Lemma 1.12 Sei $A \in \mathbb{R}^{n \times n}$ mit $\|A\| < 1$. Dann konvergiert $\sum_{k\geq 0} A^k$ „Neumannsche Reihe" (mit $A^0 = E_n$). Insbesondere ist $E_n - A$ invertierbar, $\sum_{k\geq 0} A^k = (E_n - A)^{-1}$ („$= \frac{1}{1-A}$") und $\|(E_n - A)^{-1}\| \leq \frac{1}{1-\|A\|}$

Beweis: Sei $S_l = \sum_{k=0}^{l} A^k$, $l \in \mathbb{N}_0$. Für $m \geq l$ erhalten wir:

$$\|S_m - S_l\| = \|\sum_{k=l+1}^{m} A^k\| \overset{d)}{\leq} \sum_{k=l+1}^{m} \|A^k\| \overset{e)}{\leq} \sum_{k=l+1}^{m} \|A\|^k \leq \|A\|^{l+1} \cdot \sum_{m\geq 0} \|A\|^m = \frac{1}{1-\|A\|} \|A\|^{l+1}$$

und somit $\|S_m - S_l\|_\infty \xrightarrow{l\to\infty} 0$. Damit ist $(S_l)_{l\in\mathbb{N}_0}$ Cauchy-Folge, also konvergiert die Neumannsche Reihe. Ferner gilt: $(E_n - A) \cdot \sum_{k\geq 0} A^k =$
$= \sum_{k\geq 0} A^k - \sum_{k\geq 0} A^{k+1} = E_n$ und somit $\sum_{k\geq 0} A^k = (E_n - A)^{-1}$.
Zuletzt: $\|\sum_{k\geq 0} A^k\| \leq \sum_{k\geq 0} \|A\|^k = \frac{1}{1-\|A\|}$ □

Lemma 1.13 (Banach) Sei $A \in \mathbb{R}^{n\times n}$ invertierbar, $B \in \mathbb{R}^{n\times n}$ mit $\|A^{-1}B\| < 1$. Dann $A+B$ invertierbar und $\|(A+B)^{-1}\| \leq \frac{\|A^{-1}\|}{1-\|A^{-1}B\|}$, $\|(A+B)^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2 \|B\|}{1-\|A^{-1}B\|}$

Beweis: Da $\|A^{-1}B\| < 1$. mit Lemma 1.12 ist $E_n - A^{-1}B$ invertierbar mit $(E_n + A^{-1}B)^{-1} = \sum_{k\geq 0} (-A^{-1}B)^k$ Damit ist auch $A(E_n + A^{-1}B) = A+B$ invertierbar und $(A+B)^{-1} = (E_n + A^{-1}B)^{-1} A^{-1}$. Daraus: $\|(A+B)^{-1}\| \leq \|(E_n + A^{-1}B)^{-1}\| \cdot \|A^{-1}\| \leq$
$\frac{\|A^{-1}\|}{1-\|A^{-1}B\|}$. Ferner gilt: $(A+B)^{-1} - A^{-1} = (E_n + A^{-1}B)^{-1} \cdot A^{-1} - A^{-1} =$
$= ((E_n + A^{-1}B)^{-1} - E_n) A^{-1} = (\sum_{k\geq 0} (-A^{-1}B)^k - E_n) A^{-1} = (\sum_{k\geq 1} (-A^{-1}B)^k) A^{-1} =$
$= -A^{-1}B \cdot (\sum_{k\geq 0} (-A^{-1}B)^k) A^{-1} = -A^{-1} \cdot B \cdot (E_n + A^{-1}B)^{-1} A^{-1}$.
Mit Submultiplikativität von $\|\cdot\|$ erhalten wir die 2. Aussage □

Bemerkung: Wir erhalten: 25.10.18

1) Jede invertierbare Matrix hat eine Umgebung invertierbarer Matrizen:
$U(A) = \{\tilde{A} \in \mathbb{R}^{N\times N} \mid \|A - \tilde{A}\| < \delta\}$ wobei $\delta = \|A^{-1}\|^{-1}$
[Für $\tilde{A} \in U_\delta(A)$ setze $B = A - \tilde{A}$. Dann $\|B\| < \delta = \|A^{-1}\|^{-1}$ und $\|A^{-1}B\| \leq \|A^{-1}\| \|B\| < 1$]

2) Für jede Folge $A_k \to A$ invertierbarer Matrizen gilt:
$A_k^{-1} \to A^{-1}$
[Sei $B_k = A_k - A$ $\forall k \in \mathbb{N}$. Dann $\|B_k\| \xrightarrow{k\to\infty} 0$. Insbesondere für k groß genug: $\|A_k^{-1} - A^{-1}\| = \|(A+B_k)^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2 \|B_k\|}{1-\|A^{-1}B_k\|} \xrightarrow{k\to\infty} 0$]

Lemma 1.14 Sei $F: \mathbb{R}^n \to \mathbb{R}^n$ stetig diff.bar, $\tilde{x} \in \mathbb{R}^n$ mit $F(\tilde{x}) = 0$, $F'(\tilde{x})$ invertierbar. Dann $\exists \gamma, \delta > 0$ mit $\|F(x)\| \geq \gamma \|x - \tilde{x}\|$, $x \in B_\delta(\tilde{x})$. Insbesondere ist $\tilde{x}$ die einzige Nullstelle in $B_\delta(\tilde{x})$.

Beweis: Aus Eigenschaften der Matrixnorm: $\|x - \tilde{x}\| = \|F'(\tilde{x})^{-1} F'(\tilde{x})(x - \tilde{x})\| \leq$
$\leq \|F'(\tilde{x})^{-1}\| \cdot \|F'(\tilde{x})(x - \tilde{x})\|$. Mit $\gamma = \frac{1}{2} \|F'(\tilde{x})^{-1}\|^{-1}$ gilt dann:
$2\gamma \|x - \tilde{x}\| \leq \|F'(\tilde{x})(x - \tilde{x})\|$. Da F stetig diff.bar:
$\|F(x) - F(\tilde{x}) - F'(\tilde{x})(x - \tilde{x})\| \leq \gamma \|x - \tilde{x}\|$ in einer Umgebung von $\tilde{x}$.
Mit $F(\tilde{x}) = 0$: $2\gamma \|x - \tilde{x}\| \leq \|F'(\tilde{x})(x - \tilde{x})\| = \|F(x) - F'(\tilde{x})(x - \tilde{x}) - F(x)\| \leq$
$\leq \|F(x) - F'(\tilde{x})(x - \tilde{x})\| + \|F(x)\| \leq \gamma \cdot \|x - \tilde{x}\| + \|F(x)\| \Rightarrow \|F(x)\| \geq \gamma \|x - \tilde{x}\|$ □

Satz 1.15 Sei $F: \mathbb{R}^n \to \mathbb{R}^n$ stetig diff.bar, $\tilde{x} \in \mathbb{R}^n$ mit $F(\tilde{x}) = 0$, $F'(\tilde{x})$ invert. Dann $\exists \delta, C > 0$ mit:

a) $\forall x \in B_\delta(\tilde{x})$ ist $F'(x)$ invertierbar und $\|F'(x)^{-1}\| \leq C$

b) Für $x_0 \in B_\delta(\tilde{x})$ liegt die Gesamte Folge $(x_k)_{k \in \mathbb{N}_0}$ in $B_\delta(\tilde{x})$.

c) Für $x_0 \in B_\delta(\tilde{x})$ gibt es $k \in \mathbb{N}$ mit $x_k = \tilde{x}$ oder $(x_k)_{k \in \mathbb{N}_0}$ konvergiert superlinear gegen $\tilde{x}$.

d) Ist F' Lokal Lipschitz-stetig, d.h. für ein $L > 0$: $\|F'(x) - F'(y)\| \leq L \cdot \|x - y\|$ für $x, y \in B_\delta(\tilde{x})$, dann ist die Konvergenz sogar quadratisch.

Beweis: a): folgt direkt aus Banach's Lemma 1.13 und der Stetigkeit von F'.

b), c): Sei $x_k \in B_\delta(\tilde{x})$, $F'(x_k)$ invertierbar. Dann gilt:

$$x_{k+1} - \tilde{x} = x_k - F'(x_k)^{-1} F(x_k) - \tilde{x} = F'(x_k)^{-1} [\underbrace{F(\tilde{x})}_{=0} - F(x_k) + F'(x_k)(x - \tilde{x})] =$$

Taylor $\rightarrow$ $= F'(x_k)^{-1} \cdot h(x_k - \tilde{x}) \cdot (x_k - \tilde{x})$ mit $h(y) \to 0$ mit $y \to 0$.

Wähle $\delta > 0$ so klein, dass $h(x_k - \tilde{x}) \leq \frac{1}{2C}$ $\forall x_k \in B_\delta(\tilde{x})$.

Damit erhalten wir: $\|x_{k+1} - \tilde{x}\| / \|x_k - \tilde{x}\| \leq \underbrace{\|F'(x_k)^{-1}\|}_{\leq C} \cdot \frac{1}{2C} \leq \frac{1}{2}$

Somit $x_k \to \tilde{x}$ für $k \to \infty$ und die gesamte Folge $(x_k)_{k \in \mathbb{N}_0}$ liegt in $B_\delta(\tilde{x})$. Ferner ist: $\|x_{k+1} - \tilde{x}\| / \|x_k - \tilde{x}\| \leq \|F'(x_k)^{-1}\| \|h(x_k - \tilde{x})\| \xrightarrow{k \to \infty} 0$

$\Rightarrow$ superlineare Konvergenz.

d) Wie bereits gezeigt: $x_{k+1}-\tilde{x} = F'(x_k)^{-1}[F(\tilde{x})-F(x_k)+F'(x_k)(x-\tilde{x})]$

Benutze Hauptsatz der Integralrechnung:

$H: U \to \mathbb{R}^n$ stetig diff.bar, $x \in U$, $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq U$. Dann gilt
$\forall d \in B_\varepsilon(0): H(x+d) = H(x) + \int_0^1 H'(x+td)dt \cdot d$

Setze im Hauptsatz $x := \tilde{x}$ und $x+d = x_k$, also $d = x_k - \tilde{x}$. Wir erhalten:

$x_{k+1}-\tilde{x} = F'(x_k)^{-1} \int_0^1 F'(x_k) - F'(\tilde{x}+t(x_k-\tilde{x}))\, dt \cdot (x_k-\tilde{x})$

Aus der Annahme: $\|x_{k+1}-\tilde{x}\| \leq C \int_0^1 L(1-t)\|x_k-\tilde{x}\|dt \cdot (x_k-\tilde{x}) =$
$= \frac{L \cdot C}{2}\|x_k-x\|^2 \Rightarrow$ quadratische Konvergenz $\square$

Optimierungsprobleme: $f: \mathbb{R}^n \to \mathbb{R}$, $\min\{f(x) \mid x \in \mathbb{R}^n\}$.

Bestimme stationäre Punkte über $\nabla f(x) = 0$, also benutze $F(x) = \nabla f(x)$.

Annahme: f 2x stetig diff.bar

Newton-Verfahren: Geg.: Startwert $x_0 \in \mathbb{R}^n$. Für $k = 0,1,2,\dots$

a) Falls $\nabla f(x_k) = 0$ STOP

b) $d_k = -\nabla^2 f(x_k)^{-1} \cdot \nabla f(x_k) = -F'(x_k)^{-1} \cdot F(x_k)$

$x_{k+1} = x_k + d_k$

Aus Satz 1.15: lokale Konvergenz gegen stationäre Punkte mit invertierbarer Hessematrix.

2 Konvexe Mengen & Funktionen

2.1 Konvexe Mengen

Definition 2.1: Eine Menge $M \subseteq \mathbb{R}^n$ heißt konvex, falls $\forall x_1, x_2 \in M$, $\lambda \in [0,1]$: $\lambda x_1 + (1-\lambda)x_2 \in M$

Bsp 2.2
1) $\emptyset$, $\mathbb{R}^n$ sind konvex, einelementrige Mengen sind konvex
2) Hyperebenen sind konvex: $H = \{x \in \mathbb{R}^n : y^T x = b\}$, $y \in \mathbb{R}^n \setminus \{0\}$, $b \in \mathbb{R}$
3) Halbräume sind konvex: $H^a = \{x \in \mathbb{R}^n : y^T x \geq b\}$, $H^o = \{x \in \mathbb{R}^n : y^T x > b\}$