3: Konditionierung der Lösung linearer Gleichungssysteme

3.1 Lösbarkeit

Gegeben Matrix A, Vektor b:

$$A = (a_{ij})_{i,j=1}^{m,n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad b = (b_i)_{i=1}^{m} = \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_m \end{bmatrix}$$

Gesucht ist $x = (x_j)_{j=1}^{n} = (x_1, \dots, x_n)^T$ so dass

$$\forall i : \sum_{j=1}^{n} a_{ij} x_j = b_i ,$$

bzw.

$$Ax = b. \qquad (3.1)$$

Die Zahlen a_{ij}, b_i, x_j können aus $\mathbb{R}$ oder $\mathbb{C}$ sein.

Das Gleichungssystem (3.1) heißt

- unterbestimmt falls $m < n$
- quadratisch falls $m = n$
- überbestimmt falls $m > n$

bei maximalem Rang von A:
i. allg. viele Lösungen
genau eine Lösung
$b \in \text{Bild}(A)$.

Es gibt mindestens eine Lösung falls

$$\text{Rang}(A) = \text{Rang}([A|b]) = \text{Rang}\left(\begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{bmatrix} \right)$$

Für <u>quadratische</u> Matrizen sind folg. Aussagen äquivalent:

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

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

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

(iv) A hat keinen Eigenwert Null.

3.2 Vektornormen

Quantifizieren von Fehlern erfordert Normen.

Im Folgenden sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ der zugrunde liegende Körper und $\mathbb{K}^n$ der darauf aufbauende Vektorraum.

Definition 3.1 (Vektornorm)

Eine Abbildung $\|\cdot\| : \mathbb{K}^n \to \mathbb{R}_+$ heißt Norm falls gilt

(N1)	$\|x\| > 0$	$x \neq 0$	(Definitheit)
(N2)	$\|\alpha \cdot x\| = \|\alpha\| \|x\|$	$x \in \mathbb{K}^n, \alpha \in \mathbb{K}$	(positive Homogenität)
(N3)	$\|x+y\| \leq \|x\| + \|y\|$	$x, y \in \mathbb{K}^n$	(Subadditivität, Dreiecksungl.) ∎

Beispiel 3.2 Häufig verwendete Normen sind:

$$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$$ Euklidische Norm, ℓ_2-Norm

$$\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$$ Maximumnorm, ℓ_∞-Norm

$$\|x\|_1 = \sum_{i=1}^n |x_i|$$ ℓ_1-Norm

Man zeigt über die Konvergenz von Folgen von Vektoren:

$$\forall i = 1,\dots,n: x_i^{(t)} \to x_i \quad (t \to \infty) \iff \|x^{(t)} - x\| \to 0 \quad (t \to \infty)$$

(Konvergenz der Normen äquivalent zu komponentenweiser Konvergenz).
↝ siehe nächster Satz

Eine wichtige Folgerung aus (N3) ist die inverse Dreiecksungleichung

$$\|x - y\| \geq \big| \|x\| - \|y\| \big| \quad \forall x, y \in \mathbb{K}^n. \qquad (3.2)$$

Beweis:

$$\|x\| = \|x - y + y\| \overset{(N3)}{\leq} \|x - y\| + \|y\| \Rightarrow \|x - y\| \geq \|x\| - \|y\|$$

$$\|y\| = \|y - x + x\| \leq \|y - x\| + \|x\| \Rightarrow \|x - y\| \geq \|y\| - \|x\| = -(\|x\| - \|y\|)$$

∎

Aus (3.2) folgt (Stetigkeit der Funktion $\|\cdot\| : \mathbb{K}^n \to \mathbb{R}_+$):

$$\|x^{(t)} - x\| \to 0 \quad (t \to \infty) \Rightarrow \|x^{(t)}\| \to \|x\| \quad (t \to \infty)$$

Satz 1.3 (Äquivalenz aller Normen)

Auf $\mathbb{K}^n$ sind alle Normen im folgenden Sinne äquivalent:
Zu $\|.\|$, $\|.\|'$ gibt es Zahlen $m, M > 0$ aus $\mathbb{R}$ so dass gilt

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

Beweis: Es genügt dies zu zeigen für $\|.\|' = \|.\|_\infty$.
Seien $e_1, \dots, e_n$ die kartesischen Einheitsvektoren. Jedes $x \in \mathbb{K}^n$ hat die Darstellung $x = \sum_{i=1}^n x_i e_i$, also $\|x\| = \|\sum_{i=1}^n x_i e_i\| \leq \sum_{i=1}^n |x_i| \|e_i\| \leq \left(\max_{1\leq i\leq n} |x_i|\right) \sum_{i=1}^n \|e_i\| = \|x\|_\infty \underbrace{\sum_i \|e_i\|}_{=\gamma}$

wg. $\|x\| \leq \gamma \|x\|_\infty$ ist $\|.\|$ stetig bezügl. komponentenweiser Konvergenz.
Punktmenge $S := \{x \in \mathbb{K}^n : \|x\|_\infty = 1\} \subset \mathbb{K}^n$ ist – beschränkt (bzgl. $\|.\|$ da $\|x\| \leq \gamma$)
– abgeschlossen, d.h. Grenzwerte von Folgen aus S sind in S (da $\mathbb{K}^n$ metrischer Raum).
$\Rightarrow$ kompakt.

Die Funktion $\|.\| : S \to \mathbb{R}_+$ nimmt auf S ihr Minimum und Maximum an, d.h.:
Es gibt $\underline{x}, \bar{x} \in S$:

$$\underbrace{0 < \|\underline{x}\|}_{\text{da } 0 \notin S,\ \underline{x} \in S \text{ und N1.}} \leq \|x\| \leq \|\bar{x}\| \leq \gamma \underbrace{\|\bar{x}\|_\infty}_{\substack{=1 \\ \text{da } \bar{x} \in S}} = \gamma \qquad \forall x \in S$$

Nun sei $y \in \mathbb{K}^n \setminus \{0\}$ beliebig.
Dann ist $y/\|y\|_\infty \in S$ (da Norm 1) und somit

$$\|\underline{x}\| \leq \|y/\|y\|_\infty\| \underset{(N2)}{=} \frac{1}{\|y\|_\infty}\|y\| \leq \|\bar{x}\|$$

$$\Longleftrightarrow \quad \underbrace{\|\underline{x}\|}_{=m} \|y\|_\infty \leq \|y\| \leq \underbrace{\|\bar{x}\|}_{=M} \|y\|_\infty \qquad \boxtimes$$

Vorsicht: m, N hängen i.d.R. von der Dimension n ab.
Aufpassen wenn man etwas für $n \to \infty$ beweist.

Nochmal zur Konvergenz:
$x_i^{(t)} \to x_i \ (t \to \infty) \Longleftrightarrow \|x^{(t)} - x\|_\infty \to 0 \ (t \to \infty)$ überlegt man leicht.
Satz 1.3 zeigt dann, dass man auch eine beliebige Norm nehmen darf.

3.3 Matrixnormen

Der $\mathbb{K}^{m\times n}$ stellt auch einen Vektorraum dar und kann mit dem $\mathbb{K}^{m\cdot n}$ identifiziert werden.

Damit definiert jede Norm auf $\mathbb{K}^{m\cdot n}$ auch eine Norm auf $A\in\mathbb{K}^{m\times n}$.

$A^{(t)}\to A \ (t\to\infty)$ meint dann $a_{ij}^{(t)}\to a \ (t\to\infty) \ \forall\, ij \in \{1,\dots,m\}\times\{1,\dots,n\}$

Es zeigt sich, dass folgende Eigenschaften hilfreich sind:

Definition 3.4 Eine Norm $\|.\|$ auf $\mathbb{K}^{n\times n}$ ($m\neq n$ würde zwei V-Normen erfordern) heißt verträglich mit einer Vektornorm $\|.\|$ auf $\mathbb{K}^n$ falls gilt

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

Sie heißt Matrixnorm, wenn sie submultiplikativ ist:

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

⊠

Beispiel 3.5 (Die Frobenius-Norm)

$$\|A\|_{Fr} := \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{1/2}$$

ist eine Matrixnorm, die verträglich mit der euklidischen Norm ist.

Beweis: (als Ü?) C.S. $|\langle x,y\rangle|^2 \le \|x\|^2\|y\|^2$

$$\|Ax\|_2^2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2 \le \sum_{i=1}^{n}\left(\sum_{j=1}^{n}|a_{ij}|^2\cdot \underbrace{\sum_{j=1}^{n}|x_j|^2}_{=\|x\|_2^2}\right) = \|x\|_2^2 \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}_{=\|A\|_{Fr}^2}.$$ ⊠

Definition 3.6 (Zugeordnete Matrixnorm)

Es sei $\|.\|$ eine beliebige (Vektor-)Norm auf $\mathbb{K}^n$. Dann heißt

$$\|A\| := \sup_{x\in\mathbb{K}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|} \overset{(N2)}{=} \sup_{x\in\mathbb{K}^n,\ \|x\|=1} \|Ax\|$$

die $\|.\|$ zugeordnete (oder natürliche) Matrixnorm. $\|.\|$ ist verträglich und submultiplikativ. ⊠

(Übung): Beweis dazu: $y\neq 0$ beliebig $\frac{\|Ay\|}{\|y\|} \le \sup_{x\neq 0}\frac{\|Ax\|}{\|x\|} = \|A\| \Rightarrow \|Ay\| \le \|A\|\,\|y\|$

$$\|AB\| = \sup_{x\neq 0}\frac{\|ABx\|}{\|x\|} \le \sup_{x\neq 0}\frac{\|A\|\,\|Bx\|}{\|x\|} = \|A\| \sup_{x\neq 0}\frac{\|Bx\|}{\|x\|} = \|A\|\,\|B\|.$$

Hilfssatz 3.7

Die zugeordneten Matrixnormen zu $\|\cdot\|_\infty$ und $\|\cdot\|_1$ sind

$$\|A\|_\infty := \max_{i=1,\dots,n} \sum_{j=1}^{n} |a_{ij}| \qquad \text{(Zeilensummennorm)},$$

$$\|A\|_1 := \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}| \qquad \text{(Spaltensummennorm)}.$$

Beweis: siehe Rannacher, Numerik 0 S. 104. Teile hieraus:

a) $\|Ax\|_\infty = \max_{i=1..n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leq \max_{i=1..n} \sum_{j=1}^{n} |a_{ij}| |x_j|$

$$\leq \max_{i=1..n} \sum_{j=1}^{n} |a_{ij}| \underbrace{\left(\max_{k=1..n} |x_k|\right)}_{= \|x\|_\infty}$$

$$\leq \|x\|_\infty \underbrace{\max_{i=1..n} \sum_{j=1}^{n} |a_{ij}|}_{= \|A\|_\infty} = \|A\|_\infty \|x\|_\infty$$

Dies zeigt die Verträglichkeit

b) Somit gilt auch: bzw a) gilt für jedes x d.h. $\frac{\|Ax\|_\infty}{\|x\|_\infty} \leq \|A\|_\infty \Rightarrow \sup \ldots \leq \|A\|_\infty$

$$\sup_{\|x\|_\infty = 1} \|Ax\|_\infty \overset{(a)}{\leq} \sup_{\|x\|_\infty = 1} \|A\|_\infty \underbrace{\|x\|_\infty}_{=1} = \|A\|_\infty$$

c) zu zeigen ist nun dass ebenfalls gilt

$$\sup_{\|x\|_\infty = 1} \|Ax\|_\infty \geq \|A\|_\infty$$

konstruiere Vektor $\|x\|_\infty = 1$ mit $\|Ax\|_\infty = \|A\|_\infty$.

Für $A = 0 \Leftrightarrow \|A\|_\infty = 0$ ist $\sup_{\|x\|_\infty = 1} \|Ax\|_\infty = \|A\|_\infty = 0$

Also sei $A \neq 0$ und $\|A\|_\infty > 0$. Nehme einen Index $m \in \{1, \dots, n\}$ mit

$$\sum_{j=1}^{n} |a_{mj}| = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}| = \|A\|_\infty \qquad m\text{-te Zeile von } A.$$

und definiere den Vektor z mit $z_k = \begin{cases} \frac{|a_{mk}|}{a_{mk}} & a_{mk} \neq 0 \\ 0 & \text{sonst.} \end{cases}$

d.h. $z_k \in \{-1, 0, 1\}$, also $\|z\|_\infty = 1$.

Für $v = Az$ gilt dann für die m-te Komponente:

$$v_m = \sum_{k=1}^{n} a_{mk} z_k = \sum_{k=1}^{n} |a_{mk}| = \|A\|_\infty$$

und somit

$$\|A\|_\infty = v_m \leq \|v\|_\infty = \|Az\|_\infty \underset{\text{da } \|z\|_\infty = 1}{\leq} \sup_{\|y\|_\infty = 1} \|Ay\|_\infty \quad \blacksquare.$$

3.4 Eigenwerte und Eigenvektoren

Sei $A \in \mathbb{K}^{n\times n}$, $\lambda \in \mathbb{K}$ und $e \in \mathbb{K}^n$ so dass

$$Ae = \lambda e,$$

dann heißt λ Eigenwert von A und e zugehöriger Eigenvektor.
(λ, e) heißt auch Eigenpaar.
$(A-\lambda I)e = 0$ ist für nichttriviales $e \neq 0$ nur für

$$p(\lambda) = \det(A - \lambda I) = 0$$

erfüllbar.

Das charakteristische Polynom $p(\lambda)$ hat genau n Nullstellen in $\mathbb{C}$ (Vielfachheiten mitgezählt). (Fundamentalsatz der Algebra)
Zu jedem Eigenwert λ gibt es mindestens einen Eigen~~wert~~vektor.

Mit den Normregeln folgt für ein Eigenpaar (λ, e), $\|e\| = 1$:

$$|\lambda| \underset{\|e\|=1}{\overset{\uparrow}{=}} |\lambda|\,\|e\| \underset{(N2)}{=} \|\lambda e\| = \|Ae\| \leq \|A\|\,\underbrace{\|e\|}_{=1} = \|A\| \qquad (3.3)$$

Also: alle Eigenwerte $\lambda \in \mathbb{C}$ liegen innerhalb eines Kreises um 0 mit Radius $\|A\|$.
Z.B. mit $\|A\|_\infty$ erhält man so eine konkrete Abschätzung für den größten Eigenwert.
Folgerung aus (3.3): $\|A\| \geq \varrho(A) = \max\{|\lambda| : \lambda \text{ ist EW von } A\}$ „Spektralradius"

Definition 3.8 (Spezielle Matrizen)

Zu $A \in \mathbb{K}^{m\times n}$ heißt $A^T \in \mathbb{K}^{n\times m}$ transponierte Matrix und es ist $(A^T)_{ij} = (A)_{ji}$
Zu $A \in \mathbb{K}^{m\times n}$ ist $\bar{A} \in \mathbb{K}^{m\times n}$ gegeben durch $(\bar{A})_{ij} = \overline{(A)_{ij}}$ ← konjugiert komplexe

a) $A \in \mathbb{K}^{n\times n}$ heißt hermitesch falls $A = \bar{A}^T$ d.h. $a_{ij} = \bar{a}_{ji}$.
Manche Autoren schreiben auch A^H für $\bar{A}^T$.
Reelle hermitesche Matrizen heißen symmetrisch (dann ist $A = A^T$)
also A reell dann ist hermitesch ⇔ symmetrisch in $\mathbb{C}$ nicht!

b) Für komplexe Matrizen $A \in \mathbb{C}^{n\times n}$ heißt

$$A\bar{A}^T = \bar{A}^T A \qquad \text{normal}$$

$$A\bar{A}^T = \bar{A}^T A = I,\ \text{also } A^{-1} = \bar{A}^T \qquad \text{unitär}$$

c) Für reelle Matrizen $A \in \mathbb{R}^{n\times n}$ heißt

$$AA^T = A^TA = I,\ \text{also } A^{-1} = A^T \qquad \text{orthogonal}.$$

Definition 3.9 (Skalarprodukt)

Eine Abbildung $(\cdot,\cdot) : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ heißt <u>Skalarprodukt</u>, falls gilt:

(S1) $\quad (x,y) = \overline{(y,x)} \qquad x,y \in \mathbb{K}^n$ (Symmetrie)

(S2) $\quad (\alpha x + \beta y, z) = \alpha (x,z) + \beta (y,z) \qquad x,y,z \in \mathbb{K}^n,\ \alpha,\beta \in \mathbb{K}$ (Linearität)

(S3) $\quad (x,x) > 0 \qquad x \in \mathbb{K}^n - \{0\}$ (Definitheit)

aus (S1) folgt $(x,x) = \overline{(x,x)}$ also $\mathrm{Im}(x,x) = 0$!
und: $(x,\alpha y) = \overline{(\alpha y, x)} = \overline{\alpha}\,\overline{(y,x)} = \overline{\alpha}(x,y)$

Ein Skalarprodukt erzeugt immer eine zugehörige Norm

$$\|x\| := \sqrt{(x,x)}, \qquad x \in \mathbb{K}^n.$$

Es gilt die Cauchy-Schwarzsche Ungleichung

$$|(x,y)| \le \|x\|\,\|y\|.$$

$$\begin{tikzcd}
\text{Vektorraum} & \\
\text{normierter R.} \arrow[u, "\text{Norm}"] & \text{VR mit SP} \arrow[l] \arrow[ul, "{(\cdot,\cdot)}"'] \\
\text{Banach} \arrow[u, "\text{vollst}"] & \text{Hilbert} \arrow[l] \arrow[u, "\text{vollst}"']
\end{tikzcd}$$

Das Euklidische Skalarprodukt (wird im Folg. oft verwendet) lautet

$$(x,y)_2 = \sum_{i=1}^{n} x_i \overline{y_i} = x^T \overline{y}$$

Damit gilt (andere Schreibweise für hermitesch):

$$A = \overline{A}^T \iff (Ax,y)_2 = (x,Ay)_2 \quad \forall x,y \in \mathbb{K}^n.$$

3.5 Die Spektralnorm

Die der Euklidischen Norm zugeordnete Matrixnorm $\|A\|_2$ heißt <u>Spektralnorm</u>.

<u>Hilfssatz 3.10</u> Für die Spektralnorm <u>hermitescher</u> Matrizen gilt

$$\|A\|_2 = \max\{|\lambda| : \lambda \text{ Eigenwert von } A\} =: \rho(A)$$

Für jede Matrix $A \in \mathbb{K}^{m \times n}$ gilt

$$\|A\|_2 = \max\{|\lambda|^{1/2} : \lambda \text{ ist Eigenwert von } \overline{A}^T A\}$$

Beweis: a) Sei A hermitesch. Dann hat A n reelle Eigenwerte und einen vollständigen Satz von orthonormalen Eigenvektoren:

$$\{w^1, \dots, w^n\} \subset \mathbb{K}^n : A w^i = \lambda_i w^i, \quad (w^i, w^j)_2 = \delta_{ij} \quad i,j = 1, \dots, n$$

↑ euklidisches SP!

Jedes $x \in \mathbb{K}^n$ lässt sich in der Basis $\{w^1, \dots, w^n\}$ darstellen:

$$x = \sum_{i=1}^n \alpha_i w^i \quad \text{mit} \quad \alpha_i = (x, w^i)_2 \in \mathbb{K} \text{ (Einsetzen!)}$$

Und es gilt:

$$\|x\|_2^2 = (x,x)_2 = \Big(\sum_{i=1}^n \alpha_i w^i, \sum_{j=1}^n \alpha_j w^j\Big)_2$$

$$= \sum_{i,j=1}^n \alpha_i \bar{\alpha}_j \underbrace{(w^i, w^j)_2}_{=\delta_{ij}} = \sum_{i=1}^n |\alpha_i|^2$$

$$\overline{(a+ib)}\,\overline{(c+id)} = (a-ib)(c-id) = (ac-bd) - i(ad+bc)$$

$$\overline{(a+ib)(c+id)} = \overline{(ac-bd)+i(ad+bc)} = (ac-bd) - i(ad+bc)$$

$$(a+ib)(a-ib) = a^2+b^2$$

$$\|Ax\|_2^2 = (Ax, Ax)_2 = \Big(\sum_i \underbrace{A\alpha_i w^i}_{\alpha_i \lambda_i w^i}, \sum_j \underbrace{A\alpha_j w^j}_{\alpha_j \lambda_j w^j}\Big)_2$$

$$= \sum_{i,j=1}^n \alpha_i \lambda_i \overline{\alpha_j \lambda_j} \underbrace{(w^i, w^j)_2}_{=\delta_{ij}} = \sum_{i,j=1}^n \lambda_i^2 |\alpha_i|^2$$

↑ Betrag, da $\alpha \in \mathbb{K}$

Also
$$\|A\|_2^2 = \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\|Ax\|_2^2}{\|x\|_2^2} = \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\sum_{i}^n \lambda_i^2 |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2} \le \max_{1 \le i \le n} |\lambda_i|^2$$

Andererseits zeigt (3.3) $\|A\| \ge |\lambda|$ für jede Norm und jeden EW, also muss $\|A\|_2 = \max_{1 \le i \le n} |\lambda_i|$ sein.

b) Sei $A \in \mathbb{K}^n$ eine beliebige Matrix. Dann gilt

Rechenregeln für konjug. komplex

$$\|Ax\|_2^2 = (Ax, Ax)_2 = (Ax)^T (\overline{Ax}) = x^T A^T (\overline{Ax}) = x^T (\overline{\bar{A}^T A x}) = (x, \bar{A}^T A x)_2$$

$\bar{A}^T A$ ist hermitesch (Ü) und habe EW λ_i sowie EV w^i orthonormal wie oben

also
$$\|Ax\|_2^2 = \Big(\sum_{i=1}^n \alpha_i w^i, \sum_{j=1}^n \bar{A}^T A \alpha_j w^j\Big)_2 = \sum_{i,j=1}^n \bar{\alpha}_i \underset{\text{reell}}{\lambda_j} \bar{\alpha}_j \underbrace{(w^i, w^j)_2}_{=\delta_{ij}} = \sum_{i=1}^n \lambda_i |\alpha_i|^2$$

und damit
$$\|A\|_2^2 = \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\|Ax\|_2^2}{\|x\|_2^2} = \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\sum_{i=1}^n \lambda_i |\alpha_i|^2}{\sum_i |\alpha_i|^2} \le \max_{1 \le i \le n} |\lambda_i|$$

∎

3.6 Positiv definite Matrizen

Wichtige Klasse von Matrizen mit vorteilhaften Eigenschaften.

(in $\mathbb{K}$!)

Definition 3.11 Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt positiv definit wenn

a) $(Ax,x)_2 \in \mathbb{R} \quad \forall x \in \mathbb{K}^n \qquad (\mathbb{K}=\mathbb{C})$

b) $(Ax,x)_2 > 0 \quad \forall x \in \mathbb{K}^n \setminus \{0\}$ (eigentliche Bedingung)

Ziel: Charakterisierung positiv definiter Matrizen.

Im Fall $\mathbb{K}=\mathbb{C}$ bedeutet die Bedingung a) eine Einschränkung an die Matrix A.

Eigenschaft 3.12. Für $A \in \mathbb{C}^{n\times n}$ gilt A hermitesch genau dann wenn $(Ax,x)_2 \in \mathbb{R} \; \forall x \in \mathbb{C}^n$.

Lemma 3.13 Eine hermitesche Matrix A ist genau dann positiv definit, wenn alle ihre (reellen) Eigenwerte positiv sind. Alle Hauptdiagonalelemente sind (reell und) positiv.

Beweis:

"⇐" A sei hermitesch mit lauter positiven Eigenwerten, z.z. A pos. def.

$x \in \mathbb{C}^n$ hat Darstellung $x = \sum \alpha_i w^i$ $\quad w^i$: Eigenvektor zum EW λ_i (reell)

$$(Ax,x)_2 = \sum_{i,j=1}^{n} \lambda_i \alpha_i \bar{\alpha}_j \underbrace{(w^i,w^j)_2}_{=\delta_{ij}} = \sum_{i=1}^{n} \lambda_i |\alpha_i|^2 > 0$$

"⇒" - A sei hermitesch und positiv definit. A hat also Eigenpaare (λ_i, w^i) mit $\lambda_i \in \mathbb{R}$.

Zu zeigen ist λ_i positiv

$$0 \underset{\text{Vor.}}{<} (Aw^i,w^i)_2 = (\lambda_i w^i, w^i)_2 = \lambda_i \underbrace{(w^i,w^i)_2}_{=1} = \lambda_i$$

- setze $e^i \in \mathbb{R}^n$, $e^i_j = \delta_{ij}$ kartesische Einheitsvektoren.

$$0 < (Ae^i, e^i)_2 = a_{ii} \in \mathbb{R} \text{ (vgl. 3.11 (a))}$$

Die Aussage gilt natürlich auch für reelle symmetrische und positiv definite Matrizen, da auch diese hermitesch sind.

Speziell für reelle Matrizen gilt:

<u>Lemma 3.14</u> Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit. Dann liegt das betragsmäßig größte Element auf der Hauptdiagonalen.

Beweis: e^i sei wieder der i-te kartesische Einheitsvektor und a_{ij}, $j \neq i$, das betragsmäßig größte Element. Dann zeigt man den Widerspruch.

$$0 < (A(e^i - \text{sign}(a_{ij})e^j), e^i - \text{sign}(a_{ij})e^j)_2$$
$$= (Ae^i, e^i)_2 - 2\,\text{sign}(a_{ij}) \underbrace{(Ae^i, e^j)_2}_{a_{ij}} + \text{sign}(a_{ij})^2 (Ae^j, e^j)$$
$$\overset{A=A^T}{=} a_{ii} - 2|a_{ij}| + a_{jj} \leq 0 \quad ↯ \qquad \text{bzw } |a_{ij}| < \frac{a_{ii}+a_{jj}}{2}$$

da $|a_{ij}| \geq a_{ii}$, $|a_{ij}| \geq a_{jj}$ $(a_{ii}, a_{jj} > 0\,!)$ ∎

- Manche Autoren (auch Rannacher) verlangen, dass reelle positiv definite Matrizen symmetrisch sind. Wir fordern das extra

Ü $A \in \mathbb{R}^{n \times n}$ A ist positiv definit genau dann wenn $A_S = \frac{1}{2}(A + A^T)$ positiv definit ist.

A pos. definit, dann sind alle Hauptuntermatrizen (definieren) positiv definit

Def. 3.15 (Kondition einer Abb.)

Sei $f: X \to Y$, $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$

Dann heißt

$$K_{abs}(x) = \sup \left\{ \frac{\| f(x+\Delta x) - f(x) \|_Y}{\| \Delta x \|_X} : \Delta x \neq 0, x + \Delta x \in X \right\}$$

absolute Kondition von f im Punkt x und

$$K(x) = \sup \left\{ \frac{\| f(x+\Delta x) - f(x) \|}{\| f(x) \|_Y} \frac{\| x \|_X}{\| \Delta x \|_X} : \Delta x \neq 0, x + \Delta x \in X \right\}$$

relative Kondition von f in x.

Bem: (1) $K_{abs}(x)$ ist die Lipschitz-Konstante (in x).

(2) Für $f: \mathbb{R} \to \mathbb{R}$ entspricht das dem Verstärkungsfaktor

$$k(x) = \frac{df}{dx}(x) \frac{x}{f(x)}$$

wobei $\frac{df}{dx}(x)$ durch den Differenzenquotienten genähert wurde (dann kann man auf die Differenzierbarkeit verzichten).

3.7 Störungstheorie

Sei $A \in \mathbb{K}^{n\times n}$ regulär und $x, b \in \mathbb{K}^n$. ← Def. 3.15

$F(A,b) = A^{-1}b$ ist der „Lösungsoperator" zur Gleichung $G(x) = Ax - b = 0$.

Betrachte die relative Kondition von F, wobei wir zunächst nur Änderungen in b zulassen wollen:

$$\frac{\|F(A, b+\delta b) - F(A,b)\|}{\|\delta b\|} \frac{\|b\|}{\|F(A,b)\|} = \frac{\|A^{-1}(b+\delta b) - A^{-1}b\|}{\|\delta b\|} \frac{\|\overbrace{AA^{-1}}^{I}b\|}{\|A^{-1}b\|}$$

$$\leq \frac{\|A^{-1}\| \|\delta b\|}{\|\delta b\|} \frac{\|A\| \|A^{-1}b\|}{\|A^{-1}b\|} = \|A^{-1}\| \|A\|.$$

↑ $\|\cdot\|$ sei verträglich

Definition 3.16 Die Zahl

$$\mathrm{cond}(A) = \|A\| \|A^{-1}\|$$

für irgendeine verträgliche Matrixnorm heißt Konditionszahl von A.

submultiplikativ ist Teil der Def. von Matrixnorm.

Nun wollen wir auch Änderungen in A selbst zulassen. Sei $A \in \mathbb{K}^{n\times n}$ regulär. Frage: Wann ist $A + \delta A$ regulär?

Hilfssatz 3.17 $B \in \mathbb{K}^{n\times n}$ habe Norm $\|B\| < 1$. Dann ist $I+B$ regulär und es gilt

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

Beweis: i) Für alle $x \in \mathbb{K}^n$ gilt

$$\|x\| = \|x + Bx - Bx\| \leq \|(I+B)x\| + \|Bx\|$$

$$\Leftrightarrow \|(I+B)x\| \geq \|x\| - \|Bx\| \underset{\|Bx\| \leq \|B\|\|x\|}{\geq} \|x\| - \|B\|\|x\| = \underbrace{(1-\|B\|)}_{>0 \text{ da } \|B\|<1}\|x\|$$

Damit gilt für $x \neq 0$, dass $(I+B)x \neq 0$ also $I+B$ regulär.

ii) : Zugeord. Matrixnorm; erste Klammer auflösen; $\|I\| = \sup_{x\neq 0} \frac{\|Ix\|}{\|x\|} = \sup_{x \neq 0} \frac{\|x\|}{\|x\|} = 1$

$$1 = \|I\| = \|(I+B)(I+B)^{-1}\| = \|(I+B)^{-1} + B(I+B)^{-1}\|$$

wie oben: $\|V+W-W\| \leq \|V+W\| + \|W\|$ $\Leftrightarrow \|V+W\| \geq \|V\| - \|W\|$

$$\geq \|(I+B)^{-1}\| - \|B\|\|(I+B)^{-1}\| = \|(I+B)^{-1}\|(1-\|B\|) > 0$$

da $\|B\| < 1$

Daraus folgt die Behauptung. ∎

Damit gilt das Folgende

Satz 3.18 (Störungssatz) $A \in \mathbb{K}^{n\times n}$ sei regulär und $\|\delta A\| < \frac{1}{\|A^{-1}\|}$.
Dann ist $\tilde{A} = A + \delta A$ ebenfalls regulär und es gilt für den relativen Fehler des gestörten Systems $(A+\delta A)(x+\delta x) = b + \delta b$:

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

Beweis:

i) $A + \delta A = A(I + \underbrace{A^{-1}\delta A}_{„B"})$ und $\|A^{-1}\delta A\| \leq \|A^{-1}\|\|\delta A\| \overset{\text{Vor}}{<} \frac{\|A^{-1}\|}{\|A^{-1}\|} = 1$

nach Hilfssatz 3.17 ist $A + \delta A$ regulär.

ii) $(A+\delta A)(x+\delta x) = b + \delta b$

$\Leftrightarrow Ax + \delta A x + (A+\delta A)\delta x = b + \delta b$

x Lösung von $Ax = b$

$\Leftrightarrow (A+\delta A)\delta x = \delta b - \delta A x$

$\Leftrightarrow \delta x = (A+\delta A)^{-1}(\delta b - \delta A x)$

$$\|\delta x\| \leq \|(A+\delta A)^{-1}\| \left\{ \|\delta b\| + \|\delta A\| \|x\| \right\}$$

Normregeln

$$= \| [A(I+A^{-1}\delta A)]^{-1} \| \left\{ \|\delta b\| + \|\delta A\| \|x\| \right\}$$

$$= \|(I+A^{-1}\delta A)^{-1} A^{-1}\| \left\{ \|\delta b\| + \|\delta A\| \|x\| \right\}$$

$$\leq \underbrace{\|(I+A^{-1}\delta A)^{-1}\|}_{\text{Hilfssatz } 3.18} \|A^{-1}\| \left\{ \|\delta b\| + \|\delta A\| \|x\| \right\}$$

$$\leq \frac{\|A^{-1}\|}{1-\|A^{-1}\delta A\|} \left\{ \|\delta b\| + \|\delta A\| \|x\| \right\}$$

$\|A^{-1}\delta A\| \leq \|A^{-1}\| \|\delta A\| < 1$ nach Vor.!

$$\leq \frac{\|A^{-1}\|}{1-\|A^{-1}\| \|\delta A\|} \left\{ \|\delta b\| + \|\delta A\| \|x\| \right\}$$

Mehr abziehen macht $1-x$ kleiner, also den Bruch größer!

ausgeklammert

$$= \frac{\|A^{-1}\| \|A\| \|x\|}{1-\|A^{-1}\| \|\delta A\| \underbrace{\|A\| \|A\|^{-1}}_{=1}} \left\{ \frac{\|\delta b\|}{\|A\| \|x\|} + \frac{\|\delta A\|}{\|A\|} \right\}$$

hinzugefügt

$$\leq \|x\| \frac{\mathrm{cond}(A)}{1-\mathrm{cond}(A)\frac{\|\delta A\|}{\|A\|}} \left\{ \frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|} \right\} \quad \blacksquare$$

$\|b\| = \|Ax\| \leq \|A\| \|x\|$

Nenner verkleinern macht Bruch größer

Beispiel 3.19

Es sei $\frac{\|\delta A\|}{\|A\|} \approx 10^{-k}$ und $\frac{\|\delta b\|}{\|b\|} \approx 10^{-k}$ sowie $\text{cond}(A) \approx 10^{s}$ $(s,k>0)$

Weiter nehmen wir an, dass $10^{s} \cdot 10^{-k} \ll 1$ also etwa $s-k \leq -3$.

Dann gilt

$$\frac{\|\delta x\|}{\|x\|} \leq \underbrace{\frac{10^{s}}{1-\underbrace{10^{s}10^{-k}}_{\ll 1}}}_{\text{Nenner} \approx 1} \cdot 2 \cdot 10^{-k} \approx 10^{s-k} = 10^{-(k-s)}$$

$\Rightarrow$ man verliert s Stellen an Genauigkeit!

Eingabefehler ist in der k-ten Nachkommastelle, Fehler im Ergebnis ist in der $k-s$ ten Stelle. ☒

Man kann zeigen, dass diese Abschätzung im wesentlichen „scharf" ist [Ra]. S. 117.

Beispiel 3.20 (Kondition + Determinante) Evtl als Übung.

(a). Betrachte die 2×2 Matrix

$$A = \begin{bmatrix} -1 & 1 \\ 1-\varepsilon & -1 \end{bmatrix}, \quad A^{-1} = \frac{1}{\varepsilon}\begin{bmatrix} -1 & -1 \\ -(1-\varepsilon) & -1 \end{bmatrix}, \quad \varepsilon = \det(A), \ |\varepsilon| < 1$$

Es gilt

$$\|A\|_\infty = \max(2, 2-\varepsilon), \ \|A^{-1}\|_\infty = \frac{1}{\varepsilon}\max(2, 2-\varepsilon), \ \text{cond}_\infty(A) = \frac{(2+\varepsilon)^2}{\varepsilon} = \begin{cases} \frac{4}{\varepsilon} & \varepsilon > 0 \\ \frac{(2+|\varepsilon|)^2}{\varepsilon} & \text{sonst} \end{cases}$$

Also $\text{cond}_\infty(A) = O\left(\frac{1}{\det(A)}\right)$.

Dies ist aber nicht immer so:

(b)

$$B = \begin{bmatrix} 10^{-10} & 0 \\ 0 & 10^{-10} \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 10^{10} & 0 \\ 0 & 10^{10} \end{bmatrix}$$

$\text{cond}_\infty(B) = \|B\|_\infty \|B^{-1}\|_\infty = 10^{-10} \cdot 10^{10} = 1$ aber

$\det(B) = 10^{-20}$!

(c) $\text{cond}(\varepsilon A) = \text{cond}(A) \ \forall \varepsilon \neq 0$, aber $\det(\varepsilon A) = \varepsilon^{n} \det(A)$ ☒

Beispiel 3.21 (oder Übung)

$$A = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} d_1^{-1} & 0 \\ 0 & d_2^{-1} \end{bmatrix}, \qquad d_1, d_2 \neq 0$$

$$\operatorname{cond}_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = \frac{\max(|d_1|, |d_2|)}{\min(|d_1|, |d_2|)}$$

d.h. $\operatorname{cond}\left(\begin{bmatrix} 10^{10} & 0 \\ 0 & 1 \end{bmatrix}\right) = 10^{10}$!

Andererseits sind alle Gleichungen in $Ax = b$ unabhängig und $x_i = b_i / d_i$ ist gut konditioniert.

Problem liegt hier in der Verwendung von Normen, die Abschätzung für einzelne Komponenten ist relativ schlecht.

Man lernt: schlechte Kondition muss nicht unbedingt Probleme beim Lösen des LGS bewirken, es ist jedoch „wahrscheinlich".