# TEMATIME - 43

## NUMERISK DERIVASJON

Den deriverte til en funksjon $f$ er

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

dersom grenseverdien eksisterer. Uttrykket

$$\frac{f(x+h) - f(x)}{h}$$

er stigningstallet til sekanten til $f$ mellom punktene $x$ og $x+h$. For små $h$ er dette en grei tilnærming til $f'(x)$:

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

På papiret er det slik at jo mindre $h$, desto bedre tilnærming. La $f(x) = e^x$, og la $h = 0.1$. Vi beregner

$$f'(1.5) \approx \frac{e^{1.6} - e^{1.5}}{0.1} = 4.7134.$$

Merk at

$$f'(1.5) = e^{1.5} = 4.4817,$$

så denne tilnærmingen bommer med rundt $2 \cdot 10^{-1}$.

1 Gjenta eksperimentet med $h = 0.01$, $h = 0.001$ osv, og se hva som skjer. Hvor liten $h$ kan du ta før det går åt skogen?

Feilen i forrige oppgave er tydelig proporsjonal med $h$ - deler du $h$ på 10, deler du feilen på 10. En god illustrasjon av lineær feil. Men dette gjelder bare inntil et visst punkt, blir $h$ liten nok blir presisjonen dårligere igjen. Dette er fordi datamaskinen din kun regner med 16 desimaler.

2 Gjenta eksperimentet i forrige oppgave, men nå med formelen

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

og se hva som skjer. Hvor liten $h$ kan du nå ta før det går åt skogen?

Legg merke til hvordan feilen i forrige oppgave er nærmest perfekt kvadratisk - vi får to nye desimaler hver gang vi deler $h$ på 10. Feilen deles altså på 100 når $h$ deles på 10. Den første formelen kalles en **foroverdifferanse** og den andre en **sentraldifferanse**. Nå skal vi finne ut hvorfor den ene har lineær feil og den andre kvadratisk. Verktøyet vi trenger, er Taylors teorem.

## TAYLORS TEOREM

Her er analysens fundamentalteorem, skrevet på en rar måte:

$$f(x) = f(a) + \int_a^x f'(s)\, ds$$

Nå tar vi en delvis integrasjon på integralet på høyre side. Men ikke slik du er vant til. Integranden er $1 \cdot f'(s)$, og som antiderivert til 1 velger vi ikke $s$, men $s - x$. Det er lov, for vi integrerer med hensyn på $s$, og da er $x$ bare en konstant. Vi får

$$\int_a^x f'(s)\,ds = (s-x)f'(s)\Big|_{s=a}^{s=x} - \int_a^x (s-x)f''(s)\,ds$$

$$= f'(a)(x-a) + \int_a^x (x-s)f''(s)\,ds$$

slik at

$$f(x) = f(a) + f'(a)(x-a) + \int_a^x (x-s)f''(s)\,ds.$$

Vi gjentar prosessen, nå med $-\frac{1}{2}(x-s)^2$ som antiderivert til $x-s$:

$$\int_a^x (x-s)f''(s)\,ds = -\frac{1}{2}(x-s)^2 f''(s)\Big|_{s=a}^{s=x} + \frac{1}{2}\int_a^x (x-s)^2 f'''(s)\,ds$$

$$= \frac{f''(a)}{2}(x-a)^2 + \frac{1}{2}\int_a^x (x-s)^2 f'''(s)\,ds$$

Da ender vi opp med

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{1}{2}\int_a^x (x-s)^2 f'''(s)\,ds.$$

Holder du tungen beint i munnen, kan du fortsette slik så lenge $f$ lar seg derivere. **Taylors teorem** sier at dersom $f$ er en $n+1$ ganger kontinuerlig deriverbar funksjon på et intervall som inneholder $a$ og $x$, er

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

$$+ \frac{1}{n!}\int_a^x (x-s)^n f^{(n+1)}(s)\,ds$$

Under de samme forutsetningene på $f$ finnes en $s$ mellom $a$ og $x$ slik at feilen er lik $\frac{f^{(n+1)}(s)}{(n+1)!}(x-a)^{n+1}$. For $n = 0$ er dette sekantsetningen. La $M$ være den største verdien $|f^{(n+1)}|$ tar mellom $a$ og $x$. Absoluttverdien til et integral er aldri større enn integralet av absoluttverdien, så for $x > a$ er

$$\left|\frac{1}{n!}\int_a^x (x-s)^n f^{(n+1)}(s)\,ds\right| \le \frac{M}{n!}\int_a^x (x-s)^n\,ds = \frac{M}{(n+1)!}(x-a)^{n+1}.$$

For $x < a$ går regningen på samme måte, så uansett er feilen i absoluttverdi høyst

$$\frac{M}{(n+1)!}|x-a|^{n+1}.$$

## FEILEN I DIFFERANSEFORMLENE

Nå kan vi forklare eksperimentene. Vi lar $h > 0$ og bruker Taylors teorem med $n = 1$, der vi setter $x$ inn for $a$ og $x + h$ inn for $x$:

$$f(x+h) = f(x) + f'(x)h + R$$

Her er $R$ feilen, og den er i absoluttverdi høyst $\frac{M}{2}h^2$. Stigningstallet til sekanten kan derfor skrives slik:

$$\frac{f(x+h)-f(x)}{h} = f'(x) + \frac{R}{h}$$

Dette stigningstallet består av den eksakte verdien for $f'(x)$ pluss en feil, og feilen er ikke større enn $\frac{M}{2}h$:

$$\left|\frac{f(x+h)-f(x)}{h} - f'(x)\right| \leq \frac{M}{2}h$$

For $f(x) = e^x$, $x = 1.5$ og $h = 0.1$ er $M = e^{1.6}$, og estimatet sier at vi bommer med høyst $e^{1.6} \cdot 0.1/2 \approx 0.25$. I eksperimentet bommet vi med $0.23$. Når $h$ er liten, er $f''$ nesten konstant mellom $x$ og $x+h$, og da er feilen nesten nøyaktig $\frac{f''(x)}{2}h$. Det er derfor feilen deles på 10 når $h$ deles på 10.

For sentraldifferansen tar vi med ett ledd til i taylorpolynomet, altså $n = 2$, og bruker teoremet to ganger - en gang med $x+h$ og en gang med $x-h$:

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2}h^2 + R_1$$

$$f(x-h) = f(x) - f'(x)h + \frac{f''(x)}{2}h^2 + R_2$$

Begge feilene er i absoluttverdi høyst $\frac{M}{6}h^3$, der $M$ nå er den største verdien $|f'''|$ tar mellom $x-h$ og $x+h$. Trekker vi den nederste likningen fra den øverste, forsvinner både $f(x)$ og leddet med $f''(x)$:

$$f(x+h) - f(x-h) = 2f'(x)h + R_1 - R_2$$

Vi deler på $2h$ og får

$$\left|\frac{f(x+h)-f(x-h)}{2h} - f'(x)\right| = \frac{|R_1 - R_2|}{2h} \leq \frac{M}{6}h^2.$$

Feilen er altså $O(h^2)$. Det er symmetrien som gjør jobben - leddet med $h^2$ er likt i de to likningene, og forsvinner når vi trekker dem fra hverandre. For $f(x) = e^x$, $x = 1.5$ og $h = 0.1$ sier estimatet at vi bommer med høyst $e^{1.6} \cdot 0.01/6 \approx 0.0083$, og sentraldifferansen bommer med $0.0075$. Når $h$ er liten, er feilen nesten nøyaktig $\frac{f'''(x)}{6}h^2$, og derfor deles den på 100 når $h$ deles på 10.

Begge estimatene sier at feilen går mot null når $h$ går mot null, så hvorfor går det åt skogen når $h$ blir for liten? Datamaskinen regner som sagt bare med 16 desimaler, så både $f(x+h)$ og $f(x)$ kommer med en avrundingsfeil på rundt $10^{-16}|f(x)|$. Når differansen deles på $h$, deles avrundingsfeilen også på $h$, og den totale feilen i foroverdifferansen blir omtrent

$$\frac{M}{2}h + \frac{2 \cdot 10^{-16}|f(x)|}{h}$$

der $M$ igjen er den største verdien $|f''|$ tar mellom $x$ og $x+h$. Det første leddet krymper når $h$ krymper, men det andre vokser. Summen er minst når de to leddene er like store, og for eksponensialfunksjonen skjer det når $h$ er rundt $10^{-8}$. Mindre $h$ enn dette gjør bare vondt verre. Tilsvarende regnestykke for sentraldifferansen gir at den beste $h$ er rundt $10^{-5}$.

3 Dersom du trenger en tilnærming for $f''(x)$, kan du bruke den andre ordens sentraldifferansen

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

Bruk Taylors teorem med $n = 3$ til å vise at feilen er høyst $\frac{M}{12}h^2$, der $M$ er den største verdien $|f^{(4)}|$ tar mellom $x - h$ og $x + h$.

4 Hvis du virkelig vil slå på stortrommen, kan du prøve eksperimentet med formelen

$$f'(x) \approx \frac{f(x-2h) - 8f(x-h) + 8f(x+h) - f(x+2h)}{12h}.$$

Hva ser feilen ut til å være proporsjonal med?

## TRAPESMETODEN

Vi deler $[a, b]$ inn i $n$ like lange delintervaller med gitteravstand $h = (b-a)/n$ og gitterpunkter $x_k = a+kh$. Tar du gjennomsnittet av høyre og venstre riemannsum får du noe som kalles **trapesmetoden**. På hvert delintervall erstattes arealet under grafen med arealet av et trapes:

$$\int_{x_k}^{x_{k+1}} f(x)\,dx \approx \frac{h}{2}\left(f(x_k) + f(x_{k+1})\right)$$

Hvor mye bommer vi? Anta at $f$ er to ganger kontinuerlig deriverbar. Vi bruker samme triks som for Taylors teorem - delvis integrasjon med et uvant valg av antiderivert. La $c = x_k + h/2$ være midtpunktet i delintervallet. Integranden er $1 \cdot f(x)$, og som antiderivert til 1 velger vi $x - c$:

$$\begin{aligned}\int_{x_k}^{x_{k+1}} f(x)\,dx &= (x-c)f(x)\Big|_{x=x_k}^{x=x_{k+1}} - \int_{x_k}^{x_{k+1}} (x-c)f'(x)\,dx \\ &= \frac{h}{2}\left(f(x_k) + f(x_{k+1})\right) - \int_{x_k}^{x_{k+1}} (x-c)f'(x)\,dx\end{aligned}$$

Arealet av trapeset detter ut av seg selv, og integralet som står igjen, er feilen. Vi delvisintegrerer en gang til. De antideriverte til $x - c$ er $\frac{1}{2}(x-c)^2$ pluss en konstant, og konstanten velger vi slik at den antideriverte er null i begge endepunktene:

$$w(x) = \frac{1}{2}(x-c)^2 - \frac{h^2}{8}$$

Da blir leddet utenfor integralet null, og vi står igjen med

$$\int_{x_k}^{x_{k+1}} f(x)\,dx = \frac{h}{2}\left(f(x_k) + f(x_{k+1})\right) + \int_{x_k}^{x_{k+1}} w(x)f''(x)\,dx.$$

Mellom endepunktene er $w(x)$ negativ, og

$$\int_{x_k}^{x_{k+1}} |w(x)|\,dx = \int_{x_k}^{x_{k+1}} \left(\frac{h^2}{8} - \frac{1}{2}(x-c)^2\right)dx = \frac{h^3}{12}.$$

La nå $M$ være den største verdien $|f''|$ tar på $[a, b]$. Feilleddet er da i absoluttverdi høyst $M$ ganger dette integralet, så vi bommer med høyst $\frac{M}{12}h^3$ på hvert delintervall. Trapesmetoden på hele $[a, b]$ er summen av alle trapesene,

$$T = h\left(\frac{f(x_0)}{2} + f(x_1) + \cdots + f(x_{n-1}) + \frac{f(x_n)}{2}\right).$$

Det er $n$ delintervaller, og $nh = b - a$, så

$$\left| \int_a^b f(x)\, dx - T \right| \leq n \cdot \frac{M}{12} h^3 = \frac{(b-a)M}{12} h^2.$$

Feilen er altså $O(h^2)$, akkurat som for sentraldifferansen. Siden $w$ er negativ, forteller feilleddet også hvilken vei vi bommer: Krummer grafen nedover, er $f''$ negativ og feilleddet positivt - trapesene ligger under grafen, og trapesmetoden gir for lite.

La oss teste estimatet på integralet fra innlevering 4:

$$\frac{1}{\sqrt{2\pi}} \int_0^1 e^{-x^2/2}\, dx \approx 0.3413447460685429485852$$

Her er $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, og deriverer vi to ganger, får vi

$$f''(x) = \frac{1}{\sqrt{2\pi}} \left(x^2 - 1\right) e^{-x^2/2},$$

som er størst i absoluttverdi i $x = 0$, så $M = 1/\sqrt{2\pi} \approx 0.4$. Med $n = 1000$ er $h = 10^{-3}$, og estimatet sier at trapesmetoden bommer med høyst

$$\frac{0.4}{12} \cdot 10^{-6} \approx 3.3 \cdot 10^{-8}.$$

I virkeligheten bommer den med $2.0 \cdot 10^{-8}$, mens venstre riemannsum bommer med $7.8 \cdot 10^{-5}$.

5 Venstre riemannsum erstatter integralet over hvert delintervall med $hf(x_k)$. Gjør én delvis integrasjon, nå med $x - x_{k+1}$ som antiderivert til $1$, og vis at venstre riemannsum bommer med høyst

$$\frac{(b-a)M}{2} h$$

der $M$ er den største verdien $|f'|$ tar på $[a, b]$. Stemmer dette med det du så i innlevering 4?