Kapitel 1 - Normierte Vektorräume

Vektorräume

Eine Menge V heißt $\mathbb{K}$-Vektorraum, wenn:
1. $\forall \mathbf{v}_1, \mathbf{v}_2 \in V : \mathbf{v}_1 + \mathbf{v}_2 \in V$
2. $\forall \mathbf{v} \in V \forall k \in \mathbb{K} : k \cdot \mathbf{v} \in V$
3. $\forall \mathbf{v}_1, \mathbf{v}_2 \in V : \mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1$
4. $\forall \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V : (\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 = \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3)$
5. $\forall \mathbf{v} \in V \exists \mathbf{0} \in V : \mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}$
6. $\forall \mathbf{v} \in V \exists \mathbf{v}' \in V : \mathbf{v} + \mathbf{v}' = \mathbf{v}' + \mathbf{v} = \mathbf{0}$
7. $\forall \mathbf{v} \in V \exists 1 \in \mathbb{K} : 1 \cdot \mathbf{v} = \mathbf{v}$
8. $\forall \mathbf{v} \in V \forall k, l \in \mathbb{K} : (k \cdot l) \cdot \mathbf{v} = k \cdot (l \cdot \mathbf{v})$
9. $\forall \mathbf{v}_1, \mathbf{v}_2 \in V \forall k \in \mathbb{K} : k \cdot (\mathbf{v}_1 + \mathbf{v}_2) = k \cdot \mathbf{v}_1 + k \cdot \mathbf{v}_2$
10. $\forall \mathbf{v} \in V \forall k, l \in \mathbb{K} : (k + l) \cdot \mathbf{v} = k \cdot \mathbf{v} + l \cdot \mathbf{v}$

Vektorräume im $\mathbb{K}^n$

Ist $V \subseteq \mathbb{K}^n$, dann ist V ein $\mathbb{K}$-VR, g.d.w.:
1. $\forall \mathbf{v}_1, \mathbf{v}_2 \in V : \mathbf{v}_1 + \mathbf{v}_2 \in V$
2. $\forall \mathbf{v} \in V \forall k \in \mathbb{K} : k \cdot \mathbf{v} \in V$

(Die Addition und Skalarmultiplikation ist hierbei komponentenweise).

Wichtige Vektorräume

$\mathbb{R}^n, \mathbb{C}^n$
$\mathcal{C}^k(D)$ für $k \in \mathbb{N}_0 \cup \{\infty\}$ - Raum der k-mal stetig diff.baren Funktionen mit Definitionsbereich D.
$\mathcal{L}^p(\Omega, \mathcal{A}, \mu) :=$
$\{f : \Omega \to \mathbb{K} | f \text{ ist messbar}, \int_\Omega |f(x)|^p d\mu(x) < \infty\}$

Normierte Vektorräume

Sei X ein $\mathbb{K}$-VR. Eine Fkt. $\|\cdot\| : X \to \mathbb{R}$ heißt Norm auf X, wenn:
1. $\forall \mathbf{x} \in X : \|\mathbf{x}\| = 0 \iff \mathbf{x} = 0$
2. $\forall \mathbf{x} \in X \forall \lambda \in \mathbb{K} : \|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$
3. $\forall \mathbf{x}, \mathbf{y} \in X : \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

$(X, \|\cdot\|)$ heißt normierten Vektorraum

Wichtige Normen

$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p}$ heißt allgemein p-Norm.
$\|\mathbf{x}\|_\infty := \max\{|x_1|, \ldots, |x_n|\}$ heißt Maximumsnorm.
$\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i|$ heißt Summennorm.
$\|\mathbf{x}\|_2 := \sqrt[2]{\sum_{i=1}^n |x_i|^2}$ heißt Euklidische Norm.
$\|A\|_F := \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$ heißt Frobeniusnorm.
$\|[f]\|_{L^p(\Omega)} := (\int_\Omega |f(x)|^p dx)^{\frac{1}{p}}$ heißt L^p-Norm.
$\|f\|_\infty := \sup\{|f(x)| | x \in D\}$ heißt Supremumsnorm

Offene und abgeschlossene Mengen

Sei $(X, \|\cdot\|)$ ein normierter VR. Wir definieren:
$B_\varepsilon(\mathbf{x}) := \{y \in X | \|\mathbf{x} - \mathbf{y}\| < \varepsilon\}$ (offener Ball um $\mathbf{x}$).
$\overline{B}_\varepsilon(\mathbf{x}) := \{y \in X | \|\mathbf{x} - \mathbf{y}\| \leq \varepsilon\}$ (abgesch. Ball um $\mathbf{x}$).
$S_\varepsilon(\mathbf{x}) := \{y \in X | \|\mathbf{x} - \mathbf{y}\| = \varepsilon\}$ (Sphäre um $\mathbf{x}$).

$\mathbf{x}$ ist innerer Punkt von A, wenn $\exists \varepsilon > 0 : B_\varepsilon(\mathbf{x}) \subseteq A$.
Die Menge aller inneren Punkte von A heißt das Innere von A (A° oder int A).
$\mathbf{x}$ heißt Randpunkt von A, wenn $\forall \varepsilon > 0 : A \cap B_\varepsilon(\mathbf{x}) \neq \emptyset$ und $A^c \cap B_\varepsilon(\mathbf{x}) \neq \emptyset$. Die Menge aller Randpunkte von A heißt der Rand von A (∂A).
Die Menge $\bar{A} = A \cup \partial A$ heißt Abschluss von A.
$\mathbf{x}$ heißt Häufungspunkt von A, wenn
$\forall \varepsilon > 0 : |A \cap B_\varepsilon(\mathbf{x})| = \infty$
$\mathbf{x}$ heißt isolierter Punkt von A, wenn
$\exists \varepsilon > 0 : B_\varepsilon(\mathbf{x}) \cap A = \{\mathbf{x}\}$.

A heißt offen in X, wenn $A = A^\circ$.
A heißt abgeschlossen in X, wenn A^c offen ist.

Eigenschaften von offenen/abges. Mengen

Sei $(X, \|\cdot\|)$ ein normierter VR, $A \subseteq X$, $A_i \subseteq X$ offen $\forall i \in \mathbb{N}$ und $B_j \subseteq X$ abgeschlossen $\forall j \in \mathbb{N}$. Dann gilt:

1. A ist offen $\iff$ A^c ist abgeschlossen.
2. X und $\emptyset$ sind offen und abgeschlossen.
3. $\forall \mathbf{x} \in X : \{\mathbf{x}\}$ ist abgeschlossen.
4. $\bigcup_{i=1}^\infty A_i$ offen.
5. $\bigcap_{i=1}^n A_i$ offen.
6. $\bigcap_{j=1}^\infty B_j$ abgeschlossen.
7. $\bigcup_{j=1}^n B_j$ abgeschlossen.

Beispiel von offenen/abges. Mengen

1. Für $X = \mathbb{R}$ und $A = (0, 1]$ ist $A^\circ = (0, 1)$, $\partial A = \{0, 1\}$, $\bar{A} = [0, 1]$
2. Für $X = \mathbb{R}$ und $A = \mathbb{Q}$ ist $A^\circ = \emptyset$, $\partial A = \bar{A} = \mathbb{R}$

Konvergenz in normierten VR

Sei $(\mathbf{x}^k)_{k \in \mathbb{N}}$ eine Folge im normierten VR $(X, \|\cdot\|)$

1. $(\mathbf{x}^k)_{k \in \mathbb{N}}$ heißt beschränkt, wenn $\{\|\mathbf{x}^k\| | k \in \mathbb{N}\}$ in $\mathbb{R}$ beschränkt ist.
2. $(\mathbf{x}^k)_{k \in \mathbb{N}}$ heißt konvergent mit Limes $\mathbf{y} \in X$, wenn $\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall k > N : \|\mathbf{x}^k - \mathbf{y}\| < \varepsilon$
3. $(\mathbf{x}^k)_{k \in \mathbb{N}}$ heißt Cauchyfolge in X, wenn $\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall k, l > N : \|\mathbf{x}^k - \mathbf{x}^l\| < \varepsilon$
4. $\mathbf{y} \in X$ heißt Häufungspunkt von $(\mathbf{x}^k)_{k \in \mathbb{N}}$, wenn $\forall \varepsilon > 0 : |\{\mathbf{x}^k | k \in \mathbb{N}\} \cap B_\varepsilon(\mathbf{y})| = \infty$

Konvergenzsätze

Sei $(\mathbf{x}^k)_{k \in \mathbb{N}}$ eine Folge im normierten VR $(X, \|\cdot\|)$
1. $(\mathbf{x}^k)_{k \in \mathbb{N}}$ konvergiert $\implies$ $(\mathbf{x}^k)_{k \in \mathbb{N}}$ ist beschränkt.
2. $\mathbf{y} \in X$ ist Häufungspunkt von $(\mathbf{x}^k)_{k \in \mathbb{N}}$ $\iff$ $(\mathbf{x}^k)_{k \in \mathbb{N}}$ hat eine Teilfolge die gegen $\mathbf{y}$ konvergiert.
3. $(\mathbf{x}^k)_{k \in \mathbb{N}}$ konvergiert $\implies$ $(\mathbf{x}^k)_{k \in \mathbb{N}}$ ist Cauchyfolge.
4. Sei $(\mathbf{x}^k)_{k \in \mathbb{N}}$ eine Folge in $\mathbb{K}^n$:
$(\mathbf{x}^k)_{k \in \mathbb{N}}$ konvergiert $\iff$ $\forall j : (x_j^k)_{k \in \mathbb{N}}$ konvergiert.
$(\mathbf{x}^k)_{k \in \mathbb{N}}$ ist Cauchy $\iff$ $\forall j : (x_j^k)_{k \in \mathbb{N}}$ ist Cauchy.
5. $A \subseteq X$ ist abgeschlossen $\iff$ Der Grenzwert jeder Folge in A, die in X konvergiert, hat ihren Grenzwert in A.

Banachraum

Ein normierten VR $(X, \|\cdot\|)$ heißt vollständig, wenn jede Cauchyfolge in X konvergiert. Ein vollständiger normierter VR heißt Banachraum.

Banachraum Sätze

1. Sei $(\mathbf{x}^k)_{k \in \mathbb{N}}$ eine Folge im Banachraum $(X, \|\cdot\|)$.
$(\mathbf{x}^k)_{k \in \mathbb{N}}$ konvergiert $\iff$ $(\mathbf{x}^k)_{k \in \mathbb{N}}$ ist Cauchyfolge.
2. Ist $(X, \|\cdot\|)$ ein endlichdimensionaler normierter VR, dann ist X ein Banachraum.

Banachraum Beispiel

1. $(\mathbb{K}^n, \|\cdot\|_p)$ ist ein Banachraum.
2. Kein Banachraum
3. Bsp. 1.33 (b) im Skript ist kein Banachraum.

Ab jetzt bezeichnen wir die normierten Vektorräume $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ als X und Y.

Grenzwerte und Stetigkeit

Sei $A \subseteq X$ und $f : A \to Y$ eine Fkt. Ist $\mathbf{x}^0 \in X$ ein HP von A, dann sagen wir f hat den Grenzwert/Limes $\mathbf{y}^0 \in Y$ für $\mathbf{x} \to \mathbf{x}^0$, wenn $\forall \varepsilon > 0 \exists \delta > 0 \forall \mathbf{x} \in A \cap B_\delta(\mathbf{x}^0) \backslash \{\mathbf{x}^0\} : \|f(\mathbf{x}) - \mathbf{y}^0\| < \varepsilon$.

Ist $\mathbf{x}^0 \in A$, so heißt f stetig in $\mathbf{x}^0$, wenn $\forall \varepsilon > 0 \exists \delta > 0 \forall \mathbf{x} \in A \cap B_\delta(\mathbf{x}^0) : \|f(\mathbf{x}) - f(\mathbf{x}^0)\| < \varepsilon$.
f heißt stetig in A, wenn $\forall \mathbf{x}^0 \in A : f$ ist stetig in $\mathbf{x}^0$.

Folgenkriterium für Stetigkeit

Sei $A \subseteq X$ und $f : A \to Y$ eine Fkt. Ist $\mathbf{x}^0 \in X$ ein HP von A, dann gilt:

1. $\lim_{\mathbf{x} \to \mathbf{x}^0} f(\mathbf{x}) = \mathbf{y}^0 \iff \forall (x^k)_{k \in \mathbb{N}}$ in $A \backslash \{\mathbf{x}^0\}$: $\lim_{k \to \infty} \mathbf{x}^k = \mathbf{x}^0 \implies \lim_{k \to \infty} f(\mathbf{x}^k) = \mathbf{y}^0$
2. f ist in $\mathbf{x}^0$ stetig $\iff \forall (x^k)_{k \in \mathbb{N}}$ in A : $\lim_{k \to \infty} \mathbf{x}^k = \mathbf{x}^0 \implies \lim_{k \to \infty} f(\mathbf{x}^k) = f(\mathbf{x}^0)$

Koordinatenfunktionen und Projektionen

1. Die Funktion $\pi_j : \mathbb{K}^n \to \mathbb{K}, \pi_j(z_1, \ldots, z_n) := z_j$ heißt Projektion.
2. Für $A \subseteq X$ und die Fkt. $f : A \to \mathbb{K}^n$ heißen die Funktionen $f_j := f \circ \pi_j : A \to \mathbb{K}$ Koordinatenfunktionen von f.

Stetigkeit von Koord.fkt. und Proj.

1. Konstante Funktionen sind stetig.
2. Ist $(z^k)_{k \in \mathbb{N}}$ eine Folge in $\mathbb{K}^n$ mit $\lim_{k \to \infty} z^k = z^0$, so gilt $\forall j \in \{1, \ldots, n\} : \lim_{k \to \infty} z_j^k = z_j^0$ und daher sind alle Projektionen $\pi_j : \mathbb{K}^n \to \mathbb{K}, \pi_j(z_1, \ldots, z_n) := z_j$ stetig.
3. Für $A \subseteq X, \mathbf{x}^0 \in A$ und die Fkt. $f : A \to \mathbb{K}^n$ mit $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$ gilt: f ist stetig in $\mathbf{x}^0 \iff$ alle f_j sind stetig in $\mathbf{x}^0$.

Stetigkeit von Verknüpfungen

Seien $f, g : X \to Y$ Fkt. stetig in $\mathbf{x}^0$ und $\lambda \in \mathbb{K}$. Dann sind $f + g$, λf, $f \cdot g$, $\frac{f}{g}$ für $g \neq 0$, $|f|$, $\max\{g, f\}$, $\min\{g, f\}$, $f^+ = \max\{0, f\}$ und $f^- = -\min\{0, f\}$ stetig in $\mathbf{x}^0$ sofern überhaupt wohldefiniert.

Ist Z ein weiterer normierter VR und $h : Y \to Z$ eine Fkt. die stetig in $f(\mathbf{x}^0)$ ist. Dann ist $h \circ f : X \to Z$ stetig in $\mathbf{x}^0$, sofern überhaupt wohldefiniert.

Stetigkeit Beispiel

Sei $f : \mathbb{R}^+ \times \mathbb{C} \to \mathbb{C}, f(x, z) := x^z = \exp(z \log(x))$. f ist stetig, weil $f(x, z) = \exp(\pi_2(x, z) \log(\pi_1(x, z)))$ bzw. $f = \exp \circ (\pi_2 \cdot (\log \circ \pi_1))$ eine Verkettung bzw. Multiplikation stetiger Funktionen ist und somit laut dem Satz oben wieder stetig ist.

Stetigkeit bezüglich Koordinaten

Sei $f : \mathbb{K}^n \to Y$ und $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in \mathbb{K}^n$. Wir nennen f stetig in $\mathbf{x}^0$ bezüglich der j-ten Komponente x_j^0, wenn die Funktion $\phi : \mathbb{K} \to Y$ mit
$\phi(x) := f(x_1^0, \ldots, x_{j-1}^0, x, x_{j+1}^0, \ldots, x_n^0)$
stetig in $x = x_j^0$ ist.

Lemma: Ist f stetig in x^0, so ist f stetig in jeder Komponente von x^0. **!** Umkehrung gilt nicht im allgemeinen.

Stetigkeit bezüglich Koordinaten Beispiel

f kann in allen Komponenten von $\mathbf{x}^0$ stetig sein ohne in $\mathbf{x}^0$ stetig zu sein. Sei $\mathbf{x}^0 = (0, 0)$. Definiere $f : \mathbb{K}^2 \to \mathbb{K}$ mit
$f(x_1, x_2) = \begin{cases} 0 & x_1 x_2 = 0 \\ 1 & x_1 x_2 \neq 0 \end{cases}$, dann sind
$\phi_1(x_1) := f(x_1, 0), \phi_2(x_2) := f(0, x_2)$ beide konstant gleich 0 und somit stetig in $x_1 = 0$ bzw. $x_2 = 0$. Also ist f stetig in jeder Komponente von $\mathbf{x}^0 = (0, 0)$. Aber f ist nicht stetig in $\mathbf{x}^0$, weil z.B. für die Folge
$(x_1^k, x_2^k) = \begin{cases} (\frac{1}{k}, 0) & \text{für k gerade} \\ (\frac{1}{k}, \frac{1}{k}) & \text{für k ungerade} \end{cases}$ gilt, dass
$\lim_{k \to \infty} (x_1^k, x_2^k) = (0, 0)$, jedoch $f(x_1^k, x_2^k) = 1$ für alle ungeraden k. Also konvergiert $f(x_1^k, x_2^k)$ nicht gegen $f(0, 0) = 0$ und somit ist f nicht stetig in $\mathbf{x}^0 = (0, 0)$.

Monome/Polynome

Sei $n \in \mathbb{N}$ und $p = (p_1, \ldots, p_n) \in (\mathbb{N}_0)^n$.
p heißt Multiindex und $|p| := p_1 + \cdots + p_n$ heißt Grad von p.
Für $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{K}^n$ setze $x^p := x_1^{p_1} \cdot \ldots \cdot x_n^{p_n}$.
Die Funktion $f(\mathbf{x}) = x^p = x_1^{p_1} \cdot \ldots \cdot x_n^{p_n}$ heißt Monom vom Grad $|p|$.
Eine Linearkombination von Monomen heißt Polynom. D.h. P ist ein Polynom, wenn $P : \mathbb{K}^n \to \mathbb{K}$, mit $P(\mathbf{x}) = P((x_1, \ldots, x_n)) = \sum_{k \geq |p|} a_p x^p$ für $k \in \mathbb{N}_0$ und $a_p \in \mathbb{K}$.
Für zwei Polynome P, Q heißt $\frac{P}{Q}$ rationale Funktion.

Monome/Polynome Beispiel

		Grad
Monome	xy^2z^2	5
	x^2y^2	4
	x^2y	3
	y	1
	1	0
Polynome	$3xy^2z^2 + 4x^2y^2 - 2x^2y + 3y + 1$	5
	$(x + y)(x + z^2)$	3
	$0x^2 + 2y$	1

Skalarprodukt und induzierte Norm

Die Funktion $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{K}$ heißt inneres Produkt oder Skalarprodukt auf X, wenn gilt:

- $\forall \mathbf{x} \in X : \langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ und $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = 0$
- $\forall \mathbf{x}, \mathbf{y} \in X : \langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$
- $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in X, \lambda, \mu \in \mathbb{K}$: $\langle \lambda \mathbf{x} + \mu \mathbf{y}, \mathbf{z} \rangle = \lambda \langle \mathbf{x}, \mathbf{z} \rangle + \mu \langle \mathbf{y}, \mathbf{z} \rangle$

Ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf X, so definiert $\|\cdot\| : X \to \mathbb{R}_0^+$, mit $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ eine Norm auf X. Diese nennen wir die von $\langle \cdot, \cdot \rangle$ induzierte Norm.

Hilbertraum

Ist X ein $\mathbb{K}$-Vektorraum und $\langle \cdot, \cdot \rangle$ ein Skalarprodukt, so heißt $(X, \langle \cdot, \cdot \rangle)$ Prähilbertraum oder Innenproduktraum. Wenn die von $\langle \cdot, \cdot \rangle$ induzierte Norm $\|\cdot\|$ zusammen mit X ein Banachraum bildet (vollständige Norm), dann nennen wir $(X, \langle \cdot, \cdot \rangle)$ einen Hilbertraum.

Ab jetzt bezeichnet $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf dem normierten $\mathbb{K}$-Vektorraum X

Sätze für Skalarprodukte

Für $\langle\cdot,\cdot\rangle$ gilt:

- $\forall \mathbf{x},\mathbf{y},\mathbf{z} \in X, \lambda,\mu \in \mathbb{K}:$
  $\langle \mathbf{x}, \lambda\mathbf{y} + \mu\mathbf{z}\rangle = \overline{\lambda}\langle \mathbf{x},\mathbf{y}\rangle + \overline{\mu}\langle \mathbf{x},\mathbf{z}\rangle$
  Note: Für $\mathbb{K} = \mathbb{R}$ ist $\overline{\lambda} = \lambda$ und $\overline{\mu} = \mu$.
- $\forall \mathbf{x} \in X : \langle 0, \mathbf{x}\rangle = \langle \mathbf{x}, 0\rangle = 0$
- (Cauchy-Schwarz-Ungleichung) Für die von $\langle\cdot,\cdot\rangle$ induzierte Norm $\|\cdot\|$ gilt
  $\forall \mathbf{x},\mathbf{y} \in X : |\langle \mathbf{x},\mathbf{y}\rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$
  Gleichheit gilt g.d.w. $\mathbf{x}$ und $\mathbf{y}$ linear abhängig sind.

Äquivalente Normen

Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf X (beliebige Normen. Nicht verwechseln mit Notation für 1-Norm und 2-Norm).
Dann heißen $\|\cdot\|_1$ und $\|\cdot\|_2$ äquivalent, wenn
$\exists a,b \in \mathbb{R}^+ \forall \mathbf{x} \in X : a\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq b\|\mathbf{x}\|_1$.
In Worten: Normen sind äquivalent, wenn ihr Ergebnisse sich höchstens um einen konstanten Faktoren unterscheiden.

Sätze für äquivalente Normen

Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf X. Dann sind die beiden Aussagen äquivalent:

- $\|\cdot\|_1$ und $\|\cdot\|_2$ sind äquivalent.
- $\forall (\mathbf{x}^k)_{k\in\mathbb{N}}$ in X : $\lim\limits_{k\to\infty} \mathbf{x}^k = \mathbf{x}^0$ bzgl. $\|\cdot\|_1 \iff \lim\limits_{k\to\infty} \mathbf{x}^k = \mathbf{x}^0$ bzgl. $\|\cdot\|_2$

! Auf $\mathbb{K}^n$ sind alle Normen äquivalent. D.h. alle Normen erzeugen die genau die gleichen konvergenten Folgen.

Kapitel 2 - Differentialrechnung im $\mathbb{R}^n$

Sei im folgenden $A \subseteq \mathbb{R}^n$ und $\mathbf{x}^0 \in A$.

Partielle Ableitungen von $f : \mathbb{R}^n \to \mathbb{K}$

Sei $f : A \to \mathbb{K}$ und $j \in \{1, \dots, n\}$. Existiert ein $\varepsilon > 0$, s.d. $\forall h \in (-\varepsilon, \varepsilon) : \mathbf{x}^0 + h\mathbf{e}_j \in A$, so nennen wir f in $\mathbf{x}^0$ partiell differenzierbar nach x_j, wenn der Grenzwert $\lim\limits_{h \to 0} \frac{f(\mathbf{x}^0 + h\mathbf{e}_j) - f(\mathbf{x}^0)}{h}$ existiert.

Den Grenzwert nennen wir dann j-te partielle Ableitung von f in $\mathbf{x}^0$ und wird bezeichnet durch: $\partial_j f(\mathbf{x}^0)$, $\partial_{x_j} f(\mathbf{x}^0)$, $\frac{\partial f(\mathbf{x}^0)}{\partial x_j}$, $f_{x_j}(\mathbf{x}^0)$, $D_j f(\mathbf{x}^0)$

Randnotiz: Für $\mathbf{x}^0 \in \partial A$ definieren wir $\partial_j f(\mathbf{x}^0)$ so wie oben, bloss mit $h \in [0, \varepsilon)$ bzw. $h \in (\varepsilon, 0]$.

Existieren alle partiellen Ableitungen von f in $\mathbf{x}^0$, so nennen wir den (Zeilen)vektor $\nabla f(\mathbf{x}^0) = (\partial_1 f(\mathbf{x}^0), \dots, \partial_n f(\mathbf{x}^0))$ Gradient von f in $\mathbf{x}^0$.

Exisitert $\partial_j f(\mathbf{x}^0)$ für alle $\mathbf{x}^0 \in A$, so nennen wir die Funktion $\partial_j f : A \to \mathbb{K}, \mathbf{x} \mapsto \partial_j f(\mathbf{x})$ die j-te partielle Ableitung von f.

Partielle Ableitung berechnen

Um die $\partial_j f(\mathbf{x}^0)$ zu berechnen, falls diese existiert, tun wir so, als seien alle x_i mit $i \neq j$ konstant und berechnen die Ableitung der Funktion
$\phi(x) := f(x_1^0, \dots, x_{j-1}^0, x, x_{j+1}^0, \dots, x_n^0)$ in $x = x_j^0$ wie im 1-dimensionalen Fall.

Beispiel:
Sei $f : \mathbb{R}^2 \to \mathbb{R}, f(x_1, x_2) = x_1^2 x_2 e^{2x_1 + x_2} - 3x_2 + x_1$ und $\mathbf{x}^0 = (1, 0)$. Dann ist
$\partial_1 f(x_1, x_2) = 2x_1 x_2 e^{2x_1 + x_2} + 2x_1^2 x_2 e^{2x_1 + x_2} + 1$
$\partial_2 f(x_1, x_2) = x_1^2 e^{2x_1 + x_2} + x_1^2 x_2 e^{2x_1 + x_2} - 3$
$\nabla f(\mathbf{x}^0) = (1, e^2 - 3)$

Partielle Ableitungen von $f : \mathbb{R}^n \to \mathbb{K}^m$

Sei $f : A \to \mathbb{K}^m$. Existieren alle partiellen Ableitungen $\partial_j f_l(\mathbf{x}^0)$ für $j \in \{1, \dots, n\}, l \in \{1, \dots, m\}$, dann nennen wir die $m \times n$-Matrix

$$\begin{bmatrix} \partial_1 f_1(\mathbf{x}^0) & \dots & \partial_n f_1(\mathbf{x}^0) \\ \vdots & & \vdots \\ \partial_1 f_m(\mathbf{x}^0) & \dots & \partial_n f_m(\mathbf{x}^0) \end{bmatrix} = \begin{bmatrix} \nabla f_1(\mathbf{x}^0) \\ \vdots \\ \nabla f_m(\mathbf{x}^0) \end{bmatrix}$$

die Jacobimatrix von f in $\mathbf{x}^0$ und bezeichnen diese mit $J_f(\mathbf{x}^0)$. Für $m = n$ bezeichnen wir $\det(J_f(\mathbf{x}^0))$ als die Jacobideterminante von f in $\mathbf{x}^0$.

Eigenschaften von partiellen Ableitungen

Die j-te partielle Ableitung ist eine lineare Funktion. Somit sind auch der Gradient und die Jacobimatrix lineare Funktionen.
Das heißt $\forall \lambda \in \mathbb{K}, \forall l \in \{1, \dots, m\}, k \in \{1, \dots, n\}$:
$\partial_j(\lambda f + g)_l(\mathbf{x}^0) = \lambda \cdot \partial_j(f)_l(\mathbf{x}^0) + \partial_j(g)_l(\mathbf{x}^0)$
$\nabla(\lambda f + g)_l(\mathbf{x}^0) = \lambda \cdot \nabla f_l(\mathbf{x}^0) + \nabla g_l(\mathbf{x}^0)$
$J_{\lambda f + g}(\mathbf{x}^0) = \lambda \cdot J_f(\mathbf{x}^0) + J_g(\mathbf{x}^0)$

Jacobimatrix Beispiel

Die Funktion $f : \mathbb{R}^3 \to \mathbb{R}^2$ sei gegeben durch

$$f(x_1, x_2, x_3) = \begin{pmatrix} x_1^2 + x_2^2 \\ x_3^2 + x_3 \cdot \sin(x_2) \\ x_2 x_3 \end{pmatrix}$$

$$\partial_1 f(x_1, x_2, x_3) = \begin{pmatrix} 2x_1 \\ 0 \\ 0 \end{pmatrix}$$

$$\partial_2 f(x_1, x_2, x_3) = \begin{pmatrix} 2x_2^2 \\ x_3 \cdot \cos(x_2) \\ x_3 \end{pmatrix}$$

$$\partial_3 f(x_1, x_2, x_3) = \begin{pmatrix} 0 \\ 2x_3 + \sin(x_2) \\ x_2 \end{pmatrix}$$

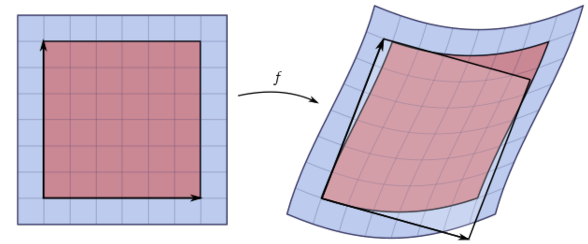
$$J_f(x_1, x_2, x_3) = \begin{pmatrix} 2x_1 & 2x_2^2 & 0 \\ 0 & x_3 \cdot \cos(x_2) & 2x_3 + \sin(x_2) \\ 0 & x_3 & x_2 \end{pmatrix}$$

$$J_f(1, 0, 1) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix}, \det(J_f(1, 0, 1)) = -4.$$

Partielle Ableitungen höherer Ordnung

Sei $f : A \to \mathbb{K}^m$ und $k \in \mathbb{N}$.
Für jedes $p = (p_1, \dots, p_k) \in \{1, \dots, n\}^k$ nennen wir
$\partial_p f(\mathbf{x}^0) := \frac{\partial^k f(\mathbf{x}^0)}{\partial x_{p_1} \dots \partial x_{p_k}} := \partial_{p_1} \partial_{p_2} \dots \partial_{p_k} f(\mathbf{x}^0)$
die partielle Ableitung k-ter Ordnung von f in $\mathbf{x}^0$ nach p (wenn existent). D.h. wir berechnen zuerst die p_k-te partielle Ableitung von f, dann davon die p_{k-1}-te partielle Ableitung, usw. und zum Schluss die p_1-te partielle Ableitung.
f selbst wird auch als partielle Ableitung 0-ter Ordnung bezeichnet.

Idee der Jacobideterminante



Eine nicht-lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ transformiert ein Rechteck (links in rot) zu einer verzerrten Fläche (rechts in rot). Würden wir für das rote Rechteck die Funktionswerte mit Hilfe der Jacobimatrix linear approximieren, dann würden wir das rote Rechteck in das weiße durchsichtige Parallelprogramm überführen. Die Jacobideterminante ist das Verhältnis von der Fläche des ursprünglichen Rechtecks zum weißen durchsichtigen Parallelprogramm und dient, in der Umgebung von dem Punkt in dem wir sie berechnet haben, als Approximation für das Verhältnis von der Fläche des ursprünglichen Rechtecks zur verzerrten roten Fläche.

Satz von Schwarz

Sei $f : A \to \mathbb{K}$. Sind alle partiellen Ableitungen 2-ter Ordnung von f in $\mathbf{x}^0$ stetig, dann gilt:
$\forall i, j \in \{1, \dots, n\} : \partial_i \partial_j f(\mathbf{x}^0) = \partial_j \partial_i f(\mathbf{x}^0)$

C^k-Funktionen

Sei $f : A \to \mathbb{K}^m$ und $k \in \mathbb{N}_0$. Wir nennen f eine C^k-Funktion, wenn für alle Koordinatenfunktionen f_l alle partiellen Ableitungen der Ordnung 0 bis k von f_l in A existieren und stetig sind.
Die Menge aller $\mathbb{K}^m$-wertigen C^k-Funktionen auf A bezeichnen wir mit $C^k(A, \mathbb{K}^m)$.
Haben alle f_l stetige partielle Ableitungen beliebiger Ordnung, so nennen wir f eine C^∞-Funktion oder glatte Funktion.

Totales Differential/ Totale Ableitung

Sei A offen und $f : A \to \mathbb{R}^m$ und $k \in \mathbb{N}_0$. Wir nennen f differenzierbar in $\mathbf{x}^0$, wenn eine lineare Abbildung $L_{\mathbf{x}^0} : \mathbb{R}^n \to \mathbb{R}^m$ existiert, so dass $\lim_{\mathbf{h}\to 0} \frac{f(\mathbf{x}^0+\mathbf{h})-f(\mathbf{x}^0)-L_{\mathbf{x}^0}(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0}$
Ist f differenzierbar in $\mathbf{x}^0$, so nennen wir L das totales Differential oder totale Ableitung von f in $\mathbf{x}^0$ und wir schreiben dann $Df(\mathbf{x}^0)$ statt L.

Idee/Intuition Totales Differential

Im eindimensionalen haben wir die Ableitung bestimmt, indem wir zwischen $f(x^0)$ und $f(x^0 + h)$ eine Linie/Gerade (lineare Funktion) zeichnen und davon die Steigung bestimmen. Die Ableitung ist dann diese Steigung für $h \to 0$. Im mehrdimensionalen reicht eine Gerade dafür nicht mehr aus, da es nun unendlich viele Punkte um $\mathbf{x}^0$ gibt, die Abstand $\|\mathbf{h}\|$ zu $\mathbf{x}^0$ haben. Das heißt wir brauchen eine Verallgemeinerung von "Linie"/"Gerade". Diese wäre im mehrdimensionalen eine lineare Abbildung L für die gilt, dass für $\mathbf{h} \approx \mathbf{0}$ gilt: $f(\mathbf{x}^0 + h) \approx f(\mathbf{x}^0) + L(\mathbf{h})$. Von der linearen Abbildung L können wir analog zu der Gerade im eindimensionalen die "Steigung"/Ableitung in alle Richtungen bestimmen. Wenn wir also den Grenzwert dazu bilden und dieser existiert, dann nennen wir diesen die totale Ableitung von f in $\mathbf{x}^0$.
Geometrisch bedeutet differenzierbar, so wie im eindimensionalen, dass f keine Sprünge und keine Knicke hat. Bei einer mehrdimensionalen Funktion kann man sich "Knick"als Falte oder Zacken vorstellen. Die Existenz der partiellen Ableitungen bedeutet nur, dass es keine Knicke in Richtung der Achsen gibt. Dennoch kann es in eine andere Richtung betrachtet Knicke geben. Deswegen folgt aus der Existenz der partiellen Ableitungen in $\mathbf{x}^0$ **nicht**, dass f in $\mathbf{x}^0$ differenzierbar ist.

Totales Differential Satz

Sei $f : A \to \mathbb{K}^m$ differenzierbar in $\mathbf{x}^0$. Dann ist f stetig in $\mathbf{x}^0$, alle $\partial_k f_l(\mathbf{x}^0)$ existieren und es gilt $Df(\mathbf{x}^0) = J_f(\mathbf{x}^0)$. Im Fall von $m = 1$ gilt dann also $Df(\mathbf{x}^0) = \nabla f(\mathbf{x}^0)$

Totales Differential Satz

Sei $f : A \to \mathbb{K}^m$. Existieren alle $\partial_j f_l$ auf ganz A und sind diese stetig in $\mathbf{x}^0$, so ist f differenzierbar in $\mathbf{x}^0$.

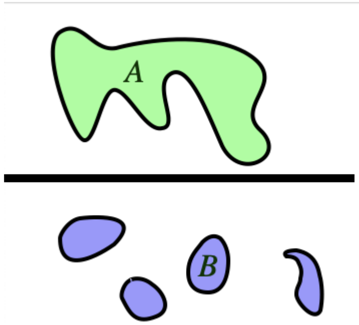
Kettenregel

Sei $A_f \subseteq \mathbb{R}^n$ offen und $f : A_f \to \mathbb{R}^m$.
Sei $f(A_f) \subseteq A_g \subseteq \mathbb{R}^m$ offen und $g : A_g \to \mathbb{K}^p$. Ist f differenzierbar in $\mathbf{x}^0$ und g differenzierbar in $f(\mathbf{x}^0)$, so ist $g \circ f : A_f \to \mathbb{K}^p$ differenzierbar in $\mathbf{x}^0$ und es gilt die Kettenregel:
$D(g \circ f)(\mathbf{x}^0) = Dg(f(\mathbf{x}^0)) \circ Df(\mathbf{x}^0)$
Insbesondere gilt, wenn f und g differenzierbar sind, so ist auch $g \circ f$ differenzierbar.

Weg-zusammenhängend

Wir nennen A weg-zusammenhängend, wenn für alle $\mathbf{x}, \mathbf{y} \in A$ eine stetige Funktion $p : [0, 1] \to A$ existiert mit $p(0) = x$ und $p(1) = y$.
Ist p auch noch differenzierbar, dann nennen wir A differenzierbar weg-zusammenhängend.

Weg-zusammenhängend



Die Menge A ist weg-zusammenhängend, weil du alle Punkte aus der Menge mit einem stetigen Pfad innerhalb der Menge verbinden kannst. Die Menge B ist nicht weg-zusammenhängend, da es keinen stetigen Pfad von einem Punkt aus der ersten "Bubble"zu einem Punkt in der letzten "Bubble"gibt.

Konstante Funktion

Ist A differenzierbar weg-zusammenhängend und $f : A \to \mathbb{K}$ differenzierbar mit $\nabla f \equiv 0$, dann ist f konstant.

Richtungsableitung

Sei $\mathbf{r} \in \mathbb{R}^n$ sei f stetig differenzierbar in $\mathbf{x}^0$, dann nennen wir $\frac{\partial f}{\partial \mathbf{r}}(\mathbf{x}^0) := \frac{df(\mathbf{x}^0+t\cdot\mathbf{r})}{dt}(0)$ (erste) Richtungsableitung von f in Richtung $\mathbf{r}$ an der Stelle $\mathbf{x}^0$.

Richtungsableitung Satz

Sei $f : A \to \mathbb{K}$ differenzierbar in $\mathbf{x}^0$, so existieren für alle $\mathbf{r} \in R^n$ die Richtungsableitung $\frac{\partial f}{\partial \mathbf{r}}(\mathbf{x}^0)$ und es gilt $\frac{\partial f}{\partial \mathbf{r}}(\mathbf{x}^0) = \nabla f(\mathbf{x}^0) \cdot \mathbf{r}$

Wenn der Vektor $\mathbf{r}$ zusätzlich normiert ist, d.h. $\|\mathbf{r}\| = 1$, dann ist $\frac{\nabla f(\mathbf{x}^0)}{\|\nabla f(\mathbf{x}^0)\|_2}$ die Richtung des steilsten Anstiegs an der Stelle $\mathbf{x}^0$.

Kapitel 3 - Extremwerte und stationäre Punkte

Sei im folgenden X ein normierter Vektorraum, $A \subseteq X$, $f : A \to \mathbb{R}$ eine Funktion und $\mathbf{x}^0 \in A$.

Globales/Lokales Minimum/Maximum

f hat in $\mathbf{x}^0$ ein (strenges) globales Maximum, wenn $\forall \mathbf{x} \in A \backslash \{\mathbf{x}^0\} : f(\mathbf{x}) \leq f(\mathbf{x}^0)$ (bzw. $f(\mathbf{x}) < f(\mathbf{x}^0)$)

f hat in $\mathbf{x}^0$ ein (strenges) globales Minimum, wenn $\forall \mathbf{x} \in A \backslash \{\mathbf{x}^0\} : f(\mathbf{x}) \geq f(\mathbf{x}^0)$ (bzw. $f(\mathbf{x}) > f(\mathbf{x}^0)$)

f hat in $\mathbf{x}^0$ ein (strenges) lokales Maximum, wenn $\exists \varepsilon > 0 \forall \mathbf{x} \in B_\varepsilon(\mathbf{x}^0) \cap A \backslash \{\mathbf{x}^0\} : f(\mathbf{x}) \leq f(\mathbf{x}^0)$ (bzw. $f(\mathbf{x}) < f(\mathbf{x}^0)$)

f hat in $\mathbf{x}^0$ ein (strenges) lokales Minimum, wenn $\exists \varepsilon > 0 \forall \mathbf{x} \in B_\varepsilon(\mathbf{x}^0) \cap A \backslash \{\mathbf{x}^0\} : f(\mathbf{x}) \geq f(\mathbf{x}^0)$ (bzw. $f(\mathbf{x}) > f(\mathbf{x}^0)$)

! Jedes globale Minimum/Maximum ist auch ein lokales Minimum/Maximum.

Min-Max Dualität

Wenn f in $\mathbf{x}^0$ ein lokales Minimum (Maximum) hat, dann hat $-f$ in $\mathbf{x}^0$ ein lokales Maximum (Minimum).

Kompakte Mengen

Wir nennen A kompakt, wenn jede Folge in A eine konvergente Teilfolge hat, deren Grenzwert in A liegt.

Kompakte Mengen Sätze

- Ist A kompakt, so ist A abgeschlossen und beschränkt.
- Ist A kompakt und $B \subseteq A$ abgeschlossen, so ist B kompakt.
- (Satz von Heine-Borel) Für $A \subseteq \mathbb{K}^n$ gilt: A ist kompakt $\iff$ A ist abgeschlossen und beschränkt
- Sei Y ein normierter VR, A kompakt und $f : A \to Y$ eine stetige Funktion, dann ist $f(A)$ kompakt in Y.
- Ist A kompakt und f stetig, so $\exists \mathbf{x}_m, \mathbf{x}_M \in A$, so dass f in $\mathbf{x}_m$ ein globales Minimum und in $\mathbf{x}_M$ ein globales Maximum hat.

Satz von Taylor in $\mathbb{R}$

Ist $I \subseteq \mathbb{R}$ ein offenes Intervall und $a, x \in I$ und $x \neq a$. Ist $f \in C^{m+1}(I, \mathbb{K})$ für $m \in \mathbb{N}_0$, so gilt: $f(x) = T_m(x, a) + R_m(x, a)$, wobei

$$T_m(x, a) := \sum_{k=0}^{m} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

und

$$R_m(x, a) := \int_a^x \frac{f^{(m+1)}(t)}{m!}(x-t)^m dt$$

Wir nennen $T_m(x, a)$ das m-te Taylorpolynom an der Stelle a und $R_m(x, a)$ das m-te Restglied (in Integralform)

Für $\mathbb{K} = \mathbb{R}$ gilt ausserdem: Es existiert ein θ im Intervall (x, a) bzw. (a, x), für das gilt:

$$R_m(x, a) = \frac{f^{(m+1)}(\theta)}{(m+1)!}(x-a)^{(m+1)}$$

Wir nennen das dann das m-te Restglied (in Lagrangeform)

Satz von Taylor in $\mathbb{R}^n$

Ist $A \subseteq \mathbb{R}^n$ offen und $\mathbf{x}^0, \mathbf{x} \in A$ und die Strecke zwischen $\mathbf{x}$ und $\mathbf{x}^0$ liegt komplett in A. Ist $f \in C^{m+1}(A, \mathbb{K})$ für $m \in \mathbb{N}_0$, so gilt: $f(\mathbf{x}) = T_m(\mathbf{x}, \mathbf{x}^0) + R_m(\mathbf{x}, \mathbf{x}^0)$, wobei

$$\begin{aligned}T_m(\mathbf{x}, \mathbf{x}^0) &:= f(\mathbf{x}^0) + \sum_{j=1}^{n} \partial_j \frac{1}{1!} f(\mathbf{x}^0)(x_j - x_j^0) \\ &+ \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{1}{2!}\partial_i\partial_j f(\mathbf{x}^0)(x_j - x_j^0)(x_i - x_i^0) + \dots \\ &+ \sum_{j_1,\dots,j_m=1}^{n} \frac{1}{m!}\partial_{j_1}\dots\partial_{j_m} f(\mathbf{x}^0)\prod_{k=1}^{m}(x_{j_k} - x_{j_k}^0)\end{aligned}$$

Formel für $R_m(\mathbf{x}, \mathbf{x}^0)$ wird hier weggelassen.
Wir nennen $T_m(\mathbf{x}, \mathbf{x}^0)$ das m-te Taylorpolynom an der Stelle $\mathbf{x}^0$

Es existiert ein θ auf der Verbindungsstrecke von $\mathbf{x}$ und $\mathbf{x}^0$ für das gilt:

$$R_m(\mathbf{x}, \mathbf{x}^0) = \sum_{j_1,\dots,j_{m+1}=1}^{n} \frac{1}{(m+1)!}\partial_{j_1}\dots\partial_{j_{m+1}} f(\theta)\prod_{k=1}^{m}(x_{j_k} - x_{j_k}^0)$$

Wir nennen das dann das n-te Restglied (in Lagrangeform)

Satz von Taylor für $n = 2$

Satz von Fermat Sind die Bedingungen aus dem Satz von Taylor alle erfüllt, so gilt für $n = 2$

$$\begin{aligned}T_2(\mathbf{x}, \mathbf{x}^0) :=& f(\mathbf{x}^0) + \nabla f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0) \\ &+ \frac{1}{2}(\mathbf{x} - \mathbf{x}^0)^\top H_f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0)\end{aligned}$$

wobei $\nabla f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0)$ als Skalarprodukt zu verstehen ist und $H_f(\mathbf{x}^0)$ die Hessematrix von f an der Stelle $\mathbf{x}^0$ ist:
$H_f(\mathbf{x}^0) = \begin{pmatrix} \partial_x\partial_x f(\mathbf{x}^0) & \partial_x\partial_y f(\mathbf{x}^0) \\ \partial_y\partial_x f(\mathbf{x}^0) & \partial_y\partial_y f(\mathbf{x}^0) \end{pmatrix}$

Sei im folgenden $A = (a_{ij})$ eine symmetrische $n \times n$-Matrix.

Quadratische Form

Eine quadratische Form ist eine Abbildung $Q_A : \mathbb{R}^n \to \mathbb{R}, Q_A(\mathbf{x}) := \mathbf{x}^\top A\mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j$

Frobenius/Hilbert-Schmid Norm

Für eine beliebige $(m \times n)$-Matrix A definieren wir die Frobenius-Norm als:

$$\|A\|_F := \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2} = \sqrt{trace(A^\top A)}$$

Es gilt $\forall \mathbf{x} \in \mathbb{R}^n : \|A\mathbf{x}\|_2 \leq \|A\|_F \cdot \|\mathbf{x}\|_2$

Quadratische Form Satz

- Quadratische Formen sind Polynome und somit stetig.
- Für $\lambda \in \mathbb{R}$ und $A, B \in \mathbb{R}^{n\times n}$ symmetrisch gilt $Q_{\lambda A+B} = \lambda Q_A + Q_B$.
- Q_A ist homogen vom Grad 2. D.h. $\forall\lambda \in \mathbb{R}, \forall\mathbf{x} \in \mathbb{R}^n : Q_A(\lambda\mathbf{x}) = \lambda^2 Q_A(\mathbf{x})$
- $\forall\lambda \in \mathbb{R}$ sind die folgenden beiden Aussagen äquivalent:
  1) $\forall\mathbf{x} \in \mathbb{R}^n : Q_A(\mathbf{x}) \geq \lambda\|\mathbf{x}\|_2$
  2) $\forall\mathbf{x} \in \mathbb{R}^n$ mit $\|\mathbf{x}\|_2 = 1 : Q_A(\mathbf{x}) \geq \lambda$
- $\forall\mathbf{x} \in \mathbb{R}^n : |Q_A(\mathbf{x})| \leq \|A\mathbf{x}\|_F \cdot \|\mathbf{x}\|_2^2$

Definitheit

- A und Q_A heißen positiv definit, wenn $\forall\mathbf{x} \neq \mathbf{0} \in \mathbb{R}^n : Q_A > 0$
- A und Q_A heißen positiv semidefinit, wenn $\forall\mathbf{x} \neq \mathbf{0} \in \mathbb{R}^n : Q_A \geq 0$
- A und Q_A heißen negativ definit, wenn $\forall\mathbf{x} \neq \mathbf{0} \in \mathbb{R}^n : Q_A < 0$
- A und Q_A heißen negativ semidefinit, wenn $\forall\mathbf{x} \neq \mathbf{0} \in \mathbb{R}^n : Q_A \leq 0$
- A und Q_A heißen indefinit, wenn sie weder positiv semidefinit, noch negativ semidefinit sind.

Definitheit Satz

- A ist positiv definit $\iff$ alle Eigenwerte von A sind positiv.
- A ist positiv semidefinit $\iff$ alle Eigenwerte von A sind positiv oder Null.
- A ist negativ definit $\iff$ alle Eigenwerte von A sind negativ.
- A ist negativ semidefinit $\iff$ alle Eigenwerte von A sind negativ oder Null.

Definitheit

Sei $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ und es ist
$Q_A : \mathbb{R}^2 \to \mathbb{R}, Q_A(x_1, x_2) := ax_1^2 + 2bx_1x_2 + cx_2^2$.
Es gilt:
$det(A) > 0$ und $a > 0 \iff A$ ist positiv definit
$det(A) > 0$ und $a < 0 \iff A$ ist negativ definit
$det(A) < 0 \iff A$ ist indefinit

Stationärer/Kritischer Punkt

Sei $A \subseteq \mathbb{R}^n$ offen, $f : A \to \mathbb{R}$ und existieren alle partiellen Ableitung von f in $\mathbf{x}^0$, so nennen wir $\mathbf{x}^0$ einen stationären oder kritischen Punkt, wenn $\nabla f(\mathbf{x}^0) = \mathbf{0}$

Satz von Fermat

Sei $A \subseteq \mathbb{R}^n$ offen, $f : A \to \mathbb{R}$ und existieren alle partiellen Ableitung von f in $\mathbf{x}^0$.
Hat f in $\mathbf{x}^0$ ein lokales Minimum oder ein lokales Maximum, so ist $\mathbf{x}^0$ eine stationäre Stelle von f.

Hessematrix

Sei $A \subseteq \mathbb{R}^n$ offen und exisitieren alle partiellen Ableitungen 2-ter Ordnung von f in $\mathbf{x}^0$, so definieren wir die Hessematrix von f an der Stelle $\mathbf{x}^0$ als:

$$H_f(\mathbf{x}^0) = \begin{pmatrix} \partial_x\partial_x f(\mathbf{x}^0) & \partial_x\partial_y f(\mathbf{x}^0) \\ \partial_y\partial_x f(\mathbf{x}^0) & \partial_y\partial_y f(\mathbf{x}^0) \end{pmatrix}$$

Satz für hinreichende Bedingung

Sei $A \subseteq \mathbb{R}^n$ offen, $f \in C^2(A)$ und sei $\mathbf{x}^0 \in A$ eine stationäre Stelle von f. Dann gilt:

- $H_f(\mathbf{x}^0)$ ist positiv definit $\implies f$ hat ein strenges lokales Minimum in $\mathbf{x}^0$.
- $H_f(\mathbf{x}^0)$ ist negativ definit $\implies f$ hat ein strenges lokales Maximum in $\mathbf{x}^0$.
- $H_f(\mathbf{x}^0)$ ist indefinit $\implies f$ hat kein lokales Extremum in $\mathbf{x}^0$ (In dem Fall nennen wir $\mathbf{x}^0$ einen Sattelpunkt von f).
- Ist $H_f(\mathbf{x}^0)$ negativ/positiv semidefinit, so kann keine Schlussfolgerung anhand der Hessematrix gemacht werden.

Zusammenfassung für das Vorgehen

Sei $A \subseteq \mathbb{R}^n$ offen, $f \in C^2(A)$.
1) Finde alle stationären Stellen von f, d.h. alle $\mathbf{x}$ mit $\nabla f(\mathbf{x}) = \mathbf{0}$. Alle lokalen Extrema von f müssen unter diesen bestimmten Stellen liegen (wenn A offen ist).
2) Berechne die Hessematrix von f an allen stationären Stellen $\mathbf{x}^0$. Benutze dann den "Satz für hinreichende Bedingung"um Schlussfolgerungen über diese Stellen zu machen.