Computerorientierte Mathematik I

4. Vorlesung

Carsten Gräser

Freie Universität Berlin

08.11.2019

Rückblick: Vorlesung vom 01.11.2019

Freie Universität Berlin

Darstellung rationaler und reeller Zahlen

Rationale Zahlen

- Rationale Zahlen als Brüche ganzer Zahlen
- Endliche und periodische q-adische Brüche
- Satz: Jede rationale Zahl ist ein periodischer q-adischer Bruch
- Eindeutigkeit: $0,\overline{9}$ statt $1 = 1,\overline{0}$
- Praktische Realisierung: Dynamische Ziffernzahl, Aufwand pro Addition problemabhängig (Hauptnenner, Kürzen)

Reelle Zahlen

- Reelle Zahlen als unendliche q-adische Brüche
- Satz: $\mathbb{R}$ ist nicht abzählbar.
- Folgerung: Es gibt keine Zifferndarstellung von $\mathbb{R}$
- Konsequenz: Numerisches Rechnen mit reellen Zahlen ist nicht möglich!

Fest- und Gleitkommazahlen

- Absoluter und relativer Fehler, Beispiele
- Definition von Festkommazahlen und Gleitkommazahlen, Beispiele

Festkommazahlen (q-adische Brüche)

Festkommazahlen

$$z_{n-1} \dots z_0, z_{-1} \dots z_{-m} = \sum_{i=-m}^{n-1} z_i q^i, \qquad z_i \in \{0, \dots, q-1\}$$

- $n, m \in \mathbb{N}$ fest gewählt
- $l = m + n$ Stellen verfügbar
- Fester, endlicher Speicherplatz pro Zahl

Beispiel: $q = 10$, $l = 4$, $n = 3$, $m = 1$

- $x = 0,123$, Runden: $\tilde{x} = 0,1$, relativer Fehler: $|x - \tilde{x}|/|x| \approx 0,2$
- $x = 123$, exakt darstellbar: $\tilde{x} = 123$, relativer Fehler: $|x - \tilde{x}|/|x| = 0$.

Folgerung

Um die Stellen optimal auszunutzen, sollte man m und n variable halten!

Definition (Gleitkommazahlen)

Jede in der Form

$$\tilde{x} = (-1)^s a \cdot q^e \tag{1}$$

mit Vorzeichenbit $s \in \{0, 1\}$, Exponent $e \in \mathbb{Z}$ und Mantisse $a = 0$ oder

$$a = 0, a_1 \dots a_l = \sum_{i=1}^{l} a_i q^{-i}, \qquad a_i \in \{0, \dots, q-1\}, a_1 \neq 0$$

darstellbare Zahl $\tilde{x}$ heißt Gleitkommazahl mit Mantissenlänge $l \in \mathbb{N}, l \geq 1$.

Die Menge all dieser Zahlen heißt $\mathbb{G}(q, l)$.

Die Darstellung (1) heißt normalisierte Gleitkommadarstellung.

Normalisierte Darstellung

$$x = a^* q^e, \qquad e \in \mathbb{Z}, \qquad q^{-1} \leq a^* < 1$$

Unendlicher q-adischer Bruch

$$a^* = 0, a_1 a_2 \dots a_l a_{l+1} \cdots = \sum_{i=1}^{\infty} a_i q^{-i}, \qquad a_i \in \{0, \dots, q-1\}$$

Runden

$$\tilde{x} = \mathrm{rd}(x) := aq^e$$

mit gerundeter Mantisse

$$a = \sum_{i=1}^{l} a_i q^{-i} + \begin{cases} 0 & \text{falls } a_{l+1} < \frac{1}{2}q \\ q^{-l} & \text{falls } a_{l+1} \geq \frac{1}{2}q \end{cases}$$

Satz

Zu jedem $n \in \mathbb{N}$ gibt es ein $x \in \mathbb{R}$, so dass

$$|x - \mathrm{rd}(x)| \geq q^n$$

Beweis.

Wähle $x = 0, z_1 \dots z_l z_{l+1} \cdot q^{l+1+n}$ mit $z_1, z_{l+1} \neq 0$ und $n \in \mathbb{N}$. □

Der absolute Fehler kann beliebig groß werden.

Satz

Es sei q eine gerade Zahl. Dann gilt

$$\frac{|x - \mathrm{rd}(x)|}{|x|} \leq \frac{1}{2} q^{-(l-1)} =: \mathrm{eps}(q, l), \qquad \forall x \in \mathbb{R}, x \neq 0.$$

Die Zahl $\mathrm{eps}(q, l)$ heißt Maschinengenauigkeit.

Beweisskizze.

O.B.d.A. sei $x > 0$ und $a_{l+1} \geq \frac{1}{2} q$.

...

□

Maschinengenauigkeit

Freie Universität Berlin

Der relative Rundungsfehler ist durch $\mathrm{eps}(q, l)$ beschränkt.

Mantissenlänge $l \Leftrightarrow l$ gültige Stellen $\Leftrightarrow \mathrm{eps}(q, l) = \frac{1}{2} q^{-(l-1)}$

Endlicher Exponentenbereich

$$e \in \{e_{\min}, e_{min} + 1, \ldots, e_{\max} - 1, e_{\max}\} \tag{2}$$

Endlicher Zahlenbereich

$$x_{\min} := q^{e_{\min}-1} \leq |x| \leq (1 - q^{-l})q^{e_{\max}} =: x_{\max} \tag{3}$$

- $x \in [-x_{\max}, -x_{\min}] \cup \{0\} \cup [x_{\min}, x_{\max}]$
- $|x| < x_{\min}$: underflow oder $x = 0$
- $|x| > x_{\max}$: overflow oder $x = NaN$ oder $x = inf$

IEEE 754 Standard

Freie Universität Berlin

	float	double
Länge in Bits	32	64
Vorzeichenbit s	1	1
Exponent e Bits	8	11
Mantisse a Bits	23	52
$e_{\min}$	-126	-1022
$e_{\max}$	-128	1024
$x_{\min}$	$1,2 \cdot 10^{-38}$	$2.2 \cdot 10^{-308}$
$x_{\max}$	$3,4 \cdot 10^{+38}$	$1.8 \cdot 10^{308}$

Menge aller Approximationen $\tilde{x}$ auf l gültige Stellen im q-System

$$\mathrm{rd}(x) \in \left\{ \tilde{x} \;\middle|\; \tilde{x} = x(1+\varepsilon), |\varepsilon| \leq \mathrm{eps}(x,l) \right\}, \qquad \forall x \in \mathbb{R}$$

Menge aller $x \in \mathbb{R}$, die auf $\tilde{x} = \mathrm{rd}(x) \in \mathbb{G}(q,l)$ gerundet werden

$$R(\tilde{x}) = \left\{ x \in \mathbb{R} \;\middle|\; \tilde{x} = \mathrm{rd}(x) \right\}$$

Satz

Es sei

$$\tilde{x} = aq^e \in \mathbb{G}(q,l), \qquad q^{-1} < a_0, a_1, \ldots, a_l \leq 1.$$

Dann gilt $R(\tilde{x}) = [\alpha(\tilde{x}), \beta(\tilde{x}))$ *mit*

$$\alpha(\tilde{x}) = \tilde{x} - q^{e-1}\,\mathrm{eps}(q,l), \qquad \beta(\tilde{x}) = \tilde{x} - q^{e-1-a_0}\,\mathrm{eps}(q,l)$$

Gleichheitsabfragen von Gleitkommazahlen

Folgerung

Die Abfrage $\tilde{x} == \tilde{y}$ mit $\tilde{x}, \tilde{y} \in \mathbb{G}(q, l)$ ist sinnlos!

$$\tilde{x} = \tilde{y} \qquad \not\Rightarrow \qquad x = y, \qquad \tilde{x} = \mathrm{rd}(x), \tilde{y} = \mathrm{rd}(y)$$

umgekehrt gilt

$$x = y \qquad \Rightarrow \qquad \mathrm{rd}(x) = \mathrm{rd}(y),$$

aber

$$x = a + b, y = x, \qquad \tilde{x} = \mathrm{rd}(a) + \mathrm{rd}(b), \tilde{y} = \mathrm{rd}(x) \qquad \not\Rightarrow \qquad \tilde{x} = \tilde{y}$$

Gleichheitsabfragen von Gleitkommazahlen sind verboten!

Grundrechenarten

- Führen aus $\mathbb{G} = \mathbb{G}(q, l)$ heraus!
- Addition:

$$\tilde{x}, \tilde{y} \in \mathbb{G} \qquad \not\Rightarrow \qquad \tilde{x} + \tilde{y} \in \mathbb{G}$$

- Analog: $-$,$\cdot$,/

Gleitkommaarithmetik

$$\tilde{x} \tilde{+} \tilde{y} = \mathrm{rd}(\tilde{x} + \tilde{y}), \qquad \tilde{x} \tilde{-} \tilde{y} = \mathrm{rd}(\tilde{x} - \tilde{y}), \qquad \tilde{x} \tilde{\cdot} \tilde{y} = \mathrm{rd}(\tilde{x} \cdot \tilde{y}), \qquad \tilde{x} \tilde{/} \tilde{y} = \mathrm{rd}(\tilde{x} / \tilde{y}),$$

Die Gleitkommazahlen mit Gleitkommaarithmetik sind kein Körper.

$\tilde{+}, \tilde{*}$ sind nicht assoziativ, kein Distributivgesetz, i.a., kein Inverses bzgl. $\tilde{*}$.

Äquivalente Umformungen in $\mathbb{R}$ sind in Gleitkommaarithmetik nicht äquivalent!

Beispiele

- Keine binomische Formel
- Kein Assoziativgesetz