# Kortversjon av feilanalyse for numerisk derivasjon

Knut Mørken

11. november 2015

På forelesningen mandag 9/11 gjennomgikk vi feilanalysen for numerisk derivasjon som du finner i seksjonene 11.1 og 11.2 i kompendiet. Jeg valgte å kommentere detaljene underveis i resonnementet, og da framstår resonnementet som langt. Hensikten med dette notatet er å gi en kort versjon som illustrerer at dette egentlig er ganske kort.

## Utgangspunkt

Vi har en funksjon $f$ som vi i utgangspunktet ikke kjenner utrykket til, men vi kan framskaffe funksjonsverdier, enten ved måling eller ved beregning på datamaskin. Vi ønsker å kunne beregne verdier av den deriverte $f'$, men dette må gjøres ved hjelp av funksjonsverdier. Vi baserer oss derfor på tilnærmingen

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

der $a$ er punktet der vi ønsker å regne ut den deriverte og $h > 0$ er et valgt, positivt tall.

## Kvaliteten på tilnærmingen — trunkeringsfeil

For å vurdere kvaliteten til tilnærmingen skal vi studere feilen så vi antar at $f$ kan deriveres og tilnærmes med et Taylorpolynom,

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2} f''(\xi_h),$$

der $\xi_h$ er et tall i intervallet $(a, a+h)$ som avhenger av $h$. Snur vi om på dette får vi

$$f'(a) - \frac{f(a+h) - f(a)}{h} = -\frac{h}{2} f''(\xi_h). \tag{1}$$

## Kvaliteten på tilnærmingen — avrundingsfeil

På datamaskinen kan som regel ikke tallet $f(a)$ representeres eksakt. I steden får vi det nærmeste flyttallet $\overline{f(a)}$, se lemma 5.21 i kapittel 5 i kompendiet. Da vet vi at

$$\frac{f(a) - \overline{f(a)}}{f(a)} = \epsilon_1 \quad \text{eller} \quad \overline{f(a)} = f(a)(1 + \epsilon_1),$$

der $\epsilon_1$ er et tall som i tallverdi er begrenset av $\epsilon^*$ og $\epsilon^*$ er omtrent $10^{-16}$ (hvis vi bruker 64 bits flyttall). Med andre ord er den relative feilen maksimalt omtrent $10^{-16}$ slik at vi har omtrent 16 riktige siffer i $\overline{f(a)}$.

På samme måte får vi en avrundingsfeil når vi regner ut $\overline{f(a+h)}$, der

$$\frac{f(a+h) - \overline{f(a+h)}}{f(a+h)} = \epsilon_1 \quad \text{eller} \quad \overline{f(a+h)} = f(a+h)(1 + \epsilon_2),$$

der $\epsilon_2$ er et tall som i tallverdi er begrenset av $\epsilon^*$. Merk at $\epsilon_2$ vil avhenge av $h$.

## Total feil

Vi kan nå sette sammen de to feilkildene og studere den totale feilen,

$$\begin{aligned}
f'(a) - \frac{\overline{f(a+h)} - \overline{f(a)}}{h} &= f'(a) - \frac{f(a+h)(1+\epsilon_2) - f(a)(1+\epsilon_1)}{h} \\
&= f'(a) - \frac{f(a+h) - f(a)}{h} - \frac{f(a+h)\epsilon_2 - f(a)\epsilon_1}{h} \\
&= -\frac{h}{2} f''(\xi_h) - \frac{f(a+h)\epsilon_2 - f(a)\epsilon_1}{h}.
\end{aligned}$$

Fra dette uttrykket ser vi i store trekk hvordan den totale feilen oppfører seg: For store $h$ vil det første leddet dominere, mens for små $h$ vil det andre leddet dominere. Dermed fins det en optimal $h$ som viser seg å være omtrent $10^{-8}$ hvis funksjonsverdiene er av størrelsesorden 1. Mer detaljer finner du i seksjon 11.2 i kompendiet.