§ 1 Nichtlineare Gleichungen

Problem: Finde Lösungen der Gleichung

(*) $\quad F(x) = b$

für eine Abbildung $F: \mathbb{R}^n \to \mathbb{R}^m$ und $b \in \mathbb{R}^m$.

Ist F linear, d.h. $F(\alpha x + \beta y) = \alpha F(x) + \beta F(y)$ $\forall \alpha, \beta \in \mathbb{R}$, $x, y \in \mathbb{R}^n$, so ist (*) äquivalent zu einem linearen Gleichungssystem

$$A x = b \qquad (A \in M(m \times n, \mathbb{R}),\ b \in \mathbb{R}^m).$$

$\rightsquigarrow$ Lineare Algebra (Gauß-Elimination, ...)

Nichtlineare Gleichungen lassen sich i.A. nicht exakt lösen

$\rightsquigarrow$ Numerische Methoden zur approximativen Lösung

A Fixpunktsatz von Banach

Beispiel 1.1 Gesucht: Lösung der Gleichung

$$p(x) = \frac{1}{3}x^3 - 2x + 1 \Longleftrightarrow f(x) := \frac{1}{6}x^3 + \frac{1}{2} = x$$

oder äquivalent ein Fixpunkt der Funktion f. Wegen

$$p(0) = 1 > 0 > -\frac{2}{3} = p(1)$$

gibt es nach dem Zwischensatz (Satz 10.2.4) eine Lösung $x_* \in [0,1]$. Da f monoton wächst, ist

$$f([0,1]) = [f(0), f(1)] = [\tfrac{1}{2}, \tfrac{2}{3}] \subset [0,1].$$

Gesucht: Fixpunkt der Abbildung $f: [0,1] \to [0,1]$.

Nach dem Mittelwertsatz (11.2.3) gibt es zu $x, y \in [0,1]$ eine Stelle ξ zwischen x und y mit

$$|f(x) - f(y)| = |f'(\xi)(x-y)| = \tfrac{1}{2}\xi^2 |x-y| \leq \tfrac{1}{2}|x-y|,$$

d.h. f ist (strikt!) kontrahierend.

Sei $x_0 \in [0,1]$ beliebig. Für die rekursiv durch

$$x_{k+1} = f(x_k) \quad (k = 0,1,2,\ldots)$$

definierte Folge in $[0,1]$ gilt

$$|x_{k+1} - x_k| = |f(x_k) - f(x_{k-1})| \leq \tfrac{1}{2}|x_k - x_{k-1}| \leq \ldots \leq \left(\tfrac{1}{2}\right)^k |x_1 - x_0|$$

$\Rightarrow$ (x_k) ist eine Cauchy-Folge (4.1.4), denn $\forall q \geq p$ gilt

$$|x_q - x_p| \leq \sum_{k=p}^{q-1} |x_{k+1} - x_k| \leq \left(\sum_{k=p}^{q-1} \left(\tfrac{1}{2}\right)^k\right) |x_1 - x_0| \xrightarrow{(p\to\infty)} 0$$

nach dem Cauchy-Kriterium (4.2.3) für die geometrische Reihe $\sum_{k=0}^{\infty} \left(\tfrac{1}{2}\right)^k$.

Satz 4.1.7 $\Rightarrow$ $x_* = \lim_{k\to\infty} x_k \in [0,1]$ existiert

f stetig $\Rightarrow$ $f(x_*) = f\left(\lim_{k\to\infty} x_k\right) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} x_{k+1} = x_*$

Die Iterationsfolge $(x_k)_{k\geq 0}$ approximiert den gesuchten Fixpunkt x_*.

Diese Methode funktioniert viel allgemeiner: Wir ersetzen

$(\mathbb{R}, |\cdot|) \rightsquigarrow (\mathbb{R}^n, \|\cdot\|) \rightsquigarrow$ (Vektorraum V, Norm $\|\cdot\|$),

$[0,1] \rightsquigarrow D \subset \mathbb{R}^n$ abgeschlossen

$f: [0,1] \to [0,1] \rightsquigarrow f: D \to D$ Kontraktion.

Def. 1.2 Sei V ein Vektorraum über $\mathbb{R}$ (oder $\mathbb{C}$).

(a) Eine Abbildung $\|\cdot\|: V \to \mathbb{R}$ heißt Norm, falls

$\forall\, x, y \in V$ und $\forall\, \alpha \in \mathbb{R}$ (oder $\mathbb{C}$) gilt

(i) $\|x\| \geq 0$, $\|x\| = 0 \Leftrightarrow x = 0$,

(ii) $\|\alpha x\| = |\alpha|\, \|x\|$

(iii) $\|x+y\| \leq \|x\| + \|y\|$

(b) Eine Folge $(x_k)_{k\geq 0}$ in V (vgl. § 4.1 für $(V, \|\cdot\|) = (\mathbb{R}, |\cdot|)$)

- konvergent gegen ein $x \in V$:$\Leftrightarrow$ $\forall\, \varepsilon > 0\ \exists\, n_0 \in \mathbb{N}$ mit $\|x_n - x\| < \varepsilon\ \forall\, n \geq n_0$
- Cauchy-Folge :$\Leftrightarrow$ $\forall\, \varepsilon > 0\ \exists\, n_0 \in \mathbb{N}$ mit $\|x_p - x_q\| < \varepsilon\ \forall\, p, q \geq n_0$

(c) Ein normierter Raum $(V, \|\cdot\|)$ heißt vollständig (oder Banachraum)

:$\Leftrightarrow$ Jede Cauchy-Folge (x_k) in V konvergiert.

Beispiele 1.3 (a) $V = \mathbb{R}^n$ mit den Normen

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

$\|(x_i)\|_1 := \sum_{i=1}^{n} |x_i|$ Summennorm

$\|(x_i)\|_\infty := \max_{i=1,\dots,n} |x_i|$ Maximumnorm

(b) $V = C[a,b] = \{f;\ f:[a,b] \to \mathbb{R} \text{ stetig}\}$ mit der Norm

$$\|f\|_\infty := \sup_{x\in[a,b]} |f(x)|$$

(Satz 10.2.1 $\Rightarrow$ $\|\cdot\|_\infty$ vollständige Norm)

(c) Sei $\alpha \in \{1, 2, \infty\}$ und sei $A \in M(n, \mathbb{R})$. Da

$$\{x \in \mathbb{R}^n;\ \|x\|_\alpha \leq 1\} \to \mathbb{R}^n,\quad x \mapsto \|Ax\|_\alpha$$

stetig auf einer kompakten Menge ist, ist (Satz 10.1.5)

$$\|A\|_\alpha := \sup_{\|x\|_\alpha \le 1} \|Ax\|_\alpha < \infty .$$

Durch

$$M(n,\mathbb{R}) \longrightarrow \mathbb{R}, \quad A \mapsto \|A\|_\alpha$$

wird die zu $(\mathbb{R}^n, \|\cdot\|_\alpha)$ gehörige <u>Matrixnorm</u> auf $M(n,\mathbb{R})$ definiert.
Für $x \in \mathbb{R}^n$ und $A, B \in M(n,\mathbb{R})$ gilt:

(i) $\|Ax\|_\alpha \le \|A\|_\alpha \; \|x\|_\alpha$,

(ii) $\|AB\|_\alpha \le \|A\|_\alpha \; \|B\|_\alpha$.

Alle Normen aus Beispiel 1.3 sind vollständig.

Sei V im Folgenden ein $\mathbb{R}$-Vektorraum und $\|\cdot\| : V \to \mathbb{R}$ eine vollständige Norm auf V. Sei $D \subset V$ und $f : D \longrightarrow D$ eine Abbildung.

<u>Definition 1.4</u>

(a) Ein Punkt $x \in D$ heißt <u>Fixpunkt</u> von f, falls $f(x) = x$ ist.

(b) Ein <u>Iterationsverfahren</u> erzeugt eine Folge $(x_k)_{k\ge 0}$ in D durch die Vorschrift

$$(*) \qquad x_{k+1} = \Phi(x_k) \qquad (k = 0,1,2,\dots)$$

zu gegebenem <u>Startwert</u> $x_0 \in D$ mit einer <u>Verfahrensfunktion</u>

$$\Phi : D \longrightarrow D .$$

(c) Ein Iterationsverfahren heißt <u>konvergent</u>, falls

$$x_* := \lim_{k\to\infty} x_k \in V \quad \text{existiert},$$

<u>konvergent von der Ordnung $p \ge 1$</u>, falls zusätzlich ein $c > 0$ existiert mit

$$\|x_{k+1} - x_*\| \le c \, \|x_k - x_*\|^p \quad \forall k \ge 0 .$$

Für $p = 1$ verlangt man $c < 1$.

(d) Eine Abbildung $f : D \to D$ heißt Kontraktion, falls ein $c \in [0,1)$ existiert mit

$$\| f(x) - f(y) \| \leq c \| x - y \| \quad \forall x, y \in D.$$

Wie im $\mathbb{R}^n$ nennt man eine Menge $D \subset V$ abgeschlossen, falls für jede Folge (x_k) in D, die gegen ein $x \in V$ konvergiert, auch $x \in D$ gehört.

Satz 1.5 (Banachscher Fixpunktsatz)

Seien $(V, \|\cdot\|)$ ein vollständiger normierter Raum, $D \subset V$ abgeschlossen, $f : D \to D$ eine Kontraktion mit Konstante $c \in [0,1)$ wie oben. Dann gilt:

(a) f besitzt genau einen Fixpunkt $x_* \in D$

(b) Sei $x_0 \in D$ beliebig. Dann konvergiert die Fixpunktiteration

$$x_{k+1} = f(x_k) \quad (k = 0, 1, 2, \dots)$$

gegen den Fixpunkt x_* von f.

(c) Es gelten die Fehlerabschätzungen

(i) $\| x_k - x_* \| \leq \frac{c^k}{1-c} \| x_1 - x_0 \|$ (a priori)

(ii) $\| x_{k+1} - x_* \| \leq \frac{c}{1-c} \| x_{k+1} - x_k \|$ (a posteriori).

Beweis: Existenz: $\forall k \geq 0$ gilt

$$\| x_{k+1} - x_k \| = \| f(x_k) - f(x_{k-1}) \| \leq c \| x_k - x_{k-1} \| \leq \dots \leq c^k \| x_1 - x_0 \|$$

$\Rightarrow$ (x_k) ist eine Cauchy-Folge in $(V, \|\cdot\|)$, denn $\forall q \geq p$ gilt

$$(*) \quad \| x_q - x_p \| \leq \sum_{k=p}^{q-1} \| x_{k+1} - x_k \| \leq \left(\sum_{k=p}^{q-1} c^k \right) \| x_1 - x_0 \| \leq c^p \underbrace{\left(\sum_{k=0}^{\infty} c^k \right)}_{= \frac{1}{1-c} \text{ (Satz 4.2.1)}} \| x_1 - x_0 \| \xrightarrow{(p \to \infty)} 0$$

$(V, \|\cdot\|)$ vollständig $\Rightarrow$ $x_* = \lim_{k\to\infty} x_k \in V$ existiert

$D \subset V$ abgeschlossen $\Rightarrow$ $x_* \in D$

f stetig (als Kontraktion) $\Rightarrow$ $f(x_*) = f(\lim_{k\to\infty} x_k) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} x_{k+1} = x_*$

Aus (*) folgt durch Grenzübergang $q \to \infty$:

$$\|x_* - x_p\| = \lim_{q\to\infty} \|x_q - x_p\| \le \frac{c^p}{1-c} \|x_1 - x_0\| \quad \text{(a priori)}$$

Ersetzt man x_0 durch (x_k), so folgt aus (i)

$$\|x_{k+p} - x_*\| \le \frac{c^p}{1-c} \|x_{k+1} - x_k\| \overset{p=1}{=} \frac{c}{1-c} \|x_{k+1} - x_k\| \quad \text{(a posteriori)}$$

<u>Eindeutigkeit:</u> Ist auch $f(\tilde{x}) = \tilde{x}$, so folgt

$$\|\tilde{x} - x_*\| = \|f(\tilde{x}) - f(x_*)\| \le c\,\|\tilde{x} - x_*\| \overset{c<1}{\Rightarrow} \|\tilde{x} - x_*\| = 0 \Rightarrow \tilde{x} = x_*. \quad \boxtimes$$

Wie erkennt man, ob eine Abbildung f eine Kontraktion ist?

<u>Satz 1.6</u> (Mittelwertsatz = 11.2.3)

Ist $f \in C^1[a,b]$, so $\exists\, \xi \in]a,b[$ mit $f(b) - f(a) = f'(\xi)(b-a)$.

Ist $c := \sup_{t\in[a,b]} |f'(t)| < 1$ und $f([a,b]) \subset [a,b]$, so ist $f : [a,b] \to [a,b]$ eine Kontraktion.

Ersetzt man $f'(t)$ durch die Jacobi-Matrix (17.2.4), so erhält man eine mehrdimensionale Version vom letzten Satz.

<u>Satz 1.7</u> Sei $K \subset \mathbb{R}^n$ und sei $f : \mathcal{U} \to \mathbb{R}^m$ eine C^1-Funktion auf einer offenen Menge $\mathcal{U} \subset \mathbb{R}^n$ mit $K \subset \mathcal{U}$. Für $\alpha \in \{1, 2, \infty\}$ gilt:

(a) Ist $K \subset \mathbb{R}^n$ kompakt, so ist $\sup_{x\in K} \|J_f(x)\|_\alpha < \infty$.

(b) Ist $K \subset \mathbb{R}^n$ konvex und ist $c := \sup_{x \in K} \|J_f(x)\|_\alpha < \infty$, so gilt

$$\|f(x) - f(y)\|_\alpha \leq c \, \|x - y\|_\alpha \quad \forall x, y \in K$$

Hierbei heißt $K \subset \mathbb{R}^n$ <u>konvex</u>, falls $\forall x, y \in K$ und $\forall t \in [0,1]$ gilt

$$t x + (1-t) y \in K$$

<u>Beweis</u>. (a) Ist $f: \mathcal{U} \to \mathbb{R}^m$ C^1, so ist

$$\mathcal{U} \to \mathbb{R}, \quad x \mapsto \|J_f(x)\|_\alpha$$

stetig und daher $\sup_{x \in K} \|J_f(x)\|_\alpha < \infty$ für jede kompakte Menge $K \subset \mathcal{U}$

(b) Forster, Analysis II, Satz 6.5 + Korollar zu Satz 6.5. ☒

<u>Beispiel 1.8</u> (a) Sei $f: [0,1] \to [0,1]$, $f(x) = \frac{1}{6} x^3 + \frac{1}{2}$ die Kontraktion aus Beispiel 1.1. Für $x_0 \in [0,1]$ konvergiert die Fixpunktiteration

$$x_{k+1} = f(x_k) \quad (k = 0, 1, 2, \dots)$$

gegen den eindeutigen Fixpunkt x_* von f in $[0,1]$. Etwa:

$$\begin{aligned}
x_0 &= 0 \\
x_1 &= f(x_0) = \tfrac{1}{2} \\
x_2 &= f(x_1) = \tfrac{25}{48} \\
x_3 &= f(x_2) \cong 0.523548
\end{aligned}$$

Wie viele Iterationen braucht man für eine Genauigkeit von $\varepsilon = 10^{-3}$, d.h.

$$|x_k - x_*| < 10^{-3} \ ?$$

Die a priori Fehlerabschätzung mit $c = \frac{1}{2}$ (siehe Bsp. 1.1) liefert

$$|x_k - x_*| \leq \frac{\left(\frac{1}{2}\right)^k}{1 - \frac{1}{2}} |x_1 - x_0| = \left(\frac{1}{2}\right)^k < 10^{-3}$$

$$\Longleftrightarrow \quad k > 9 \quad (2^9 = 512 < 1000 < 1024 = 2^{10}).$$

Also ist $|x_{10} - x_*| < 10^{-3}$.

(b) Sei $c \geq 0$. Die Abbildung $F: \mathbb{R}^2 \to \mathbb{R}^2$

$$F(\theta) = c \begin{pmatrix} \cos\theta_1 - \sin\theta_2 \\ \cos\theta_1 - 2\sin\theta_2 \end{pmatrix}$$

ist C^∞ mit

$$J_F(\theta) = -c \left(\begin{array}{c|c} \sin\theta_1 & \cos\theta_2 \\ \hline \sin\theta_1 & 2\cos\theta_2 \end{array}\right) \in M(2,\mathbb{R})$$

Für $x = (x_1, x_2) \in \mathbb{R}^2$ mit $\|x\|_\infty \leq 1$ ist

$$\|J_F(\theta)x\|_\infty = c \left\| \begin{pmatrix} (\sin\theta_1)x_1 + (\cos\theta_2)x_2 \\ (\sin\theta_1)x_1 + (2\cos\theta_2)x_2 \end{pmatrix} \right\|_\infty$$

$$\leq c \max\left(|\sin\theta_1| + |\cos\theta_2|, |\sin\theta_1| + 2|\cos\theta_2|\right) \leq 3c.$$

$\Rightarrow \quad \|J_F(\theta)\|_\infty \leq 3c$

$\Rightarrow$ Für $c \in [0, \frac{1}{3})$ ist $F: \mathbb{R}^2 \to \mathbb{R}^2$ nach Satz 1.7 eine Kontraktion und nach dem Banachschen Fixpunktsatz (Satz 1.5) konvergiert für $\theta_0 \in \mathbb{R}^2$ die Iterationsfolge

$$\theta_{k+1} := F(\theta_k) \quad (k \geq 0)$$

gegen den eindeutigen Fixpunkt θ_* von F in $\mathbb{R}^2$ mit

$$\|\theta_k - \theta_*\|_\infty \leq \frac{(3c)^k}{1-(3c)} \|\theta_1 - \theta_0\|_\infty \quad \forall k \geq 0.$$

Satz 1.9 Die Fixpunktiteration aus Satz 1.5 hat die Konvergenzordnung $p=1$.

Beweis: Sei $x_* \in D$ der eindeutige Fixpunkt von $f: D \to D$

$\Rightarrow \|x_{k+1} - x_*\| = \|f(x_k) - f(x_*)\| \leq c\|x_k - x_*\| \; \forall k \geq 0$ mit $c \in [0,1)$

Nach Definition 1.4 ist die Konvergenzordnung $p = 1$. ∎

B Newton Verfahren

Beispiel 1.10 (a) Man weiß, dass jedes reelle Polynom

$$p(x) = \sum_{i=0}^{n} a_i x^i$$

mit ungeradem Grad $\deg(p)$ mindestens eine Nullstelle in $\mathbb{R}$ hat (ZWS = Satz 10.2.4)

(b) Ist $p(x) = x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 - \left(\frac{b^2}{4} - c\right)$, so gilt

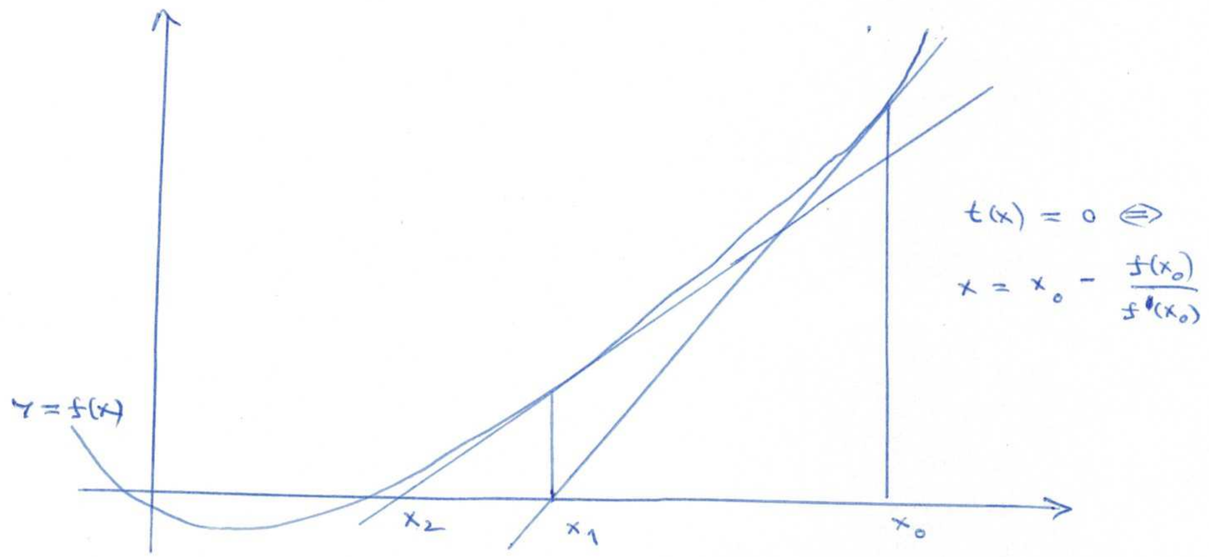
$\frac{b^2}{4} < c \Rightarrow p$ hat kein reelle Nullstelle

$\frac{b^2}{4} \geq c \Rightarrow p$ hat die N.stellen $x_\pm = -\frac{b}{2} \pm \sqrt{\frac{b^2}{4} - c}$

Für $\deg p \in \{3,4\}$ gibt es explizite Formeln für die N.stellen, für $\deg(p) \geq 5$ nicht!

Das Newtonverfahren erlaubt es, unter geeigneten Bedingungen N.stellen bis auf vorgegebene Genauigkeit zu berechnen.

Idee Ist x_0 ein erster Näherungswert für eine Lösung der Gleichung $f(x) = 0$, so erhält man eine bessere Näherung, indem man den Graphen von f ersetzt durch die Tangente $t(x) = f(x_0) + f'(x_0)(x - x_0)$



an f im Punkt $(x_0, f(x_0))$ und x_0 durch den Schnittpunkt

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

von t mit der x-Achse.

Beachte: Ist $f \in C^1[a,b]$ mit $0 \notin f'([a,b])$ und konvergiert die Iterationsfolge

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}, \quad x_0 \in [a,b]$$

gegen ξ, so ist $\xi \in [a,b]$ und

$$\xi = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty}\left(x_k - \frac{f(x_k)}{f'(x_k)}\right) = \xi - \frac{f(\xi)}{f'(\xi)} \Rightarrow f(\xi) = 0$$

Aber: Das Verfahren braucht nicht zu konvergieren!

Positive Resultate:

Satz 1.11 Sei $I = [a-r, a+r]$, $f \in C^2(I)$ mit $0 \notin f'(I)$ und $c \in [0,1)$ mit

(i) $$\left|\frac{f(x)\, f''(x)}{f'(x)^2}\right| \le c \quad \forall x \in I,$$

(ii) $$\left|\frac{f(a)}{f'(a)}\right| \le (1-c)\, r.$$

$\Rightarrow$ f hat eine eindeutige Nullstelle $x_* \in I$. Für jedes $x_0 \in I$ konvergiert die Folge

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)} \quad (k \ge 0)$$

gegen x_*. Es gibt eine Konstante $K > 0$ mit

$$(*) \quad |x_{k+1} - x_*| \le K\, |x_k - x_*|^2 \quad \forall k \ge 0$$

Beweis Die Fkt $g: I \to \mathbb{R}$, $g(x) = x - \frac{f(x)}{f'(x)}$ ist stetig differenzierbar mit

$$|g'(x)| = \left|1 - \frac{f'(x)^2 - f(x) f''(x)}{f'(x)^2}\right| = \left|\frac{f(x)\, f''(x)}{f'(x)^2}\right| \overset{(i)}{\le} c < 1 \quad \forall x \in I.$$

$\Rightarrow$ (MWS = 1.6) $\quad |g(x) - g(y)| \le c\,|x-y| \quad \forall x \in I.$ Wegen

$$|g(a) - a| = \left|\frac{f(a)}{f'(a)}\right| \overset{(ii)}{\le} (1-c)\, r$$

ist für alle $x \in I$

$$|g(x) - a| \le |g(x) - g(a)| + |g(a) - a| \le c\, r + (1-c)\, r = r.$$

$\Longrightarrow$ (Banachscher FPS = 1.5)

$g: I \to I$ hat einen eindeutigen Fixpunkt $x_* \in I$ (f eine eindeutige N.stelle!)

Sei $x_0 \in I$ beliebig und $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$ $(k \geq 0)$

$\Longrightarrow$ (Taylorentwicklung von f in x_k mit Restglied 2. Ordnung (Kor. 13.1.1))

$$0 = f(x_*) = f(x_k) + f'(x_k)(x_* - x_k) + \frac{f''(\tilde{x}_k)}{2}(x_* - x_k)^2$$

mit einer geeigneten Stelle $\tilde{x}_k$ zwischen x_k und x_*

$$\Longrightarrow \quad x_* - x_{k+1} = \frac{f(x_k)}{f'(x_k)} + (x_* - x_k) = -\frac{f''(\tilde{x}_k)}{2 f'(x_k)}(x_* - x_k)^2$$

Also folgt (*) mit

$$K = \sup_{x \in I} |f''(x)| \Big/ 2 \inf_{x \in I} |f'(x)| .$$

Eine 2× mal differenzierbare Funktion $f: I \to \mathbb{R}$ auf einem offenen Intervall $I \subset \mathbb{R}$ ist konvex bzw. konkav (Definition 11.2.2) genau dann, wenn

$f''(x) \geq 0 \ \forall x \in I$ bzw. $f''(x) \leq 0 \ \forall x \in I$ (Satz 16.6 in [Forster, Analysis I])

Man kann zeigen: ([Forster, Analysis I], Satz 17.2):

Satz 1.12 Sei $f: [a,b] \to \mathbb{R}$ 2× differenzierbar mit $f(a) < 0 < f(b)$ und $f'' \geq 0$ (oder $f(a) > 0 > f(b)$ und $f'' \leq 0$). Dann gilt:

(a) Es gibt genau ein $x_* \in]a,b[$ mit $f(x_*) = 0$.

(b) Ist $x_0 \in [a,b]$ mit $f(x_0) \geq 0$ ($f(x_0) \leq 0$), so konvergiert die Folge

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)} \quad (k \geq 0)$$

monoton fallend gegen x_*.

(c) Ist $|f''| \leq K$ auf $[a,b]$, so gilt:

$$|x_{k+1} - x_*| \leq \frac{K}{2|f'(x_*)|} |x_{k+1} - x_k|^2 \leq \frac{K}{2|f'(x_*)|} |x_k - x_*|^2 .$$

Mehrdimensionales Newton-Verfahren

Sei $f: \mathbb{R}^n \to \mathbb{R}^n$ eine C^1-Funktion und $x_* \in \mathbb{R}^n$ mit

$$f(x_*) = 0$$

so, dass

$$J_f(x) \in M(n, \mathbb{R}) \text{ regulär ist } \forall x \in \mathbb{R}^n.$$

Für x_0 nahe genug bei x_* approximiere man f linear durch

$$f(x) \approx f(x_0) + J_f(x_0)(x - x_0).$$

Definiere x_1 als Lösung der Gleichung

$$f(x_0) + J_f(x_0)\,(x_1 - x_0) = 0 \iff x_1 = x_0 - J_f(x_0)^{-1} f(x_0)$$

und iteriere

$$x_{k+1} := x_k - J_f(x_k)^{-1} f(x_k) \qquad (k \geq 0)$$

Falls $x := \lim\limits_{k \to \infty} x_k$ existiert, gilt

$$x = \lim_{k \to \infty} x_{k+1} = \lim_{k \to \infty} \left(x_k - J_f(x_k)^{-1} f(x_k)\right) = x - J_f(x)^{-1} f(x) \implies f(x) = 0$$

(Grund: Man kann zeigen, dass die Funktion $\mathbb{R}^n \to \mathbb{R}^n$, $y \mapsto y - J_f(y)^{-1} f(y)$ stetig ist.)

Satz 1.13 Seien $D \subset \mathbb{R}^n$ offen, konvex, $f \in C^1(D, \mathbb{R}^n)$ und $x_* \in D$ mit

$$f(x_*) = 0 \quad \text{und} \quad J_f(x) \text{ regulär } \forall x \in D.$$

Sei $L > 0$ mit $B_{\frac{2}{L}}(x_*) \subset D$ und

$$\text{(L)} \qquad \| J_f(x)^{-1} \left(J_f(y) - J_f(x)\right) \| \leq L \, \|y - x\| \quad \forall x, y \in D.$$

Für jedes $x_0 \in D$ mit $\|x_0 - x_*\| < \frac{2}{L} =: r$ definiert

$$x_{k+1} := x_k - J_f(x_k)^{-1} f(x_k) \quad (k \geq 0)$$

eine Folge $(x_k)_{k \geq 0}$ in $B_r(x_*)$ mit $\lim\limits_{k \to \infty} x_k = x_*$ und

$$\| x_{k+1} - x_* \| \leq \frac{L}{2} \|x_k - x_*\|^2 \quad \forall k \geq 0.$$

Ferner ist x_* die einzige Nullstelle von f in $B_r(x_0)$.

<u>Ideen</u>: Für $x, y \in D$ ist die Funktion

$$\varphi : [0,1] \to \mathbb{R}^n, \quad \varphi(t) = J_f(x)^{-1} f(x + t(y-x))$$

nach der Kettenregel differenzierbar mit

$$\varphi'(t) = J_f(x)^{-1} J_f(x + t(y-x))(y-x) \quad \forall t \in [0,1].$$

$$\Rightarrow \quad \varphi'(t) - \varphi'(0) = \left[J_f(x)^{-1} \left(J_f(x+t(y-x)) - J_f(x) \right) \right] (y-x)$$

$$\overset{(L)}{\Rightarrow} \quad \| \varphi'(t) - \varphi'(0) \| \leq L \, t \, \|y-x\|^2$$

Also gilt

$$\| J_f(x)^{-1} \left(f(y) - f(x) - J_f(x)(y-x) \right) \| = \| \varphi(1) - \varphi(0) - \varphi'(0) \|$$

$$(*) \quad = \| \int_0^1 (\varphi'(t) - \varphi'(0)) \, dt \| \leq \int_0^1 \| \varphi'(t) - \varphi'(0) \| \, dt \leq \frac{L}{2} \|y-x\|^2.$$

Nach Definition der Folge $(x_k)_{k \geq 0}$ ist

$$x_{k+1} - x_* = - J_f(x_k)^{-1} f(x_k) - (x_* - x_k)$$

$$= J_f(x_k)^{-1} \left(f(x_*) - f(x_k) - J_f(x_k)(x_* - x_k) \right)$$

Ist $\|x_k - x_*\| < r = \frac{2}{L}$, so folgt aus $(*)$ mit $y = x_*$, $x = x_k$

$$(**) \quad \| x_{k+1} - x_* \| \leq \underbrace{\left(\frac{L}{2} \|x_k - x_*\| \right)}_{<1} \| x_k - x_* \| \leq \| x_k - x_* \| < r.$$

Folglich konvergiert für $x_0 \in B_r(x_*)$ die Folge $\left(\|x_k - x_*\| \right)_{k \geq 0}$ monoton fallend gegen eine Zahl $d \in [0, r) = [0, \frac{2}{L})$.

$$\Rightarrow \quad 0 \leq d = \lim_{k \to \infty} \|x_{k+1} - x_*\| \overset{(**)}{\leq} \lim_{k \to \infty} \frac{L}{2} \|x_k - x_*\|^2 = \frac{L}{2} d^2$$

Hieraus folgt $d = 0$, denn sonst wäre $\frac{2}{L} \leq d$.

Also konvergiert $(x_k)_{k\geq 0}$ gegen x_* und die Konvergenzordnung ist $p=2$.

Wäre $y \neq x_*$ eine weitere Nullstelle von f in $B_r(x_*)$, so wäre

$$\|y-x_*\| = \| J_f(x_*)^{-1}\big(f(y)-f(x_*)-J_f(x_*)(y-x_*)\big)\|$$
$$\overset{(*)}{\leq} \left(\frac{L}{2}\|y-x_*\|\right)\|y-x_*\| < \|y-x_*\|.$$

Dieser Widerspruch zeigt, dass x_* die einzige Nullstelle von f in $B_r(x_*)$ ist.

Die Voraussetzungen von Satz 1.13 sind lokal für jede C^2-Funktion automatisch erfüllt. Dies folgt aus Teil (b) der nachfolgenden Bemerkung

Bemerkung 1.14

(a) Satz 1.13 bleibt richtig mit demselben Beweis, wenn man überall die euklidische Norm $\|\cdot\| = \|\cdot\|_2 : \mathbb{R}^n$ ersetzt durch $\|\cdot\|_\alpha : \mathbb{R}^n \to \mathbb{R}$ $(\alpha = 1,2,\infty)$ und die zugehörige Matrixnorm $\|\cdot\|_\alpha : M(n,\mathbb{R}) \to \mathbb{R}$

(b) Seien $U \subset \mathbb{R}^n$ offen, $f \in C^1(U,\mathbb{R}^n)$, $x_* \in U$. Dann gilt:

(i) Ist $J_f(x_*)$ regulär, so gibt es ein $r>0$ mit $\overline{B}_r(x_*) \subset U$ und

$$\det J_f(x) \neq 0 \quad \forall x \in \overline{B}_r(x_*).$$

Für jedes solche r ist

$$c := \sup_{x\in\overline{B}_r(0)} \|J_f(x)^{-1}\| < \infty.$$

(ii) Ist $f \in C^2(U,\mathbb{R}^n)$ und ist $\overline{B}_r(x_*) \subset U$, so gibt es ein $\tilde{c}>0$ mit

$$\|J_f(y)-J_f(x)\| \leq \tilde{c}\,\|y-x\| \quad \forall x,y \in \overline{B}_r(x_*).$$

Für r wie in (a) folgt dann

$$\|J_f(x)^{-1}\big(J_f(y)-J_f(x)\big)\| \leq L\|y-x\| \quad \text{mit } L = c\tilde{c}.$$

Ideen für (♭):

(i) Die Funktionen

$$\overline{B}_r(x_*) \ni x \mapsto \det J_f(x) \quad \text{und} \quad x \mapsto \| J_f(x)^{-1} \|$$

sind stetig und $\overline{B}_r(x_*)$ ist kompakt.

(ii) Wende den Mittelwertsatz (1.7) an auf die C^1-Funktionen

$$\partial_j f_i : U \to \mathbb{R} \qquad (1 \le i, j \le n)$$

und benutze, dass für alle $A = (a_{ij})_{1 \le i,j \le n} \in M(n, \mathbb{R})$ gilt

$$\| A \| \le n \max_{1 \le i,j \le n} |a_{ij}| .$$

§ 2 Ausgleichsrechnung

Wir messen den Wert einer Größe

$$y = f(x_1, \dots, x_n, t).$$

Für m verschiedene $t_1, \dots, t_m \in \mathbb{R}$ messen wir die Werte

$$y_1, \dots, y_m \in \mathbb{R}.$$

Wir suchen die am besten zu diesen Beobachtungen passenden Parameter $x_1, \dots, x_n \in \mathbb{R}$, das heißt das Tupel $x_* = (x_1^*, \dots, x_n^*) \in \mathbb{R}^n$, so dass

$$m := \| (f(x_*, t_i))_{i=1}^m - (y_i)_{i=1}^m \|$$

minimal wird. Optimale Kompatibilität zwischen den Daten und dem zu Grunde gelegten Modell liegt vor, wenn $m = 0$ ist.

A Hängt f linear von $(x_1, \dots, x_n) \in \mathbb{R}^n$ ab, führt dies zu dem Problem, zu einer gegebenen Matrix $A \in M(m \times n, \mathbb{R})$ und gegebenem $y \in \mathbb{R}^m$ das Residuum

$$r(x) = \| Ax - y \|$$

zu minimieren, d.h. ein $x_* \in \mathbb{R}^n$ zu finden mit

$$(*) \qquad r(x_*) \leq r(x) \qquad \forall x \in \mathbb{R}^n.$$

<u>Satz 2.1</u> (Lineare Ausgleichsrechnung) Sei $A \in M(m \times n, \mathbb{R})$ eine Matrix mit $m \geq n$ und $\operatorname{rg}(A) = n$. Dann gilt:

(i) $A^t A \in M(n, \mathbb{R})$ ist regulär,

(ii) $x_* = (A^t A)^{-1} A^t y$ ist die eindeutige Lösung von $(*)$.

<u>Idee</u> (i) Für die lineare Abbildung $M_A : \mathbb{R}^n \to \mathbb{R}^m$, $x \mapsto Ax$ gilt

$$\dim \operatorname{Kern} M_A = n - \operatorname{rg}(A) = 0 \quad \text{(Dimensionsformel = Satz 8.2.4)}$$

$\Rightarrow$ $M_A : \mathbb{R}^n \to \mathbb{R}^m$ ist injektiv. Wegen

$$\langle A^t A x, x \rangle = \langle Ax, Ax \rangle = \| Ax \|^2$$

ist dann auch $\mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto A^t A x$ injektiv.

Da auf beiden Seiten der $\mathbb{R}^n$ steht, ist diese Abbildung also auch bijektiv

$\Rightarrow$ $A^t A \in M(n, \mathbb{R})$ ist regulär

(ii) Definiere $x_* := (A^t A)^{-1} A^t y$. Dann gilt für alle $x \in \mathbb{R}^n \setminus \{0\}$:

$$\|A(x_* + x) - y\|^2 = \| Ax_* - y \|^2 - 2 \langle Ax_* - y, Ax \rangle + \| Ax \|^2$$

$$= \|Ax_* - y\|^2 - 2 \langle \underbrace{A^t A x_* - A^t y}_{= 0}, x \rangle + \underbrace{\| Ax \|^2}_{\neq 0} > \| Ax_* - y \|^2$$

$\square$

B Nichtlineare Abhängigkeit

Statt $\mathbb{R}^n \to \mathbb{R}^m$, $x \mapsto Ax - y$ betrachten wir eine C^2-Funktion

$$F : \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto \begin{pmatrix} F_1(x) \\ \vdots \\ F_m(x) \end{pmatrix} \quad \text{mit } m \geq n$$

Gesucht: Ein Vektor $x_* \in \mathbb{R}^n$ mit

$$\| F(x_*) \| \leq \| F(x) \| \quad \forall x \in \mathbb{R}^n.$$

Wir suchen zunächst nur lokale Minima der C^2-Funktion

$$g : \mathbb{R}^n \to \mathbb{R}, \quad g(x) = \frac{1}{2} \| F(x) \|^2 = \frac{1}{2} \sum_{i=1}^{m} F_i(x)^2.$$

Notwendige Bedingung (Satz 17.3.2):

$$\operatorname{grad} g(x_*) = \nabla g(x_*) = 0 \qquad (x_* \text{ ist kritischer Pkt von } g)$$

Hinreichende Bedingung (Satz 17.3.3):

$$\nabla g(x_*) = 0 \text{ und } (\operatorname{Hess} g)(x_*) \text{ ist positiv definit}$$

(1) $$\operatorname{grad} g(x) = \frac{1}{2} \sum_{i=1}^{m} \operatorname{grad}\left(F_i(x)^2\right) = \sum_{i=1}^{m} F_i(x)\left(\operatorname{grad} F_i(x)\right)$$

$$= \left(\sum_{i=1}^{m} F_i(x)(\partial_1 F_i)(x), \ldots, \sum_{i=1}^{m} F_i(x)(\partial_n F_i)(x) \right) = F(x)^t J_F(x)$$

(2) $$(\operatorname{Hess} g)(x) = \left(\frac{\partial^2 g}{\partial x_j \partial x_k}(x) \right)_{1 \leq j,k \leq n} = \left(\frac{\partial}{\partial x_j} \left(\sum_{i=1}^{m} F_i(x)(\partial_k F_i)(x) \right) \right)_{1 \leq j,k \leq n}$$

$$= \left(\sum_{i=1}^{m} (\partial_j F_i)(x)(\partial_k F_i)(x) \right)_{j,k} + \left(\sum_{i=1}^{m} F_i(x)(\partial_j \partial_k F_i)(x) \right)_{j,k}$$

$$= J_F(x)^t J_F(x) + \underbrace{\sum_{i=1}^{m} F_i(x)(\operatorname{Hess} F_i)(x)}_{=: \psi(x)}$$

Da $g : \mathbb{R}^n \to \mathbb{R}$ eine C^2-Funktion ist, ist das Gradienten-Feld

$$f : \mathbb{R}^n \to \mathbb{R}^n, \quad f(x) = \left(\operatorname{grad} g(x)\right)^t$$

eine C^1-Funktion mit

$$J_f(x) = \left((\partial_j \partial_i g)(x) \right)_{1 \leq i,j \leq n} = \operatorname{Hess} g(x).$$

Ist $F : \mathbb{R}^n \to \mathbb{R}^m$ eine C^3-Funktion, so ist $f = \operatorname{grad} g : \mathbb{R}^n \to \mathbb{R}^n$ eine C^2-Funktion und man kann das mehrdimensionale

Newton-Verfahren (Satz 1.13 + Bemerkung 1.14) benutzen, um die kritischen Stellen von g, das heißt Stellen $x_* \in \mathbb{R}^n$ mit $\operatorname{grad} g(x_*) = 0$ zu berechnen:

Ist $\operatorname{grad} g(x_*) = 0$ und $\operatorname{Hess} g(x_*)$ positiv definit ($\Rightarrow$ regulär), so gibt es nach Satz 1.13 + Bemerkg 1.14 ein $r > 0$, sodass $\forall x_0 \in B_r(x_*)$ die Iterationsfolge

$$
\begin{aligned}
x_{k+1} &= x_k - J_f(x_k)^{-1} f(x_k) \\
&= x_k - \left(\operatorname{Hess} g(x_k)\right)^{-1} \left(\operatorname{grad} g(x_k)\right)^t \\
&= x_k - \left(J_F(x_k)^t J_F(x_k) + 4(x_k)\right)^{-1} J_F(x_k)^t \, F(x_k)
\end{aligned}
$$

gegen x_* konvergiert.

Unter geeigneten Bedingungen (siehe etwa Satz 4.15 in [Deuflhard - Hohmann I]) gilt dasselbe Ergebnis für das Gauß-Newton Verfahren, das heißt für die Iterationsfolge

$$x_{k+1} = x_k - \left(J_F(x_k)^t J_F(x_k)\right)^{-1} J_F(x_k)^t \, F(x_k) \qquad (k \geq 0)$$

und Startwerte $x_0 \in B_r(x_*)$ mit genügend kleinem $r > 0$.

Idee dabei: Ersetze F in der Nähe von x_k durch die lineare Approximation

$$x \mapsto F(x_k) + J_F(x_k)(x - x_k)$$

und wähle x_{k+1} so dass die Norm der rechten Seite für $x = x_{k+1}$ minimal wird. Nach Satz 2.1 gilt dies für

$$x_{k+1} - x_k = \left(J_F(x_k)^t J_F(x_k)\right)^{-1} J_F(x_k)^t \left(-F(x_k)\right).$$

§3 Iterative Löser für lineare Gleichungssysteme

Sei V ein $\mathbb{R}$-Vektorraum (analog für $\mathbb{C}$-Vektorräume)

A Matrixnormen

Definition 3.1 Sei $\|\cdot\| : V \to \mathbb{R}$ eine Norm, das heißt eine Abbildung mit

(i) $\|x\| \geq 0$, $\|x\| = 0 \Leftrightarrow x = 0$,

(ii) $\|\alpha x\| = |\alpha| \, \|x\|$,

(iii) $\|x+y\| \leq \|x\| + \|y\|$

für alle $x, y \in V$, $\alpha \in \mathbb{R}$. Für $x_0 \in V$, $r > 0$ schreiben wir

$$B_r(x_0) = \{x \in V;\ \|x - x_0\| < r\}, \quad \overline{B}_r(x_0) = \{x \in V;\ \|x - x_0\| \leq r\}.$$

Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf V heißen äquivalent, falls $c, \tilde{c} > 0$ existieren mit

$$\|x\|_1 \leq c \|x\|_2 \quad \forall x \in V \quad \text{und} \quad \|x\|_2 \leq \tilde{c} \|x\|_1 \quad \forall x \in V.$$

Beispiele 3.2 Durch

$$\|(x_i)\|_1 = \sum_{i=1}^{n} |x_i| \quad \text{(Summennorm)}$$

$$\|(x_i)\| = \|(x_i)\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \quad \text{(Euklidische Norm)}$$

$$\|(x_i)\|_\infty = \max_{i=1,\dots,n} |x_i| \quad \text{(Maximumnorm)}$$

werden äquivalente Normen auf $\mathbb{R}^n$ definiert. Für alle $x \in \mathbb{R}^n$ gilt

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\, \|x\|_\infty,$$

$$\|x\|_\infty \leq \|x\|_1 \leq n \, \|x\|_\infty.$$

Wie in Definition 1.2 definiert man konvergente Folgen und Cauchy-Folgen in V.

Ein Banachraum ist ein vollständiger normierter Raum (Def. 1.2).

Man nennt eine Menge

$\mathcal{U} \subset V$ <u>offen</u> $:\Leftrightarrow \forall x \in \mathcal{U}\ \exists \varepsilon > 0$ mit $B_\varepsilon(x) \subset \mathcal{U}$

$A \subset V$ <u>abgeschlossen</u> $:\Leftrightarrow$ Ist (x_k) eine Folge in A, die in V konvergiert gegen ein $x \in V$, so ist auch $x \in A$.

Sind V, W normierte Räume und ist $D \subset V$, so heißt eine Abbildung

$$f : D \to W$$

<u>stetig</u>, falls

$$(x_k) \xrightarrow{k} x \text{ in } D \Rightarrow (f(x_k)) \xrightarrow{k} f(x) \text{ in } W$$

In der Mathematik zeigt man (Eschmeier, Skript zur Funktionalanalysis, Satz 4.15):

<u>Satz 3.3</u> Ist $\dim V < \infty$, so sind alle Normen auf V äquivalent. Insbesondere gilt dies für $V = \mathbb{R}^n$ und $V = M(n, \mathbb{R})$.

<u>Folgerung 3.4</u> Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$.
$\Rightarrow$ Für jede lineare Abb. $T : \mathbb{R}^n \to \mathbb{R}^n$ gibt es eine Konstante $c > 0$ mit

$$\|T(x)\| \le c\,\|x\| \quad \forall x \in \mathbb{R}^n$$

<u>Beweis.</u> Wegen <u>Satz 3.3</u> genügt es, dies für die Summennorm zu zeigen:

$$\|T(x_i)\|_1 = \Big\|\sum_{i=1}^n x_i T(e_i)\Big\|_1 \le \sum_{i=1}^n |x_i|\,\|Te_i\|_1 \le \Big(\max_{i=1,\dots,n}\|Te_i\|_1\Big)\,\|(x_i)\|_1.$$

Hierbei seien $e_i = (0,\dots,0,1,0,\dots,0) \in \mathbb{R}^n$ die Einheitsvektoren. ☒

<u>Folgerung 3.5</u> Ist $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$, so definiert

$$(*) \quad \|\cdot\| : M(n,\mathbb{R}) \to \mathbb{R}, \quad \|A\| = \sup_{\|x\| \le 1} \|Ax\|$$

eine Norm auf $M(n,\mathbb{R})$. Es gilt:

(i) $\|A\| = \min\{c \geq 0;\ \|Ax\| \leq c\|x\| \ \forall x \in \mathbb{R}^n\}$,

(ii) $\|AB\| \leq \|A\|\,\|B\| \quad \forall A, B \in M(n,\mathbb{R})$

Normen der Form (*) nennen wir <u>Matrixnormen</u> auf $M(n,\mathbb{R})$

<u>Ideen</u>:

$$\|A+B\| = \sup_{\|x\|\leq 1} \|(A+B)x\| \leq \sup_{\|x\|\leq 1} \underbrace{(\|Ax\| + \|Bx\|)}_{\leq \|A\|+\|B\|} \leq \|A\| + \|B\|$$

(i) folgt aus:

- $\|Ax\| = \|x\|\,\|A\frac{x}{\|x\|}\| \leq \|A\|\,\|x\|$
- Ist $c \geq 0$ mit $\|Ax\| \leq c\|x\| \ \forall x \in V$, so ist $\|A\| = \sup_{\|x\|\leq 1}\|Ax\| \leq c$.

(ii) folgt wie in <u>Aufgabe 2</u>. ☒

<u>Definition 3.6</u> Für eine Matrix $A \in M(n,\mathbb{R})$ nennt man die Zahl

$$\varrho(A) := \max\{|\lambda|;\ \lambda \text{ Eigenwert von } A \text{ in } \mathbb{C}\}$$

den <u>Spektralradius</u> von A.

Man zeigt

<u>Satz 3.7</u> Ist $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$ und ist $\|\cdot\|: M(n,\mathbb{R}) \to \mathbb{R}$ die zugehörige Matrixnorm, so gilt

$$\varrho(A) = \lim_{k\to\infty} \|A^k\|^{\frac{1}{k}}.$$

Insbesondere ist der Limes rechts unabhängig von der Wahl der Norm $\|\cdot\|$ auf $\mathbb{R}^n$, und es gilt

$$\varrho(A) \leq \|A\| \quad \forall A \in M(n,\mathbb{R})$$

<u>Warnung</u> Der Spektralradius ist keine Norm, z.B. ist

$$\varrho\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) = 0.$$

<u>Beispiele 3.8</u> Die Matrixnormen zu

$$\left\{\begin{matrix} \|\cdot\|_1 \\ \|\cdot\|_\infty \\ \|\cdot\|_2 \end{matrix}\right\} \text{ auf } \mathbb{R}^n \text{ sind } \left\{\begin{array}{lll} \|A\|_1 = \max\limits_{j=1,\dots,n} \sum\limits_{i=1}^{n} |a_{ij}| & (\underline{\text{Spaltensummennorm}}) \\ \|A\|_\infty = \max\limits_{i=1,\dots,n} \sum\limits_{j=1}^{n} |a_{ij}| & (\underline{\text{Zeilensummennorm}}) \\ \|A\|_2 = \sqrt{\varrho(A^tA)} & (\underline{\text{Spektralnorm}}) \end{array}\right.$$

Hierbei sei $A = (a_{ij})_{1\le i,j\le n} \in M(n,\mathbb{R})$.

<u>Beweis</u>: <u>Für $\|\cdot\|_\infty$</u>: Aufgabe 2 (b). <u>Für $\|\cdot\|_1$</u>: ähnlich.

<u>Für $\|\cdot\|_2$</u>: Da A^tA symmetrisch und positiv semidefinit ist, gibt es eine Orthonormalbasis $\{b_1,\dots,b_n\}$ von $\mathbb{R}^n$ und $\lambda_1,\dots,\lambda_n \in \mathbb{R}_{\ge 0}$ mit

$$A^tA\, b_i = \lambda_i b_i \quad (i=1,\dots,n) \quad (\text{Satz } 15.2.3 \text{ und Beweis}).$$

$\Rightarrow$ Für alle $(x_i)_{i=1}^n \in \mathbb{R}^n$ gilt

$$\Big\| A\Big(\sum_{i=1}^n x_i b_i\Big)\Big\|_2^2 = \Big\langle A^tA\Big(\sum_{i=1}^n x_i b_i\Big), \sum_{i=1}^n x_i b_i \Big\rangle = \sum_{i=1}^n \lambda_i x_i^2$$
$$\le \Big(\max_{i=1,\dots,n} \lambda_i\Big) \sum_{i=1}^n x_i^2 = \varrho(A^tA)\, \Big\|\sum_{i=1}^n x_i b_i\Big\|^2$$

Also ist $\|A\|_2 \le \sqrt{\varrho(A^tA)}$. Andererseits ist

$$\|A\|_2^2 \ge \|A b_i\|_2^2 = \langle A^tA\, b_i, b_i\rangle = \lambda_i \quad (i=1,\dots,n)$$

$$\Rightarrow \|A\|_2^2 \ge \varrho(A^tA)$$ ☒

<u>Beispiel 3.9</u> Für

$$A = \left(\begin{array}{c|c} 0 & -1 \\ \hline 2 & -3 \end{array}\right) \in M(2,\mathbb{R})$$

ist

$$\|A\|_1 = \max\left(\left\|\begin{pmatrix}0\\2\end{pmatrix}\right\|_1, \left\|\begin{pmatrix}-1\\-3\end{pmatrix}\right\|_1\right) = \max(2,4) = 4$$

$$\|A\|_\infty = \max\left(\left\|\begin{pmatrix}0\\-1\end{pmatrix}\right\|_1, \left\|\begin{pmatrix}2\\-3\end{pmatrix}\right\|_1\right) = \max(1,5) = 5$$

$$A^tA = \left(\begin{array}{c|c} 0 & 2 \\ \hline -1 & -3 \end{array}\right)\left(\begin{array}{c|c} 0 & -1 \\ \hline 2 & -3 \end{array}\right) = \left(\begin{array}{c|c} 4 & -6 \\ \hline -6 & 10 \end{array}\right)$$

Die Eigenwerte von A^tA sind die Lösungen von

$$(4-x)(10-x)-36 = \det\left(A^tA - x E_2\right) = 0$$

$$\Leftrightarrow \quad x^2 - 14x + 4 = 0$$

$$\Leftrightarrow \quad (x-7)^2 - (49-4) = 0$$

$$\Leftrightarrow \quad x_\pm = 7 \pm 3\sqrt{5} \quad (\geq 0)$$

Also ist

$$\|A\|_2 = \sqrt{\varrho(A^tA)} = \sqrt{\max(x_+, x_-)} = \sqrt{7+3\sqrt{5}}$$

$$\cong 3.70$$

Zum Vergleich berechnen wir auch den Spektralradius von A.
Die komplexen Eigenwerte von A sind genau die komplexen Lösungen von

$$(-x)(-3-x)+2 = \det(A - x E_2) = 0$$

$$\Leftrightarrow \quad x^2 + 3x + 2 = 0$$

$$\Leftrightarrow \quad \left(x+\tfrac{3}{2}\right)^2 - \left(\tfrac{9}{4}-2\right) = 0$$

$$\Leftrightarrow \quad x_\pm = -\tfrac{3}{2} \pm \sqrt{\tfrac{1}{4}} = -\tfrac{3}{2} \pm \tfrac{1}{2}$$

Also ist

$$\varrho(A) = \max\left(\left|-\tfrac{3}{2}+\tfrac{1}{2}\right|, \left|-\tfrac{3}{2}-\tfrac{1}{2}\right|\right) = 2.$$

<u>Definition 3.10</u> Für eine Matrix $A \in M(n,\mathbb{R})$ nennt man die Menge

$$\sigma(A) = \{\lambda \in \mathbb{C};\ \lambda \text{ ist Eigenwert von } A\}$$

das <u>Spektrum von A</u> (in $\mathbb{C}$).

<u>Bemerkung 3.11</u> Für $A \in M(n,\mathbb{C})$ und $k \in \mathbb{N}$ ist

$$\sigma(A^k) = \{\lambda^k;\ \lambda \in \sigma(A)\} \text{ und daher } \varrho(A^k) = \varrho(A)^k.$$