Vorlesung: Angewandte Mathematik für LAK
WS 2014/2015

Priv.-Doz. Dr(USA) Maria Charina

1. Fehleranalyse, Kondition, Rundungsoperator, Stabilität

Problem: Auswertung von $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ an der Stelle $x \in \Omega$.

Ziele:
Kondition: Vergleiche den Ausgabefehler $|f(x) - f(\widetilde{x})|$ mit dem Eingabefehler $\|x - \widetilde{x}\|$, wobei die Werte $f(x)$ und $f(\widetilde{x})$ ohne Rundungsfehler (exakt) berechnet werden;

Stabilität: Vergleiche den Ausgabefehler $|f(x) - f(\widetilde{x})|$ bei der exakten Berechnung von f mit dem Ausgabefehler $|f(x) - \widetilde{f}(x)|$ bei der numerischen Berechnung (mit Rundungsfehler) von $\widetilde{f}(x)$.

Sei V ein $\mathbb{K}-$Vektorraum über dem Körper $\mathbb{K}$.

1.1 Definition (Vektornorm): Eine Abbildung $\|\cdot\| : V \to \mathbb{R}_+$ heißt eine Vektornorm auf V, falls

(*i*) Für alle $v \in V$ gilt $\|v\| \geq 0$ und $(\|v\| = 0 \Leftrightarrow v = 0)$;

(*ii*) Für alle $\alpha \in \mathbb{K}$ und $v \in V$ gilt $\|\alpha v\| = |\alpha|\|v\|$;

(*iii*) Für alle $v, w \in V$ gilt $\|v + w\| \leq \|v\| + \|w\|$.

1.2 Satz (Normäquivalenz): Auf dem endlichdimensionalen Vektorraum $\mathbb{K}^n$ sind alle Vektornormen äquivalent, d.h.: Zu je zwei Normen $\|\cdot\|$ und $\|\cdot\|'$ gibt es positive Konstanten m, M mit denen gilt

$$m\|x\|' \leq \|x\| \leq M\|x\|', \quad x \in \mathbb{K}^n.$$

1.3 Relative Konditionszahlen eines mathematischen Problems

Seien $\Omega \subset \mathbb{R}^n$ offen, konvex, und $f : \Omega \to \mathbb{R}$ zweimal stetig differenzierbar.

Die Taylorentwicklung von f bei $\widetilde{x} = x + \varepsilon_x \in \Omega$ in der 1.Näherung ergibt

$$\left| \frac{f(x) - f(\widetilde{x})}{f(x)} \right| \quad \lesssim \quad \max_{j=1,\ldots,n} |k_j(x)| \cdot \left\| \frac{x - \widetilde{x}}{x} \right\|_1$$

mit <u>den relativen Konditionszahlen</u>

(1) $$k_j(x) := \frac{\partial f}{\partial x_j}(x) \cdot \frac{x_j}{f(x)}, \quad j = 1, \ldots, n, \quad x \in \Omega.$$

1.4 Definition: Man nennt das Problem der Auswertung von $f(x)$, $x \in \Omega$,

- (i) schlecht konditioniert, wenn ein $|k_j| >> 1$ ist;
- (ii) andernfalls gut konditioniert.

1.5 Kondition arithmetischer Grundoperationen

- Die Addition $y = f(x_1, x_2) = x_1 + x_2, \;\; x_1, x_2 \in \mathbb{R} \setminus \{0\}$, mit

$$\begin{aligned} k_1 &= \frac{\partial f}{\partial x_1}(x) \cdot \frac{x_1}{f(x)} = 1 \cdot \frac{x_1}{x_1 + x_2} \\ k_2 &= \frac{\partial f}{\partial x_2}(x) \cdot \frac{x_2}{f(x)} = 1 \cdot \frac{x_2}{x_1 + x_2} \end{aligned}$$

  ist schlecht konditioniert für $x_1 \approx -x_2$.

- Die Multiplikation $y = f(x_1, x_2) = x_1 \cdot x_2$ mit

$$\begin{aligned} k_1 &= \frac{\partial f}{\partial x_1}(x) \cdot \frac{x_1}{f(x)} = x_2 \cdot \frac{x_1}{x_1 \cdot x_2} = 1 \\ k_2 &= \frac{\partial f}{\partial x_2}(x) \cdot \frac{x_2}{f(x)} = x_1 \cdot \frac{x_2}{x_1 \cdot x_2} = 1 \end{aligned}$$

  ist gut konditioniert.

1.6 Rundungsfehler und Gleitkommaarithmetik

Definition: Die Menge $fl = fl(B, m, E_s)$ von Gleitkommazahlen zur Basis $B \in \mathbb{N}$, $B \geq 2$, mit Mantissenlänge $m \in \mathbb{N}$ und Exponentenbereich $E_s \subset \mathbb{N}$ besteht aus allen reellen Zahlen der Form

$$\pm \left(\sum_{j=1}^{m} d_j \cdot B^{-j} \right) \times B^{\pm e}, \qquad e = \sum_{k \in E_s} e_k \cdot B^k,$$

mit $d_j, e_k \in \{0, \ldots, B-1\}$, $d_1 \neq 0$.

Bemerkung: Moderne Rechner verwenden $B = 2, 10, 16$.

Definition (Rundungsoperator): Es seien $B \in \mathbb{N}$, $B \geq 2$, gerade, $m \in \mathbb{N}$, $E_s \subset \mathbb{N}$, und

$$x = \pm \left(\sum_{j=1}^{\infty} d_j \cdot B^{-j} \right) \times B^{\pm e} \in \mathbb{R}.$$

Definiere den Rundungsoperator $\mathrm{rd} : \mathbb{R} \to fl(B, m, E_s)$ durch

$$\mathrm{rd}(x) = \begin{cases} \pm \left(\sum_{j=1}^{m} d_j \cdot B^{-j} \right) \times B^{\pm e} & \text{, falls} \quad d_{m+1} < \frac{B}{2}, \\ \pm \left(\sum_{j=1}^{m} d_j \cdot B^{-j} + B^{-m} \right) \times B^{\pm e} & \text{, sonst.} \end{cases}$$

Satz: Es seien $B \in \mathbb{N}$, $B \geq 2$, gerade, $m \in \mathbb{N}$, $E_s \subset \mathbb{N}$. Dann gilt

$$\left| \frac{x - \mathrm{rd}(x)}{x} \right| \leq \frac{1}{2} B^{-m+1}, \quad x \neq 0.$$

Bemerkung: Die Zahl eps $:= \frac{1}{2} B^{-m+1}$ wird relative Rechengenaugkeit der $m-$stelligen Gleitkommaarithmetik genannt.

Definition: Die Gleitkommaoperationen

$$\{\oplus, \ominus, \otimes, \oslash\} : fl \times fl \to fl$$

sind definiert durch

$$\begin{aligned} x \oplus y &:= \mathrm{rd}(x + y), \\ x \ominus y &:= \mathrm{rd}(x - y), \\ x \otimes y &:= \mathrm{rd}(x \cdot y), \\ x \oslash y &:= \mathrm{rd}(x / y). \end{aligned}$$

1.7 Numerische Stabilität von Algorithmen

Definition: Ein Algoritmus heißt numerisch stabil, wenn der im Verlaufe der Ausführung akkumulierte Rundungsfehler den durch die Kondition des Problems

$$\textit{Berechne} \quad y = f(x), \quad x \in \Omega,$$

bedingten unvermeidbaren Problemfehler nicht übersteigt.

Beispiel: Berechne $f(x_1, x_2) = x_1^2 - x_2^2$, $x_1 = 0.334$, $x_2 = 0.333$.

Unvermeidbarer Problemfehler:

$$\left|\frac{f(x_1,x_2) - f(\tilde{x}_1, \tilde{x}_2)}{f(x_1,x_2)}\right| \lesssim \max\left\{\frac{2}{1-\left(\frac{x_2}{x_1}\right)^2}, \frac{2}{1-\left(\frac{x_1}{x_2}\right)^2}\right\} \cdot \mathrm{eps} \approx 7 \cdot 10^3 \cdot \mathrm{eps}.$$

Akkumulierter Rundungsfehler des 1.Algoritmus

$$\begin{aligned} x_1^2 - x_2^2 &\approx \widetilde{f}(x) = (x_1 \otimes x_1) \ominus (x_2 \otimes x_2) = \mathrm{rd}(\mathrm{rd}(x_1^2) - \mathrm{rd}(x_2^2)) \\ &= 0.1 \cdot 10^{-2} \end{aligned}$$

ist gegeben durch

$$\left| \frac{f(x) - \widetilde{f}(x)}{f(x)} \right| \approx 0.499 \approx 100\mathrm{eps} \quad \text{für} \quad \mathrm{eps} = 0.5.$$

Akkumulierter Rundungsfehler des 2.Algoritmus

$$\begin{aligned} x_1^2 - x_2^2 &\approx \widetilde{f}(x) = (x_1 \oplus x_2) \otimes (x_1 \ominus x_2) = \mathrm{rd}(\mathrm{rd}(x_1 + x_2) \cdot \mathrm{rd}(x_1 - x_2)) \\ &= 0.667 \times 10^{-3} \end{aligned}$$

ist gleich 0.

Nach Definition 1.7 sind diese beiden Algoritmen numerisch stabil.