LaTeX- Grundlagen

von

Dominik Wille

22 Oktober 2013

Freie Universität Berlin
Zentraleinrichtung für Datenverarbeitung
Betriebsysteme und Programmieren

Dozent:
Dr. Herbert Voß

1 ...

1.1 ...

1.2 ...

1.3 Rundungsfehler

Rechenoperationen mit reelen zahlen im Computer $\rightarrow$ Rundungsfehler.

1.3.1 Gleitkommaaritmetik

Im Vergleich zum Festpunktformat: geringerer Speicherplatzbedarf.

n-stellige Gleitkommazahl, Basis B:

$$x = \pm (0, z, z_2, ..., z_n)_B \cdot B^E = \pm \sum_{i=1}^{n} Z_i \cdot Z_i! = 0 \tag{1}$$

(Normalisierte Gleitkommadarstellung) Exponent: $Ee\mathbb{Z} : m <= E <= M$
Bsp: $+1234,567 = +(0,1234567)_{10} \cdot 10^4$
($B = 10, n = 7$) Die Werte n, B, m, M maschienenabhängig (Hardware und Compiler) Übliche Basen:

- $B = 2$ (Dualzahlen, im Computer)
- $B = 8$ (Oktalzahlen)
- $B = 10$ (Dezimal)
- $B = 16$ (Hexdezimal)

Bsp: binäre Darstellung:

$$(5,0625)_{10} = 0,50625 \cdot 10^1 \tag{2}$$

$$= 1 \cdot 2^2 + 1 \cdot 2^0 + 0 \cdot 2^{-1} + 0 \cdot 2^{-2} + 0 \cdot 2^{-3} + 1 \cdot 2^{-4} \tag{3}$$

$$= (101,0001)_2 = (0,1010001)_2 \cdot 2^3 \tag{4}$$

manche Zahlen lassen sich nur schwer als Dualzahlen darstellen:

- $(3)_{10} = (11)_2$ geht
- $(0,3)_{10} = 0 \cdot 2^{-1} + 1 \cdot 2^{-2} + ... = (0,010011001...)_2$ geht nicht

Genauigkeit der Darstellung

23 Stellen $11111111111111111111111 = 2^{23} - 1 = 8.388.608$
$\Rightarrow$ 6 Ziffren können unterschieden werden.

52 Stellen $2^{52} = 4.503.599.627.370.496$
$\Rightarrow$ 15 Stellen können unterschienden werden. Die größte darstellbare Zahl entspricht der größten Maschienenzahl.

$$x_{max} = (0, [B-1][B-1]...[B-1])_B \cdot B^M = (1 - B^{-n}) \cdot B^M \tag{5}$$

kleinste darstellbare Zahl

$$x_{min} = (0, 1000000)_B \cdot B^m = (1 - B^{-n}) \cdot B^{m-1} \tag{6}$$

$\Rightarrow$ *Die menge der Maschienenzahlen ist endlich*

Bsp:
$x_{max} + x_{max} = \infty$
$x_{min} \cdot B^{-1} = 0$

1.3.2 Rundungsfehler

Beim runden einer Zahl x wird eine Näherung $rd(x)$ unter den Maschienenzahlen geliefert, so dass der absolute Fehler $|x - rd(x)|$ minimal ist, der unvermeidbare Fahler ist der Rundungsfehler. Eine n-stellige Dezimalzahl im Gleitkommaformat

$$x = \pm(0, z_1, ..., z_n)_{10} = rd(x) \tag{7}$$

hat einen maximalen absoluten Fehler :

$$|x - rd(x)| <= 0,000..005 \cdot 10^E \tag{8}$$
$$= 0,5 \cdot E^{E-n} \tag{9}$$

, für allgemeine Basis B:

$$|x - rd(x)| <= \frac{B \cdot 1}{2 \cdot B} B^{E-n} = \frac{1}{2} B^{E-n} \tag{10}$$

Rundungsfehler werden durch die rechnung getragen!

n-stellige Gleitkommaaritmetik:
jede einzelne Rechenoperation ($+, -, \times, \div$)wird auf $n + 1$ Stellen genau berechnet und dann auf n stellen gerundet. Jedes Zwischenergebnis, nicht Endergebnis!

Bsp:
rechne $2590 + 4 + 4$ in 3 stelliger dez G.P.A.

links 1. $2590 + 4 \rightarrow 2590$

2. $2590 + 4 \rightarrow 2590$

rechts 1. $4 + 4 \rightarrow 10$

2. $2590 + 10 \rightarrow 2600$

$\Rightarrow$ Rechenwege unterscheiden sich!

Regel: *beim Addieren Summanden in der Reihenfolge aufsteigender Beträge addieren.*

Maß für der Rechenzeit eines Computers: “flops” floating point operations per second
(typisch Multiplikation oder Division)
(top500.org) 1 Tiake-2 3 Mio Cores, 54.000 T Flops, 17 MW

relative Fehler wichtiger aks absoluter Fehler:
Näherung $\tilde{x}$ zu exaktem wert x, rel. fehler $E = \left|\frac{x-x}{x}\right| \approx \frac{x-x}{\tilde{x}}$ für duale rechniungen am Computer B=2 $\rightarrow E_{max} = 2^{-n}$
E_{max} wird auch maschienenzahlgenauigkeit genannt, und gibt die kleinste potentielle Zahle an, für die gilt $|E_{max}|$; E_{max} kann aus Rechenergebnissen errechnet werden (ÜB1)

Bsp: mit 4 mantissenziffern und Exponentenziffern

Addieren/Subtrahieren von zahlen mit stark unterschiedlichem Exponenten: kleine Zahl kann durch Rundungsfehler verloren gehen.

$$1234 + 0,5 = 0,1234 \cdot 10^4 + 0,5 \cdot 10^0 \tag{11}$$
$$= 1234,5 \rightarrow 1235 Fehler \tag{12}$$

Multiplikation/Division (underflor/ oder flor möglich!)

$$0,2 \cdot 10^{-5} \times 0,3 \cdot 10^{-6} = 0,6 \cdot 10^{-12} \rightarrow 0 \tag{13}$$
$$0,6 \cdot 10^{5} \div 0,3 \cdot 10^{-6} = 0,2 \cdot 10^{12} \rightarrow \infty \tag{14}$$

Fehler des Assoziativgesetzes a)

$$x + (y + z) = (x + y) + z \tag{15}$$
$$0,1111 \cdot 10^{-3} + (-0,1234 + 0,1243) = 0,1111 \cdot 10^{-3} + 0,0009 \tag{16}$$
$$= 0,10111 \cdot 10^{-2} \rightarrow 0,1011 \cdot 10^{-2} \tag{17}$$

b)

$$(0,1111 \cdot 10^{-3} - 0,1234) + 0,1243 = 0,1233 + 0,1243 \tag{18}$$
$$= 0,0010 = 0,100 \cdot 10^{-2} \tag{19}$$

a) Fehler: $0,00001 \cdot 10^{-2} \rightarrow$ relativer fehler $\epsilon = 0,0001 = 0,01\%$

b) Fehler: $0,00111 \cdot 10^{-2} \rightarrow$ relativer Fehler $\epsilon = 0,01 = 1\%$

$\epsilon_{max} = \frac{1}{2} B^{1-4} = 0,0005$; im Fall b) ist ϵ also deutlich größer als ϵ_{max} !

1.3.3 Fehlerfortpflanzung bei Rechenoperationen

Fehler werden beim rechnen weitergetragen, selten werden Fehler dabei kleiner (meistens größer!). Durch Umstellen von Formeln können Fehler minimiert werden, trotzdem müssen Fehler abgeschätzt wreden.

Additionsfehler gegeben fehlerhaste Größen $\tilde{x}$ und $\tilde{y}$ und exakten Werte x, y Fehler der Summe: $\tilde{x} + \tilde{y} - (x + y = (\tilde{x} - x) + (\tilde{y} - y)$ Im ungünstigsten Fall addieren sich die Fehler:

$\rightarrow$ *bei Addition und Subtraktion addieren sich die Absolutbeträge der Fehler!*

Multiplikation $\tilde{x}\tilde{y} - xy = \tilde{x}(\tilde{y} - y) + \tilde{y}(\tilde{x} - x)(\tilde{y} - y)$

also hat das Produkt von $\tilde{y}$ mit einer maschienenzahl ohne Fehler ($\tilde{x} - x$ den $\tilde{x}$-fachen Fehler (und umgekehrt); Produkt der Fehler - typischer Weise vernachlässigbar.

$\rightarrow$ der absolute Fehler eines Prodults ist gegeben durch das Prodult des Faktors mit dem Fehler des anderen Faktors. (=2 Treme, oft ist einer der Terme dominant.)

Reative Fehler eines Produktes:

$$\frac{\tilde{x}\tilde{y} - xy}{\tilde{x}\tilde{y}} = \frac{\tilde{y} - y}{\tilde{y}} + \frac{\tilde{x} - x}{\tilde{x}} - \frac{(\tilde{x} - x)(\tilde{y} - y)}{\tilde{x}\tilde{y}} \tag{20}$$

$\rightarrow$ Beim Multiplizieren addieren sich die relativen Fehler.Division analog...

1.3.4 Fehlerfortpflanzung -> Funktionen

Funktionen auswertung $f(x)$ an Stelle $\tilde{x}$ anstatt x $\rightarrow$ großen/kleinen Fehler von f. bei zweiten Funktionsauswertungen wird der Fehler typischerweise größer...
Mittelwertsatz: $\int_x^{\tilde{x}} g\,(x')\,dx' = g(x_0)(\tilde{x} - x)$
Mittelwert der Funktion: $\frac{\int_x^{\tilde{x}} g(x')dx'}{\tilde{x}-x}$ = Funktionswert $g(x_0)$ an einer unbekannten Stelle x_0 im Intervall $(x, \tilde{x})$, (für stetige Funktionen $g(x)$....)

wähle $g(x) = f'(x) \rightarrow |f(\tilde{x}) - f(x)| = |\tilde{x} - x|\,|f'(x_0)|$
$\rightarrow$ *absoluter Fehler vergrößert sich für* $|f(x_0)| > 1$ *bzw verkleinert sich für* $|f(x_0)| < 1$

also: Ableitung bestimmt den Verstärkungsfaktor des Fehlers!

Abschätzung des absoluten Fehlers: $|f(x) - f(\tilde{x})| \leq M\,|x - \tilde{x}|$ mit $M = |f'(x_0)|$
Schätzung der Fehler: $|f(x) - f(\tilde{x})| \approx |f'(\tilde{x})|\,|x - \tilde{x}|$

Bsp.:Fortpflanzung des absoluten Fehlers für $f(x) = \sin x f'(x) = \cos x$ und damit $M = max_{x_0} f'(x_0) = 1$ d.h. für die meisten Argumente veringert sich der absolute Fehler!

Bsp.: $f(x) = \sqrt{x}$; $f'(x) = \frac{0{,}5}{\sqrt{x}}$divergiert also für $x \rightarrow \infty$

relativer Fehler bei Funktionsauswertung:

Konditionszahl: $\frac{|f'(\tilde{x})||\tilde{x}|}{|f(\tilde{x})|}$

Verhältnisfaktor für relative fehler; „qualitativ: Probleme zur Koordinatenzahl $>> 1$ ”schlech

2 Nullstellenprobleme

geg: stetige Funktion $f : \mathbb{R} \to \mathbb{R}$
ges: Nullstelle(n), also $x_0 e \mathbb{R}$ mit $f(x_0) = 0$
grundsätzlich:

- gibt es überhaupt keine Nullstelle ?
- gibt es mehrere?

Zweischensatz: $f : [a, b] \to \mathbb{R}, stetig$, für $ce\mathbb{R}$ mit $f(a)c \leq f(b))$ gibt es ein $x_0 e[a, b]$ so dass $f(x_0) = c$

für $c = 0$ ist der Satz hilfreich bei der Nullstellensuche:

suche Funktionsargumente mit unterschiedlichem Vorzeichen $f(a)f(b) < 0$ dann gibt es zwischen a und b mindestens eine Nullstelle!

2.1 Bisektionsverfahren

$f(a)f(b) < 0 =$ Nullstelle in (a, b), berechne Vorzeichen von $f\left(\frac{a+b}{2}\right)$
$\to f(x) = 0 in \left(0, \frac{a+b}{2}\right)$ oder $\left(\frac{a+b}{2}, b\right)$
weiter halbieren...

Bsp.: $f(x) = x^3 - x + 0,3 = 0$

a) wie viele Nullstellen? $x^3 - x$ hat 3 Nullstellen bei $x = \pm 1, 0$
Wir setzten also die Umgebeung von $x = \pm 1, 0$

x	-2	-1	0,5	1
f	-5,7	0,3	-0,075	0,3