

# 1 Modellering og computer fejl ved beregninger

## 1.1 Matematisk modellering

Medmindre den opstillede er meget simpel, så kan man ikke forvente at problemet har en pæn analytisk løsning -> der skal anvendes en numerisk tilgang, hvor der iterativt approksimeres en løsning (medfører yderligere fejl). Computeren har endelig hukommelse, hvilket skaber afrundings- og trunkeringsfejl.

## 1.2 Approksimationsfejl

Fejl ved approksimation:

Når vi skal bestemme præcisionen af en approksimeret  $x := \hat{x}$ , så bliver vi nødt til at kunne måle fejlen. Den absolute fejl måles ved  $|x - \hat{x}|$ , og den relative fejl måles ved  $\frac{|x - \hat{x}|}{|x|}$ . Hvilket kræver vi kender det vi forsøger at approksimere, hvilket er underligt og normalt ikke en mulighed.

**Absolute- og relative fejl:** Forskel på absolute fejl og relative fejl: Absolut fejl er ved værdier omkring 1, mens relative fejl anvendes ved store tal samt tal omkring 0.

**Trunkeringsfejl:** Når en uendelig række i en sum forkortes til en bestemt mængde af led, da er der lavet en trunkeringsfejl (afkørningsfejl).

**Ustabilitet og ill-conditioned problem:** En algoritme kaldes numerisk ustabil, hvis afrundingsfejl og trunkeringsfejl risikerer at medføre, at approksimationsfejlen vokser, så vi ikke nærmer os løsningen. Selvom en algoritme er numerisk stabil, så kan det underlæggende problem være sådan, at en lille afvigelse i algoritmens input medfører en stor afvigelse i algoritmens output. Et sådan problem kaldes *ill-conditioned*.

## 1.3 Hvordan repræsenteres tal i computeren

Antallet af betydende cifre har betydning for hvordan tallet skal repræsenteres. Der anvendes her symmetrisk afrunding eller chopping. Symmetrisk afrunding er det almindelige anvendte, mens chopping runder ned (chopper).

Computere anvender det binære talsystem. Tal repræsenteres med et fortegn, en eksponent og en "mantissa" (de betydende cifre). Helt generelt repræsenteres et positivt tal  $x$  ved en base  $\beta$  som

$$x = f\beta^E, \quad 1 \leq f < \beta,$$

hvorom  $f$  er mantissa'en for tallet. For det binære talsystem er dette

$$x = f \cdot 2^E, \quad 1 \leq f < 2.$$

## 2 Taylors formel

En funktion kan approksimeres omkring et punkt ud fra et polynomium af funktionens afledede til det punkt. Denne approksimations nøjagtighed øges ved højere grad.

### Definition 2.1 (Taylor polynomie)

Antag, at  $f$  er  $N$  gange kontinuert differentiabel på et interval  $I = (a, b)$  og lad  $x_0 \in I$  være fast. Polynomiet i  $h$  (afhængig af  $h$  da udviklingspunktet er holdt fast) givet ved

$$P_N(x_0; h) = \sum_{k=0}^N \frac{h^k}{k!} f^{(k)}(x_0)$$

kaldes Taylor polynomiet af grad  $N$ , hvor  $h$  er afstanden væk fra udviklingspunktet  $x_0$ .

Bevis findes i slides for anden kursusgang.

### Theorem 2.1 (Restleddet $R_{N+1}$ )

Restleddet af Taylors formel kan alternativt udtrykkes som

$$R_{N+1}(x_0; h) = \frac{h^{N+1}}{(N+1)!} f^{(N+1)}(x_0 + \theta h),$$

hvor  $0 < \theta < 1$  er et tal, der afhænger af både  $x_0, h$  og  $f$ .

Hvis  $f^{(N+1)}(x) \leq M \forall x \in I$  så får vi dermed

$$|R_{N+1}(x_0; h)| \leq \frac{|h|^{N+1}}{(N+1)!} M.$$

Når  $N \rightarrow \infty$  ved  $|h| < 1$ , så går  $R_{(N+1)} \rightarrow 0$ , hvilket betyder approksimationen nærmer sig funktionen da restleddet går mod 0.

### Theorem 2.2 (Taylors Formel)

Antag, at  $f$  er en  $N+1$  gange kontinuert differentiabel funktion på et interval  $I = (a, b)$  og lad  $x_0 \in I$  være fast. Så gælder

$$\begin{aligned} f(x_0 + h) &= \sum_{k=0}^N \frac{h^k}{k!} f^{(k)}(x_0) + \frac{h^{N+1}}{(N+1)!} f^{(N+1)}(x_0 + \theta h) \\ f(x_0 + h) &= P_N(x_0; h) + R_{N+1}(x_0; h), \end{aligned}$$

hvor  $0 < \theta < 1$  er et tal, der afhænger af både  $x, h$  og  $f$ .

### 3 Nulpunktsbestemmelse

Der skal bestemmes nulpunkt:

$$f(x) = 0.$$

#### 3.1 Bisektionsmetoden

Til nulpunktsbestemmelse skal der anvendes en numerisk tilgang, hvor der anvendes iterative metoder, der konvergerer mod en løsning.

##### **Theorem 3.1 (Mellemsværdisætningen)**

Antag at  $f$  er kontinuert på  $[a; b]$ . Hvis  $u$  er et tal mellem  $f(a)$  og  $f(b)$ , så eksisterer der et  $c \in [a; b]$ , sådan at  $f(c) = u$ .

Antag at  $f$  er kontinuert på et interval  $[a; b]$  og opfylder  $f(a)f(b) < 0$  (så er den ene funktionsværdi negativ og den anden positiv). Så er der ifølge mellemsværdisætningen mindst én løsning til ligningen  $f(x) = 0$ .

Der sættes nu et midtpunkt  $m = \frac{a+b}{2}$  og sætter  $b = m$  hvis  $f(a)f(m) \leq 0$  og ellers  $a = m$ . Proceduren gentages indtil der er fundet et interval, hvor længden af intervallet  $|b - a|$  er mindre end en given tolerance  $\epsilon$ .

##### 3.1.1 Konvergens af Bisektionsmetoden

##### **Theorem 3.2 (Konvergens af bisektionsmetoden)**

Antag at  $f$  er kontinuert på  $[a; b]$ . Hvis  $f(a)f(b) < 0$ , så konvergerer bisektionsmetoden mod en løsning til ligningen  $f(x) = 0$ .

Bevis kan findes i slides for tredje kursgang.

##### 3.1.2 Fejlvurdering af Bisektionsmetoden

Fejlen for den  $n$ 'te iteration defineres som

$$b_n - a_n = \frac{b_{n-1} - a_{n-1}}{2} = \dots = \frac{b_0 - a_0}{2^n}.$$

Fejlen halveres ved hver iteration, og bisektionsmetoden siges derfor at konvergere *lineært* (konvergerer langsomt sammenlignet med andre metoder).

#### 3.2 Funktionsiterationsmetoder

Antag at  $f(x) = 0$  kan omskrives til en fixpunkt ligning

$$g(x) = x.$$

En løsning  $s$  til  $f(x) = 0$  vil dermed opfylde  $g(s) = s$ . Der gættes nu på en løsning  $x_0$  og iterativt definerer

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, \dots$$

Hvis  $g$  er kontinuert, og følgen  $\{x_n\}_{n \geq 0}$  konvergerer, så vil elementerne i følgen nærme sig hinanden  $x_{n+1} \approx x_n$ . Det vil sige  $g(x_n) \approx x_n$ , hvilket betyder  $x_n$  udgør en approksimation til løsningen  $s$ .

### 3.2.1 Konvergens

#### Theorem 3.3 (Middelværdisætningen)

Antag at  $f$  er kontinuert på et interval  $I = [a; b]$  og differentiabel på  $(a; b)$ . Så eksisterer der et  $c \in I$ , sådan at

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

#### Theorem 3.4 (Konvergens af Funktionsiterationsmetoder)

Antag at  $g$  er differentiabel på intervallet  $I = [a; b]$ , hvor  $c \in I$  samt at

1.  $g(c) \subseteq I$ .
2.  $|g'(x)| \leq K \leq 1 \quad \forall x \in I$ .

Så har ligningen  $g(x) = x$  en unik løsning i intervallet  $[a, b]$  og

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, \dots$$

konvergerer mod denne løsning for alle  $x_0 \in I$ . (Dette er en tilstrækkelig betingelse)

Bevis kan findes i slides for tredje kursusgang.

### 3.2.2 Fejlvurdering af Funktionsiterationsmetoder

Ved at lave en Taylorudvikling omkring  $s$  får vi

$$\begin{aligned} e_{n+1} &= |x_{n+1} - s| = |g(x_n) - s| \\ &= \left| g(s) + (x_n - s)g'(s) + \frac{(x_n - s)^2}{2}g''(s) + \dots - s \right| \\ &= \left| (x_n - s)g'(s) + \frac{(x_n - s)^2}{2}g''(s) + \dots \right|. \end{aligned}$$

Hvis  $e_n$  er tilstrækkelig lille, så kan vi lave følgende approksimation

$$e_{n+1} \approx |(x_n - s)g'(s)| = e_n |g'(s)|.$$

Dermed konvergerer iterationsmetoden som udgangspunkt lineært. Hvis  $g'(s) = 0$ , så er der hurtigere konvergens.

## 4 Newtonsmetode og Sekantmetoden

### 4.1 Newtons Metode

En første ordens Taylor approksimation anvendes til at bestemme funktionsværdierne for punkter nær nulpunktet  $f(s) = 0$ :

$$f(x) \approx f(x_0) + (x - x_0)f'(x_0),$$

ud fra denne approksimation kan en ny første ordens Taylorudvikling findes, som nu er tættere på nulpunktet

$$\begin{aligned} f(x_0) + (x_1 - x_0)f'(x_0) &= 0 & \Updownarrow \\ x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)}, & f'(x_0) \neq 0. \end{aligned}$$

Taylor approksimationen sættes lig 0 og der findes en ny  $x$  værdi.

Dette leder frem til **Newtons iterationsformel**:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \dots$$

Newtons metode kan anvendes til at bestemme kvadratrødder, givet  $c > 0$ , så kan  $\sqrt{c}$  bestemmes ved at løse ligningen

$$f(x) = x^2 - c = 0,$$

hvilket giver iterationsformlen

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2}{2x_n} = \frac{x_n + c/x_n}{2}.$$

#### Theorem 4.1 (Tilstrækkelige krav for konvergens)

Lad  $f$  være to gange differentiabel på  $[a; b]$  og opfylde

1.  $f(a)f(b) < 0$ .
2.  $f'$  har ingen nulpunkter på  $[a; b]$ . (Newtons metode er ikke defineret)
3.  $f''$  skifter ikke fortegn på  $[a; b]$ . (Krumning)
4.  $\left| \frac{f(a)}{f'(a)} \right|, \left| \frac{f(b)}{f'(b)} \right| < b - a$ .

Så har  $f(x) = 0$  en unik løsning  $s \in (a, b)$ , og Newtons metode konvergerer mod  $s \forall x_0 \in [a, b]$ .

#### 4.1.1 Fejlvurdering af Newtons metode

Fejlen ved Newtons metode er givet ved Taylorudviklingen

$$g'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2}, \quad g(x) = x - \frac{f(x)}{f'(x)}.$$

Det vil sige  $g'(s) = 0$ , så fejlen kan approksimeres ved

$$e_{n+1} \approx \left| \frac{(x_n - s)^2}{2} g''(s) \right| = \frac{|g''(s)|}{2} e_n^2.$$

## 4.2 Sekantmetoden

For at kunne anvende Newtons iterationsmetode skal funktionens afledede eksistere og være forskellig fra nul. Sekantmetoden er langsommere end Newtons metode, men kræver ikke at den afledede eksisterer.

Antag at  $f(s) = 0$  og betragt sekanten gennem to nabopunkter  $x_0$  og  $x_1$ . Hældning er givet ved

$$a = \frac{f(x_1) - f(x_0)}{x_1 - x_0},$$

og skæringen med y-aksen er givet ved

$$b = f(x_1) - ax_1 = f(x_1) - \frac{f(x_1) - f(x_0)}{x_1 - x_0} x_1.$$

Ligningen for sekanten er dermed

$$y = ax + b = \frac{f(x_1) - f(x_0)}{x_1 - x_0} (x - x_1) + f(x_1).$$

Dette kan anvendes til at finde den næste approksimation ved at sætte højresiden lig nul. Det leder frem til **sekantmetodens iterationsformel**:

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n).$$

Sekantmetoden svarer til Newtons metode med approksimationen

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

Sekantmetoden konvergerer superlineært:

$$e_{n+1} = |x_{n+1} - s| \approx C e_n^\alpha,$$

hvor  $C$  er en konstant og  $\alpha = \frac{1+\sqrt{5}}{2} \approx 1.6$ . Sekantmetoden konvergerer altså hurtigere end bisektionsmetoden, men langsommere end Newtons metode.

## 5 Interpolation og numerisk differentiation

Givet et datasæt med  $N + 1$  datapunkter

$$(x_0, y_0), (x_1, y_1), \dots, (x_N, y_N), \quad x_j \neq x_k, j \neq k,$$

da skal der findes et interpolerende polynomium som opfylder

$$p(x_j) = y_j \quad \forall j = 0, 1, \dots, N.$$

Graden af  $p(x)$  skal holdes lav for at minimere kompleksiteten af approksimationen.

### Theorem 5.1 (Weierstrass' Approksimationssætning)

Lad  $f$  være en kontinuert funktion på  $[a; b]$ . Givet  $\epsilon > 0$ , så eksisterer der et polynomium  $p_{N(\epsilon)}$  af grad  $N(\epsilon)$ , sådan at

$$|f(x) - p_{N(\epsilon)}(x)| < \epsilon \quad \forall x \in [a; b].$$

Sætningen er en væsentlig grund til at approksimere med polynomier, da det fungerer, og polynomier nemt kan integreres og differentieres. Polynomier kan effektivt evalueres med Horner's skema. Man kan direkte opskrive en følge af polynomier med den angivne konvergensgenskab. Disse polynomier kaldes Bernstein polynomier og er beregningsmæssigt tunge at bestemme for en vilkårlig funktion.

**Horner's skema:** Et polynomie kan omskrives som

$$\begin{aligned} p(x) &= a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \\ p(x) &= a_0 + x(a_1 + x(\dots + x(a_{n-1} + x a_n))) \end{aligned}$$

### 5.1 Lagrange Polynomier

Lagrange basis polynomier defineres som

$$l_k(x) = \frac{\prod_{j=0, j \neq k}^N (x - x_j)}{\prod_{j=0, j \neq k}^N (x_k - x_j)}, \quad k = 0, 1, \dots, N.$$

Hvert polynomium  $l_k(x)$  har grad præcis  $N$  og opfylder

$$l_k(x_j) = \begin{cases} 1, & k = j \\ 0, & k \neq j \end{cases}.$$

Lagrange interpolerende polynomium er da givet som

$$p(x) = \sum_{k=0}^N y_k l_k(x).$$

### Theorem 5.2 (Entydighed)

Givet datapunkter

$$(x_0, y_0), (x_1, y_1), \dots, (x_N, y_N), \quad x_j \neq x_k, j \neq k,$$

så eksisterer der ét og kun ét polynomium  $p(x)$  af grad højst  $N$  med interpolationsegenskaben  $p(x_j) = y_j$  for alle  $j$ .

Interpolation er nyttigt til at modellere et givent datasæt. Vi kan med det interpolerende polynomium approksimere mellemliggende værdier og forudsige fremtidige værdier.

### Theorem 5.3 (Fejl ved Lagrange interpolation)

Antag at  $f$  er  $N + 1$  gange kontinuert differentiabel på intervallet  $[a; b]$ , der indeholder  $x_0, x_1, \dots, x_N$ . Lad  $p$  være Lagrange interpolation polynomiet baseret på

$$(x_0, f(x_0)), (x_1, f(x_1)), \dots, (x_N, f(x_N)).$$

Da gælder for alle  $x \in [a; b]$  at

$$f(x) - p(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_N)}{(N + 1)!} f^{(N+1)}(\xi),$$

hvor  $\xi \in [a; b]$ .

## 5.2 Numerisk differentiation

Ideen er at bestemme det interpolerende polynomium gennem et par nabopunkter, og så lave approksimationen

$$f'(x_0) \approx p'(x_0).$$

Der kigges på topunkts differensformler, hvor man anvender et enkelt nabopunkt  $x_0 + h$  eller  $x_0 - h$ . Differentiationen findes med Lagranges interpolerende polynomium, hvilket med ét nabopunkt svarer til hældningen for en ret linje.

### Theorem 5.4 (Første ordens differensformler)

Første ordens fremadgående differensformel:  $f'(x_0) \approx \frac{f(x_0+h) - f(x_0)}{h}$ .

Første ordens tilbagegående differensformel:  $f'(x_0) \approx \frac{f(x_0) - f(x_0-h)}{h}$ .

Fejlen er givet ved Taylors formel:

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2} f''(x_0 + \theta h), \theta \in (0; 1)$$

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} + \mathcal{O}(h).$$



For den tilbagegående differens gælder

$$f(x_0 - h) = f(x_0) - hf'(x_0) + \frac{h^2}{2}f''(x_0 + \theta h), \theta \in (0; 1)$$
$$f'(x_0) = \frac{f(x_0) - f(x_0 - h)}{h} + \mathcal{O}(h).$$

## 6 Numerisk differentiation og interpolerende kvadraturregler

Der undersøges de to nærmeste symmetriske nabopunkter og differentiationen approksimeres ud fra de to nabopunkter.

$$(x_0 - h, y_{-1}), (x_0, y_0), (x_0 + h, y_1).$$

Lagrange polynomiet er givet ved

$$\begin{aligned} p(x) &= y_{-1}l_{-1}(x) + y_0l_0(x) + y_1l_1(x) \\ &= y_{-1}\frac{(x - x_0)(x - x_0 - h)}{2h^2} + y_0\frac{(x - x_0 + h)(x - x_0 - h)}{-h^2} + y_1\frac{(x - x_0 + h)(x - x_0)}{2h^2}. \end{aligned}$$

### Theorem 6.1 (Differentiation af Lagrange Polynomie)

Den første afledte er givet ved

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

Den anden afledte er givet ved

$$f''(x_0) \approx \frac{f(x_0 + h) - 2f(x_0) + f(x_0 - h)}{h^2}.$$

Fejleddet findes med Taylors formel som

$$\begin{aligned} \frac{f(x_0 + h) - f(x_0 - h)}{2h} &= f'(x_0) + \frac{h^2}{12} \left( f^{(3)}(x_0 + \theta_1 h) + f^{(3)}(x_0 + \theta_2 h) \right) \\ &= f'(x_0) + \mathcal{O}(h^2), \end{aligned}$$

hvilket viser at det er en anden ordens metode, for hvilken fejlen minimeres ved lave  $h$  værdier, antaget at den tredje afledte opfører sig "pænt".

Der kan også undersøges to assymetriske nabopunkter:

$$(x_0, y_0), (x_0 + h, y_1), (x_0 + 2h, y_2).$$

De afledte er da fundet med Lagrange polynomiet og er givet i følgende sætning.

### Theorem 6.2

De afledte for fremmadgående assymetriske nabopunkter er givet ved

$$\begin{aligned} f'(x_0) &= \frac{-f(x_0 + 2h) + 4f(x_0 + h) - 3f(x_0)}{2h} \\ f''(x_0) &= \frac{f(x_0 + 2h) - 2f(x_0 + h) - f(x_0)}{h^2}. \end{aligned}$$

De afledte for tilbagegående assymetriske nabopunkter er givet ved

$$f'(x_0) = \frac{3f(x_0) - 4f(x_0 - h) + f(x_0 - 2h)}{2h}$$
$$f''(x_0) = \frac{f(x_0) - 2f(x_0 - h) + f(x_0 - 2h)}{h^2}.$$

Bemærk at det kun er de dobbelt afledte der er anden ordens metoder, mens de første ordens afledte er første ordens metoder.

## 7 Numerisk integration

Givet en kontinuert funktion  $f : [a; b] \rightarrow \mathbb{R}$  skal vi approksimere

$$\int_a^b f(x) \, dx.$$

Intervallat  $[a; b]$  inddeles i  $N + 1$  ækvidistante punkter og det Lagrange interpolerende polynomium bestemmes ved

$$p(x) = \sum_{k=0}^N f(x_k) l_k(x).$$

Der udføres herefter approksimationen

$$\int_a^b f(x) \, dx \approx \int_a^b p(x) \, dx,$$

hvilket kaldes en interpolerende kvadraturregel.

Dette udledes for ét punkt  $x_0 = \frac{a+b}{2}$ . Det interpolerende polynomium er  $p(x) = f\left(\frac{a+b}{2}\right)$  og den interpolerende kvadraturregel er

$$\begin{aligned} \int_a^b p(x) \, dx &= \int_a^b f\left(\frac{a+b}{2}\right) \, dx = f\left(\frac{a+b}{2}\right) \int_a^b 1 \, dx \\ &= f\left(\frac{a+b}{2}\right) [x]_a^b = (b-a) f\left(\frac{a+b}{2}\right). \end{aligned}$$

Dette kaldes midpunktsreglen, som betyder at

$$\int_a^b f(x) \, dx \approx (b-a) f\left(\frac{a+b}{2}\right).$$

Helt generelt for kvadraturreglen gælder:

$$\begin{aligned} \int_a^b p(x) \, dx &= \int_a^b \sum_{k=0}^N f(x_k) l_k(x) \, dx \\ &= \sum_{k=0}^N f(x_k) \int_a^b l_k(x) \, dx = \sum_{k=0}^N c_k f(x_k), \end{aligned}$$

hvor  $c_k = \int_a^b l_k(x) \, dx$ . Koefficienterne  $c_k$  kaldes for vægte.

For midtpunktsreglen er vægten givet ved  $c_0 = (b-a) = 2\tilde{h}$ , hvor  $\tilde{h} = \frac{b-a}{2}$  er skridtlængden.

## 8 Interpolerende Kvadraturregler

For en funktion  $f(x)$  skal integralet findes, men dette kan ikke gøres analytisk. Der findes nogen datapunkter, for hvilke det interpolerende polynomium  $p_N(x)$  opfylder

$$p_N(x_j) = f(x_j), \quad j = 0, 1, 2, \dots, N.$$

Der er  $N + 1$  datapunkter. Der udføres her approksimationen

$$\int_a^b f(x) dx \approx \int_a^b p_N(x) dx.$$

Når  $N = 0$  er der ét datapunkt og midtpunktssætningen anvendes, hvor  $f\left(\frac{a+b}{2}\right)$ , (Midtpunktsregel -  $M$ )

$$M = (b - a)f\left(\frac{a + b}{2}\right).$$

Når  $N = 1$  er der to datapunkter,  $a, f(a), b, f(b)$  og det interpolerende polynomium har grad 1, altså en ret linje, (Trapezregel -  $T$ )

$$T = \frac{b - a}{2}(f(a) + f(b))$$

Når  $N = 2$  er der tre datapunkter  $a, f(a), \frac{a+b}{2}, f\left(\frac{a+b}{2}\right), b, f(b)$  og det interpolerende polynomium har grad 2, altså et andengrads polynomium, (Simpsonsregel -  $S$ )

$$S = \frac{b - a}{6} \left( f(a) + 4f\left(\frac{a + b}{2}\right) + f(b) \right).$$

Det er de tre kvadraturregler der skal kendes samt deres grafer.

### 8.1 Fejl-udtryk

Midtpunktsreglen:

$$(M) \int_a^b f(x) dx - (b - a)f\left(\frac{a + b}{2}\right) = \frac{(b - a)^3}{24} f^{(2)}(\xi_M), \xi_M \in [a, b].$$

Trapezreglen:

$$(T) \int_a^b f(x) dx - \frac{b - a}{2}(f(a) + f(b)) = -\frac{(b - a)^3}{12} f^{(2)}(\xi_T), \xi_T \in [a, b]$$

Simpsonsregel:

$$(S) \int_a^b f(x) dx - \frac{b - a}{6} \left( f(a) + 4f\left(\frac{a + b}{2}\right) + f(b) \right) = -\frac{(b - a)^5}{2880} f^{(4)}(\xi_S), \xi_S \in [a, b]$$

## 8.2 Sammensat Kvadratur

Runge effekten er problematisk når man skal bestemme det interpolerende polynomium når der er mange datapunkter. Af den årsag skal der anvendes sammensat kvadratur.

Der anvendes et interval  $I = [a, b]$  med  $N + 1$  ækvivalente punkter, hvilket anvender skridtlængden  $h = x_{j+1} - x_j = \frac{b-a}{N}$ . Da er trapezreglen angivet som

$$T_N = \frac{h}{2}(f(x_0)+f(x_1))+\frac{h}{2}(f(x_1)+f(x_2))+\cdots+\frac{h}{2}(f(x_{N-1})+f(x_N)), x_0, x_1, x_2, \dots, x_N \in I.$$

Der er en fælles faktor

$$T_N = \frac{h}{2} (f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{N-1}) + f(x_N)),$$

som skrives kort som

$$T_N = \frac{h}{2} \left( f(x_0) + 2 \left( \sum_{j=1}^{N-1} f(x_j) \right) + f(x_N) \right).$$

**Fejlvurdering:**

$$\begin{aligned} \int_a^b -T_N &= -\frac{h^3}{12}f^{(2)}(\xi_0) - \frac{h^3}{12}f^{(2)}(\xi_1) - \frac{h^3}{12}f^{(2)}(\xi_2) - \cdots - \frac{h^3}{12}f^{(2)}(\xi_{N-1}) \\ &= -\frac{h^3}{12} \sum_{j=0}^{N-1} f^{(2)}(\xi_j). \end{aligned}$$

Påstand: Der findes et  $\xi \in [a, b]$ , så at

$$\sum_{j=0}^{N-1} f^{(2)}(\xi_j) = N f^{(2)}(\xi).$$

**Bevis:**  $f^{(2)}(x)$  er kontinuert på  $I = [a, b]$  (antagelse). Der findes  $\alpha, \beta$  således  $\alpha \leq f^{(2)}(x) \leq \beta, x \in [a, b]$ .

$$\alpha \leq \frac{1}{N} \sum_{j=0}^{N-1} f^{(2)}(\xi_j) \leq \beta,$$

så ifølge sætningen om mellemværdien af kontinuerte funktioner, findes der et  $\xi \in I$ , så  $f^{(2)}(\xi) = \frac{1}{N} \sum_{j=0}^{N-1} f^{(2)}(\xi_j)$ .

## 9 Sammensatte Kvadraturregler

Et integrale

$$I = \int_a^b f(x) dx$$

skal approximeres med sammensatte kvadraturregler. Hertil har vi midtpunktsreglen med  $a = x_0 < x_1 < \dots < x_N = b$ ,  $x_{j+1} - x_j = \frac{b-a}{N}$ , hvor integralet for et interval er  $(b-a)f\left(\frac{b+a}{2}\right)$ . Det sammensatte kvadraturregel  $M_N$  er da:

$$M_N = h \sum_{j=0}^{N-1} f\left(x_0 + \left(j + \frac{1}{2}\right)h\right).$$

Der er også trapezreglen med integralet for et interval  $\frac{b-a}{2}(f(a) + f(b))$ . Det sammensatte kvadraturregel er

$$T_N = \frac{h}{2} \left( f(x_0) + 2 \sum_{j=1}^{N-1} f(x_0 + hj) + f(x_N) \right).$$

Til sidst er der simpsonsreglen med den sammensatte kvadraturregel givet ved

$$S_N = \frac{h}{6} \sum_{j=0}^{N-1} \left( f(x_j) + 4f\left(\frac{x_j + x_{j+1}}{2}\right) + f(x_{j+1}) \right).$$

Bemærk at  $N$  er antallet af delintervaller.

### 9.1 Fejlvurderinger

Fejlvurderingerne er givet ved

$$\begin{aligned} I - M_N &= \frac{(b-a)h^2}{24} f^{(2)}(\xi_M) \\ I - T_N &= -\frac{(b-a)h^2}{12} f^{(2)}(\xi_T) \\ I - S_N &= -\frac{(b-a)h^4}{2880} f^{(4)}(\xi_S). \end{aligned}$$

Vurdering af nødvendigt antal  $N$  for at få  $|I - M_N| < \epsilon$  for  $\epsilon$  værende tolerancen.

$$|I - M_n| \leq \frac{(b-a)h^2}{24} |f^{(2)}(\xi_M)|$$

Antag at vi kan finde  $M_2$ , så  $|f^{(2)}(\xi)| \leq M_2$  for alle  $\xi \in [a, b]$ . Bestem  $N$ , så at

$$\begin{aligned} \frac{(b-a)^3 M_2}{24N^2} &< \epsilon \\ N^2 &> \frac{(b-a)^3 M_2}{24\epsilon} \\ N &> \sqrt{\frac{(b-a)^3 M_2}{24\epsilon}} \end{aligned}$$

**Example 9.1**

$$\int_1^2 \frac{x}{1+x^2} dx = \left[ \frac{1}{2} \ln(1+x^2) \right]_1^2 = \ln \left( \sqrt{\frac{5}{2}} \right)$$

Det integrale skal approksimeres numerisk. Funktionen er da  $f(x) = \frac{x}{1+x^2}$ ,  $f'(x) = \frac{1}{1+x^2} + x(-1)\frac{2x}{(1+x^2)^2} = \frac{1-x^2}{(1+x^2)^2}$ . Den anden afledede er

$$f^{(2)}(x) = \frac{-2x}{(1+x^2)^2} + (1-x^2)(-2)\frac{2x}{(1+x^2)^3} = \frac{-2x(1+x^2) - 4(x-x^3)}{(1+x^2)^3} = \frac{-6x+2x^3}{(1+x^2)^3}.$$

Der skal bestemmes en  $\xi$ :

$$|f^{(2)}(x)| \leq \frac{6x+2x^3}{(1+x^2)^3},$$

som skal vurderes for  $x \in [1, 2]$ :

$$|f^{(2)}(x)| \leq \frac{6x+2x^3}{(1+x^2)^3} \leq \frac{28}{8} = \frac{7}{2}.$$

Tolerancen vælges til at være  $\epsilon = 10^{-4}$  og vi husker på  $a = 1, b = 2, b - a = 1$ ,  $N$  kan da bestemmes

$$N > \sqrt{\frac{7}{2 \cdot 24 \cdot 10^{-4}}} = 10^2 \sqrt{\frac{7}{48}} \approx 54.$$

Givet  $\int_a^b f(x) dx$  som vi ikke kan finde  $M_4, |f^{(4)}(x)| \leq M_4, x \in [a, b]$ , med en tolerance  $\epsilon > 0$ . Stop med Simpson ved  $|S_{2N} - S_N| < \epsilon$ .

**Proof 9.1**

Sammenlign  $S_N$  og  $S_{2N}$ . Antallet af delintervaller er fordoblet og dvs. at ifølge fejlvurderingen er fejlen blevet en  $\frac{1}{16}$  mindre. Derfor er det rimeligt at sige at de to fejle  $I - S_N$  og  $16(I - S_{2N})$  nogenlunde er det samme.

$$\begin{aligned} I - S_N &\approx 16(I - S_{2N}) = 15(I - S_{2N}) + (I - S_{2N}) \\ S_{2N} - S_N &\approx 15(I - S_{2N}), \text{ hvis } |S_{2N} - S_N| \leq \epsilon \\ \text{så er } |I - S_{2N}| &\simeq \frac{1}{15}\epsilon \\ I &\approx \frac{16S_{2N} - S_N}{15} = Q_N. \end{aligned}$$

**Udregning:**

$$S_2 \quad S_1, \quad \tilde{h} = \frac{b-a}{4} (h = \frac{b-a}{2N}, \text{ ændrer lidt på kvadraturreglerne})$$

Der anvendes Simpsons kvadraturregel på et interval med to delintervaller ( $S_1$ ) og på et

interval med fire delintervaller ( $S_2$ ), hvortil  $x_j = x_0 + hj$ .

$$S_1 = \frac{4\tilde{h}}{6} (f(x_0) + 4f(x_2) + f(x_4))$$
$$S_2 = \frac{2\tilde{h}}{6} (f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + f(x_4))$$

Nu vil intervallet præsenteret før udregningerne, beregnes.

$$\begin{aligned} 16S_2 - S_1 &= \frac{2\tilde{h}}{3} (8f(x_0) + 32f(x_1) + 16f(x_2) + 32f(x_3) + 8f(x_4)) \\ &\quad - \frac{2\tilde{h}}{3} (f(x_0) + 4f(x_2) + f(x_4)) \\ &= \frac{2\tilde{h}}{3} (7f(x_0) + 32f(x_1) + 12f(x_2) + 32f(x_3) + 7f(x_4)) \\ Q_N &= \frac{2\tilde{h}}{45} (7f(x_0) + 32f(x_1) + 12f(x_2) + 32f(x_3) + 7f(x_4)). \end{aligned}$$

Bemærk at dette er en fempunkts kvadraturregel. Fejlvurderingen på denne kvadraturregel er

$$I - Q_N = c\tilde{h}f^{(6)}(\xi_Q),$$

og er derved en sjette ordens metode. Ved at gøre dette systematisk kaldes det Romberg integration.

## 10 Differentialligninger

Der kigges typisk på differentialligninger af typen

$$\frac{dy}{dx} = f(x, y), y(x_0) = x_0.$$

### Example 10.1

$$\frac{dy}{dx} = 3x^2y, y(0) = 1 : \text{Løsning } y(x) = e^{x^3}$$

Hvad siger en differentialligning? Hvis jeg et punkt kender jeg hældningen for kurven til det punkt. Man går ind og laver et retningsfelt – en masse vektorer med samme længde, hvis retning er angivet af punkterne. Til ét punkt  $(x_0, y_0)$  kan hældningen bestemmes, men



til punktet  $(y(x_0 + h), x_0 + h)$  skal  $y$  approksimeres med Taylor:

$$\begin{aligned}y(x_0 + h) &= y(x_0) + hy'(x_0) + \mathcal{O}(h^2) \\&= y_0 + hf(x_0, y_0) + \mathcal{O}(h^2) \\&\approx y_0 + hf(x_0, y_0) \\y_1 &= y_0 + hf(x_0, y_0) \\y_{n+1} &= y_n + hf(x_n, y_n) \quad \text{Eulers metode.}\end{aligned}$$

## 11 Runga-Kutta Metode: RK4

RK4:

$$\begin{aligned}y_{n+1} &= y_n + \frac{h}{6} (k_1 + 2(k_2 + k_3) + k_4) \\y' &= f(x, y), \quad y(x_0) = y_0 \\k_1 &= f(x_n, y_n) \\k_2 &= f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_1\right) \\k_3 &= f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_2\right) \\k_4 &= f(x_n + h, y_n + hk_3).\end{aligned}$$

Der anvendes en vægtet hældning af  $k_1, k_2, k_3$  og  $k_4$  til at beregne funktionsværdien af næste punkt. Hældningen er derved  $\frac{1}{6}(k_1 + 2(k_2 + k_3) + k_4)$ .

Den lokale trunkeringsfejl opstår grundet approximationen og er  $\mathcal{O}(h^5)$  mens den globale trunkeringsfejl er  $\mathcal{O}(h^4)$  og det er derved en fjerde ordens metode.

### Example 11.1

Vi har to differentiaalligninger

$$\begin{aligned}y_1' &= f_1(x, y_1, y_2), \quad y_1(x_0) = y_{1,0} \\y_2' &= f_2(x, y_1, y_2), \quad y_2(x_0) = y_{2,0} \\ \vec{y} &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \vec{y}(x_0) = \vec{y}_0 = \begin{bmatrix} y_{1,0} \\ y_{2,0} \end{bmatrix} \\ \vec{f}(x, \vec{y}) &= \begin{bmatrix} f_1(x, y_1, y_2) \\ f_2(x, y_1, y_2) \end{bmatrix} \\ \vec{y}' &= \vec{f}(x, \vec{y}) \\ \vec{y}(x_0) &= \vec{y}_0\end{aligned}$$

RK4 algoritmen er vektoriseret.

Løsning visualiseres som en kurve i  $y_1, y_2$ -planen.

$$x \rightarrow (y_1(x), y_2(x))$$

Variablen  $x$  svarer til tiden  $t$  og derved beskrives hastigheden af dette system.

Et autonomt (tidsinvariant) system af differentiaalligninger opstår ikke eksplicit på højresiden:

$$\begin{aligned}\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ x_1' &= ax_1 + bx_2 \\ x_2' &= cx_1 + dx_2\end{aligned}$$

## 12 Differentialligninger

Differentialligninger kan skrives som

$$\frac{dy}{dx} = f(x, y),$$

og beskriver den afledede af en ukendt funktion. Ofte er der givet et begyndelseskriterie

$$y(x_0) = y_0.$$

Typisk pålægges funktionen den betingelse at være kontinuert og have partielt afledede for at sikre eksistens.

Der kigges på differentialligningen

$$\frac{dy}{dx} = F(x)G(y),$$

hvor vi isolerer  $y$  på den ene side og  $x$  på den anden:

$$\begin{aligned}\frac{1}{G(y)} dy &= F(x) dx \\ \int G(y) dy &= \int F(x) dx + c.\end{aligned}$$

Eksempel:

$$\begin{aligned}\frac{dy}{dx} &= xy \\ \frac{1}{y} dy &= x dx \\ \log(y) &= \frac{1}{2}x^2 + \tilde{c} \\ y &= e^{\frac{1}{2}x^2 + \tilde{c}} = e^{\tilde{c}} e^{\frac{1}{2}x^2} = ce^{\frac{1}{2}x^2}.\end{aligned}$$

### 12.1 Eulers Metode

Vi vil gerne finde løsningen et skridtlængde  $h$  længere fremme og vil gerne kende

$$y(x_0 + h) = y(x_0) + hy'(x_0) + \frac{1}{2}h^2 y''(x_0) + \dots$$

givet af en Taylor udvikling. Det kan approksimeres til

$$\begin{aligned}y(x_0 + h) &\approx y(x_0) + hy'(x_0) \\ &= y_0 + hf(x_0, y_0) \\ y_1 &= y_0 + hf(x_0, y_0).\end{aligned}$$

Det kan gøres iterativt og giver derved formlerne

$$\begin{aligned}x_{n+1} &= x_n + h \\ y_{n+1} &= y_n + hf(x_n, y_n).\end{aligned}$$

Dette er en simpel og ineffektiv metode. Det skaber typisk lokale trunkeringsfejl og efter  $n$  iterationer er der akkumuleret  $n$  fejl. Den globale fejl svarer til  $Nh^2 = (b-a)h$ ,  $h = \frac{b-a}{N}$ , hvor  $a, b$  svarer til interval endepunkterne  $[a, b]$  og  $N$  er antallet af punkter i intervallet. Forveksel ikke trunkeringsfejl med afrundingsfejl. Metoden konvergerer mod den sande løsning når  $N \rightarrow \infty$ .

## 12.2 Runge-Kutta Metoder

Givet er en differentialligning

$$\begin{aligned}y' &= f(x, y) \\ y(x_0) &= y_0 \\ y(x_0 + h) &= y(x_0) + hy'(x_0) + \frac{1}{2}h^2y''(x_0) + \mathcal{O}(h^3) \\ \text{Vælger } \alpha \in [0, 1], \quad y''(x_0) &\approx \frac{y'(x_0 + \alpha h) - y'(x_0)}{\alpha h} \\ &= y_0 + hf(x_0, y_0) + ?\end{aligned}\tag{1}$$

Vi anvender

$$\begin{aligned}y'(x_0) &= f(x_0, y_0) = k_1 \text{ (bemærk at der er forskel på konstanterne)} \\ y'(x_0 + \alpha h) &= f(x_0 + \alpha h, y(x_0 + \alpha h))\end{aligned}$$

så vi approksimerer  $y(x_0 + \alpha h) \approx y_0 + \alpha h k_1$  og danner en ny konstant

$$\begin{aligned}f(x_0 + \alpha h, y_0 + \alpha h k_1) &= k_2 \\ y''(x_0) &\approx \frac{k_2 - k_1}{\alpha h}.\end{aligned}$$

Nu kan differentialligningen fra ligning(1) bestemmes som:

$$y_1 = y_0 + k_1 + \frac{1}{2}h^2 \frac{k_2 - k_1}{\alpha h} = y_0 + h \left( k_1 \left( 1 - \frac{1}{2\alpha} \right) + \frac{1}{2\alpha} k_2 \right).$$

Hvilket er en god metode.

$\alpha = \frac{1}{2}$  kaldes for korrigeret Eulers metode, da nogle led går ud og den da vil være

$$y_{n+1} = y_n + hk_2, \quad k_2 = f(x_0 + 1/2h, y_0 + 1/2hk_1)$$

$\alpha = 1$  kaldes for modificeret Eulers metode som giver

$$y_{n+1} = y_n + \frac{1}{2}h(k_1 + k_2), \quad k_2 = f(x_0 + h, y_0 + hk_1).$$

$\alpha = \frac{2}{3}$  kaldes for Heuns metode og giver

$$y_{n+1} = y_n + \frac{1}{4}h(k_1 + 3k_2), \quad k_2 = f(x_0 + 2/3h, y_0 + 2/3hk_1).$$

Den lokale fejl er  $h^3$  men den globale trunkeringsfejl er  $h^2$  og de er derved en anden ordens metode.

Anden ordens runge-kutta metoder er baseret på at lave anden ordens trunkeret taylor approksimationer.

## 12.3 Integration og differentialligninger

$$y' = g(x)$$

Hvor højresiden ikke afhænger af  $y$  ved vi at

$$y_1 = y_0 + \int_{x_0}^{x_0+h} g(s) ds \approx y_0 + hg\left(x_0 + \frac{1}{2}h\right) \text{ Midtpunktsreglen}$$

som svarer til korrigeret Eulers metode. Korrigeret Eulers metode er en generaliseret version af midtpunktsreglen.

Hvis Trapezreglen anvendes får man

$$y_1 = y_0 + \int_{x_0}^{x_0+h} g(s) ds \approx y_0 + h \frac{g(x_0 + h) + g(x_0)}{2}$$

som er det samme som modificeret Eulers metode. Modificeret Euler er da en generaliseret version af trapezreglen.

En højere ordens Runge-Kutta metode kaldes **RK4** og er givet ved

$$\begin{aligned} y' &= f(x, y) \\ y(x_0) &= y_0 \\ y_{n+1} &= y_n + \frac{h}{6} (k_1 + 2(k_2 + k_3) + k_4) \end{aligned}$$

hvor konstanterne udover  $k_1$  betyder noget andet end de tidligere:

$$\begin{aligned} k_1 &= f(x_0, y_0) \\ k_2 &= f\left(x_0 + \frac{1}{2}h, y_0 + \frac{1}{2}hk_1\right) \\ k_3 &= f\left(x_0 + \frac{1}{2}h, y_0 + \frac{1}{2}hk_2\right) \\ k_4 &= f(x_0 + h, y_0 + hk_3). \end{aligned}$$

RK4 er en fjerde ordens metode og er en generaliseret version af Simpsonsreglen.