Numerisk beräkning av derivator

Fredrik Berntsson

20 april 2022

Sammanfattning

Numerisk derivering leder till svårigheter då problemet är instabilit i den meningen att små störningar i använda funktionsvärden lätt leder till relativt stora störningar i svaret. Vi undersöker denna instabilitet bårde teoretiskt och numeriskt.

1 Introduktion

Vissa typer av problem leder till matematiska modeller som är svåra att lösa numeriskt på grund av känslighet, med avseende på tillexempel beräkningsfel. Sådana problem kallas ofta *illa-ställda*. Ett vanligt sätt att definiera detta begrepp är som följer [2]: Ett problem kallas *rätt–ställt* om följande villkor är uppfyllda

- En lösning existerar för alla tillåtna data.
- Lösningen till problemet är unik.
- Lösningen beror kontinuerligt på data.

Så snart ett matematiskt problem inte upfyller ett, eller flera, av dessa villkor kallas problemet *illa-ställt*. Givetvis så är ovanstående ingen strikt matematisk definition. Vi skulle isåfall behöva förtydliga vilka data som är tillåtna, vilket mått som skall användas för att mäta det kontinuerliga beroendet, samt vad som menas med en lösning till problemet. Detta måste göras på ett sätt som passar det konkreta problem man vill studera.

Vi skall i denna text studera ett specifikt exempel på ett illa-ställt problem. Syftet är att ge en intuitiv förståelse av vilken typ av svårigheter som dyker upp då denna typ av matematiska problem behandlas teoretiskt och numeriskt.

Denna text är strukturerad som följer: En kort teoretisk genomgång av problemställningen finns i Kapitel 2. I Kapitel 3 presenteras ett numeriskt experiment. Några korta slutord presenteras i Kapitel 4.

2 Numerisk beräkning av derivator

Derivatan av en kontinuerlig funktion definieras som

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h},\qquad h>0. \tag{1}$$

Om detta gränsvärde existerar så sägs funktionen $f(x)$ vara deriverbar [1], och derivatans värde ges av nämnda gränsvärde.

Antag nu att $f(x)$ är en deriverbar funktion. För att tydligt se att (1) kan leda till instabilitet så låter vi $\delta > 0$, och n vara ett heltal. Definiera

$$f_{n,\delta} = f(x) + \delta \sin(\frac{nx}{\delta}), \qquad -\infty < x < \infty. \tag{2}$$

Vi betraktar $f(x)$ som den exakta funktion vi vill derivera, och $f_{n,\delta}(x)$ som en approximation av $f(x)$ innehållande ett litet mätfel. Vi ser nu att även $f_{n,\delta}(x)$ är deriverbar, och kan beräkna derivatan,

$$f'_{n,\delta}(x) = f'(x) + n \cos(\frac{nx}{\delta}).$$

Observera att,

$$\max_{-\infty<x<\infty} |f(x) - f_{n,\delta}(x)| = \delta, \quad \text{och,} \max_{-\infty<x<\infty} |f'(x) - f'_{n,\delta}(x)| = n.$$

Skillnaden mellan de båda funktionerna kan alltså vara godtyckligt liten, och trots det kan skillnaden i derivata bli stor.

Slutsatsen är alltså att två funktioner som liknar varandra kan ha väldigt olika derivata. Vi kan alltså säga att i detta fall beror *lösningen*, dvs den beräknade derivatan, ej *kontinuerligt* på använda data. Detta betyder att derivering är ett illa-ställt problem. Detta måste på något sätt visa sig i praktiska beräkningar.

För att beräkna derivator numeriskt används ofta differenskvoter baserade på definitionen (1). Ofta används *central differens approximation*: Välj ett fixt $h > 0$ och sätt

$$f'(x) \approx D_0 f(x) = \frac{f(x+h) - f(x-h)}{2h}. \tag{3}$$

Vi skall nu undersöka hur stor skillnaden är mellan f' och approximationen $D_0 f$.

För att undersöka hur väl central differens approximationen överensstämmer med derivatan används Taylor's formel [3]:

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

med hjälp av Taylor's formel kan vi skriva (3) som

$$D_0 f(x) = f'(x) + \frac{h^2}{6} f^{(3)}(x) + \ldots$$

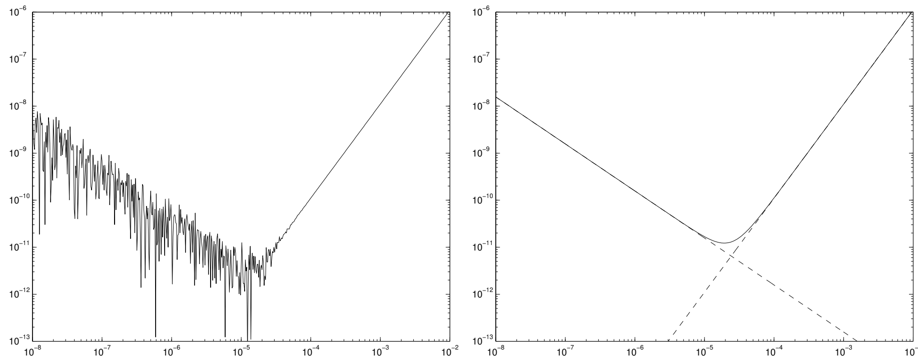
Då en derivata approximeras med hjälp av en central differens görs alltså ett trunkeringsfel

$$R_T = D_0 f(x) - f'(x) \approx \frac{h^2}{6} f^{(3)}(x). \tag{4}$$

En viktig observation är att trunkeringsfelet minskar då $h \to 0$. Denna del är alltså inte någon källa till instabilitet.

Vi skall nu undersöka vad som händer då funktionen $f(x)$ endast kan beräknas med ett litet fel. Vi antar alltså att det finns ett $\varepsilon > 0$ och att de beräknade värdena $\bar{f}(x)$ satisfierar $|\bar{f}(x) - f(x)| < \varepsilon$. Detta kommer att resultera i en felterm,

$$R_{XF} = |D_0 \bar{f}(x) - D_0 f(x)| < \frac{\varepsilon}{h}. \tag{5}$$



Figur 1: Totala felet $|D_0\bar{f}(2) - f'(2)|$ för olika värden på h visas i vänstra grafen. De teoretiskt uträknade feltermerna R_T, R_{XF}, samt R_{TOT}, visas i högra grafen.

Vi ser att indatafelet förstärks och R_{XF} alltså växer då $h \to 0$.

Det totala felet blir alltså,

$$R_{TOT} = R_T + R_{XF} < \frac{h^2}{6} f^{(3)}(x) + \frac{\varepsilon}{h}. \tag{6}$$

Det optimala valet av steglängden h kommer alltså att bli kompromiss mellan behovet att approximera derivatan noggrannt, dvs litet R_T, och att begränsa inverkan av beräkningsfel, dvs R_{XF}. Detta är typiskt för många illa-ställda problem. Ofta måste vissa parametrar i den numeriska metoden väljas på ett lämpligt sätt.

3 Ett numeriskt exempel

Vi skall nu genomföra ett praktiskt experiment där funktionen $f(x)=\sqrt{x}$ deriveras med hjälp av en central differens approximation. Vi skall försöka beräkna $f'(2)$. Beräkningarna utfördes i MATLAB med hjälp av följande korta program:

```
h=10.^-(2:0.01:8);
df=(sqrt(2+h)-sqrt(2-h))./(2*h);
Rtot=abs(df-1/sqrt(2)/2);
loglog(h,Rtot);
```

Resultatet av beräkningen visas i Figur 1. Notera att felet är stort både för stora och väldigt små h.

Eftersom $f(x) = \sqrt{x}$ så kan felgränserna (4) och (5) beräknas explicit. Vi antar att $f(x)$ kan beräknas med ett relativt fel begränsat av datorns avrundningsenhet, dvs $|\bar{f}(x) - f(x)| \leq \mu = 2^{-53}$. Totala felet ges då av,

$$R_{TOT} \leq R_{XF} + R_T \approx \frac{h^2}{6}|f^{(3)}(2)| + \frac{\mu}{h}|f(2)| = 1.10 \cdot 10^{-2} h^2 + 1.57 \cdot 10^{-16} \frac{1}{h}. \tag{7}$$

Dessa felgränser illustreras i Figur 1. Det är tydligt att den teoretiska felgränsen överensstämmer väl med det praktiska resultatet.

Det finns en optimal steglängd som kan beräknas genom att minimera felgränsen (7). Vi får att

$$h_{opt} = \left(\frac{3\mu |f(2)|}{|f^{(3)}(2)|} \right)^{\frac{1}{3}} \approx 1.9 \cdot 10^{-5}.$$

Vi ser att för detta speciella värde på h fås det minsta totala felet även i praktiken.

4 Slutsatser

Då man beräknar derivator numeriskt är det viktigt att inse att valet av steglängd innebär en kompromiss mellan dels inverkan av beräknngsfel, och dels behovet att använda en differenskvot som ger tillräckligt litet trunkeringsfel.

Vi har visat hur man kan använda härledda uttryck för dessa två felkomponenter för att välja ett bra värde på steglängden. Ett praktiskt experiemnt visar att teori och praktik stämmer väl överens.

Referenser

[1] Göran Forsling and Mats Neymark. *Matematisk Analys - En variabel.* Liber AB, Stockholm, 2ed edition, 2011.

[2] J. Hadamard. *Lectures on Cauchy's problem in linear partial differential equations.* Dover Publications, New York, 1953.

[3] L. Råde and B. Westergren. *BETA Mathematics handbook.* Studentlitteratur, Lund, 1993.