

0.1 Totale Differenzierbarkeit

Erinnerung (Analysis 1) $f: D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}$, ist genau dann in $x \in D$ differenzierbar, falls f in x "gutlinear approximierbar" ist, d.h. $\exists a \in \mathbb{R}$ mit $f(x+h) = f(x) + a \cdot h + w(h)$ wobei $\lim_{h \rightarrow 0} \frac{w(h)}{|h|} = 0$ ($f'(x) = a$).

Definition 0.1 (total differenzierbar). Es sei $D \subset \mathbb{R}^n$ offen und $f: D \rightarrow \mathbb{R}^m$ eine Abbildung. f heißt im Punkt $x \in D$ **(total) differenzierbar**, falls es eine **lineare Abbildung** $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ gibt, sodass

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(x+h) - f(x) - A \cdot h}{\|h\|} = 0 \quad (1)$$

Oft wird (1) durch eine Bedingung an den Rest $w(h)$, $w: D \rightarrow \mathbb{R}^m$ definiert $f(x+h) = f(x) + A \cdot h + w(h)$, wobei $\lim_{h \rightarrow 0} \frac{\|w(h)\|}{\|h\|} = 0$ ($\Leftrightarrow$ (1), $\Leftrightarrow w(h) = o(\|h\|)$). Da alle Normen auf $\mathbb{R}^m$ äquivalent sind, ist es gleichgültig, welche Norm man in (1) verwendet. A heißt das **Differential** von f im Punkt x . Schreibweise:

$$df(x), \left. df \right|_x, df_x, Df(x), \left. Df \right|_x, Df_x, \left. df(x) \right|_{x=x_0}, Df(x_0).$$

Bemerkung 0.2. Für $n = m = 1$ ist die Definition der totalen Ableitung äquivalent zur Definition der Ableitung von Funktionen einer Variablen.

Satz 0.3 (Differenzierbarkeit). Sei $D \subset \mathbb{R}^n$ offen. Für Funktionen $f: D \rightarrow \mathbb{R}^m$ gilt:

- 1) Sei f in $x \in D$ differenzierbar, dann ist f partiell differenzierbar und $Df(x) = J_f(x)$, $J_f(x)$ Jacobi-Matrix
- 2) Sei f partiell differenzierbar in einer Umgebung von $x \in D$ und die partiellen Ableitungen stetig in x , dann ist f differenzierbar in x .

Beweis. $n = 2$ und $m = 1$.

- 1) Sei f differenzierbar. Dann gilt

$$\begin{aligned} \lim_{h_i \rightarrow 0} \frac{f(x + h_i \cdot e^{(i)}) - f(x)}{h_i} &= \lim_{h_i \rightarrow 0} \left(Df(x) \cdot e^{(i)} + \frac{w(h_i)}{h_i} \right) \\ &= Df(x) \cdot e^{(i)} + \underbrace{\lim_{h_i \rightarrow 0} \frac{w(h_i)}{h_i}}_{\rightarrow 0} \\ &= Df(x) \cdot e^{(i)} \end{aligned}$$

$$\Rightarrow f \text{ partiell differenzierbar und } Df(x)e^{(i)} = \begin{pmatrix} \partial_i f_1 \\ \vdots \\ \partial_i f_m \end{pmatrix} \Rightarrow Df(x) = J_f(x).$$

- 2) Sei f stetig partiell differenzierbar und $h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$. Dann gilt

$$f(x+h) - f(x) = f(x_1 + h_1, x_2 + h_2) - f(x_1 + h_1, x_2) + f(x_1 + h_1, x_2) - f(x_1, x_2)$$

$\exists \theta_1, \theta_2 \in (0, 1)$ mit

$$\begin{aligned} f(x+h) - f(x) &\stackrel{\text{MWS}}{=} h_2 \partial_2 f(x_1 + h_1, x_2 + \theta_2 \cdot h_2) + h_1 \partial_1 f(x_1 + \theta_1 \cdot h_1, x_2) \\ &= h_2 (\partial_2 f(x_1, x_2) + w_2(h_1, h_2)) + h_1 (\partial_1 f(x_1, x_2) + w_1(h_1, h_2)), \end{aligned}$$

wobei

$$w_1(h_1, h_2) = \partial_1 f(x_1 + \theta_1 h_1, x_2) - \partial_1 f(x_1, x_2)$$

und

$$w_2(h_1, h_2) = \partial_2 f(x_1 + h_1, x_2 + \theta_2 h_2) - \partial_2 f(x_1, x_2).$$

$\partial_1 f(x), \partial_2 f(x)$ stetig $\implies \lim_{h \rightarrow 0} w_1(h_1, h_2) = 0, \lim_{h \rightarrow 0} w_2(h_1, h_2) = 0$. Daher gilt

$$\begin{aligned} f(x+h) - f(x) &= h_1 \partial_1 f(x) + h_2 \partial_2 f(x) + h_1 w_1(h) + h_2 w_2(h) \\ &= (\partial f_1(x), \partial f_2(x)) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + (w_1(h), w_2(h)) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\ &= Df(x) \cdot h + \tilde{w}(h) \end{aligned}$$

mit $\lim_{h \rightarrow 0} \frac{\|\tilde{w}(h)\|}{\|h\|} = 0$. Also ist f differenzierbar und $Df(x) = \nabla^T f(x)$.

□

Korollar 0.4. stetig partiell differenzierbar $\implies$ (total) differenzierbar $\implies$ partiell differenzierbar. Die umgekehrten Implikationen gelten im Allgemeinen nicht.

Lemma 0.5 (Richtungsableitung). Sei $D \in \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}$ im Punkt $x \in D$ differenzierbar. Dann gilt $\forall v \in \mathbb{R}^n$ mit $\|v\|_2 = 1$ existiert die Ableitung in Richtung v (sog. **Richtungsableitung**)

$$\frac{\partial f}{\partial v}(x) := \lim_{t \searrow 0} \frac{f(x + tv) - f(x)}{t}$$

und

$$\frac{\partial f}{\partial v}(x) = (\nabla f(x), v)_2$$

Beweis. Sei $x \in D$ und definiere die Funktion $\xi(t) := x + tv$. $\xi(t) \in D$ für $t \in [0, \varepsilon)$ für genügend kleine $\varepsilon > 0$. Betrachte die Komposition $h := f \circ \xi: [0, \varepsilon) \rightarrow \mathbb{R}$. Dann gilt

$$\begin{aligned} \frac{\partial f}{\partial v}(x) &\stackrel{\text{Def. Richtungsabl.}}{=} \lim_{t \searrow 0} \frac{f(x + tv) - f(x)}{t} \\ &\stackrel{\text{Def. Abl.}}{=} \left. \frac{d}{dt} f(x + tv) \right|_{t=0} \\ &\stackrel{f(x+tv) = (f \circ \xi)(t)}{=} \left. \frac{dh}{dt} \right|_{t=0} \\ &= h'(0) \end{aligned}$$

Kettenregel:

$$\begin{aligned} h'(t) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\xi(t)) \cdot \xi'_i(t) \\ \implies h'(0) &\stackrel{\xi(0)=x+0 \cdot v, \xi'_i(t)=v_i}{=} \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) \cdot v_i \\ &= (\nabla f(x), v)_2 \end{aligned}$$

□

Korollar 0.6. Sei $\nabla f(x) \neq 0$. Dann ist der Winkel θ zwischen zwei Vektoren $v \in \mathbb{R}^n$ und $\nabla f(x) \in \mathbb{R}^n$ definiert durch

$$\cos(\theta) = \frac{(\nabla f(x), v)_2}{\|\nabla f(x)\|_2 \cdot \|v\|_2}.$$

Damit gilt für $\|v\|_2 = 1$

$$\frac{\partial f}{\partial v}(x) \stackrel{\text{Lemma 0.5}}{=} (\nabla f, v)_2 = \|\nabla f(x)\|_2 \cdot \|v\|_2 \cdot \cos(\theta) \stackrel{\|v\|_2=1}{=} \|\nabla f(x)\|_2 \cdot \cos(\theta)$$

$\frac{\partial f}{\partial v}(x)$ wird maximal, wenn $\cos(\theta) = 1 \implies v$ und $\nabla f(x)$ die gleiche Richtung haben: d.h. Der Vektor $\nabla f(x)$ ist die Richtung des stärksten Anstiegs von f im Punkt x .

Bemerkung 0.7. 1. Es gibt Funktionen, für welche alle Richtungsableitungen existieren, die aber dennoch nicht (total) differenzierbar sind.

2. Es gibt Funktionen, die stetig und partiell differenzierbar, aber nicht (total) differenzierbar sind.

Satz 0.8 (Kettenregel). Seien $D_f \subset \mathbb{R}^n$ und $D_g \subset \mathbb{R}^m$ offen, $g: D_g \rightarrow \mathbb{R}^n, f: D_f \rightarrow \mathbb{R}^r$ Abbildungen. Falls g im Punkt $x \in D_g$ und f im Punkt $y = g(x) \in D_f$ differenzierbar sind, gilt: Die Komposition $h = f \circ g$ ist im Punkt x differenzierbar und

$$\underbrace{D_x h(x)}_{\in \mathbb{R}^{r \times m}} = \underbrace{D_y f(g(x))}_{\in \mathbb{R}^{r \times n}} \underbrace{D_x g(x)}_{\in \mathbb{R}^{n \times m}}$$

Beweis. Seien $x \in D_g$ und $y = g(x) \in D_f$. Dann gilt nach Voraussetzungen

$$\begin{aligned} g(x + \xi) &= \underbrace{g(x)}_{=: y} + \underbrace{D_x g(x)\xi + w_g(\xi)}_{=: \eta} & \text{mit } \lim_{\substack{x+\xi \in D_g \\ \|\xi\| \rightarrow 0}} \frac{\|w_g(\xi)\|}{\|\xi\|} &= 0 \\ f(y + \eta) &= f(y) + D_y f(y)\eta + w_f(\eta) & \text{mit } \lim_{\substack{x+\eta \in D_f \\ \|\eta\| \rightarrow 0}} \frac{\|w_f(\eta)\|}{\|\eta\|} &= 0 \end{aligned}$$

Dann erhalten wir

$$\begin{aligned} (f \circ g)(x + \xi) &= f(g(x + \xi)) \\ &\stackrel{g(x+\xi)=g(x)+\eta=y+\eta}{=} f(y + \eta) \\ &= f(y) + D_y f(y) \cdot \eta + w_f(\eta) \\ &= f(y) + D_y f(y)(D_x g(x)\xi + w_g(\xi)) + w_f(\eta) \\ &= f(y) + D_y f(y)D_x g(x) \cdot \xi + D_y f(y)w_g(\xi) + w_f(\eta) \\ &= (f \circ g)(x) + \underbrace{D_y f(y)D_x g(x)}_{D_x(f \circ g)} \cdot \xi + w_{f \circ g}(\xi), \end{aligned}$$

wobei hier $w_{f \circ g}(\xi) = D_y f(y)w_g(\xi) + w_f(\eta)$. Es genügt also zu zeigen, dass

$$\lim_{\|\xi\| \rightarrow 0} \frac{\|w_{f \circ g}(\xi)\|}{\|\xi\|} = 0.$$

Aus $\lim_{\|\xi\| \rightarrow 0} \frac{\|w_g(\xi)\|}{\|\xi\|} = 0$ folgt sofort $\lim_{\|\xi\| \rightarrow 0} \frac{\|D_y f(y) w_g(\xi)\|}{\|\xi\|} = 0$. Wir schließen aber auch, dass es eine Konstante $c > 0$ geben muss, sodass $\|w_g(\xi)\| \leq c \|\xi\|$. Wegen $\lim_{\|\eta\| \rightarrow 0} \frac{\|w_f(\eta)\|}{\|\eta\|} = 0$ muss es ein $\tilde{w}_f(\eta)$ mit $\lim_{\eta \rightarrow 0} \tilde{w}_f(\eta) = 0$ geben, sodass $w_f(\eta) = \|\eta\| \cdot \tilde{w}_f(\eta)$. Mit diesen Aussagen gilt:

$$\begin{aligned} \|w_f(\eta)\| &\leq \|D_x g(x) \xi + w_g(\xi)\| \tilde{w}_f(\eta) \\ &\leq (\|D_x g(x)\| \|\xi\| + \|w_g(\xi)\|) \cdot \|\tilde{w}_f(\eta)\| \\ &\leq (\|D_x g(x)\| + c) \|\xi\| \cdot \|\tilde{w}_f(\eta)\| \\ \implies \frac{\|w_f(\eta)\|}{\|\xi\|} &= (\|D_x g(x)\| + c) \cdot \|\tilde{w}_f(\eta)\| \end{aligned}$$

Wegen $\lim_{\xi \rightarrow 0} \eta = \lim_{\xi \rightarrow 0} D_x g(x) \xi + w_g(\xi) = 0$ und der Stetigkeit von $\tilde{w}_f$ gilt

$$\lim_{\xi \rightarrow 0} \frac{\|w_f(\eta)\|}{\|\xi\|} = \lim_{\xi \rightarrow 0} (\|D_x g(x)\| + c) \cdot \|\tilde{w}_f(\eta)\| = (\|D_x g(x)\| + c) \cdot \tilde{w}_f\left(\lim_{\xi \rightarrow 0} \eta\right) = 0$$

Insgesamt erhalten wir

$$\lim_{\xi \rightarrow 0} \frac{\|w_{f \circ g}(\xi)\|}{\|\xi\|} = \lim_{\|\xi\| \rightarrow 0} \frac{\|D_y f(y) w_g(\xi)\|}{\|\xi\|} + \lim_{\xi \rightarrow 0} \frac{\|w_f(\eta)\|}{\|\xi\|} = 0$$

□

Bemerkung 0.9 (Komponentenweise für $i = 1, \dots, m$ und $j = 1, \dots, r$).

$$D_x h(x) = D_y f(y) D_x g(x) \Leftrightarrow \underbrace{\frac{\partial h_j}{\partial x_i}(x)}_{\partial_i (f \circ g)_j} = \sum_{k=1}^n \frac{\partial f_j}{\partial y_k}(g_1(x), \dots, g_n(x)) \cdot \frac{\partial g_k}{\partial x_i}(x_1, \dots, x_m)$$

Spezialfall: $m = r = 1$ ($g: D_g \subset \mathbb{R} \rightarrow \mathbb{R}^n$), $f: D_f \in \mathbb{R}^n \rightarrow \mathbb{R}$

$$h'(x) = \frac{d}{dx} h(x) = \frac{d}{dx} f(g(x)) = \sum_{k=1}^n \frac{\partial f}{\partial y_k}(g_1(x), \dots, g_n(x)) \cdot \frac{d}{dx} g_k(x) = (\nabla_y f(g(x)), g'(x))_2$$