

# Sammanfattning

Michael Sörsäter

## 1 Fel

Typer av fel:

- $R_X$  fel i resultatet, som härrör från fel i indata
- $R_{XF}$  fel i resultatet, som härrör från fel i de använda funktionsvärdena
- $R_B$  avrundningsfel
- $R_T$  trunkeringsfel

Närmevärde till  $x$ :  $\bar{x}$

Absolut fel:  $\Delta x = \bar{x} - x$

Relativt fel:  $\frac{\Delta x}{x}$

### 1.1 Korrekta decimaler

Om  $|\Delta a| \leq 0.5 * 10^{-t}$  sägs  $\bar{a}$  ha  $t$  korrekta decimaler

### 1.2 Fortplantning av fel

Beräkning	Absolut Fel	Relativt fel	Felgräns
$y = x_1 + x_2$	$\Delta y = \Delta x_1 + \Delta x_2$		$ \Delta y  \leq  \Delta x_1  +  \Delta x_2 $
$y = x_1 - x_2$	$\Delta y = \Delta x_1 - \Delta x_2$		$ \Delta y  \leq  \Delta x_1  +  \Delta x_2 $
$y = x_1 * x_2$		$\frac{\Delta y}{y} \approx \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2}$	$ \frac{\Delta y}{y}  \leq \approx  \frac{\Delta x_1}{x_1}  +  \frac{\Delta x_2}{x_2} $
$y = x_1 / x_2$		$\frac{\Delta y}{y} \approx \frac{\Delta x_1}{x_1} - \frac{\Delta x_2}{x_2}$	$ \frac{\Delta y}{y}  \leq \approx  \frac{\Delta x_1}{x_1}  +  \frac{\Delta x_2}{x_2} $

Table 1: Fortplantning av fel

### 1.3 Maximalfelsuppskattning

För att uppskatta felet när en dator räknar utförs beräkningsanalys på varje del. Följande summa skrivs upp:

$$|\Delta f| \leq \approx \sum_{k=1}^n \left| \frac{\delta f}{\delta x_k} \Delta x_k \right|$$

Varje beräkning (+, -, \*, / ...) ersätts av en bokstav. Derivatans beräknas med avseende på sin bokstav:

$$|\Delta f| \leq \left| \frac{\delta f}{\delta a} \right| |\Delta a| \left| \frac{\delta f}{\delta b} \right| |\Delta b| \dots$$

För att uppskatta absoluta felet utnyttjas ( $\mu = \text{maskinkonstant}$ ):

$$\frac{\Delta a}{a} \approx \mu \Leftrightarrow \Delta a \approx a\mu$$

## 2 Talsystem och flyttalsrepresentation

Talsystem beskrivs på formen  $(\beta, t, L, U)$

- $\beta$  är basen
- $t$  är precisionen (antalet decimaler)
- $L$  är undre gränsen på exponenten
- $U$  är övre gränsen på exponenten

Med 32 bitar representeras ett flyttal som:  $(2, 23, -126, 127)$

s (1 bit) positivt/negativt

e (8 bitar) exponenten för talet

f (23 bitar) decimaldelen för talet

$$x = (-1)^s (1.f)_2 * 2^{e-127}$$

Felet från det talet man lagrar till det riktiga talet:

$$\frac{|x - x_r|}{|x|} \leq \frac{1}{2} \beta^{-t}$$

där  $x_r$  är det talet som ligger närmast  $x$ .

Jämför med formeln i 1.1.

## 2.1 Roten ur

För att beräkna roten ur av ett tal.  $x = \sqrt{a}$  används funktionen  $f(x) = x^2 - a$ .

Med Newton-Raphsons fås:  $x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n})$

I flyttalssystemet skrivs tal som:

$$x = (1.f)_2 2^e \Rightarrow \sqrt{x} = \sqrt{(1.f)_2} \sqrt{2^e}$$

Resulterar i två fall:

$$\sqrt{2^e} = \begin{cases} e \text{ jämn, } 2^{\frac{e}{2}} \\ e \text{ udda, } \sqrt{2} 2^{\frac{e-1}{2}} \end{cases}$$

Termen  $(1.f)_2$  ligger alltid i intervallet  $[1, 2]$ . Då  $e$  är udda, multipliceras tvåan in i den roten och uttrycket blir:

$$\sqrt{x} = \sqrt{(1.f)_2} 2^{\frac{e-1}{2}}$$

Vilket gör att samma term ligger i intervallet  $[1, 4]$  och det blir då det enda intervallet som måste beräknas.

## 3 Summa - restterm

Summa som är konvergent

$$S = \sum_{n=1}^{\infty} a_n, \quad S_N = \sum_{n=1}^N a_n$$

$$R_N = S - S_N = \sum_{n=N+1}^{\infty} a_n$$

Hur uppskattas  $R_N$  utan att räkna ut den?

### 3.1 Alternerande

Är serien konvergent och alternerande:

$$|R_N| \leq |a_{N+1}|$$

Alltså, felet är mindre än nästa term.

### 3.2 Postiv monotont avtagande

Om funktionen avtar, går mot 0.

$$R_N = \sum_{N+1}^{\infty} f(n) \leq \int_N^{\infty} f(x) dx$$

Alltså, integrera funktionen och beräkna integralen. Lägg märke till att man beräknar en extra term då man i integralen börjar på N istället för N+1.

## 4 Iterationsmetoder

### 4.1 Konvergerar

För att undersöka om en iterationsmetod konvergerar mot en lösning med ett givet startvärde  $x_0$  måste den uppfylla två krav:

1. Måste kunna skrivas om på originalform
2.  $|\varphi'(x_0)| < 1$  för att den ska konvergera

Exempel:

$$f(x) = xe^x - 1 = 0$$

$$x_{n+1} = \underbrace{\frac{1+x_n}{1+ex_n}}_{\varphi(x)}$$

Genom att multiplicera upp nämnaren och flytta runt fås samma orginalfunktion (krav 1). Genom att derivera och sätta in  $x_0$  fås ett värde som är  $< 1$  (krav 2).

## 4.2 Newton-Raphsons metod

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

## 4.3 Metodoberoende feluppskattning

Används för att uppskatta hur nära  $x$  är  $x^*$ .

$$|x^* - \bar{x}| \leq \frac{f(\bar{x})}{f'(\bar{x})}$$

## 5 Interpolation och splines

### 5.1 Interpolation

Beräkna polynomet:

$$p_n(x) = c_0 + c_1(x - x_1) + c_2(x - x_1)(x - x_2) + \dots$$

Beräkna konstanterna  $C$  genom att först sätta in  $x_1$  (gör alla delar till 0 förutom  $c_0$ ) och sedan vidare likadant.

Felet är den "Extra term" som inte ingår i interpolationen,  $R_T$ . För en linjär interpolation är det termen vid  $c_2$ .

Vid en fullständig feluppskattning ska tre delar tas med. Alla delar är absolutbelopp.

- $R_B$  Avrundningsfel. Efter beräkningen  $P(a)$ , har svaret  $t$  korrekta decimaler.  
Ger felet:  $R_B = 0.5 * 10^{-t}$
- $R_T$  Trunkeringsfel. (är ofta den dominerande delen) Använd termen  $c_n$  för att beräkna felet. Ex:  $R_T \leq |c_2(x - x_1)(x - x_2)|$
- $R_{XF}$  Fel i indatan. Likadant som för  $R_B$  antalet  $t$  decimaler som lägst givna i indatan.  $R_{XF} = 0.5 * 10^{-t}$

Svaret blir sedan  $R_{TOT} = |R_B| + |R_T| + |R_{XF}|$

#### 5.1.1 Fel indata

Ibland har även den punkt som uppskattas en felmarginal:

$$A = \bar{A} \pm \Delta A$$

Vad blir felet i  $f(A)$ ?

Använd maximalsfelsuppskattning.

$$|\Delta f| \leq \left| \frac{\delta f}{\delta A} \right| |\Delta A| \approx \left| \frac{\delta P_n}{\delta x} \right| |\Delta A| \approx \text{/om linjär/} \approx |c_1| |\Delta A|$$

## 5.2 Spline

Låt

$$a = x_1 < x_2 < \dots < x_n < b$$

## 5.3 Lineär spline

En funktion  $s$  sägs vara en lineär splinefunktion på intervallet  $[a, b]$  om:

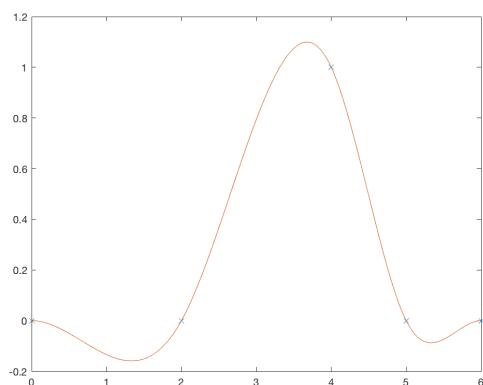
- $s$  är kontinuerlig på  $[a, b]$
- $s$  är en rät linje på varje delintervall  $[x_i, x_{i+1}]$ ,  $i = 1, \dots, n-1$

## 5.4 Kubisk spline

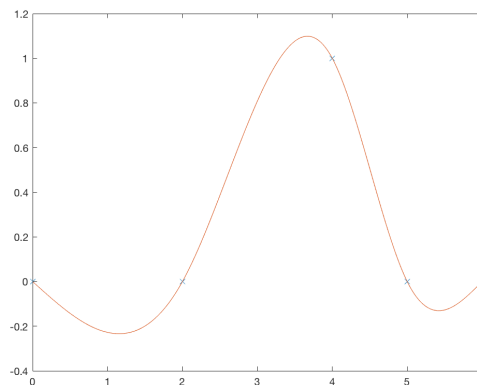
En funktion  $s$  sägs vara en kubisk splinefunktion på intervallet  $[a, b]$  om:

- $s, s', s''$  är kontinuerlig på  $[a, b]$
- $s$  är ett polynom av grad  $\leq 3$  på varje delintervall  $[x_i, x_{i+1}]$ ,  $i = 1, \dots, n-1$

Finns två typer av kubiska splines. Antingen är derivatan i ändpunkterna 0 (rak linje), eller så knyter den ihop med nästa spline. För formler om hur man formeln är, se sida 130 i kursboken.



(a) Derivatan 0



(b) Mjuk övergång till nästa spline

Figur 1: Kubiska Splines

## 6 LU-faktorisering

Från matrisen  $A$ , beräkna  $PA = LU$  där  $P$  är en permutationsmatris (ändrar om raderna).

$$P_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, L = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix}, U = \begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$$

Används för att kunna lösa  $Ax=b$ .

$$Ax = b \Leftrightarrow PAx = Pb \Leftrightarrow LUx = Pb$$

Sätt  $Ux = y$ , lös  $Ly = Pb$  och sedan  $Ux = y$

## 7 Feluppskattning $Ax = b$

$$\frac{\|\Delta x\|_{\infty}}{\|x\|_{\infty}} \leq \|A\|_{\infty} \|A^{-1}\|_{\infty} \frac{\|\Delta b\|_{\infty}}{\|b\|_{\infty}}$$

Matrisnormen definieras som:

$$\|A\|_{\infty} = \max_{1 \leq i \leq n} (\sum_{j=1}^n |a_{ij}|)$$

Alltså, för varje rad, ta absolutvärdet av varje element och addera dem. Maximala värdet för alla rader är matrisnormen.

Matrisnormen är ett mått på hur störningar i högerledet ( $b$ ) förstoras och påverkar  $x$ .

## 8 Minsta kvadratmetoden

Minsta kvadratmetoden används för att hitta en modell för data som inte har exakta lösningar.

För att lösa ekvationen  $Ax = b$  används  $A^T$  för att skapa normalekvationen:

$$A^T Ax = A^T b$$

Det går att visualisera det som att det finns ett plan som spänns upp av kolonnerna i  $A$  och en vektor  $b$ , som ej ligger i planet. Vi vill då hitta en vektor,  $x$ , som placeras ortogonalt mot  $b$  (vilket då är närmast). För att "få ner"



$\mathbf{b}$  i planet multipliceras  $\mathbf{b}$  med matrisen  $A^T$  och då även vänsterledet. Detta resulterar i normalekvationen, som alltid har en lösning.

## 9 Deriveringsregler

$$\frac{d}{dx} \frac{u}{v} = \frac{vu' - uv'}{v^2}$$

## 10 Taylor & Maclaurin

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$$

$$f(x+h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} h^n = f(x) + f'(x)h + \frac{f^{(2)}(x)}{2!} h^2 + \dots$$