

Numerische Mathematik für Maschinenbauer

Lineare Gleichungssysteme

A. Reusken

K.-H. Brakhage, I. Voulis, H. Saß

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Sommersemester 2017

Heute in der Vorlesung

Themen: Dahmen & Reusken Kap. 3.1-3.4

- ▶ Kondition und Störungssätze
- ▶ Zeilenskalierung
- ▶ Dreiecksmatrizen

Was Sie mitnehmen sollten:

- ▶ Wie ist das Problem $\mathbf{Ax} = \mathbf{b}$ konditioniert?
- ▶ Warum verwendet man Zeilenskalierung?
- ▶ Wie rechnet man mit Dreiecksmatrizen?

Motivation

- ▶ **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
- ▶ Struktur (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

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

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

Beispiel 3.2.

Diskretisierung:

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

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:

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

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

$$b_i := h^2 f(x_i)$$

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.

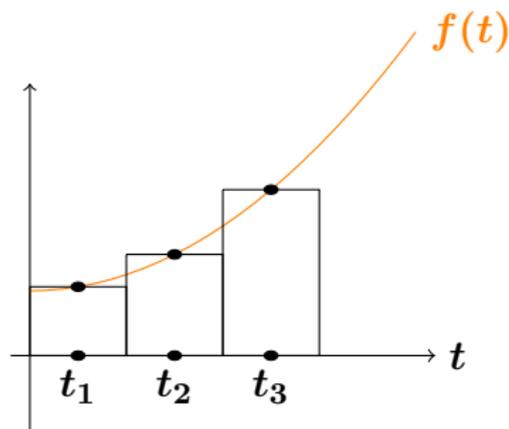
Diskretisierung (Gitterpunkte)

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

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 $u_i \approx u(t_i)$:

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

In Matrixform ergibt sich

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

Problemstellung

Notation: $\mathbb{R}^{m \times n}$ ist die Menge der Matrizen

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

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

Aufgabe

Zu gegebenem $A \in \mathbb{R}^{n \times n}$ und $\mathbf{b} = (b_1, \dots, b_n)^T \in \mathbb{R}^n$
bestimme ein $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, so dass

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

bzw. kurz $A\mathbf{x} = \mathbf{b}$ erfüllt.

Bemerkung

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

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

A heißt **regulär** oder **nichtsingulär**, wenn $\det A \neq 0$.

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

Störung in der rechten Seite \mathbf{b}

Satz 3.7

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

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \underbrace{\|\mathbf{A}^{-1}\| \|\mathbf{A}\|}_{\kappa_{\|\cdot\|}(\mathbf{A})} \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

wobei

$$\kappa_{\|\cdot\|}(\mathbf{A}) \equiv \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$$

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

Der relative Fehler der Lösung läßt sich durch das Vielfache

$$\kappa(\mathbf{A}) = \kappa_{\|\cdot\|}(\mathbf{A}) \equiv \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$$

des relativen Fehlers der rechten Seite (Eingabedaten) abschätzen.

Störung in A und b

Satz 3.9

Sie $x + \Delta x$ die Lösung von

$$(A + \Delta A)(x + \Delta x) = b + \Delta b.$$

Falls $\kappa(A) \frac{\|\Delta A\|}{\|A\|} < 1$ gilt, folgt für den relativen Fehler

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

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

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

sowie die gestörten Daten

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

Schätzen Sie den relativen Fehler in der Lösung ab.

1. Schritt: Berechne Konditionszahl von A

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

und damit erhalten wir

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

Beispiel 3.11.

2. Schritt: Abschätzen der Störungen

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

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

3. Schritt: Aus Satz 3.9 ergibt sich

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

Mit exakter Lösung $x = (1, -1)^T$ von $Ax = b$ und gestörter Lösung $\tilde{x} = (0.2229, 1.3333)^T$ von $\tilde{A}\tilde{x} = \tilde{b}$ ergibt sich der tatsächliche relative Fehler

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

Bemerkung 3.10

In einer Maschine mit Maschinengenauigkeit eps :

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

Nach Satz 3.9 ist wegen der Kondition des Problems

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

der unvermeidliche Fehler durch

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

gegeben.

Residuum als Maß für Genauigkeit

- Gegeben: – Gleichungssystem $Ax = b$
– Näherungslösung \tilde{x} .

Definition

Das Residuum \tilde{r} :

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

- Beachte: – Residuum ist ohne Kenntnis der Lösung x berechenbar
– $\tilde{r} = \mathbf{0} \Leftrightarrow \tilde{x} = x$.

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:

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

\Rightarrow hängt wieder von der Kondition ab.

Beispiel 3.12.

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

Exakte Lösung: $x = (1, 0)^T$.

Für die Annäherungen

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

gilt

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

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

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

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

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

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

Zeilenskalierung

Motivation

$\kappa(\mathbf{A})$ ist verantwortlich für die Datenfehlerverstärkung
⇒ wesentlicher Faktor in der Lösung linearer Gleichungssysteme.

Frage: Kann man die Kondition einer Matrix verbessern?
⇒ Zeilenskalierung (Zeilenäquilibrierung), d.h. Multiplikation der i -ten Zeile der Matrix mit einer Zahl $d_i \neq 0$ ($1 \leq i \leq n$).

Forme das Gleichungssystem $\mathbf{A}\mathbf{x} = \mathbf{b}$ in ein äquivalentes Systems

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

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

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

Zeilenskalierung

Sei D_z die Diagonalmatrix mit Diagonaleinträgen definiert durch

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

Für die skalierte Matrix 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

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

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

Mit der Diagonalmatrix

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

ergibt sich

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

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

Rechenaufwand

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

Wie hoch ist der Rechenaufwand zur Bestimmung von

- ▶ Skalarprodukt $x^T y$: $2n - 1$ $\mathcal{O}(n)$
- ▶ Dyadisches Produkt $x y^T$: n^2 $\mathcal{O}(n^2)$
- ▶ Matrix-Vektor Produkt Ax : $2n^2 - n$ $\mathcal{O}(n^2)$
- ▶ Matrix-Matrix Produkt AB : $2n^3 - n^2$ $\mathcal{O}(n^3)$

Beachte

Bei der Zählung der Rechenoperationen in einem Algorithmus werden

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

Beispiel 3.16.

Löse das Gleichungssystem $Rx = b$, wobei

$$R = \begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix}, \quad \text{und} \quad b = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}.$$

Die spezielle Struktur von R erlaubt eine einfache Lösung:

$$x_3 = 4/2 = 2$$

$$x_2 = 1 - 3 \cdot 2 = -5$$

$$x_1 = \frac{1}{3}(0 - (-1)(-5) - 2 \cdot 2) = -3$$

\Rightarrow Rückwärtseinsetzen

Dreiecksmatrizen, Rückwärtseinsetzen

Definition

Eine Matrix $R = (r_{i,j})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ heißt *obere Dreiecksmatrix*, falls

$$r_{i,j} = 0 \text{ für } i > j$$

gilt.

Allgemeiner Fall:

$$\begin{array}{rcccccc}
 r_{1,1} x_1 & + & r_{1,2} x_2 & + & \dots & + & r_{1,n} x_n & = & b_1 \\
 & & r_{2,2} x_2 & + & \dots & + & r_{2,n} x_n & = & b_2 \\
 & & & & \ddots & & \vdots & & \vdots \\
 & & & & & & & & & r_{n-1,n-1} x_{n-1} & + & r_{n-1,n} x_n & = & b_{n-1} \\
 & & & & & & & & & & & r_{n,n} x_n & = & b_n.
 \end{array}$$

Dreiecksmatrizen, Rückwärtseinsetzen

Lösbarkeit

Da

$$\det R = r_{1,1} r_{2,2} \cdots r_{n,n},$$

ist $Rx = b$ genau dann stets eindeutig lösbar, wenn alle Diagonaleinträge $r_{j,j}$, $j = 1, \dots, n$, von Null verschieden sind.

Vorgehen:

- ▶ Beginnend bei der letzten Gleichung

$$x_n = b_n / r_{n,n}.$$

- ▶ Einsetzen von x_n in die zweitletzte Gleichung

$$x_{n-1} = (b_{n-1} - r_{n-1,n} x_n) / r_{n-1,n-1}.$$

- ▶ ...

Dreiecksmatrizen, Rückwärtseinsetzen

Rückwärtseinsetzen

Für $j = n, n - 1, \dots, 2, 1$ berechne

$$x_j = \left(b_j - \sum_{k=j+1}^n r_{j,k} x_k \right) / r_{j,j},$$

wobei die Summe für $j = n$ leer ist und als Null interpretiert wird.

Analog: untere Dreiecksmatrix L

- ▶ $L = (l_{i,j})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ mit $l_{i,j} = 0$ für $i < j$
- ▶ Eindeutig lösbar, wenn $l_{j,j}$, $j = 1, \dots, n$, von Null verschieden
- ▶ Vorwärtseinsetzen

Rechenaufwand

Für jedes $j = n - 1, \dots, 1$:

$n - j$ Multiplikationen / Additionen,
eine Division,

und für $j = n$ eine Division. Also insgesamt

- ▶ $\sum_{j=1}^{n-1} (n - j) = \frac{n(n - 1)}{2}$ Additionen / Multiplikationen,
- ▶ n Divisionen.

Rechenaufwand für Rückwärtseinsetzen

ca. $\frac{1}{2}n^2$ Operationen

Operation = Multiplikation oder Division.

Eigenschaften 3.18

- ▶ Das Produkt von oberen (unteren) Dreiecksmatrizen ist wieder eine obere (untere) Dreiecksmatrix.
- ▶ Die Inverse einer oberen (unteren) nichtsingulären Dreiecksmatrix ist wieder eine obere (untere) Dreiecksmatrix.
- ▶ Die Determinante einer Dreiecksmatrix ist gerade das Produkt aller Diagonaleinträge.
- ▶ Die Eigenwerte einer Dreiecksmatrix sind gerade die Diagonaleinträge.

Zusammenfassung

- ▶ $\kappa(\mathbf{A})$ spielt eine zentrale Rolle bei der Kondition des Problems $(\mathbf{A}, \mathbf{b}) \rightarrow \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$, siehe Satz 3.9.
- ▶ **Residuum** als Maß für die Genauigkeit: nicht immer aussagekräftig.
- ▶ **Zeilenskalierung** verbessert $\kappa_{\infty}(\mathbf{A})$.
- ▶ Gleichungssysteme mit einer **Dreiecksmatrix** kann man effizient ($\frac{1}{2}n^2$ Operationen) lösen.

Verständnisfragen

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

- Es sei $R \in \mathbb{R}^{n \times n}$ eine reguläre obere Dreiecksmatrix. Der Rechenaufwand zur Bestimmung der Lösung y von $Ry = b$ über Rückwärtseinsetzen beträgt etwa $n^3/3$ Operationen.

Es sei $B := DA$ die zeilenäquilibrierte Matrix zu A . Geben Sie $\|B\|_\infty$ an.

- Es sei \tilde{x} eine Annäherung der Lösung x und $r := b - A\tilde{x}$ das zugehörige Residuum. Es gilt $\|r\| \leq \kappa(A)\|\tilde{x} - x\|$, mit $\kappa(A) := \|A\|\|A^{-1}\|$.
- Es sei $B \in \mathbb{R}^{n \times n}$ beliebig aber regulär und $\kappa(\cdot)$ die Konditionszahl bzgl. $\|\cdot\|$. Es gilt $\kappa(AB) \leq \kappa(A)\kappa(B)$.