## Konvergenz

Eine Funktion $f : U \to \mathbb{R}^m$ heißt stetig in $a \in U$, wenn für alle folgen $x_n \in U$ mit $x_n \to a$ die Folge $f(x_n)$ gegen $f(a)$ konvergiert. Dies ist gleichbedeutend damit, dass für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $d(f(x), L_a) < \epsilon$ gilt für jedes $x$ mit $d(x, a) < \delta$.

**Stetigkeit**

Eine Funktion $f : U \to \mathbb{R}^m$ heißt stetig, wenn sie für alle $a \in U$ stetig ist.

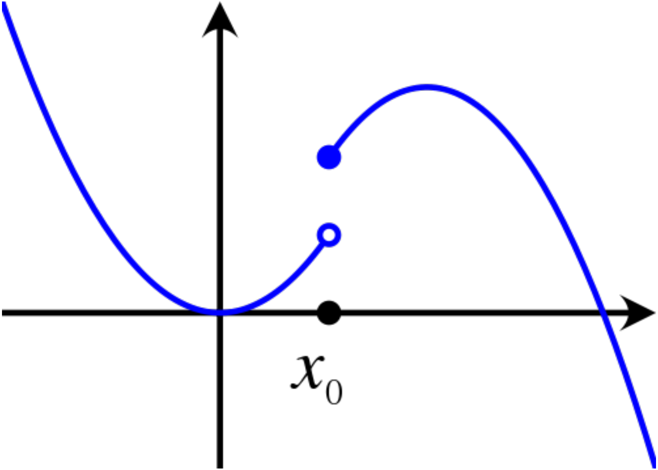



Figure: Quelle: Wikipedia: https://commons.wikimedia.org/wiki/File:Upper_semi.svg

## Landau Notation

Für eine Funktion $g : U \subset \mathbb{R}^n \to \mathbb{R}$ bezeichnen wir die Wachstumsklasse

$$o(g) := \{f : U \to \mathbb{R} \mid \lim_{x \to a} \frac{|f(x)|}{|g(x)|} = 0 \ \forall a \in U\}$$

$$O(g) := \{f : U \to \mathbb{R} \mid \lim_{x \to a} \frac{|f(x)|}{|g(x)|} < \infty \ \forall a \in U\}$$

## Lokale Linearisierung

Eine Funktion $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ heisst differenzierbar in $a \in U$ falls es eine lineare Funktion $df(a) : \mathbb{R}^n \to \mathbb{R}^m$ gibt mit

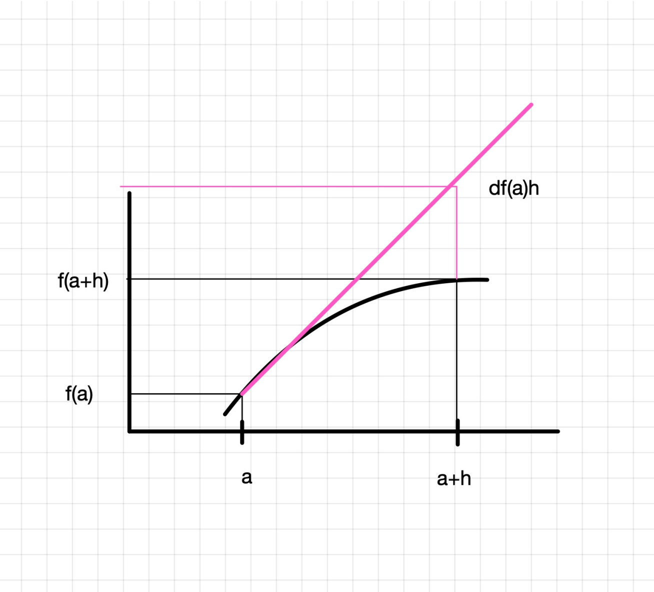
$$f(a+h) = f(a) + df(a)h + o(||h||) \tag{1}$$

$$\Leftrightarrow lim_{h \to 0} \frac{f(a+h) - f(a) - df(a)h}{||h||} = 0 \tag{2}$$

für alle $h \in \mathbb{R}^n$

## Bedeutung

Eine differenzierbare Funktion kann auf hinreichend kleinen Umgebungen beliebig genau durch eine lineare Funktion approximiert werden.


df(a)h
f(a+h)
f(a)
a
a+h

## Eindeutigkeit

Die lineare Abbildung $df(a)$ ist eindeutig bestimmt.

## Beweis

Ist $df'(a)$ eine weiter Abbildung mit Eigenschaft (1), so gilt für jeden Basisvektor $e_i$

$$\lim_{t \to 0} \frac{f(a+te_i) - f(a) - df(a)te_i}{||te_i||} = 0 \tag{3}$$

$$\lim_{t \to 0} \frac{f(a+te_i) - f(a) - df'(a)te_i}{||te_i||} = 0 \tag{4}$$

$$\Rightarrow (df(a) - df'(a))(e_i) = lim_{t \to 0} \frac{(df'(a) - df(a))(te_i)}{||te_i||} = 0 \tag{5}$$

## Beispiel

$A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$

$$f(x) := A \cdot x + b \tag{6}$$

$$df(a) := A \tag{7}$$

## Beweis

$$\lim_{h \to 0} \frac{A(x+h) - A \cdot x - A \cdot h}{||h||} = \tag{8}$$

$$\lim_{h \to 0} \frac{A \cdot x + A \cdot h - A \cdot x - A \cdot h}{||h||} = 0 \tag{9}$$

**Ab Jetzt**

Der Fall $m = 1$. Wir betrachten also Funktionen $f : U \subset \mathbb{R}^n \to \mathbb{R}$

## Ableitung

Ist $f$ differenzierter in $U$, so gilt wegen der Linearität

$$df(a)h = \sum_{i=1}^{n} (df(a)e_i) \cdot h_i \tag{10}$$

wobei $(e_1, \cdots, e_n)$ die Standardbasis des $\mathbb{R}^n$ ist. Die einzeilige Matrix

$$f'(a) := (df(a)e_1, \cdots, df(a)e_n) \tag{11}$$

heißt Ableitung von $f$ in $a$.

### Richtungsableitung

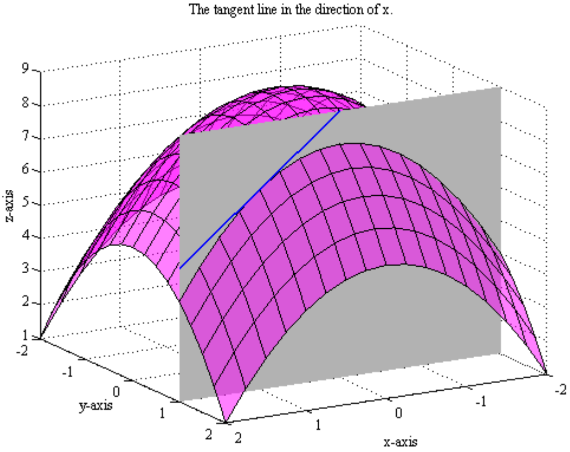
Sei $f : U \to \mathbb{R}$ eine Funktion. Für einen Vektor $h \in \mathbb{R}^n$ und einen Punkt $a \in U$ heißt der Grenzwert (falls er existiert)

$$\partial_h f(a) := \lim_{t \to 0} \frac{f(a+th) - f(a)}{t}$$

Richtungsableitung von $f$ am Punkt $a$ in Richtung $h$. Sie misst die Änderung der Funktion in Richtung $h$.
Speziell nennen wir für die Standard Basisvektoren $e_i$

$$\frac{\partial f(a)}{\partial x_i} := \partial_{e_i} f(a) := \lim_{t \to 0} \frac{f(a+te_i) - f(a)}{t}$$

die partielle Ableitung von $f$ in $a$ nach $x_i$.



The tangent line in the direction of x.
z-axis
y-axis
x-axis

### Partielle Differenzierbarkeit

Eine Funktion $f : U \to \mathbb{R}$ heißt partiell differenzierbar im Punkt $a \in U$, falls alle partiellen Ableitungen

$$\frac{\partial f(a)}{\partial x_1}, \dots, \frac{\partial f(a)}{\partial x_n}$$

existieren.

**Beispiel**

## Beispiel

Ist eine Funktion $f$ in $a$ differenzierbar, so ist sie dort partiell differenzierbar und es gilt

$$df(a)h = f'(a)h = \sum_{i=1}^{n} \partial_i f(a) \cdot h_i$$
$$f'(a) = \left(\partial_1 f(a), \cdots, \partial_n f(a)\right)$$

### Beweis

Ist $f$ differenzierter, so gilt für $t \in \mathbb{R}$ und $h \in \mathbb{R}^n$

$$f(a+th) = f(a) + df(a)th + R(||th||)$$

$$\lim_{th \to 0} \frac{||R(th)||}{||th||} = 0$$

$$\Rightarrow \lim_{t \to 0} \frac{||R(th)||}{||th||} = 0 \Rightarrow \lim_{t \to 0} \frac{||R(th)||}{|t|} = 0$$

$$\Rightarrow df(a)h = \lim_{t \to 0} \frac{f(a+th) - f(a)}{|t|} = \partial_h f(a)$$

$$\Rightarrow df(a)e_i = \partial_i f(a)$$

## Differenzierbarkeis Kriterium

Eine Funktion $f : U \to \mathbb{R}$ ist differenzierbar im Punkt $a \in U$, falls alle partiellen Ableitungen

$$\frac{\partial f(a)}{\partial x_1}, \dots, \frac{\partial f(a)}{\partial x_n}$$

existieren und stetig sind.

## Mittelwertsatz einer Veränderlichen

Sei $f : [a, b] \to \mathbb{R}$ stetig und differenzierbar für alle $x \in (a, b)$. Dann gibt es $\xi \in (a, b)$ mit $f'(\xi) = \frac{f(b)-f(a)}{b-a}$.

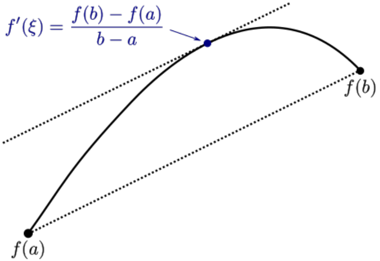



Figure: Quelle: Wikipedia: https://commons.wikimedia.org/wiki/File:Mittelwertsatz3.svg

## Beweis

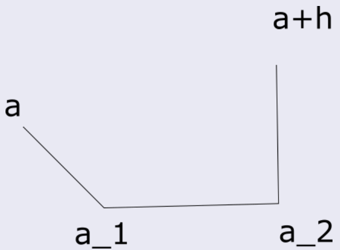



Figure: Kantenzug mit achesenparallelen Kanten

## Beweis

$$a_0 := a$$
$$a_i := a_{i-1} + h_i e_i; \quad i = 1, \cdots, n$$

## Beweis

- $f(a+h) - f(a) = \sum_{i=1}^{n} \big( f(a_i) - f(a_{i-1}) \big)$

## Beweis

- $f(a+h)-f(a)=\sum_{i=1}^{n}\big(f(a_i)-f(a_{i-1})\big)$
- Mit $\varphi_i(t):=f(a_i+te_i)$ gilt $f(a_i)-f(a_{i-1})=\varphi_i(h_i)-\varphi_i(0)$

## Beweis

- $f(a+h) - f(a) = \sum_{i=1}^{n} \big(f(a_i) - f(a_{i-1})\big)$
- Mit $\varphi_i(t) := f(a_i + te_i)$ gilt $f(a_i) - f(a_{i-1}) = \varphi_i(h_i) - \varphi_i(0)$
- Mittelwertsatz einer Veränderlichen: Es gibt $\tau_i$ mit

$$\varphi_i(h_i) - \varphi_i(0) = h_i \varphi'(\tau_i) \; .$$

## Beweis

- $f(a+h) - f(a) = \sum_{i=1}^{n} \left(f(a_i) - f(a_{i-1})\right)$
- Mit $\varphi_i(t) := f(a_i + te_i)$ gilt $f(a_i) - f(a_{i-1}) = \varphi_i(h_i) - \varphi_i(0)$
- Mittelwertsatz einer Veränderlichen: Es gibt $\tau_i$ mit

$$\varphi_i(h_i) - \varphi_i(0) = h_i \varphi'(\tau_i) \ .$$

- 

$$f(a+h) - f(a) - df(a) \cdot h = \sum_{i=1}^{n} \left( \frac{\partial f(\xi)}{\partial x_i} - \frac{\partial f(a)}{\partial x_i} \right) h_i$$

Da $\varphi_i'(t) = \frac{\partial f(a_{i-1} + te_i)}{\partial x_i}$ und mit $\xi_i := a_i + \tau_i e_i$

**Beweis**

$$|f(a+h)-f(a)-df(a)\cdot h| \leq ||h||_\infty \sum_{i=1}^{n} \left| \frac{\partial f(\xi)}{\partial x_i} - \frac{\partial f(a)}{\partial x_i} \right| .$$

Für $h \to 0$ gilt $\xi_i \to a$ und da die partiellen Ableitung stetig sind nach Voraussetzung und alle Normen äquivalent sind folgt

$$\lim_{h \to 0} \frac{f(a+h)-f(a)-df(a)\cdot h}{||h||} = 0$$

## Eigenschaften des Differentials

Für das Differential einer differenzierbaren Funktion $f : U \to \mathbb{R}$ gilt für alle $a \in U$:

- $df(a) \cdot h = \partial_h f(a)$.
- $d(f \cdot g)(a) = gdf(a) + f(a)dg$
- $d(f + g)(a) = df(a) + dg(a)$

**Beweis**

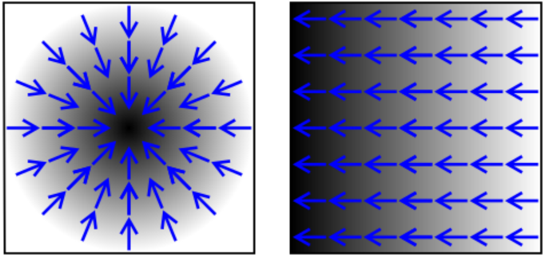
- Für die Basisvektoren ist per Definition $df(a) \cdot e_i = \partial_{e_i} f(a)$. Da jeder Vektor $h$ eine Linearkombination der Basisvektoren ist und $df$ linear ist, folgt die Behauptung.
- Folgt direkt aus der entsprechenden Eigenschaft reeller Funktionen.
- Folgt direkt aus der entsprechenden Eigenschaft reeller Funktionen.

## Gradient

Der Vektor

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

wird als Gradient bezeichnet. Es ist $df(a) \cdot h = \langle \nabla f(a), h \rangle$.



**Gradient**

Sei $f : U \to \mathbb{R}$ differenzierbare Funktion, $a \in U$ und $v := \operatorname{argmax}_{||h||=1}\{\partial_h f(a)\}$. Dann gilt

$$||\nabla f(a)||v = \nabla f(a) \, .$$

**Gradient**

Der Gradient zeigt in die Richtung des steilsten Anstiegs.

### Beweis

Für beliebiges $h$ gilt

$$\partial_h f(a) = df(a)h = \langle \nabla f(a), h \rangle = ||\nabla f(a)|| \cdot ||h|| \cdot \cos(\varphi)$$

wobei $\varphi$ den Innenwinkel zwischen $\nabla f(a)$ und $h$ bezeichnet. Für $||h|| = 1$ wird somit $\partial_h f(a)$ maximal, wenn $\varphi = 0$ und somit$h = \frac{\nabla f(a)}{||\nabla f(a)||}$ ist.