Numerische Mathematik für Maschinenbauer

Lineare Gleichungssysteme

A. Reusken

K.-H. Brakhage, Saskia Dietze, Thomas Jankuhn

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Sommersemester 2018

Heute in der Vorlesung

Themen: Dahmen & Reusken Kap 3.1-3.2

- Kondition und Störungssätze
- Zeilenskalierung

Was Sie mitnehmen sollten:

- Wie ist das Problem $\boldsymbol{Ax} = \boldsymbol{b}$ konditioniert?
- Warum verwendet man Zeilenskalierung?

Motivation

Problemstellungen (Auswahl)

- Diskretisierung von Integralgleichungen und gewöhnlichen oder partiellen Differentialgleichungen.
- In der linearen Ausgleichsrechnung entsteht auf natürliche Weise ein lineares Gleichungssystem (Normalgleichungen, siehe nächstes Kapitel).
- Zur Lösung eines nichtlinearen Gleichungssystems werden oft Linearisierungsverfahren, wie z.B. das Newton-Verfahren, eingesetzt. Bei so einem Verfahren ergibt sich eine Reihe von linearen Gleichungssystemen (siehe übernächstes Kapitel).

Wahl des Lösungsverfahrens hängt vom Problem ab

- dünnbesetztes vs. vollbesetztes Gleichungssystem
- Strukur (Tridiagonal-, Bandmatrizen)

Beispiel 3.2.

Gesucht $u(x)$, das eine Differentialgleichung vom Typ

$$-u''(x) + \lambda(x)u(x) = f(x), \quad x \in [0,1],$$

mit den Randbedingungen

$$u(0) = u(1) = 0$$

erfüllt (Randwertaufgabe).

Diskretisierung (Gitterpunkte)

$$x_j = jh, \quad j = 0, \dots, n, \quad h = \frac{1}{n},$$

Beispiel 3.2.

Ableitung $\rightarrow$ Differenzenquotienten (Taylorentwicklung)

$$u(x_j+h) = u(x_j)+hu'(x_j)+\frac{h^2}{2}u''(x_j)+\frac{h^3}{6}u'''(x_j)+\mathcal{O}(h^4),$$

$$u(x_j-h) = u(x_j)-hu'(x_j)+\frac{h^2}{2}u''(x_j)-\frac{h^3}{6}u'''(x_j)+\mathcal{O}(h^4),$$

Daraus folgt sofort

$$u(x_j+h) - 2u(x_j) + u(x_j-h) = h^2u''(x_j) + \mathcal{O}(h^4),$$

Beispiel 3.2.

Es gilt also

$$u(x_j + h) - 2u(x_j) + u(x_j - h) = h^2 u''(x_j) + \mathcal{O}(h^4),$$

und somit

$$u''(x_j) = \frac{1}{h^2}\left[u(x_{j+1}) - 2u(x_j) + u(x_{j-1})\right] + \mathcal{O}(h^2).$$

D.h., bis auf Terme 2. Ordnung in der Schrittweite h entspricht die 2. Ableitung einem **Differenzenquotienten**.

Beispiel 3.2.

Diskretisierung:

$$-\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + \lambda(x_j)u_j = f(x_j), \quad j = 1, \ldots, n-1,$$

zusammen mit den Randbedingungen $u_0 = u_n = 0$.

Für kleines h, also großes n:

$$u_j \approx u(x_j).$$

Beispiel 3.2.

Gleichungssystem in Matrixform:

$$\begin{pmatrix} a_{1,1} & -1 & 0 & \dots & 0 \\ -1 & a_{2,2} & -1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & a_{n-1,n-1} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-2} \\ u_{n-1} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-2} \\ b_{n-1} \end{pmatrix}$$

$$a_{i,i} := 2 + h^2 \lambda(x_i)$$
$$b_i := h^2 f(x_i)$$

Beispiel 3.3.

Gesucht $u(x)$, das die Integralgleichung

$$u(x) + 2\int_0^1 \cos(xt)\, u(t)\, dt = 2, \quad x \in [0,1]$$

erfüllt.

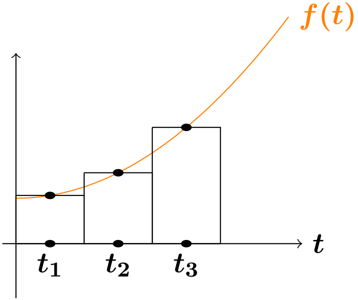
Diskretisierung (Gitterpunkte)

$$t_j = \left(j - \frac{1}{2}\right) h, \quad j = 1, \ldots, n, \quad h = \frac{1}{n}.$$

Beispiel 3.3.

Annäherung des Integrals (Mittelpunktsregel):

$$\int_0^1 f(t)dt \approx h \sum_{j=1}^{n} f(t_j)$$



Beispiel 3.3.

Gleichung nur in den Punkten $x = t_i$ betrachten.

Gleichungssystem für $u_i \approx u(t_i)$:

$$u_i + 2\,h \sum_{j=1}^{n} \cos(t_i t_j)\, u_j = 2, \quad i = 1, \ldots, n.$$

In Matrixform ergibt sich

$$\begin{pmatrix} \frac{1}{2h} + \cos(t_1 t_1) & \cos(t_1 t_2) & \cdots & \cos(t_1 t_n) \\ \cos(t_2 t_1) & \frac{1}{2h} + \cos(t_2 t_2) & & \cos(t_2 t_n) \\ \vdots & & \ddots & \vdots \\ \cos(t_n t_1) & \cdots & & \frac{1}{2h} + \cos(t_n t_n) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \frac{1}{h} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Problemstellung

Notation:

$\mathbb{R}^{m\times n}$ ist die Menge der $(m \times n)$-Matrizen

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

mit Einträgen

$$a_{i,j} \in \mathbb{R}.$$

Problemstellung

Aufgabe

Zu gegebenem $A \in \mathbb{R}^{n\times n}$ und $b = (b_1, \ldots, b_n)^T \in \mathbb{R}^n$ bestimme ein

$$x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n,$$

so dass

$$\begin{array}{ccccccc} a_{1,1}\,x_1 & + & \cdots & + & a_{1,n}\,x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n,1}\,x_1 & + & \cdots & + & a_{n,n}\,x_n & = & b_n \end{array}$$

bzw. kurz

$$Ax = b$$

erfüllt.

Bemerkung

Folgende Aussagen sind äquivalent für das System $\boldsymbol{Ax} = \boldsymbol{b}$ mit $\boldsymbol{A} \in \mathbb{R}^{n \times n}$:

- Das System hat für jedes $\boldsymbol{b} \in \mathbb{R}^{n}$ eine eindeutige Lösung $\boldsymbol{x} \in \mathbb{R}^{n}$.
- Die Matrix $\boldsymbol{A}$ hat vollen Rang n.
- Das *homogene* System $\boldsymbol{Ax} = \boldsymbol{0}$ hat nur die triviale Lösung $\boldsymbol{x} = \boldsymbol{0}$.
- Es gilt $\det \boldsymbol{A} \neq 0$.

$\boldsymbol{A}$ heißt regulär oder nichtsingulär, wenn $\det \boldsymbol{A} \neq 0$.

Annahme: Wir nehmen bei der Lösung linearer Gleichungssysteme stets an, dass $\det \boldsymbol{A} \neq 0$ gilt.

Störung in der rechten Seite $\boldsymbol{b}$

Satz 3.7.

Sei $\boldsymbol{x} + \boldsymbol{\Delta x}$ die Lösung von $\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{\Delta x}) = \boldsymbol{b} + \boldsymbol{\Delta b}$. Dann gilt:

$$\frac{\|\boldsymbol{\Delta x}\|}{\|\boldsymbol{x}\|} \leq \underbrace{\|\boldsymbol{A}^{-1}\| \; \|\boldsymbol{A}\|}_{\boldsymbol{\kappa}_{\|\cdot\|}(\boldsymbol{A})} \; \frac{\|\boldsymbol{\Delta b}\|}{\|\boldsymbol{b}\|}$$

wobei

$$\boldsymbol{\kappa}_{\|\cdot\|}(\boldsymbol{A}) \equiv \|\boldsymbol{A}^{-1}\| \; \|\boldsymbol{A}\|$$

die (relative) Konditionszahl der Matrix $\boldsymbol{A}$ (bzgl. $\|\cdot\|$) ist.

Der relative Fehler der Lösung läßt sich durch das Vielfache

$$\boldsymbol{\kappa}(\boldsymbol{A}) = \boldsymbol{\kappa}_{\|\cdot\|}(\boldsymbol{A}) \equiv \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$$

des relativen Fehlers der rechten Seite (Eingabedaten) abschätzen.

Störung in $\boldsymbol{A}$ und $\boldsymbol{b}$

Satz 3.9

Sei $\boldsymbol{x} + \boldsymbol{\Delta x}$ die Lösung von

$$(\boldsymbol{A} + \boldsymbol{\Delta A})(\boldsymbol{x} + \boldsymbol{\Delta x}) = \boldsymbol{b} + \boldsymbol{\Delta b}.$$

Falls

$$\boldsymbol{\kappa}(\boldsymbol{A}) \frac{\|\boldsymbol{\Delta A}\|}{\|\boldsymbol{A}\|} < 1$$

gilt, folgt für den relativen Fehler

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

Störung in A und b

Beachte

Falls

$$\kappa(A)\,\frac{\|\Delta A\|}{\|A\|} = \|A^{-1}\|\,\|\Delta A\| \ll 1,$$

beschreibt die Konditionszahl $\kappa(A)$ auch maßgeblich den Effekt der Störungen in den übrigen Eingabedaten.

Beispiel 3.11.

Gegeben seien

$$\boldsymbol{A} = \begin{pmatrix} \mathbf{3} & \mathbf{1.001} \\ \mathbf{6} & \mathbf{1.997} \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} \mathbf{1.999} \\ \mathbf{4.003} \end{pmatrix},$$

sowie die gestörten Daten

$$\tilde{\boldsymbol{A}} = \begin{pmatrix} \mathbf{3} & \mathbf{1} \\ \mathbf{6} & \mathbf{1.997} \end{pmatrix}, \quad \tilde{\boldsymbol{b}} = \begin{pmatrix} \mathbf{2.002} \\ \mathbf{4} \end{pmatrix}.$$

Aufgabe: Schätzen Sie den relativen Fehler in der Lösung ab.

Beispiel 3.11.

1. Schritt: Berechne Konditionszahl von $\boldsymbol{A}$

$$\boldsymbol{A^{-1} = \frac{-1}{0.015}\begin{pmatrix} 1.997 & -1.001 \\ -6 & 3 \end{pmatrix}}$$

und damit erhalten wir

$$\boldsymbol{\|A^{-1}\|_\infty = 600, \ \|A\|_\infty = 7.997 \ \Rightarrow \ \kappa_\infty(A) = 4798.2}.$$

2. Schritt: Abschätzen der Störungen

$$\boldsymbol{\Delta A = \tilde{A} - A = \begin{pmatrix} 0 & -0.001 \\ 0 & 0 \end{pmatrix} \Rightarrow \|\Delta A\|_\infty = 0.001,}$$

$$\boldsymbol{\Delta b = \tilde{b} - b = \begin{pmatrix} 0.003 \\ -0.003 \end{pmatrix} \Rightarrow \|\Delta b\|_\infty = 0.003.}$$

3. Schritt: Aus Satz 3.9. ergibt sich

$$\boldsymbol{\frac{\|\Delta x\|_\infty}{\|x\|_\infty} \le 10.49.}$$

Beispiel 3.11.

Mit

$$\boldsymbol{Ax} = \begin{pmatrix} \mathbf{3} & \mathbf{1.001} \\ \mathbf{6} & \mathbf{1.997} \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{-1} \end{pmatrix} = \begin{pmatrix} \mathbf{1.999} \\ \mathbf{4.003} \end{pmatrix} = \boldsymbol{b}$$

sowie

$$\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}} = \begin{pmatrix} \mathbf{3} & \mathbf{1} \\ \mathbf{6} & \mathbf{1.997} \end{pmatrix} \begin{pmatrix} \mathbf{0.2229} \\ \mathbf{1.3333} \end{pmatrix} \begin{pmatrix} \mathbf{2.002} \\ \mathbf{4} \end{pmatrix} = \tilde{\boldsymbol{b}}$$

ergibt sich der tatsächliche relative Fehler

$$\frac{\|\boldsymbol{\Delta x}\|_\infty}{\|\boldsymbol{x}\|_\infty} \approx \mathbf{2.333}$$

gebenüber dem geschätzten Wert von $\mathbf{10.49}$.

Bemerkung 3.10.

In einer Maschine mit Maschinengenauigkeit eps:

$$\frac{\|\boldsymbol{\Delta A}\|}{\|\boldsymbol{A}\|} \leq \mathrm{eps} \quad \text{und} \quad \frac{\|\boldsymbol{\Delta b}\|}{\|\boldsymbol{b}\|} \leq \mathrm{eps}.$$

Nach Satz 3.9 ist wegen der Kondition des Problems

$$\underbrace{(\boldsymbol{A}, \boldsymbol{b})}_{\text{Eingabe}} \rightarrow \underbrace{\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}}_{\text{Ausgabe}}$$

der unvermeidliche Fehler durch

$$\frac{\|\boldsymbol{\Delta x}\|}{\|\boldsymbol{x}\|} = \mathcal{O}(\kappa(\boldsymbol{A})\,\mathrm{eps}).$$

gegeben.

Residuum als Maß für Genauigkeit

Gegeben:

- Gleichungssystem $\boldsymbol{Ax} = \boldsymbol{b}$
- Näherungslösung $\tilde{\boldsymbol{x}}$.

Definition

Das Residuum $\tilde{\boldsymbol{r}}$:

$$\tilde{\boldsymbol{r}} := \boldsymbol{b} - \boldsymbol{A}\tilde{\boldsymbol{x}}.$$

Beachte

- Residuum ist ohne Kenntnis der Lösung $\boldsymbol{x}$ berechenbar
- $\tilde{\boldsymbol{r}} = \boldsymbol{0} \Leftrightarrow \tilde{\boldsymbol{x}} = \boldsymbol{x}$.

Residuum als Maß für Genauigkeit

Frage

Wie aussagekräftig ist die Größe des Residuums in Bezug auf den tatsächlichen Fehler?

Für die Norm des Residuums im Vergleich zu der des Fehlers gilt:

$$\boldsymbol{\kappa(A)^{-1}\,\frac{\|\tilde{r}\|}{\|b\|} \leq \frac{\|x-\tilde{x}\|}{\|x\|} \leq \kappa(A)\,\frac{\|\tilde{r}\|}{\|b\|}}$$

$\Rightarrow$ hängt wieder von der Kondition ab.

Beispiel 3.12.

Sei $\boldsymbol{b} = (\boldsymbol{3}, \boldsymbol{6})^{\boldsymbol{T}}$ und $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{3} & \boldsymbol{1.001} \\ \boldsymbol{6} & \boldsymbol{1.997} \end{pmatrix}$ $(\boldsymbol{\kappa_\infty(A) = 4798.2})$.

Exakte Lösung: $\boldsymbol{x} = (\boldsymbol{1}, \boldsymbol{0})^{\boldsymbol{T}}$.

Für die Annäherungen

$$\tilde{\boldsymbol{x}} = \begin{pmatrix} \boldsymbol{0.99684} \\ \boldsymbol{0.00949} \end{pmatrix}, \qquad \hat{\boldsymbol{x}} = \begin{pmatrix} \boldsymbol{1.000045} \\ \boldsymbol{0.000089} \end{pmatrix},$$

gilt

$$\|\tilde{\boldsymbol{r}}\|_\infty = \|\boldsymbol{b} - \boldsymbol{A}\tilde{\boldsymbol{x}}\|_\infty = \boldsymbol{1.95 \cdot 10^{-5}},$$

$$\|\hat{\boldsymbol{r}}\|_\infty = \|\boldsymbol{b} - \boldsymbol{A}\hat{\boldsymbol{x}}\|_\infty = \boldsymbol{4.48 \cdot 10^{-4}}.$$

Die Norm des Residuums für $\tilde{\boldsymbol{x}}$ ist also viel kleiner als für $\hat{\boldsymbol{x}}$:

$$\|\tilde{\boldsymbol{r}}\|_\infty \ll \|\hat{\boldsymbol{r}}\|_\infty.$$

Der Fehler in $\tilde{\boldsymbol{x}}$ ist aber viel größer als in $\hat{\boldsymbol{x}}$:

$$\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_\infty = \boldsymbol{9.49 \cdot 10^{-3}} \gg \|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_\infty = \boldsymbol{8.90 \cdot 10^{-5}}.$$

Zeilenskalierung

Motivation

$\kappa(A)$ ist verantwortlich für die Datenfehlerverstärkung

$\Rightarrow$ wesentlicher Faktor in der Lösung linearer Gleichungssysteme.

Frage

- Kann man die Kondition einer Matrix verbessern?

$\Rightarrow$ Zeilenskalierung (Zeilenäquilibrierung), d.h. Multiplikation der i-ten Zeile der Matrix mit einer Zahl $d_i \neq 0$ $(1 \leq i \leq n)$.

Zeilenskalierung

Forme das Gleichungssystem $Ax = b$ in ein äquivalentes Systems

$$\underbrace{D_z A}_{A_{\text{neu}}} x = \underbrace{D_z b}_{b_{\text{neu}}}$$

um, wobei D_z die Diagonalmatrix $D_z = \mathbf{diag}(d_1, \dots, d_n)$ bezeichnet.

Ziel: Wähle D_z so, dass die Kondition der Matrix A_{neu} (wesentlich) verbessert wird.

Sei D_z die Diagonalmatrix mit Diagonaleinträgen definiert durch

$$d_i = \left(\sum_{j=1}^{n} |a_{i,j}| \right)^{-1}$$

Die Matrix $D_z A$ nennt man skalierte Matrix.

Zeilenskalierung

Für die skalierte Matrix $D_z A$ gilt

$$\sum_{j=1}^{n} |(D_z A)_{i,j}| = 1 \quad \text{für alle } i,$$

also sind die Betragssummen aller Zeilen gleich eins. Eine Matrix mit dieser Eigenschaft heißt zeilenweise äquilibriert.

Optimalitätseigenschaft

$\kappa_\infty(D_z A) \leq \kappa_\infty(DA)$ für jede reguläre Diagonalmatrix D.

$\Rightarrow$ Zeilenskalierung mit D_z liefert die minimale Konditionszahl bezüglich der Maximumnorm.

Beispiel 3.14.

Für

$$A = \begin{pmatrix} 8 & 10000 \\ 50 & -60 \end{pmatrix} \qquad \begin{array}{l} \xrightarrow{\sum_{j=1}^{2} |a_{1,j}|} \quad 10008 \\ \xrightarrow{\sum_{j=1}^{2} |a_{2,j}|} \quad 110 \end{array}$$

erhält man $\kappa_\infty(A) = 201.2$.

Mit der Diagonalmatrix

$$D_z = \begin{pmatrix} \frac{1}{10008} & 0 \\ 0 & \frac{1}{110} \end{pmatrix}$$

ergibt sich

$$D_z A = \begin{pmatrix} 0.799 \cdot 10^{-3} & 0.999 \\ 0.455 & -0.545 \end{pmatrix}$$

und damit $\kappa_\infty(D_z A) = 3.40$.

Rechenaufwand

Gegeben seien Matrizen $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times n}$ und Vektoren $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{n}$.

Wie hoch ist der Rechenaufwand zur Bestimmung von

- Skalarprodukt $\boldsymbol{x}^T \boldsymbol{y}$: $\boldsymbol{2\,n - 1}$ $\quad \mathcal{O}(n)$
- Dyadisches Produkt $\boldsymbol{x}\,\boldsymbol{y}^T$: $\boldsymbol{n^2}$ $\quad \mathcal{O}(n^2)$
- Matrix-Vektor Produkt $\boldsymbol{A}\,\boldsymbol{x}$: $\boldsymbol{2\,n^2 - n}$ $\quad \mathcal{O}(n^2)$
- Matrix-Matrix Produkt $\boldsymbol{A}\,\boldsymbol{B}$: $\boldsymbol{2\,n^3 - n^2}$ $\quad \mathcal{O}(n^3)$

Beachte

Bei der Zählung der Rechenoperationen in einem Algorithmus werden

- (traditionsgemäß) nur Multiplikationen und Divisionen, und
- nur Terme höchster Ordnung gezählt.

Zusammenfassung

- $\kappa(A)$ spielt eine zentrale Rolle bei der Kondition des Problems $(A, b) \to x = A^{-1}b$, siehe Satz 3.9.

- Residuum als Maß für die Genauigkeit: nicht immer aussagekräftig.

- Zeilenskalierung verbessert $\kappa_\infty(A)$.

Verständnisfragen

Es seien $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ beliebig aber regulär, $\boldsymbol{b} \in \mathbb{R}^{n}$ und gesucht sei die Lösung $\boldsymbol{x}^* \in \mathbb{R}^{n}$ von $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$.

Es sei $\boldsymbol{B} := \boldsymbol{D}\boldsymbol{A}$ die zeilenäquilibrierte Matrix zu $\boldsymbol{A}$. Geben Sie $\|\boldsymbol{B}\|_\infty$ an. 1

f Es sei $\tilde{\boldsymbol{x}}$ eine Annäherung der Lösung $\boldsymbol{x}$ und $\boldsymbol{r} := \boldsymbol{b} - \boldsymbol{A}\,\tilde{\boldsymbol{x}}$ das zugehörige Residuum. Es gilt $\|\boldsymbol{r}\| \leq \boldsymbol{\kappa}(\boldsymbol{A})\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|$, mit $\boldsymbol{\kappa}(\boldsymbol{A}) := \|\boldsymbol{A}\|\,\|\boldsymbol{A}^{-1}\|$.

w Es sei $\boldsymbol{B} \in \mathbb{R}^{n \times n}$ beliebig aber regulär und $\boldsymbol{\kappa}(\cdot)$ die Konditionszahl bzgl. $\|\cdot\|$. Es gilt $\boldsymbol{\kappa}(\boldsymbol{A}\,\boldsymbol{B}) \leq \boldsymbol{\kappa}(\boldsymbol{A})\,\boldsymbol{\kappa}(\boldsymbol{B})$.