Kondition linearer Gleichungssysteme

↳ beschreibt, wie stark das Ergebnis / die Lösung einer Berechnung / eines Problems von den Eingangsgrößen abhängt.

▶ 1

exaktes Problem

$$A x = b \qquad \text{mit Input: Systemmatrix } A \in \mathbb{R}^{n \times n} \text{ & rechte Seite } b \in \mathbb{R}^n$$

$$\text{Output: Lösungsvektor } x \in \mathbb{R}^n$$

gestörtes Problem

$$\tilde{A}\,\tilde{x} = \tilde{b} \qquad \text{mit Input: } \bullet \text{ (gestörte) Systemmatrix } \tilde{A} = A + \Delta A \quad (\text{Störung: } \Delta A \in \mathbb{R}^{n \times n})$$

$$\bullet \text{ (gestörte) rechte Seite } \tilde{b} = b + \Delta b \quad (\text{Störung: } \Delta b \in \mathbb{R}^n)$$

$$\text{Output: (gestörter) Lösungsvektor } \tilde{x} = x + \Delta x \quad (\text{Störung im Ergebnis } \Delta x \in \mathbb{R}^n)$$

? Kann man Δx durch ΔA und Δb abschätzen?

Beispiel: $\begin{pmatrix} 1 & 1 \\ 0 & 0{,}01 \end{pmatrix} \cdot x = \begin{pmatrix} 1 \\ -0{,}02 \end{pmatrix}$ eindeutig lösbar

▶ 2

≀ störung

- $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \cdot x = \begin{pmatrix} 1 \\ -0{,}02 \end{pmatrix}$ nicht lösbar — Situation: nur Systemmatrix gestört
- $\begin{pmatrix} 1 & 1 \\ 0 & 0{,}01 \end{pmatrix} \cdot x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ eindeutig lösbar — Situation: nur rechte Seite gestört
- $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ unendlich viele Lösungen — Situation: beides gestört

Dr. rer. nat. M. Striebel

zum Verständnis notwendig : Vektor- und Matrix-Normen

1.) Vektor-Norm

▶ 3

Definition (Vektornorm)

Eine Abbildung $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_0^+ , x \mapsto \|x\|$ heißt Vektornorm (in $\mathbb{R}^n$), wenn gilt:

i) $\|x\| \geq 0 \quad \forall x \in \mathbb{R}^n$ und $\|x\| = 0 \Leftrightarrow x = \vec{0}$ (positiv definit)

ii) $\|\alpha \cdot x\| = |\alpha| \cdot \|x\| \quad \forall \alpha \in \mathbb{R}, \forall x \in \mathbb{R}^n$

iii) $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathbb{R}^n$ (Dreiecksungleichung)

gebräuchliche Normen:

- Standardnorm / euklidische Norm / 2-Norm im $\mathbb{R}^n$

$$\|x\| = \|x\|_2 = |x| = \sqrt{\sum_{k=1}^{n} x_k^2} = \sqrt{\langle x, x \rangle}$$

- Summennorm / 1- Norm

$$\|x\|_1 = \sum_{k=1}^{n} |x_k|$$

- Maximumnorm / unendlich Norm

$$\|x\|_\infty = \max_{k=1,\dots,n} |x_k|$$

p-Norm:

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p}$$

2.) Matrix- Norm

▶ 4

Definition (Matrixnorm)

Eine Abbildung $\|\cdot\| : \mathbb{R}^{n \times m} \to \mathbb{R}_0^+ , A \mapsto \|A\|$ heißt Matrixnorm (in $\mathbb{R}^{n \times m}$), wenn gilt:

i) $\|A\| \geq 0 \quad \forall A \in \mathbb{R}^{n \times m}$ und $\|A\| = 0 \Leftrightarrow A = \mathcal{O}_{n \times m}$ (positiv definit)

ii) $\|\alpha \cdot A\| = |\alpha| \cdot \|A\| \quad \forall \alpha \in \mathbb{R}, \forall A \in \mathbb{R}^{n \times m}$

iii) $\|A + B\| \leq \|A\| + \|B\| \quad \forall A, B \in \mathbb{R}^{n \times m}$ (Dreiecksungleichung)

Dr. rer. nat. M. Striebel

oft geforderte Zusatzforderungen:

i) Submultiplikativität: $\|A \cdot B\| \leq \|A\| \cdot \|B\| \quad \forall A \in \mathbb{R}^{n \times \ell} \quad \forall B \in \mathbb{R}^{\ell \times m}$

ii) Verträglichkeit mit einer Vektornorm

$\|A \cdot x\| \leq \|A\| \cdot \|x\| \quad \forall A \in \mathbb{R}^{n \times m}, \forall x \in \mathbb{R}^m$

($\|A \cdot x\|$: Vektornorm; $\|A\|$, $\|x\|$: Matrixnorm)

besonders mit einer Vektornorm verbundene Matrixnormen

Definition (induzierte Matrixnorm / natürliche Norm)

Eine Matrixnorm $\|\cdot\|$ heißt von einer Vektornorm $\|\cdot\|$ induziert, wenn gilt

$$\|A\| = \max_{\|x\| \neq \vec{0}} \frac{\|Ax\|}{\|x\|} \quad \forall A \in \mathbb{R}^{n \times m}$$

Bemerkungen:

i) alternativ: $\|A\| = \max_{\|x\| = 1} \|Ax\|$

ii) eine Matrixnorm ist von einer Vektornorm induziert, wenn es zu jeder Matrix A einen Vektor x gibt so dass $\|A\| = \|Ax\|$ (Matrixnorm / Vektornorm)

iii) eine induzierte Matrixnorm ist verträglich mit der sie induzierenden Vektornorm

gebräuchliche natürliche Matrixnormen

- Spaltensummennorm: $\|A\|_1 := \max_{\ell = 1,\dots,m} \sum_{k=1}^{n} |a_{k\ell}|$ (induziert von der Summennorm)
- Zeilensummennorm: $\|A\|_\infty := \max_{k = 1,\dots,n} \sum_{\ell=1}^{m} |a_{k\ell}|$ (induziert von der Maximumnorm)
- Spektralnorm: $\|A\|_2 := \sqrt{\lambda_{max}(A^T A)}$ (induziert von der Standardnorm)

(λ_{max}: größter EW von $A^T A$)

Beispiele:

i) $u = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$:

$\|u\|_2 = \sqrt{1^2 + 2^2 + (-3)^2} = \sqrt{14} \approx 3{,}742$

$\|u\|_1 = |1| + |2| + |-3| = 1 + 2 + 3 = 6$

$\|u\|_\infty = \max\{|1|, |2|, |-3|\} = 3$

Dr. rer. nat. M. Striebel

ii) $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & -2 \\ 7 & -3 & 5 \end{pmatrix}$

$\|A\|_1 = \max\{|1|+|3|+|7|, |2|+|4|+|-3|, |3|+|-2|+|5|\}$

$= \max\{11, 9, 10\} = 11$ (Spaltensummennorm)

$\|A\|_\infty = \max\{|1|+|2|+|3|, |3|+|4|+|-2|, |7|+|-3|+|5|\}$

$= \max\{6, 9, 15\} = 15$ (Zeilensummennorm)

$\|A\|_2 \approx 9{,}302$ (via Matlab)

Zurück zum Vergleich von $Ax = b$ mit $\tilde{A}\tilde{x} = \tilde{b}$

▶ 7

absolute und relative Differenz zwischen x und $\tilde{x}$ hervorgerufen durch $\Delta A := \tilde{A} - A$ und $\Delta b := \tilde{b} - b$

absolut: $\|\tilde{x} - x\|$

relativ: $\frac{\|\tilde{x} - x\|}{\|x\|}$

Grundannahmen:
- $A \in \mathbb{R}^{n \times n}$, regulär
- gewählt sei eine Vektornorm und eine damit verträgliche Matrixnorm

I.) Störung der rechten Seite – $\Delta b \neq \vec{0}$, $\Delta A = O_{n \times n}$

▶ 8

exaktes Problem: $A \cdot x = b$

gestörtes Problem: $A \cdot \tilde{x} = \tilde{b}$

mit $\tilde{b} = b + \Delta b$ und $\tilde{x} = x + \Delta x$ gilt

$A\tilde{x} = \tilde{b} \Rightarrow A(x + \Delta x) = b + \Delta b$

$\Rightarrow Ax + A \cdot \Delta x = b + \Delta b$

$\Rightarrow A \cdot \Delta x = \Delta b$

$\Rightarrow \Delta x = A^{-1} \cdot \Delta b$

Erinnerung: „Konditionszahlen"

mit $\varphi(x) = y$ ist

$|\Delta y| \doteq |\varphi'(x)| \cdot |\Delta x|$

hier: $\varphi_A(b) = x$ mit $\varphi(b) = A^{-1} \cdot b$

somit absolute Abweichung zwischen x & $\tilde{x}$:

$\|x - \tilde{x}\| = \|\Delta x\| = \|A^{-1} \cdot \Delta b\| \leq \|A^{-1}\| \cdot \|\Delta b\|$

($\|A^{-1}\|$: Matrixnorm)

mit $Ax = b$ ist $\|b\| = \|A \cdot x\| \leq \|A\| \cdot \|x\| \Rightarrow \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$

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

($\|A\| \cdot \|A^{-1}\|$: Matrixnormen)

Skalierungsfaktor für relative Abweichung

≙ relative Konditionszahl

Dr. rer. nat. M. Striebel

Definition (Konditionszahl)

Für eine reguläre Matrix $A \in \mathbb{R}^{n \times n}$ heißt

$$\kappa(A) := \|A\| \cdot \|A^{-1}\| \qquad (\kappa = \text{„Kappa“})$$

die Konditionszahl von A (bzgl. der Matrixnorm $\|\cdot\|$)

▶ 9

⚠ $\kappa(A)$ wird i.A. nicht über die Matrixinverse A^{-1} bestimmt.

Bemerkungen zur Bestimmung der Konditionszahl

i) ist A regulär (oder gilt $\operatorname{rang}(A) = n$ für $A \in \mathbb{R}^{m \times n}$ mit $m \geq n$)

▶ 10

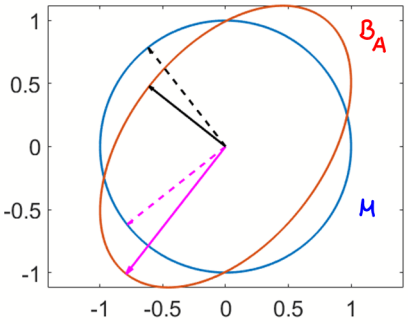
und ist die Matrixnorm von einer Vektornorm induziert dann gilt:

$$\kappa(A) = \frac{\max\limits_{\|x\|=1} \|Ax\|}{\min\limits_{\|x\|=1} \|Ax\|}$$

Interpretation:

$M := \{ x \in \mathbb{R}^n \mid \|x\| = 1 \} \;\hat{=}\; \text{„Einheitskugel“}$

$B_A := \{ Ax \mid x \in M \} \;\hat{=}\; \text{Bild der „Einheitskugel“ unter } A$



$$A = \begin{pmatrix} 1 & 0{,}01 \\ 0{,}5 & 1 \end{pmatrix}$$

$\kappa(A) = 1{,}658$ (bzgl. Spektralnorm)

↳ interpretierbar als maximale Verzerrung der Einheitskugel

ii) ist A symmetrisch (allgemeiner: ist A normal, d.h. $A^T A = A A^T$)

dann gilt hinsichtlich der Spektralnorm:

$$\kappa(A) = \left| \frac{\lambda_{max}(A)}{\lambda_{min}(A)} \right|$$

iii) allgemein gilt für $A \in \mathbb{R}^{m \times n}$ $(m \geq n)$

$$\kappa(A) = \|A\| \cdot \|A^+\|$$

↳ Pseudoinverse

iv) Konditionszahlen bzgl. unterschiedlicher Normen können gegeneinander abgeschätzt werden.

Dr. rer. nat. M. Striebel

Beispiel:

gesucht: Fehlerfortpflanzung für das Problem $A \cdot \vec{x} = \vec{b}$ mit ▶ 11

$$A = \begin{pmatrix} 3 & 1{,}01 \\ 6 & 1{,}99 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

falls jede Komponente der rechten Seite um maximal $0{,}1$ abweicht

(ie: $b_1 \in [-2{,}1; -1{,}9]$, $b_2 \in [0{,}9; 1{,}1]$)

„maximale Abweichung der Komponenten" ⇝ Maximum-Norm

d.h. $\|\vec{b}\|_\infty = \max\{|-2|, |1|\} = 2$

$\|\Delta\vec{b}\|_\infty = 0{,}1$

$\|A\|_\infty = \max\{|3| + |1{,}01|, |6| + |1{,}99|\} = 7{,}99$ (Zeilensummennorm)

mit $A^{-1} = \frac{1}{3 \cdot 1{,}99 - 6 \cdot 1{,}01} \begin{pmatrix} 1{,}99 & -1{,}01 \\ -6 & 3 \end{pmatrix}$

$= -\frac{1}{9/100} \begin{pmatrix} 1{,}99 & -1{,}01 \\ -6 & 3 \end{pmatrix} = \begin{pmatrix} -\frac{199}{9} & \frac{101}{9} \\ \frac{600}{9} & -\frac{300}{9} \end{pmatrix}$

daher $\|A^{-1}\|_\infty = \max\left\{\left|-\frac{199}{9}\right| + \left|\frac{101}{9}\right|, \left|\frac{600}{9}\right| + \left|-\frac{300}{9}\right|\right\} = \frac{900}{9} = 100$

und $\chi(A) = \chi_\infty(A) = \|A\|_\infty \cdot \|A^{-1}\|_\infty = 7.99 \cdot 100 = 799$

somit
- absoluter Fehler: $\|\Delta\vec{x}\|_\infty \leq \|A^{-1}\|_\infty \cdot \|\Delta\vec{b}\|_\infty \leq 100 \cdot 0{,}1 = 10$
- relativer Fehler:

$$\frac{\|\Delta\vec{x}\|_\infty}{\|\vec{x}\|_\infty} \leq \chi_\infty(A) \cdot \frac{\|\Delta\vec{b}\|}{\|\vec{b}\|} \leq 799 \cdot \frac{0{,}1}{2} = 39{,}95$$

Bedeutung:
- Lösung $\tilde{\vec{x}}$ von $A\tilde{\vec{x}} = \vec{b}$ weicht in jeder Komponente um maximal 10 ab von der Lösung $\vec{x}$ von $A\vec{x} = \vec{b}$ (mit $\|\vec{b} - \tilde{\vec{b}}\|_\infty \leq 0{,}1$)
- relativer Fehler beträgt maximal 39,95
- Fehlerverstärkung absoluter Fehler: maximal 100
  relativer Fehler: maximal 799

Dr. rer. nat. M. Striebel

konkrete Werte:

exakte Lösung: $\vec{x} = \begin{pmatrix} \frac{499}{9} \\ -\frac{500}{3} \end{pmatrix} \approx \begin{pmatrix} 55{,}44 \\ -166{,}7 \end{pmatrix} \Rightarrow \|\vec{x}\|_\infty = 166{,}7$

mit $\vec{b} = \begin{pmatrix} -1{,}9 \\ 0{,}9 \end{pmatrix}$ d.h. $\|\Delta \vec{b}\|_\infty = 0{,}1$, $\frac{\|\Delta \vec{b}\|_\infty}{\|\vec{b}\|_\infty} = \frac{0{,}1}{2} = 0{,}05$

ist $\tilde{\vec{x}} \approx \begin{pmatrix} 52{,}11 \\ -156{,}7 \end{pmatrix}$

somit • absoluter Fehler: $\|\Delta \vec{x}\|_\infty = \|\tilde{\vec{x}} - \vec{x}\|_\infty = \left\| \begin{pmatrix} -3{,}333 \\ 10 \end{pmatrix} \right\| = 10$

somit mit Faktor 100 verstärkt

• relativer Fehler:

$$\frac{\|\Delta \vec{x}\|_\infty}{\|\vec{x}\|_\infty} = \frac{10}{166{,}7} \approx 0{,}06$$

Verstärkung des relativen Fehlers um Faktor 12

II.) Störung der Systemmatrix – $\Delta b = \vec{0}$, $\Delta A \neq 0_{n \times n}$

▶ 12

exaktes Problem: $A \cdot x = b$

gestörte Problem: $\tilde{A} \cdot \tilde{x} = b$

mit $\tilde{A} = A + \Delta A$ und $\tilde{x} = x + \Delta x$ gilt

$\tilde{A} \cdot \tilde{x} = b \Rightarrow (A + \Delta A) \cdot (x + \Delta x) = b$

$\Rightarrow A \cdot x + A \cdot \Delta x + \Delta A \cdot x + \Delta A \cdot \Delta x = b$

$\Rightarrow A \cdot \Delta x + \Delta A \cdot x + \Delta A \cdot \Delta x = \vec{0}$

$\Rightarrow A \cdot \Delta x = -\Delta A \cdot x - \Delta A \cdot \Delta x$

$\Rightarrow \Delta x = -A^{-1} \Delta A\, x - A^{-1} \Delta A \cdot \Delta x$

$\Rightarrow \|\Delta x\| = \|+A^{-1} \Delta A \cdot x + A^{-1} \Delta A \cdot \Delta x\|$ ↘ Dreiecksungleichung Vektornorm

$\Rightarrow \|\Delta x\| \leq \|A^{-1} \Delta A \cdot x\| + \|A^{-1} \cdot \Delta A \cdot \Delta x\|$ ↘ Verträglichkeit

$\Rightarrow \|\Delta x\| \leq \|A^{-1} \cdot \Delta A\| \cdot \|x\| + \|A^{-1} \Delta A\| \cdot \|\Delta x\|$ ↘ submultiplikativität Matrixnorm

$\Rightarrow \|\Delta x\| \leq \|A^{-1}\| \cdot \|\Delta A\| \cdot \|x\| + \|A^{-1}\| \cdot \|\Delta A\| \cdot \|\Delta x\|$

$\Rightarrow \|\Delta x\| \leq \frac{\|A\|}{\|A\|} \cdot \|A^{-1}\| \cdot \|\Delta A\| \cdot \|x\| + \frac{\|A\|}{\|A\|} \|A^{-1}\| \cdot \|\Delta A\| \cdot \|\Delta x\|$

Dr. rer. nat. M. Striebel

$$\Rightarrow \|\Delta x\| \leq \underbrace{\|A\| \cdot \|A^{-1}\|}_{\kappa(A)} \cdot \frac{\|\Delta A\|}{\|A\|} \cdot \|x\| + \underbrace{\|A\| \cdot \|A^{-1}\|}_{\kappa(A)} \cdot \frac{\|\Delta A\|}{\|A\|} \cdot \|\Delta x\|$$

$\frac{\|\Delta A\|}{\|A\|}$: relative Abweichung zwischen A und $\tilde{A}$

$$\Rightarrow \|\Delta x\| \left(1 - \kappa(A) \cdot \frac{\|\Delta A\|}{\|A\|}\right) \leq \kappa(A) \cdot \frac{\|\Delta A\|}{\|A\|} \cdot \|x\|$$

somit ist die relative Abweichung zwischen x und $\tilde{x}$ abschätzbar durch:

$$\frac{\|\tilde{x} - x\|}{\|x\|} = \frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A) \cdot \frac{\|\Delta A\|}{\|A\|}} \cdot \frac{\|\Delta A\|}{\|A\|}$$

$\frac{\kappa(A)}{1 - \kappa(A) \cdot \frac{\|\Delta A\|}{\|A\|}}$: Verstärkungsfaktor der relativen Abweichung in A

⚠ Kondition von A ist von zentraler Bedeutung

III) Störung in der Systemmatrix und der rechten Seite – $\Delta b \neq \vec{0} \wedge \Delta A \neq \mathbb{O}_{n \times n}$

▶ 13

exaktes Problem: $Ax = b$

gestörte Problem: $\tilde{A}\tilde{x} = \tilde{b}$

mit $\tilde{A} = A + \Delta A$, $\tilde{b} = b + \Delta b$ und $\tilde{x} = x + \Delta x$ gelangt man zur Abschätzung des relativen Fehlers:

$$\frac{\|\tilde{x} - x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A) \cdot \frac{\|\Delta A\|}{\|A\|}} \cdot \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|}\right)$$

Verstärkungsfaktor

⚠ Maßgeblich beeinflußt durch die Kondition der Systemmatrix

Dr. rer. nat. M. Striebel

Beispiel einer schlecht konditionierten Matrix

▶ 14

Hilbert-Matrix $H_n = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots & \frac{1}{n+2} \\ \vdots & \vdots & \vdots & & \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{pmatrix} \in \mathbb{R}^{n \times n}$

Kondition $\varkappa(H_n)$ steigt mit steigendem n

```
Kondition für n= 1 : 1.000000
Kondition für n= 2 : 19.281470
Kondition für n= 3 : 524.056778
Kondition für n= 4 : 15513.738739
Kondition für n= 5 : 476607.250241
Kondition für n= 6 : 14951058.641453
Kondition für n= 7 : 475367356.911439
Kondition für n= 8 : 15257575566.627958
Kondition für n= 9 : 493153404551.012146
Kondition für n=10 : 16025028168113.175781
```

via Matlab (cond)

Spektralnorm

Ansatz zur Verbesserung der Kondition: Vorkonditionierung

▶ 15

Ziel: Umformung des Problems so, dass die Lösungsmenge erhalten bleibt, die Kondition aber verkleinert wird

Prinzip: wähle $D \in \mathbb{R}^{n \times n}$ regulär so, dass $\varkappa(D \cdot A) < \varkappa(A)$

damit geht $A \cdot x = b$ über in $(D \cdot A)\, x = (D \cdot b)$

einfaches Verfahren: Zeilenskalierung / Zeilenäquilibrierung (bzgl. Zeilensummennorm)

hier: $D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$ mit $d_k = \left(\sum_{j=1}^{n} |a_{kj}| \right)^{-1}$

Beispiel:

$A = \begin{pmatrix} 8 & 10\,000 \\ 50 & -60 \end{pmatrix} \Rightarrow \varkappa_\infty(A) \approx 201{,}2$

Zeilenskalierung: $D = \begin{pmatrix} 1/10008 & 0 \\ 0 & 1/110 \end{pmatrix}$

liefert: $D \cdot A \approx \begin{pmatrix} 0.799 \cdot 10^{-3} & 0{,}999 \\ 0{,}455 & -0{,}545 \end{pmatrix}$

mit $\varkappa_\infty(D \cdot A) \approx 3{,}40$

Dr. rer. nat. M. Striebel