Kap. I: Lösung linearer Gleichungssysteme I (direkte Verfahren)

Lineares Gleichungssystem: Gegeben sei eine Matrix A und ein Vektor b,

$$A = (a_{jk})_{\substack{j=1,\dots,m \\ k=1,\dots,n}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad b = (b_j)_{j=1,\dots,m} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Gesucht ist ein Vektor $x = (x_k)_{k=1,\dots,n}$ mit

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

in Matrixschreibweise

$$Ax = b.$$

Kriterien zur Lösbarkeit (Lineare Algebra):

a) $Ax = b$ ist lösbar genau dann, wenn $\mathrm{Rang}(A) = \mathrm{Rang}([A, b])$ gilt, mit der zusammengesetzten Matrix

$$[A, b] := \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

b) Im quadratischen Fall $m = n$ sind folgende Aussagen äquivalent:

(i) Das LGS $Ax = b$ ist für jedes b eindeutig lösbar.

(ii) Das homogone LGS $Ax = 0$ hat die einzige Lösung $x = 0$.

(iii) $\text{Rang}(A) = n$.

(iv) $\det A \neq 0$.

(v) A ist regulär, d.h. invertierbar.

(vi) Alle Eigenwerte von A sind ungleich Null.

Bemerkung: Die invertierbaren (regulären) $n \times n$-Matrizen bilden mit dem Matrixprodukt eine *Gruppe*, bezeichnet mit $GL(n)$ (general linear group). Untergruppen von $GL(n)$ bilden die invertierbaren oberen (unteren) Dreiecksmatrizen, die invertierbaren Diagonalmatrizen, die sog. Orthogonalmatrizen (im Reellen) bzw. unitären Matrizen (im Komplexen).

I.1. Normen und Fehlerabschätzungen

Im folgenden sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

1.1 Definition:
Eine *Norm* auf einem $\mathbb{K}$-Vektoraum V ist eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ mit folgenden Eigenschaften:

(i) $\|v\| \geq 0 \quad$ für alle $v \in V$, und $(\|v\| = 0 \Longleftrightarrow v = 0)$; (Definitheit)

(ii) $\|\lambda v\| = |\lambda| \cdot \|v\| \quad$ für alle $v \in V, \quad \lambda \in \mathbb{K}$; (pos. Homogenität)

(iii) $\|v + w\| \leq \|v\| + \|w\| \quad$ für alle $v, w \in V$ (Dreiecksungleichung)

Ein Vektorraum V mit einer Norm $\|\cdot\|$ heißt ein *normierter Raum.*

Sprechweise: $\|v\|$ ist die "Norm von v" (eine Art Länge von v)

Bemerkung: Eine Norm induziert eine Metrik $d(v, w) = \|v - w\|$, und damit die topologischen Begriffe wie

- Konvergenz einer Folge
- Cauchy-Folge
- offene und abgeschlossene Teilmengen von V, ϵ-Umgebungen
- Stetigkeit von Abbildungen auf V

Als Schlussfolgerung aus der Dreiecksungleichung ergibt sich

$$\|v \pm w\| \geq \Big| \|v\| - \|w\| \Big|,$$

und hieraus folgt die Stetigkeit der Norm als Funktion von V nach $\mathbb{R}$.

1.2 Definition: Zwei Normen $\|\cdot\|$ und $\|\cdot\|'$ auf V heißen *äquivalent*, wenn es positive Zahlen m, M gibt, so daß für alle $v \in V$ gilt:

$$m\|v\| \leq \|v\|' \leq M\|v\|.$$

Bemerkung: Äquivalente Normen führen zu denselben Begriffen von konvergenten Folgen, Cauchy-Folgen, offenen und abgeschlossenen Mengen, beschränkten Mengen, stetigen Abblildungen.

Beispiele:

(i) Die Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ auf $\mathbb{K}^n$ sind äquivalent.

(ii) Auf $C_{[a,b]}$ sind $\|\cdot\|_1$ und $\|\cdot\|_\infty$ **nicht** äquivalent.

Normen auf endlichdimensionalen Vektorräumen

1.3 Satz
Je zwei Normen auf einem endlich-dimensionalen Vektorraum sind äquivalent.

Es gibt also z.B. im $\mathbb{K}^n$ keine zwei verschiedenen Konvergenzbegriffe aufgrund der Verwendung verschiedener Normen! Konvergenz von Vektoren ist dasselbe wie komponentenweise Konvergenz.

Normen für Matrizen

Der Vektorraum $\mathbb{K}^{n\times n}$ der quadratischen $n\times n$-Matrizen ist isomorph zum $\mathbb{K}^{n^2}$; also sind alle Normen auf $\mathbb{K}^{n\times n}$ äquivalent, und die Konvergenz

$$A_n \to A$$

einer Folge (A_n) von Matrizen ist gleichbedeutend mit der komponentenweisen Konvergenz.

Man betrachtet vornehmlich Normen, die mit der Bedeutung von Matrizen als lineare Abbildungen von $\mathbb{K}^n$ nach $\mathbb{K}^n$ zu tun haben:

1.4 Definition: Auf $\mathbb{K}^n$ sei eine Norm $\|\cdot\|$ gegeben. Eine Norm auf $\mathbb{K}^{n\times n}$ (ebenfalls $\|\cdot\|$ geschrieben) heißt

(i) verträglich mit der Vektornorm $\|\cdot\|$, wenn

$$\|Ax\| \leq \|A\|\,\|x\|, \qquad x\in\mathbb{K}^n, \quad A\in\mathbb{K}^{n\times n}$$

(ii) *Matrixnorm* oder *submultiplikativ*, wenn

$$\|AB\| \leq \|A\|\,\|B\|, \qquad A,B\in\mathbb{K}^{n\times n}.$$

1.5 Satz und Definition: Auf $\mathbb{K}^n$ sei eine Norm $\|\cdot\|$ gegeben. Dann definiert

$$\|A\| := \max_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

eine Norm auf $\mathbb{K}^{n \times n}$, die submultiplikativ und mit der gegebenen Norm auf $\mathbb{K}^n$ verträglich ist. Sie heißt die zugehörige *Operatornorm* oder *natürliche Matrixnorm*.

1.6 Satz

(i) Die zur Maximumsnorm $\|\cdot\|_\infty$ auf $\mathbb{K}^n$ gehörende Operatornorm ist die sog. *Zeilensummennorm*

$$\|A\|_\infty := \max_{j=1,\ldots,n} \sum_{k=1}^{n} |a_{jk}|.$$

(ii) Die zur Betragssummennorm $\|\cdot\|_1$ auf $\mathbb{K}^n$ gehörende Operatornorm ist die sog. *Spaltensummennorm*

$$\|A\|_1 := \max_{k=1,\ldots,n} \sum_{j=1}^{n} |a_{jk}|.$$

(iii) Die Matrix-Norm

$$\|A\|_F := \left(\sum_{j,k=1}^{n} |a_{jk}|^2 \right)^{1/2}$$

heißt Frobenius-Norm und ist mit der $\|\cdot\|_2$-Norm auf $\mathbb{K}^n$ verträglich sowie submultiplikativ. Sie ist NICHT die zu $\|\cdot\|_2$ gehörende Operatornorm!

(iv) Die zur euklidischen Norm $\|\cdot\|_2$ gehörende Operatornorm heißt die Spektralnorm:

$$\|A\|_2 := \max_{\|x\|_2=1} \|Ax\|_2.$$

Äquivalente Ausdrücke sind

$$\|A\|_2 = \max_{\|x\|_2=1} \sqrt{\overline{x}^T \overline{A}^T A x} = \max\{\sqrt{|\lambda|} \ \mid \ \lambda \text{ Eigenwert von } \overline{A}^T A\}$$

1.7 Skalarprodukt in $\mathbb{K}^n$

1.8 Orthonormalbasen und unitäre Matrizen (C)

1.9 reell-symmetrische und hermitesche Matrizen

Beweis von 1.6(iv): erfolgt mit der Hauptachsentransformation der hermiteschen Matrix $\overline{A}^T A$.

1.10 Spezialfall: Für die Spektralnorm einer hermiteschen Matrix A gilt

$$\|A\|_2 = \max\{|\lambda| \;\mid\; \lambda \text{ Eigenwert von } A\}$$

1.11 Positiv definite Matrizen

1.12 Zush. von verträglichen Normen und dem Spektralradius einer Matrix.

Fehleranalyse bei linearen Gleichungssystemen

gegeben: reguläre Matrix $A \in \mathbb{K}^{n\times n}$, rechte Seite $b \in \mathbb{K}^n$
bestimme: Lösung des linearen Gleichungssystems $Ax = b$

Frage: Wie wirken sich **Störungen** der Daten (also von A und b) auf die Lösung x aus?

Vergleich der Lösungen von

$$Ax = b \qquad \text{und} \qquad (A + \delta A)\,(x + \delta x) = b + \delta b.$$

<u>Absolute Fehler:</u> $\|\delta A\|, \quad \|\delta b\|, \quad \|\delta x\|$

<u>Relative Fehler:</u> $\dfrac{\|\delta A\|}{\|A\|}, \quad \dfrac{\|\delta b\|}{\|b\|}, \quad \dfrac{\|\delta x\|}{\|x\|}$

1.13 Hilfssatz: Für die Matrix $B \in \mathbb{K}^{n\times n}$ gelte $\|B\| < 1$ mit einer natürlichen Matrixnorm. Dann ist die Matrix $I + B$ regulär und es gilt

$$\|(I + B)^{-1}\| \le \frac{1}{1 - \|B\|}.$$

1.14 Störungssatz: Die Matrix $A \in \mathbb{K}^{n\times n}$ sei regulär und es sei

$$\|\delta A\| < \frac{1}{\|A^{-1}\|}.$$

Dann ist auch die gestörte Matrix $A + \delta A$ regulär, und für den relativen Fehler der Lösung gilt

$$\frac{\|\delta x\|}{\|x\|} \le \frac{\operatorname{cond}(A)}{1 - \operatorname{cond}(A)\frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right)$$

mit der *Konditionszahl* von A

$$\operatorname{cond}(A) = \|A\| \, \|A^{-1}\|.$$

Beweis: wir zeigen nur für $\Delta A = 0$, dass

$$\frac{\|\Delta\vec{x}\|}{\|\vec{x}\|} \leq \operatorname{cond} A \ \frac{\|\Delta\vec{b}\|}{\|\vec{b}\|}$$

gilt. Also:

$$\begin{aligned}\frac{\|\Delta\vec{x}\|}{\|\vec{x}\|} &= \frac{\|\vec{x} + \Delta\vec{x} - \vec{x}\|}{\|\vec{x}\|} = \frac{\|A^{-1}(\vec{b} + \Delta\vec{b}) - A^{-1}\vec{b}\|}{\|A^{-1}\vec{b}\|} \\ &= \frac{\|A^{-1}\Delta\vec{b}\|}{\|A\|\ \|A^{-1}\vec{b}\|} \cdot \|A\| \leq \frac{\|A^{-1}\|\ \|\Delta\vec{b}\|}{\|AA^{-1}\vec{b}\|} \cdot \|A\| = \operatorname{cond} A \ \frac{\|\Delta\vec{b}\|}{\|\vec{b}\|}.\end{aligned}$$

Bemerkung: Es gelten

- $\operatorname{cond} A \geq 1$, denn

  $$\operatorname{cond} A = \|A\|\ \|A^{-1}\| \geq \|AA^{-1}\| = \|E_n\| = 1.$$

  (Die letzte Gleichung gilt für jede natürliche Matrixnorm.)

- $\operatorname{cond}(\alpha A) = \operatorname{cond} A$ für alle $\alpha \in \mathbb{K}$, $\alpha \neq 0$.