Inhaltsverzeichnis

Einführung

0.1 Stetigkeit in einer Dimension

f ist stetig in x_0

$$\begin{aligned} &\Leftrightarrow \quad \lim_{x \to x_0} f(x) = f(x_0) \\ &\Leftrightarrow \quad \forall\,(x_n) \text{ mit } \lim_{n\to\infty} x_n = x_0 \text{ gilt } \lim_{n\to\infty} f(x_n) = f(x_0) \\ &\Leftrightarrow \quad \forall\, \varepsilon > 0 \quad \exists\, \delta \quad \text{mit} \quad |f(x) - f(x_0)| < \varepsilon \quad \forall\, x \in (x_0 - \delta, x_0 + \delta) \end{aligned}$$

Bemerkung: Der Grenzwert von Funktionen ist über den Grenzwert von Folgen definiert und kann auch nur so überprüft werden.

0.2 Zwei Sonderfälle

Skalarfeld

Sei $f : \mathbb{R}^2 \to \mathbb{R}$
Visualisierung durch Höhenlinien: $H_c := \{x \in \mathbb{R}^n : f(x) = c\}$
Beispiel: $f(x, y) = x^2 + y^2$

Vektorfeld

Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$
Beispiel: $f(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}$

Kapitel 1

Differentialrechnung in höheren Dimensionen

1.1 Topologie

Skalarprodukt

Definition: $\langle x, y\rangle := x^\top y = \sum_{k=1}^n x_k y_k$ für $x, y \in \mathbb{R}^n$

Euklidische Norm

Definition: $\|x\|_2 := \sqrt{\langle x, x\rangle} = \sqrt{\sum_{k=1}^n x_k^2}$

1.1.1 Korollar

Sei $x \in \mathbb{R}^n$ mit $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

1.

$$\max_{1\leqslant k\leqslant n} |x_k| \leqslant \|x\| \leqslant \sqrt{n} \max_{1\leqslant k\leqslant n} |x_k|$$

2. Cauchy-Schwarz-Ungleichung:

$$\forall\ x, y \in \mathbb{R}^n \quad : \quad |\langle x, y\rangle| \leqslant \|x\| \cdot \|y\|$$

Begründung (nicht Beweis!) durch alternative Definition: $\langle x, y\rangle = \|x\| \cdot \|y\| \underbrace{\cos\alpha}_{\leqslant 1}$

Dabei ist α der Winkel der zwischen x und y eingeschlossen wird.
Daraus folgt:

$$|\langle x, y\rangle| = \|x\| \cdot \|y\| \Leftrightarrow x, y \text{ sind lin. unabhängig} : x = \lambda y \text{ oder } y = \lambda x \text{ für } \lambda \in \mathbb{R}$$

3. $\|\cdot\|$ ist eine Norm. Eine Norm hat folgende Eigenschaften:

   (i) $\|x\| \geqslant 0$ und $\|x\| = 0 \Leftrightarrow x = 0$

   (ii) $\|\lambda x\| = |\lambda| \cdot \|x\|$

   (iii) $\|x + y\| \leqslant \|x\| + \|y\|$ Dreiecksungleichung

1.1.2 Konvention

Für $A \subset \mathbb{R}^n$ gilt für das Komplement $A^c = \mathbb{R}^n \setminus A$

1.1.3 Definition der ε-Umgebung

Sei $x_0 \in \mathbb{R}^n$ und $\varepsilon > 0$, dann gilt für die ε-Umgebung $U_\varepsilon(x_0)$ von x_0:

$$U_\varepsilon(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| < \varepsilon\}$$

Bemerkung: Die punktierte ε-Umgebung ist definiert als: $\dot{U}_\varepsilon = U_\varepsilon(a) \setminus \{a\}$

1.1.4 Topologische Grundbegriffe

Sei $A \subset \mathbb{R}^n$, dann heißt ein Punkt $x_0 \in \mathbb{R}^n$

(i) ein **innerer Punkt**, wenn gilt $\exists\ \varepsilon > 0$ mit $U_\varepsilon(x_0) \subset A$
Menge aller inneren Punkte: $\mathring{A} = \{x \in \mathbb{R}^n : \exists\ \varepsilon > 0 \text{ mit } U_\varepsilon(x) \subset A\}$

(ii) ein **Berührungspunkt**, wenn $\forall\ \varepsilon > 0$ gilt $U_\varepsilon(x_0) \cap A \neq \varnothing$
abgeschlossene Hülle: $\overline{A} = \{x \in \mathbb{R}^n : \forall\ \varepsilon > 0 \text{ gilt } U_\varepsilon(x_0) \neq \varnothing\}$

(iii) ein **Häufungspunkt**, wenn $\forall\ \varepsilon > 0$ gilt $(U_\varepsilon(x_0) \setminus \{x_0\}) \cap A \neq \varnothing$
Die Menge aller Häufungspunkte wird mit A' bezeichnet.

(iv) ein **Randpunkt**, wenn $\forall\ \varepsilon > 0$ gilt $U_\varepsilon(x_0) \cap A \neq \varnothing$ und $U_\varepsilon(x_0) \cap A^c \neq \varnothing$
Menge aller Randpunkte oder auch **Rand** von A wird mit ∂A bezeichnet.

Korollar

(i) $\mathring{A} \subset A$

(ii) $\mathring{A} \subset \overline{A}$

(iii) $\partial A \subset \overline{A}$

(iv) $\overline{A} = \mathring{A} \cup \partial A$

(v) $\overline{A} = A \cup \partial A$ (schwächere Aussage als (iv))

1.1.5 Definition von offen und abgeschlossen

Eine Menge $A \subset \mathbb{R}^n$ heißt

(i) **offen**, wenn $A = \mathring{A}$ gilt (A besteht nur aus inneren Punkten)

(ii) **abgeschlossen**, wenn $\partial A \subset A$ gilt (wenn der Rand in der Menge enthalten ist)

1.1.6 Beispiele

1. Jede ε-Umgebung $U_\varepsilon(x_0 \in \mathbb{R}^n)$ ist offen

2. Sei $I \subset \mathbb{R}$, dann gilt

   (i) I ist offen, wenn $I = (a, b)$ mit $-\infty \leqslant a \leqslant b \leqslant \infty$
   für $a = b$ gilt $I = \varnothing$ mit I offen
   und für $a = -\infty, b = \infty$ ist I auch offen

(ii) I ist abgeschlossen, wenn $I = [a, b]$ mit $a, b \in \mathbb{R}$
oder $I = (-\infty, b]$ oder $I = [a, \infty)$ oder $I = (-\infty, \infty) = \mathbb{R}$

(die reellen Zahlen sind offen und abgeschlossen zugleich)

1.1.7 Satz

für $A \subset \mathbb{R}^n$ sind folgenden Aussagen äquivalent:

(i) A ist abgeschlossen $A = \overline{A}$

(ii) A enthält alle Häufungspunkte, $A' \subset A$

(iii) A enthält alle Randpunkte, $\partial A \subset A$

(iv) A^c ist offen

1.1.8 Satz

(i) $\varnothing$ und $\mathbb{R}^n$ sind offen.

(ii) Die Vereinigung beliebig vieler offene Mengen ist offen:

$$\bigcup_{j \in J} (O_j \text{ offen}) = O \text{ offen}$$

(iii) Der Durchschnitt <u>endlich</u> vieler offener Mengen ist offen:

$$\bigcap_{j=1}^{n} (O_j \text{ offen}) = O \text{ offen}$$

Bemerkung: Für unendlich viele offene Mengen gilt dies nicht immer:

$$\bigcap_{k=1}^{\infty} \left(-\frac{1}{k}, \frac{1}{k}\right) = (-1, 1) \cap \left(-\frac{1}{2}, \frac{1}{2}\right) \cap \left(-\frac{1}{3}, \frac{1}{3}\right) \cap ... = \{0\} \text{ abgeschlossen}$$

Beispiel

Seien A_1, A_2 zwei abgeschlossene Mengen, dann gilt

(i) $A_1 \cup A_2$ ist abgeschlossen
Beweisidee: A_1 ist abgeschlossen $\Rightarrow A_1^c$ ist offen

$$(A_1 \cup A_2)^c \overset{\text{De Morgan}}{=} \underbrace{A_1^c}_{\text{offen}} \cap \underbrace{A_2^c}_{\text{offen}} \text{ ist offen wegen Satz 1.1.8}$$
$$((A_1 \cup A_2)^c)^c = A_1 \cup A_2 \text{ ist abgeschlossen}$$

1.1.9 Satz

(i) $\varnothing$ und $\mathbb{R}^n$ sind abgeschlossen.

(ii) Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen:

$$\bigcap_{j \in J} (A_j \text{ abgeschlossen}) = A \text{ abgeschlossen}$$

(iii) Die Vereinigung endlich vieler abgeschlossenen Mengen ist abgeschlossen:

$$\bigcup_{j=1}^{n} (A_j \text{ abgeschlossen}) = A \text{ abgeschlossen}$$

Bemerkung: Für unendlich viele abgeschlossene Mengen gilt dies nicht immer:

$$\bigcup_{k=1}^{\infty} \left[-1 + \frac{1}{n}, 1 - \frac{1}{n}\right] = \{0\} \cup \left[-\frac{1}{2}, \frac{1}{2}\right] \cup \left[-\frac{2}{3}, \frac{2}{3}\right] \cup ... = (-1, 1) \text{ offen}$$

1.1.10 Definition von beschränkt und kompakt

Eine Menge $A \subset \mathbb{R}^n$ heißt:

(i) **beschränkt** wenn $\exists\ c > 0$ mit $\|x\| < c \quad \forall\ x \in A$

(ii) **kompakt**, wenn A abgeschlossen und beschränkt ist.

1.2 Folgen

1.2.1 Definition von Konvergenz und Beschränktheit

Eine Folge $(a_k)_{k=1}^{\infty}$ heißt

(i) **konvergent**, wenn gilt

$$\exists\ a \in \mathbb{R}^n \quad \text{mit} \quad \forall\ \varepsilon > 0 \quad \exists\ N(\varepsilon): \quad \|a_k - a\| \quad \forall\ k \geqslant N(\varepsilon)$$

Dann ist a der Grenzwert der Folge:

$$a = \lim_{k \to \infty} a_k \quad \text{oder} \quad a_k \overset{k \to \infty}{\to} a$$

(ii) **beschränkt**, wenn $\exists\ c > 0$ mit $\|a_k\| < c \quad \forall\ k$

1.2.2 Bemerkung

Wenn eine Folge $(a_k) = \begin{pmatrix} \left(a_1^{(k)}\right) \\ \vdots \\ \left(a_n^{(k)}\right) \end{pmatrix} \in \mathbb{R}^n$ konvergiert, so gilt

(i) $\Leftrightarrow$ jede Komponente $\left(a_1^{(k)}\right), ..., \left(a_n^{(k)}\right)$ konvergiert:

$$\lim_{k \to \infty} a_k = a \quad \Leftrightarrow \quad \lim_{k \to \infty} a_i^{(k)} = a_i \quad \text{für } i = 1, ..., n$$

(ii) $\Leftrightarrow (a_k)$ erfüllt das **Cauchy-Kriterium**:

$$\forall\, \varepsilon > 0 \quad \exists\, N(\varepsilon): \quad \|a_k - a_l\| < \varepsilon \quad \forall\, k, l \geqslant N(\varepsilon)$$

(iii) $\Leftrightarrow$ jede Teilfolge von (a_k) konvergiert gegen a: $a_{l_k} \overset{k\to\infty}{\to} a$ für $l_1 \geqslant 1, l_2 \geqslant 2, ...$

(iv) der Grenzwert a ist eindeutig.

1.2.3 Satz von Bolzano Weierstraß

Jede beschränkte Folge im $\mathbb{R}^n$ besitzt einen konvergente Teilfolge.

Beispiele

(i) $n = 1$: Sei $A \leqslant (a_k) \leqslant B \quad \forall\, k$. Konstruiert man eine neue Schranke mit $\frac{A+B}{2}$ so liegen wiederum ∞ viele Elemente in der oberen und/oder unteren Hälfte.

(ii) Sei $(a_k) = \begin{pmatrix} (x_k) \\ (y_k) \end{pmatrix}$ eine beschränkte Folge im $\mathbb{R}^2$
$\Rightarrow (x_k), (y_k)$ sind beschränkte Folgen
$\overset{\text{Satz von Bolzano Wierstraß}}{\Rightarrow} \exists\ (x_k), (y_k)$ sind konvergent

1.2.4 Abschließende Bemerkungen

(i) Grenzwert Rechenregeln können aus dem $\mathbb{R}$ für $\mathbb{R}^n$ übernommen werden.
z.b. $a_k \overset{k\to\infty}{\to} a, \quad b_k \overset{k\to\infty}{\to} b \quad \Rightarrow \quad a_k^\top b_k \overset{k\to\infty}{\to} a^\top b$

(ii) Es gibt viele Zusammenhänge zwischen den Eigenschaften von Folgen und den topologischen Eigenschaften von Mengen.
z.b. Sei $A \subset \mathbb{R}^n$ und $a \in \mathbb{R}^n$ ein Häufungspunkt
$\Leftrightarrow \quad \exists (a_k)_{k=1}^\infty$ mit $a_k \in A \setminus \{a\}\, \forall\, k \quad$ und $\quad a_k \overset{k\to\infty}{\to} a$

1.3 Funktionsgrenzwerte und Stetigkeit

1.3.1 Definition

Eine Funktion $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ nennt man eine Funktion mit n-Veränderlichen.

$$f(x_1, ..., x_n) = f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = \begin{pmatrix} f_1(x_1, ..., x_n) \\ \vdots \\ f_m(x_1, ..., x_n) \end{pmatrix} \quad \text{mit} \quad f_1, ..., f_m : \mathbb{R}^n \to \mathbb{R}$$

1.3.2 Definition Grenzwert/Limes

Sei $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ und $a \in \overline{A}$. Ein $b \in \mathbb{R}^m$ heißt Grenzwert von f für $x \to a$, wenn gilt:

$$\forall \varepsilon > 0 \quad \exists\ \delta(\varepsilon) > 0 : \quad \|f(x) - b\| < \varepsilon \quad \forall\ x \in \dot{U}_{\delta(\varepsilon)}(a) \cap A$$

Bemerkung: Die Funktion f muss in a nicht stetig sein, so kann z.b. gelten:
$\lim_{x\to a} f(x) = b \neq f(a)$

1.3.3 Bemerkung

Sei $f : A \subset \mathbb{R}^n \to \mathbb{R}^m, a \in \overline{A}, b \in \mathbb{R}^m$ dann sind folgende Aussagen äquivalent:

(i) $f(x) \overset{x\to a}{b}$

(ii) $\|f(x) - b\| \overset{x\to a}{0} \in \mathbb{R}^1$ (Eine Norm bildet immer auf ein Skalar ab)

(iii) $f_1(x) \overset{x\to a}{\to} b_1, ..., f_m(x) \overset{x\to a}{\to} b_m$

Zusätzlich gilt das **Cauchy-Kriterium**:

$$\lim_{x\to a} f(x) = b \quad \Leftrightarrow \quad \forall\, \varepsilon > 0\ \exists\ \delta(\varepsilon) > 0 : \quad \|f(x), f(y)\| < \varepsilon \quad \forall\ x, y \in \dot{U}_{\delta(\varepsilon)}(a) \cap A$$

1.3.4 Beispiel

Sei $f(x, y) = \frac{xy}{x^2+y^2}$

$$a_k = \begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k} \end{pmatrix}, \quad f(a_k) = \frac{\frac{1}{k^2}}{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{1}{2} \quad \forall\ k$$

$$b_k = \begin{pmatrix} x_k \\ 0 \end{pmatrix} \text{ mit } x_k \overset{k\to\infty}{\to} 0, \quad f(b_k) = \frac{0}{x_k^2} \quad \forall\ k$$

Da $\lim\limits_{k\to\infty} f(a_k) = \frac{1}{2} \neq 0 = \lim\limits_{k\to\infty} f(b_k)$ kann der Grenzwert nicht existieren.

1.3.5 Lemma Folgenkriterium

Sei $f : A \subset \mathbb{R}^n \to \mathbb{R}^m, a \in \overline{A}$

$$\underbrace{\exists b \in \mathbb{R}^m \text{ mit } \lim_{x\to a} f(x) = b}_{\text{der Grenzwert } b \text{ existiert}} \quad \Leftrightarrow \quad \underbrace{\begin{array}{l}\text{jede Folge } (x_k)_{k=1}^{\infty} \subset A \text{ mit } x_k \neq a\ \forall\ k \text{ und } x_k \overset{k\to\infty}{\to} a \\ \Rightarrow \quad f(x_k) \overset{k\to\infty}{\to} b\end{array}}_{\text{jede beliebige Folge konvergiert gegen } b}$$

1.3.6 Satz zu Grenzwerte verketteter Funktionen

Sei $A \subset \mathbb{R}^n, B \subset \mathbb{R}^m, a \in \overline{A}, f : A \to B, g : \overline{B} \to \mathbb{R}^l$

$$\exists\ b \in \overline{B} \text{ mit } \lim_{x\to a} f(x) = b, \quad \exists\ c \in \mathbb{R}^l \text{ mit } \lim_{y\to b} g(y) = c \quad \Rightarrow \quad \lim_{x\to a} \underbrace{g\left(f(x)\right)}_{(g\circ f)(x)} = \lim_{y\to b} g(y) = c$$

1.3.7 Beispiel

Sei $f(x, y) = e^{-x^2+y^2} = \exp\left(g(x, y)\right)$ mit $g(x, y) = x^2 + y^2$, dabei gilt:

$$\lim_{(x,y)^\top \to (0,0)^\top} g(x, y) = \lim_{(x,y)^\top \to (0,0)^\top} x^2 + y^2 = 0 \quad \Rightarrow \quad \lim_{z\to 0} f(z) = \lim_{z\to 0} e^z = 1$$

1.3.8 Definition der Stetigkeit

Sei $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$

(i) f ist **stetig** in $a \in A$ wenn gilt:

$$\forall\, \varepsilon > 0\ \exists \delta(\varepsilon):\quad \|f(x) - f(a)\| < \varepsilon \quad \forall\, x \in U_{\delta(\varepsilon)}(a) \cap A$$

Bemerkung: Es wird $\lim_{x \to a} f(x) = f(a)$ gefordert.
Diese Definition unterscheidet sich in der nicht punktierten ε-Umgebung und es gilt $f(a)$ anstatt b.

(ii) f ist stetig auf A, wenn f in jedem Punkt $a \in A$ stetig ist.

1.3.9 Bemerkung

(i) Kompositionen stetiger Funktionen sind wieder stetig: f, g stetig $\Rightarrow$ $f+g, f-g, ...$ stetig

(ii) Das Folgenkriterium überträgt sich:
Sei $(a_k)_{k=1}^{\infty}$ eine Folge in A mit $\lim_{k\to\infty} a_k = a \quad \Leftrightarrow \quad \lim_{k\to\infty} f(a_k) = f(a)$

(iii) Ist A kompakt, dann nimmt eine stetige Funktion $f : A \to \mathbb{R}$ immer ein Maximum und Minimum an:

$$\exists\ x_m, x_M \in A \text{ mit } f(x_m) = \min_{x\in A} f(x), f(x_M) = \max_{x \in A} f(x)$$

1.4 Partielle Ableitungen, Richtungsableitungen

1.4.1 Definition der partiellen Ableitung

Die Funktion $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ heißt **partielle differenzierbar** in $a \in A$ nach der k-ten Variable x_k mit $k \in \{1, ..., n\}$ wenn der folgender Grenzwert existiert:

$$\frac{\partial}{\partial x_k} f(a) = f_{x_k}(a) = \lim_{h \to 0} \frac{f(a + h \cdot e_k) - f(a)}{h}$$

Existieren alle partielle Ableitungen $f_{x_1}(a), ..., f_{x_n}(a)$, dann ist der **Gradient** von f wie folgt definiert:

$$\nabla f(a) = \begin{pmatrix} f_{x_1}(a) \\ \vdots \\ f_{x_n}(a) \end{pmatrix}$$

und die Funktion f heißt mindestens einmal partielle differenzierbar. Sind die partiellen Ableitungen $f_{x_1}(a), ..., f_{x_n}(a)$ zudem stetig, so heißt f einmal stetig differenzierbar: $f \in C^1(A, \mathbb{R}^m)$ oder kurz $f \in C^1(A)$.

1.4.2 Beispiel

Sei $f(x,y,z) = x^2 - xy + 3z$

$$\begin{aligned}
\frac{\partial}{\partial x} f(x,y,z) &= \lim_{h\to 0} \frac{f(x+h,y,z) - f(x,y,z)}{h} \\
&= \lim_{h\to 0} \frac{(x+h)^2 - (x+h)y + 3z - (x^2 - xy + 3z)}{h} \\
&= \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h} - \frac{(x+h)y - xy}{h} + \frac{3z - 3z}{h} \\
&= \left(\frac{d}{dx} x^2\right) - \left(\frac{d}{dx} x\right) y + \left(\frac{d}{dx} 0\right) z \\
&= 2x - y + 0 \\
&\Rightarrow \nabla f(x,y,z) = \begin{pmatrix} 2x - y \\ -x \\ 3 \end{pmatrix}
\end{aligned}$$

1.4.3 Definition der Richtungsableitung

Sei $a, r \in \mathbb{R}^n$ mit $\|r\| = 1$ (normiert), $f : \mathbb{R}^n \to \mathbb{R}^m$, dann heißt der folgende Grenzwert die Richtungsableitung von f bei a in Richtung r:

$$\frac{\partial}{\partial r} f(a) = f_r(a) = \lim_{h\to 0} \frac{f(a + h \cdot r) - f(a)}{h}$$

Bemerkung

(i) Ist $r = e_k$, dann erhalten wir gerade eine partielle Ableitung.

(ii) Es gibt Funktionen die in a in <u>jede Richtung differenzierbar</u> sind, aber in a <u>nicht stetig</u> sind!

1.5 Total Differenzierbarkeit

Idee: Differenzierbare Funktionen sind lokal im Punkt x_0 linear approximierbar:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \underbrace{r(x)\|x - x_0\|}_{\tilde{r}(x)}$$

Dabei muss der Fehler $\tilde{r}(x) = r(x)\|x - x_0\|$ *schneller gegen Null gehen als x gegen x_0* also muss $\tilde{r}(x) = \mathrm{o}(x - x_0)$ gelten (Landau-Notation: klein-oh).

1.5.1 Definition der totalen Differenzierbarkeit

Sei $f : A \subset \mathbb{R}^n \to \mathbb{R}^m, A$ offen, $x_0 \in A$

(i) Die Funktion f nennt man **total differenzierbar** bei x_0, wenn eine Matrix $A \in \mathbb{R}^{m\times n}$ existiert, mit der sich die Funktion f in einer ε-Umgebung um x_0 mittels einer Hyperebene approximieren lässt:

$$f(x) = f(x_0) + A(x - x_0) + r(x)\|x - x_0\|$$

Dann nennt man die Matrix $A = f'(x_0) = \frac{\partial}{\partial x} f(x_0)$ die total Ableitung von f in x_0.

(ii) Ist $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$ partiell diff'bar, so nennt man die Ableitung **Jacobi-Matrix**:

$$f'(x_0) = \frac{\partial}{\partial x} f(x_0) = J_f(x_0) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x_0) & \dots & \frac{\partial}{\partial x_n} f_1(x_0) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} f_m(x_0) & \dots & \frac{\partial}{\partial x_n} f_m(x_0) \end{pmatrix} \in \mathbb{R}^{m \times n}$$

Bemerkung: Es gilt: $\exists f'(x_0) \quad \Rightarrow \quad f'(x_0) = J_f(x_0)$, nicht aber die Gegenrichtung! Es kann also sein, dass die Jacobi-Matrix J_f existiert die Funktion aber nicht total diff'bar ist.

1.5.2 Beispiele

(i)

$$f(r, \varphi) = r \cdot \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} \quad \Rightarrow \quad J_f = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

(ii) $f(x) = a + b^\top(x - x_0), \quad f : \mathbb{R}^n \to \mathbb{R}, \quad a \in \mathbb{R}, \quad b, x_0 \in \mathbb{R}^n$
$\Rightarrow \quad f(x_0) = a, \quad f'(x_0) = b^\top$

(iii) $f(x) = a + A(x - x_0), \quad f : \mathbb{R}^n \to \mathbb{R}^m, \quad a \in \mathbb{R}^m, \quad A \in \mathbb{R}^{m \times n}, \quad x_0 \in \mathbb{R}^n$
$\Rightarrow \quad f(x_0) = a, \quad f'(x_0) = A$

Bemerkung: Beispiel (ii) und (iii) sind lineare Funktionen.

1.5.3 Satz

Ist $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ in jedem Punkt $x_0 \in A$ total differenzierbar, so ist f stetig in A.
Beweis:

$$f(x) = \underbrace{f(x_0)}_{\overset{x\to x_0}{\to} f(x_0)} + \underbrace{A\,\underbrace{(x - x_0)}_{\overset{x\to x_0}{\to} 0 \in \mathbb{R}^n}}_{\overset{x\to x_0}{\to} 0 \in \mathbb{R}^n} + \underbrace{r(x)}_{\overset{x\to x_0}{\to} 0 \in \mathbb{R}^m} \underbrace{\|x - x_0\|}_{\overset{x\to x_0}{\to} 0 \in \mathbb{R}} \qquad \text{mit } r(x) \overset{x\to x_0}{\to} 0$$

$\lim_{x \to a} f(x) = f(x_0) \quad \square$

1.5.4 Satz

Sei $f : A \subset \mathbb{R}^n \to \mathbb{R}^m, x_0 \in A$

a. Ist f total differenzierbar in x_0, so gilt

   (i) $f'(x_0) = J_f(x_0)$

   (ii) f ist in jede Richtung r differenzierbar mit: $\frac{\partial}{\partial r} f(x_0) = J_f(x_0) \cdot r$

b. Existieren in x_0 alle partiellen Ableitungen (also alle Komponenten der Jacobi-Matrix) und diese stetig sind $\quad \Rightarrow \quad f$ ist in x_0 total differenzierbar.

Beweis zu a: ...
Beweis zu b: ...

1.5.5 Bemerkung

Sei r eine Richtung mit $\|r\| = 1$ und $x = x_0 + r$, dann gilt:

$$\begin{aligned} f(x) &\approx f(x_0) + \nabla f(x_0)^\top \cdot r \\ &\Rightarrow \text{1. Fall}: \quad r, \nabla f(x_0) \text{ zeigen in dieselbe Richtung}: \\ &\quad f(x) - f(x_0) \approx \|\nabla f(x_0)\| \|r\| = \|\nabla f(x_0)\| > 0 \\ &\Rightarrow \text{2. Fall}: \quad r, \nabla f(x_0) \text{ zeigen in entgegengesetzte Richtungen}: \\ &\quad f(x) - f(x_0) \approx -\|\nabla f(x_0)\| < 0 \end{aligned}$$

In allen Fällen gilt Näherungsweise:

$$-\|\nabla f(x_0)\| < \nabla f(x_0)^\top r \leqslant \|\nabla f(x_0)\|$$

Fazit: Beim Reinzoomen sind die Höhenlinien parallel. Der Gradient zeigt in Richtung des steilsten Anstieges.

1.5.6 Satz zur Kettenregel

Ist $f : A \subset \mathbb{R}^n \to B \subset \mathbb{R}^m$ differenzierbar in $a \in A$ und $g : B \subset \mathbb{R}^m \to \mathbb{R}^l$ differenzierbar in $b \in B$, so gilt:

$$(g \circ f)'(a) = g'\left(\underbrace{f(a)}_{=b}\right) f'(a) = \underbrace{J_g(b)}_{\in \mathbb{R}^{l \times m}} \underbrace{J_f(a)}_{\in \mathbb{R}^{m \times n}}$$

1.6 Lokale Extremstellen und Mittelwertsätze

In einer Dimension gilt:

1. Mittelwertsatz

Ist f differenzierbar auf (a, b) und stetig auf $[a, b]$, so gilt:

$$\exists\, \xi \in (a, b) \text{ mit } f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

Satz von Rolle

Ist f differenzierbar auf (a, b) und stetig auf $[a, b]$ und gilt $f(a) = f(b)$, so gilt:

$$\exists\, \xi \in (a, b) \text{ mit } f'(\xi) = 0$$

1.6.1 Definition lokale/globale Extremstellen

(i) Eine Funktion $f : A \subset \mathbb{R}^n \to \mathbb{R}$ (Skalarfeld) hat bei $x_0 \in A$ ein lokales Minimum (Maximum) wenn in einer Umgebung $U = U_\varepsilon(x_0) \cap A$ für $\varepsilon > 0$ (offen bezüglich A) von x_0 gilt:

$$f(x_0) \overset{(\geqslant)}{\leqslant} f(x) \quad \forall x \in U$$

Ist bei x_0 ein lokales Minimum (Maximum) dann nennt man x_0 eine lokale Extremstelle.

(ii) f besitzt in x_0 ein globales Minimum (Maximum), wenn gilt:

$$f(x_0) \overset{(\geqslant)}{\leqslant} f(x) \quad \forall x \in A$$

1.6.2 Satz zur notwendigen Bedingung für eine lokale Extremstelle

Besitzt $f : \mathring{A} \subset \mathbb{R}^n \to \mathbb{R}$ bei $x_0 \in A$ eine lokale Extremstelle und f ist partiell differenzierbar, dann ist

$$\nabla f(x_0) = 0$$

Bemerkung: Der Rand ist ausgeschlossen da $\mathring{A}$ (alle inneren Punkte) gilt.
Es gilt:

$$x_0 \text{ ist eine lokale Extremstelle} \quad \overset{\not\Leftarrow}{\Rightarrow} \quad f'(x_0) = 0$$

Aus $f'(x_0)$ folgt nicht direkt die Extremstelle, denn Sattelpunkte sind keine Extremstellen.

Beweis ...

1.6.3 Definition des kritischen Punktes

Ein $x_0 \in \mathbb{R}^n$ mit $\nabla f(x_0) = 0$ heißt **kritischer** oder stationärer Punkt.

1.6.4 Mittelwertsatz

Sei $f : G \subset \mathbb{R}^n \to \mathbb{R}$ differenzierbar und sei G offen und enthalte die Menge $\overline{a, b} = \{a, b \in G \text{ mit } a + t(b - a) : t \in [0, 1]\}$ (a, b können durch eine Gerade verbunden werden). Dann:

$$\exists\ \xi \in (0, 1) \quad \text{mit} \quad f(b) = f(a) + \nabla f(a + \xi(b - a))^\top (b - a)$$

Bemerkung:

$$h(t) = a + t(b - a) \quad g(t) = f(h(t)) \quad \text{(differenzierbar)}$$
$$\Rightarrow \quad \exists\ \xi \in (0, 1) \quad \text{mit} \quad g'(\xi) = \frac{g(1) - g(0)}{1 - 0}$$

Beweis: ...

1.6.5 Definition eines Gebiets

(i) Ein Menge, die wie folgt konstruiert werden kann, heißt **Polygonzug**:

$$\overline{a_0, ..., a_k} = \bigcup_{j=1}^{k} \overline{a_{j-1}, a_j} \quad \text{mit} \quad a_0, ..., a_k \in \mathbb{R}^n$$

(ii) Eine Menge $M \subset \mathbb{R}^n$ heißt **kurvenweise zusammenhängend** wenn zu beliebigen $a, b \in M$ eine stetige Funktion $\gamma : [0, 1] \to M$ mit $\gamma(0) = a, \gamma(1) = b$ existiert.

(iii) Eine Menge $G \subset \mathbb{R}^n$ heißt **Gebiet**, wenn G offen und kurvenweise zusammenhängend ist.

1.6.6 Bemerkungen zu Gebieten

(i) Ein Gebiet G entspricht einem offenen Intervall $(a, b) \subset \mathbb{R}$ im Eindimensionalen: Der Rand ist nicht dabei, es hat keine Inseln.

(ii) Man kann zeigen, dass es reicht, wenn $a, b \in G$ mit einem Polygonzug verbunden werden kann.

1.6.7 Satz

Sei $G \subset \mathbb{R}^n$ ein Gebiet, $G \neq \varnothing$, und $f : G \to \mathbb{R}$ differenzierbar, dann gilt:

$$f(x) = \text{konst.} \quad \Leftrightarrow \quad \nabla f(x) = 0 \quad \forall\, x \in G$$

Beweis: ...

1.6.8 Definition partieller Ableitungen r'ter Ordnung

Für $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ definiert man (wenn diese auch existieren) induktiv für $x_0 \in A$ und $k_1, ..., k_r \in \{1, ..., n\}$ die partiellen Ableitungen r'ter Ordnung als:

$$\frac{\partial^n}{\partial x_{k_1} \dots \partial x_{k_r}} f(x_0) = f_{x_{k_1} \dots x_{k_r}} = \begin{cases} f(x_0) & r = 0 \\ \frac{\partial}{\partial x_{k_1}} f(x_0) & r = 1 \\ \frac{\partial}{\partial x_{k_1}} \left(\frac{\partial^{r-1}}{\partial x_{k_2} \dots \partial x_{k_r}} f(x_0) \right) & r > 1 \end{cases}$$

Existieren alle Ableitungen r'ter Ordnung und sind diese zudem stetig, so nennt man die Funktion f r-mal stetig differenzierbar: $f \in C^r(A; \mathbb{R}^m)$.

1.6.9 Definition der Hessematrix

Ist $f : A \subset \mathbb{R}^n \to \mathbb{R}$ 2-mal stetig differenzierbar bei $x_0 \in A$, dann ist die **Hessematrix** wie folgt definiert:

$$H_f(x_0) = \begin{pmatrix} f_{x_1, x_1}(x_0) & \dots & f_{x_1, x_n}(x_0) \\ \vdots & & \vdots \\ f_{x_m, x_1}(x_0) & \dots & f_{x_m, x_n}(x_0) \end{pmatrix}$$

1.6.10 Beispiele

(i)

$$f(x, y) = 2xy^3 + y \log x$$

(ii)

$$f(x, y) = \begin{cases} xy\frac{x^2 - y^2}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

$\Rightarrow$ Hessematrix ist nicht symmetrisch.

(iii) $A \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n, c \in \mathbb{R}, Q : \mathbb{R}^n \to \mathbb{R}$

$$\begin{aligned} Q(x) &= x^\top A x + b^\top x + c \\ \nabla Q(x) &= \left(A + A^\top\right) x + b \quad \overbrace{(= 2Ax + b)}^{\text{wenn } A \text{ sym.}} \\ H_Q(x) &= A + A^\top \qquad \underbrace{(= 2A)}_{\text{wenn } A \text{ sym.}} \end{aligned}$$

Q wird eine quadratische Funktion genannt.

1.6.11 Satz von Schwarz

Ist $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$ in x_0 2-mal stetig partielle differenzierbar ($f \in C^2(A)$), dann ist $H_f(x_0) \quad \forall\, x_0 \in A$ symmetrisch und es gilt:

$$f_{x_l,x_k}(x_0) = \frac{\partial}{\partial x_l \partial x_k} f(x_0) = \frac{\partial}{\partial x_k \partial x_l} f(x_0) = f_{x_k,x_l}(x_0) \quad \forall\, l, k \in \{1, ..., k\}$$

1.6.12 Satz von Taylor mit quadratischem Restglied

Seien $a, b \in G \subset \mathbb{R}^n, f \in C^2(G), G$ ein Gebiet, dann:

$$\exists\, \xi \in \overline{a, b} \quad \text{mit} \quad f(b) = f(a) + \nabla f(a)^\top (b - a) + \frac{1}{2}(b - a)^\top H_f(\xi)(b - a)$$

Beweis: Definiere $g(t) = f(h(t))$ mit $h(t) = a + t(b-a) \quad \Rightarrow \quad g(0) = f(a), g(1) = f(b)$. Bei Funktionen mit einem Skalar, gilt der eindimensionale Taylor mit einer Zwischenstelle $z \in [0, 1]$:

$$\begin{aligned} \Rightarrow \quad \underbrace{g(1)}_{f(b)} = \underbrace{g(0)}_{f(a)} + & \underbrace{g'(0)(1-0)}_{\substack{g'(t)=\frac{d}{dt}f(h(t)) \\ =f'(h(t))h'(t) \\ =\nabla f(a+t(b-a))^\top (b-a) \\ =(b-a)^\top \nabla f(a+t(b-a))}} + \underbrace{\frac{1}{2} g''(z)(1-0)^2}_{g''(t)=...=(b-a)^\top H_f(a+t(b-a))(b-a)} \\ &= f(a) + \nabla f(a)^\top (b - a) + \frac{1}{2}(b - a)^\top H_f(\underbrace{\xi}_{\xi=a+t(b-a)})(b - a) \end{aligned}$$