

$$f_i' \approx \frac{f_{i+1} - f_{i-1}}{2h}$$

که نشان دادیم از مرتبه‌ی دو است.

این دو فرمول را در یک آزمایش با هم مقایسه می‌کنیم. مسئله‌ی ساده‌ی محاسبه‌ی مشتق تابع

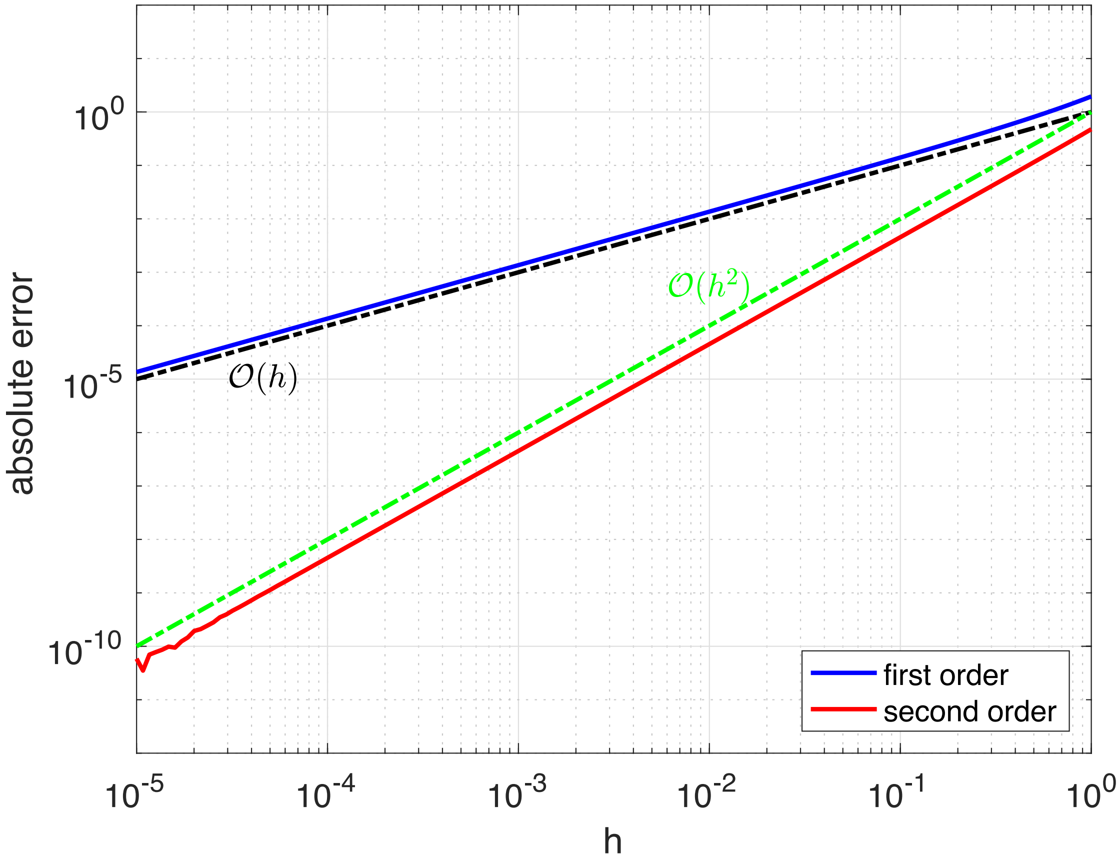
$$f(x) = \exp(x)$$

را در نقطه‌ی $x = 1$ با دو فرمول ((مشتق‌گیری عددی)) بالا را در نظر بگیرید. به این دلیل از این تابع ساده استفاده می‌کنیم که تعیین میزان خطای جواب، کاملا سرراست است چرا که می‌دانیم مقدار دقیق مشتق در

نقطه‌ی داده شده برابر است با $\exp(1)$.

۱.۶.۴ خبر خوب

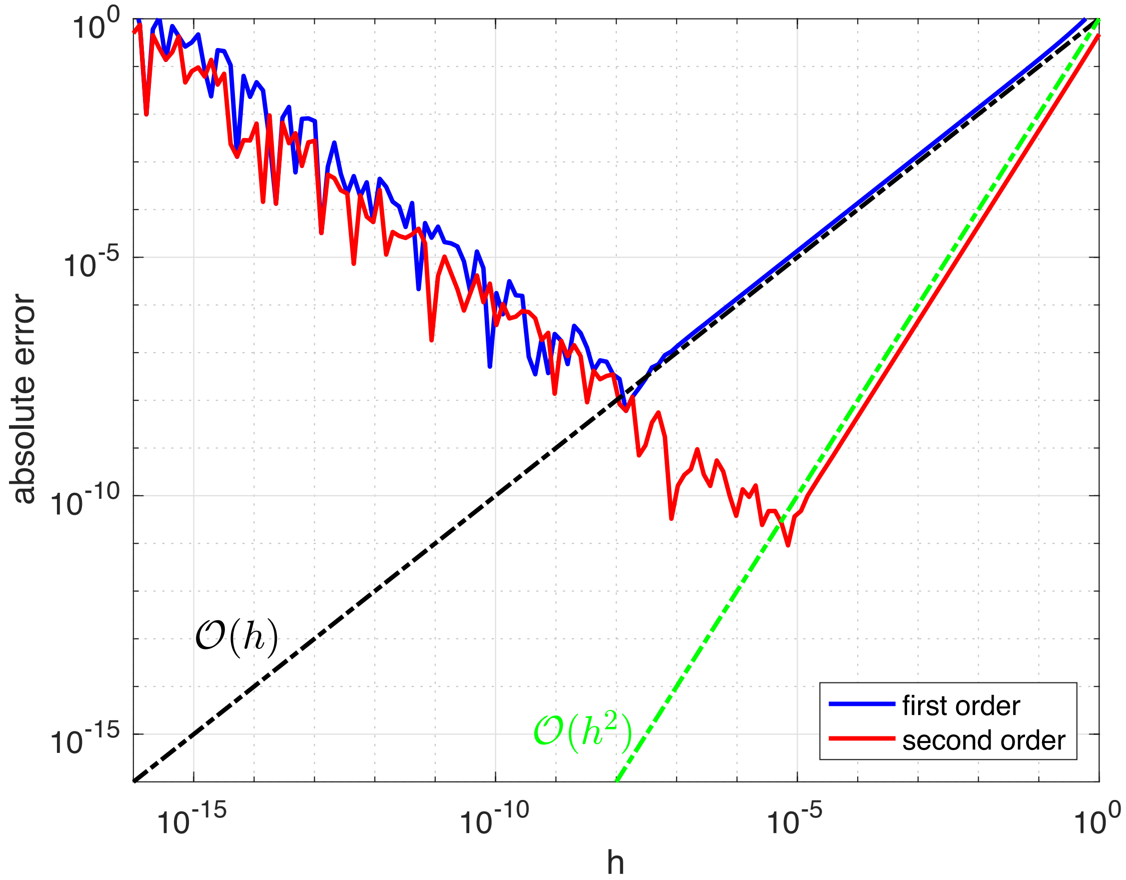
ثابت کردیم که در تئوری میزان خطای (برشی) دو فرمول قبل به ترتیب $\mathcal{O}(h)$ و $\mathcal{O}(h^2)$ است. میزان خطای (مطلق) دو فرمول را به ازای مقادیر مختلف h از 10^0 تا 10^{-5} با هم مقایسه می‌کنیم. شکل ۱.۴ حاصل را نشان می‌دهد جایی که دو خط قطعه‌قطعه دو منحنی دقیق $e_1(h) = h$ و $e_2(h) = h^2$ را نشان می‌دهند. خبر خوب اینکه همه چیز دقیقا طبق تئوری پیش رفته است!



شکل ۱.۴: مقایسه‌ی دو فرمول مشتق‌گیری عددی تفاضلات متناهی مرتبه‌ی اول و دوم

۲.۶.۴ خبر بد

این بار طول گام h را تا 10^{-16} یعنی تا نزدیکی اپسیلون ماشین کوچک می‌کنیم. یعنی میزان خطای (مطلق) دو فرمول قبل را به ازای h از 10^0 تا 10^{-16} با هم مقایسه می‌کنیم. شکل ۲.۴ را ببینید جایی که قسمت مربوط به h از 10^0 تا 10^{-5} دقیقا چیزی است که در شکل ۱.۴ نیز دیدیم. اما خبر بد اینکه با ادامه‌ی کاهش h، کاهش خطا در نمودار جدید تا حدود زیادی مطابق آنچه تئوری پیش‌بینی کرده، نیست!



شکل ۲.۴: مقایسه‌ی دو فرمول مشتق‌گیری عددی تفاضلات متناهی مرتبه‌ی اول و دوم برای طول گام‌های کوچک‌تر

- در فرمول پیشرو که مرتبه‌ی اول است با کاهش h تا حدود 10^{-8} همچنان کاهش خطا رخ داده و این فرمول نهایتا در بهترین حالت به خطای 10^{-8} می‌رسد. اما وقتی $h < 10^{-8}$ می‌شود، کاهش خطا ادامه نمی‌یابد! بجای اینکه خطا مانند شکل ۱.۴ باز هم کوچک‌تر شود، بزرگتر شده است! هیچ نشانی از پایداری عددی در رفتار روش تفاضلات پیشرو برای $h < 10^{-8}$ وجود ندارد. دلیل مشکل را می‌توان به پدیده‌ی حذف منتسب کرد که در زمان محاسبه‌ی صورت کسر مربوط به هر دو فرمول تفاضلات تقسیم‌شده رخ می‌دهد. به بیان دیگر وقتی $h < 10^{-5}$ می‌شود در فرمول

$$f_i' \approx \frac{f_{i+1} - f_i}{h}$$

مقدار $f_{i+1} = f(x_i + h)$ خیلی به مقدار $f_i = f(x_i)$ نزدیک شده و تفریق دو عدد نزدیک به هم یعنی پدیده‌ی حذف رخ می‌دهد. این مشکل برای hهای بزرگ‌تر وجود نداشته. یعنی (در این مثال خاص) وقتی $h \geq 10^{-5}$ خطای برشی بر خطاهای گردکردن غالب است و همه چیز طبق تئوری پیش رفته است. اما وقتی $h < 10^{-5}$ شود خطاهای گردکردن بر خطای برشی غالب شده و داستان

متفاوت است.

- در فرمول مرکزی که مرتبه‌ی دوم است وقتی h از 10^{-5} کمتر می‌شود کاهش خطا متوقف شده و نهایتا کمترین خطایی که با این فرمول به آن دست می‌یابیم حدود 10^{-10} است. اما وقتی $h < 10^{-5}$ می‌شود، ناپایداری عددی را در رفتار این روش تفاضلات متناهی مشاهده می‌کنیم.

۷.۴ مشتق‌گیری گام مختلط

قبل از پایان این فصل، بد نیست کمی با روش مشتق‌گیری گام مختلط [1] آشنا شویم. این، یکی از روش‌هایی است که می‌تواند بعنوان جایگزین برای مشتق‌گیری عددی با تفاضلات متناهی درنظر گرفته شود. فرض کنید تابع f برای ورودی‌های حقیقی مانند x ، حقیقی-مقدار باشد. ایده‌ی روش، گام برداشتن در صفحه‌ی مختلط بجای طول گام حقیقی است که در روش‌های تفاضلات متناهی استفاده می‌شود. به بیان دقیق‌تر بجای محاسبه‌ی $f(x+h)$ سری تیلور $f(x+ih)$ را در صفحه‌ی مختلط می‌نویسیم:

$$\begin{aligned} f(x+ih) &= f(x) + (ih)f'(x) + \frac{(ih)^2}{2!}f''(x) + \frac{(ih)^3}{3!}f''(x) + \cdots \\ &= f(x) + ihf'(x) - \frac{h^2}{2}f''(x) - i\frac{h^3}{6}f''(x) + \cdots . \end{aligned}$$

با محاسبه‌ی قسمت مختلط دو سمت تساوی داریم:

$$\mathrm{Imag}(f(x+ih)) = \mathrm{Imag}(ihf'(x)) + \mathcal{O}(h^3) = hf'(x) + \mathcal{O}(h^3).$$

در اینجا $\mathrm{Imag}(z)$ نشان‌دهنده‌ی قسمت مختلط z بوده و $\mathcal{O}(h^3)$ بجای $-i\frac{h^3}{6}f''(\zeta)$ ظاهر شده جایی که ζ متعلق به دیسکی به مرکز x و شعاع h در صفحه‌ی مختلط است. با تقسیم دو سمت رابطه‌ی قبل بر h داریم

$$f'(x) = \frac{\mathrm{Imag}(f(x+ih))}{h} + \mathcal{O}(h^2).$$

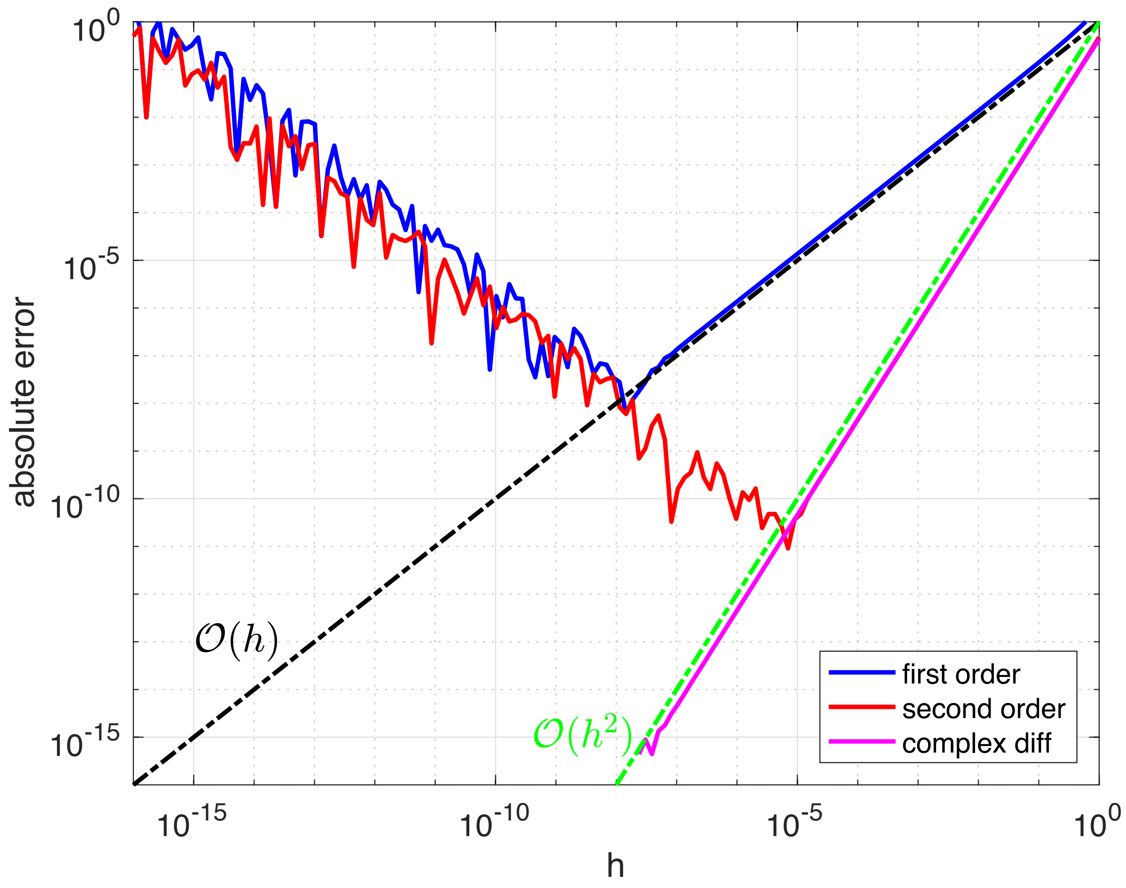
پس فرمول

$$f'(x) \approx \frac{\mathrm{Imag}(f(x+ih))}{h}. \tag{۹.۴}$$

[1] complex step differentiation

حاصل می‌شود که دارای خطای $\mathcal{O}(h^2)$ است. بنابراین خطای (برشی) این روش متناسب با فرمول تفاضلات مرکزی (۸.۴) است در حالی که در اینجا تنها نیاز به یک ارزیابی مقدار تابع f است اما در فرمول (۸.۴) نیاز به دو ارزیابی مقدار تابع f بود. بعلاوه همانطور که می‌بینیم در فرمول (۹.۴) هیچ تفریقی که احیانا بتواند در شرایط خاصی منجر به پدیده‌ی حذف شود وجود ندارد. این بدان معناست که با ارزیابی تابع f در ورودی مختلط $x+ih$ و تقسیم بر h می‌توان تقریبی از $f'(x)$ به دست آورد که تا مرتبه‌ی h^2 درست یعنی اگر h را حدود 10^{-8} انتخاب کنیم تقریبی برای $f'(x)$ خواهیم داشت که خطایی در حد اپسیلون ماشین خواهد داشت: خطایی که حتی در ذخیره‌سازی یک عدد حقیقی با یک عدد قالب دوگانه‌ی IEEE هم رخ می‌دهد.

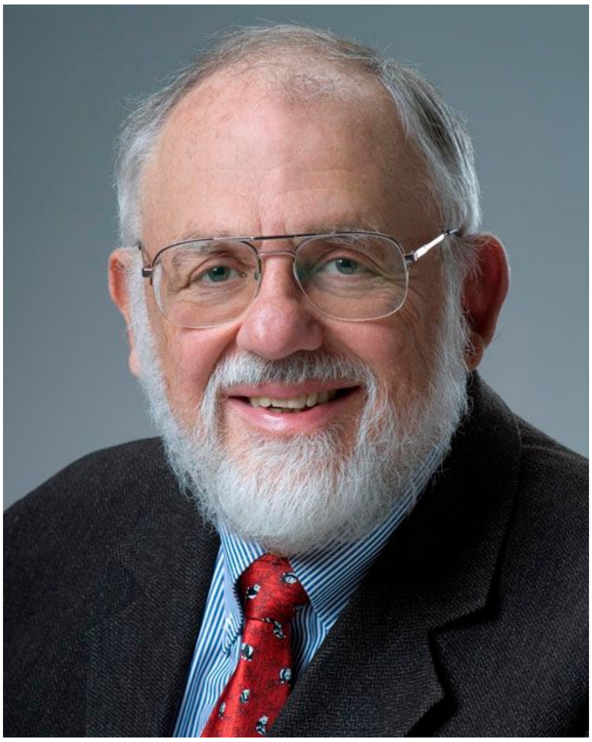
اکنون آزمایش قبل مربوط به مشتق تابع $f(x)=\exp(x)$ را با در نظرگرفتن فرمول تفاضلات مختلط (۹.۴) مجددا اجرا می‌کنیم. شکل ۳.۴ پایداری عددی روش گام مختلط را نشان می‌دهد.



شکل ۳.۴: مقایسه‌ی مشتق‌گیری عددی با تفاضلات متناهی مرتبه‌ی اول و دوم و روش گام مختلط

نکته‌ی پایانی در مورد مقایسه‌ی سه روش قبل این که هرچند فرمول‌های گام مختلط چشمگیر بوده و مشکل ناپایداری عددی تفاضلات متناهی را حل می‌کنند، اما بر این فرض استوار شده‌اند که قادر هستیم تابع

f را در صفحه‌ی مختلط ارزیابی کنیم. این بدان معناست که تابع f باید در صفحه‌ی مختلط، ((تحلیلی))[۱] باشد که شرطی قوی‌تر از شرط مشتق‌پذیری از مرتبه‌ی دوم و سوم روی تابع f است که دو فرمول تفاضلات متناهی (۳.۴) و (۸.۴) به ترتیب نیاز دارند. از این دیدگاه مقایسه‌ی این روش‌ها با هم می‌تواند تا حد زیادی غیرمنصفانه باشد.



شکل ۴.۴: کلیو مولر، فارغ‌التحصیل کَلتِک و استنفورد، روش مشتق‌گیری گام مختلط را (به همراه جیمز لینِس) در سال ۱۹۶۷ معرفی کرد. وی در اواخر دهه‌ی ۱۹۷۰ نخستین نسخه‌ی نرم‌افزار متلب را برای استفاده در کلاس درسش در دانشگاه نیومکزیکو نوشت. مولر در سال ۱۹۸۴ شرکت مَت‌وُرکز که محصول اصلی آن متلب است را راه‌اندازی کرد. متلب با حدود سه میلیون کاربری که در سرتاسر جهان دارد، درآمد شرکت مَت‌وُرکز را به حدود ۹۰۰ میلیون دلار در سال ۲۰۱۷ رسانده است. مولر جوایز مختلفی را در علوم کامپیوتر برنده شده است. (عکس از وب‌سایت موزه‌ی تاریخ کامپیوتر)

موضوع مرتبط دیگر که در این مجالِ کوتاه به آن نپرداختیم، ((مشتق‌گیری خودکار [۲])) است که به تازگی اهمیت فراوانی بخصوص در بهینه‌سازی پیداکرده، جایی که نیاز زیادی به بردارهای ژاکوبین و ماتریس‌های هسه‌ای وجود دارد.

[۱] یعنی بینهایت‌بار مشتق‌پذیر و دارای ادامه‌ی تحلیلی در صفحه‌ی مختلط. به زبان غیردقیق به لحاظ محاسباتی این بدان معناست که ضابطه‌ی تابع f به صورتی ساده مثلا در یک خط و احتمالا بدون شرط if و یا بدون یک حلقه‌ی for تعریف شده است.

[۲] automatic differentiation